

École Doctorale des Sciences de l'Environnement d'Île-de-France
Année 2008-2009

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

Olivier Talagrand
Cours 5
11 Juin 2009

Optimal Interpolation

Random field $\varphi(\xi)$

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

For one particular realization of the field, observations

$$y_j = \varphi(\xi_j) + \varepsilon_j, \quad j = 1, \dots, p, \quad \text{making up vector } \mathbf{y} = (y_j)$$

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

$$x^a = \alpha + \sum_j \beta_j y_j = \alpha + \boldsymbol{\beta}^T \mathbf{y}, \quad \text{where } \boldsymbol{\beta} = (\beta_j)$$

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

Optimal Interpolation (continued 1)

Solution

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

i. e.,

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

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

Minimized quadratic estimation error

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

Estimation made in terms of deviations from expectations x' and y' .

Optimal Interpolation (continued 2)

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

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

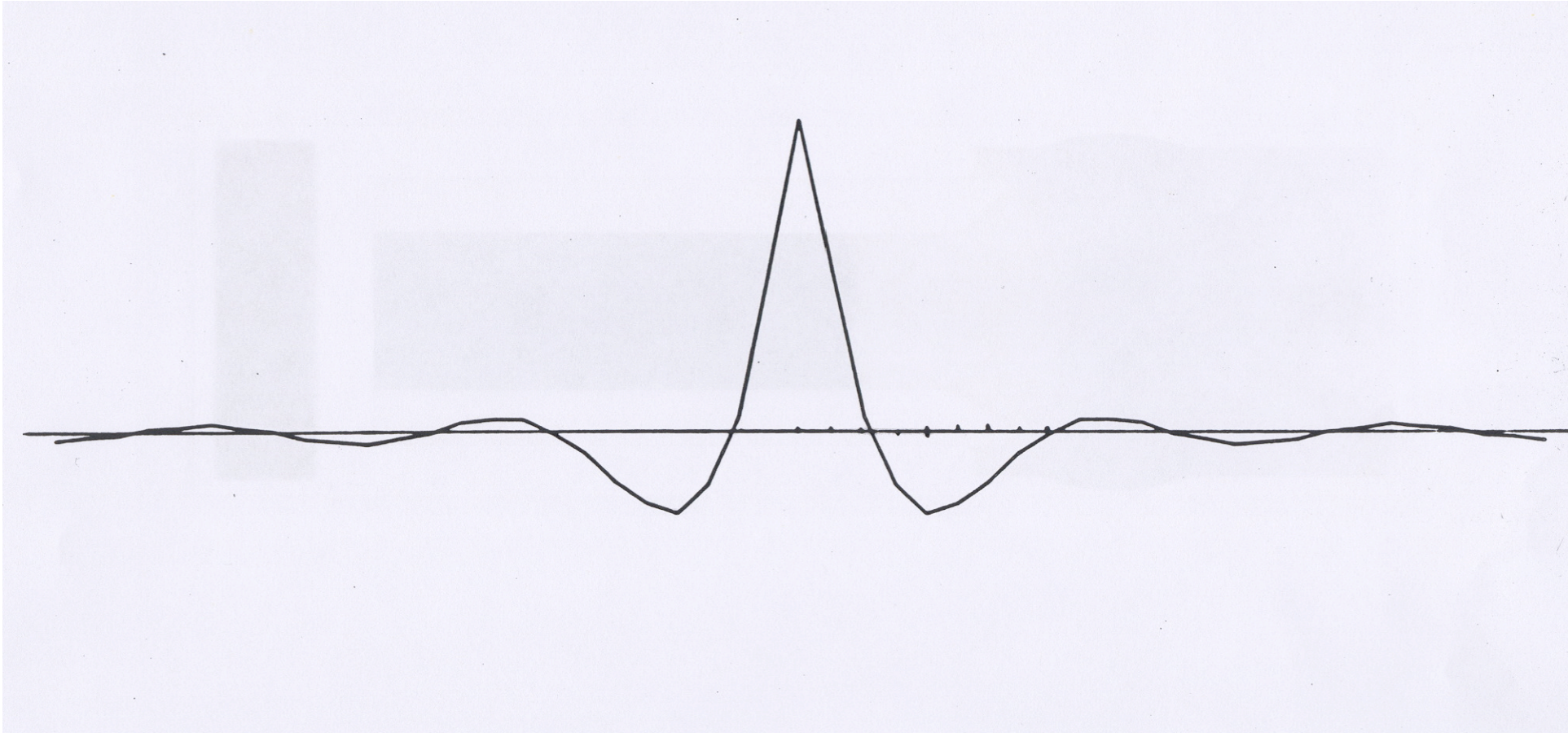
$$E(y_j'y_k') = E[\varphi'(\xi_j) + \varepsilon_j][\varphi'(\xi_k) + \varepsilon_k']$$

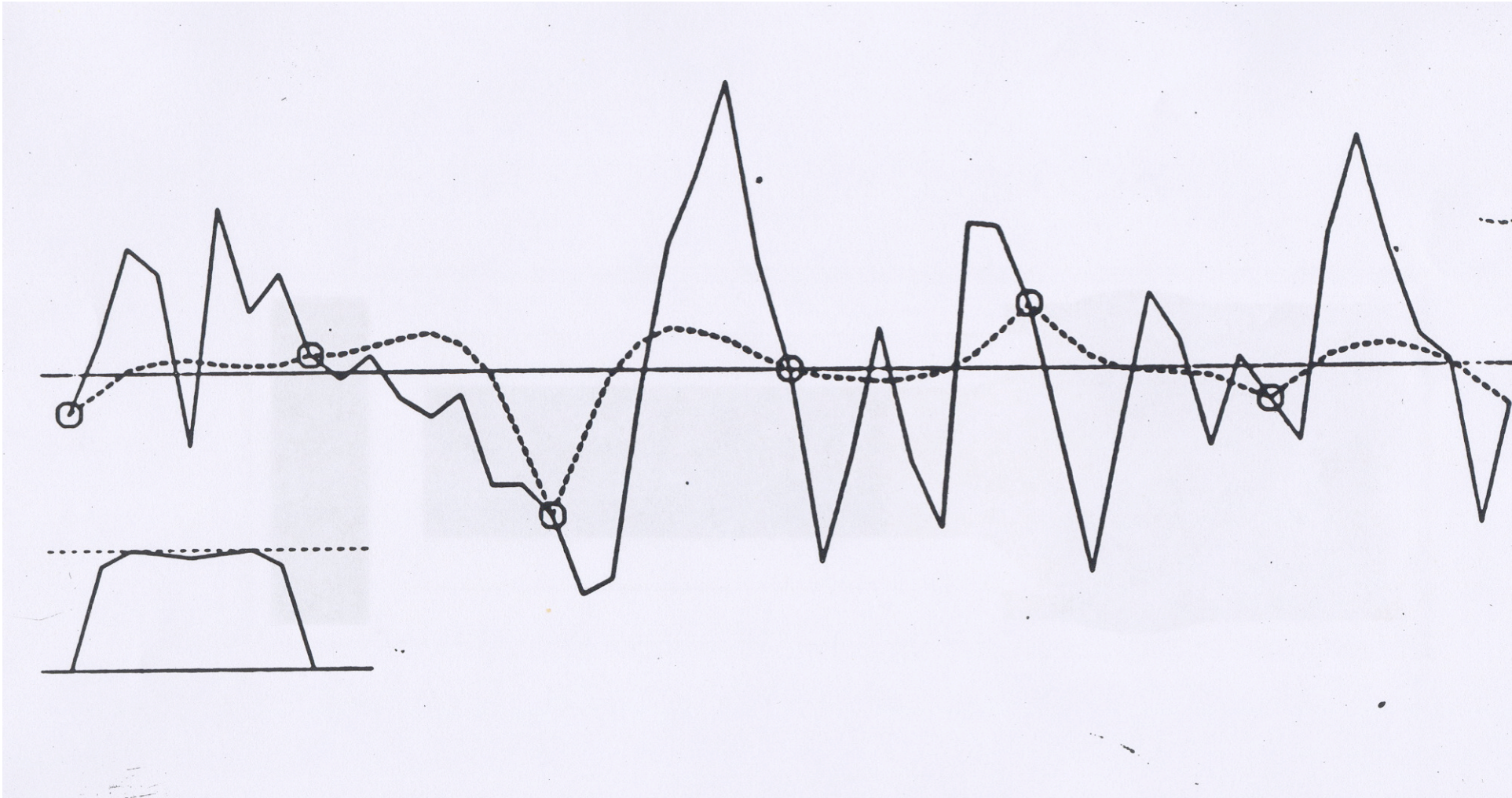
If observation errors ε_j are mutually uncorrelated, have common variance s , and are uncorrelated with field φ , then

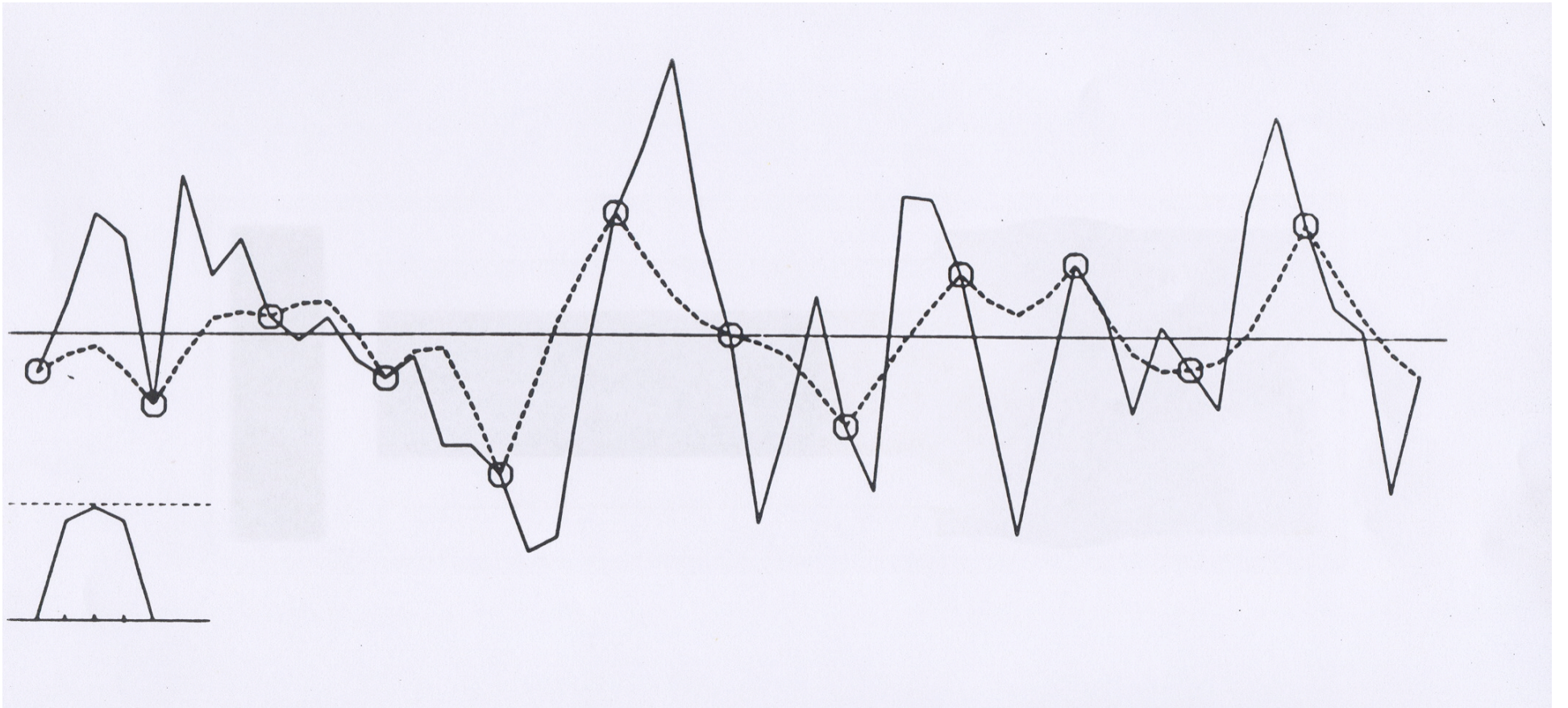
$$E(y_j'y_k') = C_\varphi(\xi_j, \xi_k) + s\delta_{jk}$$

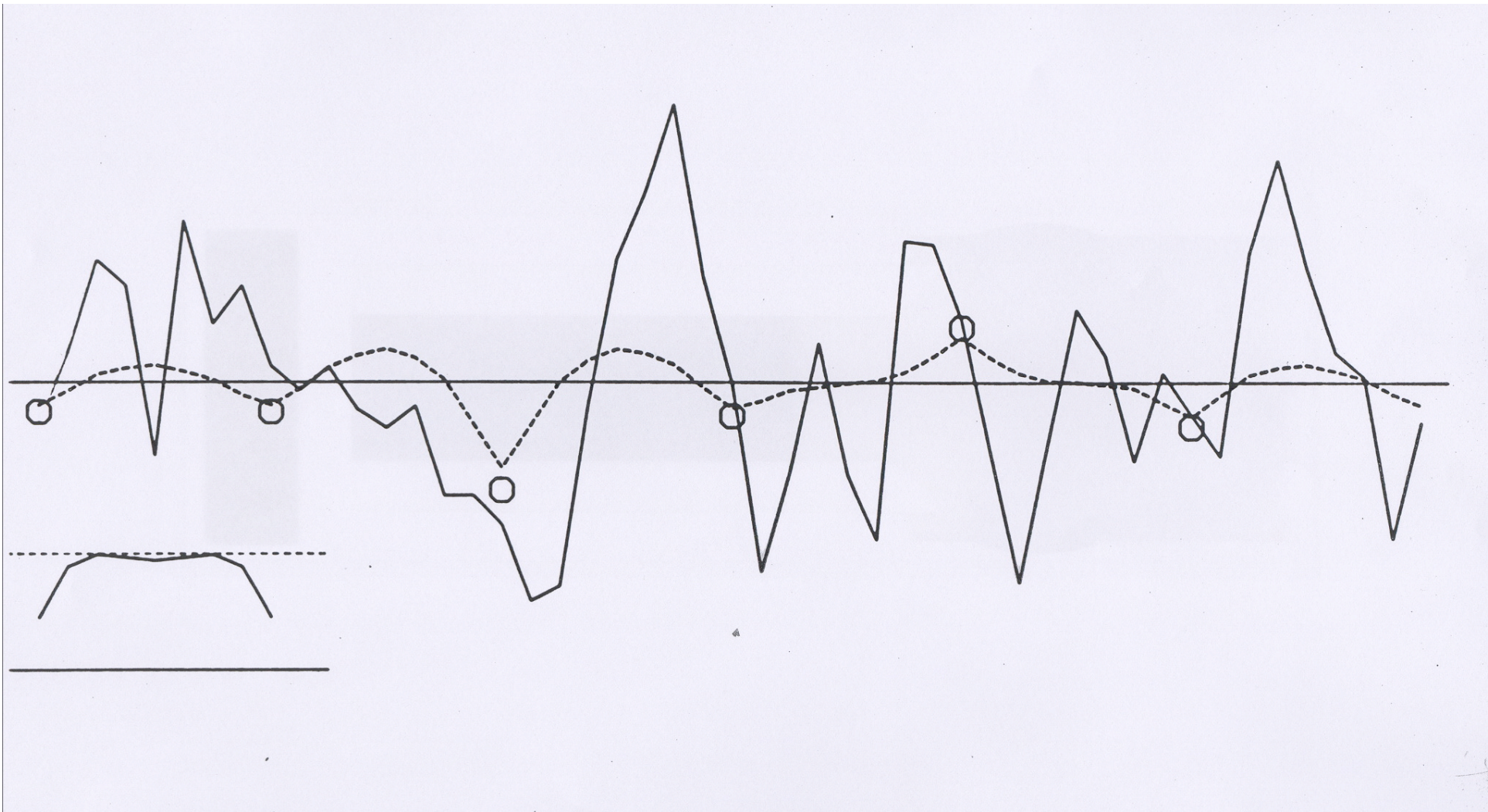
and

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









Optimal Interpolation (continued 3)

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

Vector

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

is independent of variable to be estimated

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

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

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 .

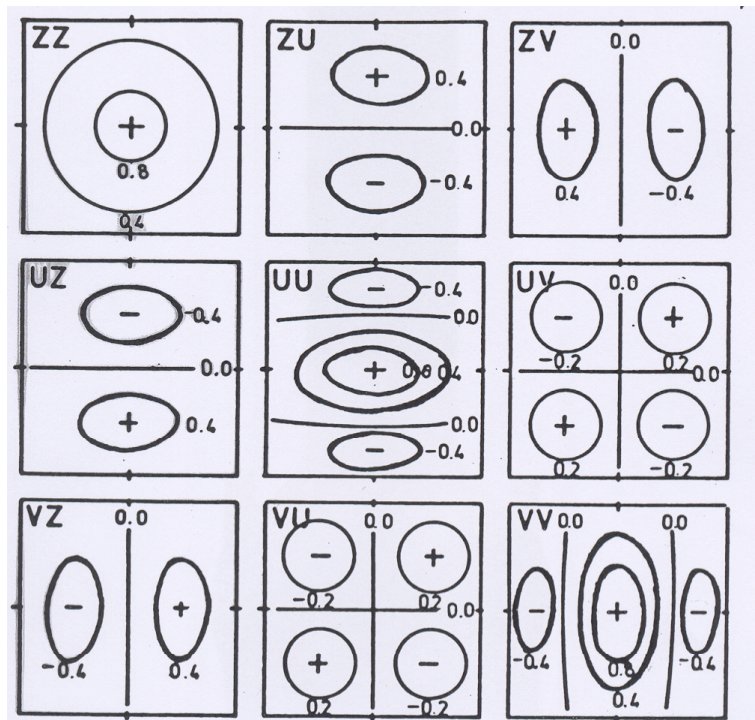
Optimal Interpolation (continued 4)

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

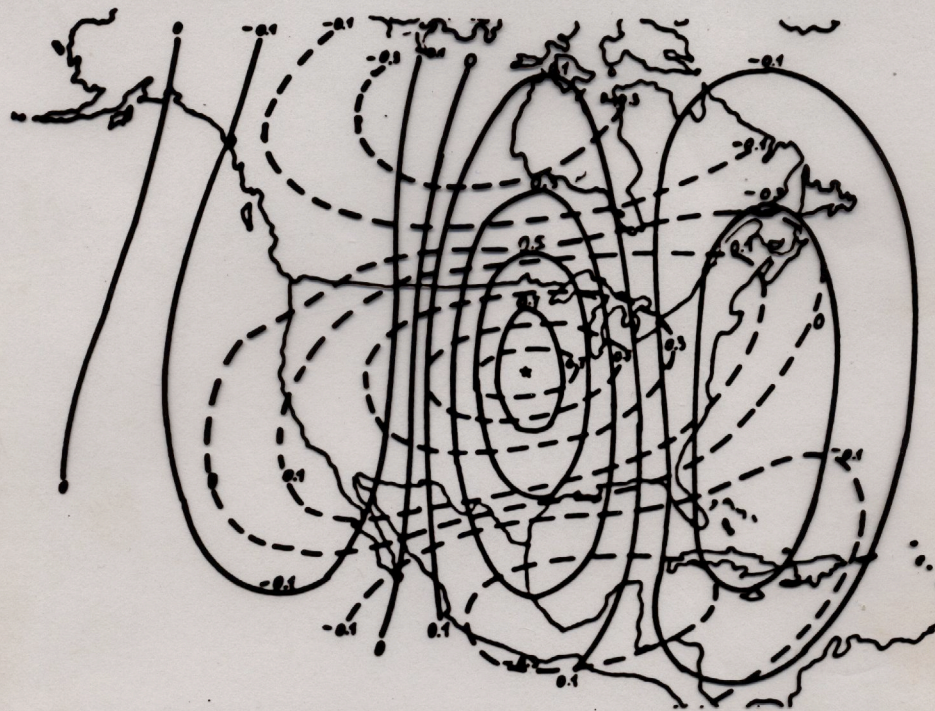
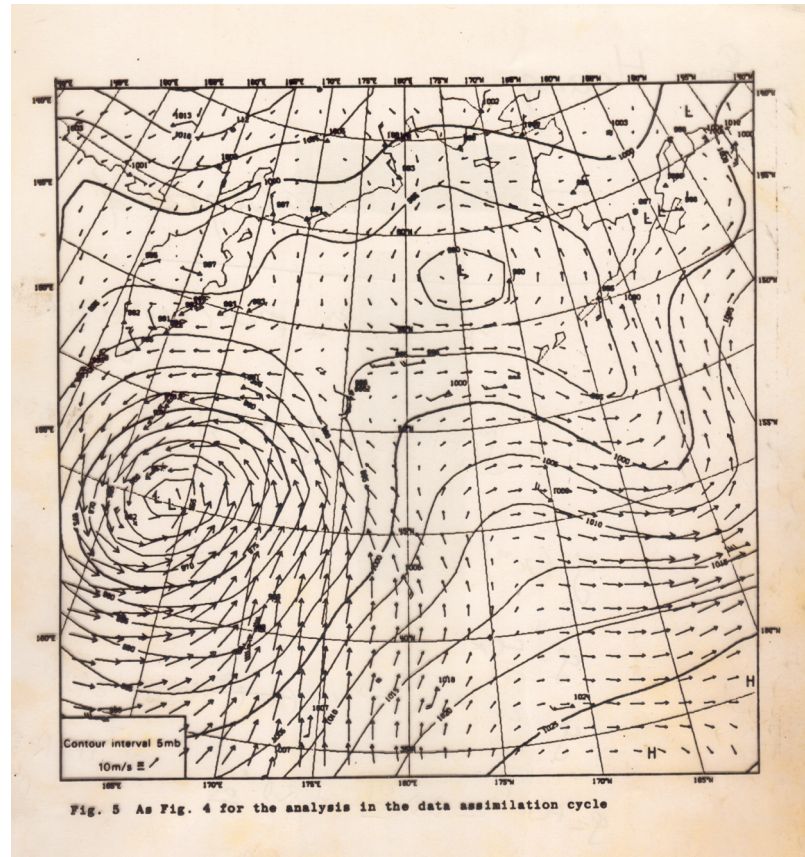
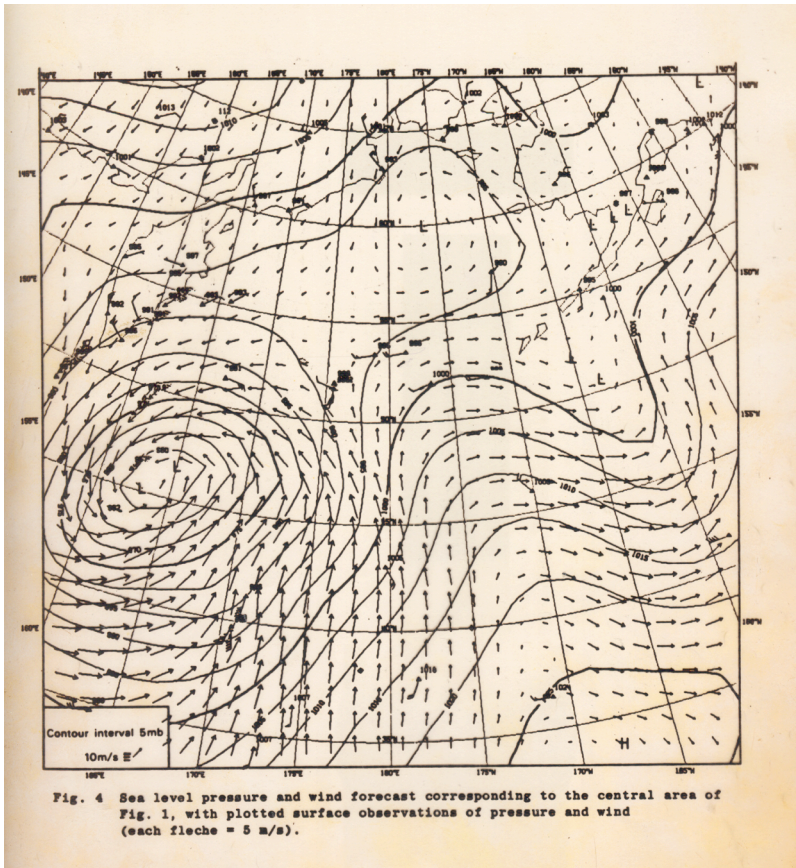


Figure 4.2.4.3: Isolines for the auto-correlation of the 500 mb u-wind component (dashed line) and the auto-correlation of the 500 mb v-wind component (full line). The "star" indicates the position of the reference station. (From Buel (1972).

After N. Gustafsson



After A. Lorenc

Optimal Interpolation (continued 5)

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

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

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

$$\mathbf{x}^a = E(\mathbf{x}) + E(\mathbf{x}'\mathbf{y}'^T) [E(\mathbf{y}'\mathbf{y}'^T)]^{-1} [\mathbf{y} - E(\mathbf{y})] \quad (3)$$

$$\mathbf{P}^a \equiv E[(\mathbf{x} - \mathbf{x}^a)(\mathbf{x} - \mathbf{x}^a)^T] = E(\mathbf{x}'\mathbf{x}'^T) - E(\mathbf{x}'\mathbf{y}'^T) [E(\mathbf{y}'\mathbf{y}'^T)]^{-1} E(\mathbf{y}'\mathbf{x}'^T) \quad (4)$$

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 \mathbf{P}^a .

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