RESEARCH ARTICLE

Ensemble Kalman filter based data assimilation for tropical waves in the MJO skeleton model

Tabea Gleiter[1] | Tijana Janjić[1] | Nan Chen[2]

[1]Meteorologisches Institut, Ludwig-Maximilians-Universität, Munich, Germany

[2]Department of Mathematics, University of Wisconsin–Madison, Madison, Wisconsin

Correspondence
T. Janjić, Meteorologisches Institut, Ludwig-Maximilians-Universität, Theresienstr. 37, Munich 80333, Germany.
Email: tijana.pfander@lmu.de

Funding information
Deutsche Forschungsgemeinschaft : Heisenberg Award (DFG JA1077/4-1); Transregional Collaborative Research Center SFB / TRR 165 "Waves to Weather"; Office of Naval Research (ONR), Grant/Award Number: N00014-21-1-2904

Abstract

The Madden–Julian oscillation (MJO) is the dominant component of tropical intraseasonal variability, with wide-reaching impacts even on extratropical weather and climate patterns. However, predicting the MJO is challenging. One reason is the suboptimal state estimates obtained with standard data assimilation (DA) approaches. These are typically based on filtering methods with Gaussian approximations and do not take into account physical properties that are important specifically for the MJO. In this article, a constrained ensemble DA method is applied to study the impact of different physical constraints on the state estimation and prediction of the MJO. The quadratic programming ensemble (QPEns) algorithm utilized extends the standard stochastic ensemble Kalman filter (EnKF) with specifiable constraints on the updates of all ensemble members. This allows us to recover physically more consistent states and to respect possible associated non-Gaussian statistics. The study is based on identical twin experiments with an adopted nonlinear model for tropical intraseasonal variability. This so-called skeleton model succeeds in reproducing the main large-scale features of the MJO and closely related tropical waves, while keeping adequate simplicity for fast experiments on intraseasonal time-scales. Conservation laws and other crucial physical properties from the model are examined as constraints in the QPEns. Our results demonstrate an overall improvement in the filtering and forecast skill when the model's total energy is conserved in the initial conditions. The degree of benefit is found to be dependent on the observational setup and the strength of the model's nonlinear dynamics. It is also shown that, even in cases where the statistical error in some waves remains comparable with the stochastic EnKF during the DA stage, their prediction is improved remarkably when using the initial state resulting from the QPEns.

KEYWORDS
conservation laws, constrained ensemble DA, MJO, non-Gaussianity, QPEns, skeleton model, stochastic EnKF, tropical waves

1 | INTRODUCTION

The Madden–Julian oscillation (MJO) is the most prominent component of tropical intraseasonal variability. It consists of a deep convective center with a surrounding dry dynamic structure that propagates slowly eastward, typically from the Indian Ocean through the western Pacific (Zhang, 2005). Nowadays, the MJO is known to be the planetary-scale envelope of multiple convectively coupled equatorial waves (CCEWs: see Kiladis et al., 2009), mostly equatorial Rossby waves of small meridional mode number and Kelvin waves (Castanheira and Marques, 2021). As the MJO influences tropical cyclones, monsoons, and the El Niño–Southern Oscillation (ENSO), as well as extratropical weather and climate patterns, it is of great societal importance. Therefore, understanding and predicting the MJO and its most important contributing tropical waves is a central problem in contemporary meteorology (Vitart and Molteni, 2010; Khouider et al., 2013; Zhang et al., 2013).

However, forecasting the MJO is a challenging task. Depending on several factors such as the MJO indices used, the ensemble size, or the season, it has been shown by Vitart (2017) and Wang et al. (2019) that the MJO and its convection can be predicted for up to 4–5 weeks in the best subseasonal operational models, in particular the ECMWF model. In their studies, the predictability limit is defined as the forecast lead time at which the bivariate correlation coefficient between forecast and observations falls below 0.6. These results are in good accordance with the prediction skills for large-scale MJO structures that are achieved by low-order statistical models (Seo et al., 2009; Kondrashov et al., 2013; Chen et al., 2014). This means that, although remarkable progress was made in the last decade, predictions of the MJO are nowadays only possible up to roughly one month, which is insufficient for an intraseasonal oscillation. The challenges in forecasting the MJO come, on one hand, from the fact that the detailed mechanisms behind the complex and hierarchical organization of tropical convection are not yet fully understood and modelled. It has been shown that better parameterizations of subplanetary-scale convective processes in general circulation models (GCMs) become crucial in characterizing the MJO (see Khouider et al., 2011; Deng et al., 2015; Yang et al., 2019). On the other hand, there is a general lack of data assimilation (DA) systems tailored to the features of the tropical atmosphere. Most operational state estimation algorithms are customized for dry dynamic geostrophic balance in the midlatitudes and based on filtering methods that are optimized for linear models, as they rest on the assumption of Gaussian probability density functions (PDFs). In the tropics, however, such comparably dominant and simple relationships are

absent and atmospheric motions are instead shaped by nonlinear processes, in particular convection. For these reasons, interesting approaches to tropical DA (see, e.g., Žagar et al., 2004; 2005; 2016) and adaptions of existing filters for non-Gaussian situations (see Bocquet et al., 2010) have been proposed in recent years. However, little focus has so far been placed on the MJO. It is thus of special importance to develop a suitable DA algorithm for the MJO.

DA methods that respect non-Gaussian moments directly, such as particle filters (PFs), are not yet sufficiently robust and computationally affordable for high-dimensional atmospheric applications. Although strategies to mitigate these issues are progressively being researched (see, e.g., the local PF described in Penny and Miyoshi, 2016; Poterjoy, 2016), operational algorithms still mostly rely on assumptions of Gaussianity. Often, ensemble Kalman filters (EnKFs) are adopted inadequately for nonlinear problems (Kalnay, 2003; Evensen, 2009; Law et al., 2015). Moreover, traditional DA approaches are not able to respect specific physical properties, as is commonly done in numerical modelling. For example, the strict positivity of certain variables is often treated only rudimentarily by truncations of undesirably negative values to zero. Therefore, the physical consistency of the assimilated states cannot be guaranteed (Janjić et al., 2014; Zeng and Janjić, 2016), which can induce artificial wave activities in the subsequent forecast or even filter divergence (Houtekamer and Zhang, 2016; Žagar et al., 2016). In order to address these problems, a constrained DA algorithm, the so-called quadratic programming ensemble (QPEns), was recently developed by Janjić et al. (2014). It extends the typical stochastic EnKF with state constraints on the analysis ensemble members and thereby enables the conservation of physical properties such as mass, energy, or positivity of specific variables. This turns the analytic update step into an ensemble of constrained numerical optimization problems with a quadratic cost function. As was shown in the first applications to several conceptual models (see Zeng et al., 2017; Ruckstuhl and Janjić, 2018), the QPEns with well-chosen constraints can especially improve the filtering skill in non-Gaussian situations, since important physical properties are then easily hurt by a filter with Gaussian assumptions. In this article, we use this algorithm to investigate which physical constraints imposed on the initial condition in the tropics could improve the prediction of large-scale, nonlinear MJO dynamics. More specifically, we test a QPEns DA setup with different constraints applied to a simplified model for tropical intraseasonal variability (hereafter called the "skeleton model").

The skeleton model was first developed by Majda and Stechmann (2009) in its basic linearized form and has

become well known for its good dynamical representation of the MJO and important associated tropical waves at the planetary scale, despite its intermediate complexity. The model uses approximated dry dynamic equations that are closely related to the standard Matsuno–Gill model and incorporates a coarse convection parameterization together with a leading-order meridional and vertical mode truncation. Since its first formulation, it has been extended by various modifications, in particular the inclusion of nonlinearity and stochasticity in the convection parameterization by Majda and Stechmann (2011) and Thual *et al.* (2014), and to versions with meridional and vertical truncation to higher order modes by Thual *et al.* (2015) and Thual and Majda (2016), respectively. The skeleton model has also been coupled to different ocean models to describe the interactions between the MJO and the ENSO (see Thual *et al.*, 2018; Yang *et al.*, 2021). Although the model's realism naturally increases with the level of complexity, and particularly with the inclusion of smaller-scale physical processes, in this study we focus on the rather basic nonlinear deterministic version as in Majda and Stechmann (2011). This version realistically simulates important large-scale features of equatorial Kelvin and Rossby waves of the smallest meridional mode number as well as MJO signals, but still allows for computationally affordable and easily interpretable DA experiments on the intraseasonal time-scale, due to its reduced state space and relatively simple mathematical modeling structure. As it furthermore incorporates a realistic non-Gaussian climatology, it provides a good framework to study new approaches for filtering MJO-related planetary-scale atmospheric dynamics. Note, however, that there are other MJO models and theories that are also widely accepted, such as the moisture mode theory for the MJO (Sobel and Maloney, 2012; 2013).

In order to study the influence of dynamical constraints in DA when applied to the skeleton model, we set up an identical twin experiment framework. This is beneficial, as it resolves the need to retrieve suitable observations for the simple model from real data and simplifies the diagnostics, as the truth is known. Moreover, any error sources in the system are reduced to the update step, in particular any neglect of non-Gaussian moments as well as imperfections in localization and inflation, which are used in our DA system to mitigate sampling errors. Our experiments test the influences of different analysis constraints on filtering and forecasting the large-scale MJO structure and important CCEW. Therefore, in addition to exploring the DA skill in physical space, diagnostics for tropical waves in the skeleton model are set up. The QPEns with different constraints is compared with its underlying stochastic EnKF, which is used here as the reference DA algorithm. In particular, we examine constraints of the two energy

quantities that are conserved in the model, that is, the moist static energy and total energy, as well as a positivity constraint of convective activity, which is also included in the model dynamics. Moreover, a constraint of dry mass is also tested, as mass is a common conservation property in many numerical weather prediction models, its constraint has proven beneficial in earlier QPEns studies (see Janjić *et al.*, 2014; Ruckstuhl and Janjić, 2018), and it is conserved in its long-time average in the skeleton model. As we find a particularly positive influence of the total energy constraint, we investigate this constraint further in combination with different observations and different amounts of non-Gaussianity in the model.

The rest of the article is organized as follows. Section 2 includes a summary of the skeleton model as it is used in this study. Section 3 describes the setup of the identical twin experiments, the DA algorithms, and the diagnostics. The DA results are presented in Section 4 and the discussion and conclusions are included in Section 5.

2 | THE SKELETON MODEL

In this section, the skeleton model is introduced in its nonlinear deterministic version as used in this study. The model's underlying physical equations as well as the projections for its numerical solution and the settings used in this article are explained in Section 2.1. In Section 2.2, the key dynamical features found in earlier works are summarized and the separation of wave signals from the model states is described. Finally, the climatological properties that are relevant for DA purposes, in particular conservation quantities and PDFs, are presented in Section 2.3.

2.1 | Model equations and projections

The skeleton model is a dynamical model with intermediate complexity that simulates tropical intraseasonal variability, especially the MJO and its relevant tropical waves, at the planetary scale. To retrieve its nonlinear deterministic version as introduced in Majda and Stechmann (2011) and used in this study, the Matsuno–Gill model for tropical large-scale dynamics (see Matsuno, 1966; Gill, 1980) is coupled via the diabatic heating term to additional equations for the evolution of moisture and convection. The latter are designed around a coarse parameterization for the growth of convective activity that is derived from theoretical as well as modelling and observational indications. Moreover, the model regards the MJO as a neutrally stable wave at the planetary scale driven only by convective instabilities at the subplanetary scale. Any large-scale dissipation is thus neglected, except for a

constant radiative cooling. Furthermore, a vertical truncation of the dry dynamics to the first baroclinic mode and the replacement of pressure by potential temperature via the hydrostatic relationship is performed. This returns the following as underlying physical equations with periodic boundary conditions along the equatorial belt and equatorial longwave scaling:

$$\frac{\partial u}{\partial t} - yv - \frac{\partial \theta}{\partial x} = 0, \tag{1a}$$

$$yu - \frac{\partial \theta}{\partial y} = 0, \tag{1b}$$

$$\frac{\partial \theta}{\partial t} - \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) = \overline{H}a - s^{\theta}, \tag{1c}$$

$$\frac{\partial q}{\partial t} + \overline{Q} \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = -\overline{H}a + s^{q}, \tag{1d}$$

$$\frac{\partial a}{\partial t} = \Gamma q a. \tag{1e}$$

Here, x and y are the zonal and meridional positions and t is the time. In the dry dynamics (Equations 1a–1c), u, v, and θ are the zonal and meridional velocities and potential temperature, respectively. Equation 1d describes the evolution of the lower-tropospheric moisture q. All variables are anomalies from radiative–convective equilibrium (RCE), except the planetary-scale envelope of synoptic-scale convective activity (in the following "convective activity" for short), a. The skeleton model contains a minimal number of parameters: $\overline{Q}$ is the mean background vertical moisture gradient; Γ is the growth rate of convective activity; $\overline{H}$ is irrelevant to the dynamics but is a scaling constant; s^{θ} and s^{q} are external sources of radiative cooling and moistening/latent heating that are prescribed in the system. The model variables, with dimensions and physical parameters, are summarized in Table 1.

Next, Equations 1 are projected on to the parabolic cylinder functions (PCFs) in the meridional direction. Thereby, the dry dynamic variables can be reformulated into model variables K and R_m, with $m \geq 1$, which are the amplitudes of the Kelvin wave and equatorial Rossby waves of different meridional mode numbers. Note, however, that these amplitudes only propagate as known from the theory of dry tropical waves if they are unforced. In the case of forcing as in the skeleton model, they are coupled to convection and each other. A subsequent truncation of the diabatic heating variables/structures (a, s^{θ}, and s^{q}) based on the PCF projection determines the number of excited Rossby-wave modes and thus reduces the model complexity. In this article, we apply the leading-order meridional truncation, which has been widely used in the literature

on the skeleton model and is sufficient to recover the key features of the MJO. In other words, only the projection on to the first PCF $\Phi_0(y) \propto \exp(-y^2/2)$ (see Majda and Stechmann, 2011) is retained in the meridional diabatic heating structure. Moreover, $s^{\theta} = s^q$ is assumed for reasons of simplicity. The resulting system then excites the Kelvin wave K and the first equatorial Rossby wave $R := R_1$. It reads

$$\frac{\partial K}{\partial t} + \frac{\partial K}{\partial x} = -\frac{1}{\sqrt{2}} \left(\overline{H}A - S \right), \tag{2a}$$

$$\frac{\partial R}{\partial t} - \frac{1}{3} \frac{\partial R}{\partial x} = -\frac{2\sqrt{2}}{3} \left(\overline{H}A - S \right), \tag{2b}$$

$$\frac{\partial Q}{\partial t} + \frac{\overline{Q}}{\sqrt{2}} \frac{\partial K}{\partial x} - \frac{\overline{Q}}{6\sqrt{2}} \frac{\partial R}{\partial x} = -\left(1 + \frac{\overline{Q}}{6} \right) \left(\overline{H}A - S \right), \tag{2c}$$

$$\frac{\partial A}{\partial t} = \gamma \Gamma A Q, \tag{2d}$$

where Q, A, and S are the projections of q, a, and s^{θ} as well as s^q on to $\Phi_0(y)$, while γ is a projection coefficient (see Table 1) for the interaction of two truncated variables in the truncated system. The numerical parameters for the solution of the skeleton model are also listed in Table 1 and the physical variables can be reconstructed as

$$u(x,y) = \left(\frac{K(x)}{\sqrt{2}} - \frac{R(x)}{2\sqrt{2}} \right) \Phi_0(y) + \frac{R(x)}{4} \Phi_2(y),$$

$$v(x,y) = \left(\frac{\partial R(x)}{\partial x} - \frac{\overline{H}A(x) + S(x)}{\sqrt{2}} \right) \frac{\Phi_1(y)}{3},$$

$$\theta(x,y) = \left(-\frac{K(x)}{\sqrt{2}} - \frac{R(x)}{2\sqrt{2}} \right) \Phi_0(y) - \frac{R(x)}{4} \Phi_2(y),$$

$$q(x,y) = Q(x) \cdot \Phi_0(y),$$

$$\overline{H}a(x,y) = \overline{H}A(x) \cdot \Phi_0(y). \tag{3}$$

Note that, corresponding to the variables K, R, Q, and A in Equation 2 projected onto the leading meridional basis, the dry dynamics components in Equation 3 also contain terms with the next PCFs $\Phi_1(y)$ and $\Phi_2(y)$ (see Majda and Stechmann, 2011). These terms are necessary to retrieve the quadrupole structure of the MJO, but are not dynamically relevant. Thus, in the following, the values of the variables in physical space consider only the components associated with the leading PCF in the meridional direction in Equation 3. Amongst other things, this means that any meridional winds are not considered here. Furthermore, the variable values cannot be considered as amplitudes in a 3D grid, but can be regarded as projection amplitudes on the structures explained above in the meridional and vertical directions.

TABLE 1 Variables, dimensions, and physical as well as numerical parameters in the skeleton model as used in this article

Variable	Meaning	Dimension/value
t	Time	$\sqrt{\frac{\pi}{NH\beta}}$
x, y	Zonal, meridional position	$\sqrt{\frac{NH}{\pi\beta}}$
u, v	Zonal, meridional velocity anomalies	$\frac{NH}{\pi}$
θ	Potential temperature anomalies	$\frac{HN^2\theta_0}{\pi g}$
q	Lower-tropospheric moisture anomalies	$\frac{c_p}{L_v}\frac{HN^2\theta_0}{\pi g}$
$\overline{H}a$	Planetary-scale envelope of synoptic-scale convective activity	$\sqrt{\frac{H^3 N^5 \theta_0^2 \beta}{\pi^3 g^2}}$
Dimensional parameter		
β	Meridional derivative of Coriolis parameter	2.28×10^{-11} (ms)$^{-1}$
H	Tropopause height	16 km
N^2	Brunt–Väisälä frequency	10^{-4} s^{-2}
g	Gravitational acceleration	9.8 m · s^{-2}
θ_0	Surface potential temperature	300 K
c_p	Specific heat of dry air at constant pressure	1006 J · kg^{-1} · K^{-1}
L_v	Latent heat of vaporization	2.5×10^6 J · kg^{-1}
Nondimensional parameter		
$\overline{Q}$	Mean background vertical moisture gradient	0.9
$\overline{H}$	Scaling constant for convective activity (without dynamical meaning)	0.22
s^θ, s^q	Background radiative cooling, background latent heating (arrays with values at all grid points)	Modelled as warm-pool $s^\theta_{WP}(x) = s^q_{WP}(x) = 0.022 \times \left(1 - 0.6\cos\left(\frac{2\pi x}{L}\right)\right)$ (Majda *et al.*, 2019)
Γ	Convective activity growth rate	1.66
Numerical parameter		
L	Zonal length of domain	$26.\overline{6}$, that is, 40,000 km
Δx	Distance between zonal grid points	0.4167, that is, 625 km
n_x	Number of zonal grid points	64
Δt	Length of time steps	0.2083, that is, 1.66 hr
γ	Projection coefficient for the truncation	$\int_{-\infty}^{\infty}\Phi_0(y)\Phi_0(y)\Phi_0(y)\,dy = \sqrt{\frac{2}{3}}\pi^{-1/4}$ (Majda *et al.*, 2019)

2.2 | Key dynamical features

The linear solutions of Equation 2 with a uniform background (see Majda and Stechmann, 2009) exhibit plane-wave eigenmodes for different zonal wavenumbers that can be associated with the MJO, dry Kelvin, and dry and moist Rossby waves of the smallest meridional mode number. It should be noted, however, that these wave modes in fact all consist of coupled Kelvin and Rossby waves, the amplitudes of which are given by K and R, that is, the naming refers only to the dominant Kelvin or Rossby component, the corresponding main large-scale

characteristics, and the amount of convective coupling ("dry" and "moist"). The MJO eigenmodes cannot be clearly assigned to either Kelvin or Rossby waves, but capture several essential observed large-scale features of the MJO, which include the following in particular:

- a deep convective center that is moving eastward at a phase speed of approximately 5 m·s^{-1};

- a peculiar dispersion relation with constant intraseasonal oscillation periods of 30–90 days for the leading few zonal wavenumbers; and

TABLE 2 Frequencies (eigenvalues) and amplitudes of the different characteristic components (eigenvectors) of the plane-wave MJO eigenmodes for zonal wavenumbers $k \in \{1, 2, 3\}$ 1/(40,000 km) in the linearized skeleton model with parameter values as in Table 1

k in $\frac{1}{40,000 \ km}$	ω in $\frac{1}{days}$	$\tilde{K}$	$\tilde{R}$	$\tilde{Q}$	$\tilde{A}'$
1	40.0	$0.3224i$	$-0.8521i$	$-0.1465i$	0.3800
2	35.4	$0.2137i$	$-0.7678i$	$-0.2472i$	0.5661
3	35.1	$0.1627i$	$-0.6728i$	$-0.2977i$	0.6771

Note: The corresponding distribution of the variables $(\tilde{X})$ in physical space is calculated as $X_{k,\mathrm{real}}(x) = \tilde{X}_k^i e^{ikx} + \tilde{X}_k^* e^{-ikx}$.

- a horizontal structure with positive moisture anomalies to the east of the convective center and a quadrupole wind circulation around the convective center.

Table 2 shows the frequencies (eigenvalues) and amplitudes of the different characteristic components (eigenvectors) of the MJO eigenmodes for the leading three zonal wavenumbers. The nonlinear nature of the model version used here (see Majda and Stechmann, 2011) allows for interactions between the different waves. These introduce more realistic occurrences of individual MJO events with variability in strength and lifetime instead of plane-wave propagations, while all important features of the MJO eigenmodes, that is, the coarse structure, the phase speed, and the dispersion relation, are retained. The same is true for the other equatorial wave types. See Figure 1 for a Hovmöller diagram of the wave amplitudes in a model simulation.

To retrieve the zonal wave structures, the data are projected in zonal Fourier space on to the waves' eigenmodes from the linear model. Therefore, in analogy to the MJO index in Stechmann and Majda (2015), we use an inner product, to which the eigenmodes are orthogonal, in order to avoid unrealistic projections of other wavetypes. The projection is done for all planetary-scale structures $(1 \leq k \leq 3)$ and a back-transform to physical space subsequently yields the amplitude distributions. The matrix **M** used for the inner product is

$$\mathbf{M} = \begin{pmatrix} 1 + \frac{1}{2}\frac{\overline{Q}}{1-\overline{Q}} & \frac{1}{4}\frac{\overline{Q}}{1-\overline{Q}} & -\frac{1}{\sqrt{2}}\frac{1}{1-\overline{Q}} & 0 \\ \frac{1}{4}\frac{\overline{Q}}{1-\overline{Q}} & \frac{3}{8} + \frac{1}{8}\frac{\overline{Q}}{1-\overline{Q}} & -\frac{1}{2\sqrt{2}}\frac{1}{1-\overline{Q}} & 0 \\ -\frac{1}{\sqrt{2}}\frac{1}{1-\overline{Q}} & -\frac{1}{2\sqrt{2}}\frac{1}{1-\overline{Q}} & \frac{1}{\overline{Q}}\frac{1}{1-\overline{Q}} & 0 \\ 0 & 0 & 0 & \frac{1}{\gamma\overline{\Gamma Q}S} \end{pmatrix}. \quad (4)$$

Notably, it has been investigated whether the stochastic Skeleton model, which is a simple extension of the deterministic model used here, is capable of reproducing observed MJO statistics such as the average duration of MJO events and the overall MJO activity (Stachnik *et al.*, 2015). It has also been illustrated that projecting reanalysis

data on to the MJO basis defined by the skeleton model leads to a time series that resembles the real-time multivariate MJO index (see Wheeler and Hendon, 2004) with a pattern correlation of 0.99 (Stechmann and Majda, 2015). All these findings justify the suggestion that the skeleton theory is suitable to represent and understand some main features of the MJO.

A preferential localization of the MJO over areas with enhanced background radiative cooling and latent heating can be observed when modelling the background RCE as a warm pool (see Table 1). This is in good accordance with the geographical confinement of the MJO to the eastern Indian and western Pacific oceans, which have warmer sea-surface temperatures. We thus use this setup in this study. Many more features, such as the complex westward-tilted vertical structure and overturning circulation, a seasonal cycle, or interactions with the extratropics can be incorporated into the skeleton model (Thual *et al.*, 2015; Thual and Majda, 2016). However, they are not the main focus of this study and are thus omitted here.

In analogy to the experiments in Majda and Stechmann (2011) (see the scenario WP-MJO), we initialize the model in our experiments with the normalized MJO eigenmode of the linearized model for the zonal wavenumber 2 multiplied by an amplitude of 0.05. The initialization is important for the simulation results in the skeleton model, which is understandable from the fact that the model is deterministic and holds two energy conservation principles approximately. After the initialization, we conduct a 10-year spin-up run to let the climatology build up due to nonlinearity and interactions with the warm-pool background (see figure 7 in Majda and Stechmann (2011) for a visualization of the typical model spin-up).

2.3 | DA relevant climatological properties

The underlying skeleton model Equations 1 have the following two domain-integrated energy conservation quantities when equalized background forcing $(s^\theta = s^q)$ is assumed:

FIGURE 1 Wave amplitudes during a one-year model run after the model's spin up, calculated by projection on to the linear model's eigenmodes for zonal wavenumbers 1–3 (from Gleiter, 2021) [Colour figure can be viewed at wileyonlinelibrary.com]
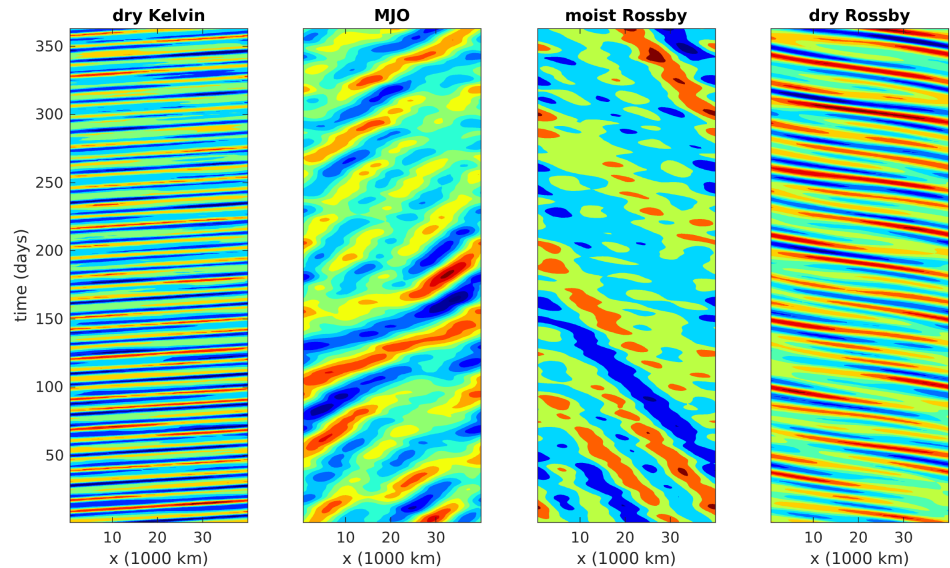
- moist static energy conservation:

$$\int (\theta + q) = \text{const},\qquad (5)$$

and

- total energy conservation (consisting of contributions from dry kinetic energy, potential energy, moist potential energy, and convective energy (Majda *et al.*, 2019):

$$\int \left(\frac{u^2}{2} + \frac{\theta^2}{2} + \frac{1}{2}\frac{\overline{Q}}{1-\overline{Q}}\left(\theta + \frac{q}{\overline{Q}} \right)^2 + \frac{\overline{H}}{\Gamma\overline{Q}}a - \frac{s}{\Gamma\overline{Q}}\ln(a) \right)$$
$$= \text{const}.\qquad (6)$$

These also hold approximately in the meridionally truncated Equation 2. On the other hand, the domain-integrated dry mass (here $\int \theta$) is not conservative in the underlying equations, but it is conserved in its long time average, since θ is only the deviation from the background mean. A further physical property in the skeleton model is the strict positivity of convective activity:

$$a > 0.\qquad (7)$$

The climatological univariate PDFs in the skeleton model with the configuration used in this article are shown in Figure 2. They have a significant amount of non-Gaussianity in the convective activity, connected to its strict positivity, while the other variables are approximately Gaussian-distributed and seem to be rather unaffected by the amount and shape of non-Gaussianity in the convective activity.

3 | METHODS

In this section, the setup of the identical twin experiments conducted in this study is detailed. The identical twin approach means that we generate a synthetic "truth" and synthetic "observations" from a nature run of the skeleton model, while the same model is used as the forecast model for the ensemble propagation in the DA algorithm and subsequent prediction. Such a setup has the advantage that no model error exists and the observational error is exactly known, which allows us to focus on studying impacts of changes in the DA update step, that is, how the different constraints in the QPEns affect the filtering skill. The precisely given truth also simplifies the error assessment and moreover enables us to implement exact state constraints, since all physical properties are known.

The section starts in Section 3.1 with a description of the identical twin setup, thus the initialization of truth and ensemble, and the generation of the observations. Subsequently, the DA algorithms, that is, the stochastic EnKF and the QPEns, are introduced in Section 3.2. The diagnostics employed when discussing the results are explained in Section 3.3.

3.1 | Identical twin setup

Initial values of the truth and the ensemble members

The initial ensemble for the ensemble DA and prediction is selected as equally distributed states from a 100-year climatological run of the skeleton model starting after
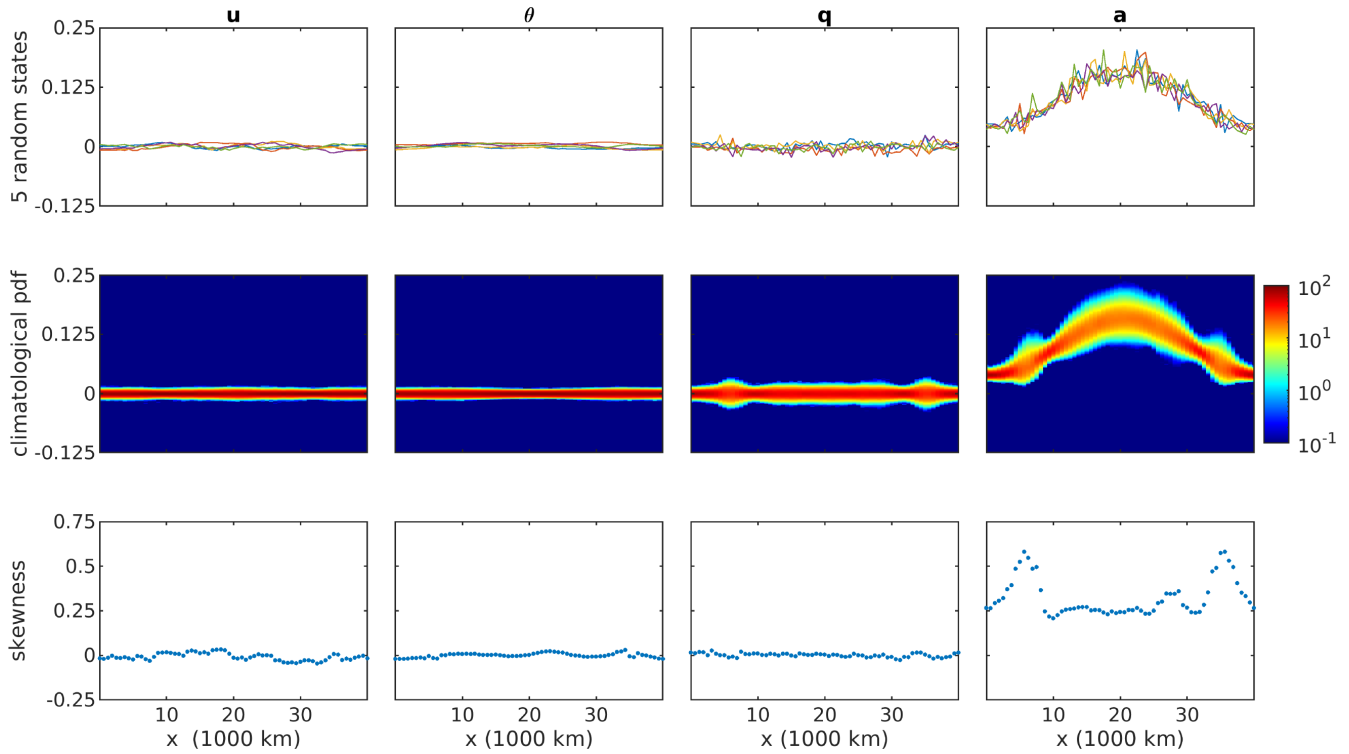
FIGURE 2 Visualizations of skeleton-model variable zonal climatologies retrieved from 12,200 (100 years, every three days) samples starting after 10 years of model spin-up: top: five random draws showing the typical spatial variability; middle: univariate PDFs calculated with 100 histogram bins and moving-average smoothing over 10 values; bottom: skewness values of the univariate climatological PDFs [Colour figure can be viewed at wileyonlinelibrary.com]

10 years of model spin-up. This is possible since the model does not have a seasonal cycle. The ensemble obtained is found to be approximately orthogonal. However, its details only play a subordinate role for the general filter quality after the filter spin-up. All experiments are conducted with an ensemble size of 50 members. The final time instant of the above climatological run is used as the initial value of the truth in the test period.

Generating the observations

The observations contain a subset of the variables in the physical domain (i.e., u, v, θ, and a) and are generated by adding uncorrelated noise to the truth. Except for the noise in the convective activity, all other noise satisfies a zero-mean Gaussian distribution, with variances being equal to 10% of the climatological variances. For the convective activity, which takes only positive values, the noisy observations satisfy a lognormal distribution with the same noise level, that is, same variance. A lognormal distribution is chosen, as it approaches a Gaussian for sufficient distance from the zero point and furthermore has $\lim_{\mathbf{x}\to 0} \text{PDF}_{\text{lognl}}(\mathbf{x}) = 0$.

3.2 | Data assimilation algorithms

3.2.1 | Stochastic ensemble Kalman filter

Algorithm

The stochastic EnKF (Evensen, 1994; Burgers *et al.*, 1998) is a DA algorithm that is widely used in the atmospheric sciences. It updates the background ensemble $\{\mathbf{x}^{\mathrm{b},i}\}_{i=1}^{N}$ with perturbed observations $\{\mathbf{y}^{i}\}_{i=1}^{N}$ using Kalman filter equations. The algorithm requires a specification of the observation-error covariance matrix $\mathbf{R}_k$, while the background-error covariance matrix $\mathbf{P}_k^{\mathrm{b}}$ at time k is calculated from the ensemble as

$$\mathbf{P}_k^{\mathrm{b}} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\mathbf{x}_k^{\mathrm{b},i} - \overline{\mathbf{x}_k^{\mathrm{b},i}}\right)\left(\mathbf{x}_k^{\mathrm{b},i} - \overline{\mathbf{x}_k^{\mathrm{b},i}}\right)^{\mathrm{T}}, \quad (8)$$

with $\overline{\mathbf{x}_k^{\mathrm{b},i}}$ representing the background ensemble mean at time k. The stochastic EnKF reproduces the Kalman filter update statistically. This means that it relies on the same assumption of Gaussianity in its analysis step, based on which the Kalman filter is derived from perfect sequential Bayesian filtering. Nevertheless, its advantages over the

exact Kalman filter are that it is very efficient, it can also be applied with a nonlinear model and/or observational operator, and its outcome can be used directly for ensemble forecasts. In the setup of identical twin experiments, the error sources in the stochastic EnKF are reduced to sampling errors and the neglect of non-Gaussian moments in the PDF. The former are mitigated in this study using inflation and localization.

Generating the perturbed observations

The perturbed observations that are used for the updates of the individual ensemble members are generated as Gaussian samples distributed around the actual observations with the diagonal observation-error covariance $\mathbf{R}_k$. In order to get an accurate representation of the observation, the error in the sample mean is corrected by subtraction. For convective activity, this procedure is again adapted, and the observation perturbations are sampled from a lognormal distribution. The sampling error correction is therefore done differently by a transform from the sampled lognormal distribution, with an erroneous mean, to a lognormal distribution, with a more accurate mean.

Tuning by localization and inflation

The unavoidable sampling errors in the stochastic EnKF introduce the well-known problems of erroneous spurious correlations in the background-error covariance and a self-enforcing decrease in ensemble spread that saturates at a level with far underestimated variance (van Leeuwen, 1999). These issues necessitate some tuning, which is done by the typical measures of localization and inflation (Houtekamer and Zhang, 2016). The details of the implementation in this study are included in the Appendix. Moreover, an additional relaxation of any occurring negative convective activity values in the analysis ensemble to a small value of 10^{-5} was implemented in the stochastic EnKF in order to enable the subsequent model propagation.

3.2.2 | Quadratic programming ensemble (QPEns)

Principle

The idea behind the QPEns introduced by Janjić *et al.* (2014) is to extend the stochastic EnKF by imposing

additional physical constraints on the atmospheric states when updating the ensemble members. This can yield more physically plausible states and also allows us to consider nonlinear relationships and therefore non-Gaussian moments in the background PDF. If used without constraints, the QPEns equals the stochastic EnKF.

Adding constraints transforms the update step into a set of N numerical minimization problems of the form

$$\mathbf{x}_k^{\mathrm{a},i} = \arg\min_{\mathbf{x}} \; \mathcal{J}_k^i(\mathbf{x})$$

$$\text{subject to} \quad c_l(\mathbf{x}) = 0, \; l \in \mathcal{E} \quad \text{and/or} \quad c_m(\mathbf{x}) \leq 0, \; m \in \mathcal{I}, \tag{9}$$

$$\text{with} \quad \mathcal{J}_k^i(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_k^{\mathrm{b},i})^{\mathrm{T}} \mathbf{P}_k^{\mathrm{b}-1}(\mathbf{x} - \mathbf{x}_k^{\mathrm{b},i})$$
$$+ \frac{1}{2}(\mathbf{y}_k^i - \mathcal{H}(\mathbf{x}))^{\mathrm{T}} \mathbf{R}_k^{-1}(\mathbf{y}_k^i - \mathcal{H}(\mathbf{x})),$$

with c_l and c_m being nonlinear or linear equality and inequality constraint functions, the multitude of which is indicated by $\mathcal{E}$ and $\mathcal{I}$. The minimizations of the cost functions/objective functions $\mathcal{J}_k^i$ in respect of these constraints yield the analysis ensemble members $\mathbf{x}_k^{\mathrm{a},i}$ at times k. In the case of a linear observation operator, that is, $\mathcal{H}_k = \mathbf{H}_k$, the cost functions are quadratic and the problems can thus be reformulated. The quadratic structure of the objective functions is then clearly visible and their Hessians and gradients can be directly read off (Janjić *et al.*, 2014):

$$\mathbf{z}_k^{\mathrm{a},i} = \arg\min_{\mathbf{z}} \; \mathcal{J}_k^i(\mathbf{z})$$

$$\text{subject to}$$

$$c_l(\mathbf{x}_k^{\mathrm{b},i} + \mathbf{X}_{kL}^{\mathrm{b}}\mathbf{z}) = 0, \; l \in \mathcal{E} \quad \text{and/or}$$
$$c_m(\mathbf{x}_k^{\mathrm{b},i} + \mathbf{X}_{kL}^{\mathrm{b}}\mathbf{z}) \leq 0, \; m \in \mathcal{I},$$

$$\text{with} \quad \mathcal{J}_k^i(\mathbf{z}) = \frac{1}{2}\mathbf{z}^{\mathrm{T}}\left(\mathbf{I} + (\mathbf{H}_k\mathbf{X}_{kL}^{\mathrm{b}})^{\mathrm{T}}\mathbf{R}_k^{-1}\mathbf{H}_k\mathbf{X}_{kL}^{\mathrm{b}}\right)\mathbf{z}$$
$$+ \left(\mathbf{H}_k\mathbf{x}_k^{\mathrm{b},i} - \mathbf{r}_k^i - \mathbf{y}_k\right)^{\mathrm{T}}\mathbf{R}_k^{-1}\mathbf{H}_k\mathbf{X}_{kL}^{\mathrm{b}}\mathbf{z}. \tag{10}$$

Here, $\mathbf{X}_{k\mathbf{L}}^{\mathrm{b}}$ signifies the square root of the localized background-error covariance matrix that is gained by Cholesky decomposition, that is, $\mathbf{P}_{kL}^{\mathrm{b}} = \mathbf{X}_{kL}^{\mathrm{b}}\mathbf{X}_{kL}^{\mathrm{bT}}$, and a variable transformation $\mathbf{x}_k^{\mathrm{a},i} \rightarrow \mathbf{z}_k^{\mathrm{a},i}$ is made with $\mathbf{x}_k^{\mathrm{a},i} = \mathbf{x}_k^{\mathrm{b},i} + \mathbf{X}_{kL}^{\mathrm{b}}\mathbf{z}_k^{\mathrm{a},i}$. As this transformation is linear, it does not change the respective linear or nonlinear nature of the constraints. Depending on the type of constraints (linear/nonlinear, equality/inequality), different numerical minimization algorithms with different computational demands and either local or global minima can be used for the solution.

Implementation

In this study, constraints on the following properties are taken into consideration:

- the total energy in the truth,
- the moist static energy in the truth,
- the dry mass in the truth, and
- the positivity of convective activity.

For the solution of the numerical optimization problems, we use the MATLAB Optimization Toolbox 2019 algorithms "quadprog" and "fmincon" with their solvers "interior-point-convex" and "interior-point", for linear and nonlinear constraints respectively. The interior-point methods solve the problem sequentially using iteratively updated penalty terms in the objective function for values close to inequality constraint boundaries. This is efficient and requires only a small amount of memory; however, the exact inequality constraint boundary can not be reached. Such an algorithm is thus a suitable choice in terms of efficiency, accuracy, and robustness for the purpose of this study, since the boundary for the only possible inequality constraint, that is, the positivity of convective activity, should not be reached anyway.

We specify the gradients and Hessians of the objective function and nonlinear constraints. Other than this, the default settings of the MATLAB minimization algorithms stay untouched. In particular, there is no change in the tolerance on constraint violations and the values of the stopping criteria. The accuracy of this setup is high and the computational cost affordable in combination with the comparably low-dimensional state space of the skeleton model. Moreover, we neglect any theoretical need for global minimization in the case of nonlinear constraints, since no significant deviations due to this were observed in earlier experiments.

3.3 | Diagnostics

To assess the DA skill, we set up the following three verification metrics. Here, $\boldsymbol{\mu}_c$ and $\boldsymbol{\sigma}_c$ represent the n-dimensional vectors of the climatological mean and the climatological standard deviation, respectively (n being the length of the atmospheric state vector, that is, u, θ, q, and a at all zonal grid points). The n-dimensional truth at time instant k is denoted by $\mathbf{x}_k^t$ and the best estimate/analysis is denoted by $\mathbf{x}_k^a$, which is the same as the ensemble mean $\overline{\mathbf{x}_k^{a,i}}$ for Gaussian distributions. The symbol ./ signifies a pointwise division. In our study, the ensemble mean is adopted as analysis throughout. Any of these verification

metrics can be calculated equally for a subset of the state variables, e.g., especially a single variable.

- The root-mean-square error (RMSE) of the best estimate with respect to the truth scaled by the climatological standard deviations is

$$\text{RMSE} = \sqrt{\frac{1}{n}\left[\left(\mathbf{x}_k^a - \mathbf{x}_k^t\right)./\boldsymbol{\sigma}_c\right]^T\left[\left(\mathbf{x}_k^a - \mathbf{x}_k^t\right)./\boldsymbol{\sigma}_c\right]}. \quad (11)$$

The possible values of the RMSE are $\in [0, \infty]$. It measures the mean error in the state estimate. When RMSE > 1, the state estimate loses its skill, since the error reaches the climatological standard deviation.

- The Pearson correlation coefficient (PCC) (often also referred to as pattern correlation) between the best estimate and the truth is given by

$$\text{PCC} = \frac{\left(\mathbf{x}_k^t - \boldsymbol{\mu}_c\right)^T\left(\mathbf{x}_k^a - \boldsymbol{\mu}_c\right)}{\sqrt{\left(\mathbf{x}_k^t - \boldsymbol{\mu}_c\right)^T\left(\mathbf{x}_k^t - \boldsymbol{\mu}_c\right)}\sqrt{\left(\mathbf{x}_k^a - \boldsymbol{\mu}_c\right)^T\left(\mathbf{x}_k^a - \boldsymbol{\mu}_c\right)}}. \quad (12)$$

The possible values of the PCC are $\in [-1, 1]$. It measures the spatial correlations between the state estimate and the truth. When PCC < 0.5, the state estimate is often said to lose its skill.

- The relative ensemble spread (RES) is

$$\text{RES} = \frac{\sum_n \left[\left|\mathbf{x}_k^{a,i} - \overline{\mathbf{x}_k^{a,i}}\right|./\boldsymbol{\sigma}_c\right]}{\sum_n \left[\left|\mathbf{x}_k^a - \mathbf{x}_k^t\right|./\boldsymbol{\sigma}_c\right]}. \quad (13)$$

The RES is the ratio between the mean absolute deviation of the ensemble members from their mean and the mean absolute deviation of the best estimate from the truth, both of which are scaled by the respective climatological standard deviations. The possible values of the RES are $\in [0, \infty]$. The RES is expected to be around 1 throughout the entire filtering process if the ensemble statistics are a good representation of the uncertainties in the analysis.

These error measures can also be applied to measure the filtering skill and the forecast quality for the different waves by using the wave amplitude distributions calculated from the eigenmode projections as state vectors.

4 | RESULTS

In this section, we present the results of our series of identical twin experiments with the skeleton model. In

Section 4.1, we first investigate the impact of the different constraints in the QPEns on the filtering skill and forecast quality for the MJO and tropical waves in a realistic observational setup. As we find in particular a benefit from the total energy constraint, we compare the QPEns further with the total energy constraint to the stochastic EnKF in different observational setups in Section 4.2. For observations that are less exploited in the EnKF, we show an enhanced positive impact of the constraint. Finally, the QPEns with total energy constraint is tested in a setting with increased non-Gaussianity in the climatological convective activity distribution in Section 4.3. In this case, the new algorithm is able to prevent the filter divergence observed for the EnKF. Thus, we are able to show the robustness and even enhanced importance of our results in the presence of more extreme convective activity caused by, for example, increased sea-surface temperatures.

4.1 | Impact of the different constraints in a realistic observational setup

In our first study, we examine the filtering skill of the QPEns with the different constraints in comparison with the stochastic EnKF. In analogy to Chen and Majda (2016), only the zonal wind and the convective activity are adopted as observational variables, since they are widely used to track the MJO in practice (Wheeler and Hendon, 2004).

Experimental design

In contrast to Chen and Majda (2016), in which the convective activity a is fully observed in space with no additional observational noise, we use sparse observations in space and time, which are generated by adding independent random noise to the truth (see Section 3.1). Specifically, we select a spacing of every fourth grid point, which corresponds to a distance of approximately 2,500 km, and every fourth time step, which means every 6.64 hr, equally for all observations. To obtain statistically significant results, we run 50 experiments with 50 different truths taken from subsequent time spans of model simulation. Each of these experiments is run over a simulation time of 2 years, containing 1 year of DA followed by 1 year of free forecast. In the individual experiments, we compare the stochastic EnKF and the QPEns with each of its possible four constraints (see Section 3.2.2), now using the same truth and observations for each of these five DA setups. The initial ensemble is the same throughout all experiments.

We evaluate the above experiments for the time evolution of the error statistics in the different QPEns configurations with respect to the stochastic EnKF. Therefore, we use the following procedure.

Step 1. Calculate the RMSEs and the pattern correlations over time for the state variables and waves in all experiments and all DA configurations.

Step 2. For each QPEns configuration, take the difference in RMSE and pattern correlation from the EnKF, that is, subtract the EnKF reference value for each of the 50 experiments individually.

Step 3. Calculate the means and 95% confidence intervals of the above differences from the 50 experiments. For the confidence intervals, we assumed Gaussian distributions such that they are equal to the standard deviation multiplied by 1.96.

Results

The results of this study are presented in Figure 3. They show that both the moist static energy constraint ("ME con") and the dry mass constraint ("DM con") mainly impact only the filtering of the potential temperature θ and their positive impact is, moreover, decreasing over time. Furthermore, including only the constraint of positivity in the convective activity ("a pos con") even has a negative influence on the filtering and forecast skill. On the other hand, there is a significant statistical benefit from the total energy constraint ("TE con") during both the DA and the free forecast period for all model variables (u, θ, q, a) and also the waves. One interesting finding is that the positive impact of the constrained filter on the waves (especially the MJO, dry, and moist Rossby waves) becomes more significant during the forecast period. During DA, the EnKF and QPEns produce similar RMSEs, since the same observations are used to constrain the state and the propagation steps are too short for significant error growth. However, during the free forecast, initial small error differences amplify, in particular due to the nonlinear nature of the model. Thus, the benefit of the physically more accurate analysis produced by the QPEns becomes evident.

In summary, the results of this study clearly indicate a benefit when including the total energy constraint. It should be noted that the total energy constraint captures a nonlinear relationship between the model variables that automatically includes the constraint of convective activity to positive values, due to the natural logarithm of a involved (see Equation 6). The results presented here therefore do not imply that the positivity constraint is not important, but rather that using it alone is not a suitable strategy. Instead, it should be combined with some means
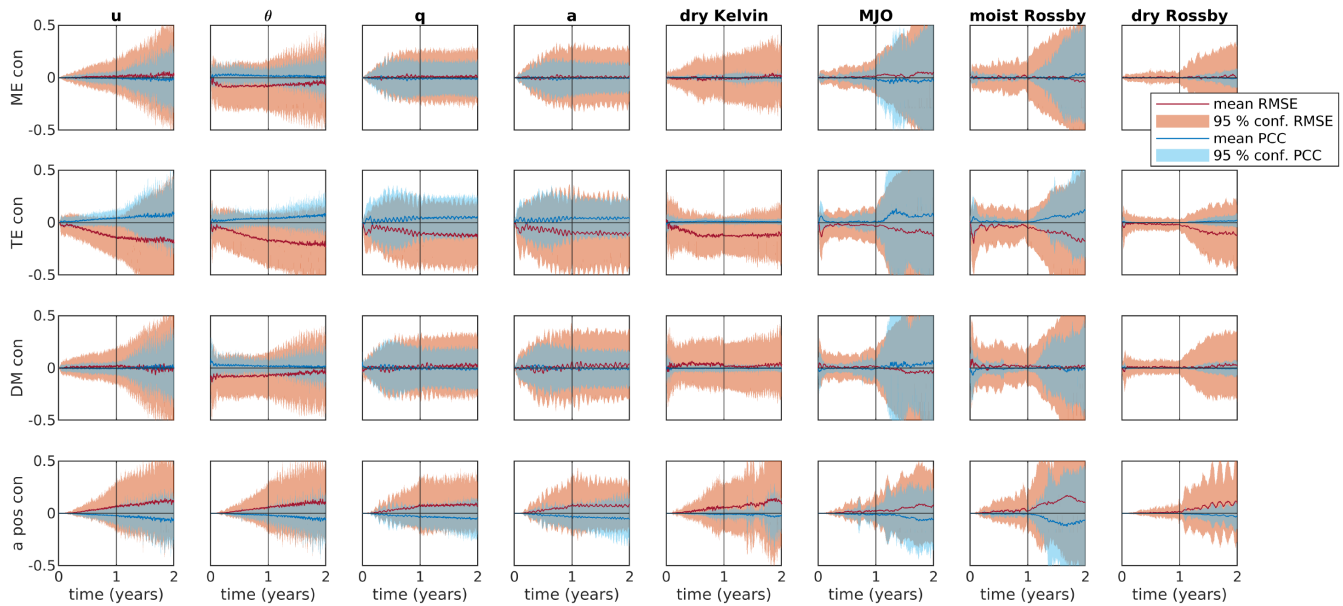
FIGURE 3 Means and 95% Gaussian confidence intervals of the changes in RMSEs and pattern correlations when using the different constraints. These are calculated by subtracting the reference value of the stochastic EnKF from those of the QPEns configurations. The statistics are calculated from 50 experiments with different truths and observations, but using the same truth and observations for each of the setups in each individual experiment. The results show 1 year of DA and 1 year of free forecast, where in the DA observations of u and a are assimilated at every fourth grid point and time step. "ME con", "TE con", "DM con", and "a pos con" stand for the moist static energy constraint, total energy constraint, dry mass constraint, and constraint of positivity in convective activity, respectively [Colour figure can be viewed at wileyonlinelibrary.com]

to respect non-Gaussian moments in the PDF shape of convective activity, as is done via the total energy relationship. Moreover, it should be stated that our results do not mean that every single DA experiment benefits from the total energy constraint. There is a finite probability of unlucky runs, which can be observed when looking at the 50 individual experiments used to calculate the above statistics (not shown here). However, the statistical influence is significantly positive.

4.2 | Comparison of the QPEns with total energy constraint with the stochastic EnKF in different observational setups

Since we have illustrated a significant statistical benefit from the total energy constraint in the experiments in Section 4.1, in this second study we concentrate on a comparison between the QPEns with total energy constraint and the stochastic EnKF when observing the different model variables. With this, we are able to show the dependence of the degree of benefit on the observational setup and draw conclusions about the impact of total energy conservation on the use of specific observations.

Experimental design

In analogy to Section 4.1, we again run 50 experiments for 1 year of DA followed by 1 year of free forecast. Also, we again use different subsequent truth simulations for the individual experiments, but take the same truth and same observations for the QPEns and stochastic EnKF when comparing the different setups within one experiment. In contrast to Section 4.1, these setups are now only comparing the QPEns with total energy constraint and the stochastic EnKF, but with four different sets of observations. These observations are again taken at every fourth grid point and time step, but contain only subsets of one physical variable, that is, u, θ, q, or a, respectively. From the 50 experiments, we again calculate the means and 95% Gaussian confidence intervals for the RMSEs and pattern correlations during the DA and forecast stages for each observational setup in each experiment. However, we do not calculate the differences between the error measures in the QPEns and the EnKF over time, but average the RMSE and pattern correlation over the last half year of DA and the first half year of the free forecast for both filters separately.
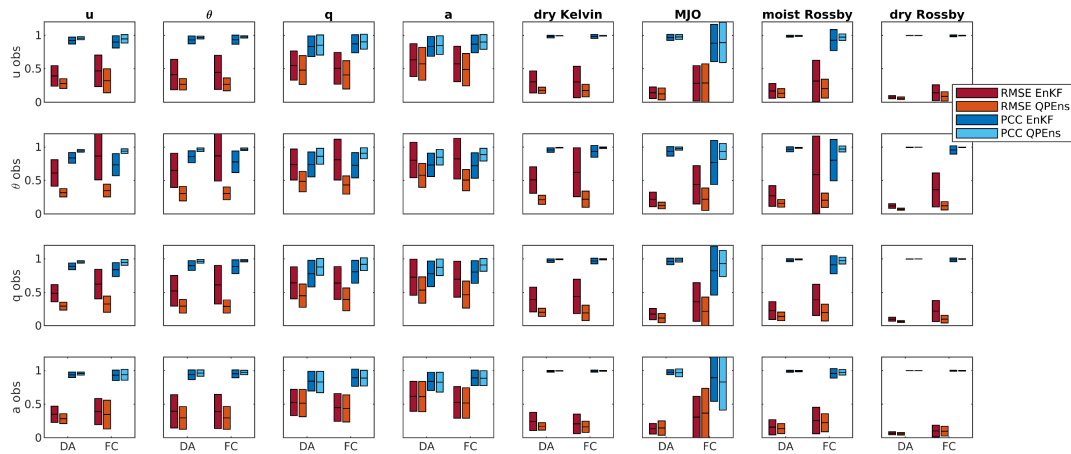
FIGURE 4 Means and 95% confidence intervals of the RMSEs and pattern correlations averaged over the last half year of DA and the first half year of free forecast, comparing the QPEns with total energy constraint and the stochastic EnKF for different variable observations. The statistics are calculated from 50 experiments with different truths and observations, but using the same truth for each of the observational setups and the same observations for the QPEns and the stochastic EnKF in each individual experiment. Each experiment includes 1 year of DA and 1 year of free forecast, with observations at every fourth grid point and time step. The upper confidence interval of the pattern correlation partly reaches values larger than 1, since the statistical distribution is not fully symmetric, but the confidence interval is calculated based on a Gaussian assumption [Colour figure can be viewed at wileyonlinelibrary.com]
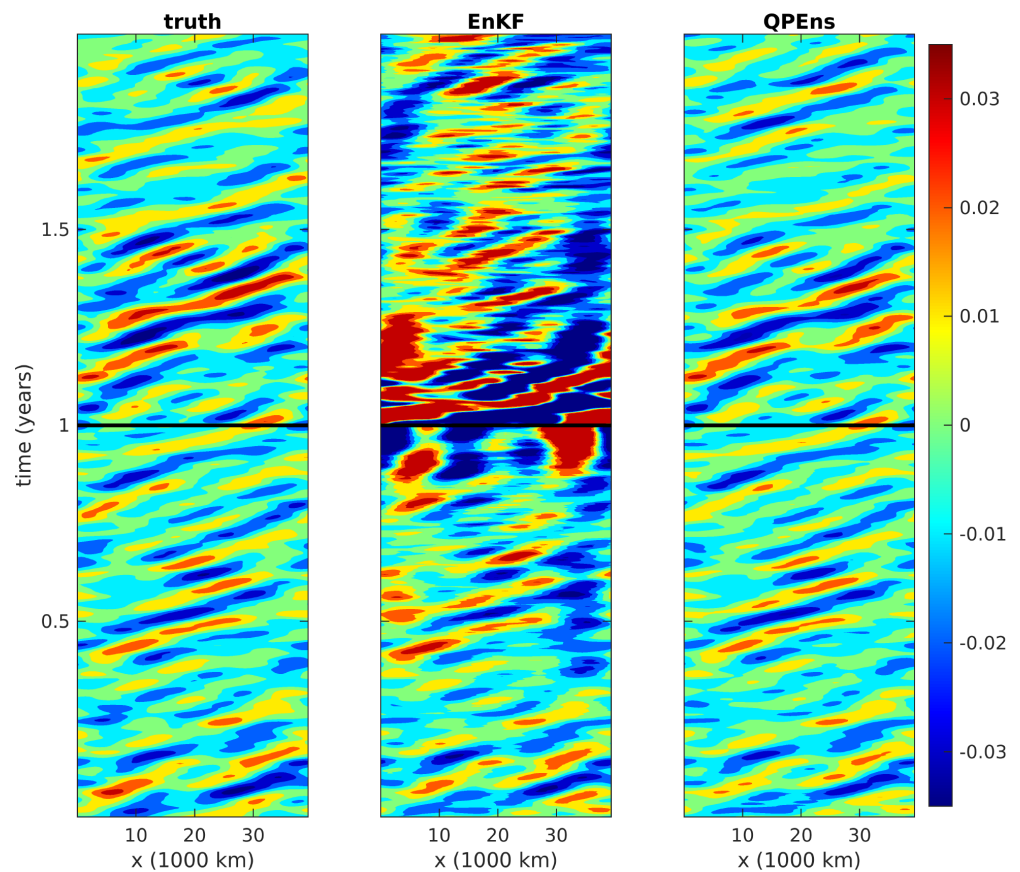
FIGURE 5 Hovmöller diagrams of the MJO wave activity in the truth and ensemble means of the stochastic EnKF and the QPEns with total energy constraint in the experiment with stronger non-Gaussianity in convective activity. Horizontal black lines: transition from filtering to forecast [Colour figure can be viewed at wileyonlinelibrary.com]
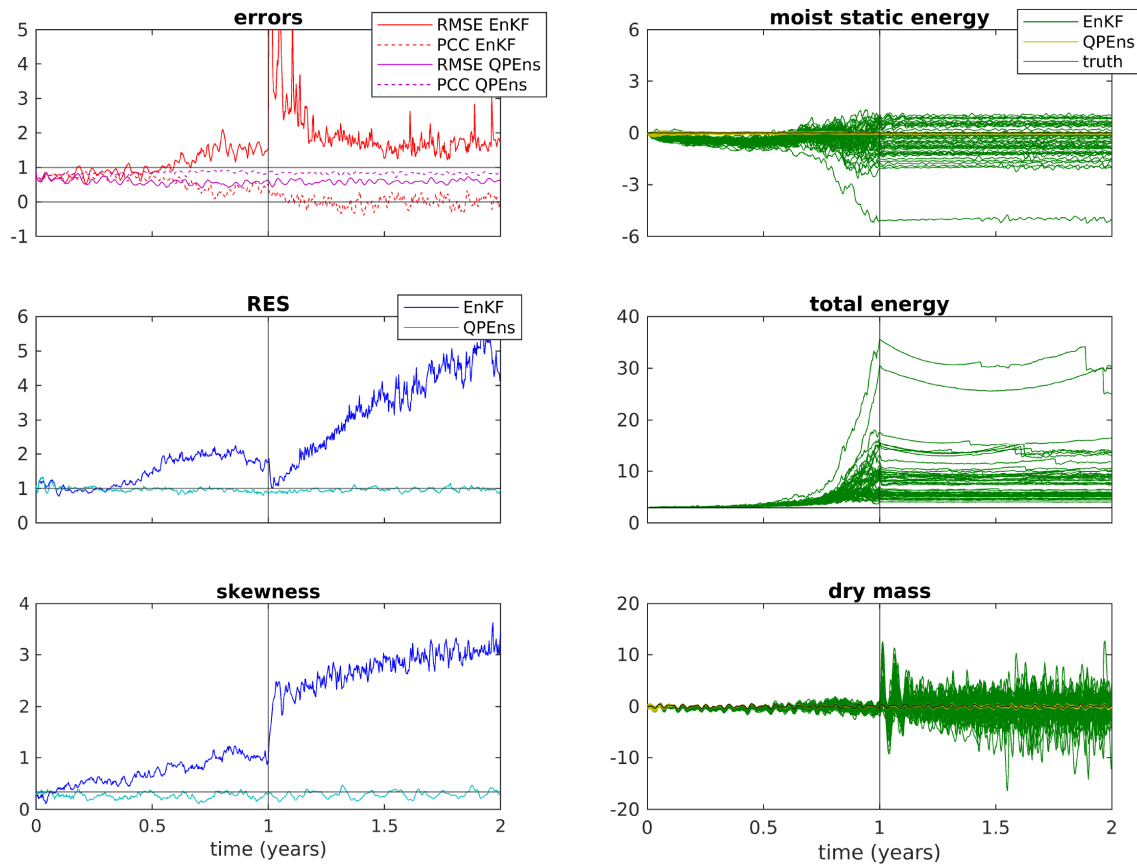
FIGURE 6   RMSE, pattern correlation in the ensemble mean of convective activity, relative ensemble spread (RES), and mean skewness in the ensemble of convective activity, conservation properties and mass in the ensemble during the experiment with strong non-Gaussianity for the stochastic EnKF and the QPEns with total energy constraint. Horizontal black lines: perfect RMSE/pattern correlation and relative ensemble spread, mean climatological skewness in convective activity (0.3406); vertical black lines: transition from filtering to forecast. The yellow line (QPEns result) is beneath the (horizontal) black line (truth) in the middle right plot (total energy), since the total energy is constrained to the truth in the QPEns. The dry mass and moist static energy can become negative (top and bottom right plots), since these are defined as deviations from equilibrium [Colour figure can be viewed at wileyonlinelibrary.com]

Results

Again, the results in Figure 4 clearly indicate the statistically positive impact of the total energy constraint throughout all observational setups and all variables, as well as all waves. The mean error measures are improved on average and the confidence intervals are narrower, indicating a better and more reliable filtering and forecast skill. Moreover, the results show that the stochastic EnKF performs worst with observations of potential temperature, which is also known to be the variable with the smallest MJO signal in the skeleton model (Majda and Stechmann, 2011), and best with observations of u or a as used in Section 4.1. In contrast, the QPEns improves the state estimates most significantly during both the DA and the forecast period if θ or q are observed, such that a filtering skill comparable to filtering with u or a observations is achieved.

Therefore, we arrive at two important conclusions. First, the total energy constraint can possibly help to improve DA when using the current operational set of observations for the MJO, that is, the zonal wind and the convection activity. Second, an even more significant benefit can possibly be seen when trying to include the observations of other variables that are not used so far, especially the potential temperature and the moisture.

4.3 | A stronger non-Gaussian test case for the QPEns with total energy constraint

In order to further explore the potential of the QPEns with total energy constraint, we study its robustness against the stochastic EnKF in an experiment that involves stronger climatological non-Gaussianity in the convective activity.

FIGURE 7 Hovmöller diagrams of the MJO wave activity in the truth and ensemble means of the stochastic EnKF and the QPEns with total energy constraint in the experiment with standard non-Gaussianity (parameters as in Table 1) in convective activity. Horizontal black lines: transition from filtering to forecast [Colour figure can be viewed at wileyonlinelibrary.com]
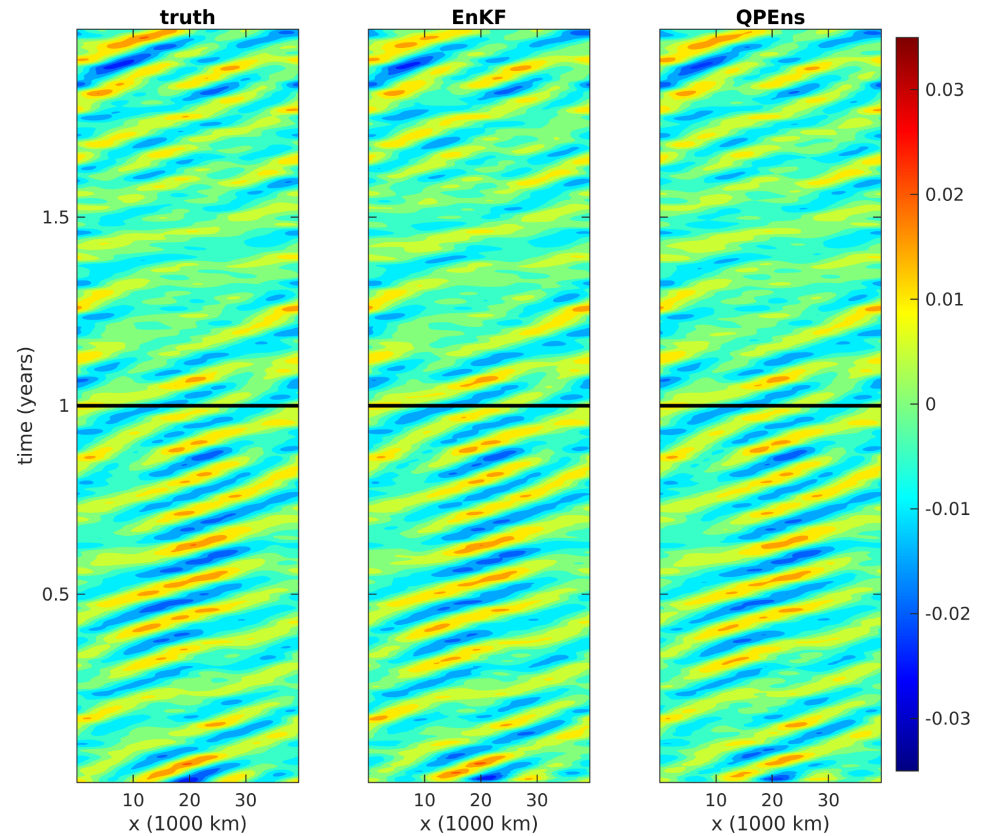
Specifically, we alter the default model setup by increasing the background warm-pool parameter from 0.6 to 0.75 (cf. Table 1). This is slightly different from the current climate, but it is a potential expected situation in the presence of the acceleration of climate change. Such a modification increases the possibility for smaller values of convective activity and enhances its climatological skewness from a mean value of 0.3251 to 0.3406. Consequently, the error due to the rudimentary treatment of non-Gaussianity in our stochastic EnKF, that is, by setting negative values of convective activity in the analysis ensemble back closely above zero, is amplified. The ability of the QPEns to handle non-Gaussian situations can therefore be tested more easily.

Experimental design

In this experiment, we again adopt the observational setup from Section 4.1, that is, we observe u and a with sparse observations at every fourth grid point and time step. Since the adjustment of the model parameters alters the model climatology, we retune the localization and inflation parameters as described in the Appendix. The new localization radius and the new constant inflation factor

are both found to be 1; the adaptive inflation stays unaffected.

The experiment is run with 1 year of DA and 1 year of free forecast for both the QPEns with total energy constraint and the stochastic EnKF, using the same truth and the same set of observations. For comparison, the experiment is repeated with the standard model settings (i.e., the background warm-pool parameter being 0.6) and the original localization and inflation parameters as used in Sections 4.1 and 4.2. In contrast to Section 4.1, this time only one single run is performed for each configuration, since this experiment is not intended to give statistical insight into the benefits of the total energy constraint, but to reveal its underlying mechanism.

Results

The results are presented in Figures 5,6 and Figures 7,8 for the stronger non-Gaussian and the standard setup, respectively. Figures 5 and 7 show the Hovmöller diagrams of the true MJO and the analysis ensemble means, that is, best estimates, of the stochastic EnKF and the QPEns with total energy constraint. Figures 6 and 8 show the RMSEs and pattern correlations, the relative ensemble spread, the
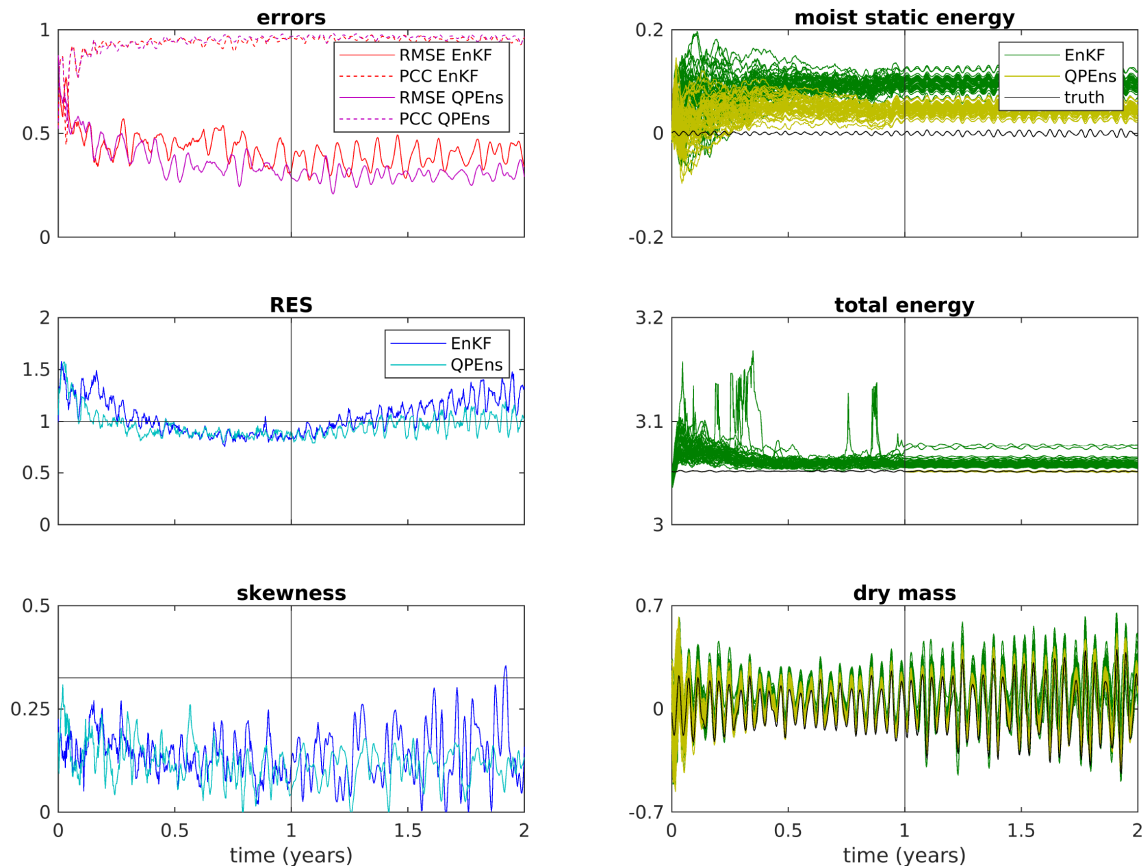
FIGURE 8 RMSE, pattern correlation in the ensemble mean of convective activity, relative ensemble spread (RES), and mean skewness in the ensemble of convective activity, conservation properties and mass in the ensemble during the experiment with standard non-Gaussianity (parameters as in Table 1) for the stochastic EnKF and the QPEns with total energy constraint. Horizontal black lines: perfect RMSE/pattern correlation and relative ensemble spread, mean climatological skewness in convective activity (0.3251); vertical black lines: transition from filtering to forecast. The yellow line (QPEns result) is beneath the (horizontal) black line (truth) in the middle right plot (total energy), since the total energy is constrained to the truth in the QPEns. The dry mass and moist static energy can become negative (top and bottom right plots), since these are defined as deviations from equilibrium [Colour figure can be viewed at wileyonlinelibrary.com]

mean skewness in the convective activity ensemble distribution, and the evolution of the conservation properties in the ensemble over time.

According to the error measures (top left in Figures 6 and 8), it is obvious that the QPEns with total energy constraint performs better in both experiments. However, in the experiment with the standard setup (Figures 7 and 8), the stochastic EnKF also converges (top left in Figure 8) and the wave amplitude pattern is recovered overall (middle column in Figure 7). Despite being worse than the QPEns, the ensemble spread and skewness in the EnKF stay within reasonable ranges (middle and bottom left in Figure 8). The moist static energy and dry mass fluctuate in comparable amounts for both filtering algorithms, while the total energy is conserved in the constrained QPEns and fluctuates only in the EnKF (right plots in Figure 8).

The situation becomes very different in the stronger non-Gaussian test case (Figures 5 and 6). Here, the stochastic EnKF diverges (top left in Figure 6), which

leads to a strong increase in the MJO amplitude, especially towards the end of the filtering period and the beginning of the forecast before the MJO wave activity degrades (middle column in Figure 5). This divergence is connected to an unrealistic increase in relative ensemble spread and skewness (middle and bottom left in Figure 6) and the conservation properties are also significantly affected (right plots in Figure 6). In contrast, the QPEns with total energy constraint keeps a filtering and forecast quality comparable with the one in the standard settings experiment.

The mechanism behind the divergence of the stochastic EnKF is illustrated further in Figure 9. It shows the analysis ensembles, analysis ensemble means, and truth of the convective activity for both DA algorithms at different time instants in the stronger non-Gaussian test experiment. In particular, it can be observed (in the upper row) that small convective activity values, which arise from the setting of negative values in the EnKF back to values closely above zero, arise at more and more grid points and
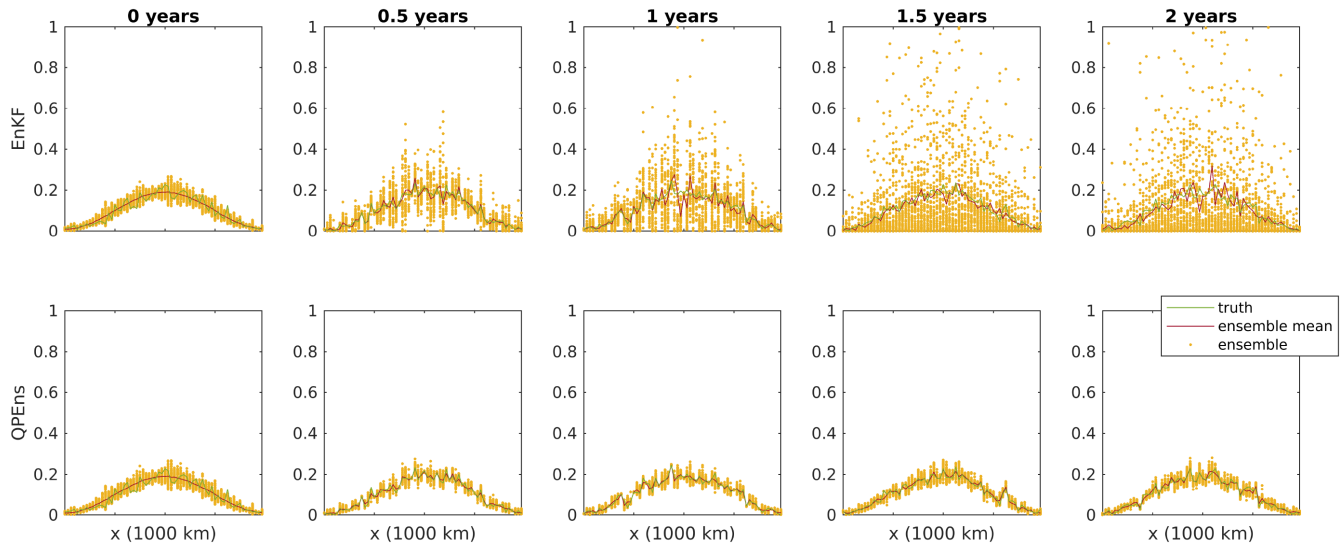
FIGURE 9 Truth, ensemble mean/analysis, and individual ensemble members in the stochastic EnKF and the QPEns with total energy constraint at different time instants during the experiment with stronger non-Gaussianity in the climatological PDF of convective activity [Colour figure can be viewed at wileyonlinelibrary.com]

ensemble members with increasing filtering time. This accumulation is specifically sustained, since small convective activity hardly evolves in the model due to the nonlinear oscillator relationship in Equation 2d. The consequential misestimation of the PDF shape is a self-amplifying effect, since the related decrease in Gaussianity makes the problem less and less suitable for a Gaussianity-based filter algorithm and thus leads to further errors in the estimated PDF. The result is an unreasonable ensemble inflation.

It is important to mention that a divergence comparable with that of the stochastic EnKF is found in the stronger non-Gaussian test case when using the QPEns with any other constraints, that is, the moist static energy constraint, positivity constraint, mass constraint, or any combinations of these. The corresponding experiments are not shown here, but highlight the fact that it is specifically the total energy constraint that can prevent filter divergence in stronger non-Gaussian situations.

5 | DISCUSSION AND CONCLUSIONS

In this article, we test a QPEns with different physical constraints against its underlying stochastic EnKF in identical twin experiments with the skeleton model. In a realistic observational setup, in particular, the constraint to the total energy in the truth, which is conserved throughout the model's nature run, has a significant positive impact on the overall filtering and forecast skill. Interestingly, even if for some waves the statistical error from the truth remains approximately the same as in the stochastic EnKF during

the DA stage, their prediction benefits remarkably from the improved initial state. The physically more accurate analysis mitigates the error growth in the free forecast, which is determined particularly by the nonlinear model dynamics. Moreover, we investigate further the QPEns with total energy constraint in combination with different observational setups and show that it has the potential to strongly increase the benefit from observations that are less useful in the stochastic EnKF. Lastly, we use a test case with increased climatological non-Gaussianity in convective activity to demonstrate that, under certain circumstances, the QPEns can even prevent filter divergence that occurs in the stochastic EnKF. The nonlinear total energy relationship, which automatically includes the positivity of convective activity, is particularly useful to reproduce non-Gaussian moments in the PDF estimates.

Our findings highlight the potential of the QPEns to outperform DA algorithms that are based on Gaussian assumptions, especially in non-Gaussian situations. This is in line with earlier works on the QPEns as by Janjić *et al.* (2014), Zeng *et al.* (2017), and Ruckstuhl and Janjić (2018). Nevertheless, the interpretation of the results here can go beyond that, since the skeleton model is considered to represent a realistic approximation of some main large-scale MJO dynamics. Knowledge about the relevance of the total energy constraint for DA in this idealized model can thus serve as a first hint for possible important real-world large-scale relationships. At least, it underlines the need to go beyond Gaussianity-based filtering when aiming for an improved DA for intraseasonal forecasts in the Tropics.

Nevertheless, it should be noted that the skeleton model in the version used for our study is much less

complex than fully developed subseasonal to seasonal numerical weather prediction (NWP) models. The skeleton model, by its design, focuses on the characterization of planetary-scale dynamics, whereas mesoscale details are highly simplified or parameterized. Thus, it is only suitable for simulating large-scale MJO features. Moreover, it is worthwhile to note that our study is based on identical twin experiments instead of comparisons with real-world observations. The predictability in this setup is unrealistically high for small errors in the initial conditions, which is directly understandable from the deterministic nature of the model version we use. The study thus cannot be compared with results from other MJO predictability studies that investigate operational forecasts based on GCMs, such as those by Vitart (2017) or Wang *et al.* (2019). Rather, all results presented should be interpreted as relative benefits of the QPEns with respect to its underlying stochastic EnKF in our simplified setup and not as absolute measures for the real-world MJO prediction skill.

In line with this, future work includes testing the total energy constrained QPEns with more sophisticated setups in order to verify the significance of our results for advances in tropical DA. A natural next step is to proceed with the available extended versions of the skeleton model, i.e., with additional stochasticity or a refined meridional or vertical structure (see Thual *et al.*, 2014; 2015; Thual and Majda, 2016). However, besides this, the proposed DA algorithm should also be tested on other MJO models, for example, the MJO moisture mode theory (Sobel and Maloney, 2012; 2013) and the model in Stechmann and Hottovy (2017). Using a more detailed model, the total energy constrained QPEns could then also be tested with real observational data projected on the model basis as shown for the stochastic skeleton model in Ogrosky and Stechmann (2015). However, for such extended investigations it will also be necessary to refine the DA system, for example, for the inclusion of a model error estimation. The ultimate goal is to understand the role of the constrained DA for MJO forecasts by testing it in a series of setups with increasing degrees of freedom, that is, approaching a full NWP system.

In addition, another important direction is to explore the operational applicability of the QPEns algorithm. The main challenges here are the increased computational demand and the necessity to estimate the constraint's value without knowing the true atmospheric state. With respect to the computational speed, there are two aspects to consider: first the ensemble size, since the minimization is done for each ensemble member, and second the number and type, that is, linearity or nonlinearity, of the constraints, which determine the complexity of the minimizations. In practice, one would not use standard Matlab routines, but a minimization algorithm tailored

to the high-dimensional system. Such algorithms are currently under development. For example, Janjić *et al.* (2021) reduced the computational cost of a QPEns to a third of its original cost by exploiting the fact that the positivity constraint and mass constraint are disjoint, which allows for the use of projection algorithms. They showed how to modify the conjugate gradient minimization to achieve faster, but accurate results. Moreover, Ruckstuhl *et al.* (2021) found that it is possible to improve the computational speed of a QPEns with the use of neural networks and Ruckstuhl and Janjić (2018) showed that a QPEns can outperform the EnKF even for small ensemble sizes. In short, the QPEns algorithm could be used in operational practice with careful consideration of the ensemble size and minimization algorithm. For the total energy constrained QPEns in particular, a first reduction of computational expense without significant accuracy loss can be achieved by a linearization of the nonlinear total energy relationship around the current background value in each update step (see Gleiter, 2021). This moreover enables an even less expensive implementation as pseudo-observation, which in addition allows us to quantify the uncertainty in the constraint.

Finally, the current study can also be extended beyond its focus on the MJO and related tropical waves, specifically to other MJO-related weather and climate phenomena. One particular interesting topic is to investigate the total energy constrained QPEns when applied to coupled MJO–ENSO models, which were developed recently by Thual *et al.* (2018) and Yang *et al.* (2021). These models involve an ocean contribution in addition to a modification of the skeleton model as the atmospheric component. A coupling of a robust DA routine for the skeleton model, which is based on our results, can thus serve to examine the influence of the MJO DA on forecasting ENSO.

ACKNOWLEDGEMENTS

aspects and technical help. Open Access funding enabled and organized by Projekt DEAL.

AUTHOR CONTRIBUTIONS

Tabea Gleiter: investigation; methodology; visualization; writing-original draft; writing-review & editing. **Tijana Janjić:** conceptualization; funding acquisition; methodology; project administration; resources; supervision; validation; writing-review & editing. **Nan Chen:** funding acquisition; methodology; resources; writing – original draft; writing – review and editing.

ORCID

Tabea Gleiter https://orcid.org/0000-0001-7807-3048
Tijana Janjić https://orcid.org/0000-0002-8837-0879

REFERENCES

Žagar, N., Andersson, E. and Fisher, M. (2005) Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors. *Quarterly Journal of the Royal Meteorological Society*, 131, 987–1011.

Žagar, N., Gustafsson, N. and Källén, E. (2004) Variational data assimilation in the Tropics: the impact of a background-error constraint. *Quarterly Journal of the Royal Meteorological Society*, 130, 103–125.

Bocquet, M., Pires, C.A. and Wu, L. (2010) Beyond Gaussian statistical modeling in geophysical data assimilation. *Monthly Weather Review*, 138, 2997–3023.

Burgers, G., van Leeuwen, P.J. and Evensen, G. (1998) Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, 126, 1719–1724.

Castanheira, J.M. and Marques, C.A.F. (2021) The equatorial wave skeleton of the Madden–Julian oscillation. *Quarterly Journal of the Royal Meteorological Society*, 147, 3778–3788.

Chen, N. and Majda, A.J. (2016) Filtering the stochastic skeleton model for the Madden–Julian oscillation. *Monthly Weather Review*, 144, 501–527.

Chen, N., Majda, A.J. and Giannakis, D. (2014) Predicting the cloud patterns of the Madden–Julian oscillation through a low-order nonlinear stochastic model. *Geophysical Research Letters*, 41, 5612–5619.

Deng, Q., Khouider, B. and Majda, A.J. (2015) The MJO in a coarse-resolution GCM with a stochastic multicloud parameterization. *Journal of the Atmospheric Sciences*, 72, 55–74.

Evensen, G. (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99, 10143–10162.

Evensen, G. (2009) *Data Assimilation: The Ensemble Kalman Filter*. Berlin Heidelberg: Springer-Verlag.

Gaspari, G. and Cohn, S.E. (1999) Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125, 723–757.

Gill, A.E. (1980) Some simple solutions for heat-induced tropical circulation. *Quarterly Journal of the Royal Meteorological Society*, 106, 447–462.

Gleiter, T. (2021). Improving data assimilation for MJO prediction based on experiments with the skeleton model for tropical intraseasonal variability. Master's thesis, München: Ludwig-Maximilians-Universität.

Houtekamer, P.L. and Zhang, F. (2016) Review of the ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 144, 4489–4532.

Janjić, T., McLaughlin, D., Cohn, S.E. and Verlaan, M. (2014) Conservation of mass and preservation of positivity with ensemble-type Kalman filter algorithms. *Monthly Weather Review*, 142, 755–773.

Janjić, T., Ruckstuhl, Y.M. and Toint, P.L. (2021) A data assimilation algorithm for predicting rain. *Quarterly Journal of the Royal Meteorological Society*, 147, 1949–1963.

Kalnay, E. (2003) *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge: Cambridge University Press.

Khouider, B., Majda, A.J. and Stechmann, S.N. (2013) Climate science in the Tropics: waves, vortices and PDES. *Nonlinearity*, 26, R1–R68.

Khouider, B., St-Cyr, A., Majda, A.J. and Tribbia, J. (2011) The MJO and convectively coupled waves in a coarse-resolution GCM with a simple multicloud parameterization. *Journal of the Atmospheric Sciences*, 68, 240–264.

Kiladis, G.N., Wheeler, M.C., Haertel, P.T., Straub, K.H. and Roundy, P.E. (2009) Convectively coupled equatorial waves. *Reviews of Geophysics*, 47, RG2003.

Kondrashov, D., Chekroun, M., Robertson, A. and Ghil, M. (2013) Low-order stochastic model and "past-noise forecasting" of the Madden–Julian oscillation. *Geophysical Research Letters*, 40, 5305–5310.

Law, K., Stuart, A. and Zygalakis, K. (2015) *Data Assimilation: A Mathematical Introduction*. Switzerland: Springer International Publishing.

van Leeuwen, P.J. (1999) Comment on "data assimilation using an ensemble Kalman filter technique". *Monthly Weather Review*, 127, 1374–1377.

Majda, A.J. and Stechmann, S.N. (2009) The skeleton of tropical intraseasonal oscillations. *Proceedings of the National Academy of Sciences*, 106, 8417–8422.

Majda, A.J. and Stechmann, S.N. (2011) Nonlinear dynamics and regional variations in the MJO skeleton. *Journal of the Atmospheric Sciences*, 68, 3053–3071.

Majda, A.J., Stechmann, S.N., Chen, S., Ogrosky, H.R. and Thual, S. (2019) *Tropical Intraseasonal Variability and the Stochastic Skeleton Method*. Switzerland: Springer Nature.

Matsuno, T. (1966) Quasi-geostrophic motions in the equatorial area. *Journal of the Meteorological Society of Japan*, 44, 25–42.

Ogrosky, H.R. and Stechmann, S.N. (2015) The MJO skeleton model with observation-based background state and forcing. *Quarterly Journal of the Royal Meteorological Society*, 141, 2654–2669.

Penny, S.G. and Miyoshi, T. (2016) A local particle filter for high-dimensional geophysical systems. *Nonlinear Processes in Geophysics*, 23, 391–405.

Poterjoy, J. (2016) A localized particle filter for high-dimensional nonlinear systems. *Monthly Weather Review*, 144, 59–76.

Ruckstuhl, Y., Janjić, T. and Rasp, S. (2021) Training a convolutional neural network to conserve mass in data assimilation. *Nonlinear Processes in Geophysics*, 28, 111–119.

Ruckstuhl, Y.M. and Janjić, T. (2018) Parameter and state estimation with ensemble Kalman filter based algorithms for convective-scale applications. *Quarterly Journal of the Royal Meteorological Society*, 144, 826–841.

Seo, K.-H., Wang, W., Gottschalck, J., Zhang, Q., Schemm, J.-K.E., Higgins, W.R. and Kumar, A. (2009) Evaluation of MJO forecast skill from several statistical and dynamical forecast models. *Journal of Climate*, 22, 2372–2388.

Sobel, A. and Maloney, E. (2012) An idealized semi-empirical framework for modeling the Madden–Julian oscillation. *Journal of the Atmospheric Sciences*, 69, 1691–1705.

Sobel, A. and Maloney, E. (2013) Moisture modes and the eastward propagation of the MJO. *Journal of Atmospheric Sciences*, 70, 187–192.

Stachnik, J.P., Waliser, D.E., Majda, A.J., Stechmann, S.N. and Thual, S. (2015) Evaluating MJO event initiation and decay in the skeleton model using an RMM-like index. *Journal of Geophysical Research: Atmospheres*, 120, 486–508.

Stechmann, S.N. and Hottovy, S. (2017) Unified spectrum of tropical rainfall and waves in a simple stochastic model. *Geophysical Research Letters*, 44, 10–713.

Stechmann, S.N. and Majda, A.J. (2015) Identifying the skeleton of the Madden–Julian oscillation in observational data. *Monthly Weather Review*, 143, 395–416.

Thual, S. and Majda, A.J. (2016) A suite of skeleton models for the MJO with refined vertical structure. *Mathematics of Climate and Weather Forecasting*, 1, 70–95.

Thual, S., Majda, A.J. and Chen, N. (2018) A tropical stochastic skeleton model for the MJO, El Niño, and dynamic Walker circulation: a simplified GCM. *Journal of Climate*, 31, 9261–9282.

Thual, S., Majda, A.J. and Stechmann, S.N. (2014) A stochastic skeleton model for the MJO. *Journal of the Atmospheric Sciences*, 71, 697–715.

Thual, S., Majda, A.J. and Stechmann, S.N. (2015) Asymmetric intraseasonal events in the stochastic skeleton MJO model with seasonal cycle. *Climate Dynamics*, 45, 603–618.

Vitart, F. (2017) Madden–Julian oscillation prediction and teleconnections in the S2S database. *Quarterly Journal of the Royal Meteorological Society*, 143, 2210–2220.

Vitart, F. and Molteni, F. (2010) Simulation of the Madden–Julian oscillation and its teleconnections in the ECMWF forecast system. *Quarterly Journal of the Royal Meteorological Society*, 136, 842–855.

Wang, S., Sobel, A., Tippett, M. and Vitart, F. (2019) Prediction and predictability of tropical intraseasonal convection: seasonal dependence and the maritime continent prediction barrier. *Climate Dynamics*, 52, 6015–6031.

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

Yang, Q., Majda, A.J. and Chen, N. (2021) ENSO diversity in a tropical stochastic skeleton model for the MJO, El Niño, and dynamic Walker circulation. *Journal of Climate*, 34, 3481–3501.

Yang, Q., Majda, A.J. and Moncrieff, M.W. (2019) Upscale impact of mesoscale convective systems and its parameterization in an idealized GCM for an MJO analog above the equator. *Journal of the Atmospheric Sciences*, 76, 865–892.

Žagar, N., Blaauw, M., Jesenko, B. and Magnusson, L. (2016) Diagnosing model performance in the Tropics. *ECMWF Newsletter*, 147, 26–33.

Zeng, Y., Janjić, T., Ruckstuhl, Y.M. and Verlaan, M. (2017) Ensemble-type Kalman filter algorithm conserving mass, total energy and enstrophy. *Quarterly Journal of the Royal Meteorological Society*, 143, 2902–2914.

Zeng, Y. and Janjić, T. (2016) Study of conservation laws with the local ensemble transform Kalman filter. *Quarterly Journal of the Royal Meteorological Society*, 142, 2359–2372.

Zhang, C. (2005) Madden–Julian oscillation. *Reviews of Geophysics*, 43, RG2003.

Zhang, C., Gottschalck, J., Maloney, E.D., Moncrieff, M.W., Vitart, F., Waliser, D.E., Wang, B. and Wheeler, M.C. (2013) Cracking the MJO nut. *Geophysical Research Letters*, 40, 1223–1230.

APPENDIX A. INFLATION AND LOCALIZATION

The localization aims to alleviate erroneous correlations in the background-error covariance matrix $\mathbf{P}_k^b$ that is used to calculate the Kalman gain. We use a covariance localization, in which we Schur-multiply $\mathbf{P}_k^b$ with a positive semidefinite matrix containing entries $\in [0, 1]$ before each DA update. This localization matrix C is constructed as the sum of climatology-based and distance-depending damping contributions. The former approximately represents the typical bivariate correlation strengths and is constructed by the following steps from the covariance matrix of 12,200 (100 years, every three days) climatological samples starting after 10 years of model spin-up:

- transformation of all matrix entries to their absolute values;
- eigenvalue decomposition of the matrix and construction of a similar positive semidefinite matrix from all eigenvectors to eigenvalues ≥ 0;
- adding the absolute value of the smallest negative matrix entry to all entries;
- scaling all entries with the standard deviations of the variables concerned.

The distance-dependent damping for the second contribution to the localization matrix is calculated according to a fifth-order compactly supported Gaspari–Cohn function (see Gaspari and Cohn, 1999). The ratio of its
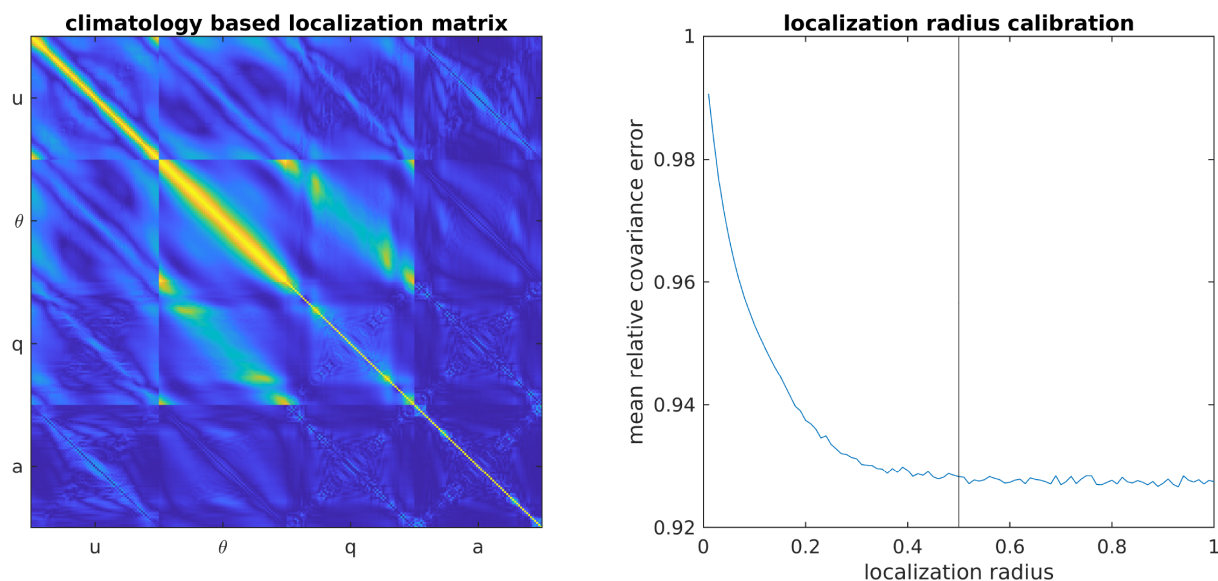
FIGURE A1 Left: climatology-based localization matrix calculated from 12,200 (100 years, every three days) samples from a nature run starting after 10 years of model spin-up; right: calibration of the distance-dependent localization for 50 ensemble members, vertical line at the adopted value of 0.5. The y-axis shows the mean relative error in 1,000 localized covariance matrices, each calculated from a different 50-member subset of the climatological 12,200 member ensemble, with respect to the full climatological covariance matrix. The error calculation is performed for different localization radii; the climatology-based localization is applied throughout [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE A1 Manually tuned constant inflation factors (additional to the adaptive inflation) for the observational setups used in this article

Observations	Constant inflation factor
u every fourth grid point and time step	1.0001
θ every fourth grid point and time step	1.0008
q every fourth grid point and time step	1.0005
a every fourth grid point and time step	0.9998
u and a every fourth grid point and time step	1.0001

corresponding Gaussian's standard deviation to the whole zonal model domain (i.e., 40,000 km) is further referred to as the localization radius. That is, a localization radius of 0.5 means that covariances of variables with a distance of 20,000 km are approximately dampened by a factor $1/e$. To determine the best localization radius for our ensemble size of 50 members, we test a number of different localization radii in combination with the climatology-based localization. We therefore apply the localization on covariance matrices calculated from 1000 different subsets of 50 members from the climatological 12,200 member ensemble. Thereafter, we calculated the mean relative error in the localized covariance matrix

entries with respect to the full climatological covariance matrix. From this, we derive a suitable localization radius of 0.5 (cf. Figure A1).

The inflation aims to increase the underestimated error covariance artificially and is applied as a second measure in combination with localization. To set up an inflation routine, in a first step, the localization is assumed to be perfect. Then, an adaptive inflation factor is calculated, which rescales the ensemble such that the trace of its analysis-error covariance $\mathbf{P}^a_{k\text{ens}}$ meets the theoretically expected variance in the analysis-error covariance $\mathbf{P}^a_{k\text{th}}$ that is retrieved from an exact (not ensemble-based) Kalman filter update with the localized background-error

covariance. This factor (β_{k+1}, since applied to the background in the next assimilation step) is calculated as

$$
\begin{aligned}
\beta_{k+1} &= \sqrt{\frac{tr(\mathbf{P}^a_{k\text{th}})}{tr(\mathbf{P}^a_{k\text{ens}})}} \\
&= \sqrt{\frac{tr\left(C \circ \mathbf{P}^b_k - [C \circ \mathbf{P}^b_k]\mathbf{H}^T_k\left(\mathbf{H}_k[C \circ \mathbf{P}^b_k]\mathbf{H}^T_k + \mathbf{R}_k\right)^{-1} \mathbf{H}_k[C \circ \mathbf{P}^b_k]\right)}{tr\left(\mathbf{P}^a_k\right)}}.
\end{aligned}
\tag{A1}
$$

In our setup, we observe a large benefit from this adaptive inflation routine; however, one should note that this procedure is only applicable in a simplified model setting due to its computational demand, and also is not robust with respect to changes in the localization technique. Such remaining errors, that is, imperfections in the localization and other neglected error sources that contribute to the ensemble collapse, are likely the reason that

we find a further improvement in the maintenance of a reasonable ensemble spread when we apply an additional heuristically tuned constant inflation factor. This factor is dependent on the observational setup and its values for the experiments in this article are listed in Table A1. All inflation is applied to the background ensemble, such that the destruction of potential nonlinear constraints in the analysis is avoided and the occurrence of negative convective activity values in the analysis ensemble is reduced.