Rainfall Spatiotemporal Correlation and Intermittency Structure from Micro-gamma to Meso-beta Scale in the Netherlands

van Leth, Thomas C.; Leijnse, Hidde; Overeem, Aart; Uijlenhoet, Remko

DOI
10.1175/JHM-D-20-0311.1

Publication date
2021

Document Version
Final published version

Published in
Journal of Hydrometeorology

SPECIAL
IPC12
COLLECTION

Rainfall Spatiotemporal Correlation and Intermittency Structure from Micro-γ to Meso-β Scale in the Netherlands

THOMAS C. VAN LETH,[a] HIDDE LEIJNSE,[a,b] AART OVEREEM,[a,b] AND REMKO UIJLENHOET[a,c]

[a] *Hydrology and Quantitative Water Management Group, Wageningen University and Research, Wageningen, Netherlands*
[b] *R&D Observations and Data Technology, Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands*
[c] *Department of Water Management, Delft University of Technology, Delft, Netherlands*

(Manuscript received 22 December 2020, in final form 28 May 2021)

ABSTRACT: We investigate the spatiotemporal structure of rainfall at spatial scales from 7 m to over 200 km in the Netherlands. We used data from two networks of laser disdrometers with complementary interstation distances in two Dutch cities (comprising five and six disdrometers, respectively) and a Dutch nationwide network of 31 automatic rain gauges. The smallest aggregation interval for which raindrop size distributions were collected by the disdrometers was 30 s, while the automatic rain gauges provided 10-min rainfall sums. This study aims to supplement other micro-γ investigations (usually performed in the context of spatial rainfall variability within a weather radar pixel) with new data, while characterizing the correlation structure across an extended range of scales. To quantify the spatiotemporal variability, we employ a two-parameter exponential model fitted to the spatial correlograms and characterize the parameters of the model as a function of the temporal aggregation interval. This widely used method allows for a meaningful comparison with seven other studies across contrasting climatic settings all around the world. We also separately analyzed the intermittency of the rainfall observations. We show that a single parameterization, consisting of a two-parameter exponential spatial model as a function of interstation distance combined with a power-law model for decorrelation distance as a function of aggregation interval, can coherently describe rainfall variability (both spatial correlation and intermittency) across a wide range of scales. Limiting the range of scales to those typically found in micro-γ variability studies (including four of the seven studies to which we compare our results) skews the parameterization and reduces its applicability to larger scales.

KEYWORDS: Rainfall; Gauges; Microscale processes/variability

1. Introduction

Quantification of small-scale rainfall variability is important for the design and operation of small-scale sensor networks for flood prediction, particularly in urban areas (ten Veldhuis et al. 2018). It is also important for the assessment of the spatial representativeness of path- or area-averaged remote rainfall measurement methods such as microwave links (Berne and Uijlenhoet 2007; van Leth et al. 2020), weather radar (Jaffrain et al. 2011; Peleg et al. 2013), and satellite remote sensing (Habib and Krajewski 2002; Villarini et al. 2008). Furthermore the spatial and temporal patterns of variability provide insight into the physical structure of rain and can be used for the construction of rain simulations (Schleiss et al. 2009).

Several measurement campaigns have been carried out in the past decades in order to assess the spatiotemporal variability of rain in a statistical sense at scales comparable to a typical weather radar pixel, corresponding to the micro-α to micro-γ scales in the framework of Orlanski (1975). Networks of laser disdrometers have been used before to measure the temporal variability of drop size distributions by Tapiador et al. (2010) and Jaffrain and Berne (2012). Jameson et al. (2015) used a similar disdrometer network with logarithmic spacing in order to investigate clustering in rainfall fields. Habib and Krajewski (2002), Ciach and Krajewski (2006), Villarini et al. (2008), Peleg et al. (2013), Mascaro (2017), and O and Foelsche (2019) presented similar analyses using rain gauges at small spatiotemporal scales combined with a modified exponential spatial correlation model. Meanwhile, van de Beek et al. (2012) used a Dutch nationwide automatic rain gauge network to analyze rainfall variability at the mesoscale at hourly to daily time steps and intergauge distances of 30–300 km. Schuurmans et al. (2007) employed the same dataset in conjunction with a microscale tipping-bucket gauge network and radar for high-resolution rainfall field prediction.

Although quite some empirical data are therefore available to investigate spatial rainfall variability at scales smaller than 1 km with subhourly aggregation intervals, much uncertainty remains regarding the applicability of these measurements for the assessment of the spatial representativeness of rainfall estimates from ground-based or spaceborne remote sensors (Habib and Krajewski 2002; Peleg et al. 2013; Sunilkumar et al. 2016;

Denotes content that is immediately available upon publication as open access.

Corresponding author: Remko Uijlenhoet, r.uijlenhoet@tudelft.nl

DOI: 10.1175/JHM-D-20-0311.1

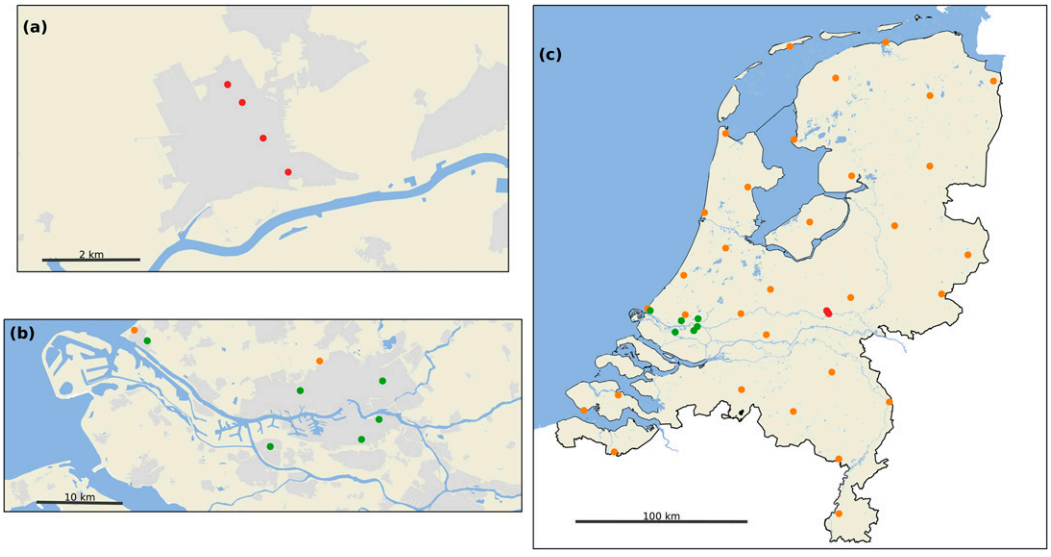

FIG. 1. (a) Relative locations of the five disdrometers in Wageningen (red dots; note that two disdrometers are installed at the northern location). (b) Relative locations of the six disdrometers in Rotterdam (green dots). (c) Relative locations of the 31 automatic rain gauges (orange dots) and 11 disdrometers (red and green dots) in the Netherlands.

Villarini et al. 2008). Different measurement campaigns find contrasting parameterizations and the number of independent measurement campaigns carried out is insufficient to determine whether these differences in parameterization are related to climatology, instrumentation, the different ranges of interstation distances or other causes. Therefore, more independent spatial variability experiments are needed as well as an effort to reconcile results from experiments with different interstation distance ranges. This paper aims to contribute toward filling that gap using disdrometer data with interstation distances ranging from 7 to 2200 m originally collected to validate rainfall measurements using microwave links (van Leth et al. 2018). We supplement these data with a medium-range dataset (with interstation distances ranging from ~3 to ~30 km) collected in the city of Rotterdam as part of an operational hydrometeorological network. Finally, we also use new data from the Dutch nationwide gauge network (with interstation distances of up to 315 km), which is now available at 10-min aggregation intervals. We aim to test the consistency of the findings at different spatiotemporal scales and as part of different measurement campaigns, in particular regarding the spatial correlation structure and intermittency of rainfall.

This paper is organized as follows. The employed rainfall datasets and associated quality control procedures are presented in section 2. In section 3 we present and motivate the employed measures for rainfall spatial correlation and intermittency. Results are presented and analyzed as functions of temporal aggregation interval and season in section 4. In section 5 the presented methodology and associated results are put in a broader perspective by comparing them to methods and results from similar studies around the globe. Finally, the conclusions of our study on small-scale space–time rainfall structure are provided in section 6.

2. Data

a. Datasets

We have used disdrometer data from a field campaign (van Leth et al. 2018) conducted in the Dutch town of Wageningen. This field campaign was set up to provide a test bed for microwave link rainfall retrieval algorithms. The five OTT Parsivel[1] disdrometers were situated roughly on a line of 2.2-km length (Fig. 1a). Two of the disdrometers were situated on top of the same building at only 7 m apart, the other three were placed on other buildings in the Wageningen urban area. The disdrometers provided integrated drop size distributions per 30 s and data from all disdrometers were available from March 2015 until December 2015. From these drop size distributions, rainfall intensities were calculated (see van Leth et al. 2018 for details). These particular disdrometers were used before in Lausanne, Switzerland (Jaffrain et al. 2011), where they had been part of the disdrometer network used by Jaffrain and Berne (2012).

Another dataset is used in order to extend the analysis to larger spatial and temporal scales. This dataset was collected in and around the Dutch city of Rotterdam, around 80 km west of Wageningen, in the period from January 2010 to October 2011. It was collected by the municipality of Rotterdam using six Thies laser disdrometers. These disdrometers are functionally quite similar to the OTT Parsivel[1] disdrometers used in Wageningen, although systematic differences in derived rainfall intensities between both types of disdrometers have been reported, both for the Parsivel[1] (Guyot et al. 2019) and for the Parsivel[2] (Angulo-Martínez et al. 2018). However, because we focus on the spatial correlations and intermittencies rather than on the intensities themselves, we do not expect these differences to have significantly affected the presented analysis. In addition, all employed data were carefully quality-controlled before further

analyses were conducted (section 2b). The locations of these measurement stations are shown in Fig. 1b. These disdrometers provided rainfall intensity estimates over 1-min aggregation intervals. We did not have access to the actual drop size distributions. This dataset is used to provide spatial correlation estimates across interstation distances ranging from ~3 to ~30 km.

To extend the range of spatial scales even further we also make use of data from a Dutch nationwide automatic rain gauge network operated by the Royal Netherlands Meteorological Institute (KNMI). This dataset provides rainfall depths aggregated to 10 min at 31 locations throughout the Netherlands (Fig. 1c), for the period from January 2010 to October 2011 (i.e., the same period as the Rotterdam disdrometer dataset). In contrast to the tipping-bucket gauges used in other studies, these gauges measure the rainfall depth using the displacement of a float placed in a reservoir and can therefore record continuously. The range of interstation distances d overlaps somewhat with the Rotterdam dataset (20 km $< d <$ 315 km). The network of automatic rain gauges is the same as the one used by van de Beek et al. (2012). However, they used a different measurement period (1979–2009) and only had access to hourly rainfall sums.

Although the employed datasets are relatively limited in duration, they are not too short for the purpose of this investigation. The (maritime) rainfall climatology of the Netherlands is such that the yearly rainfall total is never dominated by one or a few major storm events.

b. Quality control

The Wageningen disdrometer dataset included an indication of precipitation type. We used the product of this built-in Parsivel[1] algorithm to exclude anything other than liquid precipitation. Note that only 7% of the total precipitation in the Netherlands falls as solid or mixed precipitation (de Haij 2007). We also set any interval where the Parsivel[1] algorithm indicates dry weather to 0 mm, as this is not always the case for the Parsivel[1].

The Rotterdam dataset did require some quality control before it could be compared reliably to the Wageningen dataset. Because little metadata was available from the disdrometers, we used ~1 × 1 km^2 gauge-adjusted radar pixels containing the disdrometer locations as a reference to eliminate suspicious measurements, including any issue with solid or mixed precipitation. This weather radar dataset was obtained using the two C-band radars operated by KNMI and adjusted using both the automatic rain gauge network and a network of 325 manual gauges (Overeem et al. 2009a,b, 2011). The temporal resolution is 5 min.

Of course, weather radar and ground-based measurements cannot be compared directly even when the magnitude is adjusted. This is because 1) the situation at 1.5 km above the ground (the effective height of the constant-altitude plan position indicator) can be different than at ground level, 2) the radar pixel represents a volumetric average as opposed to a point measurement, and 3) the pixels do not align perfectly with the positions of the disdrometers. Hence, directly comparing the two measurements would yield many false positives for exclusion. We recognize several aspects of the radar data that could lead to intervals being erroneously categorized as faulty disdrometer measurements: intermittency of precipitation within an event, a difference in event timing between radar and disdrometer, and small-scale spatial variability in magnitude which is not captured by radar.

Since spatiotemporal variability at the small scale is exactly what we are investigating here, it is of the utmost importance that we do not unduly influence our results by erroneously excluding events that simply exhibit a high variability. To do so we have devised the following filtering algorithm. We first apply a 3-h moving average filter to both the disdrometer time series (after aggregating to 5 min) and the radar time series. This smooths out small-scale intermittency. Next, we compare the smoothed radar time series and the smoothed disdrometer time series and evaluate the following equation:

$$(R_{\text{disdro,3h}}/R_{\text{radar,3h}} > a) \wedge (R_{\text{disdro,3h}} - R_{\text{radar,3h}} > b), \quad (1)$$

where the first term represents the multiplicative error, the second term represents the additive error, and $\wedge$ indicates the Boolean AND operator. If the above equation evaluates as true for a given 5-min interval then the interval is excluded. After manually verifying the effect of this filter for different values of a and b, we finally selected $a = 2$ and $b = 1$ mm h^{-1}. Although a filter like this admittedly has a certain degree of arbitrariness (Park et al. 2019), the selected values of the filter parameters ensure that intervals are only excluded when the 3-h disdrometer moving averages substantially overestimate the gauge-adjusted radar reference. In addition, applying the filter to 3-h moving averages avoids erroneously excluding intervals [using Eq. (1)] displaying actual small-scale spatial rainfall variability. In this manner, instrumental outliers are removed while actual small-scale rainfall fluctuations are preserved.

3. Methodology

Following Ciach and Krajewski (2006), Villarini et al. (2008), Tokay and Öztürk (2012), Peleg et al. (2013), and Mascaro (2017), we use Pearson's product-moment correlation coefficient ρ_{ij} as an estimator for the interstation correlation:

$$\rho_{ij} = \frac{\overline{R_{i,t}R_{j,t}} - \overline{R_{i,t}}\,\overline{R_{j,t}}}{\sqrt{(\overline{R_{i,t}^2} - \overline{R_{i,t}}^2)(\overline{R_{j,t}^2} - \overline{R_{j,t}}^2)}}, \quad (2)$$

where $R_{i,t}$ and $R_{j,t}$ are the rainfall intensity time series of the two stations (with indices i and j) to be compared and the means (indicated by the overbars) are taken over the entire time series (zeroes included). This is then applied for every combination of two stations in a given dataset.

For random variables with strongly skewed probability distributions, such as rainfall intensity (in particular at short temporal aggregation intervals), Pearson's estimator is known to produce biased and uncertain correlation estimates (see Habib et al. 2001, and references therein). Using a simulation experiment based on a bivariate intermittent lognormal distribution Habib et al. (2001) showed that, in case of strong positive skewness, Pearson's estimator may lead to overestimates of the true correlation (they reported positive biases of up to 0.15 for the smallest sample sizes), whereas a logarithmic transformation of the data would produce essentially unbiased

estimates. However, following the reasoning of Ciach and Krajewski (2006) for rainfall in central Oklahoma and Villarini et al. (2008) for southwest England, we expect the probability density function of rainfall intensity in the Netherlands to have a much lighter tail than the mixed lognormal distribution and therefore decided not to apply the procedure proposed by Habib et al. (2001). This has the added advantage of rendering the results of our analysis directly comparable to the mentioned previous work on small-scale rainfall variability.

We use a bootstrap procedure to estimate the sampling uncertainty in the spatial correlation coefficient for all interstation distances and every temporal aggregation interval considered. Here, we randomly take N intervals (with replacement) from the N-length time series and recalculate the correlation coefficient, where N indicates the total number of samples that is available for a given dataset. This implies that for coarser temporal aggregation intervals, N decreases, hence the sampling uncertainty in the spatial correlation coefficient increases, which is exactly what the bootstrap procedure aims to quantify. We perform this procedure 100 times for every combination of stations and calculate the standard deviation from the resulting distributions. This we take as a measure of confidence in the correlation coefficients, where a small standard deviation indicates a high confidence and vice versa.

We parameterize the rainfall decorrelation as a function of interstation distance with the following two-parameter exponential spatial correlation model:

$$\rho(d) = \exp\left[-\left(\frac{d}{d_0}\right)^{s_0}\right], \tag{3}$$

where d is the interstation distance, d_0 is the decorrelation distance (the distance where the correlation reduces to $1/e$) and s_0 is the shape parameter (with $s_0 = 1$ corresponding to the classical single-parameter exponential model and $s_0 = 2$ to a Gaussian shape). This is similar to the three-parameter exponential model suggested by Krajewski et al. (2003) and later applied by Ciach and Krajewski (2006), Villarini et al. (2008), Jaffrain and Berne (2012), Tokay and Öztürk (2012), Peleg et al. (2013), Sunilkumar et al. (2016), and O and Foelsche (2019), except that we omit the nugget parameter. This assumption is consistent with previous research on spatial rainfall variability in the Netherlands (van de Beek et al. 2011, 2012) and elsewhere (Mandapaka and Qin 2013; Mascaro 2017). A further comparison with nugget parameters found for other locations around the world is provided in section 5. We take the decorrelation distance to be a reliable measure of the spatial rainfall variability for a given temporal accumulation interval. In doing so we therefore assume that the rainfall variability is isotropic (i.e., not dependent on direction), an assumption that was tested experimentally by Jaffrain and Berne (2012), Mandapaka and Qin (2013), Mascaro (2017), and O and Foelsche (2019) and implicit in all other studies using Eq. (3) as spatial rainfall correlation model and other studies dealing with spatial rainfall variability in the Netherlands (Schuurmans et al. 2007; van de Beek et al. 2011, 2012). As a consequence, the orientation of the Wageningen disdrometer array (Fig. 1) is not expected to significantly affect our results. Finally, we do not expect any long-range dependencies to be present in our relatively small-scale dataset, rendering Eq. (3) a suitable correlation model.

To fit the optimal parameters d_0 and s_0 we use a standard Levenberg–Marquardt nonlinear least squares implementation, where we have used the uncertainties calculated with the bootstrap procedure as weights in the cost function. The confidence bands for the nonlinear regression curves are determined using a Monte Carlo technique. Here, the numerically estimated parameter covariances are used to generate 100 parameter instances of a multivariate Gaussian distribution. Using Eq. (3), we obtain 100 instances of the correlation coefficient at each distance. Next, we determine the 5th and 95th percentile of the correlation coefficient for each distance. We perform the abovementioned procedure for several temporal aggregation periods ranging from 30 s (depending on the available data) up to 48 h.

We also investigate the degree of spatial intermittency of rainfall. That is, regardless of the amount of rainfall, we aim to quantify the spatial variability in dry and rainy subintervals within a rainy period. As a measure of spatial intermittency we calculate the following quantity:

$$I_{ij} = \frac{N_{\text{all\_wet}}}{N_{\text{any\_wet}}} = \frac{N_{(R_{i,t}>0 \wedge R_{j,t}>0)}}{N_{(R_{i,t}>0 \vee R_{j,t}>0)}}. \tag{4}$$

Here, $N_{\text{all\_wet}}$ is the number of intervals where both stations register rain throughout the aggregation interval, while $N_{\text{any\_wet}}$ is the number of intervals where at least one of the stations registers rain at least once during the aggregation interval. By registering rain, we mean here that the rainfall intensity at the original (smallest) aggregation level is nonzero. We will refer to I as the "intermittency ratio." This is equivalent to the conditional probability of detection, defined by Krajewski et al. (2003) as the probability of rainfall at one point, given that it rains at another point during the same interval. Hence, the intermittency ratio I defined in this manner is inversely related to the intermittency of rainfall. Given the spatial coherence of rain, I is expected to be a decreasing function, from 1 at small interstation distances to 0 at large interstation distances, in much the same way as the rainfall interstation correlation ρ [Eq. (2)]. This will allow us to investigate if the two-parameter exponential model adopted for spatial correlation [Eq. (3)] is also an appropriate model for spatial intermittency of rainfall. Another measure of rainfall intermittency was employed by Habib and Krajewski (2002) and Villarini et al. (2008), who considered the probability of a rain gauge registering zero rainfall within an area where nonzero rainfall was measured. Schleiss et al. (2011) considered the probability that a randomly chosen area of a given size remains completely dry for at least so many hours. However, in the context of the current study, we prefer the intermittency measure provided by Eq. (4), as it is defined for each disdrometer/gauge combination, much like the interstation correlation [Eq. (2)].

4. Results

a. Spatial correlation

The correlations between the rainfall intensities from two stations are plotted against the distances between them. This is

done for both the Wageningen and the Rotterdam disdrometer datasets as well as the nationwide gauge dataset. The decorrelation distances are determined using Eq. (3) for different resampled temporal aggregation intervals. The results are shown in Fig. 2. The optimal values found for the inferred parameters are given in the figure together with the standard deviations estimated with the regression procedure described in section 3.

It can be seen that rainfall estimates from the two closest stations (7 m apart) are already highly correlated at a 30-s aggregation interval ($\rho = 0.966 \pm 0.002$). At a 1-min aggregation interval and a distance of roughly 500 m (483 m and 490 m) the correlation coefficients are 0.860 ± 0.008 and 0.875 ± 0.008, respectively. This is clearly higher than what Villarini et al. (2008) found for their smallest interstation distance bin of roughly 500 m at a 1-min aggregation interval. They found correlation coefficients ranging from 0.21 to 0.38 for the Brue catchment (United Kingdom) using tipping-bucket rain gauges. Tapiador et al. (2010) only analyzed a few single days, but for this aggregation period and interstation distance found similarly high correlation coefficients (although for radar reflectivity instead of rainfall intensity), ranging between 0.90 and 0.97 depending on the date. They used instruments of the same type as used in this experiment. Including an additional nugget parameter in the correlation model does not improve the goodness-of-fit for either of the two datasets. The coefficient of determination (R^2), indicating the fraction of the observed variance explained by the regression model, is found to be 0.99 at the 30-s aggregation level for the Wageningen dataset and 0.95 for the 1-min Rotterdam dataset, with no significant difference using either model (with or without nugget).

The fitted curves based on the Wageningen dataset and the Rotterdam dataset are quite dissimilar and gradually become more so as the aggregation period increases. However, it is also clear that a direct comparison between the two is impossible given their respective interstation distance ranges do not overlap. Especially with regard to the Wageningen dataset it should be noted that the optimally fitted decorrelation distances (d_0 increases from 3.50 km at an aggregation interval of 30 s to 8.93 km at 1 h; Figs. 2a–f) are in all cases outside of the range of actually measured distances (d ranges from 7 m to 2.2 km; Fig. 1a). Hence, d is always smaller than d_0, indicating that the range of interstation distances in the Wageningen dataset is insufficient to determine the actual correlation patterns for this range of temporal aggregation periods. Actually, if $s_0 = 1$, which is very nearly the case for an aggregation interval of 1 min (Fig. 2b), the two-parameter exponential spatial correlation model [Eq. (3)] reduces to the classical single-parameter exponential model $\rho(d) = e^{-(d/d_0)}$ (Ha et al. 2007), which for small values of d approaches the straight line $\rho(d) = 1 - (d/d_0)$ with (negative) slope $-1/d_0$. This indicates that, for the Wageningen dataset at aggregation intervals around 1 min, d_0 actually represents the inverse of the slope of a linear correlation model rather than the decorrelation distance. We do see that the prediction bands overlap where the respective ranges of the Wageningen and Rotterdam datasets almost meet. For aggregation periods longer than 1 h the confidence band for the Wageningen dataset becomes so wide as to render the fit essentially meaningless. Therefore, no fitted models for the Wageningen dataset are shown in Figs. 2g–2i.

Figures 2d–2i also show the correlograms for the nationwide dataset. Shape factors are lower than for the Rotterdam dataset at all aggregation levels. Decorrelation distances are lower than for the Rotterdam dataset at aggregation intervals of 30 min and below and decorrelation distances are higher at higher aggregation levels.

Because the distance ranges covered by the Wageningen and Rotterdam datasets are complementary, we increase the representativeness of the fit by combining the two in a single regression. In doing so, we implicitly assume both datasets to be sampled from the same population, i.e., that there is no climatological difference between the sampled areas and times and that there are no instrumental differences that could in any way influence the correlations between stations. Although the two datasets are collected in two different time frames, at different locations and with different instruments, the assumption is reasonably justified: the timespans of the datasets are relatively long compared to the aggregation intervals (9 months and 1.5 years, respectively), they take place within 4 years of each other and both are collected over similar (flat, urban) terrain within 100 km of each other within a larger area which has no significant height differences or mountain ranges. Furthermore, both types of instruments operate along the same general principles: both are optical laser disdrometers (Angulo-Martínez et al. 2018; Guyot et al. 2019).

With that justification, we determine the correlation model parameters using correlations found in both datasets together and compare them to the parameters of the regression curve based on only the Rotterdam dataset. Both the shape parameter s_0 and decorrelation distance d_0 become higher for the 1-min aggregation period. For aggregation periods of 5 min and above the difference in d_0 is not significant at a 95% significance level. For 20-min aggregation intervals and longer (not shown here, tested up to 1 h) the confidence bands overlap over the entire range. Therefore, we can conclude that the rainfall correlation pattern observed from just the Rotterdam network is fundamentally in agreement with the pattern observed from both networks. To assess the representativeness of the functions fitted to the combined dataset for both long and short distances, we calculate the coefficient of determination of the function fitted to both datasets applied to the Wageningen dataset or the Rotterdam dataset alone. For 10-min aggregation intervals this yields $R^2 = 0.92$ for Rotterdam and $R^2 = 0.72$ for Wageningen (not shown in Fig. 2). We conclude that the function fitted to both datasets accurately describes the correlation pattern at both smaller ($d < 2.2$ km) and intermediate ($3 < d < 30$ km) interstation distances.

Similarly, we also fitted the correlation model using all three datasets together. The same justification applies, except that the nationwide dataset is collected with automatic gauges instead of laser disdrometers. The results are illustrated in Figs. 2d–2i. The decorrelation distances for aggregation intervals of 10 and 30 min are closer to those determined using only the Rotterdam dataset, while for aggregation intervals of 1 h and longer they are not significantly different from those determined using only the nationwide dataset. This was to be

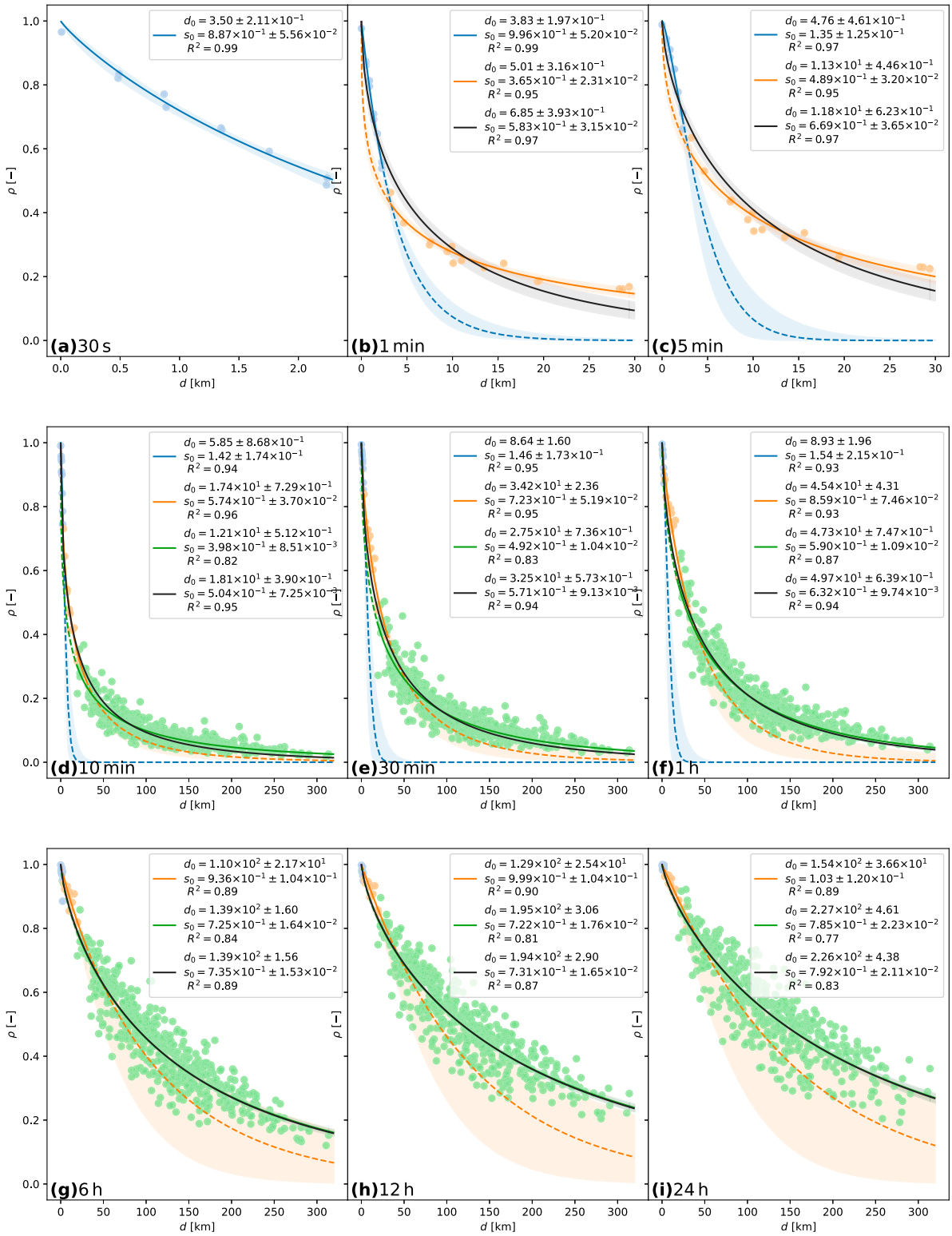

FIG. 2. Spatial correlograms for different temporal aggregation intervals. Blue dots indicate disdrometer pairs from the Wageningen dataset, while orange dots indicate disdrometer pairs from the Rotterdam dataset and green dots indicate gauge pairs from the nationwide dataset. The correspondingly colored lines indicate fitted two-parameter exponential models [Eq. (3)] using only one of the datasets. Fitted models using all datasets are shown in black [in (b) and (c) the nationwide dataset is not included]. Confidence intervals are indicated by lighter colored bands, and extrapolations of fitted models beyond the range of measurements on which they are based are indicated with dashed lines. Aggregation intervals are from left to right and from top to bottom: 30 s, 1 min, 5 min, 10 min, 30 min, 60 min, 6 h, 12 h, and 24 h. Legends show estimated values of model parameters and their confidence intervals, as well as corresponding coefficients of determination.

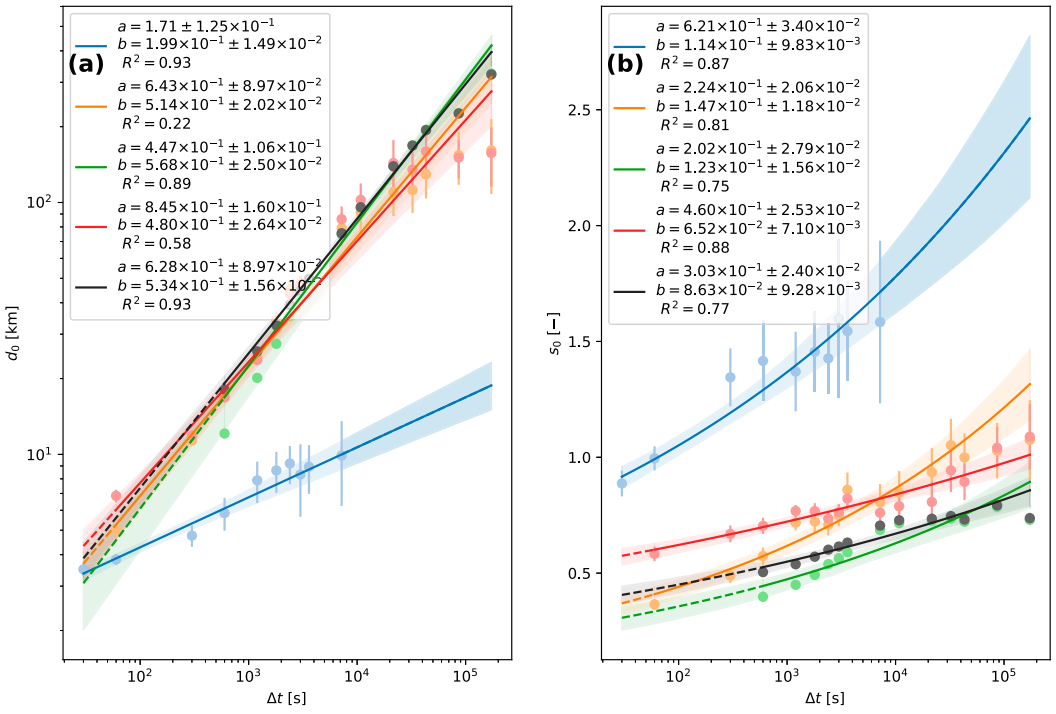

FIG. 3. (a) Decorrelation distance (and associated uncertainty) vs aggregation interval using the Wageningen dataset in blue, the Rotterdam dataset in orange and the nationwide dataset in green. Also included is the decorrelation distance fitted using both the Rotterdam and the Wageningen datasets in red and the decorrelation distance fitted using all datasets together in black. (b) As in (a), but showing the shape factor as a function of aggregation interval. Extrapolations of fitted models beyond the range of measurements on which they are based are indicated with dashed lines. Legends show estimated values of model parameters and their confidence intervals, as well as corresponding coefficients of determination.

expected, as for the longer aggregation intervals (i.e., the larger spatial scales) the nationwide dataset dominates the correlation structure, whereas for the intermediate aggregation intervals (i.e., the intermediate spatial scales) the Rotterdam dataset plays a dominant role. As before, we assess the representativeness of the parameterization at different distance ranges by calculating coefficients of determination of the fitted function applied to each individual dataset at the 10-min aggregation level. This gives a reasonable goodness-of-fit for large ($R^2 = 0.77$ nationwide) and intermediate distances ($R^2 = 0.94$ for Rotterdam), but no fit at all for small distances ($R^2 = -0.94$ for Wageningen). We conclude that there is not a single parameterization of the two-parameter exponential model that applies to the entire distance range.

Figure 3 shows the fitted model parameters for a larger number of aggregation levels. The relationship between the decorrelation distance and the aggregation interval takes the form of a power law,

$$d_0 = a\,\Delta t^b, \tag{5}$$

where Δt is the aggregation period and a and b are fitted parameters (Berne et al. 2004). The parameters are given in Fig. 3a. The decorrelation distance–aggregation interval relationships obtained using most combinations of datasets are consistent, with confidence bands overlapping over most of the range. The relationship obtained using only the Rotterdam dataset results in the lowest goodness-of-fit ($R^2 = 0.22$), while the relationship obtained using all three datasets together results in the highest coefficient of determination ($R^2 = 0.93$). The inferred exponents (b) for these datasets are all close to the value of 0.5 reported by Berne et al. (2004) for small-scale rainfall variability in the city of Marseille, southern France. The exception is the relationship obtained using only the Wageningen dataset. The decorrelation distances at the 30-s level are consistent with the extrapolation obtained from the other fits, but at higher aggregation levels the decorrelation distances are all significantly lower than would have been obtained using any other combination of datasets, up to an order of magnitude difference. The cause of this is that the range of interstation distances in the Wageningen dataset is too small to determine the actual correlation patterns for this range of temporal aggregation periods, as noted before.

A similar power law is also used to fit the shape factor as a function of the aggregation interval and is given in Fig. 3b. The shape factor–aggregation interval relationship is not as consistent between the different combinations of datasets. The difference between the fit obtained using all datasets and that obtained using only the nationwide dataset is small but still significant. In all cases and at all aggregation levels the shape

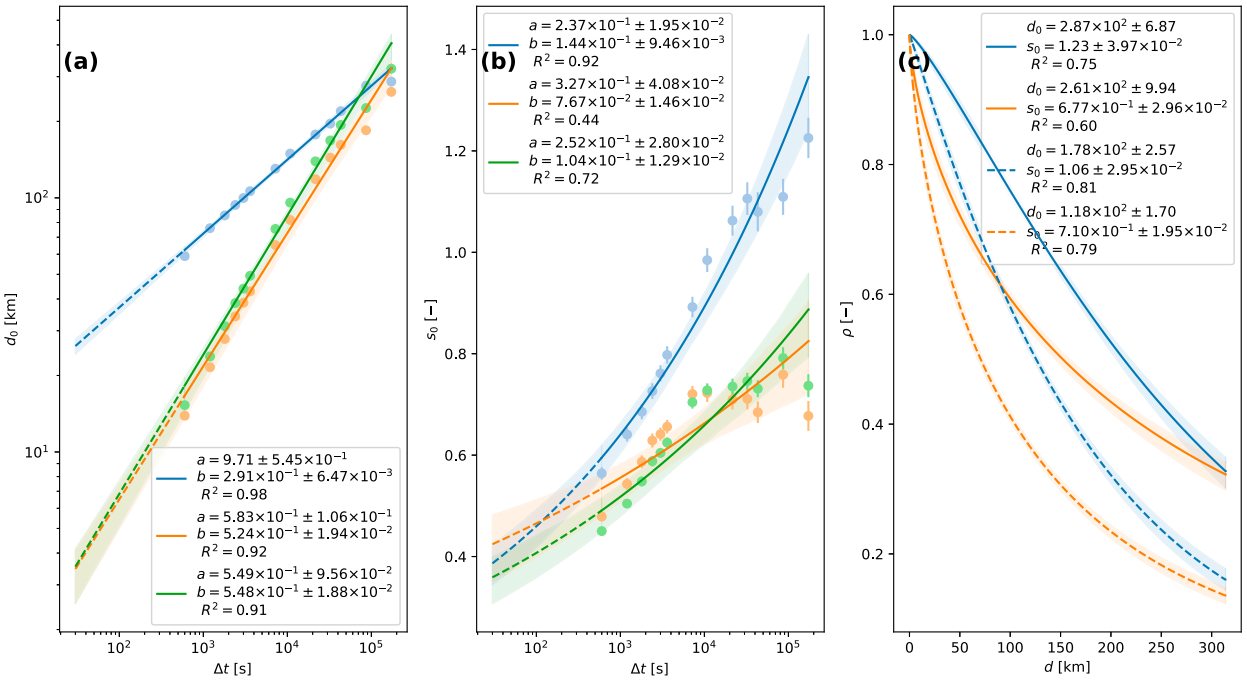

FIG. 4. (a) Decorrelation distance of the disdrometer/gauge pairs (and associated uncertainty) vs aggregation interval per season. Summer months are indicated in orange, while winter months are indicated in blue. All-year values are indicated in green. (b) Shape factors of the correlation function as a function of aggregation interval per season. Colors are the same as in (a). In (a) and (b) extrapolations of fitted models beyond the range of measurements on which they are based are indicated with dashed lines. (c) Correlation coefficient as a function of interstation distance. Solid lines are based on rainfall aggregated to 48 h, while dashed lines are based on 6-h aggregated rainfall. Colors as in (a). Legends show estimated values of model parameters and their confidence intervals, as well as corresponding coefficients of determination.

factor remains below a value of 1. Recall that for $s_0 = 1$ the two-parameter exponential spatial correlation model [Eq. (3)] reduces to a simple exponential model. The exceptions are the shape factors obtained using only the Wageningen dataset, which are all close to or above 1. In summary: using only the subpixel dataset results in anomalous correlation model parameters, which are inconsistent with the findings from a larger range of scales (both including and excluding the subpixel range itself). This could be either due to the limited number of datapoints or the limited range of scales.

We also investigated the effect of using a moving average instead of a sequential averaging scheme for aggregating to longer time intervals. This did not make a significant difference in the results. In the analyses of the rest of this paper we will only consider the results obtained using all datasets to fit the model.

b. *Seasonality of spatial correlation*

To investigate seasonal effects in a crude fashion we have performed the abovementioned analyses on two subsets of our data, one subset spanning only the months of December, January and February (winter) and another subset spanning only the months of June, July and August (summer). Because the Wageningen dataset spans only nine months, we performed the analyses only on the Rotterdam dataset and the nationwide dataset. The results for the summer months are very similar to those found for the entire year, while the results for the winter months are very different. Our results show that typical decorrelation distances are far lower in summer than in winter, as can be seen in Fig. 4a, confirming earlier findings for Dutch rainfall at coarser spatiotemporal scales (van de Beek et al. 2012). This is especially true of the shorter aggregation intervals, while the differences become gradually less toward the longer aggregation intervals. At an aggregation interval of 48 h there is no significant difference anymore and at even longer aggregation intervals (not shown here) inferred summer decorrelation distances become larger than winter decorrelation distances. However, it should be noted that at these aggregation levels the estimated d_0 becomes longer than the longest interstation distance $d_{\max}$ (315 km). Therefore, inferred values are highly uncertain. The inferred exponents of the power-law decorrelation distance–aggregation interval relationships are of the order of 0.5 for the summer months and 0.3 for the winter months (Fig. 4a). Note that van de Beek et al. (2012) found exponents of 0.29 for both summer and winter in the Netherlands at coarser spatiotemporal scales.

There is a striking difference in the values of s_0 as well as in their relation with aggregation interval for the winter months as compared to those found for the summer months (Fig. 4b). In the summer (and over the entire year) the shape factors all stay below approximately 0.8 and they generally do not increase anymore from aggregation intervals of 2 h onward. As

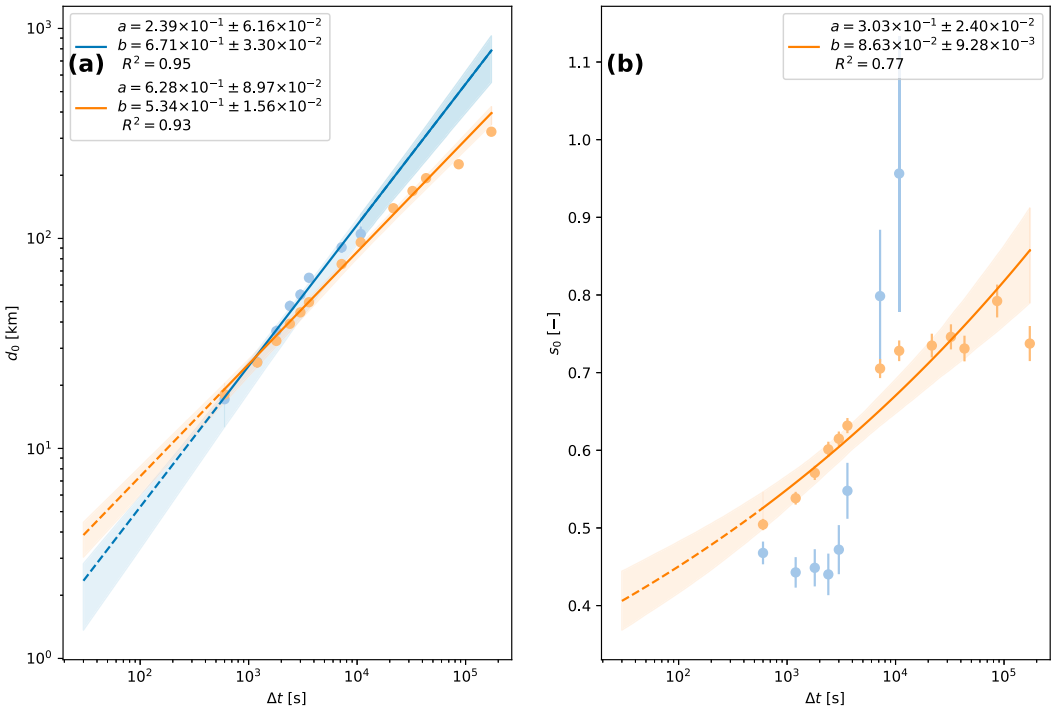

FIG. 5. (a) Dependence of decorrelation distance (and associated uncertainty) on aggregation interval using only wet intervals. Blue lines and dots indicate estimates using only purely wet intervals (per pair of stations) for all datasets, while orange lines and dots include partially wet intervals as well. (b) As in (a), but for shape factor as function of aggregation interval. Extrapolations of fitted models beyond the range of measurements on which they are based are indicated with dashed lines. Legends show estimated values of model parameters and their confidence intervals, as well as corresponding coefficients of determination.

such, the relationship in summer deviates from the assumed power-law relationship, which is also evident from the goodness-of-fit ($R^2 = 0.44$). For the winter months, the shape factor is not only larger than for the summer months, it also increases faster and does not seem to approach an asymptote. Instead, it more closely follows the power-law relationship ($R^2 = 0.92$), and it becomes larger than unity for aggregation intervals of 2 h and longer. In this context, the observation that the decorrelation distances between summer and winter appear to approach each other as the aggregation interval increases (Fig. 4a) can be a bit misleading; for aggregation intervals up to 48 h, the correlation during winter is significantly higher than during summer for the entire distance range. This is illustrated in Fig. 4c for aggregation intervals of 6 and 48 h.

c. *Intermittency*

The datasets contain many periods where rain was registered at one or more stations but some other stations were dry at the same time. In addition, there were periods where rain was registered at only certain subintervals within a longer temporal aggregation interval. To assess the relative effect of such intermittency in rainfall versus the purely quantitative variability in rainfall intensity we have performed the same analysis as described above, but with an extra filtering step. The goal is to exclude any period containing such (spatially and/or temporally intermittent) edge cases. To do so we filter out any time interval per pair of stations where at least one of the two stations for which the correlation coefficient is calculated has reported a rainfall intensity smaller than 0.1 mm h^{-1} during any subinterval. Filtering in this manner leaves 5.0% of the 1-min pair intervals in the Wageningen dataset. Pair intervals where both disdrometers register rainfall intensities smaller than 0.1 mm h^{-1} during all subintervals comprise 92.5% of the dataset. This leaves 2.5% of the pair intervals where at least one of the disdrometers registers rain during at least one subinterval but at least one disdrometer does not do so during at least one subinterval. For the Rotterdam dataset these percentages are 5.0%, 89.5%, and 5.5%, respectively.

The decorrelation distances and shape factors calculated from the filtered data are shown in Fig. 5 for several aggregation intervals together with the same parameters calculated from the unfiltered dataset. Parameters derived from the filtered dataset are not included for aggregation intervals of 6 h and longer; the number of individual intervals that are completely wet are too few to derive meaningful statistics from. When filtering the datasets in this manner it can be observed that the decorrelation distances are significantly longer as compared to the unfiltered dataset at aggregation intervals of 30 min and longer. However, the decorrelation distance–aggregation interval power laws have overlapping significance bands over almost the entire range, only diverging for

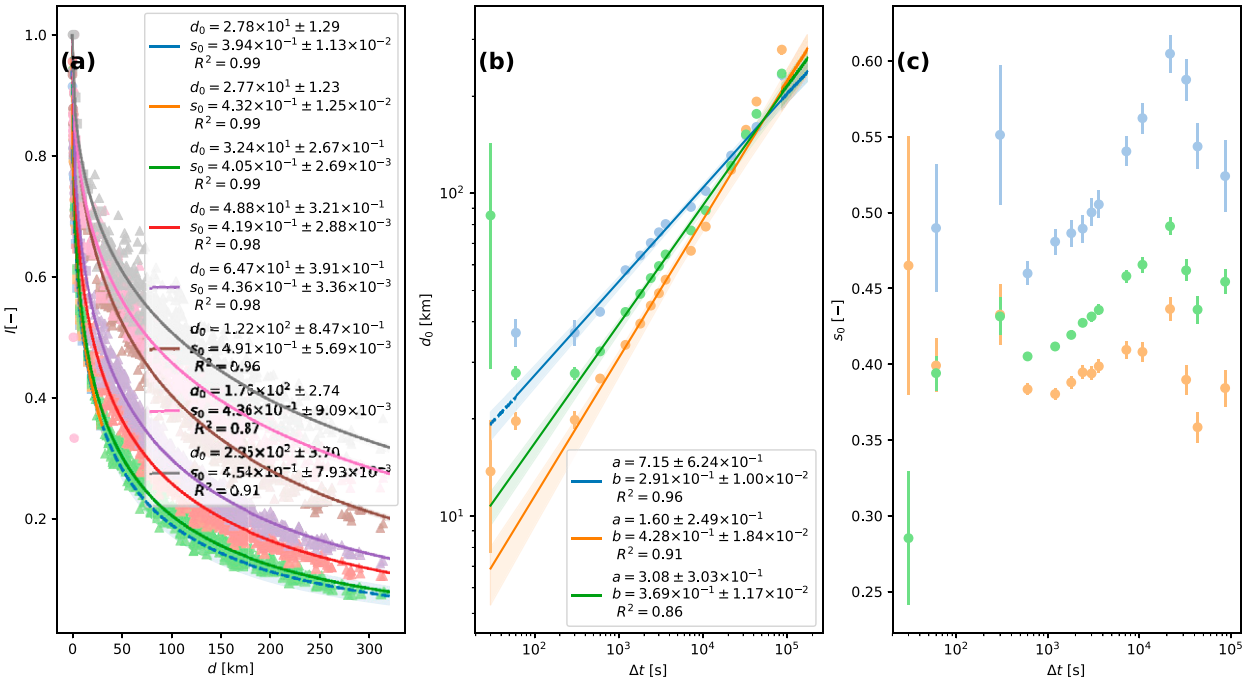

FIG. 6. (a) Intermittency ratios of disdrometer/gauge combinations as a function of interstation distance for several temporal aggregation intervals. Circles represent Wageningen disdrometer pairs, while squares represent Rotterdam disdrometer pairs, and triangles represent nationwide gauge pairs. A two-parameter exponential function [Eq. (3)] is fitted through all datasets. (b) The *e*-folding distance (and associated uncertainty) of intermittency ratios of disdrometer pairs as a function of aggregation interval. Parameters determined using all data are plotted in green. Parameters determined for only summer months are indicated in orange, while blue shows parameters determined for only winter months. (c) Shape factor (and associated uncertainty) of the intermittency ratio–interstation distance relationship as a function of aggregation interval. Colors are the same as in (b). Extrapolations of fitted models beyond the range of measurements on which they are based are indicated with dashed lines. Legends show estimated values of model parameters and their confidence intervals, as well as corresponding coefficients of determination.

intervals of 2 h and longer. In contrast, the differences in terms of the shape factors are more fundamental. The shape factor–aggregation interval relationship in the filtered case is not at all similar to the nonfiltered case and cannot be described with a power-law function.

To gain more insight into the spatial and temporal dependency of the intermittency we use the intermittency ratio I introduced in Eq. (4). The results are plotted against the distance between the stations in Fig. 6a. Following Krajewski et al. (2003), a two-parameter exponential model similar to that used to fit the correlation function [Eq. (3)] is also used to fit the distance dependency of the intermittency ratio. The same analysis is performed for different aggregation intervals. The resulting *e*-folding distances and shape factors are shown in Figs. 6b–6c. A power-law relationship between the *e*-folding distance and the aggregation interval is used. No power-law relationship was fitted between the shape factor and the aggregation interval, since it does not provide a good fit.

We see, e.g., at the 10-min scale, that the conditional probability that two stations at a relative distance d both detect rainfall, given that one of them has detected rainfall, is 1 for $d =$ 0 km, decreasing to 0.1 for $d = 300$ km. Krajewski et al. (2003) found intermittency ratios significantly smaller than 1 for different locations around the world at 5-min, 15-min, and 1-h aggregation intervals. This may have been related to local random errors (associated with the finite bucket volumes) affecting the rainfall estimates from the employed tipping-bucket rain gauges (Ciach 2003), as opposed to the optical disdrometers and float-based gauges we used (section 2a). The exponential decrease model fits the observed pattern very well ($R^2 = 0.99$). The *e*-folding distance (analogous to the decorrelation distance) is of the same order of magnitude as the decorrelation distances observed for the same aggregation interval and data selection, but still significantly longer. At longer aggregation intervals, the intermittency ratios [defined according to Eq. (4)] become gradually higher, implying that the magnitude of the intermittency decreases. For instance, at a 24-h aggregation interval the intermittency ratio at $d = 300$ km becomes 0.35. The *e*-folding distance at longer aggregation intervals becomes lower than the decorrelation distance at the same intervals, but remains still of the same order of magnitude. For all aggregation intervals, the shape factor is lower than 0.5. The power-law function fits the *e*-folding distance–aggregation interval relationship quite well ($R^2 = 0.86$). Note that the first few aggregation intervals shown in Figs. 6b and 6c provide parameters derived only from the Rotterdam and Wageningen datasets and are therefore not included in the power-law fit. The shape factor–aggregation interval relationship

TABLE 1. Estimates for the decorrelation distance d_0 (km) and shape factor s_0 of the two-parameter exponential spatial correlation model [Eq. (3)] for rainfall intensity at 1-min, 15-min, 30-min, and 3-h aggregation intervals from several studies (Ciach and Krajewski 2006; Villarini et al. 2008; Jaffrain and Berne 2012; Tokay and Öztürk 2012; Peleg et al. 2013; Mascaro 2017; O and Foelsche 2019, abbreviated as CK2006, V2008, JB2012, TO2012, P2013, M2017, and OF2019, respectively) as well as this study. The results for JB2012 are listed from left to right: convective, transitional, and frontal rain. The results for M2017 are for summer and winter, respectively. The corresponding numbers for 15 min were obtained by linearly interpolating between the reported values for 10 and 30 min. The results from this study are as follows: Wageningen and Rotterdam datasets (Wag + Rot); Wageningen, Rotterdam, and nationwide datasets combined (All); and only the Wageningen dataset (Wag). For comparison, the minimum d_{min} (m) and maximum d_{max} (km) interstation distances for each study are given.

Study	CK2006	V2008	JB2012	TO2012	P2013	M2017	OF2019	Wag + Rot	All	Wag
Country/state	Oklahoma	England	Switzerland	Virginia	Israel	Arizona	Austria	Netherlands		
d_{min}	2	500	85	1	400	500	700	7	7	7
d_{max}	4	15	0.8	2.672	5	227	23.4	30	315	2.2
d_0, 1 min	7.5	6.3	1.9/4.9/4.1	—	3	—	—	6.9	—	3.8
d_0, 15 min	10.5	20.0	2.9/5.8/16.7	43.1	7	6.8/8.1	~20	21.5	21.0	5.3
d_0, 30 min	12.0	26.0	—	70	9	10.7/23.2	~30	33.0	32.5	8.6
d_0, 3 h	19.0	72.0	—	200	21	26.8/135.2	~90	102	95.8	15
s_0, 1 min	1.10	0.41	1.32/1.11/1.01	—	1.1	—	—	0.58	—	1.00
s_0, 15 min	1.45	0.78	1.78/1.61/0.94	0.55	1.2	0.50/0.28	0.89	0.72	0.70	1.66
s_0, 30 min	1.50	0.88	—	0.62	1.2	0.57/0.37	0.91	0.75	0.57	1.46
s_0, 3 h	1.57	0.86	—	0.81	1.1	0.59/0.50	0.85	0.79	0.73	1.36

does not conform to a power law, although a consistent pattern does appear.

Finally, the seasonal effect is less pronounced in terms of the intermittency ratios (shown in Figs. 6b–c) than in terms of the rainfall intensities. However, there is still a significant increase in *e*-folding distance in winter compared to summer. Longer distances in winter are expected given the prevalence of stratiform over convective rainfall. Curiously, the distances actually appear to be shorter in winter than in summer when aggregated to intervals of 24 or 48 h. The fitted trends are not significantly different, though.

5. Discussion

In comparing our findings to the existing literature, focusing on studies employing exponential spatial correlation models of the form of Eq. (3) for short temporal aggregation intervals (a few minutes to a few hours), a few salient points can be identified: contrary to some other studies we have not included a nugget parameter in our parametric model (similar to Mascaro 2017). Given our relatively small number of point pairs we wanted to limit the number of degrees of freedom in our model. Moreover, we find that the inclusion of a nugget parameter does not improve the goodness-of-fit and we find that the nugget is not significantly different from 1 (section 4a). Ciach and Krajewski (2006) estimated the nugget parameter from collocated rain gauges and estimated a nugget for their experiment of 0.996 at a 1-min aggregation interval (and even higher at longer intervals), which supports our assertion that the nugget can be neglected. Tokay and Öztürk (2012) also used collocated rain gauges to estimate the nugget parameter. They found a value of 0.97 for their lowest aggregation interval of 5 min. Jaffrain and Berne (2012) estimated the nugget from a collection of point pairs with distances that were all greater than 100 m, which makes their estimation more prone to error. Nevertheless, they do not report nugget parameters lower than 0.92. Peleg et al. (2013) used the same method and obtained similar results. O and Foelsche (2019) estimated the nugget for aggregation intervals between 5 min and 24 h from a network with a minimum interstation distance of 700 m and reported values exceeding 0.9 for intervals of 15 min and longer. Villarini et al. (2008), however, reported far lower nugget parameters, e.g., 0.51 for 1-min aggregation intervals, which is in disagreement with the above-mentioned reports as well as this study. This may have been the result of the lack of point pairs below 500-m distance in their study. In addition, their analysis is based on tipping-bucket rain gauge observations, which are likely prone to strong local random errors (Ciach 2003) associated with the finite resolution of the bucket of such rain gauges. van de Beek et al. (2012) neglect the nugget as well, however, they did not consider aggregation intervals shorter than 1 h or short distances.

In all the existing literature a roughly power-law relationship is found between decorrelation distance and aggregation interval. We find decorrelation distances that are in good agreement with Ciach and Krajewski (2006) for short aggregation intervals, yet for longer intervals the decorrelation distances found in this study are considerably longer (Table 1). The decorrelation distances found by Villarini et al. (2008) and in particular those reported by O and Foelsche (2019) are in good agreement with our findings throughout the range of time scales. Conversely, the decorrelation distances found by Jaffrain and Berne (2012) and Peleg et al. (2013) are significantly shorter at all aggregation intervals. Jaffrain and Berne (2012) did a separate analysis for convective and frontal rainfall and found that decorrelation distances for frontal rainfall can be roughly 2–6 times longer than for convective rainfall, yet still shorter than what we find (Table 1). Of the studies we have compared only Tokay and Öztürk (2012) reported decorrelation

distances that were significantly longer than in our study (almost by a factor of two across aggregation intervals).

We find shape factors s_0 that are below 1 at all aggregation intervals and increasing with correlation distance, which is in agreement with Villarini et al. (2008), Tokay and Öztürk (2012), and Mascaro (2017), but counter to the findings of Ciach and Krajewski (2006) and Jaffrain and Berne (2012), who report shape factors above 1 in all cases and most cases, respectively (Table 1). All studies in this comparison reported an increase of the shape factor with aggregation interval, except for Peleg et al. (2013) and O and Foelsche (2019), who found shape factors of about 1.2 and 0.9, respectively, but no trend.

It should be noted that of these studies only Villarini et al. (2008), Mascaro (2017), and O and Foelsche (2019) measured across distances greater than 5 km (Table 1). As we have shown in Fig. 2, not taking into account longer interstation distances, at least of the same order as the decorrelation distance itself, can severely distort the parameter estimates of d_0 and s_0. If we take only into account the data from Wageningen, then our d_0 estimates are more in line with those found by Peleg et al. (2013), while our s_0 estimates are all around or higher than 1, increasing at first and from aggregation intervals of 15 min onward varying roughly around 1.5 (Figs. 2a–f).

The longer decorrelation distances we find in winter can be explained by the prevalence of frontal over convective systems in the Netherlands during winter; the opposite is generally true in summer (Lochbihler et al. 2017). Van de Beek et al. (2011, 2012) also investigated seasonal differences in spatial rainfall variability in the Netherlands. They analyzed the variability in terms of semivariance, so we cannot directly compare the magnitude of the model parameters. However, they found a range parameter that is much smaller during summer than during winter, which is consistent with our findings.

Except for Mascaro (2017) and O and Foelsche (2019), seasonal differences in spatial rainfall variability have not been investigated extensively at small scales. Both Mascaro (2017) and O and Foelsche (2019) report longer decorrelation distances in winter than in summer, in accordance with our results (Fig. 4a). However, Mascaro (2017) also report smaller values of the shape factor of the correlation function in winter than in summer (Table 1), opposite to our findings (Fig. 4b). This is likely due to climatological differences between Arizona and the Netherlands.

Tokay and Öztürk (2012) briefly discuss seasonal differences, but do not provide actual parameter values, making it hard to compare our results to theirs. They investigated two years and found that correlations in winter were higher than in summer during the first year, while correlations in winter were slightly lower than in summer in the second year. They also found large differences between the same seasons in different years. This is quite different than our results, part of which may be due to climatological differences between Virginia, United States, and the Netherlands. Both the Netherlands and Virginia are Northern Hemisphere coastal regions. However, while the climate in Virginia can be classified as humid subtropical, the Netherlands have a temperate maritime climate. Furthermore, the remnants of tropical cyclones can contribute to out-of-season rainfall patterns in Virginia (Tokay and Öztürk 2012), whereas when such remnants arrive in the Netherlands they are already much weaker.

6. Conclusions

We have used data from three different gauge/disdrometer networks at three different spatial scales in the same general area to assess the spatial variability of rainfall. We use spatial correlation and intermittency ratio as metrics and fit to these a two-parameter exponential model as a function of interstation distance. By combining datasets we obtain a more accurate regression that captures the rainfall correlation pattern over a broad range of scales. Although this may be a well-known result from the (geo)statistical literature, it remains important to emphasize this, since many studies (including four of the seven studies to which we compare our results) limit their analysis of rainfall spatial correlation to the extent of one typical weather radar or satellite pixel.

We find that spatial variability patterns at the small scale are in general agreement with the pattern at larger scales. We confirm that decorrelation distance d_0 (km) is related to the temporal aggregation interval Δt (s) by a power law relationship, $d_0 = 0.63\Delta t^{0.53}$. In winter, the decorrelation distance tends to be significantly larger and follows a different power-law relationship, $d_0 = 9.7\Delta t^{0.29}$. In general, our estimates of the parameters of an exponential decorrelation function of rainfall are in reasonable agreement with estimates by Villarini et al. (2008), Mascaro (2017), and O and Foelsche (2019) but not with other studies that only took into account interstation distances smaller than 5 km. We found that the correlation between stations at the sub-10-m scale is ~1. This is in agreement with other studies which measured correlations between collocated stations. We have analyzed the spatial patterns of intermittency in rain separately from the spatial pattern in rainfall intensity. In the Netherlands, the spatial correlation pattern of rainfall intensity within contiguous patches of rain is not significantly different than the correlation pattern including dry zones within a larger rain field.

Our results show that, in order to accurately determine spatial decorrelation distance of rainfall, a gauge/disdrometer network with mutual distances of only a few kilometers is not enough. This is demonstrated by the fit obtained using only the Wageningen dataset and is supported by our assessment of other studies. The inclusion of station pairs with mutual distances in the order of tens of kilometers is sufficient for subhourly time scales. Nevertheless, we also find that, when taking into account the full scale range from micro-γ to meso-β, the shape of the fitted correlogram does not adequately capture the correlation pattern at the micro-γ scale, hinting that the two-parameter exponential model might not be general enough to capture the true shape. Finding a better model could be an avenue for further research.

Finally, beyond the applications mentioned in section 1, the spatial pattern in intermittency ratio could be further employed to provide guidelines for the maximum spatial range between which measurement stations may be compared for quality control purposes. For example, in the case of personal

weather stations (PWS), the range so determined may be used to find with which surrounding stations the PWS must be compared (de Vos et al. 2017, 2019).

Acknowledgments. We thank Harry Stoel from the municipality of Rotterdam for making their data available to us. We thank Rudmer Jildera (KNMI) for providing the validated 10-min aggregated data from the nationwide network. The disdrometers in the Wageningen setup were kindly provided by Alexis Berne and colleagues from the École Polytechnique Fédérale de Lausanne (EPFL) in Switzerland. We also thank Pieter Hazenberg (Wageningen University and Research) for his help with installing the disdrometers. Funding for this research was provided by the Dutch Research Council (NWO), Domain Applied and Engineering Sciences (project 11944). We thank Nadav Peleg and an anonymous reviewer for their constructive comments.

Data availability statement. The Wageningen disdrometer data are freely available from the 4TU Centre for Research Data (https://doi.org/10.4121/uuid:1dd45123-c732-4390-9fe4-6e09b578d4ff). The Rotterdam disdrometer data have been the subject of two unpublished bachelor theses, by Nadine Janssen and Michael Methorst, but have otherwise not been published before. The automatic rain gauge data are freely available from KNMI (https://www.knmi.nl/nederland-nu/klimatologie/uurgegevens) (website is in Dutch). The gauge-adjusted weather radar data are freely available from the KNMI Data Platform (https://dataplatform.knmi.nl/dataset/rad-nl25-rac-mfbs-em-5min-2-0).

REFERENCES

Angulo-Martínez, M., S. Beguería, B. Latorre, and M. Fernández-Raga, 2018: Comparison of precipitation measurements by OTT Parsivel^2 and Thies LPM optical disdrometers. *Hydrol. Earth Syst. Sci.*, **22**, 2811–2837, https://doi.org/10.5194/hess-22-2811-2018.

Berne, A., and R. Uijlenhoet, 2007: Path-averaged rainfall estimation using microwave links: Uncertainty due to spatial rainfall variability. *Geophys. Res. Lett.*, **34**, L07403, https://doi.org/10.1029/2007GL029409.

——, G. Delrieu, J.-D. Creutin, and C. Obled, 2004: Temporal and spatial resolution of rainfall measurements required for urban hydrology. *J. Hydrol.*, **299**, 166–179, https://doi.org/10.1016/S0022-1694(04)00363-4.

Ciach, G. J., 2003: Local random errors in tipping-bucket rain gauge measurements. *J. Atmos. Oceanic Technol.*, **20**, 752–759, https://doi.org/10.1175/1520-0426(2003)20<752:LREITB>2.0.CO;2.

——, and W. F. Krajewski, 2006: Analysis and modeling of spatial correlation structure in small-scale rainfall in central Oklahoma. *Adv. Water Resour.*, **29**, 1450–1463, https://doi.org/10.1016/j.advwatres.2005.11.003.

de Haij, M., 2007: Automated discrimination of precipitation type using the FD12P present weather sensor: Evaluation and opportunities. Tech. Rep., KNMI, 77 pp., https://cdn.knmi.nl/system/data_center_publications/files/000/067/817/original/rp_pwsneerslagsoort_knmi_20071025_tr297.pdf?1495620917.

de Vos, L. W., H. Leijnse, A. Overeem, and R. Uijlenhoet, 2017: The potential of urban rainfall monitoring with crowdsourced automatic weather stations in Amsterdam. *Hydrol. Earth Syst. Sci.*, **21**, 765–777, https://doi.org/10.5194/hess-21-765-2017.

——, ——, ——, and ——, 2019: Quality control for crowdsourced personal weather stations to enable operational rainfall monitoring. *Geophys. Res. Lett.*, **46**, 8820–8829, https://doi.org/10.1029/2019GL083731.

Guyot, A., J. Pudashine, A. Protat, R. Uijlenhoet, V. R. N. Pauwels, A. Seed, and J. P. Walker, 2019: Effect of disdrometer type on rain drop size distribution characterisation: A new dataset for south-eastern Australia. *Hydrol. Earth Syst. Sci.*, **23**, 4737–4761, https://doi.org/10.5194/hess-23-4737-2019.

Ha, K.-J., E.-H. Jeon, and H.-M. Oh, 2007: Spatial and temporal characteristics of precipitation using an extensive network of ground gauge in the Korean Peninsula. *Atmos. Res.*, **86**, 330–339, https://doi.org/10.1016/j.atmosres.2007.07.002.

Habib, E., and W. F. Krajewski, 2002: Uncertainty analysis of the TRMM ground-validation radar-rainfall products: Application to the TEFLUN-B field campaign. *J. Appl. Meteor.*, **41**, 558–572, https://doi.org/10.1175/1520-0450(2002)041<0558:UAOTTG>2.0.CO;2.

——, ——, and G. J. Ciach, 2001: Estimation of rainfall interstation correlation. *J. Hydrometeor.*, **2**, 621–629, https://doi.org/10.1175/1525-7541(2001)002<0621:EORIC>2.0.CO;2.

Jaffrain, J., and A. Berne, 2012: Quantification of the small-scale spatial structure of the raindrop size distribution from a network of disdrometers. *J. Appl. Meteor. Climatol.*, **51**, 941–953, https://doi.org/10.1175/JAMC-D-11-0136.1.

——, A. Studzinski, and A. Berne, 2011: A network of disdrometers to quantify the small-scale variability of the raindrop size distribution. *Water Resour. Res.*, **47**, W00H06, https://doi.org/10.1029/2010WR009872.

Jameson, A., M. Larsen, and A. Kostinski, 2015: Disdrometer network observations of finescale spatial-temporal clustering in rain. *J. Atmos. Sci.*, **72**, 1648–1666, https://doi.org/10.1175/JAS-D-14-0136.1.

Krajewski, W. F., G. J. Ciach, and E. Habib, 2003: An analysis of small-scale rainfall variability in different climatic regimes. *Hydrol. Sci. J.*, **48**, 151–162, https://doi.org/10.1623/hysj.48.2.151.44694.

Lochbihler, K., G. Lenderink, and A. P. Siebesma, 2017: The spatial extent of rainfall events and its relation to precipitation scaling. *Geophys. Res. Lett.*, **44**, 8629–8636, https://doi.org/10.1002/2017GL074857.

Mandapaka, P. V., and X. Qin, 2013: Analysis and characterization of probability distribution and small-scale spatial variability of rainfall in Singapore using a dense gauge network. *J. Appl. Meteor. Climatol.*, **52**, 2781–2796, https://doi.org/10.1175/JAMC-D-13-0115.1.

Mascaro, G., 2017: Multiscale spatial and temporal statistical properties of rainfall in central Arizona. *J. Hydrometeor.*, **18**, 227–245, https://doi.org/10.1175/JHM-D-16-0167.1.

O, S., and U. Foelsche, 2019: Assessment of spatial uncertainty of heavy rainfall at catchment scale using a dense gauge network. *Hydrol. Earth Syst. Sci.*, **23**, 2863–2875, https://doi.org/10.5194/hess-23-2863-2019.

Orlanski, I., 1975: A rational subdivision of scales for atmospheric sciences. *Bull. Amer. Meteor. Soc.*, **56**, 527–530, https://doi.org/10.1175/1520-0477-56.5.527.

Overeem, A., T. Buishand, and I. Holleman, 2009a: Extreme rainfall analysis and estimation of depth-duration-frequency curves using weather radar. *Water. Resour. Res.*, **45**, W10424, https://doi.org/10.1029/2009WR007869.

——, I. Holleman, and A. Buishand, 2009b: Derivation of a 10-year radar-based climatology of rainfall. *J. Appl. Meteor. Climatol.*, **48**, 1448–1463, https://doi.org/10.1175/2009JAMC1954.1.

——, H. Leijnse, and R. Uijlenhoet, 2011: Measuring urban rainfall using microwave links from commercial cellular communication networks. *Water Resour. Res.*, **47**, W12505, https://doi.org/10.1029/2010WR010350.

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

Peleg, N., M. Ben-Asher, and E. Morin, 2013: Radar subpixel-scale rainfall variability and uncertainty: Lessons learned from observations of a dense rain-gauge network. *Hydrol. Earth Syst. Sci.*, **17**, 2195–2208, https://doi.org/10.5194/hess-17-2195-2013.

Schleiss, M., A. Berne, and R. Uijlenhoet, 2009: Geostatistical simulation of two-dimensional fields of raindrop size distributions at the meso-γ scale. *Water Resour. Res.*, **45**, W07415, https://doi.org/10.1029/2008WR007545.

——, J. Jaffrain, and A. Berne, 2011: Statistical analysis of rainfall intermittency at small spatial and temporal scales. *Geophys. Res. Lett.*, **38**, L18403, https://doi.org/10.1029/2011GL049000.

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

Sunilkumar, K., T. N. Rao, and S. Satheeshkumar, 2016: Assessment of small-scale variability of rainfall and multi-satellite precipitation estimates using measurements from a dense rain gauge network in Southeast India. *Hydrol. Earth Syst. Sci.*, **20**, 1719–1735, https://doi.org/10.5194/hess-20-1719-2016.

Tapiador, F., R. Checa, and M. de Castro, 2010: An experiment to measure the spatial variability of rain drop size distribution using sixteen laser disdrometers. *Geophys. Res. Lett.*, **37**, L16803, https://doi.org/10.1029/2010GL044120.

ten Veldhuis, M.-C., Z. Zhou, L. Yang, S. Liu, and J. Smith, 2018: The role of storm scale, position and movement in controlling urban flood response. *Hydrol. Earth Syst. Sci.*, **22**, 417–436, https://doi.org/10.5194/hess-22-417-2018.

Tokay, A., and K. Öztürk, 2012: An experimental study of the small-scale variability of rainfall. *J. Hydrometeor.*, **13**, 351–365, https://doi.org/10.1175/JHM-D-11-014.1.

van de Beek, C., H. Leijnse, P. Torfs, and R. Uijlenhoet, 2011: Climatology of daily rainfall semivariance in the Netherlands. *Hydrol. Earth Syst. Sci.*, **15**, 171–183, https://doi.org/10.5194/hess-15-171-2011.

——, ——, ——, and ——, 2012: Seasonal semi-variance of Dutch rainfall at hourly to daily scales. *Adv. Water Resour.*, **45**, 76–85, https://doi.org/10.1016/j.advwatres.2012.03.023.

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

——, H. Leijnse, A. Overeem, and R. Uijlenhoet, 2020: Estimating raindrop size distributions using microwave link measurements: potential and limitations. *Atmos. Meas. Tech.*, **13**, 1797–1815, https://doi.org/10.5194/amt-13-1797-2020.

Villarini, G., P. V. Mandapaka, W. F. Krajewski, and R. J. Moore, 2008: Rainfall and sampling uncertainties: A rain gauge perspective. *J. Geophys. Res.*, **113**, D11102, https://doi.org/10.1029/2007JD009214.