QPE.P25

Gauge-radar adjustment by using multivariate kernel regression and spatiotemporal Kriging

Seppo Pulkkinen, Jarmo Koistinen and Timo Kuitunen
Finnish Meteorological Institute, Helsinki, Finland

Statistical model

- We model the discrepancy between gauge rainfall accumulation G and radar rainfall accumulation R by

$$F = 10 \cdot \log_{10} \frac{G}{R}$$

- Based on empirical evidence, our model is based on the assumption that this quantity is normally distributed.

- The assumption is that the error can be separated into a *systematic* component μ and a standard normally distributed *residual* component ϵ according to

$$F(\theta) = \mu(\theta) + \sigma(\theta)\epsilon \qquad (1)$$

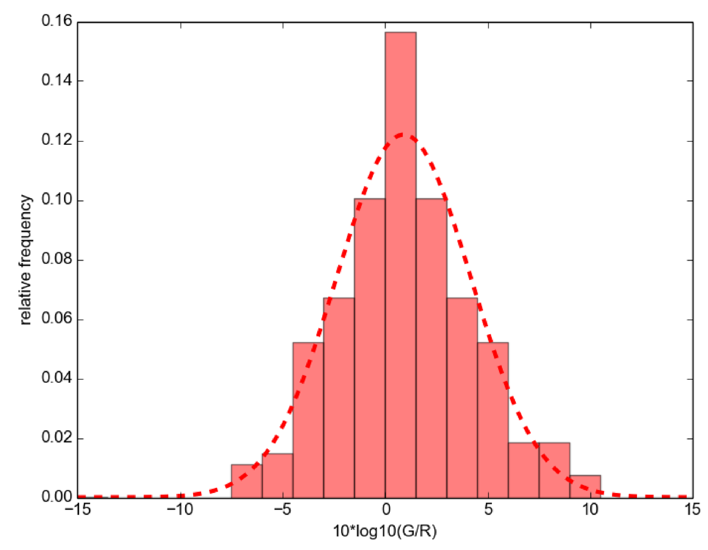

Figure 1. Hourly F-values, 179 samples.

- In practice the systematic bias μ depends on multiple factors $\boldsymbol{\theta}$ such as rainfall accumulation, distance from radar, altitude of the radar bin and precipitation type.

- To account for multiple factors affecting the systematic bias μ, we use a multivariate *kernel regression* model

$$\hat{\mu}_{\mathbf{H}}(\boldsymbol{\theta}) = \frac{\sum_{i=1}^{n} K_{\mathbf{H}}(\boldsymbol{\theta} - \boldsymbol{\theta}_i) f_i}{\sum_{i=1}^{n} K_{\mathbf{H}}(\boldsymbol{\theta} - \boldsymbol{\theta}_i)}$$

where the kernel $K_{\mathbf{H}}$ is a Gaussian function whose covariance matrix $\mathbf{H}$ is estimated from the observed variables $\boldsymbol{\theta}_i$ and F-values f_i.

- The kernel regression technique is also applied to residuals, which gives an estimate $\hat{\sigma}^2_{\mathbf{H}}$ for the regression variance σ^2.

- We use *spatiotemporal* Kriging to interpolate the residual component ϵ. Kriging gives also a variance estimate.

Main advantages:

- The regression model is capable of explaining multiple factors contributing to the systematic bias and residual error variance.

- The model is better able to separate the systematic and residual errors than simple univariate models or models assuming constant variance.

- The model accounts for transient errors by utilizing spatiotemporal correlation of residuals. Addressing temporal correlation is beneficial due to sparsity of rainfall observations and gauge network.

The model is fully probabilistic: for each bin or area-averaged radar rainfall, it gives mean and variance estimates $\hat{\mu}$ and $\hat{\sigma}^2$ determining a normal distribution for the corresponding ground rainfall.

Illustrative examples

(Ikaalinen radar, lowest elevation angle, accumulation time 1 hour, June-September 2013, attenuation+VPR corrected)

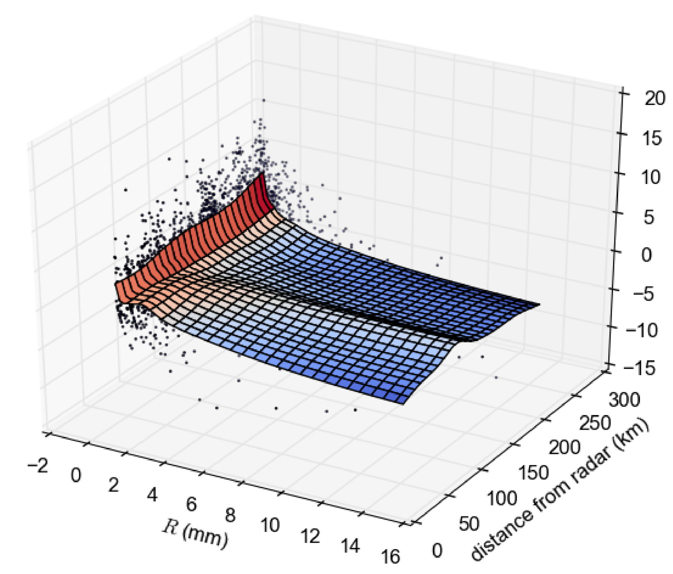

Figure 2. Regression surface $\hat{\mu}_{\mathbf{H}}$ fitted to F-values.

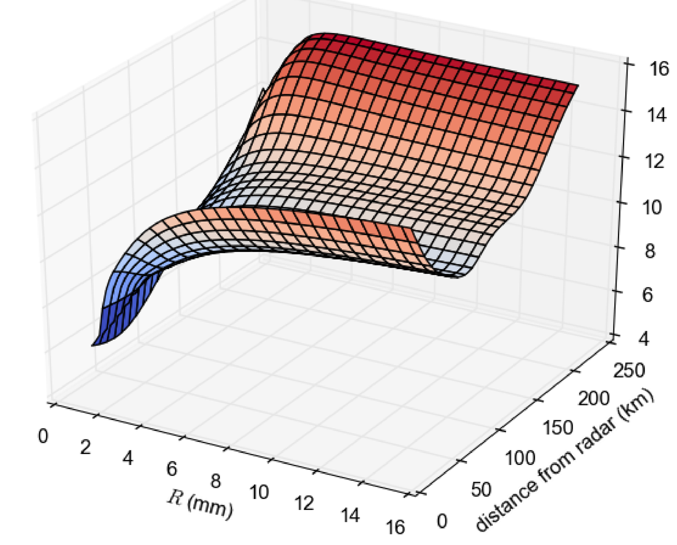

Figure 3. Residual variance $\hat{\sigma}^2_{\mathbf{H}}$ of regression surface fitted to F-values.

- The multivariate regression model reveals, for instance, different range-dependent behavior of gauge-radar error for different accumulations.

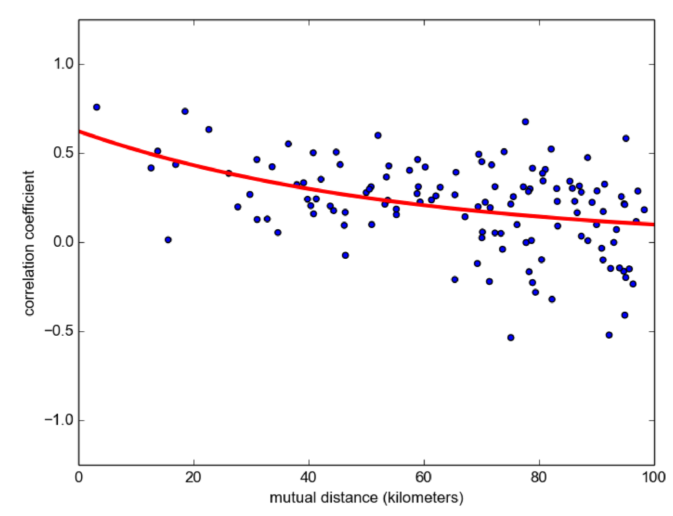

Figure 4. Spatial correlation of residual errors.

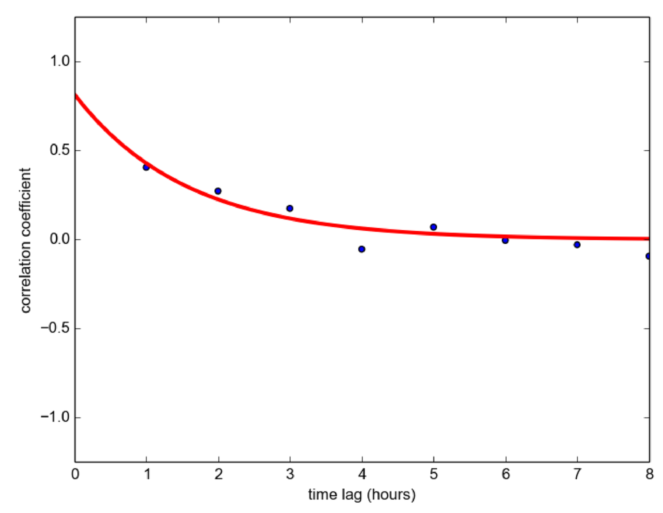

Figure 5. Temporal correlation of residual errors.

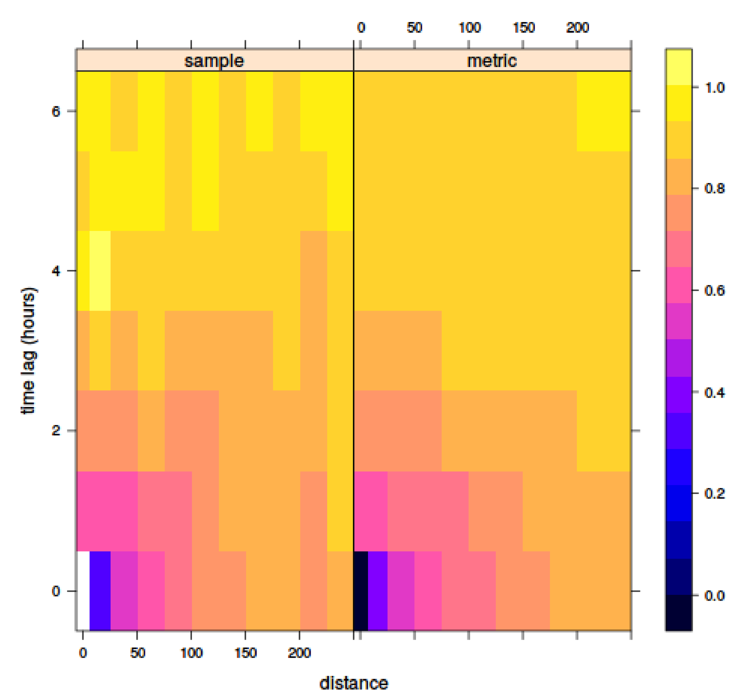

Figure 6. Variogram model fitted to residuals.

- The residual errors ϵ have both spatial and temporal correlation. This can be utilized when computing Kriging-interpolated residuals.

- A metric variogram model (time as a third spatial coordinate) is used for the Kriging interpolation.

- The model is anisotropic in space and time and consists of exponential functions.

- Observations from multiple elevation angles can be used when altitude is included as one coordinate.

Application of the method

(FMI radars, lowest elevation angle, accumulation time 1 hour, 23. October 2013 03:00-04:00 UTC, attenuation+VPR corrected)

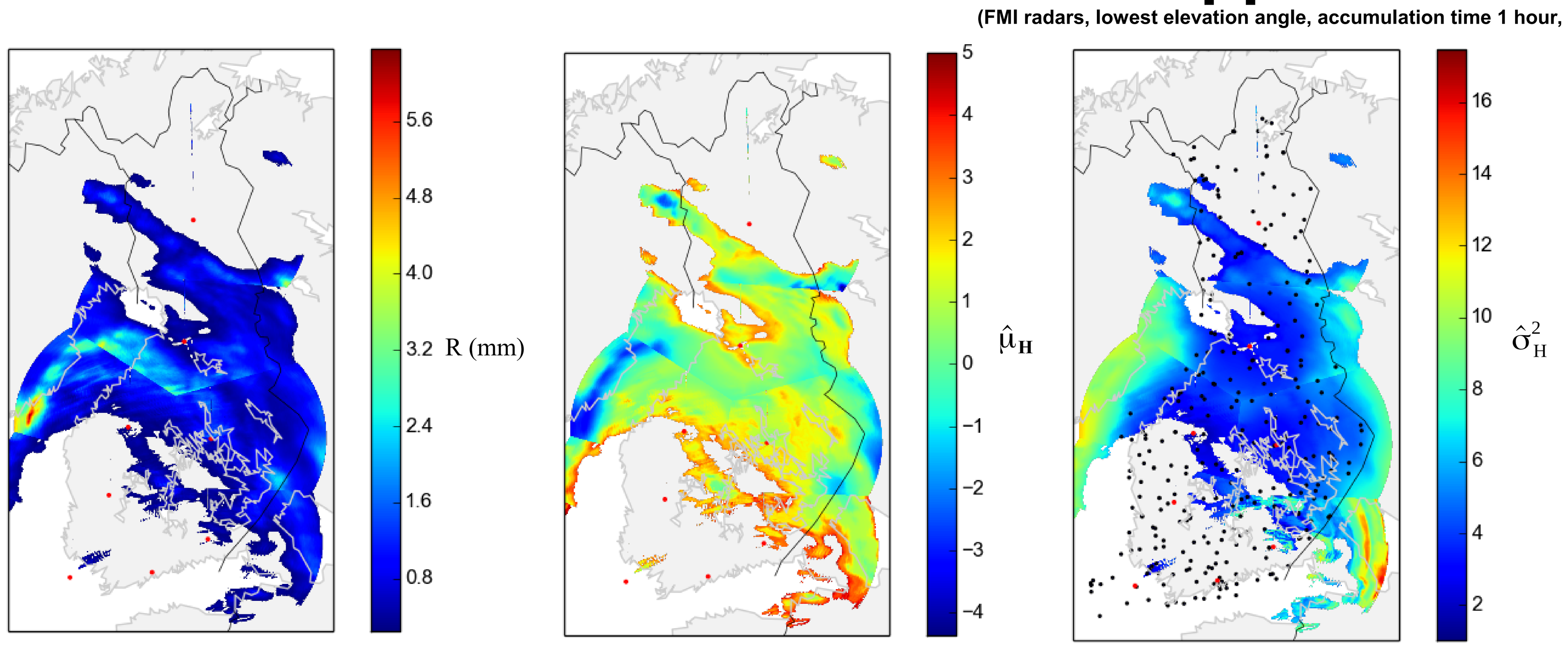

Figure 7. Original radar rainfall field and mean and variance estimates for F.

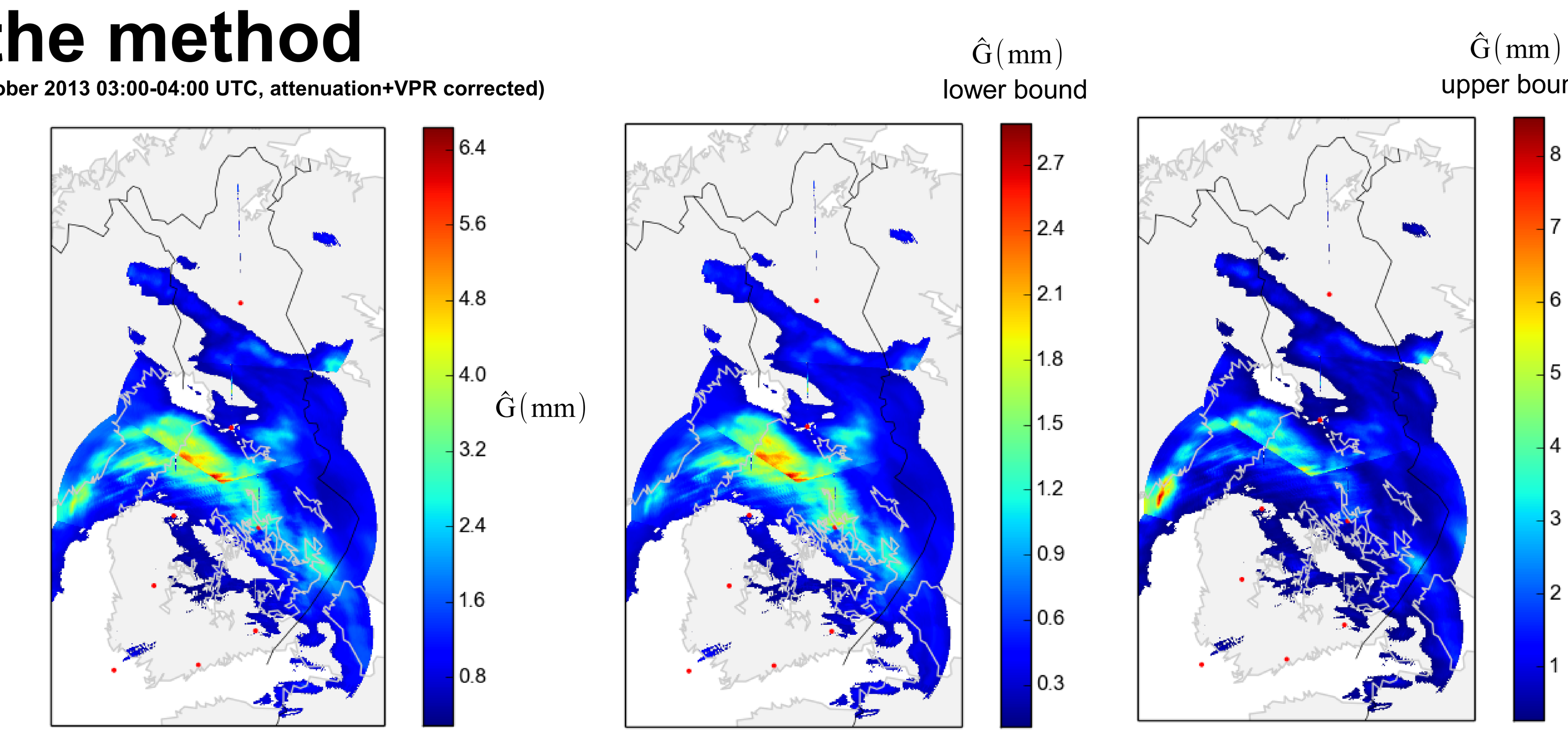

Figure 8. Regression and Kriging -corrected radar rainfall field and lower and upper bounds of 90% probability intervals.

Main observations:

- The mean F-values for small rainfall accumulations are positive (meaning radar underestimation).
- The mean F-values for large rainfall accumulations far from radars are negative (meaning radar overestimation). A likely explanation is a locally biased VPR correction and occurrence of hail.
- The model gives large variances (i.e. uncertainties) far from radars.
- The kriged estimates have small variances near gauges.

- For a given radar rainfall R, the expected ground rainfall $\hat{G}$ is given by

$$\hat{G} = 10^{\frac{\hat{\mu}}{10}} R$$

where $\hat{\mu}$ is the expected F computed from (1) by using the regression and Kriging estimates.

- By using the estimated distribution of F, we can also compute the probability that the ground rainfall accumulation lies at the given interval.

Validation results

10-fold cross validation with independent rain gauges:

- Ten iterations, at each iteration 1/10 of the gauges are not included in the model fitting or the Kriging interpolation points.
- Mean and variance of F are averaged over all gauges within the coverage of each radar.

Main observations:

- After regression correction, the systematic bias is near zero. The variance of F is also decreased.
- Using two regression variables yields the best results.
- Applying Kriging to residuals further decreases the variance of F.

	original	regression (R)	regression (distance)	regression (R+distance)
mean(F)	0.49312	-0.08725	-0.014	-0.0575
variance(F)	10.957	9.4274	10.957	9.3266

Table 1. F-values for the FMI gauge and radar networks with different regression variables.

	original	regression (R+distance)	regression (R+distance) and Kriging
mean(F)	0.88388	-0.01275	-0.0072
variance(F)	11.116	8.4351	7.6402

Table 2. F-values for the FMI gauge and radar networks with regression and combined regression and Kriging models.