Hydrology and Earth System Sciences

Areal rainfall estimation using moving cars as rain gauges – a modelling study

U. Haberlandt[1] **and M. Sester**[2]

[1]Institute of Water Resources Management, Hydrology and Agricultural Hydraulic Engineering,
Leibniz University of Hannover, Hannover, Germany
[2]Institute of Cartography and Geoinformatics, Leibniz University of Hannover, Germany

Abstract. Optimal spatial assessment of short-time step precipitation for hydrological modelling is still an important research question considering the poor observation networks for high time resolution data. The main objective of this paper is to present a new approach for rainfall observation. The idea is to consider motorcars as moving rain gauges with windscreen wipers as sensors to detect precipitation. This idea is easily technically feasible if the cars are provided with GPS and a small memory chip for recording the coordinates, car speed and wiper frequency. This study explores theoretically the benefits of such an approach. For that a valid relationship between wiper speed and rainfall rate considering uncertainty was assumed here. A simple traffic model is applied to generate motorcars on roads in a river basin. Radar data are used as reference rainfall fields. Rainfall from these fields is sampled with a conventional rain gauge network and with several dynamic networks consisting of moving motorcars, using different assumptions such as accuracy levels for measurements and sensor equipment rates for the car networks. Those observed point rainfall data from the different networks are then used to calculate areal rainfall for different scales. Ordinary kriging and indicator kriging are applied for interpolation of the point data with the latter considering uncertain rainfall observation by cars e.g. according to a discrete number of windscreen wiper operation classes. The results are compared with the values from the radar observations. The study is carried out for the 3300 km² Bode river basin located in the Harz Mountains in Northern Germany. The results show, that the idea is theoretically feasible and motivate practical experiments. Only a small portion of the cars needed to be equipped with sensors for sufficient areal rainfall estimation. Regarding the required sensitivity of the potential rain sensors in cars it could be shown, that often a few classes for rainfall observation are enough for satisfactory areal rainfall estimation. The findings of the study suggest also a revisiting of the rain gauge network optimisation problem.

1 Introduction

Rainfall is the most important input information for hydrological planning and water resources management. Especially the modelling of highly dynamic and nonlinear processes like floods, erosion or wash out of pollutants relies heavily on good information about precipitation. Due to its high variability in space and time observation of rainfall is still a challenging task. While the classical networks of non-recording rain gauges with a daily observation interval have reached a sufficient density and a good standard, the availability and density of recording rain gauges for the observation of short time step rainfall is still inadequate. Even for developed European countries like Germany the network density of recording rain gauges considering stations with longer records from the German Weather Service (DWD, MI network) is only about one station per 1800 km² compared to a density of about one station per 90 km² for non-recording rain gauges. Weather radar is a very important new data source for measuring rainfall. However, despite the high spatial resolution of radar data there is often a large space-time variable bias in radar rainfall estimates (Smith et al., 2007; Krajewski and Smith, 2002). So, a sufficient point precipitation network is still needed for calibration. Other special and innovative methods for rainfall observation use satellites (Grimes and Diop, 2003; Wardah et al., 2008), microwave

links (Leijnse et al., 2007; Messer et al., 2006) or rain gauges aboard moving ships to measure rainfall at sea (Hasse et al., 1998; Yuter and Parker, 2001). Utilising rainfall information from different sources together and applying sophisticated interpolation or merging methods can further improve precipitation estimation for hydrological applications (Goudenhoofdt and Delobbe, 2009; Chiang et al., 2007; Goovaerts, 2000; Haberlandt, 2007; Ehret et al., 2008).

The main objective of this paper is to present a new idea for measuring precipitation using moving cars as rain gauges with windscreen wipers as sensors to detect precipitation. This idea would easily be technically feasible if the cars were provided with a GPS and a small memory chip for recording time, car velocity and wiper frequency. Alternatively, also an online transmission of the sensed information via mobile phones could be realized. The potential of such a concept becomes immediately clear considering the high and ever increasing traffic density worldwide with a huge number of cars; e.g. in Germany exist more than 40 million cars (EUROSTAT, 2009).

This study explores theoretically the benefits of such an approach. A simple statistical traffic model is applied to generate car traffic on main roads in a river basin. Radar rainfall data are used as reference rainfall fields. Rainfall from these fields is sampled with a conventional rain gauge network and with a dynamic network consisting of moving motorcars. Those observed point rainfall data from the two networks are then used to calculate areal rainfall for different scales using geostatistical interpolation methods. The results are compared with the reference values from radar observations.

2 Methodology

2.1 Traffic generation model

The traffic flow of the moving cars is determined in a simple stochastic simulation process. The necessary data for the traffic flow model is geometric information about roads with associated information about traffic density at different times of the day. Furthermore, assumptions about the relative number of cars that is equipped with a rain sensor are needed. Finally, a factor determines the sampling rate, i.e. the frequency of measurements in time.

Based on these assumptions, cars are randomly generated for each road segment. From their starting position, the cars virtually drive with the given velocity towards the end-node of the road segment. Their position is interpolated according to the given sampling rate, here in intervals of 5 minutes. This leads to a collection of car positions on the roads in the form: x, y and time. Subsequently, rainfall observations are assigned to the car positions and areal rainfall is estimated using interpolation (see below).

2.2 Rainfall observation by cars

Rainfall observation by cars could be realised using the windscreen wipers as sensors. The cars need to be equipped with a GPS system to obtain the geographical position. In addition a memory chip is necessary to register time, location, car speed and wiper frequency. A crucial task for this system to work is the derivation of a relationship between wiper speed (W) and rainfall intensity (R). This relationship will be called in the following W-R relationship. In general the rainfall intensity to be estimated depends on the wipers frequency, the car speed, the specific properties of the car and the driver's preferences operating the car. The influence of the latter could be avoided if photocell rain detectors installed in modern cars for automatic set up of wiper speed are used directly as sensors.

In order to establish a W-R relationship calibration is required. The calibration of the W-R relationships using observed data from rain gauges would have some similarities to the procedure for calibration of radar data using the Z-R relationship. The main difference is that each car has its own specific W-R relationship which is primarily related to the wiper system, the driver, the cars speed and the local conditions. To test the idea of areal rainfall estimation by cars W-R relationships are needed. In the following possible ways to establish W-R relationships are briefly discussed:

a. Assuming error free W-R relationships.
   As base line scenario a correct W-R relationship with exact rainfall measurement by cars is assumed. The point error for rainfall observation by cars is zero in this case. This allows assessing the impact of point observation errors on the total uncertainty of areal rainfall estimation by comparison of this base line scenario with the following cases, where point errors are taken into account. For this reference case the total error differences in areal rainfall estimation from gauges and cars depends only on the interpolation error from the different observation networks. Note, that for this and all other cases error free observation of rainfall at the rain gauges is assumed.

b. Assuming W-R relationships with uncertainty.
   Simple W-R relationships with uncertainty will be assumed. Here, the point observation error is taken into account by considering only a limited number of rainfall classes with constant observed wiper speed for the certain ranges of rainfall. For instance rainfall observations based on 4 classes, corresponding to less accurate measurements ("more old fashioned cars") and 10 classes, corresponding to rather precise measurements ("more modern cars") could be considered. The assumed point error is uniformly distributed within the classes and can be taken into account e.g. by taking mid class values as observed car rainfall or by using indicator kriging for interpolation (see below).

c. Estimating W-R relationships from laboratory experiments.

A set of basic W-R relationships for different cars and wiper systems might be derived in laboratory experiments. Climate-wind tunnels may be employed for that allowing exact definition of rainfall intensities, car speed, etc. A simple model for the point errors could be assumed considering e.g. normal distribution for the errors. Those errors can be estimated from average and variance of the single relationships derived as outcome of the experiments. To consider this point error in the simulation experiments realisations of wiper frequencies belonging to the truth rainfall are drawn from normal distributions providing in return the uncertain point estimations of rainfall for subsequent interpolation.

d. Estimating W-R relationships from field experiments.

For field experiments some cars could be equipped with GPS and devices for recording car speed, location and wiper frequency. If rainfall intensities are available a calibration of the W-R relationships similar to Z-R relationships can be tried. Rainfall intensities can be taken from nearby rain gauges or from radar observations. However, both methods involve also significant uncertainties to estimate the reference rainfall for the car locations. When using nearby rain gauges as reference the interpolation errors have to be considered. In case of using radar as reference the uncertain transformation of reflectivity into rain rate needs to be taken into account. Also a merged precipitation product could be used as reference (e.g. Goudenhoofdt and Delobbe, 2009). The calibration can be carried out in a re-analysis mode using "historic" data or in an on-line mode using rainfall rates obtained from nearby automatic rain gauges with remote transmission capabilities. The latter would allow a dynamic updating of the transformation functions and an adaptation to different conditions like different drivers of the cars. Such a system could be further improved by allowing a communication between the cars propagating such information through the whole "sensor network" (Stefanidis and Nittel, 2004).

Investigations for establishing valid relationships between wiper frequency and rain rate from laboratory or field experiments (cases c. and d.) are beyond the scope of this pilot study and need quite an amount of further research. Here, it will be assumed, that such relationships exist, either error free as reference (case a.) or with assumed uncertainty based on a restriction of the continuous rainfall intensities to some discrete observable classes (case b.) The uncertainty in the W-R relationship is one important source or errors for areal rainfall estimation by car networks. The other important source is the interpolation error mainly related to the limited number of observation points to estimate areal rainfall, which is discussed in the following and which is the main focus of the paper.

2.3 Areal rainfall assessment

Rainfall interpolation is required to estimate raster based continuous rainfall fields in space from point observations and to calculate areal rainfall for regions or catchments. For rainfall interpolation two geostatistical methods ordinary kriging (OK) and indicator kriging (IK) are employed. OK is used for the interpolation of rainfall from the stationary gauge network and as reference method assuming continuous observation ability of the moving rain gauges. IK is used for the interpolation of rainfall measured by the cars considering their discrete observation ability in practice. It is assumed, that using the moving cars as rain gauges only a limited accuracy is possible providing rainfall observations in discrete classes. In the following a brief overview of the interpolation methods OK and IK and their implementation for this investigation is given. A more detailed description about the theory of the methods can be found in geostatistical textbooks (e.g. Goovaerts, 1997; Isaaks and Srivastava, 1989).

Ordinary Kriging (OK) is the best known and most used geostatistical interpolation method (Matheron, 1971). For using OK the requirements of the intrinsic hypothesis have to be met. That means, first, the expected value of the variable Z is constant in the whole domain

$$E[Z(u)] = m \quad \text{for all} \quad u \in D \tag{1}$$

and, second, the variance of the differences $[Z(u + h) - Z(u)]$ between two points depends only on the distance vector $\mathbf{h}$ and not on the locations u and $u + \mathbf{h}$

$$\gamma(\mathbf{h}) = \frac{1}{2}\text{Var}[Z(u + \mathbf{h}) - Z(u)] \qquad u \in D. \tag{2}$$

The function $\gamma(\mathbf{h})$ is called here variogram (exact semivariogram), characterising the spatial variability of the target variable. The linear estimator for the unknown point u_0 is a weighted sum of the observations from the n surrounding points u_i:

$$Z^*(u_0) = \sum_{i=1}^{n} \lambda_i Z(u_i). \tag{3}$$

The weights λ are calculated using the OK kriging system

$$\sum_{j=1}^{n} \lambda_j \gamma(u_i - u_j) + \mu = \gamma(u_i - u) \quad i = 1,...,n$$
$$\sum_{j=1}^{n} \lambda_j = 1 \tag{4}$$

where λ are the variogram values and μ is a Lagrange parameter. Considering that time series of precipitation need to be processed a variance weighted average experimental variogram is calculated assuming isotropy

$$\bar{\gamma}^*(\mathbf{h}) = \frac{1}{m \cdot 2n(\mathbf{h})} \sum_{t=1}^{m} \frac{1}{s_t^2(Z)} \sum_{i=1}^{n(\mathbf{h})} [Z(u_i, t) - Z(u_{i+\mathbf{h}}, t)]^2, \tag{5}$$

where m is the number of time steps t and s_t^2 is the variance for each time step. In this case only one experimental variogram is obtained, for which the fitting of the theoretical model can be done manually. The combination of a nugget effect with an exponential model is used here uniformly as theoretical variogram model:

$$\gamma(\boldsymbol{h}) = c_0 + c \cdot \left[1 - \exp\left(-\frac{3\boldsymbol{h}}{a_e} \right) \right], \tag{6}$$

where a_e is the effective range, c is the partial sill and c_0 is the nugget variance.

For indicator kriging (IK) (Journel, 1983) the observed variable $Z(u)$ is first transformed into a binary indicator variable I_α according to

$$I_\alpha = \begin{cases} 1 & \text{if} \quad Z(u) \leq \alpha \\ 0 & \text{otherwise} \end{cases}. \tag{7}$$

Using several thresholds α_k with $k = 1, ..., K$ gives a vector of indicator variables $I_{\alpha,k}$. Variograms have to be inferred for all indicator variables separately. Considering ordinary indicator kriging the interpolation is done for each indicator using the OK framework, which gives in the end an estimation of the cumulative distribution function (cdf) of $Z(u)$. Order relation deviations are corrected a posteriori following the approach of Deutsch and Journel (1992, p. 81). The mean of the cdf approximated by its discrete sum provides then an estimate for the observed variable:

$$Z^*(u) = I_{\alpha_0}^*(u)\alpha_0 + \sum_{k=0}^{K} \left[I_{\alpha_{k+1}}^*(u) - I_{\alpha_k}^*(u) \right] \cdot \frac{\alpha_{k+1} - \alpha_k}{2}, \tag{8}$$

where α_0 and α_{k+1} are the minimum the maximum values of the Z-range, respectively.

2.4 Performance assessment

To compare the performance of the moving car network with the performance of conventional rain gauges for estimation of areal precipitation space-time high resolution rainfall fields are required as reference or "true" rainfall. Stochastically generated rain fields could be employed for this purpose (e.g. Seo et al., 1990). However, here it is preferred to use weather radar rainfall, since those data are probably closer to reality regarding space-time dynamics, and avoid the additional introduction of a stochastic weather generator. Areal rainfall estimates based on moving car networks and based on a stationary gauge network are then compared against "true" areal rainfall calculated from the reference radar data. Precipitation from both networks is interpolated on a regular raster for each time step using the methods described in Sect. 2.2 and then aggregated to areal averages $Z_A^*(t)$ considering different spatial scales

$$Z_A^*(t) = \frac{1}{N} \sum_{i=1}^{N} Z^*(t, u_i), \tag{9}$$

where N is the number of raster cells within the area and t is the time step. For evaluation of the areal precipitation estimates the following performance criteria are used:

the total bias

$$\text{Bias} = \sum_{t=1}^{m} \left[Z_A^*(t) - Z_A(t) \right], \tag{10}$$

the relative standard error normalised with the average of the reference areal value

$$\text{RSE} = \frac{1}{\bar{Z}_A} \cdot \sqrt{\frac{1}{m} \sum_{t=1}^{m} \left[Z_A^*(t) - Z_A(t) \right]^2} \tag{11}$$

and the coefficient of correlation

$$\text{Cor} = \frac{\text{Cor}\left[Z_A(t), Z_A^*(t) \right]}{\sqrt{\text{Var}[Z_A(t)] \cdot \text{Var}\left[Z_A^*(t) \right]}}, \tag{12}$$

where Z_A^* is the estimated areal precipitation, Z_A the reference areal precipitation and m the number of time steps considered for error calculation.

The performance of the areal rainfall estimation depends on the network density. To quantify the density of the station network and the car network the following kernel density estimator is applied to all cells $i = 1, ..., N$ of the 1 km $\times$ 1 km grid (Silverman, 1986):

$$D_i = \frac{1}{\pi r^2} \sum_{j=1}^{n} k_j \quad \text{with} \quad k_j = \begin{cases} 3\left(1 - \left(\frac{d}{r} \right)^2 \right)^2 & \text{for } d \leq r \\ 0 & \text{for } d > r \end{cases}, \tag{13}$$

where n is the number of stations within the search radius r and k is a quartic kernel weighting the stations according to their distance d from the cell centre. The total network density for one specific area or subbasin D_{sub} is then calculated by averaging D_i over all raster cells within the considered catchment. The kernel density is chosen here as estimator to consider all appropriate stations, also those located outside the catchment boundaries and to weigh the individual stations differently according to their distance from the subbasin.

2.5 General steps for analysis

For the analysis the following steps are carried out based on rainfall time series with a temporal discretisation of 5 min:

1. Areal reference rainfall time series are calculated for selected catchments from radar rainfall fields.

2. Point precipitation time series are extracted for a stationary rainfall gauge network from the radar rainfall fields.

3. Cars are randomly generated on roads according to temporal traffic density variations but with uniform spatial density considering different sensor equipment rates.

4. Point precipitation time series are extracted for the moving car networks from the radar rainfall fields considering the discrete observation ability of the car sensors. Rainfall from that raster cell in which the gauge is located or which the car is just crossing at the observation time interval is taken as observed point value.

5. The point rainfall time series from the gauge network are interpolated using ordinary kriging (OK) on a $1\,km \times 1\,km$ raster for the whole study area and are subsequently aggregated over selected catchments to areal rainfall time series.

6. The point rainfall time series from the different car networks are interpolated using ordinary kriging (OK), indicator kriging with 4 rainfall classes (IK4) and indicator kriging with 10 rainfall classes (IK10) on a $1\,km \times 1\,km$ raster each. Then they are subsequently aggregated over the selected catchments to areal rainfall time series.

7. Error statistics are calculated by comparisons of the areal rainfall time series from gauge and car networks each with the reference areal radar rainfall time series for the selected catchments.

3 Study region and data

3.1 Study region

The study is carried out for the $3300\,km^2$ Bode river basin located in the Harz Mountains in Northern Germany (Fig. 1). The considered Bode region has elevations between 1140 m a.s.l. at the top of the Brocken Mountain and about 80 m a.s.l. Mean annual rainfall varies between 1700 mm/yr and 500 mm/yr. Four mesoscale catchments of different sizes are selected for areal rainfall estimation, which comprise the Trautenstein ($40\,km^2$), the Selke ($102\,km^2$), the Holtemme ($167\,km^2$) and the Gr. Graben ($812\,km^2$) catchments. The stationary rainfall observation network consists of 14 recording rain gauges with 6 stations operated all year and 8 stations operated only during the summer time. The all year station network represents a typical recording rain gauge network density in Germany. In addition, a weather radar station is located at Ummendorf covering the whole area within its observation range.

For the car networks road data from the topographic information system (here, the German ATKIS) are used, where the roads are given according to different road categories ranging from highways to side roads. In order to make conservative estimations, only the highest road categories, namely highways and federal roads are considered here (see Fig. 1).
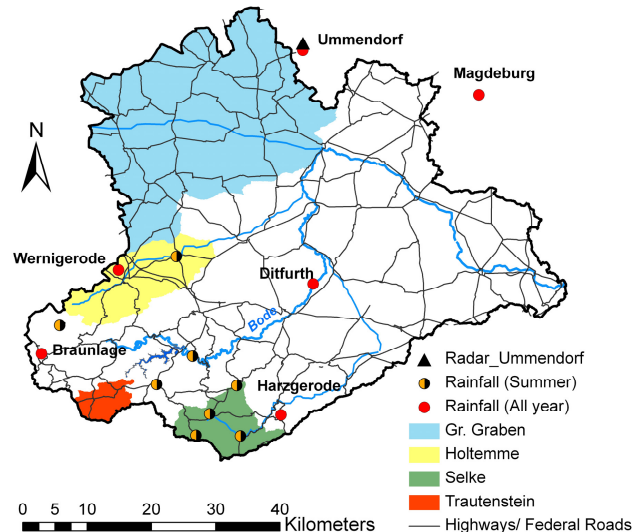
Fig. 1. Study region showing the selected four catchments for the estimation of areal rainfall, the stationary raingauge network and the location of highways and freeways used for the simulation of the car network.

3.2 Traffic data

Traffic statistics were taken from a study related to noise propagation (R.-D. Mummenthey, Trade Office Hildesheim, personal communication, September 2008). Table 1 shows the traffic data which are used in this study. Assumed average speed of the cars on major roads is 80 km/h. Three different classes of traffic density with respect to different times of the day are used. For simplification all traffic statistical data are applied uniformly within the whole study area. This is motivated by the fact, that only one type of road, namely major roads, are used neglecting all smaller roads. If other roads were also included it would increase the spatial traffic variability but it would also further increase the potential of the car network for improved rainfall estimation, i.e. our assumptions are conservative (see also Sect. 4.1). Only a certain fraction of all cars will be prepared to measure precipitation. For that four different sensor equipment rates between 0.5% and 4% are assumed here (see Table 1). The windscreen wipers are used as sensors to detect and measure precipitation. The rainfall intensity can be derived from the wipers frequency. As currently there is no known relationship between the wiper frequency and the rainfall intensity simple assumptions are made. The first one is to have an exact observation of the rainfall intensity as reference for potential maximum accuracy (case a. from Sect. 2.2). The other assumptions are to use different rainfall observation classes, e.g. corresponding to the wiper frequency intervals that can be manually set in the cars (case b. from Sect. 2.2).

Table 1. Traffic data used for car generation.

Road types	Highways, Federal Roads		
Velocity	80 km/h		
Density	Day (6 a.m.–5 p.m.) 1100 cars/h	Evening (5 p.m.–10 p.m.) 700 cars/h	Night (10 p.m.–6 a.m.) 180 cars/h
Sampling rate	5 min		
Sensor equipment rates	0.5%, 1%, 2%, 4%		

## 3.3 Rainfall data

To provide reference rainfall fields for the simulation experiments radar data from the C-band instrument at Ummendorf are utilised. The data were provided by the German Weather Service (DWD) as raw reflectivities with a spatial polar resolution of 1 km × 1° and a time discretisation of 5 min. The reflectivities are converted into rainfall intensities applying the Marshall-Palmer Z-R relationship (Marshall and Palmer, 1948)

$$Z = 200 \cdot R^{1.60}, \tag{14}$$

where Z is the reflectivity in mm^6/m^3 and R the rainfall intensity in mm/h. The rainfall intensities are interpolated to a 1 km × 1 km rectangular grid. Here a simple a nearest neighbour approach is used for those raster cells containing no more than one radar point; otherwise the mean value from all available radar points within a raster cell is taken. The radar data have been corrected for attenuation and clutter errors (Krämer, 2008), although this was not really necessary for this exercise. A bias correction of the radar data e.g. using the observations has not been applied here.

The heavy summer storm lasting from 16 July 2002 to 19 July 2002 has been selected for this analysis. It is characterised by high temporal and spatial rainfall variability with "observed" rainfall sums over the four days between 31 mm and 125 mm sampled at the 14 rainfall station locations from the radar fields. Rainfall data are used at a temporal resolution of 5 min for this analysis. Figures 2 and 3 illustrate the spatial and temporal variability of the event based on the radar precipitation data. Most rainfall occurs at the second day of the event, however with a high temporal variability. Note also that there are quite pronounced differences between the time series of areal rainfall for the four selected catchments. For further analyses only the time steps from 51 to 700 of the event are considered (see Fig. 3), which cover the most significant rainfall period.
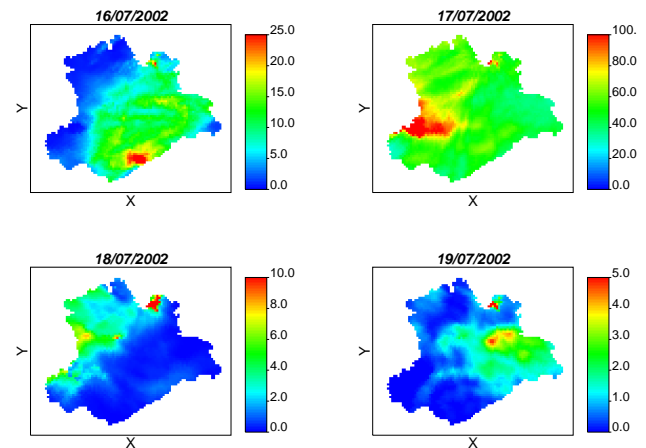

Fig. 2. Spatial distribution of the reference radar rainfall in mm/d for the Bode river basin accumulated over the four days of the storm event.
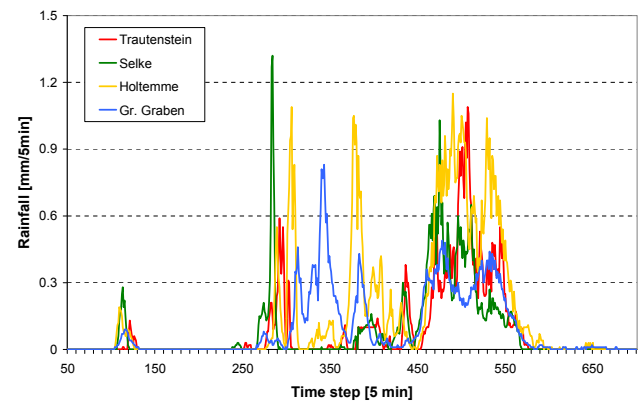
Fig. 3. Time series of areal rainfall derived from the reference radar fields for the selected four catchments considering only the time period of the event with significant rainfall over the particular catchments between time steps 51 and 700.
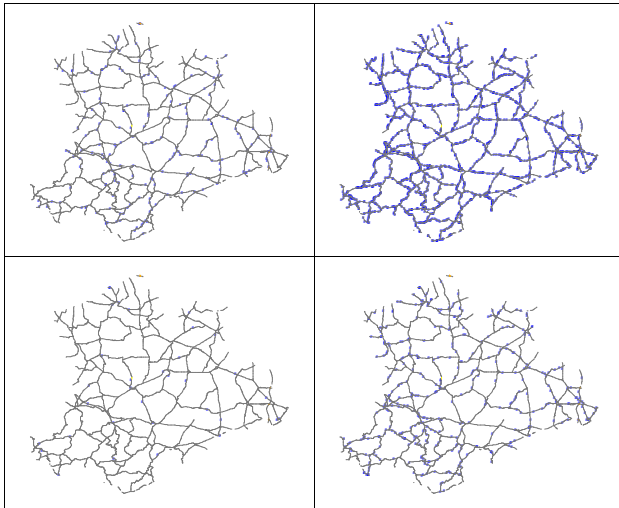
Fig. 4. Example for car distributions considering traffic with a sensor equipment rate of 1% on the road network during day (upper row) and night (lower row). On the left is the situation at one particular sampling time (5 min time step), on the right is the information of an accumulated sampling result of one hour.

4 Application and results

4.1 Car networks

Figure 4 shows examples of the distribution of cars considering a sensor equipment rate of 1% during day and night at one time step and accumulated over 1 h, respectively. The different car densities between day and night time are clearly visible. The figure also suggests that with a dynamic network the number of observation points in space can be virtually increased if measurements are integrated over time. Assuming that the change of the car locations with time (i.e. the speed of the cars) is greater than the change of the rainfall intensity with time at one location a moving network would provide advantages compared to a stationary network of the same density. In fact one car could measure rainfall at several locations before the rainfall intensity changes.

If the roads are now travelled by cars equipped with rain sensors at a rate of 1% and rainfall is sampled every 5 min, the absolute car densities for the Bode river basin can be calculated as given in Table 2. The values in this table are the result of empirical counting of discrete car positions in the whole area within a time interval of one hour for three time phases of the day (day, evening, night). The values in the second column are computed by normalizing the first values by the whole area, resulting in an average car position density per hour and km^2. For different equipment rates the values in Table 2 have to be multiplied with these rates. Note, that the figures in Table 2 are in fact the number of car locations (and not the number of cars) counted within one hour for a discrete observation time interval of 5 min. The difference is

Table 2. Number of cars (i.e. number of car positions accumulated over 5 minute intervals per hour) for the whole Bode river basin considering highways and federal roads and a sensor equipment rate of 1%.

Time of day	Absolute number of cars [h^{-1}]	Specific number of cars [h^{-1} km^{-2}]
Day (6 a.m.–5 p.m.)	2174	0.66
Evening (5 p.m.–10 p.m.)	1423	0.43
Night (10 p.m.–6 a.m.)	360	0.11

visualized in Fig. 4, where on the left panels the real number of cars is shown and on the right panels the integration of car positions used for observations over one hour.

Only the highways and federal roads having a total length of about 1300 km within the Bode river basin are used in this study. To the minor roads, which are not considered here, belong the state and district roads with a total length of 3100 km as well as the side roads with a total length of 10 300 km. Those figures give an impression what potential would be available additionally for measuring rainfall with car networks in this region.
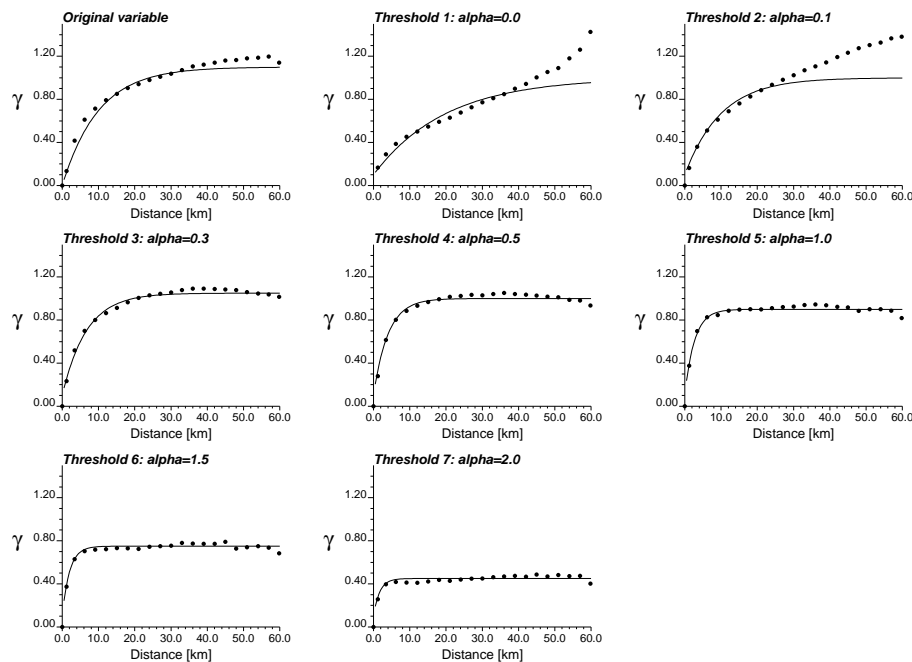
4.2 Rainfall sampling and variogram inference

Rainfall for the stationary gauge and moving car networks is sampled every 5 min from the reference radar field from those raster cells were rain gauges and cars, respectively, are located (cp. Fig. 1, Fig. 4). For the stationary gauges an exact rainfall measurement without errors is postulated. For rainfall observations by cars imprecision is assumed, which is related to various factors e.g. a restricted number of windscreen wiper frequency classes, non-unique features for different car types, calibration uncertainties etc. This is taken into account by a limited number of rainfall classes for which the sampling can be made. Rainfall observations based on 4 classes, corresponding to less accurate measurements ("more old fashioned cars") and 10 classes, corresponding to rather precise measurements ("more modern cars") are analysed here. The definition of the rainfall intervals has been done here quite subjectively based on simple assumptions about wiper frequencies and a rough examination of several observed storms in the Bode region. Table 3 lists the assumed rainfall ranges for the two classification schemes.

Prior to the application of the geostatistical interpolation methods variograms have to be estimated, which characterise the spatial variability of the rainfall fields. The variograms are inferred here a priori from the radar data. This is a certain simplification as the radar data represent the truth, which is usually not known. However, since the focus of the analysis is on relative comparisons of the performance between the stationary gauge network and different dynamic car networks

Table 3. Rainfall discretisation into classes, thresholds used for interpolation with indicator kriging (IK) and corresponding variogram parameters. For ordinary kriging (OK) the variogram No. 0 without threshold is applied. The upper bounds are included in the classes.

| No	Classes for IK with 10 intervals [mm/5 min]	Classes for IK with 4 intervals [mm/5 min]	Threshold [mm/5min]	Nugget [–]	Sill [–]	Effective Range [km]
0	–	–	None	0.0	1.1	30
1	0.0	0.0	0.0	0.1	0.9	60
2	0.0 – 0.1	–	0.1	0.1	0.9	30
3	0.1–0.3	–	0.3	0.1	0.95	20
4	0.3–0.5	0–0.5	0.5	0.1	0.9	12
5	0.5–1.0	–	1.0	0.1	0.8	8
6	1.0–1.5	–	1.5	0.1	0.65	6
7	1.5–2.0	0.5–2.0	2.0	0.1	0.35	5
8	2.0–3.0	2.0–5.0				
9	3.0–4.0					
10	4.0–5.0					

Fig. 5. Average standardized variogram of the original variable rainfall and indicator variograms for 7 rainfall thresholds given in mm/5 min estimated from observed radar data.

and the variograms are used uniquely for both networks this is tolerable.

Average experimental variograms are calculated based on Eq. (5) for rainfall as continuous variable $Z(u)$ and for different rainfall indicator variables $I_{\alpha k}(u)$ with $k = 1,...,K$. For the estimation of the indicator variograms radar rainfall is discretised into seven classes taking into account the predefined rainfall observation intervals. The combination of a nugget effect with an exponential model is fitted manually to all experimental variograms. Figure 5 shows the experimen-

tal and fitted theoretical variograms for $Z(u)$ and $I_{\alpha k}(u)$ for all selected thresholds. Table 3 lists the estimated variogram parameters and indicates which variograms are used for OK and for IK considering either 4 classes or 10 classes. Comparing the indicator variograms it can be seen that the range but also the sill decrease with increasing rainfall thresholds. Larger rainfall intensities have smaller extent which explains the decreasing range. For the highest thresholds there are only a few observed values beyond the threshold, which explains the low variance. For that reason variograms with
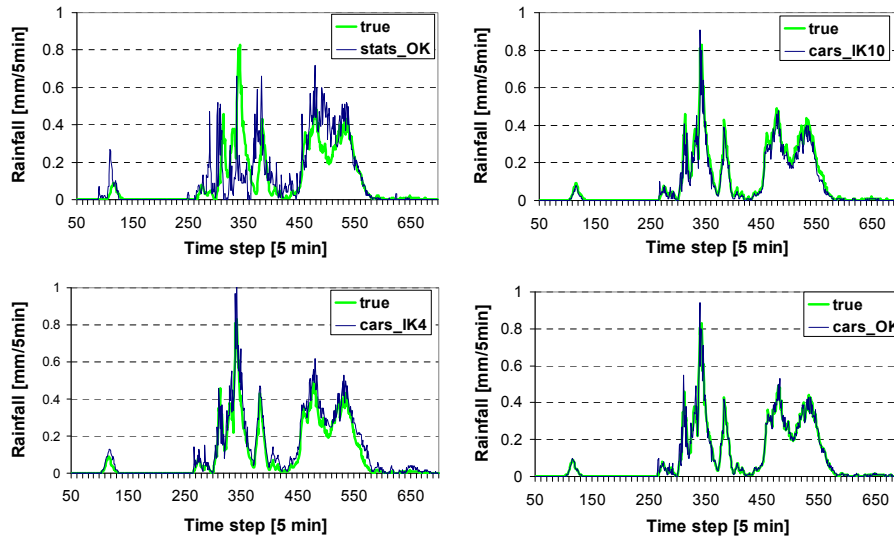
Fig. 6. Areal rainfall time series for the Gr. Graben catchment estimated from the stationary gauge network using ordinary kriging (stats_OK) and estimated from the car network with 4% sensor equipment rate using indicator kriging with 4 rainfall classes (cars_IK4), using IK with 10 rainfall classes (cars_IK10) and using OK without rainfall classification (cars_OK). Reference areal rainfall is from radar data (true).
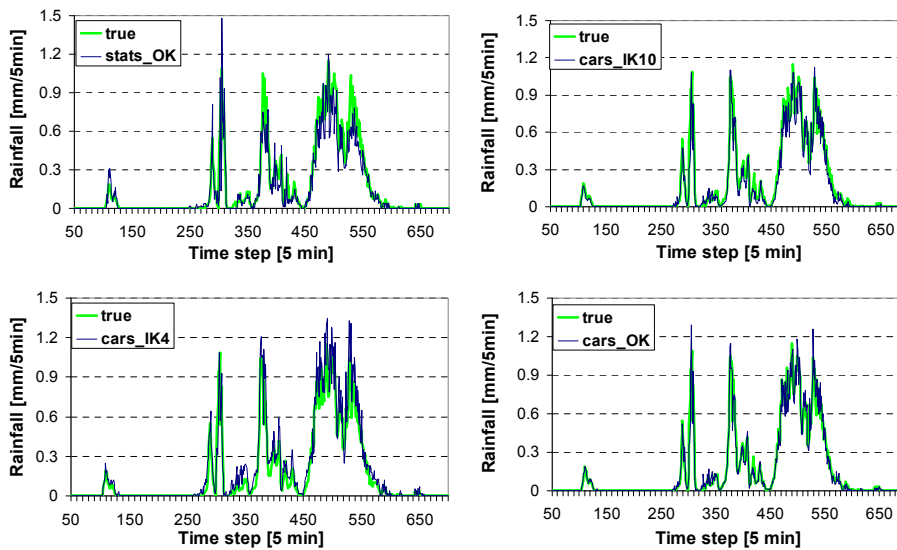

Fig. 7. Areal rainfall time series for the Holtemme catchment estimated from the stationary gauge network using ordinary kriging (stats_OK) and estimated from the car network with 4% sensor equipment rate using indicator kriging with 4 rainfall classes (cars_IK4), using IK with 10 rainfall classes (cars_IK10) and using OK without rainfall classification (cars_OK). Reference areal rainfall is from radar data (true).

higher thresholds than $\alpha = 2.0$ mm/5 min are not estimated. For interpolation of indicator variables with higher thresholds the last estimated indicator variogram is then applied (see Table 3).

4.3 Comparisons of areal rainfall estimation

Figures 6 and 7 show exemplarily areal rainfall time series for the Gr. Graben and the Holtemme catchments estimated from the rain gauge network and the car network with 4%

sensor equipment rate. For the 800 km² large Gr. Graben catchment the areal rainfall estimation using the stationary rain gauge network is poor. All approaches employing the car network provide significantly better areal rainfall estimation. This is not surprising considering that no stationary gauges are available within the basin boundaries for that catchment. For the Holtemme catchment areal rainfall estimation from the car network using IK4 provides slightly better results compared to the estimation from the rain gauge
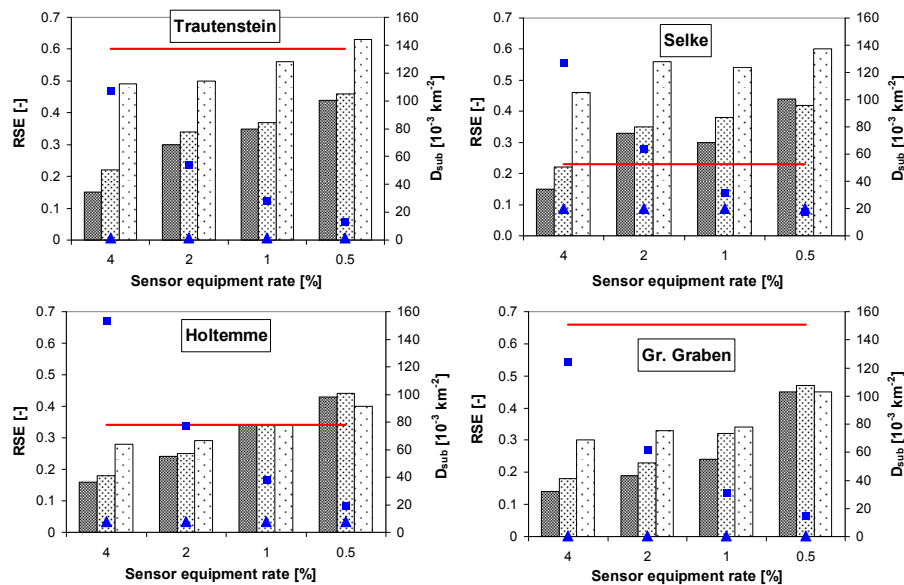
Fig. 8. Performance of areal rainfall estimation from the rain gauge network (horizontal red line) and from the 4 car networks with different sensor equipment rates (bars). For interpolation of the car rainfall observations IK4 (heavy dotted bars), IK10 (medium dotted bars) and OK (light dotted bars) are used. In addition the average network densities D_{sub} for the station network (blue triangles) and the car networks (blue squares) for each subbasin are provided.

network. Using IK10 however improves significantly the areal rainfall estimation. OK used as reference method assuming no errors in rainfall observation from car observatories gives here only an additional rather small improvement.

In Table 4 all calculated performance measures for areal rainfall estimation are listed considering the four catchments, the different car sensor equipment rates and interpolation methods. Table 5 shows the corresponding kernel density calculations of the networks. In addition, a visual comparison of the relative standard errors for areal rainfall estimation and the associated network densities is given in Fig. 8. For the largest test catchment Gr. Graben all car networks perform significantly better compared to the station network no matter what sensor equipment rate or interpolation approach is used. The error from the station network is highest (RSE = 0.66) and the station density is lowest ($D_{sub} = 0.6$) compared to the other catchments. For the smallest catchment Trautenstein the error from the station network (RSE = 0.6) is still almost always higher than the errors from the car networks, which is also related to a very low station density ($D_{sub} = 1.5$). For the Holtemme catchment the error from the station network is smaller (RSE = 0.34) as consequence of a higher rain gauge density ($D_{sub} = 7.7$). For that catchment, the car networks perform better only for the larger sensor equipment rates of 2% and 4% although the car network density is always superior compared to the station density. The car network for the Selke catchment can hardly do better than the traditional rain gauge network although the average car density is higher for all sensor equip-

ment rates which are greater or equal to 1%. One reason for that could be that the locations of the rain gauges in the Selke basin are optimal compared to the locations of the cars, which are forced on roads with possibly disadvantageous positions e.g. in valleys.

From Table 4 it can bee seen, that the bias for the indicator approaches is usually higher compared to OK. This comes likely from the suboptimal rainfall classification, required by the fixed rainfall observation classes e.g. according to the wiper frequencies. This rainfall discretisation does neither guarantee an equal number of values per class nor an equal distribution of the values within the classes. Especially the maximum limit of the highest interval is a critical assumption (cp. Table 3). All this leads to a biased estimation of the mean value from the cumulative distribution function, which has been obtained from the interpolation of the indicators using Eq. (8). This bias decreases usually with increasing number of classes (cp. results for IK4 and IK10 in Table 4).

Cars can only travel on roads. So the car precipitation network is restricted to roadways. Considering the preferred location of road lines e.g. in valleys and the relation of rainfall to elevation (i.e. on average higher rainfall for higher locations) the car network forced on roads might e.g. underestimate precipitation. For that reason one additional computer experiment is carried out generating cars anywhere within the catchment area ignoring roadways. The results are shown exemplarily for two sensor equipment rates in Fig. 9. It becomes clear, that a random positioning of the cars within the catchments does not improve the estimation performance for

Table 4. Performance of areal rainfall estimation based on the rain gauge network (Stations) and the car networks (Cars) with different sensor equipment rates (SER) in comparison with the radar reference rainfall. For interpolation ordinary kriging (OK) and indicator kriging (IK) with 4 and 10 rainfall classes are used. All values between time steps 51 and 700 with reference rainfall intensity greater or equal to 0.1 mm/5min are considered for error calculations.

	Stations OK			Cars IK4			Cars IK10			Cars OK		
SER [%]	Bias [mm]	RSE [–]	Cor [–]	Bias [mm]	RSE [–]	Cor [–]	Bias [mm]	RSE [–]	Cor [–]	Bias [mm]	RSE [–]	Cor [–]
Catchment Trautenstein (39 km^2)												
4	−6.7	0.60	0.58	15.1	0.49	0.87	−5.9	0.22	0.96	0.0	0.15	0.98
2	−6.7	0.60	0.58	12.1	0.50	0.80	−7.7	0.34	0.90	−1.4	0.30	0.91
1	−6.7	0.60	0.58	13.5	0.56	0.78	−6.0	0.37	0.87	−0.7	0.35	0.87
0.5	−6.7	0.60	0.58	12.0	0.63	0.69	−6.7	0.46	0.80	−3.2	0.44	0.81
Catchment Selke (102 km^2)												
4	0.1	0.23	0.95	16.9	0.46	0.90	−4.7	0.22	0.96	−0.9	0.15	0.98
2	0.1	0.23	0.95	17.1	0.56	0.80	−4.8	0.35	0.88	−0.9	0.33	0.89
1	0.1	0.23	0.95	14.5	0.54	0.78	−4.5	0.38	0.86	0.4	0.30	0.90
0.5	0.1	0.23	0.95	15.8	0.60	0.76	−3.1	0.42	0.81	0.4	0.44	0.81
Catchment Holtemme (167 km^2)												
4	−15.8	0.34	0.86	17.5	0.28	0.95	−10.3	0.18	0.97	−1.6	0.16	0.97
2	−15.8	0.34	0.86	12.6	0.29	0.92	−15.2	0.25	0.94	−4.9	0.24	0.93
1	−15.8	0.34	0.86	4.4	0.34	0.88	−20.2	0.34	0.89	−8.1	0.34	0.87
0.5	−15.8	0.34	0.86	−6.6	0.40	0.79	−28.1	0.44	0.82	−15.7	0.43	0.78
Catchment Gr. Graben (812 km^2)												
4	−4.1	0.66	0.16	12.9	0.30	0.91	−6.9	0.18	0.95	−1.1	0.14	0.95
2	−4.1	0.66	0.16	12.3	0.33	0.84	−7.2	0.23	0.89	−1.2	0.19	0.90
1	−4.1	0.66	0.16	7.7	0.34	0.75	−10.7	0.32	0.80	−4.1	0.24	0.86
0.5	−4.1	0.66	0.16	3.4	0.45	0.50	−13.3	0.47	0.52	−6.2	0.45	0.53

Table 5. Kernel network densities D_{sub} for the stationary rain gauge network (Stations) and averaged over time for the dynamic car networks (Cars) for the four selected catchments considering different sensor equipment rates (SER).

Catchment	D_{sub} Stations [10^{-3} km^{-2}]	D_{sub} Cars with SER of [10^{-3} km^{-2}]			
		4%	2%	1%	0.5%
Trautenstein	1.5	107.3	53.9	28.4	13.3
Selke	20.1	127.6	64.4	31.5	17.1
Holtemme	7.7	153.1	77.3	38.3	19.3
Gr. Graben	0.6	124.7	62.0	31.0	15.4

areal rainfall. For the Selke basin even the opposite seems true. For those four catchments it can be concluded that using the roadways for measuring rainfall is not suboptimal compared to a pure random network of the same density. Still, this does not disprove the possibility that the existent station

locations are optimal compared to a random distribution of gauges.

The error for areal rainfall estimation from the car networks seems not strongly related to the catchment size at least for considering the interpolation approaches IK10 and OK. The interpolation based on only 4 rainfall intervals using IK4, however, tends to give smaller estimation errors for the two larger catchments, Gr. Graben and Holtemme, compared to the two smaller ones, Selke and Trautenstein. Although for the largest basin (Gr. Graben) IK4 produces higher errors compared to the second largest basin (Holtemme).

Figure 10 shows the absolute number of cars together with observed and estimated areal rainfall time series for the Selke basin. It becomes clear, that there is a high variability in the car network density with changing levels for day, evening and night times. The figure shows also, that the car network density is poor for the largest peak of the event (during night time) which leads to a significant underestimation of rainfall. This could also contribute to the superiority of the actual station network in this catchment.
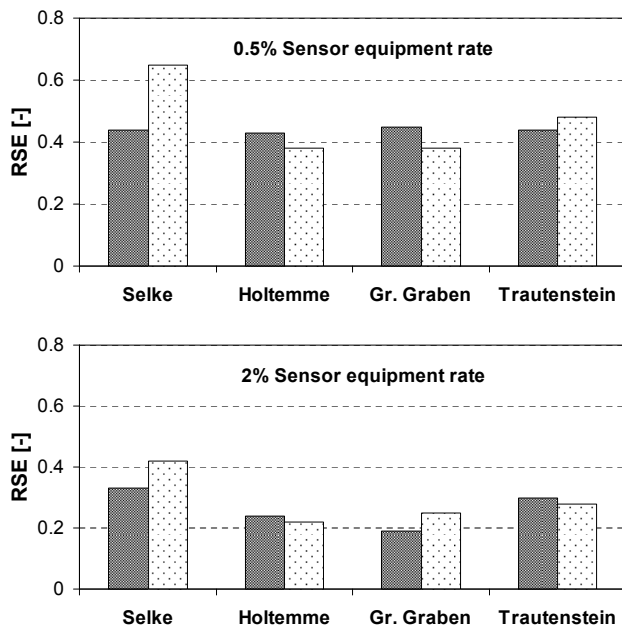
Fig. 9. Relative standard errors for areal rainfall estimation using OK interpolation with regular moving cars on roads (heavy dotted) and randomly redistributed cars (light dotted) for two sensor equipment rates.
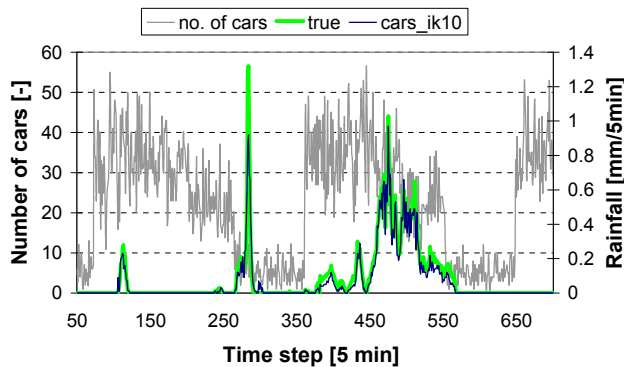
Fig. 10. Absolute number of cars (no. of cars) generated on roads in the Selke catchment with 4% sensor equipment rate in comparison with time series of observed (true) and estimated areal rainfall (cars_ik10).

5 Summary and conclusions

The main objective of this paper was to present a new approach for rainfall observation. The idea was to consider motorcars as moving rain gauges with windscreen wipers as sensors to detect precipitation. This study has explored theoretically the benefits of such an approach. A simple traffic model was applied to generate motorcars on roads in the 3300 km² Bode river basin. Rainfall from radar data fields were sampled with a conventional rain gauge network and with several dynamic networks consisting of moving motor-

cars. Calculated areal rainfall for four different catchments using ordinary kriging and indicator kriging was compared against the reference values from radar observations. The main results can be summarised as follows:

1. In general, the idea of using cars as rain gauges is theoretically feasible and would improve the assessment of areal rainfall compared to stationary gauge networks.

2. The error for areal rainfall estimation depends on the network density and decreases with growing number of cars. Considering the current traffic density obtained from statistical data a small sensor equipment rate between 0.5 and 4% was sufficient for areal rainfall estimation in the selected mesoscale catchments of the Bode river basin.

3. The imprecision in rainfall observation by the cars can be considered in areal estimation using indicator kriging based on rainfall classes. For the Gr. Graben and the Trautenstein catchments even as little as 4 discrete intervals for rainfall observation provide better areal rainfall estimation then the stationary gauge network with continuous observations. Using 10 rainfall intervals gives often almost as good results as with assumed continuous rainfall observation by the car network.

4. A problem with the imprecision of rainfall observation using cars is the prior definition of suitable rainfall intervals according to the expected range of the event. This leads likely to an additional bias in rainfall estimation.

5. A simulation study has shown that locations of the main roads have no disadvantageous effect on the position of cars regarding rainfall observation compared to a theoretical random distribution of cars in space with the same network density.

6. The results for the Selke catchment indicate that there might be optimal positions for rain gauges which reduce the required network density significantly compared to an arbitrary location of the gauges.

For this pilot study valid relationships between wiper speed and rainfall rate were assumed considering the point observation error only in a very simple way. So, the focus of the study was on interpolation uncertainty given different network densities (e.g. how many cars are needed) and regarding the required discretisation of the rainfall sensors in cars. In order to quantify point measurement errors using the cars as rain gauges and to confirm assumed discretisation accuracies for car rainfall observations practical experiments have to be carried out in the next step.

Future investigations should also consider different rainfall events, other regions and complete road networks to further test the presented ideas. Especially higher sampling rates for the cars and the integration of observations over time

need to be explored. The findings also suggest a revisiting of the rain gauge network optimisation problem. Future research might also consider a merging of rainfall information from the three sources radar, stations and cars. Eventually, practical feasibility studies are required to answer questions about applicability of the ideas in praxis.

Edited by: A. Shamseldin

References

Chiang, Y.-M., Hsu, K.-L., Chang, F.-J., Hong, Y., and Sorooshian, S.: Merging multiple precipitation sources for flash flood forecasting, J. Hydrol., 340, 183–196, 2007.

Deutsch, C. V. and Journel, A. G.: GSLIB: Geostatistical software library and user's guide, Oxford University Press, New York, 340 pp., 1992.

Ehret, U., Götzinger, J., Bárdossy, A., and Pegram, G. G. S.: Radar-based flood forecasting in small catchments, exemplified by the Goldersbach catchment, Germany, Intl. J. River Basin Management, 6, 323–329, 2008.

EUROSTAT: European Commission Statistics Database, http://epp.eurostat.ec.europa.eu, last access: 2 May 2009.

Goovaerts, P.: Geostatistics for natural resources evaluation, Oxford University Press, New York, Oxford, 483 pp., 1997.

Goovaerts, P.: Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, J. Hydrol., 228, 113–129, 2000.

Goudenhoofdt, E. and Delobbe, L.: Evaluation of radar-gauge merging methods for quantitative precipitation estimates, Hydrol. Earth Syst. Sci., 13, 195–203, doi:10.5194/hess-13-195-2009, 2009.

Grimes, D. I. F. and Diop, M.: Satellite-based rainfall estimation for river flow forecasting in Africa. I: Rainfall estimates and hydrological forecasts, Hydrol Sci. J., 48, 567–584, 2003.

Haberlandt, U.: Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event, J. Hydrol., 332, 144–157, 2007.

Hasse, L., Grossklaus, M., Uhlig, K., and Timm, P.: A Ship Rain Gauge for Use in High Wind Speeds, J. Atmos. Ocean. Tech., 15, 380–386, 1998.

Isaaks, E. H. and Srivastava, R. M.: Applied Geostatistics, Oxford University Press, New York, 1989.

Journel, A. G.: Non parametric estimation of spatial distributions, Math. Geol., 15, 445–468, 1983.

Krajewski, W. F. and Smith, J. A.: Radar Hydrology: rainfall estimation, Adv. Water Resour., 25, 1387–1394, 2002.

Krämer, S.: Quantitative Radardatenaufbereitung für die Niederschlagsvorhersage und die Siedlungsentwässerung, Mitteilungen, Heft 92, Inst. of Water Resources Management, Leibniz University of Hannover, Hannover, 391 pp., 2008.

Leijnse, H., Uijlenhoet, R., and Stricker, J. N. M.: Rainfall measurement using radio links from cellular communication networks, Water Resour. Res., 43, W03201, doi:03210.01029/02006WR005631, 2007.

Marshall, J. S. and Palmer, W. M.: The distribution of raindrops with size, J. Meteorol., 9, 327–332, 1948.

Matheron, G.: The Theory of Regionalized Variables and its Applications, Les Cahiers du Centre de Morphologie Mathématique, Fasc. 5, 1971.

Messer, H., Zinevich, A., and Alpert, P.: Environmental Monitoring by Wireless Communication Networks, Science, 312, p. 713, doi:10.1126/science.1120034, 2006.

Seo, D.-J., Krajewski, W. F., and Bowles, D. S.: Stochastic interpolation of rainfall data from raingages and radar using cokriging: 1. Design of experiments, Water Resour. Res., 26, 469–477 (489WR02984), 1990.

Silverman, B. W.: Density estimation for statistics and data analysis, Chapman and Hall, London, 1986.

Smith, J. A., Baeck, M. L., Meierdiercks, K. L., Miller, A. J., and Krajewski, W. F.: Radar rainfall estimation for flash flood forecasting in small urban watersheds, Adv. Water Resour., 30, 2087–2097, 2007.

Stefanidis, A. and Nittel, S.: GeoSensorNetworks, CRC Press, Boka Raton, Florida, 2004.

Wardah, T., Abu Bakar, S. H., Bárdossy, A., and Maznorizan, M.: Use of geostationary meteorological satellite images in convective rain estimation for flash-flood forecasting, J. Hydrol., 356, 283–298, 2008.

Yuter, S. E. and Parker, W. S.: Rainfall Measurement on Ship Revisited: The 1997 PACS TEPPS Cruise, J. Appl. Meteorol., 40, 1003–1018, 2001.