NCAR

Development of the Adjoint Model for WRF and its Prospective Applications

Qingnong Xiao, Dale Barker, John Michalakes,

Wei Huang, Hans Huang

Mesoscale and Microscale Meteorological Division

National Center for Atmospheric Research

Boulder, Colorado 80307-3000

Outlines

- Simplified Version of the WRF Model
- Automatic Derivation of the Adjoint Code for WRF Using TAF
- Adjoint Correctness Verification
- Prospective Applications
  - WRF 4D-Var
  - Parameter Estimation
  - Adjoint Sensitivity
  - Singular Vector
- Summary

Simplified Nonlinear Model for WRF

NCAR

FNM: Full Nonlinear Model of WRF

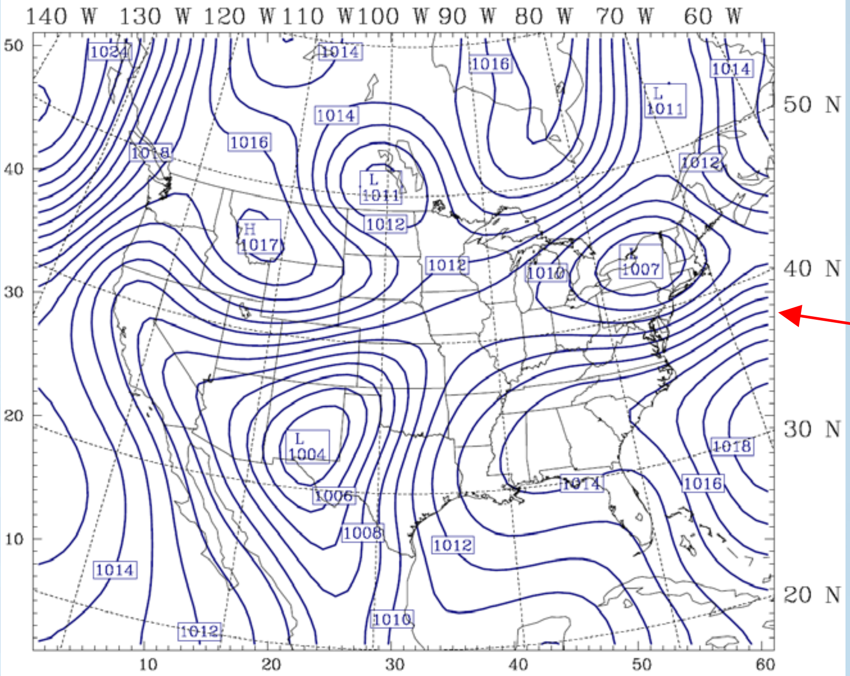
- WRF is suitable for use in a broad range of applications across scales ranging from meters to thousands of kilometers.
- WRF is a complex numerical model. There are over 900 total expanded subroutines. It has nearly 100 subroutines related to its dynamical core.

SNM: Simplified Nonlinear Model for WRF

- A basic, simplified version of WRF is necessary for the development of its tangent linear and adjoint models.
- Currently, we focus on the dry, adiabatic Euler mass coordinate WRF with simple diffusion, as the simplified nonlinear model for WRF.

Numerical Experiment with WRF SNM

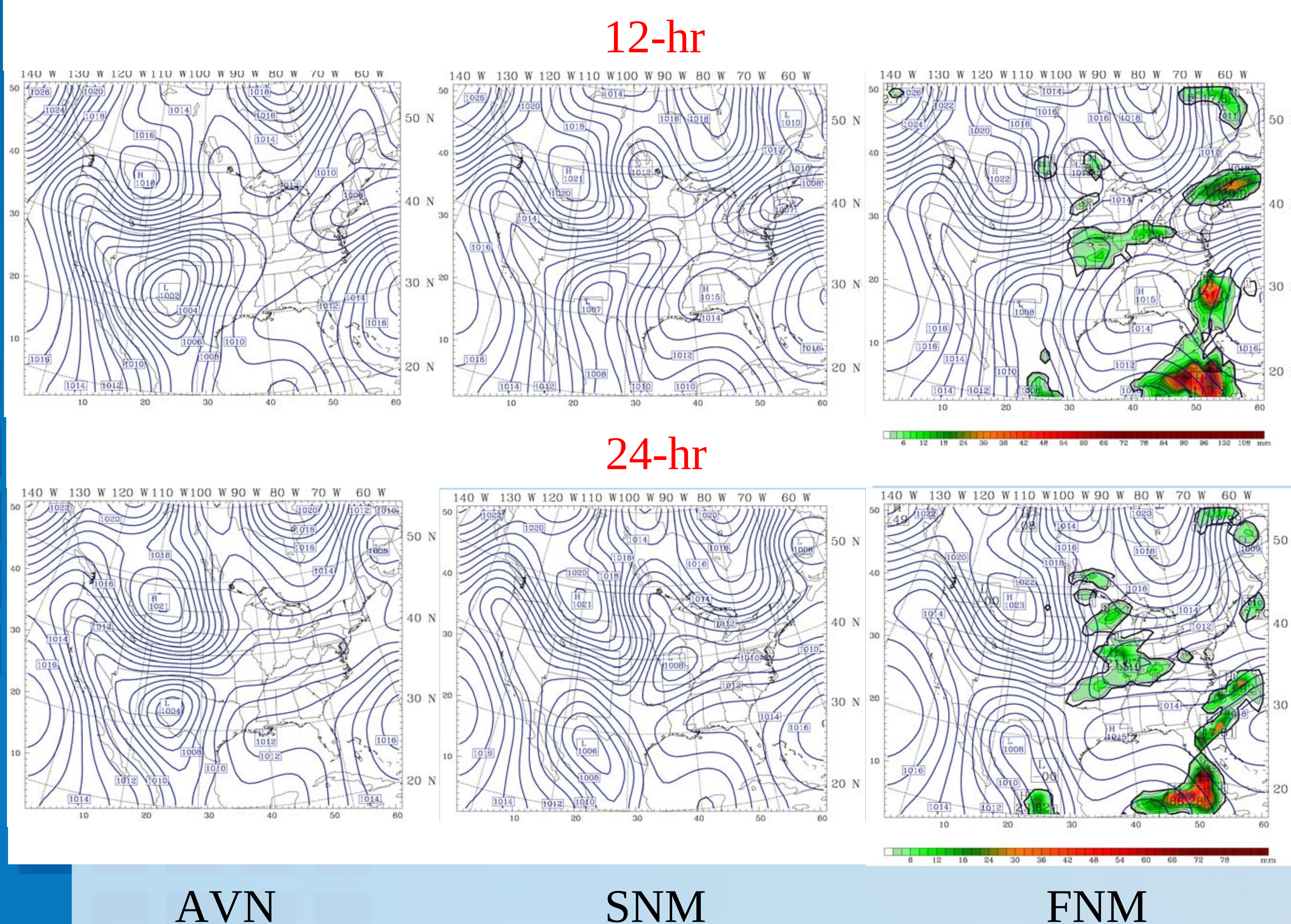
- WRF SI (Standard Initialization) at 12 UTC 12 June 2002
- NCEP AVN analyses are used for the initialization
- Model domain covers continental US (CONUS) with grids of 61X51X15 and grid-spacing of 90 km



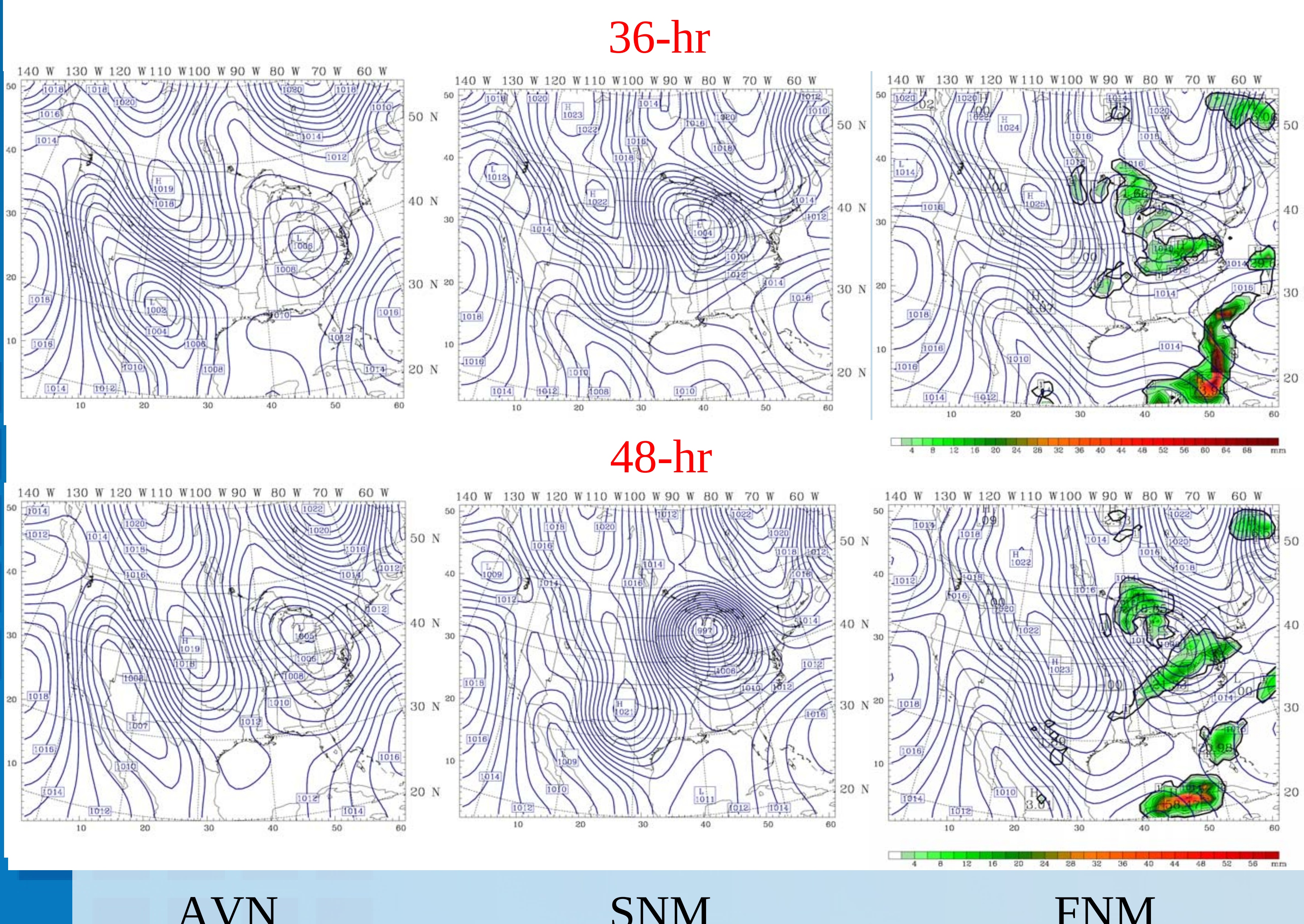
Initial condition (sea-level pressure) at 1200 UTC 12 June 2002

- 3rd order Runge-Kutta (rk_ord=3) integration scheme

SLP (interval: 1hPa) and 12-hr rainfall



SLP (interval: 1hPa) and 12-hr rainfall



WRF SNM

NCAR

- WRF simplified nonlinear model (SNM) using the dry, adiabatic, em dynamical core with simple diffusion (diff_opt=1 and km_opt=1) can produce a reasonable advection of the weather systems
- For the selected test domain with 90km resolution, the SNM, FNM 24-hr forecasts and AVN analysis at the same time are comparable
- SNM 48-hr forecast becomes apparently different from the FNM forecast and AVN analysis
- The dry, adiabatic WRF SNM is taken as an initial start for the WRF adjoint model development

Development of the Adjoint Model for WRF

NCAR

WRF Linearized Model

- The tangent linear of SNM will served as WRF linearized model. The trajectories are interpolated from the FNM

WRF Adjoint Model

- The adjoint model is developed according to the WRF linearized model with the same trajectories
- In WRF incremental 4D-Var, the outer-loop FNM is not necessary to be consistent with the inner-loop linear and adjoint models. But the linear and adjoint models must be consistent with each other to ensure the minimization convergence

Development of the Adjoint Model for WRF

NCAR

- TAF (Transformation of Algorithms in Fortran) is a source-to-source automatic differentiation tool for Fortran 90/95 codes. It is a commercial software developed by FastOpt (a German company)
- TAF has been successfully accomplished the tangent linear and adjoint model development for MM5V3 and several other numerical models
- WRF model is well-organized in its routines. All variables in each subroutine are explicitly declared. All the exchange variables between subroutines are put in the arguments list, and easy to be identified as input or output variables
- The tangent linear and adjoint models have been produced by automatic differentiation software TAF

Tangent Linear Verification

NCAR

Let *F(X)* denotes a nonlinear subroutine (function), and *g_F(X, g_X)* denotes its tangent linear subroutine (function). The correctness of the tangent linear *g_F(X, g_X)* can be tested against its nonlinear *F(X)* using the Taylor-Lagrange formula:

$$\lim_{g\_x \to 0} \frac{F(X + g\_X) - F(X)}{g\_X^T \cdot g\_F(X, g\_X)} = 1$$

where *X* is the input vector and *g_X* is the perturbation on the input vector.

Adjoint Verification

NCAR

Let $a\_F(X, a\_X)$ denotes the adjoint subroutine (function) of $g\_F(X, g\_X)$. The correctness of the adjoint $a\_F(X, a\_X)$ can be tested against its tangent linear $g\_F(X, g\_X)$ using the adjoint relation:

$$< g\_Y, g\_Y >=< a\_Y, g\_X >$$

where $g\_Y=g\_F(X, g\_X)$ and $a\_Y=a\_F(X, g\_Y)$.

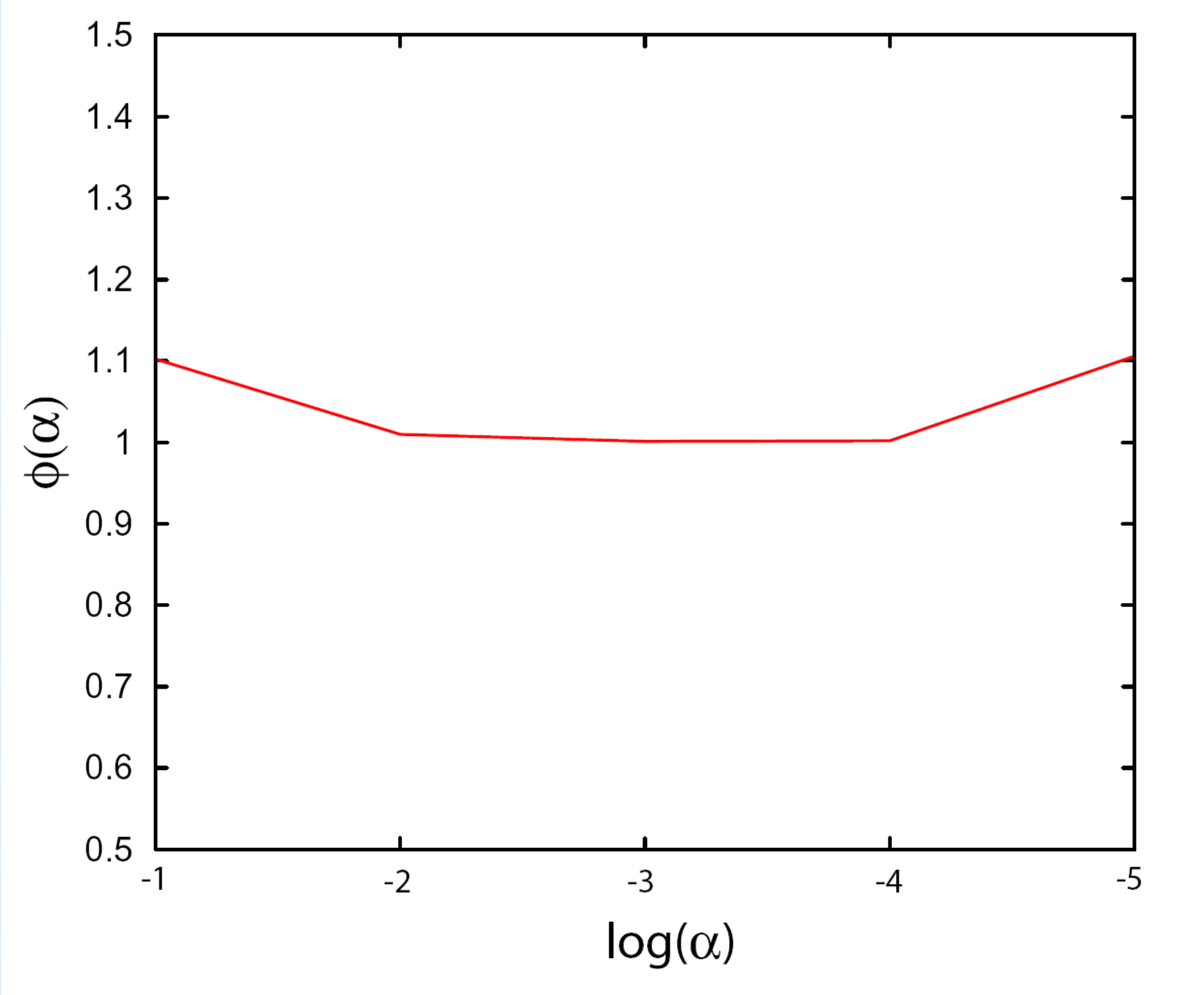
Correctness Verification

NCAR

- The correctness verification is undergoing. We have to ensure that the differentiation (linear and adjoint) for each subroutine is correct
- Here gives an example of our testing results about the advection subroutine for scalars (**Subroutine advect_scalar**)
- It has over 1000 statements in the subroutine. TAF could differentiate the subroutine and transpose the matrix correctly

Variations of the function $\phi(\alpha) = \frac{F(X+\alpha X) - F(X)}{\alpha X^T \cdot g\_F(X, \alpha X)}$ with respect to log(α). *F* denotes the subroutine **advect_scalar**

NCAR



Testing is conducted on DEC (with 32-bit precision)

Correctness Verification

- For the tangent linear testing with the value of α small but not too close to machine zero, Φ(α) is close to 1. This indicates the tangent linear code is right.

- Adjoint testing with the DEC 32-bits precision machine also passed the verification. The verification is based on the adjoint relation.

$$< g\_Y, g\_Y > = < a\_Y, g\_X)$$

0.337541E+06 0.337542E+6

It indicates the two numbers are equal with the machine accuracy.

Prospective Applications

NCAR

- WRF 4D-Var – improving the model IC
- Parameter Estimation – improving model parameter

→ Numerical prediction in operational

- Adjoint Sensitivity – defining the sensitive region for model IC and obs
- Singular Vector – defining fast-growing perturbations for a specific forecast (trajectory)

→ Targeting and predictability

WRF 4D-Var

NCAR

Cost function:

$$J = \frac{1}{2}(X(t_0) - X_b)^T B^{-1}(X(t_0) - X_b) +$$
$$\frac{1}{2}\sum_{i=0}^{N}\{Y_i - H_i[M_{t_0,t_i}\left(X(t_0)\right)]\}^T O^{-1}\{Y_i - H_i[M_{t_0,t_i}\left(X(t_0)\right)]\}$$

Gradient of the cost function:

$$\nabla_{X(t_0)} J = B^{-1}(X(t_0) - X_b) + \sum_{i=0}^{N} R^T_{t_0,t_i} H_i^{'T} O^{-1}(Y_i - H_i(X(t_0))$$

where M_{t_0,t_i} is the nonlinear model integration from t_o to t_i, R_{t_0,t_i} is the linear model and $R^T_{t_0,t_i}$ is its adjoint model, H_i is observation operator, $H_i^{'}$ is the linear operator and $H_i^{'T}$ is its adjoint.

Incremental 4D-Var

$$J = \frac{1}{2}\delta X(t_0)^T B^{-1}\delta X(t_0) +$$

$$\frac{1}{2}\sum_{i=0}^{N}\{Y_i - M_{t_0,t_i}(X_b) - H_i^{'} R_{t_0,t_i}\delta X(t_0)\}^T O^{-1}\{Y_i - M_{t_0,t_i}(X_b) - H_i^{'} R_{t_0,t_i}\delta X(t_0)\}$$

Innovations

Multi-incremental 4D-Var

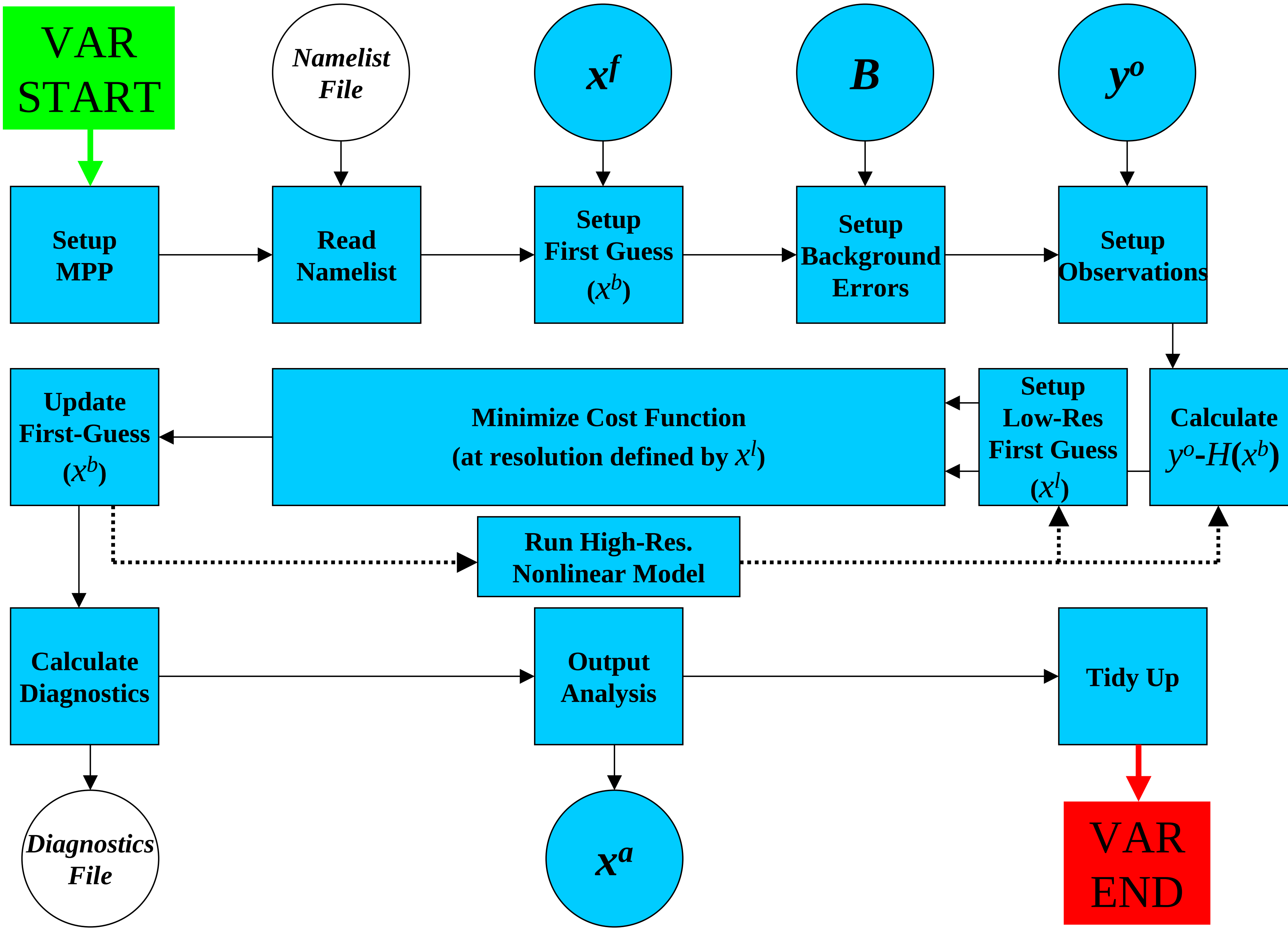
$$J = \frac{1}{2}\delta X^n(t_0)^T B^{-1}\delta X^n(t_0) +$$

$$\frac{1}{2}\sum_{i=0}^{N}\{Y_i - M_{t_0,t_i}(X_b^n) - H_i^{'} R_{t_0,t_i}\delta X^n(t_0)\}^T O^{-1}\{Y_i - M_{t_0,t_i}(X_b^n) - H_i^{'} R_{t_0,t_i}\delta X^n(t_0)\}$$

Usually $X_b^0 = X_b$, $\delta X^0(t_0) = \delta X(t_0)$, and $X_b^n = X_b^{n-1} + \delta X^{n-1}(t_0)$,

n=0, 1, 2 , … denotes outer-loops

WRF Incremental 4D-Var Mediation Layer



WRF incremental 4D-Var inner-loop

CV Transform: $x' = Uv$

SLM Integration: $x'(t) = Rx'(t_0)$

Observation Operator: $y' = Hx'$

Cost Function $J = J_b(v) + J_o(y')$

Adjoint Observation Operator: $x'_{adj} = H^T y_{adj}$

SAM Integration: $x'_{adj}(t_0) = R^T x'_{adj}(t)$

Adjoint CV Transform $v_{adj} = U^T x_{adj}$

Cost Function Gradient $= J_b(v) + J_o(v_{adj})$

Parameter Estimation

Given constrained parameters whose values vary between certain bounds, for instance, when the parameter α_i satisfies $\alpha_i \in [a_i, b_i]$, where a and b denote the lower and upper bounds, respectively, the cost function for parameter estimation may assume the following form:

$$J(\mathbf{X}, \boldsymbol{\alpha}) = \frac{1}{2}\int_{t_0}^{t_R} \langle \mathbf{W}(\mathbf{X} - \mathbf{X}^{obs}), (\mathbf{X} - \mathbf{X}^{obs})\rangle \, dt + \boldsymbol{\lambda}^{\mathrm{T}} \mathbf{g}(\boldsymbol{\alpha}), \tag{2}$$

where $\boldsymbol{\lambda}$ is the penalty coefficient vector, $\mathbf{W}$ is the diagonal weighting matrix defined as in Navon et al. (1992), $\langle \, \rangle$ denotes spatial integration, $\mathbf{X}$ represents the state variable vector, $\mathbf{X}^{obs}$ the observation vector, and t_0 and t_R denote the assimilation window. The second term consists of a penalty function. The value of $\boldsymbol{\lambda}$ is determined such that the penalty term is of the same order of magnitude as the other terms in the cost function, and $g(\alpha)$ is defined as

$$g_i(\boldsymbol{\alpha}) = \begin{cases} \frac{1}{2}(\alpha_i - b_i)^2 & \text{if } \alpha_i \geq b_i \\ 0 & \text{if } a_i < \alpha_i < b_i \\ \frac{1}{2}(\alpha_i - a_i)^2 & \text{if } \alpha_i \leq a_i, \end{cases} \tag{3}$$

Suppose that the forward model is perfect and is given in the form

$$\frac{\partial \mathbf{X}}{\partial t} = F(\mathbf{X}, \boldsymbol{\alpha}, t). \tag{6}$$

Its corresponding tangent linear model is defined as

$$\frac{\partial \delta \mathbf{X}}{\partial t} = \left(\frac{\partial F(\mathbf{X}, \boldsymbol{\alpha}, t)}{\partial \mathbf{X}}\right)\delta \mathbf{X} + \left(\frac{\partial F(\mathbf{X}, \boldsymbol{\alpha}, t)}{\partial \boldsymbol{\alpha}}\right)\delta \boldsymbol{\alpha}. \tag{7}$$

Then the adjoint model can be expressed in the form

$$-\frac{\partial \mathbf{P}}{\partial t} - \left(\frac{\partial F(\mathbf{X}, \boldsymbol{\alpha}, t)}{\partial \mathbf{X}}\right)^{\mathrm{T}} \mathbf{P} = \mathbf{W}(\mathbf{X} - \mathbf{X}^{obs}), \tag{8}$$

where $\mathbf{P}$ represents the adjoint variables. The gradients of the cost function with respect to the initial condition and the parameter α are, respectively,

$$\nabla_{\mathbf{X}_0} J = \mathbf{P}(0) \tag{9}$$

$$\nabla_{\alpha} J = \int_{t_0}^{t_R} \left[\left(\frac{\partial F}{\partial \boldsymbol{\alpha}}\right)^{\mathrm{T}} \mathbf{P}\right] dt + \boldsymbol{\lambda}^{\mathrm{T}} \frac{\partial \mathbf{g}}{\partial \boldsymbol{\alpha}}. \tag{10}$$

The adjoint model is of the same form as that where only the initial condition is considered as the control variable. Hence, the problem of parameter estimation via the adjoint method does not result in an additional computational effort when the number of parameters to be estimated is small. We may expect that the parameter estimation process will provide us with both optimally determined parameters and initial condition simultaneously. The gradient of the cost function with respect to both the initial condition and parameters is written as

$$\nabla J = (\nabla_{\mathbf{X}_0} J, \nabla_{\alpha} J)^{\mathrm{T}}. \tag{11}$$

Adjoint Sensitivity

NCAR

Define a response function

$$R(x_0) = G(H(x(t_r))),$$

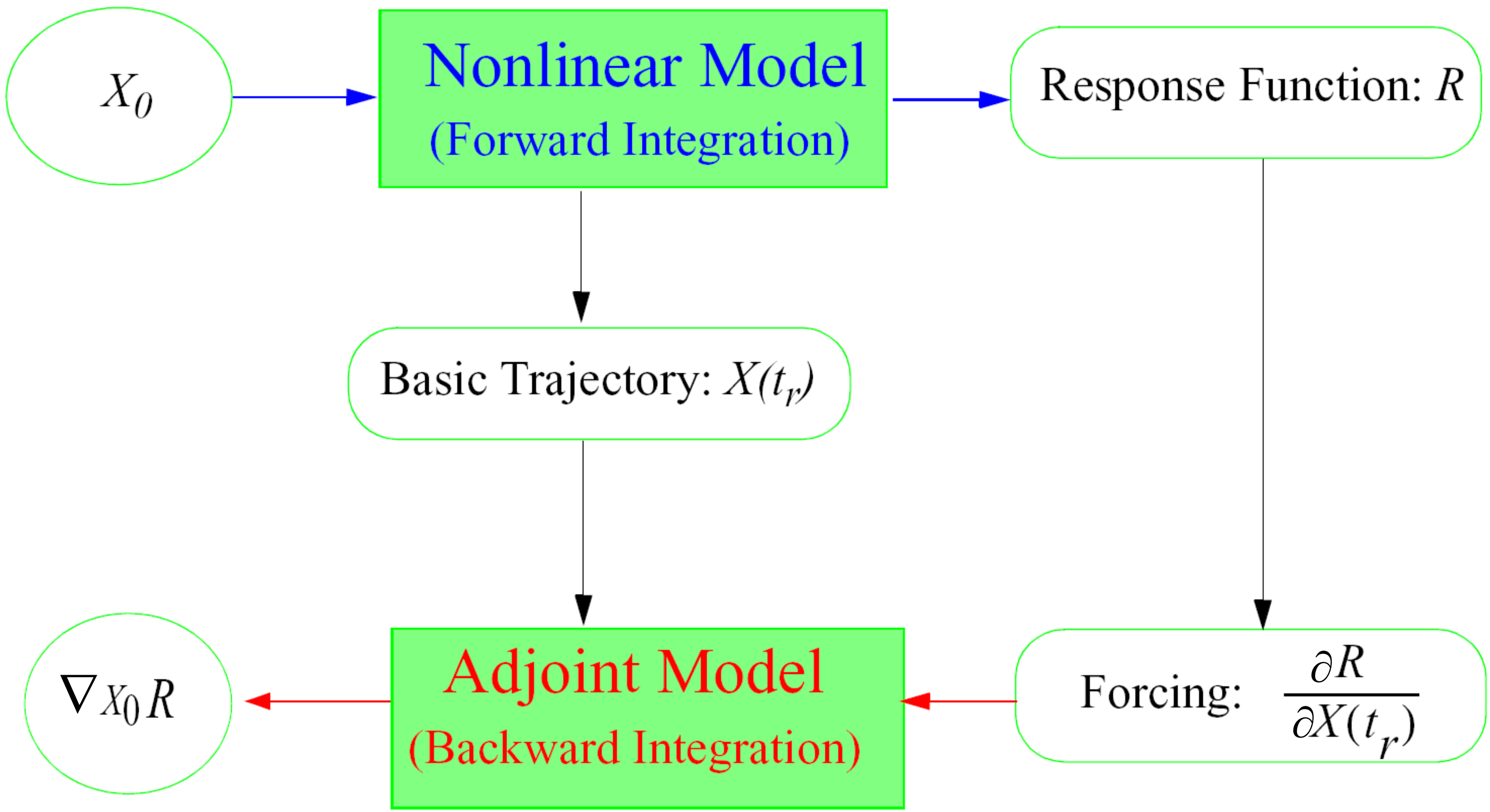
where $H(x(t_r))$ is a function of the model forecast at time t_r, and G is a scalar function of $H(x(t_r))$.

Calculate the gradient of the response function

$$\nabla x_0 R = P_r^T \left(\frac{\partial}{\partial x(t_r)} H(x(t_r))\right)^T \frac{\partial}{\partial H(x(t_r))} G(H(x(t_r))) \ ,$$

where P_r^T is the adjoint model integrated from time $t = t_r$ to $t = 0$, $\frac{\partial}{\partial x(t_r)} H(x(t_r))$ and $\left(\frac{\partial}{\partial x(t_r)} H(x(t_r))\right)^T$ are tangent linear and adjoint of the operator $H(x(t_r))$.

Flow Chart of Adjoint Sensitivity Calculation



In sensitivity analysis, the model output of interest is usually referred to as the system's *response,* instead of being called a cost function as in data assimilation. Sensitivity is a measure of the effect of changes in input variables on a selected model response. We consider a functional response $R(\mathbf{x})$ of the form

$$R(\mathbf{x}) = -\sum_{i,j} p_s(i, j; t_R) \quad (2)$$

where p_s represents the SLP at the time t_R (1200 UTC 22 Feb 1998), $\Sigma_{i,j}$ includes all model grid points in a 6

The sensitivity of the response R to variations in the system's input parameter $\mathbf{x}$ is defined by (Zou et al. 1993)

$$\nabla R(\mathbf{x}, \Delta\mathbf{x}) = \hat{\mathbf{x}}^{\mathrm{T}}\Delta\mathbf{x}, \quad (3)$$

where the value of $\hat{\mathbf{x}}$ is obtained by integrating the adjoint model backward with unit inputs of $\partial R/\partial \mathbf{x}(t_R)$ as its forcing term inside the response box, and zero values of $\partial R/\partial \mathbf{x}(t_R)$ outside the response box at the time t_R.

A NORPEX cyclone sensitivity to IC

Relativity sensitivity

Sensitivity of satellite brightness to the temperature and moisture

$$\mathrm{Var} = E_{\mathrm{fwd}}^2 + \frac{E_{\mathrm{ins}}^2}{n}, \quad (4)$$

where n is the number of samples averaged for the given IFOV. The error variance of brightness temperature is computed by multiplying the radiance error variance by $(dT_b/dR)^2$, which is the derivative of the brightness temperature T_b with respect to radiance based on the Planck function.

3. Sensitivity analysis of brightness temperatures with respect to temperature and specific humidity profiles

GOES-8 sounders measure brightness temperatures from 18 infrared spectral bands that are sensitive to the atmospheric temperature and moisture at various heights. To assess the sensitivity of the brightness temperatures with respect to perturbations in different atmospheric states, adjoint sensitivity analysis was performed. A simple response function was defined as

$$J_\alpha = J_\alpha(T, q, T_{\mathrm{sk}}, p_s) = T_b(\alpha), \quad (5)$$

where $T_b(\alpha)$ is the brightness temperature at the α channel, T is the temperature, q is the specific humidity, and T_{sk} is the surface skin temperature. In our study, T_{sk} is fixed and all the others (T, q, and p_s) are taken as input variables to the radiative transfer model and will be represented by the vector $\mathbf{x}$.

The sensitivity of J_α with respect to $\mathbf{x}$, expressed as ∇J, is usually defined as

$$\nabla J_\alpha(\mathbf{x}, \Delta\mathbf{x}) = (\nabla J_\alpha)^{\mathrm{T}}\Delta\mathbf{x} \equiv (\hat{\mathbf{x}})^{\mathrm{T}}\Delta\mathbf{x}, \quad (6)$$

where $\hat{\mathbf{x}}$ is the result of the adjoint model integration with a unit input for the adjoint variable of the brightness temperature at channel α, and zero value for the adjoint brightness temperature variables at other channels.

If a variation occurs solely in the ith component of the control variable vector $\mathbf{x}$, we denote by $\Delta\mathbf{x}^i$ the corresponding vector of variation:

$$\Delta\mathbf{x}^i = (0, \ldots, \Delta x^i, \ldots, 0)^{\mathrm{T}}, \quad (7)$$

and denote the corresponding sensitivity by ∇J_α^i. The relative sensitivity S_α^i is defined as the nondimensional quantity (Zou et al. 1993)

$$S_\alpha^i = \frac{\nabla J_\alpha^i}{J_\alpha}\left(\frac{\Delta x^i}{x^i}\right)^{-1}. \quad (8)$$

The magnitude of the relative sensitivity serves as a guide to ranking the importance of different components in the input variables $\mathbf{x}$. A plot of the vertical profile of the relative sensitivity, for example, will indicate where the most sensitive ranges of height are for brightness temperature information available at a certain channel.

Figure 1 shows the distribution of the sea level pressure (SLP) at 1200 UTC 15 August 1995 from the National Centers for Environmental Prediction (NCEP)

Singular Vectors

NCAR

Define a norm (under the definition of E sense)

$$\left\| X'(t) \right\|_E^2 = < X'(t), X'(t) >_E$$

We seek the initial perturbation X_0' that maximize the norm $\left\| X'(t) \right\|_E$.

Introducing the tangent linear model *R*, $X'(t) = RX_0'$, then

$$\left\| X'(t) \right\|_E^2 = < R^{T_E} RX_0', X_0 >_E$$

R^{T_E} is the adjoint of *R* with respect to the *E*-norm. $R^{T_E} = E^{-1} R^T E$

The problem becomes an eigenvalue problem. We seek the eigenvectors and eigenvalues of the matrix $K = E^{-1} R^T ER$.

The eigenvectors are singular vectors, and the corresponding eigenvalue is singular value.

SV's applications

- Study of the predictability of atmospheric model
- Instability study of atmospheric and oceanic flows
- Adaptive observations

Summary

NCAR

- The tangent linear and adjoint models are developed using the automatic differentiation software TAF. The correctness verification is ongoing. We expect this development will be done in the summer of 2005.

- There are a number of applications of the WRF adjoint system in the future. The basic research version of the WRF adjoint system will be ready in 2005.

  - 4D-Var
  - Parameter estimation.

  → Numerical prediction in operational

  - Adjoint sensitivity study
  - Singular vectors

  → Targeting and predictability