Real-Time Extraction of the Madden–Julian Oscillation Using Empirical Mode Decomposition and Statistical Forecasting with a VARMA Model

BARNABY S. LOVE

School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom

ADRIAN J. MATTHEWS

School of Environmental Sciences, and School of Mathematics, University of East Anglia, Norwich, United Kingdom

GARETH J. JANACEK

School of Computing Sciences, University of East Anglia, Norwich, United Kingdom

ABSTRACT

A simple guide to the new technique of empirical mode decomposition (EMD) in a meteorological–climate forecasting context is presented. A single application of EMD to a time series essentially acts as a local high-pass filter. Hence, successive applications can be used to produce a bandpass filter that is highly efficient at extracting a broadband signal such as the Madden–Julian oscillation (MJO). The basic EMD method is adapted to minimize end effects, such that it is suitable for use in real time. The EMD process is then used to efficiently extract the MJO signal from gridded time series of outgoing longwave radiation (OLR) data.

A range of statistical models from the general class of vector autoregressive moving average (VARMA) models was then tested for their suitability in forecasting the MJO signal, as isolated by the EMD. A VARMA (5, 1) model was selected and its parameters determined by a maximum likelihood method using 17 yr of OLR data from 1980 to 1996. Forecasts were then made on the remaining independent data from 1998 to 2004. These were made in real time, as only data up to the date the forecast was made were used. The median skill of forecasts was accurate (defined as an anomaly correlation above 0.6) at lead times up to 25 days.

1. Introduction

The Madden–Julian oscillation (MJO) is the dominant mode of intraseasonal tropical convective variability (Madden and Julian 1994; Zhang 2005) having a significant influence on precipitation patterns over the tropical Indian Ocean, the Maritime Continent and the western Pacific warm pool region. The MJO also accounts for approximately 50% of the variability of the "active" and "break" phases of the South Asian monsoon (Goswami 2005) and has a strong influence on the phases and intensity of the Australian summer monsoon (Wheeler and McBride 2005) and a significant influence on the genesis of tropical cyclones (Maloney and Hartmann 2000; Hall et al. 2001). The MJO can essentially be characterized as an eastward propagation of tropical deep convective precipitation anomalies over the warm pool from the equatorial Indian Ocean over the Maritime Continent into the western Pacific region. One complete cycle of the MJO lasts between 30 and 60 days.

The ability to accurately forecast such a significant tropical mode as the MJO will be crucial to the success of medium- to extended-range numerical weather prediction (Hendon et al. 2000). The limit of predictability of a quasi-periodic phenomenon may be expected to be approximately equal to its period, that is, approximately 45 days for the MJO. This concept of inherent predictability was examined by Waliser et al. (2003) through a series of twin predictability experiments in a general circulation model that approximately identified

Corresponding author address: Barnaby Love, School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, United Kingdom.
E-mail: b.love@uea.ac.uk

intrinsic predictability limits of rainfall and 200-hPa velocity potential related to the MJO to be 10–15 days and 20–30 days, respectively. Hence, statistical models may have the potential to provide an alternative or complementary approach to forecasting the MJO (Waliser et al. 1999).

However, numerical weather prediction models are not attaining anywhere near the limits of useful skill we might expect in predicting the MJO (Jones et al. 2000; Waliser 2005). Ferranti et al. (1990) and Hendon et al. (2000) have demonstrated that the medium- to extended-range forecast skill of dynamical models in the extratropics and tropics can be significantly improved if the errors associated with the shortcomings in representation of tropical convection associated with the MJO are reduced. Useful forecast skill of the MJO in operational dynamically based numerical models is currently obtained up to lead times of about 7–10 days (Jones et al. 2000). Forecast skill is improved when an atmospheric model is coupled to an ocean mixed layer model with high vertical resolution, allowing an accurate simulation of surface flux anomalies (Woolnough et al. 2007). This current limit is probably a symptom of the dynamical model's inability to correctly represent and parameterize deep tropical cumulus convection and associated diabatic heating in numerical models, rather than the reaching of some intrinsic limit of predictability of the MJO.

However, statistical models are able to produce skillful forecasts out to relatively long lead time (Waliser et al. 2006). Waliser et al. (1999) used singular value decomposition (SVD) to forecast pentad maps of outgoing longwave radiation (OLR) anomalies from the previous two pentad maps and obtained useful forecast skill out to 4 pentads. Similarly, Lo and Hendon (2000) used the first two principal components of OLR and the first three principal components of streamfunction with a lagged regression model to give useful forecast skill out to 3–4 pentads during boreal winter. Jones et al. (2004) used a combined empirical orthogonal function (EOF) analysis and multiple regression to obtain useful forecast skill out to 5 pentads over the bulk of the tropics. In contrast to the other models mentioned and the majority of atmospheric forecasts that work iteratively (i.e., forecast one step into the future from known data, then forecast the next step from this result and past data), this model used a stepped forecast where a set of models was parameterized for each forecast step into the future. In other studies, Mo (2001) used singular spectrum analysis and maximum entropy methods, while Wheeler and Weickmann (2001) used wave theory to filter and forecast convectively coupled modes in the tropics. All these statistical model studies

obtained useful forecast skill out to approximately 4 pentads. Despite the similar performance of these models, we cannot be sure that they indicate the limit of inherent predictability we can expect from empirical and numerical models.

However, many of these forecast models could not be applied in real time as they relied on filtered input data. This filtering produced unwanted end effects, such that the beginning and, more importantly, the end sections of filtered time series are distorted, or even missing. Conversely, if no filtering is applied, the MJO signal may be lost among other weather or climate "noise." Wheeler and Hendon (2004) recently developed an MJO index based on the first two EOFs of equatorially averaged OLR and 850- and 200-hPa zonal wind. These EOFs were projected onto daily maps of OLR, from which the annual cycle and a component of the interannual variability had been subtracted. Hence, the necessity for time filtering was reduced and the resulting principal component time series could be calculated in real time. A seasonally varying lagged linear regression technique was then applied to produce real-time forecasts of the two MJO EOFs (http://www.bom. gov.au/bmrc/clfor/cfstaff/matw/maproom/RMM/).

In this paper, we build on the existing literature of the statistical prediction of the MJO. The new Huang–Hilbert transform method (Huang and Shen 2005) has recently been applied to meteorological data (Wu et al. 1999; Duffy 2005; Coughlin and Tung 2005). Here we apply the empirical mode decomposition technique from the Huang–Hilbert transform to the MJO to produce a real-time intraseasonal filtered index of the MJO, and then input this to a powerful nonlinear vector autoregressive moving average (VARMA) model to skillfully forecast the MJO out to 25–40 days.

The paper is essentially divided into two sections, covering first the real-time intraseasonal data filtering and second the statistical modeling. Section 2 describes the data used in this study. The EMD methodology and its use as a real-time intraseasonal filter is described in section 3, then the selection and application of a VARMA model to forecast the MJO is described in section 4. Conclusions are presented in section 5.

2. Data

The region of interest for this study was defined as the box bounded at 60° and 180°E and between 20°N and 20°S, representing the tropical regions of major MJO activity, the Indian Ocean, Indonesia, and the western Pacific Ocean. The data used in this study are the interpolated 2.5° longitude by 2.5° latitude gridded set of daily means of OLR (Liebmann and Smith 1996).
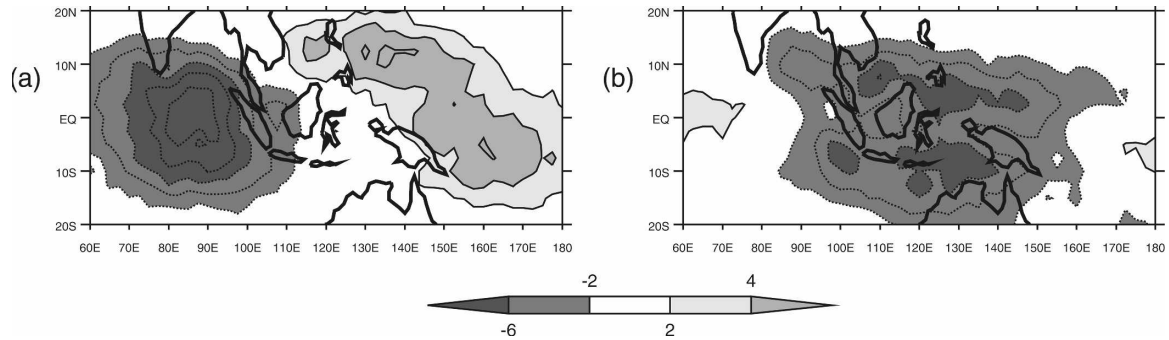
FIG. 1. (a) EOF1 and (b) EOF2 of "IMF2 filtered" OLR anomalies. Contour interval is 2 W m^{-2}. Negative contours are dashed, and the zero contour is suppressed. See legend for shading.

These satellite-measured data are a good proxy for rainfall in the tropics, and have been used many times as a measure of the state of the MJO (e.g., Lo and Hendon 2000; Waliser et al. 1999). The dataset runs for 27 yr, from 1 January 1979 to 31 December 2005. The annual cycle (calculated from the mean and the leading three annual harmonics) at each grid point was subtracted from the raw data to produce daily anomaly maps.

As a further introduction to the MJO, Fig. 1 shows the leading two EOFs of OLR after they have been passed through an empirical mode decomposition to isolate the intraseasonal variability. Further details are in section 3. EOF1 (Fig. 1a) extracts the well-known dipole phase of the MJO cycle when negative (positive) OLR anomalies, indicating a maximum (minimum) in precipitation, peak over the eastern Indian Ocean (western Pacific). EOF2 (Fig. 1b) then extracts the quadrature phase of the MJO cycle, with enhanced precipitation over Indonesia. These results are very similar to previous EOF analyses of conventional bandpass-filtered data (e.g., Matthews 2000) and confirm the effectiveness of the EMD process as a bandpass filter.

The days corresponding to the maxima in the principal component time series (PC1) of EOF1 were then found, and lagged composite maps were calculated from this list of dates. Equating negative (positive) OLR anomalies with positive (negative) precipitation anomalies, the MJO cycle as defined here has a positive OLR anomaly, that is, dry area, over the Indian Ocean at day −20 (Fig. 2a). This dry area then moves slowly eastward across Indonesia by day −15 (Fig. 2b), and a wet anomaly then appears behind it over the western Indian Ocean at day −10 (Fig. 2c). This "dipole" (wet–dry) phase of the MJO then also moves eastward through day −5 to 5 (Figs. 2d–f). Note that day 0 (Fig. 2e) corresponds to the maxima in the first principal component time series. When the dry anomaly reaches the date line it decays. Days 10–20 (Figs. 2g–i) cover the

second part of the cycle when a new dry anomaly appears over the western Indian Ocean and subsequently propagates eastward. The Hovmöller diagram in Fig. 2j summarizes this eastward propagation of the wet and dry anomalies.

3. Empirical mode decomposition

a. Rationale

The ideal input to a statistical forecast model for the MJO would be a dataset that only contained the MJO signal, with a minimum of noise due to other unrelated weather and climate systems. As the MJO is the dominant mode of variability on intraseasonal time scales, an intraseasonal (e.g., 30–70-day bandpass) filter is a simple method of doing this. However, such filters have undesirable end effects. They either lose the last section of a time series or otherwise alter the signal at the end of the time series, such that they cannot be used in real time, as the end of the time series contains the most recent information, which is necessary for a successful forecast.

Wheeler and Hendon (2004) have successfully made real-time forecasts of the MJO, which now form part of the operational output of the Australian Bureau of Meteorology. They first subtracted the annual cycle from their input data, and then removed some of the low-frequency variability, particularly that associated with ENSO. However, their input data still contained some undesirable high-frequency variability that was not associated with the MJO.

As an alternative to this approach, we utilize a recent development in data analysis, the Hilbert–Huang transform and specifically the aspect of it known as empirical mode decomposition (EMD) developed by Huang et al. (1998). This is an adaptive empirical method that has a continuously changing data-dependent basis function. It is used here as an efficient filter to extract the MJO signal, without introducing large unwanted end effects.
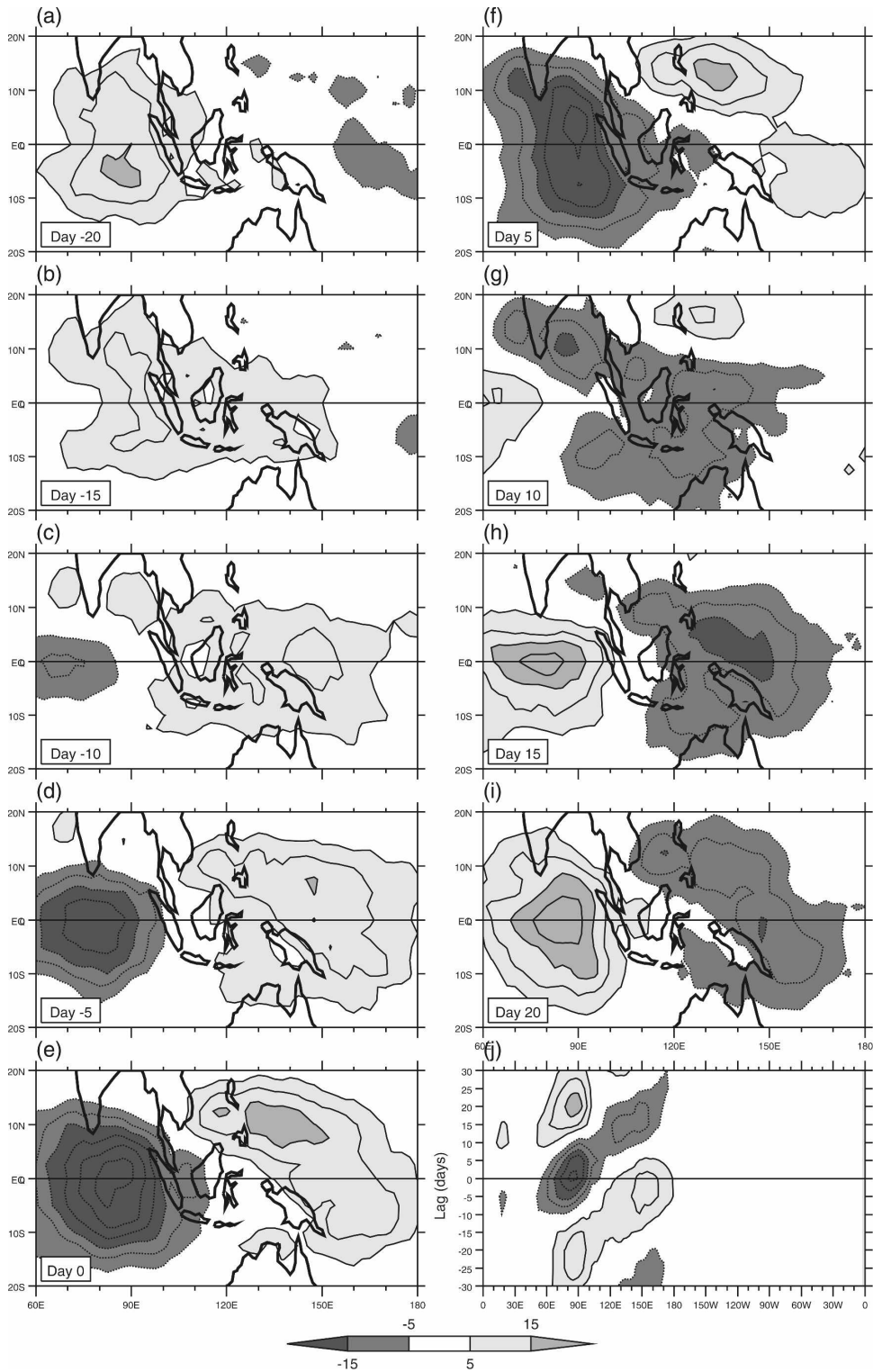
FIG. 2. Composite OLR anomaly maps, composited onto PC1 of IMF-filtered OLR, at day (a) −20, (b) −15, (c) −10, (d) −5, (e) 0, (f) 5, (g) 10, (h) 15, and (i) 20. Contour interval is 5 W m^{-2}. Negative contours are dashed, and the zero contour is omitted. (j) Hovmöller diagram of equatorial (averaged from 10°N to 10°S) OLR anomalies.

In contrast to some other popular methods, EMD does not *linearly* decompose the data into a set of modes; hence nonlinearities in the input data are preserved.

b. Standard methodology

The decomposition works on the assumption that the raw data consist of a number of simple, intrinsic modes of oscillation. These intrinsic mode functions (IMFs) have a simple empirical definition. If $x(t)$ is a time series of raw data at time t (Fig. 3a), then a cubic spline-fitted function can be found that passes through all the local maxima in this time series. The maxima are defined as those points that have a higher value than their immediate neighbors. Similarly, a second cubic spline-fitted function can be found that passes through all the minima (Fig. 3b). The mean function $m_1(t)$ of these two spline-fitted functions can then be calculated (Fig. 3c). The first IMF $e_1(t)$ is then equal to the difference between the raw time series and this mean function (Fig. 3d):

$$e_1(t) = x(t) - m_1(t). \tag{1}$$

The mean function is then recycled and becomes the raw data for the calculation of the second IMF (Fig. 3e). Hence, maximum and minimum spline-fitted functions are fitted to it (Fig. 3f), the mean function $m_2(t)$ of these two spline-fitted functions is calculated (Fig. 3g), and this mean function is then subtracted from the input time series to give the second IMF (Fig. 3h):

$$e_2(t) = m_1(t) - m_2(t). \tag{2}$$

A similar process is then carried out to calculate the third IMF (Figs. 3i–l):

$$e_3(t) = m_2(t) - m_3(t). \tag{3}$$

The remainder [$m_3(t)$, Fig. 3m] then once more becomes the input data for the next cycle, and

$$e_4(t) = m_3(t) - m_4(t), \tag{4}$$

and so on, until the remaining data are either a constant or a simple monotonic function with some additional insignificant white noise, which here typically happens after the sixth IMF. Note that, at any stage, the original time series can be reconstructed as the sum of the IMFs calculated so far, and the remainder (the last mean function). For example, after calculation of the first three IMFs,

$$x(t) = e_1(t) + e_2(t) + e_3(t) + m_3(t). \tag{5}$$

The IMF is a simple oscillatory mode that consists of the locally highest frequencies of the time series input to each decomposition, but its form is much more general than a normal oscillatory mode and can have an amplitude and frequency that varies continuously in time.

The large amplitude intraseasonal anomalies in the raw time series $x(t)$ of OLR at 0°, 80°E (Fig. 3a) can be seen to be part of eastward-propagating MJO events in a Hovmöller diagram of 20–200-day bandpass-filtered OLR anomalies (Fig. 3j). However, in the EMD analysis, these "MJO" anomalies can appear in any of the first three IMFs (Figs. 3d,h,l). For example, the positive intraseasonal OLR anomaly during March 1997 appears in IMF1 and IMF2, but the positive intraseasonal OLR anomaly in early December 1996 is largely accounted for by IMF3.

c. Adapted methodology

As the MJO is split between three IMFs, the basic EMD method is not yet an adequate intraseasonal filter. Some empirical adaptations are made to the EMD process, so it can be used practically. These are concerned with isolating the MJO signal in a single IMF with minimal additional noise, and also minimizing the degradation at the end of the data record that we refer to as the "end effect."

1) ISOLATION OF THE MJO IN A SINGLE IMF

Consider first the issue of isolating the MJO signal in a single IMF. Since the EMD method essentially selects the locally highest-frequency components from a time series to create each IMF, then our desired signal can be split between more than one IMF, as discussed in the previous section.

As the IMFs and final remainder (the last mean function) can be summed together to recreate the original signal, it is possible to add individual IMFs together to create new joint IMFs containing all of the desired signal. However, in practice such a joint IMF invariably includes unwanted high-frequency variability as well as the MJO signal, and is of little practical use in this particular scenario since the purpose of the decomposition is to provide a cleanly filtered MJO signal for a forecast application. So the primary objective must be to isolate the target oscillation in a single IMF.

This was partially achieved by making a slight change to the EMD methodology. After some empirical experimentation, the input daily mean time series was prefiltered by passing it through a 7-point running mean, then a 3-point running mean, equivalent to a single 9-point running mean with a (1, 2, 3, 3, 3, 3, 3, 2, 1) weighting (Fig. 4a). The running means were modified near the ends of the time series, to ensure that the input and filtered time series were the same length. For the 7-day running mean, the points 3 and 2 days from
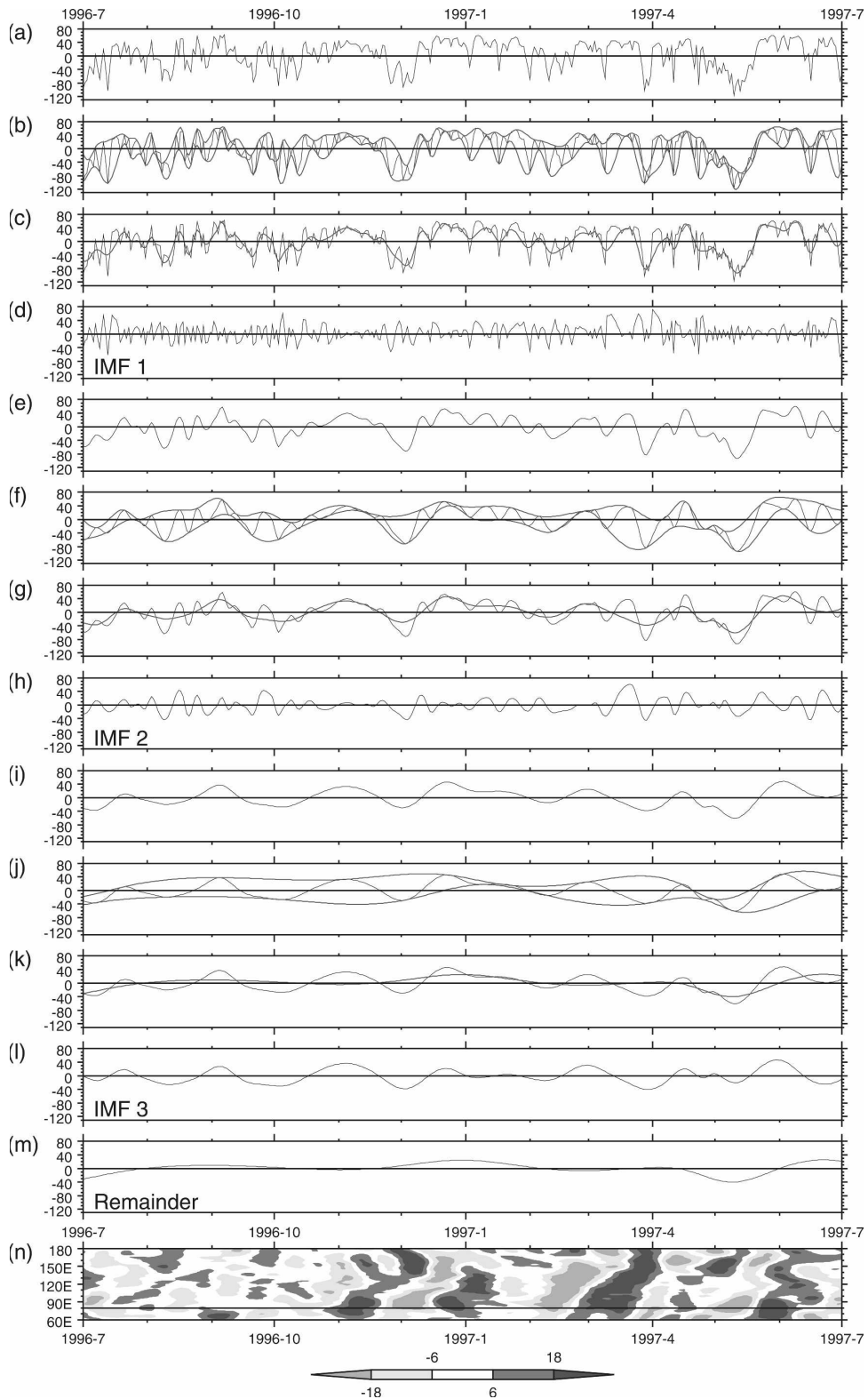
FIG. 3. The process of EMD, shown using a section of daily mean OLR time series (annual cycle subtracted) between 1 Jul 1996 and 1 Jul 1997 at 0°, 80°E. The decomposition is actually performed from 1 Jan 1996 to 31 Dec 1997, so no end point effects are shown. (a) The raw data. (b) The same data with maxima and minima spline-fitted functions. (c) The same data with the mean function. (d) IMF1. (e) The mean function/new raw dataset. (f) Maxima and minima spline-fitted functions. (g) The raw data with the mean function. (h) IMF2. (i)–(l) The process for IMF3. (m) Remaining signal. (n) Hovmöller diagram (60°–180°E) of 20–200-day filtered OLR anomalies at 0°. Shading is shown by the legend. The thick horizontal line at 80°E indicates the longitude of the raw time series in (a).
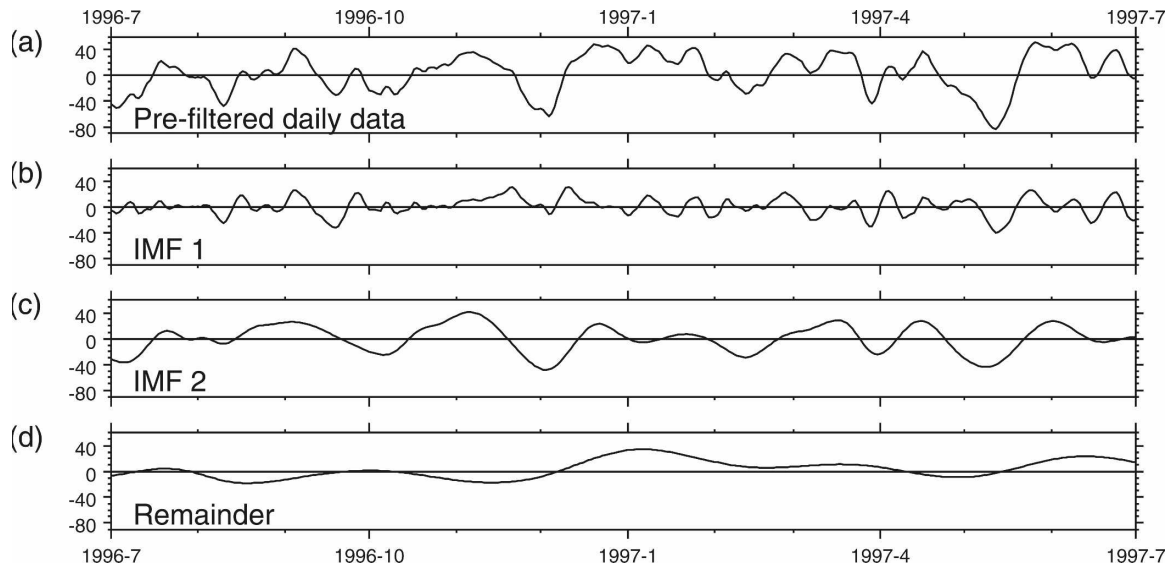
FIG. 4. (a) The daily input time series of Fig. 3a passed through a 7-point running mean and then a 3-point running mean, (b) IMF1 and (c) IMF2 of the time series in (a), and (d) remainder.

the ends were passed through 5- and 3-day running means, respectively, while the end points were just passed as they were (i.e., a 1-point running mean). Similarly, for the 3-day running mean, the end points were passed as they were. If this adaptation is made to the basic EMD methodology, then most of the high-frequency variability is contained in IMF1 (Fig. 4b), and most of the MJO variability is accounted for by IMF2 (Fig. 4c), while the remainder (Fig. 4d) accounts for the low-frequency variability.

To illustrate the effectiveness of the EMD method as an intraseasonal filter, space–time power spectra (e.g., Wheeler and Hendon 2004) of 30–70-day Lanczos bandpass-filtered OLR data (Fig. 5a) and IMF2 of the intraseasonally EMD-filtered OLR data (Fig. 5b) are compared. The spectra were calculated from a 1000-day-long section (1 January 1980 to 26 September 1982). The data were averaged from 5°S to 5°N, then passed through a Fourier transform in longitude and time. The modified EMD process has efficiently extracted the MJO signal. Its space–time power spectrum (Fig. 5b) is very similar to that of the conventional 30–70-day filtered data (Fig. 5a). Nearly all the power is concentrated in the 30–70-day (0.014–0.033 cycles per day) band in eastward-propagating zonal wavenumbers 1–4, consistent with the MJO (Salby and Hendon 1994). There is very little power outside this frequency range or in the westward-propagating waves.

At this point the data were split into two sets: a 18.5-yr training dataset from 1 January 1979 to 30 June 1997, and an independent 8.5-yr validation dataset from 1 July 1997 to 31 December 2005. All the development of
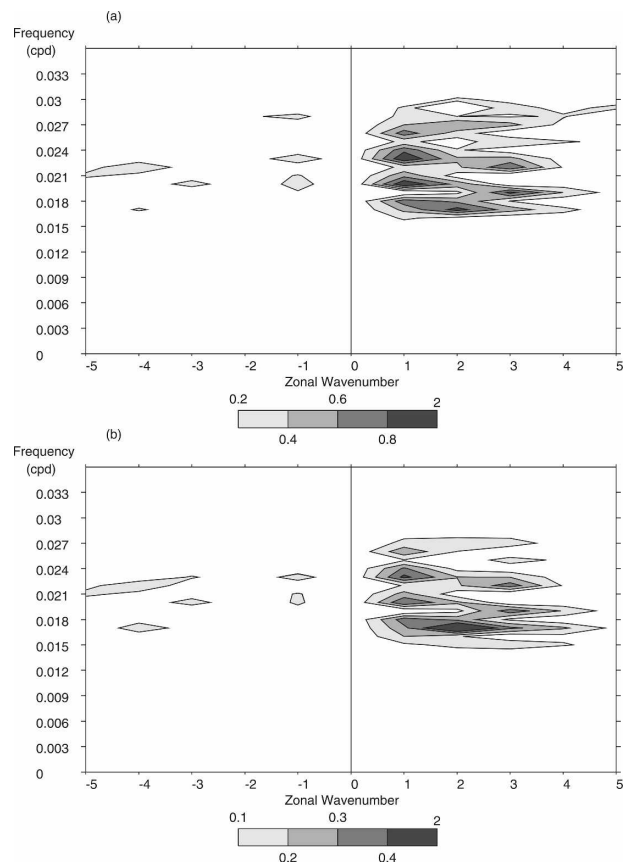

FIG. 5. 2D space–time power spectra of equatorial (averaged from 5°S to 5°N) (a) 30–70-day Lanczos bandpass-filtered OLR data, (b) EMD-filtered IMF2 OLR data. Positive (negative) zonal wavenumbers correspond to eastward- (westward) propagating waves. Frequency axes units are cycles per day.
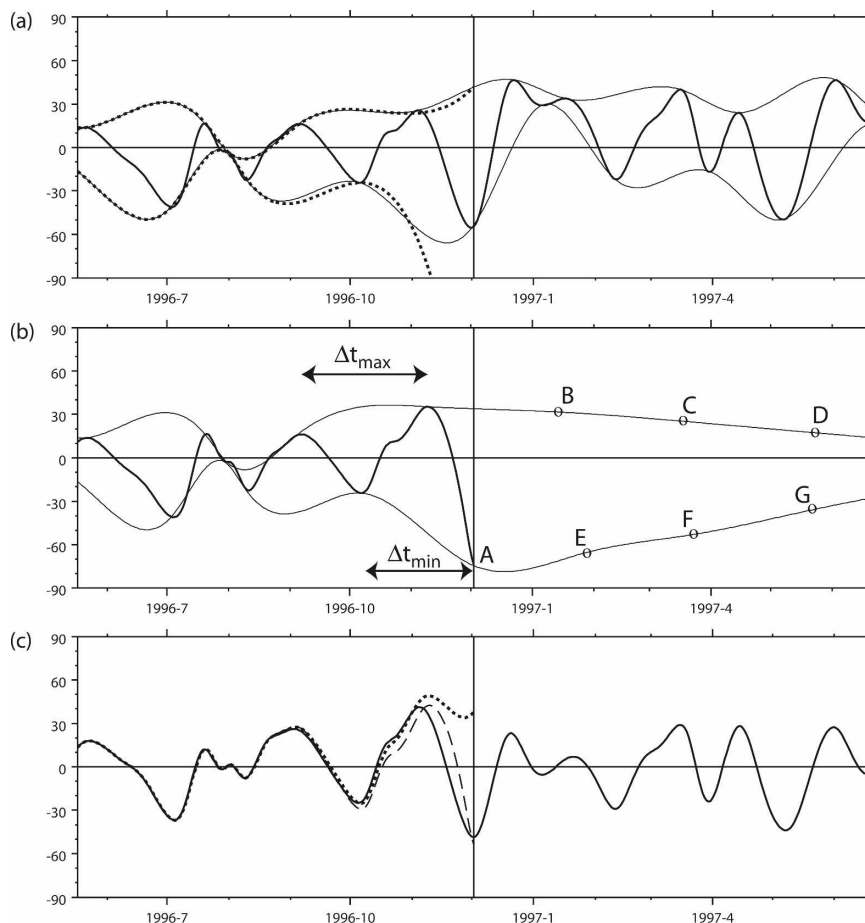
FIG. 6. The end effect of the EMD process. (a) Section of an input time series [the remainder after IMF1 has been subtracted from the prefiltered (7-day, then 3-day, running mean) daily OLR anomalies at 0°, 80°E; thick line], together with the maximum and minimum spline-fitted functions. The dotted lines show the fitted maximum and minimum spline-fitted functions when the input time series is truncated at 2 Dec 1996, shown by the thick vertical line. (b) Extension of the time series past its end point at 2 Dec 1996. See text for details. (c) IMFs of the input time series: true IMF of the full input time series (solid line), IMF of the input time series when it is truncated at 2 Dec 1996 (dotted line), IMF of the input time series when it is truncated at 2 Dec 1996, and additional maxima and minima have been created as in (b) (dashed line).

the EMD and forecasting methodology was done with the training dataset, while the validation dataset was reserved to validate the forecasts on. The details of the running means were determined pragmatically based on the skill score from the training dataset, when the IMFs were used later in conjunction with a forecast model (section 4).

2) REDUCTION OF THE END EFFECT

The EMD method also has a significant end effect at the end of the data record, even when the running mean adaption is used. This is illustrated in Fig. 6 with reference to extracting the MJO signal, by taking as an ex-

ample input the data after one cycle of EMD has been carried out. Hence, the input time series is $m_1(t)$, the remainder after IMF1 has been subtracted from the prefiltered (7-point, then 3-point, running mean) time series of daily OLR anomalies at 0°, 80°E. The IMF of this example time series will then be IMF2 of the original time series.

The thick solid line in Fig. 6a shows a section within this long input time series. High-amplitude intraseasonal (MJO) variability can clearly be seen. Maximum and minimum spline-fitted functions are then fitted as part of the EMD process (solid lines in Fig. 6a). However, if the input time series is truncated at 2 December 1996, as shown by the thick vertical line in Fig. 6a, then

the last maximum is at 9 November 1996 and the last minimum is at 7 October 1996. The cubic spline-fitted functions fitted to the maxima and minima (dotted lines in Fig. 6a) then differ significantly from hose previously fitted to the time series when information was available after 2 December 1996. This end effect can then significantly affect the value of the mean function and the IMFs.

Again, we employ a simple empirical adaptation to the basic EMD methodology to reduce the end effect. As cubic splines are sensitive to the end point over which they are fitted, we can attempt to extend the record of maxima and minima over which the splines are fitted. From a practical standpoint, if we were able to make perfect forecasts of the next few maxima and minima we would have a zero end effect. There are several possible methods that have been proposed (Huang and Shen 2005) for defining these additional extrema which can generally be grouped into three categories. Creating the new additional extrema can be based around the mean climatology, autoregressive forecasts, or linear methods. However, there is currently no consensus as to the best procedure, and research on the topic is ongoing. The choice of procedure by which the end effect is minimized is one that the users should decide, ultimately based on their intuitive knowledge of the system.

The procedure adopted in this study is as follows. If the gradient between the penultimate point and the end point is positive and the end point is also positive, then the end point is deemed to be a maximum. If this gradient is negative and the end point is also negative, then the end point is deemed to be a minimum (label A in Fig. 6b). Next, the time separation between the last two maxima, Δt_{max}, is calculated. Similarly, the time separation between the last two minima, Δt_{min}, is also calculated. For the example shown here, $\Delta t_{max} = 64$ days, and $\Delta t_{min} = 56$ days, that is, within the broad range of the period of the MJO.

The input time series is then extended beyond its end point by adding another maximum, at a time interval of Δt_{max} after the last maximum in the input time series, and with an amplitude of 0.9 of the last maximum (label B in Fig. 6b). A second additional maximum is then added at Δt_{max} after, and with an amplitude of 0.8 of, the first additional maximum (label C in Fig. 6b). Similarly, a third additional maximum is added with an amplitude of 0.7 of the second additional maximum (label D in Fig. 6b), and so on. Additional minima are then also added beyond the end point of the time series by the same method (labels E, F, and G in Fig. 6b). A cubic spline-fitted function is then fitted to all the maxima, including the new maxima beyond the end

point, and similarly another cubic spline-fitted function is fitted to all the minima (Fig. 6b). The mean function and IMF are then calculated in the usual way.

The reduction in the end effect can be clearly seen by comparing the three IMF time series in Fig. 6c. The solid line shows the "true" IMF calculated from the full example input time series. By definition, this has no end effect. The dashed line shows the IMF calculated when the input time series is truncated at 2 December 1996. The large differences in the last 30 days between this and the true IMF are the errors due to the end effect. Finally, the dotted line shows the IMF calculated when the input time series is truncated at 2 December 1996, but when additional maxima and minima have been created as in Fig. 6b. This IMF follows the true IMF much more closely, showing the substantial reduction in the error due to the end effect.

This reduction in the error due to the end effect was quantified by taking 1241 overlapping long (300 day) segments from a full 6940-day input time series (daily OLR anomalies at 0°, 80°E from the training dataset). Each successive 300-day segment was advanced from the previous one by 5 days. For each segment, prefiltering was applied, the original EMD methodology (i.e., with no additional maxima or minima) was carried out and IMF1 and IMF2 time series were calculated. In addition, an EMD analysis was carried out on the full 6940-day time series, and the IMF1 and IMF2 time series of this were calculated.

As a measure of the end effect, the correlation coefficient between the last value of the IMF2 of each 300-day segment and the corresponding IMF2 value from the full 6940-day time series were calculated, that is, the correlation coefficient between 1241 pairs of data. The correlation was effectively zero (0.03). If there was no end effect, its value would be 1. Similarly, the correlation coefficient between the 300-day segment IMF2 values 15 days before the end and the corresponding value from the full IMF2 was only 0.29. The dotted line in Fig. 7 shows the correlation coefficient as a function of time before the end of the segment. It is approximately 1 before about 50 days before the end, indicating that the end effect is negligible here, but the correlation then decreases steadily down to effectively zero (0.03) at the end, indicating the increasing size of the end effect. Given that the useful information for MJO forecasting is likely to be in the last 20 days or so of the input time series, these low correlations imply that the end effect will severely limit the skill of forecasts if it is left uncorrected.

However, the solid line in Fig. 7 shows the corresponding correlations when the adapted EMD methodology with the extra maxima and minima was used, as
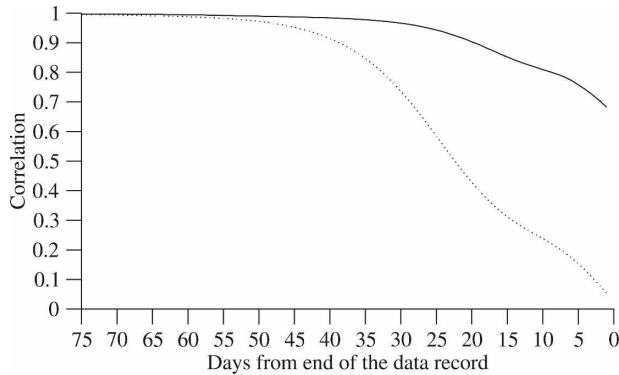
FIG. 7. Correlation coefficient as a function of time before the end of the time series, between IMF2 of a segment of the input time series (prefiltered daily OLR anomalies at 0°, 80°E), and IMF2 of the full time series. Original EMD methodology (dotted line) and adapted EMD methodology (solid line).

described earlier in this section. The correlations are much higher, with a value of 0.66 at the end of the segments, rising to near 1 (no end effect) at about 25 days before the end of the segments. Hence, the adaptations to the EMD methodology have significantly reduced the end effect.

Although significant effort has been made to reduce spurious deviations at the end of the data record, a consequence of adding the additional extrema is that the amplitude of the MJO signal in IMF2 tends to be slightly damped over the last 5–10 days. Hence, the final step in the preparation of the time series to input to the forecast model is to compensate for this effect by applying simple amplitude boosting factors. At this point the daily data are averaged into nonoverlapping pentad (5 day) means. The amplitude boosting factors are applied to the last two pentads of the time series. The amplitude boosting factors are 1.21 and 1.14 for the end and penultimate pentads, respectively. The values of the amplitude boosting parameters were optimized in a pragmatic way to give the highest mean anomaly correlations over the first six forecast pentads of "real time" forecasts (4) performed on the training dataset. Note that, although the amplitude boosting factors increased the anomaly correlations of the forecasts, they had a slightly detrimental effect on the root-mean-square errors, increasing them by 13%. This concludes the real-time intraseasonal filtering section of the paper.

4. Statistical forecast modeling

a. Vector autoregressive moving average model

The state of the MJO at any particular time can be reasonably represented by the amplitudes of the prin-

cipal component time series of the leading two EOFs of intraseasonally filtered OLR (Fig. 1). The MJO forecast problem is then reduced to forecasting the amplitudes of these two time series, PC1 and PC2. At time t, we define the bivariate vector

$$\mathbf{x}_t = (\text{PC1}_t, \text{PC2}_t). \tag{6}$$

Previously, Maharaj and Wheeler (2005) have employed a vector autoregressive (VAR) model, and Jones et al. (2004) and Wheeler and Hendon (2004) have developed multiple linear regression models to forecast the MJO. Here, we introduce a more powerful general class of statistical models: VARMA models (Box and Jenkins 1970). We apply this technique to model the bivariate MJO time series as a VARMA process with autoregressive order P and moving average order Q,

$$\mathbf{x}_t = \sum_{p=1}^{P} \phi_p \mathbf{x}_{t-p} - \sum_{q=1}^{Q} \theta_q \boldsymbol{\epsilon}_{t-q} + \boldsymbol{\epsilon}_t, \tag{7}$$

where ϕ_p and θ_q are 2×2 matrices containing VAR and vector moving average (VMA) parameters, respectively, and $\boldsymbol{\epsilon}$ is a two-dimensional vector of white noise processes. For the purpose of parameterization, $\boldsymbol{\epsilon}$ are the residuals, the difference between the predicted and actual values of $\mathbf{x}$.

The autoregressive and moving average parameters are found by a maximum likelihood method. The likelihood function is calculated using a Kalman filter algorithm (Shea 1987), and then a quasi-Newton algorithm to find the maximum of the log-likelihood function (Gill and Murray 1972). A reparameterization technique (Ansley and Kohn 1986) is used to enforce stationarity (a necessary condition of the autoregressive process) and invertibility (a necessary condition of the moving average process). An important condition on the maximum likelihood estimates being equal to their true values is that the estimates of the residual series are white noise; that is, they are uncorrelated, with zero mean and constant variance. This condition is used to help select the order of the VARMA process.

b. Model validation

To summarize the analysis so far, a total of 27 yr of OLR data were available, from 1 January 1979 to 31 December 2005. At each grid point, the daily OLR data had the annual cycle subtracted and were passed through a 7-point, then a 3-point running mean. The data were then split into the 18.5-yr training dataset from 1 January 1979 to 30 June 1997, and the 8.5-yr validation dataset from 1 July 1997 to 1 December 2005.

At each grid point in the training dataset, an EMD analysis was carried out, and IMF2 was selected, as it contains the MJO signal. The IMF2 time series at each grid point were then averaged into pentad means, and combined to make 18.5 yr of pentad maps of IMF2. These maps are essentially maps of intraseasonally filtered OLR anomalies, similar to conventional maps of bandpass-filtered data. At this point, the first year and last six months of the training dataset were discarded because of the end effects from the EMD analysis.

The remaining 17 yr (1980–96) of pentad maps of IMF2 were then subjected to an EOF analysis. Consistent with many other previous MJO studies (e.g., Hendon and Salby 1994; Matthews 2000), the two leading EOFs describe the MJO structure (Fig. 1). These were described in section 1. The spatial structures of EOF1 and EOF2 were then projected onto the pentad maps of IMF2 in the training dataset to obtain the PC1 and PC2 time series, which then form the bivariate vector time series x_t that is the input to the VARMA model. Best estimates of the autoregressive ϕ and moving average θ parameters were made using the maximum likelihood method described in section 4a, using the input x_t from the training dataset.

The same EMD analysis was then carried out on the 8.5-yr validation dataset, and the first six months and last year were discarded. The spatial structures of EOF1 and EOF2 previously calculated from the training dataset were then projected onto the pentad maps of IMF2 in the validation dataset to obtain a 7-yr (1998–2004) validation dataset of x_t.

"Real time" model predictions (hindcasts) were then made within the independent dataset using a VARMA model. For a hindcast made on a given pentad, *only data up to that date were used*. The running mean and adapted EMD methodology with extra maxima and minima were used to calculate the IMF2 time series at each grid point up to that date; then the PC1 and PC2 amplitudes (of the training dataset EOF1 and EOF2) were calculated from the maps of IMF2; then the VARMA model was used to predict values of PC1 and PC2 for the next few pentads. Then, using the training dataset, a set of monthly pairs of regression maps (one pair for each month: January, February, . . . , December) were created from a regression of the PC1 and PC2 time series to the gridded IMF2 time series. Finally the forecast values of PC1 and PC2 were projected onto their appropriate regression maps and the two maps summed to create the forecast map. Verification maps were similarly produced by projecting the respective PC1 and PC2 amplitudes, calculated from the IMF2-based validation dataset, onto the regression maps. A

total of $N = 511$ hindcasts were made over the 7-yr period of the independent dataset. The model hindcasts were then verified on the independent 7-yr validation dataset. All statistics and measures of the model performance were calculated from predictions made on the validation dataset.

c. Choice of model order

The next task is to choose the order of the VARMA model to be used, that is, the values of the integers P and Q in Eq. (7). As the order of the model is increased and more terms are included the model performance may be expected to increase, until at some point the extra parameters become unnecessary and the model performance plateaus or even decreases as the model begins to overfit to noise. Striking this balance between model performance and parsimony (i.e., not including more parameters than is necessary) is the key to the choice of model order.

There are a number of statistics that can be calculated to identify the best choice of model order. Each has their own advantages and disadvantages, but all are based on an examination of the residuals ϵ. The most commonly used statistics are variations of the information criteria (e.g., the Akaike information criterion). However, these can be very restrictive and do not always allow users to apply their knowledge of the systems' behavior (i.e., MJO quasi periodicity) and subsequently include higher-order terms. For this reason a more forecast skill–orientated approach to model selection was adopted.

An initial investigation of lagged autocorrelations and partial autocorrelations indicated that a model of autoregressive order 4–5 and a moving average order of 1–2 was probably required. This was backed up by an analysis of the variance of the residuals for a pure autoregressive model, with no moving average process ($Q = 0$). As the autoregressive order P is increased, the variance of the residuals ϵ decreased (Fig. 8), indicating a more accurate model. However, there is little improvement beyond an order of 4 or 5, implying that higher-order terms are unnecessary.

To make the selection, a measure of model skill called the anomaly correlation was used. For a given forecast lead time, a pattern correlation coefficient was calculated between the grid points of the forecast map and the grid points of the observation map. For a number of forecasts over the training dataset these anomaly correlation coefficients can be sorted and the median and upper and lower quartiles found.

The final choice of model order was made by calculating the median, lower quartile, and upper quartile
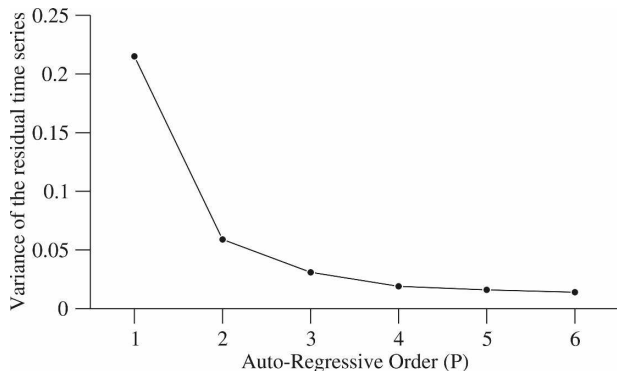
FIG. 8. Sigma values as a function of autoregressive order P of bivariate VAR models of PC1 and PC2 of the leading two EOFs of real-time IMF2.
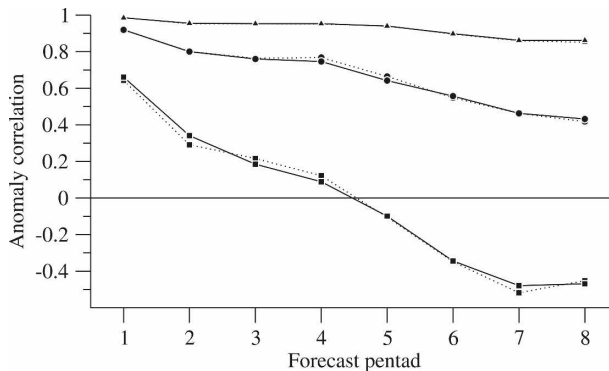

FIG. 9. The median (circles), and upper (diamonds) and lower (squares) quartiles of the distribution of anomaly correlations between forecasts and validation maps of real-time data as a function of forecast lead time. Solid lines: forecast against IMF2-based validation dataset. Dotted lines: forecast against 30–70-day-based validation dataset.

anomaly correlations between nonreal time (no end effect) model forecasts and validation observations over the first 6 forecast pentads, using the training dataset. These were compared with the Li–McLeod Portmanteau statistic (Li and McLeod 1981) for each model:

$$Q^*_{LM} = \frac{m(m+1)}{2N} + N\sum_{k=1}^{m} r_k^2, \qquad (8)$$

where r_k is the correlation coefficient between residuals (ϵ) at lag k, $N = 1241$ is the number of samples, and m is the maximum lag considered ($m = 20$ was chosen as a sufficiently high value). The Portmanteau statistic essentially tests the null hypothesis that the residuals are independent up to lag m. This independence of the residuals (i.e., they are uncorrelated) was a condition on the estimated model parameters being equal to their true values.

The simplest VARMA model has order $(1, 0)$, a pure autoregressive model of first order. This model has a low "combined" anomaly correlation (mean of the median anomaly correlations over the first 6 forecast pentads) of 0.68, indicating it has low forecast skill, and a high Li–McLeod Portmanteau statistic of 4600, indicating that the residuals (difference between model forecast and validation observations) are not independent and therefore there is still information in the residuals that this particular VARMA $(1, 0)$ model has not captured. As the orders of the autoregressive and moving average parts of the VARMA model are increased, the combined anomaly correlations consistently increase, indicating higher forecast skill, and the Portmanteau statistic decreases, indicating that there is less information left in the residuals. However, when the combined model order reaches 6—that is, VARMA $(6, 0)$, VARMA $(5, 1)$, VARMA $(4, 2)$, etc.—the Portman-

teau statistic continues to decrease but the anomaly correlations reach a plateau or decrease, indicating that the model is overfitting.

From this analysis we identify a VARMA $(5, 1)$ model (combined anomaly correlation 0.93 and Portmanteau statistic 160) as the model with the optimum balance between forecast skill and parsimony.

d. Consistency of forecasts

The frequency distribution of the anomaly correlations of the $N = 511$ hindcasts was then examined to ascertain the consistency of the forecasts. The median and the upper and lower quartiles of this distribution are shown as a function of forecast lead time by the solid lines in Fig. 9. The median anomaly correlation is 0.92 for the 1 pentad forecast. It then decreases fairly slowly, to 0.64 at 5 pentads and 0.43 at 8 pentads. The upper quartile anomaly correlation is much higher, and is still 0.86 even at a lead time of 8 pentads. By definition, a quarter of the forecasts performed even better than this at that lead time. Conversely, the anomaly correlation of the lower quartile was only 0.66 at 1 pentad, decreasing rapidly to near zero (0.19) at only 3 pentads.

At this point, we present further justification of the realism of assuming the MJO signal is contained in IMF2. A conventional 30–70-day Lanczos bandpass-filtered dataset was used as an alternative validation dataset. The PC1 and PC2 time series were again calculated from the IMF2-based validation dataset, but this time projected onto regression maps of the regression between the PC1 and PC2 time series and maps of the 30–70-day filtered data. The real-time forecasts
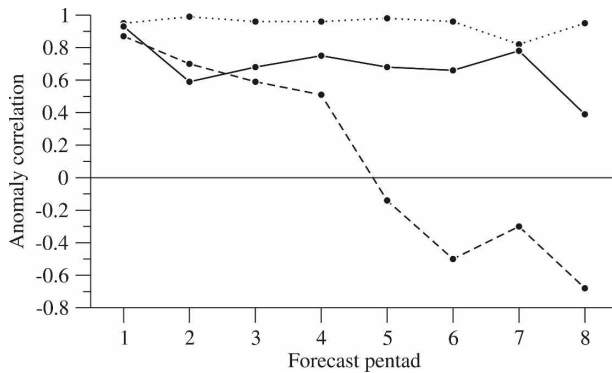
FIG. 10. Anomaly correlations as a function of forecast lead time, for three individual forecasts, beginning on 18 Dec 2002 (solid line), 11 Jul 2004 (dotted line), and 22 Jan 2004 (dashed line).

were then verified against this alternative validation dataset (dashed lines in Fig. 9). The median and the upper and lower quartiles of the anomaly correlations between the real-time hindcasts and the 30–70-day-based validation dataset are very similar to those with the IMF2-based validation dataset (solid lines in Fig. 9), confirming the effectiveness of the EMD method in extracting the MJO signal.

e. Example forecasts

While anomaly correlations are useful as an overall summary of forecast skill, they are not very intuitive. Here, we present spatial forecast maps of OLR anomalies from three individual example forecasts.

The first example forecast was made from 18 December 2002. Its anomaly correlation as a function of lead time is shown by the solid line in Fig. 10. By comparison with Fig. 9, it can be seen that this forecast is representative of a forecast with median skill. The forecast maps for this particular forecast are in the left-hand panels in Fig. 11, together with the verification maps in the right-hand panels. The high anomaly correlation of 0.93 at forecast pentad 1 is consistent with the high degree of agreement between the forecast and verification spatial maps in Figs. 11a,i. As the forecast lead time increases, the anomaly correlation tends to decrease, and the agreement between the forecast and verification maps also decreases. However, this decline in forecast skill is not monotonic for an individual forecast. In this case it increases slightly at pentad 4. There is still clearly considerable skill at pentad 6 in this forecast, as shown by the strong agreement between the forecast and verification maps (Figs. 11f,n) and summarized by the anomaly correlation of 0.66. At pentad 7 the anomaly correlation is high but the amplitudes of the forecast

and verification anomalies are very different. The model appears to be quite successful in predicting the phase of the MJO, that is, the locations of the OLR anomalies, but is not so skillful at predicting the amplitudes. In this example, the amplitude of the forecast is larger than the amplitude of the verification anomalies. Finally, at pentad 8, the model skill has decreased to the point where it is no longer useful (Figs. 11h,p).

It should be noted that the close match between the forecast and verification maps is not quite as impressive as it first appears. Even though there are over 800 grid points in each map, there are only effectively 2 degrees of freedom, from the amplitudes of PC1 and PC2. However, the forecasts can be presented in a more accessible way by the use of these maps, rather than by the use of PC time series.

The second example forecast was made from 11 July 2004. It has high anomaly correlations above 0.82 throughout the whole 8 pentad forecast (dotted line in Fig. 10) and is representative of a forecast from the upper quartile. The forecast maps of OLR (Fig. 12) closely match the verifications. This is an example of an MJO event that was accurately forecast over almost one whole period.

The third example forecast, made from 22 January 2004, is representative of a forecast from the lower quartile (dashed line in Fig. 10). Its anomaly correlation falls to near zero quickly, by pentad 5. This can be seen in the forecast map which predicts a region of enhanced convection over Indonesia (Fig. 13e), compared to the verification map that shows weakly reduced convection over the Indian Ocean and weakly enhanced convection over the western Pacific (Fig. 13m). As the forecast progresses, the forecast MJO gets further "behind" the observed event until, at pentad 8, the forecast and observed MJO are completely out of phase (Figs. 13h,p), with a strongly negative anomaly of −0.68. Clearly, this particular MJO forecast has little skill.

Hence, the overall skill of the model is very promising, with a median anomaly correlation over 0.6 at a forecast time of 5 pentads. However, there is considerable spread in the skill of individual forecasts, with the upper quartile showing very high skill (anomaly correlations over 0.85) out to 8 pentads, but the lower quartile having low skill beyond 1 pentad. It would be of use to attach some confidence to the skill of a forecast when it was made. Given that the only inputs to the statistical model are the amplitudes of PC1 and PC2, it is likely that the model will perform poorly when these amplitudes are low. At these times the next MJO event may emerge spontaneously and may not depend on the amplitude and timing of the previous cycle, or the precursor MJO signal may be weak and lost in noise. Also, as
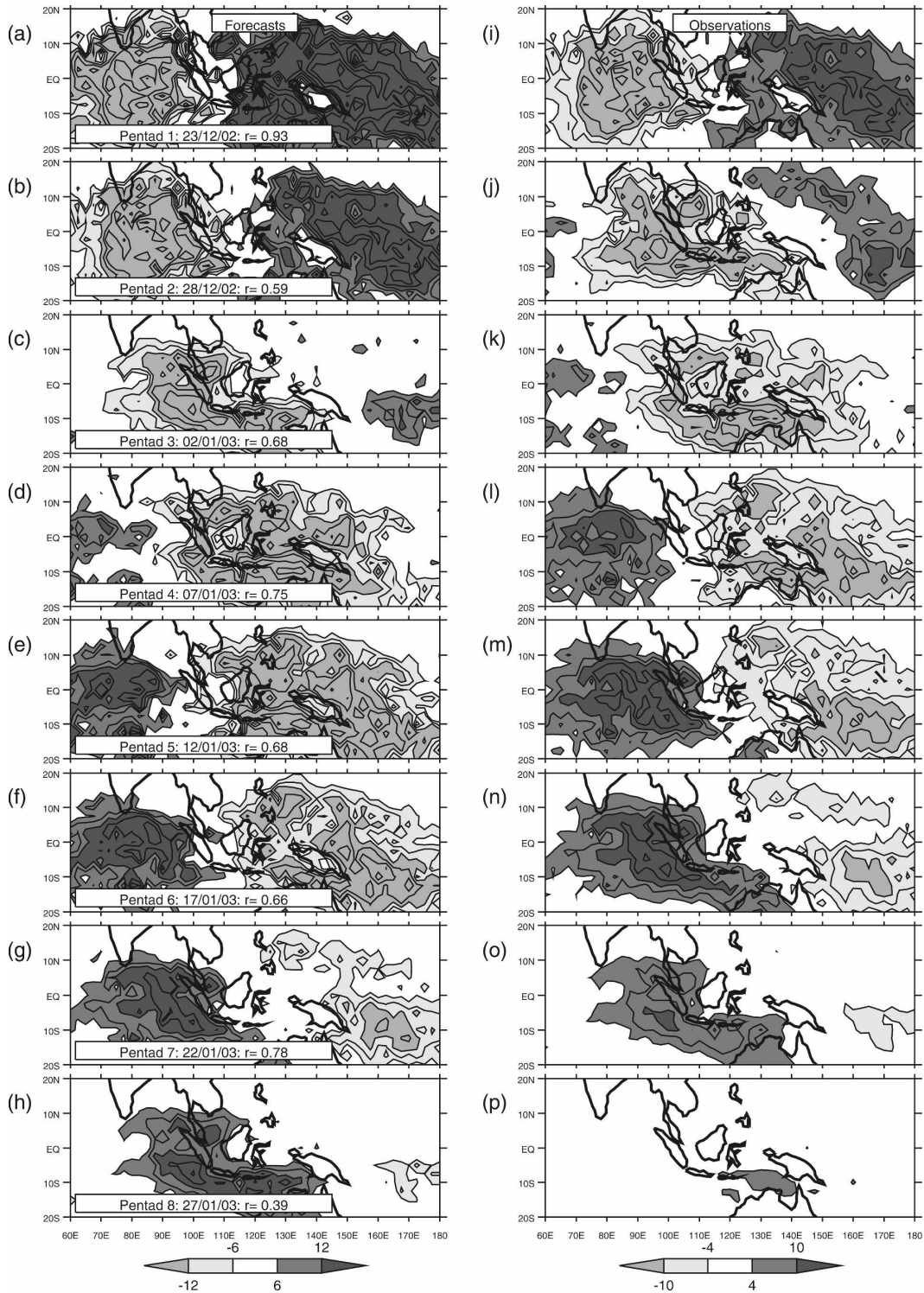
FIG. 11. Maps of OLR anomalies created by projecting the amplitudes of PC1 and PC2 onto their respective regression maps: a forecast typical of the median skill, made from 18 Dec 2002 with lead times of (a) 1, (b) 2, (c) 3, (d) 4, (e) 5, (f) 6, (g) 7, and (h) 8 pentads. (i)–(p) Corresponding validation maps. Contour interval is 3 W m^{-2}, the zero contour is suppressed, and shading is shown by the legend.
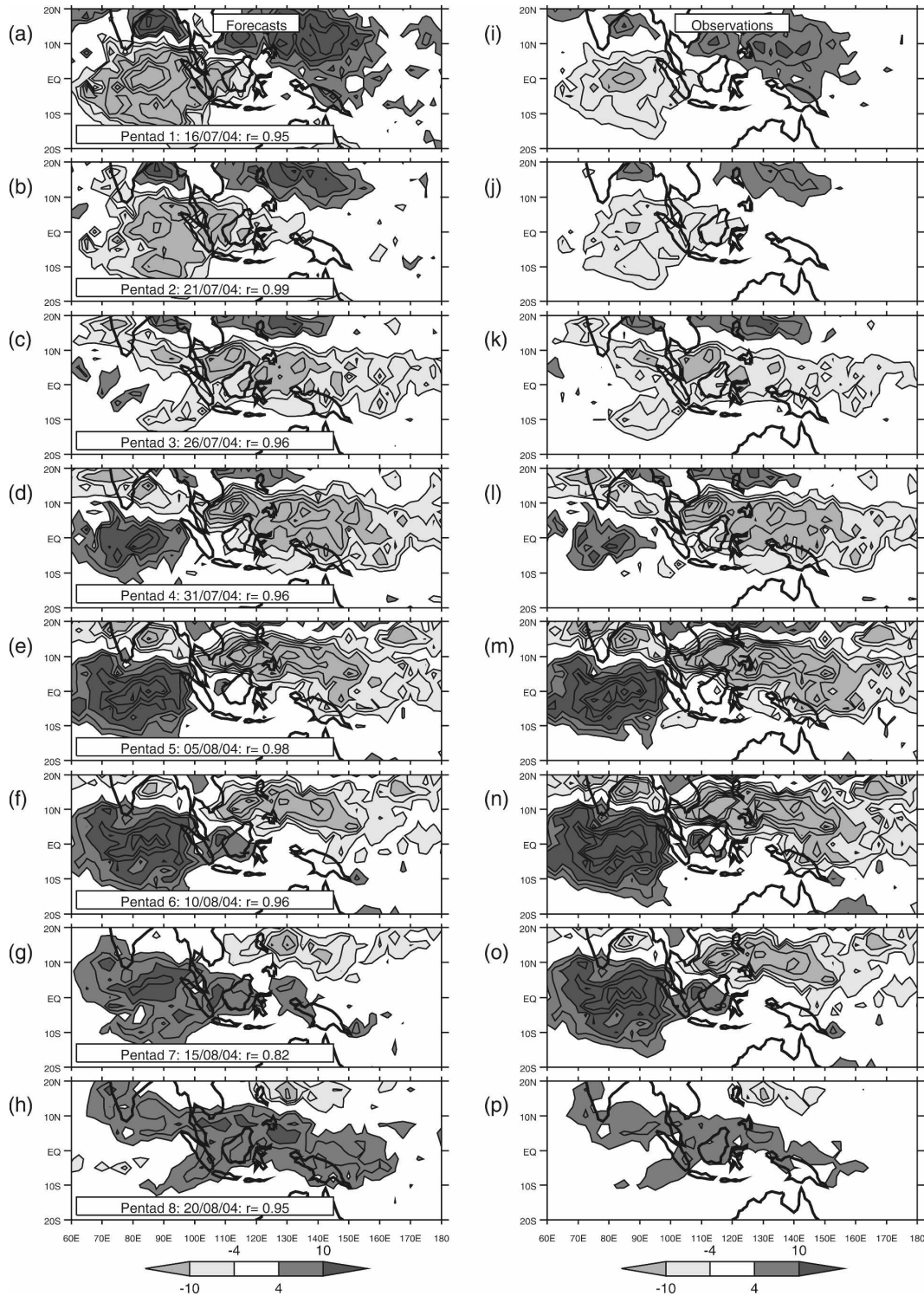
FIG. 12. As in Fig. 11 but for a forecast typical of the upper quartile skill, made from 11 Jul 2004.

the final amplitudes of the IMFs are small, the end effect will be largest at this time.

Hence, the MJO forecasts were stratified according to the initial magnitudes of PC1 and PC2. The solid lines in Fig. 14 show the median and lower quartile anomaly correlations as a function of forecast lead time, when all (100%) of forecasts are retained. These are replications of those in Fig. 9. We then define the
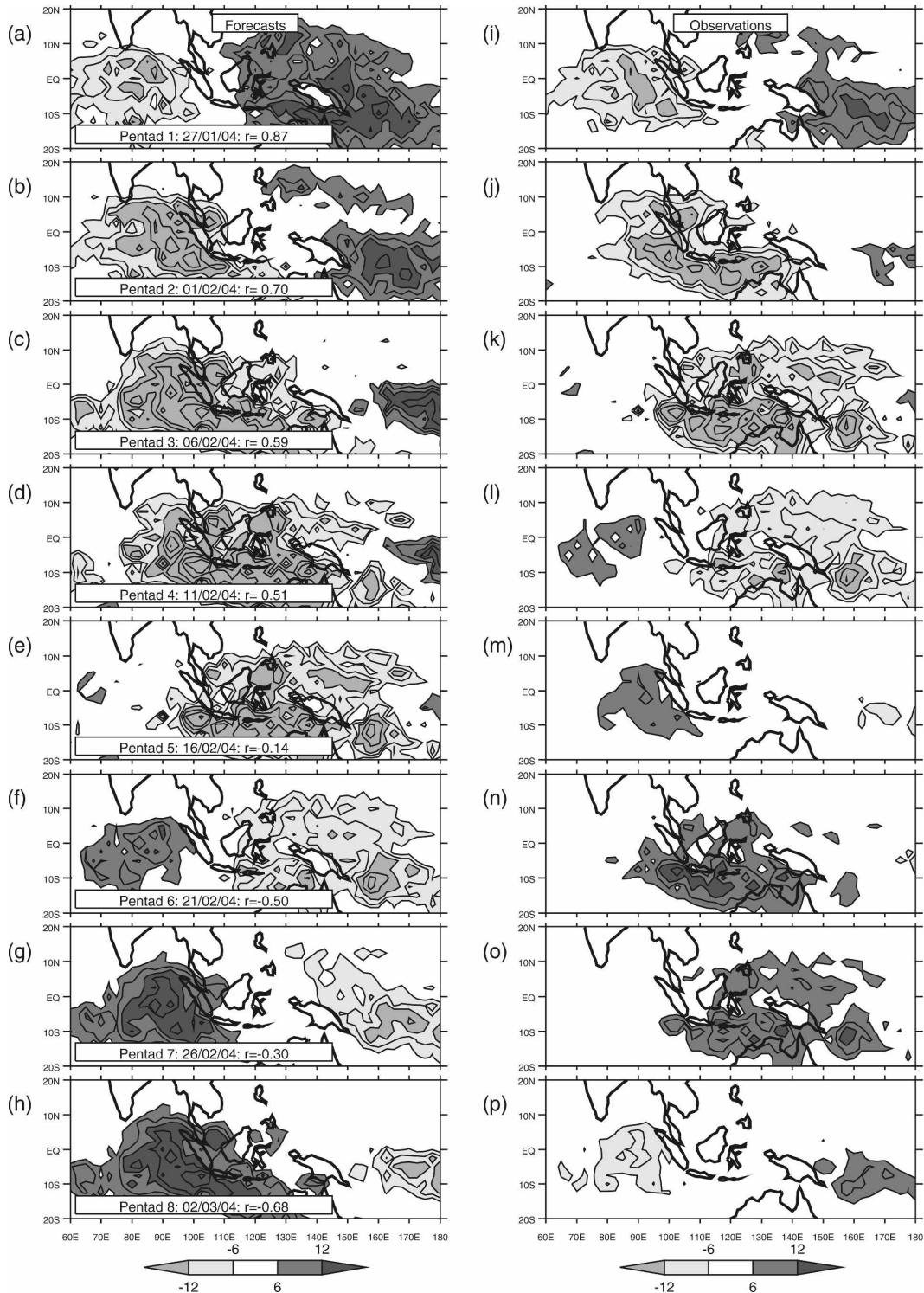
FIG. 13. As in Fig. 11 but for a forecast typical of the lower quartile skill, made from 22 Jan 2004.

initial strength of the MJO as the sum of magnitudes of the two PCs, that is, $|PC1| + |PC2|$, on the pentad the forecast is made. Strong initial MJOs are defined as those above the median value of this measure. If only those forecasts with a strong initial MJO are retained, the anomaly correlations rise considerably (dotted lines in Fig. 14). Hence, a strong MJO does generally lead to a better forecast. However, this is not a panacea. The
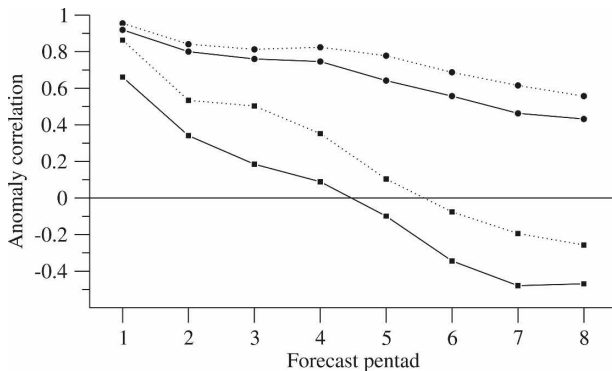
FIG. 14. Median (circles) and lower quartile (squares) anomaly correlations between forecasts and validation observations of real-time filtered anomalies of OLR against forecast pentad. Solid lines are when all forecasts (100%) are retained. Dotted lines are when only the forecasts with the strongest (50% of forecasts) initial MJOs are retained.

lower quartile of the forecasts from the initially strong MJO events still loses skill (anomaly correlation below 0.6) after 2 pentads.

5. Conclusions

The process of empirical mode decomposition (EMD), together with a modification to reduce end effects, has been shown to be a powerful technique to cleanly extract a broadband signal such as the MJO from a time series, in a form suitable for real-time monitoring and forecasting. There is much scope to apply it to other atmospheric and oceanographic phenomena, such as synoptic variability, other intraseasonal variability, and interannual variability such as ENSO.

The MJO signal, extracted using EMD, was then input to a VARMA (5, 1) statistical model to predict its future development in real time. This nonlinear model showed considerable skill, with a median anomaly correlation between forecast and verification data of above 0.6 at a forecast lead time of 25 days. The forecast skill improved when only those forecasts with a strong initial MJO were considered. This could imply that the MJO is inherently less predictable when it is weak. However, the end effects of the EMD process are more severe and would lead to larger forecast errors at these times also. Use of both (first order) moving average and higher-order autoregressive terms in the VARMA (5, 1) model gave a considerable increase in skill when compared to just using a lower-order autoregressive model. The model uses a relatively simple input, in the form of OLR data only. Hence, the model only receives information about the current and past state of deep convection or precipitation; although, because of the

thermodynamical balances in the tropics, this will also contain implicit information on vertical motion and horizontal divergence. Work is in progress to incorporate other information, such as dynamical (e.g., wind) fields from analysis data and sea surface temperatures, and to investigate other classes of statistical models such as neural networks. The intention is to implement these forecasts operationally in the near future.

The skill of statistical forecast models of the MJO is still increasing and does not appear to have reached the limit of predictability yet. These statistical models have set a high benchmark against which to measure the performance of MJO forecasts by dynamically based models.

Operational forecasts of the current MJO using this method are available in real-time at http://envam1.env. uea.ac.uk/mjo_forecast.html.

REFERENCES

Ansley, C. F., and R. Kohn, 1986: A note on reparameterising a vector autoregressive moving average model to enforce stationarity. *J. Stat. Comput. Simul.,* **24,** 99–106.

Box, G. E. P., and G. M. Jenkins, 1970: *Time Series Analysis.* Holden-Day, 553 pp.

Coughlin, K., and K. K. Tung, 2005: Empirical mode decomposition and climate variability. *Hilbert–Huang Transform and its Applications,* N. E. Huang and S. S. P. Shen, Eds., World Scientific Publishing, 149–168.

Duffy, D. G., 2005: The application of Hilbert–Huang transforms to meteorological data sets. *Hilbert–Huang Transform and its Applications,* N. E. Huang and S. S. P. Shen, Eds., World Scientific Publishing, 129–148.

Ferranti, L., T. N. Palmer, F. Molteni, and K. Klinker, 1990: Tropical–extratropical interaction associated with the 30–60-day oscillation and its impact on medium and extended range prediction. *J. Atmos. Sci.,* **47,** 2177–2199.

Gill, P. E., and W. Murray, 1972: Quasi-Newton methods for unconstrained optimization. *J. Inst. Math. Appl.,* **9,** 91–108.

Goswami, B. N., 2005: South Asian monsoon. *Intraseasonal Variability in the Atmosphere–Ocean Climate System,* W. K.-M. Lau and D. E. Waliser, Eds., Springer-Praxis, 19–62.

Hall, J. D., A. J. Matthews, and D. J. Karoly, 2001: The modulation of tropical cyclone activity in the Australian region by the Madden–Julian oscillation. *Mon. Wea. Rev.,* **129,** 2970–2982.

Hendon, H. H., and M. L. Salby, 1994: The life cycle of the Madden–Julian oscillation. *J. Atmos. Sci.,* **51,** 2225–2237.

——, B. Liebmann, M. Newman, J. D. Glick, and J. Schemm, 2000: Medium-range forecast errors associated with active episodes of the Madden–Julian oscillation. *Mon. Wea. Rev.,* **128,** 69–86.

Huang, N. E., and S. S. P. Shen, Eds., 2005: *Hilbert–Huang Trans-*

form and its Applications. World Scientific Publishing, 311 pp.

——, and Coauthors, 1998: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. Roy. Soc. London, Ser. A,* **454,** 903–995.

Jones, C., D. E. Waliser, J. K. Schemm, and W. K. Lau, 2000: Prediction skill of the Madden–Julian oscillation in dynamical extended range forecasts. *Climate Dyn.,* **16,** 273–289.

——, L. M. V. Carvalho, R. W. Higgins, D. E. Waliser, and J.-K. E. Schemm, 2004: A statistical forecast model of tropical intraseasonal convective anomalies. *J. Climate,* **17,** 2078–2095.

Li, W. K., and A. I. McLeod, 1981: Distribution of the residual autocorrelations in multivariate ARMA time series models. *J. Roy. Stat. Soc., Ser. B,* **43,** 231–239.

Liebmann, B., and C. A. Smith, 1996: Description of a complete (interpolated) OLR dataset. *Bull. Amer. Meteor. Soc.,* **77,** 1275–1277.

Lo, F., and H. H. Hendon, 2000: Empirical prediction of the Madden–Julian oscillation. *Mon. Wea. Rev.,* **128,** 2528–2543.

Madden, R. A., and P. R. Julian, 1994: Observations of the 40–50-day tropical oscillation. A review. *Mon. Wea. Rev.,* **122,** 814–837.

Maharaj, E. A., and M. C. Wheeler, 2005: Forecasting an index of the Madden–Julian oscillation. *Int. J. Climatol.,* **25,** 1611–1618.

Maloney, E. D., and D. L. Hartmann, 2000: Modulation of eastern North Pacific hurricanes by the Madden–Julian oscillation. *J. Climate,* **13,** 1451–1460.

Matthews, A. J., 2000: Propagation mechanisms for the Madden–Julian oscillation. *Quart. J. Roy. Meteor. Soc.,* **126,** 2637–2652.

Mo, K. C., 2001: Adaptive filtering and prediction of intraseasonal oscillations. *Mon. Wea. Rev.,* **129,** 802–817.

Salby, M. L., and H. H. Hendon, 1994: Intraseasonal behavior of clouds, temperature and motion in the tropics. *J. Atmos. Sci.,* **51,** 2207–2224.

Shea, B. L., 1987: Estimation of multivariate time series. *J. Time Ser. Anal.,* **8,** 95–110.

Waliser, D. E., 2005: Predictability and forecasting. *Intraseasonal Variability in the Atmosphere–Ocean Climate System,* W. K.-M. Lau and D. E. Waliser, Eds., Springer-Praxis, 389–423.

——, C. Jones, J. K. Schemm, and N. E. Graham, 1999: A statistical extended-range tropical forecast model based on the slow evolution of the Madden–Julian oscillation. *J. Climate,* **12,** 1918–1939.

——, W. Stern, S. Schubert, and K. M. Lau, 2003: Dynamic predictability of intraseasonal variability associated with the Asian summer monsoon. *Quart. J. Roy. Meteor. Soc.,* **129,** 2897–2925.

——, and Coauthors, 2006: The experimental MJO prediction project. *Bull. Amer. Meteor. Soc.,* **87,** 425–431.

Wheeler, M., and K. M. Weickmann, 2001: Real-time monitoring and prediction of modes of coherent synoptic to intraseasonal tropical variability. *Mon. Wea. Rev.,* **129,** 2677–2694.

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

——, and J. L. McBride, 2005: Australian–Indonesian monsoon. *Intraseasonal Variability in the Atmosphere–Ocean Climate System,* W. K.-M. Lau and D. E. Waliser, Eds., Springer-Praxis, 125–174.

Woolnough, S. J., F. Vitart, and M. A. Balmaseda, 2007: The role of the ocean in the Madden–Julian Oscillation: Implications for MJO prediction. *Quart. J. Roy. Meteor. Soc.,* **133,** 117–128.

Wu, M.-L. C., S. Schubert, and N. E. Huang, 1999: The development of the South Asian monsoon and the intraseasonal oscillation. *J. Climate,* **12,** 2054–2075.

Zhang, C., 2005: Madden-Julian Oscillation. *Rev. Geophys.,* **43,** RG2003, doi:10.1029/2004RG000158.