



UNIVERSITY OF LEEDS

This is a repository copy of *Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/93197/>

Version: Accepted Version

Article:

Garcia-Papani, F, Uribe-Opazo, MA, Leiva, V et al. (1 more author) (2017)
Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data. *Stochastic Environmental Research and Risk Assessment*, 31 (1). pp. 105-124. ISSN 1436-3240

<https://doi.org/10.1007/s00477-015-1204-4>

© 2016, Springer. This is an author produced version of a paper published in *Stochastic Environmental Research and Risk Assessment*. Uploaded in accordance with the publisher's self-archiving policy. The final publication is available at Springer via <https://doi.org/10.1007/s00477-015-1204-4>

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

SERRA manuscript No. (will be inserted by the editor)

Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data

Fabiana Garcia-Papani · Miguel Angel Uribe-Opazo ·
Victor Leiva · Robert G Aykroyd

Abstract Applications of statistical models to describe spatial dependence in geo-referenced data are widespread across many disciplines including the environmental sciences. Most of these applications assume that the data follow a Gaussian distribution. However, in many of them the normality assumption, and even a more general assumption of symmetry, are not appropriate. In non-spatial applications, where the data are uni-modal and positively skewed, the Birnbaum-Saunders distribution has excelled. This paper proposes a spatial log-linear model based on the Birnbaum-Saunders distribution. Model parameters are estimated using the maximum likelihood method. Local influence diagnostics are derived to assess the sensitivity of the estimators to perturbations in the response variable. As illustration, the proposed model and its diagnostics are used to analyse a real-world agricultural data set, where the spatial variability of phosphorus concentration in the soil is considered—which is extremely important for agricultural management.

Keywords asymmetric distributions · local influence · Matérn model · maximum likelihood methods · Monte Carlo simulation · non-normality · R software · spatial data analysis

1 Introduction

Spatial statistical models take into account the dependence of a variable over space using geo-referenced data. These models are essential in many fields; see, for example, Krige (1951), Mardia and Marshall (1984), Waller and Gotway (2004) and Militino et al. (2006). Recent studies include Borssoi et al. (2011), Uribe-Opazo et al. (2012) and Grzegozewski et al. (2013). All these papers make the assumption that a normal (or Gaussian) distribution is appropriate for modelling the spatial variability. However, such an assumption is not always appropriate; see Davis (1952) and Lange et al. (1989). One approach to deal with non-normality is to transform the data to achieve at least approximate normality. Nevertheless, when working with data transformations, problems, such as the difficulty of interpreting the results from the analysis on the original scale, can be introduced into the modelling; see Azzalini and Capitanio (1999). When the normality assumption is not valid, an alternative approach is to carry out the modelling with some non-normal distribution to be suitable for the data under analysis. For example, Assumpção et al. (2011, 2014) conducted a geo-statistical study using the Student-t distribution, which has heavier tails than the normal distribution. De Bastiani et al. (2015) studied spatial modelling and diagnostics based on the family of elliptic (symmetric) distributions, which has as members the Gaussian and Student-t cases. With all of this said, elliptic models are still not appropriate if the data follow a skew distribution.

Fabiana Garcia-Papani, Miguel Angel Uribe-Opazo
Postgraduate Program in Agricultural Engineering, Universidade Estadual do Oeste do Paraná, Cascavel-PR, Brazil. E-mail: fgarcia-papani@gmail.com, E-mail: mopazo@unioeste.br

Victor Leiva
Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Viña del Mar, Chile. E-mail: victorleivasanchez@gmail.com

Robert G Aykroyd
Department of Statistics, University of Leeds, Leeds, UK. E-mail: r.g.aykroyd@leeds.ac.uk

There has been little work in the literature that investigates the use of asymmetric distributions to analyse spatial data. However, in non-spatial situations, many distributions have been proposed to model phenomena that give rise to skew data, such as, the Birnbaum-Saunders (BS), exponential, gamma, log-normal and Weibull distributions; see Johnson et al. (1994, 1995). In particular, the BS distribution was proposed for modelling random variables describing processes of fatigue by Birnbaum and Saunders (1969). Applications in earth sciences of the BS distribution have been considered by, for example, Leiva et al. (2008, 2009, 2015a), Podlaski (2008), Vilca et al. (2010), Marchant et al. (2013) and Saulo et al. (2013). What makes this distribution attractive for the analysis of skew data are its properties and its relationship with the normal distribution; see Johnson et al. (1995, pp. 651–663). In contrast to its original application to fatigue processes, Leiva et al. (2015c) justified why the BS distribution is suitable for modelling earth and environmental data using theoretical arguments based on the law of proportionate effects. Rieck and Nedelman (1991) defined a relationship between the BS distribution and its logarithmic version, named the log-BS distribution. They used this relationship to propose a BS fixed effect model; whereas Villegas et al. (2011) considered a BS mixed effect model. A multivariate extension to the BS fixed effect model was studied by Marchant et al. (2015). An approach to BS spatial modelling was provided by Xia et al. (2011), who presented a methodology based on semi-Markov process to produce a spatio-temporal model for the movement of tourists. The authors considered several distributions, including the BS model. To date, however, spatial models based on the BS distribution have not been studied in practical examples.

The identification of cases that can produce substantial changes in the estimated parameters is an important step in any statistical investigation. The task of detecting possible atypical cases can be addressed by eliminating cases one-by-one from the data set and measuring the effects on estimated parameters –this is known as global influence; see Cook (1987). Another method for detecting cases that could potentially be influential was proposed by Cook (1987), which is known as local influence. This method studies the effect of small perturbations introduced into the models and/or the data on the maximum likelihood (ML) estimates. Different perturbation schemes are often considered to evaluate the sensitivity of ML estimates of the model parameters to such perturbations. The local influence method has at least two advantages over the global influence method: it has a lower computational cost, especially when the number of cases is large, and it allows us to detect groups of data exerting a joint influence. Zhu et al. (2007) proposed a methodology for choosing a perturbation scheme that is appropriate for the particular model being considered. Gimenez and Galea (2013) applied the method proposed by Zhu et al. (2007) to heteroscedastic models with functional measurement errors. Galea et al. (2004) applied the local influence method in the BS fixed effect model, whereas Leiva et al. (2014, 2015b) and Liu et al. (2015) derived diagnostic tools in accelerated life models, in fixed effect models with stochastic restrictions and in the possibly heteroskedastic linear model with exact restrictions. In spatial modelling, diagnostic techniques have been discussed for Gaussian models by Militino et al. (2006) and Uribe-Opazo et al. (2012), for Student-t models by Assumpção et al. (2014), and for elliptic models by De Bastiani et al. (2015).

The main objective of this paper is to develop a spatial log-linear model based on the BS distribution and to derive its corresponding diagnostics. This distribution can be more appropriate than the Gaussian distribution in the analysis of spatial data with positive asymmetric behaviour. ML estimators of the model parameters and local influence diagnostic tools are derived for the BS spatial model. A computational framework in R code of the developed methodology is available from the authors under request. Specifically, Section 2 provides a background on uni- and multi-variate BS and log-BS distributions and on spatial modelling. Section 3 formulates the BS spatial log-linear model, estimates its parameters using the ML method and derives an appropriate perturbation scheme for the response variable using the methodology proposed by Zhu et al. (2007). Section 4 conducts two Monte Carlo (MC) simulation studies for evaluating the performance of the corresponding ML estimators and diagnostic tools. Section 5 illustrates the potential applications of the proposed model and its diagnostics with real-world data from agricultural engineering. Section 6 discusses some conclusions and possible future work. Detailed algebra is presented in appendices.

2 Background

2.1 The Birnbaum-Saunders distribution

If a random variable T follows a BS distribution with shape parameter α and scale parameter β , we use the notation $T \sim \text{BS}(\alpha, \beta)$. The distribution can be defined by its cumulative distribution function (CDF) given by

$$F_T(t; \alpha, \beta) = \Phi \left(\frac{1}{\alpha} \left(\sqrt{t/\beta} - \sqrt{\beta/t} \right) \right), \quad t > 0, \alpha > 0, \beta > 0, \quad (1)$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution. Then, the probability density function (PDF) of T obtained from (1) is expressed as

$$f_T(t; \alpha, \beta) = \frac{1}{2\alpha} \left(\sqrt{1/\beta t} + \sqrt{\beta/t^3} \right) \phi \left(\frac{1}{\alpha} \left(\sqrt{t/\beta} - \sqrt{\beta/t} \right) \right), \quad t > 0, \alpha > 0, \beta > 0, \quad (2)$$

where $\phi(\cdot)$ is the standard normal PDF. Thus, in turn the PDF in (2) can be re-written as

$$f_T(t; \alpha, \beta) = \frac{\exp(\alpha^{-2})}{2\alpha\sqrt{2\pi}\beta} \exp \left(-\frac{1}{2\alpha^2} \left(\frac{t}{\beta} + \frac{\beta}{t} \right) \right) t^{-\frac{3}{2}} (t + \beta), \quad t > 0, \alpha > 0, \beta > 0. \quad (3)$$

Also, it may be said that a continuous random variable T has a BS distribution with parameters $\alpha > 0$ and $\beta > 0$, if and only if $Z = (1/\alpha)(\sqrt{T/\beta} - \sqrt{\beta/T}) \sim \text{N}(0, 1)$. Some properties of the BS distribution are presented as follows. If $T \sim \text{BS}(\alpha, \beta)$, then: (i) $E(T) = \beta(1 + \alpha^2/2)$ and $\text{Var}(T) = (\alpha\beta)^2(1 + 5\alpha^2)$; (ii) if $b > 0$, then $bT \sim \text{BS}(\alpha, b\beta)$, which means that the BS distribution is closed under scalar multiplication; (iii) $1/T \sim \text{BS}(\alpha, 1/\beta)$, which means that the BS distribution is closed under reciprocity; (iv) the median of the distribution of T is β , which can be directly obtained when $q = 0.5$ from its quantile function given by $t(q; \alpha, \beta) = F_T^{-1}(q; \alpha, \beta) = \beta(\alpha z(q)/2 + \sqrt{(\alpha z(q)/2)^2 + 1})^2$, for $0 < q < 1$, where $z(q)$ is the standard normal quantile function; and (v) the BS distribution is positively skewed as α increases and approximately symmetrical around β as α goes to zero; see Fig. 1(left).

2.2 The log-Birnbaum-Saunders distribution

A continuous random variable Y has a log-BS distribution with shape parameter $\alpha > 0$ and location parameter $\mu \in \mathbb{R}$, which is denoted by $\text{log-BS}(\alpha, \mu)$, if and only if $Z = (2/\alpha)\sinh((Y - \mu)/2) \sim \text{N}(0, 1)$. Then, the CDF of Y is given by

$$F_Y(y; \alpha, \mu) = \Phi \left(\frac{2}{\alpha} \sinh \left(\frac{y - \mu}{2} \right) \right), \quad -\infty < y, \mu < +\infty, \alpha > 0. \quad (4)$$

Consequently, from (4), the PDF of Y is obtained as

$$f_Y(y; \alpha, \mu) = \frac{1}{\alpha\sqrt{2\pi}} \cosh \left(\frac{y - \mu}{2} \right) \exp \left(-\frac{2}{\alpha^2} \sinh^2 \left(\frac{y - \mu}{2} \right) \right), \quad -\infty < y, \mu < +\infty, \alpha > 0. \quad (5)$$

Some properties of the log-BS distribution are presented as follows. If $Y \sim \text{log-BS}(\alpha, \mu)$, then: (i) $T = \exp(Y) \sim \text{BS}(\alpha, \beta)$, which means that the log-BS PDF given in (5) can be obtained from the standard normal PDF or from the BS PDF defined in (3); (ii) $E(Y) = \mu$; (iii) there is no closed form for the variance of Y , but based upon an asymptotic approximation for the moment generating function of the log-BS distribution, it follows that, if $\alpha \rightarrow 0$, then $\text{Var}(T) = \alpha^2 - \alpha^4/4$, and if $\alpha \rightarrow \infty$, then $\text{Var}(T) = 4(\log^2(\sqrt{2}\alpha) + 2 - 2\log(\sqrt{2}\alpha))$; (iv) if $X = \pm Y + d$, then $X \sim \text{log-BS}(\alpha, \pm\mu + d)$; and (v) the log-BS distribution is symmetric around μ , unimodal for $\alpha \leq 2$ and bimodal for $\alpha > 2$; see Fig. 1(right).

If a random vector $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$ follows an n -variate log-BS distribution, the notation $\mathbf{Y} \sim \text{log-BS}_n(\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is used. Here, the vector of shape parameters is $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)^\top$, with $\alpha_i > 0$, the vector of location parameters is $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^\top$, with $-\infty < \mu_i < +\infty$, for $i = 1, \dots, n$, and $\boldsymbol{\Sigma}$ is an $n \times n$ positive definite (non-singular) matrix of scale parameters. The CDF of \mathbf{Y} can be defined from (5) as

$$F_{\mathbf{Y}}(\mathbf{y}; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \Phi_n \left(\frac{2}{\alpha_1} \sinh \left(\frac{y_1 - \mu_1}{2} \right), \dots, \frac{2}{\alpha_n} \sinh \left(\frac{y_n - \mu_n}{2} \right); \boldsymbol{\Sigma} \right), \quad \mathbf{y} \in \mathbb{R}^n, \quad (6)$$

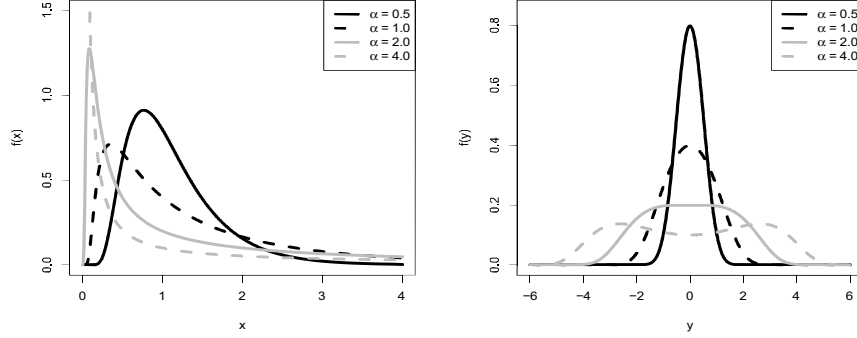


Fig. 1 PDF of (left) $\text{BS}(\alpha, 1)$ and (right) $\text{log-BS}(\alpha, 0)$ distributions for the indicated value of the shape parameter α .

where $\Phi_n(\cdot; \Sigma)$ is the CDF of the n -variate normal distribution with mean vector equal to zero and variance-covariance matrix Σ . Therefore, the PDF of \mathbf{Y} can be obtained from (6) as

$$f_{\mathbf{Y}}(\mathbf{y}; \alpha, \boldsymbol{\mu}, \Sigma) = \phi_n \left(\frac{2}{\alpha_1} \sinh \left(\frac{y_1 - \mu_1}{2} \right), \dots, \frac{2}{\alpha_n} \sinh \left(\frac{y_n - \mu_n}{2} \right); \Sigma \right) \prod_{i=1}^n \frac{1}{\alpha_i} \cosh \left(\frac{y_i - \mu_i}{2} \right), \quad (7)$$

for $\mathbf{y} \in \mathbb{R}^n$, where $\phi_n(\cdot; \Sigma)$ is the n -variate normal PDF with mean vector equal to zero and variance-covariance matrix Σ .

2.3 Spatial models

Consider a stochastic process $\{Y(s), s \in D\}$, which is defined over a region D , with $D \subset \mathbb{R}^2$, described by the spatial linear model

$$Y(s) = \mu(s) + \varepsilon(s), \quad s \in D, \quad (8)$$

where $\mu(\cdot)$ is a mean function and $\varepsilon(\cdot)$ is the model error. These errors have mean zero and common variance σ^2 , which means that $E(Y(s)) = \mu(s)$ and $\text{Var}(Y(s)) = \sigma^2$, for all $s \in D$.

If the spatial process $\{Y(s), s \in D\}$ is assumed to be stationary, then its mean function is constant, that is, $\mu(s) = \mu$, for all $s \in D$. If it is further assumed to be isotropic, then its covariance function only depends on the distance between spatial locations, that is,

$$\text{Cov}(Y(s_i), Y(s_j)) = C(s_i, s_j) = C(h_{ij}), \quad s_i, s_j \in D, \quad (9)$$

where $h_{ij} = \|s_i - s_j\|$ is the Euclidean distance between s_i and s_j .

2.4 Covariance and variogram models

Now suppose that n measurements are collected at a set of known spatial locations $\mathbf{s} = \{s_1, \dots, s_n\}$, providing the n -variate random vector $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$, where $Y_i = Y(s_i)$, for $i = 1, \dots, n$. The covariance between all pairs of random variables (Y_i, Y_j) is determined by an $n \times n$ scale matrix $\Sigma = [\sigma_{ij}]$, which must be symmetric and positive-definite, where $\sigma_{ij} = C(h_{ij})$, with $C(\cdot)$ given in (9). Then, Σ defines the spatial dependence structure for stationary and isotropic processes by a parameter vector $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \varphi_3)^\top$, where $\varphi_1 \geq 0$, $\varphi_2 \geq 0$ are parameters known as nugget effect and partial sill, respectively, whereas $\varphi_3 \geq 0$ is a parameter related to the effective range or spatial dependence radius $a = g(\varphi_3)$; see Mardia and Marshall (1984) and Uribe-Opazo et al. (2012). The nugget effect is related to an analytical error, indicating an unexplained variability from one point of the sampling grid to another. This variability may be attributed to measurement errors or to a variability not captured due to the sampling distance used; see, for example, Cambardella et al. (1994). The nugget effect can also act as a regulatory tool in spatial design for random fields making many designs feasible; see Müller and Stehlík (2009, 2010). We assume a particular parametric form for the scale matrix given by

$$\Sigma = \varphi_1 \mathbf{I}_n + \varphi_2 \mathbf{R}, \quad (10)$$

where \mathbf{I}_n is the $n \times n$ identity matrix and $\mathbf{R} = [r_{ij}]$ is an $n \times n$ symmetric matrix with diagonal elements $r_{ii} = 1$, for $i = 1, \dots, n$. Specific forms for r_{ij} given by $r_{ij} = \sigma_{ij}/\varphi_2$, with $i \neq j$ and $\varphi_2 \neq 0$, define the model used to explain the spatial dependence, with the most common forms being those obtained from the Matérn and power exponential families; see Isaaks and Srivastava (1989) and Diggle and Ribeiro (2007). In the family of Matérn models, we have

$$r_{ij} = \begin{cases} 1, & i = j, \\ \frac{1}{2^{\delta-1}\Gamma(\delta)} \left(\frac{h_{ij}}{\varphi_3}\right)^{\delta} K_{\delta}\left(\frac{h_{ij}}{\varphi_3}\right), & i \neq j, \end{cases} \quad (11)$$

where δ is a shape parameter, $\Gamma(\cdot)$ is the usual gamma function and $K_{\delta}(\cdot)$ is the modified Bessel function of third kind of order δ . From (11), we get

$$\sigma_{ij} = \begin{cases} \varphi_1 + \varphi_2, & i = j, \\ \frac{\varphi_2}{2^{\delta-1}\Gamma(\delta)} \left(\frac{h_{ij}}{\varphi_3}\right)^{\delta} K_{\delta}\left(\frac{h_{ij}}{\varphi_3}\right), & i \neq j. \end{cases} \quad (12)$$

In the family of power exponential models, for $i \neq j$, we have $r_{ij} = \exp(-(h_{ij}/\varphi_3)^p)$, where $0 < p \leq 2$ is a shape parameter, which implies

$$\sigma_{ij} = \begin{cases} \varphi_1 + \varphi_2, & i = j, \\ \varphi_2 \exp\left(-\left(\frac{h_{ij}}{\varphi_3}\right)^p\right), & i \neq j. \end{cases} \quad (13)$$

Although the models in (12) and (13) have no finite range, the effective range a can be defined as the smallest distance between two locations, such that the covariance has dropped to 5% of the maximum covariance, $C(0)$. The exponential and Gaussian models are members of the power exponential family when $p = 1$ and $p = 2$, respectively, and also of the Matérn family when $\delta = 0.5$ and $\delta \rightarrow \infty$, respectively. In stationary processes, from the covariance function given in (9), it is possible to define the variogram function by

$$\gamma(h) = C(0) - C(h), \quad h > 0, \quad (14)$$

where $C(0) = \varphi_1 + \varphi_2$ and $C(h)$ is specified from $C(h_{ij}) = \sigma_{ij}$, for $h = h_{ij}$, with a suitable member of the Matérn or power exponential families given in (12) and (13), respectively. The plot of points $(h, \gamma(h))$ obtained from the variogram function given in (14) is a useful tool in spatial statistics.

2.5 Kriging interpolation

In geo-statistical analyses, a commonly used method for interpolation is Kriging; see Krige (1951). The Kriging prediction is given by a linear combination of the observed data $\mathbf{y} = (y_1, \dots, y_n)^{\top}$ defined as

$$\hat{y}(s_0) = \sum_{i=1}^n \lambda_i y_i, \quad (15)$$

where $\hat{y}(s_0)$ is the predicted value at a new location s_0 and $\lambda_1, \dots, \lambda_n$ are weights chosen to define the best linear unbiased predictor. This can be achieved by minimizing the variance of the error with respect to the weights whilst requiring the predictor to be unbiased. It may be shown that, under the stationarity assumption, the values $\lambda_1, \dots, \lambda_n$ are given by the solution to $\mathbf{C}\boldsymbol{\lambda} = \boldsymbol{\mathcal{Y}}$, where

$$\mathbf{C} = \begin{pmatrix} C(s_1, s_1) & \dots & C(s_1, s_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ C(s_n, s_1) & \dots & C(s_n, s_n) & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix}; \boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n, \varrho)^{\top}; \boldsymbol{\mathcal{Y}} = (C(s_1, s_0), \dots, C(s_n, s_0), 1)^{\top};$$

and ϱ is a Lagrange multiplier introduced to ensure unbiasedness when minimizing the error variance.

2.6 Accuracy measures

To quantify the similarity between two maps, the global accuracy (GA) and kappa (κ) indexes can be used. Consider two maps (one called the reference map and the other the model map), both divided into the same m class, denoted by M_i , for $i = 1, \dots, m$. In addition, let N_{ij} , for $i, j = 1, \dots, m$, be the number of pixels (a pixel is defined as a single point in an image; for example, in our application of Section 5, 1 pixel $\approx 85\text{m}^2$) belonging to class M_i of the model map and to class M_j of the reference map, and N be the total number of pixels in each map. The GA index is based on those pixels that belong to the same class in both maps and is defined as $\text{GA} = (\sum_{i=1}^m N_{ii})/N$. Model and reference maps have an acceptable similarity if the GA index is greater than 0.85; see Anderson et al. (1976). The κ index is based on all of the pixels (those belonging to the same class or not) and is defined as $\kappa = (N \sum_{i=1}^m N_{ii} - \sum_{i=1}^m N_{\bullet i} N_{i \bullet}) / (N^2 - \sum_{i=1}^m N_{\bullet i} N_{i \bullet})$, where $N_{i \bullet} = \sum_{j=1}^m N_{ij}$ and $N_{\bullet i} = \sum_{j=1}^m N_{ji}$. Model and reference maps have a low similarity if $\kappa < 0.67$, a medium similarity if $0.67 \leq \kappa \leq 0.80$, and a high similarity if $\kappa > 0.80$; see Krippendorff (2004).

3 The Birnbaum-Saunders spatial model

3.1 Formulation of the model

Let $\{T(s), s \in D\}$ be a stochastic process defined over a region D , with $D \subset \mathbb{R}^2$. Suppose that n measurements $\mathbf{T} = (T_1, \dots, T_n)^\top$, where $T_i = T(s_i)$, for $i = 1, \dots, n$, are collected at a set of known spatial locations $\mathbf{s} = \{s_1, \dots, s_n\}$. Consider a spatial model of the form

$$T_i = \exp(\mu_i) \eta_i, \quad i = 1, \dots, n. \quad (16)$$

Assume stationarity, such that $\mu_i = \mu(s_i) = \mu$, and the model error $\eta_i = \eta(s_i) \sim \text{BS}(\alpha, 1)$, for $i = 1, \dots, n$. Then, $\exp(\mu)$ is the median of the model. Note that the shape parameter α is also assumed to be constant across the spatial locations; see Marchant et al. (2015). Applying a logarithmic transformation to (16), a BS spatial log-linear model is obtained as

$$Y_i = \log(T_i) = \mu + \log(\eta_i) = \mu + \varepsilon_i, \quad i = 1, \dots, n, \quad (17)$$

where $\varepsilon_i = \log(\eta_i) \sim \text{log-BS}(\alpha, 0)$, for $i = 1, \dots, n$. Note that the BS spatial log-linear model defined in (17) has a similar form to the model given in (8). For ease of notation, the model in (17) can be written in matrix form as

$$\mathbf{Y} = \mu \mathbf{1} + \boldsymbol{\varepsilon}, \quad (18)$$

with $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$, $\mathbf{1} = (1, \dots, 1)^\top$ and $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)^\top$ being $n \times 1$ vectors. Here, $\boldsymbol{\varepsilon}$ is a stationary spatial stochastic process with mean vector $\text{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$. Suppose that the covariance between all pairs (Y_i, Y_j) is determined by the $n \times n$ scale matrix $\boldsymbol{\Sigma}$ satisfying the conditions given in (10), where the elements of \mathbf{R} can be modelled with the Matérn structure given in (11).

3.2 Parameter estimation

Let $\boldsymbol{\theta} = (\alpha, \mu, \varphi_1, \varphi_2, \varphi_3)^\top$ be the vector of unknown parameters of the spatial model formulated in (18) to be estimated. Then, the likelihood function for $\boldsymbol{\theta}$, based on the observations $\mathbf{y} = (y_1, \dots, y_n)^\top$ of \mathbf{Y} , obtained from (7) is given by

$$\mathcal{L}(\boldsymbol{\theta}) = \mathcal{L}(\boldsymbol{\theta}; \mathbf{y}) = \frac{1}{\alpha^n (2\pi)^{\frac{n}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left(-\frac{2}{\alpha^2} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \mathbf{V}\right) \prod_{i=1}^n \cosh\left(\frac{y_i - \mu}{2}\right), \quad (19)$$

where $\mathbf{V} = (V_1, \dots, V_n)^\top$ is an $n \times 1$ vector with elements $V_i = \sinh((y_i - \mu)/2)$, for $i = 1, \dots, n$. The corresponding log-likelihood function for $\boldsymbol{\theta}$ obtained from (19) is then

$$\ell(\boldsymbol{\theta}) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log(|\boldsymbol{\Sigma}|) - n \log(\alpha) - \frac{2}{\alpha^2} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \mathbf{V} + \sum_{i=1}^n \log\left(\cosh\left(\frac{y_i - \mu}{2}\right)\right). \quad (20)$$

The ML method defines the estimator $\hat{\theta}$ of θ as the vector which maximises $\mathcal{L}(\theta)$, or equivalently $\ell(\theta)$, over the parameter space of θ . Thus,

$$\hat{\theta} = \arg \max_{\theta} \ell(\theta). \quad (21)$$

When the value in (21) is associated with a stationary point, it can be obtained from the solution of a homogeneous system of equations created from the score vector and given by

$$\frac{\partial \ell(\theta)}{\partial \alpha} = 0, \frac{\partial \ell(\theta)}{\partial \mu} = 0, \frac{\partial \ell(\theta)}{\partial \varphi} = \mathbf{0}; \quad (22)$$

see details of the score vector in Appendix I. Note that no analytical solution to the system of equations given in (22) can be obtained. Then, the ML estimator $\hat{\theta}$ must be computed with an iterative procedure to solve the non-linear system. Here the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton procedure (see Nocedal and Wright, 1999; Lange, 2001) may be used through the functions `optim` and `optimx` implemented in the software R; see www.R-project.org and R-Team (2015). The signs of the determinants of the Hessian matrix and of its minors were also checked to ensure that a valid maximum had been found.

Under the usual regularity conditions, the ML estimator $\hat{\theta}$ is consistent for θ and has an asymptotic normal distribution. Then, as $n \rightarrow \infty$,

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{D} N_5(\mathbf{0}, \mathbf{J}(\theta)^{-1}), \quad (23)$$

where $\mathbf{J}(\theta) = \lim_{n \rightarrow \infty} \mathbf{I}(\theta)/n$, with $\mathbf{I}(\theta)$ being the expected Fisher information matrix given in Appendix III, and \xrightarrow{D} denotes convergence in distribution. Asymptotic confidence intervals (CIs) of a $100 \times [1 - \zeta]\%$ level for μ , α and φ_i , with $i = 1, 2, 3$, can be obtained from the asymptotic normality property given in (23) as

$$\begin{aligned} \text{CI}(\mu, [1 - \zeta] \times 100\%) &= \left[\hat{\mu} - z(1 - \zeta/2)\widehat{\text{SE}}(\hat{\mu}), \hat{\mu} + z(1 - \zeta/2)\widehat{\text{SE}}(\hat{\mu}) \right], \\ \text{CI}(\theta, [1 - \zeta] \times 100\%) &= \left[\exp\left(\hat{\theta}^* - z(1 - \zeta/2)\widehat{\text{SE}}(\hat{\theta}^*)\right), \exp\left(\hat{\theta}^* + z(1 - \zeta/2)\widehat{\text{SE}}(\hat{\theta}^*)\right) \right], \end{aligned}$$

where $\hat{\theta}^*$ is the ML estimate of $\theta^* = \log(\theta)$, with $\theta = \alpha$ or $\theta = \varphi_i$, for $i = 1, 2, 3$, and $\widehat{\text{SE}}(\hat{\theta}^*)$ is the estimated asymptotic standard error (SE) of the ML estimator of θ^* ; see Leiva et al. (2015d).

3.3 Local influence

Let $\ell(\theta)$ be the log-likelihood function for the vector of model parameters $\theta = (\alpha, \mu, \varphi_1, \varphi_2, \varphi_3)^\top$ given in (18), which is referred to as the non-perturbed log-likelihood. Then, let $\omega = (\omega_1, \dots, \omega_n)^\top \in \Omega \subset \mathbb{R}^n$ be an $n \times 1$ perturbation vector, where Ω is an open set of relevant perturbations. Let $\ell(\theta|\omega)$ be the log-likelihood function perturbed by ω , called the perturbed log-likelihood, and $\hat{\theta}_\omega$ be the ML estimate of θ obtained from $\ell(\theta|\omega)$. In addition, let $\omega_0 \in \Omega$ be an $n \times 1$ non-perturbation vector, such that $\ell(\theta|\omega_0) = \ell(\theta)$. Suppose that $\ell(\theta|\omega)$ is twice continuously differentiable in a neighbourhood of $(\hat{\theta}, \omega_0)$. Now comparing the parameter estimates $\hat{\theta}$ and $\hat{\theta}_\omega$ using local influence, we are able to investigate how the inference is affected by the perturbation. Consider the likelihood displacement (LD) given by

$$\text{LD}(\omega) = 2(\ell(\hat{\theta}) - \ell(\hat{\theta}_\omega)), \quad (24)$$

which is used to assess the influence of the perturbation ω . Large values of $\text{LD}(\omega)$ in (24) indicate that $\hat{\theta}$ and $\hat{\theta}_\omega$ differ considerably related to the contours of the non-perturbed log-likelihood function $\ell(\theta)$. The method studies the local behaviour of the influence plot $a(\omega) = (\omega^\top, \text{LD}(\omega))^\top$ around ω_0 . Cook (1987) suggested invigilating the direction of maximum curvature, C_{\max} , of the surface $a(\omega)$. For $\text{LD}(\omega)$ in (24), $C_{\max} = \max_{\|d\|=1} C_d$, where $C_d = 2|d^\top \mathbf{B}d|$, with \mathbf{B} being an $n \times n$ matrix and d a unit-length direction vector. To find C_{\max} and the direction vector d_{\max} , we need to calculate $\mathbf{B} = -\Delta \ddot{\ell}(\hat{\theta})^{-1} \Delta$, where $\ddot{\ell}(\hat{\theta})$ is

the Hessian matrix obtained from (19) evaluated at $\theta = \hat{\theta}$; see details of the Hessian matrix in Appendix II. Here, Δ is a $5 \times n$ matrix obtained from the perturbed log-likelihood function and given by

$$\Delta = \frac{\partial^2 \ell(\theta|\omega)}{\partial \theta \partial \omega^\top} = \begin{pmatrix} \frac{\partial^2 \ell(\theta|\omega)}{\partial \alpha \partial \omega^\top} \\ \frac{\partial^2 \ell(\theta|\omega)}{\partial \mu \partial \omega^\top} \\ \frac{\partial^2 \ell(\theta|\omega)}{\partial \theta \partial \omega^\top} \end{pmatrix} = \begin{pmatrix} \Delta_\alpha \\ \Delta_\mu \\ \Delta_\varphi \end{pmatrix}. \quad (25)$$

Note that (25) must be evaluated at $\theta = \hat{\theta}$ and $\omega = \omega_0$. Then, \mathbf{d}_{\max} is a unit-length eigenvector associated with the maximum absolute eigenvalue C_{\max} of \mathbf{B} . A large absolute value of any element of \mathbf{d}_{\max} reveals that the case is likely to be influential. Other important directions correspond to the canonical basis vectors $\mathbf{d} = \mathbf{e}_i$, for $i = 1, \dots, n$, where \mathbf{e}_i is a vector in \mathbb{R}^n with a one (1) in its i th position and zeros (0) in the other positions. In this case, the curvature is given by $C_i = 2|b_{ii}|$, where b_{ii} is the (i, i) element of the matrix \mathbf{B} , for $i = 1, \dots, n$. The plot of $C_i = C_{\mathbf{d}_i}$ versus the index i can also be used to identify influential cases. We use index plots of C_i and $|d_{\max_i}|$ as diagnostic measures of local influence in Section 5. Although there is no consensus about a benchmark to determine an influential case, we use a value analogous to that proposed by Zhu and Lee (2001), which indicates the case i as influential if $C_i > \bar{C} + 2SE(C)$, for $i = 1, \dots, n$, where \bar{C} and $SE(C)$ denote, respectively, the mean normal curvature and the corresponding sample SE. Similarly, in the index plot of $|d_{\max_i}|$, the case i is indicated as influential if $|d_{\max_i}| > \overline{|d_{\max}|} + 2SE(|d_{\max}|)$, for $i = 1, \dots, n$, where $\overline{|d_{\max}|}$ and $SE(|d_{\max}|)$ denote, respectively, the mean of the elements in absolute value of the vector \mathbf{d}_{\max} and the corresponding sample SE. We use these benchmark values in the simulation study and in the agricultural application considered in Sections 4 and 5, respectively.

3.4 Selection of the appropriate perturbation scheme

Suppose that the perturbation of the response variable is of the form $\mathbf{Y}_\omega(\mathbf{s}) = \mathbf{Y}(\mathbf{s}) + \mathbf{A}\omega$, where \mathbf{A} is a symmetric and non-singular matrix. Hence, $\mathbf{Y}_{\omega_0}(\mathbf{s}) = \mathbf{Y}(\mathbf{s})$ and $Y_\omega(s_i) = Y(s_i) + \mathbf{a}_i \omega$, where \mathbf{a}_i is the i th row of the matrix \mathbf{A} . In this case, the perturbed log-likelihood function is given by

$$\ell(\theta|\omega) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log(|\Sigma|) - n \log(\alpha) - \frac{2}{\alpha^2} \mathbf{V}_\omega^\top \Sigma^{-1} \mathbf{V}_\omega + \sum_{i=1}^n \log \left(\cosh \left(\frac{y_i + \mathbf{a}_i \omega - \mu}{2} \right) \right), \quad (26)$$

where $\mathbf{V}_\omega = (V_{\omega_1}, \dots, V_{\omega_n})^\top$, with $V_{\omega_i} = \sinh((y_i + \mathbf{a}_i \omega - \mu)/2)$, for $i = 1, \dots, n$. Then, the corresponding score vector obtained from (26) is given by (see details in Appendix IV)

$$\mathbf{U}(\omega) = \frac{\partial \ell(\theta|\omega)}{\partial \omega} = -\frac{2}{\alpha^2} \mathbf{A} \Sigma^{-1} \mathbf{V}_\omega + \frac{1}{2} \mathbf{A} \mathbf{V}_\omega. \quad (27)$$

Consider the variance of the score vector given in (27) as a function of the perturbation vector ω , that is, $\mathbf{G}(\omega) = \text{Var}(\mathbf{U}(\omega)) = \text{E}(\mathbf{U}(\omega) \mathbf{U}^\top(\omega))$, recalling that $\text{E}(\mathbf{U}(\omega)) = \mathbf{0}$. For the BS spatial log-linear model, we have that (again see details in Appendix IV)

$$\mathbf{G}(\omega) = \mathbf{A} \left(\frac{1}{\alpha} \Sigma^{-\frac{1}{2}} - \frac{\alpha}{4} \Sigma^{\frac{1}{2}} \right)^2 \mathbf{A}. \quad (28)$$

According to Zhu et al. (2007), the perturbation ω is appropriate if and only if $\mathbf{G}(\omega_0) = c \mathbf{I}_n$, for $c > 0$, with $\mathbf{G}(\cdot)$ given in (28). In general, for an arbitrary symmetric and non-singular matrix \mathbf{A} , $\mathbf{G}(\omega) \neq c \mathbf{I}_n$. Instead, for the perturbation to be appropriate, \mathbf{A} must be found using the condition

$$\mathbf{A} \left(\frac{1}{\alpha} \Sigma^{-\frac{1}{2}} - \frac{\alpha}{4} \Sigma^{\frac{1}{2}} \right)^2 \mathbf{A} = c \mathbf{I}_n, \quad (29)$$

for some value of $c > 0$. Given that c is non-negative, considering $c = 1$ immediately shows that $\mathbf{A} = ((1/\alpha) \Sigma^{-\frac{1}{2}} - (\alpha/4) \Sigma^{\frac{1}{2}})^{-1}$ satisfies the condition given in (29). Thus, for the BS spatial log-linear model, an appropriate perturbation scheme for the response variable is given by

$$\mathbf{Y}_\omega(\mathbf{s}) = \mathbf{Y}(\mathbf{s}) + \left(\frac{1}{\alpha} \Sigma^{-\frac{1}{2}} - \frac{\alpha}{4} \Sigma^{\frac{1}{2}} \right)^{-1} \omega. \quad (30)$$

Now consider the perturbation matrix Δ defined in (25), the perturbation given in (30), and

$$\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} = -\frac{2}{\alpha^2} \mathbf{V}_\omega^\top \boldsymbol{\Sigma}^{-1} \mathbf{A} + \frac{1}{2} \mathbf{V}_\omega^\top \mathbf{A} = \mathbf{V}_\omega^\top \left(-\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} + \frac{1}{2} \mathbf{I}_n \right) \mathbf{A}. \quad (31)$$

Then, the elements of Δ obtained from (31) are expressed as

$$\Delta_\mu = \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \mu \partial \boldsymbol{\omega}^\top} = \mathbf{1}^\top \left(\frac{1}{\alpha^2} \boldsymbol{\Sigma}^{-1} - \frac{1}{4} \mathbf{I}_n \right) \mathbf{A}, \quad (32)$$

$$\begin{aligned} \Delta_\alpha &= \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial \boldsymbol{\omega}^\top} = \mathbf{D}^\top \left(-\frac{1}{\alpha^2} \boldsymbol{\Sigma}^{-1} + \frac{1}{4} \mathbf{I}_n \right) \mathbf{A}, \\ &\quad + \mathbf{V}_\omega^\top \left(\frac{4}{\alpha^3} \boldsymbol{\Sigma}^{-1} \mathbf{A} + \left(-\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} + \frac{1}{2} \mathbf{I}_n \right) \mathbf{A} \left(\frac{1}{\alpha^2} \boldsymbol{\Sigma}^{-\frac{1}{2}} + \frac{1}{4} \boldsymbol{\Sigma}^{\frac{1}{2}} \right) \mathbf{A} \right), \end{aligned} \quad (33)$$

$$\Delta_\varphi = \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial^2 \varphi \partial \boldsymbol{\omega}^\top} = \left(\frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_1 \partial \boldsymbol{\omega}^\top}, \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_2 \partial \boldsymbol{\omega}^\top}, \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_3 \partial \boldsymbol{\omega}^\top} \right)^\top, \quad (34)$$

where $\mathbf{D} = (d_1, \dots, d_n)^\top$, with $d_j = \mathbf{l}_i \boldsymbol{\omega}$, for $i = 1, \dots, n$, and \mathbf{l}_i being the i th row of the matrix $\mathbf{A} \left(\frac{1}{\alpha^2} \boldsymbol{\Sigma}^{-\frac{1}{2}} + \frac{1}{4} \boldsymbol{\Sigma}^{\frac{1}{2}} \right) \mathbf{A}$. The elements of Δ_φ given in (34) are defined as

$$\begin{aligned} \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \varphi_i \partial \boldsymbol{\omega}^\top} &= \mathbf{M}^\top \left(-\frac{1}{\alpha^2} \boldsymbol{\Sigma}^{-1} + \frac{1}{4} \mathbf{I}_n \right) \mathbf{A} + \mathbf{V}_\omega^\top \left(\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_i} \boldsymbol{\Sigma}^{-1} \right) \mathbf{A} \\ &\quad + \mathbf{1}^\top \left(-\frac{2}{\alpha^2} \boldsymbol{\Sigma}^{-1} + \frac{1}{2} \mathbf{I}_n \right) \mathbf{A} \left(\frac{1}{\alpha} \boldsymbol{\Sigma}^{-\frac{1}{2}} \frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_i} \boldsymbol{\Sigma}^{-\frac{1}{2}} + \frac{\alpha}{4} \frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_i} \right) \mathbf{A}, \end{aligned}$$

where $\mathbf{M} = (m_1, \dots, m_n)^\top$, with $m_i = \mathbf{l}_i \boldsymbol{\omega}$, for $i = 1, \dots, n$, and \mathbf{l}_i being the i th row of the matrix

$$\mathbf{A} \left(-\frac{1}{\alpha} \boldsymbol{\Sigma}^{-\frac{1}{2}} \frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_i} \boldsymbol{\Sigma}^{-\frac{1}{2}} + \frac{\alpha}{4} \frac{\partial \boldsymbol{\Sigma}^{\frac{1}{2}}}{\partial \varphi_i} \right) \mathbf{A}.$$

Detailed algebra of the above expressions is presented in Appendix V.

4 Simulation study

In this section, we conduct two MC simulation studies based on n sampling points $\mathbf{y} = (y_1, \dots, y_n)^\top$ generated from a log-BS distribution with shape parameter $\alpha \in \{0.3, 1.0\}$, scale parameter $\mu = 2$ and spatial structure described by the Matérn model with $\delta = 0.5$, $\varphi_1 \in \{0.5, 1.0\}$, $\varphi_2 \in \{0.5, 1.0\}$ and $\varphi_3 \in \{1.0, 1.5\}$. We consider a regular grid with a minimum distance between points of one unit. We evaluate the performance of the estimators of the model parameters for small and large sample sizes. Then, we assess the performance of the proposed diagnostic methodology on the detection of influential cases.

4.1 Study I: ML estimation

To assess the efficiency of the estimator $\widehat{\theta}_i$ empirically, we use the absolute relative bias (ARB) and the root mean squared error (RMSE) defined as the square root of the mean squared error given by

$$\text{ARB}(\widehat{\theta}_i) = \left| \frac{\bar{\theta}_i - \theta_i}{\theta_i} \right| \times 100, \quad \text{MSE}(\widehat{\theta}_i) = \frac{1}{p} \sum_{j=1}^p (\widehat{\theta}_{ij} - \theta_i)^2,$$

where $\bar{\theta}_i = (1/p) \sum_{j=1}^p \widehat{\theta}_{ij}$, with $\widehat{\theta}_{ij}$ being the ML estimate of $\theta_i \in \{\alpha, \mu, \varphi_1, \varphi_2, \varphi_3\}$ for the j th MC replication. Note that $\widehat{\boldsymbol{\theta}} = (\widehat{\alpha}, \widehat{\mu}, \widehat{\varphi}_1, \widehat{\varphi}_2, \widehat{\varphi}_3)^\top$ is the ML estimate of $\boldsymbol{\theta} = (\alpha, \mu, \varphi_1, \varphi_2, \varphi_3)^\top$. We consider $p = 500$ MC replications in each case. Table 1 displays the empirical ARBs and RMSEs of the corresponding ML estimators. Note that, for $n = 100$, all the ARBs are small. In particular, the ARBs of the ML estimators of the parameter α are less than 10% in fifteen of sixteen cases studied, further in 50% of these cases the ARBs are less than 2%. For the ML estimator of the parameter μ , the ARBs are less than

1% in practically all cases. For the ML estimators of the parameters describing the spatial structure, the ARBs are greater than 10% in a few cases with a maximum of 13%. However, for $n = 36$, this picture is quite different, because only the ARBs of the ML estimator of the parameter μ remain low. For the other parameters, in various scenarios, the ARBs are high. In general, the RMSEs are small for the estimators of α and μ , but are moderate for the spatial parameters. As expected the values increase for $n = 36$ compared to those for $n = 100$. Such results show the sensitivity of the ML estimators of the BS spatial model to small samples.

4.2 Study II: influence diagnostics

We evaluate the performance of the corresponding diagnostic tools in detecting influential cases. In order to carry out this evaluation, we generate two data sets using a response variable as given in (18), with scale matrix structure described by Matérn family model, considering the shape parameter $\delta \in \{0.5, 1.0\}$ and $n = 100$. After generating each data set $\mathbf{y} = (y_1, \dots, y_n)^\top$, we contaminate its maximum value to generate an outlier. This contamination is similar to that used in Ortega et al. (2003) to study the influence of a perturbation in an explanatory variable for generalized log-gamma fixed effect models. Specifically, we consider the contamination

$$y_{\max}^* = y_{\max} + 10\sqrt{\mathbf{y}^\top \mathbf{y}}. \quad (35)$$

For the first data set ($\delta = 0.5, n = 100$), the contaminated case according to (35) results to be case #59. Fig. 2(left) shows the box-plot for this data set, which identifies case #59 as an outlier. Fig. 2(right) displays the corresponding model map divided into quartiles, from which it is possible to identify where outliers are located. After the contamination, the estimates of the parameters (obtained using the BFGS method) with their estimated asymptotic SEs (in parenthesis) are: $\hat{\alpha} = 4.828(3.414 \times 10^{-1})$, $\hat{\mu} = 71.188(1.309 \times 10)$, $\hat{\varphi}_1 = 480.101(8.989 \times 10^2)$, $\hat{\varphi}_2 = 200.090(8.999 \times 10^2)$ and $\hat{\varphi}_3 = 0.346(1.472 \times 10^{-3})$, resulting in an estimated spatial dependence radius $\hat{a} = 1.038(4.416 \times 10^{-3})$. Fig. 3 displays the index plot of C_i^t (left) and $|d_{\max}|$ (right) for the spatial structure described by the Matérn family model with $\delta = 0.5$, considering the contaminated case according to (35). This study of local influence shows that case #59 is detected as potential influential.

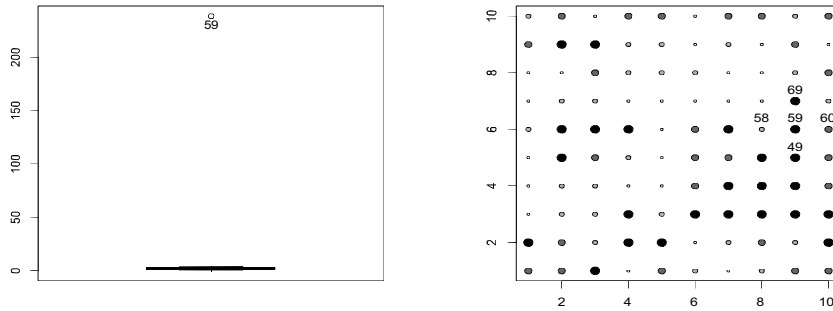


Fig. 2 Box-plot (left) and model map (right) of the simulated data for Matérn model with $\delta = 0.5$.

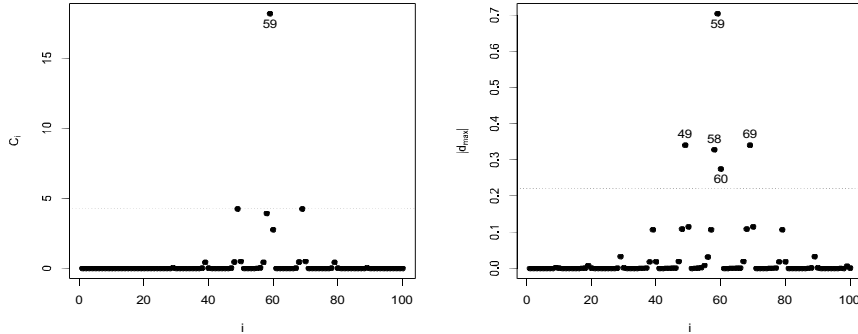


Fig. 3 Index plots using (left) C_i versus i and (right) $|d_{\max_i}|$ versus i for the simulated data from the Matérn model with $\delta = 0.5$.

Similarly to the first data set, for the second data set ($\delta = 1.0, n = 100$), the contaminated case according to (35) results to be case #59. Fig. 4(left) shows the box-plot for this second data set, whereas Fig. 4(right) displays the corresponding model map divided by quartiles. In this second data set, after the contamination, the estimates of the parameters (obtained with the BFGS method) with their SEs (in parenthesis) are: $\hat{\alpha} = 5.004(3.538 \times 10^{-1})$, $\hat{\mu} = 71.250(1.253 \times 10)$, $\hat{\varphi}_1 = 400.118(1.403 \times 10^3)$, $\hat{\varphi}_2 = 200.097(1.404 \times 10^3)$ and $\hat{\varphi}_3 = 0.222(3.937 \times 10^{-4})$, presenting an estimated maximum distance of spatial dependence $\hat{a} = 0.888(1.575 \times 10^{-3})$. Fig. 5 displays the index plots of influence local which, such as in the first data set, indicate case #59 as potential influential.

Therefore, the proposed diagnostic methodology seems to be effective for detecting outlying and potential influential cases.

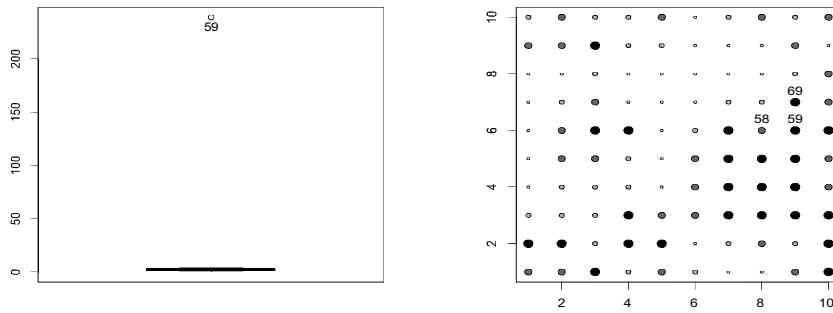


Fig. 4 Box-plot (left) and model map (right) of the simulated data for Matérn model with $\delta = 1$.

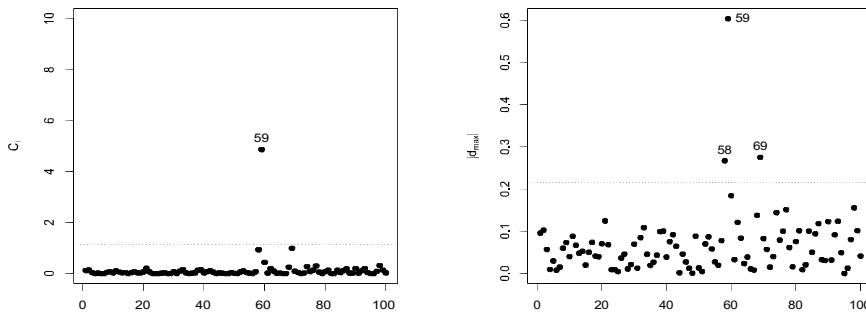


Fig. 5 Index plots using (left) C_i versus i and (right) $|d_{\max_i}|$ versus i for the simulated data for Matérn model with $\delta = 0.5$.

Table 1 ARB (in %) and RMSE of the ML estimator of the indicated parameter and n from simulated data with the BS spatial model.

Parameter	$n = 36$			$n = 100$			$n = 36$			$n = 100$		
	Value	ARB	RMSE	Value	ARB	RMSE	Value	ARB	RMSE	Value	ARB	RMSE
α	1.00	8.501	0.211	0.30	38.932	0.175	1.00	8.683	0.170	0.30	6.526	0.058
μ	2.00	0.377	0.425	2.00	1.010	0.339	2.00	0.48	0.245	2.00	0.102	0.070
φ_1	0.50	15.204	0.628	0.50	0.134	0.474	0.50	6.446	0.505	0.50	1.624	0.586
φ_2	1.00	2.879	0.549	1.00	12.064	0.546	1.00	3.702	0.492	1.00	9.129	0.200
φ_3	1.00	10.631	1.549	1.00	3.283	1.591	1.00	8.288	0.922	1.00	10.102	0.731
α	1.00	0.036	0.131	0.30	16.731	0.093	1.00	0.403	0.121	0.30	7.886	0.083
μ	2.00	0.782	0.172	2.00	0.310	0.076	2.00	0.298	0.310	2.00	0.124	0.051
φ_1	0.50	16.061	0.407	0.50	37.419	0.362	0.50	12.548	0.424	0.50	9.748	0.448
φ_2	0.50	20.495	0.357	0.50	9.631	0.285	0.50	4.074	0.579	0.50	8.521	0.418
φ_3	1.00	19.786	1.406	1.00	31.122	0.598	1.00	12.092	1.000	1.00	3.368	0.840
α	1.00	5.018	0.180	0.30	0.713	0.052	1.00	1.543	0.122	0.30	0.969	0.049
μ	2.00	0.569	0.196	2.00	0.128	0.088	2.00	0.094	0.308	2.00	0.056	0.054
φ_1	1.00	23.470	0.559	1.00	30.419	0.598	1.00	4.416	0.444	1.00	7.977	0.296
φ_2	0.50	69.708	0.812	0.50	72.644	1.077	0.50	3.011	0.712	0.50	12.575	0.594
φ_3	1.00	23.311	1.479	1.00	27.248	0.508	1.00	3.736	0.966	1.00	10.600	0.838
α	1.00	11.580	0.149	0.30	2.084	0.044	1.00	0.675	0.116	0.30	1.361	0.041
μ	2.00	1.073	0.347	2.00	0.044	0.095	2.00	0.261	0.219	2.00	0.143	0.065
φ_1	0.50	13.250	0.417	0.50	10.198	0.392	0.50	3.457	0.380	0.50	1.860	0.226
φ_2	0.50	47.626	0.420	0.50	7.230	0.409	0.50	9.557	0.600	0.50	2.624	0.265
φ_3	1.50	4.095	1.365	1.50	28.613	0.875	1.50	7.062	0.999	1.50	7.850	1.048
α	1.00	9.409	0.182	0.30	34.818	0.127	1.00	6.615	0.131	0.30	9.041	0.050
μ	2.00	0.832	0.424	2.00	0.135	0.112	2.00	0.351	0.262	2.00	0.141	0.071
φ_1	1.00	22.382	0.688	1.00	58.626	0.704	1.00	9.877	0.600	1.00	6.292	0.609
φ_2	1.00	16.801	0.698	1.00	32.248	0.494	1.00	11.614	0.649	1.00	10.743	0.399
φ_3	1.00	29.993	1.441	1.00	25.988	0.722	1.00	6.278	1.088	1.00	6.915	0.854
α	1.00	3.315	0.213	0.30	7.271	0.084	1.00	1.731	0.136	0.30	3.474	0.068
μ	2.00	1.291	0.464	2.00	0.208	0.131	2.00	1.147	0.312	2.00	0.029	0.088
φ_1	0.50	13.031	0.513	0.50	0.375	0.799	0.50	6.876	0.406	0.50	4.700	0.419
φ_2	1.00	8.044	0.385	1.00	11.701	0.383	1.00	1.374	0.477	1.00	7.779	0.324
φ_3	1.50	11.971	1.455	1.50	26.713	0.862	1.50	11.409	1.010	1.50	1.422	1.048
α	1.00	2.479	0.122	0.30	36.118	0.150	1.00	1.645	0.096	0.30	0.124	0.047
μ	2.00	1.012	0.254	2.00	0.016	0.110	2.00	0.310	0.339	2.00	0.101	0.067
φ_1	1.00	12.815	0.434	1.00	49.457	0.593	1.00	2.893	0.401	1.00	8.174	0.312
φ_2	0.50	44.539	0.610	0.50	51.302	0.840	0.50	9.810	0.648	0.50	7.033	0.827
φ_3	1.50	19.714	1.927	1.50	10.903	0.570	1.50	2.042	1.105	1.50	0.266	1.052
α	1.00	7.312	0.159	0.30	21.1904	0.094	1.00	6.414	0.176	0.30	10.473	0.051
μ	2.00	0.318	0.456	2.00	0.335	0.136	2.00	0.622	0.331	2.00	0.207	0.093
φ_1	1.00	9.491	0.691	1.00	42.096	0.615	1.00	1.245	0.466	1.00	3.463	0.560
φ_2	1.00	8.964	0.528	1.00	28.573	1.121	1.00	1.951	0.910	1.00	9.094	0.328
φ_3	1.50	24.874	1.753	1.50	27.398	0.883	1.50	10.009	1.138	1.50	13.259	0.928

5 Numerical applications

5.1 Description of the experimental area and study variables

The experimental study was conducted during the crop year of 2012/2013 in a 167.35 ha commercial area of grain production of Cascavel, a town in the western region of the state of Paraná, Brazil. The area has a geographic location of, approximately, 24.95° south/53.57° west, with an average altitude of 650 m. The soil is classified as *Red Haplortox Oxisol*, with a clay texture; see EMBRAPA (2009). The climate of the region is humid subtropical (*Koepfen climatic type: Cfa*) with an average annual temperature of 21°C. Soil sampling locations were defined at points forming a regular square lattice with an inter-point distance of about 140m. In addition, extra locations were chosen at random across the region. This leads to a total of $n = 102$ locations, which were geo-referenced in UTM coordinates using a GPS device. To analyse the chemical properties, four soil samples were collected at depths from 0.0 to 0.2m and mixed to produce a single representative sample at each location. The mixed soil samples were analysed by the laboratory of the Cooperativa Central de Desenvolvimento Tecnológico e Econômico Ltda. (Coodetec, Brazil) to determine the concentration of macro and micro-nutrients. Phosphorus concentration was chosen as the study variable, because it has a positive effect on plant growth and nutrition. Furthermore, its spatial variability is extremely important for agricultural management. According to the classification given by COAMO/COODETEC

(2001), the phosphorus concentration in clayey soil is considered as follows. For soybean planting, it is: (i) low for levels less than 3 mg/dm³; (ii) medium for levels between 3.1 and 6 mg/dm³; (iii) high for levels between 6.1 and 9 mg/dm³; and (iv) very high for levels greater than 9 mg/dm³. For corn planting, the phosphorus concentration is: (i) low for levels less than 2 mg/dm³; (ii) medium for levels between 2.1 and 4.5 mg/dm³; (iii) high for levels between 4.6 and 11 mg/dm³; and (iv) very high for levels greater than 11 mg/dm³.

5.2 Exploratory and spatial dependence analysis

The exploratory data analysis (EDA) for the phosphorus concentration is divided in two: non-spatial and spatial. From the non-spatial point of view, the sample mean of the phosphorus concentration in the soil of the area under study is 18.11 mg/dm³, whereas the corresponding sample coefficients of variation (CV), skewness (CS) and kurtosis (CK) are CV = 0.41 (41%), CS = 1.787 and CK = 5.104. These descriptive statistics indicate a reasonable degree of homogeneity around the mean, a positive skewness and a high kurtosis level, such as visualized in the box-plot of Fig. 6(left). From this box-plot, note that four outliers, identified as cases #32, #53, #57 and #59, are detected. The circled points in Fig. 6(right) identify these outlying cases, which are located at the lower part of the studied region. The non-spatial EDA supports the use of the BS distribution. From the spatial point of view, a standard analysis of sample variograms (omitted here), using the directions 0°, 45°, 90° and 135°, shows that the directional variograms have similar behaviour until a distance of about 900 m is attained. Therefore, we can assume that there is isotropy up until that distance.

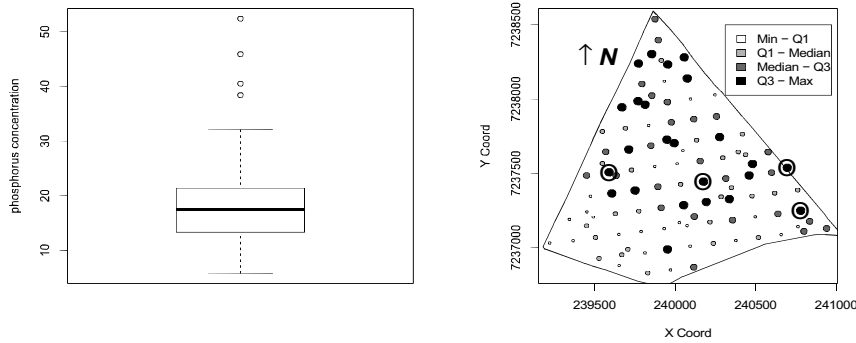


Fig. 6 Box-plot (left) and model map (right) of the phosphorus concentration data.

5.3 Parameter estimation and Kriging

To choose the best model for describing the spatial dependence structure of the phosphorus concentration, the cross-validation criterion and the maximum value of the log-likelihood function are considered. Note that the parameter δ corresponds to the order of the variogram model in the Matérn family, which it is not estimated to avoid identifiability problems in the estimation of covariance matrix parameters. Thus, several variogram models based on the Matérn family are fitted and then the model with the smallest cross-validation value is chosen. Once δ is determined, the best Matérn model with parameter δ is used to estimate θ by the ML method using a profile likelihood approach. Therefore, using this criterion, the best model corresponds to the Matérn family with parameter $\delta = 2.5$. The estimated BS model and variogram parameters (with estimated asymptotic SE in parenthesis) are $\hat{\alpha} = 0.997(3.521)$, $\hat{\mu} = 2.807(0.082)$, $\hat{\varphi}_1 = 0.134(0.946)$, $\hat{\varphi}_2 = 0.020(0.142)$, $\hat{\varphi}_3 = 177.940(0.0000014)$, and $\hat{a} = 1.053(0.0000083)$; see Table 4. Then, we obtain the fitted spatial map shown in Fig. 8(left) for the soil phosphorus concentration by using the fitted BS spatial model obtained from (16) and the ordinary Kriging interpolation described in (15). Observe that this is obtained from the ML estimates provided above, the fitted scale matrix

$\widehat{\Sigma} = 0.134\mathbf{I}_n + 0.020\widehat{\mathbf{R}}$, with $\widehat{\mathbf{R}}$ given from (11), and $\delta = 2.5$. According to the classification provided by COAMO/COODETEC (2001) (see Subsection 5.1), note that the phosphorus concentration is considered very high for both soybean and corn planting, which is suitable for planting both crops.

5.4 Model selection

We compare the spatial BS and Gaussian models using the Akaike (AIC) and Schwarz Bayesian (BIC) information criteria. These are given by $\text{AIC} = -2\ell(\widehat{\theta}) + 2d$ and $\text{BIC} = -2\ell(\widehat{\theta}) + d\log(n)$, where $\ell(\widehat{\theta})$ is the log-likelihood function for the parameter θ associated with the model evaluated at $\theta = \widehat{\theta}$, d is the dimension of the parameter space, and n the size of the data set. Both criteria are based on a penalized log-likelihood function as the model becomes more complex, that is, with more parameters. Thus, a model whose information criterion has a smaller value is better; see Ferreira et al. (2012), Leiva et al. (2015e) and references therein. Besides the AIC and BIC information criteria, the Bayes factor (BF) can also be used to compare the BS and Gaussian spatial models. The BF, denoted by B_{12} , allows us to compare M1 (model considered as correct) to M2 (model to be contrasted with M1) by $2\log(B_{12}) \approx \text{BIC}_{M_2} - \text{BIC}_{M_1}$, where BIC_{M_i} stands for the BIC associated with the model M_i , for $i = 1, 2$. The BF provides an objective value to quantify the degree of superiority of one model with respect to another. An interpretation of the BF is displayed in the Table 2. Thus, according to Tables 2 and 3, we detect that the BS model is superior to the Gaussian model with a very strong evidence in its favor when the data are not transformed by the logarithm.

Table 2 Interpretation of $2\log(B_{12})$ associated with the BF.

$2\log(B_{12})$	Evidence in favor of M_1
< 0	Negative (M_2 is accepted)
$[0, 2)$	weak
$[2, 6)$	Positive
$[6, 10)$	Strong
> 10	very strong

Table 3

Model	$\ell(\widehat{\theta})$	AIC	BIC	$2\log(B_{12})$
BS (transformed data)	-44.577	99.154	112.279	-
Gaussian (transformed data)	-43.890	95.780	106.280	-
BS	-332.576	675.152	688.276	28.224
Gaussian	-349.000	706.000	716.500	-

5.5 Influence diagnostics

To evaluate the effect of atypical cases on the fitted spatial map shown in Fig. 8(left), we carry out local influence diagnostics. By using a response variable perturbation scheme for the detection of influential cases, plots of C_i versus i and $|d_{\max}|$ versus i are considered. It is clear to see from Fig. 7 that cases #2, #48 and #94 are identified as influential by both techniques. Note that these three cases are different from the outlying cases detected by the box-plot in Fig. 6(left), which indicates the relevance of using the local influence method instead of a simple analysis using the box-plot.

Table 4 shows estimated model parameters and the corresponding estimated SEs in parenthesis, for various spatial models fitted with the complete data set and subsets of it when the influential cases are removed, either individually or jointly. As usual in the local influence method, once one or more influential cases are identified, the cases are removed from the data set to investigate how their removal affects the model selection, the estimation of parameters and the construction of maps. Note that removal of the influential cases causes dramatic changes in the estimated spatial range, \widehat{a} , especially when cases #32 and #48 are removed together. This is due to the considerable change in the estimate $\widehat{\varphi}_3$ and a change in the estimated variogram

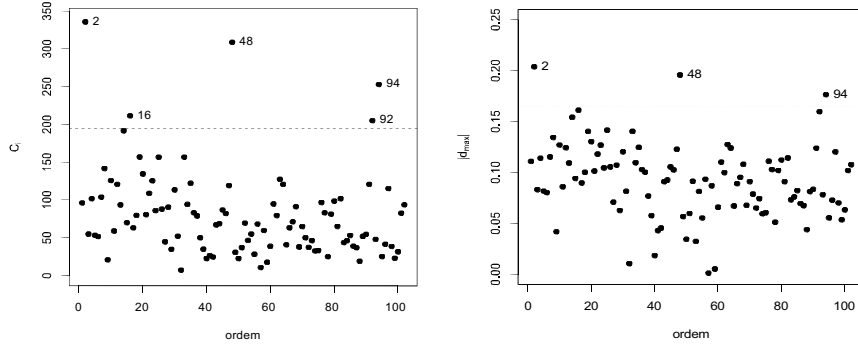


Fig. 7 Index plots using (left) C_i versus i and (right) $|d_{\max}|$ versus i for phosphorus concentration data.

model chosen to describe the spatial variability. Table 4 also shows the p-values for hypotheses of the form $H_0: \theta_j = 0$ versus $H_1: \theta_j \neq 0$, where θ_j is any of the parameters of the vector $\theta = (\alpha, \mu, \varphi_1, \varphi_2, \varphi_3)^\top$. To test H_0 , we can use the test statistic $Z = \hat{\theta}_j / \text{SE}(\hat{\theta}_j)$, where $\hat{\theta}_j$ is the ML estimator of θ_j and $\text{SE}(\hat{\theta}_j)$ the corresponding SE. The asymptotic distribution of Z is known to be normal; see results detailed in (23). From the p-values shown in this table, it follows that, in all models, the hypothesis H_0 is rejected at a 5% significance level for the parameters μ and φ_3 , whereas H_0 is not rejected in the tests for the parameters α , φ_1 and φ_2 . Note that no inferential changes are detected when removing the influential cases. Observe that, in the cases when the parameter space is the set of the positive real numbers, we should restrict H_1 to this set, which was considered accordingly in the calculation of the corresponding p-value. Note also that, for the Matérn model, the spatial dependence radius a is a function of φ_3 , that is, $a = c\varphi_3$, where c is a constant that depends on δ , so that $\hat{a} = c\hat{\varphi}_3$ and then $\text{SE}(\hat{a}) = c\text{SE}(\hat{\varphi}_3)$.

Fig. 8 displays contour maps of the soil phosphorus concentration using the BS spatial model and ordinary Kriging interpolation. The maps were created based on two scenarios: (i) using the complete data set (reference map) and (ii) removing the influential cases #2, #48 and #94, individually and jointly. To construct the maps, we consider five classes of equal size obtained by dividing the range of estimated phosphorus concentration into five equal width intervals. Note that the removal of the cases individually does not change the map significantly. However, the joint removal of the cases causes a dramatic change. A more objective comparison of the maps is carried out using the GA and κ indexes; see Table 4. Thus, according to Anderson et al. (1976), when the influential cases are removed individually, the model maps are similar to the reference map displayed in Fig. 6(right). Nevertheless, this does not occur when the cases are removed jointly. Using the classification of Krippendorff (2004) for the κ index, note the following: (i) the model map created removing the influential case #48 has a high similarity compared to the reference map; (ii) the map created removing cases #2 and #94 has a medium similarity; and (iii) the other maps have a low similarity, which suggests that the cases are only jointly influential.

6 Conclusions and future work

We have proposed a novel spatial log-linear model based on the Birnbbaum-Saunders distribution. This is an alternative to the Gaussian spatial model for describing data with spatial dependency structure and, most importantly, with a positive skew distribution. Maximum likelihood estimates of the model parameters were calculated using an iterative approach. Local influence diagnostics for the new model were derived and corresponding equations for the most appropriate perturbation obtained. We evaluated the performance of the estimation procedure and diagnostic tools by simulation. For large samples, estimation and diagnostics have a good performance. The proposed approach was also used to analyse real-world agricultural engineering data. In this application, influential cases were detected, and their removal caused a considerable change in the spatial dependence radius and on the spatial maps. Importantly, not all these influential cases would have been identified as traditional outliers.

Some possible issues to be addressed in future studies are the following. First, because the Birnbbaum-Saunders distribution is based on the normal distribution, a heavy-tailed version based, for example, on the

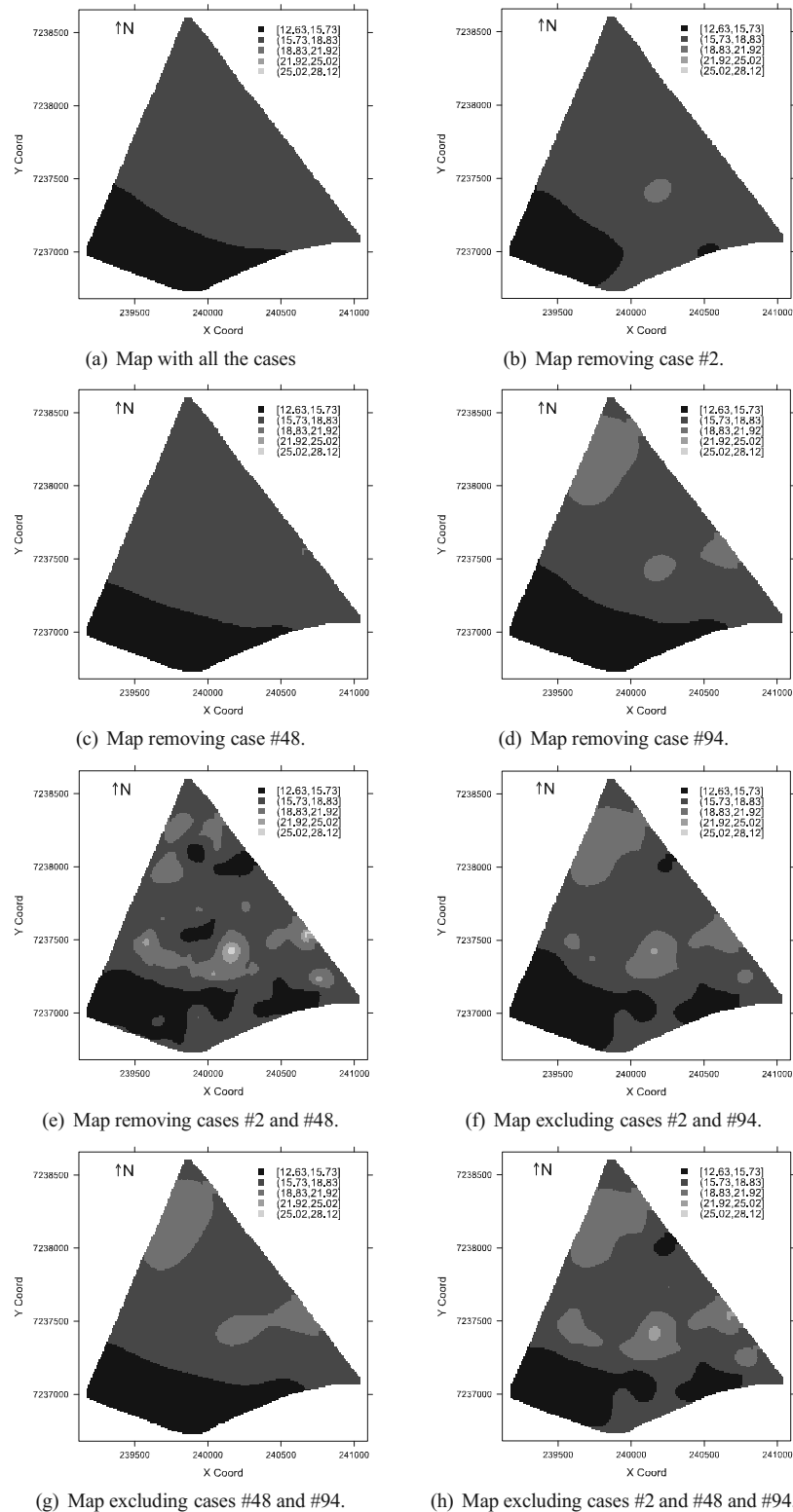


Fig. 8 Shaded contour plots showing the effects of removing the indicated case(s) for phosphorus concentration data.

Table 4 ML estimates of the indicated model parameter with (estimated asymptotic SE in parenthesis) and [p-values in brackets], and values of GA and κ indexes for the phosphorus concentration data set with indicated dropped case(s).

Dropped case(s)	Model	$\hat{\alpha}$	$\hat{\mu}$	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_3$	\hat{a}	GA	κ
None	Matérn $\delta = 2.5$	0.997 (3.521) [0.389]	2.807 (0.082) [< 0.001]	0.134 (0.946) [0.444]	0.020 (0.142) [0.444]	177.940 (0.0000014) [< 0.001]	1053.405 (0.0000083) [< 0.001]	-	-
#2	Matérn $\delta = 2.5$	0.993 (4.097) [0.404]	2.831 (0.059) [< 0.001]	0.125 (1.031) [0.452]	0.016 (0.132) [0.452]	108.655 (0.0000008) [< 0.001]	643.238 (0.0000047) [< 0.001]	0.90	0.72
#48	Matérn $\delta = 2.5$	0.996 (3.417) [0.386]	2.824 (0.073) [< 0.001]	0.125 (0.856) [0.442]	0.019 (0.133) [0.443]	152.374 (0.0000012) [< 0.001]	902.054 (0.0000071) [< 0.001]	0.97	0.92
#94	Matérn $\delta = 1$	0.997 (3.417) [0.374]	2.817 (0.073) [< 0.001]	0.122 (0.856) [0.436]	0.028 (0.133) [0.437]	308.828 (0.0000012) [< 0.001]	1235.312 (0.0009092) [< 0.001]	0.84	0.69
#2, #48	Matérn $\delta = 0.5$ (exponential)	0.991 (6.149) [0.436]	2.845 (0.046) [< 0.001]	0.071 (0.882) [0.468]	0.060 (0.746) [0.469]	81.052 (0.0000884) [< 0.001]	243.156 (0.0002652) [< 0.001]	0.65	0.31
#2, #94	Matérn $\delta = 0.5$ (exponential)	0.995 (3.308) [0.382]	2.840 (0.063) [< 0.001]	0.097 (0.644) [0.440]	0.038 (0.254) [0.441]	182.059 (0.0002974) [< 0.001]	546.177 (0.0008922) [< 0.001]	0.72	0.45
#48, #94	Matérn $\delta = 2.5$	0.994 (2.980) [0.370]	2.834 (0.077) [< 0.001]	0.114 (0.683) [0.434]	0.024 (0.145) [0.435]	144.768 (0.0000010) [< 0.001]	856.737 (0.0000059) [< 0.001]	0.80	0.60
#2, #48, #94	Matérn $\delta = 0.5$ (exponential)	0.982 (3.442) [0.388]	2.855 (0.056) [< 0.001]	0.085 (0.597) [0.444]	0.041 (0.287) [0.443]	143.699 (0.0002051) [< 0.001]	431.097 (0.0006153) [< 0.001]	0.66	0.35

Student-t distribution can be considered thereby reducing the influence of atypical data which can have an adverse effect on spatial maps. Second, explanatory variables may be considered in the spatial modelling, which can help to improve its predictive power. Third, other perturbation schemes could be considered to assess the influence of atypical data. Fourth, we could consider more than one random variable in the spatial modelling by means of multivariate structures for the Birnbaum-Saunders distribution. Work on these four issues is currently under progress and we hope to report some findings in a future paper.

Acknowledgement

The authors thank the Editors and anonymous referees for their constructive comments on an earlier version of the manuscript, which resulted in this improved version. We are grateful to Carolina Brianezi-Melchior, who translated this work into English, from its original Portuguese. This research work was partially supported by CNPq grants from the Brazilian government, and by FONDECYT 1120879 grant from the Chilean government.

References

Anderson, J., Hardy, E., Roach, J., and Witmer, R. (1976). A land use and land cover classification system for use with remote sensor data. Technical Report Paper 964, US Geological Survey Professional, Washington, US.

Assumpção, R., Uribe-Opazo, M., and Galea, M. (2011). Local influence for spatial analysis of soil physical properties and soybean yield using student-t distribution. *Revista Brasileira de Ciência do Solo*, 35:1917–1926.

Assumpção, R., Uribe-Opazo, M., and Galea, M. (2014). Analysis of local influence in geostatistics using student-t distribution. *Journal of Applied Statistics*, 41:2323–2341.

Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew-normal distribution. *Journal of the Royal Statistical Society B*, 61:579–602.

Birnbaum, Z. and Saunders, S. (1969). A new family of life distributions. *Journal of Applied Probability*, 6:319–327.

Borssoi, J., De Bastiani, F., Uribe-Opazo, M., and Galea, M. (2011). Local influence of explanatory variables in Gaussian spatial linear models. *Chilean Journal of Statistics*, 2:29–38.

Cambardella, C., Moorman, T., Novak, J., Parkin, T., Karlen, D., Turco, R., and Konopka, A. (1994). Field-scale variability of soil properties in central Iowa soils. *Soil Science Society of America Journal*, 58:1501–1511.

COAMO/COODETEC (2001). Soil fertility and plant nutrition. Technical report, Cooperativa Agropecuaria Mouraense Ltda./Development Center Technological and Economic Cooperative Ltda. (COAMO/COODETEC), Cascavel, Brazil.

- Cook, R. (1987). Influence assessment. *Journal of Applied Statistics*, 14:117–131.
- Davis, D. (1952). An analysis of some failure data. *Journal of the American Statistical Association*, 47:113–150.
- De Bastiani, F., Cysneiros, A., Uribe-Opazo, M., and Galea, M. (2015). Influence diagnostics in elliptical spatial linear models. *TEST*, 24:322–340.
- Diggle, P. and Ribeiro, P. (2007). *Model-Based Geostatistics*. Springer, New York, US.
- EMBRAPA (2009). Brazilian system of soil classification. Technical report, Brazilian Enterprise for National Agricultural Research/Centre of Soil Research (EMBRAPA/CPI, Rio de Janeiro, Brazil).
- Ferreira, M., Gomes, M., and Leiva, V. (2012). On an extreme value version of the Birnbaum-Saunders distribution. *Revstat Statistical Journal*, 10:181–210.
- Galea, M., Leiva, V., and Paula, G. (2004). Influence diagnostics in log-Birnbaum-Saunders regression models. *Journal of Applied Statistics*, 31:1049–1064.
- Gimenez, P. and Galea, M. (2013). Influence measures on corrected score estimators in functional heteroscedastic measurement error models. *Journal of Multivariate Analysis*, 114:1–15.
- Grzegorzewski, D., Uribe-Opazo, M., De Bastiani, F., and Galea, M. (2013). Local influence when fitting Gaussian spatial linear models: An agriculture application. *Ciência e Investigação Agrária*, 40:235–252.
- Isaaks, E. and Srivastava, R. (1989). *An Introduction to Applied Geostatistics*. Oxford University Press, Oxford, UK.
- Johnson, N., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions*, volume 1. Wiley, New York, US.
- Johnson, N., Kotz, S., and Balakrishnan, N. (1995). *Continuous Univariate Distributions*, volume 2. Wiley, New York, US.
- Kendrick, D. (2002). *Stochastic Control for Economic Models*. McGraw Hill, New York, US.
- Krige, D. (1951). A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, 52:119–139.
- Krippendorff, K. (2004). *Content Analysis: An Introduction to its Methodology*. Sage, Thousand Oaks, US.
- Lange, K. (2001). *Numerical Analysis for Statisticians*. Springer, New York, US.
- Lange, K., Little, J., and Taylor, M. (1989). Robust statistical modeling using the t distribution. *Journal of the American Statistical Association*, 84:881–896.
- Leiva, V., Barros, M., Paula, G., and Sanhueza, A. (2008). Generalized Birnbaum-Saunders distribution applied to air pollutant concentration. *Environmetrics*, 19:235–249.
- Leiva, V., Ferreira, M., Gomes, M., and Lillo, C. (2015a). Extreme value Birnbaum-Saunders regression models applied to environmental data. *Stochastic Environmental Research and Risk Assessment*, in press.
- Leiva, V., Liu, S., Shi, L., and Cysneiros, F. (2015b). Diagnostics in elliptical regression models with stochastic restrictions applied to econometrics. *Journal of Applied Statistics*, in press.
- Leiva, V., Marchant, C., Ruggeri, F., and Saulo, H. (2015c). A criterion for environmental assessment using Birnbaum-Saunders attribute control charts. *Environmetrics*, 26:463–476.
- Leiva, V., Rojas, E., Galea, M., and Sanhueza, A. (2014). Diagnostics in Birnbaum-Saunders accelerated life models with an application to fatigue data. *Applied Stochastic Models in Business and Industry*, 30:115–131.
- Leiva, V., Sanhueza, A., and Angulo, J. M. (2009). A length-biased version of the Birnbaum-Saunders distribution with application in water quality. *Stochastic Environmental Research and Risk Assessment*, 23:299–307.
- Leiva, V., Santos-Neto, M., Cysneiros, F., and Barros, M. (2015d). A methodology for stochastic inventory models based on a zero-adjusted Birnbaum-Saunders distribution. *Applied Stochastic Models in Business and Industry*, in press.
- Leiva, V., Tejo, M., Guiraud, P., Schmachtenberg, O., Orio, P., and Marmolejo, F. (2015e). Modeling neural activity with cumulative damage distributions. *Biological Cybernetics*, 109:421–433.
- Liu, S., Leiva, V., Ma, T., and Welsh, A. (2015). Influence diagnostic analysis in the possibly heteroskedastic linear model with exact restrictions. *Statistical Methods and Applications*, in press.
- Marchant, C., Leiva, V., Cavieres, M., and Sanhueza, A. (2013). Air contaminant statistical distributions with application to PM10 in Santiago, Chile. *Reviews of Environmental Contamination and Toxicology*, 223:1–31.
- Marchant, C., Leiva, V., and Cysneiros, F. (2015). A multivariate log-linear model for Birnbaum-Saunders distributions. *IEEE Transactions on Reliability*, in press.
- Mardia, K. and Marshall, R. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika*, 71:135–146.
- Militino, A., Palacios, M., and Ugarte, M. (2006). Outliers detection in multivariate spatial linear models. *Journal of Statistical Planning and Inference*, 136:125–146.
- Muirhead, R. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York, US.
- Müller, W. and Stehlik, M. (2009). Issues in the optimal design of computer simulation experiments. *Applied Stochastic Models in Business and Industry*, 25:163–177.
- Müller, W. and Stehlik, M. (2010). Compound optimal spatial designs. *Environmetrics*, 21:354–364.
- Nocedal, J. and Wright, S. (1999). *Numerical Optimization*. Springer, New York, US.
- Ortega, E., Bolfarine, H., and Paula, G. (2003). Influence diagnostics in generalized log-gamma regression models. *Computational Statistics and Data Analysis*, 42:165–186.
- Podlaski, R. (2008). Characterization of diameter distribution data in near-natural forests using the Birnbaum-Saunders distribution. *Canadian Journal of Forest Research*, 18:518–527.
- R-Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rieck, J. and Nedelman, J. (1991). A log-linear model for the Birnbaum-Saunders distribution. *Technometrics*, 3:51–60.
- Saulo, H., Leiva, V., Ziegelmann, F., and Marchant, C. (2013). A nonparametric method for estimating asymmetric densities based on skewed Birnbaum-Saunders distributions applied to environmental data. *Stochastic Environmental Research and Risk Assessment*, 27:1479–1491.
- Uribe-Opazo, M., Borssoi, J., and Galea, M. (2012). Influence diagnostics in gaussian spatial linear models. *Journal of Applied Statistics*, 39:615–630.

- Vilca, F., Sanhueza, A., Leiva, V., and Christakos, G. (2010). An extended Birnbbaum-Saunders model and its application in the study of environmental quality in Santiago, Chile. *Stochastic Environmental Research and Risk Assessment*, 24:771–782.
- Villegas, C., Paula, G., and Leiva, V. (2011). Birnbbaum-Saunders mixed models for censored reliability data analysis. *IEEE Transactions on Reliability*, 60:748–758.
- Waller, L. and Gotway, C. (2004). *Applied Spatial Statistics for Public Health Data*. Wiley, New Jersey, US.
- Xia, J., Zeepongsekul, P., and Packer, D. (2011). Spatial and temporal modelling of tourist movements using semi-Markov processes. *Tourism Management*, 51:844–851.
- Zhu, H., Ibrahim, J., Lee, S., and Zhang, H. (2007). Perturbation selection and influence measures in local influence analysis. *The Annals of Statistics*, 35:2565–2588.
- Zhu, H. and Lee, S. (2001). Local influence for incomplete-data models. *Journal of The Royal Statistical Society B*, 63:111–126.

Appendix I: The score vector of the BS log-linear spatial model

For the BS log-linear spatial model, the score vector is defined by

$$\mathbf{U}(\boldsymbol{\theta}) = \left(\frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha}, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \mu}, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\varphi}} \right)^\top = (U(\alpha), U(\mu), U(\varphi_1), U(\varphi_2), U(\varphi_3))^\top.$$

Using the log-likelihood function defined in (20), we have

$$\begin{aligned} U(\alpha) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} = \frac{\partial}{\partial \alpha} \left(-n \log(\alpha) - \frac{2}{\alpha^2} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \mathbf{V} \right) = -\frac{n}{\alpha} + \frac{4}{\alpha^3} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \mathbf{V}, \\ U(\mu) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \mu} = \frac{\partial}{\partial \mu} \left(-\frac{2}{\alpha^2} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \mathbf{V} + \sum_{i=1}^n \log \left(\cosh \left(\frac{y_i - \mu}{2} \right) \right) \right) \\ &= \frac{2}{\alpha^2} \mathbf{U}^\top \boldsymbol{\Sigma}^{-1} \mathbf{V} - \frac{1}{2} \sum_{i=1}^n \tanh \left(\frac{y_i - \mu}{2} \right), \\ U(\varphi_i) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \varphi_i} = \frac{\partial}{\partial \varphi_i} \left(-\frac{1}{2} \log(|\boldsymbol{\Sigma}|) - \frac{2}{\alpha^2} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \mathbf{V} \right) \\ &= -\frac{1}{2} \text{tr} \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_i} \right) + \frac{2}{\alpha^2} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_i} \boldsymbol{\Sigma}^{-1} \mathbf{V}, \quad j = 1, 2, 3, \end{aligned}$$

where $\mathbf{U} = (U_1, \dots, U_n)^\top$, with $U_i = \cosh((y_i - \mu)/2)$, and $\mathbf{V} = (V_1, \dots, V_n)^\top$ such as given in (19) with $V_i = \sinh((y_i - \mu)/2)$, for $i = 1, \dots, n$. Thus, considering that $\boldsymbol{\Sigma} = \varphi_1 \mathbf{I}_n + \varphi_2 \mathbf{R}$, we have

$$\frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_1} = \mathbf{I}_n, \quad \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_2} = \mathbf{R}, \quad \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_3} = \varphi_2 \frac{\partial \mathbf{R}}{\partial \varphi_3}.$$

Therefore,

$$\begin{aligned} U(\varphi_1) &= -\frac{1}{2} \text{tr}(\boldsymbol{\Sigma}^{-1}) + \frac{2}{\alpha^2} \mathbf{V}^\top (\boldsymbol{\Sigma}^{-1})^2 \mathbf{V}, \\ U(\varphi_2) &= -\frac{1}{2} \text{tr}(\boldsymbol{\Sigma}^{-1} \mathbf{R}) + \frac{2}{\alpha^2} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \mathbf{V}, \\ U(\varphi_3) &= -\frac{\varphi_2}{2} \text{tr} \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} \right) + \frac{2\varphi_2}{\alpha^2} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} \mathbf{V}. \end{aligned}$$

Considering the Matérn model to describe the spatial variability given in (11), we have that $\partial \mathbf{R} / \partial \varphi_3 = [\partial r_{ij} / \partial \varphi_3]$, where, for $K'_\delta(u) = \partial K_\delta(u) / \partial u = -(1/2)(K_{\delta-1}(u) + K_{\delta+1}(u))$,

$$\frac{\partial r_{ij}}{\partial \varphi_3} = -\left(\frac{1}{\varphi_3} \right) \left(\delta r_{ij} + \frac{1}{2^{\delta-1} \mu(\delta)} \left(\frac{h_{ij}}{\varphi_3} \right)^{\delta+1} K'_\delta \left(\frac{h_{ij}}{\varphi_3} \right) \right), \quad i \neq j, i, j = 1, \dots, n.$$

Appendix II: The observed information matrix for the BS log-linear spatial model

The observed Fisher information matrix for the BS log-linear spatial model is defined by $-\ddot{\ell}(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$, where $\ddot{\ell}(\boldsymbol{\theta})$ is the Hessian matrix given by

$$\ddot{\ell}(\boldsymbol{\theta}) = \begin{pmatrix} \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha \partial \alpha} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha \partial \mu} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha \partial \boldsymbol{\varphi}^\top} \\ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \mu \partial \alpha} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \mu \partial \mu} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \mu \partial \boldsymbol{\varphi}^\top} \\ \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\varphi} \partial \alpha} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\varphi} \partial \mu} & \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}^\top} \end{pmatrix} = \begin{pmatrix} \ddot{\ell}_{\alpha\alpha} & \ddot{\ell}_{\alpha\mu} & \ddot{\ell}_{\alpha\boldsymbol{\varphi}} \\ \ddot{\ell}_{\mu\alpha} & \ddot{\ell}_{\mu\mu} & \ddot{\ell}_{\mu\boldsymbol{\varphi}} \\ \ddot{\ell}_{\boldsymbol{\varphi}\alpha} & \ddot{\ell}_{\boldsymbol{\varphi}\mu} & \ddot{\ell}_{\boldsymbol{\varphi}\boldsymbol{\varphi}} \end{pmatrix},$$

with, for $\ddot{\ell}_{\alpha\mu} = \ddot{\ell}_{\mu\alpha}$,

$$\begin{aligned}\ddot{\ell}_{\alpha\alpha} &= \frac{\partial^2 \ell(\theta)}{\partial \alpha^2} = \frac{\partial}{\partial \alpha} \left(-\frac{n}{\alpha} + \frac{4}{\alpha^3} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \mathbf{V} \right) = \frac{n}{\alpha^2} - \frac{12}{\alpha^4} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \mathbf{V}, \\ \ddot{\ell}_{\mu\alpha} &= \frac{\partial^2 \ell(\theta)}{\partial \mu \partial \alpha} = \frac{\partial}{\partial \mu} \left(\frac{4}{\alpha^3} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \mathbf{V} \right) = -\frac{4}{\alpha^3} \mathbf{U}^\top \boldsymbol{\Sigma}^{-1} \mathbf{V}, \\ \ddot{\ell}_{\mu\mu} &= \frac{\partial^2 \ell(\theta)}{\partial \mu^2} = \frac{\partial}{\partial \mu} \left(\frac{2}{\alpha^2} \mathbf{U}^\top \boldsymbol{\Sigma}^{-1} \mathbf{V} - \frac{1}{2} \sum_{i=1}^n \tanh \left(\frac{y_i - \mu}{2} \right) \right) \\ &= -\frac{1}{\alpha^2} \left(\mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \mathbf{V} + \mathbf{U}^\top \boldsymbol{\Sigma}^{-1} \mathbf{U} \right) + \frac{1}{4} \sum_{i=1}^n \operatorname{sech}^2 \left(\frac{y_i - \mu}{2} \right).\end{aligned}$$

Note that $\ddot{\ell}_{\varphi\alpha} = \ddot{\ell}_{\alpha\varphi}^\top$ is a 3×1 vector with elements given by

$$\ddot{\ell}_{\alpha\varphi_j} = \frac{\partial^2 \ell(\theta)}{\partial \alpha \partial \varphi_j} = \frac{\partial}{\partial \alpha} \left(-\frac{1}{2} \operatorname{tr} \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \right) + \frac{2}{\alpha^2} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-1} \mathbf{V} \right) = -\frac{4}{\alpha^3} \mathbf{V}^\top \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-1} \right) \mathbf{V},$$

for $j = 1, 2, 3$, where $\partial \boldsymbol{\Sigma} / \partial \varphi_i$ is given in Appendix I. Then, with $\ddot{\ell}_{\alpha\varphi_1} = \ddot{\ell}_{\varphi_1\alpha}$, $\ddot{\ell}_{\alpha\varphi_2} = \ddot{\ell}_{\varphi_2\alpha}$ and $\ddot{\ell}_{\alpha\varphi_3} = \ddot{\ell}_{\varphi_3\alpha}$, we have

$$\ddot{\ell}_{\alpha\varphi_1} = -\frac{4}{\alpha^3} \mathbf{V}^\top (\boldsymbol{\Sigma}^{-1})^2 \mathbf{V}, \ddot{\ell}_{\alpha\varphi_2} = -\frac{4}{\alpha^3} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \mathbf{V}, \ddot{\ell}_{\alpha\varphi_3} = -\frac{4\varphi_2}{\alpha^3} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} \mathbf{V}.$$

Furthermore, $\ddot{\ell}_{\varphi\mu} = \ddot{\ell}_{\mu\varphi}^\top$ is a 3×1 vector with elements

$$\ddot{\ell}_{\mu\varphi_j} = \frac{\partial^2 \ell(\theta)}{\partial \mu \partial \varphi_j} = \frac{\partial}{\partial \mu} \left(-\frac{1}{2} \operatorname{tr} \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \right) + \frac{2}{\alpha^2} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-1} \mathbf{V} \right) = -\frac{2}{\alpha^2} \mathbf{U}^\top \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-1} \right) \mathbf{V},$$

for $j = 1, 2, 3$. Then, with $\ddot{\ell}_{\mu\varphi_1} = \ddot{\ell}_{\varphi_1\mu}$, $\ddot{\ell}_{\mu\varphi_2} = \ddot{\ell}_{\varphi_2\mu}$ and $\ddot{\ell}_{\mu\varphi_3} = \ddot{\ell}_{\varphi_3\mu}$, we have

$$\ddot{\ell}_{\mu\varphi_1} = -\frac{2}{\alpha^2} \mathbf{U}^\top (\boldsymbol{\Sigma}^{-1})^2 \mathbf{V}, \ddot{\ell}_{\mu\varphi_2} = -\frac{2}{\alpha^2} \mathbf{U}^\top \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \mathbf{V}, \ddot{\ell}_{\mu\varphi_3} = -\frac{2\varphi_2}{\alpha^2} \mathbf{U}^\top \boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} \mathbf{V}.$$

Moreover, $\ddot{\ell}_{\varphi\varphi} = (\ddot{\ell}_{\varphi_i\varphi_j})$ is a 3×3 symmetric matrix with elements given by

$$\begin{aligned}\ddot{\ell}_{\varphi_j\varphi_i} &= \frac{\partial^2 \ell(\theta)}{\partial \varphi_j \partial \varphi_i} = -\frac{1}{2} \operatorname{tr} \left(-\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_i} + \boldsymbol{\Sigma}^{-1} \frac{\partial^2 \boldsymbol{\Sigma}}{\partial \varphi_j \partial \varphi_i} \right) + \frac{2}{\alpha^2} \mathbf{V}^\top \left(-\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_i} \boldsymbol{\Sigma}^{-1} \right) \mathbf{V} \\ &\quad + \frac{2}{\alpha^2} \mathbf{V}^\top \left(\boldsymbol{\Sigma}^{-1} \left(\frac{\partial^2 \boldsymbol{\Sigma}}{\partial \varphi_j \partial \varphi_i} \boldsymbol{\Sigma}^{-1} + \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_i} \left(-\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \boldsymbol{\Sigma}^{-1} \right) \right) \right) \mathbf{V}, \quad i, j = 1, 2, 3.\end{aligned}$$

Then, with $\ddot{\ell}_{\varphi_1\varphi_2} = \ddot{\ell}_{\varphi_2\varphi_1}$, $\ddot{\ell}_{\varphi_1\varphi_3} = \ddot{\ell}_{\varphi_3\varphi_1}$ and $\ddot{\ell}_{\varphi_2\varphi_3} = \ddot{\ell}_{\varphi_3\varphi_2}$, we have

$$\begin{aligned}\ddot{\ell}_{\varphi_1\varphi_1} &= \frac{1}{2} \operatorname{tr} \left((\boldsymbol{\Sigma}^{-1})^2 \right) - \frac{2}{\alpha^2} \mathbf{V}^\top \left((\boldsymbol{\Sigma}^{-1})^3 + (\boldsymbol{\Sigma}^{-1})^2 \right) \mathbf{V}, \\ \ddot{\ell}_{\varphi_1\varphi_2} &= \frac{1}{2} \operatorname{tr} \left(\boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \right) - \frac{2}{\alpha^2} \mathbf{V}^\top \left(\boldsymbol{\Sigma}^{-1} \mathbf{R} (\boldsymbol{\Sigma}^{-1})^2 + (\boldsymbol{\Sigma}^{-1})^2 \mathbf{R} \boldsymbol{\Sigma}^{-1} \right) \mathbf{V}, \\ \ddot{\ell}_{\varphi_1\varphi_3} &= \frac{\varphi_2}{2} \operatorname{tr} \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} \right) - \frac{2\varphi_2}{\alpha^2} \mathbf{V}^\top \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} (\boldsymbol{\Sigma}^{-1})^2 + (\boldsymbol{\Sigma}^{-1})^2 \frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} \right) \mathbf{V}, \\ \ddot{\ell}_{\varphi_2\varphi_2} &= \frac{1}{2} \operatorname{tr} \left(\boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \mathbf{R} \right) - \frac{4}{\alpha^2} \mathbf{V}^\top \left(\boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \right) \mathbf{V}, \\ \ddot{\ell}_{\varphi_2\varphi_3} &= -\frac{1}{2} \operatorname{tr} \left(-\varphi_2 \boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} \mathbf{R} + \boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} \right) + \frac{2}{\alpha^2} \mathbf{V}^\top \left(-\varphi_2 \boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} \mathbf{R} \boldsymbol{\Sigma}^{-1} \right) \mathbf{V} \\ &\quad + \frac{2}{\alpha^2} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \left(\frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} + \mathbf{R} \left(-\varphi_2 \boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} \right) \right) \mathbf{V}, \\ \ddot{\ell}_{\varphi_3\varphi_3} &= -\frac{\varphi_2}{2} \operatorname{tr} \left(-\varphi_2 \boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} + \boldsymbol{\Sigma}^{-1} \frac{\partial^2 \mathbf{R}}{\partial \varphi_3^2} \right) - \frac{2\varphi_2^2}{\alpha^2} \mathbf{V}^\top \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} \right) \mathbf{V} \\ &\quad + \frac{2\varphi_2}{\alpha^2} \mathbf{V}^\top \boldsymbol{\Sigma}^{-1} \left(\frac{\partial^2 \mathbf{R}}{\partial \varphi_3^2} \boldsymbol{\Sigma}^{-1} - \varphi_2 \frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} \frac{\partial \mathbf{R}}{\partial \varphi_3} \boldsymbol{\Sigma}^{-1} \right) \mathbf{V}.\end{aligned}$$

Appendix III: Expected information matrix for the BS log-linear spatial model

The expected Fisher information matrix is $\mathbf{I}(\boldsymbol{\theta}) = \mathbb{E}[-\ddot{\ell}(\boldsymbol{\theta})]$. For the BS log-linear spatial model, this matrix is given by

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{pmatrix} I_{\alpha\alpha} & I_{\alpha\mu} & I_{\alpha\varphi} \\ I_{\mu\alpha} & I_{\mu\mu} & I_{\mu\varphi} \\ I_{\varphi\alpha} & I_{\varphi\mu} & I_{\varphi\varphi} \end{pmatrix}.$$

Since the model error $\varepsilon \sim \log\text{-BS}_n(\alpha\mathbf{1}, \mathbf{0}, \boldsymbol{\Sigma})$, we have that $(2/\alpha)\mathbf{V} \sim \text{N}_n(\mathbf{0}, \boldsymbol{\Sigma})$, with \mathbf{V} given in (19). Then (see Muirhead, 1982), $W = (2/\alpha)\mathbf{V}^\top \boldsymbol{\Sigma}^{-1}(2/\alpha)\mathbf{V} \sim \chi_n^2$, where χ_n^2 denotes the chi-squared distribution with n degrees of freedom, and then, $\mathbb{E}[W] = n$. Thus,

$$I_{\alpha\alpha} = \mathbb{E}[-\ddot{\ell}_{\alpha\alpha}] = \mathbb{E}\left[-\frac{n}{\alpha^2} + \frac{12}{\alpha^4}\mathbf{V}^\top \boldsymbol{\Sigma}^{-1}\mathbf{V}\right] = \frac{2n}{\alpha^2}.$$

Moreover, by using $\mathbb{E}[\mathbf{X}^\top \mathbf{A} \mathbf{X}] = (\mathbb{E}[\mathbf{X}])^\top \mathbf{A} (\mathbb{E}[\mathbf{X}]) + \text{tr}(\mathbf{A} \mathbf{C})$, where \mathbf{C} is the covariance matrix of \mathbf{X} (see Kendrick, 2002), $\mathbf{I}_{\varphi\varphi} = (I_{\varphi_i\varphi_j})$ is a symmetric 3×3 matrix with elements

$$I_{\varphi_i\varphi_j} = \mathbb{E}[-\ddot{\ell}_{\varphi_i\varphi_j}] = \frac{1}{2}\text{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_i}\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_j}\right), \quad i, j = 1, 2, 3.$$

In addition, $\mathbf{I}_{\varphi\alpha} = \mathbf{I}_{\alpha\varphi}^\top = (I_{\varphi_1\alpha}, I_{\varphi_2\alpha}, I_{\varphi_3\alpha})^\top$ is a 3×1 vector with elements given by

$$I_{\varphi_i\alpha} = \mathbb{E}[-\ddot{\ell}_{\varphi_i\alpha}] = \frac{1}{\alpha}\mathbb{E}\left[\left(\frac{2}{\alpha}\mathbf{V}\right)^\top \left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_i}\boldsymbol{\Sigma}^{-1}\right) \left(\frac{2}{\alpha}\mathbf{V}\right)\right] = \frac{1}{\alpha}\text{tr}\left(\boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_i}\right), \quad i = 1, 2, 3,$$

where $\text{tr}(\mathbf{A})$ denotes the trace of the matrix \mathbf{A} . To obtain the elements $I_{\alpha\mu}$, $I_{\mu\mu}$ and $I_{\mu\varphi}$, for $(y_i - \mu)/2$ expected to be small enough with $\cosh(\cdot) \approx 1$ and using the expansion in Taylor series for the $\cosh(\cdot)$, in a similar way to that used by Rieck and Nedelman (1991) for the $\sinh(\cdot)$, we have

$$\begin{aligned} I_{\alpha\mu} &= I_{\mu\alpha} = \mathbb{E}[-\ddot{\ell}_{\alpha\mu}] = \mathbb{E}\left[\frac{4}{\alpha^3}\mathbf{U}^\top \boldsymbol{\Sigma}^{-1}\mathbf{V}\right] \approx \mathbb{E}\left[\frac{4}{\alpha^3}\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{V}\right] = 0, \\ I_{\mu\mu} &= \mathbb{E}[-\ddot{\ell}_{\mu\mu}] \approx \frac{1}{4}\mathbb{E}\left[\left(\frac{2}{\alpha}\mathbf{V}\right)^\top \boldsymbol{\Sigma}^{-1}\left(\frac{2}{\alpha}\mathbf{V}\right)\right] + \mathbb{E}\left[\frac{1}{\alpha^2}\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1} - \frac{1}{4}\sum_{i=1}^n \mathbf{1}\right] = \frac{1}{\alpha^2}\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}. \end{aligned}$$

Furthermore, $\mathbf{I}_{\mu\varphi} = \mathbf{I}_{\varphi\mu}^\top = (I_{\varphi_1\mu}, I_{\varphi_2\mu}, I_{\varphi_3\mu})^\top$ is a 3×1 vector with elements given by

$$I_{\mu\varphi_i} = \mathbb{E}[-\ddot{\ell}_{\mu\varphi_i}] \approx \frac{1}{\alpha}\mathbb{E}\left[\left(\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\frac{\partial\boldsymbol{\Sigma}}{\partial\varphi_i}\boldsymbol{\Sigma}^{-1}\right) \left(\frac{2}{\alpha}\mathbf{V}\right)\right] = 0, \quad i = 1, 2, 3.$$

Appendix IV: Score vector $\mathbf{U}(\boldsymbol{\omega})$ and matrix $\mathbf{G}(\boldsymbol{\omega})$

Score vector used in the local influence method is given by $\mathbf{U}(\boldsymbol{\omega}) = (\partial\mathbf{V}_\omega/\partial\boldsymbol{\omega}^\top)^\top (2\boldsymbol{\Sigma}^{-1}\mathbf{V}_\omega) + (1/2)\mathbf{A}\mathbf{T}_\omega$, where $\mathbf{T}_\omega = (\mathbf{T}_{\omega_1}, \dots, \mathbf{T}_{\omega_n})$, with $\mathbf{T}_{\omega_i} = \tanh((y_i + \mathbf{A}_i\boldsymbol{\omega} - \mu)/2)$, and \mathbf{A}_i is the i th row of the matrix \mathbf{A} . For $\cosh((y_i + \mathbf{a}_i\boldsymbol{\omega} - \mu)/2) \approx 1$, with $i = 1, \dots, n$, we get

$$\left(\frac{\partial\mathbf{V}_\omega}{\partial\boldsymbol{\omega}^\top}\right) = \frac{1}{2}\begin{pmatrix} \cosh\left(\frac{y_1 + \mathbf{A}_1\boldsymbol{\omega} - \mu}{2}\right)\sigma_{11} & \cdots & \cosh\left(\frac{y_1 + \mathbf{A}_1\boldsymbol{\omega} - \mu}{2}\right)\sigma_{1n} \\ \vdots & \ddots & \vdots \\ \cosh\left(\frac{y_n + \mathbf{A}_n\boldsymbol{\omega} - \mu}{2}\right)\sigma_{n1} & \cdots & \cosh\left(\frac{y_n + \mathbf{A}_n\boldsymbol{\omega} - \mu}{2}\right)\sigma_{nn} \end{pmatrix} = \frac{1}{2}\mathbf{A},$$

where σ_{ij} is the (i, j) element of the matrix \mathbf{A} , for $i, j = 1, \dots, n$. Furthermore, for $\cosh((y_i + \mathbf{a}_i\boldsymbol{\omega} - \mu)/2) \approx 1$, with $i = 1, \dots, n$, \mathbf{T}_ω can be approximated by \mathbf{V}_ω , from which it follows that $\mathbf{U}(\boldsymbol{\omega}) = (1/2)\mathbf{A}\mathbf{V}_\omega - (2/\alpha^2)\mathbf{A}\boldsymbol{\Sigma}^{-1}\mathbf{V}_\omega$. Thus,

$$\begin{aligned} \mathbf{U}(\boldsymbol{\omega})\mathbf{U}^\top(\boldsymbol{\omega}) &= \left(\frac{1}{2}\mathbf{A}\mathbf{V}_\omega - \frac{2}{\alpha^2}\mathbf{A}\boldsymbol{\Sigma}^{-1}\mathbf{V}_\omega\right) \left(\frac{1}{2}\mathbf{A}\mathbf{V}_\omega - \frac{2}{\alpha^2}\mathbf{A}\boldsymbol{\Sigma}^{-1}\mathbf{V}_\omega\right)^\top \\ &= \frac{\alpha^2}{16}\mathbf{A}\left(\frac{2}{\alpha}\mathbf{V}_\omega\right) \left(\frac{2}{\alpha}\mathbf{V}_\omega\right)^\top \mathbf{A} - \frac{1}{4}\mathbf{A}\left(\frac{2}{\alpha}\mathbf{V}_\omega\right) \left(\frac{2}{\alpha}\mathbf{V}_\omega\right)^\top \boldsymbol{\Sigma}^{-1}\mathbf{A} \\ &\quad - \frac{1}{4}\mathbf{A}\boldsymbol{\Sigma}^{-1}\left(\frac{2}{\alpha}\mathbf{V}_\omega\right) \left(\frac{2}{\alpha}\mathbf{V}_\omega\right)^\top \mathbf{A} + \frac{1}{\alpha^2}\mathbf{A}\boldsymbol{\Sigma}^{-1}\left(\frac{2}{\alpha}\mathbf{V}_\omega\right) \left(\frac{2}{\alpha}\mathbf{V}_\omega\right)^\top \boldsymbol{\Sigma}^{-1}\mathbf{A}. \end{aligned}$$

Therefore, $\mathbf{G}(\boldsymbol{\omega}) = \mathbb{E}[\mathbf{U}(\boldsymbol{\omega})\mathbf{U}^\top(\boldsymbol{\omega})] = \mathbf{A}\left(\frac{\alpha}{4}\boldsymbol{\Sigma}^{(1/2)} - (1/\alpha)\boldsymbol{\Sigma}^{-(1/2)}\right)^2 \mathbf{A}$. To find the appropriate perturbation, according to the methodology proposed by Zhu et al. (2007), it is necessary to find \mathbf{A} , such that $\mathbf{G}(\boldsymbol{\omega}) = c\mathbf{I}_n$, for $c > 0$. Considering $c = 1$, then \mathbf{A} must satisfy $((\alpha/4)\boldsymbol{\Sigma}^{(1/2)} - (1/\alpha)\boldsymbol{\Sigma}^{-(1/2)})^2 = (\mathbf{A}^{-1})^2$. A solution of equation above is $\mathbf{A} = ((\alpha/4)\boldsymbol{\Sigma}^{1/2} - (1/\alpha)\boldsymbol{\Sigma}^{-1/2})^{-1}$. Then, $\tilde{\boldsymbol{\omega}} = ((\alpha/4)\boldsymbol{\Sigma}^{1/2} - (1/\alpha)\boldsymbol{\Sigma}^{-1/2})^{-1}\boldsymbol{\omega}$ is an appropriate perturbation for the BS log-linear spatial model.

Appendix V: The perturbation matrix

The perturbation matrix for the BS log-linear spatial model obtained from (20) is given by

$$\Delta_\mu = \frac{\partial \ell(\theta|\omega)}{\partial \mu \partial \omega^\top} = \frac{\partial \mathbf{V}_\omega^\top}{\partial \mu} \left(\frac{-2}{\alpha^2} \Sigma^{-1} + \frac{1}{2} \mathbf{I}_n \right) \mathbf{A} = -\frac{1}{2} \mathbf{U}_\omega^\top \left(\frac{-2}{\alpha^2} \Sigma^{-1} + \frac{1}{2} \mathbf{I}_n \right) \mathbf{A},$$

where $\mathbf{U}_\omega = (\mathbf{U}_{\omega_1}, \dots, \mathbf{U}_{\omega_n})$, with $\mathbf{U}_{\omega_i} = \cosh((y_i + \mathbf{a}_i \omega - \mu)/2)$, for $i = 1, \dots, n$. Thus, the results presented in (32), (33) and (34) are obtained, for $\cosh((y_i + \mathbf{a}_i \omega - \mu)/2) \approx 1$, as

$$\begin{aligned} \Delta_\mu &= \frac{\partial \ell(\theta|\omega)}{\partial \mu \partial \omega^\top} \approx \mathbf{1}^\top \left(\frac{1}{\alpha^2} \Sigma^{-1} - \frac{1}{4} \mathbf{I}_n \right) \mathbf{A}, \\ \Delta_\alpha &= \frac{\partial \ell(\theta|\omega)}{\partial \alpha \partial \omega^\top} = \frac{\partial \mathbf{V}_\omega^\top}{\partial \alpha} \left(-\frac{2}{\alpha^2} \Sigma^{-1} + \frac{1}{2} \mathbf{I}_n \right) \mathbf{A} + \mathbf{V}_\omega^\top \frac{\partial}{\partial \alpha} \left[\left(\frac{-2}{\alpha^2} \Sigma^{-1} + \frac{1}{2} \mathbf{I}_n \right) \mathbf{A} \right] \\ &= \mathbf{D}^\top \left(\frac{-1}{\alpha^2} \Sigma^{-1} + \frac{1}{4} \mathbf{I}_n \right) \mathbf{A} + \left(\frac{4}{\alpha^3} \Sigma^{-1} \mathbf{A} + \left(-\frac{2}{\alpha^2} \Sigma^{-1} + \frac{1}{2} \mathbf{I}_n \right) \mathbf{A} \left(\frac{1}{\alpha^2} \Sigma^{-\frac{1}{2}} + \frac{1}{4} \Sigma^{\frac{1}{2}} \right) \mathbf{A} \right), \end{aligned}$$

where $\mathbf{D} = (D_1, \dots, D_n)^\top$, with $D_i = \mathbf{l}_i \omega$ and \mathbf{l}_i being the i th row of the matrix

$$\mathbf{L} = \mathbf{A} \left(\frac{1}{\alpha^2} \Sigma^{-\frac{1}{2}} + \frac{1}{4} \Sigma^{\frac{1}{2}} \right) \mathbf{A}.$$

In addition,

$$\begin{aligned} \Delta_{\varphi_i} &= \frac{\partial \ell(\theta|\omega)}{\partial \varphi_i \partial \omega^\top} = \frac{\partial \mathbf{V}_\omega^\top}{\partial \varphi_i} \left(-\frac{2}{\alpha^2} \Sigma^{-1} + \frac{1}{2} \mathbf{I}_n \right) \mathbf{A} \\ &\quad + \mathbf{V}_\omega^\top \left(\frac{2}{\alpha^2} \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_i} \Sigma^{-1} \mathbf{A} + \left(-\frac{2}{\alpha^2} \Sigma^{-1} + \frac{1}{2} \mathbf{I}_n \right) \frac{\partial \mathbf{A}}{\partial \varphi_i} \right) \\ &= \mathbf{M}^\top \left(\frac{-1}{\alpha^2} \Sigma^{-1} + \frac{1}{4} \mathbf{I}_n \right) \mathbf{A} + \mathbf{V}_\omega^\top \left(\frac{2}{\alpha^2} \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_i} \Sigma^{-1} \right) \mathbf{A} \\ &\quad + \left(\frac{-2}{\alpha^2} \Sigma^{-1} + \frac{1}{2} \mathbf{I}_n \right) \mathbf{A} \left(\frac{1}{\alpha} \Sigma^{-\frac{1}{2}} \frac{\partial \Sigma^{\frac{1}{2}}}{\partial \varphi_i} \Sigma^{-\frac{1}{2}} + \frac{\alpha}{4} \frac{\partial \Sigma^{\frac{1}{2}}}{\partial \varphi_i} \right) \mathbf{A}, \end{aligned}$$

where $\mathbf{M} = (m_1, \dots, m_n)^\top$, with $m_i = \mathbf{l}_i \omega$ and \mathbf{l}_i being the i th row of matrix

$$\mathbf{L} = \mathbf{A} \left(-\frac{1}{\alpha} \Sigma^{-\frac{1}{2}} \frac{\partial \Sigma^{\frac{1}{2}}}{\partial \varphi_i} \Sigma^{-\frac{1}{2}} + \frac{\alpha}{4} \frac{\partial \Sigma^{\frac{1}{2}}}{\partial \varphi_i} \right) \mathbf{A}, \quad j = 1, \dots, n, i = 1, 2, 3.$$

Details about $\partial \Sigma^{\frac{1}{2}} / \partial \varphi_i$ can be found in De Bastiani et al. (2015).