

## Progress Report (05/01/2011 – 04/30/2012)

Title: Improved Extended-Range Prediction through a Bayesian Approach: Exploiting the Enhanced Predictability Offered by the Madden-Julian Oscillation

Award Number: NA10OAR4310250

Program Officer: Annarita Mariotti

PIs: Shang-Ping Xie, Nathaniel Johnson, Steven Feldstein

Objective: *Our goal is to implement a Bayesian framework in a multi-model ensemble (MME) approach for the purpose of enhancing current NCEP/CPC extended-range (6-10-day, 8-14-day, and Weeks 3-4) forecast products over North America.*

### 1. Results and Accomplishments

Our overall strategy employs a framework that incorporates both dynamical forecast model skill and the expected influence of the Madden-Julian Oscillation (MJO) for generating enhanced extended-range forecasts. Over the past year we have made substantial progress in diagnosing the MJO influence on the extratropical circulation over the North American region and in assessing the performance of NCEP's Climate Forecast System, version 2 (CFSv2) forecast model in capturing this MJO influence. These efforts have resulted in a paper (Riddle et al., 2012) that has been submitted for peer review. In addition, through diagnostic analyses and idealized modeling, we have gained insights into the physical mechanisms that link the MJO tropical heating to the midlatitude circulation, which provide the promise of developing new indices that may enhance extended-range forecasts. These efforts have put us in position to test a framework for incorporating model performance and MJO information into extended-range forecasts, which we are pursuing currently.

#### *1.1. Analysis of the MJO impact on the wintertime circulation over North America*

During the past year we have refined analyses of the MJO impact on the wintertime North American circulation, which has resulted in a paper submitted to *Climate Dynamics* (Riddle et al., 2012). In this study, we approximate the continuum of circulation patterns with a finite number of representative patterns through cluster analysis. In particular, we perform k-means cluster analysis on 7-day running mean wintertime (December – March), North American 500 hPa geopotential height anomalies from the NCEP/NCAR reanalysis for the period of 1979-2011. We developed a new method for optimizing the total number of cluster patterns described in Riddle et al. (2012), and determine that seven unique cluster patterns characterize the wintertime circulation over North America (Figure 1). The geopotential height patterns in Figure

1 bear a strong resemblance to well known teleconnection patterns that have a known influence on North American extended-range forecasts.

We find that three of the cluster patterns, clusters 4, 6, and 7, are strongly influenced by the MJO. These three clusters are associated with substantial upper tropospheric zonal wind, precipitation, and surface temperature anomalies over regions of the continental United States (Figure 2). Therefore, understanding how the MJO modifies the frequency of occurrence of these patterns offers potentially valuable information for probabilistic forecasts for lead times beyond one week.

To examine how the frequency of cluster occurrence is modulated in the days and weeks following an MJO event and during different phases of the El Niño/Southern Oscillation (ENSO), we follow a very similar approach to that of Cassou (2008). Like Cassou, we examine how the frequency of a cluster occurring under certain external conditions  $E$  (e.g., 7 days after the MJO is active in phase 1) is elevated or suppressed with respect to the cluster's climatological frequency of occurrence over all 3962 days. The percent change in frequency,  $C$ , is a function of the external conditions,  $E$ , and the cluster number  $i$ :

$$C(i, E) = 100 * \left[ \frac{\frac{N_{i,E} - N_i}{N_E - N_T}}{\frac{N_i}{N_T}} \right], \quad (1)$$

where  $N_T$  is 3962, the total number of days in the study,  $N_i$  is the number of times in the study that cluster  $i$  occurs,  $N_E$  is the number of days when the external conditions,  $E$ , are met, and  $N_{i,E}$  is the number of times that cluster  $i$  occurs under the conditions  $E$ . The percentage  $C(i, E)$  is equal to 100 if cluster  $i$  occurs twice as frequently under the conditions  $E$  as it does in the full record, and is equal to -100 if the cluster never occurs under the conditions  $E$ .  $C$  is calculated for a range of external conditions,  $E$ , including an active MJO during each of the 8 phases and at lead times ranging from zero to 40 days, and for La Niña and El Niño MJO days only. In all of these cases, the full reference climatology is always used for comparison.

Figure 3 shows  $C$  as a function of lag with respect to MJO phase, as defined in Wheeler and Hendon (2004), for clusters 4, 6, and 7. For each cluster pattern in Fig. 3, we consider all MJO episodes and MJO episodes during El Niño and La Niña episodes only. Statistical significance of  $C$  is assessed with a Monte Carlo test whereby 10,000 synthetic, first-order Markov chain cluster pattern time series are generated with transition probabilities that follow the observed transition probabilities, as discussed more thoroughly in Riddle et al. (2012).

Consistent with previous studies, Figure 3 reveals that the MJO exerts a significant influence on the dominant teleconnection patterns of the Pacific/North America region over lags of a few weeks. The occurrence frequency of cluster 4, which resembles the negative phase of the Arctic Oscillation (AO), is elevated significantly with respect to climatology following active MJO episodes in phases 6 and 7 (Fig. 3a). Figure 3b shows that the frequencies of cluster 6, which resembles the positive phase of the Pacific/North American (PNA) pattern, are elevated

significantly with respect to climatology following active MJO episodes in phases 5 and 6. The anomalous frequencies of the negative PNA-like cluster 7 (Fig. 3c), however, are generally opposite to those of cluster 6.

The most salient feature of Figure 3 is the statistically significant anomalous frequencies that exceed 10-20 and sometimes even 30 days after an active MJO in a particular phase, confirming the potential value of the MJO in extended-range forecasts. Moreover, Figure 3 reveals that these anomalous frequencies are strongly modulated by ENSO. In fact, the enhanced probabilities of cluster 6 (cluster 7) following phase 6 (phase 2) of the MJO are completely absent in La Niña (El Niño) years. This may be expected given that convection anomalies over the Pacific are in phase between El Niño (La Niña) and phase 6 (phase 2) of the MJO.

### *1.2. Analysis of CFSv2 forecast performance*

Because our proposed forecasting framework aims to combine information on the expected influence of the MJO with dynamical model forecast performance, we evaluated the ability of the CFSv2 in capturing the expected MJO/cluster pattern relationships. We use a set of 45-day retrospective forecast simulations of 500 hPa geopotential height, which start at six-hour intervals from the period between 1999 and 2010. As in the preceding analysis, each December-March 7-day CFSv2 forecast height field is assigned to one of the seven clusters in Fig. 1 by finding the nearest cluster centroid. Figure 4 shows the anomalous cluster frequencies ( $C$  in equation (1)) for the CFSv2 hindcasts during active MJO phases, as in Fig. 3 for the reanalysis data. In Fig. 4, however, the x-axis for the CFSv2 anomalous frequencies represents the model lead in addition to the lag with respect to the MJO. Figure 4 suggests that the model correctly captures the approximate timing of observed enhanced/suppressed probabilities of Clusters 4, 6, and 7. For example, the model shows a near doubling of the occurrence of Cluster 4 several weeks after an active MJO in phase 6. The model response, however, is somewhat later than in the observations, peaking approximately 28-32 days after the MJO phase 6 episode as opposed to 21-24 days after in the reanalysis. A doubling of the occurrence of Cluster 6 also occurs after the MJO is active in phase 6 in both the model and the reanalysis. In this case, both show the largest probability enhancements occurring at approximately 10-12 days after the MJO event, more than a week before the enhanced probabilities of Cluster 4 (negative AO). The timing of enhancements and reductions in the occurrence of Cluster 7 are also similar between the model and the reanalysis, though the anomalies are weaker and persist for longer in the model.

Though the data do not provide conclusive evidence due to the relatively short record, these results suggest that the CFSv2 model can capture the approximate timing of the tropical/extratropical connections between the MJO and northern hemisphere flow regimes.

We also have begun to examine how CFSv2 performance in extended ranges depends on the initial state of the MJO and ENSO. We find evidence that the CFSv2 performance in weeks 2-4 does, in fact, vary with the initial state of the tropics. We constructed weekly ensemble mean

CFSv2 500 hPa height forecast fields, where the ensemble consisted of 20 consecutive forecasts (five consecutive days of four-times-daily forecasts) for each available winter forecast field. We then calculated the anomaly correlations between the ensemble mean field and the verified height field for weeks 1-5. Finally, we calculated the mean anomaly correlations sorted by lead time and initial state of the MJO and ENSO. In week three, for example, we find that the mean anomaly correlation between the forecast and verified height field varies substantially with the state of the tropics. In particular, during El Niño episodes and when the MJO is in phase 7 the week three mean anomaly correlation is 0.58, which suggests considerable skill under these conditions. In contrast, during La Niña episodes and when the MJO is in phase 6, the mean anomaly correlation is only 0.11, which suggests little model skill.

These preliminary calculations suggest that there may be opportunities beyond week 2 when skillful forecasts are possible, and these opportunities may correspond to particular initial states of the tropics. Current efforts are focused in harnessing this information to quantify model skill information that can be combined with the expected MJO/ENSO influence into a Bayesian extended-range forecasting framework.

### *1.3. Development of new forecast indices and idealized modeling*

There have been two aspects to this research. The first is to develop an index that can be used for extended range (2-4 week) probabilistic forecasts of extratropical teleconnection patterns. One aim of this research is to improve the level of forecast skill beyond that which was obtained by Johnson and Feldstein (2010) who used the Wheeler and Hendon MJO index (Wheeler and Hendon 2004) (MJO index hereafter). The second aim is to perform research with an idealized climate model and with observational data to obtain a physical basis for successful probabilistic forecasts and to further refine the probabilistic forecast methodology.

#### *1.3.1. 2-4 Week Forecasts*

The probabilistic forecast methodology combined both the MJO index and the Rossby wave source (RWS) of Sardeshmukh and Hoskins (1988). It was found that probabilistic forecasts of midlatitude teleconnections that are conditioned upon the phase and amplitude of the MJO index along with a temporally and spatially smoothed RWS substantially improves the 2-4 week forecast skill of the midlatitude teleconnection patterns. After the smoothing, an empirical orthogonal function (EOF) analysis of the RWS was performed for a range of latitudes centered on the subtropics. As shown in Sardeshmukh and Hoskins (1988), the RWS can be separated into four components. A series of tests found that the component associated with the advection of the climatological absolute vorticity by the anomalous divergent wind resulted in a major improvement in forecast skill, whereas the other three components had very limited influence on the forecast skill. (Other tests involving Sverdrup balance, various metrics of tropical convection, transient eddy driving, and other indices based upon the RWS did not yield significant improvement in forecast skill.) With regard to the above smoothing of the RWS, and

the selection of a range of latitudes for the EOF analysis, both steps were found to be crucial. This is because both the smoothing and the selection of the appropriate latitudinal band filter out the impact of midlatitudes transient eddies and isolate the excitation of slowly propagating midlatitude Rossby wave trains by tropical convection.

### *1.3.2. Modeling*

For the modeling component of this study, the dynamical core of NOAA/GFDL climate model was used. The background flow corresponded to that of the Northern Hemisphere winter climatology, upon which MJO-like tropical heating was added. The response to the tropical heating for all eight phases of the MJO index was evaluated, with the focus being on the transient evolution during the first few days. For all eight MJO phases, it was found for the first two days that the response to the model's tropical heating was dominated by two quantities; (1) the advection of the climatological absolute vorticity by the anomalous divergent wind (term 1 hereafter) and (2) the horizontal divergence term, i.e., the climatological absolute vorticity multiplied by the divergence of the anomalous horizontal wind vector (term 2 hereafter). For the first 12 hours of each model integration, term 2 was larger than term 1. During the following 6 hours, term 2 declined relative to term 1, which resulted in term 1 remaining dominant until about 48 hours into the integration. The decline of term 2 was found to occur through two different processes both driven by term 1. These are (i) the alteration of the absolute vorticity field by term 1 and (ii) the inducement of Rossby wave dispersion followed by the reduction of term 2 via thermal wind adjustment, also by term 1. After day 2, the rotational component of the wind field becomes large, and Rossby wave dispersion leads to the excitement of midlatitude teleconnection patterns such as the PNA and North Atlantic Oscillation (NAO). The results of these idealized climate model integrations show that the tropical heating impacts midlatitude teleconnections via term 1. These results also motivated our choice to use term 1 in our 2-4 week probabilistic forecast model of midlatitude teleconnections, as described above. Furthermore, the spatial and temporal filtering that was applied to the RWS, along with the chosen range of latitudes for the EOF analysis, was based on the structure of the RWS in the climate model integrations.

### *1.3.3. Diagnostics with Observational Data*

A diagnostic analysis with NCEP-NCAR reanalysis data was performed to examine the spatial and temporal evolution of the RWS associated with each MJO phase. Term 1 was found to play a dominant role. Furthermore, the spatial structure of term 1 closely matched that of the climate model calculations. This finding provided further support for our choice to use term 1 in the probabilistic forecast model.

### *References*

Johnson, N. C., and S. B. Feldstein, 2010: The continuum of North Pacific sea level pressure patterns: Intraseasonal, interannual, and interdecadal variability. *J. Climate*, **23**, 851-867.

Sardeshmukh, P. D. and Hoskins, B. J. 1988: The generation of global rotational flow by steady idealized tropical divergence. *J. Atmos. Sci.*, **45**, 1228-1251

Wheeler, M. C., and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917-1932

## **2. Peer reviewed publications**

Riddle, E. E., M. Stoner, D. Collins, S. B. Feldstein, N. C. Johnson, and M. L'Heureux, 2012: The impact of the MJO on wintertime circulation anomalies over the North American region. Submitted to *Climate Dynamics*.

## **3. Seminars, presentations, and conference papers**

### *Conference papers:*

Johnson, N., E. Riddle, M. Stoner, S. Feldstein, D. Collins, and M. L'Heureux, 2011: Toward a framework for incorporating MJO and ENSO information into CPC probabilistic extended range forecasts. *Extended summary, 36<sup>th</sup> NOAA Annual Climate Diagnostics and Prediction Workshop*, NOAA's National Weather Service, 204-210, [available online at <http://www.nws.noaa.gov/ost/climate/STIP/36CDPW/36cdpw-njohnson.pdf>].

### *Conference and workshop presentations:*

Riddle, Emily: "The Impact of MJO and ENSO on Wintertime Weather Regime Frequency over the Pacific/North American Region." 36<sup>th</sup> NOAA Annual Climate Diagnostics and Prediction Workshop, 3-6 October 2011, Fort Worth, TX.

Johnson, Nat: "A Framework for Incorporating MJO, ENSO, and Storm Track Variability into CPC Probabilistic Extended Range Forecasts." 36<sup>th</sup> NOAA Annual Climate Diagnostics and Prediction Workshop, 3-6 October 2011, Fort Worth, TX.

Johnson, Nat: "Toward a framework for incorporating MJO and ENSO information into probabilistic extended range forecasts." Global Research Laboratory (GRL) Korea-US joint workshop on Global Monsoon Variability and Change, 9-10 January 2012, Honolulu, HI.

Johnson, Nat: "Toward a framework for incorporating MJO and ENSO information into probabilistic extended range forecasts." 92<sup>nd</sup> AMS Annual Meeting, 22-26 January 2012, New Orleans, LA.

### *Seminars:*

Johnson, Nat: "The impact of the MJO on wintertime circulation anomalies over the North American region." 3 April 2012, Department of Earth and Ocean Sciences, University of British Columbia, Vancouver, British Columbia, Canada.

#### 4. Future Work

- We now are in position to combine the information of CFS forecast skill and the expected MJO influence on the midlatitude circulation through Bayes' theorem to generate extended-range temperature and precipitation forecasts for the North America region. Currently, we are working on a prototype framework on the basis of the CFSv2 hindcasts from 1999-2010. We expect to complete testing of this framework within the next few months.
- Upon satisfactory refinement of this prototype, we plan to extend the framework for multi-model ensemble (MME) forecasting. We shall first test this multi-model forecast system with archived hindcasts from up to eight different model hindcast datasets from the Intraseasonal Variability Hindcast Experiment (ISVHE), currently available at the University of Hawaii, courtesy of Dr. June-Yi Lee. We anticipate completion of testing and refinement of this framework in the winter of 2013.
- After testing the Bayesian MME forecast system with the ISVHE models, we plan to apply the same system to NCEP forecast models (CFSv2 and GEFS). In this phase, transitioning the system to operations shall become the priority.
- Within the next year we plan to write up two papers for publication based on this research. One paper will present the probabilistic forecast model. The other paper will present results from the climate model calculations along with the supporting results from the diagnostic analysis of observational data.

#### 5. PIs Contact Information

Dr. Shang-Ping Xie  
IPRC, Univ. of Hawaii  
[xie@hawaii.edu](mailto:xie@hawaii.edu)  
Ph: (808) 956-6758

Dr. Nat Johnson  
IPRC, Univ. of Hawaii  
[natj@hawaii.edu](mailto:natj@hawaii.edu)  
Ph: (808) 956-2375

Dr. Steven Feldstein  
Penn State Univ.  
[sbf1@meteo.psu.edu](mailto:sbf1@meteo.psu.edu)  
Ph: (814) 865-7042

#### 6. Budget for Coming Year

UH portion: \$74,100  
PSU portion: \$?????

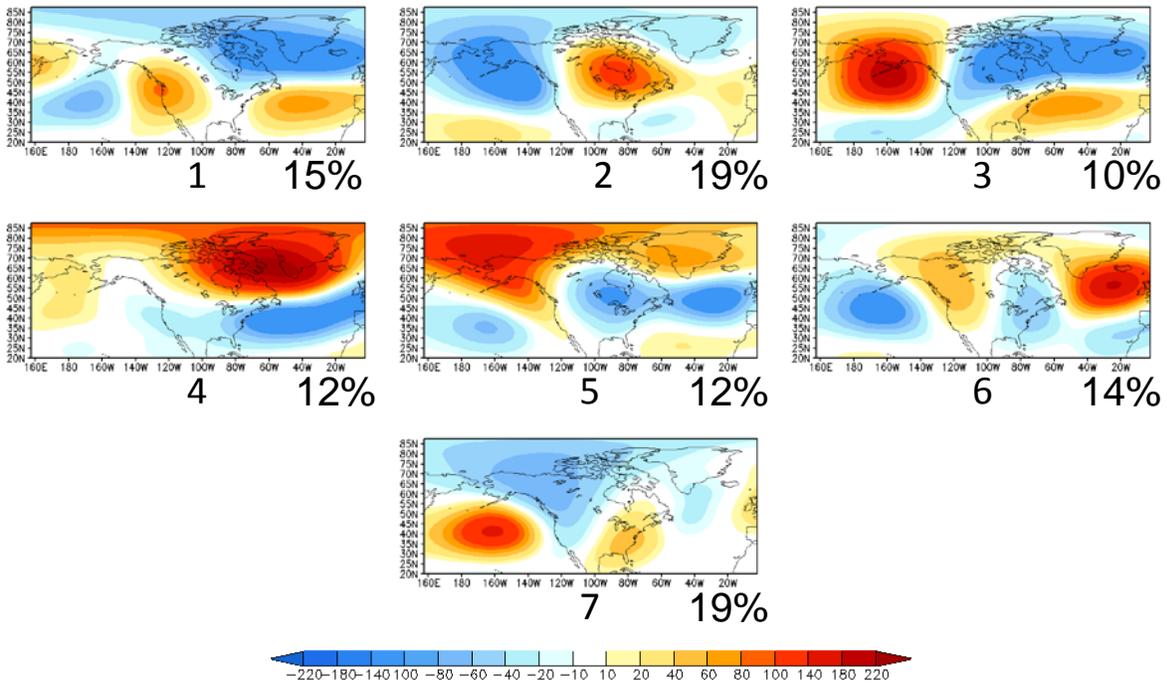
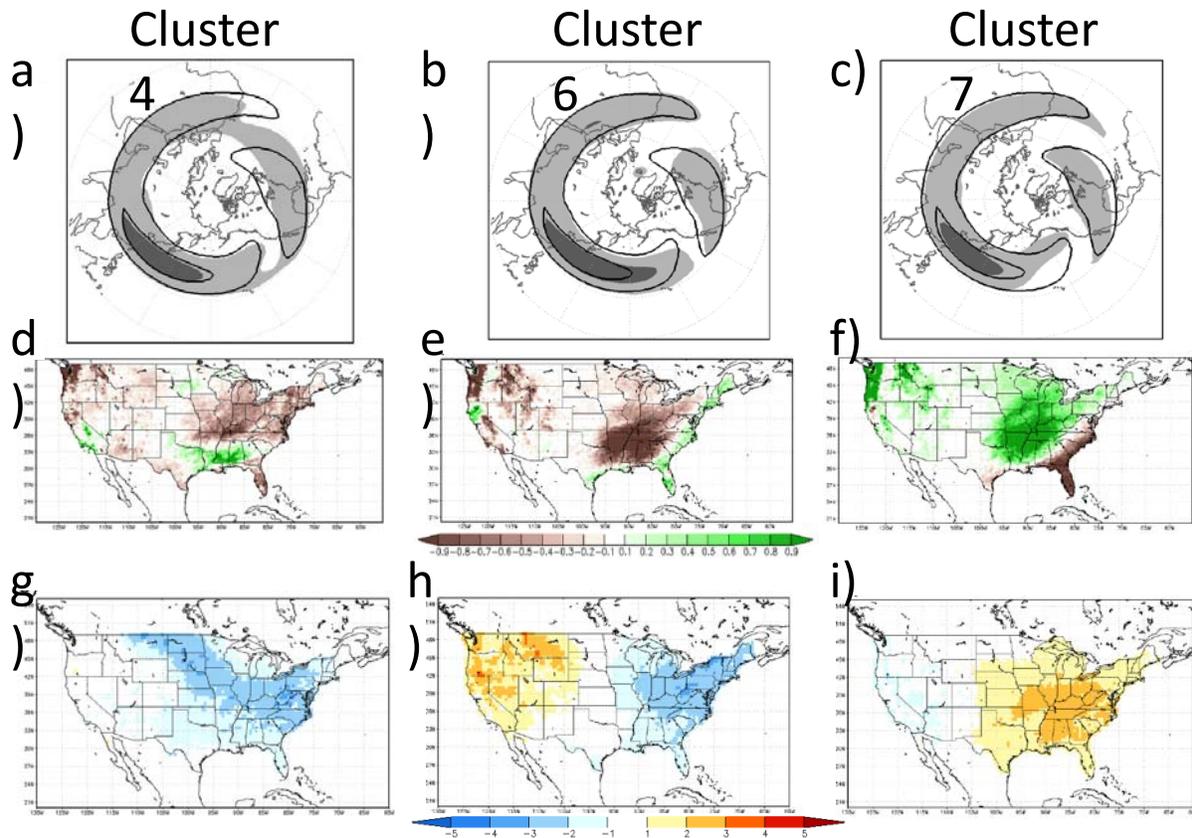
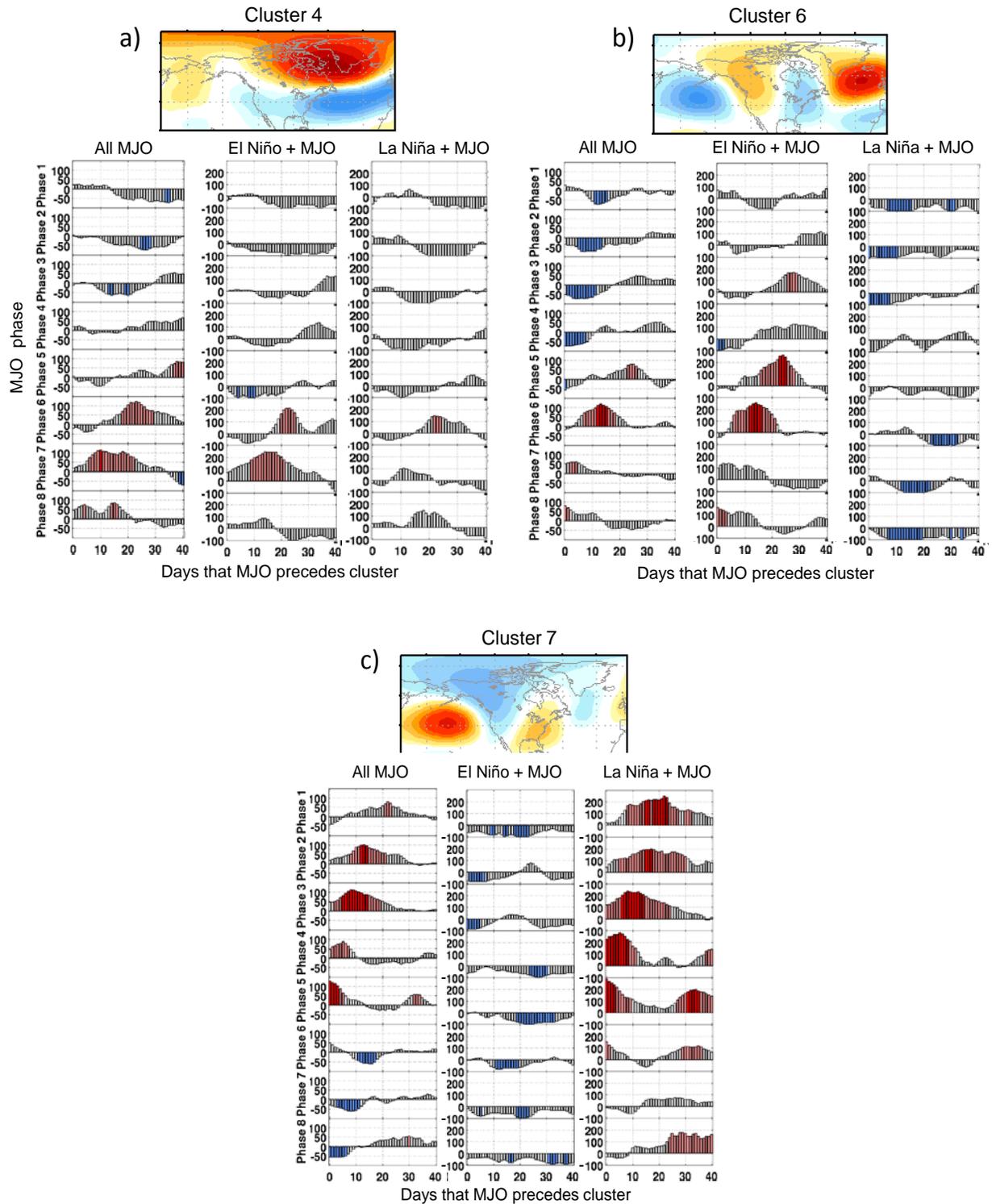


Figure 1. K-means cluster patterns of 500 hPa geopotential height anomalies (m) over the extended North America region. The percentages at the bottom right of each map correspond to the frequency of occurrence of the particular cluster pattern over the 1979-2011 period.

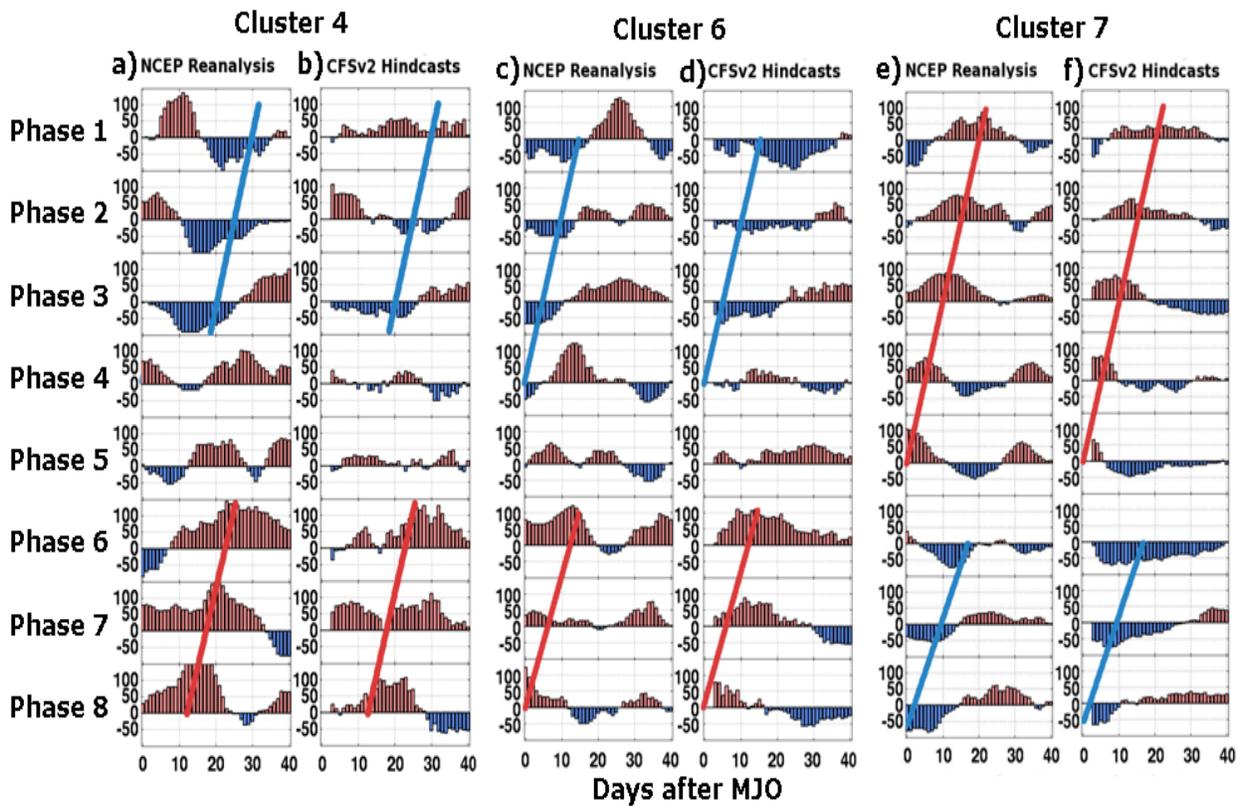


**Figure 2.** (a) – (c) Shading shows 25 m/s (light grey) and 50 m/s (dark grey) contours for a composite of 200 hPa zonal winds for all days in Clusters 4, 6, and 7 compared with the same two contours for the winter (DJFM) climatology (solid lines). (d) – (i) Composite of (d,e,f) precipitation and (g,h,i) temperature anomalies over the United States for all winter days in Clusters 4, 6, and 7. Zonal wind data are from the NCEP reanalysis, surface temperature composites are derived from the gridded daily cooperative dataset of Janowiak et al. (1999), and the precipitation composites are derived from the Climate Prediction Center Merged Analysis of Precipitation (CMAP) (Xie and Arkin 1997).



**Figure 3.** Anomalous frequency of occurrence (%) of cluster patterns (a) 4, (b) 6, and (c) 7 as a function of MJO phase (rows) and time lag (x-axis of each column). Values of 100 mean twice as frequent as climatology, whereas values of -100 mean no occurrence of the cluster pattern. Anomalous frequencies are calculated for all MJO episodes (left column), and MJO episodes

during El Niño (center column) and La Niña (right column) episodes. Note that the scaling for the El Niño and La Niña anomalous frequencies is twice that of all MJO days. Filled bars signify anomalous frequencies that are statistically significant above the 95% confidence level based on a first-order Markov chain null hypothesis. Dark red shading represents enhanced probabilities that are locally significant at a 98.5% level which is the level needed to control the false discovery rate at 15%.



**Figure 4.** (a) Anomalous frequency of occurrence for Cluster 4 based on the NCEP reanalysis 1999-2010. (b) Same as (a) except for the CFSv2 hindcasts. (c) and (d) Same as (a) and (b), except for Cluster 6. (e) and (f) Same as (a) and (b) except for Cluster 7. CFSv2 hindcasts are for 0-40 days lead after the model initialization on day 0 during an active MJO event. For example, day 11 refers to the week-2 forecast. Changes in occurrence are calculated with respect to the lead-dependent model climatological frequencies (as in Figure 12 of Riddle et al. 2012). All positive days are shaded light red, while all negative days are shaded light blue. Slanted red (blue) lines approximate maxima in enhanced (suppressed) probabilities for the full reanalysis record (1981-2010).