ATMOSPHERIC TRANSPORT

Forces in the atmosphere and basic equation(s)

What should be considered to understand the role of transport on chemistry and aerosol ?

Timescales for chemistry and dynamics

Physical processes in an Eulerian model

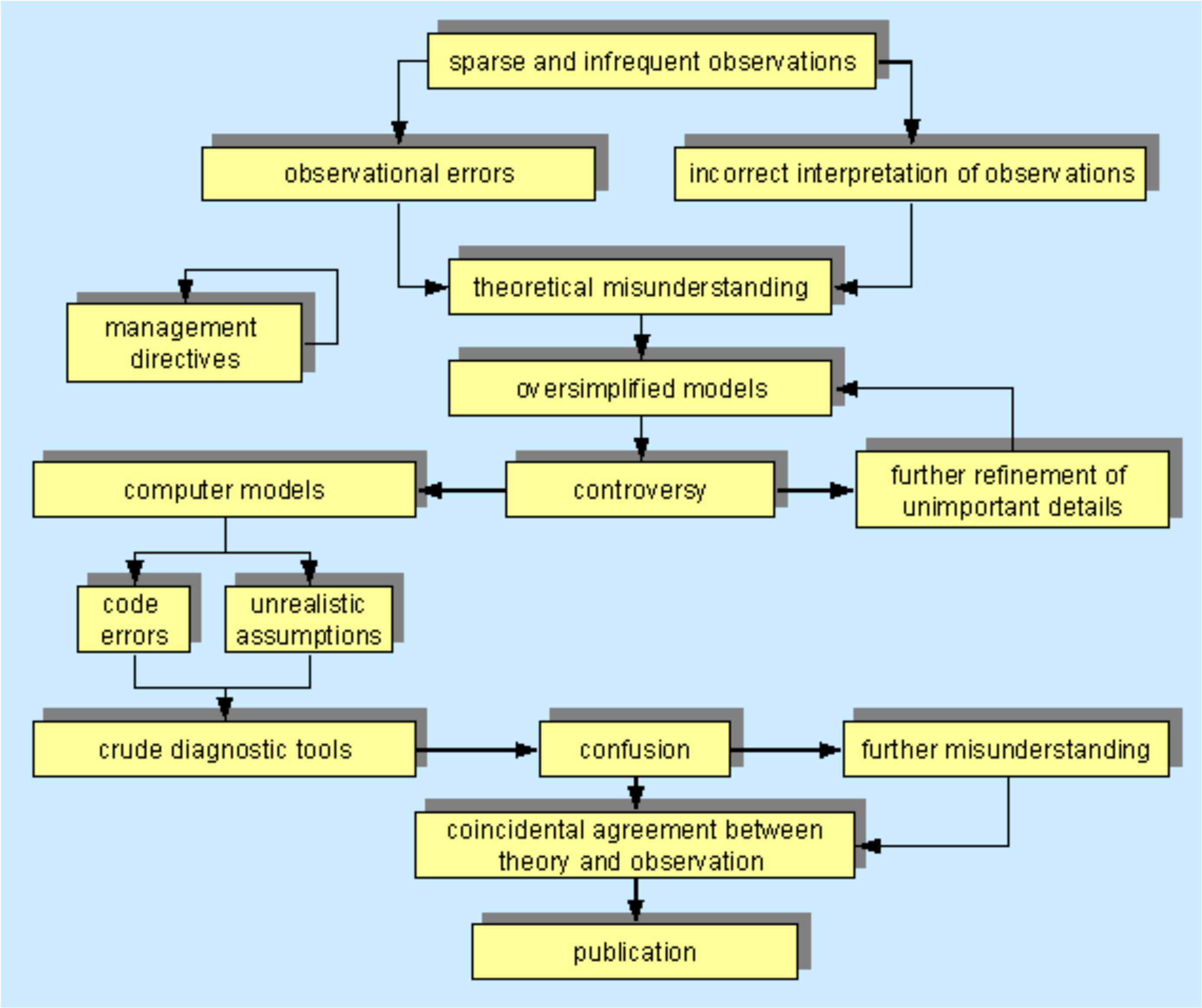
Lagrangian approach
What processes are included ? (Use the FLEXPART example)

Which information can be gathered ?

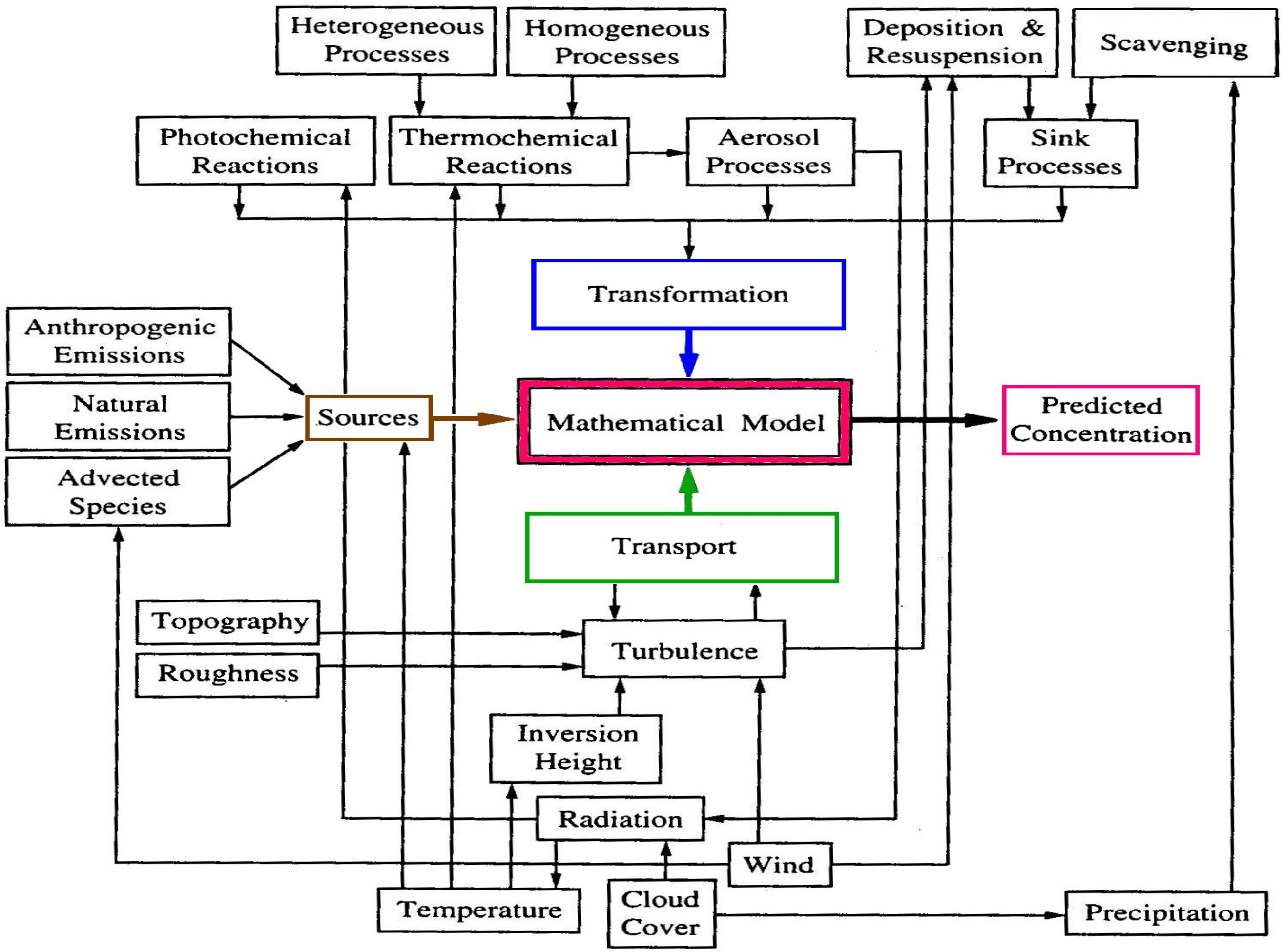
Federico Fierli – f.fierli@isac.cnr.it
Institute for Atmospheric Sciences and Climate, Italy
University of Rome (my course webpage: http://www.isac.cnr.it/~utls/?q=node/243)

Thanks to M. Jacob / M. Arnold for wonderful teaching material
http://acmg.seas.harvard.edu/people/faculty/djj/book/

Uliasz


sparse and infrequent observations
observational errors
incorrect interpretation of observations
theoretical misunderstanding
management directives
oversimplified models
computer models
controversy
further refinement of unimportant details
code errors
unrealistic assumptions
crude diagnostic tools
confusion
further misunderstanding
coincidental agreement between theory and observation
publication

Seinfeld and
Pandis
1998



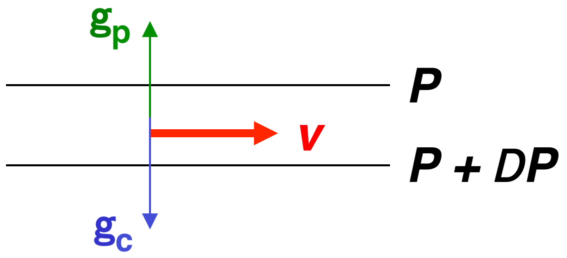
Forces in the atmosphere:

- Gravity $\mathbf{g}$
- Pressure-gradient $\boldsymbol{\gamma}_{\mathbf{p}} = -(1/\rho)\nabla P$
- Coriolis $\gamma_c = 2\omega v \sin\lambda$ to R of direction of motion (NH) or L (SH)
- Friction $\boldsymbol{\gamma}_{\mathbf{f}} = -k\mathbf{v}$

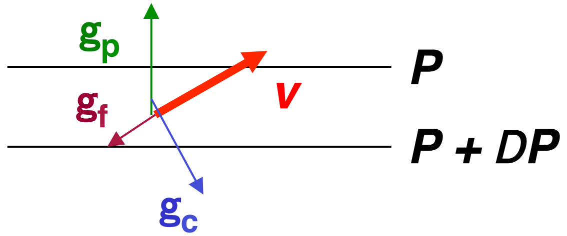
Equilibrium of forces:

In vertical: barometric law

In horizontal: *geostrophic* flow parallel to isobars



In horizontal, near surface: flow tilted to region of low pressure



$$d\overline{\mathbf{V}}/dt + fk \times \overline{\mathbf{V}} + \nabla\overline{\phi} = \mathbf{F}, \qquad \textit{(horizontal momentum)}$$

$$d\overline{T}/dt - \kappa\overline{T}\omega/p = Q/c_p, \qquad \textit{(thermodynamic energy)}$$

$$\nabla \cdot \overline{\mathbf{V}} + \partial\overline{\omega}/\partial p = 0, \qquad \textit{(mass continuity)}$$

$$\partial\overline{\phi}/\partial p + R\overline{T}/p = 0, \qquad \textit{(hydrostatic equilibrium)}$$

$$d\overline{q}/dt = S_q. \qquad \textit{(water vapor mass continuity)}$$

- Equation of motion F
- **Turbulent transport, generation and dissipation of momentum**
- Thermodynamic energy equation, Q
- **Sources, Sinks (radiation/convective-scale phase change)**
- Water vapour mass continuity S
- **Sources / sinks of water mass**

ZONAL GEOSTROPHIC FLOW AND THERMAL WIND RELATION

$\Phi = gz$ geopotential height

λ = latitude

a = Earth radius

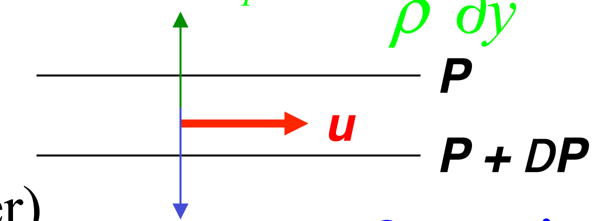
ω = angular vel. of Earth

$f = 2\omega \sin\lambda$ (Coriolis parameter)

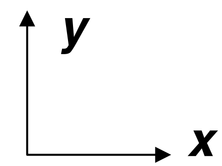
$z_* = -H \ln(p / p_o)$ log-P coordinate

$H = \dfrac{RT_o}{Mg}$ scale height

$$\gamma_p = -\frac{1}{\rho}\frac{\partial P}{\partial y} = -\frac{\partial \Phi}{\partial y} = -\frac{1}{a}\frac{\partial \Phi}{\partial \lambda}$$



$$\gamma_c = 2\omega u \sin\lambda = fu$$



Geostrophic balance:

$$fu = -\frac{1}{a}\frac{\partial \Phi}{\partial \lambda}$$

Thermal wind relation:

$$f\frac{\partial u}{\partial z_*} = -\frac{R}{aH}\frac{\partial T}{\partial \lambda}$$

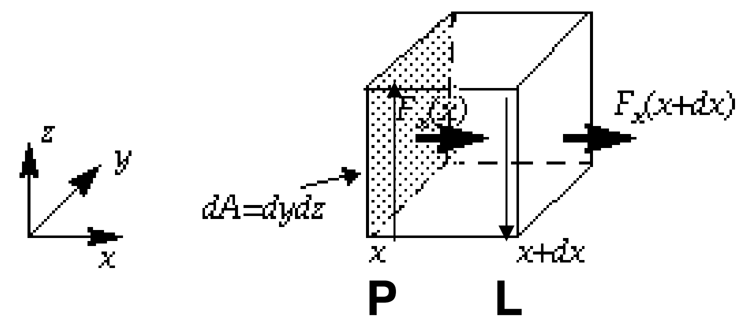
Transport

$$\frac{\partial n}{\partial t} = -\frac{\partial F_x}{\partial x} - \frac{\partial F_y}{\partial y} - \frac{\partial F_z}{\partial z} + P - L = -\nabla \cdot \boldsymbol{F} + P - L$$

N=number of molecules of …

F=nUdxdy / dxdy (normalized)

$$\frac{\partial n}{\partial t} = -\nabla \bullet (n\boldsymbol{U}) + P - L$$



$$\frac{\partial}{\partial t}\bar{n} = -\nabla \bullet (\overline{\boldsymbol{F}_A} + \overline{\boldsymbol{F}_T}) + \bar{P} - \bar{L}$$

$$= -\nabla \bullet (\bar{n}\overline{\boldsymbol{U}}) + \nabla \bullet (\boldsymbol{K} n_a \nabla \bullet \bar{C}) + \bar{P} - \bar{L}$$

Fick's law: F_T=-nD grad(n)

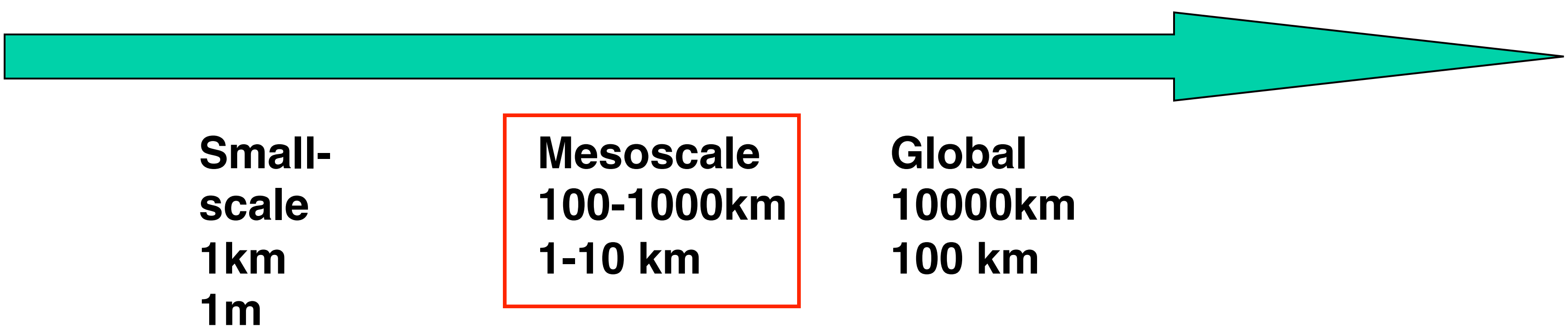
A question of scales

Impossible to explicitly resolve all physical processes

Necessary to parametrize it function of model variables

The parametrization depends on the spatial and temporal scales

Issue: complexity and computing time !



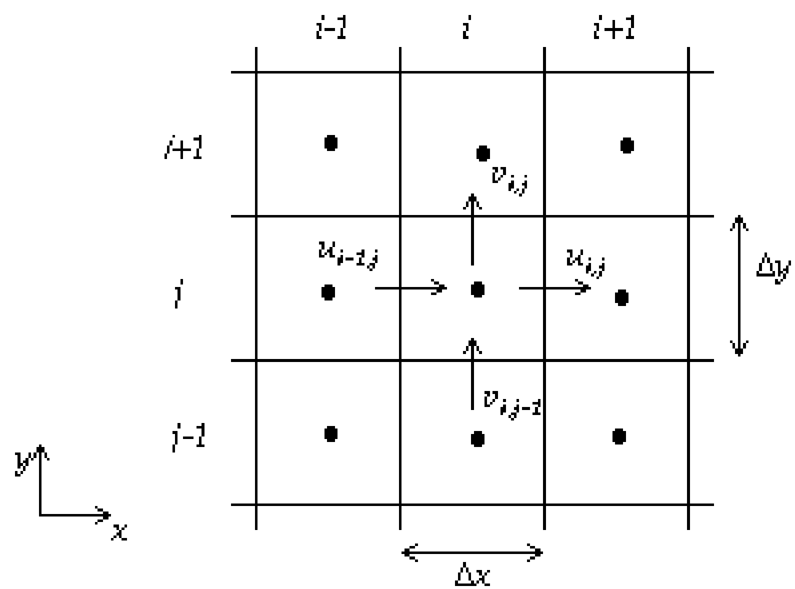
Discretization

$$A \cdot n(\mathbf{X}, t_o) = n(\mathbf{X}, t_o) + \int_{t}^{t_o+\Delta t} \left[\frac{\partial n}{\partial t}\right]_{advection} dt$$

$$n(\mathbf{X}, t_o + \Delta t) = C \cdot T \cdot A \cdot n(\mathbf{X}, t_o)$$

A= advection
T= turbulent transport
C= Chemistry

$$\begin{aligned} n(i, j, k, t_o + \Delta t) &= n(i, j, k, t_o) \\ &+ \frac{u(i-1, j, k, t_o)n(i-1, j, k, t_o) - u(i, j, k, t_o)n(i, j, k, t_o)}{\Delta x}\Delta t \\ &+ \frac{v(i, j-1, k, t_o)n(i, j-1, k, t_o) - v(i, j, k, t_o)n(i, j, k, t_o)}{\Delta y}\Delta t \\ &+ \frac{w(i, j, k-1, t_o)n(i, j, k-1, t_o) - w(i, j, k, t_o)n(i, j, k, t_o)}{\Delta z}\Delta t \end{aligned}$$



Hypothesis:
C, T, A can be separated

The atmospheric evolution of a species X is given by the *continuity equation*

$$\frac{\partial[X]}{\partial t} = E_X - \nabla \bullet (\mathbf{U}[X]) + P_X - L_X - D_X$$

- local change in concentration with time
- emission
- transport (flux divergence; U is wind vector)
- chemical production and loss (depends on concentrations of other species)
- deposition

This equation cannot be solved exactly ⇨ need to construct *model* (simplified representation of complex system)

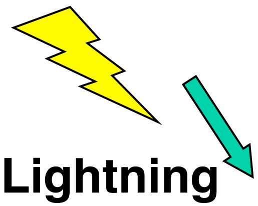
Define problem of interest → Design model; make assumptions needed to simplify equations and make them solvable → Design observational system to test model → Evaluate model with observations

Improve model, characterize its error (Evaluate model with observations → Design model)

Design model → Apply model: make hypotheses, predictions

HOW TO MODEL ATMOSPHERIC COMPOSITION?

Solve continuity equation for chemical mixing ratios $C_i(\mathbf{x}, t)$



Transport

Chemistry
Aerosol microphysics

Eulerian form:

$$\frac{\partial C_i}{\partial t} = -\mathbf{U} \bullet \nabla C_i + P_i - L_i$$

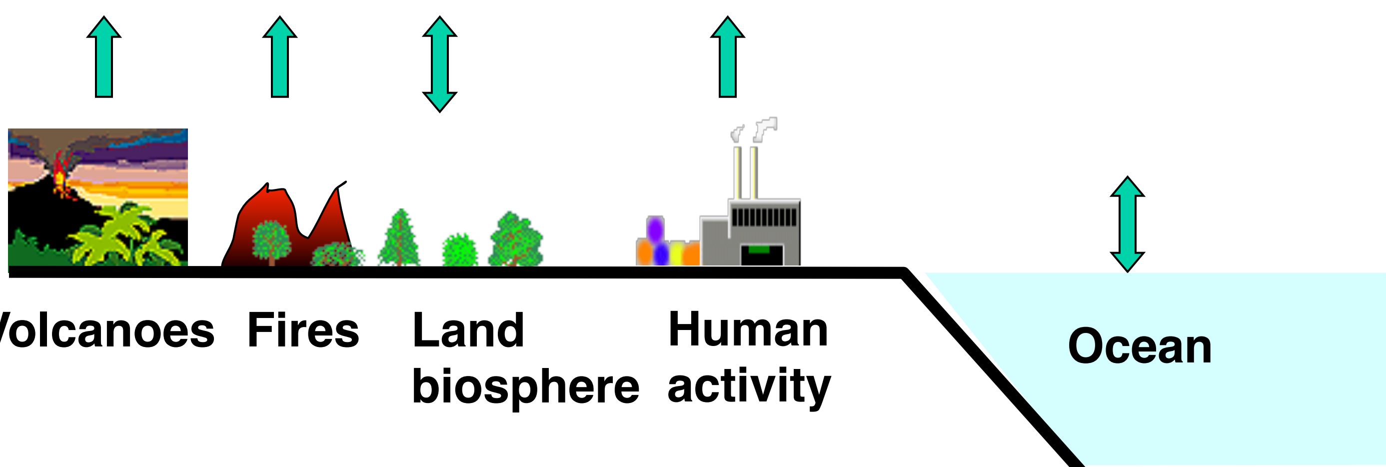
Lagrangian form:

$$\frac{dC_i}{dt} = P_i - L_i$$

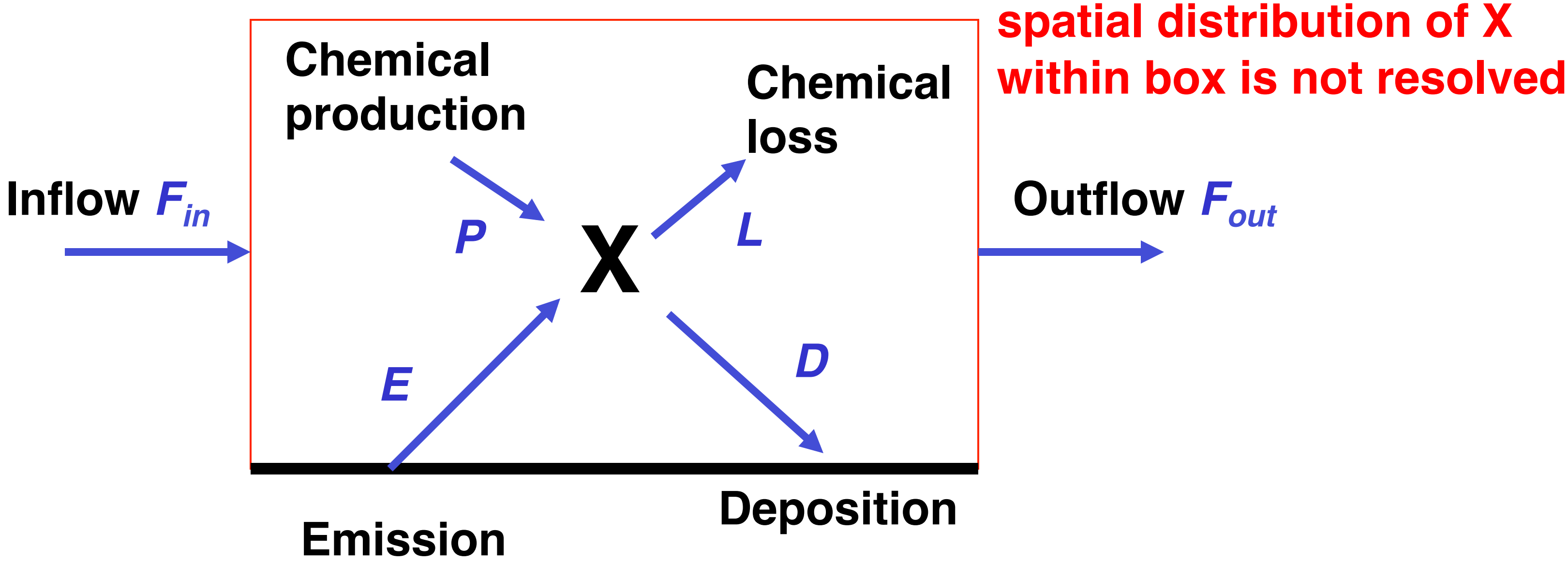
$\mathbf{U}$ = wind vector

P_i = local source of chemical *i*

L_i = local sink



ONE-BOX MODEL



Mass balance equation: $$\frac{dm}{dt} = \sum \text{sources} - \sum \text{sinks} = F_{in} + E + P - F_{out} - L - D$$

Atmospheric lifetime: $$\tau = \frac{m}{F_{\text{out}} + L + D}$$

Fraction lost by export: $$f = \frac{F_{out}}{F_{out} + L + D}$$

Lifetimes add in parallel: $$\frac{1}{\tau} = \frac{F_{out}}{m} + \frac{L}{m} + \frac{D}{m} = \frac{1}{\tau_{\text{export}}} + \frac{1}{\tau_{\text{chem}}} + \frac{1}{\tau_{\text{dep}}}$$

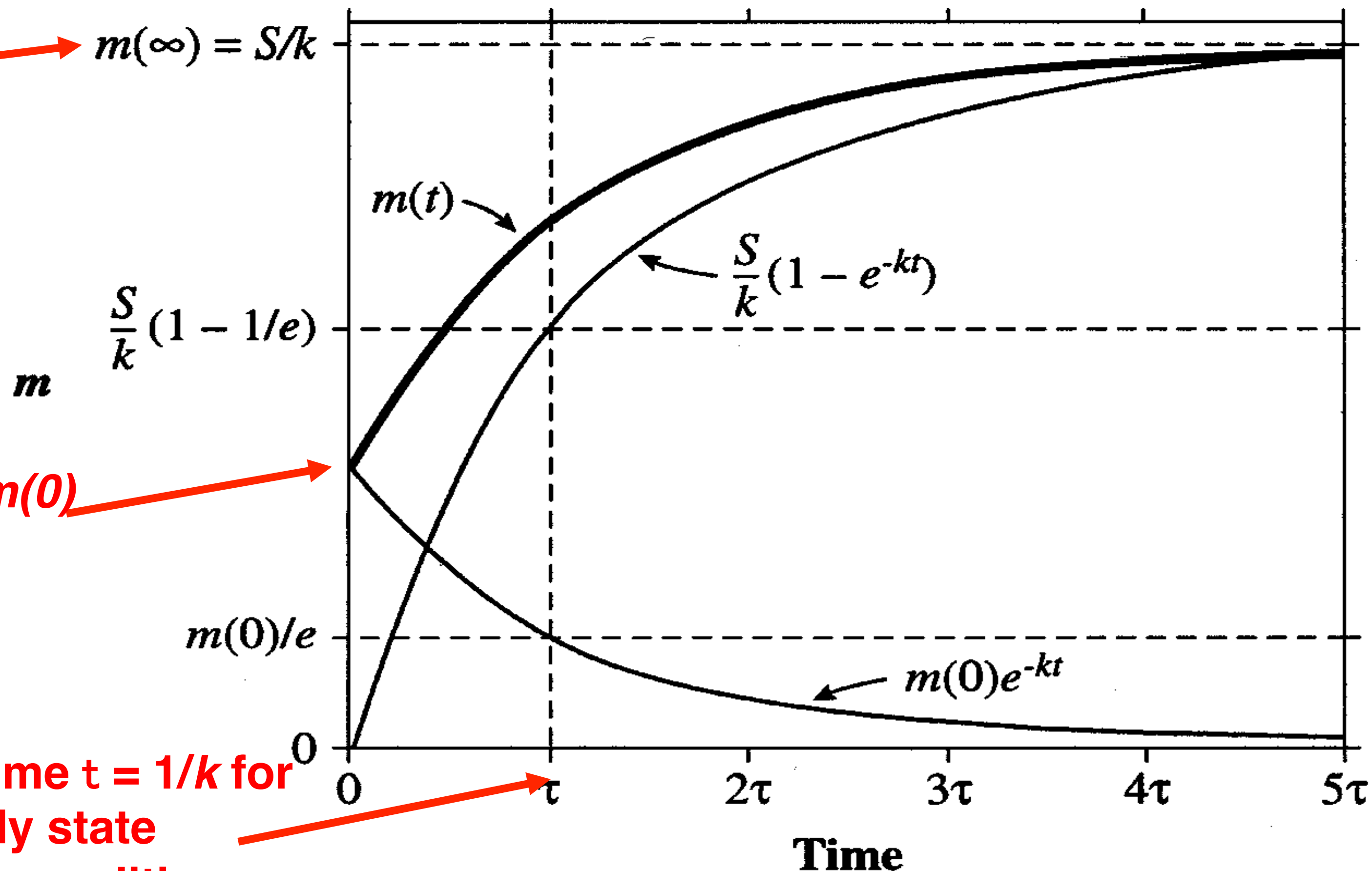
Loss rate constants add in series: $$k = \frac{1}{\tau} = k_{\text{export}} + k_{\text{chem}} + k_{\text{dep}}$$

SPECIAL CASE:
SPECIES WITH CONSTANT SOURCE, 1st ORDER SINK

$$\frac{dm}{dt} = S - km \quad \Rightarrow \quad m(t) = m(0)e^{-kt} + \frac{S}{k}(1 - e^{-kt})$$

Steady state solution (*dm/dt = 0)*

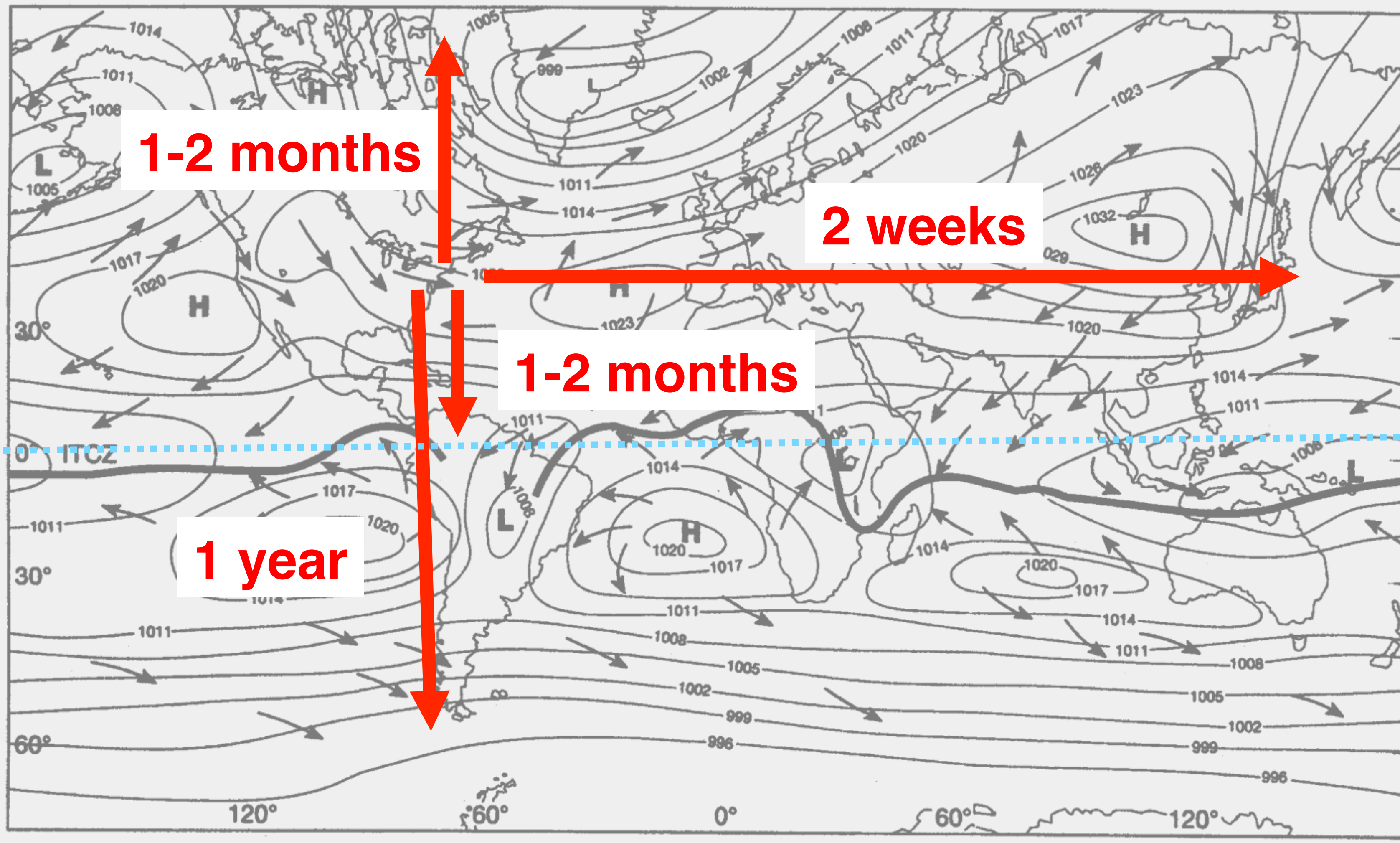
Initial condition *m(0)*



Characteristic time t = 1/*k* for
- **reaching steady state**
- **decay of initial condition**

If *S, k* are constant over *t* >> *t*, then *dm/dt* → 0 and *m*→ *S/k:* quasi steady state

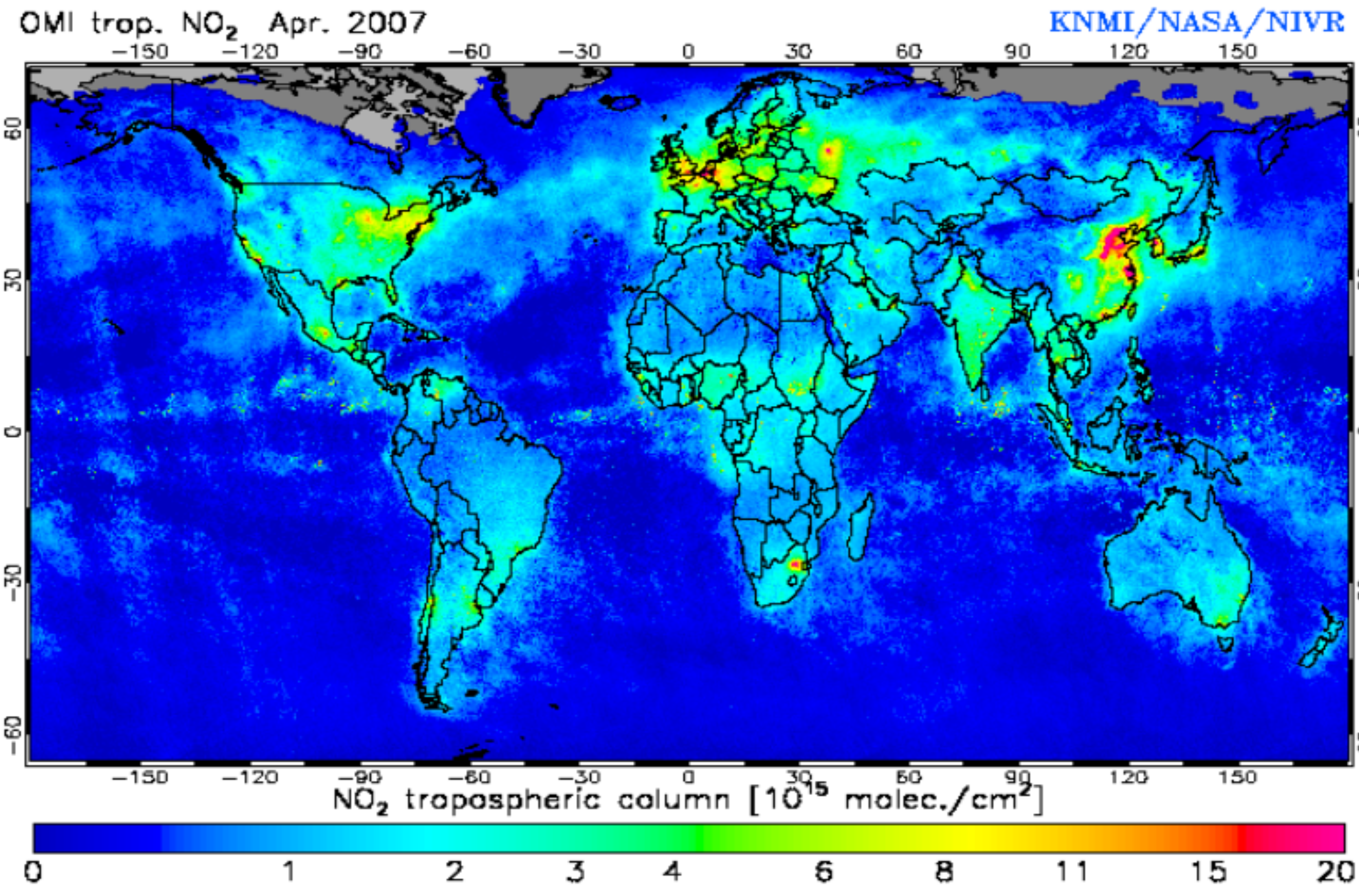
TIME SCALES FOR HORIZONTAL TRANSPORT (TROPOSPHERE)



(a) January

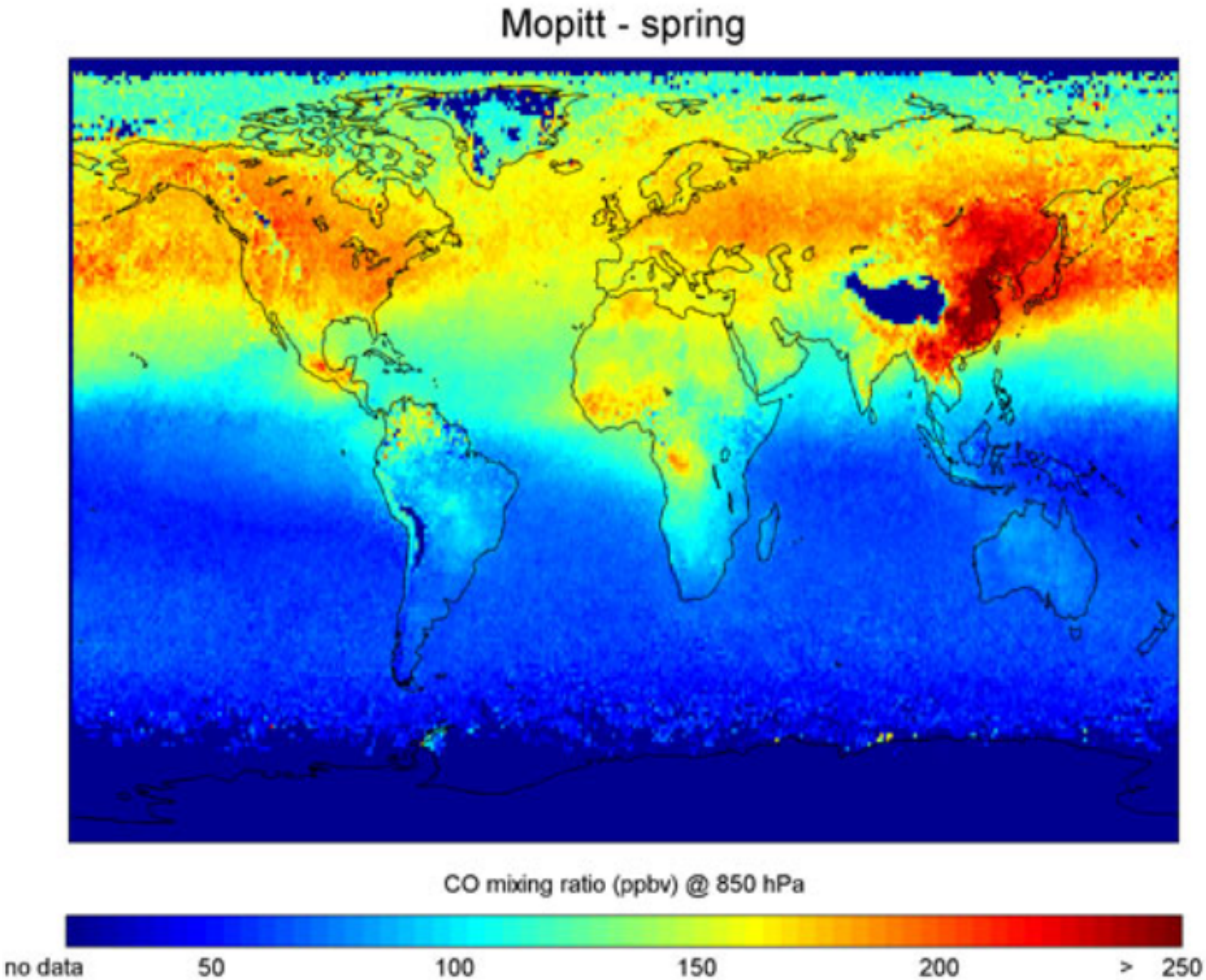
NO_2 emitted by combustion, has atmospheric lifetime ~ 1 day: strong gradients away from source regions

Satellite observations of NO_2 columns



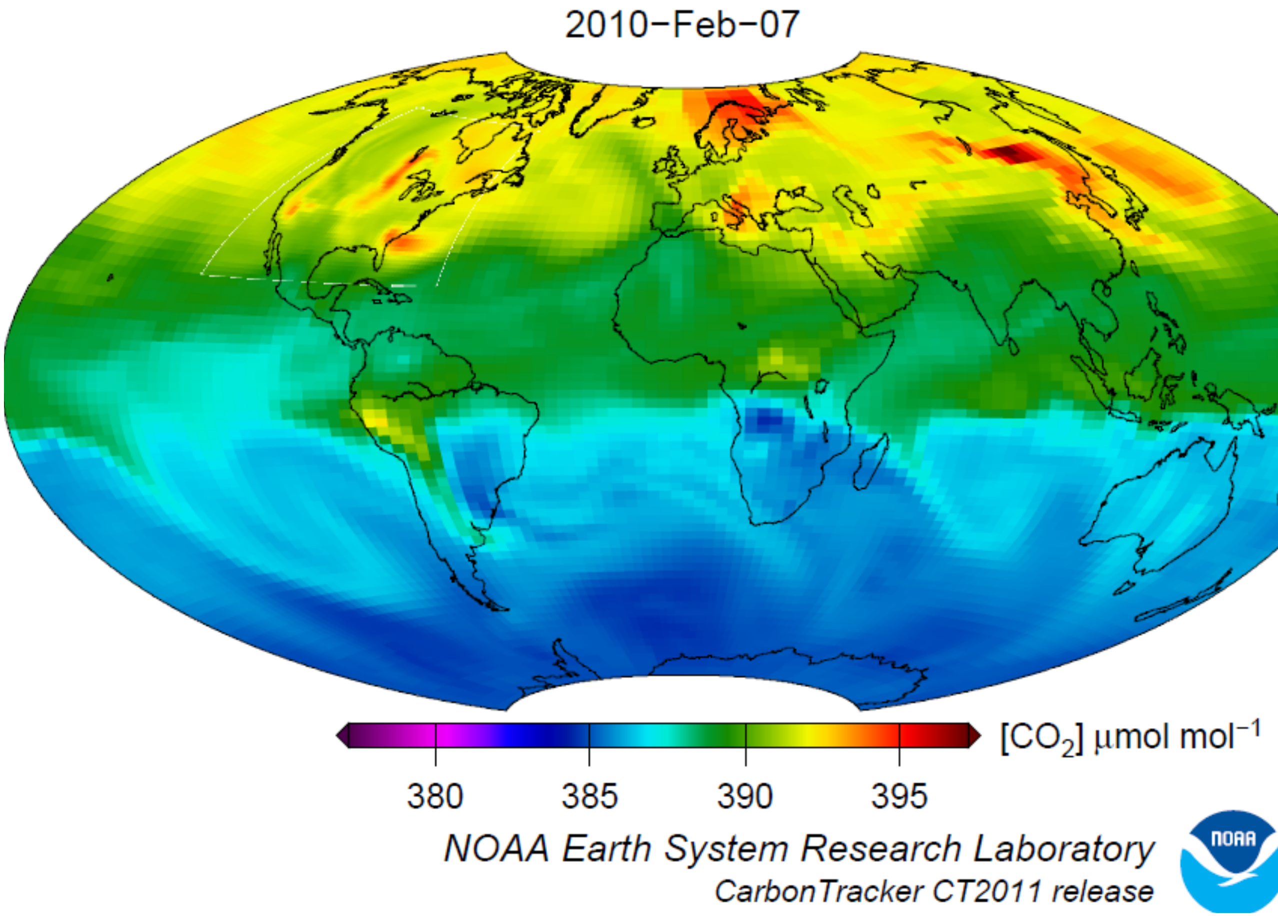
CO emitted by combustion, has atmospheric lifetime ~ 2 months: mixing around latitude bands

Satellite observations



CO_2 emitted by combustion, has atmospheric lifetime ~ 100 years: global mixing

Assimilated observations



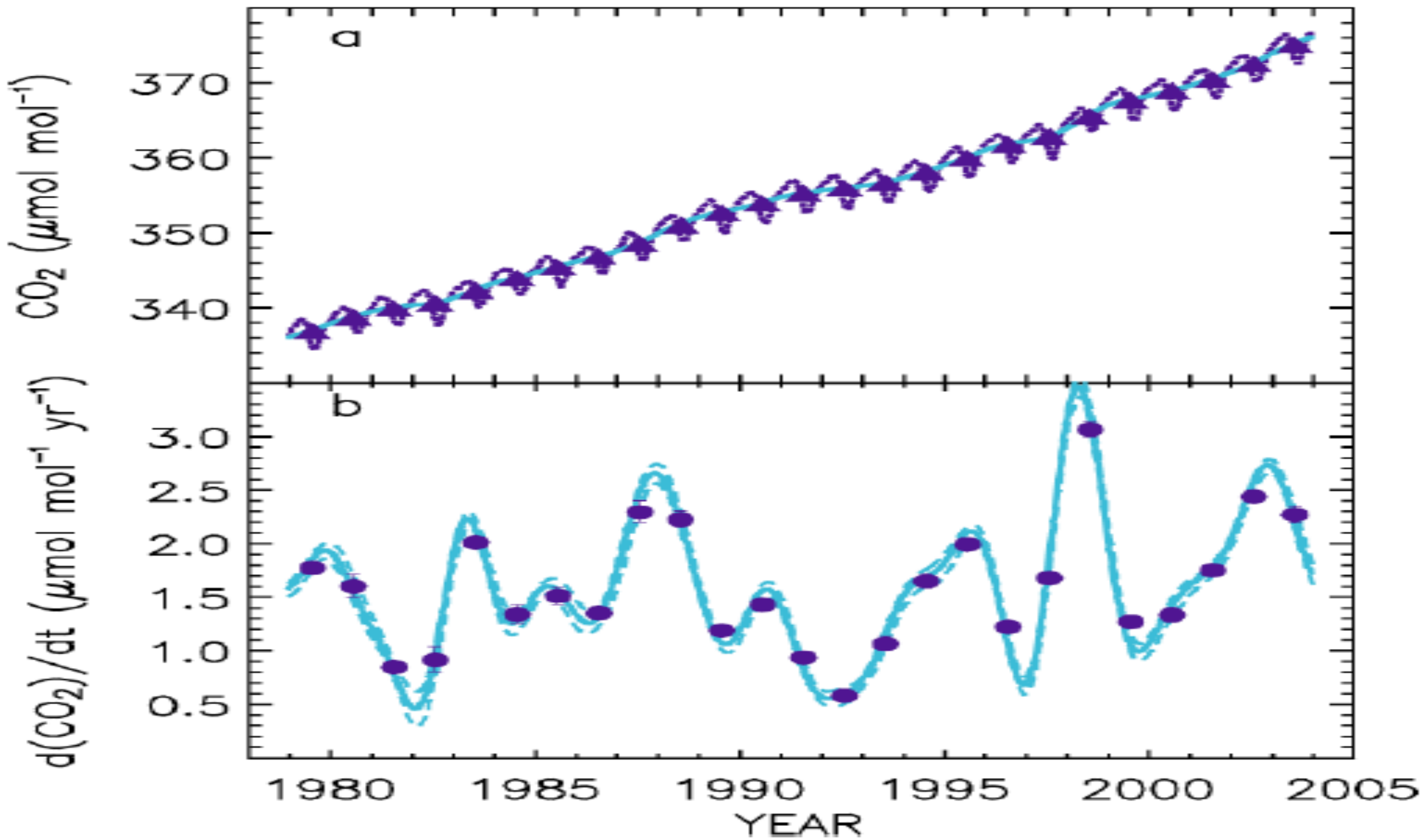
GLOBAL BOX MODEL FOR CO_2 (Pg C yr^{-1})

	1980s	1990s
Atmospheric increase	3.3 ± 0.1	3.2 ± 0.1
Emissions (fossil fuel, cement)	5.4 ± 0.3	6.3 ± 0.4
Ocean-atmosphere flux	−1.9 ± 0.6	−1.7 ± 0.5
Land-atmosphere flux*	−0.2 ± 0.7	−1.4 ± 0.7
**partitioned as follows:*		
Land-use change	*1.7 (0.6 to 2.5)*	*NA*
Residual terrestrial sink	*−1.9 (−3.8 to 0.3)*	*NA*

IPCC [2001]

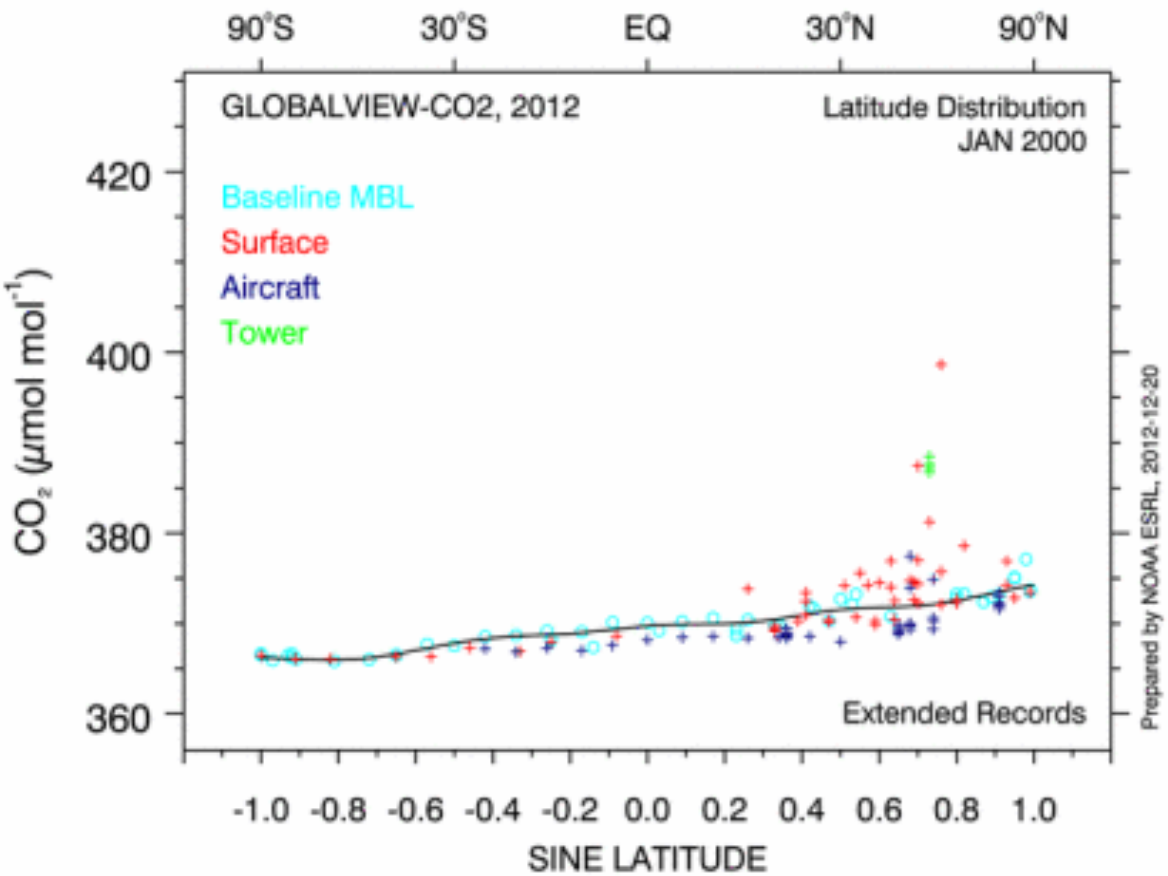
ATMOSPHERIC CO_2 TREND OVER PAST 25 YEARS

IPCC [2007]



mmol mol^{-1} is the proper SI unit; ppm, ppmv are customary units

LATITUDINAL GRADIENT OF CO_2 , 2000-2012

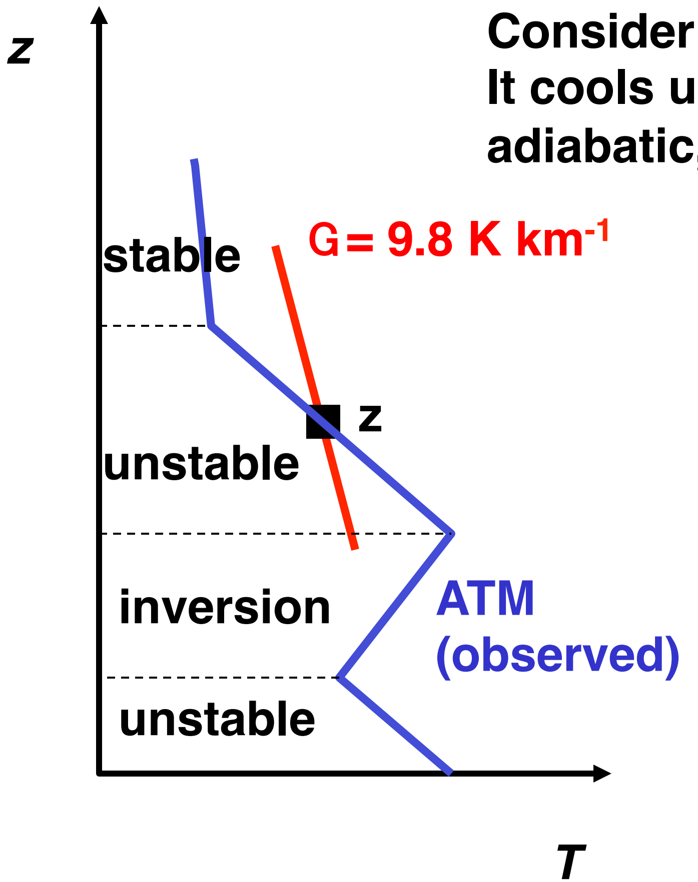


Illustrates long time scale for interhemispheric exchange;
use 2-box model to constrain CO_2 sources/sinks in each hemisphere

http://www.esrl.noaa.gov/gmd/ccgg/globalview/

ATMOSPHERIC LAPSE RATE AND STABILITY

"Lapse rate" = -*dT/dz*



Consider an air parcel at *z* lifted to *z+dz* and released. It cools upon lifting (expansion). Assuming lifting to be adiabatic, the cooling follows the adiabatic lapse rate G :

$$\Gamma = -dT/dz = \frac{g}{C_p} = 9.8 \text{ K km}^{-1}$$

What happens following release depends on the local lapse rate $-dT_{ATM}/dz$:

- $-dT_{ATM}/dz > G$ ⇨ upward buoyancy amplifies initial perturbation: atmosphere is *unstable*
- $-dT_{ATM}/dz = G$ ⇨ zero buoyancy does not alter perturbation: atmosphere is *neutral*
- $-dT_{ATM}/dz < G$ ⇨ downward buoyancy relaxes initial perturbation: atmosphere is *stable*
- $dT_{ATM}/dz > 0$ ("inversion"): very stable

The stability of the atmosphere against vertical mixing is solely determined by its lapse rate.

TYPICAL TIME SCALES FOR VERTICAL MIXING

$$\text{Turbulent flux} = -K_z n_a \frac{\partial < C >}{\partial z}$$

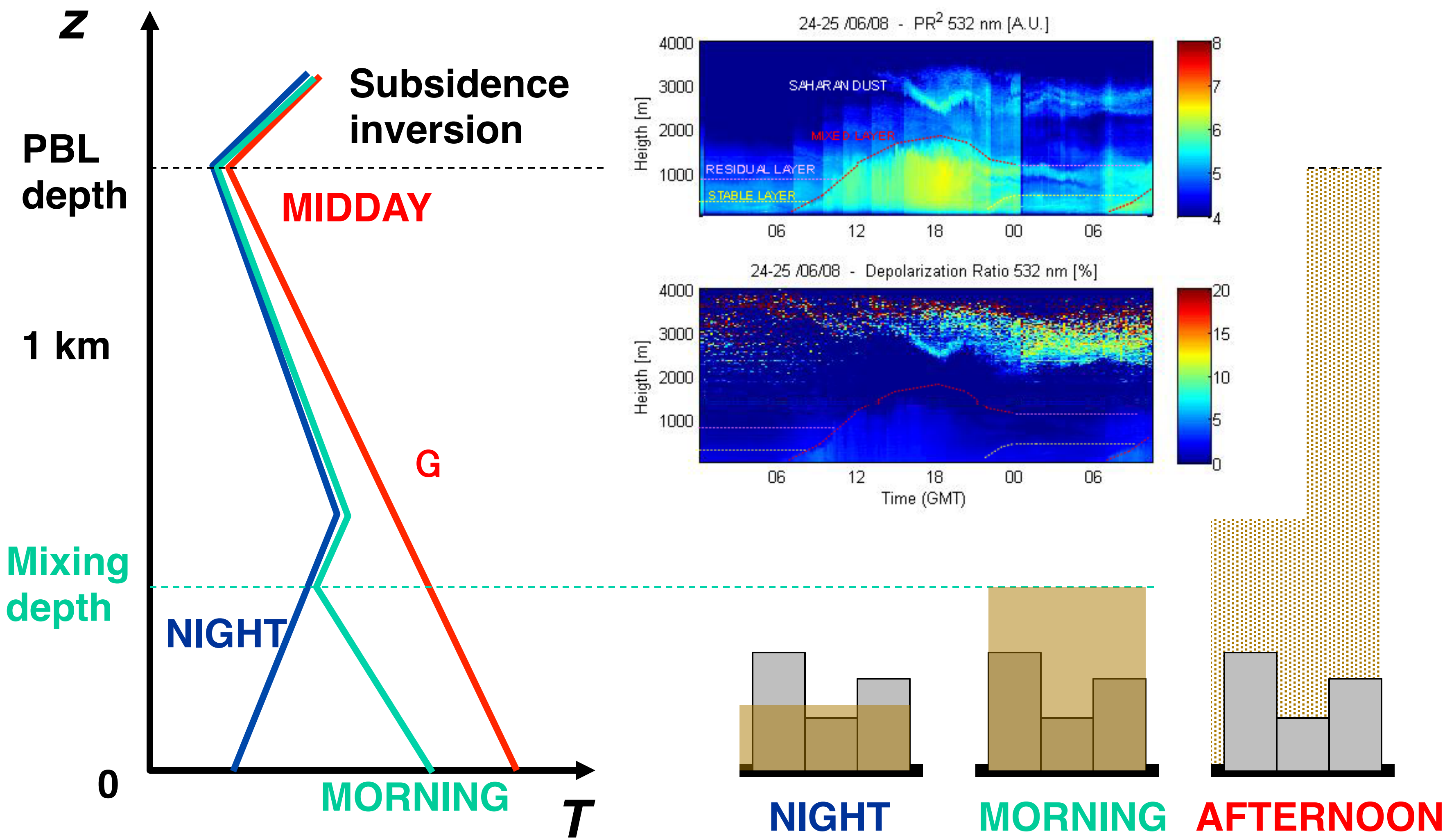
• **Typical values of K_z: 10^2 cm^{2}s^{-1} (very stable) to 10^7 cm^2 s^{-1} (very unstable); mean value for troposphere is ~ 10^5 cm^2 s^{-1}**

• **Same parameterization (with different K_x, K_y) is also applicable in horizontal direction but is less important (mean winds are stronger)**

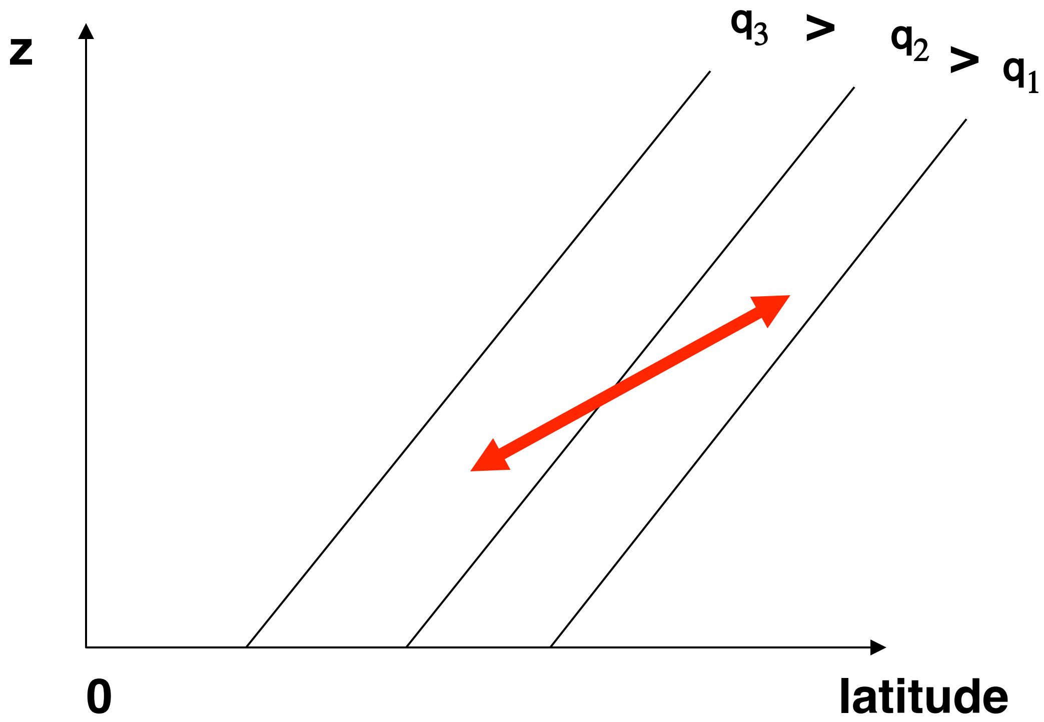
- **Estimate time D*t* to travel D*z***
- **by turbulent diffusion:**

$$\Delta t = \frac{(\Delta z)^2}{2K_z} \quad \text{with } K_z : 10^5 \text{cm}^2\text{s}^{-1}$$

DIURNAL CYCLE OF SURFACE HEATING/COOLING: ventilation of urban pollution



BAROCLINIC INSTABILITY



Buoyant vertical motion Is possible even when

$$\partial\theta / \partial z > 0$$

Dominant mechanism for vertical motion in extratropics

LATITUDINAL STRUCTURE OF TROPOPAUSE REGION

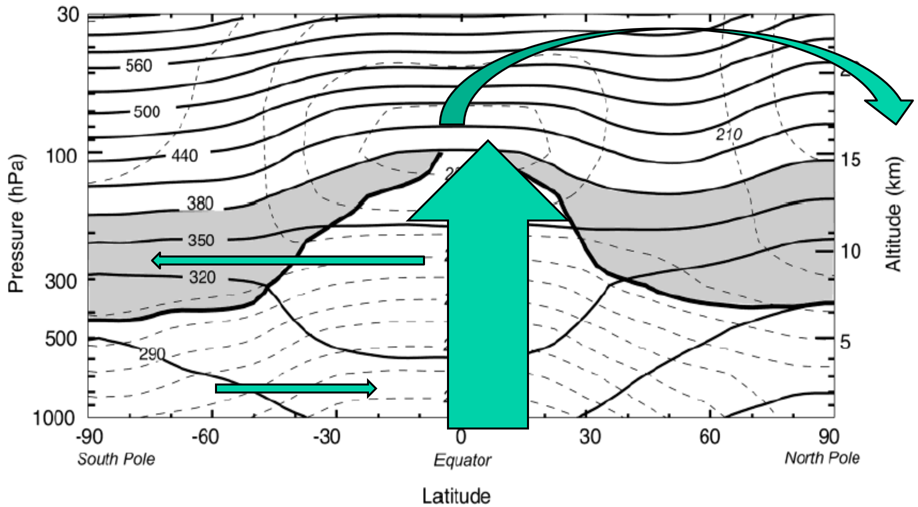
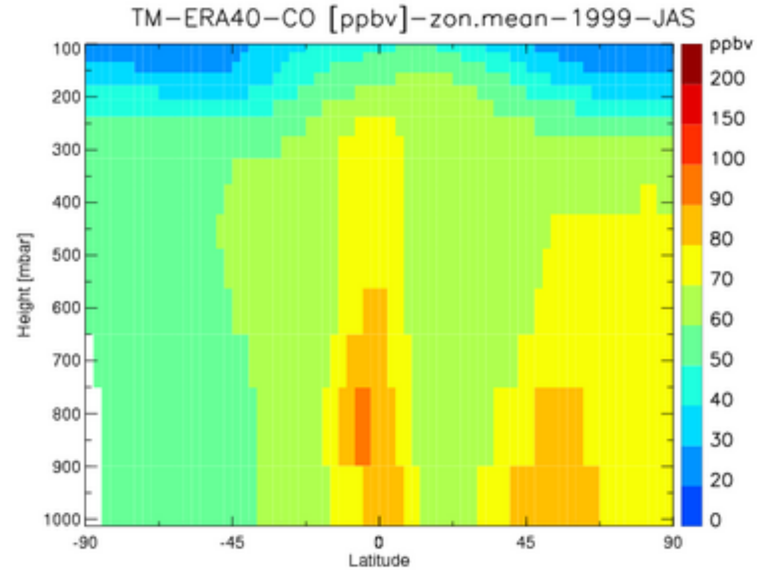
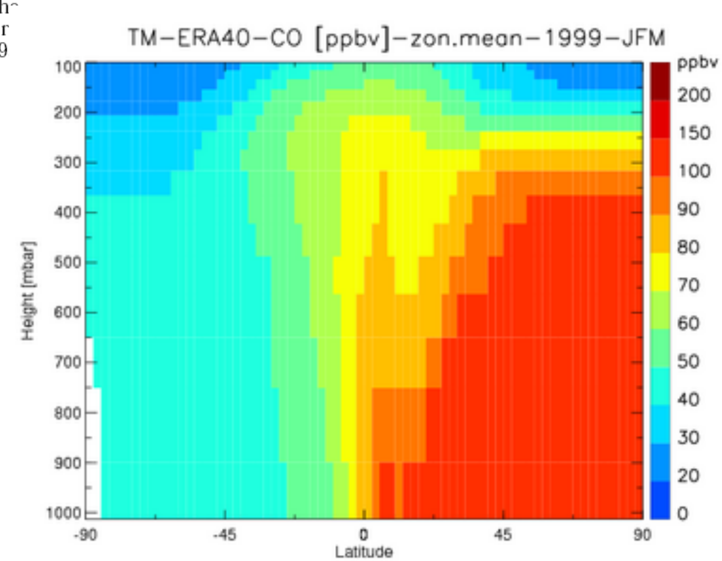
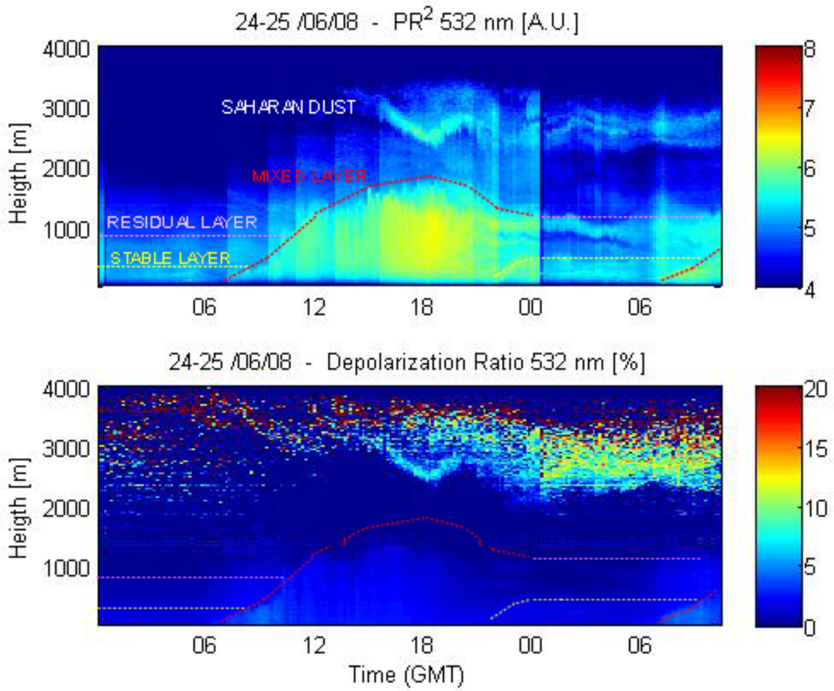


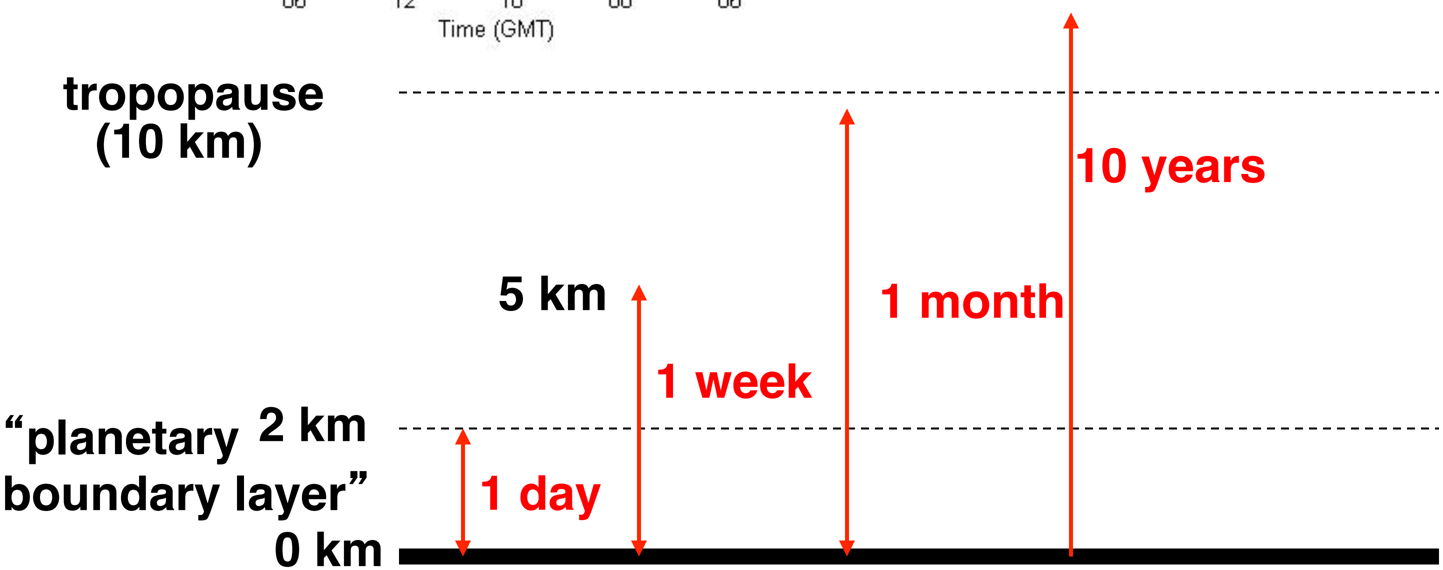
Figure 4. Annual and zonal mean distribution of potential temperature (solid) and temperature (dashed), in Kelvi thick line denotes the lapse-rate tropopause. Features to note are the weak stratification in the troposphere (and strong lapse rate, close to moist adiabatic), the strong stratification in the stratosphere, and the temperature minimum at th tropical tropopause. The shaded regions denote the "lowermost stratosphere", consisting of that part of the stratospher that is connected to the troposphere along isentropic surfaces. (Reprinted with permission from ref 4. Copyright 199 American Geophysical Union.)



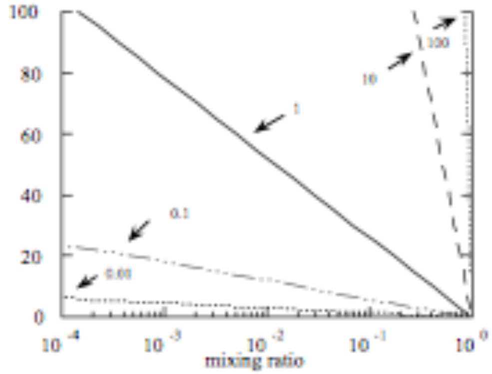


TYPICAL TIME SCALES FOR VERTICAL MIXING





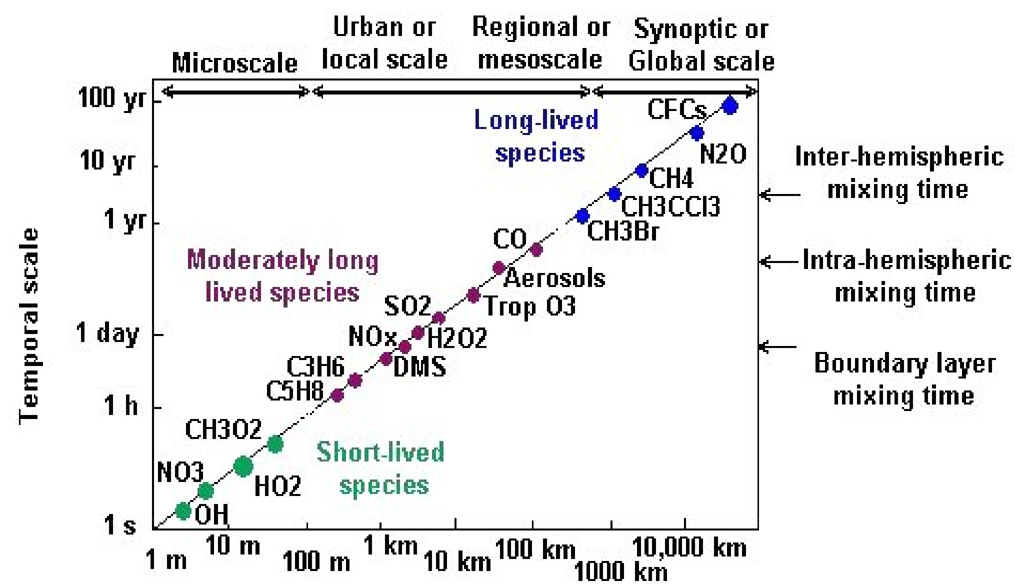
Chemical vs. transport lifetime



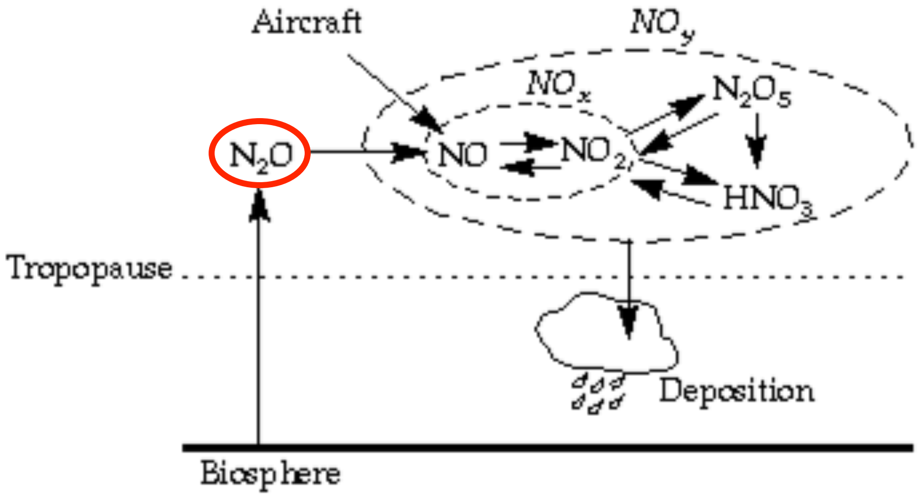
- When the chemical lifetime is 100 times larger than the dynamical lifetime, materials will have an almost constant mixing ratio to nearly 100 km altitude.
- However, when the chemical lifetime is 1% of the dynamical lifetime the mixing ratio falls very rapidly in the troposphere.

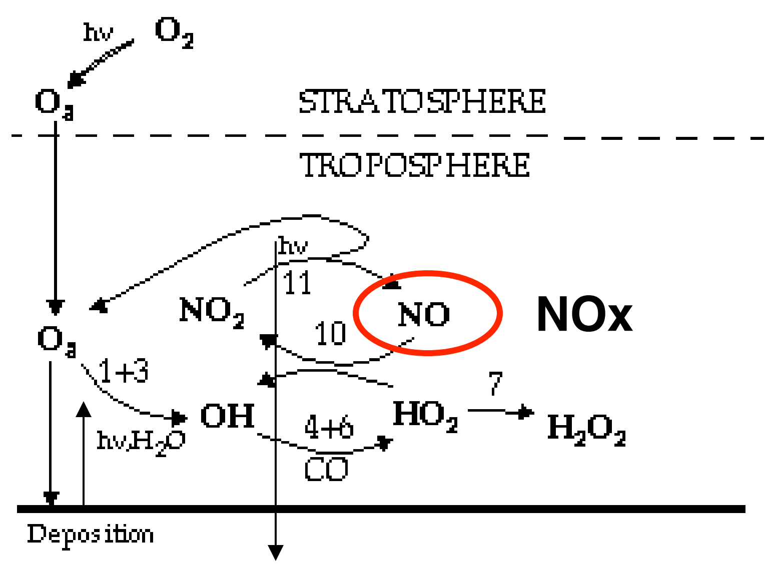
Lifetimes of some interesting materials

Material	M_w, Abundance (Tg)	P_w, Source (Tg/yr)	t_w, Lifetime (yr)
H_2O	1.3×10^7	5×10^8	0.025
CH_4	5×10^3	515	10
COS	5.2	1.2	4.3
SO_2	0.6-0.9	200	.003-.005
N_2O	2.5×10^3	12-21	120
CFC-11	8.2	0.25	50
CFC-12	10.3	0.37	102
CH_3Cl	5	3.5	1.5
NaCl	3.8	1300	0.003



From Toon et al.

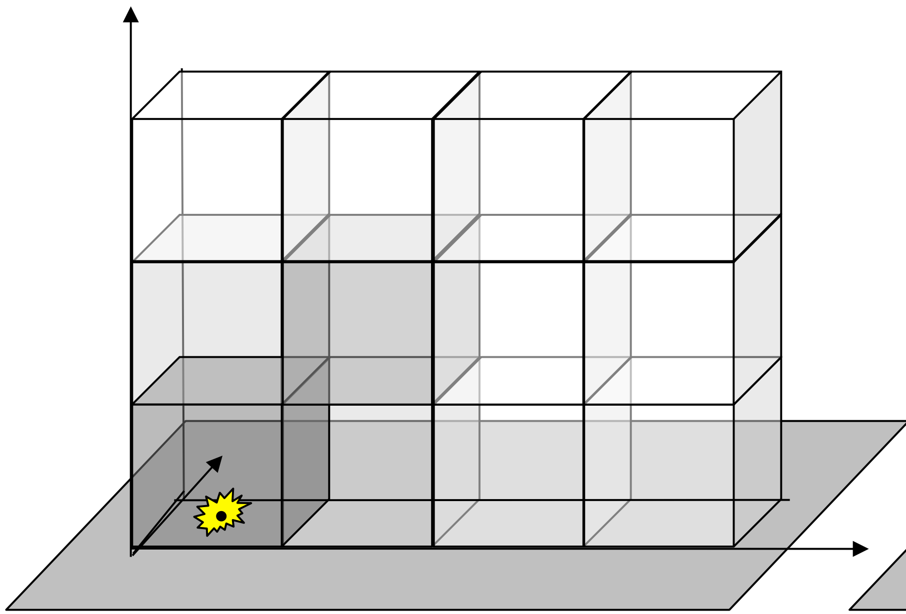


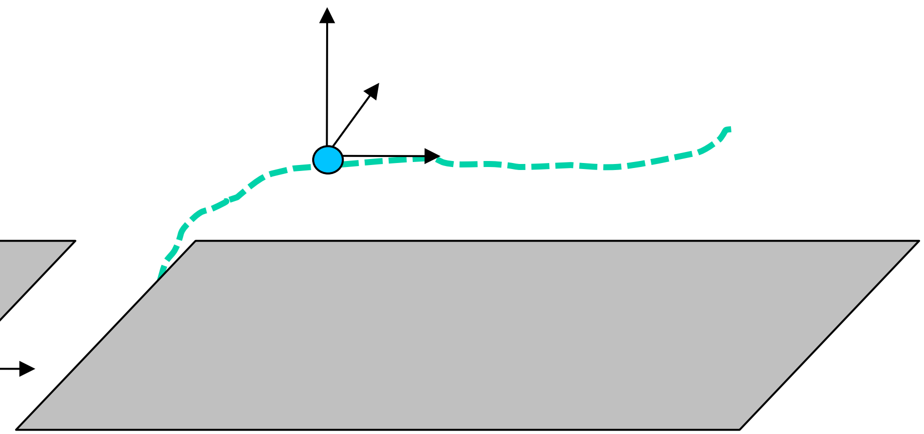


Example:
NO cycle
Troposphere

Eulerian

Lagrangian





$$\frac{\partial c_i}{\partial t} + u_x \frac{\partial c_i}{\partial x} + u_y \frac{\partial c_i}{\partial y} + u_z \frac{\partial c_i}{\partial z} = \frac{\partial}{\partial x}\left(K_{xx}\frac{\partial c_i}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_{yy}\frac{\partial c_i}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_{zz}\frac{\partial c_i}{\partial z}\right) + R_i(c_1, c_2 \cdots, c_n) + E_i(x,y,z,t) - S_i(x,y,z,t)$$

Divergence of the advected flux

Chemical reactions

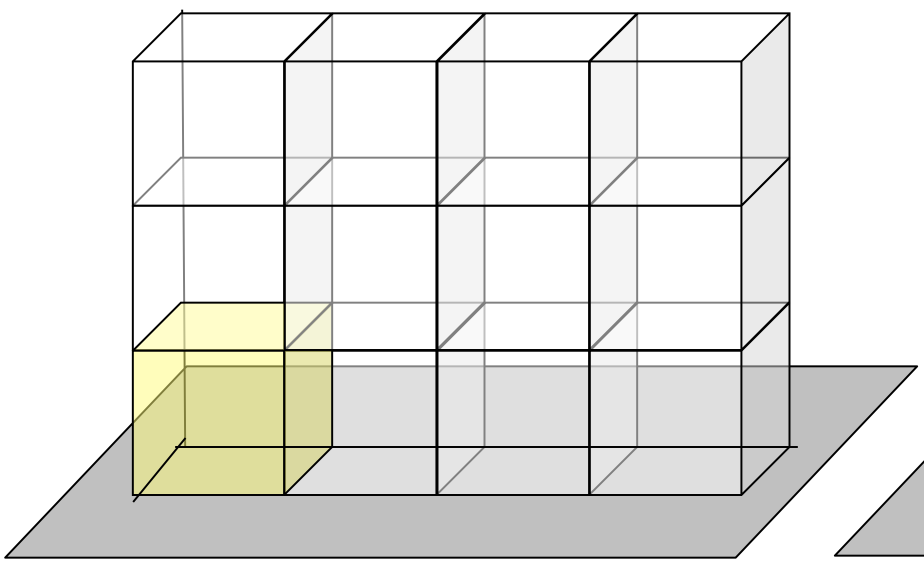
Sinks

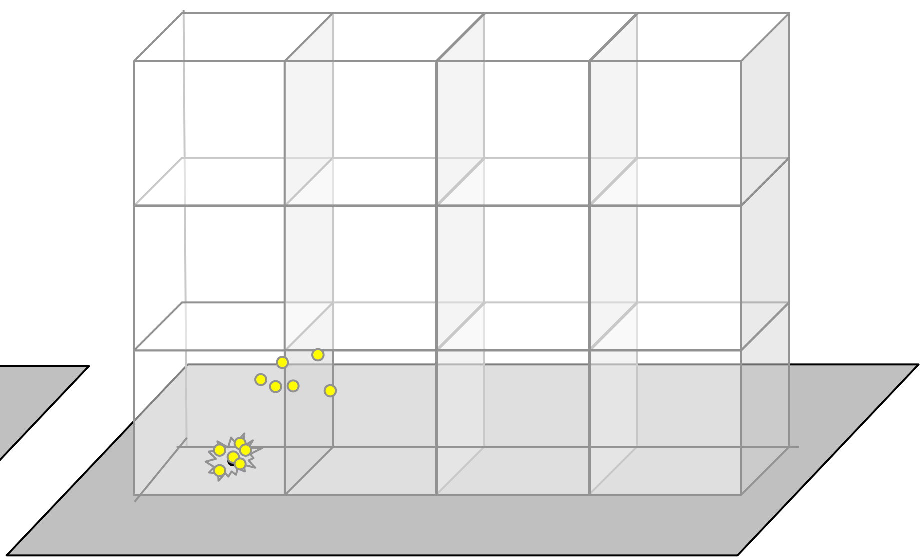
Divergence of the turbulent fluxes

Emissions

Eulerian

Lagrangian





Immediate dilution in the grid cell

Point source sub-model then needed

LPDM can deal naturally with point sources

The grid is only applied to output fields

Eulerian

Fig. 10a

Initial isolated puff

MSC-W Note 2/92, August 1992.EMEP "An Evaluation of Eulerian Advection Methods for the Modelling of Long Range Transport of Air Pollution". By Erik Berge and Leonor Tarrasón. EMEP_1992_N2.pdf

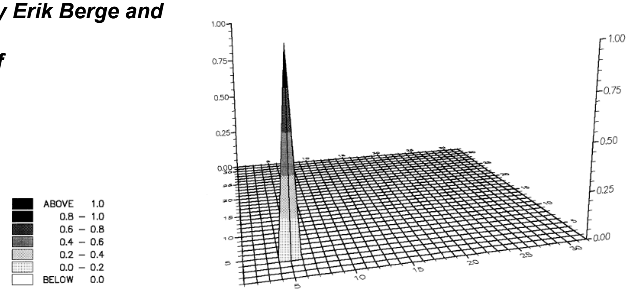


Fig. 10d

BOS: Diagonal puff

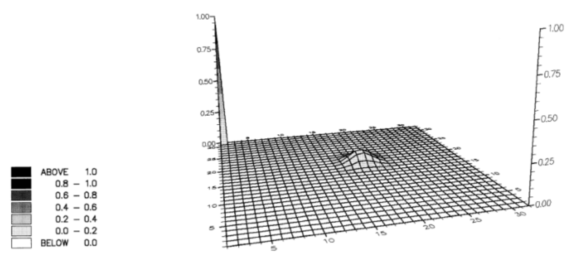
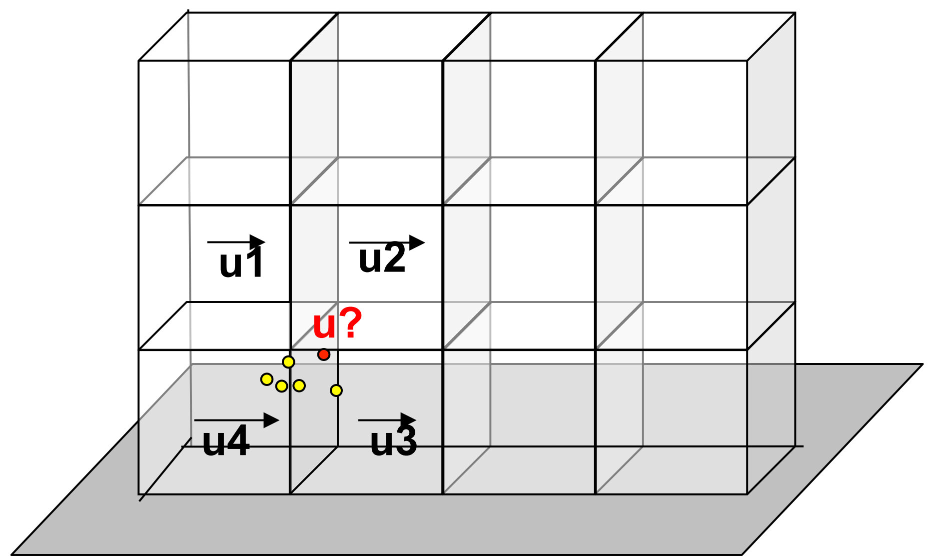


Fig. 10e

PSS: Diagonal puff

Numerical diffusion in the advection

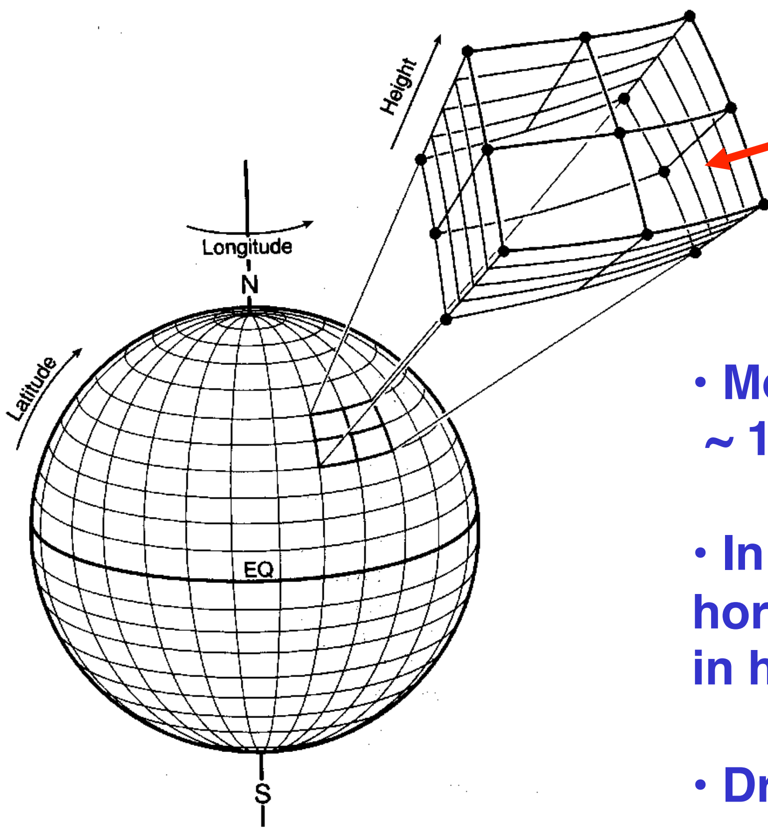
Lagrangian



Interpolation errors (of all variables to particle position)

EULERIAN RESEARCH MODELS SOLVE MASS BALANCE EQUATION IN 3-D ASSEMBLAGE OF GRIDBOXES

The mass balance equation is then the finite-difference approximation of the continuity equation.

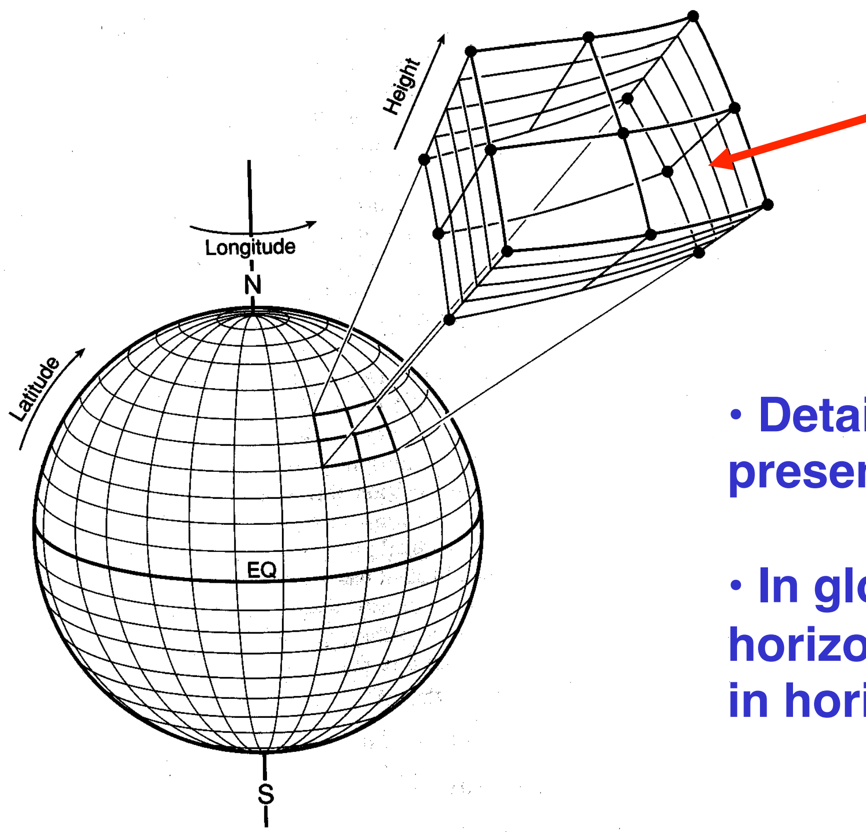


Solve continuity equation for individual gridboxes

- Models can presently afford ~ 10^6 gridboxes
- In global models, this implies a horizontal resolution of 100-500 km in horizontal and ~ 1 km in vertical
- Drawbacks: “numerical diffusion”, computational expense

EULERIAN MODELS PARTITION ATMOSPHERIC DOMAIN INTO GRIDBOXES

This discretizes the continuity equation in space



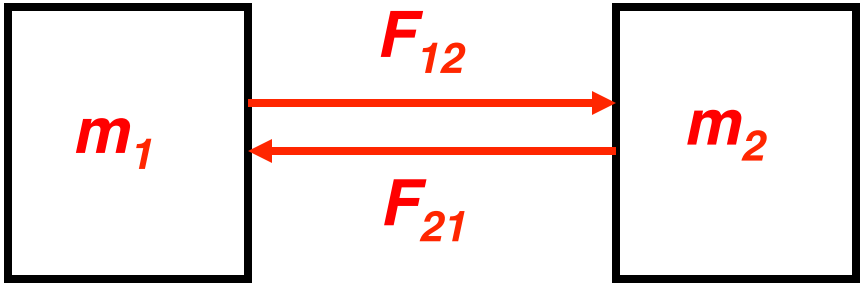
Solve continuity equation for individual gridboxes

- Detailed chemical/aerosol models can presently afford -10^6 gridboxes

- In global models, this implies a horizontal resolution of ~ 1° (~100 km) in horizontal and ~ 1 km in vertical

- Chemical Transport Models (CTMs) use external meteorological data as input
- General Circulation Models (GCMs) compute their own meteorological fields

TWO-BOX MODEL
defines spatial gradient between two domains



Mass balance equations:

$$\frac{dm_1}{dt} = E_1 + P_1 - L_1 - D_1 - F_{12} + F_{21}$$

(similar equation for dm_2/dt)

If mass exchange between boxes is first-order:

$$\frac{dm_1}{dt} = E_1 + P_1 - L_1 - D_1 - k_{12}m_1 + k_{21}m_2$$

⇨ system of two coupled ODEs (or algebraic equations if system is assumed to be at steady state)

OPERATOR SPLITTING IN EULERIAN MODELS

Reduces dimensionality of problem

- **Split the continuity equation into contributions from transport and local terms:**

$$\frac{\partial C_i}{\partial t} = \left[\frac{\partial C_i}{\partial t}\right]_{TRANSPORT} + \left[\frac{dC_i}{dt}\right]_{LOCAL}$$

$$\text{Transport} \equiv \text{advection, convection: } \left[\frac{dC_i}{dt}\right]_{TRANSPORT} = -\mathbf{U} \bullet \nabla C_i$$

$$\text{Local} \equiv \text{chemistry, emission, deposition, aerosol processes:}$$

$$\left[\frac{dC_i}{dt}\right]_{LOCAL} = P_i(\mathbf{C}) - L_i(\mathbf{C})$$

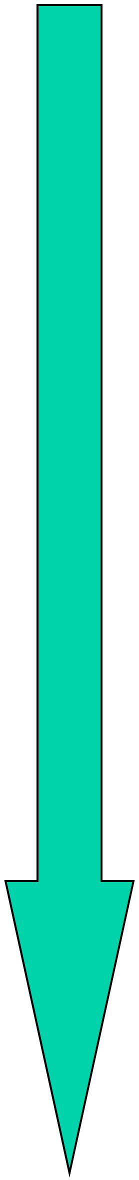
… and integrate each process separately over discrete time steps:

$$C_i(t_o + \Delta t) = (\text{Local})\bullet(\text{Transport}) \bullet C_i(t_o)$$

These operators can be split further:

- **split transport into 1-D advective and turbulent transport for *x, y, z* (usually necessary)**
- **split local into chemistry, emissions, deposition (usually not necessary)**

Time



For a grid of atmospheric columns:

1. 'Dynamics': Iterate Basic Equations
   Horizontal momentum, Thermodynamic energy,
   Mass conservation, Hydrostatic equilibrium,
   Water vapor mass conservation
2. Transport 'constituents' (water vapor, aerosol, etc)
3. Calculate forcing terms ("Physics") for each column
   Clouds & Precipitation, Radiation, etc
4. Update dynamics fields with physics forcings
5. Chemistry
6. Gravity Waves, Diffusion (fastest last)
7. Next time step (repeat)

SOLVING THE EULERIAN ADVECTION EQUATION

$$\frac{\partial C_i}{\partial t} = -u\frac{\partial C_i}{\partial x}$$

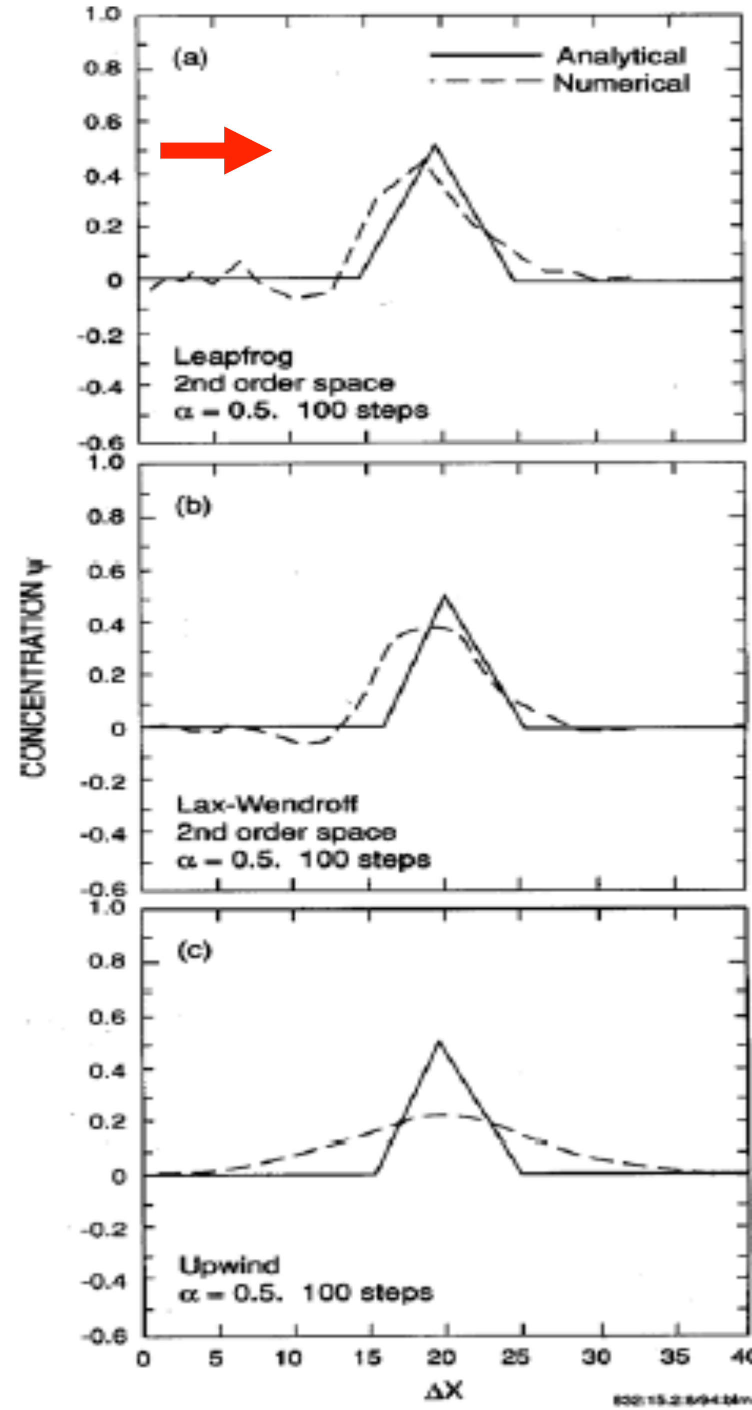
- Equation is *conservative:* need to avoid diffusion or dispersion of features. Also need mass conservation, stability, positivity…

- All schemes involve finite difference approximation of derivatives : order of approximation → accuracy of solution

- Classic schemes: leapfrog, Lax-Wendroff, Crank-Nicholson, upwind, moments…

- Stability requires Courant number *uDt/Dx* < 1
… limits size of time step

- Addressing other requirements (e.g., positivity) introduces non-linearity in advection scheme



SPLITTING THE TRANSPORT OPERATOR

- Wind velocity $\mathbf{U}$ has turbulent fluctuations over time step Dt:

$$\mathbf{U}(t) = \mathbf{U} + \mathbf{U}'(t)$$

Time-averaged component (resolved) — Fluctuating component (stochastic)

- Split transport into advection (mean wind) and turbulent components:

$$\frac{\partial C_i}{\partial t} = -\mathbf{U} \bullet \nabla C_i + \frac{1}{\rho} \nabla \bullet \mathbf{K} \nabla C_i$$

advection — turbulence (1st-order closure)

$\rho \equiv$ air density

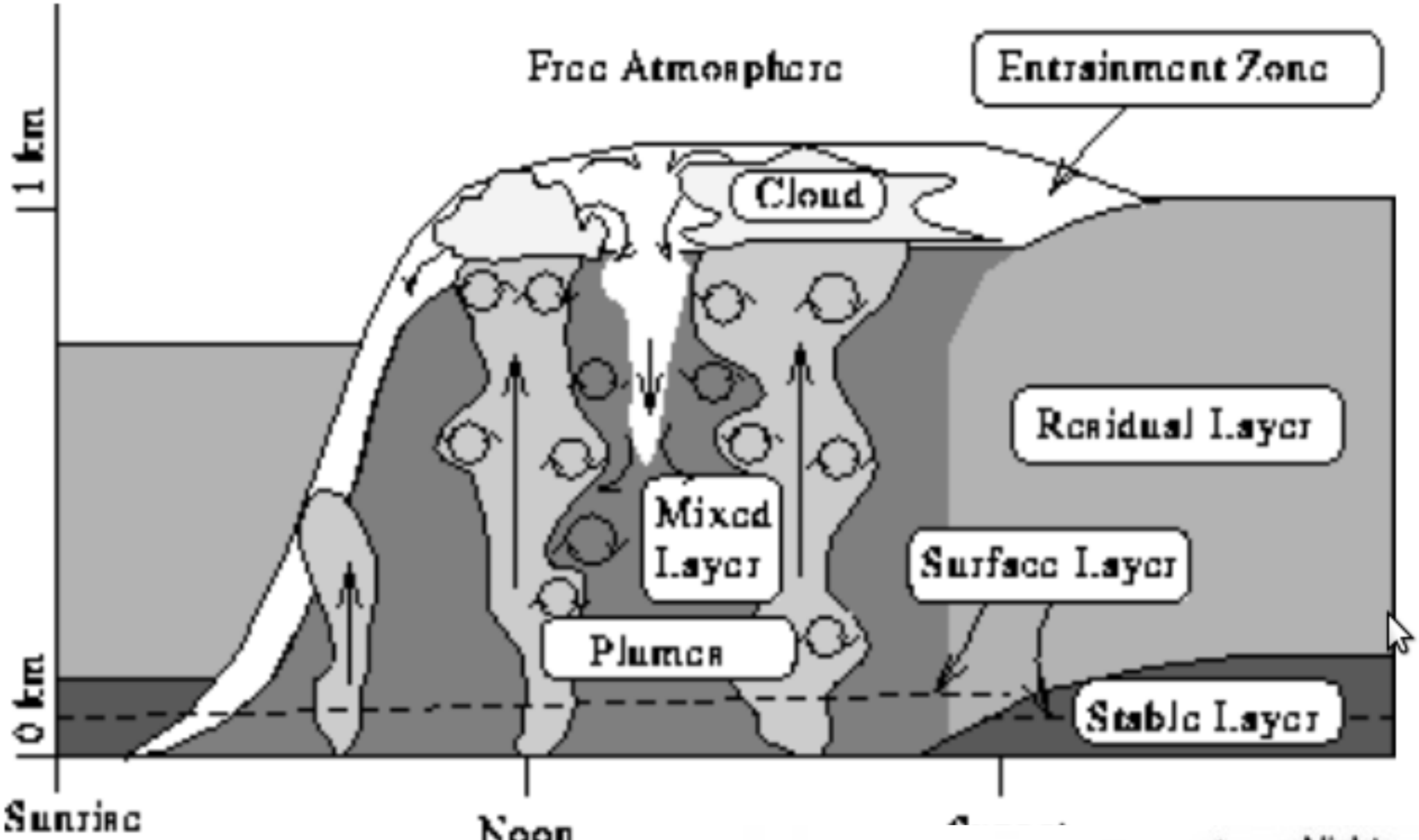
$\mathbf{K} \equiv$ turbulent diffusion matrix

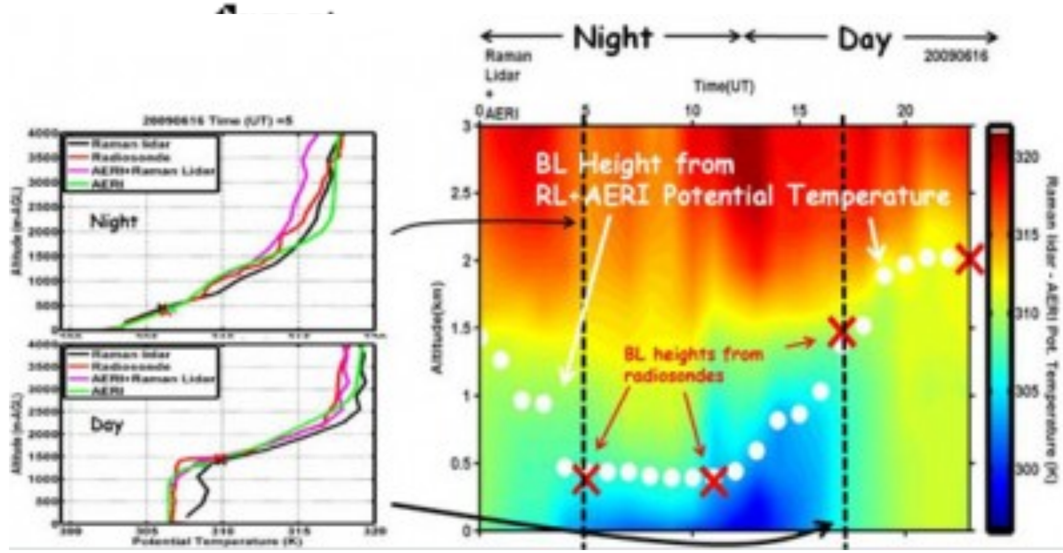
- Further split transport in *x, y, and z* to reduce dimensionality. In *x* direction:

$$\frac{\partial C_i}{\partial t} = -u \frac{\partial C_i}{\partial x} + \frac{1}{\rho} \frac{\partial}{\partial x} (K_{xx} \frac{\partial C_i}{\partial x})$$

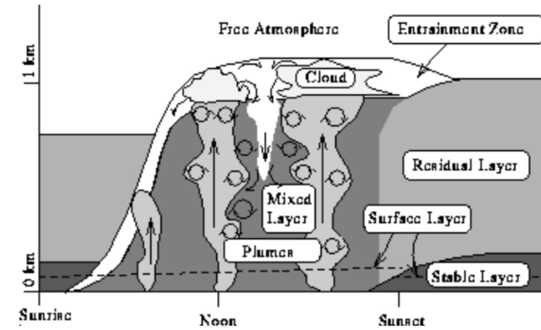
$$\mathbf{U} = (u, v, w)$$

advection operator — turbulent operator


Free Atmosphere
Entrainment Zone
1 km
Cloud
Residual Layer
Mixed Layer
Surface Layer
Plumes
Stable Layer
0 km
Sunrise
Noon


Night
Day
20090616
Raman Lidar + AERI
Time(UT)
BL Height from RL+AERI Potential Temperature
BL heights from radiosondes
Altitude(km)
Raman lidar - AERI Pot. Temperature (K)
20090616 Time (UT) =5
Raman lidar
Radiosonde
AERI+Raman Lidar
AERI
Night
Day
Altitude (m-AGL)
Potential Temperature (K)

Boundary layer height



http://lidar.ssec.wisc.edu/papers/akp_thes/node6.htm

Boundary layer height calculated using critical Ri (Vogelezang and Holtslag, 1996)

if $$Ri_l = \frac{(g/\Theta_{v1})(\Theta_{vl} - \Theta_{v1})(z_l - z_1)}{(u_l - u_1)^2 + (v_l - v_1)^2 + 100u_*^2}$$ **> 0.25 → *l* is PBLH**

- **If convective (unstable) situations then one iteration is made (max number iterations 3):**

$$\Theta'_{v1} = \Theta_{v1} + 8.5\frac{(\overline{w'\Theta'_v})_0}{w_* c_p},$$

Temp. excess from rising thermals

$$w_* = \left[\frac{(\overline{w'\Theta'_v})_0 g h_{mix}}{\Theta_{v1} c_p}\right]^{1/3}$$

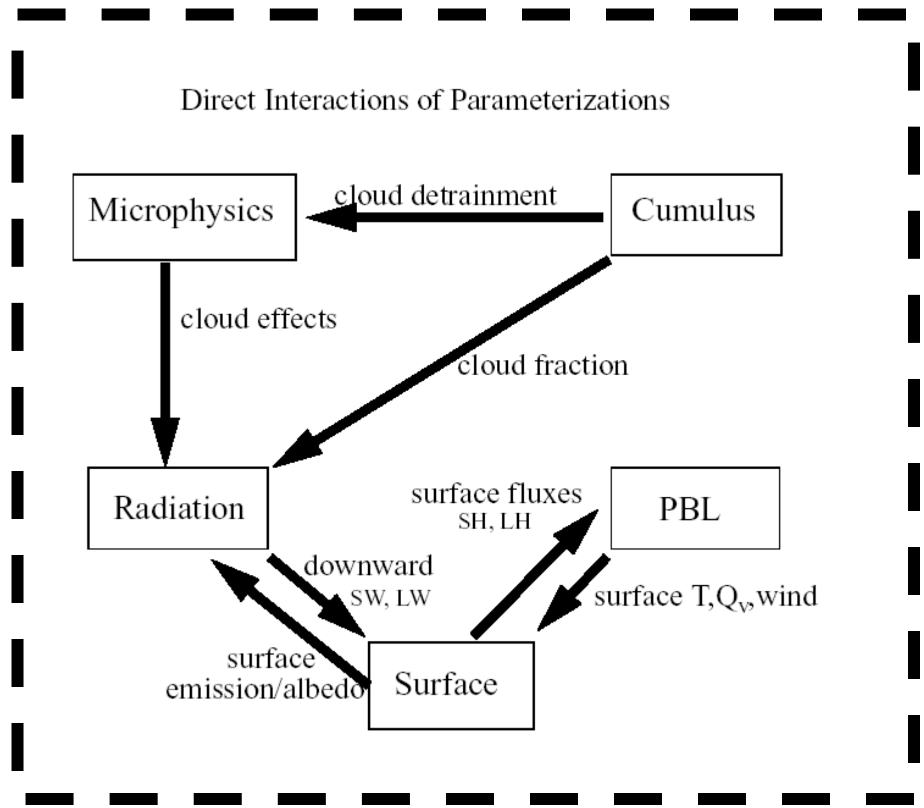
$$u_* = \frac{\kappa \Delta u}{\ln\frac{z_l}{10} - \Psi_m(\frac{z_l}{L}) + \Psi_m(\frac{10}{L})},$$

$$\Theta_* = \frac{\kappa \Delta\Theta}{0.74\left[\ln\frac{z_l}{2} - \Psi_h(\frac{z_l}{L}) + \Psi_h(\frac{2}{L})\right]},$$

$$L = \frac{\overline{T} u_*^2}{g\kappa\Theta_*},$$

Physics: feedbacks

Solve equations w/out non-explicit terms



- Parameterization computes the changes in temperature and moisture (and possibly cloud water, momentum, etc.)
- Tendency, applied at each timestep
- The tendency can be calculated each n timestep (where t_c is a convective timescale, typically 30min to 1hr)

momentum

$$\frac{du}{dt} + \frac{1}{\rho}\frac{\partial p}{\partial x} - fv = (P_{conv}) + P_{hdiff} + P_{vdiff} + P_{sfc}$$

$$\left.\frac{\partial \theta}{\partial t}\right|_{conv} = \frac{\theta_{final} - \theta_{initial}}{\tau_c} = P_{conv}$$

How Does the Feedback Occur Physics

At every grid point, predictive variables change at each time step

Different processes concur to modify temperature and water vapour

temperature

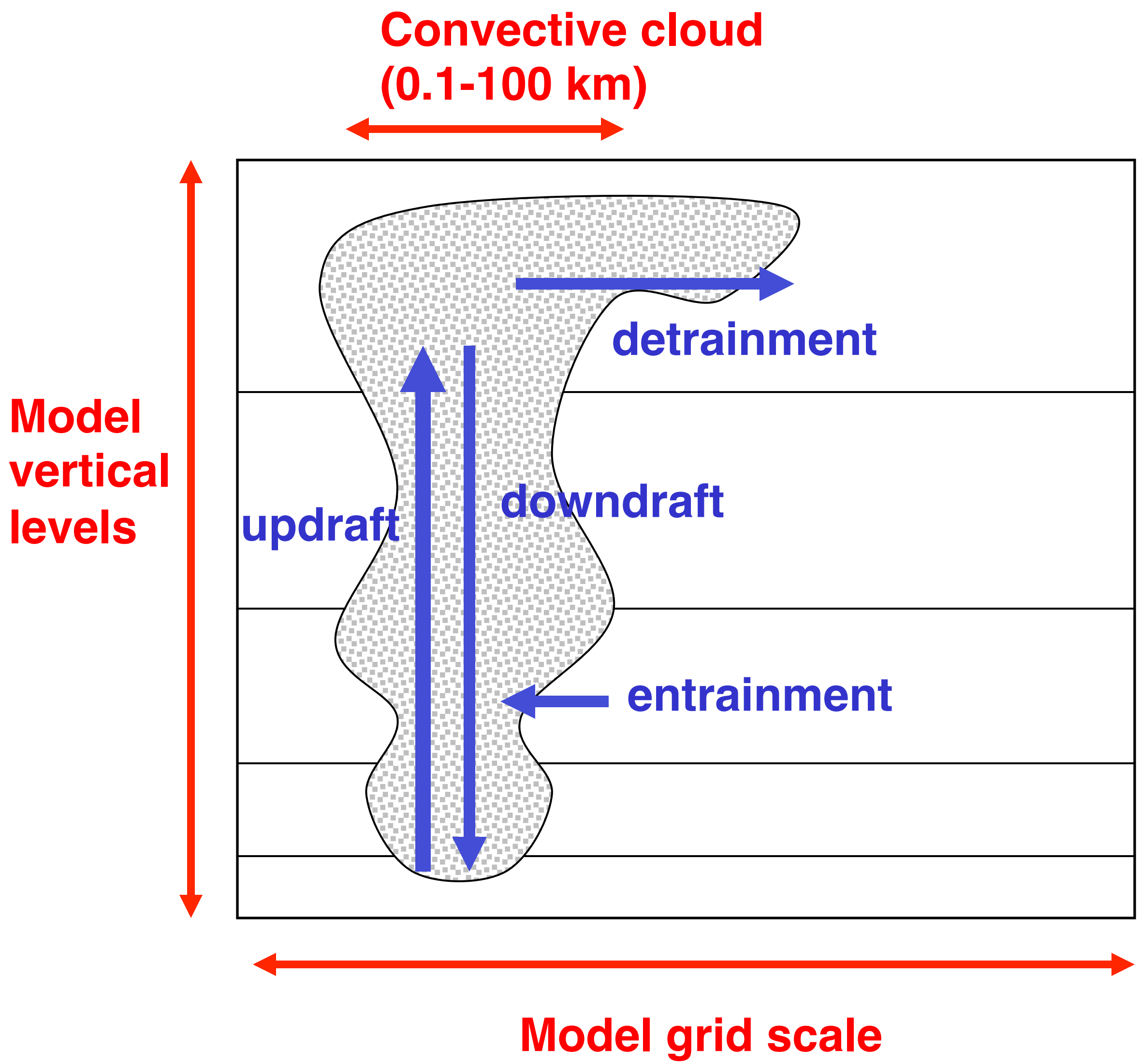
$$\frac{d\theta}{dt} = P_{rad} + P_{conv} + P_{cond/evap} + P_{hdiff} + P_{vdiff} + P_{sfc}$$

water vapor

$$\frac{dq_v}{dt} = P_{conv} + P_{cond/evap} + P_{hdiff} + P_{vdiff} + P_{sfc}$$

VERTICAL TURBULENT TRANSPORT (BUOYANCY)

- generally dominates over mean vertical advection
- K-diffusion OK for dry convection in boundary layer (small eddies)
- Deeper (wet) convection requires non-local convective parameterization



Wet convection is subgrid scale in global models and must be treated as a vertical mass exchange separate from transport by grid-scale winds.

Need info on convective mass fluxes from the model meteorological driver.

How to Parameterize convection

Relate unresolved effects to grid-scale properties using statistical or empirical techniques

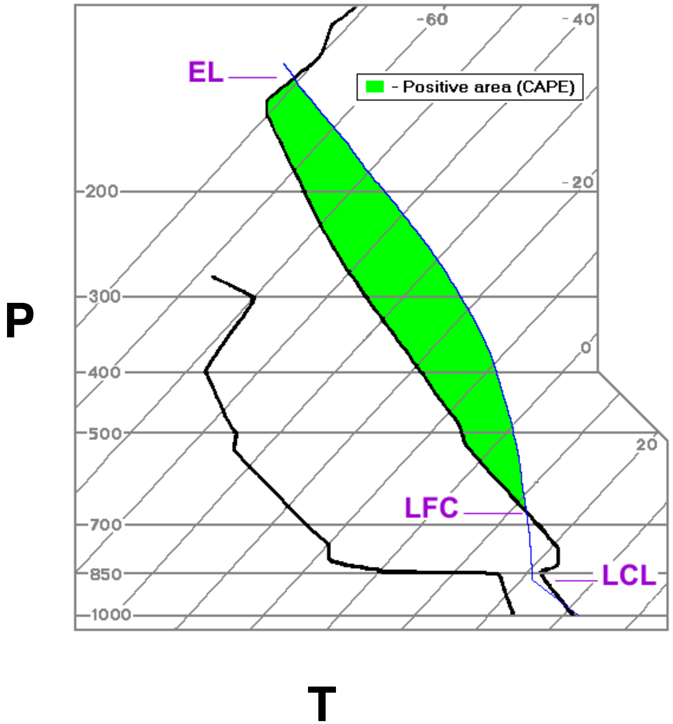
Several schemes (Grell-Pan / Kain-Fritsch / Betts-Miller / Emanuel / …)

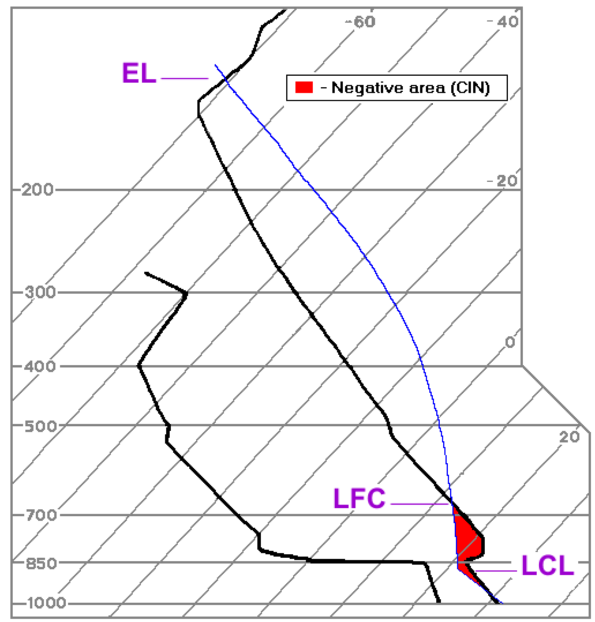
Mass-Flux: use simple cloud models to simulate rearrangements of mass in a vertical column

What properties of convection do we need to predict?

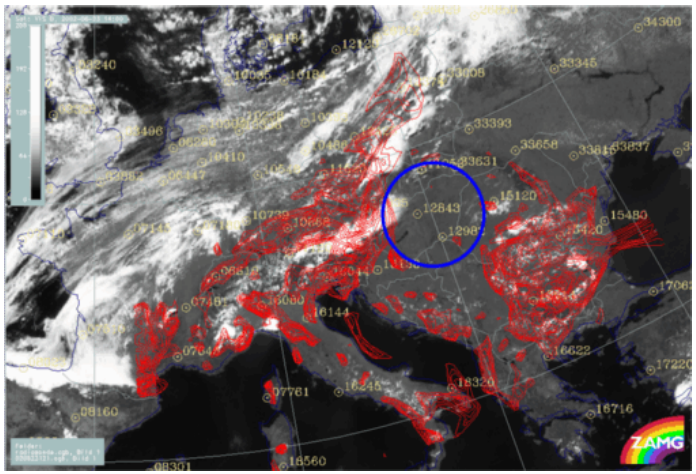
- convective triggering (yes/no)
- convective intensity (how much rain?)
- vertical distribution of heating and drying (feedback)

No scheme required if resolution high enough to reproduce updraft / downdraft (5 km)





$$CAPE = g \cdot \int_{ZLFC}^{ZLNB} \left(\frac{T'_v - T_v}{T_v} \right) dz$$



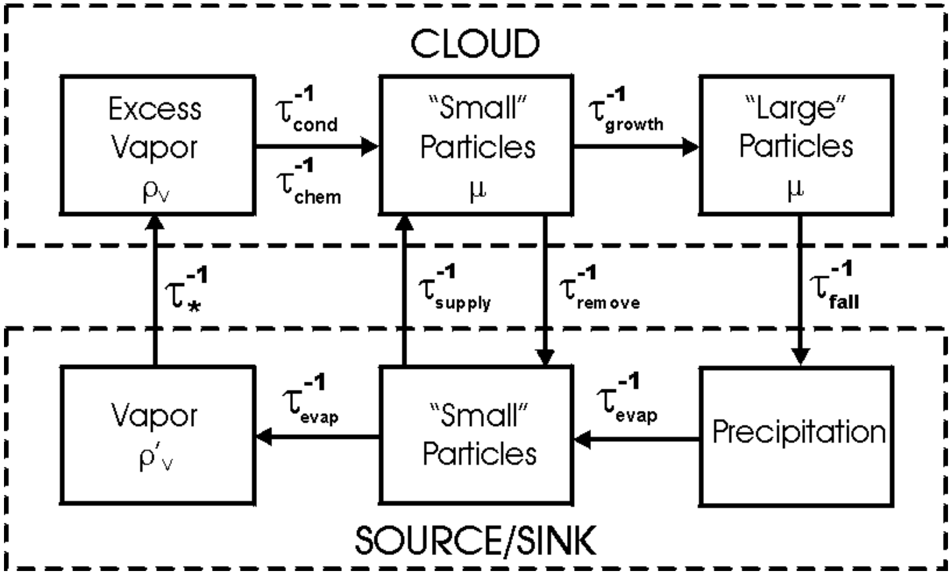
$$T_v = T\left(\frac{1 + r/\varepsilon}{1 + r}\right),$$

$$T_v = T(1 + 0.61r).$$

And how much...

- Convective intensity (net heating)
  - proportional to mass or moisture convergence
  - sufficient to offset large-scale destabilization rate
  - sufficient to eliminate CAPE (constrained by available moisture)
- Vertical distribution of heating and drying
  - determined by nudging to empirical reference profiles
  - estimated using a simple 1-D cloud model to satisfy the constraints on intensity

Cloud processes



- Cloud / Ice / Rain / Snow / Graupel

- Condensation / Collection

- Melting / Evaporation / Fall

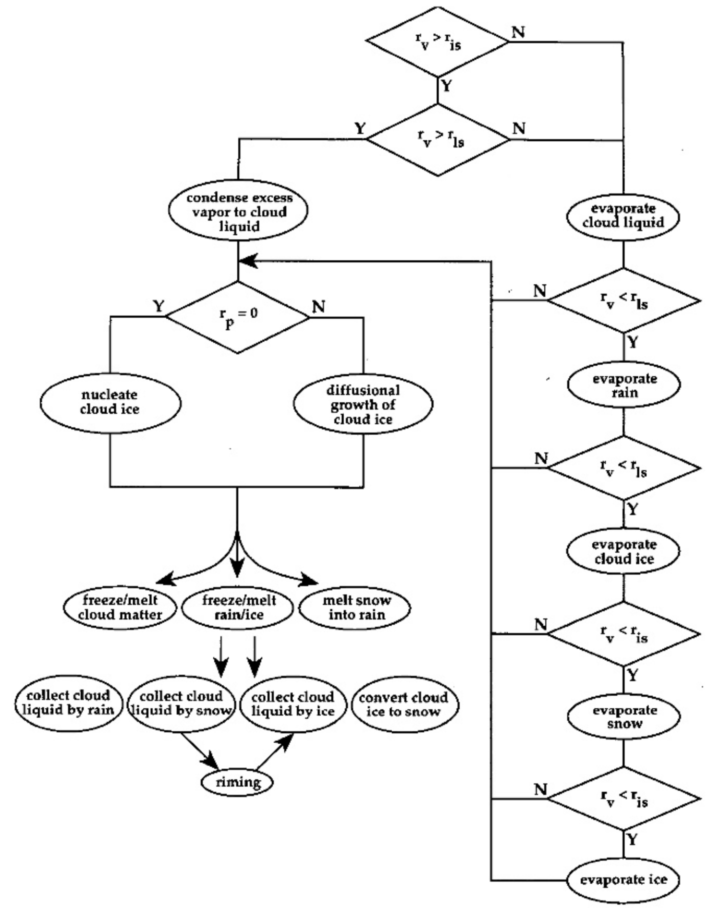


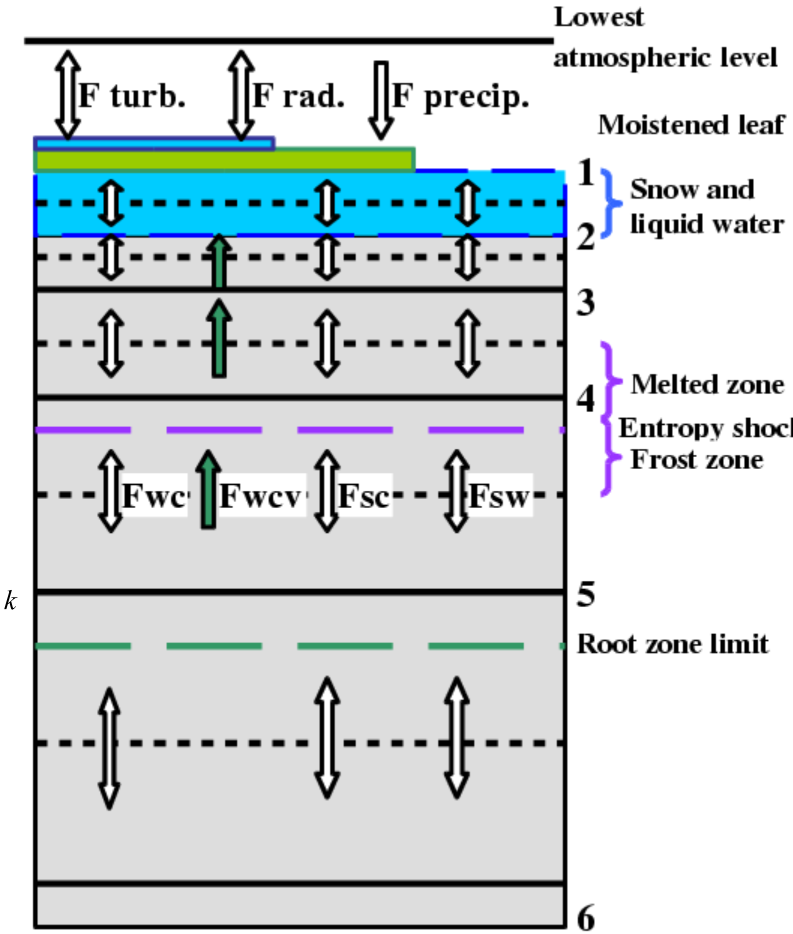
FIG. 1. Flow diagram for the NWP explicit microphysics algorithm; r is mixing ratio. The subscripts are v, vapor; p, cloud ice; ls, liquid saturation; is, ice saturation.

Soil processes ...

$$\Phi_q = \rho_S C_D \left(q_{SKIN} - q_1\right)$$

$$\Pi_1 = -P_{RES} - P_{MELT} + \Phi_{SOIL} + \sum_{k=1} \Phi_{EVTR_k}$$

$$q_1^G(t+\Delta t) = q_1^G(t) + \frac{\Delta t}{\rho_W \Delta z_1}\left(-\Pi_1 + \Pi_2\right)$$



Schematic representation of the vegetation-soil scheme

F turb. is turbulent flux of entropy and water vapour, **F rad.** is flux of shortwave and longwave radiation, **F precip.** is flux of atmospheric precipitation, **Fwc** and Fwcv are hydraulic and vegetation fluxes of soil water content, **Fsc** and **Fsw** are conductivity and hydraulic fluxes of soil entropy

IN EULERIAN APPROACH, DESCRIBING THE EVOLUTION OF A POLLUTION PLUME REQUIRES A LARGE NUMBER OF GRIDBOXES

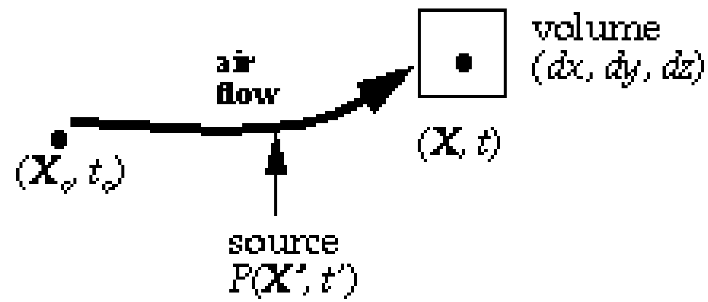
Fire plumes over southern California, 25 Oct. 2003

A Lagrangian "puff" model offers a much simpler alternative

Transport: Lagrangian

Transition probability density

$$\int_{atm} Q(\mathbf{X}_o, t_o | \mathbf{X}, t)\, dx\,dy\,dz = 1$$



$$n(\mathbf{X}, t) = \int_{atm} Q(\mathbf{X}_o, t_o | \mathbf{X}, t)\, n(\mathbf{X}_o, t_o)\, dx_o\, dy_o\, dz_o$$

$$+ \int_{atm} \int_{t_o}^{t} Q(\mathbf{X}', t' | \mathbf{X}, t)\, P(\mathbf{X}', t')\, dx'dy'dz'dt'$$

Q is difficult to estimate ---> use wind field U

Assumptions about Trajectory Transport

Parcels have no inertia (m = 0)

Parcels have no size yet "represent" their surroundings

Parcels don't know about each other except when some kind of explicit mixing is included

$$\frac{dX}{dt} = \dot{X}[X(t)] \qquad X^1(t_1) \approx X(t_0) + \frac{1}{2}(\Delta t)[X(t_0) + X(t_1)]$$

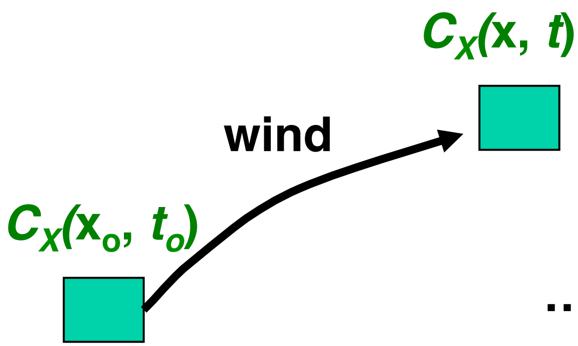
The "constant acceleration" solution

Neglects higher order terms in the Taylor series expansion of the first equation (source of truncation errors)

time resolution of wind fields, interpolation errors, vertical wind issues, wind field errors, tropospheric process errors

Stohl A., Computation, Accuracy and Applications of Trajectories - A Review and Bibliography, *Atmospheric Environment, 32*, 947-966, 1998.

PUFF MODEL: FOLLOW AIR PARCEL MOVING WITH WIND

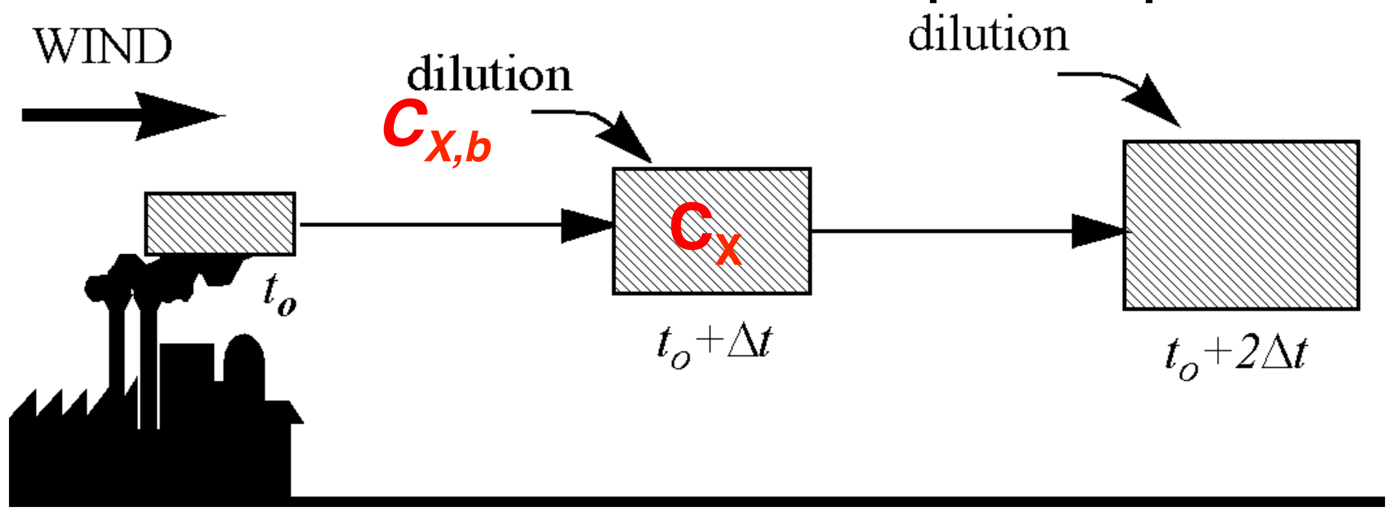


In the moving puff,

$$\frac{dC_X}{dt} = E + P - L - D$$

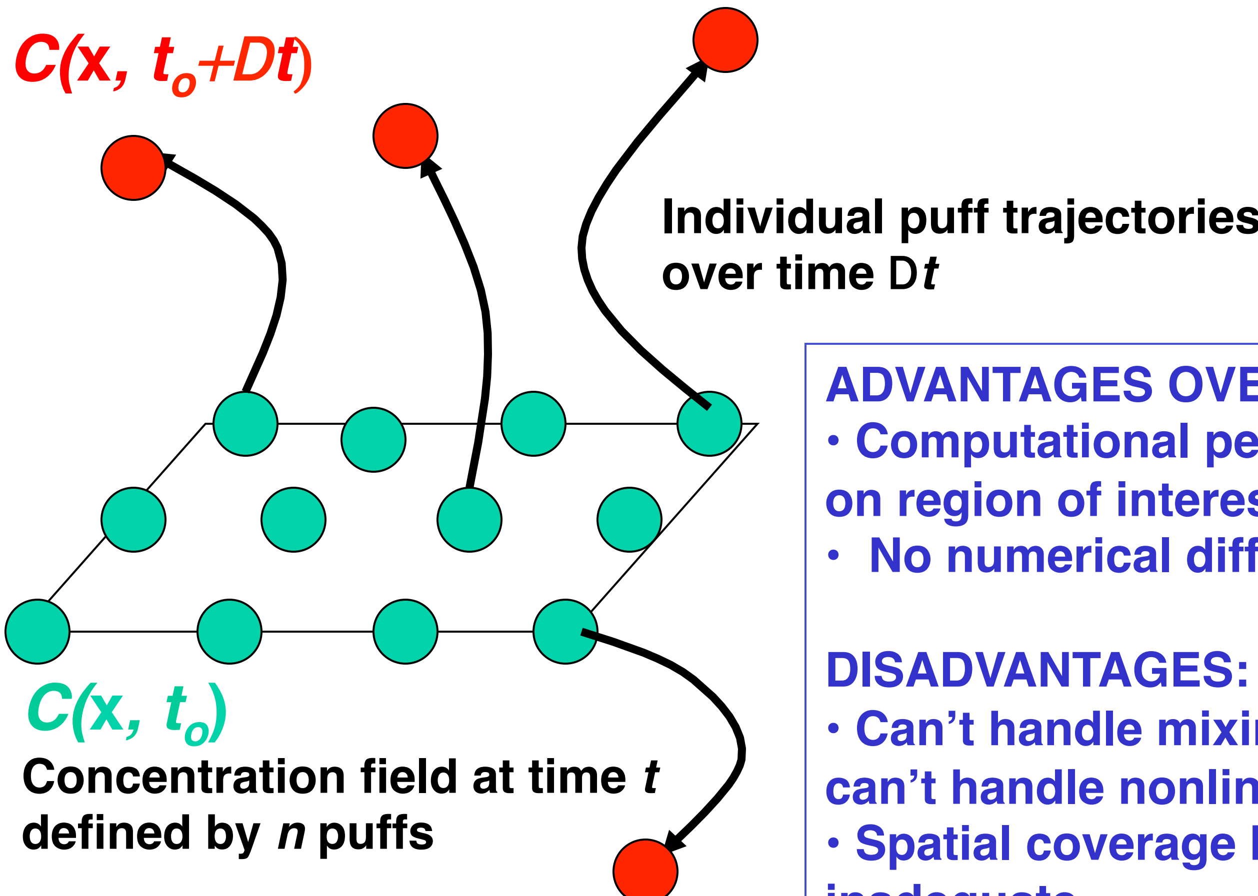
…no transport terms! (they're implicit in the trajectory)

Application to the chemical evolution of an isolated pollution plume:



In pollution plume, $$\frac{dC_X}{dt} = E + P - L - D - k_{dilution}(C_X - C_{X,b})$$

LAGRANGIAN RESEARCH MODELS FOLLOW LARGE NUMBERS OF INDIVIDUAL "PUFFS"



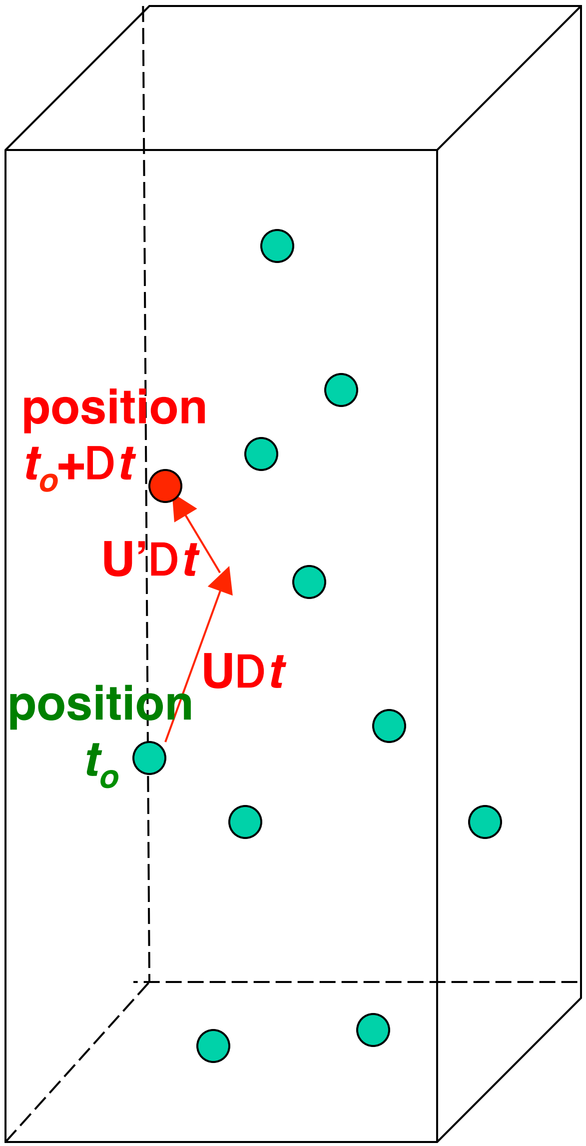
ADVANTAGES OVER EULERIAN MODELS:
- Computational performance (focus puffs on region of interest)
- No numerical diffusion

DISADVANTAGES:
- Can't handle mixing between puffs ⇨ can't handle nonlinear processes
- Spatial coverage by puffs may be inadequate

Truncation	Traj. computed with short integration time steps	Errors resulting from time step of 3 h using zero (constant) [variable] acceleration method	42	300 (100) [40] km	Walmsley and Ma
Interpolation	Zero-interpolation error traj.	Superposition of stochastic interpolation errors occurring along traj.	72	400 km	Kahl and Samson
Interpolation	Zero-interpolation error traj.	Same as above, but for more convective conditions	72	500 km	Kahl and Samson
Temporal interpolation	Calculated traj.	3-month set of 3-D traj. calculated from wind fields of 12 h (6 h) [4 h] time resolution vs. 2 h time resolution	96	730 km (410 km) [250 km]	Rolph and Draxle
Temporal interpolation	Calculated traj.	86 3-D traj. in an intense cyclone calculated from wind fields of 6 h (3 h) [1 h] time resolution vs 15 minutes time resolution	36	250 km (170 km) [30 km]	Doty and Perkey
Temporal interpolation	Calculated traj.	1-yr set of 3-D (2-D) traj. calculated from wind fields of 6 h time resolution vs 3 h time resolution	96	590 km, 20% (280 km, 9%)	Stohl *et al.* (1995)
Horizontal interpolation	Calculated traj.	3-month set of traj. calculated from wind fields of 360 km (180 km) resolution vs 90 km resolution	96	420 km (170 km)	Rolph and Draxle
Horizontal interpolation	Calculated traj.	1-yr set of 3-D (2-D) traj. calculated from wind fields of 1° resolution vs 0.5° resolution	96	411 km, 14% (111 km, 4%)	Stohl *et al.* (1995)
Forecast	Analysis traj.	1-yr set of 950 hPa forward traj. started at $T = 0$ h ($T = +36$ h)	36	245 km, 25% (720 km, 60%)	Maryon and Hea
Forecast	Analysis traj.	1-yr set of forward 3-D traj. started 500, 1000, 1500 m above ground	> 12	200 km/day	Stunder (1996)
Forecast	Analysis traj.	1-yr set of back traj. travelling 800 m above ground terminating at $T = +24$ h ($T = +48$ h) [$T = +72$ h]	96	16% (26%) [36%]	Stohl (1996a)
Wind field analysis	ECMWF traj. compared to NMC traj.	Isobaric 850 and 700 hPa traj.	120	1000 km	Kahl *et al.* (1989a
Wind field analysis	ECMWF traj. compared to NMC traj.	Isentropic traj. over the south Atlantic	120 (192)	1500 km, 60% (2500 km, 60%)	Pickering *et al.* (1
Total	Constant level balloon	26 cases, diagnostic wind field model used	< 24	25–30%	Clarke *et al.* (198
Total	Constant level balloon	16 cases in and immediately above the PBL	1–3	5–40%	Koffi *et al.* (1997a
Total	Constant level balloon	Stratospheric traj.	12–144	$\approx$ 20%	Knudsen and C Knudsen *et al.* (19
Total	Manned balloon	Single flight at a typical height of 500 hPa	100	10%	Draxler (1996b)
Total	Manned balloon	4 flights at a typical height of 2000 m	46	$< 20\%$	Baumann and Sto
Total	Tracer (CAPTEX)	6 cases, different types of traj.	24	$\approx$ 200 km	Haagenson *et al.*
Total	Tracer (CAPTEX)	6 cases	24–42	150–180 km	Draxler (1987)
Total	Tracer (ANATEX)	30 cases	< 30	20–30%	Draxler (1991)
Total	Tracer (ANATEX)	23 boundary layer traj.	24–72	$\approx$ 100 km/d^{-1}	Haagenson *et al.*
Total	Smoke plumes	112 traj. based on a fine-scale (global) analysis	< 60	10% (14%)	McQueen and Dr
Total	Saharan dust	Single case, 3-D traj.	3000 km	200 km, 7%, vertical error 50 hPa	Reiff *et al.* (1986)
Total	Potential vorticity	1-yr set of 3-D traj. based on ECMWF data	120	$< 20\%$, < 400 km,	Stohl and Seibert

Stohl A., Computation, Accuracy and Applications of Trajectories - A Review and Bibliography, *Atmospheric Environment, 32*, 947-966, 1998.

LAGRANGIAN APPROACH: TRACK TRANSPORT OF POINTS IN MODEL DOMAIN (NO GRID)



- Transport large number of points with trajectories from input meteorological data base (U) + random turbulent component (U') over time steps Dt

- Points have mass but no volume

- Determine local concentrations as the number of points within a given volume

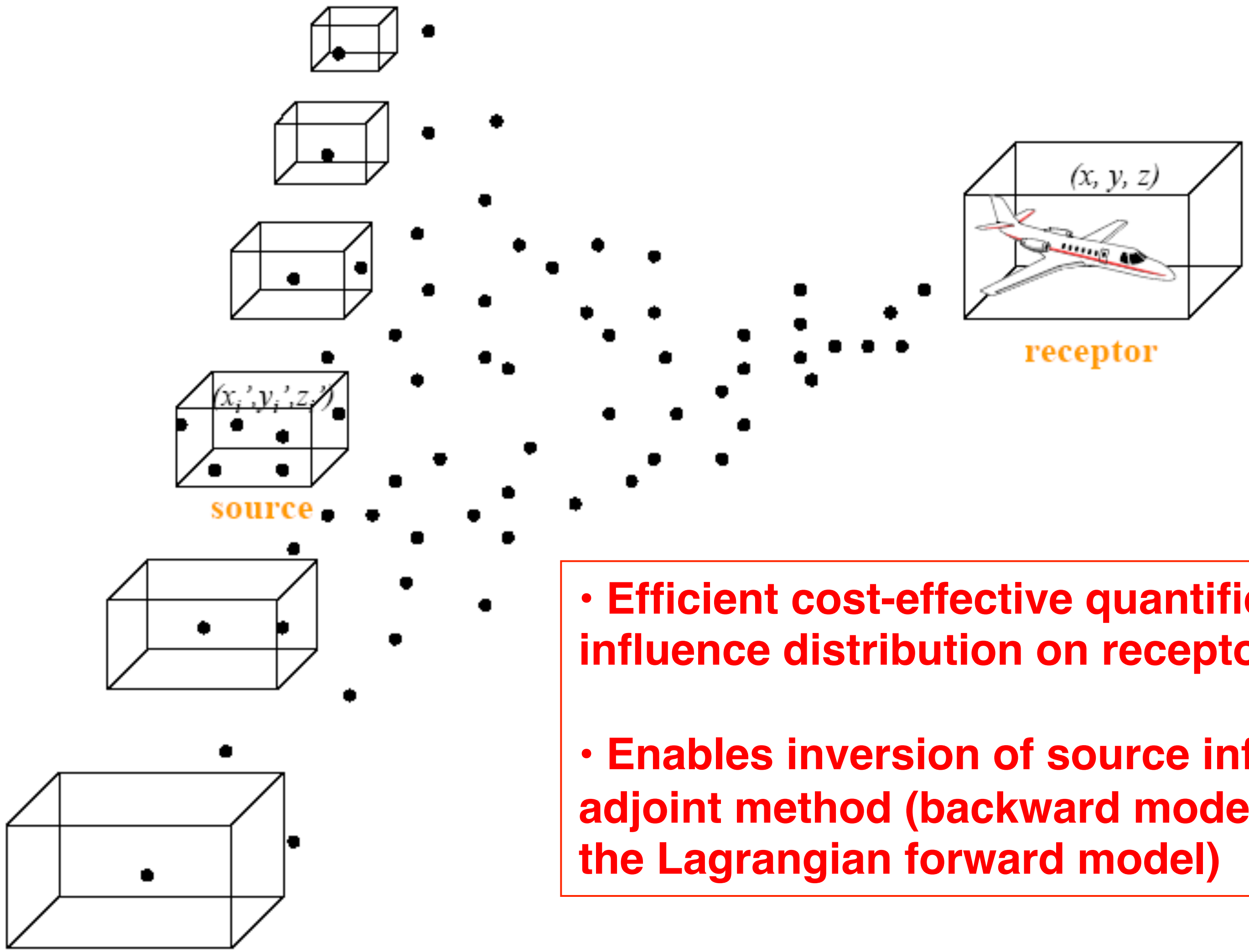
- Nonlinear chemistry requires Eulerian mapping at every time step (semi-Lagrangian)

PROS over Eulerian models:
- no Courant number restrictions
- no numerical diffusion/dispersion
- easily track air parcel histories
- invertible with respect to time

CONS:
- need very large # points for statistics
- inhomogeneous representation of domain
- convection is poorly represented
- nonlinear chemistry is problematic

LAGRANGIAN RECEPTOR-ORIENTED MODELING

Run Lagrangian model backward from receptor location, with points released at receptor location only



- **Efficient cost-effective quantification of source influence distribution on receptor ("footprint")**
- **Enables inversion of source influences by the adjoint method (backward model is the adjoint of the Lagrangian forward model)**

- **Can be computationally very efficient (depending on size of plume): only the fraction covered with particles is simulated.**
- **Turbulent processes are included in a more natural way unlike Eulerian models**
- **There is no numerical diffusion due to a computational grid**
- **Grid and/or kernels are used only for output purpose therefore no artificial diffusion is due to the averaging process**
- **Model is "self-adjoint" – can run backward in time, too.**
- **Many first order processes can be easily included with a prescribed rate: radioactive decay, dry deposition, washout, etc.**
- **One particle can carry more than one species**
- **Gravitational settling is easily included (as long as particles carry a single species)**
- ***However: it is quite difficult and computationally expensive to include non-linear chemical reactions and the process of gridding the output make as well loose some of the advantages of Lagrangian modelling.***

Stohl, A., S. Eckhardt, C. Forster, P. James,

N. Spichtinger, and P. Seibert (2002):

A replacement for simple back trajectory calculations in the interpretation of atmospheric trace substance measurements. Atmos. Environ. 36, 4635-4648

The FLEXPART is…

… a Lagrangian Particle Dispersion Model, originally developed at the University of Natural Resources and Life Sciences in Vienna, further developed by its main developer Andreas Stohl at the Norwegian Institute for Air Research in the Department of Atmospheric and Climate Research and with by group of developers in different institutions

It is released under the GNU General Public License V3.0

- Countries – 15
- Users http://transport.nilu.no/flexpart/flexpart-and-flextra-users >35

Get source code

Generate tickets that will be addressed by developers

Get updates and references

Get post-processing software

Get test data

Get course notes and data as exercise

flexpart.eu

First:

- Correctly track the particles in a given velocity field.

Second:

- Model the Sub-grid scale (SGS) unresolved physical processes that affect the particles dispersion:
  - Boundary Layer Turbulence
  - Mesoscale Turbulence
  - Cumulus turbulent convection

Third:

- Modify particles properties based on locally acting processes, e.g. radioactive decay

Fourth:

- Count particles in a volume and extract concentration value

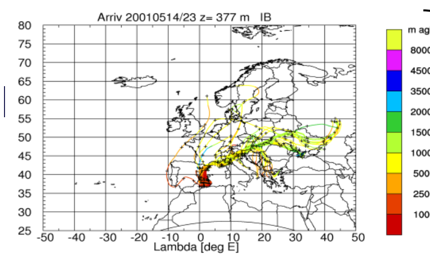
Transport and diffusion:

- **FLEXPART calculates trajectories of *computational* particles (each particle carries a certain amount of mass or mixing ratio of species – *computational* -, as defined in the releases)** ***(change of mass described later)***

Integration (1st order, zero acceleration scheme)

$$\frac{d\boldsymbol{X}}{dt} = \boldsymbol{v}[\boldsymbol{X}(t)] \longrightarrow \boldsymbol{X}(t+\Delta t) = \boldsymbol{X}(t) + \boldsymbol{v}(\boldsymbol{X},t)\Delta t$$

$\overline{\boldsymbol{v}}$ **Grid scale wind $\rightarrow$ what simpl trajectory models use (e.g. FLEXTRA)**



$\boldsymbol{v}_t$ **Turbulent wind fluctuations**

$\boldsymbol{v}_m$ **Mesoscale wind fluctuations (meandering)**

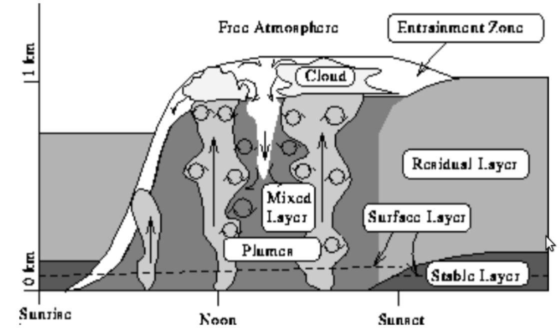
$$\boldsymbol{v} = \overline{\boldsymbol{v}} + \boldsymbol{v}_t + \boldsymbol{v}_m$$

`advance.f`

Boundary layer height

Boundary layer height calculated using critical

Ri (Vogelezang and Holtslag, 1996)



http://lidar.ssec.wisc.edu/papers/akp_thes/node6.htm

if $$Ri_l = \frac{(g/\Theta_{v1})(\Theta_{vl} - \Theta_{v1})(z_l - z_1)}{(u_l - u_1)^2 + (v_l - v_1)^2 + 100u_*^2}$$ **> 0.25 → *l* is PBLH**

- **If convective (unstable) situations then one iteration is made (max number iterations 3):**

$$\Theta'_{v1} = \Theta_{v1} + 8.5\frac{(\overline{w'\Theta'_v})_0}{w_* c_p},$$

Temp. excess from rising thermals

$$w_* = \left[\frac{(\overline{w'\Theta'_v})_0 g h_{mix}}{\Theta_{v1} c_p}\right]^{1/3}$$

$$u_* = \frac{\kappa \Delta u}{\ln\frac{z_l}{10} - \Psi_m(\frac{z_l}{L}) + \Psi_m(\frac{10}{L})},$$

$$\Theta_* = \frac{\kappa \Delta\Theta}{0.74\left[\ln\frac{z_l}{2} - \Psi_h(\frac{z_l}{L}) + \Psi_h(\frac{2}{L})\right]},$$

$$L = \frac{\overline{T} u_*^2}{g\kappa\Theta_*},$$

Vertical profiles of the turbulent quantities inside the ABL

Depend on the state of the turbulent atmosphere. Following Hanna 1982. σ_{v_i} τ_{L_i}

h, L, w_*, z_0 and u_*, i.e. ABL height, Monin-Obukhov length, convective velocity scale, roughness length and friction velocity, respectively. It is used in subroutines

1. Unstable

$$\frac{\sigma_u}{u_*} = \frac{\sigma_v}{u_*} = \left(12 + \frac{h}{2|L|}\right)^{1/3}$$

$$\tau_{L_u} = \tau_{L_v} = 0.15\frac{h}{\sigma_u}$$

$$\sigma_w = \left[1.2w_*^2\left(1 - 0.9\frac{z}{h}\right)\left(\frac{z}{h}\right)^{2/3} + \left(1.8 - 1.4\frac{z}{h}\right)u_*^2\right]^{1/2}$$

$z/h < 0.1$ and $z - z_0 > -L$:

$$\tau_{L_w} = 0.1\frac{z}{\sigma_w\left[0.55 - 0.38\left(z - z_0\right)/L\right]}$$

$z/h < 0.1$ and $z - z_0 < -L$:

$$\tau_{L_w} = 0.59\frac{z}{\;}$$

$z/h > 0.1$:

$$\tau_{L_w} = 0.15\frac{h}{\sigma_w}\left[1 - \exp\left(\frac{-5z}{h}\right)\right]$$

What about above the ABL?

In the free atmosphere turbulence is in small places coming from gravity waves, arround jet streams... it is not yet parameterized in detail.

FLEXPART treats the stratosphere with a constant vertical diffusivity (Legras et al. 2003)

$$D_z = 0.1\ \mathrm{m^2 s^{-1}}$$

And a constant horizontal diffusivity in the free troposphere

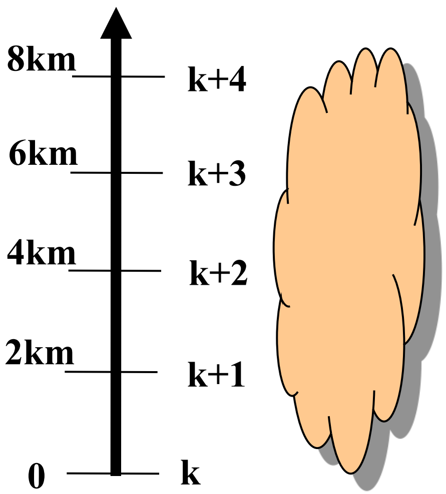
$$D_h = 50\ \mathrm{m^2 s^{-1}}$$

with an intermediate zone from free-troposphere to stratosphere. Turbulent velocity scales are then calculated by

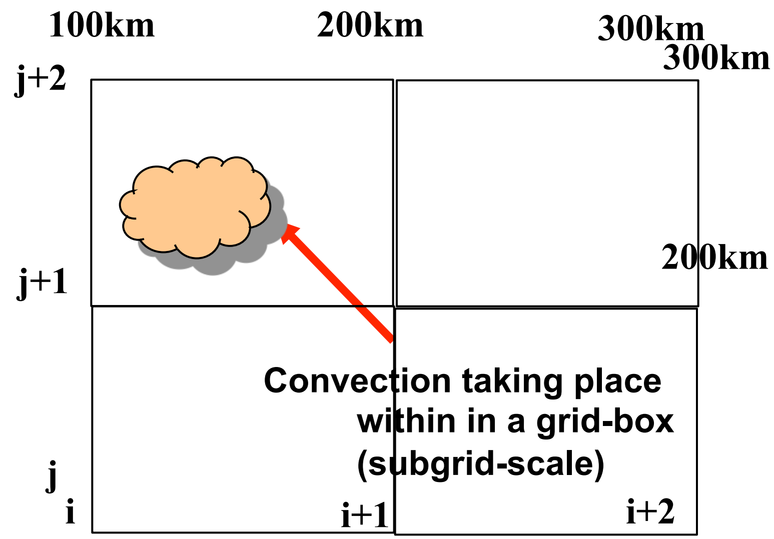
$$\sigma_{v_i} = \sqrt{D_i/dt}$$

Convection in models

convection is grid-scale in the vertical



but subgrid-scale in the horizontal

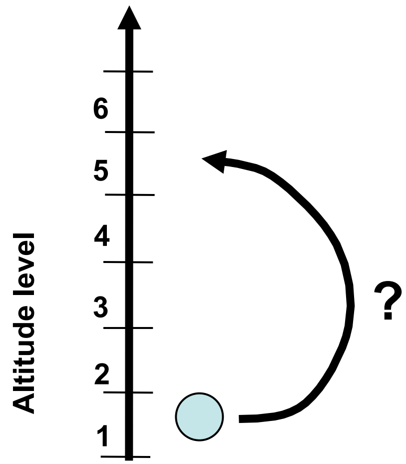


Meteorological parameters (temperature, humidity, wind etc.) given at horizonital model grid points (i,j), (i,j+1), (i,j+2) etc., but there is no information inbetween

→ convection has to be parameterized:
convection takes place under certain large-scale conditions

Convection parametrization

necessary to know how the particles shall be redistributed vertically, i.e. destination level of each particle must be known



The particles carry mass fractions in the model
→ mass fraction M displaced from level i to level j must be known

Matrix M(i,j)
i: source level
j: destination level

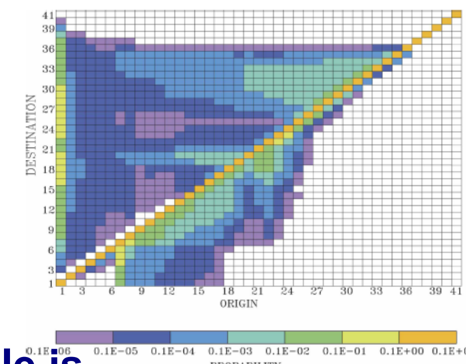


FIG. 1. Example of a mean convective redistribution matrix along 10° latitude for October 1983 calculated from the EZ99 scheme. The colors indicate the probability for a particle to be displaced from its origin to its destination level. White colors indicate probabilities below 10^{-6}. The sum of each column is 1. Origin and destination are given in numbers of model levels. For the height of these model levels, see Table 1.

FLEXPART interface:

construct a matrix of condtional probabilities P(i,j) that a particle is displaced from level i to level j given that it is in level i

P(i,j)=M(i,j)|t/m(i)

Assume that all convective fluxes (the matrix) are balanced by compensating subsidence (a downward velocity) in the environment; the subsidence acts on those particles that are not displaced by the matrix

Forster, C., A. Stohl, and P. Seibert, 2007: Parameterization of convective transport in a Lagrangian particle dispersion model and its evaluation, J. Appl. Meteorology, Vol. 46, No. 4, 403-422.

Dry deposition for gases

Calculated with the resistance method (Wesely and Hicks, 1977, 2000) – analogy to electrical resistance

$$|v_d(z)| = [r_a(z) + r_b + r_c]^{-1}$$

Aerodynamic resistance between z and a the top of the vegetation canopy

Quasilaminar sublayer resistance

Bulk surface resistance

$$r_a(z) = \frac{1}{\kappa u_*}[\ln(z/z_0) - \Psi_h(z/L) + \Psi_h(z_0/L)]$$

Profile function

Ability of the eddies to bring the material close to the surface, except for large particles, dependent on the flow

getvdep.f

raerod.f

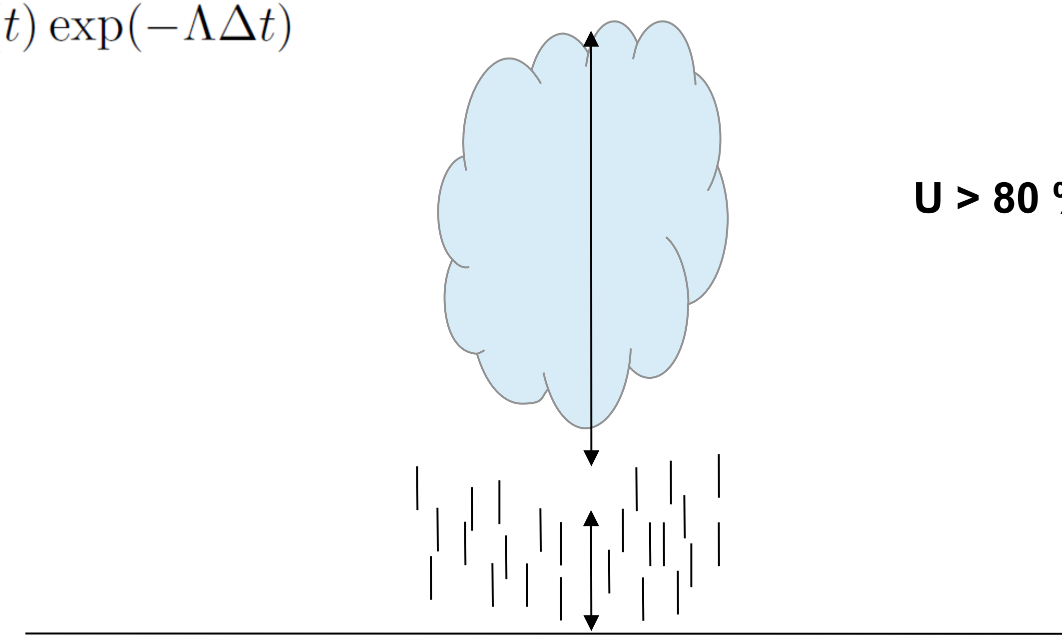
getrb.f / getrc

In FLEXPART (v 8 →) wet deposition is separated into:

1. In-cloud scavenging (also called rainout) – very efficient process
2. Below-cloud scavenging (also called washout)
3. No differences with snow scavenging processes in FLEXPART

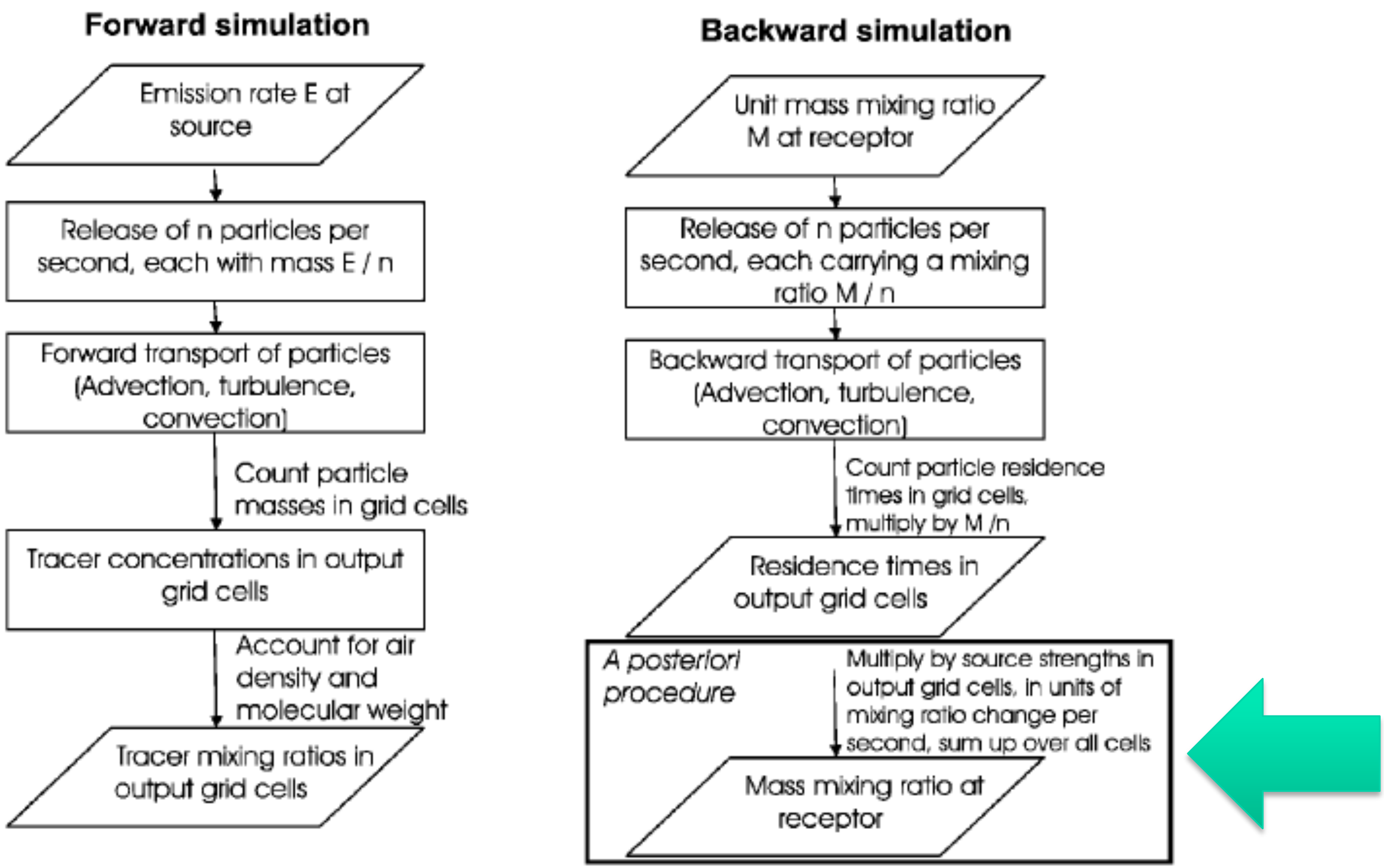
$$m(t+\Delta t) = m(t)\exp(-\Lambda \Delta t)$$

Change of mass

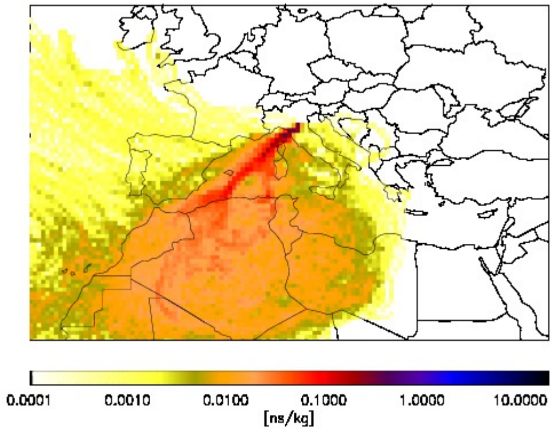


U > 80 % → CLOUD

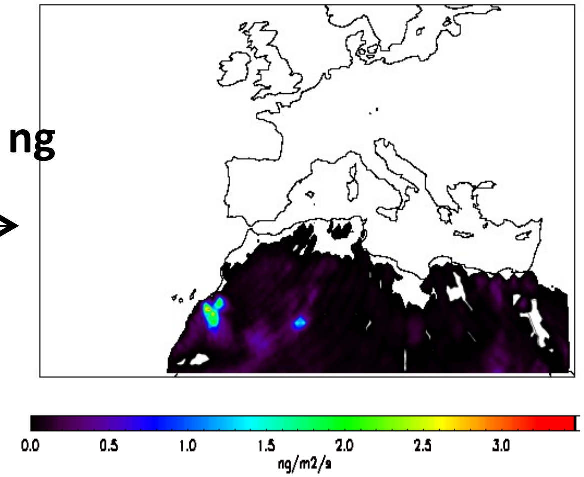
wetdepo.f


Forward simulation
Emission rate E at source
Release of n particles per second, each with mass E / n
Forward transport of particles (Advection, turbulence, convection)
Count particle masses in grid cells
Tracer concentrations in output grid cells
Account for air density and molecular weight
Tracer mixing ratios in output grid cells
Backward simulation
Unit mass mixing ratio M at receptor
Release of n particles per second, each carrying a mixing ratio M / n
Backward transport of particles (Advection, turbulence, convection)
Count particle residence times in grid cells, multiply by M /n
Residence times in output grid cells
A posteriori procedure
Multiply by source strengths in output grid cells, in units of mixing ratio change per second, sum up over all cells
Mass mixing ratio at receptor

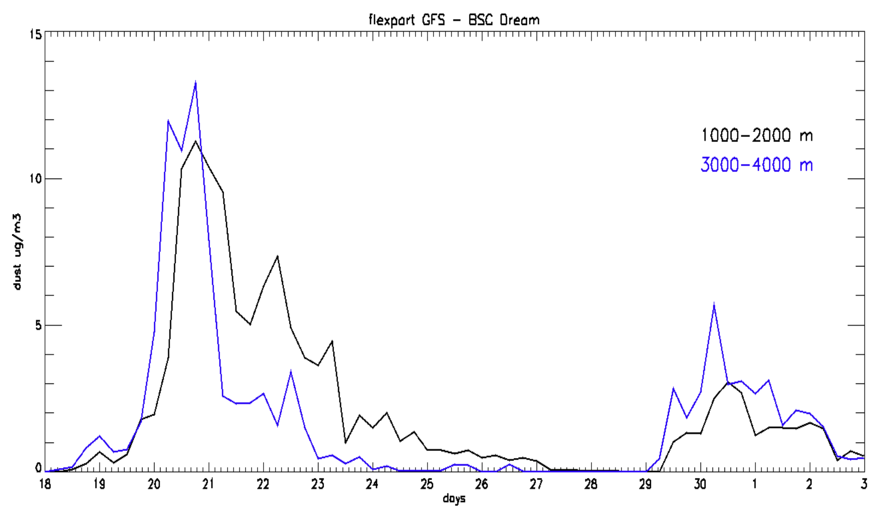
Footprint (residence time of the particles on each bin) from FLEXPART

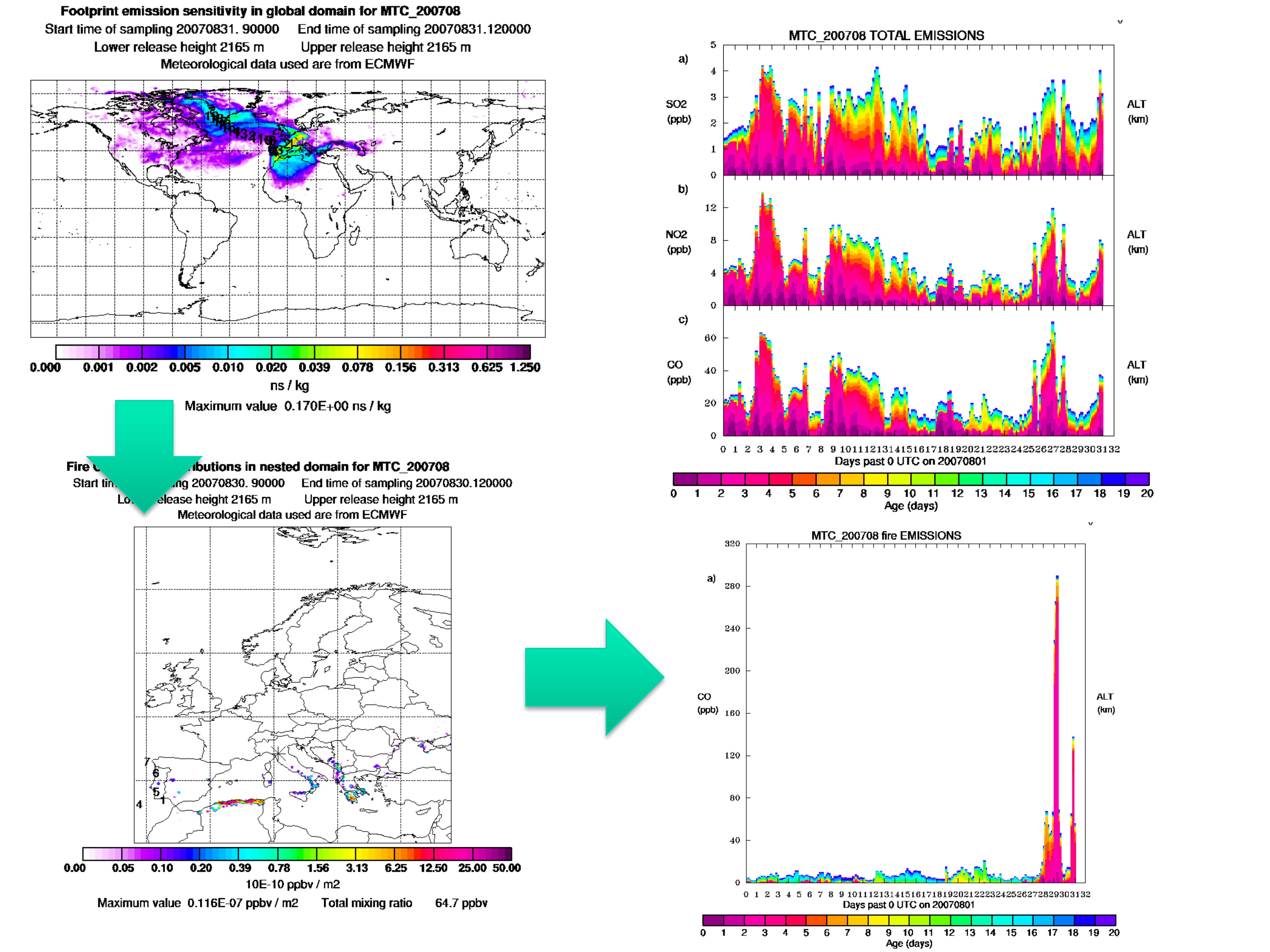


Dust emissions ng/m2/s



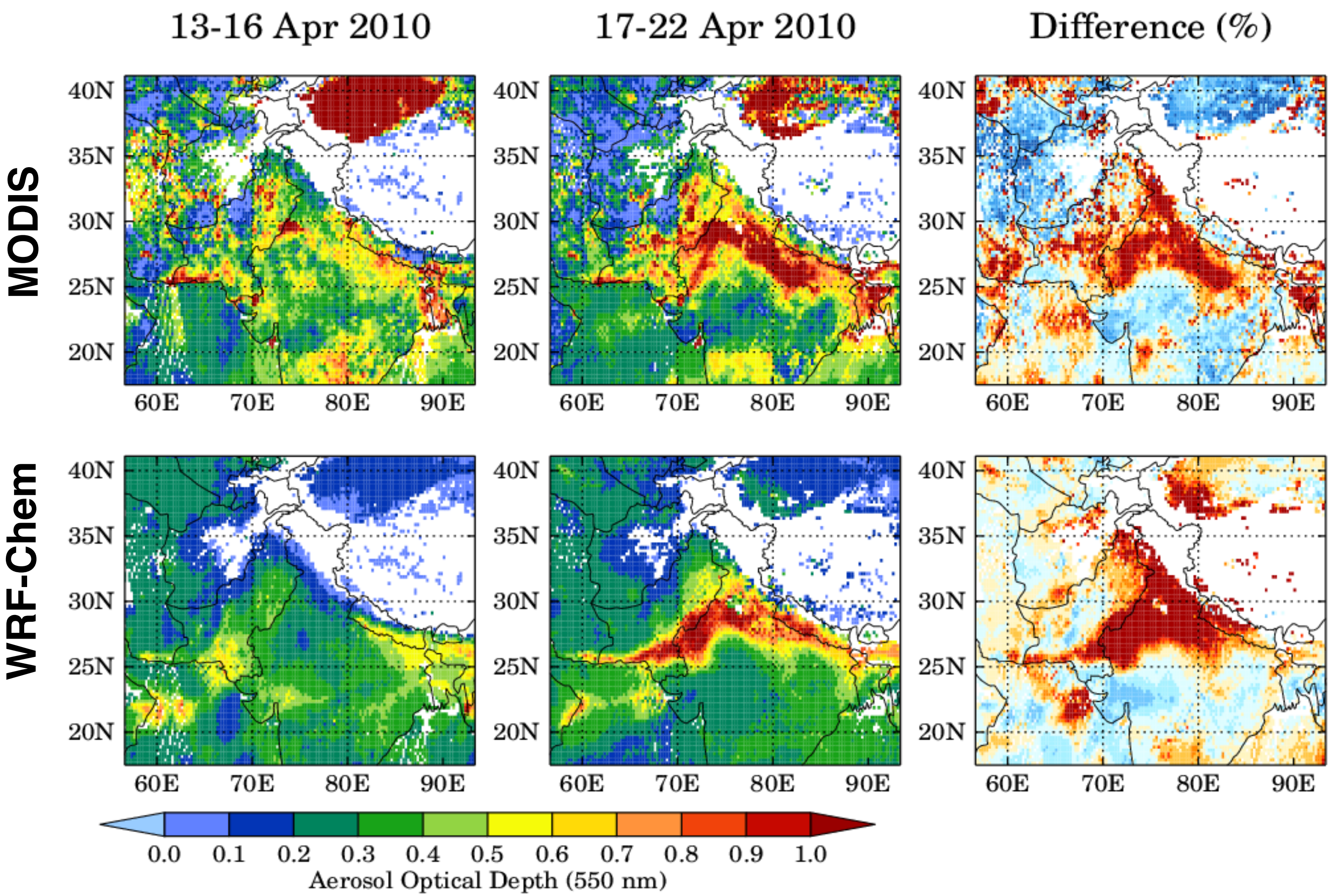
∫ (s*m2) X (ng/m2/s) = ng



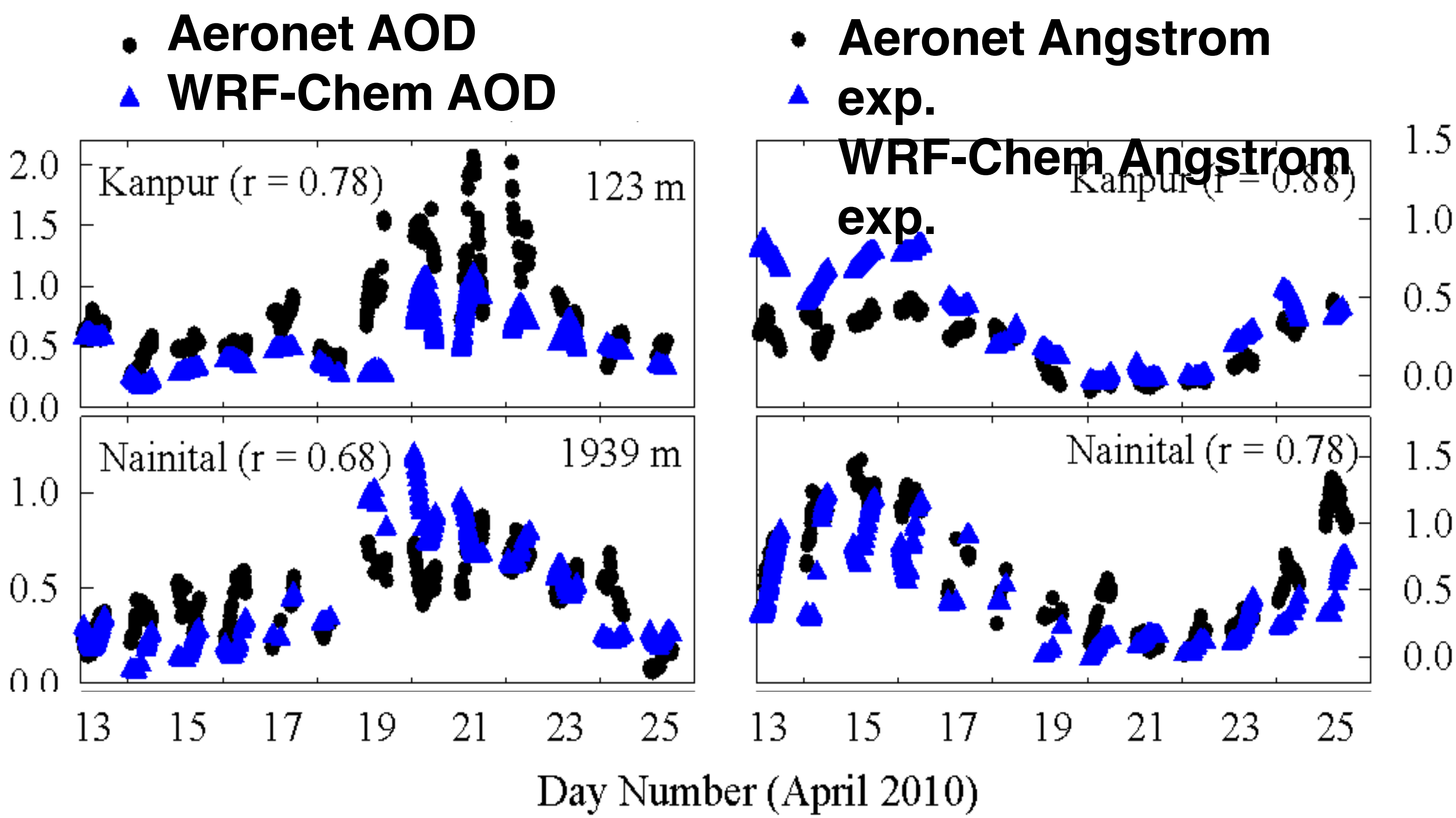


Footprint emission sensitivity in global domain for MTC_200708
Start time of sampling 20070831. 90000 End time of sampling 20070831.120000
Lower release height 2165 m Upper release height 2165 m
Meteorological data used are from ECMWF
0.000 0.001 0.002 0.005 0.010 0.020 0.039 0.078 0.156 0.313 0.625 1.250
ns / kg
Maximum value 0.170E+00 ns / kg
Fire C
butions in nested domain for MTC_200708
Start tim
ng 20070830. 90000 End time of sampling 20070830.120000
Lo
elease height 2165 m Upper release height 2165 m
Meteorological data used are from ECMWF
0.00 0.05 0.10 0.20 0.39 0.78 1.56 3.13 6.25 12.50 25.00 50.00
10E-10 ppbv / m2
Maximum value 0.116E-07 ppbv / m2 Total mixing ratio 64.7 ppbv
MTC_200708 TOTAL EMISSIONS
a) SO2 (ppb)
b) NO2 (ppb)
c) CO (ppb)
ALT (km)
Days past 0 UTC on 20070801
Age (days)
MTC_200708 fire EMISSIONS
a) CO (ppb)
ALT (km)
Days past 0 UTC on 20070801
Age (days)

Dust Storm – April 2010 (Kumar & Barth – see Mary's talk)
WRF & MODIS



WRF-Chem captures AOD and Angstrom exponent



AOD – integrated extinction coefficient over a vertical column of unit cross section.

Angstrom exponent – inverse relation with aerosol size, smaller for larger aerosols and vice versa.

[Kumar et al., ACPD, 2013]

We make use of FLEXPART on that event

The FLEXPART-WRF v3 is used
Can be downloaded and easy to compile under linux
Driven by WRF 3.2 3.3 outputs in ncdf format
Relatively fast

Backward cluster from Kanpur and Naintal

Estimate a footprint (from where air comes from)

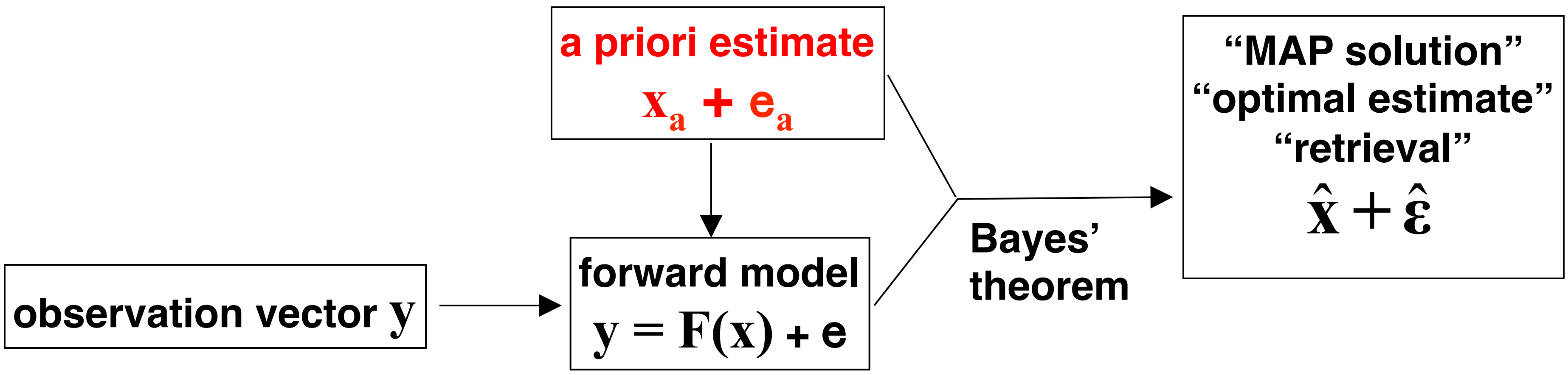
Couple with emissions (CO / Dust) to infer the emissive potential

Integrate – obtain a timeseries

Caveat: this is an exercise (put up in one day) – cannot pretend to be a scientific analysis

THE INVERSE MODELING PROBLEM

Optimize values of an ensemble of variables (*state vector* x) using observations:



THREE MAIN APPLICATIONS FOR ATMOSPHERIC COMPOSITION:

1. **Retrieve atmospheric concentrations (x) from observed atmospheric radiances (y) using a radiative transfer model as forward model**
2. **Invert sources (x) from observed atmospheric concentrations (y) using a CTM as forward model**
3. **Construct a continuous field of concentrations (x) by assimilation of sparse observations (y) using a forecast model (initial-value CTM) as forward model**

BAYES' THEOREM: FOUNDATION FOR INVERSE MODELS

$P(\mathbf{x})$ = probability distribution function (pdf) of $\mathbf{x}$

$P(\mathbf{x},\mathbf{y})$ = pdf of $(\mathbf{x},\mathbf{y})$

$P(\mathbf{y}|\mathbf{x})$ = pdf of $\mathbf{y}$ given $\mathbf{x}$

$$P(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y} = P(\mathbf{x})d\mathbf{x}P(\mathbf{y}\,|\,\mathbf{x})d\mathbf{y}$$

$$P(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y} = P(\mathbf{y})d\mathbf{y}P(\mathbf{x}\,|\,\mathbf{y})d\mathbf{x}$$

$$\Rightarrow \quad \underbrace{P(\mathbf{x}\,|\,\mathbf{y})}_{a\ posteriori\ \text{pdf}} = \frac{\overbrace{P(\mathbf{y}\,|\,\mathbf{x})}^{\text{observation } pdf}\ \overbrace{P(\mathbf{x})}^{a\ priori\ \text{pdf}}}{\underbrace{P(\mathbf{y})}_{\text{normalizing factor (unimportant)}}}$$

Bayes' theorem

Maximum *a posteriori* (MAP) solution for $\mathbf{x}$ given $\mathbf{y}$ is defined by $\max(P(\mathbf{x}\,|\,\mathbf{y}))$

$\Rightarrow$ solve for $\nabla_{\mathbf{x}} P(\mathbf{x}\,|\,\mathbf{y}) = \mathbf{0}$

SIMPLE LINEAR INVERSE PROBLEM FOR A SCALAR

use single measurement used to optimize a single source

Monitoring site measures concentration y

a priori bottom-up estimate

$x_a \pm s_a$

"Observational error" s_e { • instrument • fwd model

$y = kx \pm s_e$

Assume random Gaussian errors, let x be the true value. Bayes' theorem:

$$\ln P(x\,|\,y) : \ln P(y\,|\,x) + \ln P(x) : -\frac{(y-kx)^2}{\sigma_\varepsilon^2} - \frac{(x-x_a)^2}{\sigma_a^2}$$

Max of $P(x|y)$ is given by minimum of cost function $J(x) = \frac{(y-kx)^2}{\sigma_\varepsilon^2} + \frac{(x-x_a)^2}{\sigma_a^2}$

→ solve for $dJ/dx = 0$

Solution: $\boxed{\hat{x} = x_a + g(y - kx_a)}$ where g is a *gain factor* $g = \frac{k\sigma_a^2}{k^2\sigma_a^2 + \sigma_\varepsilon^2}$

Variance of solution: $\hat{\sigma}^2 = (\sigma_a^{-2} + (\sigma_\varepsilon / k)^{-2})^{-1}$

Alternate expression of solution: $y = kx + \varepsilon \Rightarrow \boxed{\hat{x} = ax + (1-a)x_a + g\varepsilon}$

where $a = gk$ is an *averaging kernel*

GENERALIZATION: CONSTRAINING *n* SOURCES WITH *m* OBSERVATIONS

Linear forward model: $$y_j = \sum_{i=1}^{n} k_{ij} x_i$$

A cost function defined as $$J(x_1,\ldots x_n) = \sum_{i=1}^{n} \frac{(x_i - x_{a,i})}{\sigma_{a,i}^2} + \sum_{j=1}^{m} \frac{(y_j - \sum_{i=1}^{n} k_{ij} x_i)}{\sigma_{\varepsilon,j}^2}$$

is generally not adequate because it does not account for correlation between sources or between observations. Need vector-matrix formalism:

$$\mathbf{x} = (x_1,\ldots x_n)^T \qquad \mathbf{y} = (y_1,\ldots y_m)^T \qquad \mathbf{y} = \mathbf{K}\mathbf{x} + \boldsymbol{\varepsilon}$$

Jacobian matrix

JACOBIAN MATRIX FOR FORWARD MODEL

Use of vector-matrix formalism requires linearization of forward model

Consider a non-linear forward model $\mathbf{y} = \mathbf{F}(\mathbf{x})$ and linearize it about $\mathbf{x_a}$:

$$\mathbf{y} = \mathbf{F}(\mathbf{x_a}) + \mathbf{K}(\mathbf{x} - \mathbf{x_a}) \quad (+\text{ higher-order terms})$$

$\mathbf{K} = \nabla_{\mathbf{x}}\mathbf{F} = \dfrac{\partial \mathbf{y}}{\partial \mathbf{x}}$ **with elements** $k_{ij} = \dfrac{\partial y_i}{\partial x_j}$ **is the Jacobian of $\mathbf{F}$ evaluated at $\mathbf{x_a}$**

$$\mathbf{K} = \begin{pmatrix} \partial y_1 / \partial x_1 & \mathrm{K} & \partial y_1 / \partial x_n \\ \mathrm{M} & \mathrm{O} & \mathrm{M} \\ \partial y_m / \partial x_1 & \mathrm{L} & \partial y_m / \partial x_n \end{pmatrix}$$

Construct Jacobian numerically column by column: perturb $\mathbf{x_a}$ by $\mathrm{D}x_i$, run forward model to get corresponding $\mathrm{D}\mathbf{y}$

If forward model is non-linear, $\mathbf{K}$ must be recalculated iteratively for successive solutions

$\mathbf{K}^T$ is the adjoint of the forward model (to be discussed later)

GAUSSIAN PDFs FOR VECTORS

A priori **pdf for** $\mathbf{x}$:

Scalar x

$$P(x) = \frac{1}{\sigma_a \sqrt{2\pi}} \exp[-\frac{(x - x_a)^2}{2\sigma_a^2}]$$

Vector $\mathbf{x} = (x_1, \ldots x_n)^T$

$$P(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} |\mathbf{S}_a|^{1/2}} \exp[-\frac{1}{2}(\mathbf{x} - \mathbf{x_a})^T \mathbf{S_a}^{-1}(\mathbf{x} - \mathbf{x_a})]$$

where $\mathbf{S_a}$ **is the** ***a priori*** **error covariance matrix describing error statistics on** $(\mathbf{x}\text{-}\mathbf{x_a})$

$$\mathbf{S_a} = \begin{pmatrix} \mathrm{var}(x_1 - x_{a,1}) & \mathrm{K} & \mathrm{cov}(x_1 - x_{a,1}, x_n - x_{a,n}) \\ \mathrm{M} & \mathrm{O} & \mathrm{M} \\ \mathrm{cov}(x_1 - x_{a,1}, x_n - x_{a,n}) & \mathrm{L} & \mathrm{var}(x_n - x_{a,n}) \end{pmatrix}$$

In log space:

$$-2\ln P(\mathbf{x}) = (\mathbf{x} - \mathbf{x_a})^T \mathbf{S_a^{-1}}(\mathbf{x} - \mathbf{x_a}) + c_1$$

OBSERVATIONAL ERROR COVARIANCE MATRIX

How well can the observing system constrain the *true value* of $\mathbf{x}$?

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) + \boldsymbol{\varepsilon}_{\mathbf{i}} + \boldsymbol{\varepsilon}_{\mathbf{m}}$$

observation; true value; instrument error; fwd model error

fwd model error, instrument error } observational error $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_{\mathbf{i}} + \boldsymbol{\varepsilon}_{\mathbf{m}}$

Observational error covariance matrix $\mathbf{S}_{\varepsilon} = \begin{pmatrix} \mathrm{var}(\varepsilon_1) & \mathrm{K} & \mathrm{cov}(\varepsilon_1, \varepsilon_n) \\ \mathrm{M} & \mathrm{O} & \mathrm{M} \\ \mathrm{cov}(\varepsilon_1, \varepsilon_n) & \mathrm{L} & \mathrm{var}(\varepsilon_n) \end{pmatrix}$

is the sum of the instrument and fwd model error covariance matrices:

$$\mathbf{S}_{\boldsymbol{\varepsilon}} = \mathbf{S}_{\boldsymbol{\varepsilon}_{\mathbf{i}}} + \mathbf{S}_{\boldsymbol{\varepsilon}_{\mathbf{m}}}$$

Corresponding pdf, in log space:

$$-2\ln P(\mathbf{y}\,|\,\mathbf{x}) = (\mathbf{y} - \mathbf{Kx})^T \mathbf{S}_{\varepsilon}^{-1} (\mathbf{y} - \mathbf{Kx}) + c_2$$

MAXIMUM A POSTERIORI (MAP) SOLUTION

$$-2\ln P(\mathbf{x}) = (\mathbf{x}-\mathbf{x_a})^T \mathbf{S_a}^{-1}(\mathbf{x}-\mathbf{x_a}) + c_1$$

$$-2\ln P(\mathbf{y}\,|\,\mathbf{x}) = (\mathbf{y}-\mathbf{Kx})^T \mathbf{S}_\varepsilon^{-1}(\mathbf{y}-\mathbf{Kx}) + c_2$$

Bayes' theorem:

$$-2\ln P(\mathbf{x}\,|\,\mathbf{y}) = \underbrace{(\mathbf{x}-\mathbf{x_a})^T \mathbf{S_a}^{-1}(\mathbf{x}-\mathbf{x_a})}_{\text{bottom-up constraint}} + \underbrace{(\mathbf{y}-\mathbf{Kx})^T \mathbf{S}_\varepsilon^{-1}(\mathbf{y}-\mathbf{Kx})}_{\text{top-down constraint}} + c_3$$

MAP solution: $\nabla_{\mathbf{x}} P(\mathbf{x}\,|\,\mathbf{y}) = \mathbf{0} \Rightarrow$ **miminize cost function** ***J*****:**

$$J(\mathbf{x}) = (\mathbf{x}-\mathbf{x_a})^T \mathbf{S_a}^{-1}(\mathbf{x}-\mathbf{x_a}) + (\mathbf{y}-\mathbf{Kx})^T \mathbf{S}_\varepsilon^{-1}(\mathbf{y}-\mathbf{Kx})$$

Solve for $\nabla_{\mathbf{x}} J(\mathbf{x}) = 2\mathbf{S_a}^{-1}(\mathbf{x}-\mathbf{x_a}) + 2\mathbf{K}^T \mathbf{S}_\varepsilon^{-1}(\mathbf{Kx}-\mathbf{y}) = \mathbf{0}$

Analytical solution:

$\hat{\mathbf{x}} = \mathbf{x_a} + \mathbf{G}(\mathbf{y}-\mathbf{Kx_a})$ **with gain matrix** $\mathbf{G} = \mathbf{S_a K}^T(\mathbf{KS_aK}^T + \mathbf{S}_\varepsilon)^{-1}$

$$\hat{\mathbf{S}} = (\mathbf{K}^T \mathbf{S}_\varepsilon^{-1}\mathbf{K} + \mathbf{S_a}^{-1})^{-1}$$

PARALLEL BETWEEN VECTOR-MATRIX AND SCALAR SOLUTIONS:

	Scalar problem	Vector-matrix problem
MAP solution:	$\hat{x} = x_a + g(y - kx_a)$	$\hat{\mathbf{x}} = \mathbf{x_a} + \mathbf{G}(\mathbf{y} - \mathbf{K}\mathbf{x_a})$
Gain factor:	$g = \dfrac{k\sigma_a^2}{k^2\sigma_a^2 + \sigma_\varepsilon^{2}}$	$\mathbf{G} = \mathbf{S_a}\mathbf{K}^T(\mathbf{K}\mathbf{S_a}\mathbf{K}^T + \mathbf{S}_\varepsilon)^{-1}$
A posteriori error:	$\hat{\sigma}^2 = (\sigma_a^{-2} + (\sigma_\varepsilon / k)^{-2})^{-1}$	$\hat{\mathbf{S}} = (\mathbf{K}^T\mathbf{S}_\varepsilon^{-1}\mathbf{K} + \mathbf{S_a}^{-1})^{-1}$
Averaging kernel:	$\hat{x} = ax + (1-a)x_a + g\varepsilon$	$\hat{\mathbf{x}} = \mathbf{A}\mathbf{x} + (\mathbf{I_n} - \mathbf{A})\mathbf{x_a} + \mathbf{G}\boldsymbol{\varepsilon}$
	$a = gk$	$\mathbf{A} = \mathbf{G}\mathbf{K}$

Jacobian matrix $\mathbf{K} = \partial\mathbf{y}/\partial\mathbf{x}$ **sensitivity of observations to true state**

Gain matrix $\mathbf{G} = \partial\hat{\mathbf{x}}/\partial\mathbf{y}$ **sensitivity of retrieval to observations**

Averaging kernel matrix $\mathbf{A} = \partial\hat{\mathbf{x}}/\partial\mathbf{x}$ **sensitivity of retrieval to true state**

A LITTLE MORE ON THE AVERAGING KERNEL MATRIX

$\mathbf{A}$ **describes the sensitivity of the retrieval to the true state**

$$\mathbf{A} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} = \begin{pmatrix} \partial \hat{x}_1 / \partial x_1 & \mathrm{K} & \partial \hat{x}_n / \partial x_1 \\ \mathrm{M} & \mathrm{O} & \mathrm{M} \\ \partial \hat{x}_1 / \partial x_n & \mathrm{L} & \partial \hat{x}_n / \partial x_n \end{pmatrix}$$

and hence the smoothing of the solution:

$$\hat{\mathbf{x}} = \mathbf{A}\mathbf{x} + \underbrace{(\mathbf{I_n} - \mathbf{A})\mathbf{x_a}}_{\text{smoothing error}} + \underbrace{\mathbf{G}\boldsymbol{\varepsilon}}_{\text{retrieval error}}$$

MAP retrieval gives $\mathbf{A}$ as part of the retrieval:

$$\mathbf{A} = \mathbf{G}\mathbf{K} = \mathbf{S_a}\mathbf{K}^T(\mathbf{K}\mathbf{S_a}\mathbf{K}^T + \mathbf{S}_\varepsilon)^{-1}\mathbf{K}$$

Other retrieval methods (e.g., neural network, adjoint method) do *not* provide $\mathbf{A}$

 # pieces of info in a retrieval = degrees of freedom for signal (DOFS) = $\mathrm{trace}(\mathbf{A})$

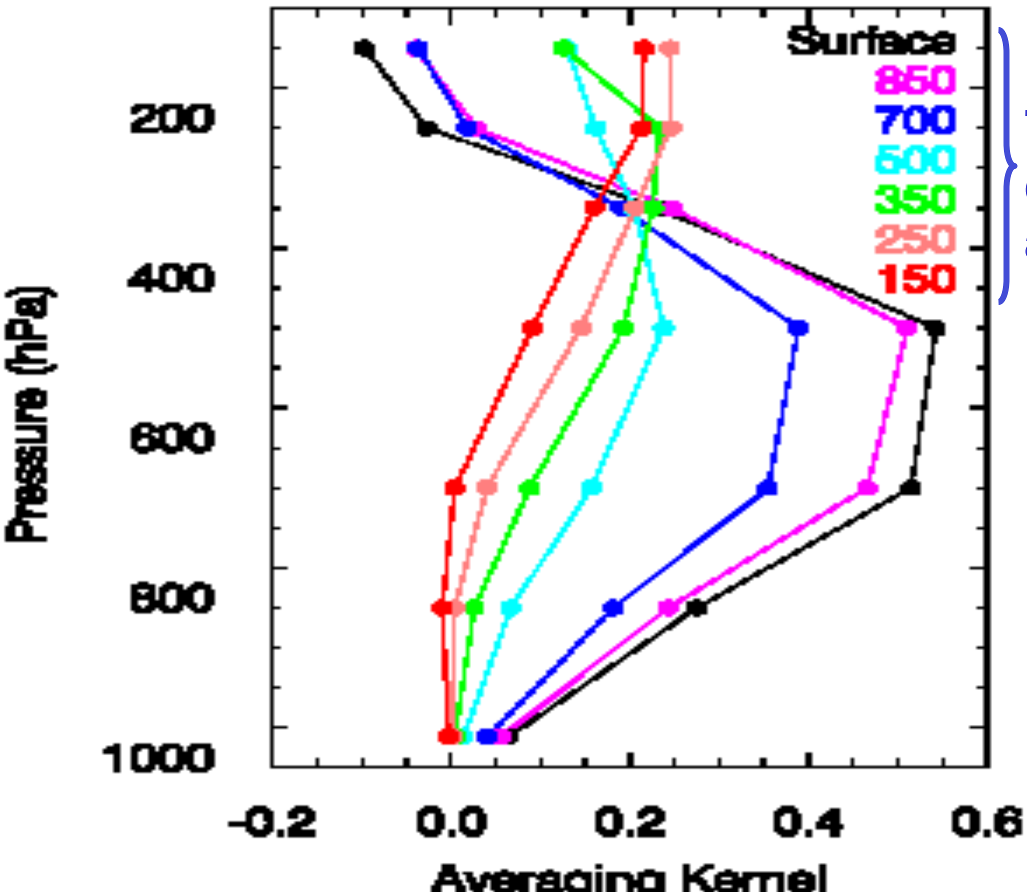
APPLICATION TO SATELLITE RETRIEVALS

Here $\mathbf{y}$ is the vector of wavelength-dependent radiances (radiance spectrum);

$\mathbf{x}$ is the state vector of concentrations;

forward model $\mathbf{y} = \mathbf{F}(\mathbf{x})$ is the radiative transfer model

Illustrative MOPITT averaging kernel matrix for CO retrieval



MOPITT retrieves concentrations at 7 pressure levels; lines are the corresponding columns of the averaging kernel matrix

trace(A) = 1.5 in this case; 1.5 pieces of information

Analytical solution to inverse problem

$$\nabla_{\mathbf{x}} J(\mathbf{x}) = 2\mathbf{S}_{\mathbf{a}}^{-1}(\mathbf{x}-\mathbf{x}_{\mathbf{a}}) + 2\mathbf{K}^T\mathbf{S}_{\varepsilon}^{-1}(\mathbf{K}\mathbf{x} - \mathbf{y}) = \mathbf{0}$$

$$\Rightarrow \quad \hat{\mathbf{x}} = \mathbf{x}_{\mathbf{a}} + \mathbf{S}_{\mathbf{a}}\mathbf{K}^T(\mathbf{K}\mathbf{S}_{\mathbf{a}}\mathbf{K}^T + \mathbf{S}_{\varepsilon})^{-1}(\mathbf{y} - \mathbf{K}\mathbf{x}_{\mathbf{a}})$$

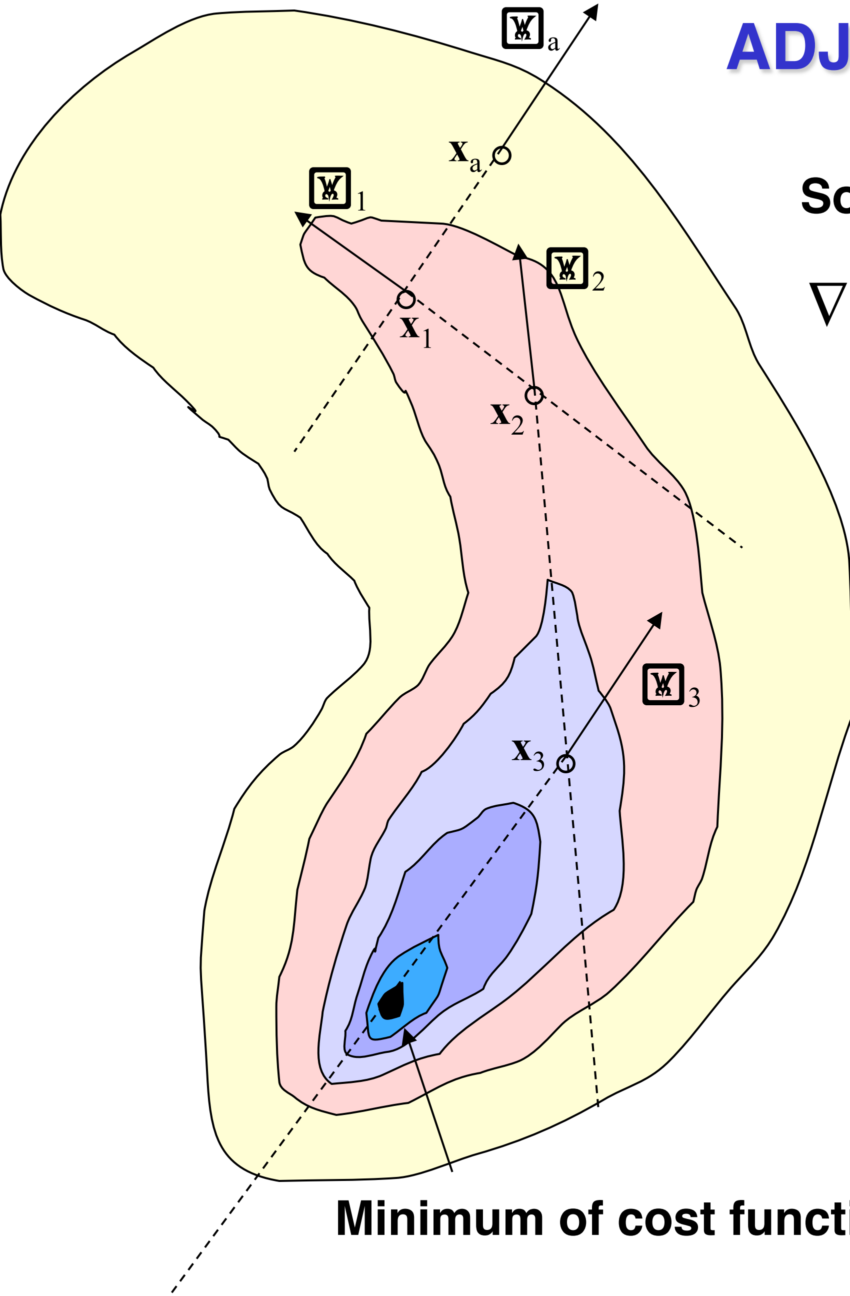
requires (iterative) numerical construction of the Jacobian matrix $\mathbf{K}$ and matrix operations of dimension *(mxn)*; this limits the size of *n*, i.e., the number of variables that you can optimize

Address this limitation with Kalman filter (for time-dependent $\mathbf{x}$) or with adjoint method

BASIC KALMAN FILTER
to optimize time-dependent state vector

time	observations	state vector
		a priori $\mathbf{x}_{a,0} \pm \mathbf{S}_{a,0}$
t_o	$\mathbf{y}_0$ →	$\hat{\mathbf{x}}_0 \pm \hat{\mathbf{S}}_0$
		Evolution model M for [t_0, t_1]: $\hat{\mathbf{x}}_{a,1} = \mathbf{M}\hat{\mathbf{x}}_0 \pm \mathbf{S}_\mathbf{M}$; $\mathbf{S}_{a,1} = \mathbf{M}\hat{\mathbf{S}}_0\mathbf{M}^T + \mathbf{S}_\mathbf{M}$
t_1	$\mathbf{y}_1$ →	$\hat{\mathbf{x}}_1 \pm \hat{\mathbf{S}}_1$
	etc.	**Apply evolution model for [t_1, t_2]…**

ADJOINT INVERSION (4-D VAR)



Minimum of cost function *J*

Solve

$$\nabla_{\mathbf{x}} J(\mathbf{x}) = 2\mathbf{S}_{\mathbf{a}}^{-1}(\mathbf{x}-\mathbf{x}_{\mathbf{a}}) + 2\mathbf{K}^T\mathbf{S}_{\varepsilon}^{-1}(\mathbf{F}(\mathbf{x}) - \mathbf{y}) = \mathbf{0}$$

numerically rather than analytically

1. **Starting from *a priori* $\mathbf{x}_{\mathbf{a}}$, calculate $\nabla_{\mathbf{x}} J(\mathbf{x}_{\mathbf{a}})$**
2. **Using an optimization algorithm (BFGS), get next guess $\mathbf{x}_1$**
3. **Calculate $\nabla_{\mathbf{x}} J(\mathbf{x}_1)$, get next guess $\mathbf{x}_2$**
4. **Iterate until convergence**

NUMERICAL CALCULATION OF COST FUNCTION GRADIENT

$$\nabla_{\mathbf{x}} J(\mathbf{x}) = 2\mathbf{S}_{\mathbf{a}}^{-1}(\mathbf{x}-\mathbf{x}_{\mathbf{a}}) + 2\mathbf{K}^T \underbrace{\mathbf{S}_{\varepsilon}^{-1}(\mathbf{F}(\mathbf{x})-\mathbf{y})}_{\text{"adjoint forcing"}}$$

adjoint → $\mathbf{K}^T$

Adjoint model is applied to error-weighted difference between model and obs

…but we want to avoid explicit construction of $\mathbf{K}$

Construct *tangent linear model* $\partial \mathbf{y}_{(i)} / \partial \mathbf{y}_{(i-1)}$ of forward model describing evolution of concentration field over time interval $[t_{i-1}, t_i]$

Sensitivity of $\mathbf{y}_{(i)}$ to $\mathbf{x}_{(0)}$ at time t_0 can then be written

$$\mathbf{K} = \frac{\partial \mathbf{y}_{(i)}}{\partial \mathbf{x}_{(0)}} = \frac{\partial \mathbf{y}_{(i)}}{\partial \mathbf{y}_{(i-1)}} \frac{\partial \mathbf{y}_{(i-1)}}{\partial \mathbf{y}_{(i-2)}} \cdots \frac{\partial \mathbf{y}_{(1)}}{\partial \mathbf{y}_{(0)}} \frac{\partial \mathbf{y}_{(0)}}{\partial \mathbf{x}_{(0)}}$$

and since $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$,

$$\mathbf{K}^T = \left(\frac{\partial \mathbf{y}_{(i)}}{\partial \mathbf{y}_{(i-1)}} \frac{\partial \mathbf{y}_{(i-1)}}{\partial \mathbf{y}_{(i-2)}} \cdots \frac{\partial \mathbf{y}_{(1)}}{\partial \mathbf{y}_{(0)}} \frac{\partial \mathbf{y}_{(0)}}{\partial \mathbf{x}_{(0)}} \right)^T = \left(\frac{\partial \mathbf{y}_{(0)}}{\partial \mathbf{x}_{(0)}} \right)^T \left(\frac{\partial \mathbf{y}_{(1)}}{\partial \mathbf{y}_{(0)}} \right)^T \cdots \left(\frac{\partial \mathbf{y}_{(n-1)}}{\partial \mathbf{y}_{(n-2)}} \right)^T \left(\frac{\partial \mathbf{y}_{(n)}}{\partial \mathbf{y}_{(n-1)}} \right)^T$$

Apply transpose of tangent linear model to the adjoint forcings; for time interval $[t_0, t_n]$, start from observations at t_n and work backward in time until t_0, picking up new observations (adjoint forcings) along the way.