Transactions of the Illinois State Academy of Science
(1990), Volume 83, 1 and 2, pp. 70-81

Use of Cooperative Weather Station Data in Contemporary Climate Research

Peter J. Lamb and Michael B. Richman
Climate and Meteorology Section
Illinois State Water Survey
Champaign, Illinois 61820

ABSTRACT

Daily precipitation observations made at thousands of cooperative weather stations in the central U.S. and southern Canada during the May-Augusts of 1949-80 are distilled into a set for a grid-like network of 557 locations which is used to delineate areas of spatial precipitation coherence for a range of time scales and periods. This delineation is achieved by applying a variety of eigenvectorial techniques and correlation analysis to network station totals for 3, 7, 15, 30, 41, and 61 day periods. All analyses are performed for May-August as a whole; those for 3-days are also conducted on a half-monthly (early May, late May,...., late August) basis. The resulting regionalizations of the vast study domain into smaller areas of precipitation coherence are compared and their implications and applications identified. This illustrates the great value of the voluntary observers who operate the cooperative weather stations.

INTRODUCTION

The two preceding papers in this series outlined the history of cooperative weather station observations in Illinois (Wendland, 1990) and used data from some of those stations to document temporal changes in the climate of both Urbana-Champaign and the state as a whole (Changnon, 1990). Their focus was thus on the *time* dimension. This final paper in the series broadens the discussion by strongly emphasizing the *space* dimension. We show here how daily precipitation observations made at thousands of cooperative weather stations in the central U. S. and southern Canada have been used to identify important spatial patterns of summer precipitation variation for that vast area. The research employed modern advanced statistical techniques and yielded results that have diverse implications and applications for predictive crop yield modeling, mesometeorological inquiry, 48 - 96 hour precipitation prediction, studies of climate variation and change, precipitation chemistry calculations, and perhaps watershed- and basin-scale surface hydrology. The work was made possible by the high spatial density of the cooperative station data -- no other available data source possessed this necessary characteristic -- which thus adds to the evidence in the two preceding papers of the great value of the voluntary observers who operate the cooperative stations.

DATA AND ANALYSIS

We created a new data set for use in the research on which this paper is based. It contains the individual daily precipitation totals for 557 locations in the central U. S. (402) and southern Canada (155) for the May-August periods of 1949-80, which total more than 2.19 million observations. The grid-like station network is displayed in Fig. 1; it is bounded to the west and east by the Rocky and Appalachian mountain systems, to the south by the Gulf of Mexico, and to the north by the limit of agriculture and/or the availability of precipitation stations with at least moderate density.

Full details of the construction and characteristics of this data set were presented in Richman and Lamb (1985, 1987), with only a brief overview being given here. The first step in the development of the data set was the acquisition of all daily precipitation totals archived by the U. S. National Climatic Data Center and the Canadian Climate Centre for the study region and period. This involved 11,634 stations, 98% of which were cooperative stations. After identifying the stations with the longest records, and carefully scrutinizing and quality controlling their data, it was concluded that a network resolution of 80-110 km was the finest practicable over most of the domain from the standpoints of spatial and temporal uniformity and data set completeness. A modified 1° latitude-longitude "target" grid (the meridional separation north of 49°N was held constant at the 49°N value) was therefore established, and for each intersection the nearest station with the most complete and seemingly most reliable record was identified. These stations were located, on average, 10 km (U.S.) and 24 km (Canada) from the grid intersection concerned, and were chosen to form the present data set (Fig. 1). For the days these stations did not possess observations (7% of total for U.S.; 23% for Canada), the value from one of two (U.S.) or three (Canada) adjacent "alternate" stations was substituted without adjustment. The alternate stations were usually situated within 30 km of their parent network station and were selected because their records, although often incomplete, covered the gaps in the records of the network station. The resulting data set thus contains actual daily precipitation totals for a regularly spaced array of stations, with no missing observations.

Richman and Lamb (1985, 1987, 1988, 1989) have reported the major early results of a comprehensive analysis of this data set that remains ongoing. In general, the work is applying a range of eigenvectorial techniques (unrotated and rotated Principal Components Analysis; Common Factor Analysis; Procrustes Target Analysis) and correlation analysis to various temporal amalgamations of the data (station totals for 3, 7, 15, 30, 41, and 61 day periods). Detailed information on the statistical methods employed appears in the aforementioned papers, Richman (1986, 1987), and Richman and Easterling (1988). A key thrust of the work to date has been the use of those methods to delineate areas of spatial precipitation coherence. We present here a selection of the results that has particularly broad application.

RESULTS

Fig. 2 provides a rather complete subdivision of the study domain into regions within which three- and seven-day precipitation tends to be spatially coherent *for May-August as a whole*. It is derived from VARIMAX-rotated Principal Component analyses that are fully described in Richman and Lamb (1985, 1987). The regions identified are reasonably uniform in size and span areas equivalent to several medium-sized U. S. states. The validity of the regionalization is enhanced by the fact that, although the U. S. and Canadian analyses were performed separately, their near-border areas of spatial precipitation coherence match closely and appear very meaningful in an agricultural systems/ecosystems sense. An essentially identical regionalization was obtained when the U. S. and Canadian data were analyzed together (Richman and Lamb, 1988, Fig. 2). Some of the most notable features of the present Fig. 2 include: (a) the identification of a region that approximates much of the eastern Corn Belt (Ohio, Indiana, Kentucky, most of Illinois); (b) the linking of Wisconsin, Michigan, and southern Ontario, despite the intervening presences of Lakes Michigan and Huron, to form a "mixed agriculture" region; (c) the existence of the domain's only north-south oriented region in its arid southwest (northwest Texas and western Oklahoma-Kansas); (d) the association of North Dakota with southern Saskatchewan-Manitoba and eastern Montana (cereal production) rather than South Dakota; (e) the linking of northwestern Ontario with northern Minnesota-Wisconsin (largely forest); and (f) the isolation of the important agricultural state of Iowa from the aforementioned eastern Corn Belt region and its fragmented distribution between three of the coherent regions.

Fig. 3 contains counterparts to Fig. 2 for each half-month period during May-August. It thus provides both the basis for the overall seasonal three-seven day precipitation coherence pattern in Fig. 2 and information on the variation of areas of such precipitation coherence during May-August. Unlike Fig. 2, the analyses summarized in Fig. 3 treated the U. S. and Canadian data together. In general, the half-monthly regionalizations in Fig. 3 are quite consistent with the seasonal pattern in Fig. 2 (cf. features (a) - (f) above). The agreement is particularly strong for May-June for the study region as a whole, and for the southern half of that region for all months. Perhaps the most striking intraseasonal variation evident in Fig. 3 concerns the orientation of the areas of precipitation coherence in the zone that contains the Great Lakes. Beginning in early July and persisting through late August, the major axes of these areas tend to steadily pivot from having northwest-southeast alignments in the west to southwest-northeast alignments in the east. This hitherto unidentified feature is not apparent for May-June. It presumably reflects a key intraseasonal variation in the behavior of the precipitation-bearing systems -- probably involving the southward sagging and northward retreat of weak frontal systems -- that invites comprehensive investigation.

Richman and Lamb (1988, Table 2) have shown that the fraction of the total domain variance that is cumulatively explained by the (first) 13 VARIMAX-rotated Principal Components that are the basis for Fig. 3 decreases steadily from early May (61%) to late July (40%), and then increases through early August (44%) and late August (46%). These changes presumably reflect a May through July decrease in organized synoptic-scale precipitation and coincident increase in less organized convective precipitation, and a reversal of those trends during August. This interpretation is strongly supported by Fig. 4, which was constructed as follows. For each of the regions shown in Fig. 3 for a given half-month, the precipitation station with the highest loading on the VARIMAX-rotated Principal Component concerned was identified. Those stations were located close to the center of their region. Their three-day precipitation totals were then separately correlated with the three-day totals for the remaining 556 stations in the full network (Fig. 1) for the half-month concerned. Next, the resulting fields of correlation coefficients were isoplethed and the +0.4 correlation isopleths corresponding to each region were plotted on a single map for each half month. Fig. 4 contains all of those half-month maps. It shows that the total area enclosed by the +0.4 correlation isopleths decreases strongly from early May (when it covers almost the entire domain) until late July (approximately one-third of domain), after which it increases steadily through early and late August. As intimated above, this provides further evidence of the decreasing (increasing) importance of synoptic-scale (convective) precipitation from early May until late July, and the opposite during August.

Fig. 5 shows how the short-period (three-seven day) regionalization pattern in Fig. 2 varies as the precipitation totalling period is extended to 15, 30, 41, and 61 days. The analyses contributing to Fig. 5 treated the U. S. and Canadian data together and were *for May-August as a whole*. The regionalizations obtained for 15 and 30 days (Figs. 5a, b) are remarkably similar to their shorter-period counterpart in Fig. 2 (cf. features (a) - (f) listed above). The major difference is that the regions along the western edge of the domain have become larger (and fewer in number) by 15 and 30 days, while those in the east have shrunk in size. This trend continues as the precipitation totalling period is lengthened to 41 and 61 days (Figs. 5c, d), with the result that most of the U. S. Great Plains are spanned by two very large regions, and some of the eastern regions are further contracted. Despite this evolution, however, several of the 61-day regions have clear three-seven day counterparts in Fig. 2 (two along Gulf Coast; eastern Corn Belt; Wisconsin-Michigan-Southern Ontario; many that span U. S.-Canadian border). This variation of the areas of precipitation coherence with increasing time-scale constitutes new information that invites further investigation. It likely reflects the degree of temporal consistency of the tracks of the precipitation producing systems.

IMPLICATIONS AND APPLICATIONS

We believe that the regionalizations presented in Figs. 2-5 have a number of interesting implications and will be of considerable utility in several contexts. Some details are now provided.

The regionalizations seem particularly relevant to the predictive crop yield modeling that is now routinely performed (often at *weekly* intervals) for the growing season of the central United States by some grain merchandisers, commodity brokers, and their consultants (Lamb et al., 1985). The modelers are uncertain about the number and morphology of the regions for which individual models should be developed and utilized -- such subdivisions need to be partially climate-based -- and also about the spatial representativeness of the year-to-date meteorological data currently used to force the models at the required short intervals. These data are largely limited to observations from the sparse network of National Weather Service (NWS) first-order stations; the national delivery system for the spatially dense NWS cooperative station data used in this study is far too slow to be helpful in the above real-time situation (Lamb et al., 1985). The foregoing uncertainties are strikingly illustrated by the discrepancy between the density of the first-order network (e.g., five stations in Illinois) and the fact that some of the predictive modeling is being performed on an individual county (e.g., 102 in Illinois) scale. We believe that Fig. 2's regionalization can substantially reduce those uncertainties, particularly since it is applicable to the weekly time-scale of much of the modeling. Furthermore, should there be a need to develop and use models for separate portions of the growing season, Fig. 3's half-monthly regionalizations would likely also be helpful.

Because the regionalizations in Figs. 2 and 3 also hold for the three-day time scale, they may have utility for the 48-96 hour prediction of precipitation. Where such prediction has a common basis (e.g., statistical) for fixed areas, as is the case with the NWS Model Output Statistics (MOS) Probability of Precipitation (PoP) forecasts (Lowry and Glahn, 1976; Erickson, 1988), it would seem desirable for those areas to possess demonstrated precipitation coherence on the time scale concerned. Figs. 2 and 3 contain the latter information. Interestingly, one of the MOS PoP forecast area delineations (Lowry and Glahn, 1976, Fig. 7) grouped northern Illinois and most of Michigan with Wisconsin and northeastern Iowa, as do Fig. 2 and many of the half-monthly panels in Fig. 3. However, the NWS and present regionalizations exhibit few other similarities.

A related potential application of Figs. 2 and 3 is in studies of climatic change over the last 100-150 years. Since this work must necessarily be based on the relatively few stations with reliable long-term records (e.g., Griffiths, 1983; Changnon, 1990), there is an urgent need for information on the spatial representativeness of such stations. Figs. 2 and 3 should provide valuable guidance in that regard for investigations of the secular change in short-period precipitation behavior. For similar reasons, they may also be of some utility in the fields of rainfall chemistry (e.g., for the estimation of total areal acidic deposition from the weekly averaged point concentration values that are now available), weather modification (e.g., the siting of experiments and operations could be guided by the regional boundaries), and perhaps watershed- and basin-scale surface hydrology. Furthermore, as already indicated in the discussions of Figs. 3-5, the regionalizations have implications concerning the lifecycles and areas of activity of individual weather systems.

Finally, the spatial variation in the extent to which the three-seven day regionalization (Fig. 2) changes as the precipitation totalling period is extended to 15, 30, 41, and 61 days (Fig. 5) provides a framework for investigating how short-period weather variations integrate to climate fluctuations. Previous climate diagnostics research has tended to neglect this important issue, and instead use monthly, seasonal, or annual data to identify patterns for only one of those time periods in a given study (reviewed in Richman and Lamb, 1985). The eigenvectorial approach employed in the research on which this paper is based particularly facilitates the inquiry into how weather integrates to climate -- it not only delineates areas of spatial precipitation coherence, but also yields formally related information on their relative precipitation abundance as functions of time.

CONCLUDING REMARKS

Our ongoing research is pursuing several of the lines of inquiry suggested by the above results and their diverse implications and applications. These opportunities are thus rooted in the valuable observations of the voluntary observers who operated thousands of cooperative weather stations during 1949-80. The continued contributions of such observers will be greatly needed as the earth and atmospheric sciences, along with Society in general, become increasingly concerned with the possibility of greenhouse gas-induced climate change and the resulting global change (National Research Council, 1986; NASA Advisory Council, 1988). Some of the scenarios for that change have suggested central North America could experience a hotter and drier summer climate by the year 2030 (Schlesinger and Mitchell, 1985; Lamb, 1987). In other words, the future summer climate of Illinois could come to resemble that of Oklahoma today. As implied by earlier discussion of (the cooperative weather station based) Fig. 5, this may well involve changes in the degree of temporal consistency of the tracks of the precipitation producing systems that affect Illinois. It is therefore likely that further research using the existing archive of historical cooperative station data will yield information that can help Society adapt to future climate change. Even more importantly, however, future cooperative station data will most certainly be needed for the difficult *research* task of detecting any such climate change.

ACKNOWLEDGEMENTS

This study was supported by NSF Grants ATM 83-12653 and 85-20877. We appreciated the opportunity to contribute to the celebration that marked a century of observations at the Urbana Cooperative Weather Station. The technical assistance of Linda Riggin, Rebecca Runge, and Phyllis Stone is gratefully acknowledged.

LITERATURE CITED

Changnon, S.A., Jr., 1990: Temporal changes in the climate of Urbana-Champaign and Illinois. Trans. Illinois State Acad. Science, 83, 57-69.

Erickson, M.C., 1988: Development and evaluation of perfect prog guidance based on output from the nested grid Proc. Eighth Conf. Numerical Wea. Pred., American Meteorological Society, Boston, 565-571.

Griffiths, J.F., 1983: The quest for climatic reference stations. The State Climatologist, 7, 9.

Lamb, P.J., 1987: On the development of regional climatic scenarios for policy oriented climatic impact assessment. Bull. Amer. Meteor. Soc., 68, 1116-1123.

Lamb, P. J., S. T. Sonka, and S. A. Changnon, 1985: Use of Climate Information by U. S. Agribusiness. NOAA Technical Report, NCPO 001, 67 pp.

Lowry, D. A., and H. R. Glahn, 1976: An operational model for forecasting probability of precipitation--PEATMOS PoP. Mon. Wea. Rev., 104, 221-232.

NASA Advisory Council, 1988: Earth System Science: A Program for Global Change: A Closer View. Report of the Earth System Sciences Committee, NASA Advisory Council, NASA, Washington, D.C., 208 pp.

National Research Council, 1986: Global Change in the Geosphere-Biosphere. National Academy Press, Washington, D.C., 91 pp.

Richman, M.B., 1986: Rotation of Principal Components. J. Climatol., 6, 293-335.

Richman, M.B., 1987: Rotation of Principal Components: A reply. J. Climatol., 7, 511-520.

Richman, M.B., and W.E. Easterling, 1988: Procrustes Target Analysis: A multivariate too for identification of climate fluctuations. J. Geophys. Res., 93, 10989-11003.

Richman, M. B., and P. J. Lamb, 1985: Climatic pattern analysis of three- and seven-day summer rainfall in the central United States: Some methodological considerations and a regionalization. J. Clim. Appl. Meteor., 24, 1325-1343.

Richman, M. B., and P. J. Lamb, 1987: Pattern analysis of growing season precipitation in southern Canada. Atmosphere-Ocean, 25, 137-158.

Richman, M.B., and P.J. Lamb, 1988: On the modes of growing-season precipitation over central North America. Proc. Twelfth Ann. Climate Diagnostics Workshop, U.S. Department of Commerce, Washington, D.C., 284-291.

Richman, M.B., and P.J. Lamb, 1989: On the interannual variability of growing season precipitation in central North America. Proc. Thirteenth Ann. Climate Diagnostics Workshop, U.S. Department of Commerce, Washington, D.C., 133-142.

Schlesinger, M.E., and J.F.B. Mitchell, 1985: Model projections of the equilibrium climatic response to increased carbon dioxide. Chapter 4 in The Potential Climatic Effects of Increasing Carbon Dioxide (MacCracken, M.C., and F.M. Luther, eds.), U.S. Department of Energy, Office of Energy Research, Report DOE/ER-0237, Washington, D.C., 81-147.

Wendland, W. M., 1989: A history of weather observations in Illinois. Trans. Illinois State Acad. Science, 83, 43-56.

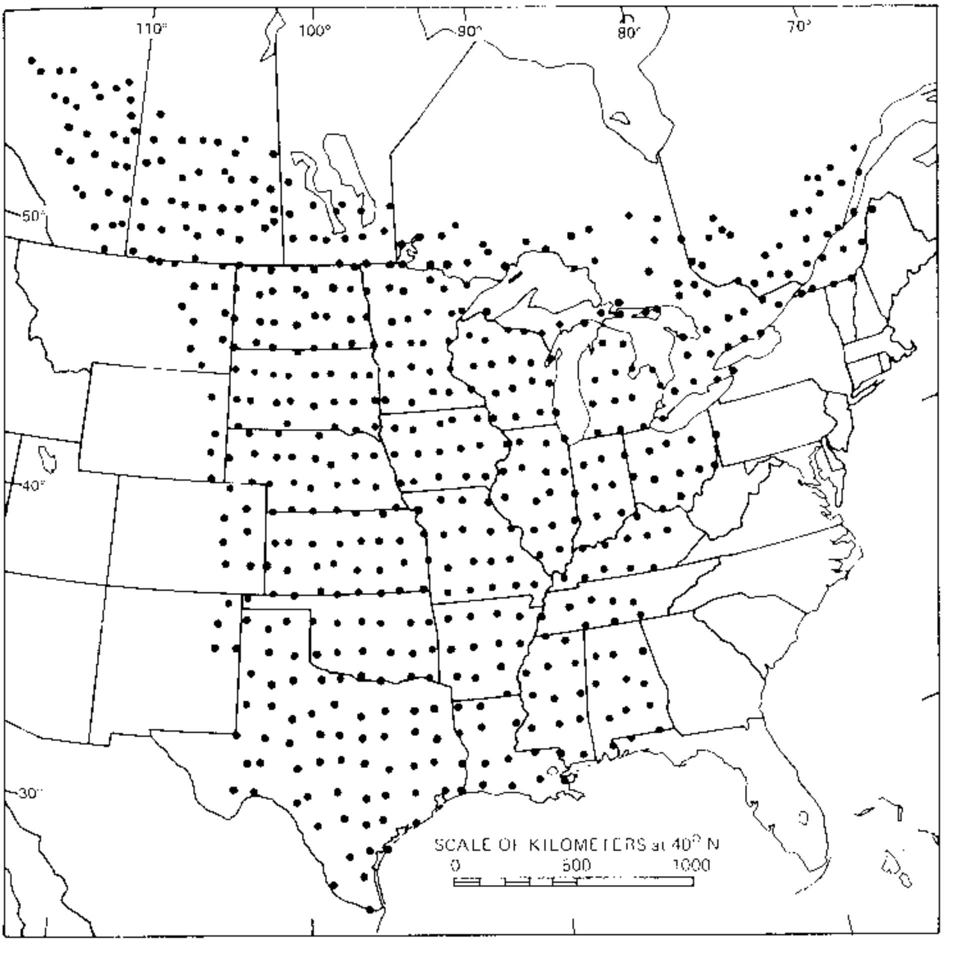

Figure 1. Station network (557) for central North American daily precipitation data set for May-August periods of 1949-80.

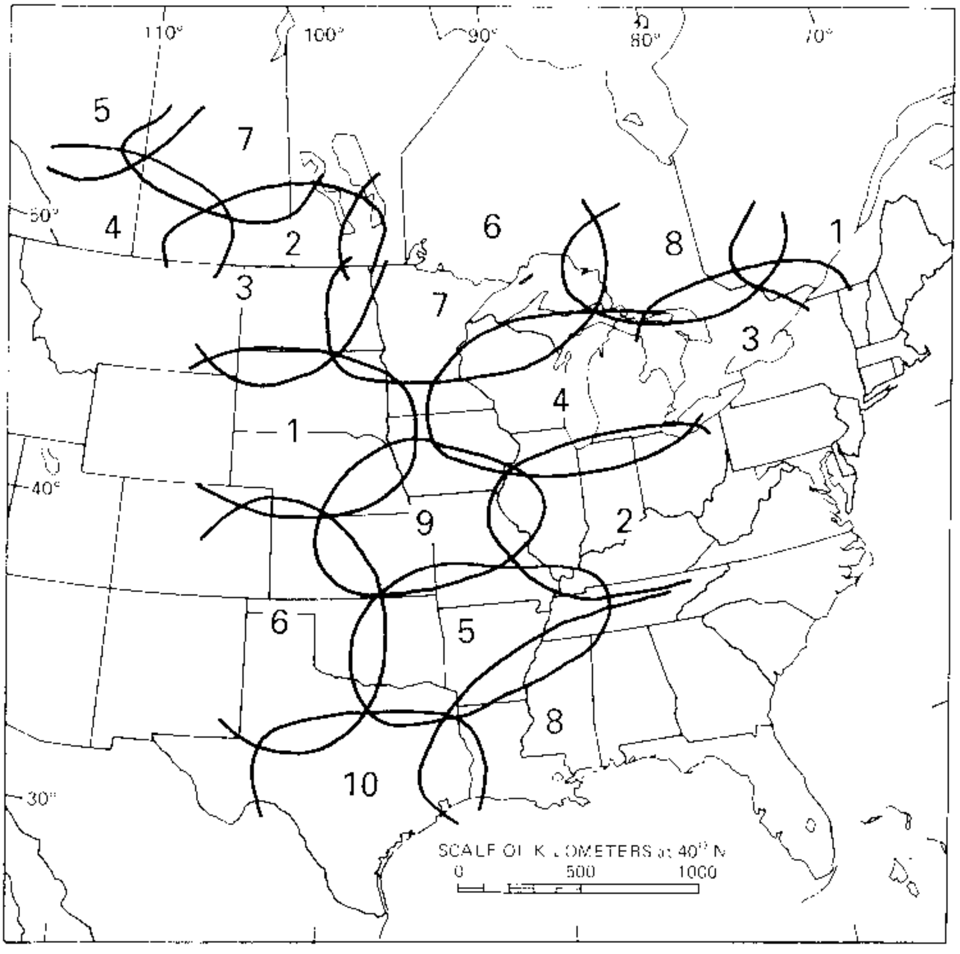

Figure 2. Combined regionalization of central U. S. with southern Canada for three-day (square-root transformed) precipitation totals during May-August. The regional boundaries are the +0.4 loading isopleths for individual VARIMAX-rotated Principal Components identified in Richman and Lamb (1985, Fig. 4; 1987, Fig. 4). The same pattern was obtained for seven-day (square-root transformed) precipitation totals (Richman and Lamb, 1985, 1987). The region numbers provide keys to associated eigenvalue and variance information in Richman and Lamb (1985, Table 2; 1987, Table 2).

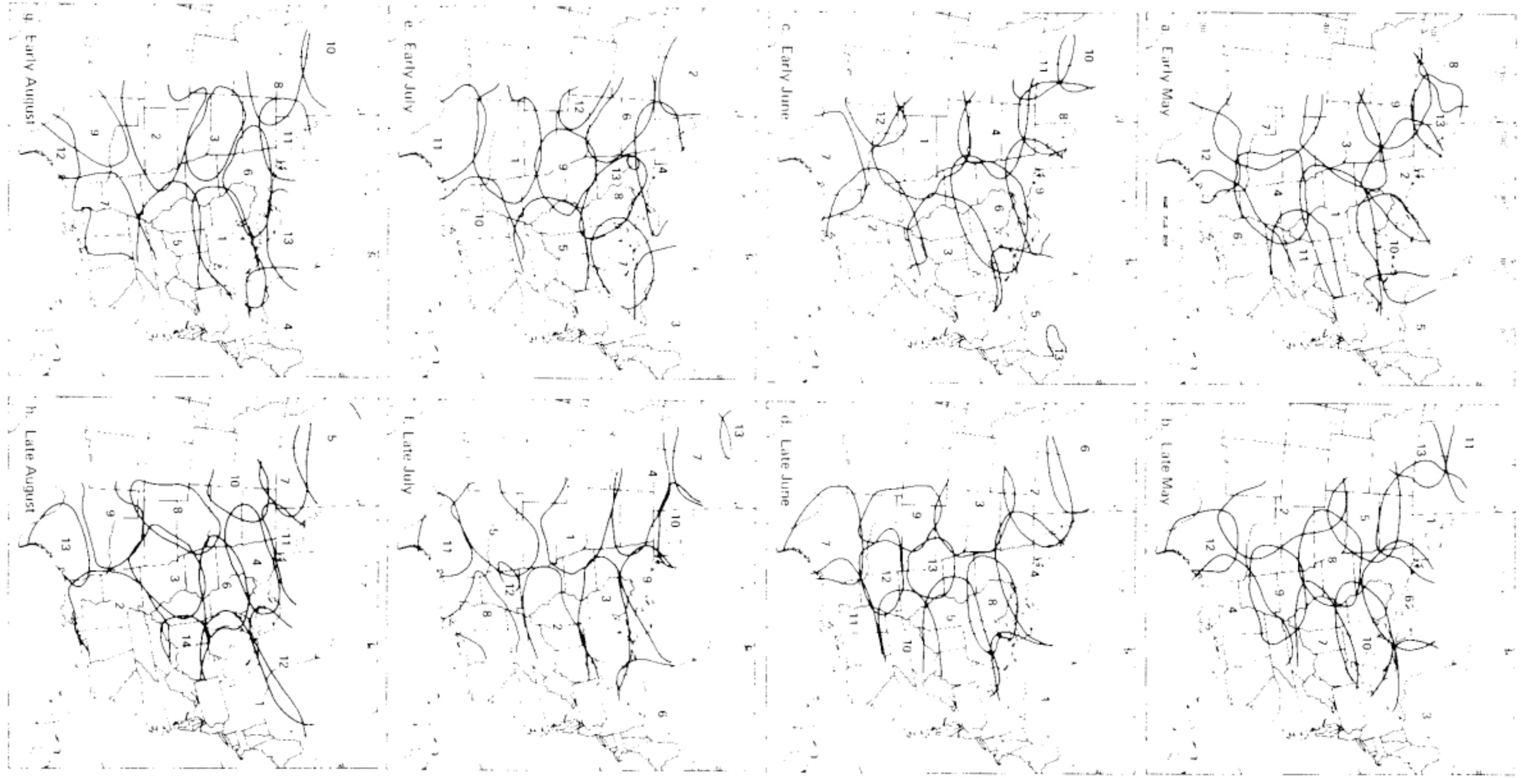

Figure 3. Regionalization of central North America for three day (square-root transformed) precipitation totals for (a) 1-15 May (early May), (b) 16-30 May (late May), (c) 31 May - 14 June (early June), (d) 15-29 June (late June), (e) 30 June - 14 July (early July), (f) 15-29 July (late July), (g) 30 July - 13 August (early August). The regional boundaries are the +0.4 loading isopleths of individual VARIMAX-rotated Principal Components, with the region numbers providing keys to associated variance information in Richman and Lamb (1988, Table 1).

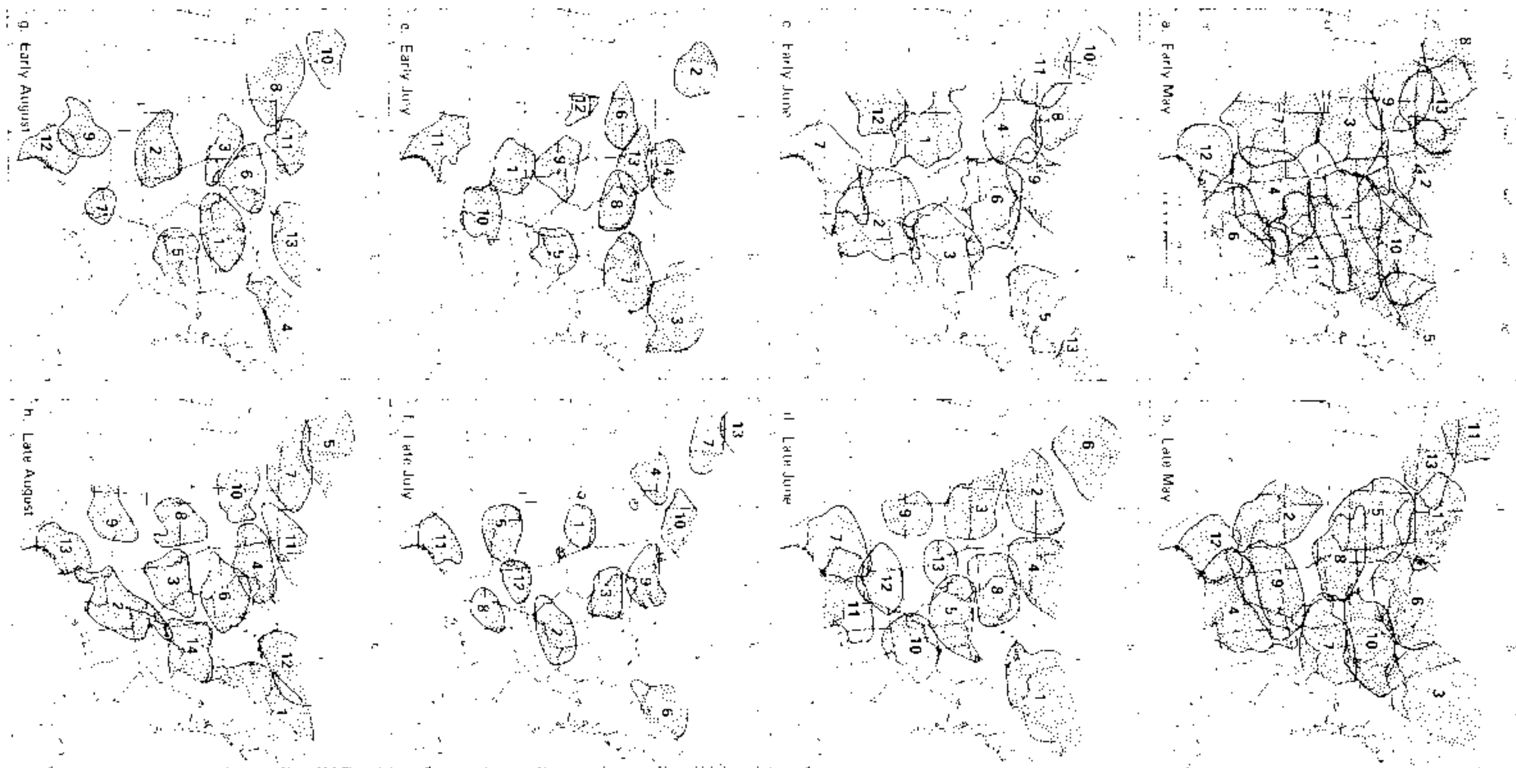

Figure 4. Correlation isopleths (+0.4) corresponding to each region delineated in each panel of Fig. 3 (see text for information on their construction). The identifying numbers are consistent with the region numbers in Fig. 3. Areas encircled by correlation isopleths are shaded.

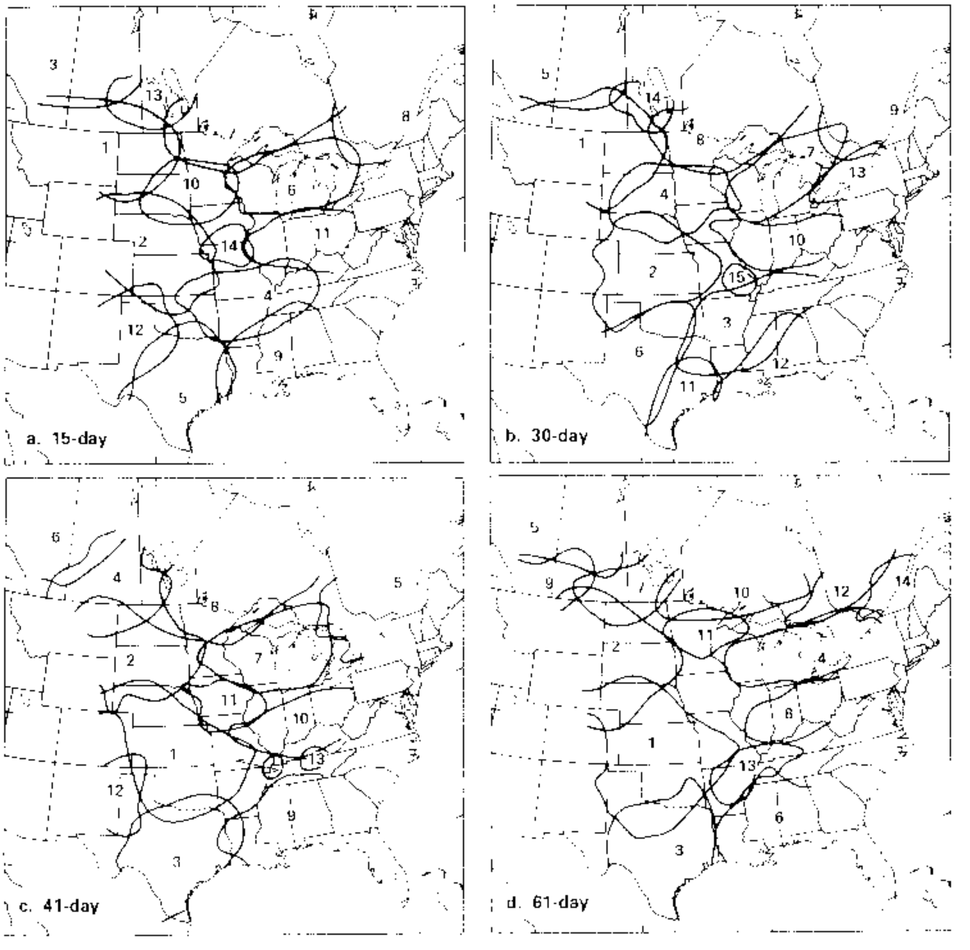

Figure 5. Regionalization of central North America for (a) 15-day (square-root transformed) precipitation totals, (b) 30-day (square-root transformed) precipitation totals, (c) 41-day precipitation totals, and (d) 61-day precipitation totals for May-August. The regional boundaries are the +0.4 loading isopleths of individual VARIMAX-rotated Principal Components, with the region numbers providing keys to associated variance information in Richman and Lamb (1988, Table 1).