Clim Dyn (2016) 46:159–176
DOI 10.1007/s00382-015-2576-3

Predictability of the Madden–Julian Oscillation index: seasonality and dependence on MJO phase

Eric C. J. Oliver[1,2,3] · **Keith R. Thompson**[2]

Received: 20 November 2014 / Accepted: 17 March 2015 / Published online: 31 March 2015

Abstract We describe here a damped harmonic oscillator model for the Wheeler and Hendon (Mon Weather Rev 132(8):1917–1932, 2004) Madden–Julian Oscillation (MJO) index in order to gain new insights into the predictability of the MJO. Building on a tradition of idealized models, the model for the MJO state consists of a bivariate autoregressive process, equivalent to a finite difference approximation to a dynamical underdamped harmonic oscillator, as represented by a second order ordinary differential equation. The statistical properties of the model, namely the ensemble mean, ensemble variance, and within-ensemble correlation, are used to develop predictability time scales for canonical MJO events. We explore the model under both white noise and coloured noise forcing and the model parameters are estimated using maximum likelihood estimation, as a function of season and initial MJO event amplitude and phase. The model provides a significantly better fit using coloured noise forcing, which is equivalent to using a higher order model, indicating that the MJO index is not a simple order-1 coupled autoregressive process. Using the fitted model we map the predictability times scales for the mean, variance, and correlation as a function of initial MJO position in phase space. It is shown that the predictability time scales, and thus MJO predictability, vary as a function of MJO phase space and season which is a novel result for empirical models of the MJO. The result that MJO predictability varies with MJO state also has relevance for the interpretation of the Maritime Continent prediction barrier.

Keywords Madden–Julian Oscillation · Statistical model · Predictability · Autoregressive process · Intraseasonal variability

1 Introduction

The Madden–Julian Oscillation (MJO) is the dominant mode of intraseasonal variability in the tropical atmosphere (Madden and Julian 1971, 1972, 1994; Zhang 2005). The MJO, like weather, has pulses of variability with a finite predictability time scale, i.e., future MJO states may be predicted up to a certain lead time based on knowledge of current and past states. In a recent review, Waliser (2012) suggested that the predictability time scale of the MJO is at least 2 to 3 weeks. Recent model-based studies have estimated the time scale to be 20–24 days using a mixed ensemble of statistical and dynamical models (Kang and Kim 2010), 20–45 days using ensembles of dynamical models (with longer predictability estimates from multi-model means rather than single ensemble members; Neena et al. (2014)), and 32 days using ensembles of modern climate forecast systems (Kim et al. 2014). An estimate based on observations, using a chaos dynamics approach with nonlinear local Lyapunov exponents, found predictability time scales of 36 days for outgoing longwave radiation (OLR), 39 days for 850 hPa zonal wind, 41 days for combined OLR, 850 and 200 hPa zonal wind, and 21 days for the daily unfiltered (Wheeler and Hendon 2004) MJO index (Ding et al. 2010). It should be noted that forecasts

---

✉ Eric C. J. Oliver
eric.oliver@utas.edu.au

[1] Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia

[2] Department of Oceanography, Dalhousie University, Halifax, NS, Canada

[3] Australian Research Council Centre of Excellence for Climate System Science, Hobart, Tasmania, Australia

of the MJO are typically skillful over time scales much shorter than the predictability time scale; the majority of forecast models show useful skill out to only 15 days with a few having skillful forecasts beyond two weeks (Zhang et al. 2013). It has also been suggested that the predictability time scale of an MJO event may depend on the amplitude or phase at which the event is initiated (Vintzileos and Pan 2007; Lin et al. 2008; Kim et al. 2009; Kang and Kim 2010). The predictability of the MJO is generally longer than the predictability limit attributed to weather (1–2 weeks) and thus provides potential predictability on time scales that bridge the gap between weather and climate. It has been argued (e.g., Zhang 2013) that a better understanding of this phenomenon, and how it influences the atmosphere-ocean-climate system, may help with the development of a seamless forecasting framework.

The MJO has been modeled statistically in many ways in order to understand its statistical behaviour and predictability. von Storch and Xu (1990a) fit a pair of POPs (von Storch et al. 1988) to the tropical 200-mb velocity potential and the resultant POP coefficient acts like a bivariate index of magnitude and phase of the intraseasonal oscillation. This statistical model provides useful forecasts out to 7 days (von Storch and Xu 1990a; von Storch and Baumhefner 1991) with greater skill in Boreal Winter (NDJF) than in Boreal Summer (MJJA) and for initial conditions of larger magnitude than of weaker magnitude. Lo and Hendon (2000) used a reduced set of EOFs of outgoing longwave radiation and the 200-mb streamfunction to predict the full set of EOFs at some forecast lead time. They found forecasts with their statistical model were skillful out to 15 days and better than those from a dynamical model out to about 1 week. Similar models were used by Waliser et al. (1999) who forecast the future state of the MJO, using present and past fields of bandpassed OLR, and by Jiang et al. (2008) who used principal components of combined OLR and 200- and 850-mb zonal velocities in a multivariate linear regression model to forecast individual fields of these variables. More recently, Maharaj and Wheeler (2005) developed a bivariate, stochastically-forced, order 1 autoregressive model for the Wheeler and Hendon (2004) MJO index. They found the model parameters to be different when fit over Boreal Summer (May–October) and Austral Summer (November–April). More recently, Cavanaugh et al. (2014) used a linear stochastic model with white noise to simulate not only a time series of the MJO but also full spatial fields of OLR and winds and Kondrashov et al. (2013) developed a low-dimensional stochastic model for the MJO which takes into account nonlinearity, seasonality, and serial correlation in the noise with useful prediction skill out to 30 days. One use of the statistical models described above is as a benchmark against which to test the performance of numerical model forecasts. For example, von Storch and Baumhefner (1991) compared the forecasts from their statistical model against those from two general circulation models, which were not overall superior to the POP forecast, and Rashid et al. (2011) showed the Predictive Ocean Atmosphere Model for Australia (POAMA) provided a better forecast of the MJO than the Maharaj and Wheeler (2005) statistical baseline.

Here we develop a forced damped harmonic oscillator (FDHO) model for the MJO index in discrete time following the autoregressive model developed by Oliver (2011) and Oliver and Thompson (2012). This model builds on the bivariate autoregressive model of by Maharaj and Wheeler (2005) and shares many features with the POP analysis performed by von Storch et al. (1988). In fact, the time-dependent behaviour of the MJO state resulting from the POP analysis is linked with the standard damped harmonic oscillator equations (von Storch et al. 1988). Those studies focused on the predictive skill of statistical models in terms of capturing the MJO index. We use our model to better understand the predictability of the MJO and test if this predictability varies with initial MJO amplitude and phase and with season.

This paper is organized as follows. The MJO index, its representation in phase space, and the tracking of ensembles of MJO events are presented in Sect. 2. The FDHO model, including its statistical properties, is outlined in Sect. 3 and fit to the observed MJO index as a function of MJO amplitude and phase in Sect. 4. A summary and discussion of the results are provided in Sect. 5.

2 The Madden–Julian Oscillation

In this section we define and describe the MJO index (Sect. 2.1) and visualize ensembles of MJO events in phase space (Sect. 2.2).

2.1 MJO index

The index developed by Wheeler and Hendon (2004) is used here to characterize the MJO. This index is bivariate and consists of the first two principal components of an EOF analysis of tropically averaged (15°S to 15°N) fields of outgoing longwave radiation (as a proxy for precipitation) and zonal wind at two heights (200 and 850 hPa). The intraseasonal signals in each of these three variables are primary indicators of the MJO and their multivariate combination into the Wheeler and Hendon (2004) index further suppresses high frequency signals which are incoherent amongst these three variables. A daily time series of the Wheeler and Hendon (2004) index, hereafter referred to as the MJO index, was obtained from 1 January 1979 to 31 December 2008 from the Government of Australia Bureau

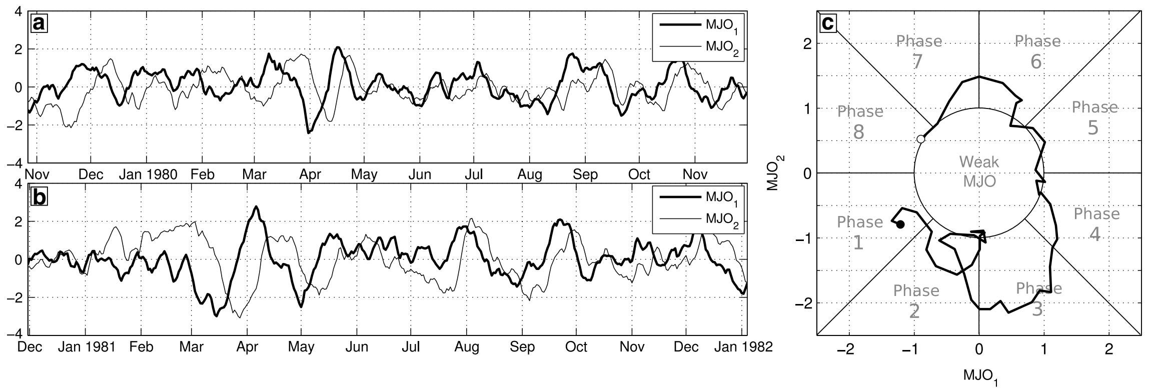

Fig. 1 Example MJO time series and phase space trajectory. Time series of the MJO index (first component: *thick line*; second component: *thin line*) over 400 days are shown for 27 October 1979 to 30 November 1980 (**a**) and for 30 November 1980 to 4 January 1982 (**b**). A trajectory in MJO phase space over 50 days is shown in **c** for 27 October 1979 (*solid dot*) to 16 December 1979 (*open dot*)

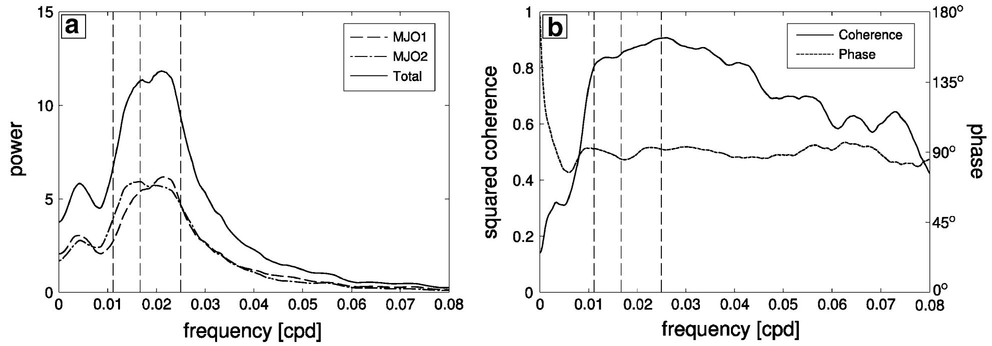

Fig. 2 Power, coherence and phase spectra of the MJO index as a function of frequency. **a** The total power spectra (*solid line*), calculated as the sum of the power spectral densities of the two MJO index components (*dashed* and *dash-dotted lines*). The coherence (*solid line*) and phase (*dashed line*) spectra between the two MJO index components are shown in **b**. The *vertical dashed lines* indicate oscillation periods of 40, 60 and 90 days

of Meteorology (the Centre for Australian Weather and Climate Research, see http://cawcr.gov.au/staff/mwheeler/maproom/RMM/). The two components have zero mean and have been normalized so that each have standard deviation of one.

The two components of this bivariate time series describe an oscillating phenomenon with periods between 30 and 90 days and they are approximately in quadrature. The oscillations and relative phase offset of the two components is clear from time series plots (e.g., Fig. 1a, b). The majority of energy lies in a band of oscillation periods between 40 and 60 days (Fig. 2a). The two components of the index are highly coherent over these time scales (squared coherence values of 0.8–0.9; Fig. 2b, solid line) and are in quadrature (90° out of phase; Fig. 2b, dashed line). (Coherence is confined to lie between 0 and 1 and can roughly be thought of as a “frequency-dependent” correlation coefficient while the phase spectra indicates the phase offset at a particular frequency.)

2.2 MJO events in phase space

The nature of the MJO index, namely that the two components are approximately oscillatory and in quadrature, leads to a natural representation of the MJO state in “phase space” (Wheeler and Hendon 2004). Phase space is a two dimensional representation of the MJO where the first component of the index defines the x_1-axis and the second component defines the x_2-axis (Fig. 1c). The radius and angle of a particular point in phase space represents the strength and phase of the MJO, respectively. The angle can also be interpreted as the region over which the active convection associated with the MJO is situated. An MJO trajectory in phase space typically follows a counterclockwise path as an event propagates eastward around the globe (e.g., Fig. 1c). Phase space is conventionally divided into 9 regions: a “weak MJO” region, defined for amplitude less than one, and eight regions (denoted “Phase 1” through “Phase 8”; see Fig. 1c) spaced equally around the remainder of phase space. The centre of MJO convection is over the Western Hemisphere and Africa during phases 8 and 1, the Indian Ocean during phases 2 and 3, the Maritime Continent during phases 4 and 5, and the Western Pacific during phases 6 and 7.

A scatter plot of all MJO index values is shown in phase space in Fig. 3a (grey points). The MJO values are approximately distributed as a bivariate normal distribution with mean zero and a variance of one in both the x_1 and x_2 directions. We wish to subset from this distribution to represent an observed initial condition for the MJO. We do this by taking a Bayesian approach. First we assume that the prior distribution is approximated by the grey points (the scatter plot of all MJO index values) and then we assume the observed initial condition is imprecise and can be represented by a probability density function (pdf). We then sample from the posterior distribution by applying a weight to the grey points according to the pdf of the observed initial condition and resampling accordingly. For example, if the observed initial condition pdf were a bivariate normal distribution centred away from the origin (Fig. 3a, black dot and circle) the posterior sample, or samples from the initial condition, are shown as the red and blue dots in Fig. 3a. The mathematical details behind this method for choosing an ensemble of MJO events, given an initial condition, are provided in Appendix 1. We consider this subset of MJO values to be a statistical representation of the "initial" state of the MJO at some time t_0.

The distribution of the state of the MJO at a future time $t_0 + k$ is approximated by simply following the evolution, according to the original MJO time series starting at time t_0, of each of the "initial" points in phase space (i.e., the red and blue dots). This procedure is demonstrated in Fig. 3 for time lag k increasing from zero to 30 days (red and blue dots). It is clear that the set of points undergoes a counterclockwise rotation and decay of the mean position of the ensemble of points towards zero (black dot). The spread of points increases through time (black circle). It can also be seen that the position of the ensemble members relative to their sample mean at time t_0 is lost with increasing time. This can be thought of as mixing of relative positions in phase space within the ensemble. For example, ensemble members that were to the clockwise side of the ensemble mean at t_0 (Fig. 3a, blue dots) are less and less likely to remain on that side of the ensemble mean as k increases.

The ensemble statistics are summarized in Fig. 4, for ensembles of trajectories with initial conditions (the initial state, or "observation") in four widely separated parts of phase space. The evolution of the ensemble mean (i.e., rotation and decay in MJO space; Fig. 4, black line) is reminiscent of the spiraling behaviour of a forced, damped harmonic oscillator, when the oscillator velocity is plotted versus oscillator position (e.g., Marion and Thornton 1995). We also note that rates of decay of the mean and increase of the variance appear to depend on initial condition (e.g., the variance (shaded area) increases faster in Fig. 4b than in Fig. 4c) and that these rates are fast compared to the oscillation frequency. This behaviour will be further demonstrated and quantified in Sect. 4.

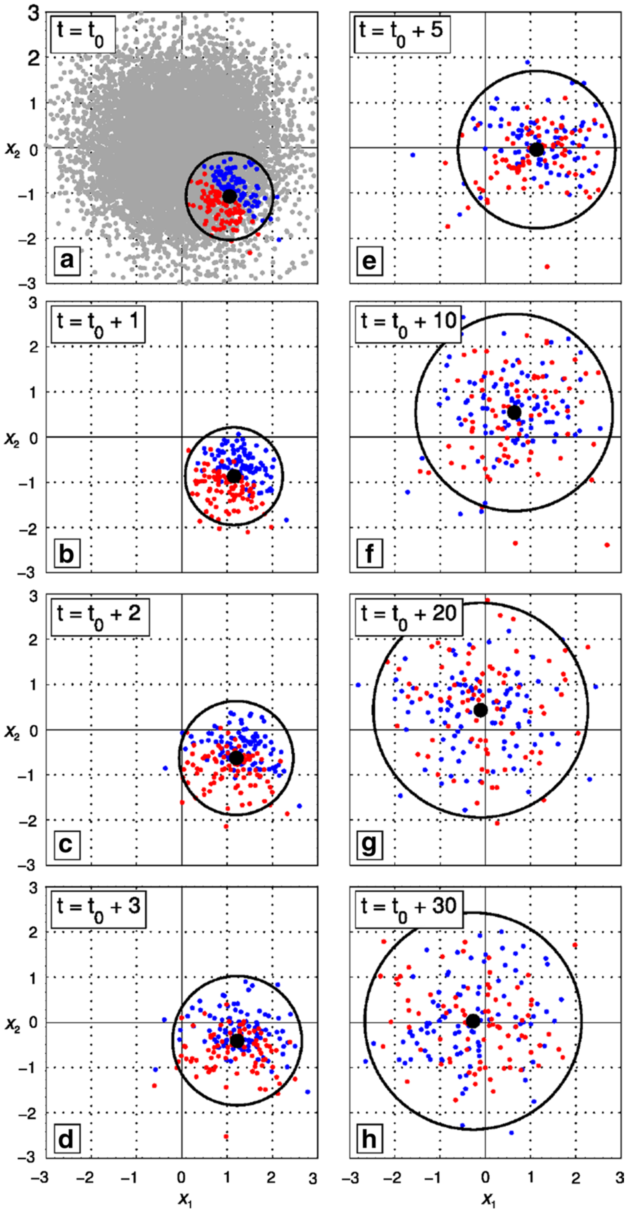

Fig. 3 Evolution of an ensemble of MJO events with initial condition in the vicinity of a specified location. The *grey points* in **a** show all values of the MJO index in phase space. The *blue* and *red points* in **a** show a sample of MJO index values representing the initial condition at time t_0 (sampled from a bivariate normal distribution with mean and 95 % enclosure indicated by the *black dot* and *circle* respectively). The *blue* and *red points* have a coordinate value to the clockwise side and the counterclockwise side of the initial mean value respectively. This ensemble of MJO trajectories are shown for **b** $k = 1$, **c** 2, **d** 3, **e** 5, **f** 10, **g** 20 and **h** 30 days, where k is the number of time steps following the initial time t_0

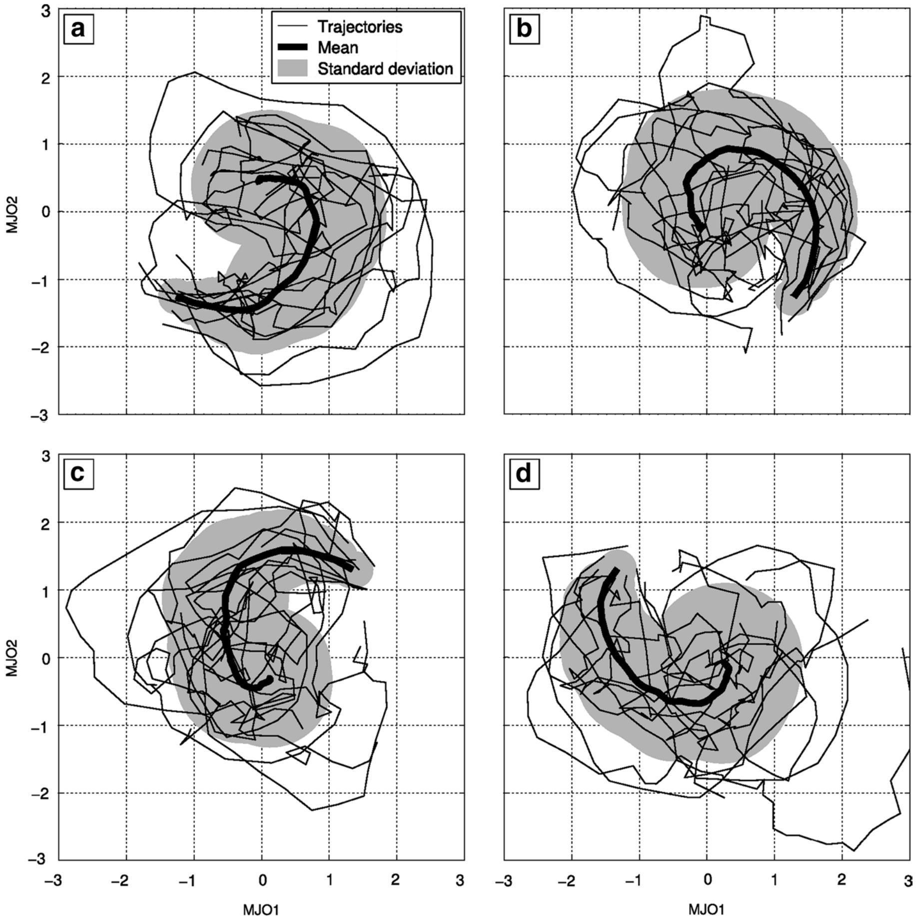

Fig. 4 Evolution of the statistics of an ensemble of MJO events. The statistical properties of a 4000-member ensemble of MJO events are shown from some initial state at time t_0 for the following 35 days. Four initial ensembles are chosen, one in each of the four corners of MJO space (**b** corresponds to the example shown in Fig. 3). The *thick black line* shows the mean, the shaded area shows the standard deviation about this mean, and the thin lines show a subset of 20 representative trajectories

Table 1 Model parameters (first column), relationship to time scales (second column), and interpretation (third column)

γ_1	$\tau_1 = -1/\log\gamma_1$	State decay rate (γ_1) or time scale (τ_1) [Models 1 and 2]
γ_2	$\tau_2 = -1/\log\gamma_2$	Forcing decay rate (γ_2) or time scale (τ_2) [Model 2 only]
θ	$P = 2\pi/\theta$	Angular rotation frequency (θ) or oscillation period (P) [Models 1 and 2]

The time step Δt is chosen to be 1 day leading to a unit of days for the time scale parameters

3 The forced damped harmonic oscillator model

A forced damped harmonic oscillator (FDHO) model is described in Sects. 3.1–3.3 and its statistical properties are then reviewed in Sects. 3.4–3.5.

3.1 Damped harmonic oscillator model in discrete time

A forced damped harmonic oscillator model (e.g., Marion and Thornton 1995) can be expressed in discrete time as follows (see Priestley 1981 for more details):

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}_1\boldsymbol{x}_t + \boldsymbol{f}_{t+1}, \tag{1}$$

where $\boldsymbol{x}_t = [x_1\ x_2]'_t$ is the bivariate state vector at time t ($'$ denotes a transpose), $\boldsymbol{A}_1$ is the 2×2 transition matrix that carries the state forward one time step, and $\boldsymbol{f}_t$ is a bivariate forcing process. For an underdamped harmonic oscillator, $\boldsymbol{A}_1$ can be written

$$\boldsymbol{A}_1 = \gamma_1 \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{2}$$

where $0 < \gamma_1 < 1$ is the damping coefficient and $\theta \neq 0$ is the rotation rate. In the absence of forcing the oscillator rotates through an angle θ over one time step, and its amplitude is scaled by a factor γ_1. The damping coefficient γ_1 and the rotation rate θ can be expressed in terms of a decay time scale τ_1 and a rotation period P (see Table 1). Hereafter we will primarily discuss the model parameters in their time scale form (using a time step of $\Delta t = 1$ day).

3.2 Damped harmonic oscillator model with white noise forcing

If the oscillator model presented in Sect. 3.1 is forced by a Gaussian white noise process, $\boldsymbol{\epsilon}_t$, then it can be represented by Eq. 1 with $\boldsymbol{f}_t = \boldsymbol{\epsilon}_t$, a model equivalent to a stochastically forced complex order one autoregressive, or AR(1), model (e.g., von Storch and Xu 1990b; von Storch and Baumhefner 1991):

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}_1\boldsymbol{x}_t + \boldsymbol{\epsilon}_{t+1}, \tag{3}$$

where the covariance of $\boldsymbol{\epsilon}_t$ is given by $\boldsymbol{\Sigma}_\epsilon = \sigma_\epsilon^2\mathrm{I}$ where σ_ϵ^2 is the variance of the white noise process and I is a 2×2 identity matrix. The noise variance σ_ϵ^2 is chosen so that the steady state variance of the two oscillator components of $\boldsymbol{x}_t$ are each unity. This model is henceforth referred to as "Model 1". This AR(1) model is nearly equivalent to the bivariate AR(1) model developed by Maharaj and Wheeler (2005) with the exception that the model transition matrix ($\boldsymbol{A}_1$) has just two parameters (γ, θ) instead of four, essentially requiring the two oscillator components to have identical statistical properties.

3.3 Damped harmonic oscillator model with autoregressive forcing

We now consider the case where the white noise stochastic forcing is replaced by coloured noise stochastic forcing $\boldsymbol{f}_t$, given by a bivariate AR(1) process. In other words, $\boldsymbol{f}_t$ in Eq. 1 is a bivariate AR(1) process and the combined oscillator/forcing system can be written as a pair of coupled bivariate processes:

$$\boldsymbol{x}_{t+1} = \boldsymbol{A}_1\boldsymbol{x}_t + \boldsymbol{f}_{t+1}, \tag{4}$$

$$\boldsymbol{f}_{t+1} = \boldsymbol{A}_2\boldsymbol{f}_t + \boldsymbol{\epsilon}_{t+1}, \tag{5}$$

where Model 1 is now nested within this model (as Eq. 4) and is supplemented by the equation to generate the autoregressive stochastic forcing (Eq. 5). The transition matrix for the autoregressive forcing is given by

$$\boldsymbol{A}_2 = \gamma_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{6}$$

and the covariance of $\boldsymbol{\epsilon}_t$ is again given by $\boldsymbol{\Sigma}_\epsilon = \sigma_\epsilon^2\mathrm{I}$. The bivariate forcing process $\boldsymbol{f}_t$ is modeled as a pair of independent AR(1) processes each with the same autoregressive parameter γ_2 (Eq. 5) and these force the oscillator (Eq. 4). As with the damping parameter γ_1, the autoregressive parameter γ_2 can be expressed as an autoregressive time scale τ_2 (see Table 1). The memory of the forcing is therefore controlled by the parameter γ_2 which is constrained to be between 0 and 1. In the present context the forcing is taken to represent the effect of short time-scale weather events on the longer time-scale MJO. The introduction of the parameter γ_2 allows us to control the auto-correlation (or, equivalently, the typical duration) of these weather events. This new model is now defined by three parameters (P, τ_1, and τ_2) and is henceforth referred to as "Model 2".

Model 2 can also be represented by the following quadrivariate stochastically forced AR(1) model:

$$\boldsymbol{y}_{t+1} = \boldsymbol{A}\boldsymbol{y}_t + \boldsymbol{\varepsilon}_{t+1}, \tag{7}$$

where $\boldsymbol{y}_t$ is now a vector of length 4 combining the oscillator and forcing components ($\boldsymbol{y}_t = [\boldsymbol{x}_t\, \boldsymbol{f}_t]'$), $\boldsymbol{A}$ is a 4×4 transition matrix, and $\boldsymbol{\varepsilon}_t$ is the quadrivariate stochastic noise process that drives the system away from a state of rest. The first two components of $\boldsymbol{y}_t$ are referred to as the oscillator components (they represent the oscillator described by Eq. 4) and the last two components are referred to as the forcing components as they represent the autoregressive forcing (desribed by Eq. 5). The transition matrix takes the partitioned form

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 & \mathrm{I} \\ \boldsymbol{0} & \boldsymbol{A}_2 \end{bmatrix}. \tag{8}$$

This quadrivariate AR(1) process is itself driven by an independent white noise process with the 4×4 covariance matrix

$$\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} = \sigma_\epsilon^2\begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \mathrm{I} \end{bmatrix} \tag{9}$$

so that the white noise process only directly influences the third and fourth components of $\boldsymbol{y}_t$ (i.e., the forcing components).

3.4 Statistical properties of the damped harmonic oscillator model

It is straightforward to calculate the statistical properties of the damped harmonic oscillator model. The following discussion focuses on Model 2; these properties also hold for Model 1 with the substitution of $\boldsymbol{x}_t$ for $\boldsymbol{y}_t$, $\boldsymbol{A}_1$ for $\boldsymbol{A}$, and $\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}$ for $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$.

The process defined by Eq. 7 is asymptotically stationary to second order if $0 \leq (\gamma_1, \gamma_2) < 1$. Its steady state covariance matrix $\boldsymbol{\Sigma}_\infty$ satisfies

$$\boldsymbol{\Sigma}_\infty = \boldsymbol{A}\boldsymbol{\Sigma}_\infty\boldsymbol{A}' + \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \tag{10}$$

with the solution given by

$$\mathrm{vec}(\boldsymbol{\Sigma}_\infty) = [\mathrm{I} - \boldsymbol{A} \otimes \boldsymbol{A}]^{-1}\mathrm{vec}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}), \tag{11}$$

where vec denotes the vectorization of a matrix, performed by sequentially stacking the columns of the specified matrix (e.g., Harvey 1991) and $\otimes$ denotes the Kronecker product. Equation 11 allows us to specify $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ in order to achieve a desired asymptotic covariance $\boldsymbol{\Sigma}_\infty$ (i.e., $\boldsymbol{\Sigma}_\infty = \mathbf{I}$, over the oscillator components, in the case of unity oscillator variance).

We are particularly interested in the evolution of the statistical properties of the model from a known initial state at time t_0. If the mean and variance of $\boldsymbol{y}_t$ at time t_0 are $\boldsymbol{\mu}_{t_0}$ and $\boldsymbol{\Sigma}_{t_0,t_0}$ respectively, the mean and variance at time t_0+k are given by

$$\boldsymbol{\mu}_{t_0+k} = \boldsymbol{A}^k \boldsymbol{\mu}_{t_0} \tag{12}$$

and

$$\boldsymbol{\Sigma}_{t_0+k,t_0+k} = \boldsymbol{A}^k \boldsymbol{\Sigma}_{t_0,t_0} \boldsymbol{A}'^k + \sum_{j=0}^{k-1} \boldsymbol{A}^j \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \boldsymbol{A}'^j \tag{13}$$

respectively. The covariance between $\boldsymbol{y}_{t_0}$ and $\boldsymbol{y}_{t_0+k}$ is given by

$$\boldsymbol{\Sigma}_{t_0+k,t_0} = \boldsymbol{A}^k \boldsymbol{\Sigma}_{t_0,t_0}. \tag{14}$$

The proportion of total variance of the oscillator state at t_0+k that can be accounted for by the initial oscillator state is given by

$$\rho_k^2 = \frac{\mathrm{tr}_2(\boldsymbol{\Sigma}_{t_0+k,t_0} \boldsymbol{\Sigma}_{t_0,t_0}^{-1} \boldsymbol{\Sigma}_{t_0,t_0+k})}{\mathrm{tr}_2(\boldsymbol{\Sigma}_{t_0+k,t_0+k})}, \tag{15}$$

where the trace acts only over the first two (oscillator) components. Finally, the cross-spectral matrix of $\boldsymbol{y}_t$ is given by Priestley (1981)

$$\boldsymbol{f}_{\boldsymbol{yy}}(\omega) = \left[e^{i\omega}\mathrm{I} - \boldsymbol{A}\right]^{-1} \boldsymbol{f}_{\boldsymbol{\varepsilon\varepsilon}}(\omega) \left[e^{i\omega}\mathrm{I} - \boldsymbol{A}\right]^{*-1}, \tag{16}$$

where $*$ denotes a conjugate transpose and $\boldsymbol{f}_{\boldsymbol{\varepsilon\varepsilon}}(\omega)$ is the cross-spectral matrix of $\boldsymbol{\varepsilon}_t$. For white noise (Model 1) $\boldsymbol{f}_{\boldsymbol{\varepsilon\varepsilon}}(\omega)$ is simply $\frac{\sigma_\epsilon^2}{2\pi}\mathrm{I}$.

Three measures of oscillator predictability are now defined. The first is based on the oscillator amplitude, given by the length of the mean $|\boldsymbol{\mu}_{t_0+k}|$ (this is analogous to the ensemble mean MJO amplitude, and is the distance to the origin in MJO phase space), which allows us to monitor the decay rate of the oscillator. The second predictability measure is based on the total oscillator variance given by $\sigma_{t_0+k}^2 = \mathrm{tr}_2(\boldsymbol{\Sigma}_{t_0+k,t_0+k})$. It measures the rate at which trajectories with similar initial conditions diverge. (The length and trace operations act only over the oscillator components of $\boldsymbol{y}_t$.) The third predictability measure is the correlation ρ_k (Eq. 15). It measures the rate at which the information on the initial relative position of an event within an ensemble of oscillator events is lost. For normal white noise forcing (Model 1) the expressions for the oscillator amplitude, total oscillator variance, and correlation can be calculated explicitly given the model time scales τ_1 and P. For autoregressive forcing (Model 2) the statistical properties involve complicated correction terms due to the initial conditions of the autoregressive forcing. Therefore, we do not state them explicitly here and simply calculate them numerically using Eqs. 12–15.

We now define three time scales of predictability given these three statistics. First, τ_μ, the rate of decay of the oscillator amplitude towards zero, given by its e-folding time scale. Second, τ_σ^2, the increase of total oscillator variance, given by the e-folding time scale towards the asymptotic variance of $\sigma_\infty^2 = 2$. Third, τ_ρ, the decrease of correlation between the oscillator state at time t_0 and t_0+k, given by the time at which the correlation drops below 0.5. Note that the three predictability time scales can also be calculated from the observed MJO index. This is done by replacing $\boldsymbol{\mu}_{t_0+k}$, $\sigma_{t_0+k}^2$, and ρ_k with classic definitions of the ensemble sample mean, ensemble sample variance, and within-ensemble correlation and simply reading off the time-lag at which these quantities reach the critical values defined above. This will be used as a form of model validation in Sect. 4.

3.5 Model behaviour

A realization of $\boldsymbol{x}_t$ for Model 1 with model parameters $\tau_1 = 15$ days and $P = 50$ days (taken from Oliver and Thompson 2012) is shown in Fig. 5a–c. The variance of the noise process is $\sigma_\epsilon^2 = 1.25 \times 10^{-1}$ and the time step Δt is chosen to be 1 day. The components of $\boldsymbol{x}_t$ are clearly in quadrature (Fig. 5a) and when plotted in phase space (i.e., $x_{2,t}$ vs. $x_{1,t}$) evidence of counterclockwise rotation (i.e., oscillation) and stochastic forcing is clear (Fig. 5c, thick line). The white noise forcing process (Fig. 5b) determines the noisy behaviour of $\boldsymbol{x}_t$ (e.g., abrupt changes in direction of the trajectory in phase space over a single time step; Fig. 5c, thin line).

A realization of $\boldsymbol{y}_t$ for Model 2 with model parameters $\tau_1 = 15$ days, $\tau_2 = 2$ days, and $P = 50$ days is shown in Fig. 5(d–f). The variance of the noise process is $\sigma_\epsilon^2 = 2.28 \times 10^{-2}$; the value of σ_ϵ^2 for Model 2 differs from that for Model 1 since the two models have different responses to stochastic forcing and so different values are required in order to achieve a variance of one for the oscillator components. The oscillator components of $\boldsymbol{y}_t$ are approximately in quadrature as expected (Fig. 5d) and when plotted in phase space, evidence of counterclockwise rotation and stochastic forcing is again clear (Fig. 5f, thick line). As expected both the oscillator and forcing components vary more smoothly when the forcing processes is autoregressive (compare Fig. 5d, e with Fig. 5a, b).

The spectral properties of the two models, for the model parameters given above, demonstrate that the power of the oscillator is confined to a narrow band about the natural oscillation frequency of $2\pi/P$ (Fig. 6a). It is also clear that the oscillator components are highly coherent with each

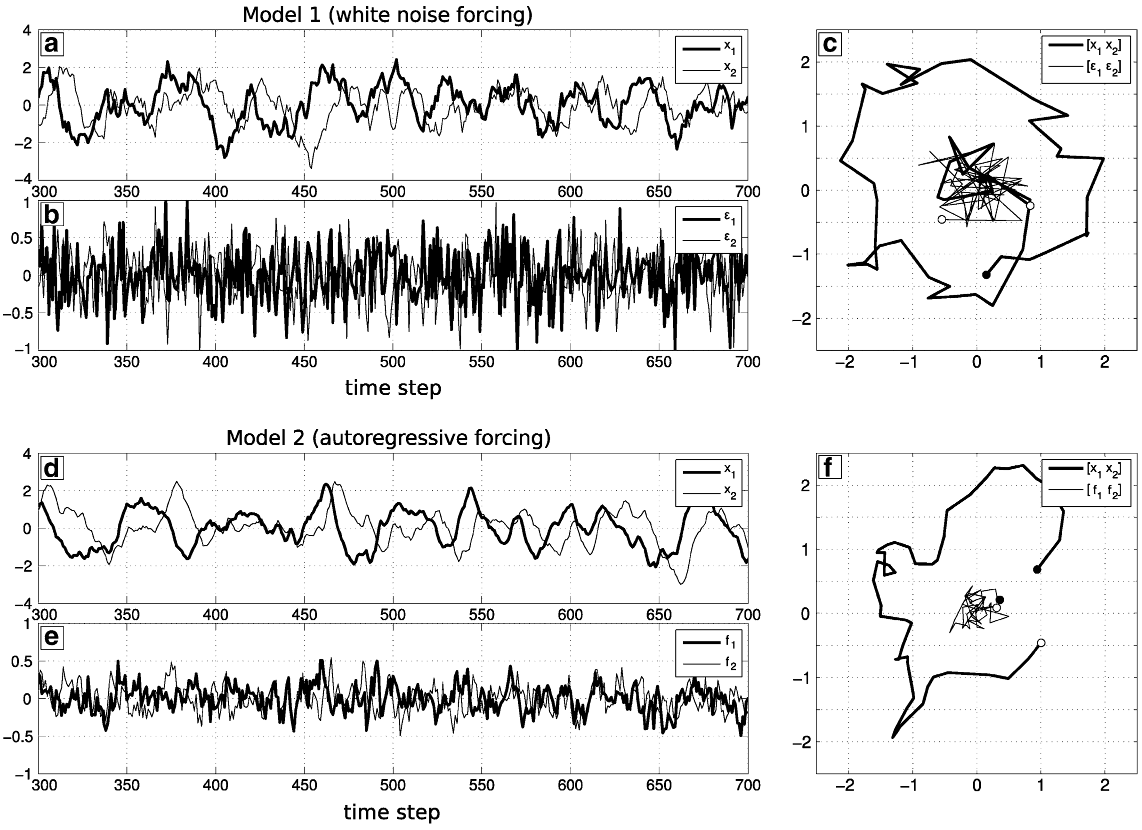

Fig. 5 Damped harmonic oscillator model in discrete time with normal white noise forcing (Model 1, upper panels) and autoregressive stochastic forcing (Model 2, lower panels). *On the left*, the oscillator components (**a**, **d**) and the forcing components (**b**, **e**) of the models are plotted for time steps 300–700. The first 300 time steps are discarded to avoid transient effects from the initial conditions $\boldsymbol{x}_0 = [0.574, -1.39]$ (equivalent to an MJO index value with amplitude 1.5 in phase 3). In **c** and **f**, the second oscillator component is plotted against the first oscillator component (*thick line*) and the second forcing component is plotted against the first forcing component (*thin line*) for time steps 300 (*closed circle*) to 350 (*open circle*)

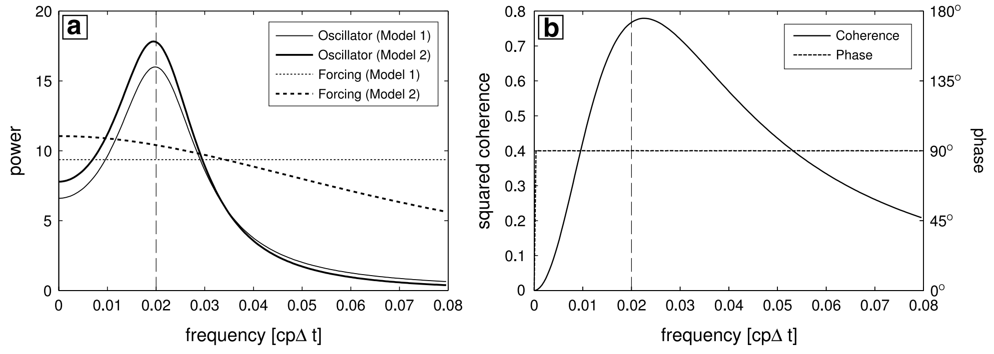

Fig. 6 Spectral properties of the damped harmonic oscillator models. **a** The power spectral densities of the first oscillator component (*solid lines*) and the first forcing component (*dashed lines*, scaled by a factor of 75) of $\boldsymbol{x}_t$ are plotted as a function of frequency for Model 1 (*thin lines*) and Model 2 (*thick lines*). **b** The coherence (*solid line*) and phase (*dashed line*) between the two oscillator components of $\boldsymbol{x}_t$. Note that the coherence and phase spectra for Model 1 and Model 2 are identical (not obvious as the *lines* are overplotted). The *dashed vertical lines* show the natural oscillation frequency ($2\pi/P$)

other and are in quadrature (90° phase shift; Fig. 6b). It is interesting to note that the coherence and phase spectra for Model 1 and Model 2 are identical. Both Model 1 and Model 2 consist of the same general process: a bivariate damped harmonic oscillator forced by bivariate stochastic noise. The only difference is the nature of the stochastic

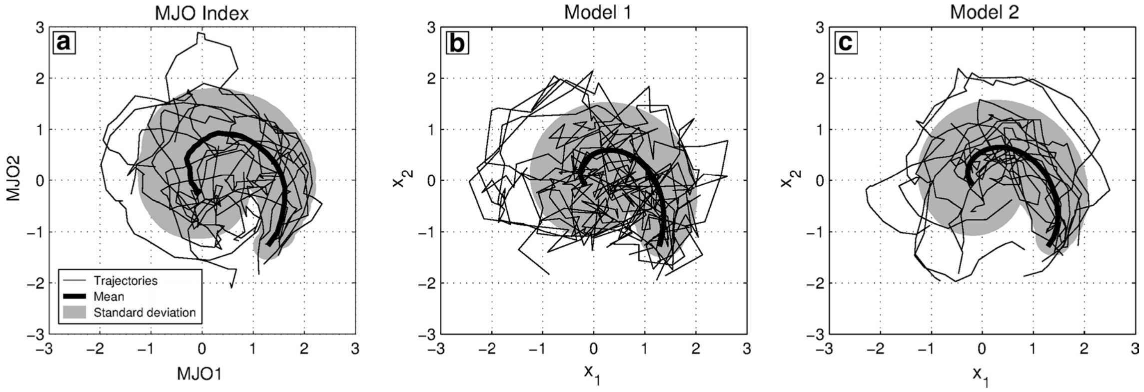

Fig. 7 Evolution of the statistics of an ensemble of MJO events. **a** The statistical properties of a 4000-member ensemble of MJO events with an initial observation at $(\sqrt{2}, -\sqrt{2})$ and with error $\sigma^2_{\text{obs}} = 0.1$ are shown for the following 35 days (as in Fig. 4b). The statistical properties of **b** Model 1 and **c** Model 2. The *thick black line* shows the mean, the *shaded area* shows the standard deviation about this mean, and the *thin lines* show 20 representative trajectories

Table 2 Estimated model parameters, in units of days

	τ_1	τ_2	P
Model 1	51.0	–	54.5
Model 2	16.2	1.58	53.4

Parameters estimated by maximum likelihood using the entire MJO index time series. Parameters are shown in units of days

noise: white noise (Model 1) or an auto-regressive process (Model 2). This difference contributes to a difference in power spectra with Model 2 having more energy at lower frequencies than Model 1 due to the “coloured” noise forcing (Fig. 6a). However, in both cases the two components of the bivariate stochastic noise are independent (uncorrelated with one another) and so will not contribute to any covariation between the two oscillator components. Therefore, the coherence and phase spectra for Model 1 and Model 2 arise from the nature of the harmonic oscillator dynamics (i.e., the $\boldsymbol{A}_1$ matrix which is identical between models) and not from the stochastic forcing.

There is clearly strong similarity in the spectral properties of the model and the observed MJO index (Figs. 6, 2). The evolution of modeled ensemble statistics also exhibit the same features as the observed MJO index (Fig. 7), namely the rotation and decay of the mean and the increase of the variance with time.

4 Model fitting and predictability

The two models described in Sect. 3 are now fit to the observed MJO index. First, we fit both Model 1 and Model 2 to the complete MJO index time series and compare the two models (Sect. 4.1). The preferred model is then fit to MJO ensembles, initiated over a broad range of MJO amplitudes and phases, in order to map the model parameters (Sect. 4.2), and predictability time scales (Sect. 4.3), in phase space.

4.1 Model selection

Parameter estimation was performed using maximum likelihood estimation with the likelihood function given in Appendix 2. Estimated model parameters are shown in Table 2. Estimated parameters for Model 1 are $\tau_1 = 51.0$ days and $P = 54.5$ days (with $\sigma_\epsilon = 0.1961$). Note that the estimated model parameters for the bivariate AR(1) model of Maharaj and Wheeler (2005) have forms nearly identical to that of Model 1 (i.e., approximately in the form of $\boldsymbol{A}_1$, see Eq. 2). Writing their model parameters into this form (by averaging cross-diagonal elements and setting them equal to $\gamma_1 \cos\theta$ and $\gamma_1 \sin\theta$) we can transform their model parameters into the parameters of Model 1, yielding $P = 49.8$ days (61.4 days) and $\tau_1 = 44.3$ days (34.7 day) for Southern Summer (Northern Summer). These values are broadly consistent with our maximum likelihood estimates.

Estimated model parameters for Model 2 are $\tau_1 = 16.2$ days, $\tau_2 = 1.58$ days and $P = 53.4$ days (with $\sigma_\epsilon = 0.1697$). Note that the estimated value of τ_1 for Model 2 is more than a factor of three smaller than for Model 1. This is because autoregressive models are effectively estimated by fitting the model's autocovariance function (ACF) to the observed ACF for small lags. Model 1, a bivariate AR(1) process, fits to the lag one ACF while Model 2, which can be expressed as a bivariate AR(2) process (Appendix 2), fits to the ACF for lags one and two. The reason the τ_1 estimates for Model 2 are much less than the corresponding estimates for Model 1 is that Model 2 was able to better

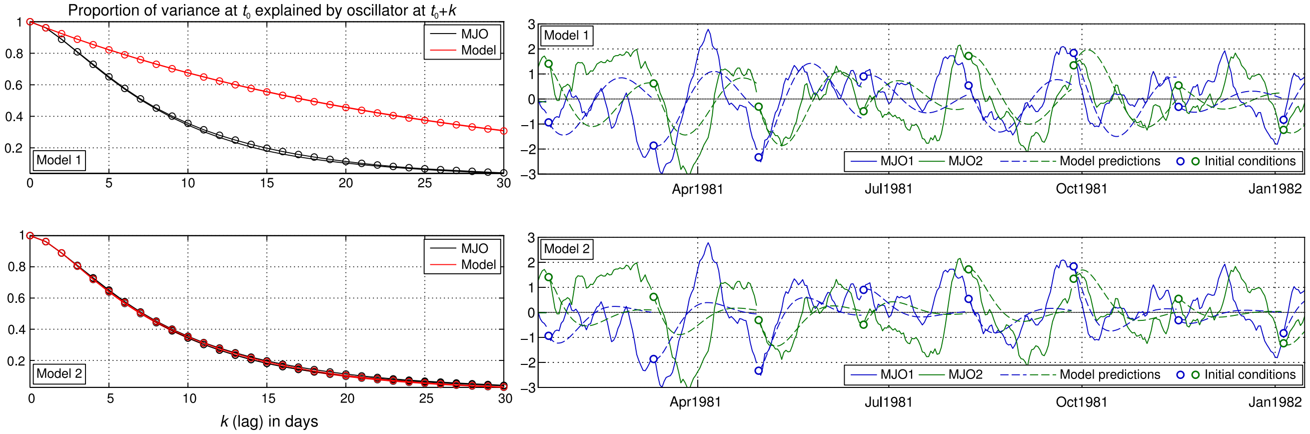

Fig. 8 Cross-validation results for Model 1 and Model 2. Left panels show the proportion of MJO index variance at time t_0 explained by the oscillator at time $t_0 + k$, for lags of $k = 0 \dots 30$, for Model 1 (*upper*) and Model 2 (*lower*). Results are shown for both validation periods; there are *two black lines* and *two red lines* in *each panel* but are so similar as to be difficult to distinguish. *Right panels* show subsets of results (over the validation period) using Model 1 (*upper*) and Model 2 (*lower*) as forecast models in which 50-day forecasts (*dashed lines*) were produced from a initial conditions (*circles*) every 50 days taken from the observed MJO index (*solid lines*)

Table 3 Cross-validation statistics for Model 1 and Model 2

m	1–50 days		1–15 days		15–30 days		31–45 days		Q_{MJO}	Q_1	Q_2
	$\sigma^2_{\mathrm{err},1}$	$\sigma^2_{\mathrm{err},2}$	$\sigma^2_{\mathrm{err},1}$	$\sigma^2_{\mathrm{err},2}$	$\sigma^2_{\mathrm{err},1}$	$\sigma^2_{\mathrm{err},2}$	$\sigma^2_{\mathrm{err},1}$	$\sigma^2_{\mathrm{err},2}$			
1	1.78	1.62	0.99	0.90	1.95	1.67	2.23	2.11	1.18	2.98	0.84
2	1.79	1.64	1.09	1.01	2.08	1.87	2.20	2.03	0.94	2.92	0.84

Results for each of the two iterations ($m = 1, 2$) of the twofold cross-validation procedure are shown by each row. The total variance of the forecast errors for Model i is denoted $\sigma^2_{\mathrm{err},i}$. The total variance is shown over the whole forecast period (1–50 days) and also over successive 15-day lead times (1–15 days, 16–30 days, and 31–45 days). The Q-factor (see text) for the MJO index (Model i) is denoted Q_{MJO} (Q_i). All statistics were calculated over the validation period

capture the curvature of the MJO index ACF near zero lag (not shown).

The models were compared using a twofold cross-validation procedure in order to choose a preferred model to use in the following predictability analysis. First both models were fit over the first half of the observed MJO index (the training period) and the properties of the fitted models were compared against those of the MJO index over the remaining period (the validation period). The procedure was then repeated with the training and validation periods swapped. Three properties were used to evaluate model performance over the validation period. First, a measure of the oscillator predictability was defined by calculating the total variance of the oscillator at time t_0 in the validation period that is explained by the oscillator at time $t_0 + k$ for lags of $k = 0 \dots 30$ (Fig. 8, left panels). It is clear from these results that Model 2 provides a much closer fit to the observed MJO index predictability than Model 1. It is also clear that Model 2 fits the observed predictability over the first two lags ($k = 1, 2$) while Model 1 only fits the first lag ($k = 1$). Second, Models 1 and Model 2 were used as forecast models in which a 50-day forecast was produced once every 50 days. All of the forecasts were then strung together as a time series and compared against the observed MJO index (Fig. 8, right panels). The initial condition for the forecasts were given by the observed MJO index immediately preceding the initial time of each forecast, simulating an operational system. The forecasts by Model 2 predicted the observed MJO index variability better than Model 1: the total variance of the forecast errors (observed index minus forecasts over all time, restricted to the validation period) were lower for Model 2 than for Model 1 (Table 3). By breaking down the total variance of the forecast errors into successive 15-day blocks of lead times it is clear that Model 2 outperforms Model 1 over all lead times but particularly over leads of 15–30 days and 31–45 days (Table 3). The difference in performance of Model 1 and Model 2 is much weaker over the first 15 days and the performance of model forecasts for lead times less than 5 days was nearly indistinguishable between Model 1 and Model 2. Third, we calculated the the Q-factor for the spectral peak of Model 1 and Model 2 and compared it against

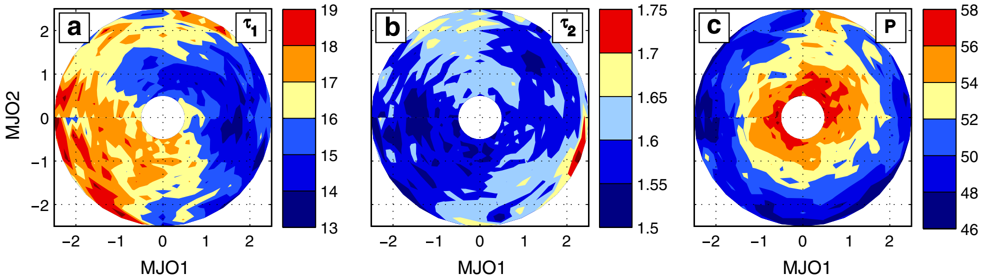

Fig. 9 Model 2 parameter time scales as a function of phase space. The model was fit to an ensemble of MJO events given an initial observation indicated by the position in phase space and an observation error of $\sigma^2_{\text{obs}} = 0.1$ (e.g., see Fig. 4 for four examples). The model time scales τ_1 (*left*), τ_2 (*centre*) and P (*right*), are shown in units of days

that for the observed index over the validation period. The Q-factor is a non-dimensional number defined as the ratio of the central frequency to the width of a spectral peak (defined as the full width at half maximum). The Q-factor has an interpretation in classical mechanics as the degree to which a harmonic oscillator is underdamped: an oscillator is underdamped for $Q > 0.5$ and critically damped or overdamped for $Q \leq 0.5$ (e.g., Marion and Thornton 1995). The observed MJO index is clearly underdamped with a Q-factor in the range 0.94–1.18 and Model 2 provides a much closer estimate of this Q-factor than Model 1 (Table 3).

The models were also compared using a likelihood-ratio test (e.g., Wilks 2011), the Akaike Information Criterion (AIC; Akaike 1974), and the Bayesian Information Criterion (BIC; Schwarz 1978). The likelihood-ratio test is based on the ratio of the likelihood functions for the two fitted models and expresses how much more likely the data are under one model or the other. Using the likelihood-ratio test the data is significantly more likely under Model 2 than Model 1 (with a p value of <0.01). The AIC and BIC are measures of relative model quality that allow a trade off between goodness-of-fit, which always improves with more free model parameters, and model complexity. When comparing two models, the one with the lower AIC or BIC values is to be favoured. The AIC and BIC values are lower for Model 2 (AIC = −55,806, BIC = −55,782) than for Model 1 (AIC = −49,480, BIC = −49,464).

Based on all the above results we conclude that Model 2 is the preferred model.

4.2 Model fitting conditioned on initial condition

Model 2 was fit to ensembles of MJO events, as described in Sect. 2.2, over a range of specified initial observations. Each ensemble consisted of 4000 observed trajectories (sampled with replacement) and each member trajectory spanned 35 days following the initial observation point. Initial observations were laid out in a grid across phase space with grid points in the amplitude direction from 0.5 to 2.5 (in steps of 0.1) and in the phase direction from 0.5 to 8.5 (in steps of 0.25). Note also that we have omitted the region within 0.5 amplitude of the origin from the analysis as these correspond to a very weak MJO which (i) cannot be expected to have much predictability, and (ii) is of little interest from the perspective of MJO dynamics and impacts. Note that an integer phase value corresponds to the angle which passes through the centre of the eight bins defined in Fig. 1c with decimal phases varying linearly between them. The model parameters were estimated by maximum likelihood estimation (see Appendix 2, which includes a discussion of the issue of fitting to relatively short trajectory lengths).

The estimated model parameters (τ_1, τ_2, and P) show considerable variation in phase space (Fig. 9a–c). The damping time scale τ_1 (Fig. 9a) is largest for phases 8 and 1–2 (τ_1 = 16–20 days) and lowest for phases 4–5 (τ_1 = 13–16 days). While there is a clear dependence of τ_1 on phase, its dependence on MJO amplitude is more complex, particularly in phases 3 (decreasing with amplitude) and 6–7 (increasing with amplitude). The autoregressive time scale τ_2 for the forcing (Fig. 9b) shows less variation in phase space, only varying between 1.5 and 1.75 days, being lowest in phases 8 and 1–2 (τ_2 = 1.5–1.6 days) and exceeding 1.75 days for the largest amplitudes in phase 4. The oscillation period P (Fig. 9c) shows very little dependence on phase but clearly decreases with MJO amplitude: it is larger for amplitudes less than one (P = 52–60 days) and smaller for larger amplitudes (P = 46–52 days). The only exception to this is in phase 6 where the dependence of P on MJO amplitude is very weak, never dropping below 52 days. The oscillator forcing magnitude, σ_ϵ, varies very little (~6 %; not shown) with lowest values in phases 6–8 (0.165–0.170) and largest values in phases 3–5 (0.170–0.177).

4.3 Predictability measures

The three oscillator statistics ($|\boldsymbol{\mu}_{t_0+k}|$, $\sigma^2_{t_0+k}$ and ρ_{t_0,t_0+k}, see Sect. 3.4) are plotted as a function of k for both the MJO index and Model 2 in Fig. 10 (solid and dashed lines

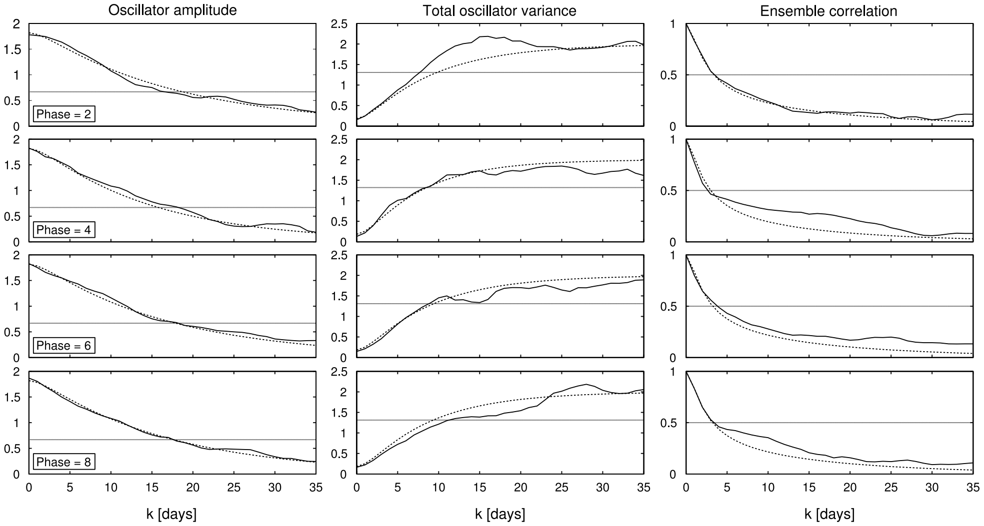

Fig. 10 Oscillator statistics for the MJO index and Model 2. The oscillator mean (*left*), total oscillator variance (*centre*), and within-ensemble correlation (*right*) are shown for both the MJO index (*solid lines*) and Model 2 (*dashed lines*). Ensembles are calculated given an initial observation of amplitude 2 and MJO phases 2 (first row), 4 (second row), 6 (third row), and 8 (fourth row) and with error $\sigma^2_{\mathrm{obs}} = 0.1$. The *horizontal lines* show the critical levels used to define the three predictability time scales that are read of the x-axis: τ_μ, τ_{σ^2}, and τ_ρ

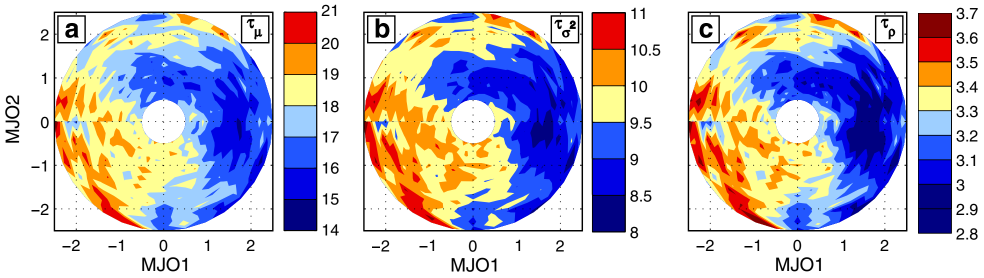

Fig. 11 Predictability time scales in phase space as a function of initial condition in phase space for Model 2. The time scales for (*left*) decay of the oscillator amplitude τ_μ, (*centre*) increase of total oscillator variance τ_{σ^2}, and (*right*) decay of the correlation τ_ρ are shown, in units of days, as a function of initial MJO value in phase space

respectively) for initial MJO amplitude 2 and phases 2, 4, 6 and 8. For each of the statistics, the initial values persist over a short period of 2–3 days before slowly relaxing to their asymptotic values (zero for the oscillator amplitude and correlation; $\sigma^2_\infty = 2$, which is the total variance of the normalized MJO index, for the total oscillator variance). Model 2 captures well the variation of the predictability statistics with lag and provides a smooth curve leading to more reliable estimates of predictability time scales. Given the critical values defined in Sect. 3.4 (Fig. 10, dashed lines) it is straightforward to read off the modeled predictability time scales τ_μ, τ_{σ^2}, and τ_ρ.

The predictability time scales show considerable variation in phase space (Fig. 11). The time scale for the decay of the oscillator mean τ_μ (Fig. 11a) is longest for phases 8 and 1–2 ($\tau_\mu = 17$–21 days). The value of τ_μ also reaches these levels in phases 6–7 for amplitudes greater than one. Elsewhere, τ_μ is in general lower (14–17 days). The time scales for increase of total oscillator variance τ_{σ^2} (Fig. 11b) and for the loss of correlation τ_ρ (Fig. 11c) exhibit the same phase space dependence as τ_μ with high values in phases 8 and 1–2 for all amplitudes and in phases 6–7 for large amplitudes ($\tau_{\sigma^2} = 9.5$–11 days; $\tau_\rho = 3.2$–3.7 days) and low values elsewhere ($\tau_{\sigma^2} = 8$–9.5 days; $\tau_\rho = 2.8$–3.2 days). For all three predictability time scales there is a minimum in phases 4–5.

The phase-space patterns of the three predictability time scales are very similar up to a scale factor (Fig. 11). In fact, a scatter plot of all estimated values of τ_{σ^2} and τ_ρ against all values of τ_μ indicates a linear relationship with slopes of 0.54 and 0.18 respectively (not shown). This is consistent with the relationships between the time scales estimated from the observed MJO index, which have slopes of 0.50 and 0.19 for τ_{σ^2} and τ_ρ, respectively, against τ_μ. This indicates that, although we have introduced three predictability time scales, they all measure the same thing, in a relative sense. It should also be noted that the phase space patterns of the three predictability time scales are strongly related to the pattern of τ_1. (A linear regression of the predictability time scales onto τ_1 over all values in phase space show a strong relationship ($R^2 > 0.97$ for all three time scales) indicating that the predictability time scales are dominated by τ_1.)

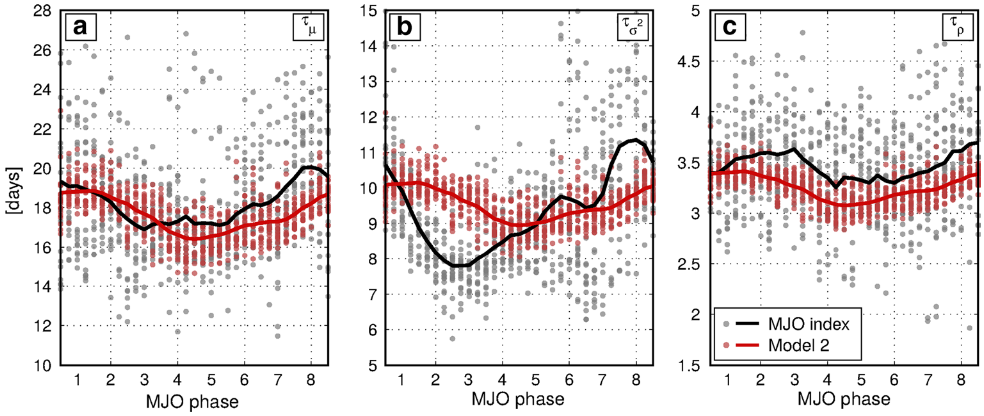

Fig. 12 Predictability time scales as a function of MJO phase. The time scales for (*left*) decay of the oscillator amplitude τ_μ, (*centre*) increase of total oscillator variance τ_{σ^2}, and (*right*) decay of the correlation τ_ρ are shown as a function of initial MJO phase for the MJO index (*grey dots*) and Model 2 (*red dots*), pooled over all MJO amplitudes. The average values across amplitudes at each phase are shown as *black* and *red lines* (MJO index and Model 2 respectively). The *black* and *red lines* have been smoothed with running average with a bandwidth of 1 MJO phase

Table 4 Estimated model parameters, in units of days, for Model 1 and Model 2 fit seasonally

	τ_1	τ_2	P
Austral summer			
Model 1	44.9	–	50.8
Model 2	18.6	1.16	49.5
Maharaj and Wheeler (2005)	44.3	–	49.8
Boreal summer			
Model 1	38.3	–	59.2
Model 2	15.3	1.20	58.3
Maharaj and Wheeler (2005)	34.7	–	61.4

Parameters estimated by maximum likelihood using the MJO index time series stratified by Austral Summer (NDJFMA) and Boreal Summer (MJJASO). The equivalent parameters from Maharaj and Wheeler (2005) are included for comparison with Model 1. Parameters are shown in units of days

The dependence of the predictability time scales on MJO phase alone was examined by pooling the estimates of τ_μ, τ_{σ^2}, and τ_ρ within each phase (Fig. 12, grey and red dots). This was performed for the time scales estimated from Model 2 (i.e., Fig. 11) and also directly from the observed MJO index (e.g., the intersection of the solid lines in Fig. 10 with the critical levels). The pattern of the predictability time scales τ_μ and τ_ρ as a function of phase shows a good correspondence between the model and the MJO index, with maximum predictability in phases 8 and 1–2 and minimum predictability in phases 4–5 (Fig. 12a, c). This pattern is also exhibited for the Model 2 estimates of τ_{σ^2} but the estimates from the observed MJO index differ slightly, with minimum values now in phases 2 and 3 (Fig. 12b). The estimated predictability times scales, averaged over all amplitudes and phases, from the Model 2 and observed MJO index are consistent: $\tau_\mu = 18.2$ days (MJO index) and 17.6 days (Model 2), $\tau_{\sigma^2} = 9.34$ days (MJO index) and 9.55 days (Model 2), and $\tau_\rho = 3.46$ days (MJO index) and 3.25 days (Model 2).

4.4 Seasonality

It is well-known that the variability of the MJO varies strongly by season. The MJO is dominant in Austral Summer and Fall during which variability is focused just south of the Equator (e.g., Zhang 2005). During Boreal Summer MJO variability is related to the Asian summer monsoon and is linked to the Boreal Summer Intraseasonal Oscillation (BSISO; e.g., Kikuchi et al. 2012; Lee et al. 2013) which consists of northward propagating intraseasonal anomalies. Previous studies demonstrated seasonal differences in the predictive skill of models at forecasting the MJO (Maharaj and Wheeler 2005; Kondrashov et al. 2013). Given this background we fit our damped harmonic oscillator model to determine if MJO predictability varies by season as well.

We first fit Model 1 and Model 2 to the observed MJO index for all phases and amplitudes, as in Sect. 4.1, but stratified by season. We define the seasons following Maharaj and Wheeler (2005): Austral Summer as November–April and Boreal Summer as May–October. Estimated model parameters are shown in Table 4. The parameters of Model 1 are consistent with those derived from the fit obtained by Maharaj and Wheeler (2005); P is shorter in Austral Summer and τ_1 is longer in Austral Summer. For Model 2, we see the same seasonal variation for P and τ_1 while the value of τ_2 does not vary greatly between seasons.

We then fit Model 2 to the observed MJO index conditioned on initial amplitude and phase, as in Sects. 4.2–4.3, but stratified by season. The fitted model was then

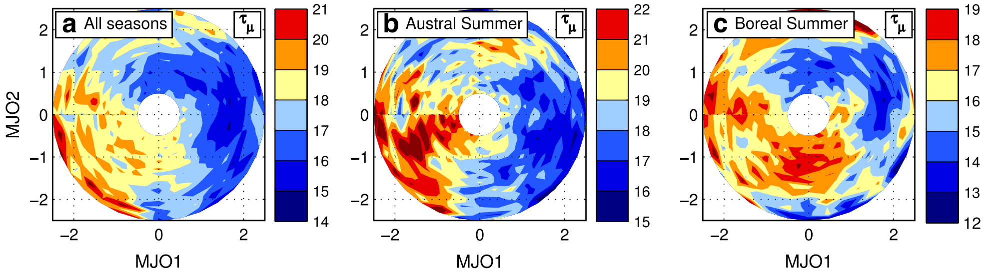

Fig. 13 Seasonally stratified predictability time scale τ_μ in phase space as a function of initial condition in phase space for Model 2. **a** Shows results for all seasons (same as Fig. 11a) while **b** and **c** show results for Austral Summer (NDJFMA) and Boreal Summer (MJJASO) respectively. The time scales are shown in units of days

used to estimate the predictability time scales τ_μ, τ_{σ^2}, and τ_ρ as a function of initial condition (Fig. 13). Here we present only τ_μ in order to illustrate seasonality. The seasonally stratified predictability time scales exhibit the same general pattern discussed previously: greater predictability in phases 8 and 1–2 and reduced predictability in phases 4–5. However, this pattern is rotated clockwise in Austral Summer and counter-clockwise in Boreal Summer by about 0.5 of an MJO phase. In other words, the timing of greatest predictability occurs earlier in the MJO cycle during Austral Summer but later during Boreal Summer. In addition, the time scales are generally larger in Austral Summer than in Boreal Summer indicating a seasonal cycle in the predictability of the MJO index. The value of τ_μ is generally 3–4 days longer in Austral Summer than in Boreal Summer (Fig. 13), τ_{σ^2} is 1–2 days longer (not shown), and τ_ρ is up to 1 day longer (not shown).

5 Summary and discussion

A forced damped harmonic oscillator (FDHO) model with autocorrelated forcing has been shown to reproduce the basic predictability features of the observed MJO index. First, it was shown that the evolution of the mean position of an ensemble of observed MJO events in phase space (given an initial observation of the MJO) evolves in a similar manner to a damped harmonic oscillator. Specifically, the ensemble mean undergoes a rotation in phase space and decay of amplitude. Additionally, the ensemble spread increases and the correlation between ensemble members decays over time (i.e., the mixing within an ensemble increases).

Motivated by this apparent damped oscillator behaviour of the observed MJO index we developed a pair of discrete time, autoregressive FDHO models for the MJO index: one with Gaussian white noise forcing (Model 1) and one with autoregressive forcing (Model 2). Model time series and phase space behaviour of the oscillator models were broadly consistent with the MJO index. Furthermore, the evolution of the model statistics (i.e., ensemble mean, variance and correlation) as well as the spectral properties (power spectral density, coherence and phase spectra) exhibit the same patterns as the MJO index.

The models were fit to the MJO index by maximum likelihood estimation. Model selection was performed by fitting both Model 1 and Model 2 to the complete MJO index time series and comparing the models by cross-validation and also using three selection criteria (a likelihood-ratio test, AIC, and BIC). It was determined that Model 2 provided a better fit to the data than Model 1. We then fit Model 2 to the MJO index for observations covering phase space in order to map out the predictability statistics of the MJO.

The predictability time scale of the oscillator depends on the choice of predictability statistics: oscillator mean, total oscillator variance, or correlation. Overall, our estimates of the predictability time scale based on the decay of the observed MJO amplitude ($\tau_\mu \simeq 14$–21 days) is broadly consistent with previous estimates (see Sect. 1). The time scale for the increase in total MJO variance (τ_{σ^2}) is typically about half of τ_μ ($\tau_{\sigma^2} \simeq 8$–11 days). The time scale for loss of correlation between individual members of an MJO ensemble (τ_ρ) is much shorter again with typical values in the range of 2.8–3.7 days. The predictability time scale found by Ding et al. (2010) using Nonlinear Local Lyapunov Exponents (NLLEs) would be similar to the time scale for our ensemble spread (σ^2) to reach 95 % of its saturation value. We do not calculate Lyapunov exponents but our approach of generating ensembles of MJO events which share the same initial condition is conceptually similar to the "local dynamic analog" used in the NLLE approach. In fact, a calculation of the time scale to 95 % saturation of σ^2 indicates values of 20–30 days (not shown) which was generally consistent with the value of 21 days found by found by Ding et al. (2010) for the Wheeler and Hendon (2004) MJO index.

The predictability time scales for the decay of the mean (τ_μ) and increase of the variance (τ_{σ^2}) vary considerably with initial MJO phase. In general, predictability is higher in phases 8 and 1–2. These phases correspond to periods when the region of active convection associated with the MJO is located over Africa and the western Indian Ocean, e.g., immediately preceding the development of an MJO event over the Indian Ocean. This may be indicative of an organized atmospheric state, preceding the development of an MJO event, which lends itself to a longer time scale for predictability as opposed to a possibly less organized atmospheric state during MJO propagation away from the generation region. It has been noted that prediction skill of the MJO in numerical models is often reduced when the initial condition consists of active MJO convection which must pass over the Maritime Continent (i.e., MJO phases 4–5, consistent with Kondrashov et al. (2013)), leading to the identification of the Maritime Continent as a potential "barrier to predictability" of the MJO (Vintzileos and Pan 2007; Seo et al. 2009; Weaver et al. 2011; Fu et al. 2011). Our results indicate that this barrier may be a property of the MJO itself rather than a failing of numerical prediction models, although our model indicates predictability is lowest for MJO events initialized in phases where anomalous convection is located over the Maritime Continent (phases 4–5) rather than immediately before (e.g., phase 2, Kim et al. 2014). Also, the fact that the time scale for the loss of correlation τ_ρ is much shorter than τ_μ indicates that, if performing an ensemble prediction of the MJO, beyond 3–4 days the ensemble mean and variance provide a better prediction than the individual ensemble members.

The predictability time scales were also found to vary with season. Greater predictability is found in Austral Summer when the MJO is strongest and most coherent. The values of τ_μ, τ_{σ^2}, and τ_ρ are 3–4 days, 1–2 days and up to 1 day longer, respectively, in Austral Summer than in Boreal Summer. We also found the Ding et al. (2010) equivalent time scale (i.e., time scale to 95 % saturation of the ensemble spread) was 4–5 days longer in Austral Summer than in Boreal Summer (not shown). In addition we found that the timing of greatest predictability occured about one MJO phase earlier in the MJO cycle during Austral Summer than Boreal Summer. Note that these are predictability time scales for the Wheeler and Hendon (2004) MJO index and not for the MJO itself. Therefore, these time scales should not be interpreted as the predictability time scales for tropical intraseasonal variability generally. In particular, during Boreal Summer predictability may be derived from the dynamics of the BSISO which are not captured by the Wheeler and Hendon (2004) MJO index.

The model captures several aspects of MJO predictability. First, estimates of the predictability time scales from the model fits are consistent with estimates directly from the observed MJO index. Second, the model captures the same relationship, in terms of relative magnitude, between τ_μ, τ_{σ^2}, and τ_ρ. Also, the best-fit model is able to capture the same dependence of τ_μ and τ_ρ on initial MJO position in phase space as is evident in the observations. The fact that model is able to reproduce the overall predictability dynamics, including oscillation, decay, time scales, and phase space variation, indicates that this model has value. This is remarkable considering the simplicity of the model and suggests it may serve as a simple but useful tool for understanding MJO predictability and as a benchmark against which to test more complex dynamical forecast systems.

Acknowledgments The authors would like to thank Adam Monahan of the School of Earth and Ocean Sciences (University of Victoria, Canada), Matthew Wheeler of the Centre for Australian Weather and Climate Research (Bureau of Meteorology, Australia), Simon Wotherspoon of the Institute for Marine and Antarctic Studies (University of Tasmania, Australia), and Ian Folkins of the Department of Physics and Atmospheric Science (Dalhousie University, Canada) for helpful comments and discussions. The authors would also like to acknowledge the two anonymous reviewers for strengthening the manuscript with their comments and suggestions. This paper makes a contribution to the objectives of the ARC Centre of Excellence for Climate System Science (ARCCCS).

Appendix 1: Calculating the initial condition distribution given an observation

Here we outline a Bayesian method for choosing an ensemble of MJO events, given an initial condition or observation. In the absence of more information, the scatter plot of all MJO index values (Fig. 3a, grey points) approximates the distribution of the MJO at some arbitrary time t_0 (e.g., Ristic et al. 2004). We will refer to this distribution as the initial condition prior (ICP) distribution (with zero mean and covariance $\sigma^2_{\mathrm{IC}}\mathbf{I}$, where $\mathbf{I}$ is the 2×2 identity matrix; note that $\sigma^2_{\mathrm{IC}} = 1$ since we have normalized the MJO index to have unity variance). Let $\boldsymbol{x}_{\mathrm{obs}}$ be an observation of the normalized MJO at time t_0, subject to observation error with zero mean and covariance $\sigma^2_{\mathrm{obs}}\mathrm{I}$ (the observation distribution is taken to be bivariate normal: proportional to $\exp(-|\boldsymbol{x}_{\mathrm{obs}} - \boldsymbol{x}|^2/2\sigma^2_{\mathrm{obs}})$ where $\boldsymbol{x}$ denotes the true state of the MJO at time t_0). We take a Bayesian approach whereby we use this observation to update the ICP distribution, yielding as the posterior distribution what we term the initial condition (IC) distribution. We approximate draws from the IC distribution by drawing samples, with replacement, from the set of observed points in the scatter plot discussed above and with weighting values proportional to the observation distribution. As $\sigma^2_{\mathrm{obs}} \to 0$ the IC distribution will reduce in spread and become centered more closely on $\boldsymbol{x}_{\mathrm{obs}}$; as $\sigma^2_{\mathrm{obs}} \to \infty$ the IC distribution will revert back to the ICP distribution. The mathematics details can be found in the following two subsections.

Observed MJO index and FDHO Model 1

Consider the observed MJO index (Sect. 2) or the FDHO Model 1 (Sect. 3.2) where the MJO index (or bivariate oscillator state) $\boldsymbol{x}_t$ at time t_0 is stored in the two dimensional state vector $\boldsymbol{x}_t$. We assume that the initial condition prior (ICP) distribution for the state at time t_0, $\boldsymbol{x}_{t_0}$, is normal with zero mean and covariance matrix $\boldsymbol{\Sigma}_{\mathrm{IC,prior}}$. ($\boldsymbol{\Sigma}_{\mathrm{IC,prior}} = \boldsymbol{\Sigma}_\infty$ for the MJO index and the models discussed in Sect. 3.)

Let an imperfect observation of the oscillator and forcing at time t_0 be denoted $\boldsymbol{x}_{\mathrm{obs}} = \boldsymbol{x}_{t_0} + \boldsymbol{\nu}$ where $\boldsymbol{\nu}$ is the observation error which is taken to be normally distributed with zero mean and covariance matrix $\boldsymbol{\Sigma}_{\mathrm{obs}} = \sigma_{\mathrm{obs}}\mathbf{I}$. In a Bayesian framework, the posterior distribution, or initial condition (IC) distribution, of the oscillator state given this observation and the ICP distribution is normal with mean and covariance

$$\boldsymbol{\mu}_{\mathrm{IC}} = \boldsymbol{\Sigma}_{\mathrm{IC,prior}}\left(\boldsymbol{\Sigma}_{\mathrm{IC,prior}} + \boldsymbol{\Sigma}_{\mathrm{obs}}\right)^{-1}\boldsymbol{\mu}_{\mathrm{obs}} \tag{17}$$

$$\boldsymbol{\Sigma}_{\mathrm{IC}} = \left(\boldsymbol{\Sigma}_{\mathrm{IC,prior}}^{-1} + \boldsymbol{\Sigma}_{\mathrm{obs}}^{-1}\right)^{-1}, \tag{18}$$

which follows directly from the product of two normal probability distributions (the observation and ICP distributions).

The exact forms of $\boldsymbol{\mu}_{\mathrm{IC}}$ and $\boldsymbol{\Sigma}_{\mathrm{IC}}$ follow from Eqs. 17 and 18 in a straightforward manner to give

$$\boldsymbol{\mu}_{\mathrm{IC}} = \frac{1}{1+\sigma_{\mathrm{obs}}^2}\boldsymbol{x}_{\mathrm{obs}} \tag{19}$$

$$\boldsymbol{\Sigma}_{\mathrm{IC}} = \frac{\sigma_{\mathrm{obs}}^2}{1+\sigma_{\mathrm{obs}}^2}\boldsymbol{\Sigma}_{\mathrm{IC,prior}}. \tag{20}$$

Note that as $\sigma_{\mathrm{obs}}^2 \to \infty$ the mean and covariance of the IC distribution revert to the ICP (prior) mean and covariance as expected; as $\sigma_{\mathrm{obs}}^2 \to 0$ the mean of the IC distribution is simply the observation itself and the covariance is zero, indicating a perfectly known observation.

FDHO Model 2

Now consider Model 2 as discussed in Sect. 3.3 where the bivariate oscillator state $\boldsymbol{x}_t$ and bivariate autoregressive forcing $\boldsymbol{f}_t$, at time t_0, are stored in the first and last pair of elements, respectively, of the four dimensional state vector $\boldsymbol{y}_t$. Again we assume that the initial condition prior (ICP) distribution for the state at time t_0, $\boldsymbol{y}_{t_0}$, is normal with zero mean and covariance matrix $\boldsymbol{\Sigma}_{\mathrm{IC,prior}}$.

Let an imperfect observation of the oscillator and forcing at time t_0 be denoted $\boldsymbol{y}_{\mathrm{obs}} = \boldsymbol{y}_{t_0} + \boldsymbol{\nu}$ where $\boldsymbol{\nu}$ is the observation error which is taken to be normally distributed with zero mean and covariance matrix $\boldsymbol{\Sigma}_{\mathrm{obs}} = \sigma_{\mathrm{obs}}\mathbf{I}$. In general, we are able to observe the oscillator (e.g., the MJO index) but not the forcing and so we must make some assumptions about the forcing observation. We assume the mean and covariance matrix of the observation distribution are of the form

$$\boldsymbol{y}_{\mathrm{obs}} = \begin{bmatrix} \boldsymbol{x}_{\mathrm{obs}} \\ \boldsymbol{f}_{\mathrm{obs}} \end{bmatrix} \boldsymbol{\Sigma}_{\mathrm{obs}} = \begin{bmatrix} \sigma_{\mathrm{obs}}^2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \alpha\mathbf{I} \end{bmatrix}, \tag{21}$$

where α represents the variance of the uncertainty on our observation of the forcing. We consider the case in which the observation error of the forcing is very large ($\alpha \to \infty$), implying that we only have observational information on the oscillator components and not on the associated forcing components. Can we write an expression for the IC distribution, including the forcing components, despite not observing the forcing?

As the uncertainty on the forcing observation becomes very large (i.e., in the limit $\alpha \to \infty$) the mean and covariance of the IC distribution is given by (using Eqs. 17 and 18)

$$\boldsymbol{\mu}_{\mathrm{IC}} = \frac{1}{1+\sigma_{\mathrm{obs}}^2}\begin{bmatrix} I \\ \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \end{bmatrix}\boldsymbol{x}_{\mathrm{obs}}, \tag{22}$$

$$\boldsymbol{\Sigma}_{\mathrm{IC}} = \frac{1}{1+\sigma_{\mathrm{obs}}^2}\left[\sigma_{\mathrm{obs}}^2\boldsymbol{\Sigma}_{\mathrm{IC,prior}} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12} \end{bmatrix}\right], \tag{23}$$

where $\boldsymbol{\Sigma}_{ij}$ refers to the ij^{th} block of $\boldsymbol{\Sigma}_{\mathrm{IC,prior}}$ when arranged in four 2×2 blocks. Now, as $\sigma_{\mathrm{obs}}^2 \to 0$ the mean of the IC distribution is simply the observation itself and an estimate of the mean forcing comes from a linear regression of the forcing components onto the oscillator components (e.g., Johnson and Wichern 2002); similarly, the posterior covariance of the forcing is simply given by the variance of the residuals from a linear regression of the forcing components onto the oscillator components of $\boldsymbol{y}_{t_0}$. Therefore, an observation of the oscillator components, in the absence of an associated observation of the forcing, can be used along with the ICP distribution to infer the mean of the forcing.

Appendix 2: Likelihood functions for Model 1 and Model 2

In this appendix we present the likelihood functions used to estimate the parameters of Model 1 and Model 2, including a generalization for ensembles of MJO events, that are selected based on their proximity to a given initial condition (Sect. 4). The method is a bivariate generalization of the method presented for a univariate process by Priestley (1981). First, consider a bivariate Gaussian white noise error process $\{\boldsymbol{\epsilon}_t\}$. The errors are independent and identically distributed with an assumed mean of zero and covariance matrix $\sigma_\epsilon^2\mathbf{I}$. It follows that the probability density of $\boldsymbol{\epsilon}_t$ is given by

$$p(\boldsymbol{\epsilon}_t) = \frac{1}{2\pi\sigma_\epsilon^2}\exp\left(-\frac{1}{2\sigma_\epsilon^2}\boldsymbol{\epsilon}_t'\boldsymbol{\epsilon}_t\right), \qquad (24)$$

where $'$ denotes a transpose, and $\boldsymbol{\epsilon}_t'\boldsymbol{\epsilon}_t$ is simply the sum of squares of the two elements of $\boldsymbol{\epsilon}_t$ (i.e., $\epsilon_{1,t}^2 + \epsilon_{2,t}^2$). The joint probability for $\boldsymbol{\epsilon}_t$ for $t = 1, 2 \ldots N$ is given by

$$\begin{aligned} p(\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2 \ldots \boldsymbol{\epsilon}_N) &= \prod_{t=1}^{N} p(\boldsymbol{\epsilon}_t) \\ &= \left(\frac{1}{2\pi\sigma_\epsilon^2}\right)^N \exp\left(-\frac{1}{2\sigma_\epsilon^2}\sum_{t=1}^{N}\boldsymbol{\epsilon}_t'\boldsymbol{\epsilon}_t\right). \qquad (25) \end{aligned}$$

A single time series

Model 1 is a bivariate AR(1) process (see Sect. 3.2) which can be rewritten $\boldsymbol{x}_t - \boldsymbol{A}_1\boldsymbol{x}_{t-1} = \boldsymbol{\epsilon}_t$ where $\boldsymbol{\epsilon}_t$ is a bivariate Gaussian white noise error process, $\boldsymbol{A}_1 = \boldsymbol{A}_1(\boldsymbol{\lambda})$ and $\boldsymbol{\lambda} = (\tau_1, P, \sigma_\epsilon)$ is the vector of unknown model parameters. Note that given observations $\boldsymbol{x}_1$ to $\boldsymbol{x}_N$ we can define

$$Q(\boldsymbol{x}_1, \boldsymbol{x}_2 \ldots \boldsymbol{x}_N, \boldsymbol{\lambda}) = \sum_{t=2}^{N}\boldsymbol{\epsilon}_t'\boldsymbol{\epsilon}_t = \sum_{t=2}^{N}(\boldsymbol{x}_t - \boldsymbol{A}_1\boldsymbol{x}_{t-1})'(\boldsymbol{x}_t - \boldsymbol{A}_1\boldsymbol{x}_{t-1}). \qquad (26)$$

The joint probability distribution of $\boldsymbol{x}_2$ to $\boldsymbol{x}_N$, conditional on the value of $\boldsymbol{x}_1$, is then given by

$$\begin{aligned} &p(\boldsymbol{x}_2, \boldsymbol{x}_3 \ldots \boldsymbol{x}_N | \boldsymbol{\lambda}, \boldsymbol{x}_1) \\ &\quad = \left(\frac{1}{2\pi\sigma_\epsilon^2}\right)^{N-1} \exp\left(-\frac{1}{2\sigma_\epsilon^2}Q(\boldsymbol{x}_1, \boldsymbol{x}_2 \ldots \boldsymbol{x}_N, \boldsymbol{\lambda})\right). \qquad (27) \end{aligned}$$

To estimate the unknown parameter vector lambda we maximize Eq. 27 in the usual way. Note that the effect of conditioning on $\boldsymbol{x}_1$ will be small if N is large enough (Priestley 1981). More discussion of this point is provided in the next subsection.

Model 2 is a quadrivariate AR(1) process (see Sect. 3.3). This process can be rewritten as a bivariate AR(2) process and expressed as $\boldsymbol{x}_t - (\boldsymbol{A}_1 + \boldsymbol{A}_2)\boldsymbol{x}_{t-1} + \boldsymbol{A}_2\boldsymbol{A}_1 = \boldsymbol{\epsilon}_t$ where $\boldsymbol{\epsilon}_t$ is a bivariate Gaussian white noise error process and $\boldsymbol{A}_2 = \boldsymbol{A}_2(\tau_2)$. It is then straightforward to follow the same procedure as above and (i) define a probability density function for $\boldsymbol{x}_3$ to $\boldsymbol{x}_N$, conditioned now on $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, and (ii) minimize the associated likelihood function with respect to the parameter vector $\boldsymbol{\lambda} = (\tau_1, \tau_2, P, \sigma_\epsilon)$.

An ensemble of trajectories

Consider now an ensemble of MJO trajectories that share an initial condition in phase space. The ensemble consists of M trajectories, each one defined over K time steps following the initial condition at time t_0. Each trajectory is thus of length $K + 1$. The log-likelihood function for an individual trajectory is equivalent to the single time series case (previous subsection), after replacing N with $K + 1$. The log-likelihood function for the ensemble of trajectories is given by summing the log-likelihood functions of the M individual trajectories which comprise the ensemble. As above, maximum likelihood estimation is used to estimate the model parameters. Concering the approximation used above for large N, or large K in the present case, we found that the log-likelihood function for the ensemble converges to a stable value for $K \geq 35$. This has informed our choice of $K = 35$ for the ensemble length used in Sect. 4.2.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723

Cavanaugh NR, Allen T, Subramanian A, Mapes B, Seo H, Miller AJ (2014) The skill of atmospheric linear inverse models in hindcasting the Madden–Julian Oscillation. Clim Dyn 44:897–906

Ding R, Li J, Seo KH (2010) Predictability of the Madden–Julian Oscillation estimated using observational data. Mon Weather Rev 138(3):1004–1013

Fu X, Wang B, Lee J-Y, Wang W, Gao L (2011) Sensitivity of dynamical intraseasonal prediction skills to different initial conditions. Mon Weather Rev 139:2572–2592

Harvey A (1991) Forecasting, structural time series models and the Kalman filter. Cambridge University Press, Cambridge

Jiang X, Waliser D, Wheeler M, Jones C, Lee M, Schubert S (2008) Assessing the skill of an all-season statistical forecast model for the Madden–Julian Oscillation. Mon Weather Rev 136:1940–1956

Johnson R, Wichern D (2002) Appl Multivar Stat Anal. Prentice Hall, Upper Saddle River

Kang I, Kim H (2010) Assessment of MJO predictability for Boreal Winter with various statistical and dynamical models. J Clim 23:2368–2378

Kikuchi K, Wang B, Kajikawa Y (2012) Bimodal representation of the tropical intraseasonal oscillation. Clim Dyn 38:1989–2000

Kim HM, Hoyos CD, Webster PJ, Kang IS (2009) Ocean-atmosphere coupling and the boreal winter MJO. Clim Dyn 35:771–784

Kim HM, Webster PJ, Toma VE, Kim D (2014) Predictability and prediction skill of the MJO in two operational forecasting systems. J Clim 27:5364–5378

Kondrashov D, Chekroun MD, Robertson AW, Ghil M (2013) Low-order stochastic model and past-noise forecasting of the Madden–Julian Oscillation. Geophys Res Lett 40(19):5305–5310

Lee JY, Wang B, Wheeler M, Fu X, Waliser D, Kang IS (2013) Real-time Multivariate Indices for the boreal summer intraseasonal oscillation over the Asian summer monsoon region, vol 40

Lin H, Brunet G, Derome J (2008) Forecast skill of the Madden–Julian Oscillation in two Canadian atmospheric models. Mon Weather Rev 136:4130–4149

Lo F, Hendon H (2000) Empirical extended-range prediction of the Madden–Julian Oscillation. Mon Weather Rev 128(7):2528–2543

Madden R, Julian P (1971) Detection of a 40–50 day oscillation in the zonal wind in the Tropical Pacific. J Atmosp Sci 28(5):702–708

Madden R, Julian P (1972) Description of global-scale circulation cells in the tropics with a 40–50 day period. J Atmosp Sci 29(6):1109–1123

Madden R, Julian P (1994) Observations of the 40–50-day tropical oscillation—a review. Mon Weather Rev 122(5):814–837

Maharaj E, Wheeler M (2005) Forecasting an index of the Madden–Julian Oscillation. Int J Climatol 25:1611–1618

Marion J, Thornton S (1995) Classical mechanics of particles and systems. Holt Rinehart & Winston, New York

Neena JM, Lee JY, Waliser D, Wang B, Jiang X (2014) Predictability of the Madden–Julian Oscillation in the Intraseasonal Variability Hindcast Experiment (ISVHE). J Clim 27:4531–4543

Oliver ECJ (2011) Ph.D. Thesis: local and remote forcing of the ocean by the Madden–Julian Oscillation and its predictability. Dalhousie University, Halifax, Canada

Oliver ECJ, Thompson KR (2012) A reconstruction of Madden–Julian Oscillation variability from 1905 to 2008. J Clim 25(6):1996–2019

Priestley M (1981) Spectral analysis and time series. New York

Rashid HA, Hendon HH, Wheeler MC, Alves O (2011) Prediction of the Madden–Julian Oscillation with the POAMA dynamical prediction system. Clim Dyn 36(3–4):649–661

Ristic B, Arulampalam S, Gordon N (2004) Beyond the Kalman filter: particle filters for tracking applications. Artech House Publishers, London

Schwarz G et al (1978) Estimating the dimension of a model. Ann Stat 6(2):461–464

Seo K-H, Wang W, Gottschalck J, Zhang Q, Schemm J-KE, Higgins WR, Kumar A (2009) Evaluation of MJO forecast skill from several statistical and dynamical forecast models. J Clim 22:2372–2388

Vintzileos A, Pan HL (2007) On the importance of horizontal resolution and initial conditions to forecasting tropical intraseasonal oscillations: the Maritime Continent prediction barrier. In: Extended abstracts, NOAA CTB-COLA joint seminar series

von Storch H, Baumhefner DP (1991) Principal oscillation pattern analysis of the tropical 30-to 60-day oscillation. Part II: the prediction of equatorial velocity potential and its skill. Clim Dyn 6(1):1–12

von Storch H, Xu J (1990a) Principal oscillation pattern analysis of the 30-to 60-day oscillation in the tropical troposphere. Clim Dyn 4(3):175–190

von Storch H, Xu J (1990b) Principal oscillation pattern analysis of the 30-to 60-day oscillation in the tropical troposphere. Part I: definition of an index and its prediction. Clim Dyn 4(3):175–190

von Storch H, Bruns T, Fischer-Bruns I, Hasselmann K (1988) Principal oscillation pattern analysis of the 30-to 60-day oscillation in general circulation model equatorial troposphere. J Geophys Res 93(D9):11,022–11,036

Waliser D (2012) Intraseasonal Variability in the atmosphere-ocean climate system, chap predictability and forecasting. 2nd edn. Springer, Berlin, pp 433–476

Waliser DE, Jones C, Schemm JKE, Graham NE (1999) A statistical extended-range tropical forecast model based on the slow evolution of the Madden–Julian Oscillation. J Clim 12(7):1918–1939

Weaver SJ, Wang W, Chen M, Kumar A (2011) Representation of MJO variability in the NCEP Climate Forecast System. J Clim 24:4676–4694

Wheeler M, Hendon H (2004) An all-season real-time multivariate MJO index: development of an index for monitoring and prediction. Mon Weather Rev 132(8):1917–1932

Wilks DS (2011) Statistical methods in the atmospheric sciences. Academic Press, New York

Zhang C (2005) Madden–Julian Oscillation. Rev Geophys 43:1–36

Zhang C (2013) Madden–Julian Oscillation: bridging weather and climate. Bull Am Meteorol Soc 94(12):1849–1870

Zhang C, Gottschalck J, Maloney ED, Moncrieff MW, Vitart F, Waliser DE, Wang B, Wheeler MC (2013) Cracking the MJO nut. Geophys Res Lett 40(6):1223–1230