The Chemistry-Forecast System at the Meteorological Service of Canada

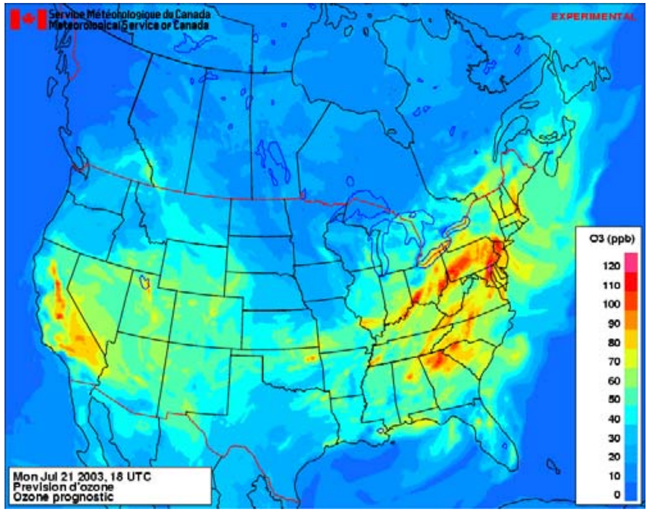
Richard Ménard
Meteorological Service of Canada

Global Earth-System Monitoring,
ECMWF Seminar, Reading, UK. September 5-9, 2005

Canadian Hemispheric and Regional Ozone and NOx System = CHRONOS

CHRONOS is a CTM used in operational air quality prediction, data assimilation and real-time scenarios

- 21 km, terrain following height coordinate (Gal-Chen) from 0-6 km
- continental domain
- ADOM-II chemical reactions
- 2 bin PM representation: PM 2.5 PM 10
- gas-phase and heterogeneous chemistry
- aerosol physics:sedimentation



CHRONOS operational version (public)

- 1 run/day (00Z) , 48 h forecast
- surface ozone objective analysis
- predicts O3, PM2.5 mass, PM 10 mass

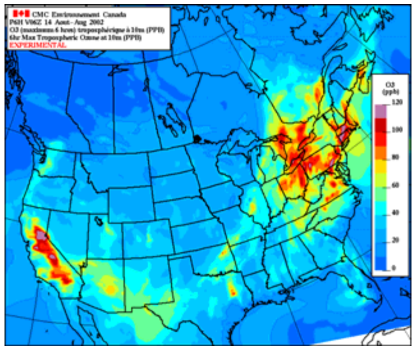
http://www.msc-smc.ec.gc.ca/aq_smog/aq_guidance_e.cfm

CHRONOS experimental version (ICARTT)

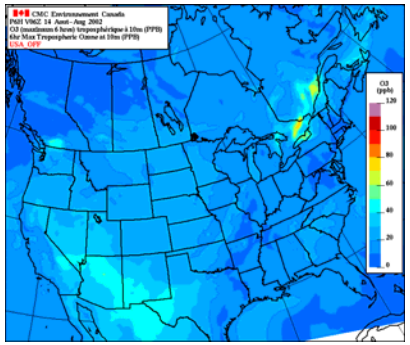
- 2 runs/day (00Z and 12Z), 48 h forecast
- assimilation of surface O3 observations

CHRONOS real-time scenarios (MSC)

- 7 runs/day (00Z), 24 h forecast
- On/Off runs for different regions



All US and Canadian emissions



Canadian emissions only

Objective analysis and assimilation of surface ozone observations

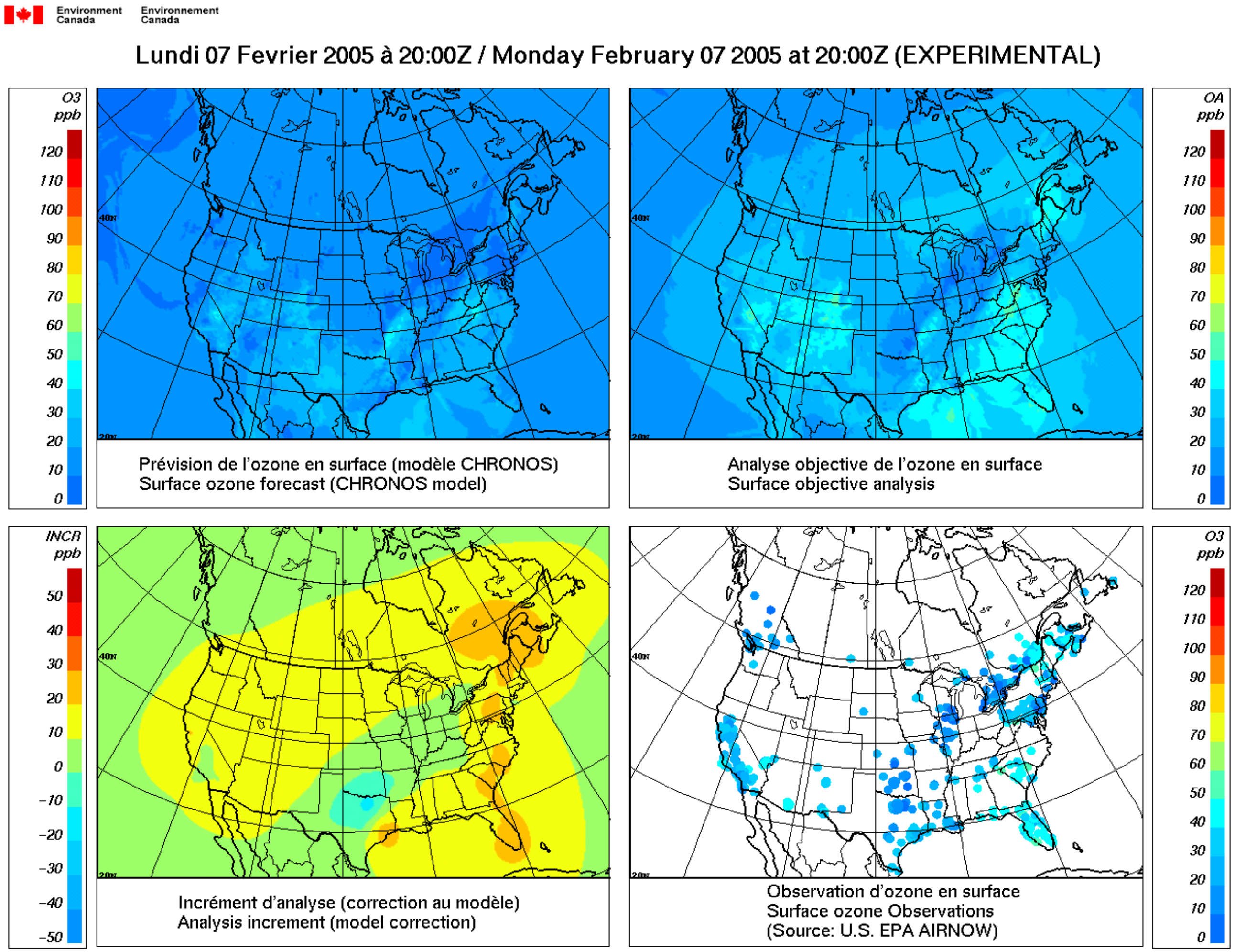
Alain Robichaud, Richard Ménard, AQMAG team

- Assimilation and objective analysis using the model CHRONOS

- Objective analysis each hour, 24/7, year round

- Operational since June 2004

- Multiyear analyses since the summer 2002

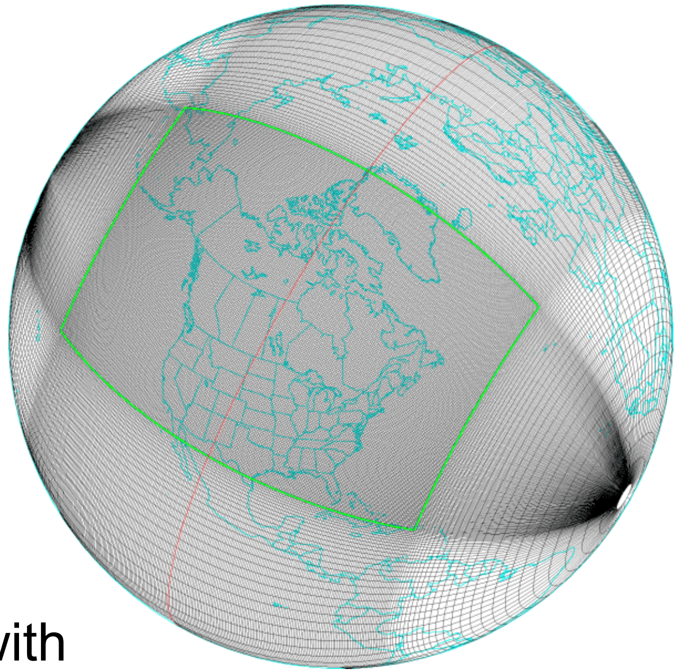


http://www.msc.ec.gc.ca/aq_smog/analysis_e.html

Global Environmental Multi-scale Air Quality model = GEM-AQ

GEM is the operational meteorology model with 4D Var capability

- semi-Lagrangian
- global/variable (100km) , limited area (15 km)
- Tropospheric 10 hPa
- Stratospheric 0.1 hPa
  non-horographic GWD
  Li and Barker k-correlated radiation



On-line chemistry

- Tropospheric chemistry (ADOM-II) with EDGAR emissions York University
- Stratospheric chemisty York University + BIRA-IASB

Online assimilation with operational model GEM

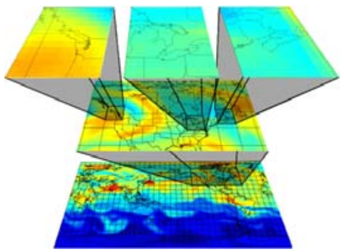
Richard Menard, Alain Robichaud,
and Pierre Gauthier

Model: GEM-AQ (J. Kaminski and L. Neary, York University, Ontario)
Tropospheric chemistry with prescribed surface emissions online with the operational meteorological model GEM-DM v3.1.2

3DVAR-CHEM: Y. Yang and Yves Rochon, MSC – Downsview
Chemical tracers analysis added (online) on the operational 3D Var assimilation system MSC-DORVAL)

MOPITT: Canadian instrument (J.Drummond, U of T) (validation V3.0) mounted on satellite EOS-TERRA

YORK
UNIVERSITÉ
UNIVERSITY



Environment Canada
Meteorological Service of Canada

Environnement Canada
Service Météorologique du Canada

Coupled chemical-dynamical data assimilation

R. Ménard, P. Gauthier, A. Robichaud, Y. Yang, A. Kallaur,
S. Ménard, M. Charron, X. Xie (Meteorological Service of Canada)

D. Fonteyn, S. Chabrillat (Belgium Institute for Space Aeronomy)

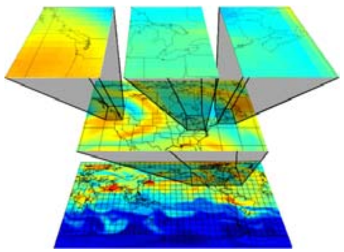
J. McConnell, J. Kaminski, L. Neary, J. Jarosz (York University)

- Two year project (ESA) to examine the benefits and drawbacks of chemistry-dynamics coupling in data assimilation

  type of coupling: Online , offline, semi-online , multivariate

- 3D and 4D Var assimilation of meteorology and chemistry observations from ENVISAT

YORK
UNIVERSITÉ
UNIVERSITY



BIRA-IASB

BASCOE

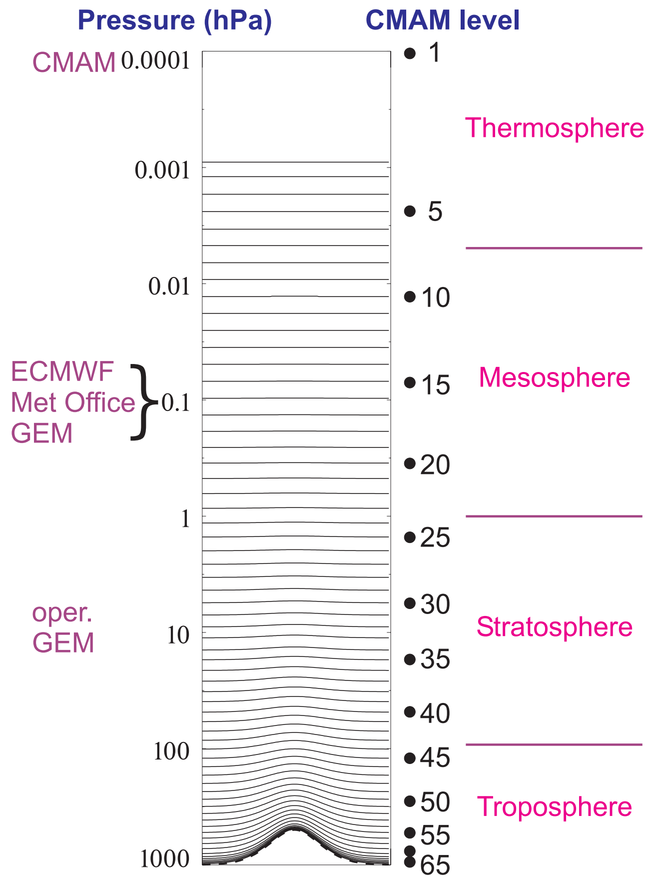
Environment Canada
Meteorological Service of Canada

Environnement Canada
Service Météorologique du Canada

Canadian Middle Atmosphere Model = CMAM

CMAM is a complex GCM with interactive chemistry, radiation and dynamics

- T47, 65 levels from 0-95 km
- 127 gas-phase chemical reactions
- heterogeneous chemistry
- Hines GWD scheme



CMAM Data Assimilation

(Polavarapu, Ren, Rochon, Sankey, Yang)

- CMC's 3DVAR on CMAM's coordinates
- obs: conventional, AMSU-A 4-14
- start-up from climate state Dec. 15, 2001

Courtesy of S. Polavarapu

Outline

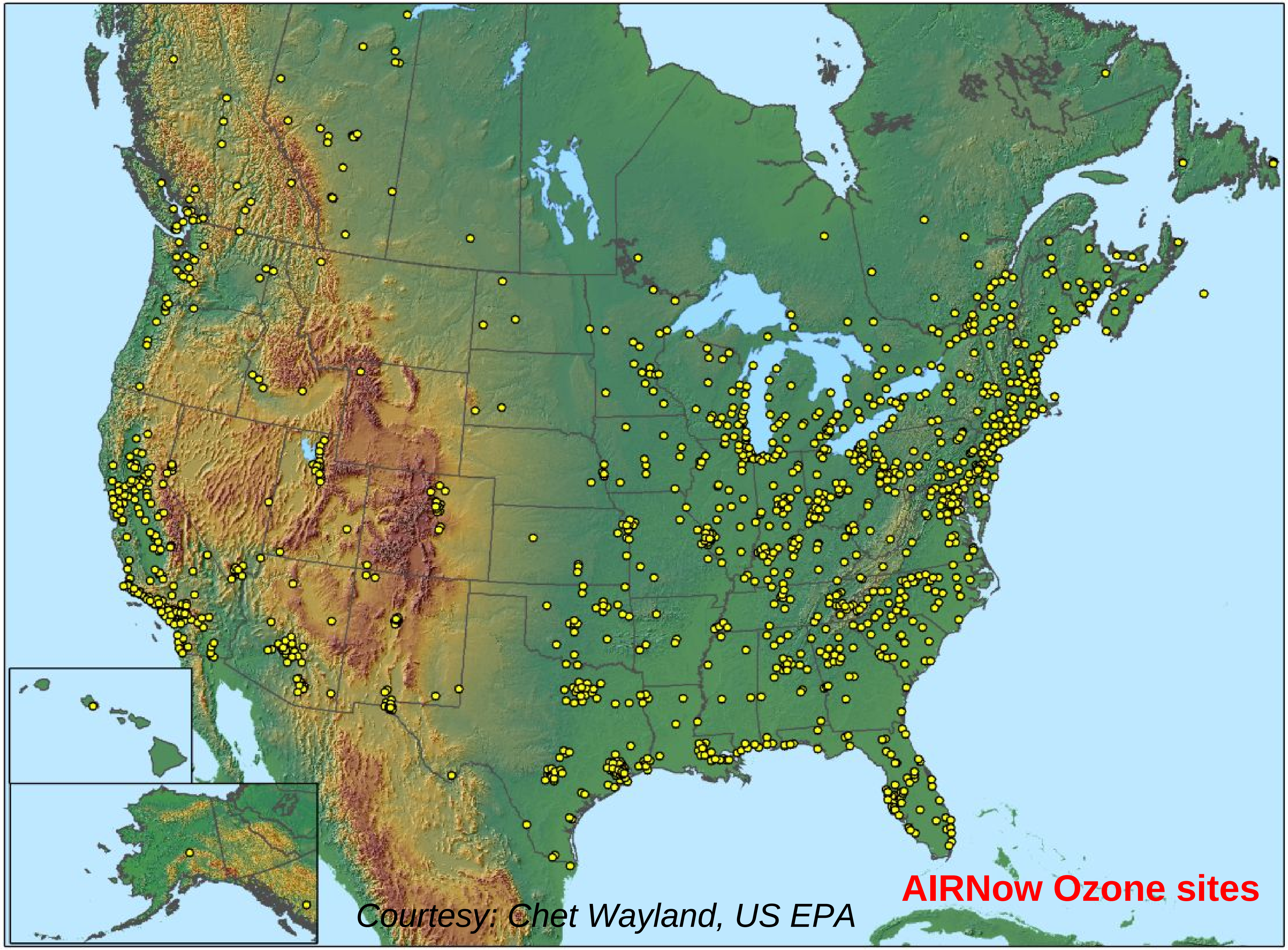
1. Surface observations
   - AIRNOW real-time observations
   - Canadian and US surface observations: current and planned
2. Kalman filtering theory
   - algorithm
   - asymptotic stability
3. Optimum interpolation using surface ozone observations
   - model improvement
   - implementation
   - covariance modelling
   - statistical consistency
   - verification
   - background error vertical correlation
4. Other species
   - why univariate ?
   - impact of O3 analyses on other species
5. Prediction
6. Environmental impact
7. How important is chemistry in assimilation and monitoring ?
8. Ongoing and future direction
9. Some outstanding problems in chemical data assimilation
   - unobserved species: analysis, error statistics

1. Surface observations

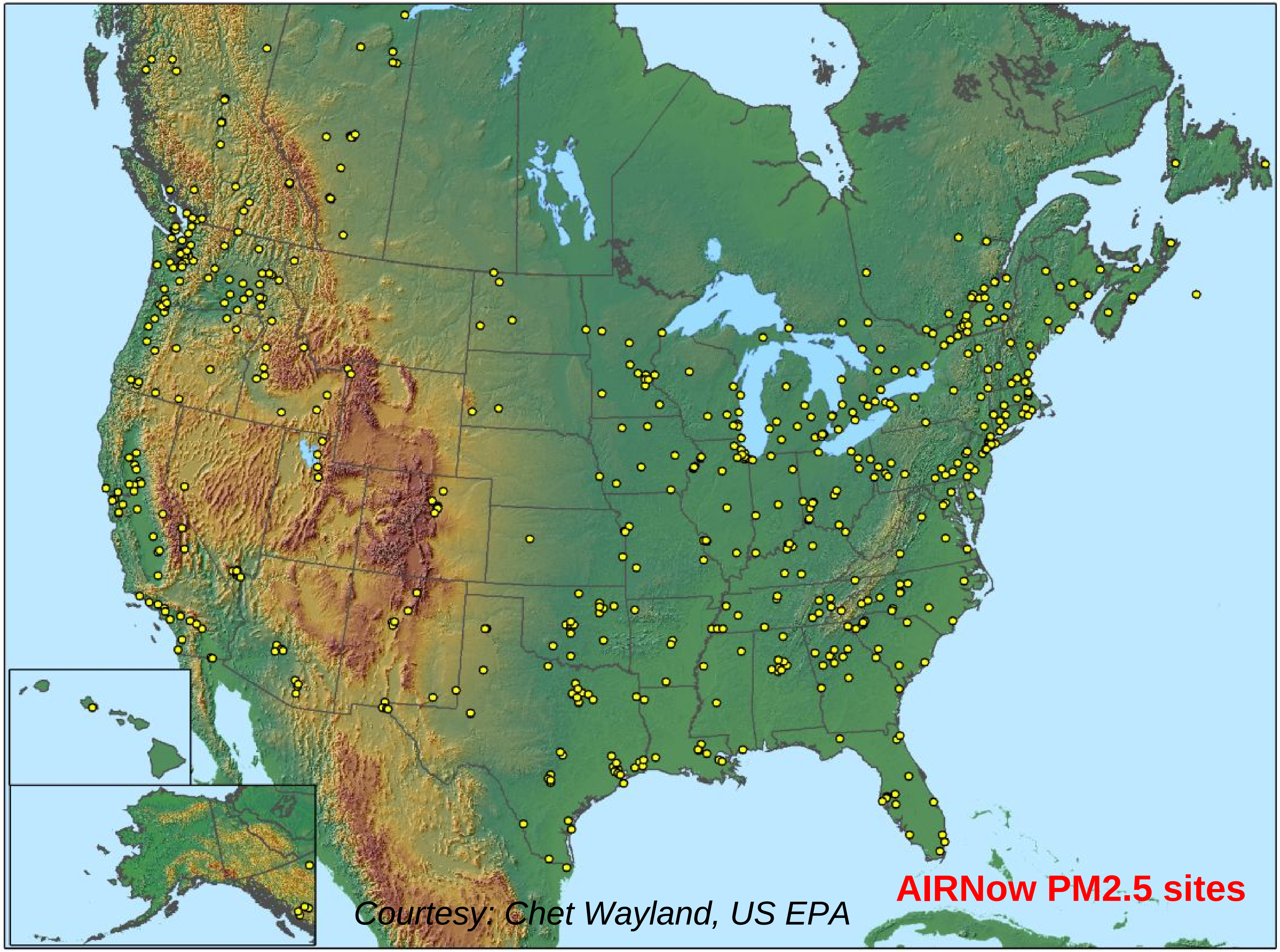
Real-time observations : US-Canada

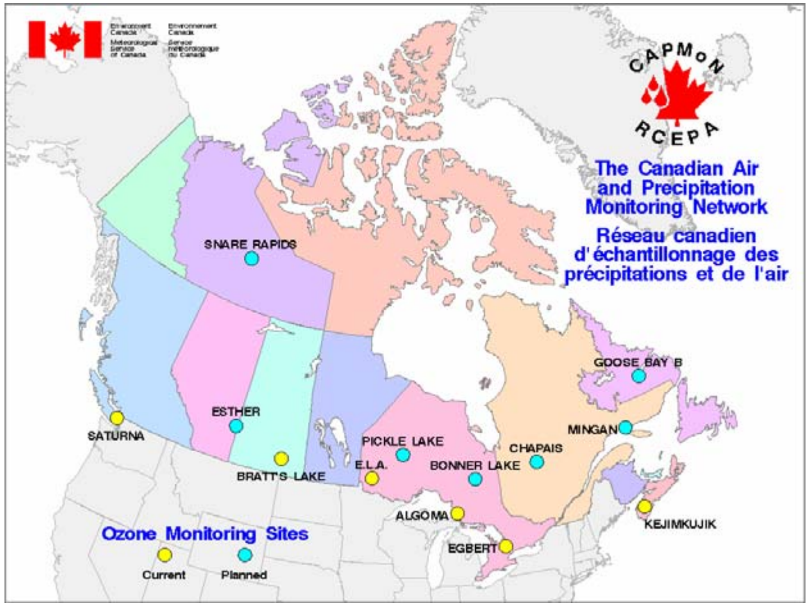
- Available from AIRNOW (US EPA) Data Management Center
  currently ftp download ASCII files , future BUFR distributed by NOAA
  currently Canadian observations collected at CMC in BUFR
- Data automatically QA/QC's each hour at US EPA
- Available 20-30 min. past the hour

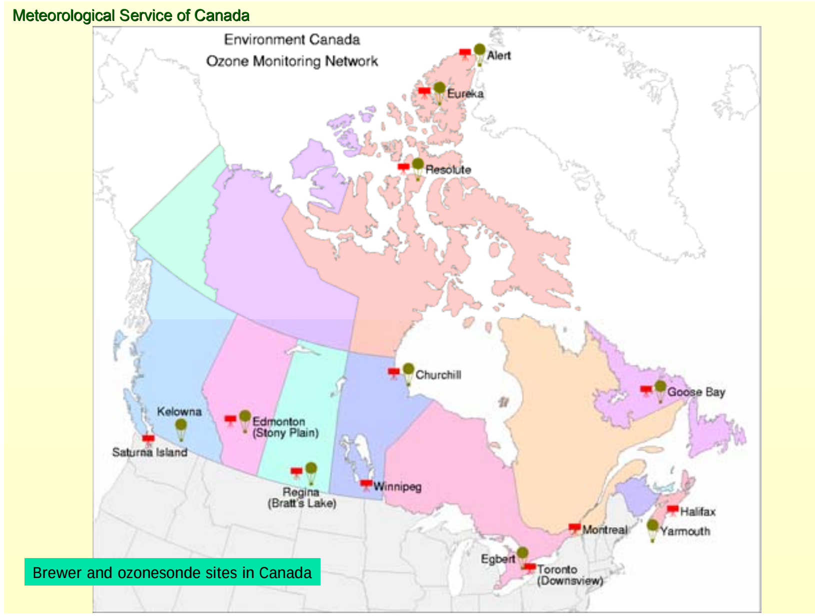
Parameter	Units	# Sites	Frequency	Coverage
Ozone	ppm	~1300	hourly	Good
PM2.5	ug/m3	~ 450	hourly	Mod-Good
PM10	ug/m3	~ 40	hourly	Limited

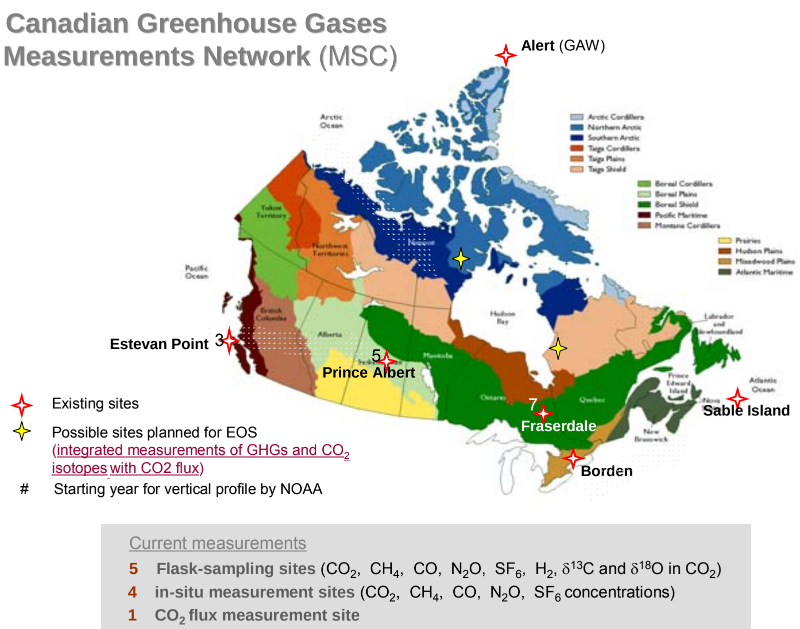


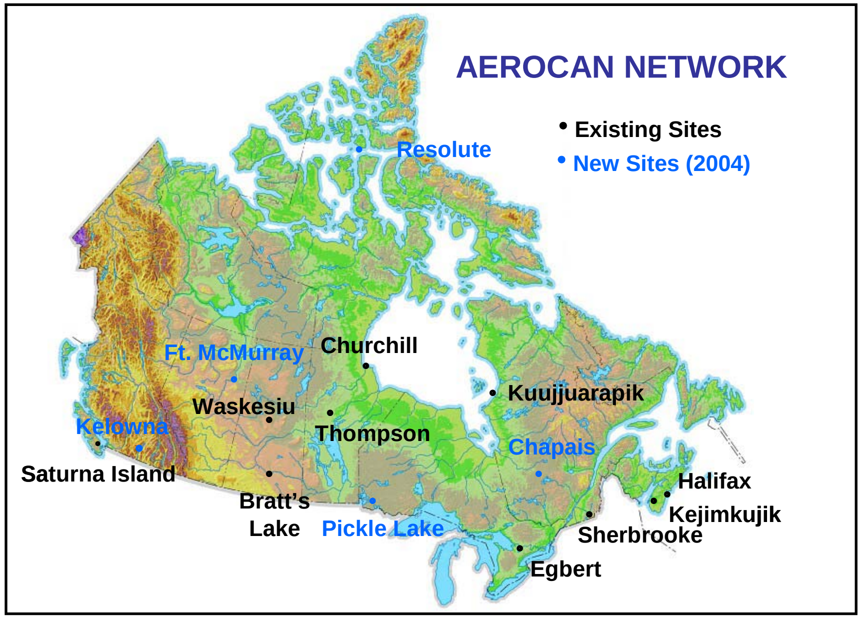
AIRNow Ozone sites
Courtesy: Chet Wayland, US EPA


AIRNow PM2.5 sites
Courtesy: Chet Wayland, US EPA


CAPMoN
RCEPA
The Canadian Air and Precipitation Monitoring Network
Réseau canadien d'échantillonnage des précipitations et de l'air
SNARE RAPIDS
GOOSE BAY B
ESTHER
SATURNA
MINGAN
PICKLE LAKE
CHAPAIS
E.L.A.
BONNER LAKE
BRATT'S LAKE
ALGOMA
KEJIMKUJIK
EGBERT
Ozone Monitoring Sites
Current
Planned


Meteorological Service of Canada
Environment Canada
Ozone Monitoring Network
Alert
Eureka
Resolute
Churchill
Goose Bay
Kelowna
Edmonton (Stony Plain)
Saturna Island
Regina (Bratt's Lake)
Winnipeg
Halifax
Yarmouth
Montreal
Egbert
Toronto (Downsview)
Brewer and ozonesonde sites in Canada

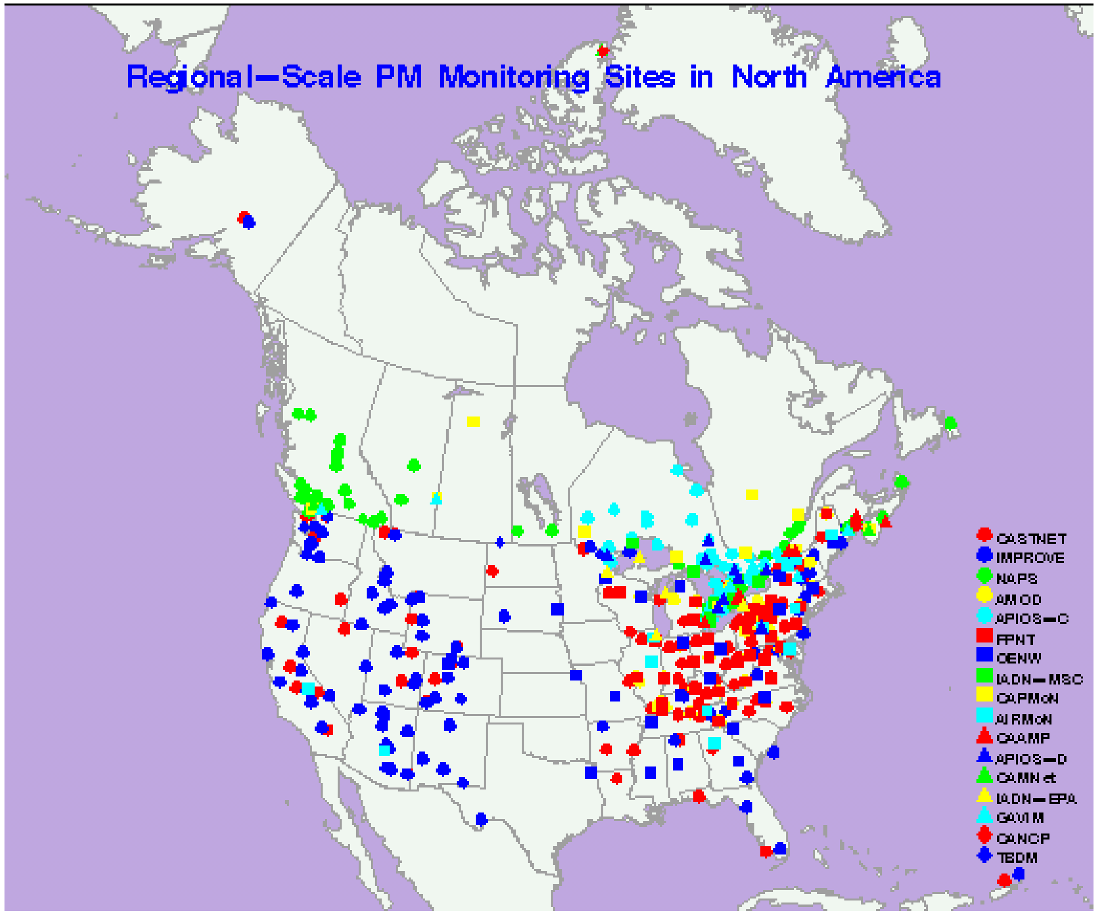

Canadian Greenhouse Gases Measurements Network (MSC)
Alert (GAW)
Estevan Point 3
5 Prince Albert
7 Fraserdale
Sable Island
Borden
Existing sites
Possible sites planned for EOS (integrated measurements of GHGs and CO_2 isotopes with CO2 flux)
Starting year for vertical profile by NOAA
Current measurements
5 Flask-sampling sites (CO_2, CH_4, CO, N_2O, SF_6, H_2, $\delta^{13}C$ and $\delta^{18}O$ in CO_2)
4 in-situ measurement sites (CO_2, CH_4, CO, N_2O, SF_6 concentrations)
1 CO_2 flux measurement site


AEROCAN NETWORK
• Existing Sites
• New Sites (2004)
Resolute
Ft. McMurray
Churchill
Kuujjuarapik
Waskesiu
Thompson
Kelowna
Chapais
Saturna Island
Bratt's Lake
Pickle Lake
Halifax
Kejimkujik
Sherbrooke
Egbert

Repository of monitoring networks (US & Canada)

NATChem www.msc-smc.ec.gc.ca/natchem/index_e.html

AQ, climate and toxics data sets



Future (proposed) of AIRNow (ref Chet Wayland, US EPA)

- US EPA DMC can accept additional parameters
  - precursor gases
  - speciation of aerosols
  - meteorological parameters

- Continuous measurements of precursor gases at 24 sites
  - NO, NO_2 (true measurement), NO_x, NO_y, SO_2, CO

- Continuous measurements of aerosols speciation at 10 sites
  - EC (elementary carbon PM2.5)
  - OC (organic carbon)
  - BC (black carbon)
  - UBV (second channel of Aethalometer)
  - NO_3 ions (nitrate)
  - SO_4 ions (sulfates)

2. Kalman filtering theory

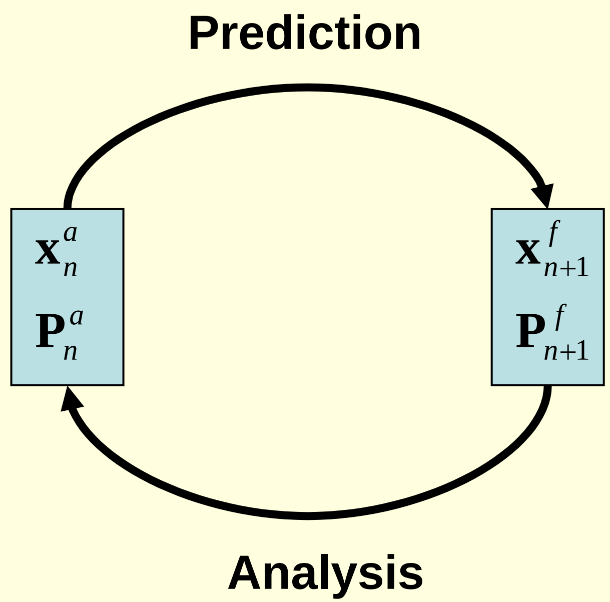
In a Kalman filter, the error covariance is dynamically evolved between the analyses, so that both the state vector and the error covariance are updated by the dynamics and by the observations.

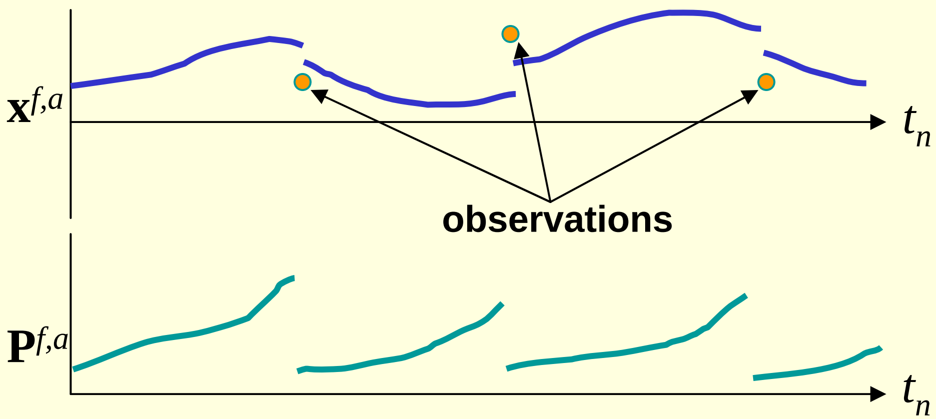
- *Prediction*

$$\mathbf{x}^f_{n+1} = \mathbf{M}_n \mathbf{x}^a_n$$
$$\mathbf{P}^f_{n+1} = \mathbf{M}_n \mathbf{P}^a_n \mathbf{M}^T_n + \mathbf{Q}_n$$

- *Analysis*

$$\mathbf{x}^a_n = \mathbf{x}^f_n + \mathbf{K}_n(\mathbf{y}_n - \mathbf{H}_n \mathbf{x}^f_n)$$
$$\mathbf{K}_n = \mathbf{P}^f_n \mathbf{H}^T_n \left(\mathbf{H}_n \mathbf{P}^f_n \mathbf{H}^T_n + \mathbf{R}_n\right)^{-1}$$
$$\mathbf{P}^a_n = \left(\mathbf{I} - \mathbf{K}_n \mathbf{H}_n\right)\mathbf{P}^f_n$$





Definition The Kalman filter produces the best estimate of the atmospheric state given *all current and past* observations, and yet the algorithm is *sequential in time* in a form of a predictor-corrector scheme.

From a Bayesian point of view the Kalman filter constructs an estimate based on

$$p(\mathbf{x}_n \mid \mathbf{y}_n, \mathbf{y}_{n-1}, \ldots, \mathbf{y}_0)$$

Time sequential property of a Kalman filter is not easy to show

Example: Estimating a scalar constant using no prior, and assuming white noise observation error with constant error variance. The MV estimate can be shown to be the simple *time averaging*

$$\hat{x}_k = \frac{1}{k}\sum_{i=1}^{k} y_i$$

and which can be re-written in a sequential scheme

$$\hat{x}_{k+1} = \frac{1}{k+1}\sum_{i=1}^{k+1} y_i = \frac{k}{k+1}\left(\frac{1}{k}\sum_{i=1}^{k} y_i\right) + \frac{1}{k} y_{k+1} = \frac{k}{k+1}\hat{x}_k + \frac{1}{k} y_{k+1}$$

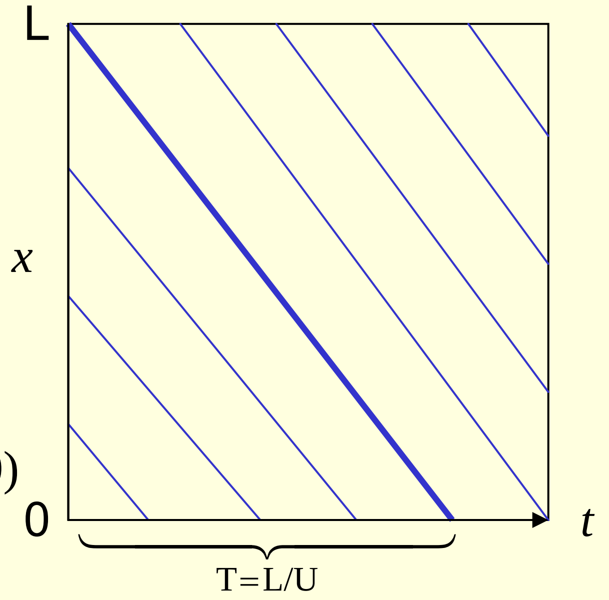
a) Asymptotic stability and observability

- We say that we have *observability* when it is possible for a data assimilation system with a perfect model and perfect observations to determine a unique initial state from a *finite* time sequence of observations.

Example: one-dimensional advection over a periodic domain

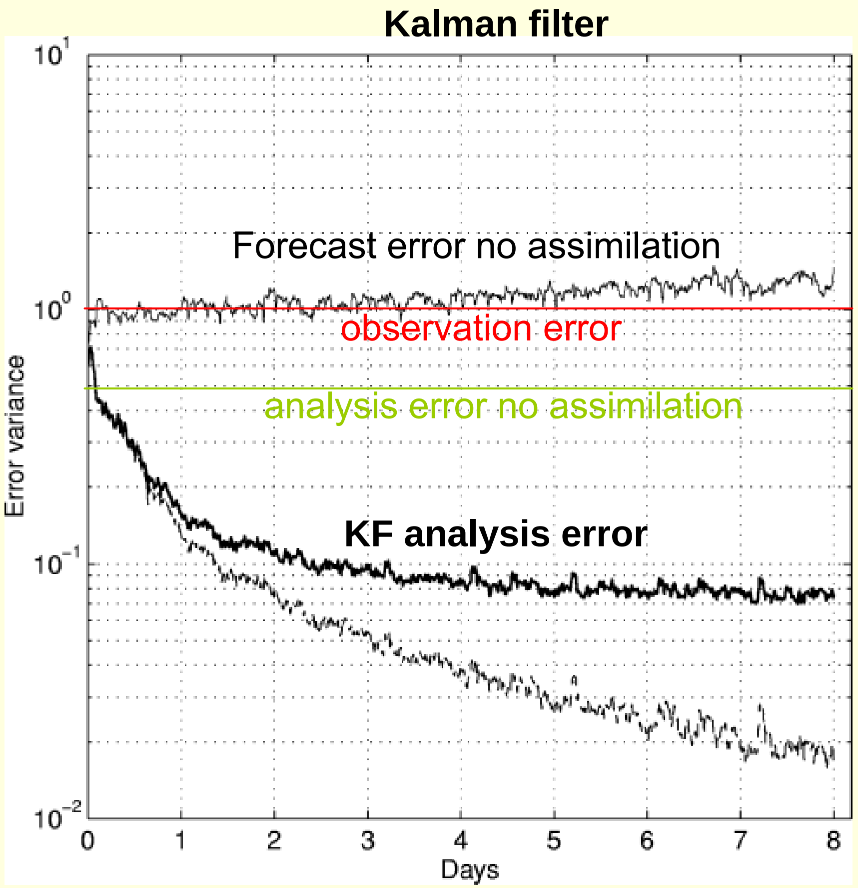
$$\frac{\partial \phi}{\partial t} = U \frac{\partial \phi}{\partial x}$$

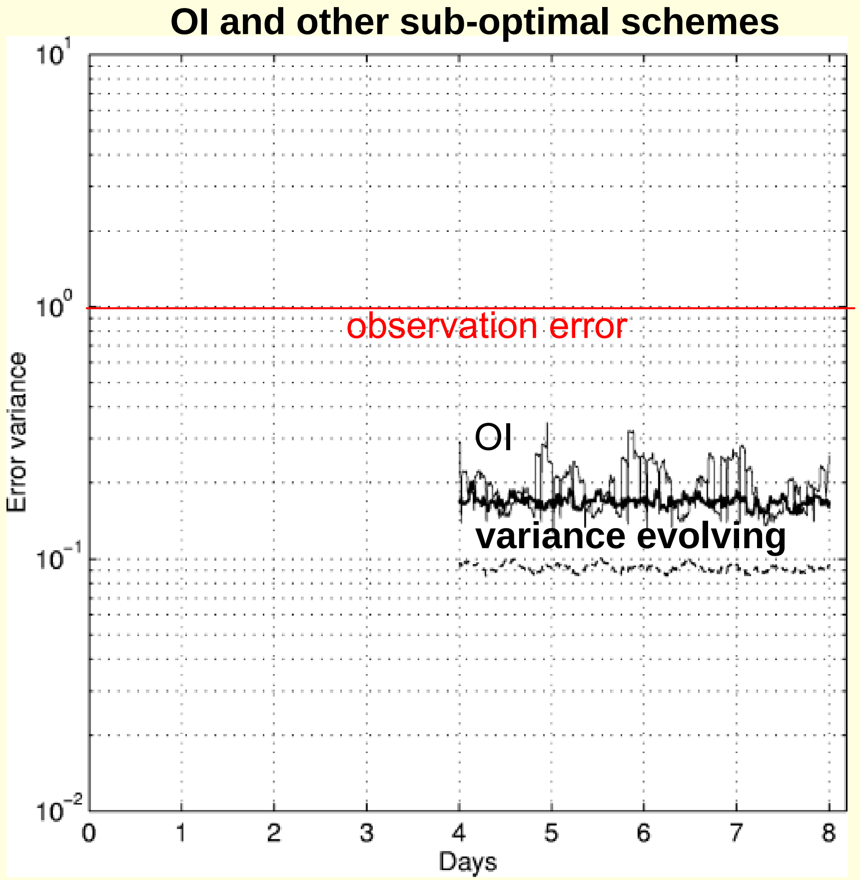
continuous observations at a single point x=0 over a time T=L/U determines uniquely the initial condition $\phi_0(x) = \phi(x,0)$



Theorem. If an assimilation system is observable and the model is imperfect, $\mathbf{Q} \neq 0$, then the sequence of forecast error covariance $\left\{\mathbf{P}_k^f\right\}$ converges geometrically to $\mathbf{P}_\infty^f$ which is independent of $\mathbf{P}_0^f$

Results from the assimilation of CH_4 observations from UARS
Menard and Chang (2000)





The accumulation of information over time is limited by model error, so that in practice we need to integrate a Kalman filter only over some finite time to obtain the optimal solution

3. Optimum interpolation using surface ozone observations

Motivations

- Tropospheric ozone analysis using an alternative approach to assimilating limb/nadir combination of satellite measurements

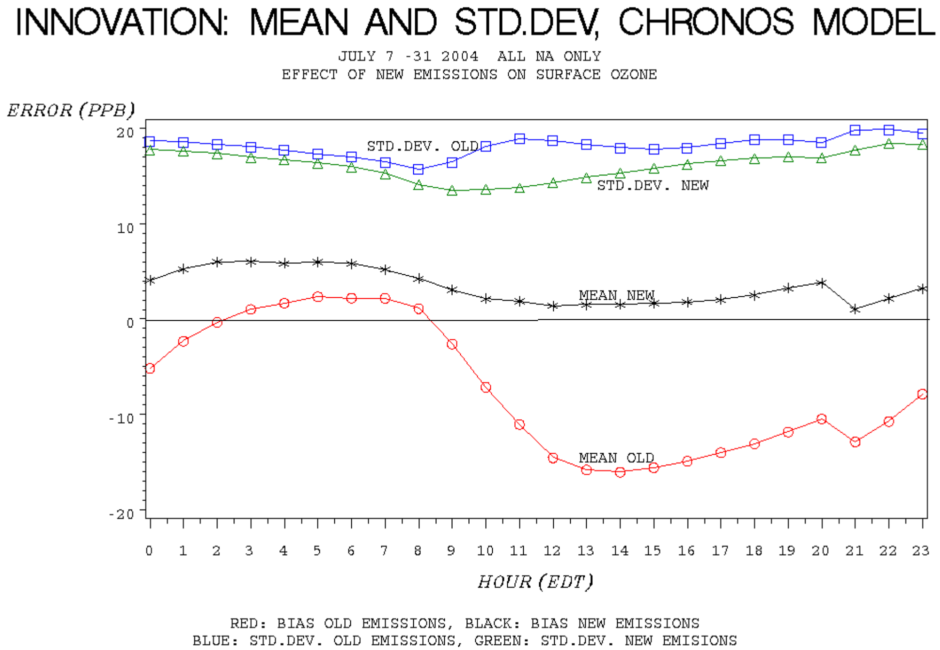
- Surface ozone measurements are accurate (<1 ppb), calibrated each night, very small bias, reports each hour in real time, fixed location, extensive spatial coverage. They are ideal to construct error statistics.

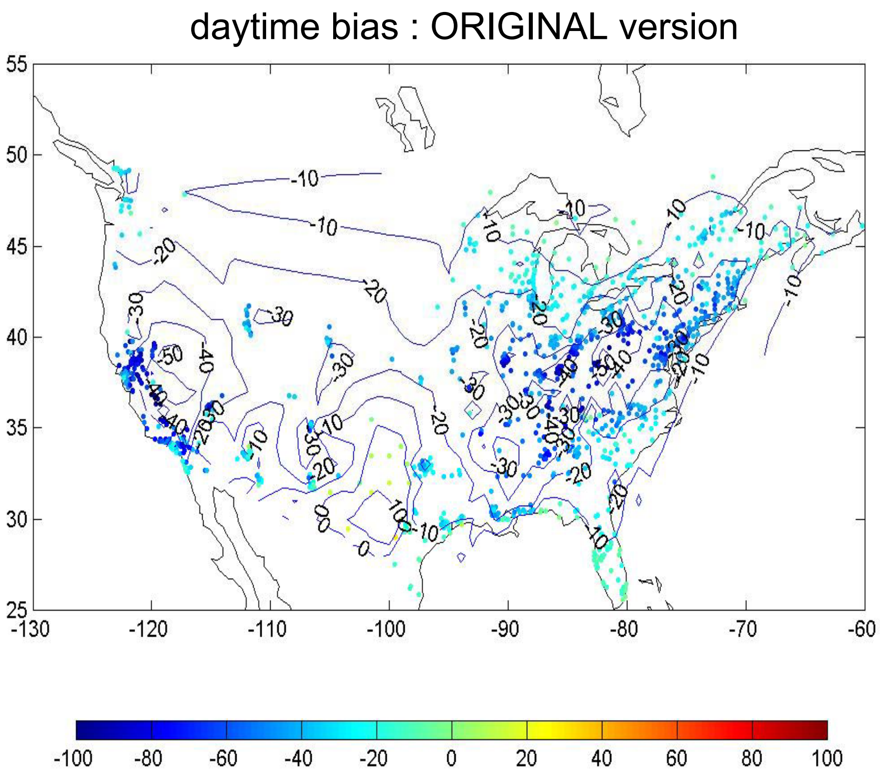
- Identify the main model problems, and hopefully correct them.

- To have quickly an operational product, which then helps to take the next step in the development of a comprehensive chemical weather system.

- Air quality analyses and derived quantities (although it may not impact significantly the air quality prediction).

a) model improvement





ORIGINAL version (pre assimilation era) of CHRONOS

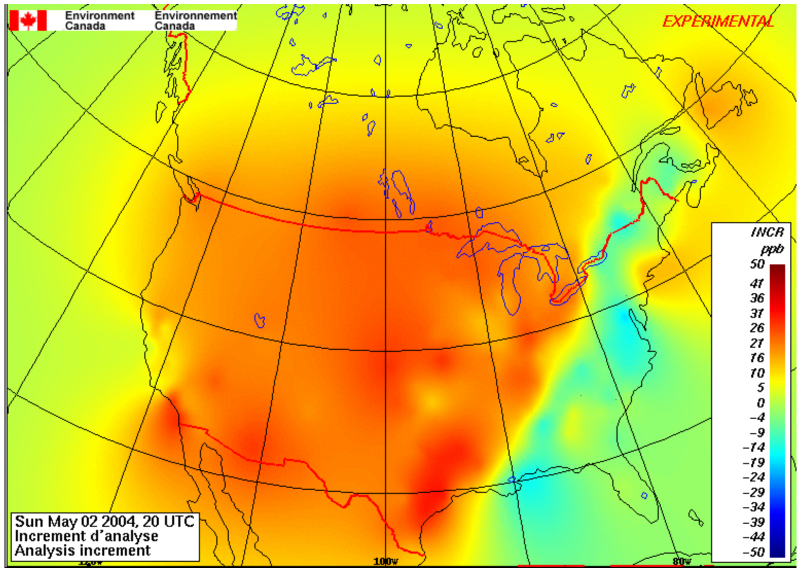
• bias (red) and error standard deviation (blue) of ~ 20 ppb
  and sometime of comparable size ⟶ no assimilation

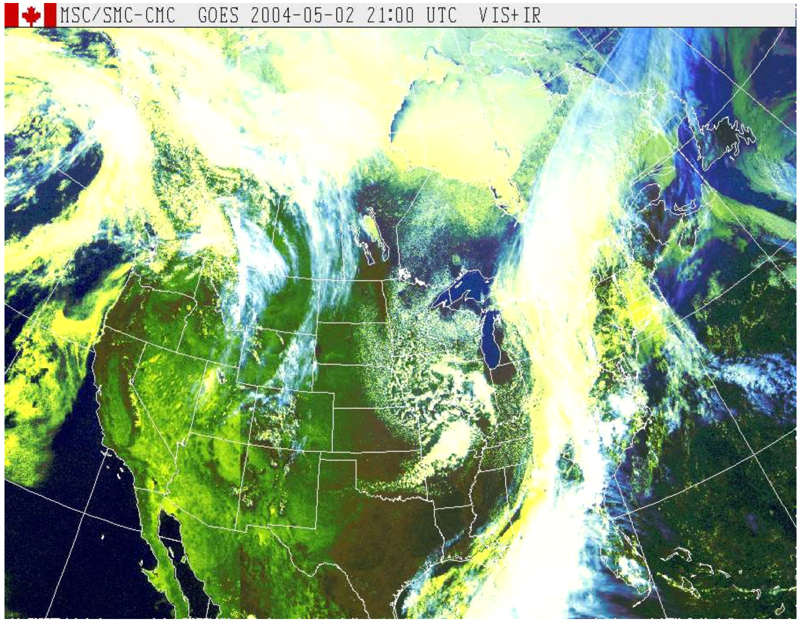
Last year (with improvement on clouds, dry deposition, and emissions)

• bias (~ 5 ppb, black curve) is smaller than the error standard deviation (green ~ 15 ppb)

In assimilation

Incremental analysis vs cloud
Case study. May 02 2004 20Z





- Photochemistry in CHRONOS uses NMP model clouds.
- Increase model top in order to better represent clouds in the CTM

b) Implementation of optimal interpolation

What is optimum interpolation ?

Approximation to the steady-state Kalman filter

- *Prediction*

$$\boxed{\mathbf{x}_{n+1}^f = \mathbf{M}_n \mathbf{x}_n^a}$$

- *Analysis*

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

$$\mathbf{K}_n = \left\{\vec{K}_n(i)\right\}$$

$$\vec{K}_n(i) = \vec{B}(r_i, \mathbf{r}_{obs}) \left[\mathbf{B}(\mathbf{r}_{obs}, \mathbf{r}_{obs}) + \mathbf{R}\right]^{-1}$$

where $i = 1, \ldots, N$ (number of model grid points)

$\vec{K}$ is a row vector of dimension p (number of observations)

r_i is the position of grid point i

$\mathbf{r}_{obs}$ is a vector of position of observations

Remarks:

1- Data selection is used to permit the inversion of the innovation covariance matrix

2- The innovations are generally used to obtain (fit) a functional form of the background error covariance

3- The background error variances are extrapolated to the whole model domain

Covariance modelling

Positive definite matrix (Horn and Johnson 1985, Chap 7)

A real *n* x *n* symmetric matrix **A** is positive definite if

$$\mathbf{c}^T \mathbf{A}\,\mathbf{c} > 0$$

for any nonzero vector **c**. **A** is said to be positive semi-definite if

$$\mathbf{c}^T \mathbf{A}\,\mathbf{c} \geq 0$$

Properties

- The sum of any positive definite matrices of the same size is also positive definite
- Each eigenvalue of a positive definite matrix is a positive real number
- The trace and determinant are positive real numbers.

Covariance matrix

The covariance matrix **P** of a random vector $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ is the matrix $\mathbf{P} = [P_{ij}]$ in which $P_{ij} = \mathbf{E}\left[(X_i - \overline{X}_i)(X_j - \overline{X}_j)\right]$ where $\overline{X}_i = \mathbf{E}\left[X_i\right]$ and **E** is the mathematical expectation.

Property: A covariance matrix is positive semi-definite

$$\mathbf{E}\left[\left(c_1(X_1-\overline{X}_1)+\cdots+c_n(X_n-\overline{X}_n)\right)^2\right]=\mathbf{E}\left[\sum_{i,j=1}^{n}c_i(X_i-\overline{X}_i)c_j(X_j-\overline{X}_j)\right]$$

$$=\sum_{i,j=1}^{n}c_i\mathbf{E}\left[(X_i-\overline{X}_i)(X_j-\overline{X}_j)\right]c_j \;=\mathbf{c}^T\mathbf{P}\,\mathbf{c}\geq 0$$

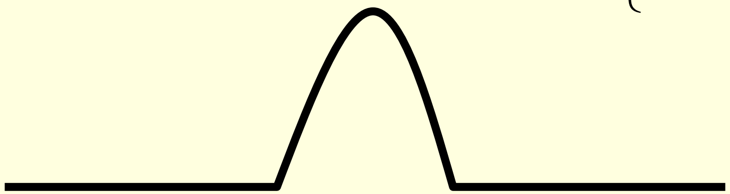
Remarks

1 - It is often necessary in data assimilation to invert the covariance matrices, and thus we need to have *positive definite* covariances

2 – The positive definite property is global property of a matrix, and it is not trivial to obtain

Examples:

a) Truncated parabola $$C(i,j) = \begin{cases} 1 - \dfrac{(i-j)^2}{d^2} & for\ |i-j| \le n \\ 0 & otherwise \end{cases}$$



for $n=4$

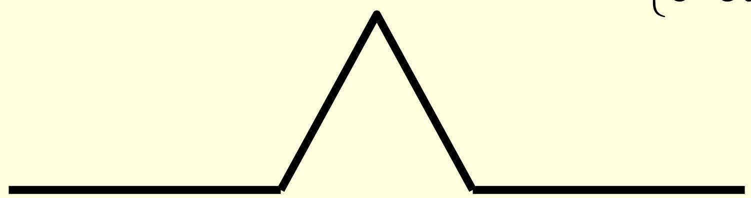
$$\begin{bmatrix} 1.000 & 0.937 & 0.750 & 0.437 & 0.000 \\ 0.937 & 1.000 & 0.937 & 0.750 & 0.437 \\ 0.750 & 0.937 & 1.000 & 0.937 & 0.750 \\ 0.437 & 0.750 & 0.937 & 1.000 & 0.937 \\ 0.000 & 0.437 & 0.750 & 0.937 & 1.000 \end{bmatrix}$$

eigenvalues

$$\begin{array}{r} 3.8216 \\ 1.2500 \\ 0.0000 \\ 0.0000 \\ -0.0716 \end{array}$$

b) Triangle

$$C(i,j) = \begin{cases} 1 - \dfrac{|i-j|}{d} & for\ |i-j| \le n \\ 0 & otherwise \end{cases}$$



for n=4

$$\begin{bmatrix} 1.000 & 0.750 & 0.500 & 0.250 & 0.000 \\ 0.750 & 1.000 & 0.750 & 0.500 & 0.250 \\ 0.500 & 0.750 & 1.000 & 0.750 & 0.500 \\ 0.250 & 0.500 & 0.750 & 1.000 & 0.750 \\ 0.000 & 0.250 & 0.500 & 0.750 & 1.000 \end{bmatrix}$$

eigenvalues

3.0646
1.3090
0.2989
0.1910
0.1365

Covariance functions (Gaspari and Cohn 1998)

Definition 1: A function $P(\mathbf{r},\mathbf{r}')$ is a covariance function of a random field X if

$$P(\mathbf{r},\mathbf{r}') = \left\langle [X(\mathbf{r}) - \left\langle X(\mathbf{r}) \right\rangle][X(\mathbf{r}') - \left\langle X(\mathbf{r}') \right\rangle] \right\rangle$$

Definition 2: A covariance function $P(\mathbf{r},\mathbf{r}')$ is a function that defines positive semi-definite matrices when evaluated on any grid.
That is, letting $\mathbf{r}_i$ and $\mathbf{r}_j$ be any two grid points, the matrix $\mathbf{P}$ whose elements are $\mathbf{P}_{i,j} = P(\mathbf{r}_i,\mathbf{r}_j)$ is defines a covariance matrix, when P is a covariance function

The equivalence between definition 1 and 2 can be found in Wahba(1990, p1-2).
The covariance function is also known as the *reproducing kernel*.

Remark Suppose a covariance function is defined in a 3D space, $\mathbf{r} \in \mathrm{R}^3$.
Restricting the value of $\mathbf{r}$ to remain on an manifold (e.g. the surface of a unit sphere) will also define a covariance function, and a covariance matrix (e.g. a covariance matrix on the surface of a sphere)

<u>Correlation function</u> A correlation function $C(\mathbf{r},\mathbf{r}')$ is a covariance function $P(\mathbf{r},\mathbf{r}')$ normalized by the standard deviation at the points $\mathbf{r}$ and $\mathbf{r}'$

$$C(\mathbf{r},\mathbf{r}') = \frac{P(\mathbf{r},\mathbf{r}')}{\sqrt{P(\mathbf{r},\mathbf{r})}\,\sqrt{P(\mathbf{r}',\mathbf{r}')}}$$

<u>Homogeneous and isotropic correlation function</u> If a correlation function is invariant under all translation and all orthogonal transformation, then the correlation function become only a function of the distance between the two points,

$$C(\mathbf{r},\mathbf{r}') = C_0(\|\mathbf{r}-\mathbf{r}'\|)$$

Smoothness properties

- The continuity at the origin determines the continuity allowed on the rest of the domain. For example, if the first derivative is discontinuous at the origin, then first derivative discontinuity is allowed elsewhere (see example with triangle)

Spectral representation

- On a unit circle

$$C(\mathbf{r},\mathbf{r}') = R(\cos\theta) = \sum_{m=0}^{\infty} a_m \cos(m\theta)$$

  where θ is the angle between the two position vectors, and where and all the Fourier coefficients a_m are nonnegative

- On a unit sphere

$$C(\mathbf{r},\mathbf{r}') = R(\cos\theta) = \sum_{m=0}^{\infty} b_m P_m(\cos\theta)$$

  where all the Legendre coefficients b_m are nonnegative.

Examples of correlation functions

1. Spatially uncorrelated model (black)

$$C_0(\|\mathbf{r}-\mathbf{r}'\|) = \begin{cases} 1 \text{ if } \mathbf{r}=\mathbf{r}' \\ 0 \text{ if } \mathbf{r}\neq\mathbf{r}' \end{cases}$$

2. First-order auto-regressive model (FOAR) (blue)

$$C_0(\|\mathbf{r}-\mathbf{r}'\|) = \exp\left(-\frac{\|\mathbf{r}-\mathbf{r}'\|}{L}\right)$$

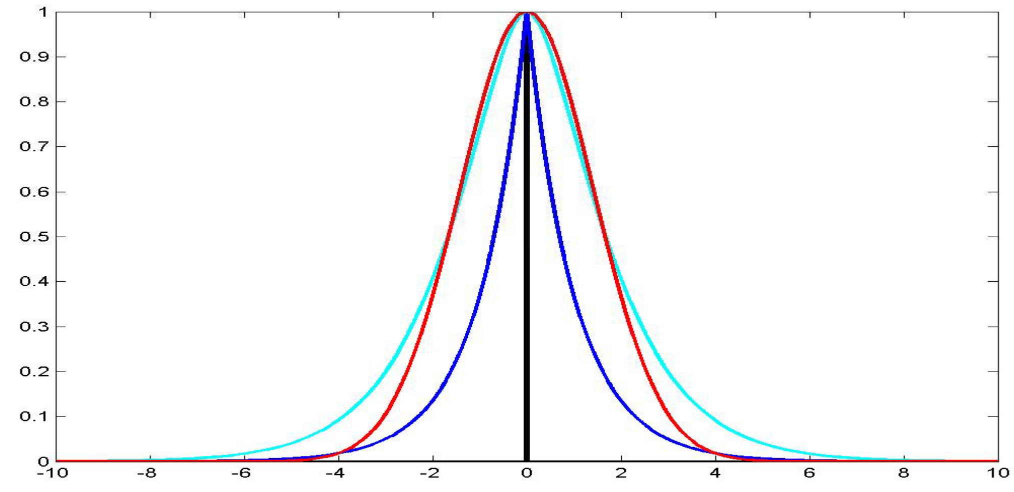
3. Second-order auto-regressive model (SOAR) (cyan)

$$C_0(\|\mathbf{r}-\mathbf{r}'\|) = \left(1+\frac{\|\mathbf{r}-\mathbf{r}'\|}{L}\right)\exp\left(-\frac{\|\mathbf{r}-\mathbf{r}'\|}{L}\right)$$

4. Gaussian model (red)

$$C_0(\|\mathbf{r}-\mathbf{r}'\|) = \exp\left(-\frac{\|\mathbf{r}-\mathbf{r}'\|^2}{2L^2}\right)$$

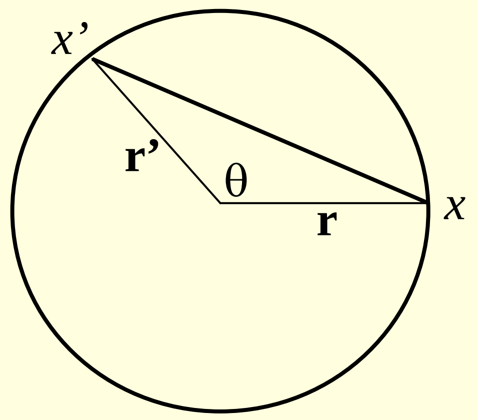
where L is the correlation length scale



Exercise: Construct a correlation matrix on a one-dimensional periodic domain

Consider a 2D Gaussian model on a plane

$$C(\mathbf{r},\mathbf{r}') = \exp\left(-\frac{\|\mathbf{r}-\mathbf{r}'\|^2}{2L^2}\right) \quad \text{where } \mathbf{r} \in R^2$$



Along the circle of radius a

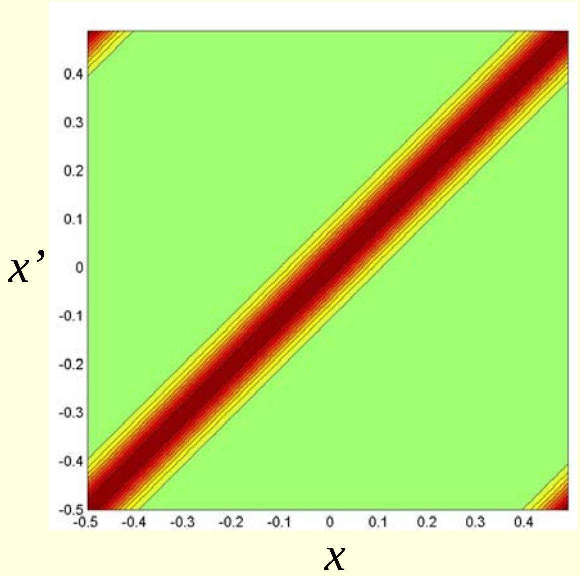
$$\|\mathbf{r}-\mathbf{r}'\|^2 = 2a^2(1-\cos\theta)$$

Now define a coordinate x along the circle ,

$$x' - x = \frac{a\theta}{2\pi} \quad \text{for } -a \le x, x' \le a$$

then we get

$$C(x,x') = \exp\left(-\frac{(1-\cos[2\pi(x-x')/a])}{(L/a)^2}\right)$$



Ways to define new covariances by transformation

Define a new spatial coordinate

If $x \mapsto f(x)$ is one-to-one, then

$$Q(x, x') = P(f(x), f(x'))$$

is a new covariance function

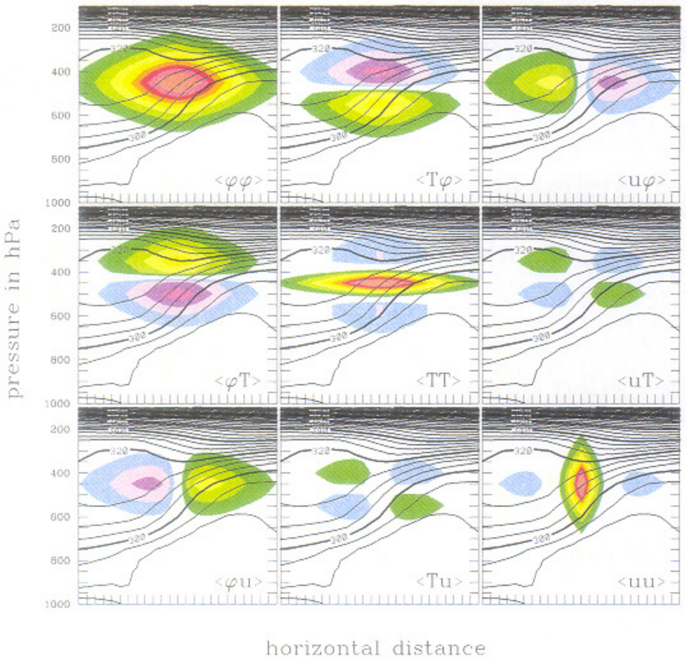


Figure 4.5
x/P cross-section of 9 correlations – pressure as vertical coordinate

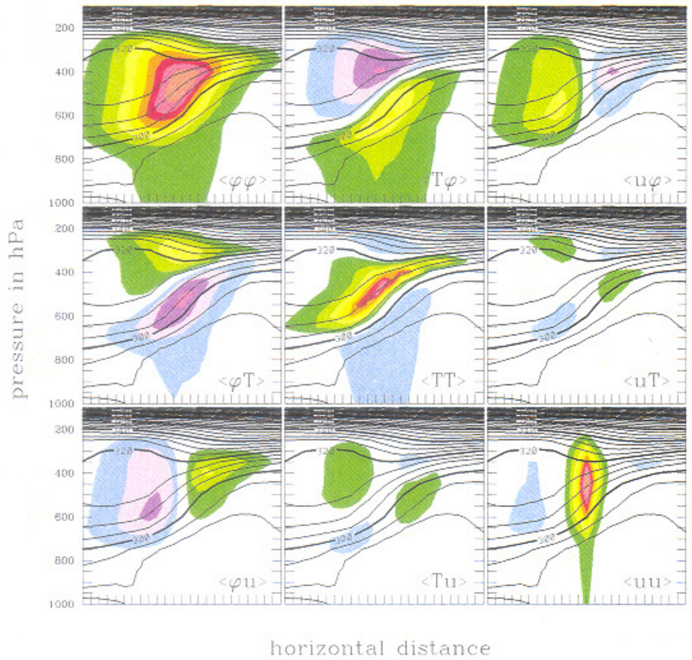


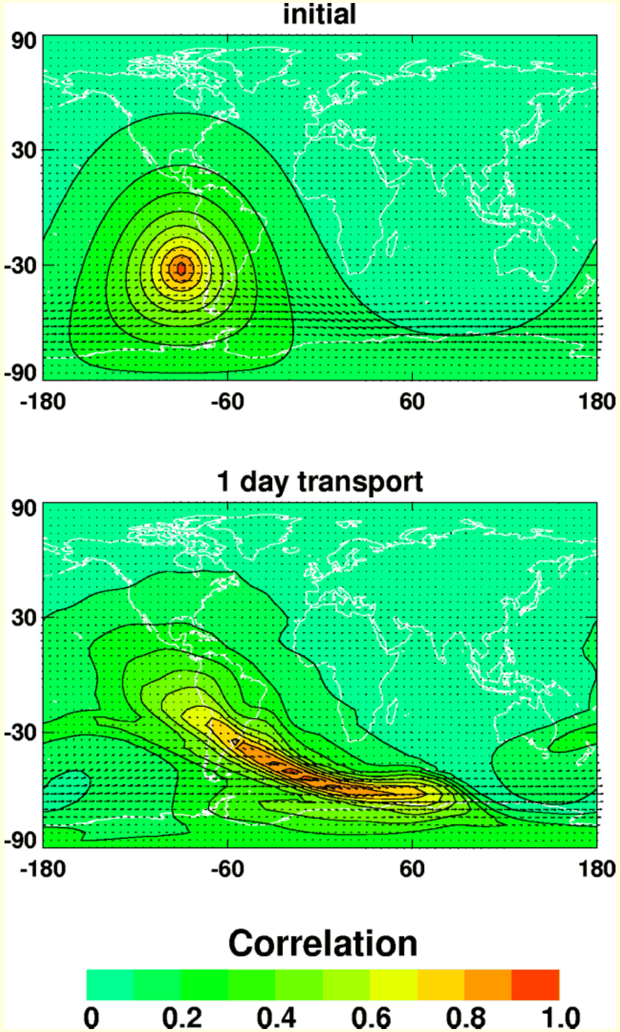
Figure 4.6
Same as Figure 4.5, except for θ as vertical coordinate

Linear transformation

If $L(x)$ is a linear operator,

$$\mu^*(x) = L(x)\ \mu(x),$$

then $\mathbf{P}_{\mu^*} = \mathbf{L}\ \mathbf{P}_\mu \mathbf{L}^T$



Hadamard product

If **A** and **B** are two covariance matrices, then the componentwise multiplication $\mathbf{A} \circ \mathbf{B}$ is also a covariance matrix

e.g. Separable correlation functions

$$C(x,y,z,x',y',z') = C^h(x,y,x',y')\, C^v(z,z')$$

Self-convolution

The self-convolution of discontinuous functions that goes to zero at a finite distance, produces compactly-supported covariance functions

- Examples in Gaspari and Cohn (1999)
- Hadamard product of an arbitrary covariance function with a compactly supported function is very useful in Ensemble Kalman filtering to create compactly supported covariances

State dependence

If f is a function of the state μ,

then a state-dependent error

$$\varepsilon^*_\mu(x) = f(\langle\mu\rangle)\ \varepsilon_\mu(x)$$

can be used for data assimilation, if

- $\langle\mu\rangle = \mu^a$ in forecast errors
- $\langle\mu\rangle = \mu^f$ in analysis errors

• When formulated this way state-dependent errors can be introduced in estimation theory for data assimilation

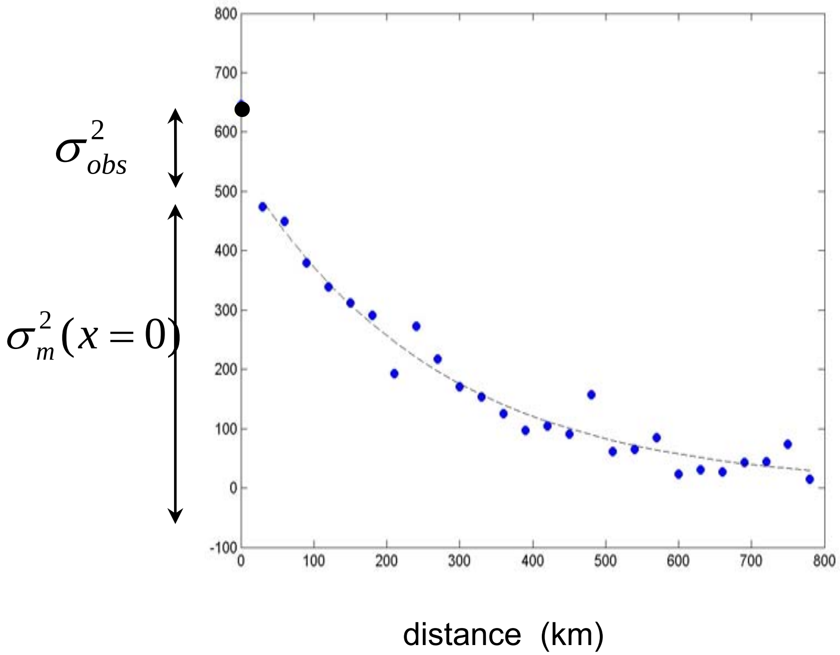
e.g. To create a relative error formulation for observation error, the relative error is scaled by the forecast interpolated at the observation location (not the observed value!)

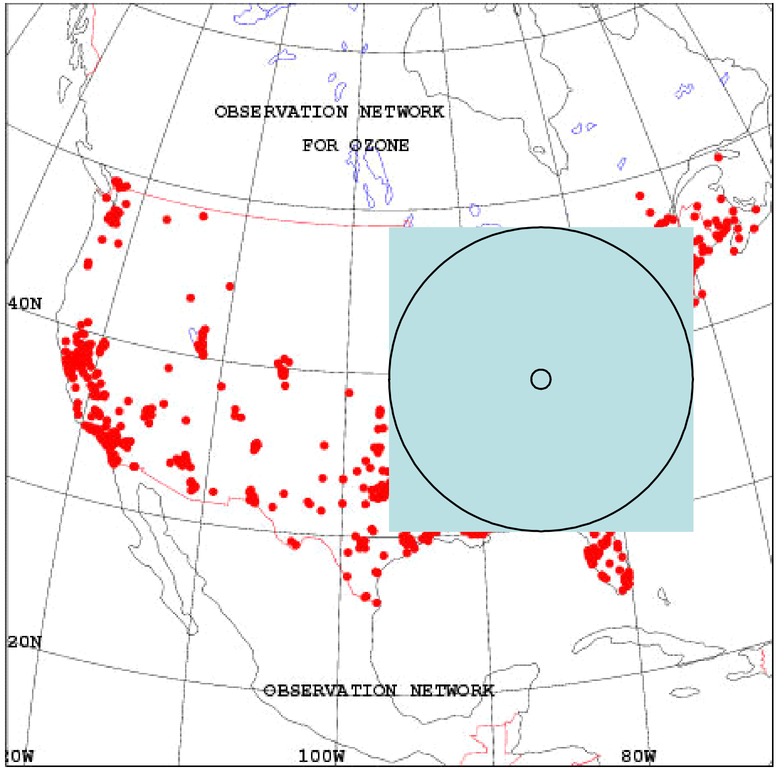
To obtain the error statistics
we use innovations

obs – model (obs loc) =

(true + obs error) -
(true + model error) =

obs error – model error



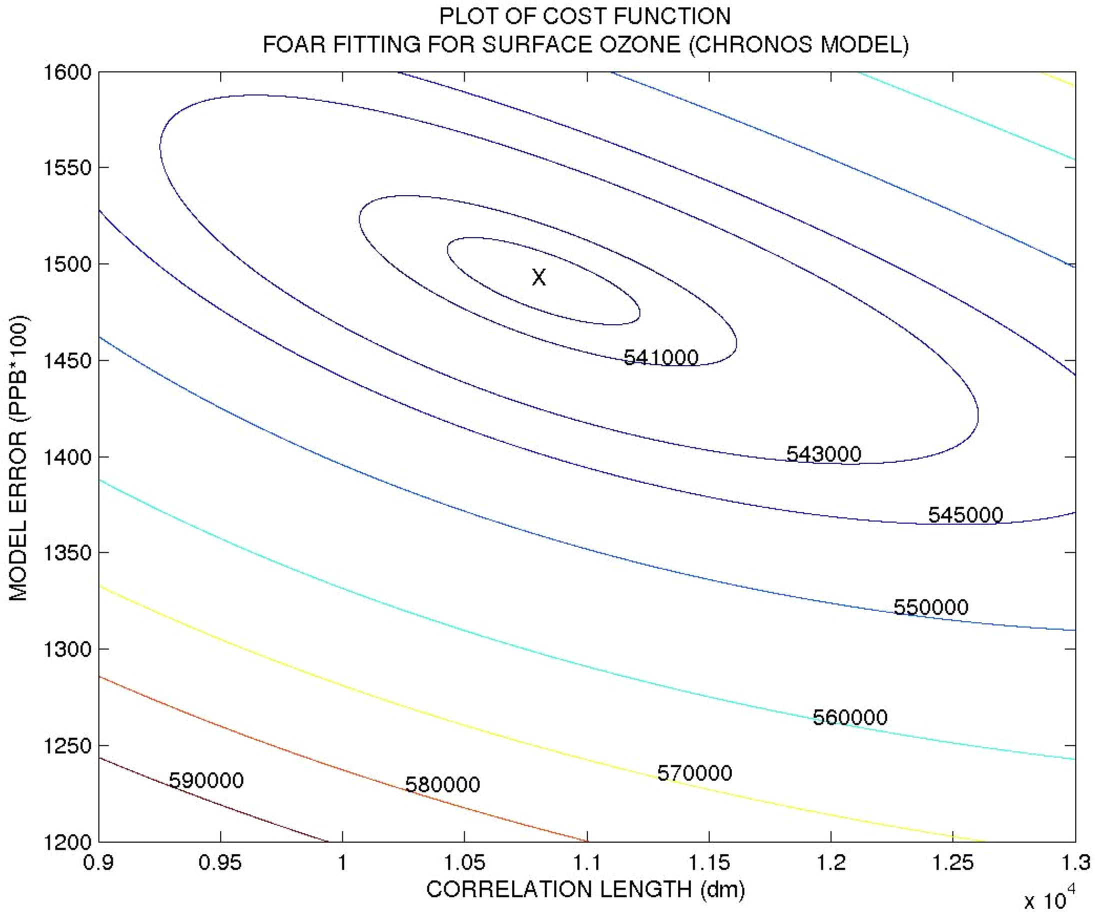


Obs. & background error variances and correlation length scale at

- each sites
- each hour

$$J\left(L(i),\, \sigma_{\mathrm{B}}^{2}(i)\right) = \sum_{j \neq i} \left(\left\langle OmF_i \,,\, OmF_j \right\rangle - \sigma_B(i)\,\sigma_B(j)\; C(i,j) \right)^2$$

$$\approx \sum_{j \neq i} \left(\left\langle OmF_i \,,\, OmF_j \right\rangle - \sigma_B^2(i)\, C(i,j) \right)^2$$



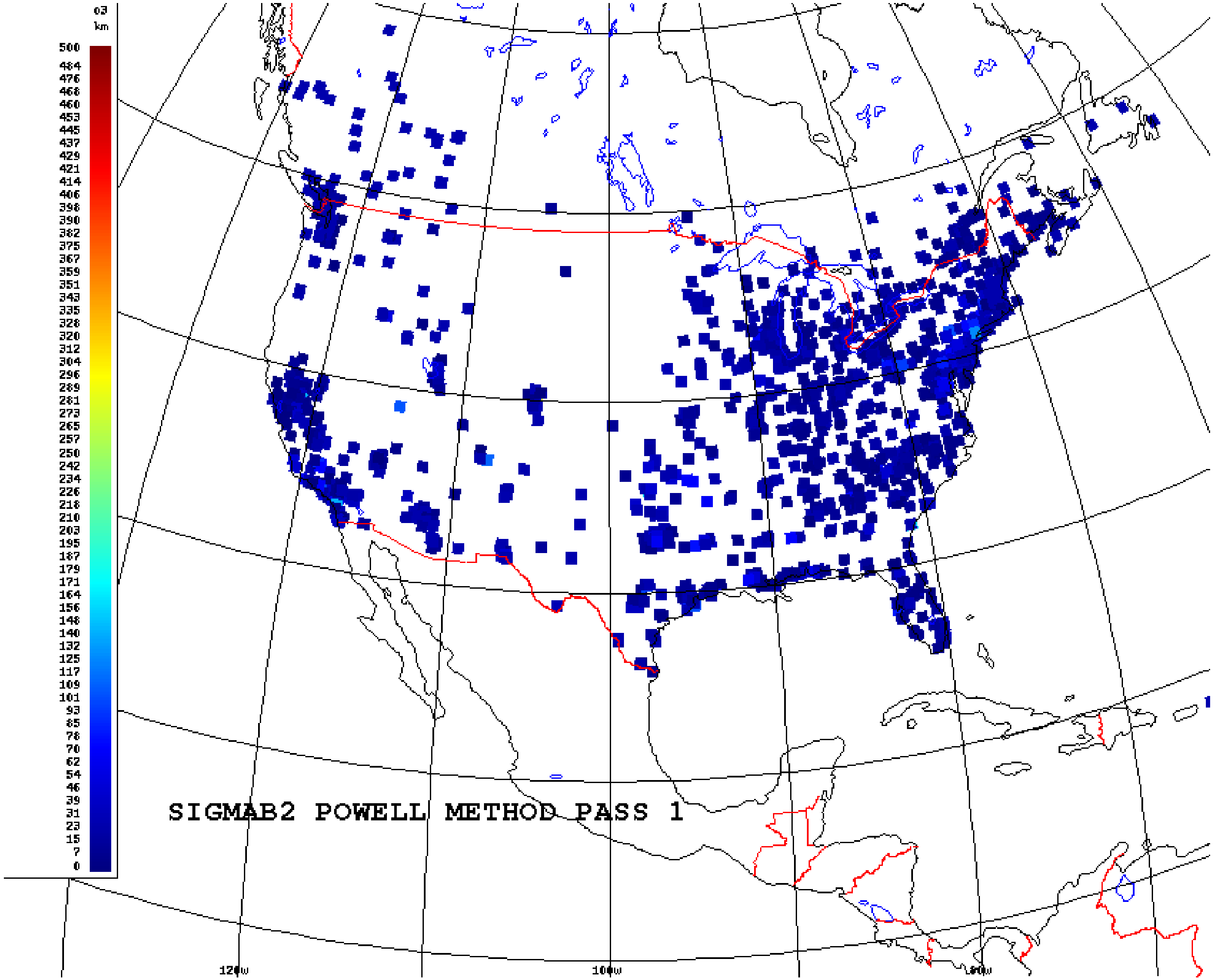
- Best fit for correlation model: FOAR (First order autoregressive)

$$C(i, j) = \exp\left(-\frac{|r_i - r_j|}{L} \right)$$

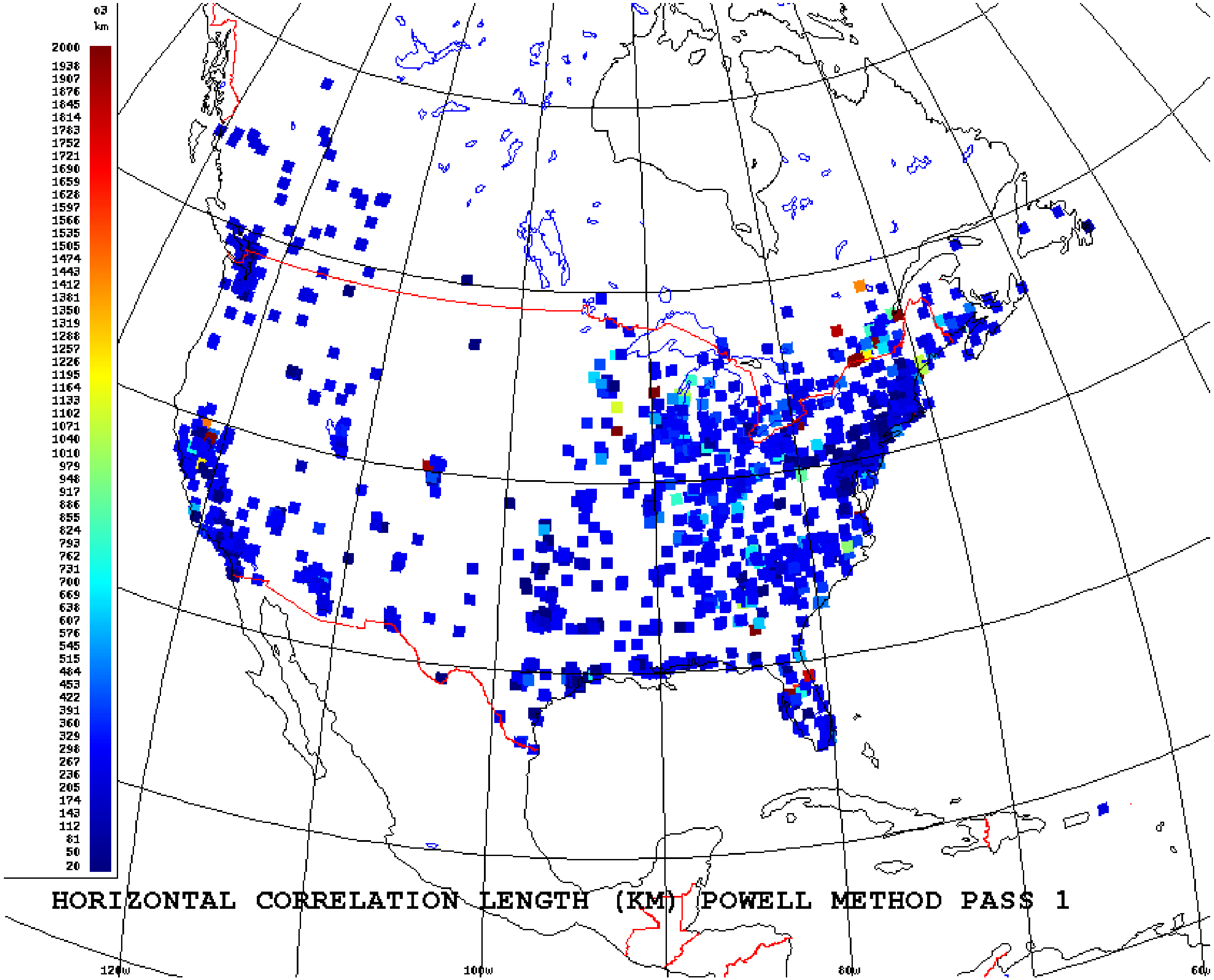
- Error statistics in terms of land use

	FOREST	COMMERCIAL	RESIDENTIAL	AGRICULTURAL	INDUSTRIAL
Total variance	270	355	363	332	400
Forecast error	213	279	286	276	297
Correlation length scale (km)	412	333	334	314	308
Observation error	57	76	77	56	103
Number of sites	20	53	85	47	12

background error variance - σ_B^2



horizontal correlation length scale - L



To compute the Kalman gain we have assumed

- $B(r_i, r_{j-obs}) = \sigma_B(r_i)\sigma_B(r_{j-obs})\ C(r_i, r_{j-obs}) \approx \overline{\sigma}_B\ \sigma_B(r_{j-obs})\ C(r_i, r_{j-obs})$

that the correlation is homogeneous

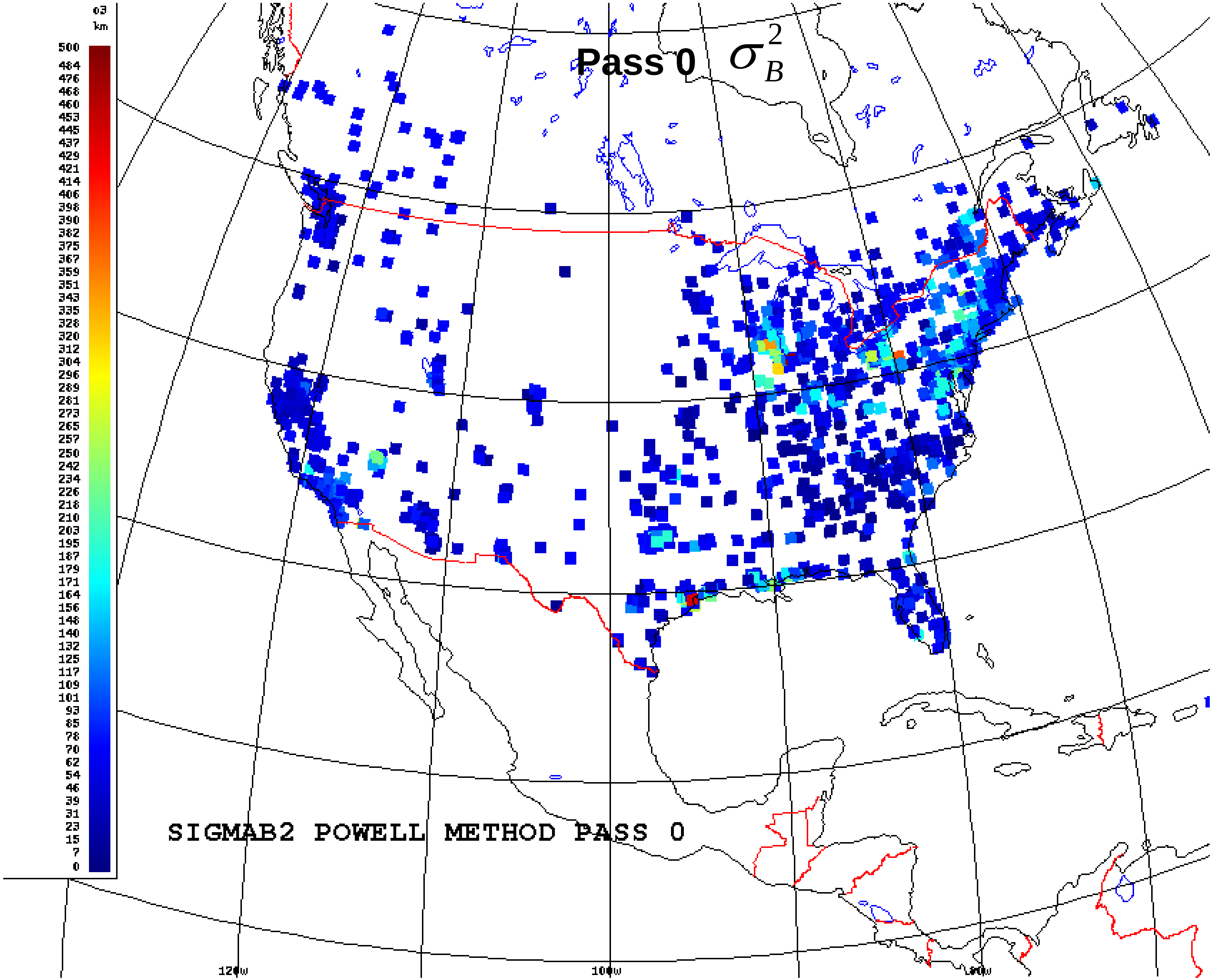
- $L = \overline{L}(r_{j-obs})$

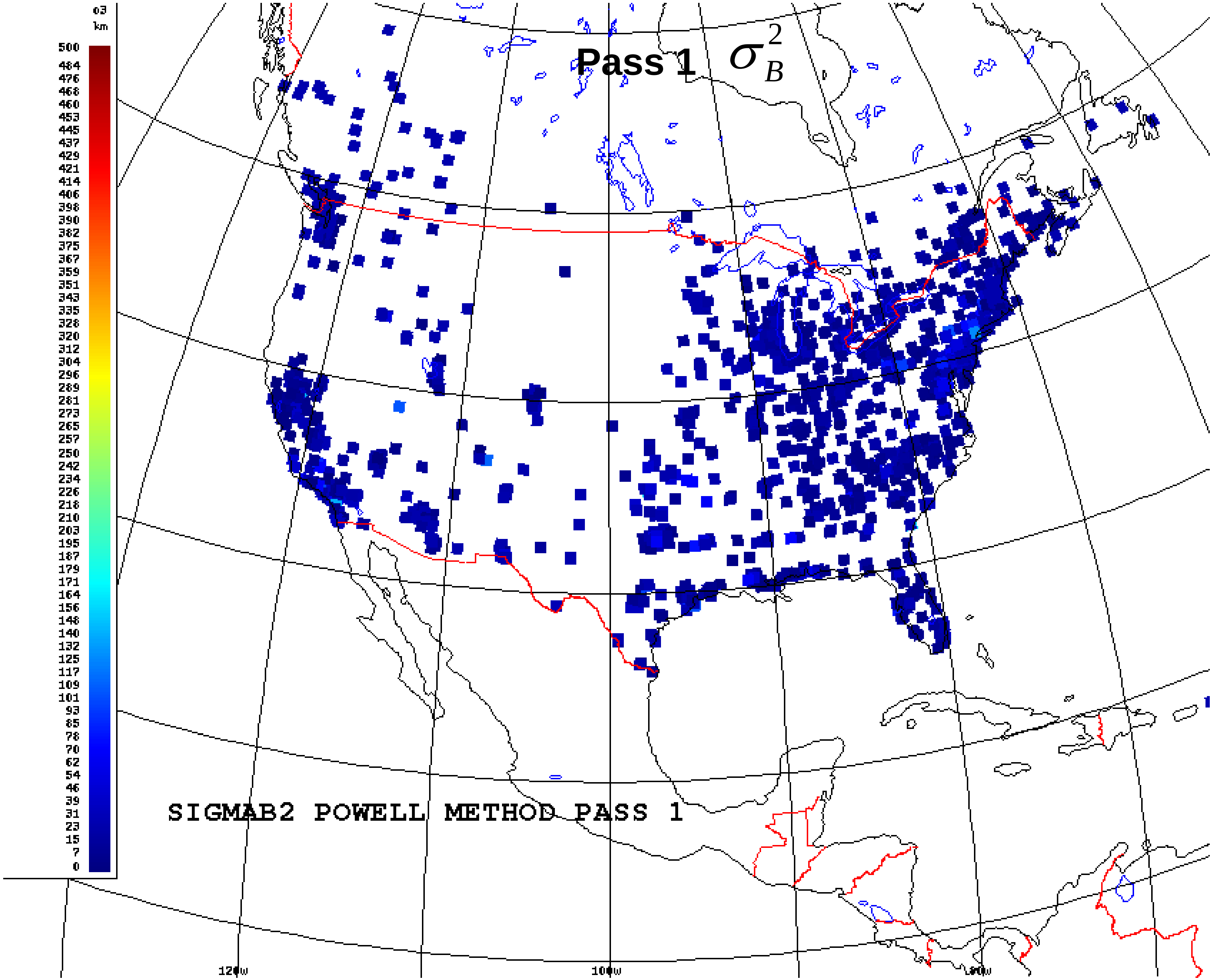
- and since $p \approx 1300$ we perform the full inversion of **B** + **R** without data selection

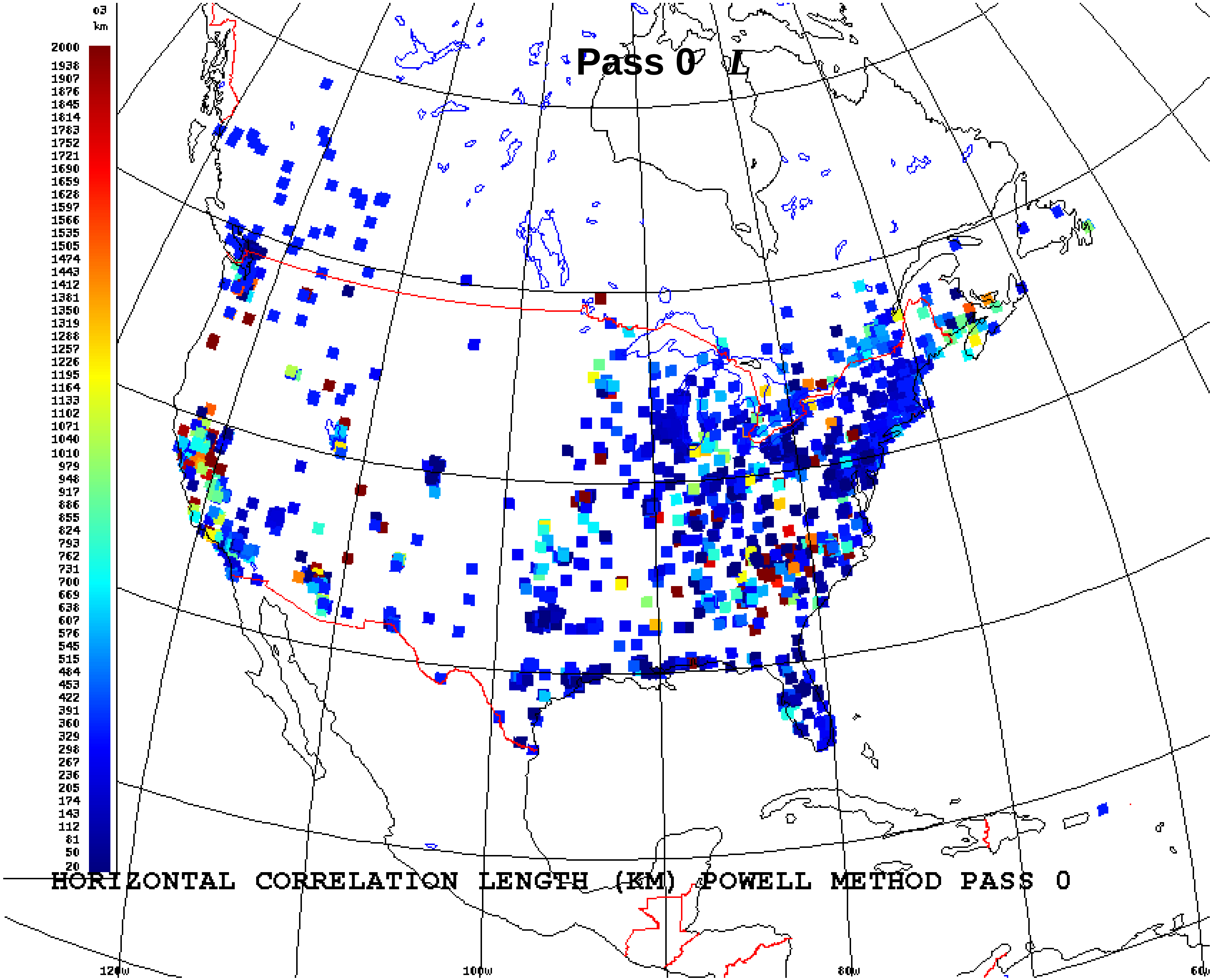
The implementation of an optimum interpolation scheme is actually an iterative process until

$$\underbrace{\left\langle OmF_i , OmF_j \right\rangle}_{\text{output assimilation}} \approx \underbrace{B(r_{i-obs} , r_{j-obs})}_{\text{input assimilation}} \quad \text{for} \quad i \neq j$$

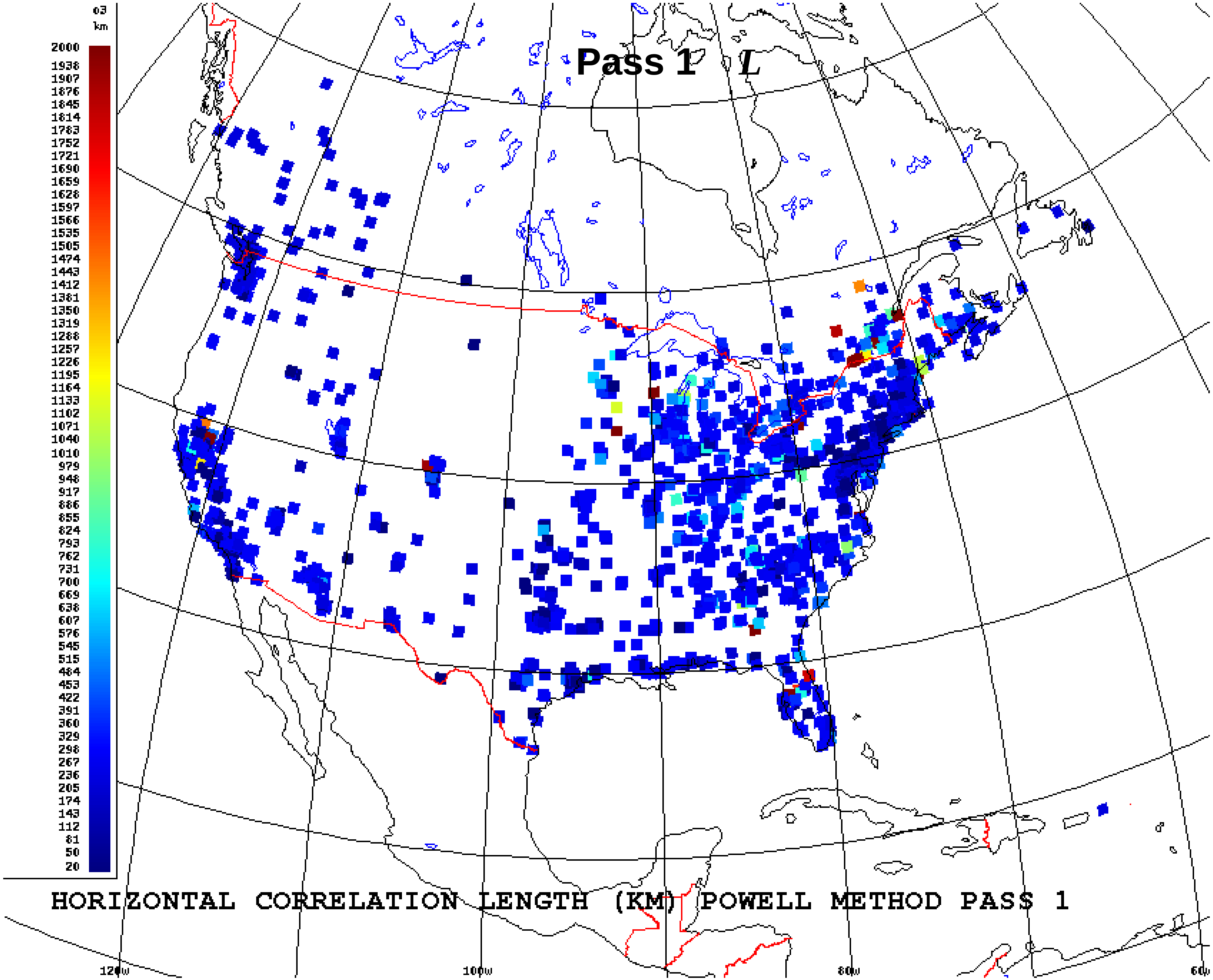
- **Pass 0** - No assimilation. Comparison of model simulation with observations in order to get a first guess on the error statistics
- **Pass 1** - Assimilation using Pass 0 error statistics
- **Pass 2** – Assimilation using Pass 1 error statistics


Pass 0 σ_B^2
o3
km
500
484
476
468
460
453
445
437
429
421
414
406
398
390
382
375
367
359
351
343
335
328
320
312
304
296
289
281
273
265
257
250
242
234
226
218
210
203
195
187
179
171
164
156
148
140
132
125
117
109
101
93
85
78
70
62
54
46
39
31
23
15
7
0
SIGMAB2 POWELL METHOD PASS 0
120W
100W
80W


Pass 1 σ_B^2
o3
km
500
484
476
468
460
453
445
437
429
421
414
406
398
390
382
375
367
359
351
343
335
328
320
312
304
296
289
281
273
265
257
250
242
234
226
218
210
203
195
187
179
171
164
156
148
140
132
125
117
109
101
93
85
78
70
62
54
46
39
31
23
15
7
0
SIGMAB2 POWELL METHOD PASS 1
120W
100W
80W



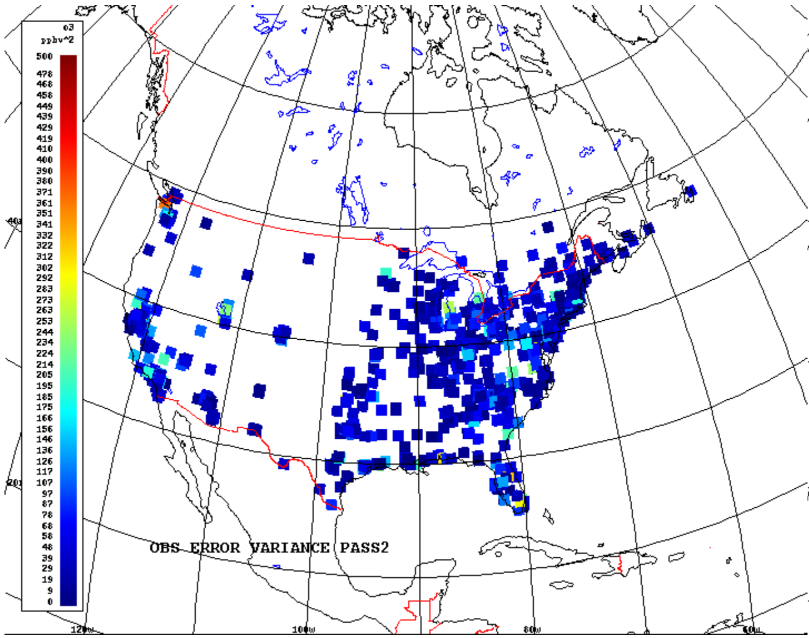
HORIZONTAL CORRELATION LENGTH (KM) POWELL METHOD PASS 0



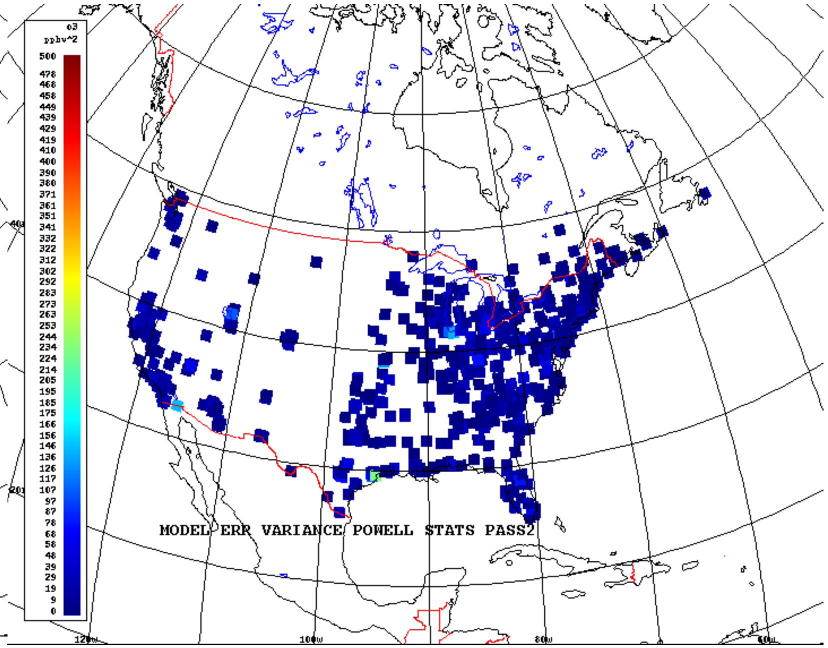
Pass 1 L
o3
km
2000
1938
1907
1876
1845
1814
1783
1752
1721
1690
1659
1628
1597
1566
1535
1505
1474
1443
1412
1381
1350
1319
1288
1257
1226
1195
1164
1133
1102
1071
1040
1010
979
948
917
886
855
824
793
762
731
700
669
638
607
576
545
515
484
453
422
391
360
329
298
267
236
205
174
143
112
81
50
20
HORIZONTAL CORRELATION LENGTH (KM) POWELL METHOD PASS 1
120w
100w
80w
60w

Pass 2 - Final

σ_B^2

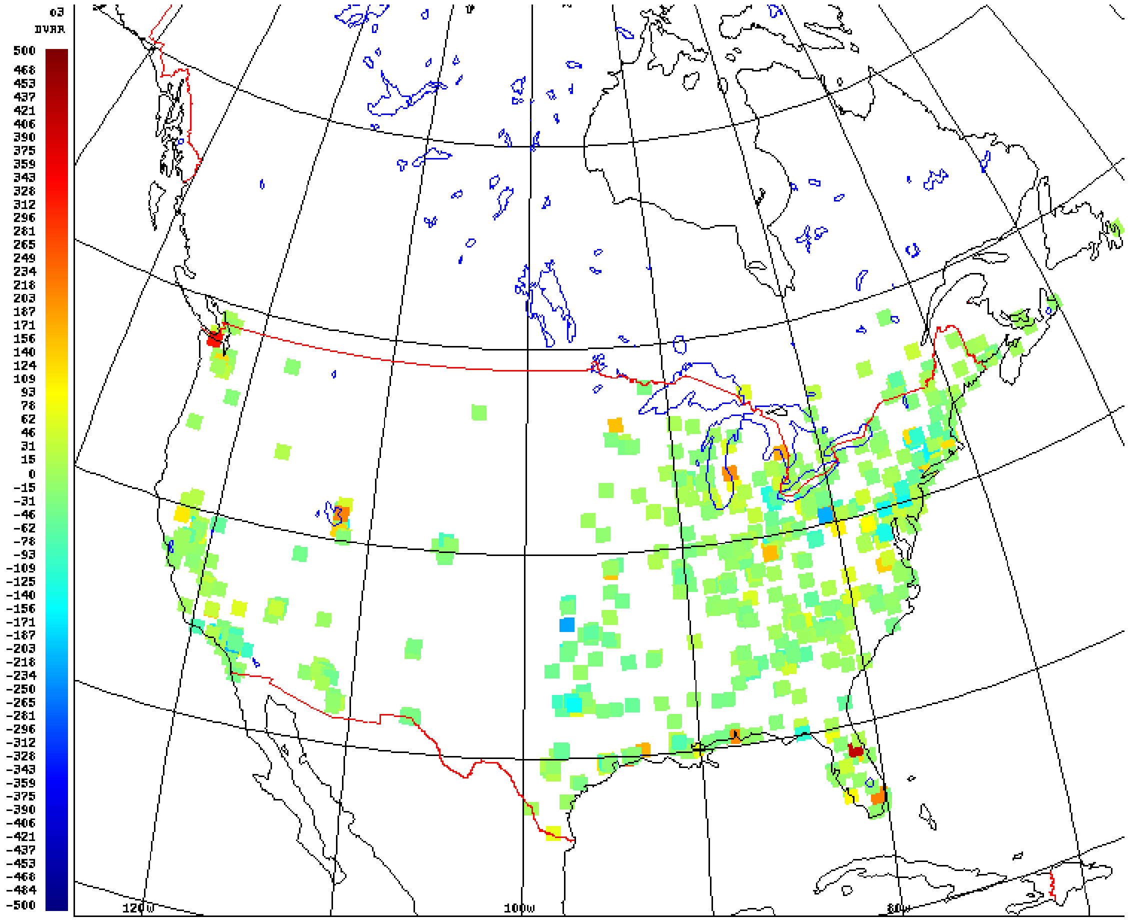


σ_O^2



c) statistical consistency

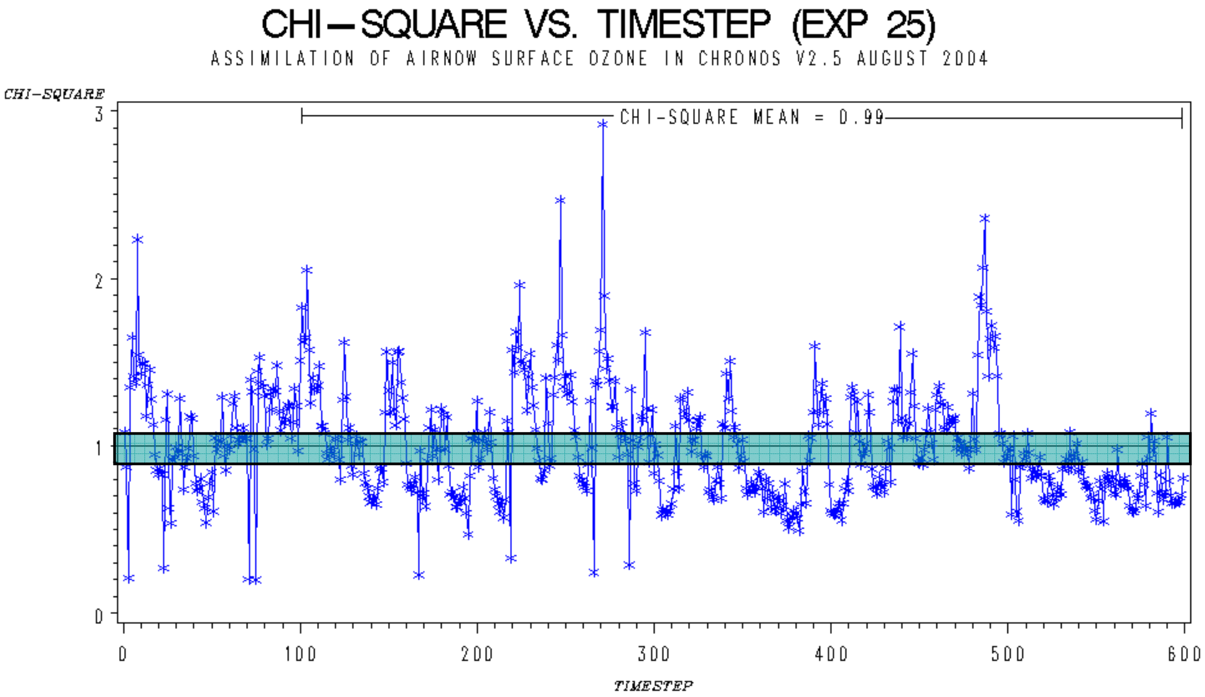
$$\left\langle (OmF)^2 \right\rangle - (\mathbf{B} + \mathbf{R})$$

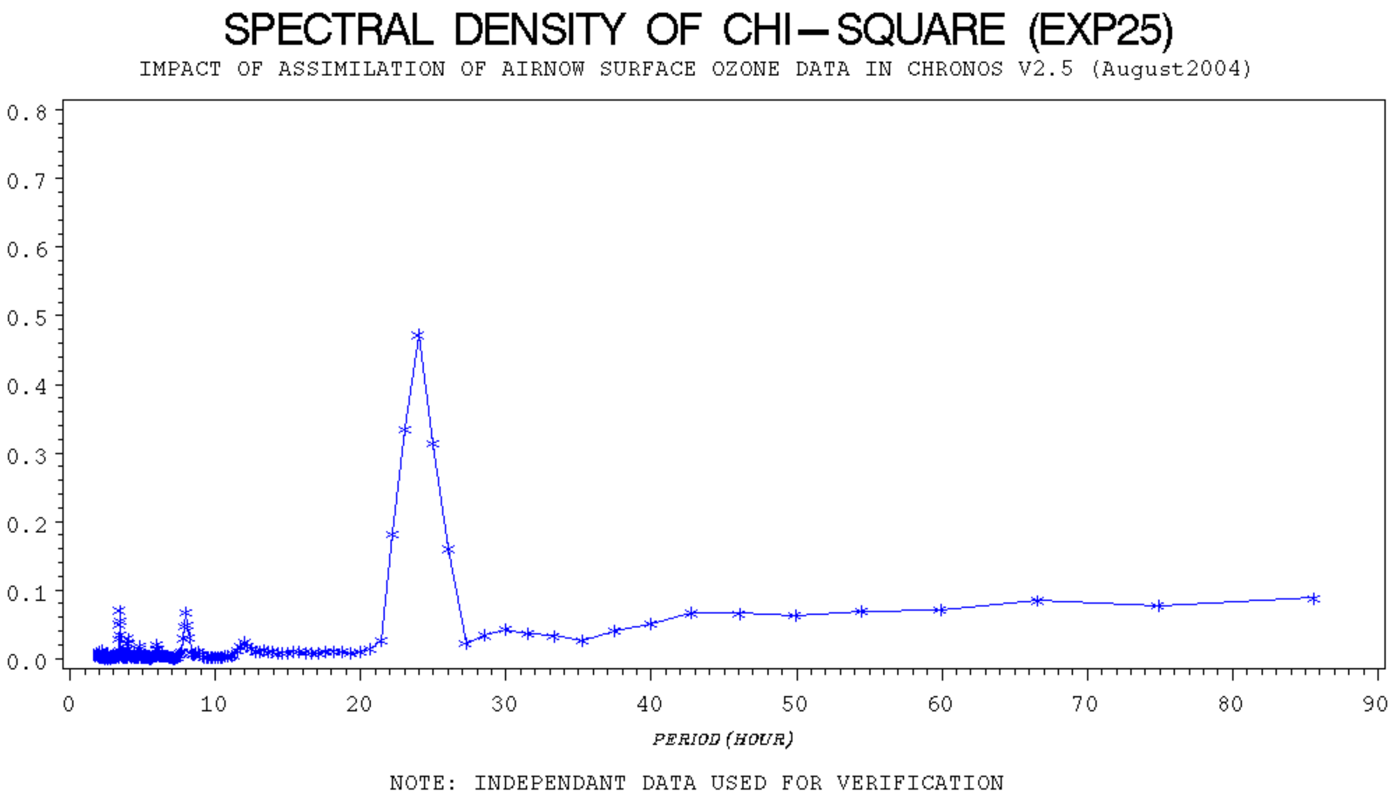


Monitoring of the error statistics

chi-square

$$\langle \chi^2 \rangle = \langle \nu^T (\mathbf{B}+\mathbf{R})^{-1} \nu \rangle = p$$





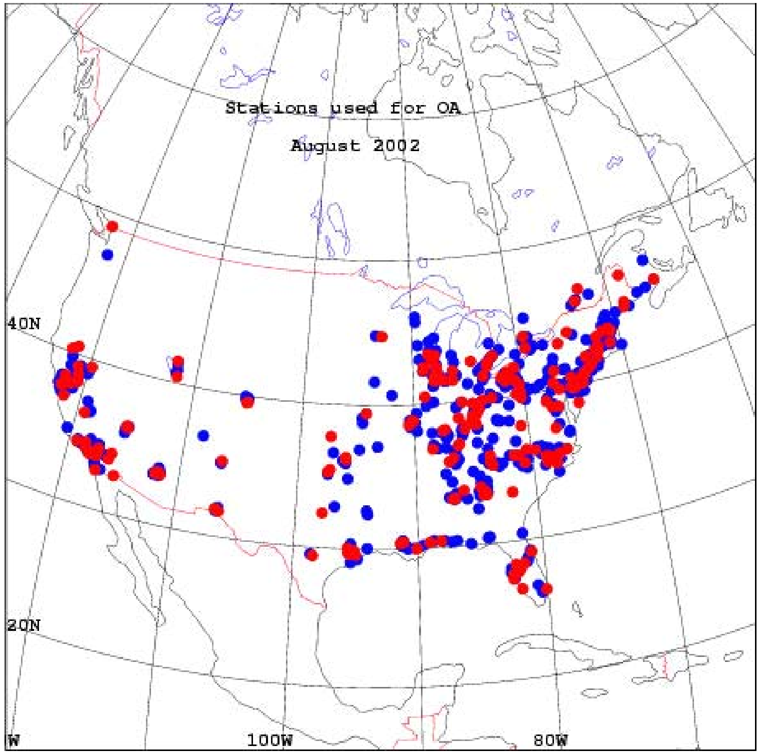
Power spectrum of χ^2 gives further insights on the weaknesses of the assimilation system.

- e.g. meteorological analyses are refreshed each 24 hours
- e.g. when error stat. were updates each 3 hours we saw a peak at 3 hours (results not shown)

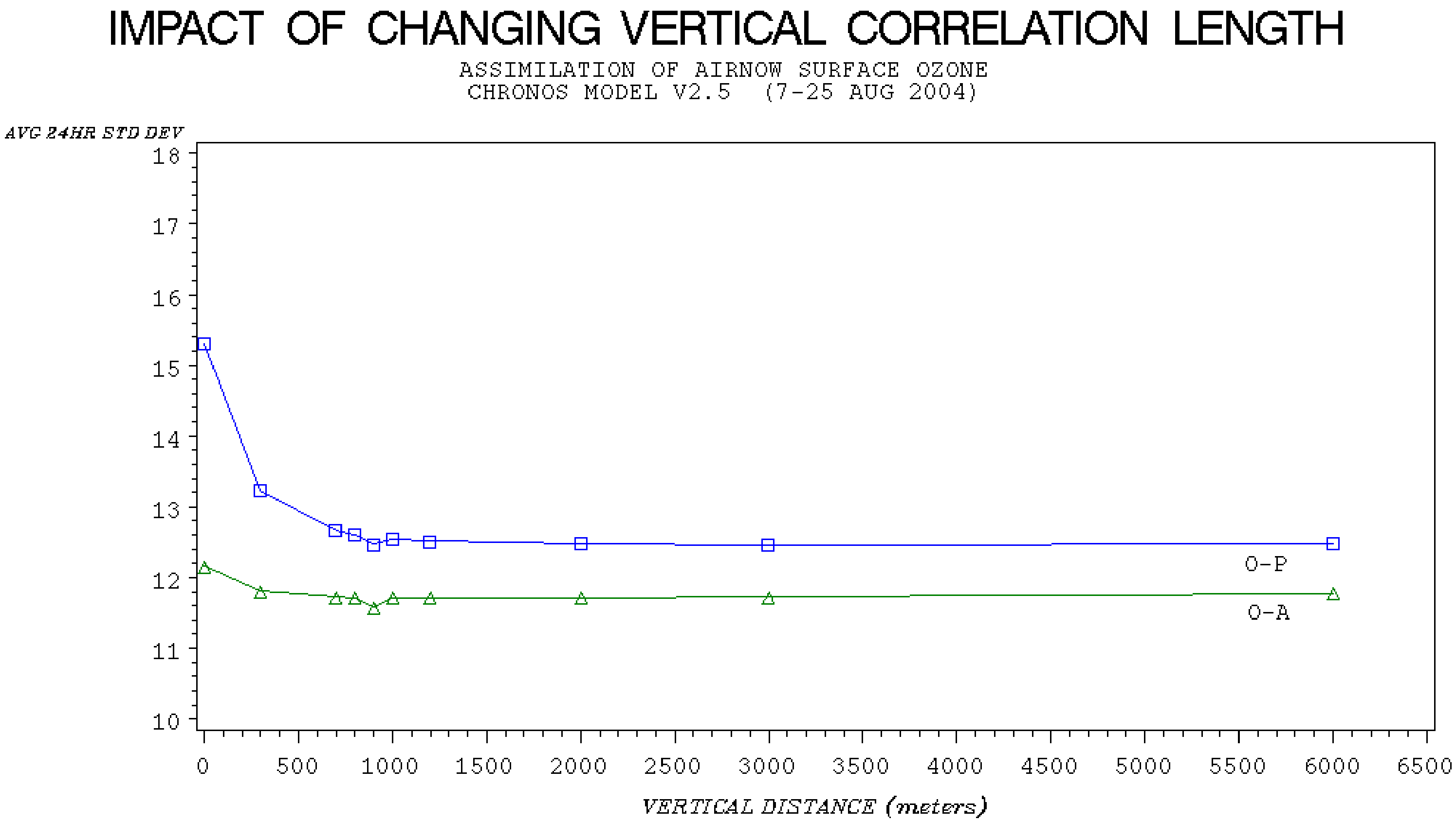
d) verification

To verify against independent observations we use

- 1/3 of observations for the verification (red)
- 2/3 of observation used to produce the analysis (blue)
- rederive the error statistics (because the background error may have changed)



e) background error vertical correlation



- The minimum error variance is achieved when the vertical correlation length scale is 900 m – i.e. correlation reaches zero just above the PBL.

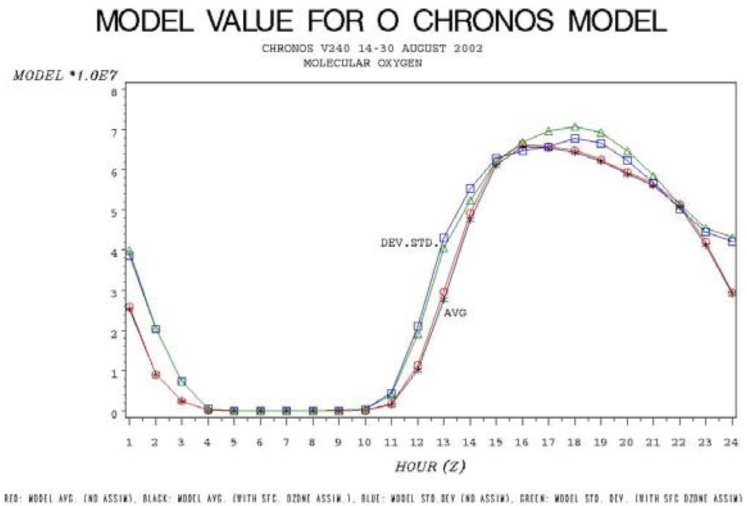
4. Other species

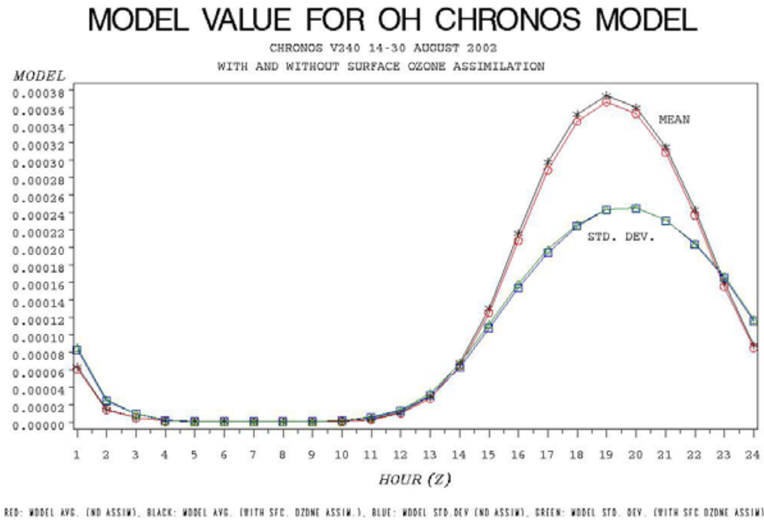
Error statistics during NARSTO 1995 campaign

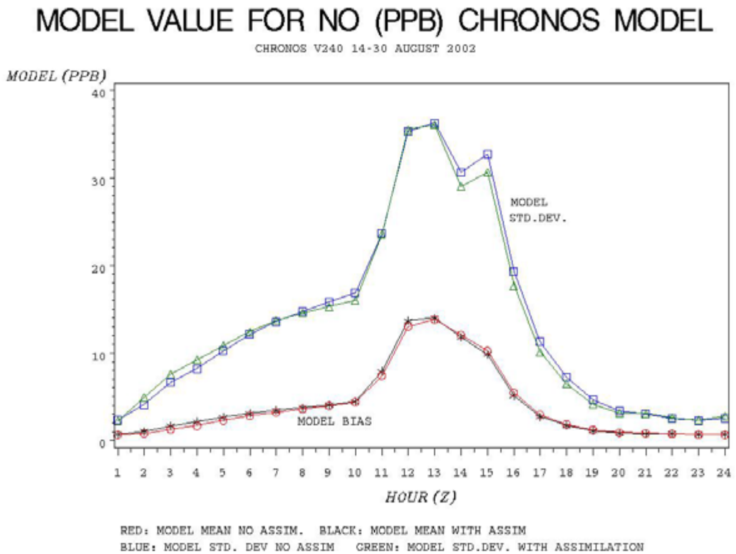
	August 2002, O_3 N=623	July 1995 O_3 N = 224	July 1995 NO_2 N=25	July 1995 NO N=39
Total variance	194	292	76	69
Forecast Error var	134	194	21	11
Obs error variance	60	98	55	58
Obs weight	0.72	0.68	0.48	0.22
Horiz corr lenth	310	370	690 ??	122 ??

→ **Too few observations to construct horizontal error statistics**

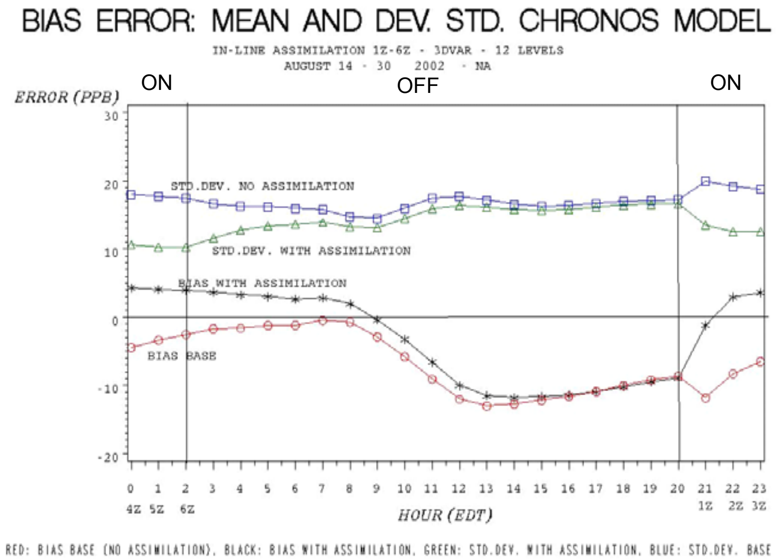
Impact of assimilating ozone on other species

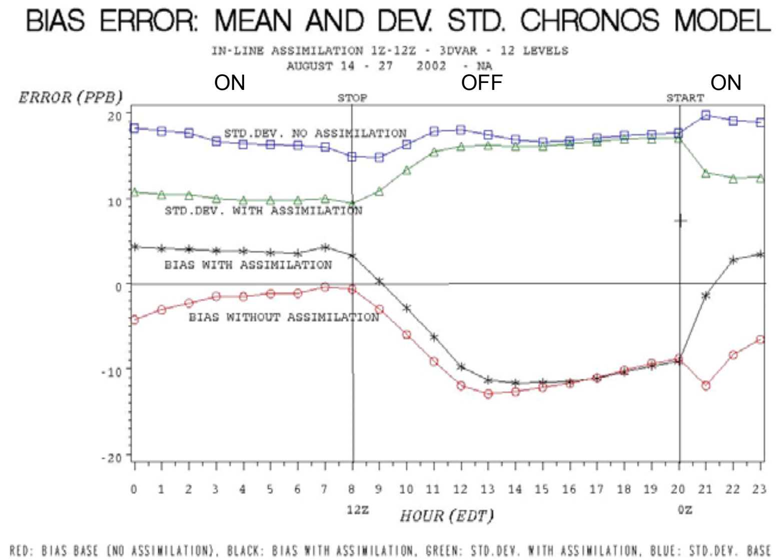




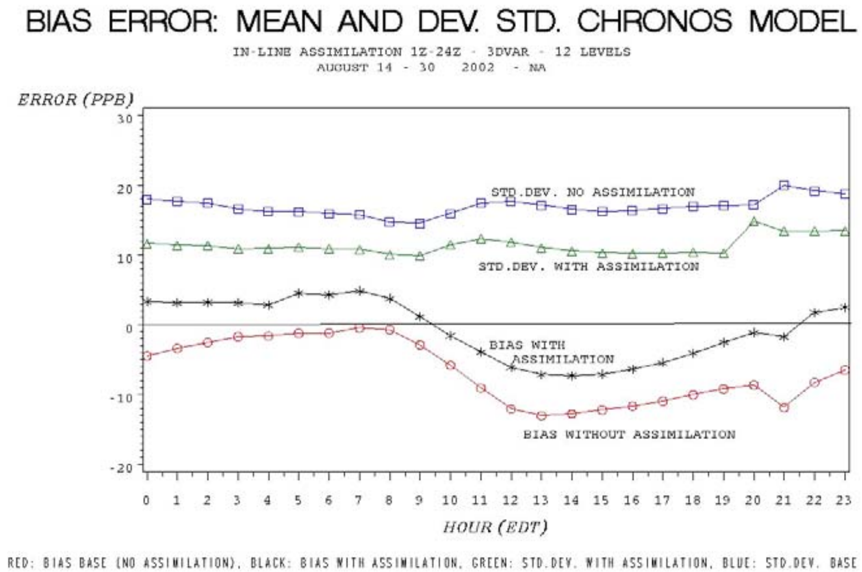


5. Prediction





continuous assimilation



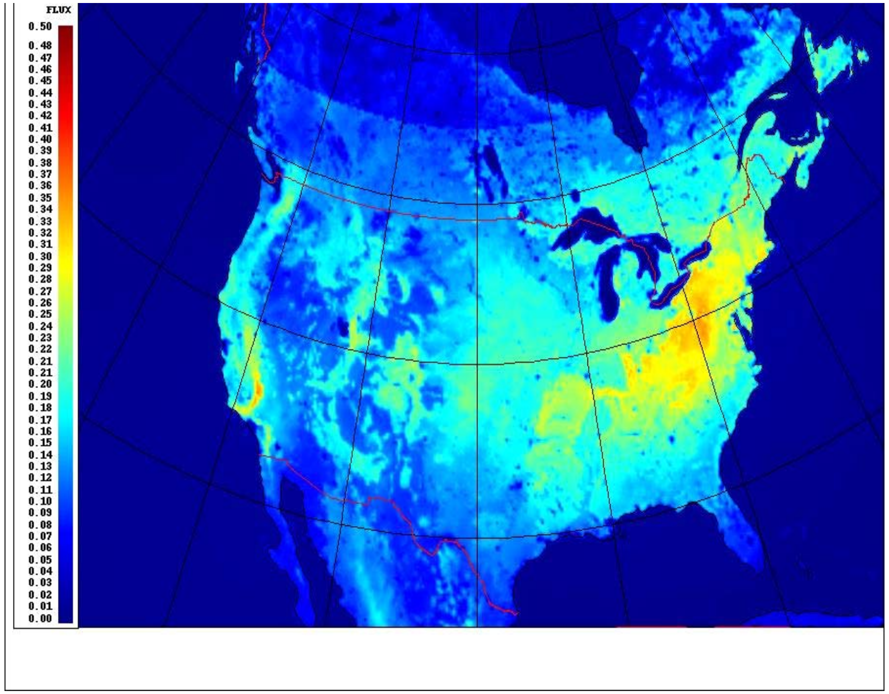
6. Environmental impact

• Dry deposition has an impact on crops and vegetation on a seasonal time scale

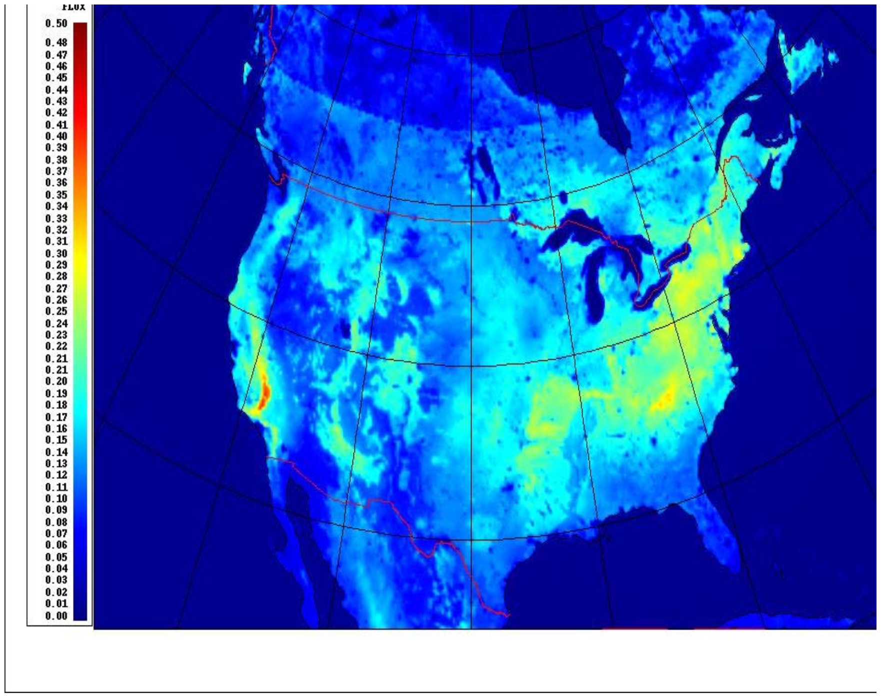
AVG. FLUX OF OZONE TO SURFACE : VD*[ozone] – Aug. 7-30 2002

ppb*m/s

ppb*m/s



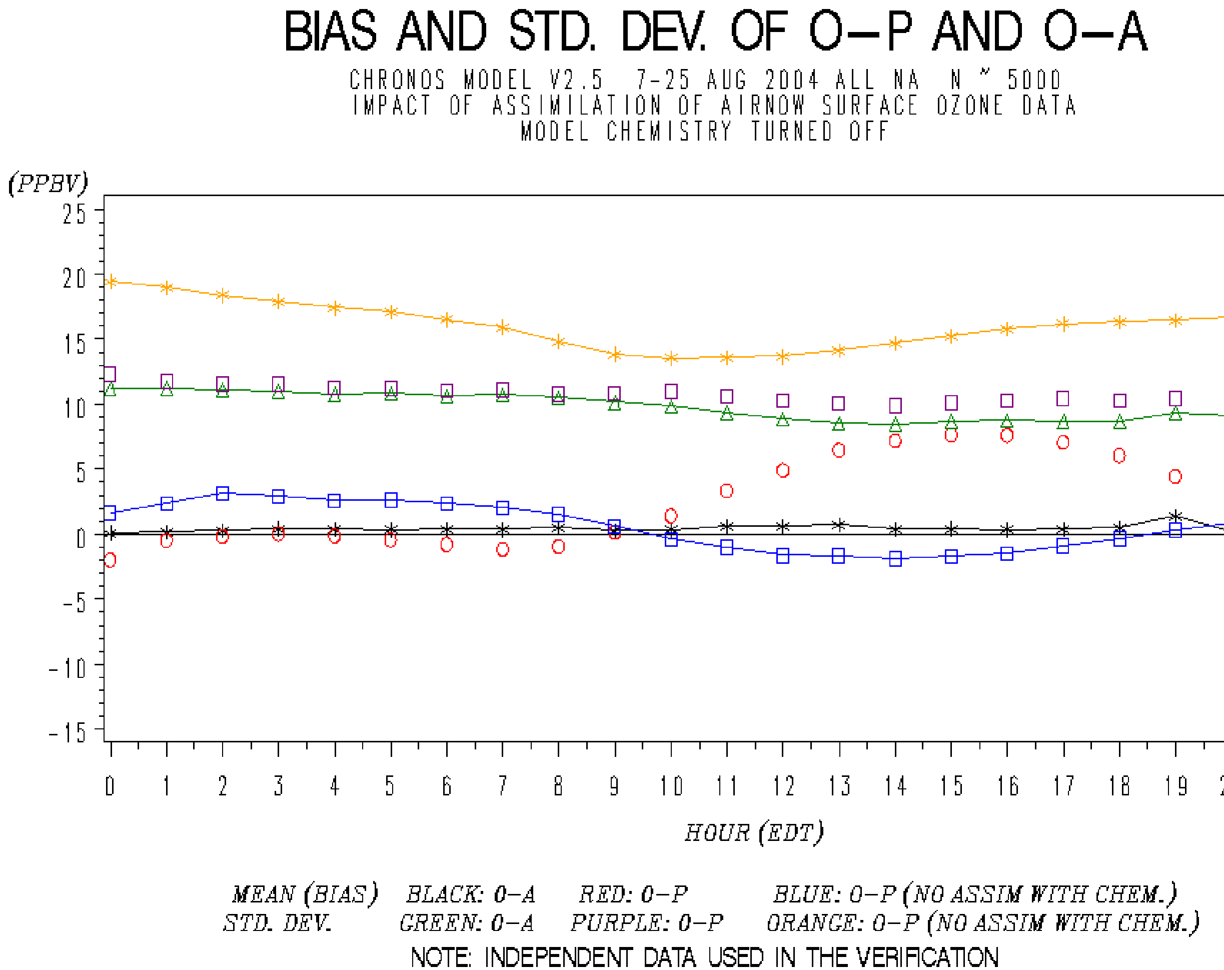
NO O_3 ASSIMILATION



WITH O3 ASSIMILATION

7. How important is chemistry in assimilation and monitoring

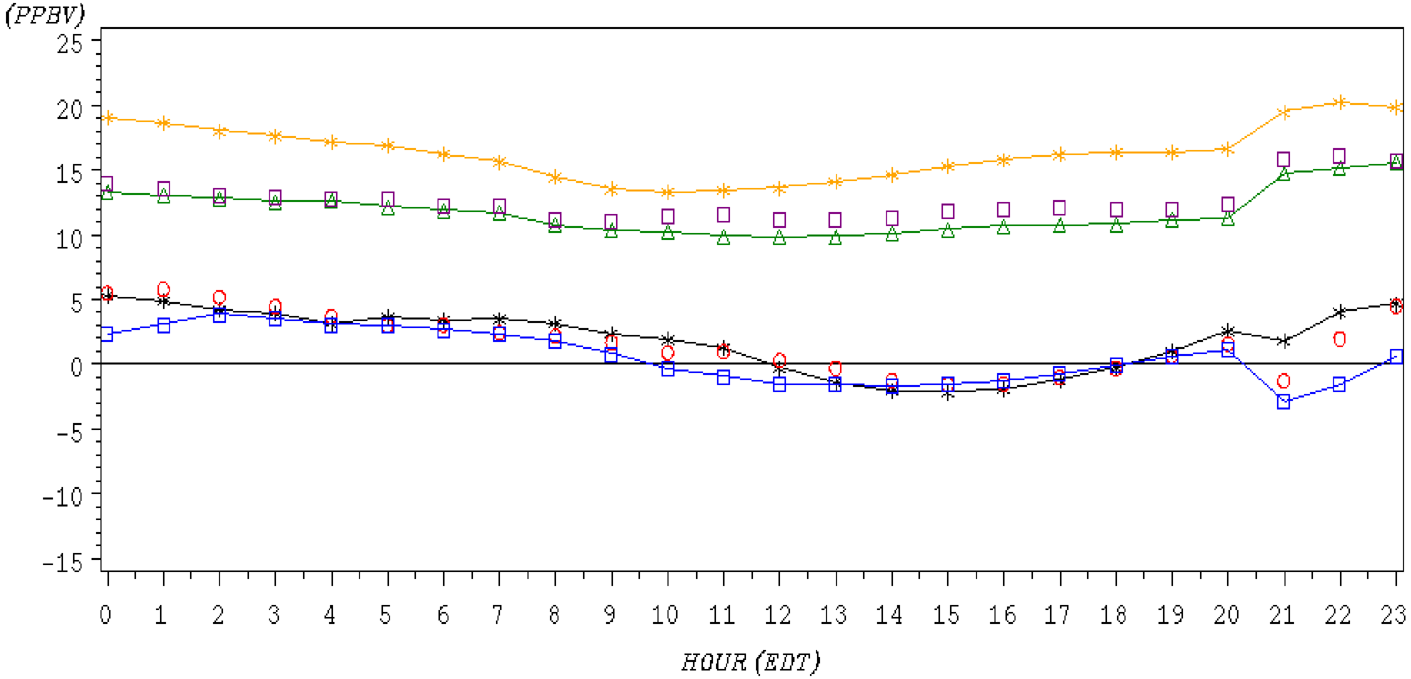
No chemistry



With chemistry

BIAS AND STD. DEV. OF O–P AND O–A

CHRONOS MODEL V2.5 7-25 AUG 2004 ALL NA N ~ 5000
IMPACT OF ASSIMILATION OF AIRNOW SURFACE OZONE DATA
POWELL METHOD FOR BACKGROUND ERROR STATISTICS



MEAN (BIAS) BLACK: O–A RED: O–P BLUE: O–P (NO ASSIM)
STD. DEV. GREEN: O–A PURPLE: O–P ORANGE: O–P (NO ASSIM)

NOTE: INDEPENDENT DATA USED IN THE VERIFICATION

In summary

- chemistry + free mode (no assimilation)
  - std 20 ppb
  - bias 2 ppb

- no chemistry + assimilation
  - std ~ 12 ppb
  - bias a) ~ 1 ppb in analyses
    b) large daytime (~10 ppb) in 1hr forecast

- chemistry + assimilation
  - std ~ 14 ppb
  - bias ~ 3 ppb (both analysis and forecast)

Some conclusions

- The reduction of error variance is largest when there is no chemistry. The bias is nearly zero in monitoring mode. The addition of chemistry reduces the effect of observations i.e. more resilient in reducing the error variance and in adapting the bias.

- The forecast and observation error variances are of comparable size.

- Reduction of forecast error variance due to surface observations is about one half.

- The impact of (univariate) assimilation of surface ozone observations on prediction last about 3hrs (e-folding time), creating very little corrections on other species.

- the vertical correlation length scale was found by minimizing the forecast error variance. Although barely detectable (i.e. observability ?), it indicates that correlation should penetrate just above the PBL.

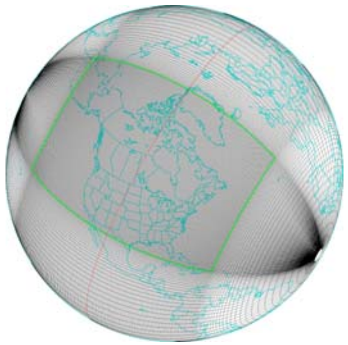
- surface flux of ozone with/without assimilation of observations is quite different on a seasonal time scale and may have an impact on the assessment of the environmental impact of ozone

8. Ongoing and future research

- Online chemistry with the operational NWP model GEM (Global Environmental Multiscale model)
- Extend the control variables of the operational 3D and 4D Var to include atmospheric compounds concentrations and surface fluxes.

Dynamics and physics

- Global Environmental Multiscale (GEM) model
  operational NWP model at Meteorological Service of Canada
  semi-Lagrangian, adjoint + TLM
  global uniform/variable resolution



- stratospheric version
  hybrid vertical coordinate $\sigma \rightarrow p$
  80 levels, top 0.1 hPa
  240 × 120 (1.5 degree)

- radiation, k-correlated method (Li and Barker 2004) uses as input
  H_2O, CO_2, O_3, N_2O, CH_4, CFC-11, CFC-12, CFC-113, CFC-114
  sulfate, sea salt, and dust aerosols.

- non-orographic gravity wave drag (Hines)

data assimilation system

- Stratospheric assimilation inherits the characteristics of the operational assimilation 3D Var and 4D Var
  - AMSU-A (*channel 10-14 added*) and AMSU-B microwave channels
  - GEOS infrared radiances
  - Data quality control with BG check and QC-Var
  - Conventional meteorological data

- Extension of operational 3D Var with an arbitrary number of of chemical species

- Chemical species BUFR format proposal to WMO (using IGACO chemical parameters + other AQ species)

chemistry

- Chemical interface to GEM (next official release)

- *Emissions handled through the physics interface (next year)*

- Kinetic PreProcessor (KPP) symbolic computation to generate
  production and loss terms
  jacobian, hessian, LU decomposition matrices.

- Look up J values and online J calculation
  (MESSy, Landgraf and Crutzen 1998).

- All species advected and gas phase chemistry solved
  with Rosenbrock or Fully implicit chemical solver (45 min time step).

- *Implementation of TLM and adjoint starting this fall.*

- *Adding surface fluxes as control variables (Next year)*

Computational Issues

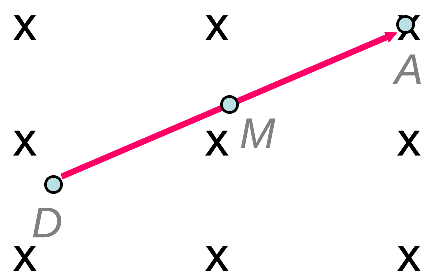
- Distributed computing / distributed memory

GCCM	OpenMP , MPI
VAR-CHEM	OpenMP , MPI (underway) , Analysis splitting

- Transport

Can save computation in semi-Lagrangian advection transport

• upstream point (*D* or *M*) is the same for all advected species



• interpolation weights $C_i(x)$ are the same for all advected species

e.g. cubic Lagrange interpolation $\varphi(x) = \sum_{i=1}^{4} C_i(x)\varphi_i$ with weights $C_i(x) = \dfrac{\prod_{k \neq i}^{4} (x - x_k)}{\prod_{k \neq i}^{4} (x_i - x_k)}$

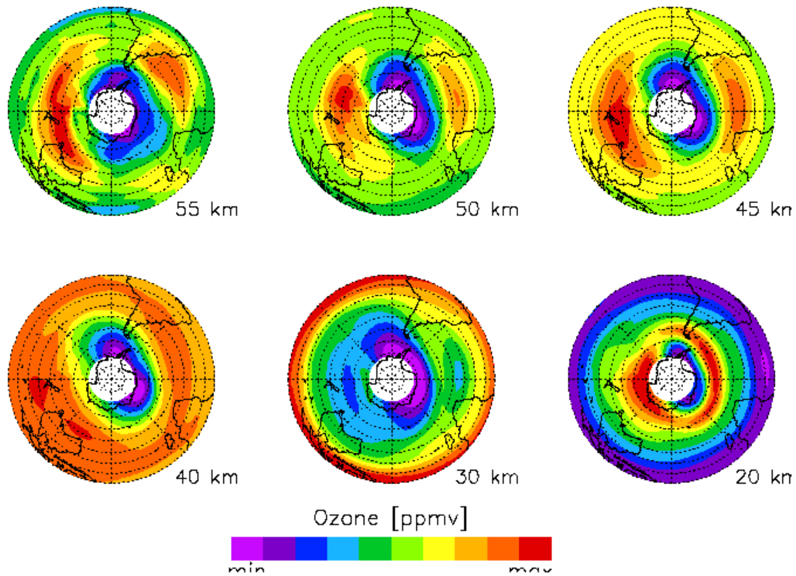
Data assimilation issues

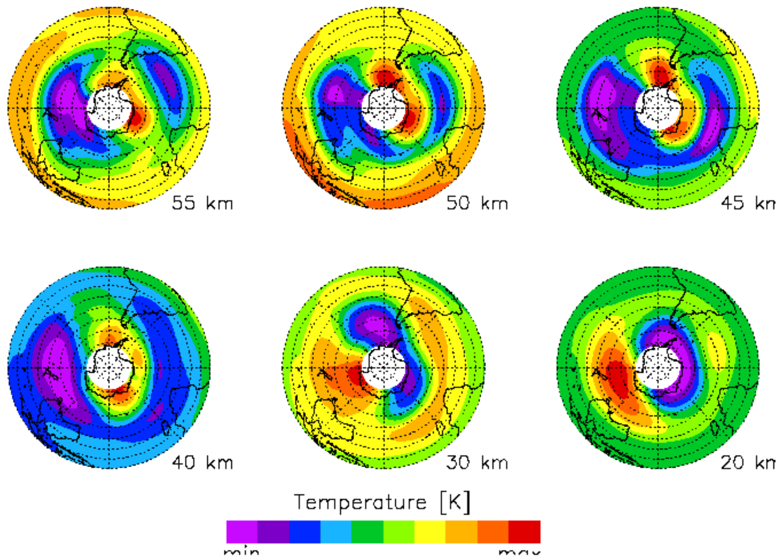
Cross-error covariance models
e.g. Temperature-Ozone

- Because the ozone production rate increases with decreasing temperatures, in regions dominated by photochemistry (above 35 km) a negative correlation between temperature and ozone would occur

- Haigh and Pyle (1982), Froideveau et al. 1989, Smith 1995, Ward 2002

$$[O_3] = B \exp\left(\frac{\Theta}{T}\right)$$





- For data at a given level, perturbations can fit an expression of the form

$$\frac{\delta[O_3]}{[O_3]} = -\frac{c}{T^2}\delta T$$

with a correlation that can be up to 0.92 above 42 km, and increase linearly from zero to 0.92 between 37 km to 42 km.

- Cross error coupling in 3D Var

$$\begin{pmatrix} \psi \\ \chi \\ (T, p_s) \\ \ln q \\ \ln[O_3] \end{pmatrix} = \begin{pmatrix} I & 0 & 0 & 0 & 0 \\ E & I & 0 & 0 & 0 \\ N & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ M & 0 & F & 0 & I \end{pmatrix} \begin{pmatrix} \psi \\ \chi_u \\ (T_u, p_{su}) \\ \ln q \\ \ln[O_3]_u \end{pmatrix}$$

Not all chemical species are observed

Analysis splitting → only observed variables in control vector

The problem of minimizing

$$J(\mathbf{x},\mathbf{u}) = \frac{1}{2}\begin{pmatrix}\mathbf{x}-\mathbf{x}^f \\ \mathbf{u}-\mathbf{u}^f\end{pmatrix}^T \begin{pmatrix}\mathbf{P}_{xx}^f & \mathbf{P}_{xu}^f \\ \mathbf{P}_{ux}^f & \mathbf{P}_{uu}^f\end{pmatrix}^{-1} \begin{pmatrix}\mathbf{x}-\mathbf{x}^f & \mathbf{u}-\mathbf{u}^f\end{pmatrix} + \frac{1}{2}(\mathbf{y}-H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}-H(\mathbf{x}))$$

*with respect to **x** and **u** is mathematically equivalent to minimizing*

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

followed by the update (Ménard et al. 2004)

$$\mathbf{u}^a = \mathbf{u}^f + \mathbf{P}_{ux}(\mathbf{P}_{xx})^{-1}(\mathbf{x}^a - \mathbf{x}^f)$$

- **4D Var extension**

Uses same solver as in 3D Var

tangent linear integration

$$J(\xi) = \frac{1}{2}(\xi-\xi^f)^T(\xi-\xi^f) + \frac{1}{2}(\mathbf{y}-\mathbf{H}(\mathbf{L}(\xi)))^T \mathbf{R}^{-1}(\mathbf{y}-\mathbf{H}(\mathbf{L}(\xi)))$$

End …