M.Sc. Course on Operational/Advanced Data Assimilation Techniques (MTMD02)

2011

Part I: Variational techniques (Ross Bannister)

(I) Introduction

- Inverse problems
- Filtering and smoothing
- Challenges of operational data assimilation
- History of data assimilation

(II) Variational techniques

- Euler-Lagrange equations
- A-priori information and the variational cost function
- Practical approximations to the cost function
- Error covariance matrices

(III) Modelling the B-matrix

- Univariate aspects
- Multivariate aspects

(IV) Measuring the B-matrix

(V) Variational algorithm, diagnostics and observation networks

Part II: Sequential techniques (Stefano Migliorini)

TBA

Further reading for part I

Bennett A.F., 2002, *Inverse Modeling of the Ocean and Atmosphere* (Euler-Lagrange equations and representers - sections 1.2, 1.3).

Daley R., 1991, *Atmospheric Data Analysis* (historical aspects and basic ideas - chapters 1, 13).

Kalnay E., 2003, *Atmospheric Modeling, Data Assimilation and Predictability* (basic aspects of data assimilation - chapter 5).

Lewis J.M., Lakshmivarahan S., Dhall S.K., 2006, Dynamic data assimilation: a Least Squares Approach (applications - chapters 3,4, data assimilation algorithms - chapter 19).

Schlatter T.W., 2000, *Variational Assimilation of Meteorological Observations in the Lower Atmosphere: a Tutorial on How it Works*, Journal of Atmospheric and Solar-Terrestrial Physics 62, pp. 1057-1070.

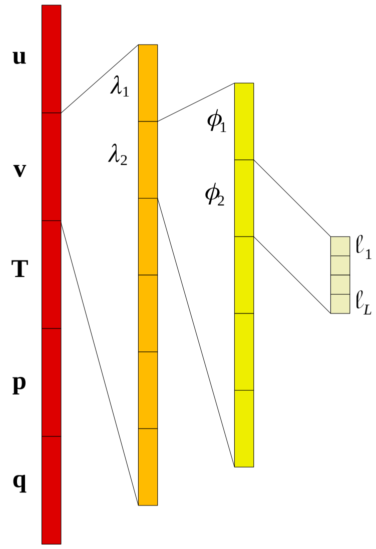
Further reading for part II

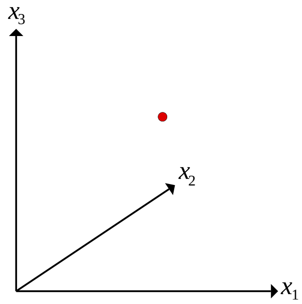
TBA

Example inverse problems

Field	Example problem to be solved
Weather forecasting	What is the global state of the atmosphere/ocean (e.g., u, v, T, p, q, cloud, SST, salinity)?
Atmospheric retrievals	What is the vertical profile of atmospheric quantities from remotely sensed observations?
Atmospheric pollution	What is the source/sink field of an atmospheric pollutant?
Parameter estimation	Determination of unknown model parameters.
Astronautics	Landing a spacecraft safely on another planet.
Astronomy	Orbit determination from observations.
Astrophysics	Determination of the internal structure of the Sun from surface observations.
Seismology	Determination of subterranean properties from seismic data (e.g. porosity, hydrocarbon content)?
Medical diagnosis	What is the 3-D structure of biological tissues from X-ray images (CAT scan)?

Data Assimilation Notation / Jargon





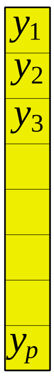
Model space, $\mathbf{x}$, $\mathbf{x}_A$, $\mathbf{x}_B$, $\delta\mathbf{x}$, etc. Left: λ, ϕ, l are longitude, latitude and level. There are n elements in total. Right: state space for $n = 3$.

$\mathbf{x}_A$ analysis state (posteriori state)

$\mathbf{x}_B$ background state (a-priori state)

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}(1) \\ \dots \\ \mathbf{x}(T) \end{pmatrix}$$ time augmented state

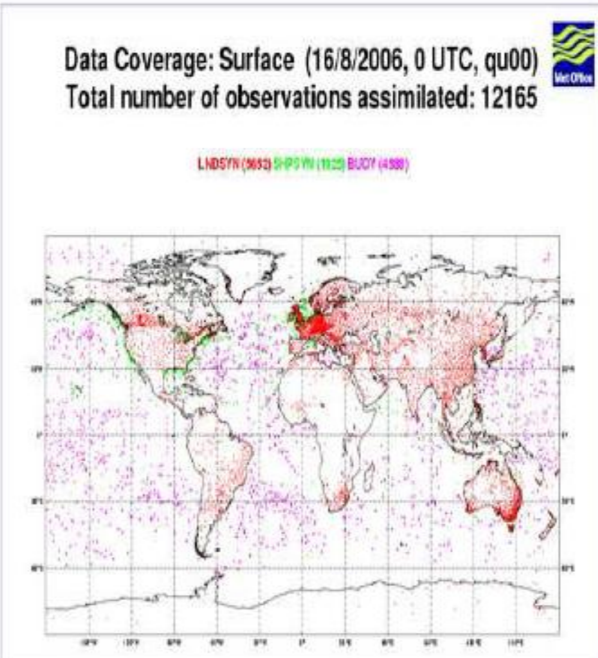
y

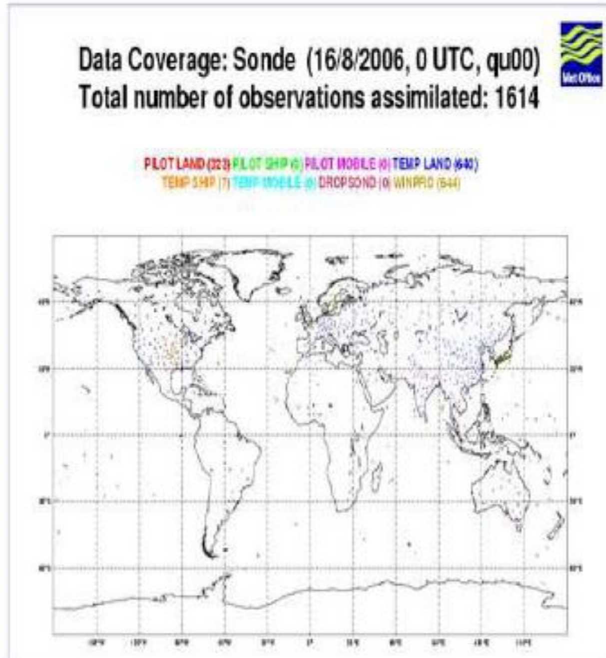


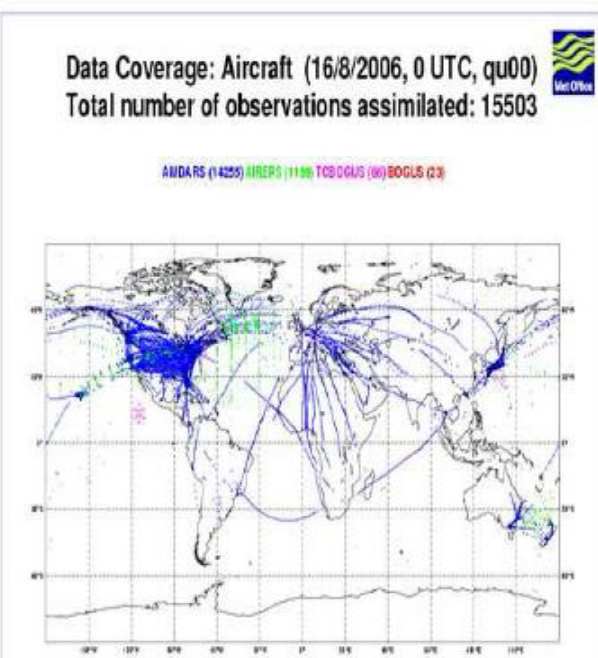
Observation space, **y**, observations may be of various types, at various positions and various times. There are p elements in total.

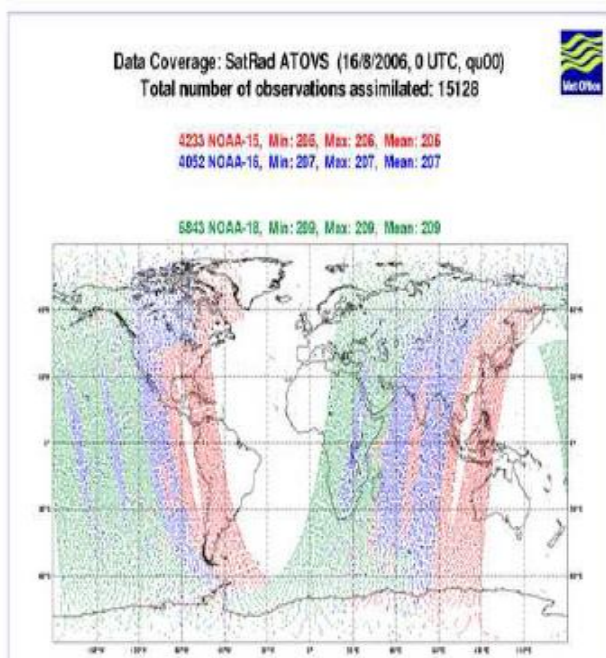
$$\mathbf{y} = \begin{pmatrix} \mathbf{y}(1) \\ \dots \\ \mathbf{y}(T) \end{pmatrix}$$ time augmented state

Example Observation Types

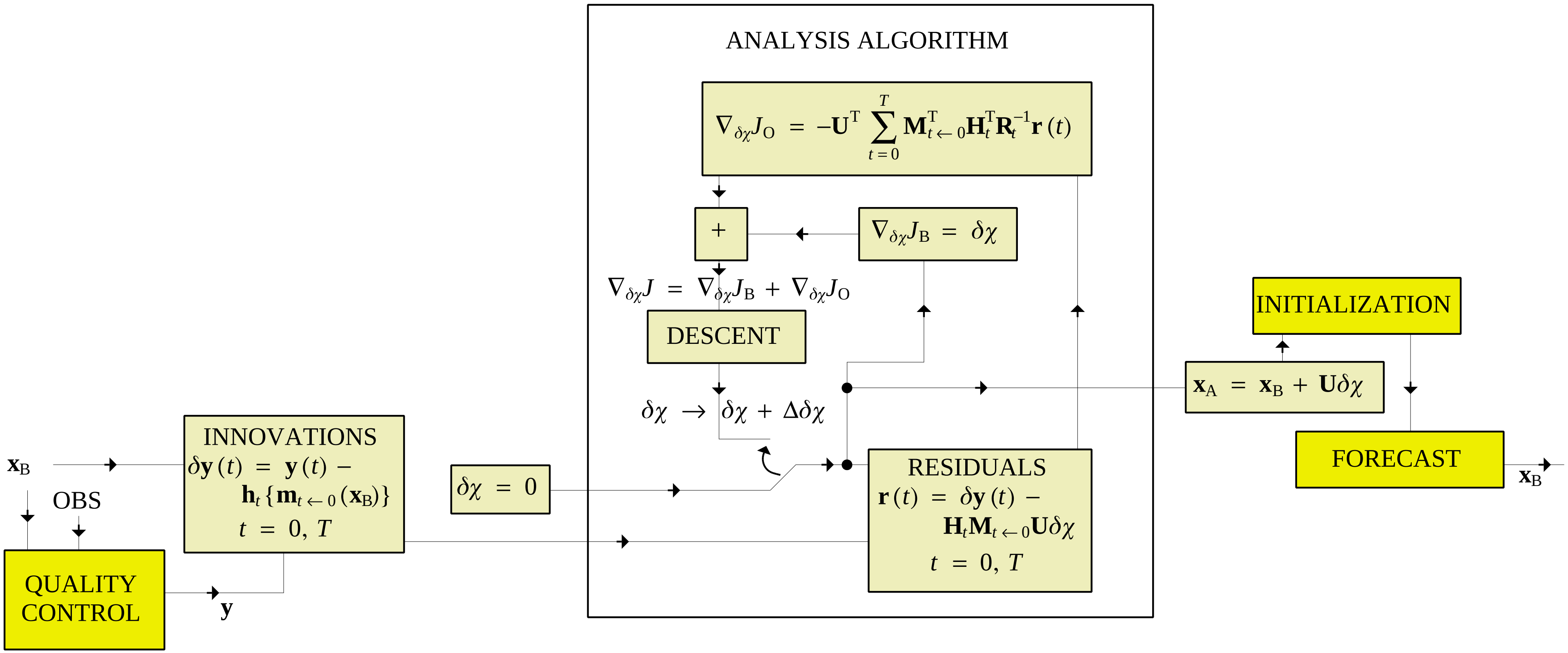








The Data Assimilation Cycle



DATA ASSIMILATION CYCLE FOR STRONG CONSTRAINT 4D-VAR

Summary of Methods of Operational Data Assimilation in NWP

	SUBJECTIVE ANALYSIS	POLYNOMIAL EXPANSION	CRESSMAN (SCM)	NUDGING	OI	3D-VAR	4D-VAR	EnKF
Objective/subjective	S	O	O	O	O	O	O	O
Account for obs accuracy?	✗	✗	✗✓	✗	✓	✓	✓	✓
Use prior knowledge?	✗	✗	✓	✓	✓	✓	✓	✓
Account for prior knowledge accuracy?	N/A	N/A	✗✓	✗	✓	✓	✓	✓
Ensure smooth fields?	✓	✓	✗	✓	✓	✓	✓	✓
Multivariate analysis ('balance')?	✗	✗	✗	✗✓	✗✓	✓	✓	✗✓
Allow global analysis in practice?	✗	✓	✓	✓	✗	✓	✓	✓
Allow indirect obs?	✗	✗	✗	✗	✗✓	✓	✓	✓
Allow non-linear obs operators easily?	N/A	N/A	N/A	N/A	✗✓	✓	✓	✓
Respect the exact obs time?	✗	✗	✗	✗	✓	✗	✓	✓
Respect physical laws?	✗	✗	✗	✓	✓	✓	✗	✗✓
Flow dependent prior error stats?	N/A	N/A	N/A	N/A	✗	✗✓	✗✓	✓
Allow non-Gaussian stats?	N/A	N/A	N/A	N/A	✗	✗	✗	✗
Sequential/non-sequential	S	S	S	NS	S	S	NS	S
Allow simultaneous parameter estimation?	✗	✗	✗	✗	✗✓	✓	✓	✓
Can allow for model error?	N/A	N/A	N/A	✗	✗	✗	✓	✓

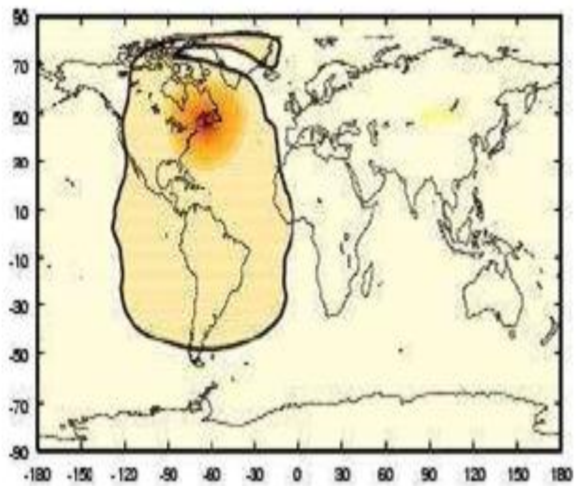
SCM = Successive Correction Method, OI = Optimal Interpolation, VAR = Variational assimilation, EnKF = Ensemble Kalman Filter.

Problems encountered in operational data assimilation

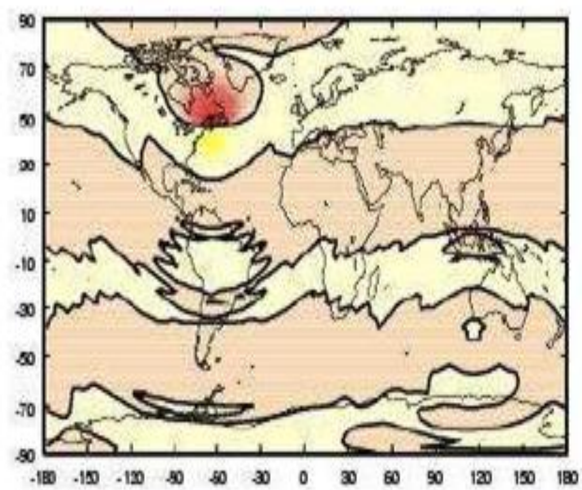
- Do not the computer power to deal with explicit matrices $O(n \times n)$.
- Do not have enough information to determine $\mathbf{P}_{\mathrm{f}}$ matrix.
- Need huge numbers of diverse observations promptly from all over the world and from space.
- A sensible analysis must be found even if observations 'go wrong' or are unavailable.
- Need to continually monitor for instrument problems (e.g. biases).
- Need to deal with complicated and potentially non-linear model and observation operators ($\mathbf{m}$, $\mathbf{h}$), their linearizations ($\mathbf{M}$, $\mathbf{H}$), and their adjoints ($\mathbf{M}^{\mathrm{T}}$, $\mathbf{H}^{\mathrm{T}}$).
- Usually don't know linearizations of operators as explicit matrices.
- Need to cope with numerical issues like bad conditioning of the minimization problem (see later).
- Sometimes need to consider imperfect operators.

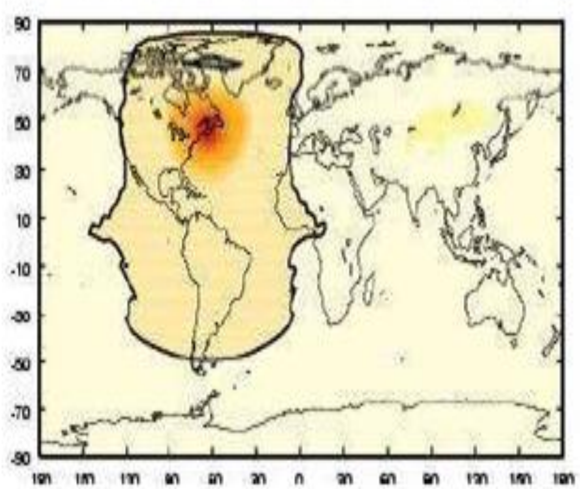
Example multivariate structure functions
(associated with pressure at the position over the E coast of N America)

Pressure part of structure function

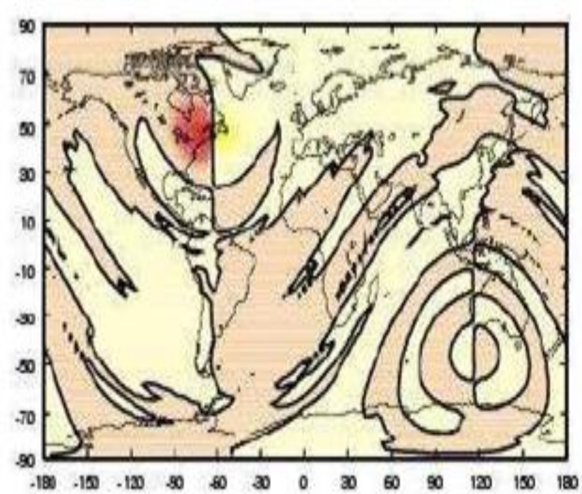


Zonal wind part of structure function





Temperature part of structure function



Meridional wind part of structure function

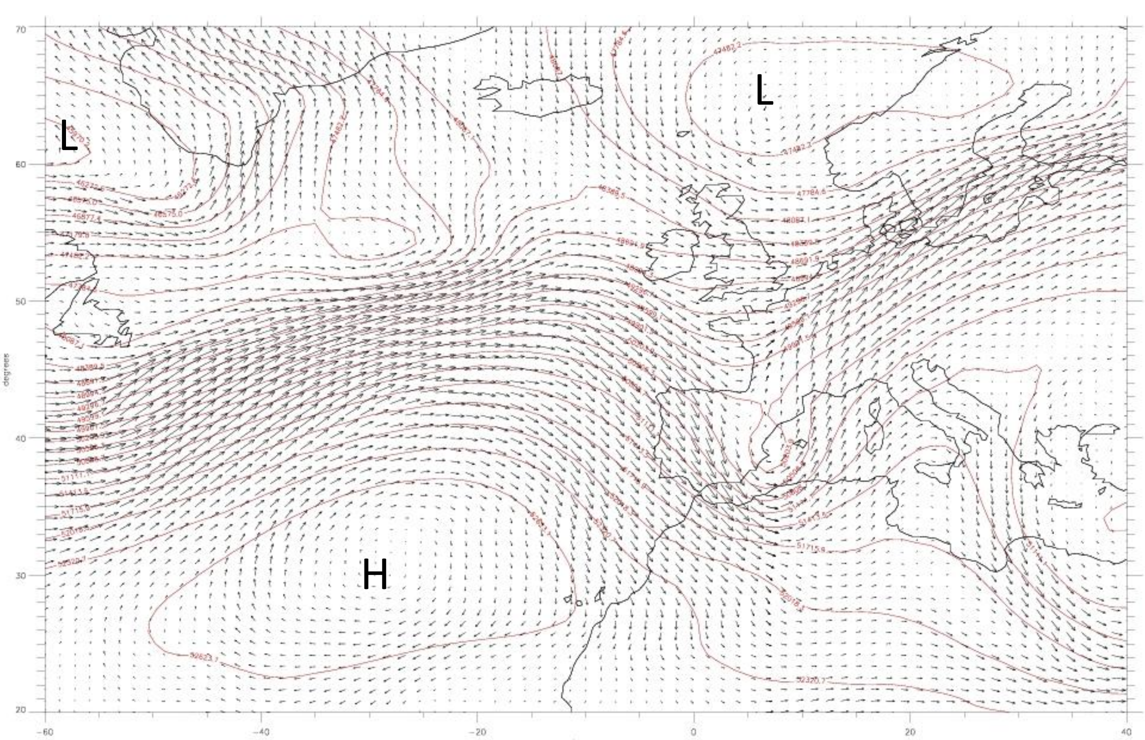
What is geostrophic balance?

Geostrophic balance is a diagnostic relationship between pressure and wind that results from dominant terms in the horizontal momentum equations (at mid latitudes).

$$\frac{\partial u}{\partial t} + \mathbf{u} \cdot \nabla u = \underbrace{fv - \frac{1}{\rho}\frac{\partial p}{\partial x}}, \qquad \frac{\partial v}{\partial t} + \mathbf{u} \cdot \nabla v = \underbrace{-fu - \frac{1}{\rho}\frac{\partial p}{\partial y}}, \qquad \mathbf{u} = \begin{pmatrix} u \\ v \end{pmatrix}.$$

Geostrophic balance:

$$v = \frac{1}{f\rho}\frac{\partial p}{\partial x}, \qquad u = -\frac{1}{f\rho}\frac{\partial p}{\partial y}.$$



Geostrophic covariances

How to construct an explicit background error covariance matrix that is subject to geostrophic balance between pressure and wind.

1. Assume pressure-pressure correlations are homogeneous and isotropic

$$\mu_{ij} = \exp - \frac{r_{ij}^2}{2L^2},$$

μ_{ij} is the correlation of pressure at positions i and j, $\sqrt{2}L$ is the horizontal length scale (approx. 750 km) and r_{ij} is the distance between positions i and j

$$r_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2 .$$

2. State approx. incremental form of geostrophic balance

$$\delta v = \frac{1}{f\rho}\frac{\partial \delta p}{\partial x}, \qquad \delta u = -\frac{1}{f\rho}\frac{\partial \delta p}{\partial y}.$$

3. Derive covariances for wind and pressure between positions *i* and *j*

Let σ_p be the (constant) background standard deviation for pressure error.

p-p covs: $\langle \delta p_i \delta p_j \rangle = \sigma^2 \mu_{ij}.$

p-u covs: $\langle \delta p_i \delta u_j \rangle = -\frac{1}{f\rho}\langle \delta p_i \frac{\partial}{\partial y_j}\delta p_j \rangle = -\frac{1}{f\rho}\frac{\partial}{\partial y_j}\langle \delta p_i \delta p_j \rangle = -\frac{\sigma_p^2}{f\rho}\frac{\partial \mu_{ij}}{\partial y_j}.$

p-v covs: $\langle \delta p_i \delta v_j \rangle = \frac{1}{f\rho}\langle \delta p_i \frac{\partial}{\partial x_j}\delta p_j \rangle = \frac{1}{f\rho}\frac{\partial}{\partial x_j}\langle \delta p_i \delta p_j \rangle = \frac{\sigma_p^2}{f\rho}\frac{\partial \mu_{ij}}{\partial x_j}.$

u-p covs: $\langle \delta u_i \delta p_j \rangle = -\frac{1}{f\rho}\langle \frac{\partial}{\partial y_i}\delta p_i \delta p_j \rangle = -\frac{1}{f\rho}\frac{\partial}{\partial y_i}\langle \delta p_i \delta p_j \rangle = -\frac{\sigma_p^2}{f\rho}\frac{\partial \mu_{ij}}{\partial y_i}.$

u-u covs: $\langle \delta u_i \delta u_j \rangle = \frac{1}{f^2\rho^2}\langle \frac{\partial}{\partial y_i}\delta p_i \frac{\partial}{\partial y_j}\delta p_j \rangle = \frac{1}{f^2\rho^2}\frac{\partial^2}{\partial y_i \partial y_j}\langle \delta p_i \delta p_j \rangle = \frac{\sigma_p^2}{f^2\rho^2}\frac{\partial^2 \mu_{ij}}{\partial y_i \partial y_j}.$

u-v covs: $\langle \delta u_i \delta v_j \rangle = -\frac{1}{f^2\rho^2}\langle \frac{\partial}{\partial y_i}\delta p_i \frac{\partial}{\partial x_j}\delta p_j \rangle = -\frac{1}{f^2\rho^2}\frac{\partial^2}{\partial y_i \partial x_j}\langle \delta p_i \delta p_j \rangle = -\frac{\sigma_p^2}{f^2\rho^2}\frac{\partial^2 \mu_{ij}}{\partial y_i \partial x_j}.$

v-p covs: $\langle \delta v_i \delta p_j \rangle = \frac{1}{f\rho}\langle \frac{\partial}{\partial x_i}\delta p_i \delta p_j \rangle = \frac{1}{f\rho}\frac{\partial}{\partial x_i}\langle \delta p_i \delta p_j \rangle = \frac{\sigma_p^2}{f\rho}\frac{\partial \mu_{ij}}{\partial x_i}.$

v-u covs: $\langle \delta v_i \delta u_j \rangle = -\frac{1}{f^2\rho^2}\langle \frac{\partial}{\partial x_i}\delta p_i \frac{\partial}{\partial y_j}\delta p_j \rangle = -\frac{1}{f^2\rho^2}\frac{\partial^2}{\partial x_i \partial y_j}\langle \delta p_i \delta p_j \rangle = -\frac{\sigma_p^2}{f^2\rho^2}\frac{\partial^2 \mu_{ij}}{\partial x_i \partial y_j}.$

v-v covs: $\langle \delta v_i \delta v_j \rangle = \frac{1}{f^2\rho^2}\langle \frac{\partial}{\partial x_i}\delta p_i \frac{\partial}{\partial x_j}\delta p_j \rangle = \frac{1}{f^2\rho^2}\frac{\partial^2}{\partial x_i \partial x_j}\langle \delta p_i \delta p_j \rangle = \frac{\sigma_p^2}{f^2\rho^2}\frac{\partial^2 \mu_{ij}}{\partial x_i \partial x_j}.$

4. Find first and second derivatives of p-p correlations

$$\frac{\partial \mu_{ij}}{\partial x_i} = -\frac{\mu_{ij}(x_i - x_j)}{L^2}, \qquad \frac{\partial \mu_{ij}}{\partial x_j} = \frac{\mu_{ij}(x_i - x_j)}{L^2},$$

$$\frac{\partial \mu_{ij}}{\partial y_i} = -\frac{\mu_{ij}(y_i - y_j)}{L^2}, \qquad \frac{\partial \mu_{ij}}{\partial y_j} = \frac{\mu_{ij}(y_i - y_j)}{L^2},$$

$$\frac{\partial^2 \mu_{ij}}{\partial x_i \partial x_j} = \frac{\mu_{ij}}{L^2}\left(1 - \frac{(x_i - x_j)^2}{L^2}\right), \qquad \frac{\partial^2 \mu_{ij}}{\partial y_i \partial y_j} = \frac{\mu_{ij}}{L^2}\left(1 - \frac{(y_i - y_j)^2}{L^2}\right),$$

$$\frac{\partial^2 \mu_{ij}}{\partial y_i \partial x_j} = -\frac{\mu_{ij}(x_i - x_j)(y_i - y_j)}{L^4}, \qquad \frac{\partial^2 \mu_{ij}}{\partial x_i \partial y_j} = -\frac{\mu_{ij}(x_i - x_j)(y_i - y_j)}{L^4}.$$

5. Example structure functions (red +, blue -)

The following structure functions give the output field (p, u, or v down the side) associated with a point in the centre of the domain (either of p, u, or v along the top). Red is positive and blue is negative.