LMD mokslo darbai, 355–358
1998, Vilnius

Comparisons of spatial prediction methods for stationary Gaussian random fields

K. Dučinskas, Ž. Vainilavičius (KU)

Suppose the spatial data $Z(s_1), \dots, Z(s_n)$ observed at spatial locations $\{s_1, \dots, sn\}$ are modelled as a collection of random variables generated by the random field

$$Z(s) = \mu + \delta(s), \quad s \in D \subset R, \tag{1}$$

where $\delta(s)$ is a second-order stationary zero mean finite-variance Gaussian random field with covariance function $\operatorname{cov}(\delta(u), \delta(s)) = C(u - s)$, for all $u, s \in D$. The function $C(\cdot)$ is usually called a covariogram.

Then $Z(s)$ is a second-order stationary Gaussian field with $E(Z(s)) = \mu$, $\operatorname{Var}(Z(s)) = C(0)$, for all $s \in D$.

Spatial prediction refers to predicting the value of $Z(s_0)$ at known spatial location $s_0 \in D$ from data $Z(s_1), \dots, Z(s_n)$ observed at known spatial locations $s_1, \dots, s_n$. Let $\Sigma = \|\operatorname{cov}(Z(s_i), Z(s_j))\|_{i,j=1,\dots,n}$, $c' = (C(s_0 - s_1), \dots, C(s_0 - s_n))$. The prime always will denote vector transpose in this paper. It has been seen, that for squared-error loss, the best predictor for Gaussian field $Z(s)$ is the linear predictor

$$p_{sk}(s_0) = \sum_{i=1}^{n} l_i Z(s_i) + k, \tag{2}$$

where

$$l_i = c'\Sigma^{-1}, \quad k = \mu(1 - c'\Sigma^{-1} 1_n) \tag{3}$$

and $1'_n = (1, \dots, 1)$.

Definition. The mean-squared prediction error (MSPE) for any spatial predictor $p(s_0)$ is defined by

$$\mathrm{MSPE}(p(s_0)) = E(Z(s_0) - p(s_0))^2. \tag{4}$$

Matheron (1967) has called such spatial prediction simple kriging (sk). This is a stochastic method of spatial prediction that minimises the MSPE. Sometimes for the linear prediction of $Z(s_0)$ the interpolation method based on inverse distance (id) are used i.e.

$$p_{id}(s_0) = \sum_{i=1}^{n} a_i(s_0) Z(s_i), \tag{5}$$

where

$$a_i(s_0) = \frac{1/(d_i)^2}{\sum_{i=1}^{n} 1/(d_i)^2}, \tag{6}$$

and d_i is a Euclidean distance between s_i and s_0. This method of nonstochastic spatial prediction was used by authors of this paper in the analysis of benthic communities of the Curonian lagoon (see Razinkovas et al., 1997).

The third investigated method of spatial prediction is the method based on measure of central tendency (ct) (see Cressie p. 370, 1991), i.e. predictor $p_{ct}(\delta_0)$ is given by

$$p_{ct}(s_0) = \sum_{i=1}^{n} Z(s_i)/n \tag{7}$$

This is simple, computationally fast but not resistant to outliers method.

THEOREM. *The mean-squared prediction errors for the three predictors of the second-order stationary Gaussian field* $Z(s)$ *defined by* (1) *are*

$$\mathrm{MSPE}(p_{sk}(s_0)) = C(0) - c'\Sigma^{-1}c, \tag{8}$$

$$\mathrm{MSPE}(p_{id}(s_0)) = C(0) + a'\Sigma a - 2a'c, \tag{9}$$

$$\mathrm{MSPE}(p_{ct}(s_0)) = C(0) + (1_n'\Sigma 1_n)/n^2 - 21_n'c/n, \tag{10}$$

where $a' = (a_1(s_0), \ldots, a_n(s_0))$.

Proof. The proof of stated theorem was completed by direct application of the definition of MSPE for considered spatial predictors of Gaussian field $Z(s)$ defined in (1).

The performance of the spatial predictors could be compared by calculating the values of the following coefficients of relative efficiency

$$k_1 = \frac{\mathrm{MSPE}(p_{sk}(s_0))}{\mathrm{MSPE}(p_{id}(s_0))}, \quad k_2 = \frac{\mathrm{MSPE}(p_{sk}(s_0))}{\mathrm{MSPE}(p_{ct}(s_0))}, \quad k_3 = \frac{\mathrm{MSPE}(p_{id}(s_0))}{\mathrm{MSPE}(p_{ct}(s_0))} \tag{11}$$

Example. Suppose that field $Z(s)$ is defined on the 2 dimensional integer rectangular lattice Z^2 and covariogram is isotropic i.e.,

$$C(h) = C(|h|) = C(0)e^{-\alpha|h|}. \tag{12}$$

The following two neighbourhood schemesare considered.

Scheme A. The first-order neighbourhood with 4 neighbours i.e., $n = 4$.

Scheme B. The second-order neighbourhood with 8 neighbours i.e., $n = 8$.

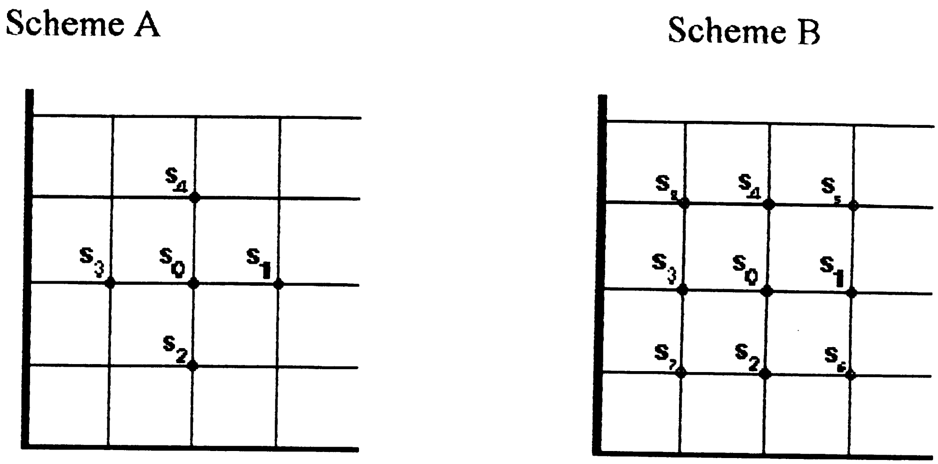

Figure 1. Two schemes of neighbourhood of s_0.

These schemes are presented in Figure 1.

Further the superscripts A and B are used to identify the considered schemes of neighbourhood.

Then, by using (8), (9), (10) the following formulas was obtained

$$\mathrm{MSPE}(p_{sk}^{A}(s_0)) = C(0) - 4e^{(-2\alpha)}/(2e^{(-\alpha\sqrt{2})} + e^{(-2\alpha)} + 1), \tag{13}$$

$$\mathrm{MSPE}(p_{id}^{A}(s_0)) = \mathrm{MSPE}(p_{ct}^{A}(s_0)) = C(0) + (1 + 2e^{(-\alpha\sqrt{2})} + e^{(-2\alpha)})/4, \tag{14}$$

$$\begin{aligned}\mathrm{MSPE}(p_{sk}^{B}(s_0)) = C(0) - 4\big\{&(2e^{(-4\alpha)} + 2e^{(-2\alpha(1+\sqrt{2}))} - 4e^{(-\alpha(\sqrt{2}+2))}\\ &+ e^{(-2\alpha)} - 4e^{(-\alpha(1+\sqrt{2}+\sqrt{5}))} + 2e^{(-3\alpha\sqrt{2})} + e^{(-2\alpha\sqrt{2})})/(-2e^{(-4\alpha)} - 4e^{(-\alpha(\sqrt{2}+2))}\\ &+ e^{(-2\alpha)} - e^{(-2\alpha(1+\sqrt{2}))} + 8e^{(-\alpha(1+\sqrt{5}))} - e^{(-2\alpha\sqrt{2})} + 4e^{(-2\alpha\sqrt{5})} - 1 - 2e^{(-\alpha\sqrt{2})}\\ &- 2e^{(-3\alpha\sqrt{2})})\big\},\end{aligned} \tag{15}$$

$$\begin{aligned}\mathrm{MSPE}(p_{id}^{B}(s_0)) = C(0) + (&-8e^{(-\alpha\sqrt{2})} - 144e^{(-\alpha)} + e^{(-2\alpha\sqrt{2})} + 18e^{(-2\alpha)}\\ &+ 16e^{(-\alpha\sqrt{5})} + 17)/100,\end{aligned} \tag{16}$$

$$\begin{aligned}\mathrm{MSPE}(p_{ct}^{B}(s_0)) = C(0) + (&2e^{(-\alpha\sqrt{2})} + 4e^{(-\alpha)} + e^{(-2\alpha\sqrt{2})} + 3e^{(-2\alpha)}\\ &+ 4e^{(-\alpha\sqrt{5})} + 2)/16,\end{aligned} \tag{17}$$

The coefficients of relative efficiency $\{k_l\}_{i=1}^{3}$ defined in (11) for two neighbourhood schemes are presented in Table 1. It is obvious that $p_{id}(s_0) = p_{ct}(s_0)$ for scheme A and so $k_3^A = 1$.

From Table 1, it can be concluded, that the relative superiority of the $p_{sk}(s_0)$ against other two spatial predictors decreases for scheme A and increases for scheme B as spatial dependency is decreasing faster i.e., when α is increasing.

Table 1

Coefficients of relative efficiency for schemes A and B.

α	k_1^A	k_2^A	k_1^B	k_2^B	k_3^B
0,50	0,998529	0,998529	0,562965	0,526143	0,934592
1,00	0,994816	0,994816	0,987056	0,945278	0,957674
1,50	0,965122	0,965122	0,984985	0,964723	0,979428
2,00	0,923112	0,923112	0,956481	0,953801	0,997198
2,50	0,884408	0,884408	0,926167	0,936050	1,010670
3,00	0,855085	0,855085	0,901827	0,920211	1,020385
3,50	0,834948	0,834948	0,884670	0,908675	1,027135
4,00	0,821814	0,821814	0,873371	0,901050	1,031692
4,50	0,813485	0,813485	0,866197	0,896255	1,034702
5,00	0,808287	0,808287	0,861733	0,893319	1,036655
5,50	0,805073	0,805073	0,858986	0,891545	1,037904
6,00	0,803098	0,803098	0,857306	0,890480	1,038696
6,50	0,801889	0,801889	0,856283	0,889843	1,039192
7,00	0,801150	0,801150	0,855661	0,889461	1,039502
7,50	0,800700	0,800700	0,855283	0,889233	1,039694
8,00	0,800425	0,800425	0,855054	0,889096	1,039812
8,50	0,800258	0,800258	0,854915	0,889013	1,039885
9,00	0,800157	0,800157	0,854831	0,888964	1,039930
9,50	0,800095	0,800095	0,854780	0,888934	1,039957
10,00	0,800058	0,800058	0,854749	0,888916	1,039974

REFERENCES

[1] Cressie N. A. C., *Statistics for spatial data*, Wiley, New York, 1993.

[2] Matheron G., Kriging or polynomial interpolation procedures? *Transactions of the Canadian Institute of Mining and Metallurgy*, **70** (1967), 240–244.

[3] Razinkovas A., Dučinskas K., and Vainilavičius Ž., Comparison of different spatial prediction methods applied to the benthic communities of the Curonian lagoon, *Aplinkos tyrimai, inžinierija ir vadyba*, **2**(5) (1997), 22–28.

Stacionarių Gauso laukų prognozavimo metodų palyginimas

K. Dučinskas, Ž. Vainilavičius

Gautos stacionarių Gauso laukų stochastinių ir nestochastinių prognozavimo metodų vidutinių kvadratinių prognozavimo klaidų analitinės išraiškos. Atliktas skaitinis prognozavimo metodų palyginimas izotropinės kovariacijos atvejui, naudojant pirmos ir antros eilės kaimynų schemas taisyklingose stačiakampėse gardelėse.