A climatological benchmark for operational radar rainfall bias reduction

Ruben Imhoff[1,2], Claudia Brauer[1], Klaas-Jan van Heeringen[2], Hidde Leijnse[1,3], Aart Overeem[1,3], Albrecht Weerts[1,2], and Remko Uijlenhoet[1,4]

[1]Hydrology and Quantitative Water Management Group, Wageningen University & Research, Wageningen, The Netherlands
[2]Operational Water Management & Early Warning, Department of Inland Water Systems, Deltares, Delft, The Netherlands
[3]R&D Observations and Data Technology, Royal Netherlands Meteorological Institute, De Bilt, The Netherlands
[4]Department of Water Management, Delft University of Technology, Delft, The Netherlands

Correspondence: Ruben Imhoff (Ruben.Imhoff@deltares.nl)

Abstract. The presence of significant biases in real-time radar quantitative precipitation estimations (QPE) limits its use in hydro-meteorological forecasting systems. Here, we introduce CARROTS (Climatology-based Adjustments for Radar Rainfall in an OperaTional Setting), a set of fixed bias reduction factors, which vary per grid cell and day of the year. The factors are based on a historical set of 10 years of 5-min radar and reference rainfall data for the Netherlands. CARROTS is both operationally available and independent of real-time rain gauge availability, and can thereby provide an alternative to current QPE adjustment practice. In addition, it can be used as benchmark for QPE algorithm development. We tested this method on the resulting rainfall estimates and discharge simulations for twelve Dutch catchments and polders. We validated the results against the operational mean field bias (MFB) adjusted rainfall estimates and a reference dataset. This reference consists of the radar QPE, that combines an hourly MFB adjustment and a daily spatial adjustment using observations from 31 automatic and 325 manual rain gauges. Only the automatic gauges of this network are available in real-time for the MFB adjustment. The resulting climatological correction factors show clear spatial and temporal patterns. Factors are higher far from the radars and higher from December through March than in other seasons, which is likely a result of sampling above the melting layer during the winter months. Annual rainfall sums from CARROTS are comparable to the reference and outperform the MFB adjusted rainfall estimates for catchments far from the radars. This difference is absent for catchments closer to the radars. QPE underestimations are amplified when used in the hydrological model simulations. Discharge simulations using the QPE from CARROTS outperform those with the MFB adjusted product for all but one basin. Moreover, the proposed factor derivation method is robust. It is hardly sensitive to leaving individual years out of the historical set and to the moving window length, given window sizes of more than a week.

1 Introduction

Radar rainfall estimates are essential for hydro-meteorological forecasting systems. In these systems, the data are used to force hydrological models (e.g., Borga, 2002; Thorndahl et al., 2017), to initialize Numerical Weather Prediction models (e.g., Haase et al., 2000; Rogers et al., 2000) or as input data for rainfall nowcasting techniques (e.g., Ebert et al., 2004; Wilson et al., 2010; Foresti et al., 2016; Heuvelink et al., 2020; Imhoff et al., 2020a). A major disadvantage of radar quantitative

25 precipitation estimations (QPE) are the considerable biases with respect to the true rainfall, caused by three main groups of errors: (1) sources of errors related to the reflectivity measurements, (2) sources of errors in the conversion from reflectivity to rainfall rate and (3) spatio-temporal sampling errors (Austin, 1987; Joss and Lee, 1995; Creutin et al., 1997; Sharif et al., 2002; Uijlenhoet and Berne, 2008; Ochoa-Rodriguez et al., 2019; Imhoff et al., 2020b). These biases can amplify when used in hydrological models (Borga, 2002; Borga et al., 2006; Brauer et al., 2016). Hence, radar QPE requires corrections before

30 operational use in hydro-meteorological (forecasting) models.

A large number of correction methods is already available. These methods range from corrections prior to the rainfall estimations, e.g. corrections for physical phenomena such as ground clutter, attenuation, the vertical profile of reflectivity and variations in raindrop size distribution (e.g., Joss and Pittini, 1991; Berenguer et al., 2006; Cho et al., 2006; Uijlenhoet and Berne, 2008; Kirstetter et al., 2010; Qi et al., 2013; Hazenberg et al., 2013, 2014), to statistical post-processing steps for bias

35 removal in the radar QPE using rain gauge data. These post-processing methods either merge rain gauge and radar QPE from the same interval or base correction factors on the total precipitation in both products over a past period, such as a number of rainy days (e.g. seven days in Park et al., 2019). An often used method is the mean field bias (MFB) correction method, which determines a spatially-averaged correction factor from the ratio between rain gauge observations and the radar QPE of the superimposed grid cells at the locations of these gauges (Smith and Krajewski, 1991; Seo et al., 1999). This method, which

40 is used operationally in the Netherlands and many other countries (Holleman, 2007; Harrison et al., 2009; Thorndahl et al., 2014; Goudenhoofdt and Delobbe, 2016), does not account for any spatial variability in the radar QPE bias, even though the bias is known to increase with increasing distance from the radar (Joss and Lee, 1995).

It is possible to account for this spatial variability with geostatistical techniques (e.g. ordinary kriging, kriging with external drift or co-kriging, Krajewski, 1987; Creutin et al., 1988; Wackernagel, 2003; Schuurmans et al., 2007; Goudenhoofdt and

45 Delobbe, 2009; Sideris et al., 2014) or Bayesian merging methods (Todini, 2001). Although these methods substantially improve the QPE in the spatial domain, all gauge-based radar QPE adjustment methods are limited by the timely availability of sufficient, and ideally quality-controlled, rain gauge observations (for an overview of methods and their limitations, see Ochoa-Rodriguez et al., 2019). The gauge networks operated by the Royal Netherlands Meteorological Institute (KNMI) are an example of this issue. Although there is approximately one station per 100 km^2, only 31 out of 356 rain gauges operate

50 automatically. The remaining 325 manual rain gauges report just once a day. Thus, only the automatic rain gauges are used for the MFB adjustment that takes place every hour in real-time (Holleman, 2007), and since recently even every five minutes.

In addition, two potential operational (forecasting) issues need to be considered when using these more advanced geostatistical and Bayesian merging methods: (1) the methods are computationally expensive, especially methods such as co-kriging and

Bayesian merging that integrate radar and rain gauges (Ochoa-Rodriguez et al., 2019), and (2) when the adjustment method

55 changes the spatial structure of the original radar rainfall fields (kriging and Bayesian methods), this may impact the continuity of the rainfall fields over time and thereby also the radar rainfall nowcasts (Ochoa-Rodriguez et al., 2013; Na and Yoo, 2018). In case of a negative impact on the nowcasts, this suggests that adjustment methods should be applied to the nowcasts as a post-processing step. To do this, the forecaster would need to estimate the future (bias) correction factors (a method for this using MFB adjustment is described in Seo et al., 1999) or simply assume that the latest correction factors are exemplary for

60 the coming hours.

Hence, operational hydro-meteorological forecasting calls for a radar rainfall adjustment approach that (1) takes the spatial variability in radar QPE errors into account and (2) is available in real time so that it can be used operationally for post-processing of radar-based rainfall forecasts, such as nowcasting. Here, we present CARROTS (Climatology-based Adjustments for Radar Rainfall in an OperaTional Setting): a set of gridded climatological adjustment factors for every day of the year, based

65 on a historical set of 10 years of 5-min radar and reference rainfall data for the Netherlands. When sufficient rain gauges are operationally available, which would allow for a robust application of more advanced geostatistical and Bayesian merging methods, CARROTS can serve as a benchmark for testing these and other more sophisticated adjustment techniques.

2 Data and methods

2.1 Radar rainfall estimates

70 The archive (2009–2018) of radar rainfall composites in this study originates from two C-band weather radars operated by KNMI (Fig. 1). Between September 2016 and January 2017, both radars were replaced by dual-polarization radars and the radar in De Bilt ('DB' in Fig. 1) was replaced by a new one in Herwijnen ('H' in Fig. 1). The radar renewals and relocation have had a limited impact on the QPE product, mainly because the operational products are not yet (fully) using the additional information from the dual-polarization (Beekhuis and Holleman, 2008; Beekhuis and Mathijssen, 2018).

75 The radar product is Doppler filtered for ground clutter. This product is then used to construct horizontal cross-sections at a nearly constant altitude of 1,500 m, called pseudo-constant plan position indicators (pseudo-CAPPI). Subsequently, range-weighted compositing is used to combine the reflectivities from both radars (Overeem et al., 2009b). Since 2013, non-meteorological echoes are removed as an additional step with a cloud-mask obtained from satellite imagery. As a final step, rainfall rates are estimated with a fixed $Z - R$ relationship (Marshall et al., 1955):

80 $$Z_{\mathrm{h}} = 200R^{1.6}. \tag{1}$$

In this equation, Z_{h} is the reflectivity at horizontal polarization (mm^6 m^{-3}, but generally given in dBZ, according to $10 \times \log_{10}[Z_{\mathrm{h}}]$) and R is the rainfall rate (mm hr^{-1}). The final product is called the unadjusted radar QPE (R_{U}) in this study.

KNMI also provides adjusted radar rainfall products, based on the aforementioned product, but adjusted with quality controlled observations from both 31 automatic hourly and 325 manual daily rain gauges (Overeem et al., 2009a,b, 2011). The

Number	Name	Size (km²)	Used model
1	Aa	836	WALRUS
2	Reusel	176	WALRUS
3	Roggelsebeek	88	WALRUS
4	Hupsel Brook	6.5	WALRUS
5	Grote Waterleiding	40	WALRUS
6	Regge	957	WALRUS
7	Dwarsdiep	83	WALRUS
8	Linde	150	Sobek RR
9	Luntersebeek	63	WALRUS
10	Gouwepolder	10	Sobek RR
11	Delfland	379	Sobek RR
12	Beemster	71	Sobek RR-CF
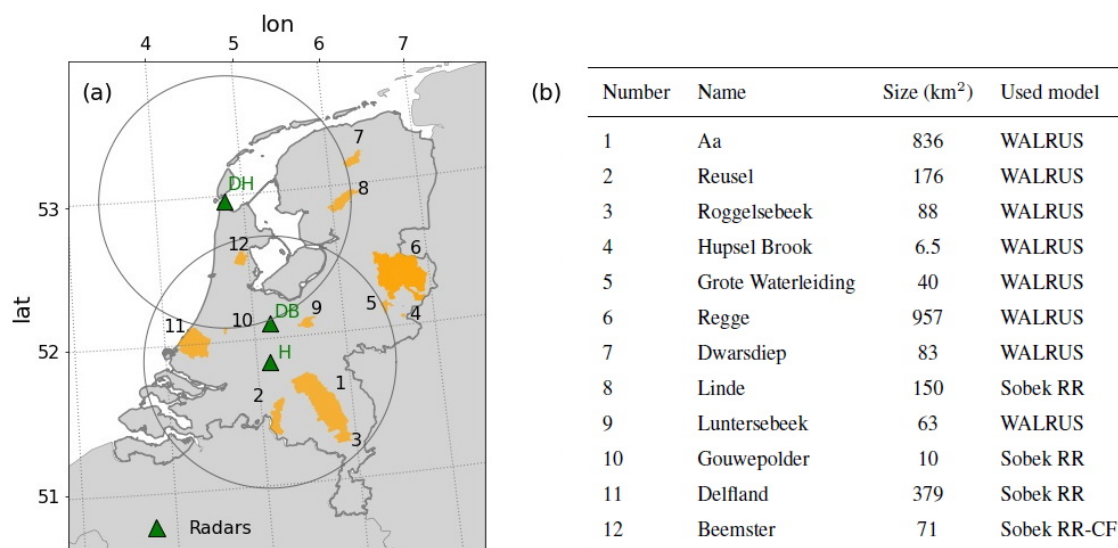
Figure 1. Overview of the basins in this study: (a) study area with the location of the three radars (green triangles) operated by KNMI and the twelve basins (orange polygons). The two grey circles indicate a range of 100 km around the radars in Den Helder (DH) and Herwijnen (H). The other radar (DB) is the radar in De Bilt, which was used until 2016 and replaced by the radar in Herwijnen; (b) list of the basin names, sizes and employed hydrological models. The numbers in the left column refer to the numbers in (a). The right column states the used model for these areas.

85   same 31 automatic rain gauges are used for the MFB adjustment method, which will be introduced in Sec. 2.2.1. In contrast to the spatially uniform MFB adjustment, the observations from the manual rain gauges are used for spatial adjustments, based on distance-weighted interpolation of these observations (Barnes, 1964; Overeem et al., 2009b). This product is considered as a reference rainfall product in the Netherlands and it is therefore also regarded as reference here (referred to as R_A in this study). The R_A data is not available in real time (available with a delay of one to two months), but it is archived and can therefore be

90   used for 'offline' methods. Both R_A and R_U have a 1-km² spatial and 5-min temporal resolution.

      The year 2008 is actually the first year in the KNMI archive of both data sets, but it was left out of the analysis here. R_U for this year showed a significantly different behaviour than the other years, especially during the first half year in which the product rarely underestimated and frequently even overestimated the rainfall sums. The reason for this behaviour is not yet fully understood. KNMI (2009) reported that spring was exceptionally dry in the north of the country and that the months January

95   and May were among the warmest on record. On some days with overestimations, clear bright band effects were visible in the radar mosaic, which may have contributed to the systematic differences.
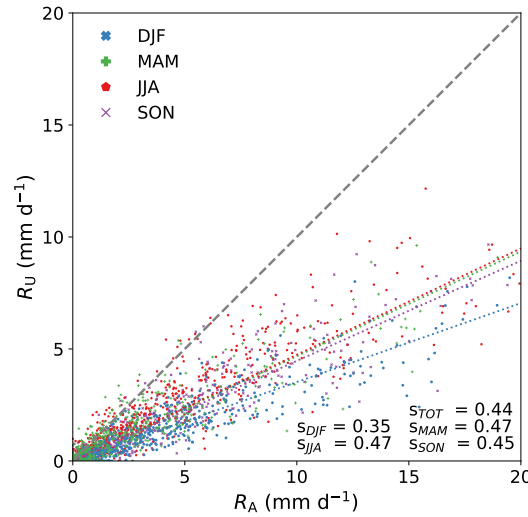
Figure 2. Scatter plot indicating the systematic discrepancy between the reference rainfall (R_A) and the unadjusted radar QPE (R_U). Shown are the daily country-average rainfall sums based on ten years (2009–2018), classified per season. The slope of a linear fit between the two rainfall products is shown in the bottom right for all observations together (indicated with 's_{TOT}') and per season.

2.2 Bias correction factors

Figure 2 indicates the need for correction of the radar rainfall product. R_U systematically underestimates the true rainfall amounts, averaged for the land surface area of the Netherlands, with 56%. This bias is not uniform in space, as will be highlighted in Sec. 3, and in time with higher underestimations during winter (on average 65%) than during the other seasons (53 – 55 %). In the following two subsections, the operationally used MFB adjustment method and the in this study proposed CARROTS method will be introduced.

2.2.1 Mean field bias adjustment

The mean field bias (MFB) adjustment method is the operational adjustment technique in the Netherlands and it was used in this study for comparison with the proposed climatological bias reduction method (Sec. 2.2.2). This method provides a spatially uniform multiplicative adjustment factor that is applied to R_U. The adjustment factor (F_{MFB}) was calculated as (Holleman, 2007; Overeem et al., 2009b):

$$F_{\mathrm{MFB}} = \frac{\sum\limits_{n=1}^{N} G(i_n, j_n)}{\sum\limits_{n=1}^{N} R_{\mathrm{U}}(i_n, j_n)}, \tag{2}$$

with $G(i_n, j_n)$ the hourly rainfall sum for gauge n at location (i_n, j_n) and $R_U(i_n, j_n)$ the unadjusted hourly rainfall sum for the
110   corresponding radar grid cell. The calculation of F_{MFB} only took place when both the rainfall sum of all rain gauges together
and the sum of all corresponding radar grid cells was at least 1.0 mm. In all other cases, $F_{MFB} = 1.0$.

The MFB adjustment factors were determined from the 1-hr accumulations of both R_U and the 31 automatic rain gauges, as
only the automatic gauges were operationally available every hour (Holleman, 2007; Overeem et al., 2009b). The adjustment
factors at the temporal resolution of the radar QPE (5 min) were assumed to equal the 1-hr adjustment factors for a given hour.
115   Moreover, this analysis took place with archived datasets, which were validated and consisted of quality-controlled rain
gauge observations. It is noteworthy that the same quality control is absent and that missing data occurs in real-time, which
can lead to deteriorating results when the MFB adjustment is applied in an operational test case.

2.2.2   CARROTS method

To derive the climatological bias correction factors for the CARROTS method, both R_U and R_A were used for the years 2009–
120   2018. The use of the reference data for this method was possible, because the CARROTS method did not require a real-time
availability of the data. The bias correction factors were determined per grid cell in the radar domain according to the following
three steps:

1. For every day in the period 2009–2018, all 5-min rainfall sums (both R_U and R_A) within a moving window of 31 days
   (the day of interest plus the fifteen days before and after it) were summed. The purpose of the moving window was to
125   smooth the systematic day-to-day variability of the estimated rainfall in the 10-year data. Sections 2.4 and 3.4 describe
   the sensitivity of the method to the moving window size.

2. For every day of the year, the 31-day sums around that day were averaged over the ten years. Thus, the value for e.g. 16
   January consisted of the average 31-day sum for the period 1 to 31 January over the ten years.

3. Finally, gridded climatological adjustment factors (F_{clim}) were calculated per day of the year as:

130   $$F_{clim}(i,j) = \frac{R_A(i,j)}{R_U(i,j)}, \tag{3}$$

with $R_A(i, j)$ the reference rainfall sum and $R_U(i, j)$ the unadjusted (operational) radar rainfall sum at grid cell (i, j) for
the ten years.

2.3   Application to twelve basins

Both bias adjustment methods were applied to the ten years (2009–2018) of R_U. In order to provide a hydro-meteorological
135   testbed, both the CARROTS and MFB adjusted QPE products (from here-on referred to as R_C and R_{MFB}, respectively) were
validated against the reference rainfall. In addition, R_C and R_{MFB} were used as input for the rainfall-runoff models of twelve
basins (a combination of catchments and polders) in the Netherlands (Fig. 1). Most of the involved water authorities use
these (lowland) rainfall-runoff models either operationally or for research purposes, often embedded in a Delft-FEWS system

(Werner et al., 2013). For this reason, most models were already calibrated (e.g. Brauer et al., 2014b; Sun et al., 2020). SOBEK
140   RR(-CF) (Stelling and Duinmeijer, 2003; Stelling and Verwey, 2006; Prinsen et al., 2010) is semi-distributed and therefore
we used sub-catchment averaged rainfall sums from the gridded radar QPE. WALRUS (Brauer et al., 2014a) is lumped, so the
radar QPE was catchment averaged and used as input. A more detailed description of both rainfall-runoff models is outside the
scope of this paper. All twelve model setups were run with a 5-min time step for the period 2009–2018. The resulting discharge
simulations were validated for the same period using the Kling-Gupta Efficiency (KGE) metric (Gupta et al., 2009).

2.4   Sensitivity analysis

As mentioned in Sec. 2.2.2, the purpose of the 31-day moving window in the factor derivation of CARROTS was to smooth
the day-to-day variability of rainfall. To test the sensitivity of the method to the employed moving window size, the adjustment
factors were re-derived for a range of moving window sizes (1 day, 1 week, 2 weeks, 6 weeks and 2 months). The derived
factors were then compared to the original factor in this study, which was based on a moving window size of 31 days, and used
150   to derive adjusted QPE products. Subsequently, these QPE products served as input for one of the 12 catchments, namely the
WALRUS model for the Aa catchment (Fig. 1), to test the effect on the simulated discharges. The Aa catchment was chosen
because the unadjusted QPE product (R_U) for this catchment has one of the highest biases of the twelve studied catchments
(see Sec. 3 and Fig. 4).

Besides the moving window choice, the length of the radar rainfall archive (ten years) was finite. To test whether or not this
155   archive length was sufficient for reaching a stable factor derivation, individual years in the ten-year archive were left out of
the CARROTS method. Hence, the adjustment factors were recalculated ten times, applied to R_U and used as input for the
WALRUS simulations for the Aa catchment.

3   Results

3.1   Seasonal variability

160   The adjustment factors from CARROTS present the spatial variability in the radar QPE errors, with generally higher adjustment
factors towards the edges of the radar domain (Fig. 3). This difference is most pronounced from December through March,
with more than two times higher factors in the south and east of the country than in the central and northwestern parts (Fig. 3a,
b and l). Figure 3 demonstrates a clear annual cycle of the adjustment factors, with higher adjustment factors from December
through March than in the other months. Figure 4a shows similar results for the catchment-averaged adjustment factors, with
165   factors ranging from 2.1 for the Beemster polder to 3.2 for the Hupsel Brook catchment in January, whereas adjustment factors
range from 1.3 for the Grote Waterleiding catchment to 1.6 for the Roggelsebeek catchment in June.

An explanation for these higher adjustment factors from December through March is that radar QPE often severely underes-
timates the rainfall amounts for stratiform systems, which regularly occur during the Dutch winter. This especially holds when
the QPE is constructed from reflectivities sampled above the melting layer (Fabry et al., 1992; Kitchen and Jackson, 1993;
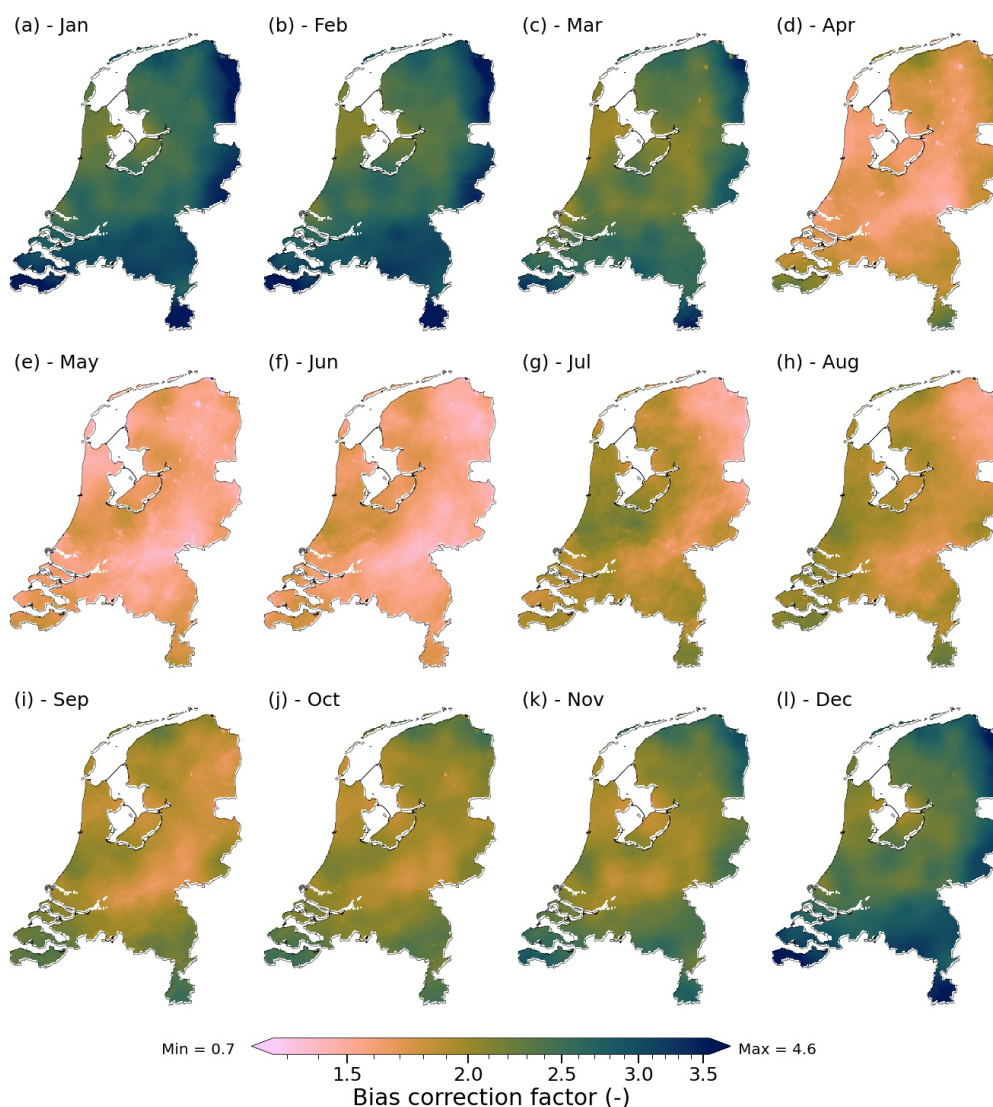
Figure 3. Spatial variability of the CARROTS factors, as derived from the archived radar and reference data for the period 2009–2018. Shown are monthly averages of the daily factors.

170 Bellon et al., 2005; Hazenberg et al., 2013). This seems to be the case here as well. A simple first-order estimation of the 0°C isotherm level, using a constant wet adiabatic lapse rate of 5.5 °C km^{-1} with ground temperature data for all rainy hours in the ten years (Fig. 4b), indicates that the 1500 m pseudo-CAPPI is generally above the 0°C isotherm level from December through March. This coincides with the months with higher adjustment factors (Fig. 4c) and could thus explain the winter effect on the adjustment factors. This effect is presumably even stronger further away from the radars, because the QPE product consists

175 of samples at even higher altitudes than 1500 m for locations at more than 120 km from the radars. Besides, an additional
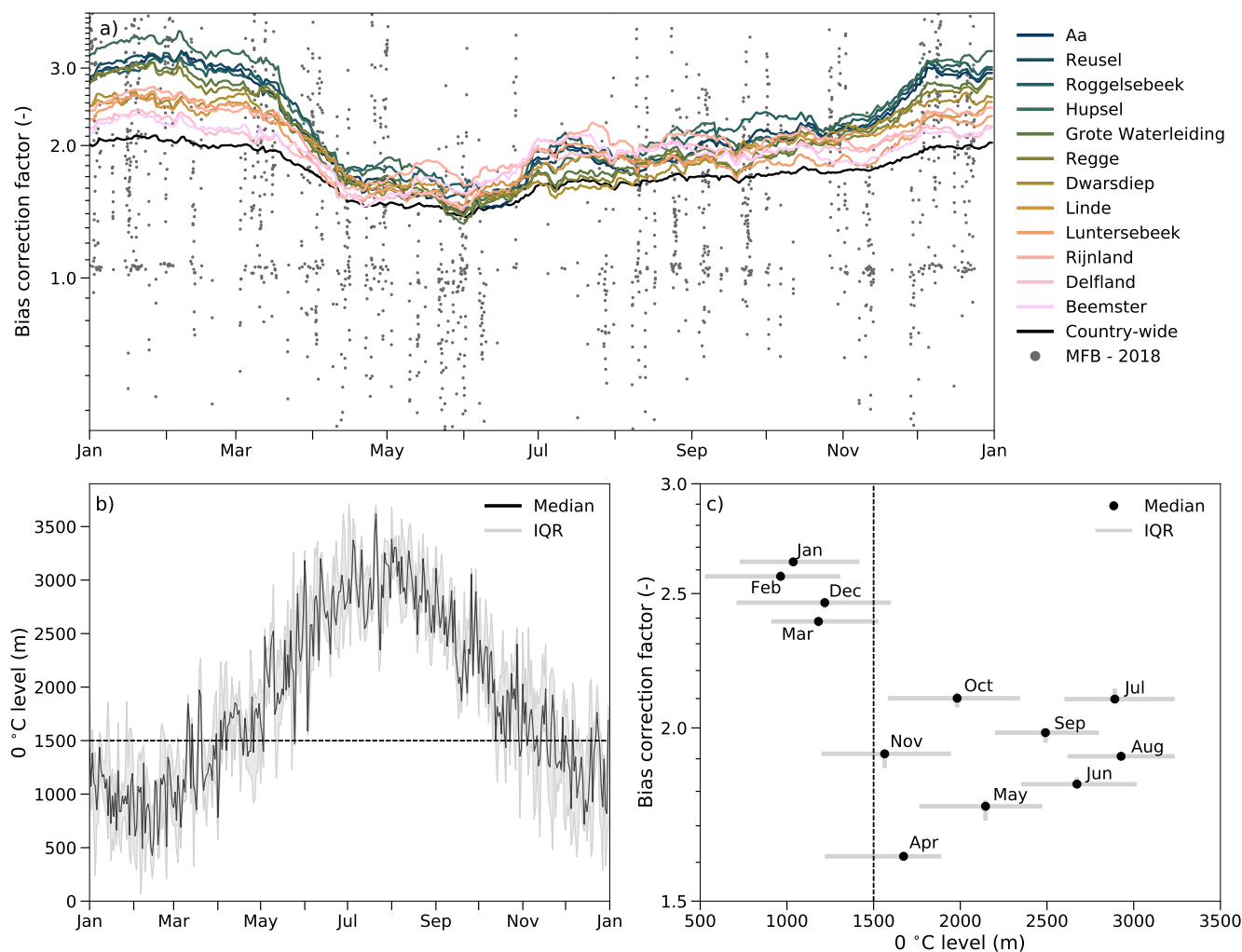
Figure 4. Seasonal dependency of the CARROTS factors and comparison with the operational MFB adjustment factor. (a) Temporal variability of the climatological daily adjustment factors for the twelve basins (colours, catchment-averaged), the country-average (black line) and of the country-wide hourly MFB factor for the (example) year 2018 (grey dots, some also fall outside the indicated range). (b) Estimate of the height of the $0°C$ isotherm at KNMI station De Bilt for all rainy hours in the ten year period, based on a constant wet adiabatic lapse rate of $5.5°C$ km^{-1}. (c) Dependency of the monthly adjustment factor on the estimated $0°C$ isotherm level for KNMI station De Bilt and the superimposed grid cell of this station. Note that for this analysis, the adjustment factor was based on only the rainfall sums within that month, the effective adjustment factor for that month, which roughly coincides with the factor for the 15th of the month in the CARROTS method. The grey bars indicate the interquartile range (IQR) for that month, based on the spread in hourly $0°C$ isotherm level estimates (the horizontal bars) and the sensitivity to leaving out individuals years in the ten year period for the factor derivation (vertical bars).

dependence of the monthly factor on the time of year that cannot be explained by temperature, seems to be present with lower adjustment factors during spring and early summer and higher factors for the subsequent period (Fig. 4c).
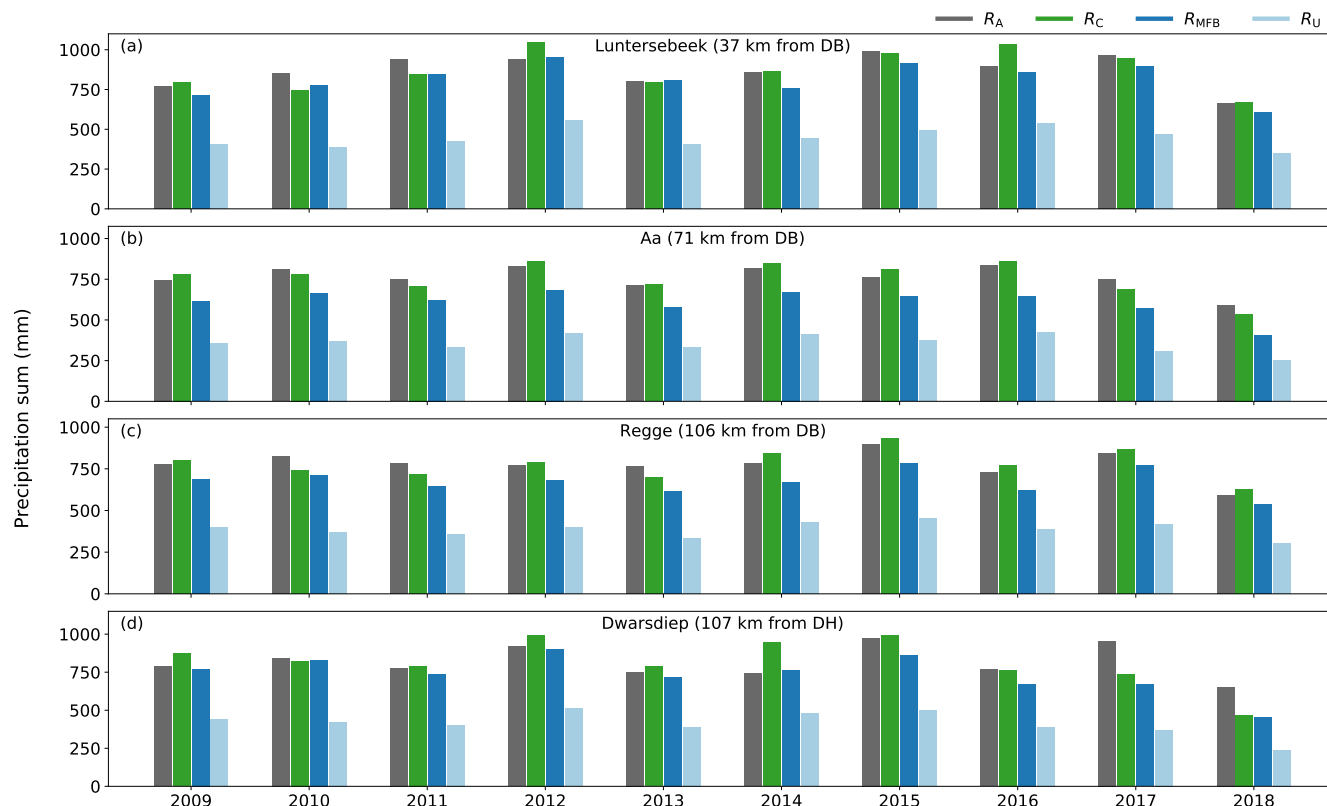
Figure 5. Effect of the adjustment factors on the catchment-averaged annual rainfall sums for a sample of four catchments that are spread over the country (and thus the radar domain): (a) Luntersebeek, (b) Aa, (c) Regge and (d) Dwarsdiep. Shown are the reference rainfall sum (R_A; grey), the estimated rainfall sum after correction with the CARROTS factors (R_C; green), the estimated rainfall sum after correction with the MFB adjustment factors (R_{MFB}; dark blue) and the rainfall sum with the unadjusted radar rainfall estimates (R_U; light blue). The distance between the catchment center and the closest radar in the domain is given in the title (DH is Den Helder and DB is De Bilt). The radar in Herwijnen, which replaced the radar in De Bilt, is not included here, because this radar was operational for the shortest time in this analysis.

3.2 Annual rainfall sums

An advantage of the MFB adjustment is that it corrects for the circumstances during that specific day and thus also for instances
180 with overestimations (Fig. 4a). The negative effect of the spatial uniformity of the factor, however, becomes apparent in Fig. 5, which compares the annual precipitation sums of the two adjusted radar rainfall products with the reference and R_U for four of the catchments. Both adjusted products manage to significantly increase the QPE towards the reference, which is clearly needed, as R_U underestimates the rainfall amounts by 50% or more for these four catchments. The CARROTS QPE (R_C) outperforms the MFB adjusted QPE (R_{MFB}) for the Aa and Regge catchments, which are located in the far south and east of

185   the country, respectively. R_{MFB} still underestimates the annual reference rainfall sums with on average 20% for the Aa and

13% for the Regge, while this is on average only 5% (both under- and overestimations) for R_{C}.

The MFB adjusted QPE performs better for the Dwarsdiep polder (10% underestimation) and Luntersebeek catchment (6%

underestimation) due to their location in the radar mosaic. The Luntersebeek catchment (central Netherlands, Fig. 1) is located

closer to both radars. There, R_{MFB} generally performs better and sometimes even overestimates the true rainfall, which is

190   consistent with Holleman (2007). Holleman (2007) also indicates that the MFB adjustment performs well in the north of the

country, where the Dwarsdiep catchment is located, even though this region is closer to the edge of the radar domain. The

CARROTS QPE tends to overestimate the rainfall amount of both basins for some years (e.g. with 16% for the Luntersebeek

in 2016). So, the annual rainfall sums are not necessarily better estimated by R_{C} for these two catchments.

Summarizing, the CARROTS factors have a clear annual cycle, with generally higher adjustment factors further away from

195   the radars (Sec. 3.1). This results in estimated annual rainfall sums that are closer to the reference than with the MFB adjusted

QPE for regions close to the edges of the radar domain. This effect is expected to become more pronounced when the adjusted

QPE products are used for discharge simulations.

3.3   Effect on simulated discharges

The severe underestimations of R_{U} have a considerable effect on the discharge simulations for the twelve basins (Fig. 6). This

200   leads to hardly any discharge response and thus negative KGE values for most basins as compared to discharge simulations with

the reference rainfall data. The effect is most pronounced for the freely draining catchments in the east and south of the country.

These catchments are more driven by groundwater flow than the polders in the west of the country. Groundwater flow gets

hardly replenished, because of similar estimated annual evapotranspiration and R_{U} sums, resulting in too low baseflows. The

polders, especially Delfland and Beemster, are an exception to this, because they are less driven by groundwater-fed baseflow

205   and more by direct runoff from greenhouses or upward seepage flows, which makes them more responsive to individual rainfall

events leading to higher KGE values (with R_{U} as input) compared to the other basins.

The model runs using R_{MFB} as input significantly improve the simulated discharges, compared to the runs with R_{U}. Nev-

ertheless, the model runs still strongly underestimate the simulated discharges compared to those from the reference runs for

the catchments in the south and east of the country (Fig. 6a–f). This is particularly noticeable for the catchments Reusel (KGE

210   = 0.26) and Roggelsebeek (KGE = 0.04). The spatial uniformity of the MFB factors is identified as the cause of these effects,

because the MFB method can not correct for the sources of errors leading to the biased QPE in space. This already led to clear

underestimations in the annual rainfall sums for these regions (Fig. 5).

The CARROTS QPE outperforms R_{MFB}, when this product is used as input for the twelve rainfall-runoff models. This is

not exclusively the case for the six catchments in the east and south of the country (Fig. 6a–f), but also for the other polder and

215   catchment areas. The exception to this is the Beemster polder (which is mostly upward seepage driven), although the difference

in performance is small, with a KGE of 0.92 (using R_{C}) versus 0.96 for R_{MFB}, as compared to the reference run.
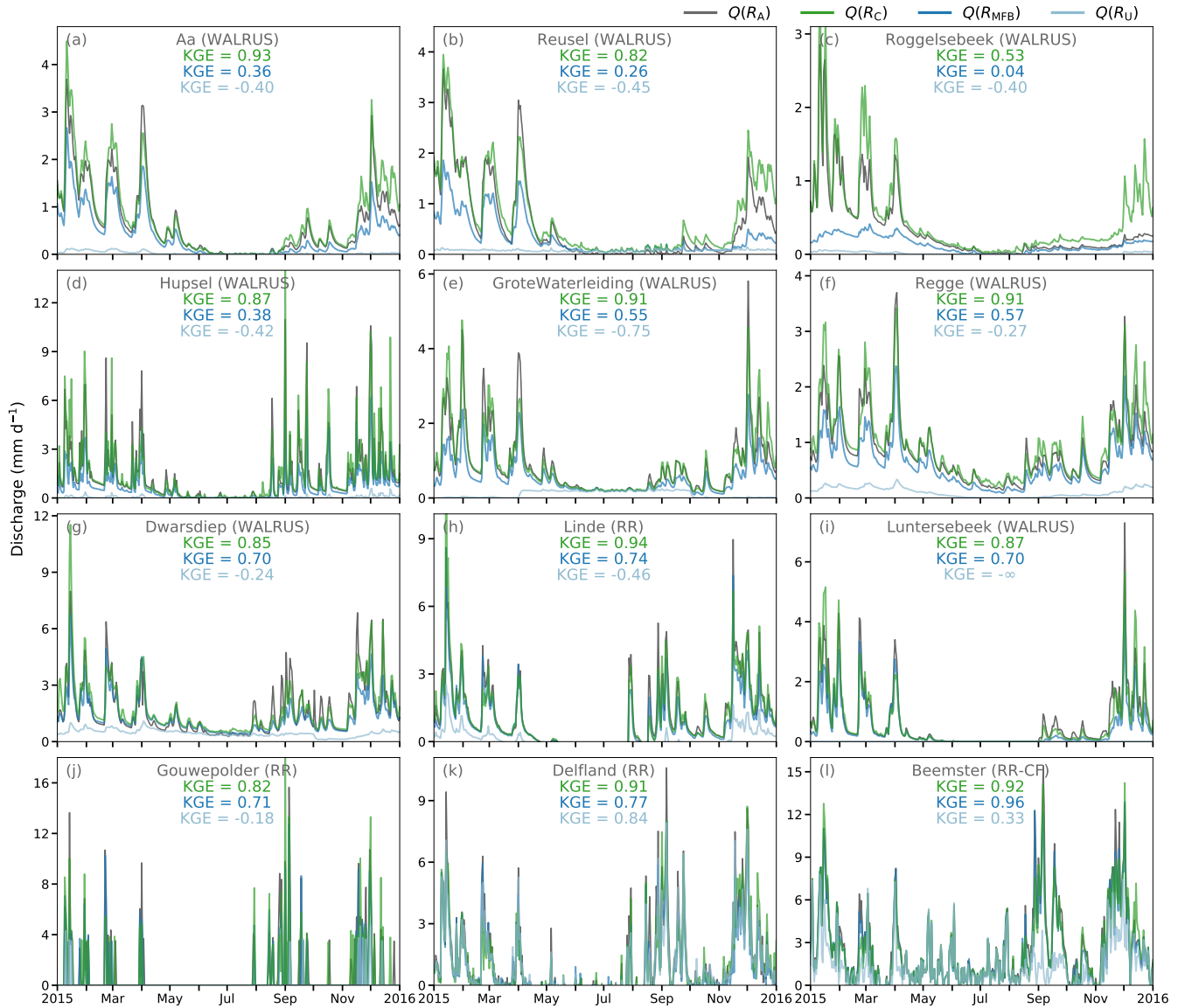
Figure 6. Differences in simulated discharges for the twelve basins (a–l) as a result of the differences between rainfall estimates. The models are run for the period 2009–2018 with the following rainfall products as input: the reference (R_A; grey), the QPE corrected with the CARROTS factors (R_C; green), the MFB adjusted QPE (R_{MFB}; dark blue) and the unadjusted radar rainfall estimates (R_U; light blue). Only the simulated discharges for 2015 are shown here for clarity; the KGE is based on all years.

3.4 Sensitivity analysis

The use of a different moving window size hardly influences the CARROTS factors for moving window sizes of two weeks or longer, but this does not hold for moving window sizes of a day or, to a lesser extent, one week (Fig. 7a). The factor derived
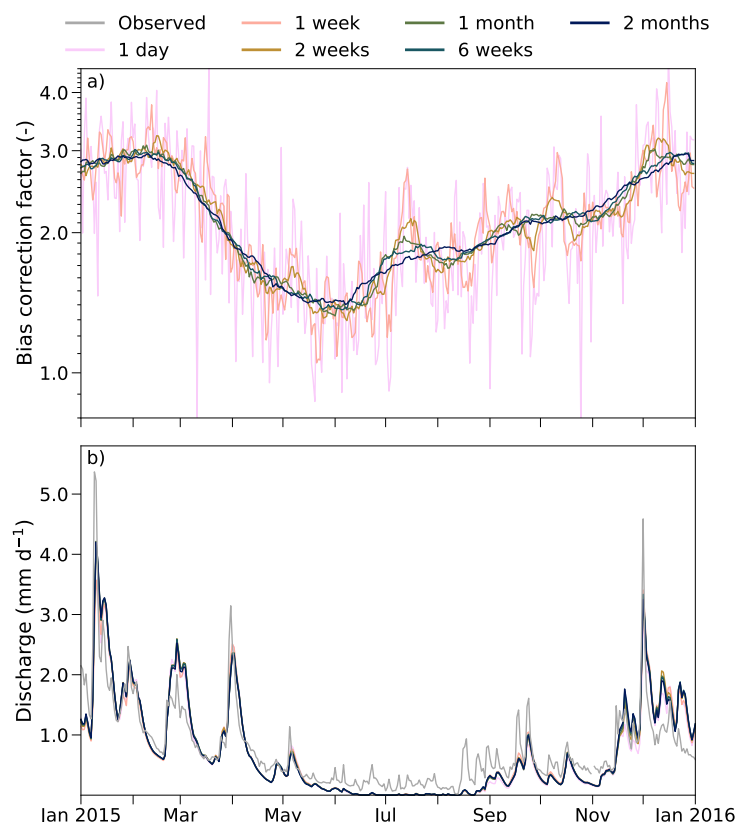
Figure 7. Sensitivity of the CARROTS factor derivation to the moving window size. (a) The adjustment factors for the Aa catchment for six different moving window sizes. The moving window size of 31 days was used in the methodology of this study. (b) The effect of the six moving window sizes in (a) on the simulated discharges for the Aa. Similar to Fig. 6, discharge was simulated for the full period (2009–2018), but only 2015 is shown here. The grey line indicates the observed discharge.

with a moving window size of one day fluctuates heavily from day to day. This suggests that the adjustment factor is still quite sensitive to individual events in the 10-year period, when a moving window size of seven days or less is used. In contrast to this, the differences between these six sets of CARROTS factors (Fig. 7a) lead to minimal variations in the simulated discharges for the Aa catchment, when these factors are used to adjust the input QPE (Fig. 7b). Differences in timing and magnitude (0.2–0.3 mm d^{-1}) are visible during peaks and recessions, for instance in early April. However, these are small compared to the differences between the model runs with R_{C} and R_{MFB} (Fig. 6). Concluding, a 31-day smoothing of the climatological adjustment factor is warranted.

In addition, leaving individual years out of the ten-year archive has a limited impact on the CARROTS factors (see also the vertical bars in Fig. 4c). Similar to the aforementioned results for the moving window size analysis, it leads to hardly any variations in the simulated discharges for the Aa catchment (not shown here). This suggests that the ten-year archive length was sufficiently long for the factor derivation.

4 Discussion

In this study, we introduced the CARROTS method to derive adjustment factors that reduce the bias in radar rainfall estimates. We derived these factors using 10 years of 5-min radar and reference rainfall data for the Netherlands. The method and resulting QPE product generally outperformed the mean field bias (MFB) adjustment that is used operationally in the Netherlands, especially when the QPE products were used as input for hydrological model runs.

The main difference with the MFB adjustment is the presence of a high-density network of (manual) rain gauges in the reference dataset, a dataset that is not available in real-time. This allows for spatial adjustments. Overeem et al. (2009b) demonstrate that this reference dataset mostly depends on the daily spatial adjustments from the manual rain gauges, while the higher-frequency MFB adjustment based on the automatic gauges plays a smaller role in the adjustments of this reference product. According to Saltikoff et al. (2019), at least 40 countries have an archive of historical radar data for a period of ten years or more. The proposed CARROTS method is potentially valuable for these countries, especially when the density of their network of automatic rain gauges is, similar to the Netherlands, significantly smaller than the total network of rain gauges.

MFB adjustment of radar rainfall fields is still the most frequently applied adjustment method (Holleman, 2007; Harrison et al., 2009; Thorndahl et al., 2014; Goudenhoofdt and Delobbe, 2016). The results indicate that this choice may be reconsidered, at least for the Netherlands. This could also hold for other regions, especially mountainous regions where the uniformity of the MFB adjustment factor is likely not sufficient to correct for all orography-related errors (Borga et al., 2000; Anagnostou et al., 2010).

However, the proposed CARROTS method has to be recalculated for every change in the radar setup, calibration, additional post-processing steps (e.g. VPR corrections, Hazenberg et al., 2013) or final composite generation algorithm. For instance, including a new radar in the composite would require a recalculation of the adjustment factors, thereby assuming the presence of an archive of the new composite product. This could potentially limit the usefulness of the proposed method.

Although the results are promising, this method is not expected and meant to outperform more advanced spatial QPE adjustment methods, such as geostatistical and Bayesian merging methods (for an overview of methods and their limitations, see Ochoa-Rodriguez et al., 2019). A major advantage of these methods is the real-time derivation of spatial adjustment factors, in contrast to the proposed method in this study, which was solely based on historical data. The MFB adjustment factors can also be derived in near real-time, but are uniform in space, which can explain the worse performance as compared to the proposed method in this study. A disadvantage of geostatistical and Bayesian merging methods is that they are computationally expensive and require the real-time availability of a dense network of rain gauges. Instead, we consider the proposed climatological radar rainfall adjustment method as a benchmark for the development and testing of operational radar QPE adjustment techniques. In essence, this means that a QPE adjustment method should at least be able to outperform the proposed method in this study.

Another possible option would be to combine the CARROTS method with the real-time application of the MFB adjustment, i.e. CARROTS is applied and the resulting QPE is then adjusted in time with real-time MFB adjustment factors. This would allow for real-time temporal corrections of the QPE, without the need for a high density of rain gauges in real-time, while the corrections in space are based on the (historical) CARROTS factors.

265   As mentioned in the previous paragraph, the climatological adjustment factor is not calculated for the current meteorological conditions and resulting QPE errors, which could lead to considerable errors during extreme events. Nonetheless, this is also the case for the MFB adjustment technique (Schleiss et al., 2020). The absolute errors for the 10 highest daily sums in this study for the Aa and Hupsel Brook catchments (one of the largest and the smallest catchment in the study) are similar for the MFB and climatological adjustment methods, with on average a 20% difference with the reference (this would have been

270   50–60 % without corrections). Note that for individual events in these twenty extremes, the errors can still reach 48% for the QPE adjusted with CARROTS and 64% for the MFB adjusted QPE.

        Finally, the CARROTS factors were derived with the reference rainfall data for the Netherlands. The same data was used as reference in this study. Although the use of the same data as training and validation set is sub-optimal, leaving out individual years has had a limited impact on the estimated adjustment factors and the resulting QPE and discharge simulations (see also

275   the vertical bars in Fig. 4c).

5   Conclusions

A known issue of radar quantitative precipitation estimations (QPE) are the significant biases with respect to the true rainfall amounts. For this reason, radar QPE adjustments are needed for operational use in hydro-meteorological (forecasting) models. Current QPE adjustment methods depend on the timely availability of quality-controlled rain gauge observations from dense

280   networks. This especially applies to methods that correct for the spatial variability in the QPE errors. To overcome this issue and to provide a benchmark for future QPE algorithm development, we have presented CARROTS (Climatology-based Adjustments for Radar Rainfall in an OperaTional Setting): a set of gridded climatological adjustment factors for every day of the year. The factors were based on a historical set of 10 years of 5-min radar rainfall data and a reference dataset for the Netherlands. The climatological adjustment factors were compared with the mean field bias (MFB) adjustment factors, which

285   are used operationally in the Netherlands. For the period 2009–2018, both adjustment factors were validated on the estimated annual rainfall sums and the effect of the adjusted QPE products on simulated discharges with the rainfall-runoff models for twelve Dutch basins.

        The CARROTS factors show clear spatial and temporal patterns, with higher adjustment factors towards the edges of the radar domain. This is caused by larger QPE errors further away from the radars. The factors are also higher from December

290   through March than in other seasons. This is likely a result of sampling above the melting layer during these months, which causes higher underestimations in the unadjusted radar rainfall product.

        Although the MFB factors are based on the current over- or underestimations in the QPE, the factor is spatially uniform and does not correct for spatial errors. This directly impacts the adjusted QPE. The MFB adjusted QPE leads to annual rainfall sums that still underestimate those of the reference for the catchments in the east and south of the country (towards the edge

295   of the radar domain). This bias is almost absent for the CARROTS factors (up to 5% over- and underestimation for the same catchments). For basins closer to radars, this effect decreases and both adjustment methods perform well.

The effects of both adjustment methods on the QPE is amplified when they are used as input for the rainfall-runoff models of the twelve studied basins. The discharge simulations with the CARROTS QPE outperforms those using the MFB adjusted QPE for all but one basin. For the Netherlands, these results indicate that the operationally used MFB adjustment performs

300 worse than the proposed climatological adjustment factor for hydrological applications.

Despite the aforementioned results, the CARROTS method has two main limitations: (1) for every change in the radar setup, the radar calibration, post-processing algorithms or the final composite generation method, the adjustment factors have to be recalculated; (2) the factor is not calculated for the actual meteorological conditions and resulting QPE errors, which could lead to considerable errors during extreme events. Nonetheless, the latter is also the case for the MFB adjustment technique

305 (Schleiss et al., 2020), even though the MFB factors are derived in real-time.

The main advantage of the introduced method is the continuous availability of spatially distributed adjustment factors, due to the independence of timely rain gauge observations. This is beneficial for operational use. In addition, the CARROTS factors are shown to be robust, as the derivation is not found to be sensitive to leaving out individual years or the used moving window, especially when this window is longer than a week.

310 Finally, this method is not expected and meant to outperform more advanced spatial QPE adjustment methods (which require data from dense rain gauge networks for robust application), but it can serve as a benchmark for the development and testing of more advanced operational radar QPE adjustment techniques. QPE adjustment methods (including CARROTS) greatly benefit from a denser, frequently-available rain gauge network. From that perspective, crowd-sourced personal weather stations hold a promise for improving radar rainfall products, given their direct surface measurements and dense networks (Vos et al.,

315 2019). This also holds for rain gauge observations from other governmental or third parties, e.g. the water authorities in the Netherlands. Hence, we think that this could further improve radar rainfall products in the near future.

Data availability. The archived gauge-adjusted (reference) and unadjusted radar QPE are available via https://dataplatform.knmi.nl/catalog/ datasets/index.html?x-dataset=rad_nl25_rac_mfbs_em_5min&&x-dataset-version=2.0 and

320 https://doi.org/10.4121/uuid:05a7abc4-8f74-43f4-b8b1-7ed7f5629a01. The daily climatological bias adjustment factors for the Netherlands can be found at: https://doi.org/10.4121/13573814. The used parameter values for WALRUS and SOBEK RR are operationally used by the water authorities and should therefore be requested via them. Interested readers are invited to contact the authors about this. The used color schemes in Fig. 3 and 4 are described in Crameri (2018) and Crameri et al. (2020), and are available via: https://doi.org/10.5281/zenodo. 4153113.

325

Video supplement. The supplement contains a visualisation of the daily spatial variability of the CARROTS factors.

Author contributions. All authors were involved in the design of the study lay-out. RI carried out the analyses with contributions from CB and KJvH and input from HL, AO, AW and RU. RI prepared the manuscript and all co-authors contributed to the content and improvement
330   of the manuscript.

Competing interests. The authors declare that they have no competing interests.

Acknowledgements. We are thankful for the catchment data, model parameters, the operational Delft-FEWS systems and information that were provided by the involved Dutch water authorities: Hoogheemraadschap Delfland, Hoogheemraadschap Hollands Noorderkwartier,
335   Hoogheemraadschap Rijnland, Waterschap Aa en Maas, Waterschap De Dommel, Wetterskip Fryslân, Waterschap Limburg, Waterschap Noorderzijlvest, Waterschap Rijn en IJssel, Waterschap Vallei en Veluwe and Waterschap Vechtstromen. In addition, we would like to thank Xiaohan Li and Pieter Hazenberg (Deltares) for answering our questions and their interest in our work. This study was supported by funding from the DAISY2-project, supported by the European Regional Development Fund (Grant PROJ-00581) and Deltares' Strategic Research Program.

340 **References**

Anagnostou, M. N., Kalogiros, J., Anagnostou, E. N., Tarolli, M., Papadopoulos, A., and Borga, M.: Performance evaluation of high-resolution rainfall estimation by X-band dual-polarization radar for flash flood applications in mountainous basins, Journal of Hydrology, 394, 4–16, https://doi.org/10.1016/j.jhydrol.2010.06.026, 2010.

Austin, P. M.: Relation between measured radar reflectivity and surface rainfall, Monthly Weather Review, 115, 1053–1070,
345   https://doi.org/10.1175/1520-0493(1987)115<1053:RBMRRA>2.0.CO;2, 1987.

Barnes, S. L.: A technique for maximizing details in numerical weather map analysis, Journal of Applied Meteorology, 3, 396–409, 1964.

Beekhuis, H. and Holleman, I.: From pulse to product, highlights of the digital-IF upgrade of the Dutch national radar network, in: Proceedings of the Fifth European Conference on Radar in Meteorology and Hydrology (ERAD 2008), Helsinki, Finland, https://cdn.knmi.nl/system/data_center_publications/files/000/068/061/original/erad2008drup_0120.pdf?1495621011, 2008.

350 Beekhuis, H. and Mathijssen, T.: From pulse to product, Highlights of the upgrade project of the Dutch national weather radar network, in: 10th European Conference on Radar in Meteorology and Hydrology (ERAD 2018) : 1-6 July 2018, Ede-Wageningen, The Netherlands, edited by de Vos, L., Leijnse, H., and Uijlenhoet, R., pp. 960–965, Wageningen University & Research, Wageningen, the Netherlands, https://doi.org/10.18174/454537, 2018.

Bellon, A., Lee, G. W., and Zawadzki, I.: Error statistics of VPR corrections in stratiform precipitation, Journal of Applied Meteorology and
355   Climatology, 44, 998–1015, https://doi.org/10.1175/JAM2253.1, 2005.

Berenguer, M., Sempere-Torres, D., Corral, C., and Sánchez-Diezma, R.: A fuzzy logic technique for identifying nonprecipitating echoes in radar scans, Journal of Atmospheric and Oceanic Technology, 23, 1157–1180, https://doi.org/10.1175/JTECH1914.1, 2006.

Borga, M.: Accuracy of radar rainfall estimates for streamflow simulation, Journal of Hydrology, 267, 26–39, https://doi.org/10.1016/S0022-1694(02)00137-3, 2002.

360 Borga, M., Anagnostou, E. N., and Frank, E.: On the use of real-time radar rainfall estimates for flood prediction in mountainous basins, Journal of Geophysical Research: Atmospheres, 105, 2269–2280, https://doi.org/10.1029/1999JD900270, 2000.

Borga, M., Esposti, S. D., and Norbiato, D.: Influence of errors in radar rainfall estimates on hydrological modeling prediction uncertainty, Water Resources Research, 42, https://doi.org/10.1029/2005WR004559, 2006.

Brauer, C. C., Teuling, A. J., Torfs, P. J. J. F., and Uijlenhoet, R.: The Wageningen Lowland Runoff Simulator (WALRUS):
365   a lumped rainfall–runoff model for catchments with shallow groundwater, Geoscientific Model Development, 7, 2313–2332, https://doi.org/10.5194/gmd-7-2313-2014, 2014a.

Brauer, C. C., Torfs, P. J. J. F., Teuling, A. J., and Uijlenhoet, R.: The Wageningen Lowland Runoff Simulator (WALRUS): application to the Hupsel Brook catchment and the Cabauw polder, Hydrology and Earth System Sciences, 18, 4007–4028, https://doi.org/10.5194/hess-18-4007-2014, 2014b.

370 Brauer, C. C., Overeem, A., Leijnse, H., and Uijlenhoet, R.: The effect of differences between rainfall measurement techniques on groundwater and discharge simulations in a lowland catchment, Hydrological Processes, 30, 3885–3900, https://doi.org/10.1002/hyp.10898, 2016.

Cho, Y.-H., Lee, G., Kim, K.-E., and Zawadzki, I.: Identification and removal of ground echoes and anomalous propagation using the characteristics of radar echoes, Journal of Atmospheric and Oceanic Technology, 23, 1206–1222, https://doi.org/10.1175/JTECH1913.1, 2006.

375 Crameri, F.: Geodynamic diagnostics, scientific visualisation and StagLab 3.0, Geoscientific Model Development, 11, 2541–2562, https://doi.org/10.5194/gmd-11-2541-2018, 2018.

Crameri, F., Shephard, G. E., and Heron, P. J.: The misuse of colour in science communication, Nature Communications, 11, 5444, https://doi.org/10.1038/s41467-020-19160-7, 2020.

Creutin, J. D., Delrieu, G., and Lebel, T.: Rain measurement by raingage-radar combination: A geostatistical approach, Journal of Atmo-
spheric and Oceanic Technology, 5, 102–115, https://doi.org/10.1175/1520-0426(1988)005<0102:RMBRRC>2.0.CO;2, 1988.

Creutin, J. D., Andrieu, H., and Faure, D.: Use of a weather radar for the hydrology of a mountainous area. Part II: radar measurement
validation, Journal of Hydrology, 193, 26–44, https://doi.org/10.1016/S0022-1694(96)03203-9, 1997.

Ebert, E. E., Wilson, L. J., Brown, B. G., Nurmi, P., Brooks, H. E., Bally, J., and Jaeneke, M.: Verification of nowcasts
from the WWRP Sydney 2000 forecast demonstration project, Weather and Forecasting, 19, 73–96, https://doi.org/10.1175/1520-
0434(2004)019<0073:VONFTW>2.0.CO;2, 2004.

Fabry, F., Austin, G. L., and Tees, D.: The accuracy of rainfall estimates by radar as a function of range, Quarterly Journal of the Royal
Meteorological Society, 118, 435–453, https://doi.org/10.1002/qj.49711850503, 1992.

Foresti, L., Reyniers, M., Seed, A., and Delobbe, L.: Development and verification of a real-time stochastic precipitation nowcasting system
for urban hydrology in Belgium, Hydrology and Earth System Sciences, 20, 505–527, https://doi.org/10.5194/hess-20-505-2016, 2016.

Goudenhoofdt, E. and Delobbe, L.: Evaluation of radar-gauge merging methods for quantitative precipitation estimates, Hydrology and Earth
System Sciences, 13, 195–203, https://doi.org/10.5194/hess-13-195-2009, 2009.

Goudenhoofdt, E. and Delobbe, L.: Generation and verification of rainfall estimates from 10-Yr volumetric weather radar measurements,
Journal of Hydrometeorology, 17, 1223–1242, https://doi.org/10.1175/JHM-D-15-0166.1, 2016.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Impli-
cations for improving hydrological modelling, Journal of Hydrology, 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Haase, G., Crewell, S., Simmer, C., and Wergen, W.: Assimilation of radar data in mesoscale models: Physical initialization and latent heat
nudging, Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, 25, 1237–1242, https://doi.org/10.1016/S1464-
1909(00)00186-6, 2000.

Harrison, D. L., Scovell, R. W., and Kitchen, M.: High-resolution precipitation estimates for hydrological uses, Proceedings of the Institution
of Civil Engineers - Water Management, 162, 125–135, https://doi.org/10.1680/wama.2009.162.2.125, 2009.

Hazenberg, P., Torfs, P. J. J. F., Leijnse, H., Delrieu, G., and Uijlenhoet, R.: Identification and uncertainty estimation of vertical reflectivity
profiles using a Lagrangian approach to support quantitative precipitation measurements by weather radar: VPR estimation and uncertainty,
Journal of Geophysical Research: Atmospheres, 118, 10,243–10,261, https://doi.org/10.1002/jgrd.50726, 2013.

Hazenberg, P., Leijnse, H., and Uijlenhoet, R.: The impact of reflectivity correction and accounting for raindrop size distribution variability
to improve precipitation estimation by weather radar for an extreme low-land mesoscale convective system, Journal of Hydrology, 519,
3410–3425, https://doi.org/10.1016/j.jhydrol.2014.09.057, 2014.

Heuvelink, D., Berenguer, M., Brauer, C. C., and Uijlenhoet, R.: Hydrological application of radar rainfall nowcasting in the Netherlands,
Environment International, 136, 105 431, https://doi.org/10.1016/j.envint.2019.105431, 2020.

Holleman, I.: Bias adjustment and long-term verification of radar-based precipitation estimates, Meteorological Applications, 14, 195–203,
https://doi.org/10.1002/met.22, 2007.

Imhoff, R. O., Brauer, C. C., Overeem, A., Weerts, A. H., and Uijlenhoet, R.: Spatial and temporal evaluation of radar rainfall nowcasting
techniques on 1,533 events, Water Resources Research, 56, e2019WR026 723, https://doi.org/10.1029/2019WR026723, 2020a.

Imhoff, R. O., Overeem, A., Brauer, C. C., Leijnse, H., Weerts, A. H., and Uijlenhoet, R.: Rainfall nowcasting using commercial microwave
links, Geophysical Research Letters, 47, e2020GL089 365, https://doi.org/10.1029/2020GL089365, 2020b.

415   Joss, J. and Lee, R.: The application of radar–gauge comparisons to operational precipitation profile corrections, Journal of Applied Meteorology, 34, 2612–2630, https://doi.org/10.1175/1520-0450(1995)034<2612:TAORCT>2.0.CO;2, 1995.

Joss, J. and Pittini, A.: Real-time estimation of the vertical profile of radar reflectivity to improve the measurement of precipitation in an Alpine region, Meteorology and Atmospheric Physics, 47, 61–72, https://doi.org/10.1007/BF01025828, 1991.

Kirstetter, P.-E., Andrieu, H., Delrieu, G., and Boudevillain, B.: Identification of vertical profiles of reflectivity for correc-
420   tion of volumetric radar data using rainfall classification, Journal of Applied Meteorology and Climatology, 49, 2167–2180, https://doi.org/10.1175/2010JAMC2369.1, 2010.

Kitchen, M. and Jackson, P. M.: Weather radar performance at long range - simulated and observed, Journal of Applied Meteorology and Climatology, 32, 975–985, https://doi.org/10.1175/1520-0450(1993)032<0975:WRPALR>2.0.CO;2, 1993.

KNMI: KNMI - Jaar 2008: Twaalfde warme jaar op rij, https://www.knmi.nl/nederland-nu/klimatologie/maand-en-seizoensoverzichten/
425   2008/jaar, 2009.

Krajewski, W. F.: Cokriging radar-rainfall and rain gage data, Journal of Geophysical Research: Atmospheres, 92, 9571–9580, https://doi.org/10.1029/JD092iD08p09571, 1987.

Marshall, J. S., Hitschfeld, W., and Gunn, K. L. S.: Advances in radar weather, in: Advances in Geophysics, edited by Lansberg, H. E., vol. 2, pp. 1–56, Academic Press Inc., New York, NY, 1955.

430   Na, W. and Yoo, C.: A bias correction method for rainfall forecasts using backward storm tracking, Water, 10, 1728, https://doi.org/10.3390/w10121728, 2018.

Ochoa-Rodriguez, S., Rico-Ramirez, M., Jewell, S. A., Schellart, A. N. A., Wang, L., Onof, C., and Maksimović, v.: Improving rainfall nowcasting and urban runoff forecasting through dynamic radar-raingauge rainfall adjustment, in: 7th International Conference on Sewer Processes & Networks, http://spiral.imperial.ac.uk/handle/10044/1/14662, 2013.

435   Ochoa-Rodriguez, S., Wang, L.-P., Willems, P., and Onof, C.: A review of radar-rain gauge data merging methods and their potential for urban hydrological applications, Water Resources Research, 55, 6356–6391, https://doi.org/10.1029/2018WR023332, 2019.

Overeem, A., Buishand, T. A., and Holleman, I.: Extreme rainfall analysis and estimation of depth-duration-frequency curves using weather radar, Water Resources Research, 45, W10 424, https://doi.org/10.1029/2009WR007869, 2009a.

Overeem, A., Holleman, I., and Buishand, A.: Derivation of a 10-year radar-based climatology of rainfall, Journal of Applied Meteorology
440   and Climatology, 48, 1448–1463, https://doi.org/10.1175/2009JAMC1954.1, 2009b.

Overeem, A., Leijnse, H., and Uijlenhoet, R.: Measuring urban rainfall using microwave links from commercial cellular communication networks, Water Resources Research, 47, W12 505, https://doi.org/10.1029/2010WR010350, 2011.

Park, S., Berenguer, M., and Sempere-Torres, D.: Long-term analysis of gauge-adjusted radar rainfall accumulations at European scale, Journal of Hydrology, 573, 768–777, https://doi.org/10.1016/j.jhydrol.2019.03.093, 2019.

445   Prinsen, G., Hakvoort, H., and Dahm, R.: Neerslag-afvoermodellering met SOBEK-RR, Stromingen, 15, 8–24, 2010.

Qi, Y., Zhang, J., Zhang, P., and Cao, Q.: VPR correction of bright band effects in radar QPEs using polarimetric radar observations, Journal of Geophysical Research: Atmospheres, 118, 3627–3633, https://doi.org/10.1002/jgrd.50364, 2013.

Rogers, R. F., Fritsch, J. M., and Lambert, W. C.: A simple technique for using radar data in the dynamic initialization of a mesoscale model, Monthly Weather Review, 128, 2560–2574, https://doi.org/10.1175/1520-0493(2000)128<2560:ASTFUR>2.0.CO;2, 2000.

450   Saltikoff, E., Friedrich, K., Soderholm, J., Lengfeld, K., Nelson, B., Becker, A., Hollmann, R., Urban, B., Heistermann, M., and Tassone, C.: An overview of using weather radar for climatological studies: successes, challenges, and potential, Bulletin of the American Meteorological Society, 100, 1739–1752, https://doi.org/10.1175/BAMS-D-18-0166.1, 2019.

Schleiss, M., Olsson, J., Berg, P., Niemi, T., Kokkonen, T., Thorndahl, S., Nielsen, R., Ellerbæk Nielsen, J., Bozhinova, D., and Pulkkinen, S.: The accuracy of weather radar in heavy rain: a comparative study for Denmark, the Netherlands, Finland and Sweden, Hydrology and Earth System Sciences, 24, 3157–3188, https://doi.org/10.5194/hess-24-3157-2020, 2020.

Schuurmans, J. M., Bierkens, M. F. P., Pebesma, E. J., and Uijlenhoet, R.: Automatic prediction of high-resolution daily rainfall fields for multiple extents: The potential of operational radar, Journal of Hydrometeorology, 8, 1204–1224, https://doi.org/10.1175/2007JHM792.1, 2007.

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.

Sharif, H. O., Ogden, F. L., Krajewski, W. F., and Xue, M.: Numerical simulations of radar rainfall error propagation, Water Resources Research, 38, 15–1–15–14, https://doi.org/10.1029/2001WR000525, 2002.

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

Smith, J. A. and Krajewski, W. F.: Estimation of the mean field bias of radar rainfall estimates, Journal of Applied Meteorology and Climatology, 30, 397–412, https://doi.org/10.1175/1520-0450(1991)030<0397:EOTMFB>2.0.CO;2, 1991.

Stelling, G. S. and Duinmeijer, S. P. A.: A staggered conservative scheme for every Froude number in rapidly varied shallow water flows, International Journal for Numerical Methods in Fluids, 43, 1329–1354, https://doi.org/10.1002/fld.537, 2003.

Stelling, G. S. and Verwey, A.: Numerical flood simulation, in: Encyclopedia of Hydrological Sciences. Part 2: Hydroinformatics, John Wiley & Sons, Ltd, https://doi.org/10.1002/0470848944.hsa025a, 2006.

Sun, Y., Bao, W., Valk, K., Brauer, C. C., Sumihar, J., and Weerts, A. H.: Improving forecast skill of lowland hydrological models using ensemble kalman filter and unscented kalman filter, Water Resources Research, 56, e2020WR027 468, https://doi.org/10.1029/2020WR027468, 2020.

Thorndahl, S., Nielsen, J. E., and Rasmussen, M. R.: Bias adjustment and advection interpolation of long-term high resolution radar rainfall series, Journal of Hydrology, 508, 214–226, https://doi.org/10.1016/j.jhydrol.2013.10.056, 2014.

Thorndahl, S., Einfalt, T., Willems, P., Nielsen, J. E., ten Veldhuis, M.-C., Arnbjerg-Nielsen, K., Rasmussen, M. R., and Molnar, P.: Weather radar rainfall data in urban hydrology, Hydrology and Earth System Sciences, 21, 1359–1380, https://doi.org/10.5194/hess-21-1359-2017, 2017.

Todini, E.: A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements, Hydrology and Earth System Sciences, 5, 187–199, https://doi.org/10.5194/hess-5-187-2001, 2001.

Uijlenhoet, R. and Berne, A.: Stochastic simulation experiment to assess radar rainfall retrieval uncertainties associated with attenuation and its correction, Hydrology and Earth System Sciences, 12, 587–601, https://doi.org/10.5194/hess-12-587-2008, 2008.

Vos, L. W. d., Leijnse, H., Overeem, A., and Uijlenhoet, R.: Quality control for crowdsourced personal weather stations to enable operational rainfall monitoring, Geophysical Research Letters, 46, 8820–8829, https://doi.org/10.1029/2019GL083731, 2019.

Wackernagel, H.: Multivariate geostatistics: An introduction with applications, Springer, Berlin Heidelberg, Germany, 3 edn., https://doi.org/10.1007/978-3-662-05294-5, 2003.

Werner, M., Schellekens, J., Gijsbers, P., van Dijk, M., van den Akker, O., and Heynert, K.: The Delft-FEWS flow forecasting system, Environmental Modelling & Software, 40, 65–77, https://doi.org/10.1016/j.envsoft.2012.07.010, 2013.

490   Wilson, J. W., Feng, Y., Chen, M., and Roberts, R. D.: Nowcasting challenges during the Beijing Olympics: Successes, failures, and impli-
        cations for future nowcasting systems, Weather and Forecasting, 25, 1691–1714, https://doi.org/10.1175/2010WAF2222417.1, 2010.