Smart Interpolation of Annually Averaged Air Temperature in the United States

CORT J. WILLMOTT AND KENJI MATSUURA

Department of Geography, Center for Climatic Research, University of Delaware, Newark, Delaware

(Manuscript received 6 December 1994, in final form 5 June 1995)

ABSTRACT

Two "smart" interpolation procedures are presented and assessed with respect to their ability to estimate annual-average air temperatures at unsampled points in space from available station averages. Smart approaches examined here improve upon commonly used procedures in that they incorporate spatially high-resolution digital elevation information, an average environmental lapse rate, and/or another higher-resolution longer-term average temperature field. Two other straightforward or commonly used interpolation methods also are presented and evaluated as benchmarks to which the smart interpolators can be compared. Interpolation from a spatially high-resolution, long-term-average air temperature climatology serves as a first approximation, while "traditional" interpolation (from a single realization of annual average air temperature on a single station network) is the other benchmark. Traditional interpolation continues to be the most commonly used interpolation approach within many of the atmospheric and environmental sciences.

Smart approaches are significantly more accurate than either traditional methods or estimates spatially interpolated from a high-resolution climatology alone. A smart interpolation method that makes combined use of a digital elevation model (DEM) and traditional interpolation was nearly 24% more accurate than traditional interpolation by itself. Average error associated with this DEM-assisted interpolation algorithm, for interpolating yearly average air temperatures in the United States, was 0.44°C. The other smart method that was evaluated combines DEM information with a high-resolution average air temperature field. It was even more accurate, as expressed in an overall average interpolation error of only 0.38°C per year, which makes it some 34% more accurate than traditional interpolation. It is likely that the performance of smart interpolation, relative to traditional interpolation, will be even better when used with relatively sparse station networks.

1. Introduction

Weather station networks with long-term records and good spatial coverage are uncommon (Willmott et al. 1994). Two problems arise from this paucity of well-conditioned observational networks. Estimating a time-averaged weather or climate variable (e.g., annual-average air temperature) at unsampled locations by spatial interpolation is relatively unreliable, and in turn, areal averages made from the network observations can be biased. It is necessary, therefore, to develop spatial estimation methods that are more reliable than traditional interpolation methods. Our interest here then is to present and test spatial interpolation algorithms that can be used to estimate time-averaged (e.g., annual average) point air temperatures from relatively low resolution station networks. More specifically, we consider straightforward approaches that make use of additional information, information that is spatially correlated with the interpolate (annually averaged air temperature, in this instance). Simple meteorological (lapse rate) theory also informs our interpolation procedure. Related multivariate interpolation methods, such as Daly et al.'s (1994) somewhat more complicated approach to interpolating precipitation, also are beginning to appear in the meteorological and climatological literature. Guided both by theory and by spatial correlations, such approaches are termed "smart" interpolators.

2. Spatial estimation of annually averaged air temperature

a. Traditional interpolation

Traditional interpolation methods include applications of spatial regression, thin-plate splines, kriging, and inverse-distance weighting. While each performs somewhat differently, all estimate average air temperature with accuracies on the same order (Robeson 1994; Ishida and Kawashima 1993). At its most generic level, traditional interpolation can be expressed as

$$\hat{T}_k = I[\mathbf{T}_{n(k)}, \lambda, \phi], \quad (1)$$

where $\hat{T}_k$ is the interpolated annual-average air temperature at any location k, $\mathbf{T}_{n(k)}$ is a vector that contains n_k [$n_k = n(k)$] neighboring annual-average station observations—each of which influences $\hat{T}_k$, λ and ϕ are the corresponding n_k-element longitude and latitude

Corresponding author address: Dr. Cort J. Willmott, Department of Geography, Center for Climatic Research, University of Delaware, Newark, DE 19716-2541.

vectors, and $I(\mathbf{x}, \lambda, \phi)$ is the interpolation function. While a number of traditional interpolation functions would adequately represent $I(\mathbf{x}, \lambda, \phi)$, a previously presented inverse-distance algorithm is used here (Shepard 1968; Willmott et al. 1985a). It can be written

$$\hat{T}_k = \frac{\sum_{i=1}^{n_k} w_{ik}(T_i + \Delta T_i)}{\sum_{i=1}^{n_k} w_{ik}}, \quad (2)$$

where i is a station location, k is the location of the estimate, n_k is the number of nearby stations that influence the estimate at location k, T_i is an annual average air temperature at station i, ΔT_i is a gradient estimate at station i, and $\hat{T}_k$ is the spatially interpolated temperature at location k. Inverse-distance weighting means that each w_{ik} is some function of the inverse distance between station i and location k.

Based on spherical rather than planar geometry, the Willmott et al. (1985a) traditional interpolation algorithm [used as $I(\mathbf{x}, \lambda, \phi)$ herein] weights (w_{ik}) the station temperatures primarily based on powers (≤ 2) of inverse distance. The weights also account for spatial autocorrelation among clusters of the n_k nearby stations, using an interstation cosine weighting function. A simple (local) extrapolator additionally allows $\hat{T}_k$ to exceed the range within $\mathbf{T}_{n(k)}$, when estimated spatial gradients (ΔT_i) at the n_k nearby stations warrant it. The number of locally influential stations (n_k) is, on the average, 7—although this number can be as low as 4 (when the station network is sparse locally) or as high as 10 (when the network is dense locally). None of the parameters (e.g., n_k or the powers of inverse distance) are optimized (although they can be) in order to maintain a high-level of computational efficiency. Even without optimization, as Willmott and Robeson (1995), Bussières and Hogg (1989), Weber and Englund (1992), and others have shown, this and related interpolators are surprisingly accurate. For a fuller description of the algorithm, the reader is referred to Willmott et al. (1985a).

b. *Topographically informed interpolation*

One way to enhance the traditional interpolation of annual-average air temperature is to exploit the relationship between air temperature and elevation. Recently available, spatially high-resolution digital elevation models (DEMs) make this feasible (National Geophysical Data Center 1993). Elevational effects on air temperature interpolation have been implicitly (statistically) included by several authors (e.g., Ishida and Kawashima 1993). Our implementation attempts to include elevational influences in a more physically meaningful, yet straightforward, way.

Station elevations first are estimated from a DEM (discussed in section 3c). Estimated, rather than actual, station elevations are used for consistency and because actual station elevations are unavailable for many of the world's historical station records. An annual-average sea level air temperature at station i then can be estimated from

$$_sT_i = T_i + \Gamma z_i, \quad (3)$$

where Γ is an average environmental lapse rate ($\Gamma \approx 6.5 \times 10^{-3}$ °C m^{-1}), z_i is the elevation at station i, and $_sT_i$ is the estimated annual-average air temperature at sea level for station i. Annual-average sea level air temperature at grid point j then becomes

$$_s\hat{T}_j = I[_s\mathbf{T}_{n(j)}, \lambda, \phi], \quad (4)$$

where $_s\mathbf{T}_{n(j)}$ is a vector of n_j [$n_j = n(j)$] nearby annual-average sea level air temperatures that influence $_s\hat{T}_j$, and grid point j replaces generic location k. Finally, air temperature at grid point j is estimated from the DEM according to

$$\hat{T}_j = {}_s\hat{T}_j - \Gamma z_j, \quad (5)$$

where z_j is a DEM-estimated elevation at grid point j.

c. *Climatologically aided interpolation (CAI)*

Another approach is to make use of a second air temperature field, one observed over a different time period but on a much higher spatial resolution station network (Willmott and Robeson 1995). Termed climatologically aided interpolation (CAI), this approach begins by interpolating climatologically averaged station air temperatures observed on the higher-resolution station network to the locations of the stations in the lower-resolution network. Interest is in estimating the annually averaged air temperature field observed on the lower-resolution network, but at unsampled locations. Our primary purpose in presenting CAI here is not to evaluate it [as Willmott and Robeson (1995) have done]; rather, our goal merely is to describe CAI, as CAI is a component of the "topographically and climatologically informed" algorithm discussed below (section 2d).

A difference between the observed annually averaged temperature of interest (on the lower-resolution station network) and an interpolated higher-resolution climatological average temperature (at a lower-resolution station) then is obtained from

$$\delta T_i = T_i - \hat{\bar{T}}_i, \quad (6)$$

where δT_i is the difference between the lower-resolution annual-average station temperature of interest (T_i) and $\hat{\bar{T}}_i$—the climatological average temperature interpolated from the higher-resolution network to lower-resolution station i, according to $\hat{\bar{T}}_i = I[\mathbf{T}_{n(i)}, \lambda, \phi]$. The vector $\mathbf{T}_{n(i)}$ contains nearby temperatures observed on the higher-resolution station network that influence the estimate of $\hat{\bar{T}}_i$. Nearby differences (δT_i), as well as the climatologically averaged air temperatures from the

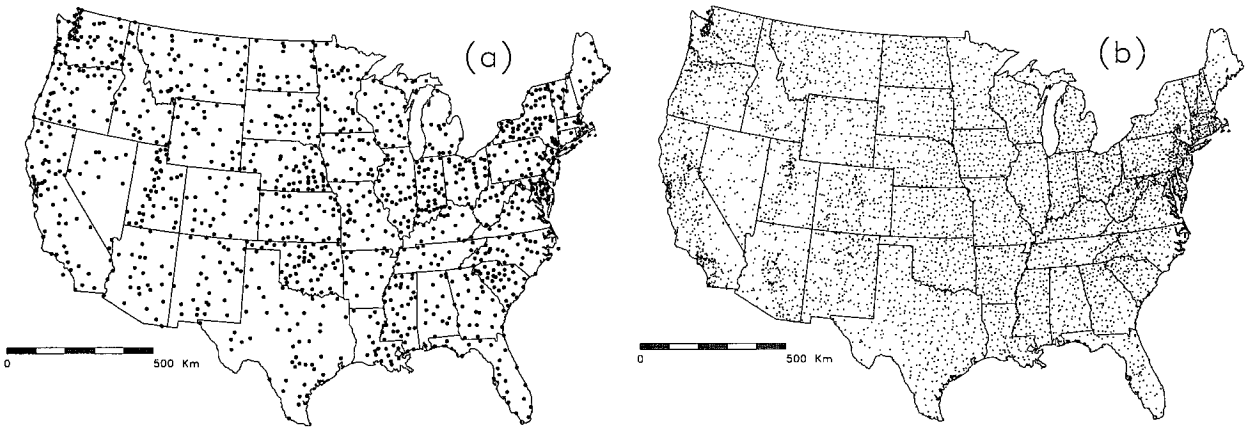


FIG. 1. Station locations of (a) all 1228 stations of the low-resolution HCN network and (b) the 5685-station high-resolution Legates and Willmott (1990) network.

higher-resolution station network, then are interpolated to grid point j. Expressed generically, the estimation is

$$\delta\hat{T}_j = I[\delta\mathbf{T}_{n(j)}, \lambda, \phi] \quad (7)$$

and

$$\hat{\bar{T}}_j = I[\mathbf{T}_{n(j)}, \lambda, \phi], \quad (8)$$

where $\delta\hat{T}_j$ and $\hat{\bar{T}}_j$ are the interpolated temperature difference and higher-resolution climatologically averaged air temperature at grid point j, respectively. The CAI annual-average air temperature estimate at grid point j then is

$$\hat{T}_j = \delta\hat{T}_j + \hat{\bar{T}}_j. \quad (9)$$

Willmott and Robeson (1995) demonstrated that CAI can reduce interpolation errors associated with interpolating average temperatures over continental-scale areas by more than 50%. A somewhat subtle aspect of CAI, which contributes to its accuracy, is bias-correction information about the underlying high-resolution climatology that resides within the δT_i field and consequently within the $\delta\hat{T}_j$ field. When $\delta\hat{T}_j$ is added back onto $\hat{\bar{T}}_j$, the estimated $\hat{T}_j$ field then approaches the T_i field in accuracy and "unbiasedness." Willmott and Robeson (1995) elaborate on this aspect of CAI.

d. Topographically and climatologically informed interpolation

As DEM-aided interpolation and CAI explain correlated but somewhat different components of the spatial variance, their combined application should further reduce interpolation error. Our approach is to first reinterpolate the higher-resolution climatologically averaged air temperature field to the nodes of the DEM as well as to the lower-resolution stations, using our DEM/mean environmental lapse rate algorithm. This has the effect of improving the spatial accuracy of the higher-resolution (climatological) air temperature field.

As before, the higher-resolution (climatologically averaged) station temperatures are reduced to sea level at the Γ rate. Estimated higher-resolution (climatologically averaged) sea level air temperatures then are interpolated to the DEM grid, as well as to the lower-resolution air temperature station network. Interpolated higher-resolution sea level air temperatures at each DEM grid point j, and at lower-resolution station i, are then raised to the DEM grid point and DEM-estimated station (i) heights at the Γ rate.

With the new (DEM aided) higher-resolution (climatologically averaged) air temperatures estimated at the desired DEM nodes ($\hat{\bar{T}}_j$) and at the lower-resolution stations ($\hat{\bar{T}}_i$), CAI is performed. The calculation is

$$\delta T_i = T_i - \hat{\bar{T}}_i \quad (10)$$

and

$$\hat{T}_j = I[\delta\mathbf{T}_{n(j)}, \lambda, \phi] + \hat{\bar{T}}_j, \quad (11)$$

where the temperature-difference vector $[\delta\mathbf{T}_{n(j)}]$ has the same interpretation as in Eq. (7) except that the differences are taken with respect to the DEM-aided interpolations of higher-resolution temperature rather than with respect to the traditionally interpolated higher-resolution temperature field (Willmott and Robeson 1995).

3. Air temperature and elevation data for the United States

Our illustrations and tests of the above-described interpolation methodologies are made by estimating yearly average temperatures at unsampled locations within the United States Historical Climatology Network (HCN) (Karl et al. 1990). This is our lower-resolution network. Legates and Willmott's (1990) average air temperature climatology serves as our higher-

resolution observational (station) network. A very high resolution DEM also is used to obtain the georeferenced heights (z_i and z_j) for station and grid networks. Pertinent properties of each of these data bases are outlined below.

a. *HCN time series of air temperature*

Air temperature data within the HCN include time series of monthly maximum, minimum, and mean air temperature for 1228 stations in the contiguous United States (Fig. 1a). The majority of these stations have serially complete records, back through 1920 (Fig. 2). Our analyses (below) are based on the annually averaged station air temperatures only. While both original (unadjusted) and adjusted HCN monthly temperatures are encoded (with confidence estimates for the adjusted temperatures), we use the original data, primarily because the adjustments involve interstation correlations (an implicit spatial interpolation method). If we had used the adjusted station records, our interpretations of interpolator efficacy would have been less clear. Inasmuch as the unadjusted HCN temperatures are more representative of averages contained in most historical station-record archives, interpolating the original temperatures also is a more typical application. Use of the original HCN temperatures additionally provides a somewhat more challenging test for the interpolation algorithms.

b. *Legates and Willmott's mean air temperature fields*

Legates and Willmott (1990) compiled a spatially high-resolution global archive of mean monthly and annual surface air temperature from some 10 sources. The data were screened for a variety of errors, and redundant station records were discarded. It remains one of the highest spatial-resolution global air temperature archives in existence. Since Legates and Willmott's main purpose was to compile a spatially high-resolution station air temperature database, long-term station averages from different averaging periods were included. Including stations with variable long-term record lengths was and is desirable because more of the unexplained variance in annual and longer-term average air temperature is spatial than is temporal. More specifically, at the spatial resolutions of most available longer-term station-record archives, more of the uncertainty in estimating annual and longer-term average air temperature at unsampled locations (between the stations) derives from between-station spatial variability than from unsampled temporal variability (Willmott and Robeson 1995). Long-term station means from within the contiguous United States (5685 of them) are used here (Fig. 1b) to represent the spatially correlated higher-resolution temperature field (Fig. 3), drawn from the climatological record.

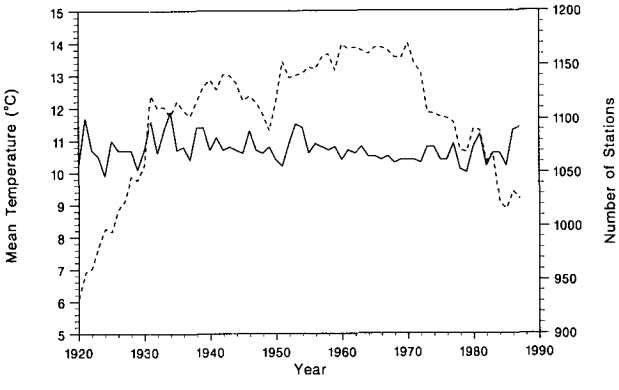


FIG. 2. Number of HCN stations available for each year (dashed line). Time series of estimated annual average air temperature over the United States from 1920 through 1987 is also plotted (solid line). Each yearly average was obtained by spatially integrating an interpolated field of the annual HCN averages. Interpolations were made by the topographically and climatologically informed interpolation algorithm (section 2d).

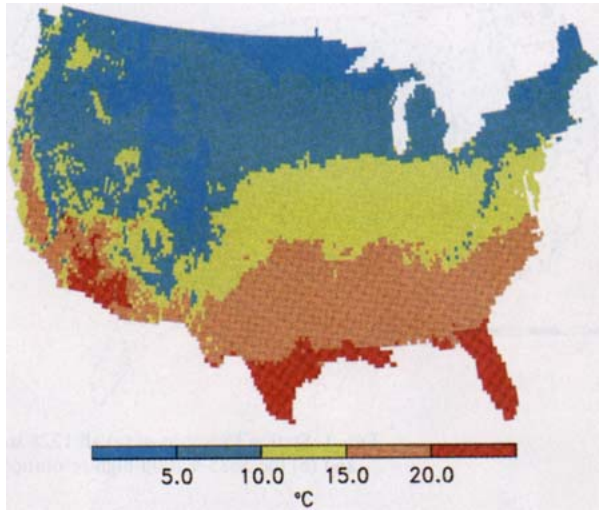


FIG. 3. Annual average air temperature over the United States based upon a DEM-enhanced interpolation from the Legates and Willmott (1990) station archive.

c. *Digital elevation data*

Georeferenced heights for the stations and grid points were drawn from a DEM cooperatively compiled by the U.S. Geological Survey (USGS), the National Geodetic Survey (NGS), the Defense Mapping Agency (DMA), and the National Geophysical Data Center (NGDC) (1993). The DEM elevations that we use are encoded as 30″ of latitude by 30″ of longitude tile averages. The 30″-average DEM elevation for the cell within which a station or grid point falls is used as the station or gridpoint height.

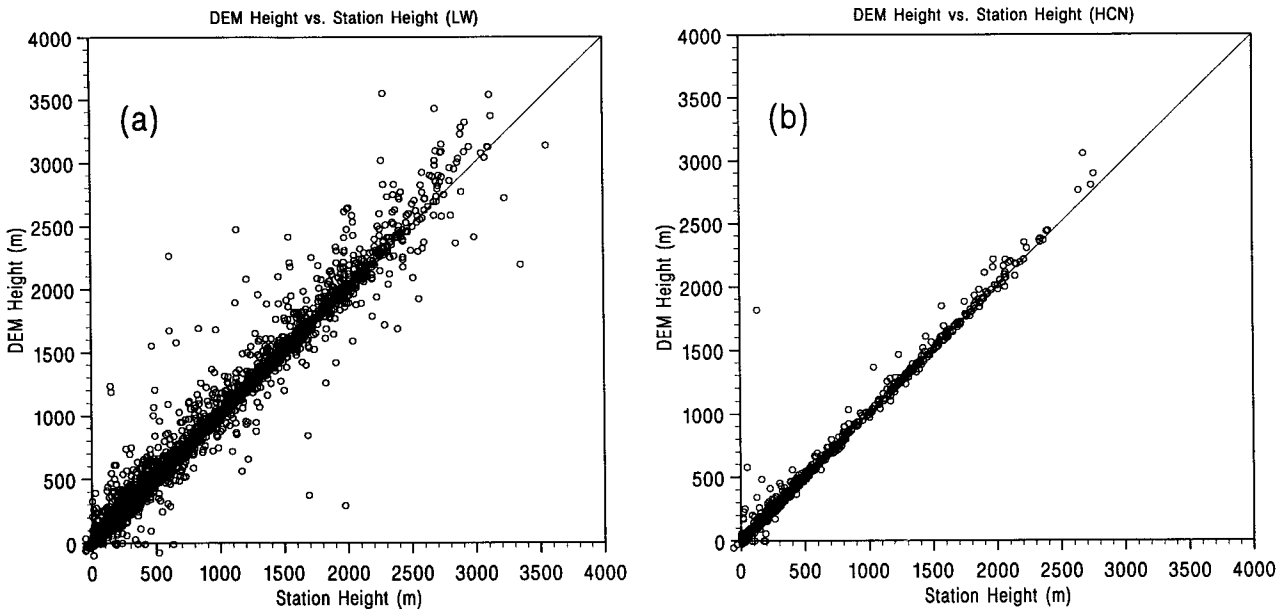


FIG. 4. Scatterplots of the DEM estimates (30″ averages) of station elevation vs the actual recorded station elevations for (a) the Legates and Willmott climatology and (b) the HCN station network.

Encoded documentation (on the source CD-ROM) indicates that this 30″ DEM was derived from a DMA digitization of 1:250 000 scale topographic maps. The resulting digital terrain map (DTM) contained point data at a spatial resolution of 3″. These elevations then were sampled at a 30″ resolution and each selected elevation was rounded to the nearest 20 ft. A copy of this 30″ DTM then was transmitted to the NGS, where "corrections were made." It also was reformatted by NGS and the elevations were rounded to the nearest 10 m. The 30″ tile averages that we use (each is an average of four NGS corner elevations) were made by the USGS and distributed by the NGDC.

4. Performance of interpolators

Simple cross validation is employed to evaluate comparatively the performance of each interpolation method. Errors are estimated by removing one station from the HCN network at a time and then interpolating the annual-average air temperature observed on it from the corresponding average temperatures observed on the other (remaining) stations. This process is repeated for all n stations, with each removed station being reinserted back into the network after it has been interpolated. Interpolation error at each HCN station then can be represented by $(\hat{T}_i - T_i)$. When Legates and Willmott's station climatology informs the interpolator, and the HCN station being cross validated occurs at

TABLE 1. Summary of cross-validation errors for the HCN network.

Interpolation method	MAE	rmse
DEM-assisted LW climatology	0.73	1.00
Traditional with HCN	0.58	0.98
DEM-assisted with HCN	0.44	0.71
DEM-assisted LW + CAI	0.38	0.64

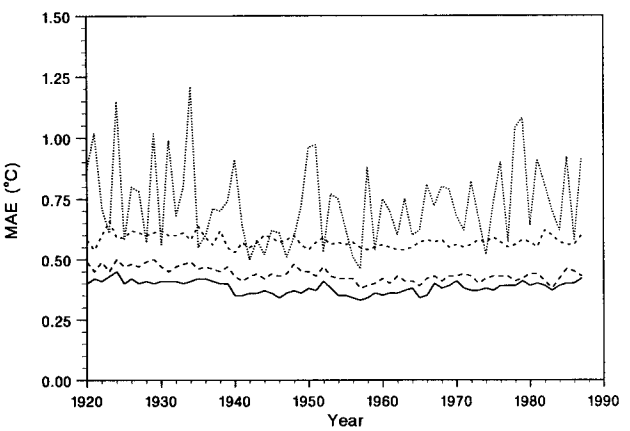


FIG. 5. Time series of spatially averaged MAEs (for each year) associated with the yearly average HCN station temperatures interpolated to and cross validated at the HCN stations. Interpolators include (a) the DEM-enhanced version of Legates and Willmott's long-term-average field, which assumes no temporal variability (dotted line), (b) traditional interpolations from each of the yearly HCN station networks alone (short-dashed line), (c) topographically informed (by the DEM) interpolations from each of the yearly HCN station networks (longer-dashed line), and (d) topographically and climatologically (by the DEM-enhanced Legates and Willmott average fields) assisted interpolations from each of the yearly HCN station networks (solid line).

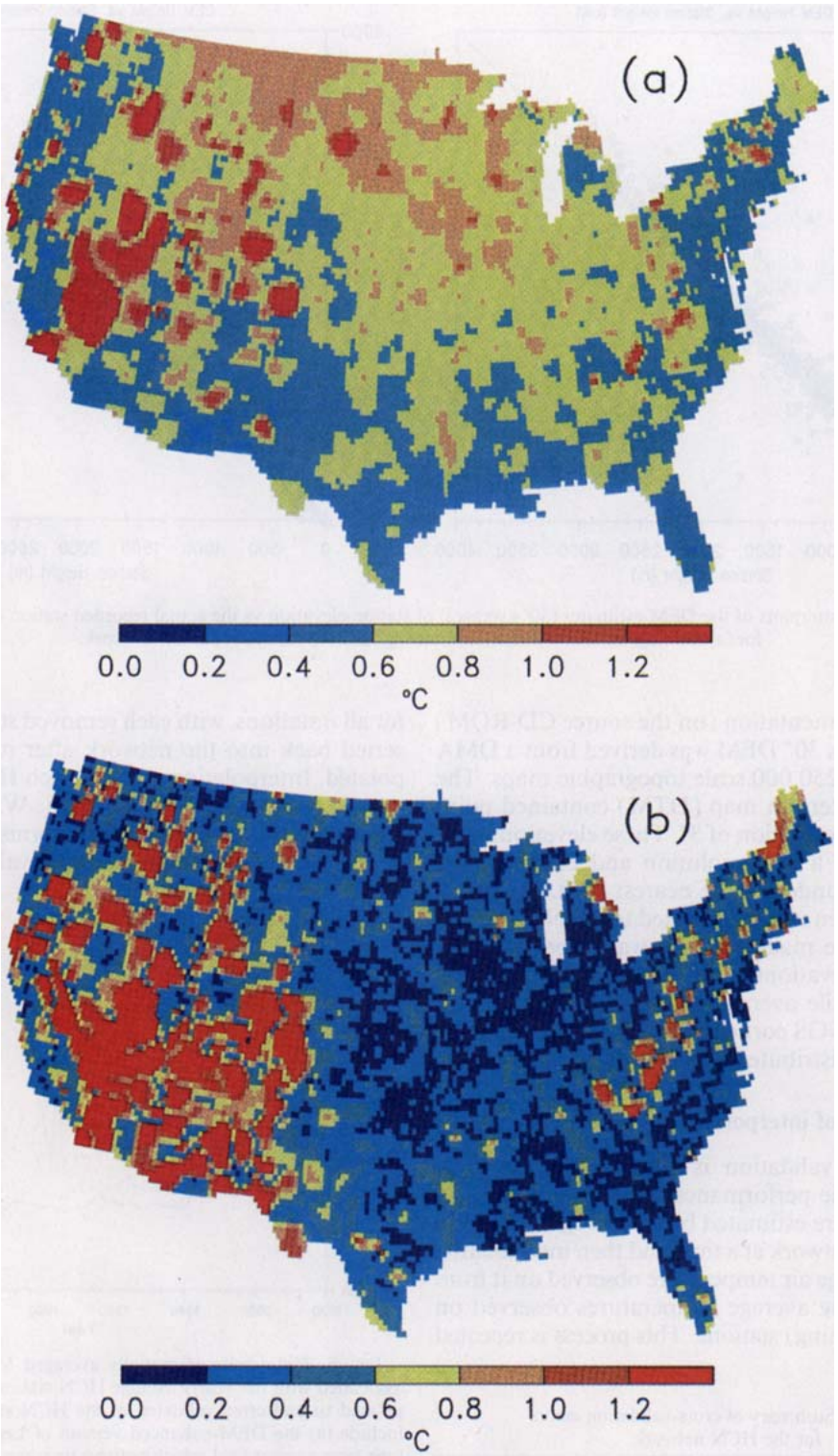
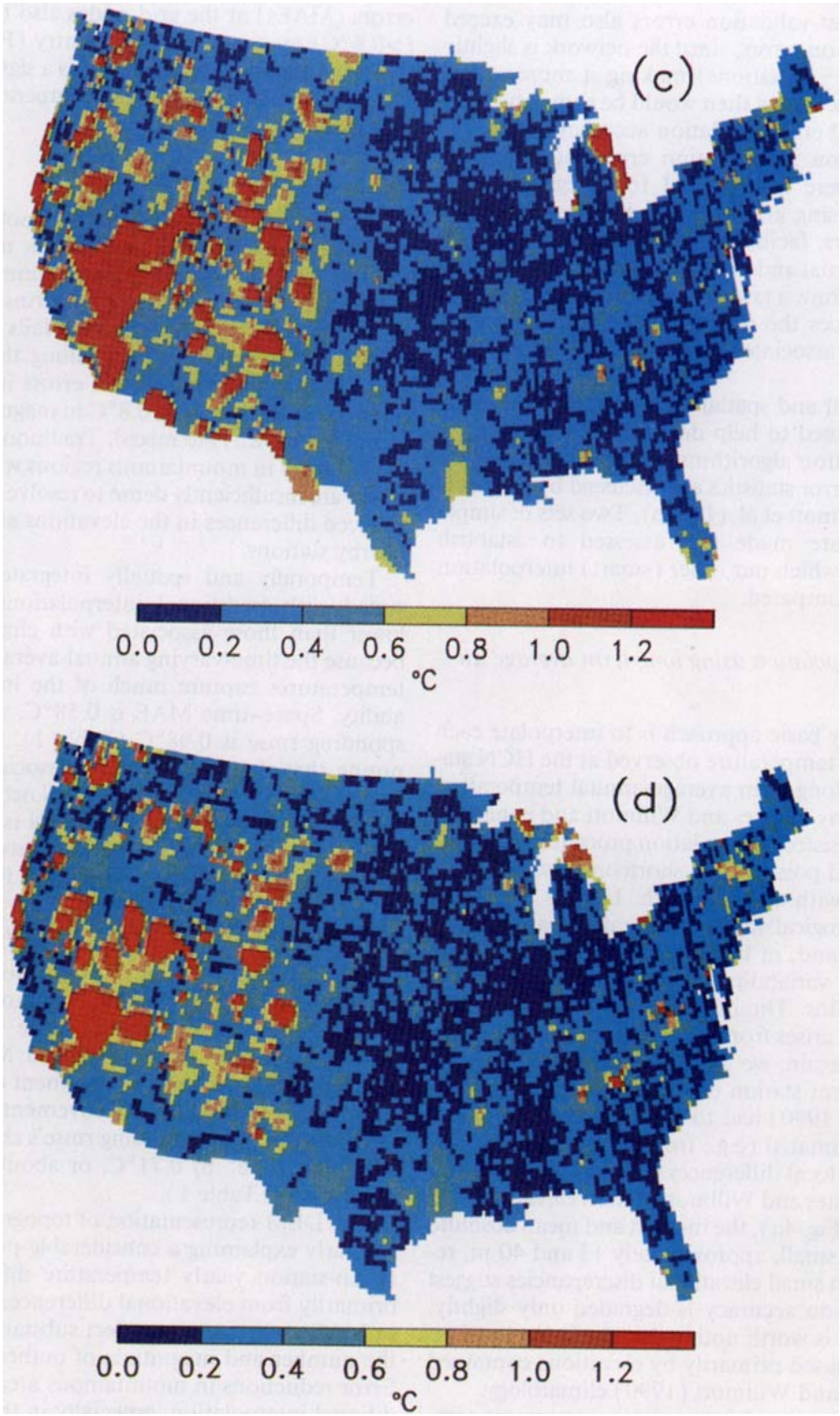


FIG. 6. Estimated spatial distributions of time-averaged (1920–87) yearly cross-validated interpolation errors (MAEs) at the HCN stations. Interpolators include (a) the DEM-enhanced version of Legates and Willmott's long-term-average field, which assumes no temporal variability, (b) traditional interpolations from each of the yearly HCN station networks alone, (c) topographically informed (by the DEM) interpolations from each of the yearly HCN station networks, and (d) topographically and climatologically (by the DEM-enhanced Legates and Willmott average fields) assisted interpolations from each of the yearly HCN station networks.



the same location as a Legates and Willmott station, the collocated Legates and Willmott station is temporarily removed along with the HCN station. Two other characteristics of cross validation also should be mentioned. As with interpolation in general, ill-conditioned station networks may bias the cross-validated

error fields. Cross-validation errors also may exceed actual interpolation errors, since the network is slightly degraded (to $n - 1$ stations) making it more sparse locally. "Real" accuracy then would be somewhat better than reported cross-validation accuracy.

Cross-validation interpolation errors at the HCN stations then were interpolated to a regular 0.25° × 0.25° grid. Using gridded error fields, rather than the station errors, facilitates mapping. Perhaps more importantly, spatial and temporal averages of the error fields obtained from a gridded rather than station-network field reduces the deleterious impacts of spatial sampling biases associated with irregularly spaced station networks.

Both temporal and spatial integrations of the error fields are examined to help determine the accuracies of the interpolation algorithms. Computation and interpretation of error statistics are discussed by Willmott (1984) and Willmott et al. (1985b). Two sets of simple interpolations are made and assessed to establish benchmarks to which our other (smart) interpolation results can be compared.

a. *Spatial interpolation using long-term average air temperature*

A particularly basic approach is to interpolate each annual-average temperature observed at the HCN stations from the long-term average-annual temperature field presented by Legates and Willmott and enhanced by our DEM-assisted interpolation procedure (Fig. 3). At least one and possibly two shortcomings, however, are associated with this approach. Legates and Willmott's climatological mean temperatures are temporally invariant and, in turn, cannot resolve the year-to-year spatial variability in the annually averaged temperature fields. The other, albeit relatively minor source of error, arises from using estimated station elevations. Once again, we use 30″-average DEM elevations to represent station elevations, whereas Legates and Willmott (1990) had to encode a mixture of recorded and estimated (e.g., from atlases) station elevations. While local differences (outliers) between the DEM and Legates and Willmott station elevations can be quite large (Fig. 4a), the median and mean absolute differences are small, approximately 13 and 40 m, respectively. Such small elevational discrepancies suggest that interpolation accuracy is degraded only slightly. Nonetheless, it is worth noting that the larger local errors may be caused primarily by elevations contained in the Legates and Willmott (1990) climatology.

When the associated interpolation errors are temporally and spatially integrated across the contiguous United States, the mean absolute interpolation error (MAE) is 0.73°C, while the corresponding root-mean-square error (rmse) is 1.00°C (Table 1). Spatially averaged MAEs (Fig. 5a) can be temporally variable, as well as quite high (e.g., >1.0°C). Temporally averaged errors (MAEs) at the grid nodes also tend to be high (>0.6°C) over most of the country (Fig. 6a). Nonetheless, climatology alone explains a significant portion of the space–time variance in temperature (Willmott and Robeson 1995).

b. *Traditional interpolation*

Traditional interpolation is another relatively straightforward approach, and the one most commonly used in the atmospheric and environmental sciences. While it [Eqs. (1) and (2)] performs relatively well (Figs. 5b and 6b), it commonly fails in the western mountain regions, as well as along the Appalachian range (Fig. 6b). Interpolation errors in mountainous terrain frequently exceed 0.8°C in magnitude, and their signs (not shown) are mixed. Traditional interpolation methods fail in mountainous regions when station networks are insufficiently dense to resolve topographically induced differences in the elevations and exposures of nearby stations.

Temporally and spatially integrated MAEs associated with traditional interpolations are generally lower than those associated with climatology alone, because the time-varying annual-average HCN station temperatures capture much of the interannual variability. Space–time MAE is 0.58°C, while the corresponding rmse is 0.98°C (Table 1). It also is worth noting that the average errors associated with traditional interpolations from the HCN network are among the lowest in the world, as the HCN is among the spatially highest-resolution station networks available (Willmott et al. 1994; Willmott and Robeson 1995).

c. *Topographically informed interpolation*

Topographically (DEM) informed interpolation [Eqs. (3)–(5)] from the HCN stations considerably reduces interpolation errors (Table 1 and Fig. 5c). Spatially and temporally averaged MAE is 0.44°C, which represents a 40% improvement over climatology alone, and nearly a 25% improvement over traditional interpolation. Corresponding rmse's also dropped from 1.00° and 0.98° to 0.71°C, or about 29% and 28%, respectively (Table 1).

The DEM representation of topographic variability is clearly explaining a considerable portion of the between-station yearly temperature differences, arising primarily from elevational differences. Improvements in MAE and rmse also reflect substantial decreases in the number and magnitude of outliers (large errors). Error reductions in mountainous areas relative to traditional interpolation, especially in the Appalachians, are clearly visible (Figs. 6b,c). Median and mean absolute differences between DEM- and HCN-station elevations (about 9 and 21 m, respectively) also are quite small; there are few outliers as well (Fig. 4b). This correspondence indicates that the DEM well represents the actual (HCN) station elevations.

d. Topographically and climatologically informed interpolation

Our combined DEM-assisted and climatologically aided interpolation [Eqs. (6)–(11)] uses the DEM-improved Legates and Willmott (1990) archive as the higher-resolution climatological field. This method exhibits the best overall performance (Table 1, Figs. 5 and 6d). Overall error (MAE) is 0.38°C, while the corresponding rmse is 0.64°C (Table 1). Relative to traditional interpolation, MAE and rmse are reduced by nearly 35%. Error reductions are dramatic, especially in mountainous areas (Fig. 6d). Improvements are marked over much of the country because CAI is explaining a significant portion of the annual-average air temperature variance left unaccounted for by DEM-improved climatology. It additionally is worth noting that, in most other parts of the world where time series networks (such as the HCN) are sparse, improvements over traditional methods would be even more impressive.

One reason why topographically and climatologically informed interpolation works well is that much of the spatial variability in short-term lapse rates is filtered out when longer-term time averages are made. Our use of a constant lapse rate (Γ), therefore, appears to introduce relatively little error; that is, when interpolations are made on annually or longer-term averaged temperature fields. Potential errors associated with lapse-rate spatial variability are reduced further when the station networks (the HCN and Legates and Willmott's network, in this instance) are sufficiently dense to resolve the spatial lapse-rate variability. If the station networks resolve the spatial lapse-rate variability, in other words, it will be accounted for by the interpolator. Application of our techniques to short-term averaged temperature fields, however, may not work as well owing to the increased spatial variability of short-term lapse rates.

It is difficult to precisely identify the relative importance of the topographical (DEM) and climatological (Legates and Willmott) contributions because they are correlated and comparisons are dataset dependent. Nonetheless, while CAI by itself [Eqs. (6)–(9)] performs well (Willmott and Robeson 1995), DEM-assisted interpolation (again by itself) generally performs better, at least for annually or longer-term averaged air temperature observed on networks like the HCN. Neither alone, however, performs as well as the two used together.

An intriguing aside is that our estimates of nationwide yearly average temperatures suggest a slight decline over the last seven decades (Fig. 2). While this may be real, it also may be due to improved station networks (Figs. 2, 5, and 6), notably in cooler environments. Changes in observation time (Schaal and Dale 1977) and station relocations from built environs, for instance, also may contribute to a cooling "bias" in recent decades. Although our "smart" interpolators probably reduce some of the adverse station-location effects, absence of a warming signal in this curve may be spurious.

5. Summary and conclusions

Four interpolation procedures have been described and evaluated using annual-average air temperature from the United States and a cross-validation methodology. Two of the interpolation approaches were "smart" in that they incorporated spatially high-resolution digital elevation information, an average environmental lapse rate, and/or another higher-resolution longer-term-average temperature field. Two simpler approaches served as benchmarks to which the smart interpolators could be compared—to help gauge improvements realized by smart methods. Climatology alone was used as one of the benchmark procedures, while traditional interpolation (from a single realization of annual-average air temperature on a single station network) was the other.

Both smart approaches exhibited significant reductions in interpolation error, relative to interpolation from climatology alone or by traditional means. Of the two, however, the one that combined DEM information and a higher-resolution climatologically averaged air temperature field performed the best. It posted an overall average interpolation error of only 0.38°C per year for the entire United States, which makes it nearly 35% more accurate than traditional interpolation. Inasmuch as the spatial resolutions of most station time series archives are lower than the resolution of the HCN archive, it is likely that the relative (to traditional interpolation) performance of smart interpolation in other applications will be even better than reported here. It is clear that the incorporation of higher-resolution spatially correlated information can reduce time-averaged air temperature interpolation errors significantly.

Acknowledgments. Portions of this research were funded by NASA Grants NAGW-1884 and NAGW-4355, and U.S. EPA Cooperative Agreement CR 816 278. A number of helpful discussions with Scott Robeson (Indiana University) and Scott Webber (University of Delaware) also are greatly appreciated. An anonymous reviewer additionally suggested several worthwhile improvements.

REFERENCES

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

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical–topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.,* **33,** 140–158.

Ishida, T., and S. Kawashima, 1993: Use of cokriging to estimate surface air temperature from elevation. *Theor. Appl. Climatol.,* **47,** 147–157.

Karl, T. R., C. N. Williams Jr., and F. T. Quinlan, 1990: United States historical climatology network serial temperature and precipitation data. NDP—019/R1, 83 pp. [Available from Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN 37831.]

Legates, D. R., and C. J. Willmott, 1990: Mean seasonal and spatial variability in global surface air temperature. *Theor. Appl. Climatol.,* **41,** 11–21.

National Geophysical Data Center (NGDC), 1993: Global Relief Data (Compact Disc). [Available from National Geophysical Data Center, 325 Broadway, E/GC4, Boulder, CO 80303-3328.]

Robeson, S. M., 1994: Influence of spatial sampling and interpolation on estimates of air temperature change. *Climate Res.,* **4,** 119–126.

Schaal, L. A., and R. F. Dale, 1977: Time of observation temperature bias and "climatic change." *J. Appl. Meteor.,* **16,** 215–222.

Shepard, D., 1968: A two-dimensional interpolation function for irregularly spaced data. *Proc. of the 23d Natl. Conf.,* ACM, 517–523.

Weber, D., and E. Englund, 1992: Evaluation and comparison of spatial interpolators. *Math. Geol.,* **24,** 381–389.

Willmott, C. J., 1984: On the evaluation of model performance in physical geography. *Spatial Statistics and Models,* G. L. Gaile and C. J. Willmott, Eds., D. Reidel, 443–460.

——, and S. M. Robeson, 1995: Climatologically aided interpolation (CAI) of terrestrial air temperature. *Int. J. Climatol.,* **15,** 221–229.

——, C. M. Rowe, and W. D. Philpot, 1985a: Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *Amer. Cartogr.,* **12**(1), 5–16.

——, S. G. Ackleson, R. E. Davis, J. J. Feddema, K. M. Klink, D. R. Legates, J. O'Donnell, and C. M. Rowe, 1985b: Statistics for the evaluation and comparison of models. *J. Geophys. Res.,* **90**(C5), 8995–9005.

——, S. M. Robeson, and J. J. Feddema, 1994: Estimating continental and terrestrial precipitation averages from rain-gauge networks. *Int. J. Climatol.,* **14,** 403–414.