A 36-yr Climatological Description of the Evaporative Sources of Warm-Season Precipitation in the Mississippi River Basin

KAYE L. BRUBAKER
Department of Civil and Environmental Engineering, University of Maryland, College Park, Maryland

PAUL A. DIRMeyer
Center for Ocean–Land–Atmosphere Studies, Calverton, Maryland

ARIEF SUADRJAT
Department of Civil and Environmental Engineering, University of Maryland, College Park, Maryland

BENJAMIN S. LEVY
MITRE Corporation, McLean, Virginia

FREDRIC BERNAL
Institute for Software Research, Inc., Fairmont, West Virginia

(Manuscript received 25 October 2000, in final form 20 June 2001)

ABSTRACT

The terrestrial and oceanic sources of moisture that supply warm-season rainfall to the Mississippi River basin and its subbasins are examined over a 36-yr period (1963–98). Using hourly observed precipitation, National Centers for Environmental Prediction (NCEP) reanalyses at 6-h intervals, and a back-trajectory algorithm, the water falling during observed precipitation events is probabilistically traced to its most recent surface evaporative source, terrestrial or oceanic. Maps of these sources generally show dual maxima, one terrestrial and one oceanic, in spring and a dominance of terrestrial sources in summer. Pentad time series averaged over the 36 years show a late-summer maximum of precipitation recycling in all but the Missouri subbasin. During the 36 years analyzed, 32% of warm-season precipitation in the entire Mississippi basin originated as evaporation within the basin (recycled). About 20% of warm-season precipitation was contributed directly by evaporation from the Gulf of Mexico and Caribbean. The Midwest flood year, 1993, represents a positive outlier in terms of July precipitation supplied to the Upper Mississippi directly by evaporation from the Caribbean. The monthly recycling ratios for warm-season precipitation during the drought year, 1988, represent extreme values in the time series but are not identified as outliers. A positive trend in precipitation recycling in the Upper Mississippi and Missouri subbasins and accompanying decrease in Gulf of Mexico/Caribbean–supplied precipitation to those regions are statistically significant but may reflect changes in the observational data stream assimilated by the NCEP model. Perturbation analysis demonstrates that the source fractions and recycling ratios are somewhat sensitive to systematic errors but not to random errors in the model-derived evapotranspiration (ET), arguably the largest source of uncertainty in the back-trajectory approach. Systematic errors in terrestrial ET on the order of 20% introduce errors of about 0.02 in land source fractions (including recycling ratios) that are themselves on the order of 0.10–0.30.

1. Introduction

In the context of the Global Energy and Water Cycle Experiment Continental-Scale International Project, it has been suggested that knowledge of source areas of precipitation to the Mississippi basin could contribute to seasonal or interannual precipitation forecasting for that region. Higgins et al. (1997a,b) have suggested that not only the Great Plains low-level jet but also the North American monsoon system over southwestern North America play significant roles in determining warm-season rainfall over the central United States. Meanwhile, Findell and Eltahir (1997) have shown observational evidence for local feedbacks between antecedent soil moisture and rainfall over the central United States.
States. Theories based on dynamics and thermodynamics have been proposed to explain the observed relationships, but there also may be a more direct role played by the reallocation of atmospheric moisture between regions.

The amount of precipitation over a given region is determined by hemispheric and mesoscale dynamics and moisture supply, described by the velocity vector and the humidity state of the air. Direction and humidity are not independent; how much moisture the air contains is generally a function of where the air has been. Systems that supply moisture to continental regions are driven by synoptic or large-scale systems and by local surface temperature differences. Over land, the moisture state of the soil is the controlling factor in surface temperature variations, as well as affecting the water content of the air through evapotranspiration. Thus, evapotranspiration from land surfaces plays a role in both aspects of precipitation formation: dynamics and moisture supply. Local evapotranspiration controls local moisture supply as well as the thermodynamic state of the atmosphere, whereas variations in circulation may alter the remote sources of moisture for rainfall over a particular region on intraseasonal to interannual timescales.

It has been hypothesized that regions receiving much of their precipitation moisture from local, terrestrial sources may be subject to more sustained dry periods in comparison with those supplied mostly by oceanic sources (Entekhabi et al. 1992). In terms of their moisture supply, such systems have a positive feedback. Bosilovich and Sun (1999) used a regional model to study the impact of terrestrial moisture sources on the Great Plains flood of 1993. They found that by suppressing local evaporation from the surface water budget, rainfall over the Great Plains was reduced by as much as 20%, although there existed some compensation by increased moisture flux convergence in the atmosphere. The influence of evaporation upwind of the flood region was slightly less than the local impact. In Dirmeyer and Brubaker (1999), it was found that the moisture sources contributing precipitation to the Mississippi basin during drought and flood years were very different.

Estimating the relative contributions of water evaporated from various regions to rainfall over a specific area is not trivial. Brubaker et al. (1993) performed bulk calculations of precipitation recycling over a number of specific regions around the globe using a bulk approach and monthly mean data. Dirmeyer and Brubaker (1999) showed that the bulk approach has inherent problems owing to the nonlinear nature of the moisture advection term. Trenberth (1999) calculated recycling and the relative magnitudes of evaporation and precipitation to moisture advection using monthly data and a bulk approach but applied globally as a function of a predefined length scale. This method, when combined with calculations of moisture advection, can give a rough idea of moisture sources and sinks for any area of the globe but suffers from the same shortcomings as the work in Brubaker et al. (1993).

Models that allow for the tracking of atmospheric tracers have been used to diagnose moisture sources and sinks (e.g., Druyan and Koster 1989; Numaguti 1999). These approaches investigate a purely model climate, which may not well represent the real world. Perhaps the ideal approach would be to use such a tracer model to track water vapor from evaporation to precipitation event in a data assimilation mode, so that the model circulation could be constrained by observations. How-
ever, no tracer analysis is currently being done with an operational analysis model over the Mississippi basin. Thus, the next best approach would be to approximate the tracer method by using high-temporal-resolution atmospheric analyses in tandem with observed precipitation data to calculate trajectories that are close to what a tracer model would give. Dirmeyer and Brubaker (1999) applied such an approach to two hydrologic extreme years in the Mississippi basin: the drought summer of 1988 and the floods of 1993. The summer of 1988 showed a higher recycling ratio, and the map of sources in 1993 showed a significantly increased and elongated source region from the Gulf of Mexico and Caribbean Sea. One motivation for the current study was to determine whether these years were anomalous in terms of the source regions supplying the different amounts of seasonal rainfall.

The current study extends the methodology of Dirmeyer and Brubaker (1999) to a continuous 36-yr period to establish a “climatology,” or description of the long-term mean behavior, of moisture sources for the warm season over the Mississippi basin, and to investigate the seasonal and interannual variability in these sources. Section 2 describes the data and methods used. Results are presented in section 3. Section 4 presents an investigation into the sensitivity of the results to errors and uncertainties in evapotranspiration. In section 5, the time series of precipitation source fractions are analyzed for statistical outliers and trends. A summary and conclusions are given in section 6.

### Table 1. Target and source region definitions.

<table>
<thead>
<tr>
<th>Subbasins of the Mississippi</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>LM</td>
<td>Lower Mississippi</td>
</tr>
<tr>
<td>MO</td>
<td>Missouri</td>
</tr>
<tr>
<td>OH</td>
<td>Ohio</td>
</tr>
<tr>
<td>RA</td>
<td>Red/Arkansas</td>
</tr>
<tr>
<td>UM</td>
<td>Upper Mississippi</td>
</tr>
<tr>
<td>Entire Mississippi Basin</td>
<td>LM + MO + OH + RA + UM</td>
</tr>
<tr>
<td>Remote land source regions</td>
<td></td>
</tr>
<tr>
<td>ARC</td>
<td>Arctic</td>
</tr>
<tr>
<td>NNA</td>
<td>Northern North America</td>
</tr>
<tr>
<td>WES</td>
<td>Western North America</td>
</tr>
<tr>
<td>SOW</td>
<td>Southwest North America</td>
</tr>
<tr>
<td>TMX</td>
<td>Texas/Mexico</td>
</tr>
<tr>
<td>SOE</td>
<td>Southeast North America</td>
</tr>
<tr>
<td>CSA</td>
<td>Central and South America</td>
</tr>
<tr>
<td>Remote ocean source regions</td>
<td></td>
</tr>
<tr>
<td>TEP</td>
<td>Temperate Pacific</td>
</tr>
<tr>
<td>TRP</td>
<td>Tropical Pacific</td>
</tr>
<tr>
<td>TEA</td>
<td>Temperate Atlantic</td>
</tr>
<tr>
<td>TRA</td>
<td>Tropical Atlantic</td>
</tr>
<tr>
<td>GOM</td>
<td>Gulf of Mexico</td>
</tr>
<tr>
<td>CAR</td>
<td>Caribbean</td>
</tr>
</tbody>
</table>

### Table 2. Relative contributions of the Gulf of Mexico/Caribbean and the Mississippi basin to subbasin warm-season precipitation.

<table>
<thead>
<tr>
<th>Subbasin</th>
<th>Percent of warm-season precipitation (1963–98) contributed directly by evapo(transpi)ration from:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower Mississippi (LM)</td>
<td>M (Gulf of Mexico/Caribbean) = 26 (36)</td>
</tr>
<tr>
<td>Ohio (OH)</td>
<td>M (Mississippi basin) = 39 (22)</td>
</tr>
<tr>
<td>Upper Mississippi (UM)</td>
<td>M (Mississippi basin) = 44 (17)</td>
</tr>
<tr>
<td>Red/Arkansas (RA)</td>
<td>M (Mississippi basin) = 19 (19)</td>
</tr>
<tr>
<td>Missouri (MO)</td>
<td>M (Mississippi basin) = 34 (9)</td>
</tr>
</tbody>
</table>

2. Data and methods

Using gridded hourly observed precipitation (Higgins et al. 1996), National Centers for Environmental Prediction–National Center for Atmospheric Research (hereinafter referred to as NCEP) reanalyses at 6-hourly intervals (Kalnay et al. 1996), and a back-trajectory algorithm, the water falling during each observed precipitation event is probabilistically traced backward in time to obtain a spatial distribution of that water’s most recent evaporative sources, whether terrestrial or oceanic. Repeating this analysis over all precipitation events in a given time period produces a map of the most likely evaporative sources of water that became precipitation at a given location. Details of this algorithm are given in Dirmeyer and Brubaker (1999). The advantage of the back-trajectory method over linear bulk methods is that it can track the movement of air parcels associated with precipitation events and is not based on time- and space-averaged values of atmospheric variables.

The NCEP reanalysis 2D and 3D fields used in this work consist of 1) winds, temperature, and specific humidity on sigma levels; 2) surface pressure; and 3) surface evaporation rate, at T62 (about 1.9° × 1.9°) horizontal resolution and 28 vertical levels (the model’s original sigma levels). The NCEP/Climate Prediction Center precipitation dataset of Higgins et al. (1996) is on a 2° × 2.5° grid. Our calculations are performed on the NCEP reanalysis grid, and the precipitation data are interpolated onto that grid. Dirmeyer and Brubaker (1999) discuss the quality of the input data and conclude that the combined reanalysis dynamics and observed precipitation are adequate for this application. The accuracy of the model evapotranspiration remains of concern, and therefore we take a closer look at the sensitivity of these results to uncertainty in the evapotranspiration (ET) inputs (section 4).

Sources are analyzed for 1963–98 by pentad for precipitation beginning 16 March. The algorithm traces the
evaporative sources up to 15 days backward in time; therefore, evaporation contributing to rainfall in the 16 March pentad may be traced back as early as 1 March. Of the 31 pentads analyzed, pentads 1–15 (16 March–29 May) are defined as spring; 16–31 (30 May–17 August) are defined as summer.

Figure 1 shows the Mississippi basin and its subbasins outlined on the grid of the NCEP reanalyses. The contribution by selected source regions was computed for precipitation over the whole Mississippi basin and separately over the five target subbasins by pentad for March–August. (The target basins are identified by two-letter codes in Fig. 1 and the codes are defined in Table 1.) The analysis gives a grid-based accounting of the most recent evaporative location of water falling as rain during year $y$ ($1963–98$) and pentad $p$ ($1–31$). For target region $t$ ($t =$ subbasin UM, MO, OH, RA, or LM or the entire basin M), the contribution of a particular cell $x$ on the grid is expressed as a spatially distributed source function $e_{x,t,s,y,p}$ ($\text{kg m}^{-2}$). The total warm-season contribution of cell $x$ to precipitation in target region $t$ is the sum over all pentads:

$$E_{x,t,s,y} = \sum_{p} e_{x,t,s,y,p} \quad (\text{kg m}^{-2}).$$  

(1)

In addition to the basin and its subbasins, 14 additional source regions are identified (three-letter codes on Fig. 1). A source region’s evaporative contribution to rainfall in a target region is the area-weighted sum of the contributions by each of the grid cells that make up that region:

$$e_{x,t,s,y,p} = \sum_{x \in s} e_{x,t,s,y,p} A(x) \quad (\text{kg}),$$  

(2)

where $e_{x,t,s,y,p}$ ($\text{kg}$) is the mass of precipitation in target region $t$ contributed by evaporation in source region $s$ during year $y$ and pentad $p$, and $A$ is the area of the grid cell ($\text{m}^2$). The total mass of precipitation in target region $t$ during pentad $p$ of year $y$ is given by

$$p_{t,s,y,p} = \sum_{x \in s} e_{x,t,s,y,p} \quad (\text{kg}).$$  

(3)

(The definitions of the measures used in this study are summarized in the appendix.) Over the entire warm season, the total mass of precipitation supplied to target region $t$ by source region $s$ is given by

$$M_{t,s,y} = \sum_{p} p_{t,s,y,p} \quad (\text{kg}).$$  

(4)

and the total warm-season precipitation to that target region is the sum of precipitation over pentads, or the sum of mass contributions over source regions,

$$P_{t,s,y} = \sum_{p} p_{t,s,y,p} = \sum_{s} M_{t,s,y} \quad (\text{kg}).$$  

(5)

In any given warm season, we are interested in what fraction of the target region’s precipitation was supplied by a given source region. This fraction is defined as

$$F_{t,s,y} = \frac{M_{t,s,y}}{P_{t,s,y}}.$$  

(6)

We may wish to consider the fraction of precipitation contributed to target $t$ by source $s$ during a general time period $T$, consisting of one or more pentads (for example, a month), which can be obtained as follows:

$$F_{t,s,y,T} = \frac{\sum_{p \in T} e_{x,t,s,y,p}}{\sum_{p \in T} p_{t,s,y,p}}.$$  

(7)

For a climatological perspective, any of these quantities may be averaged over years to obtain a 36-yr average, indicated by an overbar. For example, a 36-yr average source fraction for target $t$, source $s$, pentad number $8$ would be

$$\overline{F}_{t,s,8} = \frac{\sum_{p \in 8} F_{t,s,y,p}}{36}.$$  

(8)

3. Results

The evaporative source for rain in the basin and in each of five subbasins, averaged over the 36 years analyzed, is presented in map form in Figs. 2–7. The grid boxes composing the target region $t$ are outlined in each figure. In the top half of each figure is the spring (March–May) spatial distribution $E_{t,MAM}(x)$, and in the bottom is the summer (June–August) distribution $E_{t,JA}(x)$. In each case, the contours show locations of equal contribution ($\text{kg m}^{-2}$) to precipitation in the target region; the areal integral of the function defined by the contours equals the average seasonal precipitation to that region $P_t$ ($\text{kg}$) over the 36 yr. Grid cells in which the 36-yr mean contribution is at least 2 times as large as the interannual standard deviation are shaded (an approximation to a 95% confidence level). The volume described by the contours in Fig. 2 (entire Mississippi basin M) is the sum of the volumes for the subregions in Figs. 3–7.

For the Mississippi (Fig. 2) in spring, two primary maxima appear, one over the southern part of the basin and one over the gulf/Caribbean region. During spring, rainfall in the basin is derived largely from evaporation from the Gulf of Mexico (source region GOM) and from the southern part of the basin itself (LM); there is also a secondary source region off Baja California (BAJ). In summer, the continental and gulf maxima merge into a single maximum centered over the south-central United States. With the establishment of anticyclonic flow around the Bermuda high, the fetch of the Gulf of Mexico/Atlantic oceanic source region lengthens into the Caribbean (CAR) and tropical Atlantic (TRA). The importance of the BAJ region is reduced in summer, when evaporated moisture from that source region evidently rains out in northern Mexico and the U.S. Southwest with the onset of the North American monsoon.

The Lower Mississippi subbasin (LM, Fig. 3) shows a similar pattern to the entire basin, with the center of mass of the source map shifting to the south-central United States in summer. For the Missouri (MO, Fig.
FIG. 2. The evaporative source for rain falling over the Mississippi basin (M, indicated by double-ruled outline) during spring (Mar–May) and summer (Jun–Aug), averaged over 1963–98. Contours are at 1, 2, 4, 8, 16, 32, and 64 kg m\(^{-2}\) (equivalent to millimeters of liquid water). Shading indicates grid cells in which the 36-yr mean contribution is at least 2 times as large as the interannual standard deviation (an approximation to a 95% confidence level).

Over the 36 yr studied, the fraction of total precipitation contributed by the Mississippi basin from its own evaporation (i.e., \(\Sigma_{\text{all}} M_{M,M,y}/\Sigma_{\text{all}} P_{M,y}\)) was about 32\%. The precipitation contributed directly by evaporation from the Gulf of Mexico and Caribbean [i.e., \(\Sigma_{\text{all}} (M_{M,GOM,y} + M_{M,CAR,y})/\Sigma_{\text{all}} P_{M,y}\)] was 20\%. As one moves generally from southeast to northwest, the within-basin contribution increases and the fraction contributed di-
rectly from GOM/CAR decreases (Table 2), consistent with the general direction of air and moisture flow over this region. As discussed below, the RA subbasin is unique in receiving a significant fraction of its precipitation from Pacific sources.

Figure 8 provides additional detail on the seasonal variation of total and recycled rain for the entire basin and its five subbasins. For each target region, the panels show a time series of 36-yr average total pentadwise precipitation \( (P_{Tg}, T_g) \) and recycled precipitation \( (M_{Tg}, T_g) \). The 36-yr average pentadwise recycling ratios \( f_{Tg} \) are given for all of the target regions in Fig. 9. In these figures, recycled precipitation is defined as rainwater having its most recent evaporative origin within the target region.

During the late summer, a decline in total precipitation to the Mississippi basin is accompanied by a general increase in recycling (Fig. 8a). This analysis reflects the dominance of convective precipitation, with corresponding shorter distances of moisture transport between evaporation and precipitation events. In the Lower Mississippi (Fig. 8f), there is a marked increase in the recycling ratio from spring to summer, accompanying lower total rainfall, decreased transport from the Gulf of Mexico and Caribbean sources, and increased terrestrial evapotranspiration. The Ohio and Red/Arkansas basins (Figs. 8d,e) also show increased recycling in the summer. The Upper Mississippi (Fig. 8b) has a suggestion of a similar pattern but with more consistent recycling ratios throughout the year. The Missouri subbasin (Fig. 8c) is unique, in that its time of maximum rainfall coincides with the time of maximum recycling (June).
The fractions of the 36-yr total precipitation supplied by each of the 19 sources to each subbasin are tabulated by season (Table 3). For each subbasin and season, the source regions are ranked from high to low. In all cases during the summer season, the subbasin itself is among the top three sources, with recycling ratios ranging from 0.11 (RA, third source) to 0.22 (MO, top source). In spring, recycled moisture is among the top four sources for all but Red/Arkansas. GOM and CAR are the top oceanic contributors to LM, OH, and UM, with GOM ranking as the highest contributor to LM in both seasons and to OH in spring. The southerly components of the mean flow and transient cyclonic systems carry evaporated moisture from the land regions Texas–Mexico (TMX) and Southeast (SOE) northward into the Mississippi basin, placing TMX among the top five sources for all but OH in summer and SOE as the fifth most important source for the easternmost subbasins, LM and OH. The westernmost and largest subbasin, MO, has only western regions in its top four sources in both seasons. The northern North America (NNA) source contributes a greater share of precipitation to UM than to MO, despite the fact that MO is a larger target region and shares a longer boundary with source region NNA; this difference is likely due to warm-season storm tracks.

The RA is unique in that its dominant springtime moisture source is a Pacific Ocean region: BAJ. The BAJ region contributes 14% of spring season precipitation while GOM, the third source, contributes 11%; in summer, GOM continues to contribute 11% while BAJ is reduced to 9% and the TMX land region becomes the primary source. Figure 10 shows the abrupt drop-off in the BAJ source contribution to Red/Arkansas dur-
Fig. 5. As in Fig. 2, but for the Red/Arkansas subbasin (RA, indicated by double-ruled outline).

In June, accompanied by a parallel drop in total precipitation to the RA region. This analysis suggests an early-summer switching mechanism in the rain-out site of evaporated moisture from the Baja region: in spring, the moisture is transported well into the continental interior and reaches the Red/Arkansas before raining out; in summer, with the onset of the North American monsoon, that moisture flow is interrupted by precipitation-forming mechanisms over northern Mexico and the U.S. Southwest.

Although recycling is not the only focus of this paper, it is interesting to consider the fraction of precipitation to a given region that is supplied by that region itself [i.e., \( t = s \) in Eqs. (2) and (4)]. As discussed elsewhere (Brubaker et al. 1993; Eltahir and Bras 1994), the value of the recycling ratio is 0 for a point, 1 for the entire earth, and between 0 and 1 for a region of intermediate size. Figure 11 shows, on a log-log plot, the dependence of the recycling ratio on the size of the domain for a sequence of subregions within the Mississippi basin, from a single grid cell at the Gulf coast to the 99 cells of the entire basin, during the warm season of a randomly selected year, 1992. Also included are the five subbasins defined in Fig. 1; the results for the entire basin and the subbasins are shown as circled crosses. The recycled fraction is approximately proportional to the square root of area in this example. The warm season of 1992 was arbitrarily selected for this analysis as an example; some interannual variation in the details could be expected, but physical reasoning dictates that the recycling ratio should increase with target region size. In a study of recycling...
in the Amazon region, Eltahir and Bras (1996) also report a square root relationship.

4. Sensitivity to uncertainty in model evapotranspiration

Of the gridded datasets being used, model-derived ET is considered to be the least reliable. It is classified as a category-"C" variable, meaning the quantity is solely from the model, constrained only by the model states and parameterizations (Kalnay et al. 1996). Therefore, it is of interest to evaluate the sensitivity of the source-region estimates and recycling ratios to uncertainty in ET. Although the true error bars on ET are unknown, the sensitivity analysis can demonstrate how strongly the results may be affected by errors. Errors in model ET could be systematic, or random, or have components of both. Therefore, two perturbation sensitivity analyses were conducted, one with systematic and one with randomly distributed perturbations. Each perturbation study required running the entire source-region estimation procedure as described in section 2, with the model's surface ET adjusted by some amount at every point in space and time. For the systematic perturbations, all terrestrial ET values were adjusted by the same fraction of their value, ranging from $-50\%$ to $+50\%$. For the random perturbations, the fractional perturbation of ET at any point in space–time was drawn from a normal probability distribution function with a mean of 0, $-20\%$, and $20\%$ and a constant standard deviation of 10%. The contributions of all the source regions to rain falling over UM, MO, OH, RA, and LM during
spring and summer of 1992 were computed under these perturbed ET scenarios, and were compared with the unperturbed solution.

Results for the Upper Mississippi subbasin are presented, because they demonstrate well the sensitivity of the model to uncertainty in ET and evaporation $E$. Perturbed evaporative source regions are shown for marine $E$ perturbations and terrestrial ET perturbations (Figs. 12a and b, respectively). The results in Fig. 12 implicitly reflect the back-trajectory moisture-accounting mechanism. The algorithm starts from the inner terrestrial target region (UM in this case) and moves outward to other regions until the total mass represented by the precipitation event is accounted for. Perturbations to marine $E$ do not affect the mass of evaporative moisture for UM supplied by the inner terrestrial regions but do change the extent of the oceanic source region (Fig. 12a). When terrestrial ET is decreased (increased), however, both terrestrial and marine source regions must expand (contract) to account for the total water supplied as precipitation to the target region (Fig. 12b).

Relative sensitivity $R_s$ is a dimensionless value that can be used to measure the importance of any term in a mathematical model:

$$R_s = \frac{\Delta y}{\Delta e \cdot Y} = \frac{\Delta y}{Y} \cdot \frac{\Delta e}{E},$$

where $E$ and $Y$ are, respectively reference values of ET and any output variable, $\Delta e$ represents a perturbation in ET, and $\Delta y$ is the responding change in the output quantity (McCuen 1974). Equation (9) is a finite-dif-
difference approximation to the slope of the curve when fractional change in $y$ is plotted against fractional change in ET. Such curves are shown for the five subbasins’ source contributions to target region UM in the 1992 sensitivity experiments (Fig. 13). For positive systematic perturbations of ET (right-hand side of Fig. 13a), the response is linear, with $R_s$ ranging from 0.4 for the OH and LM sources to 0.7 for the UM source.
However, the source fraction is increasingly sensitive to negative perturbations of ET, as seen by the changing slope of the line for $\Delta e/E$ less than 0; For LM, an ET perturbation of $-0.25$ causes a change of $-0.2$ ($R_s = 0.8$) in $F_{UMLM,1992}$, and an ET perturbation of $-0.5$ results in a decrease of $-0.48$ in that variable ($R_s = 0.96$). Necessarily, as $\Delta e$ approaches $-E$, the fractional change in $F_{UMLM,1992}$ approaches $-1$ because of the dependence of $F$ on $E$. Negative errors (underestimates) in ET would lead to underestimation of the recycling ratio and land-region source fractions and positive errors to overestimation. An interpretation of the perturbation analysis is that the results are more sensitive to negative errors than to positive ones and are increasingly sensitive as the errors grow more negative.

The addition of a spatially random component to the perturbations has little effect on the sensitivity response (Fig. 13b). Errors due to random perturbations alone, without a systematic bias in ET, were found to be negligible, suggesting that “noise” in ET does not affect the calculations of recycling or the relative importance of various source regions. The results shown here are for the entire warm season; the summer results are more sensitive to uncertainty in ET than those for spring, because of higher reference values of terrestrial ET during summer [reference ET appears in the numerator in Eq. (9)].

Warm-season ET is believed to be positively biased in the reanalysis, likely driven by excessive precipitation over southeastern North America and nudging of soil moisture to climatological values (Trenberth and Guillemot 1998). Assuming ET to be the major source of error or uncertainty in the source fraction and recycling ratios, this would imply that the terrestrial source fractions and recycling ratios may be somewhat overestimated. If we assume a bias on the order of 20%, then the resulting overestimates in the source fraction ratios are on the order of 10% or 15% of their nominal values, a difference of about 0.01-0.02 in the recycling ratio $F_{UMLM,1992}$, for example.

Most of the inner terrestrial source regions’ contributions $F_{r,s,1992}$ have an $R_s$ value above 0.3 (considered sensitive) for systematic terrestrial ET perturbations. The recycling ratio $F_{r,t,1992}$ is most sensitive to errors in terrestrial ET ($R_s > 1$ for some subbasins). The recycling ratios of all subbasins are changed within the range of about ±40% of their reference value when terrestrial ET is systematically perturbed from −50% to +50% (Fig. 14).

Precipitable water from the NCEP reanalyses agrees very well with radiosonde observations (raobs) over land areas where the raob data are assimilated such as...
Fig. 12. Summary maps for sensitivity to systematic error in (a) marine evaporation \( E \) and (b) terrestrial evapotranspiration \( ET \) in the calculated evaporative sources of precipitation to Upper Mississippi subbasin (target UM). Only the 1 and 20 kg m\(^{-2}\) contours are shown. Five curves are plotted for each contour: the thick center contour indicates the base case (no perturbation in \( E \) or \( ET \)), outer contours reflect surface \( E \) or \( ET \) systematically reduced by 25% and 50%, and inner contours reflect surface \( E \) or \( ET \) systematically increased by 25% and 50%. Shading is between the 25% and 50% contours.

Over the continental United States, whereas satellite-derived observational products such as the National Aeronautics and Space Administration Water Vapor Project (NVAP; Randel et al. 1996) have mean values that are slightly higher, with reduced temporal variability (Jedlovec et al. 1998). Over areas with few observations, such as over ocean and at low latitudes, the reanalysis estimates of precipitable water are generally higher than NVAP, particularly in deep convective regions. This result may indicate a wet bias in the reanalysis when it is not constrained by observations. Given the alleged high bias in evaporation in the NCEP reanalysis and that source regions in our back-trajectory method are determined as a function of the ratio of \( ET \) to precipitable water,
there may be a tendency to have an offsetting compensation of errors over remote source areas.

5. Outlier detection and time series analysis

Dirmeyer and Brubaker (1999) focused on the hydrometeorologically outstanding years of 1988 and 1993. The recycling ratio for the Mississippi River basin was found to be higher in the drought year of 1988 (41%) than in the flood year of 1993 (33%). The contribution of the Gulf of Mexico and Caribbean source regions were unusually large in 1993, increasing enormously during the June–July peak of that year’s flood. A major motivation for the study reported here was to determine whether those years represented anomalies with respect to precipitation recycling and source contributions. If a dry spell is caused by suppressed remote precipitation sources while local sources continue to contribute moisture, the recycling ratio increases simply because the denominators in Eqs. (6) and (7) are decreased disproportionately to the numerators. If all sources, both local and remote, are suppressed in a dry year, then little variation in the recycling ratio would be expected. In a similar way, if wet spells are driven largely by unusual remote or oceanic contributions of moisture, then increased total precipitation would be associated with decreased recycling ratios.

The 36-yr time series of selected source-region contributions to the entire M basin and the combined UM and MO subbasins are examined on a monthly basis, focusing on the M, GOM, and CAR sources, as in Dirmeyer and Brubaker (1999). For each target region (M or UM + MO), time series plots and box plots are presented in Figs. 15 and 16. The box plot is a compact data descriptor, as well as a nonparametric technique to detect suspect outliers in a dataset. An outlier is an observation that is unusually large or small relative to the other data. The box-plot method of outlier detection is based on the interquartile range IQR, defined as follows:

\[ IQR = Q_U - Q_L, \]  

(10)

where \( Q_U \) and \( Q_L \) are, respectively, the sample upper quartile (75th percentile) and lower quartile (25th percentile); data points that lie outside a distance of 1.5IQR from either \( Q_U \) or \( Q_L \) are suspect or moderate outliers, and any data points beyond 3IQR are considered to be highly suspect outliers. The box plot technique is an alternative to the the Z-score test for outliers, based on the sample mean and standard deviation:

\[ Z = \frac{y - \bar{y}}{s}, \]  

(11)

where \( y \) is a particular observation, \( \bar{y} \) is the sample mean, and \( s \) is the sample standard deviation; a Z score greater than 3 in absolute value indicates a suspect outlier. Both techniques are discussed in textbooks on descriptive statistics, such as Mendenhall and Sincich (1995). For small datasets, the presence of outliers may inflate the sample standard deviation used to calculate the Z scores; outliers do not affect the interquartile range used in constructing the box plot.

For the Mississippi basin as a whole, the monthly
recycling ratios are plotted as time series (Fig. 15a) and as box plots showing the medians, means, quartiles, and suspect outliers (Fig. 15b). The recycling ratio $f_{M,GOM,\text{month}}$ is higher during the summer months than in April; it is most variable in July, least variable in April, most skewed in June, and most symmetrical in May. April 1968, 1987, 1989, and 1998; May 1972; and July 1984 are identified as outliers. The drought year, 1988, has near-average recycling in April, above median in May and June, and below median in July. The flood year of 1993 has above-average recycling in April and June, and below median in July. April 1988, has near-average recycling in April, above median in May and June, and below median in July. The flood year of 1993 has above-average recycling in April and June, and below median in May, and well below median (although not identified as an outlier) in July.

Figures 15c and 15d present the fraction of monthly Mississippi basin precipitation traced to GOM. Of the four months studied, April shows the largest mean value of $f_{M,GOM,\text{month}}$; this fraction declines in May and June and rises again in July. No outliers were identified in these datasets. In 1988, the contribution from GOM to M is below median in April, May, and July and is above median in June. In 1993, the GOM source fraction is high but well within the inner fence defined by $Q_L + 1.5IQR$ (the highest value of $f_{M,GOM,\text{Apr}}$ appears in 1974). The 1993 GOM source fraction is below average in May and July and is slightly above average in June. The time series of monthly source fractions from CAR to M are shown in Figs. 15e and 15f. In 1988, April is on the low side of average, May represents a minimum value (but not an outlier), and June and July return to median values. Dirmeyer and Brubaker (1999) identified the flood year of 1993 as having an unusually large contribution from this source region. That year’s fraction begins well below median in April, moves to median in May, above median in June, and climbs to a maximum value in July. Nonetheless, there is enough variability in the $f_{M,GOM,\text{Jun}}$ time series that July of 1993 does not represent an outlier in the box plot. Like the GOM source fractions to M, April shows the largest relative contribution from CAR, with maxima in 1989 and 1991.

For the combined Upper Mississippi and Missouri basins (UM + MO), the Mississippi basin source (M) contributions are shown in Figs. 16a and 16b. (Because the UM + MO target region is a subset of the M source

<table>
<thead>
<tr>
<th>TABLE 3. (Continued)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Red/Arkansas (RA)</strong></td>
</tr>
<tr>
<td>Springer source</td>
</tr>
<tr>
<td>BAJ</td>
</tr>
<tr>
<td>TMX</td>
</tr>
<tr>
<td>GOM</td>
</tr>
<tr>
<td>TEP</td>
</tr>
<tr>
<td>CAR</td>
</tr>
<tr>
<td>WES</td>
</tr>
<tr>
<td>RA</td>
</tr>
<tr>
<td>TRP</td>
</tr>
<tr>
<td>MO</td>
</tr>
<tr>
<td>SOW</td>
</tr>
<tr>
<td>LM</td>
</tr>
<tr>
<td>TEA</td>
</tr>
<tr>
<td>NNA</td>
</tr>
<tr>
<td>CSA</td>
</tr>
<tr>
<td>TRA</td>
</tr>
<tr>
<td>UM</td>
</tr>
<tr>
<td>SOE</td>
</tr>
<tr>
<td>ARC</td>
</tr>
<tr>
<td>OH</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

* Percentile is percent of dataset less than or equal to the selected value (rounded to whole numbers); Z score is deviation from the sample mean normalized by sample standard deviation.
Fig. 13. Sensitivity to errors in model ET of subbasins' fractional contribution to precipitation in the Upper Mississippi region, this analysis does not represent a recycling ratio. In these time series, April of 1980, June of 1992, and July of 1989 are identified as outliers, all on the positive tail of their respective distributions. In 1993, the contribution of this local, terrestrial source to precipitation in UM + MO is above median in April, May, and June, and only July shows a large suppression of basin-supplied rainfall. For 1988, the in-basin supply is above median for all four months. The only other year showing all four months above median for this variable is 1979, a year of near-normal summer precipitation.

The monthly contribution by GOM to UM + MO (Figs. 16c,d) is generally less than 10%, with the lowest average in June. Only May of 1977 is identified as an outlier. In 1988, the GOM contribution to UM + MO is below median for April and June, near median in May, and just above median in July. In 1993, the GOM source fraction to UM + MO is near median in all months but June, for which it is somewhat above median, but not unusually so.

The average fraction of UM + MO precipitation contributed directly by CAR is less than 5% for all four months (Figs. 16e,f). In contrast to the result for the entire M basin, July of 1993 is identified as an outlier, with an unusually high fraction of precipitation (9%) in these northern subbasins originating directly from remote Caribbean evaporation. Note that the April contribution was well below median that year and that the May and June contributions of CAR to UM + MO are progressively higher when compared with their respective 36-yr records. For 1988, the CAR source was below median in April and June, a record low (but not an outlier) in May, and near median in July.

Two different measures of relative standing were computed for the 1988 and 1993 analysis results: the percentile represented by each monthly source fraction and its Z score (or departure from normal), with respect to the entire time series for that month. These statistics are presented in Table 4 for the entire Mississippi basin and in Table 5 for the combined Upper Mississippi and Missouri basins. Values are calculated on the basis of 36 data points for April and July; a computer system problem unfortunately caused a loss of intermediate calculations for the months of May and June in 1978; therefore, these two months have only 35 data points. No outliers are identified by the Z-score method, although the July-1993 CAR source to UM + MO comes close to the rule-of-thumb cutoff with a Z score of 2.8. Outliers identified by the box-plot method have been discussed above. For the hydrologically anomalous years of 1988 and 1993, only July of 1993 appears as an outlier in this analysis.

In general, an outlier may be attributed to one of three causes: errors are introduced in data collection or transmission, the observation is drawn from a different population than the rest of the sample, or the observation is correct and represents a rare event in the population (e.g., Mendenhall and Sincich 1995). We are confident in our analysis technique, and July of 1993 is a member of the population by definition, leading to the conclusion that the large July-1993 CAR source fraction was a rare occurrence, which possibly would be observed again in a longer time series.

For a symmetrical distribution, a Z score of 0 corresponds to the 50th percentile. Because the source-fraction datasets have skewed distribution, there is not a perfect agreement between the two measures of relative standing. Nonetheless, certain patterns are evident in Table 4 (M): the above-average recycling in all four months in 1988, the change from below-average GOM and CAR contributions in April–June to above-average contributions in July of 1988, and the April-to-July decline in relative share of recycling and the corresponding increase to record-high CAR source in 1993. Corresponding patterns appear in Table 5 (UM + MO). The return to normal contributions from both GOM and CAR sources to M...
and UM + MO in July of 1988 corresponds to the ending of the meteorological drought in that year.

Time series of the seasonal source-region mass contribution ($M_{\text{s,t,y}}$) and mass fraction ($F_{s,t,y}$) were tested for significant temporal trends. The significance of a temporal trend in either variable was tested with a lag-1 serial correlation coefficient $r_x$, a univariate test statistic. The null hypothesis of no trend ($\rho_x = 0$) was evaluated against the alternative hypothesis of a nonzero trend ($\rho_x > 0$), with an appropriate sample statistic and prescribing a 10% level of significance. For cases in which the hypothesis test indicated a significant trend at the 10% level, the slope of the trend was determined. Table 6 summarizes the analysis results for $F_{s,t,y}$; 17 source–target pairs were found to have significant trends in the fraction of seasonal rainfall contribution.

Any trends in time series based on the NCEP reanalyses are likely to reflect the major change in the model’s data assimilation in the late 1970s, at which time satellite data began to be included among the observational datasets (Fiorino 1999). A few studies have investigated trends in variables related to the moisture
Fig. 15. 36-yr time series and summary box plots of Mississippi basin (target region M) monthly precipitation fraction traced to (a), (b) the Mississippi basin itself (recycling), (c), (d) the Gulf of Mexico, and (e), (f) the Caribbean. In the box plot, boxes indicate upper and lower quartiles (UQ, LQ), the solid center line is the median, the dashed center line is the mean, and the whiskers (vertical lines) are the highest and lowest data values within inner fence [1.5 times the interquartile range (IQR) from UQ and LQ]. Solid circles indicate moderate outliers, defined as data points lying outside the inner fence. Extreme outliers (none found) would lie outside the outer fence (3IQR from UQ and LQ). For Apr and Jul, \(n = 36\); for May and Jun, \(n = 35\) (loss of 1978 data).
source fractions reported here. Roads et al. (1998) found no trends in global atmospheric water cycling rate in the NCEP reanalysis. White (1999) found trends in global \( P, E, \) and radiation in the NCEP reanalysis and evidence for the continued overestimation of \( E \) relative to \( P. \) Given the nonstationarity in the assimilated data fields and the attendant uncertainties, any conclusions about the trends shown here would be premature. With these caveats, it is interesting that, for the interior continental subbasins, UM and MO, the trends shown are
negative for the oceanic sources and positive for the nearby continental source regions, implying that, over the 36-yr period, the reanalysis-derived contribution of oceanic moisture (relative to total observed precipitation) in these regions has been decreasing. A sample time series, the fraction of UM rainfall contributed by the GOM source (significant negative trend with the greatest slope) is shown (Fig. 17).

6. Summary and conclusions

This study has produced a 36-yr climatology for the atmospheric branch of the hydrologic cycle over central North America. We have mapped the evaporative sources of the moisture that becomes rainfall over the Mississippi basin and its subbasins, based on gridded observed precipitation and back-trajectory modeling, using NCEP reanalysis dynamic fields. Interannual time series of the moisture source fractions for the Mississippi and its subbasins have been constructed, for total warm-season precipitation (March–August) and for each month separately. Maps of the precipitation sources generally show dual maxima, one terrestrial and one oceanic, in spring and a dominance of terrestrial sources in summer. Pentadwise time series averaged over the 36 years show a late-summer maximum of precipitation recycling in the basin as a whole and in all but the Missouri subbasin.

According to this analysis, about 32% of total warm-season Mississippi basin rainfall from 1963 to 1998 had its most recent evaporative source within the basin, as compared with 20% originating directly from evaporation off the Gulf of Mexico and Caribbean. The Midwest flood year, 1993, represents a positive outlier in terms of the fraction of July rainfall supplied to the Upper Mississippi/Missouri subbasins directly by evaporation from the Caribbean. In April–June 1993, local source fractions (or recycling) were above average. In other words, the anomalous wet summer of 1993 was fed by extreme values of both local terrestrial and distant oceanic sources during different months. The recycling ratio for the Mississippi basin in 1993 was unusually low only during July. The degree of warm-season precipitation recycling during the drought year of 1988 was above average but was not particularly anomalous with respect to the 36 years; no months of that year are identified as outliers.

The computed recycling ratios and source contributions are affected by uncertainty in the input data. We examine the effect of uncertainty in model evapotranspiration, on the assumption that it is the least well known of the forcing variables for our scheme. Systematic errors in ET on the order of 20% are shown to introduce errors of the forcing variables for our scheme. Systematic errors are affected by uncertainty in the input data. We examine the effect of uncertainty in model evapotranspiration, on the assumption that it is the least well known of the forcing variables for our scheme. Systematic errors in ET on the order of 20% are shown to introduce errors of the forcing variables for our scheme.

Table 5. Measures of relative standing for selected source fractions of combined Upper Mississippi and Missouri basin precipitation ($F_{um-basins}$) in 1988 and 1993.

<table>
<thead>
<tr>
<th>Source</th>
<th>Apr ($n = 36$)</th>
<th>May ($n = 35$)</th>
<th>Jun ($n = 35$)</th>
<th>Jul ($n = 36$)</th>
<th>Apr ($n = 36$)</th>
<th>May ($n = 35$)</th>
<th>Jun ($n = 35$)</th>
<th>Jul ($n = 36$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>60/0.11</td>
<td>82/0.82</td>
<td>82/0.87</td>
<td>80/0.85</td>
<td>94/1.32</td>
<td>56/0.18</td>
<td>56/0.12</td>
<td>3/1.55</td>
</tr>
<tr>
<td>GOM</td>
<td>14/−0.96</td>
<td>41/−0.33</td>
<td>18/−0.82</td>
<td>63/0.12</td>
<td>37/1−0.47</td>
<td>47/−0.21</td>
<td>70/0.22</td>
<td>40/−0.28</td>
</tr>
<tr>
<td>CAR</td>
<td>31/−0.57</td>
<td>0/−1.82</td>
<td>65/−0.04</td>
<td>63/0.17</td>
<td>6/1.47</td>
<td>79/0.85</td>
<td>88/1.31</td>
<td>100/2.80</td>
</tr>
</tbody>
</table>

* See note to Table 4.

Fig. 17. 36-yr time series of the fraction of warm-season precipitation in the Upper Mississippi (target region UM) contributed by the Gulf of Mexico (source region GOM).
of about 0.02 in source fraction and recycling ratios that are themselves on the order of 0.10–0.30. Spatially random perturbations in ET did not affect the results.

Several trends have been identified in the time series of warm-season precipitation target–source pairs developed in this study. Although statistically significant, these trends must be interpreted with caution because of step changes in the observational data assimilated by the NCEP model during the 1970s. Bearing this caveat in mind, of particular interest for further investigation is an apparent increase in the share of terrestrial sources (including recycling) and attendant decrease in the contribution of the Gulf of Mexico/Caribbean to rain in the Upper Mississippi cycling) and attendant decrease in the contribution of the Amazon Basin to rain in the Upper Mississippi and Missouri subbasins over the 36 years. Continuing work will investigate patterns related to the El Niño–Southern Oscillation and Pacific decadal oscillation signals and will examine interannual variability in these source fractions on seasonal and monthly scales.

As of the time of writing, an Internet site had been established to display the exhaustive collection of maps and figures generated by this analysis, of which only a few examples could be included here (http://rainsrc.umd.edu/GCIP/).

Acknowledgments. The authors thank D. Lettenmaier, R. Oglesby, and two anonymous reviewers for their thoughtful critiques and suggestions, which substantially improved the paper. This research is supported under the NOAA–NASA GCIP Program. Participation by KLB, AS, BL, and FB is supported by Grant NA96-GP0268. PAD’s participation is supported by Grant NA96-GP0353.

APPENDIX

Notation

\begin{align*}
t & \quad \text{Index for target region} \\
s & \quad \text{Index for source region} \\
y & \quad \text{Index for year} \\
p & \quad \text{Index for pentad} \\
p_{t,y,p} & \quad \text{Total mass (kg) of precipitation in target region} \\
& \quad \text{during year } y \text{ and pentad } p \\
P_{t,y} & \quad \text{Total mass (kg) of precipitation to target region} \\
& \quad \text{in year } y \\
M_{t,y} & \quad \text{Total mass (kg) of precipitation to target region} \\
& \quad \text{supplied by source region } s \text{ in year } y \\
F_{t,y,s} & \quad \text{Fraction of precipitation contributed to target} \\
& \quad \text{region } t \text{ by source region } s \text{ during year } y \\
& \quad \text{(if } t = s \text{, this represents a recycling ratio)} \\
f_{t,y,s,T} & \quad \text{Fraction of precipitation contributed to target} \\
& \quad \text{region } t \text{ by source region } s \text{ during year } y \\
& \quad \text{and period } T \text{ (consisting of a number of pentads; if } t = s \text{, it is a recycling ratio)}
\end{align*}

REFERENCES


Fiorino, M., 1999: The impact of observing system changes on the climate-scale variability and temperature in the ECMWF and NCEP reanalyses. Proc. Second WCRP Int. Conf. on Reanalyses, Reading, United Kingdom, WMO, 65–68.


Trenberth, K. E., 1999: Atmospheric moisture recycling: Role of advection and local evaporation. J. Climate, 12, 1368–1381.
