Multi-model Communication and Data Assimilation for Mitigating Model Error and Improving Forecasts

Yian CHEN[1]   Samuel N. STECHMANN[2]

(*Dedicated to Professor Andrew J. Majda on the occasion of his 70th birthday*)

Abstract Models for weather and climate prediction are complex, and each model typically has at least a small number of phenomena that are poorly represented, such as perhaps the Madden–Julian Oscillation (MJO) or El Niño–Southern Oscillation (ENSO) or sea ice. Furthermore, it is often a very challenging task to modify and improve a complex model without creating new deficiencies. On the other hand, it is sometimes possible to design a low-dimensional model for a particular phenomenon, such as the MJO or ENSO, with significant skill, although the model may not represent the dynamics of the full weather–climate system. Here a strategy is proposed to mitigate these model errors by taking advantage of each model's strengths. The strategy involves inter-model data assimilation, during a forecast simulation, whereby models can exchange information in order to obtain more faithful representations of the full weather–climate system. As an initial investigation, the method is examined here using a simplified scenario of linear models, involving a system of stochastic partial differential equations (SPDEs) as an imperfect tropical climate model and stochastic differential equations (SDEs) as a low-dimensional model for the MJO. It is shown that the MJO prediction skill of the imperfect climate model can be enhanced to equal the predictive skill of the low-dimensional model. Such an approach could provide a route to improving global model forecasts in a minimally invasive way, with modifications to the prediction system but without modifying the complex global physical model itself.

Keywords MJO, Multi-model communication, Data assimilation, Kalman filter algorithm
2000 MR Subject Classification 60H10, 60H15, 86A10, 62M20

1 Introduction

In this paper, in brief, we investigate an approach to weather and climate forecasting that uses multiple dynamical models. Such an approach becomes more relevant as prediction models have advanced from being local, regional models to being global models, since global models aim to represent the entire weather–climate system as a whole, but doing this accurately with a single model is difficult.

[1]Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706, USA. E-mail: chen689@wisc.edu
Current Affiliation: Department of Statistics, University of Chicago, Chicago, IL 60637, USA. Email: yianc@uchicago.edu
[2]Department of Mathematics and Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison, WI 53706, USA. E-mail: stechmann@wisc.edu

Furthermore, predictions are advancing beyond the short-term forecasts of only a few days, and now a major challenge is to make forecasts at the intersection of weather and climate, at lead times of weeks, months, or seasons in advance [78, 21, 39, 97, 69, 54, 95, 50, 5, 82, 94]. Such subseasonal to seasonal (S2S) predictions are related to some of the strongest signals of intraseasonal and interannual variability, such as monsoons, the Madden–Julian Oscillation (MJO), El Niño–Southern Oscillation (ENSO), sea ice, etc., and these phenomena can be global themselves or linked globally through various teleconnections.

While global models such as global climate models (GCMs) are powerful tools for making predictions, they can have significant deficiencies, and the deficiencies cause prediction skills that are below the estimated predictability limits. For instance, GCMs commonly have deficient representations of the MJO [84, 58, 34, 36], and their predictions of the MJO are likewise below the estimated theoretical predictability [99, 97, 25, 95, 49, 72]. Other phenomena that have deficient representations in GCMs include the intertropical convergence zone (ITCZ) (for which there is a serious double-ITCZ problem) [57, 35] and the El Niño–Southern Oscillation (ENSO) [37, 27].

Improving a complex dynamical model such as a GCM is a very challenging task. Sometimes improving one aspect could cause a worsening of another aspect of the simulation. For example, such an issue has been noticed in attempts to improve the MJO in GCMs, where improvements in the MJO have sometimes been seen to be accompanied by a worsened climatological mean state [48, 1]. Such attempts to modify GCMs to improve their simulations of the MJO are undoubtedly worthwhile. Nevertheless, the difficulties of such a challenging task can also be taken as motivation for alternative methods for simulating and predicting the MJO (as well as other challenging phenomena such as ENSO and sea ice).

Besides global models such as GCMs, an alternative approach to forecasting is to use simple models. For example, the MJO could be crudely represented by two indices, $x_1(t)$ and $x_2(t)$, which represent the amplitudes of the two empirical orthogonal functions (EOFs) of the MJO's structure from principal component analysis (PCA) [102]. Then stochastic/statistical models can be used for predicting these types of indices [10, 9]. Other types of simple models might incorporate aspects of atmospheric and/or oceanic fluid dynamics, thereby providing some added physical realism while still maintaining a simplified modeling approach with, say, $O(100)$ degrees of freedom; examples of this type include the skeleton model for the MJO [65, 92, 93, 74, 85, 11, 12, 13, 76] and models for ENSO [108, 91]. Despite the simplicity of these types of models, they can often produce forecasts with significant skill. In fact, forecast skill with simple models sometimes meets or exceeds the skill of forecasts with comprehensive physical dynamical models such as GCMs [98, 38, 104, 95, 10, 9]. Such results have been seen for not only the MJO but also other aspects of weather/climate variability, such as ENSO [80, 81] and sea ice [59, 100, 7].

Given the background described above, the main goal of the present paper is to investigate a multi-model communication strategy for forecasting with deficient models. As a concrete example, the case of the MJO will be used (although the general idea should be applicable to other scenarios described above as well). It is assumed that two models are available: (i) a deficient tropical climate model, which is meant to be an adequate model of climate but deficient in its representation of climate variability (the MJO), and (ii) a low-order model for the MJO,

which has no representation at all of the rest of the weather–climate system. The goal is to utilize both models together to achieve the best possible forecast of tropical weather/climate variability, and to do so in a way that will emphasize each model's strengths and mitigate each model's deficiencies.

To carry out the multi-model communication strategy, the models will be allowed to assimilate data from each other while in forecast mode. In other words, each model will be run forward in time for a forecast, and occasionally each model will consider the other model's forecast to be providing an "observation," and this data will be assimilated. Schematically, the strategy could be described for two models as follows:

$$
\begin{array}{ccc}
u_0 & & v_0 \\
\downarrow & & \downarrow \\
u_1 = Fu_0 + \sigma_1 & \longleftrightarrow & v_1 = \tilde{F}v_0 + \tilde{\sigma}_1 \\
\downarrow & & \downarrow \\
u_2 = Fu_1 + \sigma_2 & \longleftrightarrow & v_2 = \tilde{F}v_1 + \tilde{\sigma}_2 \\
\downarrow & & \downarrow \\
\vdots & & \vdots
\end{array}
$$

Here, u and v represent the state vectors of the two models, F and $\tilde{F}$ represent the dynamical operators, σ and $\tilde{\sigma}$ represent (potentially random) forcing, the subscripts $(0, 1, 2, \cdots)$ represent points in time, the downward arrows represent evolution in time, and the horizontal arrows indicate the exchange of information via data assimilation.

For the specific scenario described above, the deficient tropical climate model will occasionally assimilate information about the MJO from the low-order MJO model forecast. It is hoped that the assimilated forecast data will help the deficient climate model to, e.g., maintain and propagate an MJO with an enhanced skill that would otherwise be impossible if the climate model were run on its own. Such a strategy is plausible given that reanalysis products [40, e.g.] (which assimilate actual observational data into deficient climate models in order to provide a best estimate of the actual atmospheric state and its evolution) are commonly assumed to have reasonable (although not perfect) representations of the MJO and used to study its behavior [47, e.g.], whereas free-running climate models that do not assimilate observational data are commonly seen to have deficient MJOs [84, 58, 34, 36]. Here, it will be the low-order MJO model forecast that provides the "observations" of the MJO that could potentially drive a reasonable MJO in the deficient climate model's forecast. On the other hand, in the opposite direction, one can imagine that the low-order MJO model would benefit from assimilating data from the climate model, which could provide information about the rest of the climate system and its influence on the MJO.

If such a strategy could be successful, it could allow one to leverage advancements in simplified models and predictions [65, 10, 92, 9, 74, 8, 76] to improve predictions with comprehensive global models. Furthermore, such an approach will potentially allow improvements in a minimally invasive way, since it is the prediction system that would be modified rather than the dynamics of the climate model itself.

Several simplifications will be utilized in the present paper. The main simplification will be

the use of linear, Gaussian models for both the deficient climate model and the MJO model. On the one hand, such a simplification is crude; on the other hand, it allows the use of the Kalman filter for the inter-model data assimilation, which will help provide a more transparent view of the workings of the methods here. Another simplification is that the deficient tropical climate model will use a crude vertical structure and no meridional structure (so it represents variability directly above the equator at $y = 0$). On the one hand, such a simplification will restrict the variety of modes of climate variability; on the other hand, this restriction can be taken as one way in which this climate model is deficient.

Another approach that involves the use of multiple models is the multi-model ensemble mean, which has been in use for some time [51, 53, 52, 42, 28, 103, 17], but it involves multiple models only in post-processing, not through any type of direct communication between models. In allowing the coupling of multiple models, the multi-model communication approach is, to some extent, reminiscent of the topic of synchronization of chaotic dynamical systems [16, 106] and the approach of superparameterization methods which couple the dynamics of macro-scale and micro-scale models [26, 41]. Also, there has been some use of artificial intelligence, machine learning, neural networks, multifidelity methods [79], and other data science methods for weather and climate applications [68, 83]. To an extent, the method of the present paper could be viewed as a type of machine learning, where one model learns of some components of the weather–climate system from another model. While this idea is implemented here through a data assimilation algorithm, it is possible that other machine learning techniques could potentially be investigated for similar purposes in the future.

This article is dedicated to Andrew J. Majda to celebrate his 70th birthday, and it has drawn on his influence through his many contributions to topics of models for tropical weather and climate variability [61, 2, 44, 45, 65, 4, 105], new convective parameterizations for GCMs [46, 14, 15, 22, 23, 24], data assimilation and filtering [62, 6, 29, 30, 31, 20, 19, 64, 32, 8, 55], predictions of tropical weather and climate variability [10, 9], and multi-model ensemble strategies [3], to name only a few.

The paper is organized as follows. The observational data and data analysis methods are described in Section 2. The two models—the deficient tropical climate model and the simple MJO model—are described in Section 3 and Section 4, respectively, including descriptions of the models themselves as well as their individual skill as models for data assimilation and forecasting. Finally, the model communication strategy is then investigated in Section 5, followed by conclusions in Section 6. Many details of the methods are presented in the appendices in order to streamline the presentation in the main text.

2 Data Source and Data Processing Methodology

Our tropical climate model (which will be presented below in Section 3) involves three variables: wind (zonal velocity) u, potential temperature θ, and a variable q that represents moisture or water vapor or cloudiness. To assess the model's skill, observation-based estimates of the three model variables are needed. This section provides a brief overview of the data sources and data processing methodology, and a more detailed description is given in Appendix 7.

Two data sources are used here. First, the National Oceanic and Atmospheric Administra-

tion (NOAA) interpolated outgoing longwave radiation (OLR) is used [56]. OLR is commonly used as a surrogate for cloudiness, and it is also somewhat well correlated with moisture anomalies [47, e.g.]. For the purposes of the present paper and its simple and deficient model, the distinction between moisture and cloudiness is not a major consideration, although in other frameworks, such as the MJO skeleton model, such a distinction is important [65, 66]. Second, National Centers for Environmental Prediction-National Center for Atmospheric Research (NCEP-NCAR) reanalysis daily zonal winds and geopotential height are used [40]. Despite the fact that reanalysis data are a combined product of observational data and model dynamics, we sometimes refer to it as observational for simplicity. Both datasets have a horizontal spatial resolution of $2.5° \times 2.5°$ and a daily temporal resolution from 1 January 1979 to 31 December 2011. Since the connections between the model variables and observations are largely based on previous work [90, 88, 73, 75], further details are relegated to Appendix 7.

Here the influence of the seasonal cycle is removed via subtracting the time mean and first three harmonics of the annual cycle, based on the 1979-2011 observation record. And to further remove low-frequency variability, an 120-day-mean of the previous 120 days is subtracted, as is recommended by the Climate Variability and Predictability [96].

An example of the observed evolution of the atmosphere is shown in Figure 1. Inspection of Figure 1 reveals some features occurring in MJO, especially the typical eastward-propagating waves, which are most visible in the wind u and also in the moisture/cloudiness q.

3 Deficient Tropical Climate Model

Our main goal of this paper is to illustrate an idea for multi-model communication, using inter-model data assimilation, as a strategy to improve forecasts. The improvements will come from mitigating the effects of model errors. As an illustration in a simplified setting, we utilize a simplified tropical climate model that has some significant deficiencies (Section 3.1). We intend to treat it as a complex climate model that can simulate the outlines of the climate but is deficient in its description of some important climate phenomena, such as the MJO. To finish the section, we describe the model's free-running climate and variability (Section 3.2) as well as the model's performance in data assimilation and prediction (Section 3.3 and Section 3.4, respectively). In later sections, we will use this deficient climate model as part of the multi-model communication strategy to improve its forecasting performance.

3.1 Model description

The deficient tropical climate model (DTCM) is a three-variable system of stochastic partial differential equations (SPDEs) with the following form:

$$\frac{\partial u}{\partial t} - \frac{\partial \theta}{\partial x} = -\frac{1}{\tau_u} u \tag{3.1}$$

$$\frac{\partial \theta}{\partial t} - \frac{\partial u}{\partial x} = \frac{1}{\tau_c} q - \frac{1}{\tau_\theta} \theta \tag{3.2}$$

$$\frac{\partial q}{\partial t} + \widetilde{Q} \frac{\partial u}{\partial x} = -\frac{1}{\tau_c} q + b \frac{\partial^2 q}{\partial x^2} + D_* \dot{W} \tag{3.3}$$

where u is the zonal wind velocity anomaly, θ is the potential temperature, q is the water vapor anomaly and $\dot{W}$ is spatiotemporal white noise. The equations have been nondimensionalized
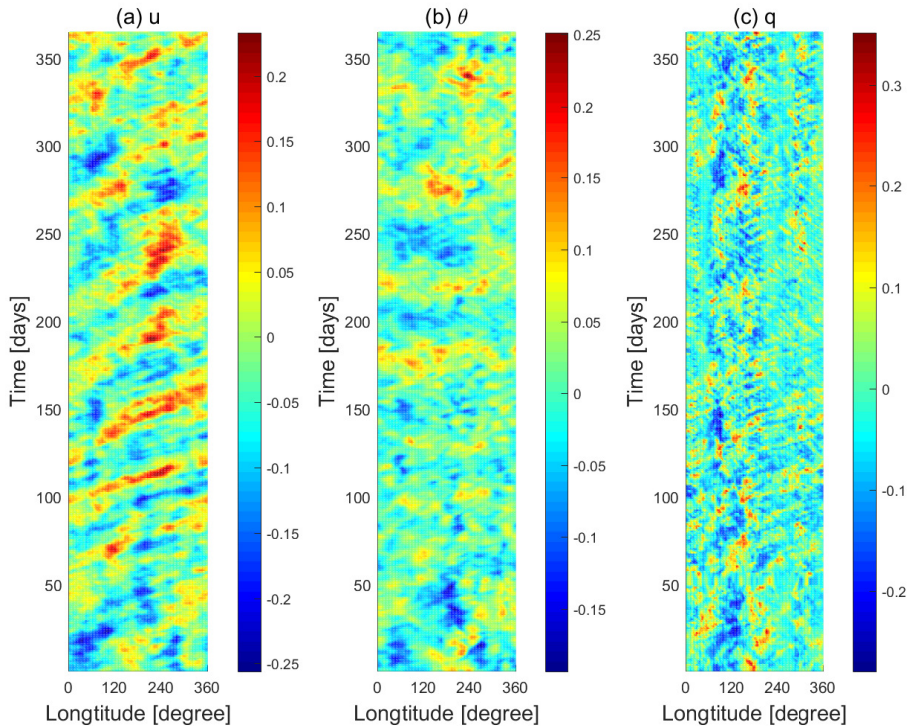
Figure 1. Plots of dimensionless (a) u, (b) θ, (c) q variables derived from observational data. The time period is 1 Jan 1979−31 Dec 1979.

using the standard equatorial reference scales shown in Table 1 [60, 89, 90, 88].

In the model in (3.1)–(3.3), the left-hand side terms represent a linearization of hydrostatic atmospheric dynamics [89, e.g.]. On the right-hand side, $-u/\tau_u$ represents momentum damping (Rayleigh friction), and $-\theta/\tau_\theta$ represents Newtonian cooling. The term q/τ_c represents a heat source and moisture sink associated with deep convection and precipitation. Finally, $b\partial_{xx}q$ and $D_*\dot{W}$ represent turbulent advection–diffusion of water vapor, represented as eddy diffusion plus stochastic forcing.

To put the model in perspective, it can be compared with other models that are similar in spirit. For instance, it is similar to convective adjustment models [70, 107, 71, 18, 87], although a nonlinear switch is not used here (for simplicity), and eddy diffusion of moisture has been added here (for allowing a somewhat realistic representation of the background spectrum of tropical convection). From another perspective, one could view this as an extension of a model for stochastic advection–diffusion of tropical moisture [33], where the present model has been extended to include equations for u and θ in order to represent waves. Only a single moisture variable has been included here, as opposed to a similar model that distinguishes between lower and middle tropospheric moisture [86]; and the use of a single moisture variable here will lead to deficiencies in the present model's representation of the MJO and convectively coupled equatorial waves (CCEWs).

Table 1. Model parameters and reference scales.

Parameter	Description	Dimensional	Nondimensional
L	Equatorial length scale	1500 km	1
T	Equatorial time scale	8 h	1
τ_u	Rayleigh drag coefficient	384 h	48
τ_θ	Newtonian cooling coefficient	384 h	48
τ_c	Convective adjustment time	96 h	12
$\widetilde{Q}$	Background vertical moisture gradient		0.9
b	Diffusion coefficient	8.33 km^2 s^{-1}	0.1066
D_*	Magnitude of stochastic forcing	12.87 g kg^{-1} km$^{1/2}$ h$^{-1/2}$	0.15
C	Velocity scale	52.08 m s^{-1}	1
	Potential temperature scale	15.6 K	1
	Geopotential height scale	265 m	1
	Water vapor scale	6.27 g kg^{-1}	1
P_e	Circumference of Earth at the equator	40132.8 km	26.7552
H_{OLR}	OLR-to-heating-rate conversion factor	0.06 K day^{-1}(W m^{-2})$^{-1}$	

The values of the parameters are chosen to obtain the best (although quite deficient) match of the climatological statistics between observational data and free-running model simulations, as described in more detail below in Section 3.2. Such an approach for parameter estimation is somewhat common [33, 86, e.g.]. The parameter values are summarized in Table 1.

The model only has two independent coordinates, x and t, which represent the longitude and the time, respectively. Therefore the model describes the tropical climate along the equator ($y = 0$) for all longitudes. The dependence on the vertical (z) coordinate would come from associating u, θ, and q with vertical basis functions that represent the first baroclinic mode, as is common practice [89, 65, 88, e.g.].

To facilitate comparison between the model and the observational data, the same spatial resolution should be used in the model. As described above in Section 2, the observational dataset has a spatial resolution of $2.5° \times 2.5°$. Since the circumference of the Earth at the equator is 40132.8 km (26.7552 in nondimensional units), there are in total $N = 144$ data points along the equator, and the grid spacing is Δx=278.7 km (0.1858 in nondimensional units).

The model is a system of stochastic partial differential equations as functions of x and t, which are linear with constant coefficients and can be solved using Fourier transforms [62, 63, 33, 86]. Some details of the solutions methods are described in Appendix 8. In the following paragraphs we refer to the model as the DTCM.

3.2 Free-running simulations

In this section, we present free-running simulations of the DTCM. The simulations can be compared versus observational data to assess the deficiencies in the model, and the comparison also provides a way to choose model parameter values (in order to obtain the best possible comparison of climatological statistics).

A free-running model simulation is shown in Figure 2. The plots are set up for comparison with the observational data in Figure 1. Both Figure 1 and Figure 2 represent only a small section from statistical steady states, and therefore the weather cannot be compared on a

day-by-day basis, but the variability of the two datasets can be compared in a statistical sense.

Note the deficiency in the waves of the climate model. In particular, note that DTCM appears to have equal amounts of waves that propagate in the eastward and westward directions (see Figure 2), whereas in observational data the eastward-propagating waves clearly dominate over the westward-propagating waves (see Figure 1). (One can identify eastward-propagating waves in these figures as streaks that are tilted from lower-left to upper-right in the plot. Conversely, one can identify westward-propagating waves in these figures as streaks that are tilted from lower-right to upper-left in the plot.) In addition to this visual appearance, this behavior can be quantified statistically by considering the power spectral density, as shown in Figure 3 and Figure 4. For observational data, the power spectra show larger power for positive wavenumbers ($k > 0$), which corresponds with eastward-propagating waves (as the phase speed is ω/k and power is plotted for $\omega > 0$). In contrast, for the deficient climate model, the spectral power has equal magnitude for positive ($k > 0$) versus negative ($k < 0$) wavenumbers, and therefore has equal magnitude for eastward- and westward-propagating waves, respectively. In short, the climate model is deficient in the sense that it fails to properly simulate the different wave types (MJO, CCEWs, etc.) and their different propagation directions. (This deficiency is also evident from the model equations, which do not include the Coriolis term, which would introduce east–west asymmetry and plays a major role in determining the propagation of the MJO and CCEWs.)

Despite its deficiency in simulating waves, note that the model does at least present a good simulation of the climate variables in two ways. First, the model and observations have similar order of magnitudes of the anomalies (roughly 0.1 to 0.3, as seen by comparing Figure 1 and Figure 2). Second, the model and observations have the same broad structure of the power spectral density (see Figure 3 and Figure 4), which, to a first approximation, resembles the power spectrum of spatiotemporal red noise [33] (although the model is deficient in that it looks like the "background spectrum" and lacks the detailed additional shape imparted by the MJO and CCEWs [101, 73, 86]).

The comparisons between model and observations (Figure 1 through Figure 4) were also used to determine the parameter values in the model. As a starting point, parameter values similar to [86] were used, and then slight modifications were made in order to obtain the best comparison between model and observations. Note that θ is noisier and more difficult to filter or predict accurately than the other two variables, so we mainly focused on fitting u and q better. This method of parameter tuning is somewhat ad hoc, although it is also quite common, and it serves the purposes of the present paper.
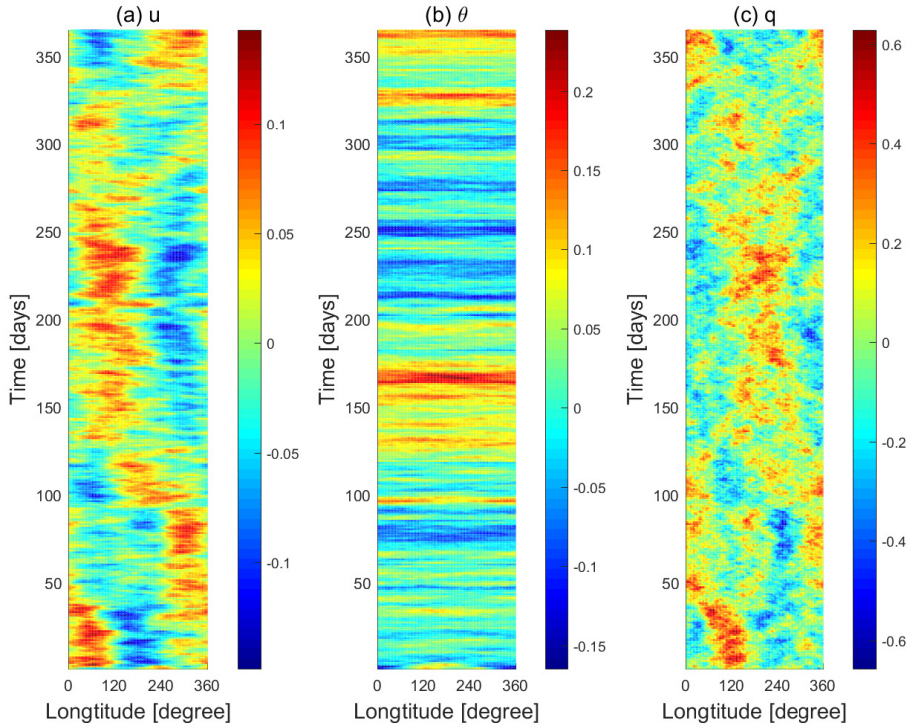
Figure 2. Plots of free-running model simulation of (a) u, (b) θ, (c) q variables. The time evolution is sampled one time every 8 hours. A one-year free-running simulation is plotted.
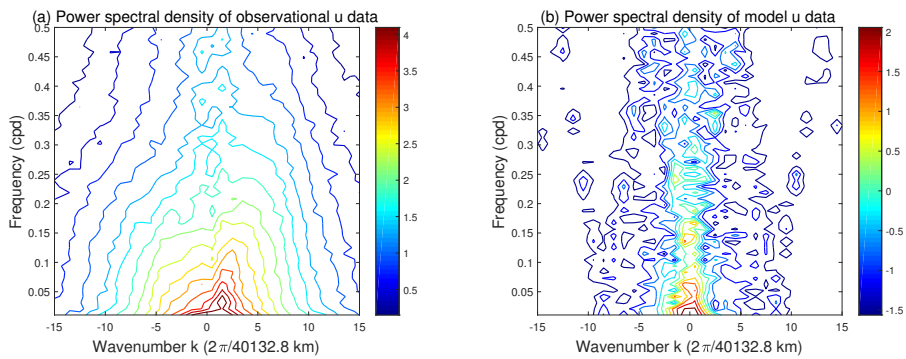

Figure 4. (a) Power spectral density plot of observational u data. (b) Power spectral density plot of model-simulated u data.

3.3 Filtering

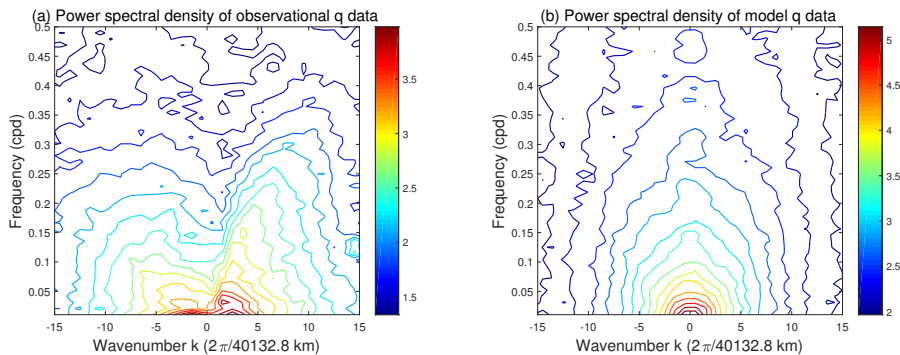In this section, we describe the Kalman filter method that is used for data assimilation,

Figure 3. (a) Power spectral density plot of observational q data. (b) Power spectral density plot of model-simulated q data.

which will be used in later sections to provide the initial conditions for forecasts. The Kalman filter is an algorithm which uses two estimates together—an estimate from the physical model and an estimate from observational data—to produce a best estimate of the state of the system (u, θ, q). The details of the Kalman filter and the method of applying it to DTCM can be found in Appendix 8.

To assess the DTCM's performance as a filter model, we use the following approach. We treat the observational data as a "true signal", and then we add artificial observation noise to the "true signal" to get artificial "corrupted observations". The Kalman filter is then applied using the artificial "corrupted observations" as one of the inputs. The strength (standard deviation) of the artificial observation noise is chosen to be 15% of the climatological standard deviation of the observational data ("true signal"), similar to other previous studies [8, e.g.].

An example of filtering performance of q in Fourier wavenumber 1 is shown in Figure 5 (a), and overall RMS error comparison for each longitude is shown in Figure 5 (b). (Notice that the corrupted observations in panel (a) appear to be quite close to true signal, despite the presence of 15% noise in the corrupted observations; the closeness is due to the fact that this figure illustrates the signal for Fourier wavenumber 1, whereas the 15% noise is added at each grid point in physical space; so the impact of the noise on Fourier wavenumber 1 will appear smaller than the noise's impact in physical space.) As a baseline for comparison in panel (b), the corrupted observations are seen to differ from the true signal with a RMS error of approximately 0.01. The filter result, on the other hand, appears to provide an estimate of the true signal that is actually slightly worse; its RMS error is roughly in the range of 0.010 to 0.013. Hence the DTCM with Kalman filter does not help us to get closer to the true signal, which is another indication of the model's deficiencies.

3.4 Forecasting

In this subsection, the forecasting performance of DTCM is tested. The criterion we use to assess the prediction performance is the model's predictability of two RMM indices. RMM 1 and 2 indices, whose full name is Real-time Multivariate MJO (RMM) series 1 and 2, are
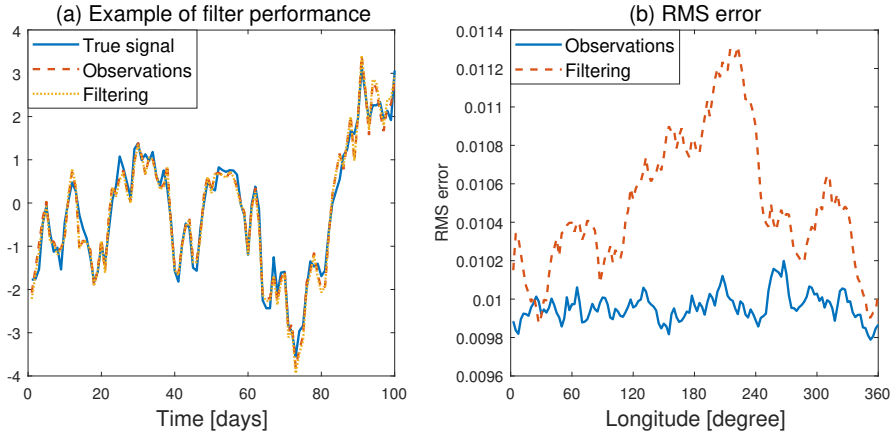
Figure 5. (a) Example of SPDE filtering compared with true signal and (corrupted) observations. The data we show in (a) is Fourier mode 1 of variable q with time period 1 Jan 1979−10 Apr 1979. (b) The averaged RMS error of (corrupted) observations and filter result for each longitude. The RMS error data are obtained by averaging independent experiments of time period 1 Jan 1979−24 May 1984.

first introduced by [102]. They are a pair of principal component (PC) time series based on projecting the observational data onto a pair of empirical orthogonal functions (EOFs). In many previous studies, the leading two EOFs generally appear as a pair and they are shown to be useful and skillful indices of monitoring the MJO. The spatial patterns of atmospheric variability, especially the eastward propagating signal, are well captured by RMM 1 and 2. Here we use these indices to aid in assessing the model's ability of simulating and forecasting the MJO.

To derive RMM indices, EOF analysis is applied to the observational data. Here the combined field of (u, q) observational data is used to determine the EOF basis. This is essentially the same choice as in [102], except the step of normalizing each field by its global variance has been neglected here because, in our model, each variable has been nondimensionalized by its reference scale and therefore the nondimensional data for u and q already have very similar global variance. The seasonal cycle and further interannual variablity have been removed in data processing procedure. The spatial structure of two leading EOFs are shown in Figure 6. The variance explained by EOF 1 and 2 is 12.02% and 11.87%, respectively. Therefore, leading pair of EOFs are used to be low-dimensional approximations which desire to capture a considerable amount of variance. By projecting u, q combined field to leading EOF basis, we can get RMM 1 and 2 time series.

The forecasting method follows standard procedures. The Kalman filter described above is used for obtaining the initial conditions for the forecast. Ensembles of DTCM simulations can be run forward in time to provide the forecast.

The MJO prediction skill (pattern correlation) as a function of prediction lead time is shown in Figure 7. Each data point is averaged among independent experiments of time period 1 Jan 1979−24 May 1984. As we can observe in the figure, the prediction skill of DTCM decays dramatically with the lead time. As a baseline, we consider that if the averaged pattern
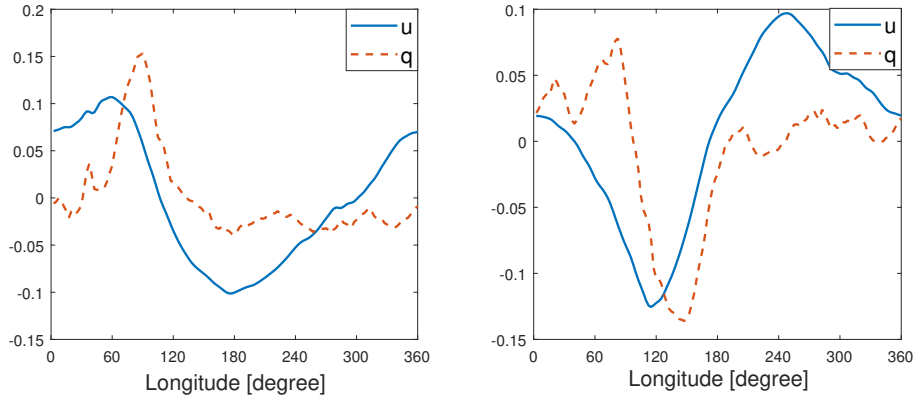
Figure 6. Spatial EOF structure of leading two EOFs.

correlation is greater than 0.5, we deem that the model can provide effective predictions under this given lead time. Therefore, DTCM can only provide effective predictions within lead time about 3 days. The prediction skill decays to nearly 0 after lead time 6 days. There is little difference in the prediction skill of the two indices RMM 1 and 2.
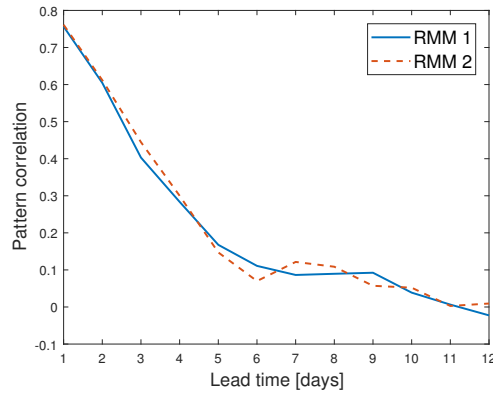

Figure 7. Prediction skill of the deficient tropical climate model. Pattern correlation data are averaged among independent tests of time period 1 Jan 1979−24 May 1984.

4 A Simple Model for Intraseasonal Oscillations

4.1 Model description

With the stated DTCM as complex climate model which requires improvements, our next step is to construct a simple model which can capture the skeleton of the key features of MJO and use the simple model to improve the performance of DTCM. More specifically, since our main focus is on predicting RMM indices, the simple model should at least have better prediction

performance for RMM indices. As stated in Section 3.4, RMM 1 and 2 indices are a pair of time series of leading EOFs which can express a considerable amount of atmospheric variability of MJO. Therefore, a complex-valued stochastic differential equation (SDE) is introduced in this section. The real part and imaginary part of the equation are aimed at representing the time evolution dynamics of the RMM 1 and 2 indices, respectively. In order to distinguish it from DTCM, we refer to the SDE model as MJO model in the following text. The complex-valued SDE has form

$$dX(t) = (-\gamma + i\omega)X(t)dt + \sigma dW(t) \tag{4.1}$$

where $\gamma, \sigma > 0$ and ω are real, $dW(t)$ is a complex Gaussian white noise with form

$$dW(t) = \frac{dW_1(t) + i\, dW_2(t)}{\sqrt{2}} \tag{4.2}$$

Each component is an independent Gaussian white noise. The remaining problem is how to determine the parameters by the given observation time series. The idea is to choose parameters to let the MJO model simulation and the observational data have the same (or close to the same) statistics, especially the autocorrelation function. The MJO model is exactly solvable. Thus a formula of autocorrelation function in terms of the 3 unknown parameters can be derived. We can compare the formula with the climatological statistics to make them match as closely as possible. Detailed calculations can be found in Appendix 9. Here we choose $\gamma = 8.8$, $\omega = 12.1$, $\sigma = 1.84$, in nondimensional units; in dimensional units, these parameters are $\gamma = 0.088 \text{ day}^{-1}$, $\omega = 0.121 \text{ day}^{-1}$, $\sigma = 0.184 \text{ day}^{-1/2}$. Figure 8 shows that there is a reasonably good fit, for such a simple model, between the autocorrelation function of the observational RMM series and the SDE model simulation.
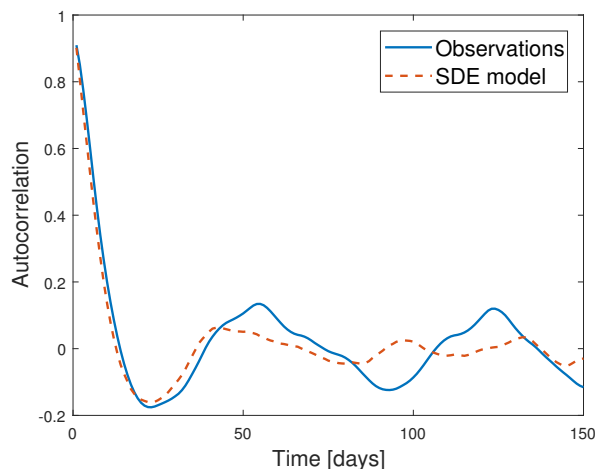

Figure 8. Real parts of autocorrelation function of observational RMM indices (solid) and SDE simulated indices (dashed).

4.2 Free-running simulation

In this section, the connection is drawn between (i) the MJO model for the RMM indices and (ii) DTCM. The connection is made by utilizing the EOF structures to associate spatial structures $u(x,t)$ and $q(x,t)$ to the MJO model time series, and this connection will be used in the multi-model communication strategy in Section 5.

A free-running simulation of the MJO model will be used to illustrate the model's variability and the reconstruction of $u(x,t)$ and $q(x,t)$. Let $X(t) = X_1(t) + iX_2(t)$. Divide the equation into its real part and imaginary part, the equation (4.1) is also equivalent to the two dimensional system

$$dX_1(t) = [-\gamma X_1(t) - \omega X_2(t)]\, dt + \frac{\sigma}{\sqrt{2}} dW_1(t) \tag{4.3}$$

$$dX_2(t) = [+\omega X_1(t) - \gamma X_2(t)]\, dt + \frac{\sigma}{\sqrt{2}} dW_2(t) \tag{4.4}$$

The system is then solved numerically using the standard Euler-Maruyama method with a time step of 0.01 in nondimensional units.

Given time series of $X_1(t)$ and $X_2(t)$, we can reconstruct a corresponding combined field $(u(x,t), q(x,t))$ by the determined EOF basis functions, which were presented earlier in Figure 6. An example of reconstructed field is shown in Figure 9. The result is similar to the OLR reconstruction shown in [102]. Even though it is a simple model, the typical eastward propagation signal is well simulated. Through these reconstructed signals of $u(x,t)$ and $q(x,t)$, the MJO model can be connected with DTCM in the multi-model communication strategy in Section 5.

4.3 Filtering

Data assimilation will be used to provide the initial conditions for the MJO model forecasts (described below in Section 4.4). As a data assimilation strategy, since the MJO model is linear, we use the Kalman filter to estimate the values of the RMM indices. Before proceeding with forecasts, we assess the filter's performance in the present section.

The detailed filter setup can be found in Appendix 10. An example of filter performance is shown in Figure 10. The filter typically provides a smoothed signal, and it often misses the maximum amplitudes of the strongest events (such as the strong event in Figure 10 from day 60 to day 120). While the filter does not appear to offer an improved estimate of the state of the RMM indices, it will still be used to provide initial conditions for forecasts, since the idea is to create a setup that is similar to an actual operational forecast scenario.

4.4 Forecasting

The simple MJO model has much greater forecast skill than the DTCM, as illustrated in Figure 11. As we can observe in the figure, the MJO model can provide effective predictions up to lead time 5-6 days, which is much better than DTCM. Therefore, it is hoped that, in the model communication strategy presented below, the information from this simple MJO model will be able help improve the DTCM's MJO forecasts.

Note that the skill of this simple MJO model is less than other simple MJO models that have been presented in other studies [95, 10, 9]. Nevertheless, it is useful in that it is easy to formulate and simulate, and it will still serve the purposes of the present paper, since its forecast skill is greater than the DTCM's forecast skill.
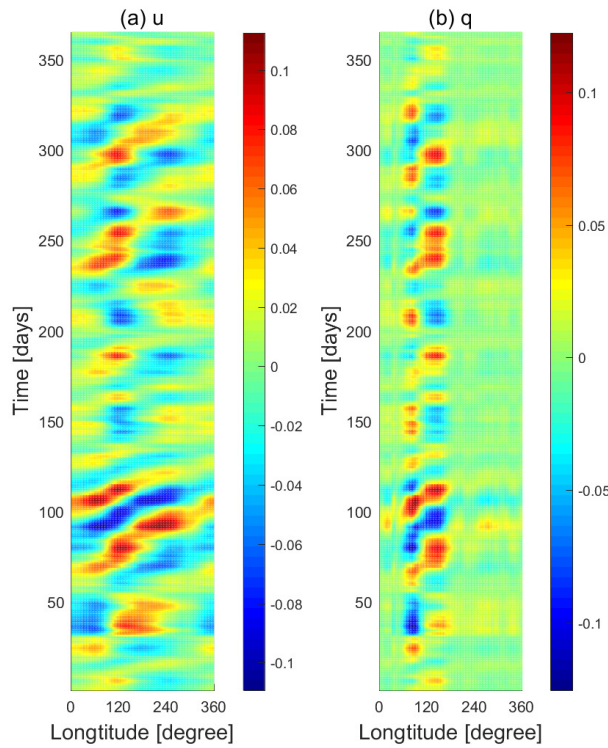
Figure 9. Signals of (a) u and (b) q reconstructed from the MJO model simulated RMM 1&2 time series, using the model in (4.3)–(4.4). The duration of the plotted data is 1 year. The duration of the plotted data is 1 year.

5   Model Communication for Improved Forecasts

In previous sections, two different types of models were described with different sets of pros and cons: the DTCM has certain defects and it has very limited prediction performance of the RMM indices, whereas the simple MJO model can only simulate two RMM time series but its prediction performance is much better.

Now, in this section, the key idea is to allow the two models to 'communicate' with each other during the forecast. By letting the comprehensive model assimilate information from the simple model, we want to improve the prediction performance of the comprehensive model to reach or even possibly exceed the performance of the simple model. More specifically, we want the DTCM to assimilate the good MJO predictions given by the MJO model.

Three main advantages of this method are as follows. The first one is that by assimilating data from simple model, we find a way to improve the comprehensive model and avoid to make any changes of the model, as comprehensive models may be too complex to modify and if modifying some parts of the model, it's hard not to touch anything else. The second one is we now have more options to make specific improvement of our model. By choosing several targeted indices and constructing focused simple models, we are able to work on improve the
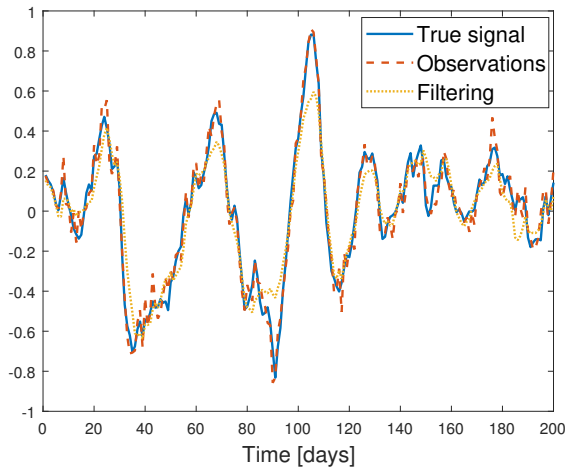
Figure 10. Example of filtering comparing with true signal and (corrupted) observations. The
data shown are RMM 1 index with time period 1 Jan 1980−19 Jul 1980.

model's performance of these specific indices. The third one is we can also have controls of
the data assimilating process. Choose observation errors to represent the degree of trust to the
simple model and choose proper assimilation frequencies to involve more information or reduce
computing costs. More details are discussed in the following subsections.

In this first investigation of model communication, a simple Kalman filter method is used to
achieve communications between models. When making forecasts, we run the MJO model pre-
dictions of RMM indices independently, and then these RMM index predictions are considered
"observations" by the DTCM. The DTCM assimilates the RMM index predictions to hopefully
get better prediction results. The details about how to apply the filter algorithm in the process
is provided in Appendix 12.

5.1 Forecast skill with model communication

The forecast skill with model communication is shown in Figure 12. Four cases of model
communication are shown, and they are labeled by the strength of artificial "observational
noise" that is added to "observed" RMM indices supplied by the simple MJO model. The
four noise strengths are 0%, 5%, 10%, and 15% of the climatological standard deviation of the
RMM time series. It is reasonable to include some noise, partly because the MJO model's
predictions are themselves not perfect, and partly as a way to assess the impact of errors in the
prediction system. The noise strength could also be taken as an indicator of the trustworthiness
of the simple MJO model predictions, where the extreme case of 0% indicates that the MJO
model predictions are trusted completely relative to the DTCM predictions of the RMM indices.
While the noise strength is simply assigned a fixed value in the tests shown here, it would be
interesting in the future to allow the different models to learn the trustworthiness of the other
models' information based on the skill of past forecasts. Also shown for comparison in Figure
12 are the forecast skills of the MJO model (SDE model) and the DTCM (SPDE model).
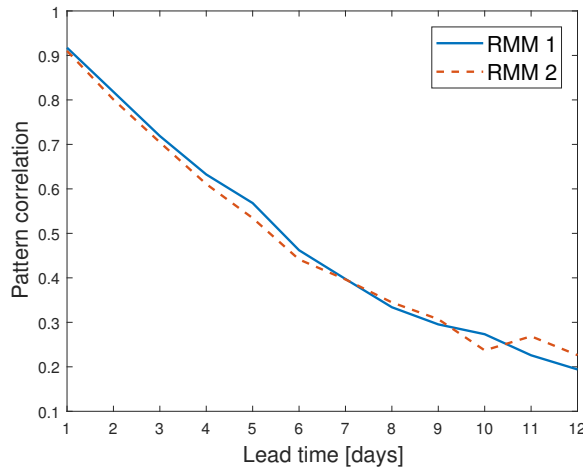
Figure 11. MJO model prediction skill of RMM indices 1 and 2. The pattern correlation data are averaged among 5000 independent tests. The testing time window is 1 Jan 1980−9 Sept 1993.

In brief, one can see in Figure 12 that the skill of model communication can become equal to that of the MJO model (SDE). Compared to the skill of the DTCM (SPDE model), this represents a significant improvement. While it is perhaps not surprising that the model communication skill is equal to the MJO model (SDE) skill in the case of 0% noise (since the forecast target is the same as the assimilated RMM index data, and the models here are both linear), it is interesting to note the changes in forecast skill for different noise strengths. In particular, relative to 0% noise, the case of 5% noise has essentially the same skill, and the case of 5% noise has only a somewhat small reduction. Furthermore, beyond these points of note for RMM 1, notice that RMM 2 is forecast skillfully via model communication for all considered noise strengths. Differences in forecast skill for RMM 1 versus RMM 2 have also been noted in some past studies with other models [95, e.g.]. Overall, these results suggest the possibility of increasing the DTCM's forecast skill through a model communication strategy.

5.2 Choice of assimilation frequencies

In the results described above, each model communication result is obtained by assimilating "observations" every 1 day (in nondimensional units, $\Delta t = 3$). Since the MJO model provides better predictions of the RMM indices, we would expect that assimilating information from the MJO model more often would generate better results. Also note that assimilation frequencies lead to no difference when the artificial observational noise is 0. In Figure 13, the observation error used is fixed to be 15%. Take RMM index 1 as an example. The pattern correlations of model communication results of different assimilation frequencies are shown in Figure 13. As we can observe in the figure, assimilation frequencies do affect the forecast skill of model communication. Higher "observation"/communication frequencies tend to provide better results but more calculation may be involved in the process.
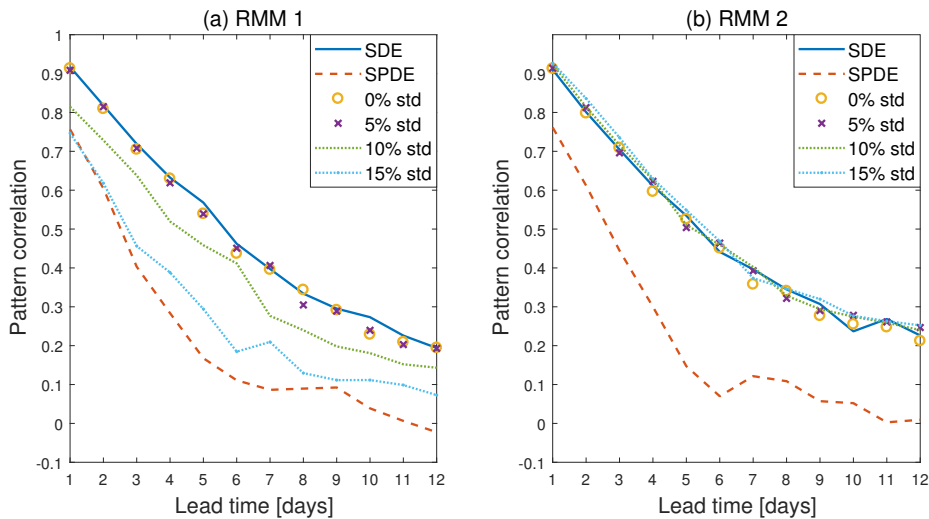
Figure 12. Prediction skills of (a) RMM 1 and (b) RMM 2. The options provided are MJO predictions (SDE), DTCM predictions (SPDE), and model communication performance with observation error 0%, 5%, 10%, 15% of climatological standard deviation.

5.3 Comparison with multi-model mean

Another way to utilize multiple models in making a forecast is called the multi-model mean (MMM). To create the multi-model mean, each forecast model is run independently, with no communication between the different models. For instance, here the result would be a value x_{simple} predicted by the simple MJO model alone, and also a value x_{DTCM} predicted by the DTCM alone. The multi-model mean prediction, x_{MMM}, is then defined simply as the average: $x_{MMM} = (x_{simple} + x_{DTCM})/2$.

The MMM prediction skill is shown in Figure 14, in comparison with the prediction skills of the DTCM (SPDE), the MJO model (SDE), and model communication (with 5% standard deviation artificial observation error). Multi-model mean prediction skill falls in between the MJO model and DTCM. It does provide some improvement over the DTCM by itself, and it does not require any communication between forecast models during the forecast stage. However the data assimilation-based model communication strategy we stated in Section 5.1 provides better results.

6   Conclusions

A model communication strategy was investigated for allowing multiple forecast models to assimilate information from each other in order to overcome individual model deficiencies. The results show that a deficient climate model's skill at forecasting the MJO can be significantly enhanced through communication with a skillful, simple MJO model. The multi-model communication strategy is also able to outperform the multi-model mean, which utilizes multiple models only in post-processing and not in any type of direct communication during the forecasts.
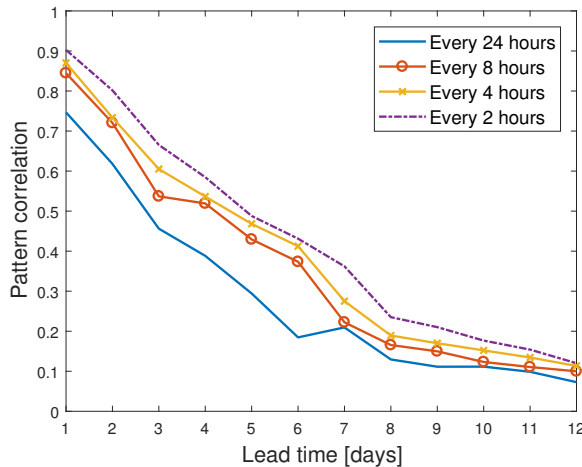
Figure 13. Model communication prediction skill of RMM 1 for different assimilation frequencies, all with relatively strong artificial observational error of 15%.

In the future, it would be interesting to test the model communication strategy in a setup that is closer to an operational forecast. In the present study, a simplified setup was used in order to allow transparent evaluation of different components of the strategy. A future next step could utilize an actual global forecast model rather than the simplified linear Gaussian model used here, and it could utilize more sophisticated data assimilation methods for the inter-model communication rather than the simple Kalman filter used here. Also, the nonlinearity of an operational forecast model would cause the assimilation of MJO EOF information to be spread to other components of the weather–climate system other than the MJO, and it would be interesting to see whether such spreading of information is beneficial or detrimental to the forecast of other components of the weather–climate system other than the MJO.

7 Appendix: Data processing

7.1 Estimate the first baroclinic component of u and θ

The zonal wind velocity data and geopotential height data are used to estimate u and θ, respectively. Two pressure levels, 850hPa and 200hPa, are used to represent the bottom and the top of the troposphere, respectively. Then associate the data with contributions from two vertical modes, a barotropic contribution and a first baroclinic mode contribution. Take zonal wind velocity u as an example. The velocity of different pressure level can be expressed as

$$u(z) = u_{\mathrm{BT}} + u_{\mathrm{BC}}\sqrt{2}\cos(z) \tag{7.1}$$

where u_{BT} represents the barotropic contribution and u_{BC} represents the first baroclinic mode contribution. Given the data from two pressure levels, the first baroclinic component can
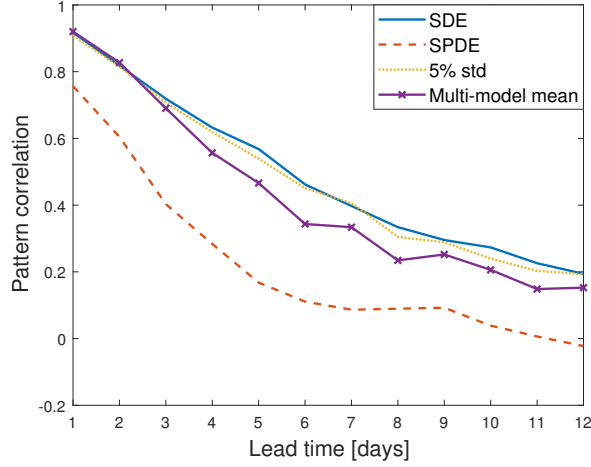
Figure 14. Multi-model mean prediction skill (crosses) in comparison with the MJO model (SDE, solid), the DTCM (SPDE, dashed), and the DTCM with 5% observation error.

be estimated by

$$u_{\text{BC}}(x, y, t) = \frac{u(850\text{hPa}) - u(200\text{hPa})}{2\sqrt{2}} \tag{7.2}$$

Similarly, the first baroclinic component of the geopotential height can be estimated as following

$$Z_{\text{BC}}(x, y, t) = \frac{Z(850\text{hPa}) - Z(200\text{hPa})}{2\sqrt{2}} \tag{7.3}$$

And consider the hydrostatic balance, which has form $\partial z / \partial p = -\theta$. It results in the expression of the potential temperature anomaly θ from the model in terms of the geopotential height anomaly Z from the data

$$\theta_{\text{BC}}(x, y, t) = -\frac{Z(850\text{hPa}) - Z(200\text{hPa})}{2\sqrt{2}} \tag{7.4}$$

Since we use the standard reference scale to nondimensionalize the model, the reference scale for u and θ can be directly derived. The reference scales for u and θ are 52.08 m/s and 15.6 K, respectively.

7.2 Estimate of the water vapor anomaly q

The OLR data are used to estimate water vapor anomaly q. Since water vapor can absorb certain wavelengths of OLR adding heat to the atmosphere. High concentration of water vapor tends to reduce OLR below clear sky values. Thus we can use OLR data as simple indicators to estimate the water vapor anomaly.

$$q(x, y, t) = -\text{OLR} \tag{7.5}$$

The reference scale for water vapor anomaly q is deduced by the following steps. In a recent study [88], the water vapor reference scale is 6.27 g/kg. We use the same reference scale here. However unit conversions are needed to apply the reference scale to the OLR dataset above. The observational OLR dataset has unit wm^{-2}. Firstly, we use a OLR-to-heating-rate conversion factor which is 0.06 $\mathrm{Kday}^{-1}(\mathrm{wm}^{-2})^{-1}$. This conversion factor is the same as [88] which is based on a recent study [90]. Now the unit becomes Kday^{-1}. We want to further convert it to $\mathrm{mm} \cdot \mathrm{day}^{-1}$. Assume an area of 1 m^2, now we estimate the mass of air in the atmospheric column. Consider the pressure of air is decaying exponentially with the height and the scale height is $\mathrm{H_p} = 10$ km. The function of pressure in terms of height z can be expressed by

$$P(z) = P_0 e^{-\frac{z}{\mathrm{H_p}}} \tag{7.6}$$

Note that the scale height depends on the air temperature, by turning the scale height to a function of air temperature we can get a better estimate of the pressure. But here we only want to get a rough estimate so a constant value of scale height is use in the computation. According to the ideal gas law, assume the temperature is constant, the density of air is proportion to the air pressure. Thus the density of air can be expressed by

$$\rho(z) = \rho_0 e^{-\frac{z}{\mathrm{H_\rho}}} \tag{7.7}$$

Use $\rho_0 = 1$ $\mathrm{kg/m}^3$. Integrate the density of air from $z = 0$ to infinity and times the area. An estimation of the mass of air in the column is 10000 kg. Assume a depth of 1 mm of liquid water (after the water vapor is condensed to liquid), we use that the density of water is 10^3 $\mathrm{kg/m}^3$ and the latent heat factor of water is 2500 kJ/kg. If this amount of water is converted from water vapor to liquid, the heat release in the process is 2.5×10^6 J. If this amount of heat is used to warm the air in the atmospheric column and we use the specific heat of air is 1004 $\mathrm{J/(kg \cdot K)}$, the temperature of air will raise 0.249 K. Thus we get the unit conversion relation: 1 K/day $\approx$ 4 mm/day.

Now the unit becomes mm/day. Since the reference scale we use is 6.27 g/kg and the 1 m^2 column has 10000 kg air, that equals to 62.7 kg water vapor, which also equals to 62.7 mm liquid water in the column. This value can only apply to the air in low altitude or in other words, near the surface. If taking the vertical structure of water vapor into account, a more accurate expression of the ratio is needed. By many other studies [77, 43, 71], the water vapor ratio of the vertical atmospheric column also decays exponentially with the height. Thus the water vapor ratio can be expressed by

$$q(z) = q_0 e^{-\frac{z}{\mathrm{H_q}}} \tag{7.8}$$

Here we also use $\mathrm{H_q} = 10$ km as scale height and $q_0 = 62.7$ mm. Note that the reference scale is the ratio of water vapor to the total humid air. And the total amount of humid air decays with height according to (7.7). Thus the reference scale used in our model can be calculated by

$$q_{\mathrm{ref}} = \int_0^\infty q(z)\rho(z) \cdot 1 \ \mathrm{m}^2 \cdot dz$$
$$= 31.35 \ \mathrm{mm} \tag{7.9}$$

To associate the water vapor reference scale with the observational OLR data, divide the reference scale by $\tau_c = 96$ h $= 3$ days. Thus the heating-rate reference scale is 10.45 mm/day $=$ 2.61 K/day $= 43.54$ wm^{-2}. By using this reference scale, we can nondimensionalize the OLR data.

7.3 Meridional projections

By previous step, we find proper surrogates for all three variables (u, θ, q) in our model. Next step is to use the meridional mode decomposition method to capture the skeleton of the data. By utilizing the meridional basis functions which take forms of the parabolic cylinder functions. Each variable can be expressed as linear combinations of parabolic cylinder functions. Take u_{BC} as an example,

$$u_{\text{BC}}(x, y, t) = \sum_{m=0}^{\infty} u_m(x, t)\phi_m(y) \tag{7.10}$$

where $\phi_m(y)$ are basis functions and u_m represents the m-th meridional mode. Use the 0 meridional mode $u_0(x, t)$ as the skeleton of the whole data/ This quantity can be obtained by the projection

$$
\begin{aligned}
u_0(x, t) &= \int_{-\infty}^{\infty} u_{\text{BC}}(x, y, t)\phi_0(y)dy \\
&= \int_{-\infty}^{\infty} u_{\text{BC}}(x, y, t)\frac{1}{\pi^{1/4}}e^{-\frac{y^2}{2}}dy
\end{aligned}
\tag{7.11}
$$

This meridional projection also reduces the 2D (x, y) dataset to an 1D (x) dataset. Same formula can also apply to θ and q. We can project θ_{BC} and q to $\phi_0(y)$ and get $\theta_0(x, t)$, $q_0(x, t)$ respectively. We use (u_0, θ_0, q_0) as an estimate of variables (u, θ, q) in our model. To sum up, we apply the above data processing method to OLR, zonal wind velocity and geopotential height dataset. Then nondimensionalize the data by their corresponding reference scales and project them to meridional basis functions to get the observation data we use in our model. The observation data of three variables derived from the dataset are shown in Figure 1. Eastward propagating zonal wind can be clearly observed from Figure 1 (a).

8 Appendix: Solving the tropical climate model

Define the discrete Fourier transform for u as following,

$$\hat{u}_k = \sum_{n=0}^{N-1} u_n e^{-2\pi ikn/N}, \quad k = 0, 1, \cdots, N-1 \tag{8.1}$$

Analogous formulas can also be apply to θ and q. Then we get the expression of the system in Fourier space,

$$\frac{du_k}{dt} = -\frac{1}{\tau_u}u_k + \frac{2\pi ik}{L}\theta_k \tag{8.2}$$

$$\frac{d\theta_k}{dt} = \frac{2\pi ik}{L}u_k - \frac{1}{\tau_\theta}\theta_k + \frac{1}{\tau_c}q_k \tag{8.3}$$

$$\frac{dq_k}{dt} = -\widetilde{Q}\frac{2\pi ik}{L}u_k - \left(\frac{1}{\tau_c} + \frac{4\pi^2 k^2 b}{L^2}\right)q_k + D_*\widehat{\widetilde{W}}_k \tag{8.4}$$

where $\widehat{\dot{W}}_k = \sum_{n=0}^{N-1} \dot{W}_n e^{-2\pi i k n / N}$ is the k-th Fourier mode of the white noise. The white noise $\dot{W}$ satisfies,

$$\begin{cases} E[\dot{W}_n(t)] = 0 \\ E[\dot{W}_n(t)\dot{W}_n(s)] = \delta(t - s) \end{cases} \tag{8.5}$$

By using the expression of $\widehat{\dot{W}}_k$ stated above, the properties of $\widehat{\dot{W}}_k$ can be directly derived,

$$\begin{cases} E\left[\widehat{\dot{W}}_k(t)\right] = \sum_{n=0}^{N-1} E[\dot{W}_n(t)]e^{-2\pi i k n / N} = 0 \\ E\left[\widehat{\dot{W}}_k(t)\widehat{\dot{W}}_k^*(s)\right] = E\left[(\sum_{n=0}^{N-1} \dot{W}_n(t)e^{-2\pi i k n / N})(\sum_{m=0}^{N-1} \dot{W}_m(s)e^{2\pi i k m / N})\right] = N\delta(t - s) \end{cases} \tag{8.6}$$

Now the system has been transformed to a three-dimensional linear system. Denote $(u_k, \theta_k, q_k)^T$ by V_k. The system can be rewrote by matrices form,

$$\frac{dV_k}{dt} = \begin{pmatrix} -\frac{1}{\tau_u} & \frac{2\pi i k}{L} & 0 \\ \frac{2\pi i k}{L} & -\frac{1}{\tau_\theta} & \frac{1}{\tau_c} \\ -\widetilde{Q}\frac{2\pi i k}{L} & 0 & -\frac{1}{\tau_c} - \frac{4\pi^2 k^2 b}{L^2} \end{pmatrix} V_k + \begin{pmatrix} 0 \\ 0 \\ D_*\widehat{\dot{W}}_k \end{pmatrix} = F_k V_k + D_k \tag{8.7}$$

Note that the coefficient matrix depends on the Fourier wavenumber k. But each Fourier mode is independent with others which makes it possible to focus on solving one system and apply the same method to all other Fourier modes. Next, diagonalization is performed to the system. Let A_k be the eigenvector matrix (whose column are eigenvectors) of F_k and Λ_k be the eigenvalue matrix (whose diagonal entries are eigenvalues). And let $U_k = A_k^{-1}V_k = (U_{k,1}, U_{k,2}, U_{k,3})^T$. Then U_k satisfies the diagonalized system,

$$\frac{dU_k}{dt} = \Lambda_k U_k + A_k^{-1}D_k \tag{8.8}$$

Each variable can be solved analytically by integrating exactly and applying the properties of $\widehat{\dot{W}}_k$. The exact solution is

$$U_{k,j}(t) = e^{\Lambda_k(j,j)t}U_{k,j}(0) + \int_0^t A_k^{-1}(j,3)D_* e^{\Lambda_k(j,j)(t-s)}d\hat{W}_k(s) \quad j = 1, 2, 3 \tag{8.9}$$

Note that the integral in (8.9) is a Gaussian random variable at each time t. The mean and variance can be easily derived. Then $U_{k,j}(t)$ has mean $e^{\Lambda_k(j,j)t}U_{k,j}(0)$ and variance

$$\text{Var}[U_{k,j}(t)U_{k,j}^*(t)] = \frac{D_*^2 N A_k^{-1}(j,3)A_k^{-1}(j,3)^*}{-2\text{Real}(\Lambda_k(j,j))}\left(1 - e^{2\text{Real}(\Lambda_k(j,j))t}\right) \tag{8.10}$$

This method is called semi-analytical solution method because the solution is derived analytically which means no numerical integration error occur, but the solution contains random term which may cause sampling error in realization. This formula can be used to numerically realize the model. Each Fourier mode can be evolved forward independently. For Fourier mode k, in each time step, given the initial data vector $V_{k,n} = (u_n, \theta_n, q_n)^T$, use the matrix A_k to diagonalize the system and get U_k. Then draw an Gaussian random variable with mean and variance given in (8.10). Then use matrix A_k again to change back and get numerical solution $V_{k,n+1}$. Treat it as initial condition in next time step to move forward.

9 Appendix: Kalman filter for the tropical climate model

A detailed description of the algorithm is introduced in [67]. In the process, two kinds of information is involved—an estimation from the previous time step and an estimation from current measurement. Assume u is an M-dimensional unknown variable which requires estimation. And assume the filter model has form,

$$u_{m+1} = Fu_m + \sigma_{m+1} \tag{9.1}$$

where F is a complex dynamic matrix and σ_{m+1} is a M-dimensional complex Gaussian noise vector with zero mean and covariance $R = <\vec{\sigma}_m \otimes \vec{\sigma}_m^T>$. Thus we assume that the equation (9.1) can give an unbiased estimation of the current state u_{m+1} from the previous state. Two states of the unknown variable are provided in the process. The prior state $u_{m+1|m}$ is an estimation of the true state prior to the knowledge of the observation at time m+1. The posterior state $u_{m|m}$ is an estimation of the true state after considering the observation at time m. The relation is given by

$$u_{m+1|m} = Fu_{m|m} + \sigma_{m+1} \tag{9.2}$$

Define the mean state $\vec{\bar{u}}_{m+1|m}$ by taking the mean of $\vec{u}_{m+1|m}$ and the error covariance

$$R_{m+1|m} = <(\vec{u}_{m+1} - \vec{\bar{u}}_{m+1|m}) \otimes (\vec{u}_{m+1} - \vec{\bar{u}}_{m+1|m})^T> \tag{9.3}$$

And assume noisy observations satisfy

$$v_m = Gu_m + \sigma_m^0 \tag{9.4}$$

where G is the complex observation matrix and σ_m^0 is a M-dimensional Gaussian noise vector with zero mean and covariance $R^0 = <\vec{\sigma}_m^0 \otimes \vec{\sigma}_m^{0T}>$. Kalman filter is a two-step algorithm. In the prediction step, estimate prior mean state and prior error covariance by

$$\begin{cases} \vec{\bar{u}}_{m+1|m} = F\vec{\bar{u}}_{m|m} \\ R_{m+1|m} = FR_{m|m}F^T + R \end{cases} \tag{9.5}$$

In the correction step, the kalman gain matrix is defined as

$$K_{m+1} = R_{m+1|m}G^T(GR_{m+1|m}G^T + R^0)^{-1} \tag{9.6}$$

And the estimations of the posterior state are

$$\begin{cases} \vec{\bar{u}}_{m+1|m+1} = \vec{\bar{u}}_{m+1|m} + K_{m+1}(\vec{v}_{m+1} - G\vec{\bar{u}}_{m+1|m}) \\ R_{m+1|m+1} = (I - K_{m+1}G)R_{m+1|m} \end{cases} \tag{9.7}$$

The posterior state is used as the kalman filter estimate of the true state. To associate the filter algorithm with DTCM, several choices should be made. Note that the model can be linearized in Fourier space. The true signal we want to estimate is $V_k = (u_k, \theta_k, q_k)^T$, for each Fourier wavenumber k. Since there is an easy transition between V_k and U_k stated in Appendix

8, we will estimate U_k instead. The time evolution dynamics of U_k are given by equation (8.9). For each time step, the dynamic matrix

$$F_{k,\text{DTCM}} = \begin{pmatrix} e^{\lambda_k(1,1)\Delta t} & & \\ & e^{\lambda_k(2,2)\Delta t} & \\ & & e^{\lambda_k(3,3)\Delta t} \end{pmatrix} \tag{9.8}$$

And the model error covariance matrix

$$R_{k,\text{DTCM}} = \begin{pmatrix} \text{Var}\left[U_{k,1}(\Delta t)U_{k,1}^*(\Delta t)\right] & & \\ & \text{Var}\left[U_{k,2}(\Delta t)U_{k,2}^*(\Delta t)\right] & \\ & & \text{Var}\left[U_{k,3}(\Delta t)U_{k,3}^*(\Delta t)\right] \end{pmatrix} \tag{9.9}$$

where $\text{Var}\left[U_{k,1}(\Delta t)U_{k,1}^*(\Delta t)\right]$ is given by equation (8.10). And the observations we use are data derived from dataset in Section 2. Thus the observation matrix is

$$G_{\text{DTCM}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{9.10}$$

Note that the observation error covariance matrix is a measurement of the extent of how we trust the observation data. We choose 15% of the standard deviation of each variable from the observation data as observation error, as is common in other similar studies [32, 8]. Thus the observation error matrix is

$$R_{\text{DTCM}} = \begin{pmatrix} (15\% \text{ of std of } u)^2 & & \\ & (15\% \text{ of std of } \theta)^2 & \\ & & (15\% \text{ of std of } q)^2 \end{pmatrix} \tag{9.11}$$

Apply Kalman filter algorithm to each time step and provide the filter results given by DTCM.

10 Appendix: Kalman filter for simple MJO model

Implement the standard Kalman filter to simple MJO model. For consistency, we use the same notation as in Appendix 8. As the filter model, we use the complex MJO model (4.1), and we create complex-valued observational RMM indices by using $X(t) = $RMM $1 + i \times$RMM 2. To choose the filter parameters such as the model's evolution operator and noise magnitude, consider the analytic solution of the MJO model:

$$X(t) = e^{(-\gamma+i\omega)t}X(0) + \sigma \int_0^t e^{(-\gamma+i\omega)(t-s)}dW(s) \tag{10.1}$$

Such a formula will be used to evolve the model from one observation time to the next, with observation time interval Δt. Therefore, if using X^T as the signal we want to estimate, the filter model dynamic matrix will be

$$F_{\text{MJO}} = e^{(-\gamma+i\omega)\Delta t} \tag{10.2}$$

The model error matrix

$$R_{\mathrm{MJO}} = \frac{\sigma^2}{2\gamma}(1 - e^{-2\gamma\Delta t}) \tag{10.3}$$

and the observation matrix $G_{\mathrm{MJO}} = 1$. To create a scenario with "observational error", we treat the observed RMM indices as the "true signal" and add artificial observation error to get artificial "observations". For this purpose, we use 15% of the RMM indices' standard deviation as the "observation error" ($R^0 = 15\%(\mathrm{std}(\mathrm{RMM}\ 1) + i \times \mathrm{std}(\mathrm{RMM}\ 2))$).

11 Appendix: Determining parameters in the MJO model

Autocorrelation function is a useful tool to analyze the repeating patterns in a signal. In order to determine the parameter values, we assume the RMM indices satisfy the form of the MJO model (4.1). We first derive the autocorrelation function of the given SDE equation. Note that the MJO model is a complex Ornstein-Uhlenbeck process. The analytic solution is

$$X(t) = e^{(-\gamma+i\omega)t}X(0) + \sigma \int_0^t e^{(-\gamma+i\omega)(t-s)}dW(s) \tag{11.1}$$

That means at any time t, the solution is a random Gaussian variable with mean and variance

$$\mathrm{E}[X(t)] = e^{(-\gamma+i\omega)t}X(0) \tag{11.2}$$

$$\mathrm{Var}\,[X(t)] = \frac{\sigma^2}{2\gamma}(1 - e^{-2\gamma t}) \tag{11.3}$$

Define the autocorrelation function,

$$R(\tau) = \lim_{T\to\infty} \frac{1}{T\mathrm{Var}[X(\infty)]} \int_0^T X(s)X(s+\tau)^* ds \tag{11.4}$$

Then we can calculate the autocorrelation function of the SDE,

$$R(\tau) = \lim_{T\to\infty} \frac{1}{\mathrm{Var}[X(\infty)]}\mathrm{E}\left[\left(\sigma \int_0^T e^{(-\gamma+i\omega)(T-t')}dW(t')\right)\left(\sigma \int_0^{T+\tau} e^{(-\gamma+i\omega)(T+\tau-s')}dW(s')\right)^*\right]$$
$$= \lim_{T\to\infty} \frac{1}{\mathrm{Var}[X(\infty)]}\sigma^2 e^{-2\gamma T}\int_0^T \int_0^{T+\tau} e^{\gamma(t'+s'-\tau)-i\omega(t'-s'+\tau)}\delta(t'-s')dt'ds'$$
$$= e^{-\tau(\gamma-i\omega)} \tag{11.5}$$

We denote the integral of the autocorrelation function as

$$\int_0^\infty R(\tau)d\tau = \frac{\gamma-i\omega}{\gamma^2+\omega^2} = S_1 + iS_2 \tag{11.6}$$

Thus the integral of the autocorrelation function can be used to calculate the parameters. And furthermore, by equation (11.3), σ can also be calculated.

$$\gamma = \frac{S_1}{S_1^2 + S_2^2} \tag{11.7}$$

$$\omega = -\frac{S_2}{S_1^2 + S_2^2} \tag{11.8}$$

$$\sigma = \sqrt{2\gamma \mathrm{Var}[\mathrm{X}(\infty)]} \tag{11.9}$$

12 Appendix: Kalman filter for model communication

To implement standard Kalman filter for model communication, we now use RMM time series simulated by the simple SDE model. The following equations summarize the relations between signals.

$$
\begin{cases}
(u_k, \theta_k, q_k)^T_{m+1} = F_{k,\mathrm{DTCM}} \cdot (u_k, \theta_k, q_k)^T_m + \sigma_{m+1,\mathrm{DTCM}} \\
(\mathrm{RMM\ 1} + i\mathrm{RMM\ 2})_{m+1} = F_{\mathrm{MJO}} \cdot (\mathrm{RMM\ 1} + i\mathrm{RMM\ 2})_m + \sigma_{m+1,\mathrm{MJO}}
\end{cases}
\tag{12.1}
$$

$$
(\mathrm{RMM\ 1}, \mathrm{RMM\ 2})^T_m = G \cdot (u_k, q_k)^T_m + \sigma^0_m
\tag{12.2}
$$

The true signal we want to estimate is (u_k, θ_k, q_k) and the "observations" are RMM time series $(\mathrm{PC}_1, \mathrm{PC}_2)$. These two signals can be evolved by equations (12.1). However the true signal is not fully observed since only two components (u_k, q_k) contribute to the time series. In this case, the relation between true signal and observation can be expressed in equation (12.2).

Now the problem is to find observation matrix G and properties of the error term σ^0_m. To determine the observation matrix G, an alternative representation of time series in terms of Fourier modes is required. Firstly, by EOF analysis, the (u, q) combined field can be decomposed by the EOF basis. Let $\phi_1, \phi_2, \cdots, \phi_N$ denote the EOF basis vector and $\mathrm{PC}_1(t), \mathrm{PC}_2(t), \cdots, \mathrm{PC}_N(t)$ denote their corresponding time series. Note that $\mathrm{PC}_1(t), \mathrm{PC}_2(t)$ are exactly RMM 1 and 2 indices. Let u and q be of form time$\times$longitude. Then we can write

$$
(u, q)^T = \mathrm{PC}_1(t)\phi_1 + \mathrm{PC}_2(t)\phi_2 + \cdots + \mathrm{PC}_N(t)\phi_N
\tag{12.3}
$$

Since EOF basis vectors are pairwise orthogonal, the PC time series can be calculated by inner products,

$$
\mathrm{PC}_j(t) = <(u, q)^T, \phi_j >
\tag{12.4}
$$

Let $u = (u_1, u_2, \cdots, u_N)$, $q = (q_1, q_2, \cdots, q_N)$ and $\phi_j = (\phi_{j,1}, \phi_{j,2}, \cdots, \phi_{j,2N})^T$. By discrete Fourier transformation,

$$
u_n = \frac{1}{N} \sum_{k=0}^{N-1} \hat{u}_k e^{i2\pi kn/N}
\tag{12.5}
$$

$$
q_n = \frac{1}{N} \sum_{k=0}^{N-1} \hat{q}_k e^{i2\pi kn/N}
\tag{12.6}
$$

Then plug in equation (12.4).

$$
\begin{aligned}
\mathrm{PC}_j(t) &= <(u, q)^T, \phi_j > \\
&= \sum_{k=0}^{N-1} \left[\frac{1}{N} \sum_{n=0}^{N-1} e^{i2\pi kn/N} \phi_{j,n} \right] \hat{u}_k + \sum_{k=0}^{N-1} \left[\frac{1}{N} \sum_{n=0}^{N-1} e^{i2\pi kn/N} \phi_{j,n+N} \right] \hat{u}_k \\
&= \sum_{k=0}^{N-1} \alpha_{j,k} \hat{u}_k + \sum_{k=0}^{N-1} \beta_{j,k} \hat{q}_k
\end{aligned}
\tag{12.7}
$$

Therefore, by equation (12.7), we associate the EOF time series with Fourier modes. Consequently, we can bring the Fourier representation of the time series in the filter algorithm. According to the diagonalization in Appendix 8, $V_k = A_k U_k$. Denote

$$A_k = \begin{pmatrix} A_{k,1,1} & A_{k,1,2} & A_{k,1,3} \\ A_{k,2,1} & A_{k,2,2} & A_{k,2,3} \\ A_{k,3,1} & A_{k,3,2} & A_{k,3,3} \end{pmatrix} \tag{12.8}$$

Then

$$\hat{u}_k = \sum_{l=1}^{3} A_{k,1,l} U_{k,l} \tag{12.9}$$

$$\hat{q}_k = \sum_{l=1}^{3} A_{k,3,l} U_{k,l} \tag{12.10}$$

Plug in equation (12.7),

$$\mathrm{PC_j}(t) = \sum_{k=0}^{N-1} \sum_{l=1}^{3} (\alpha_{j,k} A_{k,1,l} + \beta_{j,k} A_{k,3,l}) U_{k,l} \tag{12.11}$$

With the above preparations, now we now to the filter algorithm. For each time step t_m, the true signal which requires estimation is $u_m = (U_{0,1}, U_{0,2}, U_{0,3}, \cdots, U_{N-1,1}, U_{N-1,2}, U_{N-1,3})^T$. Note that we omit the time subscript m in the vector to avoid redundant subscripts. Note that the observations (RMM 1, RMM 2) are time series $v_m = (\mathrm{PC_1}(t_m), \mathrm{PC_2}(t_m))$. According to equation (58), the observation matrix is

$$G = \begin{pmatrix} \alpha_{1,0} A_{0,1,1} + \beta_{1,0} A_{0,3,1} & \alpha_{1,0} A_{0,1,2} + \beta_{1,0} A_{0,3,2} & \cdots & \alpha_{1,N-1} A_{N-1,1,3} + \beta_{1,N-1} A_{N-1,3,3} \\ \alpha_{2,0} A_{0,1,1} + \beta_{2,0} A_{0,3,1} & \alpha_{2,0} A_{0,1,2} + \beta_{2,0} A_{0,3,2} & \cdots & \alpha_{2,N-1} A_{N-1,1,3} + \beta_{2,N-1} A_{N-1,3,3} \end{pmatrix} \tag{12.12}$$

And assume RMM time series 1 and 2 are decorrelated with each other. Then the observation error matrix is a diagonal matrix. Different choices of observation error are discussed in Section 5.1.

References

[1] James J Benedict, Eric D Maloney, Adam H Sobel, Dargan M Frierson, and Leo J Donner. Tropical intraseasonal variability in version 3 of the gfdl atmosphere model. *Journal of Climate*, 26(2):426–449, 2013.

[2] J. A. Biello and A. J. Majda. A new multiscale model for the Madden–Julian oscillation. *J. Atmos. Sci.*, 62:1694–1721, June 2005.

[3] M. Branicki and A. J. Majda. An information-theoretic framework for improving imperfect dynamical predictions via multi-model ensemble forecasts. *J. Nonlin. Sci.*, 25(3):489–538, 2015.

[4] ND Brenowitz, Y Frenkel, and AJ Majda. Non-local convergence coupling in a simple stochastic convection model. *Dynamics of Atmospheres and Oceans*, 74:30–49, 2016.

[5] Gilbert Brunet, Melvyn Shapiro, Brian Hoskins, Mitch Moncrieff, Randall Dole, George N Kiladis, Ben Kirtman, Andrew Lorenc, Brian Mills, Rebecca Morss, et al. Collaboration of the weather and climate communities to advance subseasonal-to-seasonal prediction. *Bulletin of the American Meteorological Society*, 91(10):1397–1406, 2010.

[6] E. Castronovo, J. Harlim, and A. J. Majda. Mathematical test criteria for filtering complex systems: plentiful observations. *Journal of Computational Physics*, 227(7), 2008.

[7] D. Chen and X. Yuan. A Markov model for seasonal forecast of Antarctic sea ice. *J. Climate*, 17(16):3156–3168, 2004.

[8] N. Chen and A. J. Majda. Filtering the stochastic skeleton model for the Madden–Julian oscillation. *Mon. Wea. Rev.*, 144(2):501–527, 2016.

[9] N. Chen and A. J. Majda. Predicting the cloud patterns for the Boreal summer intraseasonal oscillation through a low-order stochastic model. *Math. Clim. Weather Forecast.*, 1:1–20, 2016.

[10] N. Chen, A. J. Majda, and D. Giannakis. Predicting the cloud patterns of the Madden–Julian Oscillation through a low-order nonlinear stochastic model. *Geophys. Res. Lett.*, 41(15):5612–5619, 2014.

[11] S. Chen, A. J. Majda, and S. N. Stechmann. Multiscale asymptotics for the skeleton of the Madden–Julian oscillation and tropical–extratropical interactions. *Math. Clim. Weather Forecast.*, 1:43–69, 2015.

[12] S. Chen, A. J. Majda, and S. N. Stechmann. Tropical–extratropical interactions with the MJO skeleton and climatological mean flow. *J. Atmos. Sci.*, 73(10):4101–4116, 2016.

[13] S. Chen and S. N. Stechmann. Nonlinear traveling waves for the skeleton of the Madden–Julian oscillation. *Comm. Math. Sci.*, 14:571–592, 2016.

[14] Q. Deng, B. Khouider, and A. J. Majda. The MJO in a coarse-resolution GCM with a stochastic multicloud parameterization. *J. Atmos. Sci.*, 72:55–74, 2015.

[15] Q. Deng, B. Khouider, A. J. Majda, and R. S. Ajayamohan. Effect of stratiform heating on the planetary-scale organization of tropical convection. *J. Atmos. Sci.*, 73(1):371–392, 2016.

[16] Gregory S Duane, Joseph J Tribbia, and Jeffrey B Weiss. Synchronicity in predictive modelling: a new view of data assimilation. *Nonlinear processes in Geophysics*, 13(6):601–612, 2006.

[17] Chris Fraley, Adrian E Raftery, and Tilmann Gneiting. Calibrating multimodel forecast ensembles with exchangeable and missing members using bayesian model averaging. *Monthly Weather Review*, 138(1):190–202, 2010.

[18] D. M. W. Frierson, A. J. Majda, and O. M. Pauluis. Large scale dynamics of precipitation fronts in the tropical atmosphere: a novel relaxation limit. *Commun. Math. Sci.*, 2(4):591–626, 2004.

[19] B. Gershgorin, J. Harlim, and A. J. Majda. Improving filtering and prediction of spatially extended turbulent systems with model errors through stochastic parameter estimation. *J. Comput. Phys.*, 229(1):32–57, 2010.

[20] B. Gershgorin, J. Harlim, and A. J. Majda. Test models for improving filtering with model errors through stochastic parameter estimation. *J. Comput. Phys.*, 229(1):1–31, 2010.

[21] Lisa Goddard, Simon J Mason, Stephen E Zebiak, Chester F Ropelewski, Reid Basher, and Mark A Cane. Current approaches to seasonal to interannual climate predictions. *International Journal of Climatology*, 21(9):1111–1152, 2001.

[22] B. B. Goswami, B. Khouider, R. Phani, P. Mukhopadhyay, and A. Majda. Improving synoptic and intraseasonal variability in CFSv2 via stochastic representation of organized convection. *Geophys. Res. Lett.*, 44(2):1104–1113, 2017.

[23] B. B. Goswami, B. Khouider, R. Phani, P. Mukhopadhyay, and A. J. Majda. Implementation and calibration of a stochastic multicloud convective parameterization in the NCEP Climate Forecast System (CFSv2). *J. Adv. Model. Earth Syst.*, 9:1721–1739, 2017.

[24] B. B. Goswami, B. Khouider, R. Phani, P. Mukhopadhyay, and A. J. Majda. Improved tropical modes of variability in the NCEP Climate Forecast System (Version 2) via a stochastic multicloud model. *J. Atmos. Sci.*, 74(10):3339–3366, 2017.

[25] J. Gottschalck, M. Wheeler, K. Weickmann, F. Vitart, N. Savage, H. Lin, H. Hendon, D. Waliser, and K. Sperber. A framework for assessing operational Madden–Julian Oscillation forecasts: A CLIVAR MJO Working Group project. *Bull. Amer. Meteor. Soc.*, 91(9):1247–1258, 2010.

[26] W. W. Grabowski and P. K. Smolarkiewicz. CRCP: a Cloud Resolving Convection Parameterization for modeling the tropical convecting atmosphere. *Physica D: Nonlinear Phenomena*, 133:171–178, 1999.

[27] E. Guilyardi, A. Wittenberg, A. Fedorov, M. Collins, C. Wang, A. Capotondi, G. J. Van Oldenborgh, and T. Stockdale. Understanding El Niño in ocean–atmosphere general circulation models: Progress and challenges. *Bull. Amer. Meteor. Soc.*, 90(3):325–340, 2009.

[28] Renate Hagedorn, F. J. Doblas-Reyes, and TN Palmer. The rationale behind the success of multi-model ensembles in seasonal forecasting–i. basic concept. *Tellus A*, 57(3):219–233, 2005.

[29] J. Harlim and A. J. Majda. Filtering nonlinear dynamical systems with linear stochastic models. *Nonlinearity*, 21(6):1281–1306, 2008.

[30] J. Harlim and A. J. Majda. Catastrophic filter divergence in filtering nonlinear dissipative systems. *Comm. Math. Sci.*, 8(1):27–43, 2010.

[31] J. Harlim and A. J. Majda. Filtering turbulent sparsely observed geophysical flows. *Monthly Weather Review*, 138(4):1050–1083, 2010.

[32] J. Harlim and A. J. Majda. Test models for filtering and prediction of moisture-coupled tropical waves. *Q. J. Royal Meteorol. Soc.*, 139(670):119–136, 2013.

[33] S. Hottovy and S. N. Stechmann. A spatiotemporal stochastic model for tropical precipitation and water vapor dynamics. *J. Atmos. Sci.*, 72:4721–4738, 2015.

[34] M.-P. Hung, J.-L. Lin, W. Wang, D. Kim, T. Shinoda, and S. J. Weaver. MJO and convectively coupled equatorial waves simulated by CMIP5 climate models. *J. Climate*, 26(17):6185–6214, 2013.

[35] Y.-T. Hwang and D. M. W. Frierson. Link between the double-intertropical convergence zone problem and cloud biases over the southern ocean. *Proc. Natl. Acad. Sci. USA*, page 201213302, 2013.

[36] Xianan Jiang, Duane E Waliser, Prince K Xavier, Jon Petch, Nicholas P Klingaman, Steven J Woolnough, Bin Guan, Gilles Bellon, Traute Crueger, Charlotte DeMott, et al. Vertical structure and physical processes of the madden-julian oscillation: Exploring key model physics in climate simulations. *Journal of Geophysical Research: Atmospheres*, 120(10):4718–4748, 2015.

[37] E. K. Jin, J. L. Kinter, B. Wang, C.-K. Park, I.-S. Kang, B. P. Kirtman, J.-S. Kug, A. Kumar, J.-J. Luo, J. Schemm, et al. Current status of ENSO prediction skill in coupled ocean–atmosphere models. *Climate Dynamics*, 31(6):647–664, 2008.

[38] C. Jones, L. M. V. Carvalho, R. Wayne Higgins, D. E. Waliser, and J. K. E. Schemm. A statistical forecast model of tropical intraseasonal convective anomalies. *J. Climate*, 17(11):2078–2095, 2004.

[39] C. Jones, D. E. Waliser, K. M. Lau, and W. Stern. The Madden–Julian oscillation and its impact on Northern Hemisphere weather predictability. *Mon. Wea. Rev.*, 132(6):1462–1471, 2004.

[40] E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, et al. The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, 77(3):437–471, 1996.

[41] M. F. Khairoutdinov and D. A. Randall. A cloud resolving model as a cloud parameterization in the NCAR Community Climate System Model: Preliminary results. *Geophys. Res. Lett.*, 28(18):3617–3620, 2001.

[42] Viatcheslav V Kharin and Francis W Zwiers. Climate predictions with multimodel ensembles. *Journal of Climate*, 15(7):793–799, 2002.

[43] B. Khouider and A. J. Majda. Multicloud convective parameterizations with crude vertical structure. *Theor. Comp. Fluid Dyn.*, 20:351–375, 2006.

[44] B. Khouider and A. J. Majda. A simple multicloud parameterization for convectively coupled tropical waves. Part I: Linear analysis. *J. Atmos. Sci.*, 63:1308–1323, 2006.

[45] B. Khouider and A. J. Majda. A simple multicloud parameterization for convectively coupled tropical waves. Part II: Nonlinear simulations. *J. Atmos. Sci.*, 64:381–400, 2007.

[46] B. Khouider, A. St-Cyr, A. J. Majda, and J. Tribbia. The MJO and convectively coupled waves in a coarse-resolution GCM with a simple multicloud parameterization. *J. Atmos. Sci.*, 68:240–264, 2011.

[47] G. N. Kiladis, K. H. Straub, and P. T. Haertel. Zonal and vertical structure of the Madden–Julian oscillation. *J. Atmos. Sci.*, 62:2790–2809, August 2005.

[48] Daehyun Kim, Adam H Sobel, Eric D Maloney, Dargan MW Frierson, and In-Sik Kang. A systematic relationship between intraseasonal variability and mean state bias in agcm simulations. *Journal of Climate*, 24(21):5506–5520, 2011.

[49] H.-M. Kim, P. J. Webster, V. E. Toma, and D. Kim. Predictability and prediction skill of the MJO in two operational forecasting systems. *J. Climate*, 27(14):5364–5378, 2014.

[50] Ben Kirtman and Anna Pirani. The state of the art of seasonal prediction: Outcomes and recommendations from the first world climate research program workshop on seasonal prediction. *Bulletin of the American Meteorological Society*, 90(4):455–458, 2009.

[51] Ben P Kirtman, Dughong Min, Johnna M Infanti, James L Kinter III, Daniel A Paolino, Qin Zhang, Huug Van Den Dool, Suranjana Saha, Malaquias Pena Mendez, Emily Becker, et al. The north american multimodel ensemble: phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bulletin of the American Meteorological Society*, 95(4):585–601, 2014.

[52] Tiruvalam Natarajan Krishnamurti, CM Kishtawal, Zhan Zhang, Timothy LaRow, David Bachiochi, Eric Williford, Sulochana Gadgil, and Sajani Surendran. Multimodel ensemble forecasts for weather and seasonal climate. *Journal of Climate*, 13(23):4196–4216, 2000.

[53] TN Krishnamurti, CM Kishtawal, Timothy E LaRow, David R Bachiochi, Zhan Zhang, C Eric Williford, Sulochana Gadgil, and Sajani Surendran. Improved weather and seasonal climate forecasts from multimodel superensemble. *Science*, 285(5433):1548–1550, 1999.

[54] W. K. M. Lau and D. E. Waliser, editors. *Intraseasonal Variability in the Atmosphere–Ocean Climate System*. Springer, Berlin, 2nd edition, 2012.

[55] Y. Lee, A.J. Majda, and D. Qi. Preventing catastrophic filter divergence using adaptive additive inflation for baroclinic turbulence. *Mon. Wea. Rev.*, 145(2):669–682, 2017.

[56] B. Liebmann and C. A. Smith. Description of a complete (interpolated) outgoing longwave radiation dataset. *Bull. Amer. Meteor. Soc.*, 77:1275–1277, 1996.

[57] J.-L. Lin. The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean-atmosphere feedback analysis. *J. Climate*, 20(18):4497–4525, 2007.

[58] J.-L. Lin, G. N. Kiladis, B. E. Mapes, K. M. Weickmann, K. R Sperber, W. Lin, M. Wheeler, S. D. Schubert, A. Del Genio, L. J. Donner, S. Emori, J.-F. Gueremy, F. Hourdin, P. J. Rasch, E. Roeckner, and J. F. Scinocca. Tropical intraseasonal variability in 14 IPCC AR4 climate models Part I: Convective signals. *J. Climate*, 19:2665–2690, 2006.

[59] RW Lindsay, J Zhang, AJ Schweiger, and MA Steele. Seasonal predictions of ice extent in the Arctic Ocean. *Journal of Geophysical Research: Oceans*, 113(C2), 2008.

[60] A. J. Majda. *Introduction to PDEs and Waves for the Atmosphere and Ocean*, volume 9 of *Courant Lecture Notes in Mathematics*. American Mathematical Society, Providence, 2003.

[61] A. J. Majda and J. A. Biello. A multiscale model for the intraseasonal oscillation. *Proc. Natl. Acad. Sci. USA*, 101(14):4736–4741, 2004.

[62] A. J. Majda and M. J. Grote. Explicit off-line criteria for stable accurate time filtering of strongly unstable spatially extended systems. *Proc. Natl. Acad. Sci. USA*, 104:1124–1129, 2007.

[63] A. J. Majda and M. J. Grote. Mathematical test models for superparametrization in anisotropic turbulence. *Proc. Natl. Acad. Sci. USA*, 106(14):5470–5474, 2009.

[64] A. J. Majda and J. Harlim. *Filtering Turbulent Complex Systems*. Cambridge University Press, 2012.

[65] A. J. Majda and S. N. Stechmann. The skeleton of tropical intraseasonal oscillations. *Proc. Natl. Acad. Sci. USA*, 106(21):8417–8422, 2009.

[66] A. J. Majda and S. N. Stechmann. Nonlinear dynamics and regional variations in the MJO skeleton. *J. Atmos. Sci.*, 68:3053–3071, 2011.

[67] Andrew J Majda and John Harlim. *Filtering complex turbulent systems*. Cambridge University Press, 2012.

[68] A. McGovern, K. L. Elmore, D. J. Gagne, S. E. Haupt, C. D. Karstens, R. Lagerquist, T. Smith, and J. K. Williams. Using artificial intelligence to improve real-time decision-making for high-impact weather. *Bull. Amer. Meteorol. Soc.*, 98(10):2073–2090, 2017.

[69] M. W. Moncrieff, M. Shapiro, J. Slingo, and F. Molteni. Collaborative research at the intersection of weather and climate. *WMO Bull.*, 56:204–211, 2007.

[70] J. David Neelin and Jia-Yuh Yu. Modes of tropical variability under convective adjustment and the Madden–Julian oscillation. Part 1: Analytical theory. *J. Atmos. Sci.*, 51:1876–1894, 1994.

[71] J. David Neelin and Ning Zeng. A quasi-equilibrium tropical circulation model—formulation. *J. Atmos. Sci.*, 57:1741–1766, 2000.

[72] J. M. Neena, J. Y. Lee, D. Waliser, B. Wang, and X. Jiang. Predictability of the Madden–Julian oscillation in the intraseasonal variability hindcast experiment (ISVHE). *J. Climate*, 27(12):4531–4543, 2014.

[73] H. R. Ogrosky and S. N. Stechmann. Assessing the equatorial long-wave approximation: asymptotics and observational data analysis. *J. Atmos. Sci.*, 72:4821–4843, 2015.

[74] H. R. Ogrosky and S. N. Stechmann. The MJO skeleton model with observation-based background state and forcing. *Q. J. Roy. Meteor. Soc.*, 141(692):2654–2669, 2015.

[75] H. R. Ogrosky and S. N. Stechmann. Identifying convectively coupled equatorial waves using theoretical wave eigenvectors. *Mon. Wea. Rev.*, page in press, 2016.

[76] H. R. Ogrosky, S. N. Stechmann, and A. J. Majda. Boreal summer intraseasonal oscillations in the MJO skeleton model with observation-based forcing. *Dyn. Atmos. Oceans*, 78:38–56, 2017.

[77] Angel C Otarola, Richard Querel, and Florian Kerber. Precipitable water vapor: Considerations on the water vapor scale height, dry bias of the radiosonde humidity sensors, and spatial and temporal variability of the humidity field. *arXiv preprint arXiv:1103.3025*, 2011.

[78] Tim N Palmer and David LT Anderson. The prospects for seasonal forecastinga review paper. *Quarterly Journal of the Royal Meteorological Society*, 120(518):755–793, 1994.

[79] B. Peherstorfer, K. Willcox, and M. Gunzburger. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Review*, 60(3):550–591, 2018.

[80] C. Penland and T. Magorian. Prediction of Niño 3 sea surface temperatures using linear inverse modeling. *J. Climate*, 6(6):1067–1076, 1993.

[81] C. Penland and P. D. Sardeshmukh. The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, 8(8):1999–2024, 1995.

[82] Andrew W Robertson, Arun Kumar, Malaquias Peña, and Frederic Vitart. Improving and promoting subseasonal to seasonal prediction. *Bulletin of the American Meteorological Society*, 96(3):ES49–ES53, 2015.

[83] T. Schneider, S. Lan, A. Stuart, and J. Teixeira. Earth system modeling 2.0: A blueprint for models that learn from observations and targeted high-resolution simulations. *Geophys. Res. Lett.*, 44(24), 2017.

[84] J. M. Slingo, K. R. Sperber, J. S. Boyle, J.-P. Ceron, M. Dix, B. Dugas, W. Ebisuzaki, J. Fyfe, D. Gregory, J.-F. Gueremy, J. Hack, A. Harzallah, P. Inness, A. Kitoh, W. K.-M. Lau, B. McAvaney, R. Madden, A. Matthews, T. N. Palmer, C.-K. Park, D. Randall, and N. Renno. Intraseasonal oscillations in 15 atmospheric general circulation models: results from an AMIP diagnostic subproject. *Climate Dyn.*, 12:325–357, 1996.

[85] J. P. Stachnik, D. E. Waliser, A. J. Majda, S. N. Stechmann, and S. Thual. Evaluating MJO event initiation and decay in the skeleton model using an RMM-like index. *J. Geophys. Res. Atmos.*, 120(22), 2015.

[86] S. N. Stechmann and S. Hottovy. Unified spectrum of tropical rainfall and waves in a simple stochastic model. *Geophys. Res. Lett.*, page in press, 2017.

[87] S. N. Stechmann and A. J. Majda. The structure of precipitation fronts for finite relaxation time. *Theor. Comp. Fluid Dyn.*, 20:377–404, 2006.

[88] S. N. Stechmann and A. J. Majda. Identifying the skeleton of the Madden–Julian oscillation in observational data. *Mon. Wea. Rev.*, 143:395–416, 2015.

[89] S. N. Stechmann, A. J. Majda, and B. Khouider. Nonlinear dynamics of hydrostatic internal gravity waves. *Theor. Comp. Fluid Dyn.*, 22:407–432, 2008.

[90] S. N. Stechmann and H. R. Ogrosky. The Walker circulation, diabatic heating, and outgoing longwave radiation. *Geophys. Res. Lett.*, 41:9097–9105, 2014.

[91] S. Thual, A. J. Majda, N. Chen, and S. N. Stechmann. Simple stochastic model for El Niño with westerly wind bursts. *Proceedings of the National Academy of Sciences*, 113(37):10245–10250, 2016.

[92] S. Thual, A. J. Majda, and S. N. Stechmann. A stochastic skeleton model for the MJO. *J. Atmos. Sci.*, 71:697–715, 2014.

[93] S. Thual, A. J. Majda, and S. N. Stechmann. Asymmetric intraseasonal events in the stochastic skeleton MJO model with seasonal cycle. *Clim. Dyn.*, 45:603–618, 2015.

[94] F Vitart, C Ardilouze, A Bonet, A Brookshaw, M Chen, C Codorean, M Déqué, L Ferranti, E Fucile, M Fuentes, et al. The subseasonal to seasonal (s2s) prediction project database. *Bulletin of the American Meteorological Society*, 98(1):163–173, 2017.

[95] D. Waliser. Predictability and forecasting. In W. K. M. Lau and D. E. Waliser, editors, *Intraseasonal Variability in the Atmosphere–Ocean Climate System*. Springer, Berlin, 2012.

[96] D. Waliser, K. Sperber, H. Hendon, D. Kim, E. Maloney, M. Wheeler, K. Weickmann, C. Zhang, L. Donner, J. Gottschalck, et al. MJO simulation diagnostics. *J. Climate*, 22:3006–3030, 2009.

[97] D. Waliser, K. Weickmann, R. Dole, S. Schubert, O. Alves, C. Jones, M. Newman, H.L. Pan, A. Roubicek, S. Saha, et al. The experimental MJO prediction project. *Bull. Amer. Meteor. Soc.*, 87(4):425–432, 2006.

[98] D. E. Waliser, C. Jones, J. K. E. Schemm, and N. E. Graham. A statistical extended-range tropical forecast model based on the slow evolution of the Madden–Julian oscillation. *J. Climate*, 12(7):1918–1939, 1999.

[99] D. E. Waliser, K. M. Lau, W. Stern, and C. Jones. Potential predictability of the Madden–Julian oscillation. *Bull. Amer. Meteor. Soc.*, 84(1):33–50, 2003.

[100] L. Wang, X. Yuan, M. Ting, and C. Li. Predicting summer Arctic sea ice concentration intraseasonal variability using a vector autoregressive model. *J. Climate*, 29(4):1529–1543, 2016.

[101] M. Wheeler and G. N. Kiladis. Convectively coupled equatorial waves: analysis of clouds and temperature in the wavenumber–frequency domain. *J. Atmos. Sci.*, 56(3):374–399, 1999.

[102] M. C. Wheeler and H. H. Hendon. An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, 132(8):1917–1932, 2004.

[103] Jeffrey S Whitaker, Xue Wei, and Frédéric Vitart. Improving week-2 forecasts with multimodel reforecast ensembles. *Monthly weather review*, 134(8):2279–2284, 2006.

[104] Zhiwei Wu, Bin Wang, Jianping Li, and Fei-Fei Jin. An empirical seasonal prediction model of the east asian summer monsoon using enso and nao. *Journal of Geophysical Research: Atmospheres*, 114(D18), 2009.

[105] Q. Yang and A. J. Majda. Upscale impact of mesoscale disturbances of tropical convection on synoptic-scale equatorial waves in two-dimensional flows. *J. Atmos. Sci.*, 74(9):3099–3120, 2017.

[106] Shu-Chih Yang, Debra Baker, Hong Li, Katy Cordes, Morgan Huff, Geetika Nagpal, Ena Okereke, Josue Villafane, Eugenia Kalnay, and Gregory S Duane. Data assimilation as synchronization of truth and model: Experiments with the three-variable lorenz system. *Journal of the atmospheric sciences*, 63(9):2340–2354, 2006.

[107] J.-Y. Yu and J. D. Neelin. Modes of tropical variability under convective adjustment and the Madden–Julian oscillation. Part II: Numerical results. *J. Atmos. Sci.*, 51(13):1895–1914, 1994.

[108] S. E. Zebiak and M. A. Cane. A model El Niño–Southern Oscillation. *Mon. Wea. Rev.*, 115(10):2262–2278, 1987.