École Doctorale des Sciences de l'Environnement d'Île-de-France
Année 2009-2010

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

Olivier Talagrand

Cours 5

21 Mai 2010

Random vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T = (x_i)$ (*e. g.* pressure, temperature, abundance of given chemical compound at n grid-points of a numerical model)

- Expectation $E(\boldsymbol{x}) \equiv [E(x_i)]$; centred vector $\boldsymbol{x}' \equiv \boldsymbol{x} - E(\boldsymbol{x})$

- Covariance matrix

$$E(\boldsymbol{x}'\boldsymbol{x}'^T) = [E(x_i' x_j')]$$

dimension nxn, symmetric non-negative (strictly definite positive except if linear relationship holds between the x_i'' s with probability 1).

- Two random vectors

$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$

$\boldsymbol{y} = (y_1, y_2, \ldots, y_p)^T$

$$E(\boldsymbol{x}'\boldsymbol{y}'^T) = E(x_i' y_j')$$

dimension nxp

Random function $\Phi(\xi)$ (field of pressure, temperature, abundance of given chemical compound, … ; ξ is now spatial and/or temporal coordinate)

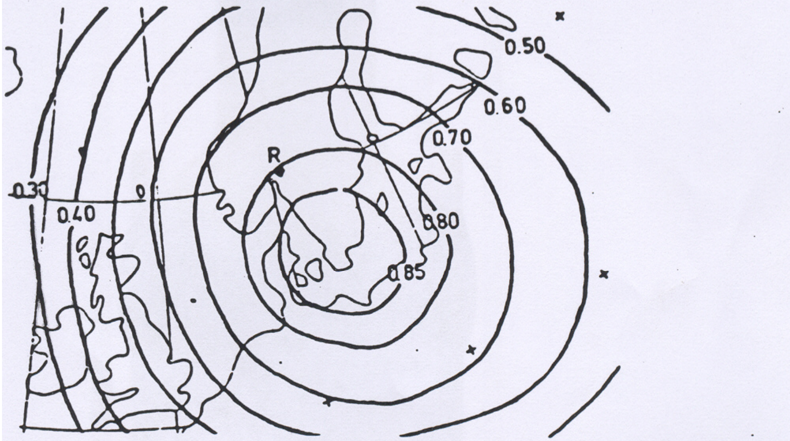
- Expectation $E[\Phi(\xi)]$; $\Phi'(\xi) \equiv \Phi(\xi) - E[\Phi(\xi)]$
- Variance $Var[\varphi(\xi)] = E\{[\varphi'(\xi)]^2\}$
- Covariance function

$$(\xi_1, \xi_2) \rightarrow C_\varphi(\xi_1, \xi_2) \equiv E[\Phi'(\xi_1)\ \Phi'(\xi_2)]$$

- Correlation function

$$Cor_\varphi(\xi_1, \xi_2) \equiv E[\Phi'(\xi_1)\ \Phi'(\xi_2)] \,/\, \{Var[\Phi(\xi_1)]\ Var[\Phi(\xi_2)]\}^{1/2}$$

.: **Isolines for the auto-correlations of the 500 mb geopotential between the station in Hannover and surrounding stations.**
From Bertoni and Lund (1963)



Isolines of the cross-correlation between the 500 mb geopotential in station 01 384 (R) and the surface pressure in surrounding stations.

After N. Gustafsson

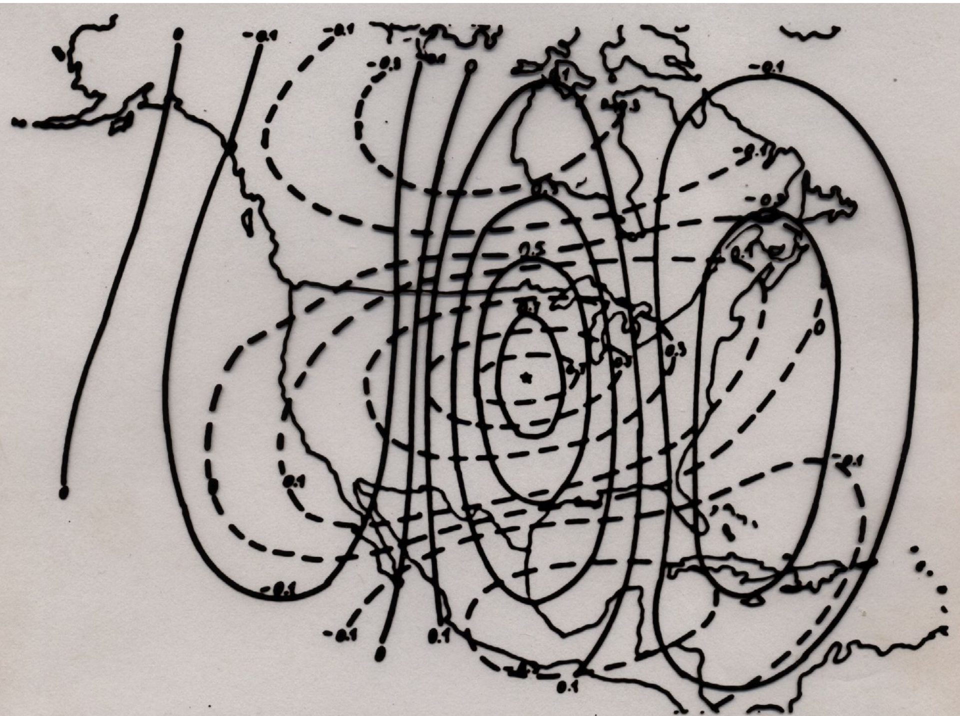


Figure 4.2.4.3: Isolines for the auto-correlation of the 500 mb u-wind component (dashed line) and the auto-correlation of the 500 mb v-wind component (full line). The "star" indicates the position of the reference station. (From Buel (1972).

After N. Gustafsson

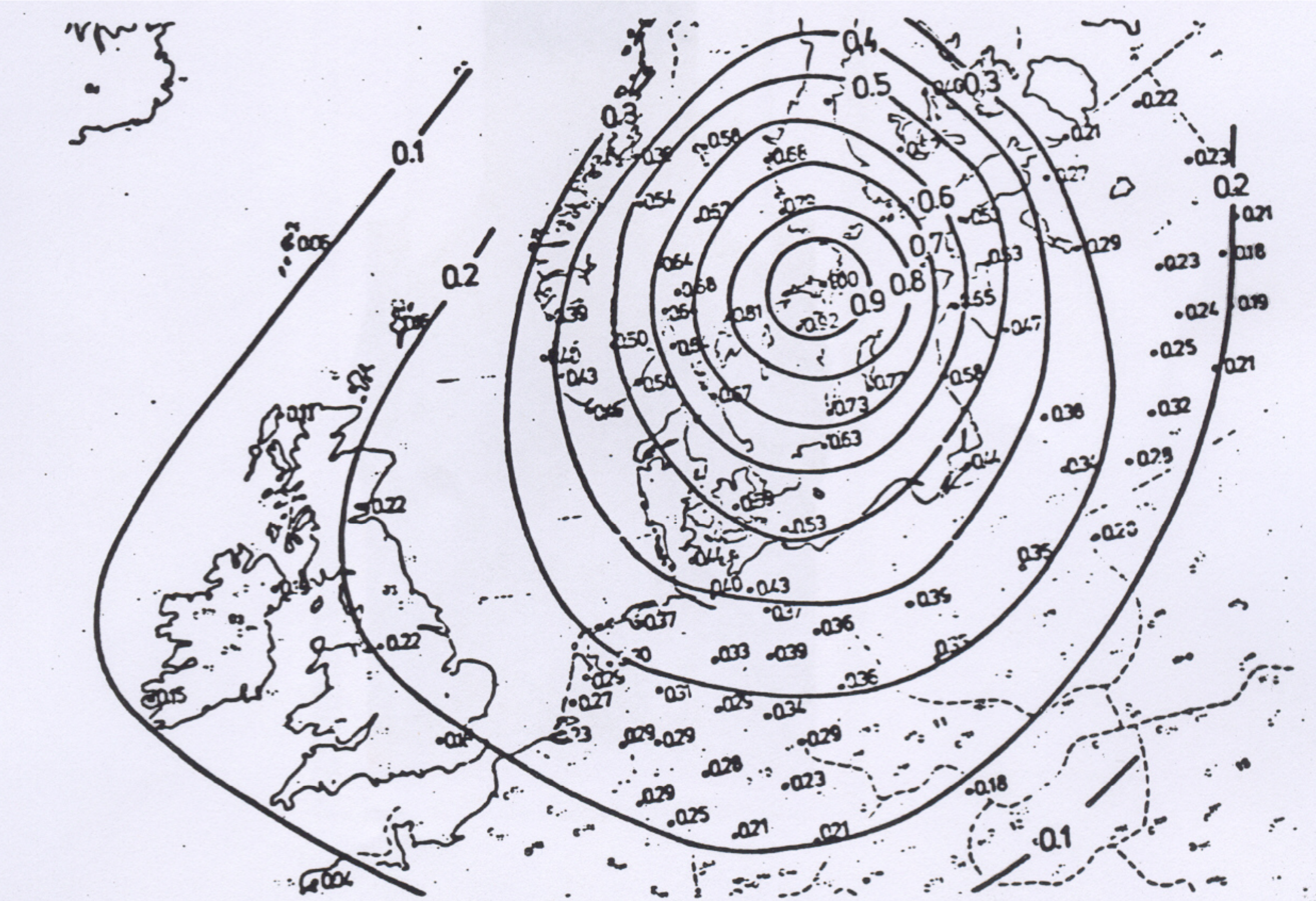


Figure 5.1.1.4.1 Auto-correlation of errors in 12h numerical forecasts of surface pressure in a reference station (Stockholm) and other stations.

After N. Gustafsson

Optimal Interpolation

Random field $\Phi(\xi)$

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

For one particular realization of the field, observations

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

Estimate $x = \Phi(\xi)$ 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

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

i. e.,

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

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

Estimate is unbiased $E(x - x^a) = 0$

Minimized quadratic estimation error

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

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

Optimal Interpolation (continued 2)

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

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

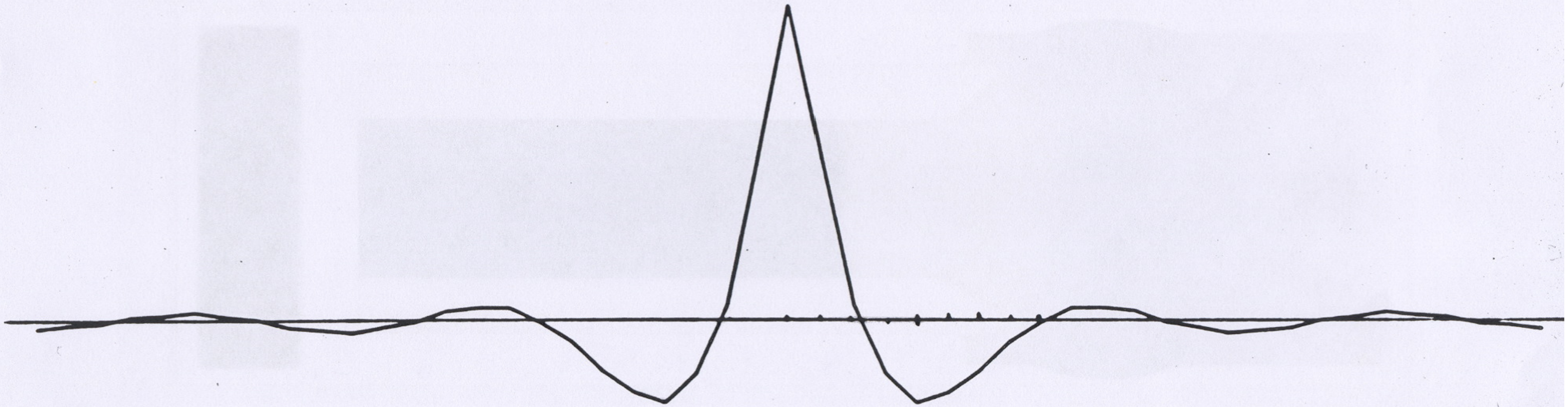
$$E(y_j'y_k') = E[\Phi'(\xi_j) + \varepsilon_j'][\Phi'(\xi_k) + \varepsilon_k']$$

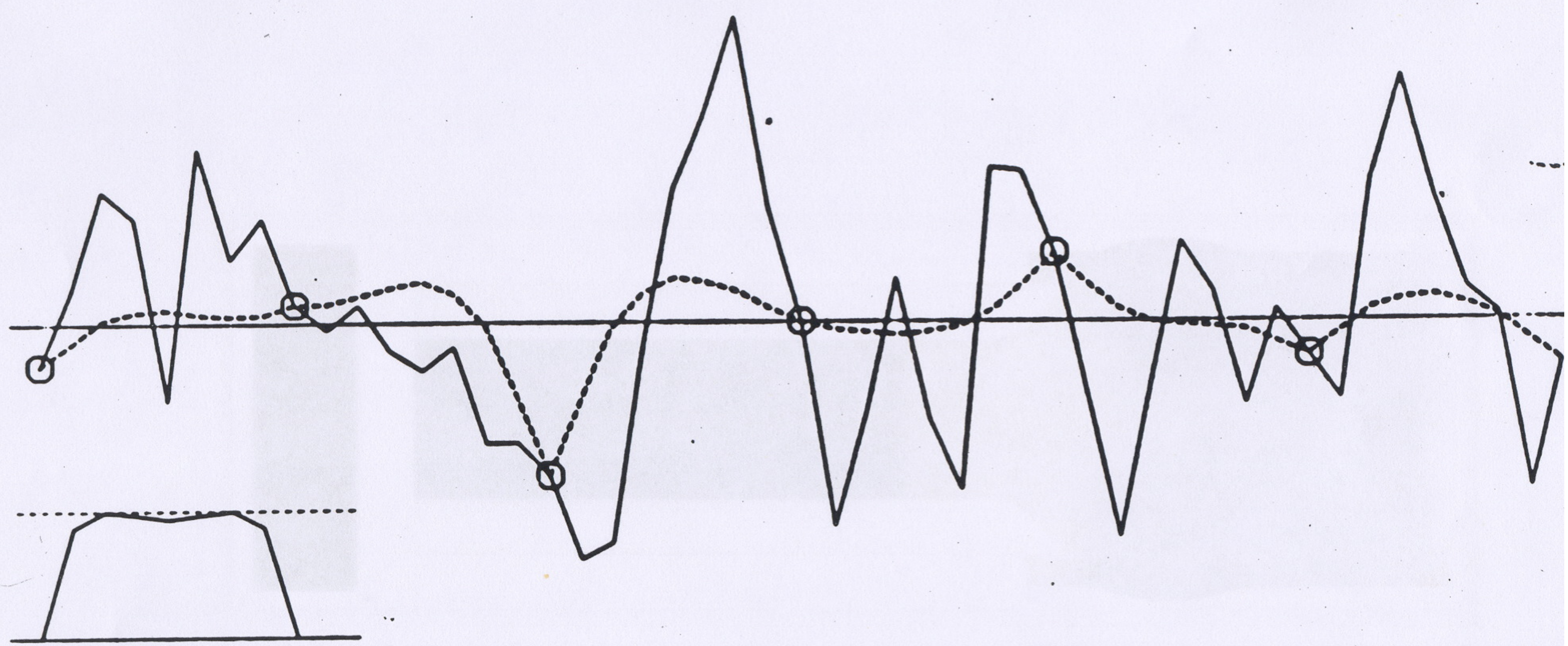
If observation errors ε_j are mutually uncorrelated, have common variance s, and are uncorrelated with field Φ, then

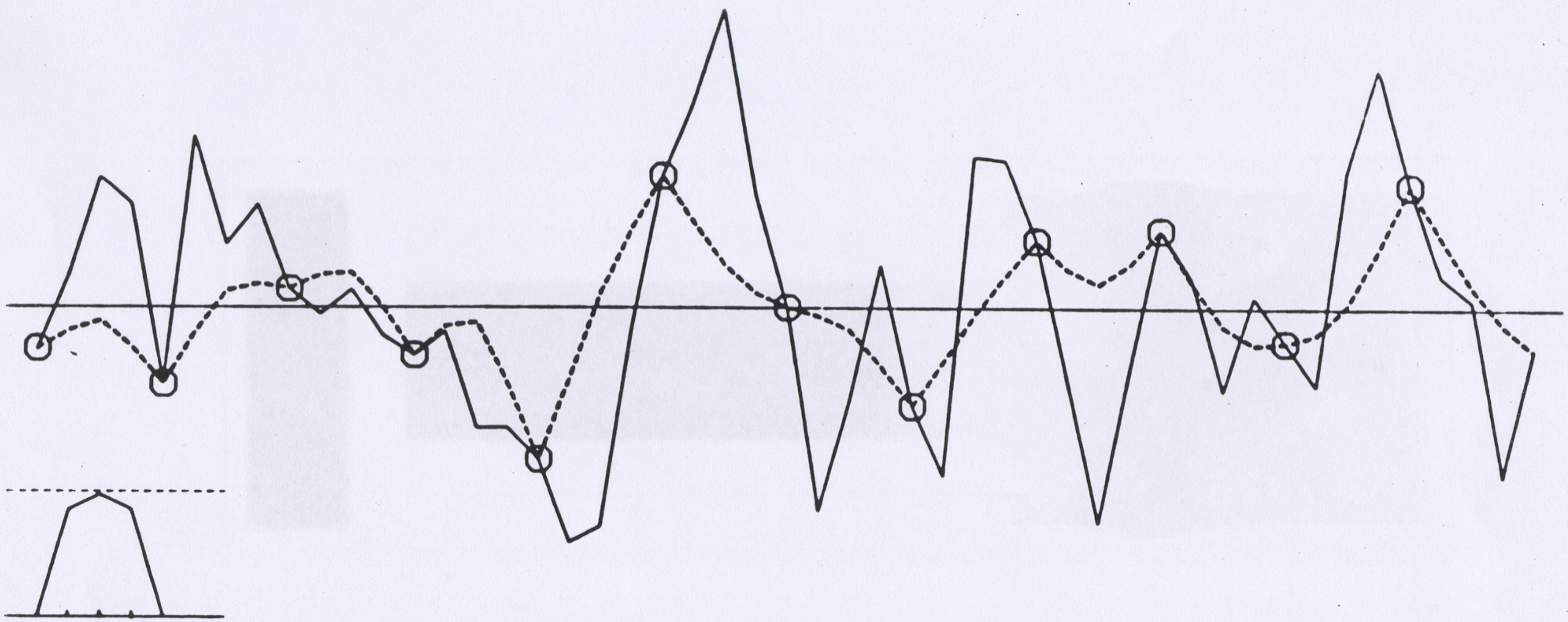
$$E(y_j'y_k') = C_\Phi(\xi_j, \xi_k) + s\delta_{jk}$$

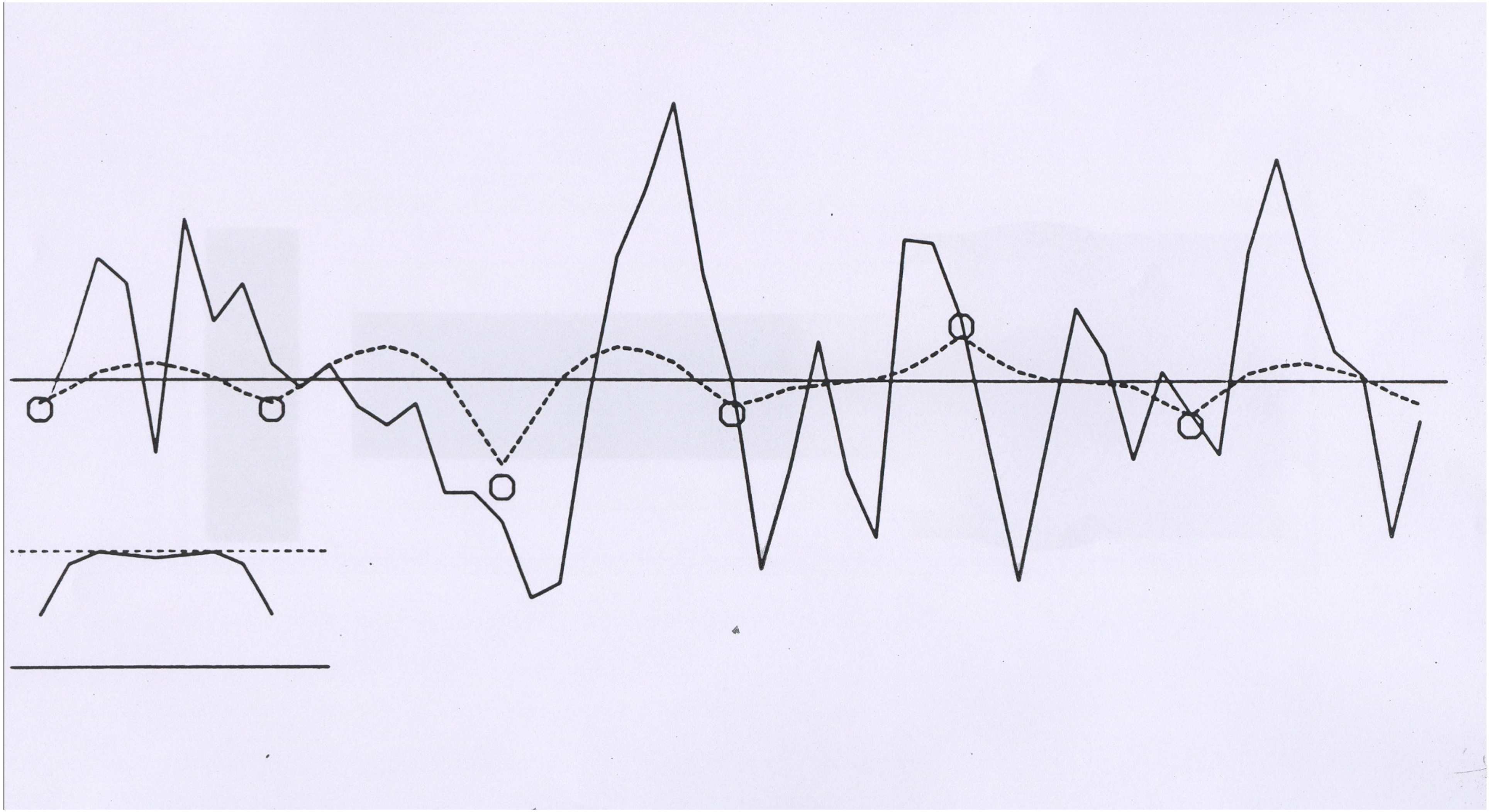
and

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









Optimal Interpolation (continued 3)

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

Vector

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

is independent of variable to be estimated

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

$$\begin{aligned} \Phi^a(\xi) &= E[\Phi(\xi)] + \Sigma_j\, \mu_j\, E[\Phi'(\xi)\, y_j'] \\ &= E[\Phi(\xi)] + \Sigma_j\, \mu_j\, C_\Phi(\xi, \xi_j) \end{aligned}$$

Correction made on background expectation is a linear combination of the p functions

$$E[\Phi'(\xi)\, y_j'].\ E[\Phi'(\xi)\, y_j']\ [\, = C_\Phi(\xi, \xi_j)\,]$$

considered as a function of estimation position ξ, is the *representer* associated with observation y_j.

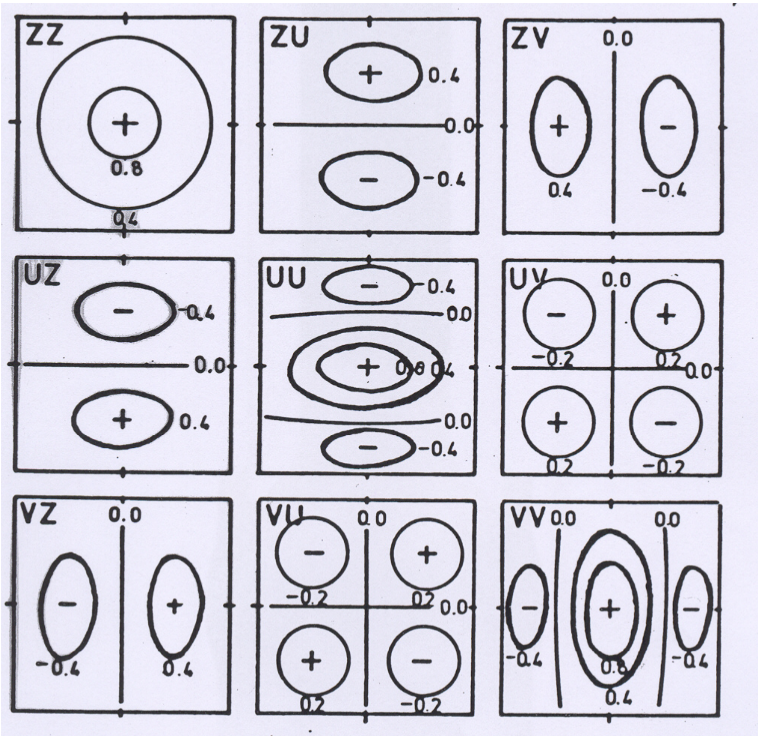
Optimal Interpolation (continued 4)

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

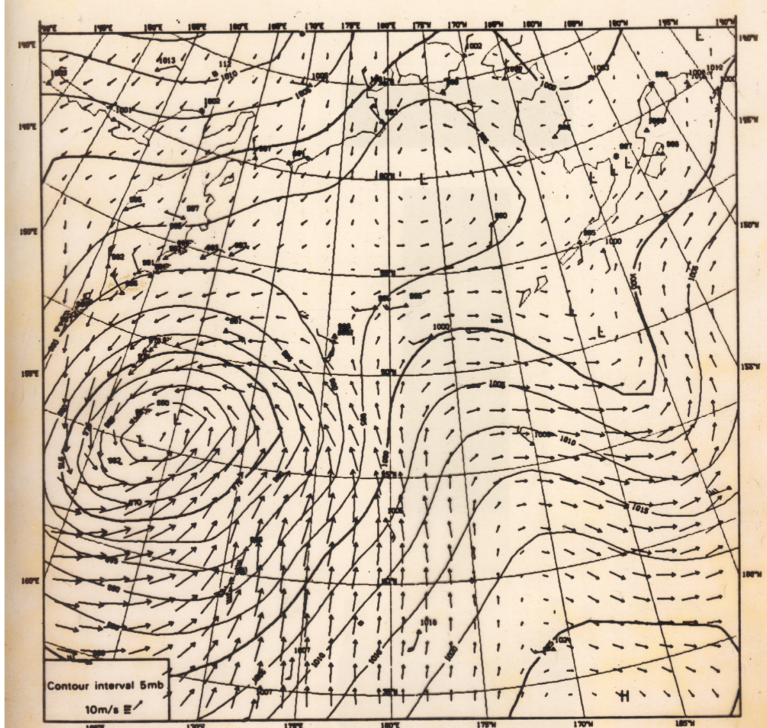


Fig. 4 Sea level pressure and wind forecast corresponding to the central area of Fig. 1, with plotted surface observations of pressure and wind (each fleche = 5 m/s).

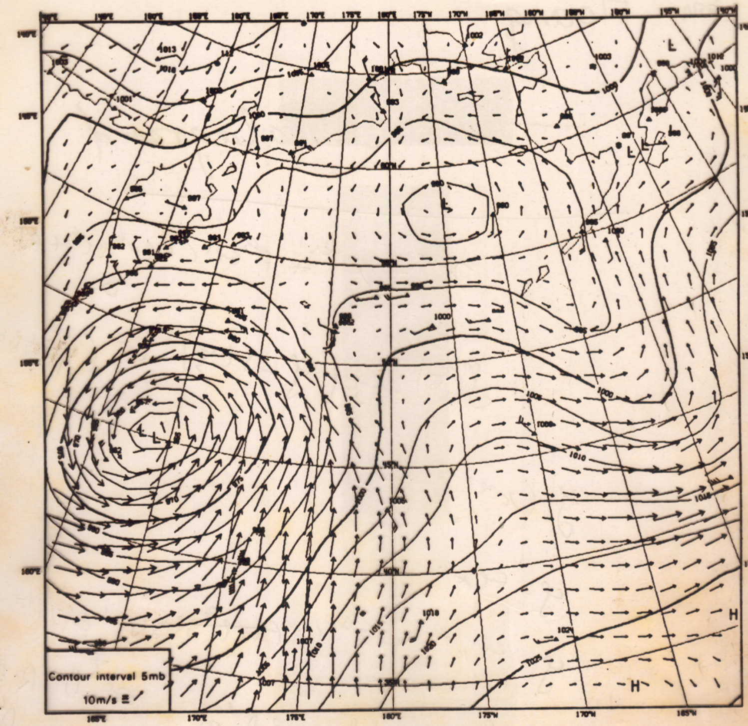


Fig. 5 As Fig. 4 for the analysis in the data assimilation cycle

After A. Lorenc