

Spatial interpolation: an overview

Donald E. Myers

Department of Mathematics, University of Arizona, Tucson, AZ 85721, USA

(Received June 18, 1993; accepted after revision October 20, 1993)

ABSTRACT

The interpolation of spatial data has been considered in many different forms. The various forms of kriging are among the best known in the earth sciences although techniques such as inverse distance weighting were and are in use for spatially located data. In the numerical analysis literature various forms of splines and more recently radial basis functions have been developed and used. Because these techniques have been developed in very different contexts the relationship between them has not always been apparent. Various forms of kriging are considered as well as kernel estimators, splines and radial basis functions. By using the dual form of kriging and the positive definiteness property of the variogram connections are shown between splines, kriging and radial basis functions. One of the distinctions between kriging and other interpolators is the incorporation of the support of the samples and explicit estimation of linear functionals such as spatial integrals.

INTRODUCTION

Interpolation is a technique that is well-known to users of tables tabulating the values of a function, generally a function not easily computed. Interpolation in such cases is ordinarily used to obtain values with an additional decimal place of accuracy and the interpolation is linear. Interpolation may include two facets and often these are not distinguished. For example, suppose that the values of a function $f(x)$ are known (or tabulated) at points x_1, x_2, \dots, x_n . In the case of points in one-dimensional space one is interested in the value of $f(x)$ for $x_{i-1} < x < x_i$. Continuity of $f(x)$ is sufficient to ensure that linear interpolation is adequate when $x_i - x_{i-1} = a_i$ is small. That is:

$$f^*(x) = [(x - x_{i-1})/a_i]f(x_{i-1}) + [(x_i - x)/a_i]f(x_i)$$

will be very close to $f(x)$. The error of interpolation/estimation is $f^*(x) - f(x)$ but there may be other errors. The data may also incorporate errors, in this instance the data is being smoothed and the function is being interpolated at the same time. While the two nearest data points may be sufficient for simple interpolation, estimation of the errors in the data (smoothing) will usually require the use of additional data locations. The simple form of interpolation illustrated above does not easily extend to functions defined

on higher dimensional spaces, particularly when the data locations are irregularly spaced. In particular it is usually more appropriate to use additional data locations even when the data are error free. Although soil data are often collected on transects and along any one transect the problem is one-dimensional, the application is more commonly two-dimensional. For example, if the objective is to characterize sodium concentrations in a field, as in the work by Burgess and Webster (1980a-c), then it is not sufficient to interpolate only along the transects.

More generally, interpolation occurs in at least two contexts and is sometimes called estimation or prediction. In the first instance there is a known function given either by an analytic representation or as the solution of an integral or differential equation. The equation might be difficult to solve, the function might be difficult to evaluate or the objective might be to optimize the function. Finding the optimum will in general require solving differential equations, even when evaluation of the function is straightforward the solving differential equation(s) may not be.

Note that when interpolation is viewed as a form of approximation it is generally approached in a very different way. In the earth sciences there is another form of the interpolation problem that is more familiar. There is an unknown function f , values of f are known at a finite number of points in space, and the objective is to produce an approximation to the value of f at one or more points where the functional value is not known. Alternatively, the "values" might be linear functionals such as spatial integrals or derivatives, the objective might be to estimate a different linear functional, e.g., a spatial integral over a different sized or shaped region. When the function is unknown it is usually necessary to begin with a model, it may be that what appear to be different models produce the same results. The model in turn will ordinarily include one or more parameters that must be chosen or estimated from the data. Either explicitly or implicitly the model will lead to an interpolating or approximating function, and often the model is chosen to ensure that the interpolating function has certain desirable properties. The model may be either deterministic or stochastic, and the degree to which the interpolating function approximates the unknown function will have to be quantified in a corresponding manner, but in some fashion the interpolator must incorporate the extent to which the function is unknown. The interpolation problem arises in many contexts, it has been considered in statistics in several different forms, in numerical analysis and with various ad hoc solutions in application fields. When the function is unknown it may be more difficult to directly quantify the error or the uncertainty of estimation. The connection between estimation and quantification of the uncertainty may be approached in at least two ways. First, perhaps the most common, the estimator produces an estimate of the value of the function at a non-data point and there is also a measure of the uncertainty associated with that estimate

(for example, the variance of the error of estimation). Second, the uncertainty may itself be estimated but in such a manner that it is associated with an estimated value of the function. The first approach is typified by ordinary kriging and the second by indicator kriging. The unknown functional value may be estimated by representing it as the mean of a random variable and then estimating the mean. It should also be noted that some interpolation schemes such as inverse distance weighting do not provide any measure of precision or accuracy.

In general one should draw a distinction between interpolating spatial data and contouring spatial data although software packages that contour data usually provide an interpolation scheme. Contouring might be best thought of as the algorithm that starts with data on a regular grid and produces the graphical contour plots whereas interpolation (and smoothing) is used when the data are not available on a regular grid. Note that if interpolation is used as a first step in producing a contour plot then the choice of the grid mesh for interpolation is of some importance. Because a software package may use different interpolation schemes different contour plots may result, the uncertainty in the contour plot is a combination of the uncertainties associated with the interpolation step and the uncertainties associated with the contouring step.

DETERMINISTIC METHODS

Kernel approximation

Consider a sample from a random variable, u_1, \dots, u_n . If the cumulative distribution function is given by $F(u)$ then the empirical distribution function is given by:

$$F^*(u) = (1/n) \sum \Delta(u_i) \quad (1)$$

where $\Delta(u) = 1$ if $u \geq u_i$ and is 0 otherwise.

This approximation function is not very smooth as can be seen by examining the corresponding approximation to the density:

$$f^*(u) = (1/n) \sum \delta(u - u_i) \quad (2)$$

where $\delta(u) = 1$ if $u = 0$ and is 0 otherwise. The δ function places a "spike" at each data value. Unfortunately, as interpolators neither of these functions is very good in that F^* simply uses the value at the nearest data point. Both of these can be improved by replacing the δ and Δ functions by functions which smooth out the spikes. For example if:

$$\theta(u) = \exp(- (u/b)^2) \quad (3)$$

then f^* could be replaced by:

$$K \sum a_i \theta(u - u_i) \quad (4)$$

where K is a normalizing constant. Now there is an approximation or interpolation for f between data values. As the constant b increases the θ function approaches the δ function. This idea is easily extended to data in one-dimensional space by a slight modification. Simply let $a_i = f(u_i)$. If in the θ function the variable u is taken to be the magnitude then an isotropic interpolator is obtained that extends to 2, 3 or higher dimensional spaces. The constant b is called the band width of the kernel function and approximately represents the radius of an interval about a point which will include the data points with non-zero weights. By adjusting the band width the interpolator gives greater or lesser weights to data in terms of how close the data locations are to the point where the interpolation is desired. The constant K is simply the sum of the values of the kernel function at the different data locations. As the constant b changes, the interpolator changes from a smoother to a strict interpolator. Kernel estimation is well-known in the statistical literature for fitting probability distributions but is less well-known for spatial interpolation. While the function being fitted or interpolated can be thought as defined in one dimension, the data are not really spatial, i.e., there is no "location" for each data value. There are practical problems if the method is used for spatial data. The Gaussian kernel (eq. 3), is somewhat arbitrary although it has as a limit case the δ function (as b goes to zero). There is little in the literature to suggest how to choose the kernel for a particular data set. The connection between kernel estimation and splines is discussed by Silverman (1984).

Inverse distance weighting (IWD) can be seen to be a special case of kernel estimation. Let $\theta(u) = 1/u^a$ (where $a > 0$) in lieu of eq. (3), then eq. (4) is the IWD interpolator. Usually $a = 1$ and $u - u_i$ is interpreted as the distance between u and u_i . In contrast to the Gaussian kernel above, the inverse distance weighting kernel will have infinite band width. Note that the only real restriction on the kernel is that it be positive valued; hence the choice is not unique. Since the parameter a is not unique it is appropriate to use the data to select a value. This can be done by using a form of cross-validation. Kane et al. (1982) used IWD to interpolate hydrogeochemical data in two different regions. In each region a subset of the data locations was used as a test set, the parameter was optimized by estimating the values at the locations in the test set using the remaining locations. This is a form of cross-validation. It was found that the sensitivity of the results to the parameter value varied considerably, depending on the geochemical variable and on the region. While IWD is relatively easy to apply, once the value of a has been selected, it has at least two disadvantages. First, the weights depend only on the distances between the data locations and the particular location to be estimated, i.e., the relative locations of the data locations to themselves is not incorporated into the interpolation scheme. Second, IWD is not exact, i.e., because the location to be

estimated must be different from all of the data locations. Moreover when the location to be estimated is close to a data location the weight at that location will dominate completely. There does not seem to be a simple way to convert the IWD interpolator into a functional form nor to estimate linear functionals such as spatial averages or derivatives.

Splines

The thin plate spline f^* can be derived by imposing two conditions:

$$\sum_{i=1,n} [f(x_i) - f^*(x_i)]^2 = 0 \tag{5a}$$

and:

$$\int_R [Df^*(x)]^2 dx \text{ is minimized} \tag{5b}$$

where D is a second-order differential operator and R is the region of interest. Equation (5a) corresponds to exactness and eq. (5b) to a smoothness condition. In one dimension this variational problem is easily solved since D is just a second derivative and R is an interval. It is somewhat more complicated in higher dimensions and in particular one must choose the degree of anisotropy, i.e., the relative weighting of the various partial derivatives in D . The smoothing spline is obtained by allowing a compromise between smoothness and exactness, that is:

$$\sum_{i=1,n} [f(x_i) - f^*(x_i)]^2 + \lambda \int_R [Df^*(x)]^2 dx \text{ is minimized} \tag{5c}$$

The constant λ is called the smoothing parameter and is related to the variance of the error term. Thin plate and smoothing splines are implemented in a number of commercial software packages. The connections between splines and kriging was noted as early as 1970 by Kimmeldorf and Wahba, and it was made even more explicit by Matheron (1981). It has also been discussed by Dubrule (1983) and the use of splines is discussed by Hutchinson and Gesler (1994).

Trend surfaces

For a point x_i in k -space, let the coordinates be denoted by x_{ij} , $j = 1, \dots, k$. A monomial of degree q in the coordinates is of the form $(x_{i1})^{a1} (x_{i2})^{a2} \dots (x_{ik})^{ak}$, where the exponents are non-negative integers whose sum is q . The trend surface interpolator f^* is a linear combination of such monomials:

$$f^*(x) = \sum \sum \dots \sum b_j (x_{i1})^{a1j} (x_{i2})^{a2j} \dots (x_{ik})^{akj} \tag{6}$$

where the coefficients are obtained by least squares. It is both an interpolator and a smoother. In one dimension it is simply the best fitting polynomial. More generally the x_{ij} 's could simply be independent variables and eq. (6) is simply the regression equation. This is implemented in standard statistical packages and if f is multivariate normal with mean f^* then statistical inference can be applied to the coefficients. Applied to spatial data, the trend surface interpolator provides regional information but not local information. In part this is because the interpolator is too smooth and generally not exact (unless a very high degree polynomial is used). It has been shown by Marcotte and David (1988) that the trend surface interpolator is the same as universal kriging (except at the data points) when a pure nugget variogram is used.

STOCHASTIC METHODS

Models

When the function is unknown and there are no state equations sufficient even to place the function in a particular class of functions (such as polynomials or functions with continuous second derivatives) it is necessary to choose a class in a different manner. Let $h_0(x), \dots, h_p(x)$ be known linearly independent functions (defined on k -dimensional space). For example, a common choice would be to let these functions be monomials in the coordinates of x . Then the unknown function is considered to be of the form:

$$f(x) = Y(x) + d_0 h_0(x) + \dots + a_p h_p(x) \quad (7)$$

Moreover the data may be of the form:

$$\hat{f}(x) = f(x) + \epsilon(x) \quad (8)$$

where $\epsilon(x)$ represents error or noise. The objectives include removal of the noise term from the data and estimation or prediction of $f(x_0)$. One way to choose a class of functions is for f to be considered as a "value", i.e. a realization, of a random function $Z(x)$. In addition $Y(x)$ is assumed to have mean zero and to have a known spatial structure function. The mean of the unknown random function $Z(x)$ is then represented by the linear combination of the known linearly independent functions. In the simplest case only h_0 is present and is a constant function. The simplest form of an estimator would be:

$$f^*(x) = \sum_{i=1, n} \lambda_i(x) f(x_i) \quad (9)$$

The model given by eqs. (7) or (8), or both, is not complete without specifying the spatial structure function. This is a rather different representation

than given for the other interpolators, the interpolating function is expressed as a sum of functions but those functions are not explicitly known nor specified a priori. In normal practice it is common to write the coefficients in eq. (9) without indicating the dependence on x although it is clear that they are dependent. This reflects an emphasis on estimating or predicting the value of $f(x_0)$ in lieu of estimating the function f . The question then is how to (uniquely) determine the coefficients in eq. (9)? This should be done in a way that allows the incorporation of the information contained in the data, in particular the degree of spatial correlation. It is accomplished by a combination of assumptions pertaining to the random function $Z(x)$ and imposing certain conditions on the interpolator. These are given in statistical form, the most important ones pertain to the spatial structure function.

Spatial structure functions

The covariance $C_x(h)$ and the variogram $\gamma_x(h)$ are two common forms of spatial structure functions. They are defined as follows:

$$C_x(h) = \text{Cov}[Z(x+h), Z(x)] \quad (10)$$

$$\gamma_x(h) = 0.5 \text{Var}[Z(x+h) - Z(x)] \quad (11)$$

These are characteristics of the random function $Z(x)$ and not characteristics of the data values. Both of these quantify the interdependence of pairs of random variables, $Z(x+h)$ and $Z(x)$. The basic underlying intuitive premise is that values at locations that are close together are more similar and values at locations far apart are relatively independent. In order to infer the covariance or variogram from data it is necessary to assume that they do not depend on x . These assumptions will be made more explicit. If $Z(x)$ is second-order stationary then the covariance will exist and not depend on x (in that case the subscript is unnecessary). The variogram can exist under slightly weaker conditions, commonly known as the *intrinsic hypothesis*. If $Z(x)$ is second-order stationary then $\gamma(h) = C(0) - C(h)$. Although the covariance is always bounded the variogram need not be. The most difficult problem is the inference of the covariance or variogram, in part this is because the data is not replicative, i.e., when using the model above the data is a non-random sample from one realization of the random function and most standard statistical inference methods are not applicable (not even when additional strong assumptions such as multivariate normality are made). In addition all the known standard models satisfy those conditions, hence it is for practical reasons rather than for theoretical reasons that one of the usual stationarity conditions is imposed. Positive definiteness is much more important however since it ensures that the coefficient matrix in the kriging system is invertible, i.e., that the solution is unique. One of the criticisms of kriging is that the estimation

and modeling of the spatial structure function is at least partially subjective, although automatic methods such as weighted least squares and maximum likelihood have been used to estimate the parameters of the models.

The dual form

Recall that the system of equations for the coefficients in the kriging estimator can be written in the form:

$$\begin{bmatrix} \mathbf{G} & \mathbf{H} \\ \mathbf{H}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{F} \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_0 \\ \mathbf{H}_0 \end{bmatrix} \quad (12)$$

where \mathbf{F} is the vector of coefficients in eq. (8) and \mathbf{M} is the vector of Lagrange multipliers (all of which depend on x). The interpolator (8) can be re-written in the form:

$$f^*(x) = \sum b_i g(x-x_i) + \sum a_j h_j(x) \quad (13)$$

where the coefficients are obtained from a system with the same coefficient matrix as in eq. (12) but the right-hand side will be replaced by a vector of the data values and a vector of zeros. That is:

$$\begin{bmatrix} \mathbf{G} & \mathbf{H} \\ \mathbf{H}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{B} \\ \mathbf{A} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_0 \\ \mathbf{0} \end{bmatrix} \quad (14)$$

If the spatial structure function is a variogram which only satisfies a weak positive definiteness condition the submatrix \mathbf{G} might not be invertible, but the coefficient matrix will be invertible. That this is a sufficient condition is shown in Myers (1988a) The form given by eq. (14) appears elsewhere in the literature and in particular includes the thin plate and smoothing splines as special cases (Cressie, 1989a, b).

Radial basis functions

Hardy (1971) proposed an estimator of the form given by eq. (13) (but without the sum of the h 's) using a special choice of the g function known as the multiquadratic. If that simplified form of interpolator is required to be exact then the equations in (14) become:

$$\mathbf{GB} = \mathbf{F}_0 \quad (15)$$

and positive definiteness of the g function is sufficient to ensure a unique solution for this system. Other functions include Euclidean distance (which would correspond to a linear variogram) but it was noted that the coefficient matrix was sometimes not invertible. It was also noted that by adding the

additional terms in eq. (13) and using eq. (14) instead of eq. (15) that invertibility was assured. In the numerical analysis literature, interpolators given by eq. (13) are known as radial basis function interpolators (Powell, 1990). The function g is a radial basis function (called radial because only isotropic models were used). As seen above, however, there is a duality between such estimators and universal kriging. By using the stochastic formulation it is easy to see why the additional terms in (12) are needed, they represent the unbiasedness conditions.

Bayesian models

It is well known that the (unique) minimum variance unbiased estimator of $Z(x)$ given the data $Z(x_1), \dots, Z(x_n)$ is the conditional expectation. Moreover if Z is assumed to have a multi-Gaussian probability distribution then the conditional expectation is linear in the data and coincides with the simple kriging estimator. The same results can be obtained in a slightly different manner by assuming a multigaussian distribution but deriving the a posteriori distribution given a prior on the mean and covariance. If the prior is non-informative then the mean of the a posteriori distribution is the simple kriging estimator and the variance of that distribution is the kriging variance. Note that the kriging estimator is only the minimum variance linear estimator.

Non-linear extensions

Several less than desirable properties of the kriging estimator were noted very early by various authors. These include an over-smoothing and non-applicability for certain problems such as the estimation of probabilities. At least two extensions have been proposed, disjunctive kriging (Matheron, 1976) and indicator kriging (Journel, 1985). The disjunctive kriging estimator can be thought of as a generalization of (8) as follows. Consider the term $\lambda_i(x)f(x_i)$, this is a special form of a function given by:

$$A_i(x, f(x_i)) \quad (16)$$

The disjunctive kriging interpolator would be given by:

$$f_{DK}^*(x) = \sum A_i(x, f(x_i)) \quad (17)$$

Then the functions $A_i, i=1, \dots, n$ must be estimated or modeled. In disjunctive kriging this is done by assuming a bivariate distribution for Z that is factorable in a certain sense and which allows the representation of the A 's in terms of orthogonal polynomials (whose weight function is the univariate distribution). For example, if a bivariate Gaussian is used then the orthogonal functions are the Hermite polynomials. The problem reduces to solving multiple simple kriging systems to obtain the coefficients in the Hermite expansions. Note that the disjunctive kriging estimator is in a sense the same kind as the

ordinary, simple and universal estimators, i.e., it directly produces an estimate of the value of the unknown function.

Indicator kriging proceeds in a different manner. A probability distribution is associated with each location. The estimate of the value at a location is obtained by producing an estimate of the corresponding probability distribution. At data locations the probability distribution is like a Heaviside function, that is, a pure jump. Indicator kriging was introduced by Journel (1985) as an alternative to disjunctive kriging. It does invoke strong assumptions, however, namely strong stationarity (as contrasted with second order stationarity or the intrinsic hypothesis for other kriging estimators).

LINEAR FUNCTIONAL ESTIMATION

There are at least three linear functionals (of a function) that are of interest. The point value is one, the integral over a prescribed area or volume is another and the value of the derivative (or a partial derivative) is another. One of the advantages of the kriging estimators is that the formulation leads naturally to a slight modification allowing the estimation of any one of these functionals. Only the right-hand side of eq. (12) needs to be changed. In contrast if interpolators or estimators of the form given by eq. (13) are used then only the functional applied to the interpolating function is obtained. However, by using the duality between the two different formulations it can be shown that this is equivalent to estimation of the functional applied to the unknown function.

MULTIVARIATE EXTENSIONS

If the function f is replaced by a vector function then one must ask in what sense the vector function is being interpolated and what kind of data is available. Both ordinary and disjunctive kriging have obvious extensions as shown by Myers (1988a) wherein the form of the interpolator remains the same, the kriging equations have the same form and the univariate case is subsumed under the vector case. Modeling the spatial structure function is much more complicated. The positive definiteness is defined in a completely analogous fashion and is sufficient to ensure the invertibility of the coefficient matrix.

SUMMARY

Interpolation from spatial data can be performed in various ways by invoking different assumptions and models. In some instances equivalent interpolators are obtained. None of these approaches, however, avoids the problem that is inherent, namely that if only data are available then the problem is ill-posed and there is no unique interpolator.

REFERENCES

- Bastin, G. and Gevers, M., 1985. Identification and optimal estimation of random fields from scattered point-wise data. *J. Autom. Control*, 151: 139–155.
- Boufassa, A. and Armstrong, M., 1989. Comparison between different kriging estimators. *Math Geol.*, 21: 331–346.
- Burgess, T.M. and Webster, R., 1980a. Optimal interpolation and isarithmic mapping of soil properties. I. The semivariogram and punctual kriging. *J. Soil Sci.*, 31: 315–331.
- Burgess, T.M. and Webster, R., 1980b. Optimal interpolation and isarithmic mapping of soil properties. II. Block kriging. *J. Soil Sci.*, 31: 333–341.
- Burgess, T.M. and Webster, R., 1980. Optimal interpolation and isarithmic mapping of soil properties. III. Changing drift and universal kriging. *J. Soil Sci.*, 31: 505–524.
- Cressie, N., 1989a. *Geostatistics*. *Am. Stat.*, 43: 197–202.
- Cressie, N., 1989b. The many faces of spatial prediction. In: M. Armstrong (Editor), *Geostatistics*, 1. Kluwer, Dordrecht, pp. 163–176.
- Cressie, N., 1990. The origins of kriging. *Math. Geol.*, 22: 237–252.
- Christakos, G., 1990. A Bayesian/maximum entropy view to the spatial estimation problem. *Math. Geol.*, 22: 763–778.
- Delfiner, P., 1976. Linear estimation of non stationary spatial phenomena. In: M. Guarascio, M. David and C. Huijbrechts (Editors), *Advanced Geostatistics in the Mining Industry*. Reidel, Dordrecht, pp. 221–236.
- Dubrule, O., 1983. Two methods with different objectives: splines and kriging. *Math. Geol.*, 15: 245–258.
- Foley, T.A., 1987. Interpolation and approximation of 3-D and 4-D scattered data. *Comput. Math. Applic.*, 13: 711–740.
- Goldberger, A.S., 1962. Best linear unbiased prediction in the generalized linear regression model. *J. Am. Stat. Assoc.*, 57: 369–375.
- Hardy, R.H., 1971. Multiquadratic equations of topography and other irregular surfaces. *J. Geophys. Res.*, 76: 1905–1915.
- Hutchinson, M.F. and Gessler, P.E., 1994. Splines — more than just a smooth interpolator. *Geoderma*, 62: 45–67, this issue.
- Journel, A.G., 1985. The deterministic side of geostatistics. *Math. Geol.*, 17: 1–16.
- Journel, A.G., 1986. Constrained interpolation and qualitative information — the soft kriging approach. *Math. Geol.*, 18: 269–286.
- Kane, V., Begovich, C., Butz, T. and Myers, D.E., 1982. Interpretation of regional geochemistry. *Comput. Geosci.*, 8: 117–136.
- Kimmeldorf, G. and Wahba, G., 1970. A correspondence between Bayesian estimation of stochastic processes and smoothing by splines. *Ann. Math. Stat.*, 41: 495–502.
- Marcotte, D. and David, M., 1988. Trend surface analysis as a special case of IRF-k kriging. *Math. Geol.*, 20: 821–824.
- Matheron, G., 1973. The intrinsic random functions and their applications. *Adv. Appl. Probability*, 5: 439–468.
- Matheron, G., 1976. A simple substitute for conditional expectation: the disjunctive kriging. In: M. Guarascio, M. David and C. Huijbrechts (Editors), *Advanced Geostatistics in the Mining Industry*. Reidel, Dordrecht, pp. 221–236.
- Matheron, G., 1981. Splines et krigeage: leur equivalence formelle. *Int. Rep. No. N-695*, Centre de Geostatistique, Fontainebleau, 36 pp.
- Myers, D.E., 1988a. Interpolation with positive definite functions. *Sci. Terre*, 28: 251–265.
- Myers, D.E., 1988b. Multivariate geostatistics for environmental monitoring. *Sci. Terre*, 27: 411–427.

- Myers, D.E., 1989. To be or not to be... stationary: that is the question. *Math. Geol.*, 21: 347–362.
- Myers, D.E., 1991a. Multivariate, multidimensional smoothing. In: A. Posolo (Editor), *Spatial Statistics and Imaging. Proc. AMS–IMS–SIAM Joint Summer Research Conf.*, June 18–24, 1988, Bowden College, Maine. *Lecture Notes — Monograph Series. Inst. of Mathematical Statistics*, Hayward, CA, pp. 275–285.
- Myers, D.E., 1991b. Interpolation and estimation with spatially located data. *Chemometrics and Intelligent Laboratory Systems*, 11: 209–228.
- Myers, D.E., 1992. Kriging, cokriging, radial basis functions and the role of positive definiteness. *Comput. Math. Applic.*, 24: 139–148.
- O’Hagan, A., 1978. Curve fitting and optimal design for prediction. *J.R. Stat. Soc. B*, 40: 1–42.
- Powell, M.J.D., 1990. The theory of radial basis function approximation in 1990. University of Cambridge, *Numerical Analysis Reports, Department of Applied Mathematics and Theoretical Physics*, 102 pp.
- Silverman, B.W., 1984. Spline smoothing: the equivalent variable kernel method. *Ann. Stat.*, 12: 898–916.
- Thiebaux, H.J., 1990. Spatial objective analysis. In: *Encyclopedia of Physical Science and Technology, 1990 Yearbook*. Academic Press, New York, pp. 535–540.
- Weber, D. and Englund, E., 1992. Evaluation and comparison of spatial interpolators. *Math. Geol.*, 24: 381–392.