

Statisztikai alapú valószínűségi előrejelzés készítése a szélerőművek által termelt energia becslésére

SZAKDOLGOZAT
FÖLDTUDOMÁNYI ALAPSZAK
METEOROLÓGUS SZAKIRÁNY



Készítette:
Lencse Róbert
(Földtudományi alapszak, Meteorológus szakirány)

Témavezető:
Brajnovits Brigitta
(Országos Meteorológiai Szolgálat)

Tanszéki konzulens:
Dr. Matyasovszky István
(ELTE TTK, Meteorológiai Tanszék)

Eötvös Loránd Tudományegyetem, Természettudományi Kar
Földrajz- és Földtudományi Intézet
Meteorológiai Tanszék

Budapest, 2015

Tartalomjegyzék

I. Bevezetés.....	3
II. Megújuló energiaforrások.....	5
III. Szélenergia szerepe.....	6
III. 1. A szélenergia szerepe a világ energiaellátásában.....	6
III. 2. A szélenergia szerepe Magyarország energiaellátásában.....	9
III. 3. Szélenergia becslése.....	11
IV. Valószínűségi előrejelzések.....	13
IV. 1. Hibaforrások.....	13
IV. 2. Bizonytalanságok számszerűsítésének módszerei.....	15
V. Analog ensemble.....	18
VI. ALADIN.....	19
VII. Verifikációs módszerek.....	21
VIII. Eredmények.....	24
VIII. 1. Felhasznált adatok.....	24
VIII. 2. A vizsgált kísérletek.....	24
VIII. 3. Eredmények kiértékelése.....	26
IX. Összefoglalás.....	36
X. Köszönetnyilvánítás.....	37
XI. Irodalomjegyzék.....	38
XI. 1. Internetes hivatkozások.....	40

I. Bevezetés

Napjaink egyre jelentősebb problémája a légköri széndioxid koncentráció növekedésének megállítása. A II. világháború után drasztikusan megugró növekedés az emberiség egyre nagyobb energiaigényének volt köszönhető, amelyet túlnyomórészt fosszilis tüzelőanyagok elégetéséből fedeztek. Ebből kifolyólag az egyik legfontosabb kérdésfelvetés az, hogy milyen módon csökkenthető a fosszilis tüzelőanyagok elégetéséből származó CO₂ kibocsátás.

Több oldalról közelíthetjük meg ezt a problémát. Lehetséges megoldásként tekinthetünk a megújuló energiaforrásokban rejlő lehetőségek kiaknázására. Ha csökkentjük a hagyományos erőművek által termelt energiát, akkor a keletkező energiahiányt pótolhatjuk olyan megújuló energiaforrások által, amelyek kevésbé, vagy egyáltalán nem termelnek CO₂-ot. Ez azonban azt a problémakört veti fel, amely a megújuló energiaforrások energiahálózatba való beintegrálásával jár. A vízenergia, a szélenergia és a napenergia a legfontosabb megújuló energiaforrások. A jelen dolgozatban a szélenergia jelentőségével, valamint teljesítménybecslésének egyik lehetséges módjával kívánok részletesen foglalkozni.

A szélenergia integrálhatósága szempontjából nagyon fontos előre ismerni az általuk leadott teljesítményt, nagy időbeli felbontásban (15 percenként) mivel a hálózatra kapcsolandó energiamennyiséget csak így lehet előre megtervezni. Fontos hogy a tervezéshez a lehető legpontosabb becslés álljon rendelkezésre, mert az előrejelzettnél kevesebb vagy több beérkező energia az áramszolgáltatásban fennakadásokat okozhat. Éppen ezért a szélenergia-üzemeltetőknek menetrendet kell készíteniük a MAVIR Zrt. (Magyar Villamosenergia-ipari Átviteli Rendszerirányító Zártkörűen Működő Részvénytársaság) számára, arról hogy várhatóan mekkora teljesítményt adnak le. Ezt jelentősen segítheti az, ha adott időpontban minél pontosabb a szélsébség előrejelzése a rotorok magasságában, mivel ebből számítható ki a szélenergia-üzemeltetők által leadott teljesítmény. Ahhoz hogy egy-egy időintervallumnak az átlagos szélsébség értékét a legjobb megközelítéssel adják, szükséges a szél minél pontosabb és minél gyakoribb mérése, mert ez az egyik legváltozékonyabb és legnehezebben előrejelezhető meteorológiai paraméter. Jelenleg 15 percenként küldi az OMSZ (Országos Meteorológiai Szolgálat) az adott időpillanatra előrejelzett rotorszintű szélsébségekből számított teljesítmény előrejelzést.

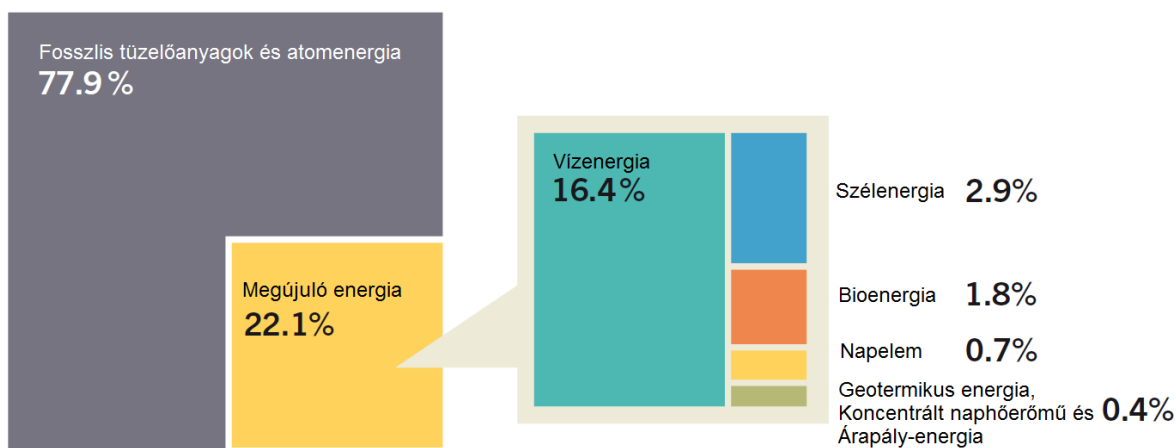
Dolgozatom célja az eddigi determinisztikus alapú előrejelzések valószínűségi alapra helyezése. Erre a célra kiválasztott módszer az *analog ensemble* (AnEns) (Delle Monache et al., 2013), aminek segítségével egy olyan modell állítható fel, amelyben lehetőség nyílik statisztikai összefüggések keresésére a múltban előrejelzett és mért szélsőségek között (Alessandrini et al., 2014). Eredményként az várható hogy ez által jobb előrejelzés adható a vizsgált szélerőmű üzemeltetőinek.

A dolgozatom első részében bemutatom a megújuló energiaforrások szerepét a világban, azon belül is kiemelten a szélerőműveket, majd bemutatásra kerülnek a valószínűségi előrejelzések és az ALADIN (*Aire Limitée Adaptation Dynamique Développement International* – korlátos tartományú modellezés dinamikai adaptációval nemzetközi együttműködésben) modell, amelyből származnak az előrejelzett értékek.

II. Megújuló energiaforrások

Az elmúlt évszázadok során jelentősen növekedett az emberiség energiafelhasználása. Ahhoz, hogy ez a növekedés továbbra is fenntartható legyen, különböző energiaforrásokra van szükségünk, mivel a primer energiaforrások (fosszilis tüzelőanyagok) felhasználása jelentős környezetterhelést eredményez. Ennek mérséklése úgy érhető el, ha csökkentjük az egy főre jutó energiafelhasználás mértékét, illetve növeljük a megújuló energiaforrások felhasználásának arányát. Megújuló energiaforrásnak tekintjük a természeti folyamatok során folyamatosan rendelkezésre álló, vagy újratermelődő energiaforrásokat (nap-, szél-, víz-, geotermikus és a biomasszából nyert energia) (Bartholy et al., 2013a).

A REN21 (2014) jelentése szerint 2013-ban a teljes villamosenergia-termelés több mint 22%-át megújuló energiaforrásokból állították elő (1. ábra). Ezen megújuló energiaforrások közül a legtöbb beépített villamosenergia kapacitással rendelkező erőművek a víz (1000 GW), a szél (318 GW) és a Nap 139 (GW) energiáját hasznosítják. A vízerőművek mellett a nap- illetve a szélerőművek telepítésére történnek a legnagyobb befektetések a világon. 2013-ban a szélerőművek esetében 80 milliárd dollár értékű beruházás valósult meg, amiből látszik, hogy napjaink egyik legfontosabb megújuló energiaforrása, amely 2014 közepére már elérte a világ villamosenergia-termelésének 4%-át (WWEA, 2014).



1. ábra: a globális villamosenergia-termelés becslés megoszlása energiaforrásonként (REN21 nyomán, 2014).

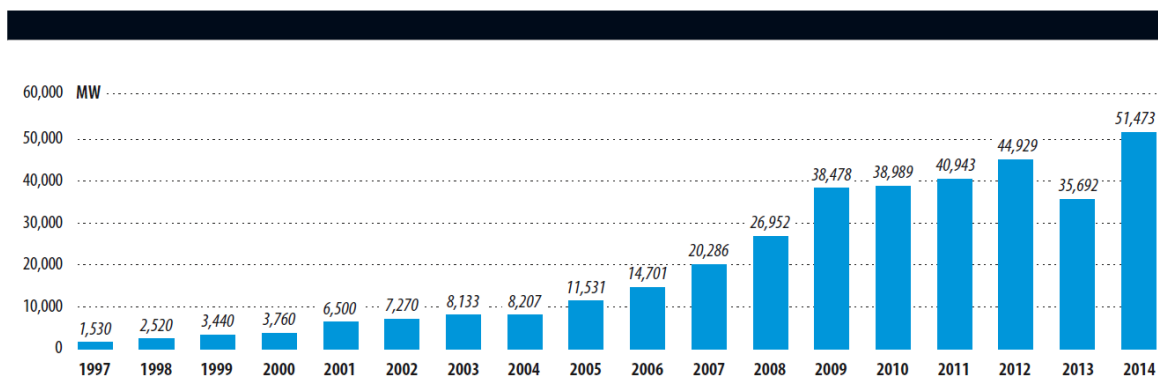
III. A szélenergia szerepe

III. 1. A szélenergia szerepe a világ energiaellátásában

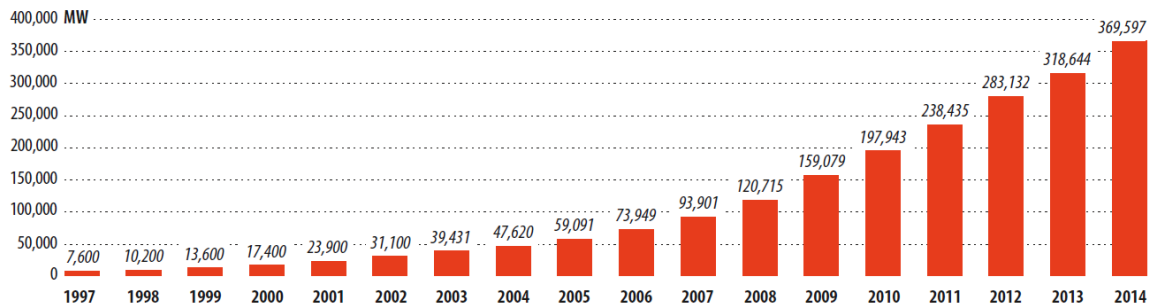
A szélenergia-ipar áttörése az 1990-es évekre tehető. Számos tényező együttes hatásának volt ez köszönhető: többek között a gyártástechnológia fejlődésének, a turbinák teljesítményének növekedése, valamint a megújuló energiaforrások hasznosítását támogató pénzügyi források bővülése hozta meg az előrelépést (Hartmann, 2012). 2015 elején a *Global Wind Energy Council* már publikálta éves jelentését (GWEC, 2015). Főként ennek adatai alapján kívánom részletesen bemutatni a szélenergia helyzetét napjaink villamosenergia-ellátásában.

A szélenergia-ipar számára 2014 egy kiváló év volt, hiszen több mint 51 GW-nyi kapacitást telepítettek az adott évben, ami rekordnak számít. Ez a szám az előző évhez képest 44%-kal több, amikor csak 35,6 GW-ot telepítettek. Továbbá 2014-re a világon összesen 369,6 GW-nyi névleges kapacitású szélenergia-üzemelt, amely 16%-os növekedést jelent az előző évi adathoz képest, ami alacsonyabb, mint az átlagos 23%-os növekedési ráta az elmúlt 10 évben.

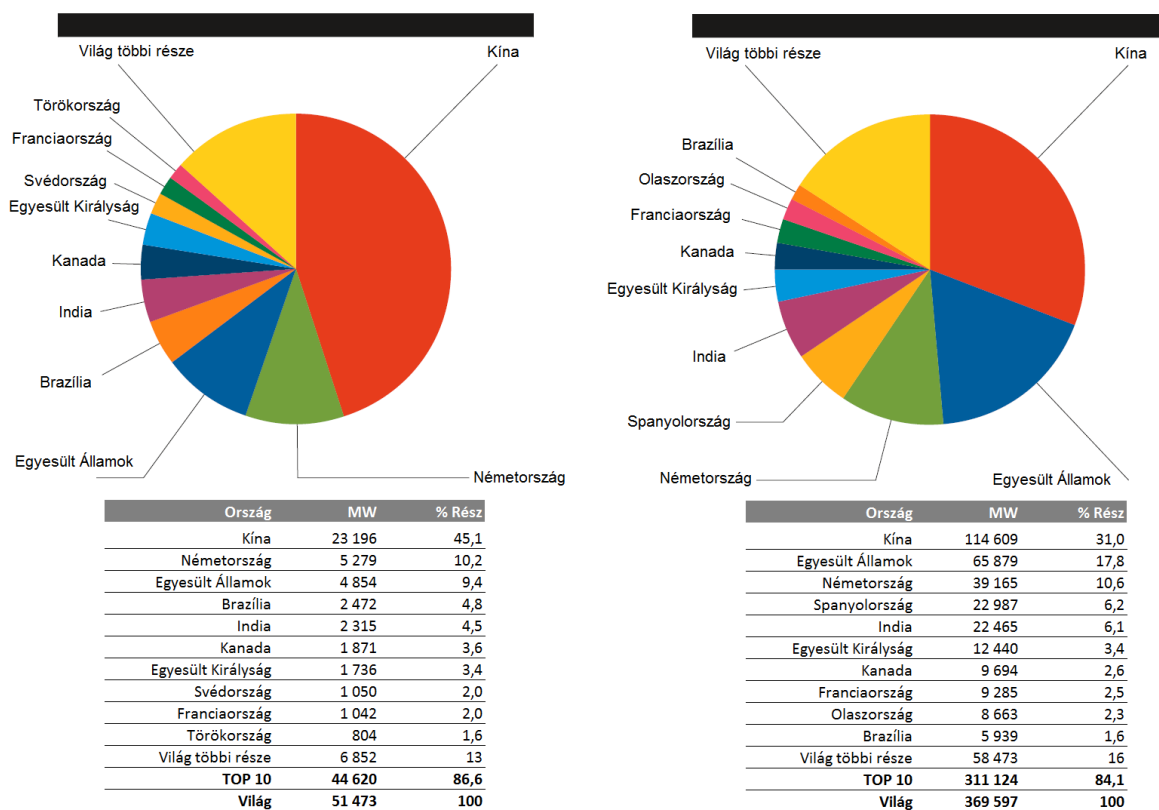
Az adott évben üzembe helyezett, illetve az adott év végén üzemben álló szélenergia-üzemek kapacitásai a 2. és 3. ábrán láthatóak 1997-ig visszamenőleg.



2. ábra: adott évben a világon üzembe helyezett szélenergia-üzemek kapacitásai (GWEC nyomán, 2015).



3. ábra: adott évben a világon üzemelő szél erőművek összkapacitásai (GWEC nyomán, 2015).



4. ábra: Balra: a 2014-es évben üzembe helyezett, valamint (jobbra) az összesített szél erőművi kapacitás (GWEC nyomán, 2015).

Továbbá 2014-re 11 ország rendelkezik több mint 5000 MW-nyi beépített kapacitással és 24 ország rendelkezik legalább 1000 MW-nyi kapacitással, amiből 16 ország Európában található.

A 4. ábra ismerteti a 2014-es évben üzembe helyezett szél erőművi kapacitások országok szerinti eloszlását, illetve az év végéig beépítésre került, összesített kapacitásokat.

A világpiacon tekintve Kína jelenti a fő hajtóerőt, ahol 2014-ben rekord méretű 23,1 GW-nyi újonnan telepített kapacitás valósult meg, ami a valaha volt legtöbb éves befektetés a világon. Ráadásul Kína már átlépte a 100 GW-os határt, ami a világ 31%-át jelenti. Ázsia második legnagyobb piaca India 22 GW-tal, ahol több mint 2 GW-nyit telepítettek újonnan. A sorban Japán (2789 MW), Tajvan (633 MW) és Dél-Korea (609 MW) következik.

Észak-Amerikát tekintve, újra az Egyesült Államok telepítette a legtöbbet 2014-ben, pontosan 4854 MW-ot, ami több a 2013-as adathoz képest (1084 MW), de még mindig messze elmarad a 2012-es szinttől és az az előtti évektől is. Ennek ellenére az Egyesült Államok továbbra is a világ második legnagyobb szélenergiás kapacitásával bír. Az Egyesült Államok után Kanada következik, ahol a 2014-es évben rekord nagyságú (1871 MW) szélenergiás kapacitást telepítettek, amivel a világ hatodik legnagyobb (9694 MW) szélenergiás ipari piacával rendelkezik.

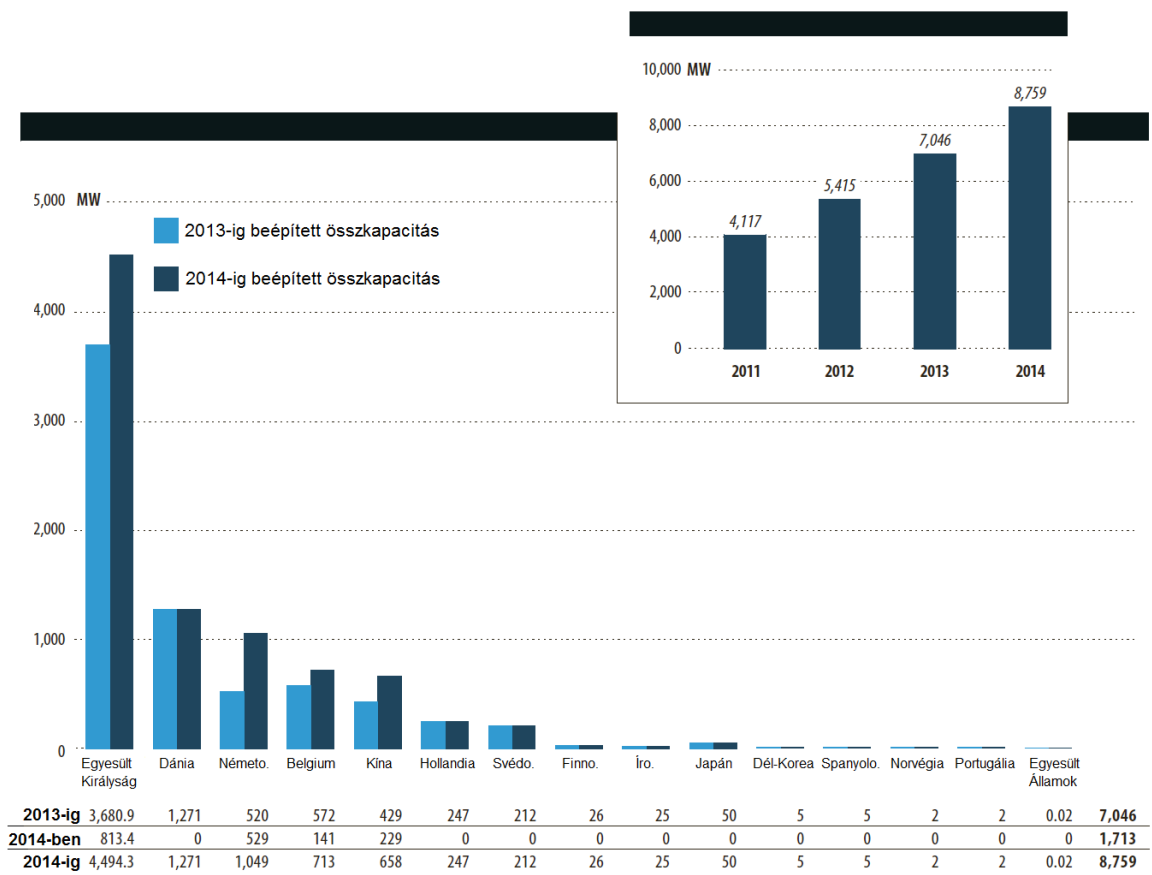
Közép- és Dél-Amerika szintén jó évet zárt, ahol Brazíliában található a legtöbb beépített kapacitás (5939 MW), ami leginkább az elmúlt 3 évnek köszönhető. Brazíliát Mexikó (2551 MW), Chile (836) és Uruguay (464 MW) követi.

Óceániát tekintve Ausztráliában található a legtöbb, összesen 3,8 GW, míg Új-Zélandon a 2014-es évben nem építettek újabb szélenergiás kapacitásokat, így összességében 623 MW-nyi kapacitás található ott.

Afrikában is a 2014-es év rendkívüli év volt, hiszen Dél-afrikai Köztársaság a meglévő 10 MW-hoz további 560 MW-nyi szélenergiás kapacitást telepített, továbbá Marokkó Afrika legnagyobb, 300 MW-os szélenergiás kapacitását hozta létre, így összesen 787 MW-al rendelkezik, amivel Afrika legnagyobb (787 MW) szélenergiás ipari piacával rendelkezik. Egyiptomban csekély mértékű, összesen 60 MW-ot telepítettek 2014-ben, amivel lecsúszott a második helyre (610 MW). A sorban Tunézia (245 MW) és Etiópia (171 MW) következik.

Amennyiben Európát nézzük, a 2014-es évben 12.858 MW-nyi szélenergiás kapacitás üzembe helyezése valósult meg, ebből 11 829 MW az EU 28 tagállamának valamelyikében. Az összeurópai beépített kapacitás jelenleg 134 GW, ebből 128,7 GW található az EU országaiban. A vezető szerepet továbbra is Németország (39,1 GW), Spanyolország (22,9 GW) és az Egyesült Királyság (12,4 GW) tölti be. Az előbbieket követi Franciaország (9285 MW), Olaszország (8663 MW) és Svédország (5425 MW).

A közép- és kelet-európai régiót tekintve, Lengyelországban található a legtöbb beépített kapacitás (3834 MW), ezt követi Törökország (3763 MW), Románia (2952 MW), Ausztria (2095 MW) és Görögország (1980 MW).



5. ábra: adott évben a világon üzemelő és üzembe helyezett offshore szélenergia kapacitások (GWEC, nyomán 2015).

Az *offshore* (tengerre telepített szélenergia) beruházások 2013-hoz képest nagymértékben elmaradtak, hiszen 2014-ben összesen 1713 MW-nyi kapacitást telepítettek. Ezen beépített kapacitás jelentős része továbbra is az Egyesült Királyság partjainál található.

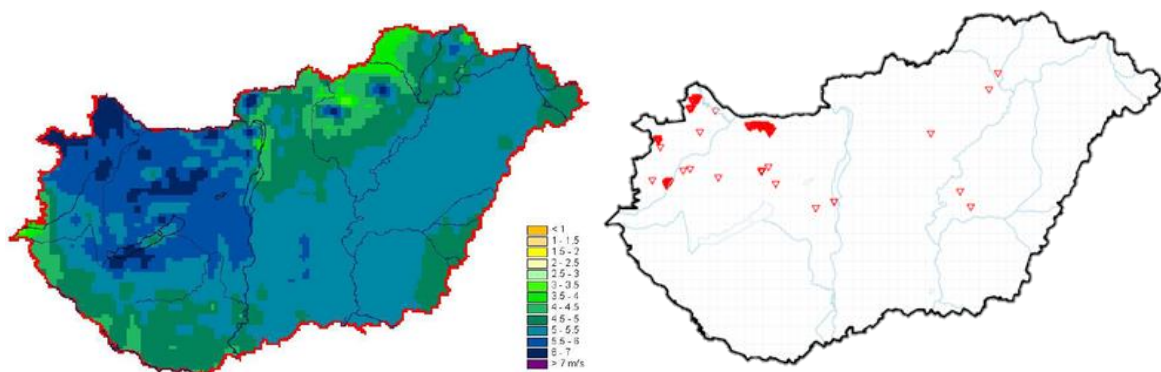
III. 2. A szélenergia szerepe Magyarország energiaellátásában

Az OMSZ által 2006-ban publikált kutatásai szerint hazánk igen jelentős szélenergia potenciállal rendelkezik (Hunyár et al., 2006). Ennek ellenére Jánosi (2012) szerint Magyarország nem számít kedvező területnek szélenergia potenciál szempontjából, mivel a szárazföld felett a légáramlatok erősen lefékeződnek, illetve a Kárpát-medencét körülvevő hegyláncok árnyékoló hatása negatívan befolyásolja a szél sebességét. Mindazonáltal hazánk szélenergia potenciáljára vonatkozó becsléseket az OMSZ végezte el. A kapott eredmények alapján elmondható, hogy a legnagyobb átlagos szélességek a felszíntől mért 10-100 méteres magasságban a Dunántúl északi területein fordulnak elő. Az átlagos értékek

helyenként itt meghaladják az 5 m/s-ot (egyes források szerint e fölött érdemes az adott helyen szélérőmű-beruházást kezdeményezni, de más forrás szerint 7 m/s körüli átlagértékekkel rendelkező területen valósítható meg gazdaságos szélenergia hasznosítás).

Ennek tükrében nem meglepő, hogy a legtöbb szélérőmű az ország északnyugati területén létesült. Az első szélérőművet Inotán állították fel 2000-ben, ami gyakorlatilag csak demonstrációs célokat szolgált az alacsony névleges teljesítménye (250 kW) miatt, valamint rossz elhelyezkedésének köszönhetően (Hartmann, 2012). Az első sikeres vállalkozás Kulcs településén történt 2001-ben. Innentől kezdve egészen 2011-ig minden évben telepítettek szélérőműveket az ország különböző pontjaiban, és az ország szélérőművi kapacitása elérte a 330 MW-ot, amit 172 db széltoronnyal valósítottak meg.

Magyarországon a villamosenergia-rendszer stabilitásának biztosítása miatt jelenleg kapacitáskorlát vonatkozik a szélérőművekre, azaz csak korlátozott összkapacitású szélérőmű csatlakozhat a hálózatra. Ennek oka több tanulmány szerint elsősorban a szabályozási, és kevésbé a műszaki okokra vezethető vissza (Hartmann, 2012).



6. ábra: Balra: átlagos szélesség 75 méteres magasságban. Jobbra: szélérőművek elhelyezkedése (Hunyár et al., 2006).

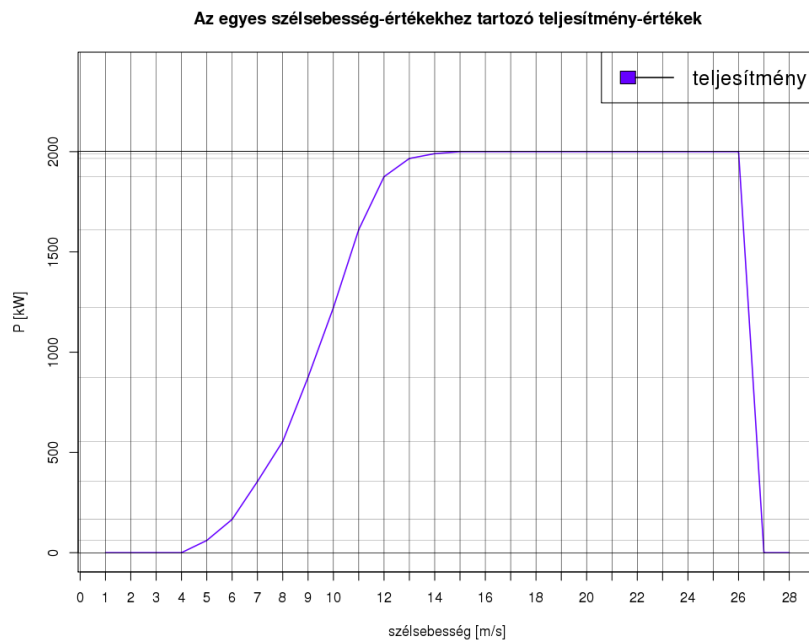
1. táblázat: magyarországi szélérőművek listája (Hartmann, 2012).

Helyszín	Tornyok száma	Egységteljesítmény [kW]	Összteljesítmény [kW]	Típus	Üzembe helyezés
Inota	1	250	250	Nordex N-250	2000
Kulcs	1	600	600	Enercon E-40	2001
Mosonszolnok	2	600	1200	Enercon E-40	2002
Mosonmagyaróvár	2	600	1200	Enercon E-40	2003
Bükkaranyos	1	225	225	Vestas V-27	2004
Erk	1	800	800	Enercon E-48	2005
Újrónafő	1	800	800	Enercon E-48	2005
Szápár	1	1 800	1 800	Vestas V-90	2005
Vép	1	600	600	Enercon E-40	2005
Mosonmagyaróvár	5	2 000	10 000	Enercon E-70	2005
Mezőtúr	1	1 500	1 500	Fuhrlander MD-77	2006
Törökszentmiklós	1	1 500	1 500	Fuhrlander MD-77	2006
Mosonmagyaróvár	5	2 000	10 000	Vestas V-90	2006
Felsőlolca	1	1 800	1 800	Vestas V-90	2006
Csetény	2	2 000	4 000	Vestas V-90	2006
Ostffyasszonyfa	1	600	600	Enercon E-40	2006
Levél	12	2 000	24 000	Gamesa G-90	2006
Mosonszolnok	1	800	800	Enercon E-48	2007
Csorna	1	800	800	Enercon E-48	2007
Mecsér	1	800	800	Enercon E-48	2007
Bakonycsernye	1	2 000	2 000	Vestas V-90	2007
Sopronkövesd	4	3 000	12 000	Vestas V-90	2008
Nagylőzs	3+1	3 · 3 000+1 · 2 000	11 000	Vestas V-90	2008
Levél	12	2 000	24 000	Repower MM-82	2008
Jánossomorja	5	4 · 2 000+1 · 1 800	9 800	Vestas V-90	2008
Ács	1	2 000	2 000	Vestas V-90	2008
Pápakovácsi	1	2 000	2 000	Vestas V-90	2008
Vönöck	1	850	850	Vestas V-52	2008
Kisigmánd	19	2 000	38 000	Gamesa G-90	2009
Nagyigmánd	6	2 000	12 000	Gamesa G-90	2009
Böny	8+4+1	8 · 2 000+4 · 1 800+1 · 1 800	25 000	Vestas V-90	2009
Csém	6	2 000	12 000	Gamesa G-90	2010
Nagyigmánd	7	2 000	14 000	Gamesa G-90	2010
Ács	6	2 000	12 000	Gamesa G-90	2010
Lövő	1	2 000	2 000	Vestas V-90	2010
Ács	16	2 000	32 000	Gamesa G-90	2010
Nagyigmánd	2	2 000	4 000	Gamesa G-90	2010
Bábolna	6+1	6 · 2 000+1 · 3 000	15 000	Vestas V-90	2010
Jánossomorja	1	2 000	2 000	Vestas V-90	2010
Ikervár	4+13	4 · 2 000+13 · 2 000	34 000	Gamesa G-90	2011
Összesen	172		329 325		

III. 3. Szélenergia becslése

Egy szélérőmű energiatermelésének becslésére az egyes szélérőmű típusokra rendelkezésre álló teljesítménygörbét használják, amely a leadott teljesítmény szélsősebességtől való függését mutatja. Általában a generátorok meghajtásához szükséges minimális szélsősebesség jellemzően 2,5-5 m/s, (*cut-in* érték) azaz ekkor indul meg a termelés. Ezt az értéket bekapcsolási vagy indulási sebességnek is hívják. Efelett a kimenő teljesítmény a szélsősebesség közel harmadik hatványával növekszik mindaddig, amíg eléri a névleges szélsősebességet, ami 11-13 m/s körüli (egyész szélturbináknál ez az érték akár 14-17

m/s is lehet), ekkor az aktív turbina-kontroll bekapcsolásra kerül. Ez után a kimenő teljesítmény egy konstans érték (névleges teljesítmény plató) körül mozog, ami a maximális értéket jelenti. Ekkor a leadott teljesítményt a vezérlőrendszer határozza meg a szélérőmű védelmének érdekében. A tovább növekvő széleseesség esetén is a rendszer megközelítőleg a névleges teljesítményen marad. Túlságosan erős szélben viszont a tornyot és a lapátokat erős mechanikai terhelés érheti, ezért biztonsági okok miatt a leállási sebesség elérésekor, azaz 25 m/s körül (*cut-out* érték) a turbinát leállítják. A szélérőmű a *cut-in* és *cut-out* értékek közelében valójában összességében több energiát emészt fel, mint amennyit termel, de ezzel a becslés során nem kell foglalkoznunk.



7. ábra: egy átlagos szélérőmű teljesítménygörbéje.

IV. Valószínűségi előrejelzések

A minél pontosabb és minél megbízhatóbb előrejelzések készítése érdekében a modellezés számos területén folynak kutatások. Ide tartozik a modell dinamikai magja, a tér- és időbeli diszkretizáció, az adatasszimiláció és a fizikai parametrizáció [1 – OMSZ]. Ahol a modell dinamikai mag: a légkör folyamataira felírt matematika egyenletrendszerek közelítő megoldását előállító programrészt általában a modell „dinamikai magjának” vagy egyszerűen modell-dinamikának nevezik (Bartholy, 2013b). Az adatasszimiláció: a modell kezdeti feltételeinek előállítását jelenti a légkör aktuális állapotát jellemző rendelkezésre álló információk figyelembevételével. A fizikai parametrizáció: a rácsfelbontásnál kisebb skálájú folyamatoknak a nagyobb skálájú folyamatokra gyakorolt összegzett hatásának leírását értjük.

Közismert tény, hogy az időjárás-előrejelzések nem tökéletesek, hiszen az előrejelzések készítése során korlátokba ütközünk, mivel nem lehet tetszőleges időtávra, tetszőleges részletességgel és pontossággal számolni. Továbbá különböző hibák jelentkeznek a mérések illetve modell futtatások során (Szűcs & Horányi, 2013). A modell futtatások során létrejövő hibáknak több fajtája is van.

IV. 1. Hibaforrások

A hibaforrások egyike a külső hiba, ami abból ered, hogy a gyakorlatban a természet összetett folyamatait csak közelítésekkel tudjuk leírni, azaz a modellek nem képesek teljes részletességgel kifejtetni, így bizonytalanságokat hordoznak, amelyek hatása megjelenik az előrejelzésekben. Tehát a számszerű előrejelző modellek eredményei mindig hibával terhelték. Ezen modellezési hibákból származó bizonytalanságok a következőkből adódhatnak:

- adatasszimilációból, amely a modell kezdeti feltételeinek előállításából adódik;
- határfeltételek megválasztásából;
- diszkretizációból, amely a modell által leírt légköri folyamatok rácson való megjelenítéséből adódik;
- fizikai parametrizációból, amely a kis skálájú folyamatok kezeléséből származhat.

Egy másik hibaforrás a belső hiba, aminek megértéséhez nézzük egy egyszerű, nemlineáris rendszer kaotikus tulajdonságait. Edward Norton Lorenz amerikai meteorológus

alacsony dimenziószámú légköri modellek vizsgálata során érzékelte, hogy két hasonló kezdeti feltételből indított modellintegrálás eredményei az idő előrehaladtával egyre nagyobb mértékben eltérnek egymástól mindaddig, amíg a kezdeti különbségekből eredő bizonytalanságok el nem uralják teljesen az előrejelzést, így pedig korlátot szabva az előrejelezhetőségnek (Götz, 2001). Az időjárás modellek természetesen nem egyszerű rendszerek, hanem bonyolult, nemlineáris parciális differenciálegyenletekből állnak, melyekben számos változó szerepel. Ezek a rendszerek magukon viselik az egyszerű nemlineáris rendszerek kaotikus tulajdonságait, azaz a kezdeti feltételekre való érzékenységet. Ez az érzékenység és az ebből fakadó hiba tehát a légköri előrejelzés problémájának természetéből fakad. A kaotikus rendszerek legfontosabb tulajdonságai:

- kezdeti feltételekre való érzékenység (kis kiindulási hiba, később nagy eltérésekhez vezethet az előrejelzésben);
- korlátozott előrejelezhetőség;
- aperiodikusság (a rendszer sohasem veszi fel kétszer ugyanazt az állapotot).

A fentiek alapján felmerülhet a kérdés, hogy miként tudjuk kezelni az időjárás korlátozott előrejelezhetőségének problémáját. A problémára megoldást jelenthet, ha az egyes időjárás eseményekhez kategorikus igen-nem válaszok helyett valószínűségi értékeket rendelnek. Ezt a múltban már felismerték, hogy egyetlen meteorológiai előrejelzés sem lehet teljes értékű, amely csak magát az előrejelzés értékét adja meg. Éppen ezért készítenek valószínűségi előrejelzéseket, amelyek számszerűsítik az előrejelzésekben rejlő bizonytalanságokat (Götz, 2001). Ennek segítségével az egy-egy időjárás paraméterre történő előrejelzés nem csak egy szám lesz, hanem egy eloszlásfüggvény. E módszer során nem csak egyetlen előrejelzést futtatnak, hanem többet eltérő modellintegrálással, azaz a fizikai folyamatok leírásában térhetnek el egymástól vagy valamilyen perturbációt ültethetnek a kezdeti feltételekre. Emellett egy korlátos tartományú modellben szükség van globális modellekre az oldalsó peremfeltételek meghatározásához, így ezen keresztül is érkezhettek perturbációk egy globális *ensemble* (együttes) rendszerből. Tehát előrejelzések sokaságát kapjuk, ahol minden *ensemble* tag beválási valószínűségét egyenlőnek tekintjük. Egyébként maguk az *ensemble* előrejelzés tagjai is determinisztikusak, hiszen adott kezdeti feltételhez egyetlen végállapot kapható, csak úgy, mint egyetlen nagy felbontású előrejelzésnél. Előrejelzések sokaságát futtatni költséges, ezért az *ensemble* tagok tér- és időbeli felbontása rosszabb, mint az ún. determinisztikus előrejelzéseké. A gyakorlatban

általában tíz és ötven között mozog az *ensemble* előrejelzések tagjainak száma (Szűcs & Horányi, 2013). Habár egy *ensemble* rendszer minél több tagból áll, annál jobban képes megfogni az előrejelzés bizonytalanságát, a gyakorlatban ennek határt szab az adott számítógép teljesítménye.

IV. 2. Bizonytalanságok számszerűsítésének módszerei

Az egyes *ensemble* tagok eltérése függ, a fent említett belső és modellezési bizonytalanságokból (Szűcs & Horányi, 2013). Ezek számszerűsítése a következőképpen lehetségesek.

Kezdeti feltétel perturbációk:

A perturbációk kezdeti feltételekre való ültetésével, azaz a kezdeti feltételek kis hibáiból származó bizonytalanságok becslése. Ezen perturbációk nem lehetnek akármilyen nagyok, azaz hibahatáron belül kell maradniuk. Továbbá az előrejelzés során fejlődő hibával összhangban is kell növekedniük, hogy jól mutassák meg a leírni kívánt bizonytalanságokat. A perturbációk létrehozására többféle módszer is létezik, amelyek közül a legfontosabbak a következők:

Szinguláris vektorok módszere:

Ezt az eljárást az ECMWF-nél (*European Centre for Medium-Range Weather Forecasts* – Európai Középtávú Előrejelző Központ) vezették be először az 1990-es évek elején. A módszer során olyan kezdeti feltétel perturbáció előállítására törekcsenek, mely az előrejelzés egy korai, rövid szakaszán valamilyen norma szerint a legnagyobb mértékben növekszik (Kertész, 2006), így ez az eljárás a fázistér leginkább instabili irányait azonosítja. Ezeket a perturbációkat optimális perturbációknak is szokás nevezni.

Breeding módszer:

Az amerikai NCEP-nél (*National Centers for Environmental Prediction*) fejlesztették ki az ún. breeding („tenyésztés”) módszert (Toth & Kalnay, 1997). Ennek során kezdetben az előrejelzéseket véletlenszerű perturbációkkal látják el. Ezek idővel, a belső bizonytalanságoknak köszönhetően növekedni kezdenek, ezért ezeket a megnövekedett nagyságú perturbációkat visszaskálázzák a kezdeti feltételek becsült hibájának megfelelően egy bizonyos idő után. A visszaskálázott perturbációkból újabb előrejelzést indítanak. Ezt

többször megismétlik, amíg a modell az állapottér azon irányait nem mutatja, amelyek az előző modellintegrálások során a legtöbb bizonytalanságot hordozták.

Ensemble adatasszimilizáció:

Ennek során egymással párhuzamosan több adatasszimilációt is futtatnak, azaz adatasszimilációk sokasága, *ensemble*-je készíthető el (*Ensemble of Data Assimilation – EDA*, Isaksen et al., 2010). Az egyes adatasszimilációs ciklusok különbsége abból ered, hogy a megfigyeléseket véletlenszerűen perturbálják a becsült megfigyelési hibák határán belül. Ezen perturbációk az asszimiláció során tovább fejlődnek, köszönhetően a belső bizonytalanságoknak. Ezáltal a módszer az állapottér azon irányait mutatja meg, amelyek az analízis során a legtöbb bizonytalanságot hordozzák.

Modell hibák jellemzése

A modell integrálása során alkalmazott eljárások is számos bizonytalanságot hordoznak. Ezek a bizonytalanságok is hasonló hibanövekedéshez vezethetnek az előrejelzés során, mint ahogyan a kezdeti feltételek hibájánál láttuk. Így az integrálás során a modell hibáit is számszerűsítik. Leginkább a fizikai parametizációkból fakadó bizonytalanságokat számszerűsítik.

Multi-fizika

Ennek során az egyes előrejelzésekhez különböző parametizációs csomagokat vagy azok különböző beállításait használják. Így ezek mindegyike hozzárendelhető az *ensemble* előrejelzés egy-egy tagjához, azaz reprezentálható a modellekben rejlő bizonytalanságok.

Sztochasztikus fizika

A különböző parametizációkból származó tendenciákra ültetett véletlenszerű perturbációk segítségével számszerűsíti a parametizációk bizonytalanságát. Természetesen a különböző rácsponatokban illetve időlépcsőkben alkalmazott perturbációk mérete nem lehet független egymástól (Palmer et al., 2009; Bouttier et al., 2012).

Multi-rendszerek

A bizonytalanságnak egy másik módszere a következő: sok előrejelző központ futtat numerikus modelleket, amelyek mind felépítésben, mind beállításokban eltérnek egymástól. Így ezen különböző modellek segítségével is létrehozhatóak *ensemble* előrejelzések.

Multi-modell ensemble

Ez egy egyszerű módszer, ahol több előrejelző központ, adott pontra vagy területre való független előrejelzéseinek sokasága képezi az *ensemble* tagokat, amelyek így egy *ensemble* rendszert képeznek.

Multi-analízis

Több előrejelző központ a kezdeti feltételeit használva mindegyikből adott modellel készítenek előrejelzéseket.

Multi-LBC

A regionális vagy korlátos előrejelzések futtatása során szükség van peremfeltételekre. A Multi-LBC (*lateral boundary conditions* – oldalsó peremfeltételek) módszernél több globális modell oldalsó peremfeltételével készítenek előrejelzési együttest, s így számszerűsítjük a határfeltételekben rejlő bizonytalanságokat.

A fentieket általában nem önmagukban, hanem valamilyen kombinációban szokták alkalmazni. Ilyen lehet a kezdeti feltétel perturbációs módszerek ötvözése (Buizza et al., 2010), a kezdeti feltétel és a modell hiba szemléltetésének együttes használata, vagy a korlátos tartományú rendszerekben a multi-módszerek elegye (Heizenreder et al., 2006; García-Moya et al., 2011). A gyakorlat során leginkább a kezdeti feltételekben rejlő hibák számszerűsítésén van a hangsúly.

V. Analog ensemble

Az *analog ensemble* módszere egy statisztikai alapú valószínűségi előrejelzés, ami hasonlít az *ensemble* előrejelzésekre olyan értelemben, hogy itt is több, azaz sokasági előrejelzést készítünk. Eredetileg *ensemble* előrejelzésekre alkalmazták, de a determinisztikus előrejelzésekre is használható (Delle Monache et al., 2013). A fő különbség, hogy nem a légkör jövőbeli állapotára adunk előrejelzést, hanem a már meglévő adathalmazunkban keresünk analógiát az egyes értékek között. A módszer alkalmazása során az előrejelzési értékekre analógiát keresünk adott helyen illetve időpontra a múltban történő megfigyelt értékekkel. Az adatsorunkat két részre osztjuk fel. Az egyik a tanuló (*training period*), a másik a teszt (*test* vagy *verification period*) időszak. Az *ensemble* tagok számának meghatározására általában a tanuló időszak napjainak számának négyzetgyökét (kerekítve) használják (Vanvyve, 2014). Az *ensemble* tagok létrehozásánál a következő lépésekre van szükség:

A teszt időszak adott előrejelzéséhez (itt szélsőérték előrejelzés) megkeressük a tanuló időszakban lévő megfigyelt szélsőérték értékek közül azt az n darabot, amely a legkisebb mértékben tér el az előrejelzéstől, és az ezekhez tartozó prediktandusz (itt teljesítményértékek) értékek adják az *ensemble* n tagját (Nissen, 2012). A következő metrika szerint hajtható végre a tagok megkeresése (Delle Monache, 2011, 2013):

$$\|F_t, A_{t'}\| = \sum_{i=1}^{N_v} \frac{w_i}{\sigma_{f_i}} \sqrt{\sum_{j=-\tilde{t}}^{\tilde{t}} (F_{i,t+j} - A_{i,t'+j})^2} \quad (1)$$

ahol F_t az előrejelzés, melyre a t' időpontban az $A_{t'}$ analógiát keressük. N_v a változók száma, súlyuk w_i , míg σ_{f_i} az adott helyszínre a változó szórását jelöli a múltbéli előrejelzések esetén (Alessandrini et al., 2014).

VI. ALADIN

Az [2 – OMSZ] forrás szerint az Országos Meteorológiai Szolgálat számára az ALADIN projekt egy fontos nemzetközi együttműködés, a korlátos tartományú numerikus előrejelzések terén, amit még 1990 végén hozott létre a *Météo-France*. Napjainkra 16 nemzeti meteorológiai szolgálat tagja az együttműködésnek, aminek az OMSZ a kezdetektől fogva aktív tagja. Az együttműködés célja egy korszerű korlátos tartományú numerikus előrejelző rendszer megalkotása és fejlesztése. Ennek eredményeképpen, két regionális numerikus előrejelző modell az ALADIN és az AROME (*Applications of Research to Operations at Mesoscale*) fut operatíván a Szolgálatnál. A dolgozatom során felhasznált előrejelzett pillanatnyi szél és teljesítmény értékek az ALADIN modellből származnak.

Az ALADIN modell egy három-dimenziós rácson, naponta négyszer a főterminusokban előállítja a meteorológiai paraméterek (nyomás, szél, hőmérséklet, csapadékmennyiség, stb.) jövőben várható értékeit. Az előrejelzés alapvető módszertanát a folyadék-, vagy légkör-dinamika, a termodinamika, a sugárzástan és a felhőfizika képezi (a Navier-Stokes egyenletek numerikus, közelítő megoldását számítják ki, kiegészítve a víz fázisátalakulásának figyelembevételével). Az ALADIN modell futtatásával a nagy kiterjedésű (100-1000 km) időjárási folyamatok előrejelzésére helyezik a hangsúlyt, ezért egy 8 km-es horizontális rácsfelbontással és 49 db vertikális szinttel futtatják egy kontinentális Európát lefedő tartományon. Továbbá az ALADIN modell dinamikai magja lehetőséget ad mind hidrosztatikus (vertikális gyorsulások elhanyagolása), mind nem-hidrosztatikus integrálásra. A modell felbontása miatt a hidrosztatikus közelítést használják, mivel az ekkora vagy ennél nagyobb skálájú folyamatokban jó közelítéssel elhanyagolhatóak a vertikális gyorsulások a horizontális gyorsulásokhoz képest. Az ALADIN modellben használt fizikai parametrizációs csomag, az ARPEGE (*Action de Recherche Petite Echelle Grande Echelle*) globális modell, fizikai parametrizációinak regionális adaptációján és annak továbbfejlesztésén alapszik, így a modell parametrizálja a konvektív folyamatokat. Az ALADIN modell determinisztikus változatában egy lokális adatasszimilációs rendszert (3d-var) használnak, míg *ensemble* változatában adatasszimiláció nélküli leskálázás történik. Továbbá a gravitációs hullámokat digitális szűrővel távolítják el a kezdeti feltételekből. Az ALADIN modell korlátos tartománya miatt a futtatáshoz a tartomány határain kívül játszódó folyamatokról (kényszerekről) információt kap az előrejelzés során. Ezeket az oldalsó peremfeltételeket a determinisztikus ALADIN

modell futtatásához az ECMWF-ben létrehozott globális modell futtatások szolgáltatják. Az ALADIN modell valószínűségi előrejelző rendszere az ARPEGE modellre épülő PEARP (*Prévision d'Ensemble ARPEGE*) francia globális *ensemble* rendszerből kapja a kezdeti- és peremfeltételeit. Ez egy 35 tagú rendszer, amiből az OMSZ regionális rendszere az első 11 tagot használja peremfeltételként.

VII. Verifikációs módszerek

A meteorológiai előrejelzések bevalásának minőségét statisztikai mérőszámok segítségével határozhatjuk meg. Ezeket a mérőszámokat az előrejelzések és a megfigyelések számszerű összehasonlítása során kapjuk. Ezt a kiértékelési eljárást nevezzük verifikációnak [3 – OMSZ].

Determinisztikus és valószínűségi előrejelzések esetén különböző verifikációs módszereket szokás használni, mivel utóbbinál nem egyetlenegy előrejelzett értéket adunk meg, hanem előrejelzések sokaságát. Azonban, ha a valószínűségi előrejelzés tagjainak valamilyen középértékét (számtani közép vagy medián) vesszük, akkor már alkalmazhatóak rá egyes determinisztikus jellegű verifikációs módszerek (Feddersen, 2014). Ezek közé tartoznak például a különböző verifikációs hibastatisztikák, ahol arról kapunk információt, hogy az egyes meteorológiai paraméterek előrejelzett értéke mennyivel tér el a megfigyelt értéktől (Nurmi, 2003). Természetesen a valószínűségi előrejelzésekre számos más verifikációs módszer létezik ezen kívül.

Hibastatisztikák

Az előrejelzések és a megfigyelések számszerű összehasonlításának egyik legegyszerűbb módja az átlagos hiba vagy más néven szisztematikus hiba (*Mean Error* vagy *Bias*) kiszámítása, ami az előrejelzett és a mért értékpárok különbségének átlagából adódik.

$$Bias = \frac{1}{n} \sum_{i=1}^n (fc_i - obs_i), \quad (2)$$

ahol fc_i az i -edik előrejelzés középértékét, az obs_i az i -edik megfigyelést, és n pedig az esetek számát jelöli. Értéke elméletben mínusz végtelentől plusz végtelenig terjedhet, és egy tökéletes előrejelzésnél nulla értéket vesz fel. Azonban, ha a nagy hibák ellenkező előjellel szerepelnek, akkor a szisztematikus hiba értéke ekkor is közel lehet a nullához. Ennek az esetleges félreértésnek a tisztázására több megoldás is létezik, ezek közé tartozik az átlagos abszolút hiba (*Mean Absolute Error*), az átlagos négyzetes hiba (*Mean Squared Error*) és ennek négyzetgyöke (*Root Mean Squared Error*). Ezek értéke elméletben nullától plusz végtelenig terjedhet, és szintén nullát vesz fel egy tökéletes előrejelzésnél.

$$MAE = \frac{1}{n} \sum_{i=1}^n |fc_i - obs_i|, \quad (3)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (fc_i - obs_i)^2, \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (fc_i - obs_i)^2}, \quad (5)$$

ahol a jelölések megegyeznek a szisztematikus hibánál alkalmazott jelölésekkel. A különbségek négyzetre emelésének köszönhetően az átlagos négyzetes hiba a kiugró értékeknél érzékenyebb a hibákra, azaz jobban „büntet”, mint az (2), (3) és (5) hibafajta.

Rank histogram

A valószínűségi előrejelzéseknél használt egyik diagram típus a *rank histogram* (Jolliffe & Stephenson, 2012), mely az EPS (*Ensemble Prediction System*) szórását (*EPS spread*) értékeli. A feltételezés az, hogy egy ideális EPS rendszerben az *ensemble* minden tagja ugyanakkora valószínűséggel következik be, azaz a mért érték a sorba rendezett két szomszédos *ensemble* tag közé egyenlő gyakorisággal esik, így az x-tengely minden *bin*-jébe („rekesz”) ugyanannyi megfigyelésnek kell esnie. Tehát egy ideális EPS rendszernek a *rank histogram*-ja lapos és szisztematikus hibától mentes, így minden intervallumban egyenlő gyakoriságot mutat. Azonban a valóságban általában a legelső és a legutolsó *bin*-be sokkal több megfigyelés esik, mint a többibe, így jellegzetes U-alakot kölcsönözve a hisztogramnak, ami azt jelenti, hogy az *ensemble* szórása nem elég nagy, nem fedi le az előrejelzett értékek sokasága a megfigyelést. A szélső, nyílt intervallumokba eső megfigyeléseket *outlier* eseteknek nevezzük. Ha a diagram alakja J vagy L-alakú, akkor ez az EPS rendszerben fölé- vagy alábecslést jelent.

CRPS

Egy másik módszer a CRPS (*Continuously Ranked Probability Score*), ami megadja az integrált négyzetes eltérést az előrejelzés és a megfigyelés kumulatív eloszlás-függvényei között (Szamoránsky, 2009; Wilks, 2011; Jolliffe & Stephenson, 2012).

$$CRPS = \frac{1}{n} \sum_{i=1}^n \int_{x=-\infty}^{x=\infty} [F_i^f(x) - F_i^o(x)]^2 dx, \quad (6)$$

ahol $F_i^f(x)$ és $F_i^o(x)$ az előrejelzés és megfigyelés eloszlásfüggvényei.

Megbízhatósági diagram

Egy másik diagramtípus a megbízhatósági diagram (*reliability diagram*), ami összegzi az átlagos előrejelzett valószínűség és az esemény bekövetkezésének, relatív gyakoriságának négyzetes eltérését intervallumonként (Wilks, 2011; Jolliffe & Stephenson, 2012). Az előrejelzett valószínűség résztartományként az x-tengelyen, a megfigyelés relatív gyakorisága pedig az y-tengelyen vannak feltüntetve. Egy tökéletes előrejelzés megbízhatósági görbéje a diagram átlóját képezi, azaz ekkor pont annyiszor következett be a megfigyelés minden intervallumban, amekkora valószínűséggel előrejeleztünk. Ha a görbe az átló felett halad, akkor az előrejelzésünk alábecsül, ha alatta, akkor fölébecsül, tehát túl nagy valószínűséggel jelzi az esemény bekövetkezését.

VIII. Eredmények

VIII. 1. Felhasznált adatok

Az ALADIN modell 00 UTC-s futásából nyert előrejelzett pillanatnyi szélesség értékeket és az ezek alapján számított pillanatnyi teljesítmény értékeket hasonlítottam össze egy magyarországi szélérőmű mért egy óras átlagos szélesség értékeivel és mért egy óras átlagos teljesítmény értékeivel (a kapott megfigyelt adataim pontos értékéről és a szélérőmű elhelyezkedéséről titoktartás miatt nem számolhatok be). Az előrejelzett adataim 15 perces felbontásban 2012.06.06. 07:00 UTC és 2013.04.01. 15:00 UTC közötti intervallumból származnak, a megfigyelt értékek pedig óras felbontásban, magyar idő szerint 2012.06.01. 00:00 és 2013.06.01. 23:00 közötti időszakból állnak rendelkezésre. Az adatok típusából eredően, több hiba is jelentkezik a modellezés során. Elsők között az, hogy pillanatnyi értékeket próbálok összehasonlítani az óras átlagos értékekkel, később ennek kiküszöbölésére az előrejelzett pillanatnyi szélesség értékekből óras átlagos értékeket készítettem. Továbbá az előrejelzett értékeket az OMSZ-nek interpolálni kell, hiszen a rácspont és a megfigyelési hely különbözik. A fentiekén kívül a szélesség értékek mellett nem szerepeltek a szélirány adatok, pedig az *analog ensemble* fejezetben részletezett képletben ezeket mind lehet használni.

Vizsgálataim során *ensemble* előrejelzések különböző verifikációs módszereit használtam az eredményeim kiértékelésére, melyek leírását bemutattam az VI. fejezetben. Ezen módszerek létrehozására és ábrázolására, illetve a megfelelő adatok kiválogatására az R statisztikai programozási nyelvet választottuk (R Foundation for Statistical Computing, 2010; Kabacoff, 2011), ahol az általam készített programot használtam. A programozást az OMSZ-nél használt Debian GNU/Linux operációs rendszerrel ellátott, valamint a saját Windows 8.1 operációs rendszerrel ellátott számítógépen végeztem el.

Az *ensemble* tagok létrehozásánál fellépő problémák illetve a pillanatnyi szélesség értékeim kezelésére több verziót is létrehoztam.

VIII. 2. A vizsgált kísérletek

Az egyes EPS tagok létrehozására a V. fejezetben leírt (1) képletet használtam. Összesen három kísérleti futtatást végeztem négy féle beállítással (verzióval). Az első kísérletben (az első négy verzióban) az előrejelzett pillanatnyi szélesség értékekkel számoltam. Tehát a teszt időszakban lévő egyes előrejelzett szélesség értékeimet a tanuló időszakban lévő

összes megfigyelt szélesség értékkel összehasonlítottam, és a 12 legkisebb különbséghez tartozó megfigyelt teljesítmény értékek képezték az EPS tagokat a teszt időszakban lévő egyes sorokban.

Az első verzióban a teszt időszak első sorára (2. táblázat szerint a 3557. sor) az EPS tagjainak számítása:

$$\sqrt{(fc_v_{3557} - obs_v_{1-i})^2}, \quad (7)$$

ahol $i=1, \dots, 3556$. Ennek a távolságnak nézzük az első k legkisebb értékét.

Mivel a megfigyelt szélesség értékeim egy tizedesjegy pontossággal szerepeltek, ezért többször is előfordult, hogy több mint 12 azonos különbségem keletkezett, így a hozzá tartozó megfelelő teljesítmény értékek kiválogatása ebben az esetben nem volt egyértelmű.

Ezért a második verzióban figyelembe vettem mind a teszt időszakban, mind a tanuló időszakban az egyes szélesség értékek előtti és utáni értékeket is. Például a 3558. sorra:

$$\sqrt{(fc_v_{3557} - obs_v_{i-1})^2} + \sqrt{(fc_v_{3558} - obs_v_i)^2} + \sqrt{(fc_v_{3559} - obs_v_{i+1})^2}, \quad (8)$$

ahol $i=2, \dots, 3555$.

2. táblázat: az általam használt táblázat demonstrálása, fiktív adatokat tartalmaz. Az első oszlop a sorok száma, a második-harmadik az időpontot UTC-ben, a negyedik-ötödik az előrejelzett szélességet és teljesítményt, a hatodik-hetedik a mért szélesség és teljesítményt, az utolsó tizenkettő pedig az *ensemble* tagok értékét mutatja.

	Date	UTC0	fc_v	fc_P	obs_v	obs_P	1st	2nd	11th	12th
1	2012.06.06	8:00	7,251537	641,8321	7,5	758,7				
2	2012.06.06	9:00	6,924255	537,1671	7,1	600				
3	2012.06.06	10:00	6,720432	509,1243	6,9	550,2				
4	2012.06.06	11:00	6,221012	401,3354	5,8	391,2				
5	2012.06.06	12:00	7,543211	720,2056	7,9	860,3				
			:	:						
3552	2012.11.02	22:00	9,424214	1403,02	9,2	1330,3				
3553	2012.11.02	23:00	13,9421	1823,865	13,8	1815,2				
3554	2012.11.03	0:00	15,92381	1871,234	13,2	1804,5				
3555	2012.11.03	1:00	12,43921	1791,091	13,6	1811,3				
3556	2012.11.03	2:00	9,881701	1557,231	9,5	1456,7			..	
3557	2012.11.03	3:00	9,394182	1321,541	8,9	1251,2			..	
3558	2012.11.03	4:00	9,821091	1541,183	9,8	1550,4				
3559	2012.11.03	5:00	10,23841	1714,492	10,1	1700,3				
3560	2012.11.03	6:00	10,30831	1725,744	9,9	1651,1				
3561	2012.11.03	7:00	9,695281	1471,121	9,2	1315,6				
			:	:						
7108	2013.03.31	22:00	11,01321	1770,311	13,1	1799,5				
7109	2013.03.31	23:00	8,942821	1201,431	8,5	1000				
7110	2013.04.01	0:00	8,221801	937,931	8,1	950,4				
7111	2013.04.01	1:00	7,823219	827,3155	7,8	840				
7112	2013.04.01	2:00	7,11284	599,1292	7,5	760,3				

Ezzel sikerült elérni, hogy ne legyen több, mint 12 azonos legkisebb különbség, így egyértelmű volt a megfigyelt teljesítményértékek (EPS tagok) kiválogatása.

Továbbá készítettem még két verziót ennek analógiájára, az eltérés csupán annyi, hogy az előbbi két verziót időlépcsőnként tekintettem, viszont ennek az egyik hátránya, hogy sokkal kevesebb értéket hasonlítok össze, ami meglátásom szerint túl kevés a megfelelő kiértékelés elvégzéséhez, hiszen a teljes időszak így sem éri el az egy év hosszúságot.

Mivel mind a négy esetben túl nagy negatív szisztematikus hibát kaptam, így a második kísérletben a teszt időszakban lévő egyes előrejelzett szélesség értékeket a tanuló időszakban lévő összes előrejelzett szélesség értékkel hasonlítottam össze. Tehát, újabb négy verziót készítettem el. Ettől a módszertől azt vártam, hogy az előbb említett negatív szisztematikus hibát ki lehet küszöbölni. Ezen feltevés alapja, hogy a modellben vizsgált időszak alatt nem történtek változtatások, így a negatív szisztematikus hiba a tanuló időszakban is feltehetően jelen volt.

A harmadik kísérletben ezen kétféle módon előállított verziókat készítettem el, azzal a különbséggel, hogy a 15 percenként megadott pillanatnyi szélesség értékeimből egy óras átlagos szélességet hoztam létre három különböző módon. Az első módon előállított óras átlagos előrejelzett szélesség értékeknél öt pillanatnyi szélesség értéket vettem. Egy példán illusztrálva a 2. táblázat első sorára: 07:00, 07:15, 07:30, 07:45 és 08:00 perckor előrejelzett pillanatnyi szélesség értékek átlagát vettem. A második és harmadik módnál pedig négy értéket vettem figyelembe az óras átlagok előállításánál: második esetben a 07:15, 07:30, 07:45 és 08:00 percben előrejelzett pillanatnyi szélesség értékek átlagát vettem, harmadik esetben pedig a 07:00, 07:15, 07:30 és 07:45 percben előrejelzett pillanatnyi szélesség értékek átlagát vettem.

Tehát három kísérletet készítettem el, az első kettőt négy verzióval, a harmadikat pedig az első kettőre három módon, ami összesen 32 esetet jelent.

VIII. 3. Eredmények kiértékelése

Az előző alfejezetben részletezett verziók nagy száma miatt (32) kiválasztottam a legrosszabb és a legjobb eredményre vezetőt, és ezeknek az eredményeit mutatom be részletesen a továbbiakban.

Elsőként a 8-13. ábrákon a legrosszabb, azaz az első kísérlet első verziójának (a pillanatnyi előrejelzett szélességet a múltbeli mért szélesség értékekkel hasonlítottam össze) eredményeit részletezem.

A 8. ábrán látható, hogy az ALADIN modell szisztematikus hibája minden időlépcsőben jóval nagyobb, mint az *analog ensemble* átlagának és mediánjának szisztematikus hiba értékei. Ez az általában negatív szisztematikus hibát tekintve azt jelenti, hogy ezen futtatás eredményei ennek a verifikációs *score*-nak a tükrében általában véve rosszabbak a referenciának tekintett modellnél.

Az RMSE ábráján (9. ábra) megjelenítettem az *ensemble* szórását is, hiszen ideális esetben az átlag és a medián RMSE értékei megegyeznek az *ensemble* szórásával, valós esetben erős korrelációt és minél kisebb különbséget várunk ezen értékek között. Ezen az ábrán viszont azt láthatjuk, hogy a korreláció értéke igen kicsi és az átlag illetve medián RMSE-i a szórástól számított különbsége nagy, legalább 250 kW. Az ábra alapján elmondhatjuk, hogy ez az előrejelzés a determinisztikus *score*-ok összképe alapján, nem tűnik jobbnak az ALADIN előrejelzésénél. A továbbiakban megnézzük, hogy milyen képet mutatnak az eredményeim a valószínűségi előrejelzések kiértékelése során alkalmazott verifikációs eljárások tükrében.

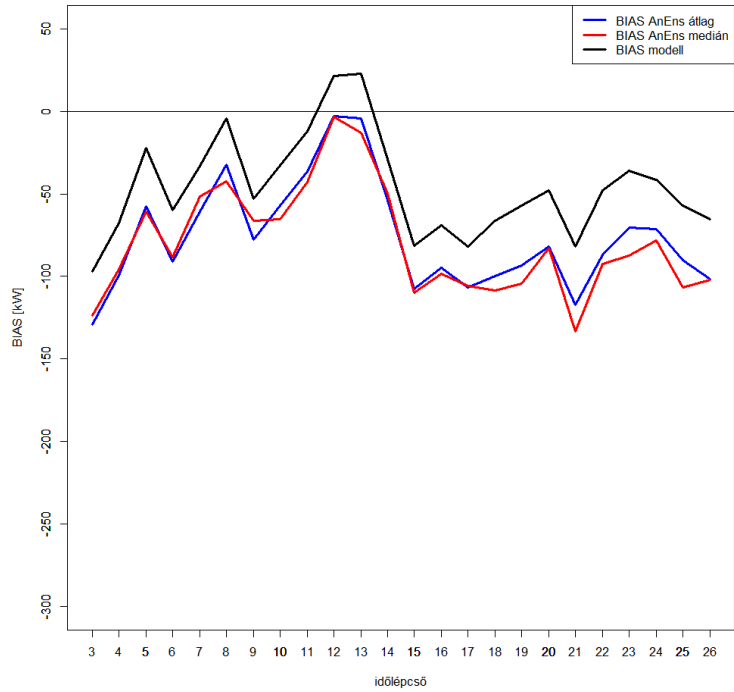
A *rank histogram* (10. ábra) a jellegzetes U-alakú képet veszi fel, ami azt mutatja, hogy az előállított statisztikai EPS szórása nem elég nagy, mint ahogyan azt a 9. ábrán is láthattuk. Így a legtöbb érték vagy kisebb, mint a sorba rendezett EPS tagok legkisebbje, vagy nagyobb a legnagyobbbnál.

A következő ábrán (11. ábra) a CRPS-t jelenítettem meg, amely determinisztikus esetben az átlagos abszolút hiba értéket adja, emiatt referenciaként ábrázoltam az ALADIN modell átlagos abszolút hibáját is. A két érték különbsége a 10 kW-os nagyságrendbe esik és ez esetben az *analog ensemble* értékei esnek közelebb a kívánatos nulla értékhez, de ezzel együtt is legalább 200 kW-os eltéréseket tapasztalhatunk az előrejelzett és megfigyelt értékek eloszlásfüggvényei között.

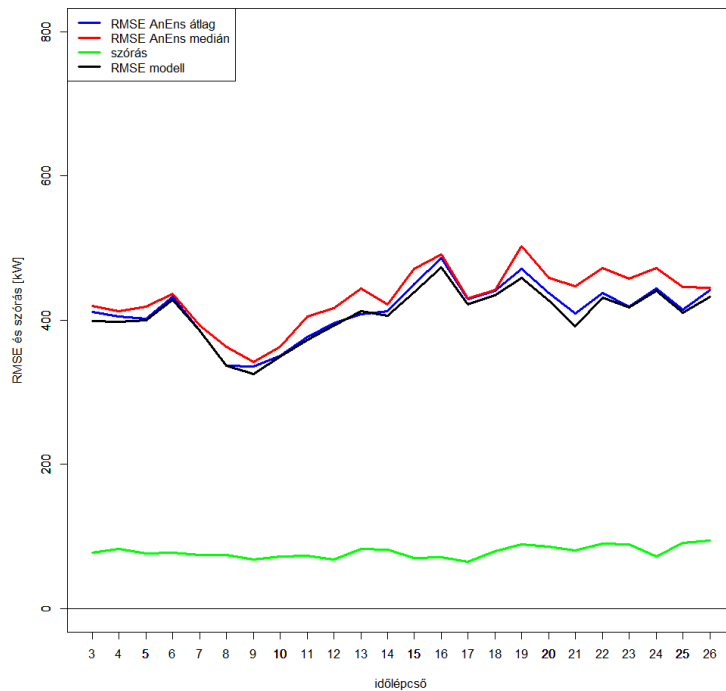
A megbízhatósági diagramon ideális esetben az értékek az átló mentén helyezkednek el, hiszen ekkor egyezik meg a vízszintes tengelyen ábrázolt előrejelzési valószínűség a függőleges tengelyen megjelenített megfigyelések relatív gyakoriságával. Ehhez az ábra típusához különböző küszöbértékeket definiálhatunk, és az ezeket meghaladó előrejelzési értékek valószínűségének függvényeként jelennek meg a megfigyelési gyakoriságok. Jelen esetben a küszöbértékeket 1000 illetve 1500 kW-nak választottam (12. és 13. ábra), mert a nagyobb teljesítmény értékeken nagyobb a hangsúly integrálhatóság szempontjából. Mindkét ábrán azt láthatjuk, hogy az előrejelzés bizonyos valószínűségi értékeket nem is adott ezekre a küszöbértékekre. A kisebb küszöbérték esetében a kis valószínűségi

értékeknél alábecsült, míg a nagy valószínűségű értékekhez tartozó megfigyelések relatív gyakorisága az előrejelzetnél kisebb értékű, az átló alatt helyezkedik el. Hasonló megfigyelést tehetünk a 13. ábra esetén, de ott a 75%-os valószínűséghez tartozó függvény érték pont az átlón helyezkedik el, ezzel az erre a valószínűségre jól előrejelzést indikálva.

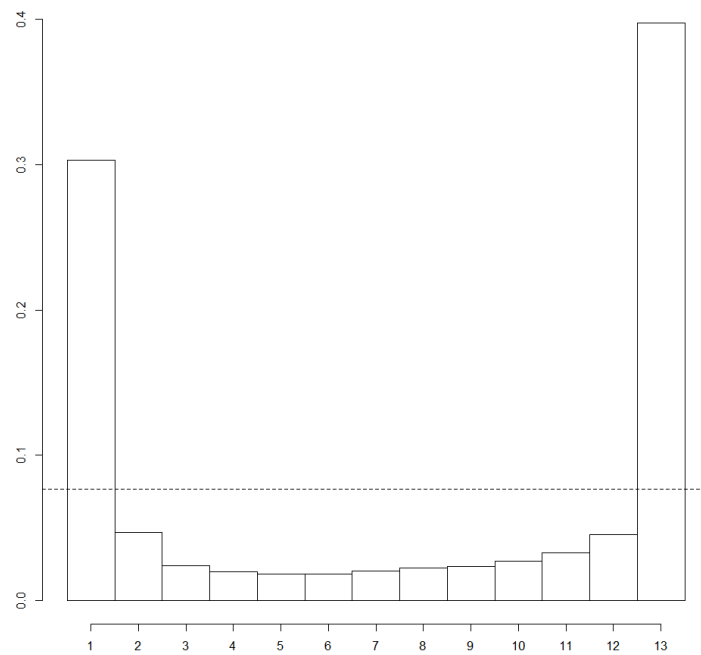
Összességében elmondható, hogy ezzel a verzióval kapott előrejelzés nem megfelelő mértékben reprezentálja a megfigyelt értékeket.



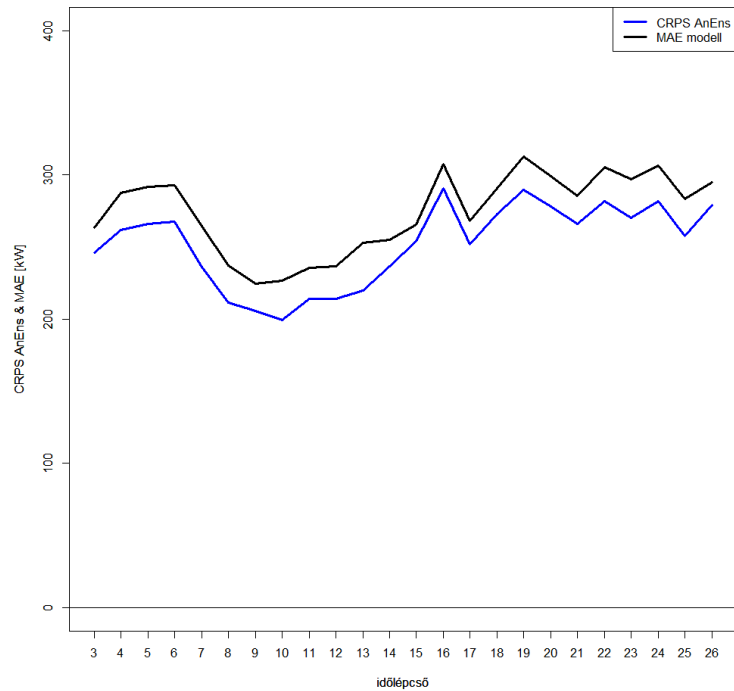
8. ábra: az ALADIN modell, az AnEns átlag és medián szisztematikus hibája az időlépcső függvényében.



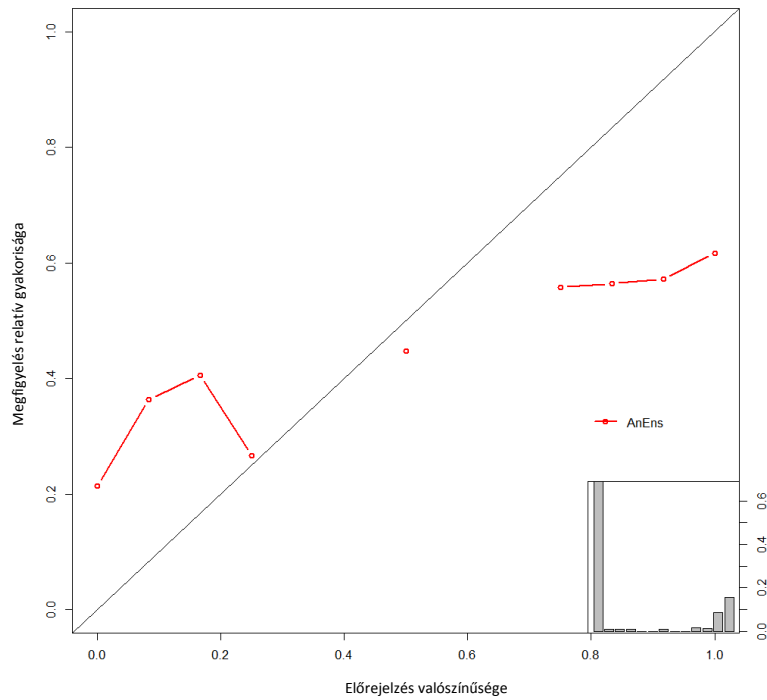
9. ábra: az ALADIN modell, az AnEns átlag és medián RMSE-je az időlépcső függvényében, valamint az *ensemble* tagok szórása.



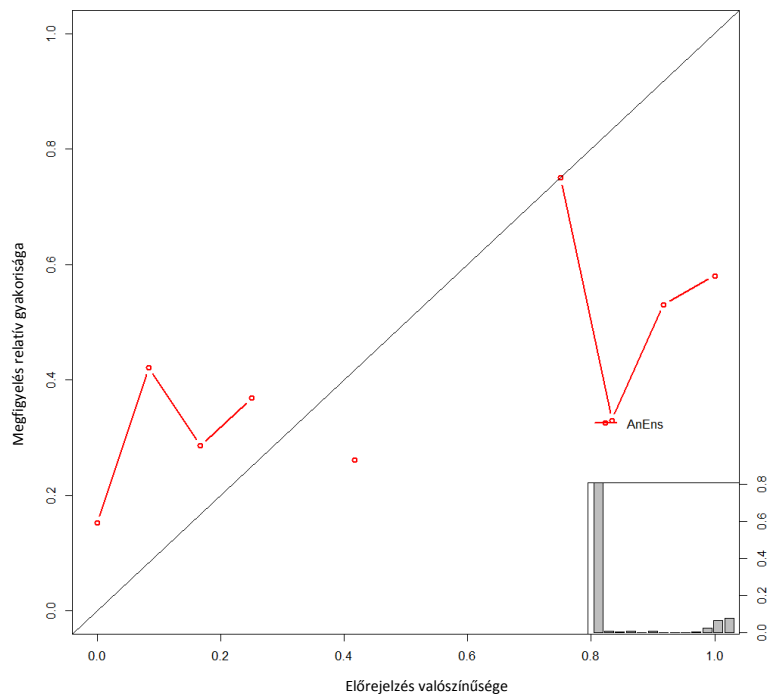
10. ábra: az EPS-ek rank histogram-ja.



11. ábra: az EPS-ek CRPS mutatói az időlépcső függvényében és az ALADIN modell átlagos abszolút hibája.



12. ábra: az AnEns megbízhatósági diagramja az előrejelzési valószínűség függvényében. A nominális teljesítmény fele (1000 kW) a küszöbérték.



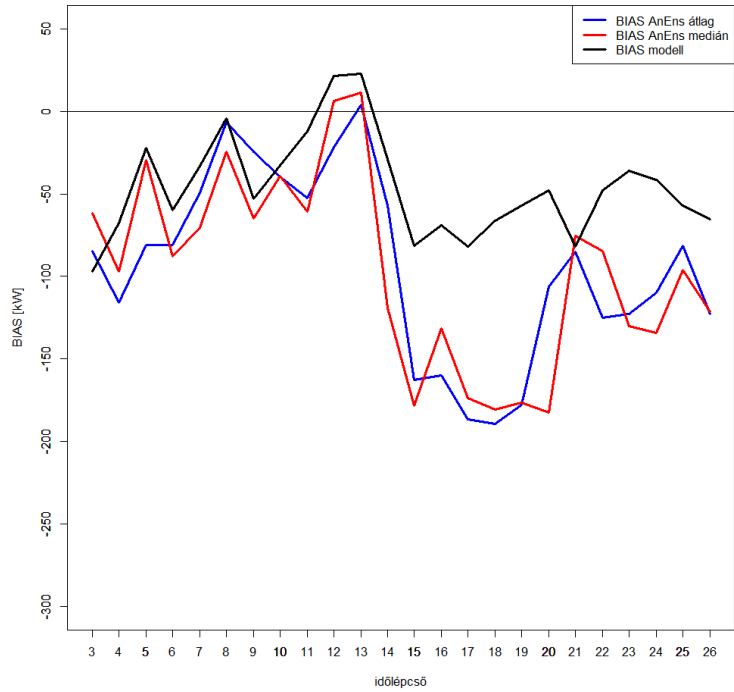
13. ábra: az AnEns megbízhatósági diagramja az előrejelzési valószínűség függvényében. A nominális teljesítmény harmomgyede (1500 kW) a küszöbérték.

Másodikként a 14-19. ábrákon a legjobb, azaz a második kísérlet negyedik verziójának (a pillanatnyi előrejelzett szélsőértéket a múltbeli pillanatnyi előrejelzett szélsőértékekkel hasonlítottam össze időlépcsőnként), eredményeit a továbbiakban közlöm.

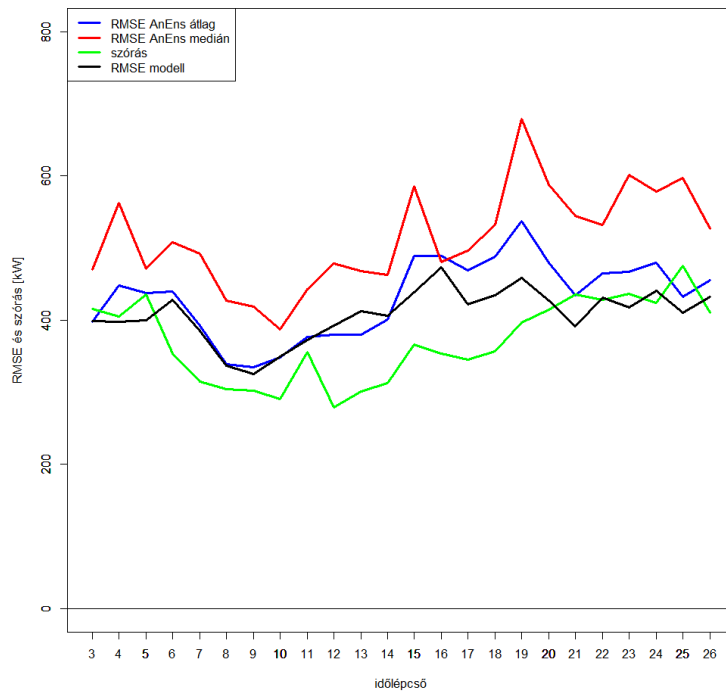
A 14. ábrán megfigyelhető, hogy a modell szisztematikus hibája bizonyos időlépcsőkben (12. és 13. időlépcső) már rosszabb, mint az *ensemble* mediánjának és átlagának az adott időlépcsőkre vett szisztematikus hiba értékei, viszont a többi időlépcsőben jóval nagyobb az eltérés a referencia és az *analog ensemble* minősége között, az ALADIN javára. Összességében itt is elmondható, hogy ezen futtatás eredményei ennek a verifikációs *score*-nak a tükrében rosszabbak, mint a referenciaként tekintett modellnél. Az RMSE ábráján (15. ábra) láthatjuk, hogy itt már a korreláció értéke nagyobb az RMSE értékek és az *ensemble* szórása között és az átlag illetve a medián RMSE értékeinek a szórástól számított különbsége a legtöbb esetben kisebb, mint 250 kW (az *ensemble* szórása nagyobb, mint az előző esetben). Elmondhatjuk továbbá, hogy összességében (a 13. időlépcső környékét leszámítva) az előrejelzés a determinisztikus *score*-ok alapján, nem tűnik jobbnak az ALADIN előrejelzésénél. A *rank histogram* (16. ábra) ebben az esetben már nem veszi fel a jellegzetes U-alak képét, azaz kiegyensúlyozott az egyes *bin*-ekbe eső megfigyelt teljesítmény, ami azt jelenti, hogy az EPS tagok szórása elég nagy ahhoz, hogy lefedje a egyes mért teljesítményértékeket. A következő ábrán (17. ábra) az látható, hogy az *ensemble* CRPS görbéje és a modell átlagos abszolút hiba értékeinek a különbsége nagyobb, azaz kb. 50 kW-os, azaz jóval jobb, a nullához közelebbi értékeket mutat az *analog ensemble* előrejelzésére. Jelen esetben legalább 150 kW-os eltéréseket tapasztalhatunk az előrejelzett és megfigyelt értékek eloszlásfüggvényei között.

A megbízhatósági diagramon (18. és 19. ábra) láthatjuk, hogy szintén a kisebb valószínűségi értékeknél alábecsült, és a nagy valószínűségű értékekhez tartozó megfigyelések relatív gyakorisága az előrejelzettnél kisebb értékű, az átló alatt helyezkedik el. Azonban a kettő közötti tartományon már az előrejelzési valószínűség a megfigyelt relatív gyakoriságával közel azonos.

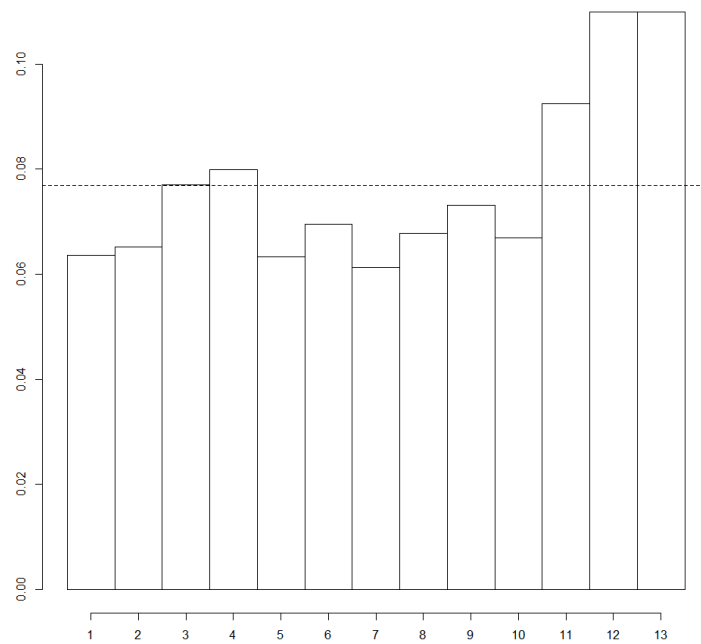
Összességében elmondható, hogy a determinisztikus *score*-okat tekintve ez a kísérlet valamivel rosszabbat mutat, mint a referenciaként tekintett ALADIN modell, de ha valószínűségi értelemben vizsgáljuk az együttes előrejelzést, akkor az *ensemble* által adott valószínűségek a valóságot jobban lefedik. Így, hogy ha az egyes értékek bekövetkezésének valószínűségére vagyunk kíváncsiak, akkor ez a kísérlet pontosabb eredményt ad, mint az általam készített összes többi.



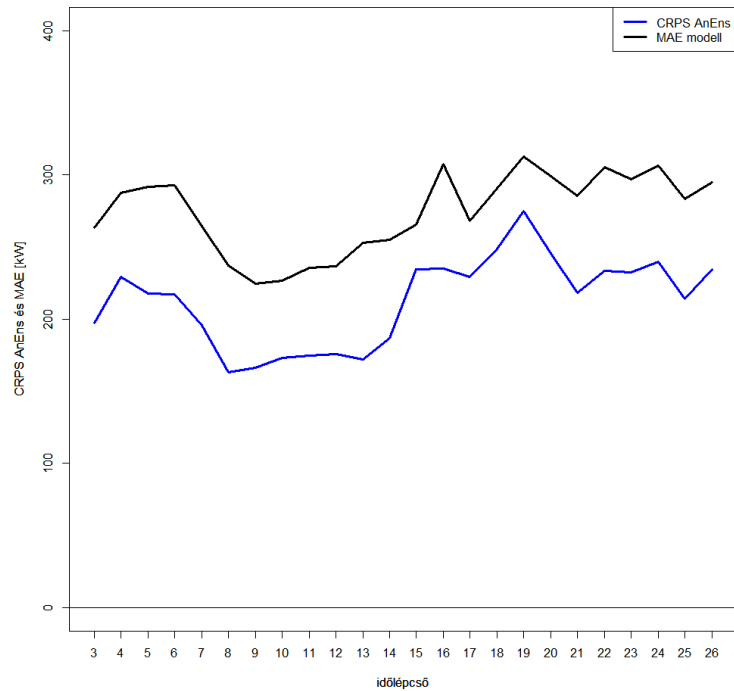
14. ábra: az ALADIN modell, az AnEns átlag és medián szisztematikus hibája az időlépcső függvényében.



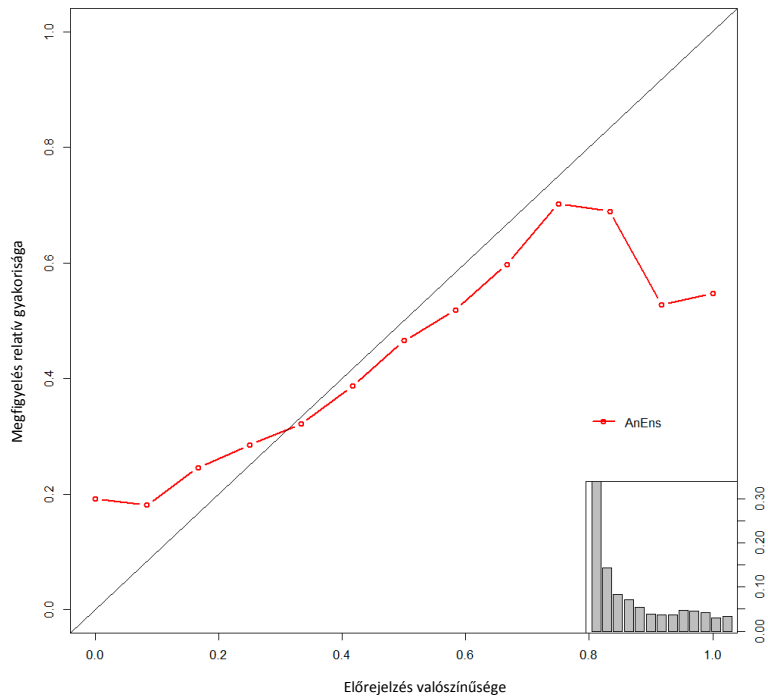
15. ábra: az ALADIN modell, az AnEns átlag és medián RMSE-je az időlépcső függvényében, valamint az ensemble tagok szórása.



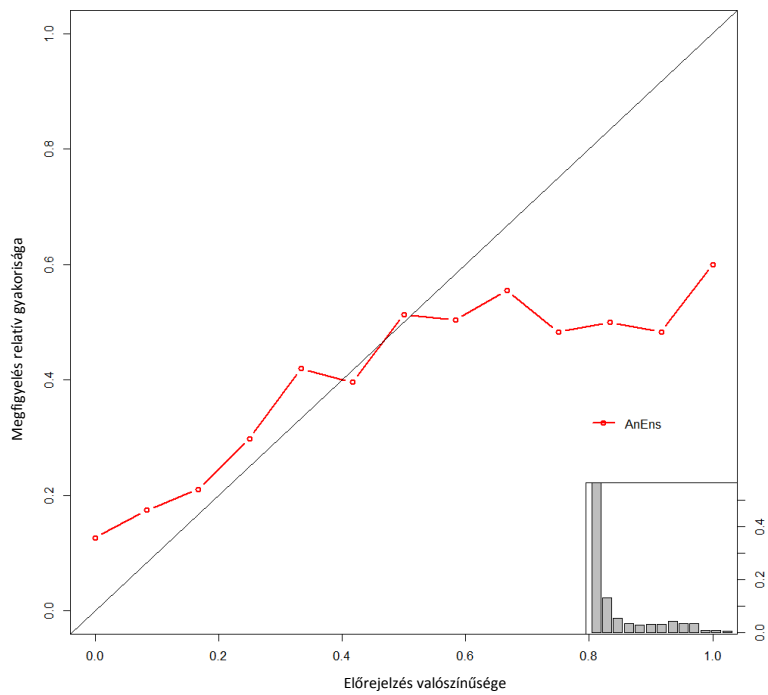
16. ábra: az EPS-ek rank histogram-ja.



17. ábra: az EPS-ek CRPS mutatói az időlépcső függvényében és az ALADIN modell átlagos abszolút hibája.



18. ábra: az AnEns megbízhatósági diagramja az előrejelzési valószínűség függvényében. A nominális teljesítmény fele (1000 kW) a küszöbérték.



19. ábra: az AnEns megbízhatósági diagramja az előrejelzési valószínűség függvényében. A nominális teljesítmény harmomgyede (1500 kW) a küszöbérték.

IX. Összefoglalás

A dolgozatom célja a magyarországi szélérőművek helyzetének ismertetése, valamint a szélérőművekre készített, OMSZ által futtatott ALADIN modell determinisztikus alapú előrejelzésének valószínűségi alapra helyezése. Az erre a célra kiválasztott módszer az *analog ensemble* technikáján alapszik, aminek segítségével egy olyan modellt állítottam fel, amelyben statisztikai összefüggéseket kerestem a múltban előrejelzett és mért szélsőségek között. Ez a módszer új és eddig kevesen foglalkoztak vele.

A vizsgálataimat egy magyarországi szélérőmű 2012.06.06. 07:00 UTC és 2013.04.01. 15:00 UTC időintervallumra vonatkozó 15 perces felbontású előrejelzett értékein, valamint magyar idő szerint 2012.06.01. 00:00 és 2013.06.01. 23:00 közötti időszak órás felbontású mért értékein végeztem. A vizsgálataim során különböző verifikációs módszereket használtam az eredményeim kiértékelésére. Ezen módszerek létrehozására, ábrázolására és a megfelelő adatok kiválogatására az R statisztikai programozási nyelvet választottam, ahol az általam készített programot használtam.

Az első kísérlet során a múltban előrejelzett pillanatnyi szélsőségek értékeit hasonlítottam össze a megfigyelt átlagos szélsőségek értékekkel négy verzióban. Mivel mind a négy esetben túl nagy negatív szisztematikus hibát kaptam és az ALADIN modellben a vizsgált időszak alatt nem történtek változtatások, így a második kísérletben a teszt időszakban lévő előrejelzett szélsőségek értékeit a tanuló időszakban lévő előrejelzett szélsőségek értékkel hasonlítottam össze szintén négy verzióban. Továbbá az előző két kísérletet elvégeztem egy újabb módon, ahol a pillanatnyi előrejelzett szélsőségek értékeimet háromféleképpen átlagoltam. Tehát összesen 32 esetet jelent ez.

Ebből a 32 esetből bemutatásra került az általam kiválasztott legrosszabb és legjobb eredményre vezető megoldás. Összességében elmondható, hogy az ALADIN modell determinisztikus előrejelzését az *ensemble* előrejelzéseim determinisztikus értelemben maximum csak megközelíteni tudják, viszont valószínűségi értelemben többlet információt adhatnak. Így ez által hasznos, a tervezést még inkább elősegítő információt nyújthatnak.

Mindezek alapján elmondható, hogy az ilyen jellegű vizsgálatoknak van létjogosultsága és további módszerekkel lehet az előrejelzéseken javítani.

A jövőben tervezem e módszert felhasználni több szélérőmű hosszabb időszakra vonatkozó adatain, illetve más technikákkal további javítást végezni az eredményeimben.

X. Köszönetnyilvánítás

Szeretnék köszönetet mondani mindazoknak, akik segítettek és biztattak a szakdolgozatom írása során.

Köszönöm témavezetőm, Brajnovits Brigitta szakmai segítségét, türelmét, biztatását.

Köszönöm tanszéki konzulensem, Dr. Matyasovszky István segítségét, amire bármikor számíthattam, amikor szükségem volt.

Köszönettel tartozom Szűcs Mihálynak OMSZ-ban tett javaslataiért.

Köszönöm Szűcs Dominikának az apróbb részletekben nyújtott segítségét, valamint családomnak és barátaimnak, hogy mellettem álltak.

XI. Irodalomjegyzék

- ALESSANDRINI, S., DELLE MONACHE, L., SPERATI, S. & NISSEN, J. N. (2014) – A novel application of an analog ensemble for short-term wind power forecasting. *Renewable Energy* **76** (2015): 768-781.
- BOUTTIER, F., VIÉ, B., NUISSIER, O. & RAYMAND, L. (2012) – Impact of Stochastic Physics in a Convection-Permitting Ensemble. *Monthly Weather Review* **140**, 3706-3721.
- BUIZZA, R., LEUTBECHER, M., ISAKSEN, L. & HASELER, J. (2010) – Combined use of EDA- and SV-based perturbations in the EPS. *ECMWF Newsletter* **123**, 22-28.
- DELLE MONACHE, L., NIPEN, T., STULL, R., LIU, Y. & ROUX, G. (2011) – Kalman Filter and Analog Schemes to Postprocess Numerical Weather Predictions. *Monthly Weather Review* **139**, 3554-3570.
- DELLE MONACHE, L., ECKEL, F. A., RIFE, D. L., NAGARAJAN, B. & SEARIGHT, K. (2013) - Probabilistic Weather Prediction with an Analog Ensemble. *Monthly Weather Review* **141**, 3498-3516.
- FEDDERSEN, H. (2014) – A Short-Range Limited Area Ensemble Prediction System. *Danish Meteorological Institute Technical Report 09-14*, 1-47.
- GARCÍA-MOYA, J.-A., CALLADO, A., ESCRIBÁ, P., SANTOS, C., SANTOS-MUNOZ, D. & SIMARRO, J. (2011) – Predictability of short-range forecasting: a multimodel approach. *Tellus A* **63**, 550-563.
- GÖTZ, G. (2001) – Ensemble előrejelzések. In: Káosz és prognosztika. *Országos Meteorológiai Szolgálat*, Budapest. 411-459.
- GWEC. (2015) – Global Wind Report 2014, Global Wind Energy Council, 80p.
- HARTMANN, B. (2012) – Szélerőművek integrálása a villamosenergia-rendszerbe. Doktori disszertáció, BME, Budapest (Dr. Dán András). 105 p.
- HEIZENREDER, D., TREPTE, S. & DENHARD, M. (2006) – A regional multi-model ensemble in Europe. *The European forecaster. Newsletter of the WGCEF* **11**, 29-35.
- HUNYÁR, M., VESZPRÉMI, K. & SZÉPSZÓ, G. (2006) – Újdonságok Magyarország szélenergia potenciáljáról. In: Magyarországi szél és napenergia kutatás eredményei (Dobi, I.). *Országos Meteorológiai Szolgálat*, Budapest. 94-113.
- ISAKSEN, L., HASELER, J., BUIZZA, R. & LEUTBECHER, M. (2010) – The new Ensemble of Data Assimilation. *ECMWF Newsletter* **123**, 17-21.
- JOLLIFFE, I., T. & STEPHENSON, D., B. (2012) – Forecast Verification: A Practitioner's Guide in Atmospheric Science, Second edition, pp. 134-136, 146-149.

- KABACOFF, R., I. (2011) – R in Action, Data analysis and graphics with R. ISBN: 9781935182399
- KERTÉSZ, S. (2006) – A perturbációk fejlődésének rövid jellemzése. Kézirat, Országos Meteorológiai Szolgálat.
- NISSEN, J., DELLE MONACHE, L., HAUPT, S., E., MARIC, T. & GULSTAD, L. (2012) – Analog Ensemble based power forecasting. *ES1002: Workshop 22-23 May, 2012*. 5 pp.
- NURMI, P. (2003) – Recommendations on the verification of local weather forecasts. *ECMWF Technical Memoranda No. 430*, 1-19.
- PALMER, T., N., BUIZZA, R., DOBLAS-REYES, F., JUNG, T., LEUTBECHER, M., SHUTTS, G., J., STEINHEIMER, M. & WEISHEIMER, A. (2009) – Stochastic Parametrization and Model Uncertainty. Technical Memorandum 598, ECMWF, Reading, UK
- R: a language and environment for statistical computing. R Foundation for Statistical Computing. ISBN: 3-9000051-07-0
- REN21. (2014) – Renewables 2014 Global Status Report (Paris: REN21 Secretariat). 216p.
- SZAMORÁNSKY, J. (2009) – Valószínűségi előrejelzések. Diplomamunka, ELTE, Budapest (Arató Miklós). 59 p.
- TOTH, Z. & KALNAY, E. (1997) – Ensemble forecasting at NCEP the breeding method. *Monthly Weather Review* **125**, 3297-3319.
- VANVYVE, E., DELLE MONACHE, L., MONAGHAN, A., J. & PINTO, J., O. (2014) – Wind resource estimates with an analog ensemble approach. *Renewable Energy* **74** (2015): 761-773.
- WILKS, D., S. (2011) – Statistical methods in the atmospheric sciences, Third edition, pp. 334-340, 351-354.
- WWEA. (2014) – Half-year Report 2014, The World Wind Energy Association, 8p.

XI. 1. Internetes hivatkozások

[1 – OMSZ]

<http://www.met.hu/omsz/tevekenysegek/idojarasmodellezes/valoszinusegi/>
[olvasva: 2014.10.10-én]

[2 – OMSZ]

<http://www.met.hu/omsz/tevekenysegek/idojarasmodellezes/modellek/#>
[olvasva: 2014.10.10-én]

[3 – OMSZ]

http://www.met.hu/omsz.php?almenu_id=homepages&pid=numprog&pri=20&mpx=0
[olvasva: 2014.10.09-én]

BARTHOLY, J., BREUER, H., PIECZKA, I., PONGRÁCZ, R. & RADICS, K. (2013a) - Megújuló energiaforrások. Internetes jegyzet.

<http://elte.prompt.hu/sites/default/files/tananyagok/MegujuloEnergiaforrasok/book.pdf>

BARTHOLY, J., PONGRÁCZ, R., BARCZA, Z., BIHARI, Z., LAKATOS, M., MÉSZÁROS, R., PIECZKA, I., PRÁGER, T. & RADICS, K. (2013b) – A modern globális éghajlatmodellek áttekintése. In: Klímaváltozás. Internetes jegyzet.

<http://elte.prompt.hu/sites/default/files/tananyagok/Klimavaltozas/book.pdf>

JÁNOSI, I. (2012) – A szélenergia hasznosításának hazai perspektívái. Internetes cikk.

http://mta.hu/koztestuleti_strategiai_programok_energia

SZÜCS, M., & HORÁNYI, A., (2013) – Valószínűségi előrejelzések. In: Alkalmazott számszerű előrejelzés: numerikus időjárási és csatolt modellek a gyakorlatban. Internetes jegyzet.

<http://elte.prompt.hu/sites/default/files/tananyagok/AlkalmazottNumerikusPrognosztika/book.pdf>