P 1.12

A STATION DENSITY STRATEGY FOR THE UNITED STATES CLIMATE REFERENCE NETWORK (USCRN)

Michael J. Janis *
Southeast Regional Climate Center, SC Department Of Natural Resources, Columbia, SC

Kenneth G. Hubbard
High Plains Regional Climate Center, University Of Nebraska, Lincoln, NE

Kelly T. Redmond
Western Regional Climate Center, Desert Research Institute, Reno, NV

1. INTRODUCTION

A Climate Reference Network (CRN) is being deployed to monitor temporal trends in temperature and precipitation across the US. Initially, a network of about 250 CRN stations was thought to be sufficient for capturing the climatic signal for the contiguous US. This network density was projected from study that estimated a network of 182 stations could reasonably reproduce the 1910-1996 trend in annual precipitation computed from the climate division data set (Karl and Knight 1998). The purpose of the present study is to refine the estimate of the spatial density and total number of stations required over the contiguous US. This study estimates the number of stations that reproduce, within predetermined monitoring goals, annual temperature and precipitation variability.

2. AIR TEMPERATURE AND PRECIPITATION DATA FOR THE UNITED STATES

Temperature and precipitation data are drawn from the 1971-2000 Climatography of the US sequential database (CLIM81, NCDC 2002). CLIM81 contains 5313 stations with monthly average temperature and 7507 stations with monthly total precipitation over the period 1971-2000. CLIM81 metadata are screened for flags associated with estimated or adjusted data. Stations containing more than three years with more than three flags each are culled from analyses for this study. The resulting networks consist of 3642 temperature stations and 5156 precipitation stations. Sequential temperature and precipitation data are transformed into thirty-year annual temperature anomalies and percent-of-median annual total precipitation based on the 1971-2000 reference period.

Instead of performing a global spatial sampling of all stations in CLIM81, these analyses are carried out on stations stratified into 2.5° latitude $\times$ 3.5° longitude grid cells (this grid-cell resolution has been recommended for spatial analysis of US HCN temperature and precipitation data). Of the 115 grid cells covering the contiguous US, only 6% of the grid cells contain fewer than 10 stations. Local within-grid-cell spatial sampling strategies overcome awkward network arrangements that can result from a global random sampling across the entire network.

3. METHODOLOGY

The method described here is a systematic decrease in network resolutions from a baseline network (e.g., CLIM81). Data for a number of stations are selectively removed from the initial full-density network configuration. Each step involves generating measures of similarity between the generated networks of lower spatial resolution and the full-density network. The number of stations that reproduce temperature or precipitation trends in the baseline networks within predetermined climate monitoring goals is an ideal density for climate monitoring. For example, a monitoring system may have the goal: "temperature change for any location within an area can be represented by a single station with an average mean-absolute-error less than 0.1°C per decade."

Ensemble average time series of temperature anomalies and precipitation percent of median from either baseline or "lower resolution" networks are computed within each grid cell. The density can be variable from region to region although the goal is the same from region to region. These sort of local spatial analyses help to determine where higher or lower network densities are needed to satisfy climate-monitoring goals.

Within each grid cell, networks of lower spatial resolution (LSR) are derived by randomly selecting subsets of stations. The number of stations within each subset is incremented by one so that all subset sizes from one to the total number of stations within a grid cell are examined. Average or 'ensemble' time series, derived from combinations of individual station data, are generated for each LSR network. A Monte Carlo resampling procedure is applied to produce 100 different realizations of each LSR network size. This procedure addresses the effect of multiple network configurations, reduces the effect of poorly distributed LSR networks, and reduces the influence of undetected inhomogeneous station records. Similar resampling approaches have been successfully employed to examine the influence of sample size on spatial interpolation of annual total precipitation (Willmott et al. 1996) and monthly average temperature (Robeson and Janis 1998).

* *Corresponding author address:* Dr. Michael J. Janis, Southeast Regional Climate Center, SC Dept. Natural Resources, 2221 Devine St., Suite 222, Columbia, SC 29205. E-mail: janis@dnr.state.sc.us. Tel: 803-734-9569. Fax: 803-734-9573

Grid-cell ensemble trends of temperature and precipitation are computed from annual average temperature and annual total precipitation. Climate monitoring goals for temperature and precipitation trends are defined (in practice these will depend on the threshold of detection desired) and the corresponding station density is identified. A Monte Carlo procedure is applied to each grid cell as follows (von Storch and Zwiers 1999):

1. Randomly sample, without replacement, an N_s-station LSR network ($1 \le N_s \le N-1$, where N is the total number of stations in a grid cell and N_s is the number of stations in the grid cell for the subset).

2. Generate ensemble time series of spatially averaged temperature or precipitation from LSR network. For example, an ensemble time series from the first realization of a randomly drawn without replacement 20-station LSR network is computed as:

$$\hat{T}_1 = \frac{1}{20}\sum_{j=1}^{20} T_j \, ,$$

   where $\hat{T}_1$ is an ensemble time series derived from N_s = 20 randomly selected station time series.

3. Compute linear temperature or precipitation trends for each realization of an N_s-station ensemble time series. Trends are computed over the period of record in the CLIM81 dataset, 1971-2000.

4. Repeat steps 1 through 3, choosing from among all N stations, 100 times to generate multiple realizations of LSR network trends. For any N_s-station LSR network, mean absolute error (MAE_s) for trends is computed from 100 Monte Carlo realizations as:

$$MAE_s = \frac{1}{100}\sum_{k=1}^{100}\left|\frac{\Delta\bar{T}}{\Delta t} - \frac{\Delta\hat{T}_k}{\Delta t}\right| .$$

   where $\Delta\bar{T}/\Delta t$ is the trend for the baseline full-resolution time series and $\Delta\hat{T}_k/\Delta t$ is the trend for k^{th} realization of an N_s-station LSR network.

5. Repeat steps 1 through 4 for all possible LSR network sizes (N_s = 1, 2, …, N-1).

6. Perform polynomial regression between MAE and LSR network size (N_s) by:

$$N_s = a_0 + a_1 MAE_s^1 + \cdots + a_4 MAE_s^4 \, .$$

7. Set monitoring goals (e.g., MAE = 0.1°C/decade) then solve polynomial model for N_s. Inside every grid cell, network densities required to meet the monitoring goals are determined. Solving a polynomial equation for N_s with this goal can be interpreted as identifying the network resolution necessary to reproduce annual air temperature trends from a baseline network to within 0.05°C/decade.

4. RESULTING STATION DENSITIES

Values of N_s for each grid cell corresponding to climate monitoring goals are mapped to identify regions of the country that are most sensitive to network configuration and network density. Grid cells characterized by high spatial variability in temporal variability characteristics will need more stations to estimate trends for that cell to within the desired tolerances. Climate monitoring goals are typically expressed as degrees Celsius per decade (°C/decade) or percent of median annual precipitation per decade. It is important to remember trends are calculated as degrees Celsius per year (or % median precipitation per year) over 30 years of data (i.e., 1971-2000). Although expressed as trends per decade, they are computed over thirty years of data and contain more confidence than trends computed over ten years of data.

4.1 Annual Air Temperature Trends

Four temperature-trend monitoring goals are examined and evaluated relative to a simple one-station-per-grid-cell network (Table 1). A national network meeting a monitoring goal of 0.05°C/decade consists of 622 stations with an average of 5.5 stations per grid cell and an average station separation of 149 km. The western US (west of the 100^{th} Meridian) requires 310 stations or an average of 6.2 stations per grid cell (Fig. 1). The eastern US requires 311 stations or 4.9 stations per grid cell. Approximately 55% of grid cells require 5 or more stations per cell to meet this monitoring goal, including three grid cells that require ten or more stations. Approximately 60% of grid cells require a 112-162 km spatial separation between stations to meet this monitoring goal.

Table 1 Integrated number of stations required to meet various temperature monitoring goals.

Temperature Trend (°C/decade)	Number of Stations
0.050	622
0.075	338
0.100	233
0.125	167
1 station per grid cell	114

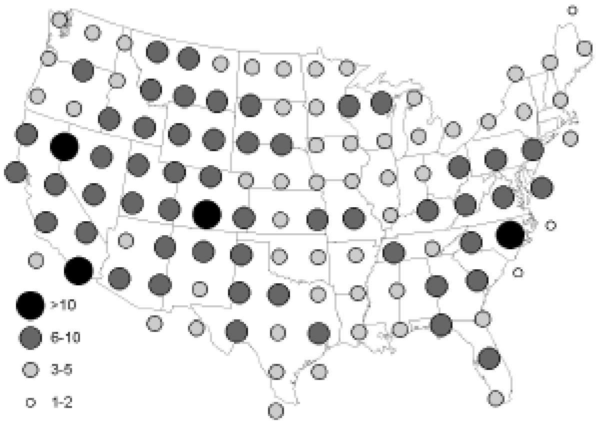

Figure 1 Grid-cell densities of 622 stations satisfying an annual temperature-trend monitoring goal of MAE < 0.05°C per decade. Graduated circles represent the number of stations within each grid cell.

A national network meeting a monitoring goal of 0.10°C/decade consists of 233 stations with an average station separation of 242 km. With nearly 100% of all cells requiring fewer than 4 stations to meet this monitoring goal, little spatial variability in network density results (Fig. 2). The western and eastern US require an average of 2.1 and 2.0 stations per grid cell, respectively, to meet a 0.10°C/decade monitoring goal.

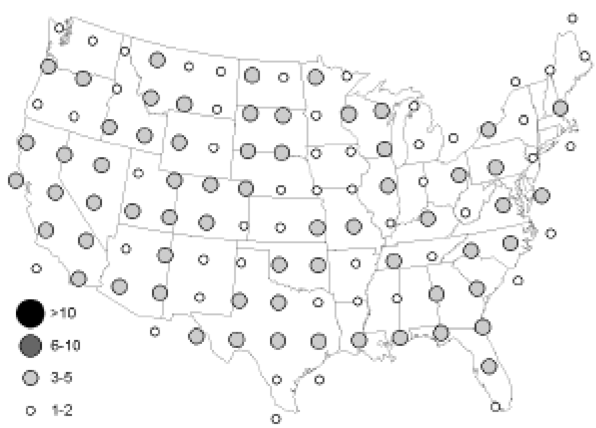

Figure 2 Grid-cell densities 233 stations satisfying an annual temperature-trend monitoring goal of MAE < 0.10°C per decade.

4.2 Annual Precipitation Trends

Four precipitation-trend monitoring goals are examined (Table 2). Precipitation trend monitoring goals are based on observed annual trends less than ±5% per decade across the contiguous US. A national network meeting a monitoring goal of 1.5% of median annual precipitation per decade consists of 490 stations with an average of 4.3 stations per grid cell. Approximately 45% of grid cells require a 137-187 km (173 km average) spatial separation between stations to meet this monitoring goal. Higher station densities are generally required west of the 100th Meridian, with the highest station densities occurring in the southwestern US (Fig. 3). The western US requires 275 stations or an average of 5.4 stations per grid cell while the eastern US requires 215 stations or an average of 3.4 stations per grid cell.

Table 2 Integrated number of stations required to meet various precipitation monitoring goals.

Precipitation Trend (% median/decade)	Number of Stations
1.0	900
1.5	490
2.0	293
2.5	218
1 station per grid cell	115

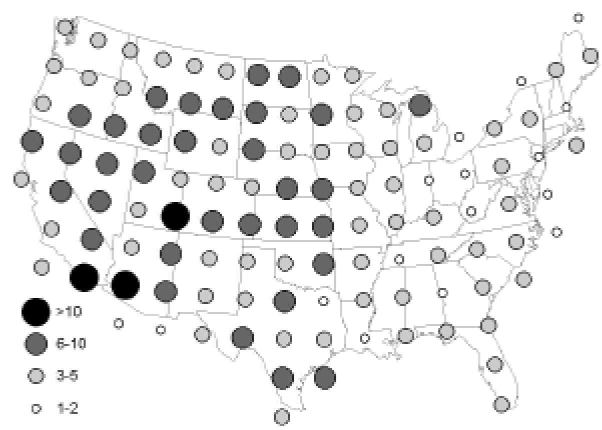

Figure 3 Grid-cell densities of 490-station LSR network satisfying an annual precipitation-trend monitoring goal of MAE < 1.5% median precipitation per decade.

A national network meeting a monitoring goal of 2.0% of median annual precipitation per decade consists of 293 stations with an average of 2.5 stations per grid cell and an average station separation of 225 km (Table 2). Thirty percent of grid cells require only one station to meet this monitoring goal (Fig. 4). The most common spatial separation per grid cell is 162-212 km, but higher resolutions (e.g., >300 km) are necessary for some grid cells. To meet this monitoring goal, higher network densities are necessary in the western US (168 stations or an average of 3.3 stations per grid cell), while lower network densities are necessary in the eastern US (125 stations or an average of 1.9 stations per grid).

Figure 4 Grid-cell densities of 293-station LSR network satisfying an annual precipitation-trend monitoring goal of MAE < 2.0% median precipitation per decade.

5. SUMMARY AND DISCUSSION

A goal of this work is to provide a recommendation for US CRN station density. Spatial density requirements are examined relative to measured precipitation and temperature from a spatially denser network of existing weather stations. A stratified local sampling strategy with Monte Carlo resampling techniques applied within each grid cell is used to build information on how trend estimates may diverge with decreasing network density. Regions of the country that require higher station densities to meet climate-monitoring goals are identified.

Network densities resulting from 0.10°C per decade and 2.0% of median annual precipitation per decade monitoring goals are superimposed. The larger number of stations within each grid cell for either monitoring goal determines the combined network density. The result provides a minimum density estimate for national climatic change networks (Fig. 5). A national network meeting the combined monitoring goals consists of 327 stations with an average of 2.9 stations per grid cell. The western US (west of the 100th Meridian) requires 168 stations or an average of 3.4 stations per grid cell while the eastern US requires 159 stations or 2.5 stations per grid cell. Comparing the combined network density map (Fig. 5) to the maps for individual monitoring goals (Figs. 2 and 4) shows that precipitation monitoring drives higher densities in the west. Densities in the east are slightly higher under temperature monitoring goals.

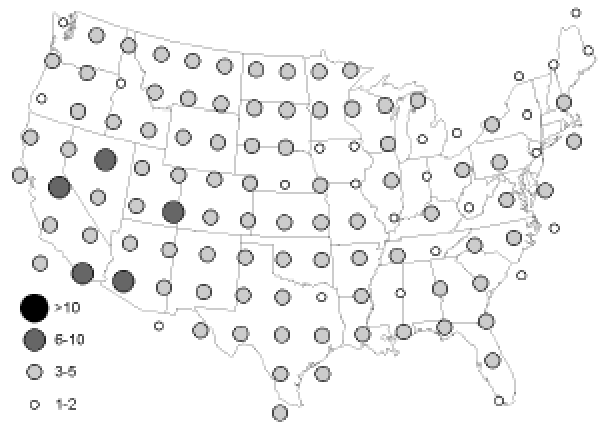

Figure 5 Grid-cell densities of combined 327 stations satisfying an annual temperature-trend monitoring goal of MAE < 0.10°C per decade and a precipitation-trend monitoring goal of 2% median annual precipitation per decade.

Acknowledgements. This work was partially supported under the NOAA Cooperative Agreement NA17RJ1222.

6. REFERENCES

Karl, T. R., and R. W. Knight, 1998: Secular trends of precipitation amount, frequency, and intensity in the United States. *Bull. Amer. Meteor. Soc.,* **79**, 231–241.

NCDC, 2002: Climatography of the U.S. No. 81: Monthly Station Normals of temperature, Precipitation, and Heating and Cooling Degree Days 1971-2000. National Climatic Data Center, CD-ROM. [Available online at http://nndc.noaa.gov/onlinestore.html.]

Robeson, S. M., and M. J. Janis, 1998: Comparison of temporal and unresolved spatial variability in multiyear time-averages of air temperature. *Climate Res.,* **10**, 15–26.

von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge, UK, 484 pp.

Willmott, C. J., S. M. Robeson, and M. J. Janis, 1996: Comparison of approaches for estimating time-averaged precipitation using data from the USA. *Int. J. Climatol.,* **16**, 1103–1115.