An examination of the accuracy of the linearization of a mesoscale model with moist physics

By RONALD M. ERRICO* and KEVIN D. RAEDER

National Center for Atmospheric Research, USA

(Received 29 August 1997; revised 26 May 1998)

SUMMARY

The accuracy of tangent linear and adjoint versions of a primitive-equation model with moist physics is examined with respect to growing perturbations having significant initial magnitudes. The Jacobians for the convective parametrizations are approximated using a perturbation method. These Jacobians are then quality controlled to ensure that the approximations are suitable. Results show that: (1) linearization of the diabatic moist physics can have a significant impact; (2) even where such impacts are large, the linearized versions of the model can yield good approximations to the nonlinear behaviour for significant perturbations, especially if there is sufficient dynamical influence; (3) poor approximations can be obtained when convection dominates the results; and (4) a straightforward linearization of some parametrization schemes may be inadequate. The results are encouraging for quantitative applications of some moist adjoint models to extratropical cyclones in the winter, but suggest some tangent linear approximations may be unsuitable in the tropics or over continents in the summer, except if qualitative agreements with nonlinear results are sufficient. Detailed comparisons of linear and nonlinear results should be made, particularly using optimal perturbations, prior to any applications of tangent linear or adjoint models.

KEYWORDS: Adjoint Primitive equations Tangent linear model

1. INTRODUCTION

The use of tangent linear and adjoint models continues to expand in meteorology (Errico 1997). When they are determined from nonlinear models without any approximations, they describe the exact behaviours of small perturbations introduced in the nonlinear models in the limit as the perturbation size becomes small (aside from the numerical limitation of round-off errors). Determination of such linearized behaviour is useful or necessary in applications of sensitivity analysis (e.g. Errico and Vukicević 1992), 4-dimensional variational data assimilation (e.g. Courtier and Talagrand 1990), synoptic studies (e.g. Langland et al. 1995), and stability analysis (e.g. Ehrendorfer and Errico 1995).

The behaviour of infinitesimal perturbations as explicitly revealed by tangent linear and adjoint models, however, is seldom of actual concern. Instead, almost all practical interests concern the behaviour of perturbations of similar size to uncertainties in the model’s formulation and application (e.g. uncertainties in its initial and boundary conditions, and its specification and formulation of parameters). The utility of tangent linear and adjoint models is, therefore, practically determined by how well they describe behaviour of perturbations that correspond to ones in the nonlinear model as the perturbation size increases to the size of actual uncertainties, rather than in the infinitesimal limit. In particular, in the four applications mentioned above, interest is focused on the behaviour of perturbations of similar size to uncertainties in initial condition (i.e. analysis).

When we are concerned with perturbations of finite size, tangent linear model (TLM) and adjoint model (AM) results must be assumed to yield only approximations to the behaviour of true interest, i.e. those in the nonlinear model (NLM). The accuracy of the linearization employed by the TLM and AM should therefore be a critical concern. This accuracy is distinct from measures of how correctly the linearization has been performed.

* Corresponding author: National Center for Atmospheric Research, P.O. Box 3000, Boulder, Colorado 80307, USA.
† The National Center for Atmospheric Research is sponsored by the US National Science Foundation.
i.e. whether the correct results are obtained in the infinitesimal limit. Instead this accuracy is a measure of the expected utility of the linearized models in some of their actual, finite-amplitude applications.

(a) Background

Very rarely have the accuracies of TLMs and AMs been described in the literature. Some notable exceptions are: linearizations of shallow water models by Lacarra and Talagrand (1988); quasi-geostrophic models by Oortwijn and Barkmeijer (1995); a dry, primitive-equation, mesoscale model by Errico et al. (1993); a column-cloud model by Park and Droegemeier (1995); and a micro-physical cloud model by Verlinde and Cotten (1993). In the dry models with perturbations initially the size of analysis uncertainties, corresponding perturbations in TLM and NLM forecasts typically remain strongly correlated throughout periods up to 2 or 3 days. Thereafter, perturbations typically have grown sufficiently large that their nonlinear effects in the NLM cause divergence from the TLM solutions.

When moist physical processes are simulated by a model, the accuracy of the linearization can be reduced. In the cloud model used by Park and Droegemeier (1995), this disagreement occurred once the mature stage of the cloud developed at about one hour into the forecast. Unlike nonlinearity in the formulation of advection, which is often quadratic, precipitation processes can have strong nonlinearity due both to the Clausius–Clapeyron equation, which determines saturated specific humidity from temperature and pressure, and to discontinuities associated with critical conditions of stability, saturation, mass flux, etc. Furthermore, parametrizations of precipitation processes are often designed to relax solutions toward critical points for these conditions, e.g. toward neutral stability or 100% relative humidity. This implies that at many locations and times, the forecast fields may lie close to values where small perturbations can cross a threshold, changing a critical condition, and thereby discontinuously changing the response due to a perturbation. If the neighbourhoods of such critical values were rarely visited, the fact that the parametrization had discontinuities could a priori be neglected, but since the opposite is true and the nonlinearities are strong, a careful comparison of linearized and nonlinear responses is necessary.

The first application of adjoints in a meteorological model of moist convection was by Hall et al. (1982). They developed an adjoint of a convective–radiative column model and applied it over periods much longer than that of the convection. The AM-derived estimates of the effects of perturbations agreed well with actual NLM perturbed results for many of the perturbations examined. Analogous results would be extremely unlikely in a numerical weather prediction model where the dynamics are chaotic: even the smallest perturbations eventually grow in size so that their nonlinear effects become significant (Lorenz 1963; Simmons et al. 1995). Starting from perturbations the size of analysis errors, the linearizations employed in TLM or AM calculations should, therefore, eventually become unreliable for estimating the behaviour of perturbations in the NLM.

It has sometimes been assumed that a linearization of moist physics is satisfactory if, when applied in a 4-dimensional, variational, data-assimilation procedure, the minimization algorithm can reduce the cost function by a factor of 2 or so (Zupanski and Mesinger 1995; Zou and Kuo 1996; Tsuyuki 1996). This conclusion is unwarranted, however, without sufficiently detailed examination of the results; e.g. perhaps the cost function has decreased simply by fitting slow or fast gravity waves irrespective of the accuracy of the moist linearization. In fact, there are strong hints of problems that may or may not be due to linearization issues; e.g. in the paper by Tsuyuki (1996), the norm of the gradient of the cost function changes erratically with iteration number, and its net reduction at the final
iteration is only by a factor of approximately 2 compared to its first-iteration value. When
the adjoint-determined gradient fields do not adequately describe the effects of significant
perturbations, the usual algorithms used in variational analysis (e.g. conjugate-gradient
or quasi-Newton methods) may be inadequate, because the strong nonlinearity implied
indicates that the cost function may actually have multiple minima.

A series of papers by Xu (1996a,b; 1997a,b) discuss the effects of discontinuously
defined functions on TLM and AM results. The first three papers discuss primarily the
behaviour of temporally continuous models. The models investigated are simple so that
analytical solutions can be determined, but they are highly illustrative of behaviour that
would be found in more complicated, temporally continuous models. In temporally discrete
models, however, conditionals are unaltered by perturbations in the limit of infinitesimal
perturbations, unless the trajectory passes through a critical point. It has therefore been
common to ignore the changing sense of conditionals in the corresponding TLM and
AM (e.g. Zou et al. 1993). Xu (1997b) notes that such discrete models with conditionals
actually do not have correct asymptotic behaviour as the time step is reduced, implying that
the corresponding TLM and AM versions still have incorrect and undesirable behaviour
with regard to the corresponding temporally continuous model. Undesirable behaviour
may be present even with more common large time steps, however, for the reasons already
discussed.

(b) Outline

Attention here is focused on the accuracy of a linearization of moist physics applied to
a temporally discrete, nonlinear, primitive-equation model. This accuracy is primarily de-
scribed by comparisons of TLM output with differences between corresponding reference
(commonly called "control") and perturbed NLM output, when the TLM and NLM initial
conditions are perturbed identically. The accuracy of some adjoint applications will also
be described. In order not to bias the perturbation evolutions to the quasi-linear physics
describing gravity-wave propagation and small-scale diffusion (Errico et al. 1993), the
perturbations are determined using singular vectors (SVs). The determination of SVs also
highlights some problems with straightforward construction of moist TLMs.

The accuracy of TLM or AM results depends on many factors, including the exact
size and shape of the perturbation, the synoptic features in the reference state, and the
forecast duration, as well as details of the formulation of the models. Furthermore, how
accuracy should be assessed depends on the purpose of the application. For these reasons,
it is impossible to generalize results from only a few experiments. Therefore, instead
of general results, a range of accuracy and illustration of significant difficulties will be
presented here. These are intended to serve as counter examples to both pessimists (who
believe linearizations of moist physics that ignore the changing branches of conditionals
to be essentially useless) and optimists (who believe linearizations of moist physics are
straightforward with no significant restrictions to their applications).

The model used in this study is described in section 2, followed by a description of the
linearization of precipitation and convection in section 3. The winter case studied with the
current model version is presented in section 4. The two kinds of optimal perturbations used
are described in section 5. Sections 6–8 have examples of the significance of moist effects,
accurate TLM and AM results, and inaccurate results, respectively. In section 9 examples
from a convectively active summer case are presented. Examples of the inadequacy of
using straightforward linearizations of some parametrizations of moist convection are
presented in section 10, followed by a summary, general discussion, and recommendations
in section 11.
2. Description of Model

The model used here is version 2 of the limited-area, primitive-equation, Mesoscale Adjoint Modelling System (MAMS2) developed at the National Center for Atmospheric Research (NCAR) as described by Errico et al. (1994) with modifications as listed here. Its vertical coordinate is $\sigma = (p - p_\text{t})/(p - p_\text{s})$, where $p$ is pressure, $p_\text{t}$ is the model top at 10 mb, and $p_\text{s}$ is surface pressure. In all the experiments here, the wind components $(u, v)$, temperature $(T)$ and mixing ratio $(q)$ are defined on 10 levels equally spaced in values of $\sigma$. The vertical finite differencing follows the energy conserving formulation of the NCAR Community Climate Model (CCM2: Hack et al. 1993) and the vertical diffusion follows Kiehl et al. (1996). The adiabatic, vertical-mode initialization scheme of Bourke and McGregor (1983) is applied to the external and first internal modes. Radiative effects on the ground only are considered. These are influenced by column total precipitable water and fractional coverage by three levels of clouds determined as piece-wise linear functions of relative humidity.

Large-scale (i.e. non-convective) precipitation has been modelled based on a critical value of 100% relative humidity. Excess moisture is precipitated after approximately accounting for the accompanying increases in temperature and saturation vapour pressure due to condensation. Values for the specific heat of air and latent heat of evaporation are treated as $T$-independent for this calculation. No evaporation of precipitation falling below the level of condensation is considered.

Moist convection is modelled using the Relaxed Arakawa–Schubert (RAS) scheme developed by Moorthi and Suarez (1992), plus a consideration of evaporation of falling precipitation (provided by Moorthi and Suarez) and depth-dependent relaxation timescales. These time-scales range between 2.5 h for deep clouds to 1 h for shallow clouds (here a minimum depth of approximately 200 mb).

The TLM and AM linearizations of MAMS2 are exact, including all the physics, except for two approximations used in this study. The first is that the reference state about which the linearization is performed is only archived every 3 time steps (15 minutes). This state is then held fixed during the corresponding period in the TLM or AM. This approximation has negligible impact on the TLM solution, except when the perturbation size is smaller than the typical change in the reference state between consecutive time steps. The vertical-mode initialization is particularly useful for limiting high-frequency behaviour that would otherwise exceed this condition. The second approximation is that the Jacobians for moist and dry convection are computed using the perturbation method described in section 3. They are averaged every 3 time steps (as is the Jacobian exactly computed for the non-convective precipitation) for use by the TLMs and ALMs. The Jacobians for moist convection are also filtered as described in section 3.

The NLM is started from analyses produced by the European Centre For Medium-Range Weather Forecasts (ECMWF); they are archived at NCAR on a 128 × 64, global, Gaussian grid and subsequently interpolated to the MAMS2 grid. The lateral boundary conditions at subsequent times are obtained from fields linearly interpolated in time between subsequent 12-hourly analyses produced in the same way.

3. Determination and Application of MAMS2 Jacobians

The Jacobian of a parametrization scheme in any discrete model is the matrix $M$ whose elements are:

$$ m_{i,j} = \partial y_i / \partial x_j, \quad (1) $$
where $y_i$ is an element of the output of the scheme, $x_j$ is an element of its input, and the derivatives are evaluated for a reference state $\tilde{x}_j$ given by the control NLM forecast. For example: input to the RAS scheme are the 21 values of $p_s$, $T_k$, and $q_k$ and output are 21 values of convective heating rates ($H_k$) and moistening rates ($Q_k$) and convective precipitation $R_c$, where $1 \leq k \leq 10$ is the grid's vertical index. If $x$ and $y$ represent the vectors with respective elements $x_j$ and $y_i$, then TLM perturbations are related by:

$$y' = Mx' + y'(\text{previous}),$$

(2)

and AM (i.e. gradient) fields by:

$$\frac{\partial J}{\partial x} = M^\dagger \frac{\partial J}{\partial y} + \frac{\partial J}{\partial x} (\text{previous}),$$

(3)

where $J$ is a scalar measure of a forecast aspect investigated using the AM, the dagger ($\dagger$) denotes a matrix transpose, and 'previous' refers to possible contributions by previously considered physics.

The TLM and AM versions of MAMS2 apply (2) and (3) only implicitly throughout most of their software, relying instead on a line-by-line development (e.g. Giering and Kaminski 1996). Three of its parametrization schemes are exceptions; these use explicit representations of their corresponding Jacobians. In the non-convective precipitation scheme, the Jacobian is computed exactly using (1), and in both moist and dry convection schemes, the Jacobians are computed approximately using a perturbation method. These Jacobians are computed during the NLM control forecast at each time step only at grid points where the respective scheme is active (i.e. where a non-zero Jacobian is possible). The Jacobians are then averaged over three time steps and saved for subsequent use by the TLM or AM. Storing at shorter intervals produced insignificant change to the results presented here, but greatly increased both storage and input/output operations (but not the expense of calculation).

In the perturbation method, at each active point at each time, temporary vectors

$$x_j^{(p)} = \tilde{x}_j + s_j e_j$$

are defined, where $e_j$ is the vector whose elements are all zero except for the $j$th that is set to 1, $s_j$ is a scaling that depends on which field is being perturbed (so the perturbed field has an appropriately small size; in the case of $q$, perturbation size also decreases with height). The number of such perturbations is the number of elements of $x$. These $x^{(p)}$ are then input into the nonlinear parametrization scheme to compute $y^{(p)}$ and the Jacobian is computed as:

$$m_{i,j} \approx \frac{y_i^{(p)} - \bar{y}_i}{s_j}.$$  

(5)

If all the $s$ are chosen sufficiently small, (5) is usually a good approximation to (1). For nonlinear schemes the perturbation method can be cheaper for computing approximate Jacobians than computing them exactly (i.e. than by computing (1) explicitly).

The approximation (5) can result in unacceptable discrepancies if the scheme is highly nonlinear and $s$ is too large. Also, if $s$ is too small perturbed results can be sensitive to round-off errors in the calculations. The $s$ used in MAMS2 are 0.001 K and 0.01 mb for $T$ and $p_s$, respectively. The $s$ for $q$ is 0.01% of the saturation mixing ratio determined by the initial $\sigma$-surface mean $\tilde{q}_k$ and $\bar{q}_k$. Even these values, however, are at times unacceptable.
In particular, the RAS scheme has conditional functions (i.e. functions having different formulations depending on conditions 'if... then...'). Perturbations that are even smaller than the chosen $s$ can alter the sense of a condition, causing a discontinuous change in the function. A change in $y_j^{(p)} - \bar{y}_j$ can thereby be quite large even if $s_j$ is very small, and (5) can even grow as smaller $s_j$ are considered. Although for some specific applications, (5) can more accurately describe the scheme's response to a finite-sized perturbation, that will not be true in general (e.g. if the sign of the perturbation changes). Some quality control of (5) is therefore necessary to ensure that it is a reasonable approximation to (1).

At some points and times, Jacobians for the RAS scheme have eigenvalues significantly greater than 1, as described in section 10 (but for results using the Hack (1994) scheme). Since these Jacobians can describe remarkable growth, SV determinations tend to accentuate their presence; more generally, it is difficult to ascertain what impact such eigenvalues have had in a TLM or AM calculation without very close inspection. Since the applicability of TLM or AM results to NLM behaviour can be greatly reduced when such Jacobians are present, it is useful to consider additional 'quality control' of the Jacobians based on their eigenvalues.

There are two quality controls of (5) implemented for the RAS scheme. The first is to test the (time-averaged) Jacobian elements for excessive size. If any Jacobian element has a size that exceeds an implied perturbed heating or moistening rate of greater than 400 K day$^{-1}$ or 400 g kg$^{-1}$day$^{-1}$, respectively, for input perturbations of 1 K, 1 g kg$^{-1}$, or 1 mb, the Jacobian is considered suspect and it is replaced by zero. In the winter and summer cases examined, less than 0.5% and 2%, respectively, of the Jacobians were filtered by this test. Elements of some accepted Jacobians are close to the limits permitted. By recomputing the Jacobians using perturbations of different sizes and signs, these large but acceptable values have been determined to be accurate, and not an artifact of the perturbation method.

Furthermore, filtering these large but acceptable values had a detrimental effect on the agreement between TLM and NLM perturbed results. In contrast, when no filtering of the unacceptable values was performed the TLM produced extraordinary perturbation growth, beyond anything observed in the NLM for reasonably sized perturbations. The second quality control is to determine the eigenvalues of the matrix $\hat{L} = I + \Delta t\hat{M}$, where $I$ is the identity matrix, and $\hat{M}$ is the portion of the scheme's Jacobian that describes the effects of only $T$ and $q$ perturbations on only $\partial T/\partial t$ and $\partial q/\partial t$. The resolvent $\hat{L}$ describes how new values of $T$ and $q$ would be determined from previous values using a forward time scheme and considering only the physics represented by $\hat{M}$. If the modulus of the largest eigenvalue $\lambda_m$ of $\hat{L}$ is greater than a critical value $\lambda_c$, then the scheme's Jacobian is replaced by zero, because it is considered to be affected by either a discontinuity or a strong nonlinearity that renders it a poor approximation or otherwise unacceptable linearization. This filtering of the Jacobians introduces errors in the TLM, since the neglected effects of perturbations at these point are actually not only non-zero but possibly large. As long as the percentage of such discarded Jacobians is very small, however, the errors so introduced may be much less than the errors or discrepancies introduced if these Jacobians were retained. The great success of this quality control strategy will be demonstrated by the results to be presented. The Jacobian quality control is only applied to the temporally averaged Jacobians to reduce computational expenses.

Better than simply replacing unacceptable or undesirable Jacobians by zero, would be to isolate the causes of large eigenvalues in the scheme's formulation and then make appropriate changes to the scheme itself. This can be as difficult, however, as it was originally designing and tuning the scheme. Alternatively, one can simply try modifying the Jacobian; e.g. by replacing suspect large elements by the maximum permissible ones. This too, however, should not be done without a thorough evaluation of characteristics of
the changed response. Modifying a small set of Jacobian elements can easily remove the effects of constraints implied in the physics, thereby creating other problems. Although the replacement of undesirable Jacobians by zero seems drastic, the number of such replacements is sufficiently small and the results so much more useful, that we consider it adequate for the present MAMS2 applications.

A disadvantage of using (2) or (3) explicitly is that the computation of either (1) or (5) may be very expensive, and that, since they are matrices rather than vectors, many storage or data transfers may be required. These costs have not been prohibitive in MAMS research applications thus far, however, because there are fewer operating constraints than for daily weather-forecasting applications. Also several steps have been taken to make the operations as efficient as possible (including the time averaging and bit packing of Jacobian values when rigorous tests are not required).

The advantages of using the explicit Jacobian representation for some physics greatly outweigh its disadvantages in most MAMS applications. First, it allows easy application of an appropriate quality control. Second, once Jacobians are computed there is negligible computational expense in applying them using (2) or (3), and for many applications, such as for determination of SVs, the TLM and AM are repeatedly used with the same set of Jacobians. Third, when the parametrization scheme is changed, little or no redevelopment is required for the TLM or AM software. Furthermore, if the perturbation method is used, little or no redevelopment is required for the Jacobian generation software in the NLM. Often it is also useful to examine the Jacobians themselves, since their structure and properties (e.g. eigen-solutions) can reveal otherwise hidden aspects of the scheme's formulation and underlying physics (e.g. see section 10 and Fillion and Errico 1997).

4. WINTER CASE EXAMINED

The winter case examined is one of explosive marine cyclogenesis during the 24-hour period beginning 0000 UTC 14 February 1982 (hereafter denoted as $t = 0$). This case was selected because it was considered likely that perturbations the size of analysis uncertainties would not alter the cyclone to the extent that the precipitation regime would change dramatically (e.g. greatly change the location of the significant precipitation). This is important since the TLM uses the precipitation regime given by the control forecast, and does not consider, for example, changes from precipitating to non-precipitating states caused by perturbations. The moist physics, however, greatly alters the forecast, and therefore this case provides a real test of the physical linearization. This is one of the same cases investigated by Ehrendorfer and Errico (1995) and Vukičević and Errico (1993).

The model uses a $46 \times 61$, 120 km grid for this case. The 50 kPa geopotential height ($z$) for the $t = 0$ analysis and sea-level pressure ($p_a$) for both initial and final ($t = 24$-hour verification) analyses appear in Fig. 1. Note that even at this resolution the cyclone's central pressure drops from 99.0 kPa to 96.4 kPa during this period at this resolution. The corresponding 24-hour forecast for $p_a$ and accumulated precipitation appears in Fig. 2. At this time the forecast cyclone's centre is displaced about 270 km south-west of the analysed cyclone, with a pressure 96.9 kPa. The 24-hour accumulated non-convective precipitation is as large as 3.8 cm near the core of the cyclone, while convective precipitation is active (maximum accumulation 2.3 cm) behind the front trailing to the south-west. There are also regions of strong convective and non-convective precipitation in the south-west corner of the domain. This is likely to be associated with the strong topographic gradients along the coast of Mexico, but may also be influenced by some inadequacy in the numerical treatment of such strong gradients. When forecast using the dry version of MAMS2,
the central pressure of the Atlantic cyclone is only 98.6 kPa and its associated pressure gradients and geostrophic winds are about 50% weaker.

5. Selection of Perturbations

The accuracies of TLM or AM results with respect to those from NLM depend on many factors. Certainly the magnitude of the perturbation is critical, since perturbations the size of the basic state variables should yield nonlinear functions of perturbations that...
are non-negligible in the perturbed NLM. The perturbation structure is equally important, because whether it projects on growing modes or neutral modes (such as external gravity waves) or damping modes (such as very small horizontal scales) will determine if the perturbation magnitudes increase or decrease in time. This projection condition can depend significantly on the basic state; e.g. where the stable and unstable regions are. If perturbations do grow at some rate, forecast duration is also important since that helps determine how large perturbations become. Some parametrization schemes can so decrease accuracy that a linearization becomes useless, as shown in section 10. Finally, an assessment of accuracy depends on the model’s intended applications and the measures used.

All the perturbations used here are optimal perturbations in the senses to be described. These optimal perturbations have large impact relative to their initial magnitudes, therefore their size grows in some sense. Hence, this choice biases the tests on the side of obtaining low accuracy, and therefore provides a rather stringent assessment.

(a) Perturbations for TLM tests

Initial perturbations for the TLM tests are scaled SVs. These are produced using a Lanczos algorithm (provided as a pre-release by the Numerical Applications Group (NAG)) applied to MAMS2 for each case using the dry energy norm, as described in Ehrendorfer et al. (1998; their experiments number 2 and 3). Use of this norm was unsuitable in an older version of MAMS (Ehrendorfer and Errico 1995), but the current energy conserving vertical differencing scheme, and a correction to a software error in the determination of $p_s$, adjustments by the vertical-mode initialization scheme, have successfully diminished the influence of gravity waves upon it, and now render its use appropriate. For each case investigated, the first SV (denoted SV1) defines the perturbation structure of $u'$, $v'$, $T'$, and $p_s'$ (with $q' = 0$ initially) that grows the most rapidly during a prescribed period (either 24 or 48 hours here) when measured using this norm. This structure, therefore, projects only weakly on linearly behaving gravity waves or dissipating modes, and the maximum amplitudes of the perturbations increase greatly with time. Subsequent ordered SVs grow by successively smaller amounts, and are orthogonal to all other SVs with respect to this norm at both the initial and final times.

Associated with each SV is a singular value that is the square root of the corresponding eigenvalue computed by the Lanczos algorithm. That eigenvalue is denoted as $\lambda$ here and is equal to the ratio of the energy norms at the final and initial forecast times in the TLM. It is therefore a mean-squared measured of growth.

The SVs are re-normalized here by multiplying all fields by a single factor, preserving the structures of all fields as well as all linear relationships between fields. This factor is chosen to yield the result that the maximum absolute value of a perturbation of either $T$, $p_s$, $u$ or $v$ at some point is 2 K, 1 mb, or 4 m s$^{-1}$ respectively (NB only one of these conditions can be satisfied, since the shape of the SV is preserved). The normalized SVs are then scaled by a second factor, $\epsilon$, to further change the magnitude of the perturbation while preserving its shape. This $\epsilon$ may therefore be considered as a ratio of maximum perturbation size to a ‘typical’ analysis error size. Since the SVs are very localized, this $\epsilon$ is a gross overestimate of the ratio of the square root of the domain-mean squared (r.m.s.) perturbations with respect to the r.m.s. analysis uncertainty. More importantly, it is likely to be an underestimate of the ratio of the perturbation scaling with respect to an expected amplitude determined by projection of analysis error onto the SV; i.e. although $\epsilon = 1$ means the maximum value is a typical (r.m.s.) error size, only a small portion of an error would be expected to project onto any single SV, therefore the maximum for a projected error should be smaller than its typical size. In this sense, even an SV scaled with $\epsilon = 0.2$
may be considered a significant perturbation size. The maximum absolute values of all initially perturbed fields for each test will be described in the result sections.

\( (b) \) Perturbations for AM tests

Perturbations for AM tests are produced by determining structures that are optimal in the specific sense described here. First a measure \( J_i \) of a forecast aspect valid at \( t = 24 \) h is chosen, where distinct \( i \) denote different measures or aspects. Next a linearized expression for its perturbations:

\[
J'_i = \sum_j \frac{\partial J_i}{\partial y_j} y'_j,
\]

is determined, where the \( y_j \) are components of the model-output fields, and the gradient is evaluated along the trajectory of the control (basic-state) forecast. The corresponding expression:

\[
J'_i = \sum_j \frac{\partial J_i}{\partial x_j} x'_j,
\]

with respect to the model input components \( x_j \) is then determined, using the AM to compute \( \partial J_i/\partial x_j \) from \( \partial J_i/\partial y_j \). Finally, we determine initial conditions that maximize (7) given the constraint that:

\[
I = \frac{1}{2} \sum_j w_j x'_j^2
\]

for some weights \( w_j \). The solution to this is (e.g. see Rabier et al. 1996 or Oortwijn and Barkmeijer 1995):

\[
x'_i\text{(optimal)} = \frac{s}{w_j} \frac{\partial J_i}{\partial x_j},
\]

where \( s \) is determined by applying (9) to (8):

\[
s = (2I)^{-\frac{1}{2}} \left\{ \sum_j \frac{1}{w_j} \left( \frac{\partial J_i}{\partial x_j} \right)^2 \right\}^{-\frac{1}{2}}.
\]

Identically shaped structures are obtained if \( J' \) is constrained and \( I \) is minimized instead. Analogous to SVs, but for a linearized rather than quadratic \( J' \), these structures are optimal in the sense that they produce the biggest \( J'_i \) expressed by (7) for the smallest (weighted mean-squared) initial perturbation size expressed by (8).

In all the AM tests here only \( q \) is perturbed, since attention is focused on effects of moist physics. Specifically, in (8) only \( q' \) on the bottom three \( \sigma \)-levels (i.e. below approximately 70 kPa) are therefore considered, and all the \( w_j = 1 \). For this simple weighting, it is sufficient to multiply the structures \( \partial J_i/\partial q \) by a location-independent scaling constant, \( \eta \), in order to obtain the optimal perturbation. At any point where the perturbation initially yields a \( q < 0 \), the perturbation is then reduced so that \( q = 0 \) there, but no constraint is applied to initially restrict supersaturation. By perturbing only the bottom three \( \sigma \)-levels, negative or large supersaturated mixing ratios are less likely to be introduced than they would be by perturbing higher in the atmosphere. The \( \eta \) are chosen so that the maximum magnitude of \( q' \) at any point is some \( q'_{\text{max}} \), prior to application of the non-negative constraint. Although the adjustment of \( q' \) to ensure non-negative \( q \) destroys the strict optimality of the structures, those adjustments are small and \( q' \) therefore remains close to optimal.
Figure 3. Squared singular values for the leading 10 singular vectors computed using dry (D) and moist (M) versions of the linearization for the 24-hour moist control forecast of the winter case.

6. IMPORTANCE OF MOIST PHYSICS

Two examples of very significant qualitative and quantitative impacts of including moist physics in TLM and AM calculations are presented here. All the results in this section use the same basic state produced by the 24-hour, moist NLM forecast for the winter case. Comparisons are made between results produced by moist and dry versions of the linearized models (TLM or AM). The dry linearization excludes consideration of $q$ perturbations on virtual temperature, in addition to exclusion of diabatic precipitation physics. Further experiments show that the differences between moist and dry versions are due overwhelmingly to the effects of moist diabatic processes.

(a) Influence on SV determination

The squared singular values ($\lambda$) determined for the first 10 SVs appear in Fig. 3. Labels 'D' and 'M' indicate results for dry and moist linearizations, respectively. The maximum $\lambda$ for the dry linearization is 54.9, but for the moist is 322. In fact, the first 7 SVs computed using moist physics all have larger $\lambda$ than the first dry SV. An even larger discrepancy was obtained for a case of summer precipitation (Ehrendorfer et al. 1998). Since the moist SVs here exclude moisture perturbations initially, and ignore measurements of $q$ perturbations at the final time, the total possible impact of moisture is underestimated. These other impacts are examined in greater detail in Ehrendorfer et al. (1998).

The structures of the $T$ components on the $\sigma = 0.65$ level are shown in Fig. 4 for 4 distinct SVs. Figures 4(a) and (b) show results for SVs 1 and 2, respectively, produced by the dry TLM and AM; Figs. 4(c) and (d) show results for SVs 1 and 4 produced by the moist versions. The leading two SVs in the dry case are both associated with the Atlantic cyclone. For the moist versions, the leading SV that is associated with the cyclone is SV4, and the first three are associated instead with the precipitation along the coasts of Mexico.
Figure 4. The temperature perturbation, $T'$, fields on the $\sigma = 0.65$ surface at time $t = 0$ for singular vectors (SVs): (a) SV1 and (b) SV2 computed using the dry linearization; (c) SV1 and (d) SV4 computed using the moist linearization in the 24-hour moist control forecast of the winter case. The fields are scaled such that the area-mean perturbation energy is 1. The contour interval is 2 K.

These three therefore grow much less when linearization of the precipitation physics is excluded.

Even when modes associated with the same synoptic features are compared (e.g. dry SVs 1 and 2 with moist SV4), differences are notable. Moist SV4 is only weakly correlated with either dry SV1 or dry SV2, although their structures have somewhat similar shapes; its associated $\lambda$ is also double those for the dry SVs.

(b) Influence on sensitivity fields

As an example of the effects of moist physics on AM solutions, the sensitivities of a forecast aspect $J_0$ with respect to $t = 0$ initial conditions have been determined using both moist and dry AM calculations. This $J_0$ is defined as:

$$J_0 = \frac{\sum_{i,j} \sum_{k=6}^{10} \zeta_{i,j,k}(\Delta\sigma)_k}{N \sum_{k=6}^{10} (\Delta\sigma)_k},$$

(11)

where $i$, $j$ are horizontal grid-point indices of the vertical component of the relative vorticity $\zeta$ at $t = 24$ h; the sum over $i$, $j$ is computed only in the subdomain shown in Fig. 5 centred on the forecast Atlantic cyclone, and $N = 49$ is the number of grid points in that subdomain. $J_0$ is thus the vertical component of the area-mean relative vorticity at $t = 24$ h averaged in the lower half of the atmosphere (using $\Delta\sigma$ as a fractional mass weighting).
Figure 5. Sensitivity fields for $J_0$, the mass-weighted vorticity in the lower half of the atmosphere averaged within the indicated box, valid 24 hours prior to the verification time (the latter being at the end of the 24-hour, moist winter forecast): (a) $\partial J_0 / \partial T$ on $\sigma = 0.65$ computed using the dry AM; (b) $\partial J_0 / \partial T$ on $\sigma = 0.65$; (c) $\partial J_0 / \partial v$ on $\sigma = 0.85$; and (d) $\partial J_0 / \partial q$ on $\sigma = 0.95$, computed using the moist AM. Contour intervals are (a) $1 \times 10^{-8}$ s$^{-1}$K$^{-1}$, (b) $2 \times 10^{-9}$ s$^{-1}$K$^{-1}$, (c) $1 \times 10^{-8}$ s$^{-1}$ (m s$^{-1}$)$^{-1}$, and (d) $1 \times 10^{-5}$ s$^{-1}$ (kg kg$^{-1}$)$^{-1}$. See text for further explanation.

Four sensitivity fields are presented in Fig. 5. Frame (a) is for $\partial J_0 / \partial T$ computed using the dry AM. The remainder (b)-(d) are for $\partial J_0 / \partial T$, $\partial J_0 / \partial v$, and $\partial J_0 / \partial q$ computed using the moist AM. The first three are shown on the $\sigma$-levels where the respective fields have maximum magnitude, but $\partial J_0 / \partial q$ is shown on a level where perturbations are introduced in section 7(b).

First compare $\partial J_0 / \partial T$ for moist and dry AM results. Note that the structures are very similar (correlation above 0.95). The corresponding amplitudes for the moist result, however, are approximately 25% larger, indicating that the moist diabatic physics will amplify the impact of perturbations.

Next compare the three sensitivity fields produced by the moist AM. If perturbations of 2 m s$^{-1}$, 1 K, and 1 g kg$^{-1}$ for $v$, $T$, and $q$ are considered equivalent magnitudes (each a typical size for a respective analysis error), then perturbations of those sizes at the single grid-points and levels where the respective maximum sensitivities are found would yield impacts on $J_0$ of $2.6 \times 10^{-8}$ s$^{-1}$, $8.7 \times 10^{-8}$ s$^{-1}$, and $13.3 \times 10^{-8}$ s$^{-1}$, respectively. In this sense, the sensitivities with respect to $q$ and $T$ are greater than that with respect to $v$. The possible effects of $q$ perturbations should not therefore be neglected.

Additionally, only in the moist AM can a sensitivity of results explicitly dependent on $q$, such as precipitation, be determined. An example of such results, and their accuracy, will be discussed in sections 7 and 8.
Figure 6. The sea-level pressure perturbation, $p'_s$, at hour 24 for singular vector SV4 scaled by $\epsilon = 1$ produced by: (a) the moist TLM; (b) the NLM difference. Contour intervals are 0.05 kPa. See text for further explanation.

7. EXAMPLES OF ACCURATE RESULTS

Results with optimal perturbations associated with the Atlantic cyclone are described here. Although moist diabatic physics has a large effect in this region (greatly increasing the cyclone intensity), there is significant dynamical control of the behaviour.

(a) Accurate TLM results

The first example here uses SV4 with a scaling $\epsilon = 1$ in the moist TLM. The structure of the $T$ field on $\sigma = 0.65$ has been shown in Fig. 4(d) for a different scaling. This optimal structure is localized in the vicinity of the 50 kPa trough, has a westward tilt with height (not shown), and is associated with approximately balanced winds. Maximum, initial absolute values of $u'$, $v'$, $T'$, and $p'_s$ fields are 2.8 m s$^{-1}$, 0.9 m s$^{-1}$, 2.0 K, and 0.07 kPa, respectively, for the scaled SV.

In the TLM, the energy norm grows approximately exponentially in time by a factor of 97 during the 24-hour forecast for SV4. At $t = 24$ hours, the maximum absolute values of $u'$, $v'$, $T'$, $p'_s$, and $q'$ fields are 26 m s$^{-1}$, 18 m s$^{-1}$, 4.6 K, 0.42 kPa, and 0.0037 kg kg$^{-1}$, respectively. In the perturbed NLM, the energy norm grows by only a factor of 71, indicating that the corresponding perturbation amplitudes are approximately 15% smaller than in the TLM. The perturbation size is far from infinitesimal at all times.

An example of well-matched fields at hour 24 is shown in Figs. 6(a) and (b) for $p'_s$. Both structures and magnitudes agree very well: the correlation is 0.95 and the r.m.s. error due to linearization is only 30% of the r.m.s. value of $p'_s$ in the NLM. In earlier versions of MAMS, the $p_s$ field was especially prone to decorrelations when high-frequency forcing acted on the NLM perturbations due to nonlinear effects of conditionals (cf. Fig. 7 in Vukičević and Errico 1993), but in MAMS2 both structure and magnitude match well.

The least correlated fields for SV4 at hour 24 are those for the perturbed, non-convective, 24-hour accumulated precipitation $R'_p$ shown in Figs. 7(a) and (b) (correlation 0.71). These fields would be better correlated if the region of $R'_p > 0$ south of the Atlantic cyclone were present in the TLM as well as the NLM. The corresponding fields for $R'_c$ are presented in Figs. 7(c) and (d). Here the main structural discrepancy is the TLM absence of the negative perturbation that appears south-east of the maximum positive perturbation in the NLM. The discrepancies in both sets of precipitation fields are due to a change from convective to non-convective precipitation in the perturbed NLM forecast, as confirmed by examining time series of precipitation at the regions in question. These discrepancies
disappeared when the test was repeated with a smaller ($\epsilon = 0.2$) initial perturbation. Even with these discrepancies, however, the agreement between TLM and NLM results for all these fields are very good for these significant perturbations.

This experiment was also repeated with dry versions of the TLM and NLM (both control and perturbed). For these, the agreement was even better, but the amplitudes of the perturbations at $t = 24$ hours were typically half as large as in the moist case (the energy norm grows by only a factor 25 in the TLM in this dry case). The smallest correlation of any corresponding TLM and NLM fields on any $\sigma$-level is 0.94, with 0.96 a more typical value. The correlations are higher in this dry experiment primarily because the physics is more quasi-linear, but there is an additional enhancement due to the smaller size of the perturbations near the end of the forecast.

Perturbations using SV8 ($\lambda = 45$) and SV12 ($\lambda = 35$) with $\epsilon = 1$ produced correlations above 0.9 in the moist tests for almost all fields at nearly all levels. Like SV4, these are associated with the cyclone and cold front over the Atlantic. The correlations are likely to be higher than for SV4 because the perturbation size does not increase as much with time.

\textbf{(b) Accurate AM results}

Three examples of accurate AM-determined sensitivity estimates in a moist diabatic context are presented in this section. The accuracy is measured by comparing changes $\Delta J_i$.
determined as differences between corresponding $J_i$ produced by a perturbed and a control NLM forecast, with estimated changes $J'_i$ determined by (7). The $x_j$ used in (7) and in the perturbed NLM forecast are identical. The perturbations are produced as described in section 5(b).

The first example considers the $J_0$ defined by (11): the mean vorticity in the box described in section 6(b). The perturbation structure on the $\sigma = 0.75$ surface was shown in Fig. 5(d) (excluding some minor changes to ensure non-negative perturbed $q$). The ratios $\Delta J_0/J'_0$ are shown in Table 1 as a function of maximum perturbation size $q'_m$. If the small adjustments to make $q \geq 0$ initially are ignored, $J'_0 = 0.012q'_m$, where $J_0$ has units of s$^{-1}$ if $q'_m$ has units kg kg$^{-1}$.

Table 1 indicates that perturbations to $J_0$ are estimated fairly accurately for significant perturbations of optimal structures. Even a perturbation of maximum size 0.001 kg kg$^{-1}$ in a small region can change the vorticity by 11%, but the error in the AM estimate is only 15% of that (determined from $(\Delta J/J')^{-1} - 1$). The perturbed moisture primarily affects the vorticity indirectly, through moist diabatic processes, rather than directly as it would if precipitation were measured.

The two other examples considered in this section are the area-mean, 24-hour accumulated precipitation ($R_a + R_c$), denoted as $J_1$ and $J_2$ in the respective areas indicated by squares in Figs. 8(a) and (b). The respective sensitivities $\partial J_i/\partial q$ on $\sigma = 0.75$ at $t = 0$ are also shown in Figs. 8(a) and (b). They are predominantly positive valued and located near their forecast-verification regions. The ratios $\Delta J_i/J'_i$ for $i = 1, 2$ are shown in Table 1 as a function of maximum perturbation size $q'_m$. The AM-determined estimates are $J'_1 = 88.1q'_m$ and $J'_2 = 95.0q'_m$, where the $J$ have units of cm when $q'_m$ has units kg kg$^{-1}$. These estimates are accurate to the indicated precision even accounting for the adjustments to satisfy $q \geq 0$ initially.

For $J_1$, the precipitation is primarily non-convective with the required supersaturation driven by large-scale vertical advection of $q$. Even when the maximum size of a perturbation is 0.003 at the most sensitive location, the error in the AM estimate is less than 5% of the true change in $J_1$.

For $J_2$, the precipitation is primarily convective. Even when $q'_m = 0.003$ and the resulting $J'_2$ is 20% of the control $J_2$, the AM estimate has less than an 11% error with respect to the true change.

<table>
<thead>
<tr>
<th>$q'_m$</th>
<th>$-0.002$</th>
<th>$0.001$</th>
<th>$0.002$</th>
<th>$0.003$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta J_0/J'_0$</td>
<td>0.91</td>
<td>0.86</td>
<td>0.70</td>
<td>0.59</td>
</tr>
<tr>
<td>$\Delta J_1/J'_1$</td>
<td>0.94</td>
<td>1.02</td>
<td>1.03</td>
<td>1.03</td>
</tr>
<tr>
<td>$\Delta J_2/J'_2$</td>
<td>0.87</td>
<td>1.05</td>
<td>1.12</td>
<td>1.16</td>
</tr>
<tr>
<td>$\Delta J_3/J'_3$</td>
<td>0.73</td>
<td>2.18</td>
<td>2.59</td>
<td>3.03</td>
</tr>
</tbody>
</table>

Ratios are given as functions of the maximum size $q'_m$ (units of kg kg$^{-1}$) of the moisture perturbation at any point. The perturbation structures are optimal ones determined as described in section 5(b). The measure $J_0$ is defined in section 6, $J_1$ and $J_2$ in section 7(b), and $J_3$ in section 8(b).
Figure 8. The sensitivity of 24-hour accumulated precipitation, averaged within the boxed areas shown in the figure, with respect to possible mixing ratio, \( q \), perturbations on the \( \sigma = 0.75 \) surface at the initial time (24 hours prior to the verification time): (a) for the region of predominately non-convective precipitation, \( J_1 \), associated with an Atlantic storm; (b) for the region of predominately convective precipitation, \( J_2 \), associated with an Atlantic cold front. Contour intervals are 0.2 cm. See text for further details.

(c) Test for a 48-hour period

The accuracy of the moist TLM for structures associated with the Atlantic cyclone and optimized over a 24-hour period is sufficiently good for a 48-hour forecast to be attempted starting 24 hours earlier (and therefore ending the same time). The initial perturbation was provided by SV1 determined for this 48-hour period, but scaled with \( \epsilon = 0.5 \). This smaller scaling was considered necessary because, due to the longer forecast period, the energy norm grows by a factor of 637, and the maximum NLM-perturbed magnitudes (including the scaling factor of 0.5) at \( t = 48 \) hours for \( p', u', v', T' \) and \( q' \) are 0.4 kPa, 10 m s\(^{-1}\), 15 m s\(^{-1}\), 4 K, and 0.003 kg kg\(^{-1}\), respectively. The correlations of the corresponding TLM fields are all greater than 0.6, except for \( u' \) at level \( \sigma = 0.55 \) (not shown) for which the correlation is only 0.5. Although the \( u' \) fields are associated with the cyclone in both the TLM and NLM cases, they are notably different in structure as well as magnitude. Indeed, the magnitudes of most TLM fields at most levels are approximately 60% larger than the corresponding NLM fields.

If \( \epsilon = 1 \) is used for the 48-hour SV1, then the correlations drop by approximately 0.1 for all fields and levels, and the TLM magnitudes are typically twice the corresponding ones in the NLM perturbation. On the other hand, if \( \epsilon = 0.2 \) is used, the correlations typically increase by approximately 0.1 for most fields and the magnitude differences are less than 30%.

These discrepancies for the 48-hour test can be considered significantly large or small, depending on the applications intended for the TLM (or corresponding AM) results. For many quantitative applications, these linearized estimates of behaviour are probably inadequate, but qualitatively they may remain extremely useful.

8. Examples of inaccurate results

Results with optimal perturbations associated with the precipitation near Mexico are described here. Moist diabatic forcing appears to dominate the physics there.

(a) Inaccurate TLM results

The structure of the \( T' \) field on level \( \sigma = 0.65 \) for SV1 has been shown in Fig. 4(c). For this test, a scaling of \( \epsilon = 0.2 \) is used. The maximum initial magnitudes of the \( T' \),
p', u', and v' fields are 0.26 K, 0.02 kPa, 0.1 m s⁻¹, and 0.1 m s⁻¹, respectively. This small initial scaling was chosen because the energy norm grows by a factor of 322 for this perturbation in the moist TLM so that, even with ε = 0.2, by hour 24 the maximum values of the respective perturbation fields are 3.5 K, 0.09 kPa, 5 m s⁻¹, and 5 m s⁻¹. Most of the perturbation growth occurs during the final 6 hours, and near the end of the forecast the energy norm is increasing by as much as 7% each time step. This rapid growth is not inconsistent with possible amplification of the energy norm due to precipitation physics (Ehrendorfer et al. 1998).

The $R'_n$ and $u'$ (at $\sigma = 0.55$) fields produced by the NLM differences at hour 24 for this scaled SV1 appear in Fig. 9. Note that $R'_n$ is dominated by a single point. $R'_c$ (not shown) is also dominated by a single point, located just to the west of that for $R'_n$, with the maximum values for both fields approximately the same (2.3 cm in 24 hours). In the lower troposphere, both the $u'$ and $v'$ fields are associated with wind convergence into the region where the perturbed precipitation is concentrated. In the middle to upper troposphere (e.g. Fig. 9(b)), these fields describe wind divergence. These wind perturbations are therefore properly structured to strongly perturb precipitation.

Correlations between corresponding TLM and NLM perturbation fields on individual $\sigma$-surfaces for this SV1 test vary between 0.95 and 0.40, except for the $T'$ field on $\sigma = .45$ which is correlated at −0.77. An abrupt transition from dominantly positive to negative values in the TLM $T'$ field occurs at 50 kPa as one moves upwards, but in the NLM perturbation it occurs at 40 kPa, resulting in negatively correlated fields at $\sigma = .45$. Lower and upper tropospheric winds are best correlated but their TLM magnitudes are half those of the corresponding NLM fields. The $R'_n$ field in the TLM looks like that in the NLM (Fig. 9(a); a correlation of 0.97) but with one tenth of the magnitude. The $R'_c$ field for the TLM has large values at two adjacent points where $R'_c$ and $R'_n$ are largest in the NLM, with magnitudes at both points similar to that of $R'_c$ in the NLM, making the correlation 0.5. In the NLM, the energy norm grows by a factor three times that in the TLM.

Consideration of the patterns and values of $R'_n$ and $R'_c$ in the TLM and NLM, as well as examination of time series of precipitation at individual points, reveals the reasons for the discrepancy between TLM and NLM results. In the control NLM forecast, non-
convective precipitation actually ceases after hour 20 at the point shown in Fig. 9(a). The TLM non-convective precipitation, therefore, also ceases to operate there after that time since it depends on the sense of the conditionals prescribed by the control forecast. The NLM perturbation, on the other hand, continues to produce non-convective precipitation there at even faster rates than earlier. The NLM perturbation of \( R_k \) is thereby several times larger than the TLM one. In the NLM perturbed forecast, convection at this same point does not occur until the final 30 minutes, when it is stronger than in the control so that the accumulated change is near zero. The NLM perturbation there is, therefore, primarily due to the sense of a conditional changing, rather than a modulation of existing precipitation as modelled by the TLM. At the point just to the west, where \( R_k \) is largest in the NLM, convection occurs strongly in both perturbed and control forecasts and therefore in the TLM. The vertical structure and intensity of the convection are very different, however, with heating rates 5-times larger and peaking lower in the atmosphere in the perturbed NLM. The TLM also creates an enhancement of convection there, but only a third as great. It is possible that the changes in convection in the perturbed NLM forecast there are also due to conditionals changing sense, aside from those determining the existence and depth of convection, but this has not been investigated.

Forecasts of the precipitation near Mexico may be very inaccurate since the lateral boundary is so close in this winter case, although inadequate discretization in the presence of large topographic gradients is likely to be an equal or worse source of error. Neither of these conditions, however, necessarily affects the accuracy of TLM and NLM, or AM and NLM, perturbation comparisons. In particular, unless some critical character of the reference state is affected, the proximity of the lateral boundary is likely to have no impact on the SV1 result, since its horizontal-scale during its growth phase extends only a few grid points.

Perturbations using SV2 and SV3 were expected to yield similar results to those of SV1, since they also grow considerably and are associated with precipitation over Mexico. Perturbations using SV6 and SV11 with \( \epsilon = 1 \) were examined. These are associated with the same synoptic feature, but their norms grow by factors of only 60 and 35, respectively. Correlations with the corresponding NLM results were similar to, or better than, those using SV1 having \( \epsilon = 0.2 \). Unlike for SV1, the magnitudes of corresponding TLM and NLM perturbation fields were similar for these other SVs.

One test was performed in which the initial perturbation was specified as one-fourth the difference between the \( t = 0 \) analysis and a MAMS2 forecast from 12 hours earlier valid at the same time. The magnitude of this perturbation was like that for the SV tests with \( \epsilon = 1 \), but the perturbation was significant throughout the domain rather than localized as for the SVs. The 24-hour TLM and NLM comparison indicated very good agreement everywhere except in the south-west quadrant of the domain. The fact that TLM forecasts with some SVs in this region yield poor accuracy, reveals the difficulty of obtaining good TLM results with more general, but still significant, perturbations.

(b) An inaccurate AM result

The example chosen here considers a \( J_2 \) equal to the area-mean 24-hour accumulated precipitation, where the averaging is done in the box shown in Fig. 10. This is the same \( J \) as considered in section 7(b), except for the location of the verification box. The sensitivity \( \partial J_2 / \partial q \) on \( \sigma = 0.75 \) is also shown in Fig. 10. The sensitivity is predominately positive.

The AM estimates are \( J_2 = 105q_m \). Their ratios with respect to the NLM values appear in Table 1. For \( q_m \) as large as 0.002, errors are greater than 50% of the NLM values. This agrees with the general result that TLM comparisons with NLM perturbations were poor when the leading SVs in this area were investigated.
9. RESULTS FOR A SUMMER CASE

Linearization of a summer case has also been examined. This case begins at 1200 UTC 16 June 1989, and ends 24 hours later. The model uses a $54 \times 43$, 80 km grid for this case. The $p_d$ and $R$ fields at $t = 24$ hours appear in Figs. 11(a) and (b), respectively. The dynamic fields are much weaker and more stationary than in the winter case. The values and extent of precipitation agree very well with observations (not shown). Most of it is convective, except for two weak regions of non-convective precipitation in southern Manitoba and south-east Ontario provinces.

For the summer case with moist linearization, $\lambda = 2548$ for SV1. The $t = 0$ fields (not shown) are localized over north-western Wyoming, south-western Montana, and south-eastern Idaho. With a scaling $\epsilon = 0.2$, the maximum initial magnitudes of the $T'$, $p'_d$, $u'$, and $v'$ fields are 0.4 K, 0.02 kPa, 0.3 m s$^{-1}$, and 0.2 m s$^{-1}$, respectively. By hour 24 in the NLM, these have increased to 3 K, 0.037 kPa, 8 m s$^{-1}$, and 7 m s$^{-1}$. The value of $\lambda$ determined for this perturbation evolving in the NLM is only 1404, indicating the TLM has over-estimated the growth, although it is still considerable in the NLM.

Corresponding $R'_c$ fields at $t = 24$ hours for this perturbation are shown in Fig. 12. Both structures are parallel (positive and negative) bands oriented 30° east of north, but
their correlation is only 0.06 because neither their band axes nor their maxima and minima actually coincide. The maximum absolute values of $R'_c$ for the TLM and NLM results are 1.4 cm and 1.9 cm, respectively, unlike the relationship for the dynamic fields that have the NLM results 60% smaller than for the TLM. Unlike for the convectively driven SV1 in the winter case, the final wind perturbation is not dominated by high-level divergence above and low-level convergence.

For this summer case, SV3 having $\lambda = 292$ was also examined using $\epsilon = 0.2$. At $t = 24$ hours corresponding values for the dynamic fields in the TLM are approximately 30% smaller than in the NLM, which is opposite to the relationship for SV1. The $R'_c$ TLM field is approximately twice as large as its NLM counterpart, but their correlation is 0.71.

The accuracies measured for this summer case with initial scalings of $\epsilon = 0.2$ are very good considering the extreme growth rates and large effects of moist physics. For much larger values of $\epsilon$, the agreement of perturbation magnitudes at hour 24 must get worse, since the TLM results are proportional to $\epsilon$ and the perturbations at the end of the forecast with $\epsilon = 0.2$ are already nearly the size of variations in the reference state. Even when correlations or magnitudes are poor, however, the shapes and locations of structures often appear qualitatively similar in TLM and NLM results (shifted enough, however, to decorrelate them). For some applications this qualitative agreement may be sufficient; e.g. when determining significantly (rather than necessarily optimally) growing structures. For applications where accurate magnitudes are critical, however, the moist linearization may be inadequate.

10. Previous results

The first version of MAMS was based on the model described by Anthes et al. (1987) denoted as MM4. An attempt to use the linearized version of the MM4 Anthes (1977) version of the Kuo (1974) convection scheme in MAMS was quickly abandoned. In that scheme the conditionals were of the worse kind, determining the moisture tendency according to:

$$\frac{\partial q}{\partial t} = \begin{cases} \text{advection} & \text{if not convective,} \\ \text{cloud model} & \text{if convective,} \end{cases}$$

(12)

where convection only occurred when several conditions were satisfied simultaneously. The cloud model used no information about the advection other than the moisture accession
(vertical integrated moisture convergence). The vertical profiles of moisture tendencies, therefore, became radically different when the branch in (12) changed.

Often the branch changing in (12) occurred very frequently, as when a two time-step oscillation between these branches occurred (Vukićević and Errico 1993; Zou and Kuo 1996). The primary detrimental effect caused by this oscillation was that it excited gravity waves in the nonlinear model, the phases of which depended on the specific time steps at which convection occurred. Since a linearized version of the scheme that did not explicitly account for the changing sense of conditionals did not model the change of these phases, it also did not simulate the change of the phases of the excited gravity waves. A significant portion of the NLM difference field was therefore missing from the TLM integrations (Vukićević and Errico 1993).

The first documented version of MAMS (MAMS1—version 1; Errico et al., 1994) used the Hack (1994) moist convective, mass flux scheme that was extensively used at NCAR. It first appeared that the TLM and corresponding AM versions of the scheme behaved well with respect to the NLM. In particular, no two time-step oscillations were observed. Some ‘strange’ behaviour was observed, however, such as intermittent, very fast changes of norms in TLM forecasts with SVs. Then, as more cases were examined, extremely pathological behaviour was encountered; e.g. increases in components of adjoint fields by factors of $10^4$ in 3 hours!

An example of the large TLM growth in MAMS1 with Hack’s scheme appears in Fig. 13. It shows time series of $p'_0$ at a single point, produced by three forecasts using a singular vector from a moist version of MAMS1 applied to the same case as in section 4. Note that one of the curves has a different scale. The black solid and dashed curves are for corresponding TLM and NLM forecasts, respectively, using $\epsilon = 1$. For the first 8 hours the curves are very similar, showing equal decay. At hour 9 the TLM value increases so
dramatically that it was originally considered indicative of an error. Although the TLM peak dropped by hour 11, the result of this and some large growth periods at other locations was to increase the magnitude of the TLM fields by a factor of two in 24 hours compared with the NLM result. When the NLM was run with the same SV and scaling reduced to $\varepsilon = 0.25$, however, it yielded a curve (not shown) that also had a steep increase in $p_i$ near hour 9. Yet another NLM forecast with $\varepsilon = 0.017$, shown as the grey curve in Fig. 13, yields very close agreement with the TLM. The TLM result was therefore not an error, and its Jacobians for the Hack (1994) scheme were accurately produced using the perturbation method.

The only way this large growth could have occurred, barring model software errors or numerical instabilities, was due to a property of the convective Jacobians, either themselves, or in conjunction with Jacobians of other processes. Once the eigen-structures of the convective Jacobians for the Hack (1994) scheme were examined, it was quickly determined that the large growth rates were indeed a property of the linearized convective scheme itself. Essentially, the Jacobians sometimes had eigenvalues as large as 1.5 in the cases examined, which meant that in every 2-minute time step in this experiment there were structures that could grow by 50%. This Jacobian property would sometimes persist for an hour or more, leading to a growth factor of more than $10^4$. Although this accurately depicted the behaviour of very small perturbations, the same growth was applied to perturbations of all sizes in the TLM and AM applications. In contrast, once these perturbations were sufficiently large in the NLM the scheme responded, both nonlinearly and by changing the sense of conditionals, so that the growth was abated and fields remained realistic. Since these nonlinear or discontinuous adjustment processes were absent in the TLM and AM, when these Jacobians were encountered the behaviours of relevant perturbations were described very poorly. These undesirable Jacobians were rare (approximately 5% of the convective points had Jacobians with eigenvalues greater than 1.05 at any time) and usually only one (and rarely 2) eigenvalues for the Jacobian at a single location caused problems. The size of a perturbation generated by these large eigenvalues depended on the initial amplitude of the corresponding eigenvectors, so that with random perturbations these bad effects were usually hardly discernable. At other times, however, the results were useless; e.g. when values of TLM or AM fields were made many orders of magnitude greater than any possible corresponding NLM result would imply.

Rather than considering the Hack scheme further, it was replaced by the RAS scheme of Moorthi and Suarez (1992), which was also used at NCAR. Modification of the Hack scheme to reduce the largest eigenvalues is likely to have been difficult because there was nothing discernable in the Jacobian itself, such as unusually large elements, that indicated an unacceptably large eigenvalue; furthermore any changes to the scheme would have had to be made without degrading either its accuracy in NLM predictions or the physical basis of its formulation. It was hoped that the RAS scheme of Moorthi and Suarez (1992) would have better properties. Additionally, it was expected that a relaxed scheme would tend to preserve the sense of its conditionals over a longer period of time than the non-relaxed Hack (1994) scheme, and therefore also not have the kind of oscillations found with the Anthes (1977) scheme. The results in sections 4 to 9 show the successful use of the RAS scheme.

11. CONCLUSIONS AND RECOMMENDATIONS

Inclusion of linearized, moist diabatic physics essentially modulated some moist TLM or AM results with respect to dry ones. For example, SVs computed for the same reference state grew twice as much in the moist TLM as in the dry one when measured by an energy
norm, although they were associated with the same synoptic feature and had similarly shaped structures. Also, corresponding sensitivities of forecast vorticity averaged within an explosive cyclone with respect to fields 24 hours earlier had nearly identical structures but 30% greater amplitude in the moist AM than in the dry AM.

Other significant moist TLM or AM results were found, however, that had no counterpart in the dry results. In particular, perturbation structures were found that grew many times more rapidly in the moist than the dry, TLM. These were sometimes in regions far removed from the dominant dry results, and were clearly a consequence of moist diabatic physics (see also Ehrendorfer et al. 1998). In the examination of forecast vorticity using the moist AM, the maximum sensitivity with respect to initial moisture was relatively larger than magnitudes for any other fields for single, grid-point perturbations the size of analysis uncertainty. Of course, only in the moist TLM or AM could measures of the direct effects of moist physics, such as precipitation amounts, be explicitly estimated. Since the inclusion of a linearization of moist diabatic physics can so greatly affect many TLM or AM results, beyond a simple modulation, it should be included in operational forecast models to accurately estimate behaviours.

The results demonstrate that it is possible to obtain useful tangent linear and adjoint results in a model with active, highly nonlinear, moist physics when perturbations are the size of analysis errors and project onto growing structures. Even when moist effects are important, as in a midlatitude explosive cyclone fuelled by moist diabatic processes, TLM or AM results can be highly accurate. Counter examples were also obtained, however, particularly when the synoptics were governed primarily by convection and the underlying dynamics was weak. Equally poor results should be expected with many convection schemes applied to the tropics or other, continental summer cases. Even the poor results, however, may not be useless in some applications; e.g. in the summer cases examined the TLM and NLM perturbation structures looked similar, although sometimes shifted a small distance which formally decorrelated them. Also, although the convectively driven, leading SVs usually grew at different rates in the TLM and NLM, growth was substantial in both cases. Although the maximum growth in the NLM may not have been found by the linear analysis, exceptionally large growth was found.

Significant discrepancy between linearized and nonlinear results can obviously occur due to conditionals changing their senses (e.g. convection turning off or on at different locations and times). This can be exacerbated by quasi-linear aspects of convection that sometimes yield very large effects. In the nonlinear model these large effects lead to changing conditionals, unlike in the linearized models. In fact, with one convection scheme we found it necessary to limit this strong quasi-linear behaviour in the linearized models by filtering the convective Jacobians. Without this filtering the results were sometimes worthless (e.g. adjoint-derived sensitivities $10^4$ times the maximum magnitudes otherwise obtained). Although these Jacobian pathologies are not ubiquitous, they can occur frequently enough with some schemes for any results obtained without appropriate examination and quality control of the Jacobians to be suspect.

Since a successful physical parametrization scheme must pass many tests for its applications in the nonlinear model, acceptable modifications to also produce useful Jacobians may be difficult or impossible. For this reason, such modifications should be part of the general development effort for the scheme. Analysis of a scheme's Jacobian requires careful examination of details that may otherwise escape attention. The net result should, therefore, at least be a better understanding of a scheme's behaviour, if not a significant improvement of the scheme itself; e.g. the Anthes (1977) scheme would be unlikely to have been implemented as it was in Anthes et al. (1987) if the behaviour and effects of the conditionals from time step to time step had been examined beforehand.
In operational applications at numerical weather prediction centres, the perturbation method used to explicitly compute Jacobians of the MAMS2 convection schemes may be deemed too computationally inefficient. Even storage of the Jacobians may be impossible on some machine architectures. For development purposes, the perturbation method should be incorporated first if possible, however, since the linearized scheme may be effectively useless. Examining the Jacobian of the scheme and its responses using this simple approach may reveal problems before any significant development time is wasted.

When any linearized version of a complicated parametrization scheme is developed, the eigen-structure of its Jacobian should be examined for a wide range of reference states. Also, results with linearized models (TLM or AM) should be compared with corresponding results in the nonlinear model using perturbations having the structures and magnitudes of the character that will be used in the models' applications. SVs are particularly useful for ferreting out problems, since any linearized aspect that causes inordinate growth will be found. Care should be taken to examine actual fields, not just global measures of agreement. There are enough reasons to suspect the usefulness of linearizations of moist physics, to assert that their accuracy with respect to meaningful perturbations should be reported prior to any presentation of their applications.

ACKNOWLEDGEMENTS

The authors thank Martin Ehrendorfer, Rolf Langland, and Luc Fillion for their reviews of the manuscript plus two anonymous reviewers for their comments. The SV results presented here were from a study the authors conducted with Martin Ehrendorfer. The Lanczos algorithm used for the eigencomputations was written by Professor Beresford N. Parlett, University of California, Berkeley and provided by the Numerical Algorithms Group through the courtesy of Karl E. Knapp and Jeremy Du Croz, with significant modifications provided by Martin Ehrendorfer. Kevin Raeder is partially funded by US Navy contract N00173-96-MP-00152 and NASA contract W-18,077, Mod. 6.

REFERENCES


Courtier, P. and Talagrand, O. 1990 Variational assimilation of meteorological observations with the direct and adjoint shallow-water equations. *Tellus* 42A, 531–549


Errico, R. M., Raeder, K. and Vukičević, T.

1994  ‘Mesoscale adjoint modeling system version 1’. NCAR Technical Note, NCAR/TN-410+IA (available from the National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000, USA)

Fillion, L. and Errico, R. M.


Giering, R. and Kaminski, T.

1996  ‘Recipes for adjoint code construction’. Internal Report 212, Max-Planck Institute fur Meteorologie, Hamburg, Germany

Hack, J. J.


1993  ‘Description of the NCAR community climate model (CCM2). NCAR Technical Note, NCAR/TN-382+ STR (available from NCAR, P.O. Box 3000, Boulder, CO 80307-3000, USA)

Hall, M. C. G., Cacuci, D. G. and Schlesinger, M. E.


1996  Description of the NCAR Community Climate Model (CCM3). NCAR Technical Note, NCAR/TN-420+ STR (available from NCAR, P.O. Box 3000, Boulder, CO 80307-3000, USA)

Kuo, H. L.

1974  Further studies of the parameterization of the influence of cumulus convection on large-scale flow. J. Atmos. Sci., 31, 1232–1240

Kuo, Y.-H. and Low-Nam, S.


Lacarra, J.-F. and Talagrand, O.

1988  Short-range evolution of small perturbations in a barotropic model. Tellus 40A, 81–95

Langland, R. H., Ellisberry, R. L. and Errico, R. M.


Lorenz, E. N.


Moorthi, S. and Suarez, M. J.


Oortwijn, J. and Barkmeijer, J.

1995  Perturbations that optimally trigger weather regimes. J. Atmos. Sci., 52, 3952–3944

Park, S. K. and Droegemeier, K. K.

1995  ‘Effect of a microphysical parameterization on the evolution of linear perturbation in a convective cloud model’. Pp. 474–479 in Preprints from the Conference on cloud physics, Dallas, American Meteorological Society, Boston, USA

Rabier, F., Klinker, E., Courtier, P. and Hollingsworth, A.


Simmons, A. J., Murcray, R. and Petrihagić, T.


Tsuyuki, T.


Verlinde, J. and Cotton, W. R.

1993  Fitting microphysical observations of nonsteady convective clouds to a numerical model: An application of the adjoint technique of data assimilation to a kinematic model. Mon. Weather Rev., 121, 2776–2793

Vukičević, T. and Errico, R. M.

1993  Linearization and adjoint of parameterized moist diabatic processes. Tellus, 45A, 493–510

Xu, Q.


