Gépi tanulásos módszerek alkalmazása a modellezésben, előrejelzésben

Szépszó Gabriella
Klimatológiai és Kutatás-fejlesztési Igazgatóság

HungaroMet
Magyar Meteorológiai Szolgáltató Nonprofit Zrt.

A (fizikai) modellfejlesztés folyamata

Célkitűzés:
Fejlesztési cél kitűzése (pl. felbontás finomítása)

Előkészítés:
Modellkód és programok módosítása, előkészítése, tesztkísérletek megtervezése, beállítások áttekintése, szükséges adatok gyűjtése, futási idő felmérése

Kísérletek:
Rövidebb és hosszabb tesztszimulációk elvégzése (általában 1-hónapos időszak 2 évszakból)

Értékelés:
Verifikáció és esettanulmányok
(verifikációs mutatók, térképes áttekintés)

Felhasználói teszt:
Az előrejelzők, felhasználók legalább 1 hónapig tesztelik a fejlesztést a napi feladataik elvégzésekor, „kvázi-operatív" futtatás

Operatív:
Bevezetés az operatív gyakorlatba, felügyelet

Valós fejlesztési példa 1.

Az IFS 49-es modellváltozatának verifikációs eredményei

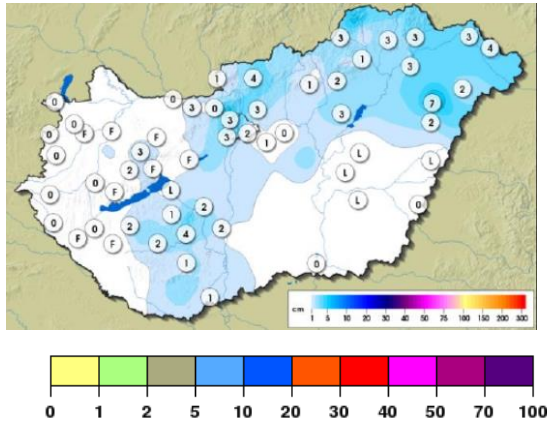
- ECMWF tagországi visszajelzések: 2-méteres hőmérséklet előrejelzések hibái
- Globális középtávú modell → régiónként változhat, hogy mi a hiba oka
- **Fejlesztési területek kijelölése (2020)**: az asszimilációban (pl. új mérés), a felszíni folyamatok leírásában (pl. új hóséma) → a kutatók által kitalálandó, bekódolandó új modellrészek
- Sok-sok fejlesztői kísérlet
- Kísérletek legalább nyári és téli évszakon, verifikáció, esettanulmányok
- Összefésülés a többi fejlesztéssel, együttes hatás vizsgálata
- Tagországi teszt visszajelzésekkel, legalább 3 hónap
- **Operatív kb. 3-4 év múlva (2024)**

https://sites.ecmwf.int/ifs/scorecards

Valós fejlesztési példa 2.

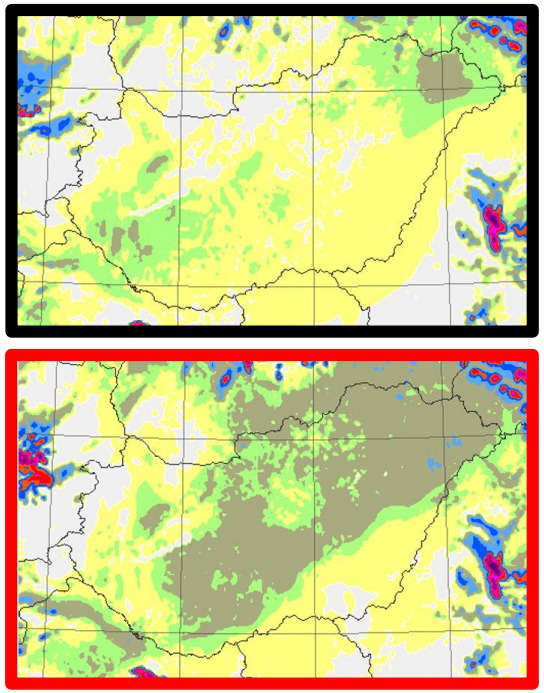
- Hazai előrejelzői visszajelzés, **2023, 2024 telei**: sokáig megmaradó hótakaró az AROME modellben, emiatt gyenge hőmérséklet előrejelzés
- Ismert ok: kezdeti hóvastagság megadásának módja adatasszimiláció hiányában
- **Hóvastagság mérések asszimilációja**
- Beállítások hangolása esettanulmányokon, kísérletek téli hónapokon, verifikáció
- Előrejelzői teszt visszajelzésekkel, legalább 2 hónap (jelen esetben november–február)
- Operatív: várhatóan **2026 márciusától**

2025. január 16. 6 UTC



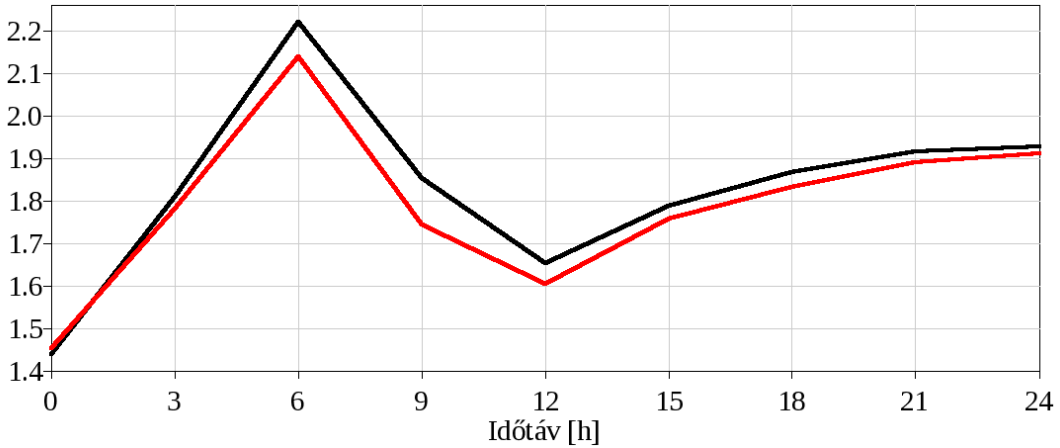
©Tóth H.

Hóvastagság [mm]



2-méteres hőmérséklet négyzetes hibája [°C]; 2025/1/7–2/7

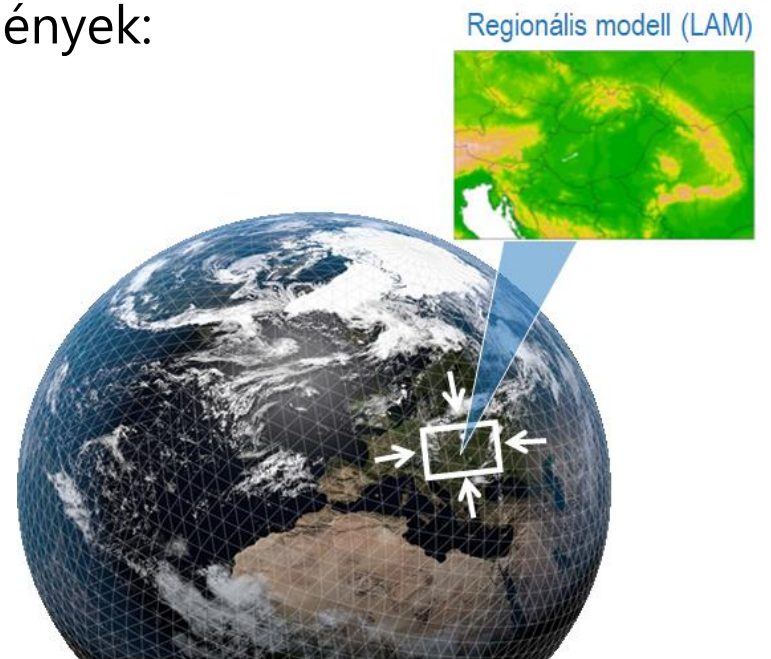
Operatív AROME, **AROME hőmérésekkel**, 6 UTC-s futtatások



HungaroMet Nonprofit Zrt.

Az előrejelzések felhasználói

- Operatív előrejelző szakember (pl. repülésmeteorológus, veszélyjelző), igényei a modell előrejelzéssel szemben:
  - A légköri folyamatok 3-dimenziós leírása (pl. nyomási szinteken)
  - Fizikai konzisztencia és a fizikai folyamatok nyomon követhetősége (még az ensemble előrejelzés egyes tagjai esetében is)
  - Általánosan megbízható előrejelzés, ismert hibákkal
  - Napon belül többszöri frissítés, egyes esetekben valósidejű felhasználás
- Korlátos tartományú előrejelző modell (egyéb dinamikus modell), igények:
  - 3-dimenziós leírás a teljes modellrácson (pl. 90 szinten)
  - Fizikai konzisztencia
- Produktum használó (pl. energia szektor), igények:
  - Nem a teljes 3-dimenziós előrejelzést használja, csak néhány meteorológiai változót, pl. globálsugárzást
  - Nagy pontosság
  - Gyakori frissítés, valósidejű felhasználás



Operatív előrejelzések ütemezése (UTC-ben)

Határfeltételek

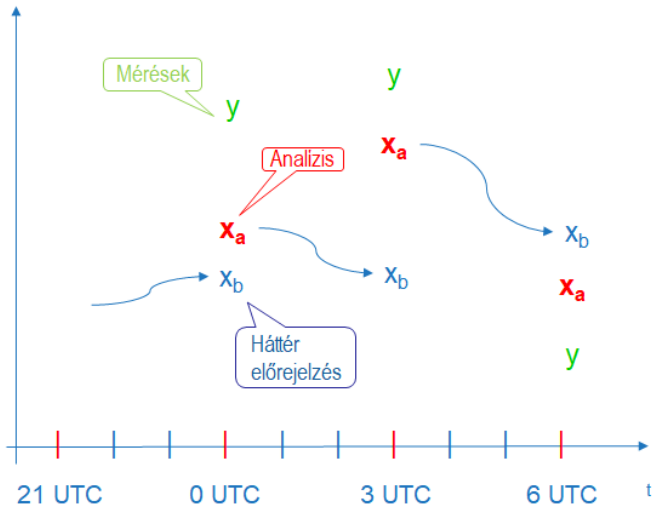
0:00–0:20

(Korábbi) globális előrejelzésből készült kivágat letöltése

Adatasszimiláció

AROME: 1:30–1:35

Mérések gyűjtése, ellenőrzése, objektív analízis → kezdeti feltétel előállítása



Modellintegrálás

AROME: 1:35–2:10

Az egyenletrendszer közelítő megoldása a modell saját vetületén és rácsán, nyers modelloutputok előállítása 15-percenként/óránként

Utó-feldolgozás

AROME: 1:44–1:57

A nyers modelloutputokból speciális paraméterek számítása a kívánt területre, szintre, felbontással

Produktumok továbbítása felhasználóknak

2:00–

Gépi tanulásos módszerek alkalmazása
Példák

- Felszíni jellemzők leírása
- A modellezés egyes részein:
  - Adatasszimiláció: nem-lineáris megfigyelési operátor előállítása
  - Parametrizáció: paraméter becslés, emulátor
- Nyers modelleredmények utó-feldolgozása
- Előrejelzés:
  - Nowcasting
  - Rövid- vagy középtávú előrejelzés
  - Ensemble tagok generálása

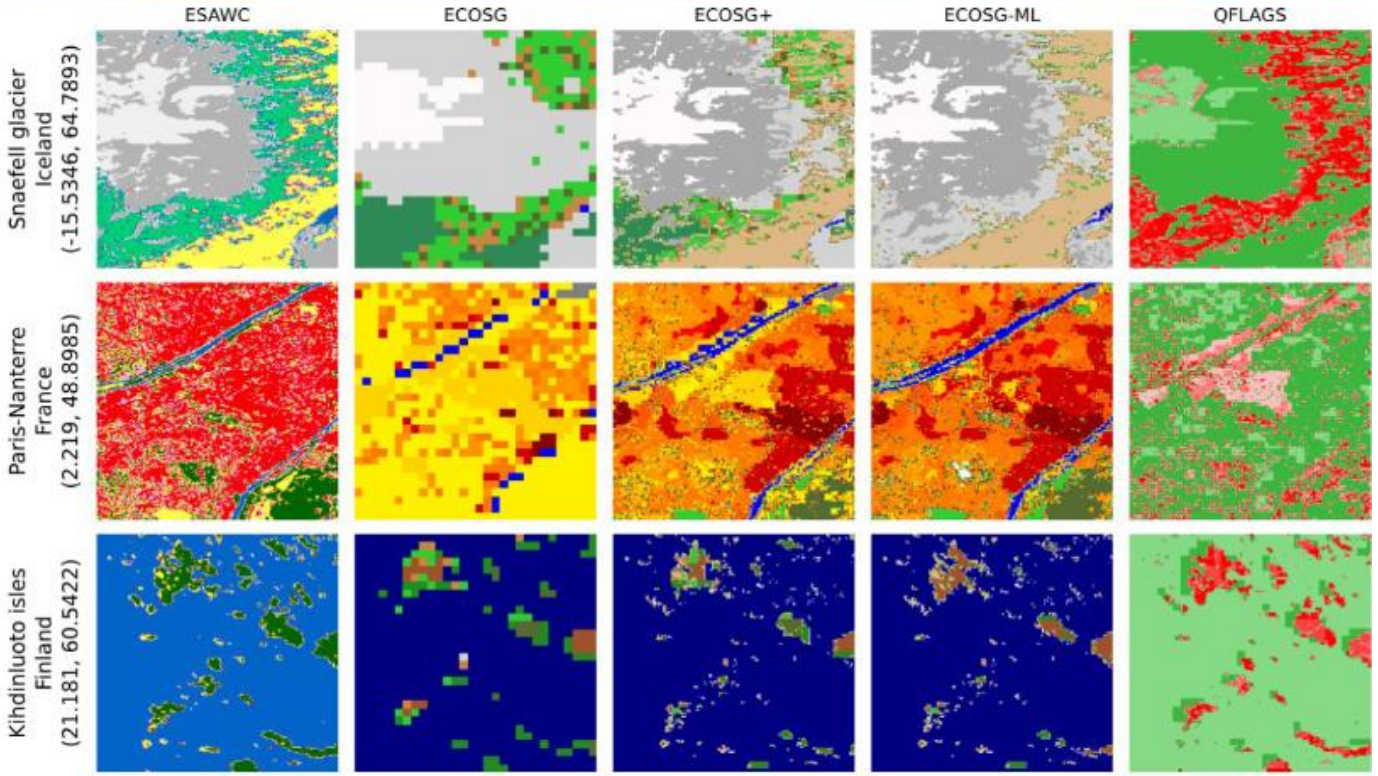
Felszíni jellemzők leírása

Next generation of physiography data is based on Machine Learning

Here, physiography generated by these two methods are compared to ESA CCI landcover Word Cover map (ESAWC) and ECOSG, accompanied by quality flags (QFLAGS).

Links to documentation:
Geoffrey Bessardon et al. (2024),
doi: 10.20944/preprints202409.0953.v1
Thomas Rieutord et al. (2024),
doi: 10.20944/preprints202409.0942.v1

Work now continues at FMI by Rudolf Mård and Panu Maalampi.



Kiindulás:

- 300 m felbontású ECOCLIMAP-SG
- 10 m felbontású ESA WorldCover
- …

Módszer:
auto-encoder

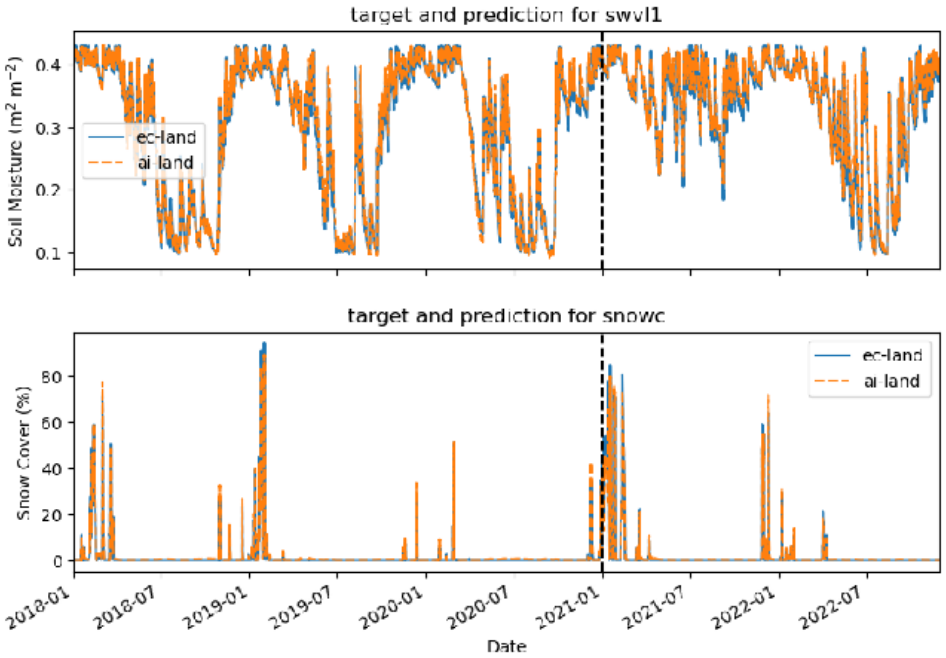
Eredmény:
60 m felbontású ECOSG-ML

ACCORD
A Consortium for COnvection-scale modelling Research and Development

Adatasszimiláción belüli alkalmazások

Emulating the land surface for Data Assimilation

- Developing ML emulator of land surface model (ECLand), providing Jacobian and ensemble information at much reduced cost
  - 500 milliseconds vs 20 mins for 1-month Global run 🚀
- Next steps to test this in "offline" LDAS system for generation of Jacobians
- Emulator more stable than expected, no sign of divergence even after multiple years



Ewan Pinnington et al.

Analízis egyenlet:

$$\mathbf{x}_a^i = \mathbf{x}_b^i + \mathbf{K}\left(y^i - H\left(x_b^i\right)\right)$$

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T\left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)^{-1}$$

Felszíni asszimilációban:

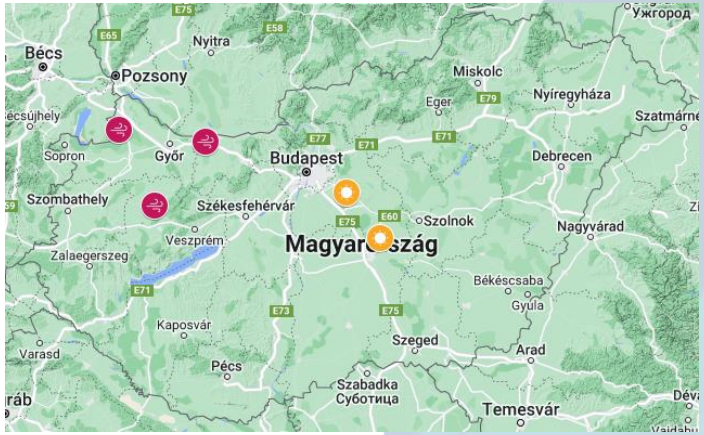
$$\mathbf{H} = \begin{matrix} \frac{\partial T2M}{\partial TG1} & \frac{\partial T2M}{\partial TG2} & \frac{\partial T2M}{\partial WG1} & \frac{\partial T2M}{\partial WG2} \\ \frac{\partial HU2M}{\partial TG1} & \frac{\partial HU2M}{\partial TG2} & \frac{\partial HU2M}{\partial WG1} & \frac{\partial HU2M}{\partial WG2} \end{matrix}$$

Kiszámítása a modellezés során: a felszíni modell (pl. SURFEX) futtatása a kontroll változó kis perturbációival

ECMWF

Nyers előrejelzések utó-feldolgozása

- Közös munka matematikusokkal (Kornyik Miklós) az AROME eredményeinek utó-feldolgozására
- Célparaméter: **100-méteres szélsebesség**, **globálsugárzás**
- Felhasznált információk
  (szél + hőmérséklet + nyomás, sugárzás + hőmérséklet):
  - 15-percenkénti partneri mérések 3/2 pontra
  - 15-percenkénti AROME előrejelzések 36/48 órára (környezet)
- Módszer: **4/3-rétegű perceptron** (neurális háló), időtávonként
- Optimalizálás **négyzetes hibára**
- 8- és12-hónapos időszak sorsolt 80/20 %-án tanítás/validáció



RMSE változás [%]			
	1. pont	2. pont	3. pont
tanítás	9	16	12
validáció	7	14	9

RMSE változás [%]		
	1. pont	2. pont
tanítás	21	11
validáció	21	10

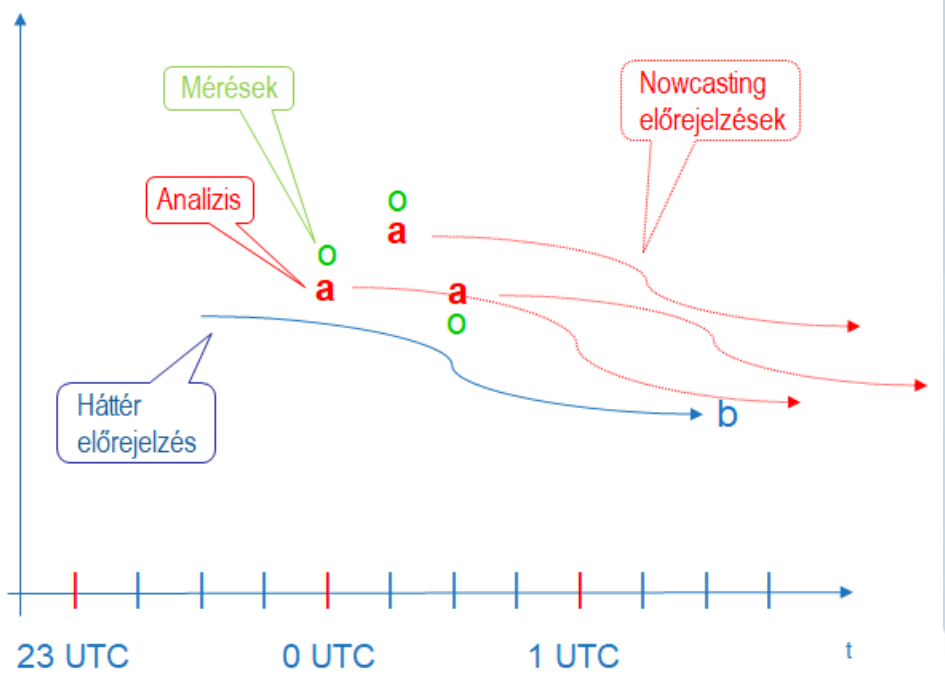
nagarolvet

További információ:
Szépszó et al., 2023

Ultra-rövidtávú előrejelzések (nowcasting)

- Gyakori analízisből **1-3 órára szóló** ultra-rövidtávú előrejelzések

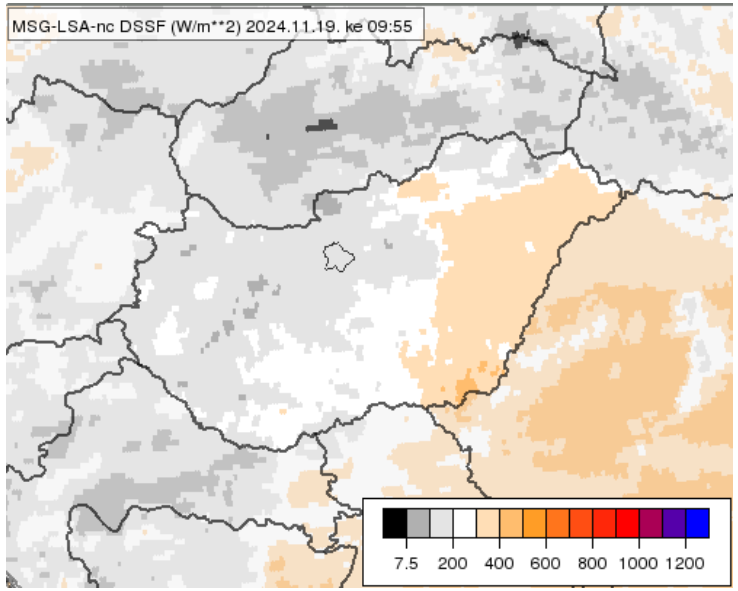
- Az analízishez a háttér (modell) előrejelzést ötvözzük mérésekkel, szempontok utóbbi kiválasztásában:
    - Nincs hosszú várakozás a mérésekre, gyors elérésű adatokat használnak
    - Az alkalmazott mérések közvetlen információt adnak a légköri eseményekről
    - A mérések is „rácson” állnak rendelkezésre

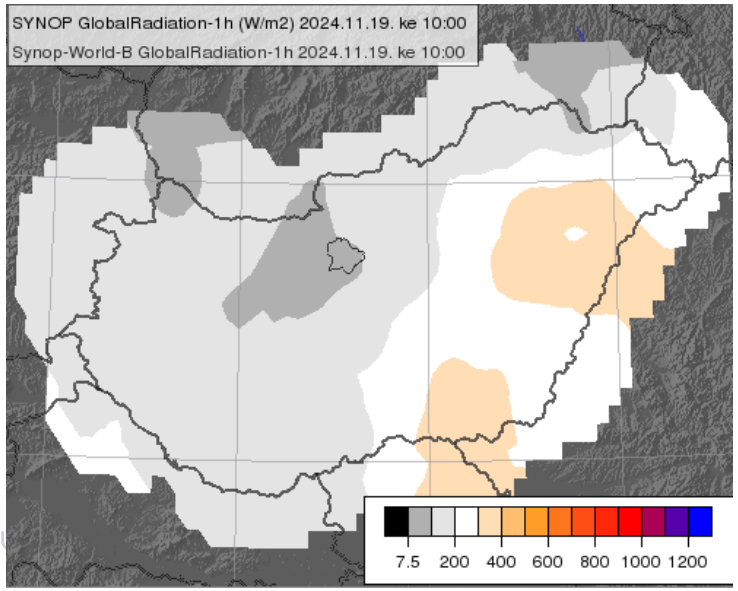


- Kézenfekvően felhasználható méréstípusok:
  radar (reflektivitás, szél),
  műhold (sugárzás, származtatott mennyiségek)

- Az analízisből egy **egyszerűsített és hatékony módszerrel** objektum „áthelyeződést” számítanak
  → **valósidejű felhasználás**

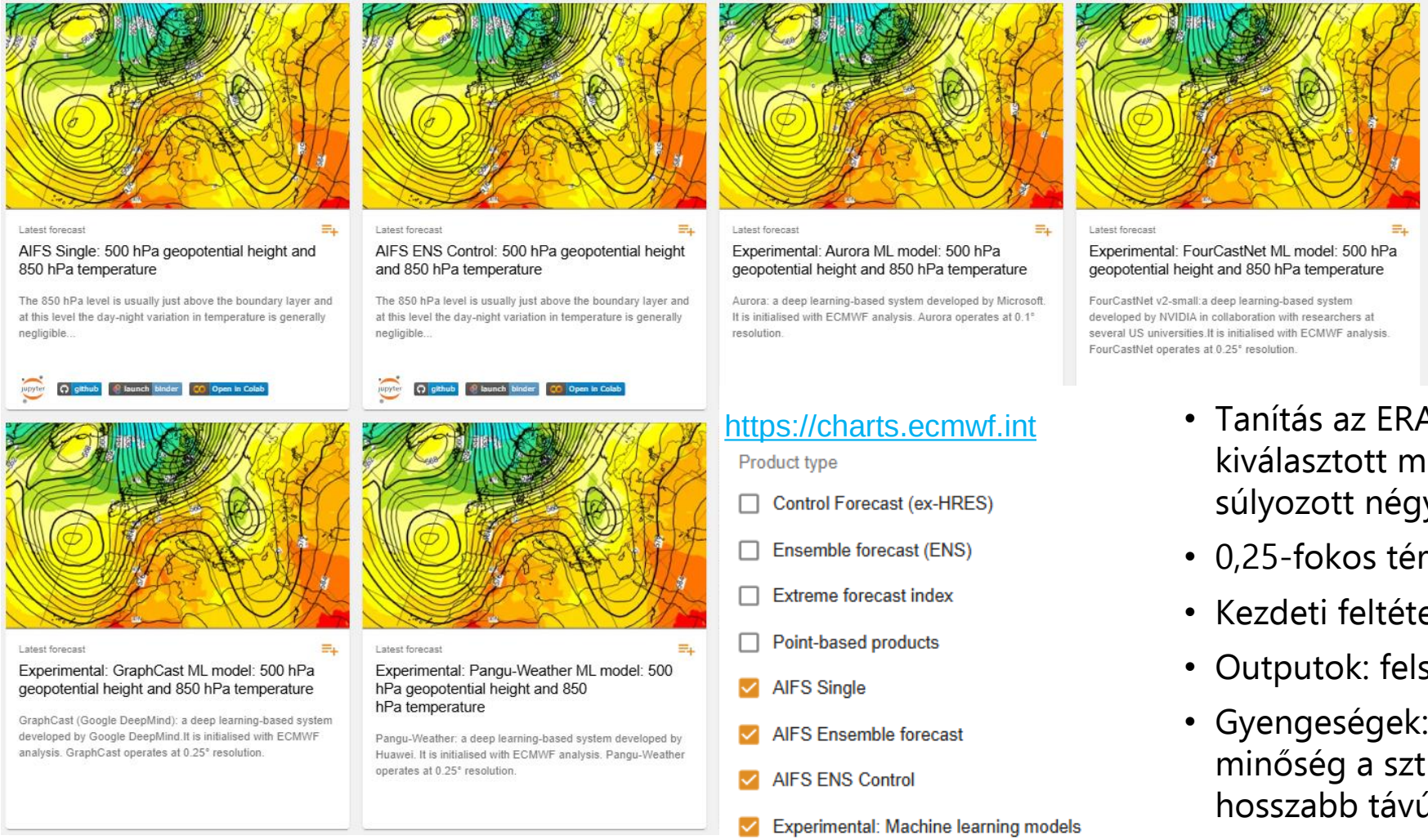
Sugárzás nowcasting – HungaroMet





- A naperőművek termelését befolyásolja az időjárási helyzet
- Az üzemeltetők számára kötelező a negyedórára bontott termelésbecslés, ami a nap során többször frissíthető (pontosítható)
- A termelés túlbecslése esetén kiegyenlítő energia → megvásárlását ráterhelik a pontatlan becslést adókra
- Intraday menetrend pontosításához közel valósidejű becslések szükségesek
- Lehetséges kiindulási adat: műholdas sugárzásadatokból számított globálsugárzás produktum 15-percenként
- (Állomási mérésekkel összevetve: felülbecslés, vastagabb felhőzet esetén nagyobb hiba)
- Neurális hálón alapuló módszer, mely felismeri a térbeli mintázatokat és időbeli összefüggéseket (pl. a felhőzet esetében)
- Néhányéves tanítási adatsor → 1-3 órás előrejelzés

Középtávú előrejelzés



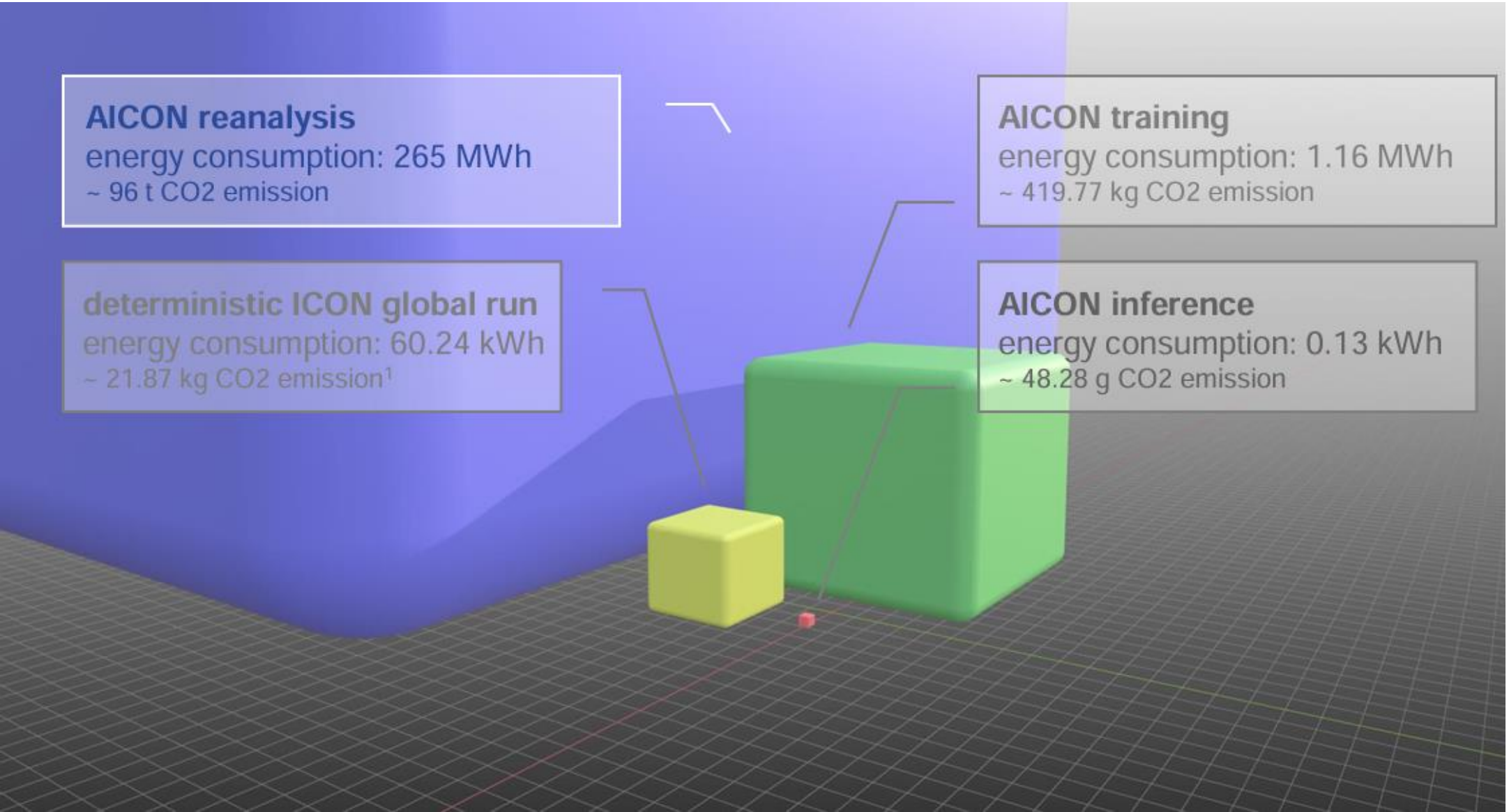
https://charts.ecmwf.int

Az ECMWF-nél elérhető 6 gépi tanulásos módszer operatív előrejelzése néhány meteorológiai paraméterre, de a kód is letölthető

Az **AIFS** perturbálatlan és ensemble előrejelzéseit a tagországok operatívan megkaphatják

- Tanítás az ERA5 adatbázison, optimalizáció kiválasztott mérőszámokra: AIFS – területtel súlyozott négyzetes hiba
- 0,25-fokos térbeli és 6/12-órás időbeli felbontás
- Kezdeti feltétel: ECMWF analízis
- Outputok: felszínközeli és nyomásszint változók
- Gyengeségek: simított előrejelzés, gyengébb minőség a sztratoszférában, extrém események, hosszabb távú előrejelzések
- Hasonló egyre több országban: AICON, BRIS …

Fizikai és AI-előrejelzés (valós) költsége



Összefoglalás

- Fizikai modellfejlesztés területei:
  - Felbontás növelése: km-es/hektometrikus skálájú globális/regionális előrejelzések
  - Módszertani fejlesztések, részfolyamatok részletesebb leírása, új folyamatok figyelembevétele
- Mesterséges intelligencia használata:
  - Az előrejelzés folyamatában (pl. asszimiláció, parametrizáció, ensemble előrejelzések)
  - Nyers modelleredmények utó-feldolgozásában
  - Nowcastingban
  - **Előrejelzés esetén alapvető hozzávaló a jó fizikai modell (pl. re-analízis)**
  - Kérdések: operatív használat és fejlesztés, kutatás viszonya, fizikai megértés