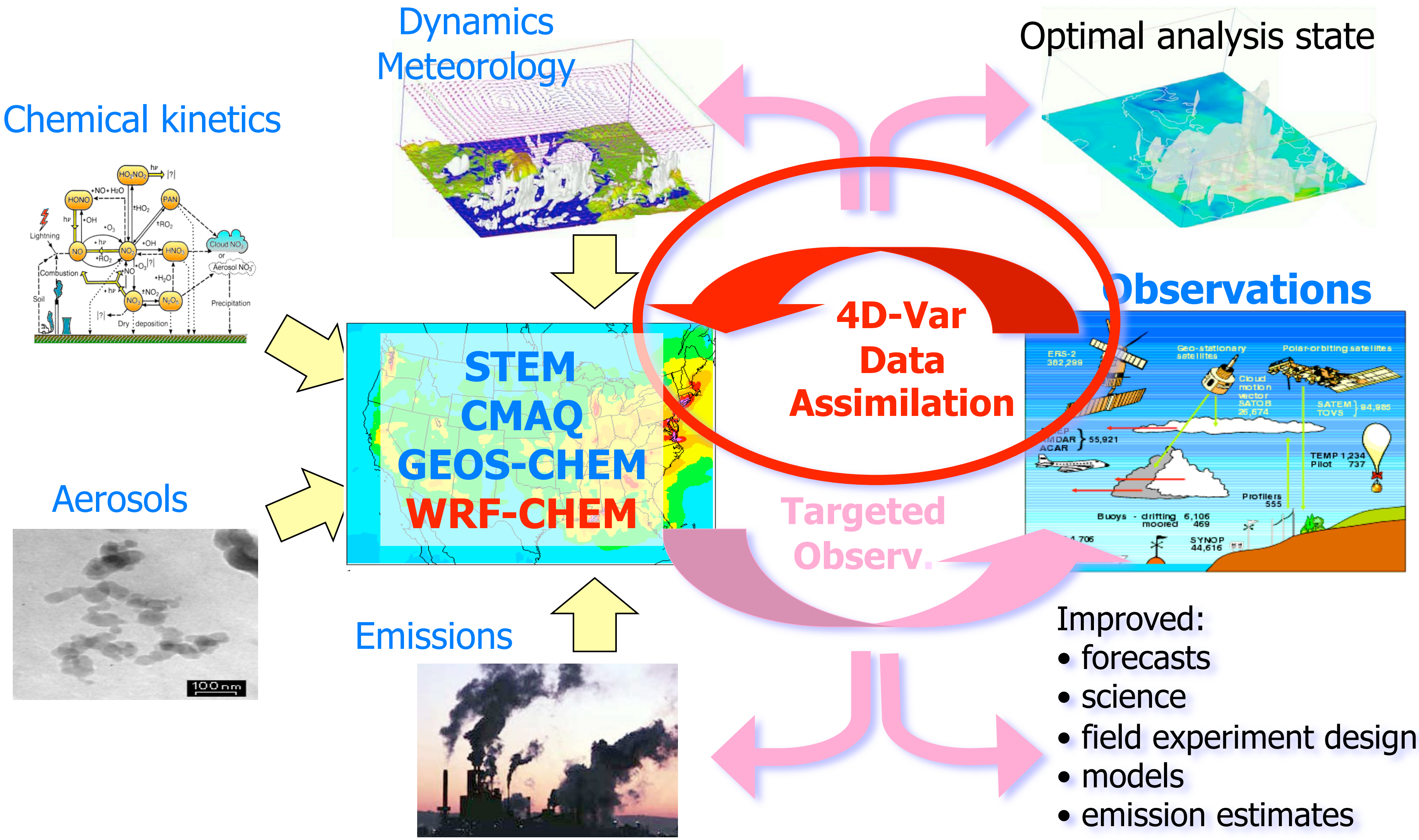
Chemical Data Assimilation

Adrian Sandu
Virginia Tech

Information feedback loops between CTMs and observations: data assimilation and targeted meas.



Virginia Tech

AR model of background errors accounts for flow-dependent correlations and is inexpensive

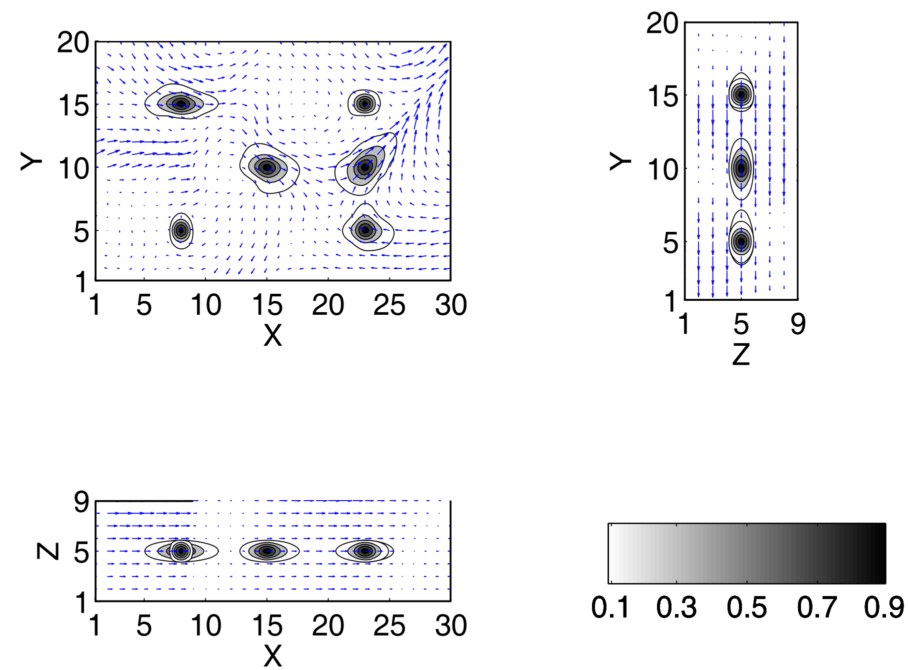
$$\psi\left(\mathbf{y}^0\right) = \frac{1}{2}\left(\mathbf{y}^0 - \mathbf{y}^b\right)^T \mathbf{B}^{-1}\left(\mathbf{y}^0 - \mathbf{y}^b\right) + \cdots$$

- Background error repres. considerably impacts the assimilation results
- Typically estimated empirically from multiple model runs (NMC)
- "Correct" mathematical models of background errors are of great interest

$$\delta\mathbf{y}' = \mathbf{M}'\delta\mathbf{y}$$

$$\left(\mathbf{I} - \Delta t\,\mathbf{M}'\right)^N \delta\mathbf{y} = \xi$$

$$\mathbf{B}^{-1} = \left(\mathbf{I} - \Delta t\,\mathbf{M}'^*\right)^N \left(\mathbf{I} - \Delta t\,\mathbf{M}'\right)^N$$

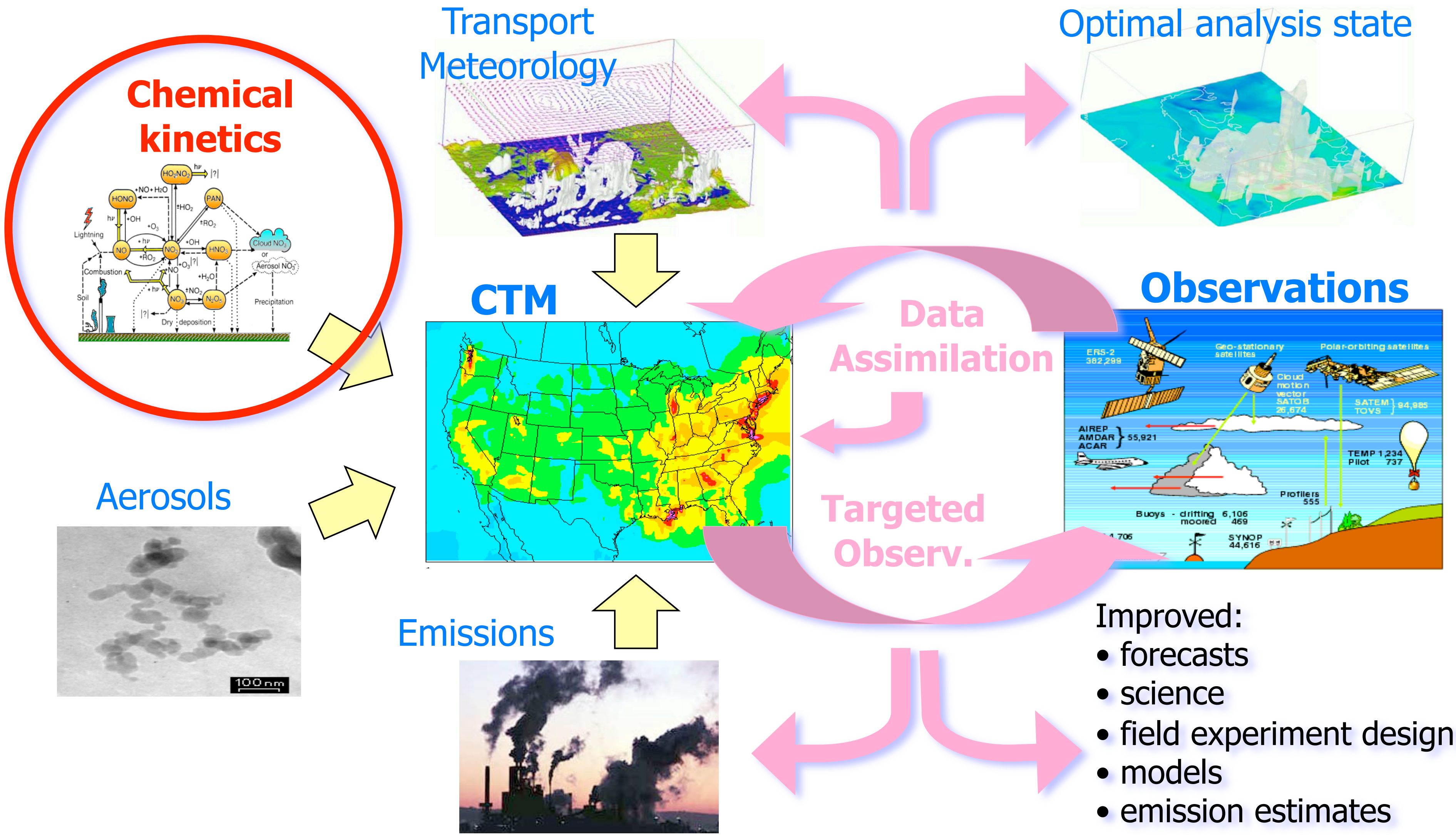


- "Monotonic TLM discretization"
- AR model of background errors
- $N\Delta t \approx$ lifetime of the species
- B is flow dep., cheap, full rank

[Constantinescu et.al., 2007]

Virginia Tech

Adjoints of stiff chemical kinetics: formulation, challenges, and automatic implementation



KPP automatically generates simulation and direct/adjoint sensitivity code for chemistry

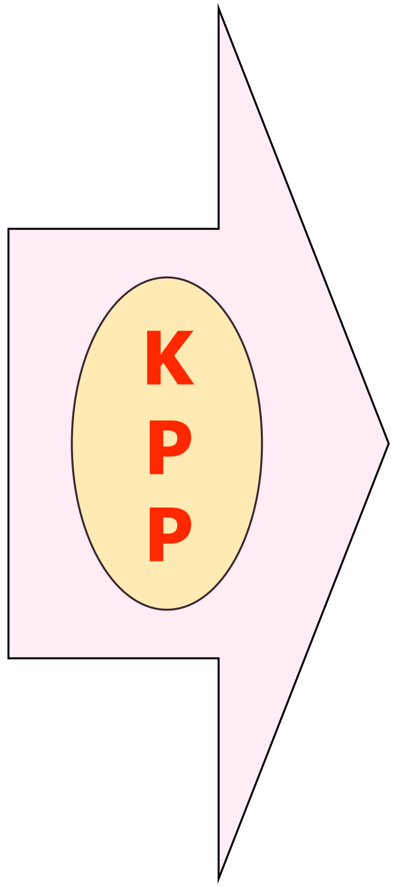
Chemical mechanism

```
#INCLUDE atoms

#DEFVAR
O  =  O;  O1D  =  O;
O3  =  O + O + O;
NO  =  N + O;
NO2  =  N + O + O;

#DEFFIX
O2  =  O + O;  M  =  ignore;

#EQUATIONS  { Small Stratospheric }
O2  + hv = 2O          :  2.6E-10*S;
O   + O2 = O3          :  8.0E-17;
O3  + hv = O  + O2     :  6.1E-04*S;
O   + O3 = 2O2         :  1.5E-15;
O3  + hv = O1D + O2    :  1.0E-03*S;
O1D  + M  = O  + M     :  7.1E-11;
O1D  + O3 = 2O2        :  1.2E-10;
NO   + O3 = NO2 + O2 :  6.0E-15;
NO2  + O  = NO  + O2   :  1.0E-11;
NO2  + hv = NO  + O    :  1.2E-02*S;
```



Simulation code

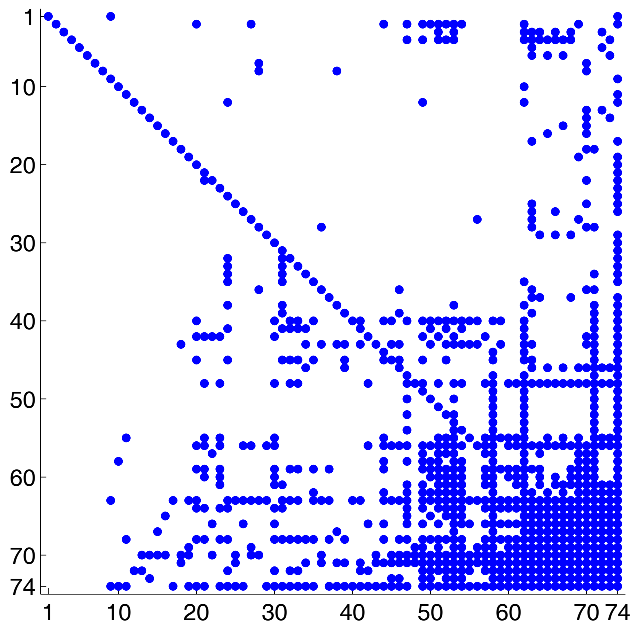
```
SUBROUTINE FunVar ( V,  F, RCT, DV )
   INCLUDE 'small.h'
    REAL*8 V(NVAR), F(NFIX)
    REAL*8 RCT(NREACT), DV(NVAR)
C A - rate for each equation
    REAL*8 A(NREACT)
C Computation of equation rates
    A(1) = RCT(1)*F(2)
    A(2) = RCT(2)*V(2)*F(2)
    A(3) = RCT(3)*V(3)
    A(4) = RCT(4)*V(2)*V(3)
    A(5) = RCT(5)*V(3)
    A(6) = RCT(6)*V(1)*F(1)
    A(7) = RCT(7)*V(1)*V(3)
    A(8) = RCT(8)*V(3)*V(4)
    A(9) = RCT(9)*V(2)*V(5)
    A(10) = RCT(10)*V(5)
C Aggregate function
    DV(1) = A(5)-A(6)-A(7)
    DV(2) = 2*A(1)-A(2)+A(3)-A(4)+A(6)-&A(9)+A(10)
    DV(3) = A(2)-A(3)-A(4)-A(5)-A(7)-A(8)
    DV(4) = -A(8)+A(9)+A(10)
    DV(5) = A(8)-A(9)-A(10)
    END
```

[Damian et.al., 1996; Sandu et.al., 2002]

Virginia Tech

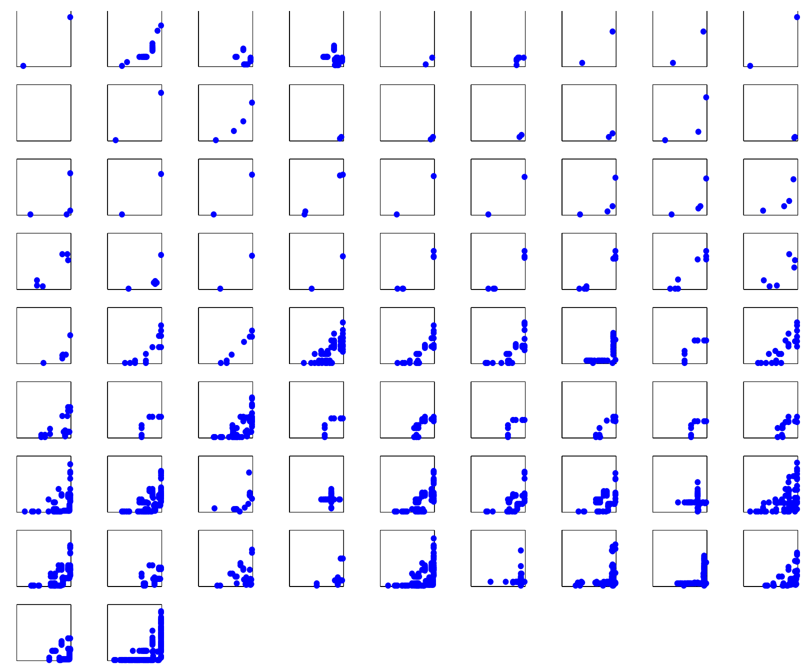
Sparse Jacobians , Hessians, and sparse linear algebra routines are automatically generated by KPP

#JACOBIAN [ON | OFF | SPARSE]
JacVar(…), JacVar[TR]_SP_Vec(…)
KppDecomp(…), KppSolve[TR](…)



SAPRC-99
79 spc./211 react. NZ=839, NZLU=920

#HESSIAN [ON | OFF]
HessVar(…), HessVar[TR]_Vec(…)

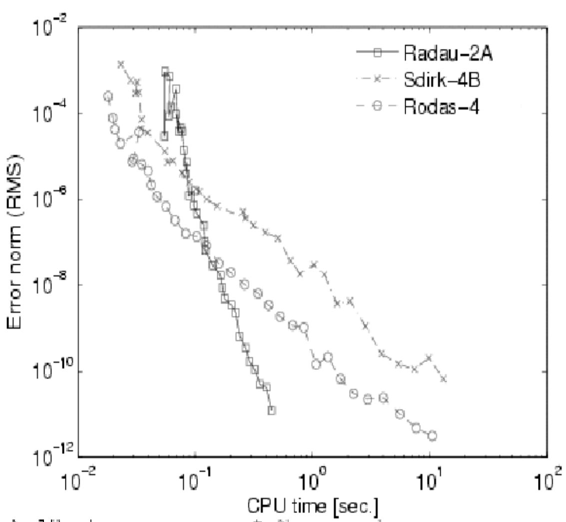


SAPRC-99.
NZ = 848x2 (0.2%)

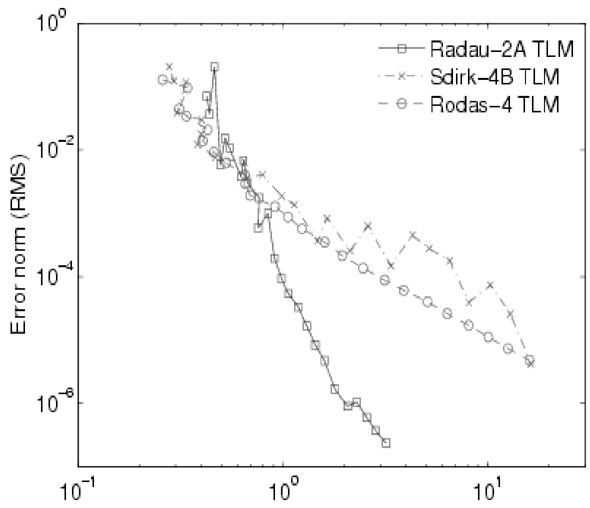
Methods available in the KPP numerical library

- FIRK 3-stage: Radau-2A (ord.5), Radau-1A (ord.5), Lobatto-3C (ord.4), Gauss (ord.6)
- SDIRK: 2a, 2b (2s, ord.2), 3a (3s, ord.2), 4a, 4b (5s, ord.4)
- Rosenbrock: Ros2, Ros3, ros4, Rodas3, Rodas4.

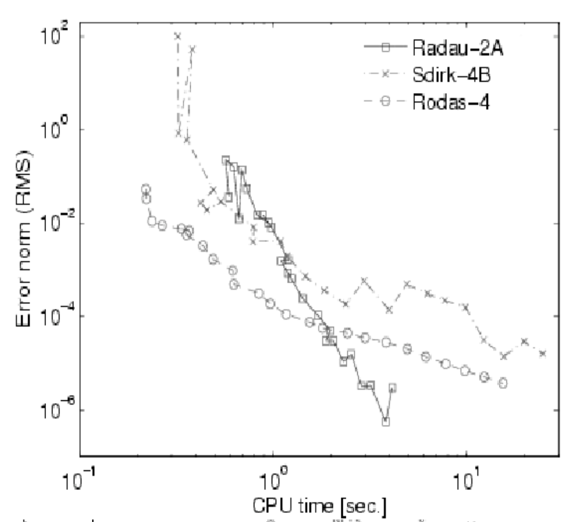
Forward



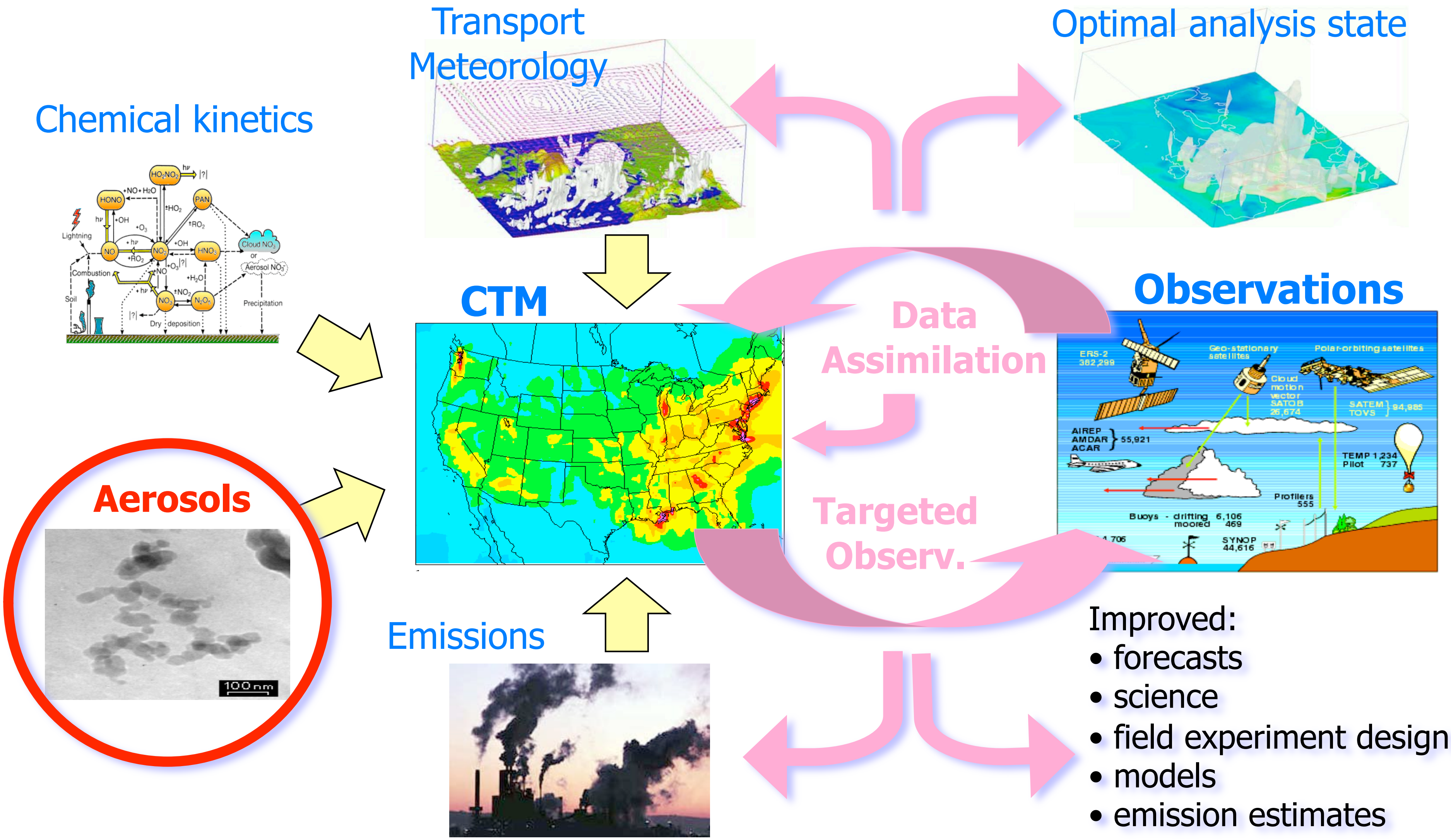
TLM (DDM)



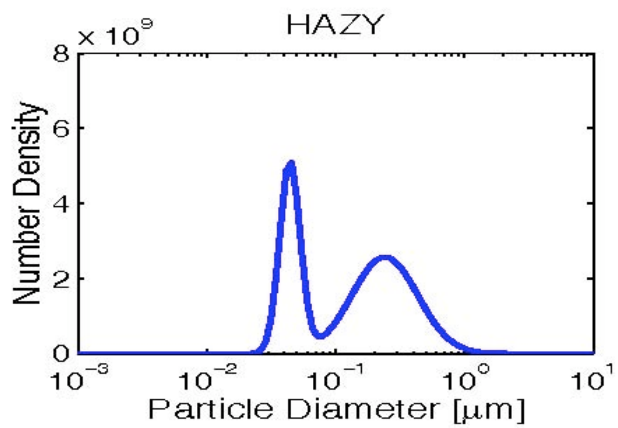
Discrete ADJ



Adjoints for Integral-PDE aerosol dynamic equations: formulation and challenges



Populations of aerosols (particles in the atmosphere) are described by their mass density



Aerosol dynamic equation - IPDE

$$\frac{\partial q_i}{\partial t} = \int_0^m \beta(m', m-m')\, q_i(m',t)\, \frac{q(m-m',t)}{m-m'}\, dm'$$

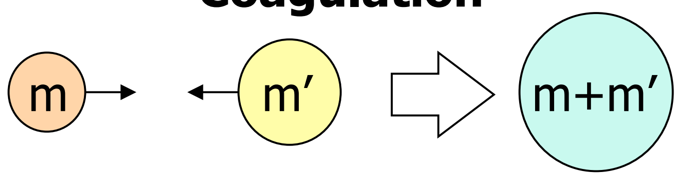
$$- q_i \int_0^\infty \beta(m,m')\, \frac{q(m',t)}{m'}\, dm'$$

$$+ H_i q - \frac{\partial}{\partial m}\left(m H q_i\right) + m_i S - L q_i + R_i(q)$$

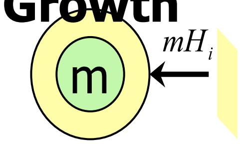
$$q_i\left(m, t=t^0\right) = q_i^0(m), \qquad 1 \le i \le n,$$

$$q_i\left(m=0,t\right) = 0, \quad q_i\left(m=\infty,t\right) = 0\ .$$

Coagulation



Growth



Virginia Tech

Adjoint aerosol dynamic models are needed to solve inverse problems

Continuous adjoint equation

$$\frac{\partial \lambda_i}{\partial t} = -\int_0^\infty \beta(m,m')(m')^{-1}\left[\lambda_i(m+m',t)-\lambda_i(m,t)\right]q(m',t)\,dm' + L\lambda_i \qquad t_+^{k-1} \le t \le t_-^k$$

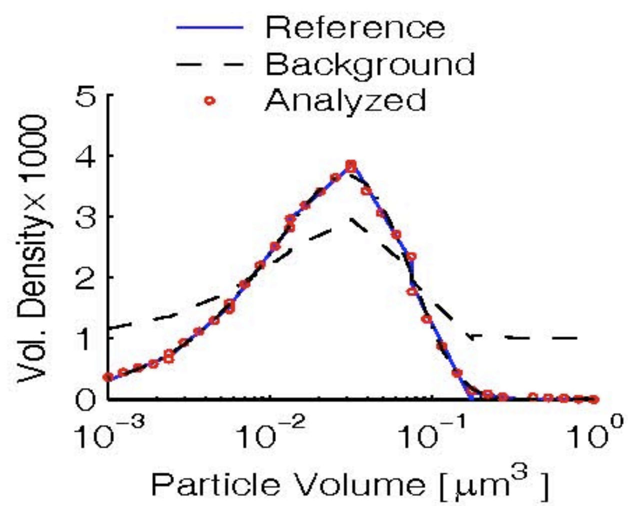
$$-\int_0^\infty \beta(m',m)\,m^{-1}\sum_{j=1}^{n}\left[\lambda_j(m+m',t)-\lambda_j(m,t)\right]q_j(m',t)\,dm' - \sum_{j=1}^{n} H_j\lambda_j - mH\frac{\partial \lambda_i}{\partial m}$$

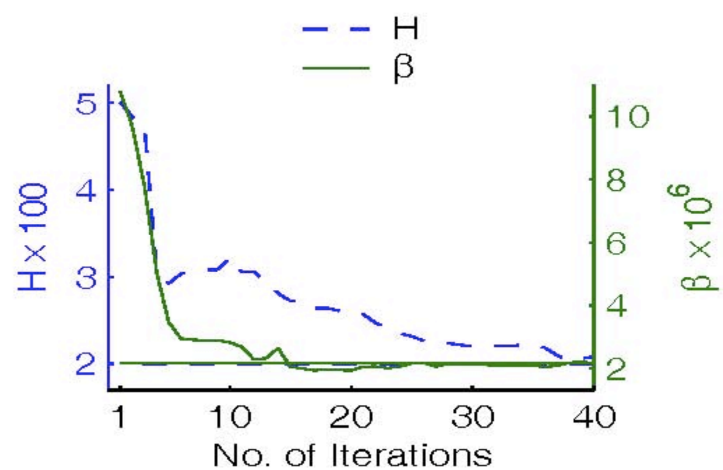
$$\lambda_i\left(m,t^N\right)=0, \quad \lambda_i(m,t_-^k) = \lambda_i(m,t_+^k) + h_{q_i}^T R_k^{-1}\left(y^k - h(q^k)\right)$$

$$\lambda_i\left(m,t^0\right)=\lambda_i\left(m,t_+^0\right)+ p_{q_i}^T B^{-1}\left(p-p^B\right) \quad \lambda_i\left(0,t\right)=0\ .$$

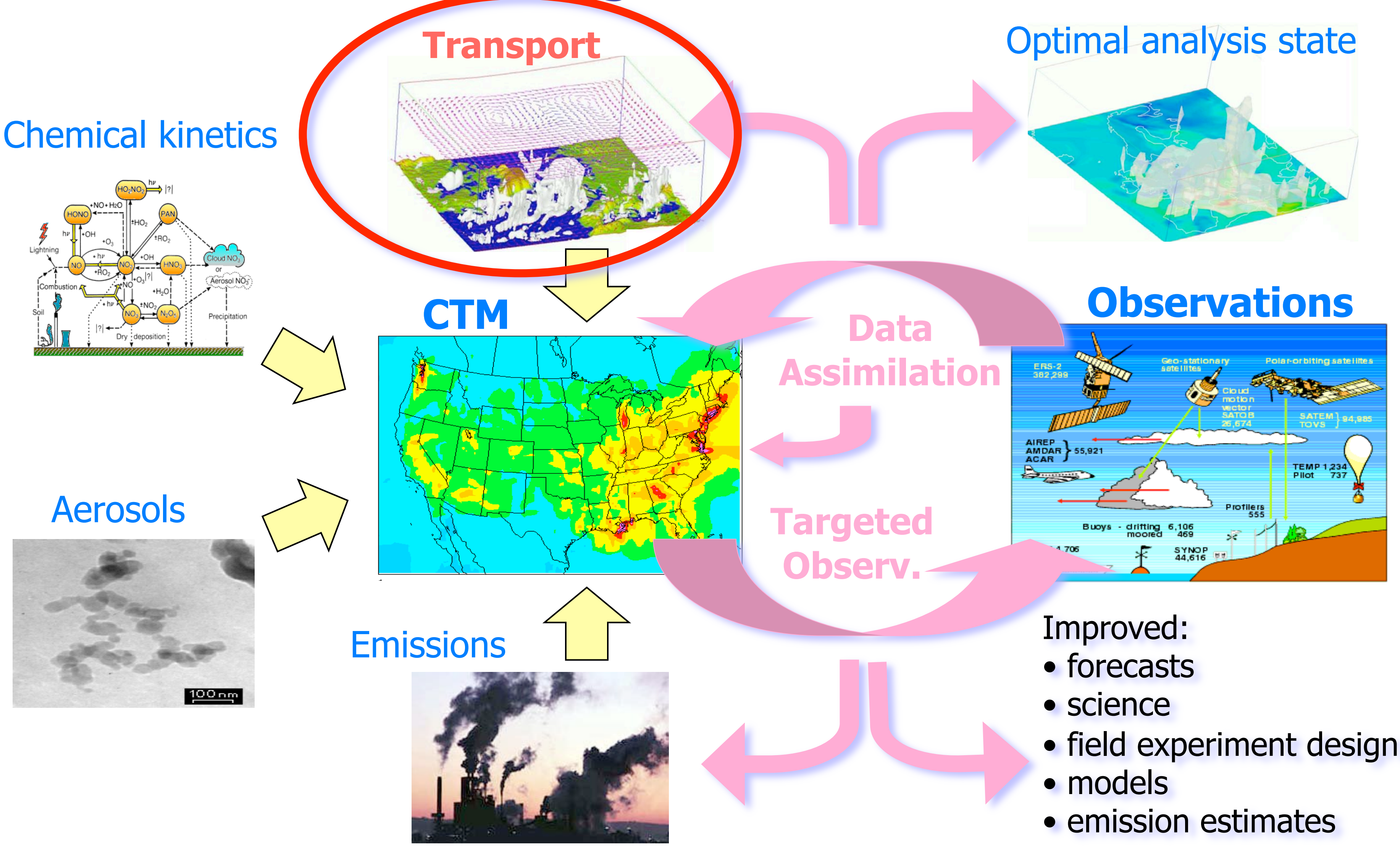
Observations of density in each bin allow the recovery of initial distribution and of parameters

[Sandu et. al., 2005; Henze et. al., 2004]



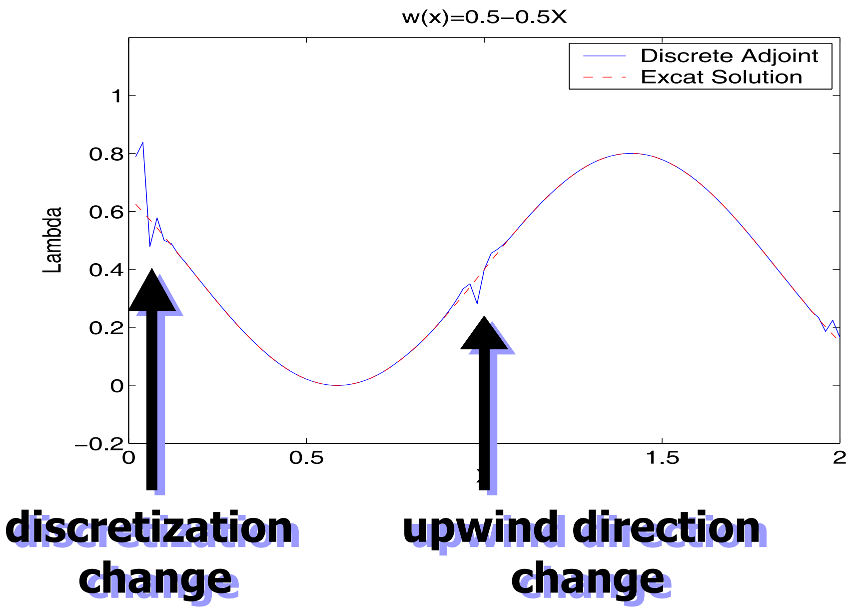


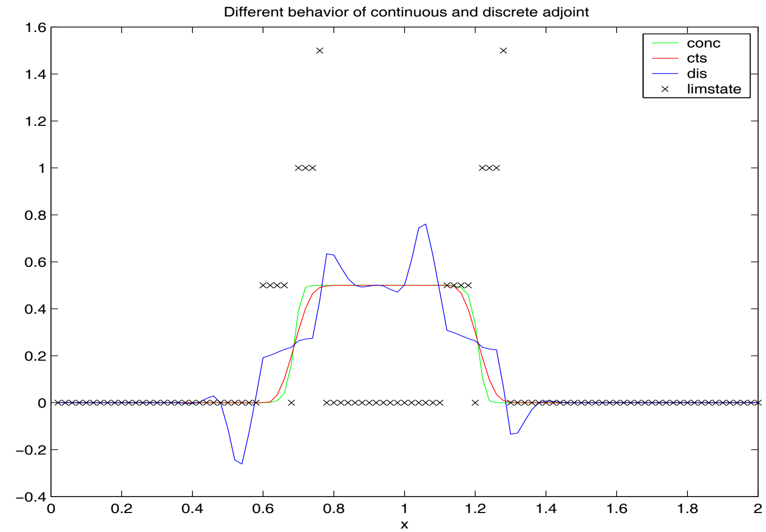
Discrete adjoint models for numerical advection: formulation and challenges



Virginia Tech

Discrete adjoints of advection numerical schemes can become pointwise inconsistent with the adjoint PDE





Change of forward scheme pattern:
- Change of upwinding
- Sources/sinks
- Inflow boundaries scheme

Example: 3rd order upwind FD

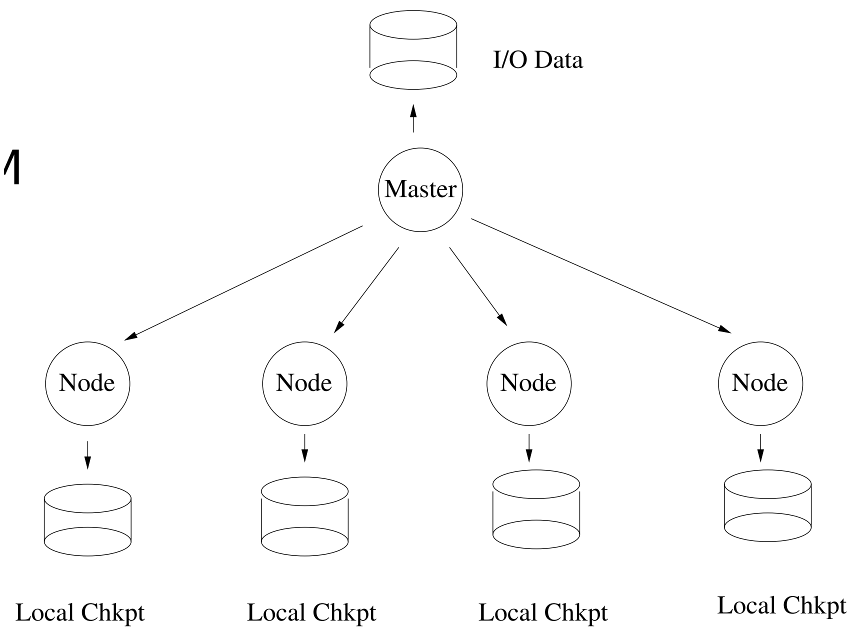
Active forward limiters act as pseudo-sources in adjoint

Example: minmod

[Liu and Sandu, 2005]

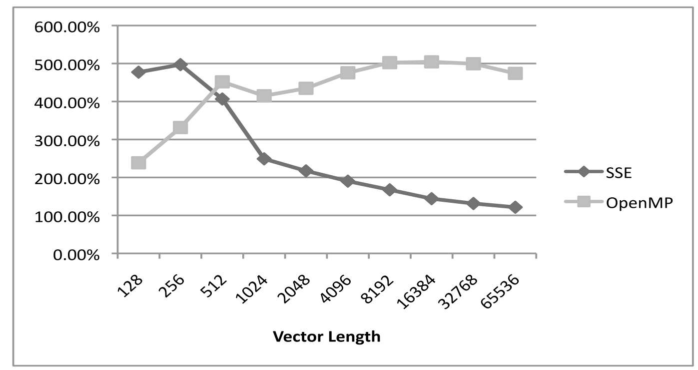
Parallelization important to speed up 4D-Var

Distributed, 2-level checkpointing scheme

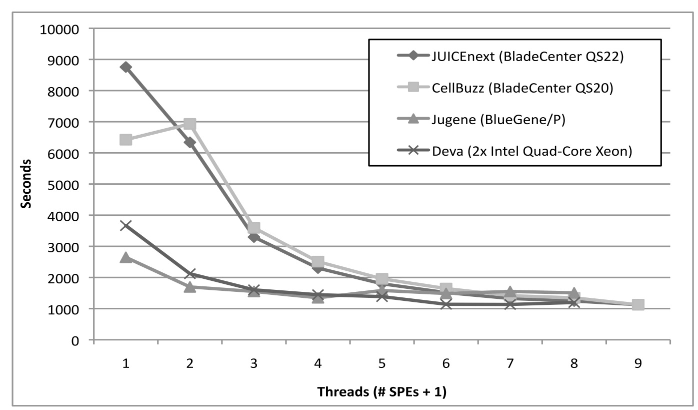


[Sandu et.al., 2003-2008]

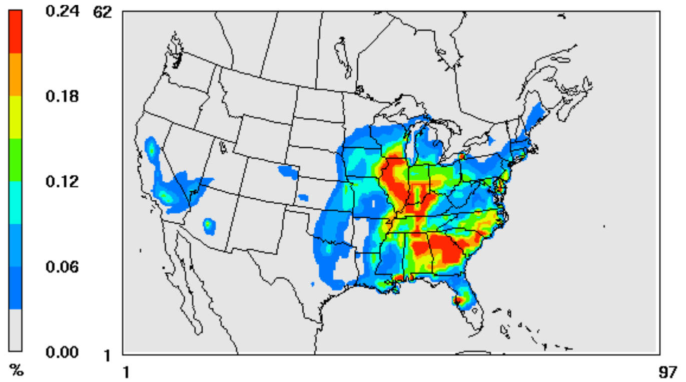
Chemistry w/ abstract vectors



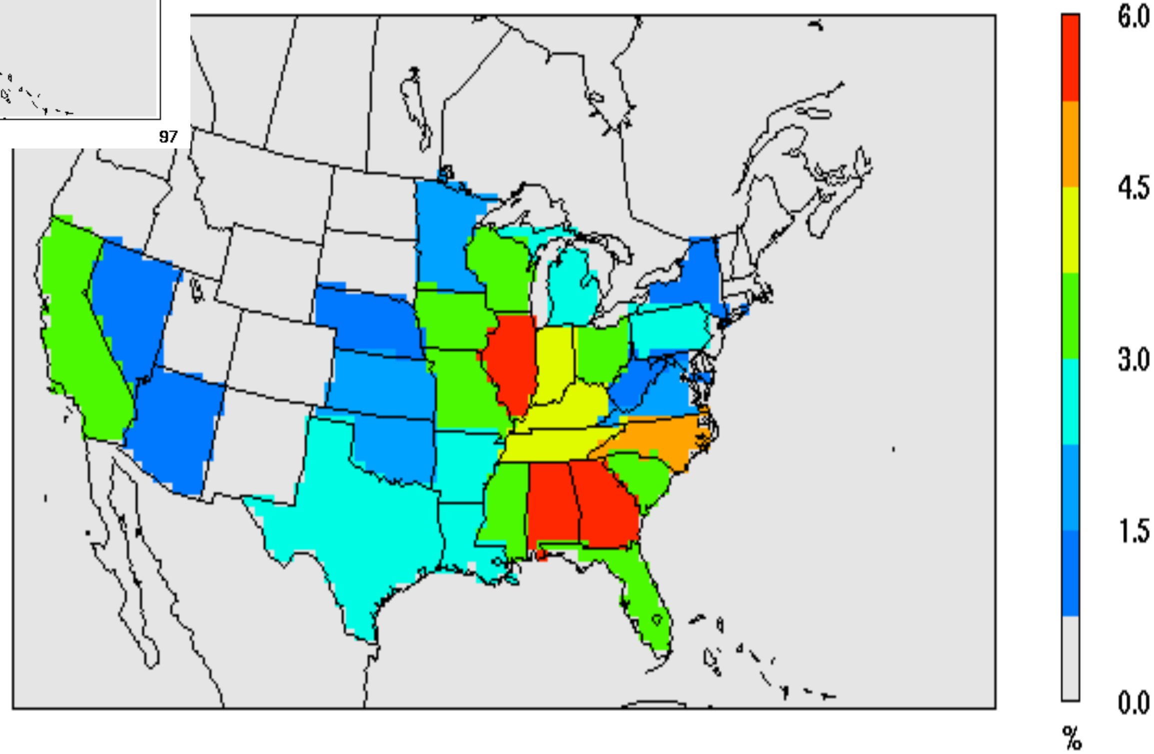
Transport on accelerators



STEM: Adjoint sensitivity analysis of non-attainment metrics can help guide policy decisions

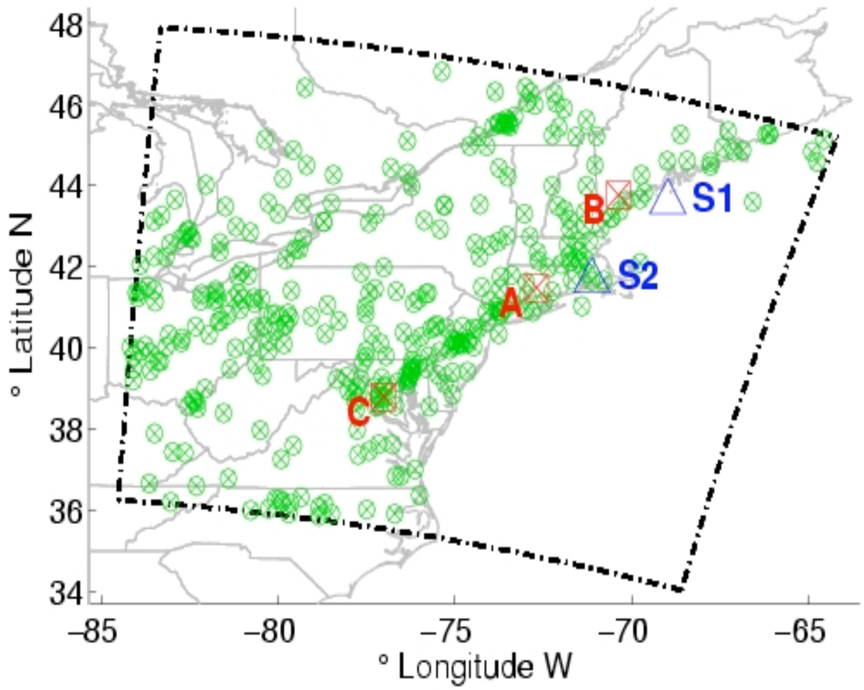


Estimated contributions by state to violating U.S. ozone NAAQS in July 2004



[Hakami et al., 2005]

Assimilation of ozone data from the ICARTT field campaign in Eastern U.S., July 2004



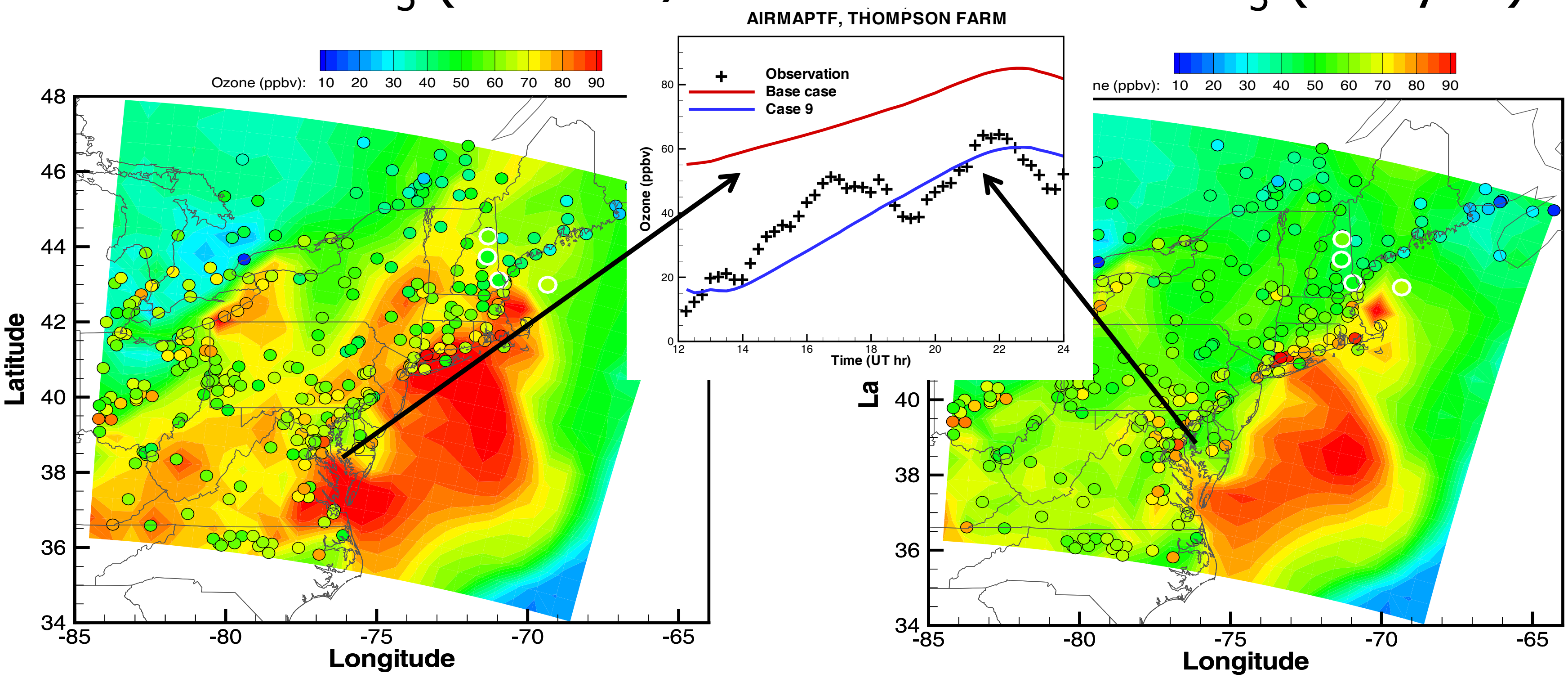
Observations	Description
AIRNOW	EPA surface stations, hourly averaged data used
DC3	Vertical profile of ozone mixing ratio from lidar
MOZ-FN	MOZAIC, Frankfurt-New York flight
MOZ-NF	MOZAIC, New York-Frankfurt flight
P3	NOAA P3-B measurement
AIRMAP	UV SPECTROSCOPY measurement at 4 sites
DC8-In	NASA In Situ Ozone via Nitric Oxide Chemiluminescence
DC8-Li	DC-8 Composite Tropospheric Ozone Cross-Sections
RHODE	Ozonesonde/Radiosonde data from Narragansett, RI
RONBR	Ozonesonde/Radiosonde data from the R/V Ronald H. Brown

[Chai et al., 2006]

STEM: Assimilation adjusts O_3 predictions considerably at 4pm EDT on July 20, 2004

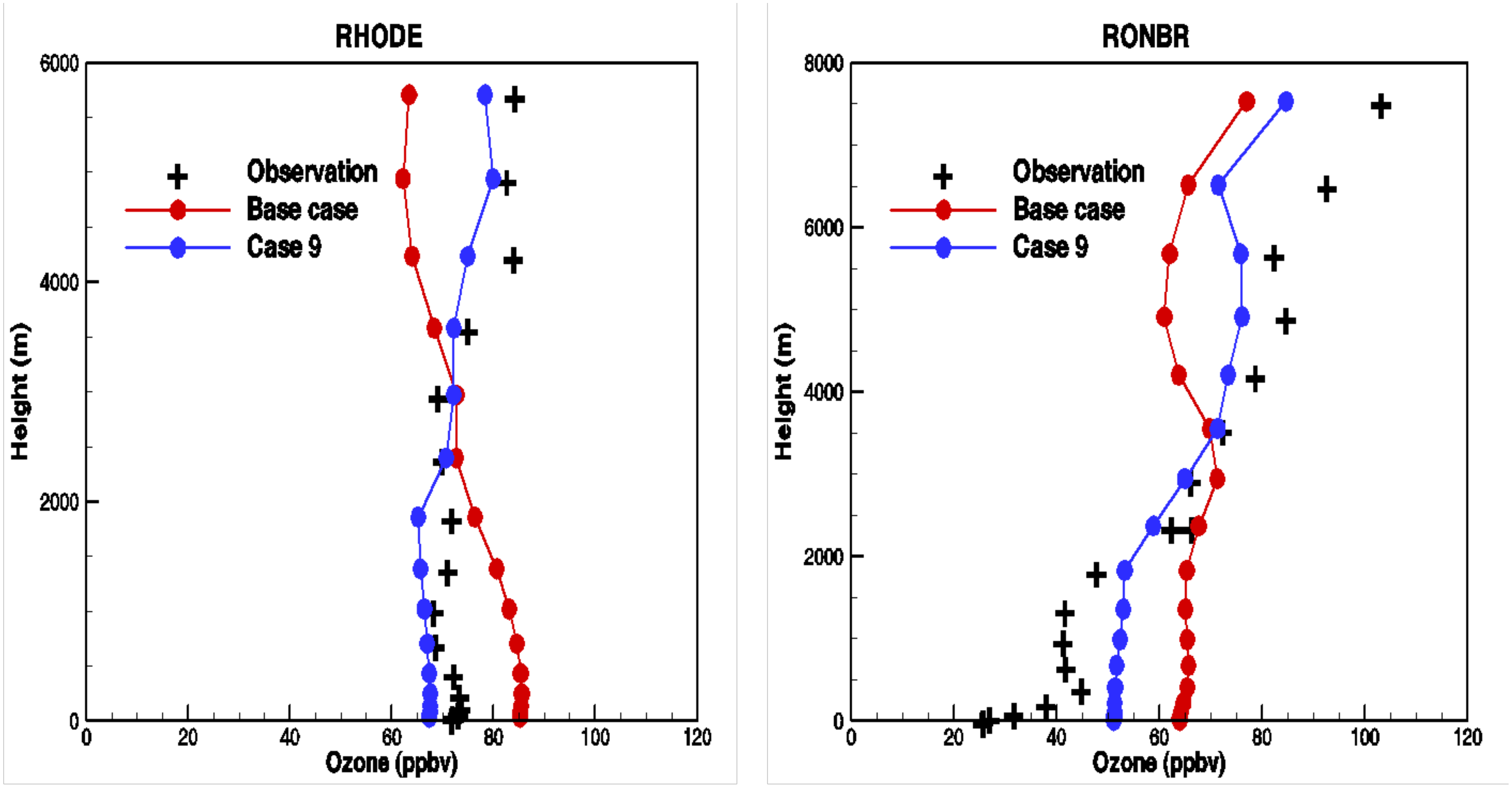
Observations: circles, color coded by O_3 mixing ratio

Surface O_3 (forecast) Surface O_3 (analysis)



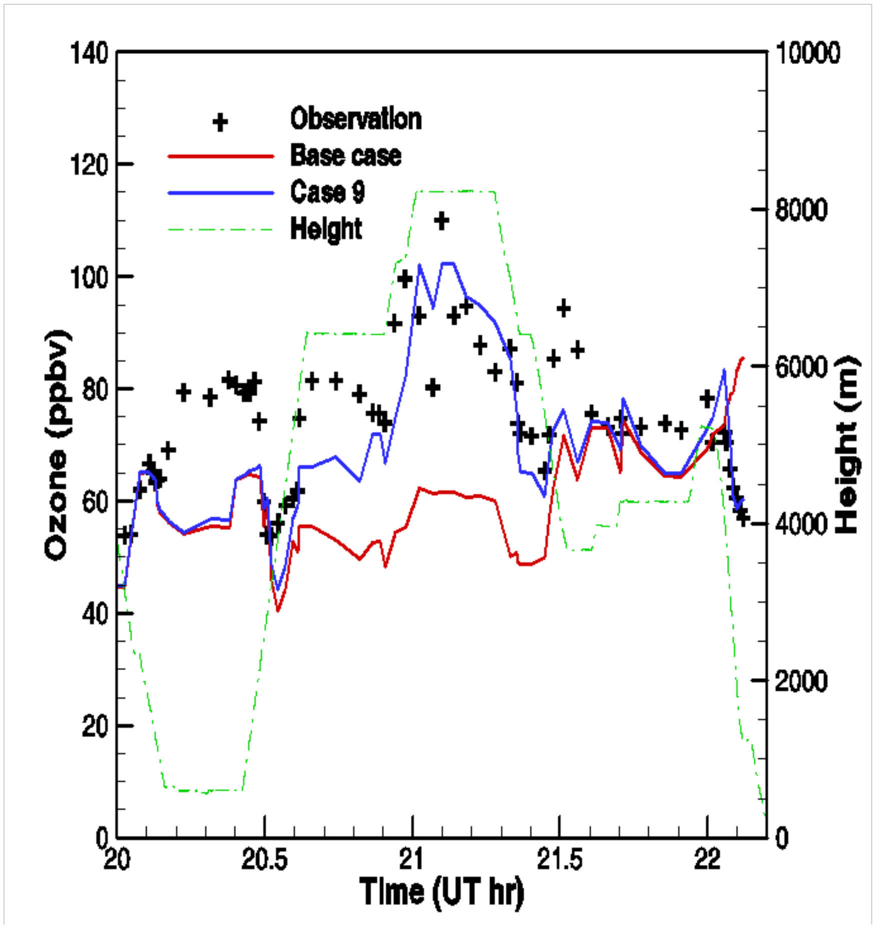
[Chai et al., 2006]

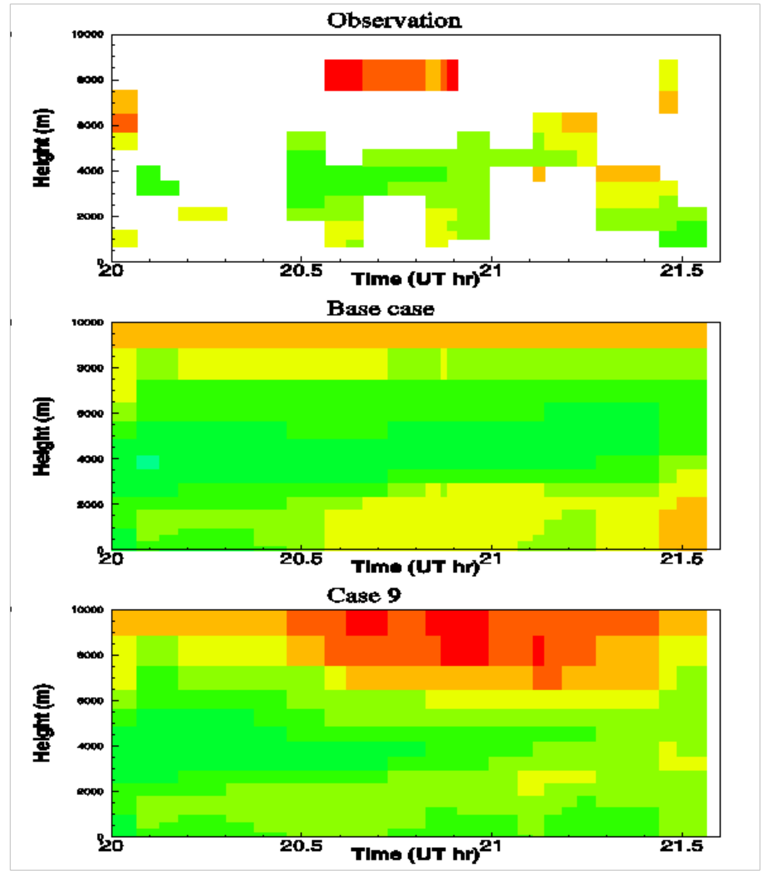
Assimilation of ozonesonde observations for July 20, 2004, show importance of vertical information



[Chai et al., 2006]

Assimilation of of DC-8 in-situ and lidar observations for July 20, 2004



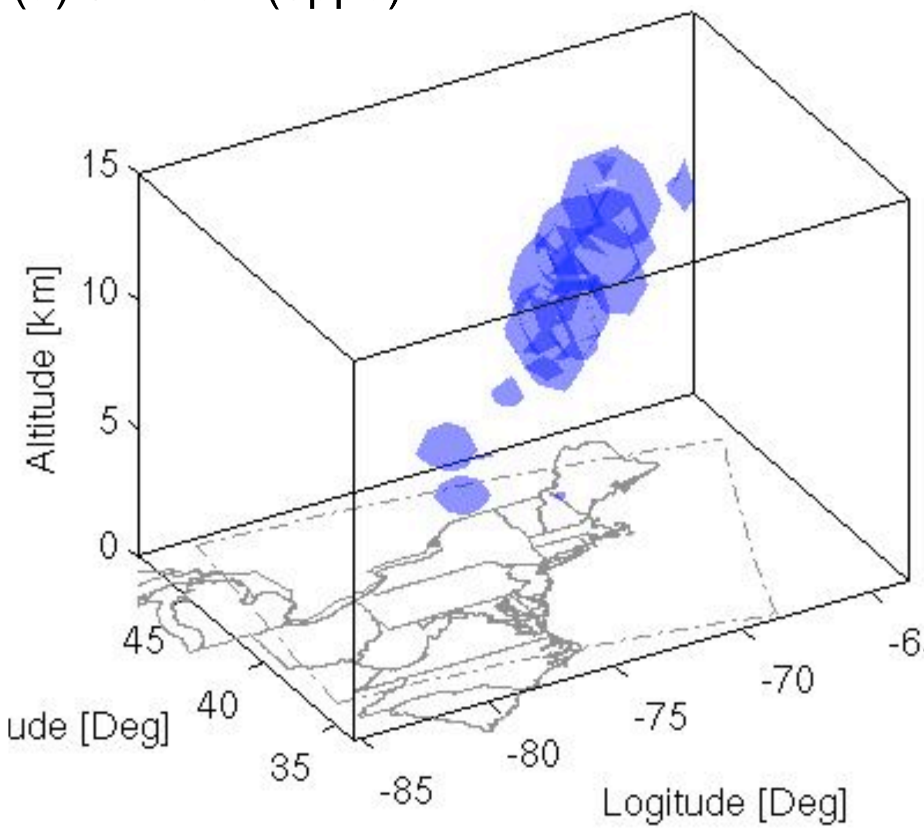


[Chai et al., 2006]

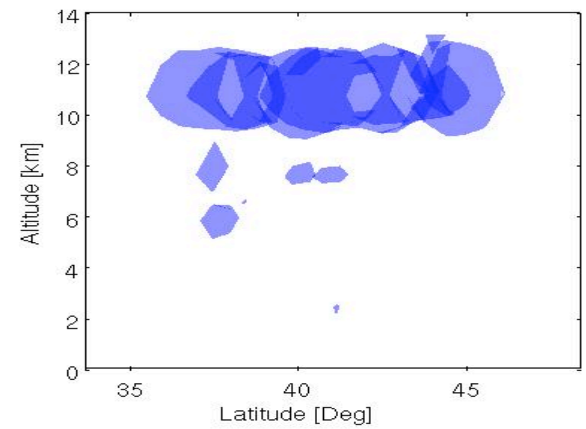
SOA: The smallest Hessian eigenvalues (vectors) approximate the principal error components

$$\left(\nabla^2_{y^0,y^0}\Psi\right)^{-1} \approx \mathrm{cov}\left(y^0\right)$$

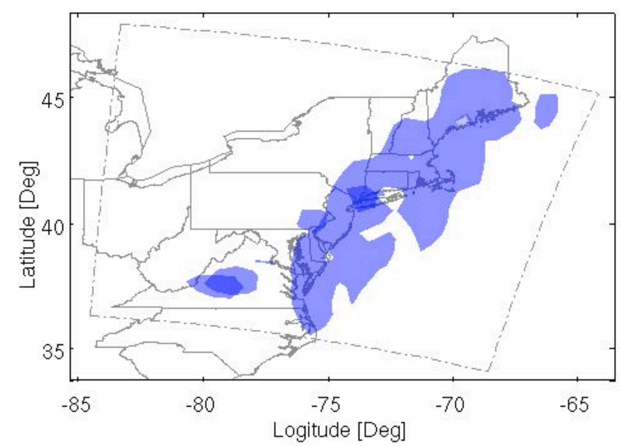
(a) 3D view (5ppb)



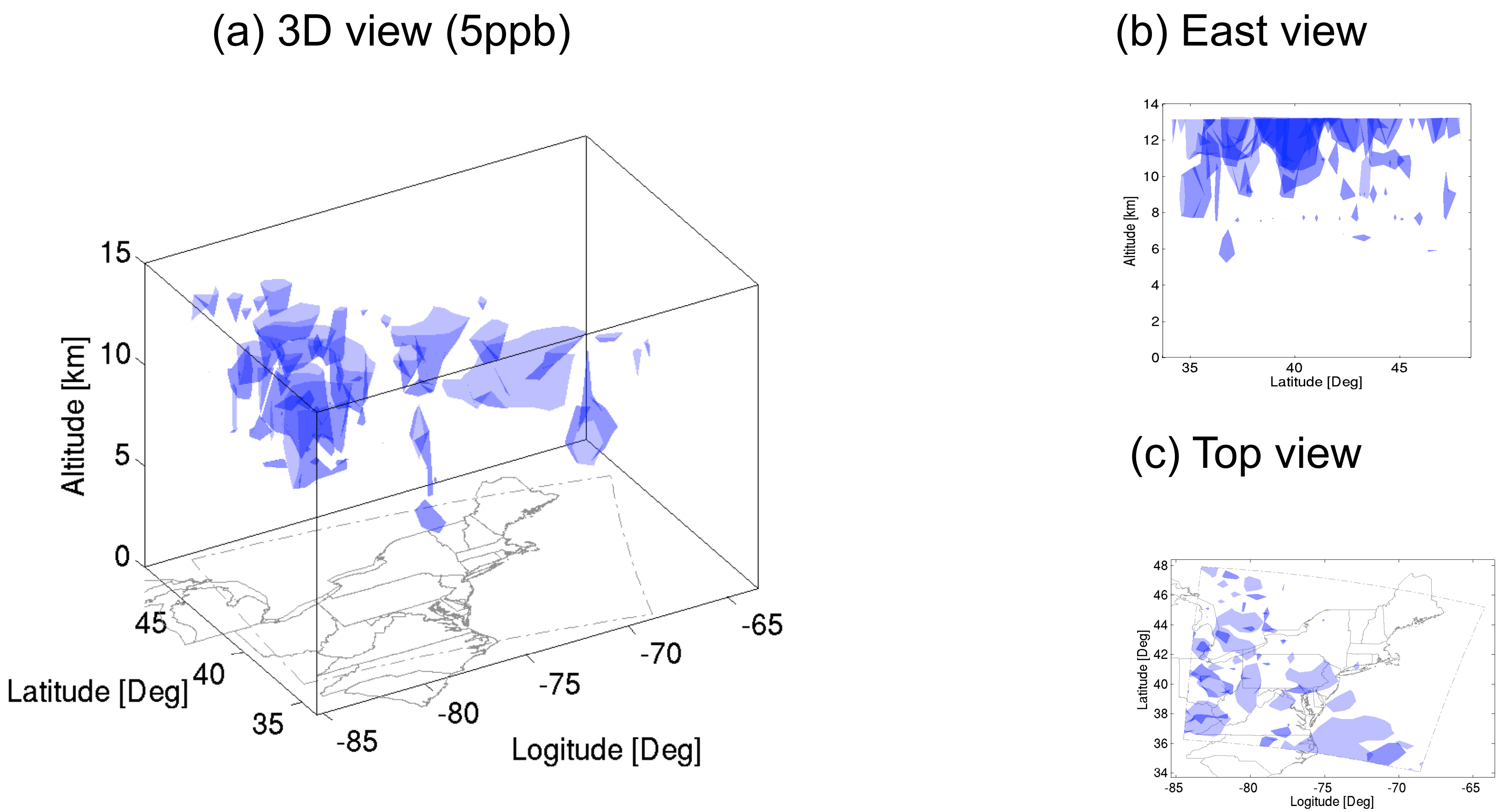
(b) East view



(c) Top view



SOA: Hessian singular vectors approximate the directions of maximum error growth – in finite time



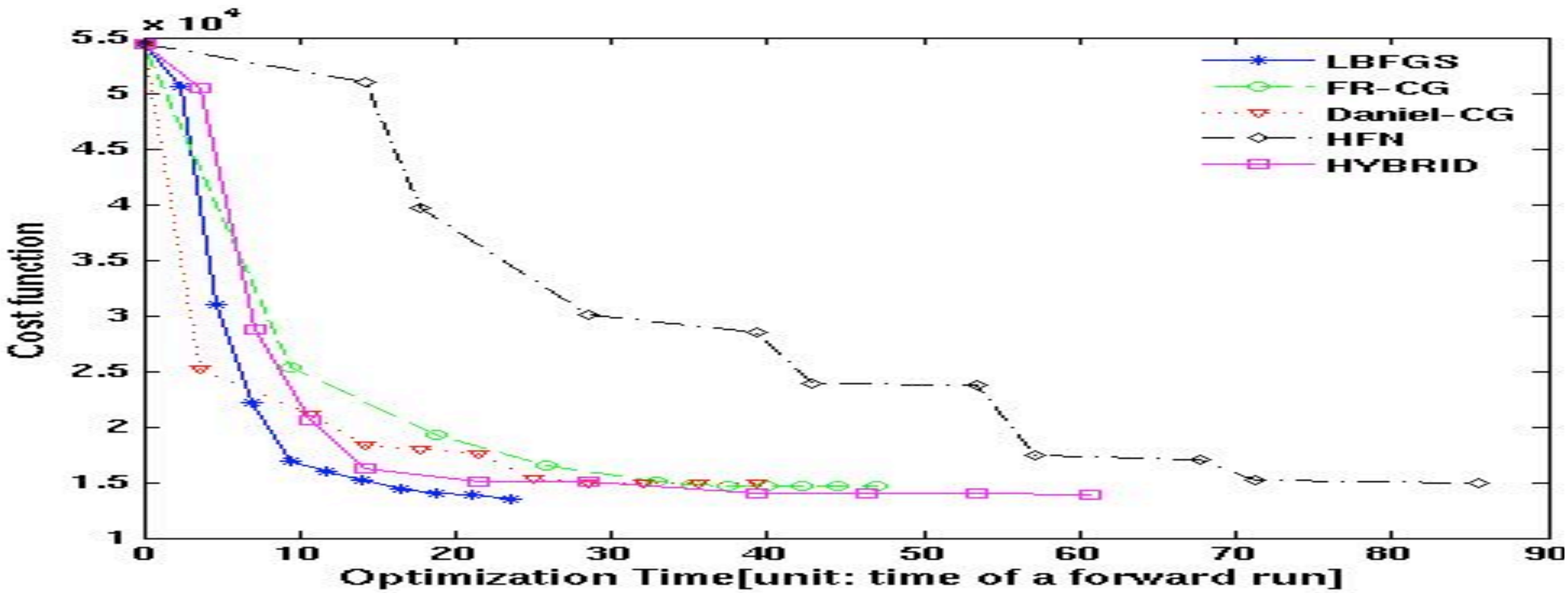
Virginia Tech

Large scale optimization methods use first and second order adjoint derivatives. BFGS performs well.

	BG	L-BFGS	FR-CG	Daniel	HFN	HYB
Grad.	4147	493	758	491	796	559
RMS	24.7	11.9	12.7	12.7	12.9	12.2
R^2	0.15	0.68	0.65	0.64	0.64	0.67

$$\boldsymbol{\lambda}^0 = \nabla_{\mathbf{y}^0} \psi$$

$$\boldsymbol{\sigma}^0 = \left(\nabla^2_{\mathbf{y}^0,\mathbf{y}^0} \psi\right) \delta y^0$$

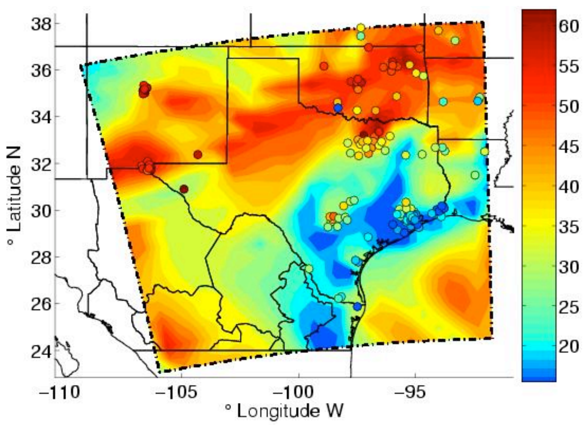


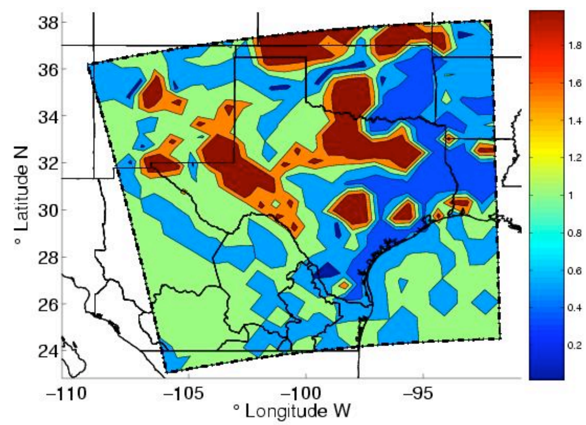
[Zhang and Sandu, 2007]

STEM: The inversion procedure can be extended to emissions, boundary conditions, etc.

Texas: 4am CST July 16 to 8pm CST on July 17, 2004.

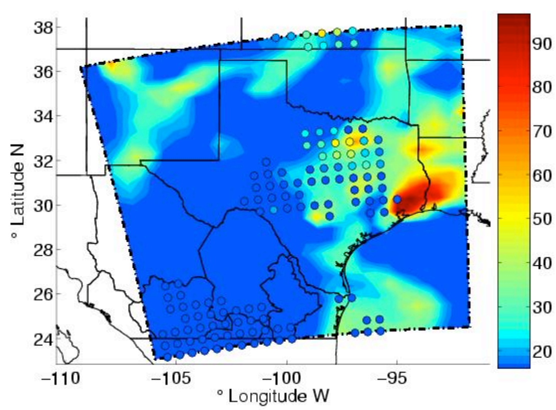
O_3
AirNow

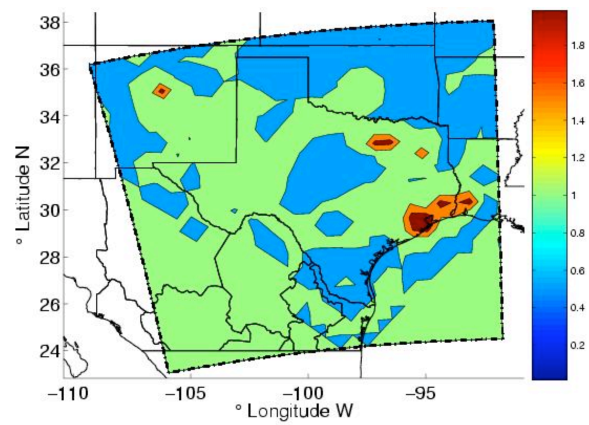




NO_2
emission
corrections

NO_2
Schiamacy





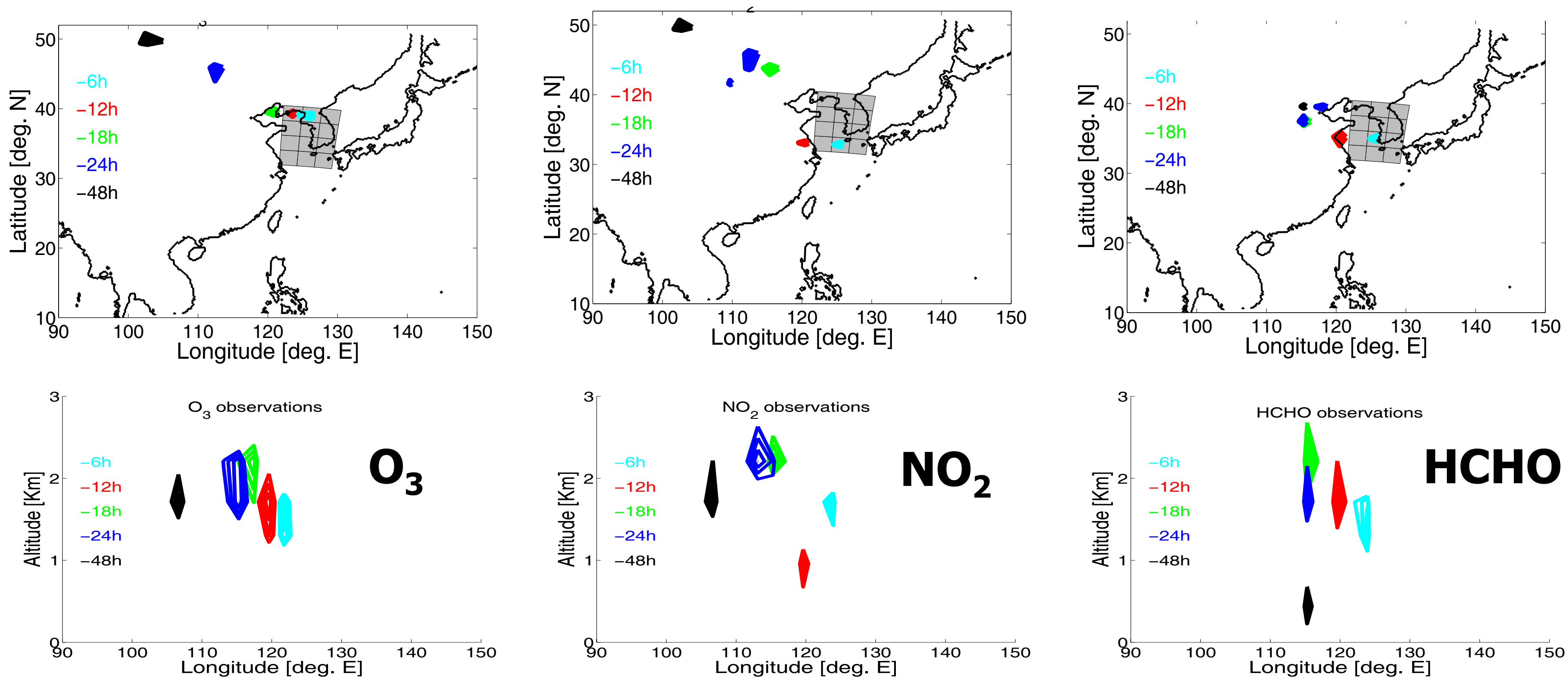
HCHO
emission
corrections

STEM: Best observation locations are different for different chemical species

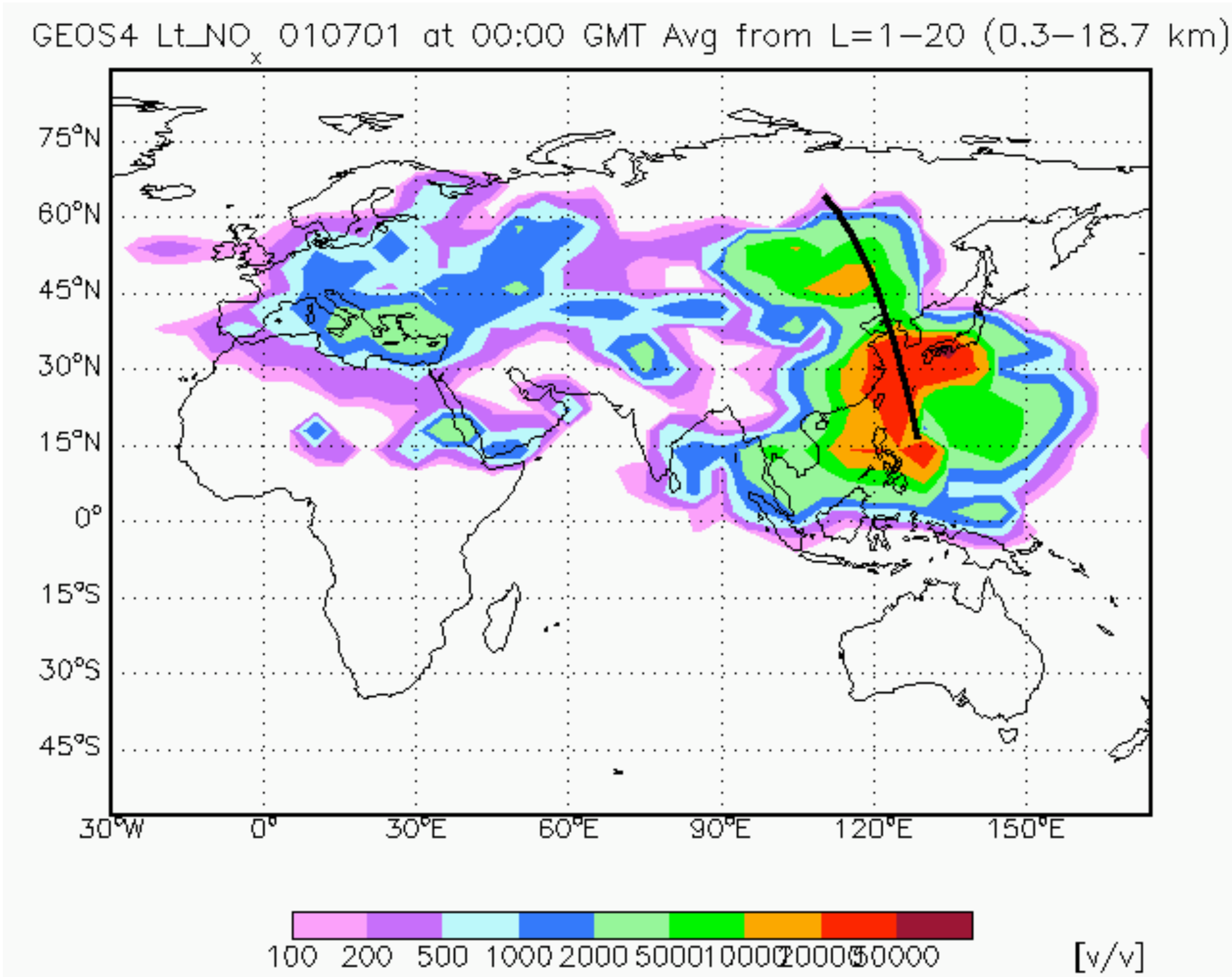
$$T = \sum_{k \geq 1} \frac{\sigma_k^2}{\sigma_{\max}^2} s_k^2$$

(criterion based on SVs)

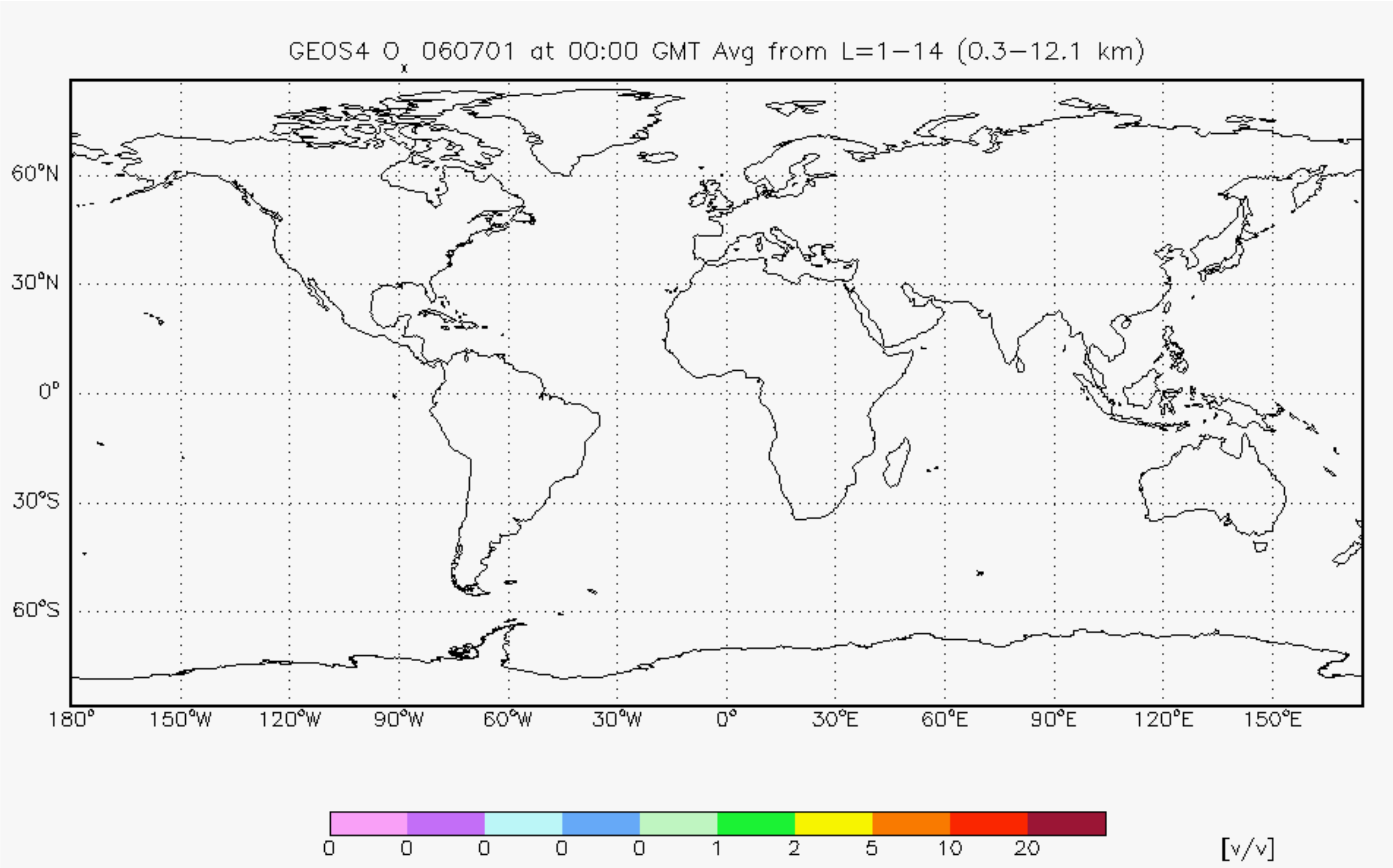
Verification:
Korea, ground O_3
0 GMT, Mar/4/2001



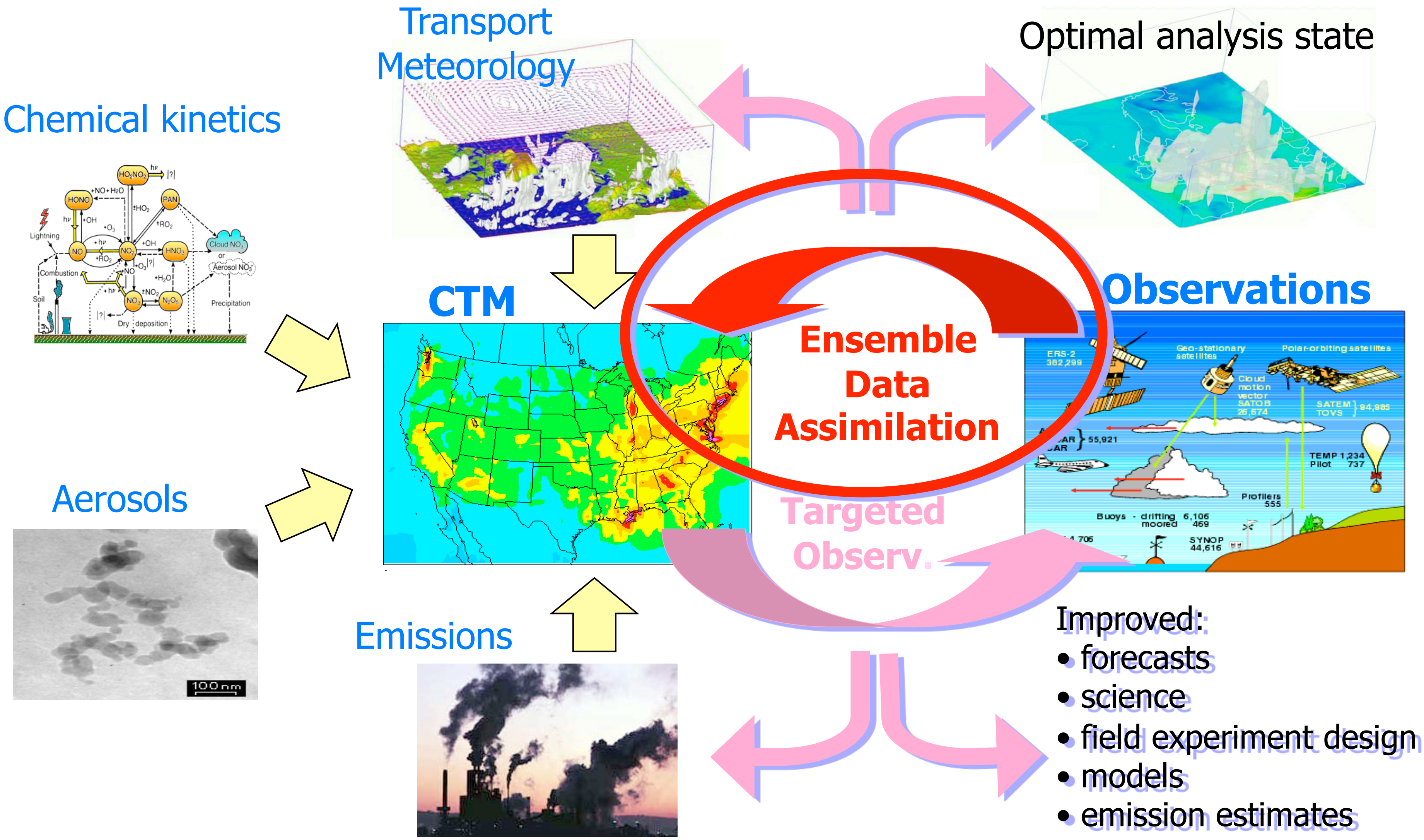
GEOS-CHEM-ADJ: Adjoints of satellite-observed ozone with respect to lightning NOx emissions in April 2004



Assimilation of TES ozone observations in GEOS-Chem



Ensemble-based chemical data assimilation is an alternative to variational techniques



Virginia Tech

The Ensemble Kalman Filter (EnKF) popular in NWP but not extensively used before with CTMs

$$\mathbf{y}_f^k = M\left(t^{k-1}, \mathbf{y}_a^{k-1}\right)$$

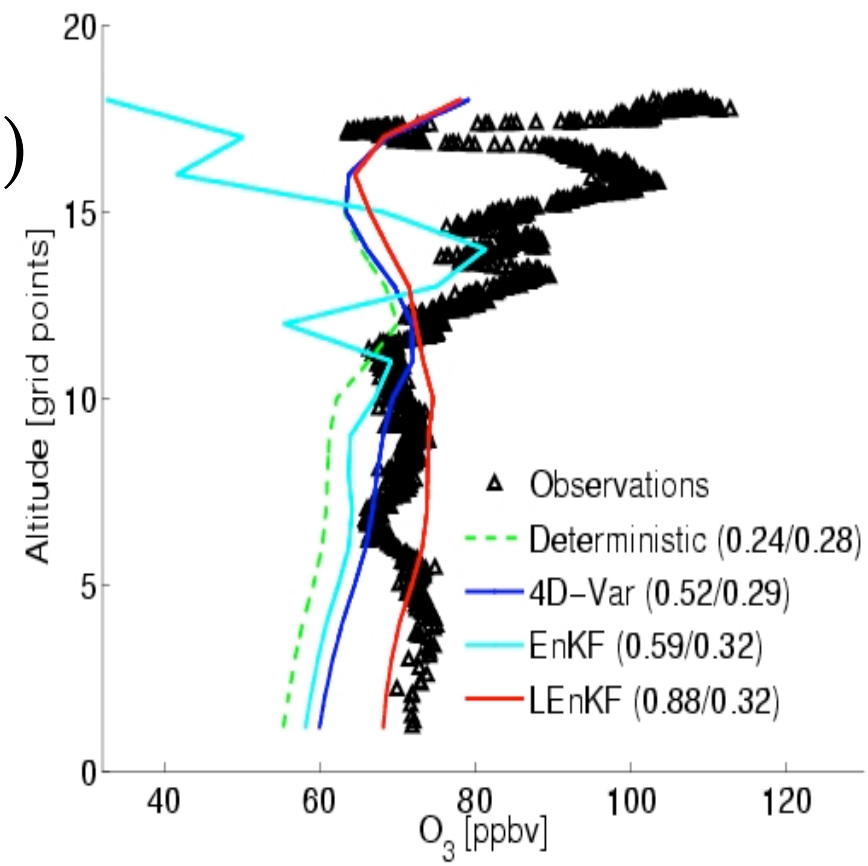
$$\mathbf{y}_a^k = \mathbf{y}_f^k + \mathbf{P}_f^k \mathbf{H}_k^T \left(\mathbf{R}_k + \mathbf{H}_k \mathbf{P}_f^k \mathbf{H}_k^T\right)^{-1} \left(\mathbf{z}_{obs}^k - \mathbf{H}_k \mathbf{y}_f^k\right)$$

Specify initial ensemble (sample B)

Covariance inflation: Prevents filter divergence (additive, multiplicative, model-specific)

Covariance localization (limit long-distance spurious correlations)

Correction localization (limit increments away from observations)



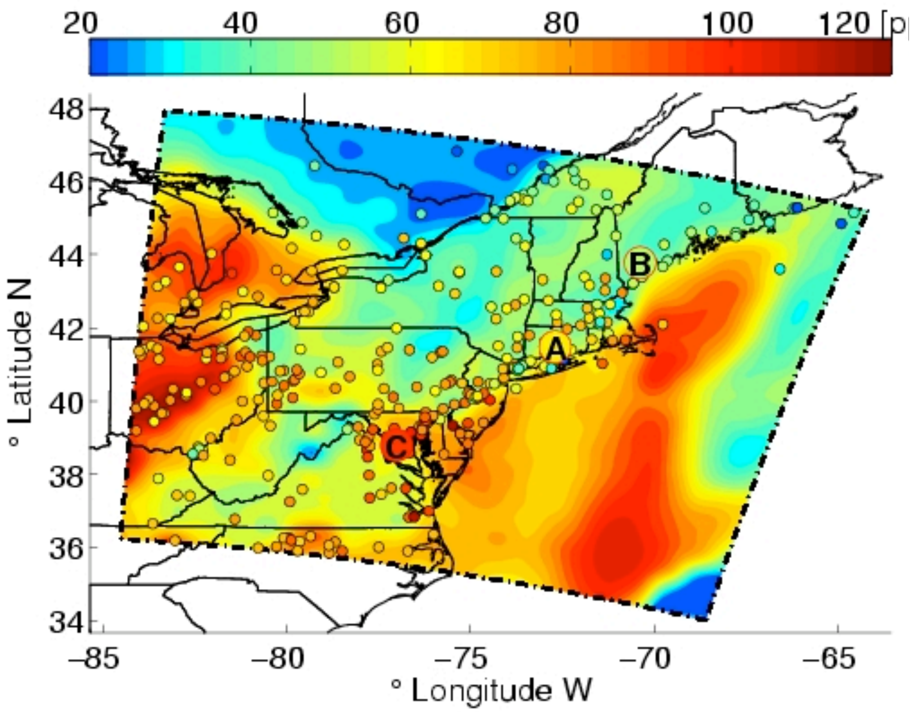
Ozonesonde S2 (18 EDT, July 20, 2004)

[Constantinescu et al., 2007]

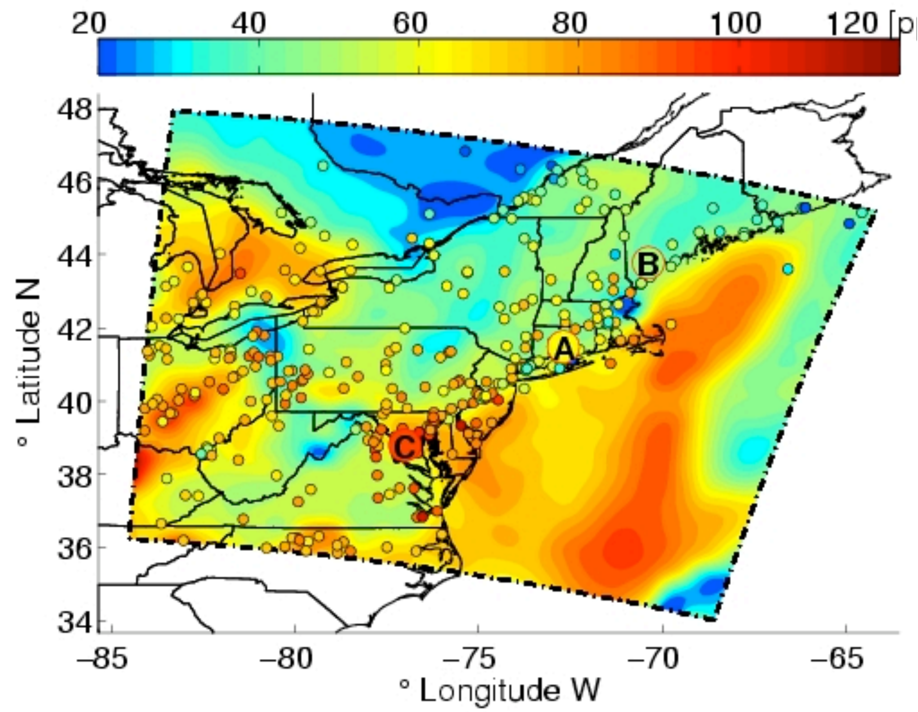
LEnKF assimilation of emissions and boundaries together with the state can improve the forecast

Ground level ozone at 2pm EDT, July 21, 2004 (in forecast window)

LEnKF (R^2=0.88/0.32)
[state only]



LEnKF (R^2=0.88/0.42)
[state + emissions + boundary]



Hybrid DA combines 4D-Var and ensembles. Obtain analysis covariance at the end of assimilation window

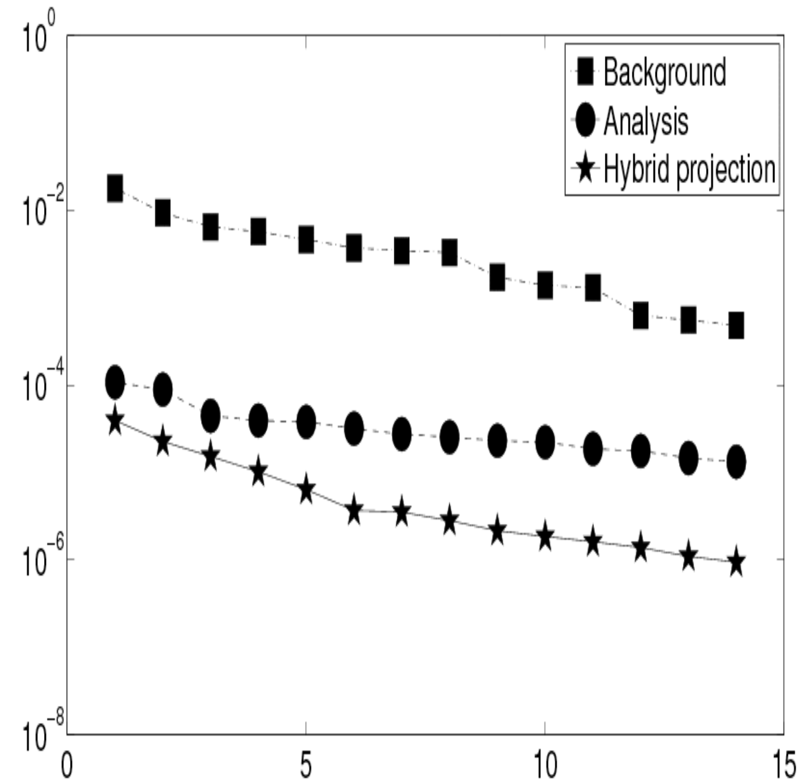
$$S_0 = range\left[x_0^{(1)} - x_0^{B}, \cdots, x_0^{(K)} - x_0^{(K-1)}\right] \approx range\left[u_1, \cdots, u_m\right]$$

$$S_F = range\left[x_F^{(1)} - x_F^{B}, \cdots, x_F^{(K)} - x_F^{(K-1)}\right] \approx range\left[v_1, \cdots, v_m\right]$$

$$\Delta x_F^H[i] = \left(I - VV^T\right)\Delta x_F^B[i] \approx \Delta x_F^A[i], \quad i = 1, \cdots, Nens$$

Lorenz 96:

Covariance eigenvalues



Dynamic integration of chemical data and atmospheric models is an important, growing field

The tools needed for 4d-Var chemical data assimilation are in place:

- adjoints for stiff systems, aerosols, transport;
- theoretical and computational understanding of discrete and continuous adjoints;
- second order adjoints
- optimization methods
- singular vectors,
- parallelization, multi-level checkpointing schemes,
- models of background errors

- their strengths demonstrated using real (field campaign) data; ambitious science projects are ongoing

Continuous and discrete adjoints of mass balance equations lead to different computational models

$$\nabla_{\mathbf{y}^0}\psi = \cdots + \sum_{k=1}^{N} \left(\partial \mathbf{y}^k / \partial \mathbf{y}^0\right)^{\mathbf{T}} \left(\mathbf{H}^k\right)^{\mathbf{T}} \mathbf{R}_k^{-1} \left(\mathbf{H}^k \mathbf{y}^k - \mathbf{z}^k\right)$$

Continuous forward model

$$\frac{dC_i}{dt} = -\vec{u}\cdot\nabla C_i + \frac{1}{\rho}\nabla\left(\rho K \cdot \nabla C_i\right) + \frac{1}{\rho} f_i(\rho C) + E_i$$

$$C_i\left(t^0, x\right) = C_i^0(x), \qquad t^0 \le t \le t^F$$

$$C_i(t,x) = C_i^{IN}(t,x) \quad on \quad \Gamma^{IN}$$

$$K\frac{\partial C_i}{\partial n} = 0 \quad on \quad \Gamma^{OUT}$$

$$K\frac{\partial C_i}{\partial n} = V_i^{DEP} C_i - Q_i \quad on \quad \Gamma^{GROUND}$$

adj

Continuous adjoint model

$$\frac{d\lambda_i}{dt} = -\nabla\cdot\left(\vec{u}\lambda_i\right) - \nabla\cdot\left(\rho K \cdot \nabla \frac{\lambda_i}{\rho}\right) - \left(F^T(\rho C)\cdot\lambda\right)_i - \phi_i$$

$$\lambda_i\left(t^F, x\right) = \lambda_i^F(x), \qquad t^F \ge t \ge t^o$$

$$\lambda_i(t,x) = 0 \quad on \quad \Gamma^{IN}$$

$$\vec{u}\lambda_i + \rho K \frac{\partial\left(\lambda_i/\rho\right)}{\partial \vec{n}} = 0 \quad on \quad \Gamma^{OUT}$$

$$\rho K \frac{\partial\left(\lambda_i/\rho\right)}{\partial \vec{n}} = V_i^{DEP}\lambda_i \quad on \quad \Gamma^{GROUND}$$

discr

discr

Discrete forward model

$$C^{k+1} = N_{[t,t+\Delta t]} \circ C^k$$

$$N_{[t,t+\Delta t]} = T_{HOR}^{\Delta t} \circ T_{VERT}^{\Delta t} \circ R_{CHEM}^{\Delta t} \circ T_{VERT}^{\Delta t} \circ T_{HOR}^{\Delta t}$$

adj

Computational adjoint model

$$\lambda^k = N'^{*}_{[t,t+\Delta t]} \circ \lambda^{k+1} + \phi^{k+1}$$

$$N'^{*}_{[t,t+\Delta t]} = \left(T_{HOR}^{\Delta t}\right)^* \circ \left(T_{VERT}^{\Delta t}\right)^* \circ \left(R_{CHEM}^{\Delta t}\right)^* \circ \left(T_{VERT}^{\Delta t}\right)^* \circ \left(T_{HOR}^{\Delta t}\right)^*$$

Second order adjoints provide Hessian-vector products useful in optimization and analysis

$$\min_{\mathbf{y}^0} \psi(\mathbf{y}^0) = \tfrac{1}{2}(\mathbf{y}^0 - \mathbf{y}^b)^T \mathbf{B}^{-1}(\mathbf{y}^0 - \mathbf{y}^b) + \tfrac{1}{2}\sum_{\mathbf{k}=1}^{\mathbf{N}} (\mathbf{H}^k \mathbf{y}^k - \mathbf{o}^k)^T \mathbf{R}_k^{-1}(\mathbf{H}^k \mathbf{y}^k - \mathbf{o}^k)$$

$$\boldsymbol{\lambda}^0 = \nabla_{\mathbf{y}^0}\psi \qquad \boldsymbol{\sigma}^0 = \left(\nabla^2_{\mathbf{y}^0}\psi\right)\delta y^0 \qquad \mathrm{cov}(\mathbf{y}^0) \approx \left(\nabla^2_{\mathbf{y}^0}\psi\right)^{-1}$$

RK & TLM Methods (KPP)

$$\mathbf{y}^{n+1} = \mathbf{y}^n + h\sum_{i=1}^{s} b_i \mathbf{f}(\mathbf{Y}^i), \qquad \mathbf{Y}^i = \mathbf{y}^n + h\sum_{i=1}^{s} a_{i,j}\mathbf{f}(\mathbf{Y}^j)$$

$$\boldsymbol{\delta}\mathbf{y}^{n+1} = \boldsymbol{\delta}\mathbf{y}^n + h\sum_{i=1}^{s} b_i \mathbf{J}(\mathbf{Y}^i)\boldsymbol{\delta}\mathbf{Y}^i, \quad \boldsymbol{\delta}\mathbf{Y}^i = \boldsymbol{\delta}\mathbf{y}^n + h\sum_{i=1}^{s} a_{i,j}\mathbf{f}(\mathbf{Y}^j)\boldsymbol{\delta}\mathbf{Y}^j$$

First and Second Order RK Discrete Adjoints (KPP)

$$\boldsymbol{\lambda}^n = \boldsymbol{\lambda}^{n+1} + \sum_{i=1}^{s}\mathbf{u}^i + \frac{\partial\boldsymbol{\Phi}}{\partial\mathbf{y}^n}, \quad \mathbf{u}^i = h\,\mathbf{J}^T(\mathbf{Y}^i)\cdot\left(b_i\boldsymbol{\lambda}^{n+1} + \sum_{j=1}^{s} a_{j,i}\mathbf{u}^j\right)$$

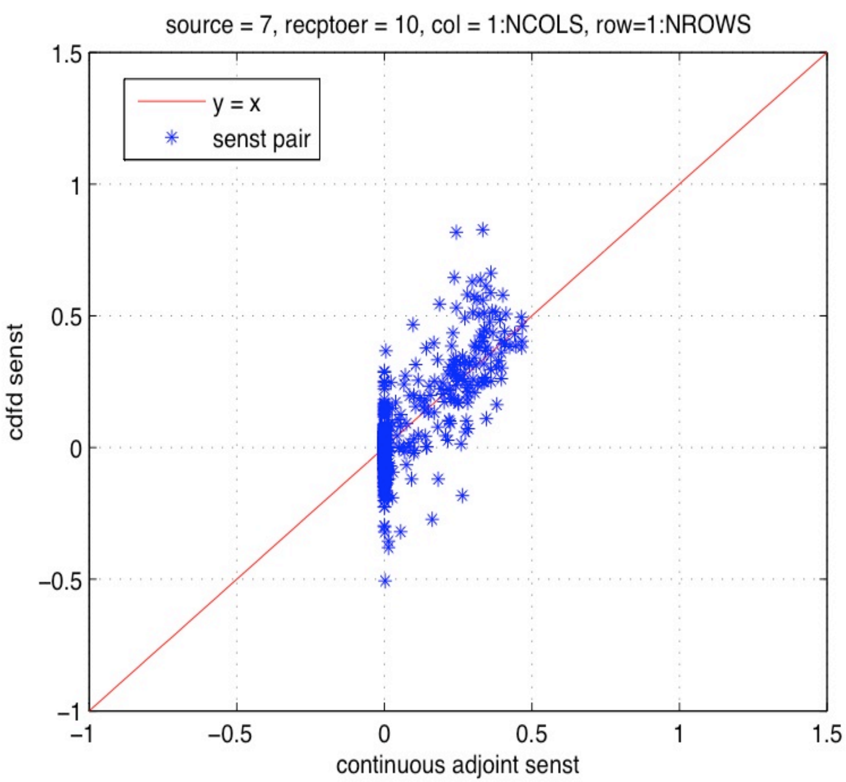
$$\boldsymbol{\sigma}^n = \boldsymbol{\sigma}^{n+1} + \sum_{i=1}^{s}\mathbf{w}^i + \frac{\partial^2\boldsymbol{\Phi}}{\partial^2\mathbf{y}^n}\boldsymbol{\delta}\mathbf{y}^n,$$

$$\mathbf{w}^i = h\,\mathbf{J}^T(\mathbf{Y}^i)\cdot\left(b_i\boldsymbol{\sigma}^{n+1} + \sum_{j=1}^{s} a_{j,i}\mathbf{w}^j\right) + h\left(\mathbf{H}(\mathbf{Y}^i)\times\boldsymbol{\delta}\mathbf{Y}^i\right)^T\cdot\left(b_i\boldsymbol{\lambda}^{n+1} + \sum_{j=1}^{s} a_{j,i}\mathbf{u}^j\right)$$

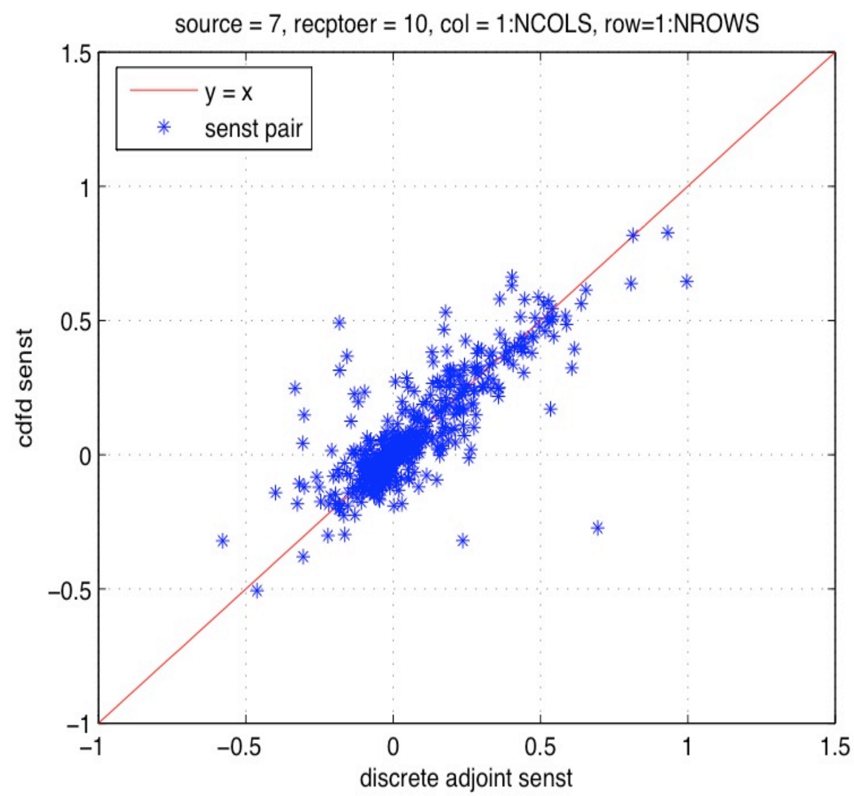
[Sandu et. al., 2005]

CMAQ: discrete advection adjoints match better finite differences than continuous advection adjoints

Continuous z-advection adjoint

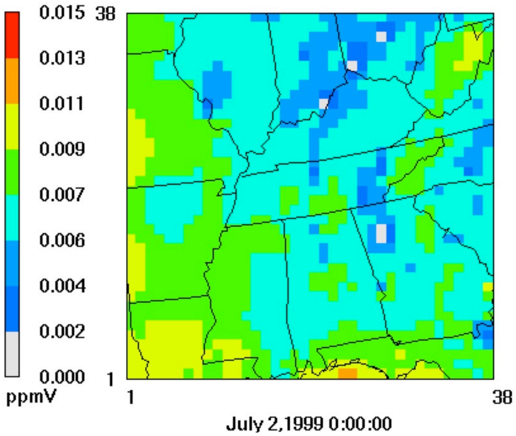


Discrete z-advection adjoint

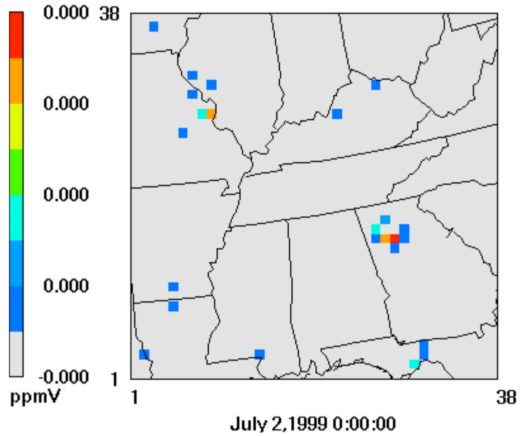


CMAQ: Optimization converges faster with continuous advection adjoints

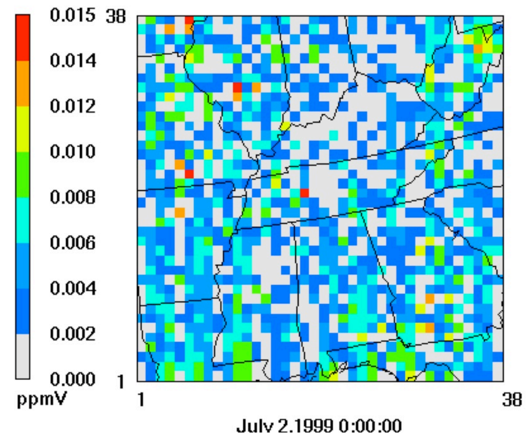
Background-Ref



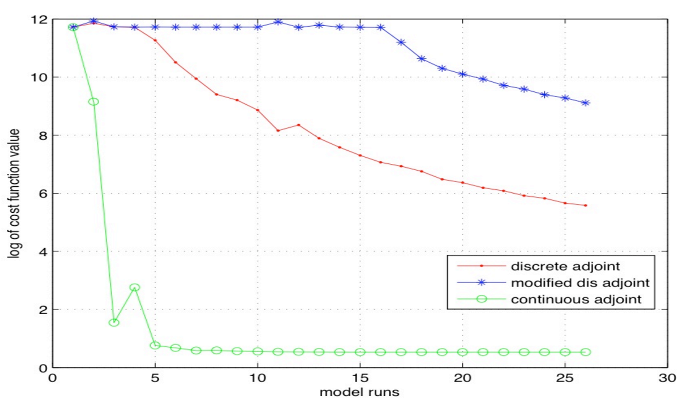
Opt-Ref (Cts)



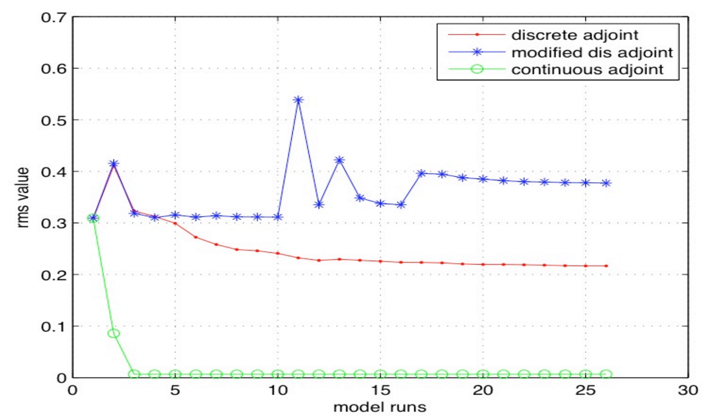
Opt-Ref (Dis)



Cost Function



RMS error



Virginia Tech