Computational Aspects of Chemical Data Assimilation in Air Quality Models

Greg Carmichael (Dept. of Chem. Eng., U. Iowa)

Adrian Sandu (Dept. of Comp. Sci., Mich. Inst. Tech.)

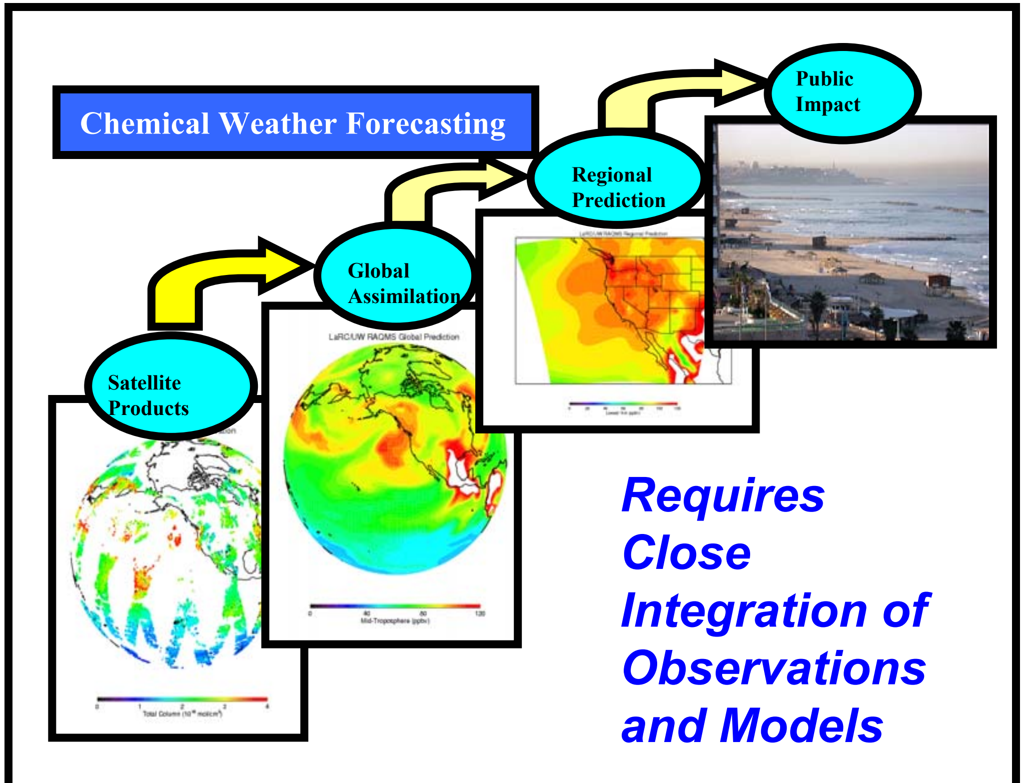
John Seinfeld (Dept. Chem. Eng., Cal. Tech.)

Tad Anderson (Dept. Atmos. Sci., U. Washington)

Peter Hess (Atmos. Chem., NCAR)

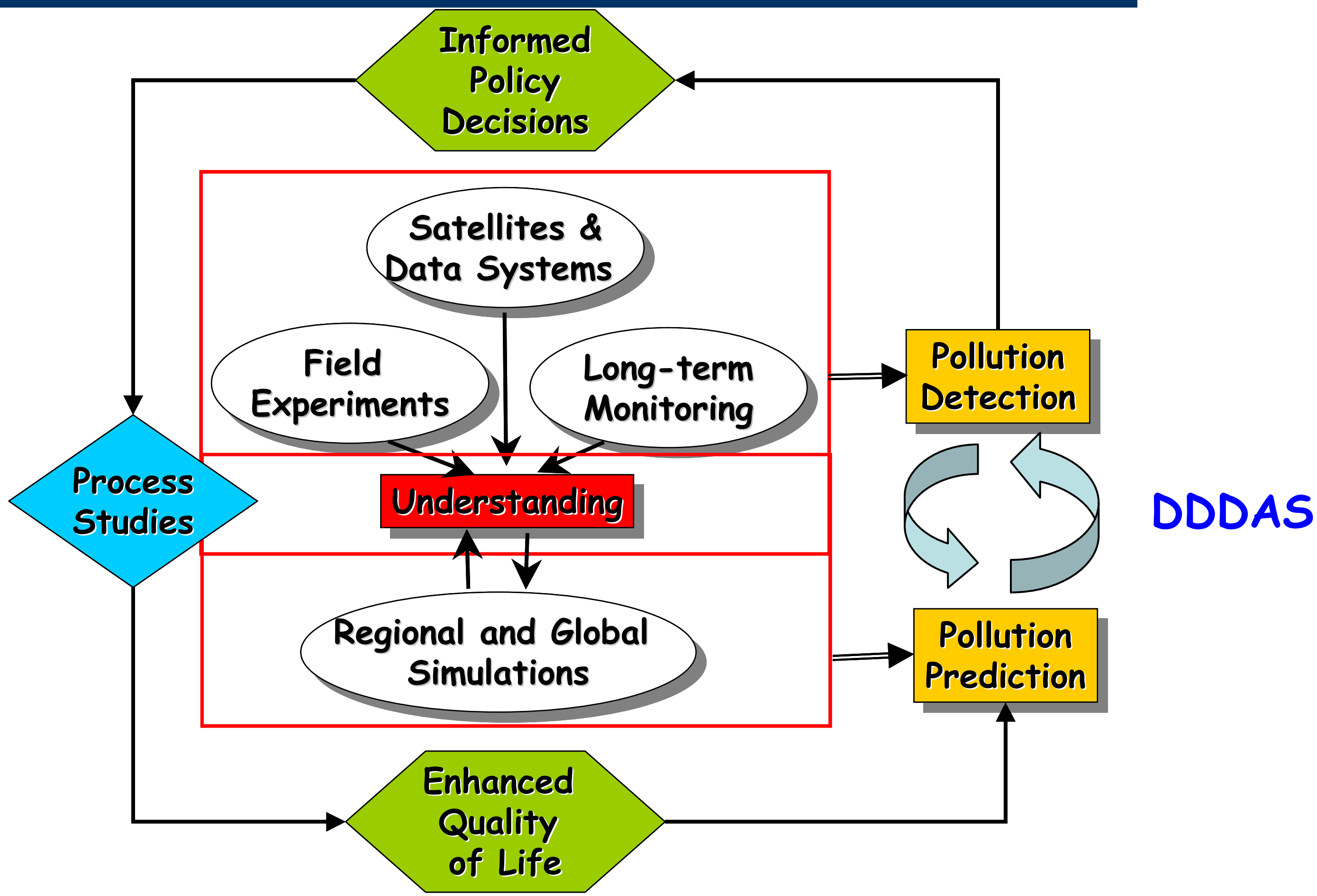
Dacian Daescu (Inst. of Appl. Math., U. Minn.)

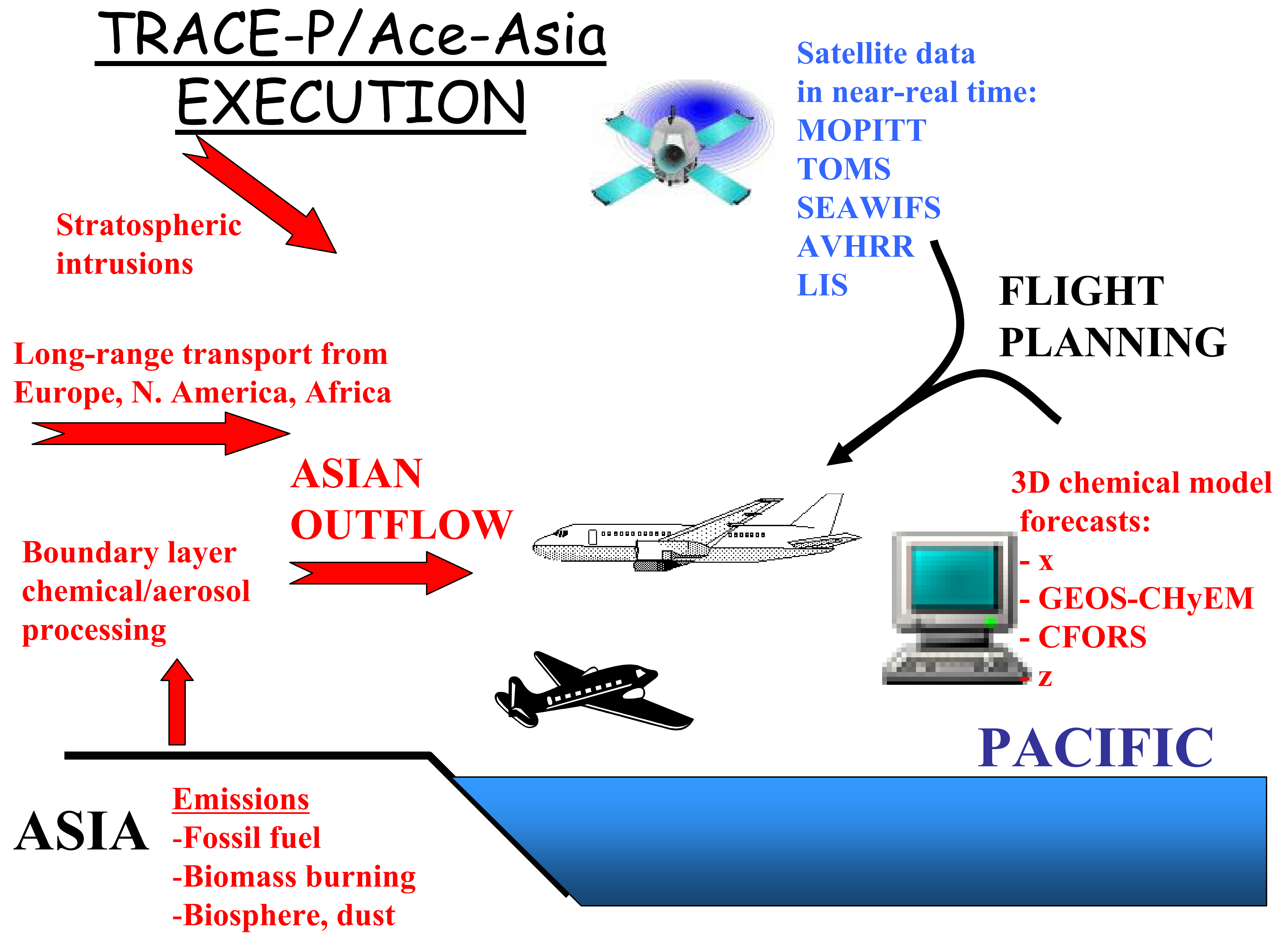
Tianfeng Chai, Center for Global & Regional Environmental Research, UIowa



ICCS2003 Melbourne Australia, June 3, 2003

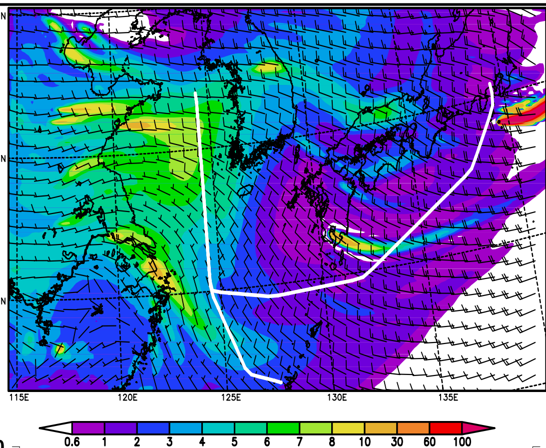
Science Support to Policy



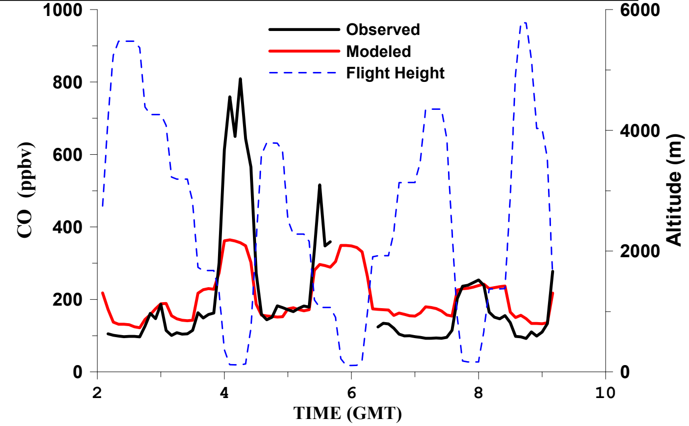

TRACE-P/Ace-Asia
EXECUTION
Stratospheric intrusions
Satellite data in near-real time:
MOPITT
TOMS
SEAWIFS
AVHRR
LIS
FLIGHT PLANNING
Long-range transport from Europe, N. America, Africa
ASIAN OUTFLOW
Boundary layer chemical/aerosol processing
3D chemical model forecasts:
- x
- GEOS-CHyEM
- CFORS
- z
PACIFIC
ASIA
Emissions
-Fossil fuel
-Biomass burning
-Biosphere, dust

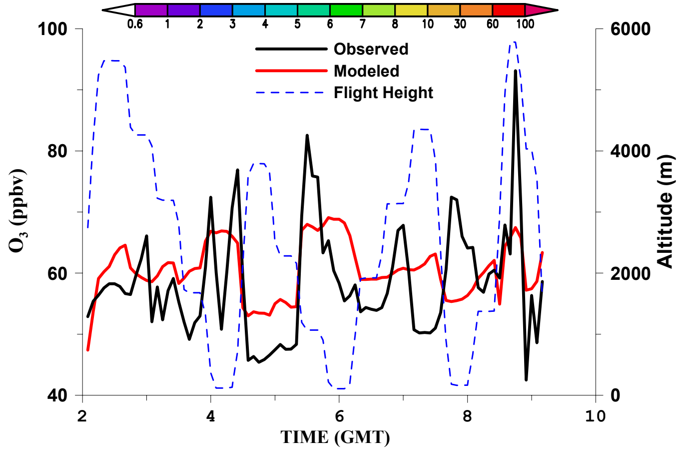
How Well Do Models Capture the Observed Features?

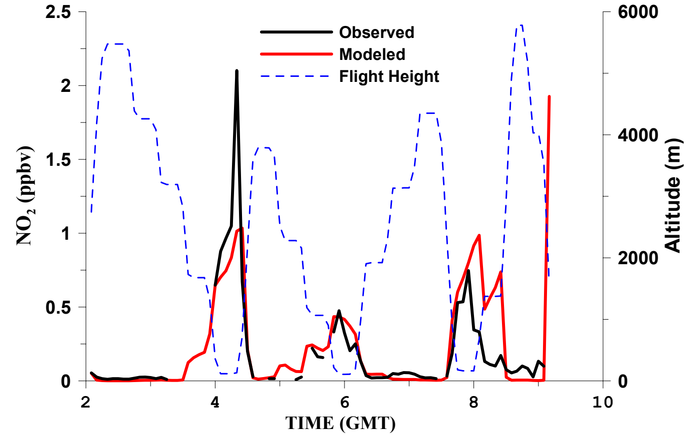
SO2

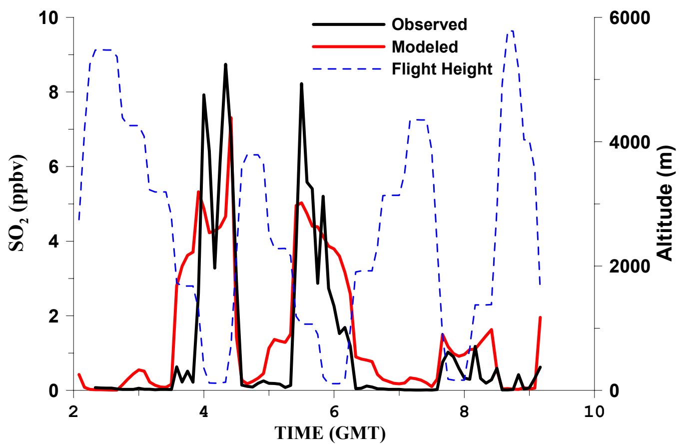


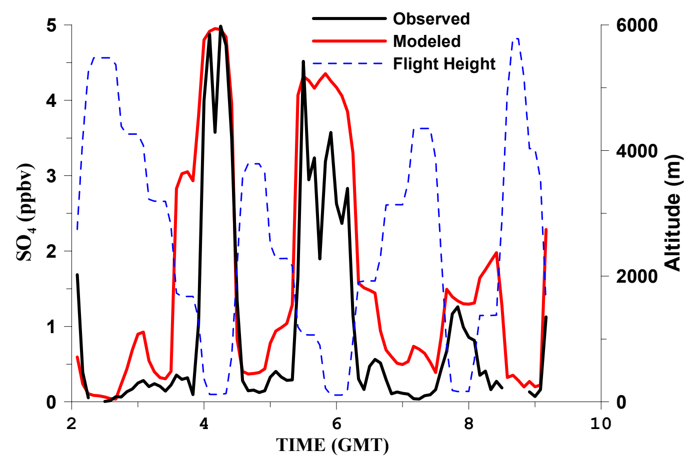
P-3B





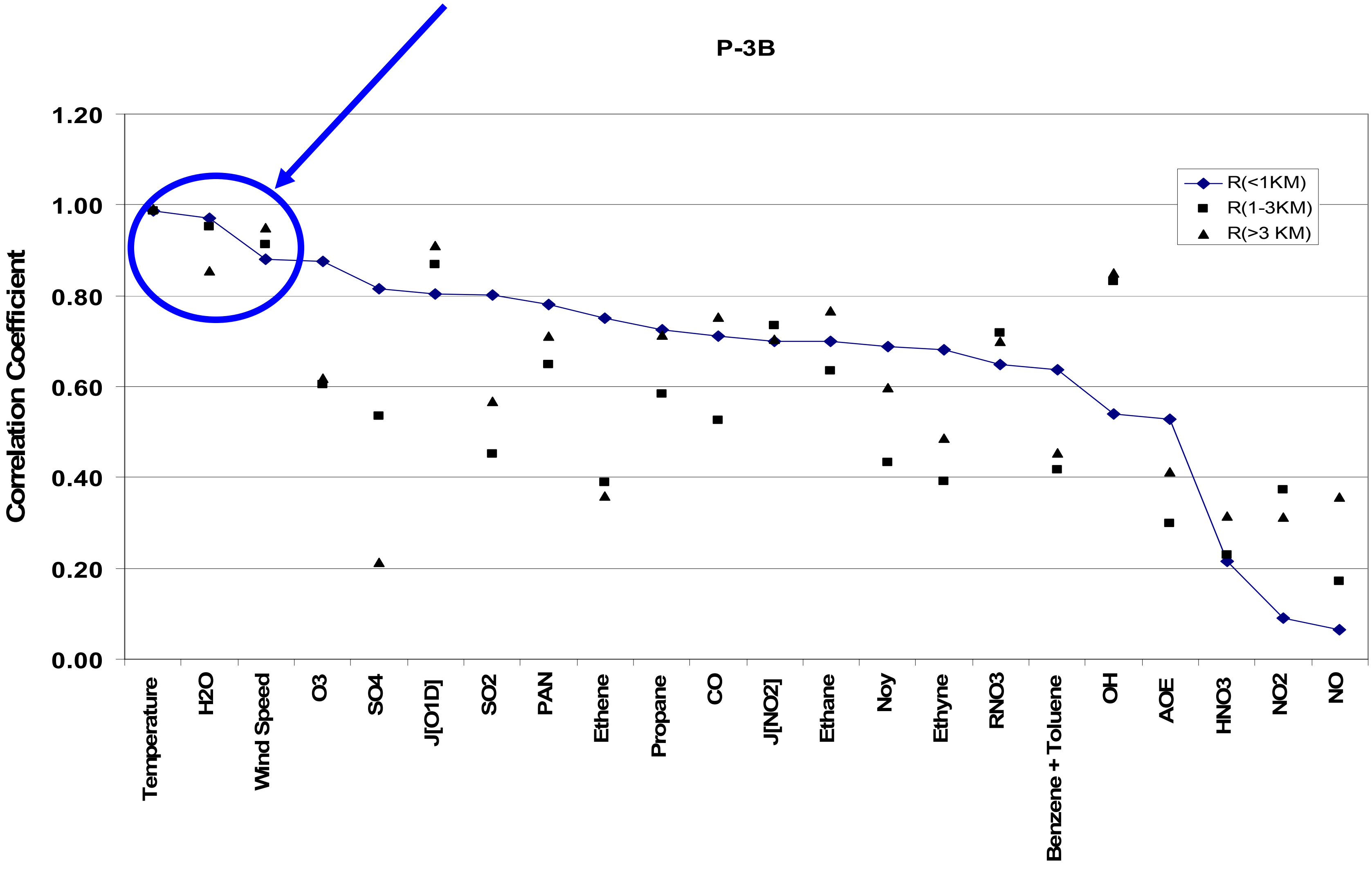






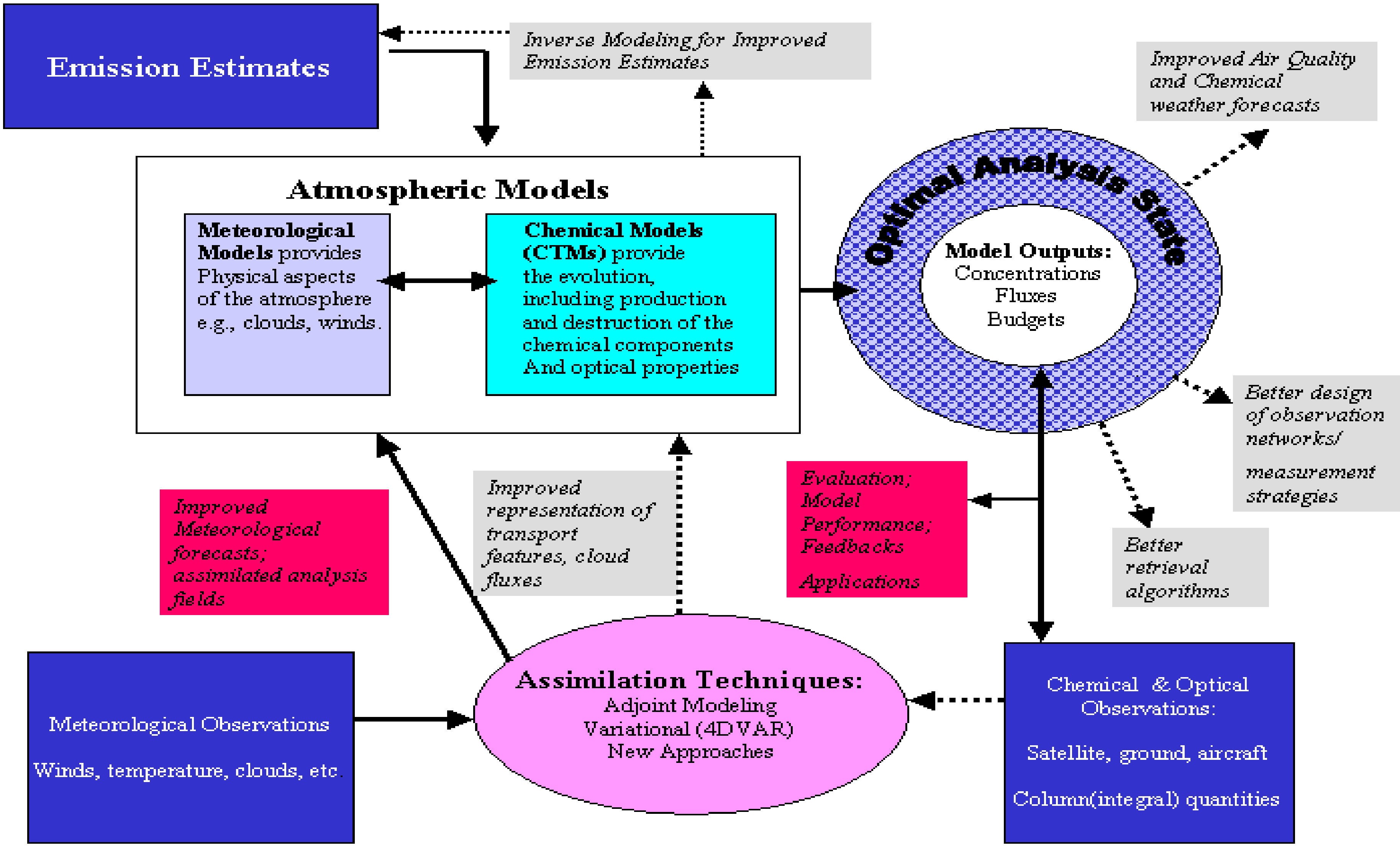
Predictability – as Measured by Correlation Coefficient

Met Parameters are Best



Overview of Research in Data Assimilation for Chemical Models.

Solid lines represent current capabilities. Dotted lines represent new analysis capabilities that arise through the assimilation of chemical data.



Four Dimensional Variational Data Assimilation

- Time-discrete deterministic model

$$\mathbf{y}_k = M_{k-1,k}(\mathbf{y}_{k-1}), \quad k = 1, 2, \ldots, n$$

- Consider the set of observations up to time t_n

$$\mathcal{O}_n = \{\hat{\mathbf{y}}_1, \hat{\mathbf{y}}_2, \ldots, \hat{\mathbf{y}}_n\}, \; \hat{\mathbf{y}}_k = \mathbf{H}_k \mathbf{y}_k + \epsilon_k^o, k = 1, 2, \ldots n$$

- Least-squares approach: minimize the cost functional

$$\mathcal{J} = \frac{1}{2}(\mathbf{y}_0 - \mathbf{y}^b)^T \mathbf{B}^{-1}(\mathbf{y}_0 - \mathbf{y}^b) + \frac{1}{2}\sum_{k=1}^{n}(\hat{\mathbf{y}}_k - \mathbf{H}_k \mathbf{y}_k)^T \mathbf{R}_k^{-1}(\hat{\mathbf{y}}_k - \mathbf{H}_k \mathbf{y}_k)$$

  with respect to $\{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_n\}$ subject to the model constraints.

- Reduced problem

$$\mathbf{y}_k = M_{k-1,k}(M_{k-2,k-1}(\ldots\ldots M_{0,1}(\mathbf{y}_0))) \Rightarrow \min_{\mathbf{y}_0} \mathcal{J}(\mathbf{y}_0)$$

- Large-scale problems $\longrightarrow$ critical need for accurate and efficient gradient evaluation $\nabla_{\mathbf{y}_0} \mathcal{J}(\mathbf{y}_0)$

Advantages of Adjoint Sensitivities

(+) Very efficient for many in $\rightarrow$ few out

(+) Computational cost of the same order as a forward model integration

(+) Provides backward sensitivity at no additional cost

(-) Inefficient for few in $\rightarrow$ many out

(-) High memory storage requirements $\Rightarrow$ trade-off CPU/memory

(-) Implementation may not be trivial

Adjoint Sensitivity

Want: Stiff ODE (e.g. chemistry); scalar functional

$$y' = f(t,y)\,,\ t^0 \le t \le t^F\,;\ \mathcal{F}(y) \quad \Rightarrow \quad \nabla_{y^0}\mathcal{F}(y(t^F))$$

Continous: Solve numerically

$$\begin{aligned}
\lambda' &= -J^T(t,y)\lambda\,, \quad \lambda(t^{\mathrm{F}}) = \nabla_y \mathcal{F}\big|_{y(t^{\mathrm{F}})} \\
\nabla_{y^0}\mathcal{F} &= \lambda(t^0)\,. \quad \text{(store fwd traj. y(t))}
\end{aligned}$$

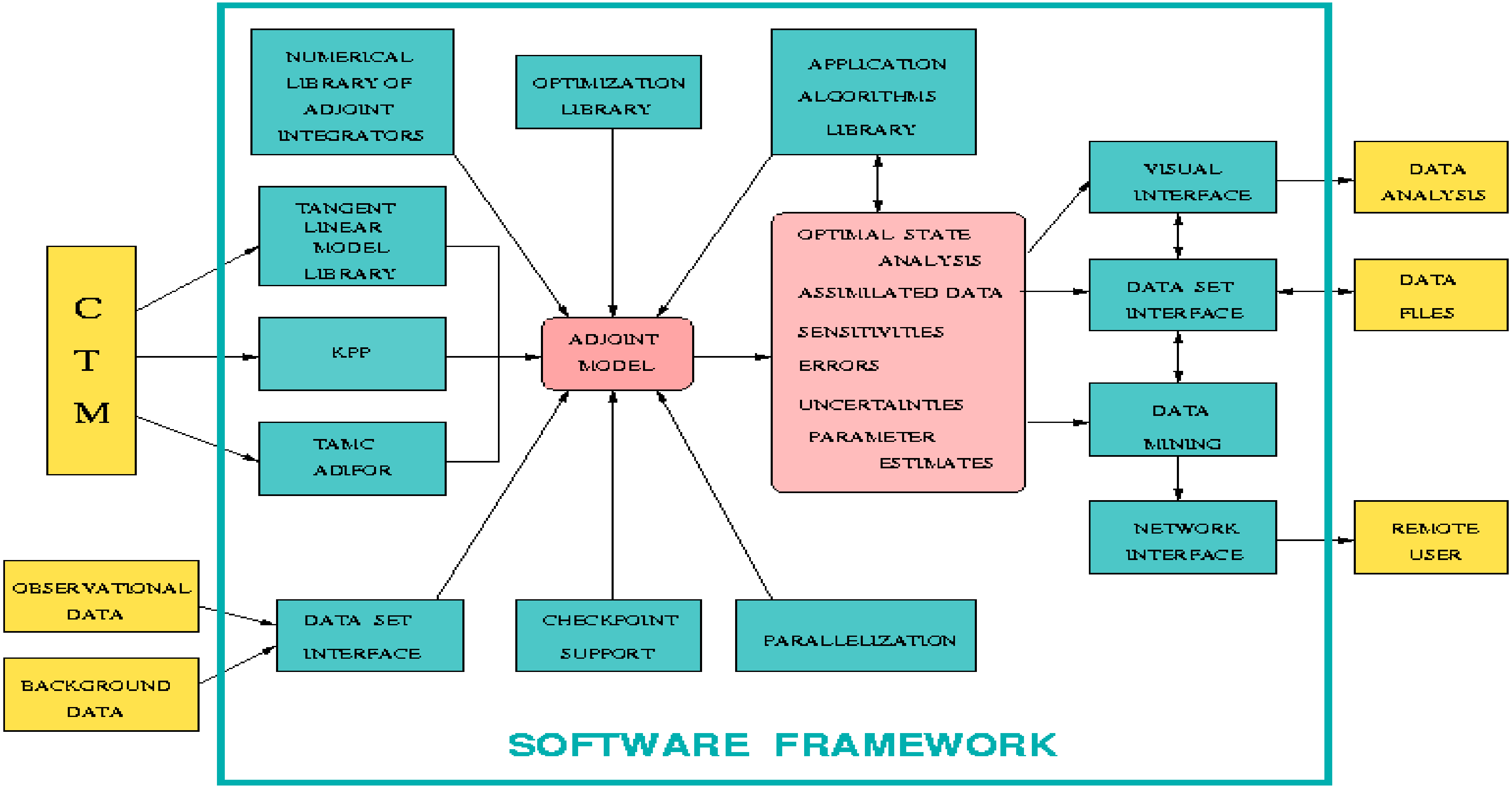
Discrete: Derivative of numerical method ($\sim$ AD)

$$\begin{aligned}
y^{i+1} &= F^i(y^i)\,, \quad i = 0,\ldots,N-1\,, \\
\lambda^i &= F^i_y(y^i,p^i)^T\lambda^{i+1}\,, \quad \lambda^N = \nabla_y\mathcal{F}\big|_{y^N} \\
\nabla_{y^0}\mathcal{F}(y^N) &= \lambda^0\,. \quad \text{(store fwd traj. y}^{\mathrm{i}}\text{)}
\end{aligned}$$

We are Developing General Software Tools to Facilitate the Close Integration of Measurements and Models

The framework will provide tools for: 1) construction of the adjoint model; 2) handling large datasets; 3) checkpointing support; 4) optimization; 5) analysis of results; 6) remote access to data and computational resources.

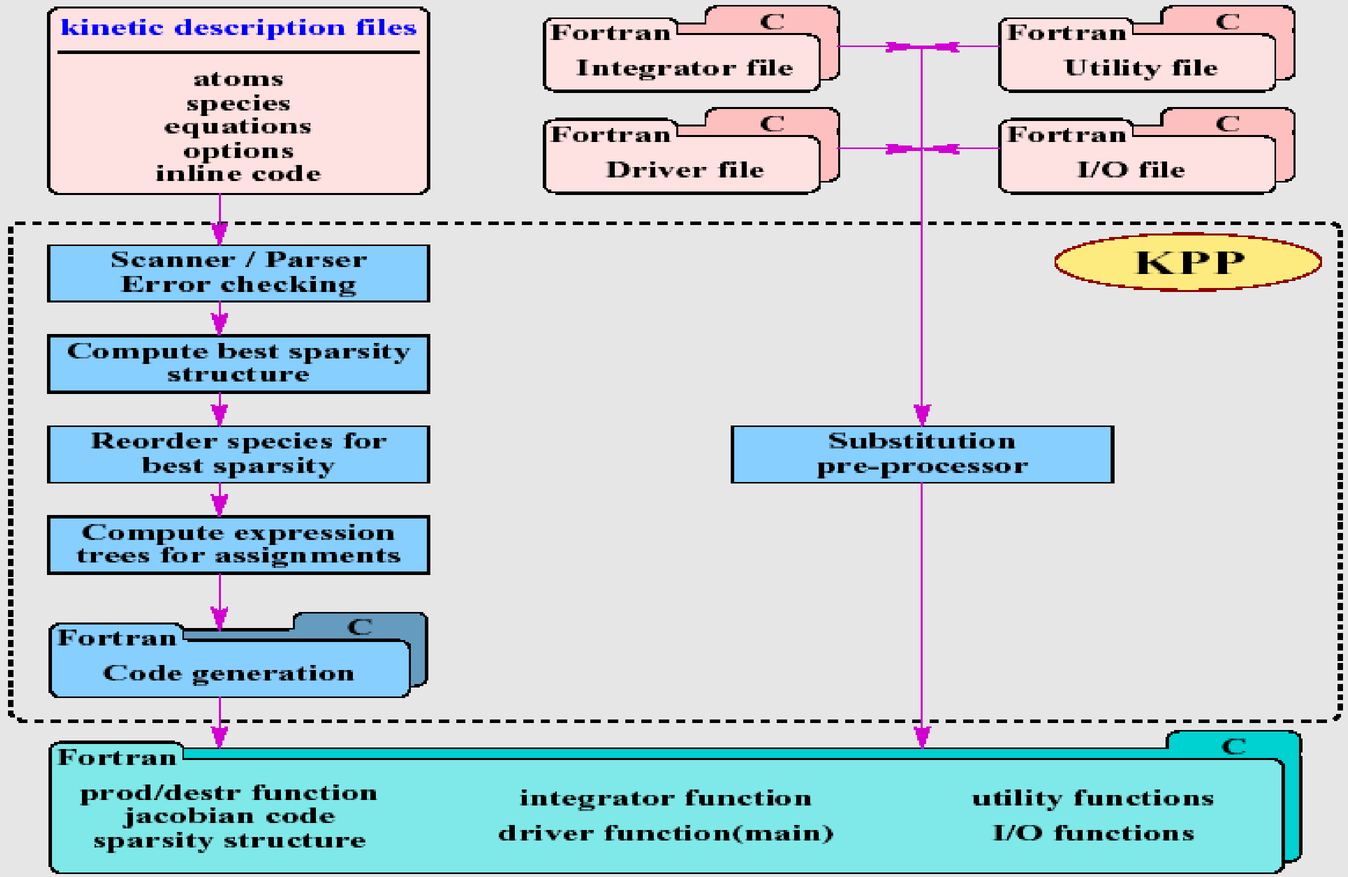
(www.cs.mtu.edu/~asandu/Software/Kpp)



The Kinetic PreProcessor (KPP)

(Damian, Sandu, et. al. 1996); (Sandu et. al. 2001, 2002)

http://www.cs.mtu.edu/∼asandu/Software/Kpp



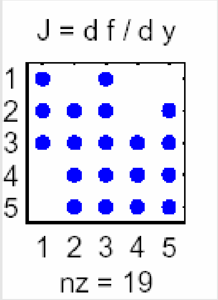
KPP-1.2 Options for Adjoints

#JACOBIAN [ON | OFF | SPARSE]

JacVar(), JacVar_SP()

JacVar_SP_Vec(), JacVarTR_SP_Vec()

KppDecomp(), KppSolve(), KppSolveTR()

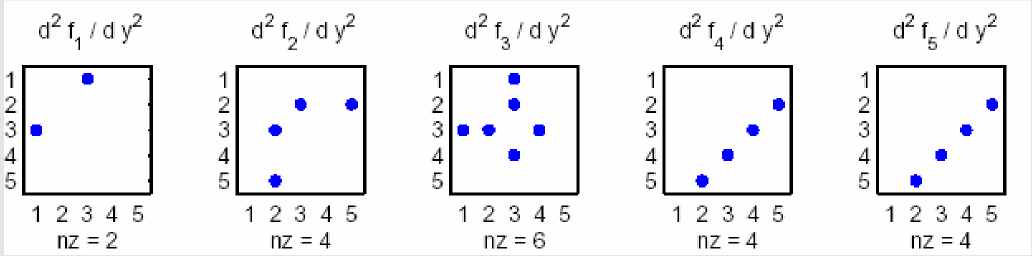


#HESSIAN [ON | OFF]

HessVar()

HessVar_Vec()

HessVarTR_Vec()



Parallelization

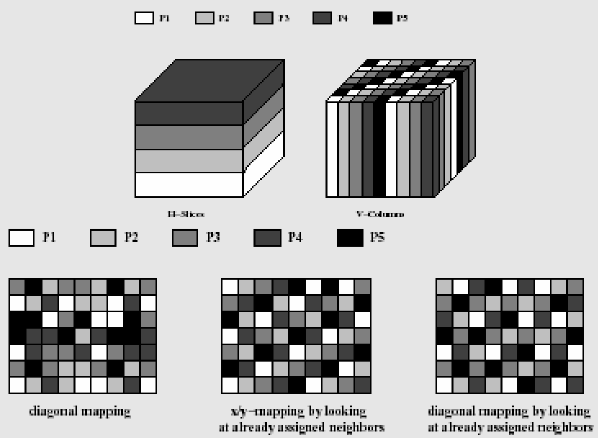
(Miehe and Sandu, 2001). http://www.cs.mtu.edu/∼asandu/Parallel

PAQMSG: Communication Library (F90/MPI)

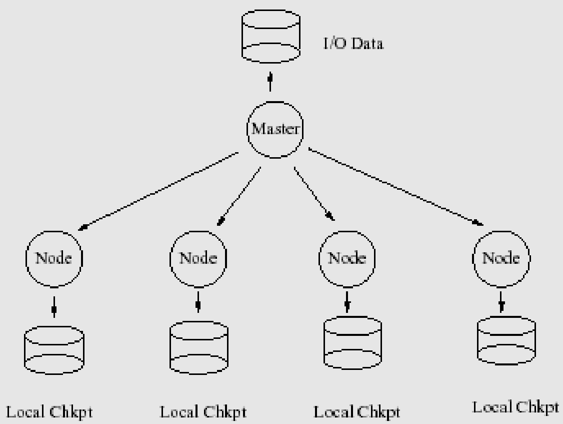
– Master/worker, domain decomp., static partitioning;

– 4D/3D/2D arrays; Alloc (local/global); Distrib; Gather; Shuffle.

Example Dom. Dec. & Mapping



Parallel Checkpointing



SAPRC Data Assimilation

(Daescu, Sandu, Carmichael, 2002)

- SAPRC (Carter, 1999): 79 spec., 211 react. Emissions considered.
- Assim. window = first 24 hrs (of 120 hrs run). Hourly observations.

$$\mathcal{F}(\ln c^0, \ln E) = \frac{1}{2} \sum_{t=1}^{24\ h} \sum_{i \in Obs} \left[\ln c_i^k - \ln(c^{\mathrm{obs}})_i^k \right]^2$$

- Twin experiments (Em.+50%). LBFGS to decrease $\mathcal{F}$ 100 times.
- $T_{\mathrm{Adj.Grad}}$ (KPP/Rodas-3) $\{$ Griewank: $T_{\mathrm{dadj}} \le 5 \times T_{\mathrm{fwd}}$ $\}$

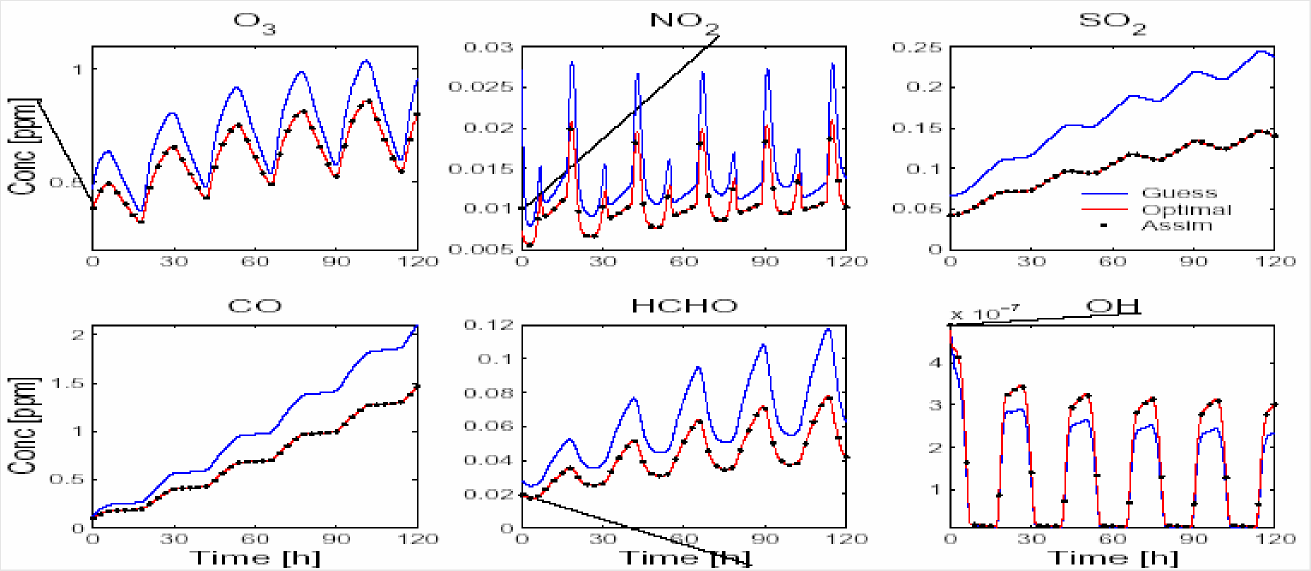
$T_{\mathrm{cadj}}/T_{\mathrm{fwd}}$	$T_{\mathrm{cadj}}^{\nabla\mathcal{F}}/T_{\mathrm{fwd}}$	$T_{\mathrm{dadj}}/T_{\mathrm{fwd}}$	$T_{\mathrm{dadj}}^{\nabla\mathcal{F}}/T_{\mathrm{fwd}}$
1.2	3.31	2.3	4.43

SAPRC Assimilation, cont'd

Assimilation of Emission Rates

Species	NO	NO_2	SO_2	HCHO	ALK1
$\Delta E^0/E$	50%	50%	50%	50%	50%
$\Delta E^A/E$	0.13%	0.07%	0.26%	1.05%	0.94%

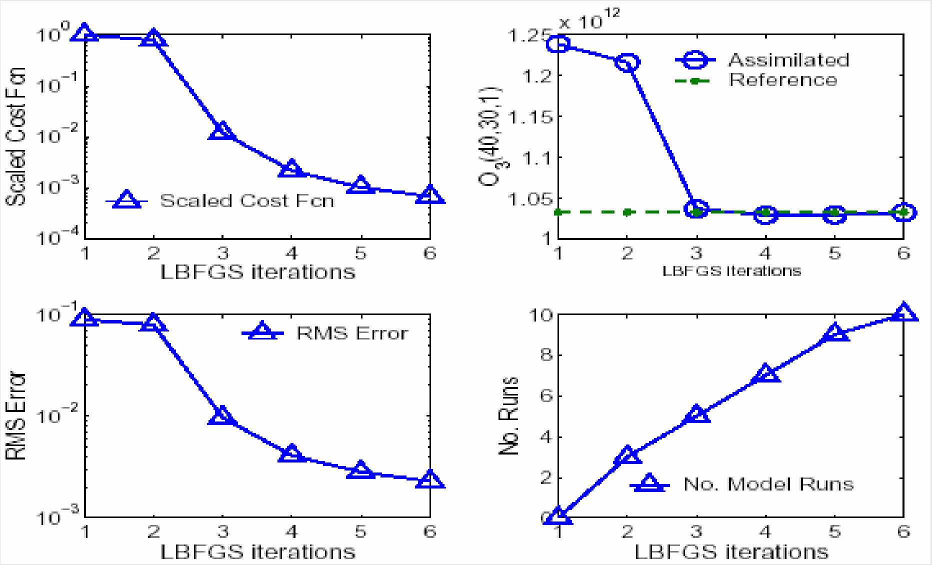
Assimilation of Initial Values



3D Assimilation Results

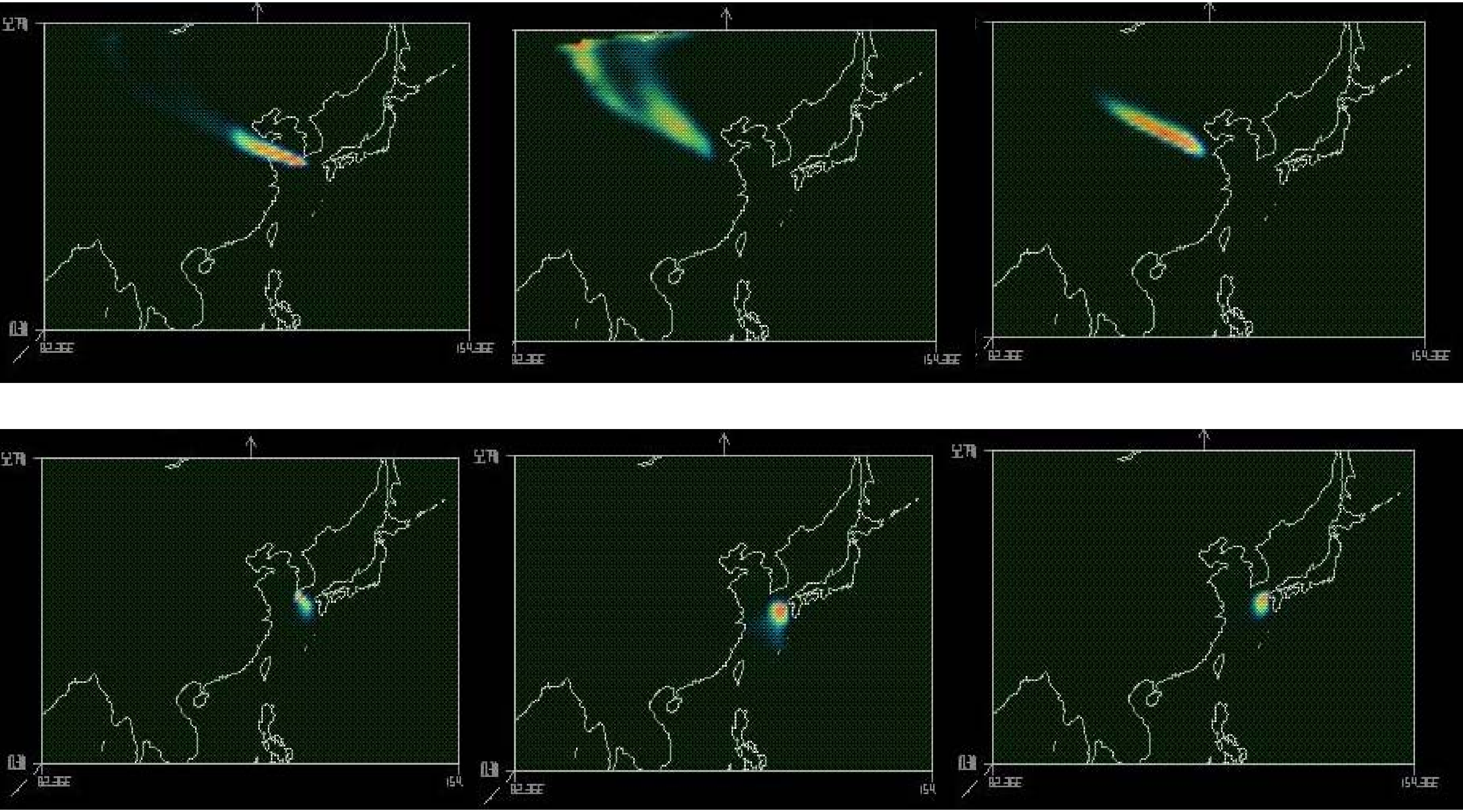
- STEM III; ○ SAPRC 99 (Ros-2); ○ 3rd upwind FD, Crank-Nich;
- TraceP: [0,6] GMT March 01, 2001;
- Twin Experiments (O_3 +20%).

$$\mathcal{F} = \frac{1}{2} \sum_{\text{gridpoints}} [O_3(t^F) - O_3^{\text{obs}}(t^F)]^2$$



Advantage Of Adjoints Is That We Get Sensitivities; e.g., Influence Functions (top view)

(top:Mar.1-3, bottom:Mar.22-24; from left to right: O_3, NO_2, HCHO)

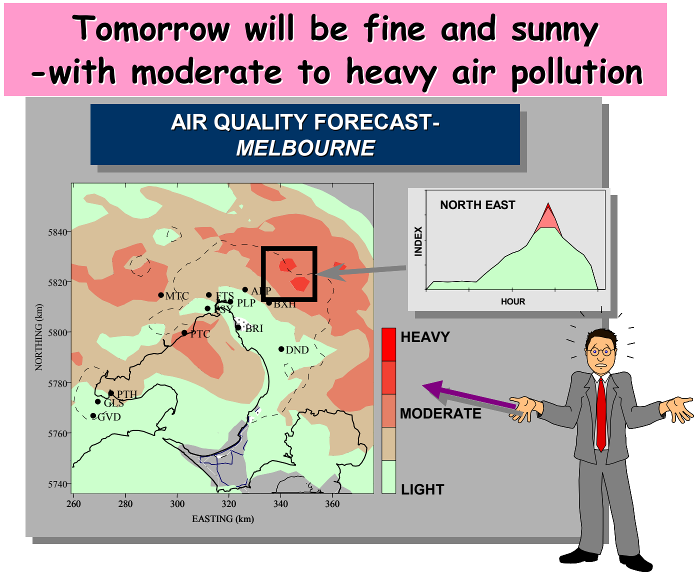


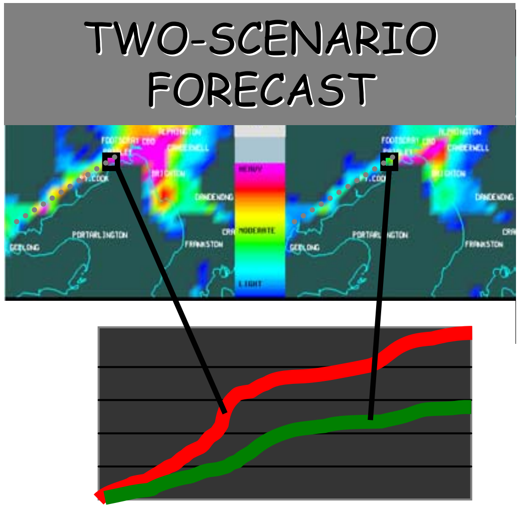
Summary

- Nonlinear, stiff chemistry important part of chemical transport models;
- One-step methods (RK, Ros) seem preferable w/ adjoint sensitivity;
- KPP: efficient code for simulation and direct/adjoint sensitivities;
- Preliminary 3D assimilation results with parallel adjoint STEM-III;
- Developing/exploring new optimization techniques (using 2nd order sensitivities), adaptive grids…..

Target Applications

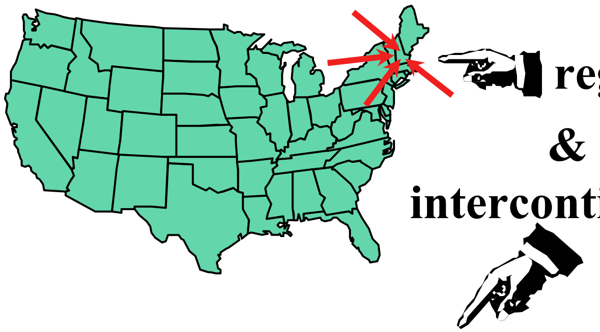
- ✓ Air quality forecasting in urban environments;
- ✓ Integration of measurements and models to produce a consistent/optimal analysis data set for atmospheric chemistry field experiments (e.g., Ace-Asia);
- ✓ Inverse analysis to produce a better estimate of emissions;
- ✓ Design of observation strategies to improve chemical forecasting capabilities.





Application: The Design of Better Observation Strategies to Improve Chemical Forecasting Capabilities.

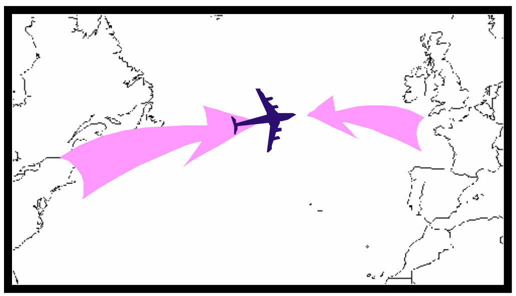
Data Assimilation Will help us Better Determine Where and When to Fly and How to More Effectively Deploy our Resources (People, Platforms, $$s)



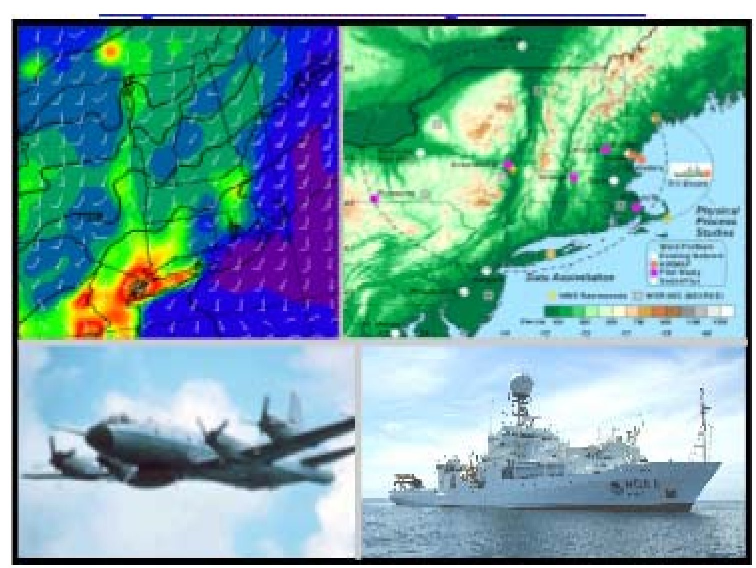
regional

&

intercontinental



We Plan to Test These Tools Including Adaptive Measurements in the Summer of 2004



- *Ability of forecast models to represent the individual processes controlling air pollution formation and transport.*