Dévényi Dezső utolsó szakdolgozójaként szerzett „életre szóló" tapasztalatom

Izsák Beatrix

Dévényi Dezső az oktatóm

- 1999/2000 I. félév:

  Dinamikus modellezés

  Modern adatasszimiláció spec.

- 1999/2000 II. félév:

  Dinamikus modellezés

- 2000/01 I. félév:

  Szaklaboratórium

- 2000/01 II. félév:

  Statisztikus meteorológia

  Szaklaboratórium

Tenzorok ☺
Optimális interpoláció
Clusterezés
Adatasszimiláció
Tanuló algoritmusok…

A matematikánál csak egy szebb dolog van..?

Diplomamunka

A Kálmán-szűrő elmélete

- A Kálmán-szűrő
- A diszkrét Kálmán-szűrő
- Optimális becslés
- Utófeldolgozási eljárások
- A Kálmán-szűrő az utófeldolgozásban

A Kálmán-szűrő gyakorlata

- Történet, modell, adatok, QC
- Numerikus kísérletek és eredmények

Elmélet (érdekességek)

- Az optimális becslése vonatkozik Sherman tétele:

Adott y_1, y_2, …,y_n megfigyelési sorozat és skalár, konvex, szimmetrikus J(x) veszteségfüggvény mellett az E(J(x)) várható veszteséget minimalizáló optimális becslés az $\hat{\boldsymbol{x}}_{t,t} = \boldsymbol{E}(\boldsymbol{x}_t|\boldsymbol{y}_t)$ feltételes várható érték.

(**Woodbury formula**) let E, G be square invertible matrices of dimensions $n_E \times n_E$ and $n_G \times n_G$ respectively, let F and H be matrices of size $n_E \times n_G$ and $n_G \times n_E$ respectively, then the following identity holds:

$$(E + FGH)^{-1} = E^{-1} - E^{-1}F(G^{-1} + HE^{-1}F)^{-1}HE^{-1} \quad (11)$$

Veszteség függvény

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y} - \boldsymbol{H}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y} - \boldsymbol{H}(\mathbf{x}))$$

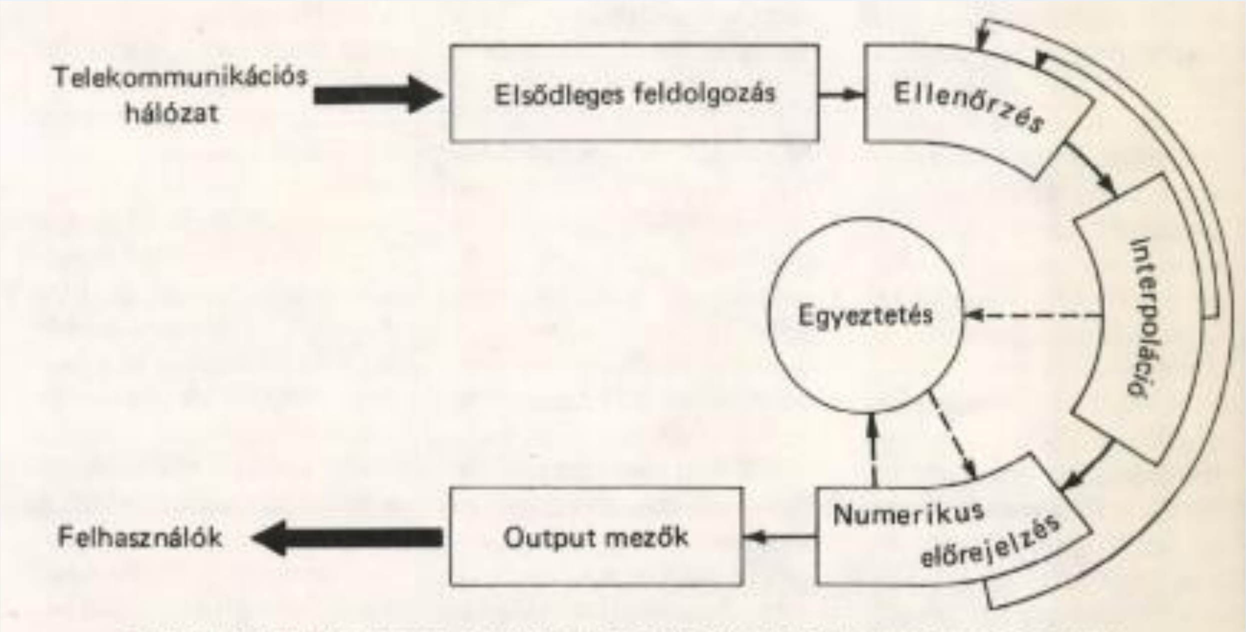
Veszteségfüggvény?

$$J(\mathbf{x})=\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)+\frac{1}{2}(\mathbf{y}-\boldsymbol{H}(\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}-\boldsymbol{H}(\mathbf{x}))$$

$$\begin{aligned}J(\mathbf{x}) = {} & (\mathbf{x}-\mathrm{E}(\mathbf{X} \mid \mathbf{X}_b=\mathbf{x}_b))^T\mathbf{B}^{-1}(\mathbf{x}-\mathrm{E}(\mathbf{X} \mid \mathbf{X}_b=\mathbf{x}_b)) + \\ & + (\mathbf{y}_0-\mathrm{E}(\mathbf{Y}_0 \mid \mathbf{X}=\mathbf{x}))^T\mathbf{R}^{-1}(\mathbf{y}_0-\mathrm{E}(\mathbf{Y}_0 \mid \mathbf{X}=\mathbf{x})),\end{aligned}$$

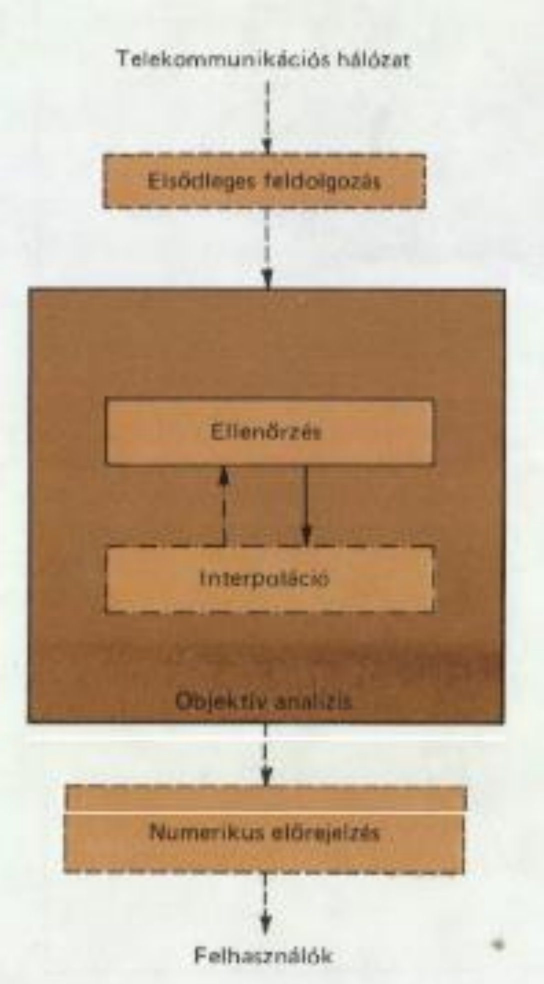
Szentimrey, T. (2016): Analysis of the data assimilation methods from the mathematical point of view. (pp. 193-205), In: Bátkai, A., Csomós, P., Faragó, I., Horányi, A., Szépszó, G. (ed.): *Mathematical Problems in Meteorological Modelling*, Springer International Publishing, Switzerland, 193–205, DOI:10.1007/978-3-319-40157-7_10.

Dévényi Dezső korábbi cikkei



1. ábra: Az operatív meteorológiai információk feldolgozásának hagyományos algoritmusa

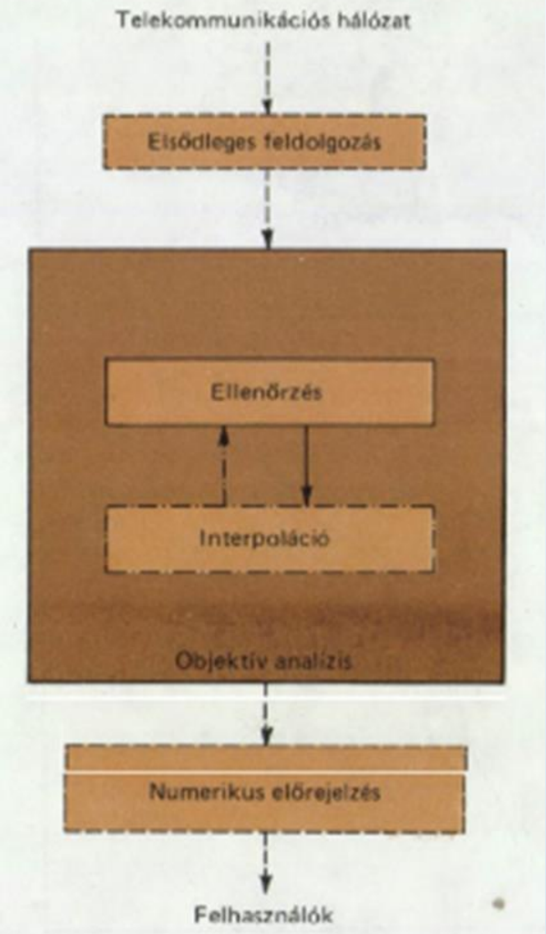
Légkör, 1979. 3 szám



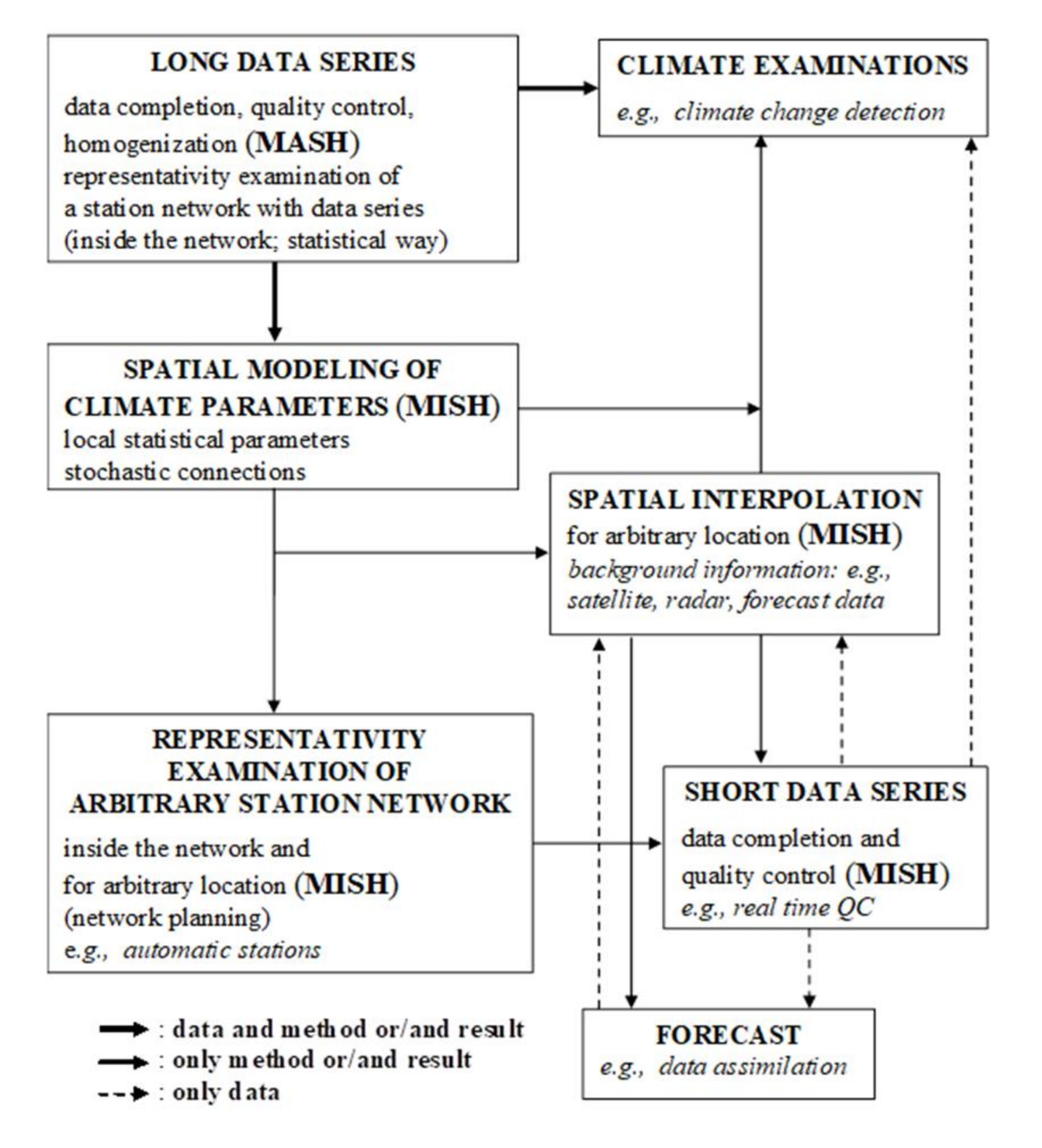
1. ábra:
A hagyományos adatfeldolgozás algoritmusa

Légkör, 1981. 2 szám

Távol vagy közel?



1. ábra:
A hagyományos adatfeldolgozás algoritmusa



A lineáris regressziós formula

A $Z(\boldsymbol{s}_0, t)$ prediktandusnak a $\boldsymbol{Z}(t)$ prediktorokra vonatkozó lineáris regressziója az alábbi formában írható fel:

$$\hat{Z}_{LR}(\mathbf{s}_0, t) = \mathrm{E}(Z(\mathbf{s}_0, t)) + \mathbf{c}^{\mathrm{T}} \mathbf{C}^{-1} (\mathbf{Z}(t) - \mathrm{E}(\mathbf{Z}(t)))$$

A különböző lineáris interpolációs formulák mindegyike visszavezethető a többváltozós lineáris regresszióra.

Determinisztikus vagy lokális paraméterek:

$$\mathrm{E}\left(Z\left(\mathbf{s}_i, t\right)\right) \left(i = 0, \ldots, M \right)$$

: várható értékek

$$\mathrm{E}\left(\mathbf{Z}(t)\right)^{\mathrm{T}} = \left[\mathrm{E}\left(Z\left(\mathbf{s}_1, t\right)\right), \ldots, \mathrm{E}\left(Z\left(\mathbf{s}_M, t\right)\right)\right]$$

: a prediktorok várható értékeinek vektora

Sztochasztikus paraméterek:

$$\mathbf{c} = \left[c_{01}, \ldots, c_{0M}\right]^{\mathrm{T}}$$

: prediktandus-prediktor kovariancia vektor,

$$\mathbf{C} = \left[\, c_{ij} \,\right]_{i,j=1}^{M}$$

: prediktor-prediktor kovariancia mátrix,

- MISH programrendszer (*Szentimrey, Bihari*)

Lineáris interpolációs formula (normál eloszlású elemekre pl. hőmérséklet):

$\sum_{i=1}^{M}(\lambda_i)=1$ feltétel miatt az alábbi formulát kapjuk a regressziós formulát és a várható értékre vonatkozó modellt felhasználva:

$$\hat{Z}_{MI}(\mathbf{s}_0, t) = E(\mathbf{s}_0) + \sum_{i=1}^{M} \lambda_i \left(Z(\mathbf{s}_i, t) - E(\mathbf{s}_i) \right)$$

$$= \sum_{i=1}^{M} \lambda_i \left(E(\mathbf{s}_0) - E(\mathbf{s}_i) \right) + \sum_{i=1}^{M} \lambda_i Z(\mathbf{s}_i, t)$$

The Optimal Interpolation Parameters are known functions of statistical parameters!

<u>Optimal constant term:</u> $\lambda_0 = \sum_{i=1}^{M} \lambda_i \left(E(\mathbf{s}_0) - E(\mathbf{s}_i)\right)$

<u>Vector of optimal weighting factors:</u> $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_M]^{\mathrm{T}}$

$$\boldsymbol{\lambda} = \mathbf{C}^{-1}\left(\mathbf{c} + \frac{\left(1 - \mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{c}\right)}{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{1}}\mathbf{1}\right) \qquad \text{(covariance form)}$$

and vector $\boldsymbol{\lambda}$ can be written as function of parameters:

$D(\mathbf{s}_0)/D(\mathbf{s}_i)$ $(i = 1, \ldots, M)$, $\mathbf{r}$, $\mathbf{R}$.

Interpolation with Background Information

Background information can decrease the interpolation error.
For example: forecast, satellite, radar data

$Z(\mathbf{s}_0, t)$: predictand

$\hat{Z}(\mathbf{s}_0, t) = \lambda_0 + \sum_{i=1}^{M} \lambda_i Z(\mathbf{s}_i, t)$: interpolation

$\mathbf{G} = \{ G(\mathbf{s}, t) \mid \mathbf{s} \in D \}$: background information on a dense grid

<u>Principle of interpolation with Background Information</u>

$$\hat{Z}_G(\mathbf{s}_0, t) = \hat{Z}(\mathbf{s}_0, t) + \mathrm{E}\left(Z(\mathbf{s}_0, t) - \hat{Z}(\mathbf{s}_0, t) \mid \mathbf{G} \right)$$

where $\mathrm{E}\left(Z(\mathbf{s}_0, t) - \hat{Z}(\mathbf{s}_0, t) \mid \mathbf{G} \right)$ is the conditional

expectation of $Z(\mathbf{s}_0, t) - \hat{Z}(\mathbf{s}_0, t)$, given $\mathbf{G}$.

Adatasszimiláció, reanalízis (a gyakorlatban)

Minimalizálandó veszteségfüggvény:

$$J(\mathbf{z}) = (\mathbf{z}-\mathbf{g})^{\mathrm{T}}\mathbf{Q}^{-1}(\mathbf{z}-\mathbf{g}) + (\mathbf{y}_0 - \mathbf{Fz})^{\mathrm{T}}\mathbf{P}^{-1}(\mathbf{y}_0 - \mathbf{Fz})$$

$\mathbf{z}$: analízis értékek (a valóság becslése)

$\mathbf{g}$: háttérinformáció (előrejelzések)

$\mathbf{y}_0$:megfigyelési értékek, $\mathbf{Q},\mathbf{P}$: kovariancia-mátrixok

Matematikai levezetés a Bayes-elmélet alapján

$$J(\mathbf{z}) = (\mathbf{z}-\mathrm{E}(\mathbf{z}|\,\mathbf{g}\,))^{\mathrm{T}}\mathbf{Q}^{-1}(\mathbf{z}-\mathrm{E}(\mathbf{z}|\,\mathbf{g}\,)) + (\mathbf{y}_0 - \mathbf{Fz})^{\mathrm{T}}\mathbf{P}^{-1}(\mathbf{y}_0 - \mathbf{Fz})$$

Általános: $\mathrm{E}(\mathbf{z}|\,\mathbf{g}\,) = \mathrm{E}(\mathbf{z}) + \mathbf{L}(\mathbf{g}-\mathrm{E}(\mathbf{g}))$, $\mathbf{Q} = \mathbf{Cov}[\mathbf{z}] - \mathbf{LCov}[\mathbf{g}]\mathbf{L}^{\mathrm{T}}$

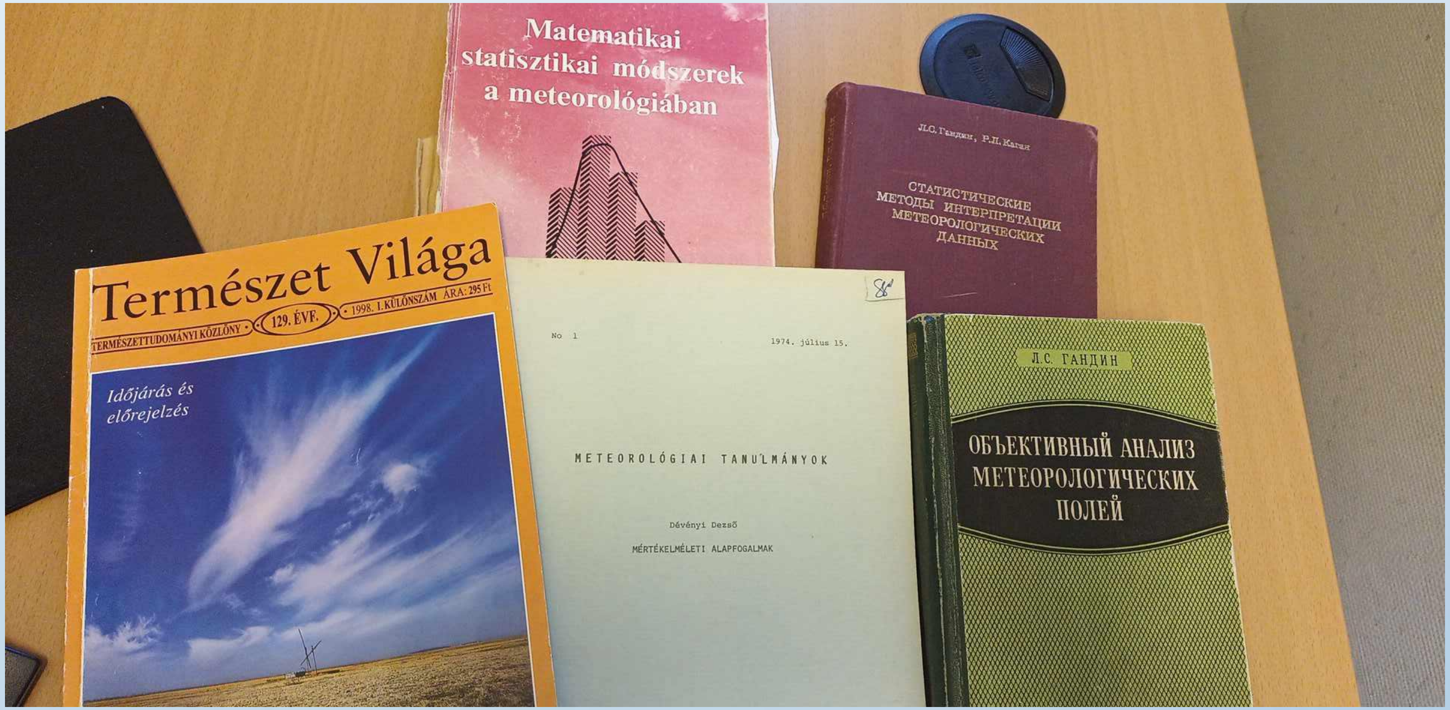
$$\mathrm{E}(\mathbf{z}|\,\mathbf{g}\,) = \mathrm{E}(\mathbf{z}) + \beta\cdot(\mathbf{g}-\mathrm{E}(\mathbf{z}))\ ,\quad \mathbf{Q} = \gamma\,\mathbf{Cov}[\mathbf{z}]$$

Éghajlati statisztikai paraméterek: $\mathrm{E}(\mathbf{z})$, $\mathbf{Cov}[\mathbf{z}]$

Köszönöm Dévényi Dezsőnek, hogy támogatott, bíztatott, segített!

Tisztelem a szakmaszeretetét és a tudását, elhivatottságát!

Dezső az adatellenőrzés, adatasszimiláció, interpoláció és minden téma esetén az elméleti alapokra helyezte a hangsúlyt. Akár a fizikai tartalomról, akár a matematikai modellről volt szó, a pontos és precíz levezetést helyezte előtérbe.



Köszönöm a figyelmet!