

Meteorológiai adatasszimiláció 1. rész

2025.11.03.

Kardos-Várkonyi Anikó

varkonyi.a@met.hu

A számszerű előrejelzés lépései

Adat-asszimiláció

mérési információk
gyűjtése, ellenőrzése,
modell rácsra történő
előállítás

Modell integrálás

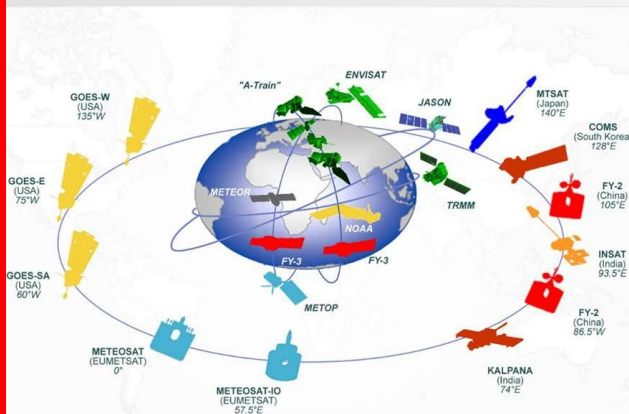
a légkör hidro-
termodinamikai
egyenletrendszerének
közelítő megoldása

Utófeldolgozás

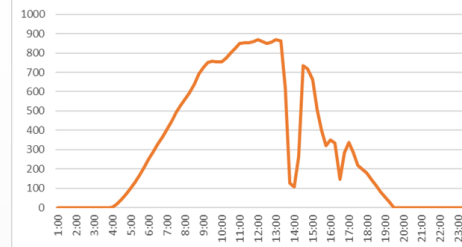
speciális paraméterek
származtatása,
megjelenítés

Verifikáció

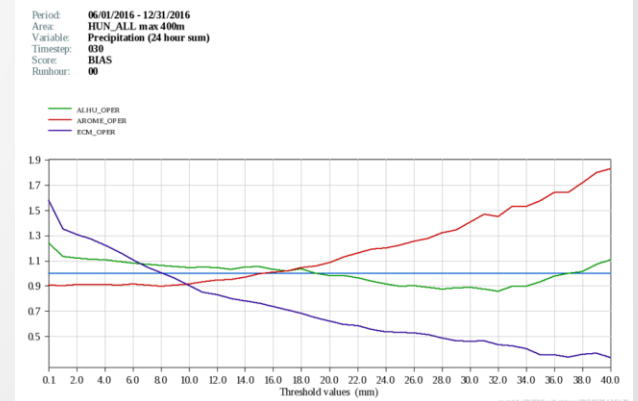
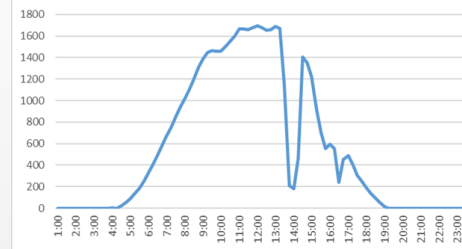
az előrejelzések
minőségének
vizsgálata



Globálisugrás előrejelzés (W/m²) - 2018.06.11.



Negyedórás teljesítmény (kW) - 2018.06.11.



Bevezetés

Mi az adatasszimiláció célja?

- Kezdeti feltételek biztosítása a numerikus modellek számára
→ Lényeges a pontosság a HTER megoldásának kezdeti feltételekre való érzékenysége miatt (a légkör kaotikus viselkedése)
- A bizonytalanságok meghatározása

Tehát az adatasszimiláció során analízist állítunk elő.

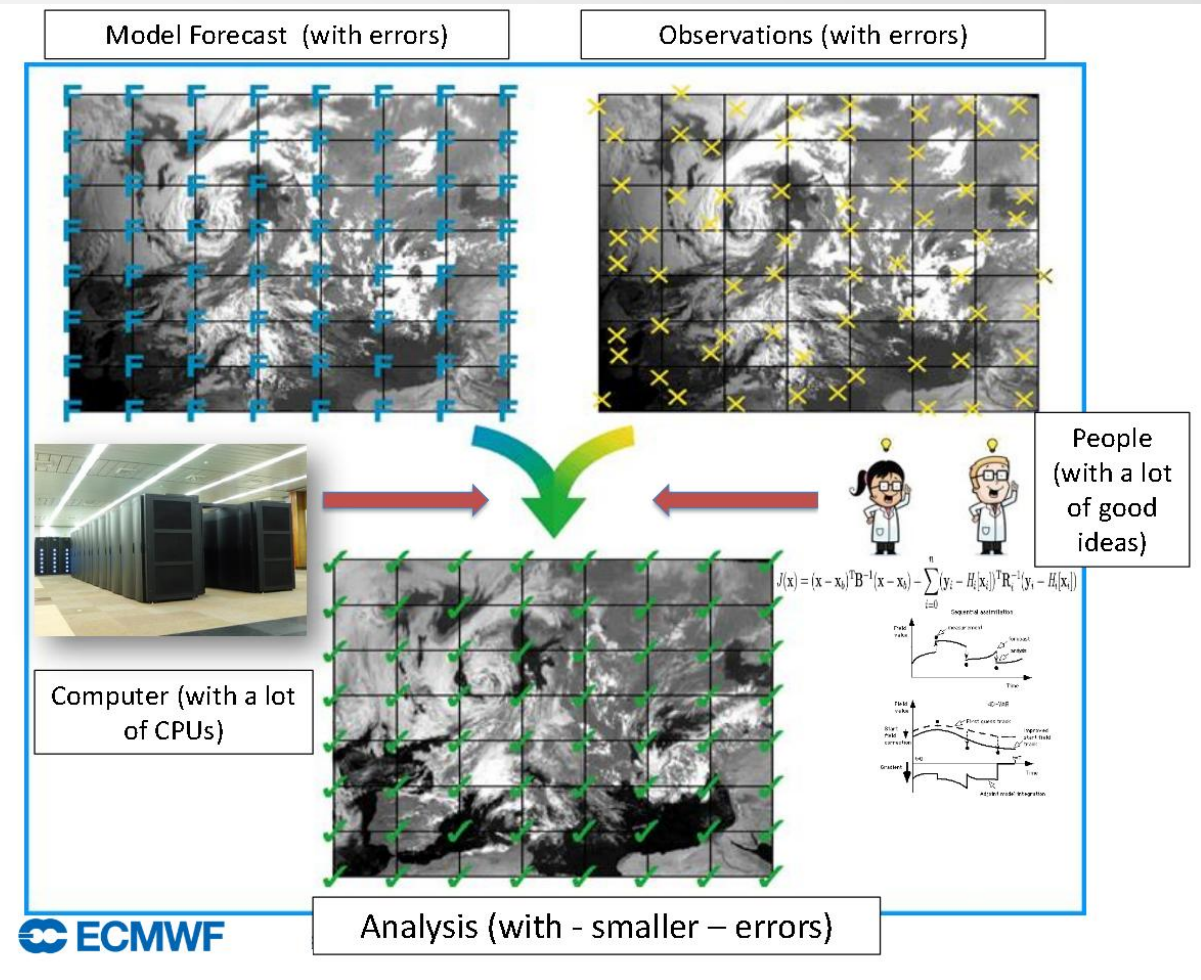
Analízis: Egy adott időpillanatban a légkör valós állapotáról a rendelkezésre álló információk alapján alkotott lehető legpontosabb kép

Mik szükségesek az adatasszimilációhoz?

- Modell előrejelzés (background)
- Megfigyelések
- Számítási kapacitás
- Emberi erőforrás

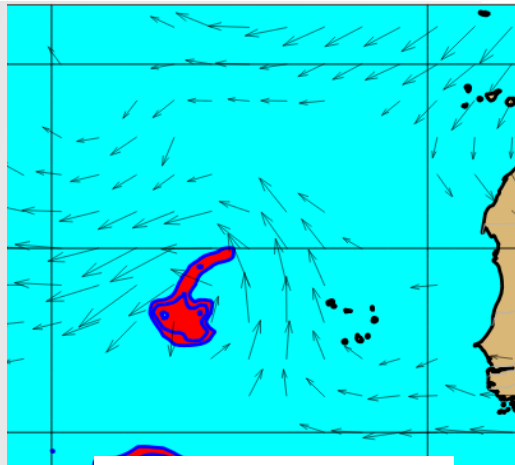
Mik a legnagyobb nehézségek?

- A megfigyelések nem a modell rácspontjaiban helyezkednek el + indirekt és gyakran bonyolult változók mérése → megfigyelési operátor
- A modell előrejelzés és a megfigyelések is hibával terheltek → optimálisan kell ezeket ötvözni



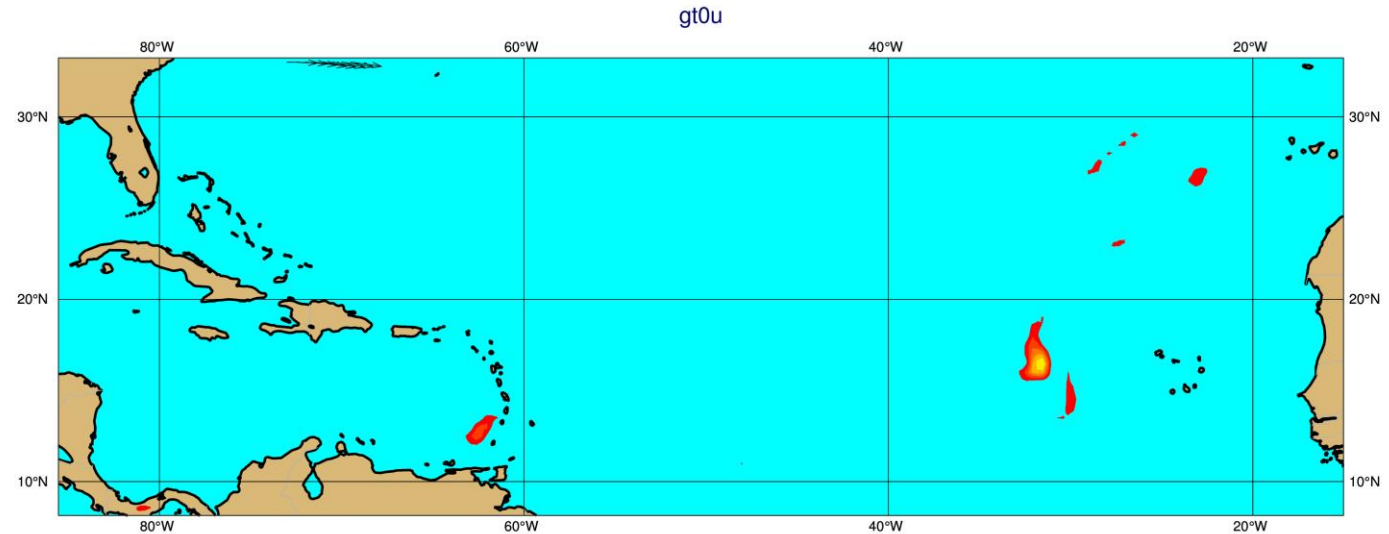
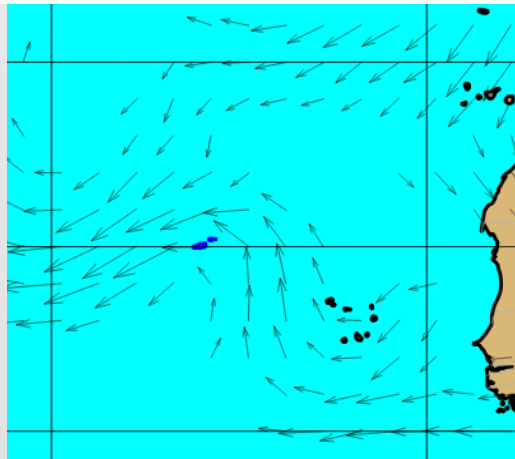
Az adatasszimiláció jelentősége

700 hPa relatív nedvesség & szél
kezdeti feltétel műholdas adatokkal

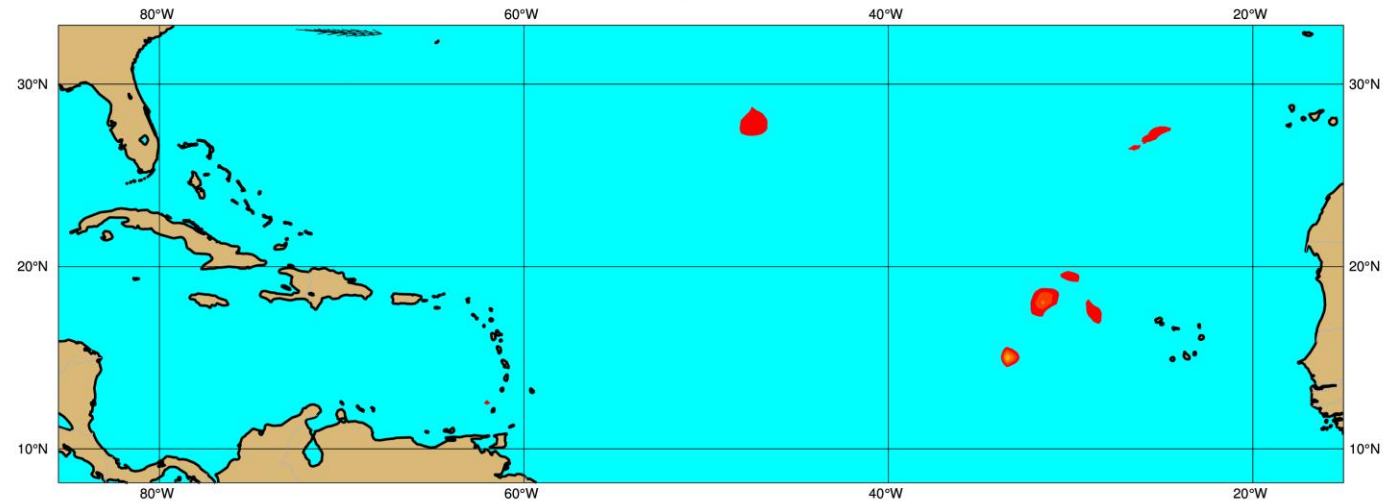


Nedvesség > 95%

700 hPa relatív nedvesség & szél
kezdeti feltétel műholdas adatok nélkül



Thursday 31 August 2017 00 UTC ecmf 500 hPa Vorticity (relative)
Thursday 31 August 2017 00 UTC ecmf 500 hPa U component of wind/V component of wind
gt0u



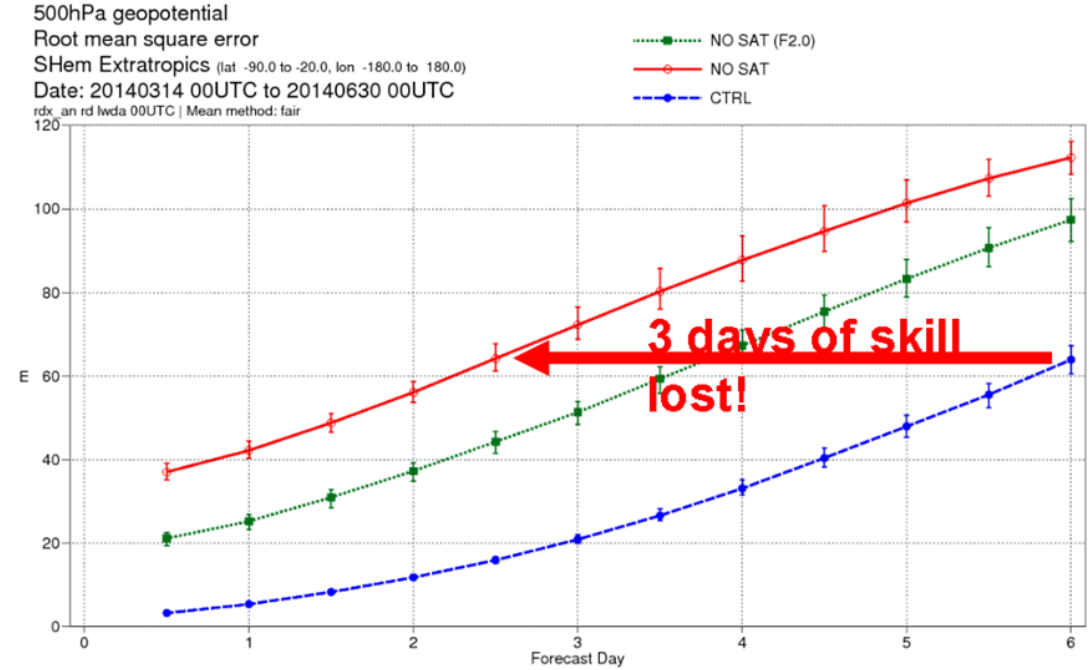
Az adatasszimiláció jelentősége

Műholdas mérések hatása

Északi félteke



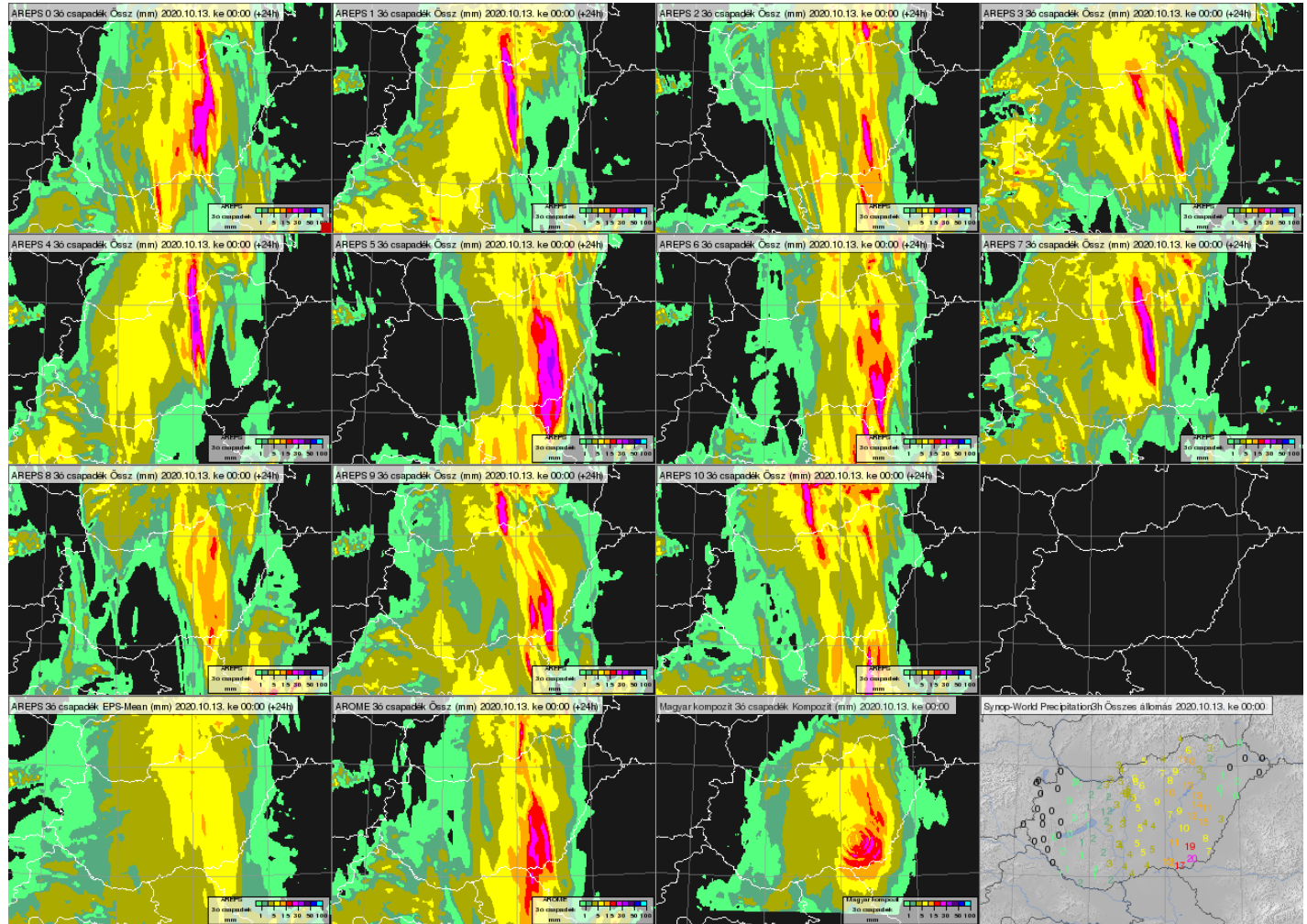
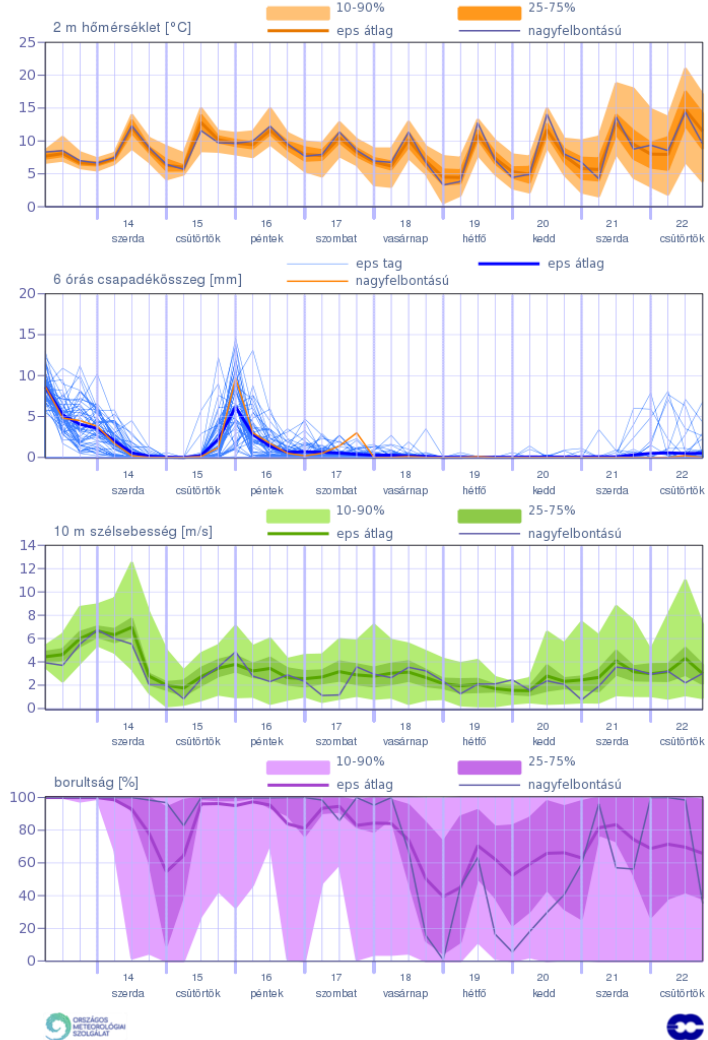
Déli félteke



Az adatasszimiláció jelentősége

ECMWF valószínűségi előrejelzés: Budapest

Készült: 2020.10.13 00 UTC-s futtatásból



Milyen igényeket támasztunk?

1. A kezdeti feltétel ugyanazon a **rácson** álljon rendelkezésre, mint amin az előrejelzést elkészítjük
2. Használjunk fel **minden** rendelkezésünkre álló **információt** – mik ezek az információk és hogyan használjuk fel ezeket?

Miért nem elegendő csak a méréseket felhasználnunk?

- A modell rácspontjaihoz képest kevesebb mérés áll rendelkezésre
- Nem elegendőek ahhoz, hogy megfelelő pontossággal az összes rácspontra meghatározzuk a modell prognosztikai változóit
- Alulhatározott lenne a feladat
- Még több információra van szükség (korábbi előrejelzések felhasználása)

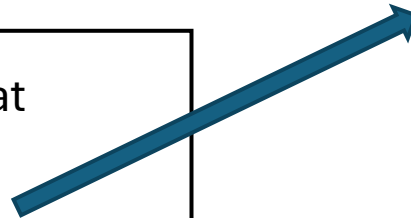
- Mérések és megfigyelések
- Korábbi előrejelzések
- Tudás a légkörről

Milyen igényeket támasztunk?

1. A kezdeti feltétel ugyanazon a **rácson** álljon rendelkezésre, mint amin az előrejelzést elkészítjük
2. Használjunk fel **minden** rendelkezésünkre álló **információt** – **mik ezek az információk és hogyan használjuk fel ezeket?**



- **Optimálisan** ötvözzük a rendelkezésre álló információkat (minimális becslési hiba)
- **Vegyük figyelembe** a felhasznált **információk hibáját** (ezeket is csak becsülni tudjuk)
- A becslés legyen **összhangban a HTER-rel** (dinamikai konzisztencia)



Milyen hibái lehetnek a felhasznált információknak?

- Műszerek hibái
- Reprezentativitási hiba
- Átszámítási hibák
- Modell előrejelzések hibája



Mi történik ha nem megfelelő a dinamikai konzisztencia?

- Zajok megjelenése
- Modellinstabilitás

Idealizált eset

- 1 térbeli pont
- 1 változó: x
- x_a becslést adunk x_t -re (a valós állapotra)
- rendelkezésünkre állnak y_1 és y_2 megfigyelések ε_1 és ε_2 hibával terhelve:

$$y_1 = x_t + \varepsilon_1$$

$$y_2 = x_t + \varepsilon_2$$

$$x_a = \hat{x}_t = f(y_1, y_2)$$

Szeretnénk a terem hőmérsékletét megbecsülni a jelen pillanatban – T_t

Rendelkezésünkre áll két mérés:

1. Digitális hőmérő: T_1



2. Analóg hőmérő: T_2



A rendelkezésre álló információk hibája:

$$\varepsilon_1 = T_1 - T_t \quad \text{és} \quad \varepsilon_2 = T_2 - T_t$$

Idealizált eset

Feltesszük, hogy:

- a mérések torzítatlanok (nincs szisztematikus hiba):

$$E(\varepsilon_1) = E(\varepsilon_2) = 0$$

- a becslések (mérések) hibáinak szórásnégyzete ismert:

$$\sigma_1^2 = E(\varepsilon_1^2) \quad \sigma_2^2 = E(\varepsilon_2^2)$$

- a becslések (mérések) hibái korrelálatlanok:

$$E(\varepsilon_1 \varepsilon_2) = 0$$

Idealizált eset

Kérdés:

$$x_a = f(y_1, y_2) = ?$$

Több megközelítés létezik, például:

- Legkisebb négyzetek módszere
- Maximum likelihood módszer

Legkisebb négyzetek módszere

Az **analízist** a becslések lineáris kombinációjaként állítjuk elő:

$$\left. \begin{array}{l} y_1 = x_t + \varepsilon_1 \\ y_2 = x_t + \varepsilon_2 \\ x_a = k_1 y_1 + k_2 y_2 \end{array} \right\} \begin{array}{l} k_1 = ? \\ k_2 = ? \end{array}$$

Szeretnénk, hogy:

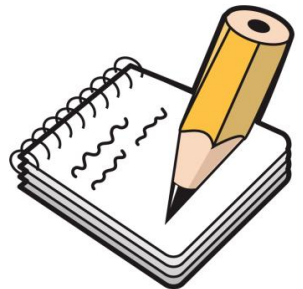
- a becslés torzítatlan legyen: $E(x_a - x_t) = 0 \longrightarrow k_1 + k_2 = 1$
- a becslés négyzetes hibája minimális legyen: $\sigma_a^2 = E((x_a - x_t)^2) = \min$

1. feladat:

$$\left. \begin{aligned} y_1 &= x_t + \varepsilon_1 \\ y_2 &= x_t + \varepsilon_2 \\ x_a &= k_1 y_1 + k_2 y_2 \end{aligned} \right\} \begin{aligned} k_1 &= ? \\ k_2 &= ? \end{aligned}$$

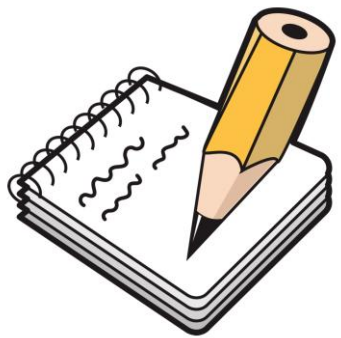
Mutassuk meg, hogy

$$E(x_a - x_t) = 0 \quad \Rightarrow \quad k_1 + k_2 = 1$$



1. feladat:

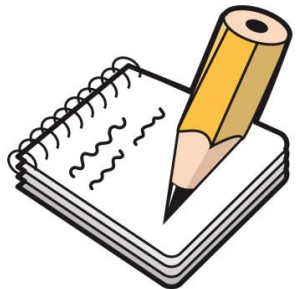
$$\left. \begin{aligned} y_1 &= x_t + \varepsilon_1 \\ y_2 &= x_t + \varepsilon_2 \\ x_a &= k_1 y_1 + k_2 y_2 \end{aligned} \right\}$$



2. feladat:
$$\left. \begin{array}{l} x_a = k_1 y_1 + k_2 y_2 \\ k_1 + k_2 = 1 \end{array} \right\} x_a = k_1 y_1 + (1 - k_1) y_2$$

Feladat: $\min \sigma_a^2(k_1)$

$$\left. \begin{array}{l} y_1 = x_t + \varepsilon_1 \\ y_2 = x_t + \varepsilon_2 \\ x_a = k_1 y_1 + k_2 y_2 \end{array} \right\}$$

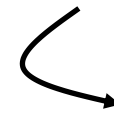


A becslés:

$$x_a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} y_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} y_2$$

A becslés megbízhatósága:

$$\frac{1}{\sigma_a^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} = \sum_{i=1}^p \frac{1}{\sigma_i^2}$$



Minden megfigyelés
növeli a
megbízhatóságot!

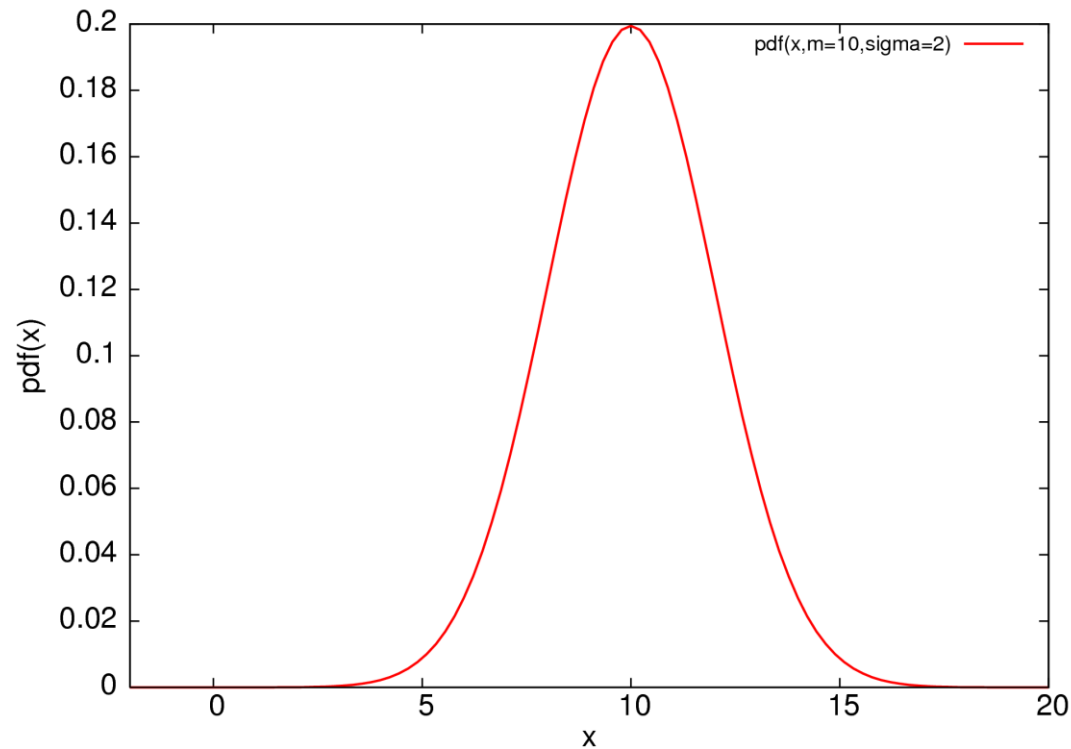


Maximum likelihood módszer

A meteorológiai változók „jól” modellezhetők normális eloszlású valószínűségi változókként

$$x \in N(m, \sigma^2)$$

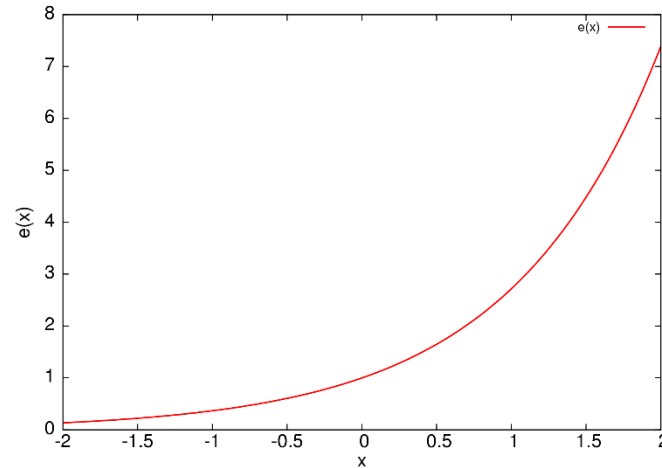
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$$



Az y_1 és y_2 mérések független becslést adnak x_t -re x várható értékkel, σ_1 és σ_2 szórással:

$$f_1(x) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{(x - y_1)^2}{2\sigma_1^2}\right)$$

$$f_2(x) = \frac{1}{\sigma_2 \sqrt{2\pi}} \exp\left(-\frac{(x - y_2)^2}{2\sigma_2^2}\right)$$



Becslés: x_a a minta együttes sűrűségfv-ének x szerinti maximuma

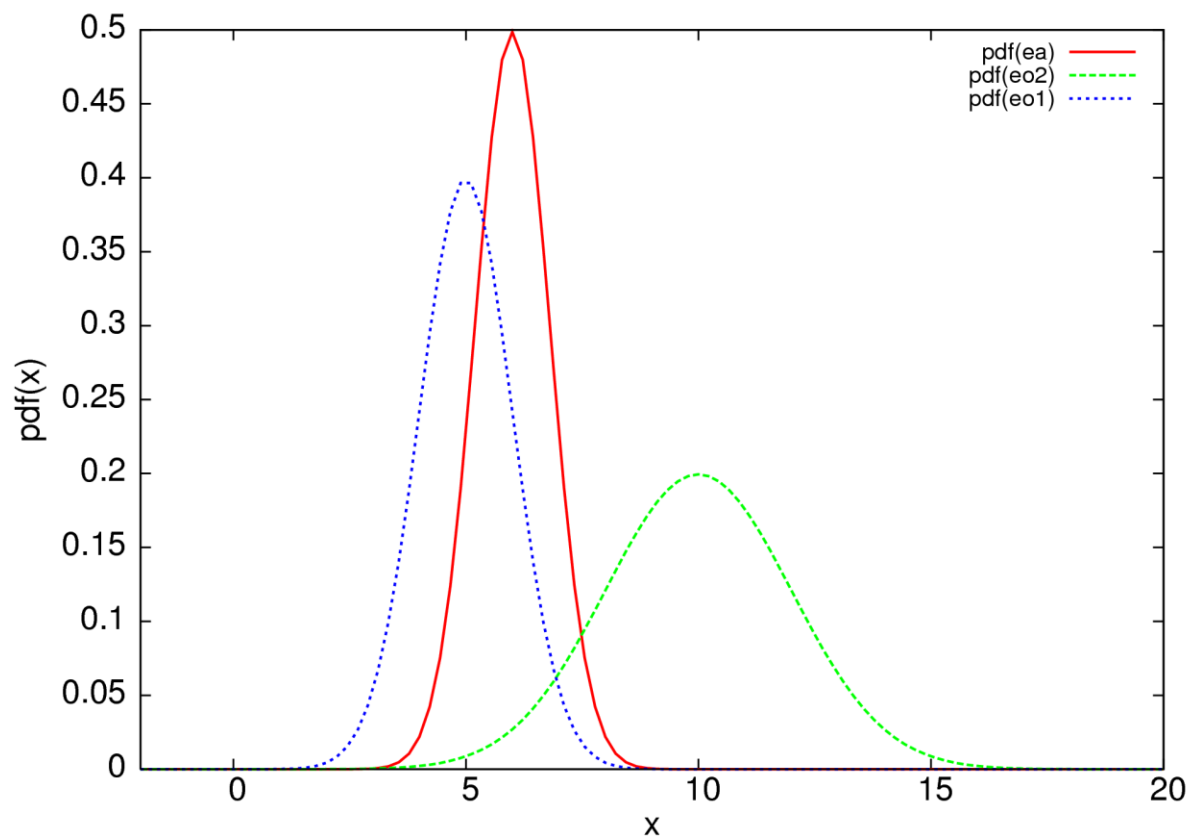
$$x_a \rightarrow \max_x f_1(x)f_2(x) = \max_x \frac{1}{\sigma_1 \sigma_2 2\pi} \exp\left(-\left(\frac{1}{2} \frac{(y_1 - x)^2}{\sigma_1^2} + \frac{1}{2} \frac{(y_2 - x)^2}{\sigma_2^2}\right)\right) =$$

$$= \min_x \left(\frac{1}{2} \frac{(y_1 - x)^2}{\sigma_1^2} + \frac{1}{2} \frac{(y_2 - x)^2}{\sigma_2^2} \right)$$

$$x_a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} y_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} y_2$$

$$\frac{1}{\sigma_a^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} = \sum_i \frac{1}{\sigma_i^2} \quad \Rightarrow \text{Minden megfigyelés növeli a megbízhatóságot!}$$

Példa:



$$y_1=5; \sigma_1=1$$

$$y_2=10; \sigma_2=2$$

$$x_a=6; \sigma_a=0,89$$



Összefoglalás

Normális eloszlású hibák esetén a legkisebb négyzetek módszere és a maximum likelihood módszerek ekvivalensek

Legkisebb négyzetek módszere
Pl. optimális interpoláció
(lineáris regresszió)

$$x_a = \sum_{i=1}^p k_i y_i$$

$$k_i = ?$$

Maximum likelihood módszer
Pl. variációs módszer
(veszteség függvény)

$$J(x) = \frac{1}{2} \sum_{i=1}^p \frac{(x - y_i)^2}{\sigma_i^2}$$

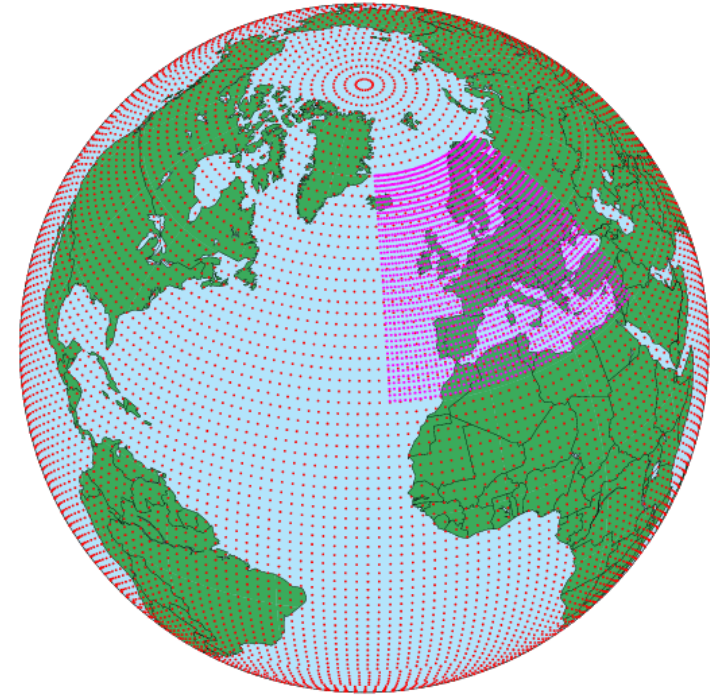
$$\operatorname{argmin} J(x) = ?$$

(iteratív minimum kereső algoritmusok)

Adatasszimiláció a gyakorlatban

A numerikus modellek egy 3D **rácson** oldják meg a HTER-t. A rácson értelmezett meteorológiai változók együttese az **állapotvektor (továbbiakban $\mathbf{x}$)** amely a mai modellek esetében $n \sim 10^7$ nagyságrendű.

Adatasszimiláció: az állapotvektor megadása a kezdeti időpillanatban (**analízis**) úgy, hogy az minél közelebb legyen a valósághoz.



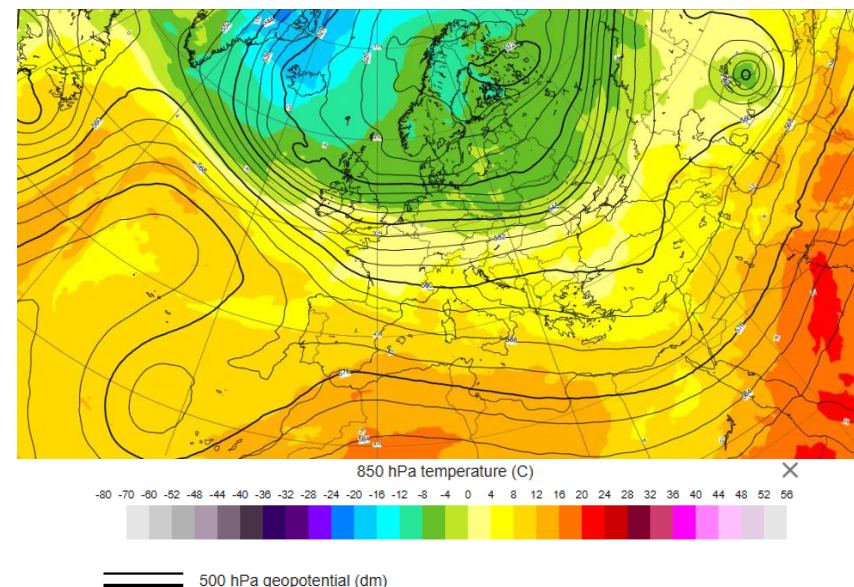
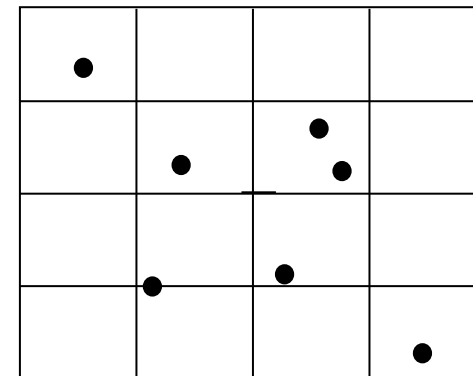
Felhasználható információk

Megfigyelések:

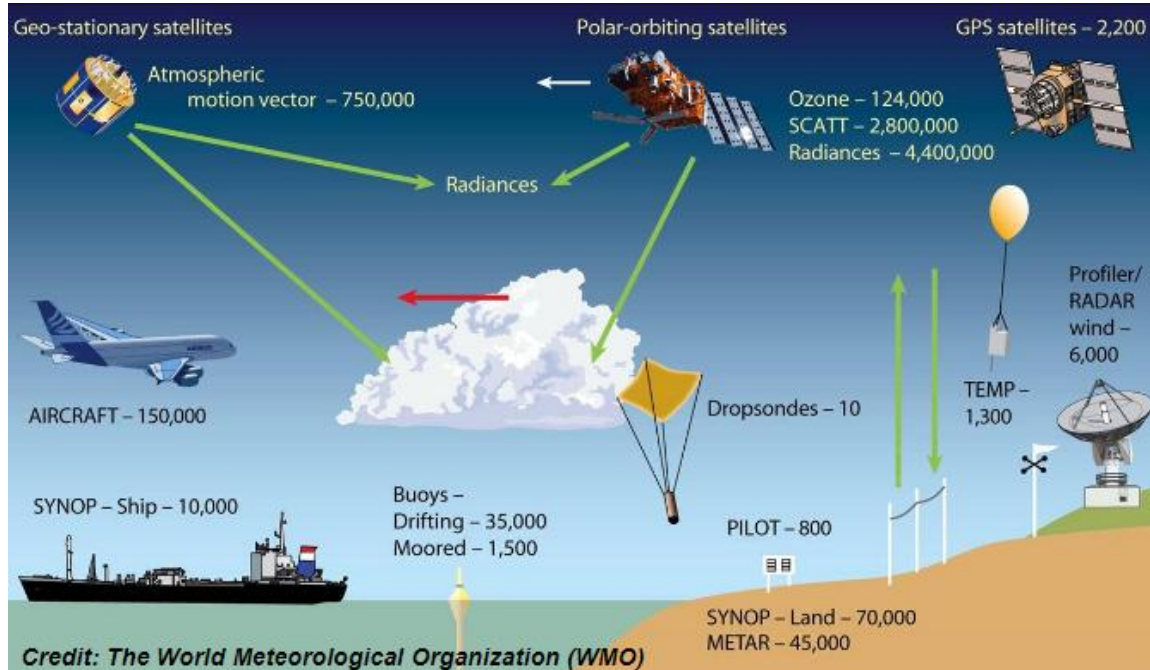
- Időben rendszeres számszerű mérések
- Térben **szabálytalan rács**
- Összetett összefüggés (lehet) a mért és a modell változók között
- A felhasznált megfigyelések száma kevesebb, mint a modellrács pontjainak száma ($p \sim 10^5$)

Background/first guess:

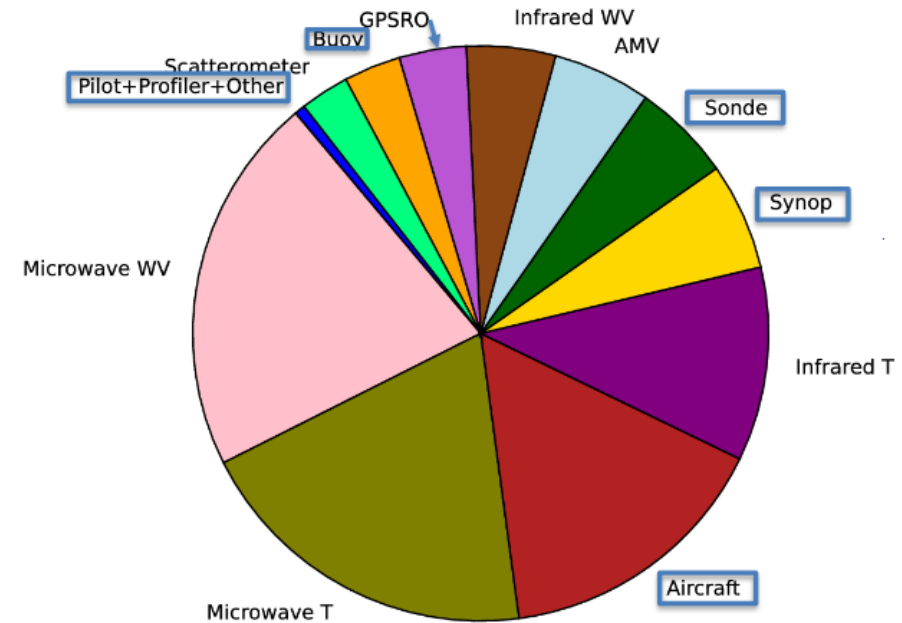
- A modell által előrejelzett állapotvektor
- Térben szabályos rács
- Közvetlenül modell változók
- Dimenzió: $n \sim 10^7$



Megfigyelések globális rendszere



Forrás: Tony McNally, ECMWF NWP training, 2015

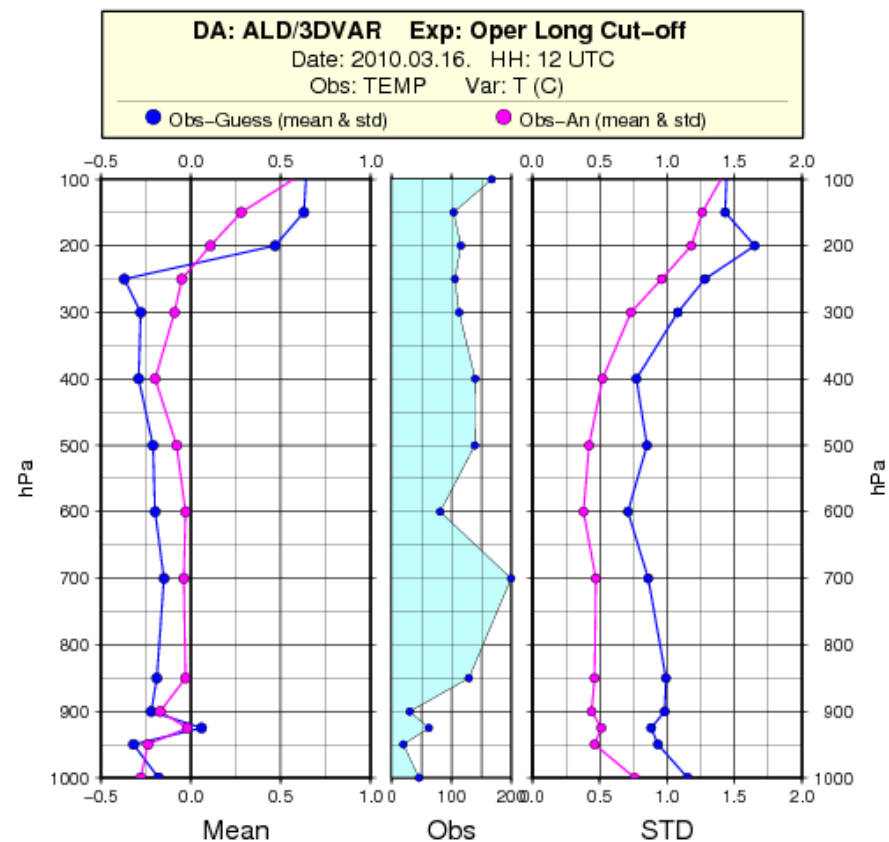


24 órás előrejelzés esetén:

- 70 % : műholdas adatok
- 30 %: in situ megfigyelések

Konvencionális megfigyelések

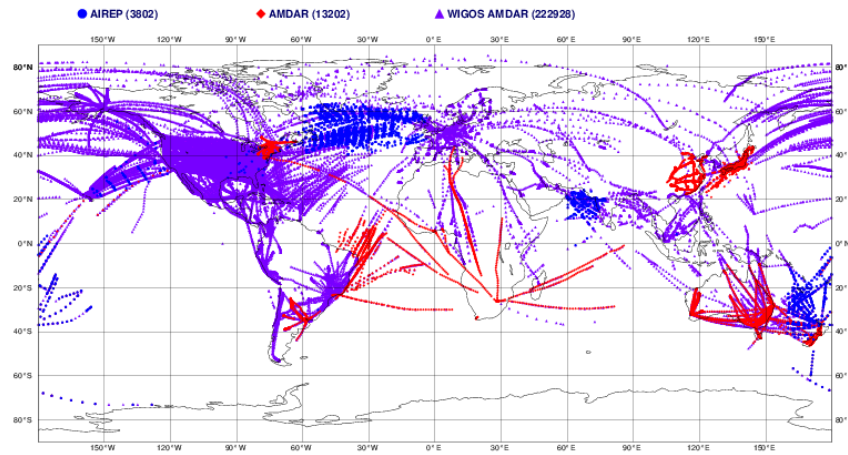
Eszköz	Változó	Szint
SYNOP SHIP METAR	hőmérséklet, harmatpont, szél, légnymás	Felszín, 2m, 10m, 25m
Bóják	hőmérséklet, légnymás, szél	2m
Rádiószondák TEMPSHIP Ejtőszondák	hőmérséklet, légnymás, szél, nedvesség	Profil
Profilerek	szél	Profil
Repülőgépek	hőmérséklet, légnymás, szél	Profil, repülési szint



Konvencionális/in-situ megfigyelések

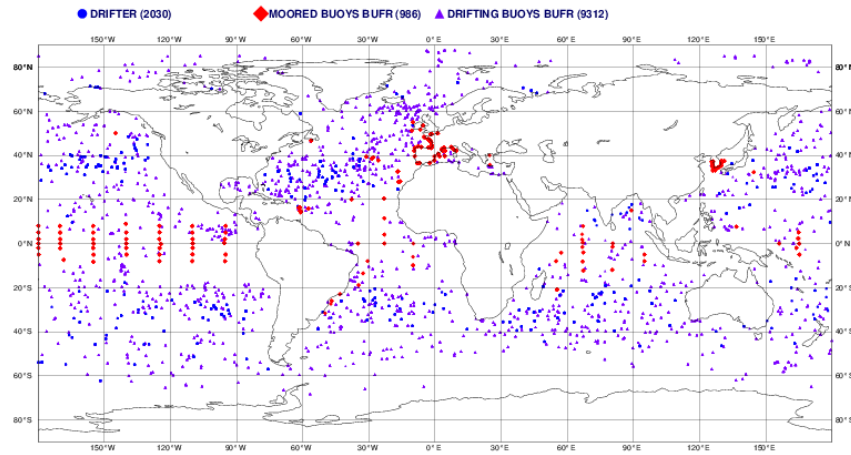
ECMWF data coverage (all observations) - AIRCRAFT
13/11/2018 00

Total number of obs = 239932



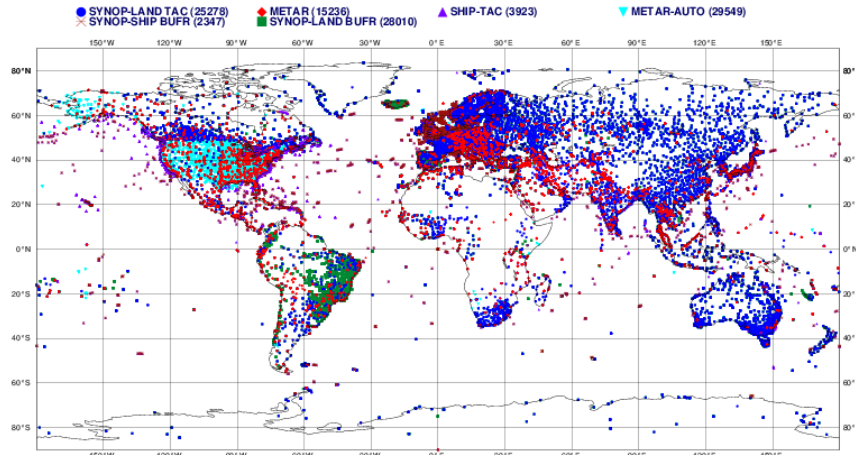
ECMWF data coverage (all observations) - BUOY
13/11/2018 00

Total number of obs = 12328



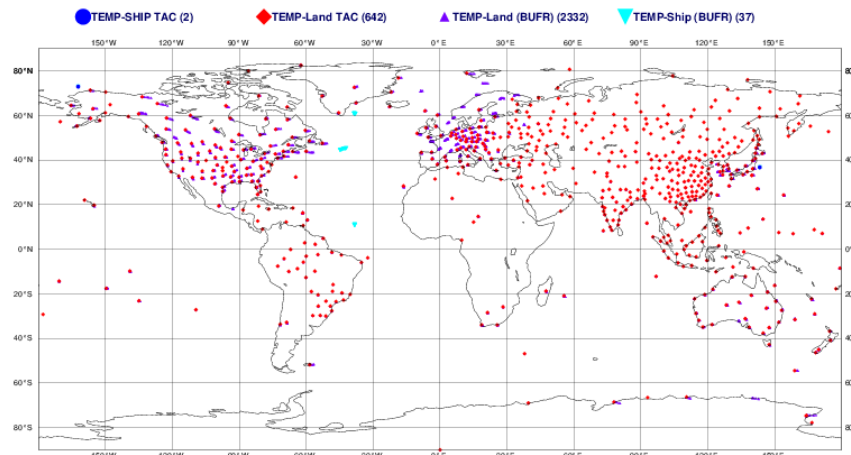
ECMWF data coverage (all observations) - SYNOP-SHIP-METAR
13/11/2018 00

Total number of obs = 104343



ECMWF data coverage (all observations) - RADIOSONDE
13/11/2018 00

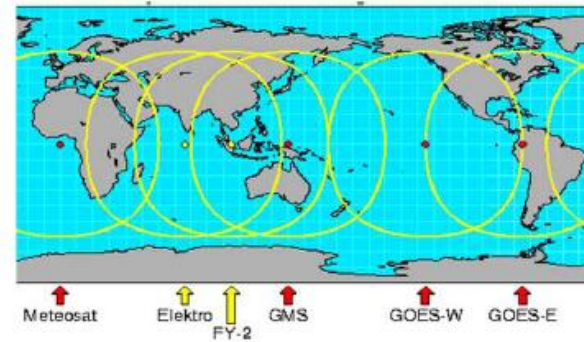
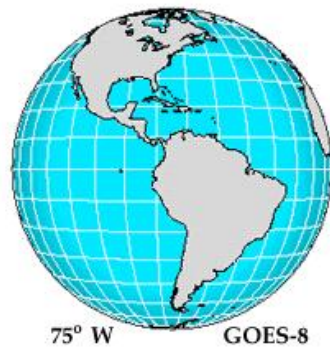
Total number of obs = 3013



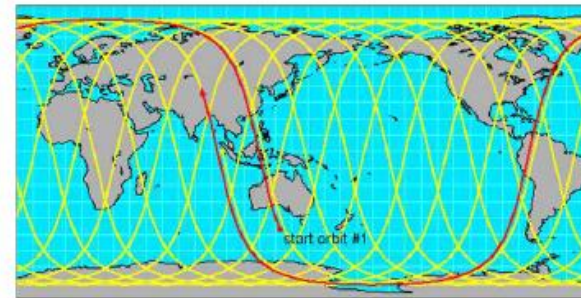
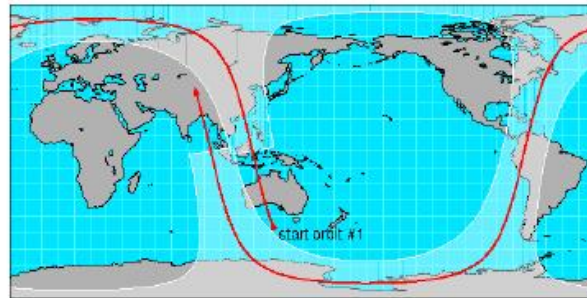
Műholdas mérések

Geostacionárius és kvázipoláris (low Earth orbit) műholdak

GEO

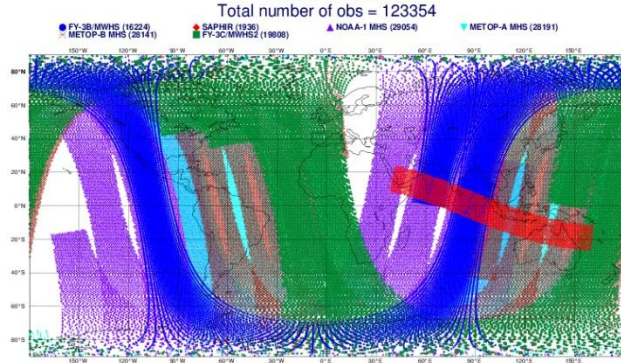


LEO

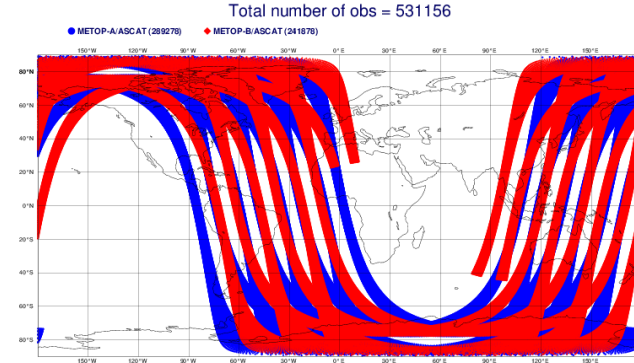


Műholdas mérési lefedettség

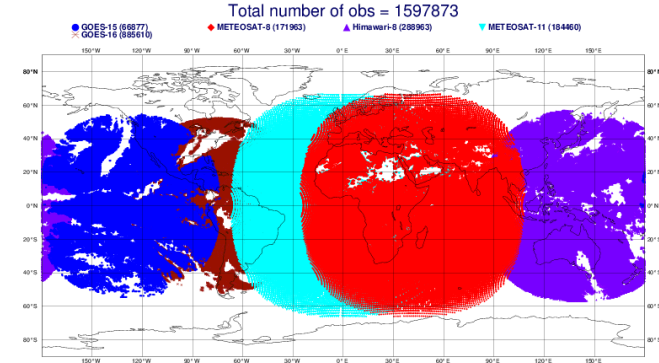
ECMWF data coverage (all observations) - MICROWAVE HUMIDITY SOUNDERS
13/11/2018 00



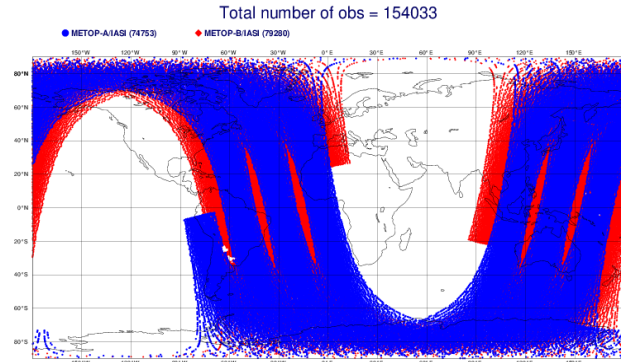
ECMWF data coverage (all observations) - SCATTEROMETER
13/11/2018 00



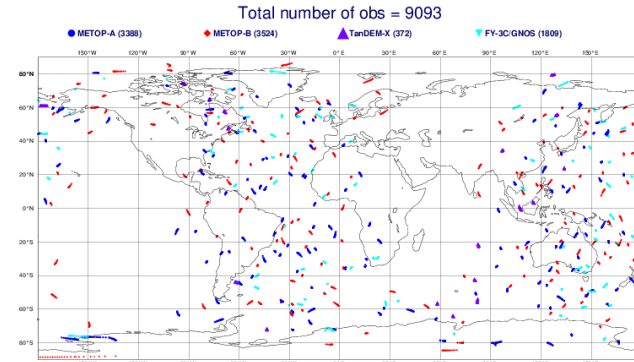
ECMWF data coverage (all observations) - GEOSTATIONARY RADIANCES
13/11/2018 00



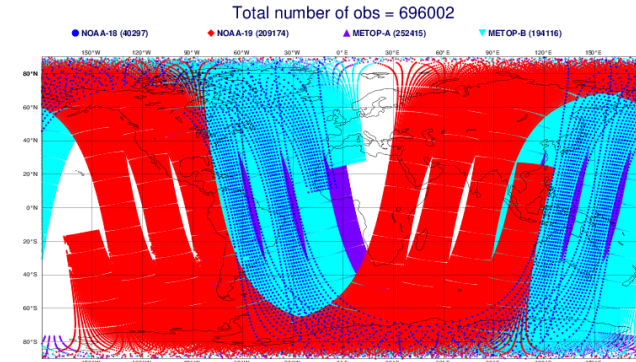
ECMWF data coverage (all observations) - IASI
13/11/2018 00



ECMWF data coverage (all observations) - GPSRO
13/11/2018 00



ECMWF data coverage (all observations) - HIRS
13/11/2018 00



Megfigyeléstípusok jellemzői

Konvencionális mérések

Nem egyenletes térbeli és időbeli lefedettség. Vannak olyan területek, ahol nincsenek konvencionális mérések.

Pontszerű mérések

Nem egységes adatminőség

Közvetlenül azt mérik, amit a modellben is használunk

Értelmezésük és használatuk az adatasszimiláció során „egyszerű”

Műholdas mérések

Indirekt és gyakran bonyolult változók mérése, ami a modell számára NEM azonnal felhasználható → átalakítások szükségesek → **radiatív transzfer egyenlet** segítségével megoldható a probléma.

Nehezítő körülmények (pl. felhőzet)

Gyenge vertikális felbontás

Bár a legtöbb műholdas megfigyelés nagyon jó minőségű, ez hirtelen megváltozhat a műszer meghibásodása vagy szoftveres probléma miatt

Hatalmas adatmennyiség

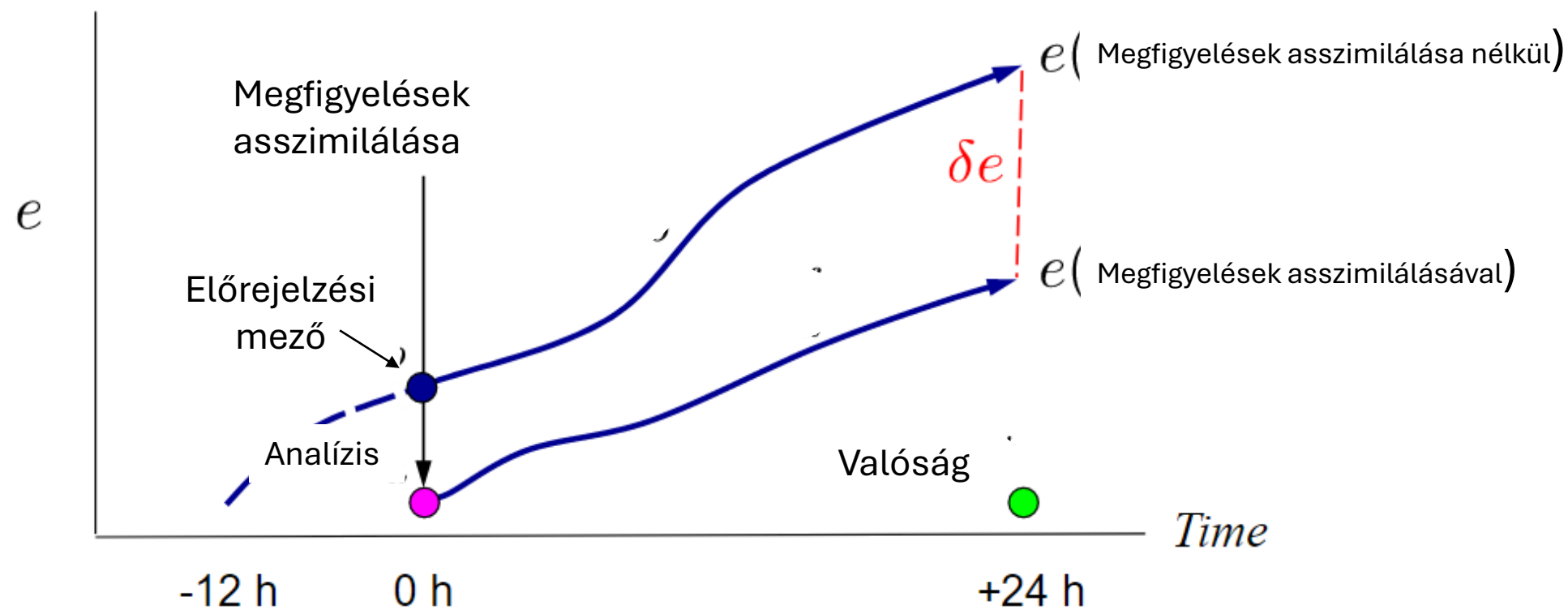
Egyenletes, globális térbeli lefedettség

Nagy időbeli sűrűség

Adatasszimiláció hatékonyságának az ellenőrzése

FSOI - Forecast Sensitivity Observation Impact

Fcst Error





To access this activity, students
go to zzi.sh and insert

ake94796

 Copy class code



Irodalom

Berre, L., 2000: Estimation of synoptic and meso scale forecast error covariances in a limited area model. *Mon Weather Rev* **128**, 644-667.

Bouttier, F. and Courtier, P., 1999: Data Assimilation concepts and methods. *ECMWF Lecture Series (Data assimilation and use of satellite data, http://www.ecmwf.int/newsevents/training/lecture_notes/pdf_files/ASSIM/Ass_cons.pdf)* 72.

Dévényi, D. és Gulyás, O., 1988: Matematikai statisztikai módszerek a meteorológiában, *Tankönyvkiadó*, Budapest, 443 pp.

Gandin, L.S., 1963: Objective analysis of meteorological fields, *Gidrometeorologicheskoe Izdatelstvo*, Leningrad, English translation by Israeli Program for Scientific Translations, Jerusalem, 1965

Gilbert, J. and Lemarechal, C., 1989: Some numerical experiments with variable-storage quasi-Newton algorithms, *Mathematical Programming* **45**, 407-435.

Kalnay, E., 2003: Atmospheric Modeling, Data Assimilation and Predictability, *Cambridge University Press*, Cambridge, 341 pp.

ECMWF Data Assimilation NWP training, 2019

ECMWF eLearning, An introduction to Data Assimilation

https://learning.ecmwf.int/mod/scorm/player.php?a=179¤torg=An_introduction_to_data_assimilation_ORG&scoid=458