Hydrology and Earth System Sciences Discussions

Interactive comment on "Spatial interpolation of hourly rainfall – effect of additional information, variogram inference and storm properties" *by* A. Verworn and U. Haberlandt

A. Verworn and U. Haberlandt

a.verworn@iww.uni-hannover.de

Received and published: 24 November 2010

GENERAL REMARKS (BOTH REFEREES):

We are very grateful to the referees for their comments and suggestions on our manuscript. We think that the reviews helped us a lot to improve the manuscript. We will respond to all specific comments in the following. All technical corrections have been considered in the revision of the manuscript. Only those comments, which require a special response, are listed below.

RESPONSE TO THE COMMENTS OF REFEREE P. GOOVAERTS:

1. Although the averaging of theoretical variogram parameters is rather unusual it provided a straightforward way to estimate mean variograms. Comparisons have shown though, that a mean experimental variogram calculated over all events matches well with the overall averaged theoretical variogram.

2. Using the semivariogram of original rainfall data instead of residuals for KED is certainly a simplification. The difficulty lies in the simultaneously unknown trend and residuals. To overcome this problem, the trend component could be computed with a slightly modified KED system (Deutsch and Journel, 1992), while calculating the residuals and the residual semivariogram afterwards. However, this iterative process would be very demanding. Another simpler approach could be to infer the semivariogram only from data pairs, that are unaffected by the trend. Investigations of Haberlandt (2007) have shown though, that there were no differences in interpolation performance, which was the reason, why we applied the simplified approach in our study as well.

3. The spherical variogram model was fitted for all experimental approaches including for automatic variogram estimation (see Eq. 2). The minimisation of Eq. (5) was achieved with the Nelder and Mead optimisation method (see e.g., Press et al., 1989) assuming that the variance is equal to the sum of nugget c_0 and sill c. Therefore, only the two parameters range a and the ratio c_0/c needed to be fitted. The nugget was forced to zero only, if the calculated value was below zero. The sill was then set equal to the variance. The nugget values in Table 3 were manually estimated. Small nugget values below 0.1 were forced to zero. The low nugget effect is related to the used dense network of radar data (one thousand randomly selected radar cells). As an example please refer to Figs. 2 and 3. The three decimal digits in Table 3 indicate a precision, which is not true. This has been corrected now.

4. The statement that log-transformed secondary data can yield a lower variance of the interpolated values was made based on experimental evidence achieved during the investigations. It has been removed from this section and will be better included in the section 'analyses and results'.

5. We agree that theoretically KED should be equivalent to OK if there is no correlation between primary and secondary variables. However, experimental evidence showed that for some time steps with low correlation KED resulted in significant worse performance than OK. With the conditional versions this could be avoided. One reason for this effect might be due to instabilities in the kriging system for time steps with many zero values leading to large interpolation errors.

6. The entire study area was covered with a fixed window size of 230 km x 230 km. The raster resolution for interpolation was set to 1 km x 1 km. Precipitation is usually affected by larger-scale orographic characteristics than the grid resolution. To find the optimal resolution of the digital elevation model regarding derivation of topographic indices, the cellsize was changed, but with the condition of keeping the same extent. This procedure resulted in a cell size of 5.75 km x 5.75 km for the luv/lee index showing the highest correlation to observed precipitation. The subsequent downscaling of the index to the 1 km x 1 km resolution for the precipitation interpolation task was carried out with a bilinear resampling technique.

7. In order to use the daily rainfall data as secondary information for KED, the data were spatially interpolated with the inverse distance method (IDW). Another alternative would have been the application of ordinary kriging (OK). However, due to the dense network of daily stations the differences were expected to be low and the simpler procedure has been applied here.

8. The different thresholds for rainfall intensities were established as follows: The low threshold of 0.1 mm/h was used for the events statistics in order to exclude time steps with non-significant rainfall. For precipitation cross-validation the focus was on time steps with heavy precipitation regarding floods. Therefore, the threshold was raised to 1.0 mm/h. The estimation of semivariograms using radar data followed the same objective as precipitation cross-validation. However, the threshold had to be lowered to 0.5 and 0.25, respectively, due to the systematic rainfall underestimation by radar. Otherwise, not enough time steps would have been selected for some events.

9. The large RMSE values show the overall high uncertainty in hourly precipitation interpolation. While in Table 4 only the semivariogram inference was assessed, the interpolation performance as main objective is referred to in Tables 5 and 6. Comparing the RMSE values from both tables it can be seen, that for summer events the RMSE is significantly larger than for winter events.

10. The expression "works best" means, that the RMSE of precipitation cross-validation results was lowest for the respective correlation thresholds. Higher thresholds reduced the advantage of the multivariate methods compared to the univariate ordinary kriging (OK), while at lower thresholds instabilities in the kriging system prevented better results.

11. In case of interpolation on log-transformed values y, the estimated values were back-transformed taking exp(y) before calculating performance measures. With interpolation on log-transformed values, problems with biased results may occur. The exponentiation in the back-transformation process exponentiates any errors, which is very sensitive to interpolation outliers. However, in our case the results seemed not be affected of that as the cross-validation results showed values of the bias in the same magnitude as for interpolation methods without log-transformed data.

RESPONSE TO THE COMMENTS OF REFEREE G. PEGRAM:

6412: 7 to 9: The exact matching of the experimental and theoretical variogram is considered less important than the choice of variogram model itself.

6413: 1 to 2: The term "zonal" means, that not only geometric anisotropies with different ranges but same sills were taken into consideration, but also more complex structures with sills not being the same in all directions.

6414: 2 to 3: Thank you for this information!

6416: 8 to 9: Actually, the multi-step procedure was carried out as suggested: The determined rainy cells of the first step were set to one, while the no-rainy cells obtained

a zero. Therefore, multiplying both fields resulted basically in interpolating just over the wet areas.

6418: 28: The 'Brocken' will be included in Fig. 1.

6423: 15 to 16: Using elevation as additional information for KED led only for one third of all events to slight improvements in comparison with ordinary kriging. Those events were characterized by different precipitation types. A distinct classification, which precipitation types are favourable for KED with elevation, could not be made in this study. One reason for the weak correlation between precipitation and elevation is certainly the short hourly time step.

6427: 1 to 2: Geostatistical interpolation methods usually lead to a spatial smoothing of the rainfall distribution with underestimation of high values und overestimation of low values. Especially, no-rain areas become rainy areas with low precipitation intensities after interpolation. These smoothing effects are not wanted for flood simulation. Seen from this angle a larger no-rainfall fraction is in most cases desirable.

6428: 1 to 4: The mentioned conclusion is revised as follows: Using KED with other variables like elevation or luv/lee index could hardly improve the interpolation performance according to precipitation cross-validation compared to the univariate method ordinary kriging (OK). However, when including topography, spatial distributions of precipitation showed shaper differentiations of spatial structures, which seemed more plausible and the use of additional information from topography is therefore recommended.

CITED REFERENCES:

Deutsch, C. V., and Journel, A. G.: GSLIB: Geostatistical Software Library and User's Guide, Oxford University Press, New York, 1992.

Haberlandt, U.: Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event, Journal of Hydrology, 332, 144-157,

2007.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T.: Numerical Recipes in Fortran 77: The Art of Scientific Computing, Cambridge University Press, New York, 1989.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 7, 6407, 2010.