

## A Reconstruction of Madden–Julian Oscillation Variability from 1905 to 2008

ERIC C. J. OLIVER AND KEITH R. THOMPSON

*Department of Oceanography, Dalhousie University, Halifax, Nova Scotia, Canada*

(Manuscript received 19 March 2011, in final form 29 August 2011)

### ABSTRACT

The most widely accepted characterization of the Madden–Julian oscillation (MJO) is the bivariate index developed by Wheeler and Hendon. This index relies in part on satellite-based observations of outgoing longwave radiation and thus is not defined for the presatellite era. The MJO is known to have a strong signature in surface pressure, and daily measurements of this variable are available as far back as the late nineteenth century. This study undertakes a statistical reconstruction of the Wheeler and Hendon MJO index from 1905 to 2008 based on tropical surface pressures estimated recently by the twentieth-century reanalysis project. The temporal and spectral properties of the reconstructed index are first shown to be consistent with the Wheeler and Hendon index over the common period (1979–2008). The reconstructed index is then validated over the earlier period (1905–1978) by examining its relationship with cloud cover, surface wind, precipitation, and sea level. These relationships are shown to be consistent with corresponding results obtained from the Wheeler and Hendon index over the shared period and stable over the earlier period. Finally, a simple damped harmonic oscillator model is used to gain new insights into the predictability of the MJO index and also demonstrate consistency between the reconstructed index and the Wheeler and Hendon index. These results give confidence in the validity of the historical reconstruction of the MJO index over the last century.

### 1. Introduction

The Madden–Julian oscillation (MJO) is the dominant mode of intraseasonal variability of the tropical atmosphere. The original studies on what would later become known as the MJO identified a “40–50-day oscillation” in sea level pressure and upper- and lower-level zonal wind in the tropics (Madden and Julian 1971, 1972). Since then, the MJO has been quantified in many different ways. For example, Shinoda et al. (1998) used EOF analysis of tropical outgoing longwave radiation (OLR) and Jones et al. (2004) used EOF analysis of global OLR and zonal wind at low and high levels. Filtering of OLR over a select band of wavenumbers and frequencies has also been used to characterize the MJO (Wheeler and Kiladis 1999; Wheeler and Weickmann 2001). The index presented by Wheeler and Hendon (2004, hereafter WH04), is now widely used to quantify the MJO. It is based on the first two principal components of filtered tropical fields of OLR and zonal wind at heights of 200 and 850 hPa.

The Wheeler and Hendon index, henceforth  $I_t^{\text{WH}}$  where  $t$  is a time index, is not defined for the presatellite era (i.e., before June 1974). As noted above, sea level pressure has a strong relationship with the MJO (e.g., Donald et al. 2006) and observations have been made by land- and ship-borne sensors for centuries. This raises the possibility of extending the Wheeler and Hendon MJO index to cover the presatellite era. Such a reconstruction would be useful in quantifying the long time-scale (i.e., decadal) variability of the MJO and its connections with other indices (e.g., ENSO) and environmental variables (e.g., Australian precipitation). Pohl and Matthews (2007) generated an MJO index from 1950 to 2005 using an EOF analysis of upper- and lower-level zonal winds from the National Centers for Environmental Prediction (NCEP) reanalysis fields. In addition, Jones and Carvalho (2009) inferred the intensity of MJO activity over the period 1880 to 2008 using a simple Markov chain model forced with monthly SST but did not provide an MJO index of the same form as the Wheeler and Hendon index.

We have reconstructed  $I_t^{\text{WH}}$  from 1905 to 2008 by using time series of surface air pressure from a long reanalysis (Compo et al. 2011) as inputs to a multivariate linear regression model. We based the selection of the pressure inputs to the regression model on (i) the strength of their

---

*Corresponding author address:* E. C. J. Oliver, Department of Oceanography, Dalhousie University, 1355 Oxford St., Halifax B3H 4R2, Canada.  
E-mail: eric.oliver@dal.ca

relationship with the MJO, (ii) the decorrelation length scales of the reanalysis pressure, and (iii) the quality of the reanalysis pressures. We show that the historical reconstruction  $\mathbf{I}_t^{\text{HR}}$  has similar temporal and spectral properties to  $\mathbf{I}_t^{\text{WH}}$  over the shared period, including measures of predictability based on a stochastically forced, damped harmonic oscillator model. The variability of  $\mathbf{I}_t^{\text{WH}}$  over the earlier period is also validated by comparing  $\mathbf{I}_t^{\text{WH}}$  with independent environmental variables including its relationship with cloud cover, surface wind, precipitation over Australia, and sea level from tide gauges in the western Pacific, the Gulf of Carpentaria, and along the Pacific coast of the Americas. We have attempted to recreate the analyses presented in WH04 for the pre-1979 period where possible (e.g., using measurements of cloud cover that date back to 1952 where WH04 used measurements of OLR that date back to 1974).

We chose to include only pressure in the regression model. It could be argued that including zonal wind at upper and lower levels could provide better model results. We chose surface pressure for two reasons. First, to the extent that the relationship between reanalysis surface pressure and wind is linear, adding winds will have little impact on the performance of the regression model. Second, we argue that any gains provided by the nonlinear relationship of winds with pressure would be offset by the introduction of additional errors due to the unreliability of reanalysis winds (especially upper-level winds) considering the scarcity of surface observations that have been assimilated into the reanalysis.

The outline of the paper is as follows. The data are described in section 2, and the regression model used to estimate the MJO is presented in section 3. The internal variability of the reconstructed index, including connections with several independent environmental variables, is examined in section 4. The predictability of the MJO is examined in section 5 using a stochastically forced, damped harmonic oscillator model. Conclusions and discussion are presented in the final section.

## 2. Data and preliminary analysis

The Wheeler and Hendon index and the surface pressure reanalysis used for its historical reconstruction are described in this section. This is followed by a description of long records of observed cloud cover, surface wind, Australian precipitation, and sea level from various tide gauge locations that are used to assess the reliability of the reconstructed index over the last century. These variables were chosen because they have a demonstrated connection with the MJO and have record lengths extending before 1979 (i.e., the period over which the reconstructed index is trained on  $\mathbf{I}_t^{\text{WH}}$ ).

### a. The Madden–Julian oscillation index $\mathbf{I}_t^{\text{WH}}$

Daily values of the bivariate index  $\mathbf{I}_t^{\text{WH}}$  were obtained from the Government of Australia Bureau of Meteorology for the inclusive period 1974–2008. Because of a gap in the instrumental record for part of 1978, we have focused on values since 1979 to have a continuous time series. The two components of the bivariate index  $\mathbf{I}_t^{\text{WH}}$  have most of their energy at periods between about 30 and 90 days. Some variability does exist at higher frequencies, which is believed to be due to the signature of Kelvin waves and equatorial Rossby waves in the wind and OLR fields used to generate  $\mathbf{I}_t^{\text{WH}}$  (Roundy et al. 2009). Given that our goal is to reconstruct the MJO index from surface pressure measurements that may not share these higher-frequency components, we have low-pass filtered  $\mathbf{I}_t^{\text{WH}}$  with a cutoff period of 10 days.

### b. Global surface pressure

Daily fields of surface pressure were obtained from the twentieth-century reanalysis project (20CR) (Compo et al. 2011). For reasons related to data scarcity, which will be discussed in more detail later, we have restricted our reconstruction to the period 1905–2008 inclusive.

The reanalysis was carried out on a global grid with a resolution of  $2^\circ$  and was based on assimilating surface pressure observations, monthly sea surface temperatures, and sea ice distribution using an ensemble Kalman smoother. Other than the inclusion of satellite-derived sea surface temperature (SST) observations from October 1978 (Rayner et al. 2003) there is no abrupt change in data quality over the analysis period. The ensemble has 56 members. We have estimated the pressure at time  $t$  at grid point  $i, j$  (with zonal and meridional indices ranging from  $i = 1 \dots 180$  and  $j = 1 \dots 91$ , respectively) by the ensemble mean pressure  $p_{ijt}$ . The reanalysis also supplies the standard deviation of the ensemble  $\Delta p_{ijt}$ , which is used later to assess the quality of the reanalysis data.

Seasonal and interannual variability of the  $p_{ijt}$  series were removed following the procedure outlined in WH04. Specifically, the annual cycle and its first two harmonics were removed using least squares. The time mean of the previous 120 days was also removed. Following the recommendation of Gottschalck et al. (2010), the only difference between the method applied here for removing low-frequency variability and that presented in WH04 is that the linear contribution of El Niño–La Niña was not removed. The pressure series were also low-pass filtered with a cutoff period of 10 days. The deseasonalized and filtered pressures are henceforth denoted by  $p_{ijt}^a$ .

### c. Cloud cover and surface wind

Observations of global cloud cover and surface wind over the ocean were obtained from the Extended Edited

Synoptic Cloud Report (EESCR) dataset from 1952 to 2008 (Hahn and Warren 1999). This dataset consists of 71 million ship-based synoptic weather reports from the International Comprehensive Ocean–Atmosphere Dataset [ICOADS; Worley et al. (2005)] that have been quality controlled to remove or correct erroneous and incomplete reports. Each data record includes a date (year, month, day, and hour) and measurements of total cloud cover, wind speed, and wind direction. All records include measurements of cloud cover but some contain missing values for the other variables.

#### d. Precipitation over Australia

Daily time series of precipitation over Australia recorded by 37 stations north of 30°S were obtained from the National Oceanic and Atmospheric Administration (NOAA) National Climatic Data Center (NCDC) Global Daily Climatology Network from 1905 to 2001. In addition, a daily time series of precipitation was obtained for Booby Island, Queensland, Australia (1908–2001), from the Australian Bureau of Meteorology. Only stations with at least 50% data availability over each of three predefined periods (1905–1949, 1950–1978, and 1979–2001) were used. This reduced the number of stations from 38 to 30 (see Table 1 for details).

#### e. Sea level from tide gauges

Hourly time series of sea level measured by tide gauges in the western Pacific, the Gulf of Carpentaria, and along the west coast of the Americas were obtained (see Table 2 for details). These locations were chosen in part based on data availability over the last century and also because they lie in regions known to have a connection with the MJO (Oliver and Thompson 2010, 2011, and references therein). For each tide gauge, the inverse barometer effect was removed using the daily twentieth-century reanalysis mean sea level pressure interpolated to the tide gauge location. The adjusted sea level was detided using a simple Doodson X0 filter (Doodson 1928), detrended, and averaged to daily means. The annual cycle and its first two harmonics were removed using least squares.

### 3. Reconstructing the MJO

The MJO index has been reconstructed from 1905 to 2008 based on the following multivariate linear regression of  $\mathbf{I}_t^{\text{WH}}$  onto  $m$  predictors:

$$\mathbf{I}_t = \boldsymbol{\beta} \mathbf{p}_t + \boldsymbol{\epsilon}_t, \quad (1)$$

where  $\mathbf{I}_t$  is the observed bivariate index (a column vector of length two),  $\boldsymbol{\beta}$  is a  $2 \times m$  matrix of regression coefficients,

TABLE 1. The long Australian precipitation records used for validation of the reconstructed MJO index. The numbers in the final three columns indicate the percent completeness of the time series over 1905–49 (P1), 1950–78 (P2), and 1979–2001 (P3).

Station	Lon (E)	Lat (S)	P1	P2	P3
Booby Island	141°54'	10°36'	100	92	87
Darwin Airport	130°52'	12°23'	100	100	96
Weipa	141°54'	12°40'	94	72	95
Broome Airport	122°13'	17°57'	100	99	96
Halls Creek Airport	127°39'	18°13'	100	100	96
Tennant Creek Airport	134°10'	19°37'	100	100	96
Burketown	139°31'	17°43'	92	100	89
Georgetown	143°33'	18°17'	100	100	96
Cairns Airport	145°45'	16°52'	100	100	96
Willis Island	149°58'	16°17'	71	62	96
Carnarvon Airport	113°40'	24°52'	100	97	96
Port Hedland Airport	118°37'	20°22'	99	96	96
Newman	119°42'	23°21'	100	75	96
Alice Springs	133°53'	23°47'	100	100	96
Mt. Isa Airport	139°28'	20°40'	99	52	96
Richmond	143°7'	20°43'	100	99	95
Longreach Airport	144°16'	23°25'	100	100	96
Mackay	149°13'	21°7'	100	100	96
Gladstone	151°15'	23°51'	100	100	88
Geraldton Airport	114°41'	28°46'	100	95	96
Meekatharra Airport	118°31'	26°36'	96	95	96
Marree	138°3'	29°38'	100	99	95
Birdsville	139°21'	25°53'	100	94	96
Tibooburra	142°1'	29°25'	96	100	95
Thargomindah	143°49'	27°58'	100	100	96
Charleville Airport	146°16'	26°23'	100	100	96
St. George	148°34'	28°1'	100	100	96
Inverell	151°7'	29°46'	100	100	96
Tewantin	153°1'	26°22'	99	100	95
Yamba	153°21'	29°25'	100	100	96

$\mathbf{p}_t$  is a  $m \times 1$  vector of predictors at time  $t$ , and  $\boldsymbol{\epsilon}_t$  is an error term corresponding to measurement error and effects not included in the model. The estimated regression coefficients  $\boldsymbol{\beta}$  were based on least squares fitting the model to the Wheeler and Hendon index; the MJO index was reconstructed from 1905 to 2008 using  $\mathbf{I}_t^{\text{HR}} = \boldsymbol{\beta} \mathbf{p}_t$ .

The MJO is a propagating phenomenon and so including lagged pressures as predictors in the regression model would reduce the mean square error. However, allowing for lags of up to 20 days or more could lead to an overly complex regression with too many predictors and an overfit model. Instead of adding lagged variables we have simply included the Hilbert transform of each gridded pressure in the set of predictors. [The Hilbert transform was calculated in the frequency domain using the algorithm provided by Oppenheim et al. (1989).] For quasi-periodic predictors with regression coefficients that do not vary much over the frequency band that contains the spectral peak (i.e., the MJO band), it can be shown that lagged predictors can be replaced by the

TABLE 2. The long sea level records used for validation of the reconstructed MJO index. The numbers in the fourth, fifth, and sixth columns indicate the percent completeness of the time series over 1905–43 (P1), 1944–78 (P2), and 1979–2008 (P3). Dashes indicate no data. The numbers in the final column indicate the time lag with respect to the MJO index used when generating the composites in section 4b. The Weipa record was obtained from Maritime Safety Queensland, Queensland Government, Australia, and all other records were obtained from the Hawai'i Sea Level Center.

Tide Gauge	Lon (E)	Lat (N)	P1	P2	P3	Lag (days)
Malakal	134°28'	7°19'	33	23	98	0
Guam	144°39'	13°26'	—	83	87	0
Kwajalein	167°43'	8°43'	—	89	100	0
Pago Pago	189°18'	−14°16'	—	82	95	0
Kanton Island	188°16'	−2°48'	—	71	79	0
Kiritimati	202°31'	1°59'	—	54	79	0
Weipa	141°52'	−12°39'	—	19	74	0
La Jolla	242°45'	32°52'	48	89	95	88
San Diego	242°49'	32°43'	94	96	100	88
Ensenada	243°22'	31°51'	—	55	34	92
Quepos	275°49'	9°24'	—	36	52	61
Balboa	280°26'	8°58'	94	100	58	58
Buenaventura	282°53'	3°53'	—	66	66	59
Tumaco	281°15'	1°50'	—	66	60	60
La Libertad	279°4'	−2°12'	—	80	97	58
Antofagasta	289°36'	−23°38'	—	90	84	70
Valparaiso	288°22'	−33°1'	—	84	78	74

predictors at zero lag and their Hilbert transforms. This approach effectively provides sine- and cosine-like time series for each pressure series, which gives freedom to the regression to choose a linear combination to achieve the best lag relationship with  $\mathbf{I}_t^{\text{WH}}$ . Including the Hilbert transforms doubles the number of predictors, but it does improve significantly the skill of the regression model (see appendix A).

#### a. Location of pressures chosen for inclusion in the regression model

Much of the tropics exhibits a strong relationship between surface pressure and the MJO index (Fig. 1). To limit the number of predictors in the regression model,

and thus avoid overfitting, we have taken into account the strength of the linear relation of the candidate predictor and the MJO as well as the quality of the hindcast pressure (see appendix A). These considerations led to an exclusion zone, restricting the available locations from which predictors could be chosen (Fig. 1, thick black line). Twelve locations were then chosen (Fig. 1, white circles) with preference given to regions with a stronger relationship with  $\mathbf{I}_t^{\text{WH}}$  and higher data quality while also taking into account meridional and zonal decorrelation length scales. Details are given in appendix A. Pressures from these 12 locations (and their corresponding Hilbert transforms) form the set of predictors for the regression model (thus  $m = 12 \times 2 = 24$ ).

#### b. The reconstructed historical index

The regression of  $\mathbf{I}_t^{\text{WH}}$  onto the 24 predictors for 1979–2008 accounts for 69% of the total variance of  $\mathbf{I}_t^{\text{WH}}$ , i.e., the sum of the variances of its two components. Note there remains a certain degree of collinearity amongst the predictors despite the precautions taken to choose them. Therefore, the magnitude of the regression coefficients do not by themselves indicate the importance of an individual predictor.

To reconstruct the MJO index from 1905 to 2008 we used the regression coefficients estimated by fitting the model over the 1979–2008 period. The individual components of the index were then normalized to each have unit variance over the 1905–2008 period. This long standardized reconstructed index is denoted by  $\mathbf{I}_t^{\text{HR}}$  where HR stands for “historical reconstruction.” Plots of  $\mathbf{I}_t^{\text{WH}}$  and  $\mathbf{I}_t^{\text{HR}}$  over a common 3-yr period (Fig. 2) show the new index predicts well the amplitude and timing of  $\mathbf{I}_t^{\text{WH}}$ . For the 1979–2008 period, over which the regression is trained, the correlation between the first component of  $\mathbf{I}_t^{\text{WH}}$  and the first component of  $\mathbf{I}_t^{\text{HR}}$  is 0.83; for the second component the correlation is 0.84. For the 1974–78 period, over which the regression is not trained, and satellite SSTs were not assimilated by the reanalysis, the correlation between the first component of  $\mathbf{I}_t^{\text{WH}}$  and the first component of  $\mathbf{I}_t^{\text{HR}}$  is 0.84; for the second component the correlation is 0.82.

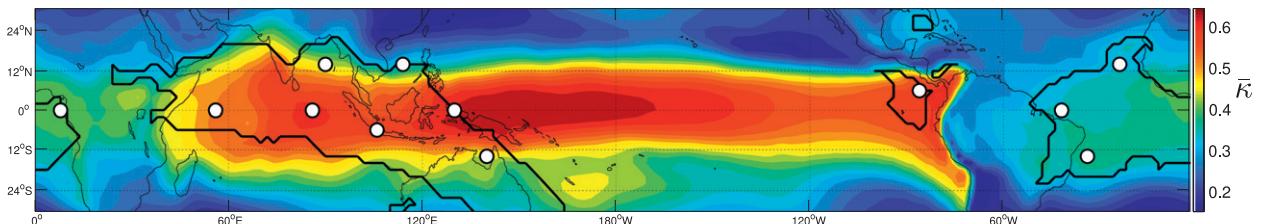


FIG. 1. Relationship between the MJO index and surface pressure.  $\bar{\kappa}_{ij}$  represents the fraction of the total standard deviation of  $\mathbf{I}_t^{\text{WH}}$  that can be statistically accounted for by surface pressure  $p_{ij}^a$  for the period 1979–2008. It is calculated using a frequency-dependent regression [see appendix A and Oliver and Thompson (2010) for details]. White circles show the locations of predictors used in the regression model, and the thick black line defines the exclusion area discussed in section 3.

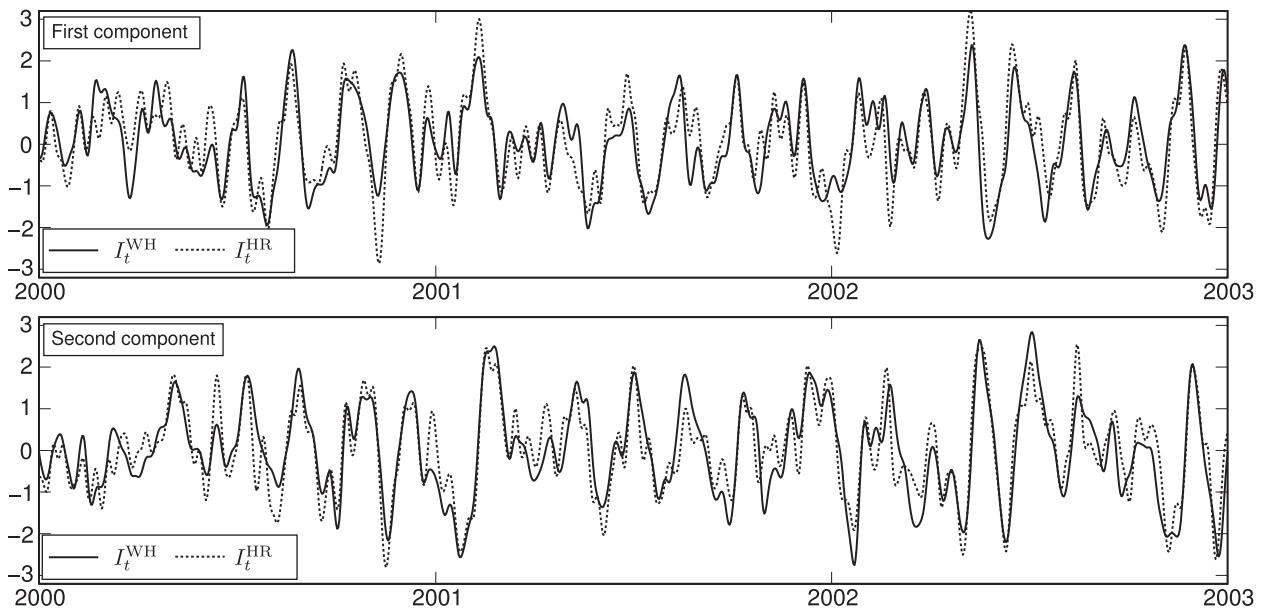


FIG. 2. Comparison of  $I_t^{\text{WH}}$  and  $I_t^{\text{HR}}$  over the period 2000–02. (top) The first component of the indices and (bottom) the second component of each index are shown.

Plots of  $I_t^{\text{HR}}$  for 2000–02 and 1920–22 (Fig. 3) show the qualitative behavior of the reconstructed index is consistent over these two, widely separated periods. These periods were chosen because they both exhibit strong MJO events but similar behavior is found for other periods.

The power spectra of  $I_t^{\text{WH}}$  and  $I_t^{\text{HR}}$  over 1979–2008 and  $I_t^{\text{HR}}$  over the 1905–78 period have the same shape (Fig. 4, analogous to WH04's Figs. 2 and 3). The reconstructed

index is highly coherent and in phase with  $I_t^{\text{WH}}$  over their common period (Fig. 4). In addition, the two components of  $I_t^{\text{HR}}$  are highly coherent over the MJO band (30–90 days) and the phase relationship indicates that the two components are in quadrature as expected.

To test whether the peaked spectrum of  $I_t^{\text{HR}}$  was due to the high- and low-pass filtering of the pressure series we generated a second long index using the same regression

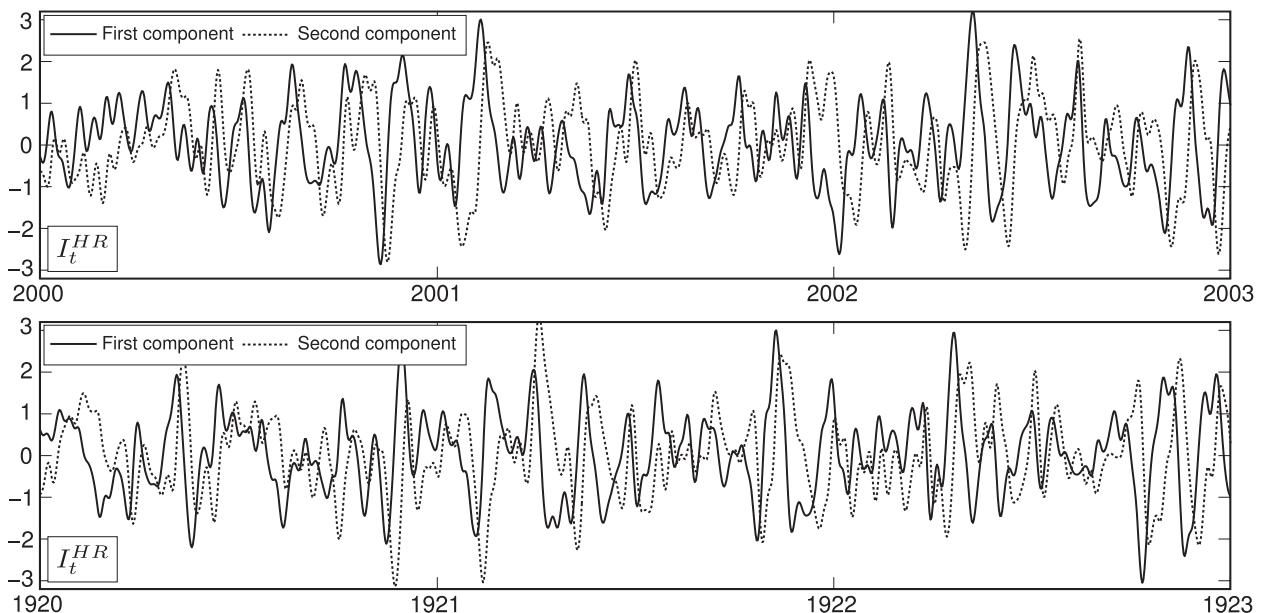


FIG. 3. Comparison of the two components of  $I_t^{\text{HR}}$  for (top) 2000–02 and (bottom) 1920–22.

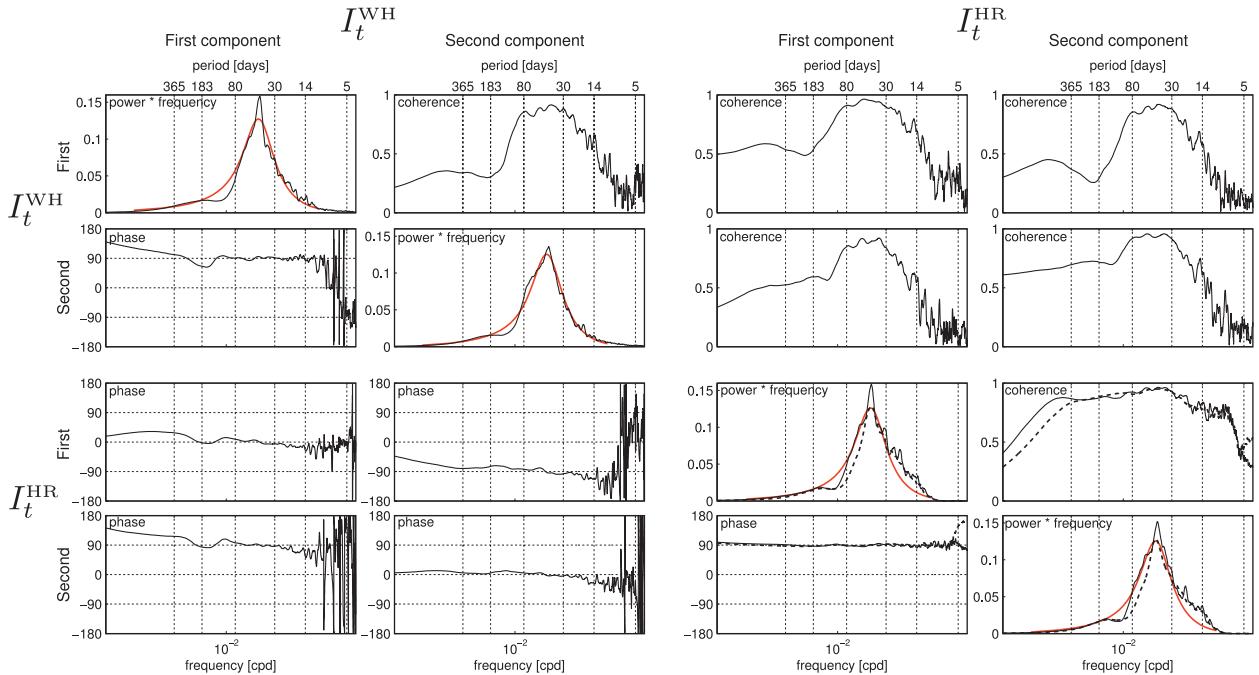


FIG. 4. Cross-spectral analysis of  $\mathbf{I}_t^{\text{WH}}$  and  $\mathbf{I}_t^{\text{HR}}$ . The power spectral densities are shown in the panels on the diagonal and the coherence and phase spectra in the above-diagonal and below-diagonal panels, respectively. Solid lines indicate that the spectra were calculated from data over the 1979–2008 period, and dashed lines indicate that they were performed from data over the 1905–78 period. The power spectral density for the autoregressive model (with  $\tau_1 = 15$  days,  $\tau_2 = 2.5$  days, and  $P = 50$  days) is shown as a red line in the diagonal panels.

coefficients and pressure locations. The pressure predictors used to generate this second index were deseasonalized but otherwise unfiltered. Notably, the power spectra of this index exhibits a spectral peak over the MJO band of frequencies with the same shape as  $\mathbf{I}_t^{\text{HR}}$  (not shown). This indicates that the peaked spectral nature of  $\mathbf{I}_t^{\text{HR}}$  is not due to the bandpass filtering of pressure variability but due to MJO-related variability present in the original pressure series.

The 3-yr running average of MJO amplitude (i.e.,  $\sqrt{I_1^2 + I_2^2}$  where  $I_i$  is the  $i$ th component of the MJO index) of  $\mathbf{I}_t^{\text{HR}}$  and  $\mathbf{I}_t^{\text{WH}}$  (Fig. 5, top panel) agree well over the 1979–2008 period. Using  $\mathbf{I}_t^{\text{HR}}$  it is possible to calculate low-frequency changes in the amplitude of the MJO over most of the last century. To first order, the amplitude of  $\mathbf{I}_t^{\text{HR}}$  is remarkably stable over the 1905–2008 period. Typical changes in MJO amplitude of about 7% about the mean value (maximum changes of about 20%) were found on time scales of decades and longer (thin line, top panel of Fig. 5), along with a weak linear trend (13% increase over the last century).

The 3-yr running power spectra of  $\mathbf{I}_t^{\text{HR}}$  and  $\mathbf{I}_t^{\text{WH}}$  (Fig. 5, middle and bottom panels) again show good agreement over the 1979–2008 period as expected. Much of the variability in the spectral density of  $\mathbf{I}_t^{\text{HR}}$  is associated with changes in the overall power levels. This is confirmed by a principal component analysis of the running

spectra of  $\mathbf{I}_t^{\text{HR}}$  over the 1905–2008 period, which shows that most (59%) of the variability in the spectra can be accounted for by a time-varying modulation of a fixed spectral shape. (The time-varying principal component agrees well with the running average MJO amplitude as expected.) More interestingly, the second principal component accounts for an additional 24% of the variability in spectral shape and effectively shifts power with periods of 30–50 days to periods longer than 100 days. The effect of this second mode of spectral variation is evident in the extended low-frequency tails of spectral energy of  $\mathbf{I}_t^{\text{HR}}$  at certain times (e.g., 1915–21, 1995–2001, middle panel of Fig. 5). We were encouraged to find that a similar analysis of  $\mathbf{I}_t^{\text{WH}}$ , and  $\mathbf{I}_t^{\text{HR}}$  over the pre-1979 period, gave essentially the same results, that is, over 80% of the variability in spectral shape could be accounted for by the two modes described above.

The reconstructed index has been shown to be consistent with  $\mathbf{I}_t^{\text{WH}}$  over the shared period (1979–2008). In the next section we focus on validating the timing of  $\mathbf{I}_t^{\text{HR}}$  over the pre-1979 period.

#### 4. Validity of reconstructed MJO variability

The reconstructed index  $\mathbf{I}_t^{\text{HR}}$  was generated from a regression of reanalysis pressure series trained on  $\mathbf{I}_t^{\text{WH}}$  over the 1979–2008 period. The number of observations

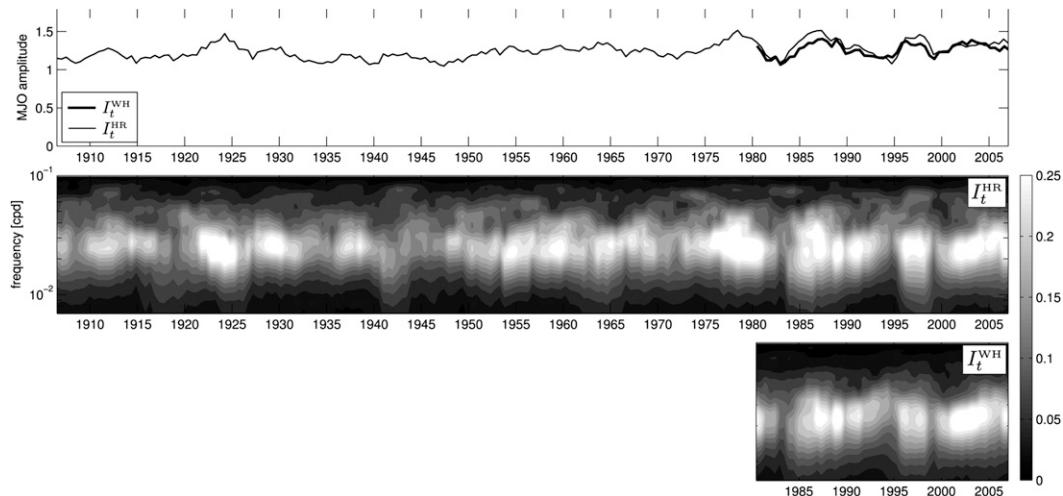


FIG. 5. Amplitude and spectral content of  $I_t^{HR}$  over most of the last century. (top) Running 3-yr-average MJO amplitudes of  $I_t^{HR}$  (thin line) and  $I_t^{WH}$  (thick line) are shown. (middle) Running 3-yr spectral density of  $I_t^{HR}$  and (bottom)  $I_t^{WH}$  are shown. The spectra of the two components of each index have been summed in order to generate these figures.

available for assimilation increases over time and, in addition, satellite SSTs were assimilated by the reanalysis after October 1978. It is important to (i) validate the index over the nontraining period and (ii) explore the sensitivity of the reconstructed index to the changes in the observing system. In this section the validity of  $I_t^{HR}$  over the pre-1979 period assessed using two approaches.

First, the space–time and frequency–wavenumber properties of tropical surface pressure fields reconstructed by  $I_t^{WH}$  and  $I_t^{HR}$  are presented and compared over the 1905–78 and 1979–2008 periods (section 4a). This is done by calculating MJO-reconstructed pressure from  $I_t^{WH}$  and  $I_t^{HR}$  and constructing their Hovmöller diagrams and frequency–wavenumber spectra.

Second, comparisons with independent environmental observations were performed (section 4b). The MJO index can be represented by its position in a two-dimensional space defined by its two components (WH04). MJO events typically follow counterclockwise paths in this space with the radius representing the strength of the MJO, and the angle representing the longitudinal position of enhanced convection along the equator. The angular position is usually divided into eight discrete  $45^\circ$  phases. The center of convection is over the Western Hemisphere and Africa during phases 8 and 1, the Indian Ocean during phases 2 and 3, the Maritime Continent during phases 4 and 5, and the western Pacific during phases 6 and 7. Weak MJOs are defined as having an MJO amplitude less than one. The second approach is based on composites of cloud cover, surface wind, probability of extreme precipitation over Australia, and sea level according to MJO phase during events with amplitude greater than one. The

results are presented over the pre-1979 period (as far back as data availability permits) and the 1979–2008 period.

#### a. Reconstructed surface pressure fields and frequency–wavenumber spectra

To demonstrate that the reconstructed index captures the eastward-propagating, intraseasonal surface pressure signal, “MJO-reconstructed pressure” was generated by regressing surface pressure  $p_{ijt}$  onto  $I_t^{WH}$  or  $I_t^{HR}$ . (The annual cycle of pressure and its first two harmonics were removed by linear regression prior to the analysis.) A Hovmöller diagram for  $I_t^{WH}$ -reconstructed pressure shows that the eastward-propagating signature of the MJO is strongly present in surface pressure (Fig. 6, left panel; analogous to WH04’s Fig. 10 for OLR). Additionally, Hovmöller diagrams for  $I_t^{HR}$ -reconstructed pressure indicate that the reconstructed index captures this propagating pressure signal for both the modern and historical periods (Fig. 6, middle and right panels).

Frequency–zonal wavenumber ( $\omega$ – $k$ ) spectra of  $p_{ijt}$  (after removing the annual cycle and its first two harmonics) have been calculated over 96-day segments for nine equi-spaced latitudes between  $16^\circ\text{S}$  and  $16^\circ\text{N}$ . These spectra were then averaged over all segments and summed over  $16^\circ\text{S}$ – $16^\circ\text{N}$ , following Wheeler and Kiladis (1999), to generate an  $\omega$ – $k$  diagram for tropical surface pressure (Fig. 7, top-left panel; analogous to WH04’s lower panel of Fig. 11 for OLR). This diagram indicates variability on time- and space scales predicted by equatorial wave theory is represented in surface pressure. The dispersion curves for equatorially trapped Kelvin waves (equivalent depths of 25 and 70 m) and

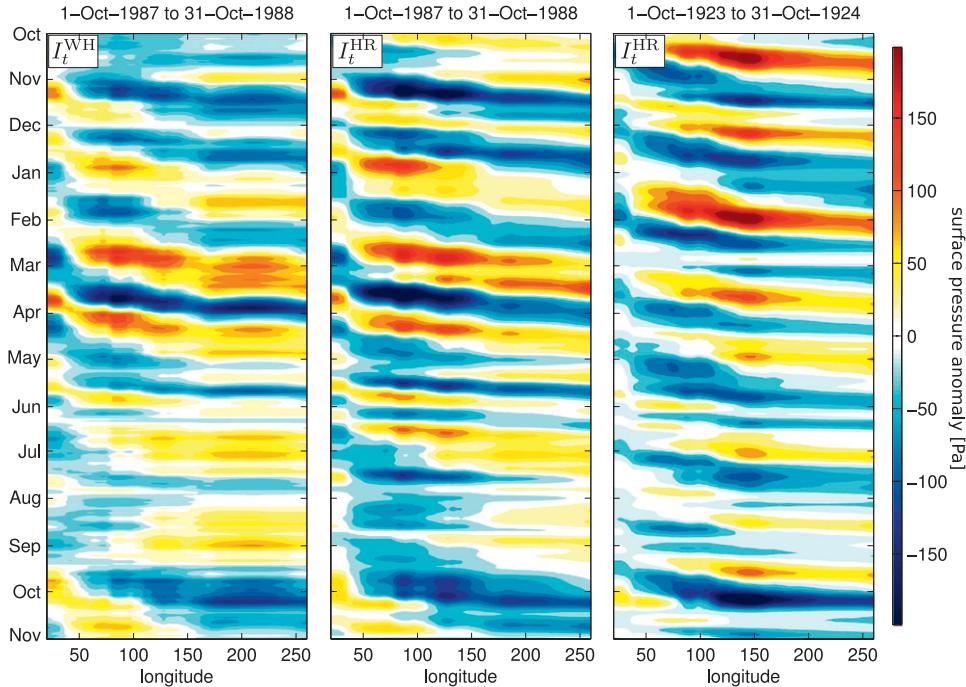


FIG. 6. Hovmöller diagram of MJO-reconstructed tropical ( $16^{\circ}\text{S}$ – $16^{\circ}\text{N}$ ) surface pressure for (left) 1 Oct 1987–31 Oct 1988 using  $\mathbf{I}_t^{\text{WH}}$ , (middle) 1 Oct 1987–31 Oct 1988 using  $\mathbf{I}_t^{\text{HR}}$ , and (right) 1 Oct 1923–31 Oct 1924 using  $\mathbf{I}_t^{\text{HR}}$ . Note that time is increasing downward.

equatorially trapped Rossby waves (equivalent depths of 12 m, 30 m, and 10 km) are shown as thin lines. The shallow equivalent depths are consistent with moist convectively coupled waves (Wheeler et al. 2000) and the 10-km equivalent depth represents the gravest Rossby wave mode permitted in the atmosphere (Diky and Golitsyn 1968; Kasahara 1976). Variability with zero wavenumber is also present indicating variability in the zonal mean pressure on a variety of time scales.

The  $\omega$ – $k$  diagram of  $\mathbf{I}_t^{\text{WH}}$ -reconstructed pressure (Fig. 7, top-right panel; analogous to WH04’s top panel of Fig. 11 for OLR) indicates the dominance of intraseasonal, zonal wavenumber-1, eastward-propagating pressure variability, which is consistent with the observed behavior of the MJO. The frequency–wavenumber diagrams of  $\mathbf{I}_t^{\text{HR}}$ -reconstructed pressure over the 1905–78 and 1979–2008 periods indicate that the reconstructed index captures this propagating pressure signal equally as well as  $\mathbf{I}_t^{\text{WH}}$  over both periods (Fig. 6, bottom-left and bottom-right panels respectively).

### b. Composites of independent environmental variables

The spatial patterns of atmospheric and oceanic variables and their relationships with the MJO are now examined using composites.

### 1) CLOUD COVER AND SURFACE WIND OVER THE OCEAN

We first test the validity of the reconstructed index by examining the stability of its relationship with cloud cover and surface wind over the ocean. WH04 compared gridded OLR and 850-hPa wind by conditionally averaging based on the phase of  $\mathbf{I}_t^{\text{WH}}$ . Here, we perform a similar analysis using ship measurements of cloud cover and surface wind over the ocean (see section 2c). These variables are used because (i) cloud cover is closely related to OLR (increased cloud cover will reduce OLR) and surface wind will behave similarly to 850-hPa wind, (ii) they are available before and after 1979 (1952–2008), and (iii) they were not used to generate the reanalysis pressures on which  $\mathbf{I}_t^{\text{HR}}$  is based.

Cloud cover and surface wind observations from  $4^{\circ} \times 4^{\circ}$  grid cells were composited with the MJO index by taking averages of observations falling on days within each of the eight MJO phases defined above and with amplitude greater than one. The composites were performed over different periods thereby allowing a comparison of the two indices over their common period and also an examination of the long-term stability of the composites based on the reconstructed index. The composites are not stratified by season because of the limited data availability.

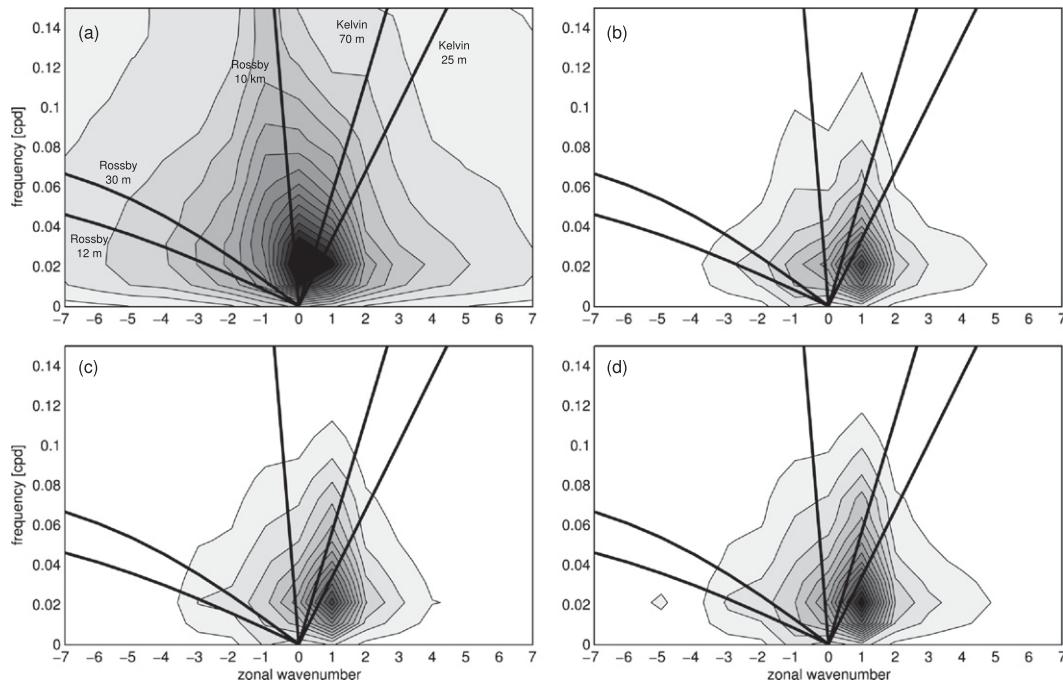


FIG. 7. Frequency–wavenumber spectra of tropical ( $16^{\circ}\text{S}$ – $16^{\circ}\text{N}$ ) surface pressure. Spectra are shown for (a) the raw pressure field  $p_{ijt}$  (with the annual cycle and first three harmonics removed) over 1905–2008, (b) the  $\mathbf{I}_t^{\text{WH}}$ -reconstructed pressure over 1979–2008, (c) the  $\mathbf{I}_t^{\text{HR}}$ -reconstructed pressure over 1979–2008, and (d) the  $\mathbf{I}_t^{\text{HR}}$ -reconstructed pressure over 1905–1978. This diagram was generated using the same calculation for OLR performed by Wheeler and Kiladis (1999). Lines represent the dispersion curves for equatorially trapped Kelvin waves (equivalent depth of 25 and 70 m) and equatorially trapped Rossby waves (equivalent depths of 12 m, 30 m, and 10 km).

Composites of cloud cover and surface wind with  $\mathbf{I}_t^{\text{WH}}$  (Fig. 8) are consistent with composites of OLR and 850-hPa wind with  $\mathbf{I}_t^{\text{WH}}$ . Compare with the mean of Figs. 8 and 9 of WH04 for boreal winter and boreal summer, respectively. During phases 1–3 cloud cover increases by up to 15% over the Indian Ocean. During phases 4–6 this disturbance propagates eastward over the Maritime Continent into the western Pacific and by phases 7 and 8 it moves over the central Pacific. A strong increase in cloud cover is also evident over the Pacific coast of Mexico (this is a boreal summer feature of the MJO, WH04). Westerly surface wind anomalies of up to  $2 \text{ m s}^{-1}$  propagate eastward along with the cloud cover anomalies; easterly wind anomalies follow a half-cycle behind along with a negative cloud cover disturbance which repeats the pattern described above.

Composites of cloud cover and surface wind with  $\mathbf{I}_t^{\text{HR}}$  over the 1979–2008 period (Fig. 9) are consistent with the results using  $\mathbf{I}_t^{\text{WH}}$  as expected. More importantly, composites with  $\mathbf{I}_t^{\text{HR}}$  over the 1952–78 period (Fig. 10) are also consistent with composites of  $\mathbf{I}_t^{\text{WH}}$ , indicating that the relationship between the MJO and both cloud cover and surface wind is stable over the second half of the last century.

## 2) PROBABILITIES OF EXTREME RAINFALL OVER AUSTRALIA

We now test the validity of the reconstructed index by examining the stability of its relationship with extreme rainfall over Australia. Using gridded weekly rainfall amounts, WH04 calculated the probability of DJF rainfall exceeding the highest quintile conditioned by MJO phase. They found a higher probability of extreme rainfall in a band across northern Australia during MJO phases 4–6 and speculated that this is due to a modulation of the monsoon by the MJO. Here, we perform a similar analysis using December–February (DJF) station rainfall data during the 1905–49, 1950–78, and 1979–2001 periods. (Given the long precipitation records, it is possible to split the pre-1979 period into two sections for validation as has been done here.)

For each station (see Table 1) and time period, the highest quintile of DJF rainfall was calculated. Twenty percent of all of the observed rainfall values are larger than this threshold; rainfall higher than this value was defined as “extreme rainfall.” Then, samples of observed rainfall stratified according to MJO phase were formed (ignoring MJO events with amplitude less than

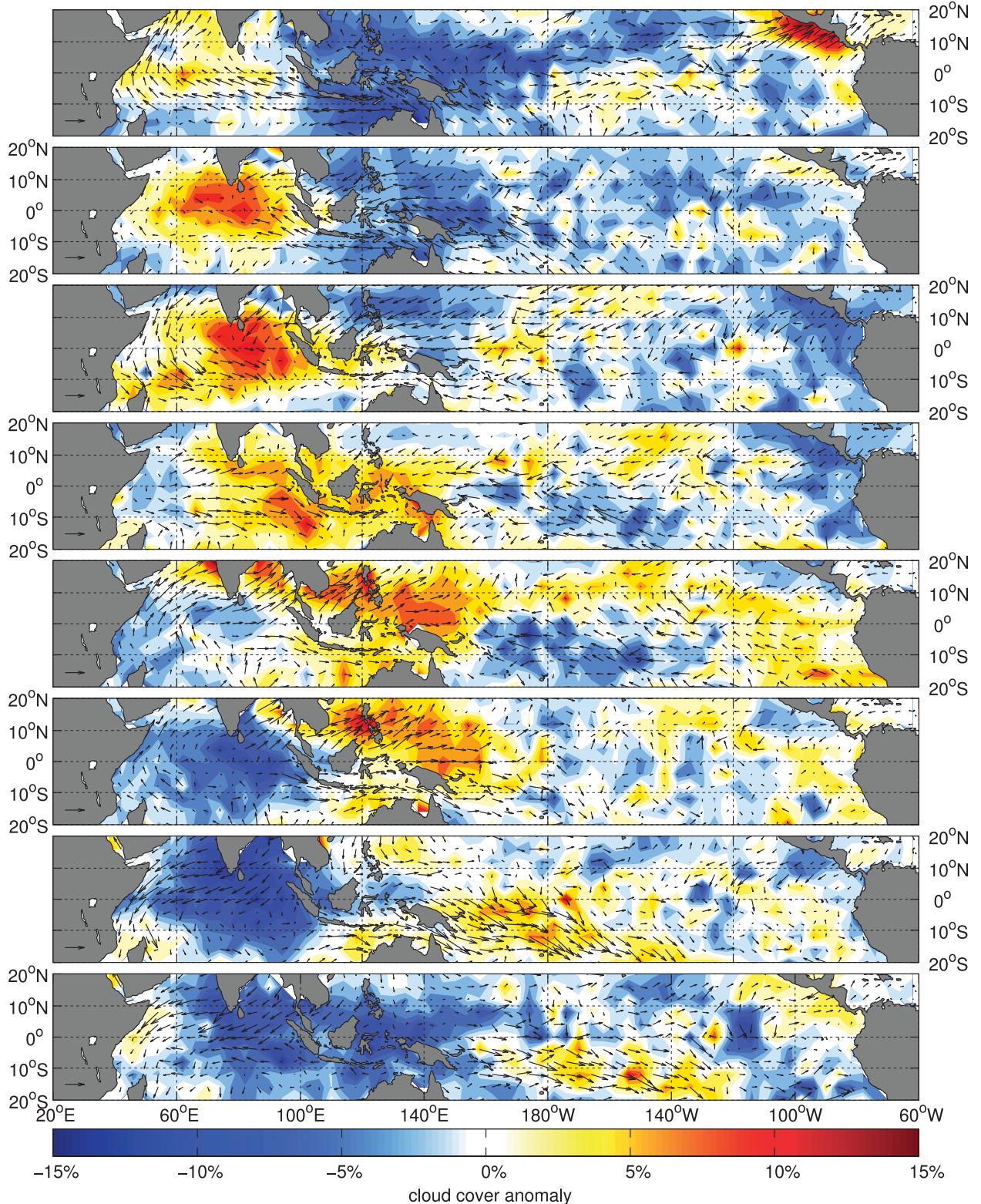


FIG. 8. Relationship between the MJO and cloud cover and surface velocity over the ocean. Color shading represents composites of cloud cover with phase of  $I_r^{WH}$ ; the arrows represent composites of surface wind. Cloud cover is given as a percent change from the mean state; the arrow in the lower left corner of each panel is a reference wind speed of  $2 \text{ m s}^{-1}$ . Results for grid cells with fewer than 20 observations are not shown. The analysis was performed over the 1979–2008 period.

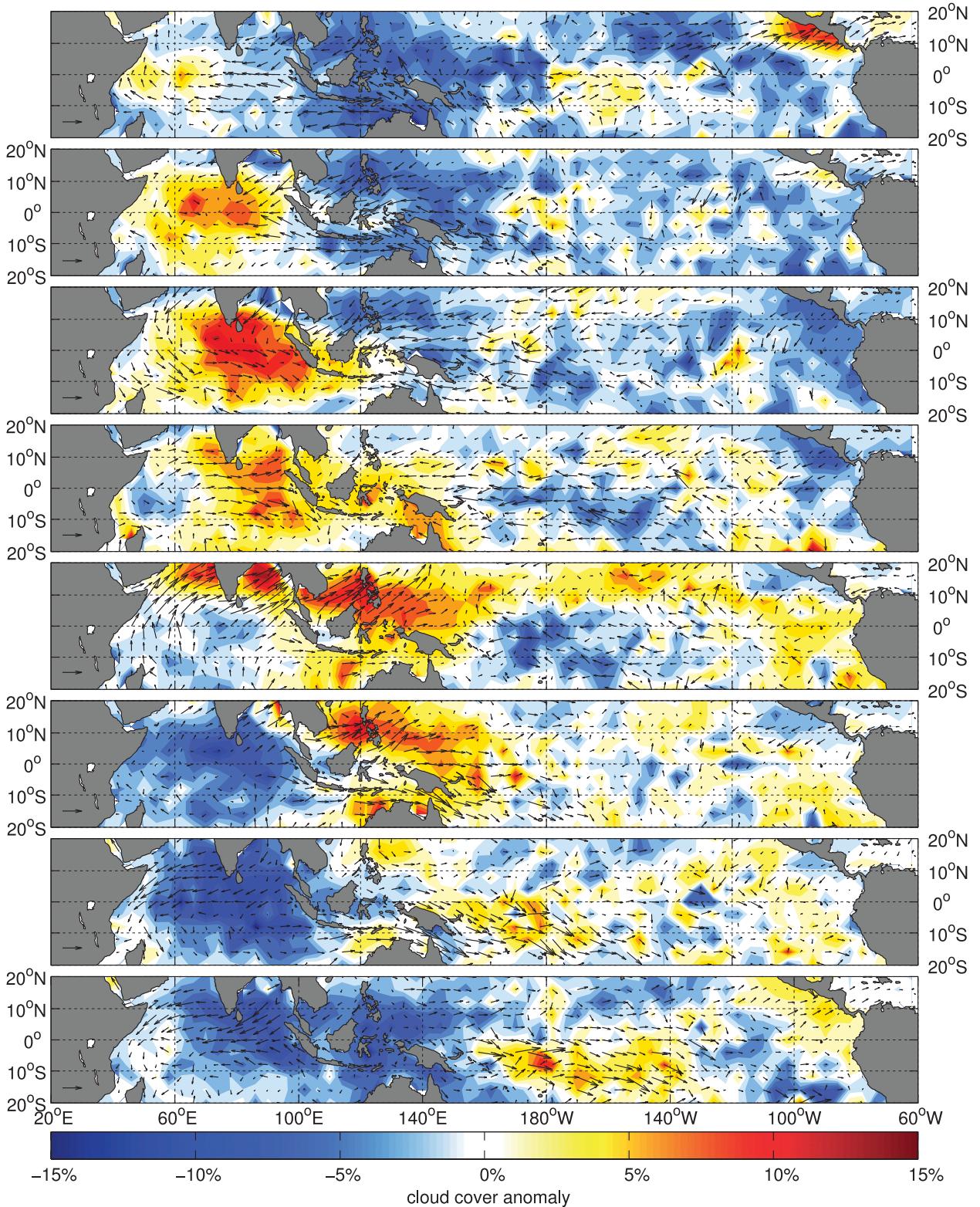


FIG. 9. As in Fig. 8, but for  $I_t^{HR}$  over the 1979–2008 period.

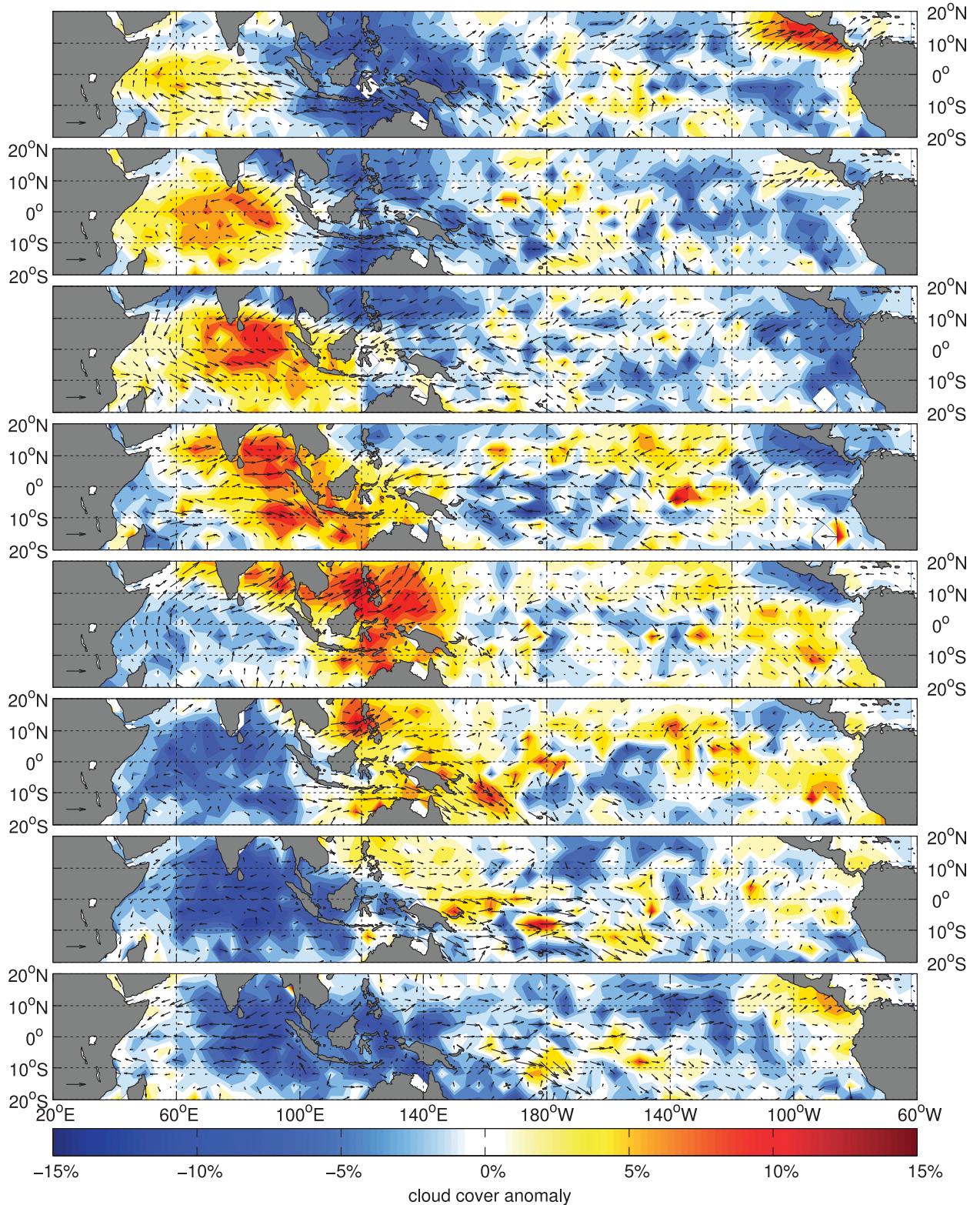


FIG. 10. As in Fig. 8, but for  $I_t^{\text{HR}}$  over the 1952–1978 period.

one) and the proportion of extreme rainfall within each sample was calculated. This proportion represents the probability of extreme rainfall for each MJO phase at each station. The analysis was performed over the three periods specified above thereby allowing a comparison of the two indices over their common period and also an examination of the long-term stability of the relationship.

The probabilities of extreme DJF rainfall over Australia conditioned on  $\mathbf{I}_t^{\text{WH}}$  phase (Fig. 11, first set of bars; analogous to Fig. 16 of WH04) are consistent with WH04. Throughout the MJO cycle these probabilities vary about 0.2; if there was no relationship with the MJO they would be exactly 0.2 everywhere apart from sampling variability. The probabilities begin to increase across northwestern Australia and the Cape York Peninsula in phase 4, peak with probabilities of 0.3–0.55 during phases 5 and 6, and begin to diminish by phase 7. A general northward propagation of high probabilities can also be seen for phases 3–6 especially near the eastern and northwestern coasts.

The probabilities of extreme DJF rainfall over Australia conditioned on  $\mathbf{I}_t^{\text{HR}}$  phase over the 1979–2001 period (Fig. 11, second set of bars) are consistent with the results for  $\mathbf{I}_t^{\text{WH}}$ . More importantly, the results for  $\mathbf{I}_t^{\text{HR}}$  over the 1950–78 period (third set of bars), and the 1905–49 period (fourth set of bars) are also generally consistent with  $\mathbf{I}_t^{\text{WH}}$  indicating that the relationship between extreme rainfall over Australia and the MJO is stable and captured by  $\mathbf{I}_t^{\text{HR}}$ .

### 3) SEA LEVEL

The MJO has been shown to account for a significant proportion of intraseasonal sea level variability in the western Pacific, the Gulf of Carpentaria, and along the Pacific coast of the Americas (Oliver and Thompson 2010, 2011, and references therein). Daily sea level from several tide gauges in these regions (see Section 2e) was composited with MJO phase. Sea level along the coast of the Americas is remotely forced by the MJO and so these time series were lagged with respect to the MJO index before compositing (see Table 2). The lags are given by the lag at maximum correlation with sea level with  $\mathbf{I}_t^{\text{WH}}$ . Note that they increase poleward from the equator consistent with the propagation of coastal trapped waves. The response of sea level from the western Pacific and the Gulf of Carpentaria is due to local forcing by the MJO and so these series were not lagged.

Composites of sea level with  $\mathbf{I}_t^{\text{WH}}$  according to MJO phase (Figs. 12 and 13, thick black lines) indicate that the MJO is related to sea level variations of 1–4 cm in the western Pacific, ~10 cm in the Gulf of Carpentaria, and 1–3 cm along the Pacific coast of the Americas. Composites of sea level according to  $\mathbf{I}_t^{\text{HR}}$  phase over the

1979–2008 period (Figs. 12 and 13, thin black lines) and the earlier periods (blue and red lines) are generally consistent with the composites using  $\mathbf{I}_t^{\text{WH}}$ .

## 5. Predictability of the MJO Index

The predictability of the MJO is now examined to further assess the realism of the reconstructed index. In the first subsection we introduce three straightforward empirical measures of predictability and estimate them for both  $\mathbf{I}_t^{\text{WH}}$  and  $\mathbf{I}_t^{\text{HR}}$ . In the second subsection a simple, physically based model is used to interpret the empirical measures and also gain new insights into the predictability of the MJO.

### a. Empirical measures of predictability

Consider a scatterplot of a time-varying MJO index with the first component on the  $x_1$  axis and the second component on the  $x_2$  axis. This set of realized states can be considered as an approximation of the probability density function of the true state of the MJO at some arbitrary time  $t_0$  given no other information (e.g., Ristic et al. 2004). We will refer to this distribution as the prior distribution (with zero mean and variance of  $\sigma_{\text{prior}}^2$ ). Suppose that a bivariate observation of the standardized MJO  $\mathbf{y}_{\text{obs}}$  becomes available at time  $t_0$ . We assume  $\mathbf{y}_{\text{obs}}$  is subject to a normally distributed, zero mean observation error with covariance  $\sigma_{\text{obs}}^2 \mathbf{I}$ . To update the prior distribution to take into account this new information we take a Bayesian approach. The posterior distribution is approximated by sampling, with replacement, from the set of observed points in the above scatterplot using weights that are proportional to the likelihood function, that is,  $\exp(-|\mathbf{y}_{\text{obs}} - \mathbf{x}|^2/2\sigma_{\text{obs}}^2)$ , where  $\mathbf{x}$  denotes the true state of the MJO at time  $t_0$ . As  $\sigma_{\text{obs}}^2 \rightarrow 0$  the posterior distribution will reduce in spread and become centered more closely on  $\mathbf{y}_{\text{obs}}$ ; as  $\sigma_{\text{obs}}^2 \rightarrow \infty$  the posterior will revert back to the prior distribution.

It is straightforward to approximate the posterior distribution of the true state of the MJO at a future time  $t_0 + k$  by simply following the observed evolution of each point in the posterior sample according to the original MJO time series starting at time  $t_0$ . This procedure is illustrated in Fig. 14 for  $k$  increasing from 0 to 35 days. (The two components of the MJO index have been scaled by their respective sample standard deviations over each period shown.) The thin lines in each panel are a representative selection of observed trajectories of the MJO with the initial positions at time  $t_0$  reflecting the posterior distribution. In all of these panels the observation is  $\mathbf{y}_{\text{obs}} = (\sqrt{2}, -\sqrt{2})$  and the observation error variance is taken to be  $\sigma_{\text{obs}}^2 = 0.07$ . It is clear from Fig. 14 that the ensemble undergoes a counterclockwise

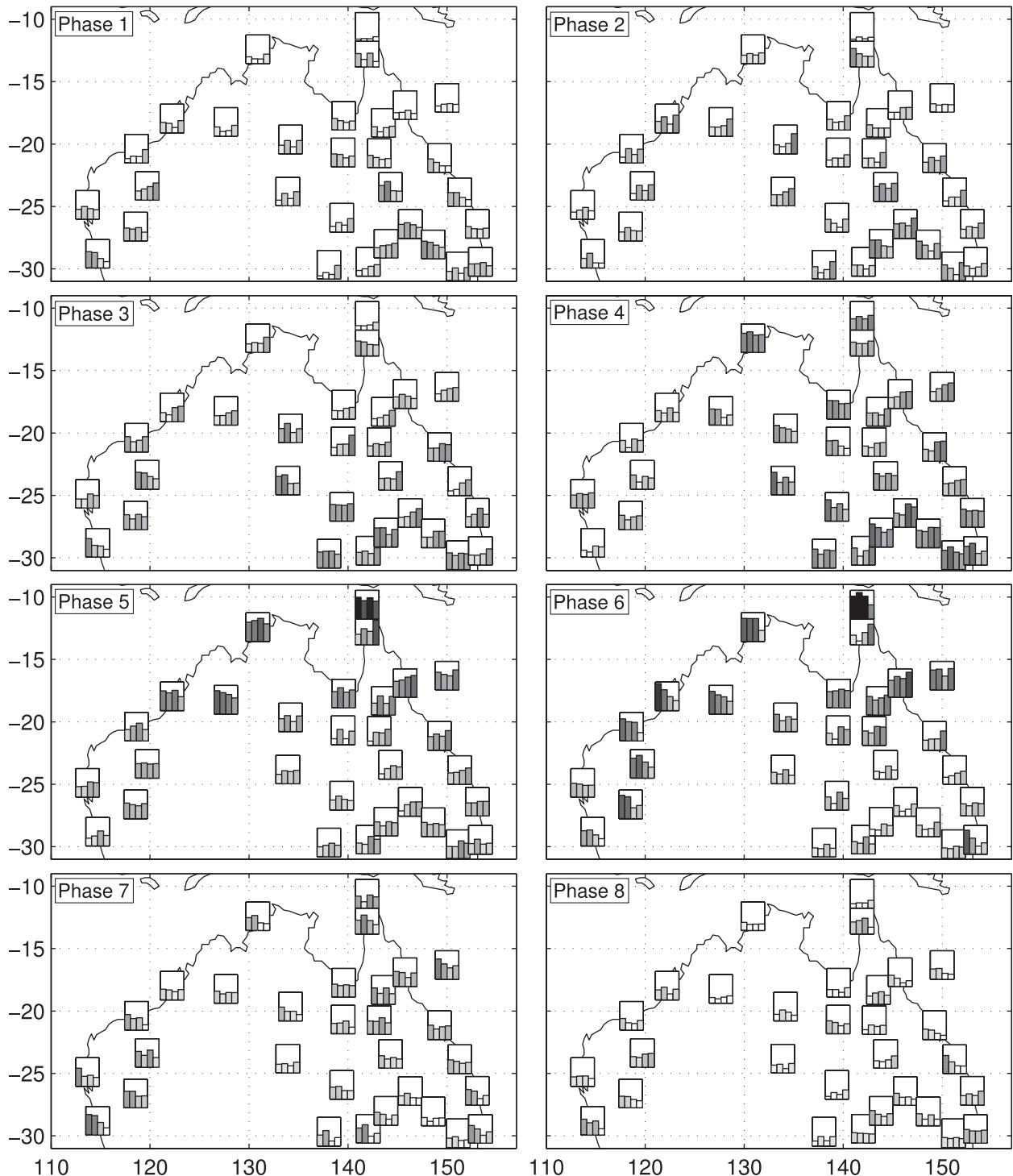


FIG. 11. Relationship between the MJO and extreme rainfall over Australia. The location of each bar plot represents the station location where rainfall data were available (see Table 1). The vertical axis of each bar plot represents the probability of weekly DJF rainfall in the highest quintile for that station, conditioned on MJO phase (the range of each bar plot is from 0 to 0.45 except 0.6 for the northernmost station, Booby Island). The shading level of each bar is proportional to its height. The first bar in each bar plot is the result using  $I_t^{\text{WH}}$  over 1979–2008, the second bar is  $I_t^{\text{HR}}$  over 1979–2008, the third bar is  $I_t^{\text{HR}}$  over 1950–73, and the fourth bar is  $I_t^{\text{HR}}$  over 1905–49.

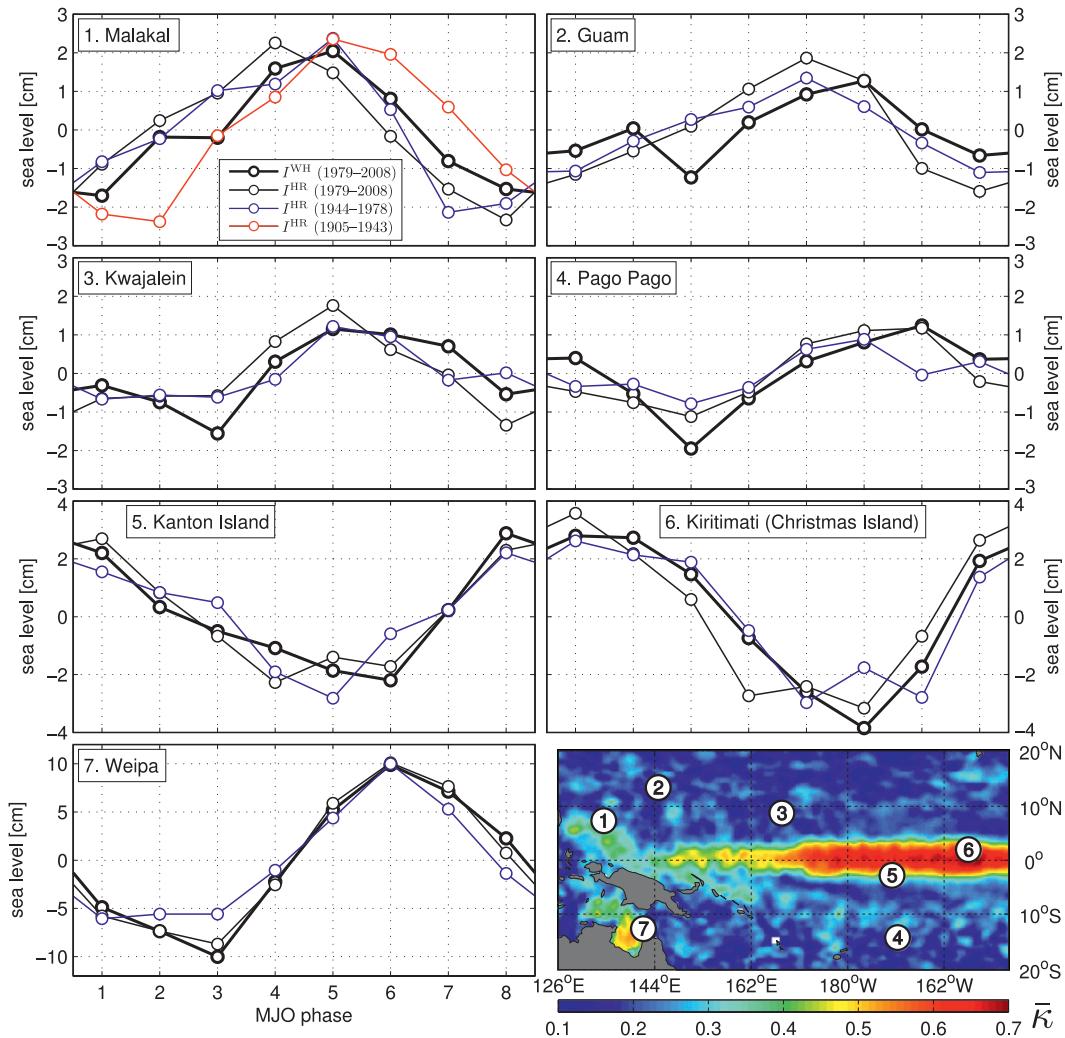


FIG. 12. Relationship between the MJO and sea level in the western Pacific and the Gulf of Carpentaria. Composites of sea level from tide gauges (see Table 2) are shown for  $I_t^{\text{WH}}$  and  $I_t^{\text{HR}}$  over the modern and historical periods. (bottom right) The locations of these tide gauges are shown along with a map of  $\bar{\kappa}$  [the proportion of satellite-derived sea level which is accounted for by the MJO, regional subset of Fig. 1 from Oliver and Thompson (2010)].

rotation, the amplitude of the ensemble mean (thick line) decreases, and the ensemble spread (shaded area) increases through time. It can also be seen that the position of the ensemble members relative to their sample mean at time  $t_0$  is lost with increasing time. For example, if an ensemble member is to the left of the ensemble mean at  $t_0$ , then the probability of the member remaining to the left of the mean decreases as  $k$  increases.

Based on the above discussion we propose three empirical measures of predictability: the rate of decay of the posterior mean toward zero, the increase of the ensemble variance, and the decrease of the correlation between the MJO at state time  $t_0$  and  $t_0 + k$ . Let  $\mu_{t_0+k}$  denote the mean of the posterior distribution at time  $t_0 + k$ . The first measure of predictability is taken to be

the time required for  $|\mu_{t_0+k}|/|\mu_{t_0}|$  to drop below a specified critical value (henceforth  $\tau_\mu$ ). Let  $\sigma_{t_0+k}^2$  denote the total variance of the posterior distribution at time  $t_0 + k$ . The second measure of predictability is the time required for  $\sigma_{t_0+k}^2/\sigma_{\text{prior}}^2$  to exceed a specified critical value (henceforth  $\tau_{\sigma^2}$ ). Let  $\rho_{t_0,t_0+k}$  denote the proportion of standard deviation of the MJO at time  $t_0 + k$  that can be accounted for by the MJO at time  $t_0$  according to the evolved posterior distribution. The final measure of predictability is the time for  $\rho_{t_0,t_0+k}$  to fall below a specified critical value (henceforth  $\tau_\rho$ ).

To compare  $I_t^{\text{WH}}$  and  $I_t^{\text{HR}}$  over different periods we have calculated their corresponding predictability measures. There have been suggestions that the time scale for the loss of predictability of a given MJO event depends

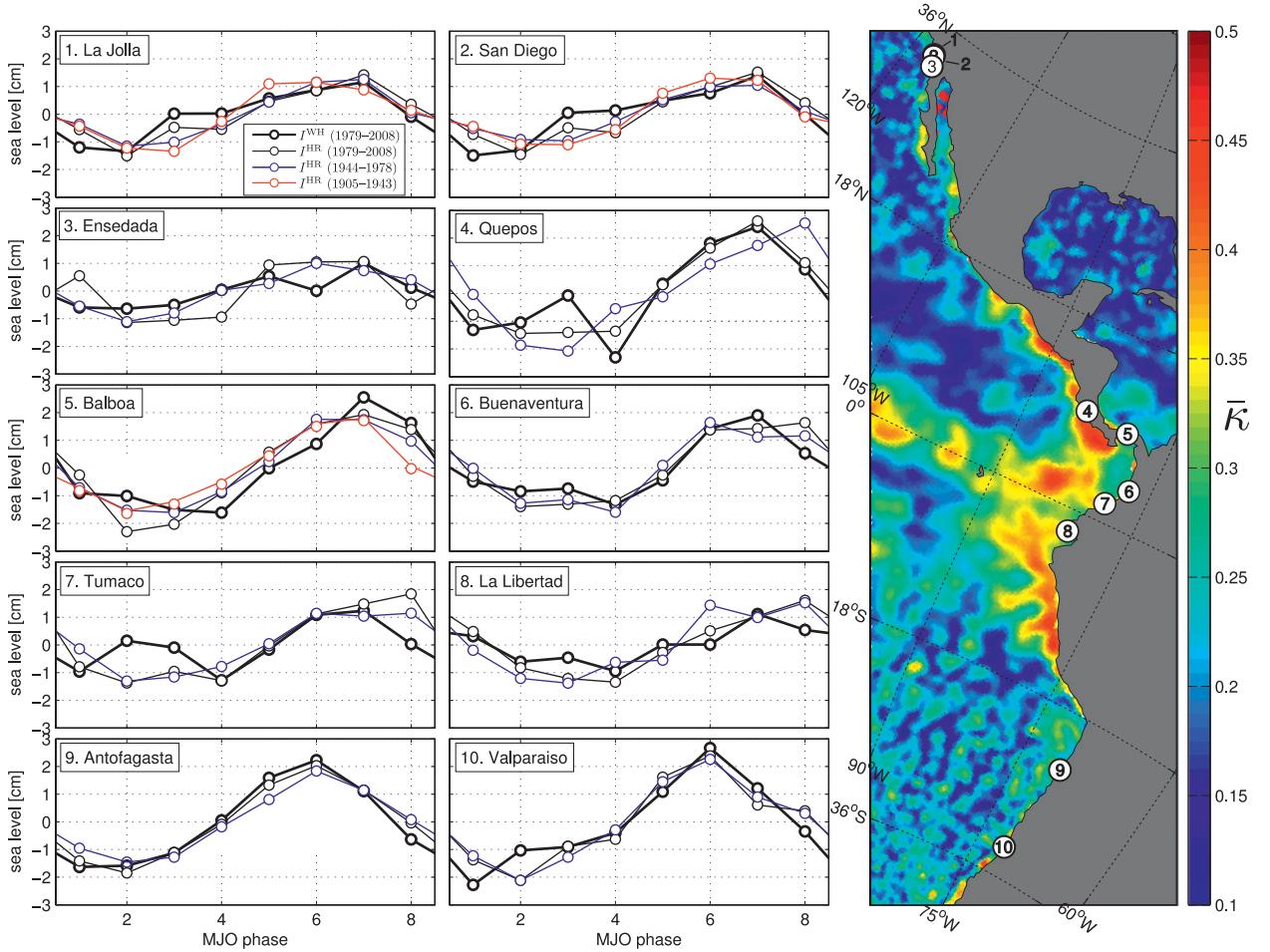


FIG. 13. As in Fig. 12, but for sea level from tide gauges along the Pacific coast of the Americas. Sea level time series have been lagged with respect to the MJO index (see Table 2 for details).

on the phase at which the event is initially defined (Lin et al. 2008; Kim et al. 2009; Kang and Kim 2010). We have therefore performed an additional set of experiments with  $|\mathbf{y}_{\text{obs}}|$  equal to 1.5 and the phase angle of  $\mathbf{y}_{\text{obs}}$  taking on eight values centered on each of the MJO phases. For these additional runs we kept the observation error variance fixed at 0.25. All of the results are discussed below.

The variation of the three predictability measures ( $|\boldsymbol{\mu}_{t_0+k}|/|\boldsymbol{\mu}_{t_0}|$ ,  $\sigma_{t_0+k}^2/\sigma_{\text{prior}}^2$  and  $\rho_{t_0,t_0+k}$ ) is plotted as a function of  $k$  for  $\mathbf{I}_t^{\text{WH}}$  in Fig. 15. For all three measures there is a short period (2–3 days) over which the initial values persist followed by a slower relaxation to their asymptotic values (the mean and correlation decay to zero and the variance increases toward 2 which is the total variance of the normalized MJO index). The critical values are also shown in this figure; once they are defined it is straightforward to read off  $\tau_{\mu}$ ,  $\tau_{\sigma^2}$ , and  $\tau_{\rho}$ .

The time scales for loss of predictability are shown in Fig. 16 as a function of MJO phase. The most important feature of Fig. 16 is that the predictability scales are

consistent among  $\mathbf{I}_t^{\text{WH}}$  and  $\mathbf{I}_t^{\text{HR}}$  over the WH04 period and  $\mathbf{I}_t^{\text{HR}}$  over the pre-WH04 period. Overall, the predictability scales for  $\mathbf{I}_t^{\text{WH}}$  and  $\mathbf{I}_t^{\text{HR}}$  are independent of phase (Fig. 16) apart from a slight dependence of  $\tau_{\mu}$  for  $\mathbf{I}_t^{\text{WH}}$  that is not reproduced by  $\mathbf{I}_t^{\text{HR}}$ .

The time evolution of the angular position of the ensemble means is approximately linear especially in the early part of each event (bottom right of Fig. 15). The slope of the phase line versus time leads to an estimate of the period of the MJO. We found that the period is stable with respect to the MJO phase of the observation and is between 48 and 54 days for  $\mathbf{I}_t^{\text{WH}}$  and  $\mathbf{I}_t^{\text{HR}}$  over their shared period. Over the pre-WH04 period the dominant period of  $\mathbf{I}_t^{\text{HR}}$  was estimated to be about 4 days shorter. Note that  $\tau_{\mu}$ ,  $\tau_{\sigma^2}$ , and  $\tau_{\rho}$  are smaller than the MJO period indicating that an average lifetime of an event is short compared to the dominant oscillation time scale.

Note that the three predictability measures are all different: the time scale depends on the chosen statistic. We found typical time scales of 18 days for the mean and

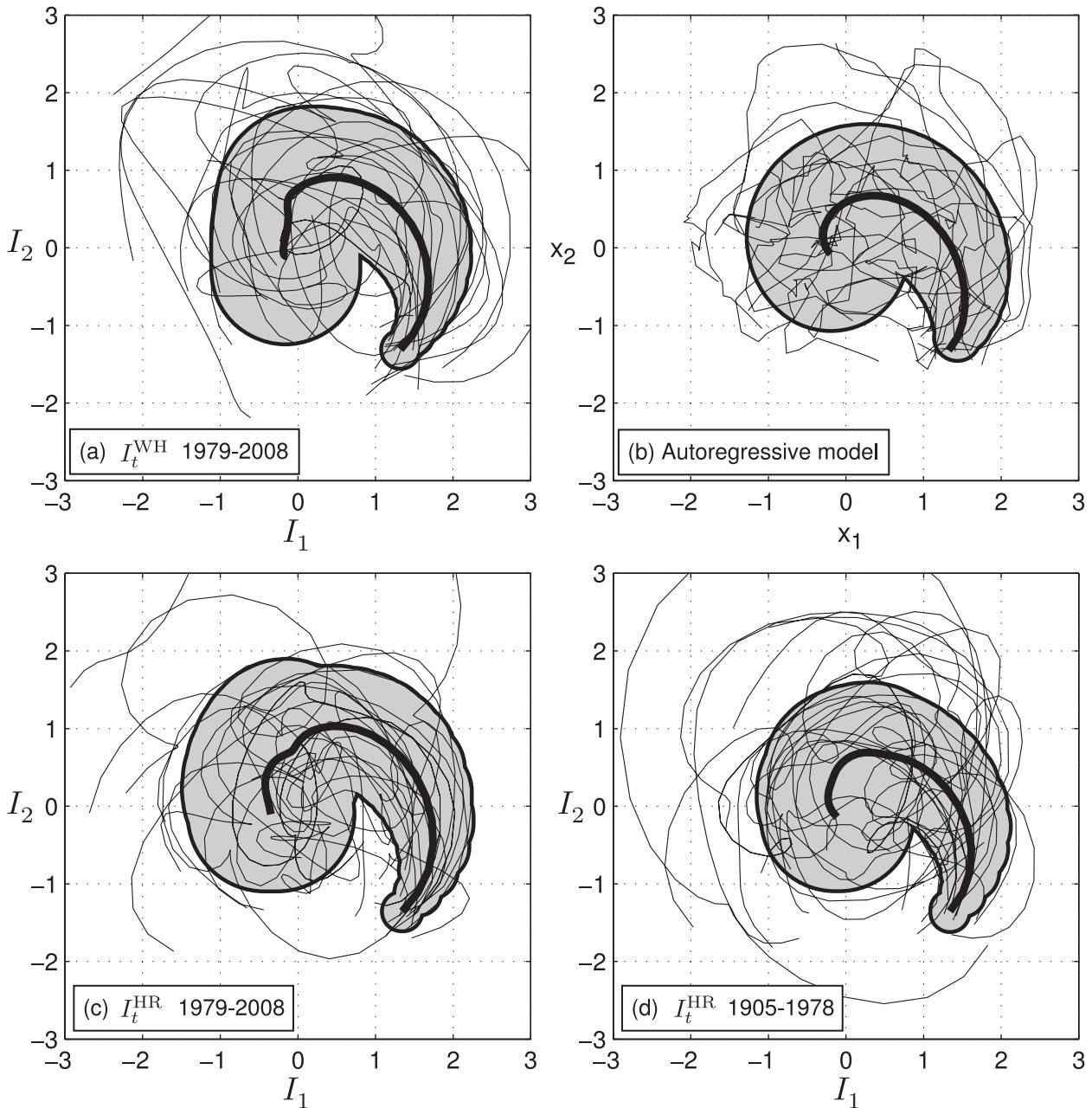


FIG. 14. Evolution of the MJO in phase space given an imperfectly observed initial condition. The statistical properties of a 4000-member ensemble with  $\mathbf{y}_{\text{obs}} = (\sqrt{2}, -\sqrt{2})$  and  $\sigma_{\text{obs}}^2 = 0.07$  are shown over the following 35 days for (a)  $\mathbf{I}_t^{\text{WH}}$ , (c)  $\mathbf{I}_t^{\text{HR}}$  (1979–2008), and (d)  $\mathbf{I}_t^{\text{HR}}$  (1905–78). (b) Representative trajectories calculated using the autoregressive model are shown for 35 time steps. (Model parameters are  $\tau_1 = 15$  days,  $\tau_2 = 2.5$  days, and  $P = 50$  days.) For clarity, only 20 representative trajectories are shown in each panel. The thick black line shows the mean of each 4000-member ensemble, and shaded region shows the standard deviation of each about this mean. Note that each index component has been scaled by its standard deviation over the period shown.

10 days for the variance. These two results are not sensitive to changes in  $|\mathbf{y}_{\text{obs}}|$  and  $\sigma_{\text{obs}}^2$  (tested by doubling and halving the values given above). We found a typical time scale of 5 days for the correlation. The dependence of  $\tau_p$  has a more complex relationship with the observation error and is discussed in section 6.

The predictability time scale based on the mean ( $\tau_\mu$ ) is consistent with previous estimates. Waliser (2005) suggested that this time scale is between 20 and 30 days. Other studies have estimated the time scale to be 2 weeks using a coupled ocean–atmosphere circulation model (Vitart et al. 2007), 15 days using autoregressive models

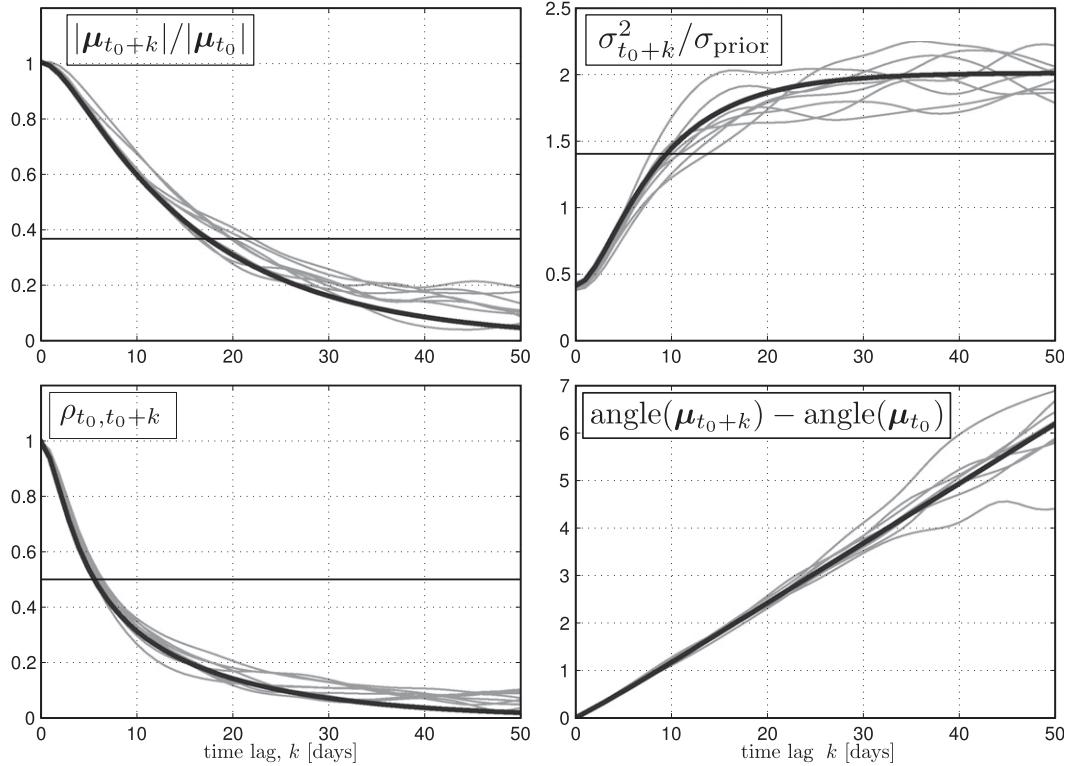


FIG. 15. Predictability measures for the Wheeler and Hendon MJO index and damped harmonic oscillator model. (top left) Decay of the mean, (top right) increase of variance, (bottom left) decrease of correlation, and (bottom right) increase in angular position are shown. Results are shown for ensembles of  $\mathbf{I}_t^{\text{WH}}$  associated with each of the eight MJO phases (thin lines), that is, eight ensembles are initialized for  $\mathbf{y}_{\text{obs}}$  with magnitude of 1.5 and centered in each of the eight MJO phases. Corresponding measures calculated using the damped harmonic oscillator model ( $\tau_1 = 15$  days,  $\tau_2 = 2.5$  days, and  $P = 50$  days) are shown by the thick lines. Critical values used to calculate predictability times scales are  $e^{-1}$  for the mean and the variance and 0.5 for the correlation (shown as horizontal lines).

(Lo and Hendon 2000; Maharaj and Wheeler 2005), and 15 days using a multivariate lag–regression model Jiang et al. (2008).

#### b. Interpretation of predictability using a damped harmonic oscillator

The evolution of the posterior distribution described above (i.e., rotation and decay in MJO space) is reminiscent of the behavior of a forced, damped harmonic oscillator (e.g., Marion and Thornton 1995). We now check this idea and also attempt to explain the different values of the three predictability measures.

We modeled the damped harmonic oscillator as a multivariate autoregressive process (see appendix B). The state of the oscillator is defined by the first two elements of a four-dimensional state vector; the last two elements define the autocorrelated stochastic forcing that drives the oscillator away from a state of rest.

The damped harmonic oscillator model depends on three parameters (appendix B). The physical interpretation of the three parameters is straightforward as follows:

$P$  is the period of the dominant oscillation;  $\tau_1$  is the decay scale of the oscillation; and  $\tau_2$  is the decay scale of the forcing. Once they are defined it is straightforward to calculate the three predictability time scales defined above. We were encouraged to find that the time scales calculated from the oscillator model with a single choice of parameters ( $P = 50$  days,  $\tau_1 = 15$  days, and  $\tau_2 = 2.5$  days) were remarkably similar to the empirical measures (thick line, Fig. 15 and dashed line, Fig. 16). We were also encouraged to find that the same set of parameters predicts a spectral density of the MJO that agrees well with the spectral shape of  $\mathbf{I}_t^{\text{WH}}$  and  $\mathbf{I}_t^{\text{HR}}$  (diagonal panels of Fig. 4).

## 6. Summary and discussion

Using long records of surface pressure from the twentieth century reanalysis project (Compo et al. 2011) and a simple regression model, the Madden–Julian oscillation index of Wheeler and Hendon has been reconstructed over the period 1905–2008. Wheeler and

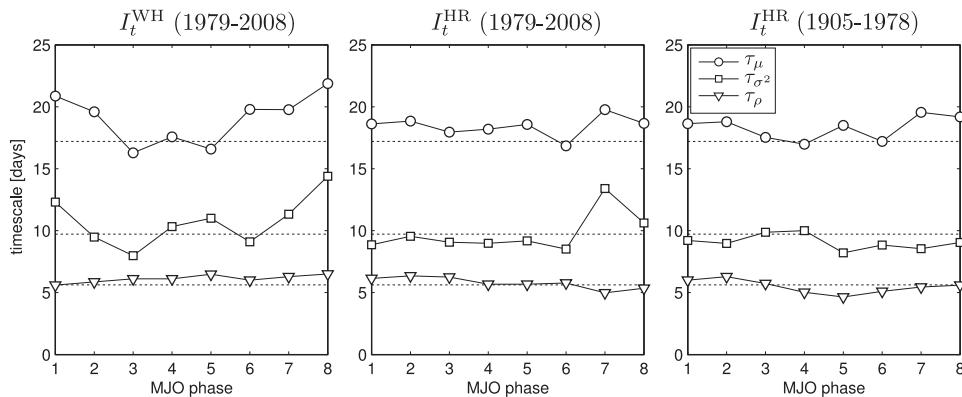


FIG. 16. Time scales for loss of predictability of a canonical MJO event. Ensembles were initialized for the same parameters as Fig. 15. The  $e$ -folding times for the mean (circles), variance (squares), and lag at which the correlation drops to 0.5 (triangles) as a function of MJO phase for (left)  $I_t^{\text{WH}}$ , (middle)  $I_t^{\text{HR}}$  over the 1979–2008 period, and (right)  $I_t^{\text{WH}}$  over the 1905–78 period. The dashed lines represent predictions from the damped harmonic oscillator model using  $\tau_1 = 15$  days,  $\tau_2 = 2.5$  days, and  $P = 50$  days.

Hendon generated their index using tropical fields of OLR and zonal wind. We chose surface pressure because of the availability of pressure observations over much of the last century and the long recognized connection of pressure to the MJO (Madden and Julian 1971, 1972).

The number of pressure predictors was limited by taking into account (i) the strength of their relationship with the MJO, (ii) decorrelation length scales of pressure, and (iii) the quality of the reanalysis. Twelve pressure predictors, and their Hilbert transforms, were chosen to reconstruct the MJO. A seasonally dependent model was also considered, but it was found that its impact on the skill of the regression was negligible. This is not surprising considering that the algorithm used by Wheeler and Hendon did not seasonally stratify their input data.

The reconstructed index accounts for 69% of the variance of Wheeler and Hendon’s index, has a similar spectral shape and lead–lag relationship between components, and captures the low-frequency changes in amplitude over their shared period (1979–2008). The reconstructed index was validated for the pre-1979 period by (i) examining the behavior of the MJO-reconstructed pressure fields using  $I_t^{\text{WH}}$  and  $I_t^{\text{HR}}$  and (ii) comparing the relationships between  $I_t^{\text{WH}}$  and  $I_t^{\text{HR}}$  and independent observations of environmental variables. First, it was found that the reconstructed index captures the intraseasonal, wavenumber-1, eastward-propagating signature in surface pressure that is known to be associated with MJO variability. The reconstructed index was also shown to capture the timing and strength of the connections between the MJO and cloud cover and surface wind over the ocean, extreme precipitation over Australia, and sea level in the western Pacific, the Gulf of Carpentaria, and the Pacific coast of the Americas. These connections were shown to

be consistent between  $I_t^{\text{WH}}$  and  $I_t^{\text{HR}}$  over the shared period (1979–2008) and stable over the historical period (1905–78, subject to data availability). These results give us confidence in the realism of our MJO reconstruction over the last century.

We also tested our reconstruction by examining time scales for the loss of predictability based on the decay of the mean  $\tau_\mu$ , increase in the variance  $\tau_{\sigma^2}$ , and loss of the correlation  $\tau_\rho$  of an ensemble of MJO events centered on a prescribed observation of the MJO index. We found that the time scales for the measures were quite different:  $\tau_\mu \simeq 18$  days,  $\tau_{\sigma^2} \simeq 10$  days, and  $\tau_\rho \simeq 5$  days. These results applied to both the Wheeler and Hendon index and the reconstructed index. The prediction from  $I_t^{\text{HR}}$  agreed over both the 1905–78 and 1979–2008 periods. For the Wheeler and Hendon index, we did observe some dependence of  $\tau_\mu$  on the phase of the MJO: greater predictability was found for events initialized in phases 1, 2, 6, 7, and 8. This phase dependence was not reproduced by the historical reconstruction.

The fact that the correlation time scale  $\tau_\rho$  is short compared to the other properties indicates that, for time scales more than a few days, there is no useful information in the initial position of an ensemble member relative to the initial ensemble mean and the best prediction is simply given by the mean and variance alone. It should be noted that  $\tau_\rho$  is strongly dependent on the observation error variance  $\sigma_{\text{obs}}^2$ . More specifically,  $\tau_\rho$  increases with  $\sigma_{\text{obs}}^2$ . When  $\sigma_{\text{obs}}^2$  is large, the initial ensemble fills a relatively large part of MJO space and an individual ensemble member will retain its position relative to the mean for a longer period of time than when  $\sigma_{\text{obs}}^2$  is small. Therefore, the size of  $\sigma_{\text{obs}}^2$  will determine the time scale over which a prediction of individual ensemble members

relative to the mean is useful. Beyond this time scale there is only value in forecasting the mean and variance.

The complex behavior of the three predictability measures, and their dependence on the observation error variance, can all be explained in terms of a simple damped harmonic oscillator model specified by three parameters. The damped harmonic oscillator model fits the original and reconstructed indices well with a single set of parameters:  $P = 50$  days,  $\tau_1 = 15$  days, and  $\tau_2 = 2.5$  days. Note, the ratio  $\tau_2/\tau_1 = 0.17$  is much less than one, and this indicates that the memory of the atmospheric forcing is much shorter than the lifetime of the MJO. More interestingly, the ratio  $\tau_1/P = 0.3$  is also less than one and thus the lifetime of an MJO event is predicted to be shorter than its period, suggesting that the MJO can be thought of as a sequence of pulse-like events rather than a quasi-periodic oscillation.

There are large variations in the density of surface pressure observations over the 1905–2008 period, and this led to the exclusion of many potential predictors from the regression model. If we were interested in creating an index spanning a more recent period, 1950–2008 for example, then we could have created a more accurate reconstruction for that period by including predictors that have been excluded here (e.g., pressure predictors from the western Pacific or other variables such as zonal wind). Although we recognize that there is an inevitable degree of subjectivity in the selection of predictors, we are confident, based on the validation studies that we have performed, that a different set of predictors that respect the restrictions outlined in this paper would yield similar results for the 1905–2008 period.

Future work may include the comparison of the long reconstructed MJO index with other climate indices and the examination of relationships with the MJO in other regions. For example, the long reconstruction could be used to examine low-frequency changes in the observed MJO connections in long station records (e.g., extratropical records, where the connection with the MJO is weak and unstable due to complex teleconnections). The reconstructed index may also shed light on interactions, possibly nonlinear, between the MJO and other climate phenomena such as El Niño/La Niña, the North Atlantic Oscillation, or the Pacific North American teleconnection pattern.

*Acknowledgments.* ECJO would like to thank Gil Compo, Daryl Metters, and Chidong Zhang for their generous assistance with provision of data and their scientific discussions. This study would not have been possible without access to results from the Twentieth Century Reanalysis Project. Support for the Twentieth Century Reanalysis Project dataset is provided by the

U.S. Department of Energy, Office of Science Innovative and Novel Computational Impact on Theory and Experiment (DOE INCITE) program, the Office of Biological and Environmental Research (BER), and the National Oceanic and Atmospheric Administration/Climate Program Office. ECJO and KRT acknowledge financial support from the GOAPP research network funded by the Canadian Foundation for Climate and Atmospheric Sciences. KRT also acknowledges financial support from the Discovery Grant program of the Natural Sciences and Engineering Research Council of Canada.

## APPENDIX A

### Restriction of Pressures for Inclusion in the Regression Model

Three restrictions limiting the number of predictors in the regression model are outlined below.

#### a. Strength of linear relationship with the MJO

The strength of the linear relationship between surface pressure and  $\mathbf{I}_t^{\text{WH}}$  is measured by

$$\bar{\kappa}_p^2 = \frac{\int [\kappa_{1,p}^2(\omega)f_1(\omega) + \kappa_{2,p}^2(\omega)f_2(\omega)] d\omega}{\int [f_1(\omega) + f_2(\omega)] d\omega}, \quad (\text{A1})$$

where  $\kappa_{1,p}^2$  and  $\kappa_{2,p}^2$  are the squared coherence between  $p_{ij}^a$  and the first and second components of  $\mathbf{I}_t^{\text{WH}}$  respectively, and  $f_1$  and  $f_2$  are the corresponding power spectral densities of the first and second components of  $\mathbf{I}_t^{\text{WH}}$ . The integral of  $f_1$  is the variance of the first component of  $\mathbf{I}_t^{\text{WH}}$ ; the integral of the product of  $f_1$  and  $\kappa_{1,p}^2$  is the variance of the first component of  $\mathbf{I}_t^{\text{WH}}$  that can be accounted for by  $p_{ij}^a$ . The statistic  $\bar{\kappa}$  is a number between zero and one representing the proportion of the total standard deviation of  $\mathbf{I}_t^{\text{WH}}$  that is accounted for by  $p_{ij}^a$  [see Oliver and Thompson (2010) for details].

This statistic is a function of latitude and longitude and is strongest in the tropical Indo-Pacific where  $\bar{\kappa}$  is about 0.65 (Fig. 1). There are also indications of the eastward-propagating nature of the MJO in the broad region of influence across the open Pacific Ocean as well as the blocking produced by the western slopes of the Andes and the East African Highlands. This pattern is consistent with Donald et al. (2006) who showed similar global behavior using composites of sea level pressure based on MJO phase. We have required all potential pressure predictors to have  $\bar{\kappa} \geq 0.33$ .

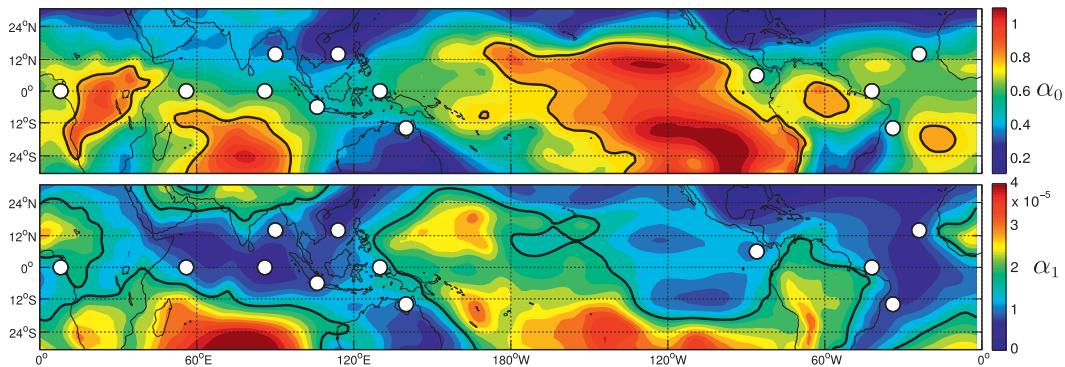


FIG. A1. Mean and trend in the relative error of surface pressure  $\Delta_{ijt}$ . (top) Mean of  $\Delta_{ijt}$  ( $\alpha_{0ij}$ ) with thick contour showing the critical value (0.8). (bottom) Trend of  $\Delta_{ijt}$  ( $\alpha_{1ij}$ ) with thick contour showing critical value ( $1.76 \cdot 10^{-5}$  days $^{-1}$ ) is shown. White circles show the chosen locations of predictors used in the regression model.

### b. Accuracy of the reanalysis pressure

The standard deviation of the 56-member ensemble at a given grid point and time  $\Delta p_{ijt}$  is a measure of the quality of the hindcast pressure. We have calculated the following relative measure of the quality of the hindcast pressure:

$$\Delta_{ijt} = \frac{\Delta p_{ijt}}{s_{ij}}, \quad (\text{A2})$$

where  $s_{ij}$  represents the standard deviation of  $p_{ijt}^a$  through time. This normalization allows for the fact that pressure is more variable at some locations.

To measure the overall level, and trend, of  $\Delta_{ijt}$  we approximated it by the model

$$\Delta_{ijt} = \alpha_{0ij} - \alpha_{1ij}(t - t_h) + \epsilon_t, \quad (\text{A3})$$

where  $t$  denotes time,  $t_h$  represents the midway point between 1905 and 2008, and  $\epsilon_t$  is an error term representing effects not included in the model. (The negative sign for  $\alpha_{1ij}$  was introduced because we anticipate that the trend in relative error will be negative, reflecting more accurate hindcasts in later years.) The coefficient  $\alpha_0$  measures the relative error in the middle of the last century (i.e.,  $t = t_h$ ). As expected it is generally low in regions with a large number of observations over 1905–2008 (top panel, Fig. A1). The coefficient  $\alpha_1$  reflects local changes in the observing system and tends to be low where the number of observations has been consistent over time (bottom panel, Fig. A1). We rejected predictors from regions where  $\alpha_0 > 0.8$  and  $\alpha_1 > 1.76 \times 10^{-5}$  days $^{-1}$  (thick contours, Fig. A1). The critical value for  $\alpha_1$  represents a change in  $\Delta_{ijt}$  of 67% over 1905–2008. The thick contour in Fig. 1 shows the intersection of the individual exclusion regions based on the critical values of  $\bar{\kappa}$ ,  $\alpha_0$ , and  $\alpha_1$ .

### c. Pressure decorrelation lengthscales

A set of sensitivity studies was carried out to determine the effect of spatial separation on the choice of pressures for inclusion as predictors in the regression model [Eq. (1)]. In one experiment the zonal and meridional separation of equi-spaced pressures along the equator, and two adjacent zonal lines, were varied and the impact on the fit of the regression was assessed. The regressions were trained onto the first half of the 1979–2008 period (1 January 1979–28 February 1994) and validated on the second half (1 March 1994–31 December 2008). A two-dimensional plot of the total error variance over the validation period as a function of longitudinal and latitudinal separation distance was constructed (not shown). As the longitudinal spacing decreases, more predictors are included; the regression is eventually overfit over the training period, and the skill reduces over the validation period. The error variance over the validation period reached a minimum with a longitudinal separation of  $18^\circ$  and a latitudinal separation of  $14^\circ$  (see Fig. A2, thick line). Excluding the Hilbert transform of each predictor significantly reduced the skill of the regression model for both components (Fig. A2, thin line). Switching the training and validation periods yielded essentially the same results.

Although the above sensitivity study is straightforward, it does not allow for geographically varying decorrelation scales. If we define the pressure decorrelation length scale as the distance required for the correlation of pressure to drop to 0.5 then the north–south decorrelation scale is fairly homogeneous with values between  $14^\circ$  and  $18^\circ$ . The east–west decorrelation scale, however, depends on both latitude and longitude (Table A1). When choosing the predictors included in the model, no two locations (white circles, Figs. 1 and A1) were chosen closer than the decorrelation scales defined in Table A1.

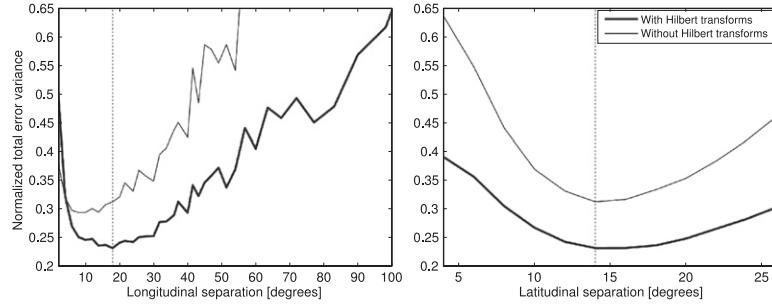


FIG. A2. Fit of the regression model to  $\mathbf{I}_t^{\text{WH}}$  over the validation period (1 Mar 1994–12 December 2008). (left) The total error variance is shown as a function of east–west spacing assuming a north–south spacing (of three lines) of  $14^\circ$ . (right) The total error variance is shown as a function of north–south spacing assuming an east–west spacing of  $18^\circ$ . Thin lines include only the pressure series as predictors in the regression model; thick lines include the pressure series and their Hilbert transforms as predictors. Note that the total error variance has been normalized by the total variance of  $\mathbf{I}_t^{\text{WH}}$  over the validation period.

## APPENDIX B

### Damped Harmonic Oscillator Model in Discrete Time

Consider the following multivariate autoregressive model of order 1:

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \boldsymbol{\varepsilon}_t, \quad (\text{B1})$$

where  $\mathbf{x}_t$  is a  $4 \times 1$  state vector at time step  $t$ ,  $\mathbf{A}$  is a  $4 \times 4$  transition matrix that carries the state vector forward one time step, and  $\boldsymbol{\varepsilon}_t$  is the noise that drives the system away from a state of rest.

The first two components of  $\mathbf{x}_t$  are taken to represent the MJO index, and the last two components are taken to represent the “forcing.” The transition matrix is assumed to be of the form

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{I} \\ 0 & \mathbf{A}_2 \end{bmatrix}, \quad (\text{B2})$$

where

$$\mathbf{A}_1 = \gamma_1 \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \mathbf{A}_2 = \gamma_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (\text{B3})$$

In the absence of forcing the MJO index rotates through an angle  $\theta$  over one time step, and its amplitude is scaled by a factor  $\gamma_1$ . For  $\theta \neq 0$ ,  $0 < \gamma_1 < 1$ , and  $\gamma_2 = 0$  this model represents damped harmonic oscillations (Priestley 1981). The rotation rate  $\theta$  can be expressed in terms of a rotation period  $P$  by writing it as  $\theta = 2\pi/P$  and the  $\gamma_i$  can be expressed in terms of decay time scales  $\tau_i$  by writing them in the form  $\gamma_i = \exp(-1/\tau_i)$ , that is,  $\tau_i = -1/\ln\gamma_i$ .

The forcing drives the MJO away from zero and is modeled as a pair of independent first-order autoregressive

[AR(1)] processes with autoregressive parameter  $\gamma_2$ . The forcing itself is driven by a bivariate, zero mean white noise process with variance:

$$\boldsymbol{\Sigma}_\varepsilon = \sigma_\varepsilon^2 \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{I} \end{bmatrix}. \quad (\text{B4})$$

The parameter  $\gamma_2$  controls the memory of the forcing, and we constrain it to be between 0 and 1. It too can be expressed as a decay time scale  $\tau_2$  by writing  $\gamma_2 = \exp(-1/\tau_2)$ .

The noise variance  $\sigma_\varepsilon^2$  was chosen to make the steady-state variance of the MJO components of  $x_n$  each unity. The model and its covariance structure are now defined by three model parameters ( $\theta$ ,  $\gamma_1$ , and  $\gamma_2$ ) or equivalently their time scales ( $P$ ,  $\tau_1$ , and  $\tau_2$ ).

Given these three parameters, it is straightforward to calculate the predictability properties of the above model. First, note that the process is asymptotically stationary to second order with mean zero and a covariance matrix  $\boldsymbol{\Sigma}_\infty$  that satisfies

$$\boldsymbol{\Sigma}_\infty = \mathbf{A}\boldsymbol{\Sigma}_\infty\mathbf{A}' + \boldsymbol{\Sigma}_\varepsilon, \quad (\text{B5})$$

TABLE A1. Longitudinal decorrelation length scales of  $p_{ijt}^a$  in each basin along the equator,  $14^\circ\text{N}$ , and  $14^\circ\text{S}$ . The length scale is a function of latitude and ocean basin and is given as a single value or range for each region. The basins are defined as lying between the three major topographic features in the tropics: the Andes, the East African Highlands, and the Maritime Continent.

	Indian	Pacific	Atlantic
$14^\circ\text{N}$	20°	20°–30°	20°
Equator	25°–30°	40°–50°	50°
$14^\circ\text{S}$	20°	20°–30°	40°

where ' denotes transpose. The solution is given explicitly by (e.g., Harvey 1991)

$$\text{vec}(\Sigma_\infty) = [\mathbf{I} - \mathbf{A} \otimes \mathbf{A}]^{-1} \text{vec}(\Sigma_\varepsilon). \quad (\text{B6})$$

Given an observation of the MJO at  $t_0$  this information can be used along with the prior distribution (i.e.,  $\Sigma_\infty$ ) to estimate a posterior distribution at  $t_0$  (see appendix C). If the mean and variance of the posterior distribution are  $\boldsymbol{\mu}_{t_0}$  and  $\Sigma_{t_0, t_0}$ , respectively, the mean and variance of  $\mathbf{x}_{t_0+k}$  are given by

$$\boldsymbol{\mu}_{t_0+k} = \mathbf{A}^k \boldsymbol{\mu}_{t_0} \quad (\text{B7})$$

and

$$\Sigma_{t_0+k, t_0+k} = \mathbf{A}^k \Sigma_{t_0, t_0} \mathbf{A}^{k'} + \sum_{j=0}^{k-1} \mathbf{A}^j \Sigma_\varepsilon \mathbf{A}^{j'}, \quad (\text{B8})$$

respectively, and the covariance between  $\mathbf{x}_{t_0}$  and  $\mathbf{x}_{t_0+k}$  is given by

$$\Sigma_{t_0+k, t_0} = \mathbf{A}^k \Sigma_{t_0, t_0}. \quad (\text{B9})$$

The squared correlation, calculated as the proportion of variance of the MJO state at  $t_0 + k$  that can be accounted for by the initial state, is given by

$$\rho_k^2 = \frac{\text{tr}(\Sigma_{t_0+k, t_0} \Sigma_{t_0, t_0}^{-1} \Sigma_{t_0, t_0+k})}{\text{tr}(\Sigma_{t_0+k, t_0+k})}, \quad (\text{B10})$$

where the trace refers only to the first two components. Finally, the  $4 \times 4$  cross-spectral density matrix of the process is given by (Priestley 1981)

$$\mathbf{h}_{\mathbf{xx}}(\omega) = [e^{i\omega} \mathbf{I} - \mathbf{A}]^{-1} \mathbf{h}_{\varepsilon\varepsilon}(\omega) [e^{i\omega} \mathbf{I} - \mathbf{A}]^{*-1}, \quad (\text{B11})$$

where \* denotes conjugate transpose.

## APPENDIX C

### Calculating the Posterior Distribution of the MJO Given an Imperfect Observation

Consider the multivariate autoregressive model discussed in appendix B. Assume the true MJO index and associated forcing at time  $t_0$  are stored in the first and last pair of elements of the four-dimensional state vector  $\mathbf{x}$ , respectively. Further assume that the prior distribution for  $\mathbf{x}$  at time  $t_0$  is normal with zero mean and covariance matrix  $\Sigma_{\text{prior}}$ . ( $\Sigma_{\text{prior}} = \Sigma_\infty$  for the model discussed in appendix B.)

Assuming a multivariate observation becomes available at time  $t_0$  and it is related to the state vector by  $\mathbf{y}_{\text{obs}} = \mathbf{x} + \nu$  where  $\nu$  is the observation error, which is taken to be normally distributed with zero mean and covariance matrix  $\Sigma_{\text{obs}}$ . It is straightforward to show that the posterior distribution of the state given this observation is also normal with mean and covariance

$$\boldsymbol{\mu}_{\text{post}} = \Sigma_{\text{prior}} (\Sigma_{\text{prior}} + \Sigma_{\text{obs}})^{-1} \boldsymbol{\mu}_{\text{obs}} \quad \text{and} \quad (\text{C1})$$

$$\Sigma_{\text{post}} = (\Sigma_{\text{prior}}^{-1} + \Sigma_{\text{obs}}^{-1})^{-1}. \quad (\text{C2})$$

These results follow directly from the product of a pair of normal probability densities, which is also normal.

Assume  $\mathbf{y}_{\text{obs}}$  and the observation error covariance matrix take the form

$$\mathbf{y}_{\text{obs}} = \begin{bmatrix} \mathbf{y}_{\text{obs}}^{\text{MJO}} \\ \mathbf{y}_{\text{obs}}^{\text{for}} \end{bmatrix}, \quad \Sigma_{\text{obs}} = \begin{bmatrix} \sigma_{\text{obs}}^2 \mathbf{I} & 0 \\ 0 & \alpha \mathbf{I} \end{bmatrix}, \quad (\text{C3})$$

where  $\alpha$  represents the variance of the uncertainty of the observed forcing. We will consider the case in which the observation error of the forcing becomes very large ( $\alpha \rightarrow \infty$ ). This implies that we only have observational information on the MJO and not on the associated forcing. What are the implications of the new observation? Can we say something about the forcing despite not observing it?

In the limit  $\alpha \rightarrow \infty$  the mean of the posterior distribution is given by

$$\boldsymbol{\mu}_{\text{post}} = \frac{1}{1 + \sigma_{\text{obs}}^2} \begin{bmatrix} \mathbf{I} \\ \Sigma_{21} \Sigma_{11}^{-1} \end{bmatrix} \mathbf{y}_{\text{obs}}^{\text{MJO}}, \quad (\text{C4})$$

where  $\Sigma_{ij}$  refers to the  $ij$ th block of  $\Sigma_{\text{prior}}$  when arranged in four  $2 \times 2$  blocks. Note that as  $\sigma_{\text{obs}}^2$  tends to infinity the posterior mean reverts to the prior mean as expected; as  $\sigma_{\text{obs}}^2$  tends to zero the posterior mean for the MJO is simply the observation itself and the mean forcing comes from a linear regression of the forcing components onto the MJO components. This tells us that, even in the absence of any observation of the forcing, an observation of the MJO can be used along with the prior to infer the mean of the forcing.

The covariance of the posterior distribution is given by

$$\Sigma_{\text{post}} = \frac{1}{1 + \sigma_{\text{obs}}^2} \begin{bmatrix} \sigma_{\text{obs}}^2 \Sigma_{\text{prior}} + \begin{pmatrix} 0 & 0 \\ 0 & \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} \end{pmatrix} \end{bmatrix}. \quad (\text{C5})$$

In the limit  $\sigma_{\text{obs}}^2$  tends to infinity and the posterior covariance reverts to the prior covariance as expected; in

the limit  $\sigma_{\text{obs}}^2$  tends to zero and the posterior covariance of the forcing is simply the variance of the residuals from a linear regression of the forcing components onto the MJO elements of the state vector.

## REFERENCES

- Compo, G., and Coauthors, 2011: The twentieth century reanalysis project. *Quart. J. Roy. Meteor. Soc.*, **137**, 1–28.
- Diky, L., and G. Golitsyn, 1968: Calculation of the Rossby wave velocities in the earth's atmosphere. *Tellus*, **20**, 314–317.
- Donald, A., H. Meinke, B. Power, M. Wheeler, A. Maia, R. Stone, J. Ribbe, and N. White, 2006: Near-global impact of the Madden–Julian oscillation on rainfall. *Geophys. Res. Lett.*, **33**, L09704, doi:10.1029/2005GL025155.
- Doodson, A., 1928: The analysis of tidal observations. *Philos. Trans. Roy. Soc. London*, **227**, 223–279.
- Gottschalck, J., and Coauthors, 2010: A framework for assessing operational model MJO forecasts: A project of the CLIVAR Madden–Julian Oscillation Working Group. *Bull. Amer. Meteor. Soc.*, **91**, 1247–1258.
- Hahn, C., and S. Warren, 1999: *Extended Edited Synoptic Cloud Reports From Ships and Land Stations Over the Globe*. Environmental Sciences Division, 76 pp.
- Harvey, A., 1991: *Forecasting, Structural Time Series Models, and the Kalman Filter*. Cambridge University Press, 554 pp.
- Jiang, X., D. Waliser, M. Wheeler, C. Jones, M. Lee, and S. Schubert, 2008: Assessing the skill of an all-season statistical forecast model for the Madden–Julian oscillation. *Mon. Wea. Rev.*, **136**, 1940–1956.
- Jones, C., and L. Carvalho, 2009: Stochastic simulations of the Madden–Julian oscillation activity. *Climate Dyn.*, **36**, 229–246.
- , —, R. Higgins, D. Waliser, and J. Schemm, 2004: A statistical forecast model of tropical intraseasonal convective anomalies. *J. Climate*, **17**, 2078–2095.
- Kang, I., and H. Kim, 2010: Assessment of MJO predictability for boreal winter with various statistical and dynamical models. *J. Climate*, **23**, 2368–2378.
- Kasahara, A., 1976: Normal modes of ultralong waves in the atmosphere. *Mon. Wea. Rev.*, **104**, 669–690.
- Kim, H., C. Hoyos, P. Webster, and I. Kang, 2009: Ocean–atmosphere coupling and the boreal winter MJO. *Climate Dyn.*, **35**, 771–784.
- Lin, H., G. Brunet, and J. Derome, 2008: Forecast skill of the Madden–Julian oscillation in two Canadian atmospheric models. *Mon. Wea. Rev.*, **136**, 4130–4149.
- Lo, F., and H. Hendon, 2000: Empirical extended-range prediction of the Madden–Julian oscillation. *Mon. Wea. Rev.*, **128**, 2528–2543.
- Madden, R., and P. Julian, 1971: Detection of a 40–50-day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702–708.
- , and —, 1972: Description of global-scale circulation cells in the tropics with a 40–50-day period. *J. Atmos. Sci.*, **29**, 1109–1123.
- Maharaj, E., and M. Wheeler, 2005: Forecasting an index of the Madden–Julian oscillation. *Int. J. Climatol.*, **25**, 1611–1618.
- Marion, J., and S. Thornton, 1995: *Classical Dynamics of Particles and Systems*. Holt Rinehart and Winston, 638 pp.
- Oliver, E., and K. Thompson, 2010: Madden–Julian oscillation and sea level: Local and remote forcing. *J. Geophys. Res.*, **115**, C01003, doi:10.1029/2009JC005337.
- , and —, 2011: Sea level and circulation variability of the Gulf of Carpentaria: Influence of the Madden–Julian oscillation and the adjacent deep ocean. *J. Geophys. Res.*, **116**, C02019, doi:10.1029/2010JC006596.
- Oppenheim, A., and Coauthors, 1989: *Discrete-Time Signal Processing*. Vol. 1999, Prentice Hall, 870 pp.
- Pohl, B., and A. Matthews, 2007: Observed changes in the lifetime and amplitude of the Madden–Julian oscillation associated with interannual ENSO sea surface temperature anomalies. *J. Climate*, **20**, 2659–2674.
- Priestley, M., 1981: *Spectral Analysis and Time Series*. Academic Press, 890 pp.
- Rayner, N., D. Parker, E. Horton, C. Folland, L. Alexander, D. Rowell, E. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, doi:10.1029/2002JD002670.
- Ristic, B., S. Arulampalam, and N. Gordon, 2004: *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House Publishers, 318 pp.
- Roundy, P., C. Schreck III, and M. Janiga, 2009: Contributions of convectively coupled equatorial Rossby waves and Kelvin waves to the real-time multivariate MJO indices. *Mon. Wea. Rev.*, **137**, 469–478.
- Shinoda, T., H. Hendon, and J. Glick, 1998: Intraseasonal variability of surface fluxes and sea surface temperature in the tropical western Pacific and Indian Oceans. *J. Climate*, **11**, 1685–1702.
- Vitart, F., S. Woolnough, M. Balmaseda, and A. Tompkins, 2007: Monthly forecast of the Madden–Julian oscillation using a coupled GCM. *Mon. Wea. Rev.*, **135**, 2700–2715.
- Waliser, D., 2005: Predictability and forecasting. *Intraseasonal Variability in the Atmosphere–Ocean Climate System*, Springer, 389–423.
- Wheeler, M., and G. Kiladis, 1999: Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber–frequency domain. *J. Atmos. Sci.*, **56**, 374–399.
- , and K. Weickmann, 2001: Real-time monitoring and prediction of modes of coherent synoptic to intraseasonal tropical variability. *Mon. Wea. Rev.*, **129**, 2677–2694.
- , and 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.
- , G. Kiladis, and P. Webster, 2000: Large-scale dynamical fields associated with convectively coupled equatorial waves. *J. Atmos. Sci.*, **57**, 613–640.
- Worley, S., S. Woodruff, R. Reynolds, S. Lubker, and N. Lott, 2005: ICOADS release 2.1 data and products. *Int. J. Climatol.*, **25**, 823–842.