Planar Contact Contour Plotting Using Spatial Methods

Peter B Dixon

School of Computing & Technology
Nottingham Trent University, Nottingham NG1 4BU, United Kingdom.
peter.dixon@ntu.ac.uk

***Abstract* – To analyse problems of contact between surfaces, finite element methods can be employed to produce continuous contour plots at interpolated as well as grid points, but it is not possible to indicate the extent to which the estimated contact values are subject to error variation. Methods of spatial statistics are investigated here to provide contour plots for contact, together with accompanying plots indicating the varying prediction error. An example is given in the planar contact case and techniques are explored for the case of contact on a three-dimensional surface in the form of a gear tooth using spheroidal projections.**

Keywords* - *Contact, contour plotting, spatial methods

I. INTRODUCTION

A. Finite Element Method

Contact readings obtained by automatic means may be used to produce contact contour plots via Finite Element Methods. See [Su, 2000]. However, with this computational method of trend surface fitting there appears to be no way of measuring the extent to which predicted values for the formation of the contour are subject to error variation.

B. Least Squares Estimation

Least squares or generalised least squares may be used, in which not only can the trend surface be estimated and a contour plot obtained, but a global standard error of estimation will provide a measure of the prediction error across the fitted surface. See [Venables, 1994]. Unfortunately, this does not allow for the fact that the prediction error is almost certain to vary across the surface. In order to examine this problem it is necessary to turn to methods of spatial statistics.

C. Spatial Statistics

In the case of contact studies, contact readings are taken at regular intervals across and down the field, and eventually measurements are obtained at all or most of the points sampled. How do the readings relate to the whole field? Indeed, what can be said about the expected contact at places not sampled at, given the knowledge gained at the places sampled? And how densely should one have sampled in order for such information to be reliable?

The contact environment is continuous, but properties are measured at only a finite number of places. Elsewhere, the best that can be done is to estimate or predict in a spatial sense. This is the justification for spatial statistics.

Spatial statistics can help to estimate probabilities that contact values exceed specified thresholds.

II. NOMENCLATURE

E	expected value
γ	semi-variogram
Z	contact
ϕ	latitude
h	spatial resolution
λ	longitude
z	contact observation
λ_i	Kriging weight
ρ	autocorrelation function
σ^2	process variance
θ	polar angle

III. SPATIAL METHODS

A. Autocorrelation

An important aspect in applying statistical methods to surface contact problems is to realise that the problems are spatial, in the sense that contact measurements at points close together tend to be similar (they are closely related; ie. autocorrelated) whereas those further apart differ more (their autocorrelation is lower, in fact the readings might be independent). It is possible to express this knowledge quantitatively, and uses it for prediction. For example, in predicting the contact level at a nonsampled point, greater weighting would be given to the measured contact level at nearby sampled points than at distant sampled points. A further point for consideration is that the relationship between contact levels at various distances apart may be different in a direction taken across the field as opposed to down the field, so any summary measure of this property needs to be directional.

B. The Variogram

In preference to the autocorrelation function, the key statistic for measuring the way in which contact levels Z at n points x1, x2, x3 ,…relate to each other at different spatial resolutions h in a particular direction, say across the field, is the n-termed sample variogram

$$2\gamma(h) = \sum_{i=1}^{n} (z(x_i) - z(x_i + h))^2 \quad (1)$$

which can be plotted against h on a graph. This is an important tool in spatial statistics for understanding spatial variation in data. Variograms computed for different directions can be used to show whether there is anisotropy, and what form it takes. Note that frequently $\gamma(h)$, the semi-variogram is used (with no loss of meaning).

It is useful for comprehension to know that, for second-order stationary processes Z(x),

$$\gamma(h) = Cov(0) - Cov(h)$$

And

$$\gamma(h) = \sigma^2(1 - \rho(h)) \quad (2)$$

where Cov( ) and ρ( ) are the covariance and autocorrelation at lag ( ), and σ^2 is the process variance.

C. Variogram Models

Variograms are usually monotonic increasing. They usually reach an upper bound called the sill, and level off. The lag h at which the sill is reached is called the range of influence; contact levels will be uncorrelated at this and greater separations. Spatial variation is not necessarily the same in all directions; the profile of contact levels across the field could be different from those down; this is called anisotropy.

Here are some possible models.

C1. Spherical Model:

$$\gamma(h) = C\left(\frac{3h}{2a} - \frac{h^3}{2a^3}\right)$$

$$\gamma(h) = C \quad (3)$$

according to whether:

$0 < h \le a$

or

h>0

C2. Exponential Model

$$\gamma(h) = C(1 - e^{-h/a}) \text{ for } h>0 \quad (4)$$

A range of computer software permits the plotting of data postings, sample variograms and the fitting of an appropriate variogram model. Different variogram models might be required for different directions. The spatial variation in a field could be different in the North-South direction as opposed to the East-West direction. Variogram models may be fitted uni-directionally, multi-directionally or omni-directionally.

D. Estimation at Non-Sampled Points: Kriging

Suppose now that a data set has been examined, sample variograms have been plotted and variogram models fitted for various orientations (directions). It is possible to estimate the value of the study variable across the whole range of study, at non-sampled points. In predicting, say, the contact at a nonsampled point, greater weight would be given to the measured contact at nearby sampled points than at distant sampled points.

Estimation can be carried out by traditional regression and/or trend surface methods but a technique known as kriging (after D G Krige, see [Krige, 1966]) is preferred as it provides contact point estimates together with estimates of the local variation at the estimated points.

An estimate of the contact level at A could be:

$$\hat{z}_A = w_1 z_1 + w_2 z_2 + \ldots w_n z_n \quad (5)$$

where the weights wi could be some inverse function of distance. Estimation at each point x_i and a measure of the local variation at that point are functions of distance and variogram values through the data. The Method of Lagrange multipliers is used to find the weights that minimise the estimation variances. A set of 'kriging' equations is formed, and solved simultaneously. Calculations are usually performed on appropriate computer software, such as Ecosse or S-Plus. A contour plot or 3 dimensional trend surface plot of the estimated response values and their standard errors in relation to a two-way grid (eg. North/South versus East/West) can be presented.

The contour plots in Figures 1 and 2 were obtained using Ecosse software [Geostokos, 1998] for appropriate data.

Kriging provides estimates based on a continuous model of spatial variation.

It makes best use of existing knowledge by taking account of the way a property varies in space indicated by the variogram. In its simplest form a kriged estimate is just a linear sum or weighted average of the data in its own

DOI 10.5013/IJSSST.a.22.03.02 ISSN: 1473-804x online, 1473-8031 print

neighbourhood. Model selection for the variogram(s) is carried out by Ecosse in the routines for the above plots.

For details see [Webster, 2001]. The theory is briefly discussed below.

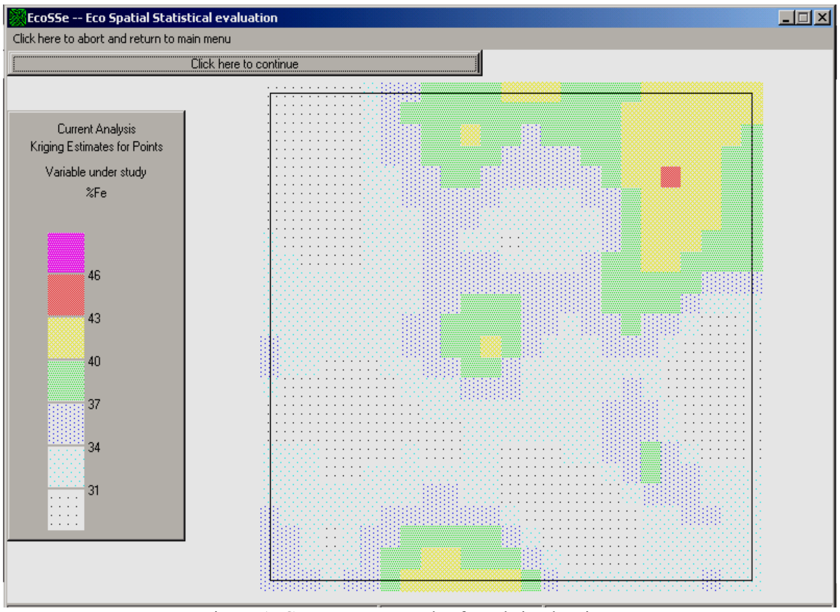

Figure 1. Contact contour plot from kriged estimates

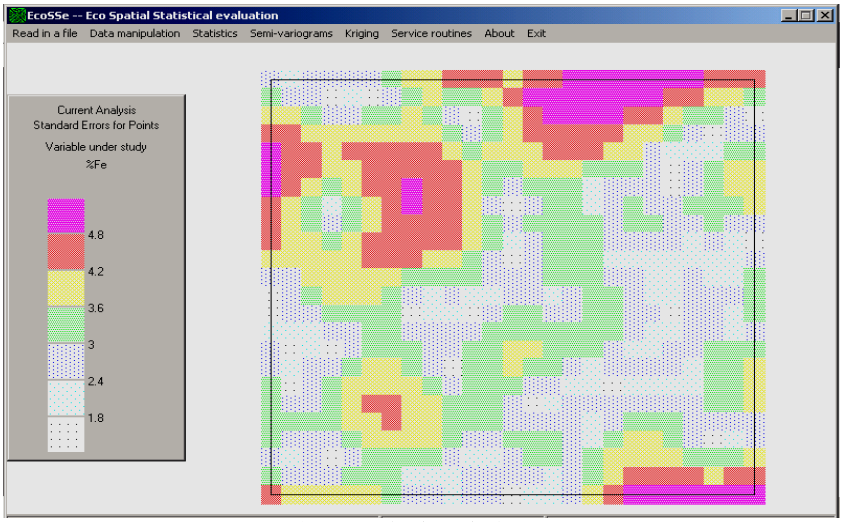

Figure 2. Kriged standard errors

IV. THEORY OF ORDINARY KRIGING

A kriged estimate at x0:

$$\hat{z}(x_0) = \sum_{i=1}^{N} \lambda_i z(x_i) \tag{6}$$

where the λ_i are made to sum to 1, is accompanied by a kriging variance

$$\begin{aligned}\operatorname{var}(\hat{z}(x_0)) &= E((\hat{z}(x_0) - z(x_0))^2) \\ &= 2\sum_{i=1}^{N} \lambda_i \gamma(x_i, x_0) - \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j \gamma(x_i, x_j)\end{aligned} \tag{7}$$

with $\gamma(a,b)$ denoting semivariance between a,b. The next step is to find the weights that minimise these variances, using Lagrange multipliers. This leads to a set of N+1 equations in N+1 unknowns. The solution gives the weights that will minimise the estimation variance. By inserting the weights into the above expressions for:

DOI 10.5013/IJSSST.a.22.03.02 ISSN: 1473-804x online, 1473-8031 print

$$\hat{z}(x_0)$$

the estimate at x0 and its error variance are obtained.

V. GEAR TOOTH CONTACT

A. Sphere-to-Plane Projections

It is clear from [Dabnichki,1999] that reasonably full solutions to gear tooth contact problems may be obtained from a 2-D approach. Three-dimensional approaches via kriging are now outlined.

In the field of gear tooth contact, the surface of a three-dimensional gear tooth may be assumed to a reasonable degree of approximation to resemble that of a half-sphere or parabola. With notation borrowed from geodesics, suppose a point on the surface of a sphere has spherical coordinates (ϕ, λ), equivalent to latitude and longitude. Projection onto a plane can be defined by 'easting' E and 'northing' N given by:

$$E = f_1(\phi, \lambda)$$

And

$$N = f_2(\phi, \lambda) \tag{8}$$

Where f_1 and f_2 are specified functions. These functions are chosen so that some property or properties is/are preserved after projection. Here are a few instances. In the case where distances and angles are distorted minimally, the projection is orthomorphic. Examples are the stereographic, mercator and transverse projections. It not possible to determine a single equidistant projection over which all distances are preserved, but one method that is used is the conic equidistant projection. In the case where areas are distorted minimally, the projection is authalic.

B. Azimuthal Projections

In the case of the half-sphere, an argument may be applied to the case where contact data on the gear tooth surface may be considered in relation to its azimuthal projection, as shown in Figure 3.

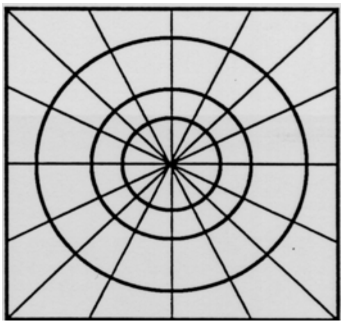

Figure 3. Contact data on the gear tooth surface in relation to its azimuthal projection

It is easy to imagine the gear tooth protruding outwards towards the viewer. See [Jackson, 1987].

Whatever the projection adopted, it is not essential in spatial statistics to use a rectangular, regular grid. Kriging calculations on projection-transformed contact data can now be carried out in relation to the nonrectangular grid. Then, kriged contact predictions and their standard errors over the plane can be reverse-transformed one-to-one onto the curved surface of the half-spherical shell.

C. The Case of the Azimuthal Equal-Area Projection

Here, any plane angle Θ at the origin of polar coordinates is equal to the angle of longitude for a globe of radius R, so $\Theta = \lambda$. It can be shown that, for equal-area:

$$r = 2R\sin\left(\frac{\pi/2 - \phi}{2}\right)$$

The inverse expressions are:

$$\phi = \frac{\pi}{2} - 2\sin^{-1}\left(\frac{r}{2R}\right)$$

and $\lambda = \Theta$

These may be written in terms of eastings and northings. Analytical derivations of a range of possible projections are given in [Maling, 1992].

D. Three-Dimensional Kriging

As an alternative to the above, kriging in three dimensions is commonly carried out on volume regions in many application areas of spatial statistics. However, on a gear tooth surface there is a paucity of contact data in relation to the complete 3-D volume space of the gear tooth. Three-dimensional kriging on a shell requires software with the capability to grid and interpolate sparse measured data in three dimensions, such as that produced by [Scientific Software Group, 1998] or [Boss Group, 2003], but the Ecosse software and S-Plus-based techniques outlined in [Venables, 1994] are adequate for the purpose.

E. Continuous Global Surfaces

[Billings, 2002] gives details of a number of methods for interpolation, including kriging and splines, and addresses in particular the case of irregular spatial distribution so that data are scattered in an irregular fashion. Computational methods are employed wherein there are automatic adjustments to varying data density. The case for 2-D problems is covered in depth but generalisation to many-dimensional space is catered for.

D. Review of Software

A useful review of software for the computational methods of spatial statistics is given in [Deutsch, 1998].

VI. CONCUSIONS

Analysis of contact problems between surfaces have employed finite element methods to produce continuous contour plots at interpolated and grid points, but were unable to indicate the extent to which the estimated contact values are subject to error variation. We outlined methods of spatial statistics to provide contour plots for contact, together with accompanying plots to indicate the varying prediction error. An example was given in the planar contact case and techniques were explored for the case of contact on a three-dimensional surface in the form of a gear tooth using spheroidal projections.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the interest of Professors D Su and A Sackfield in this work.

REFERENCES

[1] Billings S D, Beatson R K & Newsam G N, 2002, Interpolation of geophysical data using continuous global surfaces', Geophysics 67, 6
[2] Boss Group, 2003, 'FEMwater', GMS Deutsch C V & Journel A G, 1998, 'Geostatistical Software Library and User's Guide', Oxford University Press.
[3] Dabnichki P & Crocombe A, 1999, 'Finite element modelling of local contact conditions in gear teeth', Journal I Mech E
[4] Geostokos Ltd, 1998, Ecosse (Software) Jackson J E, 1987, 'Sphere, Spheroid and Projections for Surveyors', BSP Books.
[5] Krige D G, 1966, 'Two-dimensional weighted moving average trend surfaces for ore-evaluation', Journal Sth Afr Inst Mining/Metallurgy Scientific Software Group, 1998, Drill Guide
[6] Maling D H, 1992, 'Coordinate Systems and Map Projections', London/George Philip
[7] Su D, Yang F and Gentle C R, 2000, 'Finite element analysis for ZI worm gears with localised tooth contact', Proc. 8th International Power Transmission and Gearing Conference, ASME Design Engineering Technical Conferences, Baltimore, USA, (DETC 2000/PTG-14402)
[8] Venables W N & Ripley B D, 1994, 'Modern Applied Statistics with S-Plus', Springer
[9] Webster R & Oliver M A, 2001, 'Geostatistics for Environmental Scientists', Wiley

BIOGRAPHY

Peter Dixon is based in the School of Computing and Technology at the Nottingham Trent University. His interests are in the teaching and application of statistical methods across a variety of fields. His published work covers aspects of software-driven statistical applications in science, engineering, business and survey methodology.