RMetS
Royal Meteorological Society

Test models for filtering and prediction of moisture-coupled tropical waves

John Harlim[a]* and Andrew J. Majda[b]

[a]*Department of Mathematics, North Carolina State University, Raleigh, NC, USA*
[b]*Department of Mathematics and Center for Atmospheric and Ocean Science, Courant Institute of Mathematical Sciences, New York University, NY, USA*

*Correspondence to: J. Harlim, Department of Mathematics, North Carolina State University, Box 8205, Raleigh, NC 27695, USA. E-mail: jharlim@ncsu.edu

The filtering/data assimilation and prediction of moisture-coupled tropical waves is a contemporary topic with significant implications for extended-range forecasting. The development of efficient algorithms to capture such waves is limited by the unstable multiscale features of tropical convection which can organize large-scale circulations and the sparse observations of the moisture-coupled wave in both the horizontal and vertical. The approach proposed here is to address these difficult issues of data assimilation and prediction through a suite of analogue models which, despite their simplicity, capture key features of the observational record and physical processes in moisture-coupled tropical waves. The analogue models emphasized here involve the multicloud convective parametrization based on three cloud types (congestus, deep, and stratiform) above the boundary layer. Two test examples involving an MJO-like turbulent travelling wave and the initiation of a convectively coupled wave train are introduced to illustrate the approach. A suite of reduced filters with judicious model errors for data assimilation of sparse observations of tropical waves, based on linear stochastic models in a moisture-coupled eigenmode basis is developed here and applied to the two test problems. Both the reduced filter and 3D-Var with a full moist background covariance matrix can recover the unobserved troposphere humidity and precipitation rate; on the other hand, 3D-Var with a dry background covariance matrix fails to recover these unobserved variables. The skill of the reduced filtering methods in recovering the unobserved precipitation, congestus, and stratiform heating rates as well as the front-to-rear tilt of the convectively coupled waves exhibits a subtle dependence on the sparse observation network and the observation time. Copyright © 2012 Royal Meteorological Society

Key Words: tropical data assimilation; reduced stochastic filters; multicloud models; Madden–Julian Oscillation

Received 9 January 2012; Revised 28 February 2012; Accepted 26 March 2012; Published online in Wiley Online Library

Citation: Harlim J, Majda AJ. 2012. Test models for filtering and prediction of moisture-coupled tropical waves. *Q. J. R. Meteorol. Soc.* DOI:10.1002/qj.1956

1. Introduction

Observational data indicate that, through the complex interaction of heating and moist convection, tropical atmosphere flows are organized on a hierarchy of scales (Nakazawa, 1988) ranging from cumulus clouds of a few kilometres to mesoscale convective systems (Houze, 2004) to equatorial synoptic-scale convectively coupled Kelvin waves and two-day waves (Kiladis, *et al.*, 2009) to planetary-scale intraseasonal organized circulations such

Copyright © 2012 Royal Meteorological Society

as the Madden–Julian Oscillation (MJO; Zhang, 2005). These moisture- coupled tropical waves like the MJO exert a substantial influence on intraseasonal prediction in the Tropics, Subtropics, and midlatitudes (Moncrieff, *et al.*, 2007). Despite the continued research efforts by the climate community, the present coarse-resolution general circulation models (GCMs), used for prediction of weather and climate, poorly represent variability associated with tropical convection (Lau and Waliser, 2005; Zhang, 2005; Lin, *et al.*, 2006). Given the importance of moisture-coupled tropical waves for short-term climate and medium- to long-range weather prediction, new strategies for the filtering or data assimilation and prediction of moisture-coupled tropical waves are needed and this is the topic of the present article.

The approach proposed here is to address the issues of data assimilation and prediction through a suite of analogue models which, despite their simplicity, capture key features of the observational record and physical processes in moisture-coupled tropical waves. This approach is analogous to the use of various versions of the Lorenz-96 model (Lorenz, 1996; Wilks, 2005; Majda, *et al.*, 2005; Abramov and Majda, 2007; Crommelin and Vanden-Eijnden, 2008; Harlim and Majda, 2008a, 2010a; Majda and Harlim, 2012, and references therein) to gain insight into basic issues for midlatitude filtering, prediction, and parametrization. The viability of this approach for moisture-coupled tropical waves rests on recent advances in simplified modelling of convectively coupled tropical waves and the MJO which predict key physical features of these waves such as their phase speed, dispersion relation, front-to-rear tilt (Kiladis, *et al.*, 2005, 2009), and circulation in qualitative agreement with observations (Khouider and Majda, 2006a,b, 2007, 2008a,b; Majda, *et al.*, 2007; Majda and Stechmann, 2009a,b, 2011) through simplified moisture-coupled models. The analogue models emphasized here involve the multicloud convective parametrization based on three cloud types (congestus, deep, and stratiform) above the boundary layer (Khouider and Majda, 2006a,b, 2007, 2008a,b). The convective closure of the multicloud model takes into account the energy available for congestus and deep convection and uses a nonlinear moisture switch that allows for natural transitions between congestus and deep convection as well as for stratiform downdraughts which cool and dry the boundary layer. As a simplified model with two vertical baroclinic modes, the multicloud model is very successful in capturing most of the spectrum of convectively coupled waves (Kiladis, *et al.*, 2009; Khouider and Majda, 2008b; Han and Khouider, 2010), as well as the nonlinear organization of large-scale envelopes mimicking across-scale interactions of the MJO and convectively coupled waves (Khouider and Majda, 2007, 2008a). Furthermore, the multicloud parametrization has been used in the next-generation GCM (High-Order Method Modeling Environment, HOMME) of the National Center for Atmospheric Research and is very successful in simulating the MJO and convectively coupled equatorial waves, at a coarse resolution of 170 km in the idealized case of a uniform sea-surface temperature (SST) aquaplanet setting (Khouider, *et al.*, 2011). A stochastic version of the multicloud model has been utilized recently as a novel convective parametrization to improve the physical variability of deficient deterministic convective parametrizations (Khouider, *et al.*, 2010; Frenkel, *et al.*, 2012).

The filtering skill for the recovery of troposphere moisture, heating profiles, precipitation, and vertical tilts in circulation and temperature from sparse noisy partial observations is studied here for a turbulent MJO-like travelling wave (Majda, *et al.*, 2007) and for the temporal development of a convectively coupled wave train. A suite of filters with judicious model errors, based on linear stochastic models (Harlim and Majda, 2008a, 2010a; Majda and Harlim, 2012) on a moisture-coupled eigenmode basis is developed here and applied to the two test problems as well as related 3D-Var algorithms with a full moist background covariance matrix or a dry background covariance (Žagar, *et al.*, 2004a,b). These results are the first demonstration of the utility of the analogue multicloud models for gaining insight for data assimilation and prediction of moisture-coupled tropical waves.

The plan for the remainder of the article is as follows. In section 2, the suite of simplified tropical models for filtering and prediction is reviewed; section 3 illustrates two simplified cases, an MJO analogue wave (Majda, *et al.*, 2007) and the temporal development of a convectively coupled tropical wave train which illustrate phenomena in the models and also serve as examples for filtering in subsequent sections of the article. The suite of filters with judicious model errors for moisture-coupled tropical waves is introduced in section 4. Filtering skill for these algorithms applied to the MJO analogue wave and the development of a convectively coupled wave train is reported in section 5. Section 6 is a concluding discussion and summary.

2. Test models with moisture-coupled tropical waves

The test models proposed here begin with two coupled shallow-water systems: a direct heating mode forced by a bulk precipitation rate from deep penetrative clouds (Neelin and Zeng, 2000) and a second vertical baroclinic mode forced by the upper-level heating (cooling) and lower-level cooling (heating) of stratiform and congestus clouds, respectively (Khouider and Majda, 2006a). Below, for simplicity in exposition, we present these equations without explicit nonlinear advection effects and coupling to barotropic winds. This allows us to emphasize moisture-coupled tropical waves here, but we comment later in this section about how nonlinear advection and barotropic winds enrich the dynamics of the test models. Thus, the test models begin with two equatorial shallow-water equations

$$\left.\begin{aligned} &\frac{\partial \mathbf{v}_j}{\partial t}+\overline{\mathbf{U}}\cdot\nabla\mathbf{v}_j+\beta y\mathbf{v}_j^{\perp}-\theta_j=-C_d u_0\mathbf{v}_j-\frac{1}{\tau_{\mathrm{w}}}\mathbf{v}_j,\\ &\frac{\partial \theta_1}{\partial t}+\overline{\mathbf{U}}\cdot\nabla\theta_1-\operatorname{div}\mathbf{v}_1=P+S_1,\\ &\frac{\partial \theta_2}{\partial t}+\overline{\mathbf{U}}\cdot\nabla\theta_2-\frac{1}{4}\operatorname{div}\mathbf{v}_2=-H_{\mathrm{s}}+H_{\mathrm{c}}+S_2, \end{aligned}\right\} \tag{1}$$

for $j=1,2$. The equations in (1) are obtained by a Galerkin projection of the hydrostatic primitive equations with constant buoyancy frequency onto the first two baroclinic modes. More details of their derivation may be found in (Neelin and Zeng, 2000; Frierson, *et al.*, 2004; Stechmann and Majda, 2009). In (1), $\mathbf{v}_j=(u_j,v_j)_{j=1,2}$ represent the first and second baroclinic velocities assuming $G(z)=\sqrt{2}\cos(\pi z/H_{\mathrm{T}})$ and $G(2z)=\sqrt{2}\cos(2\pi z/H_{\mathrm{T}})$

Copyright © 2012 Royal Meteorological Society

vertical profiles, respectively, while $\theta_j, j = 1, 2$ are the corresponding potential temperature components with the vertical profiles $G'(z) = \sqrt{2}\sin(\pi z/H_T)$ and $2G'(2z) = 2\sqrt{2}\sin(2\pi z/H_T)$, respectively. Therefore, the total velocity field is approximated by

$$\mathbf{V} \approx \overline{\mathbf{U}} + G(z)\mathbf{v}_1 + G(2z)\mathbf{v}_2,$$
$$w \approx -\frac{H_T}{\pi}\Big[G'(z)\mathrm{div}\,\mathbf{v}_1 + \frac{1}{2}G'(2z)\mathrm{div}\,\mathbf{v}_2\Big],$$

where $\mathbf{V}$ is the horizontal velocity and w is the vertical velocity. The total potential temperature is given approximately by

$$\Theta \approx z + G'(z)\theta_1 + 2G'(2z)\theta_2.$$

Here $H_T \approx 16$ km is the height of the tropical troposphere with $0 \le z \le H_T$ and $\mathbf{v}_j^{\perp} = (-v_j, u_j)$ while $\overline{\mathbf{U}}$ is the incompressible barotropic wind which is set to zero hereafter, for the sake of simplicity. In (1), $P \ge 0$ models the heating from deep convection while H_s, H_c are the stratiform and congestus heating rates. Conceptually, the direct heating mode has a positive component and serves to heat the whole troposphere and is associated with a vertical shear flow. The second baroclinic mode is heated by the congestus clouds, H_c, from below and by the stratiform clouds, H_s, from above and therefore cooled by H_c from above and by H_s from below. It is associated with a jet shear flow in the middle troposphere (Khouider and Majda, 2006a, 2007, 2008a,b). The terms S_1 and S_2 are the radiative cooling rates associated with the first and second baroclinic modes respectively.

The system of equations in (1) is augmented by an equation for the boundary-layer equivalent potential temperature, θ_{eb}, and another for the vertically integrated moisture content, q.

$$\left.\begin{aligned} \frac{\partial \theta_{eb}}{\partial t} &= \frac{1}{h_b}(E - D), \\ \frac{\partial q}{\partial t} + \overline{\mathbf{U}}\cdot\nabla q + \mathrm{div}(\mathbf{v}_1 q + \widetilde{\alpha}\mathbf{v}_2 q) & \\ + \widetilde{Q}\,\mathrm{div}(\mathbf{v}_1 + \widetilde{\lambda}\mathbf{v}_2) &= -\frac{2\sqrt{2}}{\pi}P + \frac{1}{H_T}D. \end{aligned}\right\} \quad (2)$$

In (2), $h_b \approx 500$ m is the height of the moist boundary layer while $\widetilde{Q}$, $\widetilde{\lambda}$, and $\widetilde{\alpha}$ are parameters associated with a prescribed moisture background and perturbation vertical profiles. According to the first equation in (2), θ_{eb} changes in response to the downdraughts, D, and the sea surface evaporation E. A detailed pedagogical derivation of the moisture equation starting from the equations of bulk cloud microphysics is presented in Khouider and Majda (2006b). The approximate numerical values of $\widetilde{\lambda} = 0.8$ and $\widetilde{\alpha} = 0.1$, follow directly from the derivation, while the coefficient $\widetilde{Q}$ arises from the background moisture gradient. We use the standard value $\widetilde{Q} \approx 0.9$ (Neelin and Zeng, 2000; Frierson, *et al.*, 2004).

In full generality, the parametrizations in (1) and (2) automatically have conservation of an approximation to vertically integrated moist static energy. Notice that, the precipitation rate in (2), balances the vertical average of the total convective heating rate in (1), therefore leading to the conservation of the vertical average of the equivalent potential temperature

$$\langle\theta_e\rangle = \langle Q(z)\rangle + q + \langle\Theta\rangle + \frac{h_b}{H_T}\theta_{eb}$$

when the external forces, namely, the radiative cooling rates, S_1, S_2, and the evaporative heating, E, are set to zero. Also note that the sensible heating flux has been ignored in (1) for simplicity since this is a relatively small contribution in the Tropics. Here and elsewhere in the text

$$\langle f\rangle = \frac{1}{H_T}\int_0^{H_T} f(z)\,\mathrm{d}z.$$

The equations in (1) and (2) for the prognostic variables $q, \theta_{eb}, \theta_j, \mathbf{v}_j, j = 1, 2$, are written in non-dimensional units where the equatorial Rossby deformation radius, $L_e \approx 1\,500$ km is the length-scale, the first baroclinic dry gravity wave speed, $c \approx 50\,\mathrm{m\,s^{-1}}$, is the velocity scale, $T = L_e/c \approx 8$ h is the associated time-scale, and the dry-static stratification $\overline{\alpha} = H_T N^2\theta_0/\pi g \approx 15$ K is the temperature unit scale. For convenience, the basic bulk parameters of the model are listed in Table 1.

2.1. *The convective parametrization*

The surface evaporative heating, E, in (2) obeys an adjustment equation toward the boundary-layer saturation equivalent potential temperature, θ^*_{eb},

$$\frac{1}{h_b}E = \frac{1}{\tau_e}(\theta^*_{eb} - \theta_{eb}), \quad (3)$$

with τ_e is the evaporative time-scale. The middle tropospheric equivalent potential temperature anomaly is defined approximately by

$$\theta_{em} \approx q + \frac{2\sqrt{2}}{\pi}(\theta_1 + \alpha_2\theta_2). \quad (4)$$

Notice that the coefficient $2\sqrt{2}/\pi$ in (4) results from the vertical average of the first baroclinic potential temperature,

Table 1. Bulk constants in the two-layer mode model.

Symbol	Value	Definition
C_d	0.001	Boundary-layer turbulent momentum friction
h_b	500 m	Boundary-layer height
H_T	16 km	Height of the tropical troposphere.
L_e	$\approx$ 1500 km	Equatorial deformation radius, length-scale
N	0.01 $\mathrm{s^{-1}}$	Brunt–Väisälä buoyancy frequency
$\widetilde{Q}$	0.9	Moisture stratification factor
$T = L_e/c$	$\approx$ 8 h	Time-scale
$\widetilde{\alpha}$	0.1	Baroclinic contribution to the moisture (nonlinear) convergence associated with the moisture anomalies
$\overline{\alpha}$	$\approx$ 15 K	Dry static stratification, temperature scale
α_2	0.1	Relative contribution of θ_2 to the middle troposphere θ_e
θ_0	300 K	Reference temperature
$\widetilde{\lambda}$	0.8	Baroclinic contribution to the moisture convergence associated with the moisture background
τ_R	50 days	Newtonian cooling relaxation time
τ_w	75 days	Rayleigh-wind friction relaxation time

Copyright © 2012 Royal Meteorological Society

$\sqrt{2}\theta_1 \sin(\pi z/H_T)$, while the small value for α_2 adds a non-zero contribution from θ_2 to θ_{em} to include its contribution from the lower middle troposphere although its vertical average is zero. The multicloud model closure is based on a moisture switch parameter Λ (Khouider and Majda, 2006a, 2008a,b), which serves as a measure for the moistness and dryness of the middle troposphere. When the discrepancy between the boundary layer and the middle troposphere equivalent potential temperature is above some fixed threshold, θ^+, the atmosphere is defined as dry. Moist parcels rising from the boundary layer will have their moisture quickly diluted by entrainment of dry air, hence losing buoyancy and ceasing to convect. In this case, we set $\Lambda = 1$ which automatically inhibits deep convection in the model (see below). When this discrepancy is below some lower value, θ^-, we have a relatively moist atmosphere and we set $\Lambda = \Lambda^* < 1$. The function Λ is then interpolated (linearly) between these two values. More precisely, we set

$$\Lambda = \begin{cases} 1 & \text{if } \theta_{eb} - \theta_{em} > \theta^+ \\ A(\theta_{eb} - \theta_{em}) + B & \text{if } \theta^- \le \theta_{eb} - \theta_{em} \le \theta^+ \\ \Lambda^* & \text{if } \theta_{eb} - \theta_{em} < \theta^-. \end{cases} \quad (5)$$

The value of θ^- represents a threshold below which the free troposphere is locally moist and 'accepts' only deep convection while the value of θ^+ defines complete dryness.

Therefore, the precipitation, P, and the downdraughts, D, obey

$$P = \frac{1-\Lambda}{1-\Lambda^*} P_0 \quad \text{and} \quad D = \Lambda D_0, \quad (6)$$

while the stratiform and congestus heating rate, H_s and H_c, solve the relaxation-type equations

$$\frac{\partial H_s}{\partial t} = \frac{1}{\tau_s}(\alpha_s P - H_s) \quad (7)$$

and

$$\frac{\partial H_c}{\partial t} = \frac{1}{\tau_c}\left(\alpha_c \frac{\Lambda - \Lambda^*}{1-\Lambda^*}\frac{D}{H_T} - H_c\right), \quad (8)$$

respectively. The dynamical equations in (1), (2), (7), and (8) define the multicloud model. Notice that, as anticipated above, when the middle troposphere is dry, $\Lambda = 1$, deep convection is completely inhibited, even if P_0, i.e. convective available potential energy (CAPE) is positive, whereas congestus heating is favoured. Other variants of the equation in (8) for H_c can be utilized where changes in H_c respond to low-level CAPE (Khouider and Majda, 2008a,b).

The quantities P_0 and D_0 represent respectively the maximum allowable deep convective heating/precipitation and downdraughts, independent of the value of the switch function Λ. Notice that conceptually the model is not bound to any type of convective parametrization. A Betts–Miller relaxation type parametrization as well as a CAPE parametrization can be used to set up a closure for P_0. Here we let

$$P_0 = \frac{1}{\tau_{conv}}\left[a_1\theta_{eb} + a_2(q - \widehat{q}) - a_0(\theta_1 + \gamma_2\theta_2)\right]^+, \quad (9)$$

where $f^+ = \max(f, 0)$ and $\widehat{q}$ is a threshold constant value measuring a significant fraction of the tropospheric saturation and τ_{conv}, a_1, a_2, a_0 are parameters specified below. In particular, the coefficient a_0 is related to the inverse buoyancy relaxation time of Fuchs and Raymond (2002).

The downdraughts are closed by

$$D_0 = \frac{m_0}{\overline{P}}\left[\overline{P} + \mu_2(H_s - H_c)\right]^+(\theta_{eb} - \theta_{em}), \quad (10)$$

where m_0 is a scaling of the downdraught mass flux and $\overline{P}$ is a prescribed precipitation/deep convective heating at radiative convective equilibrium. Here μ_2 is a parameter allowing for stratiform and congestus mass flux anomalies (Majda and Shefter, 2001; Majda, *et al.*, 2004). Finally the radiative cooling rates, S_1, S_2 in (1) are given by a simple Newtonian cooling model

$$S_j = -Q^0_{R,j} - \frac{1}{\tau_R}\theta_j, \qquad j = 1, 2, \quad (11)$$

where $Q^0_{R,j}, j = 1, 2$ are the radiative cooling rates at radiative convective equilibrium (RCE). This is a spatially homogeneous steady-state solution where the convective heating is balanced by the radiative cooling. The basic constants in the model convective parametrization and the typical values utilized here are given in Table 2. The physical features incorporated in the multi-cloud model are discussed in detail in (Khouider and Majda, 2006a, 2007, 2008a,b).

Table 2. Parameters in the convective parametrization. The parameters in the middle panel will be chosen differently for the MJO-analogue case in section 3.1 and the temporal development of a convectively coupled wave train in section 3.2. The parameters in the lower panel are determined at the RCE state.

Symbol	Value	Definition
A, B		Linear fitting constant interpolating the switch function Λ
$Q^0_{R,1}$	1 K day^{-1}	Second baroclinic radiative cooling rate
α_s	0.25	Stratiform heating adjustment coefficient
γ_2	0.1	Relative contribution of θ_2 to convective parametrization
$\theta^*_{eb} - \overline{\theta}_{eb}$	10 K	Discrepancy between boundary layer θ_e at its saturated value and at the RCE state
$\theta^\pm$	10, 20 K	Temperature threshold used to define the switch function Λ
a_0		Inverse buoyancy time-scale of convective parametrization
a_1		Relative contribution of θ_{eb} to convective parametrization
a_2		Relative contribution of q to convective parametrization
α_c		Congestus heating adjustment coefficient
$\overline{\theta}_{eb} - \overline{\theta}_{em}$		Discrepancy between boundary and middle troposphere potential temperature at RCE value
Λ^*		Lower threshold of the switch function Λ
μ_2		Relative contribution of stratiform and congestus mass flux anomalies to the downdraughts
τ_c		Congestus heating adjustment time
τ_{conv}		Deep convection adjustment time
τ_s		Stratiform heating adjustment time
m_0		Scaling of downdraught mass flux
$\widehat{q}$		Threshold beyond which condensation takes place in the Betts–Miller scheme
$Q^0_{R,2}$		Second baroclinic radiative cooling rate
τ_e	≈ 8 h	Evaporation time-scale in the boundary layer

Copyright © 2012 Royal Meteorological Society

2.2. Moisture-coupled phenomena in the test models

As already noted in the introduction, the dynamic multicloud models in (1), (2), (7), (8) capture a number of observational features of equatorial convectively coupled waves and the MJO. These phenomena occur in multi-wave dynamical models with strong moisture coupling through (2), nonlinear on–off switches like (5), (9), (10) and nonlinear saturation of moisture-coupled instabilities (Khouider and Majda, 2006a, 2007, 2008a,b; Khouider, *et al.*, 2011). All of these features present major challenges for contemporary data assimilation and prediction strategies. Two detailed analogue examples are presented in section 3.

As described in detail in Khouider and Majda (2006b), the multicloud models in a limiting regime also include the quasi-equilibrium models (Neelin and Zeng, 2000; Frierson, *et al.*, 2004; Pauluis, *et al.*, 2008) which mimic the Betts–Miller and Arakawa–Schubert parametrizations of GCMs. Such models arise formally by keeping the first baroclinic mode in (1), retaining the moisture equation in (2) with $D = 0$, setting $\Lambda = 1$ in (6), and using P_0 in (9) with $a_1 = 0$ while ignoring all remaining boundary-layer and cloud equations. There are many interesting exact solutions of the nonlinear dynamics with moisture switches in this quasi-equilibrium regime, e.g. large-scale precipitation fronts, which serve as interesting test problems for filtering with nonlinear switches and moisture-coupled waves (Frierson, *et al.*, 2004; Pauluis, *et al.*, 2008; Stechmann and Majda, 2006); the behaviour of finite-ensemble Kalman filters (Evensen, 1994; Anderson, 2001; Bishop, *et al.*, 2001; Hunt, *et al.*, 2007) and particle filters (van Leeuwen, 2010; Anderson, 2010) are particularly interesting in this context with moisture-coupled switches and exact solutions. Žagar (2012) give other interesting uses of similar models as tests for tropical data assimilation. However, rigorous mathematical theory establishes that these quasi-equilibrium models have no instabilities or positive Lyapunov exponents (Majda and Souganidis, 2010), unlike realistic tropical convection and the full multicloud models. More realism in the quasi-equilibrium tropical models can be achieved by allowing active barotropic dynamics and coupled nonlinear advection which allows for tropical–extratropical wave interactions (Lin, *et al.*, 2000; Majda and Biello, 2003; Biello and Majda, 2004). Examples with these features are developed by Khouider and Majda (2005a,b).

3. Examples of moisture-coupled tropical waves in the test model

In this section, we describe two concrete examples with solutions which will be used as the truth for generating synthetic observations (as we will describe in section 4). The two specific examples include an MJO-like travelling wave (Majda, *et al.*, 2007) and the initiation of a convectively coupled wave train that mimics the solutions of explicit simulations with a cloud-resolving model (Grabowski and Moncrieff, 2001). Following the basic set-up in Khouider and Majda (2006a, 2007), we consider the multicloud model in (1), (2), (7), (8) on a periodic equatorial ring without rotation, $\beta = 0$, without barotropic wind, $\overline{\mathbf{U}} = 0$, and with a uniform background SST given by constant $\theta^*_{\rm eb}$. With this set-up, the wind velocity in (1), (2) has only the zonal wind component, $\mathbf{v}_j = u_j$, resolved at every 40 km on an equatorial belt of 40 000 km.

3.1. An MJO-like turbulent travelling wave

In our first example, we consider the parameter regime for an intraseasonal MJO-like turbulent travelling wave. Following Majda, *et al.* (2007), we set the bulk parameters in Table 1, $\widetilde{Q} = 1$, $\widetilde{\lambda} = 0.6$, $C_d = 10^{-5}$, $\tau_{\rm w} = 150$ days, $\tau_{\rm R} = 50$ days and the convective parameters in Table 2, $\overline{\theta}_{\rm eb} - \overline{\theta}_{\rm em} = 12$ K, $a_0 = 12$, $a_1 = 0.1$, $a_2 = 0.9$, $\mu_2 = 0.5$, $\alpha_{\rm c} = 0.5$, $\Lambda^* = 0.2$. The intraseasonal time-scale is generated through $\tau_{\rm conv} = 12$ h which is consistent with the current observational estimates for large-scale consumption of CAPE, and $\tau_{\rm s} = \tau_{\rm c} = 7$ days which is also consistent with the current observational record for low-level moistening and congestus cloud development in the MJO.

The linear stability analysis for this parameter regime has been studied in detail in Majda, *et al.* (2007). Here, we summarize for convenience some of the important features for eastward propagating waves: the unstable wavenumbers 2 and 3 have growth rates of roughly $(30\,\text{days})^{-1}$ and phase speed of 6.9 and 5.8 m s^{-1}, respectively. These unstable modes have westward-tilted vertical structure for heating, velocity, and temperature, with clear first and second baroclinic mode contributions and low-level higher potential temperature leading and within the deep convection (Figure 9 below). In Figure 1, we show the contour plot of the precipitation P (which is exactly the deep heating rate for this model) at the statistical steady state from a numerical simulation between 5000 and 5200 days. The main feature here is an eastward moving wavenumber 2 MJO-like wave with phase speed 6.1 m s^{-1}. Within the envelope of this wave are intense westward moving small-scale fluctuations. These fluctuations occur irregularly and there are often long breaks between intense deep convective events. All of these features are observed in the MJO (Zhang, 2005).

3.2. Initiation of a convectively coupled wave train

In this second example, we consider the three-cloud model with enhanced congestus heating (Khouider and Majda, 2008a) with slightly different parametrization from the

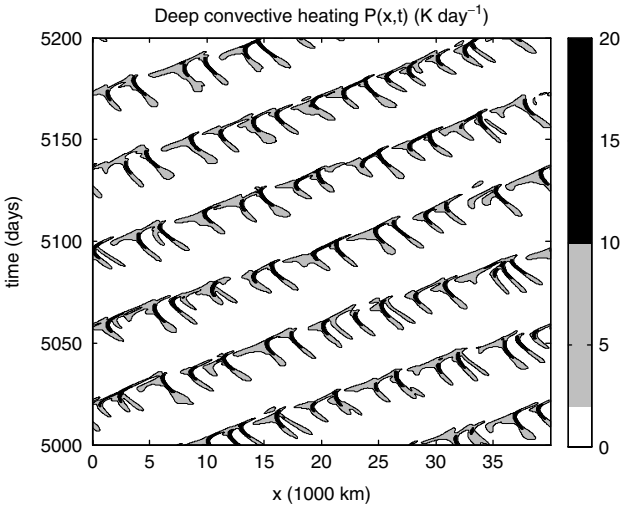

Figure 1. Contour plot of the deep convective heating $P(x, t)$ from a numerical simulation of the multicloud model with parameter values in section 3.1, Tables 1 and 2. Heating values > 2 K day^{-1} are shaded in grey and > 10 K day^{-1} in black.

Copyright © 2012 Royal Meteorological Society

above. In particular, the total precipitation, P, is different from the deep convection heating rate, $H_{\rm d}$, and is defined as

$$P = \frac{2\sqrt{2}}{\pi}(H_{\rm d} + \xi_{\rm s} H_{\rm s} + \xi_{\rm c} H_{\rm c}), \tag{12}$$

allowing for stratiform and congestus rain. The key feature in this new parametrization is attributed to the asymmetric heating rate contribution in the upper- and lower-level atmosphere with non-zero $\xi_{\rm s}$ and $\xi_{\rm c}$, respectively. This new feature replaces the first baroclinic heating equation in (1) with

$$\frac{\partial \theta_1}{\partial t} - \frac{\partial u_1}{\partial x} = H_{\rm d} + \xi_{\rm s} H_{\rm s} + \xi_{\rm c} H_{\rm c} + S_1. \tag{13}$$

The moisture equation in (2) remains unchanged, except that now we remove the scale factor $2\sqrt{2}/\pi$ in front of P since it is already included in (12).

The new congestus parametrization uses exactly the same switch function Λ in (5) with the middle-troposphere equivalent potential temperature approximation in (4). The precipitation, P, in (6) is replaced by

$$H_{\rm d} = (1 - \Lambda) Q_{\rm d}, \tag{14}$$

with bulk energy available for deep convection given by

$$Q_{\rm d} = \left[\overline{Q} + \frac{1}{\tau_{\rm conv}}\{a_1 \theta_{\rm eb} + a_2 q - a_0(\theta_1 + \gamma_2 \theta_2)\}\right]^+. \tag{15}$$

In (15), parameter $\overline{Q}$ is the bulk convective heating determined at the RCE state. The downdraught in (6) is also replaced by

$$D = \frac{m_0}{\overline{Q}}\left[\overline{Q} + \mu_2(H_{\rm s} - H_{\rm c})\right]^+ (\theta_{\rm eb} - \theta_{\rm em}). \tag{16}$$

Compared to (6), this new parametrization assigns $\Lambda^* = 0$ for the moisture switch minimum threshold and ignores the factor Λ in the original downdraught equation. The corresponding dynamical equations for the stratiform and congestus heating are

$$\frac{\partial H_{\rm s}}{\partial t} = \frac{1}{\tau_{\rm s}}(\alpha_{\rm s} H_{\rm d} - H_{\rm s}), \tag{17}$$

$$\frac{\partial H_{\rm c}}{\partial t} = \frac{1}{\tau_{\rm c}}(\alpha_{\rm c} \Lambda Q_{\rm c} - H_{\rm c}), \tag{18}$$

where

$$Q_{\rm c} = \left[\overline{Q} + \frac{1}{\tau_{\rm conv}}\{\theta_{\rm eb} - a_0'(\theta_1 + \gamma_2 \theta_2)\}\right]^+ \tag{19}$$

denotes a 'bulk energy' for congestus heating.

In our numerical experiment, we use the same parameter values as in Khouider and Majda (2008a). The bulk constants in Table 1 are not changed. The convective parameters in Table 2 are used with $\Lambda^* = 0$, $\mu_2 = 0.25$, $\alpha_{\rm c} = 0.1$, $\tau_{\rm s} = 3$ h, $\tau_{\rm c} = 1$ h, $a_0 = 5$, $a_1 = a_2 = 0.5$, $\tau_{\rm conv} = 2$ h, and $\overline{\theta}_{\rm eb} - \theta_{\rm em} = 14$ K. The additional new parameters for the enhanced congestus parametrization include: the coefficients representing contributions of stratiform and congestus clouds to the first baroclinic heating, $\xi_{\rm s} = 0.5$ and $\xi_{\rm c} = 1.25$, respectively; the inverse convective buoyancy time-scale associated with congestus clouds, $a_0' = 2$; and the bulk convective heating $\overline{Q}$ that is determined at RCE. Interested readers should consult Khouider and Majda (2008a) for the details of the linear stability analysis.

Here, we are interested in the initiation of a convectively coupled wave train to mimic the high-resolution two-dimensional explicit cloud-resolving model solutions in Grabowski and Moncrieff (2001). In particular, we integrate the model with a localized piece of a single unstable linear wave of small amplitude centred at 20 000 km as the initial condition. (Figure 2 shows a space–time plot of the first two baroclinic velocities, potential temperatures, congestus and deep heating rates, moisture, and precipitation.) Note that this set-up is exactly the regime analyzed in Frenkel, *et al.* (2011), in which they focused on understanding the effect of the diurnal cycle, and we neglect the diurnal cycle here. Notice there are fast-moving waves (q, $H_{\rm d}$, and P in Figure 2) during the first 2 days moving away from the 20 000 km mark. After about 8–10 days, additional waves appear; this wave initiation is partly due to the convectively coupled wave interactions with faster-moving gravity waves. After about 100 days, these waves mature to a wave train of six individual eastward moving waves with a wave speed of approximately 14.5 m s^{-1} (Figure 3). Such wave structure and wave train organization resembles the structure found in the explicit simulations with the cloud-resolving model of Grabowski and Moncrieff (2001). Moreover, the mature waves have a total convective heating pattern (with backward

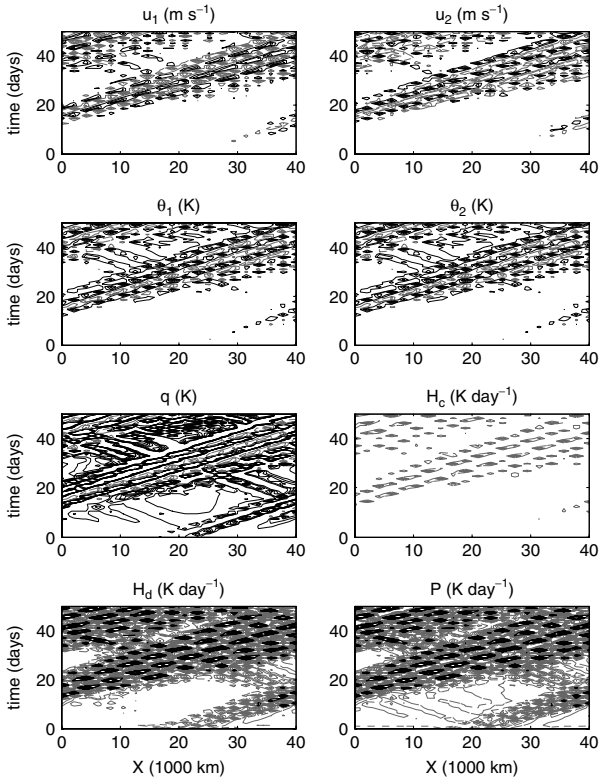

Figure 2. Initiation of a convectively coupled wave train. The space–time plot here is constructed with coarse spatial and temporal resolutions at every 2000 km and 24 h. (This coarse dataset is sampled from solutions with higher resolutions at every 40 km and 3 h.) The contour intervals are 0.25 m s^{-1} for the zonal wind, 0.025 K for the potential temperature and humidity, and 0.05 K day^{-1} for the heating rates and precipitation. For $u_1, u_2, \theta_1, \theta_2, q$, solid black (dashed grey) contours denote positive (negative) values. For $H_{\rm c}, H_{\rm d}, P$, solid black (dashed grey) contours denote heating rates greater (smaller) than 1 K day^{-1}.

Copyright © 2012 Royal Meteorological Society

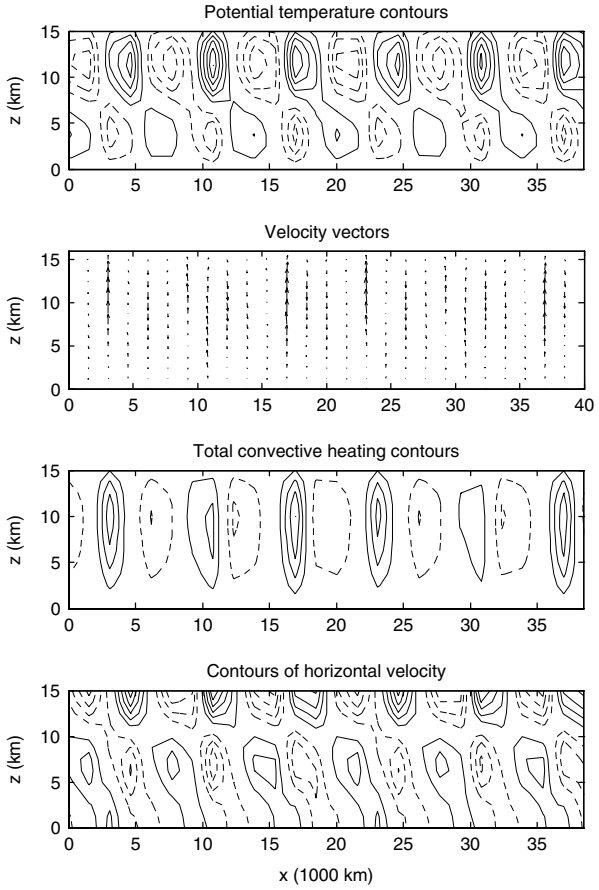

Figure 3. Moving average of the vertical structure in a reference frame of 14.5 m s^{-1} from a time period 500–1000 days. The contour intervals are 0.07 K for the potential temperature, 0.54 K day^{-1} for the total convective heating, and 0.35 m s^{-1} for the horizontal velocity. Solid (dashed) contours denote positive (negative) values.

and upward tilt in the wind and temperature fields, upper-tropospheric positive temperature anomalies slightly leading the region of the upward motion, which is in phase with the heating anomalies, with low-level convergence) which is very similar to convectively coupled Kelvin waves observed in nature (Wheeler and Kiladis, 1999; Wheeler, *et al.*, 2000; Straub and Kiladis, 2002).

4. Algorithms for filtering moisture-coupled waves from sparse observations

In this section, we first describe the sparse observation networks and then discuss in detail the reduced stochastic filtering algorithms.

4.1. Sparse observation networks

In the present article, we consider horizontally sparse observations at every 2 000 km. This means we only have $M = 20$ observations at $x_j = jh$, $h = 2\pi/40\,000$ km in a non-dimensionalized unit assuming that the equatorial belt circumference is 40 000 km. For compact notation, we define $\Psi_{j,m} = (u_1, u_2, \theta_1, \theta_2, \theta_{\rm eb}, q, H_{\rm s}, H_{\rm c})^{\rm T}$; we use subscripts j and m to specify that each component in Ψ is evaluated at grid point x_j and discrete time t_m, respectively. We define a general observation model

$$\mathbf{G}\Psi^{\rm o}_{j,m} = \mathbf{G}\Psi_{j,m} + \mathbf{G}\sigma_{j,m}, \quad \sigma_{j,m} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}^{\rm o}), \tag{20}$$

where $\mathbf{G}$ is an observation operator that maps the model state to the observation state space and $\sigma_{j,m}$ are eight-dimensional independent Gaussian white noises with mean zero and diagonal covariance matrix $\mathbf{R}^{\rm o}$. Vertically, we consider four observation networks with specific $\mathbf{G}$:

SO (Surface observations). Here, we consider observing the wind, potential temperature at a surface height $z_{\rm s} = 100$ m, and the equivalent boundary-layer potential temperature $\theta_{\rm eb}$. The corresponding observation operator is a 3×8 matrix $\mathbf{G}$ with non-zero components

$$\left.\begin{aligned} &\mathbf{G}_{1,1} = G(z_s), \quad \mathbf{G}_{1,2} = G(2z_s),\\ &\mathbf{G}_{2,3} = G'(z_s), \quad \mathbf{G}_{2,4} = 2G'(2z_s),\\ &\mathbf{G}_{3,5} = 1, \end{aligned}\right\} \tag{21}$$

where G, G' are the vertical baroclinic profiles defined in section 2.

SO+MT (Surface observations + middle troposphere temperature). This observation network includes temperature at middle-troposphere height $z_{\rm m} = 8$ km in addition to SO. The corresponding observation operator is a 4×8 matrix $\mathbf{G}$ with non-zero components

$$\mathbf{G}_{4,3} = G'(z_{\rm m}), \quad \mathbf{G}_{4,4} = 2G'(2z_{\rm m}), \tag{22}$$

in addition to (21).

SO+MTV (Surface observations + middle troposphere temperature and velocity). This observation network includes velocity at middle-troposphere height $z_{\rm m} = 8$ km in addition to SO+MT. The corresponding observation operator is a 5×8 matrix $\mathbf{G}$ with non-zero components

$$\mathbf{G}_{5,1} = G(z_{\rm m}), \quad \mathbf{G}_{5,2} = 2G(2z_{\rm m}), \tag{23}$$

in addition to (21) and (22).

CO (Complete observations). This vertically complete observation network is defined with $\mathbf{G} = \mathbf{I}$ for diagnostic purposes.

4.2. Filtering algorithms

In this article, we consider the simplest version of our reduced stochastic filters, the Mean Stochastic Model (MSM; Harlim and Majda, 2008a, 2010a,b; Majda and Harlim, 2012). The new feature in the present context is that we have multiple variables Ψ_j as opposed to a scalar field, and therefore we need to design the MSM in an appropriate coordinate expansion to avoid parametrizing various coupling terms.

As in Harlim and Majda (2008a), our design of the filter prior model is based on the standard approach for modelling turbulent fluctuations (Majda, *et al.*, 1999, 2008; Majda and Timofeyev, 2004; DelSole, 2004), i.e. we introduce model errors through linearizing the nonlinear models about a frozen constant state and replacing the truncated nonlinearity with a dissipation and spatially correlated noise (white in time) to mimic rapid energy transfer between different scales. In the present context, we consider the linearized multicloud model about the RCE,

$$\frac{\partial \Psi'}{\partial t} = \mathcal{P}(\partial_x)\Psi', \tag{24}$$

Copyright © 2012 Royal Meteorological Society

where Ψ' denotes the perturbation field about the RCE and $\mathcal{P}$ denotes the linearized differential operator of the multicloud model at RCE. A comprehensive study of the linear stability analysis of (24) involves solving eigenvalues of an 8×8 dispersion matrix, $\omega(k)$, and was reported in Majda, *et al.* (2007) for the MJO-like wave and in Khouider and Majda (2008a) and Frenkel, *et al.* (2011) for the multicloud model with enhanced congestus heating.

Consider a numerical discretization for (24) with spatial mesh size of $\Delta x = 2000$ km such that the model state space is essentially similar to the observation state space. With this approximation, the PDE in (24) becomes

$$\frac{\mathrm{d}\widehat{\Psi}_k}{\mathrm{d}t} = \mathrm{i}\omega(k)\widehat{\Psi}_k, \qquad |k| \le M/2 = 10, \tag{25}$$

where $\{\widehat{\Psi}_k\}_{|k|\le M/2}$ are the discrete Fourier components of $\{\Psi'_j\}_{j=1,\ldots,M}$. Now consider an eigenvalue decomposition, $\mathrm{i}\omega(k)\mathbf{Z}_k = \mathbf{Z}_k\Lambda_k$, where Λ_k is a diagonal matrix of the eigenvalues and $\mathbf{Z}_k$ is a matrix whose columns are the corresponding eigenvectors. Then we can write (25) as a diagonal system,

$$\frac{\mathrm{d}\widehat{\Phi}_k}{\mathrm{d}t} = \Lambda_k\widehat{\Phi}_k, \qquad |k| \le M/2 = 10, \tag{26}$$

with the transformation

$$\widehat{\Phi}_k = \mathbf{Z}_k^{-1}\widehat{\Psi}_k. \tag{27}$$

4.2.1. *The MSM filter*

The MSM is defined through the stochastic differential system,

$$\mathrm{d}\widehat{\Phi}_k = \Big[(-\Gamma_k + \mathrm{i}\Omega_k)\widehat{\Phi}_k + \mathbf{f}_k\Big]\mathrm{d}t + \Sigma_k\,\mathrm{d}W_k, \tag{28}$$

for $|k| \le M/2$. In (28), Γ_k, Ω_k, and Σ_k are diagonal matrices with diagonal components obtained through regression fitting to the climatological statistics, while the forcing term is proportional to the climatological mean field, $\mathbf{f}_k = (\Gamma_k - \mathrm{i}\Omega_k)\langle\widehat{\Phi}_k\rangle$; here, the angle bracket $\langle\cdot\rangle$ denotes an average. Notice that the realizability of this stochastic model (referred to as MSM-1; Majda, *et al.*, 2010; Harlim and Majda, 2010b; Majda and Harlim, 2012) is guaranteed since Γ_k is always positive definite as opposed to the alternative approach which sets $\Omega_k = -\mathrm{i}\Lambda_k$ (Penland, 1989; DelSole, 2000). Throughout this article, the climatological statistics are computed from solutions of the full multicloud model resolved at 40 km grid points with different temporal resolutions for the two cases: the MJO-like travelling wave and the initiation of a convectively coupled wave train (sections 5.1 and 5.2).

The discrete-time Kalman filtering problem with the MSM as the prior model is defined for each horizontal wavenumber k as

$$\widehat{\Psi}_{k,m} = \mathcal{F}_k(\Delta t)\widehat{\Psi}_{k,m-1} + \mathbf{g}_{k,m} + \eta_{k,m}, \tag{29}$$

$$\mathbf{G}\widehat{\Psi}^o_{k,m} = \mathbf{G}\widehat{\Psi}_{k,m} + \mathbf{G}\widehat{\sigma}_{k,m}, \tag{30}$$

where the observation model in (30) is the discrete Fourier component of the canonical observation model in (20) with Gaussian noises, $\widehat{\sigma}_{k,m} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}^{\mathrm{o}}/M)$. The discrete filter model in (29) has coefficients

$$\mathcal{F}_k(t) = \mathbf{Z}_k \exp\{(-\Gamma_k + \mathrm{i}\Omega_k)t\}\,\mathbf{Z}_k^{-1}, \tag{31}$$

$$\mathbf{g}_{k,m} = -\{\mathbf{I} - \mathcal{F}_k(t_m)\}(-\Gamma_k + \mathrm{i}\Omega_k)^{-1}\mathbf{f}_k, \tag{32}$$

and unbiased Gaussian noises $\eta_{k,m}$ with covariance matrix

$$\mathbf{Q}_k = \frac{1}{2}\mathbf{Z}_k\Sigma_k^2\Gamma_k^{-1}\left(\mathbf{I} - |\mathcal{F}_k(\Delta t)|^2\right)\mathbf{Z}_k^*. \tag{33}$$

These coefficients are obtained by evaluating the analytical solutions of the stochastic differential system in (28) at observation time interval $\Delta t = t_{m+1} - t_m$ and applying the transformation in (27).

The MSM filter in (29)–(30) is computationally very cheap since it only involves $M/2 + 1$ independent 8×8 Kalman filtering problems, ignoring cross-correlations between different horizontal wavenumbers. Such a diagonal approximation may seem to be counterintuitive since it generates severe model errors, but we have shown that it provides high filtering skill beyond the perfect model simulations in various contexts including the regularly spaced sparse observations (Harlim and Majda, 2008b), irregularly spaced sparse observations (Harlim, 2011), strongly chaotic nonlinear dynamical systems (Harlim and Majda, 2008a, 2010a), and midlatitude baroclinic wave dynamics (Harlim and Majda, 2010b).

Applying the Kalman filter formula on each wavenumber in (29)–(30) provides the following background (or prior) mean and error covariance estimates,

$$\widehat{\Psi}^{\mathrm{b}}_{k,m} = \mathcal{F}_k(\Delta t)\widehat{\Psi}^{\mathrm{a}}_{k,m-1} + \mathbf{g}_{k,m}, \tag{34}$$

$$\mathbf{R}^{\mathrm{b}}_{k,m} = \mathcal{F}_k(\Delta t)\mathbf{R}^{\mathrm{a}}_{k,m-1}\mathcal{F}_k(\Delta t)^* + \mathbf{Q}_k, \tag{35}$$

and analysis (or posterior) mean and error covariance estimates

$$\left.\begin{aligned}
\widehat{\Psi}^{\mathrm{a}}_{k,m} &= \widehat{\Psi}^{\mathrm{b}}_{k,m} + \mathbf{K}_{k.m}(\mathbf{G}\widehat{\Psi}^{\mathrm{o}}_{k,m} - \mathbf{G}\widehat{\Psi}^{\mathrm{b}}_{k,m}),\\
\mathbf{R}^{\mathrm{a}}_{k,m} &= (\mathbf{I} - \mathbf{K}_{k.m}\mathbf{G})\mathbf{R}^{\mathrm{b}}_{k,m},\\
\mathbf{K}_{k.m} &= \mathbf{R}^{\mathrm{b}}_{k,m}\mathbf{G}^*\{\mathbf{G}(\mathbf{R}^{\mathrm{b}}_{k,m} + \mathbf{R}^{\mathrm{o}}/M)\mathbf{G}^*\}^{-1},
\end{aligned}\right\} \tag{36}$$

where $\mathbf{K}_{k.m}$ is the Kalman gain matrix.

4.2.2. *The complete 3D-Var*

For diagnostic purposes, we also consider a 3D-Var version in the MSM framework above. That is, we simply set the background-error covariance matrix to be independent of time,

$$\mathbf{B}_k \equiv \lim_{\Delta t\to\infty}\mathbf{R}^{\mathrm{b}}_{k,m} = \frac{1}{2}\mathbf{Z}_k\Sigma_k^2\Gamma_k^{-1}\mathbf{Z}_k^*, \tag{37}$$

and repeat the mean prior and posterior updates in (34), (36) with a constant Kalman gain matrix,

$$\mathbf{K}_k = \mathbf{B}_k\mathbf{G}^*\{\mathbf{G}(\mathbf{B}_k + \mathbf{R}^{\mathrm{o}}/M)\mathbf{G}^*\}^{-1}.$$

We called this approach the complete 3D-Var because the forward model parameters in (31), (32), (33) and the background covariance matrix in (37) are determined from complete solutions of the multicloud model in (1), including the moisture and heating variables from (2), (7), (8). This formulation is significantly different from an earlier approach with variational techniques (Žagar, *et al.*, 2004b,a) in which the background covariance matrix is parametrized in an eigenmode basis constructed from the dry equatorial waveguide.

Copyright © 2012 Royal Meteorological Society

4.2.3. *The 'dry and cold' 3D-Var*

To mimic the approach in Žagar, *et al.* (2004b,a), we consider only using the wind and temperature data, $u_1, u_2, \theta_1, \theta_2, \theta_{\rm eb}$, to construct the 'dry and cold' eigenmode basis and background covariance matrix $\mathbf{B}_k$. Technically, we still use the MSM model in (28), but replace the transformation in (27) by

$$\widehat{\Phi}_k^{\rm dc} = \mathbf{Z}_k^{-1} \begin{bmatrix} \mathbf{I}_{5\times 5} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \widehat{\Psi}_k. \tag{38}$$

In this sense, the parameters Γ_k, Ω_k, and Σ_k in (28) are fitted to climatological statistics of $\widehat{\Phi}_k^{\rm dc}$ based on only the wind and temperature variables. Repeating the 3D-Var algorithm described above in this set-up provides an honest 'dry and cold' version analogous to the earlier approach in Žagar, *et al.* (2004b,a).

Besides the eigenmode basis difference, we should note that the 'dry and cold' 3D-Var here is computationally much cheaper than that in Žagar, *et al.* (2004b,a) since we perform both the prior and posterior updates in the diagonalized Fourier basis with reduced stochastic filters through (34)–(36) as opposed to their approach that propagates the nonlinear dry shallow-water equations in physical space and applies the analysis step in the spectral diagonal basis. On each data assimilation step, their approach requires back-and-forth transformations in between the physical and spectral spaces with a rotational transformation matrix that is quite often ill-conditioned, as reported in Žagar, *et al.* (2004b).

For diagnostic purposes, we will also consider the 'moist and cold' 3D-Var in the numerical simulations in section 5.1; this model is constructed exactly like the 'dry and cold' model described above with moisture q in addition to the wind and temperatures, $u_1, u_2, \theta_1, \theta_2, \theta_{\rm eb}$.

5. Filtering skill for moisture-coupled tropical waves

In this section, we report the numerical results of implementing the filtering algorithms in section 4.2 to assimilate the synthetic sparse observation networks defined in section 4.1 on the two examples discussed in section 3.

In the numerical simulations below, we consider the precise observations case with $\mathbf{R}^{\rm o} = \mathbf{0}$ and small observation noises with positive definite covariance matrix $\mathbf{R}^{\rm o} > \mathbf{0}$. In the non-zero noise case, we choose the observation noise variance to be 10% of the climatological variance of each variable. In this sense, the noise variances are less than both the peak of the energy spectrum and the smallest average signal amplitude.

5.1. *MJO-like turbulent travelling wave*

Our goal here is to check the filtering skill in recovering the structure of the MJO-like travelling wave (section 3.1) with the MSM forward model in (28) with parameters (31)–(33), which are specified from a time series of 8 000 days with temporal resolution of 6 h at the climatological state.

First, we compare the moving average of $u_1, u_2, \theta_1, \theta_2, \theta_{\rm eb}$, q, H_s, H_c, P obtained from the true solutions of the test model in section 3.1 and the posterior mean estimates in (36). The moving average is taken in a reference frame at 6.1 m s^{-1} from a time period of 750–1000 days. In Figures 4–8, we show the moving average from assimilations with observation time interval of 24 h for complete observations (CO) with $\mathbf{R}^{\rm o} = \mathbf{0}$, and for all observation networks discussed in section 4.1, CO, SO+MTV, SO+MT, SO with small observation noise covariance, $\mathbf{R}^{\rm o} > \mathbf{0}$. For observation network CO without observation errors, $\mathbf{R}^{\rm o} = \mathbf{0}$ (Figure 4), the three schemes (MSM filter, complete 3D-Var and 'dry and cold' 3D-Var) are identical and they perfectly recover the averaged MJO structure except for slight overestimation on the stratiform heating and precipitation.

In the presence of observation noise, we include results with 'moist and cold' 3D-Var (detailed discussion appears at the end of section 4.2.3). We find that all the four schemes are able to recover u_1 and $\theta_{\rm eb}$ with any observation network. When middle-troposphere wind observation is absent (SO+MT and SO in Figures 7, 8), the estimate for u_1 slightly degrades but is completely wrong for u_2. The MSM filter overestimates θ_2 by roughly 0.1 K, even with surface and middle-troposphere potential temperature observations; we find that this poor estimation is attributed to an inaccurate mean estimate (on the zeroth horizontal mode) of θ_2. The MSM filter, the complete and 'moist and cold' 3D-Var are able to recover the oscillating structure of the moisture q with any observation network (with slight errors for the MSM filter with SO) reflecting the active and suppressed convective phases of the MJO-like wave. On the other hand, the 'dry and cold' 3D-Var cannot produce q accurately even with observation network CO, and simply predicts a dry atmosphere (with zero moisture profile) when the moisture is unobserved. All the four filters are not able to reproduce the stratiform and congestus heating profiles when they are not observed.

Except for the SO network, both the complete and 'moist and cold' 3D-Var are able to reasonably recover

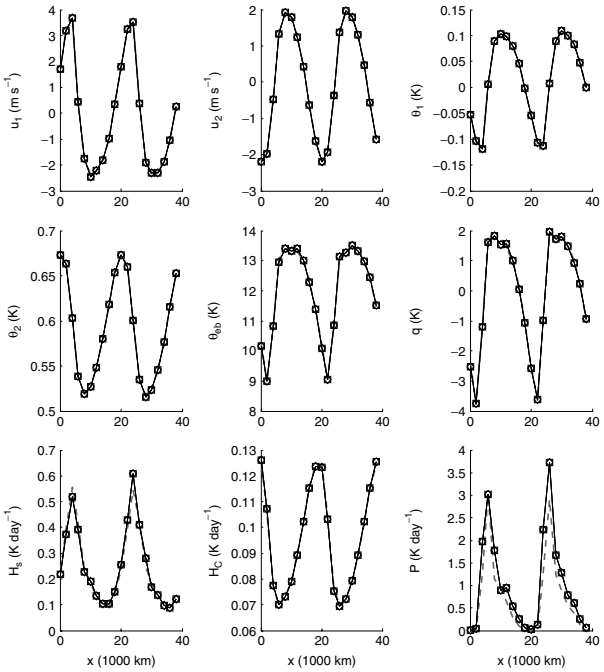

Figure 4. MJO-like waves with $\Delta t = 24$ h, $\mathbf{R}^{\rm o} = \mathbf{0}$ and complete observations (CO). The moving average is in a reference frame at 6.1 m s^{-1} of the model variables. True (grey dashes), posterior mean state of the complete 3D-Var (circles), MSM filter (squares), and the 'dry and cold' 3D-Var (diamonds).

Copyright © 2012 Royal Meteorological Society

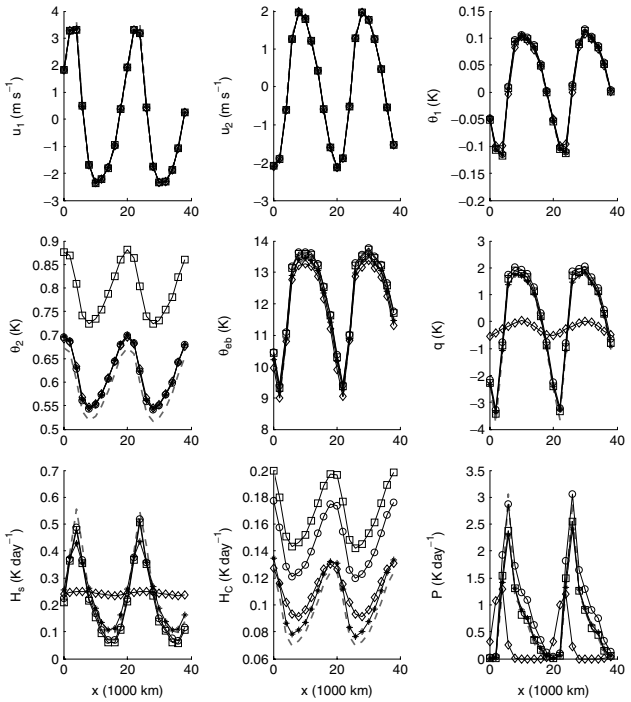

Figure 5. MJO-like waves with $\Delta t = 24\,\mathrm{h}$, $\mathbf{R}^{\mathrm{o}} > \mathbf{0}$ and complete observations (CO). The moving average is in a reference frame at $6.1\,\mathrm{m\,s^{-1}}$ of the model variables. True (grey dashes), posterior mean state of the complete 3D-Var (circles), MSM filter (squares), the 'dry and cold' 3D-Var (diamonds), and the 'moist and cold 3D-Var (asterisks).

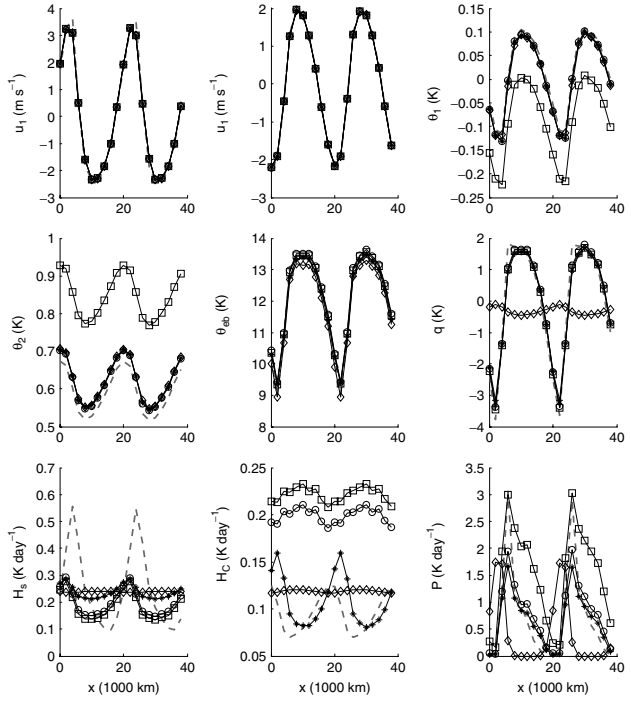

Figure 6. As Figure 5, but for surface observations plus middle-troposphere potential temperature and velocity (SO+MTV).

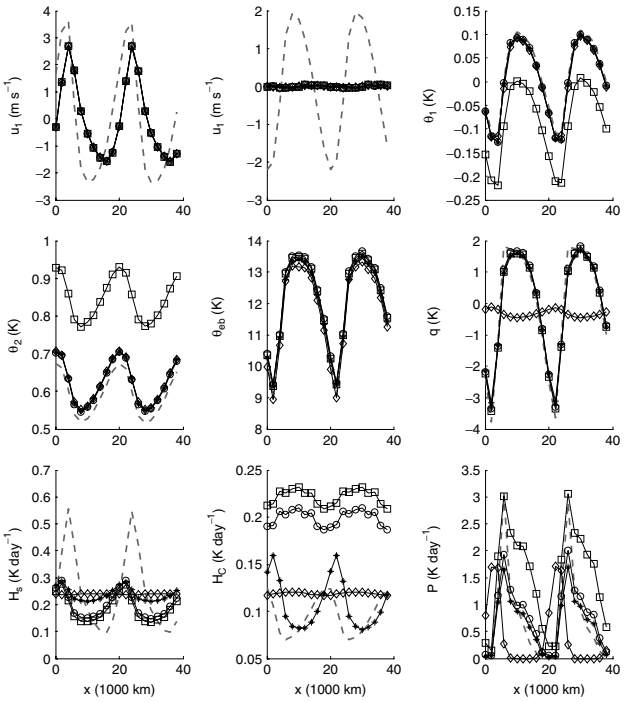

Figure 7. As Figure 5, but for surface observations plus middle-troposphere potential temperature (SO+MT).

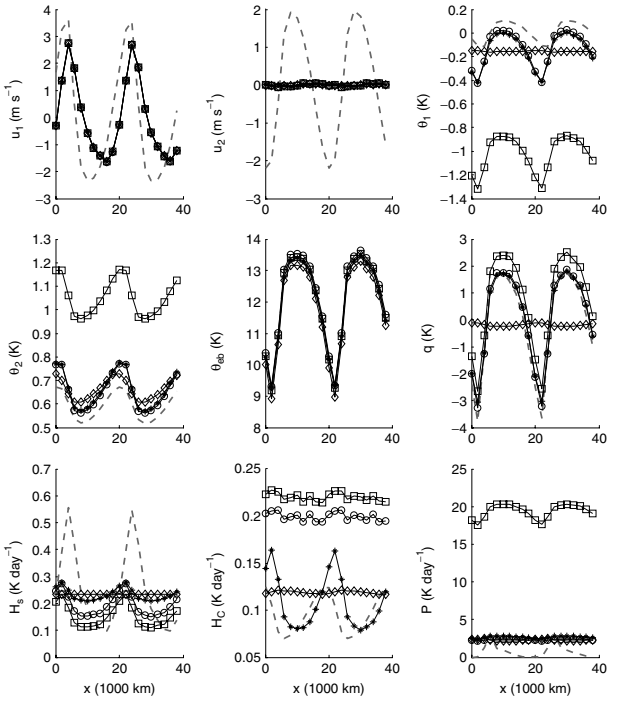

Figure 8. As Figure 5, but for surface observations (SO).

the precipitation rate, P, which in this model is exactly the deep convection heating rate; here, the 'cold and dry' 3D-Var precipitation estimate is very inaccurate (Figures 5–7). On the other hand, the MSM filter captures the peak of the precipitation on all the three observation networks (CO, SO+MTV and SO+MT), but overestimates the profile on the last two observation networks. This overestimation of the precipitation (as well as those observed when we assimilate only the SO network; Figure 8) can be explained as follows. From the precipitation budget in (9), it is obvious that the contributions of θ_{eb}, q, and θ_2 to the convective parametrization are small (with scale factors $a_1 = 0.1, a_2 = 0.5$, $a_0\gamma_2 = 1.2$, respectively) relative to θ_1 (with scale factor $a_0 = 12$). Therefore, the wet filtered state (with large precipitation estimates as seen in Figure 8) is attributed to the slight underestimation of the first

Copyright © 2012 Royal Meteorological Society

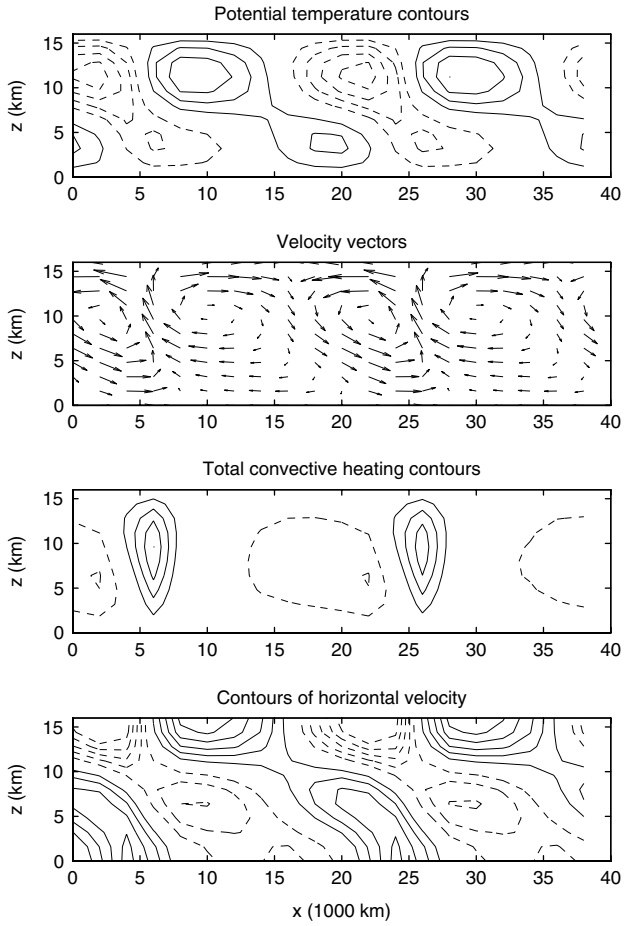

Figure 9. The true vertical profile of the MJO-like waves computed with moving average in a reference frame at 6.1 m s^{-1}. The contour intervals are 0.07 K for the potential temperature, 0.29 K day^{-1} for the total convective heating, and 1 m s^{-1} for the horizontal velocity. Solid (dashed) contours denote positive (negative) values.

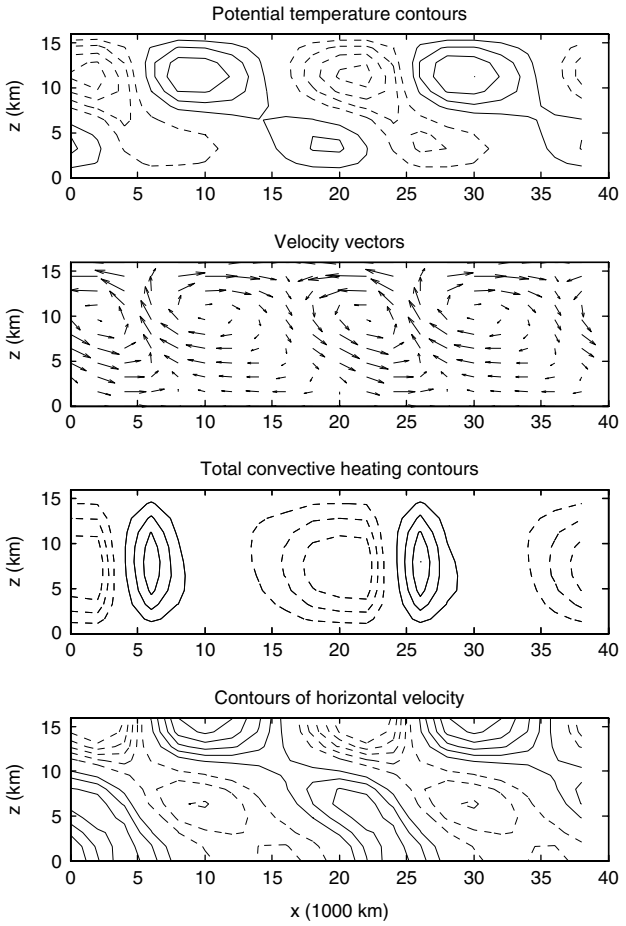

Figure 10. The vertical profile from complete 3D-Var estimate with observation network SO+MTV and $\mathbf{R}^{o} > \mathbf{0}$, and $\Delta t = 24$ h. Contours are as in Figure 9.

baroclinic potential temperature, θ_1. The complete 3D-Var underestimates θ_1 by as much as 0.5 K; this yields a spatially uniform precipitation rate of about 2.3 K day^{-1}. The MSM filter underestimates θ_1 by as much as 1.5 K and its corresponding precipitation estimate is about 20 K day^{-1}.

In Figures 9–12, we show the detailed vertical structure of the total potential temperature Θ, the velocity vector field ($\mathbf{V}$, w), the total convective heating, and horizontal velocity from the MJO-like wave in section 3.1 and the complete 3D-Var estimates with observation networks SO+MTV, SO+MT, and SO, respectively. In particular, the vertical tilted structure in the potential temperature is recovered with any of these three observation networks; similar recovery (not shown) is also obtained with the MSM filter and the 'moist and cold' 3D-Var; the 'dry and cold' 3D-Var also recovers this tilted structure except with observation network SO. On the other hand, the tilted structure in the horizontal velocity with low-level convergence that is in phase with the deep convective heating is not recovered whenever the middle-troposphere wind observation is absent. Notice also that the deep convective heating is recovered except with observation network SO; similar recovery (not shown) is also attained with the MSM filter and the 'moist and cold' 3D-Var but not with the 'dry and cold' 3D-Var.

We also find that both the complete and 'moist and cold' 3D-Var are able to reconstruct the detailed precipitation structure in Figure 1 except when assimilated with observation network SO (results are not shown). The MSM filter is able to capture the peak but overestimates the detail profile. The 'moist and cold' 3D-Var reproduces the eastward MJO-like signal, but fails to capture the westward intermittent moist fluctuations within the MJO envelope as shown in Figure 1.

We also repeated the numerical experiments above with different observation time intervals ranging from 6 h to 8 days with the complete 3D-Var and MSM filter (Figure 13 shows the average RMS errors on the MSM filter case). Particularly noteworthy is that the posterior estimates have roughly similar RMS errors for the observed variables independent of the observation times; for the unobserved variables, the RMS errors for the shorter observation times are larger than those for the longer observation times! This latter result can be understood as follows. The dynamical operator $\mathcal{F}_k$ in (31) is essentially marginally stable (with largest eigenvalue 0.9899) for $\Delta t = 6$ h and is strictly stable (with largest eigenvalue 0.8836) for longer $\Delta t = 72$ h. The observability condition, which is a necessary condition for accurate filtered solutions when the dynamical operator is marginally stable (Anderson and Moore, 1979; Majda and Harlim, 2012), is practically violated here; our test with SO+MT observation network suggests that the observability matrix is ill-conditioned with $\det\left([\mathbf{G}^{\mathrm{T}}\,(\mathbf{G}\mathcal{F}_k)^{\mathrm{T}}]\right) \approx 10^{-20}$. This explains why the longer observation times produce more accurate filtered solutions. Thus, with the crude spatial observation network and the inefficient behaviour of

Copyright © 2012 Royal Meteorological Society

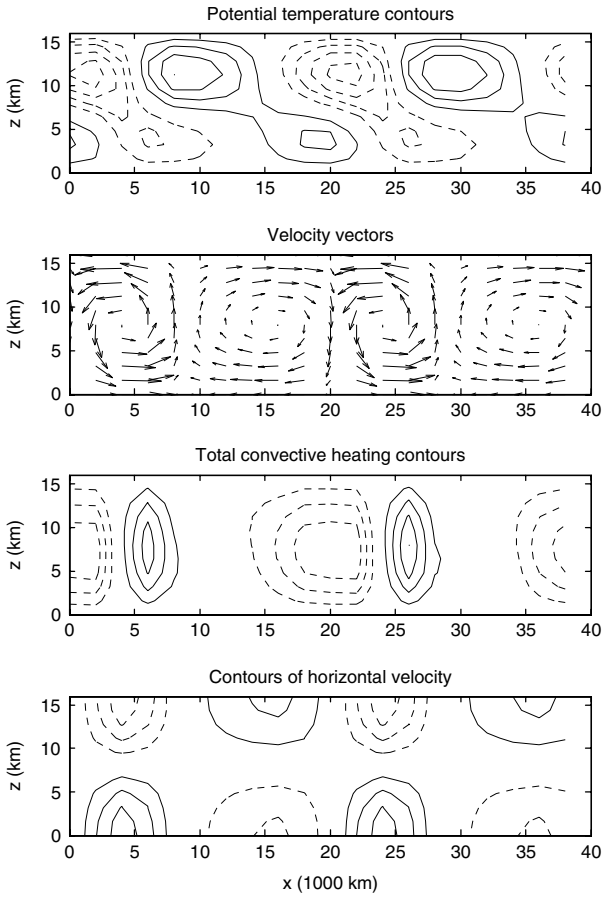

Figure 11. As Figure 10, but for observation network SO+MT.

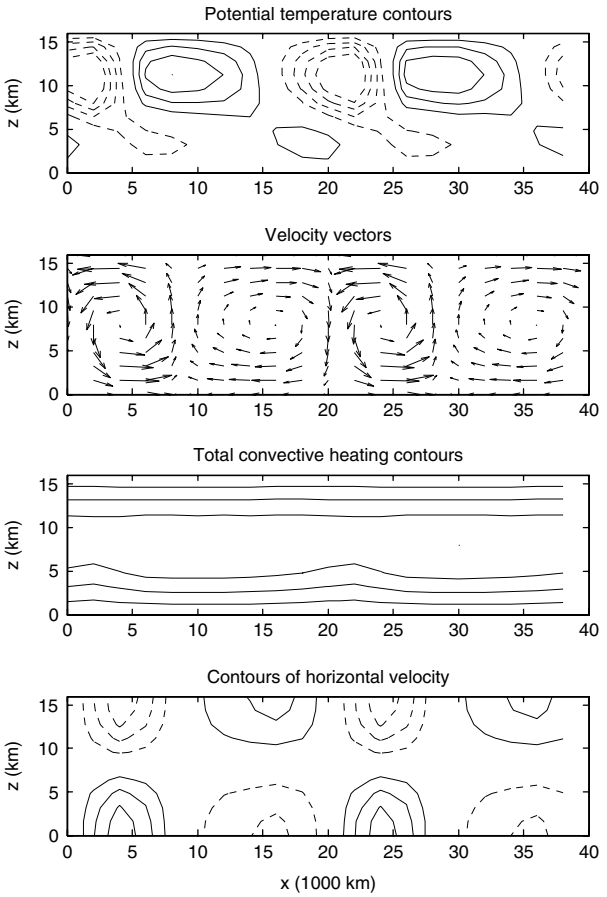

Figure 12. As Figure 10, but for observation network SO.

MSM at short times, this simple filtering strategy necessarily cannot capture subgrid-scale features of the wave with high skill; by design this is also true for 3D-Var. We encounter similar behaviour of filtered solutions in the next example in section 5.2.

In Figure 13, we include the climatological errors (dash-dotted line) and observation errors (thin dashes) for diagnostic purposes. Recall that the observation-error covariance R^{o} in our experiments is 10% of the climatological variances and the observation errors are only relevant for diagnostic purposes when the corresponding variable is observed. So, in real time, the MSM filter with sparse observation networks SO+MTV, SO+MT has reasonable skill as long as its RMS errors are below the climatological errors. In this sense, we observe that the MSM filter is very skilful for variables $u_1, \theta_{\mathrm{eb}}$ and q for any observation network as well as for θ_1 for observation networks other than SO. Our conjecture is that, on these variables, the RMS errors will increase as the observation time interval is near its slowest decaying time (70 days for this model). For the other variables, the filtering skill is not better than the climatological variability and further improvement will be addressed in future work.

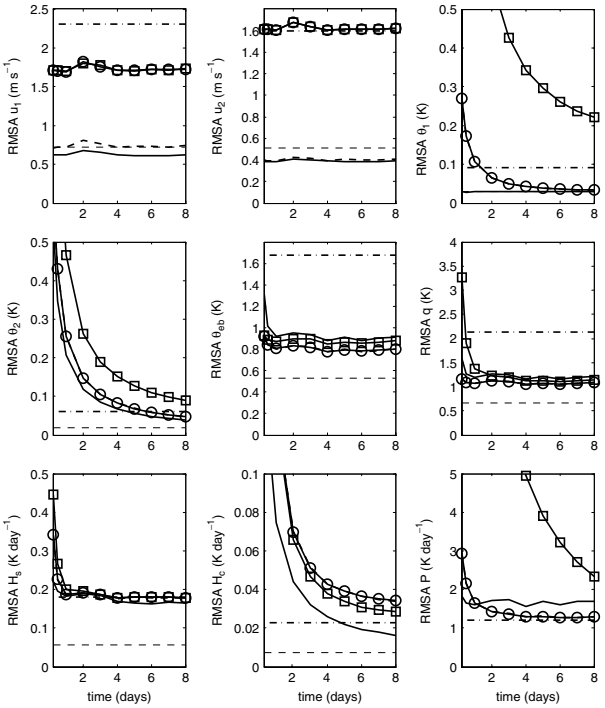

Figure 13. Average RMS errors as functions of observation time interval (in days): observation error (thin dashed line), climatological errors (dash-dotted line), CO (bold solid line), SO+MTV (bold dashes), SO+MT (circles) and SO(squares).

5.2. *Initiation of a convectively coupled wave train*

Here, our goal is to check the filtering skill in recovering the transient behaviour of initiation of a convectively coupled wave train (section 3.2) with the MSM forward model in (28) where parameters (31)–(33) are specified from a time series at the climatological state for the period of time 500–1000 days with temporal resolution of 3 h.

Copyright © 2012 Royal Meteorological Society

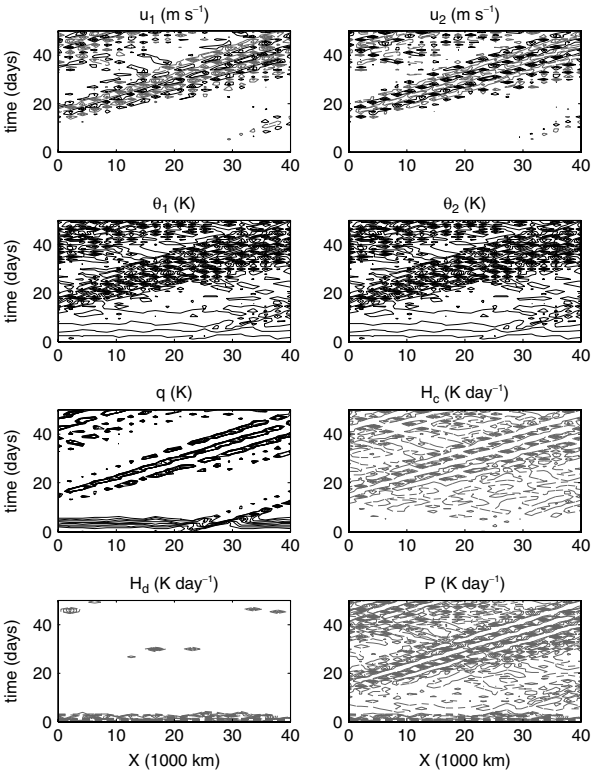

Figure 14. Space–time plot from the complete 3D-Var estimate with observation network CO and $\mathbf{R}^o > \mathbf{0}$, and $\Delta t = 24$ h. The contour intervals are 0.25 m s^{-1} for the zonal wind, 0.025 K for the potential temperature and humidity, and 0.05 K day^{-1} for the heating rates and precipitation. For $u_1, u_2, \theta_1, \theta_2, q$, solid black (dashed grey) contours denote positive (negative) values, and for H_c, H_d, P solid black (dashed grey) contours denote heating rates greater (smaller) than 1 K day^{-1}.

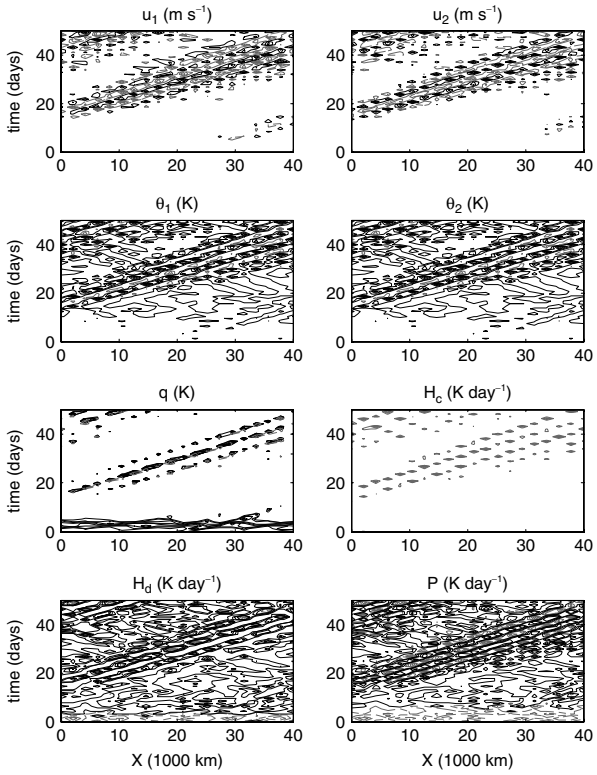

Figure 15. As Figure 14, but for observation network SO+MTV, and with contour intervals 0.05 K day^{-1} for the congestus heating rate, and 0.25 K day^{-1} for the deep convective heating and precipitation.

In Figures 14–17, we report the space–time plot of the filtered estimates at the initial period of time 0–50 days from the complete 3D-Var with observation time $\Delta t = 24$ h, observation noise variance $\mathbf{R}^o > \mathbf{0}$, and observation networks SOMTV, SOMT, and SO. By eye, we can see that the emerging pattern in Figure 2 is recovered for all variables except for the deep convection heating rate with the complete observation network! This poor estimate is attributed to an overestimation of θ_1 (which sets the available convective heating Q_d in (15) to zero). In this case, the precipitation budget in the filtered solution is dominated by the stratiform and congestus heating rates. On the other hand, even if the pattern of H_d is always captured with networks SO+MTV, SO+MT, SO, its accuracy is questionable as we will see below.

To be more precise, we quantify the filter skill with the average RMS error and pattern correlation (PC) between the posterior mean estimate and the truth at the initiation period of time 0–75 days before these waves lock into a wave train of six waves as shown in Figure 3. In Figures 18–25, we plot these two performance measures as functions of observation times for observation networks CO, SO+MTV, SO+MT, and SO. In each panel, we compare four numerical experiments including the MSM filter with $\mathbf{R}^o = \mathbf{0}$ and $\mathbf{R}^o > \mathbf{0}$, and the complete 3D-Var with $\mathbf{R}^o = \mathbf{0}$ and $\mathbf{R}^o > \mathbf{0}$.

From the average RMS errors (Figures 18, 20, 22, 24), we find that the filtering skill of the MSM filter and the complete 3D-Var are not different at all except for the wind variables when the middle-troposphere wind is not observed and $\mathbf{R}^o > \mathbf{0}$; there, the RMS errors of the complete

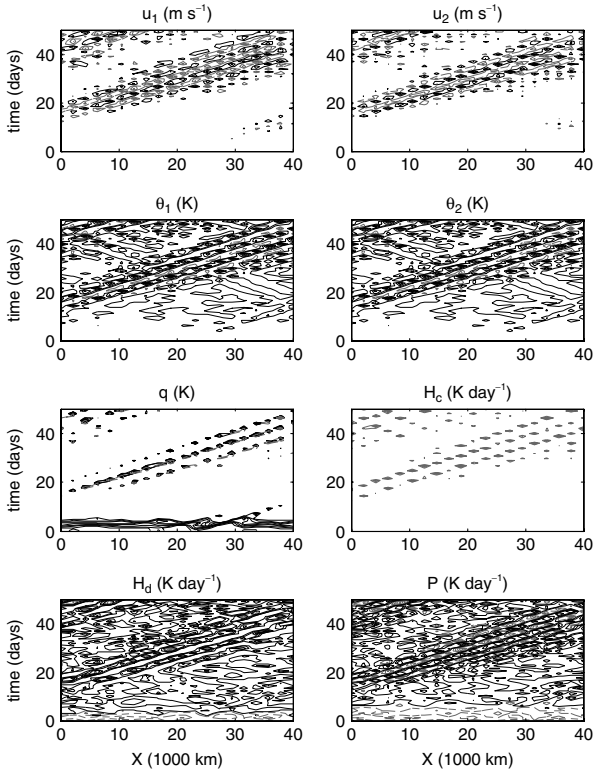

Figure 16. As Figure 15, but for observation network SO+MT.

3D-Var are smaller than those of the MSM filter (Figures 22, 24), but their PCs are identical (Figures 23, 25). When observations are complete (CO) and $\mathbf{R}^o = \mathbf{0}$, both the MSM

Copyright © 2012 Royal Meteorological Society

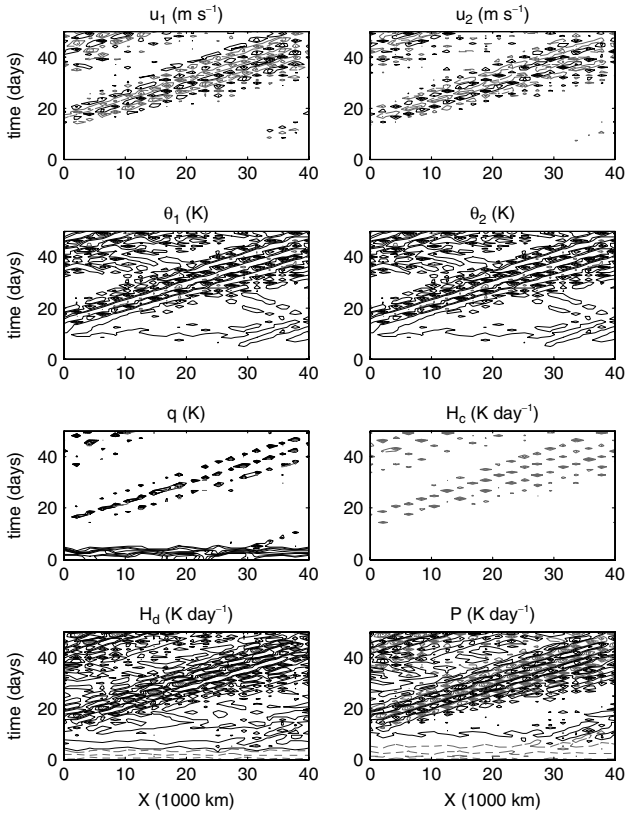

Figure 17. As Figure 15, but for observation network SO.

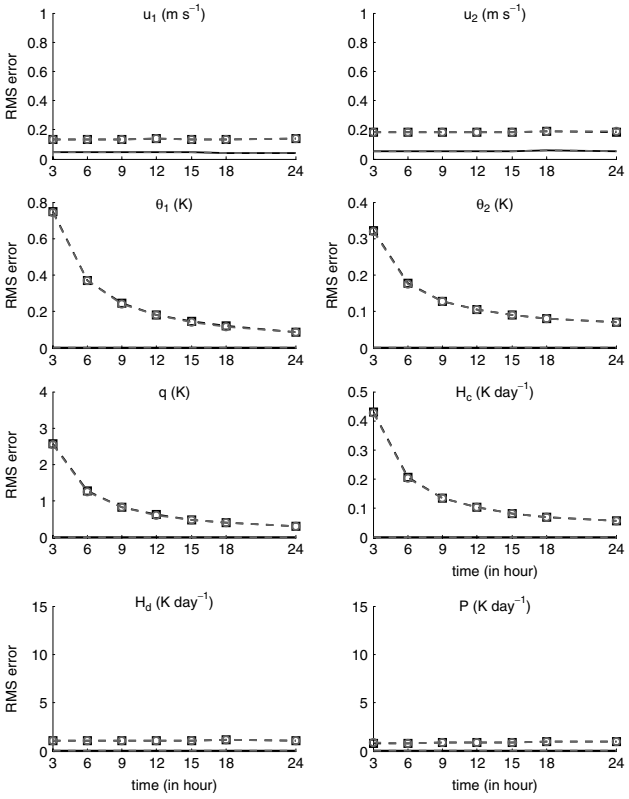

Figure 18. RMS errors as functions of observation time interval for observation network CO: MSM filter (grey), complete 3D-Var (black), $\mathbf{R}^{o} = \mathbf{0}$ (solid line and dashes with no markers), $\mathbf{R}^{o} > \mathbf{0}$ (dashes with square/circle markers).

filter and complete 3D-Var are able to recover the truth (with negligible errors and PC one).

The average RMS errors for simulations with $\mathbf{R}^{o} > \mathbf{0}$ (dashes with markers) for variables $\theta_1, \theta_2, q, H_c$ decay as functions of observation time even with complete observation network (Figure 18). We find that the larger errors with shorter observation times here are attributed to the violation of practical controllability (Anderson and Moore, 1979; Harlim and Majda, 2008b; Majda and Harlim, 2012) which is also a necessary condition for optimal filtering when the system is marginally stable (here $\mathcal{F}_k$ has a maximum eigenvalue close to 1). Additionally, we observe a similar decaying pattern for the error as function of observation time with SO+MTV, SO+MT, and SO for the unobserved variables q, H_c, H_d, and P even when the observed wind and potential temperatures have no errors ($\mathbf{R}^{o} = \mathbf{0}$). Here, the larger errors in the unobserved variables for shorter observation times are attributed to the violation of practical observability as explained in section 5.1.

When $\mathbf{R}^{o} > \mathbf{0}$, the RMS errors of the deep convection heating rate H_d are roughly $1\,\mathrm{K\,day^{-1}}$ (Figure 18) with observation network CO but the PCs are roughly zero (the PC curves are below 0.5 in Figure 19). The PC confirms the inability to recover H_d as shown in Figure 14. Both filtering schemes with the other observation networks (SO+MTV, SO+MT, SO) recover the structure of H_d (with PC of roughly 0.6 from Figures 21, 23, 25) but their errors are very large (as much as 10 K from Figures 20, 22, 24). The failure even to capture the deep convection pattern with CO is attributed to overestimation of θ_1 as explained earlier by contrasting the detailed space–time structure of θ_1 in Figures 2 and 14. The average RMS errors show such a tendency for failure with larger error with CO compared to those with the other networks, but they do not inform us whether the potential temperature estimates are above or below the truth, which is important for accurate precipitation estimation.

Finally, notice that with networks SO+MTV, SO+MT, and SO, the RMS errors of the unobserved variables for assimilation with $\mathbf{R}^{o} = \mathbf{0}$ are larger than those with $\mathbf{R}^{o} > \mathbf{0}$; for example, see the errors for variables θ_1, q, H_c, H_d, P in Figure 24. We find that these larger errors with $\mathbf{R}^{o} = \mathbf{0}$ are due to an ill-conditioned Kalman gain matrix in (36) with sparse observation networks with operator $\mathbf{G} \in \mathbb{R}^{S\times 8}$, $S < 8$.

6. Summary and concluding discussion

In this article, we use multicloud models (Khouider and Majda, 2006a, 2007, 2008a; Majda, *et al.*, 2007) as the test models for filtering moist tropical convection. In particular, we aim to establish guidelines for the future design of filtering schemes in assimilating and predicting tropical atmospheric dynamics. We view the multicloud model, with convective parametrization which includes three cloud types (congestus, deep, and stratiform) above the boundary layer, as a candidate for the simplest toy model for moisture-coupled tropical waves (analogous to the Lorenz-96 for midlatitude weather dynamics) for the following reasons: it is very successful in capturing most of the spectrum of the convectively coupled waves (Khouider and Majda, 2008a,b; Kiladis, *et al.*, 2009) as well as the nonlinear organization of large-scale envelopes mimicking across-scale interactions of the MJO (our first example in section 3.1) and convectively coupled waves (our second example in section 3.2). More importantly, this model also captures the vertical profile

Copyright © 2012 Royal Meteorological Society

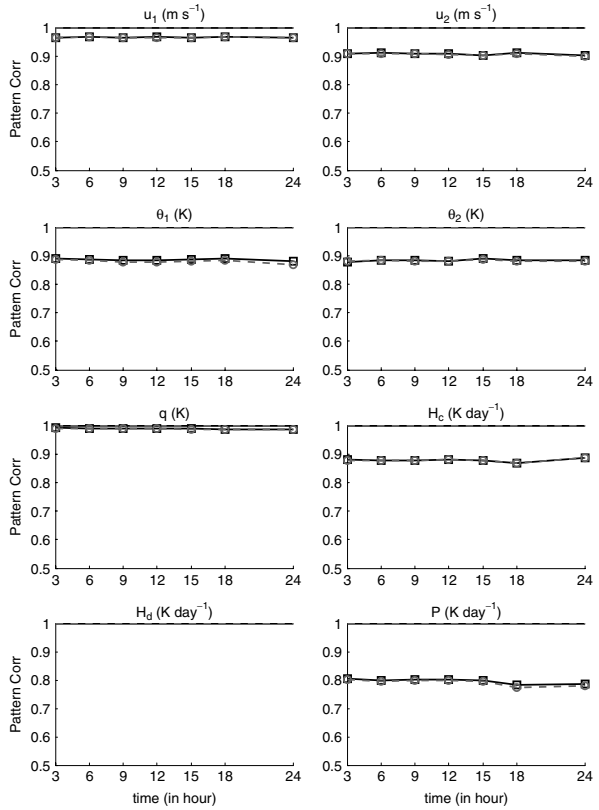

Figure 19. Pattern correlations as functions of observation time interval for observation network CO: MSM filter (grey), complete 3D-Var (black), $\mathbf{R}^o = \mathbf{0}$ (solid line and dashes with no markers), $\mathbf{R}^o > \mathbf{0}$ (dashes with square/circle markers).

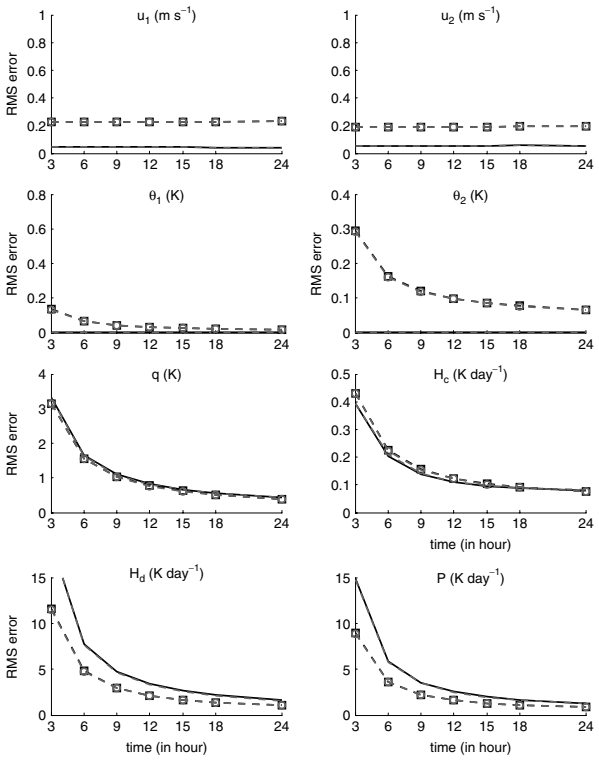

Figure 20. As Figure 18, but for observation network SO+MTV.

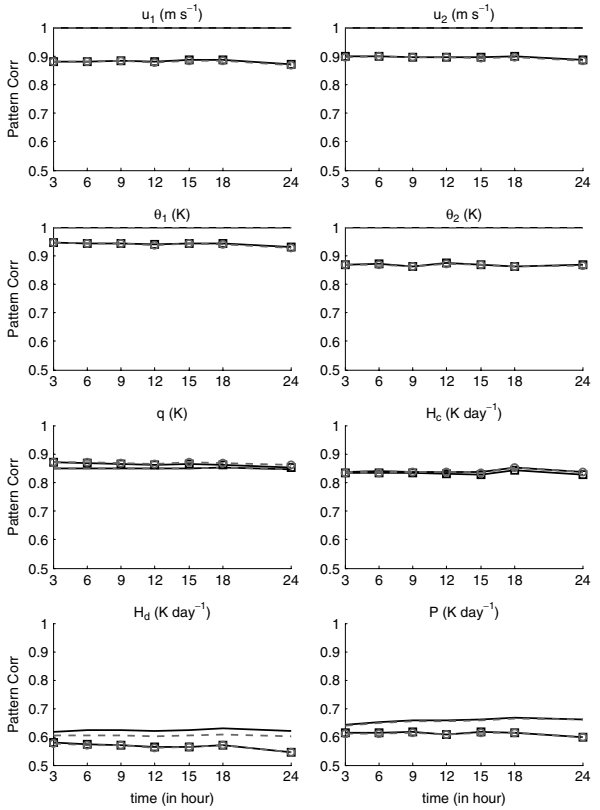

Figure 21. As Figure 19, but for observation network SO+MTV.

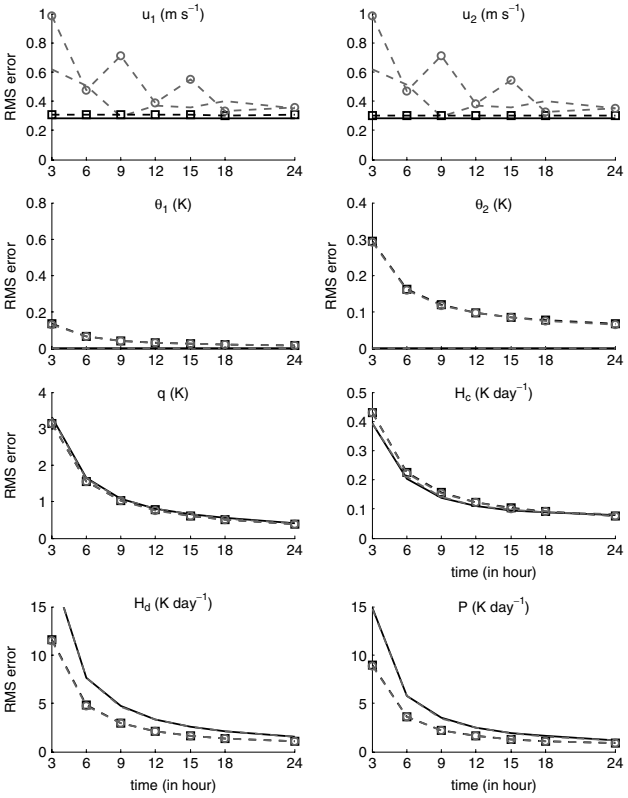

Figure 22. As Figure 18, but for observation network SO+MT.

with front and rear tilting, the phase speed, and dispersion relations that match the observational record (Kiladis, *et al.*, 2005, 2009).

Here, we demonstrate the filtering skill with a suite of reduced stochastic filters with model errors, based on linear stochastic models (Harlim and Majda, 2008a, 2010a; Majda and Harlim, 2012) in capturing the intraseasonal MJO-like

Copyright © 2012 Royal Meteorological Society

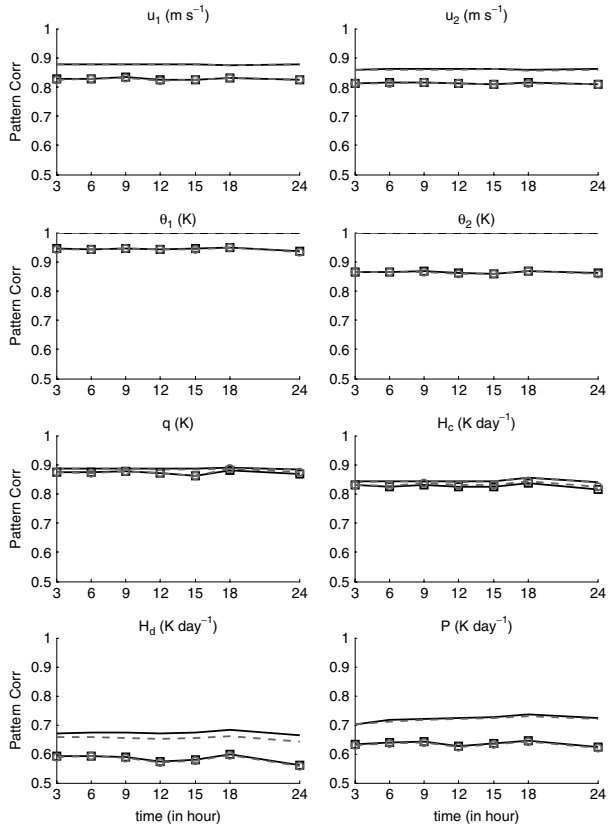

Figure 23. As Figure 19, but for observation network SO+MT.

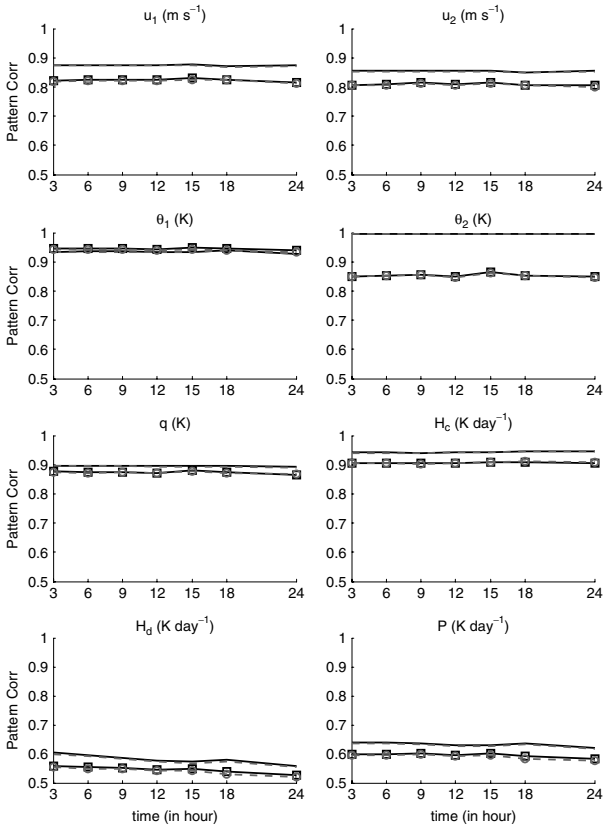

Figure 25. As Figure 19, but for observation network SO.

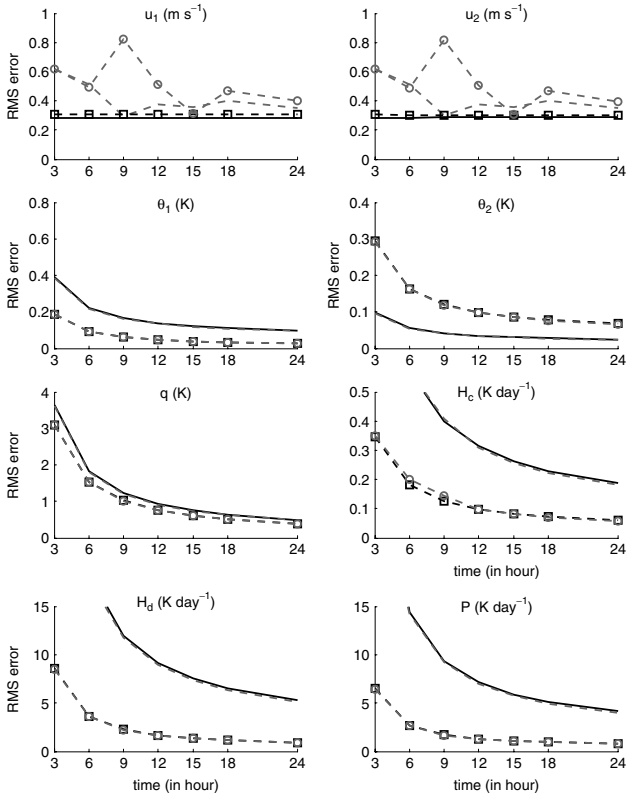

Figure 24. As Figure 18, but for observation network SO.

wave (Majda, *et al.*, 2007) and the transient initiation of a convectively coupled wave train that resembles the results from simulations with a cloud-resolving model (Grabowski and Moncrieff, 2001). From these numerical experiments, we find the following:

(i) The key factor for accurate precipitation estimates is an accurate estimation of the first baroclinic potential temperature. Our test problems suggests that slight overestimation in θ_1 produces a dry atmosphere with no rain at all and slight underestimation in θ_1 produces a wet atmosphere with unrealistic high precipitation rate;

(ii) Our simple reduced stochastic filters are able to recover moisture and precipitation field profiles (even when online observations of these variables are not available) provided that the filter forward prior model is designed in a moisture-coupled eigenmode basis. This result suggests that the future design of tropical data assimilation algorithms should account for a moisture-coupled eigenmode basis instead of dry eigenmode basis as in Žagar, *et al.* (2004a,b);

(iii) A better estimate for the tropical convection wave patterns requires more than surface wind and potential temperature observations;

(iv) The skill of the reduced filtering methods with horizontally and vertically sparse observations suggests that more accurate filtered solutions are achieved with less frequent observation times. Such a counter-intuitive finding is justified through an analysis of the classical observability and controllability conditions which are necessary for optimal filtering, especially when the observation time-scale is too short relative to the time-scale of the true signal.

Copyright © 2012 Royal Meteorological Society

We hope that the encouraging results in this article can convince researchers who are interested in tropical data assimilation to investigate:

(i) The potential of improving the estimate accuracy with more sophisticated filtering schemes through the test models here; this includes testing with the ensemble Kalman filters (Evensen, 1994; Anderson, 2001; Bishop, *et al.*, 2001; Hunt, *et al.*, 2007), particle filters with small ensemble sizes (Anderson, 2010; van Leeuwen, 2010), and other reduced stochastic filters (Gershgorin, *et al.*, 2010a,b; Majda and Harlim, 2012) or the Gaussian closure filter (Branicki, *et al.*, 2012) with stochastic parametrizations that account for model errors 'on the fly';
(ii) Exploration of the filtering skill for other parameter regimes in multicloud models with more realistic SST profile, active barotropic dynamics and coupled nonlinear advection which allows for tropical–extratropical wave interactions (Lin, *et al.*, 2000; Majda and Biello, 2003; Biello and Majda, 2004);
(iii) Use of the test models to design appropriate filters to cope with various observation networks with irregularly spaced sparse observations (Harlim, 2011) and satellite measurements.

Acknowledgements

The authors thank Boualem Khouider, Samuel Stechmann, and Yevgeniy Frenkel for sharing the relevant multicloud model source codes. The research of JH is partially supported by the Office of Naval Research grant N00014-11-1-0310, the NC State start-up fund, and the NC State Faculty Research and Professional Development fund. The research of AJM is partially supported by the National Science Foundation grant DMS-0456713 and the Office of Naval Research grants ONR DRI N00014-10-1-0554 and N00014-11-1-0306.

References

Abramov R, Majda AJ. 2007. Blended response algorithm for linear fluctuation-dissipation for complex nonlinear dynamical systems. *Nonlinearity* **20**: 2793–2821.

Anderson B, Moore J. 1979. *Optimal filtering*. Prentice-Hall Englewood Cliffs, NJ, USA.

Anderson J. 2001. An ensemble adjustment Kalman filter for data assimilation. *Mon. Weather Rev.* **129**: 2884–2903.

Anderson JL. 2010. A non-Gaussian ensemble filter update for data assimilation. *Mon. Weather Rev.* **138**: 4186–4198.

Biello JA, Majda AJ. 2004. The effect of meridional and vertical shear on the interaction of equatorial baroclinic and barotropic rossby waves. *Studies Appl. Math.* **112**: 341–390, DOI: 10.1111/j.0022-2526.2004.01518.x

Bishop C, Etherton B, Majumdar S. 2001. Adaptive sampling with the ensemble transform Kalman filter. Part I: the theoretical aspects. *Mon. Weather Rev.* **129**: 420–436.

Branicki M, Gershgorin B, Majda AJ. 2012. Filtering skill for turbulent signals for a suite of nonlinear and linear extended Kalman filters. *J. Comput. Phys.* **231**: 1462–1498, DOI: 10.1016/j.jcp.2011.10.029

Crommelin D, Vanden-Eijnden E. 2008. Subgrid-scale parameterization with conditional Markov chains. *J. Atmos. Sci.* **65**: 2661–2675.

DelSole T. 2000. A fundamental limitation of Markov models. *J. Atmos. Sci.* **57**: 2158–2168.

DelSole T. 2004. Stochastic model of quasi-geostrophic turbulence. *Surveys Geophys.* **25**: 107–149.

Evensen G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99**: 10143–10162.

Frenkel Y, Majda AJ, Khouider B. 2011. Simple models for the diurnal cycle and convectively coupled waves. *Theor. Comput. Fluid Dyn.* submitted.

Frenkel Y, Majda AJ, Khouider B. 2012. Using the stochastic multicloud model to improve tropical convective parameterization: A paradigm example. *J. Atmos. Sci.* **69**: 1080–1105.

Frierson D, Majda AJ, Pauluis O. 2004. Large-scale dynamics of precipitation fronts in the tropical atmosphere: A novel relaxation limit. *Comm. Math. Sci.* **2**: 591–626.

Fuchs E, Raymond DJ. 2002. Large-scale modes of a nonrotating atmosphere with water vapor and cloud-radiation feedbacks. *J. Atmos. Sci.* **59**: 1669–1679.

Gershgorin B, Harlim J, Majda AJ. 2010a. Improving filtering and prediction of spatially extended turbulent systems with model errors through stochastic parameter estimation. *J. Comput. Phys.* **229**: 32–57.

Gershgorin B, Harlim J, Majda AJ. 2010b. Test models for improving filtering with model errors through stochastic parameter estimation. *J. Comput. Phys.* **229**: 1–31.

Grabowski WW, Moncrieff MW. 2001. Large-scale organization of tropical convection in two-dimensional explicit numerical simulations. *Q. J. R. Meteteorol. Soc.* **127**: 445–468.

Han Y, Khouider B. 2010. Convectively coupled waves in a sheared environment. *J. Atmos. Sci.* **67**: 2913–2942.

Harlim J. 2011. Interpolating irregularly spaced observations for filtering turbulent complex systems. *SIAM J. Sci. Comput.* **33**: 2620–2640.

Harlim J, Majda AJ. 2008a. Filtering nonlinear dynamical systems with linear stochastic models. *Nonlinearity* **21**: 1281–1306.

Harlim J, Majda AJ. 2008b. Mathematical strategies for filtering complex systems: Regularly spaced sparse observations. *J. Comput. Phys.* **227**: 5304–5341.

Harlim J, Majda AJ. 2010a. Catastrophic filter divergence in filtering nonlinear dissipative systems. *Comm. Math. Sci.* **8**: 27–43.

Harlim J, Majda AJ. 2010b. Filtering turbulent sparsely observed geophysical flows. *Mon. Weather Rev.* **138**: 1050–1083.

Houze Jr RA. 2004. Mesoscale convective systems. *Rev. Geophys.* **42**: RG4003, DOI: 10.1029/2004RG000150

Hunt B, Kostelich E, Szunyogh I. 2007. Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Physica D* **230**: 112–126.

Khouider B, Majda AJ. 2005a. A non-oscillatory balanced scheme for an idealized tropical climate model: Part I: Algorithm and validation. *Theor. Comput. Fluid Dyn.* **19**: 331–354.

Khouider B, Majda AJ. 2005b. A non-oscillatory balanced scheme for an idealized tropical climate model: Part II: Nonlinear coupling and moisture effects. *Theor. Comput. Fluid Dyn.* **19**: 355–375.

Khouider B, Majda AJ. 2006a. A simple multicloud parametrization for convectively coupled tropical waves. Part I: Linear analysis. *J. Atmos. Sci.* **63**: 1308–1323.

Khouider B, Majda AJ. 2006b. Multicloud convective parameterizations with crude vertical structure. *Theor. Comput. Fluid Dyn.* **20**: 351–375.

Khouider B, Majda AJ. 2007. A simple multicloud parametrization for convectively coupled tropical waves. Part II: Nonlinear simulations. *J. Atmos. Sci.* **64**: 381–400.

Khouider B, Majda AJ. 2008a. Multicloud models for organized tropical convection: Enhanced congestus heating. *J. Atmos. Sci.* **65**: 895–914.

Khouider B, Majda AJ. 2008b. Equatorial convectively coupled waves in a simple multicloud model. *J. Atmos. Sci.* **65**: 3376–3397.

Khouider B, Biello JA, Majda AJ. 2010. A stochastic multicloud model for tropical convection. *Comm. Math. Sci.* **8**: 187–216.

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

Kiladis GN, Straub KH, Haertel PT. 2005. Zonal and vertical structure of the Madden–Julian oscillation. *J. Atmos. Sci.* **62**: 2790–2809.

Kiladis GN, Wheeler MC, Haertel PT, Straub KH, Roundy PE. 2009. Convectively coupled equatorial waves. *Rev. Geophys.* **47**: RG2003, DOI: 10.1029/2008RG000266

Lau WKM, Waliser DE (eds). 2005. *Intraseasonal variability in the atmosphere–ocean climate system*. Springer: Berlin.

Lin JL, Kiladis GN, Mapes BE, Weickmann KM, Sperber KR, Lin W, Wheeler MC, Schubert SD, Del Genio A, Donner LJ, Emori S, Gueremy JF, Hourdin F, Rasch PJ, Roeckner E, Scinocca JF. 2006. Tropical intraseasonal variability in 14 IPCC AR4 climate models. Part I: Convective signals. *J. Climate* **19**: 2665–2690.

Lin JWB, Neelin JD, Zeng N. 2000. Maintenance of tropical intraseasonal variability: Impact of evaporation–wind feedback and midlatitude storms. *J. Atmos. Sci.* **57**: 2793–2823.

Copyright © 2012 Royal Meteorological Society

Lorenz E. 1996. Predictability – a problem partly solved. In *Proceedings on predictability*, 4–8 September 1995. ECMWF: Reading, UK. 1–18.

Majda AJ, Biello JA. 2003. The nonlinear interaction of barotropic and equatorial baroclinic Rossby waves. *J. Atmos. Sci.* **60**: 1809–1821.

Majda AJ, Harlim J. 2012. *Filtering Complex Turbulent Systems*. Cambridge University Press: Cambridge, UK.

Majda AJ, Shefter MG. 2001. Waves and instabilities for model tropical convective parameterizations. *J. Atmos. Sci.* **58**: 896–914.

Majda AJ, Souganidis PE. 2010. Existence and uniqueness of weak solutions for precipitation fronts: A novel hyperbolic free boundary problem in several space variables. *Comm. Pure Appl. Math.* **63**(10): 1351–1361.

Majda AJ, Stechmann SN. 2009a. A simple dynamical model with features of convective momentum transport. *J. Atmos. Sci.* **66**: 373–392.

Majda AJ, Stechmann SN. 2009b. The skeleton of tropical intraseasonal oscillations. *Proc. Nat. Acad. Sci.* **106**(21): 8417.

Majda AJ, Stechmann SN. 2011. Nonlinear dynamics and regional variations in the MJO skeleton. *J. Atmos. Sci.* **68**: 3053–3071.

Majda AJ, Timofeyev I. 2004. Low-dimensional chaotic dynamics versus intrinsic stochastic chaos: A paradigm model. *Physica D* **199**: 339–368.

Majda AJ, Timofeyev I, Vanden-Eijnden E. 1999. Models for stochastic climate prediction. *Proc. Nat. Acad. Sci.* **96**: 15687–15691.

Majda AJ, Khouider B, Kiladis GN, Straub KH, Shefter MG. 2004. A model for convectively coupled tropical waves: nonlinearity, rotation, and comparison with observations. *J. Atmos. Sci.* **61**: 2188–2205.

Majda AJ, Abramov R, Grote M. 2005. *Information theory and stochastics for multiscale nonlinear systems*. CRM Monograph Series **25**: Amer. Math. Soc: Providence, RI, USA.

Majda AJ, Stechmann S, Khouider B. 2007. Madden–Julian Oscillation analog and intraseasonal variability in a multicloud model above the Equator. *Proc. Nat. Acad. Sci.* **104**(24): 9919–9924.

Majda AJ, Franzke C, Khouider B. 2008. An applied mathematics perspective on stochastic modelling for climate. *Phil. Trans. A Math. Phys. Eng. Sci.* **366**(1875): 2429–2455.

Majda AJ, Gershgorin B, Yuan Y. 2010. Low frequency response and fluctuation-dissipation theorems: Theory and practice. *J. Atmos. Sci.* **67**: 1181–1201.

Moncrieff M, Shapiro M, Slingo J, Molteni F. 2007. Collaborative research at the intersection of weather and climate. *WMO Bulletin* **56**(3): 1–9.

Nakazawa T. 1988. Tropical super clusters with intraseasonal variations over western Pacific. *J. Meteorol. Soc. Japan* **66**: 823–839.

Neelin J, Zeng N. 2000. A quasi-equilibrium tropical circulation model formulation. *J. Atmos. Sci.* **57**: 1741–1766.

Pauluis O, Frierson DMW, Majda AJ. 2008. Precipitation fronts and the reflection and transmission of tropical disturbances. *Q. J. R. Meteorol. Soc.* **134**: 913–930.

Penland C. 1989. Random forcing and forecasting using principal oscillation pattern analysis. *Mon. Weather Rev.* **117**: 2165–2185.

Stechmann S, Majda AJ. 2006. The structure of precipitation fronts for finite relaxation time. *Theor. Comput. Fluid Dyn.* **20**(5–6): 377–404.

Stechmann SN, Majda AJ. 2009. Gravity waves in shear and implications for organized convection. *J. Atmos. Sci.* **66**: 2579–2599.

Straub K, Kiladis G. 2002. Observations of a convectively coupled Kelvin waves in the eastern Pacific ITCZ. *J. Atmos. Sci.* **59**: 30–53.

van Leeuwen P. 2010. Nonlinear data assimilation in geosciences: an extremely efficient particle filter. *Q. J. R. Meteorol. Soc.* **136**: 1991–1999.

Wheeler M, Kiladis G. 1999. Convectively coupled equatorial waves: Analysis of clouds and temperature in the wavenumber–frequency domain. *J. Atmos. Sci.* **56**: 374–399.

Wheeler M, Kiladis GN, Webster PJ. 2000. Large-scale dynamical fields associated with convectively coupled equatorial waves. *J. Atmos. Sci.* **57**: 613–640.

Wilks DS. 2005. Effects of stochastic parameterizations in the Lorenz '96 model. *Q. J. R. Meteorol. Soc.* **131**: 389–407.

Žagar N. 2012. Multivariate data assimilation in the tropics by using equatorial waves. *Pure Appl. Geophys.* **169**: 367–379. 10.1007/s00024-011-0375-2.

Žagar N, Gustafsson N, Kallén E. 2004a. Dynamical response of equatorial waves in four-dimensional variational data assimilation. *Tellus A* **56**: 29–46.

Žagar N, Gustafsson N, Kallén E. 2004b. Variational data assimilation in the Tropics: The impact of a background-error constraint. *Q. J. R. Meteorol. Soc.* **130**: 103–125.

Zhang C. 2005. Madden–Julian oscillation. *Rev. Geophys.* **43**: RG2003, DOI: 10.1029/2004RG000 158

Copyright © 2012 Royal Meteorological Society