Influence of the Madden–Julian Oscillation on Forecasts of Extreme Precipitation in the Contiguous United States

CHARLES JONES
Earth Research Institute, University of California, Santa Barbara, Santa Barbara, California

JON GOTTESCHALCK
Climate Prediction Center, NOAA/National Centers for Environmental Prediction, Camp Springs, Maryland

LEILA M. V. CARVALHO
Earth Research Institute and Department of Geography, University of California, Santa Barbara, Santa Barbara, California

WAYNE HIGGINS
Climate Prediction Center, NOAA/National Centers for Environmental Prediction, Camp Springs, Maryland

(Manuscript received 24 May 2010, in final form 1 October 2010)

ABSTRACT

Extreme precipitation events are among the most devastating weather phenomena since they are frequently accompanied by loss of life and property. This study uses reforecasts of the NCEP Climate Forecast System (CFS.v1) to evaluate the skill of nonprobabilistic and probabilistic forecasts of extreme precipitation in the contiguous United States (CONUS) during boreal winter for lead times up to two weeks.

The CFS model realistically simulates the spatial patterns of extreme precipitation events over the CONUS, although the magnitudes of the extremes in the model are much larger than in the observations. Heidke skill scores (HSS) for forecasts of extreme precipitation at the 75th and 90th percentiles showed that the CFS model has good skill at week 1 and modest skill at week 2. Forecast skill is usually higher when the Madden–Julian oscillation (MJO) is active and has enhanced convection occurring over the Western Hemisphere, Africa, and/or the western Indian Ocean than in quiescent periods. HSS greater than 0.1 extends to lead times of up to two weeks in these situations. Approximately 10%–30% of the CONUS has HSS greater than 0.1 at lead times of 1–14 days when the MJO is active.

Probabilistic forecasts for extreme precipitation events at the 75th percentile show improvements over climatology of 0%–40% at 1-day lead and 0%–5% at 7-day leads. The CFS has better skill in forecasting severe extremes (i.e., events exceeding the 90th percentile) at longer leads than moderate extremes (75th percentile). Improvements over climatology between 10% and 30% at leads of 3 days are observed over several areas across the CONUS—especially in California and in the Midwest.

1. Introduction

Extreme precipitation events are among the most devastating weather phenomena since they are frequently accompanied by hazardous weather such as lightning, hail, heavy snow, strong surface winds, and intense vertical wind shear. Oftentimes, these events are accompanied by flash floods and landslides, which increase the potential for loss of life and property (e.g., Pielke and Downton 2000). For instance, the U.S. National Weather Service registered an average of 65 fatalities per year associated with floods in a 10-yr period (1999–2008). In this study, we refer to extreme precipitation events as those observed at the tails of the frequency distribution regardless of whether or not damage occurs.

Although numerical weather prediction (NWP) has been routine for more than 50 years, tremendous improvements continue to be achieved in the medium range ~ (3–7) days (e.g., Tribbia 1997; Kalnay et al.)
1998). Moreover, advances in NWP model physics and parameterizations, and progress in computing technology and ensemble forecasting have all led to significant improvements at longer leads (Zhu and Toth 2001; Legg et al. 2002; Young and Carroll 2002; Zhu et al. 2002; Lalaurette 2003; Legg and Mylne 2004; Vitart 2004; Zhu 2005; Saha et al. 2006). Despite this, significantly more work is necessary to understand the variability of extreme events and hazards as well as develop useful tools to monitor and predict their impacts (Zhu and Toth 2001; Higgins et al. 2008).

The Climate Forecast System (CFS) is a state-of-the-art global coupled model in operational use at the National Centers for Environmental Prediction (NCEP). The CFS has contributed to a number of significant advances in seasonal prediction (e.g., Saha et al. 2006). An important endeavor was the completion of the CFS reforecasts covering a 28-yr period (1981–2008) and consisting of 15 reforecasts per calendar month out to nine months into the future. While the reforecasts were limited to 15 initial conditions (ICs) per month because of computational constraints, they were designed to make best use of operational data each month and account for the evolution of both the atmosphere and the ocean. While the primary goal of the reforecasts was to develop proper calibrations for operational seasonal forecasts, they also offer the opportunity to investigate additional problems such as forecasts of extreme events.

There are several goals of this study. First, the CFS reforecasts were used to evaluate the forecast skill of extreme precipitation in the contiguous United States using the CFS version 1. This provides a useful benchmark on the forecast skill of extreme precipitation events by the CFS, which can be compared to the skill in future versions of the CFS. The focus is on the boreal winter season and for lead times up to two weeks. Second, the large sample of CFS reforecasts allows the comparison of skill for both nonprobabilistic and probabilistic forecasts of extreme precipitation. Third, previous studies have demonstrated that the Madden–Julian oscillation (MJO; Madden and Julian 1994; Lau and Waliser 2005; Zhang 2005) influences the occurrences of extreme precipitation in the tropics and extratropics of both hemispheres (Mo and Higgins 1998a; Higgins et al. 2000a; Jones 2000; Bond and Vecchi 2003; Carvalho et al. 2004; Jones et al. 2004a; Liebmann et al. 2004; Barlow et al. 2005; Donald et al. 2006; Jeong et al. 2008). Since the MJO is associated with large-scale convective heating in the tropics, tropical–extratropical teleconnections are substantial during its life cycle and, therefore, some studies have detected noticeable impacts on the skill of medium-to-extended range weather forecasts (Ferranti et al. 1990; Lau and Chang 1992; Hendon et al. 2000; Jones and Schemm 2000) and potential predictability (Jones et al. 2004a,b). Thus, this study examines the importance of the MJO on forecast skill of extreme precipitation.

The paper is organized as follows. Section 2 describes datasets, and section 3 compares the occurrences of extreme precipitation in observations and a climate run of the CFS model. Nonprobabilistic forecasts of extreme precipitation are presented in section 4, while probabilistic forecasts are examined in section 5. Discussion and conclusions are presented in section 6.

2. Datasets

The observed occurrence of extreme precipitation events was analyzed with the NCEP Climate Prediction Center (CPC) unified precipitation dataset (Higgins et al. 2000b). Daily averages of precipitation over the contiguous United States (CONUS) were used for the period 1 November–31 March 1981–2008. Bilinear interpolation was used to degrade the original resolution (latitude, longitude) = (1°, 1°) to (latitude, longitude) = (2.5°, 2.5°) for consistency with the grid spacing of the NCEP CFS model.

A comprehensive discussion of the NCEP CFS model and reforecasts is provided by Saha et al. (2006). The NCEP–Department of Energy (DOE) reanalysis and the NCEP Global Ocean Data Assimilation System (GODAS; Behringer 2007) provided atmospheric and oceanic ICs for the reforecasts. NCEP uses the reforecasts to calibrate and evaluate the skill of seasonal forecasts. Each run consists of a full nine-month integration. The reforecasts cover the entire year and are available from 1981 to 2008.

Each ensemble run is based on 15 ICs spanning each month. The first five ICs are on the 9th–13th, the second five ICs on the 19th–23rd, and the last five ICs on the second-to-last day of the month, the last day of the month, and the first, second, and third days of the next month. The dates of ICs were selected to coincide with the generation of real-time atmospheric and oceanic fields and stay within computing limitations. Forecasts of daily precipitation (latitude, longitude) = (2.5°, 2.5°) were used for the period 1 November–31 March 1981–2008. Additional discussions on the development of reforecasts and applications can be found in (Hamill et al. 2004a,b; Hamill and Whitaker 2006).

The climatological characteristics of extreme precipitation in the CFS model were also examined in a Coupled Model Intercomparison Project (CMIP) run of the CFS. Daily averages of precipitation in a 32-yr run were analyzed. Additional details about the CMIP run were discussed in Wang et al. (2005), while the variability and systematic errors in the CFS model were examined in Higgins et al. (2008).
Last, the MJO activity was determined with NCEP–National Center for Atmospheric Research (NCAR) reanalysis (Kalnay et al. 1996; Kistler et al. 2001). Daily averages of zonal wind at 850 hPa (U850) and 200 hPa (U200) were used for the global tropics (15°S–15°N). To complement the analysis, daily averages of outgoing longwave radiation (OLR; Liebmann and Smith 1996) were used to characterize the convective signal associated with the MJO. MJO events were identified following the method described in Jones (2009) and Jones and Carvalho (2009). A total of 70 events were recorded during 1 November–31 March 1981–2008.

3. Extreme precipitation in observations and the NCEP CFS model

To characterize the observed climatological characteristics of extreme precipitation events in the United States during winter, gamma distributions (Wilks 2006) were fitted to each time series of nonzero precipitation in the domain. Two levels of extreme precipitation are defined: daily precipitation exceeding the 75th and 90th percentiles of the gamma frequency distributions. Calculations were completed for each winter month separately and are displayed here as the mean for November–March. The spatial variability and relative magnitudes of the two percentiles is shown in Fig. 1.

The climatological probability of a “75th percentile” (90th percentile) extreme was estimated by counting the number of days when the precipitation exceeded the percentile (November–March mean) divided by the total number of dry and wet days in the period November–March 1981–2008). We recall that the 75th and 90th percentiles were determined from the nonzero values of precipitation and, therefore, the climatological probabilities of extreme precipitation are not necessarily equal to 0.25 and 0.10, respectively. The climatological probability of a 75th percentile extreme (Fig. 2, top) varies across the CONUS, ranging from 0.20–0.25 in the western states and over the Southeast to 0.15–0.20 over...
the central and eastern United States. It is also interesting to note that the climatological probability is 0.10–0.15 in a few regions. The climatological probability of a 90th percentile extreme is 0.075–0.10 over most states east of ~100°W and a few areas in the western states, while it ranges from 0.05 to 0.075 over a large area of CONUS.

The forecast skill of extreme precipitation largely depends on how well the NCEP CFS model simulates the magnitudes and frequency of these events. To evaluate this issue, gamma frequency distributions were also applied to the nonzero time series of precipitation from the 32-yr CMIP run with the NCEP CFS model (1 November–31 March). Figure 3 (top) displays the ratio of the 75th percentile between observations and the CMIP CFS run and indicates that the magnitudes of the extremes in the model are much larger than in the observations (i.e., the CFS model has longer tails in the gamma probability distribution functions than the observations). Only over a few locations does the ratio exceed values of 0.5. Likewise, the ratio of the 90th percentile shows a similar spatial pattern, although it is interesting to note that the ratio is higher than that indicated by the 75th percentile and nears 1.0 in some areas.

The ratios of percentiles described above (observations–CMIP CFS run) were used to adjust the CFS reforecasts. In this case, the percentiles in the CFS were adjusted for each month during the winter season. This method assumes that the percentiles of extreme precipitation in the reforecasts do not vary with lead time and that the statistical properties of the CMIP CFS run and CFS reforecasts are the same. It is important to note that an alternative approach could have been used. Adjustments for the percentiles could have been obtained using the reforecasts and made lead-time dependent. However, the sample sizes available for the estimation of the percentiles were markedly different. About 960 samples were available for each winter month in the CMIP CFS run (~30 days × 32 years), whereas there were about 405 data points in the CFS reforecasts for each month (~15 ICs × 27 seasons).
4. Nonprobabilistic forecasts of extreme precipitation

Nonprobabilistic forecasts of extreme precipitation were developed with the following approach. First, the mean CFS model bias (CFS minus observations) was estimated for each month separately and for lead times from 1 to 28 days. To facilitate the display, Fig. 4 shows the mean model bias averaged during November–March at lead times from 1 to 4 weeks. While the model bias does not change significantly during weeks 1–4, it varies considerably over the CONUS. The largest positive values are observed over the western and northeastern United States, indicating that the CFS overestimates the mean precipitation. In contrast, the CFS underestimates the mean precipitation by about 1 mm day$^{-1}$ or less over the southeast United States. These results are consistent with Higgins et al. (2008), who carried out a detailed comparison of seasonal numbers of wet and dry days in the CFS model and observations.

Based on the results above, two simple adjustments were made to the CFS reforecasts: 1) systematic error correction and 2) percentile adjustment. For each winter season (November–March) during 1981–2008, the mean model bias was subtracted from the forecasts of precipitation with a “one out approach” (i.e., that season was excluded from the computation of mean model bias). Note that the model bias removed varies with the month of the forecast initialization and the lead time. Next, each forecast was analyzed and the extreme event percentile adjusted using the relationship $P_{X_k} = P_{C_k} \times R_{AT_k}$, where $P_{X_k}$ is the adjusted percentile, $P_{C_k}$ is the observed percentile, $R_{AT_k}$ is the ratio of percentiles (Fig. 3), and $k = 1, 2$ for the 75th, 90th percentiles, respectively. A forecast of an extreme event was then identified if the precipitation exceeded the adjusted extreme event percentile.

The nonprobabilistic forecasts therefore consisted of binary pairs of forecasts ($y_i$) and observations ($o_i$), such that occurrences (nonoccurrences) of extremes assumed...
values of 1 (0). For each grid point in the CONUS, the total number of forecasts–observations pairs for each lead time of 1–14 days was 2025 because the CFS reforecasts were initialized in groups of five consecutive days every other five days. The pairs of forecasts–observations were aggregated into a $2 \times 2$ contingency table and several forecast skill measures computed (Table 1; see Wilks 2006 for details).

In the interest of space, only three forecast skill measures are discussed here. Figure 5 shows the forecast bias for the occurrences of 75th percentile extremes at two lead times and indicates that nonprobabilistic forecasts derived from CFS reforecasts tend to underforecast the occurrences of extreme precipitation (unbiased forecasts have bias equal to 1). Although the spatial patterns of the forecast bias do not change appreciably on lead times of 1–7 days, the overall magnitudes do change. Similar spatial patterns and magnitudes are noted in the forecast bias of 90th percentile extremes (not shown).

The false alarm rate ($F$) for the occurrences of 75th percentile extremes at 1-day lead ranges from 0.00 to 0.12 over the CONUS (Fig. 6, top). At 7-day lead, $F$ grows quickly—ranging between 0.02 and 0.18 and maximizing over the western and eastern parts of the CONUS. The Heidke skill score (HSS) is a forecast skill metric frequently used to summarize the results of nonprobabilistic forecasts; it measures the normalized proportion of correct forecasts after eliminating those forecasts that would be correct just by chance (Wilks 2006). Perfect forecasts have $HSS = 1$, $HSS = 0$ for forecasts with no skill, and forecasts worse than the reference have negative values. The HSS of 90th percentile extremes (Fig. 7) is positive over the entire CONUS during week 1 lead time, but drops quickly from $\sim (0.30–60)$ at 1-day lead to $\sim (0.1–0.2)$ at 7-day lead. Similar patterns are seen in HSS for 75th percentile extremes (not shown).

To summarize the results, the meridian of 100°W was chosen to divide the CONUS into two parts (i.e., western and eastern halves). Although arbitrary, this choice consistently separates the maxima in forecast bias and HSS into two regions. Figure 8 shows the mean HSS of 75th and 90th percentile extremes on 1–14-day lead

### Table 1. Contingency table of forecasts and observations and measures of forecast skill.

<table>
<thead>
<tr>
<th>Forecast</th>
<th>Observed</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$a$</td>
<td>$b$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$c$</td>
<td>$d$</td>
<td></td>
</tr>
<tr>
<td>Forecast bias</td>
<td>$B = (a + b)(a + c)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>False alarm rate</td>
<td>$F = b(b + d)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heidke skill score</td>
<td>$HSS = 2(ac - b^2)[(a + c)(c + d) + (a + b)(b + d)]$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of pairs</td>
<td>$n = a + b + c + d$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
times over the two regions. The minimum and maximum HSS values for each lead time and region are also shown. On average, HSS in the western CONUS is about 0.3 at 1-day lead and decreases to 0.1 at 8-day lead for both 75th and 90th percentile extremes (left panels). The spread of HSS in the western CONUS, however, shows that the CFS model has significantly more forecast skill of extremes within the region. In contrast, an interesting behavior is noted in the HSS over the eastern CONUS (right panels). The mean HSS of 75th and 90th percentile extremes are higher than in the western CONUS (differences in mean HSS are statistically significant at a 5% level). On the other hand, the rate of decrease in HSS over the eastern CONUS is higher than in the western half, especially for 75th percentile extremes.

If one arbitrarily considers a level of useful skill when HSS $\approx 0.1$, then the results above suggest that the overall skill of nonprobabilistic forecasts of extreme precipitation derived from CFS reforecasts extend to about 7–8 days’ lead time on average, although substantial spread is noted across the CONUS.

To investigate the importance of the MJO in modulating the forecast skill of extreme precipitation, MJO events were identified with combined empirical orthogonal function (EOF) analysis of equatorially averaged ($15^\circ$S–$15^\circ$N) intraseasonal zonal wind anomalies (200 hPa and 850 hPa) from NCEP–NCAR reanalysis (Jones 2009; Jones and Carvalho 2009). The phase diagram of the first two normalized principal components (PC1, PC2) was used to characterize MJO events according to 1) the phase angle between PC1 and PC2 systematically rotated anticlockwise, indicating eastward propagation at least to the Maritime Continent (phase 5); 2) the amplitude ($PC1^2 + PC2^2$) was always larger than 0.35; 3) the mean amplitude during the

![Forecast Bias (75th Extr) Lead: 1-day](image)

![Forecast Bias (75th Extr) Lead: 7-day](image)

**FIG. 5.** Forecast bias of occurrence of extreme precipitation (75th percentile) at a (top) 1- and (bottom) 7-day lead time. Contour interval is 0.1 and heavy shading indicates bias $>0.6$. 
event was larger than 0.9; and 4) the entire duration of the event lasted between 30 and 90 days. A total of 70 events were recorded during 1 November–31 March 1981–2008. Figure 9 shows composites of OLR anomalies and demonstrates the eastward propagation of the MJO as represented by eight phases, which approximately match the phase definitions used by Wheeler and Hendon (2004).

Nonprobabilistic forecasts of extreme precipitation were validated separately on days in which the MJO was active and inactive according to the observations. In the active cases, the validation was performed on each MJO phase and for lead times from 1 to 14 days, regardless of the state of the MJO in the ICs. This approach differs from an alternative one in which the validation is performed for each MJO phase present in the IC. The former approach was chosen because of the design of the CFS reforecasts (i.e., groups of five consecutive ICs separated by five days) and sampling considerations. In the selected approach, the number of available pairs of [forecasts, observations] in each of the eight phases on 1–14-day lead times varied between 123 and 256 samples. Likewise, the number of pairs during inactive MJO varied between 580 and 594 samples.

To estimate statistical significance between HSS during active and inactive MJO, a resampling technique was developed as follows. Random subsamples of 200 pairs of [forecasts, observations] were withdrawn from the total sample of inactive MJO days. Although extreme events occur infrequently in time, their occurrences still exhibit some degree of spatial correlation. Therefore, the resampling was performed in each grid point in the CONUS separately. The process was repeated 300 times and, for each batch, HSS computed for forecasts of 75th

![Figure 6](image-url)
and 90th percentile extremes. This procedure resulted in frequency distributions of HSS during inactive MJO conditions. Statistical significance was assessed by comparing HSS during active MJO phases with the 95th percentile of the frequency distribution of HSS during inactive conditions.

Figure 10 shows HSS of forecasts of 90th percentile extremes when the MJO was active and in phases 1–8.

**FIG. 7.** The HSS of extreme precipitation (90th percentile) forecasts at 1-, 3-, 5-, and 7-day lead times. Contour interval is 0.1 and heavy shading indicates HSS > 0.3.

**FIG. 8.** The HSS over (left column) the western and (right column) eastern halves of CONUS. Western (eastern) half is defined as grid points west (east) of 100°W. (top) HSS of forecasts of extremes above 75th and (bottom) 90th percentiles. Solid lines represent the mean HSS over the western–eastern domains; upper (lower) dashed lines indicate the max (min) HSS values.
FIG. 9. Phase composites of OLR anomalies. Light (dark) shading indicates positive (negative) anomalies. Contour interval is 2.5 W m$^{-2}$; 0 contours omitted.
To facilitate the display, HSS was averaged on 1–7 days’ lead (week 1). HSS is plotted only over grid points in which statistical significance was 5% or higher relative to forecasts validated during inactive MJO conditions. Several regions in the western, central, and eastern CONUS exhibit large values of HSS (above 0.3). In particular, the results suggest that when the MJO is active and in phases 1 or 8, CFS forecasts of extreme precipitation are especially useful, since high HSS values tend to cluster into spatially connected grid points covering large areas in the western and eastern United States. Statistically significant differences in HSS tend to occur in isolated grid points in phases 2–7, although HSS is high over some large areas in the central CONUS during phases 6 and 7 as well.

While the spatial patterns are more or less the same for HSS of forecasts of 75th percentile extremes (not shown), the absolute magnitudes of HSS are higher for forecasts of 90th percentile extremes. This indicates that the CFS model has better forecast skill for more severe extreme precipitation events in some regions across the CONUS.

Figure 11 shows detailed views of the influence of the MJO on forecast skill for 90th percentile extreme precipitation. Each panel shows the mean and spread of HSS validated on each MJO phase; statistics are computed

**FIG. 10.** The HSS of extreme precipitation (90th percentile) forecasts validated on each phase of the eight MJO phases. The HSS was averaged on 1–7 days’ lead time (indicated in lower right of each panel) and only grid points exceeding the 5% significance level are plotted (see text for further details).
only over grid points that are statistically significant relative to inactive MJO conditions. The first aspect to notice is that \( \text{HSS} \geq 0.1 \) extends to 13–14-day lead times, which contrasts to the overall skill of 7–8 day shown previously (Fig. 8). Another important point is that the maximum HSS validated during MJO conditions is higher than the overall skill (Fig. 8) and extends to longer leads.

A quantitative assessment of the spatial influence of the MJO on the forecast skill of extreme precipitation is shown in Fig. 12 for each MJO phase. The computation was done for both the 75th and the 90th percentile extremes. The solid dots represent the percentage of the CONUS with \( \text{HSS} \geq 0.1 \) averaged over (top) days 1–7 and (bottom) days 8–14. The statistic was computed as the number of grid points with \( \text{HSS} \geq 0.1 \) and significance level relative to inactive MJO conditions \( \geq 5\% \) divided by the total number of points in the domain. The upper and lower tips of the bars show the spread (i.e.,

---

**Fig. 11.** The HSS of extreme precipitation (90th percentile) forecasts during each MJO phase (indicated in top right of each panel). Solid lines represent the average over grid points that are significant at 5% level. Upper (lower) dashed lines indicate the max (min) HSS values. Horizontal axis is lead times in days.
maximum and minimum percentages of the CONUS with HSS $\approx 0.1$. The results clearly show that when the MJO is active, the largest influence on forecast skill of precipitation extremes corresponds to phases 1–2 and 7–8. For 75th percentile extremes, $\sim 15\%$ of the CONUS has HSS $\approx 0.1$ for week 1 lead time (Fig. 12, top left) and up to $\sim 20\%$ if one considers 1–7-day lead times individually. For week 2 lead time, phase 8 stands out with $\sim 15\%$ average and up to $\sim 29\%$ of the CONUS with HSS $\approx 0.1$ for 8–14 days analyzed individually (Fig. 12, bottom left). For 90th percentile extremes and week 1 lead time (Fig. 12, top right), the separation between phases 1 and 2, 7 and 8, and the remaining ones is greater, such that phases 1 and 8 have approximately the same influence on forecast skill. While the percentage of CONUS with HSS $\approx 0.1$ tends to decrease for week 2 lead time (Fig. 12, bottom right), phase 8 has a substantial influence (average of $16\%$ and maximum of $32\%)$.

The results above are consistent with previous studies that identified significant linkages between the activity of the MJO and the variability of precipitation—including occurrences of extreme events. For instance, precipitation anomalies over the west coast of North America show a north–south three-cell pattern. Heavy precipitation in California is accompanied by dry conditions over Washington, British Columbia, and along the southeastern coast of Alaska and reduced precipitation over the subtropical eastern Pacific. Wet (dry) events in California are favored during the MJO phase associated with enhanced convection near 150°E (120°E) in the tropical Pacific (Mo and Higgins 1998a). Extreme events along the west coast of the United States occur at all phases of ENSO, but the largest fraction of these events occurs during neutral winters prior to the onset of El Niño, which tend to be characterized by enhanced tropical intraseasonal activity (Higgins et al. 2000a). Extreme precipitation in California is more common when tropical convective activity associated with the MJO is high, as opposed to periods of quiescent phases of the oscillation (Jones 2000). A slight preference for a higher number of events is observed when convective anomalies associated with eastward-propagating MJOs are located in the Indian Ocean. Significant relationships between the MJO and flooding events in Oregon and Washington during boreal winter have also been shown (Bond and Vecchi 2003). Indeed, the influence of the MJO on precipitation variability and extreme events extends to South America (Carvalho et al. 2004; Jones et al. 2004a; Liebmann et al. 2004), Central America (Barlow and...
Salstein 2006), Australia (Wheeler et al. 2009), and southwest Asia (Barlow et al. 2005).

Since the MJO involves substantial anomalies in tropical convective activity, teleconnections with both hemispheres are significant during its life cycle. The atmospheric responses associated with the convective phase of the MJO in the western-central Pacific are similar to responses due to sea surface temperature (SST) anomalies on interannual time scales (Horel and Wallace 1981). They are characterized by Rossby wave trains of alternating positive and negative geopotential height anomalies propagating from the tropical Pacific toward the mid-latitudes of the Americas (Wallace and Gutzler 1981; Mo and Paegle 2001; Matthews et al. 2004).

5. Probabilistic forecasts of extreme precipitation

Probabilistic forecasts of extreme precipitation were developed in the following manner. As before, for each winter season during 1981–2008, the mean model bias was subtracted from the forecasts of precipitation with a “one out approach.” Next, each group of five consecutive ICs was taken together and used to compute the forecast probability of 75th and 90th percentile extremes. For example, the ensemble members initialized on the 9th–13th of each month were taken as one group. The lead times were considered relative to the ensemble member initialized on the fifth day in the group (i.e., the 13th of each month) and probabilistic forecasts computed for lead times from 1 to 14 days. This approach essentially represents errors in the ICs by taking forecasts initialized at different times but verified at the same time (“time lagged” approach; see Du 2007 for details). The same procedure was followed for the other two groups of five consecutive days of ICs.

Each forecast in the group of five members was analyzed to determine the occurrence or nonoccurrence of 75th and 90th percentile extreme precipitation. The same procedure of percentile adjustment explained before was applied. Thus, for each group of five members, the probabilities of 75th and 90th percentile extreme precipitation were computed as

\[ y_k = \frac{\text{number of members forecasting extreme precipitation}}{\text{total number of forecast members in the group}}. \]  

(1)

Note that since each group has five members, the forecast probabilities have six possible values (0.0, 0.2, . . . , 1.0). Since the CFS reforecasts consisted of groups of five members initialized every other five days, 405 pairs of forecasts–observations were available for each lead time of 1–14 days.

The probabilistic forecasts were validated using the Brier skill score (BSS; Wilks 2006) defined as

\[ \text{BSS} = 1 - \frac{\text{BS}}{\text{BS}_{\text{Ref}}}, \quad \text{BS} = \frac{1}{n} \sum_{k=1}^{n} (y_k - o_k)^2, \]

\[ \text{BS}_{\text{Ref}} = \text{clim}(1 - \text{clim}), \]  

(2)

where BS is the Brier score, \(y_k\) the forecast probability, \(o_k\) the observation (1 = extreme event, 0 = no extreme), and clim the climatological probability of 75th or 90th percentile extreme precipitation values (Fig. 2).

Probabilistic forecasts of 75th percentile extremes at a 1-day lead show improvements of 10%–40% relative to climatological forecasts over large areas across the CONUS with maxima over California and the eastern part of the country (not shown). The forecast skill, however, drops significantly at 3–7-day leads and only a few areas exhibit positive BSS on week 2 lead times. In contrast, Fig. 13 shows BSS of forecasts of 90th percentile extremes. While the skills at a 1-day lead look similar to BSS of forecasts of 75th percentile extremes (not shown), the CFS model has better skill in forecasting severe extremes at longer leads. At a 3-day lead, BSS between 10% and 30% improvements are observed over several areas across the CONUS, especially in California and the Midwest. Most positively, BSS extends to week 2 lead time.

Figure 14 shows the percentage of the CONUS with positive values of BSS of forecasts of extreme precipitation. For a 1-day lead, the CFS model has useful forecast skills of 75th percentile extremes over ~84% of the CONUS, while the percentage is 73% for the 90th percentile extremes. An important aspect, however, is that the areas covered by positive BSS values for both levels of extreme precipitation differ substantially after a 5-day lead time. While BSS of 75th percentile extremes decreases monotonically to near zero on 5–14-day lead times, the percentage of the CONUS covered with positive values of BSS of forecasts of 90th percentile extremes remain above ~9%.

The frequency distributions of BSS values over the CONUS are displayed on Fig. 15. The largest frequency of positive BSS is found between 0% and 5% improvement over climatological forecasts of 75th percentile extremes in both week 1 and week 2 leads (top). For this level of extreme precipitation, 15% and ~7% of the
distribution shows improvements between 5%–10% and 10%–15% during week 1, respectively. For extreme precipitation exceeding the 90th percentile (bottom), the CFS model shows useful skill in a wider range of improvement over climatological forecasts. In particular, we note that improvements over climatological forecasts between 5% and 10% or higher extend to week 2.

The results above show that the skills of probabilistic forecasts of severe extreme precipitation (i.e., exceeding the 90th percentile) cover larger areas of the CONUS, have better improvements over climatology, and extend to longer lead times than forecasts of moderate extreme precipitation (i.e., above the 75th percentile). This offers some exciting perspectives, since weather-related hazards are frequently associated with very heavy precipitation.

The quality of the probabilistic forecasts was further assessed by computing attribute diagrams (Wilks 2006). The two calibration distributions were computed for each grid point in the CONUS: observed frequency of extreme precipitation conditioned on the values of the forecasts and frequency of use of forecasts (or refinement distributions). To summarize the results, the calibration distributions were averaged over all grid points with BSS > 0. Figure 16 shows the average calibration functions of forecasts of 90th percentile extremes. At a 1-day lead, the probabilistic forecasts are moderately calibrated.

Fig. 13. The BSS of forecasts of 90th percentile extreme precipitation. Shading indicates percentage improvement over climatological forecasts. Lead times are shown in the lower right of each panel.
and exhibit characteristics of conditional biases: under-forecast ("dry bias"), the frequency of extremes for forecast probabilities between 0.0 and 0.4; and overforecast ("wet bias"), the observed frequency of extremes for forecast probabilities between 0.5 and 1.0. At lead times of 3 and 5 days, the conditional biases increase and the resolution of the probabilistic forecasts decrease. At a 7-day lead time, the forecasts have poor resolution. It is important to note also that the refinement distributions $p(y_i)$ have positive skewness indicating that the forecasts are occurring most frequently in the climatological probability bin. Forecasts that deviate rarely and quantitatively little from their climatological event probability exhibit little confidence (Wilks 2006). Forecasters in this case would have low confidence to discern deviations from the climatological probabilities. An important caveat is also that the probabilistic forecasts discussed above are based on very small samples of five members per ensemble.

Last, the previous section showed important results demonstrating the influence of the MJO on nonprobabilistic forecasts of extreme precipitation. Unfortunately, the sample size of CFS reforecasts does not allow an easy way to estimate that influence on probabilistic forecasts of extremes. Since the reforecasts were initialized in groups of five ICs spaced every other five days, the number of samples available when the MJO was active is substantially reduced. The number of forecasts–observations pairs varied between 23 and 55 samples for each MJO phase and for lead times of 1–14 days. Therefore, the impact of the MJO on probabilistic forecasts of extreme precipitation was not examined.

6. Discussion and conclusions

This study examined the NCEP CFS forecast skill of extreme precipitation. The focus was on the contiguous United States during winter and for lead times of 1–14 days. A comparison between observations and a climate run of the CFS model indicated that the model realistically simulates the spatial patterns of extreme events over the CONUS, although the magnitudes of the extremes in the model are much larger than in the observations. The evaluation of nonprobabilistic forecasts of extreme precipitation revealed that the CFS model systematically underforecasts the occurrences of extreme precipitation. Nevertheless, HSS of forecasts of two levels of extreme precipitation (75th and 90th percentiles) showed that the CFS model has good skill on week 1 leads and some modest skill at week 2. Large spatial variations in nonprobabilistic forecast skills were also found with maxima over the western and eastern parts of the CONUS. An interesting result is that the mean HSS in the eastern part of the CONUS is statistically higher than in the western half of the country, although the decrease in HSS in the western half is smoother than in the eastern region, especially for 75th percentile extremes. Overall the results suggest that the skills of nonprobabilistic forecasts of extreme precipitation extend to about 7–8 days’ lead time on average.
The influence of the MJO on nonprobabilistic forecasts of extreme precipitation was examined in detail. The results show that the forecast skill is usually higher when the MJO is active than in quiescent periods. The strongest influences on forecast skill are when negative OLR anomalies associated with the MJO are in phase 1 (Indian Ocean) or in phase 8 (Africa, western Pacific, and South America). In these situations, HSS greater than 0.1 extends to 13–14-day lead times. In addition, the MJO has statistically significant influences on the spatial variations of forecast skill of extreme precipitation. Approximately between 10% and 30% of the CONUS has HSS greater than 0.1 on lead times of 1–14 days when the MJO is active.

Probabilistic forecasts of 75th percentile extreme precipitation vary between 10% and 40% improvements over climatology at a 1-day lead to about 0%–5% over a few regions at a 7-day lead. An interesting finding is that the CFS model has better skill in forecasting severe extremes (i.e., those events exceeding the 90th percentile) at longer leads than moderate extremes (75th percentile). Improvements over climatology between 10% and 30% at a 3-day lead are observed over several areas across the CONUS, especially in California and the Midwest. In fact, useful probabilistic forecasts extend to week 2 lead time in some regions.

The forecasts of extreme precipitation derived from CFS reforecasts were developed in this study by correcting systematic errors and differences in the percentiles of the frequency distribution of precipitation. Examination of attributes at individual grid points (not shown) revealed large variations on regional scales. Additional calibrations are necessary to improve the reliability of the forecasts—these will need to be developed on local-to-regional scales.

Extreme precipitation events are usually associated with high impacts on economy and society, and decision-making processes.
makers at federal, state, and local levels need reliable forecasts to assess vulnerability and risk. Ensemble forecasts and probabilistic forecasts, in particular, have great potential to help meet the needs of decision makers. The results from this study show that CFS forecasts of extreme precipitation in the contiguous United States have good skill for week 1 lead times. However, the forecast skill of extreme precipitation drops quickly beyond 7 days; significantly more work is needed to derive reliable forecasts of high impact weather in the extended range.

This study suggests that useful skill at leads of 2–4 weeks might be attained by developing forecasts of opportunity in which significant changes in the mean atmospheric state alter the likelihood of extreme precipitation. One good candidate for forecasts of opportunity is the MJO, since observational and modeling studies have demonstrated important relationships between the oscillation and precipitation variability in the tropics and extratropics (Mo and Higgins 1998b; Mo 1999; Higgins et al. 2000a; Jones 2000; Jones et al. 2004a). Recently, some studies have demonstrated improvements in the forecast skill of the MJO from about 10 days’ lead time (Jones et al. 2000; Seo et al. 2005) to about 2–3 weeks (Rashid et al. 2010; Vitart and Molteni 2010). Probabilistic forecasts of extreme precipitation for leads of 2–4 weeks will probably employ a combination of empirical and numerical forecast model approaches. The authors are currently pursuing this objective and results will be presented elsewhere.

Acknowledgments. The authors thank the support of NOAA’s Climate Program Office and specifically the Climate Test Bed program that supported this work under Grant NA08OAR4310698. NCEP-NCAR reanalysis and OLR data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov. Comments from two anonymous reviewers and Matthew Wheeler (Editor) are greatly appreciated.

REFERENCES