

Statisztikus klimatológia

HungaroMet Nonprofit Zrt.
Klimatológiai és Kutatás-fejlesztési Igazgatóság
Éghajlatkutató Osztály
2024. április 16.

Éghajlati osztály főbb témakörei

- Magyarország éghajlatának vizsgálata
 - Hosszú adatsorok elemzése
 - Megfigyelt éghajlatváltozás
 - Extrém klímaindexek változása
- Éghajlati szélsőségek vizsgálata
- Kockázat-elemzés
- Részvétel nemzetközi projektekben
- Alkalmazott klimatológiai vizsgálatok
 - Aszályindexek vizsgálata
 - Agrometeorológia
 - Éghajlati ATLASZ

Statisztikai módszerek

- Adatellenőrzés
- Homogenizálás
- Adatpótlás
- Interpoláció
- Trendvizsgálat
- Extrém indexek
- Adatbázisok: CarpatClim, HuClim, DanubeClim stb...

ADATSZERVEZÉS

Mi a probléma az adatokkal?

A minőség szempontjából: adathiányok, mérési hibák, inhomogenitások (a mérőhálózat változásából következően)

A térbeli reprezentativitás szempontjából: pontonkénti mérések, továbbá ezek és rácspontokra adott háttérinformációk (pl. radar, műhold, előrejelzési adatok) együttes kezelése.



STATISZTIKUS KLIMATOLÓGIA

MATEMATIKAI SZOFTVEREINK

http://www.met.hu/en/omsz/rendezvenyek/homogenization_and_interpolation/software/

MASHv3.03

(Multiple Analysis of Series for Homogenization; *Szentimrey, T.*)

Állomás adatsorok homogenizálása, ellenőrzése és pótlása

MISHv1.03

(Meteorological Interpolation based on Surface Homogenized Data Basis;
Szentimrey, T. and Bihari, Z.)

Éghajlati statisztikai paraméterek modellezése, meteorológiai adatok interpolációja és pótlása

A homogenizálás problematikája

Inhomogén adatsor: a mérési körülmények változásából következően, megváltozik az adatsor elemeinek valószínűségi eloszlása.

Éghajlatváltozás: időben változik az adatsor elemeinek valószínűségi eloszlása.

Homogenizálás: úgy korrigálni az elemek valószínűségi eloszlásának inhomogenitását, hogy ne rontsuk el az éghajlatváltozást.

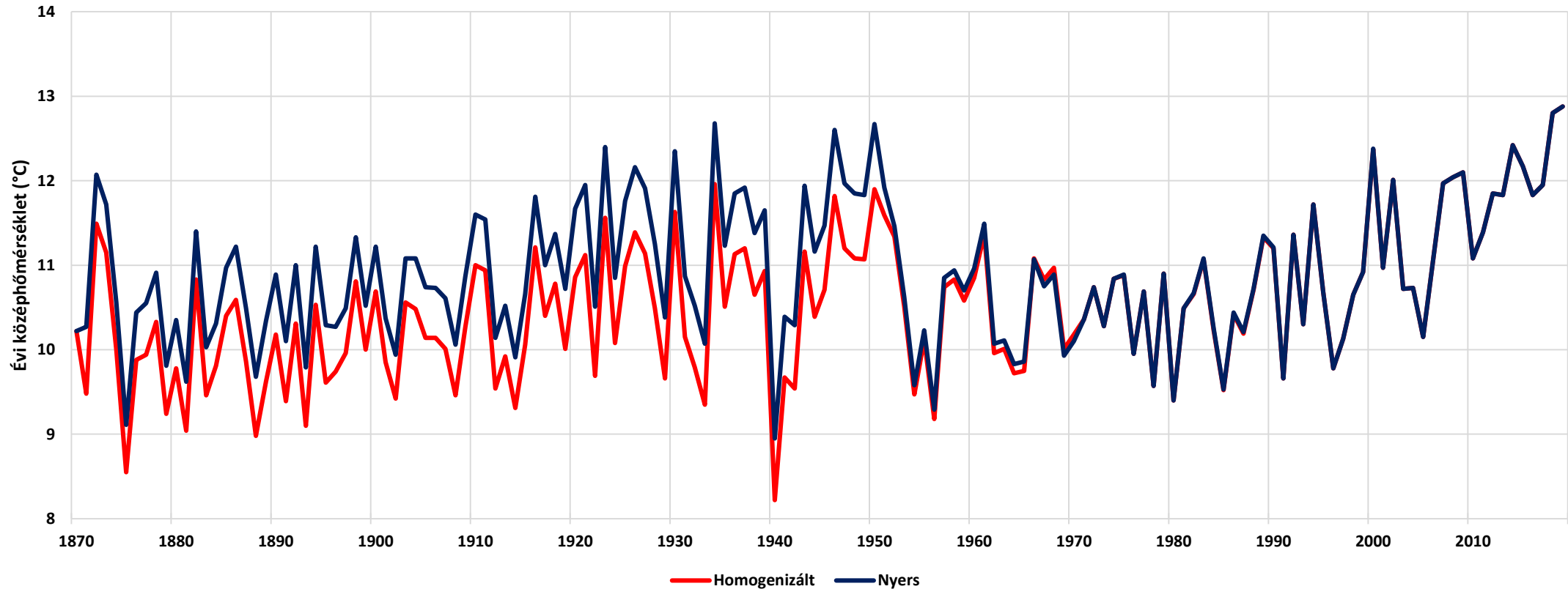
Mi lett volna, ha azonos körülmények között mértünk volna?

WMO GUIDE (**Guidelines on Homogenization**)

https://library.wmo.int/index.php?lvl=notice_display&id=21756

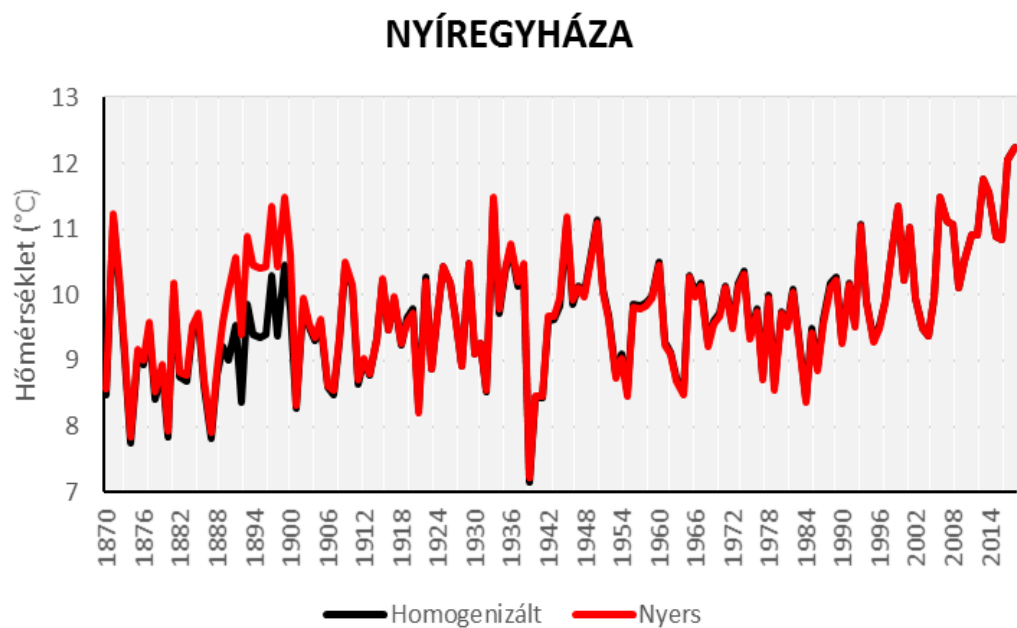
Szeged: külterületre költöztetés

Szeged



1951 repülőtérre telepítés a belvárosból. Ugyanekkor napi 3 mérés helyett napi 8 mérés.

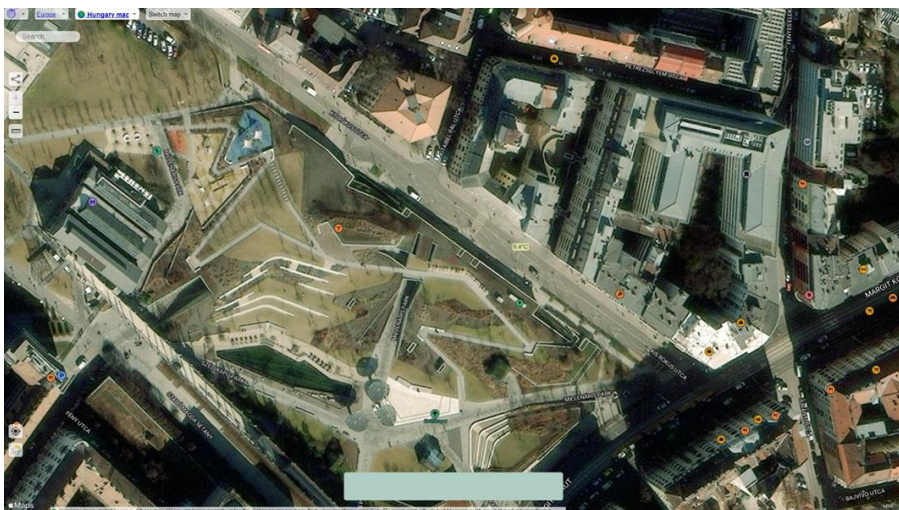
Mérési időpontok megváltozása



Nyers és **homogenizált** adatsorból számolt évi középhőmérséklet értékek, **Nyíregyháza állomáson**. A grafikonon jól látszik, hogy 1890-1901-ig a homogenizált sor jóval a nyers adatsor alatt halad, melyet az okoz, hogy az észlelések más időpontban voltak. Reggel 1 órával később, az esti mérés 1 órával előbb.

Nyíregyházán a mérési időpontokban történt váltás okoz a detektált éghajlatváltozással azonos nagyságrendű inhomogenitást.

Mérési környezet megváltozása



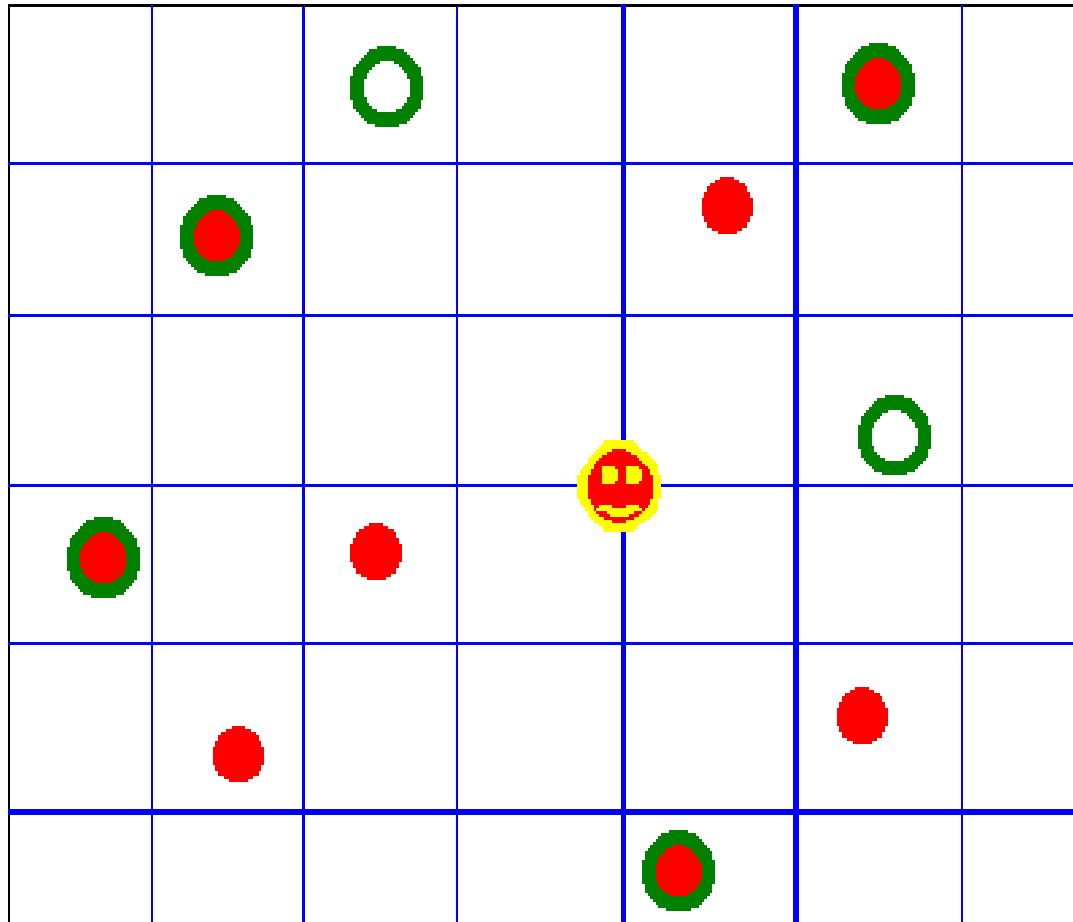
Széllkapu park jelenleg



Széllkapu park a 2010-es évek végén

A környezetben 2020-ra bekövetkező igen drasztikus változással egyértelműen magyarázható a homogenizált maximumhőmérséklet adatsorban történt nagyobb módosulások.

Földfelszíni értékekre vonatkozó információk



- : Megszűnt régi állomás, hosszú adatsorral (Térbeli és időbeli minta!)
- : Új automata állomás, rövid adatsorral (prediktor)
- : Megszűnt régi állomás és új automata állomás (prediktor) (Térbeli és időbeli minta!)
- ☺ : Tetszőleges hely, adat nélkül (prediktandus)
- + : Rácspontok háttérinformációval, pl. előrejelzés, műhold, radar

A lineáris regressziós formula

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

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

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

A térbeli interpoláció matematikai statisztikai modellje (lineáris), alapfogalmak

Meteorológiai változók

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

$Z(\mathbf{s}_i, t)$ ($i = 1, \dots, M$) : prediktorok

$\mathbf{Z}^T(t) = [Z(\mathbf{s}_1, t), \dots, Z(\mathbf{s}_M, t)]$: prediktorok vektorformában

Az \mathbf{s} helyvektor az adott D térség eleme, t az idő.

$$\hat{Z}(\mathbf{s}_0, t) = \alpha_0 + \sum_{i=1}^M \alpha_i \cdot Z(\mathbf{s}_i, t) \quad \text{Ahol } \alpha_0, \alpha_i \text{ (} i = 1, \dots, M \text{)}$$

interpolációs paraméterek.

Az interpolációs hiba jellemzése

$$MSE(\mathbf{s}_0) = E\left(\left(Z(\mathbf{s}_0, t) - \hat{Z}(\mathbf{s}_0, t)\right)^2\right)$$

$$RMSE(\mathbf{s}_0) = \sqrt{MSE(\mathbf{s}_0)}$$

Azok a matematikailag optimális interpolációs paraméterek, amelyekre az interpolációs hiba minimális.

Determinisztikus vagy lokális paraméterek:

$$E(Z(\mathbf{s}_i, t)) \quad (i = 0, \dots, M) \quad : \text{várható értékek}$$

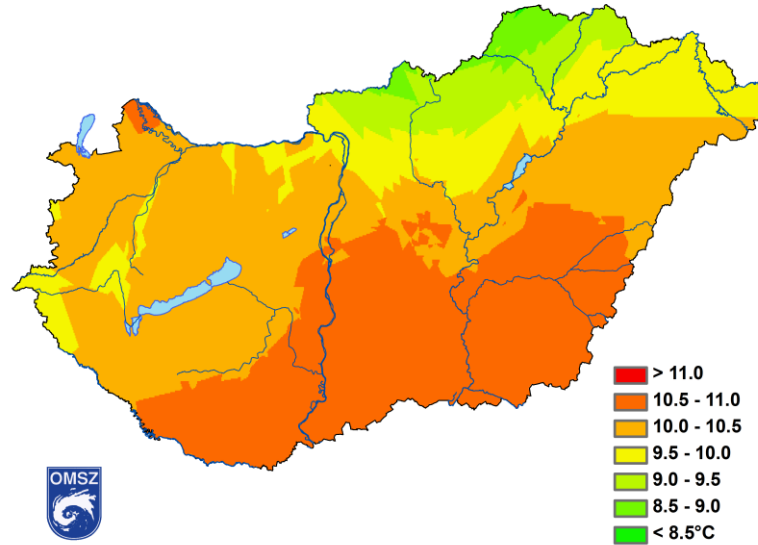
$$E(\mathbf{Z}(t))^T = [E(Z(\mathbf{s}_1, t)), \dots, E(Z(\mathbf{s}_M, t))]$$
 : a prediktorok várható értékeinek vektora

Sztochasztikus paraméterek:

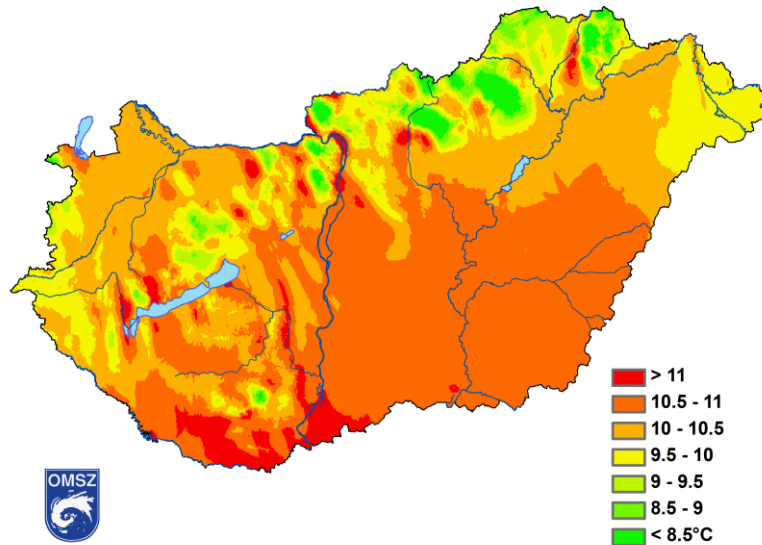
$$\mathbf{c} = [c_{01}, \dots, c_{0M}]^T \quad : \text{prediktandus-prediktor kovariancia vektor,}$$

$$\mathbf{C} = [c_{ij}]_{i,j=1}^M \quad : \text{prediktor-prediktor kovariancia mátrix,}$$

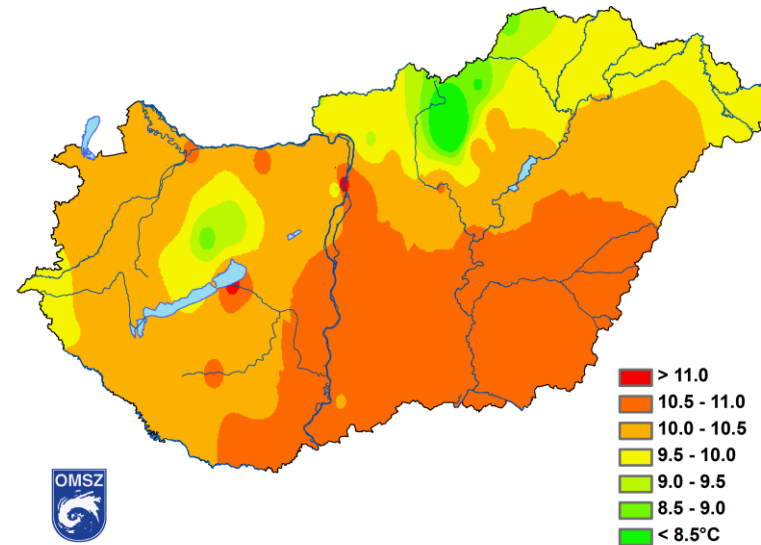
Közönséges Kriging



MISH



Inverz távolság



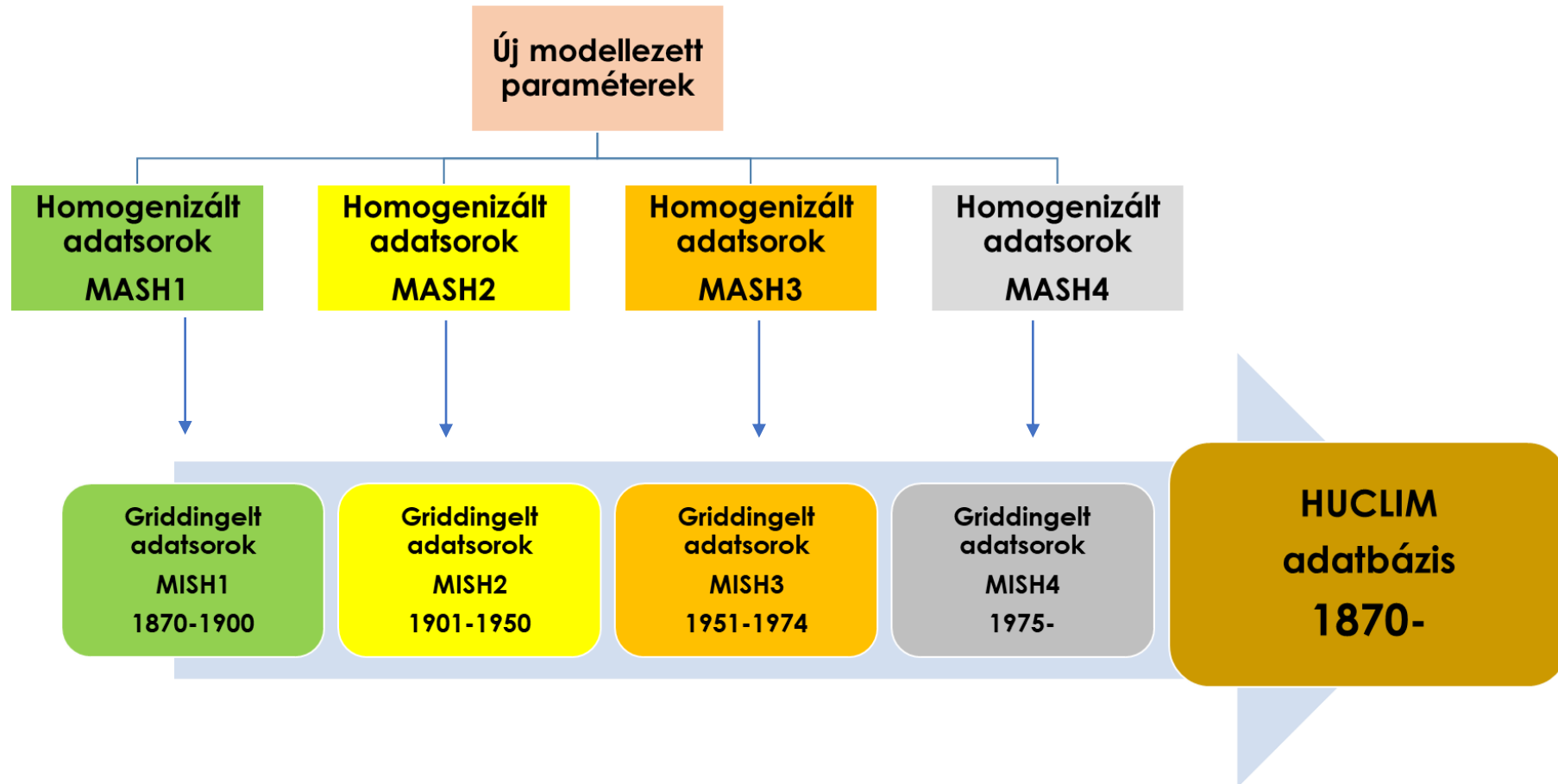
Módszer: **Cross validation**: training (tanító) és test adatbázis

Interpoláció hibája (**RMSE**)

Hőmérséklet (1981-2010)	IDW	Ordinary Kriging	Mish
Év	0.65	0.69	0.22
Tavas	0.83	0.74	0.18
Nyár	0.85	0.79	0.26
Ősz	0.69	0.70	0.18
Január	0.45	0.43	0.20
Július	1.11	1.14	0.29

Csapadék (1981-2010)	IDW	Ordinary Kriging	Mish
Év	37.48	40.39	16.93
Tavas	8.87	11.12	3.83
Nyár	12.01	13.68	4.84
Ősz	7.57	8.48	3.60
Január	2.93	3.17	1.46
Július	4.36	5.13	2.25

HUCLIM adatbázis előállításának folyamata Napi középhőmérséklet esetén



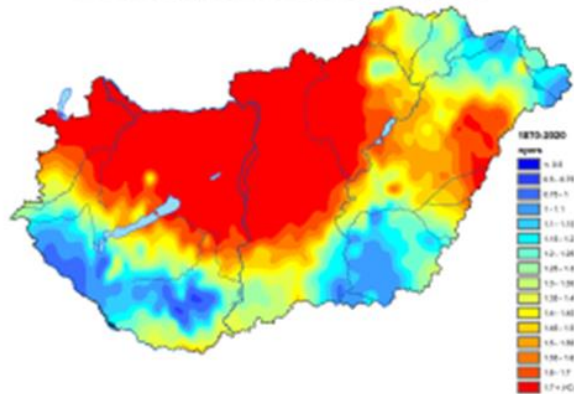
*Izsák, B., Szentimrey, T., Lakatos, M., Pongrácz, R., Szentes, O.: Creation of a representative climatological database for Hungary from 1870 to 2020, *Időjárás* 126, 1-26.*

doi:10.28974/idojaras.2022.1.1

Mennyire tér el egymástól a nyers adatsorok és a homogenizált sorok trendje? 1870-2020

Éghajlatváltozás ?

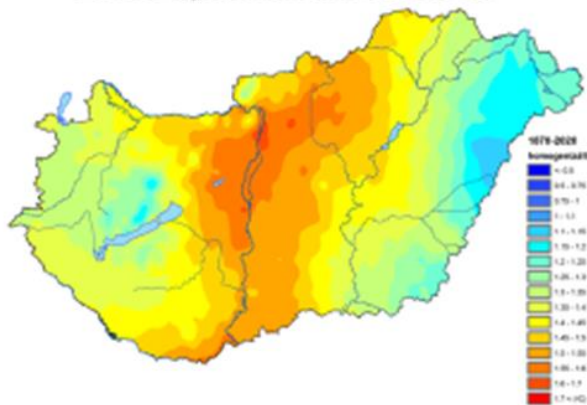
Évi középhőmérséklet változás



- *Nyers adatsorokkal dolgozom*
- *Nem baj, hogy változtak a mérési körülmények*
- *Nem kell adatellenőrzés*
- *Gyorsan publikálom az eredményeket*



Évi középhőmérséklet változás



- *Adatellenőrzést végzek*
- *Pótlom a hiányzó értékeket*
- *Homogenizált adatsorokat elemzek*
- *Adekvát matematikai módszereket és szoftvereket használok*
- *Eredményeket csak ezután publikálom*



Néhány tématerület az ÉO-n

1. **EMNL, VVNL**
2. **DIMOP**
3. **Tervezői adatszolgáltatás (csapadékintenzitás)**
4. **Danube Data Cube**
5. **ATLASZ**
6. **Adatsorok: odp.met.hu**



HUNGARIAN NATIONAL
LABORATORY




Napi rácsponti adatsorok az odp.met.hu-n

https://odp.met.hu/climate/homogenized_data/gridded_data_series/

- 1971-től:
 - hőmérséklet (közép, min, max), csapadékösszeg, légnyomás (állomásszinti), relatív nedvesség átlaga
- 2001-től:
 - globálsugárzás összege, maximális szélökés és átlagos szélesebesség



ORSZÁGOS
METEOROLÓGIAI
SZOLGÁLAT Meteorológiai Adattár

Name	Last modified	Size	Description
 Parent Directory		-	Homogenizált adatsorok
 gridded_data_series/	2021-03-30 08:11	-	Rácsponti adatsorok
 station_data_series/	2021-03-30 08:13	-	Állomási adatsorok

Köszönöm a figyelmet!

Budapest, 2024