École Doctorale des Sciences de l'Environnement d'Île-de-France
Année 2007-2008

Modélisation Numérique de l'Écoulement Atmosphérique et Assimilation d'Observations

Olivier Talagrand

Cours 5

3 Juin 2008

Optimal Interpolation

Random field $\varphi(\xi)$

Observation network $\xi_1, \xi_2, \ldots, \xi_p$

For one particular realization of the field, observations

$$y_j = \varphi(\xi_j) + \varepsilon_j \quad , \; j = 1, \ldots, p \qquad , \qquad \text{making up vector } \boldsymbol{y} = (y_j)$$

Estimate $x = \varphi(\xi)$ at given point ξ, in the form

$$x^a = \alpha + \Sigma_j \, \beta_j \, y_j = \alpha + \boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{y} \qquad , \qquad \text{where } \boldsymbol{\beta} = (\beta_j)$$

α and the β_j's being determined so as to minimize the expected quadratic estimation error $E[(x\text{-}x^a)^2]$

Optimal Interpolation (continued 1)

Solution

$$x^a = E(x) + E(x'\mathbf{y}'^T)\,[E(\mathbf{y}'\mathbf{y}'^T)]^{-1}\,[\mathbf{y} - E(\mathbf{y})]$$

i. e.,

$$\boldsymbol{\beta} = [E(\mathbf{y}'\mathbf{y}'^T)]^{-1}\,E(x'\mathbf{y}')$$

$$\alpha = E(x) - \boldsymbol{\beta}^T E(\mathbf{y})$$

Estimate is unbiased $E(x\text{-}x^a) = 0$

Minimized quadratic estimation error

$$E[(x\text{-}x^a)^2] = E(x'^2) - E(x'\mathbf{y}'^T)\,[E(\mathbf{y}'\mathbf{y}'^T)]^{-1}\,E(\mathbf{y}'x')$$

Estimation made in terms of deviations from expectations x' and $\mathbf{y}'$.

Optimal Interpolation (continued 2)

$$x^a = E(x) + E(x'\mathbf{y}'^{\mathrm{T}})\,[E(\mathbf{y}'\mathbf{y}'^{\mathrm{T}})]^{-1}\,[\mathbf{y} - E(\mathbf{y})]$$

$$y_j = \varphi(\xi_j) + \varepsilon_j$$

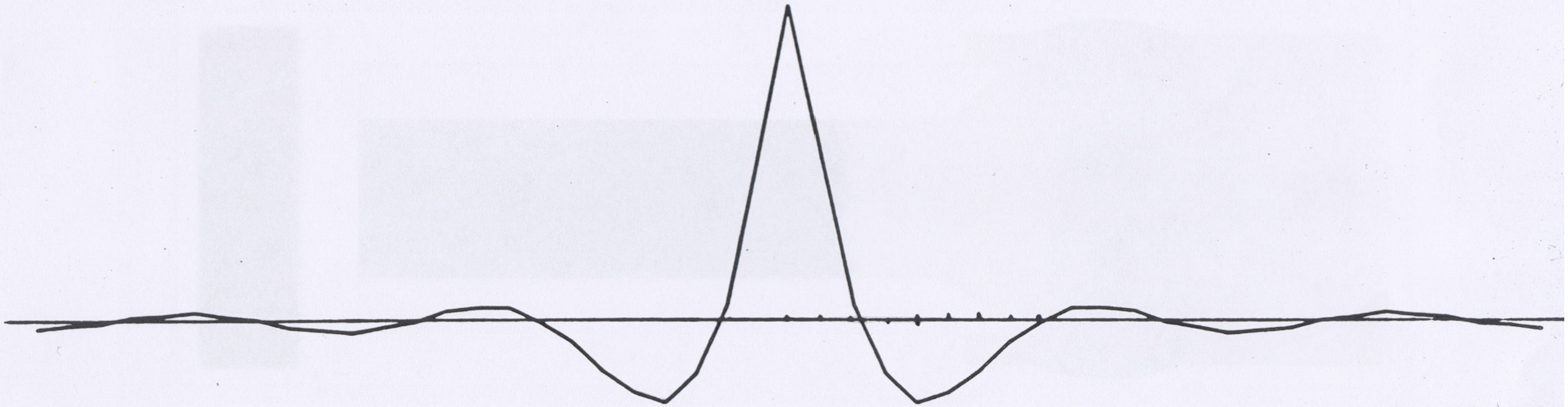
$$E(y_j'y_k') = E[\varphi'(\xi_j) + \varepsilon_j'][\varphi'(\xi_k) + \varepsilon_k']$$

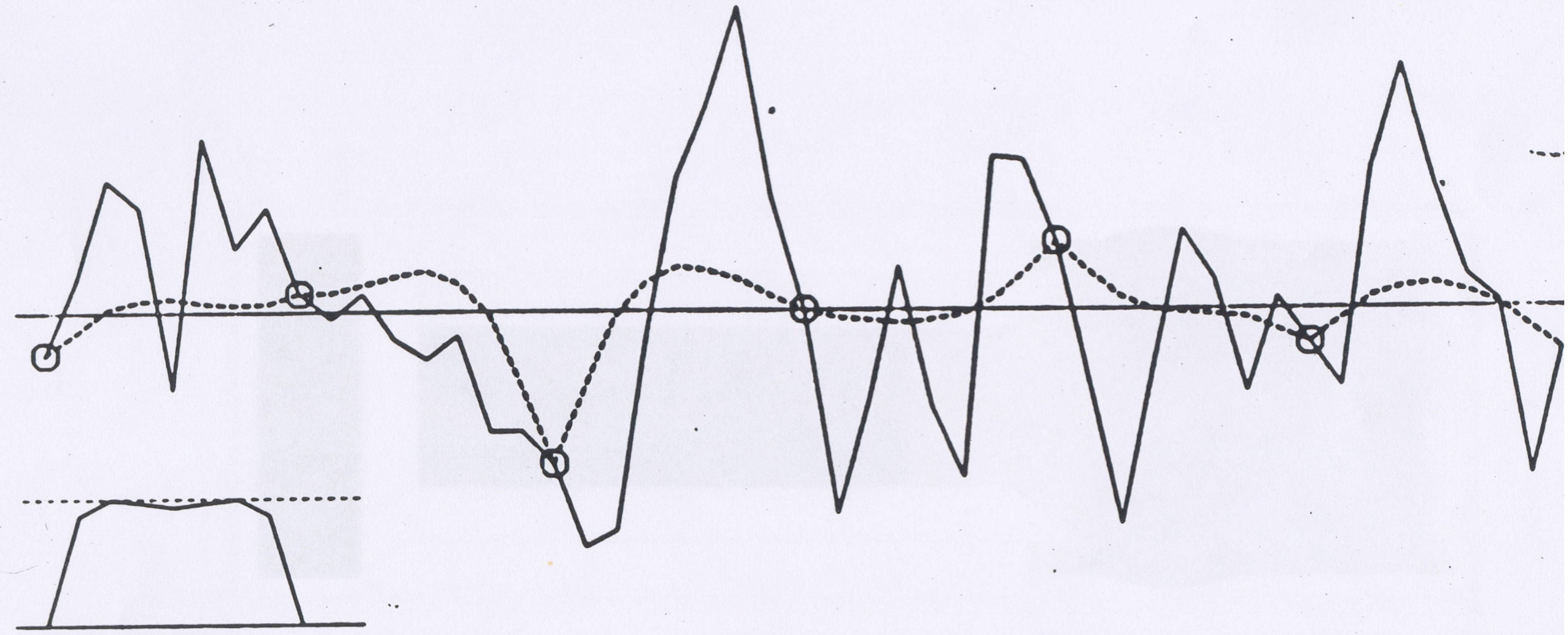
If observation errors ε_j are mutually uncorrelated, have common variance s, and are uncorrelated with field φ, then

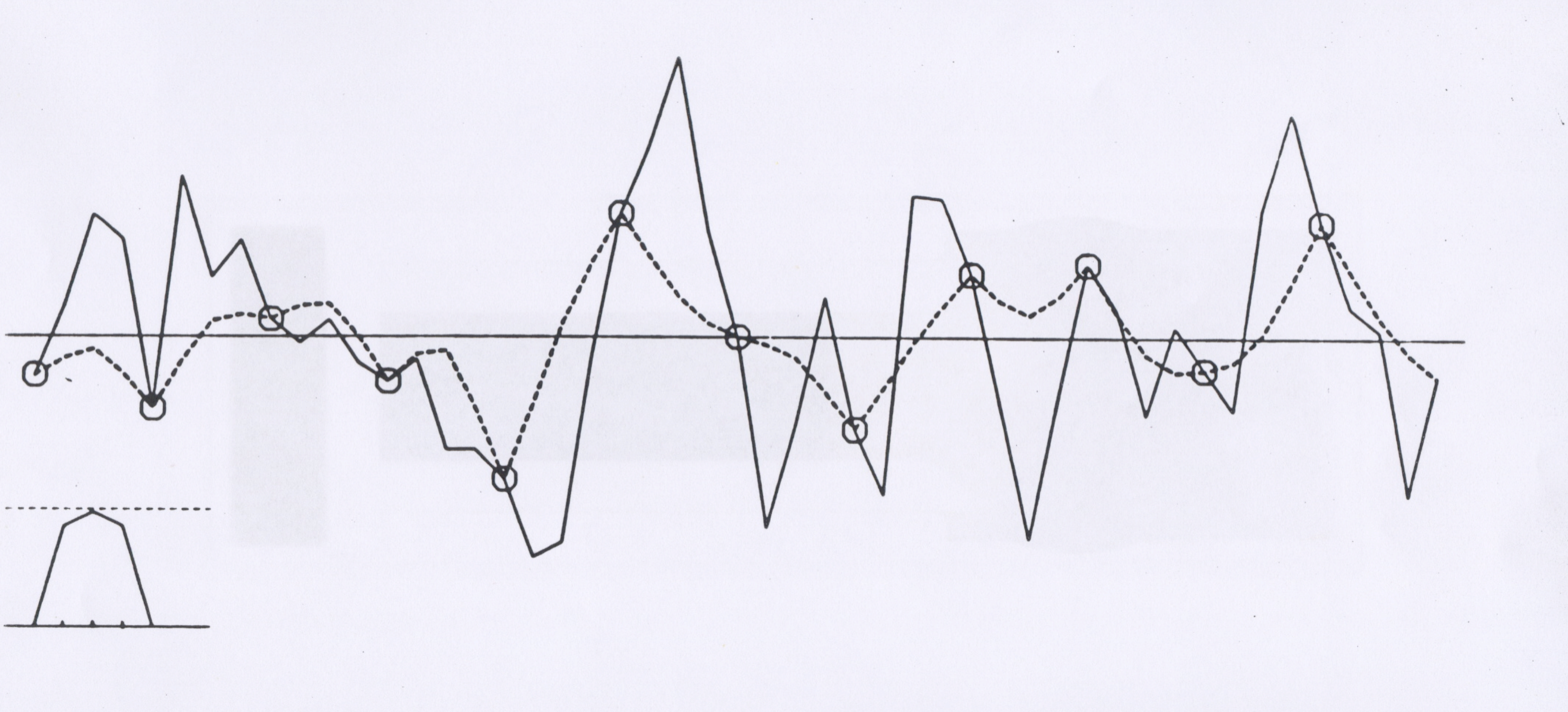
$$E(y_j'y_k') = C_\varphi(\xi_j, \xi_k) + s\delta_{jk}$$

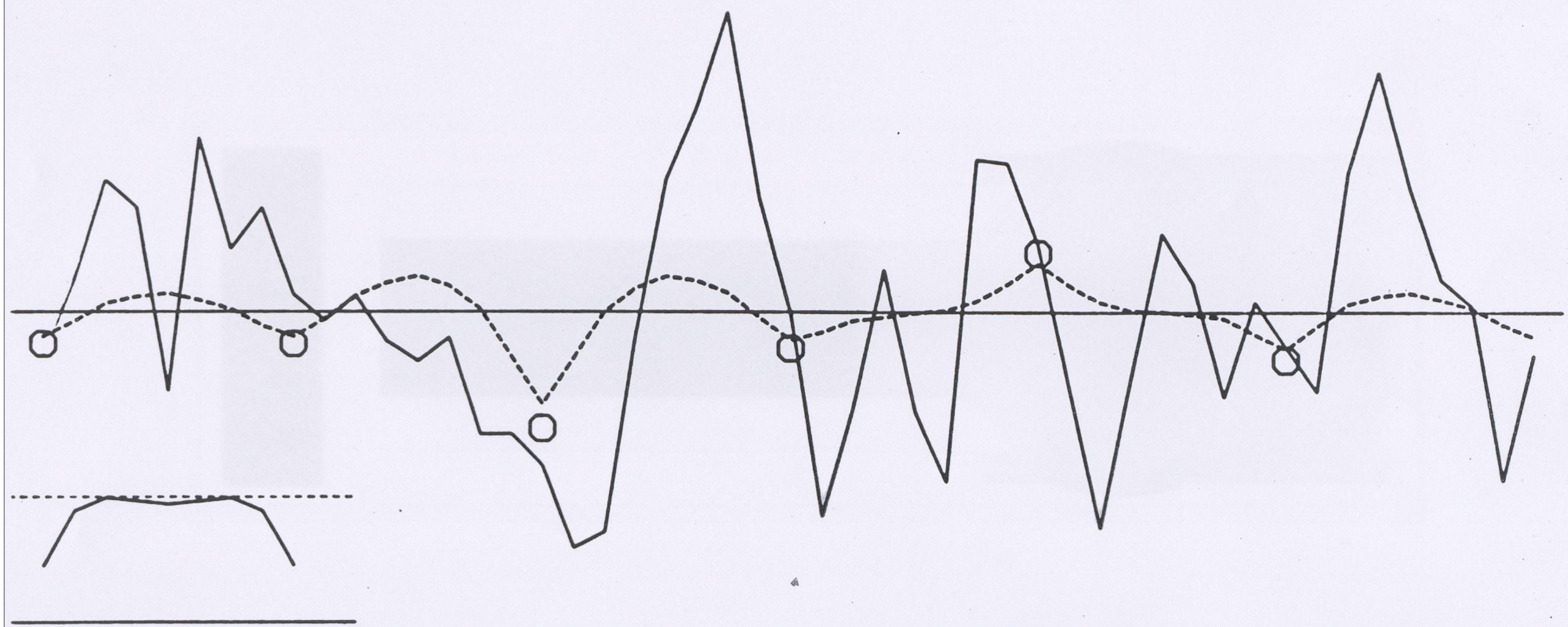
and

$$E(x'y_j') = C_\varphi(\xi, \xi_j)$$









Optimal Interpolation (continued 3)

$$x^a = E(x) + E(x'\boldsymbol{y}'^T)\,[E(\boldsymbol{y}'\boldsymbol{y}'^T)]^{-1}\,[\boldsymbol{y} - E(\boldsymbol{y})]$$

Vector

$$\boldsymbol{\mu} = (\mu_j) \equiv [E(\boldsymbol{y}'\boldsymbol{y}'^T)]^{-1}\,[\boldsymbol{y} - E(\boldsymbol{y})]$$

is independent of variable to be estimated

$$x^a = E(x) + \Sigma_j\, \mu_j\, E(x' y_j')$$

$$\varphi^a(\xi) = E[\varphi(\xi)] + \Sigma_j\, \mu_j\, E[\varphi'(\xi)\, y_j']$$
$$= E[\varphi(\xi)] + \Sigma_j\, \mu_j\, C_\varphi(\xi, \xi_j)$$

Correction made on background expectation is a linear combination of the p functions $E[\varphi'(\xi)\, y_j']$. $E[\varphi'(\xi)\, y_j']$ [$= C_\varphi(\xi, \xi_j)$], considered as a function of estimation position ξ, is the *representer* associated with observation y_j.

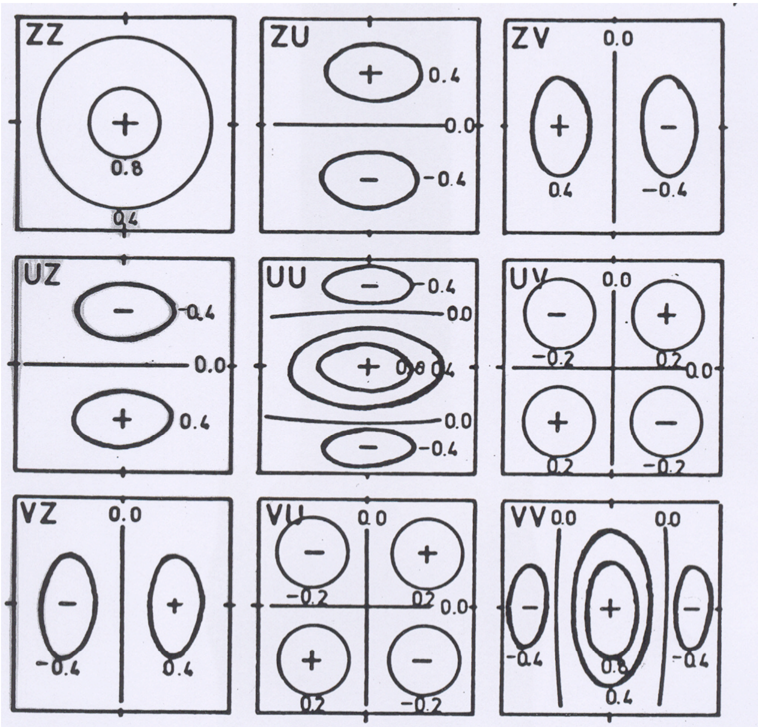
Optimal Interpolation (continued 4)

Univariate interpolation. Each physical field (*e. g.* temperature) determined from observations of that field only.

Multivariate interpolation. Observations of different physical fields are used simultaneously. Requires specification of cross-covariances between various fields.

Cross-covariances between mass and velocity fields can simply be modelled on the basis of geostrophic balance.

Cross-covariances between humidity and temperature (and other) fields still a problem.



4.: Schematic illustration of correlation functions and cross-correlation functions for multi-variate analysis derived by the geostrophic assumption.

After N. Gustafsson

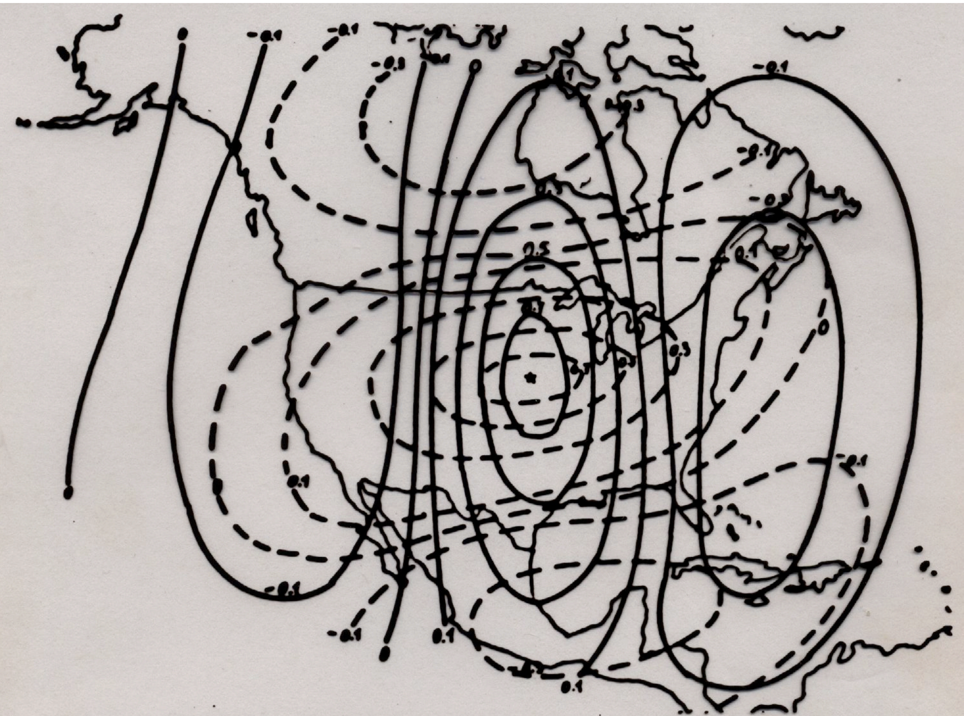


Figure 4.2.4.3: Isolines for the auto-correlation of the 500 mb u-wind component (dashed line) and the auto-correlation of the 500 mb v-wind component (full line). The "star" indicates the position of the reference station. (From Buel (1972).

After N. Gustafsson

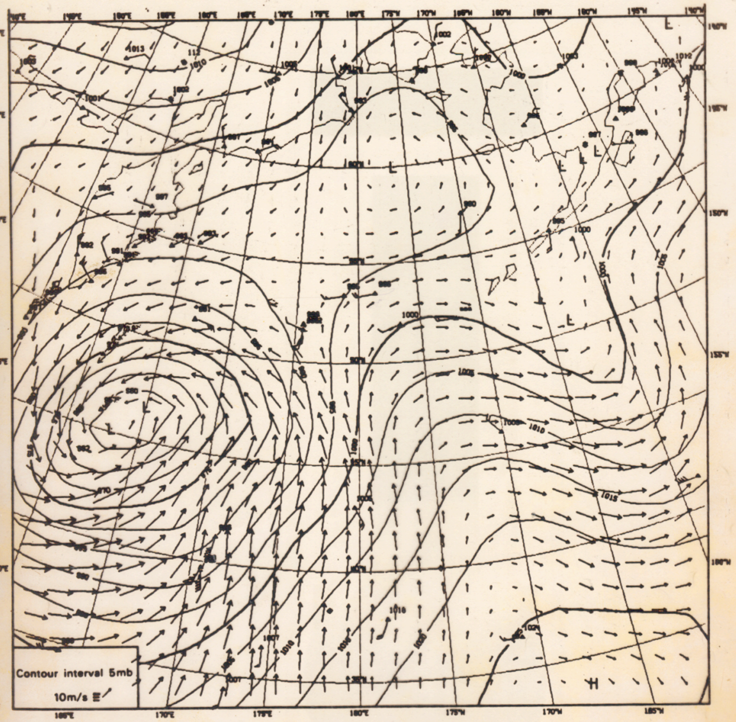


Fig. 4 Sea level pressure and wind forecast corresponding to the central area of Fig. 1, with plotted surface observations of pressure and wind (each fleche = 5 m/s).

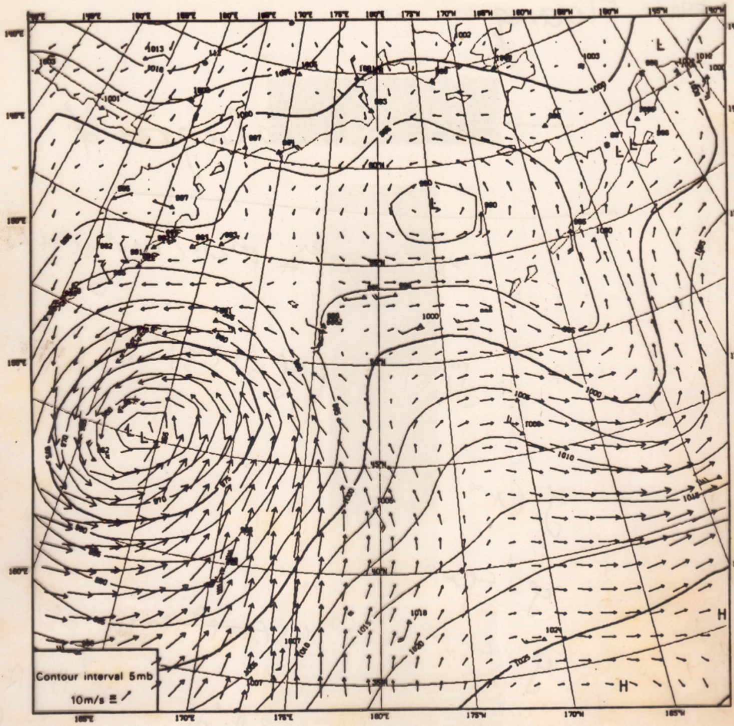


Fig. 5 As Fig. 4 for the analysis in the data assimilation cycle

After A. Lorenc

Optimal Interpolation (continued 5)

$$x^a = E(x) + E(x'\boldsymbol{y}'^{\mathrm{T}})\,[E(\boldsymbol{y}'\boldsymbol{y}'^{\mathrm{T}})]^{-1}\,[\boldsymbol{y} - E(\boldsymbol{y})] \qquad (1)$$

$$E[(x-x^a)^2] = E(x'^2) - E(x'\boldsymbol{y}'^{\mathrm{T}})\,[E(\boldsymbol{y}'\boldsymbol{y}'^{\mathrm{T}})]^{-1}\,E(\boldsymbol{y}'x') \qquad (2$$

If n-vector $\boldsymbol{x}$ to be estimated (*e*. *g*. meteorological at all grid-points of numerical model)

$$\boldsymbol{x}^a = E(\boldsymbol{x}) + E(\boldsymbol{x}'\boldsymbol{y}'^{\mathrm{T}})\,[E(\boldsymbol{y}'\boldsymbol{y}'^{\mathrm{T}})]^{-1}\,[\boldsymbol{y} - E(\boldsymbol{y})] \qquad (3)$$

$$\boldsymbol{P}^a \equiv E[(\boldsymbol{x}-\boldsymbol{x}^a)(\boldsymbol{x}-\boldsymbol{x}^a)^{\mathrm{T}}] = E(\boldsymbol{x}'\boldsymbol{x}'^{\mathrm{T}}) - E(\boldsymbol{x}'\boldsymbol{y}'^{\mathrm{T}})\,[E(\boldsymbol{y}'\boldsymbol{y}'^{\mathrm{T}})]^{-1}\,E(\boldsymbol{y}'\boldsymbol{x}'^{\mathrm{T}}) \qquad (4)$$

Eq. (3) says the same as eq. (1), but eq. (4) says more than eq. (2) in that it defines off-diagonal entries of estimation error covariance matrix $\boldsymbol{P}^a$.

If probability distributions are *globally* gaussian, eqs (3-4) achieve bayesian estimation, in the sense that $P(\boldsymbol{x} \mid \boldsymbol{y}) = \mathcal{N}[\boldsymbol{x}^a, \boldsymbol{P}^a]$.