

BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM
GÉPÉSZMÉRNÖKI KAR
ENERGETIKAI GÉPEK ÉS RENDSZEREK TANSZÉK

TDK DOLGOZAT

Globálsugárzás előrejelzések utófeldolgozása gépi tanulás
segítségével

Soós Viktória
BSc IV. éfv.

Konzulens:

Dr. Mayer Martin János

BME-GPK Energetikai Gépek és Rendszerek Tanszék



Budapest, 2023

Absztrakt

Magyarország időjárását számtalan tényező befolyásolja, többek között a Napból érkező sugárzási mennyiség, a Kárpát-medence domborzati viszonyai vagy a légtömegek nedvességtartalma. A földrajzi szélesség és a légkör elnyelési tulajdonságai, ezen belül a felhőzet mennyisége nagyban meghatározza a besugárzás területi eloszlását. Mivel Magyarország viszonylag kis területi kiterjedéssel rendelkezik, a földrajzi szélesség hatása kevésbé érvényesül, a döntő szerepet a felhőzet játssza.

A napenergia előrejelzés pontosságára rendkívül fontos a villamosenergia-hálózatot tekintve. Az egyre több beépített naperőmű kapacitás nagy kihívást jelent a hálózatnak, és a naperőművek telepítése még inkább népszerűvé válik mind a lakossági, mind pedig az ipari oldalt tekintve. A naperőművek okozta hálózati ingadozást tartalék erőművekkel kell kompenzálni, hiszen elsődleges szempont az ellátásbiztonság. A kiegyenlítő energia meglehetősen drága, azonban az előrejelzések által ezek a költségek csökkenthetőek. A napenergia előrejelzés a napsugárzás előrejelzésen keresztül történik. Pontosabb napsugárzás előrejelzéseket pedig különböző utófeldolgozási módszerek segítségével érhetünk el.

A szélsőséges időjárási események, mint például az elhúzódó hóhullámok vagy a hideg időszakok megváltoztathatják a fogyasztási szokásokat és befolyásolják a termelési kapacitásokat is. A korai előrejelzések lehetővé tennék a szélsőséges időjárásokra való felkészülést, ezáltal akár a villamosenergia-rendszer túlterhelődése is elkerülhető lenne.

A sugárzás előrejelzések legelterjedtebb módja a numerikus módszerek alkalmazása. Ezek az előrejelzések rendszerint szisztematikus hibával terhelték, amik azonban az utófeldolgozással csökkenthetőek. A dolgozatomban kétféle numerikus előrejelző modell adatait, az AROME és az ECMWF modellből származó globálsugárzás előrejelzéseket, illetve az Országos Meteorológiai Szolgálat Pestszentlőrincen található központi obszervatóriumából származó megfigyelési adatokat vizsgáltam. A kutatásom során többféle regressziós módszert használtam a gépi tanulás által, amelyeken belül különböző prediktorkombinációkat alkottam meg, ezáltal próbáltam minél több lehetőséget tanulmányozni.

A dolgozatom végén egy egységes verifikáció alapján kielemeztem az előrejelzések jóságát. A kapott 5 érték, melyek az átlagos torzítási hiba, az átlagos abszolút hiba, az átlagos négyzetes hiba négyzetgyöke, a korrelációs tényező és a varianciák aránya alapján könnyedén össze tudtam hasonlítani az eredményeket.

A dolgozatom lezárásaként pedig a kiértékelés segítségével kiválasztottam a le pontosabb módszert, mely a kutatásom szerint a legjobb előrejelzést adta.

The weather in Hungary is influenced by a number of factors, including the amount of radiation from the sun, the topography of the Carpathian Basin and the moisture content of the air masses. The latitude and the absorption properties of the atmosphere, including the amount of cloud cover, largely determine the spatial distribution of irradiance. As Hungary has a relatively small surface area, the effect of latitude is less pronounced, with cloud cover playing the dominant role.

The accuracy of solar power forecasting is crucial for the electricity grid. The increasing installed solar capacity is a major challenge for the grid and the installation of solar power plants is becoming even more popular on both the residential and industrial side. Grid fluctuations caused by solar power need to be compensated by backup power plants, as security of supply is a primary concern. Balancing power is quite expensive, but these costs can be reduced by forecasting. Solar power forecasting is done through solar radiation forecasting. More accurate solar irradiance forecasts can be obtained by various post-processing methods.

Extreme weather events such as prolonged heat waves or cold spells can change consumption patterns and affect production capacity. Early forecasts would allow for the preparation of extreme weather events, thus avoiding overloading of the electricity system.

Numerical methods are the most common way of predicting radiation. These predictions are usually subject to systematic errors, which can be reduced by post-processing. In my work I have analysed data from two different numerical forecasting models, global radiation forecasts from the AROME and ECMWF models, and observational data from the central observatory of the Hungarian Meteorological Service in Pestszentlőrinc. During my research I used several regression methods by machine learning, within which I created different predictor combinations, thus trying to study as many possibilities as possible.

At the end of my work, I analysed the goodness of the predictions based on an unitary verification. I could easily compare the results based on the 5 values obtained, which are the mean bias error, the mean absolute error, the root mean square error, the correlation coefficient and the ratio of variances.

To conclude my work, I used the evaluation to select the most accurate method that gave the best prediction according to my research.

Kulcsszavak

gépi tanulás, globálsugárzás előrejelzések, utófeldolgozás

Rövidítések jegyzéke

AROME	Applications of Research to Operations at MEscale (Kutatási alkalmazások mezoskála műveletekre)
BT	bruttó beépített villamos teljesítőképesség
DHI	Diffuse Horizontal Irradiance (vízszintes diffúz sugárzás)
DNI	Direct Normal Irradiance (direkt normál sugárzás)
ECMWF	European Centre for Medium Range Weather Forecasts (Középtávú Időjárás Előrejelzések Európai Központja)
GFS	Global Forecast System (Globális Előrejelző rendszer)
GHI	Global Horizontal Irradiance (vízszintes globálsugárzás)
HMKE	háztartási méretű kiserőmű
HRES	High Resolution Forecasts (nagyfelbontású előrejelzések)
IFS	Integrated Forecasting System (Integrált Előrejelző Rendszer)
MAE	Mean Absolute Error (átlagos abszolút hiba)
MAVIR	Magyar Villamosenergia-ipari Átviteli Rendszerirányító Zrt.
MBE	Mean Bias Error (átlagos torzítási hiba)
MEKH	Magyar Energetikai és Közmű-szabályozási Hivatal
MLP	Multi Layer Perceptron (többrétegű perceptron)
NEKT	Nemzeti Energia- és Klímaterv
NWP	Numerical Weather Prediction (numerikus időjárás előrejelzés)
OBS	observed (megfigyelt)
OMSZ	Országos Meteorológiai Szolgálat
PV	photovoltaic (fotovoltaikus)
RMSE	Root Mean Square Error (átlagos négyzetes hiba négyzetgyöke)

Köszönetnyilvánítás

Köszönettel tartozom a konzulensemnek, Dr. Mayer Martin Jánosnak, aki segítségével hozzájárult a dolgozat megalkotásához, továbbá az Országos Meteorológiai Szolgálatnak az adatok biztosításáért.

Tartalom

1. Bevezetés	1
1.1. Globális elérendő célok és problémák.....	1
1.2. Aktuális adatok.....	2
1.3. A globálsugárzás komponensei és a sugárzással kapcsolatos alapfogalmak.....	4
1.4. Magyarország globálsugárzása	6
1.5. A kutatásom áttekintése	7
2. Numerikus időjárás előrejelző modellek	8
2.1. ECMWF modell.....	8
2.2. AROME modell.....	9
3. A globálsugárzás előrejelzések korlátai	10
4. Globálsugárzás előrejelzések és az utófeldolgozásuk	12
4.1. Gépi tanulás	14
5. A módszertan	15
5.1. Lineáris regresszió.....	16
5.2. Neurális hálózat	17
5.3. Az együttes módszerek	19
5.3.1. Gradient-boosted trees	19
5.3.2. Random forest.....	20
5.3.3. Extremely randomized trees.....	20
6. Verifikáció	21
7. Konklúziók és az eredmények értékelései	24
7.1. AROME modellek eredményei	24
7.2. ECMWF modellek eredményei	26
7.3. Az AROME és az ECMWF modellek eredményeinek összehasonlítása.....	27
8. Összefoglalás	31
9. Hivatkozások	32
10. Függelék	34

1. Bevezetés

Az egyre növekvő megújuló energiaforrás részarány a villamosenergia-hálózaton jelentős kihívásokat rejt magába. Az előrejelzések és azok utófeldolgozása emiatt még hangsúlyosabb szerepet kapnak, mert kiemelten fontossá válik a tervezhetőség és a menetrendezés a megújuló energiaforrások tekintetében.

1.1. Globális elérendő célok és problémák

A jelenkor egyik legnagyobb kihívása az éghajlatváltozás. A Kárpát-medence átlaghőmérséklete várhatóan 1-2 °C-kal emelkedik 2021-2050 között és gyakoribbá válhatnak a szélsőséges természeti jelenségek, mint például az árvizek, az aszályok vagy a hóhullámok [1]. Ahhoz, hogy a környezetet és az emberi jólétet megóvjuk, át kell állni egy fenntartható modellre, melynek segítségével megállíthatók a visszafordíthatatlan folyamatok. A modell koncepcióját az Európai Zöld Megállapodás (European Green Deal) tartalmazza, melyet még 2019. december 11-én terjesztettek elő [2]. A legfőbb cél, hogy 2050-re Európa a világ első klímasemleges kontinense legyen. Ehhez 2030-ig az 1990-es szinthez képest legalább 55%-kal kell csökkenteni az üvegházhatású gázok nettó kibocsátását, amit az Európai Klímarendelet (European Climate Law) rögzít [1]. Az Európai Bizottság „Irány az 55%!” (‘Fit for 55’) csomagja olyan konkrét intézkedéseket ír elő, melyek lehetővé teszik a kitűzött célok elérését [1]. A csomag az uniós jogszabályok felülvizsgálatát és aktualizálását, valamint új kezdeményezések bevezetését célzó javaslatokat tartalmaz. Biztosítja, hogy az uniós szakpolitikák összhangban legyenek az Európai Tanács és az Európai Parlament által elfogadott éghajlat-politikai célokkal [3]. A csomag részét képezi többek között a megújuló energiaforrások arányának növelése, az energiafogyasztás csökkentése, a biológiai sokféleség és az ökoszisztémák védelme [1].

2022 májusában létrejött a REPowerEU terv, melyet az energia világpiacán kialakult zavarok miatt dolgoztak ki. A tagállamok célja ezzel a tervvel az volt, hogy megvédjék az uniós polgárokat és vállalkozásokat az energiahíánytól, hiszen az elsődleges szempont minden esetben az ellátásbiztonság, valamint felgyorsítsák a tiszta energiára való átállást, illetve diverzifikálják az energiabeszerezést, ezzel is csökkentésük az orosz importfüggőséget [4]. A REPowerEU elősegíti a zöld átállást, és ösztönzi a beruházások megvalósulását a megújuló energia területén. 2022-ben összesen 41 GW teljesítményű új napenergia-kapacitást létesítettek, a szélenergia-kapacitást pedig összesen 16 GW-al növelték a tagállamok, így jelenleg a villamos energia 39%-a megújuló energiaforrásokból származik [4]. A tervek között szerepel a

kapacitások további bővítése, ugyanis az Európai Tanács 2023 októberében elfogadta a megújuló energiáról szóló új irányelvet. Az irányelvben célként szerepel, hogy a tagállamok közösen biztosítják, hogy a megújuló energiaforrásokból származó energia részaránya az Unió bruttó végső energiafogyasztásában 2030-ban legalább 42,5% legyen. A tagállamok azonban közösen törekedni fognak arra, hogy 2030-ra 45%-ra növeljék a megújuló energiaforrásokból származó energia arányát az Unió bruttó végső energiafogyasztásában [5] [6].

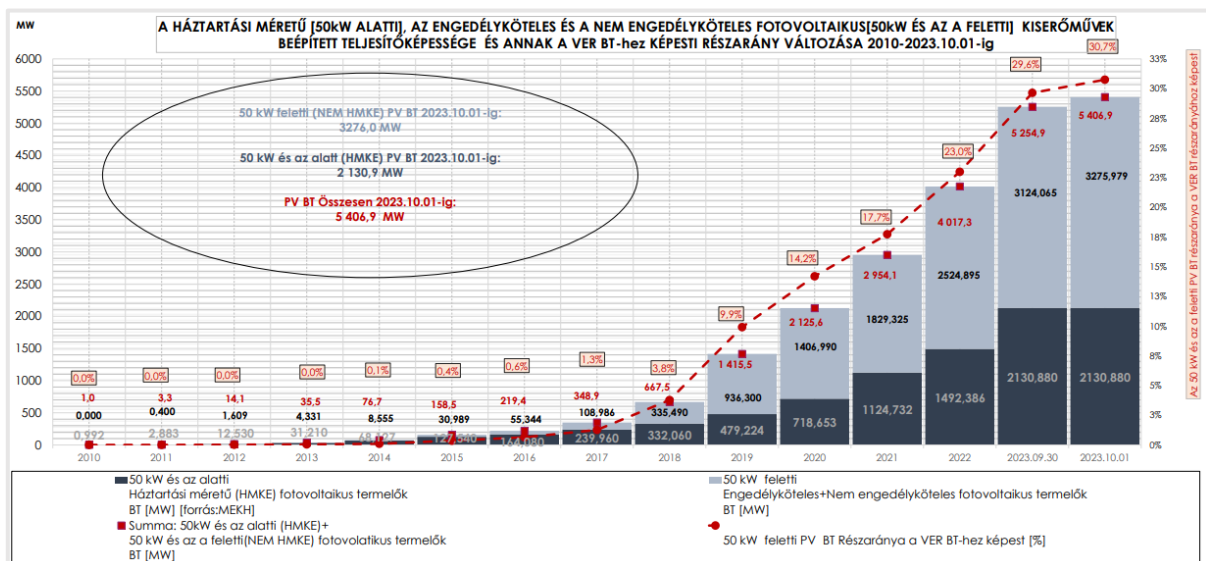
Magyarország klíma- és energiapolitikáját a Nemzeti Energia- és Klímaterv határozza meg. A jelenleg hatályos NEKT 2018-2019-ban lett kidolgozva, és 2020 elejétől vált hatályossá. A dokumentum elfogadása óta meghatározó jelentőségű változások következtek be, melyek hatással vannak a hazai energiaszektorra is, ennek következtében 2023-ban felülvizsgálatot eszközöltek. Magyarországon 2021-ben a megújuló energiaforrások bruttó végső energiafelhasználáson belüli részaránya 14,11% volt. A biomassza-hasznosítás révén még mindig leginkább a fűtési szektorban kihasználták a megújulóenergia-hasznosítási lehetőségeket, azonban a villamosenergia- és a közlekedési szektorral ellentétben a fűtési szektorban nem tudott növekedni a megújulóenergia-hasznosítás az elmúlt évtizedben. 2021-ben a megújulóenergia-felhasználás 69,2%-a fűtési-hűtési, 20,2%-a villamosenergia-termelési, 10,6%-a közlekedési célú volt. A felülvizsgált NEKT-ben új célszámot fogalmaztak meg a megújuló energiaforrásokat tekintve: a részarányukat legalább 29%-ra kell növelni 2030-ra a bruttó végsőenergia-felhasználáson belül. A korábban hatályos NEKT-ben 21% volt ez a célszám. A villamosenergia-termelésben a megújuló részarányának növelését naperőművi és szélenergia-termelési kapacitások bővítésével akarják elérni. A naperőművek esetében a tervek szerint 12 GW-ra nő a kapacitás 2030-ra, 2050-re pedig megközelítheti a 24 GW-ot. A lakossági oldalt nézve 2024 után minden új építésű ingatlan esetében kötelező, hogy 25%-ban megújuló forrásból biztosítsák fogyasztásukat, ezáltal még több háztartási kiserőmű telepítése várható a jövőben. A villamosenergia-fogyasztás részleges kiváltására a napelemes rendszerek telepítése egyre népszerűbbé válik. Az elmúlt évek tapasztalatai alapján látható, hogy 2030-ra jelentősen túlteljesül az a célkitűzés, miszerint legalább 200 ezer háztartás rendelkezzen átlagosan 4 kW teljesítményű, tetőre szerelt napelemmel [7].

1.2. Aktuális adatok

A MEKH és a MAVIR által elkészített kiadványban olvasható, hogy 9%-os kapacitásbővülés történt a hazai villamosenergia-rendszer beépített teljesítményében 2022-ben az előző évhez képest [8]. A növekedés elsősorban az 50 kW feletti új naperőművek csatlakozásával jelentkező 666,8 MW, a további 367,7 MW a növekvő háztartási méretű kiserőművek kapacitásának

eredménye [8]. Elmondható, hogy közelítőleg a naperőművek összességükben a teljes beépített teljesítmény egyharmadát teszik ki, ezzel jelentős mértékű időjárásfüggő kapacitást létrehozva. A naperőművek termelése nem csak az időjárástól függ, hanem a napszaktól is. A termelés előrejelzésének pontatlansága miatt nagy bizonytalanságot okoznak a villamosenergia-rendszerben. A bizonytalanságból adódó problémákat a rendszerirányító többlet tartalékok beszerzésével kezelte, hogy fenn tudják tartani a rendszer egyensúlyát. Ezeket a tartalékokat hagyományos erőművek, általában gyorsindítású gáztüzelésű erőművek adják, melyek termelési költségeit a földgáz ára szabja meg [8]. Ezen erőművi kapacitások szűkülése a tartalékok biztosítását egyre inkább az elvi határok felé tolja, ezzel jelentősen növelve a költségeket.

A legfrissebb adatok szerint a teljes hazai naperőművi kapacitás meghaladta az 5,4 GW-ot 2023. október elejéig, melynek közel 61%-a az ipari naperőművek, 39%-a pedig a háztartási méretű napelemes rendszerek [9]. Az idei évben eddig megközelítőleg 1400 MW-os bővülés valósult meg, mely több mint 300 MW-al nagyobb a tavalyi egész évben elért növekedéshez képest, ezzel megdöntve a 2022-es rekordot [9]. A háztartási méretű napelemes rendszerek 2030-ra elérendő célszámát (200.000 db) már most sikerült túlteljesíteni, ugyanis jelenleg több mint 236 ezer db HMKE üzemel hazánkban [9]. 2019-hez képest megnégyszereződött a háztartási méretű fotovoltaikus termelők száma. Az 50 kW feletti naperőművek maximális egyidejű teljesítménye 2023 augusztusában elérte a 2728,8 MW-ot, mellyel a beépített teljesítőképesség több mint 87%-a volt kihasználva [9]. A legjobb kihasználtsági arányt 2021 májusában jegyezték fel, amikor a beépített teljesítőképesség közel 98%-a volt üzemben [9].



1. ábra: A háztartási méretű, az engedélyköteles és a nem engedélyköteles fotovoltaikus kiserőművek beépített teljesítőképessége és annak a VER BT-hez képesti részarány változása 2010-2023.10.01-ig. (Forrás: [9])

A klímasemlegességi célok elérése érdekében növelni kell a megújuló energiaforrások részarányát, azonban az egyre nagyobb időjárásfüggő kapacitás miatt folyamatosan nő a szabályozási igény és annak költsége. Olyan fejlesztések szükségesek, melyek segítik a kiegyenlítő szabályozási piac hatékony működését, ezzel lehetővé téve a célok megvalósulását.

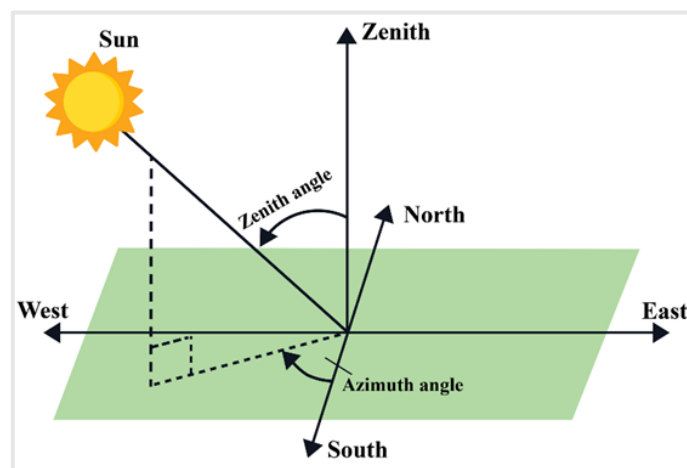
1.3. A globálsugárzás komponensei és a sugárzással kapcsolatos alapfogalmak

A direkt és a szórt sugárzás összege jelenti a globálsugárzást. A direkt sugárzás közvetlenül, míg a szórt sugárzás közvetve éri el a földfelszínt, mivel a légkör elemein, például a felhőzetben, különböző aeroszolokon szóródik. A légkör általi visszaverődés és elnyelés következtében a teljes sugárzás 40-50%-a éri el ténylegesen a földfelszínt [10].

A vízszintes globálsugárzás (GHI) a Föld egy vízszintes felületén mért, egységnyi területre jutó teljes napsugárzás, mértékegysége W/m^2 . Két komponensre bontható: direkt normál sugárzásra (DNI) és vízszintes diffúz sugárzásra (DHI). A vízszintes globálsugárzás, a direkt normál sugárzás (DNI) és a vízszintes diffúz sugárzás (DHI) közti kapcsolatot az (1) egyenlet mutatja meg [23].

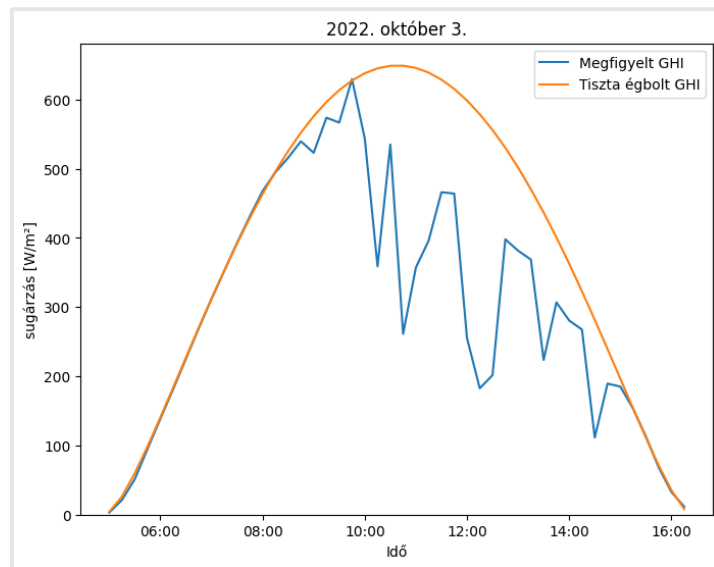
$$GHI = DHI + DNI \cdot \cos(\alpha_{zenit}) \quad (1)$$

Az (1) egyenletben az α_{zenit} a zenitszöget jelöli. A zenitszög a napsugár térbeli irányának a helyi függőleges iránnyal bezárt szöge. Az azimut szög a napsugár vízszintes felületre vetített irányának a déli iránnyal bezárt szöge. Ha az azimut szög negatív, akkor keleti, ha pozitív, akkor nyugati orientációt jelöl. A déli iránynál van a 0° , mivel a fotovoltaikus rendszerek ideális tájolása déli. Ebből az következik, hogy reggel a napfelkeltekor negatív, míg este naplementekor pozitív az azimut szög.



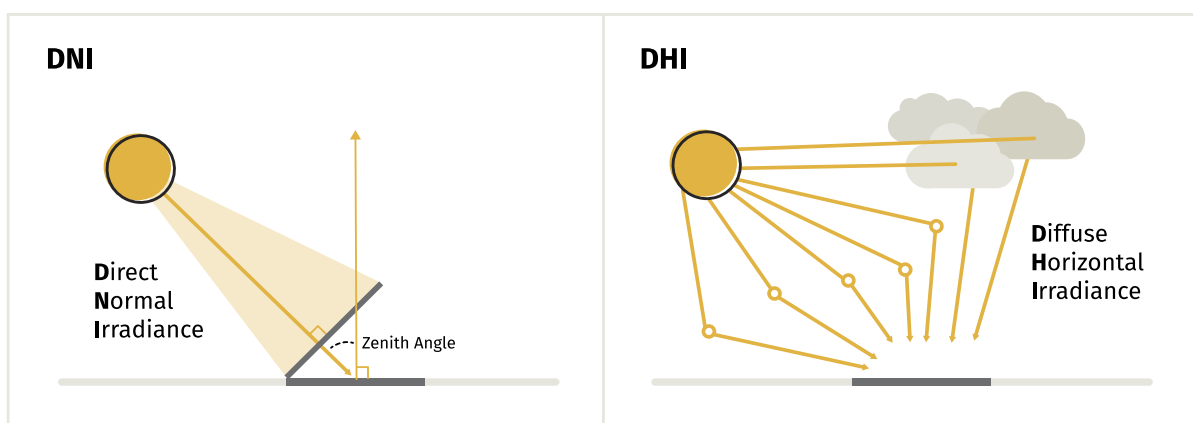
2. ábra: A zenitszög és az azimut szög.

A tiszta égbolt sugárzás a felhőzet figyelembevétele előtt számított sugárzási érték, ami több paramétertől függ. Többek között a napgeometria, a tengerszint feletti magasság, a vízgőz koncentráció, az ózon koncentráció, a szélesség és a hőmérséklet alapján számítják ki. A PV modellek esetében hasznos, ha figyelembe veszik a tiszta égbolt sugárzási adatait is, mivel segítségükkel különbséget lehet tenni a felhős időjárás és egyéb, a termelést befolyásoló körülmény között, mint például meghibásodások, szennyeződések vagy a hótakaró [23].



3. ábra: 2022. október 3-án a megfigyelt GHI és a tiszta égbolt GHI.

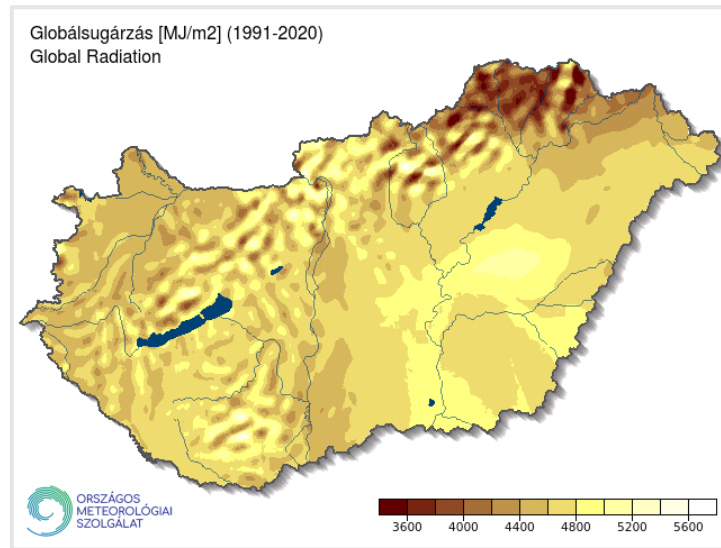
2022. október 3-ára kirajzoltattam a valós GHI-t és a tiszta égbolt GHI-t az idő függvényében, ez látható a 3. ábrán. Több időjárásjelentést is visszakeresve, ezen a napon felhőátvonulások voltak, napos és felhős időszakok váltották egymást, ami jól megfigyelhető a 3. ábrán is a valós GHI ugrásszerű változásait tekintve.



© 2021 Clean Power Research, L.L.C.

4. ábra: A DNI és a DHI. (Forrás:[23])

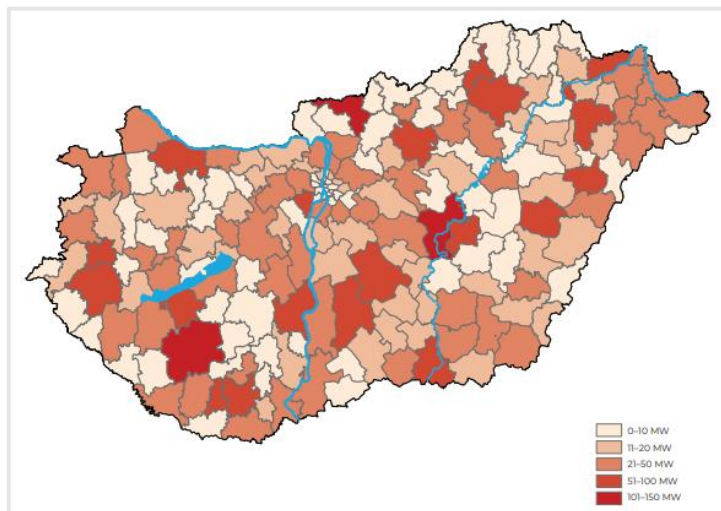
1.4. Magyarország globálsugárzása



5. ábra: A globálsugárzás átlagos évi összege Magyarországon (1991-2020). (Forrás:[11])

Hazánk nagy részén a globálsugárzás éves összege meghaladja a 4400 MJ/m²-t, de a Tiszántúl középső és déli tájain az 5000-5200 MJ/m² értéket is elérheti, míg a hegyvidéki tájakon a horizontális viszonyoknak köszönhetően akár 4000 MJ/m² alatti összegek is lehetnek [11]. Magyarország nyugati részén az átlagoshoz képest a nagyobb mennyiségű csapadékhoz köthető felhőzet és a domborzat miatt 4200-4400 MJ/m² az évi globálsugárzás [11]. Hazánkban a legtöbb besugárzás júliusban tapasztalható. A nyári napforduló júniusban van, emiatt júliusban a nappalok már rövidebbek és a Nap delelési magassága is kisebb, azonban a felhőzet mennyisége általában kisebb júniusához képest. A legkevesebb globálsugárzás decemberben van, mivel a nappalok ekkor a legrövidebbek, a felhőzet mértéke pedig nagy [11].

Az 5. és a 6. ábrát összevetve megállapítható, hogy közelítőleg azon a területen nagy a naperőművek száma, ahol nagy a globálsugárzás átlagos évi összege.



6. ábra: A naperőművek járásokonkénti eloszlása Magyarországon. (Forrás: [8])

1.5. *A kutatásom áttekintése*

A kutatásom fő célja, hogy többféle regressziós módszert megvizsgálva értékeljem az alkalmazott prediktorokat és módszereket az előrejelzések jóságán keresztül és kiválasszam a legjobb modelleket, melyek a legpontosabb előrejelzéseket adták.

A *2. fejezetben* bemutatom a numerikus előrejelző modelleket, melyeket a kutatásom során használtam. A *3. fejezetben* a globálsugárzás előrejelzések korlátait gyűjtöttem össze. A *4. fejezet* a globálsugárzás előrejelzésekről és azok utófeldolgozásáról szól. Az *5. fejezetben* a kutatásom módszertanát fejtem ki, mely során bemutatom az alkalmazott módszereket. A *6. fejezetben* a verifikációra térek ki, hogy hogyan is értékeltem ki a modellekből származó előrejelzéseket. A *7. fejezetben* pedig a konklúziók szerepelnek, illetve levonom az eredményekből a következtetéseket.

A napenergia részarányának növekedése a villamosenergia-hálózaton, valamint a naperőművek termelésének időjárásfüggő jellege miatt elengedhetetlen a napenergia előrejelzések fejlesztése, mely a napsugárzás előrejelzések fejlesztésén keresztül is történhet.

2. Numerikus időjárás előrejelző modellek

A múlt, a jelen és a jövő időjárását különböző eszközökkel kell vizsgálni. A múltról és a jelenről már rendelkezünk mért adatokkal, így a matematika statisztikai eszközeivel tudunk elemzéseket készíteni, míg a jövőt numerikus modellek által lehet tanulmányozni.

A numerikus időjárás előrejelző modellek egy háromdimenziós rácson állítják elő a különböző meteorológiai paraméterek (pl. globálsugárzás) jövőben várható értékeit matematikai számítások útján, a fizika törvényeit figyelembe véve [13]. Ilyen fizikai törvények például a folyadék-dinamikában használt Navier-Stokes egyenletek, a termodinamika első főtétele, valamint az anyag- és energiamegmaradás törvénye. Az előrejelzések kiszámítása számítógépen futtatható modellek segítségével történik [12]. A két legismertebb numerikus időjárás előrejelző modell az AROME és az ECMWF modellek.

2.1. *ECMWF modell*

Az ECMWF (European Centre for Medium Range Weather Forecasts), azaz a Középtávú Időjárás Előrejelzések Európai Központja 1975-ben jött létre 18 európai ország által, melyhez Magyarország 1994-ben csatlakozott, központja Nagy-Britanniában található. Az ECMWF modellek középtávú (2-10 napos) előrejelzések készítésére szolgálnak, havi és évszakos előrejelzéseket is készítenek, valamint a peremfeltételeket biztosítják a regionális modellek számára [15]. Az Integrált Előrejelző Rendszert fejlesztik, (Integrated Forecasting System), ami az egész Földre kiterjed, ezáltal globális előrejelzéseket készít [15]. Az IFS legnagyobb felbontású modellje a HRES (High Resolution Forecasts). Naponta kétszer futtatják, melyek 10 napra adnak részletes leírást a jövőbeli időjárásról, valamint további két időpontban 3,5 napra futtatják [36]. A modell számos fizikai kölcsönhatást is figyelembe vesz, többek között az óceán és a légkör, a talajnedvesség és a légkör vagy a hótakaró és a légkör kölcsönhatását [15]. 137 vertikális szint és 9 km-es horizontális térbeli felbontás jellemzi [15]. A modelltől származó vízszintes globálsugárzási adatokat felhasználtam a kutatásom során.

Az ECMWF előrejelzéseinek megbízhatóságát mutatja, hogy több mint 8 nappal korábban nagy biztonsággal jelezték előre, hogy 2012-ben a Sandy hurrikán eléri az amerikai kontinenst, mely egy trópusi ciklon volt a Karib-térségben és Észak-Amerika keleti partvidékén. Az ECMWF modell volt az egyetlen, amely ennyi idővel előre a legpontosabban megjósolta a hurrikán szokatlan pályáját, hiszen az amerikai meteorológiai szolgálatnál futtatott GFS modell az esemény előtt 5-7 nappal még azt jelezte előre, hogy a Sandy hurrikán várhatóan nem érinti a szárazföldi területeket [16].

2.2. *AROME modell*

Az AROME (Applications of Research to Operations at MEscale) nemzetközi projektet 2000-ben indították a francia meteorológiai szolgálatnál. Napjainkra 16 nemzeti meteorológiai szolgálat tagja az együttműködésnek, Magyarország a kezdetektől aktív résztvevője [13]. Az AROME [13] modellt a kis skálájú (10-100 km) időjárási folyamatok előrejelzésére használják, a horizontális rácsfelbontása 2,5 km, ezáltal nagy felbontású, regionális modellnek tekinthető. A földrajzi tartománya a Kárpát-medencére terjed ki, vertikális szintjeinek száma 60, valamint nagyon fejlett parametrizációs csomaggal rendelkezik. A nagy horizontális felbontás miatt szükséges a nem-hidrosztatikus dinamika, mert ebben az esetben nem elhanyagolható a vertikális sebesség időbeli változása, a vertikális gyorsulások hasonló nagyságrendűek a horizontális gyorsulásokhoz képest. A kezdeti feltételek meghatározása a numerikus előrejelzések első lépése, melyeket meteorológiai mérések és korábbi előrejelzések alapján szabnak meg. Az AROME modell esetében fontos, hogy a kezdeti feltételek minél pontosabbak legyenek a modell nagy felbontása miatt. Korlátos tartománya miatt a peremfeltételeket az ECMWF modell szolgáltatja, mivel elengedhetetlen a tartomány hatásain kívül játszódó folyamatok ismerete [13]. Az AROME modell rövidtávú előrejelzések készítésére szolgál, hiszen időbeli skálája órától pár naposig terjed (6-48 óra), ezáltal lehetőség van például az intenzív mediterrán csapadékok, a heves viharok, köd vagy a hóhullámok előrejelzésére. Az AROME modellt naponta nyolc alkalommal futtatják [14]. A modelltől származó vízszintes globálsugárzási adatokat felhasználtam a kutatásom során.

3. A globálsugárzás előrejelzések korlátai

Az előrejelzések nem lehetnek tökéletesek, az eredmények hibával terheltek. A hiba csökkentésére szolgálnak a különböző utófeldolgozási módszerek. Számtalan tényező együttesen befolyásolja a hiba mértékét, a sikeres előrejelzés több paraméteren múlik.

Tapasztalatok által megállapítható, hogy az előrejelzési időtartam növelésével nő az előrejelzési hiba értéke is, ezáltal hosszútávú előrejelzések készítése pontatlanabb. Ha szükséges a nagy pontosság, akkor célszerű rövidtávú előrejelzéseket használni. A légkör pillanatnyi állapotát nem ismerhetjük teljesen, hiszen nem tudjuk a Föld felszínét teljes mértékben meteorológiai állomásokkal lefedni. Az állomások közti területek időjárási paramétereinek értékeit interpolációval kaphatjuk meg, ami szintén bizonytalansággal járul hozzá. Az időjárási eseményeket szabályozó légkördinamikai törvényeket nem ismerjük teljes mértékben, a légkör kaotikus viselkedését nem tudjuk a fizikai törvények segítségével pontosan leírni, a bonyolult és komplex kölcsönhatásokat csak közelíteni tudjuk, illetve a leíró egyenletek megoldása során is számos egyszerűsítést alkalmazunk [15]. Az egyenletrendszerek megoldásához szükséges a kiinduló feltételek minél pontosabb meghatározása, amihez a megfigyelt adatokat használják. Ugyan a hiba valamelyest csökkenthető, ha minél több megfigyelt adat áll rendelkezésre, azonban a numerikus modellek egy szabályos háromdimenziós rácshálózat pontjaiban állítják elő az előrejelzéseket, a megfigyelések pedig térben szabálytalanul helyezkednek el. A két hálózat közötti kapcsolat létrehozásában az interpoláció segít, mely szintén lehet hibaforrás [15]. Ennek következtében megállapítható, hogy a numerikus időjárási modellek is hibával terheltek.

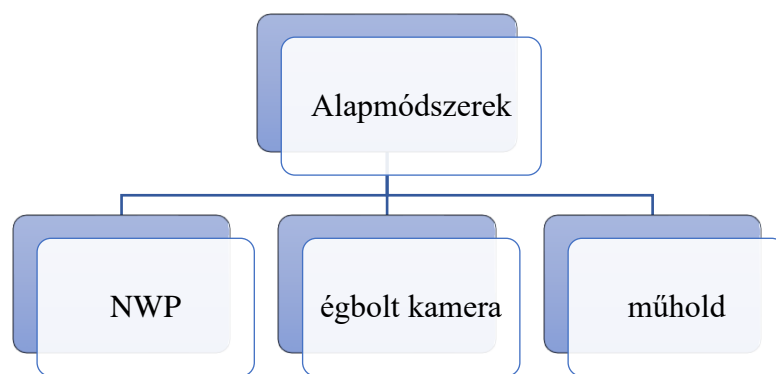
A numerikus időjárás előrejelző modellekben a napsugárzás előrejelzései erősen függenek a felhőmodellezéstől, ami viszont sztochasztikus jellegű. Az időjárás előrejelzési adatokat nem lehet közvetlenül egy PV teljesítmény előrejelzési modellbe bevinni, ugyanis a numerikus időjárás előrejelzések többsége nem azokat a változókat adja, amelyekre a PV modell használata során szükség lenne. Ahhoz, hogy maximalizálni tudják a termelést, a napelemeket a telephely szélességi fokával összemérhető dőlésszögben helyezik el, ennek következtében a globális ferde sugárzás lenne a PV modell egyik bemenete, de ez nem érhető el a numerikus időjárás előrejelzés során. Azonban a transzpozíciós modelleknek köszönhetően a globális horizontális sugárzásból, melyet a numerikus időjárás előrejelző modell ad, globális ferde sugárzást eredményez [18].

A napsugárzás, így a napenergia idősorai kettős mintázatot mutatnak a Föld keringése és forgása miatt: egy éves és egy napi ciklust. A szezonalitást az előrejelzés előtt el kell távolítani vagy külön komponensként kell kezelni. Erre szolgálnak a különböző tiszta égbolt modellek,

melyek leírják, hogy felhőmentes légkörben mekkora sugárzás éri a Föld felszínét. A tényleges sugárzás és a várható derült ég közötti arányt derült égbolt indexnek nevezzük. A felhők néhány másodperc alatt több száz W/m^2 -rel is csökkenthetik a sugárzást, ami nagy mértékben rontja az előrejelzés minőségét. A derült égbolt indexe jelentősen összefügg a felhők viselkedésével. Összességében elmondható, hogy a legjobb előrejelzés valószínűségi alapokon nyugszik. Sok más légköri változóhoz hasonlóan a napsugárzás is ötdimenziós: háromdimenziós teret, időt és valószínűséget ölel fel [18].

4. Globálsugárzás előrejelzések és az utófeldolgozásuk

A szoláris előrejelzési alapmódszereket az előrejelzési horizont alapján három nagy csoportba lehet sorolni: az égbolt kamerákon alapuló módszerek, melyek a 15 perc alatti előrejelzésre alkalmasak; a műholdas adatokon alapuló a napon belüli előrejelzésre alkalmasak; a numerikus időjárás előrejelzési módszerek pedig a másnapi előrejelzésekre alkalmasak. A három fő típuson kívül mikroméretű érzékelőhálózatokat is használnak perc alatti előrejelzésre. Azonban vannak statisztikai és gépi tanulási módszerek, melyek széles körben alkalmazhatóak regressziós problémák megoldásában, napelemes előrejelzésben is. Gyakran kombinálják ezeket NWP kimenetekkel, annak érdekében, hogy javítsák az eredményeket utófeldolgozással [18].

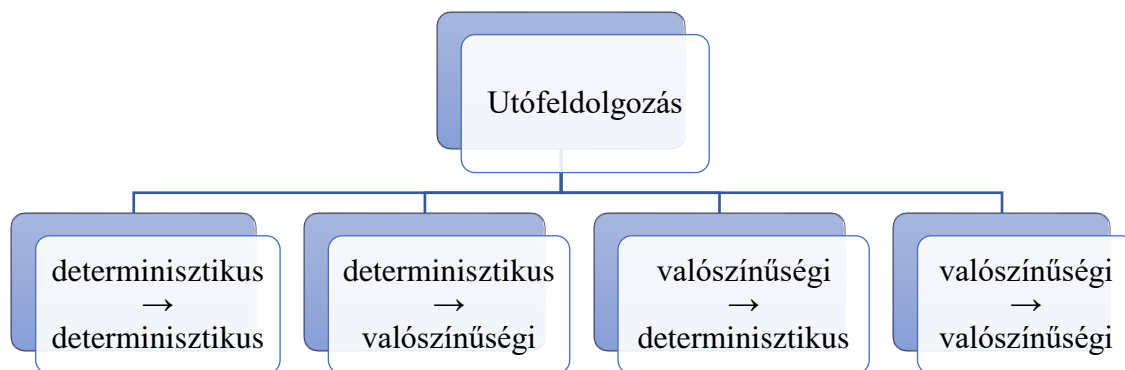


7. ábra: A szoláris előrejelzési alapmódszerek.

Az előrejelzések készítésének egyik legfontosabb lépése az utófeldolgozás, mely tulajdonképpen egyfajta optimalizálás, az előrejelző a múltbéli tapasztalatok alapján módosítja az előrejelzés értékeit. A kamera alapú, a műhold alapú vagy a numerikus előrejelzés, vagyis a fizikai alapú előrejelzések determinisztikusnak tekinthetők, mivel az előrejelzés egyszerűen csak egy szám. A determinisztikus előrejelzések átkonvertálhatóak valószínűségi előrejelzésekké és fordítva. Az utófeldolgozások módszereit többféleképpen is lehet osztályozni, de legegyszerűbb az előrejelzések jellegét figyelembe venni és az átalakítás iránya alapján csoportosítani. Négy nagy kategóriát különítenek el az utófeldolgozáson belül: determinisztikus determinisztikussá, determinisztikus valószínűségivé, valószínűségi determinisztikussá és valószínűségi valószínűségivé alakítás [17].

A determinisztikus előrejelzés csak egy értéket ad vissza, az előrejelző szerinti legjobb értéket. A determinisztikus elnevezés ebből következik, hiszen azt jelenti, hogy az események pontosan előrejelezhetőek, ha ismertek a kiinduló feltételek. Mindamelllett, hogy ezen a legjobb értéken is lehet javítani különböző utófeldolgozási technikákkal, a jövő bizonytalansága miatt nem hagyatkozhatunk csak determinisztikus előrejelzésekre. Az előrejelzések tulajdonképpen

véletlenszerű jövőbeli események, melyek mindegyikéhez rendelünk egy valószínűséget a bekövetkezéséről. Ezáltal a teljes előrejelzés egy eloszlás lesz, amit prediktív eloszlásnak szokás nevezni.



8. ábra: Az utófeldolgozás kategóriái.

A determinisztikus-determinisztikus utófeldolgozás célja, hogy korábbi minták által egy determinisztikus megfelelést hozzon létre az előrejelzések és a megfigyelések között. Az összefüggés segítségével pedig már lehetőség van a jövőbeli értékeket előrejelezni. Amikor egy új előrejelzés elérhetővé válik, a létrehozott összefüggés felhasználható az új determinisztikus előrejelzéshez [17].

A determinisztikus-valószínűségi utófeldolgozás egy jövőbeli esemény egyszeri felismerésével az ahhoz hasonló időjárási mintákat keres a múltból, majd ezeket a múltbéli megfigyeléseket a jövőbeli esemény egyformán valószínű kimeneteleként felhasználja [17].

A valószínűségi-determinisztikus utófeldolgozás a prediktív eloszlást egy statisztikai eszközzel (pl. átlag, medián vagy más kvantilisok) összefogja, ezáltal egy konkrét kimenetel lesz. A statisztikai eszköz megválasztása azonban gondos körültekintést igényel.

A valószínűségi-valószínűségi utófeldolgozás az együttes (ensemble) előrejelzéseket kalibrálja vagy különböző valószínűségi előrejelzéseket kombinál annak érdekében, hogy egy végső valószínűségi előrejelzést alkosson [17].

A napelemparkok a villamosenergia-hálózatra táplált teljesítménye függ a meteorológiai adatoktól, többek között a globálsugárzástól. Az időjárási adatokból a teljesítmény egy több elemű modellsor segítségével számítható ki, mely számos egyéb tényezőt is figyelembe vesz, mint például a napelemek porosodását vagy degradációját [19].

Dolgozatomban a globálsugárzásra és azok előrejelzéseinek utófeldolgozására fókuszálok, hiszen a napenergia előrejelzés a napsugárzás előrejelzésen keresztül valósítható meg. Kutatásom során csak determinisztikus adatokkal foglalkoztam.

4.1. Gépi tanulás

A gépi tanulás számos mérnöki területen széles körben alkalmazható, tulajdonképpen a mesterséges intelligencia egyik eszköze. A módszer előnye, hogy egy modell olyan problémákat is meg tud oldani, amelyek nem reprezentálhatók explicit algoritmusokkal, így a legtöbb esetben mindig találnak kapcsolatot a bemenetek és a kimenetek között [20]. Az előnye révén pedig a gépi tanulás kiválóan alkalmazható előrejelzési problémák megoldásában [20]. Minden esetben először megvizsgálja a rendelkezésre álló adathalmazt (tanítási adatsor), trendeket és kapcsolatokat fedez fel az egyes paraméterek között, amikből tanul. Ezt követően már el tudja végezni az előrejelzést egy tesztelési adatsoron.

A gépi tanulási technikáknak vannak paramétereik és hiperparamétereik [21]. A paraméterek a bemenetek és a kimenetek közötti kapcsolatot írják le, melyek a tanítási időszakban kerülnek beállításra. A hiperparaméterek meghatározzák az adott modell struktúráját, irányítják a tanítási időszakot, ezért ezeket még a modell futtatása előtt be kell állítani. Ebből következik, hogy mindig az adott problémára kell szabni a hiperparamétereket. A lehető legnagyobb pontosság csak a hiperparaméterek optimalizálásával érhető el. Ezt a folyamatot szokás hangolásnak nevezni. A legtöbb gépi tanulási könyvtár rendelkezik egy alapértelmezett hiperparaméter készlettel, de a pontosabb eredmény elérése érdekében ezeket a felhasználó is be tudja állítani.

A felügyelt tanulás (supervised learning) a gépi tanulás egyik megközelítése. Ebben az esetben a tanító adatbázis tartalmazza a magyarázó és magyarázott változókat, azaz a megfigyeléseket és a hozzájuk tartozó elvárt célértékeket. A cél az, hogy olyan modell jöjjön létre, amely korábban nem látott példákra is pontosan és helyesen működik, vagyis megtalálja a legjobb kapcsolatot a bemenetek és a kimenetek között [22].

A felügyelt tanuláson belül a feladattípusokat két nagy csoportba lehet sorolni: osztályozás és regresszió. Az osztályozás esetén adottak különböző kategóriák, a modell feladata, hogy az egyes elemeket besorolja a már meglévő kategóriákba [22]. Ebben az esetben a kimenetek kategorikus változók diszkrét értékekkel [21]. A regresszió esetén a célváltozó egy folytonos érték, a modell feladata, hogy az egyes elemekhez megjósolja, hogy milyen érték tartozhat [22]. A regresszió kimenetei folytonos értékek [21]. A globálsugárzás előrejelzések regressziós problémáknak tekinthetők.

5. A módszertan

Az adatok a Pestszentlőrincen található Országos Meteorológiai Szolgálat Marczell György Főobszervatóriumából származnak. Megfigyelt GHI, az AROME és ECMWF numerikus időjárás előrejelző modellek általi GHI, tiszta égbolt GHI, azimut szög és zenitszög adatokkal dolgoztam, melyek mind 2021-re és 2022-re is rendelkezésemre álltak negyedórás bontásokban, így két évnyi adatsort vizsgáltam. A megfigyelt GHI, az azimut szög és a zenitszög mért értékek, a többi pedig számított érték.

Az adatok feldolgozásához és az eredmények értékeléséhez a scikit-learn könyvtárat [35] használtam Python programozási nyelven, mely egy nyílt forráskódú gépi tanulási könyvtár. Különböző eszközökkel rendelkezik a modellillesztést tekintve, így könnyedén megvalósíthattam az ötleteimet általa.

Elsőként megalkottam a magyarázó változókat, melyeket a modellek bemeneteiként használtam fel. Az AROME és az ECMWF esetében is 12 prediktort hoztam létre a numerikus modellek GHI előrejelzéseinek, a tiszta égbolt GHI-nek, az azimut szögnek és a zenitszögnek a kombinálásával. Az egyes prediktorok az *1. táblázatban* láthatóak.

AROME prediktorok	ECMWF prediktorok
AROME GHI	ECMWF GHI
AROME GHI, tiszta égbolt GHI	ECMWF GHI, tiszta égbolt GHI
AROME GHI, azimut szög	ECMWF GHI, azimut szög
AROME GHI, zenitszög	ECMWF GHI, zenitszög
AROME GHI, azimut szög, zenitszög	ECMWF GHI, azimut szög, zenitszög
AROME GHI, azimut szög, tiszta égbolt GHI	ECMWF GHI, azimut szög, tiszta égbolt GHI
AROME GHI, zenit szög, tiszta égbolt GHI	ECMWF GHI, zenit szög, tiszta égbolt GHI
AROME GHI, azimut szög, zenitszög, tiszta égbolt GHI	ECMWF GHI, azimut szög, zenitszög, tiszta égbolt GHI
AROME GHI, zenitszög koszinusza	ECMWF GHI, zenitszög koszinusza
AROME GHI, azimut szög, zenitszög koszinusza	ECMWF GHI, azimut szög, zenitszög koszinusza
AROME GHI, zenitszög koszinusza, tiszta égbolt GHI	ECMWF GHI, zenitszög koszinusza, tiszta égbolt GHI
AROME GHI, azimut szög, zenitszög koszinusza, tiszta égbolt GHI	ECMWF GHI, azimut szög, zenitszög koszinusza, tiszta égbolt GHI

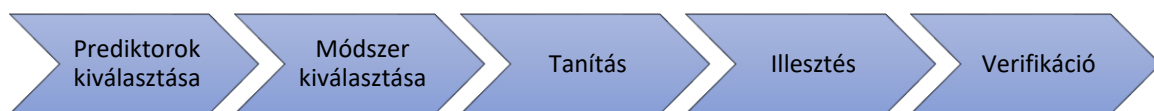
1. táblázat: Az AROME és az ECMWF prediktorok.

A két numerikus időjárás előrejelző modellből származó prediktorokat és az utófeldolgozási eredményeket a kutatásom során végig külön kezeltem, ugyanis két különböző elvű modellről van szó, így csak a hibaértékek kiszámolása után lehet azokat összehasonlítani.

A felügyelt tanulás 5 módszerét alkalmaztam. Minden esetben az alábbi pontok mentén haladtam:

1. Beolvastam az adott módszert, ezzel megalkotva az adott gépi tanulás struktúráját.
2. Az egyes prediktorok által létrejöttek a különböző modellek. A modellek a 2021-es adatsorra lettek betanítva, felhasználva a prediktorok adatsorait, valamint a megfigyelt GHI-t.
3. A tanítás után az illesztés következett. A modellek a magyarázó változók 2022-es adatsoraira illesztették a korábban megtanult összefüggéseket, mintázatokat, ezáltal létrejöttek 2022-re az előrejelzések.
4. Utolsó lépésként pedig a verifikáció, azaz az eredmények értékelése történt.

A 9. ábrán szematikusan ábrázoltam a folyamat egyes lépéseit.



9. ábra: Az utófeldolgozás lépései a kutatásomban.

A hiperparamétereket az adott módszer beolvasása előtt állítottam be a [21] cikk alapján, melyben 24 gépi tanulási módszert hasonlítottak össze a szerzők, valamint hiperparaméter hangolást is végeztek. 16 PV erőművet vizsgáltak, melyekre kiszámították a legjobb hiperparamétereket. A kutatásom során ezen legjobb értékeket vettem alapul és használtam fel, azonban fontos megjegyezni, hogy a tényleges legjobb hiperparaméterek az alkalmazástól nagyban függenek.

A kutatásom során 5 gépi tanulási módszert alkalmaztam, ezek közül egy lineáris (lineáris regresszió), a többi nemlineáris módszer.

5.1. Lineáris regresszió

A lineáris regresszió az egyik legegyszerűbb módszer, mely egyfajta lineáris kapcsolatot feltételez a függő és a független változók között.

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (2)$$

ahol $i = 1, \dots, n$ az illeszkedő minták indexe, x_i az i -edik előrejelzés, y_i az i -edik megfigyelés, ε_i az i -edik hiba, β_0 és β_1 regressziós együtthatók [17]. A β_0 és β_1 becült értékeivel kiadódik egy utófeldolgozott előrejelzés, miután egy numerikus előrejelző modell új előrejelzést adott ki.

A kutatásom első fázisában megvizsgáltam, hogy van-e lényegi különbség a hibára illesztés és a teljes sugárzásra illesztés között. Ezáltal a megfigyelés és a numerikus időjárás előrejelző modell közti hibára is betanítottam az egyes modelleket, így megkaptam az előrejelzett hibát 2022-re. A numerikus előrejelző modell 2022-es adataiból kivonva a hiba értéket kiadódott a tényleges előrejelzés. Hasonló módon jártam el, mint Lorenz, aki a numerikus időjárás előrejelző modell előrejelzéseinek torzítását regresszió formában fejezte ki [17].

$$bias_i = \beta_0 + \sum_{j=1}^m \beta_j x_i^{(j)} \quad (3)$$

ahol $bias_i$ az i -edik torzítás, $x_i^{(j)}$ a j -edik időjárási változóhoz tartozó i -edik független változó, β_0 és β_j regressziós együtthatók. Miután az ismeretlen együtthatók illesztve lettek, megjósolható a torzítás minden alkalommal, amikor új előrejelzést adnak ki a numerikus időjárás előrejelző modellek [17].

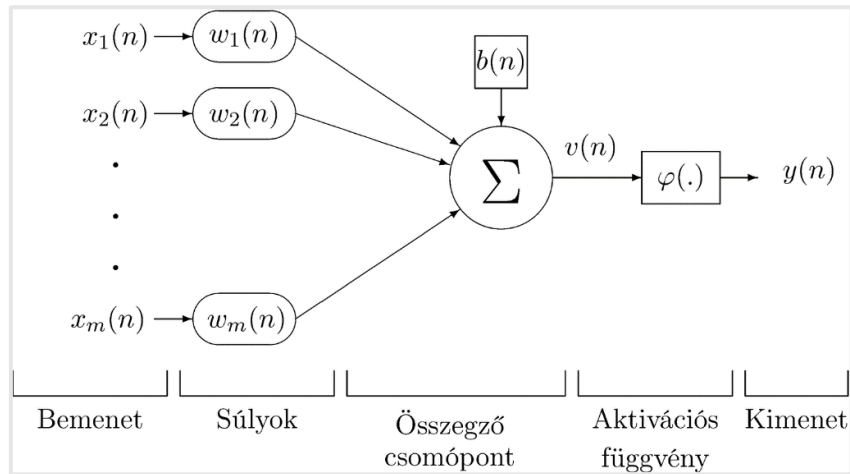
5.2. Neurális hálózat

A neurális hálózat [24] több kisebb egységből (neuron) épül fel, melyek között sok kapcsolat van. A neurális hálózat tulajdonképpen egymással összekötött neuronok összessége, a neuron pedig egy információfeldolgozó egység. Jelentős tanulási képessége van, azonban egy neurális hálózat nem akarja a jelenséget modellezni, hanem csak a bemeneti és kimeneti adatokat tekinti. Egy jó neurális hálózat ugyanarra a bemenetre hasonló kimenetet ad, mint a vizsgált jelenség.

Az egyetlen neuronból álló neurális hálózatot perceptronnak hívják. A bemenetek (x) ismertek egy adott n -edik időpillanatban, amiket a modell különböző súlyokkal (w) vesz figyelembe. A súlyok nem ismertek, éppen ezek meghatározása a feladat. A torzítás (b) értéke sem ismert, a cél ennek meghatározása is. Az összegző csomópont a bemenetek súlyozott összegét képezi.

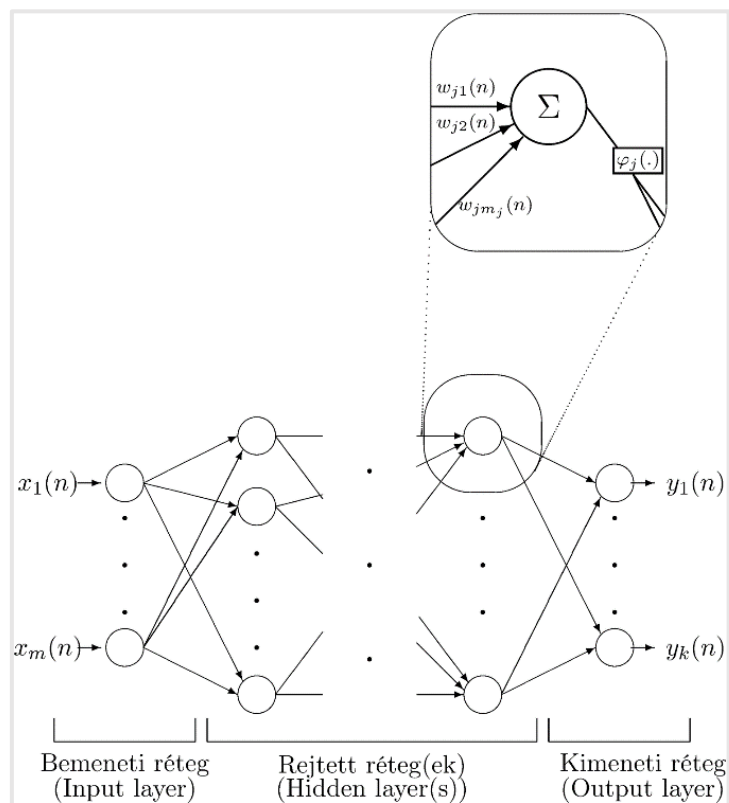
$$v(n) = b(n) + \sum_{i=1}^m w_i(n)x_i(n) \quad (4)$$

Az aktivációs, vagy más néven transzformációs függvény (φ) az összegzett értéket alakítja át, értéke pedig fontos, hogy a probléma szempontjából alkalmas intervallumba essen. A modell végeredményül egyetlen értéket (y) ad.



10. ábra: A perceptron felépítése. (Forrás: [24])

A többrétegű perceptron (MLP) [24] széles körben alkalmazott neurális hálózat, mely rétegekbe rendezett perceptronokból áll. A hálózat háromféle rétegből (layer) épül fel: bemeneti, rejtett és kimeneti rétegből. Rejtett rétegből bármennyi lehet, viszont bemeneti és kimeneti rétegből csak egy-egy. Fontos megemlíteni, hogy a többrétegű perceptron esetében minden neuronjának saját aktivációs függvénye és saját súlyai vannak.



11. ábra: A többrétegű perceptron felépítése. (Forrás: [24])

A többrétegű perceptron [25] a Pythonban egy felügyelt tanulási algoritmus, mely a fentebb ismertetett módon működik. Előnye, hogy nemlineáris összefüggéseket valós időben képes megtanulni [25]. Hátrányai közé tartozik, hogy a különböző véletlenszerű súlyozások eltérő validálási pontossághoz vezethetnek, valamint számos hiperparamétert tartalmaz, amiket az adott problémára kell hangolni a lehető legpontosabb eredmény eléréséhez [25].

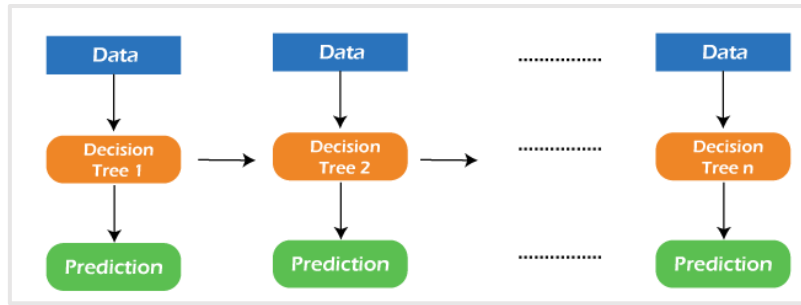
5.3. *Az együttes módszerek*

Az együttes módszerek (ensemble methods) egy adott tanulási algoritmussal létrehozott több alap modell előrejelzéseit kombinálják. Ezek a módszerek segítenek a modell robusztusságának, általánosíthatóságának javításában [26].

A döntési fa [28] egy nem parametrikus felügyelt tanulási módszer, mely a lentebb ismertetett modellek alapjául is szolgál. Ez a módszer egyaránt alkalmazható regressziós és osztályozási problémák megoldására. A cél egy olyan modell létrehozása, amely egy célváltozó értékét jósolja meg az adatjellemzőkből levezetett egyszerű döntési szabályok megtanulásával. A fa csomópontjaiban kérdések szerepelnek és a modellnek döntést kell hoznia, hogy mit válaszoljon az adott kérdésre, ezáltal eldől, hogy melyik következő csomópont felé halad tovább a folyamat. A döntések általi válaszok egyre szűkebb tartományt fednek le a lehetséges eredmények közül, így végül kialakul a végső érték.

5.3.1. *Gradient-boosted trees*

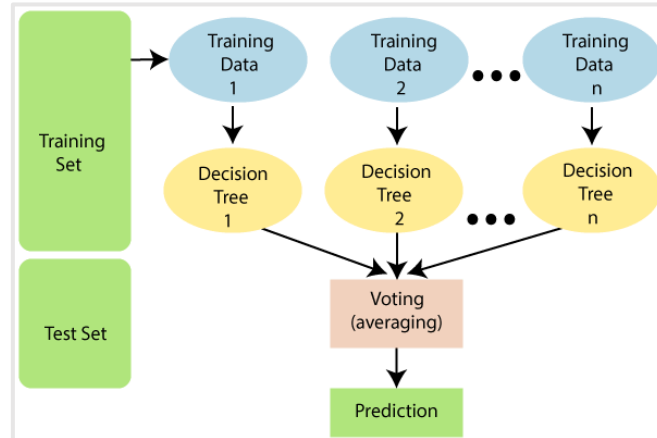
A gradient-boosting trees kiváló modell regressziós és osztályozási feladatokhoz. A regressziós gradient boosting egy additív modellt épít fel előremenő szakaszos módon [27], vagyis mindig egy fát adunk a modellhez és a korábbiakat már nem módosítjuk. Előnye, hogy lehetővé teszi tetszőleges differenciálható veszteségfüggvény optimalizálását [27]. Számos különböző veszteségfüggvényt támogat, melyet a loss argumentummal lehet megadni, alapértelmezett veszteségfüggvénye a négyzetes hiba [26]. Minden szakaszban egy regressziós fa illeszkedik az adott veszteségfüggvény negatív gradiensére [27]. Kis mintanagyság esetén előnyösebb lehet az alkalmazása [26].



12. ábra: A gradient-boosting trees módszer működése. (Forrás: [30])

5.3.2. Random forest

A random forest [29] egy fa alapú módszer, mely több döntési fát foglal magába, ezáltal az előrejelzések nem csak egy modell, hanem több modell segítségével készülnek. Minden fa a tanítási adatsorból vett mintából épül fel, majd az egyes fák valószínűségi előrejelzéseit átlagolja a módszer és megjósolja a végső kimenetet. Minél több döntési fát tartalmaz, annál pontosabb lesz az előrejelzés, ugyanis az egyes döntési fák jellemzően nagy varianciát mutatnak és hajlamosak a túlillesztésre. A véletlen (random) elnevezés onnan adódik, hogy az egyes fák a tanítási adatsorból véletlenszerűen választják ki azt a részhalmazt, melyeken a döntések alapulnak.



13. ábra: A random forest módszer működése. (Forrás: [31])

5.3.3. Extremely randomized trees

Az extra trees [29] felépítése hasonló a random forest módszeréhez azonban nagyobb véletlenszerűséggel rendelkezik. A tanítási adatsor véletlenszerű részhalmazát használják a döntési fák, de a módszer minden egyes döntéshez véletlenszerűen határozza meg a küszöbértékeket, és a véletlenszerűen generált küszöbértékek közül a legjobbat választja felosztási szabályként, így létrejönnek az újabb csomópontok. Ez általában lehetővé teszi, hogy a modell varianciája jobban csökkenjen, a torzítás némileg nagyobb mértékű növekedése árán.

6. Verifikáció

A használandó módszer, illetve a prediktorok kiválasztása befolyásolja a betanított modell pontosságát, ezáltal pedig az előrejelzés jóságát. Ahhoz, hogy a különböző eredményeket egymással össze lehessen vetni, egy egységes verifikációra, azaz értékelésre van szükség. Tulajdonképpen a prognózisok helyességét adja meg, ezáltal a gyengébben sikerült előrejelzés okait is ki lehet deríteni, ami nagyban hozzájárul a jövőbeli hasonló típusú hibák elkerüléséhez. A hosszabb időszakokra kiterjedő folyamatos vizsgálatok által lehetőség van a jellemző trendek meghatározására, illetve az adott modell fejlesztési irányainak kijelölésére is [32].

Az előrejelzés jósága ugyanakkor relatív, és függ attól, hogy milyen szempontot helyezünk előtérbe a művelet végrehajtása során. Murphy a jóság három típusát vázolta fel, melyek együttesen határozzák meg a jó előrejelzés fogalmát: következetesség, minőség, érték [33].

Egy előrejelzés akkor következetes, ha megfelel az előrejelző legjobb megítélésének. Ez az ítélet egy bizonytalansági faktort is tartalmaz, hiszen többek között a légkördinamikai törvények miatt az előrejelző ismeretei szükségszerűen hiányosak. Különböző pontozási szabályok által biztosítható a konzisztencia. A legjobb pontszámot akkor kapják az előrejelzők, ha előrejelzéseik megfelelnek az ítélőképességüknek [33].

A minőség az előrejelzések és a megfigyelések közötti megfelelésre utal. A MAE (Mean Absolute Error = átlagos abszolút hiba) és az RMSE (Root Mean Square Error = átlagos négyzetes hiba négyzetgyöke) olyan mérőszámok, amelyek az előrejelzések általános pontosságát értékelik. A pontosság az előrejelzés minőségének egyik szempontja, ami kvantitatív mérőszámokon keresztül értelmezhető. Az előrejelzés minősége az MBE (Mean Bias Error = átlagos torzítási hiba) mérőszámmal is jellemezhető. Ha két előrejelzési módszer azonos hibaértéket ad vissza, még nem jelenti azt, hogy egyformán jók, mert nem mindegy, hogy milyen körülmények uralkodnak [33].

Az érték az előrejelzésekből eredeztethető előnyökkel vagy költségekkel kapcsolatos. Az ilyen előnyök vagy költségek egy adott döntési probléma jellemzőitől függenek, a döntéshozók határozzák meg és értékelik. Az előrejelzések nem rendelkeznek konkrét belső értékkel, hanem a felhasználóik döntéseinek befolyásolásával nyernek értéket [33].

A mérőszámok összehasonlításának módja lehet pozitív orientált (minél nagyobb, annál jobb, mint például a képességpontszám), lehet negatív orientált (minél kisebb, annál jobb, mint például az RMSE) vagy lehet közép-re orientált (minél közelebb van a középértékhez, annál jobb, mint például az MBE). Meg kell jegyezni, hogy a mértékorientált megközelítés jelentős mértékben szubjektív [33].

Az eredményeim értékelésére az átlagos torzítási hibát (MBE), az átlagos abszolút hibát (MAE), az átlagos négyzetes hiba négyzetgyökét ($RMSE$), a korrelációs együtthatót (ρ) és a varianciák arányát (F) alkalmaztam.

$$MBE = \frac{1}{n} \sum_{i=1}^n (f_i - o_i) \quad (5)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |f_i - o_i| \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (f_i - o_i)^2} \quad (7)$$

ahol f_i az i -edik előrejelzett vízszintes globálsugárzás, o_i az i -edik megfigyelt vízszintes globálsugárzás, n pedig az adatok száma. Az előrejelzés annál jobb, minél közelebb van MBE értéke a 0-hoz, illetve a MAE értéke és az $RMSE$ értéke minél kisebb.

A varianciák aránya (F) megmutatja, hogy a megfigyelések változékonyságának mekkora hányadát képezik az előrejelzések.

$$F = \frac{\mathbb{V}(f)}{\mathbb{V}(o)} \quad (8)$$

ahol $\mathbb{V}()$ a variancia operátort, f az előrejelzést, míg o a megfigyelést jelöli. Ha $F < 1$, vagyis $\mathbb{V}(f) < \mathbb{V}(o)$, akkor alulszórt, míg ha $F > 1$, vagyis $\mathbb{V}(f) > \mathbb{V}(o)$, akkor túlszórt előrejelzésekről beszélhetünk [34].

A korrelációs együttható (ρ) az előrejelzés és a megfigyelés közti összefüggést méri [34]. Ha ρ nullával egyenlő, akkor a két mennyiség között nincsen lineáris kapcsolat, ha 1-gyel egyenlő, akkor együtt változnak, ha -1-gyel egyenlő, akkor ellentétesen változnak.

$$\rho = \frac{cov(f,o)}{\sigma_f \sigma_o} \quad (9)$$

ahol $cov(f,o)$ az előrejelzés és a megfigyelés közti kovarianciát jelöli, és a két mennyiség közti lineáris kapcsolatot fejezi ki. σ_f és σ_o az előrejelzés, illetve a megfigyelés szórását adja meg. A korrelációs együttható tulajdonképpen az előrejelzések potenciális jóságát tükrözi, mivel nem függ az előrejelzések kalibrációjától, azaz attól, hogy az előrejelzéseket milyen módszer alapján optimalizálták [34]. Hátránya, hogy a korrelációs együttható független az előrejelzések torzításától és varianciájától, ezért más, e tulajdonságokat mérni képes metrikákkal együtt kell használni [34].

A numerikus előrejelző modellekre, vagyis az AROME és az ECMWF vízszintes globálsugárzásaira is elvégeztem a verifikációt, ezáltal össze tudtam hasonlítani, hogy a különböző utófeldolgozási modellek milyen mértékben javítottak az előrejelzéseken.

7. Konklúziók és az eredmények értékelései

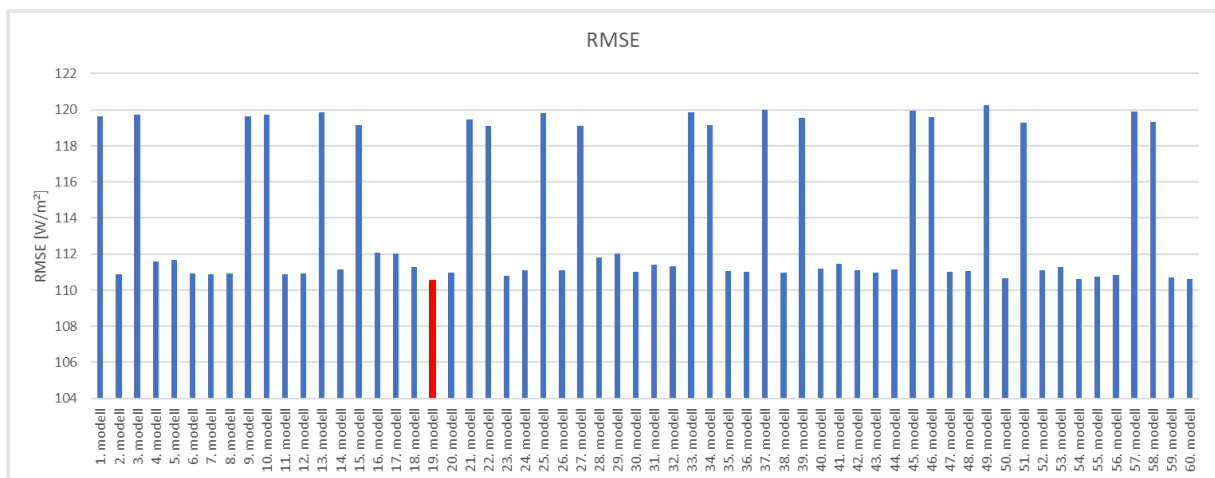
A lineáris regresszió esetében a hibára illesztés és a teljes sugárzásra illesztés között a hibaértékeket szemügyre véve nem tapasztaltam különbséget, így a többi módszer esetében a teljes sugárzásra illesztést alkalmaztam.

A lineáris regresszió módszer eredményei nagyban nem térnek el a nemlineáris módszerek eredményeitől, így megállapítható, hogy közel ugyanolyan alkalmas utófeldolgozásra, mint a komplexebbek módszerek.

A különböző modellek verifikációból származó eredményeit megvizsgálva azt tapasztaltam, hogy az értékek egybevágóak, így a következőkben az RMSE értékekre helyezem a fő hangsúlyt.

7.1. AROME modellek eredményei

Az AROME által létrejött modellek hibaértékeit elemezve kiválasztottam a legjobb modellt: az egyik neurális hálózat modell lett a legjobb, melynek prediktorai az AROME GHI, a zenitszög és a tiszta égbolt GHI. Az 1. diagramon látható, hogy ennek a modellnek (19. modell) a legkisebb az RMSE értéke.



1. diagram: Az AROME modellek RMSE értékei.

Az 1. diagramról az is egyértelműen lesűrhető, hogy bizonyos modellek jelentősen rosszabbul teljesítettek. Egyfajta trend figyelhető meg az RMSE értékeknél, a probléma okát pedig a prediktoroknál kell keresni. Azok a modellek rosszabbak, melyek prediktorai az AROME GHI vagy az AROME GHI, azimut szög vagy az AROME GHI, zenitszög koszinusza vagy az AROME GHI, azimut szög, zenitszög koszinusza.

A verifikáció eredményeinek átlagait is vettem, melyekből kiderült, hogy átlagosan az extrémely randomized trees módszer a legpontosabb. A prediktorokat illetően azok a modellek

bizonyultak jónak, melyekben az AROME GHI mellett a tiszta égbolt GHI is szerepelt a magyarázó változók között. A [34] cikkben is látszódik, hogy az AROME bias hibájának van egy délelőtt-délutáni jellegzetes eltérése, ezért lehet a tiszta égbolt GHI-nak jelentős hozzáadott értéke. Ezek az észrevételek a 2. és a 3. táblázatban is megfigyelhetők. Az is látható, hogy a modellek jóságát nem befolyásolta, hogy éppen a zenitszög vagy annak koszinusza szerepel a prediktorok között. Megállapítható, hogy vagy jobb vagy közel hasonló eredményt ad a zenitszög a zenitszög koszinuszához képest.

Módszerek	Variancia arány	Korrelációs együttható	MBE	MAE	RMSE
Lineáris regresszió	0,7975	0,9075	-1,8525	75,4820	113,9482
Neurális hálózat	0,7922	0,9077	-1,9082	74,8310	113,9635
Gradient-boosted trees	0,8206	0,9072	-0,7270	73,5606	114,0572
Random forest	0,7950	0,9075	-1,9679	74,3340	113,9996
Extremely randomized trees	0,7879	0,9080	-2,1349	74,5241	113,7785

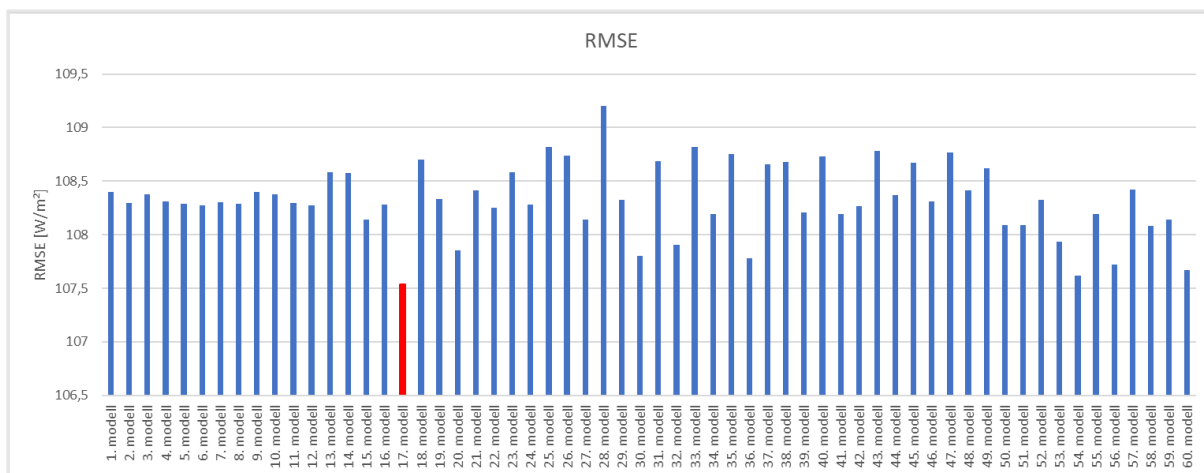
2. táblázat: Az AROME modellek verifikáció eredményeinek átlaga módszerek szerint csoportosítva.

Prediktorok	Variancia arány	Korrelációs együttható	MBE	MAE	RMSE
AROME GHI	0,7703	0,8974	-2,5924	78,2537	119,9043
AROME GHI, tiszta égbolt GHI	0,8067	0,9128	-1,1977	73,0855	110,9513
AROME GHI, azimut szög	0,7795	0,8982	-2,0147	77,1066	119,3584
AROME GHI, zenitszög	0,8049	0,9119	-4,0727	73,8701	111,5618
AROME GHI, azimut szög, zenitszög	0,8069	0,9115	-1,9582	73,2317	111,6957
AROME GHI, azimut szög, tiszta égbolt GHI	0,8092	0,9126	-1,2175	72,8378	110,9959
AROME GHI, zenit szög, tiszta égbolt GHI	0,8141	0,9127	-1,2410	72,8441	110,9127
AROME GHI, azimut szög, zenitszög, tiszta égbolt GHI	0,8123	0,9125	-1,4799	72,8502	111,0431
AROME GHI, zenitszög koszinusza	0,7810	0,8975	-1,8844	77,5810	119,7604
AROME GHI, azimut szög, zenitszög koszinusza	0,7793	0,8982	-0,7996	77,2125	119,3813
AROME GHI, zenitszög koszinusza, tiszta égbolt GHI	0,8066	0,9129	-1,6266	73,0778	110,8928
AROME GHI, azimut szög, zenitszög koszinusza, tiszta égbolt GHI	0,8132	0,9127	-0,5325	72,6051	110,9351

3. táblázat: Az AROME modellek verifikáció eredményeinek átlaga prediktorok szerint csoportosítva.

7.2. ECMWF modellek eredményei

Az ECMWF által létrejött modellek hibaértékeit is elemezve kiválasztottam a legjobb modellt: az egyik neurális hálózat modell lett a legjobb, melynek prediktorai az ECMWF GHI, az azimut szög és a zenitszög. A 2. diagramon látható, hogy ennek a modellnek (17. modell) a legkisebb az RMSE értéke.



2. diagram: Az ECMWF modellek RMSE értékei.

Ebben az esetben is vettem a verifikáció eredményeinek átlagait, melyekből kiderült, hogy átlagosan az extremely randomized trees módszer a legpontosabb, csak úgy, mint az AROME modelleknél. A prediktorokat tekintve azok a modellek bizonyultak jónak, melyekben az ECMWF GHI mellett az azimut szög is szerepelt a magyarázó változók között. Ezek a megállapítások a 4. és az 5. táblázatban is láthatóak. Ezeknél a modelleknél sem befolyásolta az előrejelzés jóságát az, hogy a zenitszög vagy annak koszinusza volt-e a prediktorok között.

Módszerek	Variancia arány	Korrelációs együttható	MBE	MAE	RMSE
Lineáris regresszió	0,8449	0,9169	1,5981	69,2474	108,3235
Neurális hálózat	0,8441	0,9172	4,6287	68,1737	108,2945
Gradient-boosted trees	0,8674	0,9170	4,6147	68,1513	108,4295
Random forest	0,8482	0,9166	1,8188	68,8028	108,5038
Extremely randomized trees	0,8392	0,9173	1,8758	68,8131	108,0755

4. táblázat: Az ECMWF modellek verifikáció eredményeinek átlaga módszerek szerint csoportosítva.

Prediktorok	Variancia arány	Korrelációs együttható	MBE	MAE	RMSE
ECMWF GHI	0,8387	0,9166	3,2005	68,9956	108,6165
ECMWF GHI, tiszta égbolt GHI	0,8506	0,9168	3,5273	68,6914	108,4741
ECMWF GHI, azimut szög	0,8479	0,9172	2,9991	68,2995	108,1927
ECMWF GHI, zenitszög	0,8503	0,9165	2,5485	68,9410	108,5687
ECMWF GHI, azimut szög, zenitszög	0,8520	0,9174	2,7819	68,4260	108,0559
ECMWF GHI, azimut szög, tiszta égbolt GHI	0,8514	0,9173	3,7072	68,4316	108,1316
ECMWF GHI, zenit szög, tiszta égbolt GHI	0,8521	0,9167	2,2920	68,9453	108,4605
ECMWF GHI, azimut szög, zenitszög, tiszta égbolt GHI	0,8572	0,9176	3,7767	68,2317	108,0262
ECMWF GHI, zenitszög koszinusza	0,8435	0,9166	1,9524	68,9275	108,5450
ECMWF GHI, azimut szög, zenitszög koszinusza	0,8448	0,9171	2,6822	68,4774	108,2425
ECMWF GHI, zenitszög koszinusza, tiszta égbolt GHI	0,8507	0,9167	3,7159	68,6330	108,5070
ECMWF GHI, azimut szög, zenitszög koszinusza, tiszta égbolt GHI	0,8460	0,9173	1,7031	68,6519	108,0838

5. táblázat: Az ECMWF modellek verifikáció eredményeinek átlaga prediktorok szerint csoportosítva.

7.3. Az AROME és az ECMWF modellek eredményeinek összehasonlítása

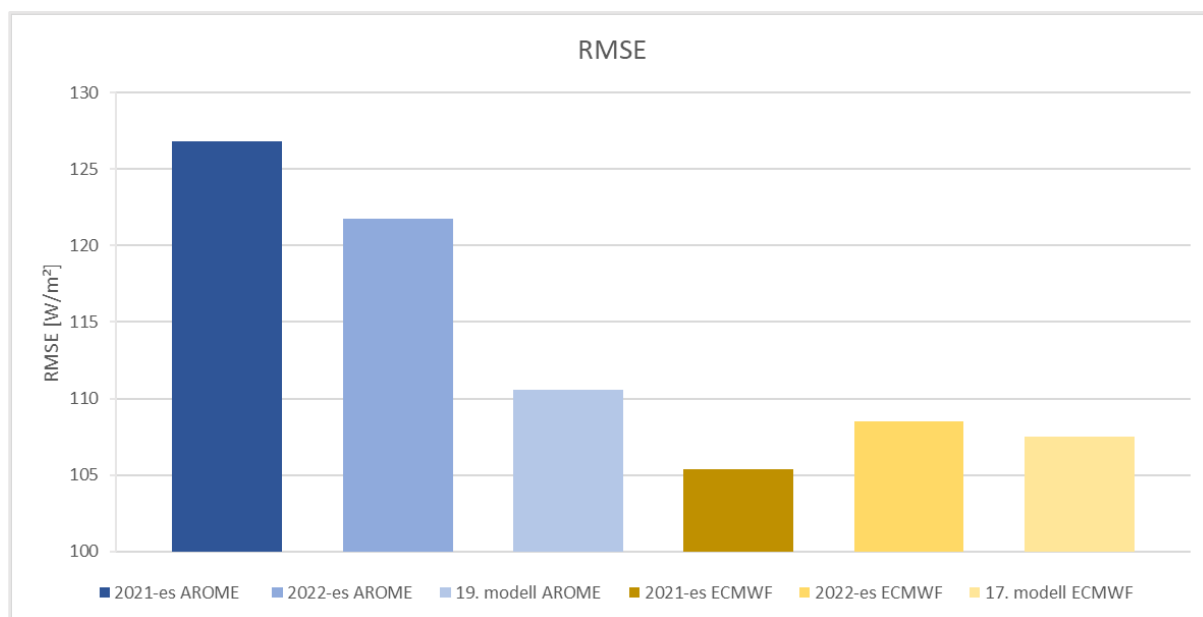
Az AROME és az ECMWF modellek verifikációból származó eredményeit is összehasonlítottam, elsősorban a korábban kiválasztott legjobb modelleket tekintve. Az AROME esetében jelentősebb javulást lehetett elérni az utófeldolgozással, például 96,84%-kal kisebb lett a legjobb AROME modell MBE értéke a 2022-es AROME adatok átlagos torzítási hibájához képest. A legjobb ECMWF modell 62,21%-os romlást mutatott a 2022-es ECMWF adatok MBE-jével összevetve. Ennek oka az esetleges túlillesztés lehet. Azonban fontos azt megjegyezni, hogy mivel kis értékekről van szó, így a %-os változás nagyobb léptékű lesz még akkor is, ha maga az eredmény még mindig jónak számít.

Bázisév	Modell	Variancia arány	Korrelációs együttható	MBE	MAE	RMSE
2021	19. modell AROME	0,84 %-os romlás	3,65 %-os javulás	96,84 %-os javulás	7,01 %-os javulás	12,84 %-os javulás
	17. modell ECMWF	5,79 %-os javulás	0,17 %-os javulás	62,21 %-os romlás	1,7 %-os romlás	2,05 %-os romlás
2022	19. modell AROME	0,64 %-os romlás	1,73 %-os javulás	96,98 %-os javulás	6,44 %-os javulás	9,23 %-os javulás
	17. modell ECMWF	5,25 %-os javulás	0,18 %-os javulás	8,8 %-os romlás	2,6 %-os javulás	0,9 %-os javulás

6. táblázat: A legjobb AROME és ECMWF modellek verifikáció eredményeinek %-os változása.

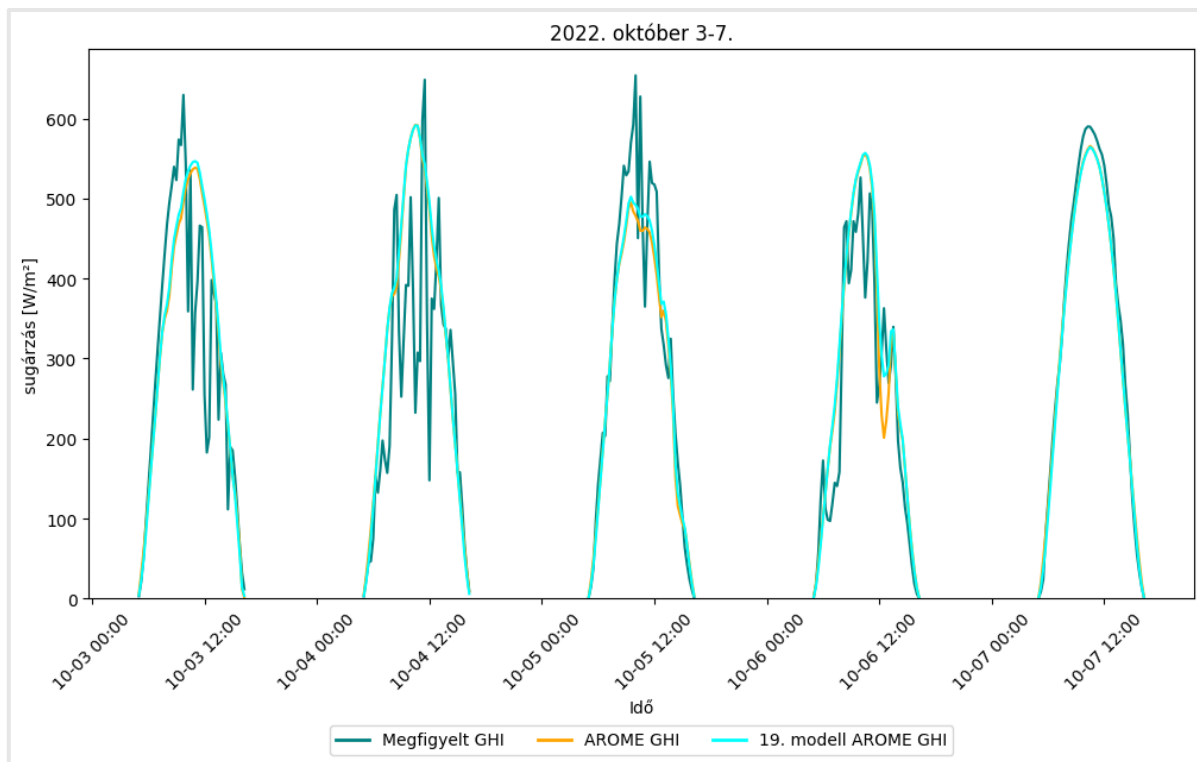
Összességében látható a 6. táblázatban, hogy a legjobb modellek majdnem minden értéken javítottak az utófeldolgozásnak köszönhetően.

A 3. diagramon látható, hogy a legjobb AROME modell látványos javulást eredményezett mind 2021-hez, mind pedig 2022-höz viszonyítva. A legjobb ECMWF modell csak a 2022 évhez képest tudott valamelyest javítani az RMSE értéken.

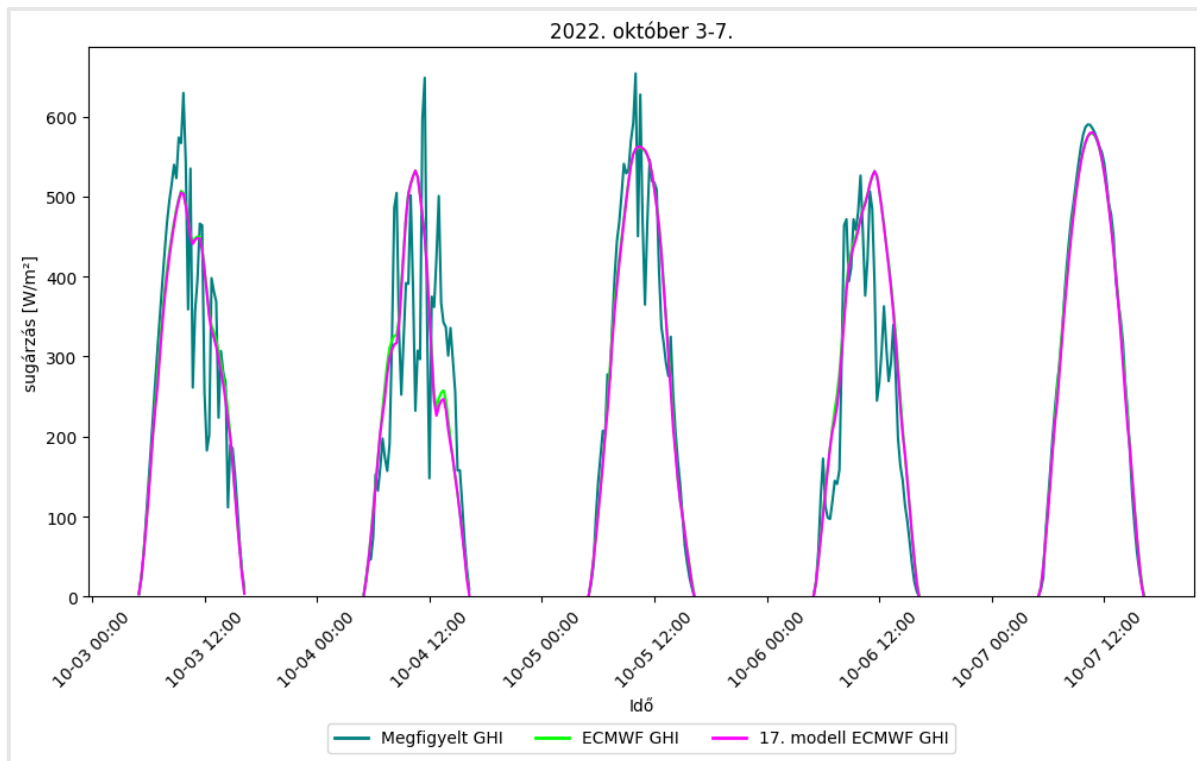


3. diagram: A legjobb AROME és ECMWF modellek, valamint a 2021-es és a 2022-es AROME és ECMWF numerikus időjárás előrejelző modellek RMSE értékei.

A verifikáció eredményeit megvizsgálva tapasztalható volt, hogy az AROME esetében az értékek nagyobb sávban szórnak, mint az ECMWF esetében, ahol általában csak kis mértékű változások voltak az utófeldolgozás következtében. Ennek a magyarázata az, hogy az ECMWF numerikus időjárás előrejelző modell jobban kalibrált, régebb óta fejlesztik, hogy az előrejelzések minél pontosabbak legyenek. Az ECMWF pontossága miatt lehetett nagyobb javulást elérni az AROME modellek esetében.



4. diagram: A megfigyelt, az AROME numerikus időjárás előrejelző modell és a legjobb AROME modell GHI értékei 2022. október 3-7. között.



5. diagram: A megfigyelt, az ECMWF numerikus időjárás előrejelző modell és a legjobb ECMWF modell GHI értékei 2022. október 3-7. között.

A 4. és 5. *diagramon* láthatóak a valós, a numerikus időjárás előrejelző modell és az általam legjobbnak ítélt modell sugárzás értékei 2022. október 3-7. között, mind az AROME, mind pedig az ECMWF esetében. Az AROME diagramon helyenként jobban kilátszódik a numerikus időjárás előrejelző modell görbéje. Ez is azt mutatja, hogy a legjobb AROME modell által pontosabb előrejelzések születtek, míg az ECMWF esetében jelentős javulást nem sikerült elérni.

8. Összefoglalás

A kutatásom során megvizsgáltam, hogy ez egyes gépi tanulási módszerek mennyire hatékonyak a globálsugárzás előrejelzések utófeldolgozásakor. Több fajta prediktort alkalmaztam különböző kombinációkban, ezáltal még több modellt tudtam létrehozni az adatok feldolgozásához. A verifikáció átlag eredményein alapulva, az AROME és az ECMWF adatok esetében is a legjobb módszernek az extremely randomized trees-t választottam. A prediktorokat tekintve az AROME modellek esetében a tiszta égbolt GHI, az ECMWF modellek esetében pedig az azimut szög eredményezett jobb előrejelzéseket. Az AROME modelleknél az abszolút legjobb modell a neurális hálózaton alapul, csak úgy, mint az ECMWF modellek esetében.

9. Hivatkozások

- [1] Európai Bizottság honlapja: *Környezet- és klímavédelem, A zöld megállapodás és Magyarország*, [Link](#), dátum: 2023. 11. 05.
- [2] Európai Bizottság honlapja: *Európai zöld megállapodás, Célunk, hogy Európa legyen a világ első klímasemleges kontinense*, [Link](#), dátum: 2023.11.05.
- [3] Európai Tanács honlapja: *Az európai zöld megállapodás*, [Link](#), dátum: 2023.11.05.
- [4] Európai Bizottság honlapja: *REPowerEU, Európa megfizethető, biztonságos és fenntartható energiaellátása*, [Link](#), dátum: 2023.11.05.
- [5] Európai Tanács honlapja: *Renewable energy: Council adopts new rules*, [Link](#), dátum: 2023.11.05.
- [6] DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL: *amending Directive (EU) 2018/2001, Regulation (EU) 2018/1999 and Directive 98/70/EC as regards the promotion of energy from renewable sources, and repealing Council Directive (EU) 2015/652*, [Link](#), dátum: 2023.11.05.
- [7] Magyarország Nemzeti Energia- és Klímaterve: *2023. évben felülvizsgált változat, Rövidített*, [Link](#), dátum: 2023.11.05.
- [8] MAVIR Zrt. és MEKH: *A MAGYAR VILLAMOSENERGIA-RENDSZER 2022. ÉVI ADATAI*, [Link](#), dátum: 2023.11.05.
- [9] MAVIR Zrt. (2023): *PV statisztika*, [Link](#), dátum: 2023.11.05.
- [10] Dobi Ildikó (2021): *NAPFÉNYTARTAM ÉS GLOBÁLSUGÁRZÁS SOKÉVI ÁTLAG TÉRKÉPEK MAGYARORSZÁGRA*, [Link](#), dátum: 2023.11.05.
- [11] OMSZ honlapja: *Magyarország napsugárzás, napfénytartam és felhőzet viszonyai*, [Link](#), dátum: 2023.11.05.
- [12] OMSZ honlapja: *Nagyfelbontású időjárás-előrejelző modellek*, [Link](#), dátum: 2023.11.05.
- [13] OMSZ honlapja: *ALADIN és AROME modellek*, [Link](#), dátum: 2023.11.05.
- [14] OMSZ honlapja: *Alaptevékenység (2018)*, [Link](#), dátum: 2023.11.05.
- [15] OMSZ honlapja: *ECMWF modellek alkalmazása*, [Link](#), dátum: 2023.11.05.
- [16] OMSZ honlapja: *Hírek a meteorológia világából, Sandy hurrikán előrejelezhetősége (2012)*, [Link](#), dátum: 2023.11.05.
- [17] D. Yang, D. van der Meer (2021): *Post-processing in solar forecasting: Ten overarching thinking tools*, [Link](#), dátum: 2023.11.05.
- [18] D. Yang, W. Wang, C. A. Gueymard, T. Hong, J. Kleissl, J. Huang, M. J. Perez, R. Perez, J. M. Bright, X. Xia, D. van der Meer, I. M. Peters (2022): *A review of solar forecasting, its dependence on atmospheric sciences and implications for grid integration: Towards carbon neutrality*, [Link](#), dátum: 2023.11.05.
- [19] M. J. Mayer, Gy. Gróf (2021): *Extensive comparison of physical models for photovoltaic power forecasting*, [Link](#), dátum: 2023.11.05.
- [20] C. Voyant, G. Notton, S. Kalogirou, M. Nivet, C. Paoli, F. Motte, A. Fouilloy (2017): *Machine learning methods for solar radiation forecasting: A review*, [Link](#), dátum: 2023.11.05.

- [21] D. Markovics, M. J. Mayer (2022): *Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction*, [Link](#), dátum: 2023.11.05.
- [22] GÉPI TANULÁS A GYAKORLATBAN (A SZEGEDI TUDOMÁNYEGYETEM, INFORMATIKAI KARÁNAK KURZUSA): *GÉPI TANULÁS ALAPFOGALMAI*, [Link](#), dátum: 2023.11.05.
- [23] SolarAnywhere honlapja: *Definitions*, [Link](#), dátum: 2023.11.05.
- [24] Fazekas István (2013): *Neurális hálózatok*, [Link](#), dátum: 2023.11.05.
- [25] scikit-learn: *Neural network models*, [Link](#), dátum: 2023.11.05.
- [26] scikit-learn: *Ensembles: Gradient boosting, random forests, bagging, voting, stacking*, [Link](#), dátum: 2023.11.05.
- [27] scikit-learn: *sklearn.ensemble.GradientBoostingRegressor*, [Link](#), dátum: 2023.11.05.
- [28] scikit-learn: *Decision Trees*, [Link](#), dátum: 2023.11.05.
- [29] scikit-learn: *Random forests and other randomized tree ensembles*, [Link](#), dátum: 2023.11.05.
- [30] javatpoint: *GBM in Machine Learning*, [Link](#), dátum: 2023.11.05.
- [31] javatpoint: *Random Forest Algorithm*, [Link](#), dátum: 2023.11.05.
- [32] OMSZ honlapja: *Előrejelzések értékelése*, [Link](#), dátum: 2023.11.05.
- [33] D. Yang, S. Alessandrini, J. Antonanzas, F. Antonanzas-Torres, V. Badescu, H. G. Beyer, R. Blaga, J. Boland, J. M. Bright, C. F.M. Coimbra, M. David, Â. Frimane, C. A. Gueymard, T. Hong, M. J. Kay, S. Killinger, J. Kleissl, P. Lauret, E. Lorenz, D. van der Meer, M. Paulescu, R. Perez, O. Perpiñán-Lamigueiro, I. M. Peters, G. Reikard, D. Renné, Y. Saint-Drenan, Y. Shuai, R. Urraca, H. Verbois, F. Vignola, C. Voyant, J. Zhang (2020): *Verification of deterministic solar forecasts*, [Link](#), dátum: 2023.11.05.
- [34] M. J. Mayer, D. Yang, B. Szintai (2023): *Comparing global and regional downscaled NWP models for irradiance and photovoltaic power forecasting: ECMWF versus AROME*, [Link](#), dátum: 2023.11.05.
- [35] scikit-learn oldala, [Link](#), dátum: 2023.11.05.
- [36] ECMWF honlapja: *Section 2.1.2.4 HRES - High Resolution Forecasts*, [Link](#), dátum: 2023.11.05.

10. Függelék

Modell	Módszer	Prediktorok
1. modell	Lineáris regresszió	NWP GHI
2. modell	Lineáris regresszió	NWP GHI, tiszta égbolt GHI
3. modell	Lineáris regresszió	NWP GHI, azimut szög
4. modell	Lineáris regresszió	NWP GHI, zenitszög
5. modell	Lineáris regresszió	NWP GHI, azimut szög, zenitszög
6. modell	Lineáris regresszió	NWP GHI, azimut szög, tiszta égbolt GHI
7. modell	Lineáris regresszió	NWP GHI, zenit szög, tiszta égbolt GHI
8. modell	Lineáris regresszió	NWP GHI, azimut szög, zenitszög, tiszta égbolt GHI
9. modell	Lineáris regresszió	NWP GHI, zenitszög koszinusza
10. modell	Lineáris regresszió	NWP GHI, azimut szög, zenitszög koszinusza
11. modell	Lineáris regresszió	NWP GHI, zenitszög koszinusza, tiszta égbolt GHI
12. modell	Lineáris regresszió	NWP GHI, azimut szög, zenitszög koszinusza, tiszta égbolt GHI
13. modell	Neurális hálózat	NWP GHI
14. modell	Neurális hálózat	NWP GHI, tiszta égbolt GHI
15. modell	Neurális hálózat	NWP GHI, azimut szög
16. modell	Neurális hálózat	NWP GHI, zenitszög
17. modell	Neurális hálózat	NWP GHI, azimut szög, zenitszög
18. modell	Neurális hálózat	NWP GHI, azimut szög, tiszta égbolt GHI
19. modell	Neurális hálózat	NWP GHI, zenit szög, tiszta égbolt GHI
20. modell	Neurális hálózat	NWP GHI, azimut szög, zenitszög, tiszta égbolt GHI
21. modell	Neurális hálózat	NWP GHI, zenitszög koszinusza
22. modell	Neurális hálózat	NWP GHI, azimut szög, zenitszög koszinusza
23. modell	Neurális hálózat	NWP GHI, zenitszög koszinusza, tiszta égbolt GHI
24. modell	Neurális hálózat	NWP GHI, azimut szög, zenitszög koszinusza, tiszta égbolt GHI
25. modell	Gradient-boosted trees	NWP GHI
26. modell	Gradient-boosted trees	NWP GHI, tiszta égbolt GHI
27. modell	Gradient-boosted trees	NWP GHI, azimut szög
28. modell	Gradient-boosted trees	NWP GHI, zenitszög
29. modell	Gradient-boosted trees	NWP GHI, azimut szög, zenitszög
30. modell	Gradient-boosted trees	NWP GHI, azimut szög, tiszta égbolt GHI
31. modell	Gradient-boosted trees	NWP GHI, zenit szög, tiszta égbolt GHI
32. modell	Gradient-boosted trees	NWP GHI, azimut szög, zenitszög, tiszta égbolt GHI
33. modell	Gradient-boosted trees	NWP GHI, zenitszög koszinusza
34. modell	Gradient-boosted trees	NWP GHI, azimut szög, zenitszög koszinusza
35. modell	Gradient-boosted trees	NWP GHI, zenitszög koszinusza, tiszta égbolt GHI
36. modell	Gradient-boosted trees	NWP GHI, azimut szög, zenitszög koszinusza, tiszta égbolt GHI

37. modell	Random forest	NWP GHI
38. modell	Random forest	NWP GHI, tiszta égbolt GHI
39. modell	Random forest	NWP GHI, azimut szög
40. modell	Random forest	NWP GHI, zenitszög
41. modell	Random forest	NWP GHI, azimut szög, zenitszög
42. modell	Random forest	NWP GHI, azimut szög, tiszta égbolt GHI
43. modell	Random forest	NWP GHI, zenit szög, tiszta égbolt GHI
44. modell	Random forest	NWP GHI, azimut szög, zenitszög, tiszta égbolt GHI
45. modell	Random forest	NWP GHI, zenitszög koszinusza
46. modell	Random forest	NWP GHI, azimut szög, zenitszög koszinusza
47. modell	Random forest	NWP GHI, zenitszög koszinusza, tiszta égbolt GHI
48. modell	Random forest	NWP GHI, azimut szög, zenitszög koszinusza, tiszta égbolt GHI
49. modell	Extremely randomized trees	NWP GHI
50. modell	Extremely randomized trees	NWP GHI, tiszta égbolt GHI
51. modell	Extremely randomized trees	NWP GHI, azimut szög
52. modell	Extremely randomized trees	NWP GHI, zenitszög
53. modell	Extremely randomized trees	NWP GHI, azimut szög, zenitszög
54. modell	Extremely randomized trees	NWP GHI, azimut szög, tiszta égbolt GHI
55. modell	Extremely randomized trees	NWP GHI, zenit szög, tiszta égbolt GHI
56. modell	Extremely randomized trees	NWP GHI, azimut szög, zenitszög, tiszta égbolt GHI
57. modell	Extremely randomized trees	NWP GHI, zenitszög koszinusza
58. modell	Extremely randomized trees	NWP GHI, azimut szög, zenitszög koszinusza
59. modell	Extremely randomized trees	NWP GHI, zenitszög koszinusza, tiszta égbolt GHI
60. modell	Extremely randomized trees	NWP GHI, azimut szög, zenitszög koszinusza, tiszta égbolt GHI

7. táblázat: A megalkotott modellek módszerei és prediktorai.