Semivariance of rainfall in The Netherlands

C. Z. van de Beek[1], H. Leijnse[1], P. J. J. F. Torfs[1], and R. Uijlenhoet[1]

[1] *Hydrology and Quantitative Water Management Group, Department of Environmental Sciences, P.O. Box 47, 6700AA, Wageningen University, Wageningen, The Netherlands, remco.vandebeek@wur.nl*

ABSTRACT

This paper shows a method for estimating the accuracy of rainfall maps created from raingauge data for hydrological modeling purposes. Using a 14 year dataset of daily rainfall gathered by 33 automatic KNMI stations the seasonal changes in semi-variance were estimated. From this semi-climatological variance areal rainfall maps can be created by kriging. The resulting kriging variance can be used to find how reliable hydrological model output is based on the rainfall data.

1. INTRODUCTION

Rain gauges can offer high quality rainfall measurements at their location [1, 2, 3]. Networks of rain gauges can offer better insight into the space-time variability of rainfall, but they tend to be too widely spaced for accurate estimates between points. While other remote sensing systems, like radar and microwave links, exist for rain measurements that offer good insight in the spatial variability on rainfall they tend to have more problems in identifying the correct rain amounts at the ground [4, 5]. A way to estimate the variability of rainfall between gauge points is to interpolate between them using fitted variograms [6]. Using such variograms to find the areal rainfall by applying kriging to gauge data the accuracy of an hydrological model can estimated. However, if a dense rain gauge network is lacking it is difficult to create accurate variograms. Using a dataset of 14-year daily rain accumulation gathered at 35 automatic weather stations operated by KNMI and a one-year data of 30 gauges in a dense network in a radius of 10 km around CESAR (Cabauw Experimental Site for Atmospheric Research), this article shows a highly seasonal variation of semi-variograms in the Netherlands and offers a way for estimating rainfall in-between rain gauge locations using the sill and range found with a fitted climatological spherical variogram for applications in (urban) catchment hydrology. Variograms at short range during winter and spring tend be underestimated, but summer and autumn are well predicted.

2. AREA AND INSTRUMENTS

Data from 35 automatic KNMI stations between January 1, 1995 and December 31, 2008 were considered (Left panel Fig. 1). During these 14 years, 29 of these stations operated for most of the time and therefore only these stations were used for a semi-climatological

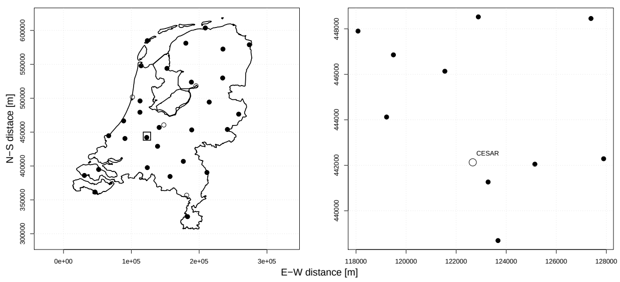

Figure 1. left panel: Station Locations of the used KNMI measurement stations. The square near the centre of the Netherlands is a detail of the dense gauge network shown in the right panel.

analysis. One-day accumulations from these data were used for this study. These data offer a good way evaluate larger scale variation of rainfall. Between March 2004 and March 2005 33 KNMI stations had rain data available and were used for a detailed one-year study.

The second data-set was gathered with a dense network of 30 tipping-bucket rain gauges around meteorological measurement site CESAR (Cabauw Experimental Site for Atmospheric Research) and were operated by a joined venture between University of Utrecht and Wageningen University. The gauges had a volumetric resolution of 0.2 mm and a time resolution of 0.5 sec and were placed within a 10 km radius around CESAR . Of these 30 gauges 10 were selected (right panel of Fig. 1) as these operated well and continuously between March 2004 and March 2005. The data was converted to one day accumulations for this study and was used for estimating the short range rainfall variation of the detailed one-year study. Discussion of this one-year study goes beyond the scope of this paper.

3. THEORY

A standard method for evaluating rainfall variability between rain gauges is to estimate variograms. Assuming stationary, isotropic data the experimental semi-variogram can be found by taking half the average of the squared difference between data pairs at equal dis-

© Proceedings of the 8th International Symposium on Tropospheric Profiling, ISBN 978-90-6960-233-2
Delft, The Netherlands, October 2009. Editors, A. Apituley, H.W.J. Russchenberg, W.A.A. Monna

tance [7]:

$$\hat{\gamma}(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} \left(z(x_i + h) - z(x_i)\right)^2 \qquad (1)$$

Here for $i = 1, 2, \ldots, n(h)$, x_i is the location and $x_i + h$ the locations at distance h from location x_i. For a data-set of measurements at n locations this means there are $N(N-1)/2$ data pairs whose separation distance can be found.

As the empirical variogram values will not offer values for each distance h one of several possible models has to be fitted to estimate these variances. While many types, like exponential, Gaussian, logarithmic and others exist [8] it was chosen to take a simple spherical variogram as this model adequately fits the variogram values while only needing few parameters:

$$\gamma(h) = \begin{cases} c_0 + c_1 \left(\frac{3}{2}\frac{h}{a} - \frac{1}{2}\left(\frac{h}{a}\right)^3\right) & \text{if} \quad h \leq a \\ c_0 + c_1 & \text{if} \quad h > a \end{cases} \qquad (2)$$

Here c_0 is the nugget and is the variance at zero distance, c_1 is the sill and is the maximum value of the fitted semi-variance function. Finally a is the range where the significant correlation between data pairs becomes zero and moves to random variations. As the variance found for two raingauges at 8 m distance was found to be near zero it was assumed that the nugget was negligible making it possible to remove the c_0 term from Eq. 2.

The data was analyzed by estimating the daily omni-directional variograms and then average the variograms using a 90-day moving window. Finally the data was binned over a distance of 5 km for both faster fitting of the spherical variogram and easier to interpret figures. With the furthest gauge pair in the data-set at 315 km it was chosen to set the maximal range at 300 km. Cases where the range is not reached before the maximal distance occur mostly around October and early November when the variogram data tends to be nearly linear over the full length.

In Section 4 pseudo-climatological data will be assessed to find the daily trend in the sill and range. To fit a function to this trend a spectral analysis is applied.

A simple time-series of a cosine function could be expressed as:

$$x_t = A\cos(2\pi\omega t + \phi) \qquad (3)$$

Where A is the amplitude, ω is an index defining the number of oscillations per time and ϕ the phase, defining the start location of the cosine function.

Using a standard trigonometry rule Equation (3) can be expressed in a linear form. This allows the problem to be solved using linear regression.

$$x_t = U_1\cos(2\pi\omega t) + U_2\sin(2\pi\omega t) \qquad (4)$$

Where $U_1 = A\cos(\phi)$ and $U_2 = -A\sin(\phi)$.

To find a signal in the fitted variogram parameters it is necessary to average over an optimal range of days to avoid the noise of to day-to-day variations. It was chosen to average over 90 days as for shorter periods it is possible that there is no rainfall, i.e. the dry April month in 2007. Also as 90 days is the length of a season it is a good length for the purpose of this study.

4. VARIOGRAM FITTING ON PSEUDO-CLIMATOLOGICAL DATA

With spherical variograms fitted to the pseudo-climatological rain data as descibed in Section 3 it is possible to find the seasonal variation of the sill and range. A frequency analysis is applied to find the best fit for the cosine function to descibe the seasonal variation.

4.1. Seasonal range

Rainfall is strongly seasonal both in type and frequency and this can also be seen in the range of the fitted variograms. The range reaches a minimum in May and a maximum in November. This can be attributed to the prevailing rain types during winter and summer. During summer the rain will tend to be convective which means that the correlation between two points quickly decreases with distance. In winter this changes with stratiform rain where over long distances rainrate can be similar. There are cases where the expected range goes beyond 400 km, which is even far beyond the furtest data pair and therefore not reliable. This occurs mostly around October when the variogram is more linear than spherical. To reduce this influence the maximum range was set to 300 km, which is also in line with the fact that the furtest data-pair is at 315 km and therefore a range beyond this would add little meaning. A square-root square-root (sqrt-sqrt) transform was applied to the data to try reduce the influence of extreme values for a better fit of the cosine function. When looking at the frequency domain in the top panel of Figure 2 it is clear there is only one meaningful frequency for the range. This is as expected the seasonal fluctuation. The resulting fit can be modelled as:

$$x_t = (A\cos(2\pi\omega t + \phi) + \text{offset})^4 \qquad (5)$$

This model is the same as Equation 3 with an added offset, but transformed with a power 4 to move back from the sqrt-sqrt transform. While the fit is not perfect the seasonal effect is followed quite well where most of the strong differences occur in October months when the variogram is more linear than spherical. The values for this fit can be found in table 1.

Another way to look at the fit is to take the average for each DOY of those 14 years. As shown in Figure 3, this results again in a clear seasonal trend. The black solid line is estimated range from the 90-day moving window spherical variograms, but with all ranges larger than 300 km removed. The grey solid line is the same, but with all ranges that would become larger than 300 km fixed at 300 km. As the data from the semi-climatological fit and that of the grey line is the same it is to be expected the fit is similar and as shown in the Figure 3 where the dotted line is the climatological fit this is indeed the case. The exception to the smooth cosine from the climatological fit is in October when the semi-variance tends to

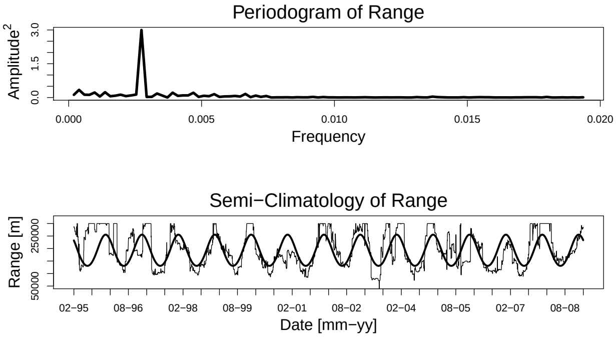

Figure 2. Top panel is the periodogram of the range with only one peak illustrating that fitting only one cosine shape is enough. Bottom panel is Seasonal variation of the range from 1995-2008. The thin line represents the found values and the thick black cosine the fitted climatological variation.

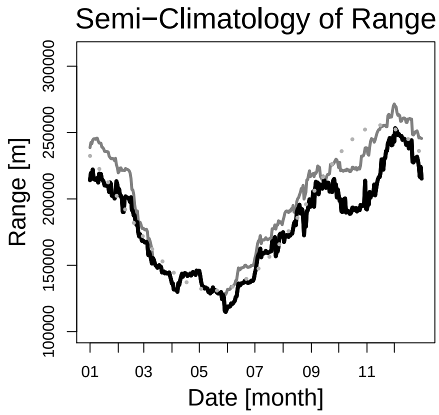

Figure 3. Seasonal variation of 90 day moving window variograms of range by daily averaged 14 year data. The black solid line is the average where range larger than 300 km is removed. The grey solid line is the average where the range larger than 300 km is set to 300 km.

Table 1. Parameters of the cosine function for the sill and range of the fitted spherical variograms on the sqrt-sqrt transformed daily rainsum of 29 stations using a 90-day moving averaging window.

	ω	A	ϕ	offset
sill	14/5114	0.253	2.945	1.807
range	14/5114	1.731	0.783	20.753

become more linear than spherical and range therefore becomes larger than 300 km. The low here is due to the fact that when the range is not linear beyond 300 km for the spherical fit it tend to find a fairly short range and this influences the average.

4.2. Seasonal sill

Like the range before, the seasonality is clearly apparent for the sill of the fitted variograms (Fig. 4). The sill data was again sqrt-sqrt transformed and fitted to a cosine model similar to the range and the corresponding values are found in table 1.

The sill reaches a maximum in July and a minimum in January. Again this can be attributed to the prevailing rain types during winter and summer. With convective rainfall the covariance between pairs wil be high, but in winter with similar rainrates over large distances the daily rain sum will be quite similar and results in low variance. As with the range exceptions occur when the fit is more linear than spherical. The reason that extreme values for sill occur less than for the range can be attributed to the fact that the times that the variogram is more linear occurs near the minimum of the sill function. The variance will tend to have very small values during this time which results in a small sill despite the range lying far away. There is a shift of about 126 days between the fitted cosines of the sill and range. A shift of approximately half a year is to be expected with strongest covariance during late summer when strong convective thunderstorms are most common resulting in a high sill and the equal amounts of daily rainsums between pairs at longer distances in winter resulting in a high range. The fact that the shift is in fact only 4 months can be explained by the transition period in autumn when the variogram decreases in strength but the range increases.

Similar to the range in Fig. 3 the average sill is also plotted. Fig. 5 shows the results, where the black line is the sill with all values more than 45 mm^2 removed and the grey line all values more than 45 mm^2 set to 45 mm^2. Again the seasonality is clear with an exception around October. Where first there was a low in the range a peak can be seen in the sill. The effect of the high sill due to linearity is far stronger here than it was for the local low range because the range was near the highest value already while the sill is near the lowest value. This results in a higher difference factor and a stronger influence of extremes. Because of this the peak for the removed high sill values is far less pronounced than the one where the high values are included. The effect of these larger sill values in October is also visible in the top panel of Fig. 4 where there is another small peak visible just to the right of the seasonal peak. While incorporating this cosine also in the function would result in slightly better fit it was chosen not to apply it as the it would go beyond the purpose of a very simple 2-parameter spherical variogram equation.

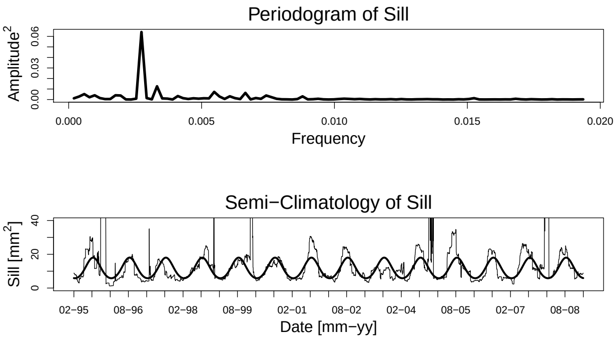

Figure 4. Top panel is the periodogram of the sill illustrating with only one peak that fitting only one cosine shape is enough. Bottom panel is Seasonal variation of the sill from 1995-2008. The thin line represents the found values and the thick black cosine the fitted climatological variation.

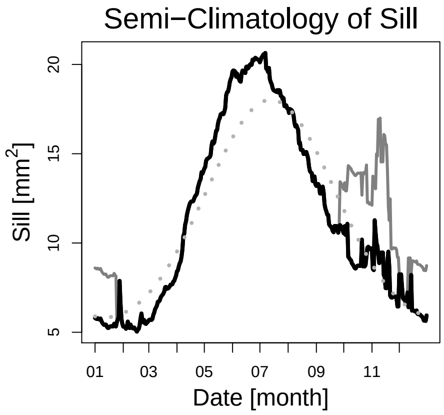

Figure 5. Seasonal variation of 90-day moving window variograms of sill by daily averaged 14-year data. The black solid line is the average where sill larger than 45 mm^2 is removed. The grey solid line is the average where the range larger than 45mm^2 is set to 45 mm^2.

5. CONCLUSIONS AND RECOMMENDATIONS

This paper has tried to prove that variograms of rainfall are strongly seasonal and that the found climatological equations for range and sill of a spherical variogram can be used for catchment hydrology in the Netherlands. The average sill and range found from the fitted spherical variogram follow a cosine function over the entire year with the exception of October where the semi-variance often is linear beyond 300 km. The yearly changes in sill and range do differ and the peaks and lows can at times be moved by as much as a month from the maximum found from a climatological fit. Also the value of the peaks and lows can be higher or lower of the climatological fit, but as mentioned before on average the fit works well.

While the results are promising the climatological variograms still have problems that need to be resolved. Some recommendations for continued research would be to:

- Evaluate if another variogram model could work better that a spherical one
- Use nested variograms for winter and spring periods
- Test variogram shape and stability for other time scales
- Compare estimated rainfall maps using the climatological-variance with maps made from daily variance values and radar measurements

ACKNOWLEDGMENTS

Financial support for this research is provided by the Space Research Organization Netherlands (SRON) through grant EO-058.

REFERENCES

[1] G. Villarini, P. V. Mandapaka, W. F. Krajewski, and R. J. Moore. Rainfall and sampling uncertainties: a rain gauge perspective. 113:D11102, 2008. doi:10.1029/2007JD009214.

[2] G. J. Ciach. Local random errors in tipping-bucket rain gauge measurements. *J. Atmos. Oceanic Technol.*, 20(5):752–759, 2003.

[3] L. C. Sieck, S. J. Burges, and M. Steiner. Challenges in obtaining reliable measurements of point rainfall. *Water Resour. Res.*, 43:W01420, 2007. doi:10.1029/2005WR004519.

[4] H. Leijnse, R. Uijlenhoet, and J. N. M. Stricker. Hydrometeorological application of a microwave link: 2. Precipitation. *Water Resour. Res.*, 43:W04417, 2007. doi:10.1029/2006WR004989.

[5] J. Joss and A. Waldvogel. Comments on “some observations on the Joss-Waldvogel rainfall disdrometer”. *J. Appl. Meteorol.*, 16(1):112–113, 1977.

[6] J. M. Schuurmans, M. F. P. Bierkens, E. J. Pebesma, and R. Uijlenhoet. Automatic prediction of high resolution daily rainfall fields for multiple extents: the potential of operational radar. *J. Hydrometeorol.*, 8(6):1204–1224, 2007. doi:10.1175/2007JHM792.1.

[7] N. A. Cressie. *Statistics for spatial data*. Wiley, NY, revised edition edition, 1993.

[8] P. J. Diggle and P. J. Ribeiro Jr. *Model-based geostatistics*. Springer, NY, 2007.