

## A Method to Determine Station Density Requirements for Climate Observing Networks

RUSSELL S. VOSE AND MATTHEW J. MENNE

*National Climatic Data Center, Asheville, North Carolina*

(Manuscript received 2 October 2003, in final form 12 February 2004)

### ABSTRACT

A procedure is described that provides guidance in determining the number of stations required in a climate observing system deployed to capture temporal variability in the spatial mean of a climate parameter. The method entails reducing the density of an existing station network in a step-by-step fashion and quantifying subnetwork performance at each iteration. Under the assumption that the full network for the study area provides a reasonable estimate of the true spatial mean, this degradation process can be used to quantify the relationship between station density and network performance. The result is a systematic “cost–benefit” relationship that can be used in conjunction with practical constraints to determine the number of stations to deploy.

The approach is demonstrated using temperature and precipitation anomaly data from 4012 stations in the conterminous United States over the period 1971–2000. Results indicate that a U.S. climate observing system should consist of at least 25 quasi-uniformly distributed stations in order to reproduce interannual variability in temperature and precipitation because gains in the calculated performance measures begin to level off with higher station numbers. If trend detection is a high priority, then a higher density network of 135 evenly spaced stations is recommended. Through an analysis of long-term observations from the U.S. Historical Climatology Network, the 135-station solution is shown to exceed the climate monitoring goals of the U.S. Climate Reference Network.

### 1. Introduction

Most in situ atmospheric observing systems were established in support of agriculture, aviation, hydrology, and weather forecasting. Data from these networks are also widely used in the detection of climatic change. Unfortunately, certain practices that may be of little significance in an operational environment, such as relocating a station or shifting its observation time, may have a profound impact on the integrity of the historical record (Conrad and Pollak 1950). As a result, data from most existing networks are poorly suited for assessing climatic change (National Research Council 1999) unless adjustments are applied to account for historical variations in observing practice (Peterson et al. 1998). Even after the careful application of such adjustments, it is often not possible to address many critical aspects of climate variability and change (Karl et al. 1995). This situation has led to increasingly frequent calls for the deployment of new observing systems that are specifically designed and managed for climate monitoring purposes (Goody et al. 2002). It has also led to the establishment of reference networks that consist of high-quality, spatially representative stations from existing

observing systems (e.g., the Global Climate Observing System Surface Network; Peterson et al. 1997).

While observations from these new networks could have many uses, they will almost certainly be employed to assess the temporal variability in the mean of a climate parameter over a particular area (e.g., the long-term temperature trend in the United States). Indeed, such considerations are likely to be intrinsic to the design specifications of a climate observing network. Consequently, this paper describes a simple empirical procedure that provides guidance in determining the number of stations required for a network to accomplish that task given a set of performance requirements. The approach involves incrementally degrading the station density in an existing network for a study area and quantifying the rate at which representativeness deteriorates. Assuming that the existing station network is sufficiently dense such that a good approximation of the true spatial mean can be calculated, this process of incremental degradation allows the cost–benefit relationship between station density and network performance to be quantified. In a real-world network application, this cost–benefit information could be weighed against stated performance requirements and practical constraints in order to determine the number of stations to deploy. The approach is illustrated using temperature and precipitation anomaly data from the Cooperative Observing Network in the conterminous United States over the

---

*Corresponding author address:* Dr. Russell S. Vose, National Climatic Data Center, 151 Patton Avenue, Asheville, NC 28801.  
E-mail: Russell.Vose@noaa.gov



FIG. 1. Locations of stations used in the investigation.

period 1971–2000. As a practical demonstration, the results from the analysis are applied to estimate the number of stations required to meet the monitoring goals of the U.S. Climate Reference Network (CRN), a new observing system being deployed by the National Oceanic and Atmospheric Administration (NOAA) to quantify climate variability and change in the United States over the next 50–100 yr.

## 2. Data

Data for the study were extracted from the U.S. Climate Normals dataset (Owen and Whitehurst 2002). This dataset contains quality-assured mean monthly maximum temperature, mean monthly minimum temperature, and total monthly precipitation time series for over 5300 stations in the conterminous United States over the period 1971–2000. To minimize the impact of data gaps, any station missing more than 15% of its record was excluded. This missing-data constraint eliminated 1288 stations from consideration, resulting in the 4012-station network depicted in Fig. 1. In general, this station network provides reasonably dense coverage of most of the United States, though a few areas in the west may be somewhat undersampled (particularly those areas at higher elevations).

While up to 15% of the record could be absent for any given station, less than 3% of the monthly values were missing for the network as a whole. To further minimize the impact of temporal data gaps, each missing month was estimated using a two-step procedure described by Karl and Williams (1987). First, an initial

estimate was computed using a weighted average of monthly observations at up to 20 neighboring stations (the weights being proportional to the correlation coefficient between the candidate series and each neighbor for the period of record for each calendar month). Second, the initial estimate was “tuned” by 1) computing the long-term mean difference between the candidate station and its neighbors and 2) adding that difference to the weighted average computed in the first step. Cross-validation errors using this approach were generally low (i.e., less than 0.5°C for temperature and 5% for precipitation approximately 75% of the time).

To help account for latitudinal and topographic variability, each monthly value at each station was expressed as a deviation from its long-term monthly mean (i.e., each monthly value was converted to an anomaly). For temperature, this involved subtracting a station’s 30-yr average for the calendar month from the station’s actual temperature in that month:

$$TA_{ij} = T_{ij} - \bar{T}_j,$$

where  $TA_{ij}$  is the station’s anomaly (in °C) for month  $j$  in year  $i$ ,  $T_{ij}$  is the station’s actual temperature for month  $j$  in year  $i$ , and  $\bar{T}_j$  is the station’s 30-yr average temperature for month  $j$ . Precipitation anomalies were computed by dividing a station’s 30-yr average for the calendar month into the actual precipitation total in that month:

$$PA_{ij} = P_{ij} / \bar{P}_j,$$

where  $PA_{ij}$  is the station’s anomaly (in percent) for month  $j$  in year  $i$ ,  $P_{ij}$  is the station’s total precipitation

for month  $j$  in year  $i$ , and  $\bar{P}_j$  is the station's 30-yr average precipitation for month  $j$ .

Both unadjusted and homogeneity-adjusted temperature data are available in the Climate Normals dataset, and both types of data were evaluated to determine whether network performance was impacted by changes in station location, time of observation, and other historical events. No significant impact was evident, suggesting that homogeneity adjustments are not required to describe the relationship between station density and network performance (i.e., the ability to capture the spatial mean) in the study area. Consequently, the results presented for temperature in this paper were obtained using raw data. The results presented for precipitation were also based upon raw data because the Climate Normals dataset does not contain homogeneity-adjusted precipitation records.

### 3. Methodology

The procedure described here assumes that the purpose of the observing system is to capture the temporal variability in the spatial mean of a climate parameter. There are four general steps in the approach. The initial step involves the development of sampling grids of various resolutions (e.g.,  $5^\circ$ ,  $4^\circ$ , etc.) to “overlay” on a map of the full station network for the study area. The grids are then applied as templates to extract a suite of spatially stratified subnetworks from the region's full station network. Next, each subnetwork is used to compute a time series of the spatial mean for the area; for comparative purposes a reference series of the spatial mean is also created using all stations in the full network. Finally, summary statistics are computed to quantify the similarity between each subnetwork time series and the full network reference series.

The procedure only quantifies the performance of spatially stratified networks for estimating the spatial mean. This distinguishes the approach from a number of recent techniques that focused on optimizing the arrangement of stations to capture spatial gradients (e.g., Collins et al. 1999; de Elía and Zawadzki 2001; Mazzarella and Tranfaglia 2000; Milewska and Hogg 2001; Schmalwieser and Schauburger 2001; Janis et al. 2004). Admittedly, a spatially stratified network may require more stations than a gradient-based distribution to achieve a comparable standard of performance because the latter can be designed to sample the “most representative” locations in a study area (Nicolis 1993). At the same time, however, the performance capabilities of a gradient-based network could deteriorate if the optimum locations in the future differ from those in the design period. In contrast, changes in climate anomaly patterns should in theory have a somewhat smaller impact on the performance of a spatially stratified network because stations are distributed throughout the study area. This is one of the reasons that the Global Climate Observing

System Surface Network, for instance, focused on a spatially stratified design (Peterson et al. 1997).

#### a. Sampling grids

The first step in the approach involves the creation of sampling grids that can be used to extract subnetworks from the full station network. Rather than traditional rectangular latitude–longitude boxes, each sampling grid consists of equilateral triangles of a specified size (e.g., triangles in which the great circle distance between each vertex is nominally  $5^\circ$  of arc). A triangular mesh is employed for three reasons. First, each node in an equilateral triangular lattice has six equidistant neighbors instead of four, resulting in a “more uniform” tessellation than a rectangular grid (Barnes 1964). Second, each node in such a triangular mesh represents approximately the same area as every other node (unlike a traditional latitude–longitude grid, in which boxes become smaller toward the poles). Finally, a triangular data sample tends to be slightly more efficient in estimating a spatial mean than its rectangular counterpart (Martin 1979; Ripley 1981). An example of a triangular lattice, one in which the great circle distance between each node is nominally  $5^\circ$  of arc, is presented in Fig. 2.

Sixteen triangular grids were developed for the conterminous United States. Grid size ranged from  $1.5^\circ$  of arc ( $\sim 160$  km) to  $9.0^\circ$  of arc ( $\sim 1000$  km) in increments of  $0.5^\circ$ . Figure 3 depicts the number of nodes associated with each grid resolution using a fixed origin of  $37^\circ\text{N}$  and  $95^\circ\text{W}$ , the approximate geographic center of the study area. The sparsest lattice contained only 10 nodes whereas the densest consisted of 360 nodes. As with rectangular grids, the relationship between grid resolution and the number of nodes is exponential.

#### b. Station subnetworks

Each triangular grid is used to extract a suite of stratified random subnetworks from the full station network. Stratified random networks are evaluated because practical constraints (e.g., lack of suitable land, presence of a water body) generally make it impossible to locate a station at the precise location of each grid node in a real-world application. The first step in this process involves assigning each station in the full network to the nearest node in each of the sampling grids. Figure 4 depicts this assignment outcome for the  $5^\circ$  sampling grid. Next, one station assigned to each node is randomly selected for inclusion in a given sample network. This second step is repeated 1000 times, yielding 1000 realizations of stratified random networks for each grid resolution. An example of a stratified random network obtained using the  $5^\circ$  sampling grid is presented in Fig. 5.

One “quasi-uniform” subnetwork is also produced for each grid. This is accomplished by matching each node in a sampling grid with the nearest station in the full network. As depicted by the example in Fig. 6, the

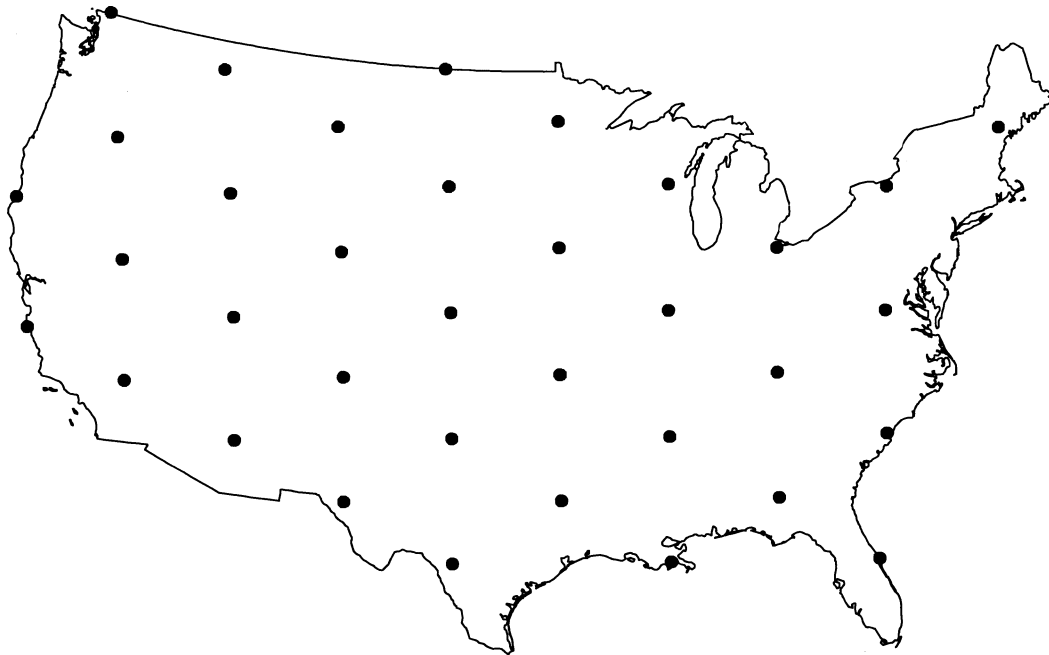


FIG. 2. Locations of nodes in the  $5^\circ$  triangular grid.

stations in these networks are relatively even in their distribution, but their arrangement is not perfectly uniform because the nearest station in the full network is rarely at the same location as the corresponding node in the sampling grid. Nevertheless, these networks are valuable because the uniform arrangement that they approximate is often more efficient in estimating areal averages when the distance–decay correlation function for the variable in question is monotonic (Haining 1993), which is largely true for temperature and precipitation anomalies in the United States (Groisman and Easterling 1994). As a result, they provide a baseline

for comparison with the stratified random networks. They also represent an obvious target strategy for station deployment.

### c. Time series of the spatial mean

Each stratified subnetwork is used to compute a time series of the spatial mean for the study area. A separate time series is prepared for each variable for the months of January and July as well as for the annual mean. Each value in a time series (e.g., the average minimum temperature anomaly across the United States in July 2000) is obtained by computing the simple arithmetic average of the anomalies at all stations in the subnetwork for the year or month in question. Because the stations in each subnetwork are spatially stratified, the anomalies are not interpolated to a grid before computing the spatial mean. The simple arithmetic average is used in place of interpolation in an effort to isolate the relationship between sample size and network performance. While interpolation might result in a slightly better estimate of the spatial mean, particularly for sparse networks, the spatial stratification by definition minimizes the potential for sampling error.

For comparative purposes, a “reference” time series is also created for each variable and month using anomalies from all stations in the full network. Because the 4012 stations used here are irregularly distributed in space, spatial means for each variable and month from 1971 to 2000 were obtained by interpolating the station-based anomalies to the nodes of a  $0.25^\circ \times 0.25^\circ$  latitude–longitude grid, then computing the cosine-weight-

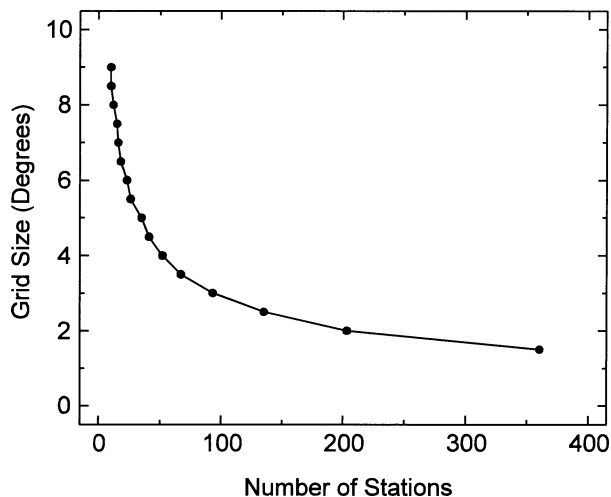


FIG. 3. Size of grid vs number of stations.

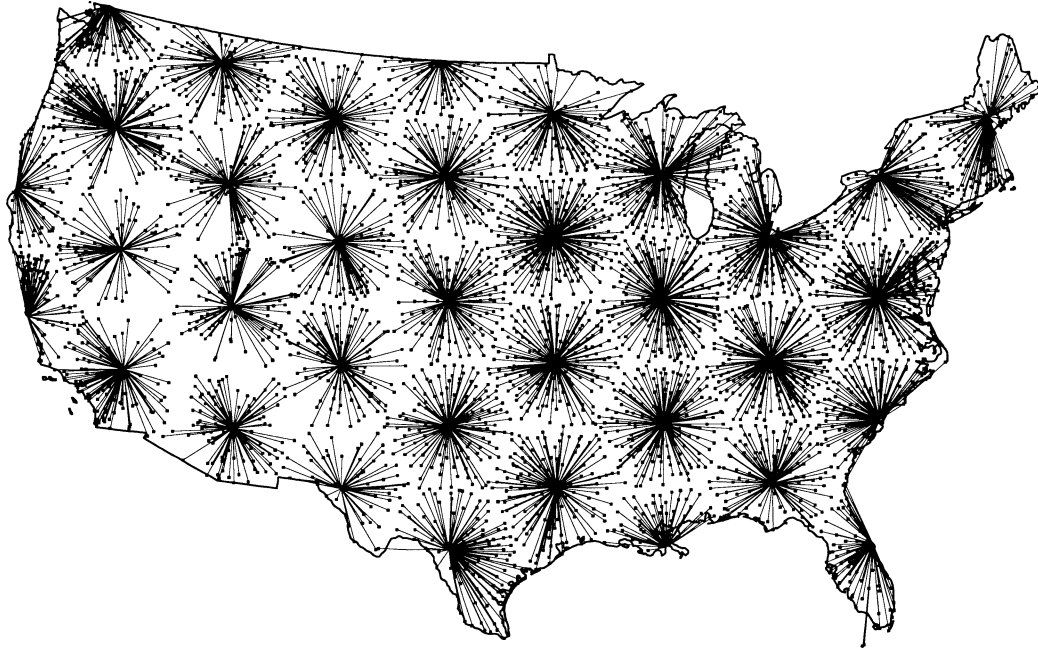


FIG. 4. The assignment of each station to the nearest node in the  $5^\circ$  triangular grid.

ed average of the gridded anomaly values. Interpolation was performed using the inverse distance weighting model of Willmott et al. (1985), which accounts for the sphericity of the earth when computing the distance between each station and grid point and permits extrapolation beyond the range of data values in the neighboring stations. The method has been used successfully for a variety of climatological problems (e.g., Bussieres and Hogg 1989; Huffman et al. 1995; Robeson 1995), and it has been shown to perform well for the interpolation of temperature and precipitation anomalies (Robeson 1997).

#### d. Network performance statistics

The last step involves evaluating the performance of each subnetwork. This is accomplished by computing statistics that quantify the similarity between each subnetwork time series and the full network reference series. Three such statistics are used as examples in this investigation: the coefficient of determination between a subnetwork series and a reference series, the mean absolute difference between the two series, and the difference in decadal trend between the two series. These statistics are employed here because they are commonly applied to describe the general similarity between time series and because the example application in section 5 (CRN) is intended to monitor changes in the long-term mean and not changes in, for instance, climate extremes. For each grid resolution, the median and the 95% confidence interval of each statistic are computed using the stratified random samples. The one instance of a quasi-

uniform sample for each grid resolution is also discussed.

## 4. Results

The approach was used to quantify the relationship between station density and network performance in the United States. In general, the results suggest that a network should contain at least 25 stations because large gains in performance occur up to that point. Similarly, there are few benefits to a network with hundreds of stations for national-scale applications because smaller configurations provide comparable performance. Finally, no “optimum” network size exists. To support these generalizations, the results for each performance measure are described in detail below.

### a. Coefficient of determination

The coefficient of determination ( $r^2$ ) quantifies the ability of a subnetwork time series to replicate the year-to-year fluctuations in a full network reference series. Figure 7 depicts the relationship between this statistic and station density on the annual time scale for both quasi-uniform and stratified random subnetworks. In general,  $r^2$  increases logarithmically as the number of stations increases, and most quasi-uniform subnetworks have  $r^2$  values that are within a few percent of the median of the stratified random samples at equivalent grid resolutions. Gains in performance are rapid to about 25 stations, at which point median  $r^2$  exceeds 95% for the maximum and minimum temperature. Modest improve-



FIG. 5. Locations of stations in a stratified random subnetwork for the  $5^\circ$  triangular grid.

ment continues, especially for precipitation, until the sample size roughly doubles. Increases in performance are small for networks in excess of 135 stations. Not surprisingly, from the standpoint of sampling variability, the width of the 95% confidence interval decreases as the density of the network increases.

Table 1 depicts the differences in  $r^2$  that are evident between the midseason months of January and July. In

general, median  $r^2$  is larger in January than in July for all variables, and the confidence interval is always wider in July than in January. Improvement in the median, and to a lesser extent the confidence interval, is gradual after the network reaches about 25 stations, though estimates of July minimum temperature and January precipitation do benefit slightly from higher resolution networks. In contrast, an extremely dense network—one

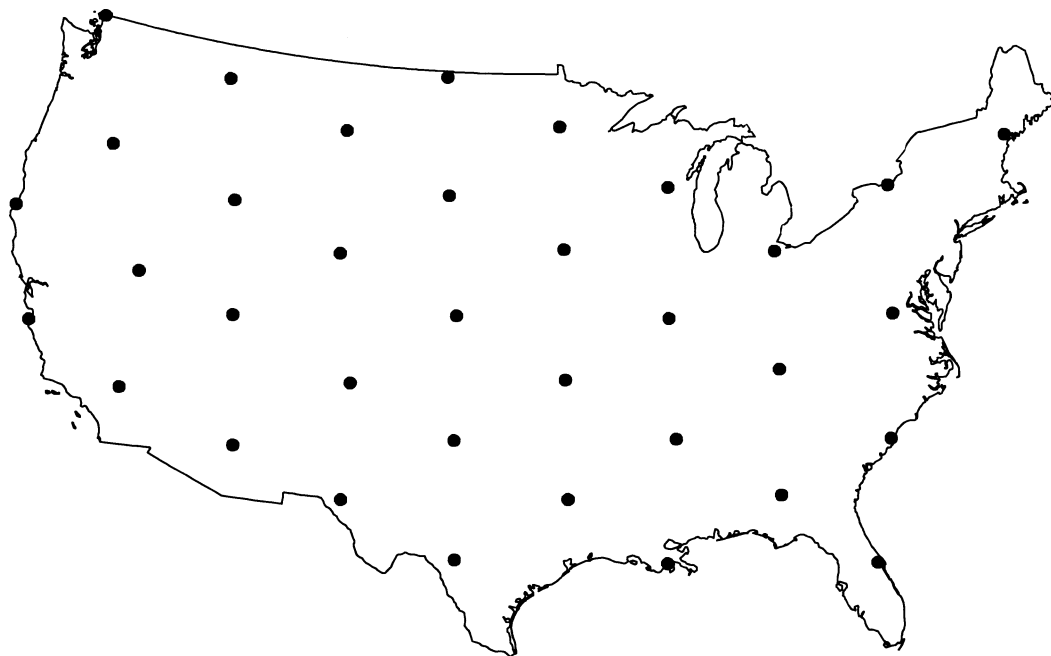


FIG. 6. Locations of stations in the quasi-uniform subnetwork for the  $5^\circ$  triangular grid.

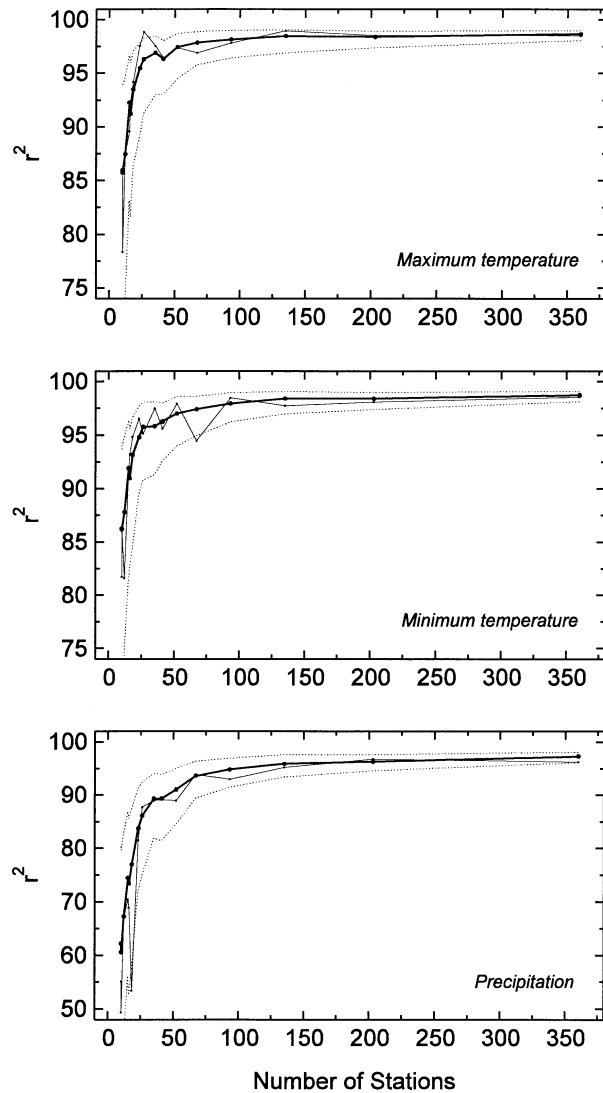


FIG. 7. Coefficient of determination ( $r^2$ ) vs number of stations for annual series. The thick line depicts the median of the random samples, the dotted lines depict the 95% confidence interval, and the thin line depicts the systematic samples.

with hundreds of stations—is needed to resolve variations in July precipitation, probably because of the high degree of spatial variability in rainfall during the convective season.

*b. Mean absolute difference*

The mean absolute difference (MAD) quantifies the average amount by which a subnetwork time series differs from a full network reference series:

$$MAD = \frac{1}{30} \sum_{i=1}^{30} |\hat{a}_i - a_i|,$$

where  $\hat{a}_i$  is the anomaly in year  $i$  in the subnetwork time series, and  $a_i$  is the anomaly in year  $i$  for the full network reference series. Figure 8 depicts the relationship between this statistic and station density on the annual time scale for both quasi-uniform and stratified random subnetworks. In general, MAD declines exponentially as the number of stations increases, and most quasi-uniform subnetworks have MAD values that are roughly coincident with the median of the stratified random samples at equivalent grid resolutions. Improvement is dramatic until the network reaches about 25 stations, at which point the average difference falls below a tenth of a degree for the maximum and minimum temperature and a few percent for precipitation. Thereafter, modest reductions continue until the network contains about 100 stations, at which point improvement levels off.

Table 2 depicts the differences in MAD that are evident between the midseason months of January and July. In general, median differences for maximum and minimum temperatures are larger in January than in July, and confidence intervals likewise tend to be smaller in July than in January. Both characteristics can probably be attributed to the larger standard deviations of temperature in winter than in summer. Reductions in MAD are small for networks in excess of about 25 stations, particularly for temperature. However, confidence intervals narrow appreciably beyond that point, dropping by about one-third each time the network roughly doubles in size.

TABLE 1. Coefficient of determination ( $r^2$ ) for Jan and Jul for select network sizes. Values in the table are derived from the stratified random samples. Also depicted are the widths of the 95% confidence interval (CI).

N	Maximum temperature				Minimum temperature				Precipitation			
	Median		95% CI		Median		95% CI		Median		95% CI	
	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul
12	93	72	12	43	92	64	15	44	72	17	37	49
26	98	93	3	10	98	87	4	18	87	47	16	51
52	99	94	2	7	98	90	2	12	91	59	11	40
93	99	96	1	4	98	93	2	8	94	70	6	29
203	99	96	1	3	98	95	1	5	95	82	4	14

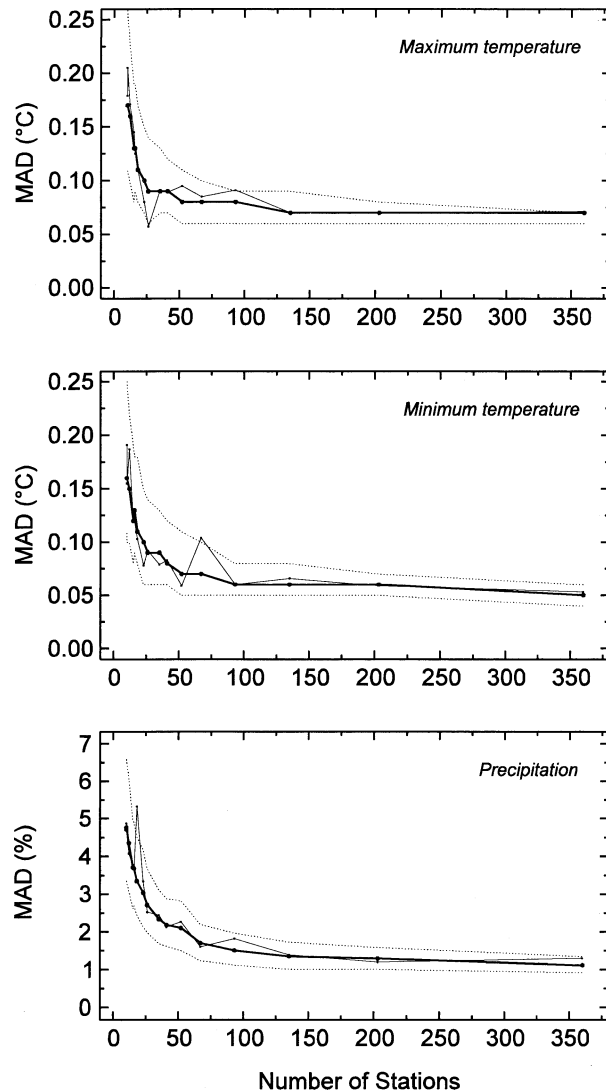


FIG. 8. MAD vs number of stations for annual series. The thick line depicts the median of the random samples, the dotted lines depict the 95% confidence interval (CI), and the thin line depicts the systematic samples.

### c. Trends in annual series

The difference in trend ( $\Delta b$ ) quantifies the ability of a subnetwork time series to replicate the time rate of change in a full network reference series:

$$\Delta b = b_s - b_f,$$

where  $b_s$  is the trend per decade (computed via least squares regression) in an annual subnetwork time series, and  $b_f$  is the trend per decade in the full network reference series. Figure 9 depicts the relationship between  $\Delta b$  and station density on the annual time scale. In general,  $\Delta b$  declines as the number of stations increases, but gains in performance occur more gradually than for  $r^2$  and MAD. For example, the confidence interval for each variable improves dramatically up to about 50 stations rather than 25. The median difference for each element also contains a small and variable bias until the network contains at least 50 stations. Furthermore, there is often a substantial difference between the performance of the coarse quasi-uniform subnetworks and the median of the stratified random samples at equivalent grid resolutions. In fact, by chance some coarse quasi-uniform subnetworks replicate trends with greater precision than their counterparts with more stations. In addition, performance characteristics differ from one element to the next (e.g., the 52-station network badly underestimates the maximum temperature trend and almost perfectly reproduces the minimum temperature trend). In general, quasi-uniform networks in excess of 100 stations are required to stabilize estimates of the “true” trend.

## 5. Example application

To illustrate their utility in a real-world situation, the findings in section 4 are employed to recommend a size for CRN. CRN is a network of climate stations now being deployed in the conterminous United States, Alaska, and Hawaii as part of a NOAA initiative (Heim 2001). The primary goal of its implementation is to provide long-term homogeneous observations of temperature and precipitation for use in the detection and attribution of climate change. Data from CRN will also be employed in operational climate monitoring activities and for placing current climate anomalies into historical

TABLE 2. MAD for Jan and Jul for select network sizes. Values in the table are derived from the stratified random samples. Also depicted are the widths of the 95% CI. Temperature values are in °C, and precipitation values are in percent.

N	Maximum temperature				Minimum temperature				Precipitation			
	Median		95% CI		Median		95% CI		Median		95% CI	
	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul	Jan	Jul
12	0.42	0.31	0.36	0.23	0.45	0.22	0.41	0.16	13.5	18.2	9.3	12.4
26	0.22	0.16	0.20	0.12	0.24	0.13	0.21	0.09	8.2	11.3	5.4	6.3
52	0.21	0.15	0.15	0.09	0.23	0.11	0.14	0.07	6.8	8.9	4.1	5.1
93	0.20	0.14	0.09	0.06	0.23	0.09	0.11	0.05	5.3	6.6	3.0	3.3
203	0.19	0.13	0.05	0.04	0.22	0.08	0.06	0.04	4.8	5.0	2.0	2.2



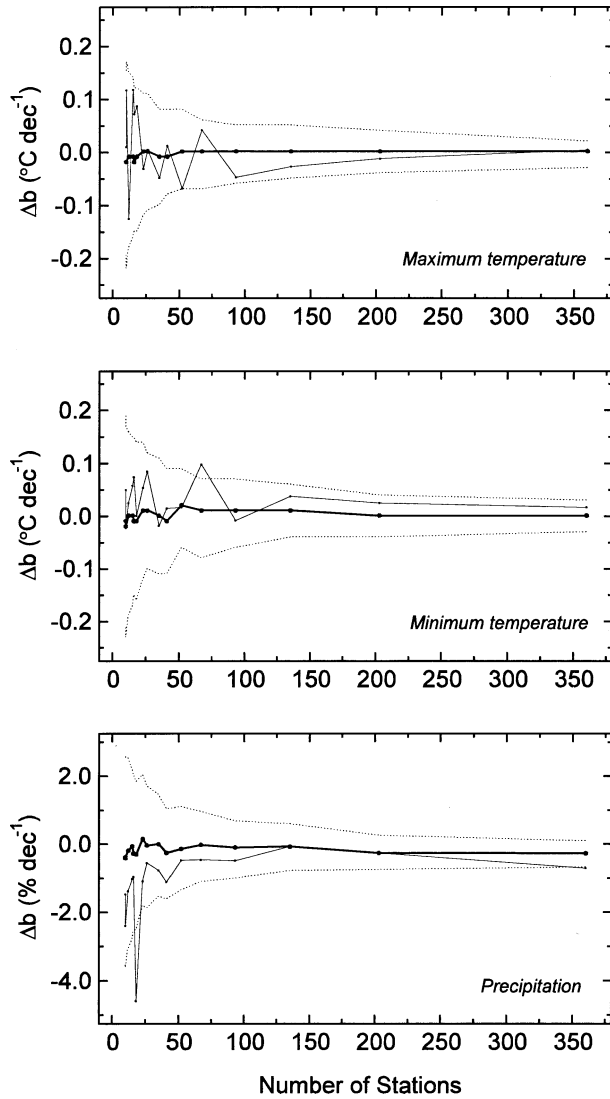


FIG. 9. Difference in decadal trend vs number of stations for annual series. The thick line depicts the median of the random samples, the dotted lines depict the 95% CI, and the thin line depicts the systematic samples.

perspective. When fully deployed, the network must capture 95% of the variance in the true annual temperature and precipitation anomaly time series of the United States (T. Karl 2003, personal communication). The purpose of these quantitative performance requirements is to ensure that the network can fulfill its mission of accurately detecting climatic change in the spatial mean of temperature and precipitation in the United States over the next 50–100 yr.

The results in section 4 indicate that CRN requires approximately 135 stations in the conterminous United States. The primary justification for this network size is the need to support CRN's quantitative performance requirements for temperature and precipitation on the annual time scale. Specifically, of the 16 network res-

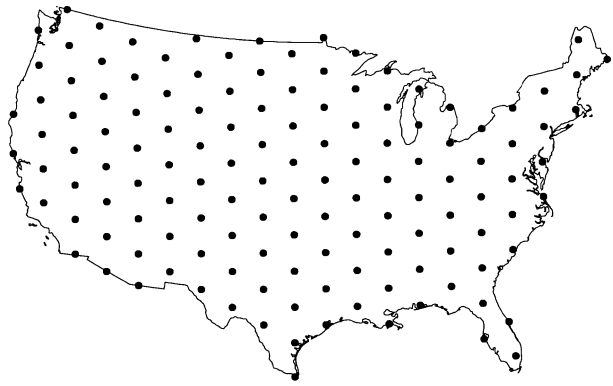


FIG. 10. Proposed station locations for CRN. Locations are based on a 2.5° triangular grid.

olutions considered in the study, the 135-station configuration is the smallest in which median  $r^2$  for the stratified random subnetworks exceeds 95% for each variable. This is also the smallest configuration in which the  $r^2$  values for the quasi-uniform subnetwork exceed 95% for each variable. While lower-density networks have performance characteristics that are nearly as good as a 135-station configuration, particularly for temperature, they do not meet the monitoring goals of CRN. Networks with more stations provide better estimates of the spatial mean, but the improvements are too small to justify the cost of additional stations.

As an independent verification of the proposed 135-station network, its performance capabilities are evaluated using data for the period 1910–70. Data for this period were extracted from the U.S. Historical Climatology Network (HCN) database (Easterling et al. 1996), which contains homogeneity-adjusted maximum and minimum temperature time series and unadjusted precipitation time series for 1221 stations in the conterminous United States. While this network contains fewer stations than the climate normals dataset, the results in section 4 indicate that the density of HCN should still be sufficient to generate a robust spatial mean for the study area. The first step in the analysis involved converting the annual mean for each variable at each station to an anomaly. Next, each node in the 135-station grid depicted in Fig. 10 was matched with the nearest HCN station. An annual time series was then prepared for each variable, each time series value being the simple arithmetic average of the anomalies at all 135 stations in that year. A full HCN reference time series was also created for each variable by interpolating the station-based anomalies at the 1221 stations to the nodes of a  $0.25^\circ \times 0.25^\circ$  latitude–longitude grid, then computing the cosine-weighted average of the gridded anomaly values. The performance of the 135-station network time series relative to the full HCN reference series was then evaluated in rolling 30-yr periods (e.g., 1910–39, 1911–40, and so on up through 1941–70).

The results confirm the suitability of a 135-station

network for CRN. For example, the lowest  $r^2$  produced by the network for any variable for any 30-yr period was 94%. Similarly, the largest MAD value for any 30-yr period was less than  $0.1^\circ\text{C}$  for the maximum and minimum temperature and less than 2% for precipitation, suggesting that annual anomalies can be estimated with precision. The network was also highly efficient in detecting linear trends. For instance, the average difference in trend between the 135-station network and the full HCN network was  $0.007^\circ\text{C decade}^{-1}$  for maximum temperature,  $-0.001^\circ\text{C decade}^{-1}$  for minimum temperature, and  $-0.177\%$   $\text{decade}^{-1}$  for precipitation. Furthermore, the worst temperature trend produced by the 135-station network differed by less than  $0.05^\circ\text{C decade}^{-1}$  from the full HCN.

## 6. Summary and conclusions

This paper describes a procedure to provide guidance in determining the number of stations required to capture changes in the spatial mean of a climate parameter over a specified region. The method entails degrading the station density of an existing network in an incremental fashion and quantifying network performance at each iteration. Assuming that the full network provides a reliable estimate of the true spatial mean, this degradation process will quantify the relationship between station density and network performance. This cost-benefit information can then be used in conjunction with network performance requirements and practical constraints to determine the number of stations to deploy.

The technique is illustrated using temperature and precipitation anomaly data from 4012 stations in the conterminous United States over the period 1971–2000, and thus the results should be useful in the design of U.S. climate observing systems. For example, if minimizing the number of stations is a high priority in a real-world application, then a network as small as about 25 stations may be a viable option because the largest gains in performance occur up to that point. If trend detection is a high priority, then a network of 135 stations is a better alternative because discernable improvement continues until the network reaches that size and because sampling variability becomes minor for higher station densities. A case study suggests that this latter network size should suit the performance requirements of CRN.

Given a sufficiently dense network, the method described in this paper could be employed to study other climate elements, time scales, periods, and areas. As an alternative approach, multiple stations could be averaged at each grid node in an effort to reduce the impact of microclimatic effects (e.g., Vucetic et al. 2000). In the absence of a high-density network, the technique could be applied to an analysis of numerical model output with a high spatial resolution. A similar approach could also be used to evaluate the relationship between

the density of an observing network and its ability to capture spatial variability.

*Acknowledgments.* The authors thank Randy Cerverny, David Easterling, Thomas Peterson, David Wuertz, and three anonymous reviewers whose comments substantially improved this manuscript.

## REFERENCES

- Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, **3**, 40–49.
- Bussieres, N., and W. Hogg, 1989: The objective analysis of daily rainfall by distance weighting schemes on a mesoscale grid. *Atmos.–Ocean*, **27**, 521–541.
- Collins, D. A., S. Johnson, N. Plummer, A. K. Brewster, and Y. Kuleshov, 1999: Re-visiting Tasmania's reference climate stations with a semi-objective network selection scheme. *Aust. Meteor. Mag.*, **48**, 111–122.
- Conrad, V., and C. Pollak, 1950: *Methods in Climatology*. 2d ed. Harvard University Press, 459 pp.
- de Elia, R., and I. Zawadzki, 2001: Optimal layout of a bistatic radar network. *J. Atmos. Oceanic Technol.*, **18**, 1184–1194.
- Easterling, D. R., T. R. Karl, E. H. Mason, P. Y. Hughes, D. P. Bowman, R. C. Daniels, and T. A. Boden, 1996: United States Historical Climatology Network (U.S. HCN) monthly temperature and precipitation data, carbon dioxide information and analysis center. Oak Ridge National Laboratory, Publ. No. 4500, Oak Ridge, TN, 83 pp.
- Goody, R., J. Anderson, T. R. Karl, R. Miller, G. North, J. Simpson, G. Stephens, and W. Washington, 2002: Why monitor the climate? *Bull. Amer. Meteor. Soc.*, **83**, 873–878.
- Groisman, P. Ya., and D. R. Easterling, 1994: Variability and trends of precipitation and snowfall over the United States and Canada. *J. Climate*, **7**, 184–205.
- Haining, R., 1993: *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press, 432 pp.
- Heim, R. R., 2001: New network to monitor climate change. *Eos, Trans. Amer. Geophys. Union*, **82**, 143.
- Huffman, G. J., R. F. Adler, B. Rudolf, U. Schneider, and P. R. Keehn, 1995: Global precipitation estimates based on a technique for combining satellite-based estimates, rain gauge analysis, and NWP model precipitation information. *J. Climate*, **8**, 1284–1295.
- Janis, M. J., K. G. Hubbard, and K. T. Redmond, 2004: Station density strategy for monitoring long-term climatic change in the contiguous United States. *J. Climate*, **17**, 151–162.
- Karl, T. R., and C. Williams, 1987: An approach to adjusting climatological time series for discontinuous inhomogeneities. *J. Climate Appl. Meteor.*, **26**, 1744–1763.
- , and Coauthors, 1995: Critical issues for long-term climate monitoring. *Climatic Change*, **31**, 185–221.
- Martin, R. J., 1979: A subclass of lattice processes applied to a problem in plane sampling. *Biometrika*, **66**, 336–338.
- Mazzarella, A., and G. Tranfaglia, 2000: Fractal characterization of geophysical measuring networks and its implication for an optimal location of additional stations: An application to a rain-gauge network. *Theor. Appl. Climatol.*, **65**, 157–163.
- Milewska, E., and W. D. Hogg, 2001: Spatial representativeness of a long-term climate network in Canada. *Atmos.–Ocean*, **39**, 145–161.
- National Research Council, 1999: *Adequacy of Climate Observing Systems*. National Academy Press, 51 pp.
- Nicolis, C., 1993: Optimizing the global observational network: A dynamical approach. *J. Appl. Meteor.*, **32**, 1751–1759.
- Owen, T. W., and T. Whitehurst, 2002: United States climate normals for the 1971–2000 period: Product descriptions and applications. *Proc. Third Symp. on Environmental Applications: Facilitating*

- the Use of Environmental Information*, Orlando, FL, Amer. Meteor. Soc., CD-ROM, J4.3.
- Peterson, T. C., H. Daan, and P. D. Jones, 1997: Initial selection of a GCOS surface network. *Bull. Amer. Meteor. Soc.*, **78**, 2145–2152.
- , and Coauthors, 1998: Homogeneity adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol.*, **18**, 1493–1517.
- Ripley, B. D., 1981: *Spatial Statistics*. Wiley, 252 pp.
- Robeson, S. R., 1995: Resampling of network-induced variability in estimates of terrestrial air temperature change. *Climatic Change*, **29**, 213–229.
- , 1997: Spherical methods for spatial interpolation: Review and evaluation. *Cartogr. Geogr. Inf. Syst.*, **24**, 3–20.
- Schmalwieser, A. W., and G. Schaubberger, 2001: A monitoring network for erythemally-effective solar ultraviolet radiation in Austria: Determination of the measuring sites and visualization of the spatial distribution. *Theor. Appl. Climatol.*, **69**, 221–229.
- Vucetic, S., T. Fiez, and Z. Obradovic, 2000: Examination of the influence of data aggregation and sampling density on spatial estimation. *Water Resour. Res.*, **12**, 3721–3731.
- Willmott, C. J., C. M. Rowe, and W. D. Philpot, 1985: Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *Amer. Cartogr.*, **12**, 5–16.