Mathematical Geology, Vol. 15, No. 2, 1983

Two Methods with Different Objectives: Splines and Kriging[1]

Olivier Dubrule[2]

When drawing a contour map from a set of irregularly spaced data points, two methods are often used: The first corresponds to a rather aesthetic criterion and consists of obtaining contour lines which will be as "smooth" as possible and will honor the data points. This generally is the objective of the draftsman, and it can be automatically performed by the method of spline interpolation. The other method, used in kriging, is to compute the Best Linear Unbiased Estimator (B.L.U.E.), that is, to obtain a map as accurate as possible. Is it possible, in practice, to predict whether the aesthetic map will also be accurate? In this paper, we first examine the theoretical point of view: Spline interpolation is equivalent to kriging with a given (generalized) covariance. We then take an example to show how this question can be answered in practice: by testing how well the spline covariance is suited to the data.

KEY WORDS: kriging, splines, measurement errors, generalized covariance, structural identification.

INTRODUCTION

In automatic cartography, the first step is to interpolate the variable onto the nodes of a regular grid. Kriging and splines have two different points of view for doing this. Kriging computes an estimate which is on average as accurate as possible. In contrast, splines work on the shape of the interpolating function.

KRIGING AND SPLINES

Kriging

Consider a linear combination λ assigning the weights λ^α to the points x_α. λ is called a generalized increment of order k (k - GI) if it filters all the mono-

[1]Manuscript received 8 February 1982; revised 2 April 1982. Paper presented at the First MGUS Conference on the Management Analysis and Display of Geoscience Data, Golden, Colorado, 27-29 January 1982.

[2]Centre de Geostatistique et de Morphologie Mathematique Ecole Nationale Superieure des Mines de Paris, Fontainebleau, France. Present address: Sohio Petroleum, 100 Pine Street, San Francisco, California 94111 U.S.A.

0020-5958/83/0400-0245$03.00/0

mials of degree up to k in the coordinates of the points $x_\alpha = (x_{\alpha 1}, x_{\alpha 2})$

$$\sum_\alpha \lambda^\alpha x_{\alpha 1}^m x_{\alpha 2}^n = 0, \qquad \text{for all } m, n \text{ such that } 0 \leqslant m + n \leqslant k \tag{1}$$

$Z(x)$ is called an intrinsic random function of order k (k - IRF) if its k - GI are stationary (in the weak sense) with zero mean. If $Z(x)$ is a k - IRF, there exists (Matheron, 1973) a function $K(h)$ such that, for all k - GI λ

$$E[Z(\lambda)] = E\left[\sum_\alpha \lambda^\alpha Z(x_\alpha)\right] = 0$$

$$\text{Var } [Z(\lambda)] = E[Z^2(\lambda)] = \sum_\alpha \sum_\beta \lambda^\alpha \lambda^\beta K(x_\alpha - x_\beta) \tag{2}$$

$K(h)$ is called the generalized covariance (GC) of $Z(x)$. So, a k - IRF is characterized by

- its order k, which is interpreted as the degree of its local drift
- its GC $K(h)$, which gives the statistical structure of the variable once its (local) drift has been filtered.

The program BLUEPACK-3D automatically determines k and $K(h)$, in the step called "automatic structure identification." The model of covariance which is fitted to the variable is isotropic and polynomial. Its expression depends on the value of k

$k = 0$ $\quad K(h) = C_o\delta(h) - b_o|h|, \qquad C_o \geqslant 0, \quad b_o \geqslant 0$

$k = 1$ $\quad K(h) = C_o\delta(h) - b_o|h| + b_1|h|^3,$

$$C_o \geqslant 0, \quad b_o \geqslant 0, \quad b_1 \geqslant 0$$

$k = 2$ $\quad K(h) = C_o\delta(h) - b_o|h| + b_1|h|^3 - b_2|h|^5,$

$$C_o \geqslant 0, \quad b_o \geqslant 0, \quad b_2 \geqslant 0, \quad b_1 \geqslant -(10/3)(b_o b_2)^{1/2}$$

$C_o\delta(h)$ is the nugget effect: $\delta(h) = 0$ if $h \neq 0$ and $\delta(0) = 1$. The constraints on the coefficients ensure that the expression (2) is always positive.

Once k and $K(h)$ are estimated, many problems can be treated.

Kriging

$Z(x)$ is estimated by a weighted average of the values at neighboring data points

$$Z_K(x) = \sum_\alpha \lambda^\alpha Z(x_\alpha)$$

The error

$$Z(x) - \sum_{\alpha} \lambda^{\alpha} Z(x_{\alpha})$$

is a k - GI if [see (1)]

$$\sum_{\alpha} \lambda^{\alpha} x_{\alpha 1}^{m} x_{\alpha 2}^{n} = x_1^m x_2^n, \qquad \text{for all } m, n, \quad 0 \leqslant m + n \leqslant k \tag{3}$$

From (2), its variance is

$$\text{Var } [Z(x) - Z_K(x)] = K(o) - 2 \sum_{\alpha} \lambda^{\alpha} K(x_{\alpha} - x) + \sum_{\alpha} \sum_{\beta} \lambda^{\alpha} \lambda^{\beta} K(x_{\alpha} - x_{\beta})$$

It is possible to compute the λ^{α} which minimize this variance under the constraints (3). We obtain an estimate which is as accurate as possible on average, and we also have the value of Var $[Z(x) - Z_K(x)]$, the "kriging variance," which informs us about the accuracy of the estimate at the point x (Delfiner, 1975, Chilès, 1977).

Cokriging

When the values at data points are observed with a random error $\epsilon(x)$, one only knows the measures

$$Y(x_{\alpha}) = Z(x_{\alpha}) + \epsilon(x_{\alpha}) \tag{4}$$

In order to estimate $Z(x)$ from the $Y(x_{\alpha})$, we "cokrige" the variable $Z(x)$ using the variable $Y(x)$. If $\epsilon(x)$ is such that

$$\begin{cases} E[\epsilon(x_{\alpha})] = 0 & \text{for all } \alpha \\ \text{Cov } [\epsilon(x_{\alpha}), \epsilon(x_{\beta})] = S_{\alpha}^2 \delta_{\alpha\beta}, & \text{for all } \alpha, \beta, \\ \text{Cov } [\epsilon(x_{\alpha}), Z(x_{\beta})] = 0 & \text{for all } \alpha, \beta \end{cases} \qquad \left(\delta_{\alpha\beta} = \begin{cases} 0, & \text{if } \alpha \neq \beta \\ 1, & \text{if } \alpha = \beta \end{cases} \right)$$

we get

$$\text{Var} \left[Z(x) - \sum_{\alpha} \lambda^{\alpha} Y(x_{\alpha}) \right] = \text{Var} \left[Z(x) - \sum_{\alpha} \lambda^{\alpha} Z(x_{\alpha}) \right] + \sum_{\alpha} \lambda^{\alpha^2} S_{\alpha}^2$$

Then it is easy to compute the λ^{α} minimizing this variance under the constraints that $Z(x) - \sum_{\alpha} \lambda^{\alpha} Z(x_{\alpha})$ is a k - GI [see (3)].

Splines

Only the "thin plate" splines, which are the most often used in practice, are considered here. The general theory of splines is presented in Laurent (1972) and Ahlberg et al. (1967).

Interpolating Splines

When using spline interpolation, one wants to obtain contour lines as "smooth" as possible; that is, a map which looks like what a draftsman would obtain manually. The method for that is to compute a function having the shape of a thin plate ("plaque mince," in French) which would be forced to pass through the data points. One assigns the value $\sigma(x)$ to each point x, where σ is among all the functions f passing through data points, the one minimizing

$$A(f) = \iint_{\mathbb{R}^2} \left[\left(\frac{\partial^2 f}{\partial x_1} \right)^2 + \left(\frac{\partial^2 f}{\partial x_2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x_1 \partial x_2} \right)^2 \right] dx_1 \, dx_2$$

$A(f)$ is equal, in a first approximation, to the bending energy of a thin plate represented by f.

Smoothing Splines

When there is a measurement error on the data [see (4)], the function $\sigma(x)$ should pass "not too far" from the experimental values. The smoothing spline is the function f which minimizes

$$A(f) + \rho \sum_{\alpha} \left[\frac{f(x_\alpha) - y(x_\alpha)}{S_\alpha} \right]^2$$

The "smoothing parameter" ρ determines the relative importance of each term

- the first term represents the "smoothness" of f
- the second represents the proximity of f to the data

Comparing Splines and Kriging

Matheron (1980) has shown that, for a given set of data, the interpolating spline is equivalent to kriging in k - IRF with[3]

$$k = 1, \qquad K(h) = |h|^2 \text{ Log } |h|$$

As for the smoothing spline, it is equal to cokriging in k - IRF with (Dubrule, 1981b)

$$k = 1, \qquad K(h) = \rho |h|^2 \text{ Log } |h|$$

---

[3]The definition of $K(h)$ for $h = 0$ does not lay any problem

$$\left\langle K(o) = \lim_{h \to 0} |h|^2 \log |h| = 0 \right\rangle$$

A practical test has also shown (Dubrule, 1981a) that both methods attain their ends: kriging is accurate, while spline interpolation gives some smooth contour lines.

At this stage, it may be asked whether it is possible to find a way to combine the qualities of both methods.

MODIFYING THE AUTOMATIC STRUCTURE IDENTIFICATION (ASI)

The Fit of the Spline Covariance

Here we suppose that the first part of the ASI has given $k = 1$. In this case the model for GC prescribed and fitted by BLUEPACK-3D is

$$K(h) = C_o\delta(h) - b_o|h| + b_1|h|^3, \qquad C_o \geqslant 0, \quad b_o \geqslant 0, \quad b_1 \geqslant 0 \tag{5}$$

In addition, it might be interesting to fit a model of the form

$$K(h) = b_s|h|^2 \text{ Log } |h|, \qquad b_s \geqslant 0 \tag{6}$$

Kriging with such a covariance will be equal to the interpolating spline, since the kriging system does not change if $K(h)$ is multiplied by a constant (only the kriging variance is multiplied by this constant). The constraint on b_s ensures that $K(h)$ is an acceptable model of GC.

If model (6) is the best one for the variable under study, it is clear that by using it for kriging, the user will obtain a map which will be both aesthetic and accurate. Moreover, he will be able to compute the kriging standard deviation map. If model (6) is not well suited to the variable, the user will know that he may lose some precision using spline interpolation.

Calculation of the Smoothing Parameter ρ

Users of smoothing splines always have the problem of choosing the value of ρ. Some methods based on cross-validation have been developed to determine ρ (Utreras, 1979, Wahba, 1979). Although they can give good results, these methods are heavy to use in practice.

When the covariance (6) is well suited to the structure of the variable $Z(x)$, the problem of choosing ρ can be solved as follows: if the $Z(x_\alpha)$ are measured with a random error [see (4)] we know that cokriging with model (6) is equivalent to the smoothing spline obtained with

$$\rho = b_s$$

So this value of ρ gives an interpolating function which is

- accurate (since it is cokriging)
- aesthetic (since it is a spline function)

Fitting a Spline Covariance with a Nugget Effect

In the previous section, we supposed that the covariance

$$K(h) = b_s |h|^2 \text{ Log } |h|$$

of $Z(x)$ was known. In fact, we implicitly assumed that some of the data had no measurement errors, and that a first ASI on them had given $K(h)$.

What happens when the error affects all the data? The ASI can only be performed on the variable:

$$Y(x) = Z(x) + \epsilon(x)$$

If the variance of $\epsilon(x)$ is the same for all the data

$$\text{Cov } [\epsilon(x_\alpha), \epsilon(x_\beta)] = C_o \delta_{\alpha\beta}$$

we have, for all k - GI λ

$$\begin{aligned}
\text{Var} \left[\sum_\alpha \lambda^\alpha Y(x_\alpha) \right] &= \text{Var} \left[\sum_\alpha \lambda^\alpha Z(x_\alpha) \right] + \text{Var} \left[\sum_\alpha \lambda^\alpha \epsilon(x_\alpha) \right] \\
&= \sum_\alpha \sum_\beta \lambda^\alpha \lambda^\beta K(x_\alpha - x_\beta) + C_o \sum_\alpha \lambda^{\alpha^2} \\
&= \sum_\alpha \sum_\beta \lambda^\alpha \lambda^\beta [K(x_\alpha - x_\beta) + C_o \delta(x_\alpha - x_\beta)]
\end{aligned}$$

So the GC of $Y(x)$ is obtained by adding the nugget effect C_o to the covariance $K(h)$ of $Z(x)$. That is why it might be interesting to apply the ASI, not only to model (6), but also to the covariance

$$K(h) = C_o \delta(h) + b_s' |h|^2 \text{ Log } |h|$$

If this model is the best one, the nugget effect can be explained in two different ways

- either the data are affected by an error having the constant variance C_o
- or there is a "microstructure" that can be modeled by a variogram with a sill C_o and a range smaller than the shortest distance between two data.

In the first case, we use cokriging with

$$\text{Var } [\epsilon(x_\alpha)] = C_o, \qquad \text{for all } \alpha \quad \text{and} \quad K(h) = b_s' |h|^2 \text{ Log } |h|$$

that is equivalent to the smoothing spline obtained with $\rho = b_s'$. In the other case, there is no error, and without additional information, we krige $Z(x)$ with

$$K(h) = C_o \delta(h) + b_s' |h|^2 \text{ Log } |h|$$

It is easy to show that both estimates are equal except at data points where the latter is equal to the experimental value, which is not the case for the former.

So they will give the same map, since a point of the estimated grid is almost never merged with a datum point.

A New Model for $K(h)$

When $k = 1$, the ASI might be performed on the model

$$K(h) = C_o\delta(h) - b_o|h| + b_s|h|^2 \text{ Log } |h| + b_1|h|^3 \tag{7}$$

Such a model is more general than (5) and, cancelling b_o and b_1, one finds the particular case of spline covariances. The constraints on the coefficients are (Dubrule, 1981b)

$$C_o \geqslant 0, \qquad b_o \geqslant 0, \qquad b_1 \geqslant 0, \qquad b_s \geqslant -\tfrac{3}{2}(b_o b_1)^{1/2} \tag{8}$$

A Practical Example

The variable under study is the thickness of a geological layer, measured at 60 wells (Fig. 1). The output of the ASI performed on this variable is given in Fig. 2.

Here we do not explain the first part of the ASI, that is the way to determine k. In short, the order k which is chosen is the one that gives the best fit of the data using local trend surface polynomials of degree k (Delfiner 1975, Chilès 1977). In our example, the value $k = 1$ has been found. So we can fit model (7).

Using (2), we get for all $1 - GI\,\lambda$

$$E\left[\sum_\alpha \lambda^\alpha Z(x_\alpha)\right]^2 = C_o\left(\sum_\alpha \lambda^{\alpha^2}\right) + b_o K_o(\lambda) + b_s K_s(\lambda) + b_1 K_1(\lambda) \tag{9}$$

where $K_o(\lambda)$, $K_s(\lambda)$, and $K_1(\lambda)$ are the variances corresponding to each elementary covariance. For instance

$$K_s(\lambda) = \sum_\alpha \sum_\beta \lambda^\alpha \lambda^\beta |x_\alpha - x_\beta|^2 \text{ Log } |x_\alpha - x_\beta|$$

(9) is a regression equation of $Z^2(\lambda)$ on the four variables C_o, b_o, b_s, and b_1. The method for determining these coefficients is to compute a large number (about 1000) $Z(\lambda_m) = \sum_\alpha \lambda_m^\alpha Z_\alpha$, for λ_m satisfying the relations (1), and to minimize

$$Q(C_o, b_o, b_s, b_1) = \sum_m w_m^2 \left[Z^2(\lambda_m) - C_o\left(\sum_\alpha \lambda_m^{\alpha^2}\right) - b_o K_o(\lambda_m) - b_s K_s(\lambda_m) - b_1 K_1(\lambda_m)\right]^2$$

The weights w_m^2 are introduced to equalize the variances of the $Z^2(\lambda_m)$ (Delfiner, 1975).

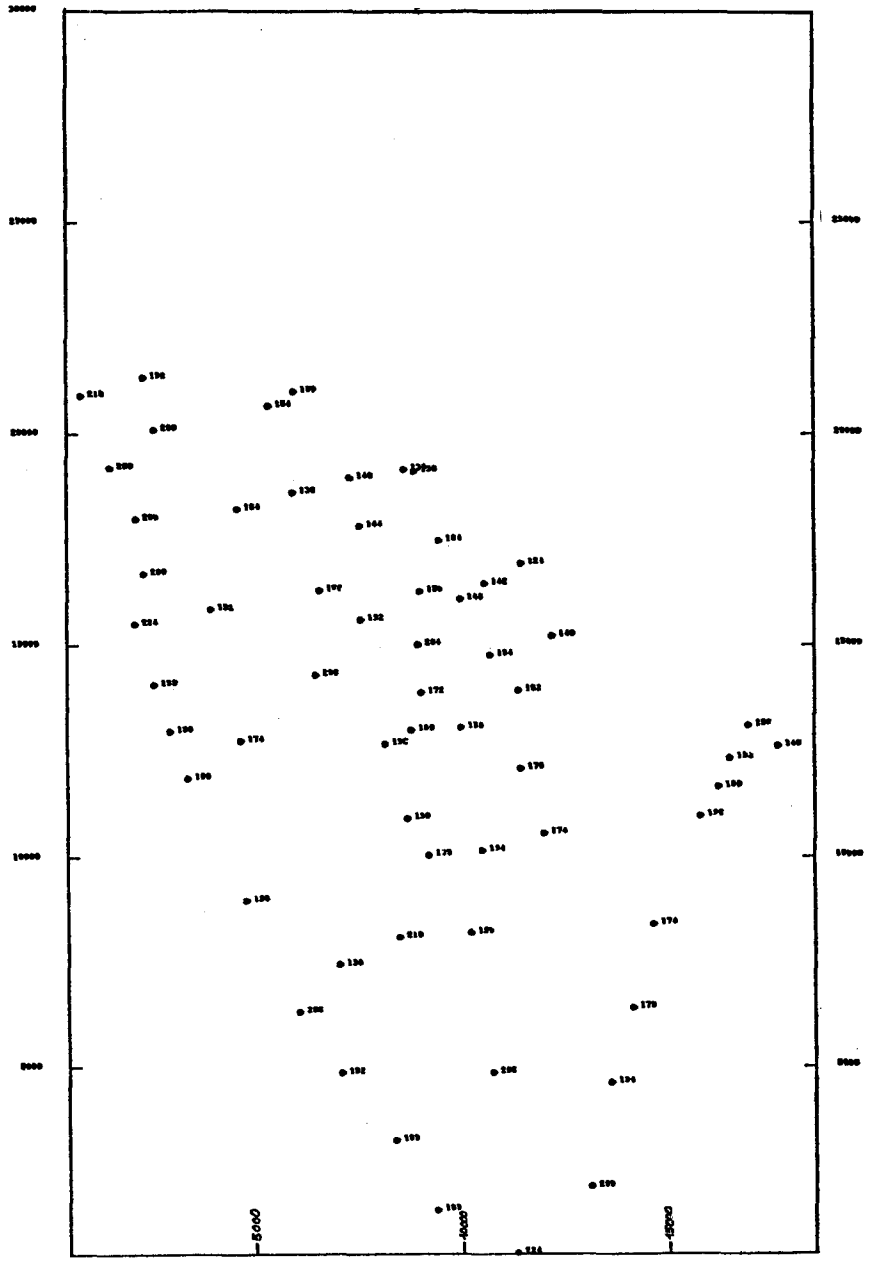

Fig. 1. Thickness of a geological layer measured at 60 wells: unit = 1 m; scale = ~1/125,000.

```
**********************************************
*  AUTOMATIC  STRUCTURE  IDENTIFICATION      *
**********************************************

1) IDENTIFICATION OF THE DRIFT
******************************
```

DEGREE	MEAN SQUARED ERRORS RING 1	MEAN SQUARED ERRORS RING 2	MEAN SQUARED ERRORS TOTAL	RANKING OF THE TRIALS AVERAGE RANKS RING 1	RANKING OF THE TRIALS AVERAGE RANKS RING 2	RANKING OF THE TRIALS AVERAGE RANKS TOTAL
2	0.2895E+04	0.1645E+04	0.2260E+04	2.29	2.08	2.19
1	0.5743E+03	0.4887E+03	0.5308E+03	1.80	1.85	1.82
0	0.5992E+03	0.6298E+03	0.6147E+03	1.91	2.07	1.99

```
DEGREE OF THE DRIFT = 1

COUNTS  RING 1:  539  RING 2:  556  TOTAL: 1095
MEAN NEIGHBORHOOD RADIUS : 0.7285E+04

2) IDENTIFICATION OF THE COVARIANCE
***********************************
```

COEFFICIENTS OF THE COVARIANCE NUGGET	LINEAR	SPLINE	CUBIC	RATIOS OF EXPERIMENTAL TO THEORETICAL VARIANCES RES/TOT	E1/T1	E2/T2	E/T	JACKKNIFE
0.1836E+03	0.0000E+00	0.0000E+00	0.0000E+00	0.6703E+00	1.7379	1.6681	1.7045	1.7066
0.0000E+00	-0.1570E+00	0.0000E+00	0.0000E+00	0.6279E+00	0.7576	0.7696	0.7632	0.7627
0.2154E+02	-0.1410E+00	0.0000E+00	0.0000E+00	0.6273E+00	0.7984	0.8085	0.8031	0.8027
0.0000E+00	0.0000E+00	0.0000E+00	0.6405E-08	0.7910E+00	0.4607	0.4677	0.4640	0.4636
0.1497E+03	0.0000E+00	0.0000E+00	0.2477E-08	0.6503E+00	0.7641	0.7599	0.7621	0.7623
0.0000E+00	0.0000E+00	0.3853E-04	0.0000E+00	0.6932E+00	0.5121	0.5270	0.5189	0.5182
0.1160E+03	0.0000E+00	0.1904E-04	0.0000E+00	0.6401E+00	0.7528	0.7597	0.7560	0.7557

```
BEST COVARIANCE FIT
-------------------

ORDER K= 1    C= 0.2154E+02    A1=-0.1410E+00    A2= 0.0000E+00    A3= 0.0000E+00
```

Fig. 2. Output of the BLUEPACK-3D automatic structure identification performed on the variable of Fig. 1.

But nothing ensures that the regression coefficients will be in accordance with the constraints (8). That is why BLUEPACK-3D runs as many regressions as there are possible models of GC, cancelling in turn one, two, or three coefficients of $K(h)$. At least, models with only one term will be acceptable; for instance, the coefficient of the elementary spline covariance will be

$$b_s = \frac{\sum_m w_m^2 K_s(\lambda_m) Z^2(\lambda_m)}{\sum_m w_m^2 K_s^2(\lambda_m)}$$

b_s is always positive, as $K(h) = |h|^2 \text{ Log } |h|$ is a GC for $k \geqslant 1$ (Dubrule, 1981b).

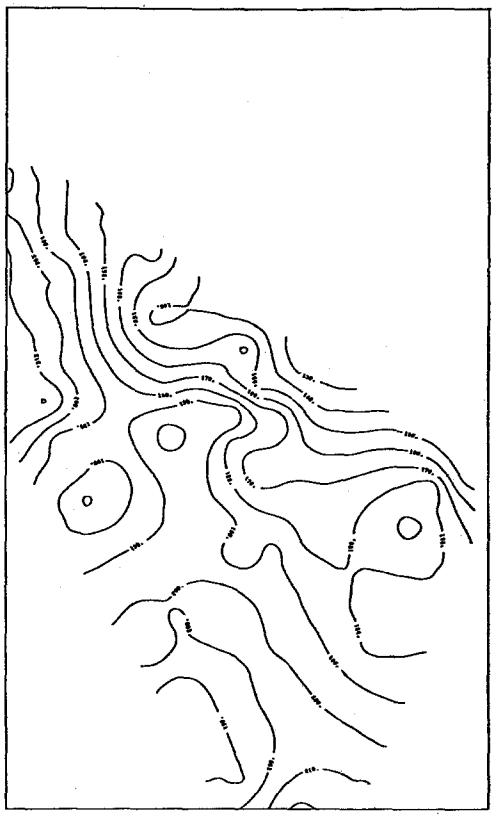

Fig. 3. Map obtained with $k = 1$ and $K(h) = 21.54\ \delta(h) - 0.141\,|h|$. Elementary mesh of the kriged grid (500 × 500 m). Each grid point is kriged using its 24 nearest neighbors: unit = 1 m; scale = ~1/125,000.

In our example, seven models are applicable. In Figure 2, each line corresponds to one model. The printed coefficients are successively: NUGGET = C_o, LINEAR = $-b_o$, SPLINE = b_s, and CUBIC = b_1. The question is: How to choose the best one among these models?

- The first criterion is the value of $Q(C_o, b_o, b_s, b_1)$. The column RES/TOT gives the value of

$$\frac{Q(C_o, b_o, b_s, b_1)}{Q(0, 0, 0, 0)}$$

This ratio is smaller than 1, and it is clear that the lower it is, the better the model.

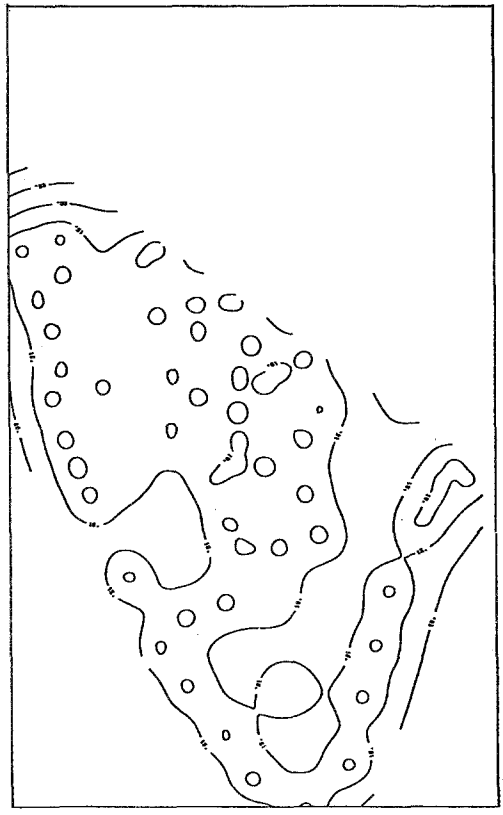

Fig. 4. Kriging standard deviation map obtained with $k = 1$ and $K(h) = 21.54\,\delta(h) - 0.141\,|h|$. Legend as in Fig. 3.

- The final criterion used in BLUEPACK-3D is more related to our final objective (that is, to obtain a good fit for the variances of the k - GI). If $\hat{K}(h)$ is the estimator of $K(h)$, we want the ratio

$$\rho = \frac{\sum_m E[Z^2(\lambda_m)]}{\sum_m E[\hat{K}(\lambda_m)]}$$

to be close to one. The JACKKNIFE is an estimate of ρ.

In our example, the model

$$K(h) = 21.54\delta(h) - 0.141|h|$$

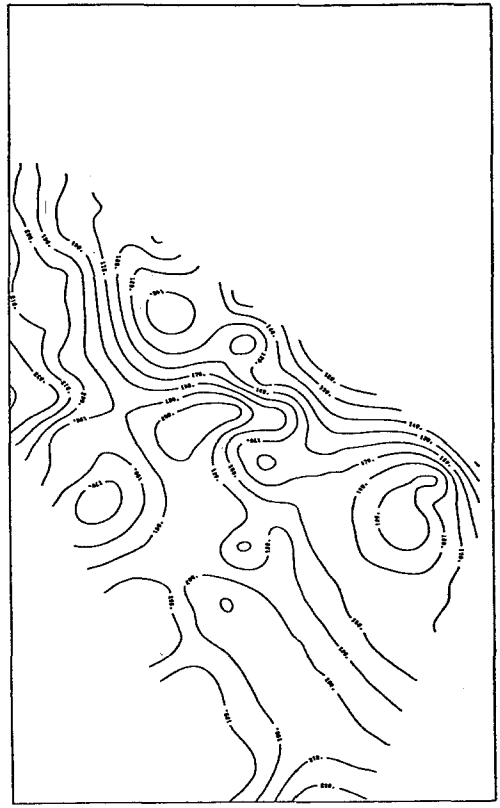

Fig. 5. Map obtained with $k = 1$ and $K(h) = 0.3853 \times 10^{-4}\ |h|^2$ Log $|h|$. Legend as in Figs. 3 and 4.

obtains the best JACKKNIFE (0.8027), and also the lowest value of RES/TOT. The map obtained with $K(h)$ is given in Fig. 3, and Fig. 4 is the corresponding map of kriging standard deviations. The spline covariance

$$K(h) = 0.3853 \times 10^{-4} |h|^2 \text{ Log } |h|$$

obtains a JACKKNIFE of 0.5182: we would expect the spline map (Fig. 5), although closer to a draftsman's map, to be less precise than the kriged one.

REFERENCES

Ahlberg, Nilson, and Walsh, The theory of splines and its applications, Academic Press, New York, 280 p.

Chilès, J. P., 1977, Géostatistique des phénomènes non stationnaires (dans le plan), "Docteur-Ingénieur" thesis: Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau, 152 p.

Delfiner, P., 1975, Linear estimation of non-stationary spatial phenomena, Proceedings of NATO-ASI advanced geostatistics in the mining industry, Rome, 1975: D. Reidel Publishing Co., Dordrecht, Holland, p. 49–68.

Dubrule, O., 1981a, Comparing "Plaque mince" splines and kriging, Internal Report, Centre de Géostatistique: Ecole des Mines de Paris, Fontainebleau, 24 p. (to be published in "Computers and Geosciences).

Dubrule, O., 1981b, Krigeage et splines en cartographie automatique–application à des exemples pétroliers. "Docteur-Ingénieur" thesis; Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau. 141 p.

Laurent, P. J., 1972, Approximation et optimisation: Enseignement des Sciences, Hermann, Paris, 531 p.

Matheron, G., 1973, The intrinsic random functions and their applications: Adv. Appl. prob.; v. 5, p. 439–468.

Matheron, G., 1980, Splines and kriging: Their formal equivalence, Internal Report, Centre de Géostatistique: Ecole des Mines de Paris, Fontainebleau. 20 p.

Utreras, F., 1979, Utilisation de la méthode de validation croisée pour le lissage par fonctions splines à une ou deux variables, "Docteur-Ingénieur" thesis: Université Scientifique et Médicale de Grenoble, 192 p.

Wahba, G., 1979, How to smooth curves and surfaces with splines and cross-validation, Proceedings, 24th Design of experiments conference, F. Dressel (Ed.): New York, Academic Press, 25 p.