An index of interannual precipitation variability
in the core of the North American monsoon region

David S. Gutzler
Department of Earth & Planetary Sciences
University of New Mexico
Albuquerque, NM 87131

Journal of Climate (in press)

Corresponding author: Dr. David S. Gutzler
address: Earth & Planetary Sciences Dept.
          University of New Mexico MSC03-2040
          Albuquerque, NM 87131
          email: gutzler@unm.edu
An index of interannual precipitation variability
in the core of the North American monsoon region

ABSTRACT
Seasonal precipitation anomalies associated with the continental North American monsoon system are characterized using a land-based data set derived from in situ observations across the southwestern United States and northwestern Mexico. Coherent regions of interannual continental precipitation variability are derived from principal components analysis, after defining separate "Early" and "Late" summer monsoon seasons. The gravest mode of Late season interannual variability captures precipitation anomalies in the core of the continental monsoon domain. A simple spatial average is developed as an index of this core variability. The seasonal separation allows examination of persistence of precipitation anomalies as an indicator of practical Late season predictability. Possible influences of large-scale oceanic interannual fluctuations (ENSO and PDO) on core index precipitation anomalies are also considered. The core precipitation index exhibits considerably more Early-to-Late season persistence than an ocean-centered precipitation index. Implications of these results for monthly/seasonal predictability of warm season precipitation are discussed.

1. Introduction
Several empirical indices have been used in the recent literature to represent fluctuations of large-scale precipitation associated the North American Monsoon System (NAMS). Such indices are useful for diagnostic studies that seek to relate monsoon precipitation to other indices of short-term climate variability, such as ocean temperature anomalies or atmospheric circulation patterns.

Higgins et al. (1997) defined an area extending across eastern Arizona and western New Mexico (AZNM) as an index of southwestern U.S. extension of the North American monsoon. A subsequent paper compared the interannual variability of AZNM precipitation to other regions in the NAMS
domain (Higgins et al. 1999). Comrie and Glenn (1998) subjected monthly mean precipitation data from individual stations in the U.S. and Mexico to a rotated principal components analysis, and derived a set of regions to describe the different seasonal cycles of precipitation across the domain. Gutzler (2000) also applied rotated principal components analysis to interannual time series of summer seasonal precipitation in southwestern U.S. climate divisions, yielding a set of three regions with spatial extent comparable to those defined by Comrie and Glenn (1998). Yu and Wallace (2000) developed a much larger scale NAMS index derived from the gravest mode of a principal components analysis of monthly precipitation anomalies over a broad domain. The index region of precipitation anomalies yielded by their analysis is centered in the eastern Pacific Ocean off the Mexican coast.

Higgins et al. (1997) noted that monsoon onset dates in the AZNM region are correlated with total seasonal precipitation; years with early (late) monsoon onset in their study tended to stay anomalously wet (dry). A subsequent study (Higgins et al. 1999) found that other regions within the North American monsoon domain did not exhibit such correlation. Yu and Wallace (2000) also found that their broad-scale NAMS precipitation anomaly index displayed very little month-to-month persistence. The finding of modest correlation between onset date and total precipitation is consistent with results of Dhar et al. (1980) for the Indian monsoon.

Clarifying these diverse observations is important for advancement of warm season precipitation prediction. Long-lead outlooks of monthly or seasonal anomalies implicitly assume that coherent anomalies exist on these time scales. Anomaly persistence is an important component of any ocean-based seasonal prediction scheme. In wetter climates, positive feedbacks associated with soil moisture are called upon to explain persistent warm season continental precipitation regimes (Eltahir 1998), and have been found to be a feature of model simulations of the NAMS domain (Small 2001). If precipitation anomalies show very little monthly or seasonal persistence during the warm season, however, then it would imply that forcing and feedbacks on these time scales are weak relative to transient forcing associated with atmospheric dynamics ("weather"). This would suggest that seasonal predictability of monsoonal precipitation associated with SST or land surface forcing may be low.
Continental warm season precipitation is dominated by deep convective processes. Even when upward motion is forced in the troposphere by large-scale dynamical processes, summertime continental airmasses are typically convectively or conditionally unstable, so that most precipitation is delivered via thunderstorms. Moist convection tends to occur rapidly and on short time scales, so it is plausible that the randomness of deep convection in the NAMS domain, modulated by atmospheric transient variability with a short time scale, limits any predictability associated with monthly/seasonal forcing or feedbacks.

This study describes an objectively defined continental NAMS precipitation index, and revisits the issue of persistence during the monsoon season. Principal components analysis is applied to "Early" and "Late" season precipitation anomalies, facilitating examination of the persistence of warm season precipitation. The gravest mode of this analysis is a region near the core of the monsoon precipitation maximum in northwest Mexico. It is shown to exhibit considerable Early-to-Late seasonal persistence. The results lead to interpretive comments that are intended to raise questions that could be addressed by special observations to be taken in the 2004 field campaign associated with international North American Monsoon Experiment (NAME 2003).

2. Data and methods

The daily gridded “Unified” U.S.-Mexican precipitation data set described by Higgins et al. (1999, 2000b) is used for the analysis in this study. This data product is based solely on rain gauge measurements and therefore covers land areas only. The analysis product covers 48 years, 1951-1998, with 1°×1° spatial resolution. The density of gauges is generally much higher in the U.S, where approximately 2500 stations nationwide (most of which reported data hourly) are used in the analysis, than in Mexico, where the analysis is based on 161 stations reporting daily (Higgins et al. 1999).

The full data set covers much of the land area of North America; for this analysis a spatial subset is analyzed covering the NAMS domain as depicted in Figs. 1-4. In regions of complicated terrain such as the NAMS domain, this data set, like any spatial analysis based on in situ weather stations, almost certainly underestimates the true spatial average. This is because gauges tend to be located preferentially
in valleys, where people live, instead of on slopes and peaks where precipitation amounts are larger. The probable systematic underestimate of precipitation does not preclude use of these data for studies of climate variability, however.

To examine intraseasonal persistence of precipitation anomalies, two monsoon seasons are defined: an “Early” season (15 May-3 Jul) and a “Late” season (5 Jul-15 Sep). Periods longer than monthly averages are desirable for this study to try to reduce the aliasing of variance on the time scale of Madden-Julian Oscillations into the time averages. The beginning and ending dates of these periods were determined from examination of precipitation statistics described by Higgins et al. (1999), who described the seasonal march of the warm season precipitation regime from south to north across western Mexico using the same Unified data set employed in this study. The Early season, from mid-May through the end of June, captures the onset of monsoonal precipitation throughout Mexico, including the region of maximum summer precipitation along the western slopes of the Sierra Madre Occidental.

The beginning of the Late season (5 Jul) is chosen based on monsoon onset dates calculated by Higgins et al. (1999) in the northern reaches of the NAMS domain (poleward of about 30°N), in the American states of Arizona and New Mexico. The 73-day Late season defined here captures the weeks of maximum warm season precipitation across the entire NAMS domain. Other averaging periods could be defined for different purposes tailored to different time scales or geographical subdomains. For example, the Early and Late seasons defined here are not well-suited to characterize variability associated with the "midsummer drought" in the southern part of the NAMS domain (Magaña et al. 1999). Higher frequency fluctuations, associated with moisture surges in the Gulf of California, have been examined recently by Higgins et al. (2004).

Daily precipitation data for each grid box are summed to generate Early and Late season averages for each year in the data record. The 48-year climatological average precipitation in the Early season (Fig. 1a) decreases sharply from the southern end of the computational domain, where it exceeds 4 mm/d, to the northwestern part of the domain in Arizona, which uniformly receives less than 0.5 mm/d. In the northern half of the domain there is an equally pronounced east-west gradient between Arizona,
where May and June are the driest months of the year, and west Texas, which receives relatively abundant springtime rains fed by moisture off the Gulf of Mexico.

Climatological precipitation in the Late season (Fig. 1b) exceeds 2 mm/d across the monsoonal part of the domain defined in previous studies (e.g. Mock 1996; Higgins et al. 1997), where a distinct spring-to-summer increase in precipitation occurs. Largest seasonal values of precipitation occur along the western slopes of the Sierra Madre Occidental, where rainfall rates exceed 4 mm/d with a small area of maximum precipitation exceeding 8 mm/day. The east-west gradient within the southwestern U.S. is gone in the Late season climatological average.

Maps of the interannual standard deviation of annual values of Early and Late season precipitation rates are shown in Fig. 2. The general pattern of Early season interannual variability is similar to the pattern of the climatological average: high in the southeastern part of the domain, very low in the northwestern part (Fig. 2a). In the Late season the variability is larger, with a distinct maximum over the Sierra Madre Occidental between 25°N and 30°N. This center is located well to the northwest of the climatological maximum average (Fig. 1b) during this season.

The correlation matrices of Early and Late season interannual variability were independently subjected to Varimax-rotated principal components analysis (Horel 1981; Richman 1986). Eigenanalysis was carried out on the correlation matrices to ensure that low-variance and high-variance grid points were weighted equally. These analyses yielded sets of spatial patterns describing regions where interannual anomalies tend to be highly correlated. The analyses were carried out several times with different numbers of retained components to examine the sensitivity of the spatial patterns to the truncation chosen. It was determined that eight spatial modes were robust. The expansion coefficients associated with the spatial modes describe the interannual variability of precipitation associated with each one.

In the next section these expansion coefficients are compared straightforwardly with several climate indices commonly used for analysis of interannual variability. Equatorial Pacific SST variations are described using the Niño3.4 index (5°N-5°S, 170°W-120°W), obtained from the NOAA Climate Prediction Center. Averages of Niño3.4 for boreal winter (January+February, when ENSO extrema tend
to reach their peak amplitude) and late spring (May+June, concurrent with the Early monsoon season defined in this paper) are considered separately. Possible effects of the Pacific Decadal Oscillation (PDO; Mantua et al. 1997) are considered by separating the period of record into two parts, before and after 1977. In that year a pronounced climate regime shift occurred which correlates with pronounced changes in the winter circulation across the Pacific (Graham 1994) and the precipitation climatology of southwest North America (Higgins et al. 2000a; Gutzler et al. 2002).

3. Intraseasonal and interannual variability of monsoon precipitation

As described in the previous section, spatial analysis of the interannual variability of Early season continental monsoon precipitation yields eight regions (Fig. 3a). Each of the regional patterns is monopolar (i.e. the analysis does not yield "teleconnection patterns"), a result that is consistent with many previous analyses of precipitation variability, including the studies of NAMS-related precipitation cited in the Introduction. Cross-correlation maps of interannual variability based on single grid points (not shown) confirm that these patterns reproduce the structure and extent of the strongest one-point correlations.

Four of the Early season regions identified lie within the United States, and four lie within Mexico. The first four regions, ranked in order of explained variance, are located in the southern and eastern parts of the domain where Early season mean and variance are large (Figs. 1a, 2a). Corresponding results for the Late season (Jul 5 - Sep 15) are shown in Fig. 3b. In this case five of the regions lie within the U.S. and three are within Mexico.

The gravest mode of variability in the Late season analysis – that is, the pattern whose temporal coefficients have the largest amplitude – is the region in northwestern Mexico. Comparison with Figs. 1b and 2b indicates that this region represents the northern extent of the climatological precipitation maximum along the western slopes of the Sierra Madre Occidental, and exhibits the largest interannual variability in the entire domain. This region will be the focus of the remainder of the analysis presented here, representing the heart of continental interannual variability of the North American Monsoon System.
A simple and convenient spatial index of Northwest Mexico precipitation variability is the continental area within the domain bounded by 24°N-30°N and 112°W-106°W, outlined by the dashed lines in Figs 1b, 2b and 3b. The spatial average of precipitation in this region will henceforth be referred to as the "core" NAMS precipitation index. Time series of core Early and Late season precipitation (Fig. 4a) show the large interannual variability inherent in this region. To some extent the very low precipitation values early in the record, prior to the large jump in precipitation rate after 1960, may be an artifact of station distribution changes with time (W. Higgins, personal communication, 2003), although the decade of the 1950s was a severe drought period throughout Southwest North America.

The same two time series are plotted as a scatter diagram in Fig. 4b. The linear correlation coefficient between annual Early and Late anomalies in the Core region for the entire period of record is 0.55, easily statistically significant. Most of the values in the lower left corner of Fig. 4b are associated with the period of the 1950s that may be somewhat suspect. However the high correlation between Early and Late season core anomalies is present even if those data points are removed from the record as will be shown in the next section. Furthermore, the scatter plot suggests that the relationship between Early and Late core precipitation is nonlinear. Specifically, a dry Early season (precipitation rate<1 mm/d) seems to offer little guidance for the subsequent Late season, but every Early season in the data record with precipitation >1 mm/d was followed by a Late season with at least 3 mm/d precipitation.

The magnitude of Early-Late season anomaly persistence is higher in the core region than in other continental regions (Fig. 3) but several other significant correlations (r>0.4) also exist among the pairs of expansion coefficients yielded by the principal components analysis, especially in the Mexican half of the domain (not shown). Thus continental precipitation associated with the NAMS appears to be substantially more persistent between Early and Late seasons as defined here than the ocean-centered monthly precipitation index described by Yu and Wallace (2001). This result also seems to suggest a higher level of early summer-to-late summer persistence in Mexican summer precipitation than the index regions discussed by Higgins et al. (1999).
4. Continental core precipitation related to other climatic indices

ENSO fluctuations, as defined by the Niño3.4 index, are poorly correlated with core precipitation when the entire 48-yr period of record is considered (Table 1). Correlation magnitudes are less than 0.3 for either winter or late spring Niño3.4. There is a slight tendency for the correlations to be negative, implying that positive precipitation anomalies are associated with La Niña years, but the correlations are so low that little practical prognostic value could be derived from them. Correlations between equatorial SST and summer continental precipitation at regions to the south of the Core area (Early mode 1 and Late mode 2 in Fig. 3) are slightly less than -0.3 but this is still modest compared with cold season correlations. These results are entirely consistent with findings reported by Yu and Wallace (2001).

Considering that ENSO extrema provide much of our current seasonal prediction skill, the slight relationship between Niño3.4 and NAMS prediction anomalies makes seasonal prediction a stiff challenge. Favorable phases of the PDO may improve this situation somewhat. The last two lines of Table 2 suggest that negative correlations between Niño3.4 and Early season core precipitation are enhanced during PDO-positive decades, such as occurred after 1976. Early season correlations decreased to -0.46 after 1976, while the average precipitation rate increased by about 15% relative to the pre-shift 1951-1976 period. The positive correlation between Early and Late seasons (i.e., anomaly persistence) is particularly pronounced after the PDO shift in the Pacific: considering just the 1977-1998 period, the correlation is 0.67.

The statistical relationship between the continental core precipitation index and the Yu and Wallace (2001) index is also shown in Table 2 for the 1979-1998 period. For purposes of comparison the monthly mean time series developed by Yu and Wallace was modified to match more closely the "seasons" defined in this study, by combining monthly values for each year as follows: (May +June*0.5)/1.5 for the Early season and (Jul + Aug + Sep*0.5)/2.5 for the Late season. Correlations are very low for the Early season, when precipitation across the core region is light. During the Late season modest, but significant, correlations above 0.3 are obtained for both antecedent (May-Jun) and concurrent (Jul-Aug-Sep) values of the Yu and Wallace index.
5. Summary and Discussion

Principal components analysis separates interannual variability of continental precipitation in the North American Monsoon domain into eight subregions, corresponding approximately in extent and location to subregions defined in different ways in previous studies (Comrie and Glenn 1998; Higgins et al. 1999; Gutzler 2000). The presence of so many distinct regions within the NAMS domain demonstrates that there is probably no single index that will be suitable for comprehensively characterizing NAMS precipitation variability on interannual time scales.

The analysis identifies a region of Northwest Mexico as the "core" of continental interannual variability of precipitation in the North American monsoon domain. A simple spatial index, covering the land area between 24°N-30°N and 112°W-106°W, captures the northern maximum of continental precipitation during the heart of the monsoon season after the beginning of July (the Late season defined in this study; Fig. 1b). This region exhibits the absolute maximum in continental interannual variability across the NAMS domain (Fig. 2b). The analysis shows that this region also exhibits significant persistence of precipitation anomalies from the "Early" summer season (before climatological monsoon onset in northwest Mexico) to the Late season, particularly in wet years (Fig. 4b).

The observation that continental precipitation exhibits more intraseasonal persistence than nearby oceanic precipitation (Yu and Wallace 2001) implies that climate-scale forcing and feedbacks may play a significant role in modulating NAMS precipitation across the North American continent. This should be considered an encouraging result from the perspective of climate predictability, given the ambiguous results described in the Introduction on warm season precipitation anomaly persistence from previous studies. Even in the absence of definitive understanding of the mechanisms responsible for persistence in the core region, some simple monthly-seasonal predictability could potentially be realized from this result.

The specific mechanism for anomaly persistence cannot be diagnosed from precipitation data alone, though the results are consistent with the following speculative interpretation. The apparent nonlinearity in the Early vs. Late season scatter plot (Fig. 4b) supports the concept that an anomalously wet land surface across the core precipitation region in the Early season helps promote the continuation of
anomalously wet conditions in the Late season (Small 2001). No such persistence of dry conditions is apparent in the data, suggesting that wet initial surface conditions are not necessary for abundant rainfall in the Late season. Perhaps the land surface in Northwest Mexico is on average so dry that a pre-onset "dry anomaly" does not provide much forcing for subsequent monsoon precipitation. In order to quantify such land surface feedbacks, large-scale observations of surface fluxes will be needed.

This analysis would not have been possible without a 50-year data record of continental precipitation, and the gridded format of the U.S.-Mexico unified data set makes spatial analysis technically simple. However, there is very little evidence of cross-border precipitation correlation in both the Early and Late season analyses (Fig. 3). This rather suspicious result may indicate a shortcoming of the merged analysis procedure that combined the spatially denser, hourly U.S. data and the coarser, daily Mexican data. In addition, the extreme decadal variability present in the core time series (Fig. 4a) may be at least partly an artifact of temporal inhomogeneity of the station data in Mexico. More detailed assessment of data quality is beyond the scope of this study, but points to the need for a fully reanalyzed precipitation data set that blends U.S. and Mexican in situ data more thoroughly for integrated cross-border analyses.

The science goals of the forthcoming international North American Monsoon Experiment (NAME 2003) include special observations, analysis and modeling that could address some of the questions remaining from this analysis. In particular, NAME seeks to measure and analyze precipitation and climatic boundary values such as land surface characteristics. By improving the observational data base and supporting diagnostic and modeling research, progress can be made on the broad questions of land surface feedbacks and seasonal predictability discussed in this note.

Acknowledgements. This research has been supported by a grant from the NOAA Office of Global Programs. Wei Shi and Wayne Higgins (NOAA NCEP) provided the gridded Unified precipitation data, helpful comments and friendly hospitality during the author's sabbatical visit to the Climate Prediction Center. Bin Yu (University of Victoria) kindly provided his NAMS precipitation index and constructive comments on the manuscript.
References


Dhar, O.N., P.R. Rakhecha, and B.N. Mandal, 1980: Does the early or late onset of the monsoon provide any clue to the subsequent rainfall during the monsoon season? *Mon. Wea. Rev.*, 108, 1069-1072.


**Table 1**

Correlation coefficients between interannual anomalies of precipitation in the continental Core region during the Early monsoon season (15 May-3 Jul) or Late season (5 Jul-15 Sep), and Niño3.4 SST anomalies for preceding winter months (Jan+Feb, denoted JF) or late spring months (May+Jun, denoted MJ). The area represented by the core region is shown by the dashed lines in Figs. 1b, 2b and 3b. Correlations are calculated over the entire period of record (1951-1998). Mean area-averaged precipitation rates (mm/d) in the Early and Late seasons are listed in the first column.

<table>
<thead>
<tr>
<th>Precipitation Index</th>
<th>mean (mm/d)</th>
<th>Core/Late</th>
<th>Niño3.4/JF</th>
<th>Niño3.4/MJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core Early Season</td>
<td>0.79</td>
<td>0.55</td>
<td>-0.29</td>
<td>-0.14</td>
</tr>
<tr>
<td>Core Late Season</td>
<td>3.40</td>
<td>1.00</td>
<td>-0.12</td>
<td>-0.09</td>
</tr>
</tbody>
</table>
Table 2

Correlation coefficients between interannual anomalies of precipitation as in Table 1, but calculated using subsets of the entire period of record divided by the 1977 shift in the Pacific Decadal Oscillation (PDO) Index (Mantua et al. 1997). Part (b) of the Table includes correlations of core region precipitation with the monthly mean index described by Yu and Wallace (2001), denoted YW, averaged over May+Jun (MJ) or Jul-Sep (JAS); see text for details.

<table>
<thead>
<tr>
<th>Precipitation Index</th>
<th>mean (mm/d)</th>
<th>Core/Late</th>
<th>Niño3.4/JF</th>
<th>Niño3.4/MJ</th>
<th>YW/MJ</th>
<th>YW/JAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) 1951-1976, pre PDO shift</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core Early Season</td>
<td>0.74</td>
<td>0.48</td>
<td>-0.19</td>
<td>0.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core Late Season</td>
<td>3.19</td>
<td>1.00</td>
<td>-0.16</td>
<td>-0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core Early Season</td>
<td>0.86</td>
<td>0.67</td>
<td>-0.46</td>
<td>-0.39</td>
<td>0.07</td>
<td>-0.06</td>
</tr>
<tr>
<td>Core Late Season</td>
<td>3.70</td>
<td>1.00</td>
<td>-0.21</td>
<td>-0.27</td>
<td>0.33</td>
<td>0.32</td>
</tr>
</tbody>
</table>
Figure Captions

1. Climatological average precipitation rates (mm/d) for “Early” and “Late” monsoon seasons. Contour interval of solid lines is 1 mm/d, with 0.5 mm/d and 1.5 mm/d contours dashed. The Core monsoon index region (defined in Section 3) is delineated by a heavy dashed line.

   (a) Early season, 15 May-3 July. (b) Late season, 5 July-15 September.

2. Interannual standard deviation (mm/d) of Early and Late season precipitation rates. Contour interval 0.25 mm/d; contours < 1 mm/d are dashed and contours ≥ 1 mm/d are solid. The Core monsoon index region (defined in Section 3) is delineated by a heavy dashed line.

   (a) Early season, 15 May-3 July. (b) Late season, 5 July-15 September.

3. Spatial patterns of interannual precipitation anomalies yielded by Varimax-rotated principal components analysis of Early and Late season averages. The patterns are numbered 1-8, ordered by the percentage of domain-wide interannual variance accounted for by each pattern. The Core monsoon index region is delineated by a heavy dashed line.

   (a) Early season, 15 May-3 July. (b) Late season, 5 July-15 September.

4. Annual values of Early and Late season precipitation rate (mm/day) in the continental Core region outlined by the heavy dashed lines in Figs. 1b, 2b, and 3b.

   (a) Time series for Early season (lower curve) and Late Season (upper curve).

   (b) Scatter plot of Early and Late season annual values. The linear correlation between the two sets of values is $r = 0.55$. 

15
Figure 1. Climatological average precipitation rates (mm/d) for “Early” and “Late” monsoon seasons. Contour interval of solid lines is 1 mm/d, with 0.5 mm/d and 1.5 mm/d contours dashed. The Core monsoon index region (defined in Section 3) is delineated by a heavy dashed line. (a) Early season, 15 May-3 July. (b) Late season, 5 July-15 September.
Figure 2. Interannual standard deviation (mm/d) of Early and Late season precipitation rates. Contour interval 0.25 mm/d; contours < 1 mm/d are dashed and contours ≥ 1 mm/d are solid. The Core monsoon index region (defined in Section 3) is delineated by a heavy dashed line. (a) Early season, 15 May-3 July. (b) Late season, 5 July-15 September.
Figure 3. Spatial patterns of interannual precipitation anomalies yielded by Varimax-rotated principal components analysis of Early and Late season averages. The patterns are numbered 1-8, ordered by the percentage of domain-wide interannual variance accounted for by each pattern. The Core monsoon index region is delineated by a heavy dashed line. (a) Early season, 15 May-3 July. (b) Late season, 5 July-15 September.
Figure 4. Annual values of Early and Late season precipitation rate (mm/d) in the continental Core region outlined by the heavy dashed lines in Figs. 1b, 2b, and 3b.

(a) Time series for Early season (lower curve) and Late Season (upper curve).

(b) Scatter plot of Early and Late season annual values. The linear correlation between the two sets of values is $r = 0.55$. 