

Doktori (PhD) értekezés tézisei

Diszkrét és folytonos változókra vonatkozó
ensemble előrejelzések egy- és többdimenziós
utófeldolgozása

Nagy-Lakatos Mária

Témavezető: Prof. Dr. Baran Sándor



DEBRECENI EGYETEM

Informatikai Tudományok Doktori Iskola

Debrecen, 2025

Tartalomjegyzék

Bevezetés	1
1. Az értekezés tézispontjai	3
2. Ensemble előrejelzések kalibrálása	5
2.1. Klaszterezés alapú térbeli interpoláció	5
2.2. Látástávolság utófeldolgozása	9
2.3. Többváltozós utófeldolgozás	12
Irodalomjegyzék	15
Publikációs lista	19

Bevezetés

Az időjárás kaotikus természete miatt a numerikus időjárás-előrejelző (NWP) modellek kezdeti feltételeiben bekövetkező apró eltérések gyorsan felerősödhetnek, ami jelentős hatással lehet az előrejelzések pontosságára. Emellett egyetlen NWP kimenet nem ad információt az előrejelzés bizonytalanságáról. Az ensemble előrejelző rendszerek bevezetése, melyek eltérő kezdeti feltételekkel és/vagy modellparaméterezéssel előrejelzések egy családját állítják elő, lehetővé tette a valószínűségi időjárás-előrejelzést (Gneiting és Raftery, 2005). Azonban az ensemble előrejelzések generálása meglehetősen számításigényes, különösen nagyobb elemszám és magasabb térbeli felbontás esetén. A Középtávú Időjárás Előrejelzések Európai Központja (ECMWF) jelentős előrelépést ért el az ensemble előrejelzések felbontásának finomításában: míg 1992-ben a felbontás még 120 km volt, 2023-ra ez 9 km-re csökkent¹.

A folyamatos fejlesztések ellenére az NWP modellek tartalmazhatnak szisztematikus hibákat, és alul- vagy túlbecsülhetik a tényleges megfigyeléseket. A 2000-es évek elején megjelent statisztikai utófeldolgozó eljárások, mint az ensemble model output statistics (EMOS; Gneiting *et al.*, 2005) és a Bayes modell átlagolás (BMA; Raftery *et al.*, 2005) ezeket a hiányosságokat úgy küszöbölik ki, hogy az ensemble előrejelzéseket előrejelző eloszlásokra képezik le. Az ensemble előrejelzések utófeldolgozása továbbra is fejlődő terület, amelyben egyre inkább helyet kapnak a gépi tanulási módszerek is (Rasp és Lerch, 2018).

A legtöbb utófeldolgozó módszert azonban egymástól függetlenül alkalmazzák minden egyes helyszínre, időpontra és időjárási mennyiségre, ami így nem veszi figyelembe az ezen dimenziók közötti összefüggéseket. A többváltozós utófeldolgozás célja éppen ezért az esetlegesen elveszett függőségek visszaállítása, ezzel javítva az előrejelzések megbízhatóságát. A többváltozós függőségi viszonyok helyreállítására már a 2010-es évek óta alkalmazzák a paraméteres és empirikus kopulákat (Scheffzik *et al.*, 2013; Möller *et al.*, 2013) valamint az együttes eloszlások közvetlen modellezését, mely módszerek azonban kombinálhatóak is (Scheffzik, 2016).

A disszertáció különböző többváltozós utófeldolgozó módszerek összehasonlítására fókuszál, kitöltve az egyes módszerek valós adatokon történő értékelésére vo-

¹<https://www.ecmwf.int/en/about/who-we-are/history>

natkozó rést. A hangsúly a kétlépcsős módszereken van, ahol a kezdeti egyváltozós utófeldolgozást követően a korrelációs struktúrák helyreállítása empirikus kopulák alkalmazásával történik. További célunk a látástávolság előrejelzések statisztikai utófeldolgozása, mely mennyiség előrejelző képessége alulmúlja más változók teljesítményét, valamint annak vizsgálata, hogyan javíthatják a Copernicus Atmosphere Monitoring Service (CAMS) kategorikus látástávolság-előrejelzései az utófeldolgozó modellek teljesítményét, és hogy hogyan járulhatnak hozzá a térbeli függőségek újradefiniálásához (Baran és Lakatos, 2023; Lakatos és Baran, 2024).

Ezen kívül a Lerch és Baran (2017) által kidolgozott szemi-lokális utófeldolgozó módszert kiegészítjük egy interpolációs lépéssel, és a kapott eljárást az ECMWF globális 10 méteres szélesség előrejelzéseinek teszteljük olyan helyszíneken, ahol a közvetlen modellezés a validáló megfigyelések hiánya miatt lehetetlen lenne.

A dolgozat felépítése a következő: Az 1. fejezet áttekintést ad az ensemble előrejelző rendszerekről és a meteorológiai változók kalibrálására szolgáló paraméteres EMOS modellekről, kitérve a diszkrét megfigyelési tartománnyal rendelkező változók utófeldolgozására szolgáló rendezett logisztikus regresszió (POLR; proportional odds logistic regression) és a többretegű perceptron (MLP; multi-layer perceptron) osztályozó algoritmusokra. A 2. fejezet részletesen tárgyalja a munkánk során tesztelt kétlépcsős többváltozós utófeldolgozási módszereket, majd a 3. fejezet a széles körben használt illeszkedési mutatókat és a tanítási folyamatot ismerteti. A 4. fejezet a Lerch és Baran (2017) által kidolgozott szemi-lokális utófeldolgozó módszer interpolációs lépéssel való kiterjesztését ismerteti, majd a bemutatott modellek teljesítményét mutatja be az ECMWF szélesség adatain (Baran és Lakatos, 2024). Az 5. fejezet a látástávolság előrejelzések statisztikai utófeldolgozásával foglalkozik az *EUPPBench* benchmark adathalmazon, valamint ECMWF és CAMS előrejelzéseket felhasználva. Itt kitérünk a térbeli függőségek helyreállítására is, illetve megvizsgáljuk, hogy a CAMS előrejelzések bevonása az utófeldolgozás folyamatába hogyan befolyásolja az alkalmazott módszerek előrejelző képességét. A 6. fejezet az ECMWF előrejelzéseinek hasonlítja össze a különféle többváltozós utófeldolgozó módszereket (Lakatos *et al.*, 2023), végül a 7. fejezet összegezi az eredményeket és további kutatási irányokat javasol.

Összefoglalva, a dolgozat bemutatja, hogy az egy- és többváltozós statisztikai utófeldolgozás kulcsfontosságú az ensemble előrejelzések javításában, és különböző kalibrálási technikákon keresztül szemlélteti, hogyan növeli az időjárási előrejelzések pontosságát és megbízhatóságát.

1. fejezet

Az értekezés tézispontjai

- **1. tézis**

A paraméteres előrejelző eloszlások interpolálására kifejlesztett módszer kiküszöböli a Gauss-mezők becslésének vagy az állomás-specifikus információk alkalmazásának számítási terhét. A módszer kizárólag az EMOS modellezéshez szükséges adatokra épít, és könnyen adaptálható tetszőleges paraméteres utófeldolgozó technikához.

- **2. tézis**

A megfigyelt helyeken a lokális és a megfigyelés-alapú szemi-lokális módszerek nyújtják a legjobb eredményeket. Ezzel szemben a megfigyelésekkel nem rendelkező helyeken az előrejelzés-alapú szemi-lokális módszerek jobban teljesítenek: a regionális modellekhez képest élesebb előrejelzési intervallumokat, alacsonyabb átlagos continuous ranked probability score (CRPS) értékeket, valamint a medián előrejelzések esetében kedvezőbb átlagos abszolút hibát (MAE) biztosítanak.

Lásd a doktori értekezés 4. fejezetét.

Kapcsolódó cikk: Baran és Lakatos (2024)

- **3. tézis**

Az utófeldolgozott előrejelzések – különösen a POLR-t alkalmazók – felülmúlják a nyers ECMWF ensemble előrejelzéseket az átlagos LogS és RMSE mutatók tekintetében. A POLR emellett az MLP-nél is jobban teljesít az egyváltozós, diszkrét látástávolság-előrejelzések utófeldolgozása során, különösen lokális tanítás esetén. A CAMS látástávolság-előrejelzések kovariánsként való bevonása tovább növeli a teljesítményt: az átlagos CRPS akár 5,25%-kal, az átlagos LogS pedig 1,37%-kal javul, míg alacsony látástávolságú

események esetén a Brier-pontszám akár 18,59%-kal is kedvezőbb lehet.

- **4. tézis**

A kétlépcsős többváltozós utófeldolgozó módszerek, mint az ensemble copula coupling (ECC) és a Schaake shuffle (SSh), sikeresen állítják helyre a látástávolság előrejelzések térbeli függőségeit, és felülmúlják a függetlenséget feltételező referencia modelleket. A CAMS előrejelzésekkel való kiterjesztés tovább javítja az energy score-t (0,58%–4,11%) és pontosítja a korrelációs struktúra leírását.

Lásd a doktori értekezés 5. fejezetét.

Kapcsolódó cikkek: Baran és Lakatos (2023); Lakatos és Baran (2024)

- **5. tézis**

A korszerű, többváltozós utófeldolgozási módszerek széles skálája javítja az előrejelzések minőségét a nyers ensemble előrejelzésekhez és az egyváltozós modellekhez képest, azonban a különböző módszerek közötti teljesítménykülönbségek minimálisak. Kiemelendő, hogy a kvantilisalapú mintavételezés (ECC-Q és SSh-Q) kissé jobb eredményeket hoz, mint a véletlenszerű mintavételezés (ECC-R és SSh-R), és az ECC-Q megközelítés robusztus, hatékony és számítási szempontból kedvező módszerként emelkedik ki. A fejlettebb eljárások, mint a dual ensemble copula coupling (dECC), minimum divergence Schaake shuffle (mdSSh) és hasonlóság alapú Schaake shuffle (simSSh) egyike sem nyújt számottevő javulást az ECC-Q-hoz képest, megerősítve, az utóbbi optimálisabb választás lehet a komplexebb modellekkel szemben.

Lásd a doktori értekezés 6. fejezetét.

Kapcsolódó cikkek: Lakatos *et al.* (2023)

2. fejezet

Ensemble előrejelzések statisztikai utófeldolgozása

2.1. Paraméteres utófeldolgozó modellek klaszterezés alapú térbeli interpolációja

Az 1. és 2. tézishez kapcsolódó esettanulmány

A disszertációban lásd: 4. fejezet

Publikáció: Baran és Lakatos (2024)

Az utófeldolgozó eljárások, mint például az EMOS, múltbéli előrejelzés-megfigyelés párokra támaszkodnak, ami a kalibrációt a megfigyelt helyszínekre korlátozza. Korábbi tanulmányok kimutatták, hogy a regionális paraméteres utófeldolgozás, majd az azt követő interpoláció 20%-kal javítja az előrejelzések pontosságát a megfigyelt és 11%-kal a nem megfigyelt helyszíneken (Baran és Baran, 2024). Azonban ha nagy az ensemble előrejelzés által lefedett tartomány, a változó éghajlati viszonyok és előrejelzési hibák korlátozhatják a regionális EMOS hatékonyságát. A disszertációban és a kapcsolódó cikkben mi egy olyan szemi-lokális utófeldolgozási módszert mutatunk be, amely Lerch és Baran (2017) munkájára építve ötvözi a klaszterezést és az interpolációt. Az interpolációs módszereket, amelyeket a dolgozat 4.1. szakasza részletez, az ECMWF 52 tagú szélsőbesség ensemble előrejelzésein teszteljük, 674 európai, ázsiai és ausztráliai SYNOP állomásnak a 2020. január 1. és 2023. április 30. közötti időszakra vonatkozó adatai alapján. Az előrejelzések 0000 UTC-kor lettek inicializálva, és az 1-től 10 napig terjedő előrejelzési horizontokat vizsgáljuk. Az adathalmaz tartalmazza a nagy felbontású (HRES; 9 km-es rács), a kontroll (CNTL; 18 km-es rács) előrejelzéseket, valamint az 50 ensemble tagot (ENS; 18 km-es rács), amelyeket perturbált kezdeti feltételekkel generáltak.

Ezek a tagok felcserélhetőnek tekinthetők, mivel statisztikailag megkülönböztethetetlenek. Az állomások 75%-át (506 állomás) a paraméterek becslésére használjuk, és az így kapott modelleket interpoláljuk a fennmaradó 168 állomásra, ahol a hiányzó adatok aránya 1,65%.

Változószelekciós módszerek

A Lerch és Baran (2017) által bemutatott szemi-lokális EMOS modell a megfigyelt helyszínek csoportosítására k -közép klaszterezést alkalmaz az egyes állomások éghajlati jellemzői és/vagy az ensemble átlagnak a megfigyelésekhez viszonyított hibái alapján. Ennek eredményeképpen az így kialakított klaszterek egy-egy közös EMOS paraméterhalmazon osztoznak. Az alábbiakban összefoglaljuk a Lerch és Baran (2017) által bemutatott, megfigyelés-alapú változószelekciós módszert, és ismertetjük az előrejelzés-alapú változószelekciós módszereinket, amelyek kiterjesztik a szerzők munkáját azáltal, hogy lehetővé teszik a prediktív eloszlások interpolációját megfigyelésekkel nem rendelkező helyszínekre.

Megfigyelés-alapú kvantilis jellemzők (OQ)

Egy állomás OQ jellemzői a tanuló időszak megfigyelt szélsőségei, valamint az ensemble átlag előrejelzési hibái empirikus eloszlásfüggvényeinek ekvidisztáns kvantiliseiből állnak (Lerch és Baran, 2017). Mivel azonban ezek a jellemzők megfigyelésekre támaszkodnak, a megfigyelésekkel nem rendelkező helyszínek meglévő klaszterekhez való hozzárendelése kihívást jelent.

Előrejelzés-alapú kvantilis jellemzők (FQ)

Ez a jellemzőkészlet az előrejelzés-alapú analógja a fentebb bemutatott, megfigyelés-alapú konstrukciónak. Jelölje $\hat{H}_{i,n}^{(1)}$, $\hat{H}_{i,n}^{(2)}$, $\hat{H}_{i,n}^{(3)}$ és $\hat{H}_{i,n}^{(4)}$ az i állomás esetén az f_{HRES} nagy felbontású előrejelzések, az f_{CNTL} kontroll előrejelzések, valamint az 50 felcserélhető ensemble tag $\bar{f}_{\text{ENS}}$ átlagának és S_{ENS} szórásának empirikus eloszlásfüggvényét az n napból álló tanuló periódusra vonatkozóan. Ebben az esetben az i állomásra vonatkozó N dimenziós jellemzővektort a $\hat{H}_{i,n}^{(j)}$ ekvidisztáns kvantilisei alkotják: $\frac{1}{N_j+1}, \frac{2}{N_j+1}, \dots, \frac{N_j}{N_j+1}$, $j = 1, 2, 3, 4$, ahol $N_1 = N_2 = N_3 = \lfloor \frac{N}{4} \rfloor$ és $N_4 = N - N_1 - N_2 - N_3$.

Előrejelzések, mint jellemzők (FF)

Az ekvidisztáns kvantilisok helyett, minden i helyszínen és n napból álló tanuló időszak esetén tekintjük egyszerűen az f_{HRES} , f_{CNTL} , $\bar{f}_{\text{ENS}}$ és S_{ENS} előrejelzéseknek (az n hosszú tanuló időszak nagy felbontású és kontroll előrejelzései, valamint a megfelelő felcserélhető tagok n -dimenziós átlag- és szórásvektorai) a felsorolás sorrendjében történő konkatenációjával kapott $4n$ -dimenziós vektort. Természetesen ez az egyszerű megközelítés meglehetősen magas dimenziójú jellemzővektorokat eredményezhet, ami növeli a k -közép klaszterezés számítási költségét. Azonban az ehhez kapcsolódó többletköltség általában kisebb, mint a előrejelzés-alapú kvantilis jellemzők kiszámításának költsége.

Interpolációs algoritmus

Mivel az előrejelzés-alapú jellemzők mind a megfigyelt, mind pedig a nem megfigyelt helyszíneken rendelkezésre állnak, a disszertáció 4.1. szakasza ismerteti az interpolációs algoritmust, amelyet a 4.1-es folyamatábra is szemléltet, illetve röviden bemutatunk a következőkben.

- 1. lépés: Végezzük el a megfigyelt helyszínek szemi-lokális EMOS modellezését (bármely fentebb felsorolt változóselektációs módszert felhasználva), meghatározva ezzel a klasztereket.
- 2. lépés: Minden egyes megfigyelt állomáshoz rendeljük hozzá az FF vagy QF jellemzőjét (ugyanazt a típust mindenütt) és számítsuk ki a választott jellemző átlagait az egyes klasztereken (klaszterközepek).
- 3. lépés: Határozzuk meg a 2. lépésben használt előrejelzés-alapú jellemzőt a nem megfigyelt helyszínekre.
- 4. lépés: Rendeljük minden egyes nem megfigyelt helyszínt a hozzá legközelebbi klaszterhez az adott helyszín jellemző vektora (FF/FQ jellemzők) és a klaszterközepek közötti Euklideszi távolság alapján.
- 5. lépés: A nem megfigyelt helyszínek előrejelző eloszlásainak meghatározásához használjuk a hozzájuk rendelt klaszterek EMOS paramétereit.

Eredmények

A megfigyelt helyszíneket 20 klaszterbe soroltuk, majd a nem megfigyelt helyszíneket az előrejelzések jellemzőkként történő felhasználásával rendeltük klaszterekhez, mivel ez a megközelítés pontosabb modelleket eredményez, mint az előrejelzés-alapú kvantilis jellemzők alkalmazása.

Referencia utófeldolgozási módszerként a regionális csonkított normális (TN) EMOS modellt alkalmazzuk (lásd a disszertáció 1.2.2. szakaszát), mely modell paramétereinek becslését a megfigyelt helyszínek adataival végezzük, majd az eloszlásokat interpoláljuk a nem megfigyelt helyszínekre. További referenciaként a nyers ECMWF ensemble előrejelzések szolgálnak. Minden EMOS modellt egy 60 napos gördülő tanuló periódussal tanítunk. Eredményeink azt mutatják, hogy a lokális és megfigyelés-alapú szemi-lokális módszerek a legjobban a megfigyelt helyszíneken teljesítenek, míg az előrejelzés-alapú szemi-lokális megközelítések a nem megfigyelt helyszíneken felülmúlják mind a megfigyelés-alapú, mind pedig a lokális módszereket (lásd a disszertáció 4.3. ábráját). A fenti eredményeket alátámasztja a disszertáció 4.4. ábrája is, amely az 1, 4, 7 és 10 napos előrejelzési horizontokon mutatja mind a megfigyelt, mind pedig a nem megfigyelt helyszíneken azon állomások arányát, ahol a szimultán DM teszt 5%-os szinten szignifikáns eltérést mutat az egyes módszerek átlagos CRPS értékei között. A regionális modellekkel összehasonlítva az előrejelzés-alapú szemi-lokális modellek élesebb előrejelzési intervallumokat, jobb átlagos CRPS-t és alacsonyabb MAE értéket adnak a medián előrejelzések esetében (lásd a disszertáció 4.6., 4.3. és 4.8. ábráit). Ezek az eredmények arra is rámutatnak, hogy nincs olyan univerzális utófeldolgozó stratégia, amely minden helyzetben ideális lenne. A megfigyelt helyszíneken a lokális vagy megfigyelés-alapú szemi-lokális módszerek előnyösebbek, míg a nem megfigyelt helyszíneken az előrejelzés-alapú szemi-lokális módszerek bizonyultak hatékonyabbnak. Az előrejelzés-alapú módszerek sikerét az magyarázza, hogy a jellemzővektorok hasonló jellegű mennyiségekből származó adatokat kapcsolnak össze. A szemi-lokális módszerek 6, 8 és 9 m/s küszöbértékekhez tartozó threshold-weighted continuous ranked probability skill score (twCRPSS) értékeit a regionális EMOS megközelítéssel hasonlítjuk össze. A megfigyelt helyszíneken a rangsorok összhangban vannak a continuous ranked probability skill score (CRPSS) alapján meghatározottakkal (lásd a disszertáció 4.5a, c, e ábráit), míg a nem megfigyelt helyszíneken az egyes modellek magas szélességeken vett teljesítménye (lásd a disszertáció 4.5b, d, f ábráit) eltér az összesített trendektől (lásd a disszertáció 4.3d ábráját). Bár az OQ előrejelzések az átlagos CRPS-ben lemaradnak, magasabb küszöbértékek esetén jobb twCRPSS értékeket mutatnak, mint a többi versengő EMOS modell. A disszertáció 4.7. ábráján látható probability integral transform (PIT) hisztogramok azt mutatják, hogy az utófeldolgozott előrejelzések kalibráltabbak a nyers ECMWF ensemble predikcióknál, valamint javítanak az alacsony szóráson és a torzításon. A regionális és szemi-lokális FQ modellek PIT értékei bírnak a legkisebb (az egyenletes eloszlás által elvárt 0,083-hoz legközelebbi) varianciával, míg a lokális és megfigyelés-alapú (OQ) modellek a nem megfigyelt helyszíneken továbbra is alacsony szórásúak maradnak.

2.2. Látástávolság előrejelzések statisztikai utófeldolgozása

A 3. és 4. tézishez kapcsolódó esettanulmányok

A disszertációban lásd: **5. fejezet**

Publikációk: Baran és Lakatos (2023); Lakatos és Baran (2024)

A Meteorológiai Világszervezet (WMO) a látástávolságot úgy határozza meg, mint „a legnagyobb távolságot, amely esetén egy megfelelő méretű fekete tárgy (földön elhelyezve) látható és felismerhető a horizonton” (WMO, 1992). A pontos látástávolság előrejelzések gazdasági és biztonsági előnyökkel járnak, azonban a látástávolság előrejelzése továbbra is kihívást jelent. A numerikus időjárás-előrejelző (NWP) modellek jellemzően a kapcsolódó időjárási változókat, például a páratartalmat, csapadékmennyiséget és aeroszol-koncentrációt használják a látástávolság becslésére, de a köd kialakulása és a mikrofizikai folyamatok bonyolítják az előrejelzéseket. A látástávolságot diszkrét értékekben jelentik, és hasonlóan a felhősödés előrejelzésekhez, a látástávolság előrejelzése is kezelhető osztályozási problémaként. Olyan módszerek, mint a rendezett logisztikus regresszió (POLR) és a többrétegű perceptron (MLP) hatékonyan alkalmazhatók erre a feladatra (részletekért az említett osztályozókról lásd a disszertáció 1.3. szakaszát).

Kutatásaink során az *EUPPBench* adatbázis és az ECMWF látástávolság ensemble előrejelzéseinek POLR és MLP módszerekkel történő kalibrálására összpontosítunk, és azt vizsgáljuk, hogy miként javíthatja a CAMS látástávolság adatainak felhasználása az ECMWF előrejelzések pontosságát, megőrizheti-e az állomások közötti térbeli összefüggéseket, valamint javíthatja-e az alacsony látástávolságú események előrejelzését. A meteorológiai állomások közötti térbeli függőségek visszaállítására a kétlépcsős ensemble copula coupling (ECC) és a Schaake shuffle (SSh) módszereket alkalmaztuk, amelyeket a disszertáció 2.1 és 2.3 szakaszaiban mutatunk be.

A *EUPPBench* adatbázis 2017 és 2018 éveket fed le, és az ECMWF összesen 42 német és francia SYNOP állomásra vonatkozó 0000 UTC-kor inicializált nagyfelbontású és 51 tagú operatív ensemble előrejelzéseit tartalmazza (kontroll és 50 felcserélhető tag), 120 órás előrejelzési horizonttal (6 órás időbeli felbontással). A második adatbázis 51 tagú ECMWF látástávolság előrejelzéseket tartalmaz (nagyfelbontású előrejelzések nélkül), 2020-2021 között, és 30 SYNOP állomást fed le Németországban, Csehországban és Lengyelországban. Az előrejelzéseket 0000 UTC-kor inicializálták és 40 előrejelzési horizont (6-240 óra) predikciói voltak elérhetőek, 6 órás lépésekkel. A CAMS előrejelzések, amelyek ugyanazokra az állomásokra vonatkoznak, de legfeljebb 120 órás előrejelzéssel, szintén szerepelnek

esettanulmányunkban. Épp ezért, a disszertáció 5.3.2. szakaszában az előrejelzések utófeldolgozása a 120 órás tartományra van korlátozva. Az ECMWF előrejelzések 240 óráig történő kalibrálásának részletei a disszertáció B mellékletében találhatóak. A modellezés során három térbeli szelekciós megközelítést alkalmazunk: lokális (MLP-L és POLR-L), regionális (MLP-R és POLR-R), valamint klaszterezés alapú szemi-lokális (MLP-C és POLR-C) módszereket. Az utóbbi megközelítés esetén a klaszterek háromdimenziós jellemzővektorokon alapulnak, amelyek a tanuló időszakon alapuló megfigyelt gyakoriságokat tükrözik a 0–5000 m, 5000–30000 m és 30000–70000 m intervallumokban. A POLR és MLP modellek bemeneti jellemzőinek kiszámítása előtt először normalizáljuk az előrejelzéseket, amihez minden előrejelzési értéket elosztunk 70000-el. Ezután az alábbi prediktorokat vesszük figyelembe: az ensemble $\tilde{f}_{CNTL} := f_{CNTL}/70000$ (normalizált) kontroll tagját, az 50 (normalizált) felcserélhető tag $\tilde{f}_{ENS}$ átlagát, az 51 tagú (normalizált) operatív ensemble s^2 szórását, az EUPPBench adatbázis esetén az $\tilde{f}_{HRES} := f_{HRES}/70000$ (normalizált) nagyfelbontású előrejelzést, valamint a látástávolságot 1000 m-ig, 1000–2000 m között, illetve 30000 m felett előrejelző ensemble tagok p_1 , p_2 és p_3 arányait. Amikor a CAMS előrejelzések hatását elemezzük az osztályozó módszerek teljesítményére, akkor a normalizált $\tilde{f}_{CAMS} := f_{CAMS}/70000$ CAMS előrejelzést is hozzáadjuk a bemeneti jellemzőkhöz. Mindezekon túl, Dabering *et al.* (2017) javaslatai alapján a szezonális változásokat úgy kezeljük, hogy a $\sin(2\pi d/365)$ és $\cos(2\pi d/365)$ függvényeket hozzáadjuk a bemeneti jellemzőkhez, ahol d az év napját jelöli. Fontos megjegyezni, hogy a normalizálás ahhoz szükséges, hogy az összes bemeneti jellemző azonos nagyságrendű legyen, ami numerikusan stabilabbá teszi a modellezést.

Eredmények

A disszertáció 5. fejezete két tanulmány eredményeit összegzi, amelyek különböző adatállományokat alkalmazva a látástávolság előrejelzések statisztikai utófeldolgozására összpontosítanak. Az eredmények azt mutatják, hogy az átlagos CRPS, az átlagos LogS és az RMSE mutatók tekintetében mind a klimatológiai, mind pedig az utófeldolgozott előrejelzések jelentősen jobban teljesítenek, mint a nyers ensemble előrejelzések, azonban az előrejelzési horizont növekedésével a teljesítménybeli különbség csökken (az ECMWF eredményekért lásd a disszertáció B mellékletét, az EUPPBench eredményekért pedig a 5.3.1-es szakaszt). A POLR modellek jobb teljesítményt mutatnak az MLP megfelelőikhez képest, megerősítve Baran *et al.* (2021) felhősödés előrejelzésekkel kapcsolatos megállapításait. A lokális POLR modell kiemelkedik az utófeldolgozó módszerek között, míg a gépi tanulási megközelítésekben a szemi-lokális modellezés gyakran felülmúlja a lokális becslést (Baran és Lakatos, 2023). A második tanulmány Lakatos és Baran (2024) munkáját folytatva a látástávolság helyspecifikus diszkrét utófeldolgozását vizsgál-

ja. Három fő kérdést érintünk: először is azt, hogy a többváltozós utófeldolgozó megközelítések képesek-e helyreállítani a látástávolság ensemble előrejelzéseinek egyváltozós kalibrálása során elveszett térbeli függőségeket; másodsor, hogyan befolyásolja a CAMS látástávolság előrejelzések, mint további magyarázó változók, bevonása az előrejelzések pontosságát; végül pedig azt, hogy az utófeldolgozott előrejelzések hogyan teljesítenek a repülésmeteorológia területén. A legjobb teljesítményű egyváltozós megközelítések (MLP-C és POLR-L) esetében a CAMS előrejelzések integrálása jelentős javulást eredményezett, 1%-tól 5,25%-ig terjedően az átlagos CRPS mutatóban, és 0,27%-tól 1,37%-ig az átlagos LogS mutatóban, de az előnyök csökkentek, ahogy az előrejelzési horizont nőtt. Az alacsony látástávolságú események, amelyek kulcsfontosságúak a repülésmeteorológiai alkalmazásokban, szintén szerepeltek az elemzésünkben. A CAMS-kibővített POLR-L modell alkalmazásával azt találtuk, hogy a Brier score átlagos javulása olyan körülmények között, amikor a látástávolság 200 m, 500 m, 1000 m és 3000 m alá esett, 14,22%-tól 18,59%-ig terjedt.

Ezen kívül a kétlépcsős ECC és SSh módszerekkel elvégeztük a 30 SYNOP állomás előrejelzéseinek együttes kalibrálását, kihasználva a nyers vektor ensemble előrejelzések és a historikus megfigyelések közötti függőségi struktúrát. Mindkét módszer jelentősen felülmúlja a referencia, függetlenül kalibrált többváltozós előrejelzéseket, amelyeket az MLP és POLR modellek által előállított minták szekvenciális összefűzésével készítenek, feltételezve az állomások függetlenségét. Az ECMWF előrejelzéseken alapuló modellekhez képest a CAMS bevonása minden egyes modell esetében 0,58%-os és 4,11%-os energy score (ES) javulást eredményezett. A korrelációs struktúra leírása szempontjából a CAMS-szel kiterjesztett előrejelzések az ECC és SSh módszerek esetén több mint 20%-kal haladják meg a klimatikus előrejelzéseket, annak ellenére, hogy a függetlenül kalibrált referencia módszerek 5%-os előnyt élveznek. A versengő többváltozós előrejelzések közül az ECC-korrigált MLP-C és POLR-L módszerek mutatták a legjobb összteljesítményt.

Az eredményeink rámutatnak, mennyire értékesek a CAMS látástávolság előrejelzések, mint a látástávolság ensemble előrejelzéseinek diszkrét utófeldolgozására használt modellek kiegészítő magyarázó változói, és demonstrálják az egyszerű kétlépcsős többváltozós módszerek hatékonyságát ezen időjárási mennyiség térbeli összefüggéseinek megragadásában.

2.3. Többváltozós utófeldolgozó módszerek összehasonlítása globális ECMWF ensemble előrejelzéseken

Az 5. tézishez kapcsolódó esettanulmányok

A disszertációban lásd: **6. fejezet**

Publikációk: Lakatos *et al.* (2023)

A többváltozós utófeldolgozás célja az, hogy helyreállítsa az előzetesen külön-külön kalibrált előrejelzések között eredetileg fennálló, de esetlegesen elveszett függőségi kapcsolatokat. A felhasználóknak olyan koherens előrejelzésekre van szükségük, amelyek figyelembe veszik ezeket a függőségeket – például a csapadék és a hőmérséklet közötti kölcsönhatásokat, különösen fagypont közeli körülmények esetén. Ezen felhasználói igények kielégítésére különböző, kopulákon alapuló megközelítéseket fejlesztettek ki. A legfejlettebb többváltozós utófeldolgozási technikák empirikus kopulát alkalmaznak, amely egy függőségi sablont felhasználva alapulhat NWP-ensemble előrejelzéseken vagy múltbéli megfigyeléseken. Az előbbire ECC-ként, míg az utóbbira SSh-ként hivatkozunk, mely módszereket a disszertáció 2. fejezetében részletesen bemutattunk, az 5. fejezetben pedig alkalmazásukra is sor került. Ezen kívül a használt módszerek rövid magyarázata szintén megtalálható az előző szakaszban.

Munkánk során úgy találtuk, hogy a múltbéli adatokat meghatározó napok számának növelése nem eredményez jelentős különbséget, ezért az egyszerűség kedvéért a függőségi sablon kialakításához a tanuló időszak megfigyeléseit használtuk fel. Az ECC és az SSh egyszerűbb változatai a mintavételi stratégiáikban térnek el egymástól. Azok a modellek, amelyek az előrejelzési eloszlások ekvidisztáns kvantilisein alapulnak, ECC-Q/SSh-Q néven ismertek, a véletlenszerű mintavételt alkalmazók ECC-R/SSh-R, míg a rétegzett mintavételt használók ECC-S/SSh-S jelölést kapnak. Az ECC egy fejlettebb változata, a dual ensemble copula coupling módszer (dECC; Ben Bouallègue *et al.*, 2016) a nyers ensemble rangstruktúráját az előrejelzési hibák korrelációjával kombinálja. A fejlettebb SSh variánsok, mint például a hasonlóság alapú SSh és a minimális divergencián alapuló SSh, rendre olyan dátumokat használnak a függőségi sablonok kialakításához, amelyek esetén a validációs napra vonatkozó ensemble előrejelzések a lehető legnagyobb hasonlóságot mutatják az azokhoz kapcsolódó előrejelzésekkel, illetve a megfigyelések eloszlásával. A alkalmazott többváltozós utófeldolgozó eljárásokkal kapcsolatos további részletekért lásd a disszertáció 2. fejezetét, vagy Schefzik és Möller (2018) kapcsolódó munkáját.

Az empirikus kopula-alapú kétlépcsős utófeldolgozó eljárások alapos összehasonlítása érdekében első lépésben EMOS-t alkalmazunk, majd azt követően a fentebb említett ECC és SSh variánsokkal újradefiniáljuk az esetlegesen elveszett korrelációs struktúrákat. Ehhez a korábban Hemri *et al.* (2014) által is felhasznált ECMWF ensemble előrejelzéseket és megfigyeléseket vettük alapul, amelyek 2 méter magasságban mért hőmérsékletre, 10 méter magasságban mért szélsőségre és 24 órás csapadékösszegre vonatkoznak, a 2002. január és 2014. március közötti időszakban. Míg az egyváltozós kalibrációt a Hemri *et al.* (2014) által vázolt EMOS modellek szerint valósítjuk meg, a második lépésben összesen négy ECC- és öt SSh-variáns által kapott predikciók teljesítményét közvetlenül az EMOS előrejelző eloszlásokon alapuló naiv megközelítésekkel hasonlítjuk össze. Munkánk során az egyszerűség érdekében kifejezetten az időbeli függőségek helyreállítására összpontosítunk. A naiv, függetlenséget feltételező előrejelzéseket az ECC és SSh második lépésének kihagyásával határozzuk meg, azonban ugyanazokat a mintavételezési stratégiákat alkalmazzuk. Ennek megfelelően a kapcsolódó modellek EMOS-Q, EMOS-R és EMOS-S jelölést kapnak.

Tudomásunk szerint ez az első olyan munka, amely többváltozós utófeldolgozó eljárások ilyen széles körét hasonlítja össze valós adatok felhasználásával, beleértve a nemrég kifejlesztett hasonlóság alapú SSh és dECC módszereket.

Eredmények

Az eredmények azt mutatják, hogy a többváltozós utófeldolgozó módszerek javítják az előrejelzések minőségét mind a nyers ensemble előrejelzésekhez, mind pedig az egyszerű, függetlenséget feltételező egyváltozós modellekhez képest. A véletlenszerű mintavételezési stratégiát alkalmazó modellek (ECC-R és SSh-R) azonban kevésbé sikeresek a kapcsolatok helyreállításában, mint a kvantilis-alapú (ECC-Q és SSh-Q) vagy rétegzett mintavételezést (ECC-S és SSh-S) alkalmazó módszerek. Az utóbbi két megközelítés között azonban nem tapasztaltunk szignifikáns teljesítménybeli különbséget, ami arra enged következtetni, hogy mindkettő stabil és megbízható választás lehet, különösen, ha a többváltozós konzisztencia kiemelten fontos szempont.

Azt tapasztaltuk, hogy az EMOS modellek jobb teljesítményt nyújtottak a nyers ensemble előrejelzéseknél, bár ez az előny a PPT24 előrejelzések esetében volt a legkevésbé jelentős, és mindhárom mennyiség esetében csökkent az előrejelzési horizont növekedésével. A disszertáció 6.2.2. szakaszában bemutattuk, hogy a T2M előrejelzések esetében minden kalibrációs módszer javította a nyers ensemble ES mutatóját, ám a teljesítménybeli különbségek minimálisak voltak. Az ECC-Q és ECC-S mutatták a legjobb előrejelzési képességet, míg a kifinomultabb függőségi sablonra épülő SSh változatok (mdSSh, simSSh) nem haladták meg az ECC-Q teljesítményét. Vizsgálataink szerint a függőségi sablon alapját képező historikus

dátumok számának növelése elhanyagolható előnyökkel járt. A megbízhatóság (*reliability*) szempontjából az ECC-S és más kopula-alapú modellek jobban teljesítettek, mint az olyan független módszerek, mint az EMOS-Q, amely túlzott szóródást mutatott. A V10 előrejelzések esetében az utófeldolgozás jelentősen csökkentette az ES mutató értékét, és az ECC-S, valamint az SSh-S nyújtotta a legjobb teljesítményt. Az empirikus kopula-alapú módszerek következetesen felülmúlták az független EMOS modelleket, amelyek – hasonlóan a T2M referencia módszerekhez – túlzott szóródást mutattak. A megbízhatóságot értékelő mutatók (RI; reliability index) azt jelezték, hogy a kopula-alapú megközelítések, különösen az SSh-Q és SSh-S, az egyenletes eloszlást jobban közelítő ranghisztogramokat eredményeztek, ami jobb kalibrációra utal. A PPT24 előrejelzések esetében a teljesítménybeli különbségek kevésbé voltak kiemelkedők. Az ECC-S érte el a legjobb teljesítményt, ezt követte az ECC-Q, a dECC és a simSSh, majd a nyers ensemble és az egyváltozós modellek. Az RI szerint az empirikus kopula-alapú módszerek mutatták a legjobb előrejelzési képességet, de a különbségek e megközelítések között statisztikailag nem bizonyultak szignifikánsnak, ami megerősíti a csapadék-előrejelzések javításának nehéz voltát. Ezek a megállapítások alátámasztják az egyszerű és alacsony számításigényű ECC-Q hatékonyságát és megbízhatóságát, mivel versenyképes eredményeket ér el a bonyolultabb modellekhez képest. Eredményeink összhangban vannak Lerch *et al.* (2020) megállapításaival, amelyek szintén rámutattak, hogy minden szituációban eltérő utófeldolgozó eljárás lehet a legoptimálisabb választás. Az összehasonlíthatóság érdekében az előrejelzések méretét a nyers ensemble elemszámára korlátoztuk, bár az SSh módszerek tetszőleges elemszámú előrejelzés generálását is lehetővé teszik. Ez a képesség jelentős hatással van a ritka és szélsőséges események modellezésére. Lerch *et al.* (2017) például arra mutatott rá, hogy a nagyobb ensemble méretek javíthatják a szélsőséges időjárási jelenségek észlelését és előrejelzését. A jövőbeli kutatások az SSh módszerek rugalmasságának hatását vizsgálhatják különböző körülmények mellett, különösen a súlyozott többváltozós illeszkedési mutatók kontextusában (Allen *et al.*, 2023).

Irodalomjegyzék

- Allen, S., Ginsbourger, D. és Ziegel, J. (2023) Evaluating forecasts for high-impact events using transformed kernel scores. *SIAM/ASA J. Uncertain. Quantif.* **11**, 906–940.
- Baran, Á. és Baran, S. (2024) A two-step machine learning approach to statistical post-processing of weather forecasts for power generation. *Q. J. R. Meteorol. Soc.* **150**, 1029–1047.
- Baran, Á., Lerch, S., El Ayari, M. és Baran, S. (2021) Machine learning for total cloud cover prediction. *Neural. Comput. Appl.* **33**, 2605–2620.
- Baran, S. és **Lakatos, M.** (2023) Statistical post-processing of visibility ensemble forecasts. *Meteorol. Appl.* **30**, paper e2157, doi:10.1002/met.2157.
- Baran, S. és **Lakatos, M.** (2024) Clustering-based spatial interpolation of parametric post-processing models. *Wea. Forecasting* **9**, 1591–1604.
- Baran, S. és Möller, A. (2015) Joint probabilistic forecasting of wind speed és temperature using Bayesian model averaging. *Environmetrics* **26**, 120–132.
- Ben Bouallègue, Z., Heppelmann, T., Theis, S. E. és Pinson, P. (2016) Generation of scenarios from calibrated ensemble forecasts with a dual-ensemble copula-coupling approach. *Mon. Weather Rev.* **144**, 4737–4750.
- Chmielecki, R. M. és Raftery, A. E. (2011) Probabilistic visibility forecasting using Bayesian model averaging. *Mon. Weather Rev.* **139**, 1626–1636.
- Dabernig, M., Mayr, G. J., Messner, J. W. és Zeileis, A. (2017) Spatial ensemble post-processing with stésardized anomalies. *Q. J. R. Meteorol. Soc.* **143**, 909–916.
- Gneiting, T. és Raftery, A. E. (2005) Weather forecasting with ensemble methods. *Science* **310**, 248–249.

- Gneiting, T., Raftery, A. E., Westveld, A. H. és Goldman, T. (2005) Calibrated probabilistic forecasting using ensemble model output statistics és minimum CRPS estimation. *Mon. Weather Rev.* **133**, 1098–1118.
- Hemri, S., Scheuerer, M., Pappenberger, F., Bogner, K. és Haiden, T. (2014) Trends in the predictive performance of raw ensemble weather forecasts. *Geophys. Res. Lett.* **41**, 9197–9205.
- Lakatos, M.** és Baran S. (2024) Enhancing multivariate post-processed visibility predictions utilizing Copernicus Atmosphere Monitoring Service forecasts. *Meteorol. Appl.* **31**, paper e70015, doi:10.1002/met.70015
- Lakatos, M.**, Lerch, S., Hemri, S. és Baran, S. (2023) Comparison of multivariate post-processing methods using global ECMWF ensemble forecasts. *Q. J. R. Meteorol. Soc.* **149**, 856–877.
- Lerch, S. és Baran, S. (2017) Similarity-based semi-local estimation of EMOS models. *J. R. Stat. Soc. Ser. C Appl. Statist.* **66**, 29–51.
- Lerch, S., Baran, S., Möller, A., Groß, J., Schefzik, R., Hemri, S. és Graeter, M. (2020) Simulation-based comparison of multivariate ensemble post-processing methods. *Nonlinear Process. Geophys.* **27**, 349–371.
- Lerch, S., Thorarinsdottir, T.L., Ravazzolo, F. és Gneiting, T. (2020) Forecaster’s dilemma: Extreme events és forecast evaluation. *Stat. Sci.* **32**, 106–127.
- Möller, A., Lenkoski, A. és Thorarinsdottir, T. L. (2013) Multivariate probabilistic forecasting using ensemble Bayesian model averaging és copulas. *Q. J. R. Meteorol. Soc.* **139**, 982–991.
- Raftery, A. E., Gneiting, T., Balabdaoui, F. és Polakowski, M. (2005) Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Weather Rev.* **133**, 1155–1174.
- Rasp, S. és Lerch, S. (2018) Neural networks for postprocessing ensemble weather forecasts. *Mon. Weather Rev.* **164**, 3885–3900.
- Schefzik, R. (2016) Combining parametric low-dimensional ensemble postprocessing with reordering methods. *Q. J. R. Meteorol. Soc.* **142**, 2463–2477.
- Schefzik, R. és Möller, A. (2018) Ensemble postprocessing methods incorporating dependence structures. In Vannitsem, S., Wilks, D. S. és Messner, J. W. (eds.), *Statistical Postprocessing of Ensemble Forecasts*. Elsevier, Amsterdam, pp. 91–125.

- Schefzik, R., Thorarinsdottir T. L. és Gneiting, T. (2013) Uncertainty quantification in complex simulation models using ensemble copula coupling. *Statist. Sci.* **28**, 616–640.
- Scheuerer, M., Switanek, M. B., Worsnop, R. P. és Hamill, T. M. (2020) Using artificial neural networks for generating probabilistic subseasonal precipitation forecasts over California. *Mon. Weather Rev.* **148**, 3489–3506.
- Taillardat, M. és Mestre, O. (2020) From research to applications – examples of operational ensemble post-processing in France using machine learning. *Nonlin. Processes Geophys.* **27**, 329–347.
- World Meteorological Organization (1992) *International Meteorological Vocabulary* (WMO-No.182) WMO, Geneva.

Publikációs lista

Folyóiratcikkek

Lakatos, M., Lerch, S., Hemri, S. és Baran, S. (2023) Comparison of multivariate post-processing methods using global ECMWF ensemble forecasts. *Q. J. R. Meteorol. Soc.* **149**, 856–877.

<https://doi.org/10.1002/qj.4436>

(**IF: 3; SJR: D1**) [Független hivatkozás: 10]

Baran, S., és **Lakatos, M.** (2023) Statistical post-processing of visibility ensemble forecasts. *Meteorol. Appl.* **30**, paper e2157.

<https://doi.org/10.1002/met.2157>

(**IF: 2.3; SJR: Q2**) [Független hivatkozás: 2]

Baran, S., és **Lakatos, M.** (2024) Clustering-based spatial interpolation of parametric post-processing models. *Wea. Forecasting* **39**, 1591–1604.

<https://doi.org/10.1175/WAF-D-24-0016.1>

(**IF: 3; SJR: Q1**)

Lakatos, M. és Baran S. (2024) Enhancing multivariate post-processed visibility predictions utilizing Copernicus Atmosphere Monitoring Service forecasts. *Meteorol. Appl.* **31**, paper e70015.

<https://doi.org/10.1002/met.70015>

(**IF: 2.3; SJR: Q2**)

Független folyóiratcikkek

Baran, S., Marín, J. C., Cuevas, O., Díaz, M., Szabó, M., Nicolis, O. és **Lakatos, M.** (2024) Machine learning-based probabilistic forecasting of solar irradiance in Chile. *Adv. Stat. Climatol. Meteorol. Oceanogr.*, in press.

(**SJR: Q2**)

Előadások nemzetközi konferencián

Nagy-Lakatos, M. és Baran, S.: Enhancing multivariate post-processed visibility predictions utilizing CAMS forecasts, *18th International Conference of the ERCIM WG on Computational és Methodological Statistics*, London, Egyesült Királyság, 2024. december 14-16. (meghívott szekció előadás)

Nagy-Lakatos, M. és Baran, S.: Machine learning-based discrete post-processing of visibility ensemble forecasts (poszter előadás), *EGU General Assembly 2024*, Bécs, Ausztria, 2024. április 14–19., EGU24-5541,
<https://doi.org/10.5194/egusphere-egu24-5541>, 2024.

Nagy-Lakatos, M. és Baran, S.: Comparison of multivariate post-processing methods using global ECMWF ensemble forecasts, *16th International Conference of the ERCIM WG on Computational és Methodological Statistics*, Berlin, Németország, 2023. december 16-18. (meghívott szekció előadás)

Nagy-Lakatos, M. és Baran, S.: Classification-based calibration of visibility ensemble forecasts, *16th German Probability és Statistics Days*, Essen, Németország, 2023. március 7-10.

Nagy-Lakatos, M. és Baran, S.: Statistical post-processing of visibility ensemble forecasts, *15th International Conference of the ERCIM WG on Computational és Methodological Statistics*, London, Egyesült Királyság, 2022. december 17-19.

Lakatos, M. és Baran, S.: Restoration of temporal dependence in post-processed ensemble forecasts, *EGU General Assembly 2022*, Bécs, Ausztria, 2022. április 23–27., EGU22-13125,
<https://doi.org/10.5194/egusphere-egu22-13125>, 2022

Lakatos, M. és Baran, S.: Restoration of temporal dependence in statistical post-processing of ensemble weather forecasts, *6th conference of the Deutsche Arbeitsgemeinschaft Statistik*, Hamburg, Németország, 2022. március 28. - április 1.

Lakatos, M.: Comparison of multivariate ensemble post-processing methods, *25th Young Statisticians Meeting*, Voralpe, Ausztria, 2021 (meghívott előadás)

Baran, S. és Lakatos, M.: Comparison of multivariate post-processing methods (poszter előadás). *The 1st Conference on Information Technology és Data Science*, Debrecen, Magyarország, 2020.



Nyilvántartási szám: DEENK/153/2025.PL
Tárgy: PhD Publikációs Lista

Jelölt: Nagy-Lakatos Mária
Doktori Iskola: Informatikai Tudományok Doktori Iskola
MTMT azonosító: 10088516

A PhD értekezés alapjául szolgáló közlemények

Idegen nyelvű tudományos közlemények külföldi folyóiratban (4)

1. Baran, S., **Nagy-Lakatos, M.**: Clustering-based spatial interpolation of parametric post-processing models.
Weather Forecast. 39 (11), 1591-1604, 2024. ISSN: 0882-8156.
DOI: <http://dx.doi.org/10.1175/WAF-D-24-0016.1>
IF: 3 (2023)
2. **Nagy-Lakatos, M.**, Baran, S.: Enhancing multivariate post-processed visibility predictions utilizing Copernicus Atmosphere Monitoring Service forecasts.
Meteorol. Appl. 31 (6), 1-19, 2024. ISSN: 1350-4827.
DOI: <http://dx.doi.org/10.1002/met.70015>
IF: 2.3 (2023)
3. **Nagy-Lakatos, M.**, Lerch, S., Hemri, S., Baran, S.: Comparison of multivariate post-processing methods using global ECMWF ensemble forecasts.
Q. J. R. Meteorol. Soc. 149 (752), 856-877, 2023. ISSN: 0035-9009.
DOI: <http://dx.doi.org/10.1002/qj.4436>
IF: 3
4. Baran, S., **Nagy-Lakatos, M.**: Statistical post-processing of visibility ensemble forecasts.
Meteorol. Appl. 30 (5), 1-18, 2023. ISSN: 1350-4827.
DOI: <http://dx.doi.org/10.1002/met.2157>
IF: 2.3





További közlemények

Idegen nyelvű tudományos közlemények külföldi folyóiratban (1)

5. Baran, S., Marín, J. C., Cuevas, O., Díaz, M., Szabó, M., Nicolis, O., **Nagy-Lakatos, M.**: Machine learning-based probabilistic forecasting of solar irradiance in Chile.
Adv. Stat. Clim. Meteorol. Oceanogr. "Accepted by Publisher", 1-17, 2025. ISSN: 2364-3579.

A közlő folyóiratok összesített impakt faktora: 10,6

**A közlő folyóiratok összesített impakt faktora (az értekezés alapjául szolgáló közleményekre):
10,6**

A DEENK a Jelölt által a Tudóstérbe feltöltött adatok bibliográfiai és tudományometriai ellenőrzését a tudományos adatbázisok és a Journal Citation Reports Impact Factor lista alapján elvégezte.

Debrecen, 2025.04.22.



Short thesis for the degree of doctor of philosophy (PhD)

Post-processing ensemble weather forecasts for discrete and continuous quantities in univariate and multivariate settings

by Mária Nagy-Lakatos

Supervisor: Prof. Dr. Sándor Baran



UNIVERSITY OF DEBRECEN

Doctoral School of Informatics

Debrecen, 2025

Contents

Introduction	1
1 Thesis statements	3
2 Calibration of ensemble forecasts	5
2.1 Clustering-based spatial interpolation	5
2.2 Post-processing of visibility	8
2.3 Multivariate post-processing	11
Bibliography	14
List of publications and conferences	19

Introduction

Weather exhibits chaotic behavior, where minor deviations in the initial conditions of numerical weather prediction (NWP) models can quickly escalate, impacting forecast accuracy. Additionally, a single NWP output provides no information on forecast uncertainty. The introduction of ensemble prediction systems (EPS) has enabled probabilistic weather forecasting by using multiple ensemble members with varying initial conditions and parameterizations (Gneiting and Raftery, 2005). However, generating ensemble forecasts is computationally demanding, especially with larger ensemble sizes and higher forecast resolution. The European Centre for Medium-Range Weather Forecasts (ECMWF) has significantly improved its ensemble forecast resolution, reducing the horizontal resolution from 120 km in 1992 to 9 km by 2023¹.

Despite improvements, NWP models often contain systematic errors and may underestimate or overestimate observations. Statistical post-processing models, such as ensemble model output statistics (EMOS; Gneiting *et al.*, 2005) and Bayesian model averaging (BMA; Raftery *et al.*, 2005), emerged in the early 2000s to address these shortcomings by mapping ensemble predictions onto full predictive distributions. Post-processing continues to evolve with increasing use of machine learning methods (Rasp and Lerch, 2018).

However, most post-processing methods are applied independently for each prediction horizon, location, and time, often neglecting inter-variable, spatial, and temporal dependencies. Multivariate post-processing incorporates these dependencies to improve ensemble calibration. Methods such as parametric and empirical copulas (Scheffzik *et al.*, 2013; Möller *et al.*, 2013), and joint distribution modeling (Baran and Möller, 2015), have been used since the 2010s to restore multivariate dependencies in ensemble forecasts, and these methods can also be combined (Scheffzik, 2016).

This thesis aims to compare various multivariate post-processing methods, filling the gap in evaluations across multiple methods using real-world data. The focus is on two-step methodologies, where initial separate post-processing is followed by restoring multivariate dependencies with copulas. Another goal is the

¹<https://www.ecmwf.int/en/about/who-we-are/history>

statistical post-processing of visibility forecasts, which are less skillful than other weather variables, and demonstrating how deterministic visibility forecasts from CAMS can improve post-processing and preserve spatial dependencies (Baran and Lakatos, 2023; Lakatos and Baran, 2024).

Additionally, we extend the semi-local parameter estimation method proposed by Lerch and Baran (2017) with an interpolation step, testing it on global ECMWF 10-m wind speed forecasts where direct model training is impossible due to the lack of verifying observations.

The thesis is structured as follows: Chapter 1 provides an overview of ensemble prediction systems and parametric EMOS models for calibrating weather variables, introducing the proportional odds logistic regression (POLR) and multi-layer perceptron (MLP) classification algorithms for post-processing discrete quantities. Chapter 2 details the two-step multivariate post-processing methods tested in this study. Chapter 3 discusses widely used evaluation metrics and the training process for post-processing. Chapter 4 extends the semi-local parameter estimation method proposed by Lerch and Baran (2017) with interpolation, tested on ECMWF wind speed data (Baran and Lakatos, 2024). Chapter 5 addresses the statistical post-processing of visibility forecasts using the *EUPPBench* benchmark dataset, ECMWF, and Copernicus Atmosphere Monitoring Service (CAMS) forecasts. For the latter two, we also consider the restoration of multivariate spatial dependencies, investigating the impact of CAMS forecasts as additional covariates on post-processing methods. Chapter 6 compares multivariate post-processing methods using ECMWF forecast-observation pairs (Lakatos *et al.*, 2023). Finally, Chapter 7 summarizes the findings and suggests directions for future research.

To conclude, the thesis highlights the crucial role of statistical post-processing in enhancing ensemble forecasts, both from univariate and multivariate perspectives. By evaluating various calibration techniques, it demonstrates how post-processing can improve the accuracy and reliability of weather predictions across different forecast scenarios.

Chapter 1

Thesis statements

- **Thesis 1.**

The novel methodology introduced for interpolating parametric predictive distributions eliminates the computational burden of estimating Gaussian fields or using station-specific information. The method relies solely on the data required for EMOS modeling and is adaptable to parametric post-processing techniques.

- **Thesis 2.**

At observed locations, local and observation-based semi-local methods yield the best results. However, at unobserved locations, forecast-based semi-local methods demonstrate superior performance, providing sharper prediction intervals, improved mean continuous ranked probability score (CRPS), and lower mean absolute error (MAE) for median forecasts compared to regional models.

For details on Theses 1 and 2, see Chapter 4 of the dissertation.
Related paper: Baran and Lakatos (2024)

- **Thesis 3.**

Post-processed forecasts, particularly those using POLR, outperform raw ECMWF ensemble predictions in mean LogS and RMSE, with POLR also surpassing MLP in univariate discrete post-processing of visibility forecasts, especially when locally trained. Adding CAMS visibility forecasts as covariates further improves performance, enhancing mean CRPS by up to 5.25% and mean LogS by 1.37%, with Brier score gains of up to 18.59% for low-visibility events.

- **Thesis 4.**

Multivariate post-processing methods, such as two-step ensemble copula coupling (ECC) and Schaake shuffle (SSh), successfully restore spatial dependencies in visibility ensemble forecasts and outperform independently calibrated multivariate approaches. Extending these models with CAMS predictions further improves energy score by 0.58%–4.11% while significantly enhancing correlation specification.

For details on Theses 3 and 4 see Chapter 5 of the dissertation.
Related papers: Baran and Lakatos (2023); Lakatos and Baran (2024)

- **Thesis 5.**

A wide range of state-of-the-art multivariate post-processing methods improves forecast quality over raw ensemble predictions and univariate models, but the performance differences are minimal across various methods. Notably, quantile-based sampling (ECC-Q and SSh-Q) yields slightly better results than random sampling (ECC-R and SSh-R), with the ECC-Q approach emerging as a robust, efficient, and computationally cost-effective method. Despite the complexity of advanced techniques like dual ensemble copula coupling (dECC), minimum divergence Schaake shuffle (mdSSh), and similarity-based Schaake shuffle (simSSh), none show significant improvements over ECC-Q, confirming it as the preferred method for future studies.

For details on Thesis 5, see Chapter 6 of the dissertation.
Related paper: Lakatos *et al.* (2023)

Chapter 2

Statistical post-processing of ensemble forecasts

2.1 Clustering-based spatial interpolation of parametric post-processing models

The details and case study for **Theses 1** and **2**.

In the dissertation see **Chapter 4**.

Publication: Baran and Lakatos (2024)

Post-processing methods like EMOS depend on past forecast-observation pairs, limiting calibration to observed locations. Prior studies have shown that regional parametric post-processing, followed by interpolation, improves accuracy by 20% at observed and 11% at unobserved sites (Baran and Baran, 2024). However, in large ensemble domains, varying climatology and forecast errors can limit regional EMOS effectiveness. We propose a semi-local method combining clustering and interpolation, building on the work of Lerch and Baran (2017). We assess the interpolation methods detailed in Section 4.1 of the thesis, using 52-member ECMWF wind-speed ensemble forecasts for 674 SYNOP stations across Europe, Asia, and Australia from January 1, 2020 to April 30, 2023. Forecasts are initialized at 0000 UTC, and we examine 10 different lead times ranging from 1 to 10 days. The dataset includes the high resolution (HRES; 9 km grid) and control (CNTL; 18 km grid) forecasts, and the 50 ensemble members (ENS; 18 km grid) obtained using perturbed initial conditions. These members can be considered as exchangeable, as they are statistically indistinguishable. We use 75% (506 stations) for parameter estimation and interpolate models to the remaining 168 stations, with about 1.65% of the cases incomplete due to missing data.

Feature selection methods

The semi-local EMOS model presented by Lerch and Baran (2017) employs k -means clustering to group observed sites based on their climatology or/and the forecast errors of the ensemble mean relative to the observations. Hence, the resulting clusters share a common set of EMOS parameters. In the following, we summarize the observation-based feature selection method introduced by Lerch and Baran (2017) and present our forecast-only feature selection methods, which extend their work by enabling the interpolation of predictive distributions to unobserved locations.

Observation-based quantile features (OQ)

OQ features for a station are derived from equidistant quantiles of the empirical CDFs of observed wind speed and forecast errors of the ensemble mean over a rolling training period (Lerch and Baran, 2017). However, since these features rely on observations, assigning unobserved locations to existing clusters is challenging.

Forecast-based quantile features (FQ)

This feature set is the forecast-based analogue of the previous, observation-based construction. Denote by $\widehat{H}_{i,n}^{(1)}$, $\widehat{H}_{i,n}^{(2)}$, $\widehat{H}_{i,n}^{(3)}$ and $\widehat{H}_{i,n}^{(4)}$ the empirical CDFs of high-resolution forecasts f_{HRES} , control forecasts f_{CNTL} and mean $\bar{f}_{\text{ENS}}$ and standard deviation S_{ENS} of the 50 exchangeable ensemble members, respectively, at location i over the training period of length n . Now the N dimensional feature vector for station i comprises the concatenation of equidistant quantiles of $\widehat{H}_{i,n}^{(j)}$ at levels $\frac{1}{N_j+1}, \frac{2}{N_j+1}, \dots, \frac{N_j}{N_j+1}$, $j = 1, 2, 3, 4$, where $N_1 = N_2 = N_3 = \lfloor \frac{N}{4} \rfloor$ and $N_4 = N - N_1 - N_2 - N_3$.

Forecasts as features (FF)

Instead of equidistant quantiles, for each location i and training period of length n , consider simply the $4n$ -dimensional concatenation of training vectors of f_{HRES} , f_{CNTL} , $\bar{f}_{\text{ENS}}$ and S_{ENS} (the n high-resolution and n control forecasts from the training window and the n -dimensional mean and standard deviation vectors of the corresponding exchangeable members) in the order as they are listed. Naturally, this simple approach might result in rather high-dimensional features, which increases the computational cost of k -means clustering. However, the additional cost is usually less than the cost of computing the equidistant quantiles comprising a forecast-based quantile feature.

Interpolation algorithm

Since forecast-based features are available at both observed and unobserved locations, Section 4.1 of the dissertation introduces an interpolation algorithm, which is summarized in the flowchart presented as Figure 4.1, and also in this subsection.

- Step 1. Estimate semi-local EMOS models for observed locations and identify clusters (using any of the feature selection methods).
- Step 2. Assign FF or FQ features to each observed location, based on which the cluster means are calculated.
- Step 3. Calculate the same forecast-based features for unobserved locations as in Step 2.
- Step 4. Unobserved locations are assigned to the nearest cluster by calculating the Euclidean distances between the feature vectors (FF/FQ features) of the unobserved locations and the cluster means. The unobserved location is then associated with the cluster whose mean has the smallest Euclidean distance to the location's feature vector.
- Step 5. Use the assigned cluster's EMOS parameters to compute predictive distributions for unobserved locations.

Results

Observed locations are arranged into 20 clusters, and unobserved locations are assigned to clusters utilizing forecasts as features, since this approach results in more skillful models than the use of forecast-based quantile features. As a reference post-processing approach, we consider the regional truncated normal (TN) EMOS model (see Section 1.2.2 of the dissertation) based only on the observed locations and interpolated to the unobserved sites, and also the raw ECMWF ensemble forecasts. All EMOS models are trained using a 60-day rolling training window. Our results indicate that local and observation-based semi-local methods perform best at observed sites, while forecast-based semi-local approaches outperform both observation-based and local methods at unobserved locations (see Figure 4.3 of the dissertation). The above findings are further supported by Figure 4.4 of the dissertation, showing the proportions of stations where simultaneous DM tests for observed and unobserved locations indicate significant difference in the mean CRPS over the verification period at a 5% level for lead times 1, 4, 7 and 10 days. Compared to regional models, forecast-based semi-local models offer sharper prediction intervals, better mean CRPS, and lower MAE of median forecasts (Figures 4.6, 4.3 and 4.8 of the dissertation). These findings highlight that

no universal post-processing method is ideal for all scenarios. For observed sites, local or observation-based semi-local methods are preferable, while forecast-based semi-local methods are more effective for unobserved locations. The success of forecast-based methods is due to the alignment of feature vectors derived from similar quantities. We compare threshold-weighted continuous ranked probability skill score (twCRPSS) values of semi-local methods to the regional EMOS approach for 6, 8, and 9 m/s thresholds. At observed locations, rankings align with continuous ranked probability skill score (CRPSS) (Figures 4.5a,c,e of the dissertation), but at unobserved locations, tail performance (Figures 4.5b,d,f of the dissertation) diverges from overall trends (Figure 4.3d of the dissertation). While OQ forecasts lag in mean CRPS, they outperform the competing EMOS models at higher thresholds in twCRPSS. The probability integral transform (PIT) histograms of Figure 4.7 of the dissertation show post-processed forecasts improve calibration over raw ECMWF ensembles, reducing underdispersion and bias. Regional and semi-local FQ models have the smallest PIT variances (closest to the desired 0.083 of the standard uniform law), while local and observation-based (OQ) models remain more underdispersive at unobserved sites.

2.2 Statistical post-processing of visibility ensemble forecasts

The details and case study for **Theses 3 and 4**.

In the dissertation see **Chapter 5**.

Publications: Baran and Lakatos (2023); Lakatos and Baran (2024)

The World Meteorological Organization (WMO) defines visibility as "the greatest distance at which a black object of suitable dimensions (located on the ground) can be seen and recognized when observed against the horizon sky" (WMO, 1992). Accurate visibility forecasts offer economic and safety benefits, though forecasting visibility remains challenging. NWP models typically estimate it using related weather variables like humidity, precipitation, and aerosol concentration, but fog formation and microphysical processes complicate predictions. Visibility is reported in discrete values, and similar to total cloud cover forecasts, post-processing visibility forecasts can be treated as a classification problem. Methods like proportional odds logistic regression (POLR) and multilayer perceptron (MLP) have been used effectively for this task (for details on the classifiers, see Section 1.3 of the dissertation).

We focus on calibrating the *EUPPBench* dataset and ECMWF visibility en-

semble forecasts using POLR and MLP, and explore how CAMS visibility data can improve ECMWF forecast accuracy, maintain spatial dependence across stations, and enhance predictions for low-visibility events. To restore spatial dependencies between meteorological stations, we used the two-step ensemble copula coupling (ECC) and Schaake shuffle (SSh) methods, where we generate samples from the calibrated forecast distributions obtained by univariate post-processing methods. Then these samples are rearranged into the raw ensemble rank structure in the case of ECC, while in the case of SSh, we use the rank structure of historical observations.

The *EUPPBench* dataset covers the years 2017 and 2018, and consists of ECMWF high-resolution and 51-member operational ensemble forecasts (control and 50 exchangeable members) initialized at 0000 UTC, with a forecast horizon of 120 hours and a temporal resolution of 6 hours for 42 SYNOP stations in Germany and France. The second dataset contains 51-member ECMWF visibility forecasts (lacking the high resolution forecasts) for 2020-2021, covering 30 SYNOP stations in Germany, the Czech Republic, and Poland. Forecasts are initialized at 0000 UTC and analyzed for 40 lead times (6-240 hours) with 6-hour steps. CAMS forecasts, available for the same stations but up to 120 hours, are also included. In Section 5.3.2 of the dissertation, post-processed forecasts are limited to this 120-hour range. For details on calibrating ECMWF forecasts up to 240 hours, see Appendix B of the dissertation. During modeling, we consider three approaches for spatial training data selection: local (MLP-L and POLR-L), regional (MLP-R and POLR-R), and clustering-based semi-local (MLP-C and POLR-C) methods. In the case of the latter approach, the clusters are based on three-dimensional feature vectors reflecting observed frequencies for visibility intervals of 0–5000 m, 5000–30000 m, and 30000–70000 m during the training period. Before calculating the input features of the POLR and MLP models, a normalization of the forecasts is performed; namely, all forecast values are divided by 70000. Then we consider the following set of predictors: the (normalized) control member of the ensemble $\tilde{f}_{CNTL} := f_{CNTL}/70000$, the mean $\bar{f}_{ENS}$ of the 50 (normalized) exchangeable members, the variance s^2 of the 51-member (normalized) operational ensemble, in the case of the EUPPBench dataset the (normalized) high-resolution forecast $\tilde{f}_{HRES} := f_{HRES}/70000$, and the proportions p_1 , p_2 and p_3 of ensemble members predicting visibility up to 1000 m, 1000 to 2000 m and more than 30000 m, respectively. When analyzing the effect of CAMS predictions on the skill of classification methods, we also add the normalized CAMS forecast $\tilde{f}_{CAMS} := f_{CAMS}/70000$ to the input features. Furthermore, following the suggestions of e.g. Dabering *et al.* (2017), seasonal variations are addressed by adding annual base functions $\sin(2\pi d/365)$ and $\cos(2\pi d/365)$ to the input features, where d denotes the day of the year. Note that normalization is required to have all input features com-

mensurate, which makes modelling numerically more stable.

Results

Chapter 5 of the dissertation synthesizes findings from two papers that focus on the statistical post-processing of visibility forecasts using distinct datasets. Results indicate that both climatological and post-processed forecasts substantially outperform raw ensemble forecasts in terms of mean CRPS, mean LogS, and RMSE, though the performance gap narrows as the forecast horizon extends (see Appendix B of the dissertation for ECMWF results, and Section 5.3.1 for *EUPPBench* outcomes). POLR models demonstrate superior performance compared to MLP counterparts, echoing the findings of Baran *et al.* (2021) regarding total cloud cover forecasts. Notably, the locally trained POLR model excels among post-processing methods, while semi-local training often outperforms local training in machine learning approaches (Baran and Lakatos, 2023). The second study continues from Lakatos and Baran (2024), investigating location-specific discrete post-processing of visibility. We address three main questions: the capability of advanced multivariate post-processing approaches to restore spatial dependencies lost during separate calibration of visibility ensemble forecasts; the impact of incorporating CAMS visibility predictions as additional covariates on forecast skill; and the predictive performance of post-processed forecasts in aeronautical meteorology. For the univariate approaches that perform the best (MLP-C and POLR-L), integrating the CAMS predictions leads to significant improvements of 1% to 5.25% in mean CRPS and 0.27% to 1.37% in mean LogS, with diminishing gains as lead time increases. Low-visibility events, which are critical for aeronautical applications, were also included in our analysis. Using the CAMS-extended POLR-L, we found that the average improvements in Brier score for conditions when visibility is below 200 m, 500 m, 1000 m, and 3000 m ranged from 14.22% to 18.59%.

We performed joint multivariate post-processing for all 30 SYNOP observation stations using two-step ECC and SSh methods, leveraging the dependence structure of the raw vector ensemble forecasts and historical observations. Both methods significantly outperform reference independently calibrated multivariate forecasts, which are generated by sequentially combining samples produced by the MLP and POLR models, thus assuming the independence of the stations. When compared to forecasts based solely on ECMWF predictions, incorporating CAMS yields an additional improvement of 0.58% to 4.11% in energy score (ES) across all models. In terms of correlation specification, the CAMS extended forecasts surpass climatology by over 20% for ECC and SSh methods, despite a 5% advantage for independent reference methods. Among the competing multivariate predictions, the ECC-corrected MLP-C and POLR-L methods exhibit the best overall performance.

Our findings highlight the value of CAMS visibility predictions as additional covariates in discrete post-processing of visibility ensemble forecasts, demonstrating the effectiveness of simple two-step multivariate methods in capturing spatial dependencies for this weather variable.

2.3 Comparison of multivariate post-processing methods using global ECMWF ensemble forecasts

The details and case study for **Thesis 5**.

In the dissertation see **Chapter 6**.

Publication: Lakatos *et al.* (2023)

The aim of multivariate post-processing is to restore correlation structures that may have been lost during univariate post-processing. End-users need coherent forecasts that account for dependencies, such as the interactions between precipitation and temperature, particularly near freezing conditions. To address this need, various copula-based approaches have been developed. Most advanced multivariate post-processing applications utilize an empirical copula derived from a dependence template based on either NWP ensembles or historical observations. The former is known as ECC, while the latter is referred to as SSh, both of which were discussed in Chapter 2 of the dissertation and also utilized in Chapter 5. A brief explanation of the methodologies applied can also be found in the previous section. Our investigations found no significant difference by increasing the number of pool dates, so for simplicity, we used the observations from the training period to form the dependence template. The simpler variants of ECC and SSh differ in their sampling strategies. Models drawing equidistant quantiles of predictive distributions are referred to as ECC-Q/SSh-Q, those employing random sampling as ECC-R/SSh-R, and those using stratified sampling as ECC-S/SSh-S. The dual ensemble copula coupling (dECC) method of Ben Bouallègue *et al.* (2016), a more advanced version of ECC, integrates the raw ensemble forecast structure with the estimated forecast error autocorrelation. Meanwhile, advanced SSh variants, such as the similarity-based Schaake shuffle and minimum divergence Schaake shuffle, select dates where ensemble forecasts align with the current prediction based on a similarity criterion or where the marginal distributions of the observations closely match the forecast marginal distributions, respectively. For more details on the applied multivariate post-processing methods, we refer to Chapter 2 of the dissertation, or see Schefzik and Möller (2018). To provide a thorough comparison of

the empirical copula-based, two-step multivariate post-processing techniques, we apply univariate EMOS in conjunction with a range of ECC and SSh variants to NWP ensemble forecasts of temperature, wind speed, and precipitation provided by the ECMWF, verifying against observations from SYNOP stations. We utilized global ECMWF ensemble forecasts of 2-meter temperature (T2M), 10-meter wind speed (V10), and 24-hour precipitation accumulation (PPT24) from January 2002 to March 2014, based on datasets from Hemri *et al.* (2014). These include ECMWF HRES forecasts, a 50-member ensemble, and the CNTL run with 1–10 day forecast horizons, alongside corresponding observations. After quality control, 4,160, 4,388, and 2,917 SYNOP stations were retained for T2M, V10, and PPT24, respectively. For dataset and quality control details see Hemri *et al.* (2014). While, we implement univariate calibration following the EMOS methods outlined by Hemri *et al.* (2014), for the multivariate step, we compare three naive approaches derived directly from univariate EMOS with four variants of ECC and five variants of SSh, focusing specifically on temporal dependence for clarity and simplicity. As reference forecasts samples are drawn omitting the second step of ECC, using the same sampling strategies. Hence, the constructed samples are denoted as: EMOS-Q, EMOS-S and EMOS-R.

To the best of our knowledge, this work is the first to compare a broad array of state-of-the-art two-step methods for multivariate post-processing, including recently proposed similarity-based Schaake shuffle approach and dECC.

Results

The findings reveal that multivariate post-processing methods consistently improve forecast quality compared to both the raw ensemble predictions and straightforward applications of univariate models that overlook multivariate dependencies. Among the basic methods evaluated, the ECC and SSh variants exhibited only slight differences in predictive performance across their respective sampling strategies. Specifically, random sampling (ECC-R and SSh-R) tended to perform less effectively than quantile-based sampling (ECC-Q and SSh-Q) or stratified sampling (ECC-S and SSh-S). Between the quantile-based and stratified sampling approaches, no significant distinctions in performance were observed, suggesting that both are robust options when multivariate consistency is critical.

We found that the EMOS models were able to surpass the skill of the raw ensemble forecasts, although this advantage was least noticeable in the case of the PPT24 forecasts and naturally decreased as the forecast horizon increased. In Section 6.2.2 of the dissertation, we illustrated, that for T2M forecasts, all calibration methods improved the raw ensemble’s ES, but differences among them were minimal. ECC-Q and ECC-S performed best, while advanced SSh variants (mdSSh, simSSh) did not surpass ECC-Q. Using longer historical datasets for de-

pendence templates provided negligible benefits. In reliability, ECC-S and other copula-based models outperformed independent methods like EMOS-Q, which suffered from overdispersion. For V10 forecasts, post-processing significantly reduced ES, with ECC-S and SSh-S emerging as top performers. Empirical copula-based methods consistently outperformed independent EMOS models, which similarly to the case of T2M reference methods showed overdispersion. Reliability metrics highlighted that SSh approaches, particularly SSh-Q and SSh-S, achieved more uniform rank histograms, reflecting better calibration. For PPT24 forecasts, performance differences were less pronounced. ECC-S had the best skill scores, followed by ECC-Q, dECC, and simSSh, while the raw ensemble and independent EMOS lagged behind. Reliability indices favored empirical copula-based methods, but differences between these approaches were not statistically significant, underscoring the difficulty in improving precipitation forecasts. This finding underscores the efficiency and reliability of ECC-Q as a benchmark, as it achieves competitive results while maintaining simplicity and low computational cost. This is consistent with the findings of Lerch *et al.* (2020), which similarly noted that no single multivariate post-processing approach was consistently superior across all scenarios. To ensure a fair comparison among methods, we restricted all approaches to a fixed sample size based on the raw ensemble. However, it is worth highlighting the flexibility offered by the SSh variants, which can generate post-processed ensemble forecasts of arbitrary sizes. This capability has significant implications for modelling rare and extreme events. For example, Lerch *et al.* (2017) pointed out that larger ensemble sizes could improve the detection and prediction of extreme weather phenomena. Future research could explore how the flexibility of SSh variants impacts forecast performance under different conditions, particularly in the context of weighted multivariate proper scoring rules, such as those recently proposed by Allen *et al.* (2023).

Bibliography

- Allen, S., Ginsbourger, D. and Ziegel, J. (2023) Evaluating forecasts for high-impact events using transformed kernel scores. *SIAM/ASA J. Uncertain. Quantif.* **11**, 906–940.
- Baran, Á. and Baran, S. (2024) A two-step machine learning approach to statistical post-processing of weather forecasts for power generation. *Q. J. R. Meteorol. Soc.* **150**, 1029–1047.
- Baran, Á., Lerch, S., El Ayari, M. and Baran, S. (2021) Machine learning for total cloud cover prediction. *Neural. Comput. Appl.* **33**, 2605–2620.
- Baran, S. and **Lakatos, M.** (2023) Statistical post-processing of visibility ensemble forecasts. *Meteorol. Appl.* **30**, paper e2157, doi:10.1002/met.2157.
- Baran, S. and **Lakatos, M.** (2024) Clustering-based spatial interpolation of parametric post-processing models. *Wea. Forecasting* **9**, 1591–1604.
- Baran, S. and Möller, A. (2015) Joint probabilistic forecasting of wind speed and temperature using Bayesian model averaging. *Environmetrics* **26**, 120–132.
- Ben Bouallègue, Z., Heppelmann, T., Theis, S. E. and Pinson, P. (2016) Generation of scenarios from calibrated ensemble forecasts with a dual-ensemble copula-coupling approach. *Mon. Weather Rev.* **144**, 4737–4750.
- Chmielecki, R. M. and Raftery, A. E. (2011) Probabilistic visibility forecasting using Bayesian model averaging. *Mon. Weather Rev.* **139**, 1626–1636.
- Dabernig, M., Mayr, G. J., Messner, J. W. and Zeileis, A. (2017) Spatial ensemble post-processing with standardized anomalies. *Q. J. R. Meteorol. Soc.* **143**, 909–916.
- Gneiting, T. and Raftery, A. E. (2005) Weather forecasting with ensemble methods. *Science* **310**, 248–249.

- Gneiting, T., Raftery, A. E., Westveld, A. H. and Goldman, T. (2005) Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Weather Rev.* **133**, 1098–1118.
- Hemri, S., Scheuerer, M., Pappenberger, F., Bogner, K. and Haiden, T. (2014) Trends in the predictive performance of raw ensemble weather forecasts. *Geophys. Res. Lett.* **41**, 9197–9205.
- Lakatos, M.** and Baran S. (2024) Enhancing multivariate post-processed visibility predictions utilizing Copernicus Atmosphere Monitoring Service forecasts. *Meteorol. Appl.* **31**, paper e70015, doi:10.1002/met.70015
- Lakatos, M.**, Lerch, S., Hemri, S. and Baran, S. (2023) Comparison of multivariate post-processing methods using global ECMWF ensemble forecasts. *Q. J. R. Meteorol. Soc.* **149**, 856–877.
- Lerch, S. and Baran, S. (2017) Similarity-based semi-local estimation of EMOS models. *J. R. Stat. Soc. Ser. C Appl. Statist.* **66**, 29–51.
- Lerch, S., Baran, S., Möller, A., Groß, J., Schefzik, R., Hemri, S. and Graeter, M. (2020) Simulation-based comparison of multivariate ensemble post-processing methods. *Nonlinear Process. Geophys.* **27**, 349–371.
- Lerch, S., Thorarinsdottir, T.L., Ravazzolo, F. and Gneiting, T. (2020) Forecaster’s dilemma: Extreme events and forecast evaluation. *Stat. Sci.* **32**, 106–127.
- Möller, A., Lenkoski, A. and Thorarinsdottir, T. L. (2013) Multivariate probabilistic forecasting using ensemble Bayesian model averaging and copulas. *Q. J. R. Meteorol. Soc.* **139**, 982–991.
- Raftery, A. E., Gneiting, T., Balabdaoui, F. and Polakowski, M. (2005) Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Weather Rev.* **133**, 1155–1174.
- Rasp, S. and Lerch, S. (2018) Neural networks for postprocessing ensemble weather forecasts. *Mon. Weather Rev.* **164**, 3885–3900.
- Schefzik, R. (2016) Combining parametric low-dimensional ensemble postprocessing with reordering methods. *Q. J. R. Meteorol. Soc.* **142**, 2463–2477.
- Schefzik, R. and Möller, A. (2018) Ensemble postprocessing methods incorporating dependence structures. In Vannitsem, S., Wilks, D. S. and Messner, J. W. (eds.), *Statistical Postprocessing of Ensemble Forecasts*. Elsevier, Amsterdam, pp. 91–125.

- Schefzik, R., Thorarinsdottir T. L. and Gneiting, T. (2013) Uncertainty quantification in complex simulation models using ensemble copula coupling. *Statist. Sci.* **28**, 616–640.
- Scheuerer, M., Switanek, M. B., Worsnop, R. P. and Hamill, T. M. (2020) Using artificial neural networks for generating probabilistic subseasonal precipitation forecasts over California. *Mon. Weather Rev.* **148**, 3489–3506.
- Taillardat, M. and Mestre, O. (2020) From research to applications – examples of operational ensemble post-processing in France using machine learning. *Nonlin. Processes Geophys.* **27**, 329–347.
- World Meteorological Organization (1992) *International Meteorological Vocabulary* (WMO-No.182) WMO, Geneva.

List of publications

Journal papers

Lakatos, M., Lerch, S., Hemri, S. and Baran, S. (2023) Comparison of multivariate post-processing methods using global ECMWF ensemble forecasts. *Q. J. R. Meteorol. Soc.* **149**, 856–877.

<https://doi.org/10.1002/qj.4436>

(**IF: 3; SJR: D1**) [Ind. cit.: 10]

Baran, S., and **Lakatos, M.** (2023) Statistical post-processing of visibility ensemble forecasts. *Meteorol. Appl.* **30**, paper e2157.

<https://doi.org/10.1002/met.2157>

(**IF: 2.3; SJR: Q2**) [Ind. cit.: 2]

Baran, S., and **Lakatos, M.** (2024) Clustering-based spatial interpolation of parametric post-processing models. *Wea. Forecasting* **39**, 1591–1604.

<https://doi.org/10.1175/WAF-D-24-0016.1>

(**IF: 3; SJR: Q1**)

Lakatos, M. and Baran S. (2024) Enhancing multivariate post-processed visibility predictions utilizing Copernicus Atmosphere Monitoring Service forecasts. *Meteorol. Appl.* **31**, paper e70015.

<https://doi.org/10.1002/met.70015>

(**IF: 2.3; SJR: Q2**)

Unrelated papers

Baran, S., Marín, J. C., Cuevas, O., Díaz, M., Szabó, M., Nicolis, O. and **Lakatos, M.** (2025) Machine learning-based probabilistic forecasting of solar irradiance in Chile. *Adv. Stat. Climatol. Meteorol. Oceanogr.*, in press.

(**SJR: Q2**)

International conferences

Nagy-Lakatos, M. and Baran, S.: Enhancing multivariate post-processed visibility predictions utilizing CAMS forecasts, *18th International Conference of the ERCIM WG on Computational and Methodological Statistics*, London, UK, 14-16 December 2024 (invited session talk)

Nagy-Lakatos, M. and Baran, S.: Machine learning-based discrete post-processing of visibility ensemble forecasts (poster presentation), *EGU General Assembly 2024*, Vienna, Austria, 14–19 Apr 2024, EGU24-5541, <https://doi.org/10.5194/egusphere-egu24-5541>, 2024.

Nagy-Lakatos, M. and Baran, S.: Comparison of multivariate post-processing methods using global ECMWF ensemble forecasts, *16th International Conference of the ERCIM WG on Computational and Methodological Statistics*, Berlin, Germany, 16-18 December 2023 (invited session talk)

Nagy-Lakatos, M. and Baran, S.: Classification-based calibration of visibility ensemble forecasts, *16th German Probability and Statistics Days*, Essen, Germany, 7-10 March 2023

Nagy-Lakatos, M. and Baran, S.: Statistical post-processing of visibility ensemble forecasts, *15th International Conference of the ERCIM WG on Computational and Methodological Statistics*, London, UK, 17-19 December 2022

Lakatos, M. and Baran, S.: Restoration of temporal dependence in post-processed ensemble forecasts, *EGU General Assembly 2022*, Vienna, Austria, 23–27 May 2022, EGU22-13125, <https://doi.org/10.5194/egusphere-egu22-13125>, 2022

Lakatos, M. and Baran, S.: Restoration of temporal dependence in statistical post-processing of ensemble weather forecasts, *6th conference of the Deutsche Arbeitsgemeinschaft Statistik*, Hamburg, Germany, March 28 - April 1, 2022

Lakatos, M.: Comparison of multivariate ensemble post-processing methods, *25th Young Statisticians Meeting*, Vorau, Austria. 2021 (invited talk)

Baran, S. and Lakatos, M.: Comparison of multivariate post-processing methods (poster presentation). *The 1st Conference on Information Technology and Data Science*, Debrecen, Hungary, 2020



Registry number:
Subject:

DEENK/153/2025.PL
PhD Publication List

Candidate: Mária Nagy-Lakatos

Doctoral School: Doctoral School of Informatics

MTMT ID: 10088516

List of publications related to the dissertation

Foreign language scientific articles in international journals (4)

1. Baran, S., **Nagy-Lakatos, M.**: Clustering-based spatial interpolation of parametric post-processing models.
Weather Forecast. 39 (11), 1591-1604, 2024. ISSN: 0882-8156.
DOI: <http://dx.doi.org/10.1175/WAF-D-24-0016.1>
IF: 3 (2023)
2. **Nagy-Lakatos, M.**, Baran, S.: Enhancing multivariate post-processed visibility predictions utilizing Copernicus Atmosphere Monitoring Service forecasts.
Meteorol. Appl. 31 (6), 1-19, 2024. ISSN: 1350-4827.
DOI: <http://dx.doi.org/10.1002/met.70015>
IF: 2.3 (2023)
3. **Nagy-Lakatos, M.**, Lerch, S., Hemri, S., Baran, S.: Comparison of multivariate post-processing methods using global ECMWF ensemble forecasts.
Q. J. R. Meteorol. Soc. 149 (752), 856-877, 2023. ISSN: 0035-9009.
DOI: <http://dx.doi.org/10.1002/qj.4436>
IF: 3
4. Baran, S., **Nagy-Lakatos, M.**: Statistical post-processing of visibility ensemble forecasts.
Meteorol. Appl. 30 (5), 1-18, 2023. ISSN: 1350-4827.
DOI: <http://dx.doi.org/10.1002/met.2157>
IF: 2.3





List of other publications

Foreign language scientific articles in international journals (1)

5. Baran, S., Marín, J. C., Cuevas, O., Díaz, M., Szabó, M., Nicolis, O., **Nagy-Lakatos, M.**: Machine learning-based probabilistic forecasting of solar irradiance in Chile.
Adv. Stat. Clim. Meteorol. Oceanogr. "Accepted by Publisher", 1-17, 2025. ISSN: 2364-3579.

Total IF of journals (all publications): 10,6

Total IF of journals (publications related to the dissertation): 10,6

The Candidate's publication data submitted to the Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

22 April, 2025

