Tellus (1988), *40A*, 173–176

SHORT CONTRIBUTION

Some real data tests of the interpolation accuracy of Bratseth's successive correction method

By R. S. SEAMAN, *Bureau of Meteorology Research Centre, P.O. Box 1289K, Melbourne 3001, Australia*

(Manuscript received 13 April; in final form 13 July 1987)

1. Introduction

The theoretical relationships between statistical interpolation (Eliassen, 1954; Gandin, 1963) and correction methods of objective analysis (Bergthorsson and Döös, 1955; Cressman, 1959, henceforth C), and the convergence properties of the latter methods, have recently been investigated by Franke and Gordon (1983), Bratseth (1986, henceforth B), Lorenc (1986), and Franke (1988). In B, there was described an iterative method, similar in some respects to that of C, but which always converges in the limit to the statistical interpolation (SI) result. The purpose of this note is to present the results of some tests of the univariate interpolation accuracy of the B method, in comparison with SI and C on a large set of real data.

2. Data base and evaluation method

The tests were performed with a data set which originally had been compiled to evaluate the performances of univariate analysis schemes, over and adjacent to data-dense areas (Seaman and Hutchinson, 1985, henceforth SH). Briefly, the data were 20 years of 09.00 (local time) observations of surface pressure from a network of up to 47 stations over southeastern Australia, covering an area of about $8 \cdot 10^5$ km^2 (Fig. 1 of SH). The observations were normalized by subtracting the climatological mean at each station.

Interpolation accuracy was assessed by randomly selecting N observations, to be used in an analysis, and interpolating to the locations of observations withheld from the analysis. Each of the N observations was used in the interpolations to the locations of every withheld datum; no data selection was employed. The process was repeated for networks of N observations on many days, until comparisons had been made with 10,000 pieces of withheld data.

The particular data set was used because it was readily available; in its form described above, it is not ideal for testing the B method. The observations are dense (average spacing about 150 km) in comparison to the characteristic length scale, and the ratio of observational error variance to background (climatology) error variance ("noise-to-signal ratio") is low. These two circumstances do not apply in many practical situations. However, such biases were partially overcome, (i) by using small values of N, and constraining the observational separation, in some of the tests, and (ii) by adding a synthetic random error component to the data in other tests.

For univariate analysis of normalized increments from a background field, and assuming uncorrelated observational errors, the basic B method uses the weights

$$a_{xj} = r_{xj}/M_j, \tag{1}$$

$$a_{ij} = (r_{ij} + \lambda_i^2 \delta_{ij})/M_j, \tag{2}$$

to iteratively interpolate the observed increments ("corrections") at points j, to grid points (x) and observing points (i). Here, r is the spatial correlation coefficient, λ_i^2 is the normalized observational error variance, and δ_{ij} is the Kronecker delta. The denominator M_j is essentially a measure of "data density" at observing point j, defined by:

$$M_j = \sum_k |r_{jk}|, \tag{3}$$

where the summation is over all observing points k.

To apply the basic B method, it is necessary to prespecify exactly the same parameters as are needed for statistical interpolation. Such parameters had already been obtained in SH, from a 5-year subset which was not used in subsequent tests. The observational and background error variances were respectively 0.834^2 and 7.94^2 hPa^2, implying $\lambda = 0.105$. The homogeneous and isotropic correlation function used was one of the two best found in SH, namely

$$r(s) = (1 + s/L)\exp(-s/L), \tag{4}$$

where s is separation and L is a length scale parameter (824 km).

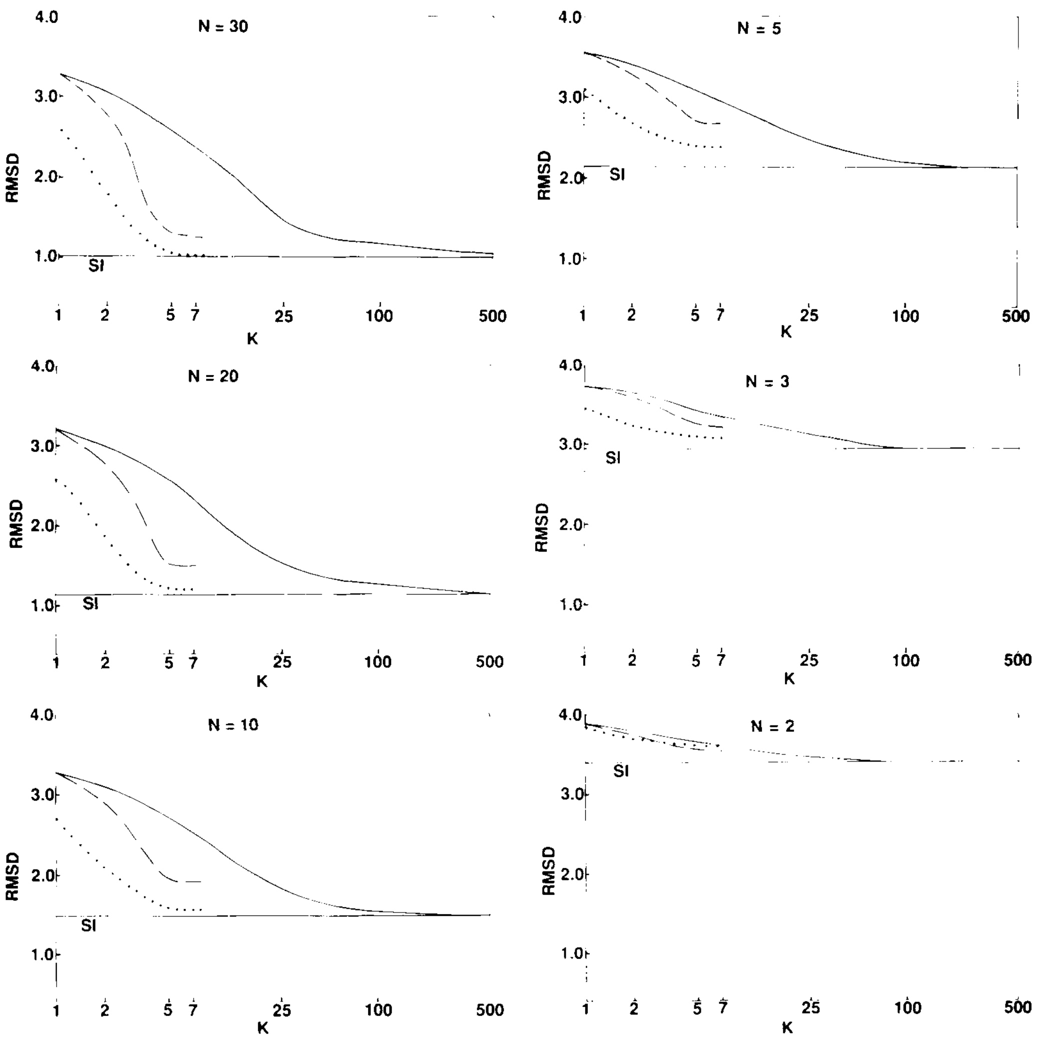

Fig. 1. Root mean square deviations RMSD (hPa) of interpolated minus withheld data for the B method (full curve), as a function of iteration number (k), and the number of observations (N) used to interpolate. Also shown are the corresponding deviations for statistical interpolation (SI), for the C method (dotted curve), and for the B method with reducing length scale (dashed curve). The noise-to-signal parameter λ was 0.105, and the background error variance was 7.94^2 hPa^2.

3. Tests of the basic method

Fig. 1 shows plots of root mean square (RMS) deviations (interpolated minus withheld data) for the basic B method (full curves), as a function of iterations and network density. Also shown are the corresponding RMS deviations for SI, and for C (dotted curves). The panel for $N = 30$ corresponds to an average observational spacing of about $0.2L$, and that for $N = 2$ to about $0.8L$. The parameters of the C method had been optimized for five iterations: see SH for details. The RMS deviations are not strictly interpolation errors, since they are augmented by random observational error. However they provide a valid basis for comparison of the methods. It is stressed that deviations are from withheld data, not from data used in the analyses.

Consistent with the theory in B, the basic B result approaches the SI result with increasing iterations. However it is only for small N, or for large numbers of iterations, that the basic B method interpolation error is close to that of SI or five-iteration C. Note in particular that the C error is much below the B error at the first iteration, for large N. Both however are well below the background error of about 8 hPa.

4. Some variations

4.1. *Tuning the B method*

In its basic form, the B method can be regarded as an alternative computational route to the SI result; for this purpose the values of λ_i and r_{ij} should be as close as possible to the true noise-to-signal ratios and increment correlation coefficients. However it was suggested in section 6 of B that a more rapid convergence, perhaps to a suboptimal limit, could be achieved by decreasing L on later iterations.

The long dashed curves in Fig. 1 were the result of reducing L by a factor β at each iteration. The value of β, typically 0.55–0.8, was optimized so as to minimize the theoretical interpolation error (eq. (2) in B) after 5 iterations. Fig. 1 confirms that such progressive reduction of L does improve the convergence rate, particularly for dense networks, although the scheme no longer converges to the SI result.

A further modest gain is achievable (results not shown) by replacing M_j (eq. (3)) by M_j/α where α is an overrelaxation factor which varies with iteration. Also, following a suggestion by Bratseth (personal communication), the values of both L and λ at each iteration, including the first, were treated as tunable parameters. As with overrelaxation, the small improvements from additional tuning of L and λ did not alter the general picture conveyed by Fig. 1. However the latter tests were not exhaustive, and further efforts may be worthwhile.

4.2. *Larger observational separations*

In view of the apparent sensitivity of the basic B method's convergence rate to data density, the experiments for $N = 2$ were repeated with the observational separation constrained to be greater than 1200 km (about $1.5L$). Fig. 2 indicates a definite improvement of the basic B method relative to the C method, and a rapid convergence to the SI result. Note that for clarity the ordinate scaling in Fig. 2 has been changed relative to Fig. 1, and that for practical purposes the performances of all three methods were still very close (3.57 to 3.64 hPa RMS). Unfortunately the available data base did not permit tests with much larger observational separations, but it seems reasonable to speculate that the trend indicated in Fig. 2 might continue.

4.3. *Higher noise-to-signal ratios*

Idealized experiments shown in B indicated an improved convergence rate when λ was larger. The value of λ is likely to be greater than the 0.105 used so far in this note, when a short-range forecast rather than climatology is used as

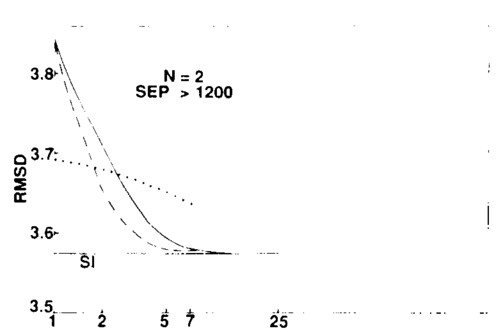

Fig. 2. As for Fig. 1, with $N = 2$, and observational separation constrained to greater than 1200 km. Note the change in ordinate scaling from Fig. 1.

Table 1. *Root mean square deviations (hPa) of interpolated minus withheld data, for networks of N observations and noise-to-signal ratio λ*

	$\lambda = 0.105$		$\lambda = 0.3$		$\lambda = 1.0$	
	$N = 30$	$N = 2$	$N = 30$	$N = 2$	$N = 30$	$N = 2$
SI	1.022	3.402	1.461	3.777	2.685	5.368
BB5	2.555	3.662	2.601	3.868	3.061	5.368
BR5	1.316	3.575	1.606	3.826	2.739	5.368
CR5	1.056	3.614	1.559	3.830	2.987	5.589

SI = statistical interpolation. BB5 = basic B method (5 iterations). BR5 = B method with reducing L (5 iterations). CR5 = C method (5 iterations).

background. To simulate such situations with the available data, a synthetic random error component was purposely added to the observations, to increase their noise-to-signal ratio to 0.3 and 1.0. These values correspond approximately to those calculated by Lonnberg and Hollingsworth (1986) for 1000 and 150 hPa geopotentials. Table 1 indicates that with λ sufficiently large, and the data sufficiently sparse, the B method becomes a little better than the C method, for the same number of iterations. However, slow convergence of the basic B method is still a problem with dense data, even when $\lambda = 1.0$.

5. Concluding remarks

The foregoing results indicate that the comparative performances of the several methods depend strongly both upon data density and upon noise-to-signal ratio. In all cases, SI produced the best result, as would have been expected since the covariance structure of the data was well known. For dense data and low noise-to-signal ratio, the performance of the C method after 5 iterations was close to SI, while the basic B method converged very slowly. On the other hand, with sparse data and noise to signal ratio close to unity, the basic B method converged rapidly to the SI result, and was slightly more accurate than the C method for the same number of iterations.

It should be emphasized that a contributory factor to the C method's performance was its *objective* tuning, by minimization of its theoretical interpolation error with respect to the tunable parameters (Seaman, 1983). Smoothing between iterations was an essential ingredient.

When the data is dense, it also appears essential to tune the B method (via L and perhaps λ). This requirement makes the B and C methods more similar. However, even in a simple univariate setting, the two approaches still differ with respect to (i) normalization of the weights, (ii) interpolation to observing points, and (iii) form of the influence function. Future research might focus on the individual impacts of these 3 factors.

6. Acknowledgement

I am grateful to Dr. Arne Bratseth for readily sharing his knowledge and experience, and in particular for suggesting a greater emphasis on noise-to-signal ratio.

REFERENCES

Bergthorsson, P. and Döös, B. R. 1955. Numerical weather map analysis. *Tellus 7*, 329–340.

Bratseth, A. M. 1986. Statistical interpolation by means of successive correction. *Tellus 38A*, 439–447.

Cressman, G. P. 1959. An operational objective analysis system. *Mon. Wea. Rev. 87*, 367–374.

Eliassen, A. 1954. Provisional report on calculation of spatial covariance and autocorrelation of the pressure field. Inst. Weath. and Climate Res. Acad. Sci. Oslo, Rep. No. 5.

Franke, R. 1988. Statistical interpolation by iteration. *Mon. Wea. Rev.*, in press.

Franke, R. and Gordon, W. S. 1983. The structure of optimum interpolation functions. Tech. Rep. NPS-53-83-005, Naval Postgraduate School, Monterey, CA, 23 pp.

Gandin, L. S. 1963. *Objective analysis of meteorological fields*. Translated (1965) by Israel Programme for Scientific Translations, Jerusalem. 242 pp.

Lönnberg, P. and Hollingsworth, A. 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus 38A*, 137–161.

Lorenc, A. C. 1986. Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc. 112*, 1177–1194.

Seaman, R. S. 1983. Objective analysis accuracies of statistical interpolation and successive correction schemes. *Aust. Meteorol. Mag. 31*, 225–240.

Seaman, R. S. and Hutchinson, M. F. 1985. Comparative real data tests of some objective analysis methods by withholding observations. *Aust. Meteorol. Mag. 33*, 37–46.