Citizen rain gauges improve hourly radar rainfall bias correction using a two-step Kalman filter

Punpim Puttaraksa Mapiam [1*], Monton Methaprayun [1], Thom Adrianus Bogaard [2], Gerrit Schoups [2], Marie-Claire Ten Veldhuis [2]

[1]Department of Water Resources Engineering, Kasetsart University, PO Box 1032, Bangkok, 10900, Thailand
[2]Department of Water Management, Delft University of Technology, PO Box 5048, 2600 GA Delft, The Netherlands

Correspondence to: Punpim P. Mapiam (punpim.m@ku.th)

Abstract. Low density of conventional rain gauge networks is often a limiting factor for radar rainfall bias correction. Citizen rain gauges offer a promising opportunity to collect rainfall data at higher spatial density. In this paper hourly radar rainfall bias adjustment was applied using two different rain gauge networks consisting of tipping buckets (measured by Thai Meteorological Department, TMD) and daily citizen rain gauges. The radar rainfall bias correction factor was sequentially updated based on TMD and citizen rain gauge data using a two-step Kalman filter to incorporate the two-gauge datasets of contrasting quality. Radar reflectivity data of Sattahip radar station, gauge rainfall data from the TMD and citizen rain gauges located in Tubma basin, Thailand were used in the analysis. Daily data from the citizen rain gauge network were downscaled to hourly resolution based on temporal distribution patterns obtained from radar rainfall time series and the TMD gauge network. Results show that an improvement of radar rainfall estimates was achieved by including the downscaled citizen observations compared to bias correction based on the conventional rain gauge network only. These outcomes emphasize the value of citizen rainfall observations for radar bias correction, in particular in regions where conventional rain gauge networks are sparse.

Keywords: Kalman filter, citizen rain gauge, radar rainfall, bias correction, downscaling, Thailand.

1 Introduction

Hydrometeorological hazards, like flash floods and landslides cause severe damage to economies, properties, and human lives worldwide. In this context, flood forecasting and warning systems are a valuable non-structural measure to mitigate damage. However, such systems require input of rainfall data at a high spatial and temporal resolution. In most regions of the world, automatic rain gauge networks are insufficient for this purpose. Weather radar, which can better capture the variation of rainfall fields at fine spatial and temporal resolutions could be used as an alternative rainfall product for improving the accuracy of flash flood estimates and warnings. (Collinge and Kirby, 1987; Sun et al., 2000; Uijlenhoet 2001; Bedient et al., 2003; Creutin and Borga, 2003; Mapiam et al., 2009a, 2014; Mapiam and Chautsuk, 2018; Corral et al., 2019). However, a weather radar provides an indirect measurement of backscattered electromagnetic waves called radar reflectivity data (Z)

and quantitative estimation of radar rainfall data (R) is acknowledged as a complex process. Various sources of errors affect radar rainfall estimates, mainly errors in reflectivity measurements and reflectivity-rainfall conversion (Jordan et al., 2000). Correction of these two sources of error is a crucial procedure to increase the accuracy of radar rainfall estimates. Ground-truthing by rain gauge data is required to calibrate the Z-R relationship ($Z=AR^b$). The calibrated parameter *A* in the *Z-R*

35 relationship will include any errors in the radar system caused by the electrical calibration of the radar (Seed et al., 2002).

The calibrated Z-R equation is used to convert the measured instantaneous reflectivity data to rainfall intensity and thereafter accumulating them into the required temporal resolution. However, the parameters A and b vary significantly, even within a single storm event depending on the rainfall characteristics which can exhibit a highly dynamic raindrop size distribution (DSD) (Ulbrich, 1983: Smith et al., 2009). Additionally, past studies found that the Z-R parameters are sensitive

40 to the temporal resolution of rain gauge rainfall data that is used for the Z-R calibration (Hitchfeld and Bordan, 1954; Smith et al., 1975; Wilson and Brandes, 1979; Klazura, 1981; Steiner et al., 1995; Mapiam and Sriwongsitanon, 2008; Mapiam et al., 2009b). Consequently, an important source of error remains associated with the Z-R conversion process (Jordan et al., 2000; Berne and Krajewski, 2013). Many researchers attempted to correct this kind of error by classification of the measured reflectivity data into different storm types and thereafter constructing the Z-R equation corresponding to the classified storm

45 characteristics. (Joss and Waldvogel, 1970; Rogers, 1971; Battan, 1973; Klazura, 1981; Austin, 1987; Rosenfeld et al., 1992, 1993; Tokay and Short, 1996; Amitai, 2000; Arai et al., 2005; Fang et al., 2018). For the effect of using rain gauge data with different temporal resolutions on Z–R relationships, Mapiam et al. (2009b) developed a universal scaling transformation function for converting the reference parameters A (obtained from using daily gauge rainfall data in the calibration) to the parameter A for sub-daily resolutions. This improved accuracy of the estimated sub-daily radar rainfall, especially in locations

50 with limited short-duration rain gauge measurements.

After Z-R conversion, bias is expected to remain between the assessed radar rainfall and the true rainfall amount at the rain gauge locations if a fixed Z-R relationship is used to estimate radar rainfall over the entire radar domain (Chumchean et al., 2006; Wang et al., 2015). An effective bias correction technique is key for enhancing the quality of radar rainfall estimates (Steiner et al., 1999) and to remove the residual errors between radar rainfall obtained from the Z-R relationship and

55 rain gauge data. Mean field bias (MFB) adjustment is the conventional method to obtain a static bias factor which assumes that the Z-R relationship is homogeneous in space but varies in time (Smith et al., 2007; Vieux and Bedient, 2004; Wilson, 1970). In this method, a multiplicative correction factor is applied uniformly across the radar coverage. Since the MFB approach does not consider noise and uncertainty of the rain gauge observations, nor spatial variability in observation bias, this can lead to large errors in radar rainfall estimates, particularly in areas where the density of rain gauge networks is limited.

60 Kalman filter (KF) is an efficient algorithm that has been applied to correct the spatially uniform mean field bias, especially in real-time by accounting for the temporal variation of the mean bias as well as uncertainties in the ground rainfall measurements (Ahnert, 1986; Smith and Krajewski, 1991; Anagnostou et al., 1998; Seo et al., 1999; Dinku et al., 2002; Chumchean et al., 2006).

Previous studies used the KF for predicting and correcting the mean field bias to mitigate the observation error variances affecting the mean field bias estimate. Chumchean et al. (2006) found that the density of the rain gauge network also plays an important role in the radar rainfall bias adjustment. They found that lowering the density of rain gauge observations in the KF process reduced accuracy of radar rainfall estimates. Additionally, the KF approach outperforms the use of MFB if rain gauge density is less than 1 per 90 km^2, and both KF and MFB produce identical performance when the rain gauge density is greater than 1 per 70 km^2.

In basins where a dense rainfall network is not available, Citizen Science (CS) offers a promising opportunity for enhancing the density of rainfall observations (Davids, et al., 2019). With the popularization of smartphones and the availability of (relatively) simple and cheap equipment, abundant mobile applications and projects have been initiated in Water Resources Management to measure hydrometeorological variables like rainfall, water level height or water quality, as well as to ground-truth remotely sensed information on e.g. land use (Srivastra et al., 2018; Davids et al., 2019; See, 2019; Seibert et al., 2019). In the current study, we focus on rainfall measured by local citizens using a network of cheap rain gauges and a specially designed mobile application. Since citizen rainfall observations are typically provided at daily scale, a temporal downscaling technique is needed for sub-daily applications. There has been a variety of temporal rainfall downscaling methods developed since the 1970s. The simplest approach is to distribute daily rainfall data to sub-daily resolutions by assuming uniform distributions. Stochastically generating sub-period data or spatially transferring finer resolution rainfall from a nearby rain gauge station to the study area based on spatial correlations are alternative approaches (Koutsoyiannis, 2003; Debele et al., 2007). However, these methods are usually not designed for real-time data disaggregation over large areas. Instead, a common approach for such scenarios is to downscale daily rainfall based on a simple fraction technique by considering the distribution patterns of high-resolution gridded rainfall products from radar or satellite sensors (Paulat et al., 2008; Wüest et al., 2010; Vormoor and Skaugen, 2013; Sideris et al., 2014; Barton et al., 2019). This study aimed to modify the KF logic by integrating hourly rain gauge data with daily citizen rain gauge data that are downscaled to hourly time scale using a simple fraction method. The radar rainfall bias correction factor was sequentially updated using a two-step Kalman filter accounting for the contrasting quality of the hourly rain gauge data and downscaled citizen rain gauge data. The question we set out to answer is to what extent the downscaled citizen rainfall observations improve the accuracy of hourly radar rainfall estimates. Several scenarios of hourly rainfall distribution patterns were applied for downscaling to investigate the most suitable technique for hourly radar rainfall assessment. Tubma basin located in Rayong province, eastern Thailand, was used as a case study area to test the approach.

2. Study Area and Data

2.1 Study Area

The study area is the Tubma basin located in Rayong province, eastern Thailand, situated between latitude 12.6789 to 12.8775, and longitude 101.0881 to 101.2975 (Fig. 1). It covers a catchment area of 197 km^2 with basin elevation ranging

from 4 to 416 m MSL. The main river, Klong Tubma, is 42 km in length and originates in Chom Hae, Kate, and Kra Bok mountains and flows downstream to the northwest before meeting the Gulf of Thailand at Pak Nam district. The Tubma watershed is susceptible to flooding, in particular Rayong. In Figure 1, we show the climatological variation across the study area and its surroundings, based on 30-year (1987-2017) annual mean rainfall from the network of 311 daily rain gauges owned by the TMD and situated within 200 km range from the Tubma basin. Spatial rainfall patterns were generated by inverse distance squared (IDS) between the gauge locations. The map shows that while there is a small gradient in mean annual rainfall (1,100 to 1,700 mm mean annual rainfall) across the area of Rayong and Chonburi provinces (within 90 km from the study area), changes are more pronounced when the distance exceeds beyond the 90 km boundary, especially to the east of the study area. This is because these areas are affected differently by the southwest monsoon. Consequently, evaluating the effectiveness of bias correction techniques was carried out within a 90 km radius from the study area with similar climatology.

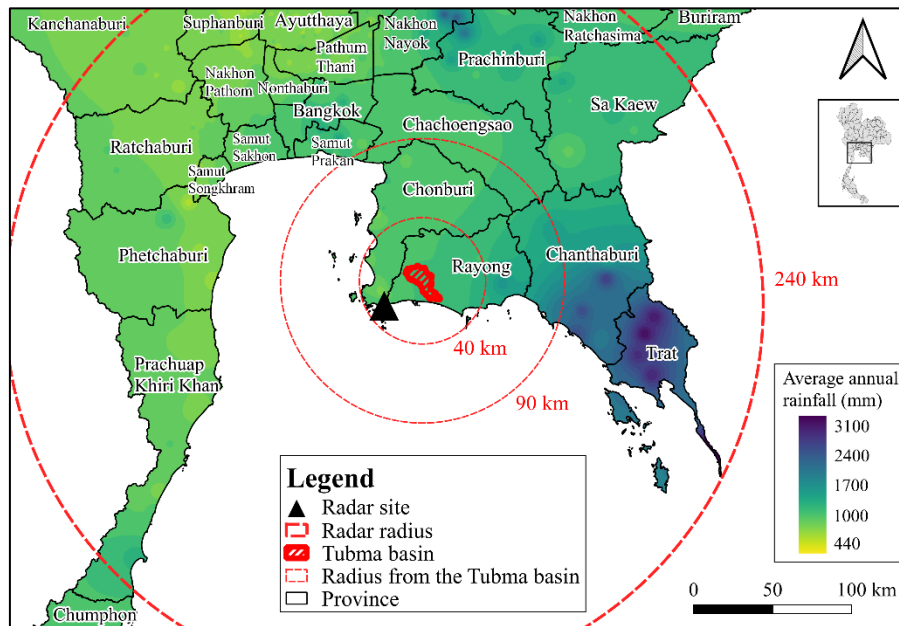



Figure 1: Climatological spatial rainfall distribution in and around the Tubma basin calculated from 30-year average annual rainfall data of 311 daily rain gauge network by using IDS method.

2.2 Radar Data

2.2.1 Reflectivity Data Collection

115   The Tubma basin is located within the coverage of the Sattahip radar station. The Sattahip radar, which belongs to the Department of Royal Rainmaking and Agricultural Aviation (DRRAA), is a S-band Doppler radar that transmits radiation with a frequency of 2.9 GHz and a half power beam width of 1.0°. The radar reflectivity product is provided in a Cartesian grid covering 240x240 km extent with 0.6x0.6 km spatial and 6-min temporal resolution. The Sattahip radar provides the reflectivity data derived from the 2.5-km constant altitude plan position indicator (CAPPI). The CAPPI reflectivity data originate from

120 below the climatological freezing level, so effects of measurement errors caused by bright band were considered negligible. The effects of ground clutter were removed from the reflectivity data by finding the clutter locations and discarding the radar measurements in these areas. Additionally, the noise power caused by various sources such as emission from space (cosmic noise), radiation from electrical sources near the radar antenna, and internally generated noise were eliminated by setting reflectivity values below 15 dBZ to zero (Doviak and Zrnic, 1984). While Fulton et al. (1998) suggested that measured

125 reflectivity that are greater than 53 dBZ are limited to 53 dBZ in radar rainfall estimation to mitigate false interpretation caused by hail. The hail cap can be seen as an adaptable threshold representing the maximum expected instantaneous rain rate which is unfortunately quite difficult to determine for a particular storm. Note that in tropical environments also slightly higher values have been reported as hail threshold. After data quality control, we separated the data into three datasets. The first dataset during May–October 2013 and May–September 2014 was used for the climatological Z-R calibration. The second dataset in

130 October 2014 were used for the Z-R verification, and the dataset for August–October 2019 was used in the bias correction processes.

2.2.2 The Z-R calibration and radar rainfall aggregation

The Z–R conversion error is a crucial source of error in radar rainfall estimates. The Z-R relationship was used to convert the measured reflectivity data (Z, mm^6/m^3) into rainfall rates (R, mm/h). The Z–R calibration and verification are

135 essential procedures to ascertain the parameters A and b in the relationship. Firstly, the instantaneous 6-minute radar reflectivity was converted to rainfall intensity using the climatological relationship $Z=200R^{1.6}$ proposed by Marshall and Palmer (1948). Secondly, the estimated 6-min initial instantaneous radar rainfall data were aggregated to hourly rainfall resolution using the accumulation algorithm proposed by Fabry et al. (1994). Thirdly, gauge rainfall was aggregated to hourly resolution. Finally, the optimal value of the parameter A was established by minimizing the mean absolute error (MAE) between the gauge and

140 radar rainfall estimates, while the exponent b was at 1.5 in our study. This is because radar rainfall estimates are relatively insensitive to b with typical values between 1.2 and 1.8 (Battan 1973; Ulbrich 1983). The value of 1.5 was generally suitable to represent the exponent b in the Z–R relation (Doelling et al., 1998; Steiner and Smith, 2000; Hagen and Yuter; 2003; Germann et al., 2006; Chantraket et al., 2016). The MAE is illustrated in Eq. (1).

$$MAE = \frac{1}{TN_G}\sum_{t=1}^{T}\sum_{i=1}^{N_G}|G_{i,t} - R_{i,t}| \tag{1}$$

where $G_{i,t}$ is the gauge rainfall (mm/h) at gauge i for hour t, $R_{i,t}$ is the hourly radar rainfall accumulation (mm) at the pixel corresponding to the ith rain gauge for hour t, N_G is the total number of rain gauges, and T is the total period used in the calculation. Results found that the optimal climatological Z-R relationship for the Sattahip radar used in this study is $Z=251R^{1.5}$. This relation is appropriate for both the calibration and validation datasets with the MAE of 1.36 mm and 1.47 mm, respectively.

2.3 Rain Gauge Data

2.3.1 Rainfall Data Collection

Data from the network of 297 continuous tipping-bucket gauge stations located within the Sattahip radar radius were collected (Fig. 2). These rain gauge data with a temporal resolution of 15 minutes are owned and operated by the Thai Meteorological Department (TMD). All continuous rain gauges used in this study have a resolution of 0.5 mm. The data quality screening was first carried out using double mass curves method of two adjacent rain gauges. To avoid no-rainfall events and systematical underreporting by the tipping-bucket rain gauge, hourly data above the tipping-bucket resolution of 0.5 mm were selected in the next step. Rain gauges with more than 80% of the recorded rainfall amounts below the 0.5 mm threshold at daily scale were excluded from the analysis. It turns out that many of these faulty gauges recorded zero rainfall throughout most of the study period. We found that rainfall data obtained from 134 rain gauges corresponding to the collected reflectivity datasets were used for the Z-R calibration and validation processes. For the bias adjustment computation, the selection of rain gauge networks with rainfall behavior similar to the study area is necessary. We selected 14 rain gauges of TMD in the region surrounding Tubma basin (Rayong and Chonburi provinces) based on spatial decorrelation analysis for this the process.

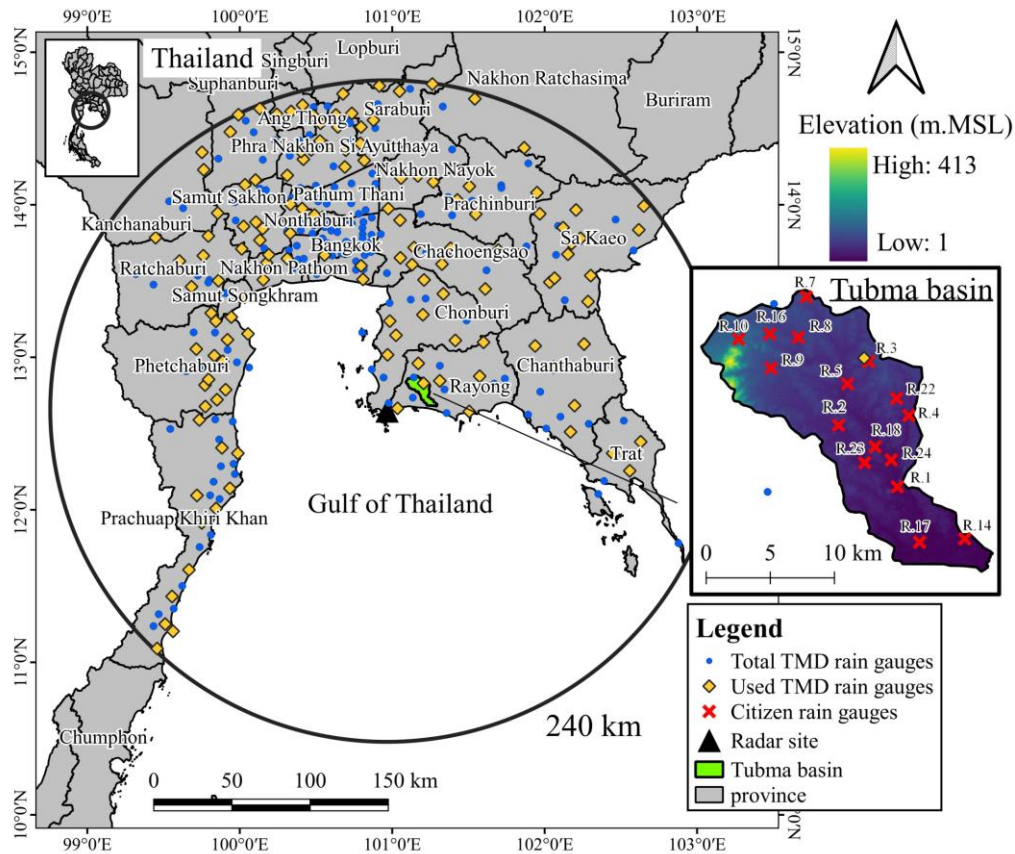



Figure 2: Location of study domain, showing Thai Meteorological Department (TMD) automatic rain gauges, citizen rain gauges, Sattahip radar, and Tubma basin.

170 2.3.2 Citizen Rain Observation

Out of the total TMD rain gauge network, only one rain gauge is located in the Tubma basin. To increase the density of the rain gauge network in the basin, low-cost citizen rain gauges were implemented in this study to capture the spatial heterogeneity of rainfall in the basin. Sixteen citizen rain gauges were installed (Fig. 2) with local residents taking daily measurements. The additional 16 citizen rain gauges with one station located at the same place as the existing TMD gauge

175 increased the density of rain gauges in the Tubma basin to 1 gauge/~12 km^2. The citizen observations were made by installing a cone-shape transparent plastic rain gauge in an open space area around a school, monastery, bridge or other building. This rain gauge is standardly used in South Africa (Fig. A1) with a diameter of 5 inches and a maximum capacity of 100 mm of rainfall. Mobile application developed by Mobile Water Management (MWM) (Mobile Water Management, 2020), the Netherlands, was used to record rainfall data for each rain gauge on a daily basis. The application has a an easily accessible

180 and user-friendly interface where participants simply fill in the observed rainfall amount, take a photo of the rain gauge and

upload this to the application. The photo and the rainfall data, together with the measuring location and time, are automatically stored in the database. Photos are used for visual validation of the recorded rainfall depth to eliminate errors.

In this study, participants were recruited amongst government officers, teachers, and local residents living close to the stations and were trained to take measurements at around 7 a.m. daily according to the TMD standards. Quality of the collected data was assured by the high photo resolution for double-checking the observations and strict requirements on measurement times to be consistent with the same standard of TMD for daily rainfall recording. Validation of the cone-shaped citizen gauges was conducted based on a citizen gauge co-located with an automatic TMD gauge located in the Tubma basin, during August – October 2019. The citizen gauge installed at the same location R.3 (Fig. A2) as a TMD gauge showed good similarity with an RMSE and Bias of 5.5 mm and 1.04, respectively.

Quality control consisted of screening all citizen rain gauge data for errors and inconsistencies using double mass curves. If citizen rain gauges reported >100 mm/day rainfall (maximum capacity of the citizen rain gauge), this data was excluded from the analysis. If days with no-rainfall data were found from all citizen rain gauges, the bias correction of that day was discarded from the dataset. By considering the data selection criteria, rainfall data recorded during August–October 2019 with rainy days, more than 80% of the whole period for the bias adjustment process was then used for further analysis.

3. Methods

The methodology for radar rainfall bias correction using tipping bucket and citizen gauges consists of the following steps. Firstly, daily citizen rain gauge data were downscaled to hourly time scale to be used as input for bias correction. The downscaling methods used in this paper are discussed in section 3.1. Next, an hourly radar bias correction model was developed combining rain gauge as well as downscaled citizen rain gauge data using a KF approach, as presented in section 3.2.

3.1 Downscaling daily to hourly rainfall

To downscale the daily citizen rain gauge data to hourly time-scale, information on the hourly storm distribution pattern is needed. Methodologies to obtain the temporal rainfall distribution patterns are outlined in Table 1.

Table 1: The four methods used in this study to downscale daily citizen rainfall amounts to hourly rainfall data.

Distribution Code	Methodologies Description	Code Description
R_P	Hourly rainfall patterns derived from radar rainfall time series of the radar pixel corresponding to citizen rain gauge location were used for downscaling.	The distribution patterns of radar rainfall at each radar pixels.

Distribution Code	Methodologies Description	Code Description
R_{MP}	Hourly radar rainfall distributions of all radar pixels corresponding to citizen rain gauge locations were averaged to represent the mean temporal distribution pattern of radar rainfall. The R_{MP} downscaling pattern was applied to all citizen rain gauges.	The mean distribution pattern of radar rainfall.
G_{MP}	Hourly gauge rainfall patterns of all 14 gauges in the region surrounding Tubma basin were averaged to construct the mean hourly distribution pattern of regional rain gauge rainfall. The G_{MP} was applied to all citizen rain gauges.	The mean distribution pattern of rain gauge rainfall.
G_{Tubma}	The hourly rainfall pattern of the single rain gauge situated in the Tubma basin was used for correction of all citizen rain gauges in the basin.	The distribution pattern of the rain gauge in the Tubma basin.

3.2 Hourly radar bias model

This section describes how radar bias is modelled (section 3.2.1), how observations are assimilated in the model to correct the bias (section 3.2.2), and how model parameters are estimated (section 3.2.3). Our approach extends a previously used Kalman
210 filter radar bias model by including two different types of rainfall observations (data from traditional and from citizen rain gauges) and by using a likelihood-based method for parameter estimation.

3.2.1 Kalman filter with two observations: model assumptions

Mean field bias (MFB) adjustment is a common technique for bias correction in radar rainfall relative to ground stations. It can be computed as the ratio of mean hourly radar rainfall estimate and rain gauge measurement (Anagnostou and Krajewski, 1999; Yoo and Yoon, 2010; Hanchoowong et al., 2013; Shi et al., 2018). However, direct application of MFB does
215 not account for uncertainty of the bias associated with each radar-gauge measurement. Alternatively, a KF has previously been used to estimate the spatially uniform MFB in real-time in several studies, including Ahnert et al. (1986), Smith and Krajewski (1991), Anagnostou et al. (1998), Seo et al. (1999), Chumchean et al. (2006), Kim and Yoo (2014), Shi et al. (2018). The KF

has the benefit of accounting for uncertainties in the observations by weighting the contribution of measurements by their respective error variances (Kalman, 1960). This is also the approach adopted here, but we extend it to include two types of observations with different error characteristics, i.e., hourly rain gauge data from TMD (y_t) and hourly downscaled citizen rain gauge data (z_t).

First, we define the logarithmic mean field radar rainfall bias β_t at hour t as (Smith and Krajewski, 1991; Anagostou et al., 1998):

$$\beta_t = log_{10}\left(\frac{\sum_{i=1}^{N_{G,t}} G_{i,t}}{\sum_{i=1}^{N_{G,t}} R_{i,t}}\right) \tag{2}$$

where $G_{i,t}$ and $R_{i,t}$ are as defined in Eq. 1. Following previous studies (e.g. Smith and Krajewski, 1991; Chumchean et al.,2006), we model the logarithmic mean field radar rainfall bias as a first-order autoregressive (AR1) process with stationary variance. Radar bias at time t is expressed in terms of the bias at previous time (β_{t-1}) and the process noise (W_t):

$$\beta_t = r_1\beta_{t-1} + W_t; \quad W_t \sim \mathcal{N}(0, \sigma_W^2) \tag{3}$$

$$\sigma_W^2 = (1 - r_1^2)\,\sigma_\beta^2 \tag{4}$$

where r_1 is lag-one correlation coefficient of the time-varying bias β_t, W_t is a zero-mean random error with variance σ_W^2, and σ_β^2 is stationary variance of the process. We consider two types of observations, y_t and z_t, of the unknown bias at time t, derived from the TMD and citizen rain gauges. Each observation is modelled as a random sample from a normal distribution conditioned on the underlying unknown bias with distinct measurement error variances $\sigma_{M_{y,t}}^2$ and $\sigma_{M_{z,t}}^2$:

$$y_t = \beta_t + M_{y,t}; \quad M_{y,t} \sim \mathcal{N}(0, \sigma_{M_{y,t}}^2) \tag{5}$$

$$z_t = \beta_t + M_{z,t}; \quad M_{z,t} \sim \mathcal{N}(0, \sigma_{M_{z,t}}^2) \tag{6}$$

where $\sigma_{M_{y,t}}^2$ and $\sigma_{M_{z,t}}^2$ are time-varying measurement error variances for the TMD and citizen rain gauges, respectively.

A factor graph representation of the radar bias and observation models is illustrated in Fig. 3 (a), with circles denoting random variables, and black squares denoting 'factors' or relations between variables in the model.

(a)   (b)

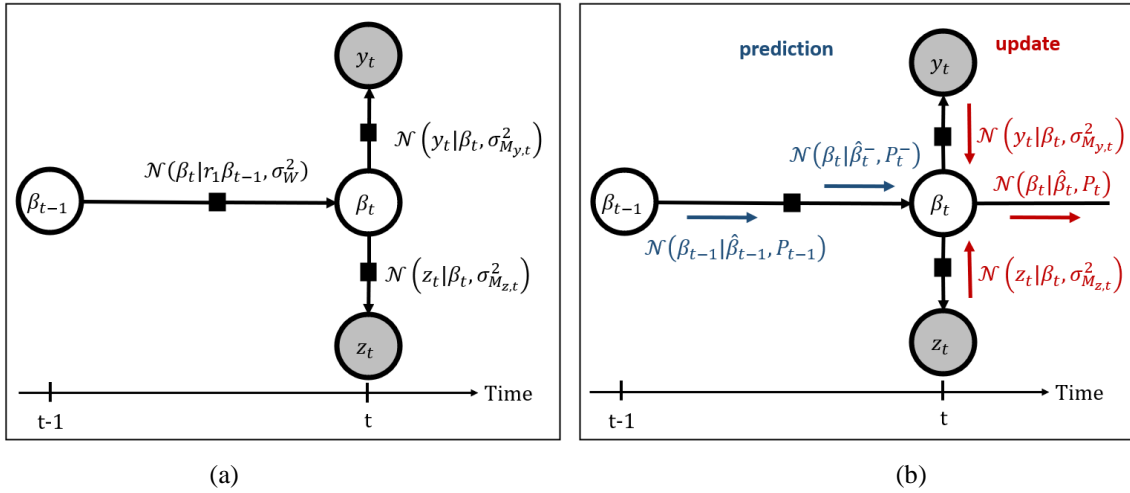



Figure 3: On the left (a), a factor graph representation of the radar bias model: white circles depict random variables (bias at each time step), grey circles are rainfall bias observations (y_t for TMD rainfall and z_t for citizen rain gauge rainfall), and black squares are relations between variables (conditional normal distributions in this case). The right figure (b) depicts uncertainty propagation along the edges of the factor graph, from previous bias to current bias (Kalman prediction step in blue) and from the observations to current bias (two Kalman updates in red).

The model contains two parameters, i.e. r_1 and σ_β^2, that need to be specified, together with values for the measurement error variances, $\sigma_{M_{y,t}}^2$ and $\sigma_{M_{z,t}}^2$. Parameter estimation will be discussed in section 3.2.3. Section 3.2.2 first describes how observations are assimilated assuming values of the parameters and the measurement error variances have been specified.

3.2.2 Kalman filter with two observations: data assimilation

Having defined the model, we describe how observations are assimilated into the model, resulting in adjustments of the estimated bias. While the regular Kalman filter has two steps (prediction and update), assimilating two observations at each time step involves three steps, i.e. a prediction followed by two updates. Figure 3 (b) shows a visual depiction of the prediction and update steps as probabilistic information propagating along the edges of the factor graph.

1) Time update step (prediction)

The first step for each hour t computes the prior mean $\hat{\beta}_t^-$ and variance P_t^- of β_t:

$$\hat{\beta}_t^- = r_1\hat{\beta}_{t-1} \tag{7}$$
$$P_t^- = r_1^2 P_{t-1} + \sigma_W^2 \tag{8}$$

where P_{t-1} is the posterior variance of β_{t-1}. For the first time step 0 ($t = 0$), we assume $\beta_0 = 0$ (climatological logarithmic mean field bias) and $P_0 = (1-r_1^2)\,\sigma_\beta^2$ (represents stationary process variance) (Smith and Krajewski, 1991; Chumchean et al., 2006).

2) The first measurement update step (1st correction)

The first update merges the bias prediction from step 1 with noisy observation y_t (with measurement error variance $\sigma_{M_{y,t}}^2$), resulting in posterior mean $\hat{\beta}_{y,t}$ and variance $P_{y,t}$ of β_t given by the following Kalman update equations:

$$K_{y,t} = P_t^- \left(P_t^- + \sigma_{M_{y,t}}^2\right)^{-1} \tag{9}$$

$$\hat{\beta}_{y,t} = \hat{\beta}_t^- + K_{y,t}(y_t - \hat{\beta}_t^-) \tag{10}$$

$$P_{y,t} = (1 - K_{y,t})P_t^- \tag{11}$$

where $K_{y,t}$ is the Kalman gain for assimilating observation y_t. If there is no observation available at time t, $K_{y,t}$ is zero (mathematically, a missing observation is equivalent to an observation with infinite variance $\sigma_{M_{y,t}}^2$ in Eq. 9), and the previous equations reduce to not performing any update, i.e. the posterior mean and variance are equal to the prior mean and variance.

3) The second measurement update step (2nd correction)

The second update is done using the posterior values from the 1st correction ($\hat{\beta}_{y,t}$ and $P_{y,t}$) as the prior values. The resulting Kalman gain and posterior mean and variance are given by:

$$K_{z,t} = P_{y,t}\left(P_{y,t} + \sigma_{M_{z,t}}^2\right)^{-1} \tag{12}$$

$$\hat{\beta}_{z,t} = \hat{\beta}_{y,t} + K_{z,t}(z_t - \hat{\beta}_{y,t}) \tag{13}$$

$$P_{z,t} = (1 - K_{z,t})P_{y,t} \tag{14}$$

If there is no observation available at time t, Kalman gain $K_{z,t}$ is zero and these equations result in no update being applied, i.e. posterior mean and variance are the same as after the first update.

Since the logarithmic mean field radar rainfall bias β_t is assumed to be normally distributed, the mean field bias (B_t) is lognormally distributed with posterior mean at time t obtained from the posterior mean and variance of β_t (Smith and Krajewski, 1991):

$$\hat{B}_t = 10^{(\hat{\beta}_{z,t}+0.5P_{z,t})} \tag{15}$$

The overall procedure of sequentially assimilating and downscaling the data is referred to as CKF (Kalman Filter combined with the citizen rain gauge data), and is summarized by the flowchart in Figure 4. Operationally it is implemented as follows:

1) Since the citizen rain gauge data were received at the last hour of day i, before receiving the downscaled hourly citizen rain gauge data of day i, the ordinary KF and observed hourly data of TMD were used to predict and correct the hourly bias adjustment factor of the day i.

2) If the citizen rain gauge data was available at the end of the day i, the citizen rain gauge data were downscaled to hourly time-scale, as explained in section 3.1.

3) The TMD and downscaled citizen rain gauge data were together used to conduct two steps of measurement update in the CKF process for all hourly time-steps of day i.

4) Bias adjustment factors were applied every hourly time step to obtain the final product of hourly radar rainfall estimation of day i. The bias factor for the last hour of the day i was used afterward as the initial value for calculating the Ordinary KF of day $i+1$.

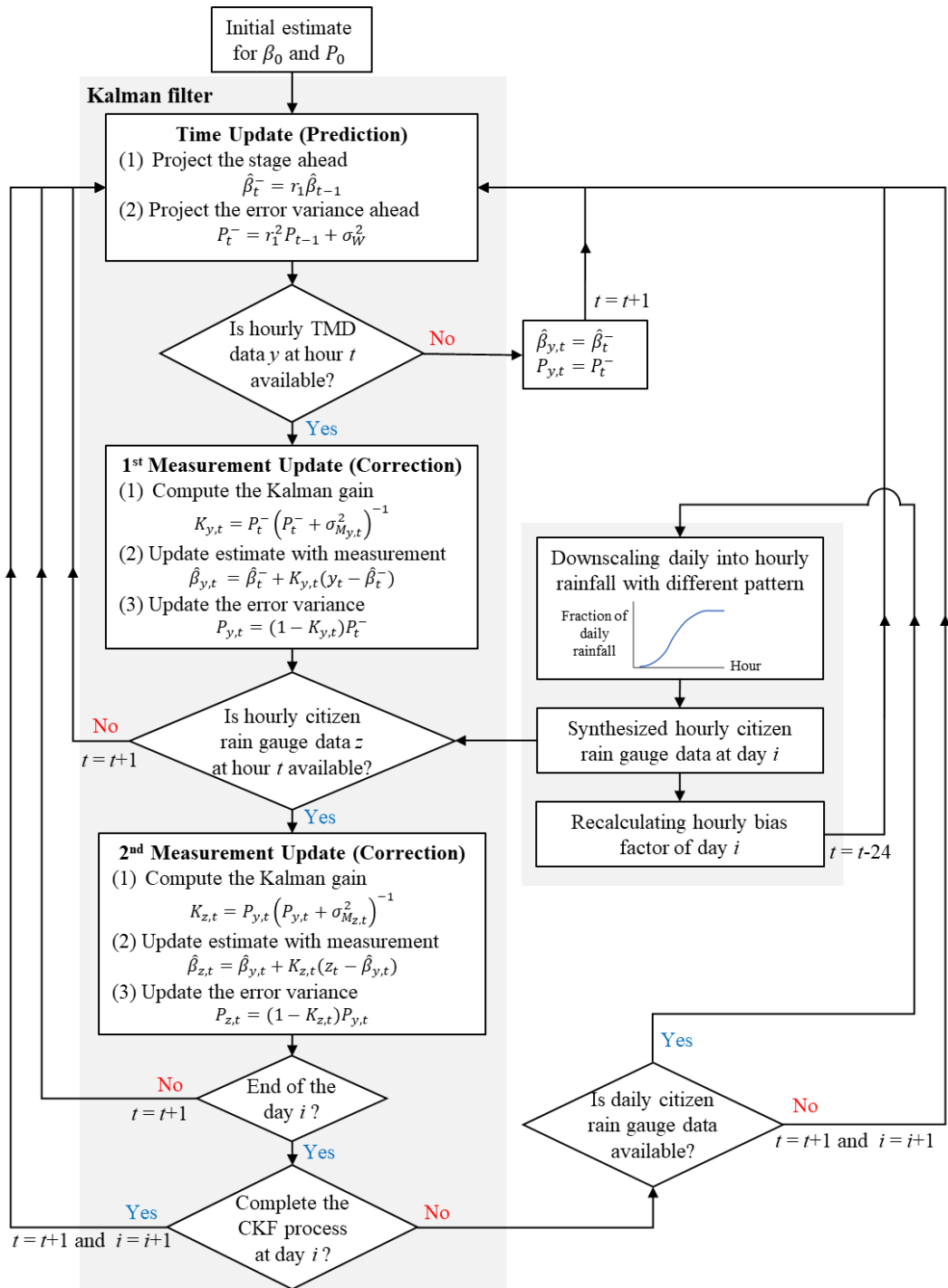



Figure 4: A diagram of the procedure of Kalman filter combined with the citizen rain gauge data (CKF)

 3.2.3 Kalman filter with two observations: parameter estimation

The model requires specification of the two parameters, r_1 and σ_β^2, and the measurement error variances, $\sigma_{M_{y,t}}^2$ and $\sigma_{M_{z,t}}^2$. We estimate the latter using formulas for the variance of the mean bias across individual TMD and citizen rain gauges:

$$\sigma_{M_{y,t}}^2 = \frac{\sigma_{\bar{y},t}^2}{n_{y,t}} \tag{16}$$

$$\sigma_{M_{z,t}}^2 = \frac{\sigma_{\bar{z}_t}^2}{n_{z,t}} \tag{17}$$

where $\sigma_{\bar{y},t}^2$ and $\sigma_{\bar{z},t}^2$ are variances that quantify spatial variability of radar bias at time t at TMD and citizen rain gauge locations, respectively, and $n_{y,t}$ and $n_{z,t}$ are the corresponding number of observations at hour t from the two networks.

Once measurement error variances are specified, values for parameters r_1 and σ_β^2 are obtained by maximizing the marginal likelihood function (Bock et al., 1981; Harvey, 1990; Proietti et al., 2013; Pulido et al., 2018). As mentioned earlier, we have two sources of observed log mean field bias at hour t, from TMD (y_t) and citizen rain gauge (z_t). When only TMD data are available, the marginal likelihood is computed as:

$$p(D) = \prod_{t=0}^{T} \int p(y_t|\beta_t)p(\beta_t|y_{t-1}, \dots, y_0)d\beta_t = \prod_{t=0}^{T} \mathcal{N}(y_t; \hat{\beta}_t^-, \sigma_{y_t}^2 + P_t^-) \tag{18}$$

where D is the data vector that contains all observed values, β_t is the true hidden state of log mean field bias at hour t, and T is the number of hourly time steps. When both TMD and citizen rain gauge data are available, the marginal likelihood is obtained as:

$$p(D) = \prod_{t=0}^{T} \int p(y_t|\beta_t)p(z_t|\beta_t)p(\beta_t|y_{t-1}, z_{t-1}, \dots, y_0, z_0)d\beta_t$$

$$= \prod_{t=0}^{T} \mathcal{N}(y_t; z_t, \sigma_{M_{y,t}}^2 + \sigma_{M_{z,t}}^2)\mathcal{N}(\tau^{-1}(\frac{y_t}{\sigma_{M_{y,t}}^2} + \frac{z_t}{\sigma_{M_{z,t}}^2}); \hat{\beta}_t^-, \tau^{-1} + P_t^-) \tag{19}$$

where $\tau = \frac{1}{\sigma_{M_{y,t}}^2} + \frac{1}{\sigma_{M_{z,t}}^2}$ and prior mean $\hat{\beta}_t^-$ and variance P_t^- are from the prediction step in section 3.2.2.

To obtain optimal values of the two parameters r_1 and σ_β^2 that maximize the (logarithm of the) marginal likelihood, the Nelder-Mead Simplex was used, which is an algorithm for searching a local optimum of a function (Lagarias et al., 1998; Luersen et al., 2004; Gao et al., 2012).

3.3 Verification of the proposed bias correction approaches

To investigate which bias adjustment technique (MFB, KF, and CKF) gives the most suitable radar rainfall estimates for the Tubma basin, the adjusted radar rainfall estimates were validated against measured rainfall data. There was only one automatic TMD rain gauge available in the basin, which was insufficient for validation purposes. Consequently, for testing the performance of hourly rainfall bias correction, data from 13 TMD stations located within a 90 km radius from the center of the Tubma basin were used, together with one TMD station in the basin. Furthermore, daily time-scale validation was conducted, using the daily rainfall data from 16 citizen rain gauges located in the Tubma basin. Leave-one-out cross-validation (LOOCV) algorithm was implemented to avoid bias in selecting the validation rain gauges. For each round of cross-validation, one rain gauge was left out for validation and the remaining rain gauges were used as the calibration rain gauges to calculate the bias adjustment factor using the three different techniques. This was repeated for all combinations and then the error of radar rainfall estimates after correcting with the estimated bias factor at each radar pixel corresponding to the held-out gauge was computed for all trials. In this study, Root Mean Square Error (RMSE) and Mean Bias Error (MBE) were applied as statistical measures to evaluate the effectiveness of the different bias correction methods at each validation rain gauge. The number of possible combinations is equal to the total number of validated gauges (N_G). Data for the period August-October 2019 were used in the evaluation. Four scenarios combining the three bias adjustment techniques were evaluated, summarized in Table 2.

Table 2: Simulation cases for evaluating the effectiveness of bias correction techniques.

Evaluation Case	Tested approaches	Number of rain gauges used for different purposes				Temporal and spatial scale of rainfall for validation
		rain gauges datasets	gauges for calculating MFB and KF	gauges for combining with KF	validation gauges (N_G)	
KF-TMD-H	MFB and KF	14 TMD	13 TMD	-	14 TMD	Hourly, Tubma plus 90 km radius
KF-TMD-D	MFB and KF	14 TMD and 16 citizen rain gauges	14 TMD	-	16 citizen rain gauges	Daily, Tubma basin
CKF-D	MFB, KF and CKF	14 TMD and 16 citizen rain gauges	14 TMD	15 citizen rain gauges	16 citizen rain gauges	Daily, Tubma basin
CKF-H*	MFB, KF and CKF	14 TMD and 16 citizen rain gauges	13 TMD	16 citizen rain gauges	14 TMD	Hourly, Tubma plus 90 km radius

*CKF-H includes four scenarios for four different hourly downscaling patterns for the citizen rain gauges, according to Table 1.

KF-TMD-H: One TMD rain gauge from the total of 14 gauges was left out for validation and the remaining 13 gauges were separated for calculating the bias adjustment factors using MFB and KF. This process was iterated 10 times until all 14 TMD rain gauges were left out for cross-validation. Aggregated hourly rainfall between the adjusted radar and gauge rainfall data were compared to obtain the RMSE and MBE.

KF-TMD-D: To identify whether MFB or KF is more accurate for daily rainfall simulation in the Tubma basin if there are only 14 TMD rain gauges available, 14 TMD and 16 citizen rain gauges were used for the analysis. All TMD gauges were used for assessing MFB and KF, and estimated bias factors were applied at the daily time-scale. Assessment of RMSE and MBE of daily rainfall at all 16 citizen rain stations was used for validation.

CKF-D: To evaluate the added value of using citizen rain gauges in the basin for bias correction, 15 citizen rain gauges (leave one citizen rain gauge out for validation) were used in addition to the TMD gauges following the CKF procedure explained in 3.2.2. Estimation of daily RMSE and MBE was carried out at the held-out citizen rain gauge.

CKF-H*: To test whether the CKF with the most suitable storm pattern could benefit radar rainfall estimates in the area further away from the Tubma basin, 14 TMD gauges were used to generate four cases of hourly rainfall distribution patterns as described in Table 1 for downscaling the selected 16 daily citizen rain gauge data to hourly time scale. The synthesized hourly citizen rain gauge data were later used as a second observation for the 2nd correction of the CKF. Thirteen TMD gauges (leave one TMD out) were used to produce estimates with MFB and KF, and all 16 citizen rain gauges were merged for CKF computation.

All bias adjustment techniques evaluated the effectiveness at the held-out gauge for all possible combinations of the LOOCV procedure.

4. Results and discussion

4.1 Simulation of bias adjustment factor

4.1.1 Parameter estimation for the KF and CKF

Five scenarios were investigated for radar bias correction using the KF, based on TMD and citizen rain gauge observations, including four scenarios comparing different hourly downscaling approaches for the citizen rain gauge data (Table 1). Parameter estimates of the KF are shown in Table 3. These results indicate that the parameter r_1, the lag-one correlation coefficient of the logarithmic mean field bias, ranges from 0.15 to 0.53, depending on the hourly downscaling approach. While σ_β^2 representing the stationary variance of the logarithmic mean field bias remains relatively invariant (ranging from 0.24-0.28) over the same period of simulation.

Table 3: The parameters of the KF estimated from different datasets of observation gauge rainfall. KF-TMD is using only TMD hourly rain gauge observations, CKF is using TMD and citizen rain gauge observations, where R_P, R_{MP}, G_{MP} and G_{Tubma} represent different strategies for hourly downscaling of the citizen rain gauge observations (Table 1)

Type of observation gauge rainfall	The KF's parameters	
	r_1	σ_β^2
KF-TMD	0.29	0.24
CKF-R_P	0.53	0.28
CKF-R_{MP}	0.33	0.24
CKF-G_{MP}	0.15	0.24
CKF-G_{Tubma}	0.38	0.25

400

4.1.2 Bias adjustment factor comparison

To test the performance of the bias adjustment techniques among KF-TMD, CKF-R_P, CKF-R_{MP}, CKF-G_{MP}, and CKF-G_{Tubma}, all approaches were used to assess the mean field bias for each hour using the data period August – October 2019. The results were compared to the MFB calculated using the 14 TMD rain gauges (MFB-TMD) in the Tubma basin and 90 km

405 radius surroundings. Results summarized in Fig. 5 show that:

- The daily observed bias is somewhat higher and shows larger variability for the citizen gauges compared to TMD gauges. The hourly observed bias based on downscaled citizen gauge data are in the same range as hourly bias based on TMD gauges, with somewhat higher median values and spread (25-75 percentile range) for the R_P and G_{Tubma} downscaling scenarios.

- Hourly observation error variance is smallest for the CKF-R_P downscaling approach and somewhat larger for the

410 other CKF approaches compared to observation error variance for the TMD gauges.

- Estimated hourly bias values based on KF-TMD show a slightly higher mean and smaller variability range compared to observations. The bias produced by the KF-TMD is close to the MFB-TMD if the observation error variance is small. In case that no measured data is available for the bias update, the computed bias factor (B_t) progressively converged to 1.3, to meet the climatological logarithmic mean field bias.

415 - Estimated bias values based on the CKF approaches are able to reproduce bias variability as observed by TMD gauges, with median values deviating by 0.2 to 0.4 and value range slightly larger for CKF-R_P and smaller for CKF-R_{MP}.

- CKF gives different bias values according to the storm distribution pattern (see Appendix B) and the availability of the daily citizen rain gauge data used in combination with the KF. In case that no citizen rain gauge data is available for updating, the bias generated by the CKF for every combination is close to the ordinary KF with small differences depending on their respective r_1 and σ_β^2 parameters.

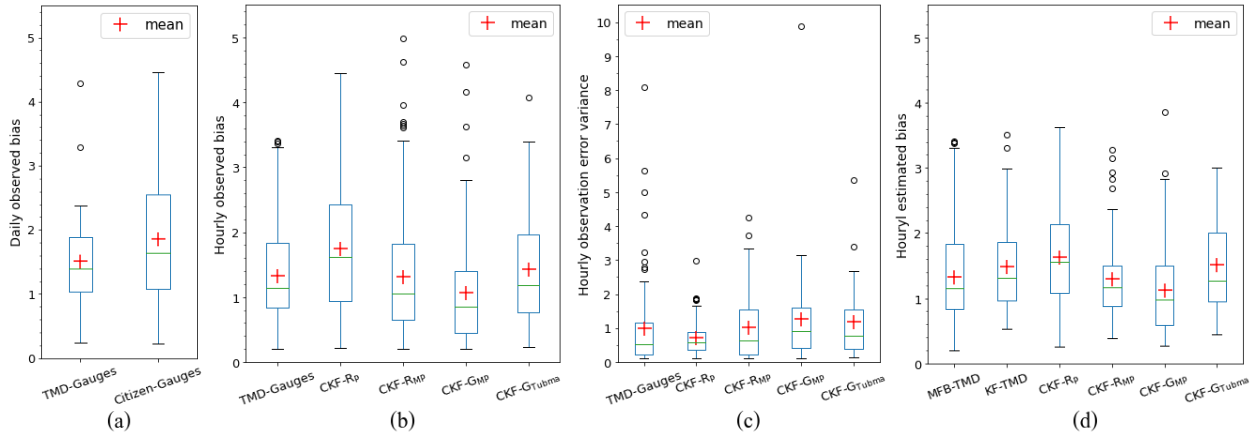



Figure 5: Comparison of (a) daily observed mean field bias based on TMD rain gauges in the region and the citizen rain gauges in the Tubma basin, (b) hourly observed mean field bias based on TMD rain gauge observations and downscaled citizen rain gauge observations, (c) hourly observation error variances and (d) hourly estimated mean field bias obtained based on MFB and the five different KF approaches. Bias calculations cover 16 citizen gauges in the Tubma basin and 14 TMD gauges within 90 km radius from the Tubma basin. Hourly scale calculations for the citizen gauges (CKF) are based on four different sub-daily interpolation scenarios (R_P, R_{MP}, G_{MP} and G_{Tubma}, Table 3).

4.2 Effectiveness evaluation of bias correction approaches

4.2.1 Hourly rainfall validation for the larger region (90 km radius) surrounding Tubma basin using MFB and KF approaches (Case 1)

Figure 6 (a) shows cross-validation results based on RMSE and MBE between TMD rain gauges and adjusted radar rainfall using MFB and KF for hourly bias adjustment. Bias adjustment reduces RMSE and especially MBE, with KF-TMD performing somewhat better than MFB-TMD especially in terms of RMSE. This confirms the ability of the KF approach that considers the error variance of observed hourly data as the weight for correcting the predicted mean bias instead of using only the calculated mean field bias (Smith and Krajewski, 1991; Chumchean et al., 2006).

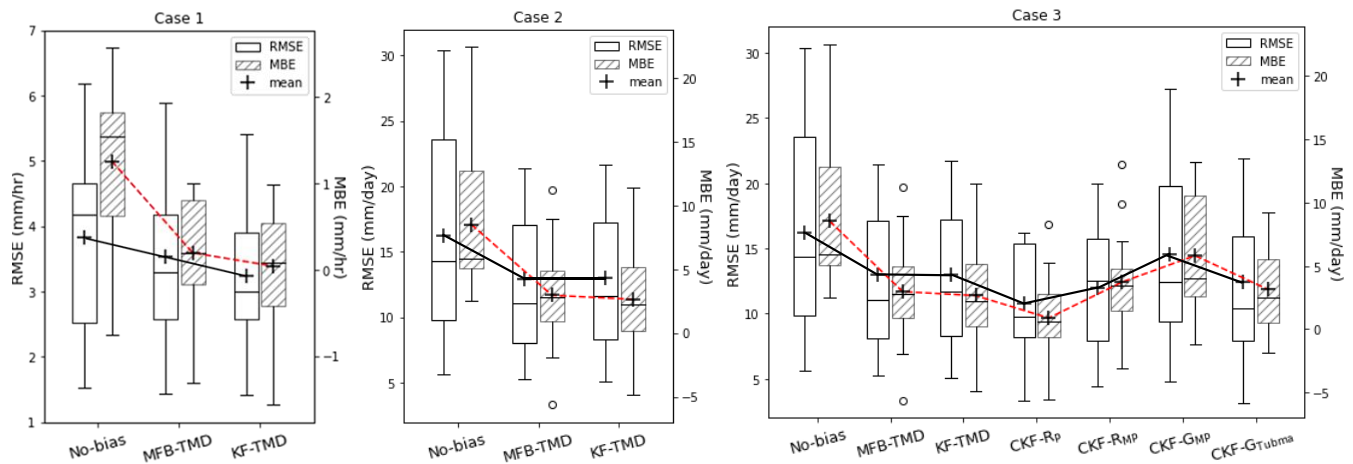



Figure 6: Variation in RMSE and MBE across the cross-validation scenarios for the various evaluation cases: (a) case 1, hourly bias updating based on MFB and KF using TMD gauges (b) case 2, daily bias updating based on MFB and KF using only TMD gauges and (c) case 3, daily bias updating using MFB, KF (TMD gauges) and CKF (TMD and citizen gauges). Validation covers 16 gauges in the Tubma basin for daily scale and 14 gauges within 90 km radius from the Tubma basin for hourly scale.

4.2.2 Daily rainfall validation in the Tubma using MFB and KF approaches, citizen gauges for validation (Case 2)

Results associated with validating the bias correction performance within the Tubma basin are presented in Fig. 6 (b). This shows bias correction performance within the Tubma basin, for MFB and KF-based daily bias adjustment. The two approaches show similar performance at the daily scale and improve RMSE by 20-30% and MBE by 50-60% (for median and upper 75percentile, respectively). The added value of a KF-based approach is limited for this case, since 14 TMD rain gauges in the region were used to compute observation variance which cannot represent the mean field bias behaviour in the Tubma basin.

4.2.3 Daily rainfall validation in the Tubma using CKF approaches (Case 3)

Figure 6 (c) shows cross-validation results at daily scale for the Tubma basin, comparing bias correction approaches using TMD only and TMD combined with citizen gauges. Following the CKF steps, citizen rain gauge data are downscaled to hourly time scale using four different approaches, resulting in variation in hourly observed bias and error variances as shown in Fig. 5 (b) and (c), respectively. Cross-validation results after accumulation to daily scale show that CKF-R_P outperforms the other approaches (CKF-R_{MP}, CKF-G_{Tubma}, MFB-TMD, KF-TMD, and CKF-G_{MP}) in terms of both RMSE and MBE. The performance of CKF techniques for radar rainfall simulation in the Tubma basin relates to the reliability of the downscaled hourly observations. This is reflected in the variation of the estimated observation error variances for CKF-R_P as shown in Fig. 5 (b) and (c). The better performance of CKF-R_P is explained by the smallest range in observation error variance, indicative

of better consistency observation bias. Comparison with No-bias, CKF-R_P can improve RMSE by 32-25 % and MBE by 90-80 % for median and upper 75 percentile, respectively. While CKF-G_{MP} exhibits the worst performance compared with the other CKF approaches with the improvement of RMSE by 13-16 % and MBE by 57-56 %, respectively. This apparently decrease in efficiency of the CKF can confirm by the highest median value of the estimated observation error variances of CKF-G_{MP} (see Fig. 5 (c)) with 33% higher than that of CKF-R_P.

4.2.4 Hourly rainfall validation using MFB, KF, and CKF approaches (Case 4)

Results for this section are presented in Figs. 7 and 8. The performance for radar rainfall estimates using CKF-RP with different distances from the Tubma basin is reported in Appendix C. Cross-validation results at hourly time-scale show a strong improvement achieved by bias adjustment using citizen gauges, in particular close to the Tubma basin where the citizen gauges are located. Figure 7 (b) and (c) show validation results based on TMD gauges for gauges close to (0-40 km radius) and further away (40-90 km radius) from the center of Tubma basin (see Fig. 7 (a)), both ranges cover the similar number of TMB gauges). Figure 7 (b) and (c) show that CKF-R_P bias adjustment significantly improves radar rainfall estimates at hourly time scale, compared to bias adjustment approaches based on TMD gauges only in the 0-40 km range closest to Tubma basin. While there is a modest improvement in mean RMSE (see the black line connecting the mean values of the box plots from MFB-TMD to CKF-R_P), the upper 75 percentile RMSE is reduced from about 6 mm/h to 3.5 mm/h. Mean MBE is changed from 0.1 to -0.15 mm/h (see the red-dotted line connecting the mean values from MFB-TMD to CKF-R_P). For the 40-90 km range, CKF-R_P performs similarly to MFB-TMD and KF-TMD.

Figure 8 illustrates that hourly rainfall distribution patterns of TMD rain gauges in the 0-40 km range, influenced mainly by the southwest monsoon, appear to be more similar to the mean citizen rain gauge data than the range beyond 40 km. Consequently, the application of CKF-R_P based on combining citizen rain gauge network to TMD rain gauge network with similarity of rainfall characteristic is a key for improving radar rainfall estimates.

Results in Fig. 7 also show that the MBE values in the 0-40 km range are explicitly lower than that in the 40-90 km range. Apparently, at shorter range, positive and negative errors represented in MBE cancel out more frequently than they do for the gauges at larger distance. In other words, gauges more or less randomly over or underestimate rainfall values as we can see similar rainfall distribution patterns among all gauges with high variation of rainfall amount during the storm period in Figure 8 (b). Conversely, in the 40-90 km range, bias correction at gauge locations consistently leads to over- or underestimation of rainfall. This can be explained by gauge at larger distance being affected by different rainfall generation patterns, associated with their location closer to the coast or mountains (see Fig. 7 (a)). The influence of the southwest monsoon strongly affects all gauges located in the coastal region on the windward side of a mountain, while rain gauge locations on the leeward side have less rainfall amount. Figure 8 (c) shows that TMD gauges located on the leeward side (e.g., 4590009 and 4590011) obviously appear a steady gradient of the mass curve reflecting light rainfall accumulation, whereas the gauges on the windward side (e.g., 4590002 and 4590003) show the mass curves with a sharper gradient. Further details on the sensitivity bias adjustment techniques accuracy to rainfall characteristics are provided in Appendix D.

(a)   (b)   (c)

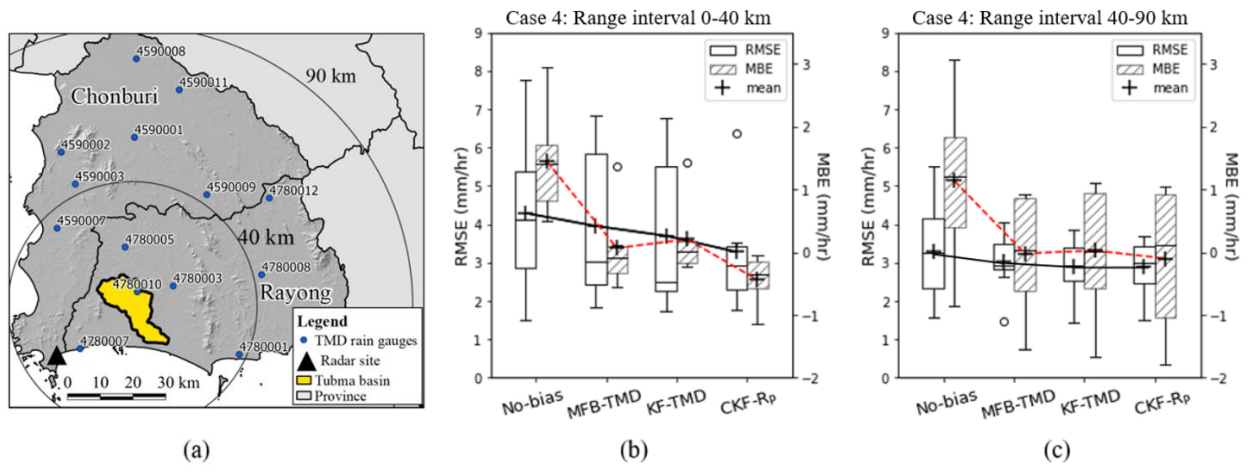



Figure 7: Comparison of the RMSE and MBE for different range interval from the centroid of the Tubma (a) Rain gauge locations at each range interval (b) the comparisons for the range 0-40 km (c) the comparisons for the range 40-90 km. For CKF, only results for the CKF-R_P approach are shown, based on its better performance at daily time-scale (shown in Figure 6c).

(a)   (b)   (c)

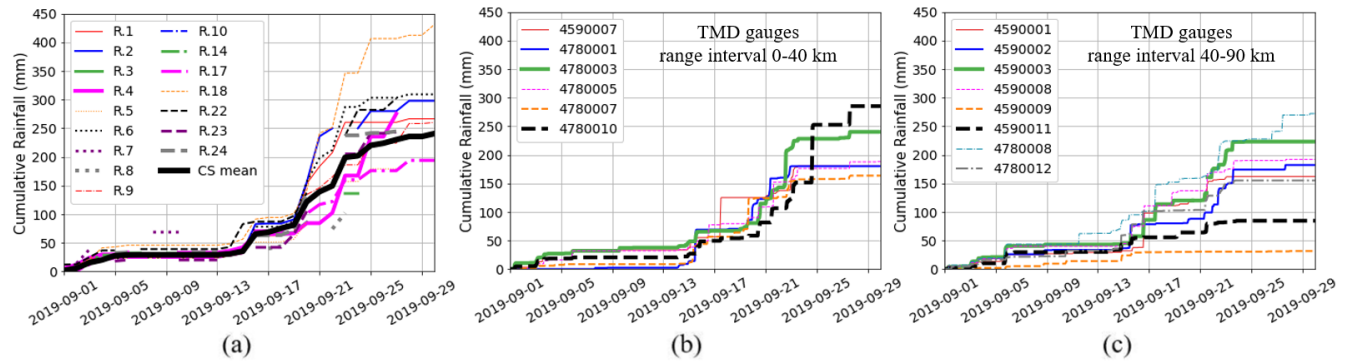



Figure 8: Comparison of mass curves of hourly rainfall among various rain-gauge locations (a) the citizen rain gauges located in the Tubma (b) TMD rain gauges within 0-40 km radius from the Tubma basin (c) TMD rain gauges within 40-90 km radius from the Tubma basin.

5. Conclusion

In this study we introduced a modified KF approach in radar bias correction in the Tubma basin, eastern Thailand. The two-step KF integrates daily data from a dense citizen rain gauge network with hourly data from a much sparser network of conventional, yet higher quality tipping bucket rain gauges. Daily citizen rain gauge observations were downscaled to hourly time scale using four different approaches. The question we aimed to answer is to what extent the downscaled citizen rainfall

observations improve the accuracy of hourly radar rainfall estimates. Results showed that citizen rain gauges significantly improve the performance of radar rainfall bias adjustment, up to a range of about 40 km from the centre of the Tubma basin (197 km^2) where the citizen rain gauge network is located. While a modest improvement in mean RMSE was obtained, the

510   upper 75 percentile RMSE was reduced from 6 mm/h to 3.5 mm/h. The mean bias error was changed from 0.1 to -0.15 mm/h across the validation period (August–October 2019). In the Tubma basin, beyond the 40 km range, no significant improvement by inclusion of the citizen gauges was found. The rainfall distribution pattern is key for downscaling the daily measured citizen rain gauge observations into hourly temporal resolution. We found that in the Tubma basin downscaling based on the rainfall patterns derived from hourly radar rainfall at overlying radar pixels corresponding to the citizen gauge location was the most

515   suitable technique, resulting in the smallest variation of observation error variances of the mean field bias. In the case of a sparse rain gauge network, the mean field bias and the Kalman filter approach both show improvement, and the degree of improvement was similar between the two approaches. In other words, in a sparse gauge network, the added value of error information represented in the Kalman filter is limited. Note that citizen rain gauge data are available only at the end of the day and consequently, the modified two-step Kalman filter as used in this study has restrictions in real-time applications.

520   However, the here proposed method has clear potential when creating high quality historical radar-rainfall time series for climatological studies and in post-event analysis. Moreover, near real-time assessment could be achievable if the citizen rain gauge data are collected at sub-daily time scale.

Appendix

Appendix A: Citizen rain observation and validation

525

An example of installing a cone-shape transparent plastic citizen rain gauge which is standardly used in South Africa at location R.22, Map Tong school, Rayong province is illustrated in Fig. A1. While validation of the cone-shaped citizen gauges was done at location R.3 by comparing the measured daily citizen gauge data with an automatic TMD gauge at same location as presented the result in Fig. A2. Root Mean Square Error (RMSE) and Bias were applied as statistical measures for

530 the validation as shown in Eq. A1 and Eq. A2, respectively.

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(G_{TMD,t} - G_{CR,t})^2} \tag{A1}$$

$$Bias = \frac{\sum_{t=1}^{T} G_{TMD,t}}{\sum_{t=1}^{T} G_{CR,t}} \tag{A2}$$

535

where $G_{CR,t}$ is daily citizen gauge rainfall at location R.3 for day t, $G_{TMD,t}$ is daily TMD gauge rainfall at the same location of R.3 for day t, *and T* is the total period used in the calculation.

540

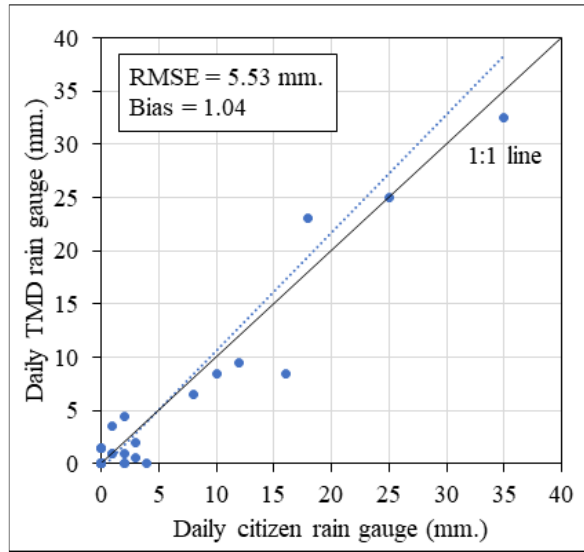
Figure A1: An example of installing a cone-shape transparent plastic citizen rain gauge at location R.22, Map Tong school.

545

Figure A2: Daily rainfall depth comparison between the TMD and citizen rain gauge at location R.3 during August – October 2019.

550 **Appendix B: Hourly rainfall distribution patterns**

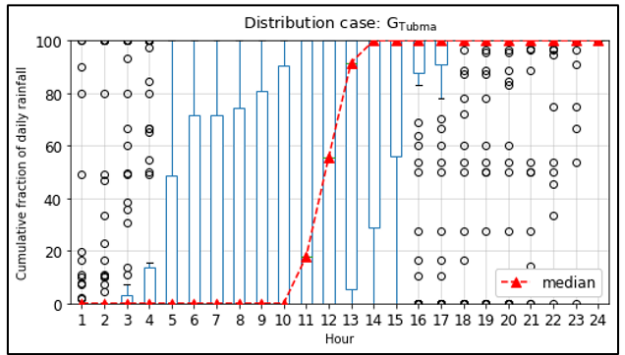
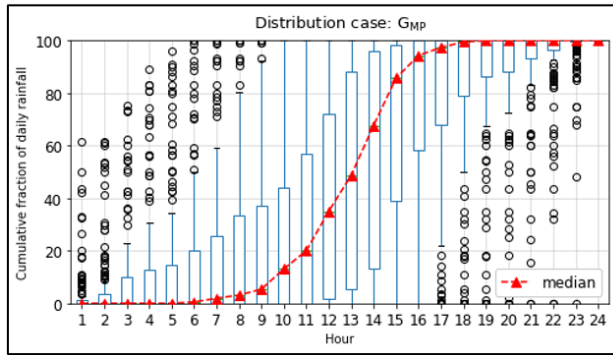
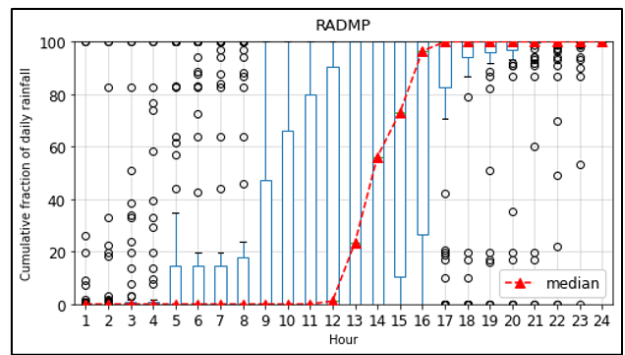
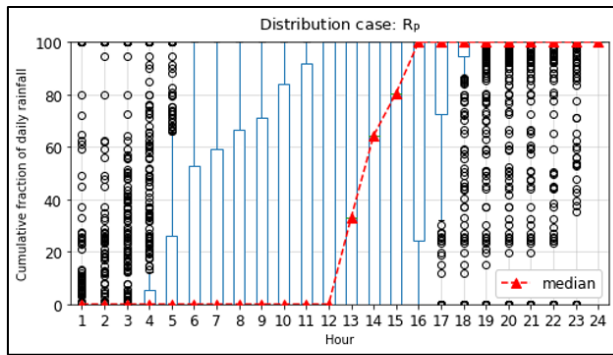
Four hourly rainfall distribution patterns were obtained as outlined in Table 1. Figure B1 illustrates the cumulative fraction of daily rainfall at hourly scale during the simulation period August-October 2019. It can be seen that most rainfall was concentrated in the afternoon hours, with very little rainfall falling at night. R_P and R_{MP} showed relatively more rainfall concentrated afternoon rainfall, while R_P and G_{MP} showed larger variability in the downscaled hourly data with substantial
555 outliers in the box plots, associated with variability in the locations underlying the rainfall distributions (multiple radar pixels within the Tubma basin for R_P versus multi TMD gauges surrounding the Tubma basin for G_{MP}). G_{MP} showed the flattest distribution with the longest rainy period of around 11 hours compared to the others having a period of heavy rainfall around 4-5 hours a day. This is explained by the larger spatial variability in the gauges covered by G_{MP}.

560

Figure B1: Variation of fraction of 24-hour rainfall for each rainfall distribution scenario.

Appendix C: The performance in radar rainfall estimates of using CKF-RP with different distances from the Tubma basin.

We chose the 40 km separation boundary to achieve an equal number of gauges in the "near" and "far" groups. Following the referee's comment, we investigated RMSE between gauge rainfall and radar rainfall without bias adjustment ($RMSE_{No-Bias}$) and with the CKF-R_P ($RMSE_{CKF-RP}$) for individual stations (located at distances 5 – 80 km from the Tubma basin). We computed the percentage improvement in radar rainfall estimates using CKF-R_P compared to No-Bias at each rain gauge, indicating the relative errors changing with distance from the Tubma basin (Fig. C1). Figure C1 shows that the improvement percentage of using CKF-R_P tends to reduce with increasing distance from the Tubma basin, where the citizen gauges are located. The percentage reduction gradually decreases beyond a distance of about 40 km.

Note that the gauge 4780010 situated at the nearest range is expected to provide the best improvement, however, the Sattahip radar temporarily stopped measuring for 3 hours (during 16:00 h - 19:00 h of 24 September 2019) which is associated with a heavy storm's center was only at the gauge 4780010. This leads to significant degradation of the radar rainfall performance. Furthermore, the lower percentage improved for station 4780005 is associated with a localized heavy rainfall event that was recorded only at this location, negatively affecting its performance as a representative station for bias correction.

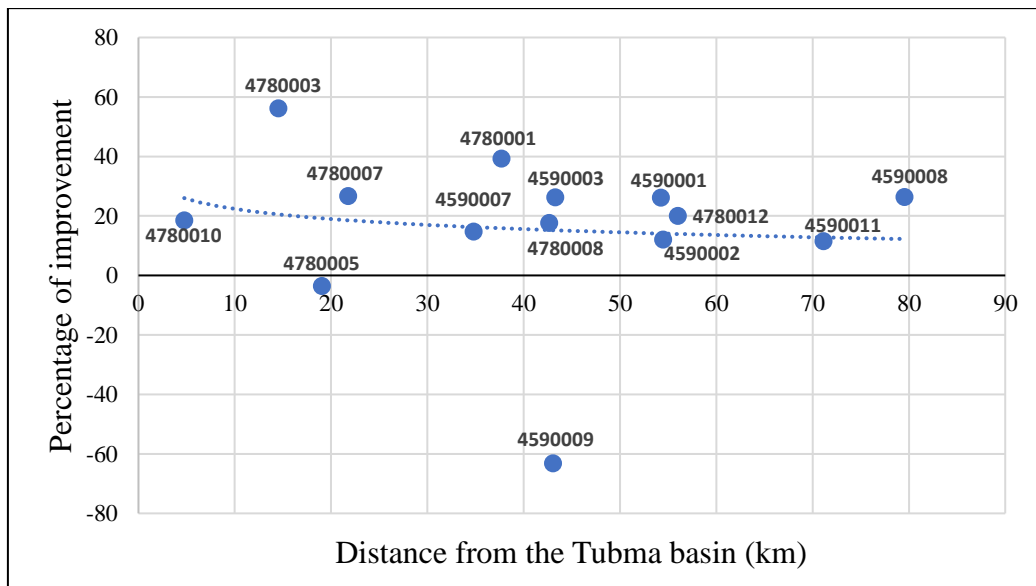



Figure C1: Alteration of the percentage improvement in radar rainfall estimates of using CKF-R_P compared to No-Bias at each rain gauge with different distance from the Tubma basin.

Appendix D: Investigation of rainfall characteristics affecting the accuracy of bias adjustment techniques

585   Figure 7 (b) in the manuscript shows that the upper 75 percentile RMSE of the shorter range is remarkably high while using only TMD gauges for the bias adjustment. These errors occurred in 3 hours at different 3-gauge locations when heavy rainfall data were only measured at the validated gauge location while there was relatively uniform light rainfall at all available surrounding TMD gauges used for the bias adjustment calculation. Consequently, the calculated bias factors from the available gauges cannot represent the heavy rainfall at the tested location leading to the significant RMSE. Analysis of hourly rainfall

590 hyetographs obtained from TMD rain gauge network compared with the validated rain gauge occurring in three different days are illustrated in Fig. D1. It shows considerable RMSE from three hours for three days comprising 15 September 2019, 12:00; 21 September 2019, 15:00; and 22 September 2019, 14:00 associated with the validated gauge 4780001, 4780005, and 4780003, respectively. However, these RMSE values decrease considerably if the CKF-R$_P$ was implemented only in the shorter range.

595

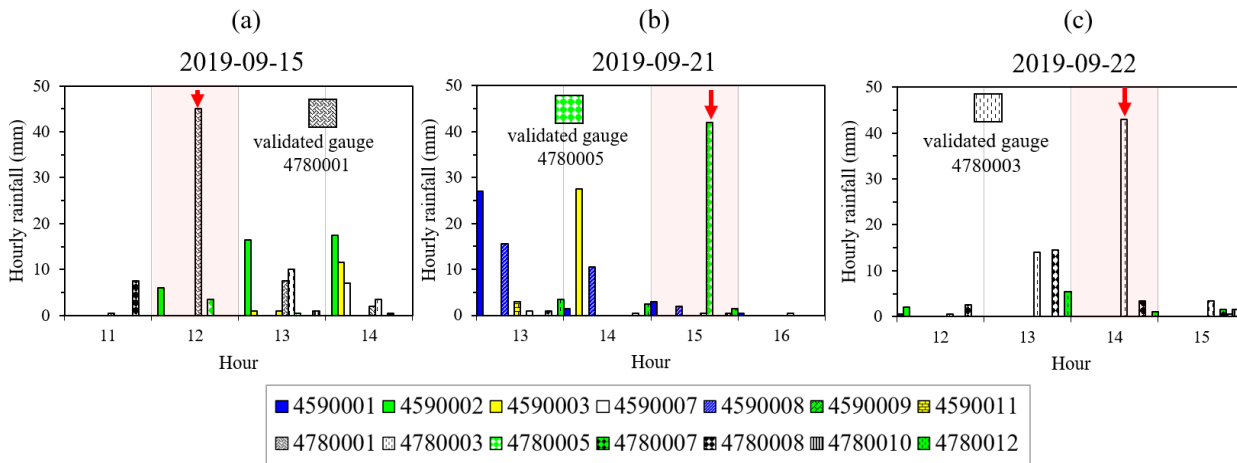



Figure D1: Hourly rainfall hyetographs obtained from TMD rain gauge network available for each hour compared with the validated rain gauge occurring in 3 different days (a) storm event during 15 September 2019 based on using 4780001 as the validated gauge, (b) storm event during 21 September 2019 based on using 4780005 as the validated gauge, and (c) storm

600 event during 22 September 2019 based on using 4780003 as the validated gauge.

        Indeed, Fig. 8 (a) and (b) show that one of the citizen gauges (R.18) collected cumulative monthly rainfall higher than the range of the other citizen and TMD rain gauges. This is associated with a storm event that occurred in September 2019, with the storm centre over R.18, while the surroundings received appreciably less rainfall. Figure D2 shows the radar

605 reflectivity field at 13.00 h on 22 September 2019, during the peak of the storm, confirming the heavy rainfall affecting gauge location R.18. This shows the citizen gauge network is able to capture local storm features thanks to the high density of the

network. The multiple reporting gaps visible in Fig. 8 (a) are caused by time errors in the observations submitted by local residents were removed from the analysis (as explained in section 2.3.2).

610

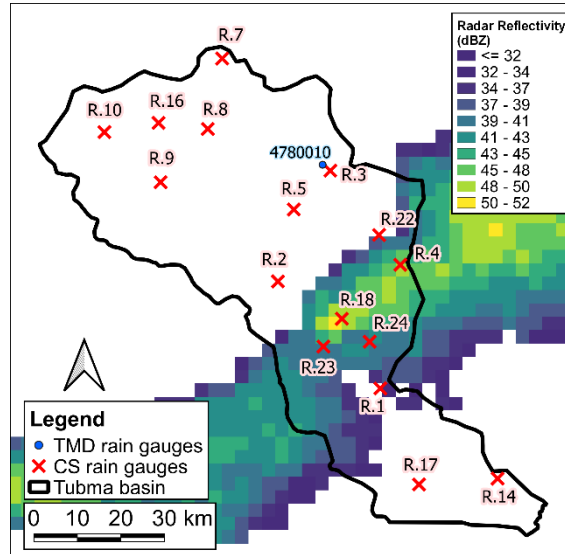



Figure D2: Overlapping between citizen rain gauges network and spatial radar reflectivity data on 22 September 2019 at 13.00 hour.

615

Acknowledgements. This research is due to the research exchange between Kasetsart University (KU), Thailand and Delft University of Technology (TU Delft), the Netherlands. The first, second and third authors gratefully acknowledge the Faculty of Engineering, Kasetsart University for financially supporting this research, 3-weeks guest visiting in TU Delft of the first author, 3-months guest visiting in TU Delft of the second author, and 2-months guest visiting in KU of the third author. The
620 authors wish to thank the Department of Royal Rainmaking and Agricultural Aviation (DRRAA) and Thai Meteorological Department (TMD) for providing radar reflectivity and rain gauge data used in this study. We also appreciate Mobile Water Management (MWM), the Netherlands for providing mobile application for measuring daily citizen rain gauge rainfall in the study area.

References

625 Amitai, E.: Systematic Variation of Observed Radar Reflectivity–Rainfall Rate Relations in the Tropics, Journal of Applied Meteorology, 39, 2198-2208, 10.1175/1520-0450(2001)040<2198:SVOORR>2.0.CO;2, 2000.

Anagnostou, E. N., Krajewski, W. F., Seo, D.-J., and Johnson, E. R.: Mean-Field Rainfall Bias Studies for WSR-88D, 3, 149-159, doi:10.1061/(ASCE)1084-0699(1998)3:3(149), 1998.

630

Anagnostou, E. N., and Krajewski, W. F.: Real-Time Radar Rainfall Estimation. Part I: Algorithm Formulation, Journal of Atmospheric and Oceanic Technology, 16, 189-197, 10.1175/1520-0426(1999)016<0189:RTRREP>2.0.CO;2, 1999.

Anhert, P., Krajewski, W. F., and Johnson, E. R.: Kalman Filter estimation of radar-rainfall mean field bias, 23rd Radar

635   Meteorology Conf. Amer. Meteor. Soc, 33-37, 1986.

Austin, P. M.: Relation between Measured Radar Reflectivity and Surface Rainfall, Monthly Weather Review, 115, 1053-1070, 10.1175/1520-0493(1987)115<1053:RBMRRA>2.0.CO;2, 1987.

640   Barton, Y., Sideris, I. V., Germann, U., and Martius, O.: A method for real-time temporal disaggregation of blended radar–rain gauge precipitation fields, 27, e1843, https://doi.org/10.1002/met.1843, 2020.

Battan, L. J.: Radar observation of the atmosphere, Rev. ed. ed., University of Chicago press, Chicago, 1973.

645   Bedient, P. B., Holder, A., Beneavides, J. A., and Vieux, B.: Radar-based flood warning system applied to tropical storm Allison, Journal of Hydrologic Engineering, 8, 308-318, 10.1061/(ASCE)10841-0699(2003)8:6(308), 2003.

Berne, A. and Krajewski, W. F.: Radar for hydrology: Unfulfilled promise or unrecognized potential?, Advances in Water Resources, 51, 357-366, https://doi.org/10.1016/j.advwatres.2012.05.005, 2013.

650

Bock, R. Darrell, and Murray Aitkin.: Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. Psychometrika 46.4, 1981.

Brandes, E. A.: Optimizing Rainfall Estimates with the Aid of Radar, Journal of Applied Meteorology (1962-1982), 14, 1339-1345, 1975.

655

Chantraket, P., Detyothin, C., Pankaew, S., and Kirtsaeng, S.: An Operational Weather Radar-Based Calibration of Z-R Relationship over Central Region of Thailand, International Journal of Engineering Issues, 2016, 92-100, 2016.

Chumchean, S., Seed, A., and Sharma, A.: Correcting of real-time radar rainfall bias using a Kalman filtering approach, Journal

660   of Hydrology, 317, 123-137, 10.1016/j.jhydrol.2005.05.013, 2006.

Collinge, V.K.: Weather radar calibration in real time: prospects for improvement. In: Collier, C.G. (Ed.), Hydrology Applications of Weather Radar. Ellis Horwood, Chinchester, UK, 1991.

665 Collinge, V.K., and Kirby, C.: Weather Radar and Flood Forecasting, Wiley, Chichester, England, 2003.

Corral, C., Berenguer, M., Sempere-Torres, D., Poletti, L., Silvestro, F., and Rebora, N.: Comparison of two early warning systems for regional flash flood hazard forecasting, Journal of Hydrology, 572, 603-619, https://doi.org/10.1016/j.jhydrol.2019.03.026, 2019.

670

Creutin, J.-D., and Borga, M.: Radar hydrology modifies the monitoring of flash-flood hazard, Hydrological Processes, 17, 1453-1456, 10.1002/hyp.5122, 2003.

Davids, J. C., Devkota, N., Pandey, A., Prajapati, R., Ertis, B. A., Rutten, M. M., Lyon, S. W., Bogaard, T. A., and van de 675 Giesen, N.: Soda Bottle Science-Citizen Science Monsoon Precipitation Monitoring in Nepal, Frontiers in Earth Science, 7, 10.3389/feart.2019.00046, 2019.

Debele, B., Srinivasan, R., and Parlange, J. J. A. i. W. R.: Accuracy evaluation of weather data generation and disaggregation methods at finer timescales, 30, 1286-1300, 2007.

680

Dinku, T., Anagnostou, E. N., and Borga, M.: Improving Radar-Based Estimation of Rainfall over Complex Terrain, Journal of Applied Meteorology, 41, 1163-1178, 10.1175/1520-0450(2002)041<1163:IRBEOR>2.0.CO;2, 2002.

Doelling, I. G., Joss, J., and Riedl, J.: Systematic variations of Z–R-relationships from drop size distributions measured in 685 northern Germany during seven years, Atmospheric Research, 47-48, 635-649, https://doi.org/10.1016/S0169-8095(98)00043-X, 1998.

Doviak, R. J. and Zmic, D. S.: Doppler Radar and Weather Observations, 1st edition, Acadamic Press Inc., San Diego, California, U.S.A., 575 pp., 1984.

690

Fabry, F., Bellon, A., Duncan, M. R., and Austin, G. L.: High resolution rainfall measurements by radar for very small basins: the sampling problem reexamined, Journal of Hydrology, 161, 415-428, https://doi.org/10.1016/0022-1694(94)90138-4, 1994.

Fulton, R. A., Breidenbach, J. P., Seo, D.-J., Miller, D. A., O'Bannon, T. J. W., and forecasting: The WSR-88D rainfall 695 algorithm, 13, 377-395, 1998.

Gao, Fuchang, and Lixing Han.: Implementing the Nelder-Mead simplex algorithm with adaptive parameters. Computational Optimization and Applications 51.1: 259-277, 2012.

700   Germann, U., Galli, G., Boscacci, M., and Bolliger, M.: Radar precipitation measurement in a mountainous region, Quarterly Journal of the Royal Meteorological Society, 132, 1669-1692, 10.1256/qj.05.190, 2006.

Hagen, M., and Yuter, S.: Relations between Radar Reflectivity, Liquid-Water Content, and Rainfall Rate during the MAP SOP, Quarterly Journal of the Royal Meteorological Society, 129, 477-493, 10.1256/qj.02.23, 2003.

705

Hanchoowong, R., Weesakul, U., and Chumchean, S.: Bias correction of radar rainfall estimates based on a geostatistical technique, ScienceAsia, 38, 373, https://doi.org/10.2306/scienceasia1513-1874.2012.38.373, 2012.

Harvey, Andrew C. Forecasting, structural time series models and the Kalman filter. Cambridge university press, 1990.

710

Hitschfeld, W., and Bordan, J.: Errors inherent in the radar measurement of rainfall at attenuating wavelengths, Journal of Meteorology, 11, 58-67, 10.1175/1520-0469(1954)011<0058:EIITRM>2.0.CO;2, 1954.

Jordan, P., Seed, A., and Austin, G.: Sampling errors in radar estimates of rainfall, Journal of Geophysical Research, 105, 715   2247-2250, 10.1029/1999JD900130, 2000.

Joss, J. and Waldvogel, A.: Raindrop size distribution and Doppler velocities, in: Proceedings 14th Radar Meteorology Conference, Boston, 1970.

720   Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems, Journal of Basic Engineering, 82, 35-45, 10.1115/1.3662552, 1960.

Kim, J. and Yoo, C.: Using Extended Kalman Filter for Real-time Decision of Parameters of Z-R Relationship, Journal of Korea Water Resources Association, 47, 10.3741/JKWRA.2014.47.2.119, 2014.

725

Klazura, G. E.: Differences between Some Radar-Rainfall Estimation Procedures in a High Rain Rate Gradient Storm, Journal of Applied Meteorology and Climatology, 20, 1376, 10.1175/1520-0450(1981)020<1376:Dbsrre>2.0.Co;2, 1981.

Koutsoyiannis, D.: Rainfall disaggregation methods: Theory and applications, in, 1–23, 10.13140/RG.2.1.2840.8564, 2003.

730

Lagarias, J., Reeds, J., Wright, M., and Wright, P.: Convergence Properties of the Nelder--Mead Simplex Method in Low Dimensions, SIAM Journal on Optimization, 9, 112-147, 10.1137/S1052623496303470, 1998.

Luersen, M. A. and Le Riche, R.: Globalized Nelder–Mead method for engineering optimization, Computers & Structures, 82,
735   2251-2260, https://doi.org/10.1016/j.compstruc.2004.03.072, 2004.

Mapiam, P. P. and Sriwongsitanon, N.: Climatological Z-R relationship for radar rainfall estimation in the upper Ping river basin, ScienceAsia, 34, 215-222, 10.2306/scienceasia1513-1874.2008.34.215, 2008.

740   Mapiam, P. P., Sharma, A., Chumchean, S., and Sriwongsitanon, N.: Runoff estimation using radar and rain gage data, The 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation, Cairns Australia, July 2009, 13-17, 2009a.

Mapiam, P. P., Sriwongsitanon, N., Chumchean, S., and Sharma, A.: Effects of Rain Gauge Temporal Resolution on the
745   Specification of a ZR Relationship, Journal of Atmospheric and Oceanic Technology - J ATMOS OCEAN TECHNOL, 26, 10.1175/2009JTECHA1161.1, 2009b.

Mapiam, P. P., Sharma, A., and Sriwongsitanon, N.: Defining the Z–R Relationship Using Gauge Rainfall with Coarse Temporal Resolution: Implications for Flood Forecasting, Journal of Hydrologic Engineering, 19, 10.1061/(ASCE)HE.1943-
750   5584.0000616, 2014.

Mapiam, P. P. and Chautsuk, S.: Improving Runoff Estimates by Increasing Catchment Subdivision Complexity and Resolution of Rainfall Data in the Upper Ping River Basin, Thailand, Chiang Mai University Journal of Natural Sciences, 17, 10.12982/CMUJNS.2018.0010, 2018.
755

Marshall, J. S. and Palmer, W. M. K.: THE DISTRIBUTION OF RAINDROPS WITH SIZE, Journal of Meteorology, 5, 165-166, 10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2, 1948.

Mobile Water Management: https://mobilewatermanagement.nl/, last access: 25 October 2020.
760

Paulat, M., Frei, C., Hagen, M., and Wernli, H.: A gridded dataset of hourly precipitation in Germany: Its construction, climatology and application, Meteorologische Zeitschrift, 17, 719-732, 10.1127/0941-2948/2008/0332, 2008.

Proietti, Tommaso, and Alessandra Luati.: Maximum likelihood estimation of time series models: the Kalman filter and beyond. Handbook of Research Methods and Applications in Empirical Macroeconomics. Edward Elgar Publishing, 2013.

Pulido, M., Tandeo, P., Bocquet, M., Carrassi, A., and Lucini, M.: Stochastic parameterization identification using ensemble Kalman filtering combined with maximum likelihood methods, Tellus A: Dynamic Meteorology and Oceanography, 70, 1-17, 10.1080/16000870.2018.1442099, 2018.

Rogers, R. R.: The effect of variable target reflectivity on weather radar measurements, Quarterly Journal of the Royal Meteorological Society, 97, 154-167, https://doi.org/10.1002/qj.49709741203, 1971.

Rosenfeld, D., Atlas, D., Wolff, D., and Amitai, E.: Beamwidth Effects on Z-R Relations and Area-integrated Rainfall, Journal of Applied Meteorology and Climatology, 31, 454-464, 10.1175/1520-0450(1992)031<0454:BEOZRR>2.0.CO;2, 1992.

Rosenfeld, D., Wolff, D., and Atlas, D.: General Probability-matched Relations between Radar Reflectivity and Rain Rate, Journal of Applied Meteorology and Climatology, 10.1175/1520-0450(1993)032<0050:GPMRBR>2.0.CO;2, 1993.

See, L.: A Review of Citizen Science and Crowdsourcing in Applications of Pluvial Flooding, 7, 10.3389/feart.2019.00044, 2019.

Seed A., Siriwardena L., Sun X., Jordan P., and Elliot J.: On the calibration of Australian weather radars, Cooperative Research Centre for Catchment Hydrology, Australia, 49 pp., 2002.

Seibert, J., Strobl, B., Etter, S., Hummer, P., and van Meerveld, H. J.: Virtual Staff Gauges for Crowd-Based Stream Level Observations, 7, 10.3389/feart.2019.00070, 2019.

Seo, D. J., Breidenbach, J. P., and Johnson, E. R.: Real-time estimation of mean field bias in radar rainfall data, Journal of Hydrology, 223, 131-147, https://doi.org/10.1016/S0022-1694(99)00106-7, 1999.

Seo, D. J. and Breidenbach, J. P.: Real-Time Correction of Spatially Nonuniform Bias in Radar Rainfall Data Using Rain Gauge Measurements, Journal of Hydrometeorology - J HYDROMETEOROL, 3, 10.1175/1525-7541(2002)003<0093:RTCOSN>2.0.CO;2, 2002.

Seo, B.-C., Krajewski, W., and Ryzhkov, A.: Evaluation of the Specific Attenuation Method for Radar-Based Quantitative Precipitation Estimation: Improvements and Practical Challenges, Journal of Hydrometeorology, 21, 10.1175/JHM-D-20-0030.1, 2020.

800 Shi, Z., Wei, F., and Venkatachalam, C.: Radar-Based Quantitative Precipitation Estimation for the Identification of Debris-Flow Occurrence over Earthquake affected Region in Sichuan, China, Natural Hazards and Earth System Sciences Discussions, 1-29, 10.5194/nhess-2017-308, 2017.

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

Smith, P. L., Myers, C. G., and Orville, H. D.: Radar Reflectivity Factor Calculations in Numerical Cloud Models Using Bulk Parameterization of Precipitation, Journal of Applied Meteorology (1962-1982), 14, 1156-1165, 1975.

810 Smith, J. A. and Krajewski, W. F.: Estimation of the Mean Field Bias of Radar Rainfall Estimates, Journal of Applied Meteorology and Climatology, 30, 397, 10.1175/1520-0450(1991)030<0397:Eotmfb>2.0.Co;2, 1991.

Smith, J., Baeck, M., Meierdiercks, K., Miller, A., and Krajewski, W.: Radar Rainfall Estimation for Flash Flood Forecasting in Small Urban Watersheds, Advances in Water Resources, 30, 2087-2097, 10.1016/j.advwatres.2006.09.007, 2007.
815
Smith, J., Hui, E., Steiner, M., Baeck, M., Krajewski, W., and Ntelekos, A.: Variability of rainfall rate and raindrop size distributions in heavy rain, Water Resources Research - WATER RESOUR RES, 45, 10.1029/2008WR006840, 2009.

Srivastava, S., Vaddadi, S., and Sadistap, S.: Smartphone-based System for water quality analysis, Applied Water Science, 8,
820 10.1007/s13201-018-0780-0, 2018.

Steiner, M., Houze, R., and Yuter, S.: Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data, Journal of Applied Meteorology, 34, 1978-2007, 10.1175/1520-0450(1995)034<1978:CCOTDS>2.0.CO;2, 1995.
825
Steiner, M., Smith, J., Kessinger, C., and Ferrier, B.: Evaluation of algorithm parameters for radar data quality control, 1999.

Steiner, M., and Smith, J. A.: Reflectivity, Rain Rate, and Kinetic Energy Flux Relationships Based on Raindrop Spectra, Journal of Applied Meteorology, 39, 1923-1940, 10.1175/1520-0450(2000)039<1923:RRRAKE>2.0.CO;2, 2000.

830

Sun, X., Mein, R. G., Keenan, T. D., and Elliott, J. F.: Flood estimation using radar and raingauge data, Journal of Hydrology, 239, 4-18, https://doi.org/10.1016/S0022-1694(00)00350-4, 2000.

Tokay, A. and Short, D. A.: Evidence from Tropical Raindrop Spectra of the Origin of Rain from Stratiform versus Convective

835   Clouds, Journal of Applied Meteorology (1988-2005), 35, 355-371, 1996.

Uijlenhoet, R.: Raindrop size distributions and radar reflectivity–rain rate relationships for radar hydrology, Hydrology and Earth System Sciences., 5, 615-628, 10.5194/hess-5-615-2001, 2001.

840   Ulbrich, C. W.: Natural Variations in the Analytical Form of the Raindrop Size Distribution, Journal of Climate and Applied Meteorology, 22, 1764-1775, 1983.

Vieux, B. E. and Bedient, P. B.: Assessing urban hydrologic prediction accuracy through event reconstruction, Journal of Hydrology, 299, 217-236, https://doi.org/10.1016/j.jhydrol.2004.08.005, 2004.

845

Vormoor, K. and Skaugen, T.: Temporal Disaggregation of Daily Temperature and Precipitation Grid Data for Norway, Journal of Hydrometeorology, 14, 989-999, 10.1175/JHM-D-12-0139.1, 2013.

Wang, L.-P., Ochoa-Rodríguez, S., Van Assel, J., Pina, R. D., Pessemier, M., Kroll, S., Willems, P., and Onof, C.:

850   Enhancement of radar rainfall estimates for urban hydrology through optical flow temporal interpolation and Bayesian gauge-based adjustment, Journal of Hydrology, 531, 408-426, https://doi.org/10.1016/j.jhydrol.2015.05.049, 2015.

Wilson, J. W. J.: Integration of radar and raingage data for improved rainfall measurement, Journal of Applied Meteorology and Climatology, 9, 489-497, 1970.

855

Wilson, J. W. and Brandes, E. A.: Radar Measurement of Rainfall-A Summary, Bulletin of the American Meteorological Society, 60, 1048, 10.1175/1520-0477(1979)060<1048:Rmors>2.0.Co;2, 1979.

Wüest, M., Frei, C., Altenhoff, A., Hagen, M., Litschi, M., and Schär, C.: A gridded hourly precipitation dataset for Switzerland

860   using rain-gauge analysis and radar-based disaggregation, International Journal of Climatology, 30, 1764-1775, 10.1002/joc.2025, 2009.

Yoo, C. and Yoon, J.-H.: A proposal of quality evaluation methodology for radar data, Journal of The Korean Society of Civil Engineers, 30, 429-435, 2010.

865