**Smooth Multidimensional Interpolation**
*by Waldo R. Tobler and Susan Kennedy*

Our primary interest is in geographical problems and the discussion focuses on examples in which the interpolation estimates are to be made in two dimensions. *We* believe that the simplest and most sensible method of geographic interpolation consists of the assignment of an average value to the location or locations for which data are required. The set over which the average is taken is obviously important, and, as weighted averages are almost invariably used, the choice of weights is also critical. For spatial variables the relevant set usually consists of values in the vicinity of the locations for which the estimates are desired. Observe here that we implicitly assume that the variable of interest is numerical, and not categorical, so that averages have meaning. Suggestions as to how to proceed when this is not the case may be found in Guptill (1975), Switzer (1975), and Tobler (1979a). We also restrict our attention to arithmetical averages, ignoring geometrical and harmonic averages and medians which may be appropriate in some cases. It should be recognized that no interpolation scheme can overcome the problem of insufficient resolution in the original observations.

We consciously avoid explicit distance-weighted averages as being computationally too cumbersome, but recognize that they are common in the literature. A rather thorough treatment of this subject is that of Gandin (1963), which includes coverage of covariance and varigram estimation approaches more recently popularized as Kriging, optimal interpolation, objective analysis, collocation, and regionalized variable techniques. Additional literature is referenced in Akima (1975), Barnhill and Nielson (1983), Besag (1974), Brady and Horn (1982), Brodlie (1980), Duchon (1975), Franke (1982), Grimson (1982), Harder and Desmaris (1971), Hardy (1971), Hessing et al. (1972), Journel and Huijbregts (1978), Kraus and Mikhail (1972), Lawson (1978), Matheron (1971), Moritz (1970), Ripley (1981), Schumaker (1976), Swain (1976), Tobler (1979c), and Wahba and Wendelberger (1980), to give only a short selection. It is here assumed that the observations are without error so that filtering of the values is not included; see the foregoing references if this is of interest.

We present three simple cases in which spatial averages can be used for interpolation. The first case involves pixels, or data on a regular mesh; in the second and third cases the known data are irregularly arranged on the plane either as resels or as point locations.

Consider first data given as square pixels (picture elements) with the value for one interior pixel missing (Fig. 1). Then (using an obvious row-column notation) the value at the missing *i, j* location is estimated as an average from its neighbors by

$$Z^*_{ij} = \frac{1}{4} ( Z_{i+1j} + Z_{i-1j} + Z_{ij-1} + Z_{ij+1} )$$

This works equally well when several interior values are missing, as shown in Figure 2, by an

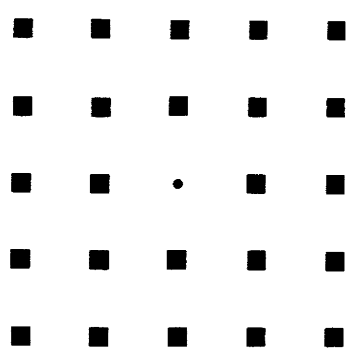

Fig. 1. Small Boxes Denote Known Values; Small Dot Indicates Location for Which an Estimate Is Desired

iteration equivalent to solving Laplace's equation by finite difference methods (Birkhoff 1972). How the missing values are initialized for the iterations is not critical but a good guess saves computational effort. In order to terminate the iterations one invokes the usual stopping rules. This, of course, is just the classical Dirichlet problem in two dimensions and the interpolated value has the harmonic property (Courant and Hilbert 1937) by the construction method. Now it is well known (Kantorovitch and Krylov 1958) that Laplace's equation arises from the least squares problem:

$$\text{min:} \iint_R \frac{\partial Z^2}{\partial X} + \frac{\partial Z^2}{\partial Y}\, dx\, dy$$

with, in the present instance, Dirichlet boundary conditions. Thus the interpolation is spatially smooth, the squared variation of the derivatives, which is minimized, providing a measure of roughness.

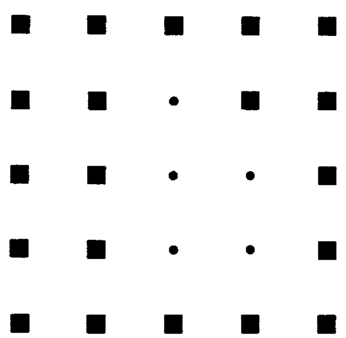

FIG. 2. Small Boxes Denote Known Values; Small Dots Indicate Locations for Which Estimates Are Desired

The foregoing simple solution has several disadvantages. One of these is that we have provided only a point estimate, without any statement of the standard error of the estimate. An obvious way around this is to sample from a distribution having the mean of the neighbors as its expectation with a variance also estimated from these neighbors. A second shortcoming of the harmonic interpolation is that the estimated value can never rise above, nor fall below, its neighbors in magnitude. This restriction can be overcome by enlarging the neighborhood and by requiring that the partial derivatives of the estimate be smooth, that is, by solving the biharmonic equation. In finite difference form this leads to

$$\begin{aligned}\hat{Z}_{ij} = \tfrac{1}{20}\big[&8(Z_{i-1j} + Z_{i-1j} + Z_{ij+1} + Z_{ij-1})\\ &- 2(Z_{i-1j-1} + Z_{i+1j-1} + Z_{i-1j+1} + Z_{i-1j-1})\\ &- (Z_{i-2j} + Z_{ij-2} + Z_{ij+2} + Z_{i+2j})\big],\end{aligned}$$

and iterative procedures are again used when several adjacent values are missing.

Now suppose that the data are given in the form of irregularly arranged resels(resolution elements): census tracts or counties in the United States, with one or more values missing. A generalization of the above results, using first-order neighbors, can be written as

$$\hat{Z}_i = \sum_{j=1}^{n} L_{ij} Z_j,$$

where $n$ is the number of neighbors of region $i$ and $L_{ij}$ are normalized neighbor weights. First-order neighbors are areas having direct contact along borders of nonzero length, second-order neighbors are the first-order neighbors of the initial neighbors, and so on. As an example, Figure 3 shows first- and second-order neighbors for Kansas, with the numerical values given in Table 1. For the population density of Kansas, using only first-order neighbors, and with normalized boundary lengths as weights, we obtain 36.05 persons per square kilometer, whereas the observed value is 27.50. Taking each individual state in turn yields an average success rate of 72 percent, which may be considered impressive in light of the simplicity of the technique (Fig. 4). The method has been extended to the case in which several interior values are estimated (Kennedy and Tobler 1983). Table 2

TABLE 1
First-Order Density Estimate for Kansas (length of border of Kansas with neighboring states, and their population densities)

| Neighbor | Km Border | Density |
|---|---|---|
| Colorado | 338 | 21.3 |
| Oklahoma | 667 | 37.2 |
| Missouri | 433 | 67.8 |
| Nebraska | 572 | 19.4 |

NOTES: Computational procedure—sum of border lengths = 2,010. Sum of border × population densities = 72,466. 72,466/2,010 = 36.05, which yields the density estimate for Kansas.

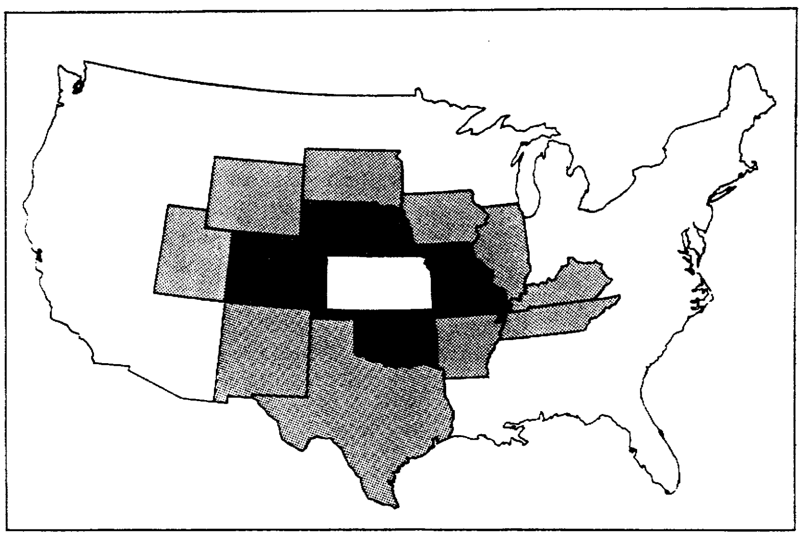

FIG. 3. First- and Second-Order Neighbors of the State of Kansas

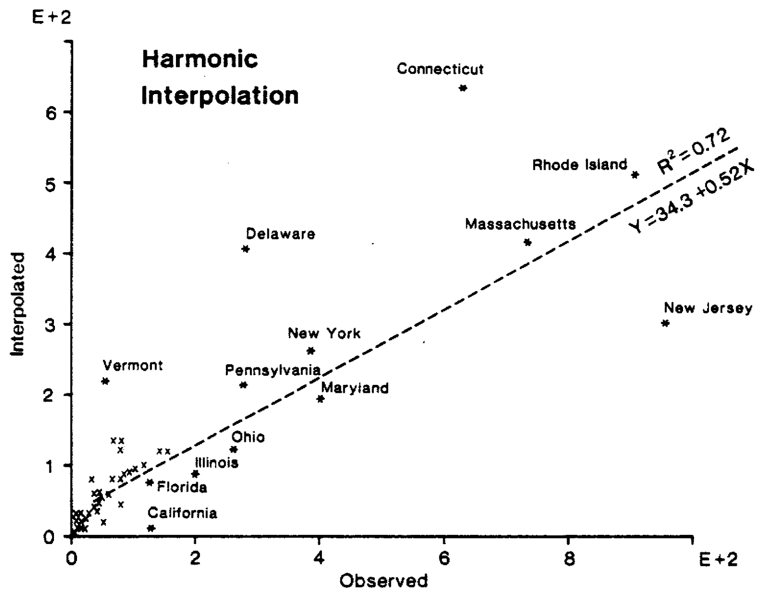


FIG. 4. Scatter Diagram Comparing Actual and Estimated Population Densities for States. Data from Kennedy and Tobler (1983, Table 2)

illustrates the comparable biharmonic density estimate for Kansas. We believe this method of adjacency weighting to be far superior to the use of arbitrary points ("centroids") to represent geographic areas.

TABLE 2
Second-Order Density Estimate for Kansas

| Length of Border of | Km Border | Density |
|---|---|---|
| *Nebraska with* | | |
| South Dakota | 641 | 8.8 |
| Wyoming | 222 | 3.4 |
| Iowa | 192 | 50.5 |
| *Colorado with* | | |
| Wyoming | 419 | 3.4 |
| Utah | 444 | 12.9 |
| New Mexico | 542 | 8.4 |
| *Oklahoma with* | | |
| New Mexico | 58 | 8.4 |
| Texas | 1,534 | 42.7 |
| Arkansas | 319 | 37.0 |
| *Missouri with* | | |
| Arkansas | 548 | 37.0 |
| Tennessee | 156 | 94.9 |
| Kentucky | 111 | 81.2 |
| Illinois | 613 | 199.4 |
| Iowa | 378 | 50.5 |
| Total | 6,177 | |

NOTES: Computational procedure—density estimate from second-order neighbors = 291.[illegible]3/6,177 = 47.11. Density estimates for Kansas = density estimate from first-order neighbors plus difference of first-order estimate and second-order estimate = 36.05 + (36.05 − 47.11) = 24.99 persons per square kilometer.

As a final example, consider the problem of interpolating a continuous scalar field from irregularly arranged point observations in two dimensions. As the first step, to reduce extrapolation, we rotate to principle axes. Thus observations that, for example, fit within an oblique rectangle are readily accommodated. We next pass one coordinate line through each observation (Fig. 5). The result is an irregular orthogonal mesh, with observations at N of the nodes and up to N * N-N nodes at which we need to make an estimate. The obvious procedure is to let the mesh define the adjacencies and then to use neighbor averaging as before. In this example we solve Laplace's equation by using

$$\hat{Z}_{ij} = W_1 Z_{i+1j} + W_2 Z_{i-1j} + W_3 Z_{ij-1} + W_4 Z_{ij+1}$$

with weights chosen from simple geometric considerations. These weights are essentially normalized inverse distances but only to immediately adjacent locations on this mesh. The grid is orthogonal so that only 2(N - 1) distances (instead of (N - 1)/2) are required, and they can be computed in advance for the entire mesh. With more neighbors, different weights, and additional boundary conditions, the method is easily extendable to the biharmonic case to obtain an interpolation with smooth derivatives. An iteration is used since most of the mesh

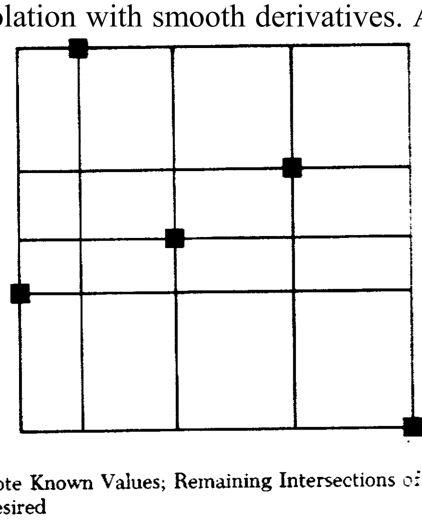

FIG. 5. Small Boxes Denote Known Values; Remaining Intersections of the Mesh Indicate Locations for Which Estimates Are Desired

points do not have observations at the adjacent mesh positions. Points that are neighbors on the mesh may not

be spatially nearest points, but the influence of all points is felt by each point, through the coupling via the mesh. The iterations start from an initial guess and end when an error tolerance is satisfied. Convergence accelerating techniques are available to speed the iterations (Graham 1983). The result is a set of smoothly varying values at the corners of the rectangles defining the mesh, and the original observations are exactly satisfied. Interpolation within the rectangles is then easily effected using conventional bilinear or splining techniques. The method, of course, bears a resemblance to the "lattice tuning" described earlier by Tobler (1979b) except that the observational values are everywhere retained which was not the case in that procedure. An advantage of the rectangular mesh over a triangulation is that it can be used directly in other computations or for display purposes. Computational experience with several extensive sets of data has reinforced our belief in the efficacy of spatial averaging for interpolation. Any interpolation scheme, of course, requires hypotheses about the phenomena under investigation and cannot be applied uncritically.

The smooth interpolation-by-averaging techniques described here can all be extended rather easily to higher-dimensional cases and to the interpolation of vector or tensor field components. An example application would be for nonparametric "rubber sheeting" in order to fit satellite images to conventional maps. It has also not escaped our notice that the methods may be reversed, in order to parse large data sets. Furthermore, the techniques are "vectorizable" for high-speed computation.

## LITERATURE CITED


Akima, H. (1975), "A Method of Bivariate Interpolation and Smooth Surface fitting for Values given at Irregularly Distributed Points." Washington, D.C.: Office of Telecommunications Report No. 75-70.

Barnhill, R., and C. Nielson, eds. (1983). "Surfaces." *Rocky Mountain Journal of Mathematics,* 14, 1-299.

Besag, J. (1974). "Spatial Interaction and the Statistical Analysis of Lattice Systems." *Journal of the Royal Statistical Society,* B, 36, 192-236.

Birkhoff C., (1972). *The Numerical Solution of Elliptic Equations.* Philadelphia, Pa.: SIAM.

Brady, J., and B. Horn (1982). "Rotationally Symmetric Operators for Surface Interpolation." MIT Artificial Intelligence Lab Memo No. 654.

Brodlie, K. (1980). *Mathematical Methods in Computer Graphics arid Design.* London: Academic Press.

Courant, R., and D. Hilbert (1937). *Methoden der Mathematischen Physik.* Berlin: Springer.

Duchon, J. (1975). "Functions-Spline du Type Plaque Mince en Dimension 2." Grenoble, University Technical Report No.231.

Frank, R. (1982). "Scattered Data Interpolation: Tests of Some Methods." *Mathematics of Computation,* 38. 181-99.

Candin, L. (1963). *Objective Analysis of Meteorological Fields.* Leningrad: CIMIZ (United States Department of Commerce translation, TT 65-50007).

Graham, N. (1983). *A Combined Algorithm for Sample Design and Interpolation.* Discussion Paper No. 5. Santa Barbara, Calif.: Department of Geography, UCSB.

Crimson, W. (1982). “A Computational Theory of Visual Surface Interpolation.” *Phil. Trans. Royal Soc. of London,* B298, 395-427.

Cuptill, 5. (1975). “Spatial Filtering of Nominal Data.” Ph.D. dissertation, University of Michigan. Ann Arbor.

Harder, R., and R. Desmaris (1971). “Interpolation Using Surface Splines.” I. *Aircraft* 9, 189-191; 869-871.

Hardy, R. (1971). “Multiquadratic Equations of Topography and Other Irregular Surfaces.” *Journal of Geophysical Research,* 76, 1905-—1915.

Hessing, R., et al. (1972). “Automatic Contouring using Bicubic Functions,” *Geophysics,* 37, 669-674.

Journel, A., and C. Huijbregts (1978). *Mining Geostatics.* New York: Academic Press.

Kantorovitch, L., and V. Krylov (1958). *Approximate Methods of Higher Analysis.* The Hague: Noordhoff.

Kennedy, S., and W. Tobler (1983). “Geographic Interpolation.” *Geographical Analysis,* 15, 151-156.

Kraus, K., and E. Mikhail (1972). “Linear Least Squares Interpolation.” *Photog,rammetric Engineering,* 40, 1016-1029.

Lawson, C. (1978). “Software for C’ Surface Interpolation.” *Mathematical Software III.* New York: Academic Press.

Matheron, C. (1971). *The Theory of Regionalized Variables and Its Applications.* Fontainbleau: Ecole Nationale.

Morita, H. (1970). “Eine Ailgemeine Theorie der Verarbeitung von Schwermessungen mach Kleinsten Quadraten.” Heft Nr. 67A. Munich: Deutsche Ceodaetische Kommission.

Ripley, B, (1981). *Spatial Statistics.* New York: Wiley.

Partial support was provided by the Fundamental Remote Sensing Research Program in Mathematical Pattern Recognition and Image Analysis of the National Aeronautics and Space Administration.

*Waldo B. Tobler is professor, and Susan Kennedy is a graduate student, Department of Geography, University of California, Santa Barbara, 93106-4060.*