

# A probabilistic approach to the prediction of area weather events, applied to precipitation

Bjoern Kriesche<sup>a,\*</sup>, Reinhold Hess<sup>b</sup>, Bernhard K. Reichert<sup>b</sup>, Volker Schmidt<sup>a</sup>

<sup>a</sup>Ulm University, Institute of Stochastics, 89069 Ulm, Germany

<sup>b</sup>Deutscher Wetterdienst, Frankfurter Straße 135, 63067 Offenbach, Germany

---

## Abstract

In meteorology it is important to compute the probabilities of certain weather events occurring. There are a number of numerical and statistical methods for estimating the probability that a weather event occurs at a fixed location (a point). However, there are no widely applicable techniques for estimating the probability of such an event occurring in a geographical region (an area). In this paper, we propose a model-based approach for the computation of area probabilities using point probabilities. We develop this approach in the context of estimating the probability of the meteorological event ‘occurrence of precipitation’. We treat the point and area probabilities as coverage probabilities of a germ-grain model, where the grains can roughly be interpreted as precipitation cells. The germ-grain model is completely characterized by a sequence of local intensities and a grain size. We compute these model characteristics using available point probabilities. A non-negative least-squares approach is used to determine the local intensities and a semivariogram estimation technique is used to find the grain size. We are then able to determine area probabilities either analytically or by repeated simulation of the germ-grain model. We validate our model, using radar observations to assess the precision of the computed probabilities.

*Keywords:* Probabilistic weather prediction, Germ-grain model, Cox point process, Semivariogram estimation, Forecasting, Area weather event

---

## 1. Introduction

Meteorological services such as Deutscher Wetterdienst (DWD) are responsible for providing weather warnings. In order to give timely and reliable warnings, they need to be able to accurately estimate the probability that a potentially harmful meteorological event, such as wind gusts or precipitation exceeding some threshold, occurs in a specific geographical area. Such an event is said to occur if it occurs somewhere in the area. We call the corresponding probability an area probability.

The existing methodologies used by DWD and other meteorological services are designed to estimate point probabilities rather than area probabilities. That is, they estimate the probability of a weather event happening at a specific point, rather than somewhere in a wider area. Area

---

\*Corresponding author

Email address: [bjoern.kriesche@uni-ulm.de](mailto:bjoern.kriesche@uni-ulm.de) (Bjoern Kriesche)

probabilities are quantitatively different from their corresponding point probabilities. For example, the probability that an event occurs somewhere in an area is greater than the probability it occurs at a given point within the area.

Available point probabilities generally correspond to geographically distinct locations. Because of their spatial correlation structure, these probabilities provide information not only about the considered locations but also about their surroundings. This means that they contain significant information about area probabilities. However, it is not clear how this information can be used most effectively. There is no known analytical formula for deriving area probabilities from a collection of geographically distinct point probabilities. In Epstein (1966) and Krzysztofowicz (1998), theoretical relationships are established between point and area probabilities. However, these approaches assume circular precipitation cells and make a number of very restrictive assumptions, including circular forecast areas and uniformly distributed precipitation cells. These assumptions imply that the point probabilities are equal for all points in the forecast area and a certain neighborhood, which makes the model applicable to a regional scale only. In addition, in order to compute area probabilities, these approaches require additional information about precipitation cell sizes to be specified by the forecaster. This, however, prevents an automated generation of weather warnings, where an algorithmic computation of area probabilities based solely on point probabilities is desired.

In recent decades, there has been an increasing interest in the analysis, modeling and simulation of meteorological data sets using spatial stochastic models, especially spatial point processes. Some current examples can be found in Elsner et al. (2013), Karpman et al. (2013) or Kriesche et al. (2014). In the present paper, we continue work in this direction by introducing a statistical method for computing area precipitation probabilities from point precipitation probabilities using a spatial stochastic model for the representation of precipitation cells. In the literature several stochastic approaches for the modeling of rain events are proposed. Besides single-site and multi-site models where mainly the temporal development of rain events is considered (Cowpertwait et al., 2011; Ramesh et al., 2012; Kaczmarska et al., 2014), some approaches for the continuous spatial modeling (Rodriguez-Iturbe et al., 1986) or the continuous spatial-temporal modeling of rain events have been introduced, too (see e.g. Cox and Isham, 1988; Cowpertwait, 1995; Northrop, 1998; Wheeler et al., 2000, 2005; Segond and Onof, 2009). A good review of the basic models can be found in Onof et al. (2000). Although none of these papers deals with area probabilities, they still provide some valuable ideas for the modeling of rain cells, e.g., the modeling of cell centers by Poisson-based point process and the consideration of circular rain cells. However, there are some strong limitations that make the provided continuous spatial models inappropriate for the purposes of the present paper. On the one hand, spatial stationarity of rain events is assumed, which restricts the modeling of rain fields to relatively small regional areas. A non-stationary approach is only provided for multi-site models, see e.g. Wheeler et al. (2005). On the other hand, these models can only be fitted to rain fields obtained from radar data, which again prevents the automated generation of warning events based solely on available point probabilities.

In the present paper we extend the work performed in previous publications, to develop a more robust and less restrictive approach. More precisely, we consider a global model where the considered meteorological objects, here precipitation cells, are distributed inhomogeneously over the entire observation window (e.g., Germany). We are able to calculate all the model characteristics algorithmically from a sequence of point probabilities. In particular, no observed

precipitation rates are necessary. We can then compute area precipitation probabilities for arbitrary areas of interest within the observation window. The method proposed in this paper is expected to play a key role in fulfilling a specific requirement of DWD, which is the calculation and dissemination of area probabilities for hazardous weather warnings (precipitation, wind gusts, thunderstorms, etc.) for customer-specific areas using semi-automated warning systems.

Our paper is organized as follows. Sect. 2 describes the computation of point precipitation probabilities, which play a fundamental role in our paper. We give a brief overview of the methods used by DWD to derive probabilities for the occurrence of precipitation exceeding certain thresholds. We also describe the data used in this paper, which includes the point precipitation probabilities. In Sect. 3, we propose a new model-based approach to computing area probabilities. We represent both point and area probabilities as coverage probabilities of a germ-grain model. The germ-grain model is characterized by a sequence of local intensities and a grain size. The grains loosely model precipitation cells. We describe how our model characteristics can be computed using point probabilities provided by DWD. Area probabilities can be determined either directly or by repeated simulation of the germ-grain model. In Sect. 4, we carry out model validation. The area probabilities computed using our model are compared with observed relative frequencies of precipitation events based on radar observations. We give a summary of our results and discuss avenues for future research.

## **2. Point precipitation probabilities**

The exact relationship between point and area precipitation probabilities is unclear. However, it is clear that point probabilities for distinct locations contain valuable information on area precipitation events due to their spatial correlation.

### *2.1. Computation of probabilities for point events*

At DWD, point probabilities of given weather events occurring, such as precipitation or wind gusts exceeding a certain warning threshold, are estimated using a multistage procedure. In the first step, a preliminary deterministic weather forecast is obtained from the numerical model GME (Majewski, 1998; Majewski et al., 2002) and the Integrated Forecasting System of the European Centre for Medium-Range Weather Forecasting (IFS/ECMWF). These data have systematic errors which arise from uncertainties about the initial weather conditions, inaccuracies in the numerical model simulating weather evolution, and the failure of the numerical model to adequately represent synoptic observations (the error of representativity). These errors are reduced in a statistical post-processing step at DWD. The applied method is model output statistics (MOS), see (Knüpfper, 1996). For a general introduction to MOS, we refer to Wilks (1995). The applied MOS system uses historical observations from about 1600 synoptic weather stations worldwide to estimate the systematic errors of GME and IFS/ECMWF at these locations and to correct the operational forecasts. More precisely, the input data are time series of both historical observations and the output of the numerical models. These time series range from 2001 up to 2014 currently. Changes in the numerical models are accounted for via binary predictors and a higher weighting of the most recent time period. All in all, roughly 150 forecast variables are computed with various reference periods and with forecasting times up to ten days ahead. Each dependent variable (predictand) is estimated based on the most relevant independent variables (predictors) via a stepwise linear regression. The predictors based on the output of the numerical models tend to have the highest explanatory power. However, recent meteorological observations

can also have considerable explanatory power, especially for small forecast ranges. For example, consider the predictand ‘precipitation amount during 1 hour’. Usually the forecast of this value from the numerical model is a very good predictor. However, other model variables and the latest observed rain rates may also have significant explanatory power. DWD uses a two stage MOS procedure. First, the forecasts from the GME model and IFS/ECWMF are post-processed individually. Then, in a second step, the resulting forecasts are merged using MOS again in order to derive statistically optimal weights for the individual forecasts.

## 2.2. Description of data

In this paper, we consider point probabilities of precipitation exceeding 0.1 mm within one hour. These are generated for 503 German and Luxembourgian synoptic weather stations, including secondary national non-WMO (World Meteorological Organization) stations. The threshold of 0.1 mm was chosen in order to allow for a consistent validation with gauge adjusted radar analyses, see Sect. 4.1. A summer period from June 1 until July 31, 2012 and a winter period from November 1 until December 31 of the same year were selected in order to address seasonal changes in precipitation patterns. Re-forecasts of the two stage MOS system of DWD were started for each day of the selected time periods and the individually post-processed forecasts of the GME and IFS/ECMWF models were mixed. Seven forecast steps from 3 UTC (Universal Time, Coordinated) every three hours up to 21 UTC are used, each providing the point probabilities that precipitation exceeding an amount of 0.1 mm occurs during the preceding time period of one hour. Altogether, 854 forecast steps of precipitation were obtained, each consisting of 503 point probabilities, corresponding the 503 considered weather stations.

## 3. Area precipitation probabilities

In this section, a statistical method for the computation of area precipitation probabilities is introduced. In particular, we present a spatial stochastic model for the representation of both point and area precipitation probabilities as coverage probabilities of a germ-grain model. In the following, we always consider an arbitrary but fixed forecast step, which means that no temporal dynamic is taken into account in our model.

### 3.1. Stochastic model for the occurrence of precipitation

The occurrence of precipitation is one of the most complex processes considered in meteorology, which makes a precise computation of probabilities extremely difficult. As already mentioned in Sect. 2.1, point probabilities computed from numerical and statistical forecast models are subject to several sources of uncertainty concerning e.g. initial weather conditions or inaccuracies in the model equations due to discretization and physical parameterization. A statistical post-processing (e.g. by using MOS) reduces systematic errors of the numerical models but, however, cannot correct for random errors. To account for that, it seems natural to model precipitation probabilities as random variables and to consider the probabilities computed by DWD as possible realizations. According to this approach we consider an arbitrary probability space  $(\Omega, \mathcal{F}, P)$ , where each elementary event  $\omega \in \Omega$  corresponds to a possible realization of the numerical weather prediction model for the considered forecast step and vice versa.

To begin with, we consider the random field  $\{P_t, t \in W\}$  with  $P_t : \Omega \rightarrow [0, 1]$  modeling the (random) point precipitation probability at location  $t$  in a bounded and convex Borel set  $W \subset \mathbb{R}^2$ . In the example of application discussed in this paper,  $W$  is a rectangle comprising the boundaries

of Germany. To allow for a reasonable modeling of  $P_t$ , various assumptions have to be made. At first, we suppose that for each  $\omega \in \Omega$  the point precipitation probabilities  $P_{s_1}(\omega), \dots, P_{s_n}(\omega)$  are available for a sequence  $s_1, \dots, s_n$  of locations of interest in  $W$ , for example the locations of synoptic weather stations. Furthermore, we suppose an intuitive relationship between precipitation probabilities and the occurrence of precipitation, which constitutes our basic modeling assumption. We say that there is precipitation at location  $t \in W$  if and only if  $t$  is covered by a certain precipitation field. In general, precipitation fields can look very irregular, see for example Fig. 3 (b) in Sect. 4.1, and thus, some simplification seems to be helpful. We assume that a precipitation field can be represented as the union of random precipitation cells, with each of them having a random cell center and a random cell shape.

We supposed that precipitation cells, or to be more precise their cell centers, are formed in a random way according to a random intensity function  $\{\Lambda_t, t \in W\}$  of cell formation with  $\Lambda_t : \Omega \rightarrow [0, \infty)$  being the random intensity at location  $t \in W$ . It is clear that there is a close relationship between the characteristics of precipitation cells, i.e., the intensity function  $\{\Lambda_t, t \in W\}$  and the cell size on the one hand, and the field of point probabilities  $\{P_t, t \in W\}$  on the other hand. Therefore, a suitable modeling approach for  $\{\Lambda_t, t \in W\}$  has to be found first. Unfortunately, both  $\{P_t, t \in W\}$  and  $\{\Lambda_t, t \in W\}$  are unknown in general, only realizations of the random variables  $P_{s_1}, \dots, P_{s_n}$  are available. To account for the fact that meteorological information is only given for the locations  $s_1, \dots, s_n$ , we make the simplifying assumption that  $\{\Lambda_t, t \in W\}$  is a piecewise constant random function, which takes a constant value within a natural neighborhood of each location  $s_i$  for  $i = 1, \dots, n$ . For that purpose, we consider the Voronoi tessellation  $\{V(s_1), \dots, V(s_n)\}$  of  $s_1, \dots, s_n$  in  $W$ , where the Voronoi cell  $V(s_i)$  of  $s_i$  in  $W$  is defined as

$$V(s_i) = \{x \in W : |x - s_i| < |x - s_j| \text{ for all } j = 1, \dots, n \text{ with } j \neq i\}, \quad (1)$$

for all  $i = 1, \dots, n$ . Consequently, we assume that the random intensity function  $\{\Lambda_t, t \in W\}$  of cell formation can be represented as

$$\Lambda_t = \sum_{i=1}^n A_i I_{V(s_i)}(t) \quad \text{for all } t \in W, \quad (2)$$

where  $I_V : W \rightarrow \{0, 1\}$  denotes the indicator of the set  $V \subset W$  and the random variables  $A_1, \dots, A_n : \Omega \rightarrow [0, \infty)$  can be interpreted as random local intensities for the formation of precipitation cells. In the next step, the modeling of the centers of precipitation cells itself is addressed, where we consider a random Cox point process  $\{X_i, i \geq 1\}$  in  $W$  with (random) intensity function  $\{\Lambda_t, t \in W\}$ , see e.g. Chiu et al. (2013) or Illian et al. (2008). The Cox process  $\{X_i, i \geq 1\}$  is defined on the extended probability space  $(\Omega \times \tilde{\Omega}, \mathcal{F} \otimes \tilde{\mathcal{F}}, Q)$ , where for each  $(\omega, \tilde{\omega}) \in \Omega \times \tilde{\Omega}$  the elementary event  $\omega$  identifies a possible realization of the numerical weather prediction model and  $\tilde{\omega}$  corresponds to the spatial distribution of precipitation cells.

After modeling the cell centers, the shape of precipitation cells has to be addressed. Due to the irregularity of precipitation fields (see e.g. Fig. 3 (b) in Sect. 4.1) it seems to be practically impossible to find a model for precipitation cells, which exactly matches real precipitation fields but is still easy to handle. Therefore, we make the simplifying assumption that there is precipitation at location  $t$  (i.e.,  $t$  belongs to a precipitation field, see above) if  $t$  is close enough to the center of at least one precipitation cell. Equivalently, we say that there is precipitation at  $t$  if  $t$  is

covered by a germ-grain model  $M$  defined on  $(\Omega \times \widetilde{\Omega}, \mathcal{F} \otimes \widetilde{\mathcal{F}}, Q)$ , which is given by

$$M = \bigcup_{i=1}^{\infty} b(X_i, R). \quad (3)$$

Here,  $b(x, r)$  denotes the two-dimensional ball with center at  $x \in \mathbb{R}^2$  and radius  $r > 0$  and the random variable  $R : \Omega \rightarrow (0, \infty)$  is interpreted as random spatial precipitation range. We want to emphasize that the germ-grain model  $M$  can only roughly be understood as a model for precipitation cells but that the modeling of precipitation probabilities as coverage probabilities can be quite accurate even if single realizations of  $M$  look atypically compared to observed precipitation fields. Note that given a particular realization  $\omega \in \Omega$ , the germ-grain model  $M$  is a well-known inhomogeneous Boolean model based on a Poisson point process with intensity function  $\{\Lambda_t(\omega), t \in W\}$  and with circular grains with radius  $r = R(\omega)$ , see e.g. Chiu et al. (2013).

### 3.2. Representation formulas for precipitation probabilities

We recall once more the key assumptions made in Sect. 3.1. We suppose that there is precipitation at location  $t \in W$  if and only if  $t$  is covered by the germ-grain model  $M$  given in (3), which is completely characterized by the sequence  $A_1, \dots, A_n$  of random intensities for the local formation of precipitation cells and the random spatial precipitation range  $R$ . Based on these assumptions, we get the following representation of  $\{P_t, t \in W\}$  as

$$P_t = P(t \in M | A_1, \dots, A_n, R) = 1 - \exp\left(-\sum_{i=1}^n A_i v_2(b(t, R) \cap V(s_i))\right) \text{ for all } t \in W, \quad (4)$$

where  $v_2(\cdot)$  denotes the two-dimensional Lebesgue measure. A particular advantage of the model proposed in this paper is the circumstance that an intuitive generalization of formula (4) to the representation of area precipitation probabilities is easily possible. Following the modeling ideas considered above, it is natural to say that there is precipitation somewhere within an arbitrary area  $B \subset W$  if and only if the distance of  $B$  to the center of at least one precipitation cell is smaller than the random precipitation range  $R$ , which is the case if and only if  $B$  intersects the germ-grain model  $M$ . To be more precise, the random area precipitation probability for any Borel set  $B \subset W$  is modeled by  $P(B \cap M \neq \emptyset | A_1, \dots, A_n, R)$ , which can be expressed as

$$P(B \cap M \neq \emptyset | A_1, \dots, A_n, R) = 1 - \exp\left(-\sum_{i=1}^n A_i v_2((B \oplus b(o, R)) \cap V(s_i))\right), \quad (5)$$

where  $o \in \mathbb{R}^2$  is the origin and  $A \oplus B = \{x + y, x \in A, y \in B\}$  denotes the Minkowski sum of any sets  $A$  and  $B$ . Note that the representation formula given in (5) can be derived as follows. We have

$$\begin{aligned} P(B \cap M \neq \emptyset | A_1, \dots, A_n, R) &= 1 - P(\#\{i : X_i \in B \oplus b(o, R)\} = 0 | A_1, \dots, A_n, R) \\ &= 1 - \exp\left(-\int_{B \oplus b(o, R)} \sum_{i=1}^n A_i \mathbb{1}_{V(s_i)}(t) dt\right) \\ &= 1 - \exp\left(-\sum_{i=1}^n A_i \int_{\mathbb{R}^2} \mathbb{1}_{(B \oplus b(o, R)) \cap V(s_i)}(t) dt\right) \\ &= 1 - \exp\left(-\sum_{i=1}^n A_i v_2((B \oplus b(o, R)) \cap V(s_i))\right), \end{aligned}$$

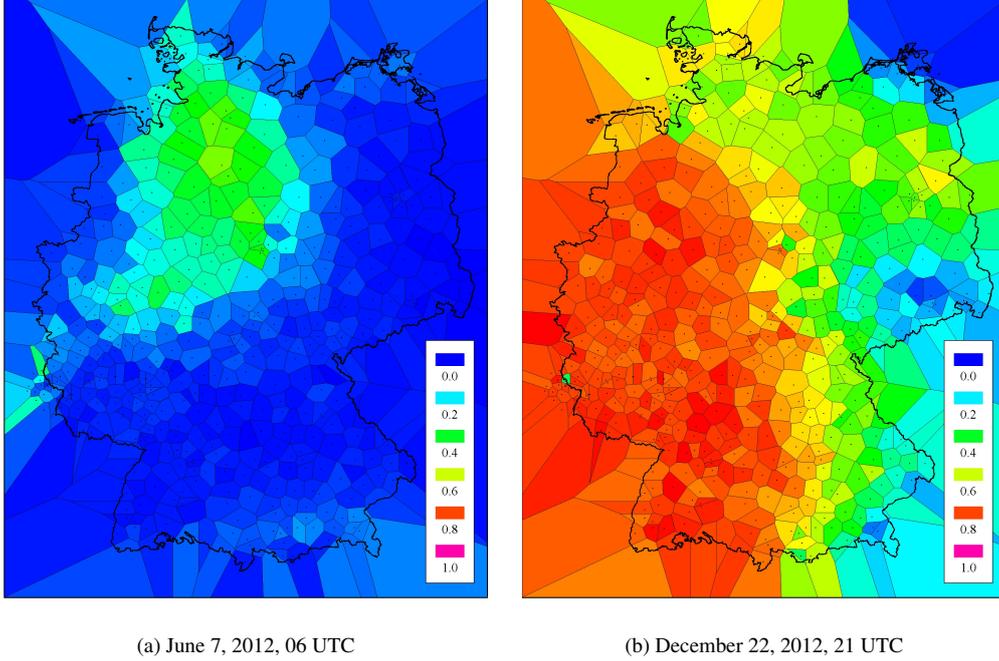


Figure 1: Available data for two selected forecast steps: the locations of the considered 503 weather stations and the corresponding Voronoi tessellation, where each Voronoi cell is colored according to the point precipitation probability at the corresponding weather station.

where  $\#A$  denotes the cardinality of a countable set  $A$ . The second equality is based on the property of  $\{X_i, i \geq 1\}$  being a Poisson point process if  $A_1, \dots, A_n$  and  $R$  are given. The representation formula of point probabilities given in (4) follows immediately from (5) by setting  $B = \{t\}$ .

### 3.3. Computation of model characteristics

In order to compute point and area probabilities according to (4) and (5) the random model characteristics  $A_1, \dots, A_n$  and  $R$  need to be determined first. For that purpose, available realizations of the point probabilities  $P_{s_1}, \dots, P_{s_n}$  are used. In applications, the computation of the random model characteristics and of area precipitation probabilities is performed in dependence of the realization of the numerical weather prediction model and therefore, an arbitrary but fixed  $\omega_0 \in \Omega$  is considered in the following. We introduce a simplifying notation for the corresponding realizations of  $P_{s_1}, \dots, P_{s_n}, A_1, \dots, A_n$  and  $R$ , that is  $p_{s_i} = P_{s_i}(\omega_0)$  and  $a_i = A_i(\omega_0)$  for  $i = 1, \dots, n$ , and  $r = R(\omega_0)$ . In our example of application,  $\omega_0$  is the particular realization of the global models GME and IFS/ECMWF that provides the basis for the computation of the available data described in Sect. 2.2. Fig. 1 illustrates the available data for two forecast steps, where each Voronoi cell  $V(s_i)$  is colored according to the given point probability  $p_i$  at location  $s_i$  of the corresponding weather station for  $i = 1, \dots, 503$ .

A simultaneous computation of the unknown characteristics  $a_1, \dots, a_n$  and  $r$  does not seem to be possible. Therefore, we propose a multi-step procedure that can be outlined as follows:

1. For each  $r' > 0$ , intensities  $a_1^{(r')}, \dots, a_n^{(r')}$  are computed from  $p_{s_1}, \dots, p_{s_n}$  under the condition that  $R$  is equal to  $r'$ .

2. The precipitation range  $r$  is computed as a function of  $p_{s_1}, \dots, p_{s_n}$  and the family of conditional intensities  $\{a_1^{(r')}, \dots, a_n^{(r')}, r' > 0\}$ .
3. Finally, the intensities  $a_1, \dots, a_n$  are computed as  $a_1^{(r)}, \dots, a_n^{(r)}$  by setting  $r' = r$  in step 1.

These steps are illustrated in detail in the following sections.

### 3.3.1. Computation of local intensities for the formation of precipitation cells

At first, the computation of conditional intensities based on a given precipitation range is proposed. We assume that for each  $r' > 0$  there is a sequence of nonnegative intensities  $a_1^{(r')}, \dots, a_n^{(r')} \geq 0$ , such that

$$p_{s_j} = P(s_j \in M | \{A_i\} = \{a_i^{(r')}\}, R = r') = 1 - \exp\left(-\sum_{i=1}^n a_i^{(r')} \nu_2(b(s_j, r') \cap V(s_i))\right) \quad (6)$$

for all  $j = 1, \dots, n$  according to (4). However, note that the characteristics  $a_1^{(r')}, \dots, a_n^{(r')}$  and  $r'$  are not necessarily suitable to describe point probabilities for locations  $t \notin \{s_1, \dots, s_n\}$ . It is clear that the system of equations stated in (6) is equivalent to

$$\log\left(\frac{1}{1-p_{s_j}}\right) = \sum_{i=1}^n a_i^{(r')} \nu_2(b(s_j, r') \cap V(s_i)) \quad \text{for all } j = 1, \dots, n, \quad (7)$$

which describes a system of  $n$  linear equations with  $n$  unknown variables  $a_1^{(r')}, \dots, a_n^{(r')}$ . Under the constraint that  $a_1^{(r')}, \dots, a_n^{(r')} \geq 0$ , however, an exact solution of (7) does not exist in general. For that reason, we suggest to compute  $a_1^{(r')}, \dots, a_n^{(r')}$  in a nonnegative least-squares sense, i.e.,

$$(a_1^{(r')}, \dots, a_n^{(r')}) = \underset{a_1, \dots, a_n \geq 0}{\operatorname{argmin}} \left\{ \sum_{j=1}^n \left( \log\left(\frac{1}{1-p_{s_j}}\right) - \sum_{i=1}^n a_i \nu_2(b(s_j, r') \cap V(s_i)) \right)^2 \right\}. \quad (8)$$

The right-hand side of (8) can be computed according to the algorithm given in Lawson and Hanson (1974), Chap. 23. In general  $a_1^{(r')}, \dots, a_n^{(r')}$  determined according to (8) are not an exact solution of (6). However, a comparison between the available probabilities  $p_{s_1}, \dots, p_{s_n}$  and probabilities computed according to (6) with intensities derived from (8) reveals only a negligible difference. In particular, no systematic deviation is observed.

In the next section, a method for the computation of the precipitation range  $r$  is proposed, such that (4) can be assumed to hold for all  $t \in W$ . When  $r$  has been determined, the intensities  $a_1, \dots, a_n$  are given by (8) with  $a_i = a_i^{(r)}$  for  $i = 1, \dots, n$ .

### 3.3.2. Iterative estimation of the semivariogram

Intuitively, the random precipitation range  $R$  should have a significant impact on the spatial correlation of  $\{P_t, t \in W\}$ , at least for pairs of points with a small distance. To quantify the degree of spatial dependence, we suggest to consider the semivariogram  $\gamma(\cdot)$  of  $\{P_t, t \in W\}$ , see e.g. Cressie (1993) and Fernández-Avilés et al. (2015). However, a meaningful estimation and analysis of  $\gamma(\cdot)$  based on realizations of  $P_{s_1}, \dots, P_{s_n}$  would only be possible if  $\{P_t, t \in W\}$  was

assumed to be second-order stationary and isotropic, in particular if  $\{P_t, t \in W\}$  had a constant mean function  $\{\mu_t, t \in W\}$  with  $\mu_t = \mathbb{E}P_t$ . This is an assumption, which can hardly be justified. Depending on the climatological area and the current weather conditions at each forecast step, the precipitation probabilities show a clear spatial trend, see e.g. the available data shown in Fig. 1.

A common approach to the estimation of the semivariogram  $\gamma(\cdot)$  in the non-stationary or non-isotropic case is to decompose  $\{P_t, t \in W\}$  into a deterministic mean function  $\{\mu_t, t \in W\}$  as introduced above and a random function of residuals  $\{\xi_t, t \in W\}$ , i.e.,

$$P_t = \mu_t + \xi_t \quad \text{for all } t \in W, \quad (9)$$

and to assume that  $\{\xi_t, t \in W\}$  is the restriction of a centered, wide-sense stationary and isotropic random field  $\{\xi_t, t \in \mathbb{R}^2\}$  to  $W$ . This implies that the semivariograms of  $\{P_t, t \in W\}$  and  $\{\xi_t, t \in W\}$  are the same and therefore, an estimation of  $\gamma(\cdot)$  based on realizations of the residual function  $\{\xi_t, t \in W\}$  seems reasonable. However, this is a certain problem since only realizations of  $P_{s_1}, \dots, P_{s_n}$  are given, whereas neither  $\mu_{s_1}, \dots, \mu_{s_n}$  nor realizations of  $\xi_{s_1}, \dots, \xi_{s_n}$  are available.

To deal with this problem we consider an iterative approach for the estimation of the semivariogram  $\gamma(\cdot)$  proposed in Neuman and Jacobson (1984). In the following, a brief review of this approach is given. Suppose that the mean function  $\{\mu_t, t \in W\}$  can be represented as

$$\mu_t = \sum_{i=1}^k f_i(t)\beta_i \quad \text{for all } t \in W, \quad (10)$$

where  $k$  is an arbitrary integer,  $f_1(\cdot), \dots, f_k(\cdot)$  are chosen to be a sequence of monomials, and  $\beta_1, \dots, \beta_k$  are certain trend coefficients. In our example of application we put  $k = 10$  and  $(f_1(t), \dots, f_{10}(t)) = (1, t_{(1)}, t_{(2)}, t_{(1)}^2, t_{(1)}t_{(2)}, t_{(2)}^2, t_{(1)}^3, t_{(1)}^2t_{(2)}, t_{(1)}t_{(2)}^2, t_{(2)}^3)$  are monomials ranging up to 3rd order for any location  $t = (t_{(1)}, t_{(2)}) \in W$ . We introduce the following simplifying notation. Let  $P = (P_{s_1}, \dots, P_{s_n})$  be the random vector of point probabilities,  $X = (x_{i,j})$  with  $x_{i,j} = f_j(s_i)$  for  $i = 1, \dots, n$  and  $j = 1, \dots, k$  be a deterministic design matrix,  $\xi = (\xi_{s_1}, \dots, \xi_{s_n})$  be the random vector of residuals,  $\Sigma$  be the covariance matrix of  $\xi$ , and  $\beta = (\beta_1, \dots, \beta_k)$  is denoted as trend vector. Using (10) and the introduced notation, (9) yields that

$$P = X\beta + \xi, \quad (11)$$

which is a well-known linear regression model with correlated residuals. Since  $\Sigma$  is unknown, generalized least squares cannot be used. However, ordinary least squares could be applied, but this would result in biased estimators for  $\beta$  and  $\Sigma$ . The iterative approach presented in Neuman and Jacobson (1984) allows for an estimation of  $\beta$  and  $\Sigma$  with drastically reduced biases, although this bias cannot be removed completely. This is due to the circumstance that the estimation of  $\Sigma$  is always biased when considering linear regression with correlated residuals, even if the true covariance matrix  $\Sigma$  of  $\xi$  is known. In Cressie (1993), Chap. 3.4.3, and Beckers and Bogaert (1998), some simple examples are given to illustrate this bias problem.

Estimators for the trend vector  $\beta$ , the covariance matrix  $\Sigma$ , and thus also for the semivariogram  $\gamma(\cdot)$  can be computed according to the following algorithm.

1. Compute an estimator  $\hat{\beta}$  for  $\beta$  using ordinary least squares, i.e., by

$$\hat{\beta} = \underset{b \in \mathbb{R}^k}{\operatorname{argmin}} (P - Xb)^\top (P - Xb) = (X^\top X)^{-1} X^\top P. \quad (12)$$

2. Based on the estimator  $\hat{\beta}$  obtained in the previous step, determine a vector of empirical residuals  $\hat{\xi}$ . Then, compute an estimator  $\hat{\gamma}(\cdot)$  of the semivariogram  $\gamma(\cdot)$  based on  $\hat{\xi}$  according to the method-of-moment estimator suggested in Fernández-Avilés et al. (2015). Finally, compute an estimator  $\hat{\Sigma}$  for the covariance matrix  $\Sigma$  of  $\xi$  using the unique relationship between the semivariogram and the covariance function of a random field.
3. Recompute the estimator  $\hat{\beta}$  for  $\beta$  using generalized least squares with covariance matrix  $\hat{\Sigma}$ , i.e., by

$$\hat{\beta} = \underset{b \in \mathbb{R}^k}{\operatorname{argmin}} (P - Xb)^\top \hat{\Sigma}^{-1} (P - Xb) = (X^\top \hat{\Sigma}^{-1} X)^{-1} X^\top \hat{\Sigma}^{-1} P. \quad (13)$$

4. Repeat steps 2 and 3 until  $\hat{\beta}$ ,  $\hat{\xi}$ , and  $\hat{\gamma}(\cdot)$  converge to stable values.

A quick convergence of this algorithm can only be ensured if a continuous semivariogram estimator is obtained in step 2. For that purpose, an exponential semivariogram model is fitted before the covariance matrix  $\hat{\Sigma}$  is computed, see Cressie (1993) and Faulkner (2002). The exponential model is chosen since it seems to fit best the semivariogram computed using the methods-of-moment estimator.

### 3.3.3. Computation of the precipitation range

Finally, we describe an approach to the computation of the precipitation range  $R$  from residual semivariograms estimated according to Sect. 3.3.2. Again, we consider a fixed  $\omega_0 \in \Omega$  and the corresponding realizations  $p_{s_1}, \dots, p_{s_n}, a_1, \dots, a_n$  and  $r$  of  $P_{s_1}, \dots, P_{s_n}, A_1, \dots, A_n$  and  $R$ . On the one hand, an estimator  $\hat{\gamma}(\cdot)$  for  $\gamma(\cdot)$  is computed based on the residuals of  $p_{s_1}, \dots, p_{s_n}$ . On the other hand, we first consider a sequence  $t_1, \dots, t_m$  of locations in  $W$  that is comparable to (but different from)  $s_1, \dots, s_n$ . Then, for any  $r' > 0$  we introduce the deterministic field  $\{p_t^{(r')}, t \in W\}$  of point precipitation probabilities that correspond to the precipitation range  $r'$ , where

$$p_t^{(r')} = P(t \in M \mid \{A_i\} = \{a_i^{(r')}\}, R = r') = 1 - \exp\left(-\sum_{i=1}^n a_i^{(r')} v_2(b(t, r') \cap V(s_i))\right) \quad (14)$$

for all  $t \in W$  with  $a_1^{(r')}, \dots, a_n^{(r')}$  as computed in Sect. 3.3.1. We suppose that  $\{p_t^{(r')}, t \in W\}$  is a possible realization of a random field  $\{P_t^{(r')}, t \in W\}$ , such that the assumptions made in (9) and (10) remain valid if  $\{P_t, t \in W\}$  is replaced by  $\{P_t^{(r')}, t \in W\}$ . Based on this, consider the semivariogram  $\gamma^{(r')}(\cdot)$  of the field of residuals  $\{\xi_t^{(r')}, t \in \mathbb{R}^2\}$  that corresponds to  $\{P_t^{(r')}, t \in W\}$  and compute an estimator  $\hat{\gamma}^{(r')}(\cdot)$  for  $\gamma^{(r')}(\cdot)$  based on the residuals of  $p_{t_1}^{(r')}, \dots, p_{t_m}^{(r')}$ . Finally, determine the precipitation range  $r$  in such a way that the squared  $\mathcal{L}_2$ -distance between  $\hat{\gamma}(\cdot)$  and  $\hat{\gamma}^{(r')}(\cdot)$  is minimized, i.e.,

$$r = \underset{r' > 0}{\operatorname{argmin}} \left\{ \int_{c_1}^{c_2} \left( \hat{\gamma}(h) - \hat{\gamma}^{(r')}(h) \right)^2 dh \right\}, \quad (15)$$

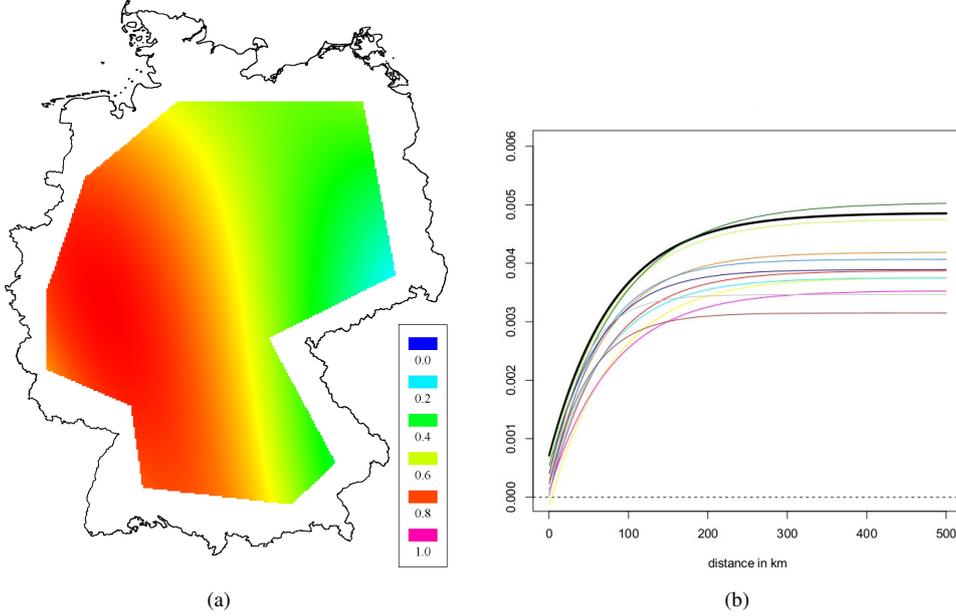


Figure 2: Results for forecast step ‘December 22, 2012, 21 UTC’. Left: estimated trend function. Right: comparison of residual semivariograms estimated from the available data (black) and from the introduced model for precipitation ranges 10, 12.5, 15, 17.5, 20, 22.5, 25, 27.5, 30 and 32.5 km (in colors).

with  $c_1, c_2 > 0$  being some suitable integration limits. Note that the precipitation range  $r$  also controls the appearance of the field  $\{P_t(\omega_0), t \in W\}$ . For a too small precipitation range  $\{P_t(\omega_0), t \in W\}$  will be close to a piecewise constant field, whereas only for large enough  $r$  the probability  $P_t(\omega_0)$  depends on more than one intensity value (see (4)) causing the field to be sufficiently smooth. Therefore, only ranges exceeding a certain minimal range (which depends on the sizes of the Voronoi cells) should be considered in (15).

To conclude this section, we illustrate the results which are obtained with the algorithm described above in our example of application. In Fig. 2 (a) the trend function  $\{\mu_t, t \in W\}$  estimated from  $p_{s_1}, \dots, p_{s_{503}}$  is depicted for one forecast step, where  $\mu_t$  is only considered for locations  $t$  within Germany to avoid boundary effects. A comparison with Fig. 1 (b) shows a good accordance. In Fig. 2 (b) the estimated residual semivariogram  $\hat{\gamma}(\cdot)$  (black) is compared to the estimated residual semivariograms  $\hat{\gamma}^{(r')}(\cdot)$  for a sequence of 10 possible precipitation ranges (in colors). It is clearly visible that the precipitation range has indeed an impact on the dependence structure of the residuals and that the semivariogram estimated from the data goes well with those estimated for different model characteristics. For the majority of forecast steps we obtained similar results.

### 3.4. Computation of area precipitation probabilities

For each sequence of point probabilities  $p_{s_1}, \dots, p_{s_n}$  the corresponding model characteristics  $a_1, \dots, a_n$  and  $r$  are computed according to (8) and (15). Then, for each Borel set  $B \subset W$ , the

corresponding area precipitation probability  $\pi(B)$  is given by

$$\pi(B) = P(B \cap M \neq \emptyset | \{A_i\} = \{a_i\}, R = r) = 1 - \exp\left(-\sum_{i=1}^n a_i v_2((B \oplus b(o, r)) \cap V(s_i))\right), \quad (16)$$

see also (5). Note that a direct computation of  $\pi(B)$  from this formula has certain disadvantages. The intersection areas of  $B \oplus b(o, r)$  and the Voronoi cells  $V(s_i)$  need to be determined numerically, which is both computationally expensive and imprecise. As an alternative approach we suggest to estimate  $\pi(B)$  as a relative frequency by repeated simulation of the germ-grain model  $M$ , which is a Boolean model for  $\omega_0 \in \Omega$ . An estimation based on 1,000 realizations of  $M$  provides a high degree of precision and is computationally more efficient than the direct computation of  $\pi(B)$  based on (16). In principle, however, both the direct and the simulation-based computation are possible.

#### 4. Model validation

The model for the computation of area precipitation probabilities proposed in this paper has been implemented in Java using classes and methods from the GeoStoch library, a Java-based software developed at Ulm University (Mayer et al., 2004). For each of the 854 available forecast steps described in Sect. 2.2 the model characteristics are determined from the 503 given point probabilities in order to allow for the computation of area precipitation probabilities. In Fig. 3 (a), a realization of the germ-grain model  $M$  is illustrated for one forecast step. A comparison with Fig. 1 (b) shows that centers of precipitation cells are mainly generated in those areas, where given point precipitation probabilities are high.

##### 4.1. Description of radar data

Computed area precipitation probabilities are compared with radar-derived precipitation analyses from the German operational radar network of DWD (Winterrath et al., 2012). The DWD radar network consists of 16 sites that cover Germany and provide precipitation scans every 5 minutes. Radar reflectivities are transformed into precipitation rates using empirical reflectivity-precipitation rate relationships. Precipitation rates are accumulated for every hour and an adjustment is performed using about 1300 rain gauges at conventional meteorological measurement sites in order to provide adjusted quantitative precipitation analyses. An additional clutter filter for hydrological applications described in Winterrath and Rosenow (2007) removes spurious pixel-scale precipitation events. The minimal observed precipitation amount per hour is 0.1 mm<sup>1</sup>, which is why the threshold value for the MOS-derived point probabilities has been set accordingly (see Sect. 2.2). An illustration of adjusted precipitation amounts derived from radar data for one example forecast step is given in Fig. 3 (b). It is clearly visible that precipitation mainly occurred in those regions of Germany, where point precipitation probabilities are high (compare with Fig. 1 (b)) and where many precipitation cells are simulated by our stochastic model (compare Fig. 3 (a)). However, we emphasize once more that the germ-grain model  $M$

---

<sup>1</sup>The consideration of a higher threshold would also be possible, where we expect that the precipitation range gets smaller and cells occur less frequently the more the threshold grows. A problem arises if the precipitation range gets too small in comparison to the sizes of the Voronoi cells  $V(s_1), \dots, V(s_n)$ , see the discussion in Sect. 3.3.3.

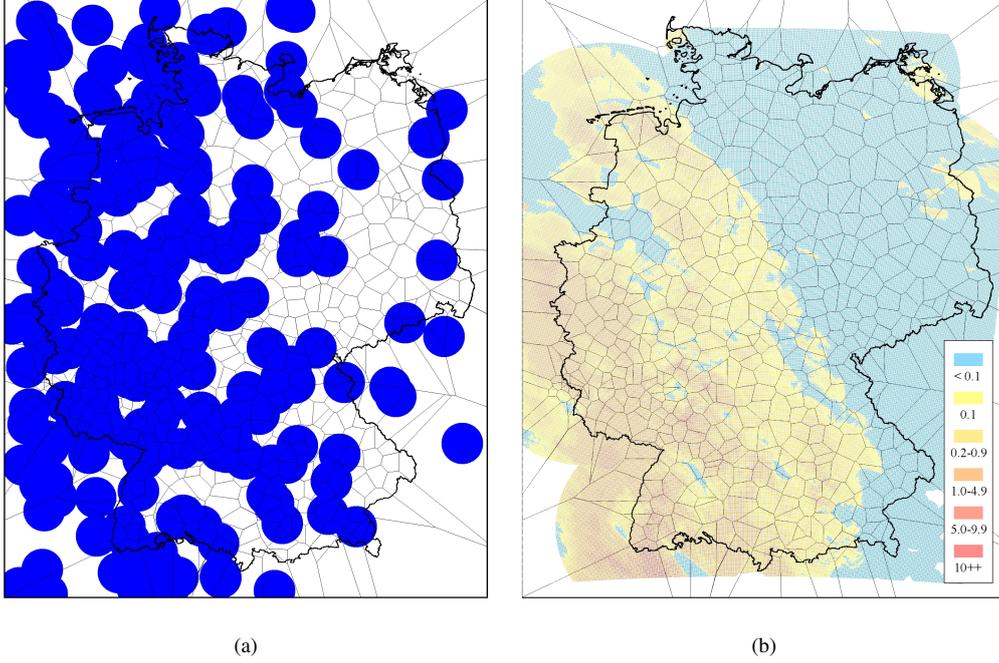


Figure 3: Results for forecast step ‘December 22, 2012, 21 UTC’. Left: realization of the germ-grain model  $M$  with characteristics computed from available point probabilities. Right: adjusted precipitation amounts in mm derived from radar data.

described in Sect. 3.1 should not be regarded as a model for real precipitation cells. In particular, single realizations of  $M$  may look atypically compared to real precipitation fields, see Fig. 3. Furthermore, a single realization of  $M$  should not be compared with radar observations since the former is based on (point) precipitation probabilities, whereas the latter shows precipitation events that can be considered to be realized by the mentioned precipitation probabilities. Therefore, a comparison seems only reasonable if a longer time period is considered, which is done in the next section.

#### 4.2. Validation of probabilistic forecast approach

A validation of our model is performed by comparing computed area precipitation probabilities with precipitation indicators derived from radar observations. For any test area  $B$  a sequence of area probabilities  $\pi_1(B), \dots, \pi_{854}(B)$  is computed according to the model introduced in Sect. 3 for all available forecast steps. Furthermore, a sequence of precipitation indicators  $I_1(B), \dots, I_{854}(B)$  is considered, where  $I_j(B)$  is equal to 1 if there is precipitation somewhere within  $B$  in forecast step  $j$  with respect to radar observations and 0 otherwise.

In order to assess the quality of computed area probabilities we consider the following scores. At first, the bias  $b$  is introduced, which is defined as the difference between the mean computed precipitation probability and the relative frequency of precipitation, i.e.,

$$b = \frac{1}{854} \sum_{j=1}^{854} \pi_j(B) - \frac{1}{854} \sum_{j=1}^{854} I_j(B). \quad (17)$$

In a good weather prediction the bias  $b$  given in (17) should be close to zero. A clearly negative or positive bias would indicate that computed probabilities are systematically too low or too high, respectively. Next, the Brier score  $bs$  (Brier, 1950) is considered, which is the mean squared difference between computed probabilities and observed precipitation indicators given by

$$bs = \frac{1}{854} \sum_{j=1}^{854} (\pi_j(B) - I_j(B))^2. \quad (18)$$

Of course, the Brier score should be as small as possible. However, there is no intuitive interpretation of this score, since the Brier score also depends on a term denoted as uncertainty (which is expressed only by the variability of  $I_j$  for  $j = 1, \dots, 854$ , see e. g. Wilks (1995)). Therefore, a closely related quantity is considered. Assume that another sequence  $\tilde{\pi}_1(B), \dots, \tilde{\pi}_{854}(B)$  of area precipitation probabilities computed from a reference method is available and consider the corresponding Brier score  $\tilde{bs}$ . Then, the Brier skill score  $bss$  defined as

$$bss = 1 - \frac{bs}{\tilde{bs}} \quad (19)$$

is considered in order to investigate whether the proposed method provides better results than another (previous) reference method. If no reference method is available, which is the case in our validation, the so-called climate mean is considered, where each probability  $\tilde{\pi}_i(B)$  is put to the relative frequency of precipitation  $\frac{1}{854} \sum_{j=1}^{854} I_j(B)$ . If the newly developed method leads to an improvement of weather prediction, the Brier skill score should be clearly positive. Finally, we suggest to consider the joint empirical correlation coefficient  $\rho$  of the computed probabilities  $\pi_1(B), \dots, \pi_{854}(B)$  and the observed precipitation indicators  $I_1(B), \dots, I_{854}(B)$ . Clearly,  $\rho$  should be as close to one as possible, where  $\rho = 1$  would indicate a perfect weather prediction and  $\rho = 0$  would suggest no relationship between computed precipitation probabilities and observed precipitation events.

We consider test areas  $B_1, \dots, B_{503}$  which are chosen to be the Voronoi cells  $V(s_1), \dots, V(s_{503})$  around the locations  $s_1, \dots, s_{503}$  of the weather stations. The Voronoi cells  $V(s_1), \dots, V(s_{503})$  are suitable for model validation since they include areas with different size, shape, and orientation. Area probabilities computed according to the model introduced in this paper are a function of model characteristics, which in turn are computed from the available point probabilities. Therefore, a validation is not only performed for area probabilities that correspond to the 503 Voronoi cells  $V(s_1), \dots, V(s_{503})$  but also for point precipitation probabilities that correspond to the locations  $s_1, \dots, s_{503}$ .

In Fig. 4 the bias of given point probabilities and computed area probabilities is illustrated. On the left-hand side the Voronoi cells  $V(s_1), \dots, V(s_{503})$  are colored according to the bias of point probabilities at the locations of corresponding weather stations, whereas on the right-hand side the bias of the area probabilities that correspond to the Voronoi cells is depicted. Again, only cells within Germany are considered to avoid boundary effects. The biases for point probabilities range from -0.04 to 0.06 for most weather stations, which are acceptably small showing a good consistency of the model output statistics with the radar data. There are three outliers with biases up to 0.14, which might be caused by problems with the radar measurements. The biases for area probabilities range from -0.06 to 0.09, which is almost in the same range as for the point probabilities. When looking at the north-east and the north-west of Germany, it seems that biases of the point probabilities are amplified when deriving area probabilities.

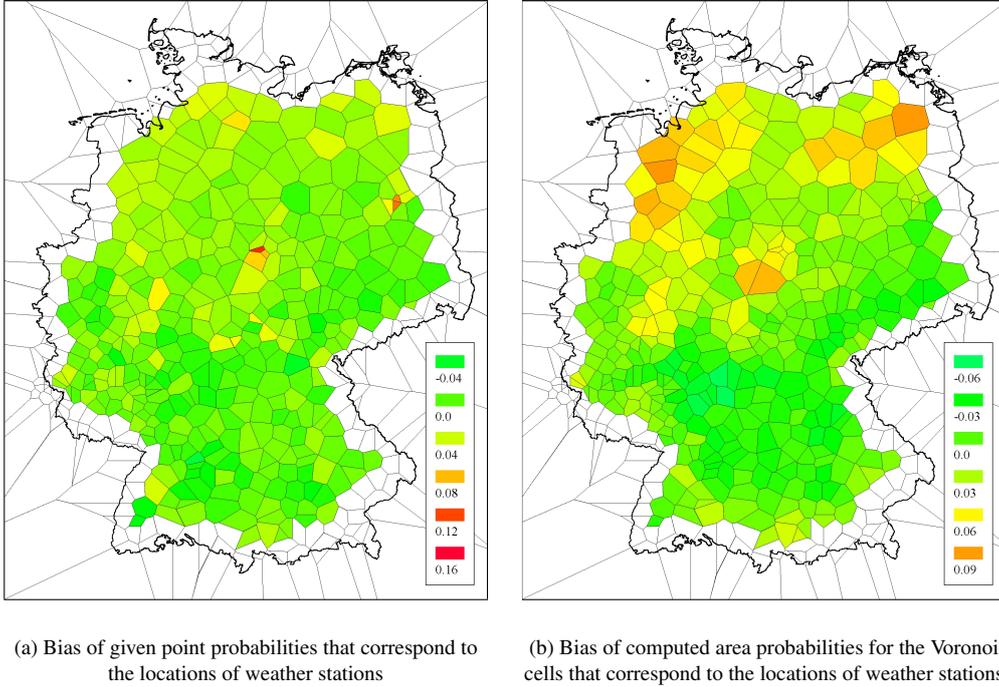


Figure 4: Bias of point and area precipitation probabilities.

Fig. 5 provides histograms of the Brier skill scores computed for point (left) and area probabilities (right). The Brier skill score is clearly positive for all considered Voronoi cells, implying that the presented method indeed provides an improvement in computing area precipitation probabilities. There are again outliers for the point probabilities, the mean value of 0.28 shows a positive signal though. It is a pleasant result that the area precipitation probabilities are not affected by these outliers and that the mean Brier skill score of 0.4 is even larger than for the point probabilities.

Similar results are obtained when considering histograms of the joint empirical correlation coefficients of precipitation probabilities and precipitation indicators, see Fig. 6. For each Voronoi cell a clearly positive correlation between computed probabilities and observed precipitation events is detected, where the mean correlation is even higher than that observed for given point probabilities. This shows impressively that area precipitation probabilities computed by the statistical method proposed in this paper correspond well with radar observations.

To conclude our validation, reliability diagrams for area precipitation probabilities are considered. Basically, a reliability diagram illustrates whether precipitation indeed occurs rarely or often for forecasts with small or high precipitation probabilities, respectively. A mathematical description is given as follows. At first, the unit interval  $[0, 1]$  is decomposed into a sequence  $U_1, \dots, U_{20}$  of 20 equal subintervals. Then, for any test area  $B$  and each fixed subinterval  $U_k$ , the reliability  $\rho(U_k)$  is defined as the relative frequency of precipitation among those forecast steps

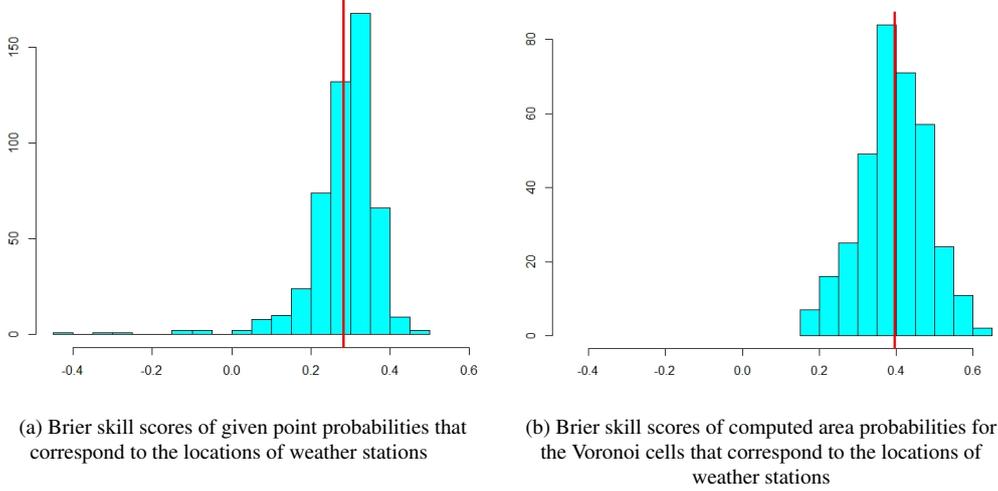


Figure 5: Histograms of Brier skill scores computed for point and area precipitation probabilities. The red line indicates the mean score.

for that the computed area precipitation probability takes a value in  $U_k$ . More precisely,

$$\rho(U_k) = \frac{\#\{1 \leq j \leq 854 : \pi_j(B) \in U_k, I_j(B) = 1\}}{\#\{1 \leq j \leq 854 : \pi_j(B) \in U_k\}} \quad \text{for } k = 1, \dots, 20. \quad (20)$$

Additionally, the midpoint  $m_k$  of each subinterval  $U_k$  is computed and the sequence of points  $(m_1, \rho(U_1)), \dots, (m_{20}, \rho(U_{20}))$  is called a reliability diagram. Ideally, the computed probabilities should be equal to the observed relative frequencies and all points of the diagram should be close to the curve  $x = y$ . In Fig. 7 reliability diagrams for two Voronoi cells are illustrated, where it is shown that precipitation occurs frequently if computed precipitation probabilities are high and vice versa. Similar results were obtained for other Voronoi cells, too.

## 5. Conclusion

In this paper a method for the computation of area precipitation probabilities has been introduced. We proposed an intuitive approach to model point and area precipitation probabilities as coverage probabilities of random precipitation fields, which are approximated by a germ-grain model. The model characteristics are computed from available point probabilities using several statistical methods. Our model has been applied to some disjoint areas within Germany and a validation has been performed. A comparison of area probabilities computed from our model with precipitation events derived from radar shows excellent results. For most of the analyzed areas no systematic bias of computed probabilities was found, there is a high correlation between precipitation probabilities and the occurrence of precipitation, and reliability diagrams show a particularly nice performance. Moreover, to run the method no further input of the forecaster is necessary and the method works in a quite reasonable computation time allowing for an automated generation of area probabilities. This means that the presented method is suitable for applications in modern weather prediction. The high impact of the presented method is also

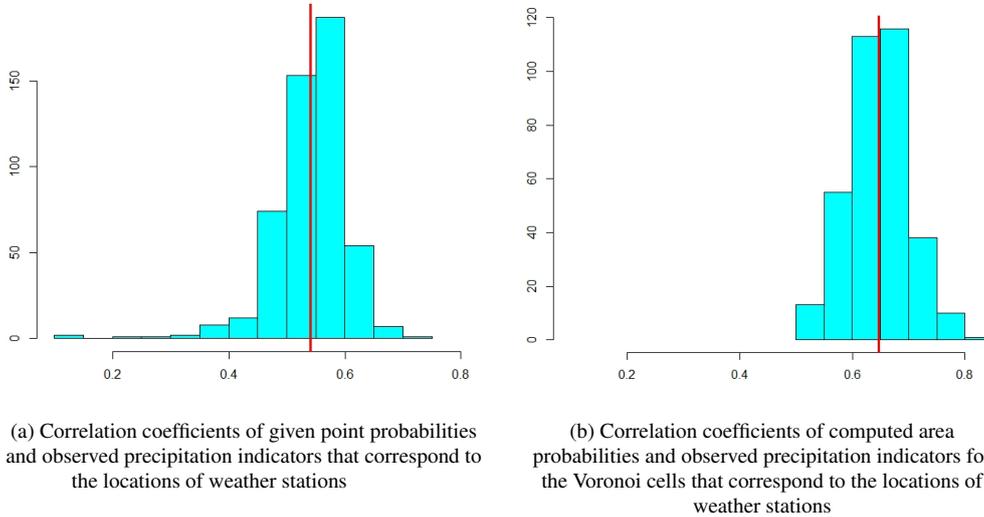


Figure 6: Histograms of joint empirical correlation coefficients of precipitation probabilities and corresponding precipitation indicators. The red line indicates the mean correlation coefficient.

shown by the fact that so far no applicable approach to the computation of area probabilities is available, whereas the number of potential problems and applications is high.

However, in this paper the performance of the model has only been validated for areas that are located within Germany and that have certain range of sizes. Although the validation has been performed for areas with different size, shape, and orientation, it is not obvious how the method works for areas that are significantly smaller or larger or have a completely different shape than those investigated here. The same holds if a region with different climatological conditions is considered. Another requirement of this method is that the typical size of precipitation cells should not be too small compared to the density of weather stations. Thus, the presented approach is not suitable for applications where a small scale weather event is considered (e.g. precipitation exceeding a high threshold) or where distances between locations with given point probabilities are too large.

In this sense, the presented approach can be considered as a first step towards the prediction of area weather events and further refinements are planned as subjects of our future research. On the one hand, a more realistic modeling of precipitation cells will be considered, including several ideas proposed in previous publications, e.g., elliptical precipitation cells (see e.g. Wheeler et al., 2000), different cell types (Cowpertwait, 1995) or clustered precipitation cells (Rodriguez-Iturbe et al., 1986). On the other hand, more sophisticated problems in probabilistic weather prediction will be considered. In particular, the computation of area probabilities for the occurrence of precipitation that exceeds a certain (high) threshold is under preparation. As mentioned before, an application of the existing model to such a small scale weather event might cause problems. Therefore, an integration of precipitation amounts into the model is currently under preparation. Furthermore, an application of our method to other weather events, e.g. the occurrence of (strong) wind gusts, is of great interest. Even a combination of several weather events, e.g. the impact of strength and direction of winds to the precipitation range in the current

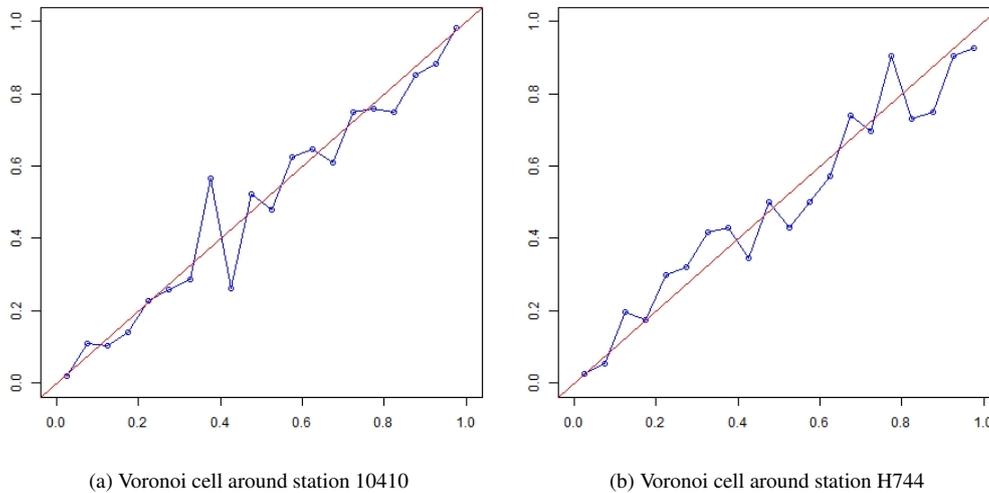


Figure 7: Reliability diagrams for area precipitation probabilities.

model, seems possible and will be considered in future research.

## Acknowledgments

We are grateful to Tanja Winterrath and Elmar Weigl for providing information on clutter correction and on gauge adjusted radar data in general. We also would like to thank Tamas Hirsch for preprocessing the radar data and Tim Brereton for a critical reading of the manuscript.

## References

- Beckers, F., Bogaert, P., 1998. Nonstationarity of the mean and unbiased variogram estimation: extension of weighted least-squares method. *Math. Geol.* 30, 499–521.
- Brier, G.W., 1950. Verification of forecasts expressed in terms of probability. *Mon. Weather Rev.* 78, 1–3.
- Chiu, S.N., Stoyan, D., Kendall, W.S., Mecke, J., 2013. *Stochastic Geometry and its Applications*. 3rd ed., J. Wiley & Sons, Chichester.
- Cowpertwait, P.S.P., 1995. A generalized spatial-temporal model of rainfall based on a clustered point process. *Proc. Royal Soc. Lond. A* 450, 163–175.
- Cowpertwait, P.S.P., Xie, G., Isham, V.S., Onof, C.J., Walsh, D.C.I., 2011. A fine-scale point process model of rainfall with dependent pulse depths within cells. *Hydrol. Sci. J.* 56, 1110–1117.
- Cox, D.R., Isham, V.S., 1988. A simple spatial-temporal model of rainfall. *Proc. Royal Soc. Lond. A* 415, 317–328.
- Cressie, N.A.C., 1993. *Statistics for Spatial Data*. 2nd ed., J. Wiley & Sons, New York.
- Elsner, J.B., Murnane, R.J., Jagger, T.H., Widen, H.M., 2013. A spatial point process model for violent tornado occurrence in the US Great Plains. *Math. Geosci.* 45, 667–679.
- Epstein, S.E., 1966. Point and area precipitation probabilities. *Mon. Weather Rev.* 94-10, 595–598.
- Faulkner, B.R., 2002. Java classes for nonprocedural variogram modeling. *Comput. Geosci.* 28, 387–397.
- Fernández-Avilés, G., Montero, J.-M., Mateu, J., 2015. *Spatial and Spatio-Temporal Geostatistical Modeling and Kriging*. J. Wiley & Sons, Chichester.
- Illian, J., Penttinen, A., Stoyan, H., Stoyan, D., 2008. *Statistical Analysis and Modelling of Spatial Point Patterns*. J. Wiley & Sons, Chichester.
- Kaczmarek, J., Isham, V.S., Onof, C.J., 2014. Point process models for fine-resolution rainfall. *Hydrol. Sci. J.* 59, 1972–1991.
- Karpman, D., Ferreira, M.A.R., Wikle, C.K., 2013. A point process model for tornado report climatology. *Stat* 2, 1–8.

- Knüpfner, K., 1996. Methodical and predictability aspects of MOS systems, in: 13th Conf. on Probability and Statistics in Atmosph. Sciences, Amer. Meteorol. Soc., San Francisco, CA. pp. 190–197.
- Kriesche, B., Weindl, H., Smolka, A., Schmidt, V., 2014. Stochastic simulation model for tropical cyclone tracks, with special emphasis on landfall behavior. *Nat. Hazards* 73, 335–353.
- Krzysztofowicz, R., 1998. Point-to-area rescaling of probabilistic quantitative precipitation forecasts. *J. Appl. Meteorol.* 38, 786–796.
- Lawson, C.L., Hanson, R.J., 1974. *Solving Least Squares Problems*. Prentice-Hall, Englewood Cliffs, NJ.
- Majewski, D., 1998. The new global icosahedral-hexagonal grid point model GME of the Deutscher Wetterdienst, in: ECMWF Seminar Proceedings: Recent Developments in Numerical Methods for Atmospheric Modelling, pp. 173–201.
- Majewski, D., Liermann, D., Prohl, P., Ritter, B., Buchhold, M., Hanisch, T., Paul, G., Wergen, W., Baumgardner, J., 2002. The global icosahedral-hexagonal grid point model GME: description and high-resolution tests. *Mon. Weather Rev.* 130-2, 319–338.
- Mayer, J., Schmidt, V., Schweiggert, F., 2004. A unified simulation framework for spatial stochastic models. *Simul. Model. Pract. Theory* 12, 307–326.
- Neuman, S.P., Jacobson, E.A., 1984. Analysis of nonintrinsic spatial variability by residual kriging with application to regional groundwater levels. *Math. Geol.* 16, 223–240.
- Northrop, P.J., 1998. A clustered spatial-temporal model of rainfall. *Proc. Royal Soc. Lond. A* 454, 1875–1888.
- Onof, C.J., Chandler, R.E., Kakou, A., Northrop, P.J., Wheeler, H.S., Isham, V.S., 2000. Rainfall modelling using poisson-cluster processes: a review of developments. *Stoch. Environ. Res. Risk Assess.* 14, 384–411.
- Ramesh, N.I., Thayakaran, R., Onof, C.J., 2012. Multi-site doubly stochastic poisson process models for fine-scale rainfall. *Stoch. Environ. Res. Risk Assess.* 27, 1383–1396.
- Rodriguez-Iturbe, I., Cox, D.R., Eagleson, P.S., 1986. Spatial modelling of total storm rainfall. *Proc. Royal Soc. Lond. A* 403, 27–50.
- Segond, M.-L., Onof, C.J., 2009. Modelling of space-time rainfall for three UK regions. *Proc. ICE - Water Manag.* 162, 147–158.
- Wheater, H.S., Chandler, R.E., Onof, C.J., Isham, V.S., Bellone, E., Yang, C., Lekkas, D., Lourmas, G., Segond, M.-L., 2005. Spatial-temporal rainfall modelling for flood risk estimation. *Stoch. Environ. Res. Risk Assess.* 19, 403–416.
- Wheater, H.S., Isham, V.S., Cox, D.R., Chandler, R.E., Kakou, A., Northrop, P.J., Oh, L., Onof, C.J., Rodriguez-Iturbe, I., 2000. Spatial-temporal rainfall fields: modelling and statistical aspects. *Hydrol. Earth Syst. Sci.* 4, 581–601.
- Wilks, D.S., 1995. *Statistical Methods in the Atmospheric Sciences*. Academic Press, San Diego.
- Winterrath, T., Rosenow, W., 2007. A new module for the tracking of radar-derived precipitation with model-derived winds. *Adv. Geosci.* 10, 77–83.
- Winterrath, T., Rosenow, W., Weigl, E., 2012. On the DWD quantitative precipitation analysis and nowcasting system for real-time application in German flood risk management, in: Moore, R.J., Cole, S.J., Illingworth, A.J. (Eds.), *Weather Radar and Hydrology* (Proceedings of a symposium held in Exeter, UK, April 2011), pp. 323–329.