

Eötvös Loránd Tudományegyetem
Földrajz- és Földtudományi Intézet
Meteorológiai Tanszék

Az ECMWF determinisztikus és valószínűségi előrejelzések összehasonlító verifikációja



Készítette:

**Tajti Dávid
Földtudomány BSc
Meteorológia szakirány**

Témavezető:

**Ihász István
Országos Meteorológiai Szolgálat**

**Tanszéki konzulens:
Bartholy Judit
ELTE Meteorológiai Tanszék**

Budapest, 2009

Tartalomjegyzék

1. A dolgozat céljáról, témaköréről.....	3
2. Verifikáció.....	4
2.1. Determinisztikus előrejelzések verifikációja	5
2.1.1. Folytonos eloszlású változók	5
2.1.2. Diszkrét eloszlású változók.....	7
2.2. A valószínűségi előrejelzések verifikációja	11
3. A verifikáció alkalmazása	16
3.1. Verifikációs eredmények ábrázolása.....	16
3.2. Az ECMWF előrejelzések verifikációja	20
4. Hőmérséklet előrejelzések verifikációja	21
4.1. A determinisztikus előrejelzés verifikációja	21
4.2. A valószínűségi előrejelzés verifikációja.....	23
5. Összefoglalás.....	26
Irodalomjegyzék.....	27
Köszönetnyilvánítás	28
Függelék	29

1. A dolgozat céljáról, témaköréről

Az időjárási előrejelzések mindennapjaink részévé váltak. Az előrejelzések numerikus időjárás előrejelző modellek alapján készülnek (Horányi, 2003). A szakdolgozatom témája a numerikus előrejelzések objektív verifikációja. A verifikáció során a különböző meteorológiai modellek bevalását vizsgáljuk. A verifikáció jelentős segítséget nyújthat mind a modell produktumok alkalmazói s felhasználói, mind a modellfejlesztők számára.

Az európai régióban az Európai Középtávú Előrejelző Központ (közismert angol rövidítése ECMWF = European Centre for Medium Range Weather Forecasts) az egyik legnagyobb előrejelző központ, amely operatíván állít elő középtávú globális előrejelzéseket. Az ECMWF 18 ország összefogásaként 1973-ban readingi (Nagy-Britannia) székhellyel alakult meg, s jelenleg 31 tagja van. Magyarország 1994-ben csatlakozott ún. együttműködő tagországgként (Kaba, 1995, Woods, 2005).

Az ECMWF determinisztikus modellje, amellyel az előrejelzés készül, a hagyományos felszíni és magas légköri megfigyelések mellett felhasználja a tengeri úszóbóják, repülőgépes megfigyelések és a műholdas megfigyelések adatait is. A modell a kezdeti feltételeit a négydimenziós variációs analízis (4DVAR) eljárással állítja elő, ami napjaink legmodernebb adatasszimilációs technikája. A modell naponta 00 és 12 UTC-s kezdeti meteorológiai mezőkből kiindulva 10 napos determinisztikus és 15 napos 51 tagú ensemble előrejelzést készít az egész Földre vonatkozóan (Jhász, 2008). Az ensemble modell felépítése hasonló a determinisztikus modelléhez, azonban a számítás igény csökkentése végett térbeli felbontása gyengébb. A modell számos fizikai kölcsönhatást vesz figyelembe, így például az óceán-légkör, a talajnedvesség és légkör, a hótakaró és légkör közötti kölcsönhatást. 1992 óta operatíván készülnek valószínűségi vagy más néven ensemble előrejelzések is.

Munkám során a determinisztikus és a valószínűségi előrejelzések verifikációjának ismertetését tűztem ki célul. Először az objektív verifikáció módszereit ismertetem, külön a determinisztikus és a valószínűségi előrejelzésekre. Ezután néhány példa kiemelésével bemutatom az Országos Meteorológiai Szolgálatnál használt OVISYS nevű program-rendszert. Végül UNIX operációs rendszer alatt FORTRAN program nyelven saját fejlesztésű programcsalád használatával a 2m-es hőmérséklet előrejelzésre 2007-es évre végzett néhány verifikációs eredményt mutatok be.

2. Verifikáció

A modell-előrejelzések bevalásának kiértékelése (verifikációja) során az előrejelzések és megfigyelések értékeiből származtatott valamilyen mérőszámot vizsgálunk. (Csima, 2003).

A verifikáció elvégzése több szempontból is fontos. Ezek a mérőszámok a modell fejlesztő szakemberek számára hasznos információt nyújtanak a modell bevalásáról, hogy mely meteorológiai elem előrejelzésén kell javítani, mely paraméternél fordulnak elő szisztematikus hibák. Emellett a fejlesztők a verifikációs mérőszámok figyelésével könnyen és gyorsan követhetik a modell hibáit és azok javulását. Szintén fontosak az előrejelző szakemberek számára ezek a mérőszámok, mivel segítségével megállapíthatják, hogy a modell előrejelzésének bevalása milyen mértékben tér el az általuk készített előrejelzés bevalásához képest.

Megkülönböztetünk objektív és szubjektív verifikációt. Az objektív verifikáció során a modell és a megfigyelés értékeiből hosszabb idősor adatainak felhasználásával statisztikai vizsgálatokat végzünk. A szubjektív verifikáció esetében az egyes meteorológiai paramétereket térképen ábrázoljuk és vizuálisan hasonlítjuk össze. A potenciálisan rendelkezésünkre álló széles objektív verifikációs eszköztárat Nurmi: *Recommendations on the verification of local weather forecast* (2003) című munkája nyomán mutatom be. A verifikációs eszközök bemutatása során Nurmi cikke mellett *Jolliefe és munkatársai* (2003), *Richardson és munkatársai* (2007) munkájára, valamint *Szépszó Gabriella* (2004) kéziratára is támaszkodtam.

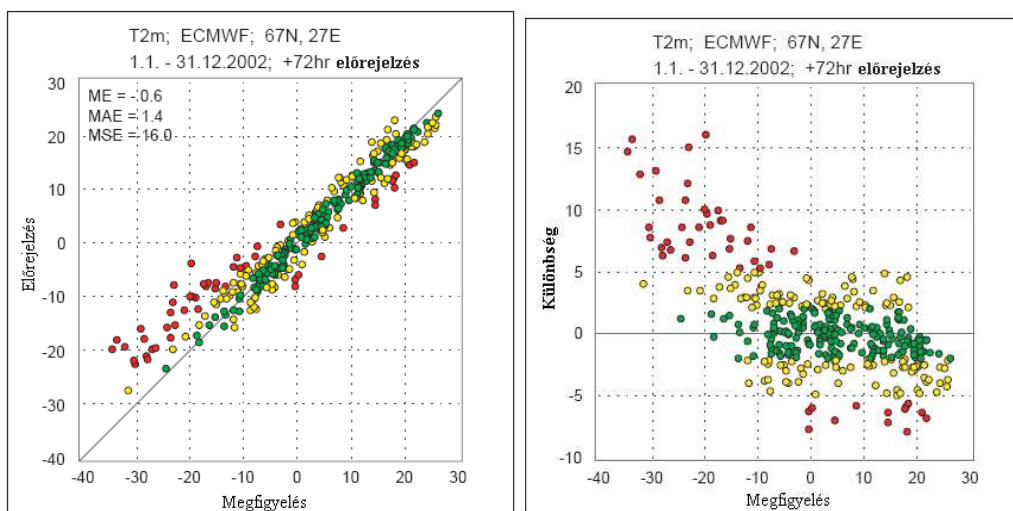
A dolgozatban lehetőség szerint a verifikációs kifejezések megszokott magyar megfelelőjét használom, mindig megadva az eredeti angol kifejezést is. Bizonyos, a hazai gyakorlatban ritkábban alkalmazott verifikációs kifejezések-indexek esetében nem ismert az egyértelmű magyar kifejezés, azonban a kifejezés tartalmát magyar nyelven szintén szeretném visszaadni. A magyar kifejezés dőlt betűvel, míg az angol kifejezés zárójelben dőlt aláhúzott betűvel jelenik meg. Ezenkívül a Függelékben jelölöm, hogy Nurmi mely mérőszámokat javasolja minimálisan szükségesnek, és melyeket optimálisnak.

2.1. Determinisztikus előrejelzések verifikációja

2.1.1. Folytonos eloszlású változók

A folytonos eloszlású meteorológiai paraméterek verifikációja jellemzően arról nyújt információt, hogy az előrejelzett érték mennyivel tér el a megfigyelt értéktől. A verifikációhoz használt leggyakoribb ilyen paraméterek a következők: 2 méteres hőmérséklet (egy időpontban nézve (pl: dél, éjfél), minimum hőmérséklet, maximum hőmérséklet, ill. átlaghőmérséklet, szélsősebesség és szélirány (egy adott időpontban vagy átlagot véve), csapadékösszeg (adott időintervallumra összegezve (pl: 6, 12, 24 óra) mérjük), felhőborítottság (rögzített időpontban, átlagolva; de tipikusan a diszkrét eloszlású paramétereknél szerepel). Azonban ezeknek az elemeknek a viselkedése meglehetősen eltérő lehet, például mikor a hőmérséklet meglehetősen „simán” viselkedik és Gauss eloszlást követ, addig a szélsősebesség tág határok között mozoghat és a csapadék időben nem folytonos meteorológiai változó időszakos, a felhőborítottság U-alakú eloszlást követ.

Az első megközelítésként a megfigyelt és előre jelzett értékek eloszlását ábrázolhatjuk, így láthatóvá válnak a különbségek és a kiugró értékek (*outliers*). Egy előrejelzés/megfigyelés diagramon (1. ábra) ábrázolva a pontos előrejelzés értékei a 45°os átló mentén helyezkednek el. Ehhez hasonlóan ábrázolhatjuk a megfigyeléseket, vagy az előrejelzéseket a különbségeikkel (előrejelzés-megfigyelés). Egy másik módszer az időszorosú (*time-series*) ábrázolás, ekkor könnyen felfedezhetőek a kiugró értékek az előrejelzési vagy a megfigyelési adatsorokban. Ezekben az esetekben az előrejelzés pontosságáról csak vizuális képet kapunk, de számszerű értéket nem.



1. ábra: Egy év háromnapos ECMWF 2m-es hőmérséklet előrejelzésének előrejelzés/megfigyelés (bal), illetve a különbségük a megfigyeléssel (jobb) való diagramja. A piros, zöld és sárga pontok három kategória hibáit választják el.

Számszerű információ előállítására a legkézenfekvőbb megoldás az *átlagos hiba* vagy más néven *szisztematikus hiba* (Mean Error vagy Bias) számítása, a mért és az előrejelzett értékpárok különbségének átlagát számítani:

$$ME = \frac{1}{n} \sum (f_i - o_i)$$

ahol az f_i az i -edik előrejelzést, az o_i az i -edik megfigyelést jelöli. ME mínusz végtelentől plusz végtelenig terjedhet, s a tökéletes előrejelzésnél 0-át venne fel. Ennek a mérőszámnak egyik hibája, hogy ha a nagy hibák ellenkező előjellel szerepelnek a kapott érték akkor is közel lehet a nullához. Ennek a kiküszöbölésére érdemes bevezetni az *átlagos abszolút hibát* (Mean Absolute Error):

$$MAE = \frac{1}{n} \sum |f_i - o_i|$$

Ennek értéke nullától végtelenig terjedhet, s megadja az eltérések átlagos nagyságát az adatsorban. Ezt a két verifikációs mérőszámot érdemes mindig egyidejűleg vizsgálni. Egy másik gyakori mérőszám az *átlagos négyzetes hiba* (Mean Squared Error):

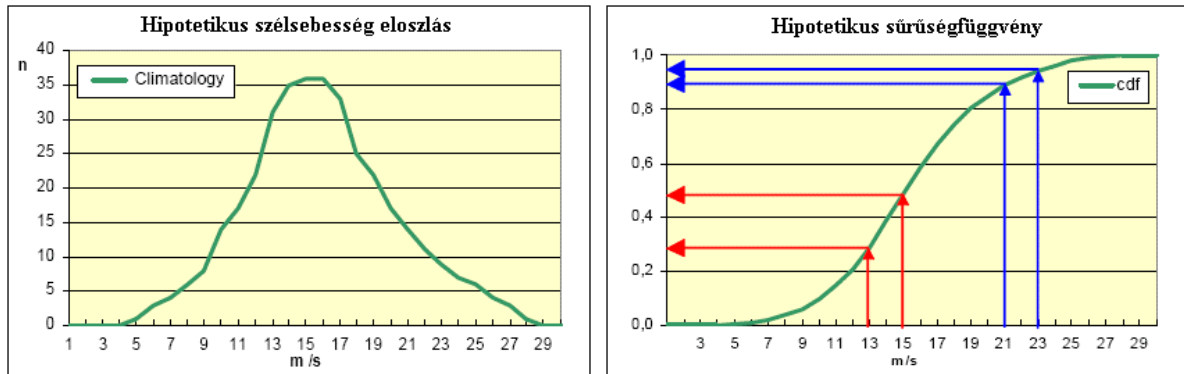
$$MSE = \frac{1}{n} \sum (f_i - o_i)^2$$

illetve ennek négyzetgyöke a (Root Mean Squared Error /RMSE/), aminek a dimenziója megegyezik a paraméter dimenziójával. Ennek értéke is nullától végtelenig terjedhet, s a tökéletes előrejelzésnél nullát venne fel. A négyzetre emelésnek köszönhetően a MSE és a RMSE sokkal érzékenyebb a nagy hibákra, mint a MAE, különösen a kiugró értékeknél, ezért kicsi vagy időben korlátozott adatsoroknál érdekesebb abszolút hibát számolni. Egy másik ajánlott, de még nem elterjedt hibastatisztika a LEPS (Linear Error in Probability Space) nevű verifikációs mérőszám:

$$LEPS = \frac{1}{n} \sum |CDF_o(f_i) - CDF_o(o_i)|$$

amely a lineáris hiba valószínűségi térbeli mértékét jellemzi. A CDF_o a *kumulatív valószínűségi sűrűségfüggvény* (Cumulative probability Density Function of the observations) egy adott meteorológiai paraméter sűrűségfüggvényéből nyert eloszlásfüggvény. Értéke nullától egyig terjedhet, a tökéletes előrejelzésnél itt is nullát venne fel. Segítségével értékelhetjük az előrejelzést, azonban szükség van hozzá a helyes klimatológiai eloszlásra minden egyes pontban. A nagyfokú klimatológiai adatigény miatt az eloszlás függvény előállítása könnyebben megvalósítható a nemzeti meteorológiai szolgálatoknál, mint a nagy nemzetközi modellközpontokban. Az LEPS eredmények azt mutatják, hogy a LEPS jobban

"bünteti" a nagyobb valószínűségű értékeknél ejtett hibát, mint a szélsőséges eseteknél adódókat (2. ábra).



2. ábra: A LEPS alkalmazása egy elméleti szélsébség eloszlásra. Baloldali ábrán az éghajlati eloszlás gyakorisága látható, a jobb oldali ábrán az ebből származtatott valószínűségi eloszlás függvény található.

Az ún. *eredményességi* vagy *beválasí tényező* (*skill score*) segítségével megállapíthatjuk egy viszonyítási/referencia szinthez képest az előrejelzés javulását (a viszonyítási szint általában a klíma állapot és/vagy a megmaradási hajlam)

$$SS = \frac{A - A_{ref}}{A_{perf} - A_{ref}}$$

ahol 'A' a felhasznált hibastatisztika (lehet MAE, MSE vagy LEPS is), A_{perf} értékét a tökéletes előrejelzéskor kapnánk, és az A_{ref} a referencia előrejelzés értéke. A negatív irányítottságú verifikációs mérőszámokra, azaz amikor a kisebb érték jobb előrejelzést jelent, mint az MAE, az MSE, vagy a LEPS esetében a beválasí tényező:

$$SS = 1 - \frac{A}{A_{ref}}$$

A beválasí tényező mínusz végtelentől az egyig terjedhet (egy a tökéletes előrejelzést jelentené). Fontos megjegyezni, hogy kis mintanagyság esetében a beválasí tényező bizonytalan lehet, különösen négyzetes hiba használata során.

2.1.2. Diszkrét eloszlású változók

A diszkrét eloszlású paramétereknél megkülönböztethetünk két- (igen/nem típusú) és több kategóriás előrejelzéseket. A kétkategóriás előrejelzéseknél azt határozzuk meg, hogy az esemény bekövetkezik-e vagy sem. Tipikusan ilyen előrejelzések az eső, a hóesés, az erős szél (ahol különböző küszöb értékeket adhatunk meg), az éjszakai fagy és a köd. Az

előrejelzések verifikációja során első lépésben egy 2*2-es kontingencia táblázatot hozunk létre, ami az előrejelzés és a megfelelő megfigyelés viszonyainak gyakoriságát szemlélteti.

Előrejelzés	Megfigyelés		
	Igen	Nem	Részösszeg
Igen	Találat <i>/Hit/</i> (a)	Téves riasztás <i>/False alarm/</i> (b)	Előrejelzés: igen <i>/Fc/</i> (a+b)
Nem	Elmaradt riasztás <i>/Miss/</i> (c)	Helyes elvetés <i>/Correct rejection/</i> (d)	Előrejelzés: nem <i>/Fc/</i> (c+d)
Részösszeg	Megfigyelés <i>/Obs/</i> igen (a+c)	Megfigyelés <i>/Obs/</i> nem (b+d)	

I. táblázat: Kategórikus események besorolásakor alkalmazott táblázat.

A táblázat alapján az előrejelzés akkor volt megfelelő, amikor *találat (Hit)* vagy *helyes elvetést (Correct rejection)* látunk, illetve nem megfelelő a *téves riasztás (False alarm)*, és a *kihagyás (Miss)* esetén. Tökéletes előrejelzésnél természetesen csak találatot és helyes elvetést tapasztalnánk, a többi cella nulla lenne. Ezekből az értékekből többféle formula alkalmazásával számíthatjuk az előrejelzés beválását.

Egy, a már említett szisztematikus hibához hasonló hibastatisztika előrejelzésekre az FBI (*Frequency Bias Index /FBI/*), amely szisztematikus hiba gyakoriságára utaló index, összehasonlítja az előrejelzett gyakoriságot (F_c igen) és a ténylegesen bekövetkezett események gyakoriságát.

$$FBI = \frac{a + b}{a + c}$$

Az FBI nullától végtelenig terjed, $FBI > 1$ esetén az előrejelzés az indokoltnál többször, $FBI < 1$ esetén kevesebbszer jelzi az eseményt.

A legegyszerűbb módszer, hogy a kategorikus előrejelzés pontosságáról információt kapjunk, a *helyes előrejelzések aránya* verifikációs mérőszám (*Proportion Correct /PC/*):

$$PC = \frac{a + d}{n}$$

Értéke nullától egyig terjed, ahol 1 az ideális érték. A PC félrevezető lehet, mivel egyenlően értékeli az igen és a nem előrejelzést.

Egy másik érték a megfigyelt események (Obs igen) és a találatok arányát vizsgálja. Ez a *találati arány* verifikációs mérőszám (*Probability Of Detection /POD/*).

$$POD = \frac{a}{a+c}$$

Értéke nullától egyig terjed, az ideális érték az egy. *Találati aránynak* (*Hit Rate*-nek /H/) is szokás nevezni. H-nak a komplementere az *elmaradt riasztások aránya* (*Miss Rate*), ami a kihagyott események relatív számát adja meg.

$$MR = 1 - H = \frac{c}{a+c}$$

A POD érzékeny a találatokra, viszont nem veszi figyelembe a téves riasztásokat. Bár mesterségesen lehet javítani „igen” előrejezések túlzott növelésével ezzel egy időben viszont a téves riasztások száma is megnövekszik. Kívánatos a találatok számát maximalizálni, és a téves riasztások számát minimalizálni, így a *találati arányt* (POD-t) célszerű a *hibás találati aránnyal*, azaz a *False Alarm Ratio*-val (FAR-ral) együtt vizsgálni.

$$FAR = \frac{b}{a+b}$$

Értéke nullától egyig terjed, az ideális a nulla. A FAR érzékeny az esemény klimatikus gyakoriságára. A POD-al ellentétben, a FAR a téves riasztásokra érzékeny, viszont nem veszi figyelembe a kihagyásokat. Hasonlóképpen ez is javítható, csak most a „nem” előrejelzés számának növelésével, ezzel csökkentve a téves riasztások számát. POD növekedésével a FAR is növekszik, s csökkenésével a FAR is csökken.

A *False Alarm Rate* (F) a téves riasztások mértékét adja meg a nem bekövetkező eseményekhez (Obs nem) képest (*Probability Of False Detection*-ként /POFD/ is ismert):

$$F = \frac{b}{b+d}$$

Értéke szintén nullától egyig terjed, ahol a nulla az ideális. F-et a H-val kombinálva kaphatjuk az ún. ROC (*Relative Operating Characteristic*) verifikációs mérőszámot, röviden ROC diagramként a valószínűségi előrejelzésekre is szokták használni (részletesebben a 2.2 fejezetben mutatjuk majd be). Azonban a ROC-ot lehet a kategorikus előrejelzésekre is alkalmazni, ezáltal közvetlenül össze lehet hasonlítani a valószínűségi előrejelzéssel. A H és F népszerű beválási tényezője, több megalkotójával, a *Hanssen-Kuipers Skill Score* (KSS), vagy *True Skill Statistic* (TSS), vagy *Peirce Skill Score* (PSS), aminek a legegyszerűbb formája:

$$KSS = H - F \text{ vagy } KSS = POD - POFD$$

Értéke mínusz egytől egyig terjed, a tökéletes előrejelzésnél egyet venne fel. Ideális esetben a KSS, képes az igen esetek (H) és a nem eseteket (F) események elkülönítésére. Ritka eseményekre a helyes elvetések száma (d) igen magas a többihez képest, ami alacsony téves

riasztáshoz vezet, amiből következik, hogy KSS közel lesz POD-hez. A ritka eseményekre széles körben használt a Threat Score (TS), vagy (Critical Success Index /CSI/):

$$TS = \frac{a}{a + b + c}$$

Értéke nullától egyig terjed, az ideális érték az egy. A TS érzékeny a találatokra, viszont nem veszi figyelembe sem a téves riasztásokat, sem a kihagyásokat. A TS szintén érzékeny az esemény gyakoriságára, rosszabb értéket adva a ritkább eseményekre, mivel néhány találat véletlen eséllyel is bekövetkezhet. Ezt a tulajdonságot nem mutatja, egy hasonló mérő szám az Equitable Threat Score (ETS) (amit Gilbert's Skill Score-ként (GSS) is ismert), ahol a találatokat összekapcsolják a véletlen esélyeikkel:

$$ETS = \frac{a - ar}{a + b + c + ar},$$

ahol ar a találatok száma a véletlen előrejelzésre.

$$ar = \frac{(a + b)(a + c)}{n}$$

Az ETS -1/3-től egyig terjedhet, az ideális érték egy.

Az egyik leggyakrabban használt beválási tényező a 2*2-es kontingencia táblázat összefoglalására a Heidke Skill Score (HSS).

$$HSS = \frac{2(ab - bc)}{(a + c)(c + d) + (a + b)(b + d)},$$

amely mínusz végtelentől egyig terjed, ideális értéke egy lenne.

Az Odds Ratio (OR) az előrejelzés találatának valószínűségét (POD-t vagy H-t) hasonlítja össze a téves riasztás valószínűségével (POFD-vel vagy F-vel).

$$OR = \frac{H/1 - H}{F/1 - F} = \frac{ad}{bc}$$

Értéke nullától végtelenig terjed, tökéletes előrejelzésnél végtelenhez tartana, ha az érték egynél nagyobb, akkor a POD (a találatok aránya) felülmúlja a FART (téves riasztások arányát). Az Odds Ratio átírható egy -1-től 1-ig terjedő beválási tényezőbe:

$$ORSS = \frac{OR - 1}{OR + 1} = \frac{ad - bc}{ad + bc}$$

Azonban ezt ritkán használják a verifikációs gyakorlatban.

Természetesen a diszkrét eloszlású változók kettőnél több kategóriába is sorolhatók, (pl. a borultság mértéke 0-2, 3-5 vagy 6-8 okta; a csapadék alakja eső, jégeső vagy hó). A több kategóriás események beválását egy multi-kategóriás esemény táblázat segítségével tanulmányozhatjuk. A Proportion Correct (PC) közvetlenül használható kettőnél több

kategóriára is. A többi mérőszám alkalmazásához, a multi-kategóriás kontingencia táblázatot 2*2-es táblázatok sorozatává kell alakítani. A KSS és a HSS skill score-okat általánosíthatjuk multi-kategóriás esetre is:

$$KSS = \frac{\sum p(f_i, o_i) - \sum p(f_i)p(o_i)}{1 - \sum (p(f_i))^2}$$

$$HSS = \frac{\sum p(f_i, o_i) - \sum p(f_i)p(o_i)}{1 - \sum p(f_i)p(o_i)}$$

ahol az i index a tábla dimenziója, $p(f_i)p(o_i)$ az előrejelzés és a megfigyelés együttes eloszlása, $p(f_i)$ és $p(o_i)$ az előrejelzés és a megfigyelés határ eloszlásának valószínűsége.

2.2. A valószínűségi előrejelzések verifikációja

A valószínűségi előrejelzés soha nem lehet "jó" vagy "rossz", kivéve mikor 0 vagy 100%-os valószínűséget jelzünk előre. Ennek következtében a valószínűségi előrejelzések verifikálásához nagy mennyiségű előrejelzés szükséges, hogy kellően pontos következtetést vonhassunk le. Érdeemes megjegyezni, hogy a valószínűségi előrejelzések - egy küszöbérték megválasztásával - átalakíthatók kategórikus előrejelzéssé, fordítva az átalakítás azonban nem egyértelmű.

A determinisztikus előrejelzésekhez hasonlóan a valószínűségi előrejelzésnél is egyszerűen definiálhatunk, a négyzetes középhiba mintájára, egy verifikációs mérőszámot, ami számszerű értéket ad az előrejelzés bevalásáról. Ezt *Brier Score*-nak (BS-nek) nevezzük, ami egy adott esemény előrejelzésének valószínűsége és tényleges előfordulásának gyakorisága közötti eltérések négyzetének átlaga, így tehát n számú előrejelzés-megfigyelés párra felírva:

$$BS = \frac{1}{n} \sum (p_i - o_i)^2$$

ahol p_i az adott esemény előrejelzett valószínűsége és o_i értéke 1 vagy 0, attól függően, hogy előfordult az esemény, vagy nem az adott megfigyelés alkalmával. Napjainkban használatos verifikációs módszerek esetén a fent leírt mennyiséget tekintik a Brier Score-nak, azonban eredetileg a Brier a komplementer esemény score-ját is hozzáadta ehhez a mennyiséghez.

1973-ban Murphy a BS új megközelítését vezette be. A valószínűségi tartományt részekre osztotta fel, s a Brier Score-ban szereplő p_i -ket e résztartományokba sorolta be, és helyettük a résztartományra jellemző átlagos valószínűséget vette. A megfigyelésekre

vonatkozóan pedig megszámlolta, hogy az egyes intervallumokba eső előrejelzések esetén hányszor valósul meg az esemény. Ebből kapta a következő képletet:

$$BS = \frac{1}{n} \sum_{k=1}^T \sum_{i=1}^{n_k} (p_k - o_{ki})^2$$

ahol T a résztartományok száma, n_k a k -adik résztartományba eső előrejelzések száma, p_k a k -adik intervallumba eső előrejelzések átlagos valószínűsége, o_{ki} pedig a k -adik intervallumban az i -edik előrejelzés esetén az előbbi o_i . Murphy igazi felfedezése azonban abban áll, hogy ezt az egyenletet átalakította, s a következőt kapta:

$$BS = \frac{1}{n} \sum_{k=1}^T \sum_{i=1}^{n_k} [(p_k - o_k) + (o_k - o) + (o - o_k)]^2$$

ahol o_k a k -adik intervallumhoz tartozóan az esemény bekövetkezésének relatív gyakorisága, o az egész mintára vonatkozóan az esemény bekövetkezésének relatív gyakorisága. Ezt az egyenletet átírva kapjuk:

$$BS = \frac{1}{n} \sum_{k=1}^T n_k (p_k - o_k)^2 - \frac{1}{n} \sum_{k=1}^T n_k (o_k - o)^2 + o(1 - o)$$

ahol az első tag a *megbízhatósági* (*reliability*), a második a *felbontási* (*resolution*), a harmadik pedig a *bizonytalansági* (*uncertainty*) tag.

A megbízhatósági tag intervallumonként összegzi az átlagos előrejelzett valószínűség és az esemény bekövetkezésének relatív gyakoriságának négyzetes eltérését. Jó Brier Score-t akkor kapunk, ha ez a tag is a lehető legkisebb, azaz az előrejelzett valószínűségek és a bekövetkezések gyakorisága közel van egymáshoz. Értéke 0 és 1 között lehet.

A felbontási tag szintén intervallumonként összegzi az intervallumra és egész mintára vonatkozó bekövetkezési gyakoriság négyzetes eltérését. Mivel ezt a tagot kivonjuk az előzőből, ezért akkor kapunk jobb előrejelzést, ha ez a tag a lehető legnagyobb. Ennek a tagnak az értéke sosem nagyobb, mint a 3. tag, s minimális értéke pedig 0.

A bizonytalansági tag az esemény adott időszak során való bekövetkezésének gyakoriságával van kapcsolatban. Ez a tag mint látható nincs összefüggésben az előrejelzéssel, csakis az esemény klimatológiai valószínűségével kapcsolatos. Tehát mindig megnöveli a mérőszám értékét, s 0-tól 0,25-ig vehet fel értékeket. Egy adott esemény verifikálása során ezt nem tudjuk kiküszöbölni, és a verifikáció során mindig figyelembe kell venni, hogy az esemény mennyire ritka, nem hasonlítható össze két jelentősen eltérő gyakoriságú esemény.

E mennyiségek grafikai ábrázolására szolgál a *megbízhatósági diagram* (*reliability diagram*) (3. ábra). Az x-tengelyen az előrejelzett valószínűségek, az y-tengelyen pedig az intervallumokra vonatkozó megfigyelt esemény-gyakoriságok vannak feltüntetve. Az így kapott megbízhatósági görbe, tehát kapcsolatot mutat az átlagos előrejelzett valószínűség és az átlagos megfigyelt gyakoriság között. Fel szokták tüntetni, hogy a pontok hány esetet reprezentálnak, például az egyes intervallumoknál az esemény megvalósulásának abszolút gyakoriságának diagramjával. A tökéletes megbízhatóságot (*perfect reliability*) akkor éri el, mikor az előrejelzett valószínűségek a megfelelő megfigyelt gyakorisággal megegyeznek, tehát a 45 fokos átló mentén helyezkednek el. Ennek a tagnak a jósága szemléletesen azt jelenti, hogy a görbe az átlóhoz közel húzódik. Ha az átló alatt halad a vonal, akkor fölébecslést, ha alatta, akkor alábecslést mutat az ábra.

A vízszintes szaggatott vonal az esemény adott időszak során mutatott gyakorisága, azaz a képlet 0 tagja. Az ezen a vonalon húzódó görbének nem lenne felbontása, hiszen az intervallumonkénti esemény-gyakoriságok megegyeznének az egész időszakra vonatkozó esemény-gyakorisággal. Ez azt jelenti, hogy bármekkora valószínűséget is adtunk, az esemény mindig ugyanazzal az 0 gyakorisággal következett be. Kissé szemléletesebben: akkor jó a felbontás tag, ha eltérő valószínűségek esetén eltérő mértékben valósul meg az esemény is. Itt nem arról van szó, hogy az előrejelzett valószínűséggel következik be az esemény, azaz egy jó felbontású előrejelzésnek nem kell megbízhatónak is lennie.

Megemlíthetjük még az *élesség* (*sharpness*) fogalmát, ami azt jelenti, hogy az előrejelzés éles, ha gyakran ad 0 és 100% körüli értékeket. Ebben az esetben mindegy, hogy ebből mennyi vált be. Az élesség jelzi, hogy a modell képes az extrém események előrejelzésére is.

Az előzőekhez hasonlóan, itt is előállíthatunk egy eredményességi tényezőt, amihez szükséges egy BS_{ref} referencia előrejelzés Brier Score-ja, ami lehet egy kontroll vagy egy tökéletes, vagy a leggyakrabban használt, az adott esemény klimatológiai előrejelzése.

$$BS_{ref} = \frac{1}{n} \sum (ref_i - o_i)^2$$

ahol ref_i legtöbbször az eseményre vonatkozó klimatikus gyakoriság. Így a Brier Skill Score (BSS) a következő

$$BSS = 1 - \frac{BS}{BS_{ref}}$$

A Brier Skill Score mínusz végtelentől egyig vehet fel értékeket, ahol is az egy a tökéletes előrejelzés esete. Ha az érték nullánál kisebb, akkor az előrejelzés nem jobb, mint a

klímaátlag előrejelzése. Ezt akkor érheti el, ha a megbízhatósági és a felbontási tag megegyezik.

A Brier Score-t általánosítható több eseményű vagy több-kategóriás helyzetekre, az RPS (*Ranked Probability Score*) verifikációs mérőszám a kategóriák szerint sorba rendezett események arányát mutatja

$$RPS = \frac{1}{k-1} \sum \{ (\sum p_i) - (\sum o_i) \}^2$$

ahol k a valószínűségi kategóriák száma A megfelelő skill score pedig

$$RPSS = 1 - \frac{RPS}{RPS_{ref}}$$

Mind a BSS, mind pedig a RPSS nagyon érzékeny az adathalmaz méretére.

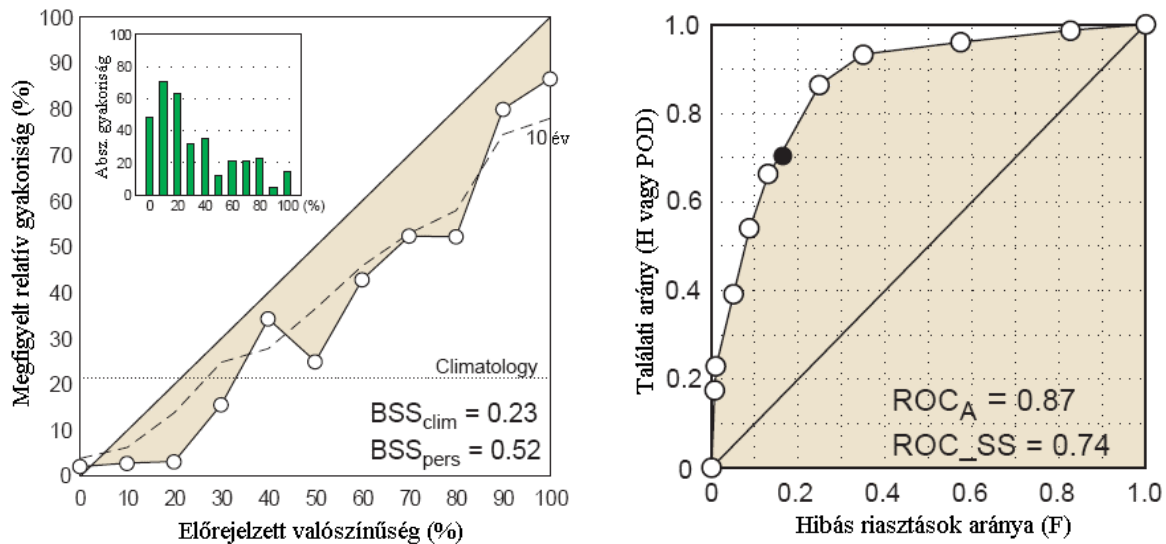
Széleskörben használt eszköz a ROC (*Relative Operating Characteristic*) diagram (3. ábra). A ROC diagram az előrejelzés megbízhatóságáról nyújt információt. A ROC görbe grafikusán ábrázolja a találatokat (H) a *False Alarm Ratio*-val (F) szemben, különböző valószínűségi küszöb értékekre. A ROC görbét valószínűségi előrejelzésből szerkeszthetjük meg, úgy hogy léptetjük a valószínűségi küszöb értéket (például 10%-os valószínűségi intervallumonként) végig a valószínűségi előrejelzésen. Ezáltal, mindegyik küszöbértékhez készíthetünk egy 2x2-es kontingencia táblázatot, így tehát kategorikus igen/nem előrejelzéssé alakítva a valószínűségi előrejelzést. Jó előrejelzésnél a H magas, míg F alacsony értékeket vesz fel, minél közelebb vannak a pontok a felső bal sarokhoz, annál jobb az előrejelzés. Tökéletes előrejelzésnél csak korrekt előrejelzések lennének téves riasztás nélkül, s így a ROC görbe a (0,0) pontból emelkedne végig az y tengely mentén a (0,1) pontig, majd egyenesen az (1,1) pontba menne.

Egy jó, széles körben használt relatív index a *ROC terület* (*ROC area* / ROC_a), ami a ROC diagramon alapul, a görbe alatti területet felhasználva, s az ehhez kapcsolódó eredményességi tényező (ROC_{SS}). A tökéletes előrejelzésnél a ROC_a 1-et venne fel. Ahogy a görbe egyre lejjebb helyezkedik el, úgy csökken a prognózis értéke, míg el nem éri a használhatatlan szintet, ami az egyenes vonal az átló mentén ($H=F$), és a terület 0,5-tel egyenlő, ugyanis ekkor az előrejelzés nem tudja megkülönböztetni az esemény bekövetkezését a nem bekövetkezésétől. A ROC_a eredményességi tényezője a következőképp definiálható:

$$ROC_{SS} = 2 * ROC_a - 1$$

Az átló alatt a ROC_{SS} negatív értékeket vesz fel, -1-nél elérve a minimumát, mikor is a $ROC_a=0$.

A ROC diagramot alkalmazhatjuk kategorikus igen/nem előrejelzésekre is, azonban ebben a speciális esetben, csak egy küszöbérték van, s csak egy pont szerepel a grafikonon. A ROC, ROC_a és a ROC_{SS} hibastatisztikák előnye, hogy közvetlenül kapcsolhatók a döntéshozatalhoz a végfelhasználó számára, ezzel gazdasági értéket is hordozva.



3. ábra: Megbízhatósági diagram (bal ábra) és ROC diagram (jobb ábra) egy év csapadék valószínűségi előrejelzésre. A megbízhatósági diagramba beágyazott kis diagram az előrejelzett valószínűségek gyakoriságát mutatja. A vízszintes pontozott vonal az éghajlati valószínűség. A megbízhatósági görbe (a körökkel) erős túlbecslést jelez, hasonlóan a 10 éves átlagot jelző görbéhez (szaggatott vonal). A jobb oldali ROC diagram az előrejelzés és a megfigyelt értékek alapján, különböző küszöb értékeket választva lett megszerkesztve. A ROC_a és a ROC_{SS} szintén fel van tüntetve.

A valószínűségi előrejelzések verifikációjához gyakran használt másik diagram a Talagrand diagram. A korlátozott számú EPS tag ellenére a verifikációs analízis túlnyúlhat az ensemble tartományon. Egy 50 tagú előrejelzésnél ez esetek 2/51-ed részében (~4%) kellene, hogy megtörténjen. A valóságban a 2000-es évek elején végzett verifikációs vizsgálatok szerint kb. 10 % esik kívül. Ez azt jelenti, hogy az EPS nem húzódik eléggé szét. Az *EPS értéktartomány* (*EPS spread*) elemzésére szolgál a Talagrand diagram. A diagram azon a feltevésen alapul, hogy az ideális EPS rendszerben a mért tag a két szomszédos ensemble tag közé egyenlő gyakorisággal esik, beleértve azt az esetet is, mikor kívül esik a tartományon. Ennek az ideális rendszernek a Talagrand diagramja lapos lenne, minden intervallumban egyenlő gyakoriságot mutatva. Azonban a valóságban a diagram némi U alakot mutat, ekkor a megfigyelés gyakrabban esik az előrejelzési tartományon kívül, és kevesebbszer esik középre, mint ahányszor előre lett jelezve. Némely paraméternél az U alak helyett J alakot figyelhetünk

meg, ami a paraméter szisztematikus hibájára utal. Az előrejelzés javításához az EPS értéktartomány (*EPS spread*) növelése szükséges, de nem elégséges; véletlenszerű mintavételezése az időjárási paramétereknek azonos időszakból az elmúlt 50 évben szintén lapos Talagrand diagramot eredményezne, de természetesen előrejelző érték nélkül.

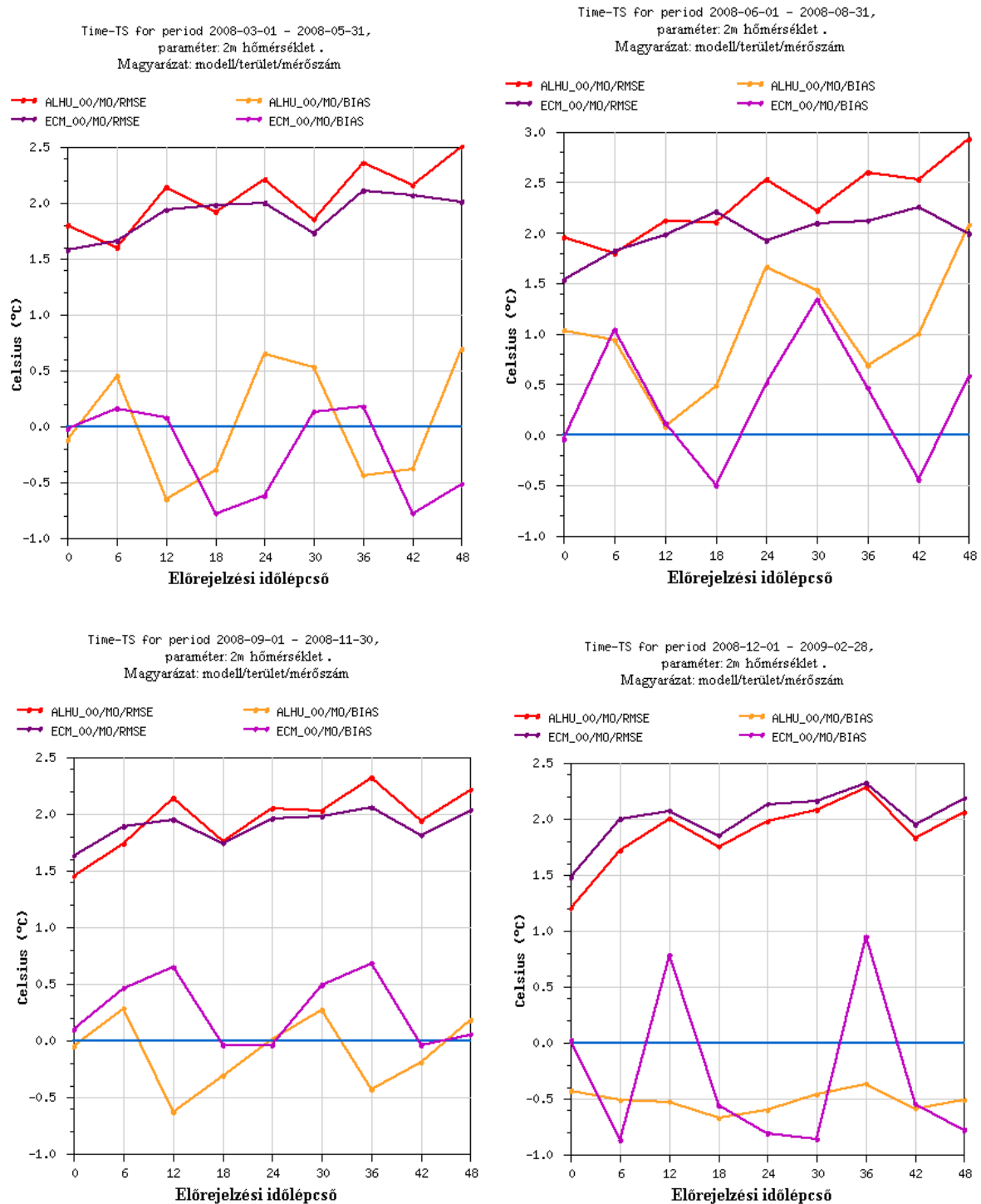
3. A verifikáció alkalmazása

3.1. Verifikációs eredmények ábrázolása

A különböző előrejelző modellek verifikációs mérőszámait grafikonon ábrázolva szemléletesen összehasonlíthatjuk egymással. Munkám során a verifikációs mérőszámok ábrázolására az Országos Meteorológiai Szolgálat által kifejlesztett szoftvert, az Objektív Verifikációs Ábrakészítő Rendszert, röviden OVISYt (*Objective Verification System*) használtam. Az OVISYS az ECMWF és az ALADIN/HU modell előrejelzéseit és szinoptikus (SYNOP) illetve rádiószondás (TEMP) megfigyelések adatait használja fel. A programmal egyszerűen és gyorsan számolhatjuk és ábrázolhatjuk a determinisztikus modellek hibastatisztikáit. A felszín közeli változók a hőmérséklet, szélirány, szélsébség, nyomás, felhőborítottság, relatív nedvesség, minimum és maximum hőmérséklet, szélökés és csapadék összeg (6, 12 és 24 órás). A TEMP főizobár szinteken pedig a választható paraméterek a hőmérséklet, geopotenciál, szélsébség, szélirány, valamint relatív nedvesség. Ezekből az adatokból négyzetes hibát, ennek a négyzet gyökét, átlagos abszolút hibát, szisztematikus hibát számolhatunk. A kontingencia táblázatnál a következő értékek kerülnek kiszámításra: ME, POD, FAR, PC, F, KSS, TSS, ETS, OR, ORSS.

Az OVISYS programmal több ábratípus is könnyen készíthető, a kiválasztott paramétertől függően, ilyen a pont diagram, vagy az idősoros diagram. A hibákat mind a napok, mind az állomások függvényében ábrázolhatjuk, a hibastatisztika értékeit elhelyezhetjük térképen, a szélirányokról is készíthetünk ábrát, illetve kontingencia táblázat is létrehozható. Az ábrázolás folyamata során először kiválasztjuk az időszakot, a modellt, az időlépcsőt, a területet vagy állomást, a verifikálni kívánt paramétert és az ábratípust is, majd a program a kiválasztott ábrának megfelelően kiolvassa az adatokat az adatbázisból, és a megadott hibastatisztika kiszámolása után ábrázolja az eredményt. Vizsgálataim során a 00 UTC-kor futtatott ECMWF és ALADIN/HU modell előrejelzéseit hasonlítottam össze. Fontos megjegyeznünk, hogy az ECMWF modell rácsfelbontása $0,5^{\circ} \times 0,5^{\circ}$, míg az ALADIN/HU modellté $0,1^{\circ} \times 0,1^{\circ}$.

A verifikációt a 2008-as időszakokra és Magyarország területére végeztem, 2m-es hőmérsékletre szisztematikus hibát és az átlagos négyzetes hiba négyzetgyökét számoltattam, majd idősoros diagramon ábrázoltam, ahol az x tengely az előrejelzési időtávot jelöli (4. ábra).

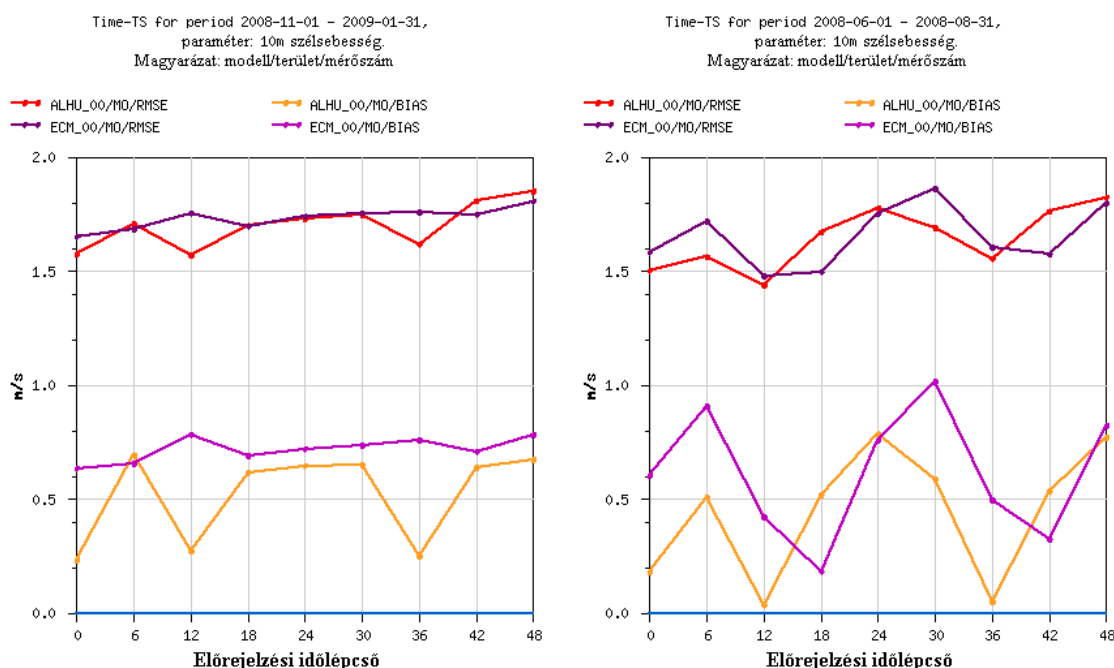


4. ábra: A 2 m hőmérsékletre vonatkozó ME és RMSE hibák évszakonkénti időbeli menete Magyarország területére a 00UTC-s ECMWF és ALADIN/HU modellek alapján, (2008 tavasz, nyár, ősz, tél).

6 órás lépcsőközöket választottam, s a 48. óráig számoltattam ki őket, mivel az ALADIN/HU modell +48 óráig terjedő előrejelzéseket tartalmaz.

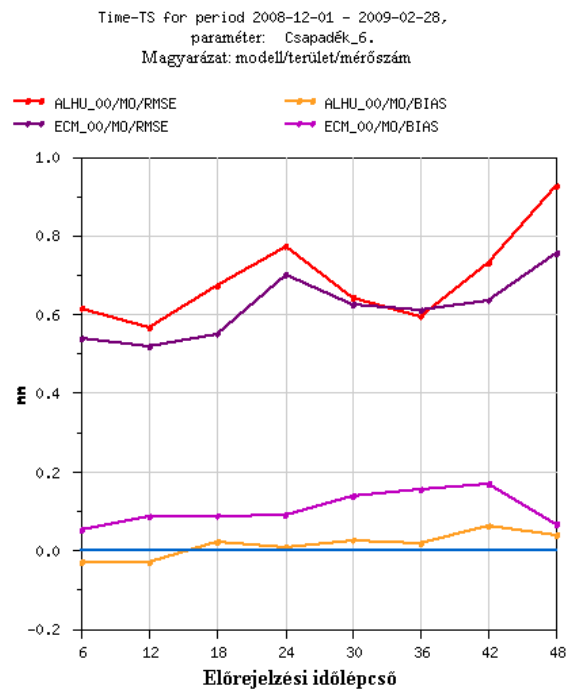
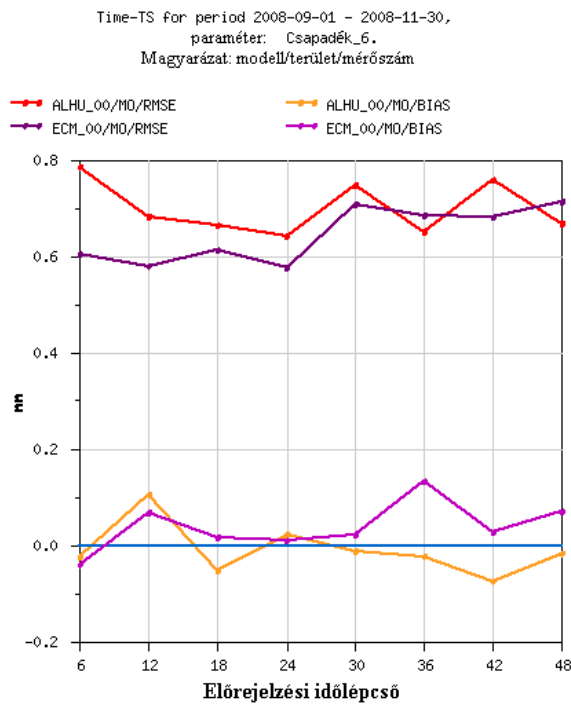
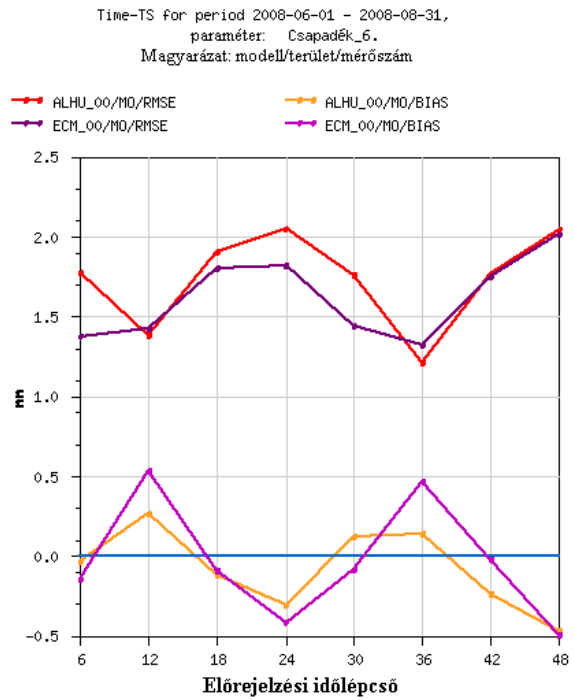
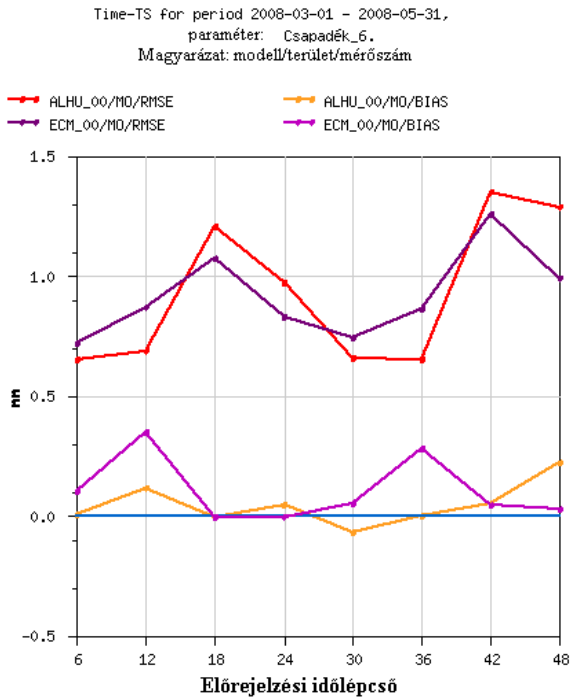
Megfigyelhető, hogy az RMSE enyhén emelkedik az időtáv előre haladtával és az ALADIN/HU modell hibája nagyobb a téli időszak kivételével. A szisztematikus hiba (ME) napi menetet, ciklikusságot mutat, kivéve télen az ALADIN/HU modellt, amikor minden időlépcsőben alábecslés jellemzi a modellt.

A 10m-es szélességre készített ábrán, viszont az figyelhető meg, hogy az ECMWF modell RMSE hibája a nagyobb, s a modellek felébecslését láthatjuk (5. ábra).



5. ábra: A 10 m-es szélességre vonatkozó ME és RMSE hibák időbeli menete Magyarország területére a 00UTC-s ECMWF és ALADIN/HU modellek alapján, (2008 nyár, tél).

Ugyanezekkel a feltételekkel elkészítettem a 6 órás csapadékösszegre vonatkozó ábrákat is. A csapadék esetében a két modell hasonló viselkedést mutat, az ECMWF modell az időlépcsők döntő többségében mind a szisztematikus, mind az RMSE hiba szerint jobb bevalású. Mindkét modell esetében az RMSE hibája az őszi, téli időszakban a legalacsonyabb, s nyáron a legmagasabb (6. ábra). Ennek az oka valószínűleg az, hogy a modellek nem tudják pontosan visszaadni a konvektív csapadékképződést.



6. ábra: A 6 órás csapadék összegre vonatkozó ME és RMSE hibák évszakonkénti időbeli menete Magyarország területére a 00UTC-s ECMWF és ALADIN/HU modellek alapján, (2008 tavasz, nyár, ősz, tél).

3.2. Az ECMWF előrejelzések verifikációja

Az ECMWF kérésére az Országos Meteorológiai Szolgálat minden évben elkészíti, az OVISYS program felhasználásával, a szolgálatnál operatívan használt ECMWF produktumok verifikációját. Az ECMWF előrejelzések időbeli skálája a rövidtávú előrejelzésektől a középtávú előrejelzéseken át a szezonális előrejelzésekig terjed. Külön vizsgálják a nyers modell eredményeket, az utófeldolgozott, és a végfelhasználóknak szánt eredményeket, illetve az ECMWF determinisztikus modell a rövidtávú skálán az ALADIN/HU modellel is összehasonlításba kerül.

A direkt ECMWF modell outputnál a 00 és a 12 órás futtatások eredményeit a magyar szinoptikus megfigyelő hálózat megfigyeléseivel összehasonlítva a következő paraméterek kerülnek verifikálásra: 2m-es hőmérséklet, 2m-es relatív nedvesség, 10m-es szélesség, teljes felhőborítottság, a 2m-es minimum és maximum hőmérséklet, valamint a napi csapadékösszeg. A verifikáláshoz szisztematikus hibát és a négyzetes hiba gyökét számolják, majd ábrázolják idősoros diagramon 168 óráig terjedő időskálán vizsgáljuk.

A numerikus modellek által szolgáltatott nyers eredmények nem mindig alkalmasak arra, hogy közvetlenül felhasználóknak bocsássák. Ennek több oka is lehet. Egyrészt a nyers modell outputok nem mindig a felhasználó igényeinek megfelelőek, vagyis más koordináta-rendszerben megadott információra, illetve más paraméterek előrejelzett értékeire lehet szüksége a felhasználónak. Másrészt előfordulhat, hogy a számítási eredmények szisztematikus hibával terheltek, és vannak olyan esetek is, amikor a felbontás nem megfelelő az elvárásoknak. Ezeket a problémákat próbálják korrigálni a különböző utó-feldolgozási eljárások révén (*Csima, 2003*). Az előrejelzések bevalásának javítása érdekében alávethetjük egy kalibrációs eljárásnak (*Mile, 2007, Üveges, 2008*). A kalibráció történhet eloszlásfüggvénnyel, ami tetszőleges meteorológiai paraméterre alkalmazható.

A külső felhasználók számára az előrejelzéseket az ECMWF modell előrejelzések felhasználásával az előrejelző szakemberek készítik, s munkájuk révén statisztikai értelemben többnyire kis mértékben képesek javítani az előrejelzés pontosságán. Legnagyobb mértékben télen tudnak javítani az eredményeken, mivel a numerikus modellek nem mindig tudják jól visszaadni a téli inverziót, és az ekkor fellépő alacsony felhőzetet. Ilyenkor az átlagos 5-10% helyett, 20-32% az előrejelzések javulása.

A jelentésben az évszakos, és a havi előrejelzésekről is olvashatunk. A szezonális előrejelzés verifikációja során szisztematikus hibát, átlagos abszolút hibát és az átlagos négyzetes hiba gyökét számolják a havi átlagos 2m-es minimum és maximum hőmérsékletre

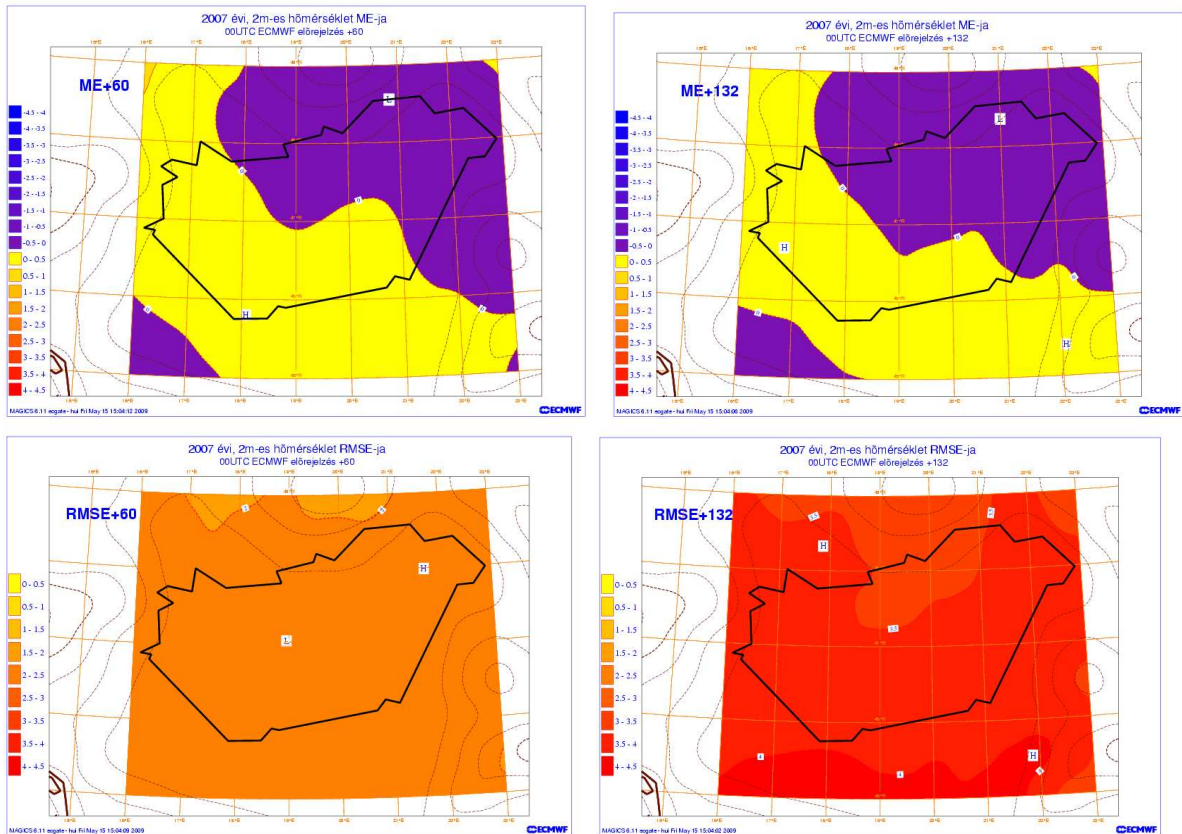
és havi csapadékösszegre vonatkozóan. A verifikációt Magyarország teljes területére és hat régiójára végzik el. A referencia adatbázis, amiből az átlagos hiba skill score-ját számolják az 1961-1990-es időszak éghajlati átlaga. Havi előrejelzések 2002 óta készülnek. A havi előrejelzések verifikációja szintén 6 régióra, és az egész ország területére készül.

4. Hőmérséklet előrejelzések verifikációja

4.1. A determinisztikus előrejelzés verifikációja

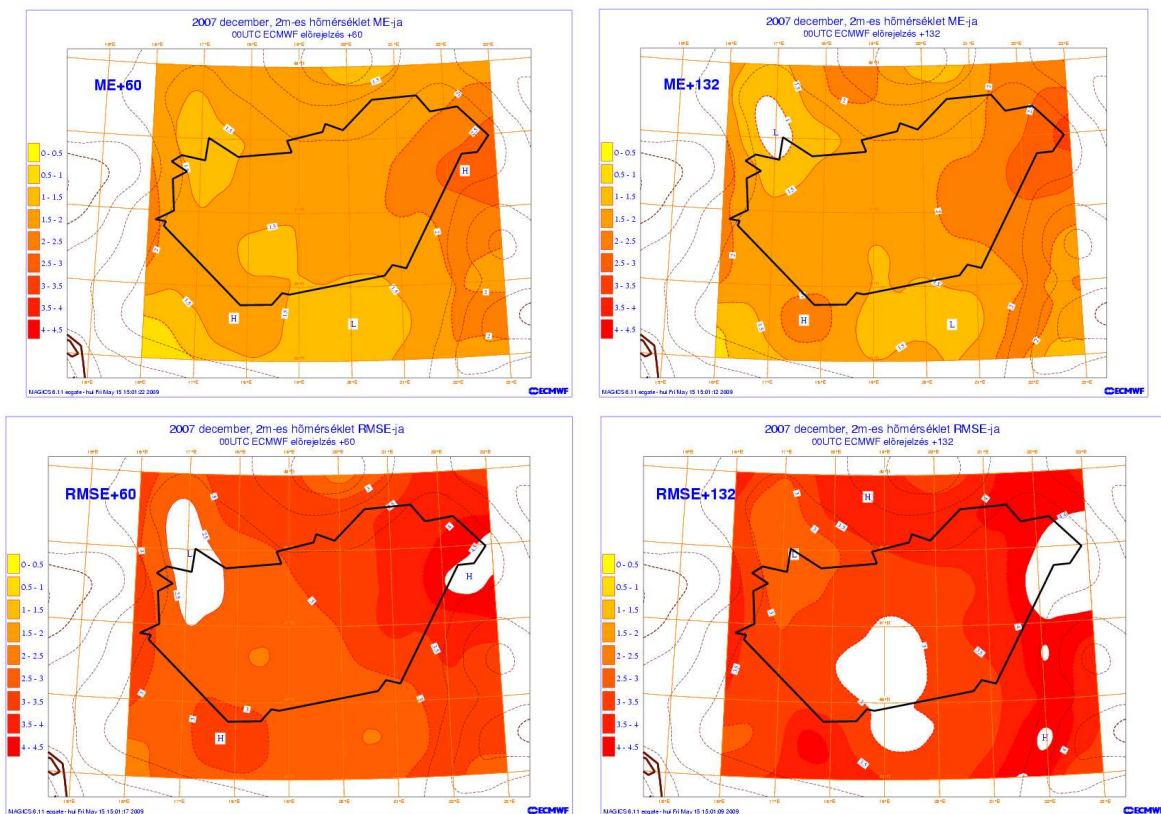
UNIX/FORTRAN program nyelvben az ECMWF determinisztikus előrejelzését és az EPS (Ensemble Prediction System) valószínűségi előrejelzéseire verifikációs programot fejlesztettem ki. Az adatok ECMWF MARS (Meteorological Archive and Retrieval System) rendszeréből származnak. Az ECMWF produktumok az operatív futtatás során azonnal archiválódnak, így rögtön elérhetők a felhasználók számára (*Raoult, 2001*). Egy 0,5x0,5 fokos földrajzi rácshálózaton Magyarország térségében egy $9*15=135$ rácspontból álló területet választottam ki, a 2007-es évre vonatkozóan, a minden nap 00 UTC-kor készített +60 és +132 órás determinisztikus és ensemble 2m-es hőmérséklet előrejelzéseket verifikáltuk, a determinisztikus modell analízissel szemben. A program segítségével minden rácspontra a szisztematikus hibát és a négyzetes hiba négyzetgyökét számoltam egész évre és havi bontásban. Ezután az adatokat, az ECMWF által fejlesztett Magics programmal térképen ábrázoltuk a könnyebb elemzés érdekében.

Már első ránézésre is szembetűnik, amint várható is volt, hogy az előrejelzési időtáv előrehaladtával nő a modell hibája. Ez a szisztematikus hiba sajátossága miatt, miszerint ellentétes előjelű nagy hibák gyengítik egymást, elsősorban az RMSE hibában tűnik fel. Az egész évre vonatkozó ábráról (7. ábra) leolvasható, hogy az RMSE a +60 órás időtávnál 2 körül van, addig a +132 óránál 4 feletti értéket is felvesz, míg ME hibája itt szinte nem is változva, nulla körüli értéket vesz fel. A szisztematikus hiba értéke a legtöbb hónapban negatív, illetve nullához közeli.



7. ábra: A 2007-es évre vonatkozó szisztematikus és RMSE hiba, 2m-es hőmérséklet, +60 és +132 órás időtávra, a determinisztikus modell alapján.

Megfigyelhető, hogy a december RMSE hibája nagyobb, mint más hónapoknál, mind a +60, mind a +132 órás időtávnál (8. ábra). A szisztematikus hiba szintén nagyobb, s az egész területen pozitív értéket vesz fel, ami a hőmérséklet felül becslését sejteti. Ezt valószínűleg a téli inverziós réteg kialakulása eredményezheti, amikor is egy alacsony felhőréteg alakulhat ki, s maradhat fent hosszabb időn keresztül. Ez az alacsony felhőzet a maximum hőmérséklet felülbecsléséhez, s a minimum hőmérséklet alulbecsléséhez vezet. Észrevehető még, mindkét időtávnál, hogy a júliusi hónappal vesz fel a szisztematikus hiba magasabb értékeket.



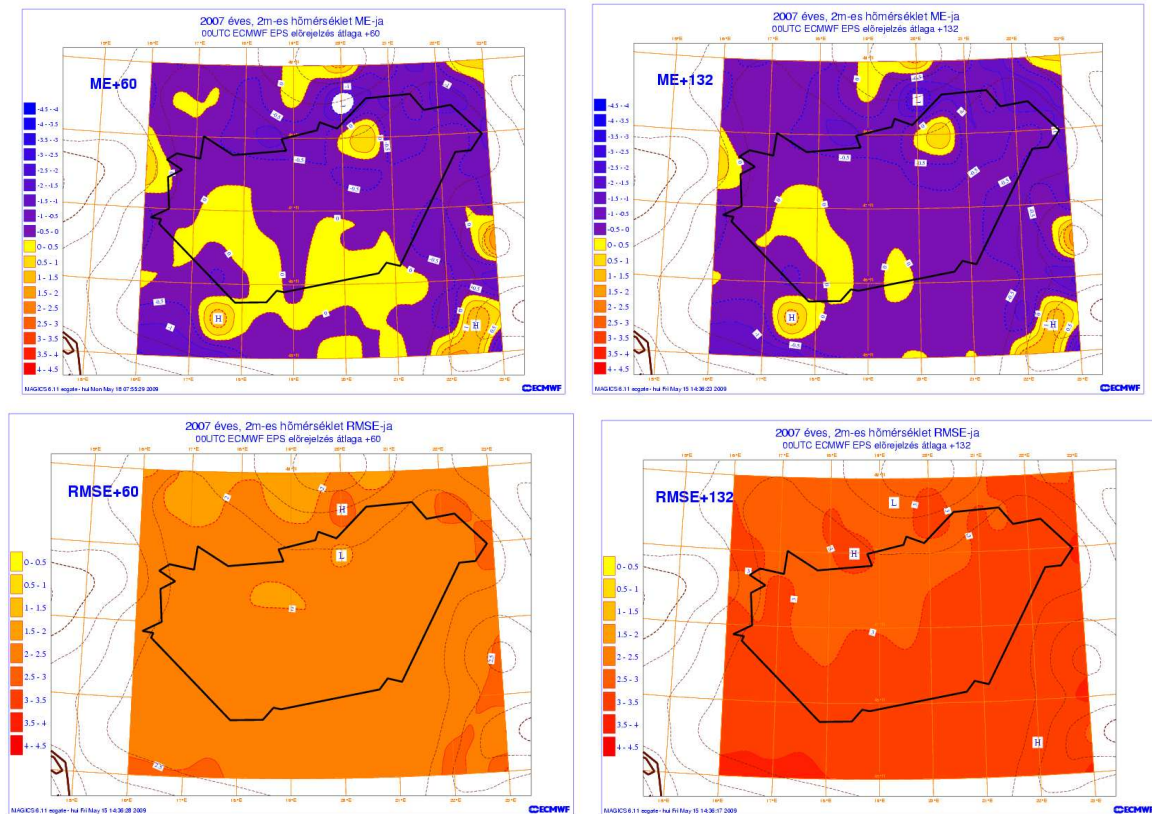
8. ábra: A 2007 decemberi szisztematikus és RMSE hiba, 2m-es hőmérséklet, +60 és +132 órás időtávra, a determinisztikus modell alapján.

4.2.A valószínűségi előrejelzés verifikációja

A determinisztikus előrejelzéshez hasonlóan, a valószínűségi előrejelzésre is elvégeztem a verifikációt. Itt is, mint a determinisztikus előrejelzésnél, hőmérsékleti adatokat verifikáltam a 2007-es évre a +60 és a +132 órás időtávra. Azonban különbség a két modell között, hogy az EPS felbontása nagyobb, míg a determinisztikus modell horizontális felbontása 25km, addig az EPSé 50km, ezáltal is csökkentve a több tagból adódó magasabb számítás igényen. Az ensemble előrejelzések esetében először az 51 ensemble tagból kiszámítottam az ensemble átlagot az összes rácspontokra, és ezekre az átlagokra végeztem el a verifikációt. A determinisztikus előrejelzéseknél alkalmazott módszerhez hasonlóan számítottam ki és térképesen ábrázoltam verifikációs mérőszámokat.

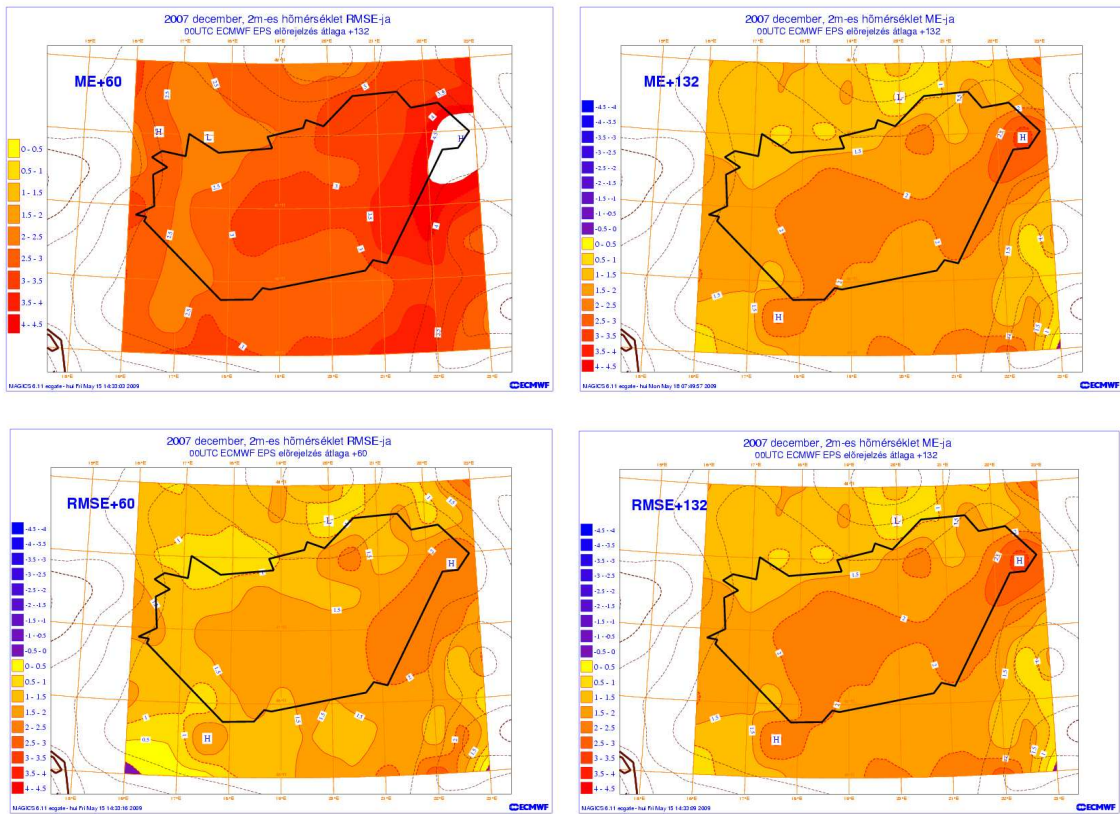
Elsőre feltűnik, hogy bizonyos hegyek (pl. Bükk) szisztematikusan feltűnnek az egyes térképeken. Ennek a magyarázata, abban lehet, hogy a verifikációhoz itt is a determinisztikus modell analízist használtam fel. Az EPS modellnél is megfigyelhető, hogy nő a modell hibája az időtáv előrehaladtával, de kisebb mértékben mint a determinisztikus előrejelzésnél.

Ugyanis az RMSE hiba +60 órás időtávnál itt is 2 körüli értéket vesz fel egész évre nézve, viszont a +132 órás időtávnál kisebb a hiba mértéke, 3-3,5 körül alakul. A szisztematikus hiba a 2007 éves átlagra nézve, mind +60 óránál, mind +132 óránál 0 körüli értékeket látunk (9. ábra).

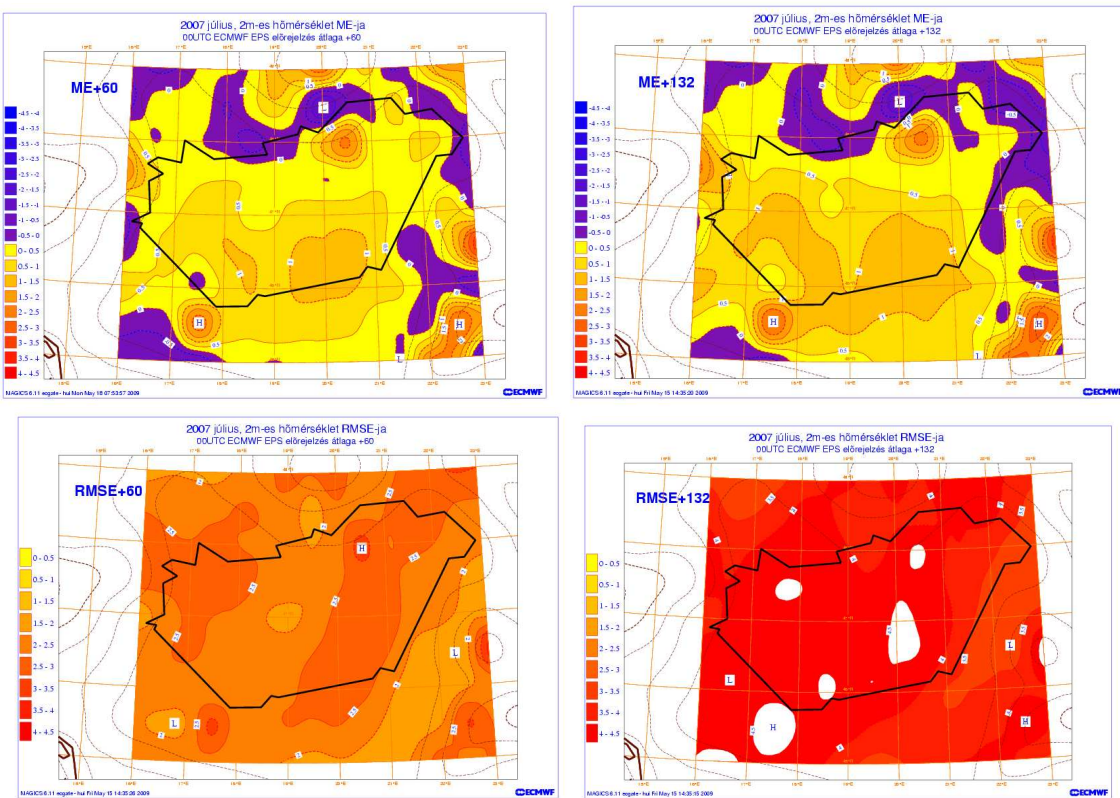


9. ábra: A 2007-es évre vonatkozó szisztematikus és RMSE hiba, 2m-es hőmérséklet, +60 és +132 órás időtávra, EPS modell alapján.

Decembernél, a determinisztikus modellhez hasonlóan, megfigyelhetjük, hogy nagyobb hibájú az előrejelzés, mint a többi hónapnál, az RMSE hiba 4 feletti értékeket is felvesz mindkét időtávnál. Ebben a hónapban a +132 órás időlépcső nem rosszabb a +60 órásnál, ami viszont már eleve nagy hibát mutat. A szisztematikus hiba is magas, 2 körüli értéket vesz fel mindkét esetben (10. ábra). Júliusban szintén magasabb hiba értékeket figyelhetünk meg, ezenfelül a szisztematikus hiba térképeken nagyon jól megfigyelhető a domborzat és a hibastatisztika együttes változása (11. ábra).



10. ábra: A 2007 decemberi szisztematikus és RMSE hiba, 2m-es hőmérséklet, +60 és +132 órás időtávra, EPS modell alapján.



11. ábra: A 2007 júliusi szisztematikus és RMSE hiba, 2m-es hőmérséklet, +60 és +132 órás időtávra, EPS modell alapján

5. Összefoglalás

A szakdolgozatot az Országos Meteorológiai Szolgálatnál készítettem. A bevezetésben rövid leírást adtam az Európai Középtávú Központtól, valamint az általuk használt determinisztikus és valószínűségi modell főbb jellemzőit ismerttettem.

A bevezetés után a különböző verifikációs mérőszámokat ismerttettem, mind az elterjedtebbeket, mind a ritkábban használtakat. Ennek a résznek az elején próbáltam rávilágítani a verifikációs munka fontosságára, felhasználási területeire. Külön alponban tárgyaltam a determinisztikus, és a több tagból álló valószínűségi előrejelzéseket, s próbáltam rávilágítani a kategorikus és valószínűségi előrejelzések verifikálása közti kapcsolatra.

Ezután a legfontosabb hibastatisztikák, mint a szisztematikus és az RMSE hiba, alkalmazására mutattam be példát. Az OVISYS rendszer rövid bemutatása után példán szemléltettem a használatát.

Végül a 2007-es évre az ECMWF determinisztikus és ensemble 2m-es hőmérséklet előrejelzésre vonatkozó objektív verifikációt végeztem el. A verifikáció során +60 és +132 órás időtávra számoltam szisztematikus és RMSE hibát, mind az egész évre vonatkozóan, mind külön-külön havonként. Ezzel a két egyszerű hibastatisztikával számszerűen összehasonlíthatóvá vált a két modell előrejelzése. Azonban ahhoz, hogy még részletesebb információt kaphassunk a modellek beválásáról érdemes lenne további mérőszámokat számolni, és további paraméterekre is elvégezni azokat.

A verifikáció alapján megállapítható, hogy a két modell hibastatisztikájának éves menete hasonlít egymáshoz, júliusban és decemberben mindkét modell esetén rosszabb a beválása az előrejelzéseknek.

Irodalomjegyzék

- Csima, G., 2003: A numerikus előrejelzés során alkalmazott utófeldolgozási és verifikációs eljárások. 29. *Meteorológiai Tudományos Napok, 2003. november 20-21.* 86-96.
- Horányi A., 2003: Numerikus előrejelzés: Nemzetközi áttekintés. 29. *Meteorológiai Tudományos Napok*, 18-24.
- Ihász I., 2008: Közép és hosszútávú előrejelzési modelltermékek és alkalmazási lehetőségeik. 34. *Meteorológiai Tudományos Napok, Budapest, 2008. november 20-21.*
http://nmo.met.hu/nmo_articles/nmo_db.php
- Jolliffe, I.J. and B., Stephenson, 2003: Forecast verification. A practitioner's guide in atmospheric science. *John Wiley and Sons*
- Kaba M., 1995: Csatlakozásunk a Középtávú Időjárás Előrejelzések Európai Központjához (ECMWF), *Légkör*, 40. 2. szám 34
- Mile, M., 2008: Az ECMWF ensemble előrejelzések utólagos kalibrációja. *ELTE TTK Diplomamunka*
- Nurmi, P., 2003: Recommendations on the verification of local weather forecasts. *ECMWF Technical Memoranda 430.*
- Richardson, D.S., J. Bidlot, L. Ferranti, A. Ghelli, M. Janousek, M. Leutbecher, F. Prates, F. Vitart and E. Zsótér, 2008: Verification statistics and evaluations of ECMWF forecasts in 2007-2008. *ECMWF Technical Memoranda 578*
- Szépszó, G., 2004, Valószínűségi előrejelzések verifikációja. *OMSZ kézirat*
- Üveges, Z., 2009: Az ECMWF havi előrejelzések kalibrációja *ELTE Diplomamunka*
- WRCP, 2008: Forecast Verification - Issues, Methods and FAQ. *World Climate Research Programme*
- Woods, A. 2006: Medium-Range Weather Prediction The European Approach. *Springer 270 pp.*

Köszönetnyilvánítás

Szeretnék köszönetet mondani a témavezetőmnek, Ihász Istvánnak az ismereteimben és a munkámban nyújtott rengeteg segítségért. Továbbá Bartholy Juditnak egyetemi konzulensemnek.

Függelék

A verifikáció elvégzéséhez minimálisan ajánlot, és az optimálishoz kibővítet, mérőszámok Nurmi alapján.

Jelölés	Angol név	Minimális	Optimális
BS	Brier Score	X	X
CDF	Cumulative Probability Space		
CR	Correct Rejection		
CSI	Critical Success Index		
ETS	Equitable Threat Score	X	X
F	False Alarm Rate	X	X
FAR	False Alarm Ratio	X	X
FBI	Frequency Bias Index		
H	Hit Rate		
HSS	Heide Skill Score	X	X
KSS	Hanssen-Kuipers Skill Score	X	X
LEPS	True Skill Score		X
MAE	Mean Absolute Error	X	X
ME	Mean Error	X	X
MR	Miss Rate		
MSE	Mean Squared Error		X
OR	Odds Ratio		X
ORSS	Odds Rate Skill Score		X
PC	Proportion Correct	X	X
POD	Probability Of Detection	X	X
POFD	Probability Of False Detection u.a. mint FAR		
PSS	Peirce Skill Score		
RMSE	Root Mean Squared Error		
ROC	Relative Operating Characteristic		X
RPS	Ranked Probability Score		X
RPSS	Ranked Probability Skill Score		X
SS	Skill Score		
TS	Threat Score	X	X
TSS	True Skill Score		

