



## RESEARCH ARTICLE

10.1002/2014JD022374

This article is a companion to Klingaman *et al.* [2015] doi:10.1002/2015JD023196; Xavier *et al.* [2015] doi:10.1002/2014JD022718; and Jiang *et al.* [2015], doi:10.1002/2014JD022375.

## Key Points:

- A strong link exists between model MJO fidelity and diabatic moistening
- The shape of a model's diabatic heating profile is not linked to its MJO skill
- No single process clearly distinguishes high- and low-fidelity models

## Correspondence to:

N. P. Klingaman,  
n.p.klingaman@reading.ac.uk

## Citation:

Klingaman, N. P., *et al.* (2015), Vertical structure and physical processes of the Madden-Julian oscillation: Linking hindcast fidelity to simulated diabatic heating and moistening, *J. Geophys. Res. Atmos.*, 120, 4690–4717, doi:10.1002/2014JD022374.

Received 4 AUG 2014

Accepted 6 APR 2015

Accepted article online 10 APR 2015

Published online 26 MAY 2015

## Vertical structure and physical processes of the Madden-Julian oscillation: Linking hindcast fidelity to simulated diabatic heating and moistening

Nicholas P. Klingaman<sup>1</sup>, Steven J. Woolnough<sup>1</sup>, Xianan Jiang<sup>2,3</sup>, Duane Waliser<sup>2,3</sup>, Prince K. Xavier<sup>4</sup>, Jon Petch<sup>4</sup>, Mihaela Caian<sup>5</sup>, Cecile Hannay<sup>6</sup>, Daehyun Kim<sup>7</sup>, Hsi-Yen Ma<sup>8</sup>, William J. Merryfield<sup>9</sup>, Tomoki Miyakawa<sup>10,11</sup>, Mike Pritchard<sup>12</sup>, James A. Ridout<sup>13</sup>, Romain Roehrig<sup>14</sup>, Eiki Shindo<sup>15</sup>, Frederic Vitart<sup>16</sup>, Hailan Wang<sup>17</sup>, Nicholas R. Cavanaugh<sup>18</sup>, Brian E. Mapes<sup>19</sup>, Ann Shelly<sup>4</sup>, and Guang J. Zhang<sup>18</sup>

<sup>1</sup>National Centre for Atmospheric Science and Department of Meteorology, University of Reading, Reading, UK, <sup>2</sup>Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, California, USA, <sup>3</sup>Jet Propulsion Laboratory, Pasadena, California, USA, <sup>4</sup>UK Met Office, Exeter, UK, <sup>5</sup>Rosby Centre, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden, <sup>6</sup>National Center for Atmospheric Research, Boulder, Colorado, USA, <sup>7</sup>Department of Atmospheric Sciences, University of Washington, Seattle, Washington, USA, <sup>8</sup>Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, California, USA, <sup>9</sup>Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria, British Columbia, Canada, <sup>10</sup>Department of Coupled Ocean-Atmosphere-Land Process Research, Japan Agency for Marine-Earth Science and Technology, Yokosuka, Japan, <sup>11</sup>Atmosphere and Ocean Research Institute, Tokyo University, Tokyo, Japan, <sup>12</sup>Department of Earth System Sciences, University of California, Irvine, California, USA, <sup>13</sup>Naval Research Laboratory, Monterey, California, USA, <sup>14</sup>CNRM-GAME, Météo-France and CNRS, Toulouse, France, <sup>15</sup>Meteorological Research Institute, Tsukuba, Japan, <sup>16</sup>European Centre for Medium-Range Weather Forecasts, Reading, UK, <sup>17</sup>National Aeronautics and Space Administration Goddard Space Flight Center, Greenbelt, Maryland, USA, <sup>18</sup>Scripps Institution of Oceanography, La Jolla, California, USA, <sup>19</sup>Rosenstiel School of Marine and Atmospheric Science, University of Miami, Coral Gables, Florida, USA

**Abstract** Many theories for the Madden-Julian oscillation (MJO) focus on diabatic processes, particularly the evolution of vertical heating and moistening. Poor MJO performance in weather and climate models is often blamed on biases in these processes and their interactions with the large-scale circulation. We introduce one of the three components of a model evaluation project, which aims to connect MJO fidelity in models to their representations of several physical processes, focusing on diabatic heating and moistening. This component consists of 20 day hindcasts, initialized daily during two MJO events in winter 2009–2010. The 13 models exhibit a range of skill: several have accurate forecasts to 20 days lead, while others perform similarly to statistical models (8–11 days). Models that maintain the observed MJO amplitude accurately predict propagation, but not vice versa. We find no link between hindcast fidelity and the precipitation-moisture relationship, in contrast to other recent studies. There is also no relationship between models' performance and the evolution of their diabatic heating profiles with rain rate. A more robust association emerges between models' fidelity and net moistening: the highest-skill models show a clear transition from low-level moistening for light rainfall to midlevel moistening at moderate rainfall and upper level moistening for heavy rainfall. The midlevel moistening, arising from both dynamics and physics, may be most important. Accurately representing many processes may be necessary but not sufficient for capturing the MJO, which suggests that models fail to predict the MJO for a broad range of reasons and limits the possibility of finding a panacea.

### 1. Introduction

The Madden-Julian oscillation (MJO) [Madden and Julian, 1972] is a key driver of tropical and extratropical circulation and precipitation variability on subseasonal scales. The MJO comprises quasiperiodic fluctuations between active and suppressed convective states in the tropics. MJO active events typically form in the equatorial Indian Ocean, grow to become large-scale envelopes of convective activity that often span several thousand kilometers zonally, then propagate east through the Maritime Continent to the West Pacific with a phase speed of approximately  $5 \text{ m s}^{-1}$  [e.g., Madden and Julian, 1994; Zhang, 2005]. Suppressed conditions (e.g., strong surface insolation, light winds, warming sea surface temperatures (SSTs)) precede and follow

©2015. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

active conditions. As these envelopes of enhanced and suppressed convection propagate, they modulate monsoon rainfall, particularly in the Australian monsoon during austral summer [e.g., *Hendon and Liebmann, 1990; Robertson et al., 2006; Wheeler et al., 2009*] and the South Asian monsoon during boreal summer [e.g., *Yasunari, 1979; Sperber et al., 2000; Annamalai and Sperber, 2005; Joseph et al., 2009; Pai et al., 2011*]. The circulation response to the MJO-associated diabatic heating drives wave trains that extend into the extratropics in both hemispheres, creating teleconnections between the MJO and the North Atlantic Oscillation [*Vitart and Molteni, 2010; Lin and Brunet, 2011*] and the Southern Annular Mode [*Flatau and Kim, 2013*]; the MJO may even influence Arctic sea ice [*Henderson et al., 2014*].

General circulation models (GCMs) used for numerical weather prediction (NWP) and climate simulations frequently suffer from weak subseasonal variability in tropical convection, poor spatial organization, and deficient eastward propagation [e.g., *Slingo et al., 1996; Lin et al., 2006; Sperber and Annamalai, 2008; Kim et al., 2009; Hung et al., 2013*]. Although MJO prediction skill in NWP has improved in some GCMs recently [e.g., *Waliser, 2012; Vitart, 2014*], the prediction skill of most GCMs lags predictability by at least 10 days [*Neena et al., 2014*]. These biases have come under increasing scrutiny recently, as modeling centers focus on the performance of their subseasonal (2–6 week) prediction systems, which fill a critical gap between medium-range and seasonal forecasts [*Gottschalck et al., 2010*]. Attempts to improve GCM representations of the MJO have centered on the subgrid-scale parameterization of convection, including adding or increasing the sensitivity of a trigger based on low-level moisture convergence [e.g., *Wang and Schlesinger, 1999*] or increasing the sensitivity of a convective parcel to environmental humidity [e.g., *Bechtold et al., 2008; Chikira and Sugiyama, 2010; Subramanian et al., 2011; Klingaman and Woolnough, 2014*]; producing more upper level heating from stratiform precipitation to increase the covariance between heating and temperature anomalies [*Fu and Wang, 2009; Seo and Wang, 2010*], based on similar instabilities found for convectively coupled waves [e.g., *Mapes, 2000*]; and introducing momentum mixing by convection [e.g., *Zhou et al., 2012*]. At a basic level, many of these studies emphasize the importance of promoting instabilities between certain components of the GCM diabatic heating profiles and the large-scale circulation, which are often damped in models with a weak MJO. However, the targeted components of the heating profiles vary among the studies. Further, the improvements found often depend on the GCM, or even on the particular GCM version, limiting their applicability and confounding parameterization development efforts.

Advances in GCM parameterization development are also complicated by a wide range of hypotheses, based on observations or on theoretical studies with reduced-complexity models, concerning which component of the MJO-associated diabatic heating profile is most critical to promoting and maintaining convective instability. These hypotheses have targeted free-tropospheric moisture convergence from low-level heating [e.g., *Lau and Peng, 1987*], frictional moisture convergence in the boundary layer [e.g., *Wang and Rui, 1990*], local cycles of moisture discharge and recharge [e.g., *Bladé and Hartmann, 1993*], and cloud-radiation interactions and the radiative heating [e.g., *Slingo and Madden, 1991; Hu and Randall, 1994; Raymond, 2001; Stephens et al., 2004*]. There are considerable disagreements among observations and reanalysis products in the shape and amplitude of MJO-associated heating profiles, even among products processed from the same Tropical Rainfall Measuring Mission instruments [*Ling and Zhang, 2011; Jiang et al., 2011*], which likely result not only from discrepancies among the data sets but also from sensitivities to MJO compositing method. While reanalyses typically show a heating structure that tilts westward with height—producing low-level heating and moisture convergence east of the strongest convection [e.g., *Lappen and Schumacher, 2012*—this tilt is missing in some observed products. Recent modeling work [*Lappen and Schumacher, 2014*] suggests that this low-level heating from shallow convection may be more critical to the representation of the MJO than the “top-heavy” heating, caused by strong stratiform components, that past studies emphasized.

Over the past few years, observational and modeling studies have focused attention on the role of the MJO-suppressed phase in maintaining the oscillation. Using observations and reanalysis, *Kim et al. [2014a]* found that Indian Ocean active phases propagated east more frequently, and more coherently, when the suppressed phase in the West Pacific was stronger, causing enhanced meridional moisture advection east of the active phase. *Woolnough et al. [2010]* analyzed cloud-resolving model (CRM) and single-column model (SCM) simulations of suppressed conditions during the Tropical Ocean–Global Atmosphere (TOGA) Coupled Ocean–Atmosphere Response Experiment (TOGA COARE) [*Webster and Yang, 1992*]. SCMs with less precipitation during the suppressed phase maintained a moister atmosphere and produced a more gradual transition from shallow to deep convection, in closer agreement with the CRMs; SCMs with more precipitation in the suppressed phase showed a sharper transition to deep convection that was delayed relative to the CRMs and

the SCMs with less precipitation. The SCMs with more precipitation in the suppressed phase also produced less precipitation in the active phase, implying weaker intraseasonal variability. Most GCMs also produce a too sharp transition to active conditions, which studies have suggested is caused by a poor relationship between convection and environmental moisture, particularly in the midtroposphere where GCMs often exhibit a mean dry bias, perhaps from too little convective moistening [e.g., *Thayer-Calder and Randall, 2009; Xavier, 2012*].

In this study, we link the fidelity of the simulated MJO in 13 GCMs (section 2.2) to their simulated diabatic heating and moistening associated with tropical convection. These 13 GCMs have participated in a novel model evaluation project (section 2.1), providing high-frequency output of heating and moistening tendencies from resolved and subgrid-scale processes (section 2.4). This study focuses on 20 day hindcast simulations of two strong MJO events (section 2.2.1) to analyze how the accuracy of models' MJO predictions relates to the vertical structure of their diabatic processes. This model evaluation project is the first to collect such detailed and high-frequency output from many GCMs, providing a unique opportunity to compare GCMs at the level of individual physical processes.

After describing the hindcast component of the evaluation project in section 2, we assess the fidelity of MJO predictions in these GCMs for these two events in section 3.1; section 3.2 examines precipitation-moisture and precipitation-vertical velocity sensitivities in these GCMs, which have been hypothesized to be important for representing the MJO [e.g., *Thayer-Calder and Randall, 2009*]; we analyze the heating and moistening profiles as a function of rain rate in sections 3.4 and 3.5, respectively. We discuss our findings and experiment design in section 4 before presenting our key conclusions in section 5.

## 2. Experiment Design and Data Sets

We analyze one of the three components of the "Vertical structure and diabatic processes of the MJO" global model evaluation project. The project is organized and supported by the Global Atmospheric Systems Studies (GASS) panel of the Global Energy and Water Exchanges (GEWEX) and the World Climate Research Program-World Weather Research Program/The Observing System Research and Predictability Experiment-Years of Tropical Convection (WCRP-WWRP/THORPEX-YOTC) MJO Task Force. (At the time of this study, the MJO Task Force was under the collective auspices of the World Climate Research Programme (WCRP), the World Weather Research Programme (WWRP), The Observing System Research and Predictability Experiment (THORPEX), and the Years of Tropical Convection (YOTC). It is currently under the auspices of the Working Group on Numerical Experimentation (WGNE).)

### 2.1. Project Design

The project aims to characterize the representation of the diabatic heating and moistening processes associated with the MJO in GCMs, including assessing those processes against analyses and observations [*Petch et al., 2011*]. The project also aims to link GCM representations of these processes to the quality of MJO simulation. Detailed analysis of diabatic processes requires a bespoke experimental design, since more general intercomparisons (e.g., the Coupled Model Intercomparison Project) do not archive subdaily parameterization tendencies.

The project comprises three-component experiments that take advantage of known links between biases in short-range weather forecasts and long-term climate simulations [e.g., *Boyle et al., 2008; Xie et al., 2012; Ma et al., 2013, 2014*]. The three experiments are (1) 20 year simulations, using either atmosphere-only GCMs forced by observed SSTs or atmosphere-ocean coupled GCMs; (2) a series of 2 day hindcasts, initialized once per day, for two strong MJO events (section 2.2) during the Year of Tropical Convection (YOTC) [*Waliser et al., 2012*]; and (3) as in (2) but for 20 day hindcasts. The goals of (1) are to determine the ability of each GCM to represent the MJO when the GCM is near its preferred climate, as well as its ability to generate MJO events [*Jiang et al., 2015*]. Component (2), which acquired time step frequency GCM output, enables a detailed investigation of the behavior of individual physics schemes when the GCM is in a near-observed state that contains an active MJO [*Xavier et al., 2015*]. Component (3) is the focus of this manuscript and seeks to link (1) and (2). By using 20 day hindcasts with a broader range of initial dates than (2), but acquiring three-hourly output (section 2.4), this component was designed to connect degradations in MJO performance with lead time to biases in simulated diabatic processes. *Klingaman et al. [2015]* synthesize and summarize the conclusions of the entire project.

**Table 1.** For Each GCM: the Name, Version, Abbreviation for the Text, Code for the Figures, Contributing Centers, Native Horizontal Resolution (Lon × Lat or Wave Number Truncation (*T*) With Degree Equivalent), Number of Vertical Points (*L*), and Reference

Model Name	Version	Abbreviation	Code	Modeling Center	Native Resolution	Reference
Community Atmospheric Model	5	CAM5	C5	NCAR <sup>a</sup>	1.25° × 0.9°, L30	Neale <i>et al.</i> [2012]
CAM5 convective microphysics	5.1	CAM5-ZM	CZ	UCSD <sup>b</sup> and LLNL <sup>c</sup>	1.25° × 0.9°, L30	Song <i>et al.</i> [2012]
Canadian Coupled Model	4	CanCM4	CC	CCCma <sup>d</sup>	T63 (1.9°), L35	Merryfield <i>et al.</i> [2013]
CNRM <sup>e</sup> Atmospheric Model	5.2	CNRM-AM	CN	CNRM <sup>e</sup>	T127 (1.4°), L31	Voltaire <i>et al.</i> [2013]
European Community Model	3	ECEarth3	E3	SMHI <sup>f</sup>	T255 (0.7°), L91	Hazeleger <i>et al.</i> [2012]
Goddard Earth Observing System	5	GEOS5	NA	NASA <sup>g</sup>	0.625° × 0.5°, L72	Rienecker <i>et al.</i> [2008]
Goddard Institute for Space Studies GCM	E2'	GISS-E2	GI	NASA <sup>g</sup>	2° × 2.5°, L40	Schmidt <i>et al.</i> [2014]
Integrated Forecast System	37r3	IFS	EC	ECMWF <sup>h</sup>	T159 (1.125° ×), L91	Vitart <i>et al.</i> [2007]
Met Office Unified Model	GA3	MetUM-GA3	MO	MOHC <sup>i</sup>	0.833° × 0.556°, L70	Walters <i>et al.</i> [2011]
Model for Interdisciplinary Research on Climate	5	MIROC5	MI	AORI, NIES, and JAMSTEC <sup>j</sup>	T85 (1.4°), L40	Watanabe <i>et al.</i> [2010]
MRI <sup>k</sup> Atmospheric GCM	3	MRI-AGCM3	MR	MRI <sup>k</sup>	TL159 (1.125°), L48	Yukimoto <i>et al.</i> [2012]
Navy Global Environmental Model	1.0	NavGEM1	NR	NRL <sup>l</sup>	T359 (0.42°), L42	Hogan <i>et al.</i> [2014]
Nonhydrostatic Icosahedral Atmospheric Model	2011	NICAM	NI	JAMSTEC, AORI <sup>j</sup> , and RIKEN <sup>m</sup>	Glevel 9 (14 km), L40	Sato <i>et al.</i> [2008]
Superparameterized CAM	3.0	SPCAM3	SP	University of Washington	T42 (2.8°), L30	Khairoutdinov <i>et al.</i> [2008]

<sup>a</sup>National Center for Atmospheric Research.

<sup>b</sup>University of California, San Diego.

<sup>c</sup>Lawrence Livermore National Laboratory.

<sup>d</sup>Canadian Centre for Climate Modelling and Analysis.

<sup>e</sup>Centre National de Recherches Météorologiques.

<sup>f</sup>Swedish Meteorological and Hydrological Institute.

<sup>g</sup>National Aeronautics and Space Administration.

<sup>h</sup>European Centre for Medium-Range Weather Forecasts.

<sup>i</sup>Met Office Hadley Centre.

<sup>j</sup>Atmosphere-Ocean Research Institute, National Institute for Environmental Studies, and Japan Agency for Marine-Earth Technology (JAMSTEC).

<sup>k</sup>Meteorological Research Institute.

<sup>l</sup>Naval Research Laboratory, USA.

<sup>m</sup>RIKEN Advanced Institute for Computational Science, Japan.

## 2.2. Experiment Design

### 2.2.1. Main Set of Dynamical Models

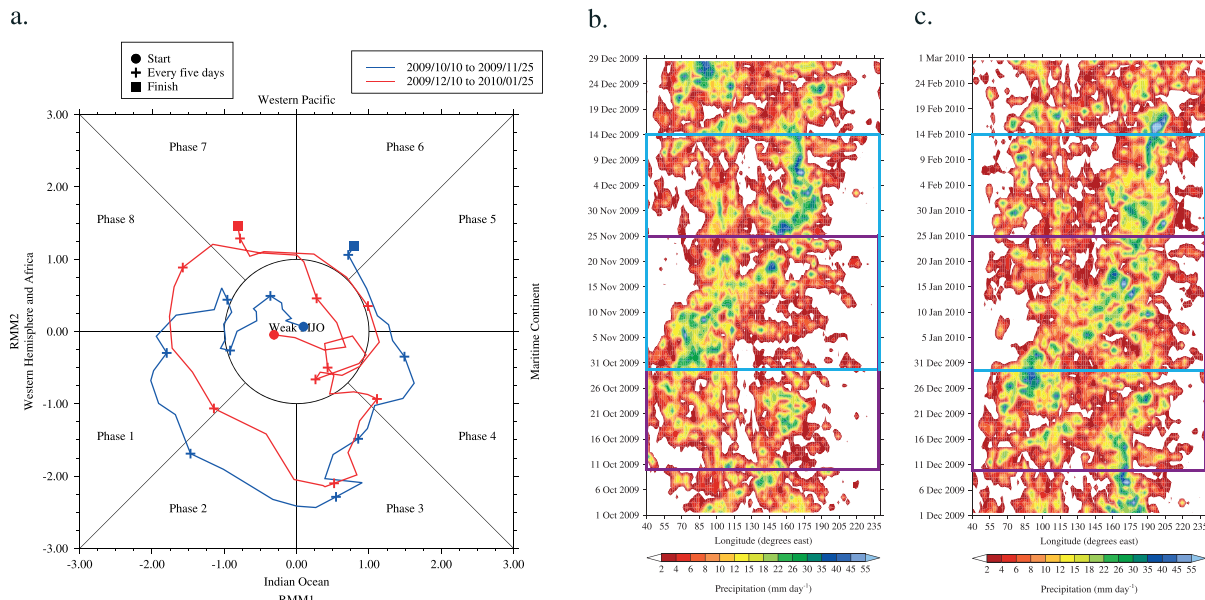
The GCMs participating in the 20 day hindcast component are listed in Table 1, along with the modeling center that submitted the simulations, the native resolutions, and a reference that provides further details on GCM formulation and parameterizations. Table 1 lists only those GCMs for which we analyze diabatic processes; other contributions are described in section 2.2.2.

Except as noted below, all GCMs performed 20 day hindcasts of two strong MJO events in boreal winter 2009–2010. For the first event (YOTC “Case E”) hindcasts were initialized every 00Z from 10 October to 25 November 2009; for the second event (YOTC “Case F”) hindcasts were initialized every 00Z from 10 December 2009 to 25 January 2010. There are 47 start dates per case. Figure 1a shows “observed” [Wheeler and Hendon, 2004] real-time multivariate MJO (RMM) indices for these periods; see section 2.5 for the computation method. Tropical Rainfall Measuring Mission (TRMM) rainfall (section 2.6) for these events is shown in Figures 1b and 1c, with the range of initial and final dates for the hindcasts indicated. The initial dates were chosen so that the genesis and lysis of the events were captured at least at 10 days lead time. All hindcasts were initialized from 00Z European Centre for Medium-Range Weather Forecasts (ECMWF) analyses produced for YOTC at 16 km horizontal resolution (ECMWF-YOTC) [Moncrieff *et al.*, 2012]. Because of the computational expense of running the Nonhydrostatic Icosahedral Atmospheric Model (NICAM), only six hindcasts were performed per event, spaced approximately every 5 days. Due to this limited sample, we do not compute MJO prediction skill (section 3.1) for NICAM; Miyakawa *et al.* [2014] estimated that NICAM displayed 27 days skill—at a threshold correlation of 0.6—using all winter MJO cases in 2003–2012.

### 2.2.2. Additional Dynamical Model Contributions

Additional contributions were received from the UK Met Office, which, as well as the standard 20 day atmosphere-only hindcasts (Table 1), also provided 15 day MetUM-GA3 hindcasts in atmosphere-only and





**Figure 1.** (a) Daily real-time multivariate MJO indices for the range of hindcast start dates for YOTC Case E (blue) and YOTC Case F (red), as computed from ECMWF-YOTC analysis winds and NOAA OLR (section 2.5), with crosses spaced every 5 days. TRMM 3B42 precipitation analyses for (b) Case E and (c) Case F. The purple boxes in Figures 1b and 1c show the range of initialization dates for the 20 day hindcasts, while the blue boxes show the dates captured at a 20 day forecast lead time.

atmosphere-ocean coupled configurations; the latter uses the Nucleus for European Modelling of the Ocean GCM with 0.25° horizontal resolution and 75 points in the vertical. The 15 day hindcasts use the same atmospheric horizontal resolution as the 20 day hindcasts but 85 vertical points instead of 70; there are also several minor physics changes between the configurations. The 15 day hindcasts were initialized from Met Office operational analyses, so they are not directly comparable to the other hindcasts in this study. We assess them for prediction skill (section 3.1) as a sensitivity test of the role of air-sea coupling. A detailed analysis of these simulations appears in *Shelly et al.* [2014].

**2.2.3. Linear Inverse Models**

Linear inverse model (LIM) forecasts were received from the University of Miami (Miami LIM) (B. Mapes, personal communication, 2013; B. Mapes, manuscript in preparation) and Scripps Institute of Oceanography (Scripps LIM) [*Cavanaugh et al.*, 2014]. We use the LIM forecasts to provide a more skillful baseline for the dynamical GCMs than simple persistence. In a LIM, anomalies propagate and decay according to a linear operator capable only of exponential (decaying sinusoidal) behavior. In the Miami LIM, this operator was built from 0 day and 2 day lagged covariance matrices of daily time-longitude sections of high-pass anomalies in 1979–2005, using all seasons. Input data were daily 15°S–15°N averages of National Centers for Environmental Prediction-National Center for Atmospheric Research (NCEP-NCAR) reanalysis 850 hPa and 200 hPa zonal wind, plus National Oceanic and Atmospheric Administration (NOAA) outgoing longwave radiation (OLR), all in 10° longitude bins. High-pass anomalies were defined by subtracting the composite seasonal cycle and the mean of the previous 120 days, as in *Wheeler and Hendon* [2004], to eliminate false skill from seasonally locked or low-frequency variability.

The Scripps LIM differs from the Miami LIM chiefly in that (a) it also uses meridional winds at 850 hPa and 200 hPa, in addition to OLR and zonal winds; (b) it uses grid point fields, rather than latitude averages; and (c) it was trained on 7 day smoothed fields, rather than daily data. Further details can be found in *Cavanaugh et al.* [2014]. The LIM used here is one of the best performing LIMs from that study: it uses 24 empirical orthogonal functions (EOFs) of wind fields, four EOFs of OLR, and a training lead time of 8 days. We use both the Miami and Scripps LIMs here because they differ in their skill over the short hindcast period in this study (section 3.1) and so provide a rudimentary estimate of the uncertainty in statistical skill.

**2.3. Further Model Details**

Here we provide further details on GCM configurations beyond those in Table 1, as well as differences in GCM configurations from the references provided.

CanCM4 is the only atmosphere-ocean coupled model that produced 20 day hindcasts. CanCM4 uses the Canadian Atmospheric Model (CanAM4) [von Salzen *et al.*, 2013] and the Canadian Ocean Model (CanOM4) [Merryfield *et al.*, 2013]. Rather than initializing from ECMWF-YOTC analyses, to maintain atmosphere-ocean coupled balance, CanCM4 hindcasts are initialized from analyses produced by assimilating six-hourly ECMWF-Interim reanalysis (ERA-Interim) atmospheric temperature, horizontal winds, and specific humidity into CanAM4, with a spectral cutoff at wave number 21 (T21). During the assimilation, daily mean CanAM4 surface forcing is applied to CanOM4; additionally, the CanOM4 SSTs are relaxed to the NCEP optimally interpolated SST analysis with a 3 day time constant. The 20 day hindcasts are initialized from this coupled assimilation system and then run freely (i.e., with no relaxation or assimilation). The same procedure is used operationally at the Canadian Centre for Climate Modelling and Analysis but assimilating operational analyses rather than ERA-Interim. We discuss the implications of the discrepancy in the initial conditions in section 4.

CAM5 was configured with a finite volume dynamical core. CAM5-ZM differs from CAM5 in the addition of the *Song and Zhang* [2011] convective microphysics to the *Morrison and Gettelman* [2008] stratiform microphysics scheme from CAM5. CNRM-AM is the atmospheric component of CNRM-CM5 used in CMIP5 simulations and described in *Volz and Coauthors* [2013]. The convection scheme in GISS-E2 was modified to improve tropical intraseasonal variability [Del Genio *et al.*, 2012; Kim *et al.*, 2012]. The NavGEM1 configuration used here, for which there is no published reference, differs from the configuration in *Hogan et al.* [2014] in that it lacks prognostic cloud water and uses the radiation scheme of *Harshvardhan et al.* [1987].

The choice of SST boundary condition was left to each modeling center. CanCM4 used its predicted SST; SPCAM3, MRI-AGCM3, NavGEM1, and CNRM-AM persisted the initial SST; CAM5, NICAM, Integrated Forecast System (IFS), ECEarth3, MetUM-GA3, and GISS-E2 persisted the initial SST anomaly with respect to a time-varying climatology; and CAM5-ZM, MIROC5, and Goddard Earth Observing System version 5 (GEOS5) used time-varying daily observed SSTs. There is no correlation between the choice of SST boundary condition and model hindcast skill (section 3.1).

#### 2.4. Model Output

All centers provided three-hourly output of a standard set of prognostic and diagnostic fields (e.g., temperature, winds, geopotential height, and surface fluxes). Most importantly for this project, all centers also provided three-hourly Eulerian tendencies for temperature, specific humidity, and zonal and meridional wind from the model dynamics and from the convection, longwave and shortwave radiation, boundary layer, and large-scale cloud and precipitation parameterizations, as applicable. A full list of the output fields requested is at [http://climate.ncas.ac.uk/pmwiki/MJO\\_Diabatic\\_Hindcast](http://climate.ncas.ac.uk/pmwiki/MJO_Diabatic_Hindcast). Prior to submission, modeling centers interpolated all data onto a  $2.5^\circ \times 2.5^\circ$  horizontal grid and 24 pressure levels (every 25 hPa from 1000 hPa to 900 hPa then every 50 hPa to 50 hPa), including the subgrid tendencies.

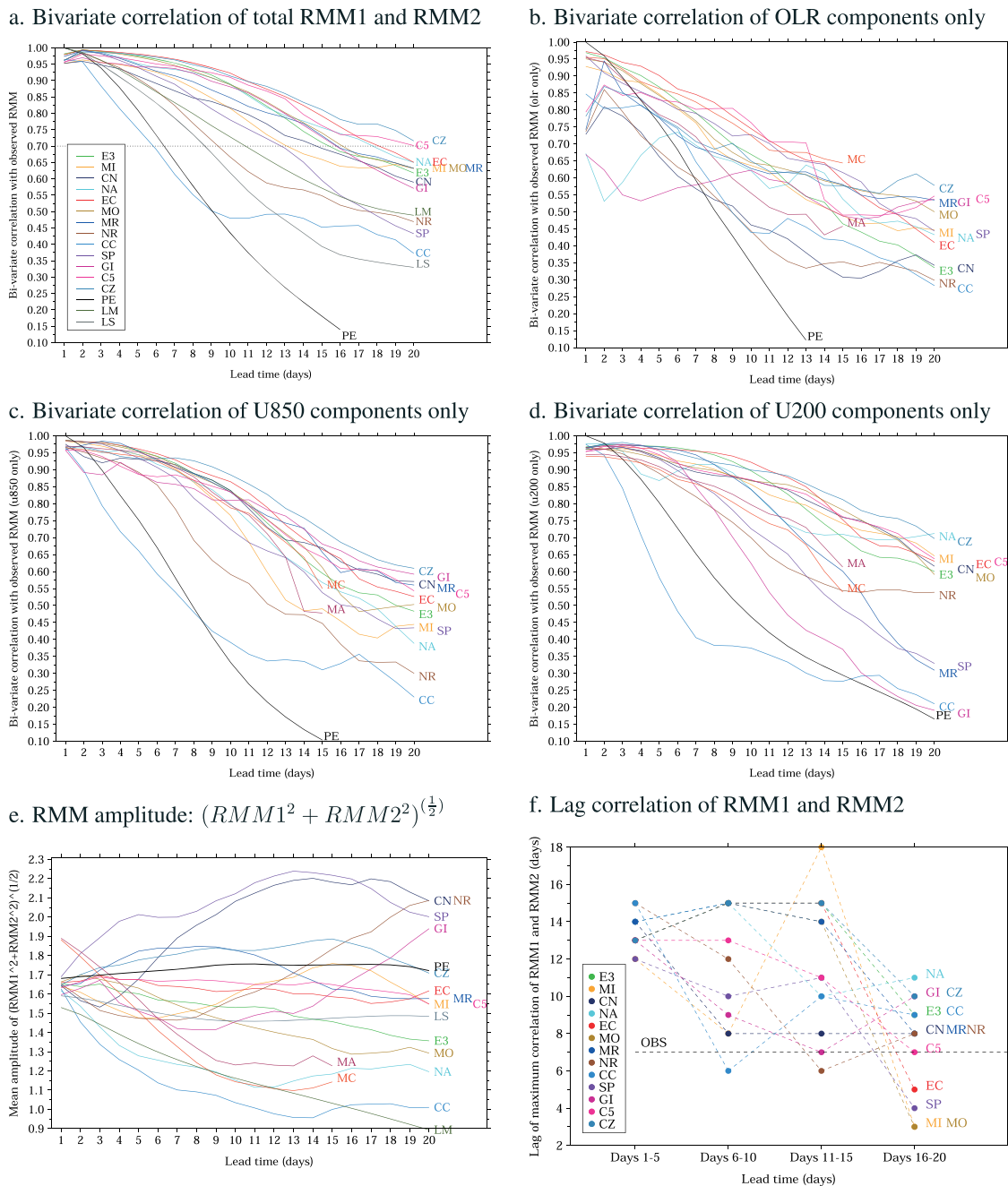
#### 2.5. Computing RMM Indices

Daily RMM indices are computed from GCM data following the procedure in *Gottschalck et al.* [2010]. From daily mean GCM data (computed from three-hourly output), we remove the mean and first three harmonics of the annual cycle of NOAA OLR and ERA-Interim reanalysis zonal wind (1979–2008). Next, we latitude average each field between  $15^\circ\text{N}$  and  $15^\circ\text{S}$ . We perform the same two steps on NOAA OLR and ECMWF-YOTC winds for the hindcast period. For each hindcast lead day  $t$  (where  $t = 1$  is the first day of data), we remove a mean of the previous 120 days of latitude-averaged data from each field. The mean is composed of  $121 - t$  days of “observations” (NOAA OLR and ECMWF-YOTC winds) and  $t$  days of hindcast data. We use NOAA OLR and ECMWF-YOTC winds, instead of the model analysis data suggested in *Gottschalck et al.* [2010], because many of the GCMs that submitted data do not have an analysis system. We then project the resulting model anomalies onto the *Wheeler and Hendon* [2004] RMM EOFs to compute the RMM indices; the *Wheeler and Hendon* [2004] EOFs are computed from 1979 to 2001 NOAA OLR and NCEP-NCAR reanalysis winds.

We also compute observed RMM indices by replacing the model data above with daily mean NOAA OLR and ECMWF-YOTC winds and repeating the procedure. We verify the GCM RMM indices against these observed values. We verify the LIM RMM indices against the LIMs’ own lagged initial (day 0) values, because the LIMs were built using different OLR and zonal wind data sets than the ones employed here. This means that the LIM day 0 error is zero, which also improves the clarity of our skill figures (section 3.1).

#### 2.6. Other Data Sets

Model rainfall is compared against three-hourly TRMM product 3B42 version 7 [Kummerow *et al.*, 1998], which is a combination of many passive microwave and infrared sensors, calibrated against and merged with gauge



**Figure 2.** Bivariate correlations with lead time of hindcast and observed (a) RMM1 and RMM2, only the (b) OLR, (c) U850 and (d) U200 components of RMM1 and RMM2, using all start dates. (e) Mean RMM amplitude with lead time. Black lines (“PE”) in Figures 2a–2e are from persistence forecasts, in which initial observed RMM values are maintained. (f) The lag of maximum correlation of RMM1 and RMM2 for models and observations (“OBS”). For models, lag correlations are computed from lead time ranges shown in the horizontal axis; for OBS, correlations are computed using only the hindcast period. Models are identified by two-letter codes (Table 1). The codes appear next to the final point of each line for a qualitative ranking.

data. Importantly for our analysis, TRMM 3B42 has been shown to underdetect very light rainfall over the tropical oceans, due to biases in one of the microwave sensors, but performs reliably for moderate and heavy rain rates [Huffman et al., 2007]. In section 3.2, we compare GCM precipitation-moisture relationships to an observed product composed of TRMM rainfall and specific humidity from the ECMWF-YOTC 0–23 h forecasts (ECMWF-YOTC/TRMM). We also compare heating and moistening tendencies from the hindcast GCMs to those from the ECMWF-YOTC 0–23 h forecasts. We stress that this biases the GCM intercomparison in favor of the IFS 20 day hindcasts, but also note that we find considerable differences between ECMWF-YOTC and the IFS

hindcasts in many diagnostics. Further, we note that the IFS 20 day hindcasts were performed with a later IFS version (37r3) than ECMWF-YOTC (35r3).

### 3. Results

In this manuscript, it is not often possible to include results for each type of analysis from all 13 GCMs. Instead, we select models from across the continuum of model skill, as discussed at the end of section 3.1. For all analyses presented, results for all GCMs are publicly available on the project website: [http://climate.ncas.ac.uk/pmwiki/MJO\\_Diabatic\\_Hindcast](http://climate.ncas.ac.uk/pmwiki/MJO_Diabatic_Hindcast).

#### 3.1. MJO Predictions for Cases E and F

Bivariate correlations and root-mean-square error (RMSE) [Gottschalck *et al.*, 2010] between hindcast and observed RMM indices were computed with hindcast lead time, using all start dates (Figure 2). RMSE values are not shown; error growth is roughly equal and opposite to the decline in the bivariate correlations. The models display a wide variety of “skill” for these events: CAM5 and CAM5-ZM (C5 and CZ in Figure 2) exhibit more than 20 days skill (using a threshold correlation of 0.7, used for all skill measures); many other GCMs reach 16–18 days skill; several GCMs and the two LIMs (LS and LM) have 8–12 days skill (Figure 2a). One model, CanCM4 (CC), has skill similar to a persistence forecast (PE; 7 days) and lower than the two LIMs. Performance for the two cases individually (not shown) strongly resembles that for the two combined. We use the term skill very loosely here, since we have a set of only 94 hindcasts over two events. This experiment does not aim to provide a thorough assessment of GCM prediction skill, but to link GCM fidelity in predicting these MJO events to representations of diabatic processes (section 2.2). For a robust assessment of the hindcast skill of contemporary GCMs, see the Intraseasonal Variability Hindcast Experiment [e.g., Neena *et al.*, 2014], which obtained 20 years of hindcasts for many models. Neena *et al.* [2014] found that most prediction systems had useful skill for the MJO to 10–20 days for individual ensemble members—in line with our skill estimates for the YoTC cases—and 15–25 days for the ensemble mean.

The set of GCM hindcasts contains two sensitivity tests: air-sea coupling, using the 15 day MetUM atmosphere-only (MA) and coupled (MC) configurations, and convective microphysics, using CAM5 and CAM5-ZM. Coupling and convective microphysics offer slight improvements in the bivariate correlation for these two cases. For air-sea coupling, this confirms the analysis of Shelly *et al.* [2014], who found small benefits to skill during YOTC.

To determine how the degradation in RMM predictions with lead time for a given GCM is associated with the three RMM components, bivariate correlations and RMSEs were computed for the contributions of OLR, U850, and U200 to the indices. For many GCMs (e.g., NavGEM1 (NR) and ECEarth3 (E3)), correlations decline more quickly, and RMSEs grow more quickly (not shown), for OLR (Figure 2b) than for U850 (Figure 2c) or U200 (Figure 2d). We note, however, that OLR contributes less to the RMM indices than U200 and U850 [e.g., Straub, 2013; Ling *et al.*, 2014]. Several GCMs, such as GEOS5 (NA) and GISS-E2 (GI), have low initial OLR correlations due to the spin-up of convection from the ECMWF-YOTC analyses. For GISS-E2, CanCM4 (CC), MRI-AGCM3 (MR), and SPCAM3 (SP), however, the correlations decline most quickly for U200. This suggests that upper level circulation biases are the limiting factors in predictions for these cases in these GCMs, which may indicate issues with either the vertical mixing of momentum by convection or with the depth of convection. SPCAM3 does not include vertical momentum mixing, producing an overly strong vertical shear of zonal wind (not shown). This behavior is consistent with a rapid drift toward the model's climatology, which exhibits excessive superrotation [Pritchard *et al.*, 2014]. The other three GCMs include parameterized convective momentum transport (CMT), as do all GCMs except SPCAM3, but biases in CMT and upper level circulation can reduce simulated MJO activity [e.g., Ling *et al.*, 2009; Klingaman and Woolnough, 2014]. Because U200 represents the largest contribution to the RMM indices of the three components [Wheeler and Hendon, 2004; Straub, 2013], using the RMM indices over an OLR-only index [e.g., Sperber and Kim, 2012] penalizes models with poor upper level wind variability. We compare RMM skill to a diagnostic of precipitation skill at the end of this section.

Correlations for CAM5-ZM slightly exceed those for CAM5 in all components, suggesting modest improvements to MJO predictions from convective microphysics. Air-sea coupling in MetUM improves predictions of the OLR components by 2 days (12 days to 14 days; Figure 2b) while slightly improving U850 and worsening U200. Shelly *et al.* [2014] also found that coupling improved OLR predictions, primarily over the Maritime Continent.



We note, without explanation, that the Miami LIM (LM) produces more skillful predictions than the Scripps LIM (LS). Our hindcast period is too short to rigorously compare these LIMs; we use them only as baseline measures of skill for the GCMs. The LIMs lie at the lower end of the envelope of GCMs, exhibiting similar skill to the least skillful dynamical models (e.g., CanCM4 and NavGEM1; Figure 2a). This represents an improvement in the skill of dynamical models over the past few years, given that in previous assessments dynamical models were shown to have similar or only slightly higher skill compared to statistical models [e.g., Seo *et al.*, 2009; Seo, 2009; Kang and Kim, 2010], although some studies have found that dynamical models could outperform simpler statistical models [e.g., Rashid *et al.*, 2011].

Most models damp RMM amplitude with lead time, either slightly or strongly (Figure 2e). This is true even for GCMs with 16–18 days skill (Figure 2a), such as GEOS5 and ECEarth3. Many models show greater and faster declines in RMM1 than RMM2 magnitude (not shown), suggesting issues predicting the MJO over the Maritime Continent. The four GCMs that maintain the observed RMM amplitude throughout the forecast also show high RMM bivariate correlations: CAM5, CAM5-ZM, IFS, and MRI-AGCM3. SPCAM3 and CNRM-AM (CN) amplify the MJO signal with lead time, but neither has skill beyond 15 days (Figure 2a).

As a measure of MJO propagation, we compute lag correlations of RMM1 and RMM2 and find the lag of the maximum correlation magnitude, similar to Sperber and Kim [2012], which used EOFs of intraseasonal OLR only. The observed RMM indices have a maximum lag correlation at 7 days (Figure 2f), which is close to the 8–9 day lag of the maximum correlation in Wheeler and Hendon [2004]. For each GCM, we calculate lag correlations between RMM1 and RMM2 using all start dates and 5 day ranges of lead times (e.g., days 1–5 and days 6–10) to reduce high-frequency noise in the metric due to the limited sample of hindcasts. All GCMs show longer-than-observed lags between RMM1 and RMM2, indicative of slow MJO propagation (Figure 2f), although in most GCMs the propagation speed increases with lead time as amplitude decays. There is little correspondence between RMM amplitude and this propagation metric: CanCM4 and GEOS5 show weak amplitude, but CanCM4 produces near-observed propagation, while GEOS5 is too slow; CNRM-AM and SPCAM3 produce overly strong amplitudes, but CNRM-AM produces near-observed propagation, while SPCAM3 is slower than observed.

For the remainder of our analysis, we compare processes among these GCMs, focusing on diabatic heating (section 3.4) and moistening (section 3.5). Our objective is to compare GCM representations of these processes to skill in predicting these cases, which requires ranking the models by some skill measure. We rank the GCMs by the lead time at which the RMM bivariate correlation first falls below the 0.7 threshold. Table 2 shows this ranking, along with a tercile-based system for classifying the models into groups of relatively “higher,” “moderate,” and “lower” MJO fidelity. These classifications are used for only two purposes: to select example GCMs from each group to display in subsequent figures and to color the GCM identification codes on several figures to provide a general indication of fidelity. When we correlate MJO fidelity with process-oriented diagnostics, we always use the values of skill shown in Table 2, not the relative classifications.

There are two issues with our approach: (a) the limited hindcast set may not represent overall GCM performance and (b) the conclusions may vary based on the skill measure chosen. For (a), we emphasize that our ranking of model performance is based solely on these two cases; it may not reflect the typical performance of these models for predicting the MJO. We discuss this further in section 4. To address (b), we present an alternate quantification of MJO fidelity, based on the method adopted by Jiang *et al.* [2015] for assessing the 20 year climate simulations from this project. Jiang *et al.* [2015] computed pattern correlations between simulated and observed rainfall Hövmoller diagrams over an equatorial band (averaged 5°S–5°N), using 20–100 day band pass-filtered rainfall and regressing the rainfall against base regions in the Indian Ocean and West Pacific. Here we form Hövmoller diagrams over the same region, but using unfiltered GCM and TRMM 3 h rainfall and without using regressions. For each of the two hindcast cases, we form Hövmollers from each GCM at fixed lead times and compute the pattern correlation with the TRMM Hövmoller diagram from the same period; we then average the pattern correlations from the two cases. Figure 3 demonstrates that models with higher MJO fidelity by the RMM bivariate correlation measure (green model codes) tend to produce higher skill in equatorially averaged rainfall; models with lower bivariate RMM correlations (red codes) tend to produce lower rainfall skill. There are several exceptions: GISS-E2 (CNRM-AM) performs better (worse) by the rainfall measure than by the RMM correlation measure. As above, the RMM indices are dominated by U200 and U850, so it is not surprising that GCMs that exhibit better performance for the OLR component than for the wind components (e.g., GISS-E2) also demonstrate better precipitation fidelity relative to their RMM skill scores. Still, the

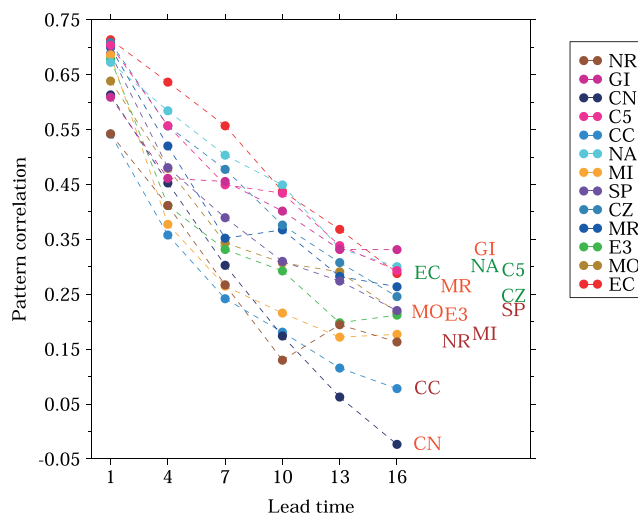
**Table 2.** For Each GCM: the Two-Letter Model Code; MJO Skill, Measured as the First Lead Time at Which the Bivariate Correlation of Simulated and Observed RMM Indices (Figure 2a) Is Less Than 0.7; the Tercile Classification of Relative MJO Fidelity Adopted in This Study; and the Color of the Model's Code in Subsequent Figures, Which Is Based on the Tercile Classification<sup>a</sup>

Model	Code	Skill	Relative Fidelity	Color
CAM5-ZM	CZ	>20 days	Higher	Green
CAM5	C5	>20 days	Higher	Green
IFS	EC	18 days	Higher	Green
GEOS5	NA	18 days	Higher	Green
ECEarth3	E3	16 days	Moderate	Orange
GISS-E2	GI	16 days	Moderate	Orange
MetUM-GA3	MO	16 days	Moderate	Orange
MRI-AGCM3	MR	16 days	Moderate	Orange
CNRM-AM	CN	15 days	Moderate	Orange
MIROC5	MI	14 days	Lower	Red
SPCAM3	SP	12 days	Lower	Red
NavGEM1	NR	10 days	Lower	Red
CanCM4	CC	6 days	Lower	Red

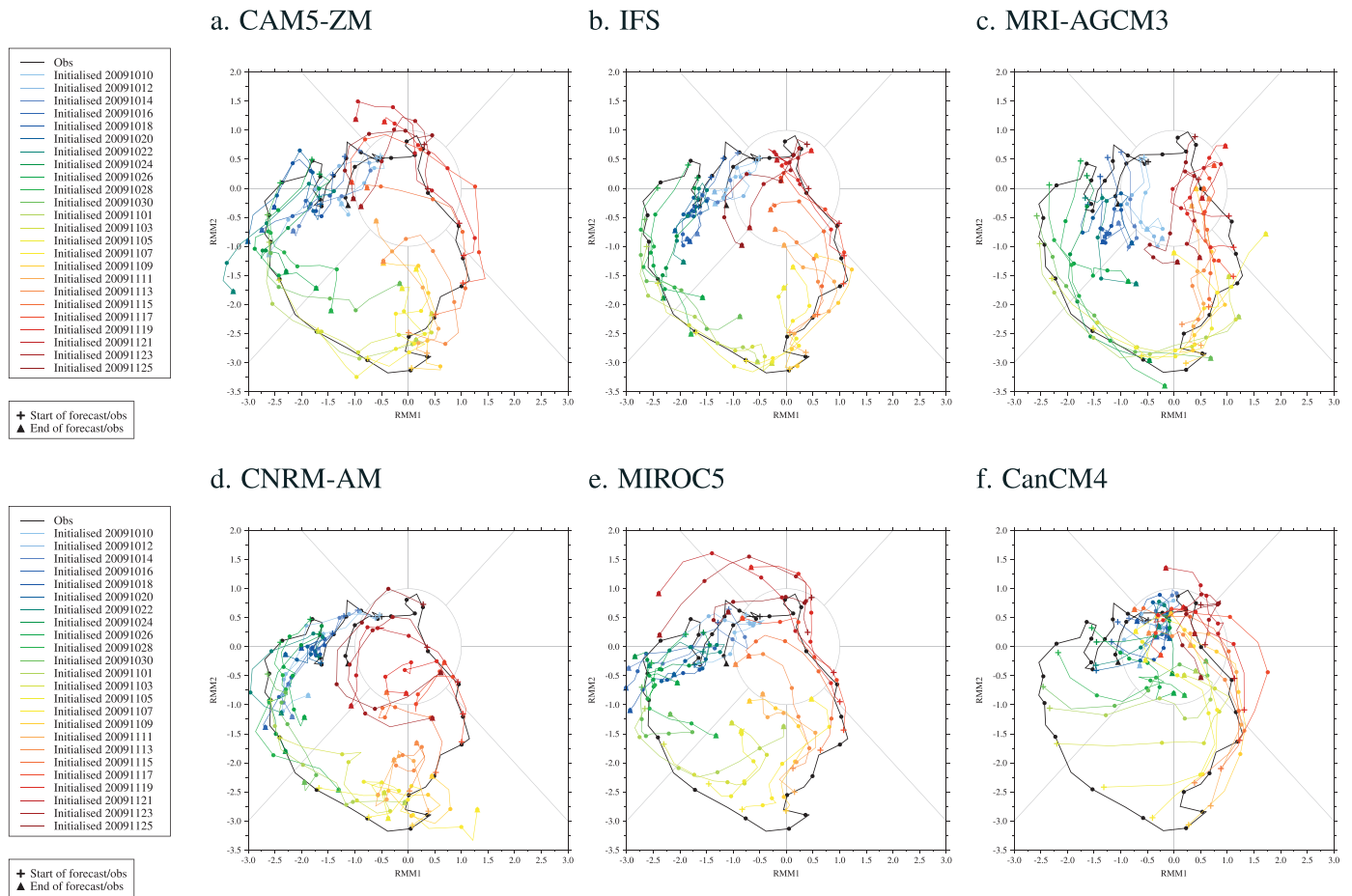
<sup>a</sup>">20 days" denotes that the bivariate correlation did not cross 0.7 within the 20 days lead time in this study.

fact that the two measures of fidelity agree reasonably well increases our confidence that the bivariate RMM correlation reliably represents overall MJO fidelity in these hindcasts.

Figure 3 also shows that all models have considerably lower skill in predicting latitude-averaged precipitation than in predicting the RMM indices, including for the OLR component of the RMM indices (Figure 2b). We note, however, that our precipitation Hövmoller diagrams were not filtered to isolate the MJO signal, unlike those in Jiang et al. [2015]; they contain many smaller-scale features in space and time that may be unrelated to the MJO and that models would be unlikely to capture at lead times beyond a few days, if at all. Again, we emphasize that this experiment is poorly suited to a robust and precise quantification of MJO performance in these models.



**Figure 3.** For each GCM, the pattern correlation of Hövmoller diagrams of latitude-averaged (10°S–10°N), daily mean rainfall, constructed across 60°E–180°, between each GCM and TRMM observations. For each of the two hindcast cases, GCM Hövmoller diagrams are constructed at the fixed lead times shown in the horizontal axis; the value shown is the mean pattern correlation over the two cases, using all start dates. Models are identified by two-letter codes (Table 1), which are colored by relative fidelity (Table 2) and placed next to the final point of each line to give a qualitative ranking.



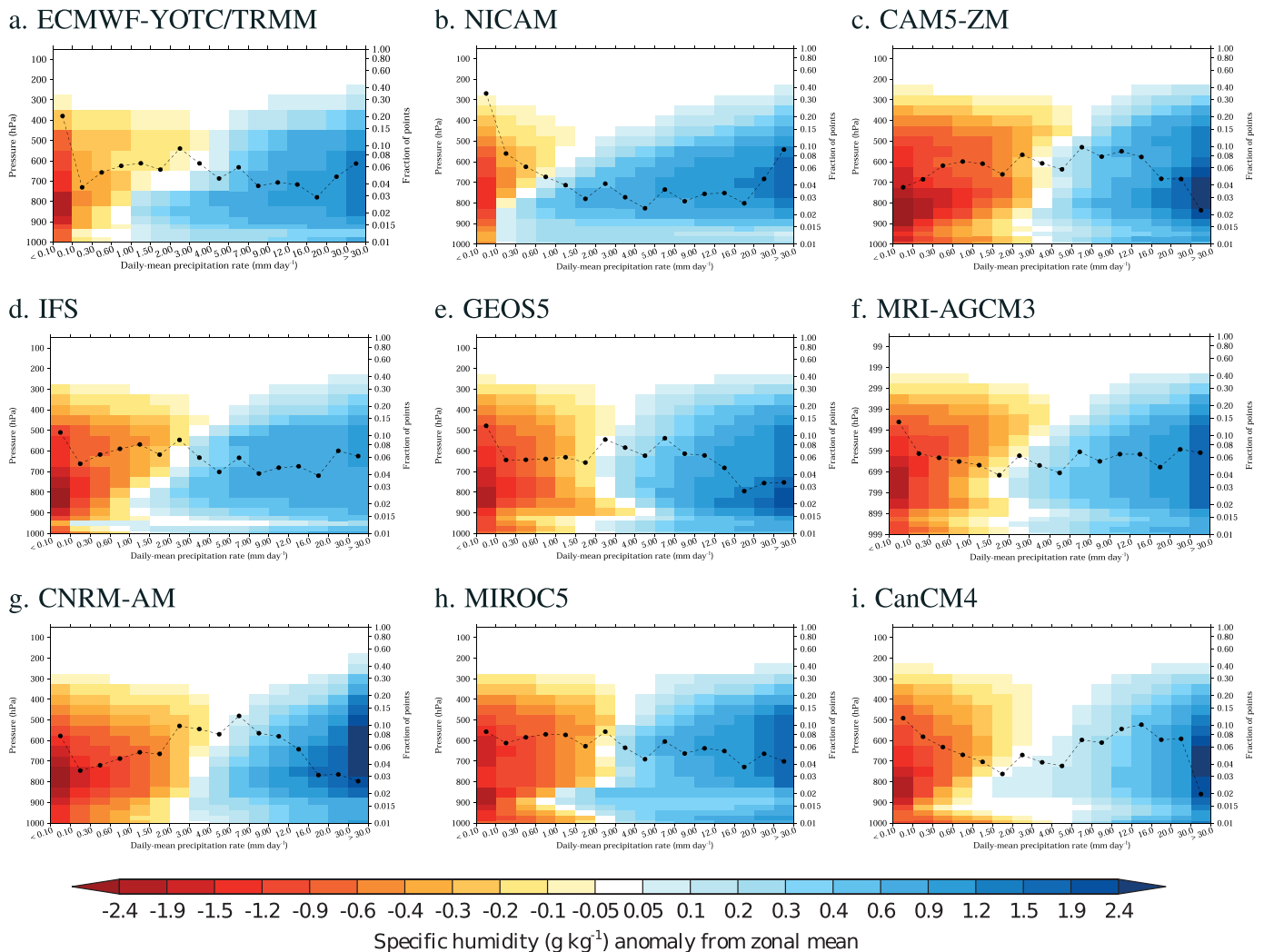
**Figure 4.** For YOTC Case E, hindcasts from (a) CAM5-ZM, (b) IFS, (c) MRI-AGCM3, (d) CNRM-AM, (e) MIROC5, and (f) CanCM4, with each color representing one hindcast start date; the black lines are the observed RMM indices for the period 10 October to 15 December 2009 (i.e., from the first day of the first hindcast to the last day of the last hindcast). For clarity, lines are drawn for only every other start date and show only the first 11 days of each hindcast; the circles are spaced every 2 days.

Figure 4 offers six characteristic examples of GCM behavior; it displays the RMM indices for Case E hindcasts, with each trace representing the first 11 days of each hindcast, with only every second start date shown to improve clarity. CAM5-ZM and IFS are examples of higher-fidelity models, which maintain the RMM amplitude well and produce reasonable propagation, if at a somewhat slower rate than observed (Figures 4a and 4b). The moderate-fidelity models, such as MRI-AGCM3 and CNRM-AM (Figures 4c and 4d), show larger errors in RMM amplitude, particularly as lead time increases. MIROC5 and CanCM4 are examples of lower fidelity models, which show still larger errors in amplitude and phase (Figures 4e and 4f). In particular, CanCM4 is strongly attracted to the unit circle in RMM phase space. We use these as examples of high-, moderate-, and low-fidelity models throughout the manuscript, adding other models that show noteworthy or distinct behavior as appropriate.

### 3.2. Relationships Between Rainfall and Specific Humidity and Vertical Velocity

Recent studies have pointed to the importance of the sensitivity of parameterized convection to environmental moisture, particularly in the midtroposphere, for GCM representations of the MJO [e.g., *Thayer-Calder and Randall, 2009; Kim et al., 2012; Xavier, 2012*]. Low sensitivity causes excessive rainfall, relative to observations, in relatively dry columns, reducing instability and weakening the MJO-suppressed phase; it also causes a peaked rain rate distribution at values close to those required to maintain radiative-convective equilibrium (6–16 mm d<sup>-1</sup>), which suppresses subseasonal variability.

To examine if the rainfall-moisture relationship is related to MJO fidelity, grid point daily mean specific humidity anomalies are composited on grid point daily mean rainfall in the warm pool (10°S–10°N, 60°E–180°).

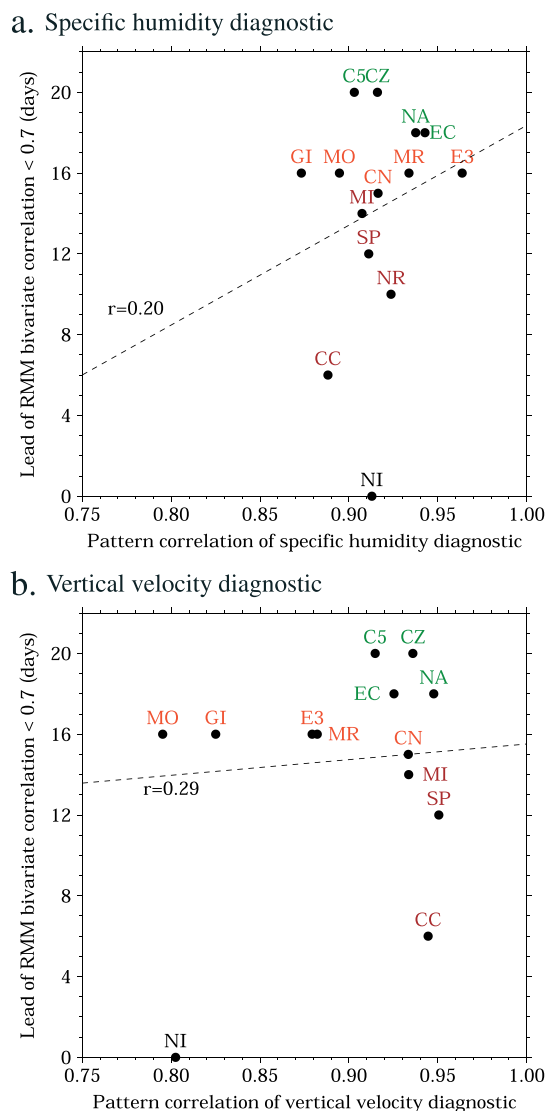


**Figure 5.** For (a) ECMWF-YOTC specific humidity and TRMM precipitation and (b–h) specific humidity and precipitation from each GCM, shaded rectangles show mean anomalies in grid point specific humidity, relative to the zonal mean, across the range of grid point precipitation rates (mm d<sup>-1</sup>) on the horizontal axis. The dashed line shows the probability distribution of precipitation rates, using the right-hand vertical axis. All panels use daily mean data for all hindcast start dates (Figures 5b–5h) and days 3–20; statistics are accumulated over 10°S–10°N and 60°E–180°; zonal means are computed using all longitudes.

Specific humidity anomalies are computed from the zonal mean, calculated daily from all longitudes (0°–360°). When a lead time-dependent climatology is not available, as in this study, using anomalies from a time-varying zonal mean limits the effects of drift in GCM humidity away from the initial analysis [Klingaman and Woolnough, 2014]. Statistics are accumulated over all start dates and lead times. Due to the use of a time-varying zonal mean, these statistics vary little with lead time in any GCM (not shown), indicating that they are intrinsic properties of the GCM and do not depend on the presence of an active MJO (i.e., several GCMs have little MJO activity after 10–15 days). The mean specific humidity anomaly is computed for each rain rate range (Figure 5); the ranges are qualitatively chosen so that TRMM is approximately evenly distributed. We also computed the mean specific humidity anomaly for ranges of rain rate percentiles in the GCMs and TRMM—ensuring a uniform rain rate distribution—and reached similar conclusions.

GCMs are compared to ECMWF-YOTC specific humidity composited on TRMM (ECMWF-YOTC/TRMM; Figure 5a); using ECMWF-YOTC, 0–23 h forecast rainfall produced similar results. ECMWF-YOTC/TRMM shows that rain rates as light as 0.8 mm d<sup>-1</sup> are associated with positive low-level humidity anomalies, with the entire column anomalously moist by 4 mm d<sup>-1</sup>. Between these rates is a gradual transition from low-level to whole-column positive anomalies. NICAM, which produces accurate forecasts, simulates a rainfall-moisture





**Figure 6.** MJO fidelity in 20 day hindcasts plotted against pattern correlations of (a) GCM specific humidity anomalies binned by rain rate (Figure 5) against the same from ECMWF-YOTC specific humidity and TRMM rainfall; (b) as in Figure a but using anomalies in vertical velocity binned by rain rate from the GCMs and ECMWF-YOTC/TRMM. Model codes are colored by relative MJO fidelity (Table 2). Least squares regression lines and correlation coefficients are also shown. We do not assess MJO fidelity in NICAM, so we place its symbol on the horizontal axis and exclude it from the correlation and regression analysis.

relationship is somewhat closer to TRMM. GEOS5 also has a “tongue” of dry anomalies near the top of the boundary layer (900 hPa) that extends out to 3–5 mm d<sup>-1</sup>, whereas ECMWF-YOTC and other models show “tongues” of moist anomalies at this height and these rain rates. Broadly similar rainfall-moisture relationships to the high-fidelity models can be seen in the moderate-fidelity models, including MRI-AGCM3 and CNRM-AM (Figures 5f and 5g), as well as in the lower fidelity models, including MIROC5 and CanCM4 (Figures 5h and 5i). With the exception of NICAM, there appears to be little variation in the rainfall-humidity relationships in this set of GCMs.

To quantify the relationship between this rainfall-humidity diagnostic and MJO hindcast fidelity, we compute the pattern correlation of the rainfall-humidity relationship shown in Figure 5 for each GCM against that in

relationship (Figure 5b) that resembles ECMWF-YOTC/TRMM, but light rain occurs in columns that are too anomalously moist, and the model has a much more gradual transition from midlevel to whole-column positive anomalies than ECMWF-YOTC, in which a sharp transition occurs around 5 mm d<sup>-1</sup>. Nearly all nonzero rain rates are associated with positive low-level humidity anomalies, which is balanced by a higher frequency of <0.1 mm d<sup>-1</sup> rain rates than in TRMM.

IFS also produces a rainfall-moisture relationship similar to ECMWF-YOTC/TRMM, which is perhaps not surprising given that the ECMWF-YOTC analyses were produced with a closely related version of the IFS (Figure 5d and section 2.6). In particular, IFS shows a sharp transition from midlevel to whole-column positive moisture anomalies around 5 mm d<sup>-1</sup>. There are substantial differences between ECMWF-YOTC/TRMM and IFS, however, which maximize 1–2 days into the IFS hindcasts (not shown), including much stronger dry anomalies at lower rain rates—indicating dry columns are associated with more rainfall—and a shift to maximum humidity anomalies above the freezing level (near 550 hPa) for rain rates >1.0 mm d<sup>-1</sup>.

Despite its high fidelity, CAM5-ZM displays a poor rainfall-moisture relationship and rain rate distribution relative to ECMWF-YOTC/TRMM (Figure 5c). Rain rates up to 2.5 mm d<sup>-1</sup> are associated with whole-column dry anomalies; there is a sharp transition from whole-column dry anomalies to whole-column moist anomalies, and the rain rate distribution peaks near 10 mm d<sup>-1</sup>. The lack of near-zero rain rates suggests that CAM5-ZM produces at least some precipitation nearly everywhere in the tropics at all times. CAM5 displays very similar results (not shown). GEOS5, another higher-fidelity model, shows many of the same features (Figure 5e), although the rain rate distribu-

**Table 3.** For ECMWF-YOTC: 0–23 h Forecasts and Each GCM, the Mean Fraction of Nonradiative Diabatic Heating ( $Q_1$ ) Above 550 hPa in Each Rain Rate Quartile, and the Mean Precipitation ( $\text{mm d}^{-1}$ ) in That Quartile<sup>a</sup>

Model	Code	Fourth Quartile	Third Quartile	Second Quartile	First Quartile	<1 mm d <sup>-1</sup>
ECMWF-YOTC		0.65, 17.08	0.59, 7.76	0.51, 4.14	0.34, 1.84	0.17, 0.52
NICAM	NI	0.50, 18.82	0.46, 9.42	0.39, 5.01	0.28, 2.05	0.10, 0.81
CAM5-ZM	CZ	0.48, 10.67	0.34, 5.68	0.26, 3.62	0.18, 1.92	0.09, 0.59
CAM5	C5	0.53, 11.52	0.39, 6.05	0.28, 3.91	0.15, 2.01	0.09, 0.58
IFS	EC	0.41, 14.18	0.32, 7.55	0.22, 4.13	0.09, 1.87	0.02, 0.53
GEOS5	NA	0.50, 12.34	0.41, 5.73	0.29, 3.44	0.17, 1.81	0.05, 0.55
ECEarth3	E3	0.50, 12.89	0.42, 6.94	0.31, 3.87	0.15, 1.80	0.04, 0.54
GISS-E2	GI	0.50, 19.27	0.41, 7.75	0.27, 3.60	0.10, 1.59	0.03, 0.61
MetUM-GA3	MO	0.45, 17.12	0.38, 9.15	0.30, 5.45	0.20, 2.46	0.06, 0.46
MRI-AGCM3	MR	0.53, 20.29	0.48, 9.74	0.39, 5.16	0.24, 2.12	0.06, 0.43
CNRM-AM	CN	0.49, 14.58	0.36, 6.30	0.33, 4.20	0.31, 2.34	0.23, 0.56
MIROC5	MI	0.47, 14.02	0.39, 6.22	0.29, 3.35	0.15, 1.62	0.05, 0.55
SPCAM3	SP	0.48, 16.78	0.39, 7.18	0.28, 3.68	0.18, 1.71	0.11, 0.45
NavGEM1	NR	0.35, 11.20	0.28, 5.71	0.21, 3.36	0.12, 1.70	0.04, 0.64
CanCM4	CC	0.48, 13.75	0.44, 8.02	0.39, 4.86	0.30, 2.21	0.12, 0.45

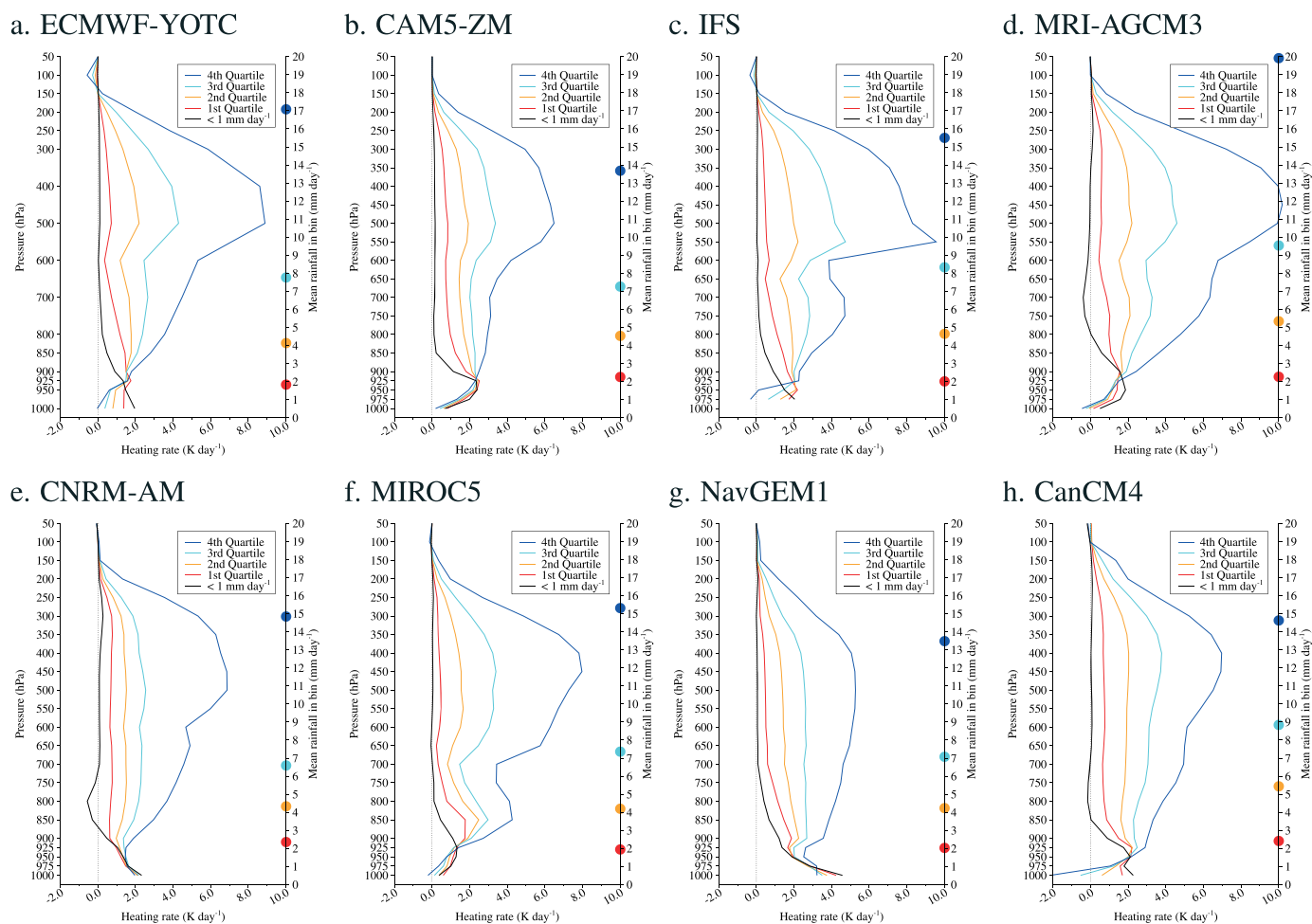
<sup>a</sup>Quartiles are constructed at all grid points in 10°S–10°N and 60°E–180° from precipitation  $\geq 1 \text{ mm d}^{-1}$ ; the final column shows values for precipitation  $< 1 \text{ mm d}^{-1}$ . Quartile values vary with lead time; see text for details. Statistics are accumulated over all hindcast dates and lead times. GCMs are listed in the order of MJO fidelity, except for NICAM, for which skill was not assessed.

ECMWF-YOTC/TRMM. There is no useful correlation ( $r = 0.20$ ,  $p > 0.20$ ) between these pattern correlations and MJO fidelity (Figure 6a). The similarity in the pattern correlation values confirms the small inter-GCM variations in the rainfall-humidity relationship discussed above. We obtain similar results for the relationship between rainfall and vertical velocity, for which we present only the pattern correlation values (Figure 6b). There is more inter-GCM spread in this diagnostic than in the rainfall-humidity diagnostic, but there is still no meaningful correlation with MJO fidelity ( $r = 0.29$ ,  $p > 0.20$ ). We note that *Jiang et al.* [2015] found that the difference in 850–500 hPa integrated relative humidity between the bottom 10% and top 5% of rainfall events [*Kim et al.*, 2014b; *Maloney et al.*, 2014] was able to better discriminate between high- and low-fidelity GCMs in the 20 year climate simulations from this project; we discuss this further in section 4.

### 3.3. Method for Producing Composite Heating and Moistening Profiles

We choose to analyze temperature and specific humidity tendencies as a function of rain rate, by compositing tendencies by quartiles of GCM rain rates over a large horizontal domain: 60°E–180° and 10°S–10°N. This method differs from that used for the companion 2 day hindcasts in *Xavier et al.* [2015], in which the authors chose a smaller ( $5^\circ \times 5^\circ$ ) region and composited tendencies by “suppressed,” “transition,” and “active” conditions according to the observed MJO phase in that region. Our method is more appropriate for the longer hindcasts in this study: it allows us to use all lead times from each GCM, regardless of the presence of an active MJO. Some GCMs damp the amplitude of the MJO sharply (Figure 2e), which would complicate analysis of tendencies by MJO phase. Choosing a small region might bias our analysis toward GCMs that happened to drift to their intrinsic climatologies less rapidly in that region than other GCMs; those same GCMs might drift more rapidly in other regions. Drift was not an issue for the 2 day hindcasts. To ensure that our analysis captured large-scale and not grid point variability, we first interpolated all rainfall and tendency data to a  $10^\circ \times 10^\circ$  horizontal grid; similar results were obtained for  $5^\circ \times 5^\circ$  and  $20^\circ \times 20^\circ$  grids.

Using all start dates and lead times, rain rate quartile thresholds were computed independently for each  $10^\circ \times 10^\circ$  box in the domain and for each 3 h period of the day (e.g., 00Z–03Z and 03Z–06Z), to avoid aliasing the diurnal cycle of convection onto the quartile thresholds. Quartiles were computed only from rain rate values  $> 1 \text{ mm}^{-1}$ . Table 3 lists the domain- and time-averaged rain rates in each quartile. In each  $10^\circ \times 10^\circ$  box, for each 3 h output time in all hindcasts, the linear trend of rain rates was computed with a 27 h sliding centered

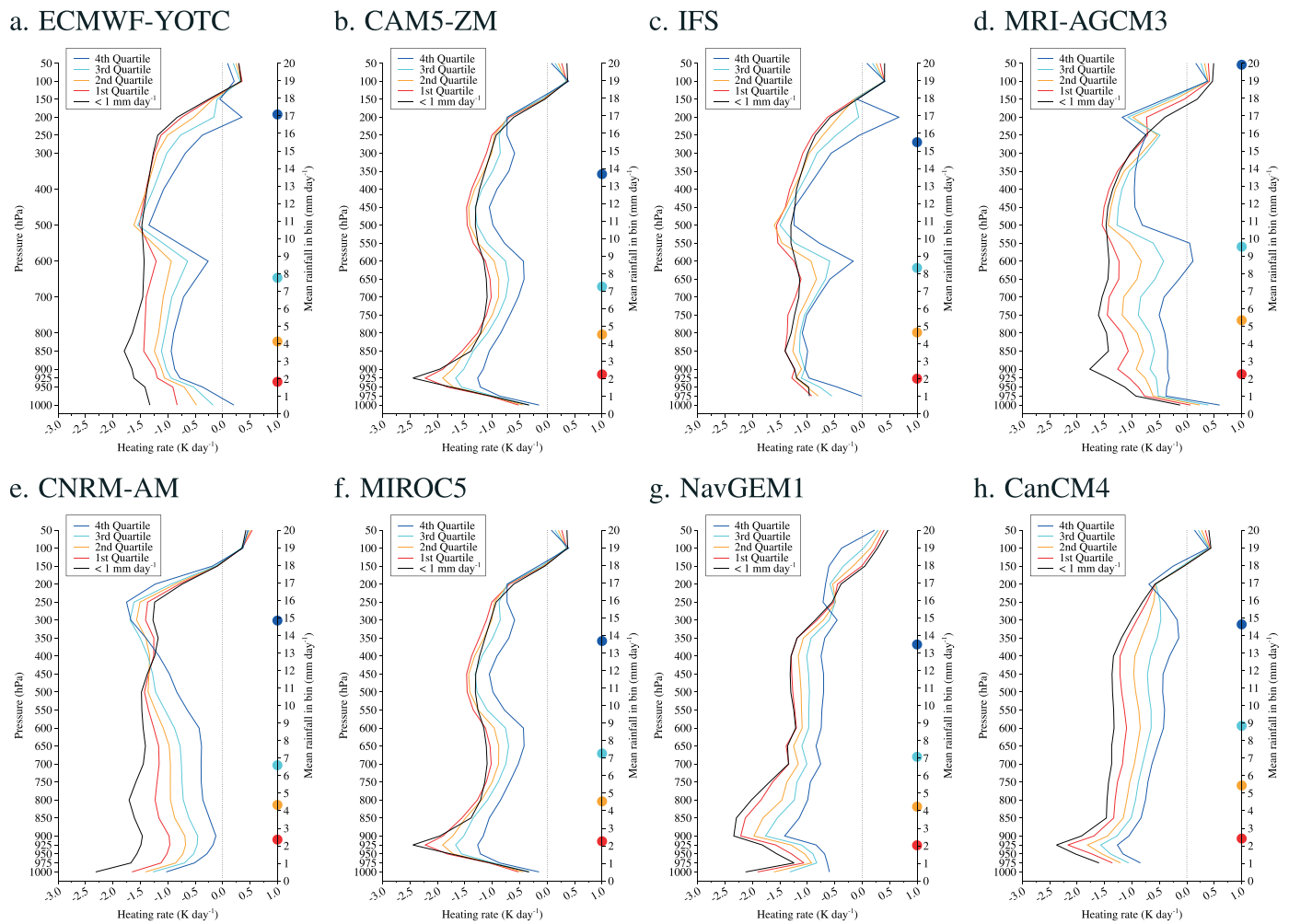


**Figure 7.** The composite total diabatic heating ( $\text{K d}^{-1}$ ) from GCM physics, excluding radiation, binned by (colors) quartiles of rain rates  $>1 \text{ mm d}^{-1}$  and (black) all rain rates  $\leq 1 \text{ mm d}^{-1}$ . Composites are produced from three-hourly precipitation and heating rates on a  $10^\circ \times 10^\circ$  horizontal grid, using data from all hindcast start dates and (b–h) all lead times; statistics are accumulated over  $10^\circ\text{S}$ – $10^\circ\text{N}$  and  $60^\circ\text{E}$ – $180^\circ$ . Rain rate quartiles are computed separately at each grid point and for each 3 h phase of the diurnal cycle. Circles on the right-hand vertical axis show the mean rain rate in each quartile. (a) Computed from ECMWF-YOTC 0–23 h forecasts for the period of the 20 day hindcasts.

window (i.e., using nine values of the three-hourly rain rate, centered on the current time). This trend was then correlated with the values in the window, to determine if the trend was statistically significant at the 5% level. This allowed us to determine, for example, whether a grid box with a rain rate in a transition (second or third) quartile was moving from heavier to lighter rain rates or from lighter to heavier rain rates. We initially created four composite profiles for each rain rate quartile, across (1) all cases in that quartile, (2) linearly increasing rain rate cases, (3) linearly decreasing rain rate cases, and (4) cases where the linear trend was not significant. These four types of composites produced highly similar results for all GCMs, indicating that the composites are insensitive to the 27 h linear trend of rainfall. We show only the composites for all cases (case (1) above); the project website contains versions of Figures 7–10 for all four types of composites. Composites were also computed for all rain rates  $\leq 1 \text{ mm d}^{-1}$ .

As for the precipitation-humidity and precipitation-vertical velocity relationships, we found only very small variations in the heating and moistening tendency composites with hindcast lead time. To maximize the sample size, we compute composites over all lead times and start dates.

We apply the same method to 3 h rainfall and temperature and specific humidity tendencies from the ECMWF-YOTC 0–23 h forecasts over the hindcast period, to compare the GCMs against the closest available approximation to observations of these relationships. We use ECMWF-YOTC rainfall, rather than



**Figure 8.** As in Figure 7 but for the total heating from the GCM longwave and shortwave radiation schemes ( $\text{K d}^{-1}$ ).

TRMM, because of the need for the rainfall to be collocated and contemporaneous with the heating and moistening tendencies.

### 3.4. Diabatic Heating

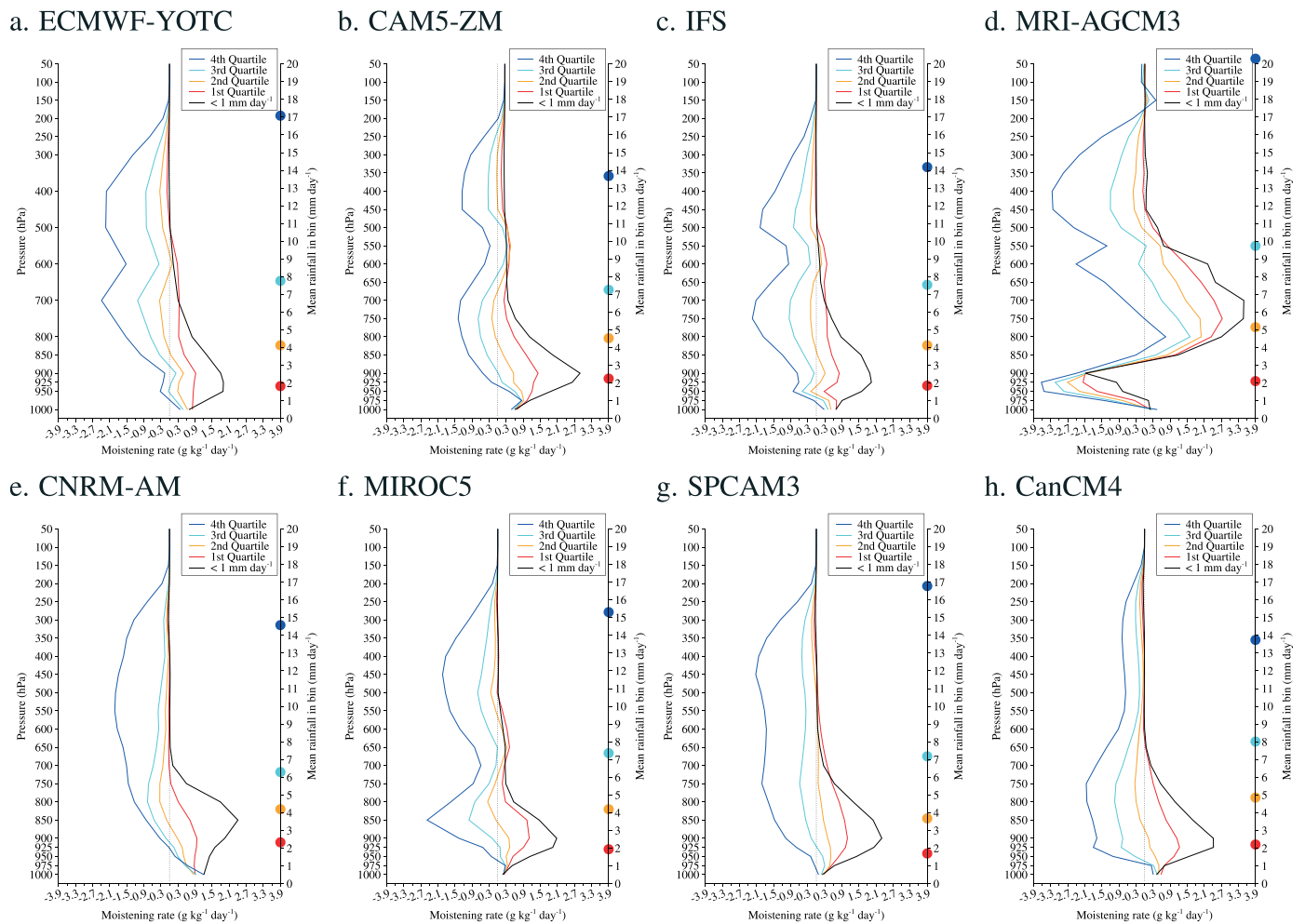
Before analyzing the diabatic heating, we note that the composite total rate of change of temperature (i.e.,  $dT/dt$ ) in all GCMs is very small, typically less than  $0.1 \text{ K d}^{-1}$ , regardless of rain rate (not shown). This is consistent with the weak temperature gradient approximation for the tropical atmosphere [e.g., Sobel *et al.*, 2001; Bretherton and Sobel, 2003]. The composite tendencies in Figures 7 and 8 are balanced by a nearly equal and opposite composite advective tendency (not shown).

#### 3.4.1. Nonradiative Diabatic Heating

We discuss first the composite total diabatic heating from the model subgrid physics, except for radiation (Figure 7); radiative tendencies are presented in section 3.4.2. Heating profiles from individual schemes for all GCMs are available on the project website. Tendencies from individual schemes are often not directly comparable across GCMs, because of intermodel variations in which physical processes are represented in which parameterizations.

ECMWF-YOTC shows mean boundary layer heating at all rain rates, associated with turbulent fluxes (Figure 7a). Above the top of the boundary layer (925 hPa), heating rates tend to decrease with height in the first (driest) quartile and for rain rates less than  $1 \text{ mm d}^{-1}$ . In the second quartile, free-tropospheric heating rates are roughly constant to 400 hPa; the profiles become progressively more top-heavy as rain rates increase through the third and fourth (wettest) quartiles. There is a local minimum in heating around the freezing level



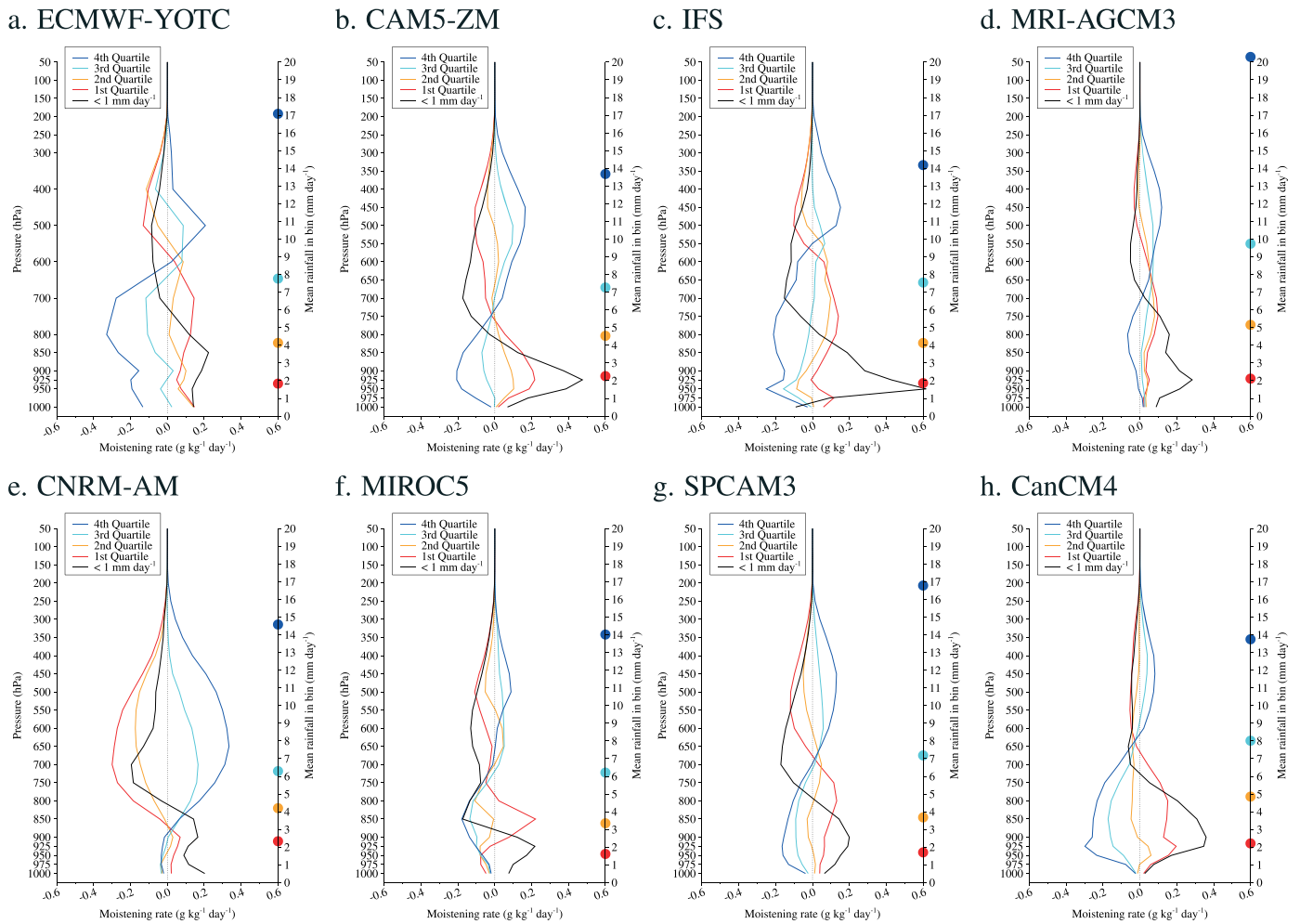


**Figure 9.** As in Figure 7 but for the total diabatic moistening ( $Q_2$ ;  $\text{g kg}^{-1} \text{d}^{-1}$ ) from the GCM subgrid physics schemes.

(500–600 hPa) from the melting of frozen precipitation, which is also present to a greater or lesser extent in several other GCMs, including the IFS (Figure 7c), MRI-AGCM3 (Figure 7d), CNRM-AM (Figure 7e), and MIROC5 (Figure 7f). The lower height of the “kink” in the MIROC5 heating profiles may be due to a warmer threshold temperature at which hydrometeors melt or to differences in the temperature profile. To quantify the “top-heaviness” of the heating profiles, Table 3 gives the fraction of heating above 550 hPa for all models and all quartiles. For all precipitation bands, ECMWF-YOTC produces the most top-heavy heating profiles of any model.

CAM5-ZM, a higher-fidelity model, produces similarly shaped heating profiles to ECMWF-YOTC for the first and second quartiles, which have similar mean rain rates (Figure 7b). In the third and fourth quartiles CAM5-ZM shows more bottom-heavy heating than ECMWF-YOTC (Table 3); in the fourth quartile the CAM5-ZM profile is upright until the freezing level, whereas the ECMWF-YOTC profile shows sharp increases associated with stronger convergence. Heating rates in the third and fourth quartiles are lower overall in CAM5-ZM relative to ECMWF-YOTC, associated with smaller mean rain rates. SPCAM3 also produces similarly shaped profiles to CAM5-ZM (not shown), despite the lower MJO fidelity for these cases in SPCAM3.

It is not surprising that the heating profiles from the IFS, another high-fidelity model, are similar to ECMWF-YOTC, given that an earlier IFS version produced the ECMWF-YOTC forecasts. However, the IFS hindcasts show much less top-heavy heating profiles than ECMWF-YOTC with a stronger minimum at the freezing level in the fourth quartile. As for the rainfall-moisture relationship, this demonstrates that the IFS hindcasts produce different behavior to ECMWF-YOTC, likely due to both the difference in model version and drift away from the ECMWF-YOTC analysis.



**Figure 10.** As in Figure 7 but for the total rate of change of specific humidity ( $dq/dt$ ;  $\text{g kg}^{-1} \text{d}^{-1}$ ). Note that the x axis has one fifth the range of the horizontal axis in Figure 9.

MRI-AGCM3 and CNRM-AM, two moderate-fidelity models, differ considerably in the structure and evolution of their heating profiles with rain rate. MRI-AGCM3 produces the most top-heavy heating profiles of any hind-cast GCM (Table 3), while CNRM-AM shows upright heating profiles until the fourth quartile. There is also substantial inter-GCM variability among the lower fidelity models. MIROC5 produces profiles that are similarly top-heavy to CAM5-ZM and more top-heavy than IFS, suggesting that top-heavy profiles are not a distinguishing feature of higher-skill models. NavGEM1 (Figure 7g) has the most bottom-heavy profiles of any GCMs in this study, with “upright” profiles for even the most intense convection. CanCM4 (Figure 7h) shows profiles similar to MRI-AGCM3 and MetUM-GA3 (not shown) for heavy and moderate precipitation, but without a reduction near the freezing level. For the first quartile, however, the CanCM4 profile is upright, with weak but constant heating between 850 hPa and 300 hPa and 30% of the heating above the freezing level (Table 3), suggesting it produces deep convection even at  $2 \text{ mm d}^{-1}$ .

Despite the variations in the profiles noted above, we find no consistent features among higher-fidelity or lower fidelity GCMs. Most GCMs produce heating profiles that are similarly top-heavy for each rainfall quartile, at least when measured by the fraction of total heating above the freezing level (Table 3). NavGEM1 produces bottom-heavy profiles and poor MJO predictions, indicating that top-heavy heating may be a necessary but not sufficient condition; it is not possible to draw firm conclusions from only one GCM, however. Correlations between model skill and the fractions of heating above 400 hPa, 500 hPa, and 550 hPa, as well as the height of peak heating, for each rainfall quartile produced no statistically significant relationships.

### 3.4.2. Radiative Heating

When composited by rain rate, the shapes, magnitudes, and occasionally even the signs of the GCM radiative heating profiles disagree, both with one another and with ECMWF-YOTC (Figure 8). This is consistent with *Xavier et al.* [2015], who analyze discrepancies in radiative heating for many of the GCMs in this study, using 2 day hindcasts and timestep data. We discuss the radiative heating profiles to confirm that the inter-GCM variations exist on broader spatial and temporal scales ( $10^\circ \times 10^\circ$ , 3 h data) and across a wider range of hindcast start dates.

ECMWF-YOTC, IFS, MRI-AGCM3, and CNRM-AM (Figures 8a and 8c–8e, respectively) generate radiative cooling profiles that, for a given precipitation quartile, are much more constant with height above the boundary layer than the other GCMs. In these GCMs, vertical variations in shortwave heating and longwave cooling roughly balance. The other models, CAM5-ZM, MIROC5, NavGEM1, and CanCM4 (Figures 8b and 8f–8h, respectively), decrease the magnitude of radiative cooling with height in the free troposphere, although there are substantial differences in the shape of the cooling profiles, and hence of the cloud profiles. These models show increasing shortwave heating and reduced longwave cooling with height, particularly at low rain rates, indicating lower cloud tops or optically thinner upper level cloud than ECMWF-YOTC, IFS, MRI-AGCM3, and CNRM-AM. ECMWF-YOTC, IFS, MRI-AGCM3, MIROC5, and, to a lesser extent, CAM5-ZM exhibit “kinks” in the radiative profiles at similar altitudes to corresponding features in the nonradiative profiles (Figure 7), suggesting the production of cloud liquid by detrainment near the freezing level. All GCMs disagree on the shape and magnitude of the upper tropospheric profiles for intense convection, presumably due to differences in cloud top heights and cloud water content.

The profiles composited on rain rates  $\leq 1 \text{ mm d}^{-1}$  (black lines in Figure 8) are effectively clear-sky profiles. ECMWF-YOTC, CAM5-ZM, IFS, and MIROC5 show relatively smaller deviations from these profiles with increasing rain rate; in SPCAM (not shown), the deviations become apparent only for the heaviest (fourth) rain rate quartile. MRI-AGCM3, CNRM-AM, and MetUM (not shown) produce somewhat stronger deviations, even in the first quartile, suggesting that they produce larger amounts of cloud during suppressed conditions. In all models, the profiles for the first, second, and third quartiles have a consistent shape, at least below 300 hPa, and change only in amplitude as rain rate increases. This implies preferred heights for cloud in these models, which do not change between relatively suppressed and active convective conditions.

As for the nonradiative profiles (section 3.4.1), the radiative profiles reveal substantial inter-GCM variations but do not distinguish between higher- and lower fidelity models. We also examined the composite total diabatic heating profiles (i.e., the sum of Figures 7 and 8), which also did not discriminate between higher- and lower fidelity GCMs.

### 3.5. Moistening Tendencies

We composite the vertical profiles of moisture (specific humidity) tendencies by rain rate as for diabatic heating (section 3.3), using all start dates and lead times. While the total rate of change of temperature ( $dT/dt$ ) was uniformly nearly zero in GCMs, due to a balance between dynamics and physics, the total rate of change of moisture ( $dq/dt$ ) is frequently nonnegligible. We analyze the implications of the net moistening and drying that results from the lack of local compensation between the dynamics and physics. We were unable to compute moisture tendencies for CAM5 due to an incomplete data set, but the results of the closely related CAM5-ZM are presented below. As for the temperature tendencies, composite moisture tendencies from all subgrid schemes for all GCMs are available on the project website.

#### 3.5.1. Diabatic Moistening Tendencies

First, we analyze the diabatic moistening from the subgrid physics (“ $Q_2$ ,” which we define to be positive for moistening; Figure 9), which in GCMs is a balance between drying by condensation and moistening from surface fluxes, convective detrainment, and evaporation of cloud water and falling hydrometeors. For rain rates less than  $1 \text{ mm d}^{-1}$  and in the driest quartile, ECMWF-YOTC displays boundary layer and lower and midtropospheric moistening (Figure 9a). In the third and fourth quartiles, the GCM physics dries all but the lowest levels of the column. Moisture tendencies are mostly negligible in the second quartile, but there is some moistening in the boundary layer and near the freezing level and slight drying in the lower and upper troposphere. Higher-fidelity models tend to produce a similar evolution of the  $Q_2$  moistening profiles to ECMWF-YOTC, as shown by CAM5-ZM and IFS (Figures 9b and 9c), however, so do two of the lower fidelity models: MIROC5 and SPCAM3 (Figures 9f and 9g).

Among the moderate-fidelity models, MRI-AGCM3 displays strong boundary layer drying and lower tropospheric moistening at all but the highest rain rates (Figure 9d). This is due to strong vertical transports of moisture by convection, even in suppressed conditions (not shown). Section 3.5.2 demonstrates that these tendencies are mostly compensated by advective drying. CNRM-AM has positive  $Q_2$  above the boundary layer for only the first quartile and very light ( $\leq 1 \text{ mm d}^{-1}$ ) rainfall, transitioning to a negative  $Q_2$  at the second quartile (Figure 9e). CNRM-AM and two of the lower fidelity models, CanCM4 (Figure 9h) and NavGEM1 (not shown), produce free-tropospheric drying from physics above the boundary layer in the second quartile, as do IFS and ECMWF-YOTC, suggesting that this behavior is not a distinguishing feature of either higher- or lower skill models.

There is little consensus among either the higher-skill or the lower skill GCMs on the shape of the  $Q_2$  profile or the rain rate at which free-tropospheric  $Q_2$  shifts from net moistening to net drying. For instance, SPCAM3 and MIROC5 produce profiles for low rain rates that resemble IFS. We note that SPCAM3 produces the deepest moistening in the second quartile of any GCM except MRI-AGCM3, which behaves quite differently to the other models.

### 3.5.2. Total Moistening

Analyzing the total rate of change of moisture ( $dq/dt$ ) by rain rate quartile produces the clearest distinction found in this study between higher-skill and lower skill GCMs. The  $dq/dt$  profiles (Figure 10) are the sum of  $Q_2$  (Figure 9) and the moisture tendencies from the GCM dynamics. All models produce net boundary layer moistening for rain rates  $\leq 1 \text{ mm d}^{-1}$ . In ECMWF-YOTC (Figure 10a) and all hindcast GCMs except MIROC5 (Figure 10f), this moistening extends into the lower troposphere (above 900 hPa). All models also show drying aloft, consistent with large-scale subsidence.

All higher-fidelity models produce net moistening in the lower free troposphere (i.e., above the boundary layer) in the first and second quartiles, as shown for CAM5-ZM and IFS (Figures 10b and 10c). The higher-fidelity models also produce midtropospheric net moistening in the second and third quartiles, similar to ECMWF-YOTC, although CAM5-ZM produces only very slight midlevel moistening in the second quartile. For IFS, a comparison with Figure 9c reveals that much of the 800–600 hPa moistening in the second quartile comes from the IFS dynamics, since the physics tendency is negligible or negative. Between the third and fourth quartiles, the higher-fidelity models transition to lower level net drying and upper level net moistening, also similar to ECMWF-YOTC.

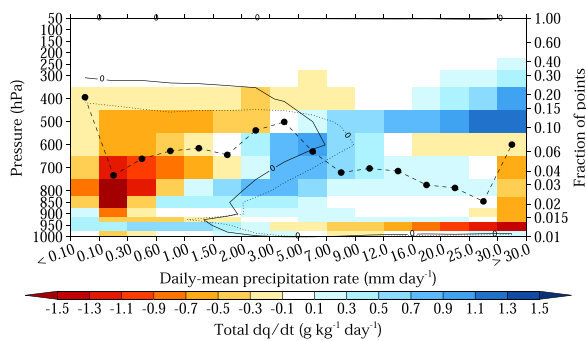
The moderate-fidelity MRI-AGCM3 produces relatively small moistening tendencies (Figure 10d), which confirm that the large  $Q_2$  values (Figure 9d) are typically compensated by opposite and nearly equal tendencies from the dynamics. MRI-AGCM3 produces moistening through much of the column in the first, second, and even third quartiles, suggesting that this behavior is not exclusive to the higher-fidelity models. In CNRM-AM, the composite profiles suggest that the dynamics consistently produces stronger moisture tendencies than the physics, with column drying in the first and second quartiles and column moistening in the third and fourth quartiles (Figure 10e). The  $dq/dt$  profiles have opposite signs to the  $Q_2$  profiles (Figure 9e), indicating the dominance of dynamics over physics. This behavior bears little resemblance to the other GCMs in this project; we discuss these results further later in this section.

Among the lower fidelity models, SPCAM3 produces lower tropospheric moistening in the first quartile, but in the second quartile the tendency becomes negligible, with weak lower tropospheric (upper tropospheric) moistening (drying; Figure 10g). The moistening in the second quartile is weaker than IFS and somewhat weaker than MRI-AGCM. Since SPCAM3 produced positive lower tropospheric  $Q_2$  in the second quartile (Figure 9g), the negligible  $dq/dt$  suggests a stronger compensation between advective drying and physics moistening in SPCAM3 than in IFS. As in IFS and MRI-AGCM, SPCAM3 produces low-level drying (from physics) and upper level moistening (from dynamics) in the third and fourth quartiles. MIROC5 (Figure 10f) and CanCM4 (Figure 10h) show very little free-tropospheric net moistening for rain rates above the first quartile. MIROC5 produces lower tropospheric net drying in the second quartile, while in CanCM4  $dq/dt$  is nearly zero. Therefore, the higher-fidelity GCMs and some moderate-fidelity GCMs produce lower and midtropospheric net moistening for light rain rates, while the lower fidelity GCMs produce net drying or negligible tendencies.

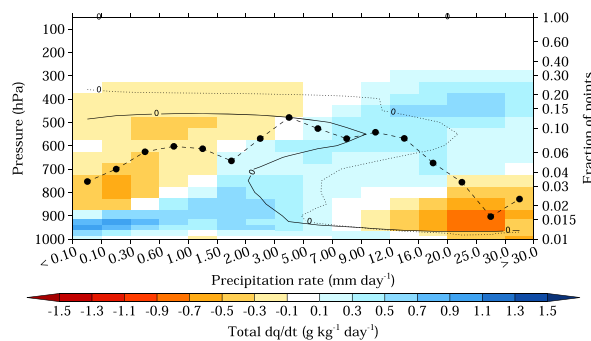
To extend these results, we compute composites of the vertical profiles of  $dq/dt$  using the narrower rain rate bins from Figure 5 to more clearly show the transition from low-level (upper level) moistening (drying) to upper level moistening (drying) with increasing rain rate (Figure 11). We also show the compensation between



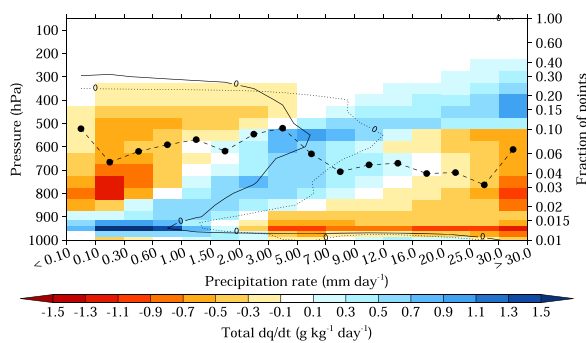
a. ECMWF-YOTC



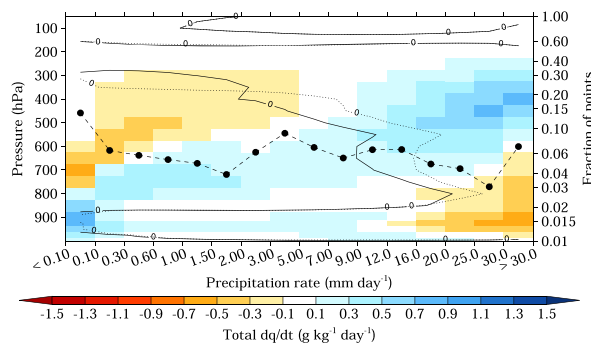
b. CAM5-ZM



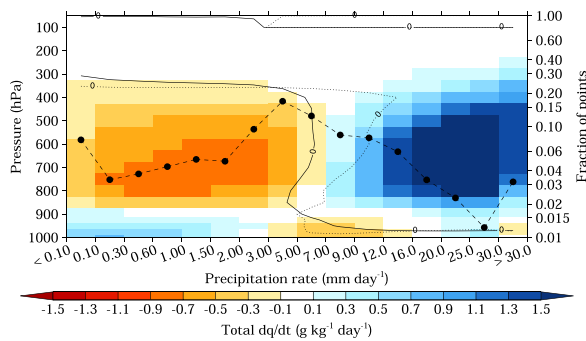
c. IFS



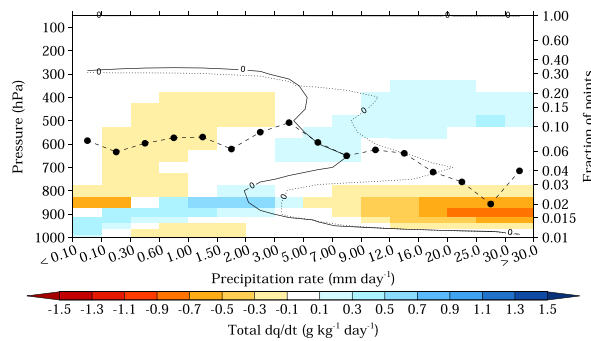
d. MRI-AGCM3



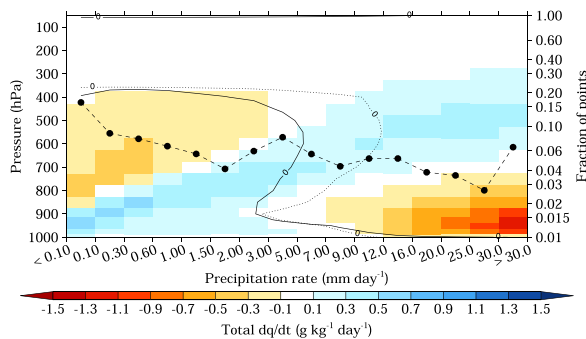
e. CNRM-AM



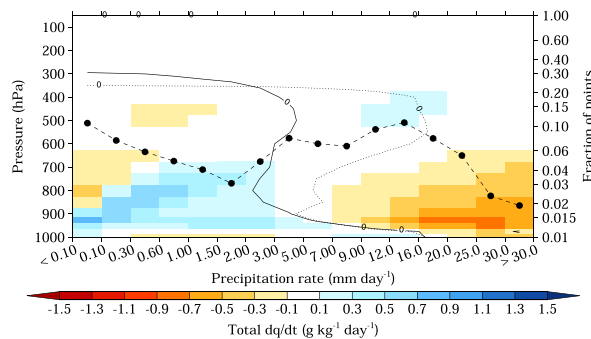
f. MIROC5



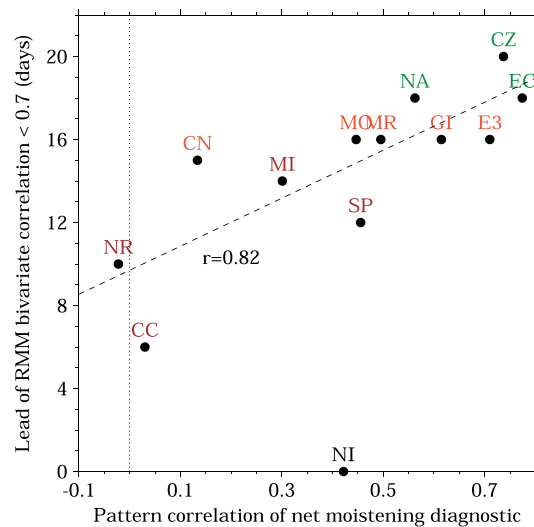
g. SPCAM3



h. CanCM4



**Figure 11.** Shading shows composites of the total rate of change of specific humidity ( $dq/dt$ ;  $\text{g kg}^{-1} \text{d}^{-1}$ ) binned by the rain rates on the horizontal axis. The solid (dotted) lines are the zero contours of the specific humidity tendencies from the GCM dynamics (physics). Dynamic tendencies are positive (moistening) above and to the right of the solid line; physics tendencies are positive below and to the left of the dotted line. The dashed line shows a PDF of rain rates, using the right-hand vertical axis.



**Figure 12.** MJO fidelity in the 20 day hindcasts plotted against the pattern correlation of composite net moistening binned by rain rate (Figure 11) against the same from ECMWF-YOTC net moistening and TRMM precipitation (Figure 11a). The least squares regression line and correlation coefficient are also shown. We do not assess MJO fidelity in NICAM, so we place its symbol on the horizontal axis and exclude it from the correlation and regression analysis.

Several GCMs with moderate or lower fidelity did not display either this progressive rising in the moistening level with rain rate or the enhancement of midlevel moistening by positive tendencies from both dynamics and physics. As discussed above, in CNRM-AM the dynamics moisture tendency is opposite to and frequently greater than the physics tendency; Figure 11e confirms this, with a sharp transition from net drying to net moistening near  $7 \text{ mm d}^{-1}$ , collocated with the zero-moistening contour of the dynamics tendency. The sharp transition may be caused by the lack of vertical tilt with rain rate in the zero-moistening contour from the dynamics; all other models show this tilt to a greater or lesser extent, as does ECMWF-YOTC. The sharp transition suggests an instability in CNRM-AM, where heavy (light) precipitation is associated with column moistening (drying). Curiously, the peak in the rain rate probability density function (PDF) (dashed line) is very close to the critical rain rate between column moistening and drying, implying that the model is finely balanced between these two unstable states. Jiang *et al.* [2015] found that 20 year CNRM simulations in atmosphere-only and coupled configurations showed large negative values of gross moist stability, implying that the large-scale circulation response to convection strongly moistens the column. MIROC5 and CanCM4 show fragmented patterns of moistening with rain rate, without much indication of a transition as rain rate increases (Figures 11f and 11h, respectively). NavGEM1, another lower skill model, displays similar behavior (not shown).

Although the association is not perfect, particularly for SPCAM3, we find that the higher-fidelity GCMs show clear and smooth progressions in the height of net moistening with increasing rain rate, including midlevel moistening at moderate rain rates of  $2\text{--}9 \text{ mm d}^{-1}$ . In most cases, this midlevel moistening comes from a combination of moistening from physics and dynamics. With the exception of SPCAM3, lower fidelity models show a nearly complete compensation between moisture tendencies from dynamics and physics, particularly at moderate rain rates, resulting in weak, if any, midlevel moistening. This results in lower fidelity models producing a fragmented progression from low-level (upper level) moistening (drying) at low rain rates to low-level (upper level) drying (moistening) at heavy rain rates.

To quantify the relationship between this diagnostic and MJO hindcast fidelity, we compute pattern correlations of the composite net moistening profiles shown in Figure 11 with ECMWF-YOTC (Figure 11a) and plot the resulting correlations against RMM bivariate skill. Figure 12 demonstrates that there is a statistically

the moisture tendencies from the GCM dynamics and physics, by including the zero contours from dynamics and physics.

Among the hindcast GCMs, CAM5-ZM, IFS, MRI-AGCM3, and SPCAM3 (Figures 11b–11d and 11g) clearly show a smooth progression from low-level moistening and upper level drying at low rain rates ( $\leq 2 \text{ mm d}^{-1}$ ), through to midlevel moistening at moderate rain rates ( $2\text{--}9 \text{ mm d}^{-1}$ ) and upper level moistening and low-level drying at heavy rain rates ( $\geq 9 \text{ mm d}^{-1}$ ). This behavior agrees strongly with ECMWF-YOTC (Figure 11a). The moisture tendencies are somewhat weaker in MRI-AGCM3, as seen in Figure 10d), and the moistening is displaced vertically relative to GISS-E2, IFS, and SPCAM3, boundary layer moisture into the free troposphere (Figure 9d). At moderate rain rates, all four models and ECMWF-YOTC display moistening from both physics and dynamics between 850 hPa and 400 hPa, which strengthens the moistening at these levels. Several other higher-fidelity and moderate-fidelity GCMs also demonstrated this feature, including GEOS5, GISS-E2, and ECEarth3 (not shown).

significant correlation ( $r = 0.82$ ,  $p = 0.001$ ) between a model's fidelity in the net moistening diagnostic, relative to ECMWF-YOTC, and that model's MJO hindcast fidelity. However, we note that the moderate-fidelity models, which produced either 15 days or 16 days of skill, show pattern correlations ranging from 0.13 (CNRM-AM) to 0.71 (ECEarth3), which confirms that the net moistening diagnostic is not a perfect indicator of MJO fidelity.

#### 4. Discussion

Since the 20 day hindcast component of the model evaluation project seeks to link the climate (20 year simulations) [Jiang *et al.*, 2015] and NWP (2 day hindcasts) [Xavier *et al.*, 2015] components (section 2.1), it is useful to relate our results to those from the other studies. Eleven GCMs from this study—all except IFS and NICAM—also produced 20 year simulations. Among those models, there is little correspondance between the hindcast skill evaluated here and the MJO fidelity assessed in Jiang *et al.* [2015]. Several models that performed well in these initialized hindcast simulations, such as CAM5, CAM5-ZM, and GEOS5, displayed weak or moderate fidelity in the 20 year climate simulations. Conversely, SPCAM3 is a lower fidelity model in the 20 day hindcasts but shows reasonable amplitude and propagation in the 20 year climate simulations.

While we found no significant correlation between MJO hindcast fidelity and the precipitation-moisture and precipitation-vertical velocity relationships (Figure 6), Jiang *et al.* [2015] identified a correlation between MJO amplitude and the difference in mass-weighted 850–500 hPa relative humidity between the bottom 10% and top 5% of daily rainfall events, as suggested by Kim *et al.* [2014b] and Maloney *et al.* [2014]. In a further manuscript that synthesizes and summarizes the conclusions of the entire project, we apply this relative humidity diagnostic to the 20 day hindcasts from the nine GCMs that submitted results to all the three components of the project [Klingaman *et al.*, 2015]. That manuscript also discusses further the relationship between MJO fidelity in the 20 day hindcasts and 20 year climate simulations.

The analysis in Xavier *et al.* [2015] focuses on GCM behavior at very short lead times (12–36 h). Several of the key conclusions of that study, such as the strong intermodel discrepancies in radiative heating profiles, rely upon this behavior remaining consistent with lead time. Figure 8 confirms that even at 20 day leads, the GCMs differ considerably in the shape, magnitude, and even sign of the radiative heating profiles, particularly above the freezing level. Further, we have found that the heating and moistening profiles, as well as the precipitation-humidity relationships (Figure 5), are robust features of the GCM physics schemes that are largely insensitive to lead time or to the presence of a strong MJO. This suggests that the time step tendencies analyzed by Xavier *et al.* [2015] are likely representative of the overall GCM behavior, although we have analyzed only 3 h averages, not time step data.

All three components have emphasized that shallow and midlevel moistening during the MJO-suppressed and transition phases may be critical to GCM representations of the MJO, which agrees with previous studies on the key role of shallow and midlevel convective clouds [e.g., Inness *et al.*, 2001; Benedict and Randall, 2009; Cai *et al.*, 2013]. Jiang *et al.* [2015] found that GCMs with strong, propagating MJOs showed heating profiles that tilted westward with height, with shallow heating and positive low-level moisture anomalies east of the active convection, similar to recent modeling studies with single GCMs [e.g., Lappen and Schumacher, 2014]. Xavier *et al.* [2015] highlighted that most GCMs produced too much rainfall in the MJO-suppressed phase and displayed strong midtropospheric dry biases in the transition phase, even at 12–36 h lead times, suggesting a lack of moistening by shallow or congestus convection. In this study, the highest-skill GCMs showed midlevel moistening at moderate rain rates, indicative of the transition phase, with positive contributions from both the GCM dynamics and subgrid physics. Other GCMs exhibited a much sharper transition from low-level (upper level) moistening (drying) at light-moderate rain rates to upper level (low-level) moistening (drying) for moderate-heavy precipitation. These results with full GCMs support the results of an SCM intercomparison [Woolnough *et al.*, 2010], in which SCMs that produced too much precipitation and too little convective moistening in the suppressed phase had a later and sharper transition to active conditions, which disagreed with the CRM results. GCMs in this study that continued to moisten the lower and midtroposphere with increasing rain rate, by both convective and dynamical processes, produced superior hindcasts to those that showed negligible or negative moisture tendencies at similar rain rates.

This study did not find any significant relationship between the shape of the GCM diabatic heating profiles and their representations of the MJO, based on these hindcasts. Despite several recent modeling

efforts that concluded a top-heavy heating profile for deep convection, produced by stratiform heating, was necessary to simulate the MJO [e.g., *Fu and Wang, 2009; Seo and Wang, 2010*], we found no correlation between the top-heaviness of the profiles and MJO fidelity. All GCMs except NavGEM1 produced similarly top-heavy profiles; the IFS produced one of the least top-heavy profiles but relatively high MJO fidelity (Table 3). The strongest conclusion possible is that a top-heavy profile is a necessary but not a sufficient GCM feature for representing the MJO. The potentially greater importance of shallow convection and low-level moistening ahead of the MJO over a top-heavy heating profile agrees with the results of detailed experiments of the sensitivity to heating structure in CAM [*Lappen and Schumacher, 2012, 2014*].

Even the association between midtropospheric moistening during the suppressed and transition phases and MJO prediction is far from perfect, however. Despite the amount of detailed model output collected, it has not been possible to identify a single, unifying feature of either the highest-skill or lowest-skill GCMs. There are two likely explanations. First, there may be several or even many factors that are individually necessary for capturing the MJO in a GCM, but which are sufficient for producing an MJO only when a GCM includes all of them. This is equivalent to hypothesizing that a GCM can fail to represent the MJO for any one of a wide range of reasons. For example, SPCAM3 produces a top-heavy heating profile (Table 3), low-level moistening from convection in the suppressed phase (Figure 9g), and a gradual increase in the height of net moistening with increasing rain rate (Figure 11g) yet has relatively low MJO fidelity for these cases (Figure 2). The lack of skill is likely due to an overly high RMM amplitude (Figure 2e) from too strong vertical shear caused by a lack of momentum transport by convection. In NavGEM1, the poor MJO may be caused by its bottom-heavy heating profile (Figure 7g); in CanCM4 it may be caused by its lack of midlevel moistening and sharp transition from suppressed to active conditions (Figure 11h). These models may simulate other diabatic heating and moistening processes well, which complicates finding a single process that only the highest-skill models capture. We recommend that these deficiencies should form the basis for further experiments by the participating modeling centers, using the experiment design developed in this project, in which one aspect of the GCM physics (e.g., the triggering of convection) is varied while all other aspects are held constant. These experiments are more likely to produce robust conclusions about the relationship between individual aspects of a GCM physics schemes and MJO fidelity than a large model intercomparison project such as this.

An alternate explanation for the lack of a single, discriminating diagnostic is that this study has not robustly assessed MJO prediction skill in these GCMs. We have examined only two MJO cases using 94 start dates. Certain GCMs that performed well or poorly for these cases may not maintain that performance over a much wider hindcast set. It is curious, for instance, that CAM5 and CAM5-ZM produce such skillful hindcasts when *Jiang et al. [2015]* found that they had weak MJO activity in a 20 year simulation. It may be that these GCMs can produce useful predictions when initialized with a strong MJO [e.g., *Miura et al., 2007; Sperber et al., 2008*], as most start dates have RMM amplitude  $\geq 1$  (Figure 1), but it may also be that Cases E and F are not a representative sample of the models' performance. A larger hindcast set was impractical because of the need to collect detailed, high-frequency output of GCM tendencies. We have tried to link the simulated diabatic processes in these cases to how accurately the GCMs predicted them, but we acknowledge that both the processes and the prediction skill may vary among MJO events.

Our conclusions about CanCM4, which we determined to be a lower skill model, may be influenced by the different initialization procedure used for that model, due to its coupled configuration (section 2.3). However, qualitative analysis of Hövmoller diagrams of CanCM4 OLR and zonal wind showed little evidence of an eastward propagating MJO beyond about 5 days lead time (not shown; available on the project website). We note that *Jiang et al. [2015]* found that CanCM4 had very weak MJO activity in its companion 20 year simulations.

## 5. Summary and Conclusions

This study has analyzed the 20 day hindcast component of the "Vertical structure and diabatic processes of the MJO" global model evaluation project. The aims of this component are (a) to link MJO hindcast fidelity to simulated diabatic heating and moistening processes and (b) to bridge the other two components of the project, which comprise 20 year climate simulations [*Jiang et al., 2015*] and 2 day hindcasts [*Xavier et al., 2015*], by examining the transition in GCMs between the YOTC analyses used as initial conditions and the models' own mean states (section 2.1). Thirteen GCMs provided 94 hindcasts, initialized once per day, of two strong MJO events during YOTC (Cases E and F; Figure 1). Additional GCM contributions were received using the high-resolution, explicit convection NICAM model—eight hindcasts per case—and MetUM—15 day coupled

### Acknowledgments

The complete archive of data produced by this project and analyzed in this manuscript is freely available for download from <https://earthsystemcog.org/projects/gass-yotc-mip/>. The authors are grateful to Chindong Zhang and three anonymous reviewers for their comments and suggestions on this manuscript. Nicholas Klingaman and Steven Woolnough were funded by the National Centre for Atmospheric Science, a collaborative center of the Natural Environment Research Council, under contract R8/H12/83/001. Xianan Jiang acknowledges support by NSF Climate and Large-Scale Dynamics program under award AGS-1228302 and NOAA MAPP program under award NA12OAR4310075. Duane Waliser acknowledges support from the Office of Naval Research under project ONRBAA12-001, and NSF AGS-1221013, and the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The National Center for Atmospheric Research is sponsored by the National Science Foundation. Daehyun Kim appreciates the NASA/GISS modeling group, especially Maxwell Kelley, Mao-Sung Yao, and Anthony Del Genio, for their invaluable and unlimited support. Daehyun Kim was supported by the NASA grant NNX13AM18G and the Korea Meteorological Administration Research and Development Program under grant CATER 2013-3142. The effort of Hsi-Yen Ma was funded by the RGC and ASR programs of the U.S. DOE as part of the CAPT. This work was performed under the auspices of the U.S. DOE by LLNL under contract DE-AC52-07NA27344. William Merryfield acknowledges Jason Cole and Mike Lazare for their roles in producing CanCM4 data and Woo-Sung Lee for performing the CanCM4 simulations. Tomoki Miyakawa acknowledges the NICAM and MIROC teams for developing the models; M. Watanabe, H. Miura, and T. Nasuno for supervising the simulations; N. Hirota and T. Hashino for assistance in data processing; and a grant of supercomputing resources from the Earth Simulator Center. Mike Pritchard was supported by a NOAA CGC postdoctoral fellowship; he thanks Marat Khairoutdinov for developing and making SPCAM3 available through the Center for Multiscale Modeling of Atmospheric Processes, a National Science Foundation (NSF) Science and Technology Center managed by Colorado State University under cooperative agreement ATM-0425247. Computing resources for SPCAM3 simulations were provided courtesy of the Extreme Science and Engineering Discovery Environment, supported by NSF grant OCI-1053575 under

and atmosphere-only hindcasts initialized from MetUM analyses. Two LIM hindcast sets were provided and used as baseline measures of MJO skill, as an improvement over simple persistence.

From the bivariate correlation of simulated and observed *Wheeler and Hendon* [2004] RMM indices (Figures 2), we ranked the GCMs in the order of hindcast skill (Table 2), which is reflected in the arrangement of the figure panels in this manuscript. We stress that this ranking is not an assessment of the overall performance of these models for predicting the MJO but reflects only their ability to hindcast these two cases. Although all but one GCM had skill above persistence, several lower skill models were broadly similar to the two LIMs. We conducted the remainder of our analysis in the context of this ranking, to associate MJO fidelity with GCM representations of particular processes.

We found no relationship between MJO skill and the sensitivity of precipitation to environmental moisture (Figure 5), which other studies had suggested was key to representing the MJO in GCMs [e.g., *Thayer-Calder and Randall*, 2009; *Xavier*, 2012; *Klingaman and Woolnough*, 2014]. We also found no link between MJO skill and the shape of the rain rate PDF or the sensitivity of precipitation to vertical velocity (section 3.2).

Likewise, most of the GCMs, regardless of skill, produced similar diabatic heating profiles from their physics schemes, when analyzed as a function of rain rate (section 3.4.1 and Figure 7). All GCMs except NavGEM1 displayed similarly top-heavy heating profiles for the heaviest quartile of precipitation; NavGEM1 had bottom-heavy profiles and low skill (Table 3). There were no significant correlations between the level of MJO skill and the fraction of heating above the freezing level (550 hPa) or above 400 hPa, despite recent studies that suggested that elevated, stratiform heating was essential to simulate the MJO [e.g., *Fu and Wang*, 2009; *Seo and Wang*, 2010]. It is possible that top-heavy heating is a necessary but not sufficient condition for producing an MJO; with only one GCM with bottom-heavy profiles, this experiment cannot confirm or reject that hypothesis. We note that *Jiang et al.* [2015] found no correlation between the stratiform rainfall fraction in GCMs and MJO fidelity. There were large discrepancies in the radiative heating profiles in these models, particularly above the freezing level, demonstrating that similar conclusions reached by *Xavier et al.* [2015] are not limited to short lead times or the presence of an active MJO in the initial conditions.

The diagnostic most able to discriminate between the highest- and lowest-skill GCMs was the net moistening (i.e.,  $dq/dt$ ) as a function of rain rate (Figures 11 and 12). While all models produced lower tropospheric moistening for low rain rates (generally  $\leq 2 \text{ mm d}^{-1}$ ) from their boundary layer and convection schemes, the highest-skill GCMs continued to moisten the lower and middle tropospheres as rain rates increased into the second and third quartiles ( $2\text{--}8 \text{ mm d}^{-1}$ , Figure 9). In the midtroposphere, the highest-skill GCMs have moistening from both their dynamics and subgrid physics, which enhances the moistening at these levels relative to the lowest-skill models, which typically moisten only from their dynamics at these rain rates (Figure 11) and often show negative or near-zero net  $dq/dt$  due to strong drying from condensation (Figure 10). This finding supports the results of a previous SCM intercomparison of the MJO-suppressed phase during TOGA COARE [*Woolnough et al.*, 2010]. It suggests that moistening from shallow or congestus convection during the MJO-suppressed and transition phases is critical to the maintenance and propagation of the active phase. This moistening is often missing in GCMs with poor MJOs, leading to midtropospheric dry biases [e.g., *Xavier*, 2012], which *Xavier et al.* [2015] found developed within the first 12–36 h in many of the GCMs used here. Our results also agree with the conclusion of *Jiang et al.* [2015] that GCMs with strong MJOs in 20 year climate simulations have heating profiles that tilt west with height, with shallow heating leading (east of) the active MJO.

Even the association between prediction skill and moistening in the MJO-suppressed and transition phases was far from perfect, however, as some models (e.g., SPCAM3) performed reasonably well by this measure but produced relatively lower MJO fidelity. From this set of hindcasts, which were limited in scope by the need to collect high-frequency output of GCM tendencies, we found no process or diagnostic which all of the higher-fidelity models captured well and all of the lower fidelity models captured poorly. While this may be due to the limited sample of hindcasts, it likely emphasizes that GCMs can fail to predict to the MJO for any one of the many reasons, which no single diagnostic can capture.

### References

- Annamalai, H., and K. R. Sperber (2005), Regional heat sources and the active and break phases of Boreal Summer Intraseasonal Variability, *J. Atmos. Sci.*, *62*, 2726–2748.



allocation TG-ATM120034. James Ridout gratefully acknowledges support from the Office of Naval Research program element 0601153N, a grant of computing time from the United States Department of Defense High Performance Computing Modernization Program, and assistance from Maria Flatau in preparing the NavGEM1 hindcasts. Frederic Vitart acknowledges Peter Bechtold for assisting in producing tendency output from the ECMWF IFS hindcasts. Guang Zhang was supported by NSF AGS-1015964, NOAA NA11OAR4310098, and DOE DE-SC0008880.

- Bechtold, P., M. Köhler, T. Jung, F. Doblas-Reyes, M. Leutbecher, M. J. Rodwell, F. Vitart, and G. Balsamo (2008), Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal timescales, *Q. J. R. Meteorol. Soc.*, *137*, 553–597.
- Benedict, J. J., and D. A. Randall (2009), Structure of the Madden-Julian oscillation in the superparameterized CAM, *J. Clim.*, *66*, 3277–3296.
- Bladé, I., and D. L. Hartmann (1993), Tropical intraseasonal oscillations in a simple nonlinear model, *J. Atmos. Sci.*, *50*, 2922–2939.
- Boyle, J. S., S. Klein, G. Zhang, S. Xie, and X. Wei (2008), Climate model forecast experiments for TOGA COARE, *Mon. Weather Rev.*, *136*, 808–832, doi:10.1175/2007MWR2145.1.
- Bretherton, C. S., and A. H. Sobel (2003), The Gill model and the weak temperature gradient approximation, *J. Atmos. Sci.*, *60*, 451–460.
- Cai, Q., G. J. Zhang, and T. Zhou (2013), Impacts of shallow convection on MJO simulation: A moist static energy and moisture budget analysis, *J. Clim.*, *26*, 2417–2431.
- Cavanaugh, N. R., T. Allen, A. Subramanian, B. Mapes, and A. J. Miller (2014), The skill of tropical linear inverse models in hindcasting the Madden-Julian oscillation, *Clim. Dyn.*, doi:10.1007/s00382-014-2181-x.
- Chikira, M., and M. Sugiyama (2010), A cumulus parameterization with state-dependent entrainment rate. Part I: Description and sensitivity to temperature and humidity profiles, *J. Atmos. Sci.*, *67*, 2171–2193.
- Del Genio, A. D., Y. Chen, Y. Kim, and Y. Mao-Sung (2012), The MJO transition from shallow to deep convection in CloudSat/CALIPSO data and GISS GCM simulations, *J. Clim.*, *25*, 3755–3770.
- Flatau, M., and Y.-J. Kim (2013), Interaction between the MJO and polar circulations, *J. Clim.*, *26*, 3562–3574.
- Fu, X., and B. Wang (2009), Critical roles of the stratiform rainfall in sustaining the Madden-Julian oscillation: GCM experiments, *J. Clim.*, *22*, 3939–3959.
- Gottschalk, J., M. Wheeler, K. Weickmann, F. Vitart, N. Savage, H. Lin, H. Hendon, D. Waliser, K. Sperber, M. Nakagawa, C. Prestrelo, M. Flatau, and W. Higgins (2010), A framework for assessing operational Madden-Julian oscillation forecasts: A CLIVAR MJO Working Group project, *Bull. Am. Meteorol. Soc.*, *91*, 1247–1258.
- Harshvardhan, R. D., D. A. Randall, and T. G. Corsetti (1987), A fast radiation parameterization for general circulation models, *J. Geophys. Res.*, *92*, 1009–1016.
- Hazeleger, W., et al. (2012), EC-Earth V2.2: Description and validation of a new seamless Earth system prediction model, *Clim. Dyn.*, *39*, 2611–2629.
- Henderson, G. R., B. S. Barrett, and D. M. Laflour (2014), Arctic sea ice and the Madden-Julian oscillation, *Clim. Dyn.*, *43*, 2185–2196.
- Hendon, H. H., and B. Liebmann (1990), A composite study of the onset of the Australian summer monsoon, *J. Atmos. Sci.*, *47*, 2227–2240.
- Hogan, T. F., et al. (2014), The Navy Global Environmental Model, *Oceanography*, *27*, 64–74.
- Hu, Q., and D. A. Randall (1994), Low-frequency oscillations in radiative–convective systems, *J. Atmos. Sci.*, *51*, 1089–1099.
- Huffman, G. J., R. F. Adler, D. T. Bolvin, G. Gu, E. J. Nelkin, K. P. Bowman, Y. Hong, E. F. Stocker, and D. B. Wolff (2007), The TRMM multi-satellite precipitation analysis: Quasi-global, multi-year, combined-sensor precipitation estimates at fine scale, *J. Hydrometeorol.*, *8*, 38–55.
- Hung, M.-P., J.-L. Lin, W. Wang, D. Kim, T. Shinoda, and S. J. Weaver (2013), MJO and convectively coupled equatorial waves simulated by CMIP5 climate models, *J. Clim.*, *26*, 6185–6214.
- Inness, P. M., J. M. Slingo, E. Guilyardi, and J. Cole (2001), Organization of tropical convection in a GCM with varying vertical resolution: Implications for the simulation of the Madden-Julian Oscillation, *Clim. Dyn.*, *17*, 777–793.
- Jiang, X., D. E. Waliser, W. S. Olson, W.-K. Tao, T. S. L'Ecuyer, K.-F. Li, Y. L. Yung, S. Shige, S. Lang, and Y. N. Takayabu (2011), Vertical diabatic heating structure of the MJO: Intercomparison between recent reanalyses and TRMM estimates, *Mon. Weather Rev.*, *139*, 3208–3223, doi:10.1175/2011MWR3636.1.
- Jiang, X., et al. (2015), Vertical structure and physical processes of the Madden-Julian oscillation: Exploring key model physics in climate simulations, *J. Geophys. Res. Atmos.*, *120*, doi:10.1002/2014JD022375.
- Joseph, S., A. K. Sahai, and B. N. Goswami (2009), Eastward propagating MJO during boreal summer and Indian monsoon droughts, *Clim. Dyn.*, *32*, 1139–1153.
- Kang, I.-S., and H.-M. Kim (2010), Assessment of MJO predictability for boreal winter with various statistical and dynamical models, *J. Clim.*, *23*, 2368–2378.
- Khairoutdinov, M., C. DeMott, and D. Randall (2008), Evaluation of the simulated interannual and subseasonal variability in an AMIP-style simulation using the CSU multiscale modeling framework, *J. Clim.*, *23*, 413–431.
- Kim, D., et al. (2009), Application of MJO simulation diagnostics to climate models, *J. Clim.*, *22*, 6413–6436.
- Kim, D., A. H. Sobel, A. D. Del Genio, Y. Chen, S. J. Camargo, M.-S. Yao, M. Kelley, and L. Nazarenko (2012), The tropical subseasonal variability simulated in the NASA GISS general circulation model, *J. Clim.*, *25*, 4641–4659.
- Kim, D., J.-S. Kug, and A. H. Sobel (2014a), Propagating vs. non-propagating Madden-Julian oscillation events, *J. Clim.*, *27*, 111–125.
- Kim, D., et al. (2014b), Process-oriented MJO simulation diagnostic: Moisture sensitivity of simulated convection, *J. Clim.*, *27*, 5379–5395.
- Klingaman, N. P., and S. J. Woolnough (2014), Using a case-study approach to improve the Madden-Julian oscillation in the Hadley Centre model, *Q. J. R. Meteorol. Soc.*, *140*, 2491–2505.
- Klingaman, N. P., X. Jiang, P. K. Xavier, J. Petch, D. Waliser, and S. J. Woolnough (2015), Vertical structure and physical processes of the Madden-Julian oscillation: Synthesis and summary, *J. Geophys. Res. Atmos.*, *120*, doi:10.1002/2015JD023196.
- Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson (1998), The Tropical Rainfall Measuring Mission (TRMM) sensor package, *J. Atmos. Oceanic Technol.*, *15*, 809–817.
- Lappen, C.-L., and C. Schumacher (2012), Heating in the tropical atmosphere: What level of detail is critical for accurate MJO simulations in GCMs?, *Clim. Dynam.*, *39*, 2547–2568.
- Lappen, C.-L., and C. Schumacher (2014), The role of tilted heating in the evolution of the MJO, *J. Geophys. Res. Atmos.*, *119*, 2966–2989, doi:10.1002/2013JD020638.
- Lau, K.-M., and L. Peng (1987), Origin of low-frequency (intraseasonal) oscillations in the tropical atmosphere. Part I: Basic theory, *J. Atmos. Sci.*, *44*, 950–972.
- Lin, H., and G. Brunet (2011), Impact of the North Atlantic Oscillation on the forecast skill of the Madden-Julian Oscillation, *Geophys. Res. Lett.*, *38*, L02802, doi:10.1029/2010GL046131.
- Lin, J. L., et al. (2006), Tropical intraseasonal variability in 14 IPCC AR4 climate models. Part I: Convective signals, *J. Clim.*, *19*, 2665–2690.
- Ling, J., and C. Zhang (2011), Structural evolution in heating profiles of the MJO in global reanalyses and TRMM retrievals, *J. Clim.*, *24*, 825–842.
- Ling, J., C. Li, and X. Jia (2009), Impacts of cumulus momentum transport on MJO simulation, *Adv. Atmos. Sci.*, *26*, 864–876.
- Ling, J., P. Bauer, P. Bechtold, A. Beljaars, R. Forbes, F. Vitart, M. Ullate, and C. Zhang (2014), Global versus local MJO forecast skill of the ECMWF model during DYNAMO, *Mon. Weather Rev.*, *142*, 2228–2247, doi:10.1175/MWR-D-13-00292.1.
- Ma, H.-Y., S. Xie, J. S. Boyle, S. A. Klein, and Y. Zhang (2013), Metrics and diagnostics for precipitation-related processes in climate model short-range hindcasts, *J. Clim.*, *26*, 1516–1534.
- Ma, H.-Y., et al. (2014), On the correspondence between mean forecast errors and climate errors in CMIP5 models, *J. Clim.*, *27*, 1781–1798.

- Madden, R. A., and P. R. Julian (1972), Description of global-scale circulation cells in the tropics with a 40–50 day period, *J. Atmos. Sci.*, *29*, 1109–1123.
- Madden, R. A., and P. R. Julian (1994), Observations of the 40–50 day tropical oscillation—A review, *Mon. Weather Rev.*, *122*, 814–837.
- Maloney, E. D., X. Jiang, S.-P. Xie, and J. Benedict (2014), Process-oriented diagnosis of East Pacific warm pool intraseasonal variability, *J. Clim.*, *27*, 6305–6324.
- Mapes, B. E. (2000), Convective inhibition, subgrid-scale triggering energy, and stratiform instability in a toy tropical wave model, *J. Atmos. Sci.*, *57*, 1515–1535.
- Merryfield, W. J., W.-S. Lee, G. J. Boer, V. V. Kharin, J. F. Scinocca, G. M. Flato, R. S. Ajayamohan, J. C. Fyfe, Y. Tang, and S. Polavarapu (2013), The Canadian seasonal to interannual prediction system. Part I: Models and initialization, *Mon. Weather Rev.*, *141*, 2910–2945, doi:10.1175/MWR-D-12-00216.1.
- Miura, H., M. Satoh, T. Nasuno, A. T. Noda, and K. Oouchi (2007), A Madden-Julian oscillation event realistically simulated by a global cloud-resolving model, *Science*, *318*, 1763–1765.
- Miyakawa, T., M. Satoh, H. Miura, H. Tomita, H. Yashiro, A. T. Noda, Y. Yamada, C. Kodama, M. Kimoto, and K. Koneyama (2014), Madden-Julian oscillation prediction skill of a new-generation global model demonstrated using a supercomputer, *Nat. Commun.*, *5*, 3769.
- Moncrieff, M. W., D. E. Waliser, M. J. Miller, M. A. Shapiro, G. R. Asrar, and J. Caughey (2012), Multiscale convective organization and the YOTC virtual global field campaign, *Bull. Am. Meteorol. Soc.*, *93*, 1171–1187.
- Morrison, H., and A. Gettelman (2008), A new two-moment bulk stratiform cloud microphysics scheme in the Community Atmosphere Model, version 3 (CAM3), *J. Clim.*, *21*, 3642–3659.
- Neale, R. B., et al. (2012), Description of the NCAR atmospheric model: CAM5.0, *Tech. Rep. NCAR/TN-486+STR*, National Center for Atmospheric Research, Boulder, Colo.
- Neena, J. M., J.-Y. Lee, D. Waliser, B. Wang, and X. Jiang (2014), Predictability of the Madden-Julian oscillation in the intraseasonal variability hindcast experiment, *J. Clim.*, *27*, 4531–4543.
- Pai, D. S., J. Bhate, O. P. Sreejith, and H. R. Hatwar (2011), Impact of MJO on the intraseasonal variation of summer monsoon rainfall over India, *Clim. Dyn.*, *36*, 41–55.
- Petch, J., D. Waliser, X. Jiang, P. K. Xavier, and S. Woolnough (2011), A global model intercomparison of the physical processes associated with the Madden-Julian oscillation, *GEWEX News*, *21*(3), 3–5.
- Pritchard, M. S., C. S. Bretherton, and C. A. DeMott (2014), Restricting 32–128 km horizontal scales hardly affects the MJO in the Superparameterized Community Atmosphere Model v3.0 but the number of cloud-resolving grid columns constrains vertical mixing, *J. Adv. Model. Earth Syst.*, *6*, 723–739.
- Rashid, H. A., H. H. Hendon, M. C. Wheeler, and O. Alves (2011), Prediction of the Madden-Julian oscillation with the POAMA dynamical prediction system, *Clim. Dyn.*, *36*, 649–661.
- Raymond, D. J. (2001), A new model of the Madden-Julian oscillation, *J. Atmos. Sci.*, *58*, 2807–2819.
- Rienecker, M. M., et al. (2008), The GEOS-5 data assimilation system: Documentation of version 5.0.1, 5.1.0, and 5.2.0, *Tech. Rep. Series on Global Modeling and Data Assimilation*, National Aeronautics and Space Administration, Greenbelt, Md.
- Robertson, A. W., S. Kirshner, P. Smyth, S. P. Charles, and B. C. Bates (2006), Subseasonal-to-interdecadal variability of the Australian monsoon over North Queensland, *Q. J. R. Meteorol. Soc.*, *132*, 519–542.
- Satoh, M., T. Matsuno, H. Tomita, H. Miura, T. Nasuno, and S. Iga (2008), Nonhydrostatic icosahedral atmospheric model (NICAM) for global cloud resolving simulations, *J. Comput. Phys.*, *227*, 3486–3514.
- Schmidt, G., et al. (2014), Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive, *J. Adv. Model. Earth Syst.*, *6*, 141–184.
- Seo, K.-H. (2009), Statistical-dynamical prediction of the Madden-Julian oscillation using NCEP Climate Forecast System, *Int. J. Climatol.*, *29*, 2146–2155.
- Seo, K.-H., and W. Wang (2010), The Madden-Julian oscillation simulated in the NCEP Climate Forecast System model: The importance of stratiform heating, *J. Clim.*, *23*, 4770–4793.
- Seo, K.-H., W. Wang, J. Gottschalck, Q. Zhang, J.-K. E. Schemm, W. R. Higgins, and A. Kumar (2009), Evaluation of MJO forecast skill from several statistical and dynamical forecast models, *J. Clim.*, *22*, 2372–2388.
- Shelly, A., P. Xavier, D. Copey, T. Johns, J. M. Rodriguez, S. Milton, and N. P. Klingaman (2014), Coupled versus uncoupled hindcast simulations of the Madden-Julian Oscillation in the Year of Tropical Convection, *J. Geophys. Res.*, *41*, 5670–5677, doi:10.1002/2013GL059062.
- Slingo, J. M., and R. A. Madden (1991), Characteristics of the tropical intraseasonal oscillation in the NCAR community climate model, *Q. J. R. Meteorol. Soc.*, *117*, 1129–1169.
- Slingo, J. M., et al. (1996), Intraseasonal oscillations in 15 atmospheric general circulation models: Results from an AMIP diagnostic subproject, *Clim. Dyn.*, *12*(5), 325–357.
- Sobel, A. H., J. Nilsson, and L. M. Polvani (2001), The weak temperature gradient approximation and balanced tropical moisture waves, *J. Atmos. Sci.*, *58*, 3650–3665.
- Song, X., and G. J. Zhang (2011), Microphysics parameterization for convective clouds in a global climate model: Description and single-column model tests, *J. Geophys. Res.*, *116*, D02201, doi:10.1029/2010JD014833.
- Song, X., G. J. Zhang, and J.-L. F. Li (2012), Evaluation of microphysics parameterization for convective clouds in the NCAR Community Atmosphere Model CAM5, *J. Clim.*, *25*, 8568–8590.
- Sperber, K. R., and H. Annamalai (2008), Coupled model simulations of boreal summer intraseasonal (30–50 day) variability. Part 1: Systematic errors and caution on the use of metrics, *Clim. Dyn.*, *31*, 345–372.
- Sperber, K. R., and D. Kim (2012), Simplified metrics for the identification of the Madden-Julian oscillation in models, *Atmos. Sci. Lett.*, *13*, 187–193.
- Sperber, K. R., J. M. Slingo, and H. Annamalai (2000), Predictability and the relationship between subseasonal and interannual variability during the Asian summer monsoons, *Q. J. R. Meteorol. Soc.*, *126*, 2545–2574.
- Sperber, K. R., J. M. Slingo, D. E. Waliser, and P. M. Inness (2008), Coarse-resolution models only partly cloudy (Comment on “A Madden-Julian oscillation event realistically simulated by a global cloud-resolving model” by Miura et al.), *Science*, *320*, 612–613.
- Stephens, G. L., P. J. Webster, R. H. Johnson, R. Engelen, and T. S. L’Ecuyer (2004), Observational evidence for the mutual regulation of the tropical hydrological cycle and tropical sea surface temperatures, *J. Clim.*, *17*, 2213–2224.
- Straub, K. H. (2013), MJO initiation in the real-time multivariate MJO index, *J. Clim.*, *26*, 1130–1151.
- Subramanian, A. C., M. Jochum, A. J. Miller, R. Murtugudde, R. Neale, and D. E. Waliser (2011), The Madden-Julian oscillation in CCSM4, *J. Clim.*, *24*, 6261–6282.
- Thayer-Calder, K., and D. A. Randall (2009), The role of convective moistening in the Madden-Julian oscillation, *J. Atmos. Sci.*, *66*, 3297–3312.
- Vitart, F. (2014), Evolution of ECMWF sub-seasonal forecast skill scores, *Q. J. R. Meteorol. Soc.*, *683*, 1889–1899.

- Vitart, F., and F. Molteni (2010), Simulation of the Madden-Julian oscillation and its teleconnections in the ECMWF forecast system, *Q. J. R. Meteorol. Soc.*, *136*, 842–855.
- Vitart, F., S. J. Woolnough, M. A. Balmaseda, and A. Tompkins (2007), Monthly forecast of the Madden-Julian oscillation using a coupled GCM, *Mon. Weather Rev.*, *135*, 2700–2715.
- Volz, A., et al. (2013), The CNRM-CM5.1 global climate model: Description and basic evaluation, *Clim. Dyn.*, *40*, 2091–2121.
- von Salzen, K., et al. (2013), The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4). Part I: Representation of physical processes, *Atmos. Ocean*, *51*, 104–125, doi:10.1080/07055900.2012.755610.
- Waliser, D. E. (2012), Predictability and forecasting, in *Intraseasonal Variability in the Atmosphere-Ocean Climate System*, 2nd ed., edited by W. K. M. Lau and D. E. Waliser, chap. 12, pp. 433–476, Springer/Praxis, Berlin.
- Waliser, D. E., et al. (2012), The “Year” of tropical convection (May 2008–April 2010), *Bull. Am. Meteorol. Soc.*, *93*, 1189–1218.
- Walters, D. N., et al. (2011), The Met Office Unified Model Global Atmosphere 3.0/3.1 and JULES Global Land 3.0/3.1 configurations, *Geosci. Model Dev.*, *4*, 919–941.
- Wang, B., and H. Rui (1990), Dynamics of the coupled moist Kelvin-Rossby wave on an equatorial  $\beta$  plane, *J. Atmos. Sci.*, *47*, 397–413.
- Wang, W., and M. E. Schlesinger (1999), The dependence on convective parameterization of the tropical intraseasonal oscillation simulated by the UIUC 11-layer atmospheric GCM, *J. Clim.*, *12*, 1423–1457.
- Watanabe, M., et al. (2010), Improved climate simulation by MIROC5: Mean states, variability and climate sensitivity, *J. Clim.*, *23*, 6312–6335.
- Webster, P. J., and S. Yang (1992), Monsoon and ENSO: Selectively interactive systems, *Q. J. R. Meteorol. Soc.*, *118*, 877–926.
- Wheeler, M. C., and H. H. Hendon (2004), An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction, *Mon. Weather Rev.*, *132*, 1917–1932, doi:10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.
- Wheeler, M. C., H. H. Hendon, S. Cleland, H. Meinke, and A. Donald (2009), Impacts of the Madden-Julian oscillation on Australian rainfall and circulation, *J. Clim.*, *22*, 1482–1498.
- Woolnough, S. J., et al. (2010), Modelling convective processes during the suppressed phase of a Madden-Julian oscillation: Comparing single-column models with cloud-resolving models, *Q. J. R. Meteorol. Soc.*, *136*, 333–353.
- Xavier, P. K. (2012), Intraseasonal convective moistening in CMIP3 models, *J. Clim.*, *25*, 2569–2577.
- Xavier, P. K., et al. (2015), Vertical structure and physical processes of the Madden-Julian Oscillation: Biases and uncertainties at short range, *J. Geophys. Res. Atmos.*, *120*, doi:10.1002/2014JD022718.
- Xie, S., H.-Y. Ma, J. S. Boyle, S. A. Klein, and Y. Zhang (2012), On the correspondence between short- and long-timescale systematic errors in CAM4/CAM5 for the Years of Tropical Convection, *J. Clim.*, *25*, 7937–7955.
- Yasunari, T. (1979), Cloudiness fluctuations associated with the Northern Hemisphere summer monsoon, *J. Meteorol. Soc. Jpn.*, *57*, 227–242.
- Yukimoto, S., et al. (2012), A new global model of meteorological research institute: MRI-CGCM3—Model description and basic performance, *J. Meteorol. Soc. Jpn.*, *90A*, 23–64.
- Zhang, C. (2005), Madden-Julian Oscillation, *Rev. Geophys.*, *43*, RG2003, doi:10.1029/2004RG000158.
- Zhou, L., R. B. Neale, M. Jochum, and R. Murtugudde (2012), Improved Madden-Julian oscillations with improved physics: The impact of modified convection parameterizations, *J. Clim.*, *25*, 1116–1136.