



Sugárzásra és magassági szélre vonatkozó rövidtávú előrejelzések operatív statisztikai utófeldolgozása

Szépszó Gabriella¹, Baran Ágnes², Baran Sándor², Jávorné Radnóczy Katalin¹, Kornik Miklós⁴, Tajti Dávid¹

¹ Országos Meteorológiai Szolgálat, szépszo.g@met.hu

² Debreceni Egyetem

³ Rényi Alfréd Matematikai Kutatóintézet,

⁴ Turbine.AI

DOI: 10.56474/legkor.2023.3.1

A nap- és szélenergia termelés becsléséhez fontos kiindulási információt szolgáltatnak az időjárás-előrejelzések. A folyamatos fejlesztés ellenére a közvetlen modell-előrejelzések hibával terhelték, ezek azonban utófeldolgozással csökkenthetők olyan meteorológiai változók esetében, amelyekre jó minőségű mérési adatsorok állnak rendelkezésre. Az OMSZ AROME modellel készülő operatív előrejelzéseire többféle paraméteres és gépi tanuláson alapuló statisztikai utófeldolgozási módszer került kifejlesztésre, melyek alkalmazása javítja a Napból közvetlenül érkező és szórt rövidhullámú sugárzásra (azaz a globálsugárzásra) és a talajközeli magassági szélre vonatkozó előrejelzések beválását. A cikk a Magyar Meteorológiai Társaság 2023. március 17-ei rendezvényén elhangzott előadások összefoglalója.

Operational statistical post-processing of short-range global radiation and low-level wind forecasts

Numerical weather predictions (NWP) provide key input information for estimation of solar and wind energy production. In spite of the continuous model developments the direct forecast model outputs are imperfect and contain some errors. However, these errors can be reduced by post-processing the raw forecasts of those meteorological variables, for which time series of good quality measurements are available. Multiple parametric and machine learning based statistical post-processing methods have been elaborated for operational forecasts of the AROME NWP model at the Hungarian Meteorological Service in order to improve the raw predictions of the global horizontal irradiance and 100-meter wind speed variables.

Bevezetés

A numerikus modellekkel készített számszerű időjárás-előrejelzések a modellfejlesztéseknek köszönhetően folyamatos fejlődést mutattak az elmúlt évtizedekben.

Például az Európai Középtávú Előrejelző Központ (ECMWF, European Centre for Medium-Range Weather Forecasts) globális előrejelzései az 5-napos időtávra Európa felett ma ugyanolyan pontosságúak, mint 3 évtizeddel ezelőtt a 2 napra szóló előrejelzések

voltak (Haiden et al., 2022) vagy az Országos Meteorológiai Szolgálatnál (OMSZ) használt AROME mezoskálájú modell előrejelzéseiben is több paraméterre (pl. az 500 hPa-os szint geopotenciál magasságára) kimutatható az átlagos bevalás fokozatos javulása.

Az előrejelzések azonban nem csak modellfejlesztés útján javíthatók. A számítógépes algoritmusok és kapacitás rohamos fejlődésével mára lehetővé vált az utófeldolgozásuk olyan paraméteres és gépi tanuláson alapuló statisztikai módszerekkel, melyek az évtizedek során felhalmozott előrejelzési és mérési adatokat kiaknázva az előrejelzés elkészülte után néhány perccel érdemben javítani tudják annak nyers eredményét. Míg a modellfejlesztés a teljes előrejelzés javítását tűzi ki célul és egy-egy szakasza általában többéves folyamat, addig az utófeldolgozás csak néhány meteorológiai változóra koncentrál, és a megfelelő módszer kiválasztása után rendkívül hatékonyan alkalmazható. Emiatt olyan ágazatokban bírhat hozzáadott értékkel, ahol rövid időtávon nagymértékű pénzügyi veszteség származhat a felhasznált előrejelzések pontatlanságából. A hazai nap- és szélenergia termelőknek a villamosenergia kapacitás tervezéséhez két napra előre meg kell becsülniük 15-perces bontásban a várható termelésüket (ezt hívják menetrendezésnek). A termelésbecslés érzékeny a kiindulási meteorológiai adatokra, azaz a felszínre érkező rövidhullámú sugárzásra és a szélturbinák rotormagasságában uralkodó szélviszonyokra vonatkozó modell-előrejelzések pontosságára. A tervezett menetrendtől való eltérés pénzügyi szankciót von maga után, így a nap- és szélenergia termelés kézenfekvő felhasználási területei az utófeldolgozási fejlesztéseknek.

Cikkünkben bemutatjuk az AROME és AROME-EPS globálsugárzás és 100-méteres szélesség előrejelzéseinek általános jellemzőit és az ezek javítására kidolgozott utófeldolgozási technikákat.

Adatok

A nemzetközi együttműködésben fejlesztett AROME korlátos tartományú modellt az OMSZ-nál 2,5 km-es horizontális rácsfelbontáson és 60 szint használatával alkalmazzuk egy Kárpát-medencét lefedő területen. A modellel naponta nyolcszor készülnek előrejelzések (Tóth et al., 2021): 0, 6, 12 és 18 UTC-kor 48-órás, míg 3, 9, 15 és 21 UTC-kor 36-órás időtávra, melyek során a tartományon kívül zajló folyamatok leírását az ECMWF előrejelzéseiből származó határfeltételek biztosítják. Az előrejelzések kezdeti feltételét lokális mérések asszimilációjával állítják elő.

Az AROME előrejelzéseit 0 és 12 UTC-kor kiegészíti egy-egy 2 napra szóló, 11-tagú ensemble előrejelzés (az AROME-EPS; Jávorné Radnóczy et al., 2020), ami lehetővé teszi az előrejelzési bizonytalanság számszerűsítését. Az AROME-EPS az ECMWF globális ensemble előrejelzések perturbálatlan (kontroll) és első 10 perturbált tagjának a Kárpát-medencére 2,5 km-es felbontással való leskálázásával áll elő, és 2023. március óta az asszimilált mérések perturbációjával a bizonytalanság helyi jellemzőit is reprezentálja. Az AROME és az AROME-EPS előrejelzéseiből több magyarországi nap- és szél erőmű üzemeltető kap globálsugárzás és 100-méteres szélesség adatokat 15-percenkénti gyakorisággal.

A fejlesztés során az AROME és az AROME-EPS globálsugárzás és 100-méteres szélesség előrejelzéseinek javítását 5 földrajzi pontban tűztük ki célul, melyekben nap- és szél erőművek működnek. A 3 szél erőmű Magyarország északnyugati részén, míg a 2 napelempark az ország középső részén helyezkedik el. A hibakorrekcióra irányuló utófeldolgozási módszereknek nagyon fontos kiindulási adatai a mérések. Az alábbi fejlesztéseknél felhasznált mérések többféle forrásból származnak (1. táblázat). A szélviszonyokra vonatkozóan az OMSZ-nak 10-méteres magasságban vannak a teljes országot lefedő mérései, ezek azonban nem alkalmasak a magasabb rétegek áramlási jellemzőinek leírására. Ezért különösen fontosak azok a szél erőművek 100-méteres rotormagasságában történő (szélesség és szélirány) mérések, melyeket az OMSZ egyik partnere bocsátott rendelkezésünkre a fejlesztéshez a fenti 3 északnyugati (továbbiakban A, B és C) pontra 15-perces gyakorisággal. A globálsugárzásra ugyancsak kaptunk a partnerünktől 15-perces sűrűségű méréseket az említett 2 közép-magyarországi (továbbiakban D és E) pontra. Ezeket elsősorban az AROME előrejelzésekre kifejlesztett utófeldolgozáshoz használtuk; az AROME-EPS esetében bevontuk az OMSZ méréseit is, amelyeket – bár térben valamivel távolabb esnek – megbízhatóbb minőségűnek tartunk. Az OMSZ sugárzásmérő hálózatából 7 állomást választottunk ki (Aszód, Budapest, Debrecen, Kecskemét, Pécs, Szeged, Tápiaiősele), melyek közül 4 közel esik a fenti napelemparkokhoz, 3 pedig az ország távolabbi tájain helyezkedik el (utóbbiak bevonásával azt vizsgáltuk, hogy az OMSZ hálózatából kiválasztott megbízható mérések felhasználásával tudjuk-e tetszőleges földrajzi pontra javítani az előrejelzést). Az OMSZ mérőállomásai 10 percenként szolgáltatnak adatokat, ezért ahhoz, hogy össze tudjuk őket hasonlítani az AROME-EPS 15-percenkénti előrejelzéseivel,

Utófeldolgozott előrejelzés	Módszerek	Bemenő előrejelzés			Bemenő mérés		
		Változó	Földrajzi hely	Magassági szint	Változó	Földrajzi hely	Magassági szint
AROME 100 m-es szélsébség	egy pontos	szélsébség, szélirány, nyomás, hőmérséklet	előrejelzési rácspont 2 rácstávolságnyi környezete	80, 90, 100, 110, 120 m	szélsébség, szélirány	mérési pont	100 m
AROME globálsugárzás	egy pontos	globálsugárzás	előrejelzési rácspont 2 rácstávolságnyi környezete	felszín	globálsugárzás	mérési pont	felszín
	egy pontos	hőmérséklet	előrejelzési rácspont 2 rácstávolságnyi környezete	2 m	hőmérséklet	legközelebbi OMSZ mérési pont	2 m
AROME-EPS 100 m-es szélsébség	TN EMOS LN EMOS TN DRN LN DRN	szélsébség	mérésekhez legközelebbi rácspont	100 m	szélsébség	összes (3) mérési pont	100 m
AROME-EPS globálsugárzás	CLO EMOS CNO EMOS	globálsugárzás	mérésekhez legközelebbi rácspont	felszín	globálsugárzás	összes (7) mérési pont	felszín

1. táblázat. Az utófeldolgozás során felhasznált adatok jellemzői

30-percre vonatkozó átlagolást végeztünk mind a méréseken, mind az előrejelzéseken. A partneri méréseket a továbbiakban – tekintettel az adatok bizalmas jellegére – anonim módon szerepeltetjük, ezért a mérési helyszínek és az adatok forrása a szövegben nem azonosítható. A közvetlenül javítani célzott meteorológiai változókon kívül az AROME előrejelzések utófeldolgozása során felhasználtuk az OMSZ mérőhálózatából az előrejelzési ponthoz legközelebb eső állomások hőmérsékleti adatait is megvizsgálva, hogy több változó figyelembevétele hozzájárul-e az előrejelzés javításához (a cikkben ennek elemzésére nem térünk ki).

AROME előrejelzések utófeldolgozása gépi tanulással módszerekkel

A gépi tanulás egyik eszköze a neurális hálózat, melynek alapjait már a 60-as években kidolgozták, ám akkoriban még nem állt rendelkezésre elegendően nagy adatmennyiség és megfelelő számítógépes kapacitás. Ma már szinte mindenhol megtalálhatók a neurális hálók, s a meteorológiában is alkalmazzák ezeket a módszereket pl. az előrejelzések készítése vagy utólagos korrekciója során. A gépi tanulás modelljére gondolhatunk úgy, mint egy bemeneti és kimeneti csatornákkal rendelkező, konfigurálható gépre. Ezeket a módszereket három nagy osztályba soroljuk. (1) A felügyelt tanulás (supervised learning) során a modell kimeneteleit az elvárt kimenettel összehasonlítva (pl. egy adott

képen kutya vagy macska látható) minimalizáljuk a hibát. (2) A nem felügyelt tanulás (unsupervised learning) esetén nincsenek elvárt kimenetek, ilyenkor a modell a bemeneti adatok belső struktúráját, felépítési logikáját, csoportosulását tanulja, találja meg (pl. azt, hogy mik a kutyákat ábrázoló képek jellemzőbb tulajdonságai). (3) A visszaigazolós tanulás (reinforcement learning) esetében sincsenek elvárt kimenetek, a modell bemenetekre adott válaszát értékeljük, és az értékelést maximalizáljuk a tanulás során (pl. egy útkereső algoritmusban a célhoz közelebb vivő lépést díjazzuk, míg a céltól távolabb vivőt büntetjük).

Az AROME előrejelzéseire kifejlesztett neurális hálózatok a felügyelt tanulás osztályába tartoznak. Ezek az emberi agyhoz és idegrendszerhez hasonlóan megtanulják adott bemenetre az előnyös választ. Építőköveik a neuronok, melyek között az információ adott irányban terjed. A bejövő információkat az egyes neuronok tanulható módon összesítik, majd továbbítják. Az elsődleges bemeneti adatok az AROME modell napi 8-szori szélsébség és sugárzás előrejelzései 15-perces felbontással. Az elvárt kimenetek az előrejelzés időtartamában mért 100-méteres szélsébség és felszínre érkező sugárzás. Az alábbiakban részletezett ún. egy pontos előrejelző modell típus előrejelzésenként minden 15-percenkénti időtávra külön tanul, így minden időtávra egy-egy almodellt kapunk, s ezek összessége teszi ki az egész intervallum utófeldolgozási modelljét. A globálsugárzás napi eloszlása miatt

az előrejelzések 4:30-tól 17:30 UTC-ig terjedő időszakában végeztünk utófeldolgozást. A modelltípus kétféle bemenet kezelésére alkalmas: alapesetben kizárólag az AROME előrejelzéseiből dolgozik, illetve az előrejelzés kezdetét megelőző időtartamból származó mérések fogadására is képes. A szélesség előrejelzések javításánál figyelembe vett változók a 2 m-es hőmérséklet, 100 m-es szélirány és szélesség, a sugárzásnál pedig a globálisugárzás és a 2 m-es hőmérséklet. Az előrejelzeni kívánt pont 2 rácsávolságnyi környezétekből is vettünk az adatokat, a szél esetében pedig a 80 és 120 m közötti 5 szint előrejelzését is megkapta a gépi tanulós modell (1. táblázat), ezáltal a változók térbeli eloszlásáról is információt nyerve. A módszer előrejelzéseket feldolgozó része a szélesség esetében egy speciálisan felépített 4, míg a sugárzás esetében 3 rejtett rétegű perceptron (multi-layer perceptron; MLP).

Az adathalmazt többnyire 2 vagy 3 részre osztják: (1) a modell a tanítóhalmazon tanul, mely általában a leghosszabb és a legtöbbféle mintát tartalmazza; (2) a validációs halmazon a modell generalizációját, vagyis általánosítási képességét méri (azaz azt, hogy a modell hogyan teljesít a tanításban részt nem vevő adatokon); (3) az opcionális teszhalmazon az optimális paraméterek hangolása után a kész modellt tesztelik. A teszhalmaz szerepe egy nagyobb szintű generalizáció mérése, ugyanis a validációs halmaz a modellparaméterek (akár kézi) optimalizálásán keresztül implicit módon kihat a tanulási folyamatra. Ezeknél a módszereknél ideális esetben többéves adatsoron történik a tanítás. A fejlesztéshez 1–2 év adatai álltak rendelkezésünkre, s ilyen rövid időszak esetén az időszak sorsolt 80%-a alkotta a tanítóhalmazt, a fennmaradó 20% pedig a validációs halmazt. A generalizáció hiteles mérése érdekében a tanítás, a validáció és az esetleges tesztelés független adathalmazokon történik. A tanítás során az átlagos négyzetes hibát (illetve annak négyzetgyökét; RMSE, azaz root mean squared error) optimalizáltuk, a hatást pedig az RMSE változás és az eredeti előrejelzés RMSE-nek hányadosával mértük.

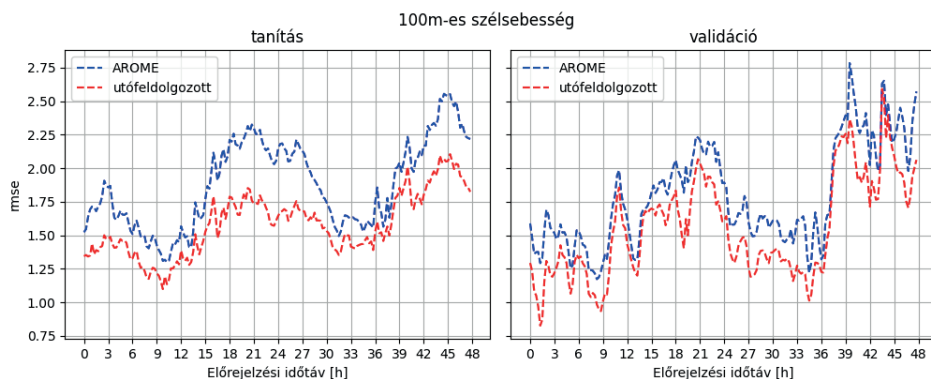
A 100-méteres szélességre vonatkozó nyers AROME előrejelzések eltérő jellemzőkkel bírnak a vizsgált pontokban (Deczki, 2021). A különbségek elsősorban a domborzati eltérésekből erednek: az A pontban szisztematikus albecslés, a B pontban fölébecslés,

a C pontban pedig nullához közeli átlagos hiba tapasztalható. Az RMSE az időszak jelentős részében 2 m/s alatti, és általában az éjszakai órákban 0,5–1 m/s-mal nagyobb, mint nappal. A gépi tanulós módszerek alkalmazásánál az esetek zömében a validációs hiba általában nagyobb, mint a tanítási időszakon (2. táblázat, 1. ábra). Ennek kézenfekvő magyarázata, hogy a modell paramétereit a tanító halmazon optimalizáljuk, a validációs halmaz a tanításban nem vesz részt, így rejtve marad a modell számára. A szélesség előrejelzés átlagos javulása a validációs halmazokon bemeneti mérésekkel 10%, míg bemeneti mérések nélkül 11% (2. táblázat). A bemeneti mérések nem javítottak szignifikánsan az előrejelzésen, melyet az időszak rövidsége okozhat. A pontszerű javulás mértéke függ a földrajzi helytől, pl. a B helyszín esetében elérte a 15%-ot, az A helyszín esetében kb. felekkora volt.

A globálisugárzás előrejelzéseket a partnerek méréseivel összevetve a D és E pontokban, az AROME fölébecsli a nagy (650 W/m² feletti) sugárzásértékeket; az RMSE nyáron a 250 W/m²-t is eléri, ősszel és télen

	A helyszín		B helyszín		C helyszín	
	mérésekkel	mérések nélkül	mérésekkel	mérések nélkül	mérésekkel	mérések nélkül
tanítás	8,8	5,6	15,6	11,3	12,0	11,3
validáció	7,0	7,2	13,8	15,0	9,3	10,8

2. táblázat. Relatív RMSE változás (%) az AROME 0 UTC-kor készülő 100-méteres szélesség előrejelzéseire az egyponos módszer mérésekkel, illetve anélkül történő alkalmazásával, 3 partneri mérési helyszínrre. A tanítás/validáció a 2020/4/15–2020/12/31 időszak 80/20%-án történt.



1. ábra. A 0 UTC-kor készülő nyers (kék) és utófeldolgozott (piros) 100 m-es szélesség előrejelzésekhez tartozó RMSE (m/s) a B helyszínrre az előrejelzési időtáv függvényében, méréseket is felhasználva a korrekció során, a tanítás (bal) és a validáció (jobb) időpontjainak átlagában.

		D helyszín		E helyszín	
		mérésekkel	mérések nélkül	mérésekkel	mérések nélkül
0 UTC	tanítás	20,7	16,7	11,4	7,8
	validáció	20,5	20,0	9,8	10,1
12 UTC	tanítás	19,5	17,3	8,9	10,1
	validáció	20,7	20,7	10,7	6,9

3. táblázat. Relatív RMSE változás (%) az AROME 0 és 12 UTC-kor készülő globálsugárzás előrejelzéseire az egyponos módszer mérésekkel illetve anélkül történő alkalmazásával, 2 partneri mérési helyszínrre. Az E helyszín 12 UTC-hez tartozó modelljei esetében a tanítási időszak 2019/2/2–2020/2/2, a validációs időszak 2020/2/2–2021/6/27. A többi esetben a 2020/4/17–2021/3/29 időszak 80/20 %-án történt a tanítás/validáció.

viszont alig haladja meg a 100 W/m²-t (Deczki, 2021). A verifikációt az OMSZ legközelebbi méréseivel is elvégezve viszont éppen a nagy értékek alábecslése jellemző, ami arra utal, hogy a partnerek mérőeszközei ebben a sugárzástartományban nem mérnek megfelelően. A 0 UTC-kor készített globálsugárzás előrejelzések esetében kb. 15% átlagos javulás érhető el a kifejlesztett módszerrel (3. táblázat). A mértéke ezúttal is függ helyszíntől: a D mérési pontban gyakran kétszeres javulást ért el az utófeldolgozás az E helyszínhez képest. Az E pontra végzett kísérletekben szimuláltuk az operatív megvalósítást (l. később): azaz az egyéves tanítási időszakot időben követte a validációs időszak. Utóbbi másfél évet ölelt fel, ez alatt a sugárzásértékek és a partneri mérésekhez viszonyított előrejelzési hibák magasabbak voltak, mint az éves átlag (2. ábra). Mindazonáltal a módszer – méréseket is használva – tudta tartani az előző évben elért javulást.

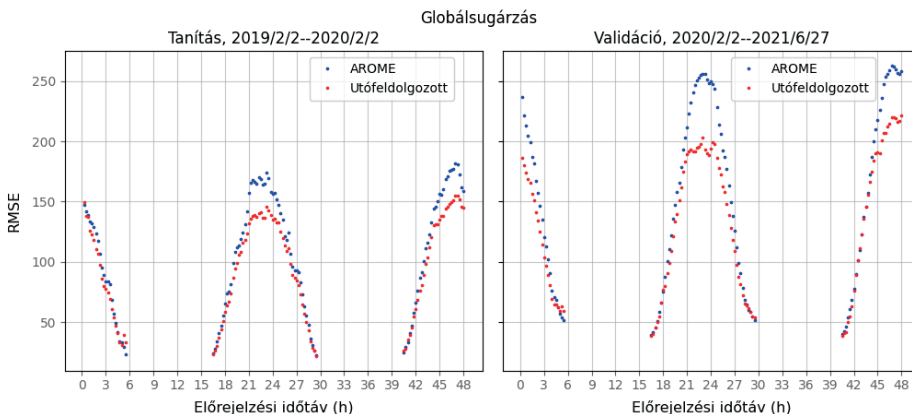
AROME-EPS előrejelzések utófeldolgozása paraméteres modellek segítségével

Az ensemble előrejelzések hibája javításának egy másik lehetséges módja az a statisztikai utófeldolgozás, ahol a paraméteres modellek az egyes légköri elemeket leíró eloszlást állítják elő. Nagyon egyszerű, de igen hatékony paraméteres technika a nem-homogén regresszió vagy ensemble model output statistics (EMOS; Gneiting et al., 2005), ahol az előrejelző eloszlás egyetlen paraméteres (pl. lognormális) eloszláscsaláddal adható meg, aminek paraméterei az ensemble előrejelzés előre megadott (pl. kontroll tagtól, ensemble szórástól függő) függvényei.

Az utófeldolgozás így lényegében ezen paraméterek becslése, amit múltbeli előrejelzés–megfigyelés párok (tanuló adatok) segítségével végeznek. Az egyes időjárási mennyiségekhez tartozó EMOS modellek csupán a vizsgált eloszláscsaládban és/vagy az előrejelzéseket az eloszlás paramétereivel összekötő függvényekben különböznek. Az EMOS modellezés általánosításaként fogható fel a distributional regression network (DRN; Rasp and Lerch, 2018), ahol ezeket a paraméteres összekötő függvényeket egy neurális háló helyettesíti, ami jóval általánosabb kapcsolatot enged meg az ensemble előrejelzések és az előrejelző eloszlás paraméterei között.

Mind az AROME-EPS 100 m-es szélesség előrejelzései esetén, mind pedig a globálsugárzás előrejelzéseknél 2 különböző előrejelző eloszlás illeszkedését vizsgáltuk. Szélességnél – annak nem-negatív volta miatt – a nullában alulról csonkított normális (truncated normal, TN; Thorarinsdottir and Gneiting, 2010) illetve a log-normális (LN; Baran and Lerch, 2015) eloszlásokon alapuló EMOS modellt teszteltük, valamint az ezen eloszlásokon alapuló DRN technikát is (TN DRN, LN DRN; Baran and Baran, 2021).

A globálsugárzás előrejelzések utófeldolgozásánál egy univerzális modellnek kezelnie kell az éjszakai órákat is, megadva a 0 W/m² sugárzás valószínűségét.



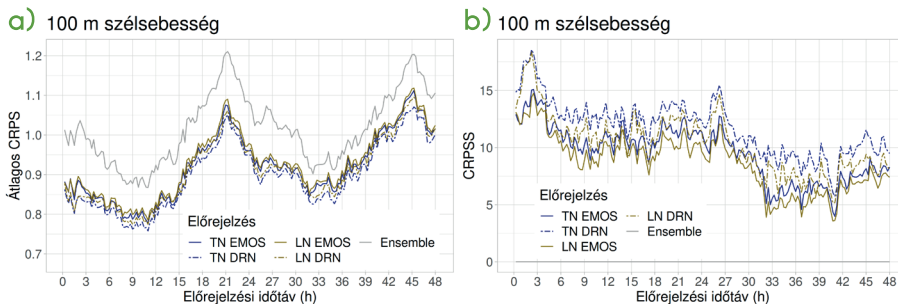
2. ábra. A 12 UTC-kor készülő nyers (kék) és utófeldolgozott (piros) globálsugárzás előrejelzéshez tartozó RMSE (m/s) az E helyszínrre az előrejelzési időtáv függvényében, méréseket is felhasználva a korrekció során, a tanítási (bal) és a validációs (jobb) időszak átlagában.

Erre a nullában alulról cenzorált logisztikus (CLO), illetve cenzorált normális (CN0) EMOS modellt (Schulz *et al.*, 2021) alkalmaztuk. Az említett EMOS modellek helyparamétere a kontroll tagtól, a perturbált tagok átlagától, valamint globálsugárzás esetén a 0 W/m^2 előrejelzések arányától, a skálaparamétere pedig az ensemble szórásától függ. A TN DRN és LN DRN modelleknél a hely- és skálaparamétereket előállító függvényeket egy egyetlen, 25 neuronból álló rejtett réteget tartalmazó MLP neurális háló helyettesíti, melynek bemenő adatai a kontroll tag, a 10 perturbált tag átlaga és szórásnégyzete.

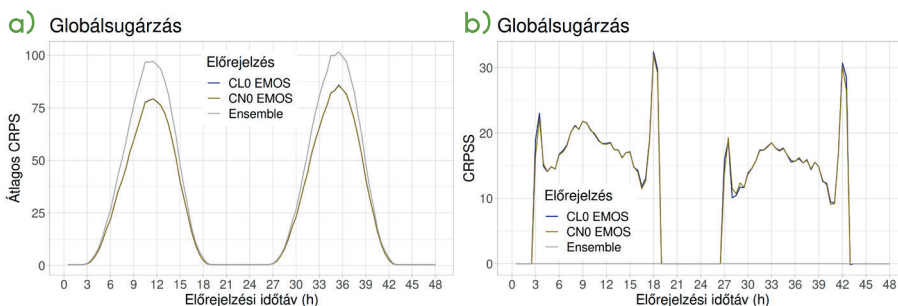
Mindkét vizsgált időjárási változó esetén gördülő tanulóperiódust használtunk, azaz egy adott időpontra vonatkozó előrejelző eloszlás paramétereit az azt megelőző napok előrejelzései és megfigyelései segítségével becsültük. Ez szél esetén 51, globálsugárzásnál pedig 31 napos tanulóperiódust jelentett (Baran and Baran, 2021; Schulz *et al.*, 2021). Az előbbi mennyiségnél lokális modelleket készítettünk mindhárom helyszínre, míg az utóbbinál az OMSZ 7 helyszínre (Aszódra, Budapestre, Debrecenre, Kecskemétre, Pécsre, Szegedre, Tápíószelére) vonatkozó 10-percenkénti mérési

adatait együtt kezeltük, egyetlen regionális modellt illesztve azokra. Ez a módszer lehetővé teszi az előrejelző eloszlások extrapolálását további helyszínekre (pl. egy tetszőleges napelemlémezés helyszínére). Ahogyan említettük, a partneri méréseket, azok gyengébb minősége miatt, itt már nem használtuk fel. A mérési és az előrejelzési adatok eltérő időbeli sűrűsége miatt a szélsőesség esetén 15-, a globálsugárzásnál pedig 30-perces időléptékű előrejelző eloszlásokat tudunk előállítani. Az EMOS modelleknél minden egyes időlépcsőt külön kezeltünk, míg a DRN modelleknél egy-egy hálót tanítottunk be a 0 és 24 óra közötti illetve a 24 és 48 óra közötti időtávra optimalizálva a tanulmányokon vett átlagos CRPS (continuous ranked probability score) értékét. A CRPS az ensemble előrejelzés és a mérési adatsor eloszlásfüggvényét hasonlítja össze, s minél alacsonyabb értéket vesz fel, annál jobbnak tekintjük az előrejelzett eloszlást. Kiszámítottuk a CRPSS (skill) mutatót is, ami a nyers ensemble előrejelzéshez viszonyítja a CRPS változását, s minél magasabb az értéke, annál nagyobb a javulás. Az egyes előrejelzéseket a 2020. július 1. – 2021. június 30. időszak adatain verifikáltuk és hasonlítottuk össze.

A 100-méteres szélsőesség esetén mindegyik utófeldolgozó modell jelentősen csökkentette a CRPS mutató értékét, de megtartotta annak napi jellemzőit (azaz az éjszakai magasabb értékeket; 3a. ábra). Az AROME-EPS előrejelzésektől való eltérések statisztikailag is szignifikánsak. Az egyes modellek egymáshoz való viszonya jobban megérthető a 3b. ábra CRPSS értékei segítségével, ahol látszik, hogy a DRN modellek mindegyik időlépcső esetén jobban teljesítettek az azonos eloszláson alapuló EMOS modellnél. A legjobban teljesítő TN DRN modell átlagos CRPS értéke az AROME-EPS átlagos CRPS értékének 88,3%-a, ezt követi az LN DRN (89,4%), TN EMOS (90,4%), LN EMOS (91,3%) és TGEV EMOS (91,7%).



3. ábra. A nyers (szürke) és utófeldolgozott (színes) 100 m-es szélsőesség előrejelzések átlagos CRPS értékei (m/s; bal) és a nyers előrejelzésekhez viszonyított CRPSS értékek (%) az előrejelzési időtáv függvényében.



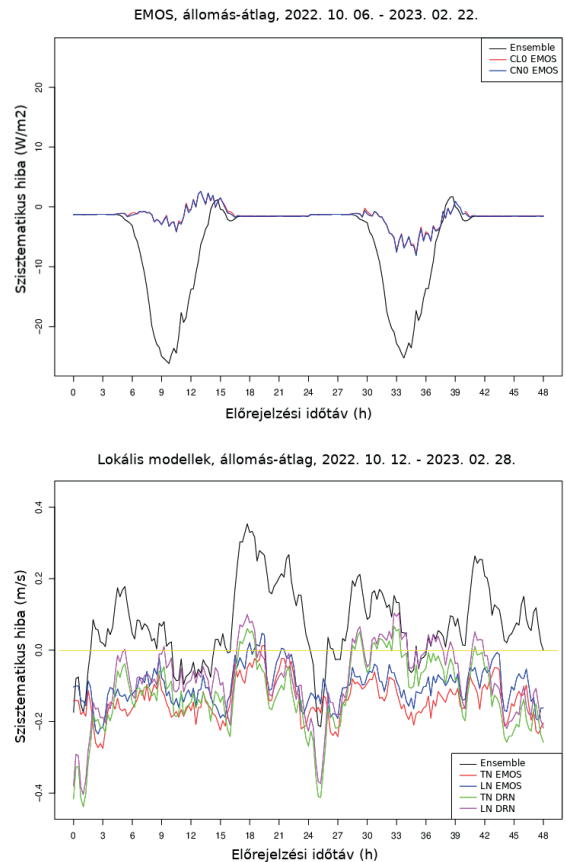
4. ábra. A nyers (szürke) és utófeldolgozott (színes) globálsugárzás előrejelzések átlagos CRPS értékei (W/m^2 ; bal) és a nyers előrejelzésekhez viszonyított CRPSS értékek (%) az előrejelzési időtáv függvényében.

A globálsugárzás utófeldolgozására használt CL0 és CN0 EMOS modellek teljesítménye között minimális az eltérés, a nyers előrejelzésekhez mért javulás azonban jóval nagyobb, mint a szélesség esetén (4. ábra). A CL0 EMOS átlagos CRPS értéke az AROME-EPS átlagos CRPS értékének százalékában 7,5 W/m² feletti megfigyelt globálsugárzásnál (azaz a nappali időszakokat figyelembe véve) 82,6%, a CN0 EMOS ugyanezen mutatója pedig 82,7%. Mindkét modell alkalmas arra, hogy kezelje az éjszakai időszakot, a CRPS értékekben történő kiugrásokat (3 és 18 UTC környékén) a kis értékekkel való osztás okozza. Teszteltük a regionális CL0 EMOS modell extrapolációját is a partneri sugárzásmérések helyszínére. Itt a javulás csupán 9,5%, de ezt okozhatják az itt telepített mérések pontatlanságai is.

Beépítés az operatív rendszerbe

Az AROME előrejelzések utófeldolgozására kifejlesztett egyponos módszert időről időre (pl. évente) újra kell tanítani annak érdekében, hogy az időjárás-előrejelző modellen időközben végbement fejlesztéseket figyelembe vegyünk. A műveletet minden egyes földrajzi pontra, meteorológiai változóra és előrejelzési időtávra el kell végezni, amire a módszert használni szeretnénk. Minél hosszabb a tanítási idősor, annál jobb eredményeket kaphatunk, ugyanakkor a rendkívüli mennyiségű adat mozgatása miatt jelenleg csak 1 évre szorítkoztunk, s az operatív implementáció előtt a 2021. szeptember 1. és 2022. augusztus 31. közötti időszakon tanítottuk újra a módszert. Az adatok előkészítése után a tanítás futtatása (CPU alapú számítógépen) 4-5 órát vesz igénybe, aminek nagy része a (NetCDF formátumú előrejelzési és ASCII formátumú partneri mérési) adatállományok megnyitásával és az adatok memóriába olvasásával telik. Ha végeztünk a tanítással, az operatív előrejelzések javítása már percek alatt elkészül. Ezt a 0 UTC-kor induló AROME futtatásra végezzük el, ennek során összegyűjtjük az aktuális AROME előrejelzést, valamint a futtatást megelőző 12 órából az OMSZ-os és partneri méréseket, amiket a módszer szintén használ a kezdeti korrigálás-hoz. Az eredményt ASCII állományként kapjuk meg, ami tartalmazza az előrejelzési időpontokra vonatkozó javított eredmény értékeket.

Az AROME-EPS előrejelzésekre minden bemutatott módszerrel egymás után lefut az utófeldolgozás a globálsugárzás esetében 7 OMSZ-os mérőállomás helyszínére, a 100-méteres szélességre pedig a 3 partneri mérési helyszínre. A gördülő tanuló periódus



5. ábra. Az ensemble átlag szisztematikus hibája globálsugárzásra (W/m²; fenn) 7 OMSZ-os mérési pontra és 100 m-es szélességre (m/s; lenn) 3 partneri mérési pontra a nyers (fekete) és a különböző módszerekkel (fenn: regionális CL0 EMOS – piros, CN0 EMOS – kék; lenn: lokális TN EMOS – piros, LN EMOS – kék, TN DRN – zöld, LN DRN – rózsaszín) utófeldolgozott AROME-EPS előrejelzésekre az előrejelzési időtáv függvényében.

előnye, hogy a módszereket nem kell előzetesen egy hosszú időszakon – számos adat mozgatásával – betanítani, azok folyamatosan, az operatív futtatás közben megtanulják a modellfejlesztéseket is. A tanítás akkor optimális, ha 31, illetve 51 napra visszamenően minden időpontra és állomásra rendelkezésre állnak a mérések és az előrejelzések. A napi adatok gyűjtése több időt vesz igénybe, mint maga az utófeldolgozás, ugyanis az OMSZ 10-perces globálsugárzás méréseit, illetve a rácsponti nyers előrejelzés értékeket (a megfelelő állomásokhoz legközelebbi rácspontokra) NetCDF formátumú, nagyméretű állományokból kell kinyerni. Ezért az eljárást az AROME-EPS 0 UTC-kor inicializált futásához úgy csatoltuk, hogy minden nap gyűjtse össze a fenti adatokat, így ezek elkészülte után az utófeldolgozást végző szerveren (8 CPU-n) néhány perc

alatt előáll a javított produktum is. Az erőművi termelésbecslés menetredezéséhez szükséges 15-perces gyakoriság miatt a 10-perces globálsugárzás méréseket a köztes időpontokra átlagoljuk. Mérési adat hiánya esetén is elkészítjük a javított előrejelzést, azonban folyamatosan figyeljük, mekkora az adathiány, mivel az befolyásolja a produktum minőségét. Ritkán előfordulhat, hogy az AROME-EPS előrejelzés nem fut le. Egy ilyen napon nem készül produktum, a későbbiekben pedig az illesztésből, illetve tanításból kihagyjuk ezt a napot, és egy nappal hátrébb lépünk az időben.

Rendszeresen készítünk verifikációs mutatókat és ábrákat, ezen felül havi szintű értékelést is az eredményekről. A 2022. októbertől 2023. januárig tartó 140-napos időszakra a CRPS-ben a várt mértékű javulást tapasztaltuk: a CN0 eloszlást használó EMOS módszerrel 7,5 W/m² feletti globálsugárzás értékekre az első napi előrejelzésekre átlagosan 11,8%, míg a második napra 7,2% a javulás. A CL0 eloszlásra ennél ugyan valamivel jobb eredményt kaptunk, azonban a CN0 EMOS numerikusan stabilabbnak bizonyult. A 100-méteres szélsőbességre a legnagyobb CRPS javulást a lokális tanítást használó TN DRN gépi tanulási módszer adta, mértéke az előrejelzés első napján 12,7%, a másodikon 9,4%. A módszerek hatékonysága látszólag kisebb, mint a fejlesztési fázisban, aminek több oka lehet: a vizsgált időszakból hiányzik a nyári szezon, amikor az ensemble tagok szórása, illetve maga a globálsugárzás nagyobb; a partneri szélmérésekben előfordulnak adathiányok; 2022 óta modellfejlesztések történtek az operatív AROME-EPS-ben. Az utófeldolgozásnak köszönhetően a szisztematikus hiba is csökkent a globálsugárzásra (5. ábra). Mindkét paraméter esetében nő az ensemble szórás, megközelítve az ensemble átlag RMSE-t, ezzel reprezentatívabb leírását adva az előrejelzési bizonytalanságnak.

Összefoglalás

Az AROME és az AROME-EPS globálsugárzás és 100-méteres szélsőbesség előrejelzéseinek javítására paraméteres és gépi tanuláson alapuló statisztikai utófeldolgozási módszereket fejlesztettünk ki. Ezek a rendelkezésre álló mérési és modell adatok felhasználásával hozzávetőlegesen 10–20%-kal javítják a nyers előrejelzések négyzetes hibáját, illetve az ensemble előrejelzések eloszlását. A bemutatott technikákat 2023. március óta tesztjelleggel alkalmazzuk az operatív előrejelzéseken. Az utófeldolgozás azonban nem helyettesíti az előrejelző modell folyamatos fejlesztését, ugyanis a fizikai folyamatok pontosabb leírása

az utóbbi révén valósítható meg, míg az előbbi segítségével egy-egy paraméter statisztikai jellemzőit közelítjük a mérésekhez. Ezek a legjobb eredményeket akkor adják, ha a felhasználó jó minőségű és nagy megbízhatósággal rendelkezésre álló mérési adatokat tud biztosítani a kívánt helyszínre vonatkozóan (például egy nap- vagy szélerőmű közelében), melyekkel az utófeldolgozási módszerek eredményesen taníthatók.

Irodalom

- Baran, S., Baran, Á., 2021: Calibration of wind speed ensemble forecasts for power generation. *Időjárás* 125, 4, 609–624. <https://doi.org/10.28974/idojaras.2021.4.4>
- Baran, S., Lerch, S., 2015: Log-normal distribution based Ensemble Model Output Statistics models for probabilistic wind speed forecasting. *Q. J. R. Meteorol. Soc.* 141, 2289–2299. <https://doi.org/10.1002/qj.2521>
- Deczki Z., 2021: Rövidtávú előrejelzések kiértékelése a megújuló energia célú felhasználásra. *MSc diplomamunka*, Eötvös Loránd Tudományegyetem, Budapest, Meteorológiai Tanszék.
- Gneiting, T., Raftery, A.E., Westveld, A.H., Goldman, T., 2005: Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Weather Rev.* 133, 1098–1118. <https://doi.org/10.1175/MWR2904.1>
- Haiden, T., Janousek, M., Vitart, F., Ben-Bouallegue, Z., Ferranti, L., Prates, F., Richardson, D., 2022: Evaluation of ECMWF forecasts, including the 2021 upgrade. *ECMWF Technical Memorandum 902*. <https://doi.org/10.21957/xqnu5o3p>
- Jávorné Radnóczy, K., Várkonyi, A., Szépszó, G., 2020: On the way towards the AROME nowcasting system in Hungary. *ALADIN-HIRLAM Newsletter* 14, 65–69.
- Rasp, S., Lerch, S., 2018: Neural networks for postprocessing ensemble weather forecasts. *Mon. Weather Rev.* 146, 3885–3900. <https://doi.org/10.1175/MWR-D-18-0187.1>
- Schulz, B., El Ayari, M., Lerch, S., Baran, S., 2021: Post-processing numerical weather prediction ensembles for probabilistic solar irradiance forecasting. *Sol. Energy* 220, 1016–1031. <https://doi.org/10.1016/j.solener.2021.03.023>
- Thorarindottir, T.L., Gneiting, T., 2010: Probabilistic forecasts of wind speed: Ensemble model output statistics by using heteroscedastic censored regression. *J. R. Stat. Soc. A.* 173, 371–388. <https://doi.org/10.1111/j.1467-985X.2009.00616.x>
- Tóth, H., Homonnai, V., Mile, M., Várkonyi, A., Kocsis, Zs., Szanyi, K., Tóth, G., Szintai, B., Szépszó, G., 2021: Recent developments in the data assimilation of AROME/HU numerical weather prediction model. *Időjárás* 125, 4, 521–553. <https://doi.org/10.28974/idojaras.2021.4.1>