Available online at www.sciencedirect.com
SCIENCE DIRECT®
Tectonophysics 376 (2003) 241–259
TECTONOPHYSICS
www.elsevier.com/locate/tecto

Geostatistics applied to best-fit interpolation of orientation data

C. Gumiaux*, D. Gapais, J.P. Brun

Géosciences Rennes, UMR 6118 CNRS, Université de Rennes 1, 35042 Rennes Cedex, France

Received 22 November 2002; accepted 25 August 2003

Abstract

Automatic methods used in geosciences to interpolate between orientation data have often limited applicability and strength, in particular where large ranges of orientations occur. In this paper, we show that geostatistical methods yield rather strong and powerful results when applied to directional data. The procedure involves the calculation of variograms, followed by a kriging interpolation of the data. In order to free from the circular property of directional data, the treatment of initial angular data sets is performed using scalar values provided by the direction cosines of double-angle values. The strength and application of the method are demonstrated by the analysis of theoretical and natural data sets. Natural examples are focused on the calculation and the analysis of cleavage trajectory maps.

Keywords: Geostatistics; Interpolation; Kriging; Directional data; Cleavage trajectories

1. Introduction

The analysis of the spatial distribution of directional data can have various applications in earth sciences, including structural geology (e.g. construction of cleavage trajectory maps) (Ramsay and Graham, 1970; Cobbold, 1979; Barbotin, 1987; Cobbold and Barbotin, 1988), sedimentology (e.g. analysis of paleocurrents) (Potter and Pettijohn, 1977), geomorphology (e.g. analysis of maximum slope orientations), paleomagnetism (e.g. analysis of principal directions of anisotropy of magnetic susceptibility), geophysics (e.g. construction of stress trajectory maps) (see Lee and Angelier, 1994), or geodesy (e.g. analysis of movement directions from GPS data).

Two main applications of the statistical analysis of orientation data are (1) the calculation of mean principal directions from a data population, and (2) the analysis of variations of data in space (e.g. on maps or cross-sections). Many automatic methods have been developed to extract best principal directions from a population of orientation data (calculation of eigenvectors and eigenvalues, see Angelier, 1994 and references therein), or to detect structural preferred orientations (e.g. Panozzo, 1984). Developments also exist for the automatic analysis of spatial variations of orientation data in geosciences (Young, 1987; Cobbold and Barbotin, 1988; Lee and Angelier, 1994; Xu, 1996; Lajaunie et al., 1997). The treatment of directional data to analyse their variations in space requires interpolation between local measurements, and this raises two main problems. First, directional data are circular, i.e. they can vary between 0° and 360°. Consequently, classical statistics, as used for algebraic

* Corresponding author. Fax: +33-2-23-23-60-97.
E-mail address: Charles.Gumiaux@univ-rennes1.fr (C. Gumiaux).

0040-1951/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.tecto.2003.08.008

data sets, cannot be straightforwardly applied to directional data (see Upton and Fingleton, 1989). Second, one can be interested in studying variations correlated at a regional scale and this can induce a difficult "scale problem" between the weight given to individual data and the degree of data filtering required. For example, most interpolation techniques tend to respect as best as possible the original data, and this may fail if one wants to extract regional-scale trends.

Data interpolation using geostatistics was first introduced by Matheron (1962) to estimate ore volume in mining engineering. Recently, this technique became more common in geosciences, with various applications, such as the interpolation between scalar values, like strain ellipticity (Mukul, 1998) or soil properties (Bourennane, 1997). Further particular applications have been developed like the treatment of vectors (Young, 1987; Lajaunie et al., 1997), or the analysis of curvilinear geometries (Xu, 1996). In particular, Lajaunie et al. (1997) have described the use of kriging interpolation to define the 3D geometry of geological surfaces (e.g. lithological boundaries or foliation envelopes). In this paper, we develop a geostatistical approach, with particular reference to the analysis of structural data. We emphasize that the method can solve problems classically encountered in the spatial analysis of orientation data, provided a careful preliminary analysis of the data set is made before proceeding to an interpolation. The general technique is first presented and then tested using both theoretical and natural examples.

2. Methodology

A set of orientation data distributed throughout a given domain can be regarded as a representative sample of a continuous regionalized variable defined for the whole area. This regionalized variable is a deterministic function relative to the chosen reference frame. However, data may also contain nonnegligible random and insignificant contributions. The kriging method developed by Matheron (1962) allows to interpolate between data points, taking into account the different contributions involved in data values. The method has been described as an "optimal prediction" for the interpolation of a variable between its location and neighbouring scattered data points (Cressie, 1990). Detailed descriptions of kriging techniques can be found in Isaaks and Srivastava (1989), Deutsch (1991), Deutsch and Journel (1992), or Wackernagel (1998). The general principle is to consider the regionalized variable as one possible outcome of a random function attached to the location of the considered data point. One random function is thus defined for each data point. The problem is then reduced to the analysis of the variations of the parameters defining these random functions, in order to estimate the spatially dependent part of the data, which is equivalent to the contribution of the regionalized variable.

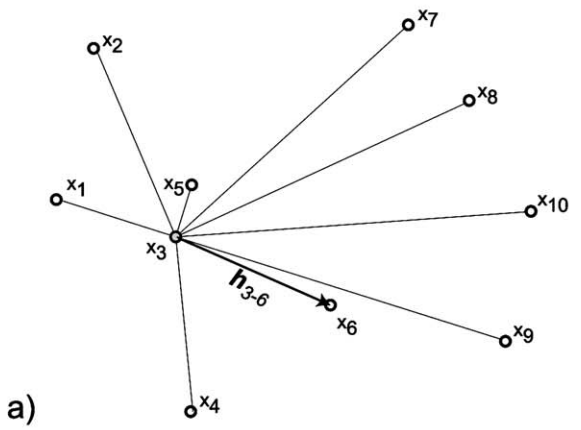
The geostatistical approach is here applied to the analysis and modelling of spatially scattered local directional data. After building a georeferenced database, the procedure used comprises two main steps, as follows (see Appendix A).

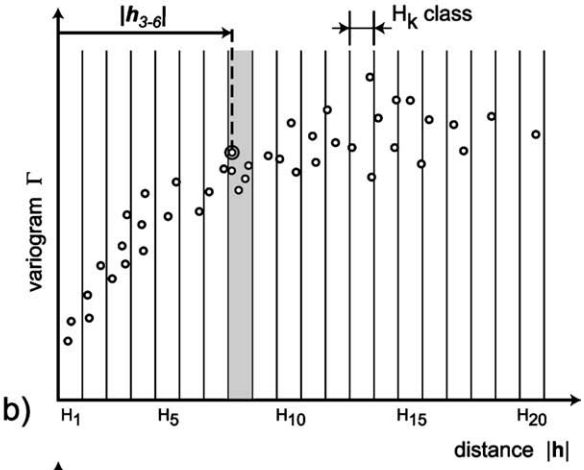
2.1. Quantification of spatial variations of the data set

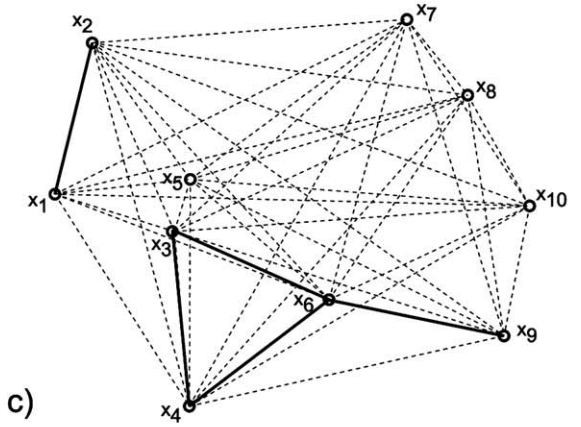
For spatial interpolation purposes, it is often admitted that two close data points may have a more similar value than two distant ones. Common techniques, such as inverse distance weighting, are based on this hypothesis. The main problems with this approach are (1) to determine the maximum distance for which data values show significant mutual dependence, (2) to estimate how and how fast this dependence decreases with increasing distance, and (3) to take into account the effects of possible anisotropies in the spatial variations of the data. The quantification of such parameters can be provided by the computation of experimental variograms.

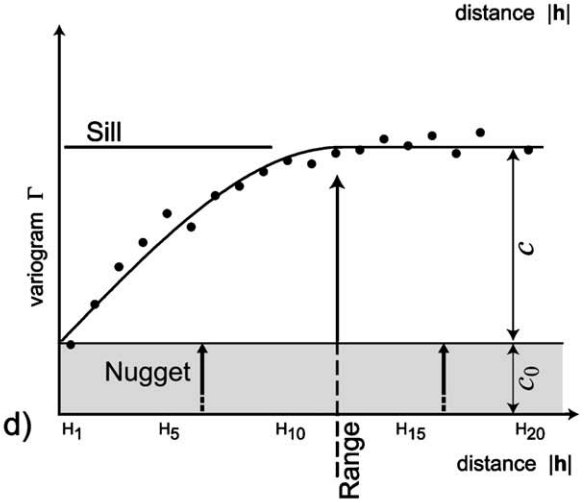
Fig. 1. General principles of variogram calculation. (a) Example of spatial data set (in 2D, unscaled). (b) Example of point cloud on an experimental variogram where the sample variance is plotted as a function of the separating distance $|\mathbf{h}|$ between data points. (c) Selection of the vectors corresponding to a given class of distance; doted lines between the points represent the vectors for all the data set; thick segments join the data pairs fitting the range of distance chosen here. (d) Sketch of the general variogram corresponding to the cloud variogram (b); the continuous line corresponds to the best-fit theoretical variogram (spherical function). (e) Selection of the vectors corresponding to a given class of direction; doted lines between the points represent the vectors for all the data set; thick segments join the data pairs fitting the range of direction chosen.

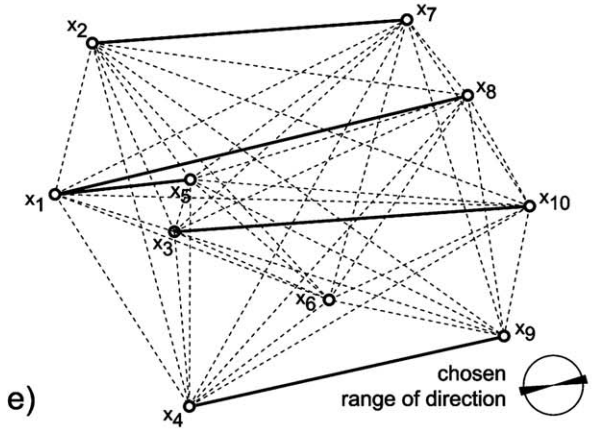
To compute an experimental angular variogram, one first calculates the differences $\Delta\alpha_{i,j}$ between directions attached to pairs of data locations (Appendix A.1). Considering a scattered set of points, the discrepancy between values at two distant locations, for example x_3 and x_6 in Fig. 1a, is expressed by the half-











Spherical theoretical variogram:

$$\Gamma(h) = c_0 + c\,\mathrm{Sph}(h)$$

$$\mathrm{Sph}(h) = \begin{cases} 1.5\,\frac{h}{a} - 0.5\left(\frac{h}{a}\right)^3 & \text{for} \mid h \mid \leq a \\ 1 & \text{for} \mid h \mid > a \end{cases}$$

h is the separating distance, a the range.

lines joining data pairs (c) and (e).

lines joining data pairs fitting the chosen range of distance (c) or direction (e)

squared difference (expressed in deg^2) Γ between the two angular values α_3 and α_6 attached to these points. This squared difference, the *sample variance*, is given by

$$\Gamma = \left(\frac{\Delta\alpha_{x_3,x_6}}{2} \right)^2 \quad (1)$$

This value is calculated for each pair of points of the data set (Fig. 1b). Considering that the sample variance is independent of the location of the considered pair of points (an assumption defined as the intrinsic hypothesis, see Wackernagel, 1998), the experimental variogram is a function of the separating vector **h** (Fig. 1a), and is expressed as:

$$\Gamma(\mathbf{h}) = \frac{1}{2}[\Delta\alpha_{i,(i+h)}]^2 \quad (2)$$

This function is symmetrical with $\Gamma(-\mathbf{h}) = \Gamma(\mathbf{h})$.

Assuming that data variations are isotropic, a general variogram can be simply imaged as the plot of the half-squared differences against the separating distance $|\mathbf{h}|$ between all the pair of points (Fig. 1b). In order to improve the analysis of the variogram, it is sliced in classes of distance H_k and the values calculated for data pairs fitting a given class are averaged (see class H_8 in Fig. 1b and the corresponding fitting vectors in Fig. 1c). The equation of the general variogram (Fig. 1d) becomes:

$$\Gamma(H_k) = \frac{1}{2n_k} \sum_{1}^{n_k} [\Delta\alpha_{i,(i+h)}]^2 \quad (3)$$

where i and j correspond to data points separated by a distance $|\mathbf{h}| \in H_k$, and n_k is the number of data pairs fitting the class H_k.

The experimental variogram defines a curve representing the increasing discrepancy (i.e. the decreasing of mutual dependence) between data, as a function of their degree of proximity $|\mathbf{h}|$ (Fig. 1d). The minimum width of classes of distance H_k and the maximum distance of computation are imposed by the minimum and maximum distances between data points, respectively. When approaching highest distances, the number of data pairs fitting the classes can be low (Fig. 1b) and the corresponding part of the variogram is thus poorly constrained.

Four important statistical parameters come up from an experimental variogram (Fig. 1d) (Matheron, 1962).

- The variogram can reach a constant value, classically defined as the *sill*, with increasing distance of calculation. This means that the variations of the variable are correlated up to distances smaller than the total size of the studied domain.
- The maximum correlation distance for which the variogram reaches the sill is defined as the *range*. For distances larger than the range, data are no longer correlated.
- The variogram can show a nonzero ordinate at the origin. From a theoretical point of view, this value comprises the random part of the random functions defined for the studied domain. It thus corresponds to the nonregionalized part of data values (in grey in Fig. 1d), which can include background noise, as well as local effects that can be significant at a scale much smaller than that of the study area. For mining purpose, Matheron (1962) defined this ordinate at the origin as the *nugget*. The regionalized part of the data is the difference between the sill and the nugget.
- The experimental trend of the variogram indicates the way the variance increases with increasing distance within the data set. It can be fitted to a theoretical function, classically chosen as spherical (Fig. 1d), exponential, Gaussian or power function, which depends on the values of the range, sill, and nugget (see Wackernagel, 1998).

For a given data set, the range of data variations can change according to the considered space direction. Such anisotropy in data variance can be detected by calculating variograms along particular space directions. The procedure is the same as for omnidirectional variograms, except that the overall computation space is sliced into orientation classes for which individual variograms are calculated. For each class, the calculation uses the data pairs separated by a vector **h** whose orientation is comprised in the orientation range of the class (Fig. 1e). During calculation, the number of orientation classes must be large enough to reveal the anisotropy, but not too large, so that the number of data pairs used for each variogram remains sufficiently large.

2.2. Interpolation between data

As for general statistical studies, geostatistics cannot be directly performed on angular values. The kriging interpolation is therefore made using direction cosines (Appendix A.2). Each individual directional data is thus expressed by its two direction cosines. The variogram analysis can then be performed on the sums of the two direction cosines (Young, 1987) or on each set of direction cosines (this study) (For the examples presented in this paper, both techniques yield similar results). Detailed descriptions of the calculation procedure can be found in Cressie (1991) or Wackernagel (1998), and a brief summary is given below.

As for other weighted moving average techniques, the interpolation is made along a regular grid covering the whole studied area. At each node location, ordinary kriging is performed using all adjacent data points p located inside a given search circle (Fig. 2). The radius of the search circle corresponds to the range estimated from the variogram analysis. From the theoretical function deduced from the variogram, covariance values are calculated between all involved data points. A set of p equations relating covariance values to distances between data points, including the considered grid node, defines the ordinary kriging system. At each point p, the direction cosines are weighted according to the solution of the corresponding equation. A weighted mean of direction cosines is then computed, and attributed to the grid node. Finally, these latter values are reconverted into angular values (Appendix A.2; Fig. 12).

From the interpolated grid, various types of representation of the distribution and variations of directional data can be made. These are in particular trajectory maps or contour maps. To draw trajectories, we used the algorithm proposed by Lee and Angelier (1994, p. 165).

3. Examples of variogram modelling and kriging of directional data

3.1. Recognition of maximum size of correlated directional changes (Fig. 3)

In this example, constant N120° striking directions occur across a given domain, and a disc-shaped

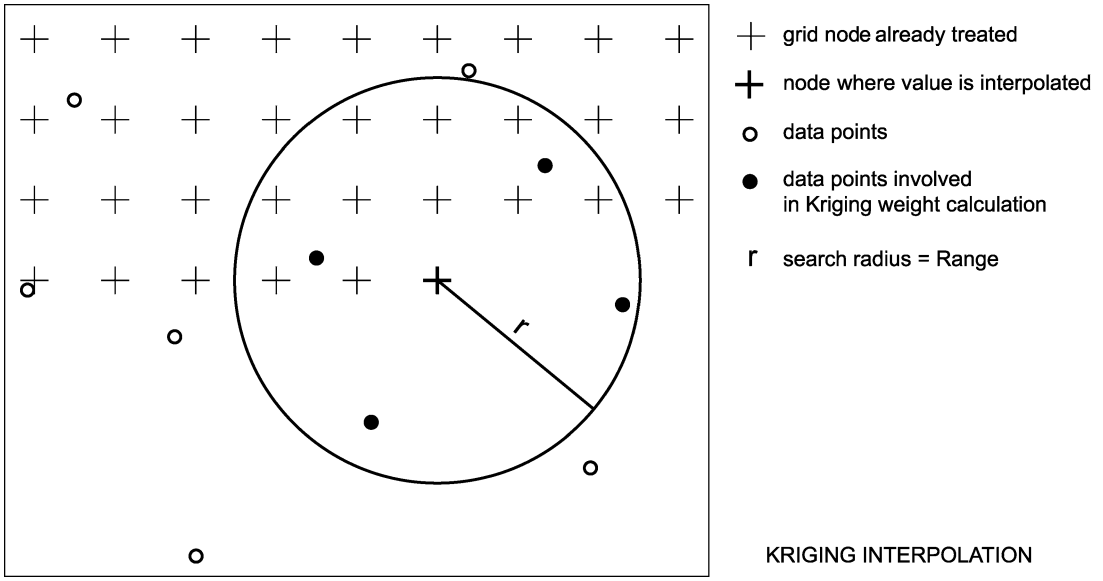


Fig. 2. Illustration of data interpolation along a regular grid. For a given node of the grid (thick cross), a circular neighbourhood is given around this point. The radius of the circle is defined by the range deduced from the variogram analysis. All data values located in the circle (black points) are implied in the calculation of the interpolated value attached to the given node (modified after Wackernagel, 1998).

perturbation of 6 units wide has been introduced within this continuous field (Fig. 3a). The maximum deviation is of 45° in the middle of the disc and linearly decreases toward zero at disc boundary. The variogram corresponding to this model has been calculated from points scattered across the whole area (Fig. 3b). The variogram reaches a clear sill for increasing separation distances, and a spherical theoretical function fits the experimental data. The range value is well defined as 3 units, which exactly equals the radius of the perturbation. For larger distances, data are not correlated.

The maximum distance of spatial correlation between directional data has already been considered as an important parameter for interpolation (Lee and Angelier, 1994). The example shown here emphasizes that this parameter can be directly constrained from the variogram before performing interpolation.

3.2. Recognition of anisotropies (Fig. 4)

Directional data can present strong anisotropies in their variation trends. To illustrate relationships between anisotropy and variograms, a regular fold-shaped pattern has been drawn, with axial planes striking N30° (Fig. 4a). The direction of the folded trajectories has been digitized for few localities (Fig. 4a) and variograms have been calculated from this data set. The omnidirectional variogram shows a bell-shaped envelope, with a maximum half-variance of about 2000 deg^2 for separation distances of about the half of the total model size (Fig. 4b). This suggests some periodical distribution. Consistently, a periodical trend appears clearly in oriented variograms approaching a WNW–ESE direction, and a maximum ratio between the amplitude and the wavelength of the signal is observed on the N120° oriented variogram (Fig. 4c). This direction actually corresponds to the one along which changes in data attitude are maximum (Fig. 4a). In contrast, the perpendicularly N30° oriented variogram is flat, with rather low values (Fig. 4d). This feature means that directions remain constant along any given N30° striking line (Fig. 4a).

This example emphasizes that oriented variogram are powerful tools to detect anisotropies in orientation distributions. Furthermore, the wavelength of the periodicity of the oriented variograms provides an estimate of the wavelength of the anisotropy. Indeed, for ideal situations, these two values should be equal (6 units, Fig. 4a,c).

3.3. Recognition of orientation domains (Fig. 5)

In this example, regularly spaced directional data have been distributed across a rectangular area (Fig. 5a). The variogram analysis and the interpolation method along nodes of a regularly spaced grid have been performed on this data set, and trajectories were constructed across the whole area. A very good consistency is observed between trajectories and initial data (Fig. 5b). The maximum range of initial directions is rather moderate (around 45°), and the segments representing directional data seem to show

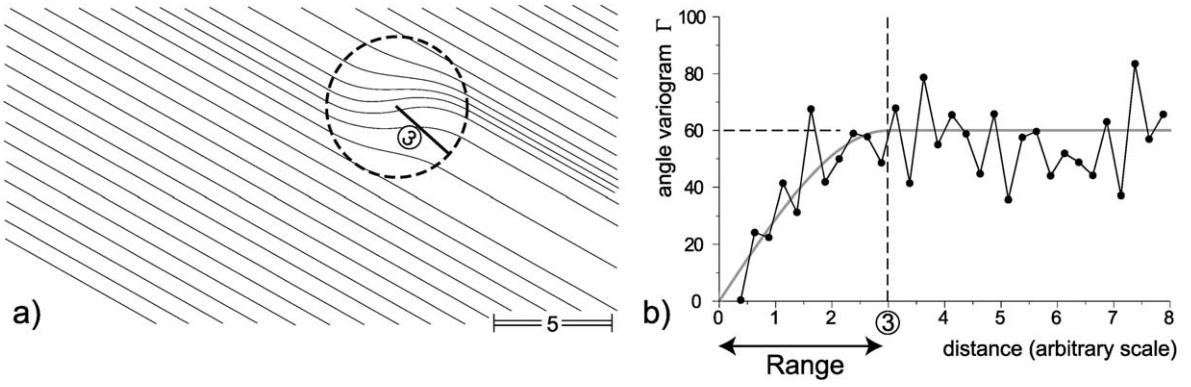


Fig. 3. Illustration of relationships between largest-scale directional variations present in the direction field (a) and the range given by the variogram (b).

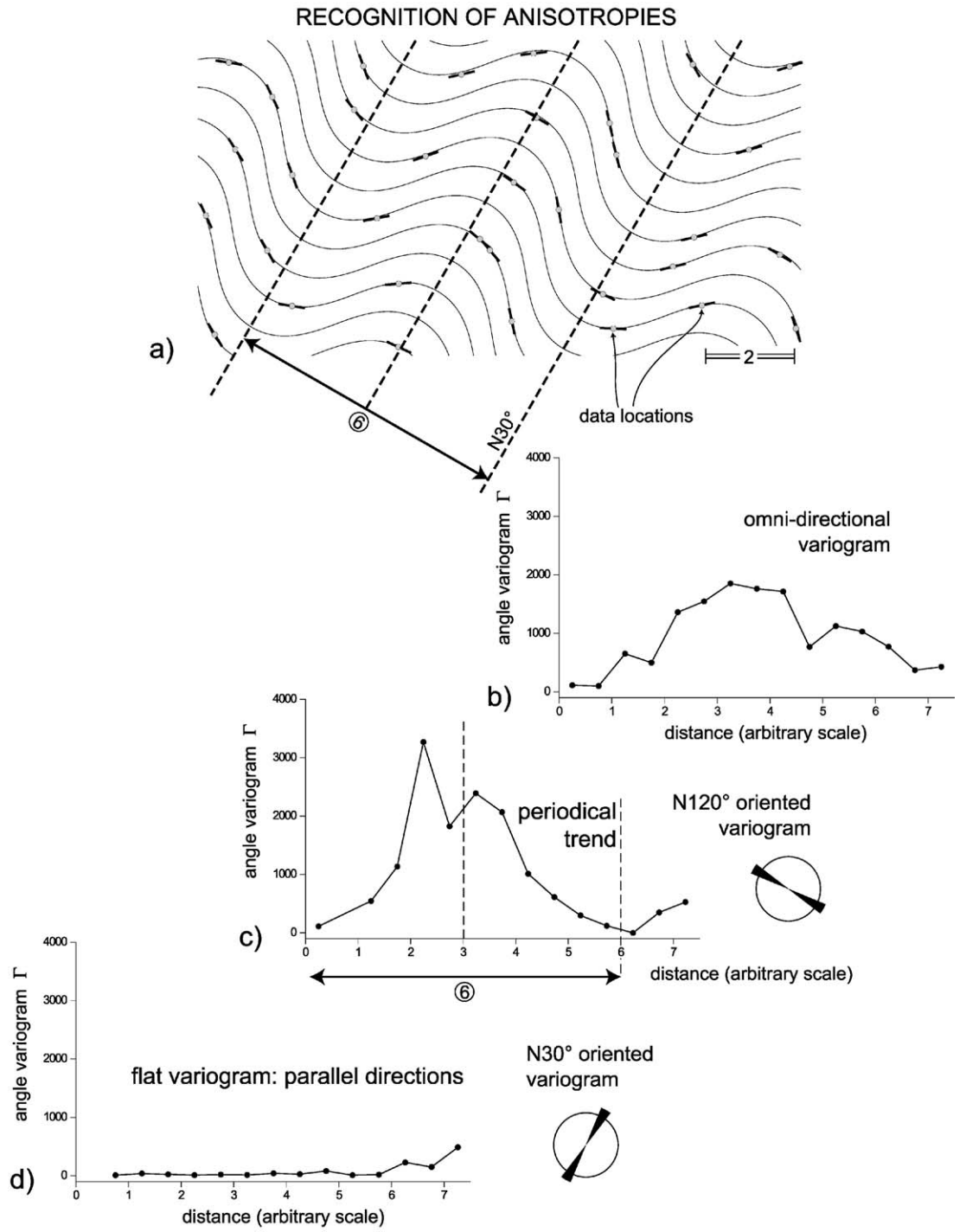


Fig. 4. Illustration of relationships between an anisotropy in direction variations (a) and variations in periodical trends of omnidirectional (b) and oriented variograms (c, d).

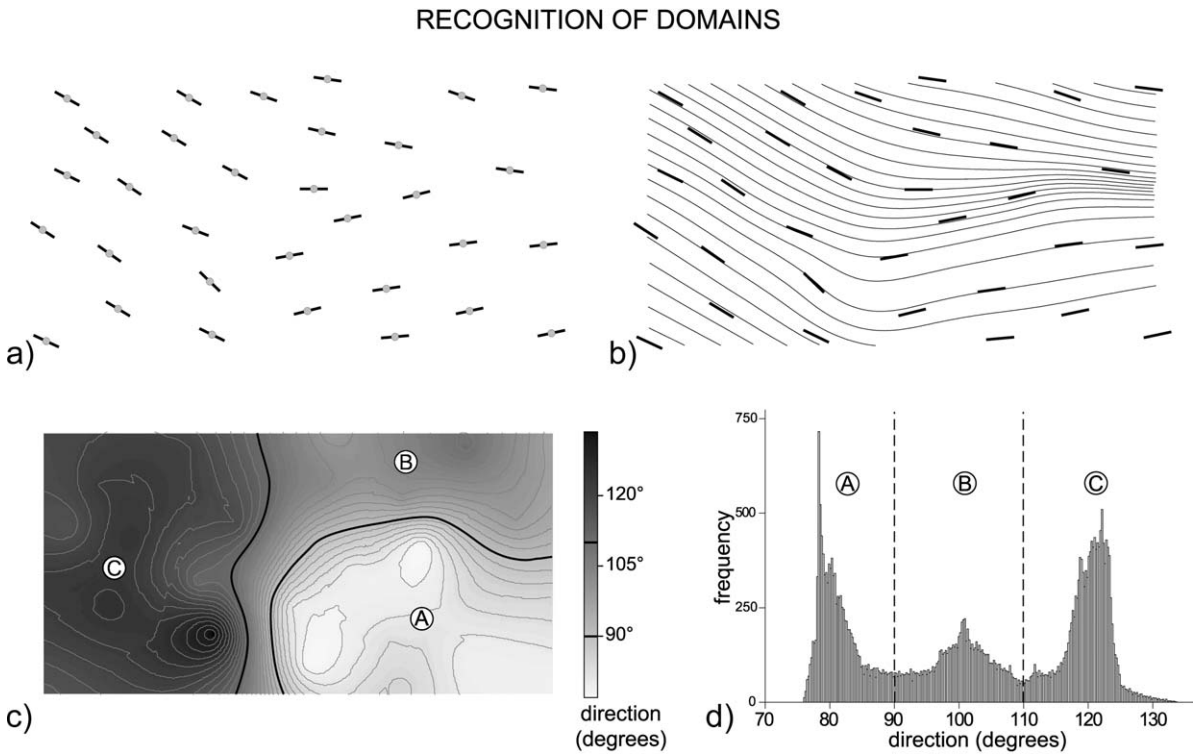


Fig. 5. Example of detection of domainal distributions. Example data set (a) and interpolated trajectories (b) showing apparently rather moderate and smooth directional variations. Direction contour map (c) and frequency diagram (d) outlining the occurrence of three distinct directional domains.

smooth variations from west to east (Fig. 7b). This is also suggested by the trajectory map. From these, we could conclude to regular variations of directions across the area. In fact, these variations are not as continuous as suggested by trajectories. Thus, a 2°-spaced contour map drawn from the interpolated grid shows local tightening of orientation contours (Fig. 5c), which could reflect different lateral variations in the gradients of orientation changes. This is confirmed by a frequency histogram of interpolated direction data that shows a multimodal distribution pattern with three well-defined maxima (Fig. 5d). On the map, these maxima correspond to the three domains separated by high lateral gradients underlined by contour attitudes (Fig. 5c). In the western domain, N120° directions are dominant. To the east, domains A and B are marked by dominant directions around N80° and N100°, respectively. The differences in the maximum frequencies of the three peaks of the histogram result from the different size of the three domains.

In this model example, the computed trajectories are in good agreement with the initial directional data and outline well the overall trend of direction variation at regional scale. It also points out that other representations of the data and additional statistical calculations (e.g. lateral gradients, frequency analysis, profile extractions) can be useful to better describe regional trends. In our example, a domainal distribution of data is outlined, a feature that was not clearly detectable by the simple trajectory map.

4. Strength and limits of the method

To test the method further, we consider an example of regularly spaced directional variables (segments in

Fig. 6a). The total angular range is 0–180°, thus covering the entire possible values for nonoriented data. The variations are continuous across the whole area and the directions show a particularly highly variable zone around the middle of the rectangle.

4.1. Comparison of kriging interpolation with other methods

Fig. 6 shows a comparison between automatic construction of trajectories using different interpolation methods, including the kriging interpolation (Lajaunie et al., 1997; this study). Fig. 6a shows the result of an interpolation technique based on the resolution of a system of quadratic functions for the entire data set (Barbotin, 1987). Fig. 6b shows the result for the fit of a single continuous polynomial function to the entire data set (Lee and Angelier, 1994). The third treatment is based on a mobile average technique where the interpolation does not consider the data set as a whole (Lee and Angelier, 1994). Instead, the method is a step-by-step one: only data lying within a disk centred on a given location are treated together. For the average calculation, data pairs are weighted as an inverse function of the separation distance. This procedure is repeated for all nodes of the grid. Results are shown in Fig. 6c. Fig. 6d shows trajectories obtained with the geostatistical method.

The two first techniques are characterized by (1) the treatment of the data set as a whole, and (2) the direct use of orientations as algebraic values. Results emphasize that these constraints can lead to irrelevant interpolations, especially if the orientation range exceeds 90°. Indeed, poor correlations are observed between data and trajectories. Thus, bulk interpolation methods do not seem to be adapted to the treatment of directional data. Even for the case of simple direction fields, as the one presented here, the smooth regional trend of the data set is not detected. The two other methods (Fig. 6c,d) can be considered as based on mobile average techniques for the interpolation. In the step-by-step technique proposed by Lee and Angelier (1994), the circularity of directions is taken into account. Indeed, data used are angles, but the reference frame is adapted to a convenient orientation with

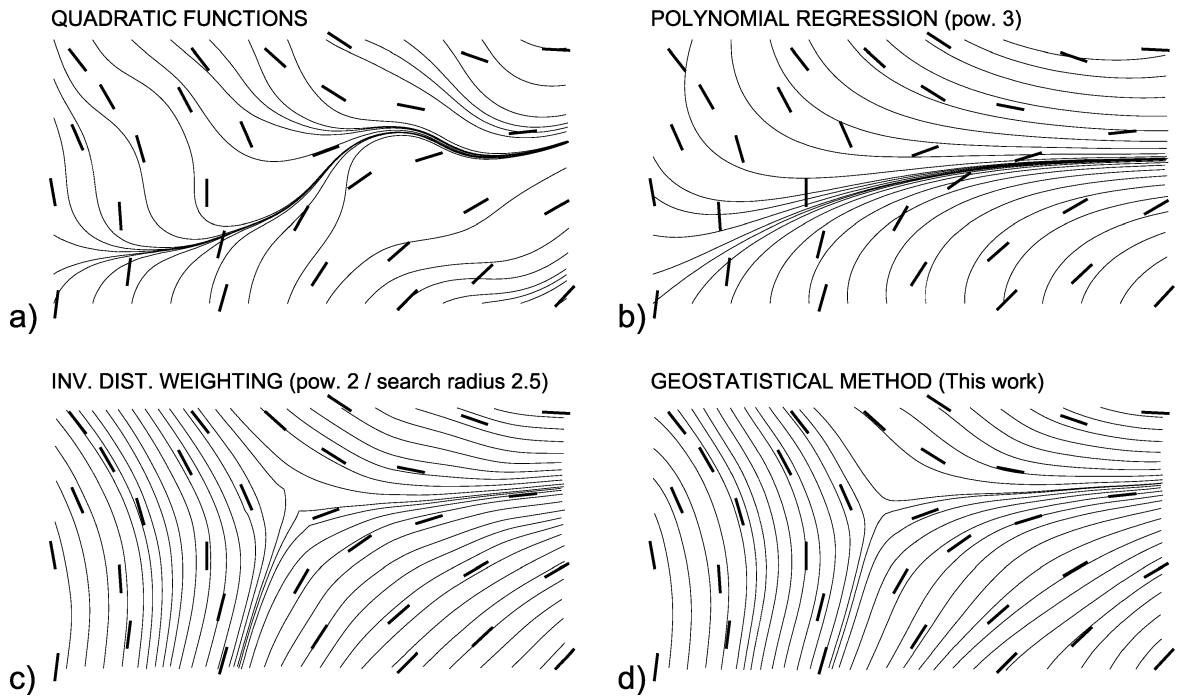


Fig. 6. Comparison between different methods of automatic interpolation and resulting direction trajectories. See text for further explanations.

respect to the considered data set averaged at each node of the grid (see "acute angle" technique of Lee and Angelier, 1994). In the present analysis, this particular problem is treated by means of the direction cosines. In the previous study of Lajaunie et al. (1997), the interpolation technique used partial derivatives of orientation data. For the example presented, marked by rather smooth local variations of direction, the two methods yield very satisfying results. The resulting trajectories are nearly the same and the regional trend of the directions is well respected. In particular, both methods detect a singular point (triple point, or neutral point) in the middle part of the map. Looking at initial data, hand-made trajectories would have drawn this singular feature.

4.2. Analysis of strongly disturbed data

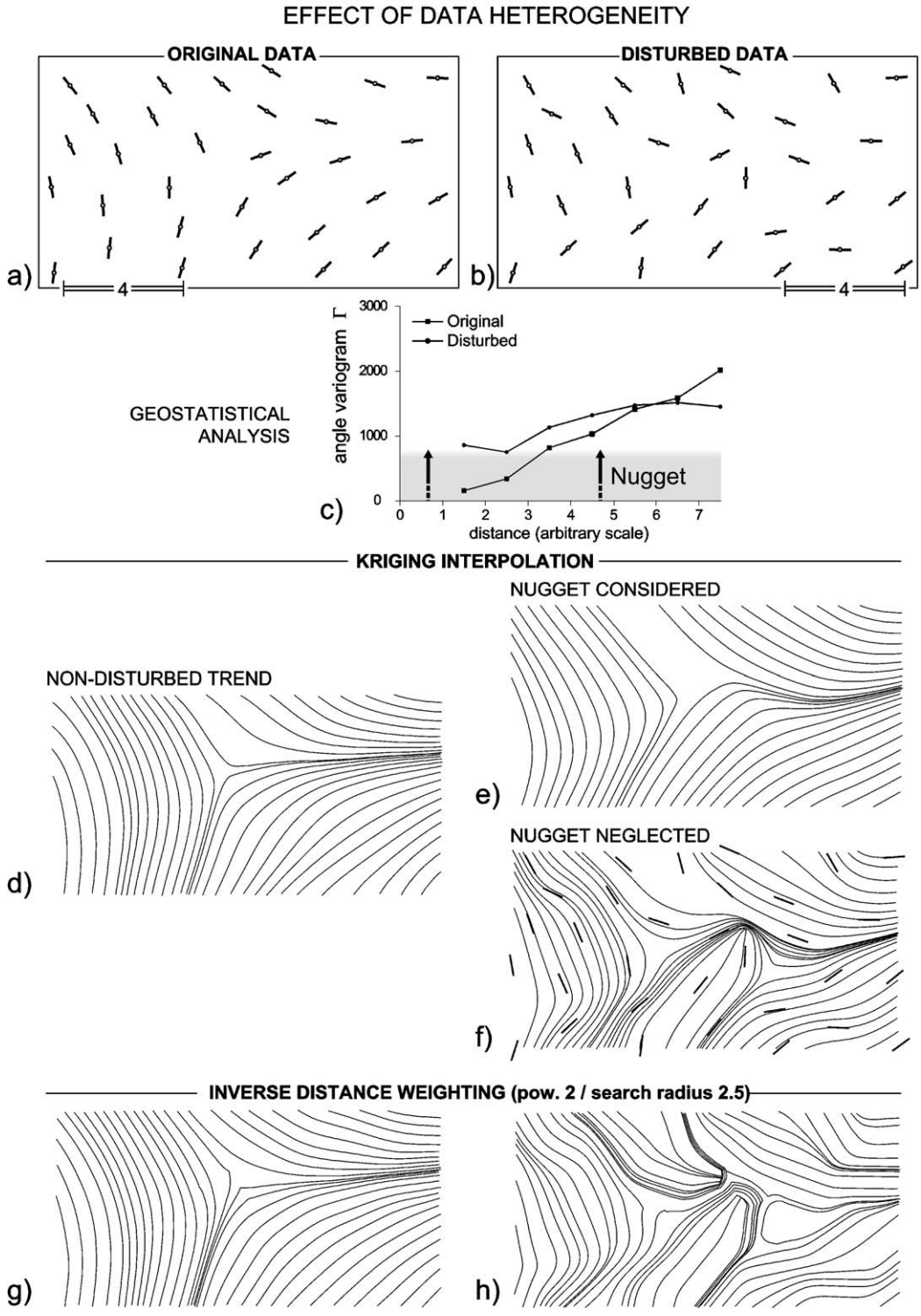
Individual data can vary strongly in space, even for close sampling locations. This can be due to the errors attached to initial measurements. It can also result from mixing of data of different origin (e.g. data from maps of different type, scale or age), and from subsequent computations (digitization, calculations). Strong local variations can also reflect very local effects that are not directly linked to the regional-scale tendencies (e.g. see Fig. 3).

In order to examine the effect of introducing noise within the data set, we start with the ideal example presented in Fig. 6. The initial data set (Fig. 7a) has been combined with a random noise (Fig. 7b). To do this, an angular positive or negative deviation comprised between zero and 40–50° was randomly applied to each initial data. Corresponding omnidirectional variograms are shown in Fig. 7c. The diagram shows that the variogram for the original data has a negligible ordinate at the origin, i.e. for short separation distances between data points. This reflects the smooth and continuous variations of directions. The variogram for the disturbed data presents a clear nugget effect of about 800 deg^2, which corresponds to a standard deviation of about 40°. This value is comparable to the maximum deviation introduced on initial individual data. With increasing distance, both variograms show more or less linear evolutions on a relatively large part of the diagram. This comparable trend suggests that both sets of data record some common regional-scale correlations. The lack of sill on the variogram for the initial data set indicates that data are correlated at distances larger than the map. In contrast, a sill around 1200 deg^2 could be present on the variogram for the disturbed data set. This suggests that the statistical analysis combines regional correlation effects with more local effects. Indeed, if very local correlations due to the background had completely obliterated the initial correlations, then only the nugget effect would appear on the variogram. The directions have been interpolated for the two data maps (Fig. 7a,b). The resulting grids are presented in the form of trajectories (Fig. 7d,e). Trajectories are very similar for both maps. This shows that, even for the highly disturbed data set, regional-scale correlations are strongly favoured with respect to local ones. This is because the nugget effect has been taken into account in the interpolation processing. If this effect is neglected, then the interpolation produces a result that strongly fits the original data set at local scale, and thus favours local correlations with respect to regional ones (Fig. 7f).

For comparison, the interpolation method based on the inverse distance-weighting is shown in Fig. 7g,h for the undisturbed and disturbed data sets, respectively. Fig. 7g (same as Fig. 6c) shows consistent trajectories. In contrast, trajectories for the disturbed data set do not reflect the trends observed for the undisturbed set. The pattern of trajectories is in fact quite similar to the one obtained by the geostatistics neglecting the nugget effect (Fig. 7f). This shows that the inverse-distance weighting tends to strongly favour local correlations, i.e. individual original data.

This example emphasizes that the geostatistical method, via the variogram analysis, allows to separate local effects and (or) errors in data from significant regional-scale tendencies.

Fig. 7. Illustration of the effect of background noise within data sets submitted to interpolation. (a) Original data set. (b) Disturbed data set. (c) Corresponding variograms; the background introduced on (b) is expressed by a nugget effect. (d, e) Trajectories from the geostatistical method applied to (a) and (b), respectively. (f) Trajectories from the geostatistical method applied to (b) without taking the nugget into account. (g, h) Trajectories from the inverse distance-weighting method applied to (a) and (b), respectively.


EFFECT OF DATA HETEROGENEITY
ORIGINAL DATA
DISTURBED DATA
a)
b)
4
4
GEOSTATISTICAL ANALYSIS
c)
angle variogram Γ
3000
2000
1000
0
Original
Disturbed
Nugget
0 1 2 3 4 5 6 7
distance (arbitrary scale)
KRIGING INTERPOLATION
NUGGET CONSIDERED
NON-DISTURBED TREND
e)
d)
NUGGET NEGLECTED
f)
INVERSE DISTANCE WEIGHTING (pow. 2 / search radius 2.5)
g)
h)

4.3. Analysis of irregularly spaced data

During field studies, measurements or calculations of directional data can be done from micro-scale to outcrop-scale, and even up to regional scale (e.g. analysis of satellite images). It is admitted that a highest density of measurements and a regular distribution of them throughout the studied area are optimal conditions to constrain the spatial variations of the considered parameters (e.g. Lee and Angelier, 1994; Mukul, 1998). Unfortunately, data sets are often heterogeneous, because of outcrop conditions and of logistics.

In order to test the strength of the geostatistical analysis with respect to sampling heterogeneity, we again use the example introduced in Fig. 6. The initial regular distribution of data (Fig. 8a) has been modified by introducing substantial sampling heterogene-

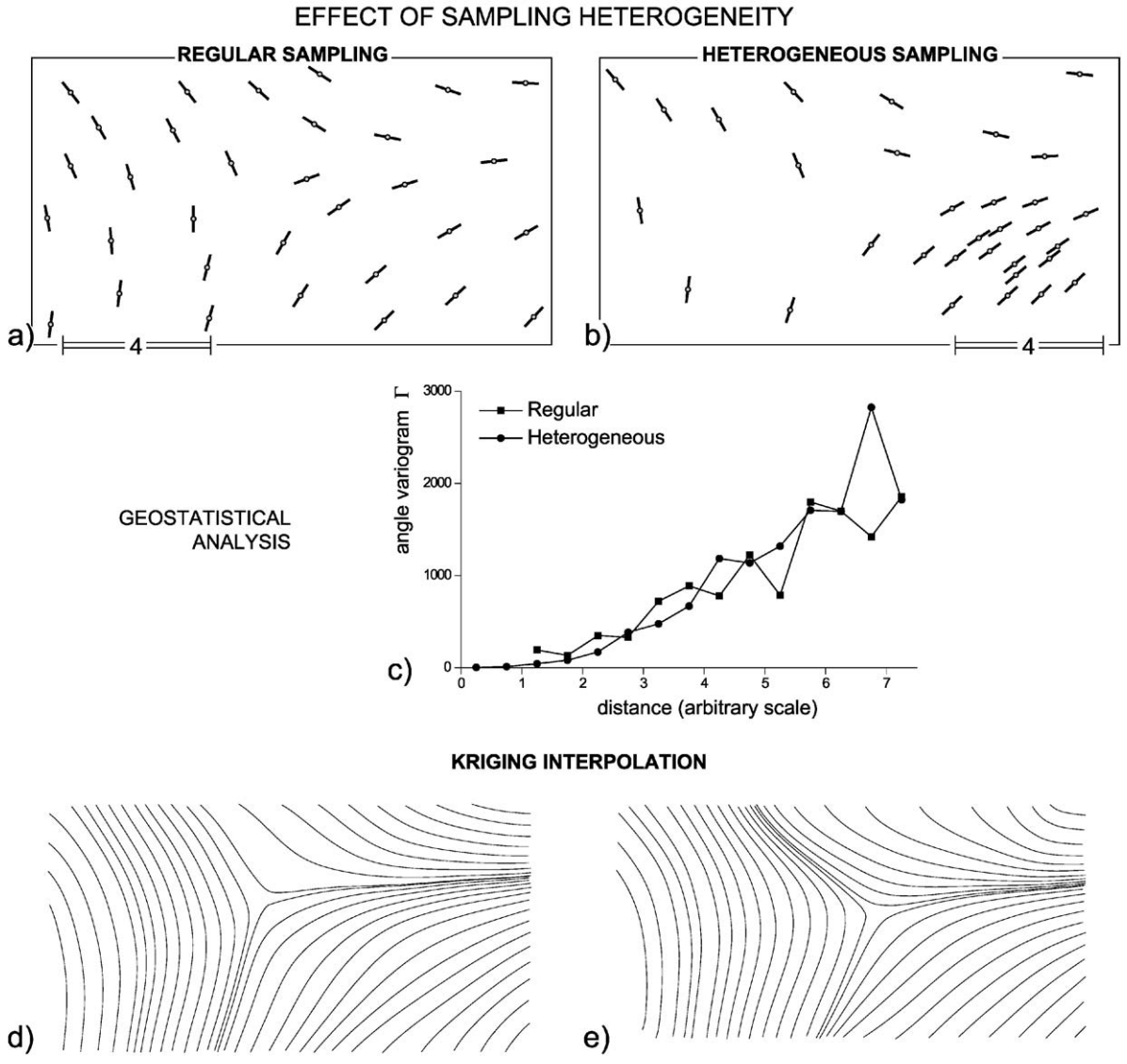


Fig. 8. Illustration of the effect of sampling heterogeneities on interpolation results. (a) Original data set. (b) Heterogeneous sampling of the same direction field as (a). (c) Corresponding variograms. (d, e) Trajectories from data sets (a) and (b), respectively.

ities (Fig. 8b). The variogram for these irregularly spaced data is reported in Fig. 8c, together with the one coming from the original regularly spaced data. Both variograms follow the same trend. This comparison highlights that the variogram analysis remains powerful even for nonideally distributed data. On the basis of the variogram analysis, the kriging interpolation has been performed for both data sets (Fig. 8d,e). The trajectory maps are again very similar. This shows that the interpolation method is not much sensitive on the initial data distribution, even where highly clustered data occur (Fig. 8b, bottom right). If we compare in detail the resulting trajectory maps, we can see that the position of the singular point, lying in the middle of the rectangle, is slightly shifted to the right for the interpolation of the irregularly spaced data points (compare Fig. 8d and e). In the latter, the singular point corresponds to a zone relatively depleted in number of data. The shift of the singular point shows that a minimum of measurements is necessary to interpolate correctly the directions within highly variable zones.

5. Examples of application to natural strain fields

5.1. Shear zones

Shear zones are very common structures where the characterization of strain gradients can be useful for the analysis of strain patterns and attached kinematics. Across shear zones, increasing shear strain is marked by the rotation of the principal strain plane (cleavage or foliation), and by a convergence of strain trajectories (Ramsay and Graham, 1970).

The geostatistical method has been applied to a famous example of small-scale shear zone described by Ramsay and Graham (1970) (Fig. 9a). To do this, a sampling of foliation directions has been made on a photograph of the shear zone (Fig. 9b). Interpolated data are presented on a trajectory map (Fig. 9c) and on a direction contour map (Fig. 9d). Consistent rotations of the foliation and tightening of trajectories with increasing shear strain are well expressed on the trajectory map. Some additional information can be found in the contour map. Thus, along-strike variations of principal strains are observed and define three domains. The central domain corresponds to the overall strike of the shear zone, and appears as the narrowest part of it (Fig. 9a). Domains to the left and to the right correspond to wider shear zone segments.

5.2. Syntectonic plutons

Strain patterns within and around syntectonic plutons emplaced in the ductile crust generally reflect the interference between strains due to pluton emplacement (e.g. ballooning effect; Ramsay, 1989) and strains due to regional tectonics (see Brun and Pons, 1981). This leads to complex foliation patterns that can in particular include singular points (neutral or triple points) and concentric fabrics around pluton boundaries.

The example presented here is a small syntectonic Hercynian pluton (Landudal leucogranite) from the western part of Central Brittany (France) (Fig. 10a). In this area, Hercynian deformations are related to regional dextral strike-slip associated with the South Armorican Shear Zone (Fig. 10a) (Gapais and Le Corre, 1980; Hanmer et al., 1982; Gumiaux et al., in press). At regional scale, strike-slip deformation resulted in the development of a subvertical cleavage bearing a subhorizontal stretching lineation.

Line segments in Fig. 10b show the strike of the cleavage measured in the deformed sediments (Hanmer et al., 1982). At distance from the pluton, fabrics strike EW to ENE–WSW. At the vicinity of the pluton, the cleavage shows rather variable orientations.

A variogram analysis and an interpolation have been performed on these direction data (Fig. 10c). On the map, the trajectories show that the mean azimuth of the cleavage is around N70°–N80°, which corresponds to the regional-scale orientation of the cleavage in this part of Central Brittany (Gumiaux et al., in press). This regional trend is consistent with dextral strike-slip along the South Armorican Shear Zone. In the middle of the study area, where the pluton is located, the interpolation procedure calculates elliptical trajectories (dashed lines). Moving away from the pluton, trajectories tend to wrap around pluton boundaries. At pluton tips, elliptical fabrics (dashed lines) and trajectories wrapping around the pluton (full lines) define cleavage triple points. This over-

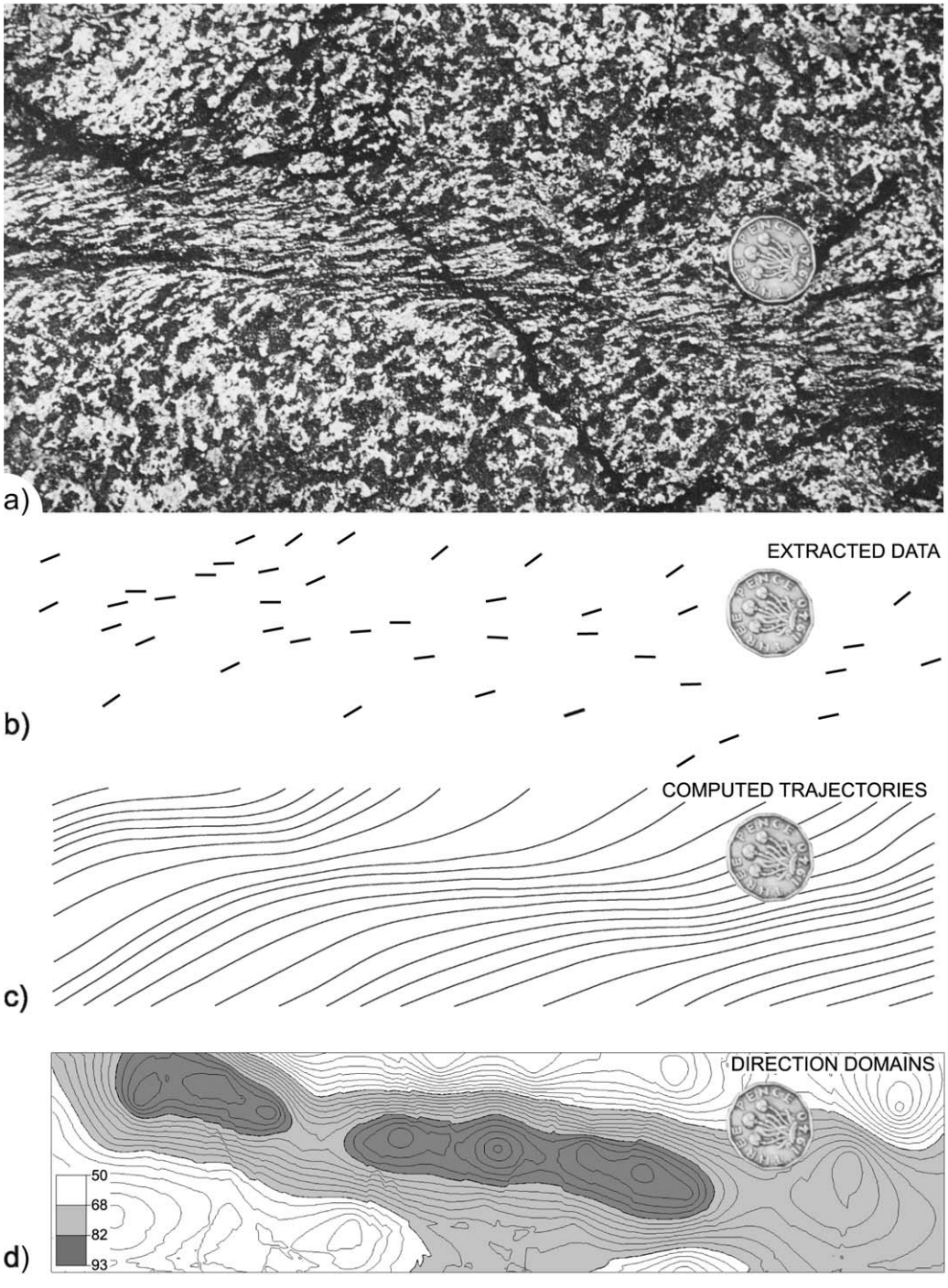


Fig. 9. Example of application of the geostatistical method on a natural shear zone (a) (photograph from Ramsay and Huber, 1983, Fig. 3.3, p. 36). (b) Sampled data set. (c) Corresponding strain trajectories. (d) Direction contour map showing the occurrence of along-strike domains. See text for further comments.

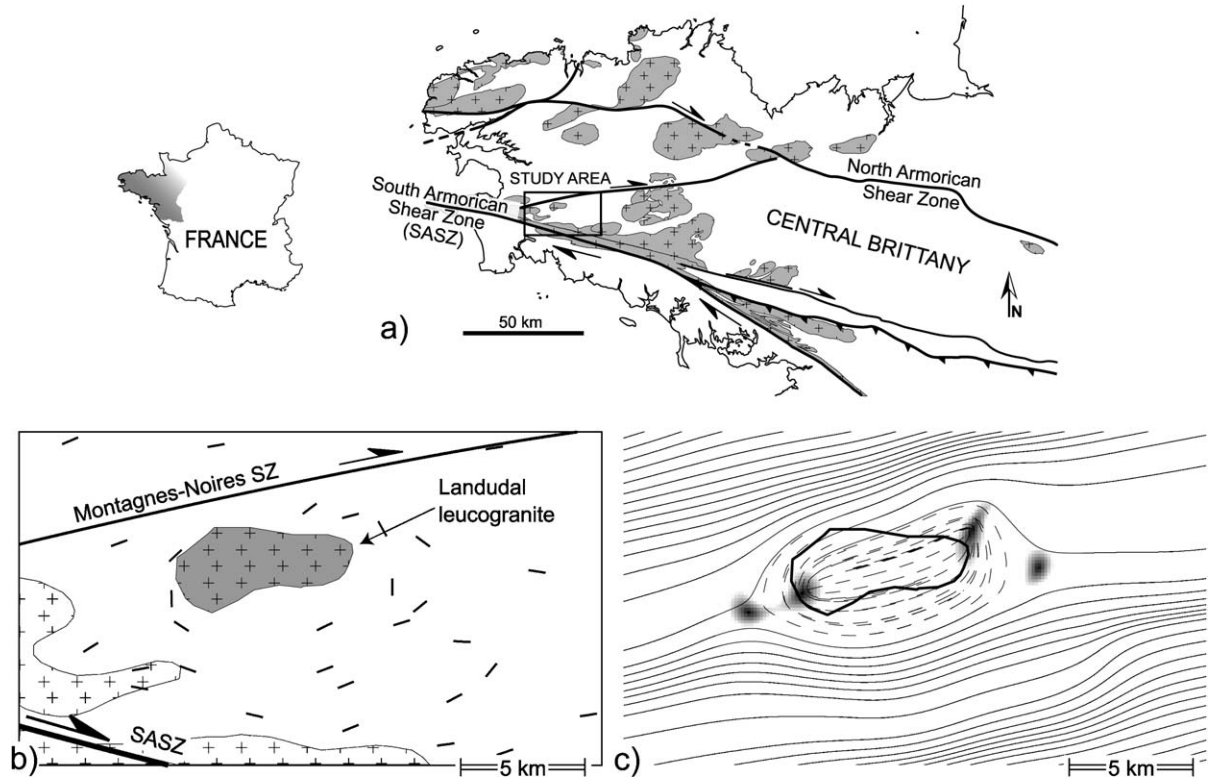


Fig. 10. Example of application of the geostatistical method to the analysis of strain patterns in and around a syntectonic pluton. (a) General sketch map of the area. (b) Cleavage direction data. (c) Computed trajectory map. See text for further comments.

all pattern of strain trajectories is consistent with some kind of pluton ballooning during regional shearing.

The occurrence of triple points can be further tested. During kriging, the direction cosines are used to calculate the interpolated direction. For a given set of orientation data, the length of the mean vector defined by the direction cosines is a direct estimate of the local direction variance. This parameter represents the degree of data concentrations around the mean vector, and has been defined as the parameter R (see Upton and Fingleton, 1989). When R equals 1, all the directions considered are parallel and the resulting mean is perfectly constrained. With decreasing R value, the resulting mean angle becomes less constrained, down to undefined when R equals zero. This is what is expected to occur within a neutral point (triple point) (see Fig. 6). On Fig. 10c, low R values have been contoured, with increasing grey levels from R equals 0.5 to zero. From numerical models, Brun (1981) demonstrated that two pairs of neutral zones, inside and outside the pluton, could develop during syntectonic ballooning. These zones are located at pluton tips, at high angle to the regional shortening direction. This is actually consistent with our interpolation results, which show four zones of low R values. Two of them are well outside the pluton and are clearly located within the cleavage triple points defined by cleavage trajectories. The two more internal ones are located at the tips of the area marked by elliptical trajectories.

Thus, the interpolation method provides consistent results for this complex example where rather few orientation data are available (Fig. 10b).

6. Concluding remarks

This study has emphasized that statistical analysis of directions cannot be made with circular data.

To avoid this problem, the use of direction cosines during automatic data treatment appears a good solution.

Geostatistics, in particular the analysis of variograms, permit to describe and quantify the spatial dependences of values from any kind of data set. The method has been here developed for directional data. The variogram analysis allows (1) to quantify the scale of data correlation, (2) to detect and quantify anisotropies in data variations, and (3) to quantify local effects, as well as inherent errors included within original data, and separate them from regional effects.

Statistical parameters provided by variograms are critical to constrain the way directional data should be interpolated. Indeed, during interpolation, data weighting depends on the best-fit equation of the variogram curve. The strength of the interpolation method is outlined by the treatment of nonideal theoretical data sets, as follows.

- The influence of data heterogeneity, like a substantial background noise, appears rather limited. This is because the nugget value obtained from the variogram analysis is taken into account during interpolation.
- The influence of heterogeneously distributed data appears also limited.

The method is convenient to draw direction trajectories. It can also provide further information. We have in particular emphasized the followings.

- Contour maps, combined with direction histograms, are convenient to detect domainal distributions of directions.
- The interpolation leads to quantify the degree of local consistency and of correlation between neighbouring data. For example, well-oriented domains can be distinguished from less-oriented ones. Thus, in the example of pluton intrusion presented in this paper, cleavage triple points have been detected.

Compared with other methods of analysis of directional data developed in geosciences, the geostatistical method appears particularly efficient, reliable, and can be applied at all scales. We have focused this paper on the analysis of strain trajectories, but the method could be applied to any kind of other directional data, like displacement vectors (e.g. from GPS measurements), fault patterns, stress fields, paleocurrents, topographic data, paleomagnetic data, or directional data extracted from geophysical maps or cross-sections (e.g. seismics, gravity and magnetic data).

Acknowledgements

This study was made in the course of the research project Armor 2 (“Géofrance 3D” Program, BRGM-CNRS). Constructive discussions with C. Walter

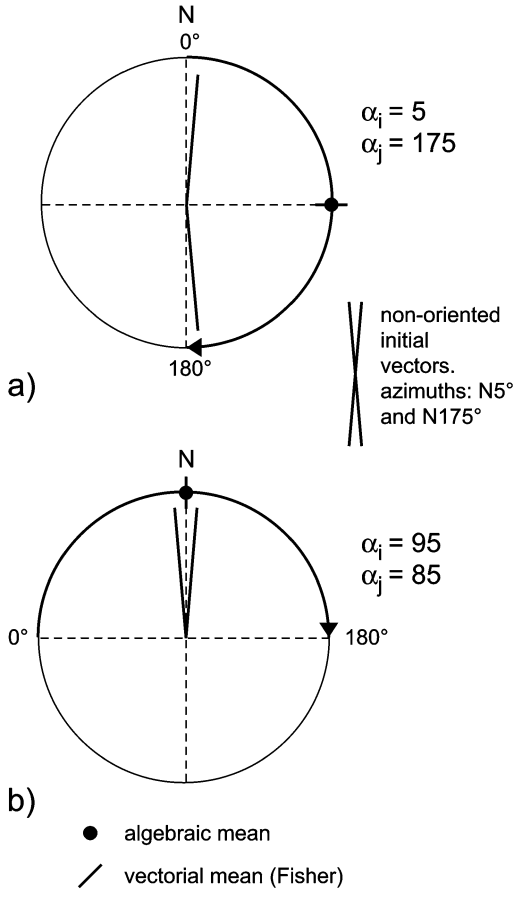


Fig. 11. Illustration of problems attached to the calculation of mean values from circular angular data.

(INRA, Rennes) strongly contributed to the methodological aspects developed in this paper. C.G. acquired a large part of the bases of the geostatistic analysis in the on-line lessons provided by D. Marcotte (Ecole polytechnique de Montréal, http://geo.polymtl.ca/~marcotte/). Reviews by J. Escuder-Viruete and D. Marcotte helped us to improve the paper.

Appendix A

Directional data are particular variables characterized by a circular property. Angles are indeed expressed modulo π for nonoriented directions or modulo 2π for vectors. It follows that none of the classical statistical calculations can be directly performed on these data. Particular treatments for directions have been developed in detail elsewhere (e.g. Mardia, 1972; Upton and Fingleton, 1989; Marcotte, 2001), and only a brief summary is given here.

A.1. Transformation of angular data for variogram construction

The construction of a variogram requires the use of angular differences $\Delta\alpha_{i,j}$ between pairs of directional data i and j (Fig. 1). In this case, the angular variable is independent of the chosen reference frame. It

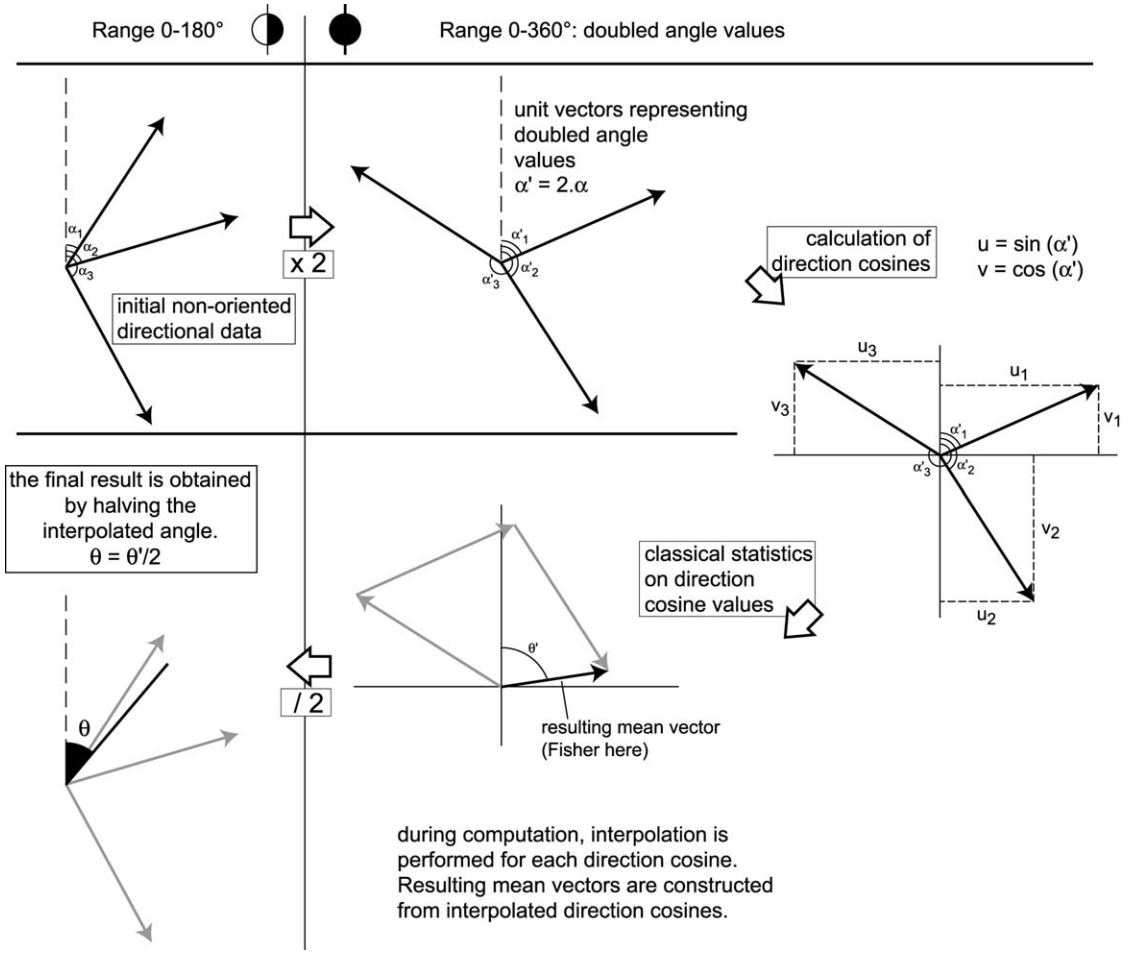


Fig. 12. Illustration of the different steps followed during processing and interpolation of directional data. See text for further explanations.

follows that problems induced by the circularity of angular variables can be solved by a simple data correction modulo π (Upton and Fingleton, 1989):

$$\Delta\alpha_{i,j} = \alpha_i - \alpha_j \quad \text{if } (\alpha_i - \alpha_j) < 90^\circ$$

and

$$\Delta\alpha_{i,j} = 180^\circ - (\alpha_i - \alpha_j) \quad \text{if } (\alpha_i - \alpha_j) \geq 90^\circ$$

For example, for a pair of directions striking N5° and N175°, the algebraic difference is $(175 - 5^\circ) = 170^\circ$ (Fig. 11), which exceeds 90° and must therefore be changed to 10°. This value then becomes the same as for angular differences falling in the 0–90° angular range (e.g. $95 - 85^\circ = 10^\circ$) (Fig. 11b).

A.2. Transformation of angular data for kriging interpolation

Problems induced by the circularity of orientation data in their statistical analysis are further critical for the interpolation between data than for the variogram calculation. Indeed, the interpolation procedure involves computations using angles and angular means defined with respect to a given reference frame. Therefore, additional transformations of original data are required. This is illustrated in Fig. 11, that shows an example of two data oriented N5° and N175°.

For this example, the resulting mean direction is intuitively estimated as striking north–south, irrespective of the reference frame (NS, Fig. 11a, and N270°, Fig. 11b). Problems arise when the arithmetic mean is directly calculated on these values $(5^\circ + 175^\circ)/2 = 90^\circ$ (Fig. 11a). Staying in the same NS reference frame, the estimate of the mean Fisher vector also yields an east–west trend, which is at 90° from the correct north–south result (Fig. 11a). For this simplistic example, one can see that if the origin is changed from north to west, the algebraic mean calculation, as well as that of the mean vector, provides the correct result (Fig. 11b). However, this transposition, which is trivial for two directions, becomes much less straightforward for a set of many values.

Different techniques have been proposed to treat such nonoriented directions (e.g. Panozzo, 1984; Lee and Angelier, 1994). A simple one, which we have followed, has been described by Upton and Fingleton (1989).

The first step consists in a preliminary doubling of the algebraic values of the nonoriented directions (Upton and Fingleton, 1989) (Fig. 12). The 0–180° range of directions is thus enlarged to the entire 0–360° range. Each direction is then considered as a unit vector. This vector is projected on each of the reference axes in order to calculate the corresponding direction cosines (Fig. 12). These latter variables are not circular, and classical calculations can be performed on them. This is for example illustrated in Fig. 12 for the estimate of a mean vector. On this basis, the kriging interpolation is performed, yielding interpolated directional data expressed by their direction cosines. The last step consists in the calculation of corresponding angles, which are halved in order to reduce the angular range from 0–360° to the initial 0–180° range (Fig. 12).

References

Angelier, J., 1994. Fault slip analysis and paleostress reconstruction. In: Hancock, P. (Ed.), Continental Deformation. Pergamon, New York, pp. 53–100. Chapter 4.

Barbotin, E., 1987. Trajectoires de déformation finie et interprétation cinématique. Modèles numériques et application à des exemples régionaux (Bretagne Centrale et Alpes Penniques). Thèse de l'Université de Rennes 1. 158 pp.

Bourennane, H., 1997. Etude des lois de distribution spatiale des sols de Petite Beauce. Thèse de l'Université d'Orléans. 229 pp.

Brun, J.P., 1981. Instabilités gravitaire et déformation de la croûte continentale: application au développement des dômes et des plutons., Thèse d'Etat, Université de Rennes 1. 211 pp.

Brun, J.P., Pons, J., 1981. Strain patterns of pluton emplacement in a crust undergoing non-coaxial deformation, Sierra Morena, Southern Spain. J. Struct. Geol. 3, 219–229.

Cobbold, P.R., 1979. Removal of finite deformation using strain trajectories. J. Struct. Geol. 1, 67–72.

Cobbold, P.R., Barbotin, E., 1988. The geometrical significance of strain trajectory curvature. J. Struct. Geol. 10, 211–218.

Cressie, N.A.C., 1990. The origins of kriging. Math. Geol. 22, 239–252.

Cressie, N.A.C., 1991. Statistics for Spatial Data. Wiley, New York. 900 pp.

Deutsch, C.V., 1991. The relationship between universal kriging, kriging with an external drift and cokriging. SCRF Report 4, May, 1991.

Deutsch, C.V., Journel, A.G., 1992. GSLIB: Geostatistical Software Library and User's Guide. Oxford Univ. Press, New York. 384 pp.

Gapais, D., Le Corre, C., 1980. Is the Hercynian belt of Brittany a major shear zone? Nature 288, 574–576.

Gumiaux, C., Brun, J.P., Gapais, D., in press. Strain removal within the Hercynian Shear Belt of Central Brittany (Western France): methodology and tectonic implications. J. Geol. Soc. (Lond.), Spec. Pub.

Hanmer, S., Le Corre, C., Berthé, D., 1982. The role of Hercynian granites in the deformation and metamorphism of Brioverian and Paleozoic rocks of Central Brittany. J. Geol. Soc. (Lond.) 139, 85–93.

Isaaks, E., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford Univ. Press, New York. 561 pp.

Lajaunie, C., Courrioux, G., Manuel, L., 1997. Foliation fields and 3D cartography in geology: principles of a method based on potential interpolation. Int. Assoc. Math. Geol. 29, 571–584.

Lee, J.C., Angelier, J., 1994. Paleostress trajectory maps based on the results of local determinations: the ''Lissage'' program. Comput. Geosci. 20, 161–191.

Marcotte, D., 2001. Géologie et géostatistique minières (partie géostatistique). Online lessons, http://geo.polymtl.ca/~marcotte/.

Mardia, K.V., 1972. Probability and Mathematical Statistics. Academic Press, London. 355 pp.

Matheron, G., 1962. Traité de Géostatistique Appliquée. Technip, Paris. 333 pp.

Mukul, M., 1998. A spatial statistics approach to the quantification of finite strain variation in penetratively deformed thrust sheets: an example from the Sheeprock Thrust Sheet, Sevier Fold-and-Thrust belt, Utah. J. Struct. Geol. 20, 371–384.

Panozzo, R., 1984. Two-dimensional strain from the orientation of lines in a plane. J. Struct. Geol. 6, 215–221.

Potter, P.E., Pettijohn, F.J., 1977. Paleocurrents and Basin Analysis. Springer-Verlag, Berlin. 425 pp.

Ramsay, J.G., 1989. Emplacement kinematics of a granite diapir: the Chindamora Batholith, Zimbabwe. J. Struct. Geol. 11, 191–209.

Ramsay, J.G., Graham, R.H., 1970. Strain variations in shear belts. Can. J. Earth Sci. 7, 786–813.

Ramsay, J.G., Huber, M.I., 1983. The techniques of modern structural geology. Strain Analysis, vol. 1. Academic Press, London. 307 pp.

Upton, G.J.G., Fingleton, B., 1989. Spatial Data Analysis by Example, vol. 2. Wiley, New York. 416 pp.

Wackernagel, H., 1998. Multivariate Geostatistics. Springer Verlag, Berlin. 291 pp.

Xu, W., 1996. Conditional curvilinear stochastic simulation using pixel-based algorithms. Math. Geol. 28 (7), 937–949.

Young, D.S., 1987. Random vectors and spatial analysis by geostatistics for geotechnical applications. Math. Geol. 19 (6), 467–479.