# Verifikáció

Tóth Boglárka
toth.b@met.hu

**Numerikus előrejelzés 2025/2026/1. félév**

***2025.11.17.***

HungaroMet
Magyar Meteorológiai Szolgáltató Nonprofit Zrt.

# Az előadás vázlata:

**Verifikáció alapjai**

- Verifikáció kérdései
- Módszertan

**“Determinisztikus” előrejelzések verifikációja**

- Objektív verifikáció
  - Pontbeli verifikáció
  - Térbeli verifikáció
- Szubjektív verifikáció

**Verifikáció a gyakorlatban**

- HungaroMet alkalmazásai
- Fejlesztési irányok

# A verifikáció alapjai

# Mi is az a verifikáció?

## Mi is az a verifikáció?

## Mi is az a validáció?

## Mi is az a verifikáció?

## Mi is az a validáció?

## Mi is az a értékelés?

## Mi is az a verifikáció?

- Előrejelzés minőségének felmérése
- Ellenőrzés, összehasonlítás: Az előrejelzett jelenség valóban bekövetkezett? → modell-előrejelzés beválását vizsgáljuk a megfigyelések adatainak felhasználásával

- "I am building the product right?"

## Mi is az a validáció?

- "Megbízhatóságot" vizsgálok
- Annak biztosítása, hogy egy modell vagy modell komponens azt csinálja, amit kell.
- Segít a gyenge teljesítmény eredetének meghatározásában

- "I am building the right product?"

## Mi is az a értékelés?

- Összetett teljesítmény vizsgálat
- Kérdés: Mennyire jó?
- Alkalmazott módszer mennyire felel meg a céloknak?

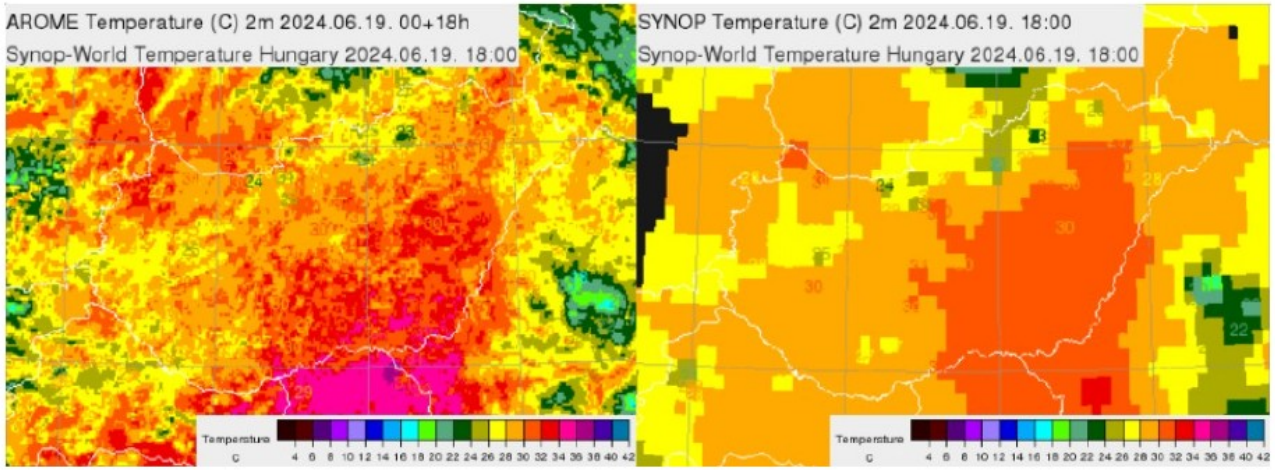
# Verifikáció kérdései

- Miért kell ez nekünk?
  - Visszacsatolás
  - Előrejelzések folyamatos monitorozása → időbeli változások nyomon követése
  - Fejlesztési irányok meghatározása → A javulás érdekében fel kell tenni a kérdést:

    *“Mit csinálunk rosszul?”*

  - Modellhibák feltárása
  - Döntéshozatal támogatása

**ACCORD User representative jelentés**



# Verifikáció kérdései

- **Előrejelzés minőségének felmérése**
  - Előrejelzések összehasonlítása → Mit hasonlítok össze mivel?

Ugyanazt értjük két változó alatt?

- Különböző előrejelzéseket
- Előrejelzéseket **megfigyelésekkel**
  - Felszíni szinoptikus adatok
  - Magaslégköri megfigyelések
  - Távérzékelési adatok (radar, műhold)

↓

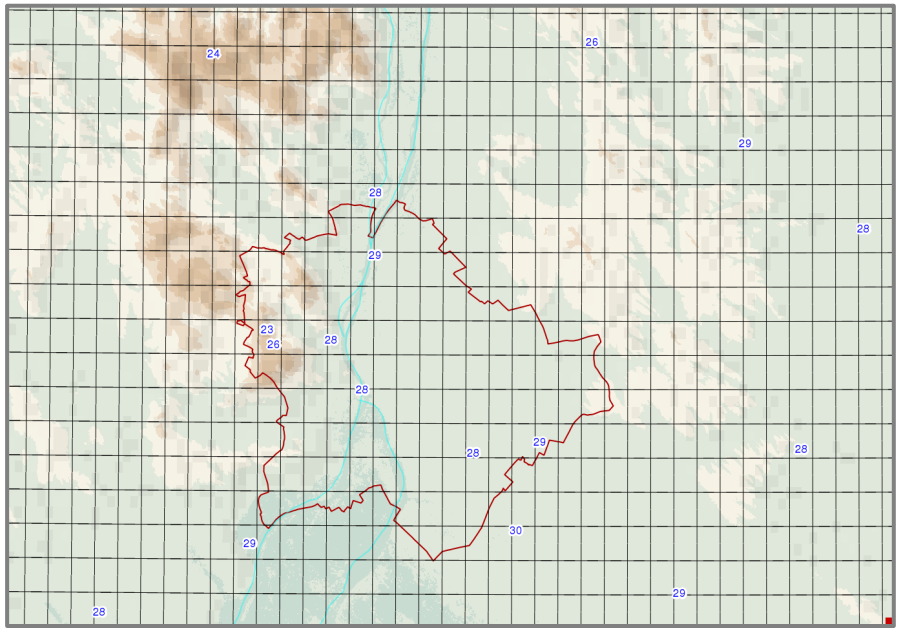
Adatellenőrzésen esnek át, de sokszor hibával terheltek

↓

Figyelmen kívül hagyunk (magas jel-zaj arány)

# Verifikáció kérdései

- **Előrejelzés minőségének felmérése**
  - Milyen idő és térbeli felbontással?
    - Reprezentativitás → Interpoláció




  - Milyen változót szeretnék verifikálni?
    - Folytonos eloszlású (pl. hőmérséklet) → **Pontszerű**
    - Diszkrét (csapadék) → **Térbeli**
    - Folytonos eloszlású kategorizált paraméter (>30 m/s széllökés)
    - Kummulált érték (24 órás csapadékösszeg)

↓

**Módszertan**

# Verifikáció módszertan

- **Csoportosítás**
  - Előrejelzés természete
    - **Determinisztikus**
    - Valószínűségi (Ensemble előrejelzések II.)
  - Megközelítés
    - **Objektív**
    - **Szubjektív**
  - Paraméter természete
    - **Pontbeli**
    - **Térbeli**

➡ **Számos megjelenítés**

Egy átfogó értékelés komplex feladat:

Fontos különböző verifikációs módszerek vegyítése

# “Determinisztikus” előrejelzések verifikációja

## Objektív verifikáció

# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

- **Pontszerű verifikáció**
- A verifikáció egy adott pontban történik, az ott végzett mérések alapján.
  1. Az előrejelzést és a megfigyelést azonos rácsra kell hozni.
  2. Előrejelzett és mért adatok közti különbségek számszerűsítése.

- **Szisztematikus hiba (bias)** $$bias = \frac{1}{n}\sum_{i=1}^{n}(f_i - o_i)$$
- **Mean Squared Error (MSE)** $$MSE = \frac{1}{n}\sum_{i=1}^{n}(f_i - o_i)^2$$
- **Root Mean Square Error (RMSE)** $$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_i - o_i)^2}$$

- $f_i$= előrejelzési érték
- $o_i$= megfigyelés
- $n$= vizsgált esetek száma

→ **Számos megjelenítés**

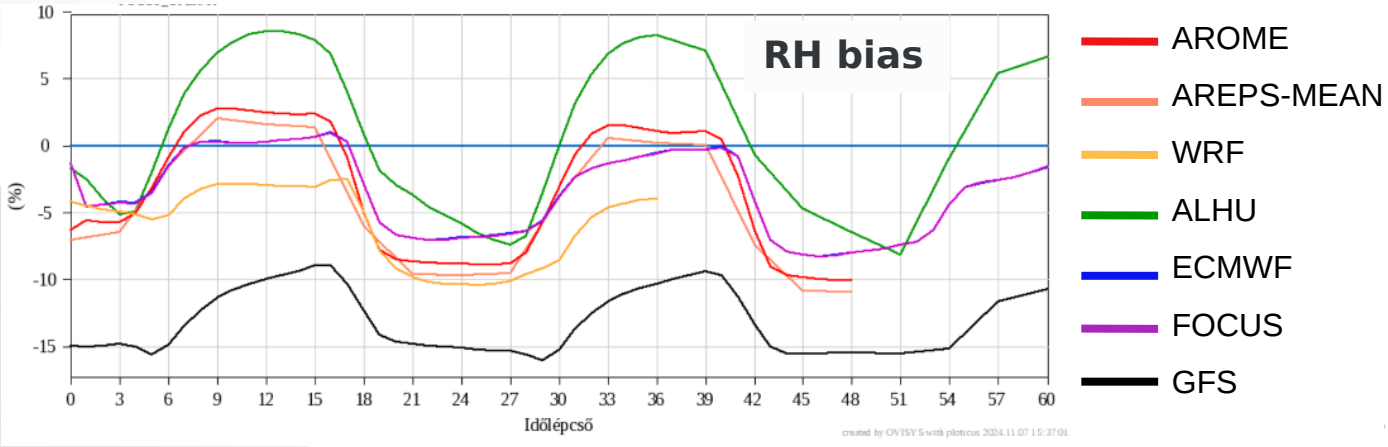
# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

- **Pontszerű verifikáció**

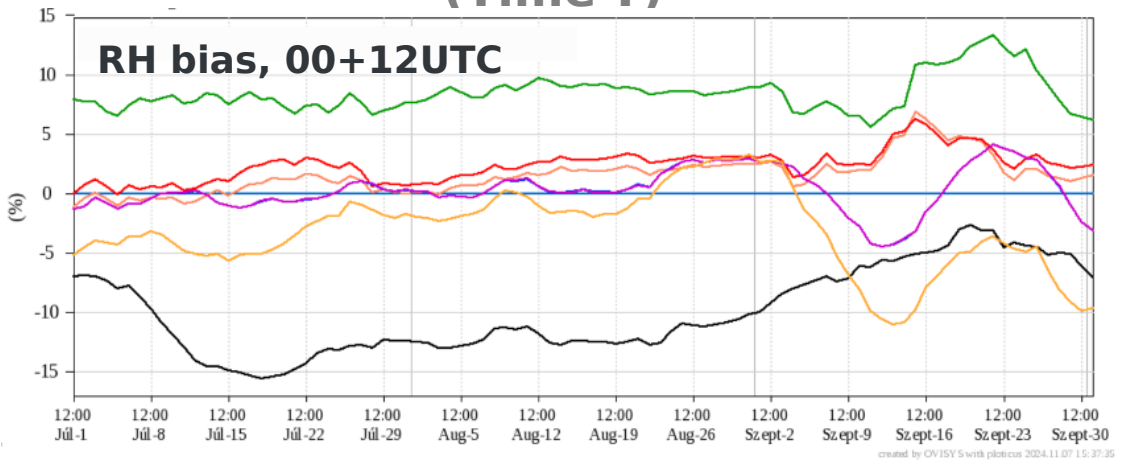
X tengelyen az időlépcső szerepel

**(Time-TS)**




X tengelyen a dátum szerepel

**(Time-T)**



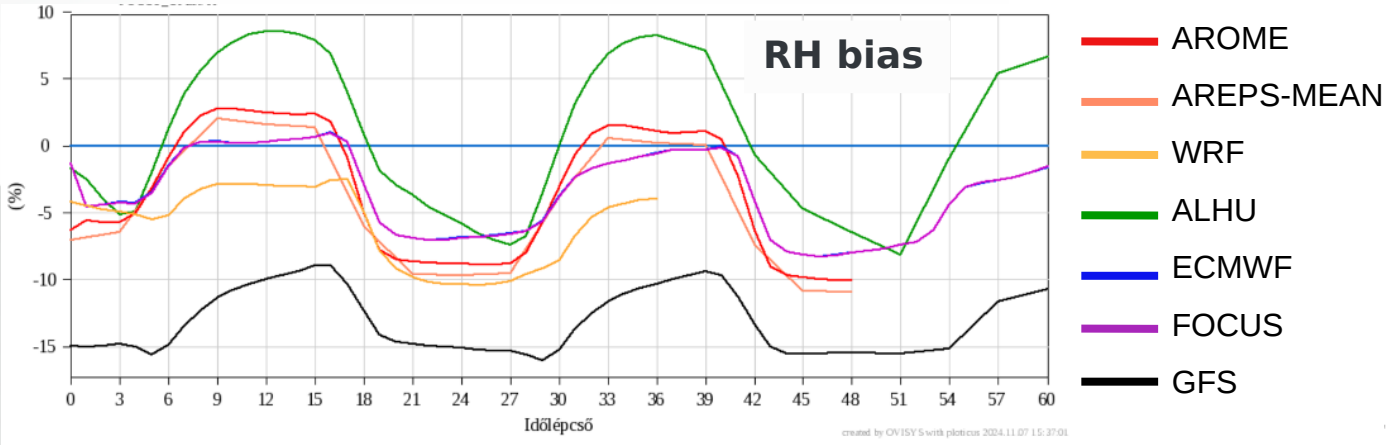
# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

- **Pontszerű verifikáció**

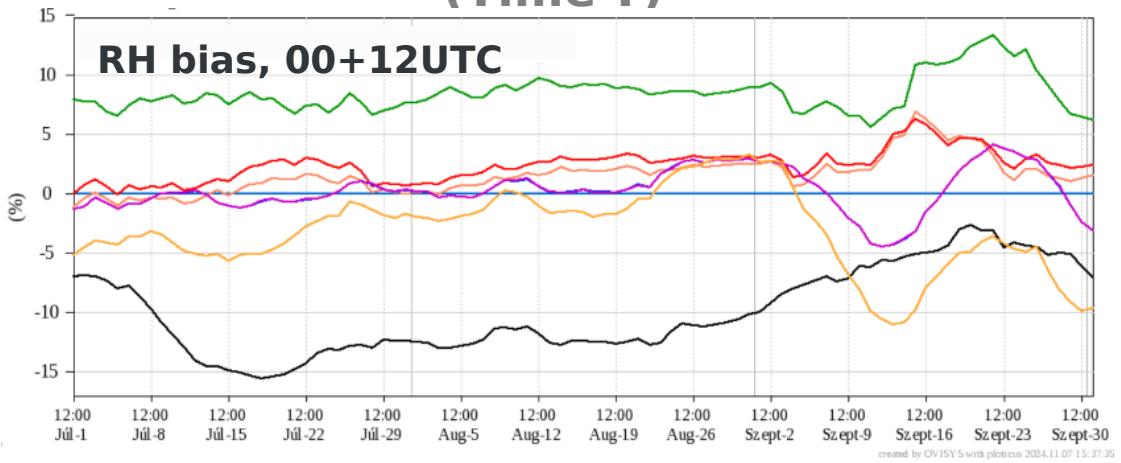
X tengelyen az időlépcső szerepel

**(Time-TS)**




X tengelyen a dátum szerepel

**(Time-T)**




### Mikor használjuk?

Napi menet változása érdekel

- Melyik napszakban jelentkezett a legnagyobb eltérés az előrejelzés és a megfigyelés között

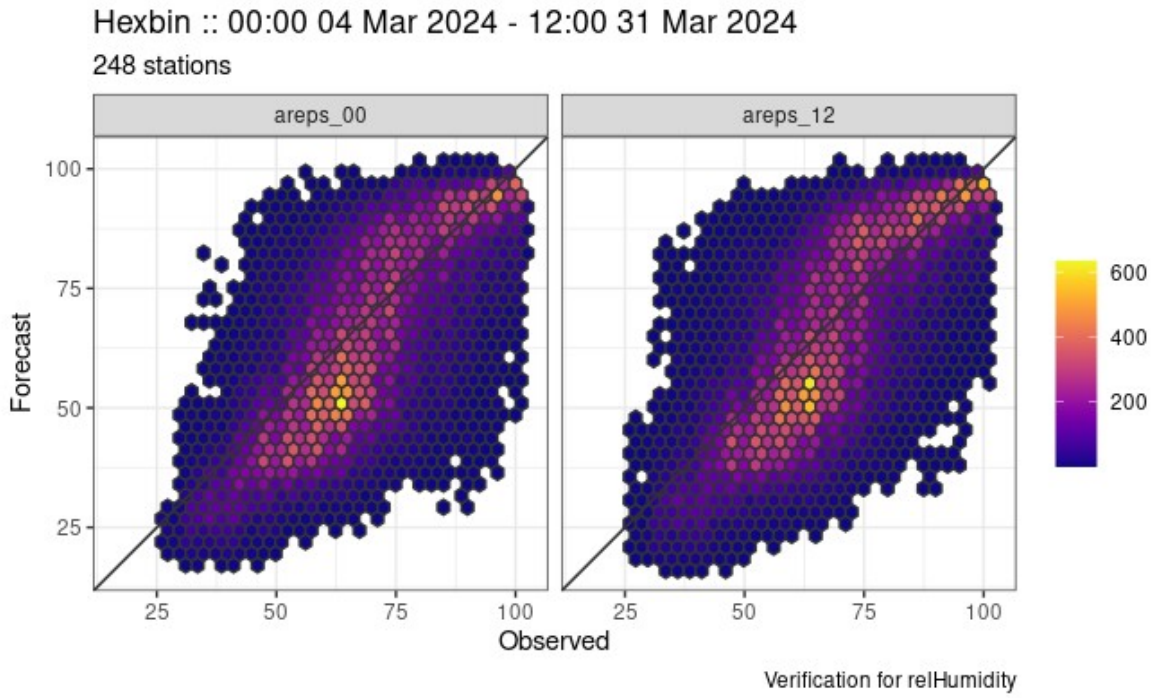
Egy adott időpontban érdekel egy hosszab időszakban bekövetkezett változás

- A relatív nedvesség 12 órás előrejelzése hogyan változott a július, augusztus és szeptember folyamán

# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

- **Pontszerű verifikáció**
- X tengelyen az megfigyelés, Y tengelyen az előrejelzés értéke szerepel → **Scatterplot**



# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

- **Pontszerű verifikáció**
- Amennyiben diszkrét előrejelzést használunk (igen-nem) → **Kontingencia táblázat**

| Determinisztikus előrejelzésre vonatkozó KONTINGENCIA TÁBLÁZAT 'n' db bináris esemény esetén | | **Megfigyelés** *(Event observed)* | | **Összes** |
|---|---|---|---|---|
| | | **Igen** *(Az esemény bekövetkezett)* | **Nem** *(Az esemény nem következett be)* | |
| **Előrejelzés** *(Event forecast(ed)* | **Igen** *(Volt előrejelzés az eseményre)* | TALÁLAT *(Hits)* **[a]** | TÉVES RIASZTÁS *(False alarms)* **[b]** | a+b |
| | **Nem** *(Nem volt előrejelzés az eseményre)* | HIBÁS ELVETÉS *(Misses)* **[c]** | HELYES ELVETÉS *(Correct rejections)* **[d]** | c+d |
| **Összes** | | a+c | b+d | n = a+b+c+d |

| Előrejelzés \ Megfigyelés | <=10.0 | >10.0 | Total |
|---|---|---|---|
| <=10.0 | 1693 | 91 | 1784 |
| >10.0 | 26 | 36 | 62 |
| Total | 1719 | 127 | 1846 |

- Volt-e 10mm-nél nagyobb csapadék az AROME modellben?
- Cél: találatok és hibák statisztikai értékelése

- Amennyiben kategorikus előrejelzést használunk → **Küszöbérték ábra**
- **Frequeny bias (FBI) =** (a+b)/(a+c)
- megadja, hogy mennyivel többször jeleztük előre az eseményt, mint ahogy az bekövetkezett. Tökéletes értéke = 1
- **Probability of detaction (POD) =** a/(a+c)
- megadja, hogy a bekövetkezett események közül mekkora részt jeleztünk előre. Tökéletes értéke = 1

# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

### Pontszerű verifikáció

Amennyiben kategorikus előrejelzést használunk → **Küszöbérték**

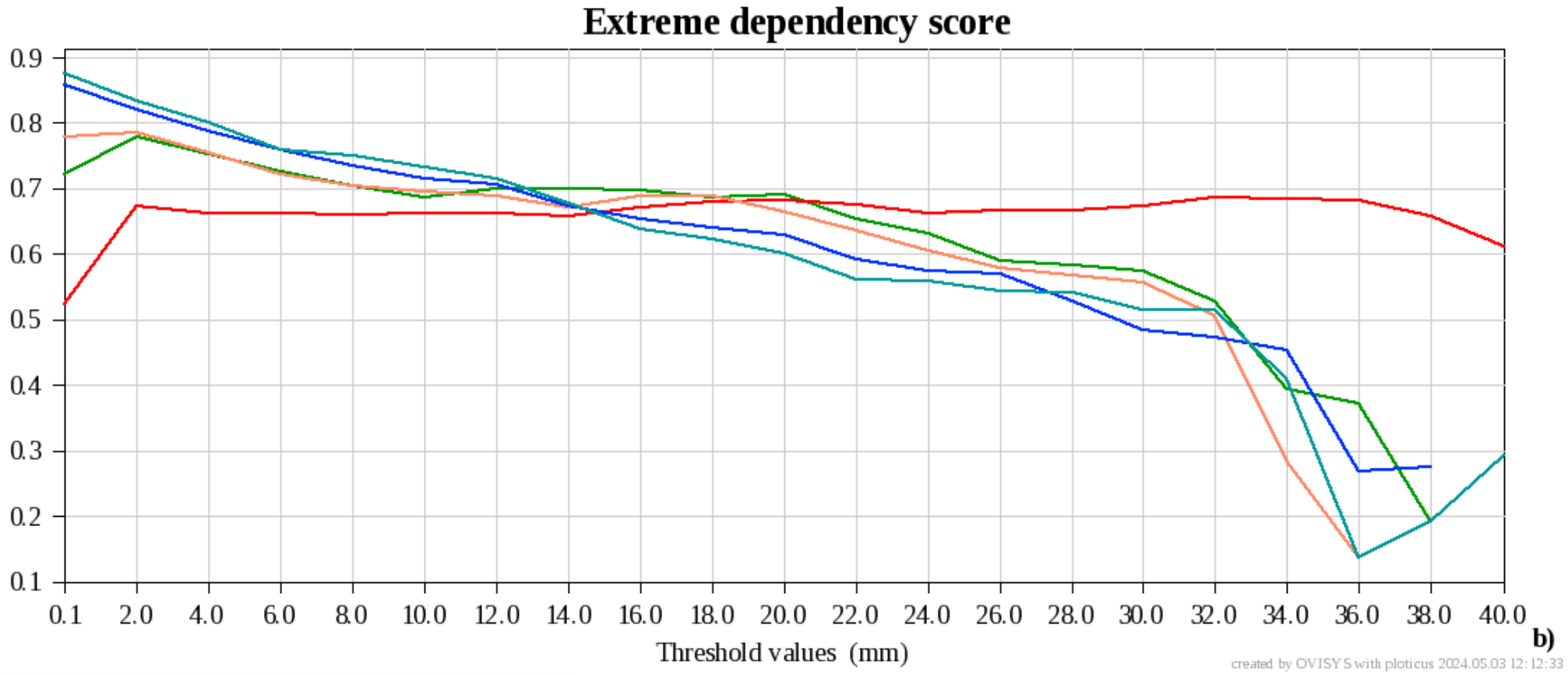
- **False Alarm Ratio (FAR) =** b/(a+b)
- megadja, hogy amikor volt előrejelzés, akkor azoknak az esetnek mekkora hányada volt téves, azaz az esemény nem következett be – más szóval az előrejelzések mekkora része volt téves riasztás. Tökéletes értéke = 0.

- **Extremitás indexek**
- **Extreme Dependency Score (EDS) =** (ln(p)-ln(H)) / (ln(p)+ln(H))
- ahol p = (a+c)/n, H = a/(a+c), n = a+b+c+d
- Egy extrém csapadékindex, amely sokkal nagyobb hangsúlyt fektet a megfigyelt eseményekre. A tökéletes értéke = 1, tehát minnél magasabb az EDS értéke, annál jobb az esemény észlelésének valószínűsége.
-

# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

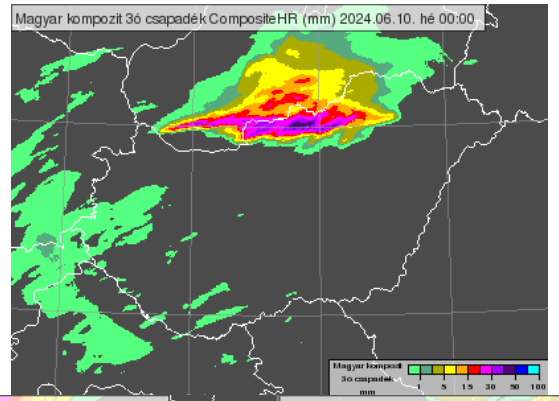
- **Extremitás indexek**
- **Extreme Dependency Score (EDS) → 2023-as eredmény, Csapadék24h, 00+30h**

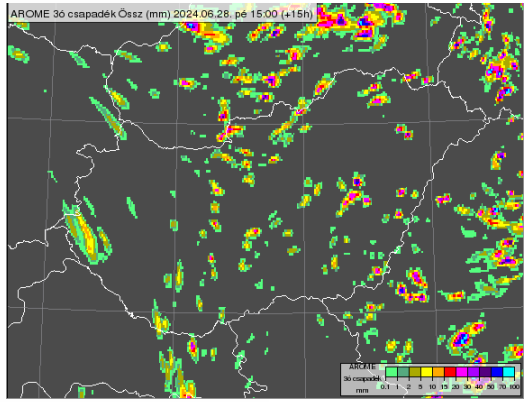


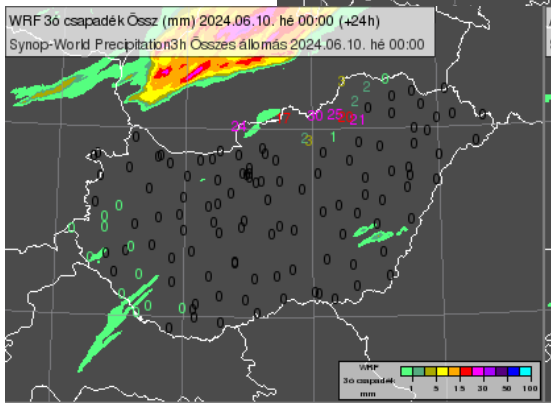
# Objektív verifikáció

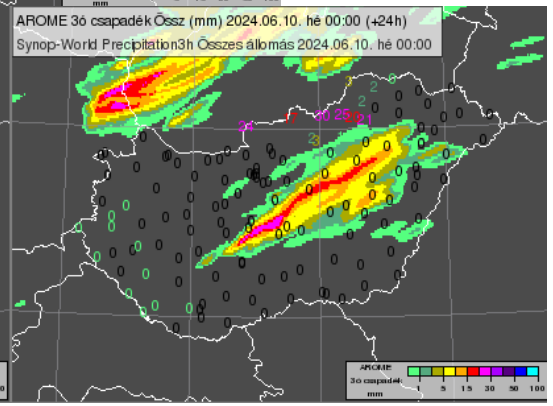
## Statisztikai mérőszámokat származtatunk

- **Térbeli verifikáció**

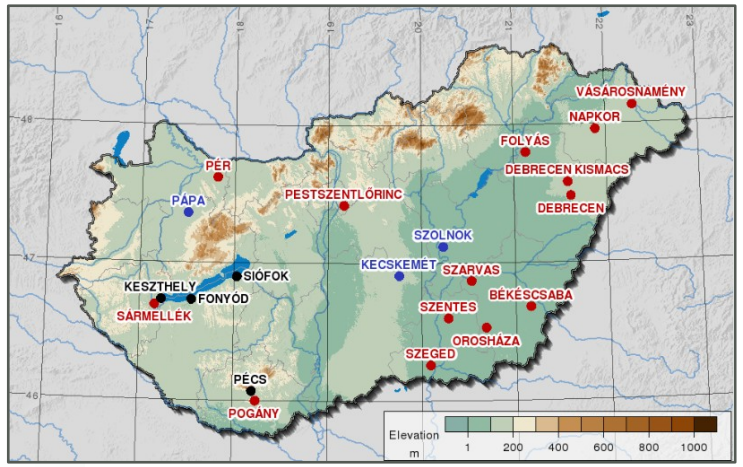


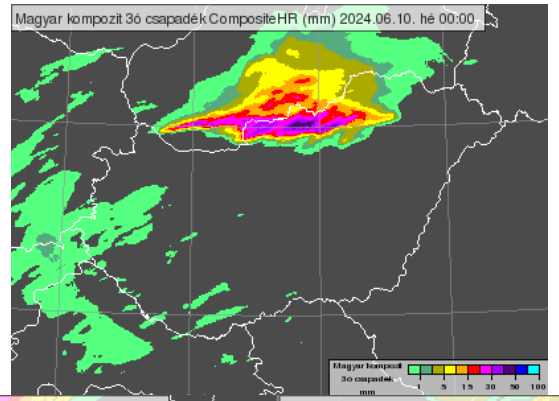

- felhőzet megfigyelési pontok



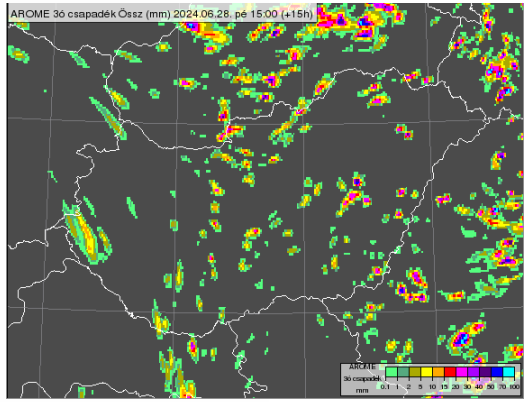
# Objektív verifikáció

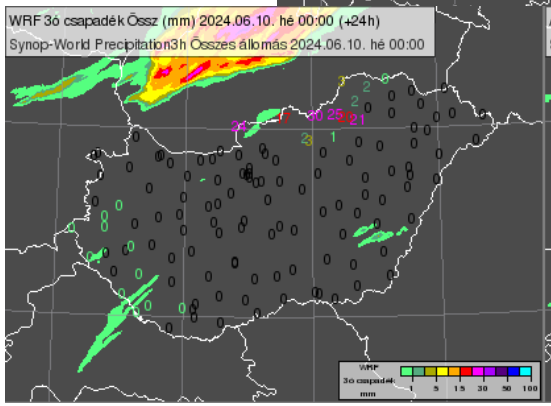
## Statisztikai mérőszámokat származtatunk

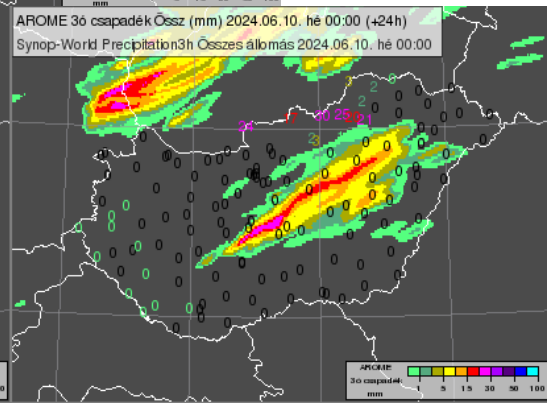
- **Térbeli verifikáció**



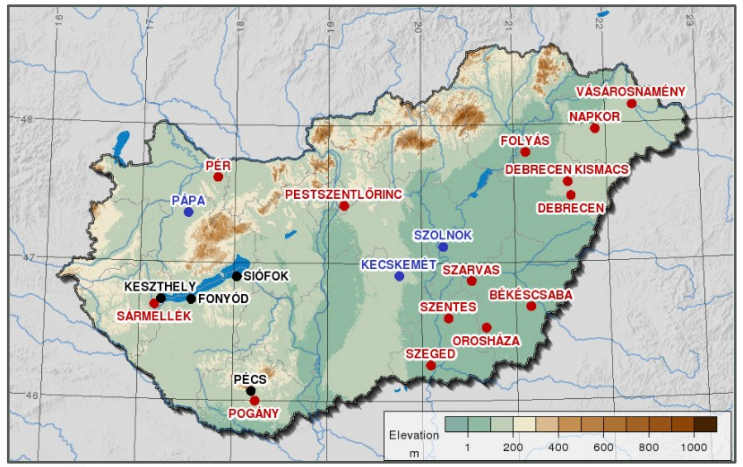

- **→ Kettős büntetés →**




- **→ Kevés megfigyelés →**
- felhőzet megfigyelési pontok



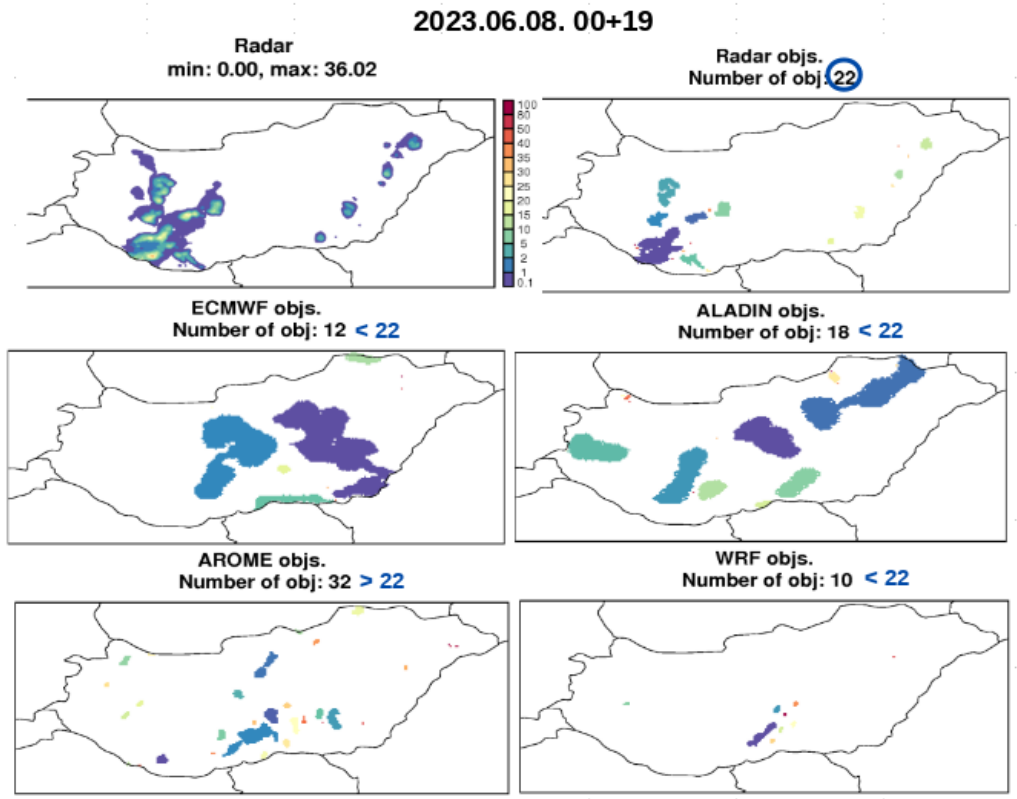

**Van megoldás!**

SAL, neighborhood technikák

# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

- **SAL (Structure, Amplitude, Location)**
  - Objektív modell verifikáció radaros csapadékösszegek használatával
  - Objektum alapú → dinamikus küszöb (Pmax* 1/15)
    rögzített küszöb (1mm)
  - Minden olyan radar pixel, ami meghaladja a csapadékmaximum 1/15-dét → 1 objektum




S = normalizált csapadék objektumok térfogatának különbsége [-2,...,**0**,...2]

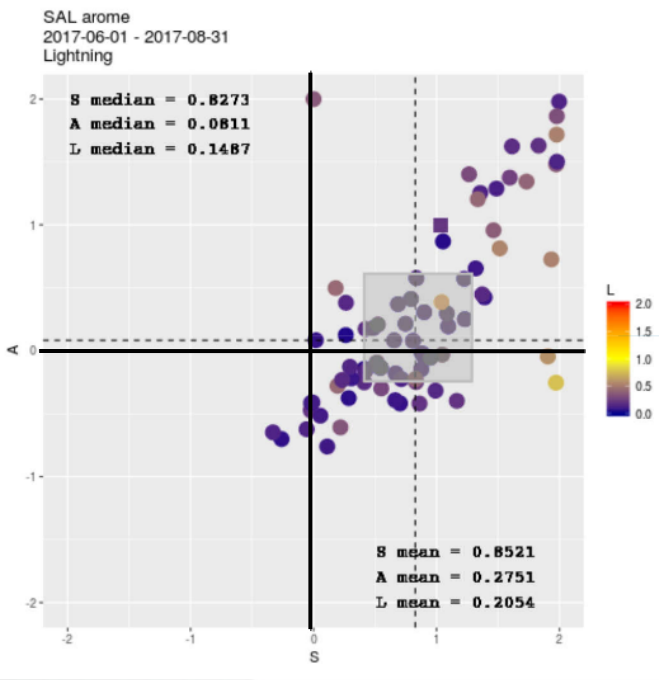
A = modell és radaros területi csapadékátlag normált különbsége [-2,...**,0,**...2]

L = objektumok súlypontjának normált távolsága + objektumok normált területi szórása [**0**,...,2]

# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

- **SAL (Structure, Amplitude, Location)**

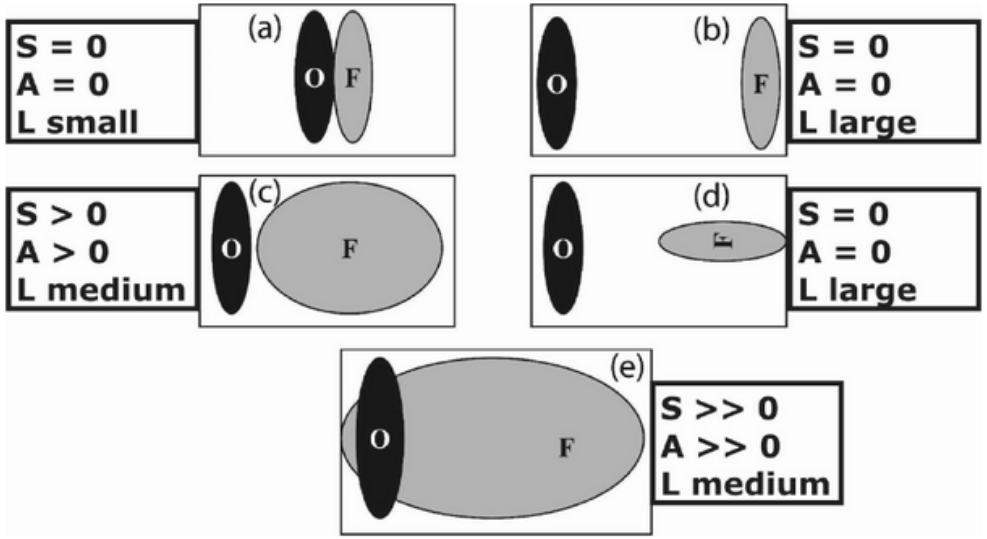



Modell felülbecslés

Location error

Modell objektum felülbecslés

- Wernli, H., Paulat, M., Hagen, M., & Frei, C. (2008). SAL—A novel quality measure for the verification of quantitative precipitation forecasts. Monthly Weather Review, 136(11), 4470-4487.




- Középponti statisztika →
- Az a sugár, amelyen belül a SAL ábrán a pontok X%-a elhelyezkedik

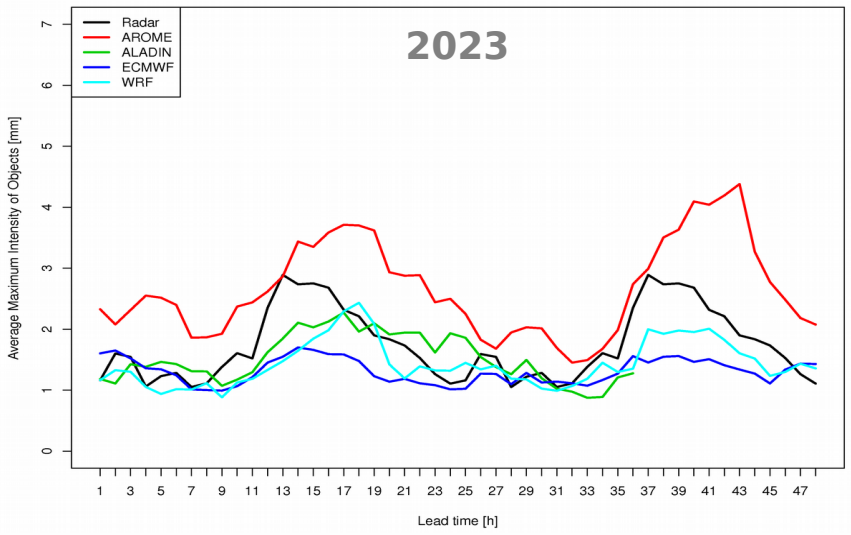
| | 5% | | 10% | | 20% | | 50% | |
|---|---|---|---|---|---|---|---|---|
| | OPER | TEST | OPER | TEST | OPER | TEST | OPER | TEST |
| 00 UTC | 0.44 | 0.36 | 0.56 | 0.49 | 0.74 | 0.68 | 1.32 | 1.20 |
| 06 UTC | 0.44 | 0.39 | 0.60 | 0.54 | 0.82 | 0.71 | 1.45 | 1.29 |
| 12 UTC | 0.44 | 0.38 | 0.57 | 0.52 | 0.81 | 0.74 | 1.32 | 1.36 |

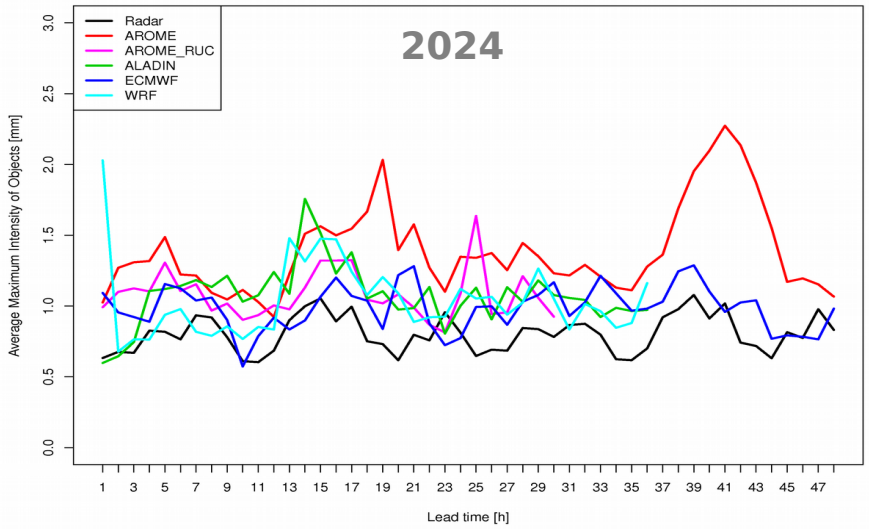
# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

- **SAL (Structure, Amplitude, Location)**
  - Egyéb csapadék karakterisztikák származtathatók
    - Területi átlag
    - Objektumok átlagos intenzitása
    - Legerősebb három objektum átlagos intenzitása

Objektumok átlagos maximum intenzitása



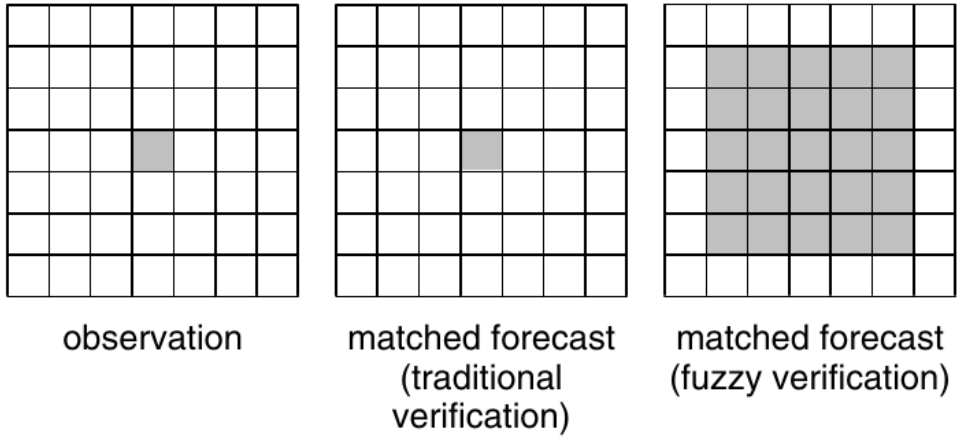
# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

- **Neighborhood technika**

  A módszerrel a kisebb tér- és időbeli hibákat eltérő módon értékelik, ugyanis a korrekt egyezés helyett a közelséget is jutalmazzák, az előrejelzési skálán hasznosnak ítélt térbeli ablakokon keresztül.

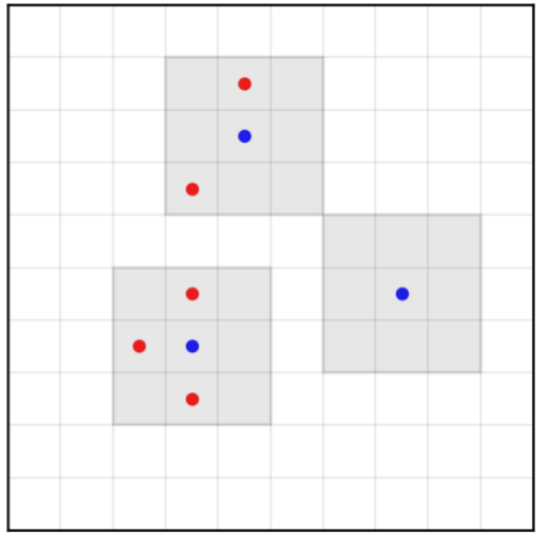
Ebert, E. E. (2008). Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling, 15(1), 51-64.




**HiRA – High Resolution Assessment**

Single observation – neighborhood forecast

- Egyetlen megfigyelt pontot hasonlítunk össze a környezetében lévő előrejelzett rácspontokkal.
- Az előrejelzés és a megfigyelés közötti kapcsolatot különböző térbeli méretekben aggregáljuk, az előrejelzési mező szomszédos pontjait felhasználva (küszöbérték).
- Pontossági metrikákat származtatunk
- Egyedi megfigyelések pontosságát hangsúlyozza → lokális



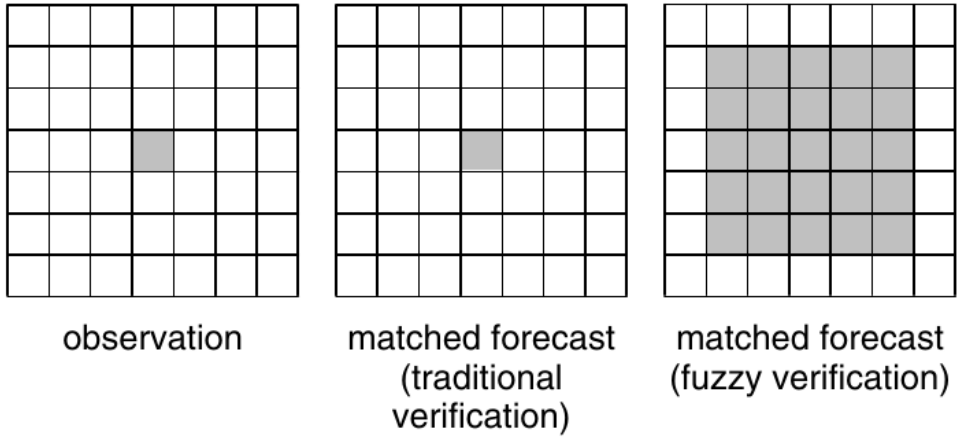
# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

- **Neighborhood technika**

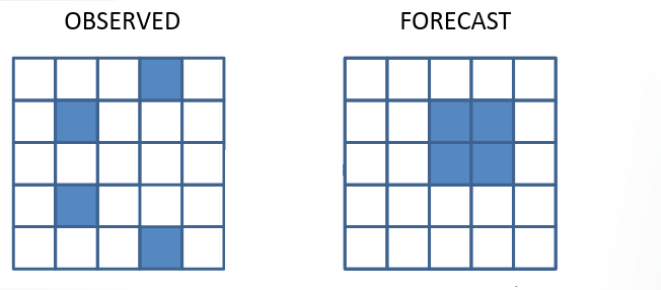
  A módszerrel a kisebb tér- és időbeli hibákat eltérő módon értékelik, ugyanis a korrekt egyezés helyett a közelséget is jutalmazzák, az előrejelzési skálán hasznosnak ítélt térbeli ablakokon keresztül.

Ebert, E. E. (2008). Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling, 15(1), 51-64.




### FSS (Fraction Skill Score)

- a küszöbértéket meghaladó pontok hány százaléka van jelen az ablakban
- Megadja, hogy az előrejelzés mennyivel jobb a teljesen véletlenszerű eloszláshoz képest.




Korábbiakban tanult MSE alkalmazandó:

$f_i$= előrejelzés frakcionális lefedettsége
$o_i$= megfigyelés frakcionális lefedettsége
n = összes vizsgált pont az ablakon belül

$$\mathbf{FSS = 1 - (MSE / MSE_{ref})}$$

$$\mathbf{MSE_{ref} = \sum (f_i^2 + o_i^2)}$$

- Nevező a lehetséges maximális hiba
- Tökéletes értéke = 1

# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

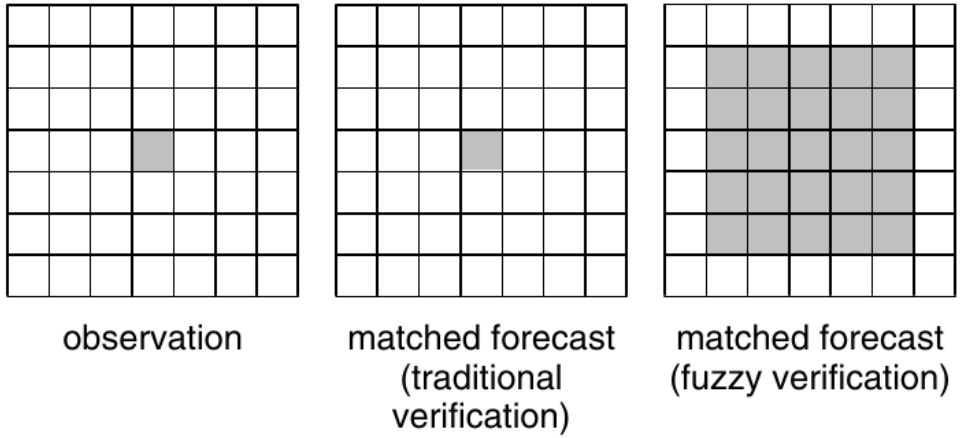
- **Neighborhood technika**

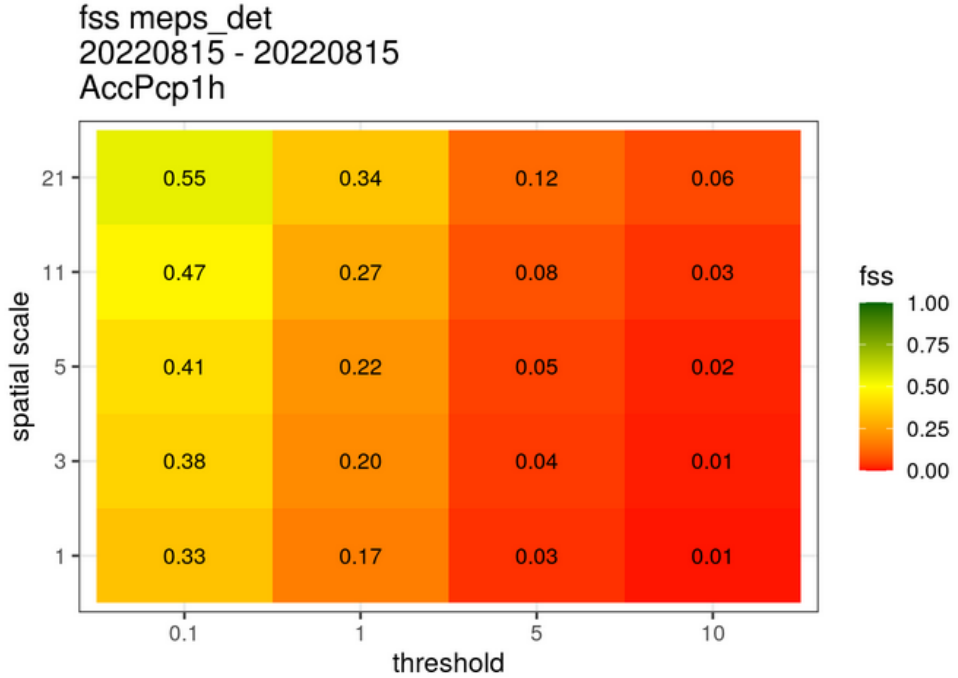
  A módszerrel a kisebb tér- és időbeli hibákat eltérő módon értékelik, ugyanis a korrekt egyezés helyett a közelséget is jutalmazzák, az előrejelzési skálán hasznosnak ítélt térbeli ablakokon keresztül.

### FSS (Fraction Skill Score)

- a küszöbértéket meghaladó pontok hány százaléka van jelen az ablakban
- Megadja, hogy az előrejelzés mennyivel jobb a teljesen véletlenszerű eloszláshoz képest.

Ebert, E. E. (2008). Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling, 15(1), 51-64.



# Objektív verifikáció

## Statisztikai mérőszámokat származtatunk

- **FSS (Fraction Skill Score)**

## Összetett verifikációs keretrendszer

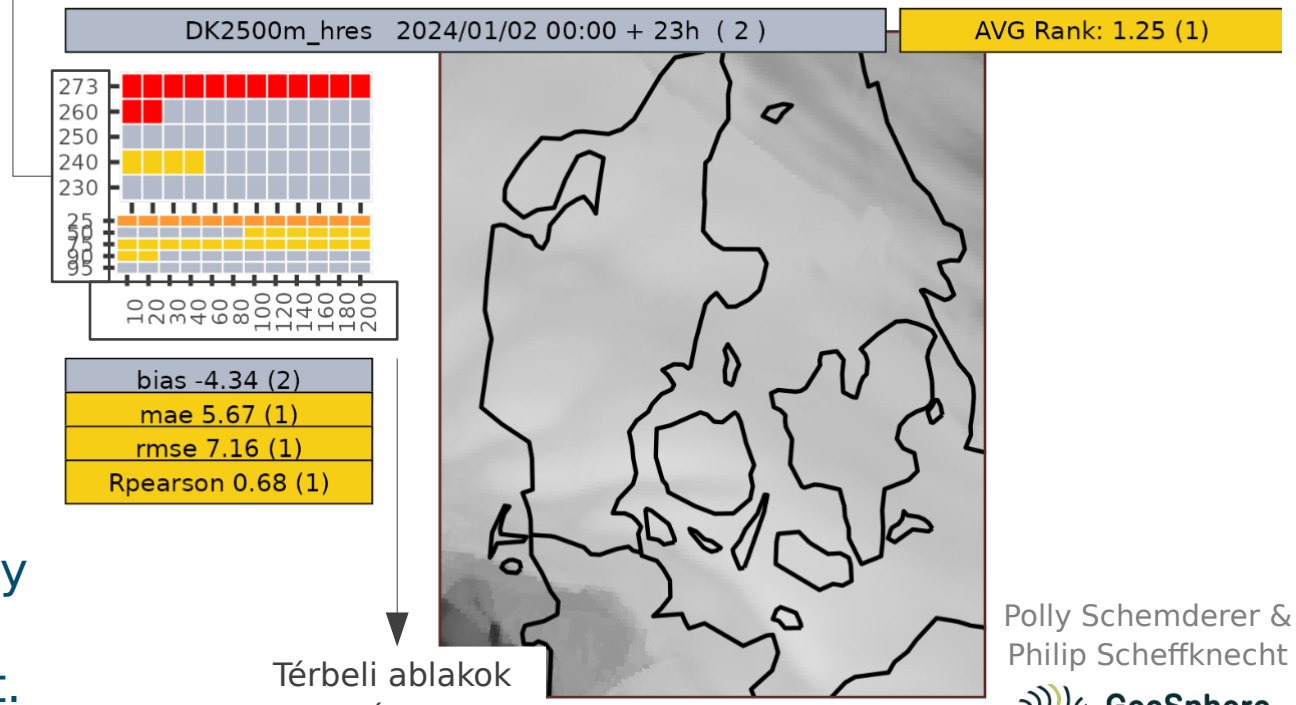
- **Panelifikáció**
  - Eltérő konfigurációk szerinti diverz összehasonlítást tesz lehetővé
  - Több verifikációs metrika, grafikon összehangolt paneles megjelenítése
  - Rangsorolás jellegű
  - Megteremti a diagnosztikai következtetések alapját azáltal, hogy rámutat a hibák struktúrájára és segít megérteni melyik modell, mely futás, mely küszöbérték és mely térbeli ablak mellett mutat eltérést a megfigyelésekhez képest.

1. hely
2. hely
3. hely
< 0.5 FSS

Összes metrika alapján (kivéve FSS) összesített rangsor!

Átalgos FSS teljesítmény rangsor

Percentilisek, küszöbértékek




Térbeli ablakok mérete

Polly Schemderer & Philip Scheffknecht
GeoSphere Austria

# “Determinisztikus” előrejelzések verifikációja

## Szubjektív verifikáció

# Szubjektív verifikáció

## Szakértői ítélőképesség alapján végzett értékelés

- Térképek vizuális elemzésével készül (HAWK) modellfejlesztő és előrejelző szakemberek által
- **Mikor használjuk?**
  - Érdekesnek vélt időjárási helyzetek elemzésénél
  - Egy adott időjárási jelenség térbeli struktúrájának vizsgálata a cél → lokális információk
  - Objektív mérőszámok kiegészítéseként, többletinformációt nyújt
  - Emberi hozzáadott érték többlettudásként jelentkezik
  - Operatív helyzetben
- **Mi a hátránya?**
  - Szubjektív → nagyban függ az értékelő szemléletmódjától
  - Időigényes, nagy mennyiségű adat feldolgozhatatlan
- **Hogyan zajlik?**
  - Érdekesnek vélt nap kiválasztása → HAWK makrók elemzése → modellek osztályozása

Prognózis megbeszélések, visszajelzések alapján

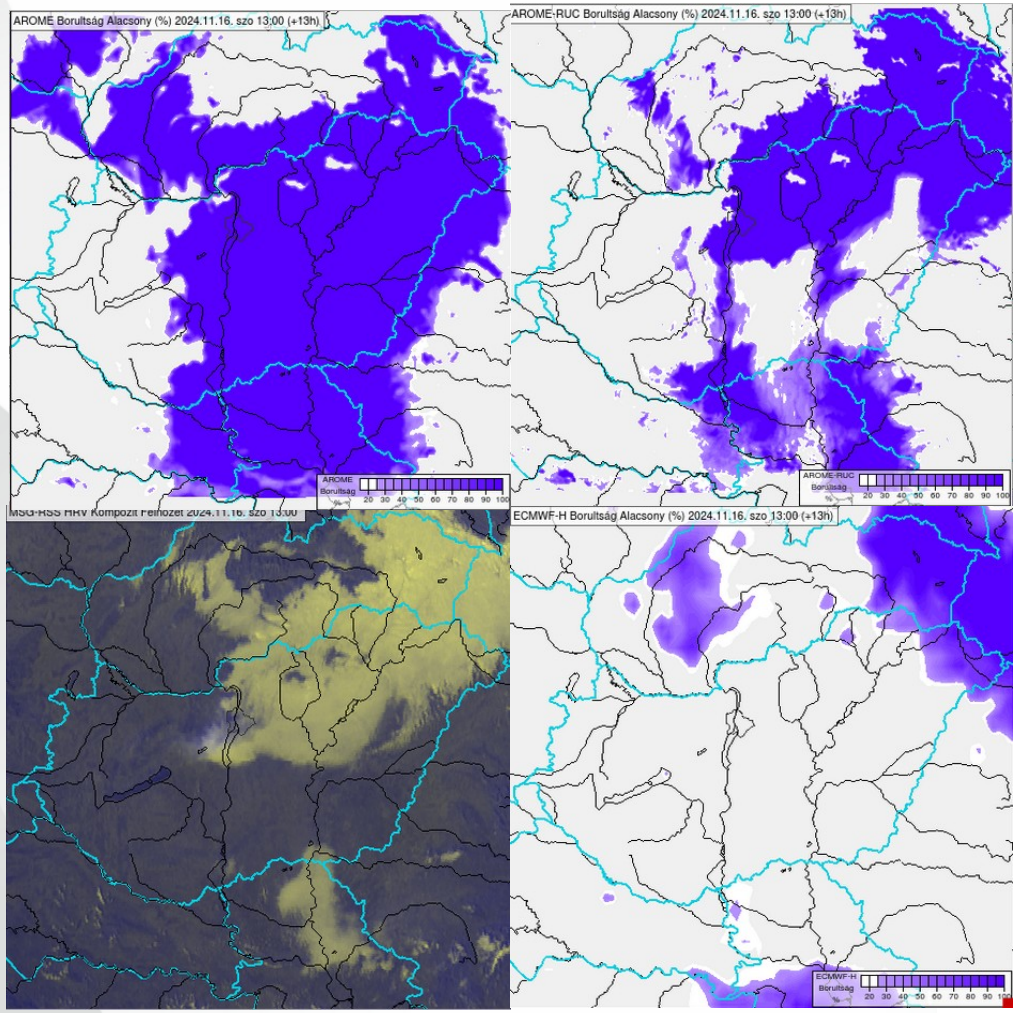
azonos színskála használat

(1-5)

# Szubjektív verifikáció

## Szakértői ítélőképesség alapján végzett értékelés

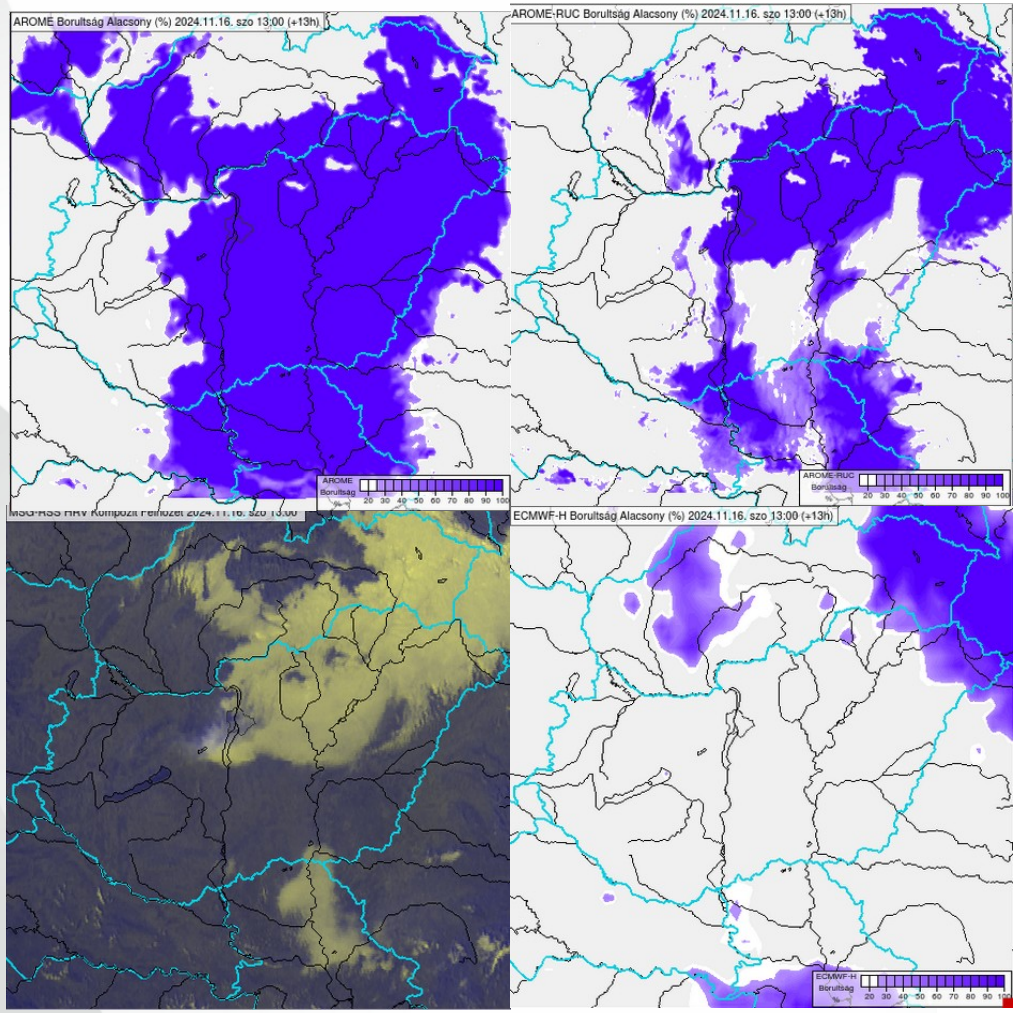
### 2024.11.16. 00+13h Felhőzet



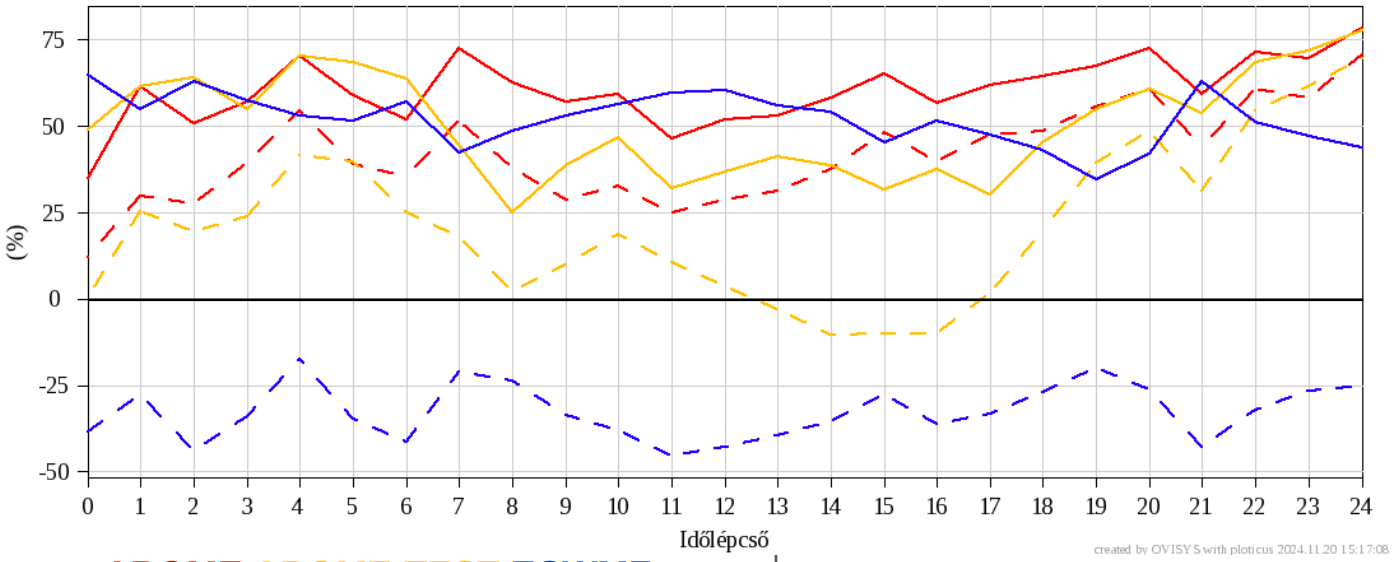
# Szubjektív verifikáció

## Szakértői ítélőképesség alapján végzett értékelés

**2024.11.16. 00+13h Felhőzet**




**RMSE, bias**




**AROME AROME-TEST ECWMF**

- Jól fogták a köd beragadását a modellek?

# Verifikáció a gyakorlatban

# Verifikáció a gyakorlatban

**Egy átfogó értékelés komplex feladat:**

**Fontos különböző verifikációs módszerek vegyítése**

**Mit szeretnénk verifikálni?** Számos paraméter esetén nem elegendő egy RMSE vagy bias, mert az átlagolás sokszor elkeni az extremitásokat, a részleteket. Célszerű ilyenkor küszöbérték ábrákat is megnézni, térbeli verifikációs score-ok és szubjektív értékelések mellett.

**Milyen időszakra?** Operatív döntéseket nem lehet 1-2 napos verifikációra alapozni. Hosszabb időszakot lefedő, akár több évszakot érintő verifikációra van szükség, hogy általános képet kapjunk egy modell viselkedéséről.

**Hogyan néz ki a gyakorlatban?**

**Objekítv értékelés (pontbeli, térbeli) + Szubjektív értékelés (modellezői, előrejelzői)**

- 1 vagy több hónapos verifikációs időszak → következtetések összefoglalása, értelmezése, átbeszélése → újabb teszt időszak meghatározása ...

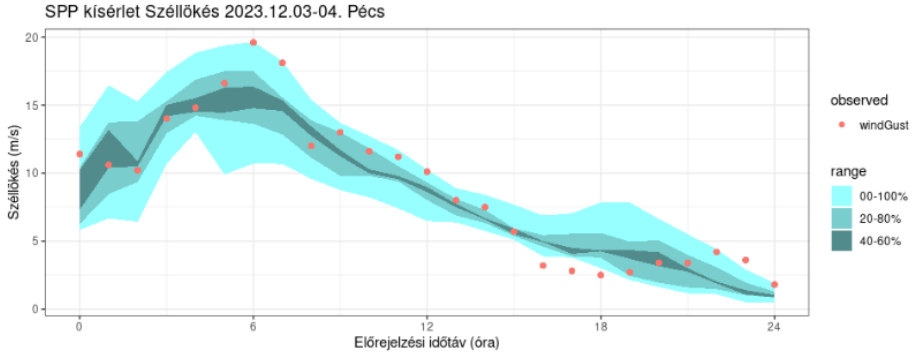
# Verifikáció a gyakorlatban

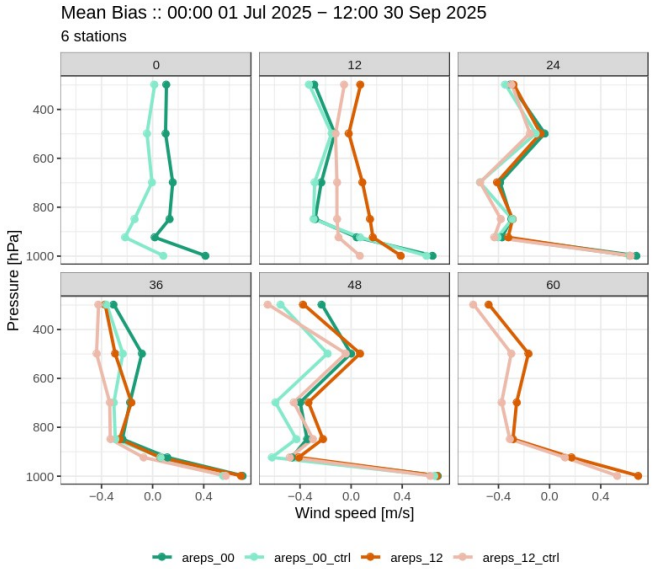
**Egy átfogó értékelés komplex feladat:**

**Fontos különböző verifikációs módszerek vegyítése**

## HungaroMet:

- **OVISYS:** sajátfejlesztésű verifikációs rendszer, kimondottan rácsponti verifikáció céljából
- **SAL**
- **HARP:** ACCORD konzorciumon belül fejlesztett verifikációs rendszer
- **Szubjektív verifikációs** munkacsoport (7 fő)




## Folyamatban lévő fejlesztések:

- Panelifikáció felhőzetre
- Belső igények mentén történő módszertani fejlesztések

# Köszönöm a figyelmet!