A comparison of surface fitting algorithms for geophysical data

T. El Abbass¹, C. Jallouil², Y. Albouy¹ and M. Diament²
¹ORSTOM, Bondy, France; ²Laboratoire de Géophysique, Université Paris Sud, Orsay, France.

ABSTRACT

Many treatments in geophysics require regular grids of data. Since the data are generally recorded irregularly (e.g. gravity measurements along roads) or even on tracks (e.g. satellite measurements), it is necessary to grid the observed data.

We present the result of a comparison of various surface fitting algorithms carried out in order to check their reliability. Two different types of sampling have been verified: (i) clouds of points and (ii) points along tracks. Five algorithms have been extensively tested: (1) polynomial fit, (2) algorithm using a combination of spline-laplace, (3) kriging (4) least-squares fitting method, and (5) finite element method. The suitability of each method for different sets of data and the limitations (in terms of amplitude and gradient) are discussed.

Terra Nova, 2, 467–475

INTRODUCTION

Geophysical data are generally recorded with a rather random distribution that needs interpolation in order to compute a regular grid. Regular grids are required either for representation purposes or for further treatments. The development of small computers has allowed the interpolation procedure to be easily performed and therefore a number of workers in earth sciences, not only in geophysics, now have to create regular grids from irregularly spaced data, for which, various algorithms have been proposed. Bhattacharyya (1969) applied the bicubic spline interpolation to geophysical data. Sandwell (1987) applied biharmonic spline interpolation in SEASAT Altimeter Data. A method which is based on the theory of random variables is described by Matheron (1970) and Olea (1974). Hayes and Halliday (1974) and Dierckx (1980) apply a least-squares fitting method of cubic spline surface.

A new method, described by Inoue (1986), is based on the definition for optimum function minimizing the L² norm composed of the data residuals and the first and second derivative. The optimum function is approximated by a cubic B-spline expansion. Many methods for solving smoothing or interpolating problems have been proposed – see reviews by Crain (1970), Brodlie (1980), McLain (1980), and Lancaster and Salkauskas (1986). We present here the results of a comparison of various methods which are commonly used. Our aim is to give some hints as to the most suitable method according to the shape of the data (important extrema or important gradients) and to their geographical distribution (random or along tracks or roads). For that purpose, we choose to study several algorithms for interpolating: (1) polynomial fitting (La Porte, 1962); (2) a combination of spline and laplace methods (plot 88); (3) kriging (Matheron, 1970); (4) least-squares fitting of bicubic spline surfaces (Dierckx, 1980); and (5) a finite-element approach using a cubic B-spline basis (Inoue, 1986).

We used theoretical data for testing these different algorithms. However, geophysical data were treated by the same algorithms for constructing the map of a Bouguer anomaly. All algorithms were tested on a personal computer (IBM compatible).

PRESENTATION OF THE ALGORITHMS CHECKED

Polynomial method

A well-known algorithm for fitting data on a surface using polynomials was proposed by La Porte (1962) in order to produce isovalue maps from a set of data using computers. For each point where an interpolated value must be computed (i.e. the 'nodes' of the grids), a fit of a polynomial surface to the neighbouring data is performed using a least-squares approach. Therefore the coefficients of the polynomial surface are computed and the interpolated value is deduced. The data on which the fit is performed are located inside a circle with a variable diameter. Additionally, each datum is weighted so that more weight is given to the closest datum to the interpolated point.

The weight is taken as (La Porte, 1962):

\[ P_i = \frac{(R^2 - D_i^2)(D_i^2 + n^2)}{D_i} \]

where \( R \) is the radius of the interpolating circle, \( D_i \) is the distance between the \( i \) data and the point to estimate, i.e. the centre of the circle, and \( n \) is a small quantity such that \( P_i \) is defined if \( D_i \) equals zero. With larger values of \( n \), the polynomial surface will be smoother.

Once \( n \) and \( R \) are fixed, in order both to avoid an over smoothing and to take into account enough data to compute the polynomial coefficients, it appears that this interpolation method depends upon the degree of the polynomial and upon the repartition of the data. Note that if the degree of the polynomial is increased, the number of data points...
needed is also increased and therefore $R$ must be enlarged.

We used a program (Albouy and Godivier, 1981) strictly following the algorithm given by La Forte (1962).

### The spline–Laplacian method

This interpolation algorithm is a combination of a Laplacian and of a spline interpolation. The equation used is:

$$[8^2X(z)+8^2Y(z)]+C[8^4X(z)+8^4Y(z)] = 0$$

The first bracket corresponds to the Laplacian interpolation and the second one to the spline interpolation. Therefore according to the value of $C$ the computed surface is more or less controlled by one interpolation or another. For example, the computed surface is more controlled by spline interpolation for high values of $C$.

It must be noted that the estimated surface has to fit the observed points. An interpolation function is computed between each two data points from their values, the slope and the curvature.

We used the Plot 88 library software (Young and Van, 1987).

### The kriging method

The kriging method, developed by G. Matheron and his school, is based on geostatistics. To estimate the value $Z$ of one variable at a point $M$, this method uses an estimator:

$$Z^* = \sum_{i=1}^{n} P_i Z_i$$

where $n$ is the number of data, $Z_i$ is the measured values, and $P_i$ is the weight.

The weights $P_i$ are not a function of distance from the data points to the estimated point only, but depend on the spatial continuity of the variable. The spatial continuity is defined by a function called a variogram which is used in the determination of the weighting factors $P_i$.

The variogram function determines the 'variance of estimate'. For a given length $(l)$ and direction $(M_iM_j)$, the variogram is defined by the following expression:

$$\gamma(l) = \frac{1}{2} \text{[variance (Z2−Z1)]}$$

and experimentally,

$$\gamma(l) = \frac{1}{2} N_{i=1}^{N} (Z2−Z1)^2$$

where $N_i$ is the number of vectors $h = \overline{M_iM_j}$.

We compute the experimental variogram and we determine a theoretical model of $\gamma(h)$ which is a linear function of the distance.

Krigage minimizes the variance of estimate $\sigma^2$ by differentiation with respect to weights $P_i$. The variance of estimate is given by the following expression:

$$\sigma^2 = 2 \sum P_i \gamma(M_iM_i) - \sum P_i \gamma(M_iM_j)$$

The minimum variance of krigage gives an idea of the error made.

Finally, we note that this method is detailed by Matheron (1970); Guillaume (1970) and Haas and Viallix (1976) applied it in order to solve the problem of estimates and contouring. Further details of variograms are given in Delhomme (1976).

### A least-squares fitting method using a cubic b-spline

Hayes and Halliday (1974) describe a computational method for fitting bicubic spline by least-squares for arbitrary data and Dierckx (1977) for data on a rectangular grid. However, automatic determination of knots is not considered, and because the goodness of fit with a least-squares spline function is considerably affected by the number of knots and their position, Dierckx (1980) developed this method where the number of knots and their positions are determined automatically.

The following problem is considered. Given the function values $f_i$ at some point $s(x_i, y_i)$, $r = 1, \ldots, m$ scattered arbitrarily in the domain $D$ with positive weights $w_{r}$, $r = 1, \ldots, m$, we determine a spline function $s(x, y)$ while trying to find a compromise between the following objectives:

(i) The prescribed value $f_i$ should be closely enough.

(ii) The approximating spline should be smooth enough in sense that the discontinuities in its derivatives are as small as possible.

The spline is given in the b-spline representation:

$$s(x, y) = \sum_{i=1}^{N} \sum_{j=1}^{M} C_i N_i(x) G_j(y)$$

where $M_i(x)$, $N_j(y)$ are normalized cubic b-spline (De Boor, 1972) and the coefficient $C_{ij}$ is determined as the solution of the following constrained minimization problem:

$$\min \sum_{i=1}^{N} w_i [f_i - s(x_i, y_i)]^2$$

subject to the constraint

$$\sum_{i=1}^{N} w_i = 1$$

where $\eta$ is a measure of the lack of smoothness of $s(x, y)$, $\eta_i$ denotes the weighted residual $w_i [f_i - s(x_i, y_i)]$, and $S$ is a non-negative number which controls the extent of smoothing and is therefore called the 'smoothing factor'. (This parameter is discussed by Reinsch (1967) and Dierckx (1980, 1981).)

We use a subroutine from NAG library (Numerical Algorithm Group, 1988).

### The finite-element approach method

An algorithm using the finite element approach method was recently proposed by Inoue (1986). This method is used to compute a surface using cubic b-spline expansion and to minimize the misfit, the fluctuations and the roughness of the function. As mentioned by Inoue (1986), the solution so computed represents the deformation of a thin elastic plate under tension. The plate being pulled laterally toward datum points by springs whose stiffnesses are inversely proportional to the variances of the data.

Indeed, Inoue proposes to minimize the following function:

$$\Pi = \sum_{i=1}^{N} w_i [f_i - s(x_i, y_i)]^2 + 1/l_i^2 \int \{ |W_i(\Phi_r + \Phi_s) + W_2(\Phi_r + 2\Phi_s + \Phi_s^2)]dx dy$$

with $\Phi^2 = [\Phi(x, y)]^2$

$$\Phi^2_{xx} = [\Phi(x, y)]^2$$_{x}$$

$$W_1 = \tau [l_i/d_i]^2$$

$$W_2 = (1-\tau) [l_i/d_i]^2$$

$$\Phi(x, y) = \sum_{i=1}^{N} \sum_{j=1}^{M} C_{ij} F_i(x) G_j(y)$$

where $F_i(x)$ and $G_j(y)$ are cubic b-spline bases, $w_i$ are the weights of the data $d_i$, $l_i$ is the unit length, $d_i$ is the unit observable, $\rho$ is a roughness parameter. It controls the tradeoff between the fitness
to data and the total smoothness, $\tau$ is a tension parameter. It controls the nature of the function by varying from a plate free from tension ($\tau = 0$) to a membrane under tension ($\tau = 1$).

These two parameters are discussed by Inoue (1986). In Fig. 1 we show the roll of each one when applied to geophysical data (we remark here that with a low value of roughness the calculated curve doesn’t fit the observed data). The interpolation is controlled also by the number of division parameters ($M_x$ and $M_y$). These parameters depend on the data distribution.

So three parameters control the interpolation: the roughness, the tension and the number of divisions of the domain.

We used the program given by Inoue (1986) in his paper.

**PRINCIPLE OF OUR STUDY**

Our approach to test various algorithms of surface fitting is shown in Fig. 2. We first choose a synthetic surface Fig. 3 which presents both high extrema and important slopes, these are difficult features to interpolate but ones that are often observed in the field. Indeed this surface is defined as the composition of three functions as:

$$f(x, y) = A \exp(ax^2 + by^2 + cx + dy + e)$$

The resulting function is computed over a square area.

Then a distribution of points is taken out of that synthetic surface and so defines our set of data. We interpolate this set of data in order to obtain a regular grid and we finally compare the computed grid with the one having the same spacing deduced directly from the synthetic surface. We can quantify the quality of the interpolation by computing the residual grid.

Several distributions have been studied: random distribution with various densities for the data and distribution along profiles varying the distances between the profiles. These two kinds of distribution were chosen in order to represent two different types of data, either data on land such as those recorded during gravity survey, or data obtained at sea or by satellites.

**RESULTS AND DISCUSSION FROM A THEORETICAL EXAMPLE**

**Polynomial fit**

As previously mentioned, the interpolation using the polynomial fit depends upon the degree of the polynomial, the interpolation radius and the distribution of experimental data.

Figure 4 shows residuals obtained for the third and fourth degrees. We show the results obtained with a random distribution of 400 data points and with a radius of a half of the diagonal. As it could be expected, the residual is smaller when the degree is higher since, for example, degree 3 gives a residual of about 5% (Fig. 4a), while degree 4 gives 0.5% (Fig. 4b). Similarly, it is obvious that the fit is improved with a smaller radius, displaying a linear fit with a radius of a fourth and an eighth of the diagonal (Fig. 5a and Fig. 5b, respectively).

We note that this method doesn’t represent the extreme very well.

Concerning the profile distribution, it
appears that the residual is controlled by the spacing between the profiles and by their orientation in respect to the features of our synthetic data as expected (Fig. 6a). But if the profiles have larger distances between them, the interpolation creates an artificial feature parallel to the profiles when the distance between them is greater than the wavelength inspected (Fig. 6b) thus there is an aliasing effect in the space domain.

It must be remarked that this much used method gives reasonably good results with a high degree of accuracy and a small interpolation radius. Nevertheless it is rather slow and respects the slopes but does not represent well the extremes.

**Spline–Laplacian method**

We first studied the influence of the parameter $C$ which controls the ratio of spline and Laplacian interpolation. The main variations are in the area with large slopes but it appears that this parameter does not act in an important way. Therefore, we kept a value of $C = 3$, i.e. a mean value, such that both spline and Laplacian interpolations act for the following tests.

Using a random distribution with 400 points, the residual is of about 5% of the signal (Fig. 7), the same range as that obtained with an identical distribution of points with a linear polynomial interpolation. It must also be noted that this method always provides poor fits in the areas of pronounced slopes.
With data along tracks, the results are always less accurate than with a random distribution of data.

In conclusion, although this method is quite fast, it seems to provide poor results and to be dependent on the data distribution. Nevertheless, it can be used for fast representations of data sets.

**Kriging method**

The first step consists in computing a variogram. Figure 8 shows the computed variogram from a random distribution of 400 data points. Interpretation of the experimental variogram is used to fix a linear model of the variogram function of distances. The first branch of the experimental variogram appears to be very linear which is logical according to the fact that data are synthetic.

Results of the interpolation are shown in Fig. 9. The high residual values are shown on the maxima and minima, so this method, like the polynomial method, doesn't represent well the extrema. Nevertheless, it requires a lot of computer time.

**A least-squares of surfaces fitting**

In this method we control the smoothness of the fit with the smoothing factor $S$.

We studied the influence of $S$ in a random distribution for a set of 1600 points with $S = 50$; Fig. 10a shows a residual about 0.5% over the slope. In Fig. 10b a residual is obtained with $S = 4000$; the high value of residual is over the extrema and it is about 1% of the signal.

Figure 11a shows a residual obtained for 400 points with $S = 50$. We have the same results for 1600 points, but with $S = 4000$ the residual is about 2%.
over the minimum and the maximum (Fig. 11b).

Using a distribution along tracks, the residuals show that the results are not controlled by the orientation of the tracks. With $S = 4000$, residual values are high over the extrema (Fig. 12a), and with $S = 50$ the low residual values are above the slope (Fig. 12b).

This method does not represent the extrema well when high values of $S$ are used but gives good results—small residuals with small $S$ parameter. However, this advantage for theoretical data is not confirmed if the method is applied to geophysical data (see later).

**Finite-element approach method**

As previously mentioned, three parameters control the quality of the interpolation, and their relative influence can be depicted. Figure 13 shows that the value of the roughness must be as high as possible in order to get the best fit. Figure 14 shows that the best fit is obtained with low tension values.

Concerning the number of divisions, it appears that it is closely related to the data distribution: the number of divisions must be increased when the grid interval is decreased. Also, as could be expected, the increase in the number of divisions yields better results; too high roughness and a great number of divisions have no influence on the results but do increase computer time.

This method gives good results, i.e. small residuals, even if data are distributed along tracks (Fig. 15). Moreover, it is a fast algorithm.

**RESULTS AND DISCUSSION FROM A GEOPHYSICAL EXAMPLE**

We have applied all these methods in gravity measurements to construct contour maps of the Bouger anomaly.

Figure 16 shows the position of the gravity measurements over a rectangle of $330 \times 330$ km containing 394 data points.

The results show that the quality of maps is not good with the spline-laplace methods (17a and b) as artificial anomalies are created.

With the polynomial method the results are not satisfactory over the boundary domain where there are few data (Fig. 18) although the slope is respected.

The maps obtained with the kriging method (Fig. 19) and finite element
approach method (Fig. 20a) are good, showing minimum curvature with respect to extrema and the slope. In addition this method gives good results in areas with limited data.

Figures 20b, c and d show the role of the three essential parameters: roughness $\rho$, tension $\tau$, and the number of divisions $M$. They give many solutions according to the desired objective and the data application.

The results show that the roughness $\rho$ and the number of divisions $M$ should be determined so that the fitted function extracts the true distribution from the noisy data as efficiently as possible.

Finally, Fig. 21 shows contour maps according to different $S$ values. Choosing $S$ too high ($S = 400000$) results in oversmoothed data (Fig. 21a). Overfitting the data results from choosing $S$ too

**Fig. 14.** These residuals show the effect of the tension parameter. (a) (top) and (b) are the results of interpolation with tension parameters equal to 0 and 1, respectively. Roughness parameter is fixed at 1000.

**Fig. 15.** Two residuals obtained along tracks. (a) (top) is the result with roughness equal to 1000 and tension equal 0; (b) is the result with roughness equal to 1000 and tension equal 1.

**Fig. 16.** The position of the gravity measurements over the studied area.

**Fig. 17.** Contour maps (a) (top) with the spline method; (b) with Laplacian method. Results are not very satisfactory as they are too rough.

**Fig. 18.** The contour map with a polynomial third degree: Result unsatisfactory near the boundaries.
Fig. 19. The contour map with Kriging method: A linear variogram is applied.

small ($S = 100000$) (Fig. 21b). Figure 21c shows the contour map with $S = 400000$, which could be accepted.

The result demonstrates that the fit produced by this method is unreliable in a region with few data. The unfavourable situation is accentuated if $S$ becomes smaller.

CONCLUSION

Our tests confirm that the interpolation technique, although widely used, is a very sensitive method and that careful attention must be paid to it. From the results presented here it is obvious that no method is universal.

Indeed the choice must be made according to the distribution of points and to the quality of interpolation needed i.e. it is more important to fit extrema or slopes, and, finally, to the computer time needed.

The polynomial method appears to be an efficient method to fit slopes if a high degree is used but can be very slow.

The spline-Laplacian method gives smoothed maps quite fast, but is far from being the best.

The kriging method is obviously a powerful one but it is heavy and expensive in calculation time.

The least-squares surface fitting method is not easy to use, because in practice this method is used with considerable interaction between the user and the computer via a graphics terminal, the reason being that the data are sensitive to the parameter $S$ which has to be chosen carefully: too small an $S$ value will result in an overfitting; too large an $S$ value results in an underfitting of data. Therefore, the user should always control the fit graphically before accepting it as satisfactory. Finally it should be mentioned that the value $S$ is not very crucial and also depends on the number of data points and the application.

The finite-element method is reasonably fast, easy to use, even if the first choice of parameters is quite subjective, and gives a good fit, so it appears to be a good choice. In addition, including the generalized tension parameter this method has more applications in order to find a suitable solution for gridding data when these data vary rapidly with distance (Smith and Wessel, 1990), e.g., using a tension parameter is good for bathymetry data.

Both kriging and finite element methods provide an estimation of the accuracy of the interpolation, and take the standard errors into account.

As previously mentioned, it is obvious that whichever interpolation algorithm is used, data would be predicted only where no measurement has been made. Therefore, interpolated grids must always be checked with respect to the original data so that possibly artificial features introduced during the interpolation may be removed before interpretation.

ACKNOWLEDGEMENTS

We thank C.I.R.C.E. (Centre Inter-Régionale de Calcul Electronique) for
permitting us to use NAG library. Critical comments by anonymous reviewers were very helpful.

REFERENCES


Received 6 June 1990; accepted 5 September 1990