Influence of spatial sampling and interpolation on estimates of air temperature change

Scott M. Robeson

Department of Geography, Indiana University, Bloomington, Indiana 47405, USA

ABSTRACT: When observed air temperatures are analyzed spatially, irregularly sampled data are usually interpolated in some fashion. As a result, methods of spatial analysis clearly play a role in determining the size and variability of estimated air temperature changes, both spatially and temporally. Through graphical and statistical analysis, 3 spherically based interpolation methods — inverse-distance weighting, triangulated surface patches, and thin-plate splines — are evaluated and compared using air temperature anomaly data. Analysis of errors resulting from spatial interpolation provides information about the strengths and weaknesses of historical station networks. Analysis of differences between the 3 interpolation methods suggests that similar spatial patterns are produced, but with some regional disparities. Mean absolute differences between interpolation methods can be over 0.4 °C for some sparse station networks and as low as 0.1 °C for dense station networks. When averaged spatially, however, the differences tend to offset one another, producing time series of terrestrial average air temperature anomalies that are largely independent of spatial interpolation method. Using cross validation to analyze spatial interpolation errors suggests that sparse station networks can produce nontrivial interpolation errors. Station networks from the late 1800s produce average interpolation errors of nearly 0.5 °C, errors that are similar in magnitude to spatial standard deviations of air temperature anomalies. Denser station networks, typical of the 1950s and 1960s, produce average interpolation errors as low as 0.2 °C. While these regional interpolation errors do not appear to influence estimates of terrestrial average air temperature, they do raise additional concerns regarding our ability to detect small climatic signals at regional scales.

KEY WORDS: Climatic change · Spatial sampling · Spatial interpolation · Network bias   Air temperature anomalies · Global warming

INTRODUCTION

Changes in near-surface atmospheric temperatures are perhaps the most common and tangible indicator of climatic change. While there still is important scientific debate regarding the magnitude and extent of changes in air temperature anomalies over the last century (e.g. Lindzen 1990, Kellogg 1991), observed air temperature trends (Houghton et al. 1990) as well as scenarios from general circulation models (e.g. Manabe & Wetherald 1986) suggest considerable potential for environmental and societal impacts at a variety of spatial scales. Clearly, it is useful to seek not only substantial evidence of climatic change in the historical record of observed air temperatures, but also to convey our relative confidence in estimates of climatic change.

Close scrutiny of data from historical air temperature records has revealed several substantial sources of error at observation station locations (e.g. Karl et al. 1989). While much research has attempted to quantify observational errors in air temperatures over the past century, several problems remain unresolved. Historical samples of climate, unfortunately, are not the result of an optimal design for the detection of climatic change. Instead, the available space-time sample of climate is a result of complex cultural and political processes that have shaped human society over the last few centuries. The result is a spatial sample that is biased towards industrialized countries and densely populated regions. Climatologists, therefore, often use a 'sample of convenience' (Freedman et al. 1978) when obtaining data.

While reliable thermometry was developed in the mid 18th century, spatial sampling largely was limited to Europe and North America until the mid to late 19th century. After the late 19th century, the limiting factors in obtaining reliable air temperature time series have been data quality and adequate spatial coverage. All studies of long-term air temperature change from the instrumental record are — and will continue to be — limited by these data inadequacies. Irregular and changing station networks have raised concerns that spatial sampling has biased estimates of air temperature change (Jones et al. 1986a, Karl et al. 1989, 1994, Willmott et al. 1991, Robeson 1993). Precisely how spatial sampling problems have influenced estimates of air temperature change remains uncertain.

To estimate spatial biases in time series of air temperature anomalies, methods for spatially analyzing irregularly sampled climate data are evaluated below. Spatial interpolation, in particular, is potentially very important, yet no objective comparisons have been performed in the context of large-scale air temperature variability. The effects of sparse, uneven station networks and spatial interpolation, however, are interrelated and difficult to separate. To some extent, both have influenced the spatial variability of air temperature as well as estimates of air temperature change.

AIR TEMPERATURE VARIABILITY

Although air temperatures can vary over a wide variety of spatial and temporal scales, spatial and temporal variability are reduced through temporal averaging. Most studies of long-term air temperature change have used monthly averages of daily air temperature data (e.g. Jones et al. 1986a, b, Hansen & Lebedeff 1987). Monthly mean air temperature data are the most readily available and are of appropriate timescale, since shorter period fluctuations usually are not directly relevant for climatic change research. When viewed as a filter, the monthly averaging procedure — a rectangular filter applied to daily values — is not an optimal way to remove high-frequency information (Hamming 1977). But, original data in the form of daily maxima and minima are not available for many stations (Karl et al. 1993).

Spatial variability is more problematic since air temperatures are measured at sparse (i.e. sparse relative to the spatial variability of air temperature) and irregularly spaced points. Sparse and irregular air temperature networks can produce samples with poor resolution and, therefore, unrepresentative and aliased spatial patterns. Unlike regular sampling intervals, the minimum resolvable scale is not well defined for an irregularly spaced sample. Therefore, there is no straightforward choice for grid size. Many studies (e.g. Yamamoto & Hoshiai 1979, Jones et al. 1986a) have opted for coarse grids, presumably in an attempt to match data resolution. While spatial aliasing within the original air temperature network cannot be avoided, interpolating from a dense station network to a coarse grid can further alias air temperature signals, distorting space-time patterns (Daley 1991). Whether a coarse or fine grid is used, the original resolution of the data should be kept in mind when further analyses of the gridded fields are performed.

Air temperature anomalies

Latitudinal and topographic influences on air temperature are well understood and, in many ways, can be considered to be independent of recent climatic change. To reduce spatial variability associated with latitudinal and topographic effects, a station mean may be removed from air temperature time series (e.g. Jones et al. 1986a, Hansen & Lebedeff 1987), creating air temperature anomalies:

$$\Delta T_j = T_j - \bar{T}_j \tag{1}$$

where ΔT_j is an air temperature anomaly at station j, T_j might be a monthly or annual mean air temperature, and $\bar{T}_j$ is a long-term average of temperature at the station.

Overall, converting air temperatures to anomalies produces a smoother spatial field while maintaining the original temporal variability at each station. Nearby stations that are located at different elevations, for instance, become more comparable (Fig. 1), making spatial interpolation an easier task. While converting air temperatures to anomalies reduces the effects of latitude and elevation, information related to interstation (i.e. spatial) variability of mean air temperature is removed with the long-term station mean. Actual air temperature space-time series are needed in their unmodified form for input to models and for verification of modeling studies. Methods for interpolating and spatially analyzing actual — as opposed to anomaly — climatic data will not be considered here; however, such methods provide an important alternative to anomaly-based analyses (e.g. Daly et al. 1994).

Air temperature data

There have been many studies of air temperature change from the instrumental record (Vinnikov et al. 1980, Jones et al. 1986a, b, Hansen & Lebedeff 1987). By combining and augmenting several climatic
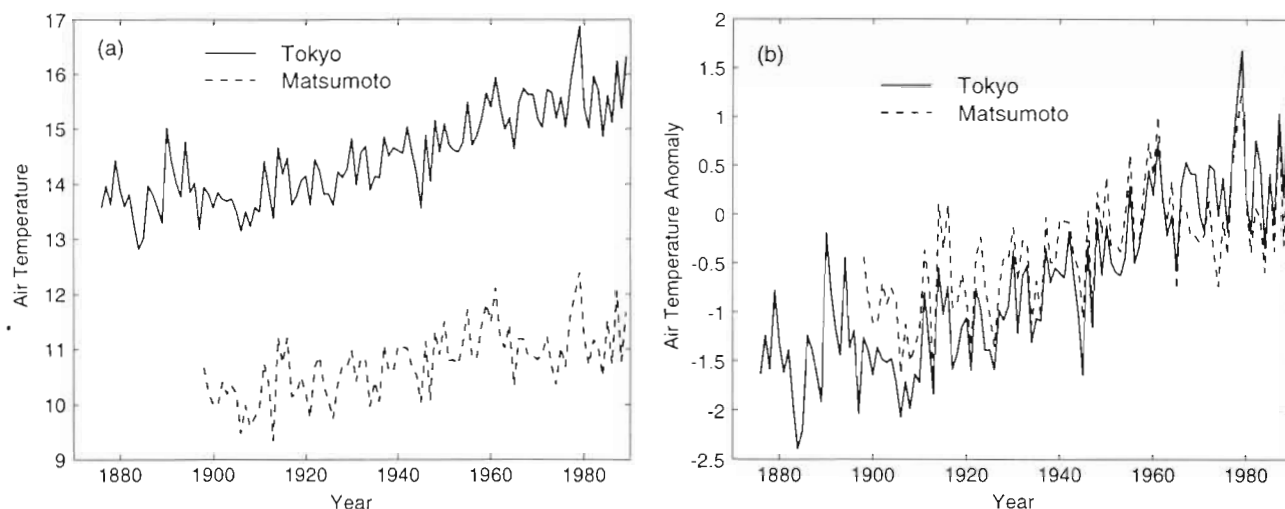
Fig. 1. Air temperature time series (in °C) at Matsumoto, Japan (elevation 611 m) and at Tokyo, Japan (elevation 5 m) for (a) actual air temperatures and (b) air temperature anomalies, illustrating the effects of removing station means. Matsumoto and Tokyo are approximately 150 km apart

archives, Jones et al. (1986a, b), however, improved the spatial coverage of several individual archives, particularly in the early years of the record (i.e. before 1920). Original data from Jones et al. (1986a, b) are available as both original station data and gridded anomaly fields (Jones et al. 1991). These well-documented and frequently updated (e.g. Jones 1988) public domain data are used here to examine the effects of uneven and changing spatial coverage. Gridded data are not of interest since only by analyzing the original spatial distribution of stations can the effects of irregular and sparse sampling be detected. To generate air temperature anomalies at each station, the procedures of Jones et al. (1986a, b) were followed. Since the greatest number of stations were available over the 1951 to 1970 time period, anomalies were created by removing the 1951 to 1970 mean value at each station (as in Jones et al. 1986a, b). To be included in the analysis, a station was required to have 15 or more years of data in the 1951 to 1970 period. Station data have been screened carefully for many types of non-spatial biases and many have been removed (e.g. discontinuities due to changes in station elevation). Others, however, are more difficult to recognize (e.g. urbanization) and still may be pervasive.

Many of the errors that affect air temperature data are well documented, some being highly systematic. Changes or discrepancies in (1) air temperature instrumentation, (2) time of observation, (3) location of measurement stations, (4) methods for constructing time-averages and (5) the local environment can induce large biases (Mitchell 1953, Schaal & Dale 1977, Karl et

al. 1986, Kukla et al. 1986, Edwards 1987, Karl & Jones 1989). Although many errors have been identified at individual air temperature stations, a great deal of uncertainty remains concerning the effects of irregular station distributions and how these distributions have changed through time. Methods for evaluating the spatial representativeness and efficacy of global and regional climate networks, however, are not well developed. As a result, studies of global and regional changes in air temperature anomalies often have not explicitly addressed the problems of uneven spatial sampling.

In their study of Northern Hemisphere air temperature change, Jones et al. (1986a) indirectly evaluated the effects of incomplete spatial coverage by comparing integrated differences between a fixed grid and a time-variable grid. Relative to the low spatial variability of air temperature anomaly data, large differences (over 0.5 °C) appear early in the record, suggesting that station distribution effects may be substantial. Using both land and ocean data, Karl et al. (1994) examined how sampling biases can affect air temperature trends. Utilizing a variety of approaches, they found small errors associated with century-long temperature trends. Decadal trends, however, contained more substantial bias. In addition, the size of errors was found to be dependent on patterns of air temperature (i.e. errors are proportional to spatial variability).

While several analyses have suggested that some level of variability is induced by irregular station networks, the effects of spatial interpolation — both the differences between methods and actual interpolation errors — have not been estimated.

METHODS OF SPATIAL INTERPOLATION

Before comparing several interpolants using air temperature anomaly data, a brief summary of spatial interpolation algorithms is presented. Several extensive reviews of spatial interpolation methods provide additional information (Shumaker 1976, Gustavsson 1981, Franke 1982, Lam 1983, Burrough 1986, Thiébaux & Pedder 1987).

Spatial interpolation is a classical estimation problem that usually employs linear combinations of observed values to determine values at unsampled locations (e.g. at the nodes of a regular grid). Most interpolation procedures, however, treat the spatial dimension as planar. A more geometrically consistent approach is to interpolate on the surface of a sphere. When interpolating over large areas of a planet's surface, planar interpolation methods (i.e. interpolation within a cartographic projection) can produce large errors (Willmott et al. 1985). Not only will planar interpolation necessarily result in interpolation errors, but each cartographic projection produces a different error field from the same data.

Most interpolation methods (spherical or planar) may be classified into 1 of 3 somewhat distinct categories: (1) distance-weighting, (2) tessellation, and (3) functional minimization. Distance-weighting involves applying a weight to each observation in inverse proportion to its distance from a point where an estimate is desired. The distance-weighting function may be derived from spatial correlation patterns (as in kriging and optimal statistical objective analysis) or may be a specified inverse function of distance. Using spatial correlation patterns does make intuitive sense; however, there is evidence that simpler methods perform just as well (Bussières & Hogg 1989, Weber & Englund 1992).

A commonly used spatial analysis tool, tessellations have wide applicability beyond interpolation (Okabe et al. 1992). After performing a spatial decomposition (often a triangulation), a surface is fit over each patch within the decomposition. Tessellation methods have long been used in climatology (e.g. Thiessen 1911). Functional minimization (e.g. least squares) establishes an objective function that requires some degree of fidelity between the observations and a weighted combination of basis functions (e.g. a truncated Fourier series).

While numerous methods of spatial interpolation are available, few spherical interpolants have been developed. A spherical interpolant from each of the above categories, however, is available. Shepard's (1968) inverse distance weighting algorithm was adapted to the sphere by Willmott et al. (1985). Renka (1984) developed a spherical triangulation routine that

applies either a C^0 (continuous and zero times differentiable, or linear) or C^1 (cubic with constraints) surface to each spherical triangle. Wahba (1981) adapted 2-dimensional thin-plate splines to the sphere.

COMPARISON OF AIR TEMPERATURE ANOMALY INTERPOLATIONS

The spherical algorithms mentioned above are compared and evaluated using air temperature anomaly data from Jones et al. (1991). As these 3 methods represent somewhat distinct approaches to spatial interpolation, comparing and contrasting the estimated temperature fields allows some generalizations to be made.

Grid point differences

The spherical spatial interpolation approaches — Renka's (1984) C^0 (hereafter referred to as RC^0) and C^1 (RC^1) methods, Willmott et al.'s (1985) distance-weighting approach (WRP), and Wahba's (1981) thin-plate splines (TPS) — are used to interpolate the air temperature anomalies to a terrestrial grid. As implemented here, the thin-plate spline interpolations provide an exact fit through the data. Thin-plate splines often are used to smooth spatial data (Wahba 1981); however, to provide a consistent comparison between the 3 methods, true interpolation (rather than approximation) will be used.

Time series of statistical differences between the interpolation methods provide a relative indication of how the interpolation procedures perform for each year. Perhaps the most useful measure of overall difference between methods is the (weighted) mean absolute difference (MAD):

$$\text{MAD} = \sum_{i=1}^{n} w'_i |\Delta\hat{T}_{1,i} - \Delta\hat{T}_{2,i}| \Big/ \sum_{i=1}^{n} w_i \qquad (2)$$

where $w_i = \cos\phi_i$ (where ϕ_i is the latitude of grid-point i) and $\Delta\hat{T}_{1,i}$ and $\Delta\hat{T}_{2,i}$ are estimates of air temperature anomalies using 2 different spatial interpolation methods. Weights are used to account for differential areas associated with the spherical lattice of n grid points.

Integrated MADs between WRP and RC^0 over the period 1881 to 1988 range from less than 0.1°C for dense networks to nearly 0.3°C for sparse networks (Fig. 2) with a strong inverse relationship between MAD and number of stations (see Fig. 3 for a time series of number of stations). Time series of MADs for WRP-TPS and RC^0-TPS have somewhat higher values than those between WRP and RC^0 Areally weighted MAD for WRP-TPS varies between 0.2 and 0.4°C until
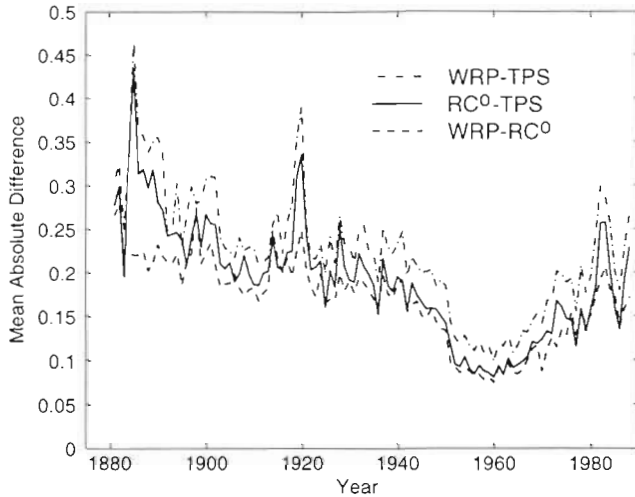
Fig. 2. Time series of weighted mean absolute difference (in °C) between terrestrial air temperature anomalies interpolated by the WRP and TPS, RC⁰ and TPS, and WRP and RC⁰ methods

Air temperature anomaly time series

Perhaps the most common and tangible indicators of air temperature change are time series of global or terrestrial average air temperature anomalies ($\overline{\Delta T}$). To develop an estimate of $\overline{\Delta T}$, gridded (interpolated) air temperature anomalies are integrated numerically. Although some regional discrepancies between interpolation procedures were shown above, integrated time series of $\overline{\Delta T}$ generated by each of the interpolation methods are virtually identical (and therefore are not shown). Mean absolute differences between the terrestrially averaged time series generated by WRP, RC⁰, and TPS are less than 0.03 °C. Regional differences between the interpolation methods appear to cancel one another when gridded values are averaged. With respect to the method of spatial interpolation, then, time series of terrestrial average air temperature appear to be very robust.

Spatial interpolation errors

Differences between interpolation methods have been presented; however, differences do not assess the relative accuracy of the interpolation methods for air temperature anomalies. One useful tool to evaluate any type of estimator is cross validation (Efron & Gong 1983). Cross validation, as implemented here, simply entails (1) removing 1 observation from the analysis, (2) attempting to estimate the removed value, and (3) repeating steps (1) and (2) for every observation. In the context of spatially interpolating air temperature anomalies, the process involves removing one station's anomaly ΔT_j from the analysis and using anomalies from surrounding stations to interpolate to the removed location:

$$\Delta \hat{T}_j = f(\Delta T_1, \Delta T_2, ..., \Delta T_k, ..., \Delta T_n) \quad j \neq k \quad (3)$$

where f represents a particular spatial interpolation method. By comparing observed and interpolated values at each station, a direct measure of interpolation error is obtained, allowing the relative merits of each procedure to be assessed.

Integrated errors

Cross validation error estimates at station locations were averaged to produce time series of mean absolute error (MAE) from 1881 to 1988 for both WRP and RC⁰ (Fig. 3). Cross validation errors for TPS are available only for several sparse networks due to computational demands. RC⁰ and TPS appear to have a consistently higher MAE than WRP. When 95% bootstrapped confi-
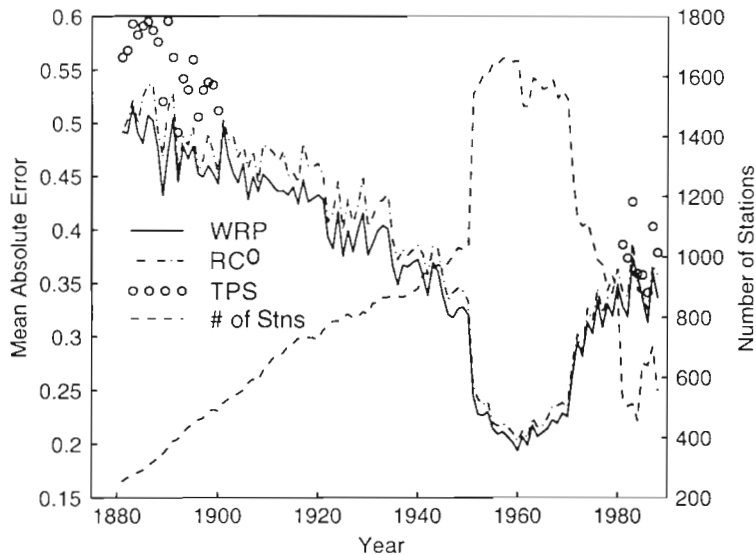
the large increase of stations in the 1951 to 1970 period reduces MAD to nearly 0.1 °C. The RC⁰-TPS time series are similar in pattern, but slightly smaller than the WRP-TPS values.

Much of the differences in the WRP-TPS and RC⁰-TPS time series can be attributed to WRP and RC⁰ being local interpolants (i.e. they use a subset of the data to interpolate to a given location) while TPS is a global method (i.e. TPS uses all of the data to fit an interpolating surface). Since TPS minimizes the 'bending energy' of an infinitely thin plate that extends over the entire planet, areas with large spatial gradients can cause undulations in the interpolated surface. The undulations do not appear when TPS smooths the data by acting as a low-pass filter (e.g. Wahba 1981). When TPS is used as a smoothing procedure, values of MAD (differenced from both WRP and RC⁰) are reduced to less than 0.2 °C for all years. In general, then, it appears that TPS should be used in a smoothing mode for data that contain large spatial gradients.

Clearly, choice of spatial interpolation method is less important for denser station networks. Without estimating differences or interpolation errors, however, it is difficult to determine what constitutes a 'dense' station network since a dense network for one variable will not necessarily be sufficient for other variables. A spatially smooth variable such as 500 mb geopotential height, for instance, would not require the sampling density needed to resolve air temperature variability. Precipitation, on the other hand, is more spatially variable than either pressure or air temperature and would likely be underrepresented by even the densest air temperature network (e.g. Willmott et al. 1991, 1994).

Fig. 3. Time series of terrestrial mean absolute error (in °C) from cross validation analysis using the WRP, RC^0, and TPS methods with air temperature anomaly data. Dashed line shows the number of stations for each year (right axis)

dence intervals are calculated, however, there often is no statistically significant difference between the cross validation errors of WRP and RC^0.

The C^1 method of Renka (1984) produced much larger cross validation errors (>1°C for all networks) than any other method. A modification to the C^1 method that applies an arbitrary amount of tension to each triangular surface (Renka pers. comm.) also was evaluated. While variable tension did improve the C^1 results, interpolation errors were still much larger than those of the C^0 method. The highly irregular station distributions appear to cause problems with the gradient estimation algorithms used within the C^1 method.

Both the magnitude of the errors and the nearly perfect inverse relationship between MAE and number of stations (correlation coefficient of 0.94) are particularly striking in the cross validation time series (Fig. 3). These 2 related features have important implications for the sparse networks of 1881 to 1950 as well as the 1980s and beyond. Only during the 1951 to 1970 period is the station network sufficiently dense to reduce interpolation error to a moderately low level (0.2°C). Prior to 1940, average interpolation errors are above 0.4°C and often are closer to 0.5°C.

While errors of 0.4 to 0.5°C are small in an absolute sense, anomaly air temperature data exhibit low spatial variability, particularly when compared to actual air temperature data. Spatial standard deviations for anomaly data (calculated over the entire terrestrial surface) are on the order of 0.6°C. Relative to data

variability, therefore, spatial interpolation of air temperature anomalies can produce large errors. Large relative errors also are evident in scatterplots of observed and cross validation-predicted air temperature anomalies for both sparse and dense networks (Fig. 4).

It should be noted that the MAEs discussed above are indicative of the absolute magnitude of *regional* interpolation errors. To examine how these regional biases affect the overall terrestrial mean, a statistic such as mean bias error (MBE = $\frac{1}{n}\sum_{i=1}^{n}[\Delta T_j - \Delta \hat{T}_j]$, or simply the difference between the mean of the observed anomalies minus the mean of the cross-validation-predicted anomalies) is useful. MBE for both WRP and RC^0 was less than 0.05°C for all years. These small MBEs corroborate the small integrated differences between interpolation methods discussed above. Time series of terrestrial *average* air temperature anomalies, therefore, appear to be unbiased with respect to spatial interpolation, despite large regional errors.

Limitations of cross validation

Cross validation has been used to compare several different interpolation methods and to evaluate network bias. Cross validation, however, is not without problems and care must be taken when interpreting results. In particular, whenever irregularly spaced data are being used, uneven spatial weighting can occur. Statistical error measures derived from cross validation (e.g. MAE) are biased towards areas that contain data. That is, error estimates are available only at each station. Gridding cross validation errors might reduce spatial bias, but an additional interpolation with its associated errors would be required. Also, interpretation of error 'fields' may not be straightforward since error is not necessarily a continuous variable. It also is not clear to what extent errors in the observed data have affected estimates of interpolation error (Burt 1985, Moore & Rowland 1990).

Interpolation errors from cross validation, nonetheless, do provide a useful measure of overall network efficacy. Conclusions regarding the 'best' interpolation procedure, however, should be limited to the particular data and domain being considered. For this reason, cross validation often is viewed as more of an exploratory tool rather than a confirmatory method (Davis 1987).
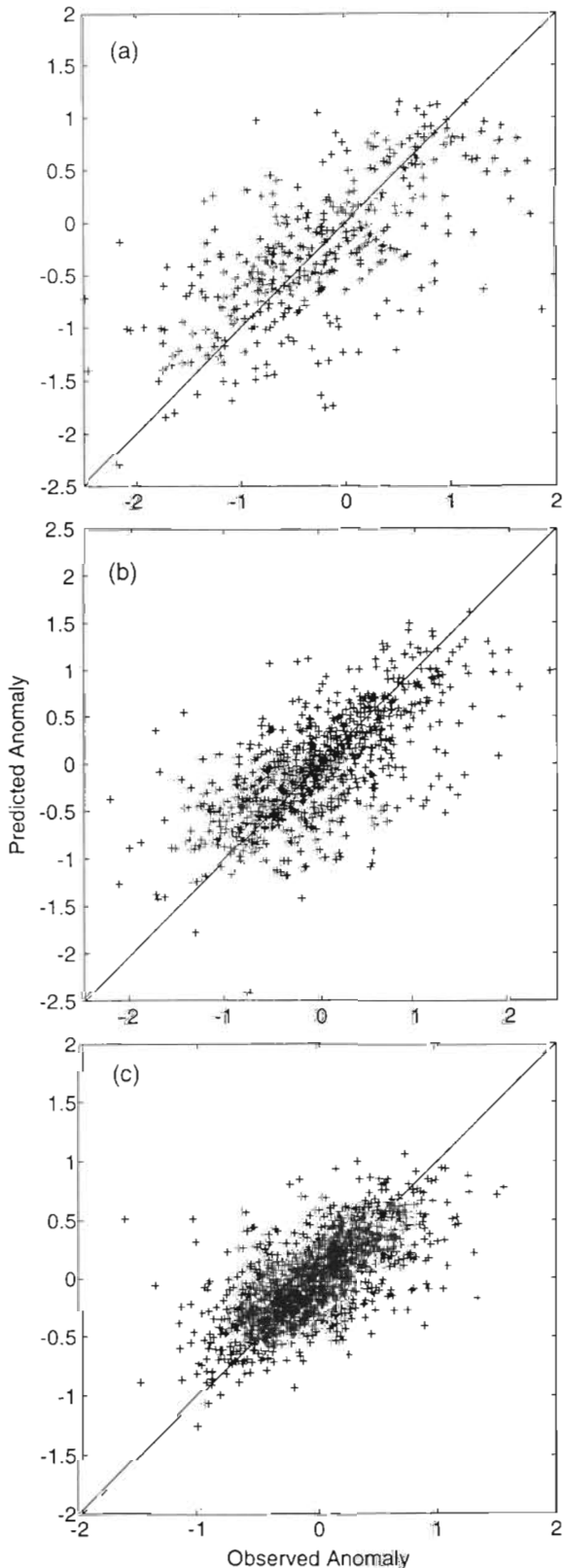
Fig. 4. Scatterplots of observed air temperature anomalies (in °C) versus cross validation-predicted anomalies using the WRP method for (a) 1890, (b) 1930, and (c) 1970

SUMMARY AND CONCLUSIONS

Since most climatic data are irregularly sampled, observations are interpolated to a spherical grid for visualization and averaging. Spatial interpolation, therefore, is a potential source of error in the analysis of air temperature data. Three spherical methods — triangulated surface patches (Renka 1984, C^0 method), inverse distance-weighting (Willmott et al. 1985), and thin-plate splines (Wahba 1981) — were used to interpolate air temperature anomalies. For sparse station networks, regional differences between interpolation methods were sometimes large (>0.5 °C). When integrated, however, the methods usually behaved in a similar fashion, particularly for denser station networks.

To examine interpolation errors, cross validation — removing one point at a time and estimating the removed value from surrounding values — was used. Cross validation errors for air temperature anomalies were similar for the 3 interpolation methods. The magnitudes of the errors, however, were large relative to air temperature anomaly variability: average errors for sparse networks were nearly 0.5 °C while denser networks produced average errors just over 0.2 °C. Since spatial standard deviations for air temperature anomalies are approximately 0.6 °C, these errors are considered large in a relative sense.

Spatially integrated time series of air temperature anomalies appear to be very robust with respect to differences between spatial interpolation methods. However, when examined at regional scales, variability induced by interpolation error and sparse station networks can, in some cases, be relatively large. Compounded with other problems (e.g. urban warming) that inject uncertainty into the historical climatic record, errors associated with spatial interpolation and inadequate spatial sampling raise additional concerns regarding our ability to detect regional climatic signals from the instrumental air temperature record.

LITERATURE CITED

Burrough, P. A. (1986). Principles of geographical information systems for land resources assessment. Oxford University Press, Oxford

Burt, J. E. (1985). Interpolation errors and spatial resolution of the United States solar radiation network. Phys. Geog. 6: 230–246

Bussières, N., Hogg, W. (1989). The objective analysis of daily rainfall by distance weighting schemes on a mesoscale grid. Atmos.-Ocean 27: 521–541

Daley, R. (1991). Atmospheric data analysis. Cambridge University Press, Cambridge

Daly, C., Nielson, R. P., Philips, D. L. (1994). A statistical-topographic model for mapping climatological precipitation over mountainous terrain. J. appl. Meteorol. 33: 140–158

Davis, B. M. (1987). Uses and abuses of cross-validation in geostatistics. Math. Geol. 19: 241–248

Edwards, H. B. (1987). Sampling theory applied to measurement and analysis of temperature for climate studies. J. Clim. appl. Meteorol. 26: 731–736

Efron, B., Gong, G.(1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. Am. Statistician 37: 36–48

Franke, R. (1982). Scattered data interpolation: tests of some methods. Math. Computation 38: 181–200

Freedman, D., Pisani, R., Purves, R. (1978). Statistics. W. W. Norton & Co., New York

Gustavsson, N. (1981). A review of methods for objective analysis. In: Bengtsson, L., Ghil, M., Källen, E. (eds.) Dynamic meteorology: data assimilation methods. Springer-Verlag, New York, p. 17–36

Hamming, R. W. (1977). Digital filters. Prentice-Hall, Englewood Cliffs, NJ

Hansen, J., Lebedeff, S. (1987). Global trends of measured surface air temperature. J. geophys. Res. 92(D11): 345–413

Houghton, J. T., Jenkins, G. J., Ephraums, J. J. (eds.) (1990). Climate change: the IPCC scientific assessment. Cambridge University Press, Cambridge

Jones, P. D. (1988). Hemispheric surface air temperature variations: recent trends and an update to 1987. J. Clim. 1: 654–660

Jones, P. D., Raper, S. C. B., Bradley, R. S., Diaz, H. F., Kelly, P. M., Wigley, T. M. L. (1986a). Northern Hemisphere surface air temperature variations: 1851–1984. J. Clim. appl. Meteorol. 25: 161–179

Jones, P., Raper, S., Cherry, S., Goodess, C., Wigley, T., Santer, B., Kelly, P. (1991). An updated global grid point surface air temperature anomaly data set, 1851–1990. CDIAC, Oak Ridge National Laboratory, Oak Ridge, TN

Jones, P. D., Raper, S. C. B., Wigley, T M. L. (1986b). Southern Hemisphere surface air temperature variations: 1851–1984. J. Clim. appl. Meteorol. 25: 1213–1230

Karl, T. R., Jones, P. D. (1989). Urban bias in area-averaged surface air temperature trends. Bull. Am. Meteorol. Soc. 70: 265–270

Karl, T. R., Jones, P. D., Knight, R. W., Kukla, G., Plummer, N., Razuvayev, V., Gallo, K. P., Lindseay, J., Charlson, R. J., Peterson, T C., et al. (1993). Asymmetric trends of daily maximum and minimum temperature. Bull. Am. Meteorol. Soc. 74: 1007–1023

Karl, T R., Knight, R. W., Christy, J. R. (1994). Global and hemispheric temperature trends: Uncertainties related to inadequate spatial sampling. J. Clim. (in press)

Karl, T. R., Tarpley, J. D., Quayle, R. G., Diaz, H. F., Robinson, D. A., Bradley, R. S. (1989). The recent climatic record: what it can and cannot tell us. Rev. Geophys. 27: 405–430

Karl, T. R., Williams, C. N. Jr, Young, P. J., Wendland, W. M.

(1986). A model to estimate time of observation bias associated with monthly mean maximum, minimum, and mean temperatures for the United States. J. Clim. appl. Meteorol. 25: 145–160

Kellogg, W. W. (1991). Response to skeptics of global warming. Bull. Am. Meteorol. Soc. 72: 499–512

Kukla, G., Gavin, J., Karl, T R. (1986). Urban warming. J. Clim. appl. Meteorol. 25: 1265–1270

Lam, N. S.-N. (1983). Spatial interpolation methods: a review. Am. Cartog. 10: 129–149

Lindzen, R. S. (1990). Some coolness concerning global warming. Bull. Amer. Meteorol. Soc. 71: 288–299

Manabe, S., Wetherald, R. T. (1986). Reduction in summer soil wetness induced by an increase in atmospheric carbon dioxide. Science 232: 626–628

Mitchell, J. M. Jr (1953). On the causes of instrumentally observed temperature trends. J. Meteorol. 10: 244–261

Moore, R. D., Rowland, J. D. (1990). Evaluation of model performance when the observed data are subject to error. Phys. Geog. 11: 379–392

Okabe, A., Boots, B., Sugihara, K. (1992). Spatial tesselations. Wiley, New York

Renka, R. J. (1984). Interpolation of data on the surface of a sphere. ACM Trans. Math. Software 10: 417–436

Robeson, S. (1993). Spatial interpolation, network bias, and terrestrial air temperature variability. Publ. Climatol. 46: 1–51

Schaal, L. A., Dale, R. F. (1977). Time of observation bias and 'climate change' J. appl. Meteorol. 16: 215–222

Shepard, D. (1968). A two-dimensional interpolation function for irregularly spaced data. Proc. 23rd Natl. Conf. ACM. Association for Computing Machinery, New York, p. 517–523

Shumaker, L. L. (1976) Fitting surfaces to scattered data. In: Lorentz, G. G., et al. (eds.) Approximation, Vol. II. Academic Press, New York, p. 203–268

Thiébaux, H. J., Pedder, M. A. (1987). Spatial objective analysis. Academic Press, New York

Thiessen, A. (1911). Precipitation averages for large areas, Mon. Weather Rev. 39: 1082–1084

Vinnikov, K. Y., Gruza, G. V., Zakharov, V. F., Kirillov, A. A., Kovyneva, N. P., Ran'kova, E. Y (1980). Contemporary variations of the Northern Hemisphere climate. Sov. Meteorol. Hydrol. 6: 5–17

Wahba, G. (1981). Spline interpolation and smoothing on the sphere. SIAM J. Sci. Stat. Comput. 2: 5–16

Weber, D., Englund, E. (1992). Evaluation and comparison of spatial interpolators. Math. Geol. 24: 381–391

Willmott, C. J., Rowe, C. M., Philpot, W. D. (1985). Small-scale climate maps: a sensitivity analysis of some common assumptions associated with grid-point interpolation and contourng. Am. Cartog. 12: 5–16

Willmott, C. J., Robeson, S. M., Feddema, J. J. (1991). Influence of spatially variable instrument networks on climatic averages. Geophys. Res. Lett. 18: 2249–2251

Willmott, C. J., Robeson, S. M., Feddema, J. J. (1994). Estimating continental and terrestrial precipitation averages from rain-gauge networks. Int. J. Climatol. (in press)

Yamamoto, R., Hoshiai, M. (1979) Recent change of the Northern Hemisphere mean surface air temperature estimated by optimum interpolation. Mon. Weather Rev. 107: 1239–1244