Reply

WILLIAM R. GOODIN

Advanced Technology Group, Dames and Moore, Los Angeles, CA 90024

GREGORY J. MCRAE AND JOHN H. SEINFELD

Environmental Quality Laboratory, California Institute of Technology, Pasadena, CA 91125

19 September 1980

The basic objective of our paper (Goodin *et al.*, 1979) was to identify and test a range of computationally efficient techniques for interpolating sparse data onto a regular mesh. We set out to discuss, in general terms, a wide variety of different interpolation procedures and their suitability to the task of generating wind and concentration fields. It was not our intention to reproduce all the details of each of the cited methods. The procedures considered were weighted interpolation, as characterized by (1), polynomial surface fitting and optimal interpolation.

$$c(p) = \frac{\sum_{k=1}^{n} c_k W_k(r)}{\sum_{k=1}^{n} W_k(r)}. \quad (1)$$

In formula (1), $c(p)$ is the interpolated value of c at the point $p = (x,y)$, c_k, $k = 1, 2, \ldots, n$ the data values at the points p_k, $W_k(r)$ the weighting function and r a suitably chosen distance metric.

Our discussion of Cressman (1959) was necessarily brief and directed primarily at the form of $W(r)$ employed as part of a procedure for generating meteorological fields. The weighting function used by Cressman is given by (2); some other examples of weighting functions are shown in Fig. 1.

$$W(r) = \begin{cases} \dfrac{R^2 - r^2}{R^2 + r^2}, & r < R \\ 0, & r \geqslant R. \end{cases} \quad (2)$$

We included a reference to his work for completeness and because of its extensive use by the National Meteorological Center. We did not describe the procedure in great detail nor use it in our comparisons because of the need for proper "tuning" to achieve good accuracy for each data set. Glahn's analysis of our data indicates how well such a method can perform when adjusted for maximum advantage. We are opposed to approaches that require extensive tuning or *a priori* knowledge for the simple reason that while the empirical constants may generate an accurate result on one data set there is no guarantee that they will produce an equally good outcome on another data set. In the application described by Glahn, a considerable wealth of experience has accumulated about the expected form and characteristics of the meteorological fields. In air pollution we do not have, as yet, such an extensive data base. It was for this reason that we rejected the Cressman method and also the class of procedures called optimum interpolation.

We were rather surprised by Glahn's statement that "probably the best way to judge the quality of the three analyses is to compare them subjectively." While this approach might be appropriate for an experienced investigator it certainly is not an ideal basis on which to objectively discriminate between different interpolation procedures. The comments by Glahn on the choice of statistical tests are well taken. Those criteria were selected because they are simple and well known; other more sophisticated tests are available (Bencala and Seinfeld, 1979) and we make extensive use of them in our work. In addition, we frequently employ graphical displays to

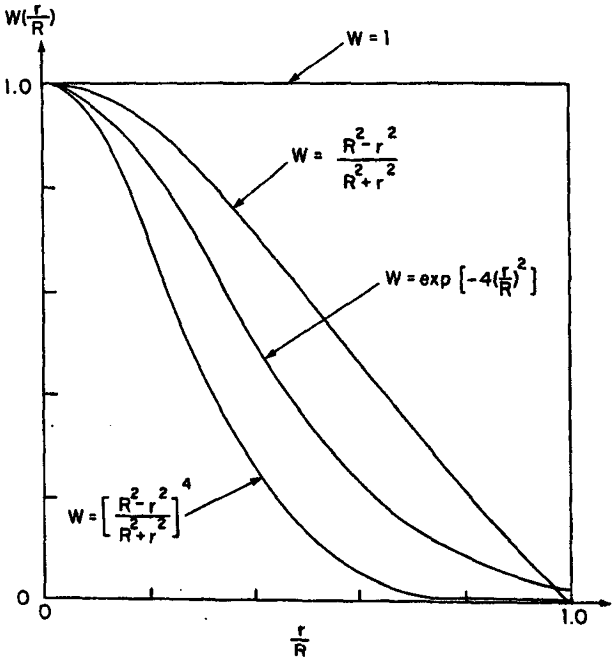

FIG. 1. Some examples of different weighting functions of the form $W(r/R)$.

0021-8952/81/010092-03$04.75

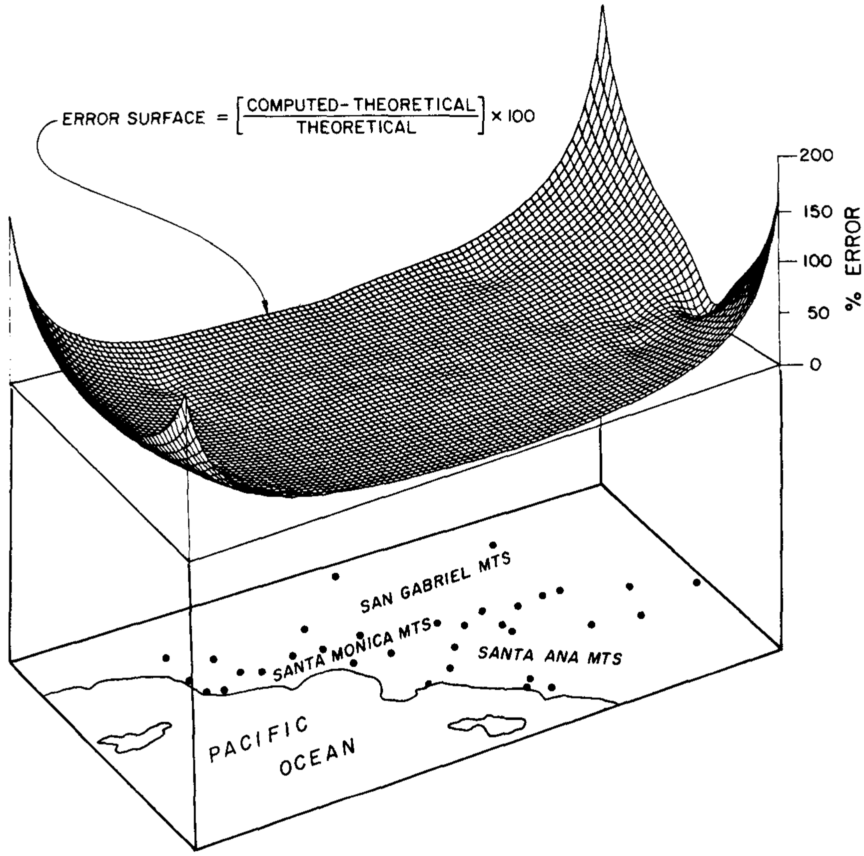

FIG. 2. Error surface for hemispherical interpolation example using a r^{-2} weighting of the data points.

depict the results or errors arising from different interpolation schemes. An example of an error surface for the simple weighting scheme $1/r^2$ is shown in Fig. 2.

It is important to keep in mind that there are two basic sources of error in an interpolation: one arises from the method itself and the other from errors in the data. When field data are used, the "withholding" procedure described by Glahn can be used. Unfortunately, in most practical applications the problem is often too little data rather than too much. One of the major reasons for using analytic functions is to test the algorithm performance independently of the quality of the data used to represent the field measurements. The choice of the hemispherical surface example was motivated by the practical need to accurately describe gradients in regions of relatively sparse data. That surface has its steepest regions at the edges of the grid. There are many other test functions that can be used. Some examples which involve drastically different scales of variation can be found in Franke (1977), Lawson (1977) and McLain (1974).

In his commentary Glahn mentioned the behavior of interpolation methods between data points. One of the advantages of using methods that do not involve empirical constants or tuning is that it is often possible to predict analytically the expected performance and properties of a particular scheme. If the distance metric in (1) is expressed in the form

$$r(p,p_i) = [(x - x_i)^2 + (y - y_i)^2]^{1/2} \qquad (3)$$

and the weighting function as

$$W(r) = r^{-a}, \qquad (4)$$

where a is a constant, then it is possible to express (1) in the equivalent form

$$c(p) = \frac{\sum_{i=1}^{n} c_i \prod_{\substack{j=1 \\ i \neq j}}^{n} r_j^a}{\sum_{i=1}^{n} \prod_{\substack{j=1 \\ i \neq j}}^{n} r_j^a} . \qquad (5)$$

In this rearranged form it is a straightforward task to evaluate the partial derivatives of $c(p)$ in the neighborhood of a data point and to show that for $a > 1$

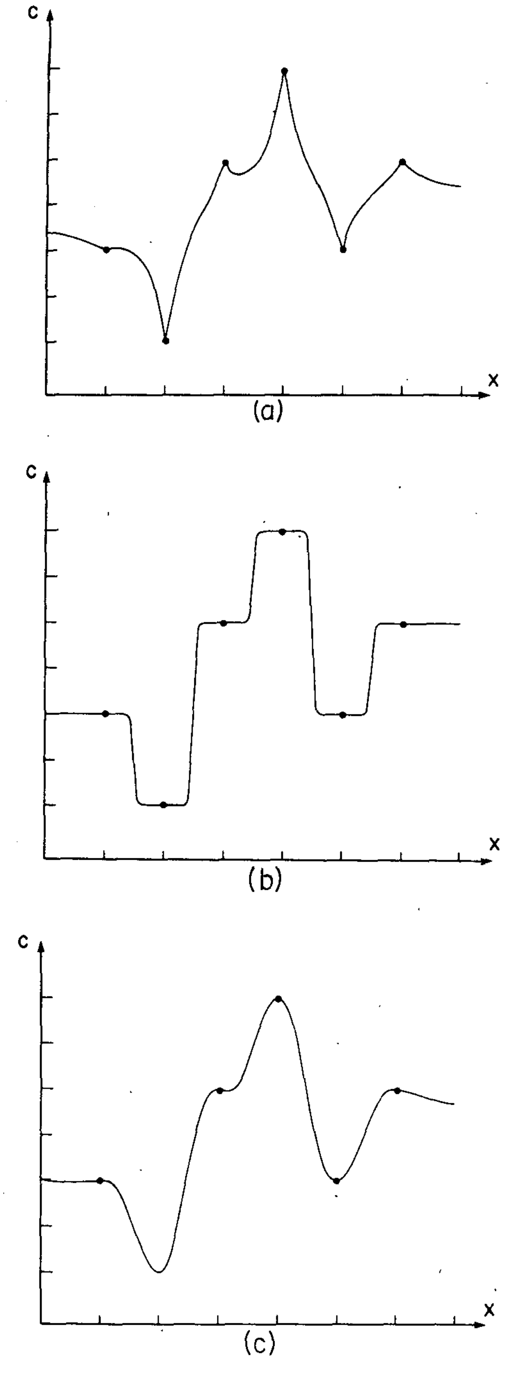

FIG. 3. Results of using different exponents in the weighting function $W(r) = r^{-a}$. The three cases correspond to (a) r^{-1}, (b) r^{-10} and (c) r^{-2} (after Gordon and Wixom, 1977).

$$\lim_{p\to p_i} \frac{\partial c}{\partial x} = \lim_{p\to p_i} \frac{\partial c}{\partial y} = 0, \qquad (6)$$

and for $0 < a \leqslant 1$, that the partial derivatives, in general, do not exist. These results have important practical consequences because it is evident that for $0 < a \leqslant 1$ there will be cusps at the data points and when $a > 1$ the slopes in the vicinity of each c_k will be zero. The effects of these conditions are illustrated in Fig. 3. Gordon and Wixom (1977) and Schumaker (1976) discuss, in considerable detail, the properties of interpolation schemes based on (1). As an example their results indicate that for the case $a = 1$, the interpolated value is bounded by the data, i.e.,

$$\min_i [c_i] \leqslant c(p) \leqslant \max_i [c_i]. \qquad (7)$$

Gordon and Wixom (1977) further show that if $c_i \geqslant 0$ for all $i = 1, 2, \ldots, n$, then $c(p) \geqslant 0$ for all p. This is a very desirable property when interpolating concentration fields. If the data are derived from a constant field, $c_i = c$ for all $i = 1, 2, \ldots, n$ then $c(p) = c$ for all p.

We appreciate the additional citations to spline methods mentioned in the commentary by Glahn. Spline interpolation methods, as a class, are employed in many different fields and an extensive discussion can be found in Schumaker (1976). As a matter of interest Wahba and Wendelberger (1980) have employed splines and cross-validation methods for developing 500 mb pressure fields.

REFERENCES

Bencala, K. E., and J. H. Seinfeld, 1979: An air quality model performance assessment package. *Atmos. Environ.*, **13,** 1181–1185.

Cressman, G. P., 1959: An Operational Objective Analysis System. *Mon. Wea. Rev.*, **87,** 367–374.

Franke, R., 1977: Locally determined smooth interpolation at irregularly spaced points in several variables. *J. Inst. Math. Appl.*, **19,** 471–482.

Goodin, W. R., G. J. McRae and J. H. Seinfeld, 1979: A comparison of interpolation methods for sparse data: Application to wind and concentration fields. *J. Appl. Meteor.*, **18,** 761–771.

Gordon, W. J., and J. A. Wixom, 1977: Shepard's method of "metric interpolation" to bivariate and multivariate interpolation. *Math. Comput.*, **32,** 253–264.

Lawson, C. L., 1977: Software for C^1 surface interpolation. *Mathematical Software III,* Academic Press, 161–193.

McLain, D. H., 1974: Drawing contours from arbitrary data points. *Comput. J.*, **17,** 318–324.

Schumaker, L. L., 1976: Fitting surfaces to scattered data. *Approximation Theory II*, G. G. Lorentz, C. K. Chui and L. L. Schumaker, Eds., Academic Press, 203–268.

Wahba, G., and J. Wendelberger, 1980: Some new mathematical methods for variational analysis using splines and cross-validation. *Mon. Wea. Rev.*, **108,** 1122–1143.