An All-Season Real-Time Multivariate MJO Index: Development of an Index for Monitoring and Prediction

Matthew C. Wheeler and Harry H. Hendon
Bureau of Meteorology Research Centre, Melbourne, Australia

(Manuscript received 11 September 2003, in final form 10 February 2004)

ABSTRACT

A seasonally independent index for monitoring the Madden-Julian oscillation (MJO) is described. It is based on a pair of empirical orthogonal functions (EOFs) of the combined fields of near-equatorially averaged 850-hPa zonal wind, 200-hPa zonal wind, and satellite-observed outgoing longwave radiation (OLR) data. Projection of the daily observed data onto the multiple-variable EOFs, with the annual cycle and components of interannual variability removed, yields principal component (PC) time series that vary mostly on the intraseasonal time scale of the MJO only. This projection thus serves as an effective filter for the MJO without the need for conventional time filtering, making the PC time series an effective index for real-time use.

The pair of PC time series that form the index are called the Real-time Multivariate MJO series 1 (RMM1) and 2 (RMM2). The properties of the RMM series and the spatial patterns of atmospheric variability they capture are explored. Despite the fact that RMM1 and RMM2 describe evolution of the MJO along the equator that is independent of season, the coherent off-equatorial behavior exhibits strong seasonality. In particular, the northward, propagating behavior in the Indian monsoon and the southward extreme of convection into the Australian monsoon are captured by monitoring the seasonally independent eastward propagation in the equatorial belt. The previously described interannual modulation of the global variance of the MJO is also well captured.

Applications of the RMM series are investigated. One application is through their relationship with the onset dates of the monsoons in Australia and India; while the onsets can occur at any time during the convectively enhanced half of the MJO cycle, they rarely occur during the suppressed half. Another application is the modulation of the probability of extreme weekly rainfall; in the “Top End” region around Darwin, Australia, the swings in probability represent more than a tripling in the likelihood of an upper-quintile weekly rainfall event from the dry to wet MJO phase.

1. Introduction

A number of recent studies have demonstrated potential for skillful empirical prediction of the MJO, with useful lead times out to about 15–20 days (von Storch and Xu 1990; Waliser et al. 1999; Lo and Hendon 2000; Mo 2001; Wheeler and Weickmann 2001). Such skill is demonstrably greater than that of a number of current and previous operational numerical weather prediction (NWP) models (e.g., Waliser et al. 1999; Jones et al. 2000; Hendon et al. 2000). Links between the MJO and synoptic weather, such as tropical cyclone activity (e.g., Maloney and Hartmann 2000; Hall et al. 2001), North American summertime and wintertime precipitation events (e.g., Mo 2000; Whitaker and Weickmann 2001), South American precipitation (e.g., Paegle et al. 2000), and variations in the Australian monsoon (Hendon and Liebmann 1990b), have also been established.

This body of work suggests that operational monitoring of the MJO, and subsequent statistical prediction of its associated weather effects, can be a beneficial addition to the activities of national meteorological centers. Here we describe development of an index for such monitoring at the Australian Bureau of Meteorology. The index is intended to efficiently describe and extract the atmospheric variability directly related to the MJO. Importantly, the index shows definite relationships with the weather effects that are known to be related to the MJO. As a bonus, the index will also be applicable for the extraction of the MJO signal from the output of global NWP models.

Development of the index draws strongly upon the aforementioned empirical work. As described in those papers, the crux of the MJO-monitoring and prediction problem is the extraction of its frequency-limited signal without the use of a typical (e.g., Lanczos) bandpass filter. Such filters are restricted from this real-time task because of their requirement for information beyond the end of the time series. Alternative approaches must be taken.

The approach taken in this study is similar to that of Lo and Hendon (2000, hereafter LH00). They showed
that much of the signal of the MJO can be isolated in minimally filtered (high-pass only) daily data simply by projection of that data onto spatial patterns characteristic of the MJO. Through this spatial projection, a large portion of the variability on other time and space scales is removed. LH00 used empirical orthogonal functions (EOFs) of either outgoing longwave radiation (OLR) or streamfunction in the global Tropics to define their spatial patterns. Their index came from the temporal coefficients of a select few of the EOFs. While the MJO has long been identified through the use of EOF analysis on bandpass-filtered data (e.g., Lau and Chan 1985; Knutson and Weickmann 1987; Maloney and Hartmann 1998; Slingo et al. 1999; Matthews 2000; Kessler 2001), LH00’s use on daily data subject to high-pass filtering has been relatively unique. We thus further explore the use of EOF analysis on such daily data for extraction of the MJO in this study.

The spectral properties of the temporal coefficients [principal components (PCs)] of the EOFs reveal how effectively the MJO extraction is performed. By defining the MJO with EOFs of a single field, as in LH00, we find that the PCs are still undesirably influenced by noisy day-to-day weather variations. This influence can be substantially reduced through the use of EOFs of combined fields, for example, OLR and winds at multiple levels combined. Projection of the daily data onto the combined multivariable patterns increases the signal-to-noise ratio simply because there are fewer high-frequency weather variations that simultaneously have a structure similar to the MJO in multiple fields. Identifying and taking advantage of this result is the main new contribution of the work presented in this paper.

After much testing, the combination of fields chosen for our EOF representation of the MJO is OLR, 850-hPa zonal wind (u850), and 200-hPa zonal wind (u200), each averaged over the latitudes of 15°S–15°N. The meridional averaging itself helps to remove some of the non-MJO, higher-frequency variability. The domain of the EOF analysis covers all longitudes in this equatorial portion of the MJO (e.g., Madden and Julian 1972) can be captured using OLR (proxy for convection) and zonal winds in the upper and lower troposphere. We use combinations of these fields to detect the MJO. The OLR data are daily averaged values from the National Oceanic and Atmospheric Administration (NOAA) polar-orbiting series of satellites (Liebmann and Smith 1996). In real time the data are obtained directly from the National Centers for Environmental Prediction (NCEP). The data are continuous in time (after some interpolation for missing data) from June 1974 to 16 March 1978, and from 1 January 1979 to the present. The zonal wind data come from the NCEP–National Center for Atmospheric Research (NCEP–NCAR) reanalysis dataset (Kalnay et al. 1996) and are available for the same period as the OLR. Both the OLR and zonal winds are analyzed on a 2.5° latitude–longitude grid. Data from the 1979 to 2001 period (8401 days) are used to define the EOFs, while data from the whole period are used for the generation of the RMM indices by projection onto the EOFs.

For real-time applications, OLR is available within a day of when it is recorded. The reanalysis zonal winds, on the other hand, run several days behind, but may be replaced by winds from operational analyses in these situations, as will be discussed in section 6.

### 2. Data and methodology

#### a. Datasets

The baroclinic, convectively driven circulation in the equatorial plane of the MJO (e.g., Madden and Julian 1972) can be captured using OLR (proxy for convection) and zonal winds in the upper and lower troposphere. We use combinations of these fields to detect the MJO. The OLR data are daily averaged values from the National Oceanic and Atmospheric Administration (NOAA) polar-orbiting series of satellites (Liebmann and Smith 1996). In real time the data are obtained directly from the National Centers for Environmental Prediction (NCEP). The data are continuous in time (after some interpolation for missing data) from June 1974 to 16 March 1978, and from 1 January 1979 to the present. The zonal wind data come from the NCEP–National Center for Atmospheric Research (NCEP–NCAR) reanalysis dataset (Kalnay et al. 1996) and are available for the same period as the OLR. Both the OLR and zonal winds are analyzed on a 2.5° latitude–longitude grid. Data from the 1979 to 2001 period (8401 days) are used to define the EOFs, while data from the whole period are used for the generation of the RMM indices by projection onto the EOFs.

For real-time applications, OLR is available within a day of when it is recorded. The reanalysis zonal winds, on the other hand, run several days behind, but may be replaced by winds from operational analyses in these situations, as will be discussed in section 6.

#### b. Removal of longer-time-scale components

The influence of the seasonal cycle is removed by subtracting from each grid point the time mean and first

---

1. The reanalysis data used was that recreated after the correction to the processing of Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) data was made, as became available in April 2002.
three harmonics of the annual cycle, based on the 1979–2001 period. As with all the preparation steps, this subtraction can be performed in real time. We denote the anomaly fields created in this manner with a superscript ‘‘A,’’ for example, OLR$^A$ and u850$^A$.

Removal of interannual variability, especially that associated with ENSO, is done in two steps. As described by LH00, this removal is necessary because time mean anomalies associated with the mature phase of El Niño/La Niña resemble the phase of the MJO in which convection is centered near the date line/Maritime Continent. Without removal of the ENSO signal, one would assume that this phase of the MJO was continually occurring during an ENSO event. The first step is the subtraction of variability that is linearly related to a measure of ENSO in sea surface temperature (SST). The particular measure used is the time series of the first rotated EOF of Indo-Pacific SSTs (SST1) as described by Drosdowsky and Chambers (2001). SST1 is well correlated with other typical measures of ENSO (e.g., the Southern Oscillation index or Niño-3). Monthly values of SST1 are calculated operationally at the Bureau of Meteorology. These values are converted to a daily basis (SST1 changes little from month to month), and a different linear regression relationship is calculated with the daily field data at each grid point, separately for each month of the year. The resulting monthly regression parameters are then interpolated to a daily basis to form a 365-day seasonally dependent regression relationship that is subtracted from the value of each field at each grid point. Finally, to remove any further aspects of interannual variability, decadal variability, and trends, a 120-day mean of the previous 120 days is subtracted. Experimentation shows that while either one of these two steps perform well at removing unwanted interannual variability from the resulting MJO indices, both steps together perform appreciably better. We denote the fields resulting after both the seasonal cycle and interannual variability removal with a prime, for example, OLR$'$ and u850$'$.

c. EOF analysis

In many previous studies the MJO has been identified by use of EOFs of a single tropically confined field that has been bandpass filtered to intraseasonal periods (e.g., Maloney and Hartmann 1998; Slingo et al. 1999; Matthews 2000; Kessler 2001). The leading two EOFs generally appear as a pair, which taken together describe the large-scale eastward-propagating signal attributed to the MJO.

Here the EOFs are calculated for daily fields that have only been subject to the filtering implied by the removal of the long time-scale variability just described. For exploratory purposes, a number of different EOF analyses are computed. In all analyses we use input fields that encompass all longitudes around the globe. To preserve the spatial (in longitude) variation of the gridpoint variance of each field, we use the covariance matrix in the eigenvector computation. In the analysis chosen for our index definition, near-equatorially averaged (15°S–15°N) fields are used, like Maloney and Hartmann (1998) and Kessler (2001). We denote the anomaly fields that are averaged over this latitude range with an asterisk, for example, OLR$^*$.

When EOFs of combined fields (e.g., OLR and winds) are computed, each field is normalized by its global variance before input. This ensures that each field contributes equally to the variance of the combined vector.$^2$

3. Results of EOF analyses

a. The chosen RMM index EOFs

The spatial structures of the leading two EOFs of the combined fields of OLR$^*$, u850$^*$, and u200$^*$ are presented in Fig. 1. Together, EOF1 and EOF2 explain 25% of the variance of the original atmospheric fields and are well separated from the remaining EOFs, based on the criteria of North et al. (1982) (EOF3 explains only 6.1% of the variance). The explained variance of the leading pair is high considering the input dataset includes variability from three fields, for all seasons, and on all time scales from daily up to seasonal. Physically, EOF1 describes the familiar situation when the MJO produces enhanced convection (negative OLR anomalies) at the longitudes of the Maritime Continent: low-level westerly wind anomalies extend throughout the Indian Ocean region and Maritime Continent, and low-level easterlies exist across the Pacific, while upper-level wind anomalies are in the opposite direction to those below. EOF2 has enhanced convection over the Pacific Ocean and wind patterns that are in close quadrature to those of EOF1. Taken as a pair, these structures are consistent with those obtained by the previous studies of EOFs of the MJO. This demonstrates the dominance of the MJO spatial structure even in data that are not bandpass filtered.

Power spectra of the PCs of the leading EOF pair, together with that of the third (PC3), are shown in Fig. 2. The spectra are displayed in the area-conserving format in which variance is proportional to area (logarithm of frequency versus power times frequency; Zangvil 1977). The bulk of the variance of PC1 and PC2 is concentrated at intraseasonal periods (30–80 days) typically associated with the MJO (e.g., Salby and Hendon 1994). The fraction of total variance occurring in the 30- to 80-day range is ~0.6 for both PCs. These fractions are much greater than if the PC time series behaved as red noise (dashed curves), which would have a corresponding fraction of ~0.12. Thus, projection of the

$^2$ This is distinctly different from using the correlation matrix in the analysis, as the latter is equivalent to normalizing each point by its own variance (e.g., Kutzbach 1967), rather than the longitudinally averaged variance.
Fig. 1. Spatial structures of EOFs 1 and 2 of the combined analysis of OLR*, u850*, and u200*. A key for the field described by each curve is given. As each field is normalized by its global (all longitudes) variance before the EOF analysis, their magnitude may be plotted on the same relative axis. Multiplying each normalized magnitude by its global variance gives the field anomaly that occurs for a 1 std dev perturbation of the PC, as given for the absolute maxima of each field. The variance explained by the respective EOFs is 12.8% and 12.2%.

Fig. 2. Power spectra of the PCs of the leading three EOFs of the combined analysis of Fig. 1, as calculated using the whole time series. The plotting format forces the area under the power curve in any frequency band to be equal to variance. The total area under each curve is scaled to equal the explained variance (Exp Var) by that EOF. The fraction of Exp Var in the 30- to 80-day band for each PC is given. The dashed curve is the red-noise spectrum computed from the lag 1 autocorrelation. Multiple passes of a 1±2±1 filter are applied to all spectra resulting in an effective bandwidth of $3.0 \times 10^{-3}$ cpd (cycles per day).

daily observed data onto the two EOFs acts as an effective filter for the intraseasonal frequencies associated with the MJO. In contrast, PC3 is not dominated by the intraseasonal time scale, and thus presumably does not contain much, if any, of the MJO signal. Instead, the MJO, as manifested by a spectral peak at 40–50 days, is predominantly contained within EOF1 and EOF2, and it accounts for the majority of the variance of these two EOFs. As will be shown, our choice of variables used as input to this chosen EOF analysis was made based partly on the desire to maximize this variance.

A further useful diagnostic of the PCs is their cross-spectra, as presented in Fig. 3. The coherence squared ($Coh^2$) between PC1 and PC2 peaks in the 30- to 80-day range, with a mean value in this range of 0.76. Thus, not only do PC1 and PC2 have a greater portion of variance in this intraseasonal range, but their fluctuations in this frequency range are also more coherent. The phase relationship in this range (PC1 leads PC2 by 1/4 cycle) is consistent with the eastward propagation of the MJO. In contrast, the cross-spectrum between PC1 and PC3 shows only weak coherence (Fig. 3b), as
AUGUST 2004  WHEELER AND HENDON

Fig. 3. (a) Coherence squared and phase between PC1 and PC2 of the EOF analysis of Fig. 1. Multiple passes of a 1–2–1 filter were applied to the co- and quadrature spectra before computing the phase and coherence resulting in an effective bandwidth of 3.0 × 10⁻² cpd. The 0.1% confidence level on the null hypothesis of no association is 0.23. The mean coherence squared in the 30- to 80-day-period range is shown. For the phase, a 90° relationship means that PC1 leads by a quarter cycle. (b) As in (a), except for the cross-spectrum between PC1 and PC3.

Fig. 4. As in Fig. 2, except for the mean spectrum of PC1 and PC2 of the EOF analysis when the input fields are OLR*, only (dotted curve), OLR* and u850*, (dashed), and OLR* and u50*, and u200* (solid). The scaling for the solid curve is identical to Fig. 2, while that for the other two curves is such that they each have the same value of maximum variance. This scaling aids the visual comparison of the relative contributions of variance coming from the different frequency bands.

b. EOFs of other fields/combinations

Our aim has been to produce an optimal method for extracting the MJO from daily observations. Hence it is important to demonstrate that the above-described EOF analysis achieves this aim.

The first point to demonstrate is how the use of EOFs of the combined fields is able to increase the signal-to-noise ratio compared to that obtained when using a single field only. To look at this we have recomputed the EOF analysis using only OLR*, and another analysis with OLR* and u850* combined. The mean power spectra of PC1 and PC2 for each of these analyses, as well as that of the three-field analysis, are shown in Fig. 4. By projecting the daily data onto a single field only, the contribution of variance coming from frequencies outside the 30- to 80-day-period range is substantially increased, especially for higher frequencies. Equivalently, the signal-to-noise ratio is reduced. The fraction of variance in the 30- to 80-day range is only 0.45 for the single field, 0.55 for the two fields, and (as previously mentioned) 0.6 for the three fields. Thus, most of the improvement comes from inclusion of just the second field, yet the inclusion of the third field, which effectively isolates baroclinic structures, also provides some improvement. The latter is especially of note around the 5-day period, where the restriction to baroclinic structures removes the influence of the external vertical structured “5-day wave” (e.g., Madden 1979). We chose to use three fields as a trade-off between maximizing the signal-to-noise ratio and simplicity of the scheme.

More results from the above-described EOF analyses, together with those of others, are summarized in Table 1. The first three analyses listed are those just described. The coherence squared between the PCs in the 30- to
properties of the RMM indices and the variability they capture

4. Properties of the RMM indices and the variability they capture

a. Time variation

Even though the EOF analysis was computed only for the 1979–2001 period, we may project data from any available period onto the EOF structures to get the RMM1 (PC1) and RMM2 (PC2) values. Normalized (by their 1979–2001 standard deviations) values for RMM1 and RMM2 are used in the subsequent analysis.

Time series of RMM1 and RMM2 for 2001 and 2002 are shown in Fig. 5. Intraseasonal variations associated with the MJO are obvious, with RMM2 lagging RMM1 by 10–15 days. Consistent with our analysis of variance, however, there is still day-to-day noise in the series, which is especially noticeable during times that the MJO is weak or nonexistent, for example, in March and April of 2001. However, MJO-induced fluctuations are relatively easy to monitor in these daily time series, as has been our experience operationally.

Lag correlations between RMM1 and itself, and with RMM2, are shown in Fig. 6. The maximum correlation between RMM1 and RMM2 is 0.56 at a lag of 9 days. This correlation may be compared with those of other studies that have identified the MJO with EOFs: compared to LH00 and Ferranti et al. (1990), both of whom used daily input data with no low-pass filtering, it is higher; compared to Maloney and Kiehl (2002), who used 30- to 90-day bandpass-filtered input data, the peak correlation is within the limit of precision they offer. This correlation may be compared with those of other studies that have identified the MJO with EOFs: compared to LH00 and Ferranti et al. (1990), both of whom used daily input data with no low-pass filtering, it is higher; compared to Maloney and Kiehl (2002), who used 30- to 90-day bandpass-filtered input data, the peak correlation is within the limit of precision they offer. The higher lag correlation of this study compared to LH00 and Ferranti et al. (1990), both of whom used daily input data with no low-pass filtering, it is higher; compared to Maloney and Kiehl (2002), who used 30- to 90-day bandpass-filtered input data, the peak correlation is within the limit of precision they offer.
Fig. 6. Lag correlations between RMM1 and itself, and with RMM2 for all seasons.

Fig. 7. (RMM1, RMM2) phase space points for all available days in DJF season from 1974 to 2003. Eight defined regions of the phase space are labeled, as is the region considered to signify weak MJO activity. Also labeled are the approximate locations of the enhanced convective signal of the MJO for that location of the phase space, e.g., the “Indian Ocean” for phases 2 and 3.

tive). This higher value is due to the greater ability of the combined EOFs at extracting just the variability associated with the MJO and removing the noise that acts to weaken the correlation. Overall, the high correlations of Fig. 6, especially those at longer lags, are an indication of the potential predictability of these indices using one or the other as a predictor.

b. (RMM1, RMM2) phase space

Given the lead–lag behavior of the RMM indices, it is convenient to diagnose the state of the MJO as a point in the two-dimensional phase space defined by RMM1 and RMM2. This representation is shown for all days in the December–January–February (DJF) season in Fig. 7. Points representing sequential days are joined by a line. Many of the sequential days trace anticlockwise circles around the origin, which signifies systematic eastward propagation of the MJO. Large-amplitude circles signify strong cycles of the MJO, while weak MJO activity appears as rather random motions near the origin.

For compositing and other applications, numbered phases are defined in Fig. 7, together with an indication of the location of the enhanced convection of the MJO for each quadrant of the diagram.

c. Composites

The full spatial patterns of atmospheric variability captured by the RMM indices can be explored through the use of composites. Here we composite by taking the average of the observed anomaly fields (i.e., OLR, \( u_{850} \), and \( v_{850} \)) occurring for the days that fall within each of the labeled phases of Fig. 7. The composites for the DJF season and May–June (MJ) season are shown in Figs. 8 and 9, respectively. These figures serve as a useful comparison to previous studies (cited below), yet it is important to emphasize that the RMM indices are somewhat unique in that they are defined independent of season. As will become apparent, however, the off-equatorial structure captured by these indices is still a strong function of season.

The DJF-season composite (Fig. 8) depicts the familiar structure and evolution of the MJO (e.g., Weickmann et al. 1985; Ferranti et al. 1990; Hendon and Salby 1994; Sperber et al. 1999; Hendon 2000; Waliser et al. 2003; Innes and Slingo 2003). The nominal time for transition between each of the numbered phases is 6 days but can vary from event to event. In phase 1, convection of a decaying MJO event is present in the central Pacific, while enhanced convection of a growing event is evident over Africa and the western Indian Ocean. At this point westerly wind anomalies exist over the Pacific, and easterlies exist over the Indian Ocean. Over subsequent phases convection in the Indian Ocean builds and moves to the east, and as it passes by the Australian landmass, shifts southward to be most concentrated at about 15°S (phase 5). Eastward movement of the 850-hPa wind anomalies is quicker. Thus, enhanced convection in phases 2 and 3 (over the Indian Ocean) is in near quadrature with the winds, while that in phase 7 is wholly within the westertlies (over the Pacific). All such aspects of this com-
Fig. 8. DJF composite OLR and 850-hPa wind vector anomalies. Shading levels denote OLR anomalies less than -7.5, -15, -22.5, and -30 W m$^{-2}$, respectively, and hatching levels denote OLR anomalies greater than 7.5, 15, and 22.5 W m$^{-2}$, respectively. Black arrows indicate wind anomalies that are statistically significant at the 99% level, based on their local standard deviation and the Student’s $t$ test. The magnitude of the largest vector is shown on the bottom right, and the number of days (points) falling within each phase category is given.

Composite are in agreement with the knowledge accumulated by the previous studies. We thus believe that the RMM indices, despite discriminating to eastward propagation along the equator, are effective for capturing the southern summer MJO. Further confidence of the effectiveness of this new MJO index is the magnitude of the composited anomalies; they are as large or larger than those of previous studies.
For the northern summer season, we show a composite for the early season months of May and June (Fig. 9). This composite can be compared directly to that displayed by Kemball-Cook and Wang (2001) in their Fig. 4. Their composite was computed for data from the same months but using an index of intraseasonal variability defined locally in the equatorial Indian Ocean. Both composites show eastward and northward propagation of the convection in the Indian Ocean sector (see also Lawrence and Webster 2002). They also both show eastward and northward movement of easterly wind anomalies into India and the Bay of Bengal ahead of the convection (phases 1 to 3), and westerlies within and behind the convection (phases 3 to 5). The mag-
nitudes of the composit anomalies are also essentially the same. Thus, RMM1 and RMM2 appear to be effective at identifying this typical intraseasonal variability of the Indian monsoon as well. Being an index that uses information at all longitudes, however, it is also effective at picking up the global signature of the MJO, for example, the variability of the eastern Pacific ITCZ as seen also in the work of Maloney and Hartmann (2000).

d. Reconstructed fields

The three fields used as input to the EOF analysis may be reconstructed by summing over the two EOFs multiplied by their respective RMM index values for any desired time. An example for the OLR field is shown in Fig. 10 for the period 1 October 1987 to 31 October 1988. Although no bandpass time filtering has been applied, Fig. 10 clearly shows the signal of the MJO during this period (cf. Hendon and Liebmann 1994; Matthews 2000). A number of relatively strong oscillations are apparent from November 1987 to April 1988. Weak activity occurs from May through early September 1988, and moderate activity occurs afterward. Interestingly, the strongest oscillations correspond to an apparent zonal shift of activity into the Pacific. For example, the 5 $W m^{-2}$ contour penetrates eastward to about $160^\circ W$ during the February to early April 1988 period. At other times (e.g., November–December 1987), the eastward penetration is only to about the date line. Changes in the eastward penetration of MJO activity into the central Pacific, which have been suggested to be related to the evolution of ENSO, have been previously discussed by Gutzler et al. (1994) and Kessler (2001). Our use of just two EOFs, however, precludes the full diagnosis of such shifts without a corresponding change in activity elsewhere.

e. Wavenumber–frequency spectra

One possible side effect of using EOF spatial structures to extract the signal of the MJO is that any other oscillating phenomenon that has the same, or similar, structure will also be extracted by projection onto those EOFs. As has already been mentioned, ENSO is one such phenomenon that, without its prior removal, would project onto the chosen RMM EOF spatial structures. Other possibilities, at the higher-frequency end of the spectrum, are the various convectively coupled equatorial waves, several of which have baroclinic circulations in the equatorial plane like that of the MJO (e.g., Wheeler et al. 2000).

Given the constraint on symmetry imposed by taking a $15^\circ S$–$15^\circ N$ meridional average, only the “symmetric” equatorial waves, as discussed by Wheeler et al. (2000), need be of concern to us here. Of the symmetric waves, only the Kelvin and $n = 1$ equatorial Rossby have large enough variance (Fig. 3 of Wheeler et al. 2000), to have the potential to significantly contribute to the variability contained within RMM1 and RMM2. We thus investigate the presence of these waves within the RMM EOF-extracted data.

Following the method of Wheeler and Kiladis (1999), wavenumber–frequency power spectra serve as a useful diagnostic of the presence of the aforementioned waves. The wavenumber–frequency power spectra of the original OLR anomaly data (OLR$^A$), and of that reconstructed from the two EOFs, is shown in Fig. 11. The signature of the MJO as a spectral peak for eastward-propagating zonal wavenumbers 1 through 3, and periods between about 30 to 80 days, is clear in both spectra. It is only in the spectrum of the original OLR data, however, that the signatures of the convectively
Coupled Kelvin and Rossby waves can be discerned. Indeed, as we desired, the projection of daily data onto the RMM EOFs effectively removes much of the signal of both the “pure” convectively coupled Rossby and Kelvin waves. Importantly though, any elements of coupled Kelvin–Rossby dynamical structures moving along with the MJO (as in the model of Wang and Rui 1990) are still retained by the RMM EOFs, as evidenced by Figs. 8 and 9.

Why the individual Kelvin and Rossby waves have been removed deserves explanation. By the analysis of Wheeler et al. (2000), the convectively coupled Rossby wave has a deeper, more barotropic structure than the MJO, which would serve to reduce its projection onto the predominantly baroclinic RMM EOFs. For the convectively coupled Kelvin wave, on the other hand, its coupled signals in convection and winds tend to appear at the same zonal wavenumbers, whereas for the MJO the signal in convection peaks at a higher wavenumber ($s \approx 1$ to 2) than the winds ($s \approx 1$). Thus the Kelvin wave also tends to have a reduced projection onto the RMM EOFs.

On the whole, the result of projecting onto the RMM EOFs produces only a small amount of variance outside the usual range ascribed to the MJO (30–80 days and eastward propagating), and that outside variance is almost exclusively all at zonal wavenumber 1 (Fig. 11a). Interestingly, some westward-propagating variance is retained, some of which likely contributes to the small apparent standing component of the MJO that sometimes appears (cf. Zhang and Hendon 1997). Westward propagation is also occasionally discerned in the (RMM1, RMM2) phase space, but in practice it has never been observed to last longer than several days. Together, these power spectra further our confidence in the RMM EOFs as an effective means for extracting the signal of the MJO.

### f. Interannual modulation

Interannual modulation of the level of MJO activity, and its possible predictability, has been the subject of several recent papers (e.g., Hendon et al. 1999; Slingo et al. 1999). Slingo et al. (1999) used as an index of MJO activity the 101-day running variance of bandpass (20–100 days)-filtered 200-hPa zonal mean zonal wind averaged for 10°N–10°S. Their use of this index was based on the postulation that changes in the zonal mean zonal wind represent a synthesis of how intraseasonal variability in the atmospheric diabatic heating translates into modification of the planetary-scale circulation patterns. A more direct measure of the global variance of the MJO would presumably come from a running variance of RMM1 and RMM2, as presented in Fig. 12.

As in Slingo et al.’s index, the tendency for the MJO to be more active during southern summer and autumn is evident in Fig. 12, as are many of the major bursts of MJO activity such as during 1979, early 1981, late 1984 to early 1985, early 1986, early 1988, etc. One major difference between this new variance series and that of Slingo et al.’s however, is during the southern summer of 1992/93. Unlike the activity series calculated from the RMM values (Fig. 12), which has a moderately large peak, no peak is seen in Slingo et al.’s index. This period happens to coincide with the much-studied intensive observation period of the Tropical Ocean Global Atmosphere Coupled Ocean–Atmosphere Response Experiment (TOGA COARE) (Webster and Lukas 1992).
FIG. 12. Time series of 91-day running mean RMM$^1$ + RMM$^2$, showing the low-frequency (primarily interannual) modulation of the variance of the global signal of the MJO.

Fig. 13. (RMM1, RMM2) phase points (as marked by the small dual numbers) for the days on which the monsoon was defined to onset at Darwin, Australia. The dual numbers refer to the “monsoon” year, being that of the nearest Dec. The monsoon onset dates are as defined and taken from (with updates) Drosdowsky (1996), covering all monsoons from 1974/75 to 2001/02. MJO phases (large single numbers) are labeled according to Fig. 7.

Authors who have studied this period have claimed that there were at least two pronounced MJO events during this time (e.g., Gutzler et al. 1994; McBride et al. 1995; Yanai et al. 2000). Thus the RMM appear to be a more direct index.

One focus of the aforementioned interannual modulation studies has been the relationship, if any, between the MJO and ENSO. Both Slingo et al. (1999) and Hendon et al. (1999) found no significant relationship between the global level of MJO activity and ENSO. This result does not change with the use of the RMM-derived activity index (Fig. 12), the correlation between it and SST1 (our measure of ENSO) being only $\sim 0.02$. As shown by Hendon et al. (1999) and Kessler (2001), however, there is a local response of the MJO to ENSO in the Pacific that appears as about a 20° longitude eastward shift of MJO activity past the date line during El Niño events. Kessler (2001) used a third EOF of OLR in the tropical strip to capture such shifts. Despite the appearance of zonal expansions of activity appearing in OLR reconstructed from the two RMM EOFs, as already presented (e.g., Fig. 10), our use of just two EOFs precludes the full extraction of such shifts.

5. Applications to synoptic weather

One of the motivations for developing the RMM index has been its potential use in statistical prediction of
the MJO’s associated weather effects. It is important to first demonstrate that a relationship of the index to such weather effects exists. Here we present a few examples.

a. Monsoon onsets

The MJO has often been related to the onset of the Indian and Australian monsoons (e.g., Lorenc 1984; Hendon and Liebmann 1990a). Here we take dates of each year’s monsoon onsets, as defined by others, and relate them to the phase of the MJO, as defined by the (RMM1, RMM2) phase space (Figs. 13 and 14).

Dates for onset of the Australian monsoon were taken from the record objectively derived (with updates) by Drosdowsky (1996). It is a locally specific onset date, being computed solely from the zonal wind recorded by balloon soundings at Darwin, Australia (12.5°S, 130.9°E), over a deep layer (surface to 500 hPa). A definite relationship with Drosdowsky’s onset dates appears (Fig. 13). Considering only onsets that occur when the MJO is nonweak (i.e., those that lie outside the unit circle), more than 80% of the dates occur in phases 4–7 (when MJO-enhanced convection is in the vicinity of northern Australia; Fig. 8) and less than 20% of the onset dates occur in the other phases (when northern Australia is under the influence of the suppressed convective phase of the MJO).

Dates for onset of the Indian monsoon, as determined by the Indian Meteorological Department (IMD), are based on rainfall over the region of Kerala (southwest India). Joseph et al. (1994) lists the dates up until 1989, and we obtained dates for subsequent years courtesy of B. N. Goswami of the Indian Institute of Science. As for the Australian summer monsoon, onset at Kerala, India, tends to occur during the enhanced-convective phase of the MJO (i.e., phases 2 and 3; Fig. 9), and it tends not to occur during the highly suppressed phases (phases 6–8; Fig. 14). Obviously, the spread of onset dates over a number of phases indicates that there are other influences besides the MJO that ultimately determine the monsoon onset (e.g., Joseph et al. 1994; Goswami and Ajaya Mohan 2001).

b. Probabilities of extreme weekly rainfall

Recent work has demonstrated a relationship between the occurrence of extreme rainfall in North America and the MJO (e.g., Jones 2000; Whitaker and Weickmann 2001). Given that the MJO is associated with variations in synoptic-scale features across Australia (e.g., Hall et al. 2001), a modulation in extreme rainfall events there is also expected. Here we demonstrate the contemporaneous relationship between the occurrence of the highest quintile of weekly rainfall across Australia and the MJO as measured by the (RMM1, RMM2) phase.

Using a 1° gridded dataset of overlapping weekly rainfall totals, the threshold value for the highest quintile of weekly rainfall for the DJF season is found to range from only 10 mm in the central Australian desert up to greater than 130 mm along parts of the northern and eastern coasts (Fig. 15). The normal probability of a weekly rainfall total exceeding this value (in the DJF season) is, by definition, 20%. The conditional probability of the weekly rainfall exceeding this threshold, stratified by the phase of the MJO, is displayed in

---

5 The rainfall dataset we use was constructed by averaging the daily data from all available rainfall stations from the Bureau of Meteorology’s National Climate Centre archives into 1° grid boxes. Some quality control was performed. Boxes with missing data (primarily in the data-sparse regions of central Australia) are filled by a series of interpolation steps (W. Drosdowsky 2003, personal communication). We have used all available data from 1974 to 1999.
Fig. 16. Probability of weekly rainfall in highest quintile conditioned upon the phase of the MJO when described by the daily (RMM1, RMM2) value. The rainfall weeks are centered on the day upon which the (RMM1, RMM2) phase is determined. The probabilities are computed for overlapping weeks in the DJF season for all available data in the 1974–99 period.

Fig. 16. The probability of such an extreme event varies greatly with the phase of the MJO. For example, in the “Top End” region around Darwin (central north of the Australian continent), the probability varies from less than 12% in phases 1 and 2 to greater than 36% in phases 5 and 6. This represents more than a tripling of the likelihood of extreme rainfall from the dry to wet phase of the MJO. In some locations (e.g., Arnhem Land and northern tip of Cape York, Australia) the relative change of probabilities is as great as 6 times.

Besides the large swings in probability of extreme rainfall across northern Australia, the MJO induces an interesting swing in probability across extratropical Australia. Increased rainfall probability in phase 3 begins in the latitude band south of about 25°S, especially in the west, and progresses to the north and east by phase 4. It is not until phase 5 that the direct tropical signal of the MJO engulfs the northwestern part of the continent (north of about 20°S). The maximum increased rainfall probability then shifts to the east be-
between phases 5 and 6. Tropical–extratropical interaction processes (e.g., Matthews and Kiladis 1999) appear to produce such increased extratropical rainfall ahead of the main tropical signal, an aspect that would be interesting to pursue in future studies.

6. Real-time calculation of RMM1 and RMM2

The NCEP–NCAR reanalysis data are produced several days (typically 3) behind real time because of the time that is waited to get a more complete set of global observations. We currently obtain the reanalyses, as they are created, from the NOAA Climate Diagnostics Center. The OLR data, on the other hand, are typically more up-to-date, being obtained directly from NCEP within a day of when they are recorded. To optimize the real-time nature of the RMM indices, we use the analyses from the Australian Bureau of Meteorology’s operational model, called the Global Assimilation and Prediction system (GASP), to calculate the latest estimates of RMM1 and RMM2. Once the reanalysis data are obtained, these estimates are replaced by the values of RMM1 and RMM2 computed with them. Even though we still subtract the reanalysis climatology when using the GASP winds, the GASP-estimated RMM1 and RMM2 values closely correspond (correlations of 0.99) to those that are calculated later with the reanalysis winds.

7. Conclusions

This study has been motivated by the need to monitor and predict the MJO in real time. For this purpose we have taken the approach of monitoring the MJO by projecting daily observations onto the leading EOFs that describe its basic baroclinic, eastward-propagating structure in the equatorial-height plane. Compared to using EOFs calculated from a single level field, EOFs of combined fields of equatorially averaged OLR and zonal winds more effectively discriminate to the MJO signal.

The principal components of the leading pair of EOFs, RMM1 and RMM2, have been shown to be useful indices of the MJO and related variability. Although they still contain day-to-day noise, intraseasonal fluctuations due to the MJO dominate the variation of RMM1 and RMM2. Three-dimensional composites, which vary greatly from season to season, closely resemble those that have been produced by other studies using indices that have been season specific, hence our confidence that they are effective for monitoring the oscillation in all seasons. Zonal expansions and contractions of MJO activity that occur from season to season and year to year are also captured by the two RMM indices, and the use of the RMM indices for a measure of the global variance and interannual modulation of the MJO is thought to be better than that of some previous studies.

It will be left to a future paper to investigate the predictability of the RMM indices themselves. Using lagged multiple linear regression in the same way as was applied by LH00, we have already looked at the prediction of RMM1 and RMM2 using themselves as predictors. Our initial results show a slight improvement over the predictions of LH00, presumably reflecting the fact that the RMM indices contain less noise than the predictors and predictands of the LH00 study. This is a good demonstration of the importance of developing as accurate an index of the MJO as possible for statistical prediction, as has been done in this study.

Acknowledgments. We thank Wasyl Drosdowsky for producing the Darwin onset dates and the daily gridded rainfall data. B. N. Goswami kindly provided us with the latest Kerala onset dates. The NCEP–NCAR reanalysis data were obtained from the NOAA Climate Diagnostics Center (CDC) (http://www.cdc.noaa.gov), and we thank John Janowiak of NCEP for his help in maintaining the daily updated OLR data.

REFERENCES


NCEP retrieves the OLR data from the National Environmental Satellite, Data, and Information Service (NESDIS).


