Journal of Statistical Planning and Inference 98 (2001) 191–209

journal of statistical planning and inference

www.elsevier.com/locate/jspi

The change-of-variance function: a tool to explore the effects of dependencies in spatial statistics

Marc G. Genton

Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA

Received 10 November 1998; received in revised form 15 September 2000; accepted 18 October 2000

Abstract

This paper presents the computation of the change-of-variance function of M-estimators of scale under general contamination for dependent observations. In this context, several results of robustness are established, and the links between B-robustness, V-robustness and $V^{\diamond}$-robustness are studied. Some more specific properties are derived for Gaussian distributions. These results are then applied to variogram estimation, which is a crucial stage of spatial prediction. The change-of-variance function is shown to be a tool to explore the effects of dependencies on the variance of variogram estimators. ARMA models are used in order to model unidirectional spatial dependencies. It is shown that the shape of the change-of-variance function under dependence is characteristic of the type of variogram estimator. However, this shape depends also on the underlying dependency structure, its intensity, as well as the lag distance of the variogram estimates. Therefore, statistical insight is provided into the sensitivity and the behavior of the variance of the variogram estimator at different spatial lags. For instance, this variance plays an important role when fitting a parametric variogram model by weighted or generalized least squares. © 2001 Published by Elsevier Science B.V.

MSC: 62G35; 62M30

Keywords: Asymptotic variance; Dependent data; M-estimator; Robustness; Scale estimation; Variogram

1. Introduction

The purpose of this article is twofold. First, we compute the change-of-variance function of M-estimators of scale under general contamination for dependent data and study the corresponding robustness properties. Second, we discuss the use of the change-of-variance function as a tool to explore the effects of dependencies in spatial statistics on variogram estimators: it will produce insight into the sensitivity and

E-mail address: genton@math.mit.edu (M.G. Genton).

0378-3758/01/$ - see front matter
PII: S0378-3758(00)00298-6

the behavior of the variance of these estimators in dependent situations. This can be best done by means of a simple example, involving ARMA models as in time-series analysis. As an introduction, let us recall some concepts of spatial statistics.

Spatial statistical methods widely known under the name *kriging* are intended to predict unobserved values of a variable in a spatial domain, on the basis of observed values (e.g. Journel and Huijbregts, 1978; Isaaks and Srivastava, 1989; Haining, 1990; Cressie, 1993). These techniques are based on a function which describes the spatial dependence, the so-called *variogram*. Estimation of the variogram is a crucial stage of spatial prediction, because it determines the kriging weights, which are used to draw maps of the variable under study. Let us consider a spatial stochastic process $\{Z(\boldsymbol{x}): \boldsymbol{x} \in D\}$, where D is a fixed subset of $\mathbb{R}^d$, $d \geqslant 1$. Assume that this process is ergodic and satisfies the hypothesis of intrinsic stationarity (e.g. Cressie, 1993, p. 60) given by

(a) $\mathrm{E}(Z(\boldsymbol{x})) = \mu = \text{constant } \forall \boldsymbol{x} \in D$,
(b) $\mathrm{Var}(Z(\boldsymbol{x}+\boldsymbol{h}) - Z(\boldsymbol{x})) = 2\gamma(\boldsymbol{h}) \quad \forall \boldsymbol{x}, \boldsymbol{x}+\boldsymbol{h} \in D$,

where $2\gamma(\boldsymbol{h})$ is the variogram. This is a very simple model which can be used in practice after detrending the data or in some cases even directly. Let $\{Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n)\}$ be a sample of such a spatial process. The classical variogram estimator proposed by Matheron (1962), based on the method-of-moments, is

$$2\hat{\gamma}(\boldsymbol{h}) = \frac{1}{N_{\boldsymbol{h}}} \sum_{N(\boldsymbol{h})} (Z(\boldsymbol{x}_i) - Z(\boldsymbol{x}_j))^2, \quad \boldsymbol{h} \in \mathbb{R}^d, \tag{1}$$

where $N(\boldsymbol{h}) = \{(\boldsymbol{x}_i, \boldsymbol{x}_j): \boldsymbol{x}_i - \boldsymbol{x}_j = \boldsymbol{h}\}$ and $N_{\boldsymbol{h}}$ is the cardinality of $N(\boldsymbol{h})$. This estimator is unbiased, but behaves poorly if there are outliers in the data. One single outlier can make this estimator arbitrarily large. For that reason, Cressie and Hawkins (1980) proposed a more "robust" estimator in the case of Gaussian independent data:

$$2\hat{\gamma}(\boldsymbol{h}) = \left[\frac{1}{N_{\boldsymbol{h}}} \sum_{N(\boldsymbol{h})} |Z(\boldsymbol{x}_i) - Z(\boldsymbol{x}_j)|^{1/2}\right]^4 \Big/ \left(0.457 + \frac{0.494}{N_{\boldsymbol{h}}}\right), \quad \boldsymbol{h} \in \mathbb{R}^d,$$

where the denominator corrects for bias under a Gaussian distribution. However, this estimator can also be made arbitrarily large by a single outlier in the data and is, therefore, not really a solution to the problem.

To view variogram estimation as a problem of identifying the scale at various lags is intuitively appealing and opens up new perspectives. By an estimator of the scale of a sample $V_1, \ldots, V_n$ we mean any function $S_n(V_1, \ldots, V_n)$ which satisfies

$$S_n(\alpha V_1 + \beta, \ldots, \alpha V_n + \beta) = |\alpha|\, S_n(V_1, \ldots, V_n)$$

$\forall \alpha \in \mathbb{R}$, $\forall \beta \in \mathbb{R}$. In effect, the stochastic process of differences at lag $\boldsymbol{h}$, $V(\boldsymbol{h}) = Z(\boldsymbol{x} + \boldsymbol{h}) - Z(\boldsymbol{x})$, has zero expectation and a variance of $2\gamma(\boldsymbol{h})$. Thus, if $V_1(\boldsymbol{h}), \ldots, V_{N_h}(\boldsymbol{h})$ is the sample of $V(\boldsymbol{h})$ corresponding to the sample $Z(\boldsymbol{x}_1), \ldots, Z(\boldsymbol{x}_n)$ of Z, the variogram estimator (1) takes the form

$$2\hat{\gamma}(\boldsymbol{h}) = \frac{1}{N_{\boldsymbol{h}}} \sum_{i=1}^{N_{\boldsymbol{h}}} V_i(\boldsymbol{h})^2, \quad \boldsymbol{h} \in \mathbb{R}^d, \tag{2}$$

i.e. it is simply the classical estimator of the sample variance of $V_1(\boldsymbol{h}),\dots,V_{N_{\boldsymbol{h}}}(\boldsymbol{h})$. The use and properties of a new, highly robust, variogram estimator based on a highly robust scale estimator is discussed by Genton (1998a). We now use the theory of M-estimators of scale, in order to derive robustness properties for a wide class of corresponding variogram estimators. Note that although variograms are in fact variances, we here focus on scale estimators because most results in the literature are formulated for scale (e.g. Hampel et al., 1986). However, conclusions are similar for variances, i.e. for variograms, because a scale S_n is simply the square root of a variance S_n^2.

2. M-estimators of scale

Recall the definition of an M-estimator of scale (Hampel et al., 1986). Suppose we have one-dimensional observations $X_1,\dots,X_n$ which are independent and identically distributed according to a distribution from the parametric model $\{F_\sigma;\ \sigma > 0\}$, where $F_\sigma(x) = F(x/\sigma)$. We need the following regularity conditions on F:

(F1) F has a twice continuously differentiable density f which is symmetric around zero and is everywhere positive.

(F2) The mapping $\Lambda = -f'/f = (-\ln f)'$ satisfies $\Lambda'(x) > 0$, $\forall x \in \mathbb{R}$, and $\int \Lambda'(x) f(x)\mathrm{d}x = -\int \Lambda(x) f'(x)\,\mathrm{d}x < \infty$.

An M-estimator $S_n(X_1,\dots,X_n)$ of σ is defined by the implicit equation

$$\frac{1}{n}\sum_{i=1}^{n} \chi(X_i/S_n) = 0,$$

and corresponds asymptotically to the statistical functional S defined by

$$\int \chi(x/S(F))\,\mathrm{d}F(x) = 0, \tag{3}$$

where χ is a real, symmetric, and sufficiently regular function. Denote by

$$A(\chi,F) = \int \chi^2(x/S(F))\,\mathrm{d}F(x), \tag{4}$$

$$B(\chi,F) = \int (x/S(F))\chi'(x/S(F))\,\mathrm{d}F(x). \tag{5}$$

We assume that χ belongs to the class Ψ of all functions satisfying the following four regularity conditions:

(R1) χ is well-defined and continuous on $\mathbb{R} \setminus D^{(0)}(\chi)$, where $D^{(0)}(\chi)$ is finite. In each point of $D^{(0)}(\chi)$ there exist finite left and right limits of χ which are different. Also $\chi(-x) = \chi(x)$ if $\{-x,x\} \subset \mathbb{R} \setminus D^{(0)}(\chi)$, and there exists $\delta > 0$ such that $\chi(x) \leqslant 0$ on $(0,\delta)$ and $\chi(x) \geqslant 0$ on (δ,∞).

(R2) The set $D^{(1)}(\chi)$ of points in which χ is continuous but in which χ' is not defined or not continuous, is finite.
(R3) $\int \chi(x)\,\mathrm{d}F(x) = 0$, i.e. $S(F) = 1$ at the model (Fisher consistency), and $0 < A(\chi,F) < \infty$.
(R4) $0 < B(\chi,F) = \int (x\Lambda(x) - 1)\chi(x)\,\mathrm{d}F(x) < \infty$.

The influence function $\mathrm{IF}(u,S,F)$ of a statistical functional S at a distribution F is defined as the kernel of a first-order von Mises (1937, 1947) derivative

$$\int \mathrm{IF}(u,S,F)\,\mathrm{d}G(u) = \frac{\partial}{\partial\epsilon}[S((1-\epsilon)F + \epsilon G)]_{\epsilon=0},$$

where G ranges over all distributions (including point masses). The influence function of an M-estimator of scale S at F is well known (Hampel et al., 1986)

$$\mathrm{IF}(u,S,F) = \frac{\chi(u)}{B(\chi,F)}.$$

An important summary value of the influence function is the gross-error sensitivity of S at F, defined by

$$\gamma^* = \sup_u |\mathrm{IF}(u,S,F)|.$$

This quantity measures the worst influence that a small amount of contamination can have on the value of the estimator. It is desirable that γ^* be finite, in which case S is called B-robust (bias-robust) at F. Analogously, the change-of-variance function (Rousseeuw, 1981) is defined by

$$\int \mathrm{CVF}(u,S,F)\,\mathrm{d}G(u) = \frac{\partial}{\partial\epsilon}[V(S,(1-\epsilon)F + \epsilon G)]_{\epsilon=0}, \tag{6}$$

where $V(S,F)$ is the asymptotic variance of S at F. The change-of-variance sensitivity κ^* is defined as $+\infty$ if a delta function with positive factor occurs in the CVF, and otherwise as

$$\kappa^* = \sup_u \frac{\mathrm{CVF}(u,S,F)}{V(S,F)}.$$

Note that large negative values of the CVF merely point to a decrease in V, indicating a better accuracy of the estimator. If κ^* is finite then S is called V-robust (variance-robust) at F. Hampel et al. (1986) simplified (6) for M-estimators of scale by considering only contaminating distributions G with $S(G) = S(F) = 1$. The case of general contaminating distributions G was studied by Genton and Rousseeuw (1995) and yields

$$\mathrm{CVF}(u,S,F) = \frac{A(\chi,F)}{B^2(\chi,F)}\left[1 + \frac{\chi^2(u)}{A(\chi,F)} - 2\frac{u\,\chi'(u)}{B(\chi,F)} + C(\chi,F)\,\frac{\chi(u)}{B(\chi,F)}\right], \tag{7}$$

where

$$C(\chi,F) = 4 - \frac{2}{A(\chi,F)}\int x\,\chi(x)\,\chi'(x)\,\mathrm{d}F(x) + \frac{2}{B(\chi,F)}\int x^2\,\chi''(x)\,\mathrm{d}F(x).$$

Note that (7) differs from the expression in Hampel et al. (1986) by the addition of the last term, the integral of which is zero when $S(G) = 1$. In the case of Gaussian observations, an alternative formula for $C(\chi,\Phi)$, which does not involve any derivative of χ, has been obtained (Genton and Rousseeuw, 1995).

Another important robustness property is the breakdown point ϵ^* of a scale estimator. This indicates, how many data points need to be replaced by arbitrary values to make the estimator explode (tend to infinity) or implode (tend to zero). In the case of M-estimators of scale, it has been shown (Huber, 1981) that

$$\epsilon^* = \min\left(\frac{-\chi(0)}{\chi(+\infty)-\chi(0)}, \frac{\chi(+\infty)}{\chi(+\infty)-\chi(0)}\right) \leqslant \frac{1}{2}.$$

The special choice $\chi(x) = |x|^q - \int |x|^q \, dF(x)$, $q > 0$, leads to the so-called L^q M-estimators of scale (Genton and Rousseeuw, 1995), which are neither B-robust, nor V-robust, for every value of $q > 0$, that is to say, $\gamma^* = \infty$ and $\kappa^* = \infty$. Moreover, it is easily seen that $\epsilon^* = 0$, for any value of $q > 0$. A closer look at Eq. (2) and the corresponding equation for the Cressie and Hawkins estimator shows that they correspond to the L^2 and $L^{1/2}$ estimators, respectively. Thus, these two estimators are neither robust in the sense of the influence function nor in the sense of the change-of-variance function and the breakdown point.

The previous results are all based on the hypothesis that the observations are independent of each other. This assumption of independence is prevalent in classical statistical theory and makes much of it tractable. However, models that involve statistical dependence are often more realistic, and are a necessity in spatial statistics where the dependence is often present in all directions and becomes weaker as data locations become more distanced. In fact, any discipline that works with data collected from different spatial locations, such as soil science, geology, mining, hydrology, forestry, atmospheric or soil pollution, meteorology, astronomy, must develop models that indicate when there is dependence between measurements at different locations.

In this paper, we intend to study the effects on the estimator resulting from dependencies between observations. Intuitively, dependence between observations is not going to modify the expectation of an estimator. On the other hand, its variance should increase because dependence expresses the fact that observations are more alike. Thus, we study the way this variance changes under dependence, i.e. the change-of-variance function under dependence. Note that the variability of variogram estimates plays an important role when fitting a parametric variogram model by weighted least squares (Cressie, 1985, 1993) or generalized least squares (Cressie 1993; Genton, 1998b, 2000). It is therefore worth understanding how this variance can be affected by outliers and by the underlying dependency structure.

3. The change-of-variance function under dependence

The behavior of robust estimators for independent and identically distributed observations has been extensively studied in the past. The case of dependent data received much less attention. It seems that pioneers in this field were Gastwirth and

Rubin (1975), with a paper investigating the effect of serial dependence in the data on the efficiency of some robust location estimators. This theme was followed up by Portnoy (1977, 1979), who studied approximately optimal estimators, in the asymptotic minimax sense of Huber (1964, 1972, 1981), in dependent situations. He also showed that the influence function remains unchanged for dependent observations and must be computed with the marginal distribution. Hössjer (1991) was the first to study the change-of-variance function for dependent observations. However, he restricted attention to the location model and his context is radically different from the one of spatial statistics. He considers an independent and identically distributed (i.i.d.) process and is interested in perturbations from a dependent process. His model is for replacement outliers and allows differentiating isolated from grouped perturbations. In fact, he is interested in the effect of unexpected correlations in the data. In spatial statistics, the data are fundamentally considered as being dependent and we rather want to study the effect of punctual, i.i.d. perturbations.

In this section, we compute the change-of-variance function under general contamination and dependent observations for M-estimators of scale. Let $V_1,\dots,V_{N_{\boldsymbol{h}}}$ be a realization of the process of differences V, with a joint distribution in $\mathbb{R}^{N_{\boldsymbol{h}}}$ denoted by $F_V^{N_{\boldsymbol{h}}}$. We suppose that each difference V_i, $i=1,\dots,N_{\boldsymbol{h}}$, is identically distributed according to a marginal distribution F_V, with zero expectation and variance equal to the variogram $2\gamma(\boldsymbol{h})$. We assume that for all $k=1,\dots,N_{\boldsymbol{h}}-1$, the bivariate distribution of the pairs (V_i,V_{i+k}) are the same for all $i=1,\dots,N_{\boldsymbol{h}}-k$. We then denote this bivariate distribution by $F_V^{(k)}$. Note that this assumption only makes sense if the original data are equally spaced along a line. An M-estimator $S_{N_{\boldsymbol{h}}}(V_1,\dots,V_{N_{\boldsymbol{h}}})$ of the scale of the stochastic process V corresponds asymptotically to the statistical functional of scale $S(F_V)$, implicitly defined by Eq. (3) with $F=F_V$. Under some regularity conditions, the M-estimator $S_{N_{\boldsymbol{h}}}$ is consistent, i.e. $S_{N_{\boldsymbol{h}}} \overset{N_{\boldsymbol{h}}\to\infty}{\to} S(F_V)$ in probability. Moreover, $\sqrt{N_{\boldsymbol{h}}}(S_{N_{\boldsymbol{h}}}-S(F_V))$ is asymptotically normal, with zero expectation and variance $V^{\diamondsuit}(S,F_V)$, given by (Portnoy, 1977)

$$V^{\diamondsuit}(S,F_V)=\frac{A(\chi,F_V)+2\sum_{k=1}^{\infty}A^{\diamondsuit}(\chi,F_V^{(k)})}{B^2(\chi,F_V)}S^2(F_V), \tag{8}$$

where A and B are defined by Eqs. (4) and (5), and

$$A^{\diamondsuit}(\chi,F_V^{(k)})=\int\int \chi(x_1/S(F_V))\chi(x_2/S(F_V))\,\mathrm{d}F_V^{(k)}(x_1,x_2).$$

Throughout this paper, the diamond ($\diamondsuit$) notation is used to emphasize situations with dependencies. Regularity conditions for consistency and asymptotic normality are given by Huber (1967) for the independent case and by Portnoy (1977, 1979) and Bustos (1982) for the dependent case. In this latter situation, mixing conditions such as α-mixing or ϕ-mixing are typically considered (Billingsley, 1968; Doukhan, 1994).

Let us now describe the contaminating process. We contaminate V by an independent and identically distributed process H. Although one could suspect that H should also be a lag $\boldsymbol{h}$ process of differences, from some other i.i.d. process contaminating the original Z process, this approach would not be tractable. The main reason is that the spatial location of a contamination in Z has a crucial importance on its effect on differences V, as noted by Genton (1998c) when trying to define a spatial breakdown point. For instance, on a regular bidimensional grid, a contamination in Z can generally affect 1, 2, 3, or 4 differences V, depending on its spatial location and the lag $\boldsymbol{h}$ considered. Thus, by contaminating the process of differences V directly, we avoid this difficulty. Let $H_1,\dots,H_{N_{\boldsymbol{h}}}$ be a realization of the process H and $B_1,\dots,B_{N_{\boldsymbol{h}}}$ a realization of a Bernoulli stochastic process, where B_i, $i=1,\dots,N_{\boldsymbol{h}}$, are i.i.d. random variables according to a Bernoulli distribution with parameter ϵ, i.e.

$$P(B_i=1)=\epsilon \quad \text{and} \quad P(B_i=0)=1-\epsilon, \tag{9}$$

for $i=1,\dots,N_{\boldsymbol{h}}$, with $0\leqslant\epsilon\leqslant1$. The contaminating process, describing replacement outliers, is then defined by

$$V_{\epsilon,i}=(1-B_i)V_i+B_iH_i,$$

for $i=1,\dots,N_{\boldsymbol{h}}$. We denote by

- F_{V_ϵ}, F_V, and F_H the univariate marginal distributions of $V_{\epsilon,i}$, V_i, and H_i, respectively.
- $F_{V_\epsilon}^{(k)}$, $F_V^{(k)}$, and $F_H^{(k)}$ the bivariate distributions of the pairs $(V_{\epsilon,i},V_{\epsilon,i+k})$, (V_i,V_{i+k}), and (H_i,H_{i+k}), respectively.
- $F_{VH}^{(k)}$, and $F_{HV}^{(k)}$ the bivariate distributions of the pairs (V_i,H_{i+k}) and (H_i,V_{i+k}), respectively.

We suppose that F_V satisfies the conditions (F1) and (F2), and $\chi\in\Psi$ satisfies the regularity conditions (R1)–(R4), given in the previous section. Moreover, we assume that $S(F_V)=1$ at the model and we add the two following assumptions:

(H1) The stochastic processes V, H and B are mutually independent.
(H2) $\sum_{k=1}^{\infty}|A^{\diamondsuit}(\chi,F_V^{(k)})|<\infty$.

From Eq. (9), we have

$$P(B_i=B_{i+k}=1)=\epsilon^2=\mathrm{O}(\epsilon^2),$$
$$P(B_i=B_{i+k}=0)=(1-\epsilon)^2=1-2\epsilon+\mathrm{O}(\epsilon^2),$$

and thus

$$\begin{aligned}A^{\diamondsuit}(\chi,F_{V_\epsilon}^{(k)})&=P(B_i=B_{i+k}=0)A^{\diamondsuit}(\chi,F_V^{(k)})+P(B_i=0,B_{i+k}=1)A^{\diamondsuit}(\chi,F_{VH}^{(k)})\\&\quad+P(B_i=1,B_{i+k}=0)A^{\diamondsuit}(\chi,F_{HV}^{(k)})+P(B_i=B_{i+k}=1)A^{\diamondsuit}(\chi,F_H^{(k)})\\&=(1-2\epsilon)A^{\diamondsuit}(\chi,F_V^{(k)})+\mathrm{O}(\epsilon^2),\end{aligned}$$

using $A^{\diamondsuit}(\chi,F_{VH}^{(k)})=A^{\diamondsuit}(\chi,F_{HV}^{(k)})=0$ by (H1). The numerator of the asymptotic variance (8) of the contaminated process is

$$\begin{aligned}&A(\chi,F_{V_\epsilon})+2\sum_{k=1}^{\infty}A^{\diamondsuit}(\chi,F_{V_\epsilon}^{(k)})\\&\quad=(1-\epsilon)A(\chi,F_V)+\epsilon A(\chi,F_H)+2(1-2\epsilon)\sum_{k=1}^{\infty}A^{\diamondsuit}(\chi,F_V^{(k)})+\mathrm{O}(\epsilon^2),\end{aligned}$$

using $A^{\diamondsuit}(\chi, F_H^{(k)}) = 0$ and (H2), and its derivative evaluated at $\epsilon = 0$ is

$$\frac{\partial}{\partial \epsilon}\left[A(\chi, F_{V_\epsilon}) + 2\sum_{k=1}^{\infty} A^{\diamondsuit}(\chi, F_{V_\epsilon}^{(k)})\right]_{\epsilon=0}$$

$$= A(\chi, F_H) - A(\chi, F_V) - 4\sum_{k=1}^{\infty} A^{\diamondsuit}(\chi, F_V^{(k)})$$

$$-\frac{2}{S(F_V)}\left(\int (x/S(F_V))\chi(x/S(F_V))\chi'(x/S(F_V))\,\mathrm{d}F_V(x)\right.$$

$$\left. +2\sum_{k=1}^{\infty} C_1^{\diamondsuit}(\chi, F_V^{(k)})\right)\int \mathrm{IF}(x, S, F_V)\,\mathrm{d}F_H(x),$$

where

$$C_1^{\diamondsuit}(\chi, F_V^{(k)}) = \int\int (x_1/S(F_V))\chi'(x_1/S(F_V))\chi(x_2/S(F_V))\,\mathrm{d}F_V^{(k)}(x_1, x_2).$$

In the same way, the denominator of the asymptotic variance (8) of the contaminated process is

$$B^2(\chi, F_{V_\epsilon}) = ((1-\epsilon)B(\chi, F_V) + \epsilon B(\chi, F_H))^2,$$

and its derivative evaluated at $\epsilon = 0$ is

$$\frac{\partial}{\partial \epsilon}[B^2(\chi, F_{V_\epsilon})]_{\epsilon=0} = 2B(\chi, F_V)\left[B(\chi, F_H) - B(\chi, F_V) - \int \mathrm{IF}(x, S, F_V)\,\mathrm{d}F_H(x)\right.$$

$$\times\left(\int (x^2/S^2(F_V))\chi''(x/S(F_V))\,\mathrm{d}F_V(x)\right.$$

$$\left.\left. +\frac{1}{S(F_V)}\int (x/S(F_V))\chi'(x/S(F_V))\,\mathrm{d}F_V(x)\right)\right].$$

Thus, the change-of-variance function $\mathrm{CVF}^{\diamondsuit}(u, S, F_V)$ under dependence is

$$\int \mathrm{CVF}^{\diamondsuit}(x, S, F_V)\,\mathrm{d}F_H(x)$$

$$= \left[\left(2S(F_V)\left(A(\chi, F_V) + 2\sum_{k=1}^{\infty} A^{\diamondsuit}(\chi, F_V^{(k)})\right)\int \mathrm{IF}(x, S, F_V)\,\mathrm{d}F_H(x)\right.\right.$$

$$\left. +S^2(F_V)\frac{\partial}{\partial \epsilon}\left[A(\chi, F_{V_\epsilon}) + 2\sum_{k=1}^{\infty} A^{\diamondsuit}(\chi, F_{V_\epsilon}^{(k)})\right]_{\epsilon=0}\right) B^2(\chi, F_V)$$

$$\left. -S^2(F_V)\left(A(\chi, F_V) + 2\sum_{k=1}^{\infty} A^{\diamondsuit}(\chi, F_V^{(k)})\right)\frac{\partial}{\partial \epsilon}[B^2(\chi, F_{V_\epsilon})]_{\epsilon=0}\right] B^{-4}(\chi, F_V).$$

Finally, we insert the previous expressions for the derivatives in the last equation and use the fact that $S(F_V) = 1$ at the model. Choosing $H = \Delta_u$, the Dirac distribution with jump at u, and normalizing the change-of-variance function such that its integral with respect to F_V is zero in order to impose uniqueness, i.e. $\int \mathrm{CVF}^{\diamond}(x,S,F_V)\,\mathrm{d}F_V(x)=0$, we obtain

$$\mathrm{CVF}^{\diamond}(u,S,F_V) = V^{\diamond}(S,F_V)\left[1 + \frac{\chi^2(u) + 2\sum_{k=1}^{\infty} A^{\diamond}(\chi,F_V^{(k)})}{A(\chi,F_V) + 2\sum_{k=1}^{\infty} A^{\diamond}(\chi,F_V^{(k)})} - 2\frac{u\,\chi'(u)}{B(\chi,F_V)}\right.$$
$$\left. + C^{\diamond}(\chi,F_V)\,\frac{\chi(u)}{B(\chi,F_V)}\right], \tag{10}$$

where

$$C_1(\chi,F_V) = \int x\chi(x)\chi'(x)\,\mathrm{d}F_V(x),$$
$$C_2(\chi,F_V) = \int x^2\chi''(x)\,\mathrm{d}F_V(x),$$
$$C_1^{\diamond}(\chi,F_V^{(k)}) = \int\!\!\int x_1\chi'(x_1)\chi(x_2)\,\mathrm{d}F_V^{(k)}(x_1,x_2),$$
$$C^{\diamond}(\chi,F_V) = 4 - 2\frac{C_1(\chi,F_V) + 2\sum_{k=1}^{\infty} C_1^{\diamond}(\chi,F_V^{(k)})}{A(\chi,F_V) + 2\sum_{k=1}^{\infty} A^{\diamond}(\chi,F_V^{(k)})} + 2\frac{C_2(\chi,F_V)}{B(\chi,F_V)}.$$

From now on, we also suppose that (H3) $\sum_{k=1}^{\infty} C_1^{\diamond}(\chi,F_V^{(k)}) < \infty$. If the process V is independent, then $A^{\diamond}(\chi,F_V^{(k)})=C_1^{\diamond}(\chi,F_V^{(k)})=0$, which yields the usual change-of-variance function under general contamination (Genton and Rousseeuw, 1995) given by (7).

4. $V^{\diamond}$-robustness

The change-of-variance function of an M-estimator of scale under dependence is a tool which allows us to study the effects of dependencies on the asymptotic variance of the estimator, as well as its variations. By analogy with the independent case, we define the notion of change-of-variance sensitivity under dependence.

Definition 1. The change-of-variance sensitivity under dependence, $\kappa^{\diamond}$, of an M-estimator of scale S at F_V is equal to $+\infty$ if a delta function with positive factor occurs in the $\mathrm{CVF}^{\diamond}$, and otherwise as

$$\kappa^{\diamond} = \sup_u \frac{\mathrm{CVF}^{\diamond}(u,S,F_V)}{V^{\diamond}(S,F_V)},$$

the supremum being taken on all u where $\mathrm{CVF}^{\diamond}(u,S,F_V)$ exists.

If $\kappa^{\diamond}$ is finite then S is called $V^{\diamond}$-robust, i.e. robust with respect to the variance under dependence. A closer look at the change-of-variance function under dependence (10) indicates that V-robustness and $V^{\diamond}$-robustness are equivalent. However, the change-of-variance sensitivity in the dependent case is different from the one in the independent case.

The following theorem shows that the concept of $V^{\diamond}$-robustness is stronger than the concept of B-robustness, and the next one, that sometimes these two notions are equivalent. A corollary gives a lower bound for the change-of-variance sensitivity under dependence $\kappa^{\diamond}$. Let us define $\gamma^- = \sup_{u\in(0,\delta)}(-\mathrm{IF}(u,S,F_V))$ and $\gamma^+ = \sup_{u\in(\delta,+\infty)}\mathrm{IF}(u,S,F_V)$. In the theorems below we will suppose that $\gamma^+ \geqslant \gamma^-$ (and hence $\gamma^* = \gamma^+$). This is a very natural requirement for scale estimators (Genton and Rousseeuw, 1995). For instance Huber (1981), when discussing breakdown properties, notes that $\gamma^+ \geqslant \gamma^-$ in the more interesting cases. The opposite situation leads to implosion of the scale estimator, as well as to lower efficiency. The expression denoted by

$$G^{\diamond}(\chi,F_V) = \frac{2\sum_{k=1}^{\infty} A^{\diamond}(\chi,F_V^{(k)})}{A(\chi,F_V) + 2\sum_{k=1}^{\infty} A^{\diamond}(\chi,F_V^{(k)})},$$

plays an important role in the following theorems. All proofs can be found in the appendix.

Theorem 1. *For all $\chi\in\Psi$ with $\gamma^+\geqslant\gamma^-$ and $C^{\diamond}(\chi,F_V)\geqslant 0$, $V^{\diamond}$-robustness implies B-robustness. We have*

$$\gamma^* \leqslant \tfrac{1}{2}[(V^{\diamond 2}(S,F_V)C^{\diamond 2}(\chi,F_V) + 4V^{\diamond}(S,F_V)(\kappa^{\diamond} - 1 - G^{\diamond}(\chi,F_V)))^{1/2} - V^{\diamond}(S,F_V)C^{\diamond}(\chi,F_V)].$$

Theorem 2. *For all $\chi\in\Psi$ with $\gamma^+\geqslant\gamma^-$, $C^{\diamond}(\chi,F_V)\geqslant 0$, and χ nondecreasing for $x\geqslant 0$, $V^{\diamond}$-robustness and B-robustness are equivalent. We have*

$$\kappa^{\diamond} = 1 + G^{\diamond}(\chi,F_V) + \frac{(\gamma^*)^2}{V^{\diamond}(S,F_V)} + C^{\diamond}(\chi,F_V)\gamma^*.$$

Corollary 2.1. *For all $\chi\in\Psi$ with $\gamma^+\geqslant\gamma^-$, $C^{\diamond}(\chi,F_V)\geqslant 0$, and χ nondecreasing for $x\geqslant 0$, we have*

$$\kappa^{\diamond} \geqslant 2 + C^{\diamond}(\chi,F_V)\gamma^*.$$

Stronger results can be proved, when the spatial stochastic process Z is assumed to be Gaussian. Denote by Φ the standard Gaussian distribution and by Φ_τ, $-1<\tau<1$, the Gaussian bivariate distribution with mean zero and covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \tau \\ \tau & 1 \end{pmatrix}.$$

A further restriction on τ may be necessary in order to insure positive definiteness of the complete covariance matrix of the stochastic process V. In this situation, Genton (1998d) shows that $A^{\diamond}(\chi,\Phi_\tau)\geqslant 0$, and consequently $V^{\diamond}(S,\Phi)\geqslant V(S,\Phi)$, i.e. the asymptotic variance of an M-estimator of scale must necessarily increase along with the dependence. Note that this result is not necessarily true if the underlying distribution is not Gaussian, with negatively correlated observations. Nevertheless, the

Gaussian assumption on the spatial stochastic process Z allows some more accurate results than those of Theorem 2 and its corollary. We need the following lemma.

Lemma 1. *For all M-estimator of scale S, defined by some $\chi \in \Psi$, and based on dependent data with marginal distribution $F_V = \Phi$ and bivariate distributions $F_V^{(k)} = \Phi_{\tau_k}$, $k \geqslant 1$, we have*

$$0 \leqslant G^{\diamondsuit}(\chi, F_V) < 1.$$

Theorem 3. *Let S be an M-estimator of scale, defined by some $\chi \in \Psi$, and based on dependent data with marginal distribution $F_V = \Phi$ and bivariate distributions $F_V^{(k)} = \Phi_{\tau_k}$, $k \geqslant 1$. If $\gamma^+ \geqslant \gamma^-$, $C(\chi, \Phi) \geqslant C^{\diamondsuit}(\chi, \Phi) \geqslant 0$, and χ nondecreasing for $x \geqslant 0$, then we have*

$$\kappa^{\diamondsuit} < \kappa^* + 1.$$

Thus, in that particular case, we are able to specify the relation between the change-of-variance sensitivity under dependence $\kappa^{\diamondsuit}$ and under independence κ^*. The analysis of some examples shows that for the same estimator, $\kappa^{\diamondsuit}$ can be larger or smaller than κ^*, depending on the kind of underlying dependency structure. However, these two quantities are generally close to each other, because it is the asymptotic variance in Eq. (10) which suffers quite a large variation due to dependencies.

5. Application to variogram estimation

In this section we model the dependency structure of the spatial stochastic process Z and study its effect on variogram estimators. We analyze the behavior of the change-of-variance function by using the family of autoregressive moving average (ARMA) stationary processes, defined on a unidimensional and equally spaced support. In spatial statistics, the variogram is often computed in one or several unidimensional directions. Therefore, we assume that observations on such a section are drawn from an ARMA process. Note that bidimensional dependency structures could also be considered, for example by using spatial autoregressive and moving average (SARMA) models, see Cressie (1993). Consider an ARMA(p,q) stationary process Z, i.e. for every location i

$$\phi(B)Z_i = \theta(B)Y_i, \tag{11}$$

where ϕ and θ are the pth and qth degree polynomials $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ and $\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$, B is the backward shift operator $B^h Z_i = Z_{i-h}$, and Y is an uncorrelated process, with expectation zero and variance σ^2. Subsequently, we assume that the parameters of the ARMA process are such that the process is causal and invertible (Brockwell and Davis, 1991). The process of differences $V(h)$ is defined by $V_i(h) = (1 - B^h)Z_i$. Defining $\theta_i = 0$ for $i > q$ and $i < 0$, and $\theta_0 = 1$, it follows from Eq. (11) that the process of differences $V(h)$ is an ARMA$(p, q+h)$ process with new coefficients $\tilde{\theta}_i = \theta_i - \theta_{i-h}$ and $\tilde{\theta}_i = 0$ for $i > q + h$. The dependency structure of the process of differences $V(h)$ changes with each value of h, and consequently also the behavior of the variogram estimator, its variance and change-of-variance function.

Let v_k^Z be the lag k covariance of the ARMA process Z and $v_k^{V(h)}$ of the process of differences $V(h)$. They are related by $v_k^{V(h)} = -(v_{k-h}^Z - 2v_k^Z + v_{k+h}^Z)$. Hence the correlation τ_k of $F_V^{(k)}$ is given by

$$\tau_k = -\frac{v_{k-h}^Z - 2v_k^Z + v_{k+h}^Z}{2(v_0^Z - v_h^Z)}. \tag{12}$$

In order to study the effects of dependencies, we can choose simple models for the spatial stochastic process Z from (11), like the AR(1) process with $v_k^Z = \phi_1^k$, or the MA(1) process with $v_0^Z = 1$, $v_{\pm 1}^Z = \theta_1/(1+\theta_1^2)$, and $v_k^Z = 0$ otherwise.

The effects of dependencies are now studied on the variogram estimator (1), which is the square of the L^2 scale estimator. Cressie and Hawkins' variogram estimator, as well as other more robust variogram estimators based on M-estimators of scale, have the same kind of behavior, although their change-of-variance functions have a different characteristic shape which depends on the function χ. Let us first present an interesting property shared by the L^q scale estimators, $q > 0$, which is the same as in the independent case (Genton and Rousseeuw, 1995).

Lemma 2. *For every marginal distribution F_V and bivariate distributions $F_V^{(k)}$, $k \geqslant 1$, the L^q scale estimator, $q > 0$, satisfies*

$$C^{\diamondsuit}(\chi, F_V) = 2.$$

As we already mentioned, the robustness properties of the L^q estimators are poor.

Theorem 4. *The L^q scale estimator is neither B-robust, V-robust, nor $V^{\diamondsuit}$-robust, at any marginal distribution F_V, and bivariate distributions $F_V^{(k)}$, $k \geqslant 1$, i.e. $\gamma^* = \infty$, $\kappa^* = \infty$, and $\kappa^{\diamondsuit} = \infty$.*

From the delta technique, the asymptotic variances under dependence of variogram and scale estimators satisfy the following relation:

$$V^{\diamondsuit}(S^2, F_V) = 4S^2(F_V)V^{\diamondsuit}(S, F_V).$$

Therefore, using Eq. (10), the change-of-variance function under general contamination and dependence for variogram estimators is

$$\mathrm{CVF}^{\diamondsuit}(u, S^2, F_V) = V^{\diamondsuit}(S^2, F_V)\left[1 + \frac{\chi^2(u) + 2\sum_{k=1}^{\infty} A^{\diamondsuit}(\chi, F_V^{(k)})}{A(\chi, F_V) + 2\sum_{k=1}^{\infty} A^{\diamondsuit}(\chi, F_V^{(k)})} - 2\frac{u\,\chi'(u)}{B(\chi, F_V)} \right.$$
$$\left. + (C^{\diamondsuit}(\chi, F_V) + 2)\frac{\chi(u)}{B(\chi, F_V)}\right], \tag{13}$$

Note that Eq. (13) differs from (10) by the addition of the constant 2 in the last term, and the factorized asymptotic variance $V^{\diamondsuit}(S^2, F_V)$. In particular, for variogram

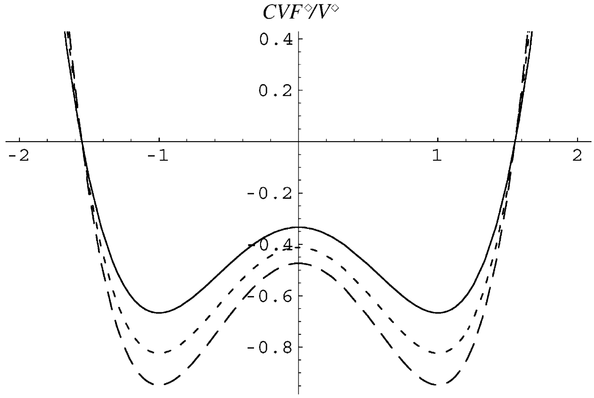

Fig. 1. The standardized change-of-variance function $\mathrm{CVF}^{\diamondsuit}(u,(L^2)^2,\Phi)/V^{\diamondsuit}((L^2)^2,\Phi)$ of the variogram estimator (1) for $h=1$, when Z is an AR(1) process. The dashed curves correspond to $\phi_1=0.4$ and $\phi_1=0.8$ (in increasing length of the dashes), whereas the solid one correspond to $\phi_1=0$, i.e. independence of Z. The curves will vary with the lag h too.

estimators based on L^q scale estimators:

$$\mathrm{CVF}^{\diamondsuit}(u,(L^q)^2,F_V)=V^{\diamondsuit}((L^q)^2,F_V)\left[\frac{\chi^2(u)+2\sum_{k=1}^{\infty}A^{\diamondsuit}(\chi,F_V^{(k)})}{A(\chi,F_V)+2\sum_{k=1}^{\infty}A^{\diamondsuit}(\chi,F_V^{(k)})}\right.$$
$$\left.+(4-2q)\frac{\chi(u)}{B(\chi,F_V)}-1\right], \tag{14}$$

using $u\chi'(u)=q\chi(u)+B(\chi,F_V)$ and Lemma 2. The variogram estimator (1) corresponds to $q=2$, i.e. $\chi(x)=x^2-1$, in which case straightforward computations yield $A(\chi,\Phi)=B(\chi,\Phi)=2$ and $A^{\diamondsuit}(\chi,\Phi_\tau)=2\tau^2$. Note that when $q=2$, the middle term of Eq. (14) is zero. Thus, the change-of-variance function under general contamination and dependence for Matheron's variogram estimator is

$$\mathrm{CVF}^{\diamondsuit}(u,(L^2)^2,\Phi)=V^{\diamondsuit}((L^2)^2,F_V)\left[\frac{u^4-2u^2-1}{2+4s}\right], \tag{15}$$

where $s=\sum_{k=1}^{\infty}\tau_k^2\geqslant 1$. For instance, consider an AR(1) process (11), where Y is Gaussian with $\sigma^2=1$. From Eq. (12), it follows that $s=(1-\phi_1)/(1+\phi_1)/4$ when $h=1$. For this situation, Fig. 1 illustrates the standardized change-of-variance function $\mathrm{CVF}^{\diamondsuit}(u,(L^2)^2,\Phi)/V^{\diamondsuit}((L^2)^2,\Phi)$ of the variogram estimator (1). The dashed curves correspond to $\phi_1=0.4$ and $\phi_1=0.8$ (in increasing length of the dashes), whereas the solid one corresponds to $\phi_1=0$, i.e. independence of Z. It is interesting to note that the accuracy of the estimator, for contaminations close to the origin, increases with increasing dependence ϕ_1. For contaminations far from the origin, the accuracy decreases with increasing dependence ϕ_1. This simple example shows that the change-of-variance function under dependence has the same characteristic shape, depending on the type of variogram estimator. However, its behavior can vary, according to the underlying

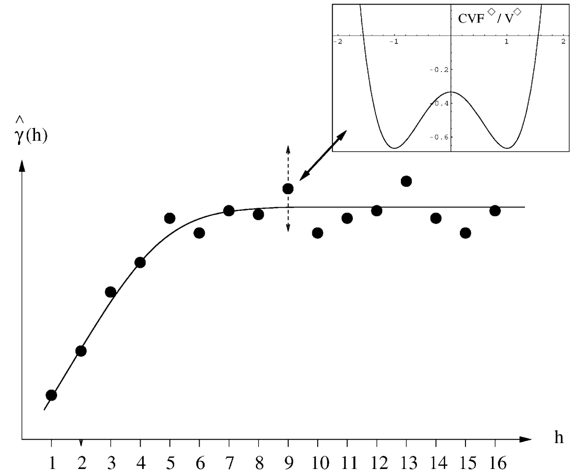

Fig. 2. Schematic illustration of the sensitivity of the asymptotic variance of the variogram estimator (1) to outliers by the standardized change-of-variance function $\mathrm{CVF}^{\diamondsuit}/V^{\diamondsuit}$ under dependence. This sensitivity depends on the underlying dependence and the value of the lag h. The dashed arrow illustrates the vertical variability of a particular variogram estimate.

dependency structure and the value of the lag h. The change-of-variance function can therefore be used to explore the effects of dependencies on variogram estimators, as is schematically illustrated in Fig. 2. In this picture, variogram estimates for 16 lags are represented (black dots), as well as a possible fit of a valid variogram model (solid line). These variogram estimates are correlated because the same observation is used for different lags (Cressie, 1993; Genton, 1998b). Therefore, each of them has a vertical variability (dashed arrow) which can be described by the asymptotic variance under dependence of the variogram estimator. Expressions for the vertical variability of each variogram estimate in situations with Gaussian, elliptically contoured, or skew-normal observations can be found in Genton (1998b, 2000), and Genton et al. (2001), respectively. Moreover, for each variogram estimate, the sensitivity of the asymptotic variance to outliers can be assessed by the standardized change-of-variance function $\mathrm{CVF}^{\diamondsuit}/V^{\diamondsuit}$ under dependence. This sensitivity depends on the underlying dependence and the value of the lag h. This means that outliers will affect the variability of variogram estimates in different ways. Thus, robust variogram estimators should be used, as suggested by Genton (1998a), in order to reduce the sensitivity of their variance to outliers.

Acknowledgements

This research is based in part on the author's Ph.D. dissertation at the Swiss Federal Institute of Technology, Lausanne, Switzerland. The author is very grateful to Prof. Stephan Morgenthaler, his Ph.D. thesis advisor, for helpful comments and suggestions, as well as to three anonymous referees and the associate editor for comments that improved the paper.

Appendix

Proof of Theorem 1. Suppose that $\kappa^{\diamond}$ is finite and that there exists some x_0 for which $|\mathrm{IF}(x_0, S, F)|$ is strictly greater than

$$\tfrac{1}{2}[(V^{\diamond 2}(S,F_V)C^{\diamond 2}(\chi,F_V)+4V^{\diamond}(S,F_V)(\kappa^{\diamond}-1-G^{\diamond}(\chi,F_V)))^{1/2} \\ -V^{\diamond}(S,F_V)C^{\diamond}(\chi,F_V)].$$

Without loss of generality, put $x_0 \notin D^{(1)}(\chi)$ and $x_0 > d$. It follows that $\chi(x_0)$ is strictly greater than

$$b=\frac{1}{2}\left[\left(\left(\frac{\tilde{A}(\chi,F_V)C^{\diamond}(\chi,F_V)}{B(\chi,F_V)}\right)^2+4\tilde{A}(\chi,F_V)(\kappa^{\diamond}-1-G^{\diamond}(\chi,F_V))\right)^{1/2} \right. \\ \left. -\frac{\tilde{A}(\chi,F_V)C^{\diamond}(\chi,F_V)}{B(\chi,F_V)}\right],$$

where $\tilde{A}(\chi,F_V)=A(\chi,F_V)+2\sum_{k=1}^{\infty}A^{\diamond}(\chi,F_V^{(k)})$.

If $\chi'(x_0)\leqslant 0$ then

$$1+G^{\diamond}(\chi,F_V)+\frac{\chi^2(x_0)}{\tilde{A}(\chi,F_V)}-2\frac{x_0\chi'(x_0)}{B(\chi,F_V)}+\frac{C^{\diamond}(\chi,F_V)}{B(\chi,F_V)}\chi(x_0) \\ \geqslant 1+G^{\diamond}(\chi,F_V)+\frac{b^2}{\tilde{A}(\chi,F_V)}+\frac{C^{\diamond}(\chi,F_V)}{B(\chi,F_V)}b=\kappa^{\diamond},$$

a contradiction. Therefore, $\chi'(x_0)>0$. Since we have $\chi(x_0)>0$, there exists $\varepsilon>0$ such that $\chi'(t)>0$ for all t in $[x_0,x_0+\varepsilon)$, so $\chi(x)>\chi(x_0)$ for all x in $(x_0,x_0+\varepsilon]$. It follows that $\chi(x)>\chi(x_0)>b$ for all $x>x_0$, $x\notin D^{(0)}(\chi)$ because only upward jumps of χ are allowed for positive x. As $D^{(0)}(\chi)\cup D^{(1)}(\chi)$ is finite, we may assume that $[x_0,+\infty)\cap(D^{(0)}(\chi)\cup D^{(1)}(\chi))$ is empty. It holds that

$$1+G^{\diamond}(\chi,F_V)+\frac{\chi^2(x)}{\tilde{A}(\chi,F_V)}-2\frac{x\chi'(x)}{B(\chi,F_V)}+\frac{C^{\diamond}(\chi,F_V)}{B(\chi,F_V)}\chi(x)\leqslant\kappa^{\diamond}.$$

Therefore

$$\chi^2(x)-2x\chi'(x)\frac{\tilde{A}(\chi,F_V)}{B(\chi,F_V)} \\ \leqslant\tilde{A}(\chi,F_V)(\kappa^{\diamond}-1-G^{\diamond}(\chi,F_V))-\frac{\tilde{A}(\chi,F_V)C^{\diamond}(\chi,F_V)}{B(\chi,F_V)}\chi(x)$$

$$\leqslant \tilde{A}(\chi, F_V)(\kappa^{\diamondsuit} - 1 - G^{\diamondsuit}(\chi, F_V)) - \frac{\tilde{A}(\chi, F_V) C^{\diamondsuit}(\chi, F_V)}{B(\chi, F_V)} b$$
$$= b^2,$$

and thus, for all $x \geqslant x_0$, $\chi^2(x) - 2x\,\chi'(x)\,\tilde{A}(\chi, F_V)/B(\chi, F_V) \leqslant b^2$. Hence

$$\frac{\chi'(x)}{\chi^2(x) - b^2} \geqslant \frac{B(\chi, F_V)}{2\tilde{A}(\chi, F_V)} \frac{1}{x}.$$

Putting

$$R(x) = -\frac{1}{b}\coth^{-1}\left(\frac{\chi(x)}{b}\right) \quad \text{and} \quad P(x) = \frac{B(\chi, F_V)}{2\tilde{A}(\chi, F_V)} \ln(x),$$

it follows that $R'(x) \geqslant P'(x)$ for all $x \geqslant x_0$. Hence $R(x) - R(x_0) \geqslant P(x) - P(x_0)$, and thus

$$\coth^{-1}\left(\frac{\chi(x)}{b}\right) \leqslant b\left[P(x_0) - R(x_0) - \frac{B(\chi, F_V)}{2\tilde{A}(\chi, F_V)} \ln(x)\right].$$

However, the left member is positive because $\chi(x) > b$ and the right member tends to $-\infty$ for $x \to \infty$, a contradiction. This proves the desired inequality. □

Proof of Theorem 2. One of the two inequalities follows from Theorem 1. For the other, assume that S is B-robust. Because χ is monotone, the CVF$^{\diamondsuit}$ can only contain negative delta functions, which do not contribute to $\kappa^{\diamondsuit}$. For all $x \geqslant 0$ it holds that $\chi'(x) \geqslant 0$, so

$$1 + G^{\diamondsuit}(\chi, F_V) + \frac{\chi^2(x)}{\tilde{A}(\chi, F_V)} - 2\frac{x\chi'(x)}{B(\chi, F_V)} + \frac{C^{\diamondsuit}(\chi, F_V)}{B(\chi, F_V)}\chi(x)$$
$$\leqslant 1 + G^{\diamondsuit}(\chi, F_V) + \frac{(\gamma^*)^2}{V^{\diamondsuit}(S, F_V)} + C^{\diamondsuit}(\chi, F_V)\gamma^*.$$

It follows that S is $V^{\diamondsuit}$-robust. □

Proof of Corollary 2.1. The asymptotic variance is

$$V^{\diamondsuit}(S, F_V) = V(S, F_V) + \frac{2\sum_{k=1}^{\infty} A^{\diamondsuit}(\chi, F_V^{(k)})}{B^2(\chi, F_V)} \leqslant (\gamma^*)^2 + \frac{2\sum_{k=1}^{\infty} A^{\diamondsuit}(\chi, F_V^{(k)})}{B^2(\chi, F_V)}.$$

Thus, the change-of-variance sensitivity under dependence is

$$\kappa^{\diamondsuit} = 1 + G^{\diamondsuit}(\chi, F_V) + \frac{(\gamma^*)^2}{V^{\diamondsuit}(S, F_V)} + C^{\diamondsuit}(\chi, F_V)\gamma^*$$
$$= 1 + \frac{1}{V^{\diamondsuit}(S, F_V)}\left(\frac{2\sum_{k=1}^{\infty} A^{\diamondsuit}(\chi, F_V^{(k)})}{B^2(\chi, F_V)} + (\gamma^*)^2\right) + C^{\diamondsuit}(\chi, F_V)\gamma^*$$
$$\geqslant 2 + C^{\diamondsuit}(\chi, F_V)\gamma^*. \qquad □$$

Proof of Lemma 1. This result is a direct consequence of $V^{\diamond}(S,\Phi)\geqslant V(S,\Phi)$ in Genton (1998d). □

Proof of Theorem 3. Using Lemma 1 and $V^{\diamond}(S,\Phi)\geqslant V(S,\Phi)$ in the result of Theorem 2, we obtain

$$\begin{aligned}\kappa^{\diamond} &= 1+G^{\diamond}(\chi,F_V)+\frac{(\gamma^*)^2}{V^{\diamond}(S,\Phi)}+C^{\diamond}(\chi,\Phi)\gamma^*\\ &< 1+1+\frac{(\gamma^*)^2}{V^{\diamond}(S,\Phi)}+C^{\diamond}(\chi,\Phi)\gamma^*\\ &\leqslant 1+1+\frac{(\gamma^*)^2}{V(S,\Phi)}+C^{\diamond}(\chi,\Phi)\gamma^*\\ &\leqslant 1+1+\frac{(\gamma^*)^2}{V(S,\Phi)}+C(\chi,\Phi)\gamma^* = 1+\kappa^*. \qquad \square\end{aligned}$$

Proof of Lemma 2. From the function $\chi(x)=|x|^q-\int|x|^q\,\mathrm{d}F_V(x)$ and its derivative $\chi'(x)=q|x|^{q-1}\,\mathrm{sign}(x)$, we deduce the relations

$$x\chi'(x)=q\chi(x)+B(\chi,F_V) \quad \text{and} \quad x^2\chi''(x)=(q-1)x\chi'(x).$$

Therefore, we have

$$\begin{aligned}C_1(\chi,F_V) &= \int x\chi(x)\chi'(x)\,\mathrm{d}F_V(x)\\ &= q\int\chi^2(x)\,\mathrm{d}F_V(x)+B(\chi,F_V)\int\chi(x)\,\mathrm{d}F_V(x)\\ &= qA(\chi,F_V),\end{aligned}$$

$$\begin{aligned}C_2(\chi,F_V) &= \int x^2\chi''(x)\,\mathrm{d}F_V(x)\\ &= (q-1)B(\chi,F_V),\end{aligned}$$

$$A^{\diamond}(\chi,F_V^{(k)}) = \int\int\chi(x_1)\chi(x_2)\,\mathrm{d}F_V^{(k)}(x_1,x_2),$$

$$\begin{aligned}C_1^{\diamond}(\chi,F_V^{(k)}) &= \int x_1\chi'(x_1)\chi(x_2)\,\mathrm{d}F_V^{(k)}(x_1,x_2)\\ &= \int\int(q\chi(x_1)+B(\chi,F_V))\chi(x_2)\,\mathrm{d}F_V^{(k)}(x_1,x_2)\\ &= q\int\int\chi(x_1)\chi(x_2)\,\mathrm{d}F_V^{(k)}(x_1,x_2)\\ &\quad +B(\chi,F_V)\int\int\chi(x_2)\,\mathrm{d}F_V^{(k)}(x_1,x_2)\\ &= qA^{\diamond}(\chi,F_V^{(k)}).\end{aligned}$$

Inserting these relations in the expression of $C^{\diamondsuit}(\chi, F_V)$ yields

$$C^{\diamondsuit}(\chi, F_V) = 4 - 2\frac{qA(\chi, F_V) + 2q\sum_{k=1}^{\infty} A^{\diamondsuit}(\chi, F_V^{(k)})}{A(\chi, F_V) + 2\sum_{k=1}^{\infty} A^{\diamondsuit}(\chi, F_V^{(k)})} + 2\frac{(q-1)B(\chi, F_V)}{B(\chi, F_V)} = 2. \qquad \square$$

Proof of Theorem 4. As χ is unbounded, the L^q estimator is not B-robust. Moreover, as the CVF$^{\diamondsuit}$ behaves like x^{2q} with a positive factor when $x \to \infty$, it is not bounded from above. Therefore, the L^q estimator is neither V-robust nor $V^{\diamondsuit}$-robust. $\square$

References

Billingsley, P., 1968. Convergence of Probability Measures. Wiley, New York.

Brockwell, P.J., Davis, R.A., 1991. Time Series: Theory and Methods. Springer, New York.

Bustos, O.H., 1982. General M-estimates for contaminated pth order autoregressive processes: consistency and asymptotic normality. Zeit. Wahr. Verw. Geb. 59, 491–504.

Cressie, N., Hawkins, D.M., 1980. Robust estimation of the variogram, I, Math. Geol. 12, 115–125.

Cressie, N., 1985. Fitting variogram models by weighted least squares. Math. Geol. 17, 563–586.

Cressie, N., 1993. Statistics for Spatial Data, 2nd Edition. Wiley, New York.

Doukhan, P., 1994. Mixing: Properties and Examples. Springer, New York.

Gastwirth, J.L., Rubin, H., 1975. The behavior of robust estimators on dependent data. Ann. Statist. 5, 1070–1100.

Genton, M.G., Rousseeuw, P.J., 1995. The change-of-variance function of M-estimators of scale under general contamination. J. Comput. Appl. Math. 64, 69–80.

Genton, M.G., 1998a. Highly robust variogram estimation. Math. Geol. 30, 213–221.

Genton, M.G., 1998b. Variogram fitting by generalized least squares using an explicit formula for the covariance structure. Math. Geol. 30, 323–345.

Genton, M.G., 1998c. Spatial breakdown point of variogram estimators. Math. Geol. 30, 853–871.

Genton, M.G., 1998d. Asymptotic variance of M-estimators for dependent Gaussian random variables. Statist. Probab. Lett. 38, 255–261.

Genton, M.G., 2000. The correlation structure of Matheron's classical variogram estimator under elliptically contoured distributions. Math. Geol. 32, 127–137.

Genton, M.G., He, L., Liu, X., 2001. Moments of skew-normal random vectors and their quadratic forms. Statist. Probab. Lett. 51, 319–325.

Haining, R., 1990. Spatial Data Analysis in the Social and Environmental Sciences. Cambridge University Press, Cambridge.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A., 1986. Robust Statistics: The Approach Based on Influence Functions. Wiley, New York.

Hössjer, O., 1991. The change-of-variance function for dependent data. Probab. Theory Related Fields 90, 447–467.

Huber, P.J., 1964. Robust estimation of a location parameter. Ann. Math. Statist. 35, 73–101.

Huber, P.J., 1967. The behavior of maximum likelihood estimates under non-standard conditions. Proceedings of the Fifth Berkeley Symposium on Mathematics, Statistics and Probability vol. 1, pp. 221–233.

Huber, P.J., 1972. Robust statistics: a review. Ann. Math. Statist. 43, 1041–1067.

Huber, P.J., 1981. Robust Statistics. Wiley, New York.

Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford Press, New York.

Journel, A.G., Huijbregts, Ch.J., 1978. Mining Geostatistics. Academic Press, London.

Matheron, G., 1962. Traité de Géostatistique Appliquée, Tome I, Mémoires du Bureau de Recherches Géologiques et Minières, Vol. 14. Paris.

Portnoy, S., 1977. Robust estimation in dependent situations. Ann. Statist. 5, 22–43.

Portnoy, S., 1979. Further remarks on robust estimation in dependent situations. Ann. Statist. 7, 22–43.
Rousseeuw, P.J., 1981. A new infinitesimal approach to robust estimation. Zeit. Wahr. Verw. Geb. 56, 127–132.
von Mises, R., 1937. Sur les fonctions statistiques. Conférences de la Réunion Internationale des Mathématiques, Gauthier-Villars, pp. 1–8.
von Mises, R., 1947. On the asymptotic distribution of differentiable statistical functions. Ann. Math. Statist. 18, 309–348.