# Development of a tangent linear model (version 1.0) for the High-Order Method Modelling Environment dynamical core

**S. Kim[1], B.–J. Jung[1] and Y. Jo[1]**

[1]{Korea Institute of Atmospheric Prediction Systems, Seoul, South Korea}

Correspondence to: B.–J. Jung (bj.jung@kiaps.org)

## Abstract

We describe development and validation of a tangent linear model for the High-Order Method Modelling Environment, the default dynamical core in the Community Atmosphere Model and the Community Earth System Model that solves a primitive hydrostatic equation using a spectral element method. A tangent linear model is primarily intended to approximate the evolution of perturbations generated by a nonlinear model, provides a computationally efficient way to calculate a nonlinear model trajectory for a short time range, and serves as an intermediate step to write and test adjoint models, as the forward model in the incremental approach to 4-dimensional variational data assimilation, and as a tool for stability analysis. Each module in the tangent linear model (version 1.0) is linearized by hands-on derivations, and is validated by the Taylor-Lagrange formula. The linearity checks confirm all modules correctly developed, and the field results of the tangent linear modules converge to the difference field of two nonlinear modules as the magnitude of the initial perturbation is sequentially reduced. Also, experiments for stable integration of the tangent linear model (version 1.0) show that the linear model is also suitable with an extended time step size compared to the time step of the nonlinear model without reducing spatial resolution, or increasing further computational cost. Although the scope of the current implementation leaves room for a set of natural extensions, the results and diagnostic tools presented here should provide guidance for further development of the next generation of the tangent linear model, the corresponding adjoint model, and 4-dimensional variational data assimilation, with respect to resolution changes and improvements in linearized physics and dynamics.

## 1 Introduction

It has long been recognized that data assimilation (DA) schemes play a key role in numerical weather prediction (NWP) systems to correctly forecast short-range predictions. Among various DA schemes, 4 dimensional variational DA (4DVar) methods have shown superior forecasting results. In addition, a recent advent of fast multiprocessor computers leads the full potential of 4DVar to be realized in more complicated systems. 4DVar schemes including Incremental 4DVar (Courtier et al., 1994), Weak 4DVar (Yannick, 2007), and Direct/InDriect Representer methods (Bennett, 2002) generally all share the common components such as a tangent linear model (TLM), its adjoint model (ADM), a background error covariance, and minimization algorithms as 4DVar drivers.

For operational NWP applications, the construction of a TLM is a very important, intermediate step in the development of the 4DVar. The TLM serves as an intermediate step to write and test the ADM, as the forward model in the incremental approach to 4DVar, and as a tool for stability analysis (Zhu and Kamachi, 2000; Ehrendorfer and Errico, 1995). It is essential for development of the 4DVar schemes to obtain consistency between the nonlinear model and its corresponding TLM that leads to the accurate development of its ADM, which plays a key role in finding a best initial condition by providing the gradient of the cost functional via minimization algorithms in the 4DVar schemes. So, the TLM has been recognized as powerful tools for analysing numerous aspects such as model sensitivity and the dynamics of flow fields, and the evolution of perturbations.

The main focus of this study is the development of a TLM for a nonlinear dynamical model that solves a primitive hydrostatic equation. The nonlinear model adopted here is the High Order Method Modeling Environment (HOMME, www.homme.ucar.edu). The HOMME is a high-order method that utilizes fully unstructured quadrilateral based finite element meshes on the sphere, and adopts a spectral element and discontinuous Galerkin method (Dennis et al., 2012). For its scalability and efficiency, the HOMME is considered as a promising dynamical core, and is the default dynamical core of the Community Atmosphere Model (CAM), and the community Earth System Model (CESM). Here, we developed a TLM for the HOMME dynamical core that can describe well the evolution of perturbations generated by the nonlinear model when the magnitude of perturbation becomes the size of actual uncertainties (Errico and Raeder, 1999).

The second section explains the TLM development for the HOMME model including the description of the HOMME, time increment with management of temporal trajectories for the nonlinear model, and linearity checks. The third section shows the numerical results of the linearity checks for all tangent linear modules, including full fields for baroclinic instabilities of time dependent zonal geostrophic flow, followed by a summary and discussion in the fourth section.

## 2 Development of tangent linear model

There are a couple of different ways to develop a TLM for a given dynamical model such as 1) a perturbation forecasting approach in which the TLM is discretized from the linearization of the given nonlinear dynamical equation, and 2) a line-by-line approach in which the TLM is linearized directly from the numerical codes of the given dynamical model. The advantage of the former is that the approach can easily deal with numerical instability than the latter, but the TLM can be more conveniently developed by the latter approach. Here, the line-by-line approach for the TLM development is adopted because of its straightforwardness of linearization for the set of the discretized nonlinear equations. The complete source codes of the described modules are available from the authors upon request.

### 2.1 HOMME dynamical core

The HOMME is a high-order element-based method to build scalable, accurate, and conservative atmospheric general circulation models that numerically solves 3-dimensional primitive equations (Nair and Tufo, 2007). HOMME employs advanced time stepping, adaptive mesh refinement and several domain decomposition strategies along with the continuous/discontinuous Galerkin (CG/DG) and spectral element (SE) methods (Thomas and Loft, 2002; Dennis et al., 2012). Also, HOMME guarantees conservation and to maintains all the attractive computational features of SE. Among the various horizontal discretization methods within HOMME, the TLM development is targeted for CG method in this study.

The numerical configuration for HOMME and its TLM share the same numerical configuration. HOMME can be configured to solve the shallow water or the dry/moist primitive equations. The baroclinic test case (Jablonowski and Williamson, 2006) configured in HOMME is utilized to appraise the evolution of baroclinic waves in the northern hemisphere using quasi-realistic initial conditions, and employs the second order explicit

Runge-Kutta time integration. The computational domain is the global sphere that is covered by six identical regions by an equiangular central projection of the faces of an inscribed cube. Each face of the cubed-sphere is free of singularities, and is partitioned into $N_e$ by $N_e$ rectangular non-overlapping elements (so, the total number of elements is 6 x $N_e^2$). For each element of the computational domain, an approximate solution is expanded by a tensor-product of Lagrange basis function of order $N_p$ defined at the Gauss-Lobatto-Legendre (GLL) points. For this study, the conservative 3 dimensional CG model is configured for the global sphere with $N_e$ = 16, $N_p$ = 4, and the horizontal resolution of 26 Lagrangian surfaces (i.e., the number of vertical levels $N_{lev}$ = 26). Then, the total number of the elements is $N_{elem}$ = 1536, and the grid resolution over the equatorial nodes is about 220 *km*, on average. A 4$^{th}$ order hyper-viscosity filter is used for spatial filtering, and the time increment is $\Delta t$ =150 *s*. Note that although the HOMME uses adaptive time stepping and adaptive mesh refinement, its TLM does not include such functions. MPI domain decomposition through the space-filling curve approach is used for parallelism (Nair et al., 2009).

The evolution of the baroclinic wave is very slow from integration day 0 to day 4. So, Fig. 1 only shows the triggering baroclinic waves and corresponding surface pressure $P_s$ and temperature field T at 850 *hPa* ($N_{lev}$ = 23) from day 6 to day 10. At days 6 and 7 the surface pressure shows few weak high and low pressure systems with shadings and the temperature field exhibits the growth of very small-amplitude waves with contours (Fig. 1-a,b). At day 8 the baroclinic instability waves are well developed in surface pressure, and the temperature waves are also clearly observed (Fig. 1-c). The baroclinic pressure waves become strong at days 9 and10, and the waves in the temperature field are almost peaked and are beginning to wrap around the trailing fronts (Fig. 1-d,e).

## 2.2 Line-by-line approach

The line-by-line approach is the easiest way to construct a TLM in that each line of the nonlinear code is rewritten to the corresponding tangent linear code via the chain rule of the implicit derivative. In general, we follow the steps below for the model linearization (Zou et al., 1997; Giering and Kaminski, 1998).

1. Determine input and output for variables and constants in the nonlinear codes.
2. Distinguish the variables for the tangent linear codes from those coefficients for nonlinear results by adding prefix "*tl_*".

3. Linearize the nonlinear codes via the chain rule of the implicit derivative (or calculus of variation).

4. Check and clean up input and output variables in the module name.

In Fig. 2, input and output for the variables in both nonlinear (NL) and tangent linear (TL) codes are indicated by intent(in) and intent(out). The variables for the NL code are ***a***, ***b*** and ***tens***, while the variables for the TL code are appended with prefix "***tl_***" and the variables ***a*** and ***b*** in the NL code are used as the coefficients in the TL code. The coefficients are generally called time varying basic states in the TL code.

In the NL code, the intrinsic *sine* function with independent variable ***a*** can be differentiated with respect to the variable ***a*** via the chain rule of the implicit derivative. Then, the *sine* function is differentiated to be the *cosine* function and its variable ***a*** becomes ***tl_a***, the variables of the tangent linear code. To complete changes from the NL code to the TL, the output variable ***tens*** in the NL code also needs to be linearized with respect to the variables *b* and ***tmp*** which depends on the variable ***a*** such that the corresponding term ***tl_tens*** in the TL code is composed of the variables ***tl_b*** and ***tl_tmp***, and constants ***b*** and ***tmp***. Note that the input coefficients ***a*** and ***b*** in the TL code should be previously read in outside of the TL code while the constant ***tmp*** must be calculated inside of the TL code by other NL variables from outside of the TL code. In certain cases, it is very important to put the tangent linear term (***tl_tmp***) before the basic state term (***tmp***), and the basic state term is not necessary if it is not associated with the nonlinear coefficient.

## 2.3 Linearization tests

The practical version of a TLM should be considered reasonably good enough if the TLM is to correctly describe time-evolving perturbations of the nonlinear model as the perturbation magnitude increases to the actual uncertainty size. The main goal in this study is to develop a TLM asymptotically that yields a similar solution as the difference between nonlinear solutions when the magnitude of perturbation approaches toward zero. So, the developed TLM can be used for various tools for the evolution of perturbations, stability analysis, and the forward model in the incremental 4DVar. We follow the method of Navon et al. (1992) below for a linearity check for the developed tangent linear model.

Assume that N(x) and M(x) respectively be the nonlinear module and its corresponding tangent linear module, respectively. Then, the correctness of the tangent linear module can be described as follows. The Taylor-Lagrange expansion of the nonlinear model is

$$N(x + a\,h) = N(x) + a\,h^{T}\,M(x) + O(a^{2}), \quad (1)$$

where x is a vector of all the input variables, $h$ is a state vector for perturbation, and the superscript T is matrix transpose. The constant $a$ is a small scalar such that the magnitude of initial perturbations is controlled by this scaling factor $a$. The Taylor-Lagrange formula in Eq. (1) can then be rewritten as

$$t(a) = \| N(x + a\,h) - N(x) \| / \| a\,h^{T}\,M(x) \| = 1 + O(a), \quad (2)$$

where $O(a)$ is the residual for the ratio of norms. When the tangent linear module is correctly developed, the above relationship $t(a)$ should hold within machine precision as the values of $a$ become small. The relationship indicates that the norm of tangent linear module in the denominator in Eq. (2) should approach to the norm of difference field between the two nonlinear models in the numerator in Eq. (2) as the magnitude of perturbations approaches zero.

We designed a practical linearity test setting, where individual variables are separately linearity-checked since the variables in the module have different magnitudes. We integrated the nonlinear model with both perturbed and unperturbed initial conditions, and the tangent linear model with the initial perturbation. Here, the constant $a$ in Eqs. (1) and (2) serves as the perturbation scaling factor of the initial perturbation and is sequentially reduced by the factor of 10 such that the magnitude of the perturbation becomes smaller by the factor.

## 2.4 Temporal increment

During the TLM time integration, the TLM requires the time-varying basic states that are provided by the nonlinear dynamical system. If the TLM requires to read these basic states every time step, then it may require huge overheads to retrieve those coefficients during input/output (I/O) due to the high dimensionality of $O(10^{7})$ or higher. This might lead the time integration of the TLM to the excess of normal NWP model integration. So, the temporal increment for the TLM is one of the critical factors for the TLM development along with linearity check in Sect. 2.3.

In the initial development of the TLM, the time step of the TLM is set identical to that of the nonlinear model, and the time-varying basic states are calculated by the nonlinear model at every time step during the TLM time evolution (Fig. 3-a). In this approach, the tangent linear model resolves the perturbation growth very well for the sufficiently high frequency of a solution trajectory, and there is no cost related to I/O due to the storage of the trajectory in memory. In this approach, the period of time integration can be extended in order of O(10) without any instability or technical issues. It is worth to note that when compared to the results of a further approximated version of TLM, it can be used as a reference solution. However, this first development still may not be practical in the operational NWP applications because of the high computational cost is extremely burdensome. So, alternate strategies for practical implementation of a TLM are required.

As seen in previous studies, many applications show the impact of less frequently updating trajectory on TLM integration, and suggest that the basic states do not have to be stored at every time step for an effective TLM (Errico et al., 1993; Yannick, 2004). One of alternate strategies is that the infrequently saved basic states are interpolated whenever the TLM requires the coefficients between the saved time steps. The strategy chosen here is first to increase the time step of the tangent linear model and second to store the nonlinear trajectory on files at the extended time. We obtained a best saving frequency of nonlinear solutions for the TLM in terms of efficiency and performance as long as the computational cost such as I/O and storage is manageable (Fig. 3-b).

## 3 Numerical results

### 3.1 Module linearity checks

Many studies employed perturbation magnitudes for wind, temperature, and surface pressure from 0.1 $ms^{-1}$, 1 $K$ and 1 $hPa$ to 1 $ms^{-1}$, 10 $K$ and 10 $hPa$ respectively for the strong and the weak perturbations (Courtier and Talagrand, 1987; Lacarra and Talagrand, 1988; Rabier and Courtier, 1992). The magnitude of perturbations changes from the strong perturbations to the weak perturbations by reducing the scaling factor $a$ by 10. For weak perturbations, the tangent linear modules are expected to well approximate the behaviour of perturbation for the nonlinear forward model and to keep the relative error small, but when the scale factor

becomes too small, the residual O($a$) for the ratio of norms in Eq. (2) is expected to be worse due to the numerical truncation errors.

For thorough linearity tests for each module, we configured different perturbations by choosing nonlinear model states at day 0, 1 and until day 8. These perturbations are initial conditions for the TLM, and reduced by the factor of 10 by multiplying the scaling factor $a$. The unperturbed nonlinear model has initial conditions at given days and the perturbed nonlinear model has initial conditions by summing the initial conditions of the unperturbed nonlinear model and the perturbations (initial conditions for the TLM).

There are two main modules to be linearized for the TLM; *compute_and_apply_rhs* calculates the dynamical tendency, and *advance_hypervis* is spatial filtering using $4^{th}$ order hyper viscosity. The module *compute_and_apply_rhs* consists of various subroutines and functions such as *divergence_sphere, gradient_sphere, vorticity_sphere, preq_hydrostatic, preq_omega_ps,* and *preq_vertadv*. The *advance_hypervis* module includes *biharmonic_wk, laplace_sphere_wk*, and *vlaplace_sphere_wk*. Prior to testing the two main modules, those subroutines and functions are directly linearized and checked individually by the linearity tests in Eq. (2).

Fig. 4 shows the results of the ratio of norms for the two major modules. The horizontal and vertical axes are respectively the values of the scaling factor $a$ and the residual O($a$) for the ratio of norms in Eq. (2). The slopes with different colors show the residual O($a$) calculated at different days. The numerical results show that for all cases, the slopes are decreased as the scaling factor $a$ is decreased, even if there are small differences of the magnitude between the slopes. As expected, when the scaling factor gets smaller, the perturbation reaches the machine precision and the slopes do not decrease anymore. With variously different perturbations and initial conditions, the similar pattern described as in Fig. 4 shows the residual O($a$) for all other modules, including the main time stepping loop module, *prim_run_subcycle* that is composed of the time stepping module *prim_advance_exp*, along with two major modules shown in Fig. 4. This implies that the linearization for all nonlinear modules is performed properly and completely. The TLM is verified to be accurate, and its solutions are therefore expected to be truly asymptotically correct.

## 3.2 Field checks

Further to verify the correctness of the TLM, we plotted the full field of V-wind components for the TLM and the corresponding difference fields between the two nonlinear model forecasts. In general, an increment produced by assimilating any DA systems is believed to represent a typical analysis error and is treated as a reasonable initial perturbation, or the increment can be constructed by a difference field between two full states in different forecast ranging (Ehrendorder and Errico, 1995). Because the magnitudes of the latter method is similar to those of the nonlinear model results at day 6 with reduced magnitude of 10% or 1%, initial perturbations are obtained by choosing nonlinear model results with 10% or 1% reduced magnitude. The initial perturbations are used as the initial condition for the TLM, and the two parallel nonlinear models are also integrated over time, one with the perturbations added to the initial condition and the other without the initial perturbation.

Fig. 5 shows the snapshots of V-wind fields to compare the difference of the two nonlinear models and the linear model evolutions at 0, 24, and 48 *hr*. The initial perturbations of 10% and 1% magnitudes of V-wind components for the TLM are respectively displayed in Figs. 5-a and 5-d (first column) with contours, and their TLM forecasts are shown with contours at day 1 (second column) and day 2 (third column). Similarly, the nonlinear evolution of the initial perturbations are evaluated by the difference fields between the two nonlinear model forecasts and displayed with shadings. In Fig. 5, both amplitudes and patterns from the TLM solutions and the differences of the two nonlinear forecasts are very similar. The amplitudes of the TLM results for both day 1 and day 2 also show linear trends between 10% and 1% magnitudes of initial perturbations, and the pattern correlation with 1% magnitude is much higher than that with 10% magnitude. These results confirm that the initial evolution is well represented by the developed TLM (version 1.0) up to at least 48 *hr* for the resolution of 220 *km* ($N_e$ =16). The similar numerical results were obtained for different model configurations with different model resolutions, initial conditions, and perturbations (Figures are not shown). These results confirm that the TLM (version 1.0) for the HOMME dynamical core is correctly developed and reasonably well represents the initial perturbation evolution.

## 3.3 Temporal increment

A time step size in tangent linear models plays an important role in numerical stability and computational cost, so it is important to choose a suitable time step size to balance between

the numerical stability and computational cost. Too short time step makes the TLM too expensive due to the I/O as seen in Sect. 2.4, and too long time step makes the model numerically instable. There are a couple of ways to determine a proper time step size for stable integration of a TLM. One is to try different time step sizes for the TLM and the other can check stability conditions for given numerical schemes.

Here, various time steps are applied to the TLM and empirically tested for numerical instabilities. Fig. 6 shows snapshots of V-wind fields at time 5 *hr* for the results of the TLM with different time step sizes from $\Delta t$=150 *s* to $\Delta t$=600 increased by 150. At the time step of $\Delta t$=300, the result shows the stable time integration of the TLM up to 48 hours, and the TLM with $\Delta t$=450 holds the numerical stability for 11 hours. The TLM with time step of $\Delta t$=600 shows the instability after 5 hour. For a given 6-hour assimilation window that is usually used for 4DVAR schemes in many NWP centres, the TLM results with time step sizes less than $\Delta t$=450 is very similar to that with default time step of $\Delta t$=150, and yields stable integration results. Thus, the expanded time step size of $\Delta t$=450 would be appropriate for a best temporal increment. This can be confirmed quantitatively by considering the relative mean error, defined, for any quantity $\boldsymbol{X}$ at the time T=5 hr, as

$$|| \boldsymbol{X}_{\mathrm{TLM}} - \boldsymbol{X}_{\mathrm{NLD}} || / || \boldsymbol{X}_{\mathrm{NLD}} ||, \quad (3)$$

where $\boldsymbol{X}_{\mathrm{TLM}}$ is a TLM field at T=5 hour, $\boldsymbol{X}_{\mathrm{NLD}}$ is the corresponding difference fields between the two nonlinear model forecasts at 5 hour, and || || is a spatial averaged norm. Table 1 gives these values for the mean of the stat variable $\boldsymbol{X}$ at time T=5 hr. And the total wallclock time is decreased, as the time step size is increased such that when $\Delta$t=150s is set to be 100%, 2$\Delta$t becomes 56%, 3$\Delta$t is 36%, and 4$\Delta$t for 33%. Although the TLM (version 1.0) developed in this study still needs further improvement for its performance, the current version is practical within a cope of a reasonable compromise between linearity, computational efficiency, and forecast performances.

## 4 Summary and discussion

In this study, modules to calculate tangent linear trajectories have been implemented into the HOMME dynamical core. The TLM describes the evolution of perturbations about time varying basic states that are provided by the nonlinear dynamical system. The TLM

accommodates a Jacobian of the dynamical operator that is tangential to a solution trajectory of the nonlinear system, and also provides a computationally efficient way to calculate the model trajectory. Since the TLM is primarily intended to approximate the evolution of perturbations in a corresponding nonlinear model, the accuracy of the TLM is considered to be a measure of the model performance. In that regard, the developed codes for the TLM are checked by the Taylor-Lagrange formula and by comparison of time-evolved perturbation fields for the TLM with the difference fields between two controlled nonlinear model runs. Overall verification of the numerical results indicates that the tangent linear model is correctly developed.

Generally, there are some major inaccuracy issues in developing TLMs (Errico et al., 1993) due to the finite magnitude of the perturbations in initial/boundary conditions, model parameters, the strong nonlinearities, discontinuities in nonlinear models, and numerical instabilities, which make difficult the development of efficient and well-behaving tangent linear codes. During the development of the tangent linear codes for the HOMME dynamical core, however, we have not experienced any significant difficulty such as a tendency to suddenly grow small perturbations due to some unintended discontinuities or ill-conditioning in the HOMME model. We believe that it is because the dynamics has good computational properties such as no singularity on both poles (Dennis et al., 2012).

Since the TLM requires nonlinear solutions as coefficients, the I/O strategy is important for the practical implication of the TLM. Two TLMs are developed with different I/O such as recalculating the basic state and storing the trajectories in file. The TLM with recalculating the basic state at every time step is extremely burdensome, but the results of the TLM well represent the evolution of perturbations, and those results can be used as reference fields in comparison with those of the approximated TLM. The extra burden leads to the alternate strategy for the TLM that is to store and read the trajectories from the file. As the time-step of the TLM is increased, the burden of I/O is decreased. Furthermore, given a time step size the instability during the TLM time integration should be carefully studied. It is an accurately developed TLM is crucially important that is because the same time step is directly used for the time step of adjoint model, and also influences on the performance of 4DVAR schemes.

Critical element in any operational prediction schemes such as 4DVar and 4 Dimensional Ensemble based Variational method (4DEnVar) will, of course, be the initialization procedure. The issue that has not been addressed by the present development is the analysis increments

in the initialization procedure that generally develop gravity waves. To filter out high-frequency waves, an incremental analysis-updating scheme (Polavarapu et al., 2004) is developed for the forecast model, and for 4DEnVar and 4DVar. The TLM (version 1.0) developed here can be another option for an internal digital filtering initialization scheme such that the high frequency in the analysis increments are filtered out by propagating the TLM forwards and backwards (with a negative time step), and then by forming a weighted average of the states in the combined trajectory. Korea Institute of Atmospheric Prediction Systems (KIAPS) is a government funded non-profit research and development institute and currently developing a 4 Dimensional Ensemble-based Variational method (4DEnVar). KIAPS will test the TLM (version 1.0) for the initialization procedure.

## 6 Code availability

All codes in the current version of TLM are available upon the request. Any potential user interested in those modules should contact B.-J. Jung, and any feedback on them is welcome. Note that one may need help to use the TLM model optimally, but we do not have the resources to support the model in an open way. Since ADM is currently being developed based on the current version of TLM, all codes of ADM are also presumably available upon the request.

## Acknowledgements

Authors would like to thank Adam Clayton at Met Office for his proof-reading and precious comments on this manuscript. Also, we would like to thank the anonymous reviewers for their careful reading of the manuscript and their thoughtful comments to make the manuscript more clarified.

## References

Bennett A.F.: Inverse modeling of the ocean and atmosphere. Cambridge University Press, Cambridge, 2002.

Courtier, P. and Talagrand O.: Variational assimilation of meteorological observations with the adjoint equation –Part I. Numerical results. Q. J. R. Meteorol. Soc., 113, 1329-1347, 1987.

Courtier P., thepaut J. -N., and Hollingworth A.: A strategy for operational implementation of 4D-Var, using an incremental approach. Q. J. R. Meteorol. Soc., 120, 1367-1387, 1994.

Dennis, J. M., Edwards J., Evans K. J., Guba O. N., Lauritzen P. H., Mirin A. A., St-Cyr A., Taylor M. A., and Worley P. H.: CAM-SE: A scalable spectral element dynamical core for the Community Atmosphere Model. Int. J. High. Perform. Comput. Appl., 26, 74-89, 2012.

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

Errico, R. M., Vukicevic, T., and Raeder, K.: Examination of the accuracy of a tangent linear model. Tellus, 45A, 462-497, 1993.

Errico, R., and K. Raeder, K.: An examination of the accuracy of the linearization of a mesoscale model with moist physics. Q. J. R. Meteorol. Soc., 120, 1367-1387, 1999.

Giering, R., and Kaminski, T.: Recipes for adjoint code construction. ACM Trans. Math. Software, 24, 437-474, 1998.

Jablonowski, C. and Williamson D. L.: A baroclinic wave test case for atmospheric model dynamical cores. Q. J. R. Meteorol. Soc., 132, 2943-2957, 2006.

Lacarra, J.-F., and O. Talagrand, O.: Short-range evolution of small perturbations in a barotropic model. Tellus, 40A, 81-95, 1988.

Nair R.D. and Tufo H. M.: Petascale atmospheric general circulation models, J. Phys., 78, 012078, doi:10.1088/1742-6596/78/1/012078, 2007.

Nair, R. D., H-W. Choi, and H. M. Tufo, 2009: Computational aspects of a scalable high-order discontinuous Galerkin atmospheric dynamical core. Computers & Fluids, Vol. 38, 309-319.

Navon, I. M., Zou, X., Derber, J., and J. Sela, J.: Variational data assimilation with an adiabatic version of the NMC spectral model. Mon. Wea. Rev., 120, 1433–1446, 1992.

Polavarapu, S. and Ren, S. andClayton A. M. and Sankey D. and Rochon Y.: On the relationship between incremental analysis updating and incremental digital filtering. Mon. Wea. Rev., 132(10), 2495-2502, 2004.

Rabier, F. and Courtier, P.: Four-Dimensional assimilation in the presence of baroclinic instability. Q. J. R. Meteorol. Soc., 118, 649-672, 1992.

Thomas, S. J., and Loft, R. D.: Semi-implicit spectral element model. J. Sci. Comput., 17, 339-350, 2002.

Yannick T.: Diagnostics of linear and incremental approximations in 4D-Var. Q. J. R. Meteorol. Soc., 130, 2233-2251, 2004.

Yannick T.: Incremental 4D-Var convergence study. Tellus, 59A, 706-718, 2007.

Zhu, J. and Kamachi, M.: The role of time step size in numerical stability of tangent linear models. Mon. Wea. Rev., 128, 1562-1572, 2000.

Zou, X., Vandenberghe, F., Pondeca, M., and Kuo, Y. -H.: Introduction to adjoint techniques and the MM5 adjoint modeling system. NCAR Technical Note, NCAR/TN-435-STR, 1997.

1 Table 1. Relative mean Errors.

| Variable | **1*Δt** | **2*Δt** | **3*Δt** | **4*Δt** 2 3 |
|---|---|---|---|---|
| **u** | 0.0124556 | 0.0128355 | 0.0135081 | 0.163502 |
| **v** | 0.0128028 | 0.0120578 | 0.0115803 | 0.13647 |
| **t** | 0.00696689 | 0.00650514 | 0.00596657 | 0.104771 |
| **ps** | 0.00697304 | 0.00639369 | 0.00547336 | 0.0750567 |

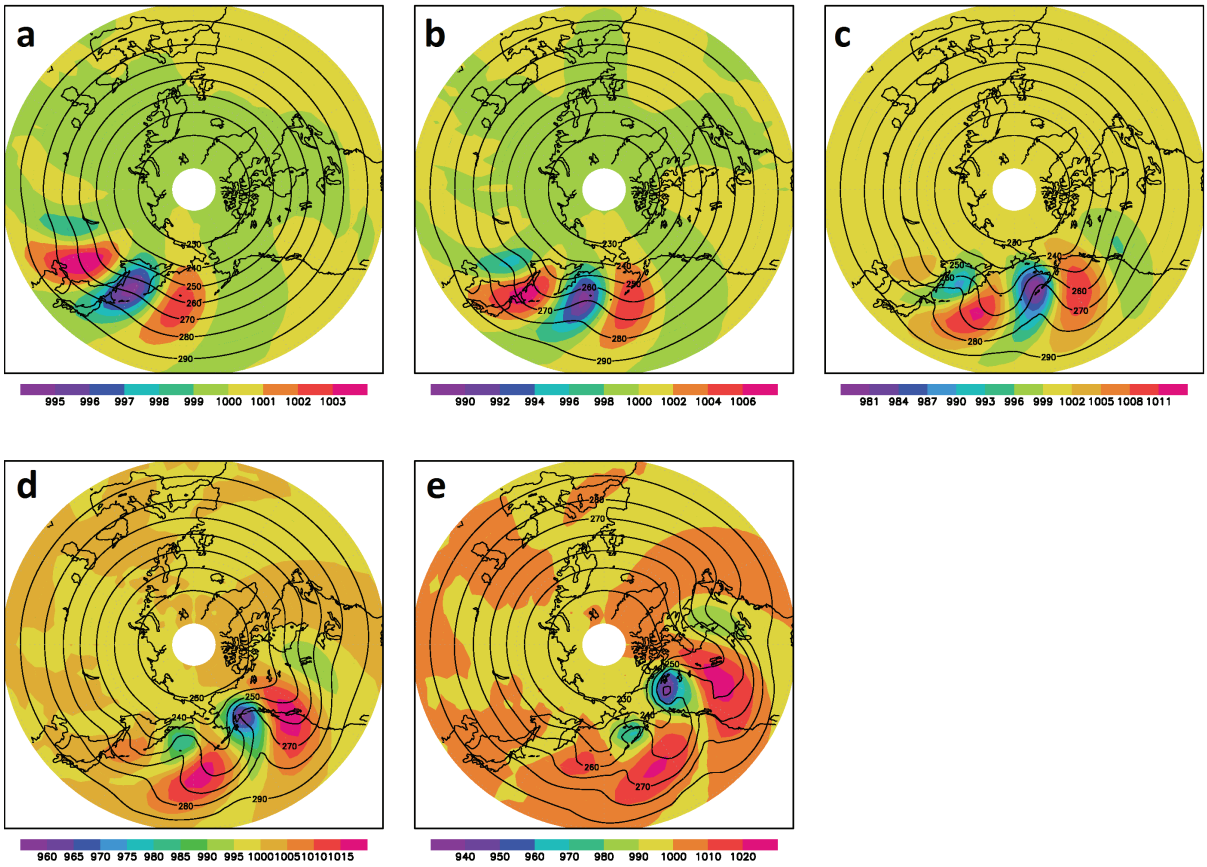


Figure 1. Evolution of the baroclinic wave from time integration with different days. The shadings and contours represents surface pressure (hPa) and temperature (K), respectively. (a) day 6, (b) 7, (c) 8, (d) 9, (e) 10.

```
Subroutine NL( a, b, tens )
real, intent(in) :: a, b
real, intent(out) :: tens
real :: tmp

tmp  = 3.0d0 * sin(a)
tens = tmp * b**2
End subroutine NL
```

```
Subroutine TL( a, b, tl_a, tl_b, tl_tens )
real, intent(in) :: a, b, tl_a, tl_b
real, intent(out) :: tl_tens
real :: tmp, tl_tmp

tl_tmp  = 3.0d0 * cos(a) * tl_a
tmp     = 3.0d0 * sin(a)
tl_tens = tl_tmp * b**2 + tmp * 2.0d0 * b * tl_b
End subroutine TL
```

Figure 2. Example of the tangent linear subroutine called TL based on the nonlinear subroutine called NL. The subroutines displays input and output with capital letters I and O in the argument variables.

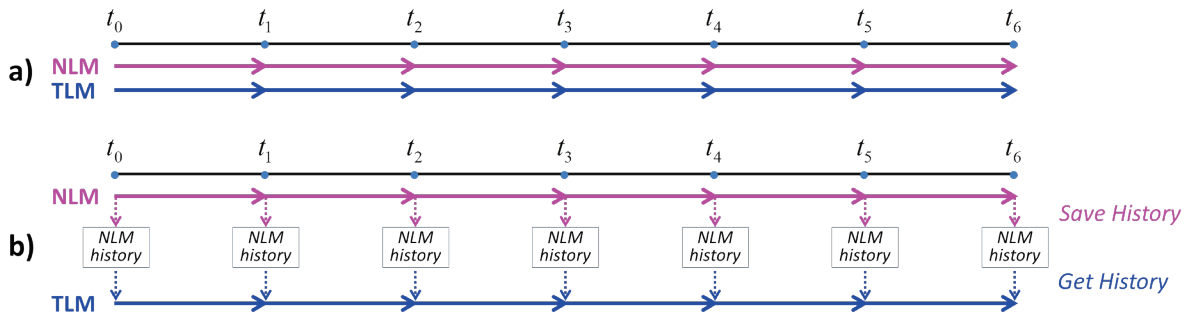


Figure 3. Nonlinear trajectory management for the tangent linear model. a) Before the tangent linear model (initial version of TLM) is integrated, the nonlinear model (NLM) is calculated every time step ahead. b) Nonlinear solutions are first saved during the time-integration of the NLM, and then the TLM is integrated over time with coefficients from the NLM run.

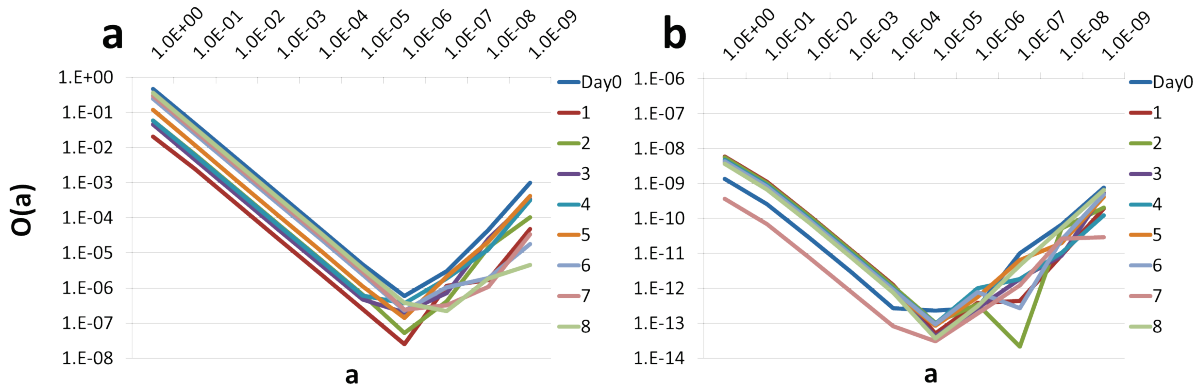


Figure 4. Linearity test for the two major modules: (a) *compute_and_apply_rhs*, and (b) *advance_hypervis*. The horizontal and vertical axes are respectively the values of the scaling factor $a$ and the residual O($a$) for the ratio of norms in Eq. (2). The slopes with different colors show the residual O($a$) calculated at different days.

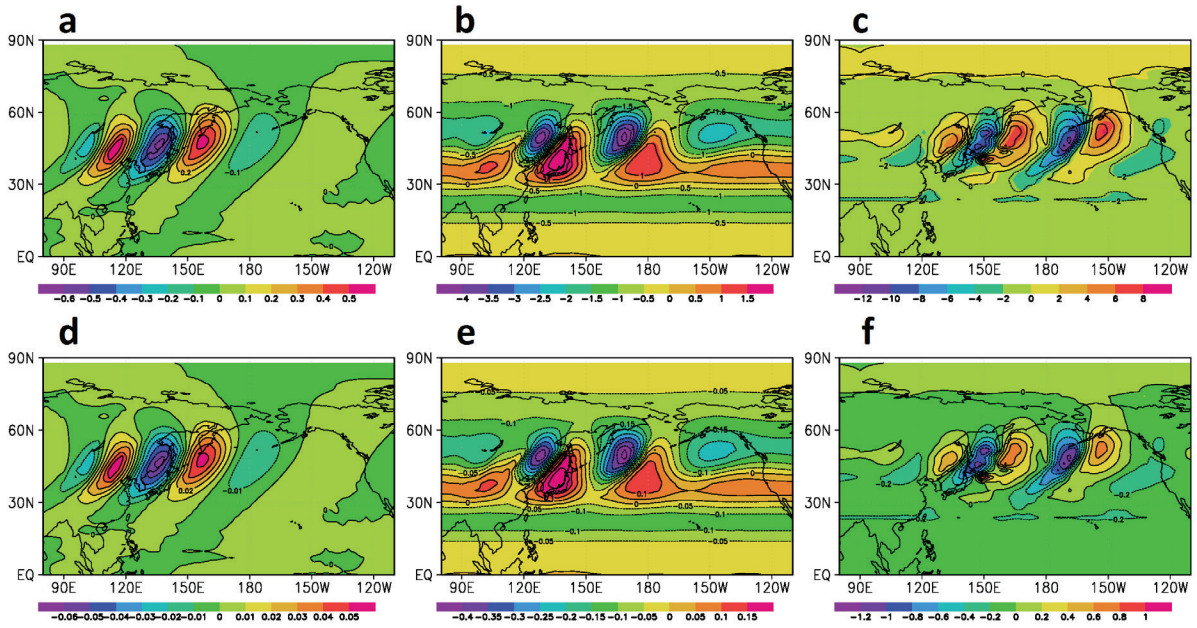


Figure 5. Evolution of different initial perturbations for the V-wind fields (m $s^{-1}$). Upper panel (a,b,c) shows wind with 10% perturbation of the initial state and lower panel (d,e,f) with 1% perturbation (see details in Sect. 3.2). The shadings represent the difference between the two nonlinear models runs with perturbed and unperturbed initial conditions. The contours illustrate the evolution of wind perturbation propagated by the tangent linear model at different times, the initial time (left column), 24h (middle), and 48h (right).

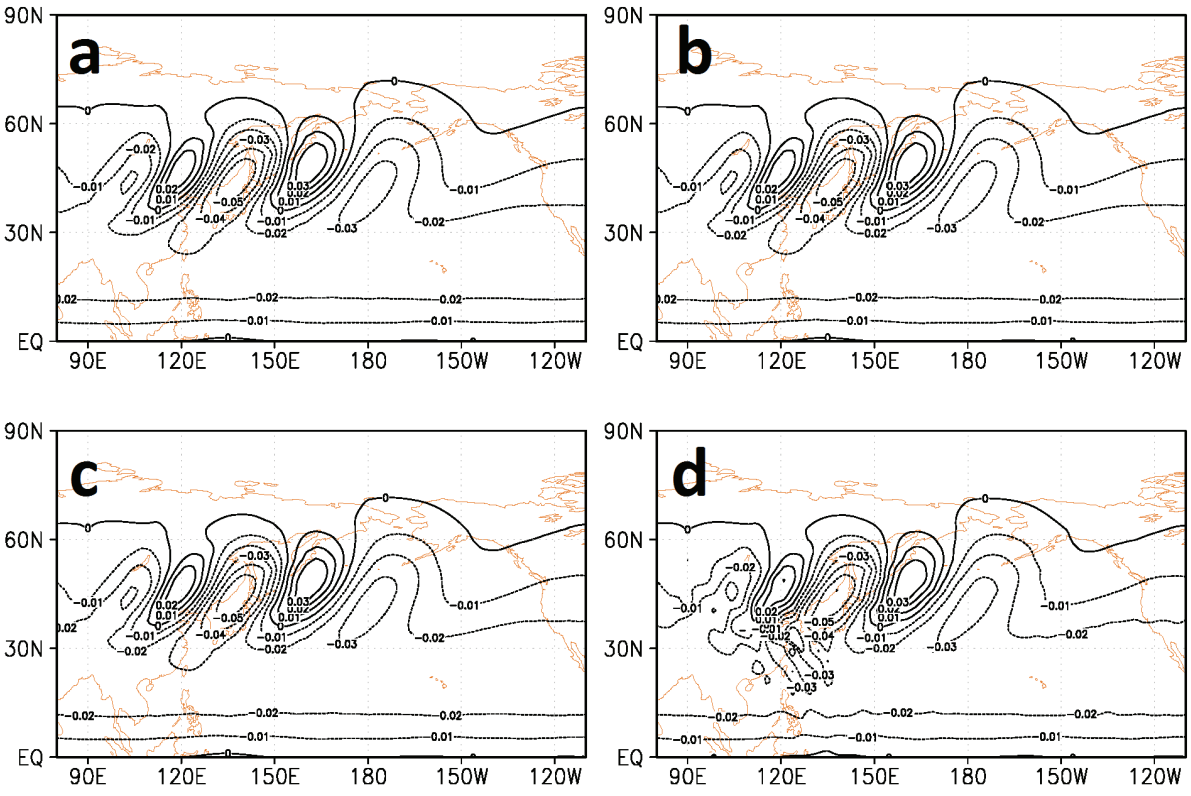


Figure 6. V-wind fields ($ms^{-1}$) of the tangent linear model with different time increments at 5 hour later. Time step size $\Delta t$ is a) 150, b) 300, c) 450, d) 600 second.