

Calibrated prediction regions for Gaussian random fields

Corrado Lagazio

Department of Economics and Business Studies, University of Genova

via Vivaldi 5, 16126 Genova, Italy.

E-mail: corrado.lagazio@unige.it

Paolo Vidoni*

Department of Economics and Statistics, University of Udine

via Tomadini 30/a, 33100 Udine, Italy.

E-mail: paolo.vidoni@uniud.it

Abstract

This paper proposes a method to construct well-calibrated frequentist prediction regions, with particular regard to highest prediction density regions, which may be useful for multivariate spatial prediction. We consider, in particular, Gaussian random fields and, using a calibrating procedure introduced by Beran (1990) and developed by Fonseca et al. (2014), we effectively improve the estimative prediction regions, since the coverage probability turns out to be closer to the target nominal value. Whenever a close form expression for the well-calibrated prediction region is not available, we may specify a simple bootstrap-based estimator. Particular attention is dedicated to the associated improved predictive distribution function, which can be usefully considered for identifying spatial locations with extreme or unusual observations. A simulation study is proposed in order to compare empirically the calibrated predictive regions with the estimative ones. The proposed method is then applied to the global model assessment of a deterministic model for the prediction of PM_{10} levels using data from a network of air quality monitoring stations.

Keywords: Air quality; Bootstrap calibration; Coverage probability; Estimative prediction region; Highest prediction density region; Spatial prediction.

*Corresponding author.

1 Introduction

In this paper we consider the problem of constructing well-calibrated frequentist prediction regions, with particular regard to highest prediction density regions, which may be useful in many geostatistical applications, whenever we aim at predicting unknown spatial observations based on the available data at a finite set of sampling locations. In this context, at least when we assume a Gaussian model, we usually define prediction intervals or prediction regions by considering the kriging predictors, with the associated kriging variance, and by replacing the unknown model parameters with the corresponding estimates. However, these so-called estimative or plug-in prediction intervals and, in general, prediction regions present a coverage probability which may differ substantially from the target nominal value, since the additional uncertainty introduced by the plug-in procedure is not adequately taken into account. This may lead to unreliable predictive conclusions, in particular when the dimension of the random vector to be predicted is high with respect to the dimension of the observed sample.

With concern to univariate prediction intervals, assuming a Gaussian random field, Sjöstedt-de Luna and Young (2003) proposes a parametric bootstrap calibration method which improves the estimative solution so that the coverage probability turns out to be closer to the nominal value. Extensions of the method for improved univariate prediction within non-Gaussian spatial models are presented in De Oliveira and Rui (2009) and Schelin and Sjöstedt-de Luna (2010). The application of these calibration methods to multivariate spatial prediction regions does not seem to have been considered so far, even if the problem of jointly predicting new observations at multiple spatial locations is of considerable interest in applications.

We acknowledge that improved multivariate procedures have been already defined in a more

general setting (see, for example, Beran, 1990; Fonseca et al., 2012) and therefore the purpose of this paper is to use one of these calibrating procedures in order to improve the coverage probability of the estimative prediction regions, and in particular of the estimative highest prediction density regions, defined within a Gaussian random field. More precisely, we apply to this spatial framework the improved predictive procedure suggested by Beran (1990) and developed also by Fonseca et al. (2014). This approach is similar that one considered in Fonseca et al. (2012), but it mainly relies on bootstrap methods instead of asymptotic arguments. Particular attention is dedicated to the predictive distribution giving improved prediction intervals and improved prediction regions, since it can be usefully considered for the specification of the thresholds identifying spatial locations with extreme or unusual observations. A simulation study, concerning some Gaussian random fields, is proposed in order to compare the coverage probabilities of the calibrated and the estimative prediction regions. Finally, the proposed method is then applied to the global model assessment of a deterministic model for the prediction of PM₁₀ levels using data from a network of air quality monitoring stations.

2 Estimative prediction regions

Let us consider a stochastic process $\{Y(s), s \in \Delta\}$ with $\Delta \subset \mathbf{R}^p$, where s indicates a spatial location and p is usually not greater than 3. In particular, we assume that the process defines a Gaussian random field with mean function $\mu(s; \theta) = E\{Y(s)\}$ and covariance function $\sigma(s, t; \theta) = Cov\{Y(s), Y(t)\}$, which both exist for all $s, t \in \Delta$, with $\theta \in \Theta \subseteq \mathbf{R}^k$, $k \geq 1$, an unknown k -dimensional parameter. We suppose that the spatial process is observed at locations $s_1, \dots, s_n \in \Delta$ and we aim at predicting the observations at the fur-

ther points $s_{n+1}, \dots, s_{n+m} \in \Delta$, with $n, m \geq 1$. Thus we observe the random vector $Y = (Y_1, \dots, Y_n)^T = (Y(s_1), \dots, Y(s_n))^T$ and our objective is to predict the not yet observed random vector $Z = (Z_1, \dots, Z_m)^T = (Y(s_{n+1}), \dots, Y(s_{n+m}))^T$.

Although prediction problems may be studied from different perspectives, we focus on the specification of a suitable α -prediction region $R_\alpha(Y) \subset \mathbf{R}^m$ for Z , such that

$$P_{Y,Z}\{Z \in R_\alpha(Y); \theta\} = \alpha,$$

for every $\theta \in \Theta$ and for any fixed $\alpha \in (0, 1)$. The above probability is called coverage probability and it is calculated by considering that (Y, Z) follows a multivariate Gaussian distribution with mean vector $\mu = \mu(\theta) = (\mu_1, \dots, \mu_{n+m})^T$ and covariance matrix $\Sigma = \Sigma(\theta) = (\sigma_{i,j})_{i,j=1,\dots,n+m}$, abbreviated as $(Y, Z) \sim N_{n+m}(\mu, \Sigma)$. Here, $\mu_i = \mu(s_i; \theta)$ and $\sigma_{i,j} = \sigma(s_i, s_j; \theta)$, $i, j = 1, \dots, n+m$. In the univariate case, namely when $m = 1$, we usually consider α -prediction intervals for $Z = Z_1$ or, in particular, α -prediction limits $l_\alpha(Y)$ such that

$$P_{Y,Z}\{Z \leq l_\alpha(Y); \theta\} = \alpha, \tag{1}$$

for every $\theta \in \Theta$ and for any fixed $\alpha \in (0, 1)$.

The simplest way for making prediction on Z is to consider the joint Gaussian distribution of (Y, Z) in order to specify a reliable prediction region, where the unknown parameter θ is replaced by an asymptotically efficient estimator $\hat{\theta}$ based on Y . However, it is well-known that prediction regions based on this estimative (plug-in) procedure are not entirely adequate predictive solutions. Their effective coverage probability differs from the required value α by a term usually of order $O(n^{-1})$ and prediction statements may be rather inaccurate for small n and/or large m or k . Hereafter, we consider the maximum likelihood estimator or any asymptotically

equivalent alternative estimator and, for the validity of the large sample results presented in what follows, we implicitly assume that the conditions such that $\hat{\theta} - \theta = O_p(n^{-1/2})$ are verified. However, we recall that, in the spatial context, the asymptotic properties of parameter estimators strongly depend on the asymptotic regime which is considered. Namely, the increasing domain framework where the spatial domain of observation is unbounded and it grows in size with the sample dimension n and the fixed (or infill) domain where the spatial domain is fixed and bounded and the density of the sampling locations increases with n . Whereas under the former regime the maximum likelihood estimators are consistent and asymptotically normal, subject to some regularity conditions (see, for example, Mardia and Marshall, 1984), under the latter analogous results do not hold and model parameters could not be consistently estimated (see, for the case of Matérn covariance functions, Zhang, 2004).

Within Gaussian random field models, it is quite usual to construct estimative prediction intervals from the kriging predictor and the associated minimum mean square error of prediction, both evaluated at $\theta = \hat{\theta}$ (see, for example, Sjöstedt-de Luna and Young, 2003). Suitable generalizations of this approach are considered also for the case of multivariate prediction where the solution is usually obtained by combining univariate kriging prediction intervals. In this case, the associated prediction region is of rectangular form and, in addition to an unsatisfactory coverage accuracy, it has the further drawback that the dependence among the observations to be predicted is not adequately taken into account.

In this paper, we shall specify estimative prediction regions directly from the conditional distribution of Z given $Y = y$, with density $f(z|y; \theta)$, which is an m -dimensional Gaussian distribution $N_m(\mu_{Z|Y}, \Sigma_{Z|Y})$, with $\mu_{Z|Y} = \mu_Z + \Sigma_{ZY}\Sigma_{YY}^{-1}(y - \mu_Y)$ and $\Sigma_{Z|Y} = \Sigma_{ZZ} - \Sigma_{ZY}\Sigma_{YY}^{-1}\Sigma_{YZ}$.

Here, the mean vector and the covariance matrix of (Y, Z) are partitioned according to the components Y and Z so that

$$\mu = \begin{pmatrix} \mu_Y \\ \mu_Z \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{YY} & \Sigma_{YZ} \\ \Sigma_{ZY} & \Sigma_{ZZ} \end{pmatrix},$$

and A^T and A^{-1} indicate, respectively, the transpose and the inverse of a matrix A .

With regard to the univariate case, where $Z = Z_1$, it is immediate to see that $Z|Y = y$ follows a univariate Gaussian distribution with conditional density and distribution functions given, respectively, by

$$f(z|y; \theta) = \frac{1}{\sigma_{Z_1|Y}} \phi\left(\frac{z - \mu_{Z_1|Y}}{\sigma_{Z_1|Y}}\right), \quad F(z|y; \theta) = \Phi\left(\frac{z - \mu_{Z_1|Y}}{\sigma_{Z_1|Y}}\right),$$

where $\mu_{Z_1|Y} = \mu_{Z_1|Y}(\theta, y)$ and $\sigma_{Z_1|Y}^2 = \sigma_{Z_1|Y}^2(\theta)$ are given, respectively, by $\mu_{Z|Y}$ and $\Sigma_{Z|Y}$ with Z_1 instead of Z . Indeed, $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal density and distribution functions and q_α indicates the corresponding α -quantile, so that $\Phi(q_\alpha) = \alpha$. Then the estimative α -prediction limit for Z corresponds to $\hat{z}_\alpha = z_\alpha(\hat{\theta}; y)$, which is the α -quantile of Z given $Y = y$ specified by $z_\alpha = z_\alpha(\theta; y) = \mu_{Z_1|Y} + q_\alpha \sigma_{Z_1|Y}$ and evaluated at $\theta = \hat{\theta}$.

In the multivariate case, as suggested by Beran (1990), a simple procedure for generating a multivariate prediction region R_α is to define a suitable real function $D(z; y, \theta)$, called the root of the prediction region, such that

$$R_\alpha = R_\alpha(\theta; y) = \{z \in \mathbf{R}^m : D(z; y, \theta) \leq d_\alpha\},$$

where $d_\alpha = d_\alpha(\theta; y)$, that is the limit of the region, satisfies the following relation

$$F(d_\alpha|y; \theta) = P_{Z|Y}\{D(Z; Y, \theta) \leq d_\alpha|Y = y; \theta\} = \alpha.$$

Here, with a slight abuse of notation, $F(d|y; \theta)$ is the conditional distribution function of $D = D(Z; Y, \theta)$ given $Y = y$ and $P_{Z|Y}(\cdot|Y = y; \theta)$ corresponds to the probability distribution of Z given $Y = y$ under the true parameter value θ . Since R_α depends on the unknown parameter θ , we consider the associated plug-in estimator $\widehat{R}_\alpha = R_\alpha(\widehat{\theta}; y)$, called estimative (plug-in) prediction region. Note that this multidimensional predictive procedure is quite general and the only practical requirement for the root $D(z; y, \theta)$ is that the resulting region turns out to be bounded.

In particular, if we are interested in highest prediction density regions, we may define a region by profiling the conditional density of Z given $Y = y$, that is by assuming $D(z; y, \theta) = -f(z|y; \theta)$, or equal to any decreasing function of $f(z|y; \theta)$. Then, for a Gaussian random field, we can equivalently consider $D(Z; Y, \theta) = (Z - \mu_{Z|Y})^T \Sigma_{Z|Y}^{-1} (Z - \mu_{Z|Y})$, which is a pivotal quantity with respect to θ , following a chi-squared distribution with m degrees of freedom, abbreviated as $\chi^2(m)$. Thus, we have that

$$R_\alpha = \{z \in \mathbf{R}^m : (z - \mu_{Z|Y})^T \Sigma_{Z|Y}^{-1} (z - \mu_{Z|Y}) \leq d_\alpha\},$$

where the limit d_α is now free of θ and corresponds to χ_α^2 , namely the α -quantile of the $\chi^2(m)$ distribution. Since R_α depends on the unknown parameter θ , we use the associated plug-in estimator $\widehat{R}_\alpha = R_\alpha(\widehat{\theta}; y)$, called estimative highest prediction density region. Although we follow here the frequentist approach, we adopt this terminology, which is commonly used in the Bayesian context for posterior distributions, also for the estimated conditional density of the future random vector Z . Note that, at least in the Gaussian case, this region for $m = 2$ has an elliptical form, while, for $m = 1$, we obtain a prediction interval which corresponds to that one derived from the estimative prediction limits defined above. Moreover, this choice for the root

$D(Z; Y, \theta)$ is very convenient, since the associated conditional distribution is known under the true parameter value θ .

Finally, we mention the fact that, also within a non-Gaussian spatial model it is possible to specify highest prediction density regions, even if the conditional distribution of the root $D(Z; Y, \theta)$ could be difficult or impossible to obtain. In this case we may approximate the unknown probability distribution using a bootstrap procedure, as suggested by Beran (1990), or we may consider an alternative expression for $D(Z; Y, \theta)$, usually a simple approximation for $-f(z|y; \theta)$ with a known probability distribution (see, for the case with Y and Z independent, Fonseca et al., 2012, Section 5.2). Furthermore, it could be useful to specify the root as a different univariate summary statistic of the not yet observed random vector Z , such as $\sum_{i=1}^m h_i(Z_i)$, with $h_i(\cdot)$ a suitable real-valued function. Although prediction may not be simple, since we could not apply directly the Gaussian process theory, in this case the problem comes back in the univariate framework and it may be solved by considering the univariate conditional distribution of the summary statistic given $Y = y$, which can be known at least approximatively or, if necessary, estimated using simulation-based methods. As a remarkable example, we may consider the problem of predicting the spatial average of a Gaussian random field over a subset of the region of interest Δ (see, for example, De Oliveira and Kone, 2015). In this case, according to the L^2 integration theory for random fields, the spatial average is a random variable defined as a stochastic integral, which can be approximated by a weighted average of (future) observations for the random field in suitable multiple locations inside the integration domain.

3 Calibrated prediction regions and distributions

The calibrating approach proposed by Fonseca et al. (2014), extending that one suggested by Beran (1990), provides predictive distribution functions which give, as quantiles, prediction limits with well-calibrated coverage probability. Thus, the coverage error turns out to be substantially reduced, with respect to that one given by the estimative solution. This proposal is here applied to the case of Gaussian random fields, by considering the estimative prediction limit \hat{z}_α and the estimative prediction region \hat{R}_α defined in the previous section.

We initially consider the univariate case. The estimative prediction limit may be interpreted as the α -quantile of the estimative distribution function $F(z|y; \hat{\theta})$, namely $\hat{z}_\alpha = F^{-1}(\alpha|y; \hat{\theta})$, where $F^{-1}(\cdot|y; \hat{\theta})$ is the inverse of function $F(\cdot|y; \hat{\theta})$. The associated coverage probability is such that

$$P_{Y,Z}(Z \leq \hat{z}_\alpha; \theta) = E_Y\{F(\hat{z}_\alpha|Y; \theta); \theta\} = C(\alpha, \theta), \quad (2)$$

where the expectation is with respect to the distribution of Y under θ . An explicit expression for the coverage probability is rarely available, but it is known that it does not match the target value α . Whenever the asymptotic behaviour of the estimator $\hat{\theta}$ is such that $\hat{\theta} - \theta = O_p(n^{-1/2})$, we find that $C(\alpha, \theta) = \alpha + Er(\alpha, \theta) + o(n^{-1})$, as $n \rightarrow +\infty$. Thus, there is a coverage error $Er(\alpha, \theta)$, usually of order $O(n^{-1})$, which is not negligible. For Gaussian models, a general expression for the first-order error term may be obtained (see, for example, Sjöstedt-de Luna and Young, 2003; Giummolè and Vidoni, 2010) but its explicit specification requires a substantial computational effort.

In this framework, Fonseca et al. (2014) prove that the function

$$F_c(z|y; \hat{\theta}, \theta) = C\{F(z|y; \hat{\theta}), \theta\},$$

obtained by substituting α with $F(z|y; \hat{\theta})$ in $C(\alpha, \theta)$, is a proper predictive distribution function, at least in regular parametric models. Moreover, it gives, as quantiles, prediction limits $z_\alpha^c = z_\alpha^c(\hat{\theta}, \theta; y)$ having coverage probability exactly α , for all $\alpha \in (0, 1)$. The proof is straightforward, since, for each $\alpha \in (0, 1)$,

$$z_\alpha^c = F_c^{-1}(\alpha|y; \hat{\theta}, \theta) = F^{-1}\{C^{-1}(\alpha, \theta)|y; \hat{\theta}\} = \hat{z}_{\alpha^c},$$

with $\alpha^c = C^{-1}(\alpha, \theta)$, where $C^{-1}(\cdot, \theta)$ and $F_c^{-1}(\cdot|y; \hat{\theta}, \theta)$ are the inverse of functions $C(\cdot, \theta)$ and $F_c(\cdot|y; \hat{\theta}, \theta)$, respectively. The limit z_α^c corresponds to the calibrated prediction limit defined by Beran (1990) and it fulfils relation (1) exactly.

Although interesting from a theoretical point of view, the calibrated predictive distribution $F_c(z|y; \hat{\theta}, \theta)$, depending on the unknown parameter θ , is in fact useless. For this reason, we may consider the corresponding plug-in estimator

$$\hat{F}_c(z|y; \hat{\theta}) = C\{F(z|y; \hat{\theta}), \hat{\theta}\},$$

which gives, for each $\alpha \in (0, 1)$, prediction limits $\hat{z}_\alpha^c = z_\alpha^c(\hat{\theta}, \hat{\theta}; y)$ improving the estimative ones, since they satisfy (1) with an error term of order $o(n^{-1})$, as proved by Beran (1990) under suitable regularity conditions.

This predictive procedure is easily applicable when an explicit expression for the coverage probability $C(\alpha, \theta)$ is available, at least to order $O(n^{-1})$. In this case, $\hat{F}_c(z|y; \hat{\theta})$ corresponds, to the relevant order of approximation, to the predictive distribution proposed by Barndorff-Nielsen and Cox (1996) and Vidoni (1998), using asymptotic calculations. However, these

asymptotic solutions are usually not easy to obtain, so that, as a useful, simpler alternative, we can define a parametric bootstrap estimator for $F_c(z|y; \hat{\theta}, \theta)$. Let y_b^* , $b = 1, \dots, B$, be parametric bootstrap samples generated from the distribution of Y , assuming $\theta = \hat{\theta}$, and let $\hat{\theta}_b^*$, $b = 1, \dots, B$, be the corresponding maximum likelihood estimates. Since $C(\alpha, \theta)$ is defined in (2) as a suitable expectation, we may calculate the bootstrap-calibrated predictive distribution as

$$\tilde{F}_c(z|y; \hat{\theta}) = \frac{1}{B} \sum_{b=1}^B F(\tilde{z}_{\alpha,b}^* | y_b^*; \hat{\theta})|_{\alpha=F_c(z|y; \hat{\theta})},$$

where $\tilde{z}_{\alpha,b}^* = z_\alpha(\hat{\theta}_b^*; y_b^*)$. The corresponding α -quantile defines, for each $\alpha \in (0, 1)$, a prediction limit having coverage probability equal to the target value α , with an error term which depends on the efficiency of the bootstrap simulation procedure.

This approach can be readily applied for improving the estimative prediction region $\hat{R}_\alpha = R_\alpha(\hat{\theta}; Y)$, which has coverage probability

$$\begin{aligned} P_{Y,Z}\{Z \in R_\alpha(\hat{\theta}; Y); \theta\} &= E_Y[P_{Z|Y}\{Z \in R_\alpha(\hat{\theta}; Y)|Y; \theta\}; \theta] \\ &= E_Y\{\hat{F}(\hat{d}_\alpha|Y; \theta); \theta\} = K(\alpha, \theta). \end{aligned}$$

Here, $\hat{F}(d|y; \theta)$ is the conditional distribution function of $\hat{D} = D(Z; Y, \hat{\theta})$ given $Y = y$, with θ the true parameter value, and the quantity, \hat{d}_α is such that $\hat{F}(\hat{d}_\alpha|Y; \hat{\theta}) = \alpha$. Note that, with the highest prediction density region, $\hat{d}_\alpha = \chi_\alpha^2$ and both the conditional distribution of $D = (Z - \mu_{Z|Y})^T \Sigma_{Z|Y}^{-1} (Z - \mu_{Z|Y})$ under θ and the conditional distribution of $\hat{D} = (Z - \hat{\mu}_{Z|Y})^T \hat{\Sigma}_{Z|Y}^{-1} (Z - \hat{\mu}_{Z|Y})$ under $\hat{\theta}$ follow a $\chi^2(m)$ distribution. On the other hand, this is not true for the conditional distribution of \hat{D} under θ . Here, as usual, a hat means evaluation at $\theta = \hat{\theta}$. Furthermore, the coverage probability $K(\alpha, \theta)$, as already mentioned, does not match the target value α , since $\hat{F}(\hat{d}_\alpha|Y; \theta)$ turns out to be different from α . Thus, as expected, we

find for \widehat{R}_α the same drawback of the estimative prediction limit introduced in the univariate case.

A simple general procedure, which gives calibrated estimative prediction regions, achieving exactly the target coverage probability α , for all $\alpha \in (0, 1)$, can be derived as in the unidimensional case. In particular, we define the calibrated conditional distribution function for \widehat{D} as

$$\widehat{F}_c(d|y; \widehat{\theta}, \theta) = K\{\widehat{F}(d|y; \widehat{\theta}), \theta\}, \quad (3)$$

which is obtained by substituting α with $\widehat{F}(d|y; \widehat{\theta})$ in $K(\alpha, \theta)$. Function (3) is a proper distribution function and the associated α -quantile $d_\alpha^c = d_\alpha^c(\widehat{\theta}, \theta; y)$ determines the calibrated estimative prediction region

$$R_\alpha^c = R_\alpha^c(\widehat{\theta}, \theta; y) = \{z \in \mathbf{R}^m : D(z; y, \widehat{\theta}) \leq d_\alpha^c\},$$

having coverage probability equal to the target value α . The proof is straightforward and it is similar to that one presented for the calibrated estimative prediction limit.

Since an explicit, exact or approximate, expression for the coverage probability $K(\alpha, \theta)$ is usually not available, also function (3) is not explicitly known. Thus, it has to be properly estimated using, for example, a parametric bootstrap simulation procedure similar to that one previously outlined for the univariate case. More precisely, let (y_b^*, z_b^*) , $b = 1, \dots, B$, be parametric bootstrap samples generated from the distribution of (Y, Z) , assuming $\theta = \widehat{\theta}$, and let $\widehat{\theta}_b^*$, $b = 1, \dots, B$, be the corresponding maximum likelihood estimates based on y_b^* . If we consider the highest prediction density region, the bootstrap estimator for (3) is

$$\widetilde{F}_c(d|y; \widehat{\theta}) = \frac{1}{B} \sum_{b=1}^B \mathbf{1}_{E_b^*}(z_b^*)|_{\alpha=\widehat{F}(d|y; \widehat{\theta})}, \quad (4)$$

with $\mathbf{1}_{E_b^*}(\cdot)$ the indicator function of the set $E_b^* = \{z \in \mathbf{R}^m : D(z; y_b^*, \hat{\theta}_b^*) \leq \chi_\alpha^2\}$. Note that $\tilde{F}_c(d|y; \hat{\theta})$ may be viewed as the empirical distribution function obtained from the bootstrap sample $d_b^* = D(z_b^*; y_b^*, \hat{\theta}_b^*)$, $b = 1, \dots, B$. The associated α -quantile defines, for each $\alpha \in (0, 1)$, a quantity which specifies a prediction region with the required coverage probability, with an error term which depends on the efficiency of the bootstrap simulation procedure. The bootstrap-based estimate (4) is different from that one considered in the univariate case and it is more involved, since the simulated values for Z are also required. As emphasized before, using the highest prediction density region, the conditional distribution of \hat{D} under θ is not a $\chi^2(m)$ distribution and then, contrary to the univariate case, the function $\hat{F}(d|y; \theta)$, which is required for calculating $K(\alpha, \theta)$, is not available.

Moreover, it is important to notice that the computation of the bootstrap estimate (4) for the calibrated conditional distribution function does not require neither assumptions on the regularity of the parametric model nor the validity of the usual asymptotic results for the estimator $\hat{\theta}$. Thus, it provides a generalization with respect to the improved predictive procedures based on asymptotic arguments and it can be usefully considered also when the spatial asymptotic regime, as in the fixed domain case, does not assure the validity of the standard large sample properties for $\hat{\theta}$. In this bootstrap-based predictive framework, the accuracy of the predictive solution mainly depends on the efficiency of the bootstrap simulation technique, which is clearly related to the quality of the estimator for the model parameter θ .

A further interesting application of these improved predictive distributions concerns the problem of identifying spatial locations where the process of interest presents an extreme or an unusual observed value. In this context, whenever a threshold defining the exceedance is not

specified or not even rationally specifiable, we could qualify a new univariate or multivariate observation z_{obs} , related to further locations, according to the associated predictive distribution for Z given the observed spatial data y . With a slight abuse of notation, we set as z_{obs} this vector of further observations even though, in this case, it is actually part of a larger observed dataset, which includes also the observations y .

More precisely, in the univariate case, we shall classify z_{obs} as an unusual high or low observation by considering the probability of observing a value for Z higher or lower than z_{obs} , which can be efficiently estimated by $1 - \tilde{F}_c(z_{obs}|y; \hat{\theta})$ or $\tilde{F}_c(z_{obs}|y; \hat{\theta})$, respectively. A low value for one of these probabilities identifies a potential extreme observation. In the multivariate case, using the highest prediction density region, the observation z_{obs} is considered as extreme with regard to the probability of observing a value for \hat{D} higher than $d_{obs} = (z_{obs} - \hat{\mu}_{Z|Y})^T \hat{\Sigma}_{Z|Y}^{-1} (z_{obs} - \hat{\mu}_{Z|Y})$, which is accurately estimated by $1 - \tilde{F}_c(d_{obs}|y; \hat{\theta})$. Here, the exceedance is interpreted as a substantial distance from the mean $\hat{\mu}_{Z|Y}$, without specifying a particular direction. In both cases, we evaluate these probabilities by considering the calibrated predictive distribution function, usually estimated using a bootstrap-based procedure, which provides more accurate predictive statements than those given by the estimative predictive distribution function.

Finally, although in this paper we consider Gaussian models and functions $D(z; y, \theta)$ which are usually pivotal quantities, we underline that the improved procedures presented in this section can be applied also to non-Gaussian models and to more general root functions D . An additional computational effort will be usually required in order to estimate, using simulation-based methods, the conditional distribution function of \hat{D} under $\hat{\theta}$, from which we obtain the estimative limit \hat{d}_α .

4 A review of further calibration procedures

The bootstrap calibrated procedure considered in the present paper, for improved prediction in Gaussian random fields, is quite general, relatively simple to apply and effective in a number of applications, including potentially also the case of observations collected under the fixed domain asymptotic regime. However, other calibration procedures have been proposed in the spatial framework, with particular regard to the one-dimensional case. All these procedures share the objective of improving the estimative predictive solution and they are based on the evaluation of the associated coverage probability, which is usually different from the target probability α . In this review section, we focus on prediction intervals and prediction limits, since the problem of specifying multivariate prediction regions reduces to a problem in one dimension, by using a suitable root function defined as a univariate summary of the future sample Z .

As emphasized at the beginning of Section 3, with concern to the univariate case, the coverage probability of the estimative prediction limit \widehat{z}_α is a quantity $C(\alpha, \theta)$, defined as in (2), which does not match the target value α . A first calibration strategy consist in finding the value $\alpha^c \in (0, 1)$ such that $C(\alpha^c, \theta) = \alpha$, exactly or approximately, and using \widehat{z}_{α^c} instead of \widehat{z}_α , since in this case we obviously achieve the target probability α . Whenever the coverage probability is not explicitly known, it can be estimated using, for example, a parametric bootstrap procedure. This simple calibration technique, called indirect by Ueki and Fueda (2007), was discussed in Beran (1990) and applied by Sjöstedt-de Luna and Young (2003) and De Oliveira and Rui (2009) to Gaussian and log-Gaussian random field models, respectively.

The improved predictive solution considered in this paper generalizes, in some sense, this first simple solution, since the calibrated predictive distribution $F_c(z|y; \widehat{\theta}, \theta)$ is defined as $C(\alpha, \theta)$

evaluated at $\alpha = F(z|y; \hat{\theta})$ and it gives, as quantiles, prediction limits $z_\alpha^c = z_\alpha^c(\hat{\theta}, \theta; y)$ having coverage probability exactly α , for all $\alpha \in (0, 1)$. In Section 3 we prove that $z_\alpha^c = \hat{z}_{\alpha^c}$ and then we conclude that the improved prediction limits defined before can be obtained as quantiles of this improved predictive distribution. Note that these two strategies do not involve asymptotic arguments and they are widely applicable. Whenever the coverage probability is unknown, the accuracy of the improved prediction limits depends on the efficiency of the simulation-based procedure which is considered for estimating $C(\alpha, \theta)$.

A further calibration strategy, called direct by Ueki and Fueda (2007), aims at defining an explicit modification for the estimative prediction limit, so that the associated coverage probability turns out to be equal to the target α with a high degree of accuracy. These solutions are based on suitable asymptotic expansions and they require regularity assumptions, assuring the usual large sample properties for the estimator $\hat{\theta}$. More precisely, assuming the \sqrt{n} -consistency for $\hat{\theta}$, we find that $C(\alpha, \theta) = \alpha + Er(\alpha, \theta) + o(n^{-1})$, where the coverage error term $Er(\alpha, \theta)$ is usually of order $O(n^{-1})$.

In order to improve the estimative solution, Barndorff-Nielsen and Cox (1996) and Vidoni (1998) define the modified estimative prediction limit $\hat{z}_\alpha - Er(\alpha, \hat{\theta})/f(\hat{z}_\alpha|y; \hat{\theta})$, which reduces the coverage error to order $o(n^{-1})$. Moreover, it is not difficult to prove that the distribution function which gives this modified limit as α -quantile is $F(z|y; \hat{\theta}) + Er(F(z|y; \hat{\theta}), \hat{\theta})$, which is an high-order approximation for the calibrated predictive distribution $F_c(z|y; \hat{\theta}, \theta)$.

Although with this strategy we improve the coverage accuracy of the estimative solution, the evaluation of the fundamental term $Er(\alpha, \theta)$ may require complicated asymptotic calculations. For Gaussian models, a general expression for this term is obtained by Sjöstedt-de Luna and

Young (2003) and Giummolè and Vidoni (2010). To overcome this computational difficulty, Ueki and Fueda (2007) define the following asymptotically equivalent expression for the modifying term of the estimative prediction limit, namely $Er(\alpha, \theta)/f(z_\alpha|y; \theta) = z_{C(\alpha, \theta)} - z_\alpha + o(n^{-1})$, and consequently the equivalent expression $2\widehat{z}_\alpha - \widehat{z}_{C(\alpha, \theta)}$ for the modified estimative prediction limit. Here the computation is greatly simplified, since we need only a simulation-based estimate for the coverage probability $C(\alpha, \theta)$. With regard to this particular predictive strategy, Fonseca et al. (2012) find a simplified expression for the associated predictive distribution function and present an application to multivariate prediction regions based on univariate summaries, by considering the case of independence between Y and Z .

5 A simple simulation study

In this section we carry out a simple simulation study in order to compare the coverage probability of the plug-in highest prediction density regions with that one of the well-calibrated highest prediction density regions specified in this paper. In particular, we repeatedly simulate realizations of a Gaussian random field at n sampling locations uniformly distributed over a region $\Delta \subset \mathbf{R}^2$. We consider a mean function $\mu(s; \theta) = \beta_0 + \beta_1 s_x + \beta_2 s_y$, with $s = (s_x, s_y) \in \Delta$, $\beta_0, \beta_1, \beta_2 \in \mathbf{R}$, an isotropic exponential covariance function $\sigma(s, t; \theta) = \sigma^2 \exp(-\|s - t\|/\phi)$, with $s, t \in \Delta$, $\sigma^2, \phi > 0$ and $\|\cdot\|$ the Euclidean norm, and a nugget effect specified by $\tau^2 \geq 0$. Thus, following the notation adopted in Section 2, we consider a particular Gaussian random field with model parameter $\theta = (\theta_1, \dots, \theta_6) = (\beta_0, \beta_1, \beta_2, \sigma^2, \phi, \tau^2)$.

We run the simulation experiments using two alternative designs that mimic, respectively, the increasing domain and the fixed domain asymptotic regimes. In the first case, we simulate

data at $n = 50, 100, 200$ sampling locations in the corresponding nested regions $\Delta_{50} = [4, 6] \times [4, 6]$, $\Delta_{100} = [2, 8] \times [2, 8]$ and $\Delta_{200} = [0, 10] \times [0, 10]$. In the second case, we generate $n = 50, 100, 200$ observation, uniformly distributed in the same fixed sampling region $\Delta = [4, 6] \times [4, 6]$. In both sampling designs, we aim at predicting $m = 1, 5, 10, 25$ observations at the further m locations uniformly distributed over the smallest nested region Δ . Notice that, whenever $m = 1$, we obtain one-dimensional centred prediction intervals.

We estimate the coverage probabilities for the estimative and the calibrated highest prediction density regions of level $\alpha = 0.9, 0.95$ using 1,000 Monte Carlo samples from the Gaussian random field previously defined with $\beta_0 = 1$, $\beta_1 = 2$, $\beta_2 = 1.5$, $\sigma^2 = 1$ and (a) $\phi = 0.3$, $\tau^2 = 0.25$, (b) $\phi = 0.3$, $\tau^2 = 0$, (c) $\phi = 0.8$, $\tau^2 = 0.25$, (d) $\phi = 0.8$, $\tau^2 = 0$. In order to obtain the calibrated prediction regions, we generate, for each simulated sample, $B = 500$ further parametric bootstrap samples. Indeed, the unknown parameters are estimated using maximum likelihood estimators, which are proved to be consistent under the increasing domain asymptotic regime. As we recall in Section 2, this is not valid under the fixed domain asymptotic regime.

The results are presented in Table 1, for the increasing domain case, and in Table 2, for the fixed domain case. They show that, in almost all the sampling experiments, the improved prediction regions remarkably improve on the estimative ones. However, when we consider $\tau^2 = 0$, we observe that, for large m values, also the improved prediction regions fail to reach the target coverage probability, even if this unsatisfactory behaviour reduces as n increases. In general, we note that both coverage probabilities deteriorate when n is low and m grows, but the deterioration is substantially reduced for calibrated prediction regions. Since the simulation

procedure is very time-consuming, we are forced to consider a moderate number simulated samples and this causes a non negligible variability in the estimation procedure.

The theoretical findings presented in the preceding sections turn out to be confirmed by this simple Monte Carlo experiment. Moreover, at least for the parameter values specified in this simulation study, the coverage results for both the estimative and the calibrated prediction regions do not seem influenced by the different asymptotic regimes. These empirical findings do not have general validity and they confirm that, in some situations, the asymptotic performance of the predictive procedure improves as n increases also within the fixed domain asymptotic regime (see, for example, Sjöstedt-de Luna and Young, 2003, Section 8).

[Table 1 about here.]

[Table 2 about here.]

Furthermore, in Figure 1, we display the plug-in conditional distribution functions of \hat{D} given $Y = y$, which corresponds in this case to $\chi^2(m)$ distributions, and the associated calibrated conditional distribution functions of \hat{D} given $Y = y$, defined by (3) and estimated using a parametric bootstrap procedure based on 10,000 simulated samples.

[Figure 1 about here.]

We base our analysis on a single sample of dimension $n = 50$ obtained from a Gaussian random field, defined as before, with $\beta_0 = 1$, $\beta_1 = 2$, $\beta_2 = 1.5$, $\sigma^2 = 1$, $\phi = 0.3$, $\tau^2 = 0.25$. We consider highest prediction density regions of dimension $m = 1, 5, 10, 25$. It is immediate to see that the two distribution functions are definitely different and that the difference notably

increases with m . In particular, given a future observation z_{obs} , the probability of observing a value for \widehat{D} not greater than $d_{obs} = (z_{obs} - \widehat{\mu}_{Z|Y})^T \widehat{\Sigma}_{Z|Y}^{-1} (z_{obs} - \widehat{\mu}_{Z|Y})$ is systematically overestimated whenever we use the plug-in distribution function instead of the calibrated one. Similar results are obtained by assuming different values for the parameters and the unsatisfactory performance of the plug-in solution substantially maintains, even if we consider observed samples with a larger dimension n .

6 An application to air quality prediction

The spatial prediction of air quality levels is a relevant issue both under the regulatory point of view (Parliament of the European Union and Council of the European Union, 2008) and in terms of public health impact evaluation (WHO, 2004, 2006a,b).

Eulerian deterministic air pollution models combine meteorology records, emission inventories, characterization of chemical and physical processes in the atmosphere to give predictions of air pollution levels usually over a fine grid of locations. They are an essential tool to evaluate the impact of pollution control strategies or of new sources of air pollution (e.g. road, power plants, etc.). A deterministic model needs to be assessed using observed pollution data (Fuentes et al., 2003) usually measured by a network of air quality monitoring stations.

The use of the improved prediction method previously described is here illustrated by considering the annual average concentration levels of PM_{10} in the Lombardy region, Italy (Cecconi et al., 2016). The Lombardy region is located in Northern Italy and it is historically characterized, at least in its flat and more densely populated part (the Po river basin), by high levels of air pollution and, more specifically, by very high levels of PM_{10} (Baccini et al., 2015).

Predicted values of PM_{10} concentration levels have been derived, using a deterministic Eulerian photochemical model (Silibello et al., 2008), over a regular 16×16 km grid (which was, originally, a 4×4 km grid) covering the flat part of the region (see Figure 2 (a)). Thus, a total of 65 prediction points have been considered.

Observed data, referred to 2007, are based on 58 monitoring stations mainly located in urban areas (see Figure 2 (b)).

[Figure 2 about here.]

The present analysis aims at assessing the prediction ability of the Eulerian photochemical model using observed data from monitor stations. Following the proposal of Fuentes et al. (2003), we will use monitor data to predict pollution levels on the same grid used for deterministic prediction. Then, using both the estimative predictive distribution and the calibrated predictive distribution proposed in this paper, we will obtain two different evaluations of the degree of consistency of the predictions given by the deterministic model.

In accordance with the literature on spatial modeling of air pollution levels (see, for example, Cameletti et al., 2013), we assume that the logarithm transformations of the PM_{10} concentration levels, measured at the monitoring stations, are a realization of a Gaussian random field with unknown covariance function. Figure 3 reports the empirical variogram and the maximum likelihood estimate of the most frequently used variogram models based on the available concentration levels. Goodness of fit of the fitted models was evaluated using cross-validation. Both the exponential and the complete Matérn models show a good fit (indeed the difference between them is negligible), and the first one is used for the specification of the Gaussian random field. We have then considered the predicted concentration levels referred to the 65

prediction points, computed according to the Eulerian photochemical model, as a new multivariate observation z_{obs} . Its agreement with monitor data has been appraised by computing the corresponding pivotal quantity d_{obs} and evaluating it under both the plug-in χ^2 distribution and the calibrated predictive distribution (3), estimated by using 10,000 parametric bootstrap simulated samples.

[Figure 3 about here.]

As can be seen from Figure 4, the value of the pivotal quantity d_{obs} can be considered high under both the distributions. However, the order of magnitude is completely different. Indeed, the probability of observing a value greater than the observed one is 0.006 under the plug-in distribution, clearly showing a very high degree of inconsistency of the multivariate value z_{obs} given by the deterministic model. On the contrary, the same probability corresponds to 0.091, when we consider the calibrated conditional distribution. Although low, this value is higher than the conventional 0.05 level, generally used as a cut-off.

Thus, the estimative and the improved procedures can bring two different results: with the first, we would conclude against the validity of the Eulerian model, while with the second we get a sufficient level of agreement between the Eulerian and the spatial statistical model. This evidence supports as-well the relevance and the necessity of a suitable calibration technique in order to improve the accuracy of predictive methods.

[Figure 4 about here.]

7 Conclusion

Spatial prediction is one of the main tasks of geostatistics. The usual procedure when a Gaussian model is defined is based on kriging prediction, using estimates in place of the unknown parameter values. As a consequence, the uncertainty on the estimated parameter values is not taken into account, and this leads to prediction intervals and regions with coverage probability usually lower than the nominal one.

We have developed a method to construct well-calibrated frequentist prediction regions based on the calibrating procedure proposed by Beran (1990) and developed by Fonseca et al. (2014). The performance of the proposed procedure has been evaluated by means of a simulation study that clearly indicates that the proposed calibration method outperforms the standard estimative procedure. The two methods have also been used to compare geostatistical predictions of PM_{10} annual mean levels based on a network of monitoring stations and those coming from a deterministic Eulerian photochemical model. Again, the results indicate the need for calibration to avoid inaccurate conclusions based on too narrow prediction regions.

This calibration method can be easily generalized to more complex situations, such as the joint prediction of more than one variable (Ver Hoef and Cressie, 1993) and prediction in spatial regression models and in a more general space-time framework (Heuvelink et al., 2015).

Acknowledgements

The comments and suggestions of the referee and of the associate editor considerably improve the presentation of the results. The authors thank the Air Quality Unit of the Lombardy Regional Agency for the Environment and the Health Directorate of the Region of Lombardy

Government for providing the data that made this work possible. The research of C. Lagazio was partially supported by the Italian Ministry for University and Research under the PRIN2015 grant No. 20152T74ZL_003. The research of P. Vidoni was partially supported by the Italian Ministry for University and Research under the PRIN2015 grant No. 2015EASZFS_003.

References

- Baccini, M., Grisotto, L., Catelan, D., Consonni, D., Bertazzi, P. A., and Biggeri, A. (2015). Commuting-adjusted short-term health impact assessment of airborne fine particles with uncertainty quantification via monte carlo simulation. *Environmental Health Perspectives*, 123:27–33.
- Barndorff-Nielsen, O. and Cox, D. R. (1996). Prediction and asymptotics. *Bernoulli*, 2:319–340.
- Beran, R. (1990). Calibrating prediction regions. *Journal of the American Statistical Association*, 85:715–723.
- Cameletti, M., Lindgren, F., Simpson, D., and Rue, H. (2013). Spatio-temporal modeling of particulate matter concentration through the spde approach. *ASTA Advances in Statistical Analysis*, 97:109–131.
- Cecconi, L., Grisotto, L., Catelan, D., Lagazio, C., Berrocal, V., and Biggeri, A. (2016). Preferential sampling and bayesian geostatistics: Statistical modeling and examples. *Statistical Methods in Medical Research*, 25:1224–1243.
- De Oliveira, V. and Kone, B. (2015). Prediction intervals for integrals of Gaussian random fields. *Computational Statistics & Data Analysis*, 83:37–51.

- De Oliveira, V. and Rui, C. (2009). On shortest prediction intervals in log-Gaussian random fields. *Computational Statistics & Data Analysis*, 53:4345–4357.
- Fonseca, G., Giummolè, F., and Vidoni, P. (2012). A note about calibrated prediction regions and distributions. *Journal of Statistical Planning and Inference*, 142:2726–2734.
- Fonseca, G., Giummolè, F., and Vidoni, P. (2014). Calibrating predictive distributions. *Journal of Statistical Computation and Simulation*, 84:373–383.
- Fuentes, M., Guttorp, P., and Challenor, P. (2003). Statistical assessment of numerical models. *International Statistical Review*, 71(2):201–221.
- Giummolè, F. and Vidoni, P. (2010). Improved prediction limits for a general class of Gaussian models. *Journal of Time Series Analysis*, 31:483–493.
- Heuvelink, G. B. M., Pebesma, E., and Gräler, B. (2015). Space-time geostatistics. In Shekhar, S., Xiong, H., and Zhou, X., editors, *Encyclopedia of GIS*, pages 1–7. Springer International Publishing, Cham.
- Mardia, K. V. and Marshall, R. (1984). Maximum likelihood estimation of models of residual covariance in spatial regression. *Biometrika*, pages 135–146.
- Parliament of the European Union and Council of the European Union (2008). Directive 2008/50/ec of the european parliament and of the council of 21 may 2008 on ambient air quality and cleaner air for europe. *Official Journal of the European Union*, L 152/1.
- Schelin, L. and Sjöstedt-de Luna, S. (2010). Kriging prediction intervals based on semiparametric bootstrap. *Mathematical Geosciences*, 42:985–1000.

- Silibello, C., Calori, G., Brusasca, G., Giudici, A., Angelino, E., Fossati, G., Peroni, E., and Buganza, E. (2008). Modelling of pm10 concentrations over milano urban area using two aerosol modules. *Environmental Modelling and Software*, 23:333–343.
- Sjöstedt-de Luna, S. and Young, A. (2003). The bootstrap and kriging prediction intervals. *Scandinavian Journal of Statistics*, 30:175–192.
- Ueki, M. and Fueda, F. (2007). Adjusting estimative prediction limits. *Biometrika*, 94:509–511.
- Ver Hoef, J. M. and Cressie, N. (1993). Multivariable spatial prediction. *Mathematical Geology*, 25:219–240.
- Vidoni, P. (1998). A note on modified estimative prediction limits and distributions. *Biometrika*, 85:949–953.
- WHO (2004). *Health aspects of air pollution. Results from WHO project “Systematic review of health aspects of air pollution in Europe”*. World Health Organization.
- WHO (2006a). *Air Quality Guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. Global update 2005. Summary of risk assessment*. World Health Organization.
- WHO (2006b). *Air Quality Guidelines. Global update 2005*. World Health Organization.
- Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistic. *Journal of the American Statistical Association*, 99:250–261.

Table 1: Gaussian random field with isotropic exponential covariance function and $\beta_0 = 1$, $\beta_1 = 2$, $\beta_2 = 1.5$, $\sigma^2 = 1$ (a) $\phi = 0.3$, $\tau^2 = 0.25$, (b) $\phi = 0.3$, $\tau^2 = 0$, (c) $\phi = 0.8$, $\tau^2 = 0.25$, (d) $\phi = 0.8$, $\tau^2 = 0$. Coverage probabilities for the plug-in and the calibrated highest prediction density regions of level $\alpha = 0.9, 0.95$, with $m = 1, 5, 10, 25$. Estimation is based on 1,000 Monte Carlo samples of dimension $n = 50, 100$ and further $B = 500$ bootstrap samples. Estimated standard errors are smaller than 0.015. Asymptotic regime: increasing domain.

			(a)		(b)		(c)		(d)		
α	n	m	Plug-in	Calib	Plug-in	Calib	Plug-in	Calib	Plug-in	Calib	
0.9	50	1	0.848	0.897	0.846	0.901	0.856	0.905	0.868	0.905	
		5	0.837	0.893	0.865	0.912	0.835	0.900	0.889	0.913	
		10	0.804	0.898	0.854	0.922	0.791	0.895	0.870	0.927	
		25	0.738	0.912	0.793	0.924	0.716	0.903	0.829	0.935	
	100	1	0.904	0.907	0.896	0.910	0.905	0.909	0.909	0.913	
		5	0.881	0.908	0.908	0.918	0.878	0.897	0.920	0.914	
		10	0.853	0.885	0.876	0.889	0.849	0.881	0.901	0.899	
		25	0.811	0.887	0.871	0.918	0.811	0.880	0.903	0.913	
	200	1	0.892	0.889	0.892	0.891	0.897	0.889	0.896	0.888	
		5	0.887	0.894	0.902	0.906	0.886	0.894	0.913	0.903	
		10	0.860	0.898	0.893	0.898	0.859	0.888	0.905	0.899	
		25	0.846	0.888	0.878	0.901	0.834	0.886	0.906	0.909	
	0.95	50	1	0.920	0.941	0.918	0.940	0.917	0.949	0.921	0.950
			5	0.903	0.953	0.919	0.951	0.904	0.953	0.926	0.960
			10	0.873	0.949	0.913	0.962	0.866	0.940	0.926	0.966
			25	0.811	0.959	0.862	0.963	0.793	0.954	0.886	0.969
100		1	0.951	0.957	0.948	0.955	0.955	0.956	0.954	0.953	
		5	0.936	0.953	0.946	0.955	0.930	0.951	0.954	0.954	
		10	0.909	0.932	0.928	0.940	0.899	0.933	0.939	0.942	
		25	0.882	0.939	0.921	0.956	0.877	0.937	0.942	0.955	
200		1	0.948	0.950	0.945	0.946	0.947	0.944	0.953	0.947	
		5	0.935	0.941	0.947	0.946	0.930	0.942	0.952	0.944	
		10	0.926	0.942	0.947	0.951	0.919	0.945	0.953	0.951	
		25	0.899	0.934	0.935	0.953	0.893	0.931	0.951	0.955	

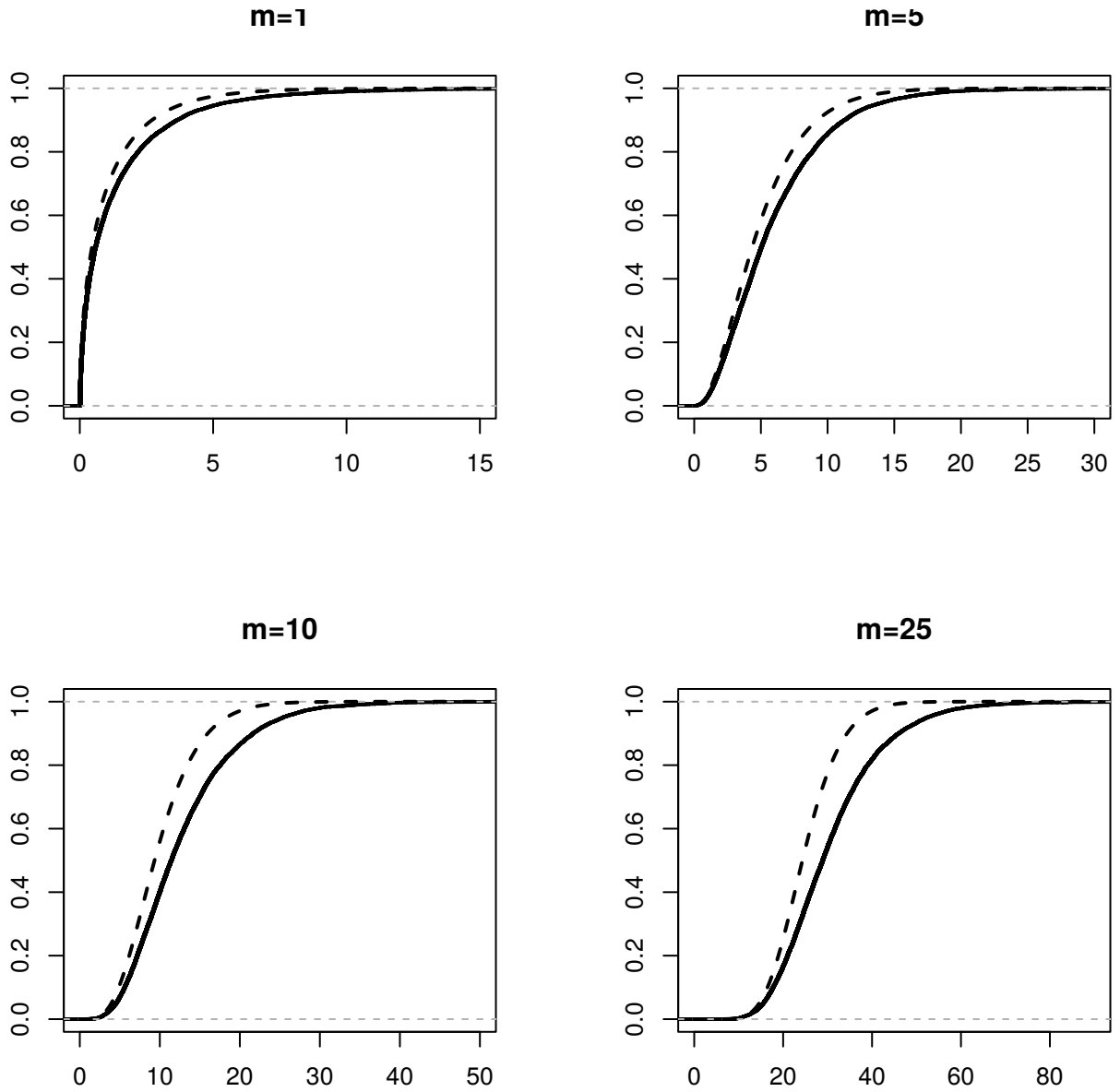


Figure 1: Conditional distribution functions of \widehat{D} given $Y = y$ based on a sample of dimension $n = 50$ obtained from a Gaussian random field with isotropic exponential covariance function and $\beta_0 = 1$, $\beta_1 = 2$, $\beta_2 = 1.5$, $\sigma^2 = 1$, $\phi = 0.3$, $\tau^2 = 0.25$: calibrated conditional distribution function (solid line), estimated using 10,000 parametric bootstrap simulated samples, and plug-in conditional distribution function (dashed line) for $m = 1, 5, 10, 25$.

Table 2: Gaussian random field with isotropic exponential covariance function and $\beta_0 = 1$, $\beta_1 = 2$, $\beta_2 = 1.5$, $\sigma^2 = 1$ (a) $\phi = 0.3$, $\tau^2 = 0.25$, (b) $\phi = 0.3$, $\tau^2 = 0$, (c) $\phi = 0.8$, $\tau^2 = 0.25$, (d) $\phi = 0.8$, $\tau^2 = 0$. Coverage probabilities for the plug-in and the calibrated highest prediction density regions of level $\alpha = 0.9, 0.95$, with $m = 1, 5, 10, 25$. Estimation is based on 1,000 Monte Carlo samples of dimension $n = 50, 100$ and further $B = 500$ bootstrap samples. Estimated standard errors are smaller than 0.015. Asymptotic regime: fixed domain.

			(a)		(b)		(c)		(d)		
α	n	m	Plug-in	Calib	Plug-in	Calib	Plug-in	Calib	Plug-in	Calib	
0.9	50	1	0.870	0.899	0.888	0.908	0.868	0.901	0.899	0.909	
		5	0.850	0.901	0.863	0.902	0.836	0.891	0.883	0.914	
		10	0.809	0.908	0.862	0.924	0.808	0.902	0.891	0.929	
		25	0.742	0.895	0.820	0.918	0.729	0.886	0.857	0.928	
	100	1	0.896	0.910	0.895	0.908	0.902	0.907	0.908	0.910	
		5	0.889	0.911	0.900	0.915	0.891	0.911	0.905	0.914	
		10	0.876	0.896	0.882	0.902	0.872	0.900	0.898	0.904	
		25	0.820	0.895	0.869	0.910	0.821	0.897	0.890	0.914	
	200	1	0.893	0.891	0.892	0.890	0.898	0.897	0.892	0.891	
		5	0.900	0.909	0.915	0.912	0.905	0.914	0.914	0.911	
		10	0.879	0.901	0.905	0.909	0.880	0.903	0.909	0.910	
		25	0.865	0.891	0.897	0.905	0.869	0.895	0.903	0.905	
	0.95	50	1	0.932	0.946	0.938	0.947	0.931	0.949	0.944	0.948
			5	0.905	0.951	0.918	0.955	0.896	0.949	0.934	0.955
			10	0.887	0.957	0.920	0.962	0.881	0.952	0.941	0.965
			25	0.815	0.941	0.872	0.960	0.803	0.931	0.906	0.967
100		1	0.952	0.961	0.954	0.960	0.953	0.957	0.953	0.960	
		5	0.936	0.954	0.938	0.954	0.941	0.957	0.946	0.957	
		10	0.923	0.944	0.927	0.945	0.922	0.944	0.936	0.948	
		25	0.887	0.948	0.918	0.953	0.890	0.948	0.934	0.956	
200		1	0.942	0.946	0.942	0.950	0.947	0.948	0.947	0.946	
		5	0.940	0.945	0.948	0.947	0.939	0.946	0.950	0.949	
		10	0.929	0.945	0.939	0.941	0.934	0.949	0.944	0.943	
		25	0.909	0.942	0.941	0.952	0.913	0.946	0.949	0.952	

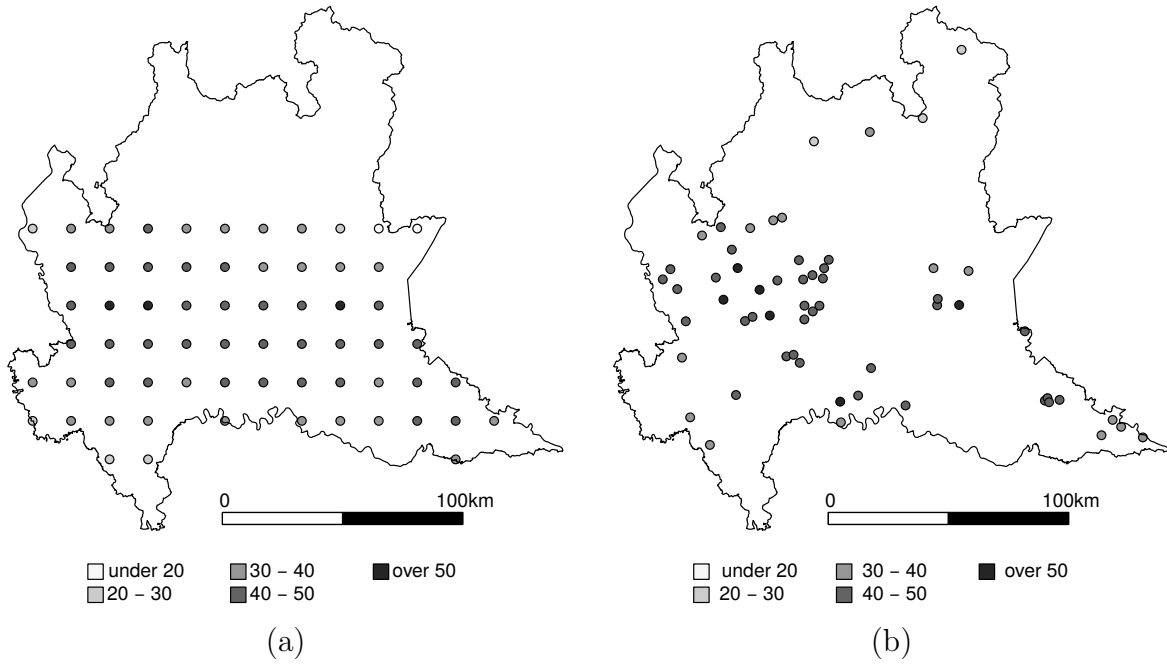


Figure 2: (a) Grid with 65 spatial locations used for prediction with PM₁₀ concentration levels derived from the Eulerian photochemical model and (b) Monitor locations and measured concentration levels of PM₁₀ ($\mu\text{g}/\text{m}^3$) in the Lombardy region in 2007.

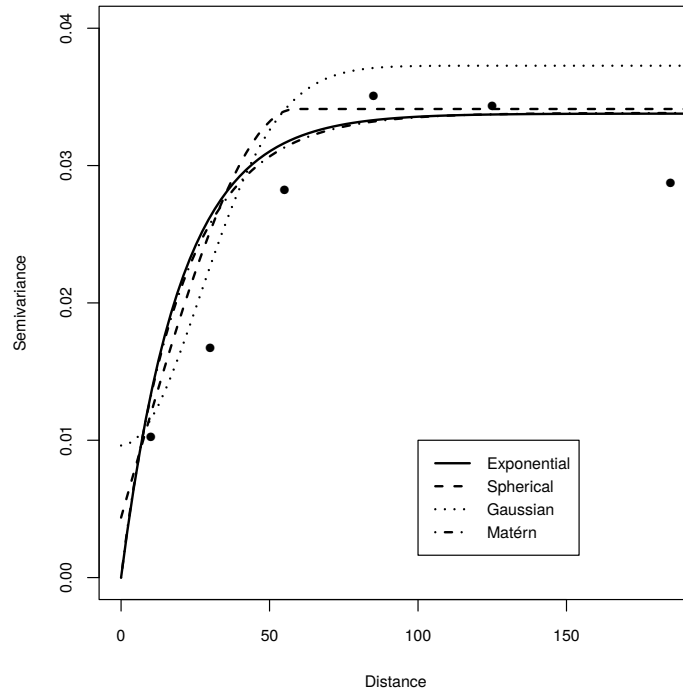


Figure 3: Empirical variogram (dots) and maximum likelihood estimates of theoretical variograms based on alternative models. Data on the concentration levels of PM_{10} ($\mu\text{g}/\text{m}^3$) in the Lombardy region in 2007.

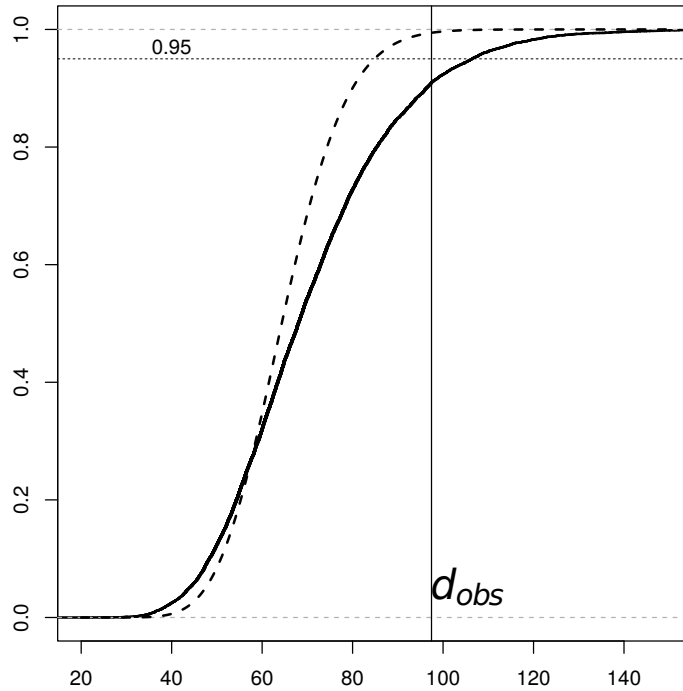


Figure 4: Conditional distribution functions of \hat{D} given $Y = y$ based on a sample of dimension $n = 58$ obtained from a Gaussian random field with isotropic exponential covariance function and parameters equal to the estimated ones: calibrated predictive distribution function (solid line), estimated using 10,000 parametric bootstrap simulated samples, and plug-in predictive distribution function (dashed line). The vertical line corresponds to the observed value of the pivotal quantity d_{obs} , computed using the predicted values given by the Eulerian photochemical model.