

Duális felbontású ensemble előrejelzések statisztikai utófeldolgozása

Szabó Marianna

Debreceni Egyetem Informatikai Kar, 4028 Debrecen, Kassai út 26.

Lektorálta: Dr. Baran Sándor, egyetemi docens

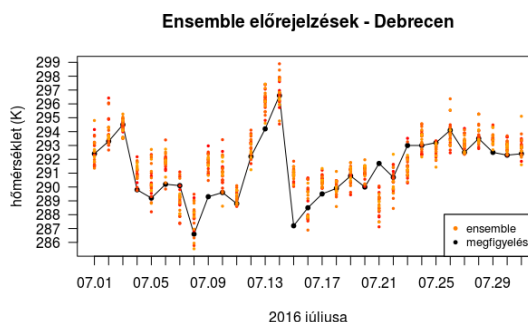
Bevezetés

Az időjárási jelenségek pontosabb előrejelzése rohamosan fejlődő világunkban egyre nagyobb szerepet kap, s ebből kifolyólag az előrejelzést szolgáltatató számos nagy szervezet folyamatosan vizsgálatokat végez azért, hogy az általuk elkészített predikciók minél közelebb legyenek a valós mérésekhez. Ennek érdekében az elmúlt években az összes nagyobb meteorológiai előrejelzéssel foglalkozó központ a különböző időjárási mennyiségek jövőbeli értékeinek prediktálásához ensemble előrejelzéseket is biztosít, amelyek különböző kezdeti feltételekkel felparaméterezett numerikus időjárási előrejelző modellekből származnak. Ez lehetővé teszi a valószínűségi időjárás előrejelzést, amit az egyesült királyságbeli European Centre for Medium-Range Weather Forecasts (ECMWF) 1992 óta alkalmaz. A 2025-ig terjedő stratégiai tervei szerint az ECMWF 18 km-ről 5 km-re növeli az ensemble előrejelzések felbontását, ami azonban jelentősen megnöveli a számítási erőforrás igényt. Ezzel a tervvel kapcsolatban az ECMWF-nél kísérleteket végeznek olyan magas és alacsony felbontású ensemble előrejelzésekből álló duális felbontású keverékek előállításával, amelyek az eddig használatban lévő modellel azonos számítási költséggel rendelkeznek, annak érdekében, hogy meghatározzák, létezik-e olyan optimális kombináció, amely a legjobb prediktív teljesítményt nyújtja. Jelen kutatás arra irányul, hogy a nyers ensemble előrejelzésekre kapott eredmény, miszerint létezik optimális keverék, akkor is igaz maradjon, ha az előrejelzéseket valamilyen statisztikai utófeldolgozó eljárással kalibráljuk.

Ensemble előrejelzés

Az időjárás előrejelzés alapvető feladata, hogy a légköri folyamatokat leíró fizikai és kémiai modellekre támaszkodva adjon egy predikciós értéket valamilyen időjárási mennyiségre, lehet ez hőmérséklet, csapadék mennyiség, szélereősség vagy légköri nyomás. Ez a determinisztikus eljárással előállított érték azonban sokfajta hibával terhelt lehet, hiszen előfordulhat, hogy a modellekhez rendelkezésre álló adatok mennyisége nem elegendő, amiből adódik, hogy a légkör aktuális kiinduló állapota sosem adható meg pontosan. Ezen bizonytalanság miatt alakult ki az úgynevezett ensemble előrejelzési módszer [3] [4]. Ennek lényege, hogy nem egy, hanem egyszerre több predikciós értéket készítenek el úgy, hogy kiválasztják a legnagyobb valószínűséggel pontosnak mondható kiinduló állapotot, majd ezt előre meghatározott hibahatáron belül módosítják, perturbálják, s az így kapott kiinduló pontokat használják fel a

megoldandó parciális differenciálegyenlet rendszerek kezdeti értékeként. Ezáltal a korábbi egyetlen előrejelzett érték helyett, az előrejelzésegysétes alapján, az időjárási változók valószínűségi eloszlása is megbecsülhető, ami lehetővé teszi a különféle időjárási események valószínűségeinek becslését is. Példaként az 1. ábrán látható az ECMWF egy ilyen ensemble felszíni hőmérséklet előrejelzése 2016 júliusára Debrecenre. A fekete vonallal jelölt megfigyeléseket jól megközelítik a predikciók, ám akadt olyan időpont is amikor jelentősebb az eltérés. Ez a különbség is mutatja, hogy a rendszer jól működik, ám kalibrálásra szorul. A kalibrálást megvalósító egyik gyakorta használt megoldás a statisztikai utófeldolgozás.



1. ábra Az ECMWF ensemble hőmérséklet előrejelzései Debrecenre.

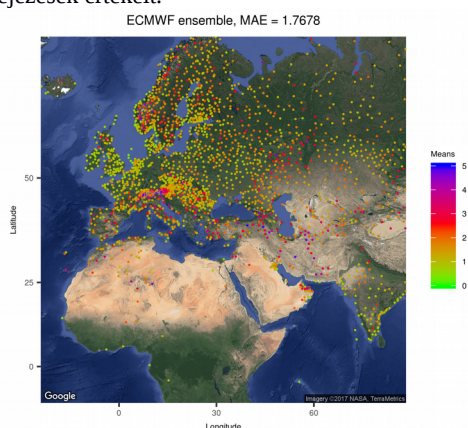
Duális rezolúció

Az ECMWF által 2025-re kitűzött egyik stratégiai cél [1], hogy a jelenleg 18 km felbontású operatív ensemble előrejelzéseinek felbontását 5 km-esre növelje, ennek azonban jelentős hátránya, hogy lényegesen nagyobb számítási kapacitást igényel, mint a kevésbé pontos predikciókra képes alacsonyabb felbontású rendszer. Ezen probléma orvoslása kapcsán kezdett el vizsgáldni Martin Leutbecher és Zied Ben Bouallègue, az ECMWF két kutatója, hogy van-e annak létjogosultsága, hogy az ECMWF-nél most operatíván használt 50 tagú 18 km felbontású (általunk a továbbiakban magas felbontásúnak nevezett) ensemble előrejelzés helyett a 200 tagot számláló alacsonyabb (29 km) felbontású ensemble előrejelzést felhasználva ugyanolyan összköltséggel rendelkező magas és alacsony felbontású előrejelzések keverékét használják. Ez a költségbeli megkötés jól kijelöli a megvizsgálandó arányokat, a kiemelt alacsony-magas párok a (0,50), (40,40), (120,20), (160,10) és a (200,0) arányokkal rendelkeznek. Vizsgálataikhoz 2016

nyarának globális hőmérséklet előrejelzéseit és a kapcsolódó validáló megfigyeléseket használták, így a mi eredményeink is ezekre az adatokra vonatkoznak.

Statisztikai utófeldolgozás

Az ensemble előrejelzések esetén általános probléma, hogy a szórásuk kicsi - mivel a perturbált értékek csak egy előre meghatározott hibatarományba eshetnek –, vagy esetleg torzítottak, ami azt eredményezi, hogy nem megfelelően kalibráltak. Az ilyen nyers adatokat terhelő hibák statisztikai alapú kiszűrésére olyan - a matematikai statisztikában ismert módszereken alapuló - utófeldolgozó eljárásokat fejlesztettek ki, amelyek a nyers ensemble előrejelzés értékeire, illetve az aktuális ténylegesen mért megfigyelésekre támaszkodva javítják az előrejelzést adott területre vonatkozóan. Tulajdonképpen az adott időjárási változéhoz megfelelő modell kiválasztásával a múltbeli adatokhoz illesztve azt, képesek vagyunk javítani a predikciót a modell segítségével módosítva a jövőre vonatkozó már előállított előrejelzések értékeit.

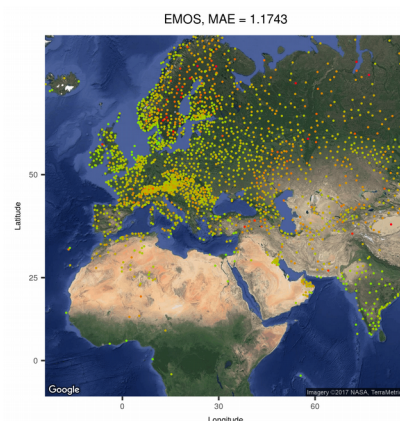


2. ábra. Az ECMWF 2016 nyarára kiadott nyers hőmérséklet előrejelzéseinek átlagos hibája

Egy ilyen statisztikai utófeldolgozást megvalósító módszer az *ensemble model output statistics* (EMOS) [2], amelynek lényege, hogy az ensemble előrejelzésekre egy valószínűségi eloszlást illesztünk, melynek paraméterei az ensemble függvényeként állnak elő.

A 2. és 3. ábrák térképein például jól összehasonlítható, hogy az európai állomásokra nézve a 2016 nyarához tartozó ensemble előrejelzések átlagos hibája nagyobb mint a EMOS technikával utófeldolgozott, s így kalibrált rendszer predikciói. Az utófeldolgozás segítségével előállított különböző arányú keverékekhez tartozó modellek predikcióinak többféle illeszkedési mutatóra kiszámított értékeit összehasonlítva egymással és a nyers adatok azonos mutatókra vett értékeivel kiderült, hogy a korábbi várakozásnak megfelelően létezik egy optimálisnak mondható (40,40) arányú duális rezolúciójú kombináció. Vizsgálataink azt is kimutatták, hogy az

utófeldolgozás jelentősen csökkenti az egyes kombinációk előrejelző képessége közötti eltéréseket.



3. ábra. Az ECMWF 2016 nyarára kiadott EMOS utófeldolgozott előrejelzéseinek átlagos hibája

Felhasznált irodalom

- [1] ECMWF strategy 2016-2025. Elérhető: https://www.ecmwf.int/sites/default/files/ECMWF_Strategy_2016-2025.pdf [Letöltve 2018. március 31.]
- [2] 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. Wea. Rev.* **133**, 1098–1118.
- [3] Leith, C. E. (1974) Theoretical skill of Monte-Carlo forecasts. *Mon. Wea. Rev.* **102**, 409–418.
- [4] Palmer, T. N. (2002). The economic value of ensemble forecasts as a tool for risk assessment: From days to decades. *Q. J. R. Meteorol. Soc.* **128**, 747–774.

Összefoglalás

Az ECMWF által biztosított 2016 nyarat felölelő 1-5 napos ensemble előrejelzésekre EMOS statisztikai utófeldolgozással végzett vizsgálatok alátámasztják a magas és alacsony felbontású előrejelzések optimális kombinációjának létezését, emellett a statisztikai kalibrálás csökkenti az illeszkedési mutatók közötti különbségeket.

Köszönetnyilvánítás

Köszönettel tartozom témavezető tanáromnak, Baran Sándornak, a kitartó támogatásáért, valamint az ECMWF két kutatójának, Martin Leutbechernek és Zied Ben Bouallègue-nek, hogy a hőmérséklet adatokat a rendelkezésünkre bocsátották, valamint útmutató tanácsaikkal segítették a munkánkat. Mindemellett a kutatást az „Integrált kutatói utánpótlás-képzési program az informatika és számítástudomány diszciplináris területein” (EFOP-3.6.3-VEKOP-16-2017-00002) című projekt támogatta. A projekt az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.