Published in final edited form as:
Comput Stat Data Anal. 2015 March 1; 83: 37–51. doi:10.1016/j.csda.2014.09.013.

PREDICTION INTERVALS FOR INTEGRALS OF GAUSSIAN RANDOM FIELDS

Victor De Oliveira[1] and **Bazoumana Kone**
Department of Management Science and Statistics, The University of Texas at San Antonio, San Antonio, TX 78249, USA

Abstract

Methodology is proposed for the construction of prediction intervals for integrals of Gaussian random fields over bounded regions (called block averages in the geostatistical literature) based on observations at a finite set of sampling locations. Two bootstrap calibration algorithms are proposed, termed indirect and direct, aimed at improving upon plug-in prediction intervals in terms of coverage probability. A simulation study is carried out that illustrates the effectiveness of both procedures, and these procedures are applied to estimate block averages of chromium traces in a potentially contaminated region in Switzerland.

Keywords

Block average; Bootstrap calibration; Change of support problem; Geostatistics; Kriging; Spatial average

1 Introduction

In this work we consider the problem of constructing prediction intervals for the integral of a spatially varying quantity over a bounded region (also called block average in the geostatistical literature), based on observations at a finite set of sampling locations. This problem is of importance in many earth sciences, such as hydrology, mining and pollution assessment, where interest often centers on spatial averages (rather than on ensemble averages). This was, in a mining context, a problem that D.G. Krige considered and that motivated G. Matheron to develop the geostatistical methodology named after him (kriging): estimating the average bock ore-grade over a mining panel based of measurements of internal core-samples (Cressie, 1990; Chilés and Delfiner, 1999). As the data support are ‘points’ while the support of the quantity of interest is a region of positive area, this is an example of what is generically called a change of support problem; see Gotway and Young (2002) for an extensive review.

[1]Corresponding author: Department of Management Science and Statistics, The University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, USA, victor.deoliveira@utsa.edu, Phone number: 1 210 4586592.

The problem of 'point prediction' of an integral of a random field over a bounded region has been considered extensively in the literature, for instance, by Cressie (1993), Chilés and Delfiner (1999), Cressie (2006), De Oliveira (2006) and Gotway and Young (2007). But the problem of 'interval prediction' has not received similar attention, and it could even be argued that it has not been adequately explored. When the model covariance parameters are *not* known, the common practice in the above works is to use a two-stage approach: the covariance parameters are first estimated and then prediction intervals are computed by treating these estimates as if they were the true covariance parameters. This is called the plug-in (or estimative) approach. It is by now well known that plug-in prediction intervals have coverage properties that differ from the nominal coverage properties and are often overly optimistic, having actual coverage probability smaller than the desired coverage probability. The main approaches to correct this drawback of plug-in prediction intervals are the Bayesian and bootstrap approaches. Both approaches have been explored for the case of inference about the quantity of interest at single locations, but similar approaches for the case of inference about spatial averages do not seem to have been explored, with the exception of the paper by Gelfand, Zhu and Carlin (2001) who proposed a Bayesian approach. This work studies bootstrap calibration approaches.

A general idea for the construction of improved prediction intervals is to calibrate plug-in prediction intervals, namely, to adjust plug-in prediction limits in such a way that the resulting prediction interval has coverage probability closer to the desired coverage probability. Two variants of this general idea have been explored that differ on how the adjustment is made. In the first variant the adjusted limit is obtained by modifying the nominal coverage probability, a variant termed as *indirect* by Ueki and Fueda (2007). This variant was initially proposed by Cox (1975), and later studied further by Atwood (1984), Beran (1990), Escobar and Meeker (1999) and Lawless and Fredette (2005). In the second variant additive adjustments are made to plug-in prediction limits, a variant termed as *direct* by Ueki and Fueda (2007). This variant was studied by Barndorff-Nielsen and Cox (1994, 1996), Vidoni (1998) and Ueki and Fueda (2007). For both variants the adjustments can be computed either analytically (Cox, 1975; Atwood, 1984; Barndorff-Nielsen and Cox, 1996; Vidoni, 1998) or by simulation (Beran, 1990; Escobar and Meeker, 1999; Lawless and Fredette, 2005; Ueki and Fueda, 2007). Analytical adjustments are often complex and difficult to obtain, while simulation-based adjustments (also called bootstrap calibration) are usually more practically feasible. The simulation-based indirect calibration variant has been studied and applied for the construction of prediction intervals in random fields at single locations by Sjöstedt-de Luna and Young (2003) and De Oliveira and Rui (2009), but bootstrap calibration does not seem to have been studied for the construction of prediction intervals for spatial averages of random fields.

In this work we study the application of both indirect and direct bootstrap calibration strategies to the construction of prediction intervals for spatial averages of Gaussian random fields over bounded regions. We extend the indirect bootstrap calibration algorithm proposed by Sjöstedt-de Luna and Young (2003) for the construction of prediction intervals for the random field at *single* locations to the construction of prediction intervals for spatial averages over bounded regions. Also, we extend the direct bootstrap calibration algorithm

proposed by Ueki and Fueda (2007) for i.i.d. data to the construction of prediction intervals for spatial averages, which relies on a 'predictive distribution' that only depends on the covariance parameters. A simulation study is carried out to illustrate the effectiveness of both types of calibrated prediction intervals at reducing the coverage probability error of plug-in prediction intervals. Finally, the proposed methodology is applied to the construction of prediction intervals for spatial averages of chromium traces in a potentially contaminated region in Switzerland.

2 Model and Problem Formulation

Consider the random field $\{Z(\mathbf{s}) : \mathbf{s} \in D\}$ representing the spatial variation of a quantity of interest, thought to vary continuously over the region of interest $D \subset \mathbb{R}^2$. It is assumed that D is compact and $|D| > 0$, where $|D|$ denotes the area of D (or more precisely its Lebesgue measure), and $Z(\cdot)$ is an L^2 random field, i.e., $E\{Z^2(\mathbf{s})\} < \infty$ for all $\mathbf{s} \in D$. The mean and covariance functions of the random field are assumed to be given by

$$E\{Z(\mathbf{s})\} = \sum_{j=1}^{p} \beta_j f_j(\mathbf{s}) =: \mu(\mathbf{s}) \quad \text{and} \quad \operatorname{cov}\{Z(\mathbf{s}), Z(\mathbf{u})\} = \sigma^2 K_\phi(\mathbf{s}, \mathbf{u}), \quad (1)$$

where $\boldsymbol{f}(\mathbf{s}) = (f_1(\mathbf{s}), \ldots, f_p(\mathbf{s}))'$ are known location-dependent covariates, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)' \in \mathbb{R}^p$ are unknown regression parameters, $\sigma^2 = \operatorname{var}\{Z(\mathbf{s})\} > 0$ is unknown, $K_\phi(\mathbf{s}, \mathbf{u})$ is a correlation function in $\mathbb{R}^2$ that is continuous on $D \times D$, and φ is an unknown correlation parameter.

The data consist of possibly noisy measurements of the random field at distinct sampling locations $\mathbf{s}_1, \ldots, \mathbf{s}_n \in D$, say $\mathbf{Z}_{\text{obs}} = (Z_{1,\text{obs}}, \ldots, Z_{n,\text{obs}})'$, where

$$Z_{i,\text{obs}} = Z(\mathbf{s}_i) + \varepsilon_i, \quad i = 1, \ldots, n; \quad (2)$$

here $\{\varepsilon_i\}_{i=1}^n \overset{\text{iid}}{\sim} \mathrm{N}(0, \tau^2)$ represent measurement errors independently distributed of the random field $Z(\cdot)$ and $\tau^2 \geq 0$ is the so-called *nugget effect*. The model parameters are then the regression parameters $\boldsymbol{\beta} \in \mathbb{R}^p$ and covariance parameters $\boldsymbol{\theta} = (\sigma^2, \varphi, \tau^2) \in \Theta \subset \mathbb{R}^q$.

The goal is to make inference about a spatial (weighted) average of the random field over a subregion of D of positive area, say $B \subseteq D$, also know as a block average in the geostatistical literature. This spatial average is the random variable given by the (stochastic) integral

$$Z_B = \frac{1}{|B|} \int_B w(\mathbf{s}) Z(\mathbf{s}) d\mathbf{s}, \quad (3)$$

where $w(\cdot)$ is known, nonnegative and piecewise continuous on D, and the integral is defined in mean square sense; its definition and some of its properties are given in the next section. For $\alpha \in (0, 1)$ we are interested in constructing $100(1 - \alpha)\%$ prediction intervals for Z_B, that is, we seek random intervals $(L(\mathbf{Z}_{\text{obs}}), U(\mathbf{Z}_{\text{obs}}))$ for which

$$P_{\beta,\theta}\{L(\boldsymbol{Z}_{\text{obs}}) \leq Z_B \leq U(\boldsymbol{Z}_{\text{obs}})\}=1-\alpha, \quad \text{for all } \boldsymbol{\beta} \in \mathbb{R}^p,\ \boldsymbol{\theta} \in \Theta,$$

where $P_{\beta,\theta}\{\cdot\}$ refers to the joint distribution of $(\boldsymbol{Z}'_{\text{obs}}, Z_B)$ when the true parameter vector is $(\boldsymbol{\beta}', \boldsymbol{\theta})$.

3 L^2 Integration

There is a large body of literature on L^2 integration for random fields in the real line (e.g., Cramér and Leadbetter, 1967; Tanaka, 1996), but a formal treatment of this for random fields in the plane is usually glossed over in the geostatistical literature. For completeness, we review in this section the definition of Z_B as well as some results that are needed in latter sections. The results hold in the Hilbert space $L^2 = L^2(\{Z(\mathbf{s}) : \mathbf{s} \in D\}, d\mathbf{s})$, with $d\mathbf{s}$ being Lebesgue measure, and convergence is considered in mean square sense. The proofs are left for the Appendix.

The definition of Z_B parallels that of Riemann integrals of deterministic bounded functions over compact subsets of $\mathbb{R}^2$ (Bartle, 1976), in which convergence of sequences of real numbers is replaced by L^2 convergence of sequences of random variables. Let $\Delta = \{J_1, \ldots, J_r\}$ denote a finite partition of B, i.e., $B = \cup_k J_k$ and $J_k \cap J_{k'} = \emptyset$ if $k \neq k'$, and $[\Delta] = \max\{|J_1|, \ldots, |J_r|\}$ denotes the 'mesh size' of Δ. A *Riemann sum* for the random field $w(\cdot)Z(\cdot)$ corresponding to the partition Δ is the random variable defined as

$$S_B(\Delta)=\sum_k w(\mathbf{t}_k)Z(\mathbf{t}_k)|J_k|,$$

where $\mathbf{t}_k$ is any point in J_k.

Definition 3.1

The random field $w(\cdot)Z(\cdot)$ is said to be L^2–integrable on B if, as $m \to \infty$, $S_B(\Delta_m)$ converges in L^2 for every sequence $(\Delta_m)_{m\geq 1}$ of finite partitions of B with the property that $\lim_{m\to\infty}[\Delta_m] = 0$. In this case the limit is also in L^2 and is denoted by the right hand side of (3).

Proposition 3.1

Suppose that B is compact, $w(\mathbf{s})$ and $\mu(\mathbf{s})$ are piecewise continuous on B and $K_\phi(\mathbf{s}, \mathbf{u})$ is continuous on $B \times B$. Then Z_B exist, i.e., $w(\cdot)Z(\cdot)$ is L^2–integrable on B.

Proposition 3.2

Let $\{Z(\mathbf{s}) : \mathbf{s} \in D\}$ be a random field with mean and covariance functions given in (1), with $w(\mathbf{s})$ and $\mu(\mathbf{s})$ piecewise continuous on B and $K_\phi(\mathbf{s}, \mathbf{u})$ continuous on $B \times B$. Also, let $\mathbf{s}_0 \in D$ and $B, C \subseteq D$ both compact. Then

$$E\{Z_B\}=\frac{1}{|B|}\int_B w(\mathbf{s})\mu(\mathbf{s})d\mathbf{s} \quad (4)$$

$$\operatorname{cov}\{Z(\mathbf{s}_0),Z_B\}=\frac{\sigma^2}{|B|}\int_B w(\mathbf{s})K_\phi(\mathbf{s}_0,\mathbf{s})d\mathbf{s} \quad (5)$$

$$\operatorname{cov}\{Z_B,Z_C\}=\frac{\sigma^2}{|B|\,|C|}\int\int_{B\times C} w(\mathbf{s})w(\mathbf{u})K_\phi(\mathbf{s},\mathbf{u})d\mathbf{s}d\mathbf{u}. \quad (6)$$

Proposition 3.3

Let $\{Z(\mathbf{s}) : \mathbf{s} \in D\}$ be a Gaussian random field with mean and covariance functions given in (1), with $w(\mathbf{s})$ and $\mu(\mathbf{s})$ piecewise continuous on B and $K_\phi(\mathbf{s}, \mathbf{u})$ continuous on $B \times B$. Also, let $\mathbf{s}_1, \ldots, \mathbf{s}_n \in D$ be distinct and $B \subset D$ compact. Then

$$\begin{pmatrix} \boldsymbol{Z}_{\text{obs}} \\ Z_B \end{pmatrix} \sim \mathrm{N}_{n+1}\left(\begin{pmatrix} X\boldsymbol{\beta} \\ \mu_B(\boldsymbol{\beta}) \end{pmatrix}, \begin{pmatrix} \sum_\theta & \sigma^2\mathbf{K}_B(\phi) \\ \sigma^2\mathbf{K}_B'(\phi) & \sigma^2 K_{BB}(\phi) \end{pmatrix}\right), \quad (7)$$

where X is the $n \times p$ matrix $(X)_{ij} = f_j(\mathbf{s}_i)$, $\mu_B(\boldsymbol{\beta})=\frac{1}{|B|}\int_B w(\mathbf{s})\mu(\mathbf{s})d\mathbf{s}=\sum_{j=1}^p \beta_j f_j(B)$, with $f_j(B)=\frac{1}{|B|}\int_B w(\mathbf{s})f_j(\mathbf{s})d\mathbf{s}$, Σ_θ is the $n \times n$ matrix $(\Sigma_\theta)_{ij} = \sigma^2 K_\phi(\mathbf{s}_i,\mathbf{s}_j) + \tau^2\mathbf{1}\{\mathbf{s}_i = \mathbf{s}_j\}$, with $\mathbf{1}\{A\}$ the indicator function of A, $K_{BB}(\phi)=\frac{1}{|B|^2}\int\int_{B\times B} w(\mathbf{s})w(\mathbf{u})K_\phi(\mathbf{s},\mathbf{u})d\mathbf{s}d\mathbf{u}$, and $\mathbf{K}_B(\phi)=\frac{1}{|B|}\left(\int_B w(\mathbf{u})K_\phi(\mathbf{s}_1,\mathbf{u})d\mathbf{u},\ldots,\int_B w(\mathbf{u})K_\phi(\mathbf{s}_n,\mathbf{u})d\mathbf{u}\right)'$.

4 Plug-in Prediction Intervals

In all that follows it is assumed that X has full rank ($= p < n$) and Σ_θ is positive definite for all $\boldsymbol{\theta} \in \Theta$. The problem of predicting a spatial average based on point-referenced data has been considered extensively in the literature, for instance by Cressie (1993), Chilés and Delfiner (1999), and Gotway and Young (2007). The best linear unbiased predictor (BLUP) of Z_B based on the data $\mathbf{Z}_{\text{obs}}$ (also know as the universal kriging block predictor) and its mean squared prediction error are given, respectively, by

$$\hat{Z}_B(\boldsymbol{\theta})=\boldsymbol{\lambda}_B'(\boldsymbol{\theta})\boldsymbol{Z}_{\text{obs}}$$
$$\hat{\sigma}_B^2(\boldsymbol{\theta})=\sigma^2 K_{BB}(\phi)-2\sigma^2\boldsymbol{\lambda}_B'(\boldsymbol{\theta})\mathbf{K}_B(\phi)+\boldsymbol{\lambda}_B'(\boldsymbol{\theta})\textstyle\sum_\theta\boldsymbol{\lambda}_B(\boldsymbol{\theta}),$$

where

$$\boldsymbol{\lambda}_B'(\boldsymbol{\theta})=\left(\sigma^2\mathbf{K}_B(\phi)+X\left(X'\textstyle\sum_\theta^{-1}X\right)^{-1}\left(\mathbf{f}_B-\sigma^2X'\textstyle\sum_\theta^{-1}\mathbf{K}_B(\phi)\right)\right)'\sum_\theta^{-1},$$

with $\mathbf{f}_B = (f_1(B), \ldots, f_p(B))'$, and $f_j(B)$, X, Σ_θ, $\mathbf{K}_B(\varphi)$ and $K_{BB}(\varphi)$ defined in Proposition 3.3; note from (2) that $\mathrm{cov}\{Z_{i,\mathrm{obs}}, Z_B\} = \mathrm{cov}\{Z(\mathbf{s}_i), Z_B\}$ for all i. The quantities $\hat{Z}_B(\boldsymbol{\theta})$ and $\hat{\sigma}^2_B(\boldsymbol{\theta})$ do not depend on $\boldsymbol{\beta}$, but both depend on the covariance parameters $\boldsymbol{\theta}$. When the random field is Gaussian, it follows from (7) and after some algebra that $\hat{Z}_B(\boldsymbol{\theta}) = E_{\hat{\beta}_\theta,\theta}\{Z_B \mid \mathbf{Z}_{\mathrm{obs}}\}$, where $\hat{\boldsymbol{\beta}}_\theta = (X'\Sigma_\theta^{-1}X)^{-1}X\Sigma_\theta^{-1}\boldsymbol{Z}_{\mathrm{obs}}$ is the best linear unbiased estimator of $\boldsymbol{\beta}$. In addition, when $\boldsymbol{\theta}$ is know, $\hat{Z}_B(\boldsymbol{\theta})$ is the best unbiased predictor (linear or not) with respect to the squared error loss function (Cox, 2004).

If the random field is Gaussian, we have from (7) and after some algebra that

$$\begin{pmatrix} Z_B \\ \hat{Z}_B(\boldsymbol{\theta}) \end{pmatrix} \sim \mathrm{N}_2\left(\begin{pmatrix} \mu_B(\boldsymbol{\beta}) \\ \mu_B(\boldsymbol{\beta}) \end{pmatrix}, \begin{pmatrix} \sigma^2 K_{BB}(\phi) & \sigma^2\boldsymbol{\lambda}'_B(\boldsymbol{\theta})\mathbf{K}_B(\phi) \\ \sigma^2\boldsymbol{\lambda}'_B(\boldsymbol{\theta})\mathbf{K}_B(\phi) & \boldsymbol{\lambda}'_B(\boldsymbol{\theta})\Sigma_\theta\boldsymbol{\lambda}_B(\boldsymbol{\theta}) \end{pmatrix}\right),$$

and hence

$$T = Z_B - \hat{Z}_B(\boldsymbol{\theta}) \sim \mathrm{N}(0, \hat{\sigma}^2_B(\boldsymbol{\theta})).$$

If the covariance parameters $\boldsymbol{\theta}$ were known, then T would be a pivot for the prediction of Z_B and a $100(1-\alpha)\%$ prediction interval for Z_B would be

$$\begin{aligned} I_B(\alpha,\boldsymbol{\theta}) &= \left(\hat{Z}_B(\boldsymbol{\theta}) - \Phi^{-1}(1-\alpha/2)\hat{\sigma}_B(\boldsymbol{\theta}), \hat{Z}_B(\boldsymbol{\theta}) + \Phi^{-1}(1-\alpha/2)\hat{\sigma}_B(\boldsymbol{\theta})\right) \\ &= (L_B(\alpha,\boldsymbol{\theta}), U_B(\alpha,\boldsymbol{\theta})), \text{ say,} \end{aligned} \quad (8)$$

where $\Phi^{-1}(\cdot)$ is the quantile function of the standard normal distribution; for now we do not make explicit the dependence of $L_B(\cdot)$ and $U_B(\cdot)$ on the data. In practice the covariance parameters $\boldsymbol{\theta}$ are not known and have to be estimated from the same data used for prediction, say by $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(\mathbf{Z}_{\mathrm{obs}})$. The simplest and most common practical solution is to use the so-called plug-in (or estimative) prediction interval obtained by replacing $\boldsymbol{\theta}$ with $\hat{\boldsymbol{\theta}}$ in (8), i.e., to use $I_B(\alpha, \hat{\boldsymbol{\theta}})$ as the prediction interval. This solution has a potentially serious drawback. Plug-in prediction intervals have coverage probability that differ from the nominal coverage probability (that is the one that holds when the true parameter values are used), because they do not take into account the sampling variability of parameter estimators. As a result, their actual coverage probability tends to be smaller than the nominal coverage probability, and the coverage probability error may range from negligible to substantial, depending on the observed data and true parameter values.

Let $1-\alpha$ be the desired coverage probability (fixed throughout) and $\boldsymbol{\theta}$ the true covariance parameters. The *coverage probability function* of the plug-in prediction interval obtained from (8) is defined as

$$\pi_B(\alpha,\boldsymbol{\theta})=P_\theta\{Z_B \in I_B(\alpha,\hat{\boldsymbol{\theta}})\}.$$

It can be shown that this coverage probability does not depend on the regression parameters $\boldsymbol{\beta}$, as long as $\hat{\boldsymbol{\theta}}$ is a translation-invariant estimator (i.e., $\hat{\boldsymbol{\theta}}(\hat{\mathbf{Z}}_{\text{obs}} + \mathbf{a}) = \hat{\boldsymbol{\theta}}(\hat{\mathbf{Z}}_{\text{obs}})$ for every $\mathbf{a} \in \mathbb{R}^n$); maximum likelihood and restricted maximum likelihood estimators are translation-invariant (Sjöstedt-de Luna and Young, 2003; De Oliveira and Rui, 2009). If $\boldsymbol{\theta}$ could be estimated without error (i.e., $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}$ with probability 1), then $I_B(\alpha, \hat{\boldsymbol{\theta}})$ would be a $100(1 - \alpha)$ % prediction interval for Z_B since $\pi_B(\alpha, \boldsymbol{\theta})$ would be identical to $1 - \alpha$. But due to estimation error (parameter uncertainty), $\pi_B(\alpha, \boldsymbol{\theta})$ in general depends on both α and $\boldsymbol{\theta}$, and $I_B(\alpha, \hat{\boldsymbol{\theta}})$ has only *nominal* coverage probability $1 - \alpha$.

5 Calibrated Prediction Intervals

The main approaches that have been considered to account for parameter uncertainty when performing predictive inference about $Z(\mathbf{s}_0)$ include the Bayesian approach (Handcock and Stein, 1993; De Oliveira, Kedem and Short, 1997) and bootstrap approaches (Sjöstedt-de Luna and Young, 2003; Wang and Wall, 2003; De Oliveira and Rui, 2009; Schelin and Sjöstedt-de Luna, 2010).

The goal of calibrating prediction intervals is to adjust the prediction bounds $L_B(\alpha, \hat{\boldsymbol{\theta}})$ and $U_B(\alpha, \hat{\boldsymbol{\theta}})$ to, respectively, $L_B^a(\alpha, \hat{\boldsymbol{\theta}})$ and $U_B^a(\alpha, \hat{\boldsymbol{\theta}})$ say, in such a way that the coverage probability of the adjusted prediction interval is *closer* to $1 - \alpha$. Two strategies have been previously explored for calibrating prediction intervals which we describe below for the construction of prediction intervals for Z_B.

5.1 Indirect Calibration

The first calibration strategy, called *indirect* by Ueki and Fueda (2007), consists of achieving the goal by adjusting α in $L_B(\alpha, \hat{\boldsymbol{\theta}})$ and $U_B(\alpha, \hat{\boldsymbol{\theta}})$, i.e., finding $\alpha_c \in (0, 1)$ for which $\pi_B(\alpha_c, \hat{\boldsymbol{\theta}}) \approx 1-\alpha$, and then use $I_B(\alpha_c, \hat{\boldsymbol{\theta}})$. This strategy was initially proposed and developed by Cox (1975), Altwood (1984) and Beran (1990), and was applied to the calibration of prediction intervals for $Z(\mathbf{s}_0)$ in Gaussian and log-Gaussian random field models by, respectively, Sjöstedt-de Luna and Young (2003) and De Oliveira and Rui (2009).

This strategy requires estimation of the function $\pi_B(x, \hat{\boldsymbol{\theta}})$ for $x \in (0, 1)$, which is carried out by simulation using Rao-Blackwellization. For that we write the coverage probability as

$$\pi_B(\alpha,\boldsymbol{\theta})=E_\theta\{P_\theta\{Z_B \in I_B(\alpha,\hat{\boldsymbol{\theta}})|\boldsymbol{Z}_{\text{obs}}\}\}$$
$$=E_\theta\left\{\Phi\left(\frac{U_B(\alpha,\hat{\boldsymbol{\theta}},\boldsymbol{Z}_{\text{obs}})-\eta_B(\mathbf{0},\boldsymbol{\theta},\boldsymbol{Z}_{\text{obs}})}{\tau_B(\boldsymbol{\theta})}\right)-\Phi\left(\frac{L_B(\alpha,\hat{\boldsymbol{\theta}},\boldsymbol{Z}_{\text{obs}})-\eta_B(\mathbf{0},\boldsymbol{\theta},\boldsymbol{Z}_{\text{obs}})}{\tau_B(\boldsymbol{\theta})}\right)\right\},$$

where the probability is computed with respect to the conditional distribution of Z_B given $\mathbf{Z}_{\text{obs}}$ and the expectation with respect to the distribution of $\mathbf{Z}_{\text{obs}}$, both when $\boldsymbol{\beta} = \mathbf{0}$ and $\boldsymbol{\theta}$ is the true value of the covariance parameters, and from (7) it follows that

$$\eta_B(\boldsymbol{\beta},\boldsymbol{\theta},\boldsymbol{Z}_{\text{obs}}) = E_{\beta,\theta}\{Z_B|\boldsymbol{Z}_{\text{obs}}\} = \mu_B(\boldsymbol{\beta}) + \sigma^2\mathbf{K}'_B(\phi)\textstyle\sum_\theta^{-1}(\boldsymbol{Z}_{\text{obs}} - X\boldsymbol{\beta})$$
$$\tau_B^2(\boldsymbol{\theta}) = \text{var}_\theta\{Z_B|\boldsymbol{Z}_{\text{obs}}\} = \sigma^2(K_{BB}(\phi) - \sigma^2\mathbf{K}'_B(\phi)\textstyle\sum_\theta^{-1}\mathbf{K}_B(\phi)).$$

The dependence of $L_B(\cdot)$ and $U_B(\cdot)$ on the data is now made explicit. Hence we have the following algorithm for the estimation of $\pi_B(x, \hat{\boldsymbol{\theta}})$:

Algorithm 1—For $B \subseteq D$ do the following:

Step 1. Compute the ML or REML estimate $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(\mathbf{z}_{\text{obs}})$ and $\tau_B^2(\hat{\theta})$ from the observed data.

Step 2. Simulate M independent and identically distributed samples $\{\boldsymbol{Z}^*_{\text{obs},j} : 1 \le j \le M\}$ from the Gaussian distribution with mean vector $\mathbf{0}$ and covariance matrix $\Sigma_{\hat{\theta}}$

Step 3. For each $j = 1, \ldots, M$ compute the ML or REML estimate $\hat{\boldsymbol{\theta}}^*_j = \hat{\boldsymbol{\theta}}(\mathbf{z}^*_{\text{obs},j})$, and compute the quantities $\hat{Z}_B(\hat{\boldsymbol{\theta}}^*_j)$, $\hat{\sigma}^2_B(\hat{\boldsymbol{\theta}}^*_j)$ and $\hat{\eta}^*_{B,j} = \eta_B(\mathbf{0}, \hat{\boldsymbol{\theta}}, \mathbf{z}^*_{\text{obs},j})$ from the simulated data.

Step 4. For $x \in (0, 1)$ estimate $\pi_B(x, \hat{\boldsymbol{\theta}})$ by

$$\pi^*_B(x,\hat{\boldsymbol{\theta}}) = \frac{1}{M}\sum_{j=1}^{M}\left[\Phi\left(\frac{U^*_{B,x,j} - \hat{\eta}^*_{B,j}}{\hat{\tau}_B}\right) - \Phi\left(\frac{L^*_{B,x,j} - \hat{\eta}^*_{B,j}}{\hat{\tau}_B}\right)\right],$$

where $L^*_{B,x,j} = L_B(x, \hat{\boldsymbol{\theta}}^*_j, \mathbf{z}^*_{\text{obs},j})$, $U^*_{B,x,j} = U_B(x, \hat{\boldsymbol{\theta}}^*_j, \mathbf{z}^*_{\text{obs},j})$, and $\hat{\tau_B} = \tau_B(\hat{\boldsymbol{\theta}})$.

The 100(1 − α)% prediction interval for Z_B obtained by indirect calibration is $I_B(\alpha_c, \hat{\boldsymbol{\theta}})$, where α_c is the root (in x) of the function $\pi^*_B(x,\hat{\boldsymbol{\theta}}) - (1-\alpha)$. For inference about $Z(\mathbf{s}_0)$ in Gaussian random fields, Sjöstedt-de Luna and Young (2003) showed that the existence and uniqueness of such root is guaranteed since the function $\pi_B(x, \hat{\boldsymbol{\theta}})$ is continuous and strictly decreasing on (0, 1), and $\lim_{x\to 0^+} \pi_B(x, \hat{\boldsymbol{\theta}}) = 1$. The proof carries over for inference about Z_B.

5.2 Direct Calibration

The second calibration strategy, called *direct* by Ueki and Fueda (2007), consists of achieving the goal by separately finding additive adjustments to the prediction limits. Specifically, $U_B(\alpha, \hat{\boldsymbol{\theta}})$ is adjusted to $U^a_B(\alpha,\hat{\boldsymbol{\theta}}) = U_B(\alpha,\hat{\boldsymbol{\theta}}) + d_U(1-\alpha/2,\hat{\boldsymbol{\theta}})$, say, where $d_U(\cdot)$ is chosen so that it holds that $P_\theta\{Z_B \le U^a_B(\alpha,\hat{\boldsymbol{\theta}})\} \approx 1-\alpha/2$. The same procedure is done to adjust the lower limit $L_B(\cdot)$, but using $\alpha/2$ in place of 1 − $\alpha/2$ in the computation of the

additive adjustment $d_L(\cdot)$, so that the coverage probability of $(L_B^a(\alpha,\hat{\boldsymbol{\theta}}), U_B^a(\alpha,\hat{\boldsymbol{\theta}}))$ is approximately $1-\alpha$. This strategy was initially proposed by Barndorff-Nielsen and Cox (1996), but the correction term was only implicitly defined. Vidoni (1998) provided an explicit expression for the correction term for models with independent and identically distributed observations, and Giummolè and Vidoni (2010) did so for a class of Gaussian models. These works carried out the computation of $d_L(\cdot)$ and $d_U(\cdot)$ analytically, in which computations are, in general, either complex or even not feasible for some models. Ueki and Fueda (2007) developed, for models with independent and identically distributed observations, an asymptotically equivalent expression for the additive adjustment that is remarkably simple and can be computed by simulation; we adapt their procedure for the model considered here.

The basis for the direct calibration strategy relies on adjusting prediction limits that are *quantiles* of a certain predictive distribution, and this holds for both independent and dependent data. For a class of Gaussian models, Giummolè and Vidoni (2010) used direct calibration based on the predictive distribution

$$F^{(1)}_{\boldsymbol{\beta},\boldsymbol{\theta}}(z|\mathbf{z}_{\text{obs}}) = P_{\boldsymbol{\beta},\boldsymbol{\theta}}\{Z_B \le z | \boldsymbol{Z}_{\text{obs}} = \mathbf{z}_{\text{obs}}\} = \Phi\left(\frac{z-\eta_B(\boldsymbol{\beta},\boldsymbol{\theta},\mathbf{z}_{\text{obs}})}{\tau_B(\boldsymbol{\theta})}\right), \quad (9)$$

and found complex closed-form expressions for the additive adjustments of the plug-in prediction limits

$$(\eta_B(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\theta}},\mathbf{z}_{\text{obs}}) - \Phi^{-1}(1-\alpha/2)\tau_B(\hat{\boldsymbol{\theta}}),\ \eta_B(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\theta}},\mathbf{z}_{\text{obs}}) + \Phi^{-1}(1-\alpha/2)\tau_B(\hat{\boldsymbol{\theta}})).$$

Note that the quantiles of $F^{(1)}_{\boldsymbol{\beta},\boldsymbol{\theta}}(z|\mathbf{z}_{\text{obs}})$ depend on both the regression and covariance parameters. Our goal is to adjust the prediction limits in (8), which have exact coverage when $\boldsymbol{\theta}$ is known. A way to achieve this is to take a (partially) Bayesian approach. If the regression parameters have the (improper) prior $p(\boldsymbol{\beta}) \propto 1$ and the covariance parameters have a point mass prior at the true value, the conditional (on $\boldsymbol{\theta}$) posterior distribution of $\boldsymbol{\beta}$ is $\mathrm{N}(\hat{\boldsymbol{\beta}}_\theta, (X'\sum_\theta^{-1} X)^{-1})$. So instead of $F^{(1)}_{\boldsymbol{\beta},\boldsymbol{\theta}}(z|\mathbf{z}_{\text{obs}})$, we use the Bayesian predictive distribution

$$F^{(2)}_{\boldsymbol{\theta}}(z|\mathbf{z}_{\text{obs}}) = \int_{\mathbb{R}^p} F^{(1)}_{\boldsymbol{\beta},\boldsymbol{\theta}}(z|\mathbf{z}_{\text{obs}}) p(\boldsymbol{\beta}|\boldsymbol{\theta},\mathbf{z}_{\text{obs}}) d\boldsymbol{\beta} = \Phi\left(\frac{z-\hat{Z}_B(\boldsymbol{\theta})}{\hat{\sigma}_B(\boldsymbol{\theta})}\right), \quad (10)$$

(O'Hagan, 1991). In addition, $F^{(2)}_{\boldsymbol{\theta}}(z|\mathbf{z}_{\text{obs}})$ also arises as the bootstrap predictive distribution proposed in Harris (1989), when $\boldsymbol{\theta}$ is assumed known. Some quantiles of $F^{(2)}_{\boldsymbol{\theta}}(z|\mathbf{z}_{\text{obs}})$ agree with the prediction limits in (8), which depend only on the covariance parameters, so the frequentist prediction intervals (8) are also Bayesian prediction intervals. Note that $\hat{Z}_B(\boldsymbol{\theta}) = \eta_B(\boldsymbol{\beta}, \hat{\boldsymbol{\theta}}, \mathbf{z}_{\text{obs}})$, but $\hat{\sigma_B}(\boldsymbol{\theta}) > \tau_B(\boldsymbol{\theta})$, so (9) and (10) have similar locations, but the latter is

always more dispersed. A benefit of using $F_{\theta}^{(2)}(z|\mathbf{z}_{\text{obs}})$ is that it accounts for the uncertainty in the regression parameters, so calibration adjusts only for the uncertainty in the covariance parameters. We compute the additive adjustments by adapting to the current model the method proposed by Ueki and Fueda (2007). In what follows we describe the method to adjust the upper prediction limit $U_B(\cdot)$, since adjusting the lower limit is similar.

Let $a_U(\theta) = P_\theta\{Z_B \le U_B(a, \hat{\theta})\}$ be the coverage probability of the plug-in prediction interval $(-\infty, U_B(a, \hat{\theta}))$. The additive adjustment obtained by applying Ueki and Fueda's method based on the predictive distribution (10) is $d_U(1-a/2, \hat{\theta}) = U_B(a, \hat{\theta}) - U_B(a_U(\hat{\theta}), \hat{\theta})$, so the adjusted prediction limit is

$$U_B^a(\alpha, \hat{\boldsymbol{\theta}}) = 2U_B(\alpha, \hat{\boldsymbol{\theta}}) - U_B(\alpha_U(\hat{\boldsymbol{\theta}}), \hat{\boldsymbol{\theta}}).$$

The quantity $a_U(\hat{\theta})$ is not available in closed-form, but can be easily and precisely estimated by simulation using Rao-Blackwellization. The resulting adjusted prediction limit is simpler to compute than the adjusted prediction limit proposed by Giummolè and Vidoni (2010). Additionally, Fonseca, Giummolè and Vidoni (2012) provided an explicit expression of a predictive distribution having quantiles approximately equal to $U_B^a(\alpha, \hat{\boldsymbol{\theta}})$. Hence we have the following algorithm for the computation of the 100(1 − a)% prediction interval for Z_B using direct calibration:

Algorithm 2—For $B \subseteq D$ do the following:

Step 1. The same as in Algorithm 1.

Step 2. The same as in Algorithm 1.

Step 3. The same as in Algorithm 1.

Step 4. Estimate $a_U(\hat{\theta})$ by

$$\alpha_U^*(\hat{\boldsymbol{\theta}}) = \frac{1}{M}\sum_{j=1}^{M} \Phi\left(\frac{U_{B,\alpha,j}^* - \hat{\eta}_{B,j}^*}{\hat{\tau}_B}\right),$$

where $U_{B,\alpha,j}^*, \hat{\eta}_{B,j}^*$, and $\hat{\tau_B}$ are the same as in Algorithm 1.

Then compute $U_B^{a*}(\alpha, \hat{\boldsymbol{\theta}}) = 2U_B(\alpha, \hat{\boldsymbol{\theta}}) - U_B(\alpha_U^*(\hat{\boldsymbol{\theta}}), \hat{\boldsymbol{\theta}})$.

Step 5. Estimate $a_L(\hat{\theta}) = P_\theta\{Z_B \le L_B(a, \hat{\theta})\}$ by

$$\alpha_L^*(\hat{\boldsymbol{\theta}}) = \frac{1}{M}\sum_{j=1}^{M} \Phi\left(\frac{L_{B,\alpha,j}^* - \hat{\eta}_{B,j}^*}{\hat{\tau}_B}\right),$$

where $L_{B,\alpha,j}^*$ is the same as in Algorithm 1.

Then compute $L_B^{a*}(\alpha,\hat{\boldsymbol{\theta}})=2L_B(\alpha,\hat{\boldsymbol{\theta}})-L_B(\alpha_U^*(\hat{\boldsymbol{\theta}}),\hat{\boldsymbol{\theta}})$.

The 100(1−a)% prediction interval for Z_B obtained by direct calibration is $\left(L_B^{a*}(\alpha,\hat{\boldsymbol{\theta}}),U_B^{a*}(\alpha,\hat{\boldsymbol{\theta}})\right)$.

6 Simulation Study

In this section we carry out a simulation study to investigate and compare the coverage properties of the prediction intervals considered in this work, namely, the plug-in, indirectly calibrated and directly calibrated prediction intervals, for integrals of Gaussian random fields over bounded regions. Although the calibrated prediction intervals are constructed to have better coverage properties than the plug-in prediction interval, it is still important to assess in practice the size of the improvement as well as how this improvement might depend on some features of the random field.

6.1 Design

We begin by simulating realizations of Gaussian random fields at n = 50 sampling locations uniformly distributed over the region D = [0, 2] × [0, 2], as displayed in Figure 1. We used as the mean function

$$\mu(\mathbf{s})=2 \quad \text{or} \quad \mu(\mathbf{s})=2+3x+4y, \quad \mathbf{s}=(x,y),$$

and the isotropic exponential covariance function

$$\sigma^2 K_\phi(\mathbf{s},\mathbf{u})=\sigma^2\exp\left(-\|\mathbf{s}-\mathbf{u}\|/\phi\right), \quad (11)$$

where σ^2 = 0.5 or 2, and φ = 0.2 or 0.8; ||·|| denotes Euclidean norm. The simulated data are obtained from (2), with τ^2 = 0 (no nugget effect) or σ^2/4 (size of nugget is 25% of signal's variance). This setup provides a wide range of data behaviors in terms of trend, variability and strength of spatial association. The quantities of interest are the spatial averages (3), with $w(\mathbf{s}) \equiv 1$ and the following integration regions of different sizes: B_1 = [0.2, 1.8] ×[0.2, 1.8] (large block), B_2 = [0.8, 1.2] × [0.8, 1.2] (medium block), and B_3 = [0.975, 1.025] × [0.975, 1.025] (small block); see Figure 1. This is done to assess whether the coverage probabilities of the prediction intervals may depend on the size of the integration region.

For each of the sixteen possible models (2 means × 2 σ^2-values × 2 φ-values × 2 τ^2-values) we simulated 1000 times the data and the three spatial averages *jointly*. This is done using the fact that ($\boldsymbol{Z}'_{\text{obs}}$, Z_{B_1}, Z_{B_2}, Z_{B_3}) has multivariate normal distribution with means, variances and covariances given in (1) and Proposition 3.2, which follows from a slight generalization of Proposition 3.3. Based on each of the simulated datasets, the three types of 95% prediction intervals (plug-in, indirectly and directly calibrated) were computed for Z_{B_1}, Z_{B_2} and Z_{B_3}. Calibrated prediction intervals were obtained using Algorithms 1 and 2 described in Section 5, with M = 500 as the bootstrap sample size, and the *same* bootstrap samples were used in both calibration algorithms. For all prediction intervals the parameters

were estimated by restricted maximum likelihood. The coverage probability of any of the aforementioned types of prediction intervals for Z_B was estimated from the simulated data by

$$\frac{1}{1000}\sum_{k=1}^{1000}\mathbf{1}\{L_{B,k} \leq Z_{B,k} \leq U_{B,k}\},$$

where B is either B_1, B_2 or B_3 defined above, $(L_{B,k}, U_{B,k})$ is the prediction interval of the same type for Z_B computed from the k^{th} simulated dataset, and $Z_{B,k}$ is the spatial average that was jointly simulated with the k^{th} dataset. The standard error of any of these coverage estimators is approximately $\sqrt{0.95 \cdot 0.05/1000} = 0.0069$.

6.2 Computation

All the computations described in this section were carried out in the R language. The joint generation of the simulated datasets and block averages was done with the function mvrnorm from the package MASS. The generation of bootstrap samples in Step 2 of Algorithms 1 and 2 was done using the function grf, and the parameter estimation in Steps 1 and 3 was done using the function likfit, both from the package geoR (Ribeiro and Diggle, 2001).

The tasks of generating the simulated block averages and carrying out Step 3 of Algorithms 1 and 2 require the computation of $\mu_{B_j}(\boldsymbol{\beta})$ given in (4), which in the present context can be done exactly. When the mean function is constant we have that $\mu_{B_j}(\beta_1) = \beta_1$, and when the mean function is not constant we have, for $B_j = [a, b] \times [c, d]$ say, that

$$\begin{aligned}\mu_{B_j}(\boldsymbol{\beta}) &= \frac{1}{|B_j|}\int_c^d\int_a^b(\beta_1+\beta_2 x+\beta_3 y)\,dxdy \\ &= \beta_1+\beta_2\frac{a+b}{2}+\beta_3\frac{c+d}{2}.\end{aligned}$$

These tasks also require, for different covariance parameters, the repeated computation of $\text{var}\{Z_{B_j}\}$, $\text{cov}\{Z(\mathbf{s}_i), Z_{B_j}\}$ and $\text{cov}\{Z_{B_j}, Z_{B_l}\}$ given by (5) and (6), but these are rarely available in closed form. These integrals could be approximated by Monte Carlo methods, but we found that numerical cubature algorithms were faster and more stable. Specifically, we used the function adaptIntegrate from the package cubature to approximate the integrals (5) and (6), which uses adaptive multivariate integration.

The above functions use algorithms that require the integration region to be a rectangle with sides parallel to the axes, which *is* the case in the present study. But in some contexts the integration region may not be a rectangle, for instance when interest centers about a spatial average over an administrative region such as a county or state. In such cases the above functions can still be used by modifying the weight function and integration region, noting that

$$\frac{1}{|B|}\int_B w(\mathbf{s})Z(\mathbf{s})d\mathbf{s}=\frac{|C|}{|B|}\cdot\frac{1}{|C|}\int_C \tilde{w}(\mathbf{s})Z(\mathbf{s})d\mathbf{s},$$

with $w(\tilde{\mathbf{s}}) = \mathbf{1}\{\mathbf{s} \in B\}w(\mathbf{s})$ and C a rectangle with sides parallel to the axes that contains B. In this case, the area $|B|$ is estimated using the hit-or-miss Monte Carlo algorithm, and $\mu_C(\boldsymbol{\beta})$, $\mathbf{K}_C(\varphi)$ and $K_{CC}(\varphi)$ are estimated as before using the weight function $w(\tilde{\mathbf{s}})$.

6.3 Results

Table 1 displays estimated coverage probabilities of the 95% plug-in, indirectly calibrated and directly calibrated prediction intervals for Z_{B_1} (large block average) under all considered models, with the top part of the table displaying the coverage for models with constant mean and the bottom part displaying the coverage for models with non-constant mean. As a check the table also displays the coverage probabilities of the prediction intervals obtained using the true values of the covariance parameters. As expected, the coverage probabilities of the plug-in prediction intervals are always smaller than nominal, with an average coverage error over all considered models of about 1.75%. On the other hand, the coverage probabilities of both calibrated prediction intervals are closer to nominal for all considered models: the average coverage errors of indirectly calibrated and directly calibrated prediction intervals are, respectively, of about 0.27% and 0.48%. The behavior of the coverage probability across all the considered models seems to be about the same for all types of prediction intervals.

Tables 2 and 3 display estimated coverage probabilities of the 95% plug-in, indirectly calibrated and directly calibrated prediction intervals for, respectively, Z_{B_2} and Z_{B_3}, under all considered models. Again, the coverage probabilities of the plug-in prediction intervals are always smaller than nominal, while the coverage probabilities of both calibrated prediction intervals are closer to nominal. For Z_{B_2}, the average coverage errors over all considered models of the plug-in, indirectly calibrated and directly calibrated are, respectively, 2.65%, 0.23% and 0.53%, and for Z_{B_3} the average coverage errors are, respectively, 2.50%, 0.22% and 0.33%. It is also observed that for the medium and small blocks, the coverage errors of the plug-in prediction intervals tend to be larger in models with measurement error ($\tau^2 > 0$), while the coverage errors of both calibrated prediction intervals tend to be about the same for all considered models. Hence, calibration appears to be most useful when the data contain measurement error.

When comparing both calibration strategies, it appears that indirect calibration has a very small edge over direct calibration in the sense of having a slightly smaller coverage error, while both calibration strategies are computationally about equally intensive. Finally, the effectiveness of both calibration strategies do not appear to depend on the size of the block, $|B|$.

Finally, we also carried out simulations using $n = 200$ sampling locations, formed by the 50 locations displayed Figure 1 plus 150 additional locations uniformly distributed over the region. The rest of the simulation design is the same as before, except that we only

considered the constant mean function. Table 4 displays estimated coverage probabilities of the 95% plug-in, indirectly calibrated and directly calibrated prediction intervals for Z_{B_1} (top table), Z_{B_2} (middle table) and Z_{B_3} (bottom table). The main findings are the same as those reported before from the simulation with $n = 50$. In addition, the estimated coverage probabilities of all prediction intervals are about the same as those of their respective counterparts for $n = 50$. These findings are discussed further in the next section.

7 Large Sample Behavior

When the data are i.i.d. and independent of the predictand, Barndorff-Nielsen and Cox (1994) showed that, provided the parameter estimators are $\sqrt{n}$–consistent, plug-in prediction intervals are first-order accurate, meaning that the coverage probability error satisfies $\pi(\alpha, \theta) - (1-\alpha) = O(n^{-1})$, as $n \to \infty$. For inference about $Z(\mathbf{s}_0)$ in Gaussian random fields, Sjöstedt-de Luna and Young (2003, Lemma 3) showed that, under some conditions, plug-in prediction intervals are first-order accurate. The asymptotic behavior of the coverage probability error of bootstrap calibrated prediction intervals were also studied in the simpler situations described above, where it was shown that, under some conditions, these prediction intervals are third-order accurate, i.e., their coverage probability error is $O(n^{-3/2})$ as $n \infty$. (Barndorff-Nielsen and Cox, 1996; Sjöstedt-de Luna and Young, 2003; Ueki and Fueda, 2007). It is conceivable that similar asymptotic behaviors may also hold for the coverage probability error of plug-in and bootstrap calibrated prediction intervals for Z_B, but no such investigation is attempted here. Also, Sjöstedt-de Luna and Young (2003, Lemma 4) showed that the asymptotic coverage probability error of bootstrap calibrated prediction intervals for $Z(\mathbf{s}_0)$ are smaller than those of plug-in prediction intervals. The simulation study in Section 6 suggests that the same holds for calibrated prediction intervals for Z_B.

For the spatial data considered here there are two types of large sample regimes, increasing domain asymptotics and infill asymptotics. Under the former, the region D grows in size with increasing sample size and the sampling design is such that the distance between nearest sampling locations is bounded away from zero. Under the latter, the region D remains fixed and the sampling locations become increasingly dense with increasing sample size. The properties of parameter estimators under these two asymptotic regimes are quite different. For increasing domain asymptotics, Mardia and Marshall (1984) showed that, under fairly general conditions holding for several types of covariance functions, maximum likelihood estimators are consistent and asymptotically normal. Similar results were shown for restricted maximum likelihood estimators by Cressie and Lahiri (1996). On the other hand, for infill asymptotics Zhang (2004) showed that, for Gaussian random fields with known mean and *isotropic* Matérn covariance function (which includes the isotropic exponential covariance (11) as a particular case), the parameters σ^2 and φ can *not* be consistently estimated.

As indicated above, Sjöstedt-de Luna and Young (2003) showed that, under some conditions, plug-in prediction intervals for $Z(\mathbf{s}_0)$ are first-order accurate and indirectly calibrated bootstrap prediction intervals are third-order accurate. Although the required conditions may hold under both asymptotic regimes, depending on the type of covariance function, they are in general difficult to check. These conditions were checked under both

increasing domain and infill asymptotic regimes for the *separable* exponential covariance function in $\mathbb{R}^2$, where in both regimes the covariance parameters can be estimated consistently. But they were not checked for the *isotropic* exponential covariance function in $\mathbb{R}^2$ (Sjöstedt-de Luna and Young, 2003, Section 7). In addition, the results reported in the last example of Sjöstedt-de Luna and Young (2003) dealing with infill asymptotics and the isotropic exponential covariance function in $\mathbb{R}^2$ show that empirical coverages of plug-in and indirectly calibrated prediction intervals for $Z(\mathbf{s}_0)$ do not seem to vary monotonically with sample size.

The simulation study carried out in Section 6 uses the isotropic exponential covariance function and a sampling design that mimics infill asymptotics, a situation in which covariance parameters can not be consistently estimated. The results in there show that both strategies for the construction of calibrated prediction intervals for Z_B are effective at reducing the coverage probability error, when compared to plug-in prediction intervals. Nevertheless, the estimated coverage probability errors for both sample sizes (n = 50 and 200) were similar, which begs the question suggested by a referee: Does it hold that $\pi_B(\alpha, \boldsymbol{\theta}) \to 1 - \alpha$, as $n \to \infty$? Although we do not have a definite answer, we conjecture that the apparent lack of consistency of prediction intervals for Z_B in the situation considered in Section 6 is due to the lack of consistent estimators for the covariance parameters.

To provide some support for the above conjecture, we run a small simulation experiment with the same model as before, but now using a design that mimics increasing domain asymptotics, for which maximum likelihood estimates of the covariance parameters are consistent. We consider Gaussian random field models with constant mean 2 and isotropic exponential covariance function (11) with $\sigma^2 = 0.5$ and $\varphi = 0.2$; we consider both situations, data with no measurement error ($\tau^2 = 0$) and data with measurement error in which $\tau^2 = 0.5/4$. We simulated data with three nested designs with sample sizes n = 25, 50 and 200, that are sampled from the nested and increasing sampling windows D = [8, 10] × [8, 10], [6, 12] × [6, 12] and [0, 18] × [0, 18], respectively. For each sampling design we simulated 1000 times the data and Z_B jointly, where B = [8.8, 9.2] × [8.8, 9.2], [8.4, 9.6] × [8.4, 9.6], and [7.2, 10.8] × [7.2, 10.8] when n = 25, 50 and 200, respectively. The block size increases with the sampling window size in a way that makes $|D|/|B|$ constant. From each of the simulated datasets, the three types of 95% prediction intervals for Z_B were computed using the same algorithms as in Section 6. Table 5 displays the estimated coverage probabilities of the three 95% prediction intervals for Z_B. It is apparent that the coverage probabilities of the plug-in prediction intervals tend to become closer to the nominal coverage as the sample size increases. This trend is less apparent for the bootstrap calibrated prediction intervals, because they seem to have coverage probabilities close to nominal even for the moderate size datasets. This suggests that bootstrap calibrated prediction intervals might be more effective in an increasing domain asymptotics scenario, although a complete study is required to support such claim.

8 Example

In 1992 a team from the Swiss Federal Institute of Technology at Lausanne collected data on contamination levels by heavy metals in a region of about 15 km^2 in the Swiss canton of

Jura. The data were measurements of traces in top soil of cadmium, chromium, cobalt, copper, lead, nickel and zinc at 359 locations scattered throughout the region; see Figure 2. The sampling protocol and an initial analysis are described in Atteia, Dubois and Webster (1994), and the datasets and geostatistical analyses appear in Goovaerts (1997)[2].

We analyze here the chromium (Cr) dataset, measured in parts per million (ppm). Exploratory data analysis suggests that the mean of the chromium traces is constant, its histogram is close to symmetric, and the data pass the Shapiro-Wilk test of normality at 5% level. Figure 3 displays the histogram of the data (left), as well as the empirical semivariogram and fitted isotropic exponential semivariogram function (right). Note that the empirical semivariogram displays an apparent discontinuity at the origin, which we interpret as measurement error. We then assume that the spatial variation of chromium traces over the region is reasonably described by a Gaussian random field with constant mean and isotropic exponential covariance function[3]. The restricted maximum likelihood estimates of the model parameter are

$$\hat{\beta}=35.38, \quad \hat{\sigma}^2=91.72, \quad \hat{\phi}=0.18, \quad \hat{\tau}^2=18.84.$$

The Jura canton can be divided in four subregions according to land use type: tillage, meadow, pasture and forest. As an illustration, we compute prediction intervals for spatial averages of chromium traces over four rectangular blocks, each of which is entirely contained into a land use type; the rectangular blocks are displayed in Figure 2 and their coordinates and respective land use type are given in Table 6.

For each rectangular block B in Table 6 we computed the 95% plug-in, indirectly calibrated and directly calibrated prediction intervals for Z_B, where for both calibration strategies we used 3000 as the bootstrap sample size. The prediction intervals are displayed in Table 7. For each rectangular block the three types of 95% prediction intervals are similar, but in all cases the plug-in prediction intervals are narrower than their corresponding calibrated prediction intervals. The increase in width of the calibrated prediction intervals ranges from about 1.8 to 3.4% of the width of the respective plug-in prediction intervals, and the estimated coverage probability of the plug-in prediction intervals range from about 94.2 to 94.5%. The benefit of calibration in these examples is modest, presumably due to the moderate sample size.

9 Conclusions

As it occurs for interval prediction of random fields at single locations, plug-in prediction intervals for spatial averages of Gaussian random fields are overly optimistic in the sense of having actual coverage probability smaller than the desired coverage probability. Depending on the data and true model, the coverage probability error may range from mild to substantial. This work has proposed two bootstrap calibration strategies to construct better

[2]Many previous analyses have divided these data into two subsets: a prediction set formed by measurements at 259 locations, and a validation set formed by measurements at the remaining 100 locations. The analysis here uses data from *all* locations.
[3]The measured traces of the other heavy metals display notably asymmetric distributions.

prediction intervals that reduce the coverage probability error, and has illustrated their effectiveness using a simulation study and a real dataset example. A problem only touched upon here, that requires further investigation, is the determination of the large sample behavior of these plug-in and bootstrap calibrated prediction intervals.

A natural follow up question is how to construct better prediction intervals for spatial averages in some types of non-Gaussian random fields. The application of the calibration methodology proposed here for non-Gaussian random fields appears challenging since in most cases the conditional distribution of Z_B given $\mathbf{Z}_{\text{obs}}$ is not known; this is the case, for instance, in log-Gaussian random fields. Rather than using calibration, Schelin and Sjöstedt-de Luna (2010) explored a semi-parametric approach to construct hybrid bootstrap prediction intervals for $Z(\mathbf{s}_0)$ that does not require distributional assumptions. We are currently exploring the implementation of this approach for the construction of prediction intervals for Z_B, which will be reported elsewhere.

Acknowledgments

We thank an Associate editor and two anonymous referees for helpful suggestions that lead to an improved paper. We also thank computational support from Computational System Biology Core at UTSA, funded by the National Institute on Minority Health and Health Disparities (G12MD007591) from the National Institutes of Health.

References

Atteia O, Dubois J-P, Webster R. Geostatistical Analysis of Soil Contamination in the Swiss Jura. Environmental Pollution. 1994; 86:315–327. [PubMed: 15091623]

Atwood CL. Approximate Tolerance Intervals, Based on Maximum Likelihood Estimates. Journal of the American Statistical Association. 1984; 79:459–465.

Barndorff-Nielsen OE, Cox DR. Prediction and Asymptotics. Bernoulli. 1996; 2:319–340.

Barndorff-Nielsen, OE.; Cox, DR. Inference and Asymptotics. Chapman & Hall; 1994.

Bartle, RG. The Elements of Real Analysis. 2. Wiley; 1976.

Beran R. Calibrating Prediction Regions. Journal of the American Statistical Association. 1990; 85:715–723.

Chilès, J-P.; Delfiner, P. Geostatistics: Modeling Spatial Uncertainty. Wiley; 1999.

Cox, DD. Best Unbiased Prediction for Gaussian and Log-Gaussian Processes. In: Rojo, J.; Pèrez-Abreu, V., editors. The First Erich L. Lehmann Symposium—Optimality. Vol. 44. 2004. p. 125-132.Institute of Mathematical Statistics Monograph Series

Cox, DR. Prediction Intervals and Empirical Bayes Confidence Intervals. In: Gani, J., editor. Perspectives in Probability and Statistics. Academic Press; 1975. p. 47-55.

Cramér, H.; Leadbetter, MR. Stationary and Related Stochastic Processes. Wiley; 1967.

Cressie N. Block Kriging for Lognormal Spatial Processes. Mathematical Geology. 2006; 38:413–443.

Cressie, NAC. Statistics for Spatial Data. Wiley; 1993. (rev. ed.)

Cressie N. The Origins of Kriging. Mathematical Geology. 1990; 22:239–252.

Cressie N, Lahiri SN. Asymptotics for REML Estimation of Spatial Covariance Parameters. Journal of Statistical Planning and Inference. 1996; 50:327–341.

De Oliveira V. On Optimal Point and Block Prediction in Log-Gaussian Random Fields. Scandinavian Journal of Statistics. 2006; 33:523–540.

De Oliveira V, Kedem B, Short DA. Bayesian Prediction of Transformed Gaussian Random Fields. Journal of the American Statistical Association. 1997; 92:1422–1433.

De Oliveira V, Rui C. On Shortest Prediction Intervals in Log-Gaussian Random Fields. Computational Statistics and Data Analysis. 2009; 53:4345–4357.

Escobar LA, Meeker WQ. Statistical Prediction Based on Censored Life Data. Technometrics. 1999; 41:113–124.

Fonseca G, Giummolè F, Vidoni P. A Note About Calibrated Prediction Regions and Distributions. Journal of Statistical Planning and Inference. 2012; 142:2726–2734.

Gelfand AE, Zhu L, Carlin BP. On the Change of Support Problem for Spatio-temporal Data. Biostatistics. 2001; 2:31–45. [PubMed: 12933555]

Giummolè F, Vidoni P. Improved Prediction Limits for a General Class of Gaussian Models. Journal of Time Series Analysis. 2010; 31:483–493.

Goovaerts, P. Geostatistics for Natural Resources Evaluation. Oxford University Press; 1997.

Gotway CA, Young LJ. A Geostatistical Approach to Linking Geographically Aggregated Data From Different Sources. Journal of Computational and Graphical Statistics. 2007; 16:1–21.

Gotway CA, Young LJ. Combining Incompatible Spatial Data. Journal of the American Statistical Association. 2002; 97:632–648.

Handcock MS, Stein ML. A Bayesian Analysis of Kriging. Technometrics. 1993; 35:403–410.

Harris IR. Predictive Fit for Natural Exponential Families. Biometrika. 1989; 76:675–684.

Lawless JF, Fredette M. Frequentist Prediction Intervals and Predictive Distributions. Biometrika. 2005; 92:529–542.

Mardia KV, Marshall RJ. Maximum Likelihood Estimation of Models of Residual Covariance in Spatial Regression. Biometrika. 1984; 71:135–146.

O'Hagan A. Bayes-Hermite Quadrature. Journal of Statistical Planning and Inference. 1991; 29:245–260.

Ribeiro PJ, Diggle PJ. geoR: A Package for Geostatistical Analysis. R-NEWS. 2001; 1:14–18.

Schelin L, Sjöstedt-de Luna S. Kriging Prediction Intervals Based on Semipara-metric Bootstrap. Mathematical Geosciences. 2010; 42:985–1000.

Sjöstedt-de Luna S, Young A. The Bootstrap and Kriging Prediction Intervals. Scandinavian Journal of Statistics. 2003; 30:175–192.

Tanaka, K. Time Series Analysis: Nonstationary and Noninvertible Distribution Theory. Wiley; 1996.

Ueki M, Fueda F. Adjusting Estimative Prediction Limits. Biometrika. 2007; 94:509–511.

Vidoni P. A Note on Modified Estimative Prediction Limits and Distributions. Biometrika. 1998; 85:949–953.

Wang F, Wall MM. Incorporating Parameter Uncertainty Into Prediction Intervals for Spatial Data Modeled via a Parametric Bootstrap. Journal of Agricultural, Biological and Environmental Statistics. 2003; 8:296–309.

Zhang H. Inconsistent Estimation and Asymptotically Equal Interpolations in Model-Based Geostatistics. Journal of the American Statistical Association. 2004; 99:250–261.

Appendix

Below we provide proofs for the results stated in Section 3. These results and their proofs are well known for random fields in the line, e.g., Cramér and Leadbetter (1967, Chapter 3) and Tanaka (1996, Chapter 2), but a treatment for random fields in the plane is difficult to find. In what follows we assume without lost of generality that $|B| = 1$ (otherwise $|B|^{-1}$ is absorbed by $w(\mathbf{s})$).

Lemma A.1

(Loève's criterion for L^2 convergence). Let $(T_m)_{m\geq 1}$ be a sequence in L^2. Then $(T_m)_{m\geq 1}$ converges in L^2 if and only $E\{T_m T_{m'}\}$ converges to a finite limit as $m, m' \to \infty$.

Proof

See Cramér and Leadbetter (1967) or Tanaka (1996).

Proof of Proposition 3.1

Assume first that $\mu(\mathbf{s}) = 0$. By the continuity hypotheses we have that the function $w(\mathbf{s})w(\mathbf{u})E\{Z(\mathbf{s})Z(\mathbf{u})\}$ is Riemann integrable on $B \times B$, so

$$\int\int_{B\times B} w(\mathbf{s})w(\mathbf{u})E\{Z(\mathbf{s})Z(\mathbf{u})\}d\mathbf{s}d\mathbf{u}<\infty. \quad (12)$$

Now, let $(\Delta_m)_{m\geq 1}$ be a sequence of finite partitions of B with the property that $\lim_{m\to\infty}[\Delta_m] = 0$. Then

$$E\{S_B(\Delta_m)S_B(\Delta_{m'})\}=E\left\{\sum_k\sum_{k'} w(\mathbf{t}_k)w(\mathbf{t}_{k'})Z(\mathbf{t}_k)Z(\mathbf{t}_{k'})|J_k|\,|J_{k'}|\right\}$$
$$=\sum_k\sum_{k'} w(\mathbf{t}_k)w(\mathbf{t}_{k'})E\{Z(\mathbf{t}_k)Z(\mathbf{t}_{k'})\}|J_k|\,|J_{k'}|,$$

and this sum converges to the left hand side in (12) as $m, m' \to \infty$. Hence, $w(\cdot)Z(\cdot)$ is L^2–integrable on B by Lemma A.1. If $\mu(\mathbf{s}) \neq 0$, the result follows by using the decomposition $Z(\mathbf{s}) = \mu(\mathbf{s}) + W(\mathbf{s})$ and noting that $w(\cdot)\mu(\cdot)$ is Riemann integrable on B.

Proof of Proposition 3.2

Recall that Z_B is defined as the L^2–limit of $S_B(\Delta_m)$, where $(\Delta_m)_{m\geq 1}$ is a sequence of finite partitions of B with the property that $\lim_{m\to\infty}[\Delta_m] = 0$. Since the order of expectation and limit is exchangeable under L^2–convergence, we have

$$E\{Z_B\}=\lim_{m\to\infty} E\{S_B(\Delta_m)\}=\lim_{m\to\infty}\sum_k w(\mathbf{t}_k)\mu(\mathbf{t}_k)|J_k|=\frac{1}{|B|}\int_B w(\mathbf{s})\mu(\mathbf{s})d\mathbf{s},$$

where the above integral exist since B is bounded and the integrand is piecewise continuous on B. Likewise, since $Z_B Z_C$ is the L^2–limit of $S_B(\Delta_m)S_C(\Delta_{m'})$, where $(\Delta_m)_{m\geq 1}$ and $(\Delta_{m'})_{m'\geq 1}$ are sequences of finite partitions of, respectively, B and C with the properties that $\lim_{m\to\infty}[\Delta_m] = \lim_{m'\to\infty}[\Delta_{m'}] = 0$, we have

$$E\{Z_B Z_C\}=\lim_{m,m'\to\infty} E\{S_B(\Delta_m)S_C(\Delta_{m'})\}$$
$$=\lim_{m,m'\to\infty}\sum_k\sum_{k'} w(\mathbf{t}_k)w(\mathbf{t}_{k'})E\{Z(\mathbf{t}_k)Z(\mathbf{t}_{k'})\}|J_k|\,|J_{k'}|$$
$$=\frac{1}{|B||C|}\int\int_{B\times C} w(\mathbf{s})w(\mathbf{u})\left(\sigma^2 K_\phi(\mathbf{s},\mathbf{u})+\mu(\mathbf{s})\mu(\mathbf{u})\right)d\mathbf{s}d\mathbf{u},$$

where again the above integral exist since B and C are bounded and the integrand is piecewise continuous on $B \times C$. This and (4) imply (6). A similar argument shows that (5) holds.

Lemma A.2

Let $(T_m)_{m\geq 1}$ be a convergent sequence in L^2, where T_m is normally distributed for all $m \geq 1$. Then its limit is also normally distributed.

Proof

By assumption, the moment generating function of T_m is $\psi_m(t)=\exp(\nu_m t+\xi_m^2 t^2/2)$ for some $\nu_m \in \mathbb{R}$ and $\xi_m \geq 0$. If T is the L^2 limit of $(T_m)_{m\geq 1}$, we have that as $m \to \infty$, $\nu_m \to \nu$ and $\xi_m \to \xi$ for some $\nu \in \mathbb{R}$ and $\xi \geq 0$, and by the continuity of the exponential function, $\psi_m(t) \to \exp(\nu t + \xi^2 t^2/2)$ as $m \to \infty$, for every $t \in \mathbb{R}$ Since convergence in L^2 implies convergence in distribution, we must have that the moment generating function of T is $\exp(\nu t + \xi^2 t^2/2)$, so T is normally distributed.

Proof of Proposition 3.3

Let $a_0 \in \mathbb{R}$, $\mathbf{a} \in \mathbb{R}^n$ and $(\Delta_m)_{m\geq 1}$ be a sequence of finite partitions of B with the property that $\lim_{m\to\infty}[\Delta_m] = 0$. Since the random field $Z(\cdot)$ is Gaussian, it holds that the random variables $T_m := \mathbf{a}'\mathbf{Z}_{\text{obs}} + a_0 S_B(\Delta_m)$ have normal distributions for $m \geq 1$. In addition

$$E\{(T_m-\mathbf{a}'\boldsymbol{Z}_{\text{obs}}-a_0 Z_B)^2\}=a_0^2 E\{(S_B(\Delta_m)-Z_B)^2\} \to 0, \quad \text{as } m \to \infty,$$

by the definition of Z_B. So $(T_m)_{m\geq 1}$ converges to $\mathbf{a}'\mathbf{Z}_{\text{obs}} + a_0 Z_B$ in L^2. By Lemma A.2, this implies that $\mathbf{a}'\mathbf{Z}_{\text{obs}} + a_0 Z_B$ is also normally distributed, which in turn implies (by definition) that $(\boldsymbol{Z}'_{\text{obs}}, Z_B)$ has a multivariate Gaussian distribution, since a_0 and $\mathbf{a}$ are arbitrary. Finally, the means, variances and covariances of this joint distribution are given by the assumed mean and covariance functions of $Z(\cdot)$, and the results in Proposition 3.2.

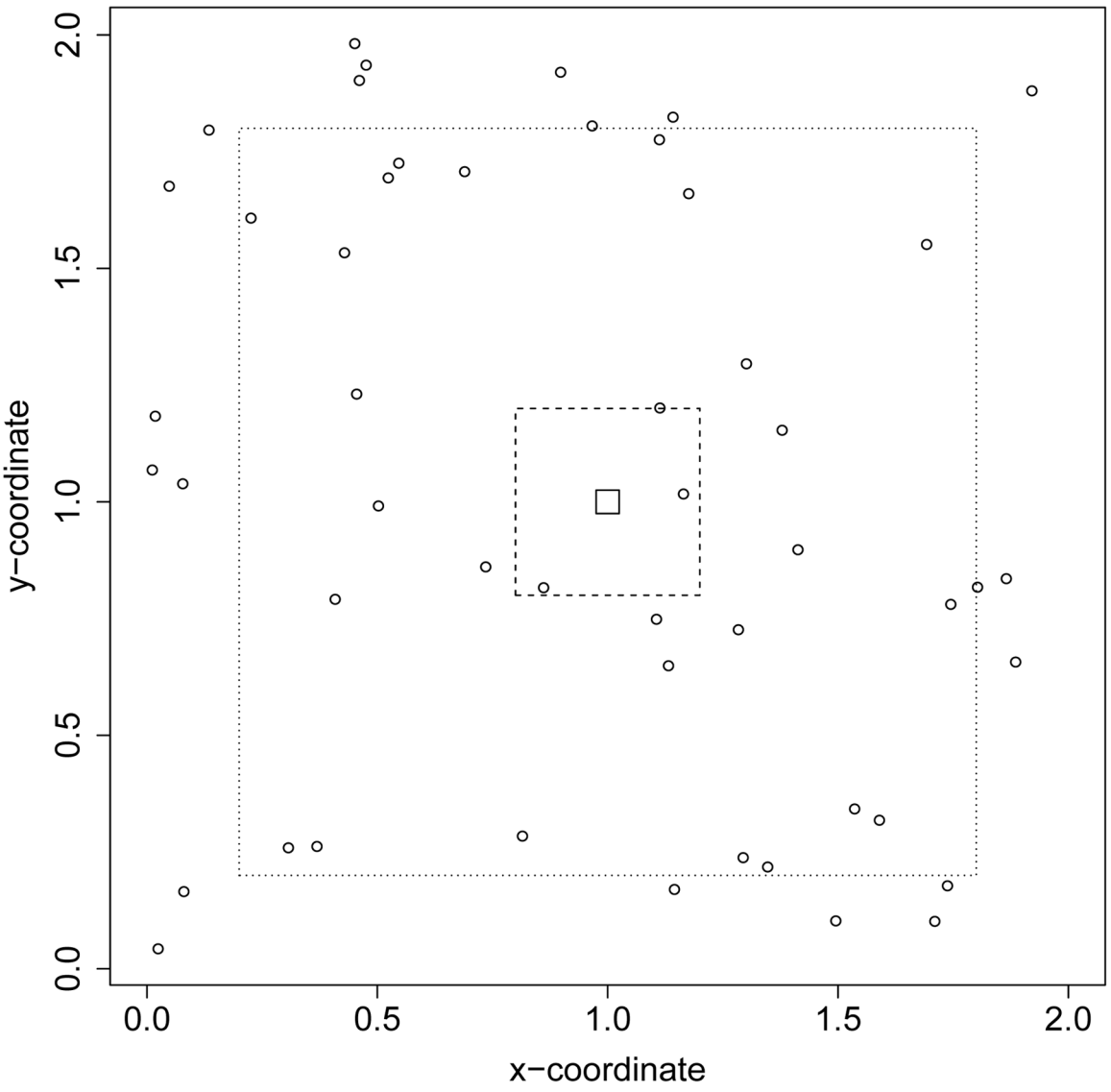

Figure 1.
Sampling locations (○) and integration regions: B_1 (large block), B_2 (medium block) and B_3 (small block).

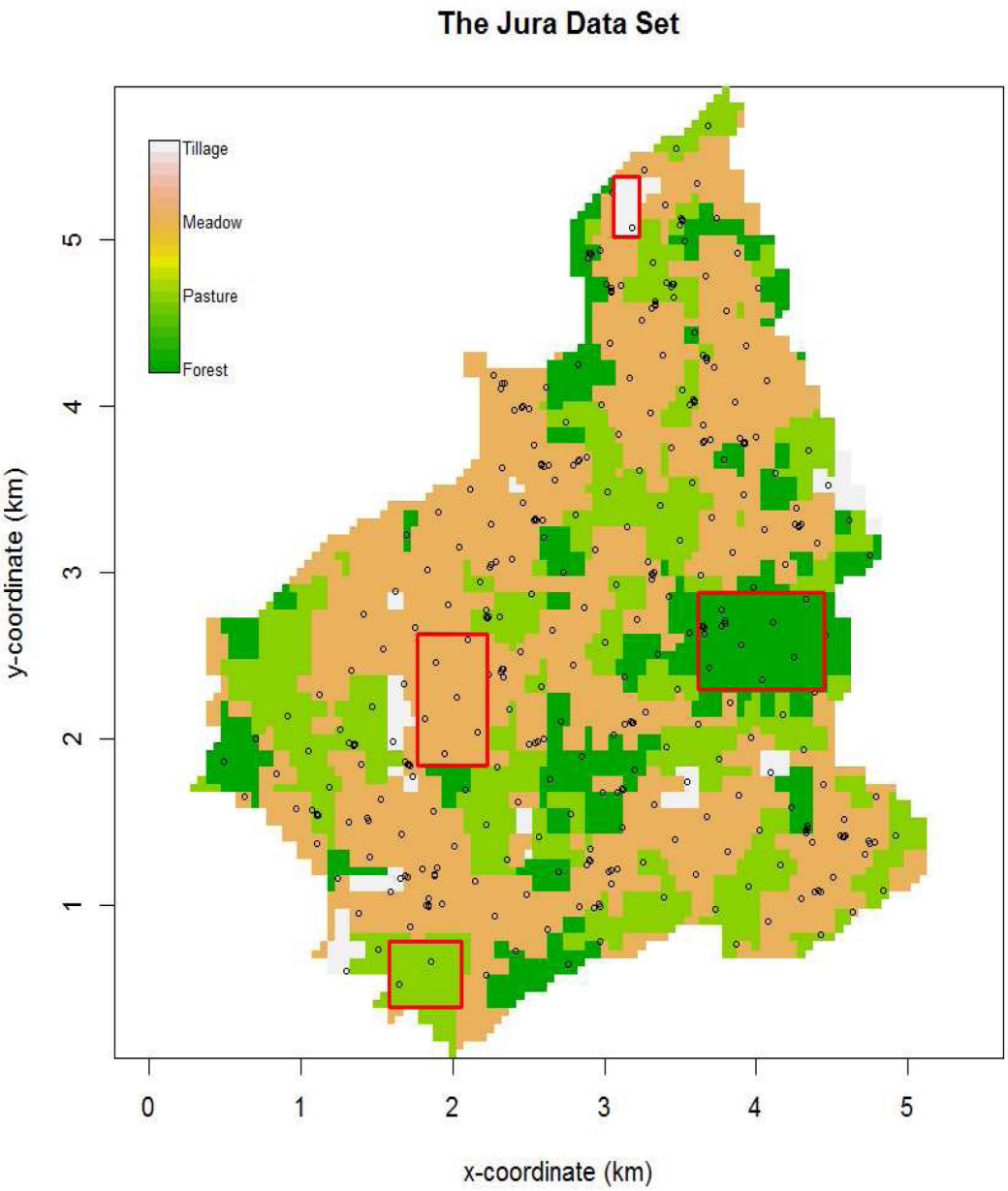

Figure 2.
Map of the Jura region with the 359 sampling locations (○), the land use type subregions (color-coded) and four rectangular blocks (red lines).

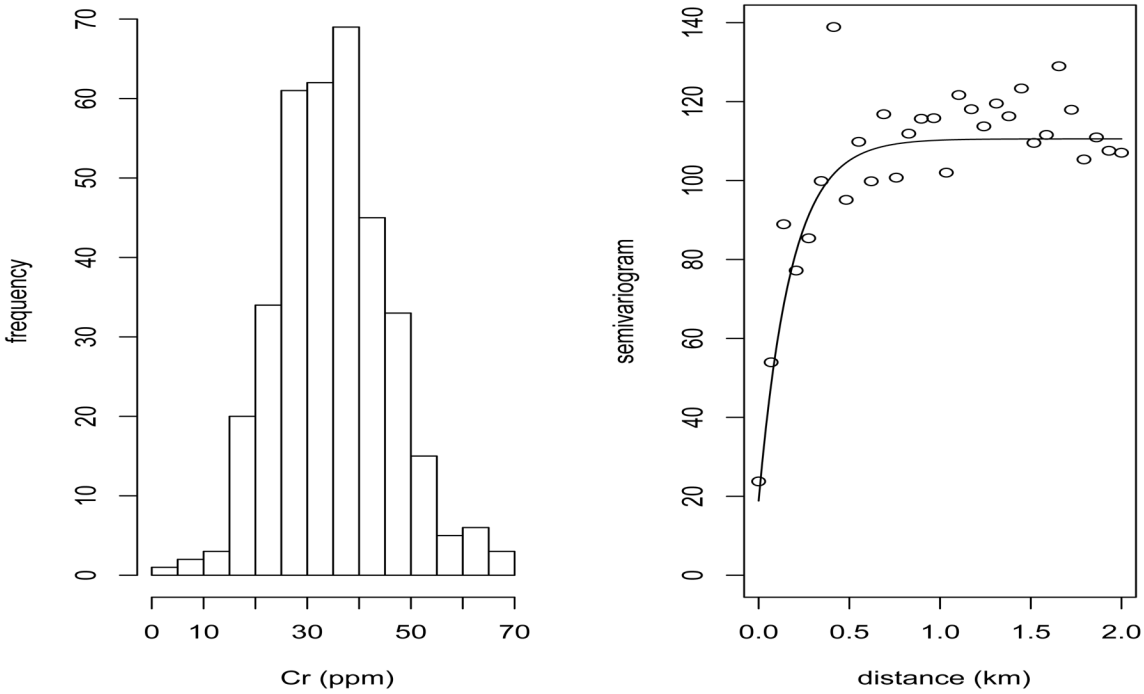

Figure 3.

Histogram of measured chromium traces (left); empirical (○) and fitted (line) semi-variograms of measured chromium traces (right).

Table 1

Coverage probabilities of the different types of 95% prediction intervals for Z_{B_1}, with $n = 50$ and $B_1 = [0.2, 1.8] \times [0.2, 1.8]$ (large block average). The top of the table displays the coverage for models with constant mean function, and the bottom displays the coverage for models with non-constant mean function.

σ^2	0.5		2.0	
φ	0.2	0.8	0.2	0.8
		$\tau^2 = 0$		
plug-in	0.930	0.937	0.930	0.940
ind calib	0.947	0.941	0.947	0.949
dir calib	0.947	0.941	0.945	0.948
true params	0.954	0.950	0.948	0.949
		$\tau^2 = \sigma^2/4$		
plug-in	0.933	0.924	0.932	0.936
ind calib	0.949	0.944	0.948	0.948
dir calib	0.947	0.945	0.941	0.947
true params	0.953	0.949	0.952	0.949
		$\tau^2 = 0$		
plug-in	0.938	0.935	0.929	0.935
ind calib	0.955	0.942	0.943	0.950
dir calib	0.951	0.941	0.944	0.948
true params	0.952	0.950	0.954	0.949
		$\tau^2 = \sigma^2/4$		
plug-in	0.928	0.937	0.923	0.939
ind calib	0.943	0.950	0.949	0.951
dir calib	0.939	0.947	0.942	0.949
true params	0.952	0.955	0.951	0.952

Table 2

Coverage probabilities of the different types of 95% prediction intervals for Z_{B_2}, with $n = 50$ and $B_2 = [0.8, 1.2] \times [0.8, 1.2]$ (medium block average). The top of the table displays the coverage for models with constant mean function, and the bottom displays the coverage for models with non-constant mean function.

σ^2	0.5		2.0	
φ	0.2	0.8	0.2	0.8
		$\tau^2 = 0$		
plug-in	0.939	0.940	0.939	0.945
ind calib	0.960	0.949	0.958	0.953
dir calib	0.960	0.949	0.954	0.953
true params	0.959	0.959	0.956	0.951
		$\tau^2 = \sigma^2/4$		
plug-in	0.914	0.912	0.905	0.925
ind calib	0.952	0.945	0.942	0.947
dir calib	0.950	0.941	0.939	0.949
true params	0.958	0.952	0.948	0.951
		$\tau^2 = 0$		
plug-in	0.912	0.935	0.915	0.931
ind calib	0.946	0.945	0.942	0.945
dir calib	0.938	0.944	0.935	0.943
true params	0.949	0.947	0.948	0.950
		$\tau^2 = \sigma^2/4$		
plug-in	0.910	0.924	0.912	0.918
ind calib	0.945	0.947	0.945	0.942
dir calib	0.940	0.941	0.940	0.939
true params	0.960	0.960	0.949	0.950

Table 3

Coverage probabilities of the different types of 95% prediction intervals for Z_{B_3}, with $n = 50$ and $B_3 = [0.975, 1.025] \times [0.975, 1.025]$ (small block average). The top of the table displays the coverage for models with constant mean function, and the bottom displays the results for coverage with non-constant mean function.

σ^2	0.5		2.0	
φ	0.2	0.8	0.2	0.8
		$\tau^2 = 0$		
plug-in	0.940	0.945	0.941	0.946
ind calib	0.951	0.950	0.952	0.949
dir calib	0.950	0.949	0.951	0.950
true params	0.953	0.952	0.957	.950
		$\tau^2 = \sigma^2/4$		
plug-in	0.902	0.910	0.903	0.923
ind calib	0.947	0.943	0.945	0.950
dir calib	0.939	0.944	0.942	0.951
true params	0.952	0.949	0.951	0.958
		$\tau^2 = 0$		
plug-in	0.925	0.933	0.935	0.931
ind calib	0.945	0.948	0.949	0.950
dir calib	0.944	0.950	0.951	0.947
true params	0.948	0.949	0.950	0.952
		$\tau^2 = \sigma^2/4$		
plug-in	0.912	0.920	0.924	0.910
ind calib	0.947	0.949	0.946	0.944
dir calib	0.945	0.947	0.942	0.945
true params	0.959	0.952	0.957	0.949

Table 4

Coverage probabilities of the different types of 95% prediction intervals for Z_{B_1} (top table), Z_{B_2} (middle table) and Z_{B_3} (bottom table), with $n = 200$ and constant mean function.

σ^2	0.5		2.0	
φ	0.2	0.8	0.2	0.8
		$\tau^2 = 0$		
plug-in	0.932	0.935	0.931	0.936
ind calib	0.949	0.945	0.942	0.951
dir calib	0.948	0.943	0.946	0.950
		$\tau^2 = \sigma^2/4$		
plug-in	0.928	0.923	0.930	0.932
ind calib	0.942	0.950	0.947	0.951
dir calib	0.944	0.946	0.945	0.951
		$\tau^2 = 0$		
plug-in	0.940	0.939	0.931	0.943
ind calib	0.952	0.949	0.951	0.955
dir calib	0.955	0.947	0.950	0.954
		$\tau^2 = \sigma^2/4$		
plug-in	0.918	0.911	0.906	0.929
ind calib	0.953	0.945	0.944	0.947
dir calib	0.953	0.943	0.940	0.948
		$\tau^2 = 0$		
plug-in	0.941	0.943	0.940	0.945
ind calib	0.952	0.951	0.950	0.949
dir calib	0.950	0.948	0.950	0.951
		$\tau^2 = \sigma^2/4$		
plug-in	0.904	0.912	0.905	0.919
ind calib	0.947	0.943	0.945	0.952
dir calib	0.940	0.946	0.947	0.953

Table 5

Coverage probabilities of the different types of 95% prediction intervals for Z_B with n = 25, 50 and 200 in an increasing domain asymptotic regime.

n	**25**	**50**	**200**
		$\tau^2 = 0$	
plug-in	0.905	0.928	0.951
ind calib	0.933	0.953	0.953
dir calib	0.929	0.955	0.953
		$\tau^2 = 0.5/4$	
plug-in	0.890	0.920	0.947
ind calib	0.950	0.944	0.952
dir calib	0.944	0.942	0.949

Table 6

Coordinates of the rectangular blocks (in km) and their respective land use type.

block	block coordinates	land use type
1	[3.06, 3.23] × [5.02, 5.38]	tillage
2	[1.77, 2.23] × [1.84, 2.63]	meadow
3	[1.58, 2.06] × [0.38, 0.78]	pasture
4	[3.62, 4.45] × [2.3, 2.88]	forest

Table 7

95% prediction intervals for the spatial averages of chromium traces over the rectangular blocks in Table 6.

block	plug-in	indirectly calibrated	directly calibrated
1	(31.04, 46.66)	(30.90, 46.79)	(30.90, 46.78)
2	(35.34, 44.04)	(35.25, 44.14)	(35.24, 44.13)
3	(32.90, 46.64)	(32.67, 49.88)	(32.66, 46.85)
4	(22.24, 29.58)	(21.16, 29.66)	(22.17, 29.65)