Inclusion of Linearized Moist Physics in NASA's Goddard Earth Observing System Data Assimilation Tools

DANIEL HOLDAWAY

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, and Goddard Earth Sciences Technology and Research, Universities Space Research Association, Columbia, Maryland

RONALD ERRICO

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, and Goddard Earth Science Technology and Research, Morgan State University, Baltimore, Maryland

RONALD GELARO

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

JONG G. KIM

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, and Science Systems and Applications Inc., Lanham, Maryland

(Manuscript received 11 June 2013, in final form 8 August 2013)

ABSTRACT

Inclusion of moist physics in the linearized version of a weather forecast model is beneficial in terms of variational data assimilation. Further, it improves the capability of important tools, such as adjoint-based observation impacts and sensitivity studies. A linearized version of the relaxed Arakawa–Schubert (RAS) convection scheme has been developed and tested in NASA's Goddard Earth Observing System data assimilation tools. A previous study of the RAS scheme showed it to exhibit reasonable linearity and stability. This motivates the development of a linearization of a near-exact version of the RAS scheme. Linearized large-scale condensation is included through simple conversion of supersaturation into precipitation. The linearization of moist physics is validated against the full nonlinear model for 6- and 24-h intervals, relevant to variational data assimilation and observation impacts, respectively. For a small number of profiles, sudden large growth in the perturbation trajectory is encountered. Efficient filtering of these profiles is achieved by diagnosis of steep gradients in a reduced version of the operator of the tangent linear model. With filtering turned on, the inclusion of linearized moist physics increases the correlation between the nonlinear perturbation trajectory and the linear approximation of the perturbation trajectory. A month-long observation impact experiment is performed and the effect of including moist physics on the impacts is discussed. Impacts from moist-sensitive instruments and channels are increased. The effect of including moist physics is examined for adjoint sensitivity studies. A case study examining an intensifying Northern Hemisphere Atlantic storm is presented. The results show a significant sensitivity with respect to moisture.

1. Introduction

The techniques and methods used in incremental variational data assimilation are largely based on the assumption that the underlying behavior of the system is linear or close to linear (Courtier et al. 1994). To apply the techniques, such as minimization of a cost function in four-dimensional variational data assimilation (4DVAR), a linearized version of the model is used.

The linear model can be thought of as providing an approximation to the trajectory of nonlinear perturbations. For example, it may be used to estimate how much an initial error (perturbation from the truth) would grow over time. If the full model is in fact linear, then the

Corresponding author address: Dan Holdaway, Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Code 610.1, Greenbelt, MD 20771.
E-mail: dan.holdaway@nasa.gov

DOI: 10.1175/MWR-D-13-00193.1

© 2014 American Meteorological Society

approximation would give the exact nonlinear perturbation trajectory.

In practice, how well a linearized model will perform for data assimilation applications depends on three main issues. First, how much detail is lacking from the linear model, for example physics that is not included; second, how close to linear the components of the nonlinear model are; and, third, how many approximations are made in the numerics—for example, use of multiple outer loops without higher than first-order linearization techniques (Trémolet 2008).

For the large-scale dynamics of the atmosphere a linearization produces a very good approximation of the quasi-linear perturbations. Indeed, for a primitive equation model, with only very simple physical parameterizations, a correlation of 0.9 or more between the linear and nonlinear perturbation trajectories is possible, even for 72-h integrations (Errico et al. 1993). Atmospheric motions that are subgrid scale, sometimes referred to as ''the physics,'' include processes such as turbulence, convection, precipitation, gravity wave drag, and radiation. The physical parameterizations that are used to model these processes can be very nonlinear and contain discontinuities. Including these processes in the linear model in a way that produces accurate representation of the actual perturbation trajectory requires much care.

The National Aeronautics and Space Administration's (NASA) Global Modeling and Assimilation Office (GMAO) is currently developing a linearization of the Goddard Earth Observing System (GEOS-5) atmospheric global circulation model (AGCM). The linear model is based on the new cubed-sphere dynamical core (Putman 2007) and replaces the old latitude–longitude-based linearized model. So far, the adjoint and tangent linear versions of the dynamical core have been developed, along with a simple vertical diffusion scheme and boundary layer. The new linearized model provides an essential upgrade to the adjoint-based observation impact tool, used for daily monitoring. It has also been designed to provide a 4DVAR-capable assimilation system.

The analysis produced by a 4DVAR system can be significantly improved if the assimilation of observations affected by moist processes such as clouds and precipitation is possible (Amerault et al. 2008; Errico et al. 2007; Errico and Raeder 1999; Janisková et al. 1999; Lopez and Moreau 2005; Mahfouf and Rabier 2000; Stiller and Ballard 2009; Stiller 2009; Tompkins and Janiskova 2004). This requires the inclusion of moist physics in the adjoint and tangent linear models. Including moist physics will also improve the estimation of the impacts coming from instruments such as the Atmospheric Infrared Sounder (AIRS), High Resolution Infrared Radiation Sounder (HIRS), and Microwave Humidity Sounder (MHS), as well as feature-tracking satellites, all of which have a sensitivity to moist processes. Further applications that can benefit from having moist physics in the linearization are adjoint-based sensitivities (Jung and Kim 2009) and singular vector calculations (Ehrendorfer and Errico 1995). Additionally, three-dimensional variational data assimilation (3DVAR), 4DVAR, and many operational ensemble methods make use of the linear model in the observation operator. Including accurate moist physics in the linear model is essential when assimilating moisture-affected observations, such as those that will be available from the upcoming Global Precipitation Satellite (Hou et al. 2008).

Convection in the nonlinear GEOS-5 AGCM is modeled using the relaxed Arakawa–Schubert (RAS) convection scheme (Arakawa and Schubert 1974; Moorthi and Suarez 1992). Large-scale condensation is modeled using the scheme developed by Bacmeister et al. (2006). The convection is computed prior to the large-scale condensation. Holdaway and Errico (2013) examined linearity and stability in the RAS convection scheme by studying Jacobian sensitivities. They found the scheme to generally exhibit good linearity and stability properties. Based on the findings of that work, a few simple modifications are applied to the RAS scheme and then an exact linearization is developed. The Bacmeister et al. (2006) scheme is rather more complex than the RAS scheme. It contains strong nonlinearities and relies on a large number of inputs that are not readily available in the linearized model. Rather than attempting to implement an exact linearization of this scheme, a reduced large-scale condensation scheme is implemented, one that simply converts supersaturation to precipitation and warming (Errico et al. 1994).

Constructing the linearized moist physics and examining its behavior against the full nonlinear system is very useful for understanding how the sensitivities in the moist schemes behave. This can assist not only in improving the analysis and associated data assimilation tools but also in developing a general understanding of how the system responds to sensitivities in the moist-physics schemes. This is important for anyone developing data assimilation systems. Even ensemble Kalman filter type methods, which may not directly require the linear model, do rely on certain assumptions about linearity in the system. The developers of the full nonlinear moist-physics schemes can also benefit from the identification of sensitivities that are not realistic and thus where the scheme may require improvement.

The development of the linearized moist-physics schemes is outlined in section 2. Validation of the linear

approximation for the moist physics is examined in section 3. The effect that moist physics has on the observation impacts is discussed in section 4. In section 5 a case study of an intensifying Atlantic storm is used to discuss the impact of linearized moist physics on adjoint sensitivity studies. Section 6 discusses the role of moisture in the norm. Section 7 offers some concluding remarks.

2. Development

In this section the general approach to modeling the linearized moist physics is outlined. Modern numerical schemes, including the RAS scheme, are complex iterative procedures so the full linearization is not presented. Instead, some key components of the schemes being used are outlined, any simplifications added to the schemes are listed, and any linearization issues are addressed.

The prognostic variables that will be used in the moist-physics schemes in the GEOS-5 linear model are the zonal and meridional wind speeds u and v ($\mathrm{m\,s^{-1}}$), the potential temperature θ (K), the specific humidity q ($\mathrm{kg\,kg^{-1}}$), and the surface pressure p_s (Pa). The profiles of temperature and specific humidity, along with the surface pressure, are used to parameterize the moist processes that are occurring. For the purposes of this work the surface pressure is not altered by the moist processes. The horizontal wind speeds can be altered, for example through vertical transport of momentum by convection, but are not explicitly used in determining the moist processes that are occurring. In the following text when linearizing variables, a superscript prime ($'$) is used to denote perturbation parts and a superscript r is used for the reference parts (e.g., $\theta = \theta^{(r)} + \theta'$).

a. Convection

The relaxed Arakawa–Schubert scheme was first introduced by Moorthi and Suarez (1992). The scheme offers a number of simplifications on the original Arakawa and Schubert (1974) scheme to make it suitable for use in general circulation models. The main simplification is that the system is relaxed toward equilibrium at each step, as opposed to being iterated to a fully equilibrated state every time step.

Convection in the atmosphere is an inherently nonlinear process. Mechanisms that need to be represented include latent heating, deposition, fast updrafts and downdrafts, and reevaporation. In addition to this the schemes used to represent convection often make use of discontinuous modeling and employ artificial processes that increase nonlinearity. A simplified explanation of the RAS scheme is given here to demonstrate the linearization limitations. See Moorthi and Suarez (1992) for a full derivation and explanation of the scheme.

The RAS scheme updates a single column of the atmosphere at one time and considers an ensemble of cloud depths within the column. Each cloud depth is assumed to have its base at the same model level, currently given by the top of the boundary layer. Different clouds have different detrainment levels, starting from the base of the cloud up to a maximum (around 30 hPa in the current formulation). The model has 72 levels in total and the lid is at 0.01 hPa. Clouds in the ensemble are characterized by an entrainment parameter λ. The normalized mass flux for each cloud depth is linear with height. Central to the algorithm is the calculation of a cloud work function for each cloud depth. The RAS scheme models all depths of convection using the same algorithm. Shallow convection just has fewer members since clouds detraining at higher levels are not found.

Before the effects of individual cloud depths are applied to the atmospheric profile, the RAS scheme determines whether convection is occurring (i.e., whether cloud is detraining at that model level). This calculation is done through a number of conditional statements based on the atmospheric profile in that column. In all, there are six conditional statements used to check whether a cloud depth should be included. The first considers whether the relative humidity at the cloud base is above a certain threshold. Second, the moist static energy at the base layer must be greater than the saturation moist static energy at the detraining level. Third, the entrainment parameter λ [Eq. (A18) in Moorthi and Suarez (1992)] must be positive. The entrainment parameter must also not be above a critical value, currently $10^{-4}\,\mathrm{kg\,m^{-2}\,s^{-1}}$. The RAS scheme computes a cloud work function [Eq. (A22) in Moorthi and Suarez (1992)], and it has to be above a critical value. The critical value itself is modeled discretely based on the profile. The final condition is that the rate of change of the cloud work function is negative and that the liquid water mixing ratio of the detraining air is positive.

In addition to these explicit conditional statements, used to determine the presence of convection, conditional statements exist in the formulas themselves. These discrete steps in the function are evident in the algorithm presented in the appendix in Moorthi and Suarez (1992) and are represented in the numerics using minimum and maximum statements. Discrete modeling will result in nonlinearity in the system and inaccuracy in a linear version of the scheme.

When attempting to represent moist physics in the linear model, a number of options exist, as outlined by Holdaway and Errico (2013). Ideally, an exact linearization of the schemes would be implemented.

However, an exact linearization may quickly diverge from the true perturbation trajectory since discontinuities and nonlinearities in the model are not properly taken into account. Alternatively, one could develop a new simple scheme to linearize (Lopez 2007). Or one could first apply simplifications or smoothing to the nonlinear scheme to refine the behavior of the subsequently derived linear model. All options come with challenges and difficulties.

To estimate the nonlinearity in the RAS scheme, Holdaway and Errico (2013) performed an extensive study examining Jacobian sensitivities. Despite the potential nonlinearity in the RAS formulation, they found that a large degree of linearity exists, especially for mid- to deep convection. In addition the linearized RAS scheme was found to be largely stable such that any solutions exhibiting growth are relatively well behaved.

The findings of Holdaway and Errico (2013) motivate an approach that involves the exact linearization of the RAS scheme, with only minor simplifications to the algorithm. These simplifications are related to the way certain constants are chosen and have a minimal effect on the overall behavior of the scheme. If the highest detraining level is less than eight model levels (approximately 80–100-hPa depth) above the cloud base layer and no convective precipitation is occurring, the linear RAS scheme is not invoked. The scheme supports the carrying of tracers, which can be neglected in the linearization.

A complication when developing a linearization of the RAS scheme is due to the use of an ensemble of cloud depths, detraining at each model level. As the scheme iterates through the cloud depths, the atmospheric profile is updated. The order is from shallowest cloud to deepest. The updated profile is used in the calculation for the next cloud depth, and so on. When the adjoint is implemented, it will work in reverse order; that is, it will start with the deepest cloud in the ensemble and finish with the shallowest. To obtain the trajectory used in the calculation of the deepest cloud requires iterating though each proceeding cloud in the nonlinear model. This embedded loop results in either extensive recomputation or large use of memory for the saving of all variables for each cloud depth. Both approaches are considered here. The optimum approach is found to be a combination of recomputation and the saving of variables that are computed for every cloud depth. Recording the results of all discrete switches avoids unnecessary recalculations when iterating the adjoint.

b. Saturation specific humidity

Moist-physics schemes rely on the computation of the saturation mixing ratio. This is an important quantity that describes the mass of water vapor contained in a unit mass of saturated air. Mathematically, it is expressed as

$$q_s = e_1 \frac{v_p(T)}{p - e_2 v_p(T)}, \tag{1}$$

where $v_p(T)$ is the saturation vapor pressure and $e_1 = 0.622$ and $e_2 = 0.378$ are dimensionless constants.

Saturation vapor pressure is computed using the Clausius–Clapeyron equation, a function of temperature that involves an exponential. The presence of the exponential makes the calculation expensive so in practice the formula is replaced with a table lookup. Values for $v_p(T)$ are interpolated from the table elements:

$$v_p(T) = v_p(T_m) + \frac{T - T_m}{\Delta T}[v_p(T_{m+1}) - v_p(T_m)]. \tag{2}$$

The table elements are denoted with subscript m; T_m and T_{m+1} are the nearest values below and above T used in obtaining v_p from the table. The resolution of the table is ΔT. This piece-wise linear table-lookup approach makes the formulation piece-wise linear in T. The linearized version of Eq. (2) is

$$v_p(T)' = \frac{T'}{\Delta T}[v_p(T_{m+1}^{(r)}) - v_p(T_m^{(r)})]. \tag{3}$$

In the current model the table lookup uses different formulations for ice and liquid phases. The table resolution is $\Delta T = 0.1$ K and values are computed between $T = 150.0$ and 333.0 K. See Murphy and Koop (2005) for a recent review of the methodologies used when calculating saturation vapor pressure tables.

c. Large-scale precipitation

For the large-scale precipitation a simple scheme that precipitates supersaturation, and as used by Errico et al. (1994), is applied. The scheme removes supersaturation at any model level by precipitating out excess water while heating the air until the relative humidity becomes 1. The amount of latent heating is applied to the potential temperature θ. The formulation for large-scale nonconvective adjustment is

$$q_{n+1} = q_n - \Delta q \quad \text{and} \tag{4}$$

$$\theta_{n+1} = \theta_n + \frac{L}{\pi c_{\mathrm{pm}}} \Delta q, \tag{5}$$

where subscript n denotes the time step. The moisture-dependent specific heat capacity is given by $c_{\mathrm{pm}} = c_p(1 + 0.887q)$, and the constant specific heat capacity is $c_p = 1004.49\ \mathrm{J\,kg^{-1}\,K^{-1}}$. The latent heat of condensation

is $L = 2.5104 \times 10^6\ \mathrm{J\,kg^{-1}}$. The Exner pressure π relates potential temperature to temperature, $T = \pi\theta$.

The profile adjustment is given by

$$\Delta q = \begin{cases} (q - q_s)\left(1 + \dfrac{\partial q_s}{\partial T}\dfrac{L}{c_{\mathrm{pm}}}\right)^{-1}, & \text{if} \quad q \geq q_s, \\ 0, & \text{otherwise}. \end{cases} \tag{6}$$

See Errico et al. (1994) for the derivation of these formulas.

The linearization of Eq. (4) is

$$q'_{n+1} = q'_n - \Delta q', \tag{7}$$

where

$$\Delta q' = \begin{cases} (q' - q'_s)\left[1 + \left(\dfrac{\partial q_s}{\partial T}\right)^{(r)}\dfrac{L}{c_{\mathrm{pm}}^{(r)}}\right]^{-1} - \dfrac{L}{c_{\mathrm{pm}}^{(r)}}\left(\dfrac{\partial q_s}{\partial T}\right)'(q^{(r)} - q_s^{(r)}) \times \left[1 + \left(\dfrac{\partial q_s}{\partial T}\right)^{(r)}\dfrac{L}{c_{\mathrm{pm}}^{(r)}}\right]^{-2}, & \text{if} \quad q \geq q_s, \\ 0, & \text{otherwise}. \end{cases} \tag{8}$$

Terms including c'_{pm} are considered small and are neglected. The $(\partial q_s/\partial T)'$ term is included here but is sometimes also neglected, as in Errico et al. (1994).

Note that the formulation above includes a switch, when $q = q_s$. This produces a nonlinearity. Away from the switch, the functions are linear and will perform well in the linearization. Close to the switch, the linear approximation may be less accurate. Smoothing could be applied near the switch, though this is not considered here.

The nonconvective precipitation rate is given by

$$R_{ls} = \frac{1}{g\Delta t}\sum_k \Delta q_k \sigma_k, \tag{9}$$

where g is the acceleration due to gravity, Δt is the model time step, σ_k is the air mass in level k, and Δq_k is the excess moisture in level k. Units are millimeters per unit area per second.

A simplification in the model is that the reevaporation of convective and nonconvective precipitation is not considered.

d. Dynamically adjusted trajectory

It is clear from Eq. (4) that supersaturation caused by the dynamics or convection is removed by the scheme. If the scheme was run twice in succession, the second run would produce no change to the variables.

An approximation in the system architecture in the linearized GEOS-5 AGCM is that the nonlinear trajectory is not passed between individual components of the linear model. For example, as the linearized dynamics is run, the trajectory could be updated and passed to the linearized physics components; however, only the perturbation trajectory is passed. The nonlinear trajectory is read in once per time step (20 min) and all components see the same trajectory profile, irrespective of the order in which components are called. This approximation is suitable when only large-scale dynamics and simple boundary layer physics are considered but could be problematic once equilibrium-seeking moist-physics schemes are included.

There is a discrepancy between the way that the physics and dynamics coupling are handled in the nonlinear model and the way they are handled in the linear model. In the nonlinear model each individual physics routine (turbulence, moist physics, gravity wave drag, radiation, etc.) produces a tendency for the temperature field. These tendencies are then weighted and combined at the end of a time step. The specific humidity is coupled consecutively; that is, each physics component sees specific humidity that has been adjusted by the previous component. In the linear model perturbation temperature and specific humidity are both handled consecutively. In a time step of the tangent linear model the dynamics is called and then the moist physics followed by the turbulence; for the adjoint model, the order is reversed.

Since the system does not pass trajectories and there is a discrepancy in the coupling, a new moist-trajectory component is added. The moist components are potential temperature and specific humidity, output by the nonlinear model just prior to the convection being called. This ensures that when the linearized moist physics is invoked, it sees unadjusted profiles and produces the correct effect. It is sufficient to use the original wind and pressure trajectory. The index of the cloud-base layer and the highest level of convection are also output into the trajectory by the nonlinear model. These quantities are required by the RAS scheme and are used to determine if convection is deep enough to be considered in the linear model.

e. Energy norm

Many applications of the linear model require a choice of metric. For example, when observation impacts are

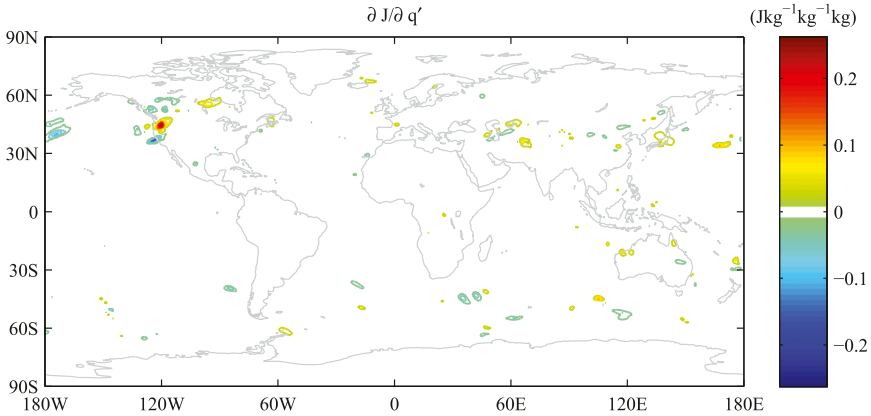

FIG. 1. Sensitivity initialized at 0000 UTC 18 Mar 2012 and integrated for 6 h using the adjoint model. Shown is the sensitivity of dry total energy with respect to the specific humidity $\partial J/\partial q$ (kg kg^{-1}) at 500 hPa. The linearized moist physics is switched on.

computed, the adjoint is used to determine the reduction of the forecast error at the observation time. That error must be measured in a certain way.

A common choice for the metric is the total perturbation energy, a sum of kinetic energy and a form of approximate available potential energy. So, observation impacts describe the ability of a certain set of instruments to reduce the error measured in terms of the total energy. The choice of metric is somewhat arbitrary and must be carefully chosen with the application in mind. For example, Hoover and Morgan (2011) use an adjoint model to examine cyclone track sensitivities. They use a vorticity error measure since that is deemed the important quantity; using total energy may not describe the sensitivities that are most relevant. In another study, Errico and Vukicevic (1992) initialize adjoint sensitivity integrations with the error in the forecast of surface pressure.

Currently, a dry total energy norm is used in GMAO's operational observation impact tools (Gelaro et al. 2010). This is a suitable choice given the dry physics in the current linear model. The interpretation of the total energy norm is presented in Errico (2000). In this study the use of a moist component in the energy norm metric is also considered.

The formulation that is used for the total energy norm follows that given by Ehrendorfer et al. (1999):

$$e = \frac{1}{2}\left[u'^2 + v'^2 + \frac{c_p}{T_0}T'^2 + RT_0\left(\frac{p_s}{p_0}\right)'^2 + \varepsilon\frac{L^2}{c_pT_0}q'^2\right]. \tag{10}$$

In Eq. (10), u' and v' are the perturbed horizontal wind components and p_s' is the perturbed surface pressure. The constants T_0 and p_0 are 270.0 K and 1000.0 hPa, respectively, and $R = 287.00\,\mathrm{J\,kg^{-1}\,K^{-1}}$ is the gas constant of dry air. The energy e is integrated to give a single value measure. Currently, the integration is over the full horizontal domain and up to approximately 125 hPa.

The moist static energy is obtained by choosing $\varepsilon \neq 0.0$. However, choosing $\varepsilon = 1.0$ gives a relatively large weighting to the q component of the norm. In fact with this choice the q component of the total energy norm is considerably larger than other components. A number of challenges remain in terms of properly utilizing moist observations and representing moisture in the linear model. It may not be sensible to give moisture its full weight in the metric before properly investigating the behavior. Doing so may lead to misinterpretation of the observation impacts, for example by heavily skewing impacts toward moist-sensitive instruments. However, since the choice of metric is largely arbitrary, it is possible to adjust ε and tune the relative weighting of the q component. Experiments presented here that employ the moist norm use a value of $\varepsilon = 0.3$; this produces approximately equal weighting between the temperature and specific humidity components of the norm. Experiments using the dry total energy use $\varepsilon = 0$.

f. Filtering of problematic profiles

Figure 1 shows an adjoint integration initialized with the dry total energy at 0000 UTC 18 March 2012. Linearized moist physics is switched on in the model. The adjoint is propagated 6 h to 1800 UTC 17 March 2012. Figure 1 shows the sensitivity with respect to specific humidity at this time (i.e., the end of the adjoint run). For the most part the model is well behaved and sensitivities to moisture are seen in reasonable places, especially where storms or fronts are occurring. However,

a large localized sensitivity is seen over northwestern North America. Sensitivity in the tropics is present, although it is masked by the dominant behavior in the plot. Further integration of the adjoint causes this dominant sensitivity to grow to a very large magnitude.

When developing a linear scheme, it is important to consider the stability. There are two issues regarding stability: first, as to whether the scheme is paired with a proper time-stepping scheme with a small enough time step and, second, as to whether the scheme contains growing modes. Generally, if the linear scheme is not paired with a suitable time step and temporal discretization, it will be apparent very quickly in a diverging solution. Further, it is generally quite rare for the nonlinear scheme to be temporally stable and the linear version of the scheme to not be. On the other hand, the presence of growing modes can often be an issue in the linearized version of the model. For example, consider baroclinic instability in the atmosphere, represented by growing modes in the nonlinear system. Later on, equilibrium is returned in the atmosphere by other processes, perhaps represented in a separate part of the nonlinear model. A simplified linear model that can represent the perturbations resulting from the baroclinic instability but that does not represent the processes responsible for the restoration to equilibrium would quickly run into trouble. In addition to this the linearization can result in large growth rates, for example when a division by something small is introduced through the differentiation.

The large gradients that are observed here appear suddenly during just one time step, can first appear at various times, and do not seem to be affected by the choice of time step; both 15 and 20 min have been tested. It seems likely that they are therefore due to the modeling of instability or large growth introduced in the linearization, rather than an issue with the choice of time step or time-stepping scheme.

A way to check for large growth is to examine the operator of the tangent linear model, as performed by Holdaway and Errico (2013). That study revealed the linearized RAS scheme to be stable or close to stable everywhere; however, the study only examined two specific times. If a large growth is encountered at just one particular time and location, it can remain part of the solution and lead to further growth, likely seen here. Further, even the rather small growth rates seen by Holdaway and Errico (2013) could amalgamate over time and lead to an issue.

It is not uncommon to encounter problematic profiles such as these when linearizing convection (Errico and Raeder 1999; Lopez and Moreau 2005). Generally, some kind of filtering must be developed, for which there are two general approaches:

- identify the cause and adapt the linearized (or, better, nonlinear) model to prevent problems occurring and
- use the trajectory to diagnose when the problem occurs and ignore or reduce the perturbation there.

Each of these approaches comes with advantages and disadvantages. Making a general change to the linear model may be more numerically efficient but can be difficult to implement. If the onset of problems is not particularly sudden, it can be difficult to identify why they occur and therefore produce a targeted enough filtering. This could result in too widespread of a correction, reducing the closeness to the actual perturbation trajectory. Identifying problem profiles as they occur can more easily produce a focused filtering but will be less numerically efficient since it will increase the number of calculations.

Here, the approach is to diagnose problematic profiles using the trajectory. Diagnosis must be done using the trajectory, rather than the perturbations themselves, so as not to cause a discrepancy between the tangent linear and adjoint models.

The tangent linear model is written as

$$\mathbf{y}' = \mathbf{M}\mathbf{x}', \tag{11}$$

where the vectors $\mathbf{y}'$ and $\mathbf{x}'$ represent the perturbation variables at the end and start times, respectively. The matrix $\mathbf{M}$ is the tangent linear model operator matrix and effectively gives the sensitivity of the scheme with respect to the input variables; $\mathbf{M}$ depends only on the trajectory, or reference, variables. The operator of the adjoint model is the transpose, $\mathbf{M}^{\mathrm{T}}$.

If large values suddenly appear in the perturbation quantities, the implication would be that the linear operator matrix $\mathbf{M}$ contains a large element for that profile. Therefore, a problematic location could be filtered by computing the linear operator matrix from the tangent linear or nonlinear model and identifying an unusually large element, or eigenvalue, as in Errico and Raeder (1999). However, computing $\mathbf{M}$ for every convective profile would be prohibitively expensive to do. Computing a column of $\mathbf{M}$ would generally require either an integration of the nonlinear model or an integration of the tangent linear model. Computing eigenvalues would be very demanding.

Holdaway and Errico (2013) showed that the structure of $\mathbf{M}$ is relatively simple and that the locations of its dominant features can be understood in terms of the profiles of temperature and moisture. Based on those findings, it should be possible to target specific columns of $\mathbf{M}$ and filter based on the values of the gradient in just those columns.

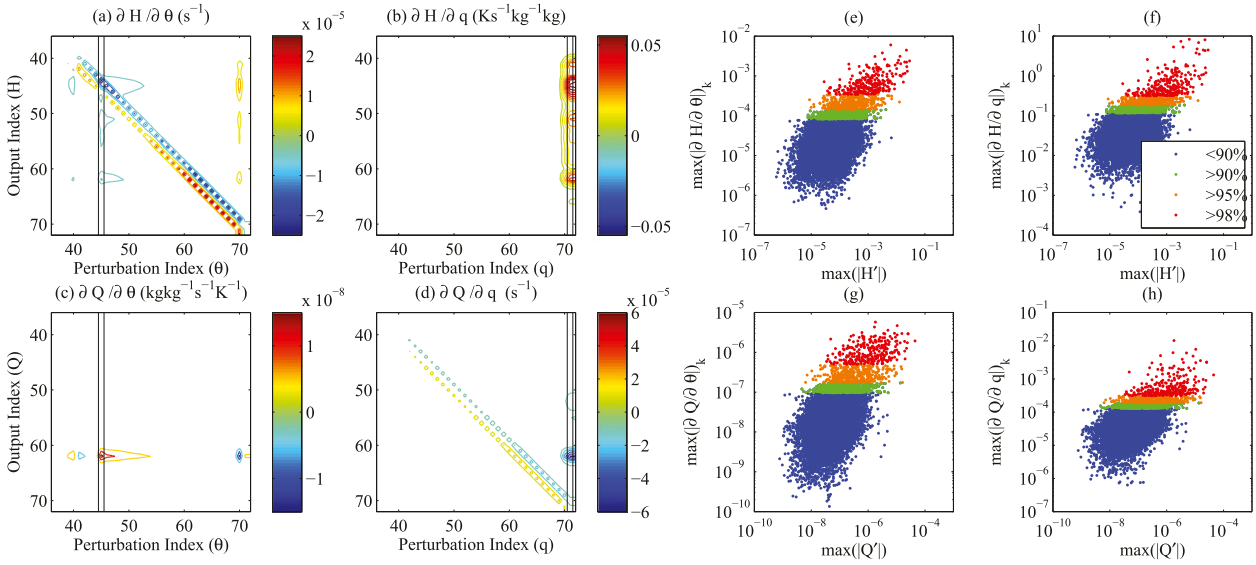

FIG. 2. The typical structure of the forward operator of the linear RAS scheme: (a) $\partial H/\partial\theta$, (b) $\partial H/\partial q$, (c) $\partial Q/\partial\theta$, and (d) $\partial Q/\partial q$. The black lines enclose the column of the operator that would be calculated for this particular profile during the filtering. The profile shown is one of deep convection. Comparisons of the maximum value in the linear heating and moistening rates throughout the layer for every profile against the maximum value in the columns of the operator that are computed for that profile: (e) H' vs the column of $\partial H/\partial\theta$, (f) H' vs the column of $\partial H/\partial q$, (g) Q' vs the column of $\partial Q/\partial\theta$, and (h) Q' vs the column of $\partial Q/\partial q$. Red points show the profiles that would be adjusted if the top 2% are filtered, orange and red if 5% are adjusted, and green, orange, and red if 10% are adjusted. Blue points are all the remaining profiles.

Consider only the potential temperature and specific humidity variables. For that case the matrix **M** in Eq. (11) is

$$\mathbf{M} = \begin{pmatrix} \dfrac{\partial H}{\partial\theta} & \dfrac{\partial H}{\partial q} \\ \dfrac{\partial Q}{\partial\theta} & \dfrac{\partial Q}{\partial q} \end{pmatrix}^{(r)}, \tag{12}$$

where $H = \partial\theta/\partial t$ and $Q = \partial q/\partial t$ are, respectively, the heating and moistening rates produced by the nonlinear scheme and t is the model time step. The full **M** would also contain components corresponding to the wind speeds and surface pressure. For the discrete model each component of **M** is itself a matrix and has dimension $n \times n$, where n is the number of model levels.

Figures 2a–d show the four components of **M**: (i) $\partial H/\partial\theta$, (ii) $\partial H/\partial q$, (iii) $\partial Q/\partial\theta$, and (iv) $\partial Q/\partial q$ for a particular atmospheric profile. The plotted operator is equivalent to Fig. 2 in Holdaway and Errico (2013), expect here it is computed using the exact tangent linear model. Each column of **M** is successively computed by initializing the tangent linear model with a vector of inputs **x** that is zero everywhere, except at the level corresponding to the column being computed, where it is set to one. The operator corresponds to a profile that exhibits deep convection. It is less efficient to compute **M** using the tangent linear model, rather than the nonlinear model, but it avoids the possibility of a switch in the numerics masking a problem and is found to produce a more reliable filtering.

Columns of **M** correspond to the level and perturbation variable being multiplied by that column. So if **M** has structure in a specific column, it describes a sensitivity to perturbations in that variable at that level. The rows of the column correspond to the levels at which the response to the perturbation occurs.

Figures 2a and 2c give the heating rate and moistening rate sensitivity with respect to temperature, respectively. For this profile the dominant sensitivity for temperature is to perturbations at levels 70 and 45 in the model, noted by the structure in those columns in Figs. 2a and 2c. Level 70 is the level of the cloud base and level 45 is where the nonlinear heating rate and upward mass flux are at their maximum. For the heating rate the response to perturbations at these two levels occurs throughout the convective region, seen by the structure across rows 45–62 in Fig. 2a. For the moistening rate the response is dominant at level 62, seen by the structure in this row in Fig. 2c. Level 62 is the location where the moistening rate is maximum. The diagonal feature in Figs. 2a and 2d represents a sensitivity to the calculation of dry and moist static energies (Holdaway and Errico 2013). The positive gradients below and

negative gradients above show how a perturbation at a given level causes an increase in heating and moistening at the level below and a decrease at the level above, changing the upward transport of temperature and moisture. Figures 2b and 2d give the sensitivity with respect to specific humidity. For moisture the dominant sensitivity is to perturbations in the subcloud layer, denoted by the structure in the columns to the right in these panels, again the response occurs throughout the convective region in the heating rate and at level 62 in the moistening rate. See Holdaway and Errico (2013) for a complete physical interpretation of these structures. The kind of behavior seen for this profile is evident for a wide variety of profiles.

The black lines in Figs. 2a and 2c show the column of **M** that corresponds to the level where the nonlinear heating rate is maximum. The black lines in Figs. 2b and 2d show the column of **M** corresponding to specific humidity one level below the cloud-base layer, inside the subcloud layer. Obtaining the sensitivity to perturbations of specific humidity somewhere in the subcloud layer and perturbations of temperature where the heating rate is maximum (i.e., the structure enclosed by the black lines) will give a large portion of the important dominant features. Two integrations of the tangent linear model will not be too computationally demanding. The filtering is performed by considering these columns and whether the values within them are of a reasonable size.

To know whether filtering based on just the two columns of **M** highlighted by the black lines in Figs. 2a–d will be successful, the magnitudes are examined against the linear growth rates for each convective profile simultaneously when running the tangent linear model. The tangent linear model is initialized using an analysis increment, which is indicative of the kinds of perturbation magnitudes that can be expected in practice. In Figs. 2e–h the maximum value in the computed columns of **M** are scattered against the maximum value in the perturbation heating and moistening rates: $H' = \partial\theta'/\partial t$ and $Q' = \partial q'/\partial t$. For one time step Fig. 2e shows the maximum absolute H', irrespective of which level the maximum occurs at, versus the maximum in the column of $\partial H/\partial\theta$ corresponding to the level where the nonlinear heating rate H is at its maximum (e.g., the column highlighted in Fig. 2a). Figure 2f shows H' versus $\partial H/\partial q$, Fig. 2g shows Q' versus $\partial Q/\partial\theta$, and Fig. 2h shows Q' versus $\partial Q/\partial q$. It is clear from these figures that the maximum value in the two columns of the operator matrix increases as the maximum value of the perturbation heating and moistening rates increases. This positive correlation means that filtering profiles based on the operator matrix, obtained with just these two perturbations, should be possible. In the figures the different colors show the profiles that are associated with the largest 2% (red), 5% (orange), and 10% (green) of elements in the reduced operator matrix.

To perform the filtering, four constants, for each quadrant of the operator matrix, are determined. If the maximum value of the computed columns is larger than any of the corresponding constants, then that profile is filtered. The four constants are chosen by examining the computed columns for a 24-h period (72 time steps) and then remain fixed for subsequent experiments. Values are chosen so as to filter on average around 4% of profiles per time step. This ensures all of the problematic profiles are dealt with while minimizing the number of profiles that are altered. In Figs. 2e–h approximately all of the red points and some of the orange points are filtered.

The filtering targets profiles for which H' is largest. Therefore, some profiles where strong convection is occurring will not be included. Since these will likely be important locations, rather than just assuming that no convection is occurring, the perturbation quantity is reduced by a factor of 10. This retains the sign of the perturbation while also preserving some sensitivity to the convection that is occurring. Fortunately, since only around 4% of profiles need to be filtered, much of the convective behavior is not affected by filtering.

The linearized moist physics has been tested with a 1° horizontal resolution (≈110 km at the equator) and with 72 levels in the vertical. Plans are under way to increase the horizontal resolution of the linear model to ½°; it is possible that the amount of filtering required will depend on the resolution and this will be tested. The linear model with moist physics has been tested with 20- and 15-min time steps. In both cases problematic profiles were encountered and had to be filtered. The time step did not impact the amount of filtering that was required in order to obtain a satisfactory solution. Different values of ε have been tested in the energy norm. Again, different choices did not impact the amount of filtering required.

3. Validation

The linear model is validated by considering how well it captures the nonlinear perturbation trajectory. The nonlinear perturbation trajectory is obtained by taking the end-time difference between two integrations of the nonlinear model. In one case the initial conditions are perturbed (by $\Delta\mathbf{x}$) and in one case they are not. The same perturbation is then used as the initial conditions of the tangent linear model to obtain the linear perturbation trajectory. By definition of the tangent linear model, as the size of the perturbation is reduced, the

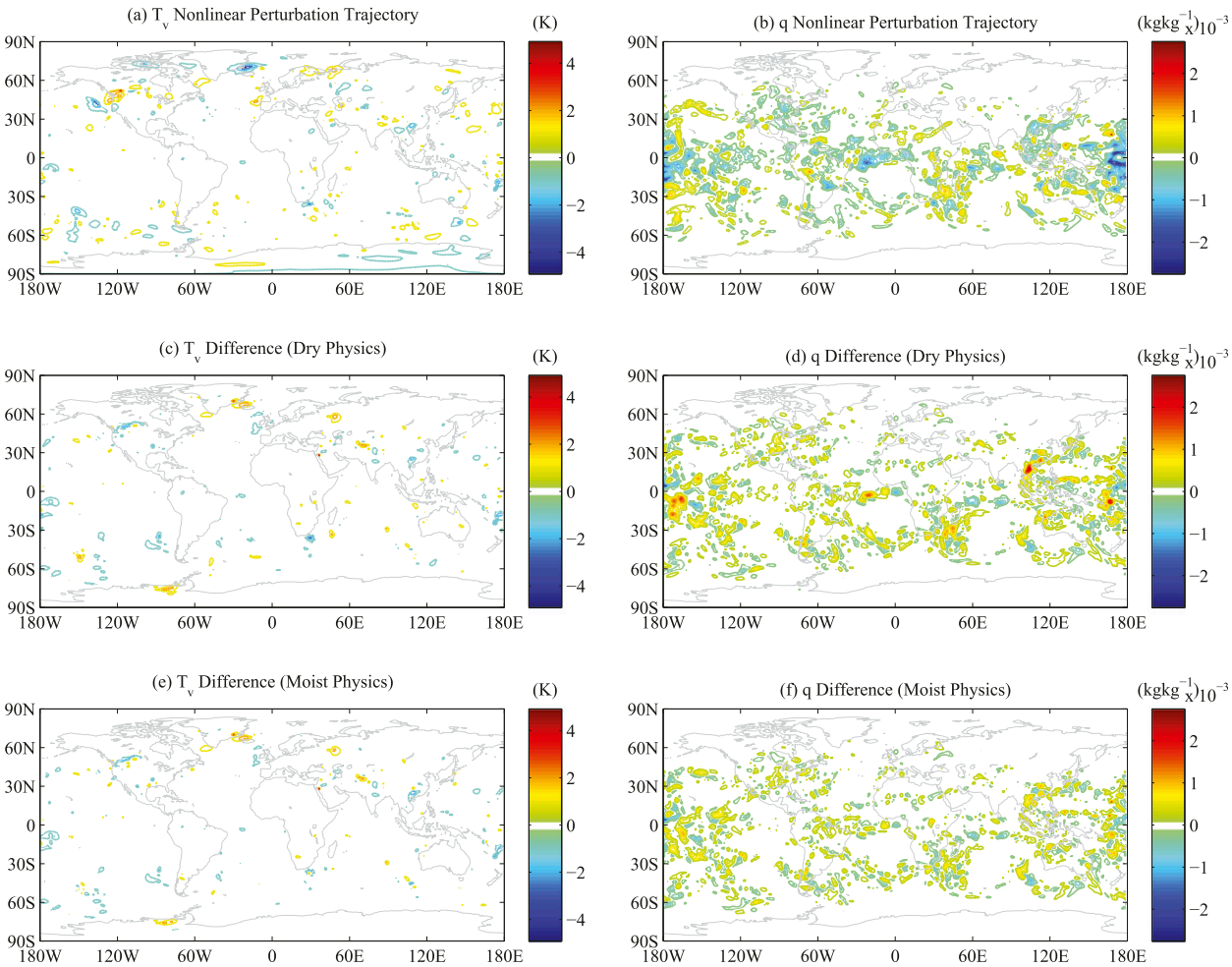

FIG. 3. The nonlinear perturbation trajectory for the (a) virtual temperature T_v and (b) specific humidity q at 500 hPa (level 50 in the model) and after a 6-h integration beginning at 0000 UTC 17 Mar. (c),(d) The difference between the nonlinear perturbation trajectory and the tangent linear model perturbation trajectory after 6 h with moist physics switched off. (e),(f) The difference when the moist physics is switched on; (left) virtual temperature and (right) specific humidity.

nonlinear and linear perturbation trajectories should converge. Mathematically this is expressed as

$$\lim_{\Delta\mathbf{x}\to 0} \frac{\mathbf{m}(\mathbf{x}+\Delta\mathbf{x}) - \mathbf{m}(\mathbf{x})}{\mathbf{M}\Delta\mathbf{x}} = 1, \tag{13}$$

where $\mathbf{m}$ represents the nonlinear model and $\mathbf{x}$ is the nonlinear model variables. The numerator gives the nonlinear perturbation trajectory and the denominator the linear perturbation trajectory.

In practice Eq. (13) will not hold, even for very small initial perturbations, due to switches and nonlinearity in the model. However, it is the behavior of the linear model in the presence of the realistic perturbation $\Delta\mathbf{x}$ that is of principle interest for almost all tangent linear and adjoint applications. The kind of perturbation that will be encountered in realistic applications is obtained by using an analysis increment (i.e., the analysis minus the background, $\Delta\mathbf{x} = \mathbf{x}_a - \mathbf{x}_b$).

For a given perturbation, the error in the linear perturbation trajectory is

$$\mathrm{TLM}_e = \mathbf{M}\Delta\mathbf{x} - [\mathbf{m}(\mathbf{x}+\Delta\mathbf{x}) - \mathbf{m}(\mathbf{x})]. \tag{14}$$

If Eq. (14) is zero, then the tangent linear model captures all the details of the nonlinear perturbation trajectory.

Figures 3a and 3b show the nonlinear perturbation trajectory for virtual temperature and specific humidity at 500 hPa (model level 50) after a 6-h integration. The integration is initialized at 0000 UTC 17 March 2012. Figures 3c and 3d show TLM_e, the tangent linear model compared to the nonlinear perturbation difference, for virtual temperature and specific humidity with the moist physics switched off in the linear model.

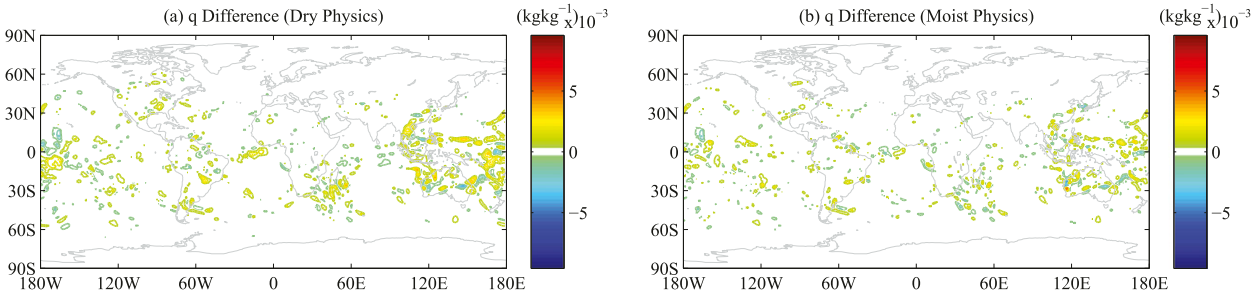

FIG. 4. As in Fig. 3, but for a 24-h integration and only showing specific humidity differences. The difference when the moist physics is switched (left) off and (right) on.

Figures 3e and 3f show TLM$_e$ with the moist physics switched on. Contours occurring in the difference plots correspond to where the structure is not captured by the tangent linear model.

Comparing Figs. 3d and 3f, it is clear that the inclusion of moist physics in the linear model significantly reduces the difference between the nonlinear and linear specific humidity perturbation trajectories. There are a number of regions, especially around the tropics, where the magnitude of the error has been reduced. The largest differences are seen over the Pacific, Indian, and Atlantic Oceans and over Southeast Asia. Regions where the dry model fails to capture aspects of the perturbation trajectory are improved once moist physics are included. Other levels are compared systematically by computing correlation coefficients between the nonlinear and linear perturbation trajectories. At every level below 100 hPa the correlation in the specific humidity field was improved; at levels above, the change is negligible. For the levels from 100 hPa to the surface the correlation between the nonlinear and linear perturbation trajectories increases from an average of 0.66 to 0.73 when going from dry to moist. At levels at around 100 hPa, where deep convection is dominant, the correlation increases by as much 35%.

Although there is significant improvement in the correlation of the specific humidity perturbation trajectory at 6 h, there is less improvement or change to the virtual temperature perturbation trajectory (Fig. 3c versus Fig. 3e). The dry-physics configuration produces a better representation of the temperature field perturbation than it does for moisture; correlations for the dry tangent linear model average around 0.73 for below 100 hPa. There are some improvements when switching on moist physics, notably off the coast of South Africa, over the western Pacific, and over the Southern Ocean. Despite seeing little change in the features, the correlation coefficient is improved at almost all model levels in the temperature field. For the moist linear model the average correlation is around 0.76 below 100 hPa. There is also a minor improvement in correlation for the wind fields with moist physics included. Positive impact on the wind fields results from the linear modeling of cumulus friction.

That the linearized moist physics model performs well for the 6-h window suggests it will prove useful in variational data assimilation applications. In the current configuration used at GMAO a 6-h window is used. In the proposed 4DVAR system the linear perturbation trajectory will be used across this window. Further, the observation operator employs the linear model to produce model space equivalents to the observations; accurate representation of moisture in the linear model over the analysis window is essential for assimilating moisture-affected observations.

Figure 4 shows TLM$_e$ at 500 hPa for a 24-h integration. As for the 6-h integration the inclusion of linearized moist physics results in the largest difference in the moisture field (temperature not shown). The most significant improvement is seen over the Pacific Ocean, over Southeast Asia, and off the southeast coast of Africa.

For the 24-h integration the correlations are much lower than they are for 6 h. For the dry model the average temperature correlation below 100 hPa is around 0.32 and the average moisture correlation is around 0.17. When including the linearized moist physics, the temperature correlation improvement is neutral and the moisture correlation increases to 0.2. These lower correlations are expected since the nonlinearities will cause a drift from the nonlinear perturbation trajectory. However, there is some improvement at almost every model level, either for the temperature or the moisture. For the wind fields the correlations are largely unchanged over 24 h.

The maximum perturbation over 24 h occurs for the moist model over Western Australia, seen in Fig. 4b. The magnitude is a little larger than expected and likely results from instability that is not captured by the filtering. However, it is clear that the overall difference between the linear and nonlinear perturbation trajectories

is decreased when including moist physics and that problematic points are being successfully filtered without the overall solution being disrupted. If filtering is switched off (not shown), a number of locations have very unrealistically steep gradients.

Having an occasional, isolated, slightly too large value should not have a negative impact on the applications of the tangent linear model. For example it would not modify the observation impacts. For the 6-h integration the largest values produced by the moist model are similar to the largest values produced by dry model, suggesting the filtering is sufficient.

The dot product test is used to ensure that the adjoint and tangent linear models have been coded in an equivalent way. This is done by checking that $\mathbf{y}^{\mathrm{T}}(\mathbf{M}\mathbf{x}) = (\mathbf{M}^{\mathrm{T}}\mathbf{y})\mathbf{x}$. The same level of similarity that is encountered for the dry model is found with the moist physics switched on.

4. Observation impacts

The GMAO routinely computes adjoint-based observation impacts to monitor the large network of observations and instruments (Gelaro et al. 2010). The adjoint-based observation impact tool is based on the work of Langland and Baker (2004). Employing adjoint-based impacts is a very powerful and useful tool. Impacts can be examined per instrument, per channel, for different regions of the globe, in a time series, and in averages. Metrics that are available include impact per analysis, impact per observation, fraction of beneficial observations, and observation count per analysis.

Impacts are computed by integrating two free-running forecasts over a 24-h window. One forecast is initialized using the analysis $\mathbf{x}_a$ and one is initialized using the background $\mathbf{x}_b$. The forecast initialized with the analysis benefits from an extra set of observations and so will have a smaller error at the end time. The error at the end time is given by

$$e_f = (\mathbf{x}^f - \mathbf{x}^t)^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{E}\mathbf{P}(\mathbf{x}^f - \mathbf{x}^t), \tag{15}$$

where $\mathbf{E}$ is a matrix that defines the energy norm [using Eq. (10)] and $\mathbf{P}$ is used to select the domain over which the error is calculated. Superscript f denotes the forecast and superscript t denotes the truth, or verification. The truth is approximated from the model analysis at the verification time. The nonlinear observation impact is then given by the difference between the errors for the two forecasts: $e_f(\mathbf{x}_a^f) - e_f(\mathbf{x}_b^f)$. It is the reduction in error due to the extra observations and analysis.

The adjoint is used to propagate the energy norm gradient backward 24 h and obtain sensitivities at the beginning of the window. The sensitivities are passed through the adjoint of the data assimilation system to convert them to observation space and the impacts. The algorithm for estimating the observation impact, $e_f(\mathbf{x}_a^f) - e_f(\mathbf{x}_b^f)$, using the linear model is described in the appendix in Langland and Baker (2004). In observation space the impact estimate is given by the vector product:

$$\left\langle (\mathbf{y} - \mathbf{H}\mathbf{x}_b), \mathbf{K}^{\mathrm{T}}\left(\frac{\partial J_a^f}{\partial \mathbf{x}_a} + \frac{\partial J_b^f}{\partial \mathbf{x}_b}\right)\right\rangle, \tag{16}$$

where $\mathbf{y}$ are the observations, $\mathbf{H}$ is the linearized observation operator, and $\mathbf{K}^{\mathrm{T}}$ is the adjoint of the data assimilation procedure. Sensitivities $\partial J_a^f/\partial \mathbf{x}_a$ and $\partial J_b^f/\partial \mathbf{x}_b$ are the gradients of cost functions describing the error in the two forecasts, then mapped to observation time using the adjoint.

How good of an approximation the linear observation impacts give of the full nonlinear observation impact depends on how good of an approximation is obtained of $\partial J_a^f/\partial \mathbf{x}_a$ and $\partial J_b^f/\partial \mathbf{x}_b$ and on any approximation made in the data assimilation system. There are a number of key factors involved. First, how much of the model is represented in the linear model: missing physics will diminish the approximation? Second, how linear is the nonlinear model? (Nonlinearity will cause the linear model to drift from the perturbation trajectory.) Third, how many approximations are made in the methodology and are aspects of the methodology accounted for in the adjoint model? For example, the GEOS-5 data assimilation uses incremental analysis updating (Bloom et al. 1996), meaning observations are applied to the state gradually over the assimilation window. The adjoint does not take this into account, reducing its accuracy.

As described by Gelaro et al. (2010), the NASA GEOS-5 model uses a double outer loop in the atmospheric data assimilation. Trémolet (2008) showed that systems utilizing multiple outer loops require the use of a second-order adjoint model in order to properly capture the observation impact. A second-order adjoint is not used in the GEOS-5 data assimilation system, so this would present a potentially large approximation in the methodology. To circumvent this, the assimilation is performed using a single outer loop and the same minimization algorithms are used for both the analysis (forward) and the sensitivity (backward) parts of the integration. The observation impacts computed using this single outer loop mode are compared with the operational observation impacts and are found to be in good agreement.

Figure 5 shows the global nonlinear and linear observation impacts for a month, from 17 March until 17 April 2012. In Fig. 5 the two positive curves show the forecast errors e_f for the forecasts initialized from the

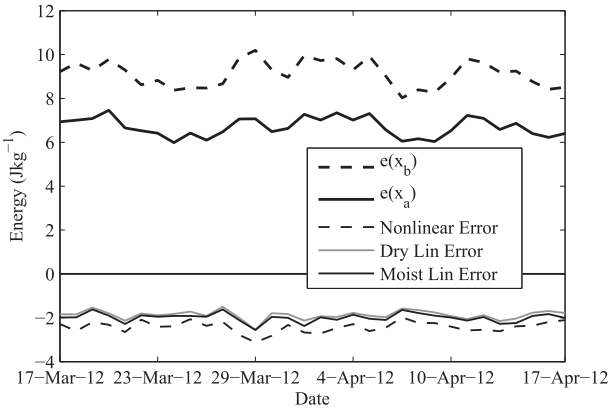

FIG. 5. The forecast error measured in the dry energy norm. The positive solid curve shows the error for forecasts initialized using the analysis, and the positive dashed curve shows error for forecasts initialized using the background. The negative dashed curve shows the difference between the forecast errors: the nonlinear observation impact. The negative gray curve shows the total adjoint impact when using the dry model; the negative black curve shows the adjoint impact when moist physics are included.

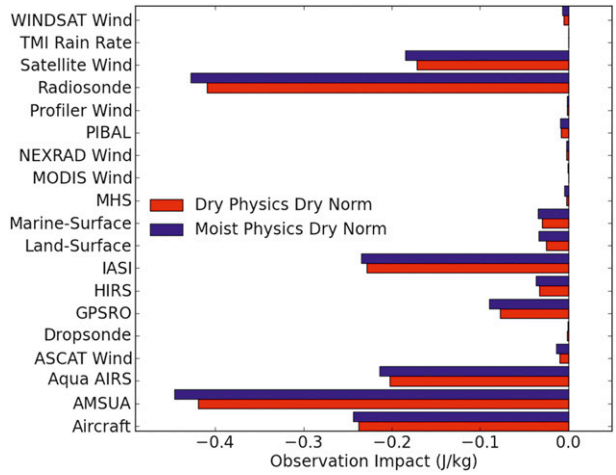

FIG. 6. The 24-h forecast observation impacts per analysis for each instrument for the period 17 Mar–17 Apr 2012. Red bars show the dry-physics configuration of the adjoint and the blue bars show the moist-physics configuration. A dry norm is used.

analysis and background states. The forecast error is computed once daily at the 0000 UTC time. Errors are computed using the standard dry energy norm. The dashed negative curve shows the nonlinear observation impact and is the difference between the two positive curves. The two solid negative curves show the linear observation impacts. The gray curve shows the linear observation impact when using only the dry physics in the adjoint, and the black curve shows the linear observation impact when including moist physics in the adjoint.

With moist physics included in the adjoint the amount of the nonlinear observation impact captured by the linear model increases from approximately 77.54% on average to 82.59%. This difference between the dry model and the moist model is statistically significant at the 95% confidence interval.

Figure 6 shows the average total impact per analysis, for each instrument currently assimilated at the GMAO. The average total impact per analysis shows how much, on average across the month-long period, each instrument contributed to the total observation impact, shown in Fig. 5. The sum of all the blue bars gives the average value of the negative black curve in Fig. 5; the sum of the red bars is the average of the negative gray curve. The observation impacts are computed for the same time period (17 March–17 April 2012). The figure compares the impact for each instrument when using the dry adjoint model (red bars) and when moist physics are included in the adjoint (blue bars). In both cases the error is measured using the dry energy norm so as to estimate only the effect of adding the linearized moist physics.

Table 1 numerically compares the impacts found using the dry physics and those found with the moist physics. It shows the percentage of the total impacts using moist physics that is found when using dry physics (i.e., 100 × red bars/blue bars in Fig. 6). Values larger than 100% in Table 1 represent a reduced reported impact once moist physics is used in the adjoint.

When linearized moist physics are included, all but three instruments are reported as having an increased positive impact on the forecast. The three instruments for which the reported impact is reduced are dropsonde, Next Generation Doppler Radar (NEXRAD) winds, and profiler winds. These are all instruments that have a very small overall impact on the forecast error. The

TABLE 1. The percentage of the total observation impact when the moist vs dry model is used. The dry norm is used in both cases. Percentages in the table effectively show the magnitude of the red bars relative to the magnitude of the blue bars in Fig. 6.

Instrument	By dry (%)		
Aircraft	97.43	MHS	49.14
AIRS	94.70	MODIS	37.35
AMSU-A	94.03	NEXRAD	111.56
ASCAT	74.02	Pilot balloon (pibal)	87.90
Dropsonde	136.42	Profiler wind	111.14
GPS radio occultation (GPSRO)	86.23	Radiosonde	95.79
HIRS	90.53	Satellite wind	93.15
IASI	97.59	TMI rain	42.80
Land surface	75.50	WindSat	78.78
Marine surface	86.58	Total	94.09

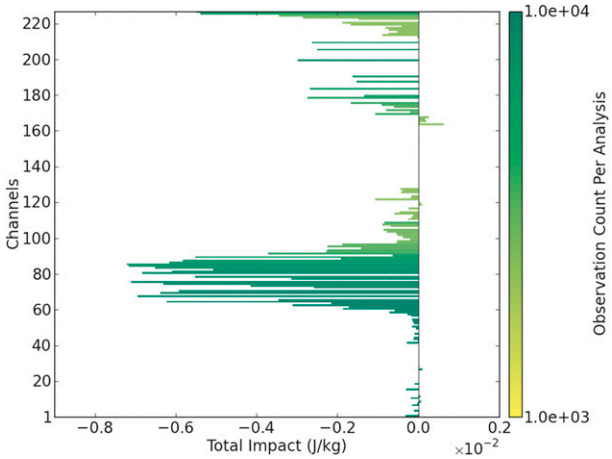

FIG. 7. The total impact per channel for the AIRS instrument. The time period is the same as in Fig. 6. The impacts are generated using dry physics and a dry norm.

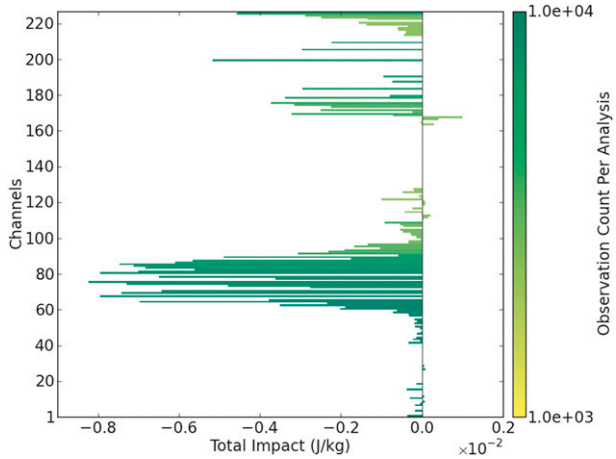

FIG. 8. As in Fig. 7, but computed using moist physics in the adjoint.

six largest overall impacts come from aircraft, the Advanced Microwave Sounding Unit-A (AMSU-A), the Atmospheric Infrared Sounder (AIRS), the Infrared Atmospheric Sounding Interferometer (IASI), radiosondes, and satellite winds. For all of these instruments the reported impact is increased when moist physics is included; the dry configuration captures between 91.15% and 97.59% of the impact. The instruments that have the largest differences between the dry and moist configurations are Advanced Scatterometer (ASCAT) winds, land surface (land stations from the Meteorological Assimilation Data Ingest System), Microwave Humidity Sounder (MHS), Moderate Resolution Imaging Spectroradiometer (MODIS) winds, and Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI) rain rate. For those instruments the configuration with dry physics in the adjoint captures between 37% and 75% of the impact. These are instruments that directly measure moist processes. Instruments such as AIRS and HIRS have moist-sensitive channels. However, the channels sensitive to temperature give the dominant impact so that the dry model captures 94.70% and 90.53% of the overall impact.

Figures 7 and 8 show the total impact per channel for the AIRS instrument. Figure 7 shows the impact when using the dry-physics model and Fig. 8 shows the impact when using the moist-physics model.

AIRS is an infrared sensor that measures over 2378 spectral channels, of which approximately 120 are assimilated. The channels that are most sensitive to moisture are between 160 and 190 in Figs. 7 and 8. When the moist configuration is implemented, the impact being reported from these channels is increased, as would be expected. However, the impact is typically 40% as large as the impact of the temperature-sensitive channels. The temperature-sensitive channels also have an increased impact when the moist physics are included in the model.

The HIRS and MHS instruments were also examined by channel (not shown). Again, the channels associated with measuring humidity in the atmosphere were found to have a larger impact when the moist configuration is implemented.

Moist norm case

So far, only the dry energy norm has been considered when computing the impacts. It is of interest to examine how the use of a moist norm would affect the observation impacts.

The above experiment is repeated for four configurations in total: dry adjoint model initialized with a dry norm (red bars in Fig. 6), dry model moist norm, moist model dry norm (blue bars in Fig. 6), and moist model moist norm. Table 2 shows the percentage of the nonlinear error captured by the four linear model configurations.

It is evident from Table 2 that including moist physics in the adjoint model increases the fraction of the observation impact captured in the linear approximation. This suggests that a better approximation of $\partial J_a^f/\partial \mathbf{x}_a$ and $\partial J_b^f/\partial \mathbf{x}_b$ in Eq. (16) is obtained. For the dry norm ($\varepsilon = 0.0$) case the percentage captured by the linear model increases by around 5% when including moist

TABLE 2. Comparison of the percentage of the nonlinear impact captured by the linear configurations.

	Dry norm (%)	Moist norm (%)
Dry-physics model	77.54	77.33
Moist-physics model	82.59	80.01

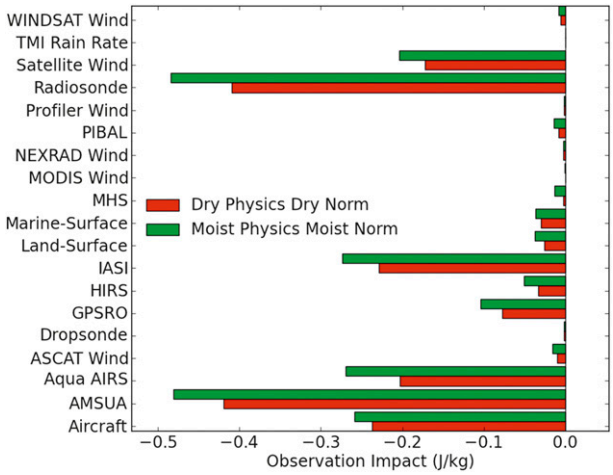

FIG. 9. As in Fig. 6, but comparing the dry-physics configuration with a configuration that has moist physics in the adjoint and uses the moist norm.

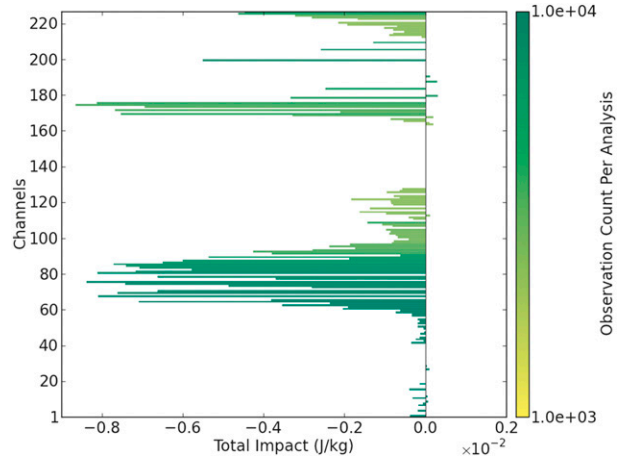

FIG. 10. As in Fig. 7, but computed using moist physics in the adjoint and using a moist norm.

physics. However, this is a somewhat unfair comparison. When including moist physics, a sensitivity will develop throughout the adjoint integration in regions where moist physics is occurring, even when the initial condition for specific humidity is zero. This means moisture is taken into account when computing the impact.

Consider an instrument that measures moisture over a location where moist physics is occurring. It is likely that a significant portion of its impact will be evident in the moisture field and the moist adjoint model can directly estimate that impact. Since the dry-physics linear model and the dry norm do not directly consider moisture in their calculations, the impact of that same instrument can only be measured indirectly through the other fields. It is therefore not surprising that a large increase is observed when comparing the dry and moist linear estimates against the nonlinear impact measured with a dry norm.

When the moist norm ($\varepsilon = 0.3$) is used, a more modest increase of around 3% is seen when going from a dry adjoint model to the moist adjoint model. This is a fairer comparison since in both cases moisture is taken into account in the nonlinear impact calculation and is included in the initial conditions of the adjoint. Even the dry linear model has some chance of modeling the impact on the moisture, through advection of the initial specific humidity field. That an increase of around 3% is observed here fairly demonstrates that the addition of linearized moist physics has increased the accuracy of the linear model.

Figure 9 compares the total impact of the various observation systems when using dry physics with the dry norm and moist physics with the moist norm for the total impact. Figure 9 shows the percentage captured by the dry–dry configuration compared with the moist–moist configuration.

When using the moist physics with the moist norm, the reported impacts are much larger than they are when using the dry–dry configuration. Of the six instruments that produce the largest impact, the dry–dry configuration only captures between 75% and 90% of the impact. The largest change for these instruments is for AIRS, for which the dry configuration captures around 75% of the moist impact. For the other moist-sensitive instruments there are also bigger differences. For HIRS around 66% is captured and for MHS only 17% is captured by the dry configuration.

Figure 10 shows the per channel total impact for AIRS when using the moist-physics moist-norm configuration. The impact from the humidity measuring channels is much larger when moisture is included in the norm. Impacts for these channels are of a similar order to those of the temperature-sensitive channels. The impact from the temperature channels in the 100–130 range is also increased.

The fourth configuration is the dry model with a moist norm. The impacts from that model are similar to those found with the moist model and the moist norm. The dry model captures 96.78% of the moist-model impact when the moist norm is used, compared to 94.09% when the dry norm was used. This smaller difference between the dry and moist models with the moist norm in terms of impacts is in agreement with the smaller difference between them in terms of the percentage of the nonlinear impact captured (Table 2).

A summary of the percentage of the impact captured by the dry model with the dry-norm configuration relative to the other configurations is shown in Table 3.

TABLE 3. The total percentage of the impact captured by the dry-model dry-norm configuration relative to the other configurations.

Configuration	%
Moist model dry norm	94.09
Dry model moist norm	85.49
Moist model moist norm	82.74

The two moist-norm configurations report the largest overall impacts. By taking moisture into account in the norm, the amount of energy that is measured is increased and this will influence the impacts. An additional mechanism by which observations can impact the forecast is being measured. For example, an instrument that measures moisture below an intensifying storm can have a much larger impact if the moisture is represented in the sensitivity, whether it is moderated by moist physical processes or not. The inclusion of moisture in the norm results in many areas where the impact due to moisture is being measured. This results in a larger relative change to the impacts.

For all of these experiments a factor of $\varepsilon = 0.3$ is included in the moisture term in the norm. If $\varepsilon = 1.0$ is used instead, then the change in impacts when alternating between the dry norm and the moist norm is very large, especially for moist-sensitive instruments. In the current operational model AIRS ranks at about fifth in terms of importance for the tropics, after AMSU-A, radiodsondes, aircraft reports, and IASI. Even with a dry model, using $\varepsilon = 1.0$ and the moist norm, AIRS is elevated to the most important set of observations here, becoming around 25% larger than even AMSU-A. This further reinforces the choice to limit the magnitude of the moisture term in the energy norm.

5. Adjoint sensitivity: Storm case study

An important research-based application of the linearized model is adjoint-based sensitivity analysis. While being only a by-product of the observation impact calculation, the sensitivity fields themselves are of interest since they contain information about sensitive regions in the initial state. The adjoint initial condition can be chosen around an area of interest so that sensitivities to just that area are obtained.

Of particular interest in numerical weather prediction is the development of cyclonic systems and the sensitivity of that development to the initial state. Understanding where the strongest sensitivities lie can be helpful in the design of future instruments and in tailoring observing networks. Most meteorological phenomena of interest involve both dry and moist dynamics, such as hurricanes or midlatitude depressions. Using the moist-adjoint model will improve the understanding gained from sensitivity studies of these systems.

As an example of such sensitivities, a strengthening depression that is tracking easterly over the North Atlantic Ocean toward Europe is identified. The adjoint is initialized as the storm reaches its maximum intensity, at 0000 UTC 25 March 2012. The energy norm is computed for a 4° × 4° box around the storm, centered at 56°N and 40°W. The adjoint is integrated backward 24 h to 0000 UTC 24 March 2012.

Figure 11 shows the adjoint sensitivity at 0000 UTC 24 March 2012 when using the dry adjoint model and dry norm; Fig. 12 shows the adjoint sensitivity when including moist physics in the adjoint and retaining the dry norm. Note that there is a change of scale between Figs. 11d and 12d. The sensitivities are shown for the 500-hPa level and are shaded red to yellow. Also shown in Figs. 11 and 12 is the mean sea level pressure and the convective precipitation rate at 0000 UTC 24 March 2012, contoured in black and blues, respectively. From 0000 UTC 24 March to 0000 UTC 25 March 2012 the center of the storm moves into the box, also shown in Figs. 11 and 12.

Comparing Figs. 11 and 12, it is clear that the sensitivity with respect to specific humidity changes significantly once moist physics are included in the adjoint. For the dry model the sensitivity with respect to q is on the order of 10^{-5} kg kg^{-1}. The specific humidity variable itself is generally three orders of magnitude smaller than the temperature and wind variables. That the sensitivity field with respect to q is of a similar order to the sensitivity with respect to other fields would mean the overall contribution to the energy would be three orders of magnitude smaller than for the other fields, giving the impression that total energy is not sensitive to specific humidity. When moist physics is included, the sensitivity with respect to specific humidity increases by three orders of magnitude. Given the relative magnitudes of the other variables, this would imply that the total energy in the box is as sensitive to specific humidity as it is to temperature and horizontal wind.

As well as the relative magnitude, the structure of the sensitivity with respect to the specific humidity also changes once moist physics is included. Overall, the sensitivity is concentrated nearer the low center than it is for other fields. There is also less variation with height (not shown). At most model levels the sensitivity remains close to the center of the storm, generally on the right-hand side of the storm. The sensitivity also tends to lie close to the region where there is convection occurring, shown by the presence of convective precipitation. Trailing behind the storm is a front where strong convection is evident. The sensitivity is not in the same location as this front for a 24-h adjoint integration.

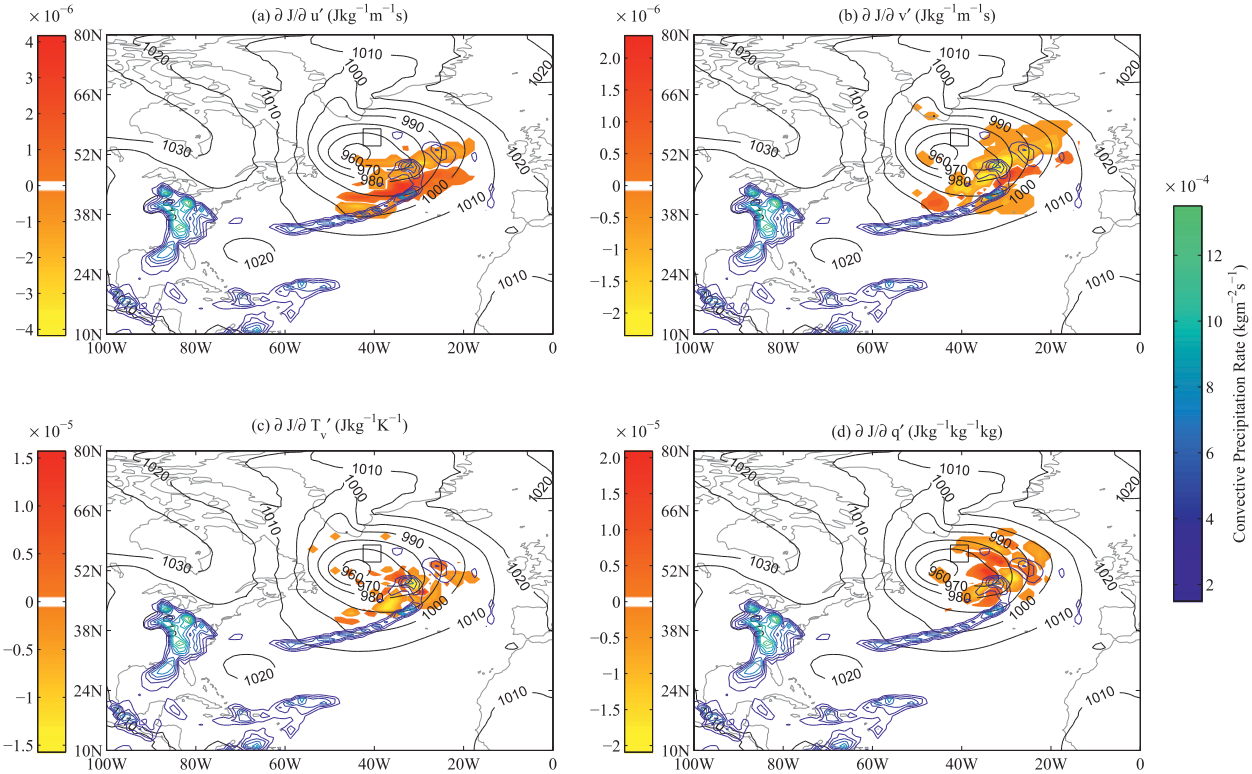

FIG. 11. Adjoint sensitivities to (a) u', (b) v', (c) T_v', and (d) q' at 24 h, shown in red-yellow shading. Sensitivities are shown for approximately 500 hPa. Adjoint initialized with energy norm forecast error in a box 54°–58°N, 38.5°–42.5°W, at 0000 UTC 24 Mar 2012. The convective precipitation at −24 h is shown in blue-green contours. In black contours is the sea level pressure at −24 h. Sensitivities are generated using the dry adjoint, i.e., without moist physics turned on. The box shows the region where the forecast error is measured. The center of the low pressure system moves into the box over the 24-h window.

For the sensitivity with respect to u, v, and T_v, the structure of the field is similar with or without moist physics included in the adjoint model. The only change is a slight increase in the magnitudes when moist physics is included. As expected, when higher levels were examined, the sensitivity was found to rotate clockwise around the storm with increasing height. For high enough levels the sensitivity with respect to the wind fields stretches out over northern Canada into the large-scale synoptic flow features (not shown).

The sensitivity experiment was repeated for both moist and dry norms with moist and dry physics in the adjoint. The sensitivities that are obtained when the moist-physics model is used are not found to be dependent on the choice of ε. Use of both dry and moist norms, with various choices for ε, produces approximately the same overall sensitivity fields. This is discussed further below.

6. Dependency on moisture in the norm

When using adjoint techniques, the choice of norm, which gives the initial conditions for the adjoint model, requires careful consideration. In choosing a specific norm a specific question is asked. Choosing the total energy norm, for example, poses the question: What is a quadratic measure of the perturbation field, that looks like energy, sensitive to? Or in the observation impact experiments: How much of a role does a particular observation play in reducing the error in the system, measured in this way? The choice of norm used to initialize the adjoint must be tailored to the problem being addressed. The energy norm is often used because it offers a convenient way of combining different model variables that have different units. Using the energy norm in the observation impact experiments allows for the quantification of the impact that observations have on all the variables simultaneously.

In both the observation impact experiments and the sensitivity study, dry and moist energy norms have been tested. Clearly, just the choice of ε in the norm has a large influence on the observation impacts that are obtained. However, it is interesting to note that for the sensitivity study the counter result is obtained, the same sensitivity is found regardless of the choice of ε.

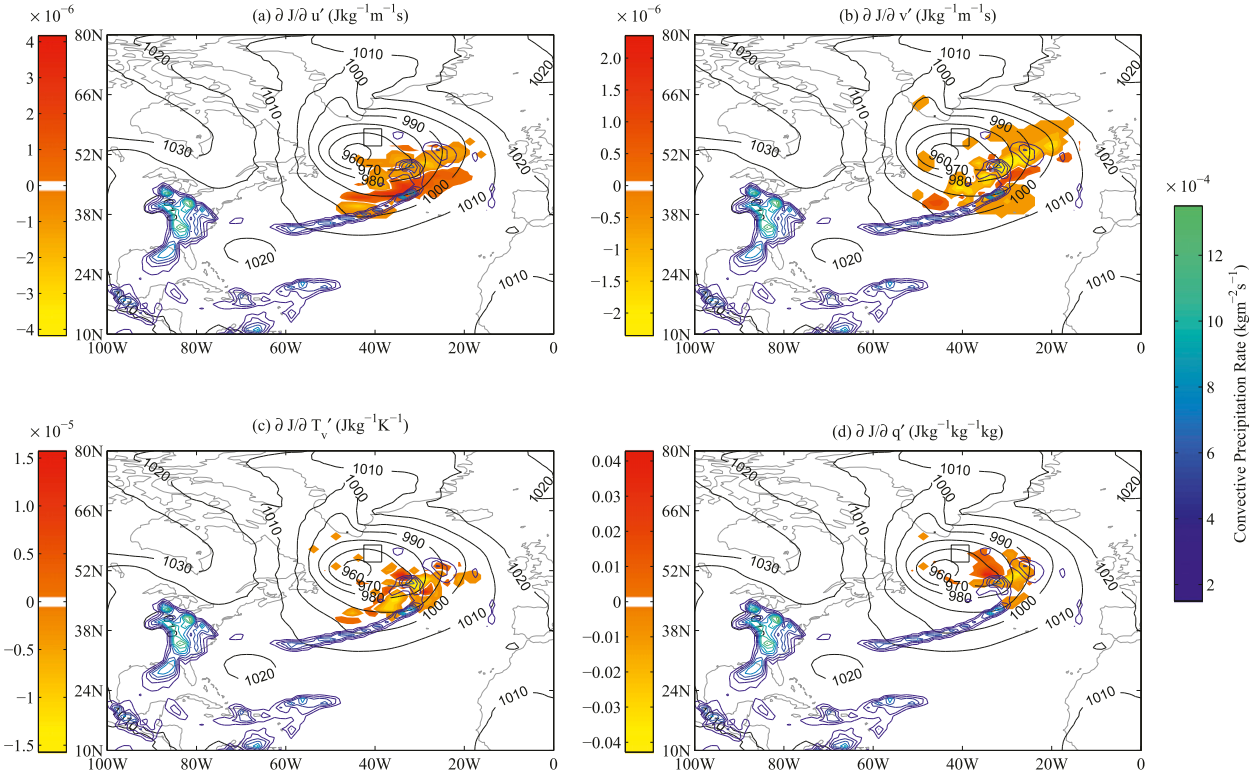

FIG. 12. As in Fig. 11, but with moist physics turned on in the adjoint model. Note the change in scale in (d) compared with Fig. 11d.

In this study it is found that when the moist norm is chosen, the modification of the sensitivity with respect to q due to the moist-physics schemes dominates over the modification due to advection. So a location that has moisture sensitivity only due to the initial conditions, or due to advection of the initial conditions, will have a smaller overall sensitivity than a neighboring location that has highly active moist physics. In addition to this, it is found that the initial conditions of moisture can be quickly forgotten when the linearized moist-physics schemes are called. Consider for example Eq. (1); in the formulation of q_s there is no dependence on specific humidity, only temperature and pressure. Much of the changes due to the moist physics are computed without considering the moisture in the initial conditions. The modification of the sensitivity by the moist physics depends much more strongly on the temperate in the initial conditions than the moisture in the initial conditions. After a few time steps at a location with active moist physics, the schemes will sufficiently modify the sensitivity to the point that the moisture initial conditions provided by the norm are forgotten.

The difference between the observation impact experiment and the limited-area sensitivity study is the relative number of locations where the moist-physics schemes are active. For the sensitivity study the adjoint is initialized with an energy norm computed only for the region close to the center of the storm. Over the temporal and spatial domain being examined both convection and large-scale condensation schemes are strongly modifying the sensitivity at every location. In addition to this the sensitivity quickly spreads out to areas where the norm was not calculated, further diluting the effect of the initial conditions.

For the observation impact study the adjoint model is initialized with a norm calculated for the whole globe. In this global situation there will be many locations where the moist-physics schemes are highly active and therefore modifying the sensitivity. However, there will be far more locations where the norm, and therefore the initial conditions, contains moisture sensitivity but where moist physics are not active. For these locations, where moist physics do not modify the sensitivity, the original choice of ε will determine the final sensitivity. Since there are a large number of locations like this, the choice of ε has a large overall impact.

That the limited-area sensitivity study is not sensitive to the choice of ε is not to say that it is not sensitive to the choice of norm entirely. If a different type of norm was chosen, for example the root-mean-square error of moisture in the box, then a very different sensitivity field would likely result. When performing this kind of

study, one has to ask what is of interest. It may be that a measure like total energy is of most interest or it may be some other metric.

7. Summary

A linearization of moist physics has been developed for use in NASA's GEOS-5 atmospheric data assimilation model. Convection is modeled using a linearization of the relaxed Arakawa–Schubert scheme. Large-scale condensation and precipitation are modeled using a linearization of a highly simplified scheme that removes supersaturation.

The new linearizations have been validated by comparing the tangent linear perturbation trajectory with the nonlinear perturbation trajectory that includes the full moist physics. For both 6- and 24-h tangent linear integrations the inclusion of moist physics is found to improve the correlation between the nonlinear and linear perturbation trajectories. In computing the observation impacts it is found that including moist physics increases the representation of the nonlinear impact by the linear model.

Since including moist physics increases the likeness between the linear and nonlinear observation impact, it should provide more realism in the impacts. Almost all of the instruments assimilated in GEOS-5 are reported as having a larger impact when the moist-adjoint configuration is considered. With the moist physics switched on, the estimate of the impact is more accurate since it is able to account for a larger fraction of the overall observation impact. Moist-sensitive observations from instruments such as MHS, TMI, infrared sounders, and some wind observations see the largest changes in the impact per analysis. Some instruments see a smaller number of their observations being beneficial but the difference is small and the total impact can still increase. Impacts from the AIRS, HIRS, and MHS instruments were examined by channel. Those channels with a sensitivity to moisture have the largest positive change. Most temperature-sensitive channels also see a small positive change.

A case study of an intensifying storm over the North Atlantic is used to examine the effect of including moist physics on sensitivity studies. With moist physics included, a large sensitivity with respect to the specific humidity is encountered. Whereas for the dry model the dominant sensitivity is with respect to horizontal wind and temperature, for the moist model the sensitivity with respect to all variables is equally large.

An important consideration when using the linearized adjoint model is the choice of norm. Currently, a dry total energy norm is used as the metric. Including moist physics in the model raises the question of whether moisture should also be considered in the metric. Four configurations were considered for both observation impacts and the storm sensitivity study. For the observation impacts there is a strong dependency on the presence of moist physics in the energy norm. Including moisture results in a large change in the observation impacts, whether the model physics are dry or moist. Choosing the natural form of moist static energy results in AIRS being reported as the instrument with the largest impact, ahead of AMSU-A. To reduce the emphasis on observations affected by moisture, a factor ($\varepsilon = 0.3$) is included in the moisture term. For the sensitivity case study the choice of ε is not as important. Whether a dry or moist norm is chosen, the same overall sensitivity field is obtained. For observation impacts there are many locations where no moist physics is occurring. Including moisture in the metric will allow instruments measuring moisture to have an impact at those locations. For the sensitivity study the locations where moist physics is active are dominant and the initial conditions are dominated by the physical behavior.

It remains to test this suite of linearized moist physics within a 4DVAR framework; it would be interesting to examine how well observations of moist processes can be assimilated with this particular linearized moist-physics package. Currently, the large-scale precipitation is represented in a very crude manner. It would likely be beneficial to try and include more complex processes such as reevaporation and to distinguish between types of precipitation.

Acknowledgments. This work is funded under the NASA–USRA GESTAR cooperative agreement. The lead author wishes to thank Ricardo Todling for his help with setting up and running the GEOS-5 data assimilation system. Thanks also to Will McCarty of NASA's GMAO for useful discussions on observation impacts and instrumentation channels.

REFERENCES

Amerault, C., X. L. Zou, and J. Doyle, 2008: Tests of an adjoint mesoscale model with explicit moist physics on the cloud scale. *Mon. Wea. Rev.,* **136,** 2120–2132.

Arakawa, A., and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. *J. Atmos. Sci.,* **31,** 674–701.

Bacmeister, J., M. Suarez, and F. R. Robertson, 2006: Rain reevaporation, boundary layer–convection interactions, and Pacific rainfall patterns in an AGCM. *J. Atmos. Sci.,* **63,** 3383–3403.

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.,* **124,** 1256–1271.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.,* **120,** 1367–1387.

Ehrendorfer, M., and R. Errico, 1995: Mesoscale predictability and the spectrum of optimal perturbations. *J. Atmos. Sci.,* **52,** 3475–3500.

——, ——, and K. Raeder, 1999: Singular-vector perturbation growth in a primitive equation model with moist physics. *J. Atmos. Sci.,* **56,** 1627–1648.

Errico, R., 2000: Interpretations of the total energy and rotational energy norms applied to determination of singular vectors. *Quart. J. Roy. Meteor. Soc.,* **126,** 1581–1599.

——, and T. Vukicevic, 1992: Sensitivity analysis using an adjoint of the PSU–NCAR mesoscale model. *Mon. Wea. Rev.,* **120,** 1644–1660.

——, and K. Raeder, 1999: An examination of the accuracy of the linearization of a mesoscale model with moist physics. *Quart. J. Roy. Meteor. Soc.,* **125,** 169–195.

——, V. Tomislava, and K. Raeder, 1993: Examination of the accuracy of a tangent linear model. *Tellus,* **45A,** 462–477.

——, K. Raeder, and T. Vukicevic, 1994: Mesoscale Adjoint Modeling System version 1. NCAR Tech. Note NCAR/TN-410+IA, 214 pp. [Available online at http://nldr.library.ucar.edu/repository/assets/technotes/TECH-NOTE-000-000-000-224.pdf.]

——, P. Bauer, and J.-F. Mahfouf, 2007: Issues regarding the assimilation of cloud and precipitation data. *J. Atmos. Sci.,* **64,** 3785–3798.

Gelaro, R., R. Langland, S. Pellerin, and R. Todling, 2010: The THORPEX Observation impact intercomparison experiment. *Mon. Wea. Rev.,* **138,** 4009–4025.

Holdaway, D., and R. Errico, 2013: Using Jacobian sensitivities to assess a linearization of the relaxed Arakawa–Shubert convection scheme. *Quart. J. Roy. Meteor. Soc.,* doi:10.1002/qj.2210, in press.

Hoover, B. T., and M. C. Morgan, 2011: Dynamical sensitivity analysis of tropical cyclone steering using an adjoint model. *Mon. Wea. Rev.,* **139,** 2761–2775.

Hou, A. Y., G. Skofronick-Jackson, C. D. Kummerow, and J. M. Shepherd, 2008: Global precipitation measurement. *Precipitation: Advances in Measurement, Estimation, and Prediction,* S. Michaelides, Ed., Springer, 131–169.

Janisková, M., F. Veersé, J.-N. Thépaut, G. Desroziers, and B. Pouponneau, 1999: Impact of a simplified physical package in 4D-Var analyses of FASTEX situations. *Quart. J. Roy. Meteor. Soc.,* **125,** 2465–2485.

Jung, B.-J., and H. M. Kim, 2009: Moist adjoint-based forecast sensitivities for a heavy snowfall event over the Korean Peninsula on 4–5 March 2004. *J. Geophys. Res.,* **114,** D15104, doi:10.1029/2008JD011370.

Langland, R., and N. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus,* **56A,** 189–201.

Lopez, P., 2007: Cloud and precipitation parameterizations in modeling and variational data assimilation: A review. *J. Atmos. Sci.,* **64,** 3766–3784.

——, and E. Moreau, 2005: A convection scheme for data assimilation: Description and initial tests. *Quart. J. Roy. Meteor. Soc.,* **131,** 409–436.

Mahfouf, J.-F., and F. Rabier, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. II: Experimental results with improved physics. *Quart. J. Roy. Meteor. Soc.,* **126,** 1171–1190.

Moorthi, S., and M. J. Suarez, 1992: Relaxed Arakawa–Schubert: A parameterization of moist convection for general circulation models. *Mon. Wea. Rev.,* **120,** 978–1002.

Murphy, D. M., and T. Koop, 2005: Review of the vapour pressures of ice and supercooled water for atmospheric applications. *Quart. J. Roy. Meteor. Soc.,* **131,** 1539–1565.

Putman, W., 2007: Development of the finite-volume dynamical core on the cubed-sphere. Ph.D. thesis, The Florida State University, 91 pp.

Stiller, O., 2009: Efficient moist physics schemes for data assimilation. II: Deep convection. *Quart. J. Roy. Meteor. Soc.,* **135,** 721–738.

——, and S. P. Ballard, 2009: Efficient moist physics schemes for data assimilation. I: Large-scale clouds and condensation. *Quart. J. Roy. Meteor. Soc.,* **135,** 707–720.

Tompkins, A. M., and M. Janiskova, 2004: A cloud scheme for data assimilation: Description and initial tests. *Quart. J. Roy. Meteor. Soc.,* **130,** 2495–2517.

Trémolet, Y., 2008: Computation of observation sensitivity and observation impact in incremental variational data assimilation. *Tellus,* **60A,** 964–978.