

# ESTIMATING CONTINENTAL AND TERRESTRIAL PRECIPITATION AVERAGES FROM RAIN-GAUGE NETWORKS

CORT J. WILLMOTT

*Center for Climatic Research, Department of Geography, University of Delaware, Newark, Delaware 19716, USA*

SCOTT M. ROBESON

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

AND

JOHANNES J. FEDDEMA

*Department of Geography, University of California, Los Angeles, California 90024, USA*

*Received 13 November 1992*

*Accepted 14 July 1993*

## ABSTRACT

Influences of varying rain-gauge networks on continental and terrestrial precipitation averages (derived from data observed on those networks) are evaluated. Unsystematically and systematically designed station networks are considered, the latter being represented by the NCAR World Monthly Surface Station Climatology, which contains hand-picked but time-varying networks that date back to the 1800s. Biases arising from spatially uneven and temporally variable precipitation-observing networks can be significant.

For all the continents, except South America, sparse rain-gauge networks produce overestimates of continental mean precipitation. Mean precipitation for South America, in contrast, is underestimated substantially by low densities of observing stations. Sampling errors tend to be large in areas of high precipitation and in regions with strong spatial precipitation gradients (e.g. in the Sahel). These patterns occur whether the station network has been selected systematically (as in the NCAR network) or unsystematically.

Systematic sampling of mean precipitation (at the NCAR station locations), however, suggests that many yearly NCAR station networks are adequate for estimating continental average precipitation. As early as 1890, NCAR networks for Australia resolve continental average precipitation accurately. Not until 1960, however, do NCAR networks for South America begin to resolve continental mean precipitation adequately. Regional and continental NCAR network errors also tend to cancel one another, often giving accurate yearly estimates of terrestrial mean precipitation.

**KEY WORDS** Precipitation averages Rain-gauge networks

## INTRODUCTION

Continental and terrestrial averages of precipitation are key components of the global hydrological cycle, yet completely reliable estimates of these spatial means are not available (Willmott and Legates, 1991). Even less is known about the seasonal and interannual variability in continental and terrestrial precipitation averages. Although global climate model (GCM) simulations and remotely sensed estimates of large-scale precipitation are improving (Legates and Willmott, 1992), historical rain-gauge records and networks continue to comprise the bases for the most credible estimates (Legates and Willmott, 1990; Hulme, 1992). Large-scale spatial averages made from historical data, in other words, comprise our best understanding of the spatial, seasonal and interannual variability in continental and terrestrial averages of precipitation. The reliability and variability of continental and terrestrial precipitation averages that have been derived from the historical rain-gauge record are the subject of this paper.

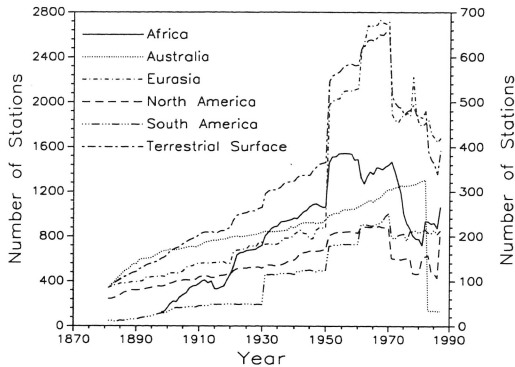


Figure 1. Number of stations per year in the NCAR World Monthly Surface Station Climatology (Spangler and Jenne, 1988) for the terrestrial surface (left axis) and for each continent (right axis) from 1881 to 1987.

Of all the climate variables, precipitation has been sampled most extensively (Mintz, 1981; Legates and Willmott, 1990). Although more station records are available for precipitation than for air temperature or pressure, it also is true that spatial variability of precipitation is much greater. Higher resolution station networks, therefore, are required to sample precipitation adequately. Rain-gauge samples of precipitation, however, have been spatially uneven and historically variable. The NCAR World Monthly Surface Station Climatology (Spangler and Jenne, 1988), for instance, contains 636 stations for 1900 whereas 2327 stations are available for 1960 (Figures 1 and 2). These changing station networks have injected considerable uncertainty into estimates of terrestrial mean precipitation (Willmott and Legates, 1991; Willmott *et al.*, 1991). Averaging over the entire terrestrial surface has reduced the influence of regional- and continental scale under- and overestimates. Terrestrial averages, consequently, are more accurate than continental and regional averages.

Our purpose here is to determine the variability and reliability of spatially averaged estimates of precipitation. Station-network influences are examined at continental and terrestrial scales through computer-intensive sampling from a high-resolution climatology (Legates and Willmott, 1990). The adequacy of numerous network densities and station distributions is described. In addition, the adequacy of the NCAR World Monthly Surface Station Climatology is investigated by resampling the Legates and Willmott high-resolution climatology at continental scales.

### SPATIAL SAMPLING, INTERPOLATION, AND AVERAGING

Spatial averages derived from any irregularly distributed station network contain biases. Bias may arise from the spatial interpolation used to estimate how the variable behaves between observations. Errors additionally may arise from the influence of ill-conditioned station networks; that is, errors result when intrinsic spatial variability is not resolved adequately. It usually is assumed that the variable is spatially continuous and that linear combinations of the observations can be used to estimate values at unsampled locations (see Lam, 1983; Thiébaux and Pedder, 1987; Daley, 1991). Since time-averaged precipitation is highly variable in space and occasionally discontinuous, sparse station networks and inadequate interpolation algorithms can combine to produce non-trivial averaging errors. Spatial averaging typically involves two steps: the interpolation from an irregularly spaced network of station values to a regular grid and then

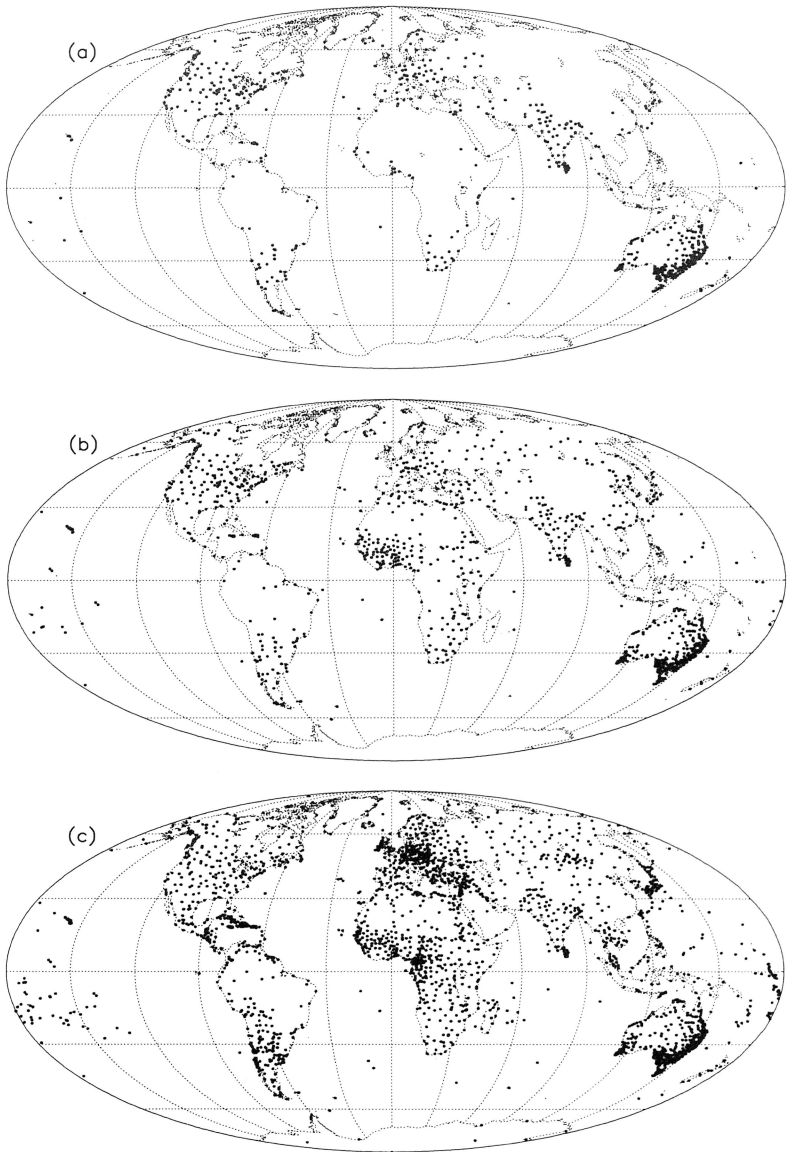


Figure 2. Spatial distribution of precipitation stations in the NCAR World Surface Station Climatology for three selected years: (a) 1900, (b) 1930, and (c) 1960.

the weighted (by area) summing of the gridded estimates. Our analyses also make use of this approach which is discussed in more detail by Willmott *et al.* (1985a) and by Willmott and Legates (1991).\*

Although a variety of spatial interpolation methods are available, when interpolating over large sections of the Earth's surface, spherical (or even ellipsoidal) geometry should be used (Willmott *et al.*, 1985a). Several spherically based interpolation procedures have been implemented (e.g. Wahba, 1981; Renka, 1984; Willmott *et al.*, 1985a); however, inverse distance-weighting algorithms have been used most frequently in precipitation studies (Bradley *et al.*, 1987; Diaz *et al.*, 1989) and several are quite reliable (Legates, 1987; Bussi eres and Hogg, 1989; Weber and Englund, 1992). One such procedure, a spherical version of Shepard's (1968) inverse-distance weighting method (Willmott *et al.*, 1985a), is used here to interpolate precipitation data to continental and terrestrial grids. The relative reliability of this interpolation method allows us to separate out and evaluate errors arising from ill-conditioned station networks—the theme of this paper. Interpolations are made to a spherical grid of 1° of latitude by 1° of longitude. This lattice was chosen as a compromise between computational efficiency and accurate spatial depiction. Once precipitation data are interpolated, cosine-of-latitude weighting is used to integrate precipitation over the continental and terrestrial domains.

### CONTINENTAL AND TERRESTRIAL PRECIPITATION AVERAGES

Influences of precipitation measurement networks on spatial averaging may be analysed at a variety of scales. Our interest here, however, is on large-scale averaging and therefore we restrict our analyses to precipitation networks at continental and terrestrial scales. Oceanic precipitation is not considered because both the *in situ* measurements and observational networks are much less reliable than their terrestrial counterparts (Legates and Willmott, 1990).

Two sampling strategies are used to assess precipitation network biases. In the first instance, subnetworks for six characteristic station densities are unsystematically (randomly) and repeatedly sampled from the Legates and Willmott (1990) high-resolution climatology (hereafter referred to as the LW climatology). This procedure essentially follows a similar experiment by Willmott and Legates (1991). Illustrating for the lowest density network (of the six), 50 subnetworks (each at a station density of 2.5 stations per 10<sup>6</sup> km<sup>2</sup>) were selected randomly (with replacement) from the LW climatology. Legates and Willmott's annual mean precipitation estimates at the station locations within each sample were then interpolated to the nodes of a 1° by 1° grid. The gridded values, in turn, were weighted by cosine-of-latitude and summed to obtain each of the 50 spatial averages. Variability among the 50 spatial means is an indication of the natural variability induced by rain-gauge networks at this station density. Differences between each of the 50 sample means and the mean computed from all station values in the LW climatology are estimates of the accuracy of the 2.5 network density. This same experiment was repeated for subnetworks of 5, 10, 15, 20, and 25 stations per 10<sup>6</sup> km<sup>2</sup>. These densities are typical of twentieth-century precipitation station networks.

The second strategy involved systematically sampling the LW climatology at all the station locations in the NCAR climatology for each year since 1881. Values of annual mean precipitation from the LW climatology at these nodes were then interpolated to the 1° by 1° grid and averaged (as above) to obtain an estimate of annual mean precipitation associated with each year's NCAR network. The difference between a spatial mean derived from a yearly NCAR network and a corresponding mean obtained from all the stations in the LW climatology is one measure of the efficacy of that year's NCAR network.

\*Several promising interpolation methodologies (e.g., that make use of anomaly or frequency-distribution parameter transformations of the raw precipitation data) are not used or evaluated in this paper because they require one or more additional realizations of the field of interest (often at a high spatial resolution). Applications—such as ours—also require estimates of precipitation magnitudes rather than of derived statistics, e.g., anomalies. At the heart of our analyses is interpolation from a single realization of a field on a single station network. More precisely, we focus on evaluating (i) how well individual station networks adequately resolve a mean precipitation field over an extensive spatial domain and (ii) how accurately large-scale spatial averages can be constructed from observations on a single station network. This involves the spatial interpolation of precipitation averages from an irregularly spaced station network to a regular grid—the most common application of interpolation in climatology—and, in turn, the numerical integration of the gridded precipitation field.

Precipitation station networks from five continents—Africa, Australia, Eurasia, North America, and South America—and from the entire terrestrial surface (excluding Greenland and Antarctica, where few stations were available) are evaluated separately. Both the unsystematic and systematic experiments were performed separately for each continent as well as for the terrestrial surface. Errors in estimating continental and terrestrial mean precipitation are analysed statistically and graphically through the use of box plots and maps of sampling errors.

### *Africa*

Precipitation station distributions within Africa are sparse, relative to well-represented areas such as Europe and the USA (see Figure 3(a)). Highest station densities are found in relatively moist areas along the Mediterranean coasts of Morocco and Algeria, in southern Africa, coastal western Africa, and the east

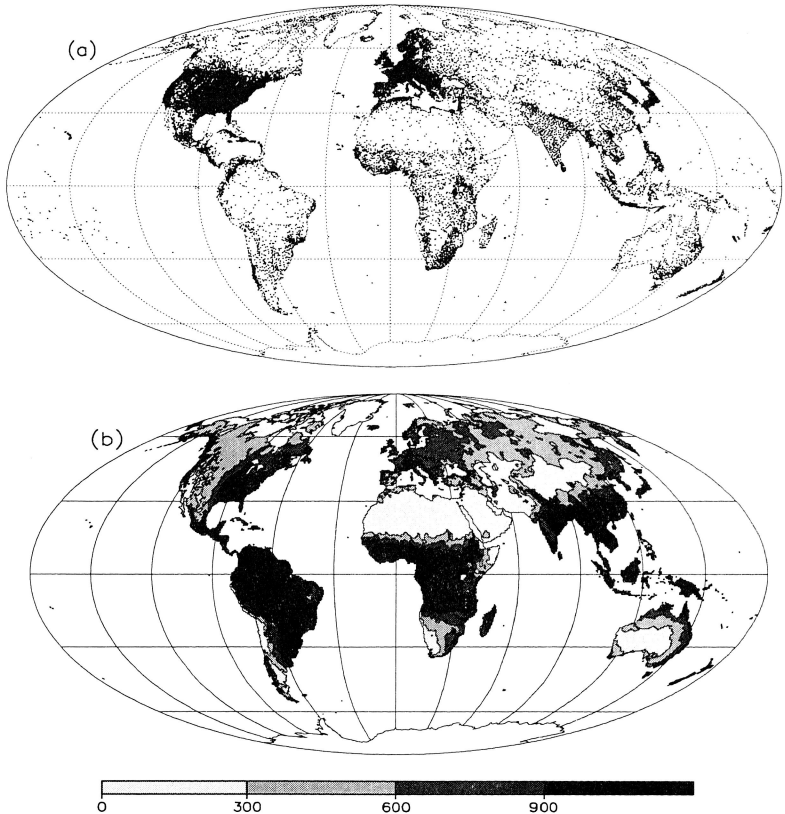


Figure 3. Spatial distribution of (a) the over 24 000 precipitation stations and (b) annual mean precipitation (mm) from the climatology of Legates and Willmott (1990).

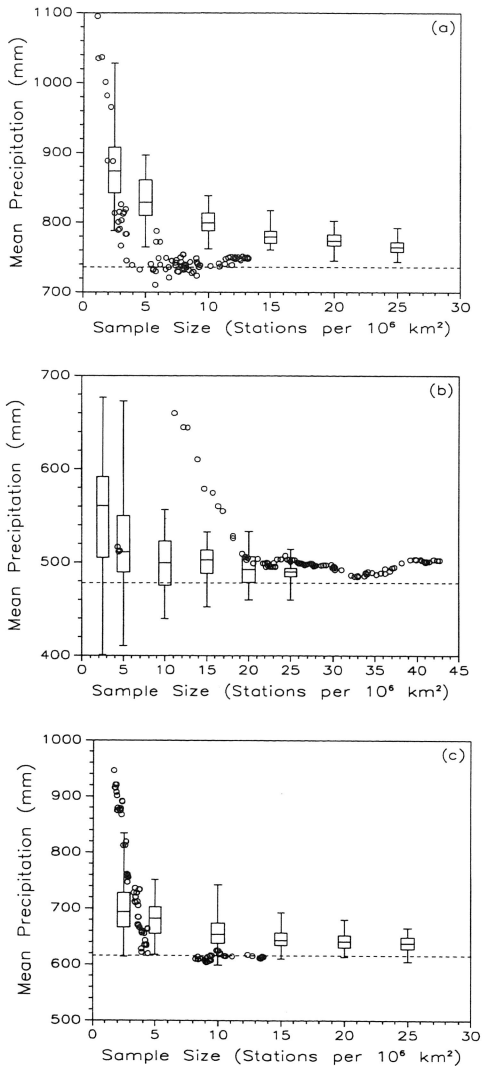


Figure 4. Annual mean precipitation derived from resampling the LW high-resolution climatology using (i) unsystematic (random) sampling (box plots) and (ii) systematic sampling at the yearly NCAR station networks (open circles) for (a) Africa, (b) Australia, (c) Eurasia, (d) North America, (e) South America, and (f) the terrestrial surface, excluding Greenland and Antarctica. Each box plot—showing minimum, maximum, lower and upper quartiles, and median—is derived from 50 random samples of the climatology. The dashed line is the value estimated from all the stations in the LW climatology.

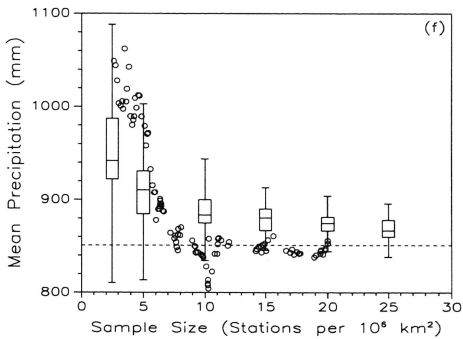
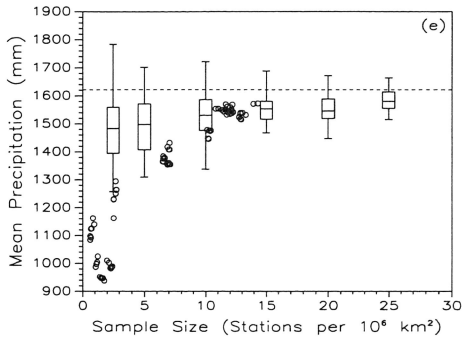
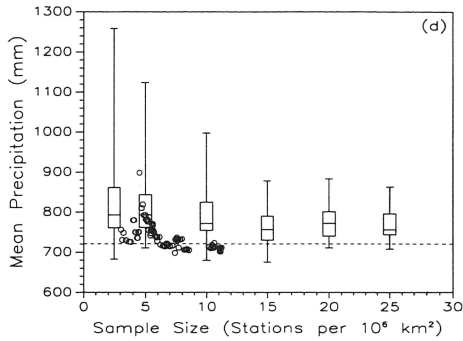


Figure 4. (Continued)

Table I. Number of stations in each continent corresponding to sampling densities (number of stations per  $10^6$  km<sup>2</sup>)

Sampling density	Africa	Australia	Eurasia	North America	South America	Terrestrial surface
2.5	73	19	125	50	44	328
5.0	146	38	250	99	88	656
10.0	292	76	500	198	175	1311
15.0	438	115	749	297	263	1967
20.0	584	153	999	396	351	2622
25.0	729	191	1249	496	439	3278

African highlands (Kenya and Uganda). Sampling problems are most prevalent in Africa's desert regions (e.g. interior Sahara, Kalahari, and in Somalia). To a lesser extent, the upper reaches of the Zaire basin also are under-represented. Overall, Africa's dry regions are undersampled.

Random sampling (with subsequent interpolation and averaging) from the African station network (Figure 3(a)), therefore, results in a wet-region bias at all sampling densities (Figures 3(b) and 4(a)—box plots). Boundaries between moist and arid areas (e.g. in the Sahel) also are misplaced (discussed below). Moist areas typically are interpolated to be too far into dry regions. For the lowest station density considered (2.5 stations per  $10^6$  km<sup>2</sup> or 73 stations for all of continental Africa; Table I), the median of the 50 random samples is 140 mm higher than the continental mean precipitation estimated from all stations in the LW climatology (the dashed line in Figure 4(a)). As the station density increases, sample-derived means asymptotically decrease and approach (but never reach) the LW climatology mean. Consider also that the LW climatology itself may overestimate continental mean precipitation because the large expanse of the Sahara Desert is sampled sparsely.

Most yearly NCAR networks, by contrast, appear to estimate African mean precipitation well (Figure 4(a)—circles), for all but the sparsest station networks (i.e. prior to 1920). Integrations using NCAR networks prior to 1900 are not shown owing to small numbers of stations (<30) in the 1800s. The NCAR station networks after 1920, however, perform better than random sampling. This is not surprising because the NCAR networks were chosen to have spatially uniform station distributions.

### *Australia*

Australia's relatively small size, low relief, and low annual mean precipitation result in less spatial variability in precipitation than occurs over any other continent (Figure 3(b)). As in Africa, areas with high station densities (Figure 3(a)) are found in relatively moist regions—along the south-east and south coasts of Australia. The wet northern coast and arid interior are sampled sparsely, with the dry interior representing a much larger land area. Yearly NCAR station networks for Australia also are not distributed evenly. The NCAR station densities in Australia, nevertheless, are much higher than for the other continents, especially in south-eastern Australia (Figure 2).

Random sampling at low station densities produces a large range of estimates for Australian mean precipitation (Figure 4(b)). Much of this variability is controlled by whether or not stations from the very wet regions in the north are sampled. When wet tropical regions are sampled, the estimated moist region extends too far into the dry interior where few stations exist. Random sampling of stations, none the less, can lead to either over- or underestimates of Australia's mean precipitation. Overestimates, however, are much more common.

The NCAR station networks represent Australian mean precipitation adequately after 1890, although estimates are always a little high (Figure 4(b)). The interior of the continent, however, is undersampled in both the LW climatology and in the NCAR networks. NCAR station densities along the Australian



coast—particularly in the south-east—are higher than in any other region in the world (well over 25 stations per  $10^6$  km<sup>2</sup> for most years). The relative wetness of this region contributes to the slight but consistent overprediction of the continental average.

### *Eurasia*

Station networks in parts of Eurasia are extremely dense, especially in western Europe, with a gradual decline towards the east (Figure 3(a)). The continental interior (particularly Siberia, Mongolia, and the Gobi Desert), the Arabian peninsula, and the northern coastal areas are not as well represented in the LW climatology.

As in Africa and Australia, sampling problems in Eurasia occur along high precipitation gradients, with moist areas being better represented than arid regions. Once again, sparse station networks give rise to overestimates of the continental mean (Figure 4(c)). A variety of factors contribute to the sampling problems along strong spatial gradients (e.g. the influence of orography, continentality, and the spatial extent of the south-east Asian monsoon). Nevertheless, station densities are sufficient to obtain adequate (albeit slightly high) estimates of Eurasian mean precipitation (Figure 4(c)). Even at moderate station densities (e.g. 10 stations per  $10^6$  km<sup>2</sup> or *ca.* 500 stations; Figure 4(c)), accurate means can be computed. Sparse NCAR networks (prior to 1950) similarly produce high estimates whereas moderate-resolution NCAR networks yield quite accurate estimates (Figure 4(c)).

### *North America*

Areas receiving relatively high amounts of precipitation within North America (e.g. central Mexico, the eastern USA, and southern Canada) contain dense station networks. Northern Canada, interior Alaska, and the northern deserts of Mexico (areas of relatively low precipitation) conversely contain rather sparse station distributions. Unsystematic sampling at stations from the LW climatology, therefore, leads to overestimation of North American mean precipitation (Figure 4(d)). Large variability in the estimated means occurs at low station densities because of the tremendous regional variability in North American precipitation (Figure 3(b)). Estimates of continental average precipitation vary by nearly 600 mm at the lowest station density (Figure 4(d)). Even the largest sample size (25 stations per  $10^6$  km<sup>2</sup>) produces a range greater than 150 mm. Interquartile ranges are much smaller. Yearly NCAR station networks produce reasonable estimates for most station densities, particularly after 1920.

### *South America*

Estimating average precipitation for South America is quite different from any of the other continents. Historical rain-gauge networks are unusually sparse in South America and the most poorly sampled areas are wet (Figure 3). North-western and southern sections of South America are better sampled than the interior, although the sampling is dense almost nowhere. Unlike virtually all other continental interiors, the interior of South America—particularly the Amazon basin—has very few rain-gauges and receives large amounts of rainfall. Tropical Africa is quite wet but relatively well-sampled.

Sparse sampling of large areas with high precipitation produces continental mean precipitation estimates that are both highly variable and grossly underestimated for all sampling densities (Figure 4(e)). At low sampling densities, underestimates as large as 200 mm are common. Even at densities of 20 stations per  $10^6$  km<sup>2</sup>, estimates of South American mean precipitation can be in error by more than 100 mm.

Sampling of South American precipitation is especially poor within the NCAR climatology (Figures 1 and 2). Underestimates range from nearly 700 mm to 100 mm (Figure 4(e)). Station densities within the NCAR archive never exceed 15 stations per  $10^6$  km<sup>2</sup>.

### Terrestrial surface

Station distributions for most parts of the terrestrial surface are described above. Oceania and several island nations, however, also are included in our analysis of the terrestrial surface. Although Oceania and island nations represent a relatively small proportion of the total land area, a few islands do contain a large number of stations. Such densely sampled areas include Japan, the Philippines, Indonesia, and Great Britain and they can exert a greater influence on terrestrially averaged precipitation estimates than is warranted by their size.

Random sampling of stations from the LW climatology (Figure 3(a)) gives rise to large variability in estimates of terrestrial mean precipitation until sampling densities of 15 stations per  $10^6$  km<sup>2</sup> are reached (Figure 4(f)). Random sampling also does considerably better than the yearly NCAR networks at low sampling densities. The NCAR networks appear to give reliable estimates when station densities exceed *ca.* 7 stations per  $10^6$  km<sup>2</sup>. Previously described continental-scale findings suggest that large underestimates in South American precipitation tend to cancel overestimates from other continents.

Additional understanding of how uneven and variable station distributions influence regional precipitation estimates can be obtained by examining error maps. Error statistics can be generated by subtracting grid-point values interpolated from the 50 random samples of the LW climatology from corresponding grid-point values obtained from all the stations in the LW climatology. More specifically, the mean absolute error (MAE) at a grid-point is

$$\text{MAE} = N^{-1} \sum_{i=1}^N |\hat{P}_i - P| \quad (1)$$

where  $P$  is the grid-point value associated with the entire LW climatology,  $\hat{P}_i$  is a grid-point value interpolated from one of the 50 random samples of the LW climatology, and  $N = 50$  (Willmott *et al.*, 1985b). Error maps are shown for the two extreme cases of 2.5 and 25 stations per  $10^6$  km<sup>2</sup> (Figure 5).

Estimates of MAE are large in areas of relatively high precipitation and in undersampled regions, as is evident in maps of both 2.5 and 25 stations per  $10^6$  km<sup>2</sup> (Figure 5). Estimates of absolute precipitation error are roughly correlated with mean precipitation because precipitation has a lower bound (zero) but no upper bound. Central tendency and variance, therefore, tend to covary. Large zones within the lower latitudes have sampling errors in excess of 400 mm, especially in South America. A number of sparsely sampled mountainous and arid areas also exhibit large precipitation errors. Desert regions sometimes have low MAEs, simply because they have low precipitation variability.

Large errors (i) occur where steep precipitation gradients are coincident with steep station-density gradients and (ii) result from the extrapolation of the relatively well-sampled region's precipitation characteristics into the adjacent, sparsely gauged region. The extent to which the higher-resolution samples (25 stations per  $10^6$  km<sup>2</sup>) better resolve the spatial variability in these transition zones is evident in the improvements in Figure 5(b) relative to Figure 5(a).

### SUMMARY AND CONCLUSIONS

Precipitation averages and spatial variability of precipitation have been investigated at continental and terrestrial scales. Effects of uneven and variable precipitation networks were investigated by sampling from a high-resolution climatology. Both randomly selected station networks and stations within the NCAR World Monthly Surface Station Climatology were evaluated. Repeated random sampling (at a variety of station densities) revealed how continental- and terrestrial-scale estimates of mean precipitation can be affected by irregular station networks. By sampling from the NCAR station network, the adequacy of systematically selected station networks also was investigated.

For all the continents except South America, low rain-gauge densities give rise to overestimates of annual mean precipitation. Annual mean precipitation in South America, in contrast, is vastly underestimated by low station densities. Maps of the mean absolute error also show that sampling errors tend to be high in areas of high precipitation, and in areas with strong precipitation gradients and station-density gradients.

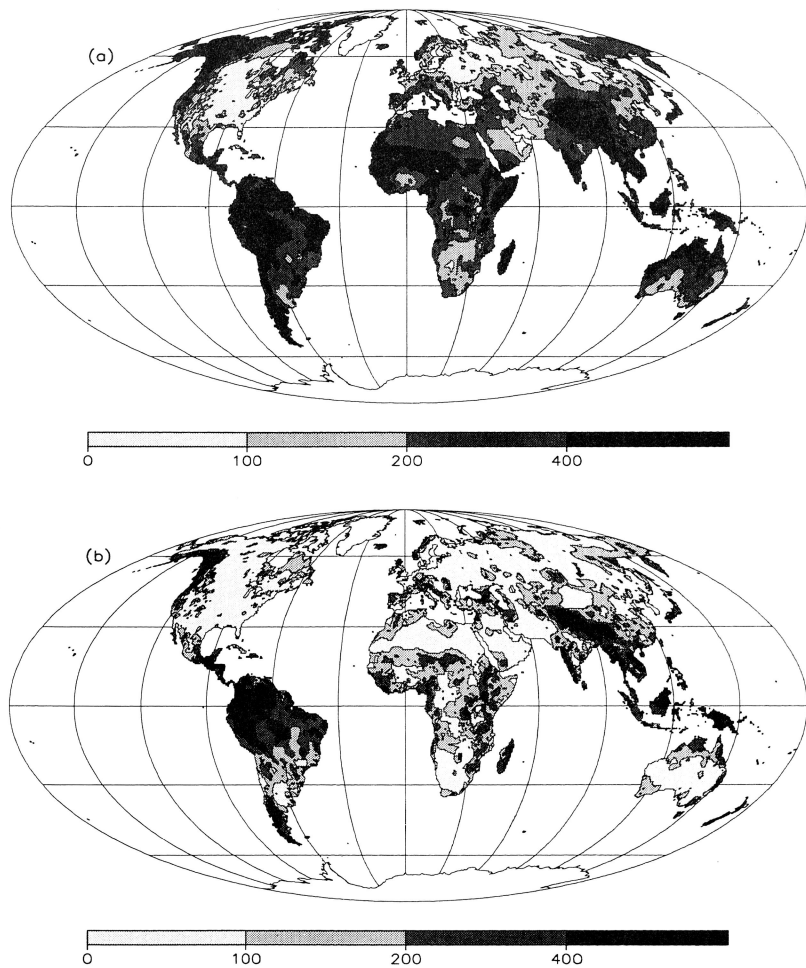


Figure 5. Mean absolute errors (MAEs) for 50 random samples of the LW high-resolution climatology for station densities of (a) 2.5 stations per  $10^6 \text{ km}^2$  and (b) 25 stations per  $10^6 \text{ km}^2$

Sampling the LW climatology at the NCAR station locations suggests that many yearly NCAR station densities (particularly the more recent ones) are adequate for estimating continental mean precipitation. Regional- and continental-scale errors, however, often cancel one another to give what appears to be reliable estimates of terrestrial average precipitation as early as 1920.

Our results also imply that future rain-gauge deployments—particularly in currently undersampled regions of the world—should carefully take into account the spatial and temporal variability of precipitation within the domain of interest. Higher resolution deployments should coincide with high-frequency (in the spatial domain) precipitation variability. Spatial shifts in the precipitation field with time (on seasonal and interannual time-scales, for example) also should be considered in the station-network design. Samples of a time-varying field obtained from a fixed network (when the network systematically underrepresents portions of the field for periods of time) can produce temporally varying biases in the estimates of the field.

#### ACKNOWLEDGEMENTS

Portions of this research were supported by NASA (grants NAG5-853 and NAGW-1884) and UCLA (Faculty Senate Grant 4-563867-1990-07).

#### REFERENCES

- Bradley, R. S., Diaz, H. F., Eischeid, J. K., Jones, P. D., Kelly, P. M. and Goodess, C. M. 1987. 'Precipitation fluctuations over Northern Hemisphere land areas since the mid-19th century', *Science*, **237**, 171–175.
- Bussi eres, N. and 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, New York.
- Diaz, H. F., Bradley, R. S. and Eischeid, J. K. 1989. 'Precipitation fluctuations over global land areas since the late 1800s', *J. Geophys. Res.*, **94**, 1195–1210.
- Hulme, M. 1992. 'A 1951–1980 global land precipitation climatology for the evaluation of general circulation models', *Climate Dynamics*, **7**, 57–72.
- Lam, N. S-N. 1983. 'Spatial interpolation methods: a review', *Am. Cartogr.*, **10**, 129–149.
- Legates, D. R. (1987) 'A climatology of global precipitation', *Pub. Climatol.*, **40**(1), 85 pp.
- Legates, D. R. and Willmott, C. J. 1990. 'Mean seasonal and spatial variability in gauge-corrected, global precipitation', *Int. J. Climatol.*, **10**, 111–127.
- Legates, D. R. and Willmott, C. J. 1992. 'A comparison of GCM-simulated and observed mean January and July precipitation', *Global Planet. Change*, **97**, 345–363.
- Mintz, Y. 1981. 'A brief review of the present status of global precipitation estimates', in Atlas, D. and Theile, O. W. (eds), *Report of the Workshop on Precipitation Measurements from Space*, NASA Goddard Space Flight Center, Greenbelt, MD.
- Renka, R. J. 1984. 'Interpolation of data on the surface of a sphere', *Association For Computing Machinery Trans. Math. Software*, **10**(4), 417–436.
- Shepard, D. 1968. 'A two-dimensional interpolation function for irregularly spaced data', *Proceedings of the 23rd National Conference, Association For Computing Machinery*, pp. 517–523.
- Spangler, W. M. L. and Jenne, R. L. 1988. *World Monthly Surface Station Climatology*, National Center for Atmospheric Research, Scientific Computing Division, Boulder, Colorado.
- Thi ebaux, H. J. and Pedder, M. A. 1987. *Spatial Objective Analysis*, Academic Press, New York.
- Wahba, G. 1981. 'Spline interpolation and smoothing on the sphere', *SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comp.*, **2**(1), 5–16.
- Weber, D. and Englund, E. 1992. 'Evaluation and comparison of spatial interpolators', *Math. Geol.*, **24**, 381–391.
- Willmott, C. J., Rowe, C. M. and Philpot, W. D. 1985a. 'Small-scale climate maps: a sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring', *Am. Cartogr.*, **12**, 5–16.
- Willmott, C. J., Ackleson, S. G., Davis, R. E., Feddema, J. J., Klink, K. M., Legates, D. R., O'Donnell, J. and Rowe, C. M. 1985b. 'Statistics for the evaluation and comparison of models', *J. Geophys. Res.*, **90**(C5), 8995–9005.
- Willmott, C. J., Robeson, S. M. and Feddema, J. J. 1991. 'Influence of spatially variable instrument networks on climatic averages', *Geophys. Res. Lett.*, **18**(12), 2249–2251.
- Willmott, C. J. and Legates, D. R. 1991. 'Rising estimates of terrestrial and global precipitation', *Climate Res.*, **1**, 179–186.