

GSI Tutorial 2013

Background and Observation Error Estimation and Tuning

Analysis system produces an analysis through the minimization of an objective function given by

$$J = \mathbf{x}^T \mathbf{B}^{-1} \mathbf{x} + (\mathbf{H} \mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{x} - \mathbf{y})$$

J^b

J^o

Where

\mathbf{x} is a vector of analysis increments: ψ , χ_u , T_u , P_{ui} , and q

\mathbf{B} is the background error covariance matrix

\mathbf{y} is a vector of the observational residuals, $\mathbf{y} = \mathbf{y}_{\text{obs}} - \mathbf{H} \mathbf{x}_{\text{guess}}$

\mathbf{R} is the observational and representativeness error covariance matrix

\mathbf{H} is the transformation operator from the analysis variable to the form of the observations.

Background Error

1. Background error covariance
2. Multivariate relation
3. Covariance with fat-tailed power spectrum
4. Estimate background error

One ob test

&SETUP

....

oneobtest=.true.

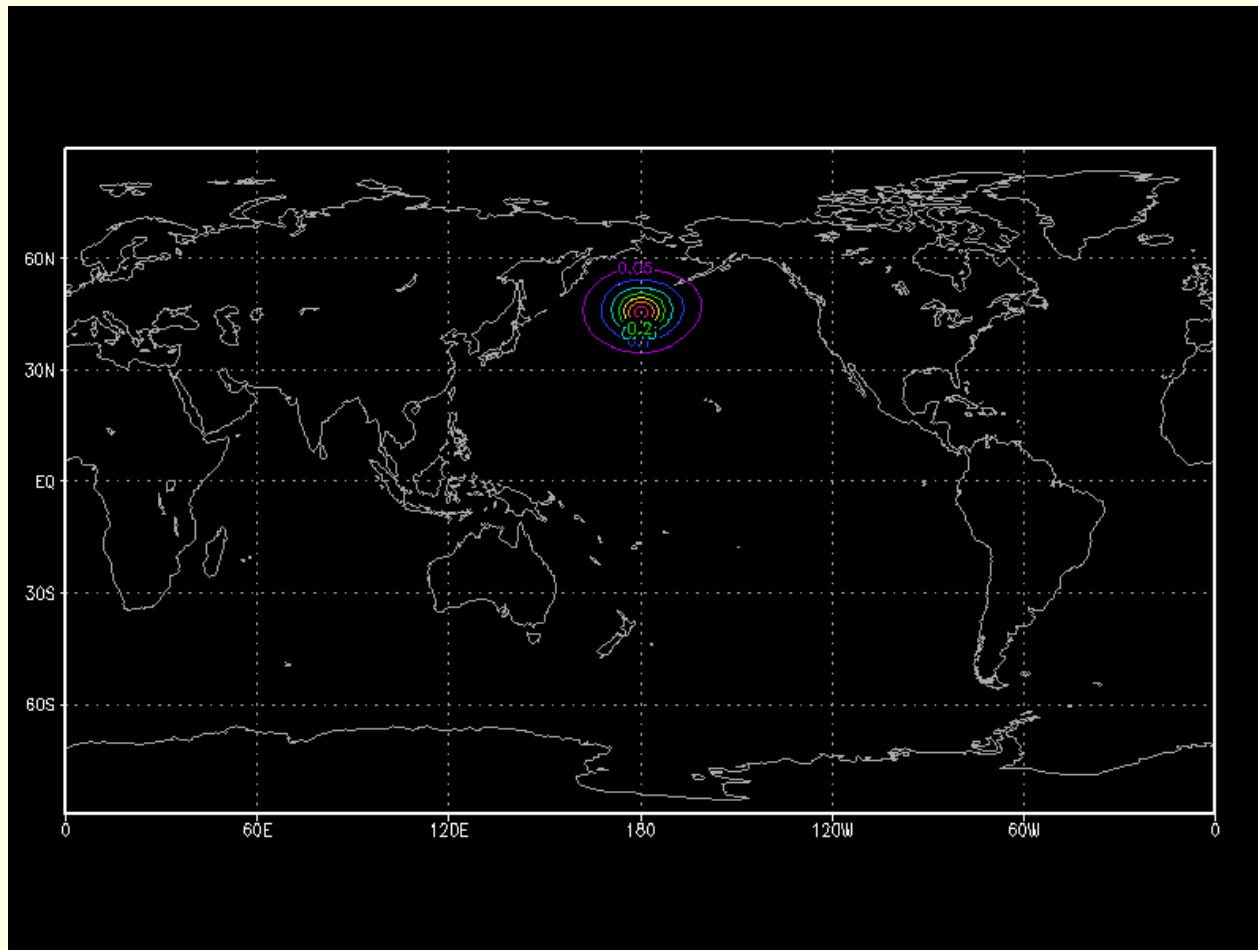
&SINGLEOB_TEST

maginnov=1.,magoberr=1.,oneob_type='t' ,

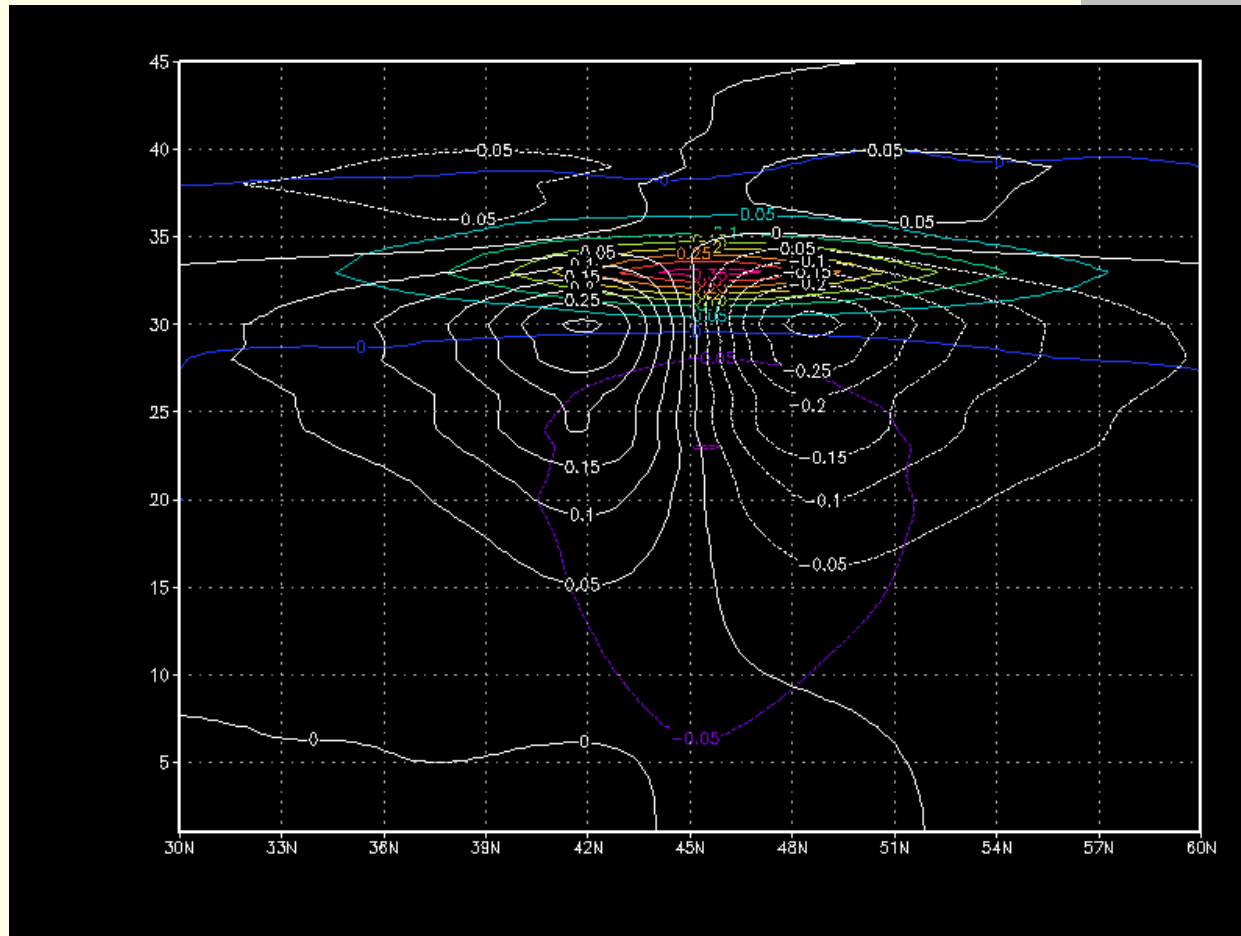
oblat=45.,oblon=270.,obpres=850 (250).,

obdattime=2013072900,obhourset=0.,

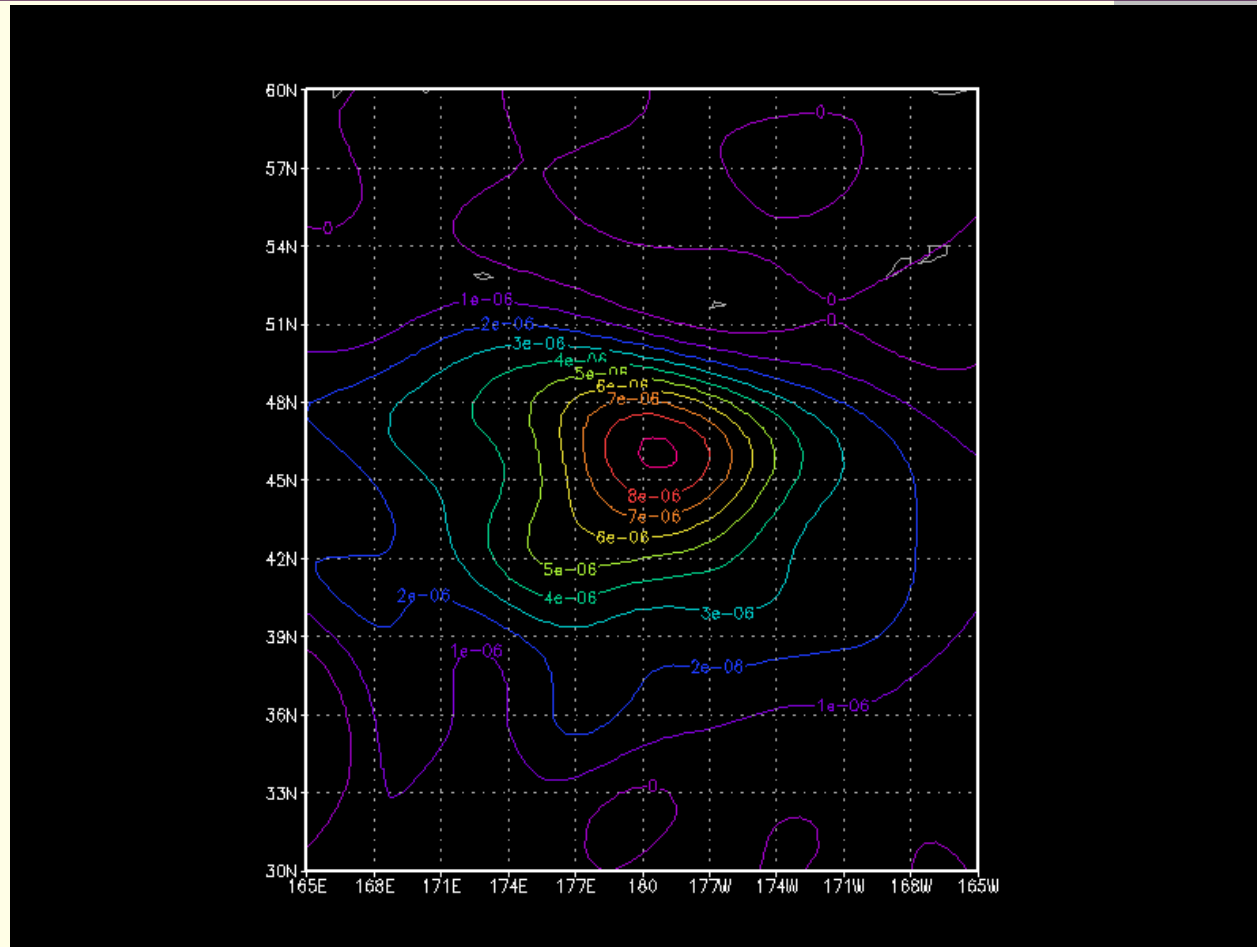
Temp Analysis Increment from Single Temp obs



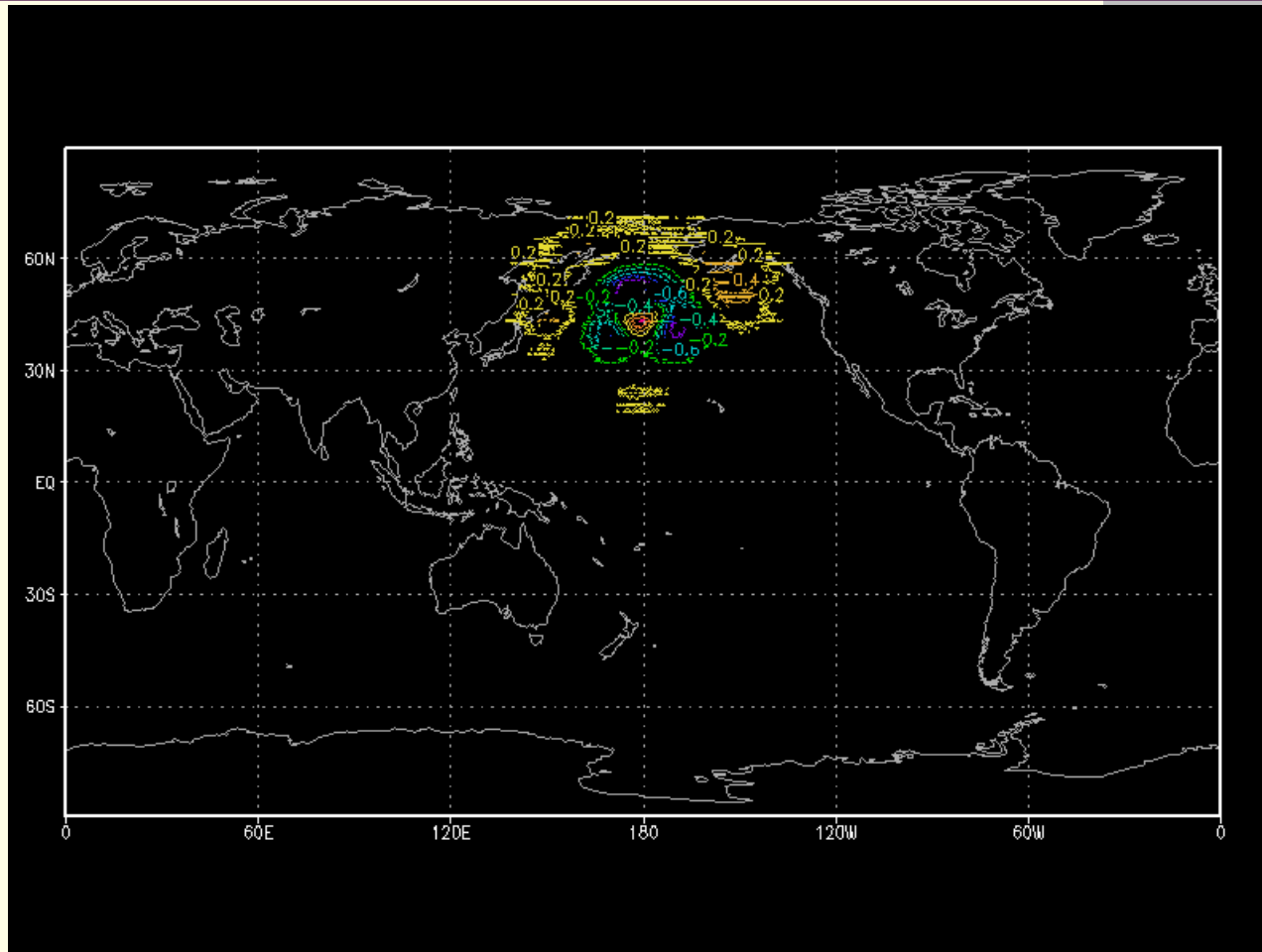
Temp and E-W Wind Analysis Increment from 1 Temp obs



Humidity Analysis Increment from Single Temp obs



Surface Pressure Analysis Increment from Single Temp obs



Multivariate relation

- ❖ Balanced part of the temperature is defined by

$$T_b = G \psi$$

where G is an empirical matrix that projects increments of stream function at one level to a vertical profile of balanced part of temperature increments. G is latitude dependent.

- ❖ Balanced part of the velocity potential is defined as

$$\chi_b = c \psi$$

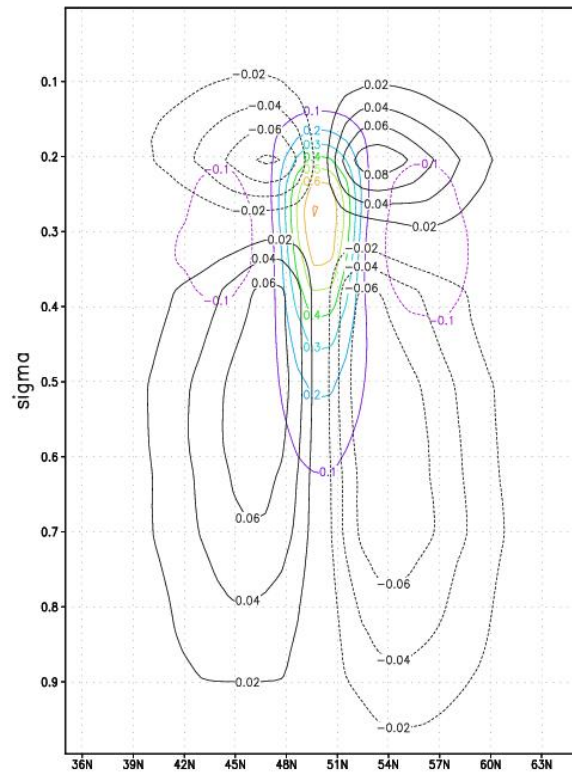
where coefficient c is function of latitude and height.

- ❖ Balanced part of the surface pressure increment is defined as

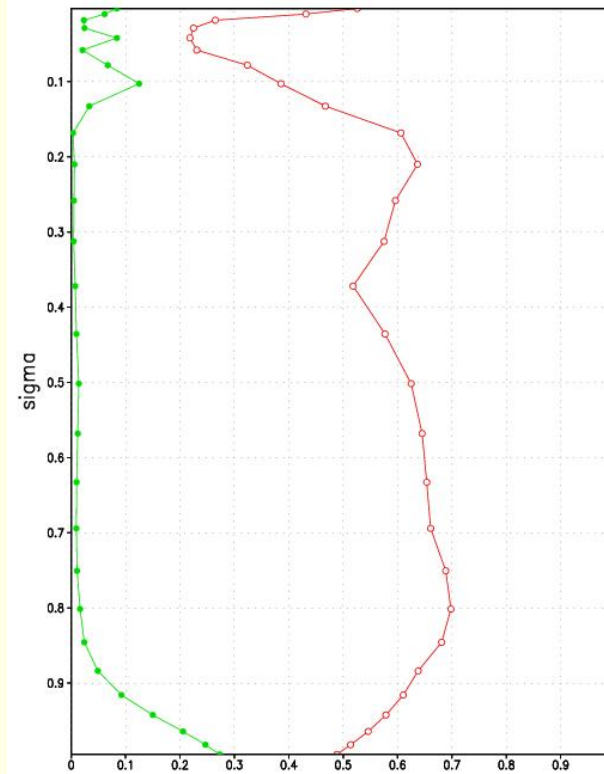
$$P_b = W \psi$$

where matrix W integrates the appropriate contribution of the stream function from each level.

Multivariate Relation



Vertical cross section of u and temp



global mean fraction of balanced temperature and velocity potential

Control Variable and Error Variances

Normalized relative humidity (qoption=2)

$$\delta RH / \sigma(RH^b) = RH^b (\delta P/P^b + \delta q /q^b - \delta T /\alpha^b)$$

where

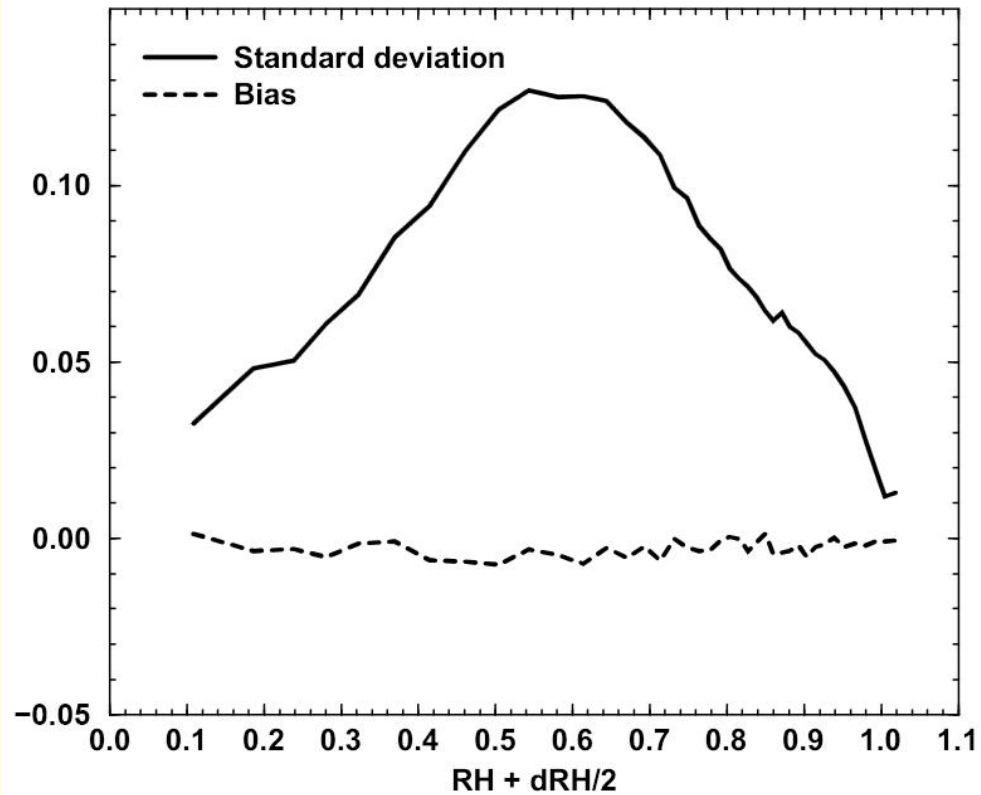
$\sigma (RH^b)$: standard deviation of background error as function of RH^b

α^b : - 1 / d(RH)/d(T)

multivariate relation with Temp and P

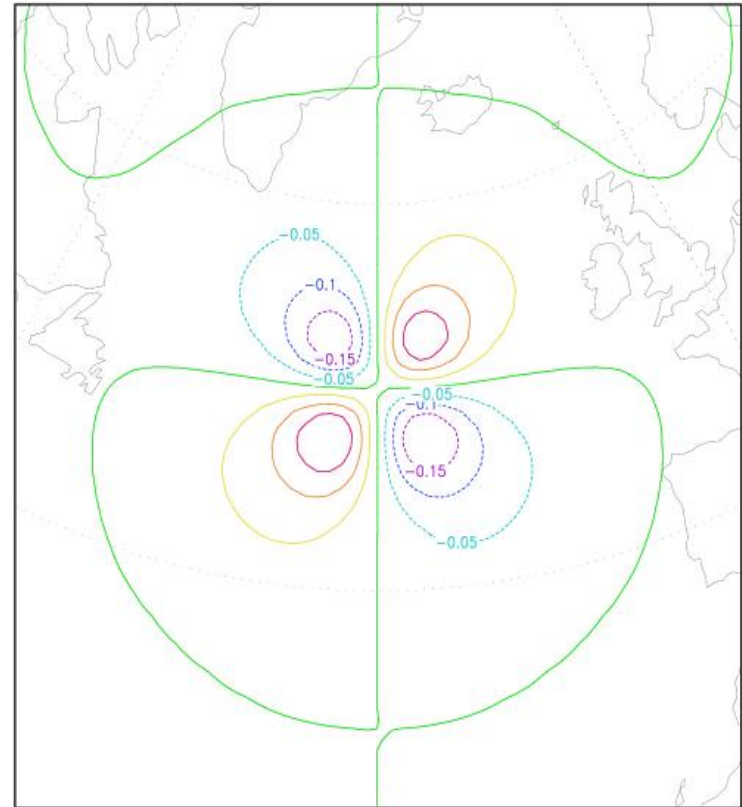
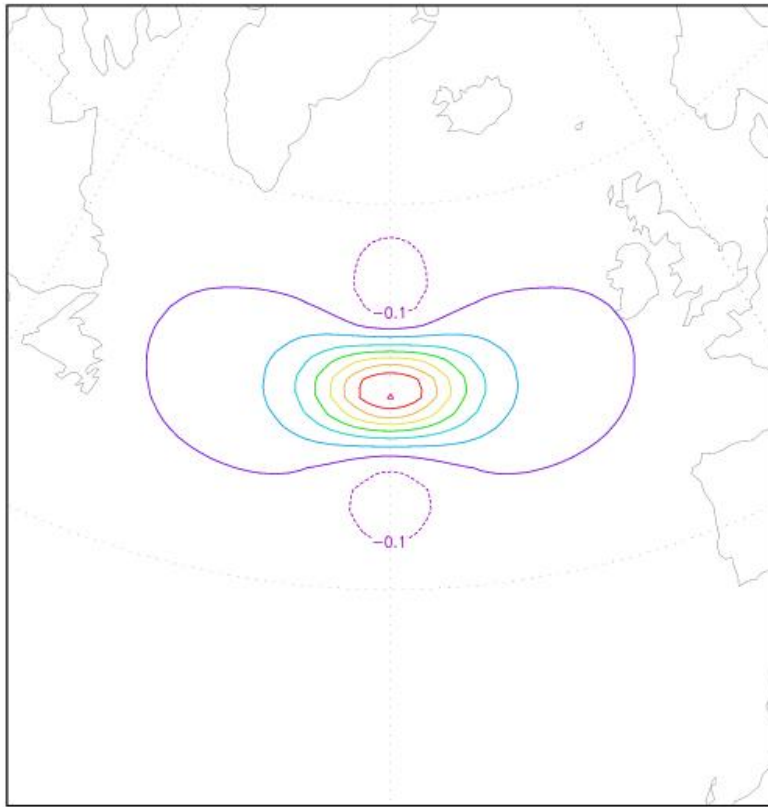
* Holm et al.(2002) ECMWF Tech Memo

Background error variance of RH



*Figure 23 in Holm et al.(2002) ECMWF Tech Memo

Control Variables and Model Variables



U (left) and v (right) increments at sigma level 0.267, of a 1 m/s westerly wind observational residual at 50N and 330 E at 250 mb.

Background error covariance

The error statistics are estimated in grid space with the ‘NMC’ method.

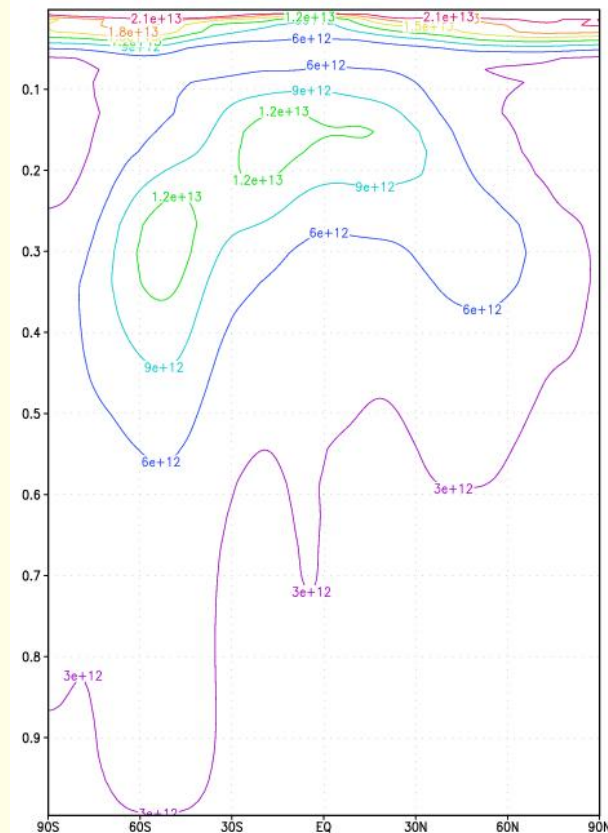
Stats of stream function are shown.

Stats are function of latitude and sigma level.

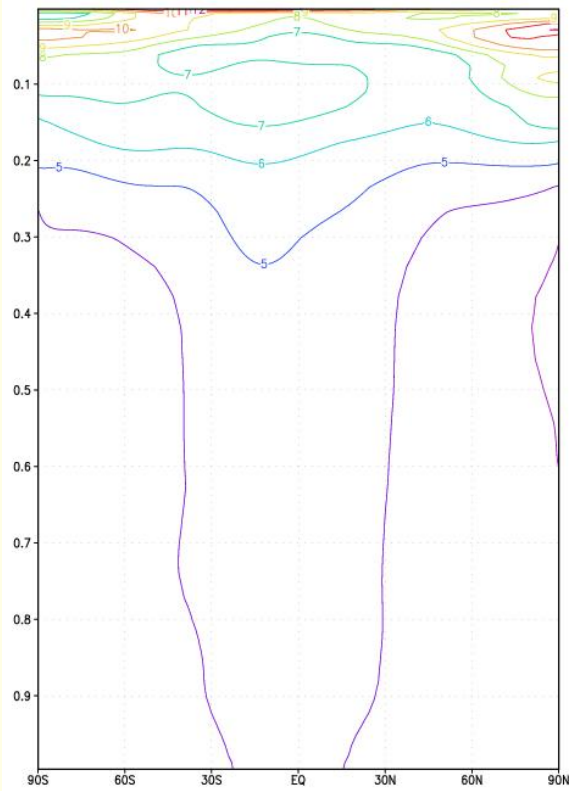
The error variance ($\text{m}^4 \text{s}^{-2}$) is larger in mid-latitudes than in the tropics and larger in the southern hemisphere than in the northern.

The horizontal scales are larger in the tropics, and increases with height.

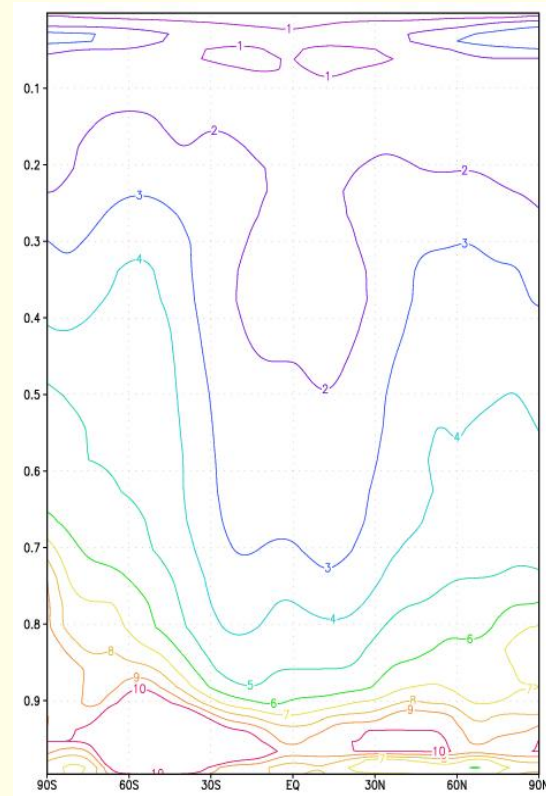
The vertical scales are larger in the mid-latitude, and decrease with height.



Background error covariance

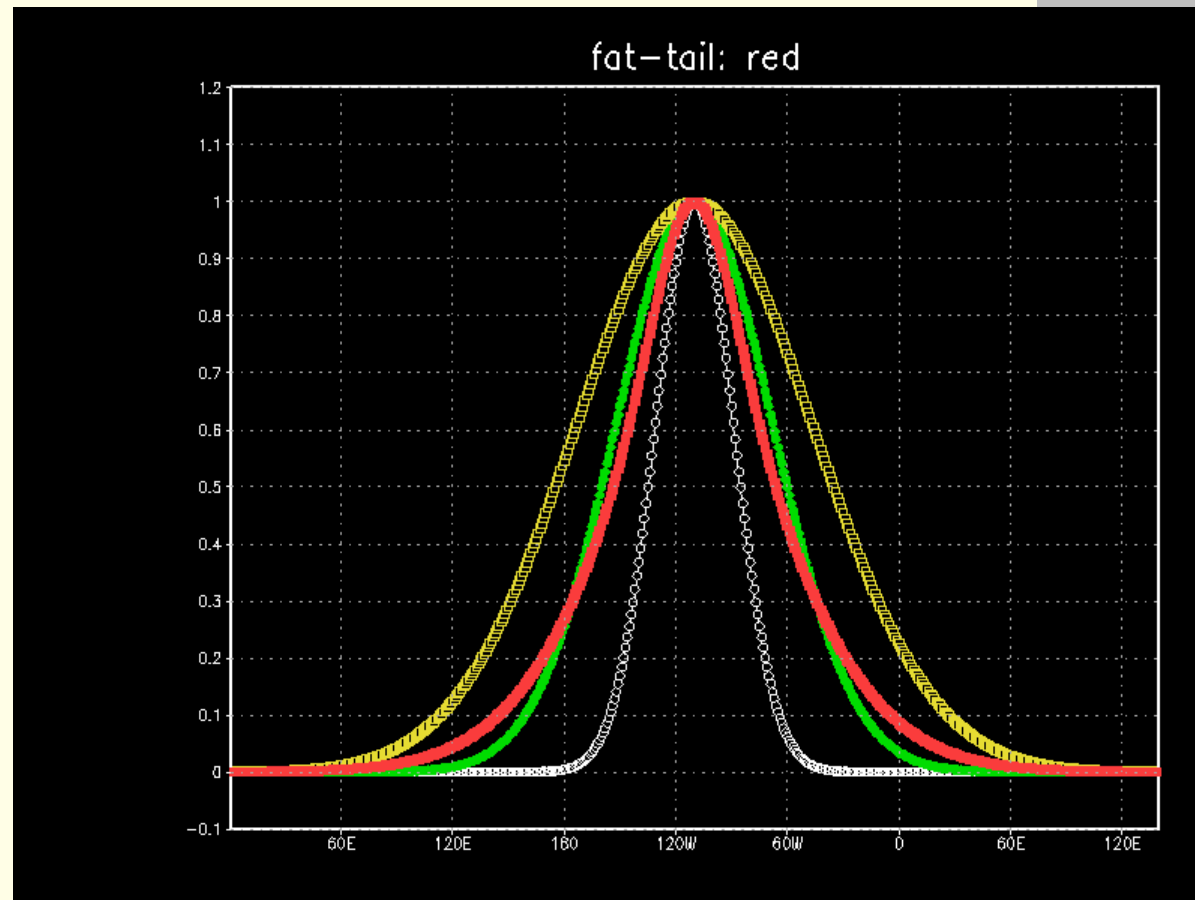


Horizontal scales in units of 100km

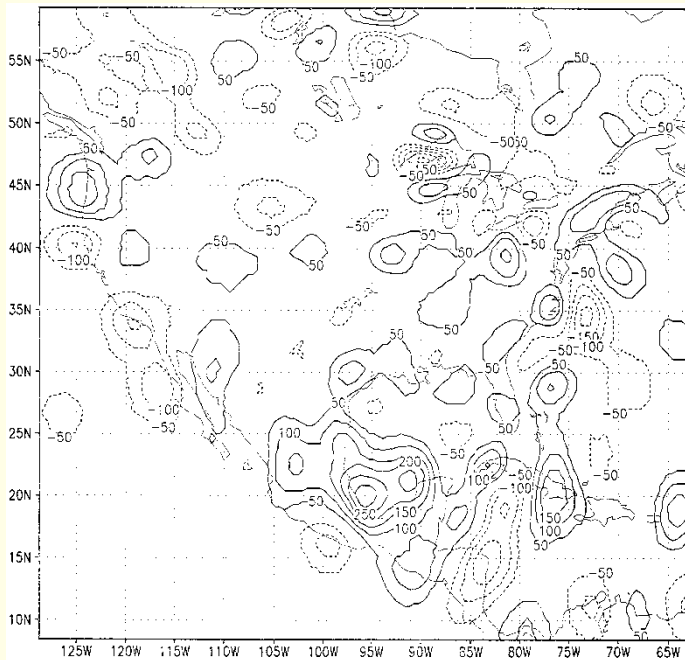


vertical scale in units of vertical grid

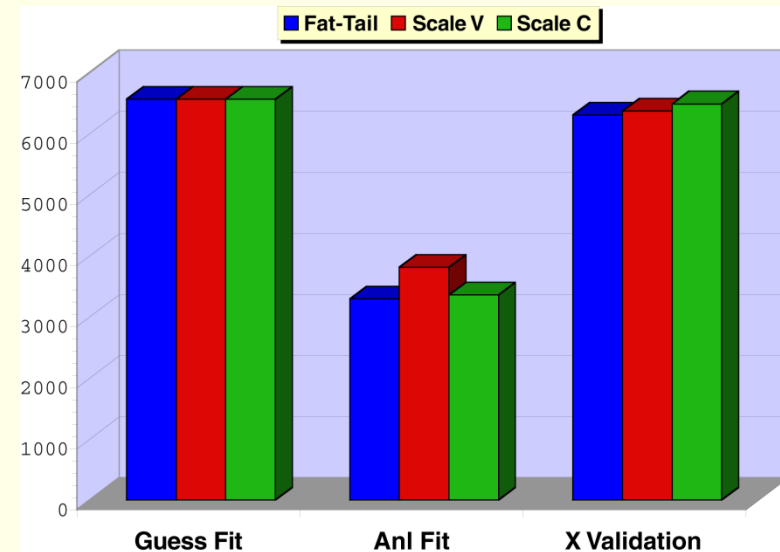
Fat-tailed Power Spectrum for B



Fat-tailed Power Spectrum

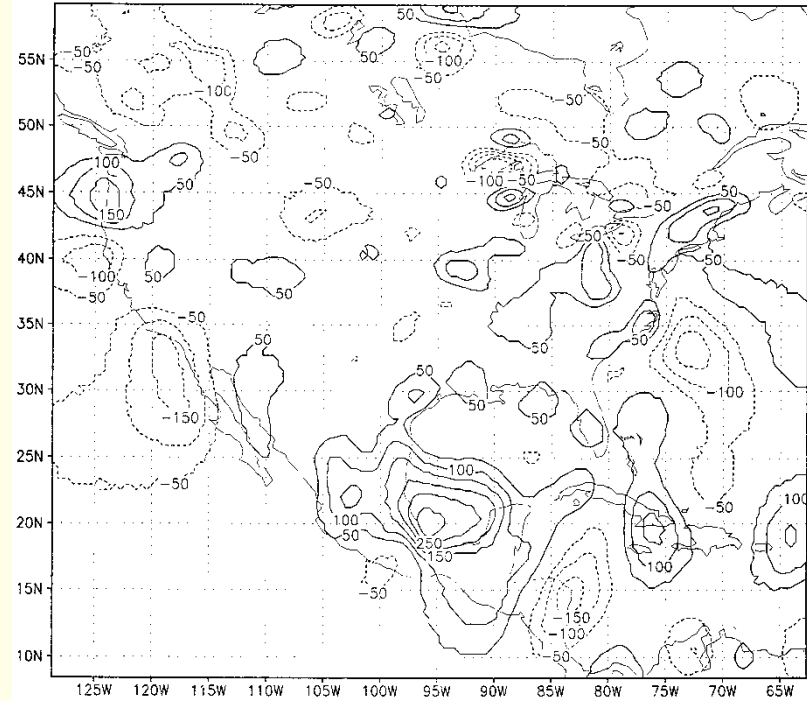
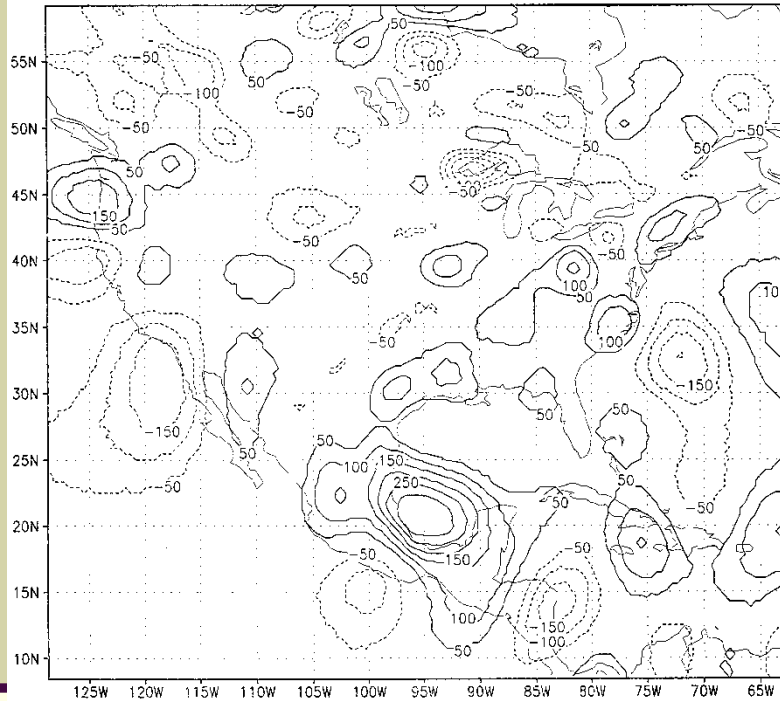


Psfc increments with single homogeneous recursive filter. (scale C)



Cross validation

Fat-tailed Power Spectrum



Psf increments with inhomogeneous scales with single recursive filter: Large scale (left) and multiple recursive filter: fat-tail (right)

Estimate Background Error

NMC method

time differences of forecasts (48-24hr)

Basic assumption: linear error growth with time

Ensemble method

ensemble differences of forecasts

Basic assumption: ensemble represents real spread

Conventional method

differences of forecasts and obs

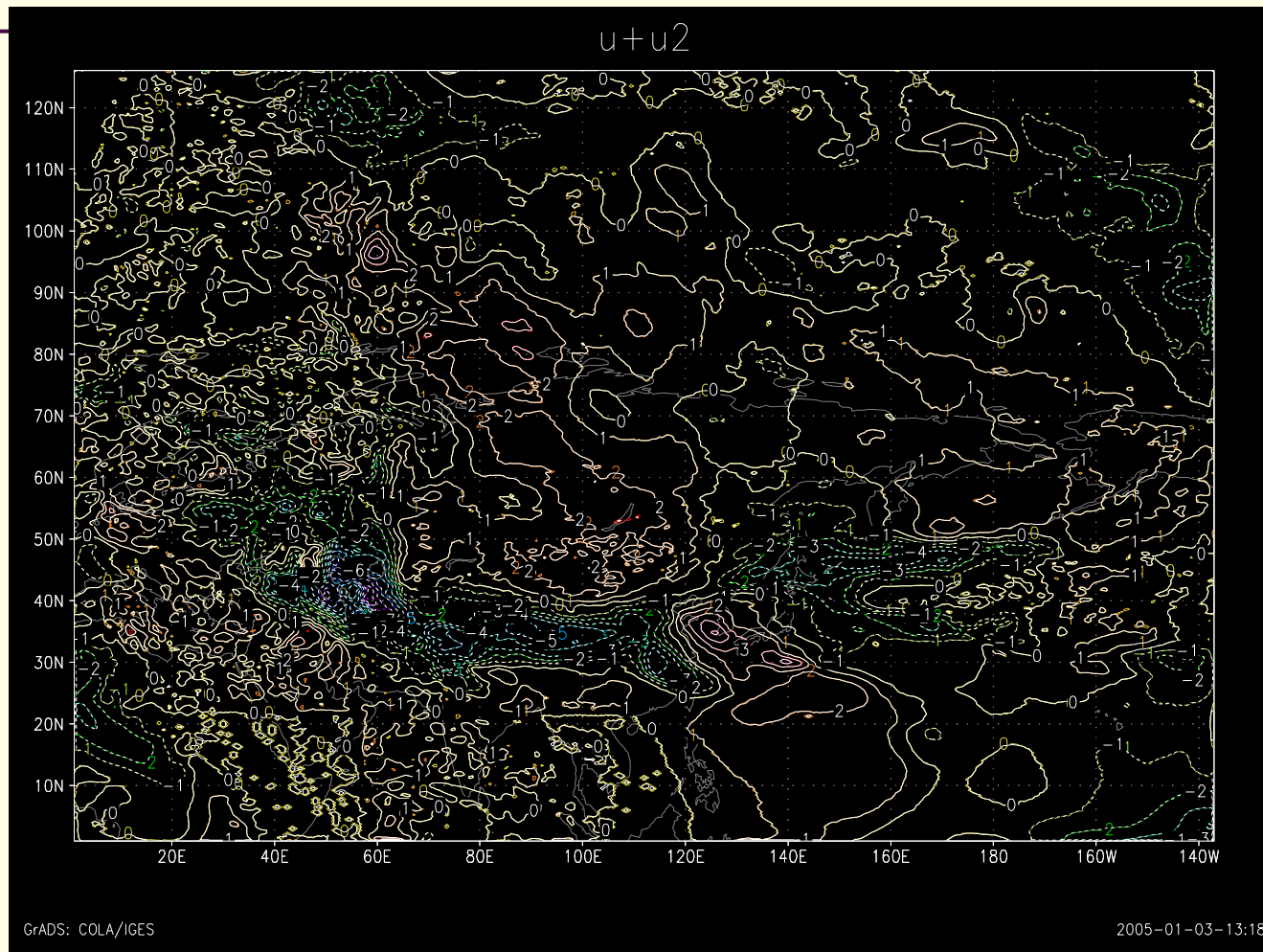
difficulties: obs coverage, multivariate...

Estimate B from Regional Forecasts

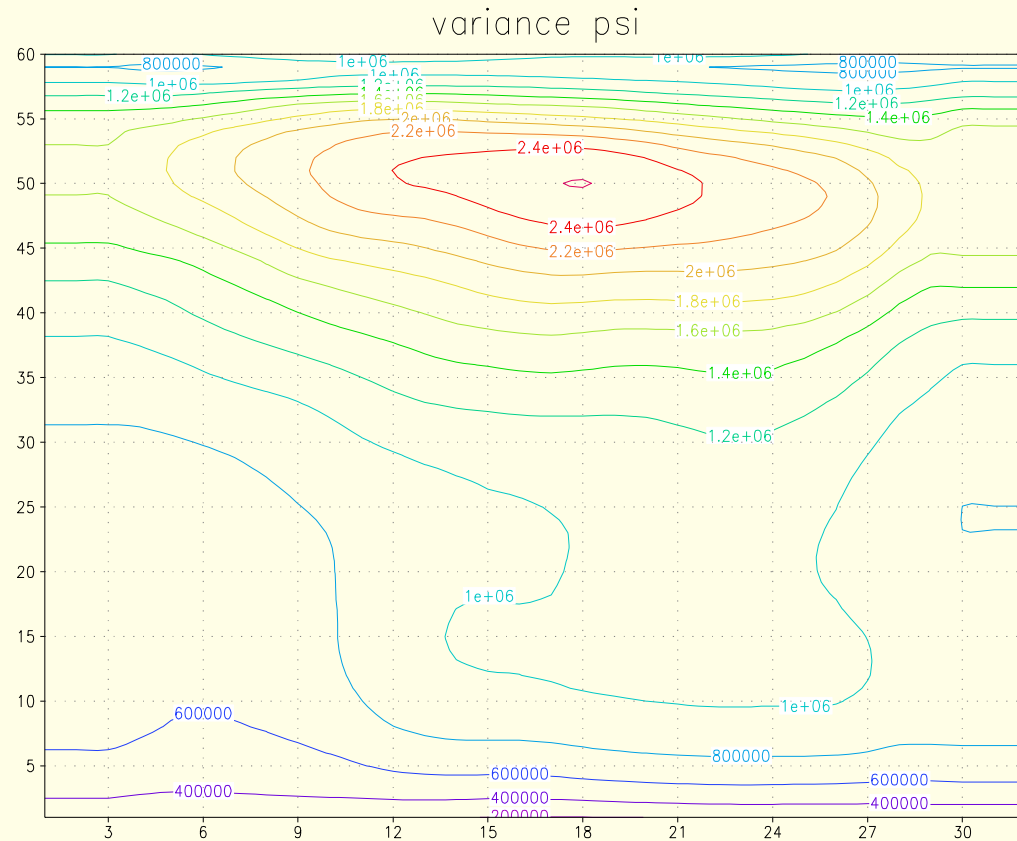
Spectral calculation of stream function and velocity potential from forecast differences of wind fields.

1. U & V : e 2 a grid
2. Fill to FFT grid number: taper & zero
3. FFT: both X & Y directions
4. Vor + Div
5. Del⁻²
6. FFT back to for Psi & Chi
7. Derivatives of Psi & Chi to find U2 & V2

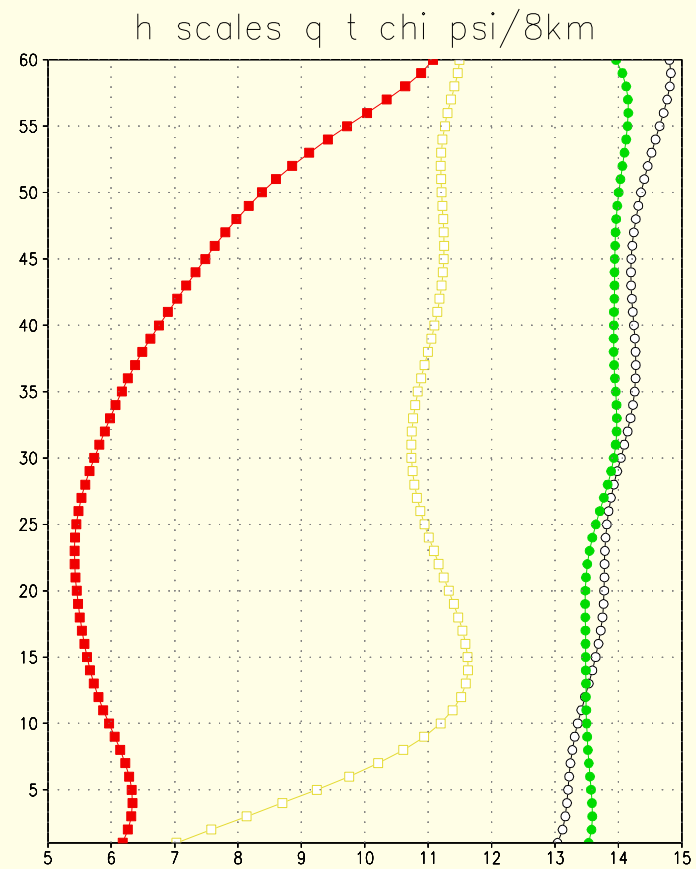
u & v from ψ & χ



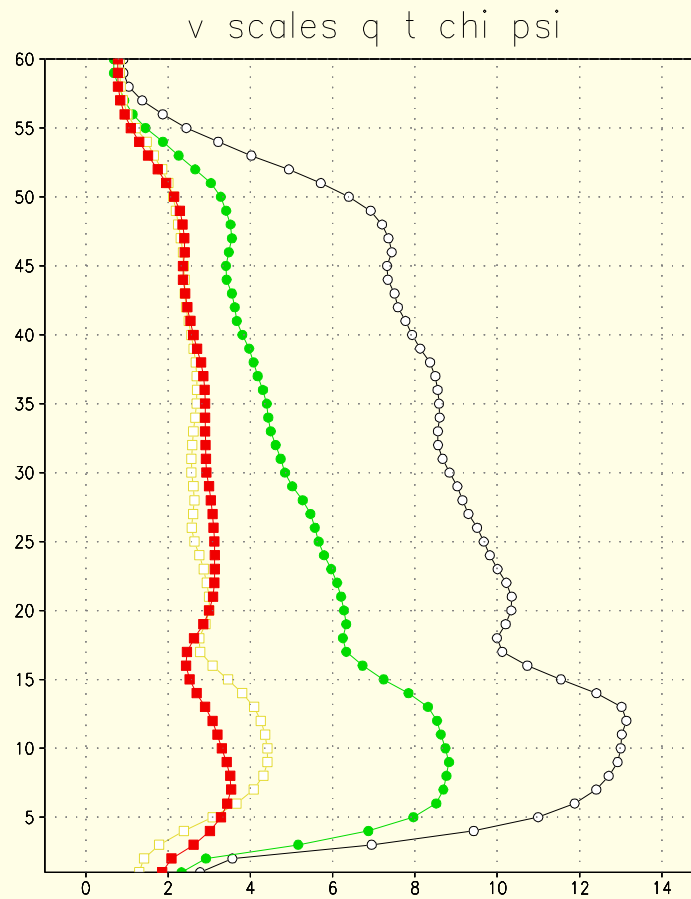
Estimate B from Regional Forecasts



Estimate B from Regional Forecasts



Estimate B from Regional Forecasts



Tuning Background Parameters

berror=\$FIXnam/nam_glb_berror.f77

&BKGERR

as=0.28,0.28,0.3,0.7,0.1,0.1,1.0,1.0,

hzscl=0.373,0.746,1.50,

vs=0.6,

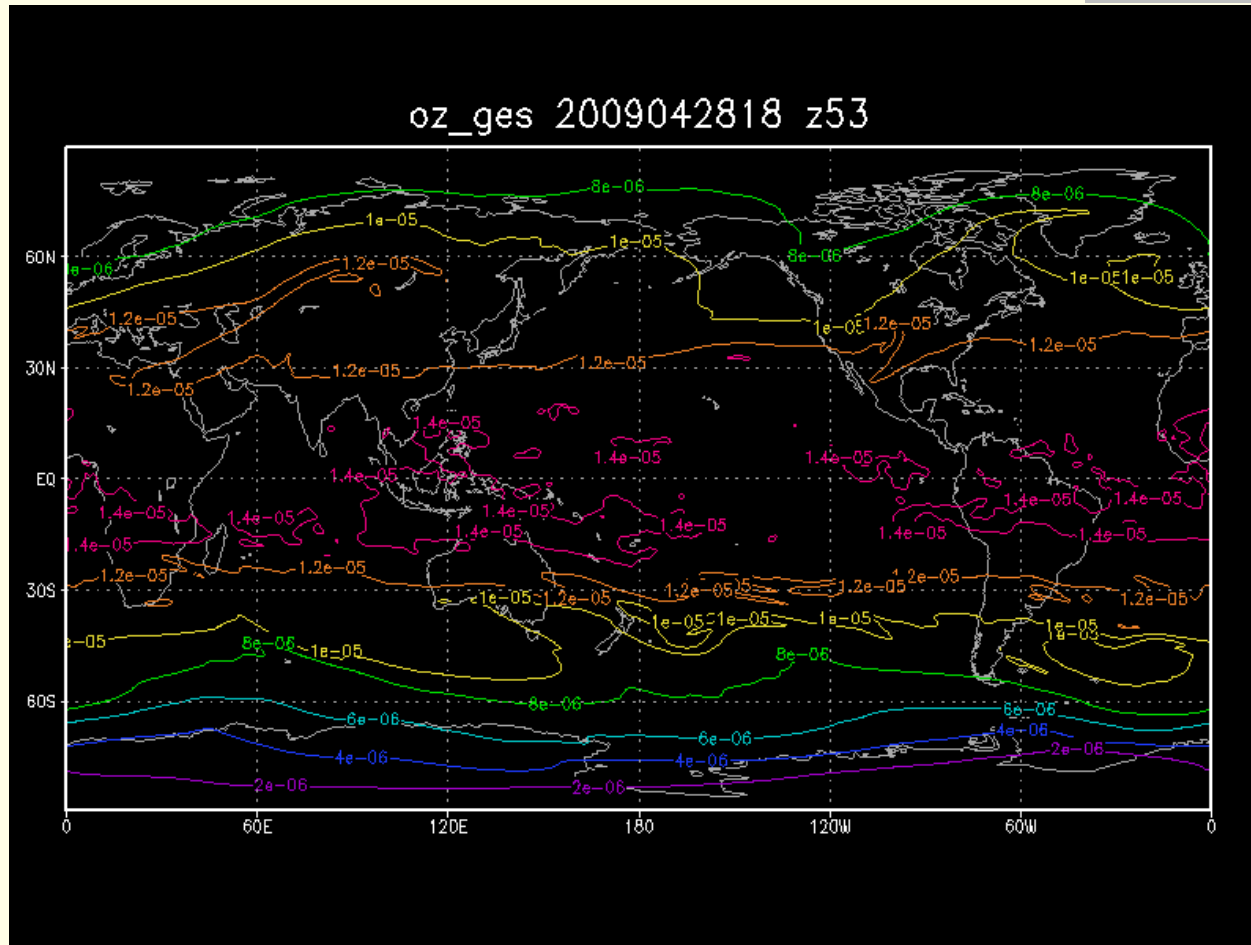
(Note that hzscl and vs apply to all variables)

Q: How to find out definitions of “as”?

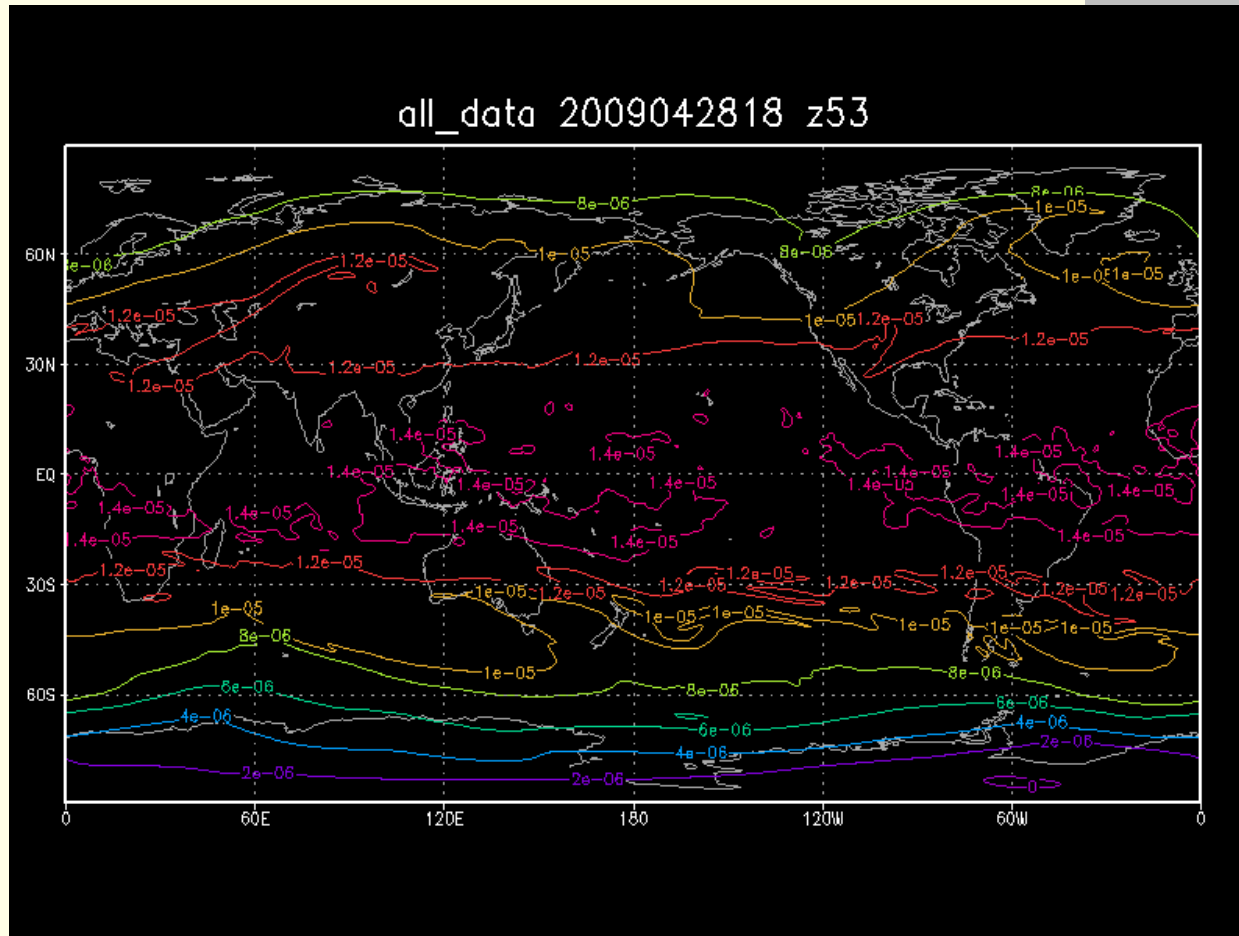
A: GSI code

(grep “as(4)” *90 to find in prewgt_reg.f90)

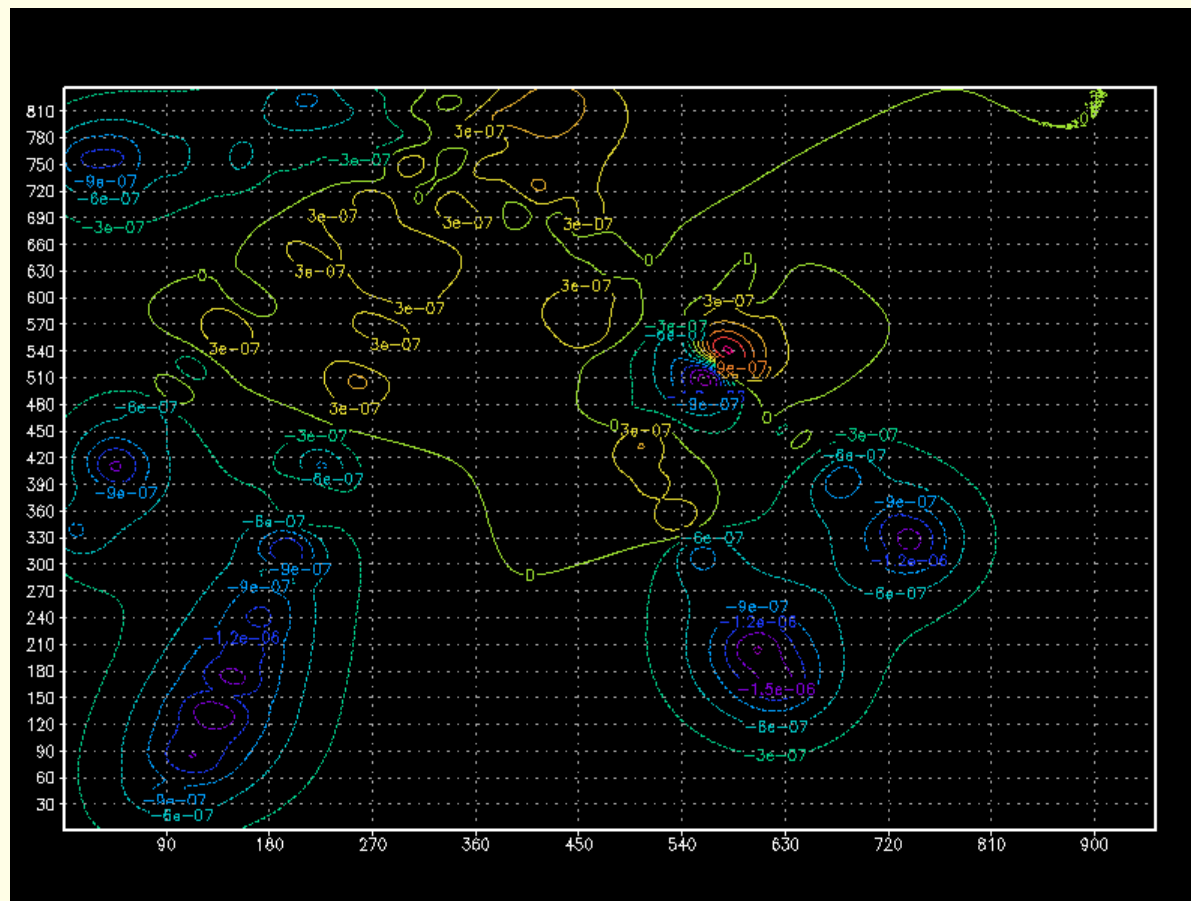
Impact of Background Error



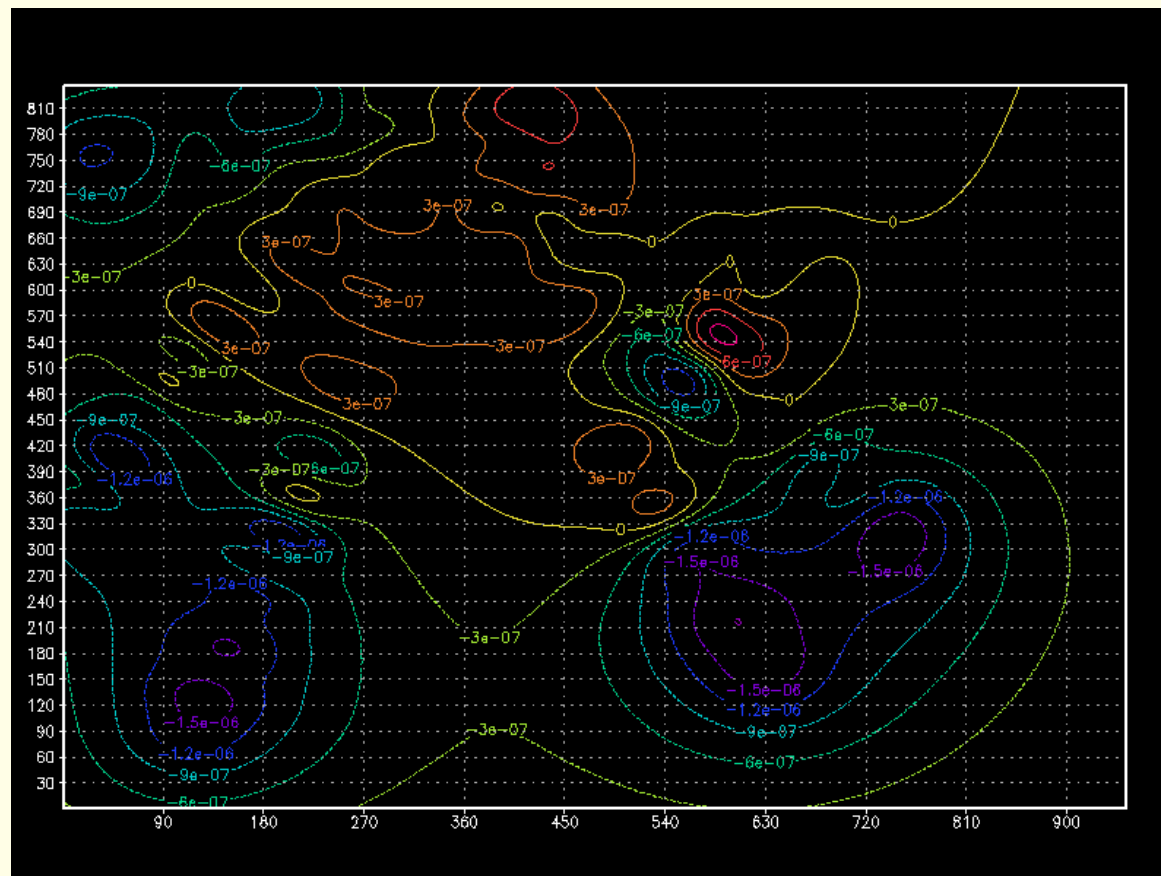
Impact of Background Error



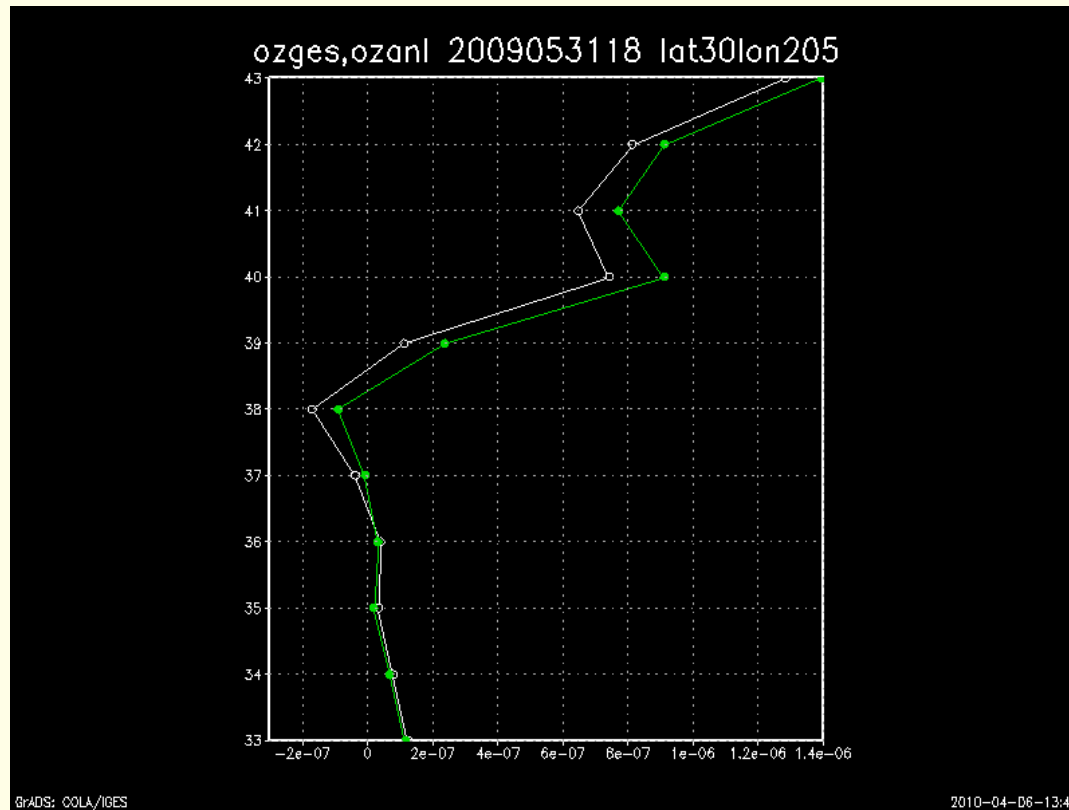
B adjustment ex1: Analysis increments of ozone



B adjustment ex2: Analysis increments of ozone

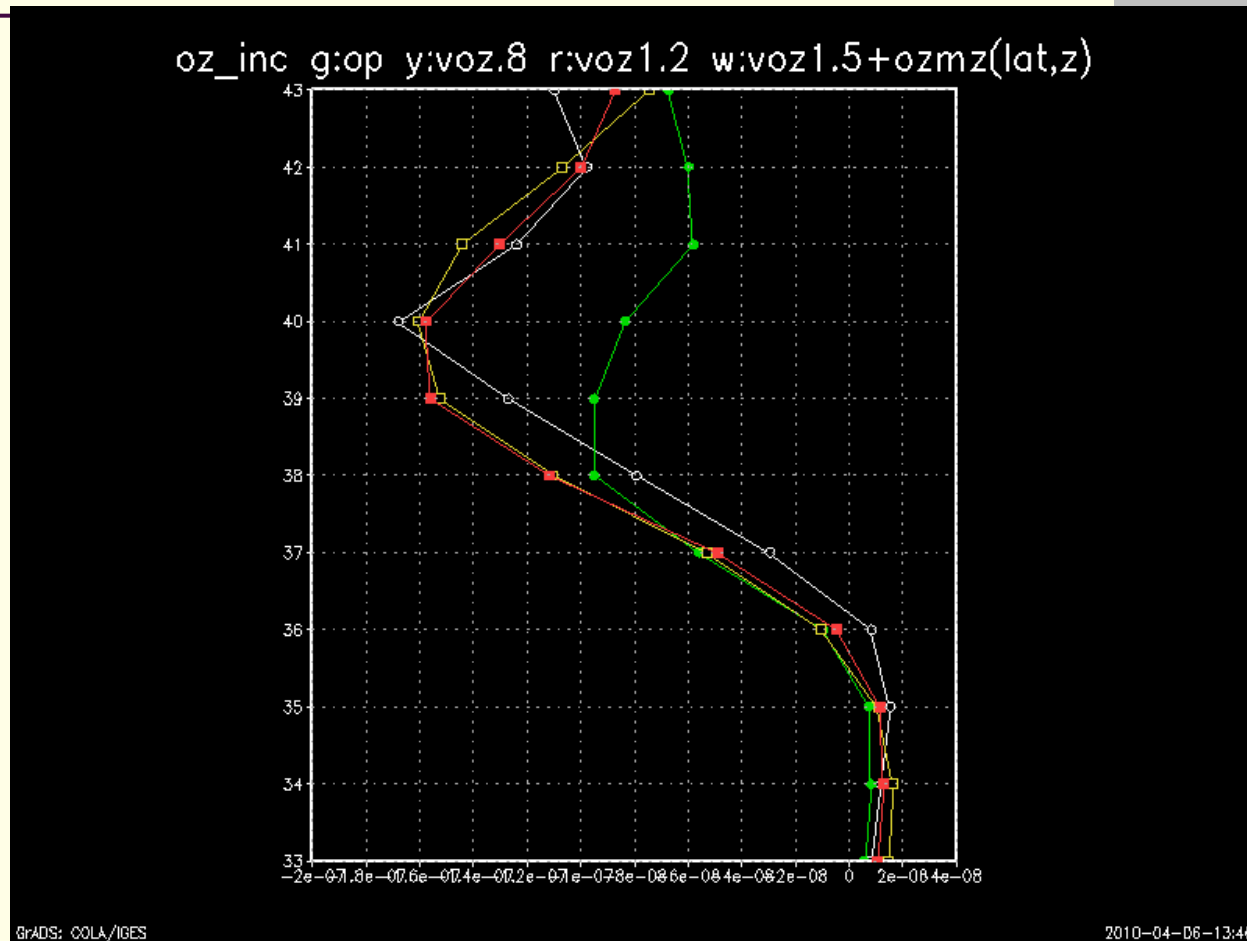


Impact of Vertical Scale

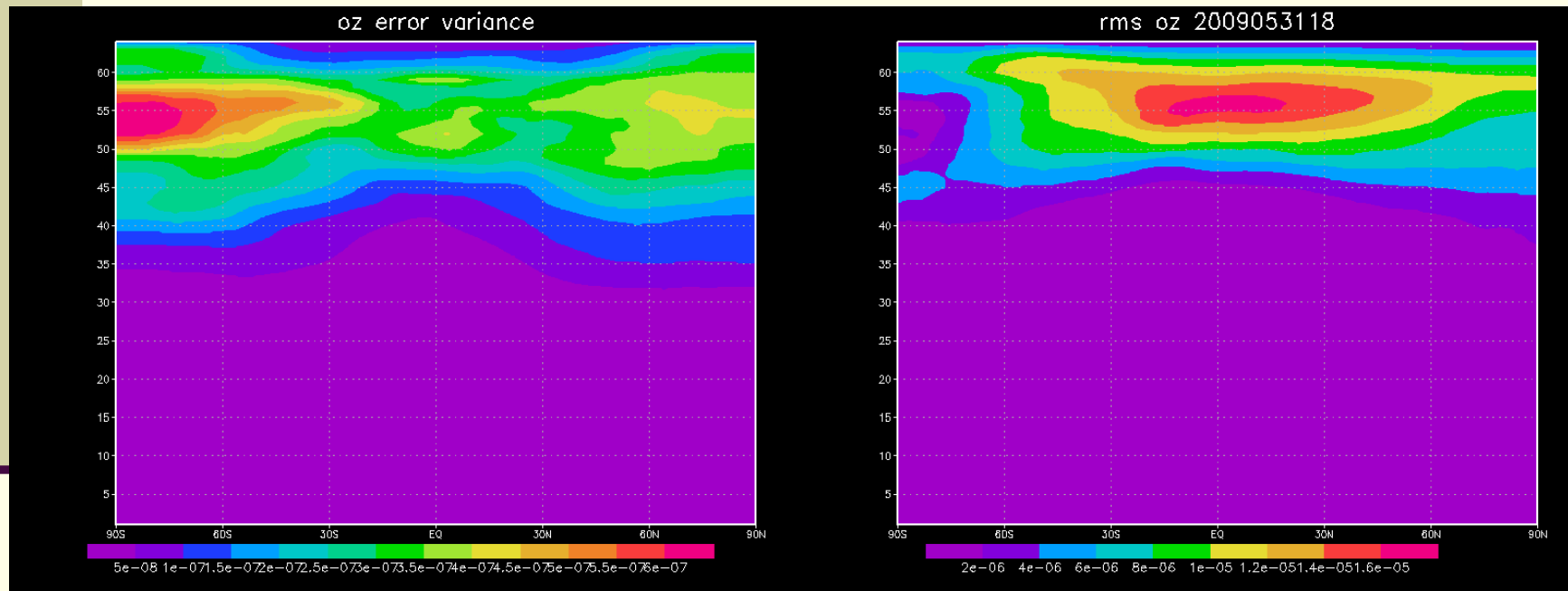


White: Analysis Green: First Guess

Impact of Vertical Scale



NMC method & Normalized Oz



Setup B for a new control variable

Uni-variate

Sparse observations

Poor first guess quality (large error variances)

With physical limit of non-negative value

Passive scalars

Ex: chemicals, aerosols, co₂, co,....

Observational Error

Conventional adjustments
Adaptive Tuning

Analysis system produces an analysis through the minimization of an objective function given by

$$J = \mathbf{x}^T \mathbf{B}^{-1} \mathbf{x} + (\mathbf{H} \mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{x} - \mathbf{y})$$

J^b

J^o

Where

\mathbf{x} is a vector of analysis increments,

\mathbf{B} is the background error covariance matrix,

\mathbf{y} is a vector of the observational residuals, $\mathbf{y} = \mathbf{y}_{\text{obs}} - \mathbf{H} \mathbf{x}_{\text{guess}}$

\mathbf{R} is **observational and representativeness** error covariance matrix

\mathbf{H} is the transformation operator from the analysis variable to the form of the observations.

Discussions

- No method is optimal; There' ll be issues to solve and subjective tuning, smoothing, and averaging.

Ex: 1) background error

Ex: 2) ob error tune

- In a semi-operational system, tune B variances so that the analysis penalties are about half of original penalties; including the scale effects.
- Conventional adjustment: Tune Oberror' s so that they are about the same as guess fit to data

Adaptive Tuning of Oberror

- Talagrand (1997) on $E (J (X^a))$

- Desroziers & Ivanov (2001)

$$E(J^o) = \frac{1}{2} \text{Tr} (I_p - HK)$$

$$E(J^b) = \frac{1}{2} \text{Tr} (KH)$$

where I_p is identity matrix with order p

K is Kalman gain matrix

H is linearized observation forward operator

- Chapnik et al.(2004): robust even when B is incorrectly specified

Adaptive Tuning of Oberror

Tuning Procedure

$$J(\delta X) = 1/s_b^2 J^b(\delta X) + 1/s_o^2 J^o(\delta X)$$

Where s_b and s_o are the background and oberr weighting parameters

$$S_o = \text{sqrt}(2J^o / \text{Tr}(I_p - HK))$$

Analysis system produces an analysis through the minimization of an objective function given by

$$J = \underbrace{\mathbf{x}^T \mathbf{B}^{-1} \mathbf{x}}_{J^b} + \underbrace{(\mathbf{H} \mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{x} - \mathbf{y})}_{J^o}$$

Where

- \mathbf{x} is a vector of analysis increments,
- \mathbf{B} is the background error covariance matrix,
- \mathbf{y} is a vector of the observational residuals, $\mathbf{y} = \mathbf{y}_{\text{obs}} - \mathbf{H} \mathbf{x}_{\text{guess}}$
- \mathbf{R} is **observational and representativeness** error covariance matrix
- \mathbf{H} is the transformation operator from the analysis variable to the form of the observations.

Randomized estimation of Tr (HK)

$$\text{Tr} (I_p - HK) = N_{\text{obs}} - \left(\sum \xi R^{-1/2} H \delta X^a(y+R^{1/2}\xi) + \sum \xi R^{-1/2} H \delta X^a(y) \right)$$

where ξ is random number with standard Gaussian distribution (mean: 0;variance:1)

2 outer iterations each produces an analysis;

output new error table

Consecutive jobs show the method converged

$$\text{Sum} = \sum (1-s_o)^2$$

S_0 for each ob type

- S_0 function of height (pressure)
rawinsonde, aircraft, aircar, profiler winds, dropsonde, satwind...
- S_0 constant with height
ship, synoptic, metar, bogus, ssm/I, ers speed, aircraft wind, aircar wind...
- The setup can be changed in penal.f90

Turn on adaptive tuning of Oberr

1) &SETUP

oberror_tune=.true.

2) If Global mode:

&OBSQC

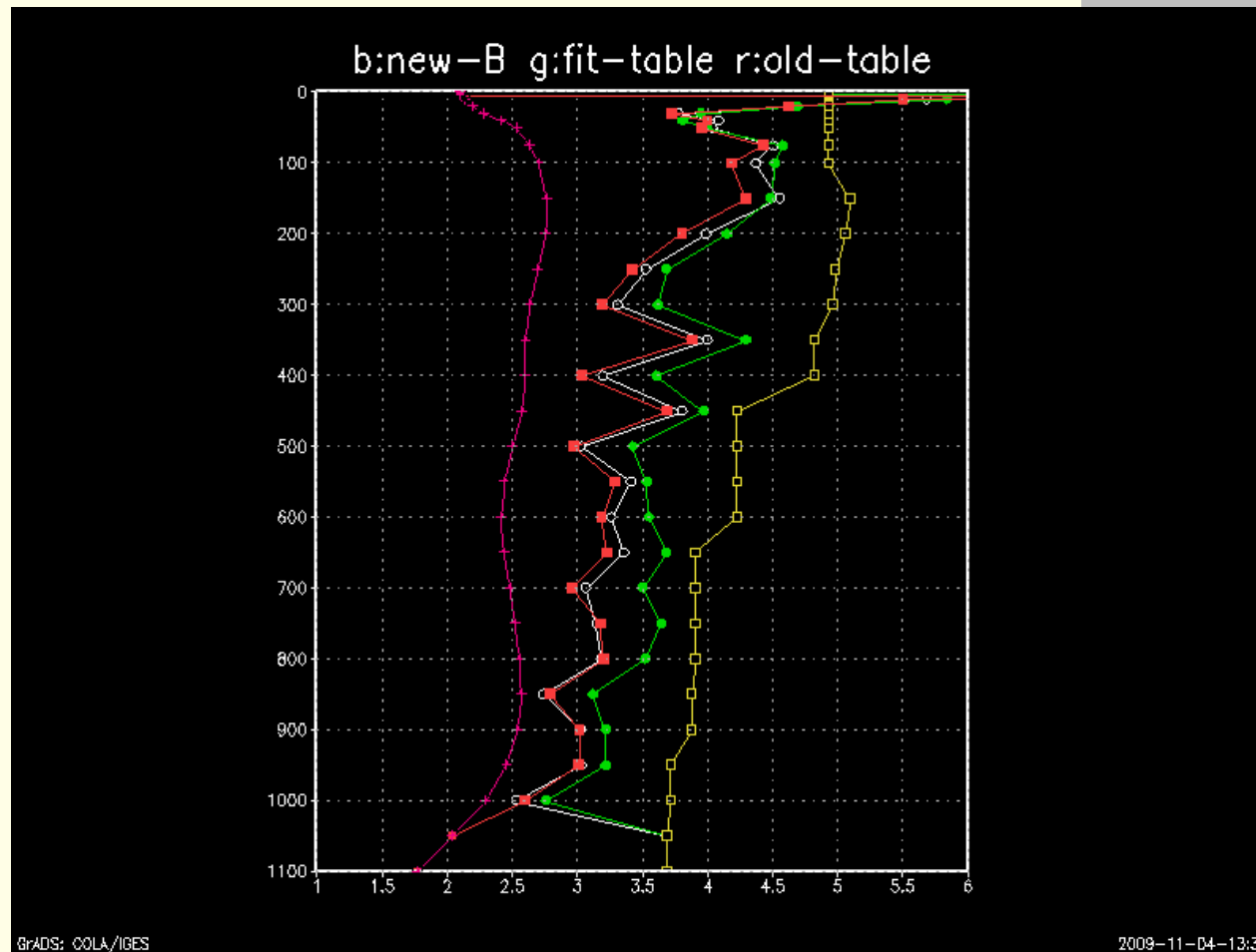
oberrflg=.true.

(Regional mode: oberrflg=.true. is default)

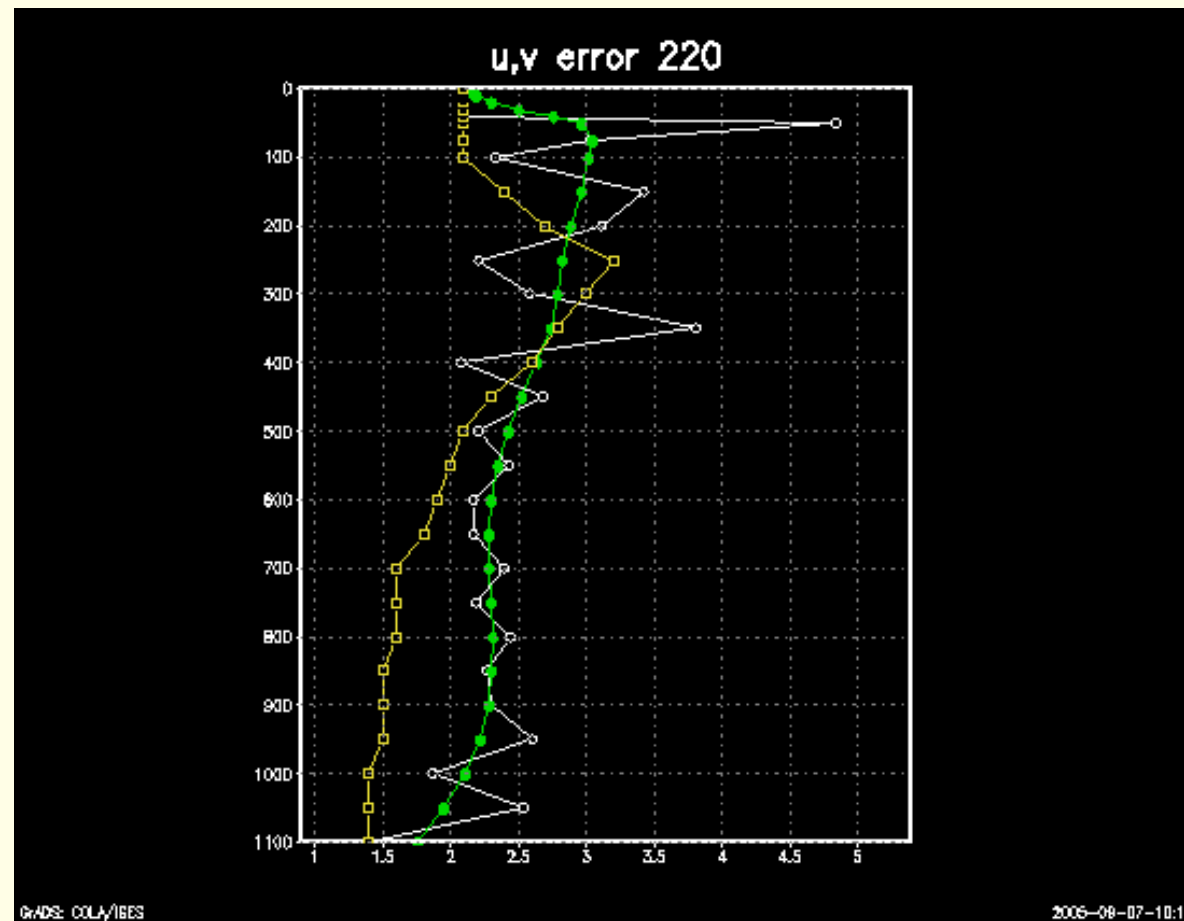
(find in file stdout: GSIMOD: ***WARNING*** reset oberrflg= T)

Note: GSI does not produce a valid analysis under the setup

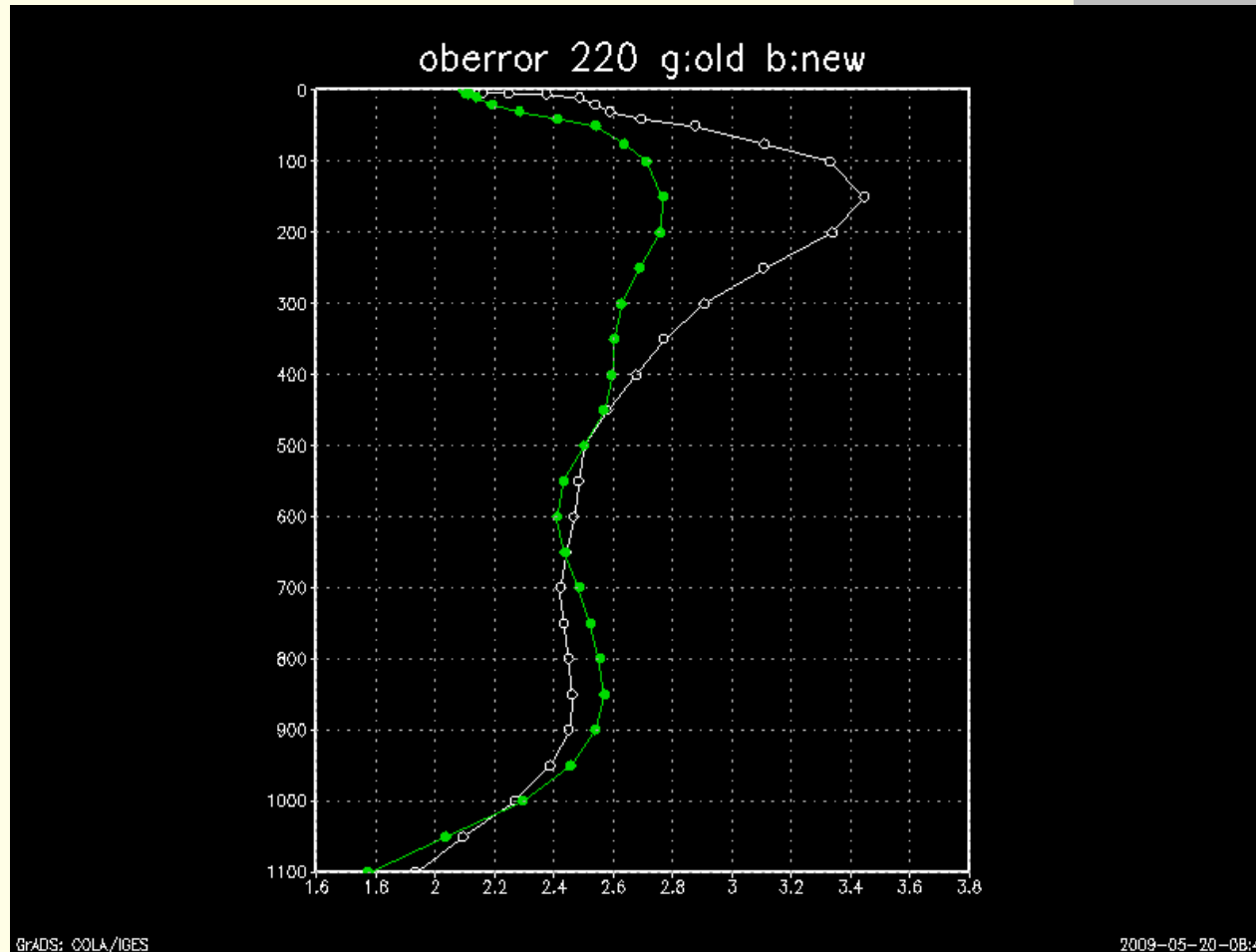
Sensitivity of adaptive tuning



Adjustment of tuned error table



Adaptive tuning of oberror of rawinsonde wind



Observational Error

Conventional adjustments
Adaptive Tuning