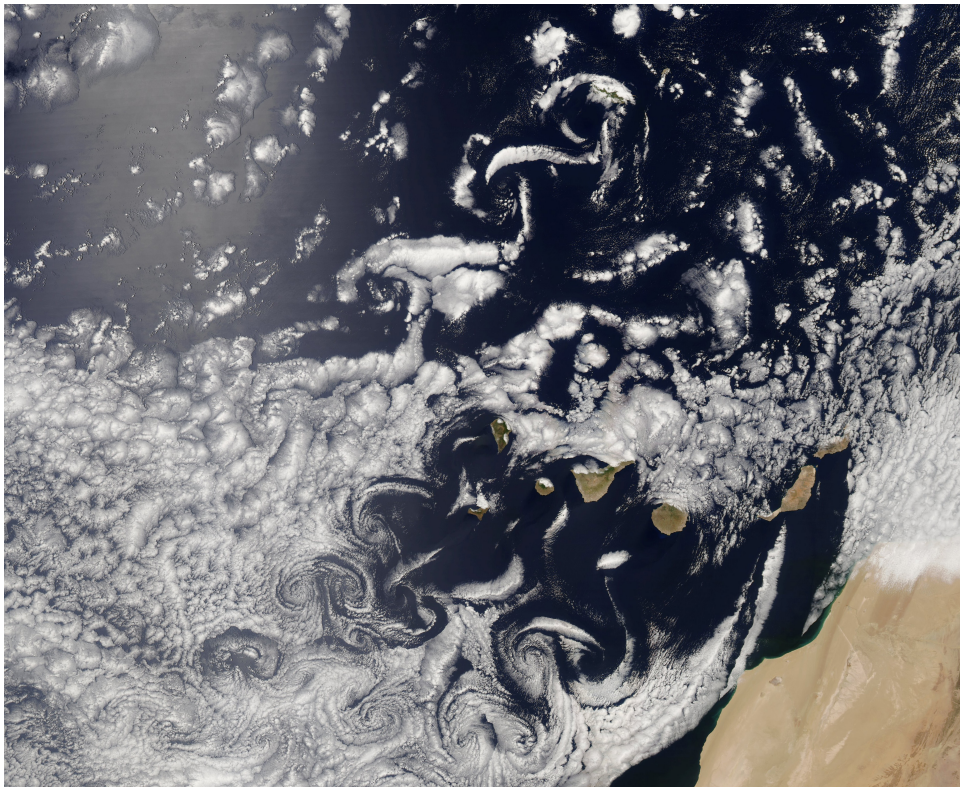
Validation and verification in global atmospheric chemistry models

Mauricio Santillana (msantill@fas.harvard.edu)
Harvard University
School of Engineering and Applied Sciences



Collaborators: Daniel J. Jacob, Michael P. Brenner, Lin Zhang.

Workshop on Verification and Validation in Computational Science, Univ. of Notre Dame, Oct 17th 2011

Terminology…

Error Quantification (Uncertainty)

$$\frac{\partial n_i}{\partial t} = -\nabla \cdot (\boldsymbol{u}(t) n_i) + R(t, \boldsymbol{n}) + C_{sub-grid}(n_i)$$

$$\begin{aligned} NO + O_3 &\longrightarrow NO_2 + O_2 & Rate_1 &= k_1[NO][O_3] \\ O + O_2 + M &\longrightarrow O_3 + M & Rate_2 &= k_2[O][O_2][M] \\ NO_2 + h\nu &\longrightarrow NO + O & Rate_3 &= J[O][O_2][M] \\ NO_2 + O &\longrightarrow NO + O_2 & Rate_{=} &k_3[NO_2][O] \end{aligned}$$

Mathematical Model

Validation

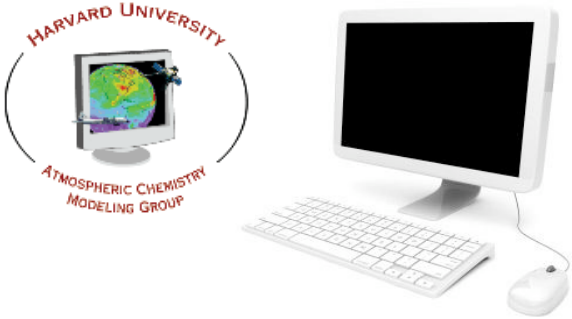
(frequently impossible)

Verification

(Numerical analysis)

Real world (Atmosphere)

In practice: Validation



Computational model
(GEOS-Chem)

Scientific papers

“Validation”

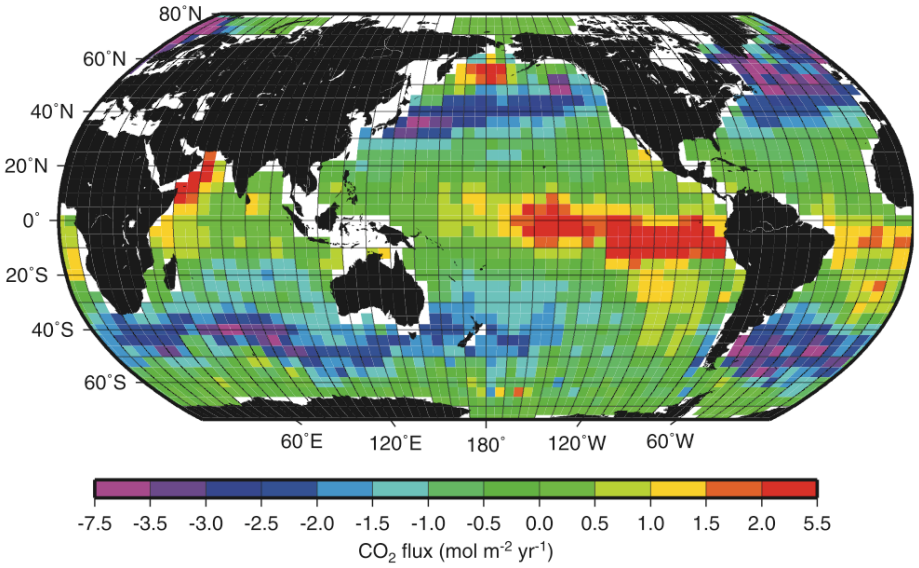
(Computational model vs measurements)

Verification

(Numerical analysis, applied math)

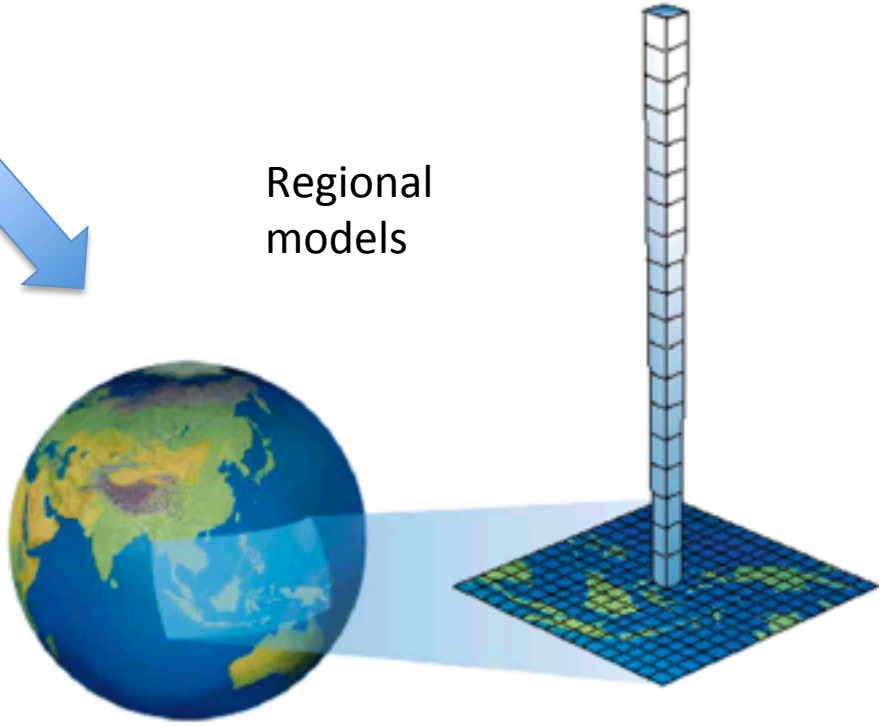
Increase of computational resources = refine mesh!

Simulation results heavily depend on temporal and spatial resolutions!



From Global (IPCC models) climate change scenarios

Regional models

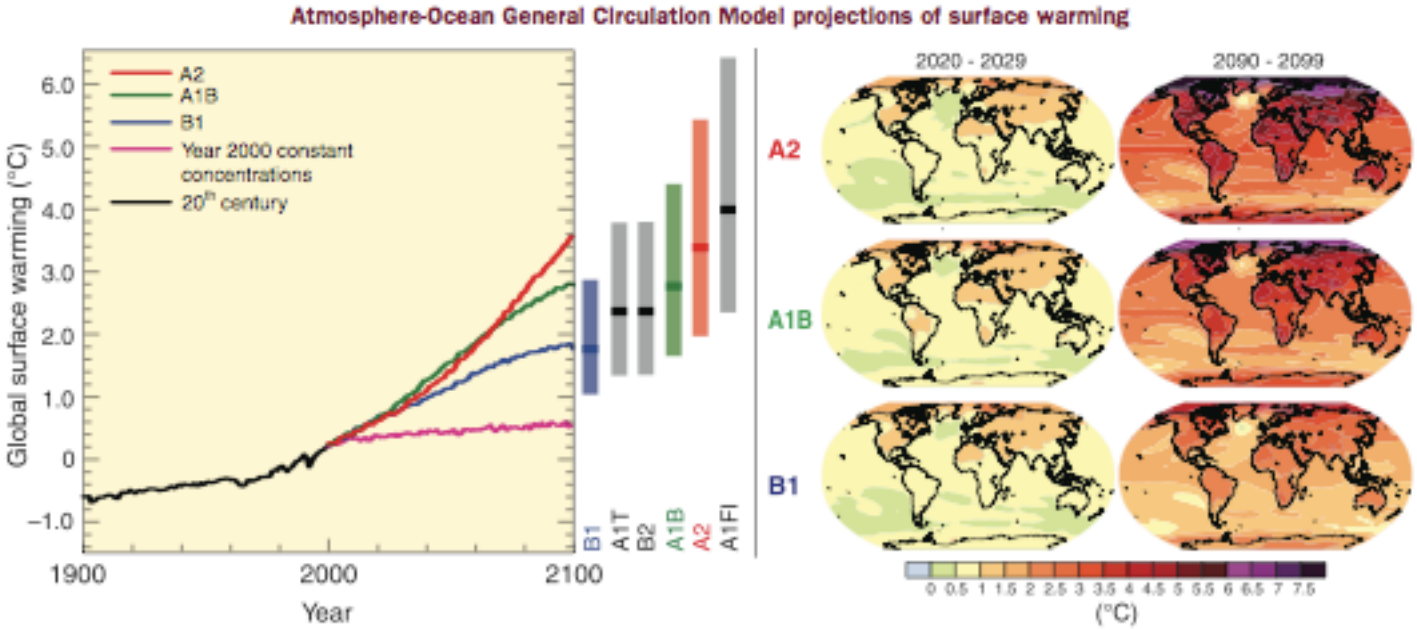


http://www.cru.uea.ac.uk/

Motivation

Why study the chemical composition of our atmosphere?

Global **climate** conditions are **changing**: Do human emissions play a key role?



Source: IPCC, 2007

Warming of the **climate** system is **unequivocal**, as is now evident from observations of increases in global average air and ocean temperatures, widespread melting of snow and ice, and rising global average sea level.

Why study the chemical composition of our atmosphere?

Anthropogenic (pollution) and natural emissions

Need to understand the fate of emitted chemical species to quantify their role in ecosystem changes and their effects on human health.

Greenhouse gasses (CO_2 and CH_4)

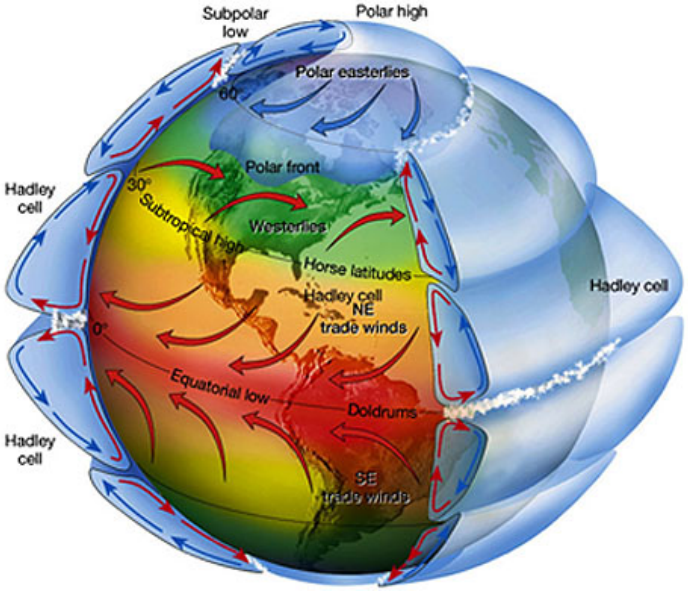
Why study the chemical composition of our atmosphere?

Understanding how gases are transported in the atmosphere helps us:

- Design appropriate **environmental policy** strategies (Kyoto Protocol, Copenhagen 2009, Cancun 2010),
- Improve air quality (London, Houston, Mexico City, Beijing),
- Monitor ***who*** **produces** ***what*** (source attribution, inverse modeling)

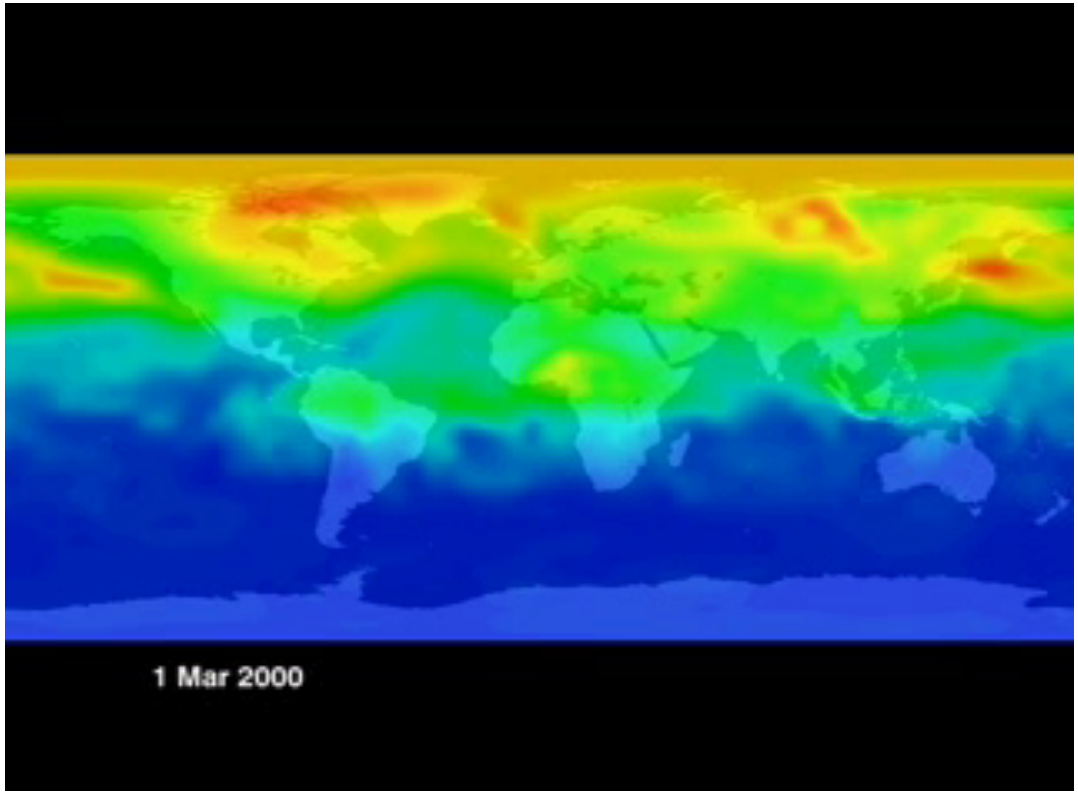
What we attempt to model is…

- Pollution is emitted,
- Transported by winds,
- Interacts with the environment.



Measurements of Carbon Monoxide (CO) in 2000 at 15,000 ft.

Source: http://www.gsfc.nasa.gov/gsfc/earth/terra/co.htm



CO levels: Red = 450 ppb, Blue =50ppb

How do we do it?

The Mathematical model

$$\frac{\partial n_i}{\partial t} = \underbrace{-\nabla \cdot (\boldsymbol{u}(t) n_i)}_{\text{resolved transport}} + \underbrace{R(t, \boldsymbol{n})}_{\text{chemistry}} + \underbrace{C_{sub-grid}(n_i)}_{\text{sub-grid processes}}$$

$$i = 1, ..., 110$$

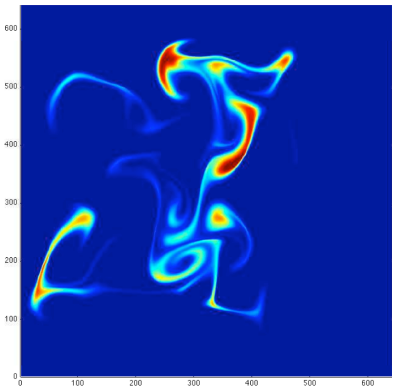
Sub-grid processes are typically **parametrizations**:

- Un-resolved transport (eddies in Planetary Boundary Layer),
- Convective transport (vertical motion),
- Micro-physical processes: dry scavenging, wet scavenging;
- Sub-grid emissions (in-grid plume dispersion models, from points to gridbox)

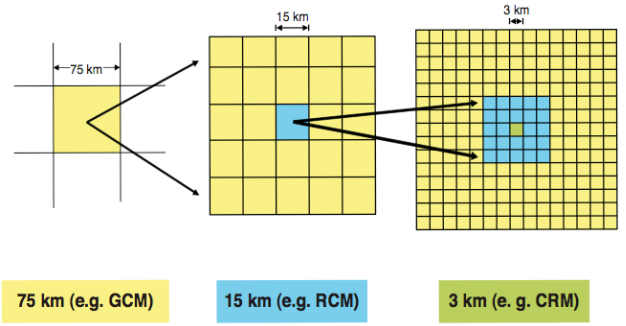
Validation and verification in global atmospheric chemistry models

(concrete examples)

1. Transport schemes are highly diffusive. Impact on **inverse modeling** (Verification and Validation)



2. Simulations exhibit heavy dependence on temporal and spatial resolutions. (Verification)

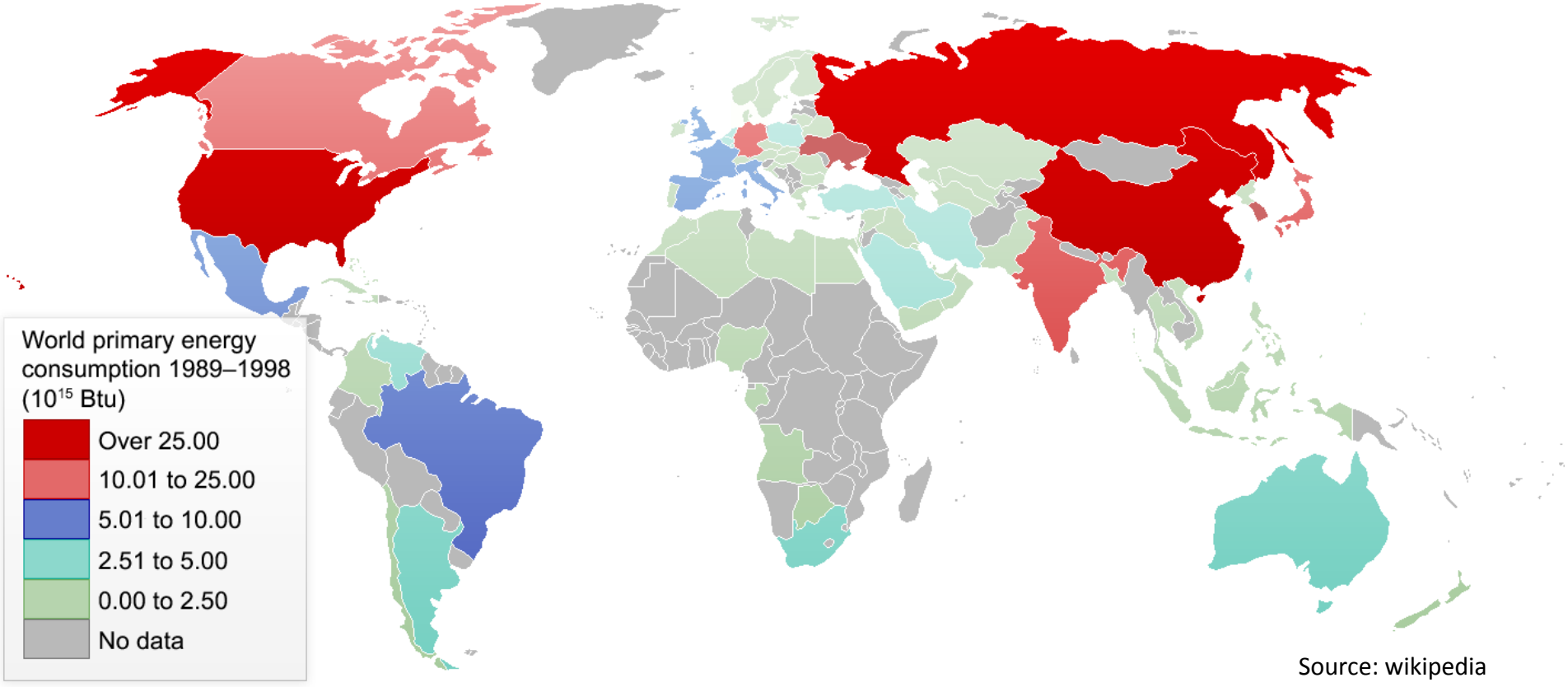


Transport schemes are highly diffusive. Impact on **inverse modeling**
(Validation and Verification)

with: Marco Iglesias (MIT) and Michael Brenner (Harvard).

What kind of questions do we need to address?
Inverse modeling

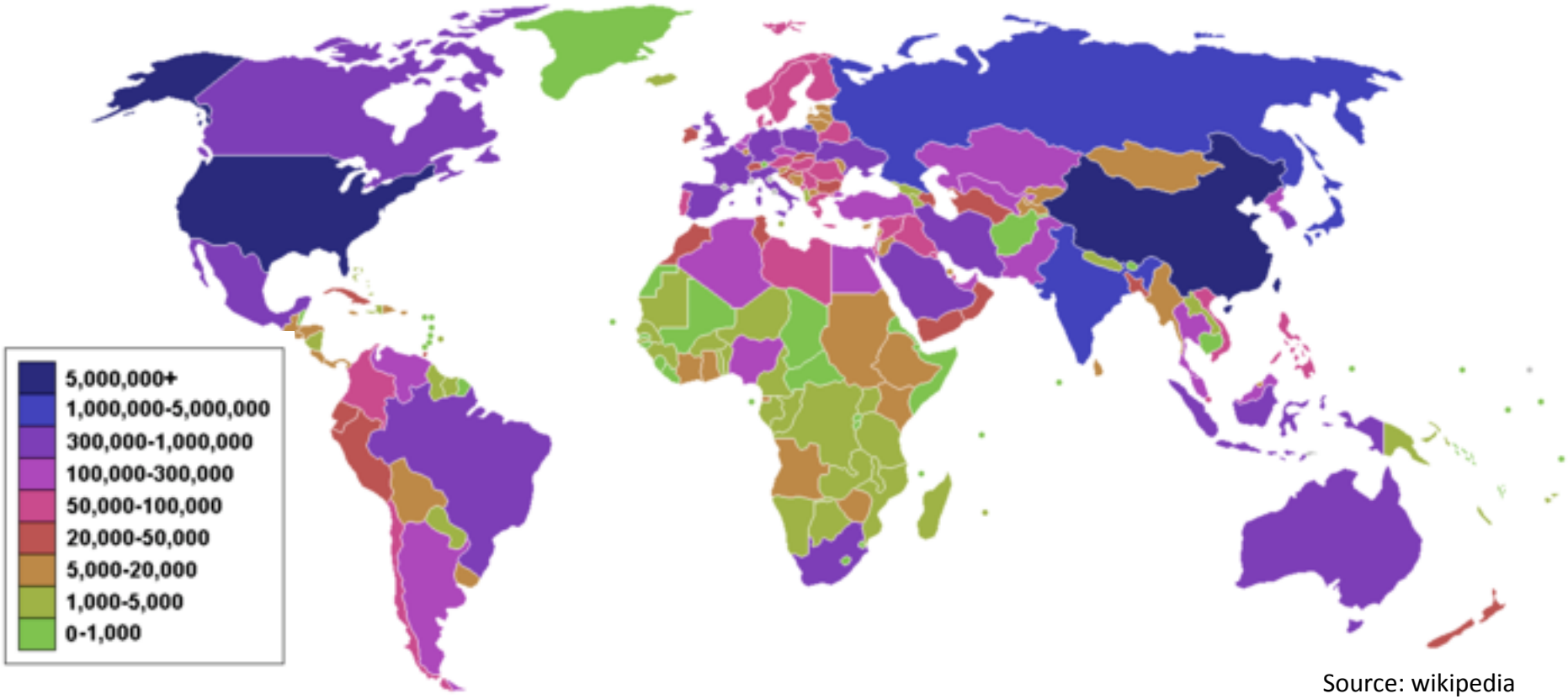
World energy consumption



Source: wikipedia

What kind of questions do we need to address?
Inverse modeling

World CO2 emissions



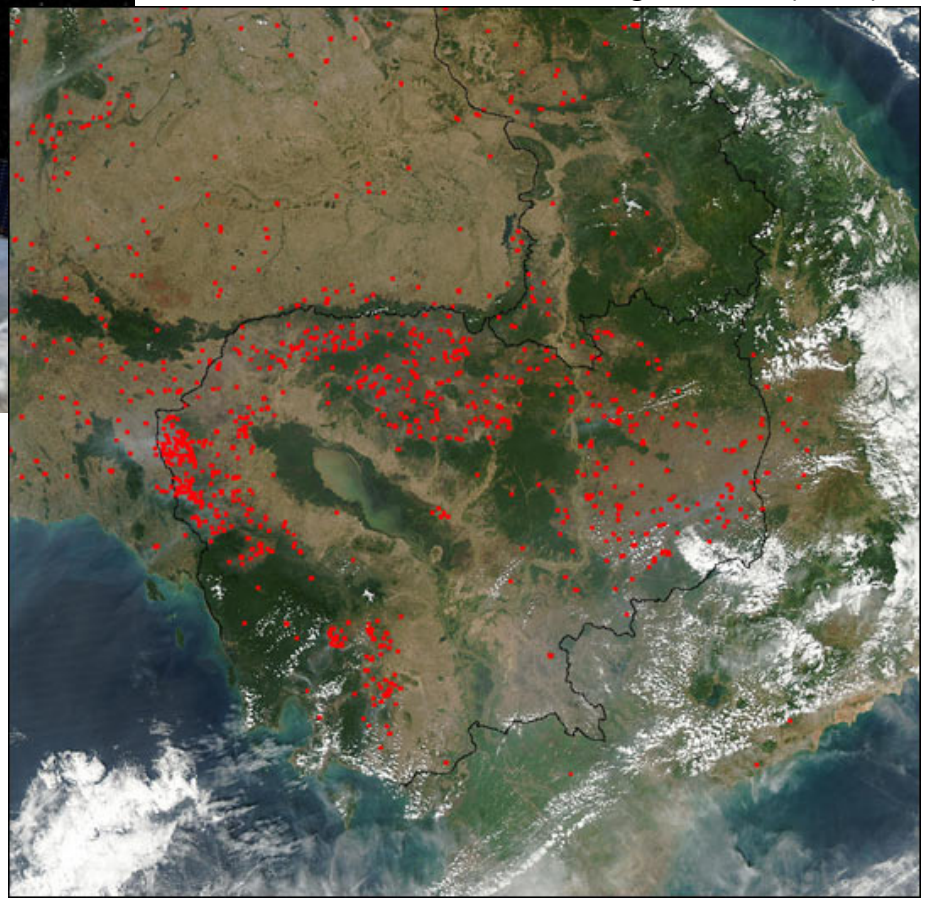
Source: wikipedia

Need reliable techniques to verify where pollution is being emitted.

What kind of questions do we need to address?
Inverse modeling

CO2 satellite

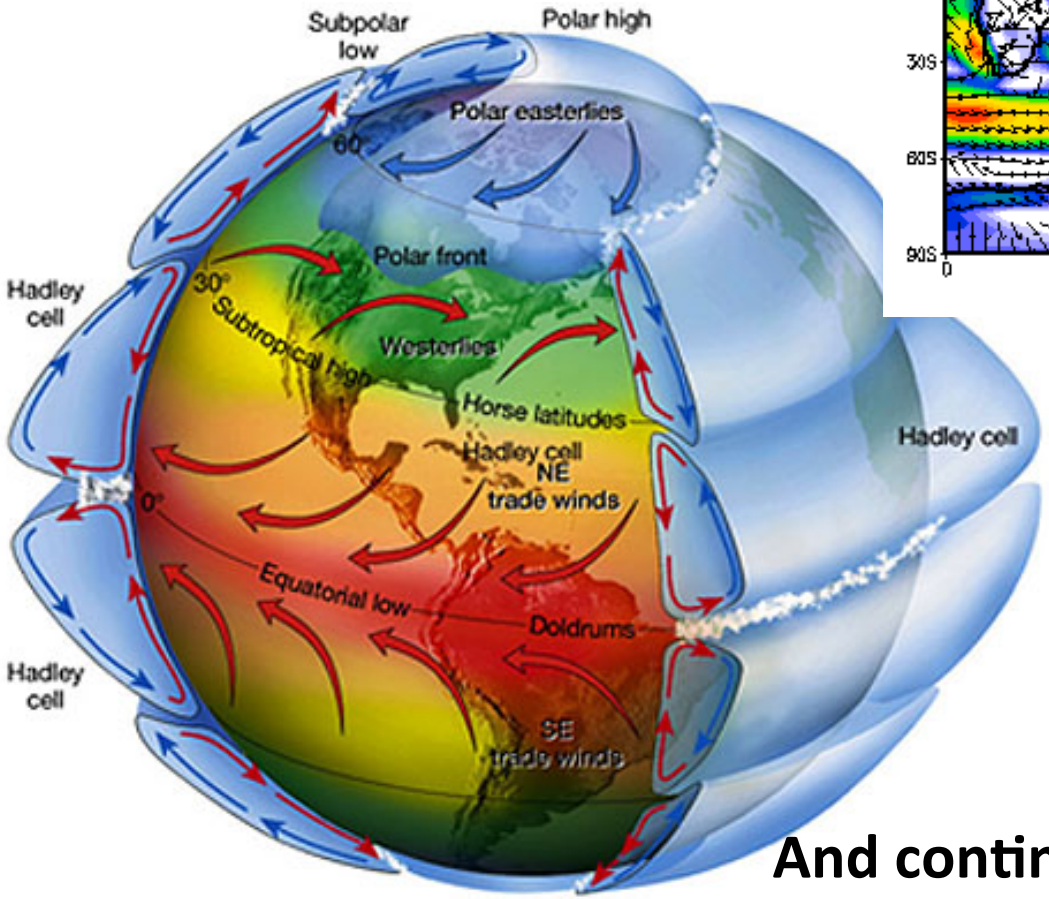
CO2 Biomass burning Indochina (NASA)

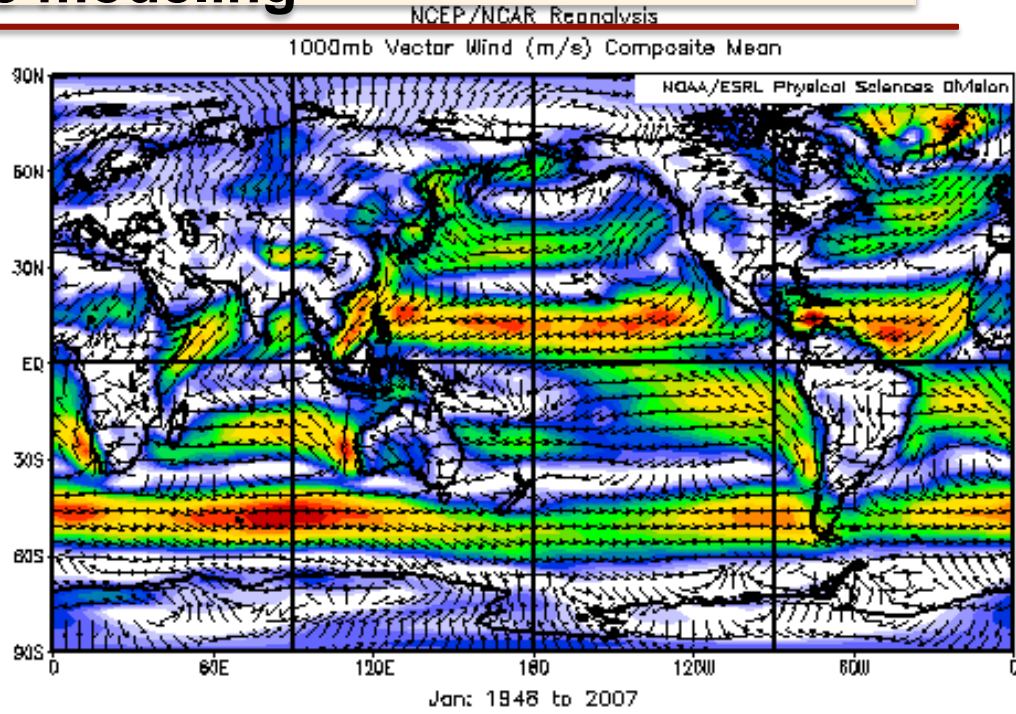


Given a set of measurements (satellites, met stations)

What kind of questions do we need to address?

Inverse modeling





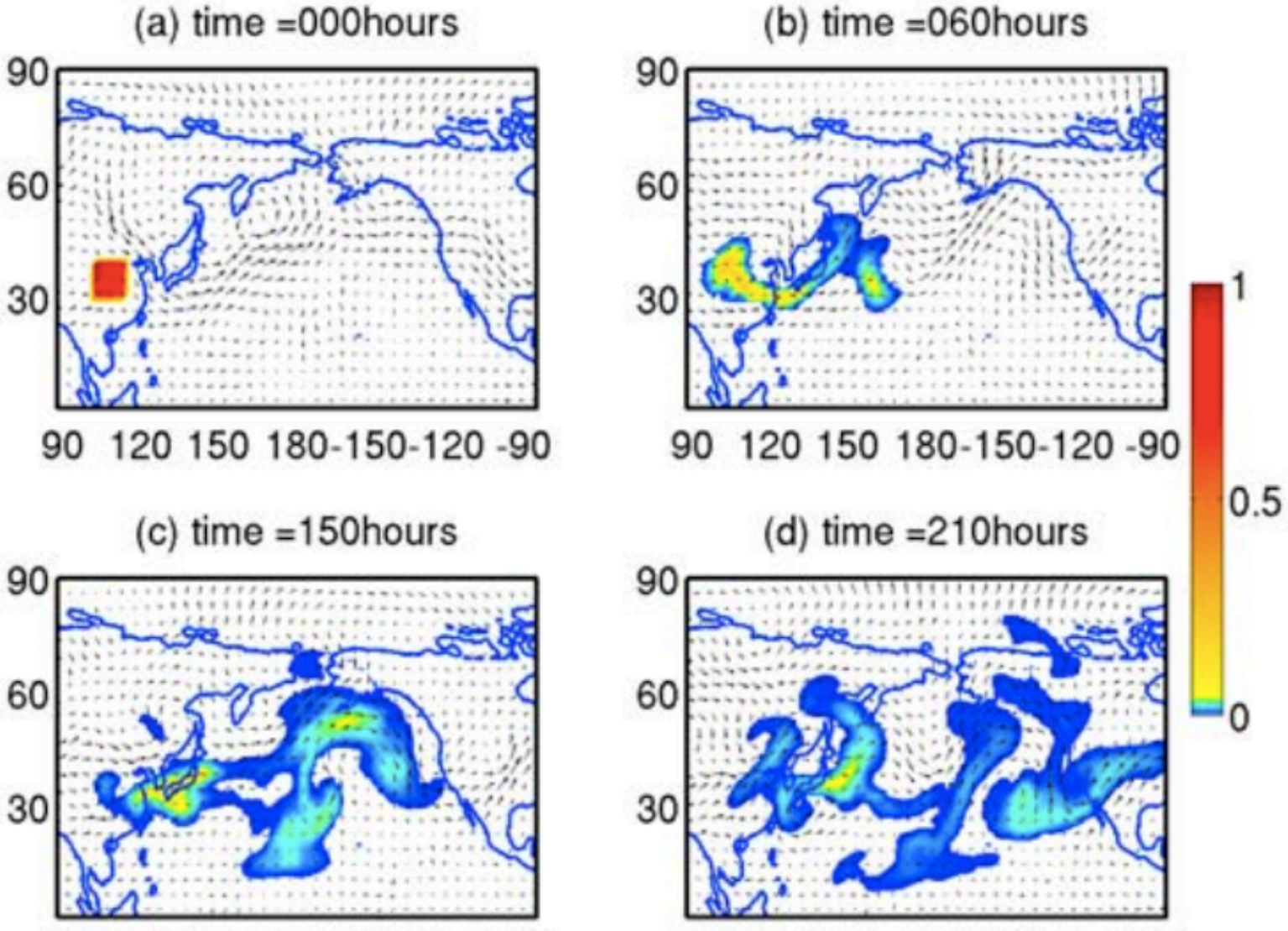
And continuous data of the Earth's winds…

Can we identify who emits what?

Is there a time scale at which information coming from an inverse modeling approach using a CTM is unreliable?

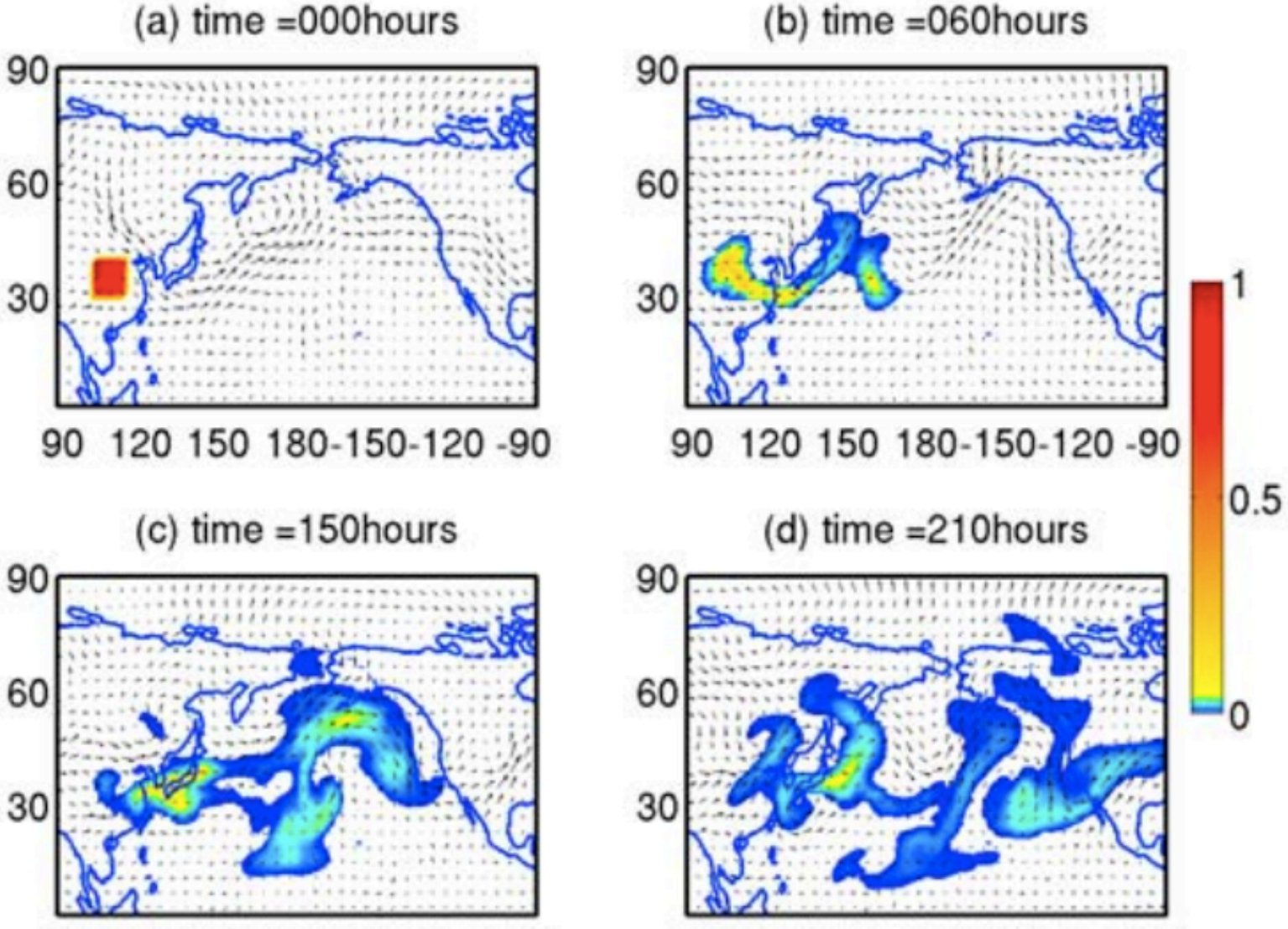
M. Santillana, M. A. Iglesias, M. P. Brenner. **Loss of information due to chaotic mixing in a time-dependent flow field. An inverse modeling perspective**. In preparation. 2011.

Numerical diffusion introduced by numerical transport schemes in CTMs has been observed to be **drastically enhanced** by **chaotic mixing,** naturally present in the atmosphere. (Rastigeyev et al., 2010)



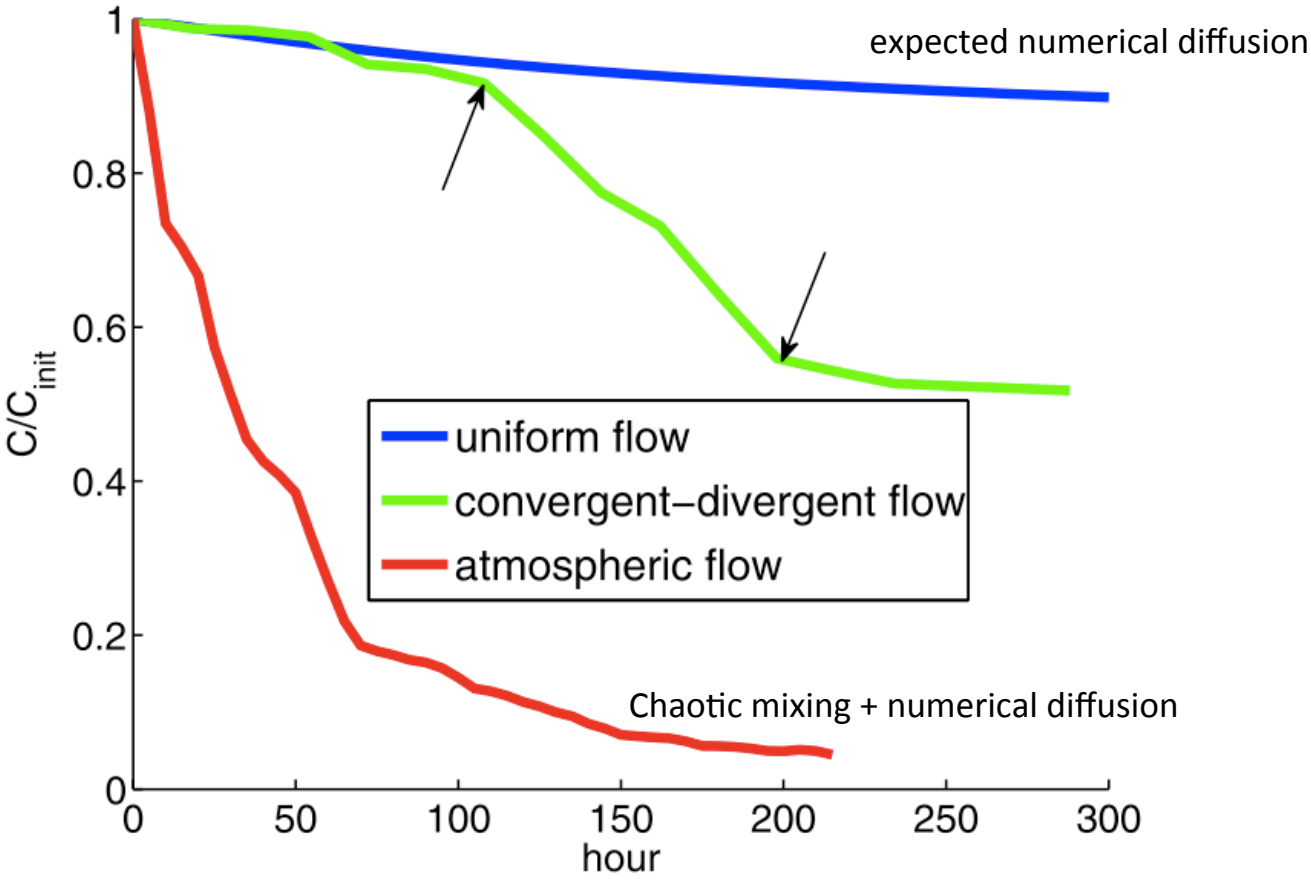
Piece-wise parabolic, flux form, eulerian, positive definite

This diffusion **reduces** the **ability** of CTMs to **reproduce** the **transport** of chemical species **accurately** as the spatial resolution is coarsened. In particular, it causes inter-continental plumes to decay much faster than expected from real observations.



Why does chaotic mixing enhance numerical diffusion?

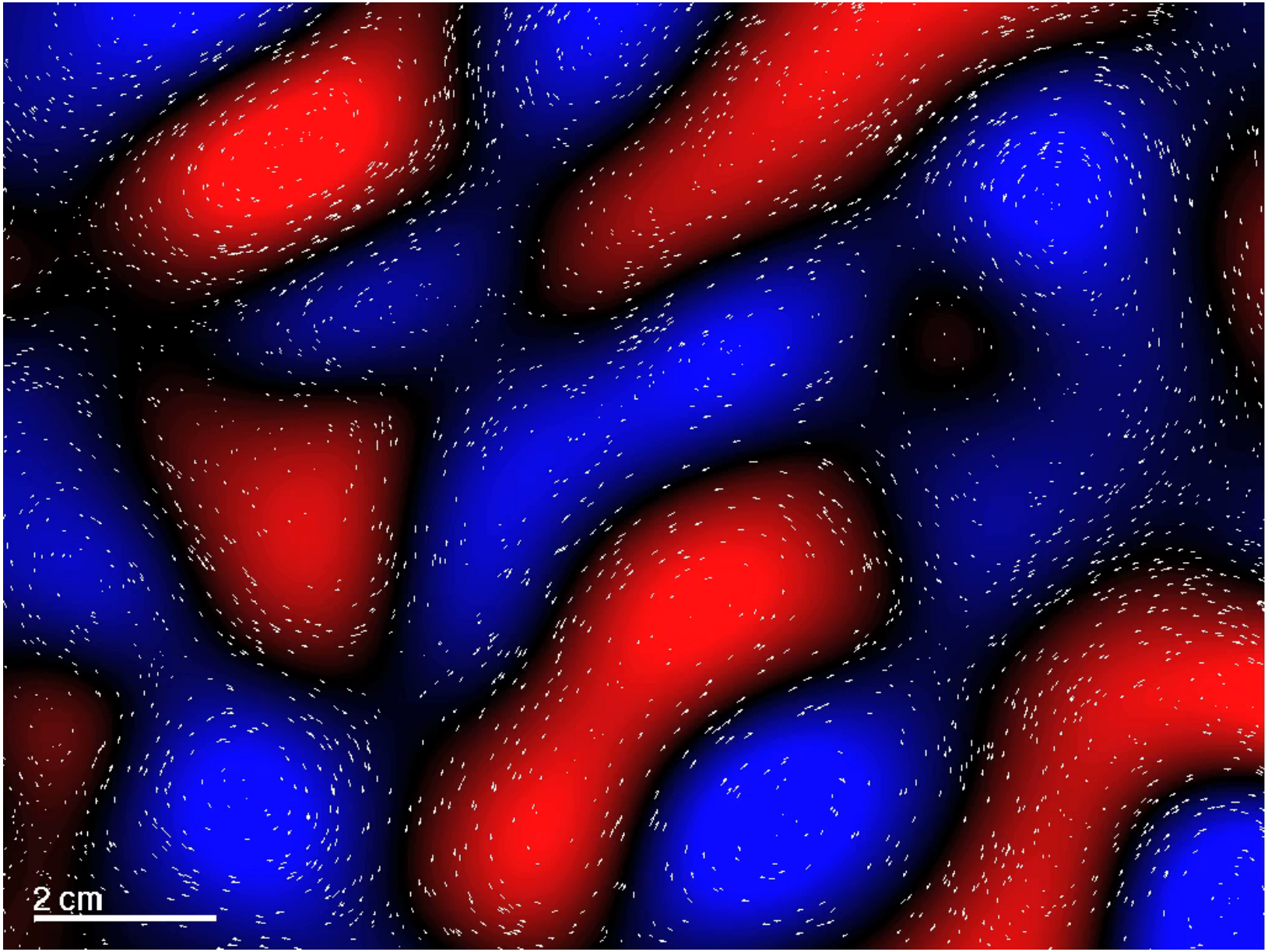
Chaotic mixing **stretches** plumes, reducing their width (filaments) and thus enhancing numerical diffusion.



Rastigeyev et al., 2010

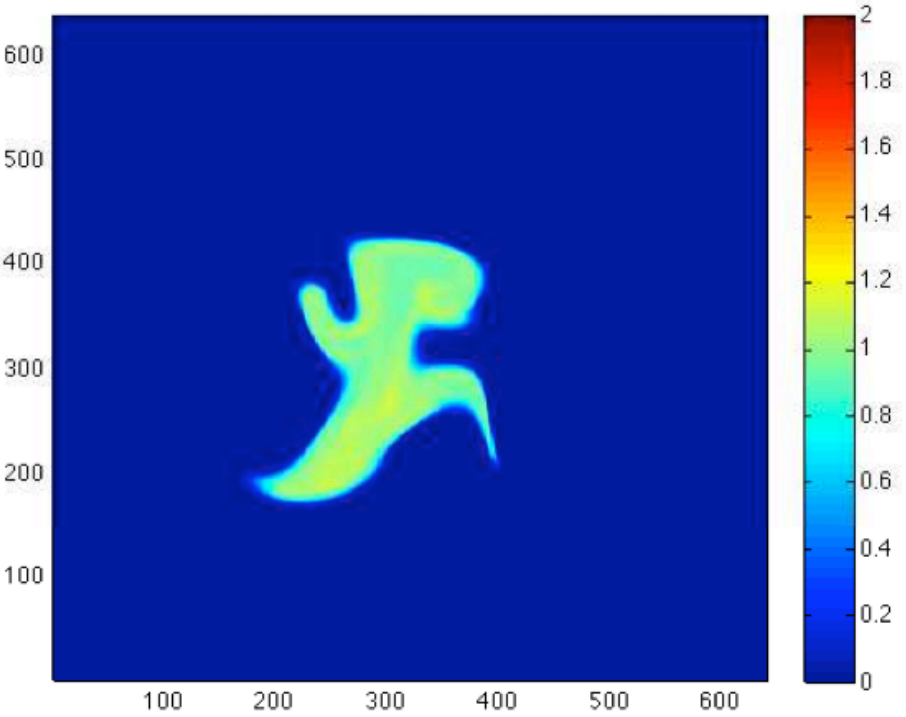
Performing numerical experiments with global models for verification purposes is involved

Alternatively we use a small scale prototype to show concept
(velocity fields from Nick Ouellete, Yale Univ.)

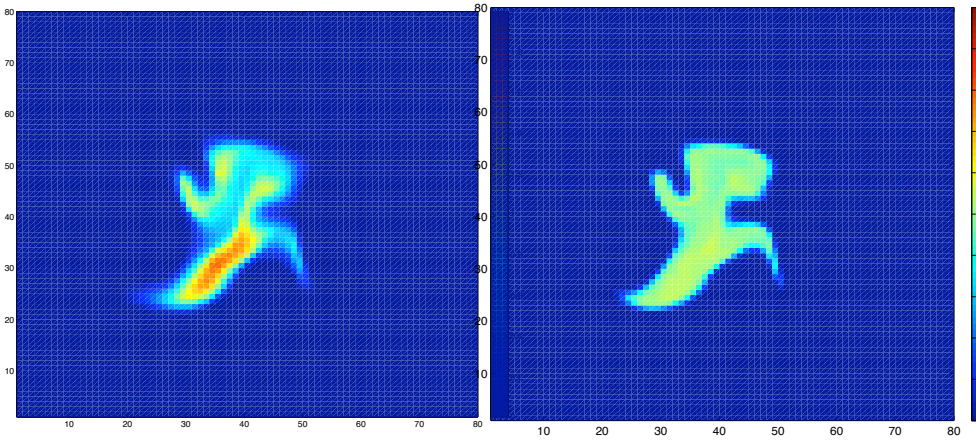


Data from Nick Ouellette, Yale Univ.

Snap shot of forward model (100 time steps)

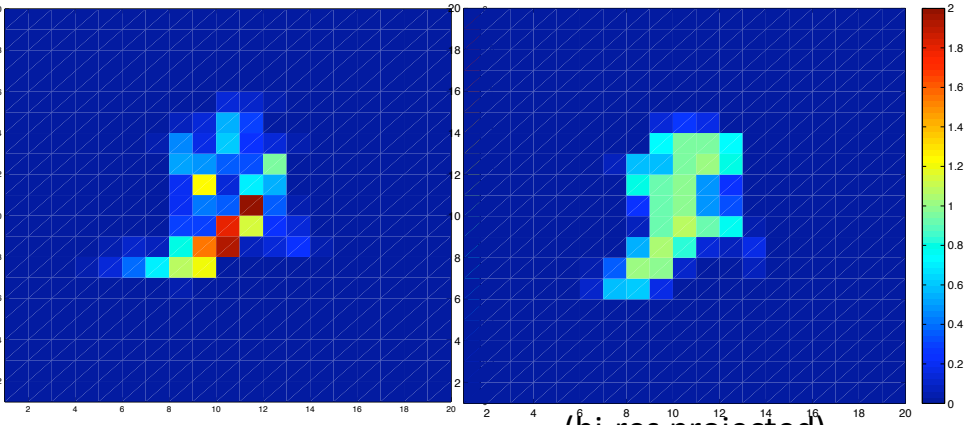


640 x 640



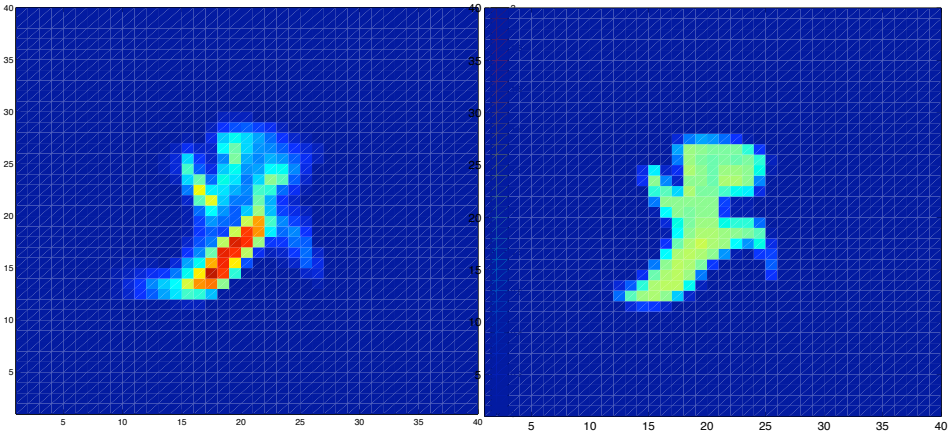
(hi-res projected)

80 x 80



(hi-res projected)

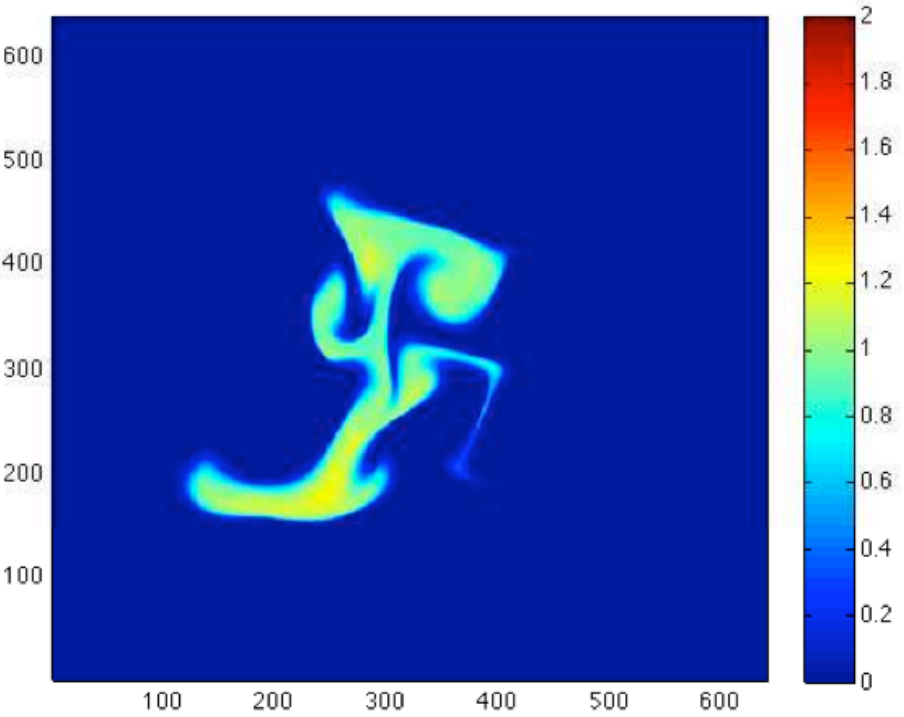
20 x 20



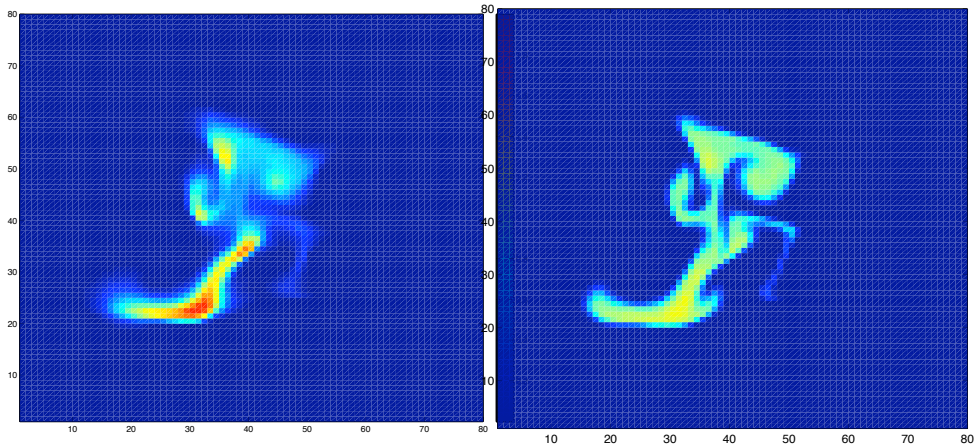
(hi-res projected)

40 x 40

Snap shot of forward model (200 time steps)

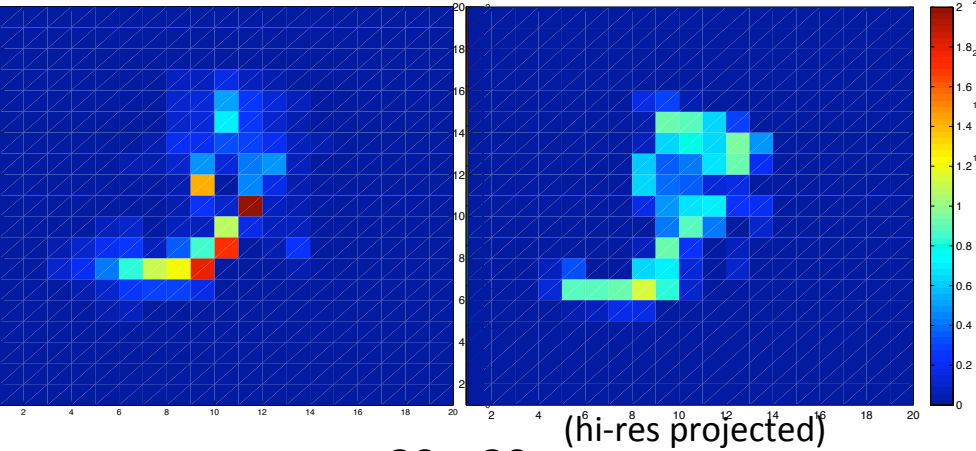


640 x 640



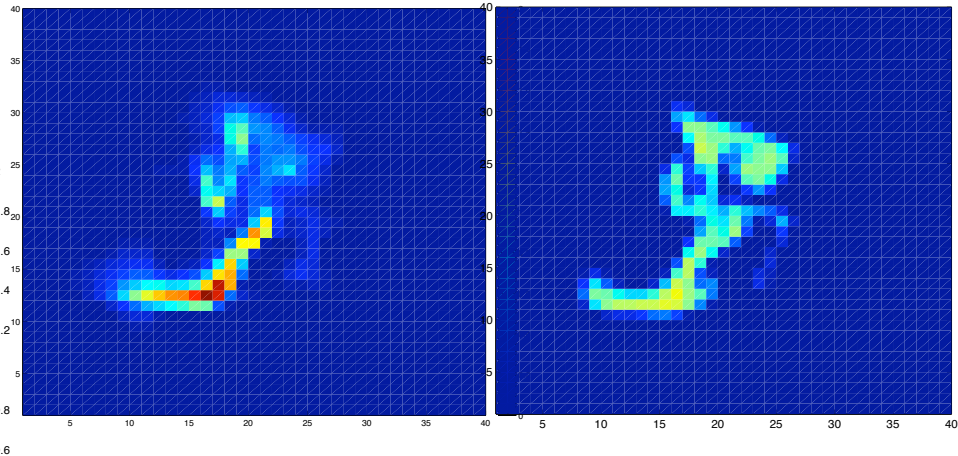
(hi-res projected)

80 x 80



(hi-res projected)

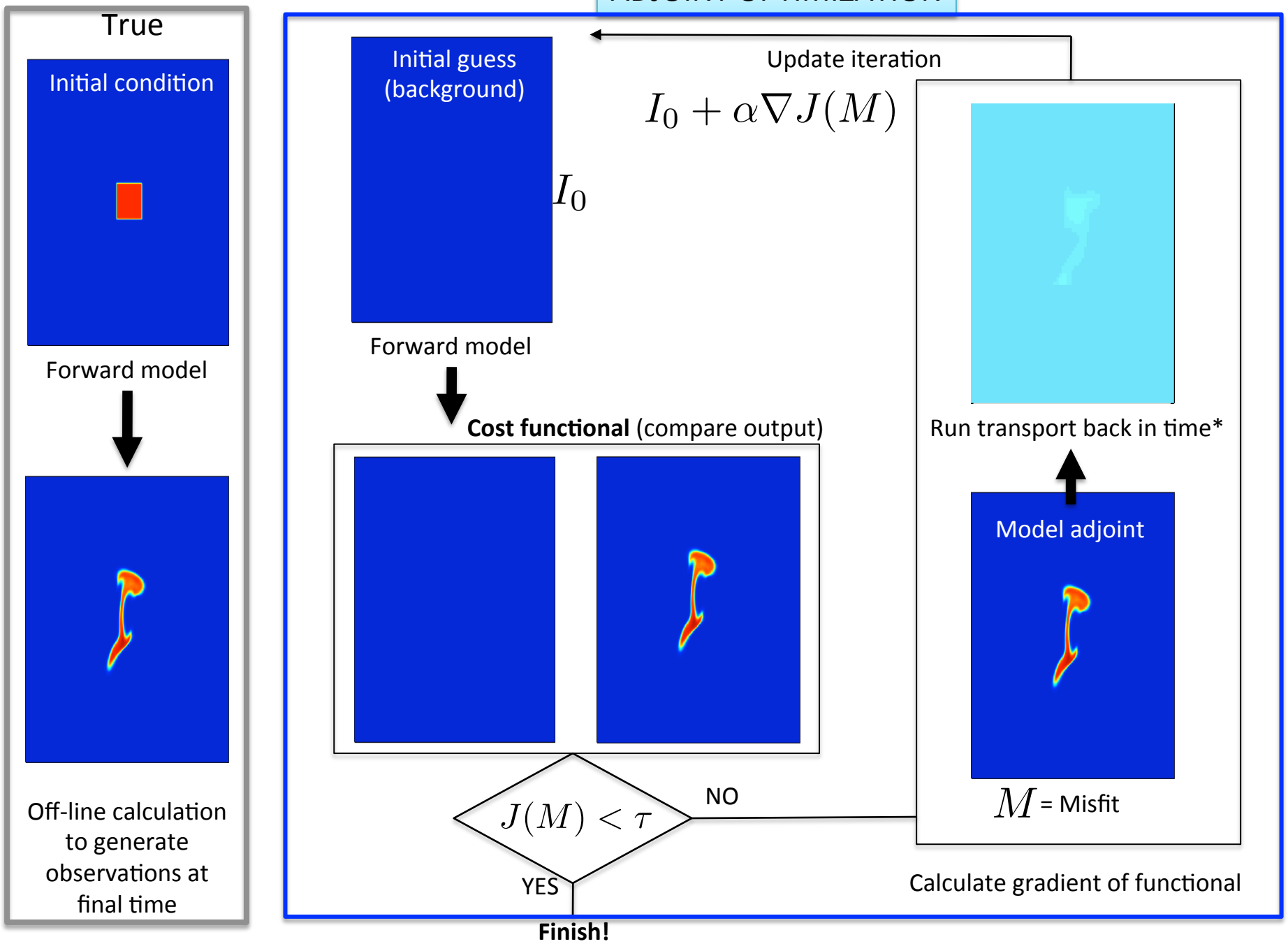
20 x 20



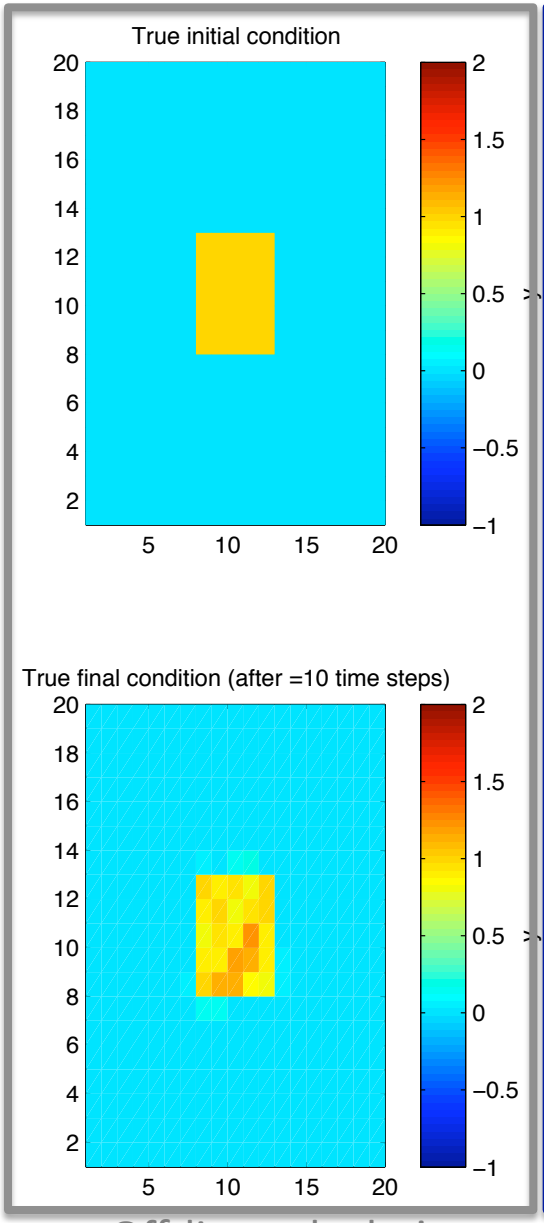
(hi-res projected)

40 x 40

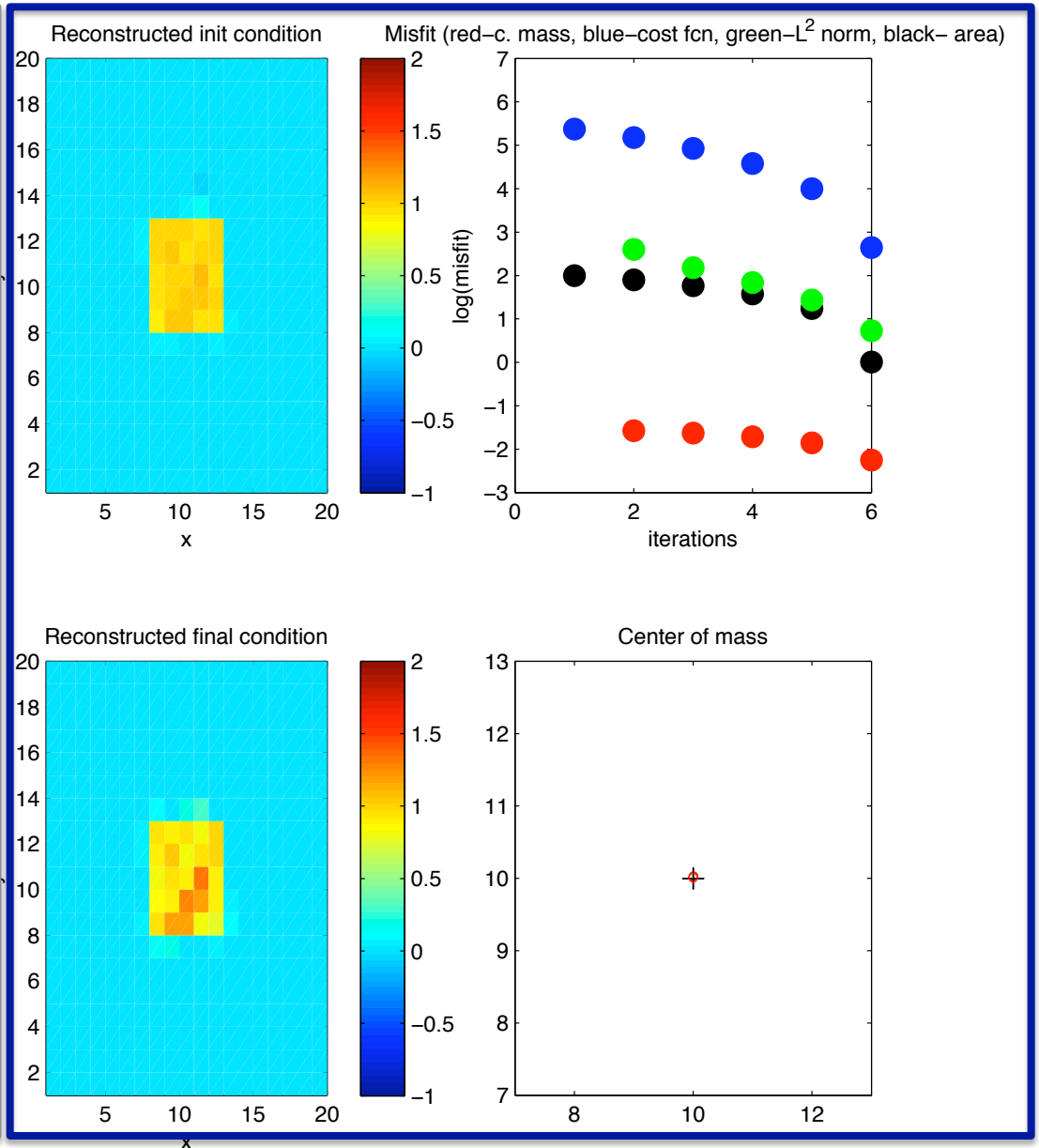
Verification: identical twin experiment



Adjoint optimization using **time-dep** winds
T= **10** time steps, resolution =**20 x 20**

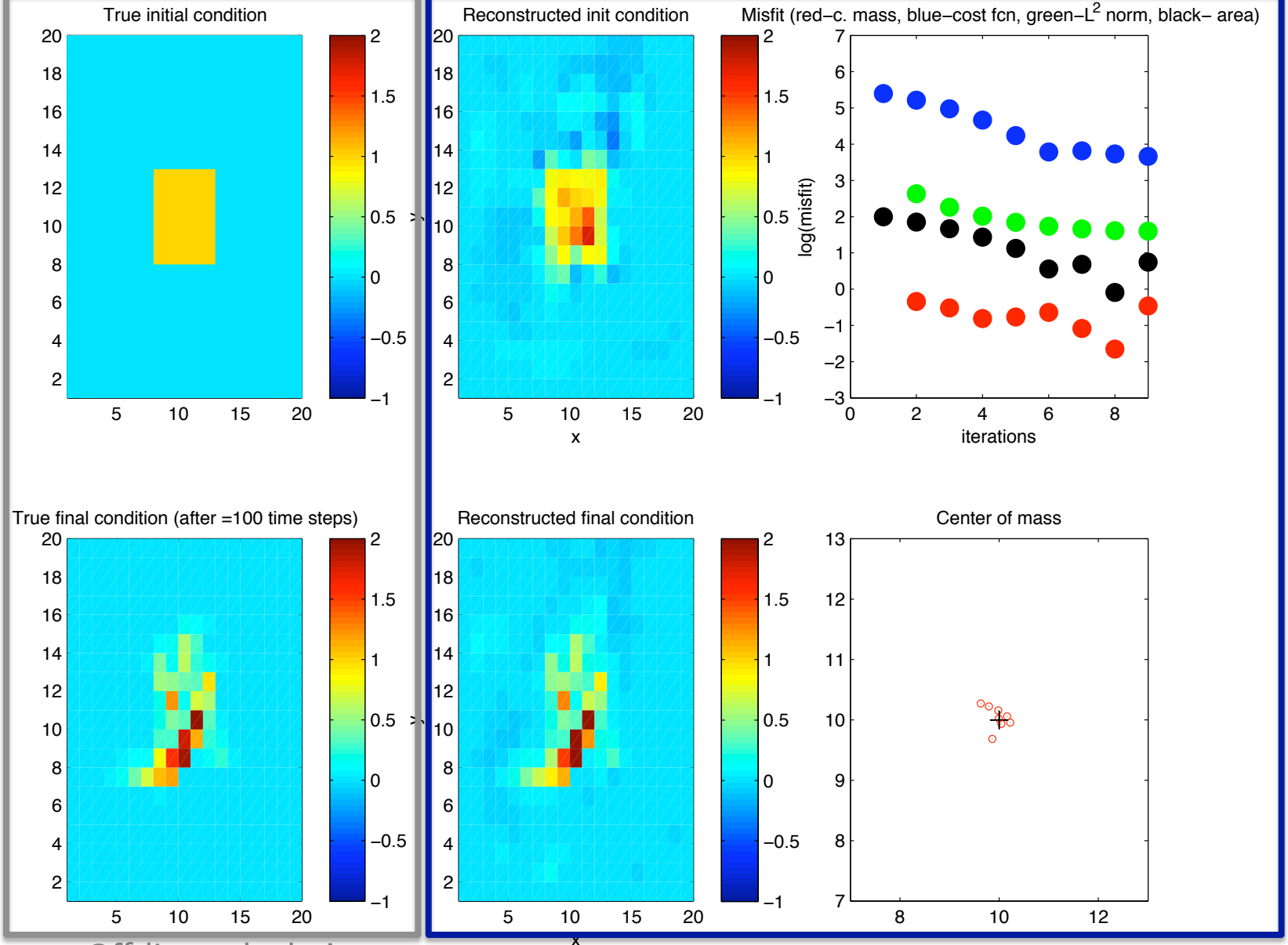


Off-line calculation

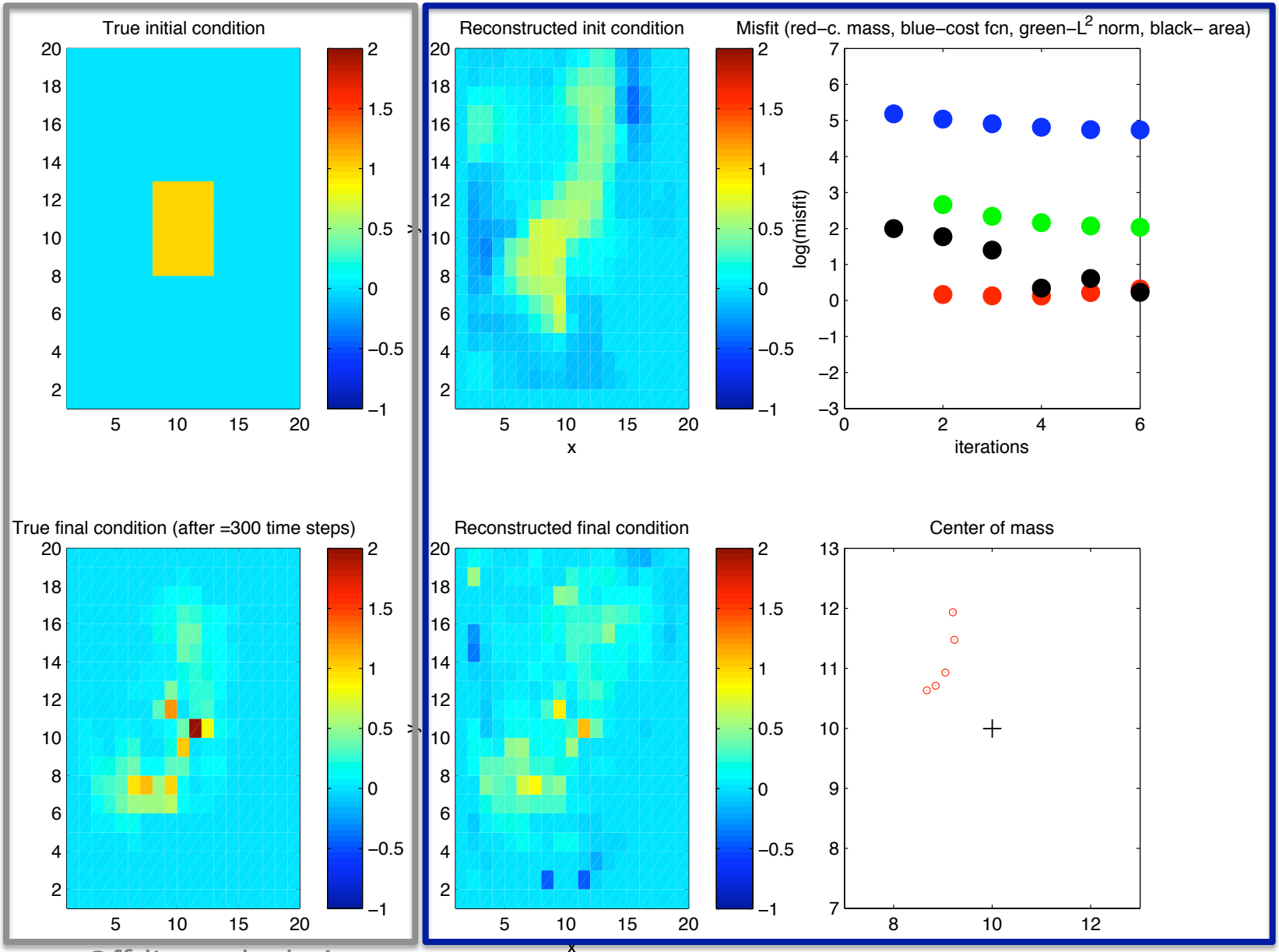


Adjoint optimization

Adjoint optimization using **time-dep** winds
T= **100** time steps, resolution =**20 x 20**



Adjoint optimization using **time-dep** winds
T= **300** time steps, resolution =**20 x 20**

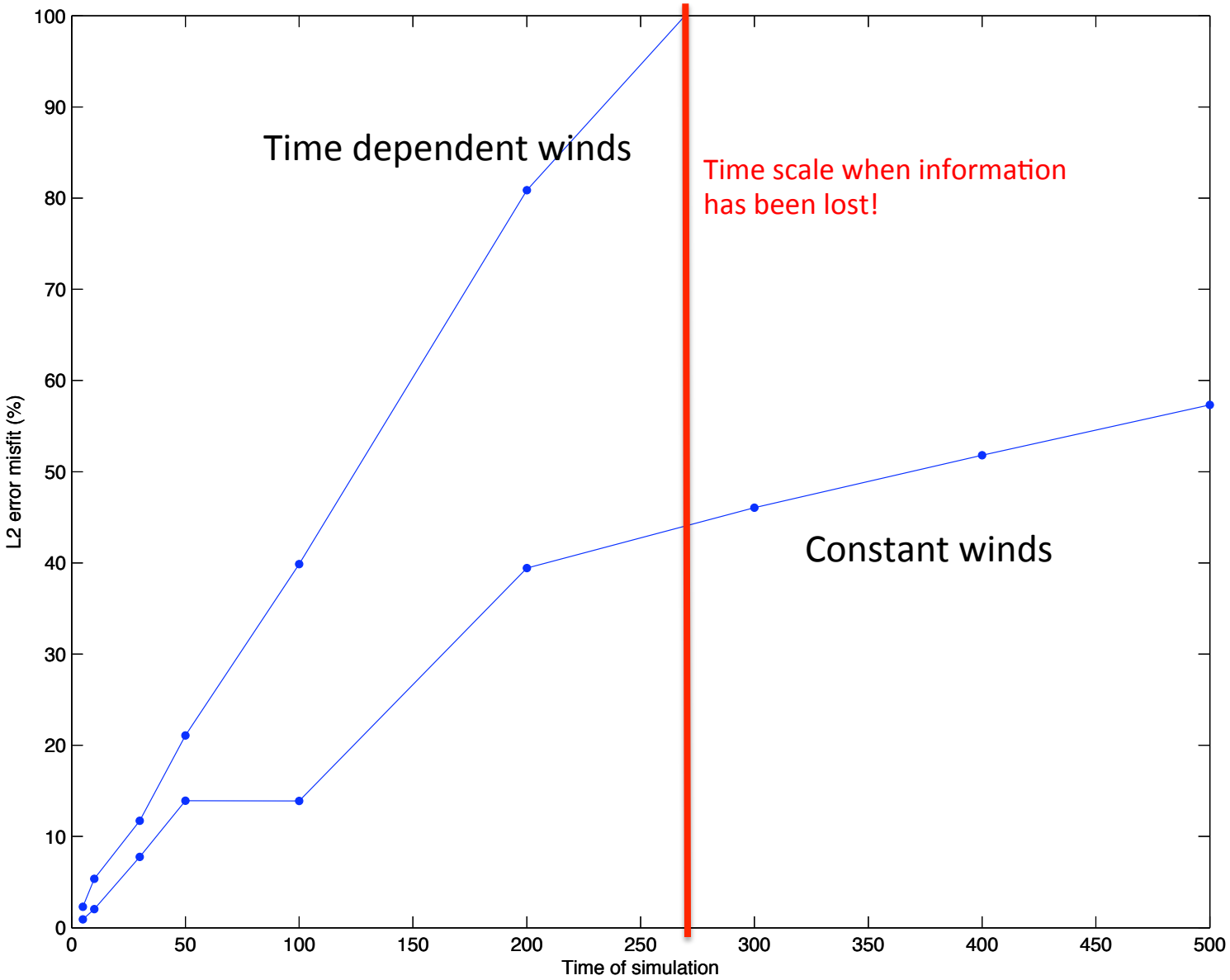


Off-line calculation

Adjoint optimization

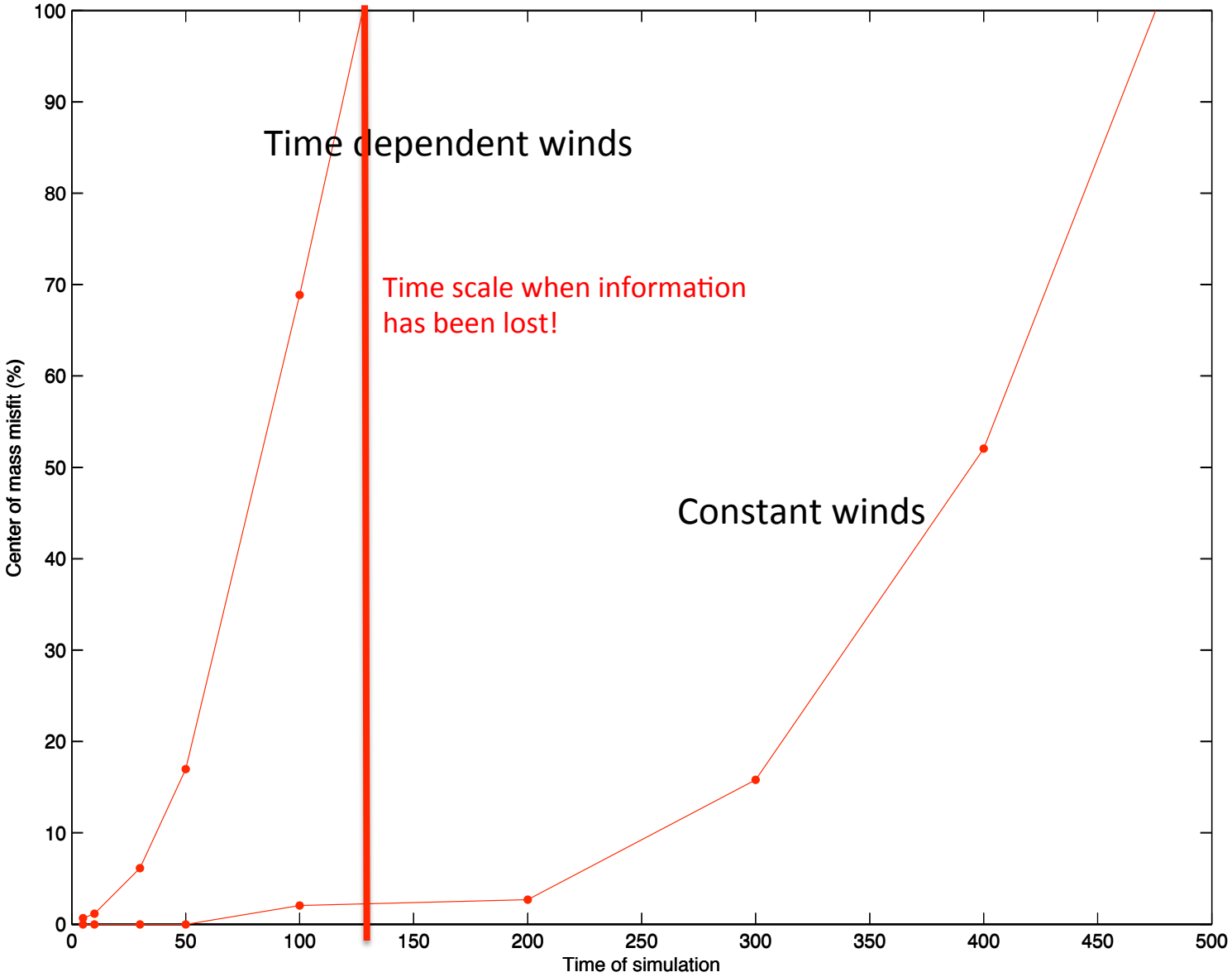
Twin experiments (observations generated at same resolution)

L^2 error for a well-resolved (big) plume at resolution **20 x 20**

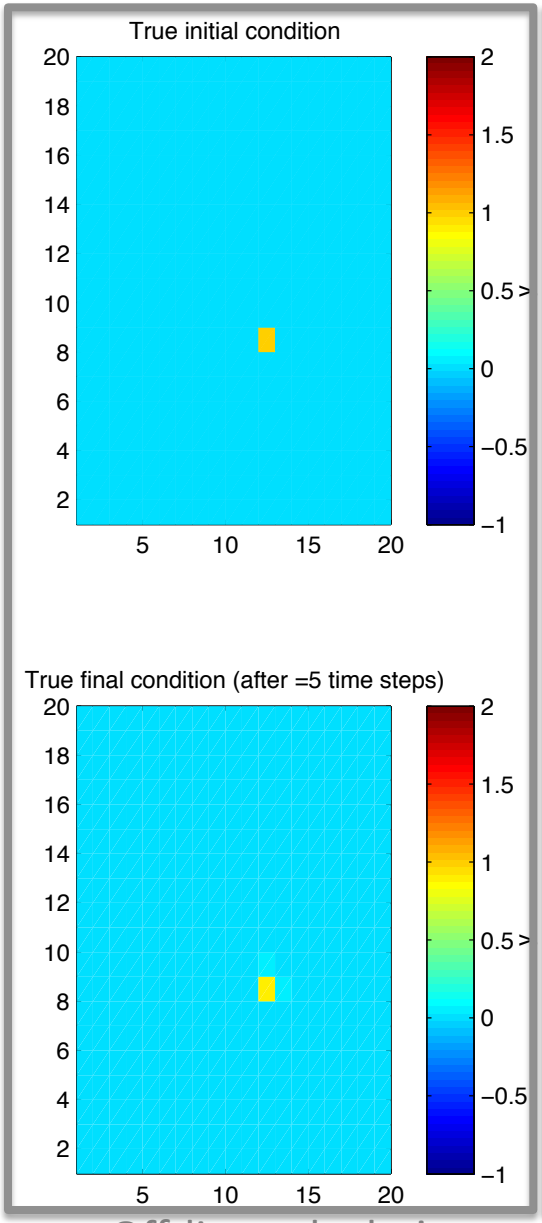


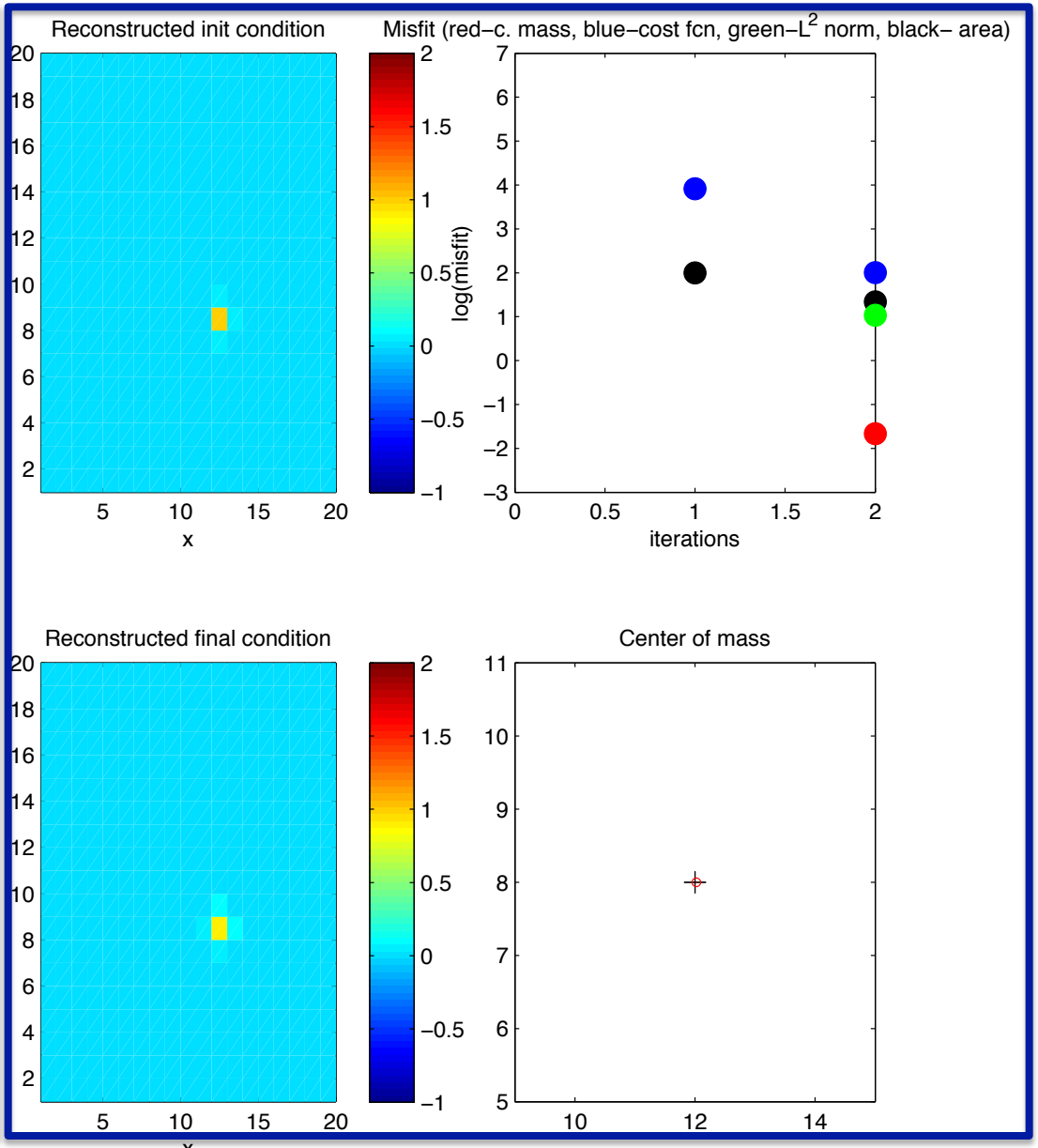
Twin experiments (observations generated at same resolution)

Center of mass error for a well-resolved (big) plume at resolution **20 x 20**



Adjoint optimization using **time-dep** winds
T= **5** time steps, resolution =**20 x 20 (small plume)**

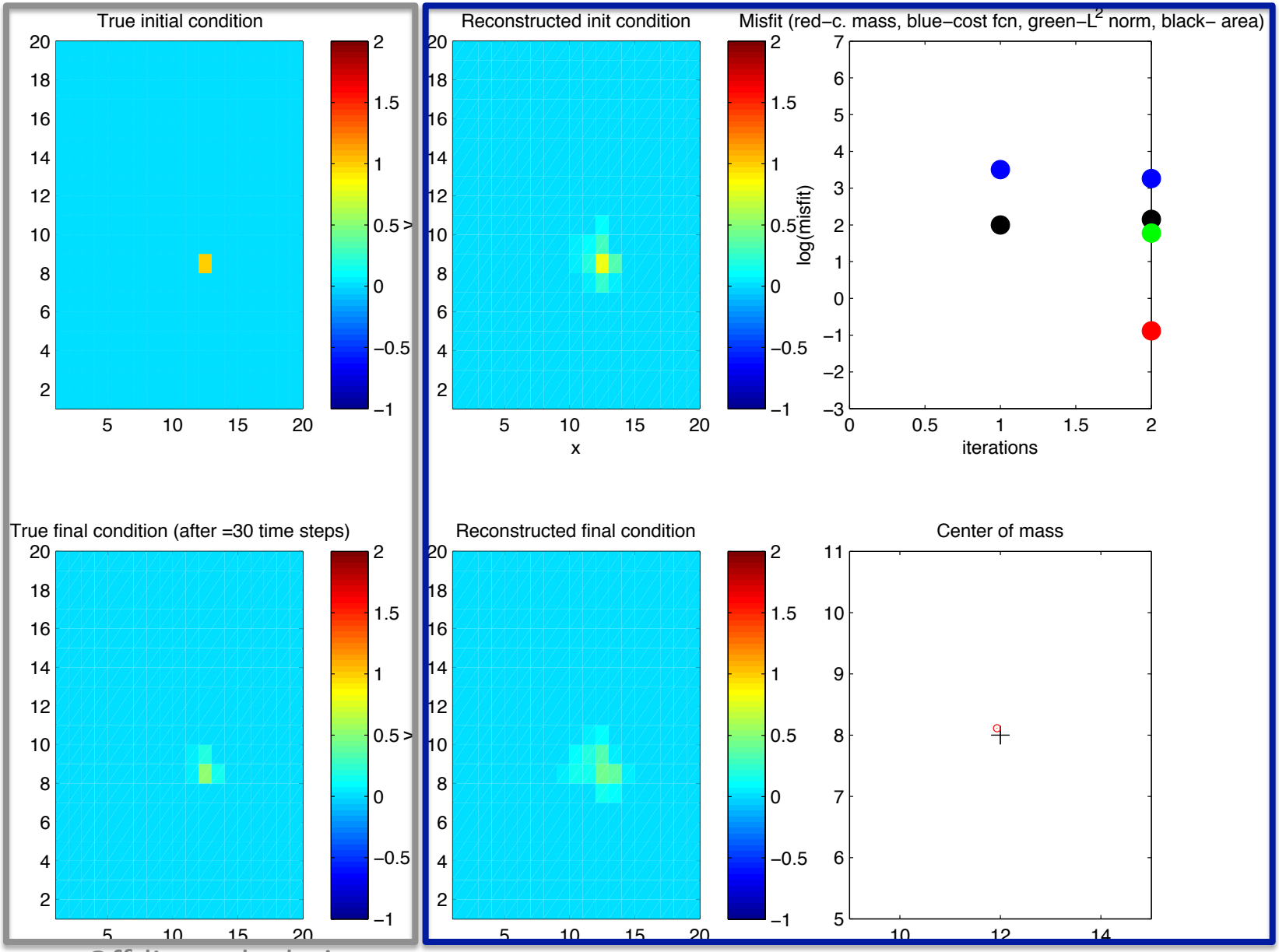




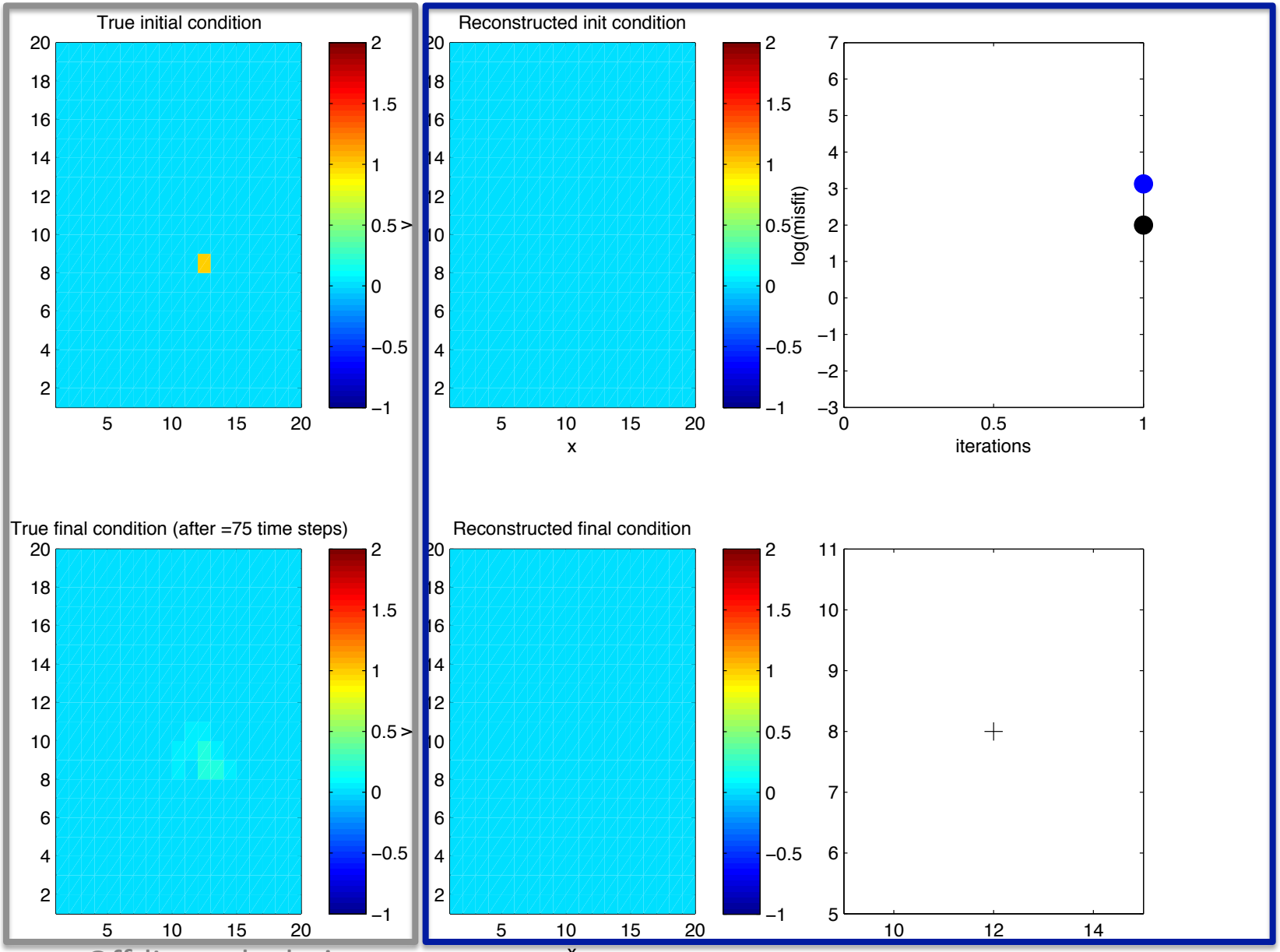
Off-line calculation

Adjoint optimization

Adjoint optimization using **time-dep** winds
T= **30** time steps, resolution =**20 x 20 (small plume)**

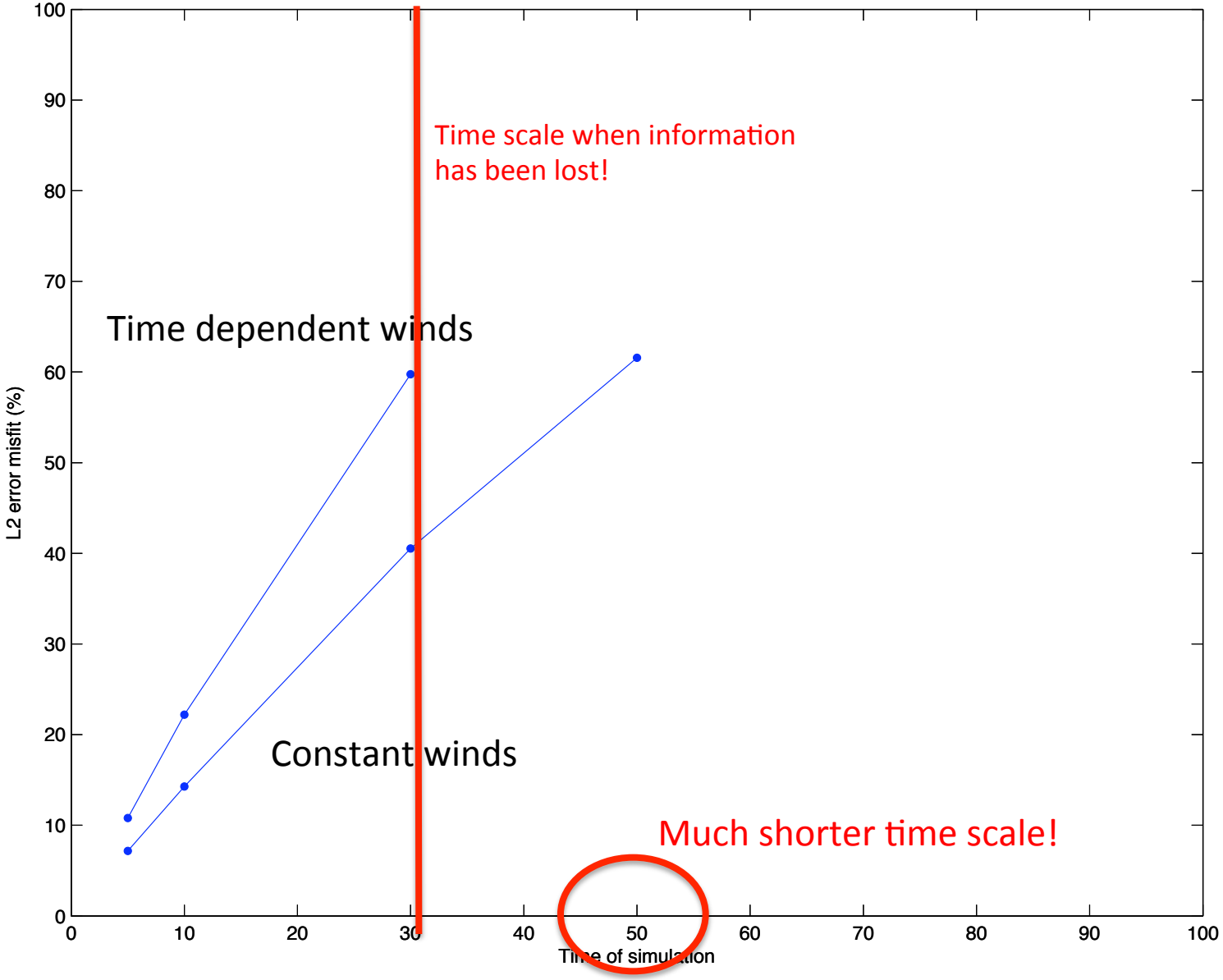


Adjoint optimization using **time-dep** winds
T= **75** time steps, resolution =**20 x 20 (small plume)**



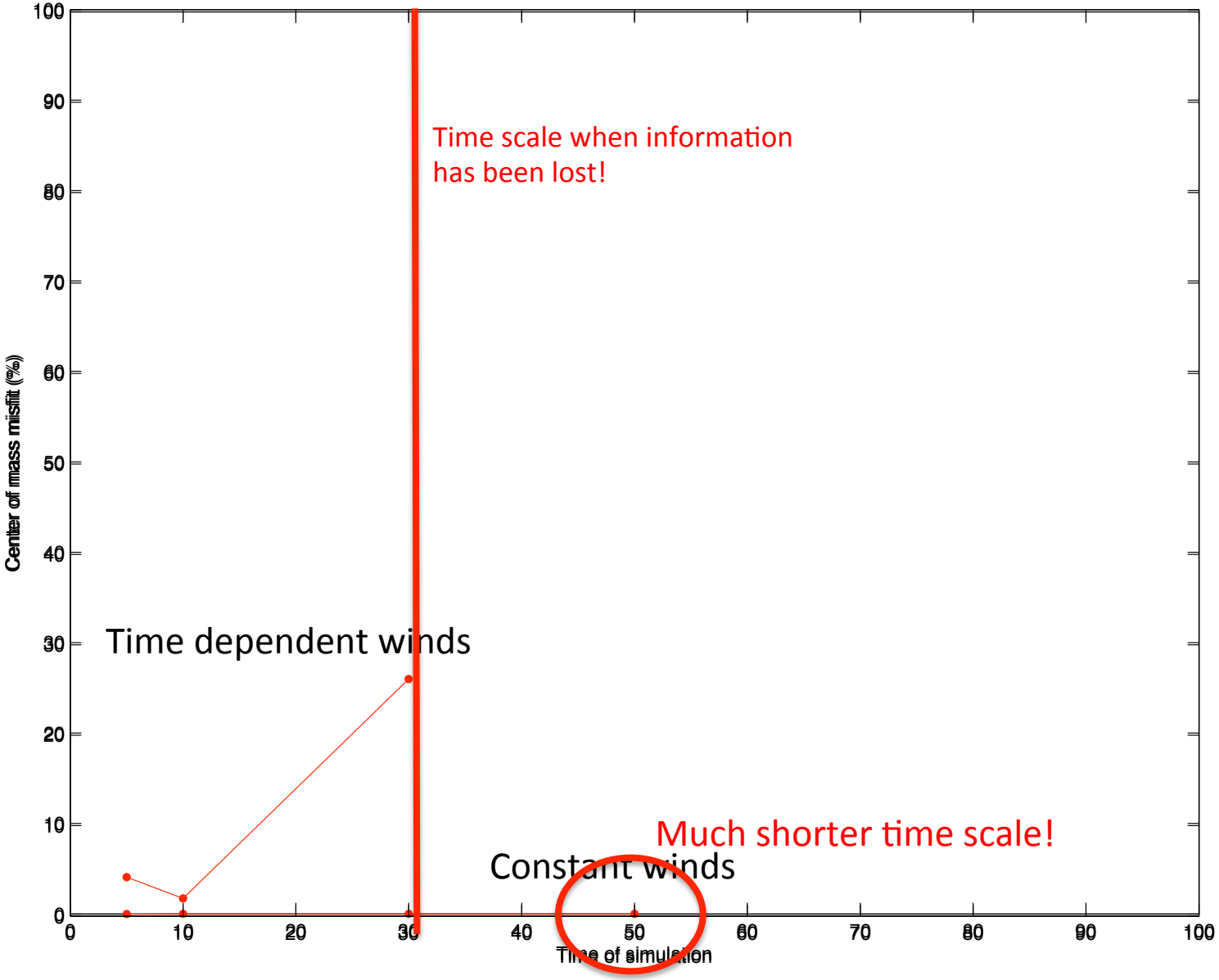
Twin experiments (observations generated at same resolution)

L^2 error for a not well-resolved plume (small) at resolution **20 x 20**

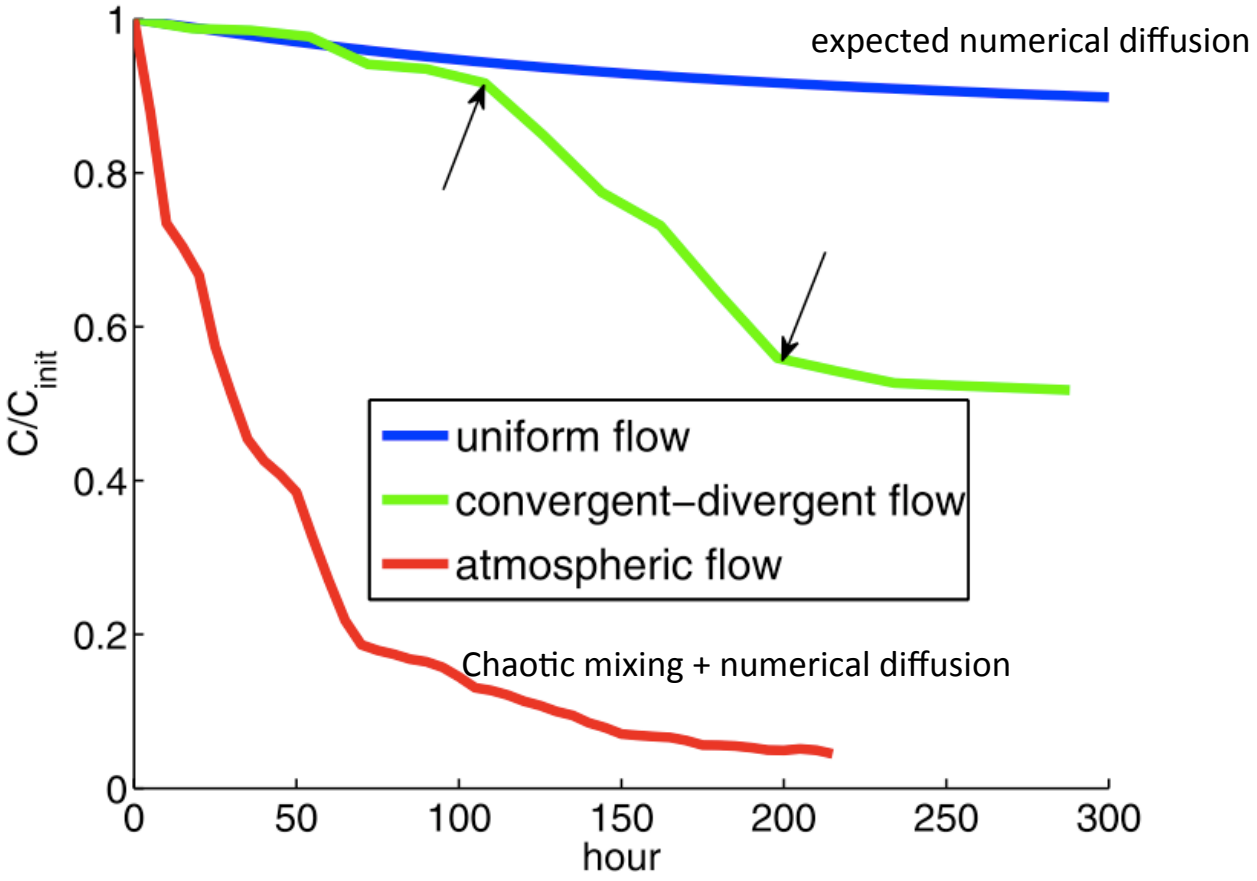


Twin experiments (observations generated at same resolution)

Center of mass error for a not well-resolved plume (small) at resolution **20 x 20**



Guess: In global models this time scale < 10 days



Rastigeyev et al., 2010

Simulations depend heavily on grid size
(Verification)

Simulations depend heavily on grid size

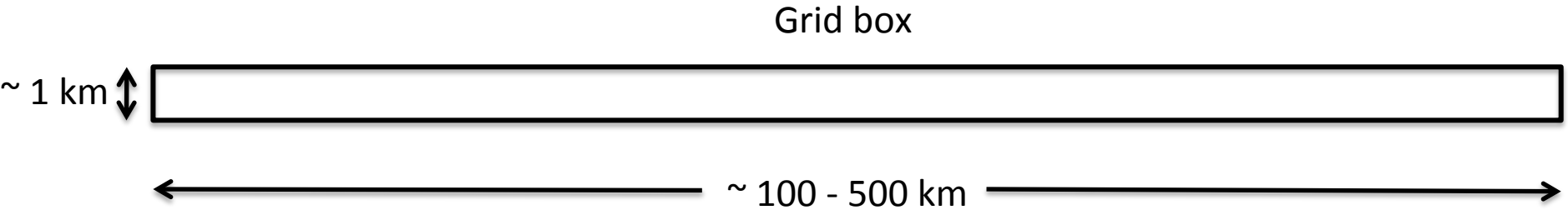
The Mathematics… Set of coupled partial differential equations

resolved transport | chemistry | sub-grid processes

$$\frac{\partial n_i}{\partial t} = \underbrace{-\nabla \cdot (\boldsymbol{u}(t) n_i)}_{\text{resolved transport}} + \underbrace{R(t, \boldsymbol{n})}_{\text{chemistry}} + \underbrace{C_{sub-grid}(n_i)}_{\text{sub-grid processes}}$$

$$i = 1, ..., 110$$

Why sub-grid processes?



Increase of computational resources = refine mesh!

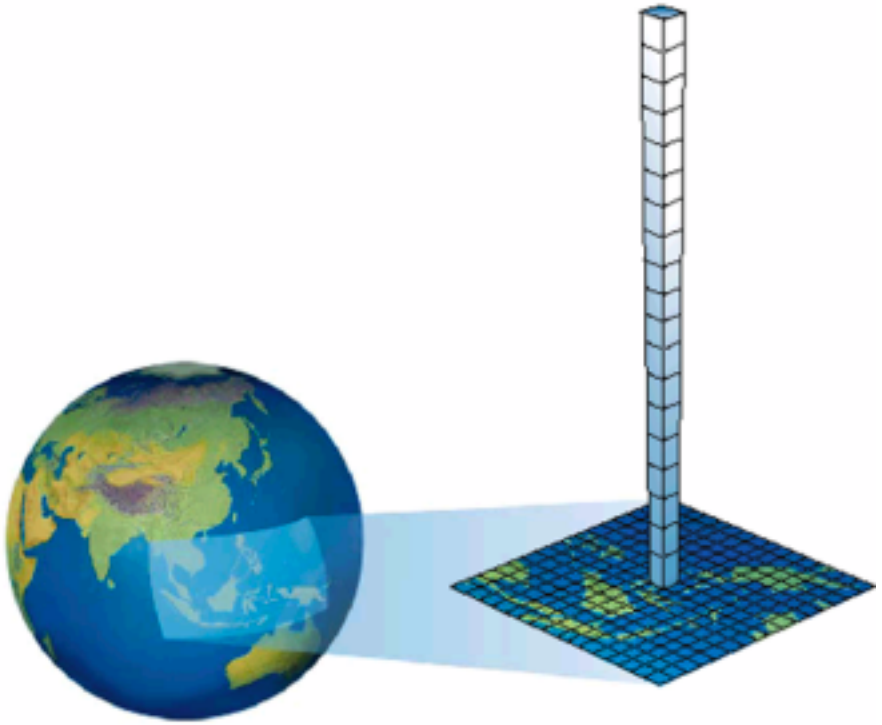
How should we go from global to local?

Results depend on temporal and spatial resolutions!

From Global (IPCC models) clim
change scenarios

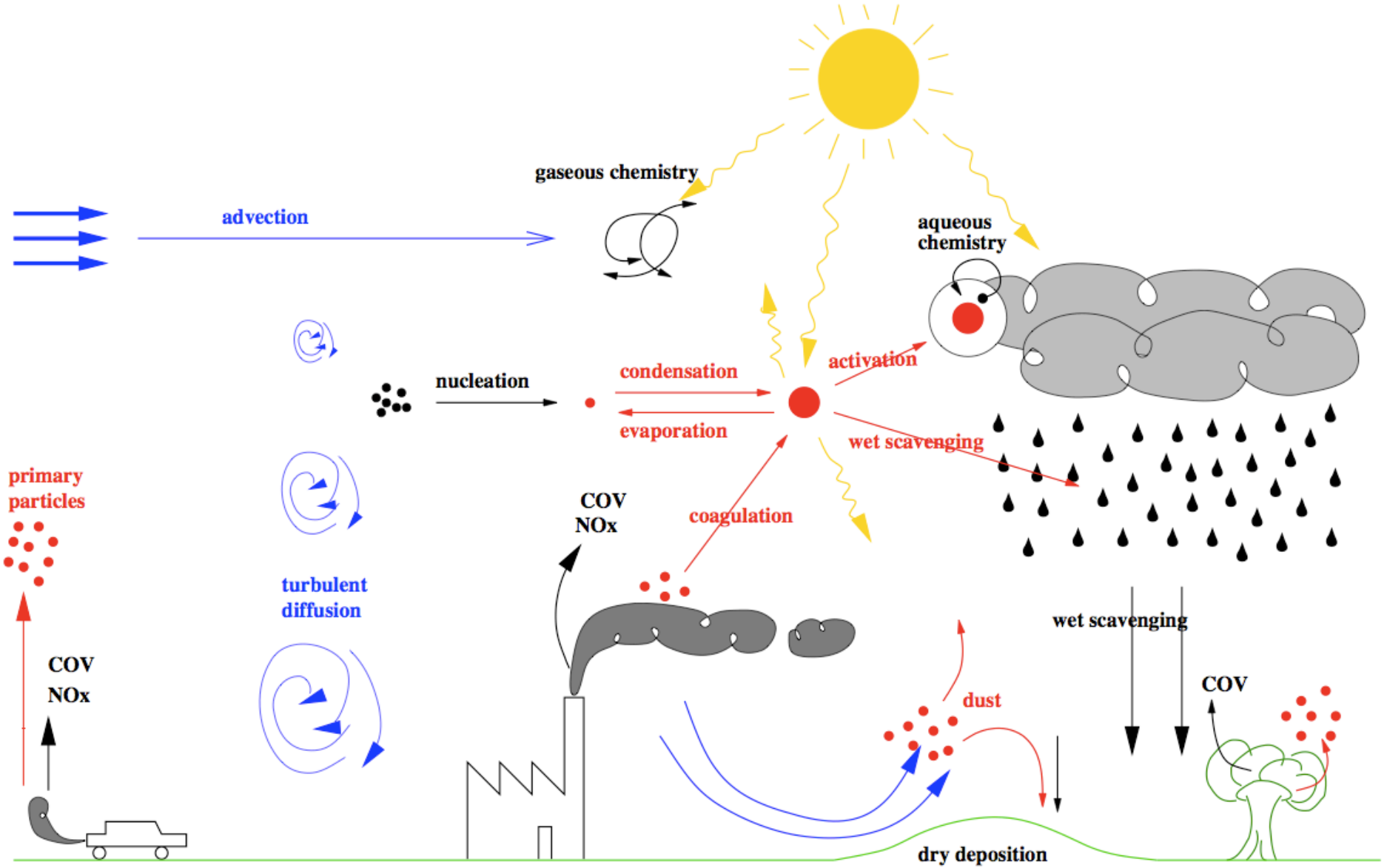
hope??

Regional
models



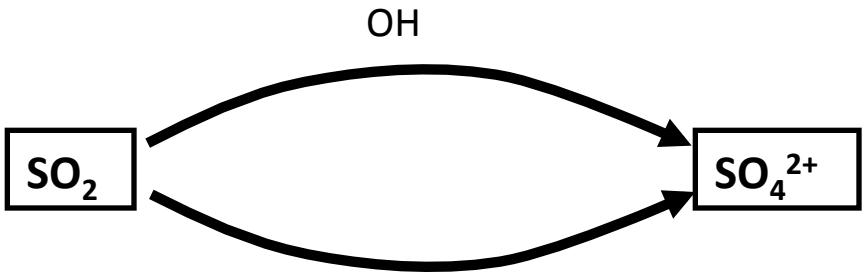
http://www.cru.uea.ac.uk/

Sub-grid processes



Sportisse 2007

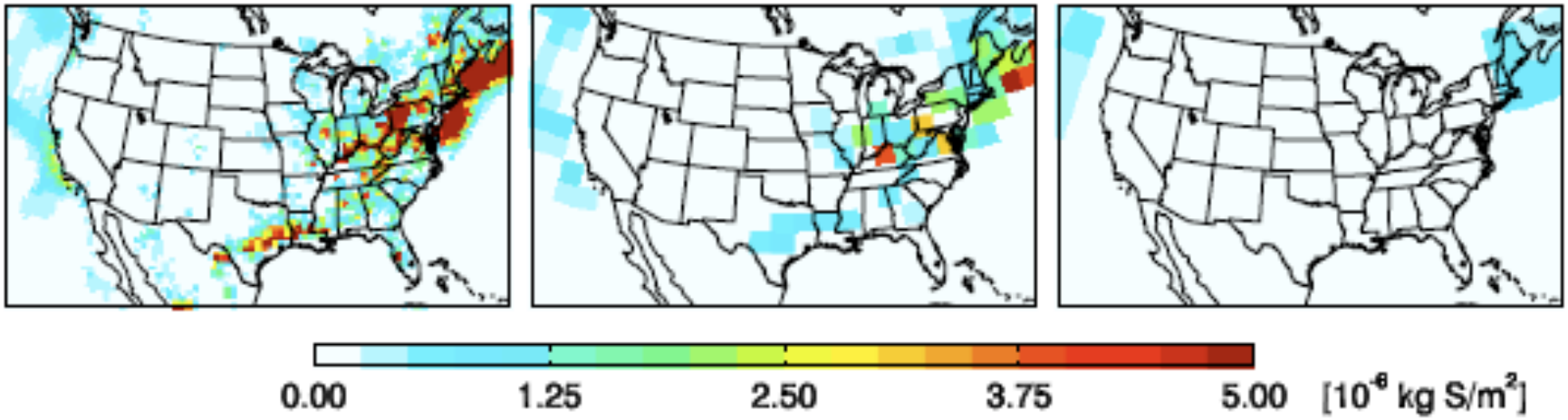
Sub-grid processes



O_3 and H_2O_2 in cloud

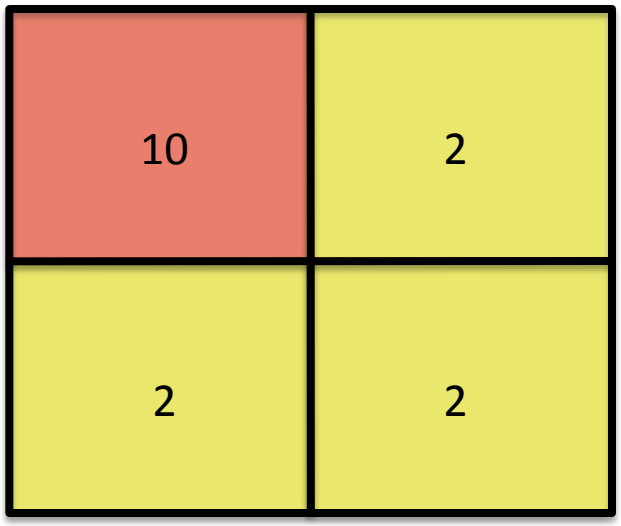
Thresholds in the model
(Cloud, T > 258 K)

Sulfate production from oxidation by H_2O_2 in cloud for July



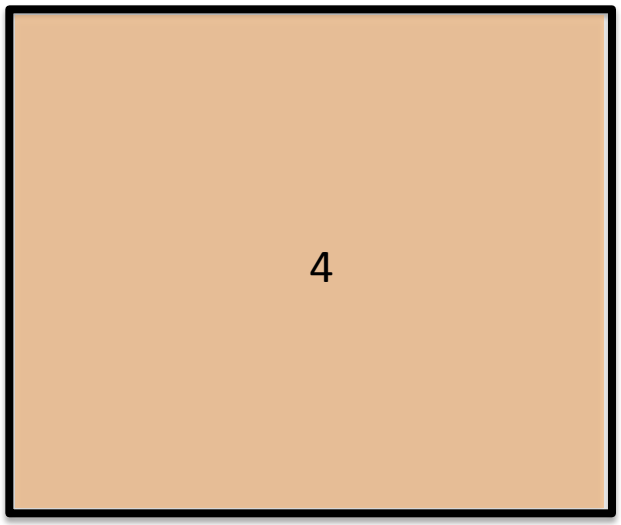
Lin et al. Work in progress. Harvard School of Engineering

Temperature (fine scale)

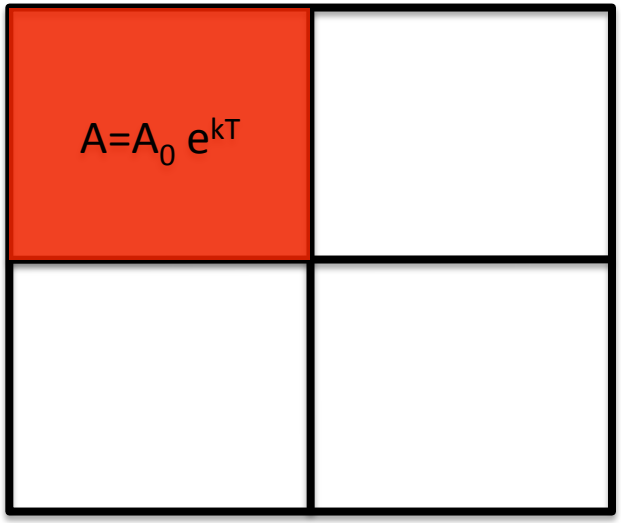


Average

Temperature (coarse scale)

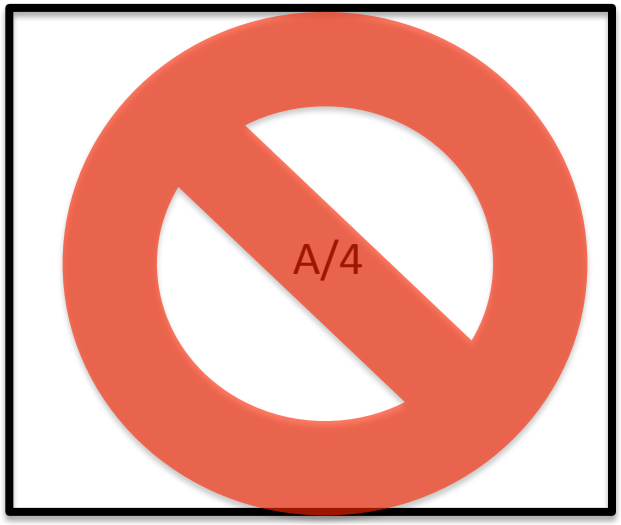


$$\frac{d[A]}{dt} = k(T > 5)[A][B]$$

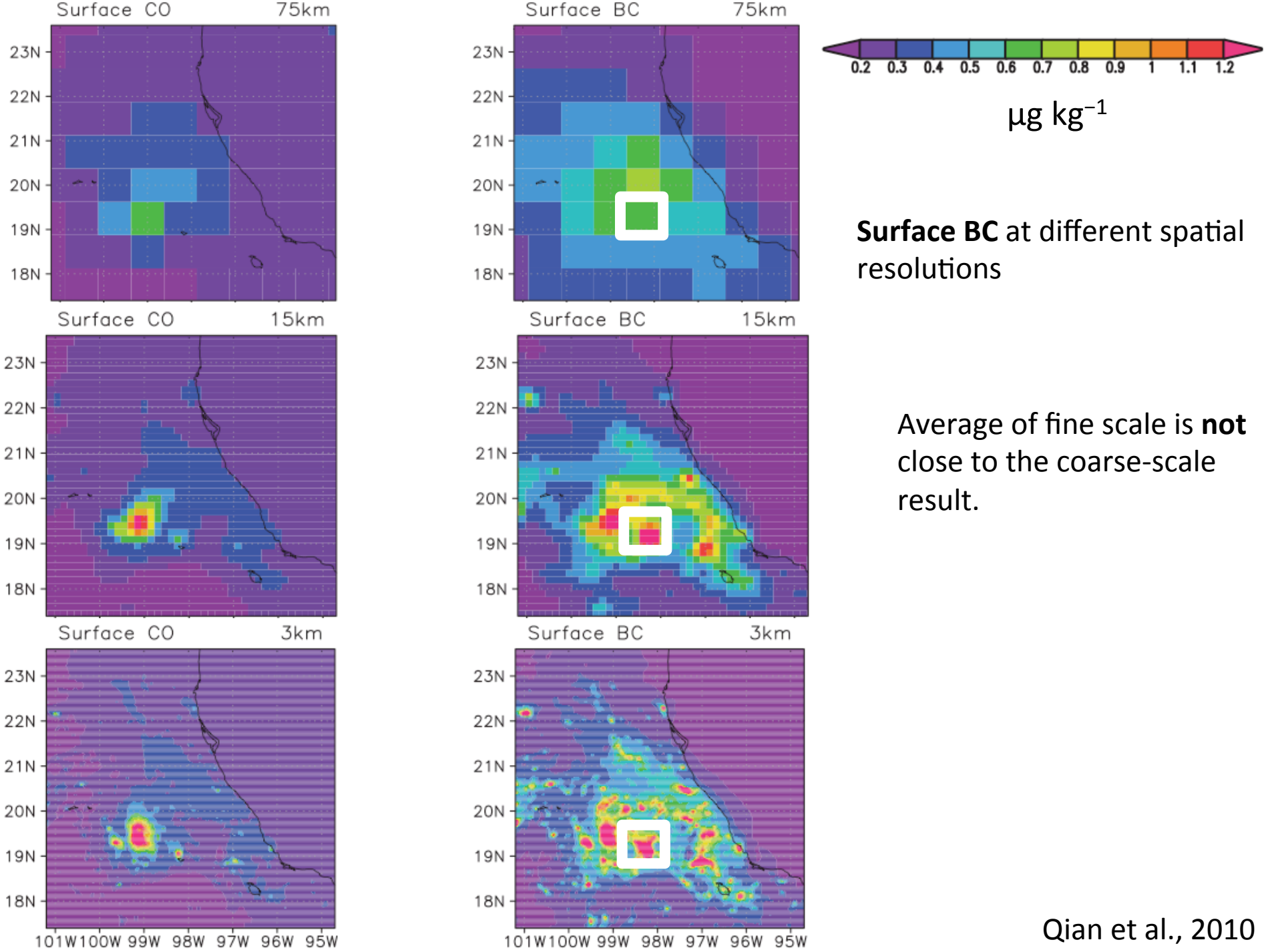


Concentration of species [A]

Average **??**



Concentration of species [A]

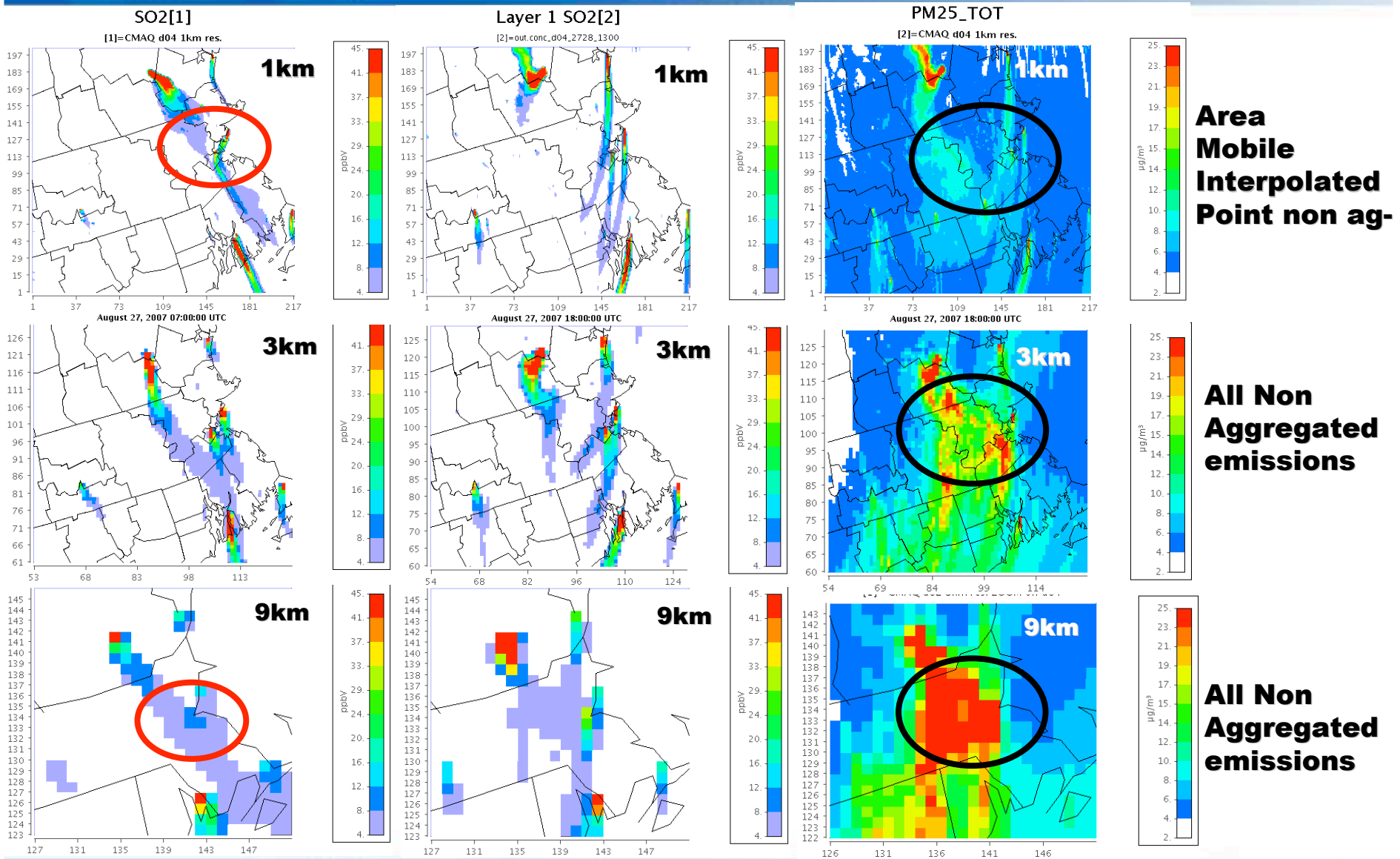


Surface BC at different spatial resolutions

Average of fine scale is **not** close to the coarse-scale result.

Qian et al., 2010

Plume. Resol. effects vs. d04. 08/27 18UT



Area Mobile Interpolated Point non ag-

All Non Aggregated emissions

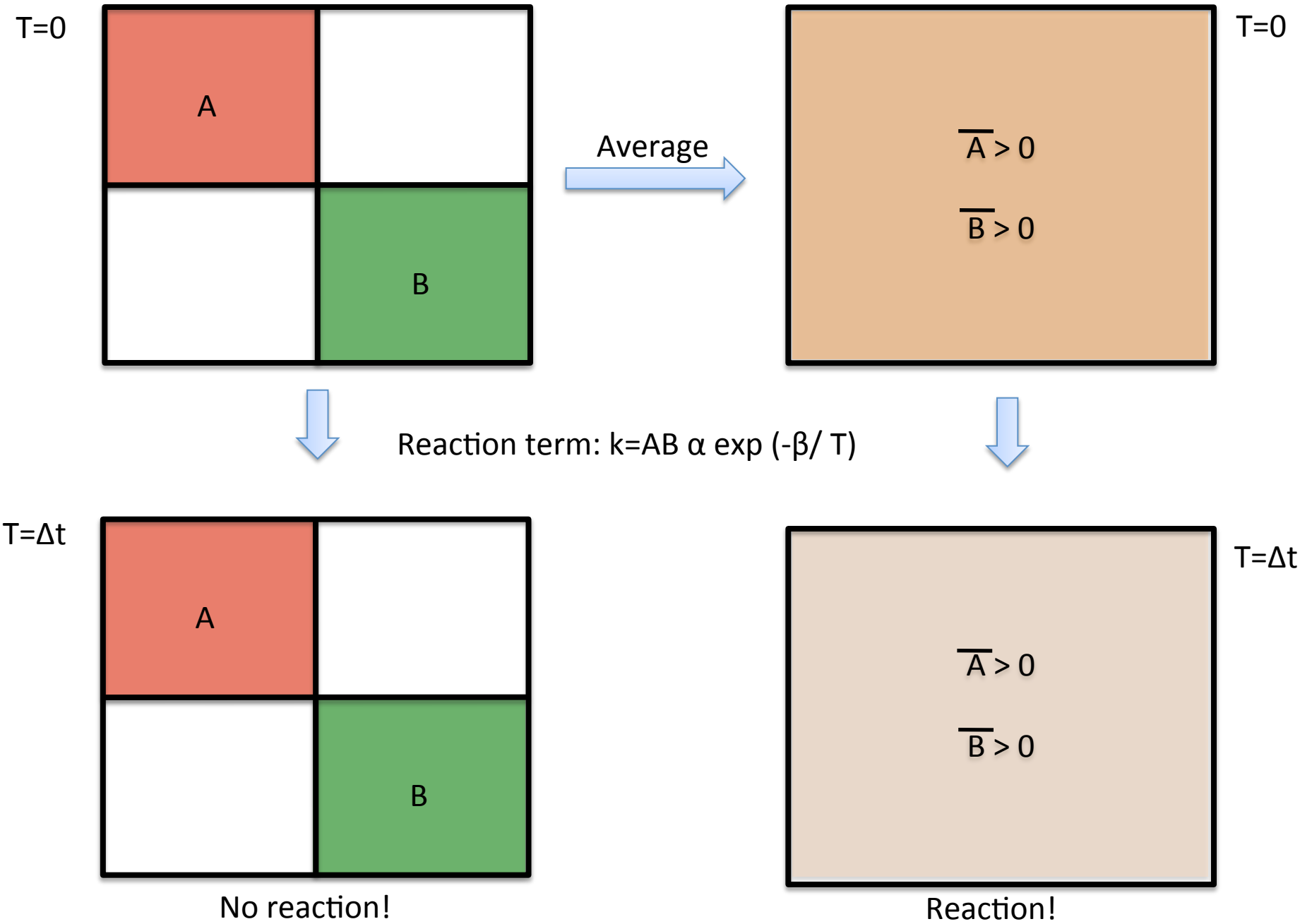
All Non Aggregated emissions

Richard Ramaroson (personal communication)

ONERA

Multi-scale dependence: No transport

2 or more species



Back to the equations…

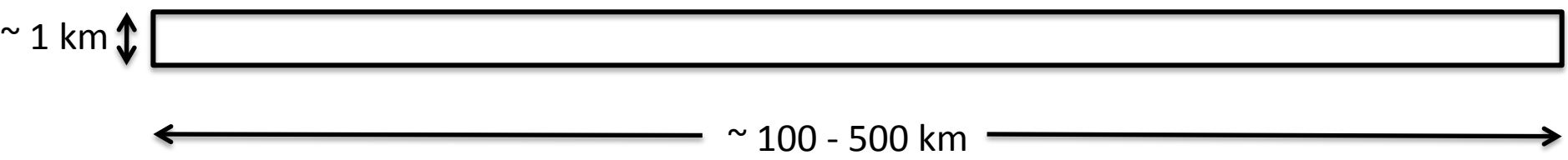
The Mathematics… a closer look

resolved transport | chemistry | sub-grid processes

$$\frac{\partial n_i}{\partial t} = \underbrace{-\nabla \cdot (\boldsymbol{u}(t) n_i)}_{\text{resolved transport}} + \underbrace{R(t, \boldsymbol{n})}_{\text{chemistry}} + \underbrace{C_{sub-grid}(n_i)}_{\text{sub-grid processes}}$$

$$i = 1, ..., 110$$

Ignoring sub-grid processes



Should we be using the same equations for any **spatial** scale?
Example:

$$\begin{cases} \dfrac{\partial n_1}{\partial t} = -\nabla \cdot (\boldsymbol{u} n_1) + k(T) n_1 n_2 \\ \dfrac{\partial n_2}{\partial t} = -\nabla \cdot (\boldsymbol{u} n_2) - k(T) n_1 n_2 \end{cases}$$

This problems have been addressed before in fluid dynamics
What are the right **averaged** equations?

Should we be using the **same equations** for any **spatial** scale?
Example:

Meteorology

$$\begin{cases} \dfrac{\partial n_1}{\partial t} = -\nabla \cdot (\boldsymbol{u} n_1) + k(T) n_1 n_2 \\ \dfrac{\partial n_2}{\partial t} = -\nabla \cdot (\boldsymbol{u} n_2) - k(T) n_1 n_2 \end{cases}$$

Chemistry

What are the right spatially averaged equations?
Write variables as (spatially):

$$n_1 = \overline{n_1} + n_1' \qquad \boldsymbol{u} = \overline{\boldsymbol{u}} + \boldsymbol{u}'$$

$$n_2 = \overline{n_2} + n_2' \qquad T = \overline{T} + T'$$

Should we be using the **same equations** for any scale?

$$\begin{cases} \dfrac{\partial n_1}{\partial t} = -\nabla \cdot (\boldsymbol{u} n_1) + k(T) n_1 n_2 \\ \dfrac{\partial n_2}{\partial t} = -\nabla \cdot (\boldsymbol{u} n_2) - k(T) n_1 n_2 \end{cases}$$

What are the right averaged equations?

Ignored

$$\begin{cases} \dfrac{\partial\, \overline{n_1}}{\partial t} = -\nabla \cdot (\overline{\boldsymbol{u}}\, \overline{n_1} + \overline{\boldsymbol{u}' n_1'}) + k(\overline{T}) \left(\overline{n_1}\, \overline{n_2} + \overline{n_1' n_2'} \right) \\ \dfrac{\partial\, \overline{n_2}}{\partial t} = -\nabla \cdot (\overline{\boldsymbol{u}}\, \overline{n_2} + \overline{\boldsymbol{u}' n_2'}) - k(\overline{T}) \left(\overline{n_1}\, \overline{n_2} + \overline{n_1' n_2'} \right) \end{cases}$$

At best parametrized as
Eddy diffusivity

Should we be using the **same equations** for any scale?

$$\begin{cases} \dfrac{\partial n_1}{\partial t} = -\nabla \cdot (\boldsymbol{u} n_1) + k(T) n_1 n_2 \\ \dfrac{\partial n_2}{\partial t} = -\nabla \cdot (\boldsymbol{u} n_2) - k(T) n_1 n_2 \end{cases}$$

Current models: use these spatially averaged equations

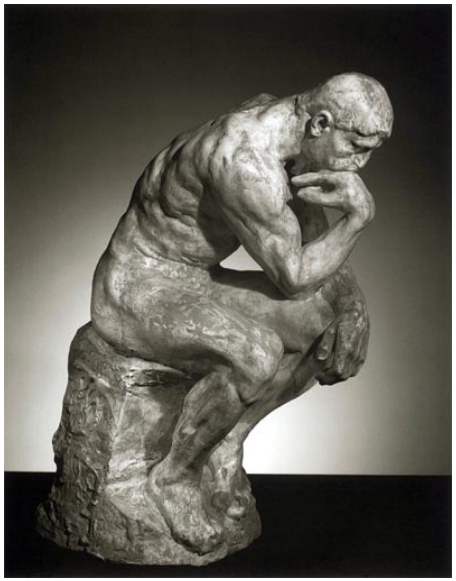
$$\begin{cases} \dfrac{\partial\, \overline{n_1}}{\partial t} = -\nabla \cdot (\overline{\boldsymbol{u}}\ \overline{n_1}) - k(\overline{T}) \overline{n_1}\ \overline{n_2} \\ \dfrac{\partial\, \overline{n_2}}{\partial t} = -\nabla \cdot (\overline{\boldsymbol{u}}\ \overline{n_2}) - k(\overline{T}) \overline{n_1}\ \overline{n_2} \end{cases}$$

Multi-scale analysis

Need closure parametrizations: kinetic theory?, **homogenization**?

Existing closure techniques in **combustion theory** need to be incorporated to atmospheric chemistry models

Thank you!



Questions?

Contact: **msantill@fas.harvard.edu**