Journal of Hydrology 576 (2019) 408–421

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

An evaluation of weather radar adjustment algorithms using synthetic data

Micha Silver[a,*], Arnon Karnieli[a], Francesco Marra[c], Erick Fredj[b,*]

[a] Ben Gurion University, Sde Boker, Israel
[b] Lev Academic Institute, Jerusalem, Israel
[c] Hebrew University, Jerusalem, Israel

ARTICLE INFO

This manuscript was handled by A. Bardossy, Editor-in-Chief

Keywords:
Weather radar
Rainfall
Gauges
Commercial microwave links
Synthetic model

ABSTRACT

Adjustment of weather radar estimates using observed precipitation has been an accepted procedure for decades. Ground observations of precipitation typically come from rain gauges, but can also include data from diverse networks of sensors, with different levels of reliability. This study presents a standardized framework for evaluating adjustment algorithms using synthetically constructed, but realistic, rain grids and weather radar rainfall. Ground observation points are randomly placed throughout the synthetic storm domain and the precipitation for each sensor is extracted from the true rain. Then a subset of the sensors are defined as unreliable, and a log-normal error factor is applied at those locations.

This double network of rain sensors could be applicable, for example, when rainfall is derived from signal attenuation between commercial microwave link (CML) antennas. Past research has tested CML observations as a source of precipitation data and validated various radar adjustment algorithms. However, a comprehensive evaluation of adjustment algorithms using accurate gauge data mixed with CML observations at different densities is lacking.

Five adjustment algorithms are applied to the synthetic radar grid: Mean Field Bias (MFB), a Multiplicative algorithm, Mixed (additive and multiplicative), Conditional Merge (CondMerge) and Kriging with External Drift (KED). Generation of the synthetic framework, and application of the adjustment algorithms is repeated for 150 realizations. Comparison of coefficient of determination (R^2), root mean square error and linear regression for all adjustment procedures over all realizations indicates the following results. Only MFB and KED adjustments performed well when using accurate gauges. The kriging based KED was able to achieve good adjustment also with the addition of error-prone sensors. CondMerge and the Mixed and Multiplicative, however, resulted in poorer adjustments.

1. Introduction

Shortly after weather radar became an accepted source of precipitation data, efforts began to improve quantitative precipitation estimates (QPE) by merging ground based observations. For example Krajewski (1987) and Seo (1998) presented early applications of kriging to correct radar rainfall with gauge data. Gjertsen et al. (2004) explained the advantages and disadvantages of both radar rainfall and gauge observations, concluding that proper adjustment of radar rainfall with ground based observations is crucial to achieving accurate QPE. Work by Velasco-Forero et al. (2009) demonstrated a unique non-parametric method of determining covariance for the gauge data before applying the kriging interpolations. A Bayesian approach was introduced by Todini (2001) and further developed in Mazzetti and Todini (2004). Co-kriging was applied successfully by Sideris et al. (2014). They prepared an accurate precipitation estimate where the primary variable was gauge observations at one time interval, and three additional secondary variables were gauge observations from an earlier time interval and radar values at gauge locations from both time steps. A different co-kriging approach, presented recently by Foehn et al. (2018), created precipitation estimates using two independent and non-collocated gauge networks. They preformed Inverse Distance Weighted, Regression Kriging and Regression Co-kriging. In all cases they corrected for non-stationarity of gauge data by calculating and removing the trend which was produced from residuals between gauges and radar. Their approach enabled successful QPE for short time intervals in a mountainous region with severe beam blockage of the available radar images.

Some of these techniques were applied operationally by Amorati et al. (2012) and Berg et al. (2016). Several national meteorological

* Corresponding authors.
E-mail addresses: silverm@post.bgu.ac.il (M. Silver), karnieli@bgu.ac.il (A. Karnieli), marra.francesco@mail.huji.ac.il (F. Marra), fredj@jct.ac.il (E. Fredj).

https://doi.org/10.1016/j.jhydrol.2019.06.064
Received 13 February 2019; Received in revised form 19 June 2019; Accepted 22 June 2019
Available online 25 June 2019

services, among them the German Deutcher Wetterdienst (DWD) (https://www.dwd.de/EN/ourservices/radar_products/radar_products.html) and the Royal Netherlands Meteorological Service (https://data.knmi.nl/datasets) publish adjusted weather radar regularly. A review of the application of radar-rain gauge adjustments across Europe appears in Gjertsen et al. (2004).

Research papers over the past decade have addressed the ability of commercial microwave link (CML) attenuation data to represent rain rates, and the uncertainty involved. Signal attenuation from CML antennas, and the derived rain rate are impacted by several sources of uncertainty:

- non-linear attenuation response
- distance between antennas
- interference resulting from water accumulating on the antennas (similar to wet radome attenuation in weather radar)
- calibration parameters for the attenuation-rain rate relationship
- microwave frequency
- drop size distribution

In Leijnse et al. (2010) the authors show that distance between CML antennas becomes a major source of uncertainty at certain microwave frequencies with distant (greater than 5 km.) antennas. They point to drop size distribution as the major cause of uncertainty in CML precipitation rates. Bianchi et al. (2013) added data from CML as a second ground-based source for radar adjustment. In an early work Berne and Uijlenhoet (2007) described the source of errors in CML derived rain rates. More recently van Leth et al. (2018) carefully analyzed the uncertainties in attenuation of a microwave link at a controlled research installation.

Evaluations of the various adjustment algorithms have appeared in many research papers. One of the early assessments was done by Goudenhoofdt and Delobbe (2009). Since then Wang et al. (2012) and McKee and Binns (2016) have covered a broad selection of the adjustment procedures. An overview of adjustment algorithm categories that appears in Hasan et al. (2016) investigated the issue of error independence between the adjusted and true rain grids. Ly et al. (2013) categorized the various adjustment methods into deterministic and geo-statistical. They concluded that geo-statistical methods out-perform deterministic in long temporal aggregations of rainfall.

Synthetic data-sets have been employed in climate and meteorology research to address complicated and multivariate systems. For example Edouard et al. (2018) used synthetic data to determine uncertainty in flash flood forecasting. They performed a sensitivity analysis by artificially applying a range of soil parameters in a combined modeling framework. Similarly, Musayev et al. (2018) recently examined the rainwater harvesting potential under several global climate models using synthetically generated daily rainfall. A new radar adjustment method, combining both gauge and crowd sourced data was presented by Yang and Ng (2019). Their method applied Bayesian regression to synthetically produced rain and radar fields as well as synthetically derived gauge and crowd sourced data.

The aim of this research is to present a standardized, fully reproducible framework for determining which adjustment procedures achieve the best correlation to the true rain using varying mixtures of reliable and error-prone ground based observations.

2. Methods

The procedure described below created 150 realizations of a synthetic framework by constructing simulated rain grids, incorporating a noise component to produce radar grids, and randomly distributing observation locations throughout the grid extent. This procedure was implemented using R (R Development Core Team, 2008) software with functions adapted from Guenzi et al. (2016), Nerini et al. (2017) and Morin and Gabella (2007). A uniform random number seed was applied throughout, making the procedure fully reproducible.

2.1. Synthetic framework

The synthetic model presented in this work does not attempt to be an exact representation of any real rain event, radar data or distribution of gauges. Rather, like any model, certain characteristics of reality were distilled to create an analysis framework under which specific aspects of radar adjustments could be examined. Refer to Yang and Ng (2019) where similar simplifications were adopted in creating synthetic rain, radar and gauge inputs. This framework included initial ‘true’ rain distributions across a predefined grid, referred to as the domain. The domain size, 100 × 100 km, was chosen to represent the range covered by a typical C-band radar.

The temporal resolution was assumed to be 24 h and the accumulated precipitation was chosen under that assumption. While operational adjustment of radar is performed on individual radar scans or short aggregations, calibration and evaluation of adjustments are applied to long time aggregations following Goudenhoofdt and Delobbe (2009). This 24 h aggregation avoids short term effects due to the high temporal variability of rainfall (as described in Marra and Morin (2018)).

Weather radar measurements involve several sources of uncertainty (summarized in Villarini and Krajewski (2010)), some of which were encapsulated into the test framework of this work by applying configurable noise structures. The ‘true’ rain grid was multiplied by three spatially varying noise grids (Fig. 1) to simulate the radar grids, similarly to the combined uncertainty model suggested by Ciach et al. (2007) and later by Villarini and Krajewski (2009b). The first was a multiplicative bias level, varying radially from the center of the domain, assumed to be the location of the radar. This noise grid was designed to mimic range degradation (Fig. 1 panel (a)).

Range degradation results from several unrelated causes: increased sampling volume due to the expanding radar beam, attenuation from atmospheric moisture or dust, and various effects attributed to the vertical profile of reflectivity (VPR). Krajewski et al. (2011) and Vignal and Krajewski (2001) show that VPR effects are the most significant cause of range degradation at long ranges. Reflectance from melting ice hydrometeors, usually at altitudes between 2 and 4 km, causes the ”bright band” effect, leading to high reflectance. At altitudes above the bright band, reflectance drops more or less linearly, due to the expanding radar beam. At lower altitude, on the other hand, VPR has very little effect on range degradation. Thus, the lowest radar elevation angle will not reach the bright band, and not be affected by VPR, but at the second and higher radar elevation angles the beam reaches the bright band altitude closer to the radar. This current work considers the storm domain to be 100 km × 100 km with the radar at the center, so the maximum range covered in this simulation is 74 km. Krajewski et al. (2011) show that up to 70 km VPR has no effect on the range dependent error at the lowest elevation angle. The second elevation angle has a small effect at 70 km, and at the third elevation angle a more substantial effect, 20% to 40% error, at 50 to 70 km respectively.

Range degradation was split into two zones by Sebastianelli et al. (2012). They considered the zone near to the radar to be affected mostly by the Z-R relationship and rain rate, with almost no range dependent decrease in radar reflectivity. In their study of an S-band radar, range dependent uncertainly appeared only beyond the 50 km. range, due to radar beam spread, and vertical profile of reflectivity. This same two zone approach appeared and was validated also in Chumchean et al. (2003). Both of these works contained a discontinuity in the error model between the zones. On the other hand, Michelson and Koistinen (2000) presented a radar degradation relation whereby the log of the gauge-radar ratio followed a second order polynomial of range r: $log\left(\frac{G_r}{R_r}\right) = a + b^* r + c^* r^2$.

Expanding on the above studies, the range dependent error

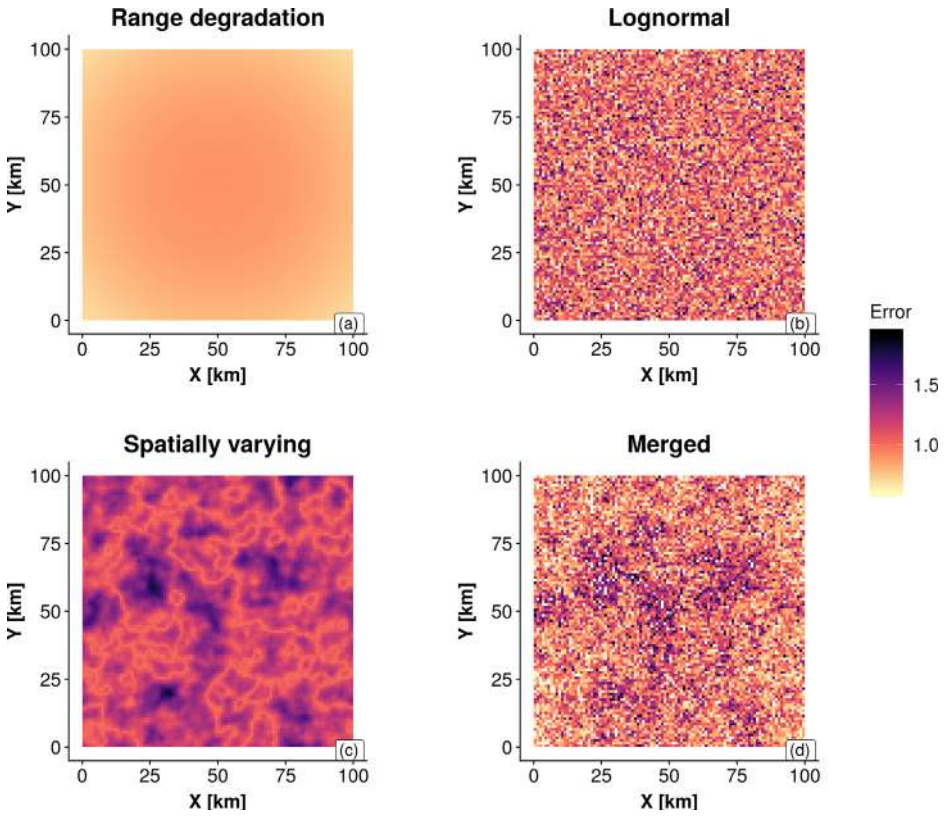

Fig. 1. Three types of errors used to construct the synthetic weather radar precipitation grid. The true rain grid is multiplied by the merged error grid (d) to produce the radar grid. Types of errors: (a) range degradation error; (b) Random noise (based on log-normal distribution); (c) Spatially correlated random noise structure; (d) the merged error grid. The legend shows multiplicative factors.

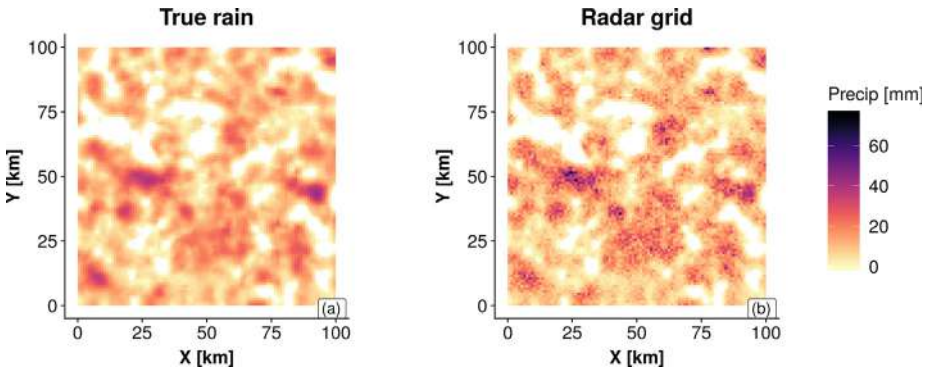

Fig. 2. True rain (a) and radar precipitation (b) values, on a synthetic Cartesian grid of 100 × 100 km.

component in this work was defined as a Gaussian function (Eq. (1)) also chosen by Villarini and Krajewski (2009a) in their error model. Thus, near the radar antenna the error factor of 1.0 did not incur any change in the 'true' rain, and at larger radial distances the error factor decreased following a Gaussian curve, approximating the two zone range degradation presented above, but avoiding discontinuity in the function. Furthermore, a Gaussian curve drops more slowly at small ranges than the polynomial used by Michelson and Koistinen (2000). Thus we maintain almost no degradation up to 15 km. from the radar, then the range degradation error grows to 25% at the maximum range of 70 km. Refer to panel (a) of Fig. 2 to see the slight effect of this error at the corners of the domain.

$$Err(r) = a^{*}e^{\left[-\frac{(r-b)^2}{2*c^2}\right]} \quad (1)$$

where: *Err* is the error factor, *r* is the distance from radar; a, b, c are empirical constants chosen to emulate almost no range degradation near the radar ($a = 1$), and a Gaussian decreasing value ($b = 4$ and $c = 100$) such that the error level at the maximum range of 70 km is 0.75.

The second radar noise grid was created from a log-normally distributed random multiplicative field, to represent uncertainty in the power law (Z-R relationship), non-uniform beam filling, non-uniform drop size distributions, etc. The third noise grid was constructed from a spatial structure defined following the methods in Nerini et al. (2017). The theoretical base for this stochastic method was developed by Seed (2002) and implemented in the Short-Term Ensemble Prediction System (STEPS) probabilistic forecast scheme by Bowler et al. (2006). This spatial structure was derived from the 'true' rain grid, thus providing a precipitation dependent error.

These three radar noise grids, once multiplied together, contained error values from 0 to above 3.0. Thus the resulting radar, the product of the 'true' rain and this error grid, acquired values both above and below the 'true' rain. Those introduced error grids are multiplicative in nature, and their combination aimed at reproducing realistic observational conditions. The noise levels used to create this radar grid, and the error levels applied to the simulated error-prone observation locations (Section 2.2) are all configurable, allowing construction of uniform and

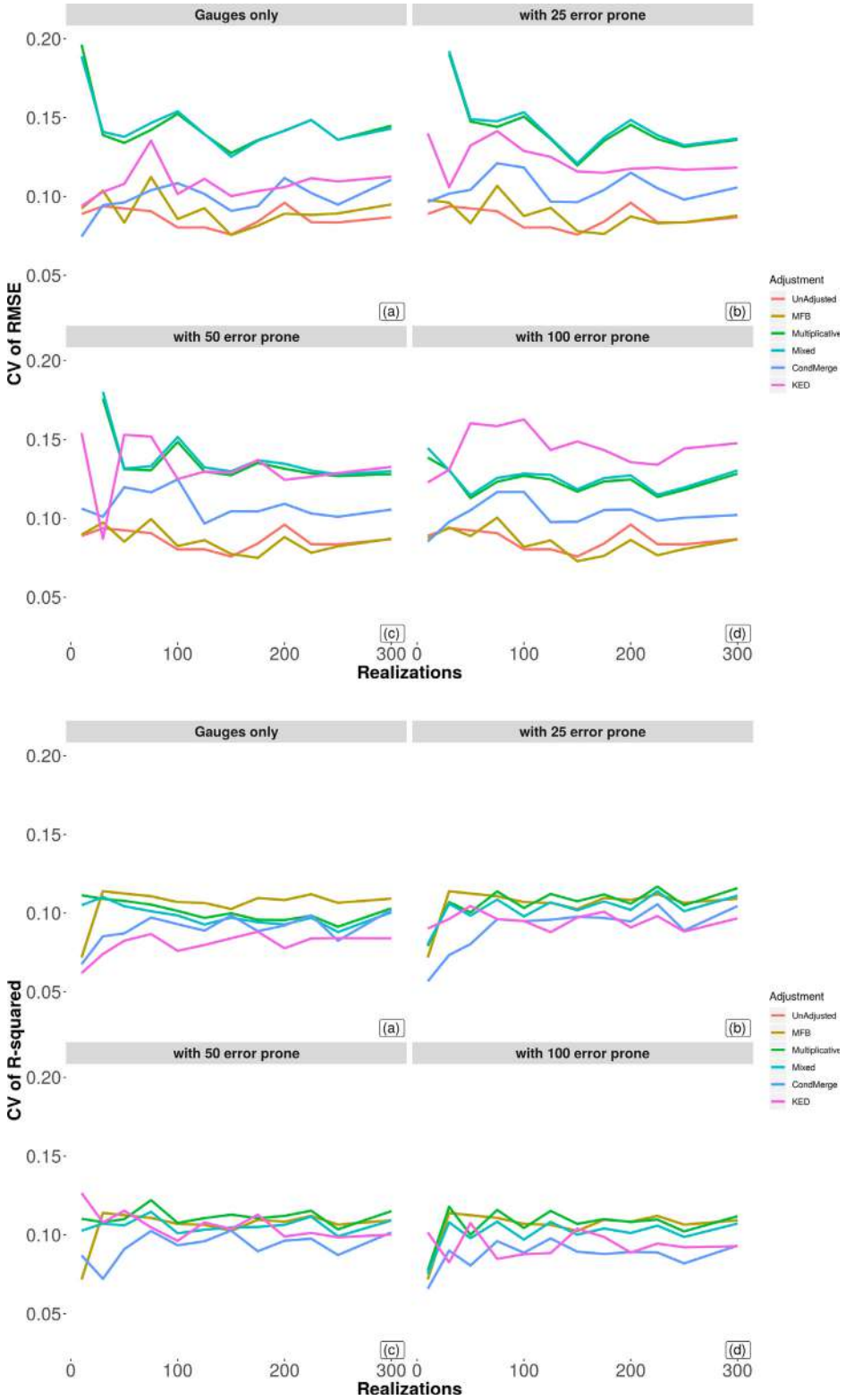

Fig. 3. Evaluation of optimal number of realizations. **Coefficient of Variation** of RMSE (top) and R^2 (bottom) for each adjustment algorithm, using a range from 10 to 300 realizations. The evaluation was repeated for all numbers of error-prone rain sensors: (a) 50 gauges only; (b) including 50 error-prone sensors; (c) including 100 error-prone; (d) including 200 error-prone.

reproducible test frameworks for the various adjustment procedures. Fig. 2 illustrates an example of the resulting radar rainfall.

2.2. Ground based precipitation sensors

Two types of ground sensors were simulated in this framework, reliable and error-prone. Observation locations were placed randomly throughout the storm domain, with 50 reliable (gauge) locations and

Table 1
Summary Statistics.

Adjustment	Err_Prone	RMSE	R.squared	Slope	Intercept
UnAdjusted	0	7.039	0.433	0.600	3.946
UnAdjusted	50	7.039	0.433	0.600	3.946
UnAdjusted	100	7.039	0.433	0.600	3.946
UnAdjusted	200	7.039	0.433	0.600	3.946
MFB	0	4.143	0.433	0.639	0.016
MFB	50	4.123	0.433	0.634	−0.033
MFB	100	4.097	0.433	0.636	−0.085
MFB	200	4.089	0.433	0.637	−0.135
Multiplicative	0	4.521	0.628	0.378	2.681
Multiplicative	50	4.779	0.608	0.362	2.835
Multiplicative	100	4.879	0.599	0.355	2.898
Multiplicative	200	4.918	0.595	0.353	2.920
Mixed	0	4.259	0.631	0.396	2.604
Mixed	50	4.515	0.617	0.377	2.753
Mixed	100	4.623	0.612	0.369	2.811
Mixed	200	4.678	0.611	0.365	2.829
CondMerge	0	3.265	0.637	0.600	0.547
CondMerge	50	3.284	0.650	0.541	1.087
CondMerge	100	3.263	0.664	0.523	1.254
CondMerge	200	3.222	0.684	0.508	1.396
KED	0	3.392	0.615	0.629	0.151
KED	50	3.271	0.636	0.615	0.187
KED	100	3.151	0.659	0.616	0.138
KED	200	2.896	0.712	0.618	0.087

three sets of error-prone: 50, 100 and 200 additional locations.

While gauges actually do sustain errors, modern rain rate measuring equipment has reduced the measurement uncertainty to approximately 5% (see, for example Colli et al. (2013)) thus gauge data were assumed to be accurate in the simulated framework; their aggregated precipitation was extracted directly from the initial 'true' grid. Throughout this work the terms 'gauges' and 'reliable sensors' are used interchangeably.

In operational weather monitoring, gauge networks are quite sparse (For a global survey of gauge densities see Kidd et al. (2017)). The influence of gauge density on radar-gauge merging was reported by Otieno et al. (2014) and Shafiei et al. (2014) further assessed the importance of gauge density and distribution. The density of gauges chosen for this work, 1 gauge per 200 km^2, attempted to mirror a typical gauge network (see further reports by Ahrens (2006), Tian et al. (2018) and Otieno et al. (2014)).

The source of error-prone observations can be a second, less reliable network of gauges. Alternatively, microwave signal attenuation in CML networks has become accepted as a source of point precipitation data. (Fencl et al. (2017); Goldshtein et al. (2009); Overeem et al. (2013)). This current work does not assert that CML derived precipitation is inherently inaccurate. However much work has been done to investigate the uncertainty associated with this source of precipitation data (i.e. Leijnse et al. (2010); Zinevich et al. (2010)). While attenuation data between CML towers represents the aggregated signal strength along the line-of-sight between towers, most research (i.e. Overeem et al. (2013) and Goldshtein et al. (2009)) simplifies this linear feature into a single observation point at the center between the towers. In their analysis of rainfall estimation from CML networks, Zinevich et al. (2010) refer to link lengths from 0.5 to 8 km, with increasing uncertainty in longer links. The report by Overeem et al. (2016), covering a long term QPE program in the Netherlands, discusses an average link length of 3.6 km for over 2000 links.

In this work, the term 'error-prone sensor' refers to point observations that could represent the midpoints along CML links. This geographic simplification together with other uncertainties (see Leijnse et al. (2010) and additional references in Section 1) in CML observations concur in creating the overall CML error. However, so far no explicit model for such an error is available. In this study, we chose to use a multiplicative error model in order to (i) deal with a reduced number of parameters and (ii) make the model easily adaptable to other remote sensing instruments (Tian et al. (2013) and Tang et al. (2015)). To do so, a random log-normally distributed multiplicative error has been

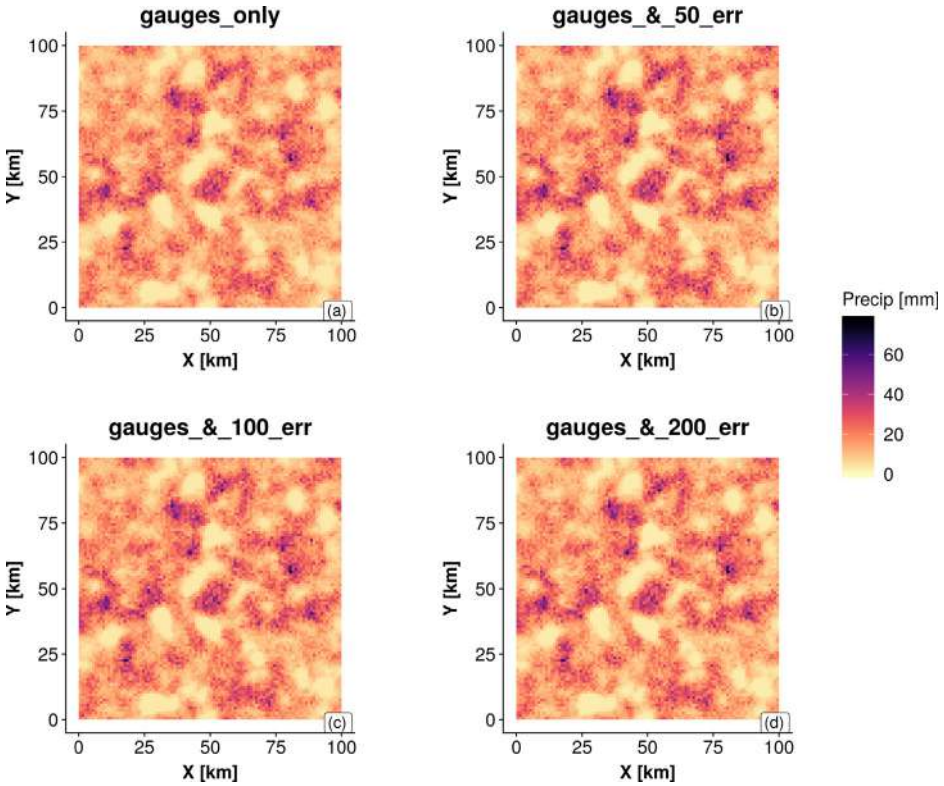

Fig. 4. Adjusted precipitation grids using **Mean Field Bias**, for different numbers of error-prone rain sensors. (a) 50 gauges only; (b) including 50 error-prone sensors; (c) including 100 error-prone; (d) including 200 error-prone.

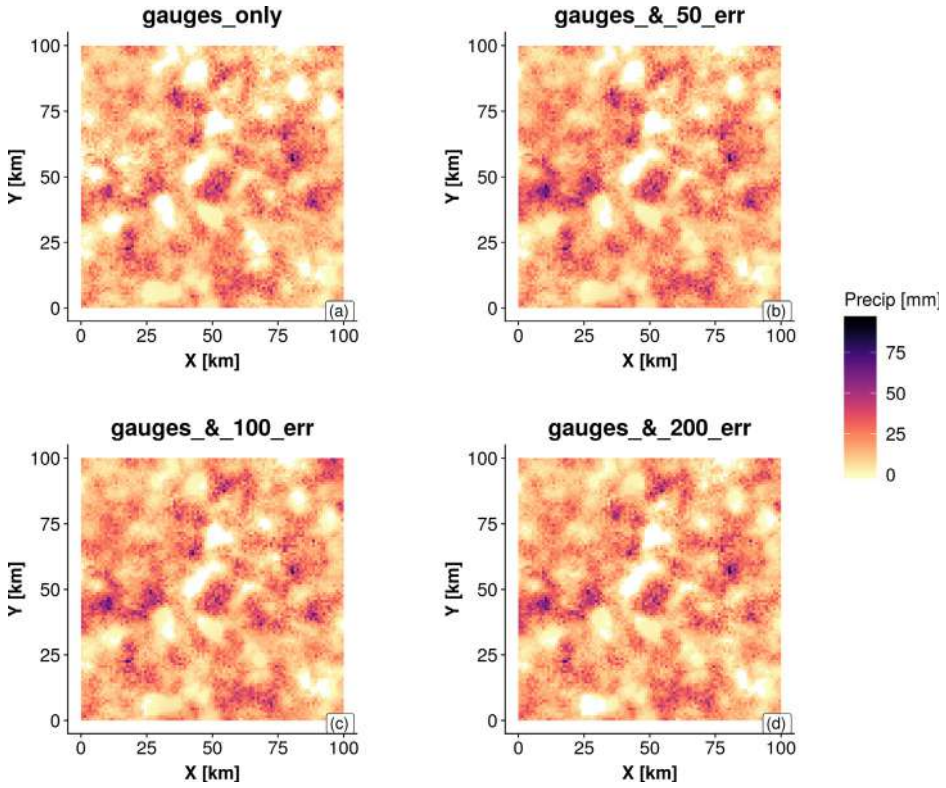

Fig. 5. Adjusted precipitation grids using **Conditional Merge**, for different numbers of error-prone rain sensors. (a) 50 gauges only; (b) including 50 error-prone sensors; (c) including 100 error-prone; (d) including 200 error-prone.

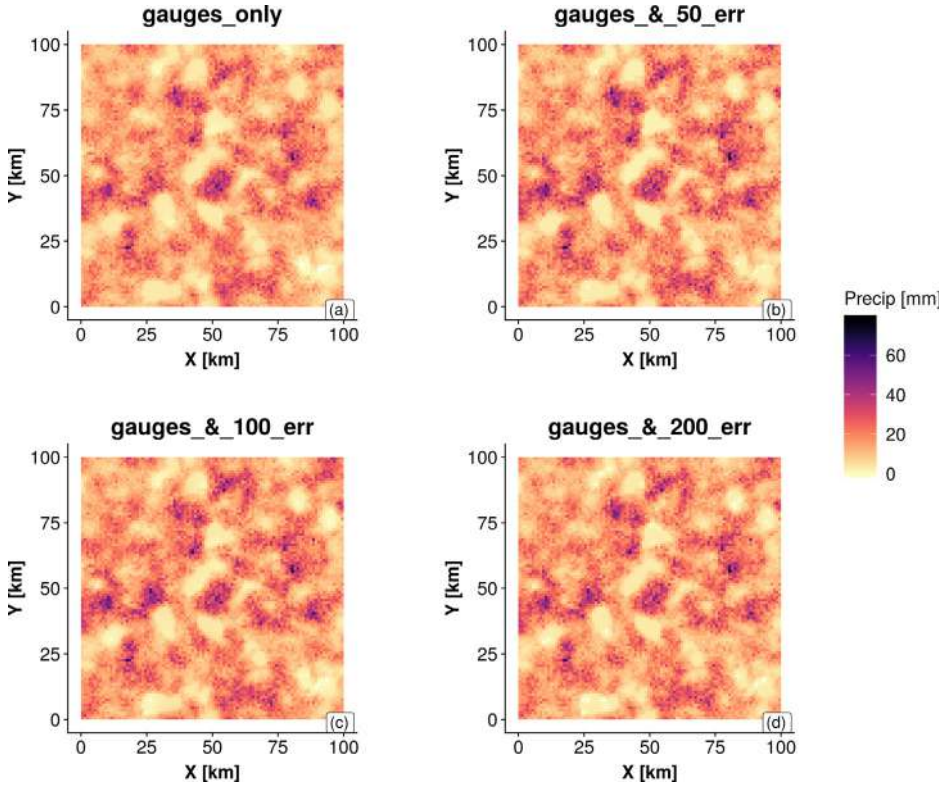

Fig. 6. Adjusted precipitation grids using **Kriging with External Drift**, for different numbers of error-prone rain sensors. (a) 50 gauges only; (b) including 50 error-prone sensors; (c) including 100 error-prone; (d) including 200 error-prone.

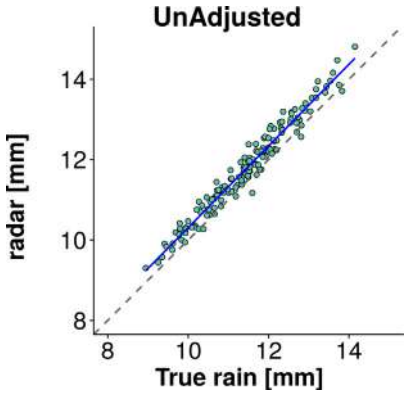

Fig. 7. Scatter plot of true rain vs. **UnAdjusted** original radar precipitation values.

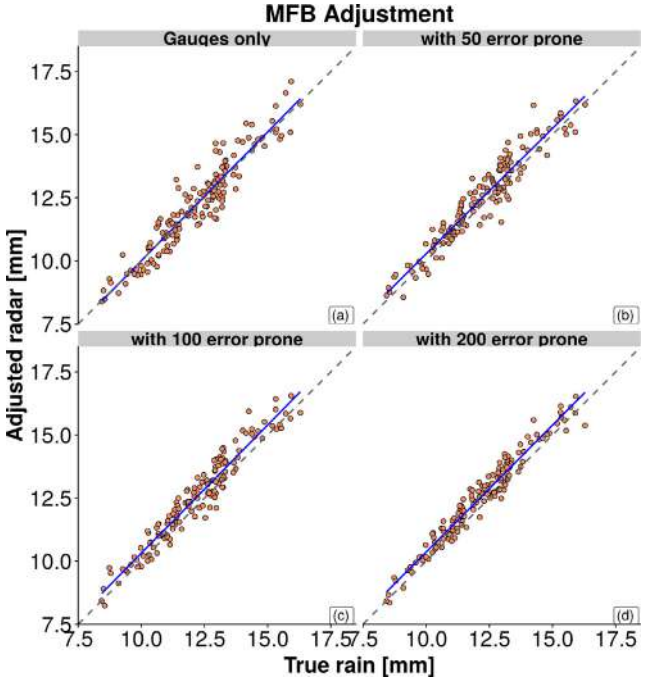

Fig. 8. Scatter plot of true rain vs. **MFB** adjusted radar precipitation values, for different numbers of error-prone rain sensors: (a) 50 gauges only, (b) including 50 error-prone sensors, (c) including 100 error-prone, (d) including 200 error-prone.

applied to the 'true' precipitation to simulate error-prone observations. This error shares the log-normal distribution parameters, and thus the error magnitude distribution, with the multiplicative error used for radar fields, but with an independent spatial pattern in each simulation.

Furthermore, CML towers are usually installed at high density compared to rain gauges (for example Chwala et al. (2012) report one CML link per 3 km^2 in southern Germany). Therefore CML derived precipitation offers a possible real world case with a mixture of few reliable (gauge) and many less reliable (CML) point precipitation sensors. Locations for error-prone data observations were therefore placed at three (configurable) densities: one observation location per 200 km^2 (identical to the gauge network) one per 100 km^2 and one per 50 km^2 (four times as many error-prone sensors as reliable). Thus the densities of reliable and error-prone sensors together were:

- 50 reliable only (representing rain gauges at 1 gauge per 200 km^2)
- Additional 50 error-prone observation locations (totaling one observation location per 100 km^2)

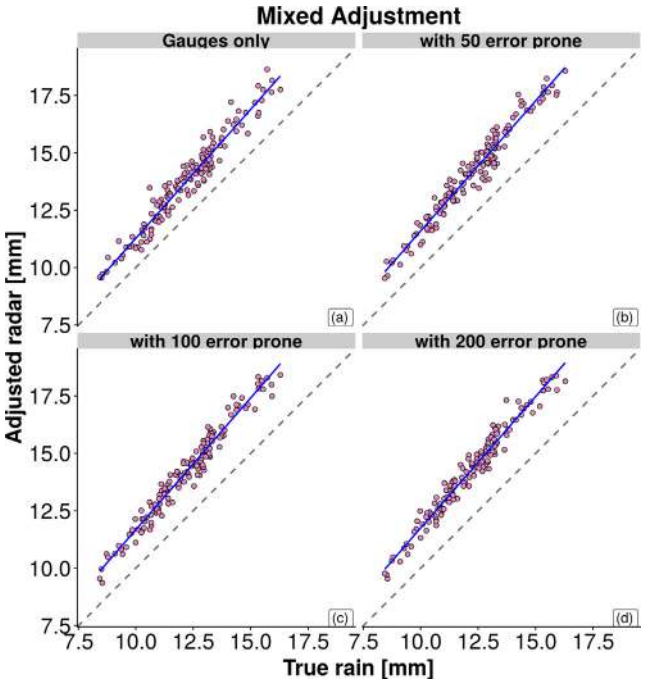

Fig. 9. Scatter plot of true rain vs. **Mixed** adjusted radar precipitation values, for different numbers of error-prone rain sensors: (a) 50 gauges only, (b) including 50 error-prone sensors, (c) including 100 error-prone, (d) including 200 error-prone.

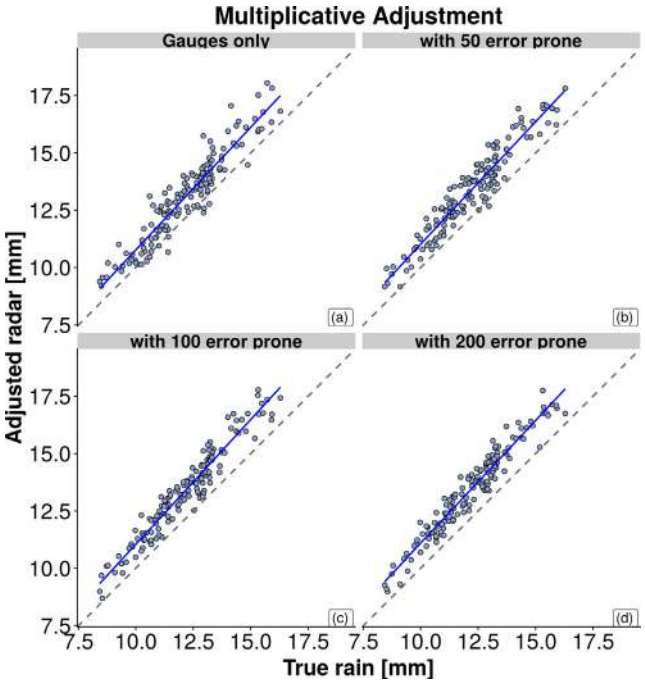

Fig. 10. Scatter plot of true rain vs. **Multiplicative** adjusted radar precipitation values, for different numbers of error-prone rain sensors: (a) 50 gauges only, (b) including 50 error-prone sensors, (c) including 100 error-prone, (d) including 200 error-prone.

- Additional 100 error-prone (totaling one observation per 66 km^2)
- Additional 200 error-prone (totaling one observation per 40 km^2)

2.3. Adjustment procedures

This work implemented and tested five adjustment algorithms:

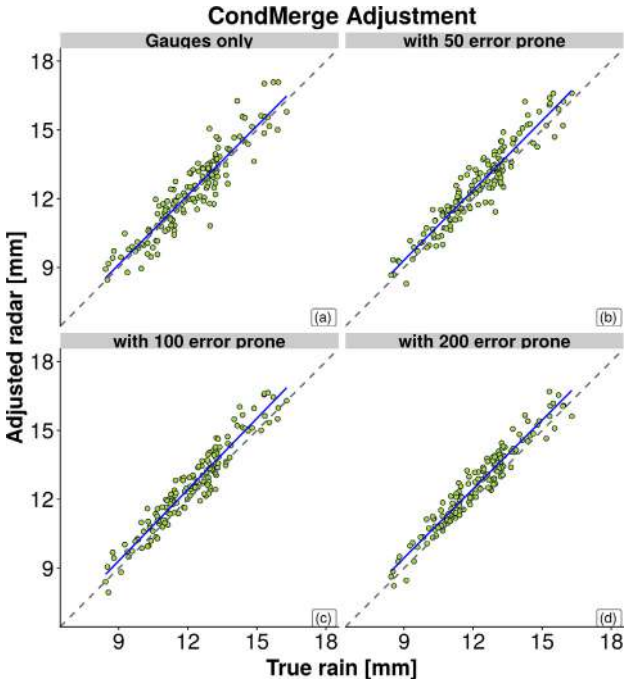

Fig. 11. Scatter plot of true rain vs. **CondMerge** adjusted radar precipitation values, for different numbers of error-prone rain sensors: (a) 50 gauges only, (b) including 50 error-prone sensors, (c) including 100 error-prone, (d) including 200 error-prone.

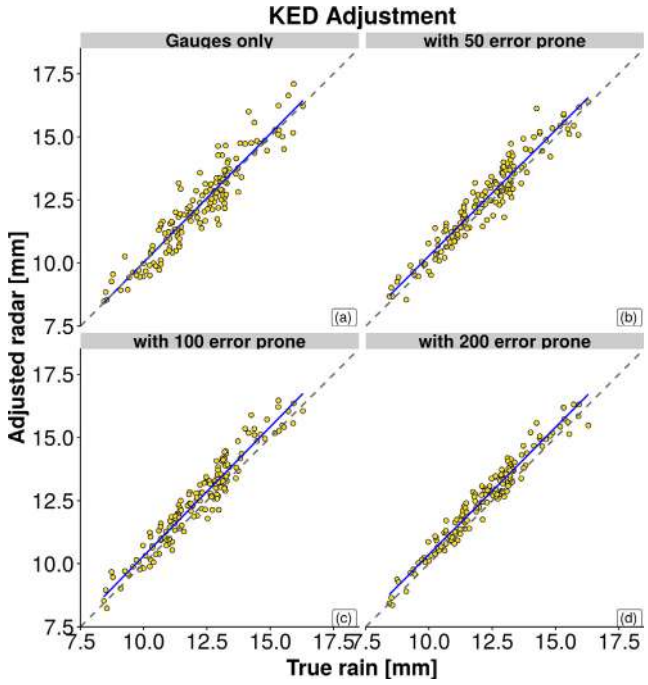

Fig. 12. Scatter plot of true rain vs. **KED** adjusted radar precipitation values, for different numbers of error-prone rain sensors: (a) 50 gauges only, (b) including 50 error-prone sensors, (c) including 100 error-prone, (d) including 200 error-prone.

Mean Field Bias (MFB), Multiplicative, Mixed, Conditional Merge (CondMerge) and Kriging with External Drift (KED). The choice of these five among the full spectrum of adjustment procedures covered the range of both deterministic (MFB) and geostatistical methods. Multiplicative, CondMerge and KED employ variations of kriging, implemented with the `gstat` package (Pebesma (2004)) in R. Furthermore, the Mixed procedure allows comparison with an additive bias method, and CondMerge, as explained below, also applies an additive, spatially varying correction. On the other hand MFB and KED adjustments adjust the radar solely by a multiplicative factor.

2.3.1. Mean Field Bias

The first, MFB, determines the average, multiplicative bias shift between the radar grid and observed precipitation, then corrects for that shift globally. This is the only adjustment algorithm that applies a correction which does not vary spatially. Each of the other algorithms calculates a spatially varying correction from the differences between the observation locations and radar pixels at those locations.

2.3.2. Mixed and Multiplicative

Both the Multiplicative and Mixed algorithms[1] that are explained in a technical report (Pfaff (2010) in German) apply spatially varying corrections by interpolating the residuals between radar and sensors at the sensor locations using ordinary kriging. Multiplicative adjustment interpolates a correction grid by ordinary kriging of the residuals between the observation locations and the radar grid. The radar grid is multiplied by this correction grid to obtain the adjusted result. The Mixed algorithm works similarly, but determines two correction grids: an additive component and a multiplicative component. The adjusted grid results from multiplying the radar by the multiplicative factor, then adding the additive component.

2.3.3. Conditional Merge

The CondMerge algorithm, initially introduced by Sinclair and Pegram (2005) and implemented by Guenzi et al. (2016), applies ordinary kriging to two different point sets: the values of the 'true' rain grid at gauge locations, and the values of the radar grid at those locations. The adjusted grid results from adding to the radar grid the kriging interpolation of the gauges, then subtracting the interpolation of the radar values, categorizing this procedure as additive. In this way, the procedure conditions the gauge adjustment by the radar precipitation values at gauge locations. This procedure was applied successfully by Berndt et al. (2014). In addition, Rabiei and Haberlandt (2015) compared Conditional Merging to Kriging with External Drift in the context of a new bias correction protocol. They reported that conditional merging performed best after applying an initial bias correction.

2.3.4. Kriging with External Drift

Among the geostatistical adjustment algorithms, KED has been applied widely. This interpolation method uses a secondary variable, the radar precipitation itself in this case, to add a trend to the gauge observations. In addition to research by Berndt et al. (2014) and Rabiei and Haberlandt (2015) cited above, Schiemann et al. (2011) also found that KED attained better correlation to the original rain than Ordinary Kriging (OK) methods. Delrieu et al. (2014) further reported that KED out-performed OK for twelve rain events in France.

2.4. Realizations

The evaluation in this work was performed on 150 realizations of synthetic data: initial 'true' storms, radar grids derived from the 'true' rain by applying noise and bias factors and randomly placed ground-based sensor locations. The 'true' storm grids were created using Sequential Gaussian Simulations (SGS), similar to the approach used by Nour et al. (2006) to choose gauge locations, and by Teo and Grimes (2007). SGS was used also by Cecinati et al. (2018) to create a realistic rainfall simulation. They then applied different gauge distributions and error models to that simulation to determine the gauge uncertainty by kriging for uncertain data, a method introduced by Mazzetti and Todini

[1] adapted in this work from https://docs.wradlib.org/en/stable/adjust.html.

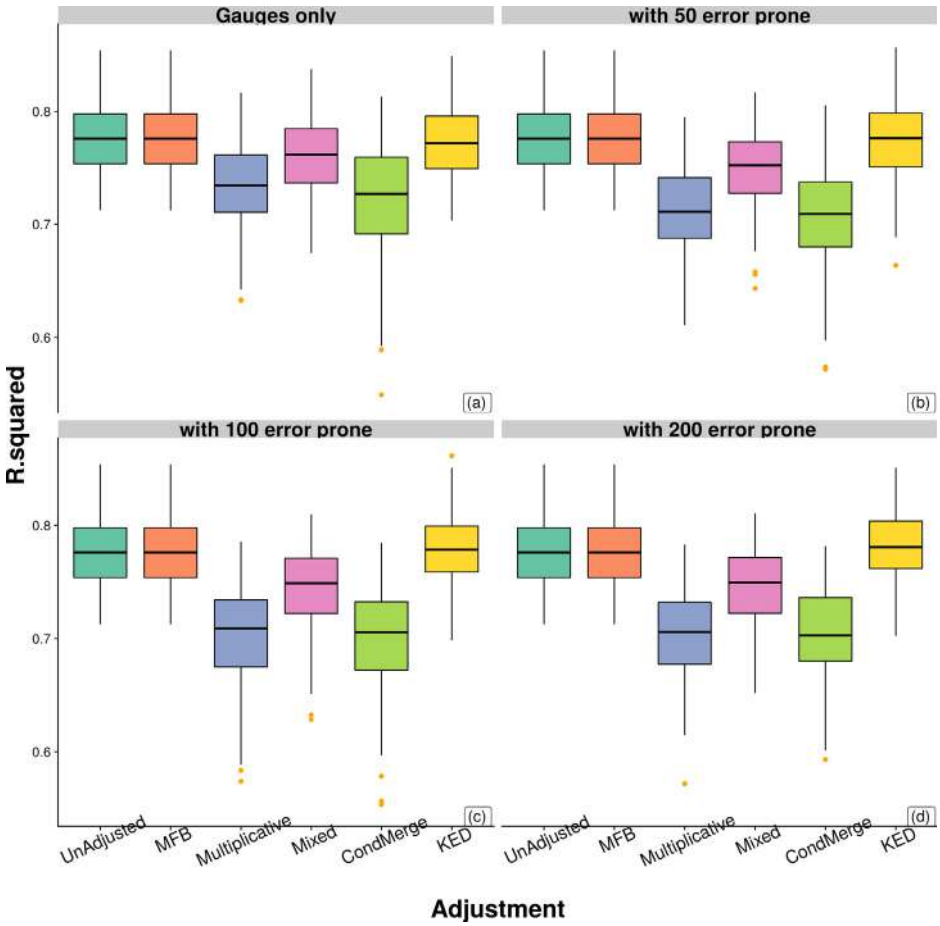

Fig. 13. Box-whisker plot of R^2 from 150 rain realizations for each adjustment with different numbers of error-prone rain sensors: (a) 50 gauges only, (b) including 50 error-prone sensors, (c) including 100 error-prone, (d) including 200 error-prone.

(2008).

In this current work a multiplicative error grid was applied to all simulations of 'true' rain to obtain 150 synthesized radar precipitation grids. The radar noise structure used to create the radar grids is described in Section 2.1 and error levels of ground based sensors in Section 2.2. Fig. 1 shows the error grids and in Fig. 2 the 'true' and resulting radar grids (for one sample realization) are shown. The five adjustment procedures, detailed in Section 2.3 were then applied to each realization. Analysis of the results (in Section 2.5) calculated the mean values of several correlation statistics for each adjustment.

The choice of 150 realizations was determined after testing from 10 to 300 iterations. Coefficients of variation (*CV*), (the ratio of variance to mean, given by $CV = \sigma/\mu$) for RMSE and R^2 were determined and averaged at different numbers of realizations (see Fig. 3). The variation of these statistical measures over the range of realizations stabilized after about 100 realizations. A larger number did not substantially change the means of *CV* for the different test statistics at the four error sensor densities, thus 150 realizations was considered sufficient.

2.5. Analysis of adjustments

In order to examine the success of each adjustment algorithm, a set of 400 validation points was randomly chosen within the domain. This number resulted from Slovin's formula to select sample size, Eq. (2):

$$n_{samp} = \frac{N}{1 + N(e)^2} \tag{2}$$

with e = 0.05 (confidence level of 95%) and N = 10,000 (total pixels in the domain)

Values of the original 'true' rain and each of the adjusted grids were extracted at these validation points for all realizations separately, and RMSE was calculated. Examining adjustment algorithms in the context of a synthesized test framework insured that results were uniform and reproducible. Each adjustment was applied to all 150 rain realizations, and with four densities of error-prone observation locations. Then average coefficient of determination (R^2) and root mean square error (RMSE), as well as linear regression slope and intercept between the adjusted grids and the initial 'true' storm were calculated and examined. Table 1 presents mean statistics for each of these tests.

Initially, in order to justify use of a linear regression, two tests were applied to check for normality of residuals: Anderson-Darling and Shapiro–Wilk. The *p-values* of both were very small ($<10^{-6}$) supplying evidence that the null hypothesis of normality should be rejected. Therefore a Box-Cox transform (Box and Cox, 1964) was applied to the set of adjusted values. This transform was used by both Erdin et al. (2012) and Cecinati et al. (2016) in the context of kriging-based radar adjustment. The optimal Box-Cox λ parameter was determined then the transformed data were again tested for normality (repeating the Anderson–Darling test) as well as homoscedasticity, using the Breusch-Pagan test (implemented in the `car` package in `R`, Fox and Weisberg (2011)). These tests showed p-values above 0.2 for all realizations (Breusch-Pagan p-values above 0.08), thus rejection of the null hypothesis was not suggested at a confidence level of 95%, and preconditions for linear analysis were upheld. Visual examination of Q-Q plots for a few realizations reinforced this conclusion. Then R^2 and RMSE were obtained from linear regressions of the Box-Cox transformed distributions to compare values of the 'true' and adjusted rain grids at the validation locations.

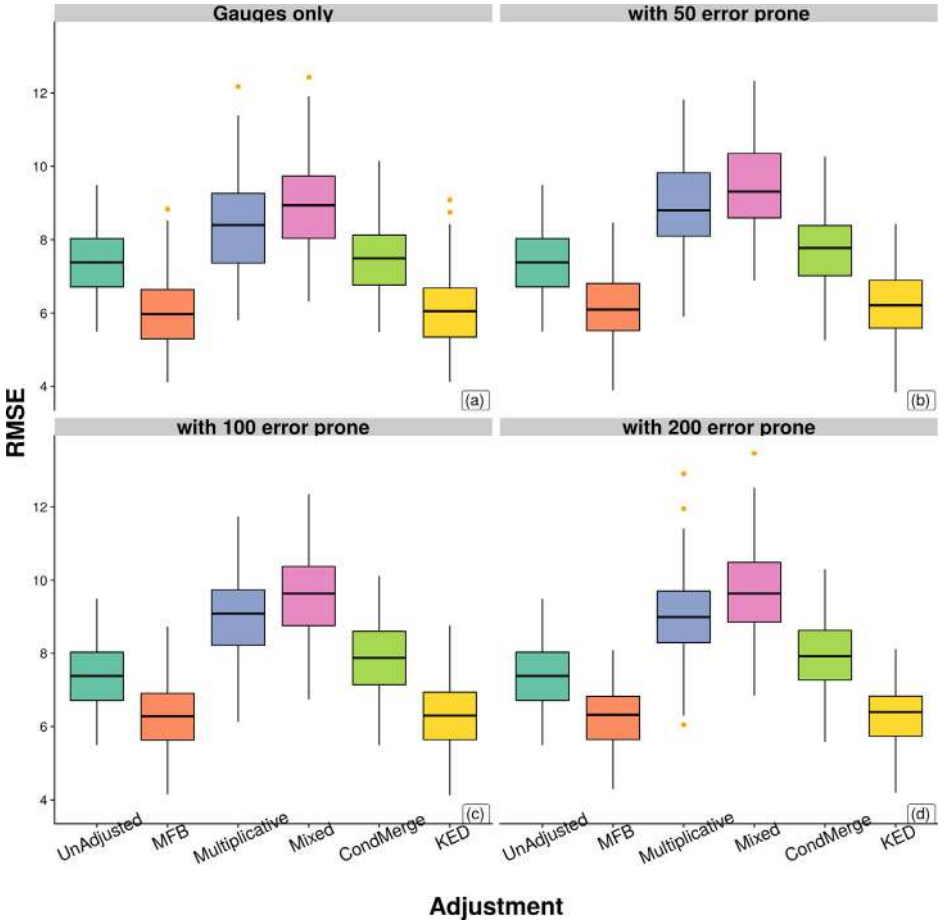

Fig. 14. Box-whisker plot of **RMSE** from 150 rain realizations for each adjustment with different numbers of error-prone rain sensors: (a) 50 gauges only, (b) including 50 error-prone sensors, (c) including 100 error-prone, (d) including 200 error-prone.

3. Results

Three sets of figures are presented below. First, images of the adjusted rain grids appear for certain representative adjustment procedures: MFB, CondMerge and KED (Figs. 4–6). Each figure includes adjustment results for all four densities of error-prone observations. These images show that the higher density of error-prone sensors (panels (d) in each Figure) are somewhat more speckled, whereas the gauges only images (panels (a)) more closely resemble the 'true' rain (refer to Fig. 2).

The next group (Figs. 7–12) show scatter plots of 'true' versus adjusted precipitation at the 400 validation points. Note that these plots represent the original precipitation value, before the Box-Cox transform (explained above in 2.5). Again each adjustment procedure appears four times for all densities of reliable/error-prone locations. Scatter plots in which the spread of points, and thus the regression line, approaches the 45° slope represent successful adjustment. Additionally, those scatter plots where dispersion of the points grows compared to the UnAdjusted plot (i.e. compare Fig. 7 to Fig. 8) indicate cases where the adjustment actually adds additional noise to the result.

The third set of figures, box-whisker plots of correlation statistics, best reveals the success of adjustment. All four statistics are presented: R^2, RMSE, slope and intercept of the fitted linear regression line. (Figs. 13–16). These plots show the mean values of each statistic (averaged over 150 realizations) as well as first and third quartiles at the height of the boxes. The whisker lines extend to 1.5 * *IQR* (interquartile range) and outliers appear as orange dots. The height of the boxes and length of the whiskers give a clear representation of the variance of each statistic across all realizations. RMSE values that appear in Fig. 14 with a low mean indicate successful adjustment, and short whisker lines mirror the dispersion of points in the scatter plots.

Results of all correlation statistics (averaged over 150 realizations) appear in Table 1.

4. Discussion

The intention in this work was to demonstrate a reproducible, synthetic test framework for evaluating weather radar adjustment procedures and to spotlight cases where additional, error-prone precipitation observations improve adjustment. The framework was constructed from multiple realizations of a synthetic rain grid using SGS. By choosing a large enough number of realizations, we ensured that no single rain or error distribution would skew the results due to randomization effects. Configurable parameters for radar noise structure, and four combinations of reliable and error-prone sensors, from zero error-prone to four times as many error-prone as reliable, were applied to complete the synthetic framework. The resulting radar grids were adjusted by five algorithms: MFB, Mixed, Multiplicative, CondMerge and KED. Then values of the 'true' rain and adjusted precipitation grids were extracted at 400 validation points. Means of several statistics comparing the adjusted and 'true' rain were then calculated.

Results were analyzed at two levels, localized and domain-wide. Higher R^2 values, and lower RMSE occur when there is a good average match at validation points between the 'true' rain and the adjusted values. The closeness of slope and intercept to the 45° "one–one" line reveals adjustment methods that achieve good bias correction, at the domain scale. However, when the slope nears 1.0 (45°), but the intercept does not go through zero, then multiplicative bias is corrected, but

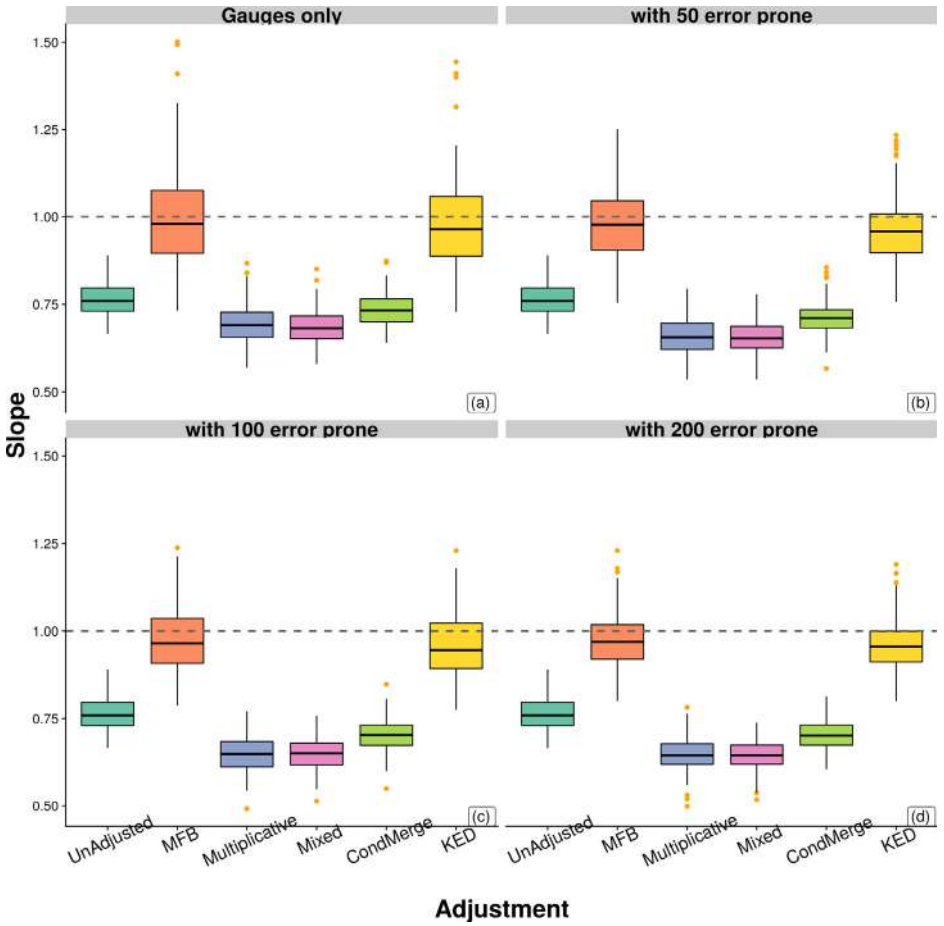

Fig. 15. Box-whisker plot of linear regression **slope** from 150 rain realizations for each adjustment with different numbers of error-prone rain sensors: (a) 50 gauges only, (b) including 50 error-prone sensors, (c) including 100 error-prone, (d) including 200 error-prone. The dotted line indicates slope of 1.0.

not additive bias. Thus methods with lower RMSE results are considered to attain better local adjustment to the 'true' rain. Slope values near 1.0 show good multiplicative bias correction, and if the intercept is also close to zero, then additive bias is also corrected.

From the scatter plots, clearly only two adjustment algorithms improved correspondence between the 'true' rain and adjusted grids: MFB and KED. Referring to the UnAdjusted scatter plot (Fig. 7), we note a somewhat high slope, and intercept value far from zero. This is expected due to the error fields applied to the true rain. The shift of regression lines in all adjustments closer to the 45° slope demonstrates correction of the multiplicative bias. However scatter plots of the adjustments show some increased dispersion of the points. Referring to the slope box-whisker plots (Fig. 15) the height of the boxes and whisker lines, especially for the MFB and KED adjustments, demonstrate this dispersion. The algorithms introduce additional noise into the adjusted grid, while correcting the overall bias.

Conversely, the Mixed, Multiplicative and CondMerge algorithms did not cause this added noise: their scatter plots and box whisker plots of slope and intercept are more compact. However the slope and intercept values for these adjustments are farther from the ideal. Furthermore, the Mixed and Multiplicative adjustments show slightly higher RMSE values and lower R^2 (Figs. 14 and 13) than the KED adjustment, also evident in the summary Table 1. We therefore determine that Mixed, Multiplicative and CondMerge perform satisfactorily at the local scale, avoiding additional noise in the correction procedure, but at the price of somewhat poorer overall correction of multiplicative bias. CondMerge displays similar, but slightly better results, with RMSE lower than both Mixed and Multiplicative, especially with the addition of error-prone sensors. Yet the slope and intercept values of CondMerge are not as good as MFB. So this "conditioning" algorithm reaches a good compromise with higher R^2, lower RMSE, but at the price of poorer bias correction.

Interestingly, the MFB adjustment achieves RMSE lower than the UnAdjusted grid for all densities of error-prone locations. The slope is slightly nearer to 45°, and the intercept close to zero. These results reflect the fact that MFB applies a domain wide uniform correction of bias, thus it is adjusting for multiplicative (but not additive) bias. We also note a slight compression of the size of the slope and intercept box whisker plots for MFB with higher density of error-prone observations. So this adjustment can gain somewhat from additional error-prone sensors.

The lowest RMSE values for all densities of error-prone sensors appear in the KED adjusted grid. Since KED takes into account radar values at the validation locations as the secondary trend variable, this result is expected. Furthermore KED shows the highest R^2 among all methods, and at all densities of error-prone sensors; in fact both RMSE and R^2 improve as the number of error-prone sensors increases. We conclude, therefore, that KED is resistant to, and actually benefits from a large number of error-prone precipitation measurements.

It is worthwhile to focus on correlation statistics that display a decrease in variation as the number of error-prone sensors grows. This is evident from the size of the box and whisker plots. The variation in R^2 decreases somewhat for KED at the 100 and 200 error-prone (Fig. 13 panels (c) and (d)) densities. KED shows a more pronounced decrease in the variation of RMSE (Fig. 14 panel (d)) at the highest density of error-prone. This result is also evident in the scatter plot (Fig. 12 panel (d)); the regression line tightly matches the "one–one" 45° slope, yet the dispersion of points decreases with more error-prone observations.

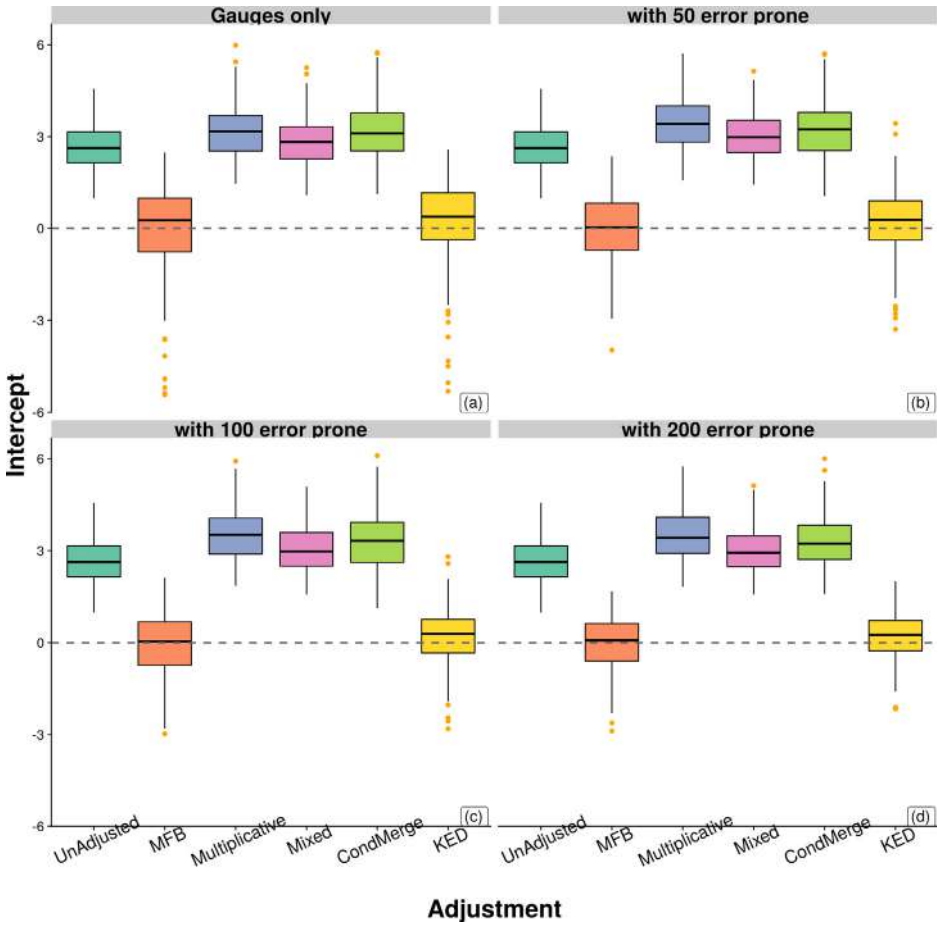

Fig. 16. Box-whisker plot of linear regression **intercept** from 150 rain realizations for each adjustment with different numbers of error-prone rain sensors: (a) 50 gauges only, (b) including 50 error-prone sensors, (c) including 100 error-prone, (d) including 200 error-prone. The dotted line indicates intercept of 0.0.

Among the kriging based algorithms, KED displays a good match to the ideal slope and intercept values, with a slope almost as high as MFB, and intercept very near to zero. Thus, KED achieves the best local adjustment as well as good correction for bias, and even improves with a large number of error-prone sensors.

These results are in good agreement with Goudenhoofdt and Delobbe (2009), Berndt et al. (2014) and Erdin et al. (2012) and others who consistently showed that KED achieved better correspondence between adjusted and observed precipitation. In parallel, research by Sinclair and Pegram (2005) and Kim et al. (2007) showed that the CondMerge algorithm successfully adjusted weather radar grids. In this current work we found that the non-linear patterns in the synthetic radar grid are not well adjusted by additive techniques such as CondMerge and Mixed.

In constructing the synthetic test framework, examination of radar adjustment procedures using different densities of error-prone sensors, with a range of error levels was made possible. Past research (Goudenhoofdt and Delobbe (2009); Wang et al. (2012); McKee and Binns (2016); Bruno et al. (2014)) conducted evaluations of adjustment procedures, but using data from a specific gauge and radar network under certain storm events. For example, work by Berg et al. (2016) was located in Sweden, while the recent research by Foehn et al. (2018) was conducted in Switzerland. However a generic evaluation framework, to the best of our knowledge, is lacking. This current work attempts to fill that gap by offering such a standardized test framework.

Furthermore, two sets of ground based observations were combined in the synthesized framework: reliable and error-prone, with tests for different densities of error-prone observations. In previous research both Foehn et al. (2018) and Sideris et al. (2014) chose co-kriging to adjust radar with two sets of ground observations. The procedure chosen by Sideris et al. (2014) merged radar and gauge observations at two time steps, but did not address different error levels. On the other hand Foehn et al. (2018) applied co-kriging to two independent gauge networks with different levels of reliability. The less reliable network in that work consisted of a smaller number of gauges than the reliable network. While this current work does not implement co-kriging, two sets of ground observations, one reliable and the other error-prone, were merged. The proportion of error-prone sensors in this work was varied up to four times the number of reliable sensors.

The procedure and results in this work demonstrate that a large number of error-prone precipitation observations can be combined with reliable gauge observations, and used for radar adjustment. The analysis presented is reproducible, and configuration parameters can be chosen to match a real world situation. Application of this method can guide research and operational QPE to chose the optimal mix of reliable and error-prone observations and the most appropriate adjustment method.

The authors are aware of additional parameters that might affect adjustment procedures, that are not included in this work. Topography can both interfere with radar signals through beam blockage, and can influence storm intensity by orographic forcing. Germann et al. (2006) discussed the challenges facing weather radar in mountainous regions, and Xu et al. (2017) compared ground based validation of satellite borne radar at different elevation levels. Yet topography was not taken into account in this work since such effects are very localized. No generic, synthetic framework could include topography without causing a dramatic local influence on the results. The underlying assumption in this work considers the radar rainfall to be corrected for

beam blockage, and we further assume that the synthetic rain grids already included effects due to orographic forcing.

Furthermore, the temporal resolution of radar data, compared to gauge or CML data can also negatively influence adjustment efforts. As discussed by Cristiano et al. (2017), Emmanuel et al. (2012) and Marra and Morin (2018), radar images capture a single point in time "snapshot" whereas gauges represent aggregated rainfall over some (albeit short) time interval. It is possible that radar misses a short lived convective storm, or that the storm moves between one radar sweep and the next. What's more, radar views the rain rate aloft while ground-based sensors collect precipitation at the surface. There could be a time lag between the two, introducing additional uncertainty into adjustment algorithms. To overcome these temporal effects, adjustment parameters are determined and methods are chosen based on very long aggregations of radar and gauge data. Within the context of this synthetic framework, the assumed time resolution was (at least) one day.

The initial 'true' rain, created as described in Section 2.4, was based on 150 realizations of Sequential Gaussian Simulations. However the final result was a single 'true' rain grid. The evaluations examined only levels of error in the error-prone gauges. No attempt was made to evaluate various types of storms.

Further research using the synthetic framework presented here could investigate three issues. First, various types of rain storms could be created by altering the parameters when preparing the initial 'true' rain. Next, different error distributions could be applied to the error-prone subset of observations. Finally, additional adjustment algorithms, such as co-kriging, as well as combinations of algorithms could be implemented and compared.

5. Conclusions

This research presented a standardized, reproducible test framework that was constructed to analyze adjustment of synthetic weather radar rainfall. Several adjustment algorithms were applied to the radar by using synthesized ground based observation data representing a mix of reliable and error-prone sensors at varying densities. Then statistical tests were used to validate the success of these algorithms in reconstructing the 'true' rain. The framework included configurable parameters: density and error level of ground observations, and the noise structure of the radar grid.

Results indicate:

- All adjustment procedures perform fairly well in correcting for multiplicative bias.
- The Conditional Merge performs satisfactorily with addition of error-prone observations, attaining lower RMSE than other additive algorithms.
- Kriging with External Drift is the most resistant to a high number of error-prone observations.
- Mean Field Bias performs fairly well at all densities of error-prone correcting only bias, as expected, but adding some noise (variance) into the adjusted grid.
- Mixed and Multiplicative adjustments avoid addition of noise into the adjusted grid, but tend to poorer overall correction (higher RMSE) than all other adjustments.

Acknowledgements

Erick Fredj thanks the DFG, German Research Foundation, and IMAP for partial support of this work.

References

Ahrens, B., 2006. Distance in spatial interpolation of daily rain gauge data. Hydrol. Earth Syst. Sci. Discussions 10, 197–208. URL: https://hal.archives-ouvertes.fr/hal-00304836/.

Amorati, R., Alberoni, P.P., Fornasiero, A., 2012. Operational Bias Correction of Hourly Radar Precipitation Estimate using Rain Gauges. http://www.meteo.fr/cic/meetings/2012/ERAD/extended_abs/QPE_007_ext_abs.pdf.

Berg, P., Norin, L., Olsson, J., 2016. Creation of a high resolution precipitation data set by merging gridded gauge data and radar observations for Sweden. J. Hydrol. 541, 6–13. https://doi.org/10.1016/j.jhydrol.2015.11.031.. URL: http://linkinghub.elsevier.com/retrieve/pii/S002216941500918X.

Berndt, C., Rabiei, E., Haberlandt, U., 2014. Geostatistical merging of rain gauge and radar data for high temporal resolutions and various station density scenarios. J. Hydrol. 508, 88–101. https://doi.org/10.1016/j.jhydrol.2013.10.028.

Berne, A., Uijlenhoet, R., 2007. Path-averaged rainfall estimation using microwave links: Uncertainty due to spatial rainfall variability. Geophys. Res. Lett. 34. https://doi.org/10.1029/2007GL029409.

Bianchi, B., Jan van Leeuwen, P., Hogan, R.J., Berne, A., 2013. A variational approach to retrieve rain rate by combining information from rain gauges, radars, and microwave links. J. Hydrometeorol. 14, 1897–1909. https://doi.org/10.1175/JHM-D-12-094.1.

Bowler, N.E., Pierce, C.E., Seed, A.W., 2006. STEPS: a probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. Q.J.R. Meteorol. Soc. 132, 2127–2155. https://doi.org/10.1256/qj.04.100.

Box, G.E.P., Cox, D.R., 1964. An analysis of transformations. J. R. Stat. Soc. Series B (Methodological) 26, 211–252. URL: http://www.jstor.org/stable/2984418.

Bruno, F., Cocchi, D., Greco, F., Scardovi, E., 2014. Spatial reconstruction of rainfall fields from rain gauge and radar data. Stoch. Env. Res. Risk Assess. 28, 1235–1245. https://doi.org/10.1007/s00477-013-0812-0.

Cecinati, F., Moreno-Ródenas, A., Rico-Ramirez, M., ten Veldhuis, M.C., Langeveld, J., 2018. Considering rain gauge uncertainty using Kriging for uncertain. Data. Atmosphere 9, 446. https://doi.org/10.3390/atmos9110446.

Cecinati, F., Wani, O., Rico-Ramirez, M.A., 2016. Evaluation and correction of uncertainty due to Gaussian approximation in radar – rain gauge merging using kriging with external drift. AGU Fall Meeting Abstracts.

Chumchean, S., Sharma, A., Seed, A., 2003. Radar rainfall error variance and its impact on radar rainfall calibration. Phys. Chem. Earth, Parts A/B/C 28, 27–39. https://doi.org/10.1016/S1474-7065(03)00005-6.

Chwala, C., Kunstmann, H., Hipp, S., Siart, U., Eibert, T., 2012. Precipitation observation using commercial microwave communication links, in: 2012 IEEE International Geoscience and Remote Sensing Symposium, IEEE. pp. 2922–2925. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6350714.

Ciach, G.J., Krajewski, W.F., Villarini, G., 2007. Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD. Data. J. Hydrometeor 8, 1325–1347. https://doi.org/10.1175/2007JHM814.1.

Colli, M., Lanza, L., La Barbera, P., 2013. Performance of a weighing rain gauge under laboratory simulated time-varying reference rainfall rates. Atmos. Res. 131, 3–12. https://doi.org/10.1016/j.atmosres.2013.04.006.

Cristiano, E., ten Veldhuis, M.C., van de Giesen, N., 2017. Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas - a review. Hydrol. Earth Syst. Sci. 21, 3859–3878. https://doi.org/10.5194/hess-21-3859-2017.

Delrieu, G., Wijbrans, A., Boudevillain, B., Faure, D., Bonnifait, L., Kirstetter, P.E., 2014. Geostatistical radar-raingauge merging: a novel method for the quantification of rain estimation accuracy. Adv. Water Resour. 71, 110–124. https://doi.org/10.1016/j.advwatres.2014.06.005.

Edouard, S., Vincendon, B., Ducrocq, V., 2018. Ensemble-based flash-flood modelling: taking into account hydrodynamic parameters and initial soil moisture uncertainties. J. Hydrol. 560, 480–494. https://doi.org/10.1016/j.jhydrol.2017.04.048.. URL: http://linkinghub.elsevier.com/retrieve/pii/S0022169417302688.

Emmanuel, I., Andrieu, H., Leblois, E., Flahaut, B., 2012. Temporal and spatial variability of rainfall at the urban hydrological scale. J. Hydrol. 430–431, 162–172. https://doi.org/10.1016/j.jhydrol.2012.02.013.

Erdin, R., Frei, C., Künsch, H.R., 2012. Data transformation and uncertainty in geostatistical combination of radar and rain gauges. J. Hydrometeorol. 13, 1332–1346. https://doi.org/10.1175/JHM-D-11-096.1.

Fencl, M., Dohnal, M., Rieckermann, J., Bare, V., 2017. Gauge-adjusted rainfall estimates from commercial microwave links. Hydrol. Earth Syst. Sci. 21, 617–634. https://doi.org/10.5194/hess-21-617-2017.

Foehn, A., Hernndez, J.G., Schaefli, B., Cesare, G.D., 2018. Spatial interpolation of precipitation from multiple rain gauge networks and weather radar data for operational applications in Alpine catchments. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2018.05.027.. URL: http://linkinghub.elsevier.com/retrieve/pii/S0022169418303548.

Fox, J., Weisberg, S., 2011. An R Companion to Applied Regression, Second ed. Sage, Thousand Oaks CA. URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.

Germann, U., Galli, G., Boscacci, M., Bolliger, M., 2006. Radar precipitation measurement in a mountainous region. Q. J. R. Meteorolog. Soc. 132, 1669–1692. https://doi.org/10.1256/qj.05.190.

Gjertsen, U., Salek, M., Michelson, D.B., 2004. Gauge adjustment of radar-based precipitation estimates in Europe. In: Proceedings of ERAD. URL: http://copernicus.org/erad/2004/online/ERAD04_P_7.pdf.

Goldshtein, O., Messer, H., Zinevich, A., 2009. Rain rate estimation using measurements from commercial telecommunications links. IEEE Trans. Signal Process. 57, 1616–1625. https://doi.org/10.1109/TSP.2009.2012554.. URL: http://ieeexplore.ieee.org/document/4749357/.

Goudenhoofdt, E., Delobbe, L., 2009. Evaluation of radar-gauge merging methods for quantitative precipitation estimates. Hydrol. Earth Syst. Sci. 13, 195–203. URL: www.hydrol-earth-syst-sci.net/13/195/2009/.

Guenzi, D., Fratianni, S., Boraso, R., Cremonini, R., 2016. CondMerg: an open source implementation in R language of conditional merging for weather radars and rain gauges observations. Earth Sci. Inf. https://doi.org/10.1007/s12145-016-0278-y.

Hasan, M.M., Sharma, A., Johnson, F., Mariethoz, G., Seed, A., 2016. Merging radar and in situ rainfall measurements: An assessment of different combination algorithms: an assessment of different combination algorithms. Water Resour. Res. 52, 8384–8398. https://doi.org/10.1002/2015WR018441.

Kidd, C., Becker, A., Huffman, G.J., Muller, C.L., Joe, P., Skofronick-Jackson, G., Kirschbaum, D.B., 2017. So, how much of the Earth's surface is covered by rain gauges? Bull. Am. Meteorol. Soc. 98, 69–78. https://doi.org/10.1175/BAMS-D-14-00283.1.

Kim, B.S., Hong, J.B., Kim, H.S., Yoon, S.Y., 2007. Combining radar and rain gauge rainfall estimates for flood forecasting using conditional merging method. Am. Soc. Civil Eng. 1–16. https://doi.org/10.1061/40927(243)414.

Krajewski, W., 1987. Cokriging radar-rainfall and rain gage data. J. Geophys. Res. 92, 9571–9580. https://doi.org/10.1029/JD092iD08p09571.

Krajewski, W.F., Vignal, B., Seo, B.C., Villarini, G., 2011. Statistical model of the range-dependent error in radar-rainfall estimates due to the vertical profile of reflectivity. J. Hydrol. 402, 306–316. https://doi.org/10.1016/j.jhydrol.2011.03.024.. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022169411002046.

Leijnse, H., Uijlenhoet, R., Berne, A., 2010. Errors and uncertainties in microwave link rainfall estimation explored using drop size measurements and high-resolution radar data. J. Hydrometeorol. 11, 1330–1344. https://doi.org/10.1175/2010JHM1243.1.

van Leth, T.C., Overeem, A., Leijnse, H., Uijlenhoet, R., 2018. A measurement campaign to assess sources of error in microwave link rainfall estimation. Atmospheric Meas. Tech. 11, 4645–4669. https://doi.org/10.5194/amt-11-4645-2018.

Ly, S., Charles, C., Degré, A., 2013. Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale. A review. Biotechnol. Agron. Soc. Environ, 15.

Marra, F., Morin, E., 2018. Autocorrelation structure of convective rainfall in semiarid-arid climate derived from high-resolution X-Band radar estimates. Atmospheric Res. 200, 126–138. https://doi.org/10.1016/j.atmosres.2017.09.020.. URL: http://linkinghub.elsevier.com/retrieve/pii/S0169809517309936.

Mazzetti, C., Todini, E., 2004. Combining Raingages and Radar Precipitation Measurements Using a Bayesian Approach. geoENV IV Geostatistics for Environmental Applications, 401–412.

Mazzetti, C., Todini, E., 2008. Combining Weather Radar and Raingauge Data for Hydrologic Applications. In: Flood Risk Management: Research and Practice Taylor & Francis Group, Samuels.

McKee, J.L., Binns, A.D., 2016. A review of gauge-radar merging methods for quantitative precipitation estimation in hydrology. Canadian Water Resources Journal/ Revue canadienne des ressources hydriques 41, 186–203. https://doi.org/10.1080/07011784.2015.1064786.

Michelson, D., Koistinen, J., 2000. Gauge-Radar network adjustment for the baltic sea experiment. Phys. Chem. Earth Part B: Hydrol., Oceans Atmosphere 25, 915–920. https://doi.org/10.1016/S1464-1909(00)00125-8.

Morin, E., Gabella, M., 2007. Radar-based quantitative precipitation estimation over Mediterranean and dry climate regimes. J. Geophys. Res. 112. https://doi.org/10.1029/2006JD008206.

Musayev, S., Burgess, E., Mellor, J., 2018. A global performance assessment of rainwater harvesting under climate change. Resour. Conserv. Recycl. 132, 62–70. https://doi.org/10.1016/j.resconrec.2018.01.023.. URL: http://linkinghub.elsevier.com/retrieve/pii/S0921344918300235.

Nerini, D., Besic, N., Sideris, I., Germann, U., Foresti, L., 2017. A non-stationary stochastic ensemble generator for radar rainfall fields based on the short-space Fourier transform. Hydrol. Earth Syst. Sci. 21, 2777–2797. https://doi.org/10.5194/hess-21-2777-2017.

Nour, M.H., Smit, D.W., Gamal El-Din, M., 2006. Geostatistical mapping of precipitation: implications for rain gauge network design. Water Sci. Technol. 53, 101–110. https://doi.org/10.2166/wst.2006.303. URL: /wst/article/53/10/101/11888/Geostatistical-mapping-of-precipitation.

Otieno, H., Yang, J., Liu, W., Han, D., 2014. Influence of Rain Gauge Density on Interpolation Method Selection. J. Hydrol. Eng. 19, 04014024. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000964.

Overeem, A., Leijnse, H., Uijlenhoet, R., 2013. Country-wide rainfall maps from cellular communication networks. PNAS 110, 2741–2745. https://doi.org/10.1073/pnas.1217961110.

Overeem, A., Leijnse, H., Uijlenhoet, R., 2016. Two and a half years of country-wide rainfall maps using radio links from commercial cellular telecommunication networks: two and a half years of radio link rainfall maps. Water Resour. Res. 52, 8039–8065. https://doi.org/10.1002/2016WR019412.

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Comput. Geosci. 30, 683–691.

Pfaff, T., 2010. Radargesttzte Schötzung von Niederschlagsensembles. Technical Report. URL: http://www.rimax-hochwasser.de/fileadmin/user_uploads/RIMAX_PUB_22_0015_Abschlussbericht.

R Development Core Team, 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: http://www.R-project.org. ISBN 3-900051-07-0.

Rabiei, E., Haberlandt, U., 2015. Applying bias correction for merging rain gauge and radar data. J. Hydrol. 522, 544–557. https://doi.org/10.1016/j.jhydrol.2015.01.020.. URL: http://linkinghub.elsevier.com/retrieve/pii/S0022169415000372.

Schiemann, R., Erdin, R., Willi, M., Frei, C., Berenguer, M., Sempere-Torres, D., 2011. Geostatistical radar-raingauge combination with nonparametric correlograms: methodological considerations and application in Switzerland. Hydrol. Earth Syst. Sci. 15, 1515–1536. https://doi.org/10.5194/hess-15-1515-2011.

Sebastianelli, S., Russo, F., Napolitano, F., Baldini, L., 2012. Assessment of an adjustment factor to model radar range dependent error. URL: https://journals.ametsoc.org/doi/pdf/10.1175/1520-0450 doi: 10.1063/1.4756523.

Seed, A., 2002. A dynamic and spatial scaling approach to advection forecasting. J. Appl. Meteorol https://journals.ametsoc.org/doi/pdf/10.1175/1520-0450.

Seo, D.J., 1998. Real-time estimation of rainfall fields using radar rainfall and rain gage data. J. Hydrol. 208, 37–52. URL: http://www.sciencedirect.com/science/article/pii/S0022169498001413.

Shafiei, M., Ghahraman, B., Saghafian, B., Pande, S., Gharari, S., Davary, K., 2014. Assessment of rain-gauge networks using a probabilistic GIS based approach. Hydrol. Res. 45, 551. https://doi.org/10.2166/nh.2013.042.

Sideris, I.V., Gabella, M., Erdin, R., Germann, U., 2014. Real-time radar-rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland: Real-time radar-rain-gauge merging. Q. J. R. Meteorological Soc. 140, 1097–1111. https://doi.org/10.1002/qj.2188.

Sinclair, S., Pegram, G., 2005. Combining radar and rain gauge rainfall estimates using conditional merging. Atmospheric Sci. Lett. 6, 19–22. https://doi.org/10.1002/asl.85.

Tang, L., Tian, Y., Yan, F., Habib, E., 2015. An improved procedure for the validation of satellite-based precipitation estimates. Atmos. Res. 163, 61–73. https://doi.org/10.1016/j.atmosres.2014.12.016.. URL: https://linkinghub.elsevier.com/retrieve/pii/S0169809515000058.

Teo, C.K., Grimes, D.I., 2007. Stochastic modelling of rainfall from satellite data. J. Hydrol. 346, 33–50. https://doi.org/10.1016/j.jhydrol.2007.08.014.. URL: https://linkinghub.elsevier.com/retrieve/pii/S002216940700460X.

Tian, F., Hou, S., Yang, L., Hu, H., Hou, A., 2018. How does the evaluation of the GPM IMERG rainfall product depend on gauge density and rainfall intensity? J. Hydrometeorol. 19, 339–349. https://doi.org/10.1175/JHM-D-17-0161.1.

Tian, Y., Huffman, G.J., Adler, R.F., Tang, L., Sapiano, M., Maggioni, V., Wu, H., 2013. Modeling errors in daily precipitation measurements: Additive or multiplicative?: modeling errors in precipitation data. Geophys. Res. Lett. 40, 2060–2065. https://doi.org/10.1002/grl.50320.

Todini, E., 2001. A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements. Hydrol. Earth Syst. Sci. Discussions 5, 187–199.

Velasco-Forero, C.A., Sempere-Torres, D., Cassiraga, E.F., Jaime Gómez-Hernández, J., 2009. A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data. Adv. Water Resour. 32, 986–1002. https://doi.org/10.1016/j.advwatres.2008.10.004.. URL: http://linkinghub.elsevier.com/retrieve/pii/S0309170808001760.

Vignal, B., Krajewski, W.F., 2001. Large-Sample Evaluation of Two Methods to Correct Range-Dependent Error for WSR-88d Rainfall Estimates. J. Hydrometeor. 2, 490–504. https://doi.org/10.1175/1525-7541(2001)002<0490:LSEOTM>2.0.CO;2. URL: https://journals.ametsoc.org/doi/full/10.1175/1525-7541%282001%29002%3C0490%3ALSEOTM%3E2.0.CO%3B2.

Villarini, G., Krajewski, W., 2009a. Inference of spatial scaling properties of rainfall: impact of radar rainfall estimation uncertainties. IEEE Geosci. Remote Sens. Lett. 6, 812–815. https://doi.org/10.1109/LGRS.2009.2025891.

Villarini, G., Krajewski, W.F., 2009b. Empirically based modelling of radar-rainfall uncertainties for a C-band radar at different time-scales. Q. J. R. Meteorological Soc. 135, 1424–1438. https://doi.org/10.1002/qj.454.

Villarini, G., Krajewski, W.F., 2010. Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. Surveys Geophys. 31, 107–129. https://doi.org/10.1007/s10712-009-9079-x. http://link.springer.com/10.1007/s10712-009-9079-x.

Wang, L.P., Ochoa, S., Simoes, N., Onof, C., Maksimovic, C., 2012. Radar-raingauge data combination techniques: a revision and analysis of their suitability for urban hydrology. Belgrade. URL: http://www.raingain.eu/sites/default/files/udm_2012_wang.pdf.

Xu, R., Tian, F., Yang, L., Hu, H., Lu, H., Hou, A., 2017. Ground validation of GPM IMERG and TRMM 3b42v7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network: validation of GPM and TRMM Over TP. J. Geophys. Res.: Atmospheres 122, 910–924. https://doi.org/10.1002/2016JD025418.

Yang, P., Ng, T.L., 2019. Fast Bayesian regression Kriging method for real time merging of radar, rain gauge, and crowdsourced rainfall data. Water Resour. Res. https://doi.org/10.1029/2018WR023857. 2018WR023857.

Zinevich, A., Messer, H., Alpert, P., 2010. Prediction of rainfall intensity measurement errors using commercial microwave communication links. Atmospheric Measure. Tech. 3, 1385–1402. https://doi.org/10.5194/amt-3-1385-2010.