

Data Assimilation  
Lecture 2  
"Advanced" Techniques  
+  
Chemical Assimilation

Mike Fisher

ECMWF

Slide 1



Variational Data Assimilation: 3d-Var & 4d-Var

- I introduced variational assimilation in my last lecture.
- The analysis is found by minimizing the "cost function":

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T (\mathbf{P}^b)^{-1} (\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{H}\mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x} - \mathbf{y})$$

- So far, we have assumed that  $\mathbf{x}$  and  $\mathbf{y}$  describe the state of the system at a single time.
  - $\mathbf{H}$  represents spatial interpolation, and transformation from state variables to observed variables.
- This method is usually called 3d-Var.
  - Because: in many applications,  $\mathbf{x}$  represents a three-dimensional state

Slide 2



Variational Data Assimilation: 3d-Var & 4d-Var

- Suppose now that the observations are not all made at the same time.
- This is not really a problem, provided we have some way to propagate the state forward in time.
  - E.g. a numerical forecast model.
- We take  $\mathbf{x}$  to be defined at some initial time, and  $\mathbf{H}$  then consists of the following steps:
  1. Use the forecast model to propagate the state to the times of the observation.
  2. Interpolate spatially to the observation location
  3. Transform from state variables (e.g. T) to observed variables (e.g. radiance).
- This is called 4d-Var: the extra dimension is time.

Slide 3



Variational Data Assimilation: 3d-Var & 4d-Var

- Formally, we can still write 4d-Var as:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^T (\mathbf{P}^b)^{-1} (\mathbf{x}_0 - \mathbf{x}^b) + \frac{1}{2}(\mathbf{H}\mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x} - \mathbf{y})$$

- But it is often written with time made explicit:

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^T (\mathbf{P}^b)^{-1} (\mathbf{x}_0 - \mathbf{x}^b) + \frac{1}{2} \sum_{k=0}^K (\mathbf{H}_k \mathbf{x}_k - \mathbf{y}_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{x}_k - \mathbf{y}_k)$$

- Where  $\mathbf{x}_k$  is valid at time  $t_k$ .
- $\mathbf{x}_k$  is determined by integrating the model:  $\mathbf{x}_k = \mathbf{M}_{0 \rightarrow k} \mathbf{x}_0$
- NB: we can evaluate  $(\mathbf{H}\mathbf{x} - \mathbf{y})$  using a single integration of the model, by interpolating to observation locations at the appropriate model timestep.

Slide 4



### Variational Data Assimilation: 3d-Var & 4d-Var

- Remember: when we developed the analysis equations, we assumed that H was error-free.
- In 4d-Var, H includes an integration of a numerical forecast model.
- Clearly, since models are imperfect, we can only regard H as error free over a limited period of time:
  - In NWP, we can get away with assuming H is perfect up to ~12 hours ahead.
- To extend the time "window" beyond 12h, we need to take account of the fact that the model is imperfect.

Slide 5



### Variational Data Assimilation: 3d-Var & 4d-Var

- The standard 4d-Var method is sometimes called "strong-constraint" 4d-Var.

- We minimize

$$J(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^T (\mathbf{P}^b)^{-1} (\mathbf{x}_0 - \mathbf{x}^b) + \frac{1}{2} \sum_{k=0}^K (\mathbf{H}_k \mathbf{x}_k - \mathbf{y}_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{x}_k - \mathbf{y}_k)$$

- Subject to the strong constraint:  $\mathbf{x}_k = \mathbf{M}_{0 \rightarrow k} \mathbf{x}_0$

- Note that J can be regarded as a function of  $\mathbf{x}_0$  only, because we can determine  $\mathbf{x}_k$  by integrating the model.
- In weak-constraint 4d-Var, we relax the constraint, and no longer require that  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_K$  is an exact "trajectory" of the model.

Slide 6



### Variational Data Assimilation: 3d-Var & 4d-Var

- In weak-constraint 4d-Var we minimize:

$$J(\mathbf{x}_0, \dots, \mathbf{x}_K) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}^b)^T (\mathbf{P}^b)^{-1} (\mathbf{x}_0 - \mathbf{x}^b) + \frac{1}{2} \sum_{k=0}^K (\mathbf{H}_k \mathbf{x}_k - \mathbf{y}_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{x}_k - \mathbf{y}_k) + \frac{1}{2} \sum_{k=1}^K (\mathbf{M}_{(k-1) \rightarrow k} \mathbf{x}_{k-1} - \mathbf{x}_k)^T \mathbf{Q}_k^{-1} (\mathbf{M}_{(k-1) \rightarrow k} \mathbf{x}_{k-1} - \mathbf{x}_k)$$

- Note that J is now a function of  $\mathbf{x}_0, \dots, \mathbf{x}_K$
- The extra term in the cost function penalises, but does not prohibit departures of  $\mathbf{x}_0, \dots, \mathbf{x}_K$  from a trajectory of the model.
- What does this term mean statistically?

Slide 7



### Weak-Constraint 4d-Var

- Remember: we derived strong-constraint 4d-Var using Bayes' theorem:

$$P(\mathbf{x} | \mathbf{y}, \mathbf{x}^b) \propto P(\mathbf{x}^b | \mathbf{x}) P(\mathbf{y} | \mathbf{x}) P(\mathbf{x})$$

- We assumed that  $P(\mathbf{x}) = \text{const}$ 
  - I.e. no prior knowledge
- In weak-constraint 4d-Var, we use our prior knowledge that  $\mathbf{x}$  should be close to a trajectory of the model.
- Define the model error at each timestep as:

$$\boldsymbol{\eta}_k = \mathbf{x}_k^t - \mathbf{M}_{(k-1) \rightarrow k} \mathbf{x}_{k-1}^t$$

Slide 8



## Weak-Constraint 4d-Var

- We will assume that model error is random, Gaussian, and “white” in time:

$$P(\boldsymbol{\eta}_k) \propto \exp\left(-\frac{1}{2}\boldsymbol{\eta}_k^T \mathbf{Q}_k^{-1} \boldsymbol{\eta}_k\right)$$

$$= \exp\left(-\frac{1}{2}\left(\mathbf{x}_k - \mathbf{M}_{(k-1) \rightarrow k} \mathbf{x}_{k-1}\right)^T \mathbf{Q}_k^{-1} \left(\mathbf{x}_k - \mathbf{M}_{(k-1) \rightarrow k} \mathbf{x}_{k-1}\right)\right)$$

- Define  $P(\mathbf{x}) = P(\eta_1) P(\eta_2) \dots P(\eta_k)$
- Taking  $-\log(P(\mathbf{x}))$  gives the extra, quadratic term on the cost function:

$$\frac{1}{2} \sum_{k=1}^K \left(\mathbf{M}_{(k-1) \rightarrow k} \mathbf{x}_{k-1} - \mathbf{x}_k\right)^T \mathbf{Q}_k^{-1} \left(\mathbf{M}_{(k-1) \rightarrow k} \mathbf{x}_{k-1} - \mathbf{x}_k\right)$$

Slide 9



## The Kalman Filter

- Let’s go back to the “OI-style” analysis equations.
- Typically, we want to make a sequence of analyses:

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^b)$$

$$\mathbf{K}_k = \mathbf{P}_k^b \mathbf{H}_k^T \left[ \mathbf{R}_k + \mathbf{H}_k \mathbf{P}_k^b \mathbf{H}_k^T \right]^{-1}$$

- For each analysis, the background will be provided by a short forecast from the previous analysis:

$$\mathbf{x}_k^b = \mathbf{M}_{(k-1) \rightarrow k} \mathbf{x}_{k-1}^a$$

- How does the covariance matrix  $\mathbf{P}_k^b$  propagate from analysis to analysis?

Slide 10



## The Kalman Filter

- The covariance matrix for analysis errors can easily be derived from:

$$\boldsymbol{\varepsilon}_k^a = \boldsymbol{\varepsilon}_k^b + \mathbf{K}_k (\boldsymbol{\varepsilon}_k^o - \mathbf{H}_k \boldsymbol{\varepsilon}_k^b)$$

- One (not very useful) expression is:

$$\mathbf{P}_k^a = \mathbf{K}_k \left[ \mathbf{R}_k + \mathbf{H}_k \mathbf{P}_k^b \mathbf{H}_k^T \right] \mathbf{K}_k^T - \mathbf{K}_k \mathbf{H}_k \mathbf{P}_k^b - \mathbf{P}_k^b \mathbf{H}_k^T \mathbf{K}_k^T + \mathbf{P}_k^b$$

- For practical purposes, it’s best to use an expression that guarantees the matrix stays positive definite and symmetric:

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^b (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T$$

- The simplest expression (no good for numerical work) is:

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^b$$

Slide 11



## The Kalman Filter

- We propagate the state using:  $\mathbf{x}_k^b = \mathbf{M}_{(k-1) \rightarrow k} \mathbf{x}_{k-1}^a$

- Subtract the true state  $\mathbf{x}_k^t$  at step  $k$  from both sides:

$$\boldsymbol{\varepsilon}_k^b = \mathbf{M}_{(k-1) \rightarrow k} \boldsymbol{\varepsilon}_{k-1}^a - \left( \mathbf{x}_k^t - \mathbf{M}_{(k-1) \rightarrow k} \mathbf{x}_{k-1}^t \right)$$

$$= \mathbf{M}_{(k-1) \rightarrow k} \boldsymbol{\varepsilon}_{k-1}^a - \boldsymbol{\eta}_k$$

- $\boldsymbol{\eta}_k$  is model error. We will assume that it is random, Gaussian, white in time, and uncorrelated with analysis error:

$$\overline{\boldsymbol{\varepsilon}_k^b (\boldsymbol{\varepsilon}_k^b)^T} = \mathbf{M}_{(k-1) \rightarrow k} \overline{\boldsymbol{\varepsilon}_{k-1}^a (\boldsymbol{\varepsilon}_{k-1}^a)^T} \mathbf{M}_{(k-1) \rightarrow k}^T + \overline{\boldsymbol{\eta}_k \boldsymbol{\eta}_k^T}$$

- i.e.  $\mathbf{P}_k^b = \mathbf{M}_{(k-1) \rightarrow k} \mathbf{P}_{k-1}^a \mathbf{M}_{(k-1) \rightarrow k}^T + \mathbf{Q}_k$

Slide 12



## The Kalman Filter

- To summarize...

$$\begin{aligned} \mathbf{x}_k^b &= \mathbf{M}_{(k-1) \rightarrow k} \mathbf{x}_{k-1}^a \\ \mathbf{P}_k^b &= \mathbf{M}_{(k-1) \rightarrow k} \mathbf{P}_{k-1}^a \mathbf{M}_{(k-1) \rightarrow k}^T + \mathbf{Q}_k \\ \mathbf{x}_k^a &= \mathbf{x}_k^b + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^b) \\ \mathbf{K}_k &= \mathbf{P}_k^b \mathbf{H}_k^T [\mathbf{R}_k + \mathbf{H}_k \mathbf{P}_k^b \mathbf{H}_k^T]^{-1} \\ \mathbf{P}_k^a &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^b \end{aligned}$$

Slide 13



## The Relationship Between 4dVar and the Kalman Filter

- It can be shown that weak constraint 4dVar gives **exactly the same analysis** at the end of the assimilation window (i.e. at time  $t_k$ ) as a Kalman filter initialized with the same covariance matrix  $\mathbf{P}^b$  at the start of the window.
  - More generally, the weak constraint 4dVar solution for  $t_0 \leq t \leq t_k$  is exactly the same as that given by the Kalman smoother.
- The proof is tedious!
  - Write  $\mathcal{J}(x_0, \dots, x_k)$  as  $\mathcal{J}(x_0, \dots, x_k; \eta_0, \dots, \eta_k)$
  - Minimize subject to the constraints:  $\eta_k = \mathbf{x}_k^i - \mathbf{M}_{(k-1) \rightarrow k} \mathbf{x}_{k-1}^i$
  - => Lagrange multipliers for the constraints:  $\lambda_k$
  - Variation w.r.t.  $\eta$ 's gives the Kalman Filter equations
  - Variation w.r.t.  $\mathbf{x}$ 's gives the backward pass of the Kalman smoother.
  - This technique is called Pontryagin minimization

Slide 14



## The Relationship Between 4dVar and the Kalman Filter

- But, we lose a lot of the advantage of the Kalman filter if we re-initialize the covariance matrix every few hours.
- It would seem that the Kalman filter is better than 4dVar.
- But, we cannot solve the Kalman filter equations for large systems.
  - The matrices are too large
  - E.g.  $\mathbf{P}^b = \mathbf{M} \mathbf{P}^a \mathbf{M}^T + \mathbf{Q}$  would require  $\sim 10^8$  integrations of  $\mathbf{M}$ .

- What can we do?

Slide 15



## Ensemble Kalman Filtering

- One attractive approach is to approximate  $\mathbf{P}^b$  by a sample covariance matrix:

$$\mathbf{P}_k^b = \frac{1}{K} \sum_{i=1}^K (\mathbf{x}_k^{b(i)} - \bar{\mathbf{x}}_k^b) (\mathbf{x}_k^{b(i)} - \bar{\mathbf{x}}_k^b)^T$$

- where  $\{\mathbf{x}^{b(i)} \dots \mathbf{x}^{b(K)}\}$  is an **ensemble** of states
- The ensemble of background states is found by propagating an ensemble of analysis states, and adding noise to represent the effect of model error.
- The ensemble of analysis states is produced by perturbing the observations by adding random noise with the statistical characteristics of observation error.

Slide 16



## The Kalman Filter

- The (basic) ensemble Kalman filter:

$$\mathbf{x}_k^{b(i)} = \mathbf{M}_{(k-1) \rightarrow k} \mathbf{x}_{k-1}^{a(i)} + \boldsymbol{\eta}_k$$

$$\mathbf{P}_k^b \approx \frac{1}{M} \sum_{i=1}^M \left( \mathbf{x}_k^{b(i)} - \overline{\mathbf{x}}_k^b \right) \left( \mathbf{x}_k^{b(i)} - \overline{\mathbf{x}}_k^b \right)^T$$

$$\mathbf{x}_k^{a(i)} = \mathbf{K}_k \left( \mathbf{y}_k + \boldsymbol{\zeta}_k - \mathbf{H}_k \mathbf{x}_k^{b(i)} \right)$$

$$\mathbf{P}_k^a \approx \frac{1}{M} \sum_{i=1}^M \left( \mathbf{x}_k^{a(i)} - \overline{\mathbf{x}}_k^a \right) \left( \mathbf{x}_k^{a(i)} - \overline{\mathbf{x}}_k^a \right)^T$$

$$\mathbf{K}_k = \mathbf{P}_k^b \mathbf{H}_k^T \left[ \mathbf{R}_k + \mathbf{H}_k \mathbf{P}_k^b \mathbf{H}_k^T \right]^{-1}$$

Slide 17 ECMWF

## Ensemble Methods

- To construct the gain matrix, we also use sample covariance matrices:

$$\mathbf{P}_k^b \mathbf{H}_k^T \approx \frac{1}{K} \sum_{i=1}^K \left( \mathbf{x}_k^{b(i)} - \overline{\mathbf{x}}_k^b \right) \left( H_k \left( \mathbf{x}_k^{b(i)} \right) - \overline{H_k \left( \mathbf{x}_k^b \right)} \right)^T$$

$$\mathbf{H}_k \mathbf{P}_k^b \mathbf{H}_k^T \approx \frac{1}{K} \sum_{i=1}^K \left( H_k \left( \mathbf{x}_k^{b(i)} \right) - \overline{H_k \left( \mathbf{x}_k^b \right)} \right) \left( H_k \left( \mathbf{x}_k^{b(i)} \right) - \overline{H_k \left( \mathbf{x}_k^b \right)} \right)^T$$

Slide 18 ECMWF

## Ensemble Methods

- Ensemble methods are very popular
  - Relatively easy to develop
  - Some advantages for nonlinear systems (but the analysis equation is still linear!)
- But:
  - They are definitely not optimal! (The covariance matrices are approximated.)
  - They require careful tweaking to avoid the worst effects of using very low-rank estimate of the covariance matrices.
    - o E.g. localization operators to remove spurious long-range correlations
    - o Ensembles sizes may need to be quite large to get analyses comparable to those given by 4dVar

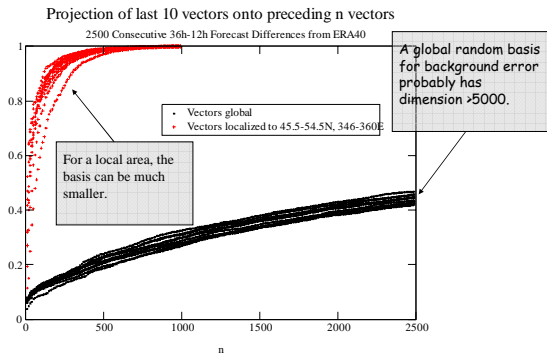
Slide 19 ECMWF

## How Big an Ensemble is Required?

- Ensemble size is usually determined by what we can afford to run.
- We can get some idea by estimating the size of ensemble required to approximate a static covariance matrix.
- The NMC method (Parrish and Derber, 1992) is a popular method for estimating background error covariances.
  - Forecast differences are used as a surrogate for background error.
- We consider a large sample of differences between pairs of 36h and 12h forecasts for the same date.
  - The forecasts were taken from the ECMWF 40-year re-analysis, and truncated to T42.

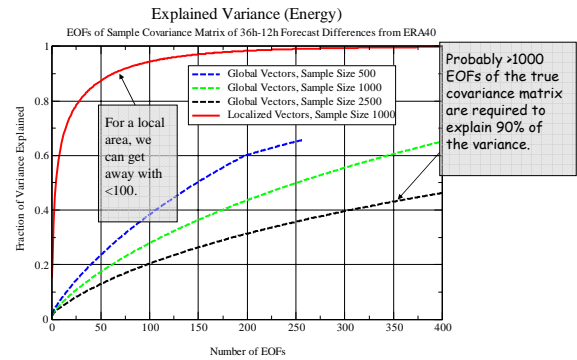
Slide 20 ECMWF

## How Big an Ensemble?



Slide 21 ECMWF

## How Big a Subspace?



Slide 22 ECMWF

## Long-Window 4dVar

- An alternative to the ensemble Kalman filter (and similar low-rank methods) is to use 4dVar.
- But, we must reduce the impact of frequent re-initialization of the covariance matrix.
- We can do this, by making the analysis “window” longer:
  - E.g. several days, instead of a few hours.
  - For NWP, we still require an analysis every few hours, so this approach implies overlapping analysis windows.

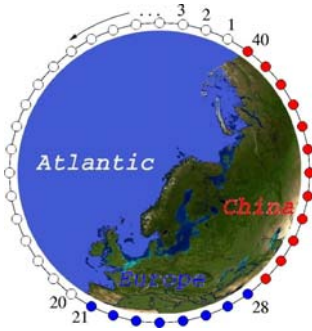
Slide 23 ECMWF

## Long-Window 4dVar

- In the Kalman filter, unstable dynamics and model error produce a limited-memory effect.
  - The analysis at time  $t$  is effectively independent of the state and observations more than a few days ago.
- But, 4dVar and the Kalman filter are equivalent over the assimilation window.
- So, for a 4dVar window of length a few days, the analysis at the end of the window is independent of the state, observations and covariance at the beginning of the window.
- Making the window even longer has no effect.
- For a long enough window, the analysis is identical to that of a long-running full-rank, un-approximated Kalman filter.

Slide 24 ECMWF

### Martin Leutbecher's "Planet L95" EKF



$$\frac{dx_i}{dt} = x_{i-2}x_{i-1} + x_{i-1}x_{i+1} - x_i + F$$

with  $i = 1, 2, \dots, 40$   
 $x_0 = x_{40}$   
 $x_{-1} = x_{39}$   
 $x_{41} = x_1$   
 $F = 8$

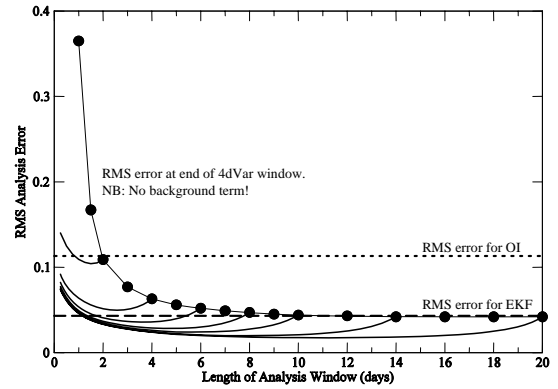
(Lorenz, 1995, ECMWF Seminar on Predictability, and Lorenz and Emanuel, 1998)

unit time ~ 5 days

Chaotic system: 13 positive Lyapunov exponents.

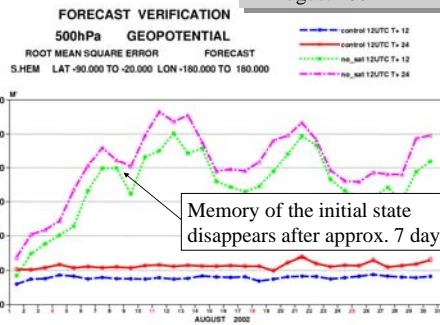
The largest exponent corresponds to a doubling time of 2.1 days.

### Mean RMS Analysis Error



### More Evidence

Analysis experiments started with/without satellite data on 1<sup>st</sup> August 2002



from: Graeme Kelly

### Chemistry

#### • What can assimilation bring to atmospheric chemistry?

- Measurement of trace-species concentrations is difficult!
- Only a subset of chemical species can be observed.
- For some unobserved species, we can use equilibrium relationships to determine their concentrations.

o E.g. daytime  $\text{NO}_x$  partitioning in the stratosphere:

$$\frac{[\text{NO}]}{[\text{NO}_2]} = \frac{J_{\text{NO}_2} + k_{\text{NO}_2+\text{O}}[\text{O}]}{k_{\text{NO}+\text{O}_3}[\text{O}_3] + k_{\text{NO}+\text{ClO}}[\text{ClO}]}$$

- But, there are plenty of longer-lived species for which equilibrium relationships cannot be used:

o Stratospheric  $\text{N}_2\text{O}_5$  production/loss has time constant of several hours. It is not in equilibrium, nor is it long-lived. Its concentration is influenced by many species.

## Chemistry

### • What can assimilation bring to atmospheric chemistry?

- For many species, time evolution provides useful information.

- o During the day: 
$$\frac{[\text{NO}]}{[\text{NO}_2]} = \frac{J_{\text{NO}_2} + k_{\text{NO}_2+\text{O}}[\text{O}]}{k_{\text{NO}+\text{O}_3}[\text{O}_3] + k_{\text{NO}+\text{ClO}}[\text{ClO}]}$$
- o At night, NO production stops, and NO is converted to NO<sub>2</sub>.
- o So, day-night variation of NO<sub>2</sub> tells us something about NO, and indirectly about the concentrations of O<sub>3</sub>, O and ClO.

- Production and loss rates also provide useful information:

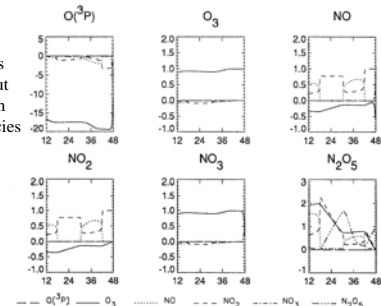
- o At night, N<sub>2</sub>O<sub>5</sub> concentration increases with a timescale of several hours. The production rate depends on O<sub>3</sub>.
- Data assimilation (e.g. 4d-Var) can make use of such information.

Slide 29



## Chemistry

Sensitivity of the analysed concentrations of all species throughout the 4dVar window to an observation of one species at t=48h (midnight)



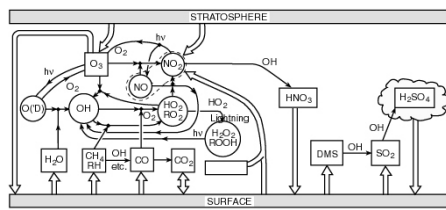
(Fisher & Lary, 1995)

Figure 5. Graphs of the influence function as for Fig. 4, but for observations at midnight.

Slide 30



## Chemistry



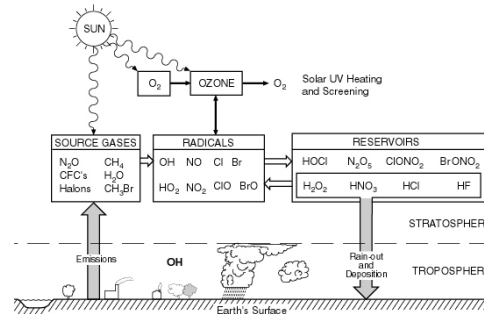
©1994 Oxford University Press

From: Atmospheric Chemistry and Global Change. Brasseur et al., 1999 & Prinn, 1994

Slide 31



## Chemistry



©1999 Oxford University Press

From: Atmospheric Chemistry and Global Change. Brasseur et al., 1999, OUP

Slide 32



## Chemistry

- **Chemical models are very expensive. It is only quite recently that it has been possible to contemplate data assimilation in reasonably comprehensive, 3d chemical models.**
- **A typical stratospheric chemistry model might contain:**
  - ~50 species
  - ~12 species/families that must be advected
  - ~40 photo-dissociation processes
  - ~100 chemical reactions

Slide 33



## Chemistry

- **Several assimilation methods have been tried:**
  - **Direct insertion of observations into a chemical model**
    - Doesn't work too well, unless equilibrium relationships are maintained. E.g. insert an observation of NO<sub>2</sub>, and the model will quickly convert NO<sub>2</sub> to NO to restore NO<sub>x</sub> equilibrium.
  - **4d-Var**
    - Can take temporal information into account
  - **Kalman Filter**
    - Too expensive for large problems
    - Can be approximated (e.g. ensemble methods)
  - **General problem: statistical analysis methods assume linearity, but chemistry is highly nonlinear.**
    - -ve concentrations, difficult minimizations, non-Gaussian statistics.

Slide 34



## Chemistry

- **Atmospheric chemistry is a very "stiff" system:**
  - Reaction timescales range from microseconds to years.
- **Two approaches are taken to address this numerically:**
  - **Family models:** species are grouped into families. The fast reactions between the species of a family are solved by assuming equilibrium. Only the slower inter-family reactions are integrated.
  - **Explicit models:** all reactions are integrated, but using solution methods designed for stiff systems to allow long timesteps.
- **Adding a new reaction or species to a family model can require a lot of changes to the code.**
  - Several equilibrium relationships may be affected. Families re-defined.
  - For 4d-Var, this effort must be duplicated for the tangent-linear and adjoint models.
- **For an explicit model, the process of adding a new reaction can be automated, even for the adjoint model.**

Slide 35



## Chemistry

- **For chemical data assimilation, it is necessary to simplify things:**
  - **Simplified chemistry:** ECMWF uses the Cariolle (1986) scheme:
$$\frac{d[O_3]}{dt} = \bar{S} + \frac{\partial S}{\partial [O_3]}([O_3] - \overline{[O_3]}) + \frac{\partial S}{\partial T}(T - \bar{T}) + \frac{\partial S}{\partial [O_3]_p}(\int [O_3]_p dp - \overline{\int [O_3]_p dp})$$
  - **No chemistry at all: Advection of long lived species.**
  - **Comprehensive chemistry, but in 2d (zonally symmetric or on an isentropic surface).**
  - **Lagrangian approach:** follow individual air parcels, applying a comprehensive "box" chemistry model to each parcel.
  - **Separate out the meteorology:** advecting winds, temperatures, etc. provided e.g. by ECMWF, and used as inputs to a comprehensive chemistry+advection model.

Slide 36



## Chemistry

- **A few references (mainly stratospheric):**

- Haggard et al. (1988, NASA tech paper 2761) used KF to analyse retrievals of water vapour and NO<sub>2</sub> from LIMS.
- Austin (1992, JGR) inserted observations of O<sub>3</sub>, HNO<sub>3</sub>, NO<sub>2</sub> and H<sub>2</sub>O into a 2D chemical model.
- Fisher and Lary (1995, QJRMMS): first application of 4d-Var to chemical data assimilation .
- Menard et al. (2000, MWR): Kalman filtering.
- Errera and Fonteyn (2001, JGR): 4dVar
- Eskes et al. (2003, QJRMMS): Sub-optimal KF (propagates variances, but correlations are fixed).
- Constantinescu et al. (2006, Atmos. Data Analysis): Ensemble Kalman filter
- Apologies to anyone not mentioned!