Rainfall field estimation using simulated microwave link information

Manuel Felipe Rios Gaona,
Aart Overeem, Hidde Leijnse,
Marc Bierkens & Remko Uijlenhoet

Utrecht Universiteit - KNMI, January 2012

Abstract

In this study simulated microwave-link information is used to compute maps of areal precipitation (rainfall fields) for the Netherlands. Simulated information from 2032 microwave links active in January 17^{th} - 2011, are constructed from radar data.Corrected radar data from the KNMI weather radar are used to validate the estimated results. Interpolations are carried out by the geostatistical technique Ordinary Kriging. For this technique, applied semivariograms are obtained on: *1*) spherical models fitted to the experimental semivariograms computed for the corresponding data sets; *2*) seasonal variogram parameterizations proposed and developed by van de Beek *et al.* (2011a,b). The estimated rainfall fields show good accuracy at regional and local scales in areas where microwave link density is high. Local cases are presented for the cities of Utrecht and Rotterdam (the Netherlands). Semivariogram temporal downscaling is carried out for time-aggregation scales of 15-minute, 1-, 3-, 6- and 12-hour. For these scales, estimated rainfall fields are obtained and filtered to account for the percentage of microwave links not registering rainfall attenuation in their signals. The Mean Error *(ME)*, Root Mean Squared Error *(RMSE)* and Variance Ratio *(VR)* are computed in order to investigate the bias, accuracy and variability of rainfall field estimations for different time scales.The highest correlation is found for the 6-h aggregated time scale. Overall, good results are obtained for the other time scales, suggesting that the developed methodology is suitable for automatic rainfall estimation at small aggregated time scales.

1 Introduction

The main scope of the present study, is to develop a methodology for the generation of interpolated rainfall fields at 24-hour and smaller time scales, based on simulated microwave link rainfall measurements. Accurate rainfall fields are the essential driving input of all processes involved in the terrestrial part of the water cycle, (Lanza *et al.*, 2001).

The state-of-the-art in rainfall measurements accounts for instruments that detect and quantify different properties of rainfall depending on their location. Michaelides *et al.* (2009) reviews the applicability and theory behind disdrometers, rain gauges, radars and satellites. Rain gauges provide the only physically direct method of collecting and measuring rainfall, generally providing relatively accurate point estimates, especially for low-to-intermediate intensity rainfalls, (Muller and Kidd, 2006). The main drawback of rain-gauge measurements is their point nature, requiring the use of interpolation techniques for rainfall field estimation, (Creutin *et al.*, 1988). Muller and Kidd (2006) also highlight some drawbacks in rain-gauge measurements such as their coarse resolution to detect significant changes in rainfall rate; and rainfall overestimation at low intensities ($< 50\text{mm} \cdot \text{h}^{-1}$), and underestimation at higher intensities ($> 100\text{mm} \cdot \text{h}^{-1}$). Wind, evaporation and the spatial and temporal variation of the drop size distribution are also considered, by Michaelides *et al.* (2009), as drawbacks in the use of rain gauges. The advantage of radar rainfall measurements is the real time monitoring of a wide area from a single point with high spatial and temporal continuity and resolution, (Sauvageot, 1994). The disadvantages of radar technologies come from the potential inaccuracies in the measurement of the reflectivity factor Z, attributed to radar calibration, attenuation, anomalous propagation, beam blockage and range effects, (Hunter, 2009). Rainfall measurements by radar still require correction to produce reliable rainfall fields. In this study corrected radar rainfall measurements are used as validation data ("ground-truth"), due to their correspondence at small time scales with microwave link measurements.

Despite recent and cutting-edge implementation of microwave link technology in rainfall estimation (Zinevich *et al.*, 2008; Goldshtein *et al.*, 2009; Leijnse *et al.*, 2007b), this technique is still in development. Given its potential, the application of microwave link technology in rainfall estimation is suitable for three main reasons: *1*) the broad use of microwave link networks all around the world, builds up the possibility for large sources of information, especially in countries with scarce conventional networks, like manual or automatic rain gauges; *2*) the spatial characteristics of such networks allow to incorporate not only many more rainfall measurements but line-averages,

being these latter more representative of areal rainfall than point measurements; *3*) this technology allows measurements at very small time intervals, which is very useful not only for urban analysis but many other cases where rainfall-runoff processes take place at very short time scales.

Diverse methods of interpolation have been developed, implemented and studied for rainfall field estimation (Cuccoli *et al.*, 2011; Schuurmans *et al.*, 2007). In this study a statistical method of interpolation is applied because apart from an estimated value, the variance of the estimation error can also be calculated. Ordinary Kriging is selected for its fundamental characteristic of considering the variation of the mean across the studied area. Daily rainfall information from January 17$^{\text{th}}$-2011 8:00 UTC to January 18$^{\text{th}}$-2011 8:00 UTC, is aggregated into six time scales: 15 minute, 1-, 3-, 6-, 12- and 24-hour. Estimated rainfall fields are computed following the methodology proposed and developed by van de Beek *et al.* (2011a,b) for downscaling semivariograms. Simulated microwave link data is obtained from the same data set used for validation. The Variance Ratio is proposed as a key metric in the analysis of the correct estimates of the variance. The accuracy of estimates is estimated by looking for the Mean Error (bias) and the Root Mean Squared Error (accuracy).

Two main components represent the aggregated value of this study in the estimation of rainfall fields: *a*) The use of a large microwave link network and the possibility for using a good validation data set; *b*) The use of simulated microwave link data and corrected radar data for temporal disaggregation of rainfall at small time scales ($<$ 1 hour). The remaining part of this thesis is divided as follows: section 2 describes the general principles and the theory behind microwave link technology and its use in rainfall estimation. Section 3 presents the mathematical concepts of semivariogram models, Ordinary Kriging and related geostatistical features. The *van de Beek* approach is also detailed in this section. The methodology developed for rainfall field estimation is presented in section 4. Results and their analyses are discussed in section 5. Conclusions and recommendations in section 6.

2 Microwave Links

The broad use of cellular communications nowadays, enables the use of fixed and commercial microwave radio networks for rainfall estimation. Wireless communication networks can be useful in measuring precipitation and complementing existing measurement systems like rain gauge networks and weather radars. The high link density in such networks, make them suitable for regional rainfall monitoring in combination with operational weather

radars (Leijnse *et al.*, 2007b). Rain gauges have a high degree in accuracy when measuring precipitation, and often few rain gauges are available; but given that these measurements are represented as points in space, they do not provide a reliable spatial description of rainfall variability. Messer *et al.* (2006) highlight some useful properties of microwave radio networks such as: their capability for measurements close to the ground, and their operating frequencies in tens of GHz. Their deployment, allows near-surface rainfall measurements. On the other hand, Zinevich *et al.* (2008) points out some sources of uncertainty when estimating rainfall intensities from attenuation measurements such as: the variability of the drop size distribution (DSD) along the link, the attenuation due to wet antennas, the determination of clear air attenuation due to water vapor-induced attenuation, and scintillation effects. Besides these uncertainties, cellular communication providers are interested in having the least possible attenuation through the network. They design and optimize their networks for communication tasks around the frequency of the signals, the geometry of the links and their lengths, and the received signal level (RSL) measurement protocols; hence, making challenging the use of such networks for hydrologic purposes (Messer, 2007).

It is not completely right to think of microwave links as straight lines (from transmitter to receiver) or to only attribute the attenuation in their signals to rainstorms. In reality, the beam of the signal is more like a narrow cone widening as it leaves the transmitter (Upton *et al.*, 2005). Rainfall is not the only source of attenuation, and although it is true that it causes the strongest variations in the signals, variations are also due to atmospheric adsorption, being dependant on temperature, pressure and humidity. Hence, the strength of the received signal is never constant, not even in dry conditions (weather). In their work, Upton *et al.* (2005), present an illustrative example in which the variations in the received signal strength are depicted for dry and rainy conditions. The power registered by the signal, when is not being attenuated by rainfall, is usually called the *reference level* or *base level*.

The temporal resolution of attenuation measurements goes from every minute to only one minimum RSL measurement per day. In between, there is equipment designed to only measure the minimum and maximum RSL in 15 minutes or one instantaneous value of power each 15 minutes (Zinevich *et al.*, 2008). Estimation of space-time rainfall intensities from RSL measurements, in microwave-links networks, has been studied by Leijnse *et al.* (2007b) and Messer *et al.* (2006). The attenuation of the signal is due to scattering and adsorption by water droplets. A larger decrease in the RSL will be registered for an increasing number and size of the raindrops present along the beam. Atlas and Ulbrich (1977) demonstrated that at frequencies of about

35 GHz, the power-law relationship, Equation (1), is approximately linear and independent on DSD and temperature. Accurate rainfall estimations, along the link, are achieved when using the difference in the attenuation at two different frequencies, due to the linearity between this difference and the rainrate (Upton *et al.*, 2005). However, often only single-frequency links are available.

Rainfall estimation through commercial microwave-links networks can be performed using a power-law relation between attenuation and rainfall rate (Atlas and Ulbrich, 1977):

$$k = a \cdot R^{\,b} \tag{1}$$

where k [dB $\cdot$ km^{-1}] is the specific attenuation and R [mm $\cdot$ h^{-1}] the rainrate. Overeem *et al.* (2011) state that in order to be able to derive rainfall intensities from received signal powers, the point-scale $k-R$ relation [Equation (1)] could be assumed as a good approximation of path-averaged rainfall intensities. Thus, Equation (2) expresses the rainfall intensity as a function of the specific attenuation, where a and b are constants depending on the frequency of the signal, temperature, size and shape of the raindrops; R [mm $\cdot$ h^{-1}] is the rainrate; P_{ref} [dB $\cdot$ km] the reference signal level or base line; P [dB $\cdot$ km] the received signal power and L [km] the microwave-link length.

$$\langle R \rangle = a \cdot \left(\frac{P_{ref}\,(L) - P\,(L)}{L} \right)^{b} \tag{2}$$

Leijnse *et al.* (2007a) present different values for coefficients a and b established in previous studies.

Wet antenna attenuation is also taken into account by Overeem *et al.* (2011) when calculating the corrected path-averaged, mean 15-min rainfall intensity [Equation (3)]. Here, A_a [dB] is the wet antenna attenuation (0 dB for dry conditions), L [km] the length of the link, A_{max} and A_{min} in [dB] the maximum and minimum rain induced attenuation in the measured interval, a and b same coefficients as in Equation (2), and α the coefficient determining the contribution of the maximum and minimum attenuation in the measured interval. k_{max} and k_{min} can be read as the maximum and minimum specific attenuations in [dB $\cdot$ km^{-1}], where H represents the Heaviside function ($H = 0$ when its argument is smaller than zero, $H = 1$ everywhere else).

$$\begin{aligned} k_{max} &= \frac{A_{max} - A_a}{L} \cdot H \cdot (A_{max} - A_a)\,, \\ k_{min} &= \frac{A_{min} - A_a}{L} \cdot H \cdot (A_{min} - A_a)\,, \\ \langle R \rangle &= \alpha \cdot a[k_{max}]^b + (1-\alpha) \cdot a[k_{min}]^b \end{aligned} \tag{3}$$

2.1 Topology of the microwave link network used in the present study

The number of links in commercial networks in the Netherlands rises up to 12000 (Leijnse *et al.*, 2007b); but for the present study, only 2032 links are used as a source of information. Simulated microwave link data come from the fully active network of T-Mobile in January 17$^{\text{th}}$, 2011. This one-sixth use of the full microwave link network capacity, in the Netherlands, is due to the fact that telecommunication service is offered by several providers, most of the times, owners of their own networks. Nevertheless, when compared to the manual rain-gauge network, 329 gauges (van de Beek *et al.*, 2011a), or even with the automatic network, 33 gauges (*ibid*)[1], the number of measurements in time and space, increase substantially the possibility for both better rainfall field estimations and aggreation of rainfall into short-time scales, 15-minute for instance. Figure 2 displays the 2032 microwave links used in the present study. From this figure, it can be observed that the highest densities occur in the cities of Amsterdam, Den Haag, Rotterdam and Utrecht.

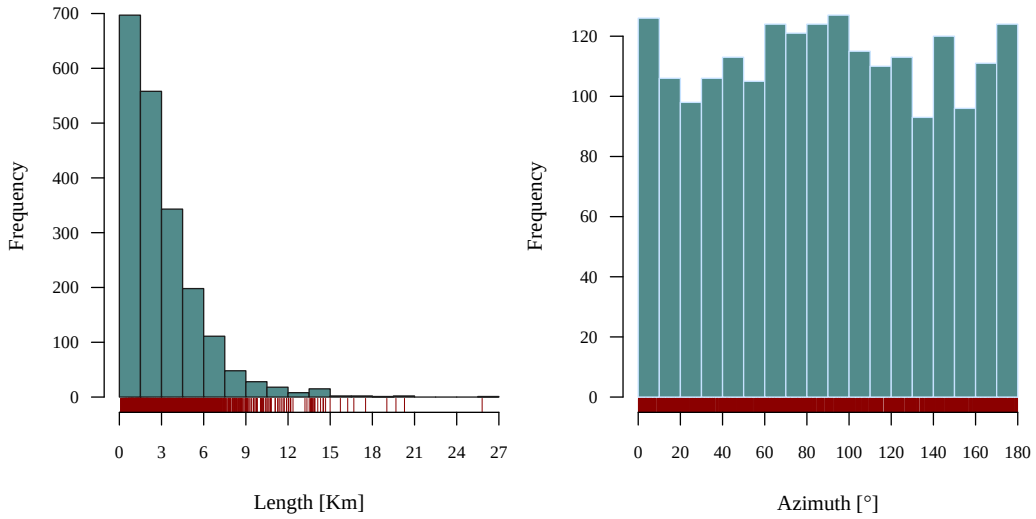

Figure 1: Histograms of microwave link length (left), and orientation with respect to the north (right).

The mean value of the microwave links orientation is 179.2 ° with respect to the north[2]. Figure 1, with frequency distribution of link length and orientation, shows that the microwave links are not orientated in any particular

[1]Overeem (2009) approximately accounts for 35 gauges for the automatic network and 325 gauges for the manual network.

[2]For an angle greater than 180 °, its supplementary angle is taken.

direction. This characteristic is important because then the performance of the rainfall field estimation is not influenced by any directionality of the microwave link network.

The average microwave link length is 3.1 km; and from what is observed in Figure 1, 88% of the links have lengths smaller than 6 km. 0.089 km and 25.809 km are the minimum and maximum link-lengths in this network.

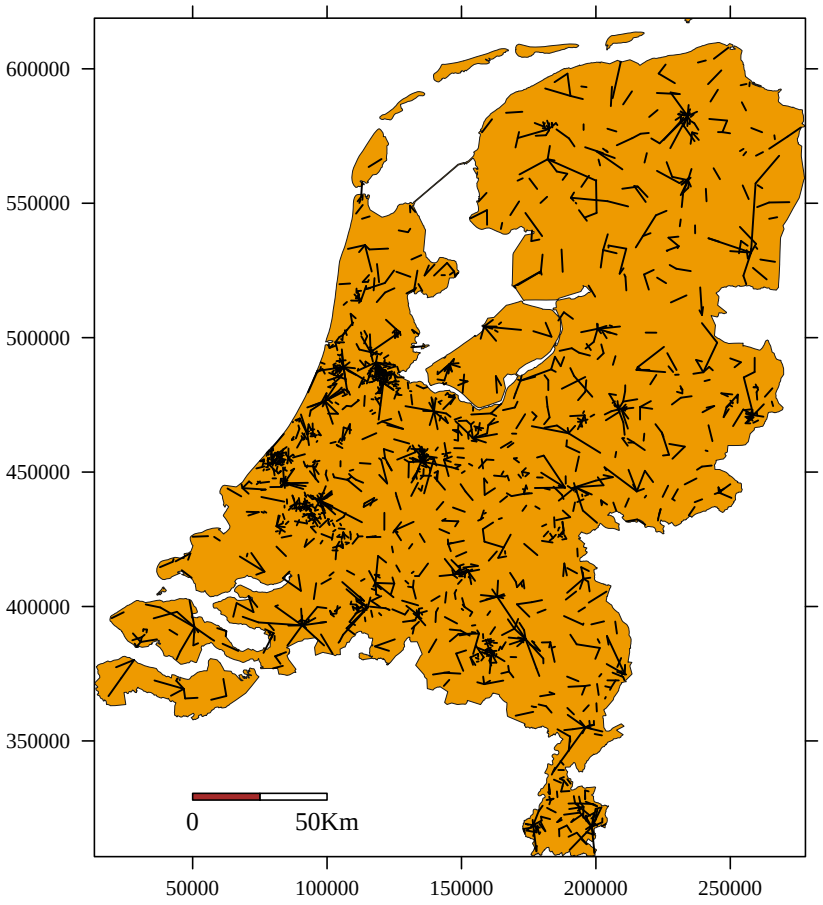

Figure 2: T-Mobile active microwave link network from January 17$^{\text{th}}$-2011 8:00 UTC to January 18$^{\text{th}}$-2011 8:00 UTC. 2032 links.

2.2 How a microwave link network functions

In communication systems, there are different types of links: *a*) overhead lines; *b*) underground copper cables; *c*) radio; and, in modern times, *d*) fiber-optic cables and lasers.

Microwave link networks are used not only for cellular communications but for many other services, like the broadcast of programs from studio to the transmitter location, high-speed Internet access or intercommunication between specific networks. The broad use of microwave links is due to their capability for moving large amounts of information at high speeds; and despite some attenuation in the signal, mainly due to rainfall, the transmission is not disrupted. There are five key components in a simple one-way microwave link:

❶ **Transmitter:** produces the signal in which the information is contained. This information can be telephone calls, text, television or radio programs, web pages, or any possible combination of these media. The transmitter not only generates the energy at the required power and frequency but also modulates the signal.

❷ **Transmission lines:** coaxial cables or waveguides (hollow pipes) transport the signal from the transmitter to the antenna or from the antenna to the receiver.

❸ **Antennas:** emit and collect the microwave signal from the transmission line into the free space, or vice versa. Their highly directional factor allows signal transmission over long distances using small amounts of power.

❹ **Clear path:** due to the signal propagation is mainly in straight lines. A clear path should be established for a successful transmission of data.

❺ **Receiver:** extracts the signal from the microwave energy used in its transmission and demodulates its codification to make the signal available in its original form. Because of the attenuation in the power that the signal experiences through its travelling from antenna to antenna, the receiver must be capable of detecting low-powered microwave signals.

The entire process of emitting and receiving the signal is carried out from antenna to antenna at the speed of light; in practical terms, data transmission is instantaneous. Figure 3 shows a simple diagram on data transmission through microwave links.

2.2.1 Making a call from a cellphone

During a cell phone call, the sound waves produced by the user are processed by the digital signal processor (DSP), embedded in the cell phone, and transformed in binary information. The binary data is splitted into four channels: a sending channel, a receiving channel, and two control channels which track the strength of the signal to and from the closest base station (BS). Binary information is sent to the BS, previously identified by the cell phone, as soon

as this latter is turned on. The BS sends the call (binary information) to the related Mobile Switching Center (MSC)[3]. It is in this part of the process that microwave links become fully active in emitting and receiving electromagnetic signals. If the call is intended to either a land-based phone or a mobile number of another service provider, the MSC passes the information to the PSTN (public switched telephone network). The information is either sent to another MSC or BS or even a land-based phone. This depends on the device (cell phone or land-based phone) and the location of the user receiving call with respect to the user making the call. Once the call has reached the final user, the binary information goes to the DSP, through the receiving channel, and is back-transformed into sound waves.

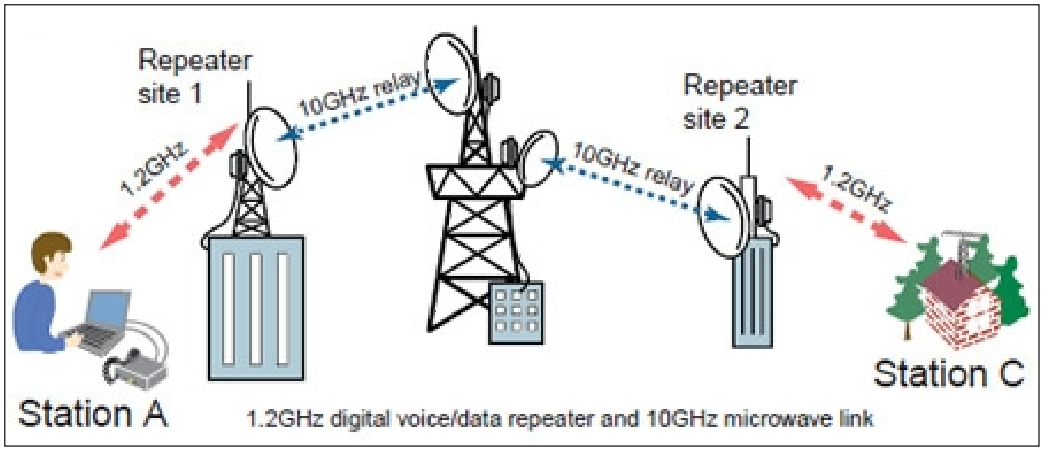

Figure 3: Data transmission using microwave links. Taken from Icom America Inc. (2011).

3 Stochastic Rainfall Modeling

Foufoula-Geourgiou and Krajewski (1995) summarize in detail all the previous work and advances done in this field. The different approaches that have been developed in this field, are presented below as a general overview of the different ways of rainfall modeling from the stochastic point of view.

❶ **Point Processes** are models based on a random collection of points, where each point represents the time and location of an event, with some random value of intensity and duration associated to them. These models are not able to describe the statistical structure of rainfall over a wide range of scales and its parameterization is rather cumbersome.

[3]An MSC is also known as mobile telephone switching office (MTSO).

❷ **Scaling Models** of rain have evolved from fractal geometry, to monofractal fields, to multifractals, to generalized scale-invariant models, to universal multifractals.

❸ **Multiplicative Cascades** have their origin in the statistical theory of turbulence, where rainfall is analyzed as an analogy to atmospheric turbulence which dictates a cascading of energy from a large input scale to smaller scales. Its parameter estimation is also not trivial.

❹ **Wavelet Transforms** offer a method for decomposing some process into "atoms" localized not only in frequency but in space. Useful for inhomogeneous fields such as rainfall events.

❺ **Other approaches:** the diffusion model (which, under some assumptions, results in a family of asymptotic probability distributions), simplified dynamics of raindrop processes, Markovian type models, aggregation models, geostatistics and some other probability distribution models.

Lanza *et al.* (2001) briefly reviews the evolution of the geostatistical approach for rainfall modeling; since the early concept of objective analysis developed by Gandin (1970) in meteorological fields, to the work done by Matheron (1971) who established an appropriate mathematical background to the previous work carried out by Krige (1951) in the field of mining engineering. The result was an optimal linear interpolation technique known as Kriging, which is characterised for being unbiased and minimum in its variance. The benefit of this technique depends on the *a priori* knowledge of the spatial covariance (for stationary random functions) or upon the knowledge of a special function: the variogram, also referred as *semivariogram*, (for non-stationary fields).

3.1 Geostatistical framework

The present work is entirely based in the geostatistical approach for two main reasons: *a*) Kriging is well-suited technique for interpolation of data points that are irregularly spaced; and *b*) interpolation by Kriging not only yields an estimate of the field itself, but also of the variance of this field.

3.1.1 Semi-Variogram

Using the notation presented by Goovaerts (1997), rainfall intensity [mm] is represented, in here, as the continuous random variable $r(\mathbf{u}_\alpha)$, where $\mathbf{u}_\alpha$ is the vector of spatial coordinates (x, y) at the α^{th} location. The distance and direction between any pair of rainfall intensities is represented by the vector

h. If homogeneity and isotropy is assumed, the vector **h** can be represented by its magnitude, *i.e.*, $h = |\mathbf{h}|$.

In geostatistics, measures of similarity are modeled by the covariance and correlation coefficients, Equations (4) and (5) respectively, where $N(h)$ is the number of data pairs within the class defined by h, m_{-h} and m_{+h} the means of the corresponding tail and head values[4], and σ^2_{-h} and σ^2_{+h} their respective variances. Computing either the covariance or the correlation coefficients for different lags $h_1, h_2 > h_1, h_3 > h_2, \ldots$ will lead to what is known as the *experimental (auto)covariance function* or the *experimental (auto)correlation function*[5]. The difference between these two functions is that the latter is a standardized unit-free measure.

$$C(h) = \frac{1}{N(h)} \cdot \sum_{\alpha=1}^{N(h)} r(\mathbf{u}_\alpha) \cdot r(\mathbf{u}_\alpha + h) - m_{-h} \cdot m_{+h} \tag{4}$$

$$\text{with} \quad m_{-h} = \frac{1}{N(h)} \cdot \sum_{\alpha=1}^{N(h)} r(\mathbf{u}_\alpha) \quad \text{and} \quad m_{+h} = \frac{1}{N(h)} \cdot \sum_{\alpha=1}^{N(h)} r(\mathbf{u}_\alpha + h)$$

$$\rho(h) = \frac{C(h)}{\sqrt{\sigma^2_{-h} \cdot \sigma^2_{+h}}} \quad \in [-1, +1] \tag{5}$$

$$\text{with} \quad \sigma^2_{-h} = \frac{1}{N(h)} \cdot \sum_{\alpha=1}^{N(h)} [r(\mathbf{u}_\alpha) - m_{-h}]^2$$

$$\text{and} \quad \sigma^2_{+h} = \frac{1}{N(h)} \cdot \sum_{\alpha=1}^{N(h)} [r(\mathbf{u}_\alpha + h) - m_{+h}]^2$$

Conversely to the covariance and correlation functions, the *experimental semivariogram*[6], Equation (6), is a measure of variability (dissimilarity) between data separated by a lag h. The semivariogram of a Random Function characterizes the degree of spatial variability between any two Random Variables, $r(\mathbf{u}_\alpha)$ and $r(\mathbf{u}_\alpha + h)$ in this case, separated by a lag h (Journel, 1989).

$$\gamma(h) = \frac{1}{2 \cdot N(h)} \cdot \sum_{\alpha=1}^{N(h)} [r(\mathbf{u}_\alpha) - r(\mathbf{u}_\alpha + h)]^2 \tag{6}$$

[4]Tail are all the values taken as initial points of reference; whereas head indicates all the values separated by the lag h from those initial points of reference.

[5]Also known as correlogram.

[6]Also known as *empirical semivariogram*

The covariance function and the semivariogram are related through:

$$C(h) = C(0) - \gamma(h) \tag{7}$$

$C(0)$ can either be set equal to the sill of the semivariogram model, *i.e.*, $\gamma(h)$ when $h \to \infty$, if the random function is a wide sense stationary one[7]; or can be set to any arbitrary large value. Equation (7) expresses the reciprocity between the semivariance and covariance [also Figure 4 (a)]. An increase in the semivariance indicates less relation among random variables through the lag h, *i.e.*, a decrease in the covariance. Stationarity is not an intrinsic property of the actual distribution of the random variables over an area but a decision made when assuming that the semivariogram, modeled after some average over the region in study, is assumed representative of the population in the whole area. This decision is implicit in all statistics and it is not particular to the geostatistical approach (Journel, 1989).

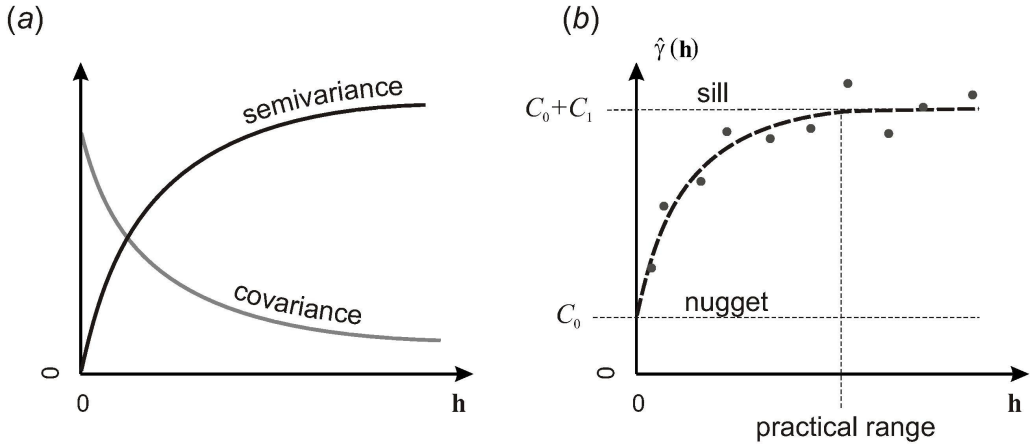

Figure 4: (*a*) Reciprocity between semivariance and covariance. (*b*) Scheme and parameter representation of a linear combination of Nugget-effect and Spherical models. Taken from Hengl (2009, p. 17).

Interpolation algorithms, *e.g.* any Kriging technique, demands the easy estimation of (semi)variance or covariance values for any possible lag h; therefore, the need for fitting continuous and positive definite functions[8] to the experimental semivariogram or covariance model. In practice, every possible

[7] In a wide sense stationary function, also called homogeneous random function, the mean is constant, the variance is constant and finite, and the covariance only depends on the lag h (Bierkens, 2010). Neither the mean nor the variance depend on space or time.

[8] A function f contained in the domain Ω such that $f : \Omega \to \Re$, where $\Omega \subset \Re^n$ and $0 \in \Omega$; is said to be **positive definite** on Ω, if and only if $f(0) = 0$ and $f(x) > 0$, for $x \in \Omega$ and $x \neq 0$ (King *et al.*, 2003, p. 381).

linear combination of basic models can be used to best-fit the empirical semivariogram; see Goovaerts (1997, p. 88), Bierkens (2010, p. 81) or Pebesma (2001, p. 38). Accounting for measurement errors and/or spatial sources of variation at distances smaller than the shortest sampling interval, will lead into a discontinuity at the origin of the semivariogram called the nugget effect, *i.e.*, $\gamma(0) \neq 0$ (Goovaerts, 1997, p. 31).

In the present study, Equation (10) is used in fitting the empirical semivariograms. This equation is a linear combination of the Nugget-effect and Spherical models, Equations (8) and (9) respectively; thus, the total semivariogram is given by: $\boldsymbol{\gamma(h)_{TOTAL} = \gamma(h)_{Nugget-effect} + \gamma(h)_{Spherical}}$.

$$\gamma(h)_{Nugget-effect} = \left\{ \begin{array}{ll} 0 & \text{if} \quad h = 0 \\ C_0 & \text{otherwise} \end{array} \right\} \quad \text{with} \quad C_0 > 0 \tag{8}$$

$$\gamma(h)_{Spherical} = \left\{ \begin{array}{ll} C_1 \cdot \left[\dfrac{3}{2} \cdot \left(\dfrac{h}{a} \right) - \dfrac{1}{2} \cdot \left(\dfrac{h}{a} \right)^3 \right] & \text{if} \quad h \leq a \\ C_1 & \text{otherwise} \end{array} \right\} \tag{9}$$

$$\text{with} \quad a \text{ and } C_1 > 0$$

$$\gamma(h)_{TOTAL} = \left\{ \begin{array}{ll} 0 & \text{if} \quad h = 0 \\ C_0 + C_1 \cdot \left[\dfrac{3}{2} \cdot \left(\dfrac{h}{a} \right) - \dfrac{1}{2} \cdot \left(\dfrac{h}{a} \right)^3 \right] & \text{if} \quad 0 < h \leq a \\ C_0 + C_1 & \text{otherwise} \end{array} \right\} \tag{10}$$

$$\text{with} \quad a, C_0 \text{ and } C_1 > 0$$

Figure 4 (b) presents a graphical description of the parameters associated with the linear combination of the Nugget-effect and Spherical models. The range a is the distance at which random variables are still correlated. Further away of this point ($h \geq a$) the semivariance reaches the sill C, that is, becomes constant and maximum. If the Random Function R is wide sense stationary, the sill C is equal to the variance, *i.e.*, $\gamma(h \geq a) = C = \sigma_R^2$.

Figure 5 shows different representations of the empirical semivariogram for 2032 simulated microwave link values of daily rainfall intensity in January 17$^{\text{th}}$, 2011. If every individual point-pair semivariance, of the whole data set, is plotted against every possible lag h, the result will look like a very dense cloud of points, Figure 5 (left). If all the possible h-lags are grouped into bins ($h_1, h_2 > h_1, h_3 > h_2, \ldots$), and the semivariances are averaged accordingly to

these bins, the semivariogram will look like a noisy (and continuous) function, Figure 5 (centre and/or right).

The fitting of a semivariogram model to the empirical semivariogram is done by iteratively reweighted least squares (WLS); minimizing Equation 11, with w_α either equal to $N(h)$ or to $N(h) \cdot \gamma(h)_\alpha^{-2}$, Pebesma (2001).

$$\sum_{\alpha=1}^{N(h)} w_\alpha \cdot [\hat{\gamma}(h)_\alpha - \gamma(h)_\alpha]^2 \tag{11}$$

After the selection of a "guess" value, each iteration is made over: *a*) a direct (least squares) fit of the partial sills; and *b*) an iterated search, using gradients, for the optimal range value(s). The iterative procedure ends after exceeding 50 iterations or when the fitted model has converged, that is, when the change in the weighted sum of squared differences between the semivariogram model and empirical semivariogram becomes less than $1.0\,\mathrm{e}^{-6}$ times the last value of this sum of squares, Pebesma (2001, 2011). The fitting is done using the Gauss-Newton method with (mostly) analytical derivative functions.

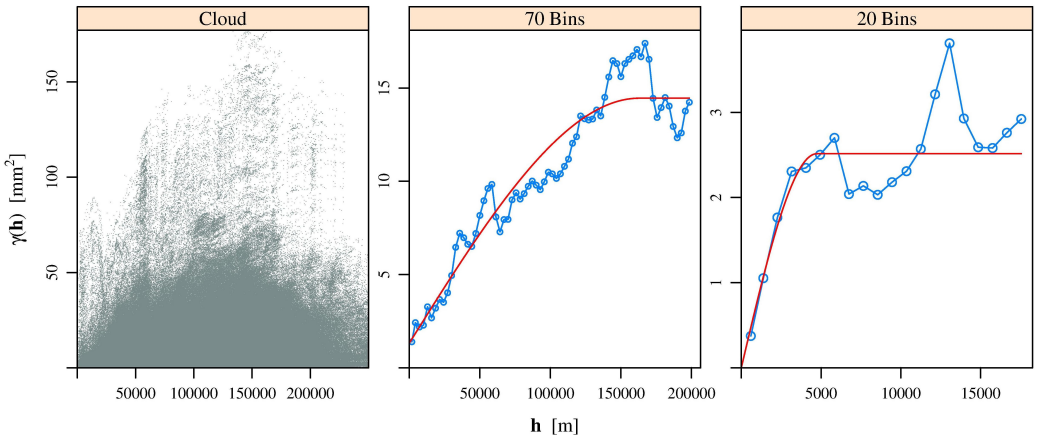

Figure 5: Empirical semivariogram models for 2032 simulated microwave link values of daily rainfall in January 17$^{\text{th}}$, 2011. The red line represents the fitted curve (semivariogram) for the empirical semivariogram using a spherical model. Parameters for the fitted semivariogram with 70 classes of **h** and maximum lag of 200 km: range 162.97 km, sill 13.13 mm^2 and nugget 1.32 mm^2 (centre). Parameters for the fitted semivariogram with 20 classes of **h** and maximum lag of 18 km: range 4.73 km, sill 2.51 mm^2 and no nugget (right).

3.1.2 Ordinary Kriging

Kriging techniques, after Krige (1951), are a family of generalized least-squares regression algorithms used to estimate the value of a continuous attribute r, at any unsampled location $\mathbf{u}$, using n sampled r-data, *i.e.*, $\{r(\mathbf{u}_\alpha), \alpha = 1, \ldots, n\}$ within a region A. The different kriging techniques: *simple, ordinary, with a trend, factorial, of residuals, with external drift(s), Poisson, etc.*; are variants of the basic linear-regression estimator $R^*(\mathbf{u})$:

$$R^*(\mathbf{u}) - m(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(\mathbf{u}) \cdot [R(\mathbf{u}_\alpha) - m(\mathbf{u}_\alpha)] \tag{12}$$

with $\lambda_\alpha(\mathbf{u})$ the weights assigned to datum $r(\mathbf{u})$ [here expressed as realizations of the Random Variable (RV) $R(\mathbf{u}_\alpha)$]; $m(\mathbf{u})$ and $m(\mathbf{u}_\alpha)$ the expected values of the RVs $R(\mathbf{u})$ and $R(\mathbf{u}_\alpha)$. Only the $n(\mathbf{u})$ data within a given neighborhood $W(\mathbf{u})$ centered on $\mathbf{u}$, are retained and used in computing the estimator $R^*(\mathbf{u})$.

The preference for using Ordinary Kriging (OK) as the interpolation algorithm in the present study, is due to its capability for accounting local variations of the mean, within the study area A, by limiting the domain of stationarity of the mean to the local neighborhood $W(\mathbf{u})$ centered on $\mathbf{u}$, the location of the estimator $R^*(\mathbf{u})$. In OK the mean $m(\mathbf{u})$ deemed constant but unknown, $\forall\ \mathbf{u} \in W(\mathbf{u})$; and is filtered from the linear estimator by forcing all the weights $\lambda_\alpha(\mathbf{u})$ to sum 1. The OK estimator $R^*(\mathbf{u})$ is unbiased since the error mean is equal to zero, *i.e.*, $E\{R^*(\mathbf{u}) - R(\mathbf{u})\} = m(\mathbf{u}) - m(\mathbf{u}) = 0$ (Goovaerts, 1997); thus, the OK estimator $\hat{R}_{OK}(\mathbf{u})$ can now be rewritten as a linear combination only of the $n(\mathbf{u})$ RVs $R(\mathbf{u}_\alpha)$:

$$\hat{R}_{OK}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{OK}(\mathbf{u}) \cdot R(\mathbf{u}_\alpha) \quad \text{with} \quad \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha^{OK}(\mathbf{u}) = 1 \tag{13}$$

The $n(\mathbf{u})$ weights $\lambda_\alpha^{OK}(\mathbf{u})$ are also determined to minimize the error variance. When minimizing the error variance $\sigma_E^2(\mathbf{u}) = Var\{\hat{R}_{OK}(\mathbf{u})\} + Var\{R(\mathbf{u})\} - 2 \cdot Cov\{\hat{R}_{OK}(\mathbf{u}), R(\mathbf{u})\}$, the definition of a Lagrangian $L(\mathbf{u})$ has to be made. The Lagrangian $L(\mathbf{u})$, also called *Lagrange multiplier*, is a function of the data weights $\lambda_\alpha^{OK}(\mathbf{u})$ and a Lagrange parameter $2 \cdot \mu_{OK}(\mathbf{u})$ (Goovaerts, 1997, p. 133). The optimal weights $\lambda_\alpha^{OK}(\mathbf{u})$ are obtained by setting to zero the $[n(\mathbf{u}) + 1]$ first partial derivatives of the Lagrangian $L(\mathbf{u})$ with respect to its variables $\lambda_\alpha^{OK}(\mathbf{u})$ and $2 \cdot \mu_{OK}(\mathbf{u})$. Hence, the Ordinary Kriging predictor $\hat{R}_{OK}(\mathbf{u})$, for the single location $(\mathbf{u})$, is a system of $[n(\mathbf{u}) + 1]$ linear equations with $[n(\mathbf{u}) + 1]$ unknowns: the $n(\mathbf{u})$ weights $\lambda_\alpha^{OK}(\mathbf{u})$ and the Lagrange

parameter $\mu_{OK}(\mathbf{u})$. This linear system can be expressed in terms of the semivariance when accounting for Equation (7):

$$\hat{R}_{OK}(\mathbf{u}) = \begin{cases} \sum_{\beta=1}^{n(\mathbf{u})} \lambda_{\beta}^{OK}(\mathbf{u}) \cdot \gamma\left(\mathbf{u}_{\alpha} - \mathbf{u}_{\beta}\right) + \mu_{OK}(\mathbf{u}) = \gamma\left(\mathbf{u}_{\alpha} - \mathbf{u}\right) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \alpha = 1, \ldots, n\left(\mathbf{u}\right) \\ \sum_{\beta=1}^{n(\mathbf{u})} \lambda_{\beta}^{OK}(\mathbf{u}) = 1 \end{cases} \tag{14}$$

The advantage in using this kind of statistical-interpolation technique(s) is that it also allows to estimate the associated uncertainty of the predicted estimator $R^{*}(\mathbf{u})$. Thus, the minimum error variance[9] of the OK predictor $\sigma_{OK}^{2}(\mathbf{u})$ is obtained by:

$$\sigma_{OK}^{2}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_{\alpha}^{OK}(\mathbf{u}) \cdot \gamma\left(\mathbf{u}_{\alpha} - \mathbf{u}\right) + \mu_{OK}(\mathbf{u}) \tag{15}$$

A unique solution of the linear OK system, Equation (14), and a positive OK variance, Equation (15), is ensured only if the semivariogram function is chosen to be "conditionally positive definite" (Bierkens, 2010, p. 130).

3.2 The "van de Beek" approach

The seasonal variation of fitted semivariogram parameters for a spherical model, *i.e.*, range and sill, is studied by van de Beek *et al.* (2011a,b). Such seasonal variation is found to be well described, and reasonably accurate, by simple cosine functions. The study is based on rainfall rainrates registered in the last 30 years; carried out not only across the entire Netherlands but also for accumulated intervals of 1, 2, 3, 4, 6, 8, 12 and 24 hours. The model describing the seasonal variation of the fitted parameters is expressed through the cosine function:

$$x_{t} = \left[x_{0} + A \cdot cos\left(2 \cdot \pi \cdot f\left(t - t_{0}\right)\right)\right]^{4} \tag{16}$$

with x_{t} the parameter to be estimated: range or sill; t $[1, 365]$ the day of the year for which the parameters have to be computed. The angle for which the cosine has to be calculated is expressed in radians. Table 3 gives the values of the different coefficients in Equation (16), according to the fitted parameter being estimated.

[9] Sometimes referred to as *the estimated variance of the prediction error.*

3.2.1 Downscaling

❶ **Parameters Ratio:** due to the similarity between the seasonal variation of the sill and range throughout the different time scales; the ratio of the p-parameter, sill or range, of a 24-hourly-based semivariogram over the same p-parameter of any other smaller time scale, can be modeled as a power-law function of the form $\boldsymbol{y = a \cdot x^{b}}$, with x [hours] the time scale for which the 24-hour parameters are to be downscaled. Table 1 gives the coefficients a and b of this power-law function for the sill and range parameters.

Table 1: Parameters of the power-law functions describing the ratio of 24-hourly sill and range parameters over any other smaller time scale(s). The corresponding standard errors are also given. Taken from van de Beek *et al.* (2011b).

p-parameter	a	b	ϵ
Range	2.92	-0.34	3×10^{-2}
Sill	0.04	0.94	3×10^{-3}

❷ **Cosine functions:** the same power-law function $y = a \cdot x^{b}$ is obtained, in here, not from the p-parameters but from the coefficients A, t_0 and x_0 of the *seasonal-variation* model [Equation (16)]. For each p-parameter (range or sill), Table 2 gives the coefficients a and b of the power-law functions used in downscaling every coefficient in Equation (16).

Table 2: Coefficients of the power-law functions of the range and sill cosine-function parameters, for any time interval between 1 and 24 hours. Corresponding standard errors are also given. Taken from van de Beek *et al.* (2011b).

parameter	coefficient	a	b	ϵ_a	ϵ_b
Range	x_0	15.51	0.09	0.26	0.009
	A	2.06	-0.12	0.09	0.023
	t_0	7.37	0.22	0.98	0.067
Sill	x_0	0.82	-0.25	0.010	0.007
	A	0.20	-0.37	0.003	0.007
	t_0	162.00	-0.03	0.400	0.001

Table 3: Cosine functions parameters of the sill and range for the daily-fitted spherical semivariograms. Taken from van de Beek *et al.* (2011a).

parameter	$1/f$	A	t_0	x_0
Range	365	1.30	2.5	19.77
Sill	365	0.31	217.9	1.83

4 Methodology

Estimation of rainfall fields was carried out mainly in two parts: *1*) for daily time scales, *i.e.*, for 24-hour; *2*) for smaller time scales, *i.e.*, from 15-min to 12-hour. In both cases the Ordinary Kriging estimator was computed for the entire area of the Netherlands, 41848 km^2 [10].

4.1 Data

Three different data sets were used in the interpolation procedure:

❶ **Corrected radar data:** for every five minutes was used as validation data. The correction is made using the manual and automatic rain gauge network of KNMI (Royal Netherlands Meteorological Institute), Overeem (2009, p. 41). Because microwave link data is given in maximum and minimum power over 15-minute intervals, the 5-minute corrected radar data are accumulated to 15-minute data to be consistent and comparable to microwave link data. Corrected radar data is consistent throughout the time scales within 24-hours (or even larger scales), *i.e.*, adding up 96 15-minute intervals is equal to rainfall values for 24 hours, and so on for other smaller time scales. Each rainfall field generated by this data consists of 38063 pixels covering the entire land surface of the Netherlands. The spatial resolution is 1 km^2. This set contains data of a stratiform rainfall event from January 17$^{\text{th}}$ 8:00 UTC, to January 18$^{\text{th}}$ 8:00 UTC, 2011.

❷ **Simulated microwave link data:** (also referred in this work as *path-averaged data*) is obtained from the corrected 5-minute radar data. Path-averaged rainfall intensities were derived from the radar pixels covering each microwave link-path for each 5-minute step. From these data, 15-minute path-averaged radar rainfall intensities are derived. Simulated microwave link data is preferred above microwave link data.

[10] Data extracted from `http://en.wikipedia.org/wiki/Netherlands`.

The main scope is not to evaluate how accurate are the microwave link measurements but to evaluate an interpolation technique, or procedure, for the generation of accurate rainfall fields when compared to the real event; this latter, represented by the corrected radar data set. As was said before, simulated microwave link data is packed in 15-minute intervals; and the value of the 15-minute path-averaged radar rainfall intensity is assigned to the middle of the link. Therefore, 2032 simulated rainfall values represent the T-Mobile microwave link network of the Netherlands for that day. Assigning the value to the middle of the link seems to be the more logical procedure, because it is unknown how the rainfall is distributed along the path of the link.

❸ **Radar data in the middle of the link:** (also referred as *centered-link pixel data*) is used to contrast the differences in rainfall estimation, when using the pixel (rainfall value) falling right below the middle of the link, instead of averaging all the pixels falling under the path of the link. Like simulated microwave link data, there are 2032 values of rainfall representing the centers of all links in the network for each 15-minute interval.

4.2 Daily rainfall field estimation

In this part of the research, all the rainfall fields are estimated for 24-hour (daily) rainfall intensities. In both parts, sections 4.2 and 4.3, estimates are performed over the spatial grid of the corrected radar data set.

For the data sets described in section 4.1, spherical-semivariogram models are fitted to the corresponding empirical semivariograms. These fitted semivariograms are computed for short- and large-range models, ~ 5 km and ~ 150 km respectively. Short-range semivariograms refer to semivariograms fitted for a maximum radius of 16 km, *i.e.*, the largest possible lag between two measurements looked for; meanwhile for large-range semavariograms a maximum radius of 200 km was looked for. Spherical semivariograms following the *van de Beek* approach (van de Beek *et al.*, 2011a) are also computed. With all these semivariance models, rainfall field estimation is carried out applying Ordinary Kriging for every corresponding data set, path-averaged and centered-link pixel values. Rainfall field estimation with permutation between semivariogram models and data sets is also done.

Although short- and large-range semivariogram models are also fitted to all pixel values from the corrected radar data set, and used in OK estimates for the other two data sets; by its nature, there is no sense in trying to apply any interpolation technique to this set of values.

4.3 Sub-daily rainfall field estimation

Rainfall fields estimations are again carried out applying OK using this time downscaled semivariograms. The downscaling is done following the *van de Beek* approach (van de Beek *et al.*, 2011b) described in section 3.2.1 for both methods, Parameters Ratio and Cosine Functions. The semivariogram to which these two methods are applied to, is the large-range spherical model for the path-averaged data set. This data set is aggregated into five different time scales: 15-minute, 1-, 3-, 6- and 12-hour; and rainfall fields are estimated for every data subset within each time scale including the 24-hour scale, *e.g.*, 96 outcomes are obtained by OK for the 15-minute scale, 24 outcomes for the 1-hour scale and so forth.

Along the 24-hour rainfall event, interpolated negative values of rainfall are expected in places where no precipitation is registered. The definition of a threshold is needed in order to surpass this situation. Such a threshold is established independently for every interpolated rainfall field in the following way: *1*) for the time interval in which the interpolation is carried out, the percentage of all simulated microwave links in the network registering any precipitation is calculated; *2*) the interpolated rainfall field (38063 pixels) is sorted from large to small rainfall values; *3*) from the highest to the lowest rainfall value, the percentage is applied to this data subset[11]; and the values that have a corresponding cumulative frequency [%] lower than this percentage are said to be below the threshold. Although the threshold does not guarantee that all the values above this latter will be positive; it does allow to account for no rainfall areas in estimated rainfall fields.

Three different procedures are applied to the estimated rainfall fields for their posterior statistical analysis: *1*) no threshold is applied to the estimated rainfall field, *i.e.*, every possible negative value is used in the analysis; *2*) all the interpolated rainfall values falling below the threshold are set to zero, and used in the analysis; *3*) all the interpolated rainfall values falling below the threshold are dropt out of the data subset, and the remaining values are used for analysis. Having done this post-processing of the interpolated rainfall fields, scattergrams of interpolated rainfall values against corrected radar data are plotted for every time scale, for both downscaling methodologies. Pearson and Spearman correlation coefficients are computed from these data sets.

[11]The term subset makes reference to the fact that the 24-hour rainfall event, here studied, could be divided into n-subsets of the i-aggregated scale, *e.g.*, there are 8 subsets of path-averaged data in a 3-hour scale.

4.4 Metrics

The following metrics were used to asses the quality of the interpolated rainfall fields.

❶ **Mean Error *(ME)*:** gives the mean bias of the estimated rainfall fields for the time scale in consideration. If the estimator is not unbiased, the difference between the estimator and the true value is called the bias of the estimator (Montgomery and Runger, 1999, p. 265). The estimated rainfall value is said to be unbiased if it is equal to the actual value of rainfall, *i.e.*, $ME = 0$. For ME larger than 0, the averaged rainfall fields overestimate the actual field ("ground-truth"), and the opposite is true for ME smaller than 0.

❷ **Root Mean Square Error *(RMSE)*:** gives the accuracy of the estimated rainfall field regarding to the actual field. The $RMSE$ measures the departure of the estimated values from the actual ones; therefore, the closer this metric is to 0, the better is considered the estimator (Helsel and Hirsch, 1995, p. 358).

❸ **Variance Ratio *(VR)*:** gives the amount of variance in the estimated rainfall field that is explained by the Kriging variance. For VR values smaller than 1, the Kriging variance overestimates the true variance, and the opposite is true for VR larger than 1. The closer this metric is to 1, the better the Kriging. The variance measures the actual uncertainty of the interpolated values.

$$ME = \frac{1}{N \cdot P\left(\mathbf{u}_\alpha\right)} \cdot \sum_{i=1}^{N} \sum_{\alpha=1}^{P(\mathbf{u}_\alpha)} \left[\left[\hat{R}_{OK}(\mathbf{u}_\alpha) \right]_i - \left[R\left(\mathbf{u}_\alpha\right) \right]_i \right] \tag{17}$$

$$RMSE = \sqrt{\frac{1}{N \cdot P\left(\mathbf{u}_\alpha\right)} \cdot \sum_{i=1}^{N} \sum_{\alpha=1}^{P(\mathbf{u}_\alpha)} \left[\left[\hat{R}_{OK}(\mathbf{u}_\alpha) \right]_i - \left[R\left(\mathbf{u}_\alpha\right) \right]_i \right]^2} \tag{18}$$

$$VR = \frac{1}{N \cdot P\left(\mathbf{u}_\alpha\right)} \cdot \sum_{i=1}^{N} \sum_{\alpha=1}^{P(\mathbf{u}_\alpha)} \frac{\left[\left[\hat{R}_{OK}(\mathbf{u}_\alpha) \right]_i - \left[R\left(\mathbf{u}_\alpha\right) \right]_i \right]^2 - \left[ME \right]^2}{\left[\sigma^2_{OK}(\mathbf{u}_\alpha) \right]_i} \tag{19}$$

The above metrics are chosen as measures of bias, accuracy and variability of the estimated rainfall fields, Equations (17), (18) and (19) respectively; where $R\left(\mathbf{u}_\alpha\right)$ is the "ground-truth" value of rainfall[12], $\hat{R}_{OK}(\mathbf{u}_\alpha)$ the estimated rainfall value by Ordinary Kriging, $P\left(\mathbf{u}_\alpha\right)$ the total number of interpolated

[12] $R\left(\mathbf{u}_\alpha\right)$ a realization of the random variable $r\left(\mathbf{u}_\alpha\right)$, where $\mathbf{u}_\alpha$ is the vector of spatial coordinates (x, y) at the α^{th} location (Section 3.1.1).

locations for every estimated rainfall field [38063 pixels], N the total number of estimated rainfall fields per time scale analyzed, *e.g.*, 96 for the 15-minute aggregated scale, and $\sigma^2_{OK}(\mathbf{u}_\alpha)$ the variance obtained after applying Ordinary Kriging [Equation (15)]. The Mean is also calculated as an absolute parameter of reference when ME*s* and RMSE*s* are analyzed. These metrics are computed for every time scale within 24 hours, for both downscaling methodologies.

5 Results and Analysis

Analysis of the results are presented following the methodology described in section 4, for the two main parts in which this study is divided: daily and sub-daily rainfall field estimation.

Table 4: Parameters of spherical and nugget-effect semivariogram models, obtained when fitting their corresponding empirical semivariograms for the three different data sets, and the *van de Beek* methodology. LR stands for large-range semivariograms, SR stands for short-range semivariograms.

Data set	Parameter	Unit	LR	SR
Corrected Radar	Range	km	219.745	54.085
	Sill	mm^2	11.645	3.371
	Nugget	mm^2	0.065	–
Path Averaged	Range	km	147.970	4.767
	Sill	mm^2	12.785	2.506
	Nugget	mm^2	1.110	–
Centered-link pixel	Range	km	154.082	4.349
	Sill	mm^2	12.907	2.474
	Nugget	mm^2	1.178	–
van-de-Beek model	Range	km	195.372	–
	Sill	mm^2	5.535	–

5.1 Daily rainfall field estimation

For the three different sets of data described in section 4.1, short- and large-range empirical semivariograms were computed for 24-hour rainfall to account for local and regional variability. Total semivariograms models (a linear combination of nugget-effect and spherical semivariograms [Equation (10)])

were fitted to all these empirical semivariograms, except for the *van de Beek* model which is a fitted model itself. Parameters of the fitted models for each data set for short and larges ranges are presented in Table 4.

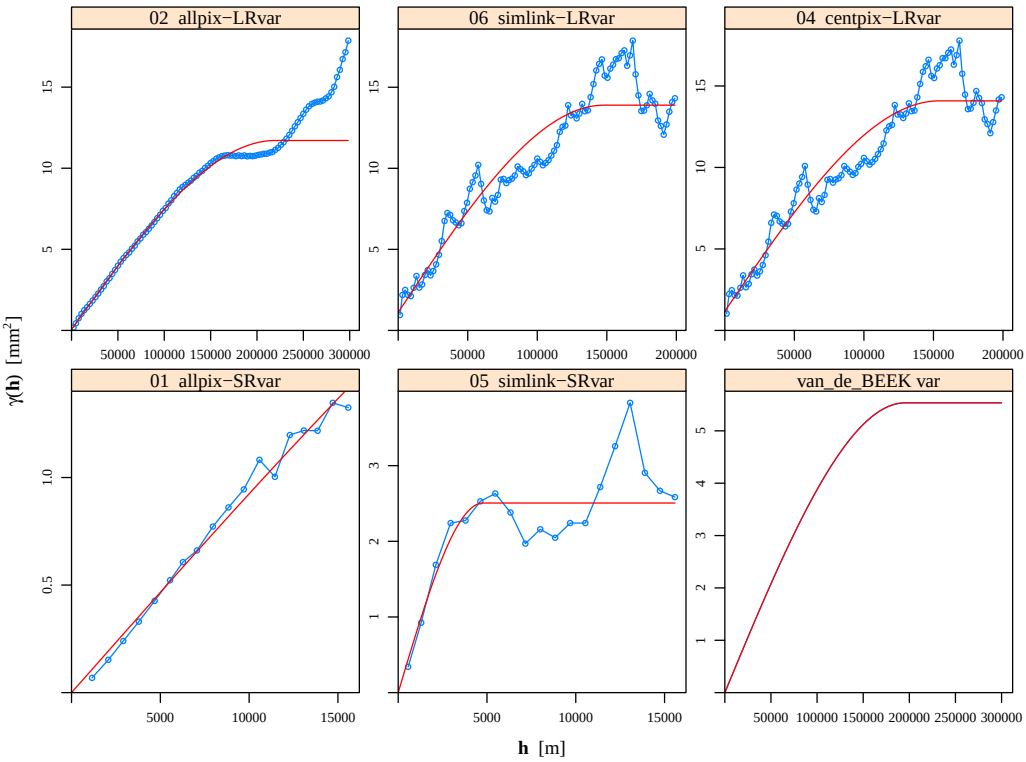

Figure 6: Empirical and fitted semivariogram models for a 24-hour rainfall event in January 17$^{\text{th}}$, 2011. The red line represents the fitted spherical model. The first part of the label indicates the data set used; the second part, whether it is a short- or large-semivariogram. Parameters for each fitted semivariogram are presented in Table 4.

Figure 6 shows six out of the seven possible models used in the present study. The large-range semivariogram for the corrected radar data set (Figure 6 top-left), represents the actual semivariogram for the rainfall event of January 17$^{\text{th}}$, 2011. In here, no nugget-effect model is fitted because of the large amount of radar pixels in the vicinity of the tail value. The "actual" semivariance increases again from approximately 220 km on, thus indicating some degree of weak correlation in areas located beyond this distance within the Netherlands, north and south-east regions (see Figure 7 left). The similarity between the path-averaged and centered-link pixel data sets (Figure 6 top- center and right); can be seen when comparing their corresponding empirical semivariograms. Because of this similarity, estimated rainfall fields

based on these two data sets are expected to have small differences among them. The "actual" and *van de Beek* (Figure 6 bottom-right) semivariogram fitted models have a range of about 200 km; but their sill values are rather different. The range of the path-averaged and centered-link pixel semivariograms are close to the "actual" range, around 50 km smaller, and with quite similar sills when the nugget-effect model is not taken into account. Large-range fitted semivariograms (LR) for path-averaged and centered-link pixel data sets show a nugget-effect. While for short-range fitted semivariograms (SR) (Figure 6 bottom-center), no nugget-effect is obtained.

5.1.1 Regional scale

For January 17$^{\text{th}}$ 2011, the regional spatial distribution of rainfall in the Netherlands, is well estimated when using LR-semivariograms as can be seen in Figure 7. This figure presents the "actual" rainfall field for January 17$^{\text{th}}$ 2011, obtained from the corrected radar data set, and two estimated rainfall fields based on the "actual" semivariogram and the *van de Beek* model, for the path-averaged data set. The interpolated fields are somewhat smoothed when compared to the radar image, which is inherent to optimal geostatistical interpolations. Nevertheless, the good results are obtained from these estimates (Figure 7 center and right) for three reasons: *a*) the characteristic of the rainfall event itself, because in this season of the year (January 17$^{\text{th}}$) stratiform rainfall events are predominant in the Netherlands; therefore, the estimation of uniform events over large areas implies low variability; *b*) the distribution of the microwave link network across the Netherlands, and its accurate simulated measurements (based on the corrected radar data set), allows to incorporate many more "points"[13] of information than any other rain-gauge network(s); *c*) the *van de Beek* methodology gives a very close estimate of the range parameter, when compared to the "actual" range, which is the key parameter in Equation (9).

In reality, there is no way to know beforehand how exactly the "actual" semivariogram of some current or future rainfall event is. Besides this, and for practical purposes, rainfall field estimations are based on simple fitted models, and although possible, finding the most accurate fit for an "actual" semivariogram is rather cumbersome (see Figure 6 top-left) and does not add accuracy to the interpolation process. What Figure 7 (center) presents, is an estimated rainfall field from a fitted spherical model of the "actual" empirical semivariogram. The *van de Beek* methodology yields and equally

[13] Although the word *points* is used to make reference to simulated microwave link measurements, this latter enclose linear distributions of the rainfall along their paths.

good approximation without any fitting procedure (right panel); therefore, is preferred here instead of using the "actual" empirical semivariogram.

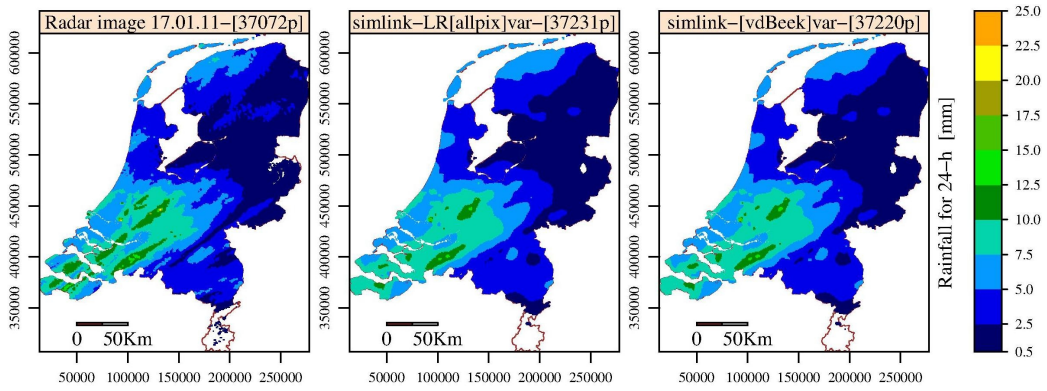

Figure 7: Corrected radar image for the 24-hour rainfall of January 17$^{\text{th}}$, 2011 (left). Estimated rainfall field for this event using a fitted spherical model of the "actual" empirical semivariogram (centre). Estimated rainfall field based on the semivariogram calculated from the *van de Beek* methodology (right). The first part of the label indicates the data set used for the interpolation; the second part indicates the semivariogram model; and the third part indicates the number of pixels with rainfall intensities larger than 0.5 mm.

When Figure 7 is compared to the microwave link network distribution presented in Figure 2, it is apparent that relative high rainfall depths are not well reproduced in areas with low microwave-links density, for example, in the south-west region. This issue becomes more important when using the fitted semivariogram models calculated from the centered-link pixel and path-averaged data sets, Figure 8 top- center and right. As expected, the results of both estimates are pretty similar, this due to the very similar semivariogram parameters they have. When the results from the fitted semivariograms are compared (Figure 8 top- center and right) to the corrected radar and the results from the *van de Beek* models (Figure 7 left and right), it can be seen that they are less accurate in that they give smaller areas where rainfall intensity is high or null. This decrease in accuracy can be related to the decrease in the range parameter these models present, around 31% smaller (Table 4), because a reduction in the semivariogram range implies a loss in correlation at larger lags. A drastic decrease in the range, *i.e.*, when semivariograms are computed for SR-semivariograms, makes it impossible to obtain accurate rainfall fields at regional scales. This can be seen by comparing Figure 8 (right- top and bottom) in which different range-semivariogram models, SR and LR, use the same data set, *i.e.*, path-averaged data set.

Examples of using different LR fitted semivariogram models for the in-

terpolation of the same data set (centered-link pixel data, Figure 8 center-top and right) were carried out as well as different data sets (centered-link and path-averaged) in which only one SR fitted semivariogram model was applied (Figure 8 bottom- left and right). The results obtained from this crossed estimation between data sets and semivariogram models, not only support the initial approximation of assigning the simulated microwave link rainfall value in the middle of the link but also indicate the low variability of the 24-h rainfall event along this path. For convective rainfall events, this low variability between path-averaged and centered-pixels is not expected.

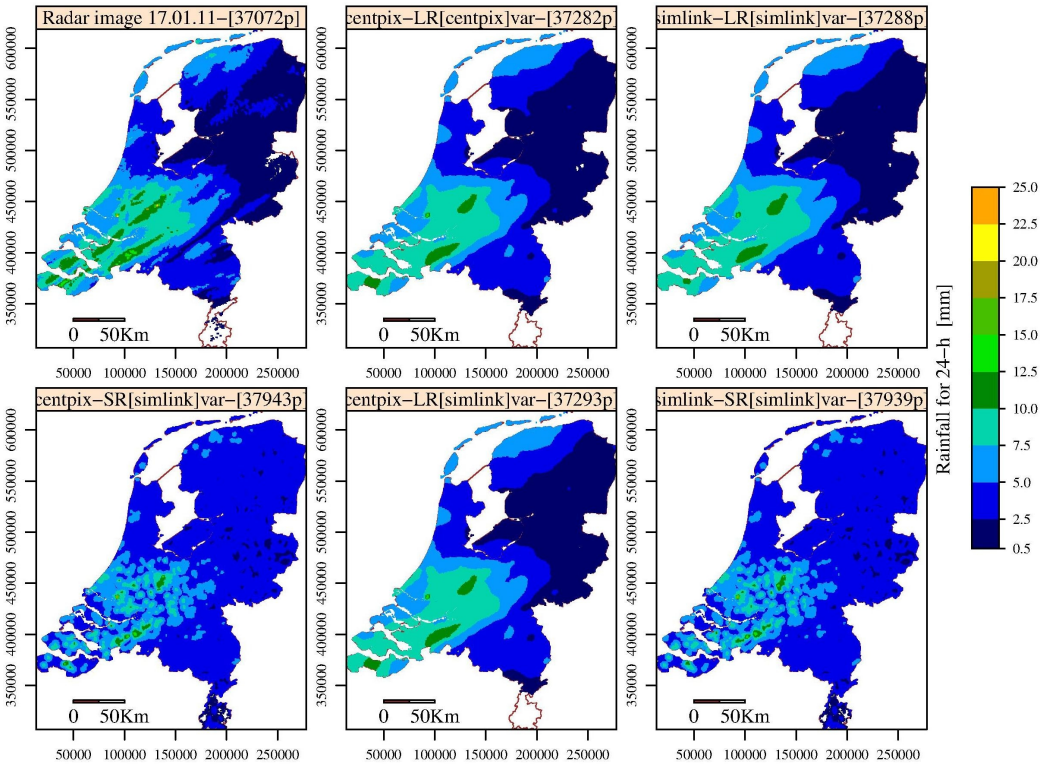

Figure 8: Corrected radar image for the 24-hour rainfall of January 17$^{\text{th}}$, 2011 (top-left). Estimated rainfall fields for the centered-link pixel data using two different LR semivariogram fitted models (center- top and bottom). Estimated rainfall fields for the simulated microwave link data using LR and SR semivariogram models (right- top and bottom respectively). Estimated rainfall field based on a SR semivariogram applied to centered-link pixel data (bottom-left). The first part of the label indicates the data set used; the second part indicates the semivariogram model applied; the third part indicates the number of pixels in the plot with rainfall intensities larger than 0.5 mm.

5.1.2 Local scale

LR fitted semivariogram models give relative good estimates for the spatial distribution of the rainfall event at local scales, as it is expected; nevertheless, the nature of the interpolation technique used here, OK, does not cope with areas where high values of rainfall are occurring, Figures 9 and 10 (top panels). Figures 9 and 10 present estimated rainfall fields at local scales, Utrecht and Rotterdam respectively, when SR and LR semivariogram models, for 24-h rainfall events, are applied to different data sets.

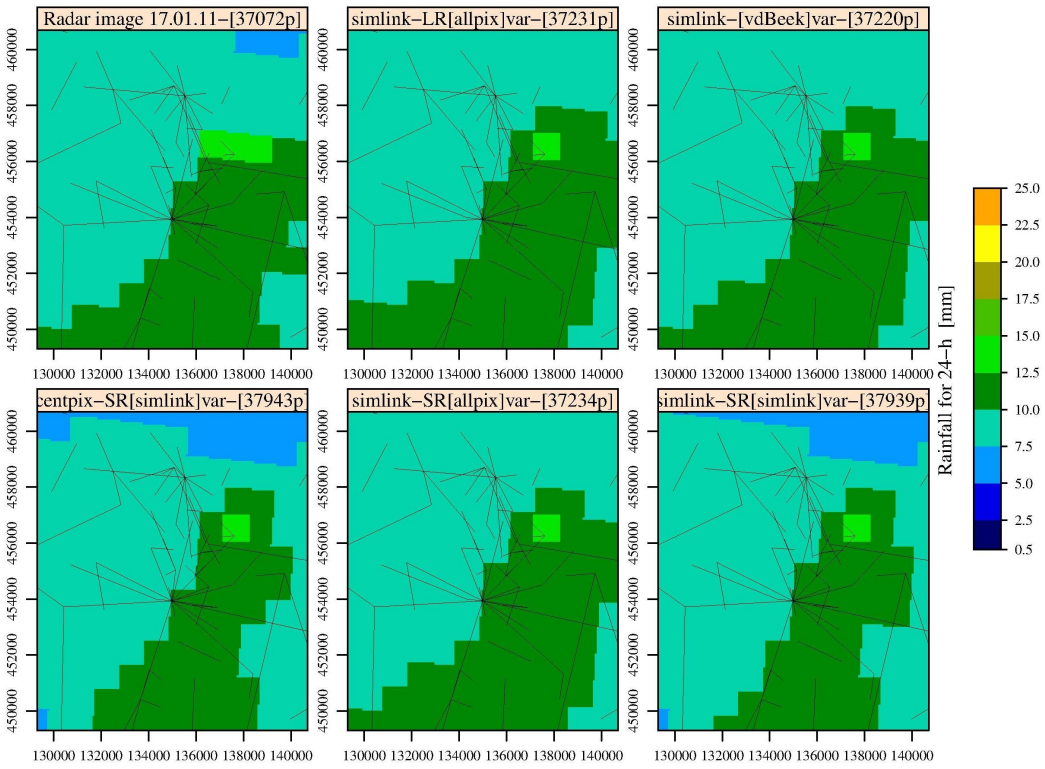

Figure 9: City of Utrecht. Local estimated rainfall fields for the 24-hour rainfall of January 17th, 2011. Corrected radar image (top-left). LR fitted semivariograms, "actual" and *van de Beek*, applied to path averaged data set (top- center and right, respectively). SR fitted semivariograms applied to centered-link pixel and path averaged data sets (bottom- left and right respectively). "actual"-SR fitted semivariogram applied to path averaged data set (bottom-center). For a detailed description of the labels, see Figure 7.

SR semivariograms used in rainfall field estimations give quite similar results in spatial variability, at very local scales, when compared to the "actual" rainfall event Figure 9 or 10 bottom- left and right. These good results in

the estimated rainfall fields are clearly visible for areas with high microwave link densities, *e.g.*, Rotterdam (Figure 10) and Utrecht (Figure 9). In areas with low microwave link densities, estimated rainfall fields underestimate the "actual" rainfall field. Relative high rainfall values were not reproduced either for SR models, when used in the OK technique. When compared to the "actual"-SR fitted semivariogram model, Figure 10 or 9 bottom-center, SR fitted models (bottom- left and right panels of these figures), yields similar results for areas with high microwave link densities, thus indicating the goodness of these SR fitted models for rainfall field estimation; at least for stratiform rainfall events.

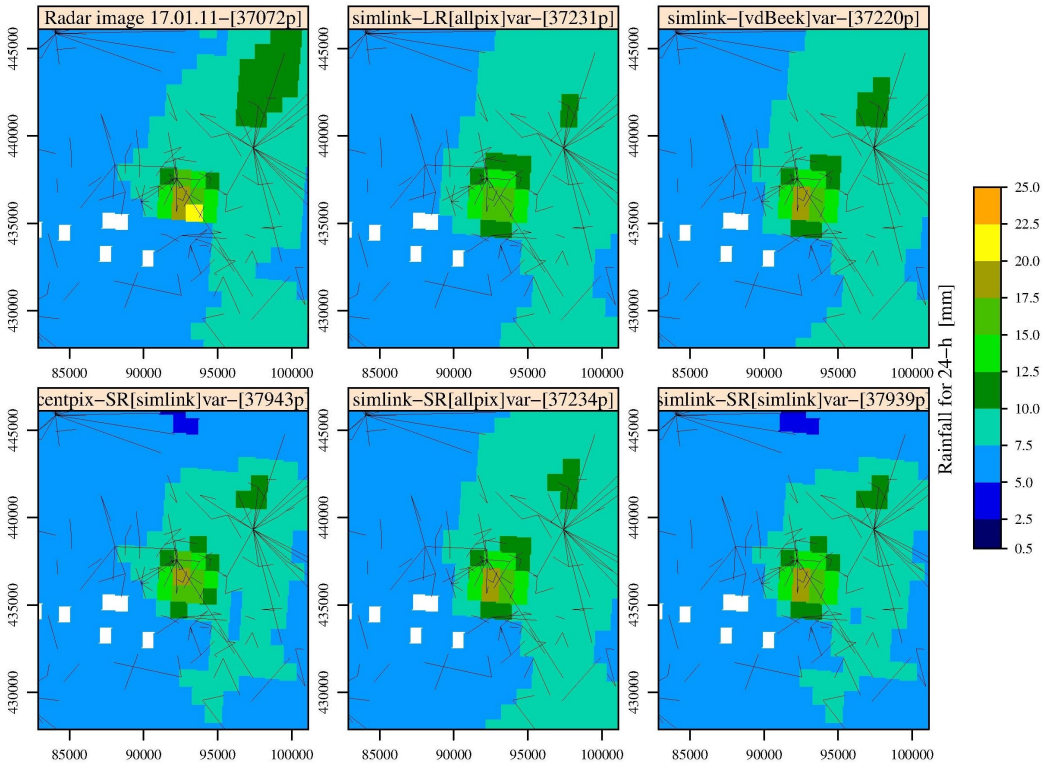

Figure 10: City of Rotterdam. Same as in Figure 9 but for Rotterdam.

Once more, the similarity in estimated rainfall fields obtained from crossed estimation between data sets and SR semivariogram models, proves the small differences among the path averaged and centered-link pixel data sets. Blank spaces in Figure 10 represent voids in the radar grid (*i.e.*, no radar estimates available) over which no interpolation was carried out. The grid for estimated rainfall fields, is given by the corrected radar data set; hence, its shifted (rotated) appearance.

5.2 Sub-daily rainfall field estimation

For 24-h rainfall events, either the *"van de Beek"* approach or LR fitted semivariogram models, not only give good results in rainfall field estimations but the fitting of the empirical semivariogram is pretty straightforward. This is because for such "large" aggregated time scales, and given the spatial characteristics of stratiform events, common in this season of the year in the Netherlands (January 17$^{\text{th}}$), enough precipitation is registered.

A large variability of rainfall, in time and space, is clearly present when rainfall is aggregated over small time scales, *i.e*, 15-minute or 1-hour. At these intervals, rainfall measurements are not as homogeneous as they are at larger aggregated intervals; and depending on the event, attenuation of the power due to rainfall is only registered by a few microwave links. When this happens, the fitting of the empirical semivariogram is not straightforward. Therefore, the task for automatic fittings of empirical semivariograms through only one model, turns out to be a cumbersome process. The downscaling parameterization developed in van de Beek *et al.* (2011b) presents a useful alternative for the semivariogram parameterization at small time scales.

The smallest interval in which van de Beek *et al.* (2011b) disaggregate rainfall events is 1 hour. The present work was based on the assumption that this downscaling methodology can also be applied to smaller intervals than 1 hour; in this case, 15 minutes. Table 5 presents the downscaled parameters for LR spherical semivariogram models, following the methodology described in section 3.2.1. None of the two 24-hour "downscaled" semivariograms for both methodologies, Parameters Ratio and Cosine Functions, were used for rainfall field estimation; instead, the LR semivariogram models used for the 24-hour scale were: *1*) the LR fitted semivariogram for 24-hour computed from the path-averaged data set; and *2*) the seasonal semivariogram obtained from the *van de Beek* model. These two models are the raw input for the two downscaling methodologies: the path-averaged LR fitted model for the Parameters Ratio methodology; and the *van de Beek* model for the Cosine Functions methodology. The parameters of these semivariograms, are not equal to their 24-hour "downscaled" counterparts. This is because the downscaling procedure for both methodologies (van de Beek *et al.*, 2011b), is based on fitted curves relating either the ratio between parameters or their seasonality, with their aggregated time scales (section 3.2.1).

The seasonal parameterization developed by van de Beek *et al.* (2011a) is intended for spherical semivariogram models only, and does not account for the nugget effect. This implies that less smoothed results are obtained when the fitted semivariogram model is applied. Although the *van de Beek*

methodology does not provide any downscaling procedure for the nugget-effect model. Nevertheless, the downscaling curve for the sill parameter used in the Parameters Ratio method (section 3.2.1), was also applied to the nugget-effect model, present on the LR fitted semivariogram model computed from the path-averaged data set. This data set was selected for all the estimations carried out in this part of the research, because for future implementations of automatic procedures for the generation of rainfall fields, this set represents the "raw" microwave link information on which these rainfall fields will be estimated.

Table 5: Downscaled parameters for LR spherical and nugget-effect semivariogram models, computed for the two methodologies proposed by van de Beek *et al.* (2011b). The parameters for the 24-hour semivariogram are the actual parameters obtained from fitting the empirical semivariogram (Parameters Ratio column), and from applying the *van de Beek* model (Cosine Functions column). Starting date: January 17th 08:00 UTC, 2011.

Time scale	Parameter	Unit	Parameters Ratio	Cosine Functions
15 minutes	Range	km	34.745	66.782
	Sill	mm^2	0.130	0.036
	Nugget	mm^2	0.014	–
1 hour	Range	km	55.666	94.688
	Sill	mm^2	0.564	0.190
	Nugget	mm^2	0.059	–
3 hours	Range	km	80.874	127.923
	Sill	mm^2	1.808	0.680
	Nugget	mm^2	0.189	–
6 hours	Range	km	102.367	156.084
	Sill	mm^2	3.769	1.496
	Nugget	mm^2	0.394	–
12 hours	Range	km	129.571	191.603
	Sill	mm^2	7.858	3.253
	Nugget	mm^2	0.822	–
24 hours*	Range	km	162.544	195.584
	Sill	mm^2	12.997	5.558
	Nugget	mm^2	1.359	–

Rainfall fields were estimated for each time scale, for each downscaled

method, using the downscaled semivariogram models summarized in Table 5. Figures 11 and 12 presents the scatter plots for the 15-minute time scale and the 24-hour aggregated time scale, respectively. The posterior analyses of estimated rainfall values, shows that the best fit between estimated and validated data, belongs to the data which is zeroed under a certain threshold. This is in accordance with the expected post-process for interpolated values; because not only negative rainfall values do not have any physical sense in reality but estimated rainfall fields, required as inputs in modeling, should also be able to reproduce null values of rainfall.

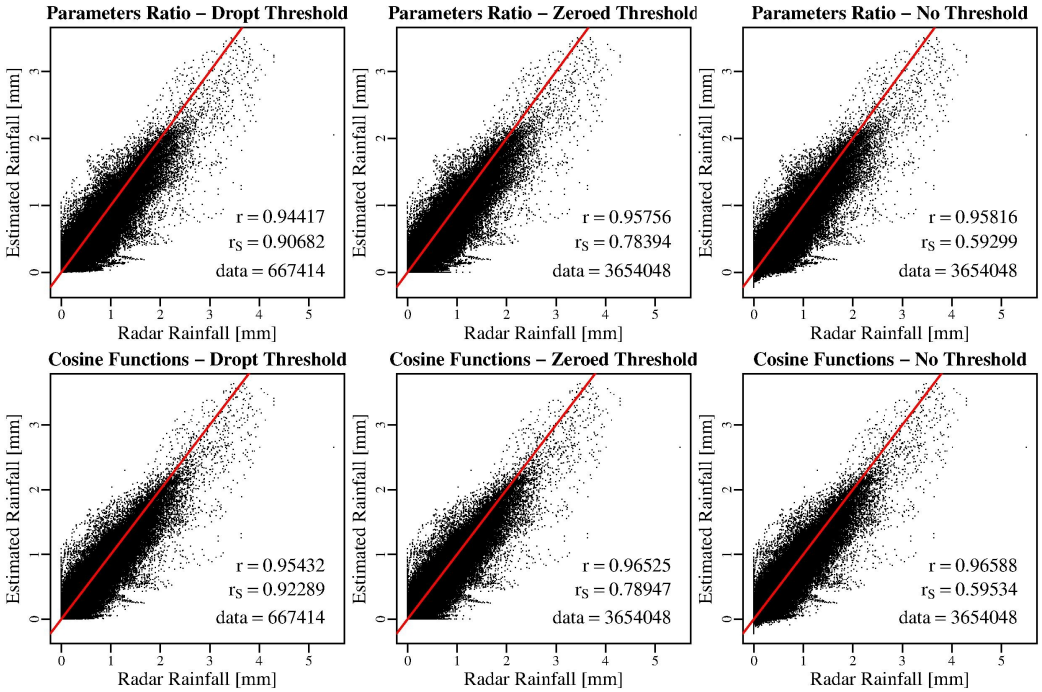

Figure 11: Scatter plots for the 15-minute aggregated time scale for the two downscaling methodologies (Parameters Ration and Cosine Functions) for the three different posterior analyses of estimated rainfall values (Dropt, Zeroed and No Threshold). The red line represents a 45° line with the origin in zero. The *data* value corresponds to the number of points used for computing the correlation coefficients.

From Figure 11 (left-panels) it can be seen how small or null values of measured rainfall influence the estimation procedure in such a way, that negative values of rainfall are more likely to be estimated at these small scales of rainfall aggregation. Relative high rainfall values are less well reproduced in the estimated rainfall fields. But this can be attributed to the fact that there is only one relative "extreme" rainfall value throughout the 96 15-minute

intervals configuring the daily event. If more relative "extreme" rainfall values were present in the data set, the estimated rainfall field will prone to reproduce this tendency. For all the outcomes, *i.e.*, all the three different posterior analyses, it can be seen that there is a slight tendency of underestimation in the interpolated values, specially for the relative large values of rainfall. When compared to large aggregated time scales, *e.g.*, 24 hours (Figure 12), the scatter plot exhibits a large variability around the 1:1 line, thus indicating the non-homogeneity of the rainfall event at such small time scales, despite the stratiform characteristics of the event under study.

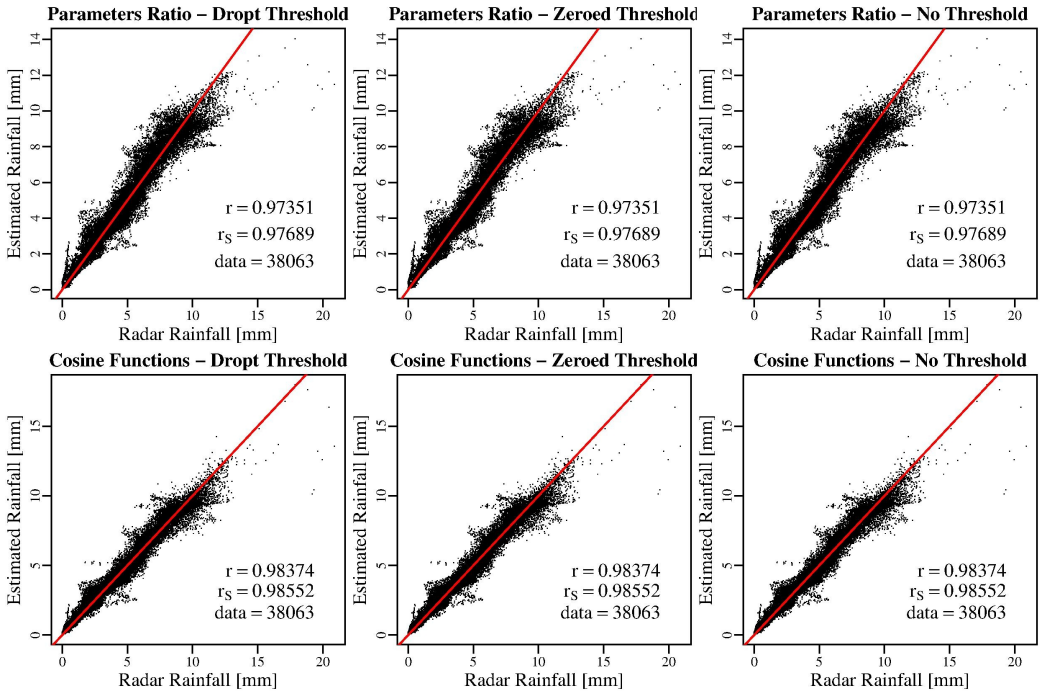

Figure 12: Same as in Figure 11 but for the 24-hour aggregated time scale.

In contrast to what was obtained for smaller aggregated time scales, the stratiform characteristics of the rainfall event, when viewed over larger aggregated time scales, drastically reduce the possibilities for negative values of estimated rainfall. The variability around the 1:1 line is also reduced for large aggregated time scales. For these scales, the interpolated values tend to be overestimated around the relative small rainfall values. Nevertheless, the tendency of underestimation towards the relative high rainfall values seen in Figure 11, is also observed in Figure 12. The few relative high values observed in this latter scatter plot, indicates once more the spatial uniformity of this stratiform rainfall event. For this aggregated time scale, the Pearson and Spearman correlation coefficients give equal results, along the three different

posterior analyses. This is because all the microwave links have non-zero registered rainfall. Therefore, in data sets with non-zero registered rainfall, there is no need for applying a threshold considering microwave links with no information.

Table 6: Non-normalized metrics for the three posterior analyses (No, Zeroed and Dropt Threshold) of estimated rainfall fields, interpolated for the two downscaling methodologies (Parameters Ratio and Cosine Functions) for six different aggregated time scales (15-minute, 1-, 3-, 6-, 12- and 24-hour).

Scale [hours]	Metric [mm]	Parameters Ratio			Cosine Functions		
		No Threshold	Zeroed Threshold	Dropt Threshold	No Threshold	Zeroed Threshold	Dropt Threshold
0.25	$\bar{R}$	0.0446	0.0446	0.2380	0.0446	0.0446	0.2384
	$\bar{R}_{OK}$	0.0448	0.0435	0.2382	0.0447	0.0439	0.2404
	M.E.	0.0003	-0.0011	0.0002	0.0001	-0.0007	0.0020
	R.M.S.E.	0.0504	0.0508	0.1160	0.0454	0.0458	0.1044
	V.R.	0.0603	0.0612	0.3206	0.6616	0.6815	3.5277
1	$\bar{R}$	0.1783	0.1783	0.6417	0.1783	0.1783	0.6422
	$\bar{R}_{OK}$	0.1794	0.1779	0.6458	0.1790	0.1780	0.6462
	M.E.	0.0012	-0.0003	0.0041	0.0007	-0.0002	0.0040
	R.M.S.E.	0.1075	0.1078	0.2031	0.0927	0.0930	0.1751
	V.R.	0.0813	0.0817	0.2917	0.7614	0.7708	2.7368
3	$\bar{R}$	0.5348	0.5348	1.2226	0.5348	0.5348	1.2227
	$\bar{R}_{OK}$	0.5391	0.5379	1.2329	0.5370	0.5364	1.2293
	M.E.	0.0043	0.0031	0.0103	0.0023	0.0016	0.0066
	R.M.S.E.	0.1976	0.1977	0.2988	0.1678	0.1679	0.2537
	V.R.	0.1039	0.1040	0.2374	0.9702	0.9721	2.2176
6	$\bar{R}$	1.0696	1.0696	1.6503	1.0696	1.0696	1.6505
	$\bar{R}_{OK}$	1.0788	1.0791	1.6658	1.0745	1.0750	1.6595
	M.E.	0.0092	0.0095	0.0156	0.0049	0.0054	0.0090
	R.M.S.E.	0.2937	0.2936	0.3646	0.2432	0.2431	0.3019
	V.R.	0.1224	0.1224	0.1886	1.1752	1.1747	1.8112
12	$\bar{R}$	2.1392	2.1392	2.1989	2.1392	2.1392	2.1991
	$\bar{R}_{OK}$	2.1603	2.1609	2.2215	2.1481	2.1484	2.2087
	M.E.	0.0212	0.0217	0.0227	0.0089	0.0092	0.0096
	R.M.S.E.	0.4303	0.4303	0.4362	0.3436	0.3436	0.3484
	V.R.	0.1362	0.1362	0.1399	1.3408	1.3407	1.3780
24	$\bar{R}$	4.2783	4.2783	4.2783	4.2783	4.2783	4.2783
	$\bar{R}_{OK}$	4.3236	4.3236	4.3236	4.2963	4.2963	4.2963
	M.E.	0.0453	0.0453	0.0453	0.0180	0.0180	0.0180
	R.M.S.E.	0.6317	0.6317	0.6317	0.4955	0.4955	0.4955
	V.R.	0.1898	0.1898	0.1898	1.6535	1.6535	1.6535

Overall, the methodology of Cosine Functions yields slightly better results for Pearson and Spearman correlation coefficients than the Parameters Ratio methodology. For all aggregated scales, the Pearson correlation coefficient increases from the Dropt Threshold case to the No Threshold case. This implies that when all estimates are taken into account, even the negative ones, these values are so small or close to zero, that the tendency for linear correspondence between corrected radar data and estimations, increases. The opposite is the case for the Spearman correlation coefficient[14], which is small when negative measurements are taken into account, and starts to increase when this measures are turned into zero values, and reaches its maximum values when these zeroed values are dropt out of the correlation. All the correlation coefficients tend to improve for larger aggregated time scales; but after reaching maximum values at 6-hour aggregated scales, these coefficients start to decrease again. This could mean that 6-hour is the aggregated scale where rainfall is more homogeneous in time and space, among the four intervals for the 24-hour event. This assumption was not tested in the present study though.

The highest Pearson correlation coefficient was found to be 0.99266 for the 6-hour aggregated scale for the Cosine Function methodology for the Zeroed Threshold case; meanwhile 0.94417 was the lowest Pearson coefficient for the 15-minute aggregated scale for the Parameters Ratio methodology for the Dropt Threshold case. Although this "lowest" correlation coefficient has a high value, close to 1, the fact that it was obtained for the smallest aggregated time scale, 15-minute, indicates a lower accuracy of the estimated rainfall field when compared against larger aggregated time scales. This is due to the absence of microwave link data between 15-minutes intervals, and the large variability of rainfall in space, despite the stratiform characteristics of the event. For the Spearman correlation coefficient, the highest value was found to be 0.99154 for the 12-hour aggreggated scale for the Cosine Function methodology for the Dropt Threshold case; meanwhile 0.59299 was the lowest coefficient for the 15-min aggregated scale for the Parameters Ratio methodology for the No Threshold case (Figure 11). This lower correlation can be attributed to the fact that the spatial scale of rainfall variability is smallest at the 15-minute time scale. As the scale of spatial rainfall variability approaches the typical scale of the separation of microwave links within the network, it is expected that the quality of rainfall retrievals decreases. The difference between Spearman coefficients along the three different Threshold cases is very small for 12-hour aggregated scales onwards; therefore, it can be

[14]Spearman correlation coefficient is a measure of how monotonic the relation between two variables is, in this case, corrected radar and estimated rainfall values.

said that for these larger scales and for the Zeroed Threshold case, there is an increasing-monotonic-linear relation between the estimated and the corrected radar rainfall values. The Zeroed Threshold case is recommended in here, for future implementations, because estimated rainfall fields must be able to represent areas where precipitation is equal to zero instead of some negative value (not realistic representation of the event) or no value at all.

OK not only returns the estimated value but also its estimate variance. If the bias (squared) is subtracted from the squared error and then divided by the OK variance, a ratio of the variance for every estimated rainfall value can be established [Equation (19)].

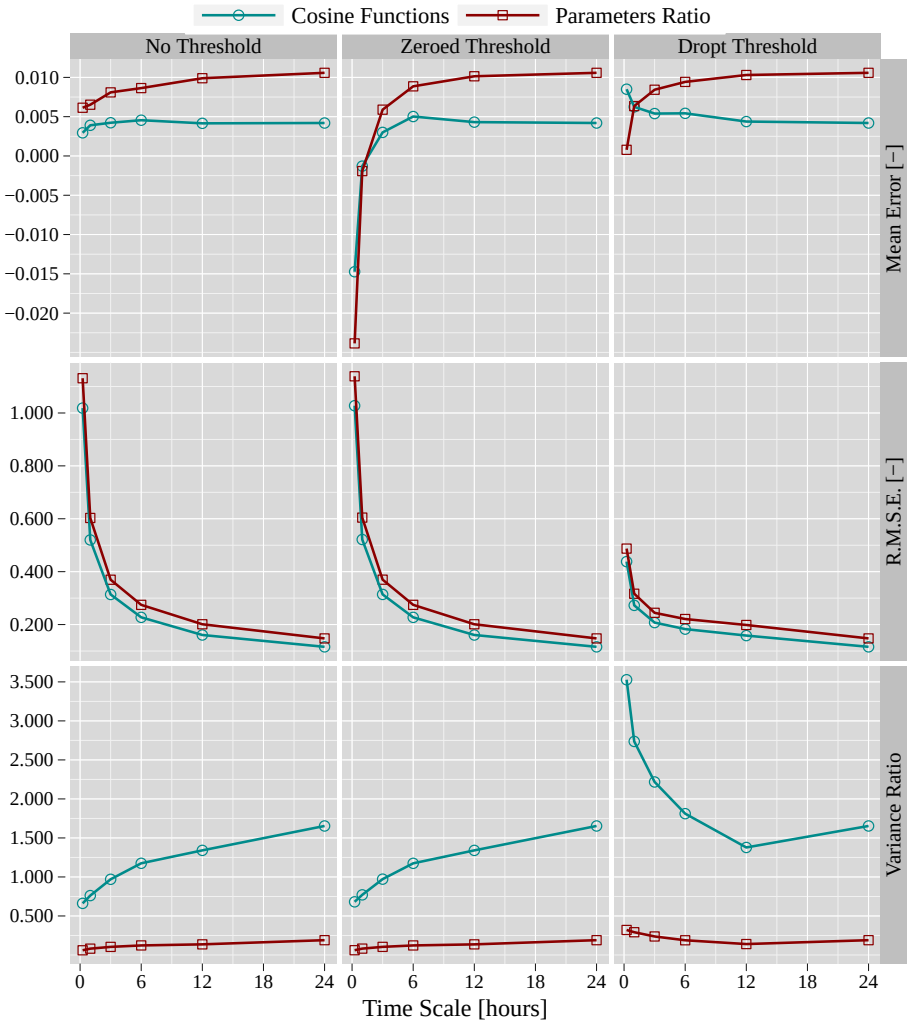

Figure 13: Normalized curves for the Mean Error (M.R.), Root Square Mean Error (R.S.M.E.) and Variance Ratio (V.R.), for each downscaling methodology, for each posterior analysis.

Table 6 presents the non-"normalized" values of the estimated metrics (mean, mean error, root mean square error and variance ratio) for both methodologies of downscaling, for each posterior analysis. Figure 13 presents the "normalized" curves for these metrics. The Variance Ratio is also presented in this figure, but given that this is already a relative measure, there is no need (nor physical meaning) for normalization. This metric is also plotted in log-normal scale, independently from the others two, in Figure 14. The normalization was carried out by dividing these parameters over the corresponding corrected radar mean value.

From Figure 13, it can be seen that the bias (ME) in the interpolated fields is very small (generally less than 1%). The decrease of the $RMSE$ while increasing temporal aggregation is also evident. This is due to the fact that there is much more variability at the 15-minute time scale. The effect of the difference between the two different semivariograms becomes apparent in Figure 13 (bottom row) and Figure 14. The Parameters Ratio methodology clearly overestimates the variance for all time scales, whereas this depends on the aggregation time scale for the Cosine Functions methodology. The VR shows slightly different behaviour when only the values above a threshold are considered. This indicates that for the part of the rainfall field that is hydrologically most interesting, *i.e.*, the part where there is significant rain, the choice of an optimal semivariogram could be different than when considering the entire field (including the zeroed threshold).

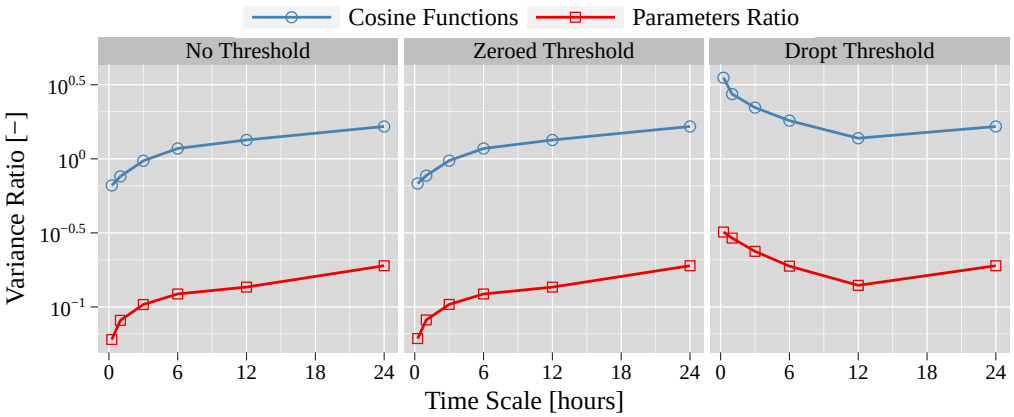

Figure 14: Normalized curves in log-normal scale, for the Variance Ratio (V.R.), for each downscaling methodology, for each posterior analysis.

6 Conclusions and Recommendations

This study showed that rainfall maps can be obtained from interpolated rainfall estimates of commercial microwave links.

For daily rainfall field estimation, large- and short-range fitted semivariograms were computed from three different data sets: corrected radar data, path-averaged data and centered-link pixel data. Also, the approach developed by van de Beek *et al.* (2011a) to obtain semivariograms for a given day in the year, was used. When used in Ordinary Kriging (OK), large-range (LR) spherical semivariogram models gave good estimates at regional and local scales for 24-hour stratiform rainfall events. These LR-semivariogram models, could be estimated either using microwave link information for the given day or the seasonal model proposed and developed by van de Beek *et al.* (2011a,b). In both cases the results were good and rather similar. The goodness of the LR-models, computed from simulated microwave link information, can be attributed to the microwave link network widely distributed across the Netherlands. Small difference were observed between estimated rainfall fields based on path-averaged and centered-link pixel data, mainly due to the close semivariogram parameters they share. Estimates based on the *van de Beek* methodology yielded accurate results, despite the fact that a predefined semivariogram was used. Short-Range (SR) spherical semivariogram models gave accurate estimations, but only at local scales in areas where local microwave link density is high. Higher values of rainfall were not completely reproduced by the OK statistical interpolation technique, *i.e.*, underestimated rainfall values were always obtained in areas where local-extreme values of rainfall are registered.

For sub-daily rainfall field estimation, 15-minute simulated microwave link data was aggregated into scales of 1-, 3-, 6- 12- and 24-hour. For each time scale, downscaled semivariograms were used for the generation of interpolated rainfall fields. The two downscaling methodologies applied to the path-averaged data set were developed previously by van de Beek *et al.* (2011b). The Parameters Ratio methodology was based on the 24-hour fitted semivariogram computed from the path-averaged data set; whereas the Cosine Functions methodology was based on the *van de Beek* seasonal model for January 17$^{\text{th}}$. Pearson and Spearman correlation coefficients, and metrics of bias, accuracy and variance reproduction were computed for all the aggregated time scales. Considering the bias, accuracy and the correlation coefficients between estimated and observed values, best results were obtained overall for the Cosine Functions downscaling methodology and post-processing the estimated fields by setting negative estimates to zero (Zeroed Threshold posterior analysis). The small ME and the substantially smaller

$RMSE$, for increasing aggregated time scales, reflect the good performance and higher accuracy in rainfall field estimates at larger aggregated time scales. When the Parameters Ratio methodology was applied to the semivariogram downscaling procedure, for all the aggregated time scales there was an overestimation in the variance of the estimated rainfall fields. When the downscaling was done by the Cosine Functions methodology, this overestimation in the variance depended on the aggregated time scale, but was generally much smaller.

In conclusion, this study showed that the Cosine Functions methodology applied to zeroed threshold posterior analyses of rainfall data, is the suggested future operational method for rainfall field estimation from link data. Not only because of its higher correlation coefficients when compared to the Parameters Ratio methodology, as was demonstrated in the present study, but also because its used semivariogram does not depend directly on the registered data itself but on the seasonal model of *van de Beek* (van de Beek *et al.*, 2011a). Hence, for dryer periods when no enough information is collected to compute a daily semivariogram, the method can be still applied to these scarce data sets.

In this study, simulated microwave link values were placed in the middle of the path for all the links within the network. For further analyses, these values could be placed at the extreme sides of the link, in order to test the influence of the path and its direction in rainfall field estimations at regional scales. However, given the directional homogeneity of the network (Figure 1), no large changes in estimated rainfall fields are expected for the Netherlands.

A different method could be applied for finding the threshold over which non-zero rainfall values are taken into account. The method developed in this study does not guarantee that for all the cases only positive values are taken into account. Even though in few cases, very small values of estimated negative rainfall were used in the different posterior analyses; in most of the cases, the "surplus" of the threshold computation used in the present work, zeroed (turned into zero) small positive values of estimated rainfall. The same procedure should also be applied for data registered around summer, where the characteristics of rainfall will differ. These will be less homogeneous in space and more intense (convective rainfall); hence, a re-validation or disregarding of the method here proposed can be the result.

It should be noted that the radar measurements account for large volumes in space, but yet, these are used in simulating path-averaged rainfall. This makes radar data not entirely representative of link data, which are real path measurements. Besides this fact, the radar data resolution used in this study, 1 km^2, is less suitable to simulate path-averaged rainfall intensities at smaller spatial scales, *e.g.*, < 1 km^2. Therefore, the use of x-band radar is highly

recommended to simulate link measurements at such smaller scales.

References

Atlas, D. and C. W. Ulbrich, 1977: Path- and area-integrated rainfall measurement by microwave attenuation in the 1-3 cm band. *J. Appl. Meteor.*, **16**, 1322–1331.

Bierkens, M. F. P., 2010: Stochastic Hydrology, *lecture notes*. Utrecht University.

Creutin, J., G. Delrieu, and T. Lebel, 1988: Rain measurement by raingage-radar combination: A geostatistical approach. *J. Atmos. Ocean. Technol.*, **5**, 102–115.

Cuccoli, F., L. Baldini, L. Facheris, S. Gori, and E. Gorgucci, 2011: Tomography applied to radiobase network for real time estimation of the rainfall rate fields, *Submitted*.

Foufoula-Geourgiou, E. and W. Krajewski, 1995: Recent advances in rainfall modeling, estimation, and forecasting. *Rev. Geophys.*, **33**, 1125–1137, *Supplement*.

Gandin, L. S., 1970: The planning of meteorological station networks. Technical Note No. 111 WMO No. 265, Secretariat of the World Meteorological Organization, Geneva, 35 pp.

Goldshtein, O., H. Messer, F. IEEE, and A. Zinevich, 2009: Rain rate estimation using measurements from commercial telecommunications links. *IEEE Trans. Signal Process.*, **57**, 1616–1625.

Goovaerts, P., 1997: *Geostatistics for Natural Resources Evaluation*. Applied Geostatistics, Oxford University Press, New York, 496 pp.

Helsel, D. R. and R. Hirsch, 1995: *Statistical Methods in Water Resources*. Elsevier Science B.V., Amsterdam, 548 pp.

Hengl, T., 2009: *A Practical Guide to Geostatistical Mapping*, 2nd edition. University of Amsterdam, Amsterdam, 291 pp.
URL `http://spatial-analyst.net/book/`

Hunter, S. M., 2009: WSR-88D radar rainfall estimation: Capabilities, limitations and potential improvements. Technical report, National Weather

Service Weather Forecast Office - Morristown, TN.
URL http://www.srh.noaa.gov/mrx/research/precip/precip.php

Icom America Inc., 2011: ID-1 1200 MHz Digital Transceiver.
URL http://www.icomamerica.com/en/products/amateur/dstar/id1/default.aspx

Journel, A. G., 1989: *Fundamentals of Geostatistics in Five Lessons*, volume 8 of *Short Course in Geology*. American Geophysical Union, Washington, 46 pp.

King, A., J. Billingham, and S. Otto, 2003: *Differential Equations - Linear, Nonlinear, Ordinary, Partial*. Cambridge University Press, Cambridge, 548 pp.

Krige, D. G., 1951: A statistical approach to some basic mine valuation problems on the witwatersrand. *J. of the Chem., Metal. and Mining Soc. of South Africa*, **52**, 119–139.

Lanza, L., J. Ramírez, and E. Todini, 2001: Stochastic rainfall interpolation and downscaling. *Hydrol. Earth Syst. Sci.*, **5**, 139–143.

Leijnse, H., R. Uijlenhoet, and J. Strickter, 2007a: Hydrometeorological application of a microwave link: 2. precipitation. *Water Resour. Res.*, **43**.

— 2007b: Rainfall measurement using radio links from cellular communication networks. *Water Resour. Res.*, **43**.

Matheron, G., 1971: *The Theory of Regionalized Variables and Its Applications*. École Nationale Suprieure des Mines de Paris, Fontainebleau, France, 218 pp., Cahiers du Centre de Morphologie Mathematique, No. 5.

Messer, H., 2007: Rainfall monitoring using cellular networks. *IEEE Signal Process. Mag.*, **24**, 144–142, doi: 10.1109/MSP.2007.361621.

Messer, H., A. Zinevich, and P. Alpert, 2006: Environmental monitoring by wireless communication networks. *Science*, **312**, 713.

Michaelides, S., V. Levizzani, E. Anagnostou, P. Bauer, T. Kasparis, and J. Lane, 2009: Precipitation: Measurement, remote sensing, climatology and modeling. *Atmos. Res.*, **94**, 512533.

Montgomery, D. C. and G. C. Runger, 1999: *Applied Statistics and Probability for Engineers*, 2nd edition. John Wiley & Sons, Inc., 944 pp.

Muller, C. and C. Kidd, 2006: Quantitative rainfall measurements: A comparison of micro rain radars (MRRs) and rain gauges. *Proceedings of the EGU General Assembly*, EGU, Vienna, 2nd-7th April 2006.

Overeem, A., 2009: *Climatology of extreme rainfall from rain gauges and weather radar*. Ph.D. thesis, Wageningen University.

Overeem, A., H. Leijnse, and R. Uijlenhoet, 2011: Measuring urban rainfall using microwave links from commercial cellular communication networks. *Water Resour. Res.*, **47**, doi: 10.1029/2010WR010350, in press.

Pebesma, E. J., 2001: *gstat user's manual*. Utrecht University, Utrecht.

— 2011: *geostatistical modelling, prediction and simulation - Package 'gstat'*. 0.9-76 edition, R package.

Sauvageot, H., 1994: Rainfall measurement by radar: a review. *Atmos. Res.*, **35**, 27–54.

Schuurmans, J. M., M. F. P. Bierkens, and E. J. Pebesma, 2007: Automatic prediction of high-resolution daily rainfall fields for multiple extents: The potential of operational radar. *J. Hydrometeor.*, **8**, 1204–1224.

Upton, G., A. Holt, R. Cummings, A. Rahimi, and J. Goddard, 2005: Microwave links: The future for urban rainfall measurement? *Atmos. Res.*, **77**, 300–312.

van de Beek, C. Z., H. Leijnse, P. J. J. F. Torfs, and R. Uijlenhoet, 2011a: Climatology of daily rainfall semi-variance in the netherlands. *Hydrol. Earth Syst. Sci.*, **15**, 171–183.

— 2011b: Seasonal semi-variance of dutch rainfall at hourly to daily scales, *Submitted*.

Zinevich, A., P. Alpert, and H. Messer, 2008: Estimation of rainfall fields using commercial microwave communication networks of variable density. *Adv. Water Resour.*, **31**, 1470–1480.