TDK-dolgozat

Günter Dalma Mária

Budapesti Műszaki és Gazdaságtudományi Egyetem
Villamosmérnöki és Informatikai Kar

Fotovillamos termelésbecslés támogatása mélyneurális háló és hagyományos képfeldolgozási módszerek kombinálásával

Tudományos Diákköri Konferencia

Intézmény: Budapesti Műszaki és Gazdaságtudományi Egyetem
Kar: Villamosmérnöki és Informatikai Kar
Tanszék: Automatizálási és Alkalmazott Informatikai Tanszék

Szerző:
Günter Dalma Mária
Évfolyam: V. évfolyam

Témavezetők:
Barancsuk Lilla, BME Villamos Energetika Tanszék, PhD hallgató
Oláhné Dr. Groma Veronika, HUN-REN Energiatudományi Kutatóközpont, tudományos munkatárs

Lezárás dátuma:
Budapest, 2023. november 1.

Kivonat

Magyarországon egyre elterjedtebbek az időjárásfüggő megújuló energiatermelők (elsősorban a fotovillamos rendszerek), melyek változékony teljesítményleadása nagyban megnehezíti a villamosenergia szolgáltatók számára a szolgáltatásminőség megfelelő szinten tartását, a szabványokban megjelölt határértékek megsértésének elkerülését.

A szolgáltatásminőség biztosításához mára elengedhetetlenné vált a fotovillamos rendszerek teljesítményleadásának pontos előrejelzése. A napelemek termelésének szempontjából a legfontosabb időjárási esemény a felhőzet változása, amely ultrarövid távon nagy megbízhatósággal előre jelezhető égboltkamera-felvételek alapján.

Az Energiatudományi Kutatóközpont (EK) Környezetfizikai Laboratórium egyik kutatási témájának keretében vizsgáljuk a napelemek termelésének ultrarövidtávú előrejelzését égboltkamera-felvételek alapján. Vizsgálataimhoz a kutatási projekt keretében kihelyezett széles látókörű, nagy felbontású égboltkamera felvételeit, valamint a kamera mellé telepített napelemek és kapcsolódó időjárás-monitoring rendszer által szolgáltatott adatokat használtam.

A projekt keretében több módszert is vizsgáltunk. Implementáltunk több hagyományos képfeldolgozási algoritmust, melyek kis számításigényűek, de nem kellően pontosak és nem tudják figyelembe venni a légköri állapotjelzőket. Kidolgoztunk továbbá egy olyan mesterséges neurális háló modellt, mely hatékonyan kombinálja az időjárás adatokból és az égboltkamera-képekből származó információkat. Ez a modell azonban nagyon erőforrásigényes és a háló tanítási igen hosszú időbe telik.

Ebben a dolgozatban bemutatom az általam kidolgozott mesterséges neurális háló alapú globálsugárzás becslő módszert, mely a hagyományos algoritmusok és az adatalapú megközelítések előnyeit kombinálja. A teljes égbolt képek használata helyett az égbolt felhőzetét egy, a hagyományos képfeldolgozó algoritmusok segítségével előállított paramétervektorral jellemzem. Ez lecsökkenti a feldolgozott adatmennyiséget és igen hatékonyá teszi a módszert, miközben továbbra is figyelembe veszi az égboltképekből származó legfőbb információkat. A jellemző vektor mellett a neurális háló az időjárás adatokat is bemeneti paraméterként feldolgozza.

A modellt pontosság és erőforrásigény szempontjából összehasonlítottam a teljes kameraképeket használó neurális háló módszerrel. A két modell átfogó összehasonlításához több meteorológiai adat-kombinációt állítottunk össze, melyek becslés pontosságára gyakorolt hatását mindkét modell esetében vizsgáltam.

Abstract

As the use of weather-dependent renewable energy sources, primarily photovoltaic systems, continues to grow in popularity in Hungary, the variable output of these systems presents significant challenges for distribution system operators. This variability makes it challenging to maintain service quality and avoid voltage limit violations.

Ensuring service quality heavily depends on accurate ultra-short-term forecasts of photovoltaic output. The primary factor affecting photovoltaic output is global irradiation, which is highly dependent on cloud cover. However, predicting cloud cover can be challenging due to its highly variable nature.

Our primary research focus at the Environmental Physics Department of the Center for Energy Research is the ultra-short-term prediction of global irradiation using all-sky camera imagery, as this approach is considered the most reliable and accurate. To achieve this, we use a high-resolution all-sky camera and a weather monitoring system.

To estimate global irradiation, multiple traditional image processing algorithms were developed, which are known for their computational efficiency but lacking in accuracy and the ability to account for atmospheric state variables. Additionally, we have developed an artificial neural network model that integrates information from both weather data and all-sky camera images. Although highly accurate, this neural network model is computationally expensive and necessitates extensive training.

This work introduces a novel model combining an artificial neural network and traditional image processing algorithms for estimating global irradiation to emphasize the strengths of both of these approaches. Instead of relying on all-sky images as an input to the neural network model, I characterize cloud cover using a parameter vector generated using traditional image processing. This approach significantly decreases computational demand, rendering the method highly efficient while still considering crucial information from the sky imagery. In addition to the feature vector, the neural network also utilizes weather data as an input.

We conducted a thorough evaluation of performance and resource utilization by comparing the combined model with the one utilizing all-sky camera images. We also assessed estimation accuracy analyzing different combinations of meteorological data to determine their impact on estimation accuracy for both models.

Tartalomjegyzék

1. Bevezetés **2**
1.1. Napsugárzás előrejelzési módszerek . . . 3
1.1.1. Napsugárzás előrejelzés Deep Learning módszerekkel . . . 3
1.2. A rövidtávú napsugárzás-előrejelzés hasznosítása a villamosenergia-rendszer üzemvitelében . . . 4
1.3. Előzmények . . . 4

2. Módszertan **7**
2.1. Égboltkamerás infrastruktúra az Energiatudományi Kutatóközpontban . . . 7
2.2. Az égboltképek színkorrekciója . . . 9
2.3. Adatok előkészítése . . . 9
2.4. Képjellemzők kinyerése hagyományos algoritmusokkal . . . 10
2.5. Időjárás szcenáriók . . . 14
2.6. Háló architektúra . . . 15
2.7. Hiperparaméter optimalizáció . . . 16

3. Eredmények **19**
3.1. A modellek pontossága . . . 19
3.2. A becslés pontossága különböző időjárási helyzetekben . . . 21
3.3. A módszer összevetése képfeldolgozó neurális hálóval . . . 21

4. Összefoglalás **23**
4.1. Kitekintés . . . 23

5. Köszönetnyilvánítás **24**

A. Képfeldolgozó algoritmusok optimális küszöbértékeinek beállítása **28**

1. fejezet

Bevezetés

Magyarország Nemzeti energia- és klímaterve [1], valamint az Európai Unió Megújulóenergia-irányelve [2] által ösztönözve, hazánk napenergia kapacitása jelentősen bővült az elmúlt néhány évben. A hazai ipari és háztartási méretű napelemes rendszerek összes beépített kapacitása idén átlépte az 5000 MW-ot [3]. A fotovillamos termelés drasztikus terjedése a teljes villamosenergia-rendszerre hatást gyakorol, rendszerszinten befolyásolja annak stabilitását, lokálisan pedig a feszültségviszonyokat és a teljesítményáramlást, veszélyeztetve a transzformátor túlterhelődését illetve a feszültséghatárérték megsértését. Ennek köszönhetően a fotovillamos termelés előrejelzése napjainkra gazdasági és stratégiai jelentőségű témává lépett elő.

A fotovillamos villamosenergia-termelés egyik legfőbb jellemzője a változékonyság, amely oka, hogy egyes légköri paraméterek – elsősorban az égbolt felhőzöttsége – közvetlen hatást gyakorolnak rá. A felhők rendszertelen és gyors mozgásának köszönhetően néhány másodperc leforgása alatt teljesen átalakulhat az égbolt képe, lényegesen megváltoztatva a fotovillamos termelést [4]. Ez a változékonyság jelentősen megnehezíti az időbeni beavatkozást a villamosenergia-rendszerbe és a termelés-fogyasztás egyensúlyának fenntartását.

A fotovillamos villamosenergia-termelés rövidtávú üzemelőkészítésének és menetrendezésének alapja a termelést közvetlenül meghatározó napsugárzás előrejelzése. Mivel a sugárzást a felhőmozgások és a légköri állapothatározók együttesen befolyásolják, ezért a legnagyobb pontosságú előrejelző modellek mindkét adattípust figyelembe veszik [5]. Ehhez azonban szükség van a légköri paraméterek és az égbolt felhőzöttségének folyamatos és pontos nyomon követésére, illetve ezen adatok közös előrejelző modellbe integrálására.

A felhőzet sugárzásra gyakorolt hatásának jellemzésére igen elterjedt módszer teljeségbolt-kamera felvételek elemzése különböző technikákkal, amelyek segítségével a felhők mozgása, kiterjedése, fényáteresztő- és fényvisszaverő képessége számszerűsíthető. Az elemzést gyakran különböző képfeldolgozásra specializált neurális háló modellek (például konvolúciós neurális hálózatok) végzik [6]. Ezeket további hálókkal összekapcsolva ún. többmódusú mélyneurális hálók hozhatók létre, amelyek többféle adattípus egyidejű feldolgozását teszik lehetővé [7], egy modellbe integrálva a rendelkezésre álló összes információt, mint az égboltkamera felvételek, légköri állapotjelzők, időbélyegek.

Bár a többmódusú neurális hálók igen nagy pontosságot érnek el a sugárzásbecslésben, ugyanakkor erőforrásigényük jelentős más módszerekhez képest [8]. Ez teszi szükségessé olyan erőforrástakarékos módszertanok kidolgozását, amelyek amellett, hogy lehetővé teszik különböző adatforrások integrációját, megfelelő becslési pontosságot biztosítanak.

Ebben a dolgozatban bemutatom az általam kidolgozott mesterséges neurális háló alapú sugárzásbecslő módszert, mely a hagyományos képfeldolgozó algoritmusok és az adatalapú megközelítés előnyeit kombinálja. A teljes égbolt képek használata helyett az égbolt felhőzetét egy, a hagyományos képfeldolgozó algoritmusok segítségével előállított paramétervektorral jellemzem. Ez lecsökkenti a feldolgozott adatmennyiséget és igen hatékonnyá teszi a módszert, miközben továbbra is figyelembe veszi az égboltképekből származó fontosabb információkat, amelyek a sugárzást befolyásolják. A neurális háló emellett a légköri

állapothatározókat és más, járulékos adatokat is felhasznál a becslés során. A modellt pontosság és erőforrásigény szempontjából összehasonlítottam a teljes kameraképeket használó neurális háló módszerrel. A két modell átfogó összehasonlításához több meteorológiai adat-kombinációt állítottunk össze, melyek becslés pontosságára gyakorolt hatását mindkét modell esetében vizsgáltam.

1.1. Napsugárzás előrejelzési módszerek

A globálsugárzás a napból érkező közvetlen és szórt (diffúz) sugárzás összessége [9]. Értéke arányos a fotovillamos termeléssel, így a technológiafüggetlen termelésbecslési eljárások általában globálsugárzás számszerűsítésére fókuszálnak [10].

A globálsugárzás becsléséhez számos ipari és tudományos megközelítés áll rendelkezésre, ideértve az összetett kereskedelmi rendszereket, amelyek felhőkamerák és szenzorok kombinációjából származó adatokat dolgoznak fel. Elérhetők továbbá olyan nagy területet lefedő szolgáltatások is, amelyek műholdfelvételek elemzésére támaszkodnak [11, 12].

A kereskedelemben elérhető rendszerek főként hagyományos képfeldolgozási, fizikai formalizáción alapuló vagy statisztikai idősor-elemzési módszereket alkalmaznak [13]. Azonban a tudományos közösségben egyre népszerűbbek a gépi tanulási modellek, mivel hatékonyan kezelik a szenzorok, kamerák és időjárási adatbázisok által generált nagy mennyiségű adatot [14].

1.1.1. Napsugárzás előrejelzés Deep Learning módszerekkel

Az adatalapú megközelítés számos előnnyel rendelkezik más, hagyományos módszerekkel szemben. Az operatív fizikai modellek ugyanis gyakran egyszerűsített formalizációkat alkalmaznak a felhőmozgás és a sugárzásváltozás kapcsolatának leírására [15], a hagyományos képfeldolgozási és statisztikai idősor-elemzési módszerek esetén pedig a különböző típusú adatforrások (képek, légköri állapotjelzők és járulékos adatok) összehangolása és kölcsönhatásaik feltárása jelenthet kihívást.

Ezzel szemben a mélytanulás alapú megközelítések eredendően alkalmasak többféle adattípus egyidejű feldolgozására. A feladatra általában teljesen összekapcsolt neurális hálózatokat használnak az időjárási adatok feldolgozására [16], az előrejelzéshez pedig az ún. Long Short-Term Memory (LSTM) hálózatokat alkalmazzák [17], amelyek a felhőmozgás rövid és hosszú távú időbeli összefüggéseinek mintázatait tárják fel.

A felsorolt architektúrák mellett a képfeldolgozó neurális hálók (például konvolúciós neurális hálók [6] vagy transformer hálók [18]) alkalmazása is igen népszerű megközelítés az égboltkamera-felvételek vagy műholdképek elemzésére, azonban ez nagy számítási kapacitású infrastruktúrát igényel, valamint igen sok tanító adatra van szükség a megfelelő pontosság eléréséhez. További hátrány, hogy maga a tanítás rendkívül hosszú időt vesz igénybe.

Ígéretesek azok a hibrid megoldások, amelyek előfeldolgozott képi adatokat használnak a neurális háló bemeneteként [7]. Ezeknél a megoldásoknál a háló nem a nyers képeket kapja bemeneti adatként, hanem olyan képjellemzőket (származtatott vagy kinyert képi adatot), amelyet egy másik, hagyományos képfeldolgozó algoritmus állít elő. Ennek eredményeként a kombinált módszerek összességében sokkal kevesebb adatmennyiséget igényelnek a tanítás során, drasztikusan csökkentve mind a tanítási időt, mind az erőforrásigényt. Bár a képeken tárolt információ egy része elveszik, ami negatívan befolyásolhatja a módszer pontosságát, ezt a hatást információban gazdagabb képjellemzők kiválasztásával csökkenthetjük. A [19] tanulmány szerzői több sugárzás-előrejelzési módszert vetettek össze, köztük a [7]-ben bemutatott hibrid modellel történő előrejelzést. A hibrid modell más módszerekkel összemérhető pontosságot ért el[1], miközben az erőforrásigénye igen alacsony. Ez

[1] A módszer Mean Average Percentage Error (MAPE) hibája 7,8 % volt, ami egy nagyságrendben van más, komplexebb módszerek hibájával, mint például a [20] tanulmányban publikált CNN-t és LSTM-et kombináló háló, amely 4 %-os RMAE hibeértéket ért el.

is rámutat, hogy a gépi tanulás és hagyományos képfeldolgozó módszerek kombinációja komoly előnyöket kínál a pontosság és erőforrásigény közötti egyensúly megtalálásában.

1.2. A rövidtávú napsugárzás-előrejelzés hasznosítása a villamosenergia-rendszer üzemvitelében

A fotovillamos termelők okozta problémák elsősorban a villamosenergia-rendszer elosztóhálózati részén jelentkeznek, ahova ezek az eszközök kapcsolódnak. Ez komoly műszaki kihívások elé állítja az eloszóhálózatok üzemeltetőit, hiszen a napelemek termelésingadozása miatt nő a feszültséghatárérték-sértések száma, a teljesítményáramlás megfordulhat, illetve jelentős túltermelés esetén az elosztóhálózati transzformátorok és vezetékek túlterhelődése is bekövetkezhet. A leadott teljesítmény hirtelen változásai miatt pedig a termelés–fogyasztás egyensúly fenntartása is nehézkessé válik.

Ezeknek a hálózati problémáknak megelőzésében kiemelt szerepet kap a napsugárzás ultrarövidtávú előrejelzése [21]. A várható sugárzás ismeretében több beavatkozási lehetőség is van az elosztóhálózat üzemvitelébe, így elkerülhetők a hirtelen letörések okozta termeléskiesések, például forgó tartalékok vagy energiatárolók vezérlésével, illetve túltermelés esetén a napelemek szükség szerinti leszabályozásával.

Az égboltkamerás rendszerek ultrarövidtávú termelés-előrejelzésre a kutatások szerint igen pontosak [21], továbbá az eljárást már a gyakorlatban is sikerrel használták az elosztóhálózati üzembiztonság növelésére. Az alábbiakban néhány olyan projektet mutatok be, amelyek mindegyikében égboltkamerás rendszer segítségével javították az elosztóhálózat energiaminőségét és csökkentették az időjárásfüggő megújuló energiatermelők okozta problémákat.

A [22] tanulmány szerzői egy napelemeket és dízelgenerátort tartalmazó rendszer támogatására alkalmaztak égboltkamerás előrejelző rendszert Nyugat-Ausztráliában. A dízelgenerátort az előrejelzés alapján vezérelték, hogy annak termelése minél jobban kiküszöbölje a felhőzet okozta termelésingadozásokat. A kísérlet eredményei alapján a módszer alkalmazásával 8 %-os üzemanyag-megtakarítás értek el, illetve biztonságosan növelhető volt a napelemek száma a hálózatban.

A „Karratha Airport Solar Project” [23] keretében a CloudCAM égboltkamerás előrejelző rendszert alkalmazták egy energiatároló vezérlésének optimalizálására, így a felőzet okozta termelésingadozások és termelésletörések megelőzhetőek voltak. A tanulmány szerint az előrejelzés használatával felére csökkenthető a szükséges akkumulátoros energiatároló kapacitása, ami jelentős anyagi megtakarítást jelent.

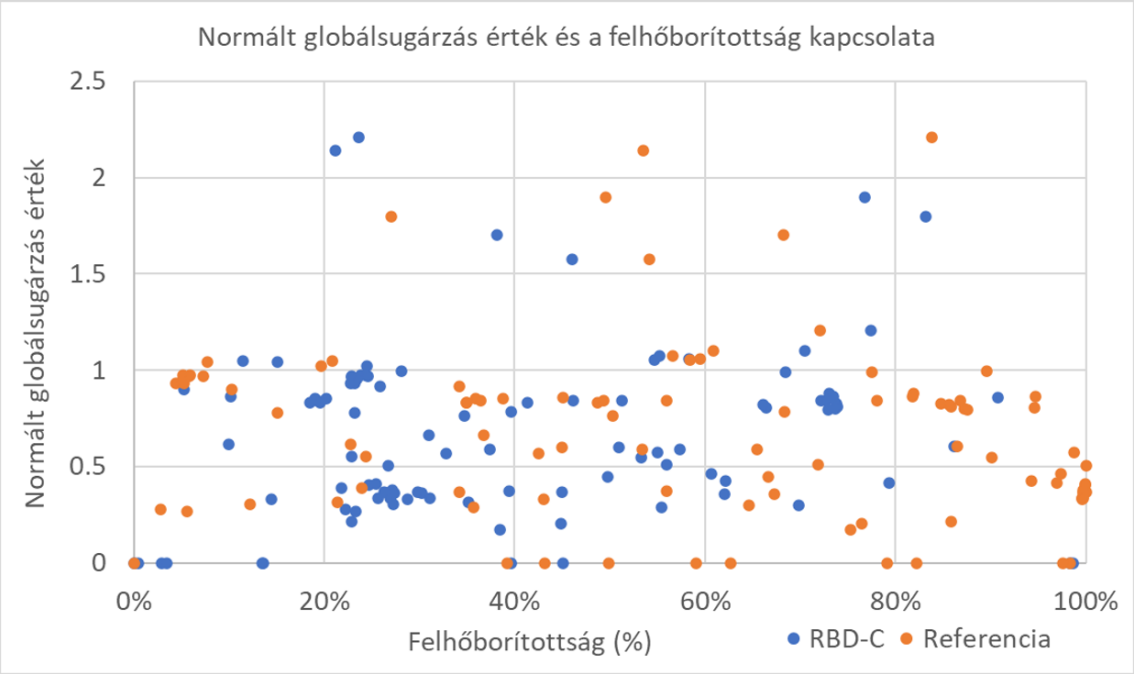
Egy spanyol projektben 50 MWp teljesítményű napkollektoros hőerőműpark teljesítményszabályzására 4 kamerából álló felhőkamera rendszert alkalmaztak [24]. A sugárzáselőrejelzés ismeretében a hőátviteli folyadék áramlását úgy szabályozták, hogy annak hőmérsékletét a kívánt értéken tartsák, ezzel növelve a naperőmű teljesítményét. A teljesítményszabályzás mellett az eszközök élettartama is növekedett, mivel a beérkező sugárzás hirtelen változásai által okozott hőterhelés csökkent. Továbbá az előrejelzéseket az intraday piaci részvétel javítására is felhasználták.

Az előző bemutatott projektek alátámasztják az égboltkamerák és a napsugárzás-előrejelzés potenciális előnyeit az elosztóhálózatok üzemeltetésében. Ugyanakkor mindegyik projekt rámutatott arra, hogy az égboltkamerás sugárzás-előrejelzés integrálása a villamosenergia-rendszerbe komoly technológiai kihívásokkal jár. Ezért a további technológiai fejlesztések elengedhetetlenek ahhoz, hogy az égboltkamerás rendszerek hatékonyan működjenek az elosztóhálózatokban.

1.3. Előzmények

Az előző évi Tudományos Diákköri Konferencia dolgozatomban [25] az égbolt felhőborítottságának hagyományos képfeldolgozó módszerekkel történő becslésére fókuszáltam. Az

égbolt százalékos felhőborítottsága alapján következtettem a globálsugárzás mértékére, ezáltal a napelemek termelésére. Munkámban több, a szakirodalomban fellelhető, szegmentálás alapú felhőfelismerő algoritmust implementáltam, optimalizáltam paramétereiket, majd a becslés pontosságát referencia adatokkal történő összehasonlítás segítségével változatos időjárási szcenáriókra numerikusan értékeltem. A detektált képeket vizsgálva arra a következtetésre jutottam, hogy az égbolt felhőzöttségének megállapítása önmagában kevés információ a globálsugárzás becsléséhez. Pontos eredmények előállításához komplexebb megközelítés szükséges, ezt támasztja alá az is, hogy az algoritmusok bizonyos időjárási körülmények között igen pontatlan eredményeket adnak. Az 1.1. ábráról leolvasható az is, hogy a referencia adatok alapján számolt felhőborítottság sincs szoros kapcsoltban a globsugárzással, illetve az algoritmusok további javításra szorulnak.



1.1. ábra. Normált globálsugárzás érték és a felhőborítottság kapcsolata

Az Energiatudományi Kutatóközpontban a kutatócsoporttal ezután kidolgoztunk egy neurális háló alapú módszert a globálsugárzás becslésére, mely az időjárás állomás által mért adatok és az égboltkamera által készített képek együttes felhasználásával végzi a becslést. A neurális hálózat használatával nagy mennyiségű időjárási adatot dolgoztunk fel, amely lehetővé tette olyan mintázatok felismerését, amelyekre a hagyományos módszerek alkalmazása nem volt hatékony, és amelyek emberi szemmel sem voltak észlelhetőek. Az időjárási adatok és a felhőkameraképek feldolgozásához többmódusú neurális hálót használtunk. A háló bemeneteként felhasználtuk a meteorológiai állomás által mért légköri adatokat, a felhőkameraképeket, valamint a képek és mérési adatok rögzítésének dátumát és időpontját is. Az időjárás adatokat egy teljesen összekapcsolt háló dolgozta fel, az égboltkameraképeket pedig egy, képfeldolgozásra specializált, úgynevezett Vision Transformer háló. Az két háló kimeneti adatait ezután egy második teljesen összekapcsolt hálóval egyesítjük. A háló kimenetét az időjárás állomás által mért globálsugárzás értékkel csatoltuk vissza; a legjobb modell pontossága 0,1 RRMSE[2] volt [26].

Kíváncsiak voltunk arra, hogy különböző időjárási adatkonfigurációk milyen mértékben befolyásolják a becslés pontosságát, ezért a felépített neurális hálózat hatékonyságát a rendelkezésre álló mérési adatokból szisztematikusan válogatott meteorológiai paraméterek alapján vizsgáltuk. E forgatókönyvek kidolgozásakor két kulcsfontosságú szempontot helyeztünk előtérbe: egyrészt a pontos globálsugárzás értékek becsléséhez nélkülözhetetlen

[2]RRMSE: Relativ Root Mean Square Error

paraméterek azonosítását, másrészt annak felmérését, hogy lehetséges-e kizárni bizonyos paramétereket, amelyek ismert fizikai törvényszerűségnek (például a légkör állapotegyenlete) köszönhetően többletinformációval nem rendelkeznek. Az időbélyeg és az égbolt kamera által szolgáltatott képek mellett a különböző meteorológiai paraméterek figyelembevételével szcenáriókat fogalmaztunk meg, melyek esetén vizsgáltuk a becslés pontosságát. Legmagasabb pontosság, amit elértünk $R^2 = 0{,}77$ volt [27].

2. fejezet

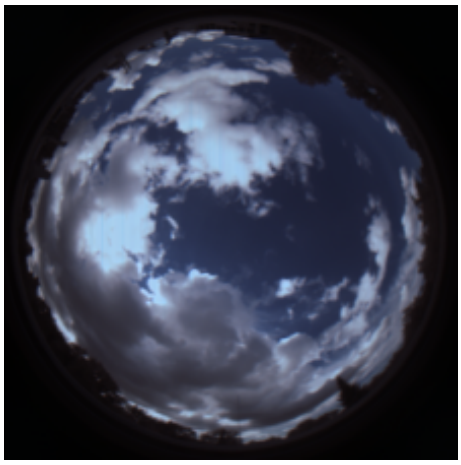
Módszertan

A korábbi kutatási fázisok során kiderült, hogy bár a képfeldolgozó neurális háló alapú módszerek ígéretesnek bizonyulnak, hátrányuk, hogy rendkívül erőforrásigényesek. Jelen TDK dolgozatomban egy mélyneurális háló modellt vizsgálok, mely bementeként kapja az időjárás állomás adatait valamint az égboltkameraképekből hagyományos algoritmusokkal kinyert égboltparamétereket. Ezzel jelentősen csökken a feldolgozott adatok mennyisége, valamint az alkalmazott háló architektúra komplexitása. Az időjárási adatokat a kutatás jelen fázisában is szcenáriókra bontottam. Az eredményeket összehasonlítom az égboltképeket közvetlenül feldolgozó neurális haló eredményeivel pontosság, futási idő és erőforrásigény szempontjából.

2.1. Égboltkamerás infrastruktúra az Energiatudományi Kutatóközpontban

Az Energiatudományi Kutatóközpont (EK) 29-es épületén található a napelemek és az időjárás-monitoring rendszer kombinációjából álló mérőállomás (2.1a. ábra). A monitoring rendszer részét képezi egy Starlight Xpress Oculus típusú nagy felbontású, 180°-os látószögű színes égboltkamera, amelynek tengelye függőlegesen felfelé néz. A kamera percenként felvételt készít az égboltról (képe a 2.1b. ábrán látható). A kamera vezérléséhez saját fejlesztésű szoftver készült, amely alkalmazkodik az adatfeldolgozás egyedi követelményeihez, lehetőséget nyújtva a fényképezés gyakoriságának és az expozíciós idő tetszőleges beállítására. A szoftver 2021 novembere óta a nappali időszakban folyamatosan rögzíti az égboltfelvételeket.

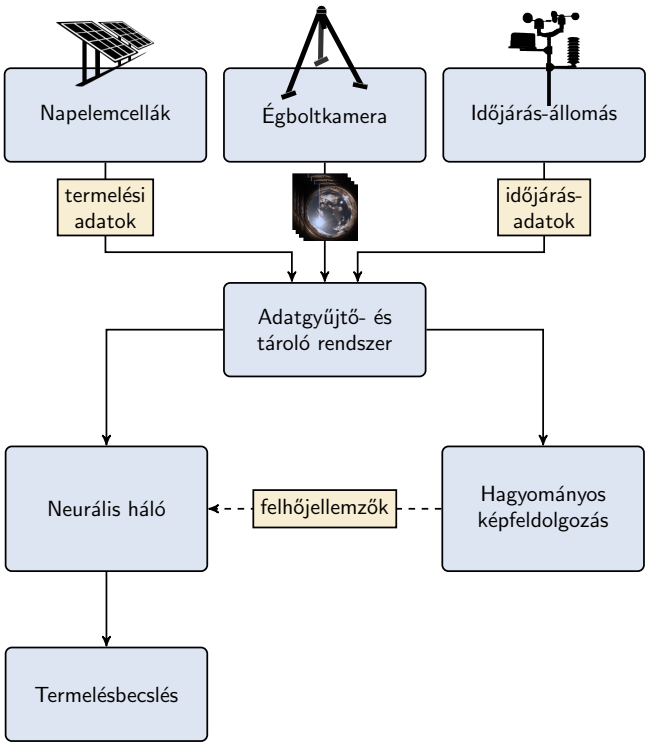
(a) Starlight Xpress Oculus típusú égbolt kamera az EK telephelyén.



(b) A nagy látószögű kamera képe. A kamera a teljes égbotról készít felvételt.

2.1. ábra. Égboltkamerás infrastruktúra, és a kamera által készített kép. A nagy látószögű kamera a teljes égboltról készít felvételt.

A kamera közvetlen szomszédságában elhelyezett időjárás-monitoring rendszer számos meteorológiai adatot mér. Ötmásodperces gyakorisággal méri a levegő hőmérsékletét, a földfelszíni szélsebességet, az atmoszférikus nyomást, a légkör páratartalmát, valamint a napelemek termelését közvetlenül meghatározó globálsugárzást, majd rögzíti ezek egyperces átlagértékét, illetve azok egy percre vonatkoztatott szórását. A képeket és adatokat egy-egy adatbázisban tároljuk további feldolgozásra. A rendszer felépítése a 2.2. ábrán látható. Az összegyűjtött adatok a mélyneurális hálós feldolgozás mellett széleskörűen felhasználhatóak további elemzésekhez is.



2.2. ábra. Az időjárás monitoring rendszer felépítése és az adatfeldolgozás folyamata

2.2. Az égboltképek színkorrekciója

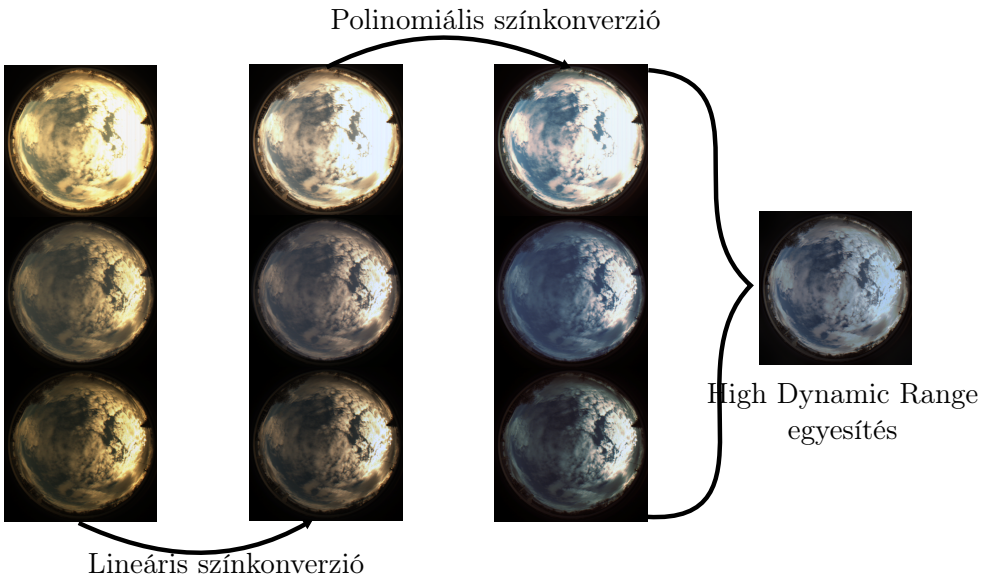
A kamerából érkező nyers képek árnyalatukban igen eltérnek a megszokott égboltképektől, amelynek oka a kamera színérzékeny felületének beállítása, ami sárgás alapszínt kölcsönöz a képeknek. Ez hozzájárult ahhoz, hogy az általunk alkalmazott algoritmusok igen pontatlanok voltak a nyers képeken. Annak érdekében, hogy a hagyományos színcsatornaküszöbözésen alapuló algoritmusok megfelelően működjenek, a képeket több lépésből álló színkonverziónak vetjük alá. Minden égboltképet három különböző záridővel készített felvételből állítunk össze az ún. High Dynamic Range (HDR) egyesítéssel. A HDR egyesítés növeli a kontrasztot, és csökkenti a képeken a nap körüli régióban jelentkező színtelítődést is. Az előfeldolgozás lépéseit a 2.3. ábra szemlélteti. A feldolgozás lépések az alábbiak:

1. Adott pixel színének kiszámítása demozaik eljárással (másnéven De-Bayer interpolációval) az RGB színtérben.

2. Lineáris színkonverzió színcsatornánként. A kék színcsatorna felskálázása (skálatényező: 1,2), a piros színcsatorna leskálázása (skálatényező: 0,95).

3. Polinomiális színkonverzió csatornánként. Empirikusan megállapítottunk egy konverziós polinomot, ami alapján a piros, zöld és kék színcsatornákat világosságfüggő konverziónak vetjük alá [28].

4. A különböző záridővel készült képek HDR egyesítése az ún. Mertens algoritmussal [29]. A művelethez az OpenCV HDRI MergeMertens algoritmusát [30] használjuk.



2.3. ábra. Kameraképek előfeldolgozási lépései. Minden égboltképet három különböző záridővel készített nyers kamerafelvételből állítunk össze, amelyeket először színkonverziónak vetünk alá.

2.3. Adatok előkészítése

A sugárzásbecslést az égboltkamera által készített képekből nyert égboltparaméter-adatok, az időjárás állomás által mért adatok, valamint további járulékos adatok alapján végeztük. Járulékos adatként a tiszta (felhőtlen) égbolthoz adott földrajzi helyszínre vonatkozó sugárzásértékeket, a mérési adatok időbélyegeit, valamint a nap zenit szögét használtuk. A kapott eredményeket összehasonlítottam az égboltkamera képeket használó háló eredményeivel. Az adatokat 2021 novembere és 2023 októbere között rögzített mérésekből állítottuk össze, hozzávetőlegesen 300 000 egyperces mintapont áll rendelkezésünkre.

A mérési adatok feldolgozása során az időjárás esetenkénti gyors változása problémát okozott, ugyanis a perces felbontású méréseknél észrevehető elcsúszást lehet felfedezni a perc elején készített kép és a perc végén rögzített sugárzásérték között. Az eltérések korrigálására a képeken és az időjárási adatokon időbeni illesztést végeztünk, valamint az úgynevezett Clear Sky adatsort is időben illesztettük a mintákhoz, mely során a téli-nyári időszámítást is figyelembe vettük. A Clear Sky sugárzásérték a felhőtlen, tiszta égbolt esetére modellezett pillanatnyi, adott földrajzi helyszínre vonatkozó globálsugárzás érték, amely ideális körülmények között meghatározza a fotovillamos termelést. Ehhez egyperces felbontású McClear típusú Clear Sky adatsort használtunk [31].

A tanító adathalmazhoz minden adatpont esetén hozzáadtuk az úgy nevezett zenit szöget, mely a Naphoz húzott egyenes által bezárt szög, vagyis a vízszintes felületre érkező sugárzás beesési szöge. Az adatot a Nap helyzetének a Föld felszínén egy adott helyről történő meghatározására alkalmazzák [32].

A hibás mérési adatokat, például az eredményeket torzító negatív páratartalom értékeket eltávolítottuk, illetve szűrtük a kiugró merési eredményeket is (alsó és felső 5%) az adatok közül.

Az adatsornak csak azt a részét használtuk fel, amely kellően világos időszakra korlátozódott, ahol ténylegesen történik villamosenergia-termelés. Ehhez az időjárás állomás szürkületérzékelőjének segítségével távolítottuk el a nem megfelelő adatpontokat. Hozzáadtuk az adatsorhoz továbbá a rögzítés dátumát és időpontját is. Ezeket a körkörös reprezentációra alakítottuk át, hogy a hálózat könnyebben felismerhesse az azonos időszakokat különböző napokon és az év során.

A tanító adathalmaz kiegyenlítettsége érdekében a különböző globálsugárzás értékeket közel azonos gyakorisággal szerepeltettük a tanító adatok között. Az így kapott adathalmazt 0,8:0,2 arányban tanító és validáló adatsorokra választottuk szét. A tanító adatok értékeit átskáláztuk, hogy 0 középértékű, 1 szórású eloszlást kapjunk minden időjárási adat esetén. A validáló adatsort a tanító adatoknak megfelelően skáláztuk át.

2.4. Képjellemzők kinyerése hagyományos algoritmusokkal

Az égboltképeken szereplő felhőzet paramétereinek kinyéréséhez a tavalyi TDK munkám során implementált, szegmentálás alapú felhőfelismerő algoritmusok továbbfejlesztett változatát használtam [25]. Az algoritmusok pixelszinten különítik el a felhőzethez és az égbolthoz tartozó részeket a képeken. 5 hagyományos algoritmus állt a rendelkezésemre: Vörös–kék különbség (Red–Blue Difference), Vörös–kék arány (Red–Blue Ratio), Fehérségdetektor (Whiteness Detector) és Fehérségdetektor átlagos intenzitással (Whiteness Detector With Average Intensity) és Hibrid Küszöbözés algoritmus (Hybrid Thresholding) [33]. Az algoritmusok küszöbértékeit egy 120 égboltképből álló referencia-adatbázis segítségével optimalizáltam. Az optimális küszöb beállításának folyamatát az A. függelékben foglaltam össze.

Az egyes algoritmusok pontosságát széleskörben használatos metrikák segítségével állapítottam meg. Definiáltam az Accuracy (Pontosság) és Recall mértékeket [34, 35]. Az Accuracy (ACC) vagyis pontosság értéke meghatározza, hogy egy adott mérési halmaz mennyire áll közel a valódi értékéhez. Esetemben ez azt jelenti, hogy a képen a pixeleket milyen arányban detektálta helyesen az algoritmus az égbolt egészére nézve. A Recall (REC) értéke meghatározza, hogy egy adott mintaterület mérési adatai mennyire állnak közel a valódi értékükhöz. Esetemben a recall értéke a képen helyesen detektált felhőpixelek arányát mutatja meg.

A 2.1 táblázatban bemutatom az algoritmusok pontosságát a korrigált képeken, összehasonlítva a tavalyi évben elért eredményekkel. Látható, hogy a színkorrekción átesett képeken sokkal jobb eredményeket érnek el az algoritmusok.

A globálsugárzás becsléshez felhasznált égboltkamera képeken lefuttattuk a legpontosabb felhőzetbecslő Red–Blue Ratio algoritmust és előállítottuk a neurális háló számá-

	2023 – Korrekció		2022	
Algoritmus	Accuracy [%]	Recall [%]	Accuracy [%]	Recall [%]
Hybrid Thresholding	67	98	–	–
Whiteness Detector	85	85	51	52
Whiteness Detector With Average Intensity	66	87	73	36
Red–Blue Ratio	**85**	**85**	52	53
Red–Blue Difference	91	73	68	71

2.1. táblázat. Algoritmusok pontossága az eredeti és a színkorrigált képeken.

ra szükséges égboltparaméteket. Mivel az egyes felhőtípusok fényáteresztőképessége és fényszórása nagy mértékben eltér egymástól, illetve a hagyományos felhőosztályozás nem alkalmas a sugárzásmódosító hatás pontos becslésére hiszen a klasszikus felhőfajták sugárzásmódosító hatása is széles tartományban mozog [36, 37], ezért olyan felhőparaméterek definiálására fókuszáltunk, amelyek sugárzásmódosító hatás tekintetében a hasonló viselkedésű felhőtípusokat jól jellemzik. Minden paramétert az égboltképek közepén található, kerek régióra számoltunk ki, a környező tereptárgyakat és a kép fekete szélét nem vettük figyelembe a számítások során.
Az általunk használt égboltparaméterek a következők:

- Felhőborítottság mértéke,

  mely megadja a kamera által készített képen az égbolt felhővel borított pixeleinek és teljes kép pixeleinek arányát. A globálsugárzás egyértelmű kapcsolatban áll ezzel az adattal, azonban a sugárzás becsléséhez további paraméterek is szükségesek. A hagyományos okta alapú felhőborítottság mérőszámának kiterjesztése [38].

- Legnagyobb kiterjedésű tiszta égbolt aránya,

  mely megadja a kameraképen a legnagyobb kiterjedésű, egybefüggő tiszta égbolt pixelszámát a teljes kép pixelszámához képest. A tiszta égbolt mérete a felhőátvonulásokról és a sugárzás mértékéről is hordoz információt.

- Felhők darabszáma,

  az összefüggő felhőrégiók száma a szegmentált képen, amely az egyes felhőfajták megkülönböztetéséhez szükséges egyik legfontosabb információ. Emellett fontos mutatója az sugárzásértékek időbeli variabilitásának.

- Felhők inhomogenitása,

  melynek értéke az egyes felhőrégiókban található sötét pixelek aránya a teljes felhőrégióhoz képest; szintén a felhő típusát jellemző adat.

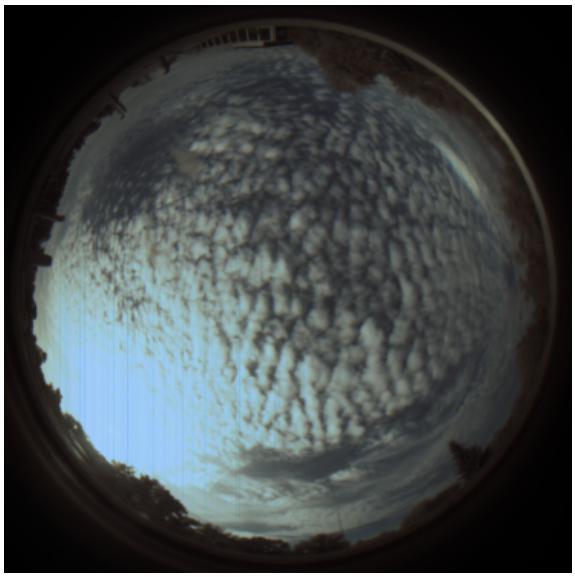
- Periodicitás mértéke,

  a periodicitást három mutatóval jellemezzük, amelyek mindegyike az egyes felhőelemek térbeli elhelyezkedésének periodikusságát jellemzi.

  - Alapharmonikus frekvenciája, mely a 2D Fourier sorfejtés során kapott alapharmonikus frekvenciája.
  - Alapharmonikus amplitúdója, mely a 2D Fourier sorfejtés során kapott alapharmonikus amplitúdója.
  - Periodikus komponensek energiája, mely egy empirikus úton meghatározott frekvenciatartományban található harmonikus komponensek energiája a teljes kép energiájához képest. Ez a mutató a periodikus altocumulus felhőket diszkriminálja (egy tipikus altocumulus felhőzet a 2.4. ábrán látható).

- Átlagos pixelintenzitás

  mely megadja a teljes kép pixeleinek átlagintenzitását.



2.4. ábra. Altocumulusok okozta periodikus felhőkép.

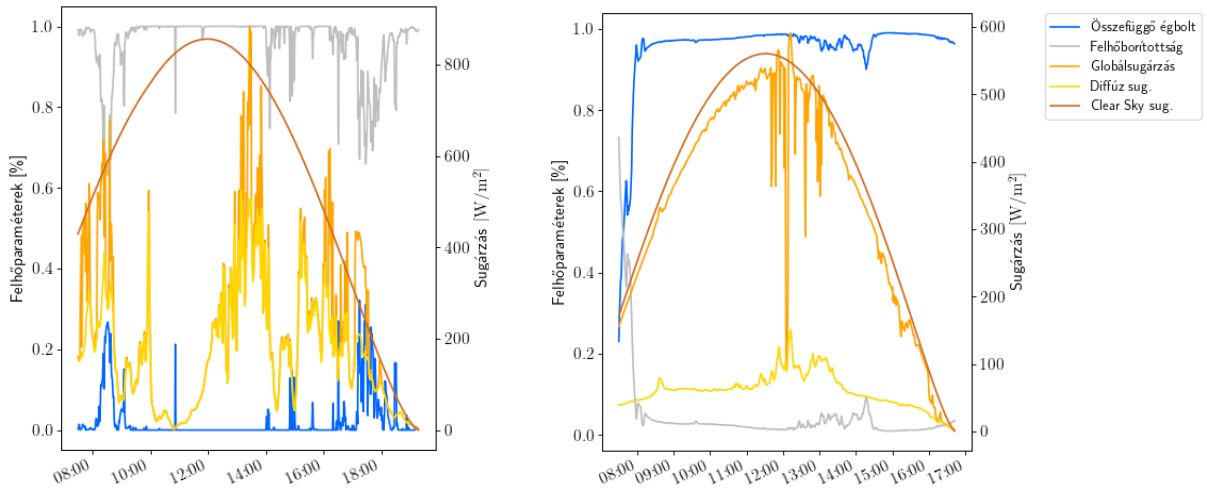
Az képekből kinyert paraméterekkel együtt vizsgáltuk a globálsugárzás, és a szórt (diffúz) sugárzás értékét. A diffúz a napból érkező sugárzás szórt része, amely a felhőkön, vagy a levegő részecskéin szóródik vagy visszaverődik. Ez a sugárzáskomponens irányában és/vagy intenzitásában különbözik a Napból közvetlenül érkező direkt sugárzástól, azonban a fotovillamos termelésben jelentős szerepet játszik [39]. A diffúz sugárzást ezért ajánlott a globálsugárzástól elkülönülten vizsgálni a fotovillamos termelésbecslő alkalmazásokban.

A 2.5a és 2.5b ábrákról leolvashatók egy felhős és egy derűs nap esetén a mért globál és diffúz sugárzás értékek, valamint a hagyományos algoritmusokkal az égboltképekből kinyert felhőborítottság és tiszta égbolt mértéke. A hagyományos algoritmusok megfelelő működésére lehet következtetni abból, hogy magas felhőzöttség mellett a tiszta égbolt mértéke alacsony, illetve ez fordított esetben is leolvasható az ábráról.

A 2.5a ábrán látható, hogy a pusztán a felhőborítottság mértékéből nem lehet következtetni a globálsugrázásra, azonban az egyes sugárzáskomponensek és a felhőborítottság között felfedezhető kapcsolat. A diffúz sugárzás jóval magasabb a felhős napon (2.5b. ábra), mint a többnyire derült égbolttal jellemezhető napon (2.5a. ábra), amelyből látható, hogy a diffúz sugárzás jelentős részét a felhőkről visszavert komponens adja. Derült napon a globálsugárzás értéke jól követi az ábrázolt Clear Sky görbét, a délutáni órákban észlelhető felhőátvonuláskor pedig láthatóan csökkent az értéke, ezzel együtt a diffúz komponens értéke nő.

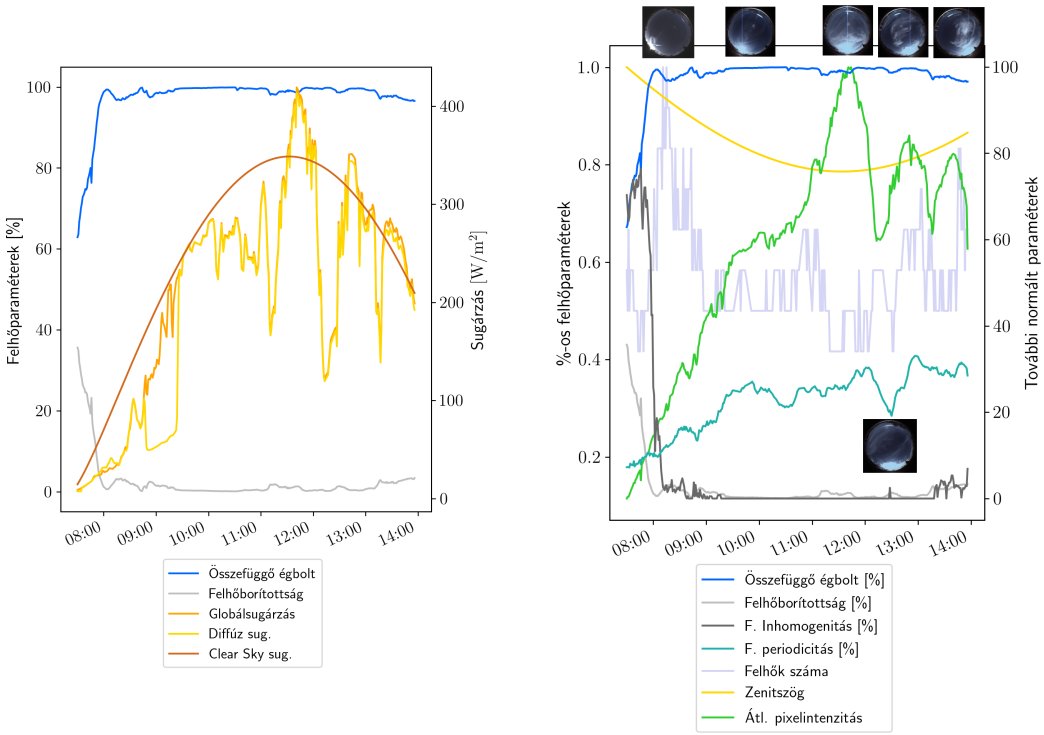
Ehhez hasonlóan a 2.5. ábráról leolvasható, hogy egy fátyolfelhős derült napos a sugárzás nagy részét a diffúz komponens adja. A felhők darabszámának, elhelyezkedésének, vastagságának változásával a sugárzás mértéke is változik a nap folyamán, annak ellenére, hogy a teljes felhőborítottság nem módosul érdemben.

Mindezekből arra lehet következtetni, hogy a különböző felhőtípusok sugárzásra gyakorolt befolyásoló hatása jelentős, a sugárzásbecslés során érdemes a különböző sugárzási komponensek értékelésére külön-külön is hangsúlyt fektetni.



(a) Felhős égbolt. (b) Tiszta égbolt cumulusokkal.

2.5. ábra. A két sugárzáskomponens (globálsugárzás, diffúz sugárzás) alakulása, valamint a felhőzöttség mértékének változása az égboltképekből kinyert paraméterek (felhőzöttség mértéke, legnagyobb összefüggő égbolt) alapján.



(a) A sugárzáskomponensek értékei a nap során. (b) Az egyes felhőparaméterek alakulása.

2.6. ábra. A különböző képjellemző paraméterek napi menete 2022. december 1-jén délelőtt, és a felhőzet változását jellemző felhőképek. A nap folyamán végig derült idő volt, néhány felhővel. A kép átlagintenzitása a délelőtt folyamán fokozatosan nő párhuzamosan a nap emelkedésével. A 12 órakor megjelenő felhőzet a visszavert sugárzás miatt jelentősen megnöveli a kép átlagintenzitását. 12:30-kor a Nap elé bekúszó cumulus hirtelen lecsökkenti a felszínre jutó sugárzást annak ellenére, hogy az égbolt felhőzöttsége ekkor igen alacsony. A délután folyamán a fátyolfelhők megjelenése miatt változó a felhők darabszáma és a változik a kép intenzitása is. A felhők textúráltságát jellemző periodicitás paraméter az égbolt felhőzöttségével arányos.

2.5. Időjárás szcenáriók

A kutatás során célunk az egyes sugárzáskomponensek minél pontosabb becslése volt, melyhez az EK-ban kihelyezett időjárás állomás által mért időjárási adatokat és a felhőkamera képekből kinyert égboltparamétereket használtunk fel.
A mérőállomás által mért, rendelkezésünkre álló adatok a következőek:

- Vízszintes felszínen mért globálsugárzás [W/m^2] – visszacsatolt adat
- Vízszintes felszínen mért diffúz sugárzás [W/m^2] – visszacsatolt adat
- Vízszintes felszínen mért direkt sugárzás [W/m^2] – nem használtuk fel
- Hőmérséklet [°C]
- Légnyomás [mBar]
- Páratartalom [%]
- Földfelszíni szélsebesség [km/h]

A neurális hálózat a 2.4 fejezetben meghatározott égboltparamétereket minden esetben bemeneti adatként megkapta, az időjárás állomás által mért adatokat viszont szisztematikusan válogattuk és különböző szcenáriókat fogalmaztunk meg. E forgatókönyvek kidolgozásakor két kulcsfontosságú szempontot helyeztünk előtérbe: egyrészt a pontos sugárzásértékek becsléséhez nélkülözhetetlen paraméterek azonosítását, másrészt annak felmérését, hogy lehetséges-e kizárni bizonyos paramétereket, amelyek ismert fizikai törvényszerűségnek (például a légkör állapotegyenlete) köszönhetően többletinformációval nem rendelkeznek. Az időbélyeg és az égbolt kamera által szolgáltatott képek mellett a különböző meteorológiai paraméterek figyelembevételével öt különböző szcenáriót fogalmaztunk meg.

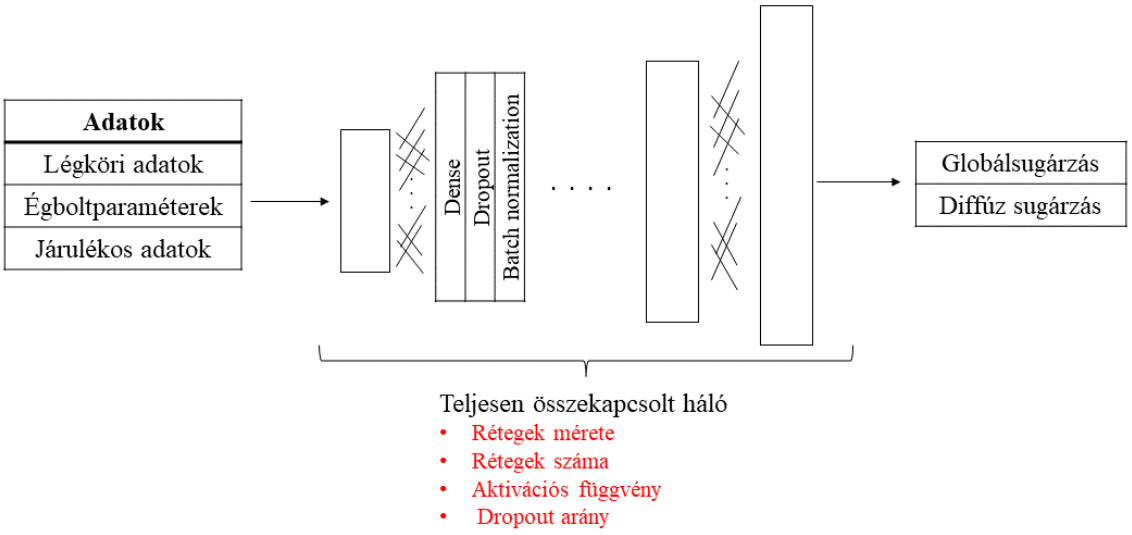
A 2.2 táblázatban láthatóak a megfogalmazott időjárási szcenáriók. A különböző szcenáriókhoz más-más adatkombinációk tartoznak, az 1-es szcenárióban szórásértékekkel is dolgozik a háló. A Clear Sky adatsort csak az első két szcenárióban kapja meg bemenetként. A Nap zenit szögét és az égboltparamétereket minden szcenárióban átadjuk bemenetként csakúgy, mint a mérési adatok időbélyegét.

Szcenárió	Felhasznált egyperces átlagértékek	Felhasznált szórásértékek	Clear Sky adatsor	Égbolt-paraméterek	Zenit szög
1. Teljes adatkészlet szórással	Hőmérséklet, légnyomás, páratartalom, szélsebesség	Hőmérséklet, légnyomás, páratartalom, szélsebesség	Igen	Igen	Igen
2. Teljes adatkészlet szórás nélkül	Hőmérséklet, légnyomás, páratartalom, szélsebesség	–	Igen	Igen	Igen
3. Teljes adatkészlet Clear Sky adat nélkül	Hőmérséklet, légnyomás, páratartalom, szélsebesség	–	Nem	Igen	Igen
4. Összefüggő légköri állapotjelzők	Hőmérséklet, légnyomás, páratartalom	–	Nem	Igen	Igen
5. Független légköri állapotjelzők	Hőmérséklet, légnyomás	–	Nem	Igen	Igen

2.2. táblázat. Időjárási szcenáriók

2.6. Háló architektúra

A becslés kivitelezéséhez egy olyan mélyneurális háló modellt hoztunk létre, mely az időjárási adatsort és a kinyert képjellemzőket egy teljesen összekapcsolt hálózattal dolgozza fel. A háló vázlatos felépítése a 2.7. ábrán látható. Minden réteg azonos neuronszámú. A tanítás során a túltanítás elkerülésére rétegenként batch normalizálást és dropout regularizációt alkalmaztunk. A teljes hálózatot az időjárásállomás által mért, perces globálsugárzás és diffúz sugárzás értékkel csatoltuk vissza, és backpropagation módszerrel tanítottuk. A tanítás során a hibát a becsült és mért sugárzás különbségének átlagos négyzetes középértékeként (root mean square error vagy RMSE) definiáltuk a két sugárzáskomponens összegére, optimalizáló algoritmusnak a változatos körülmények között is igen robosztus Adam algoritmust választottuk.



2.7. ábra. A neurális háló vázlata

2.7. Hiperparaméter optimalizáció

A modell hiperparamétereit Bayesi optimalizálással finomhangoltuk. A hangolt hiperparaméterek a teljesen összekapcsolt hálók aktivációs függvénye, a bemeneti adatokhoz (időjárási adatok, légköri adatok, járulékos adatok) tartozó jellemzővektor mérete, a kimeneti jellemzővektor mérete és az Adam becslő tanítási rátáját (learning rate) voltak. Az optimalizált paramétereket, és azok értékeit időjárási szcenáriónként a 2.4 táblázat foglalja össze.

A finomhagolás során minden időjárási szcenárióhoz 65 különböző hiperparaméter-kombinációt vizsgáltunk. Minden kombinációt 15 epoch tanítási cikluson keresztül tanítottuk 4000 darab véletlenszerűen választott mintapontra, amelyeket 0,8-0,2 arányban tanító és validáló adatsorra választottunk szét. A kísérletek sikerét a tanítás végén előállt validálási RMS hibaértékkel jellemeztük. A 65 próbálkozás során elért legkisebb hibamértéket elérő modellt további 1000 epochon keresztül továbbtanítottuk 10 000 mintapont felhasználásával. A továbbtanítást leállítottuk, ha 100 epoch után a validálási hibametrika nem mutatott javuló tendenciát. Így kaptuk a végső, becslési pontosság meghatározására szolgáló hálóarchitektúrát.

A háló az alábbi hiperparaméterekkel dolgozik, melyek lehetséges értékei a 2.3 táblázatból olvashatóak le:

- Rétegek száma

  A háló rétegeinek száma.

- Rétegek mérete

  A rétegek neuronszáma. A háló minden rétege azonos méretű.

- Aktivációs függvény

  Minden réteg kimenetén egységes az aktivációs függvény. Az aktivációs függvény típusa megadja nemlinearitás jellegét a bemeneti és kimeneti adatok között.

- Dropout arány

  Rétegenként alkalmazott regularizációs technika, a dropout arányra a 0,1-0,9 tartományból választható folytonos érték.

- Tanulási ráta

  A teljes modell optimalizálásához használt Adam algoritmus tanulási rátája, mely meghatározza, hogy milyen mértékben módosulnak a modell súlyai és paraméterei a tanítás során.

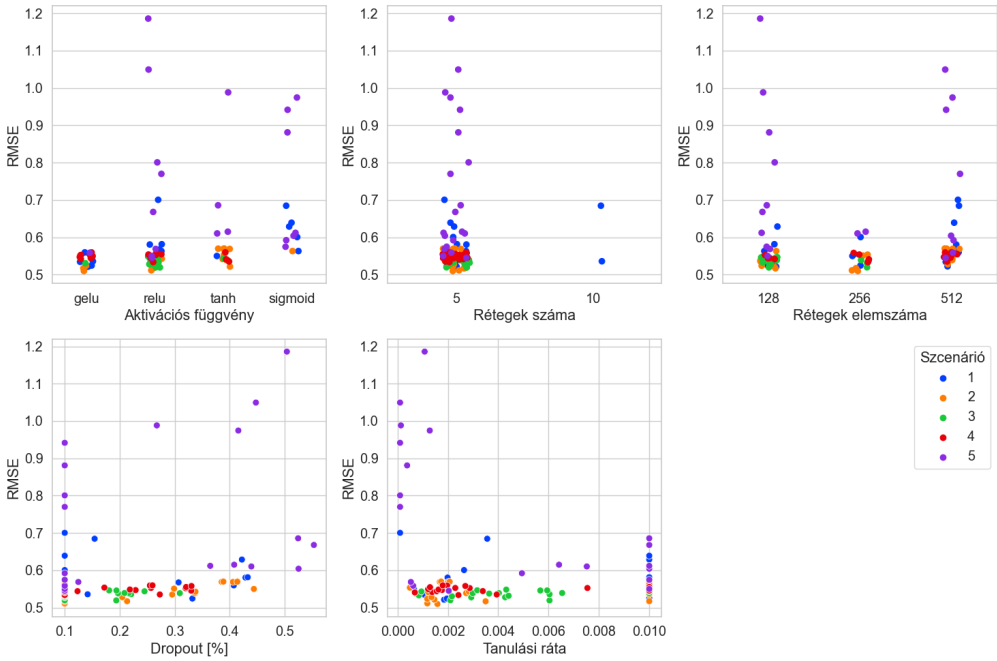
Hiperparaméter neve	Lehetséges értékek
Rétegek száma	5, 10, 50, 70
Rétegek mérete	128, 256, 512
Aktivációs függvénye	tanh, sigmoid, relu, gelu
Dropout arány	0,1-0,8
Tanulási ráta	0,0001–0,01

2.3. táblázat. Hiperparaméterek és lehetséges értékeik.

A 20 legjobb eredményt elérő hiperparaméter-kombinációkhoz tartozó validációs RMS hibaértékeket minden paraméterre, időjárási szcenáriókra bontva ábrázoltam a 2.8 ábrán.

Szcenárió	**1**	**2**	**3**	**4**	**5**
Rétegek száma	5	5	5	5	5
Rétegek mérete	128	256	128	256	512
Aktivációs függvény	gelu	gelu	relu	relu	relu
Dropout arány	0,1	0,1	0,1	0,2	0,1
Tanulási ráta	0,0018	0,0016	0,006	0,0024	0,002
Tanítható paraméterek száma	70 914	271 874	70 274	271 362	1 066 498
MFLOPS	9,1	34,9	8,9	34,5	136

2.4. táblázat. A legjobb modellekhez tartozó hiperparaméter-értékek és a legjobb modell tanítható súlytényezőinek száma, illetve a hálók MFLOPS mértéke időjárási szcenárión-ként.



2.8. ábra. A 20 legjobb hiperparaméter-kombináció szcenáriónként, és azok RMS hibája.

A hiperparaméter-optimalizáció után előálló hibák alapján megállapítható, hogy a tanítás kezdeti fázisában ha több bemeneti időjárási adat és Clear Sky adatsor is rendelkezésre áll, jobb becslési pontosság érhető el. Az 5-ös szcenárióhoz tartozó hibaértékek jóval szélesebb tartományban mozognak, mint a több adatot használó szcenáriók, bár a legjobb modellhez tartozó modell hibája egy nagyságrendbe esik a többi szcenárióéval. A pontossághoz minden időjárási szcenárió esetén a közepes tanulási ráta járul hozzá, illetve a gelu aktivációs függvény mellett teljesítenek a hálók a legjobban. Feltűnő, hogy a rétegek száma a 20 legjobb paraméterkombinációra minden esetben 5 és 10, a nagyobb rétegméretű hálók közül nem került ki megfelelően alacsony hibájú modell (az ábrán a 20 legjobb hiperparaméterkombinációhoz tartozó hálót ábrázoltuk, más, nagyobb hibájú paraméterkombinációk esetén a rétegek száma lehetett 50 és 70 is).

Az egyes szcenáriókhoz tartozó legjobb modellek (2.4. táblázat) minden esetben 5 réteggel rendelkeznek, és a rétegek elemszáma is általában a két alacsonyabb érték, 128 vagy 256. Kivétel az 5-ös szcenárió, amelynek rétegmérete 512. A legjobb hálómodellek komplexitását a rétegek mérete és száma mellett a súlytényezők számával és a háló MFLOPS (Million Floating-Point Instructions Per Second) értékével jellemeztük. Feltűnő, hogy az 5-ös szcenárióhoz tartozó háló egy nagyságrenddel nagyobb a többinél. Mivel a háló mérete jellemzi az egyes szcenáriók esetén a megoldandó következtetési probléma komplexitását,

ez utalhat arra, hogy az 5-ös szcenárióban a háló bizonyos, időjárási adatok között fennálló összefüggéseket kikövetkeztet, amelyekre a több adatot használó modelleknek nincs szüksége. Különösen érdekes, hogy míg a hiperparaméter-optimalizáció során az 5-ös szcenárióra kapjuk a legalacsonyabb pontosságú hálót, a tesztelés során, az új adatkészleten ez a szcenárió teljesít a legjobban, valószínűleg komplex felépítésének köszönhetően (3.1. ábra).

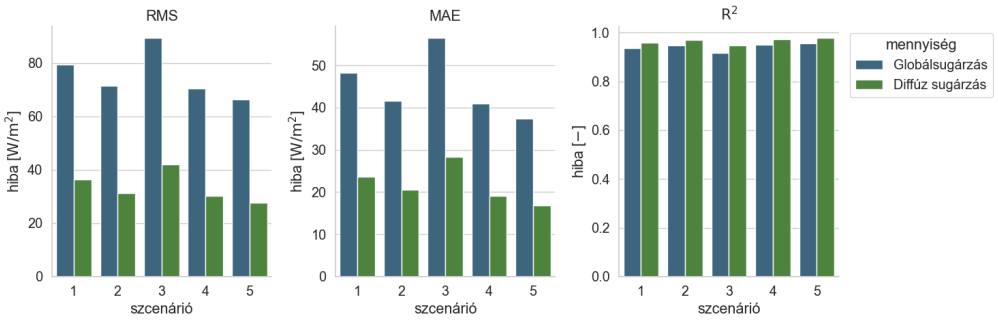
3. fejezet

Eredmények

3.1. A modellek pontossága

Az optimalizált, majd továbbtanított hálók becslési pontosságát független mintákból álló tesztadatsorral értékeltem a különböző időjárási szcenáriók esetén. A tesztadatsor összesen 152 000 adatpontból áll, és minden, a tanításhoz vagy a hiperparaméter-optimalizációhoz nem használt mintát tartalmaz. Az eredményeket a 3.2,3.1 ábrákon szemléltetem. A háló kimenetét a 3.2 regressziós ábrán 1000 tesztpontban ábrázolom, a bal oszlopban a globálsugárzáshoz, a jobb oszlopban pedig a diffúz sugárzáshoz tartozó eredmények láthatók. A sorok az időjárási szcenárióknak felelnek meg. A becsült és mért adat közötti lineáris kapcsolatot a Pearson korrelációs együtthatóval (R^2) jellemeztem.

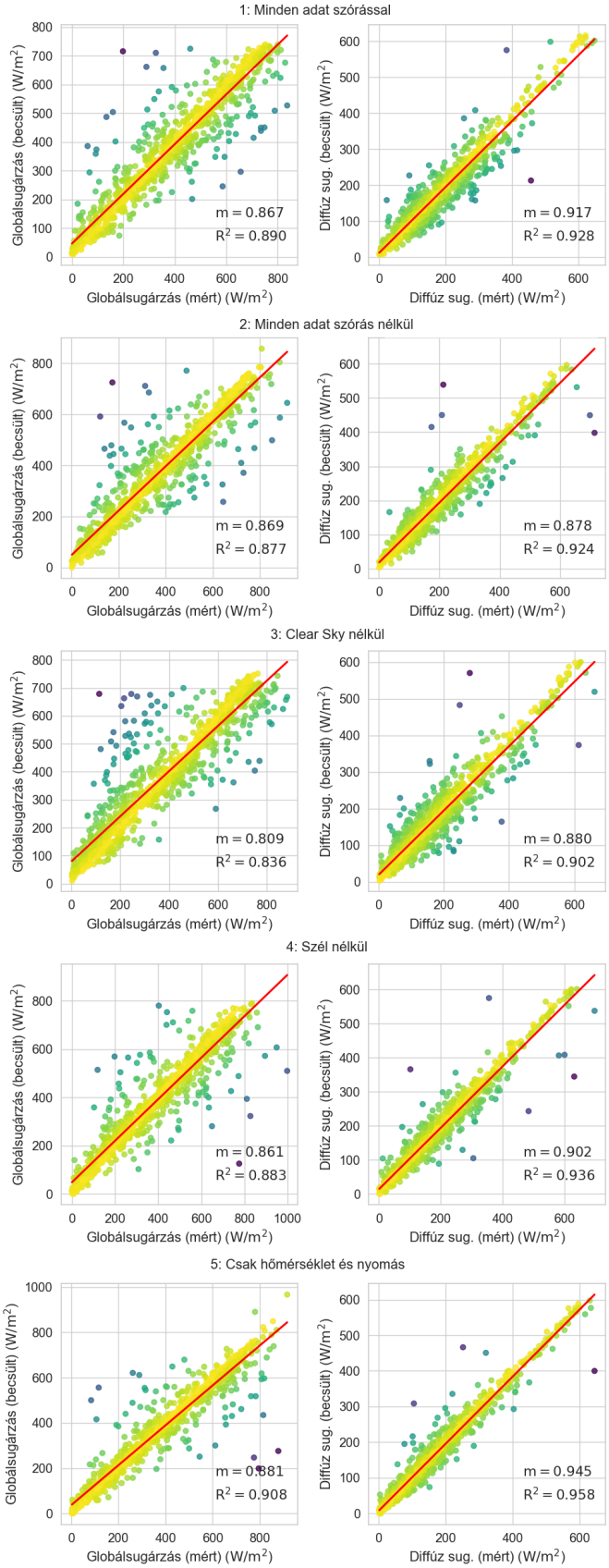
Az eredmények azt mutatják, hogy a becslés pontossága mind a globál, mind pedig a diffúz sugárzás tekintetében is az 5-ös szcenárióban a legnagyobb, ahol $R^2 = 0{,}908$. Ez az eredmény meglepő, mert az 5-ös szcenárió használja a legkevesebb meteorológiai adatot, azonban a pontosság valószínűleg a modell komplexitásának köszönhető.



3.1. ábra. A becslés pontossága időjárási szcenáriónként különböző hibametrikákkal számszerűsítve. RMS: Root Mean Square Error, MAE: Mean Average Error. A független tesztadatsorra mindkét sugárzáskomponens esetében az 5-ös szcenárióhoz tartozó háló teljesít a legjobban.

Az időjárási szcenáriók közül a 3-as számú, Clear Sky adatsorral nem dolgozó háló teljesít a leggyengébben. Ebből arra lehet következtetni, hogy a Clear Sky adat ismerete nélkül a szélsebesség torzítja a becslés pontosságát.

Összességében elmondható, hogy a háló a különböző időjárási szcenáriók esetén hasonló pontossággal teljesít, az R^2 értéke a napelemek termelését közvetlenül befolyásoló globálsugárzás esetén 0,85 körüli, míg diffúz sugárzás esetén ennél nagyobb, nagyjából 0,9. Mindezekből arra lehet következtetni, hogy a becslés pontosságának szempontjából az égboltképekből kinyert paraméterek szerepe a tanítás során jelentősebb, mint a megfelelő konfigurációjú időjárási adatok összeállítása.



3.2. ábra. A becslés pontossága időjárási szcenáriónként. A piros egyenes az adatpontokra illesztett regressziós egyenesnek felel meg, amelynek meredekségét az m paraméterrel jellemeztem. A minták színe az $y = x$ egyenestől való távolságuk függvényében változik.

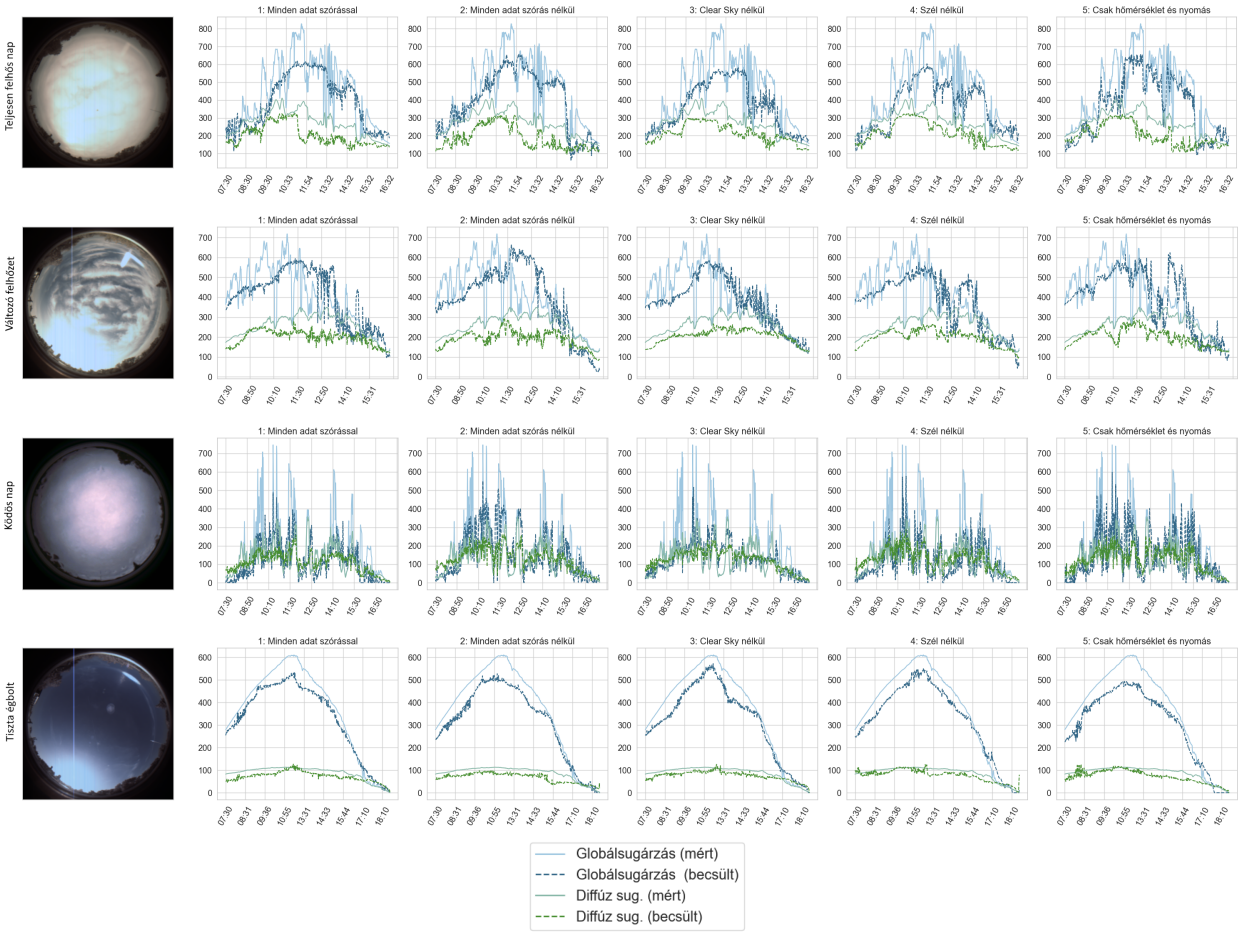
3.2. A becslés pontossága különböző időjárási helyzetekben

A 3.3 ábrán szemléltetem négy különböző időjárású napon a globál és diffúz sugárzás értékét, valamint a becslések eredményét az 5 időjárási szcenárió esetén. Az ábrázolt időjárási helyzetek: teljesen felhővel borított égbolt, változó felhőzöttség, ködös nap és tiszta égbolt. A grafikonokon összefüggő vonallal a mért, szaggatott vonallal pedig a becsült sugárzás adat olvasható le.

Az időjárási helyzetek eredményeit vizsgálva megállapítható, hogy a neurális háló különböző adatkonfigurációk mellett egységesen nagyobb hibával becsül változó felhőzöttségű napok esetén. A globálsugárzás görbén a letörések pár perces felhőátvonulások, hirtelen globálsugárzás változások, melyet a becslés nem tud megfelelő pontossággal lekövetni. Látható, hogy ilyen helyzetekben a diffúz, felhőkről visszavert sugárzás megnő, melyet a háló szintén nem becsül kellően pontosan.

Tiszta égbolt és ködös nap esetén a becslés jól követi a mért adatsort.

Összességében megállapítható, hogy a neurális háló a változékony felhőzöttségű napokon becsül legpontatlanabbul, ezért a jövőben érdemes növelni a tanítás során használt adathalmazt további felhős napok adatpontjaival.



3.3. ábra. A becslés eredménye négy különböző időjárású nap esetén: Teljesen felhővel borított égbolt egész nap, változó felhőzöttség napközben, ködös nap, tiszta égbolt egész nap.

3.3. A módszer összevetése képfeldolgozó neurális hálóval

Jelen munkám egyik célja volt összehasonlítani a képi adatokkal közvetlenül dolgozó neurális háló becslési pontosságát a képekből kinyert égboltparaméterekkel dolgozó neurális

háló eredményeivel.

Összességében elmondható, hogy a [27]-ban a kutatócsoportunk által kidolgozott neurális háló esetén a legmagasabb pontosság, amit elértünk $R^2 = 0{,}77$ volt, amely alacsonyabb a jelenleg elért legmagasabb becslési pontossághoz képest, melyet az 5. időjárási szcenárióra kaptunk. Megállapítható tehát, hogy az égboltképekből kinyert adatok segítségével a becslés pontossága magasabb mint a képfeldolgozásra specializált neurális háló esetén. A különbség oka, hogy a [27]-ben bemutatott háló a korlátos erőforrások és tanítási idő miatt nem tudott nagyobb pontosságot elérni, ez azonban a háló tanításához szükséges erőforrások növelésével javítható.

A képfeldolgozó neurális tansformer háló kb. 90 000 000 tanítható paraméterrel rendelkezik, míg ez a szám a jelenlegi, egy darab teljesen összekapcsolt háló alkalmazásával történő becslés esetén 70 000 és 1 000 000 között mozog.

A két háló közötti különbség azért is jelentős, mert az égboltképekből kinyert paraméterekkel tanított háló ezt a pontosságot egy grafikus processzor nélküli személyi számítógépen tanítva érte el. A teljes tanítási idő a hiperparaméter optimalizálással együtt szcenáriónként nagyjából 4 órát vett igénybe[1].

A hálók tanítása során mind a tanítási időkben, mind a felhasznált erőforrások mértékében jelentős különbséget tapasztaltunk. Ugyanazon infrastruktúrán a képfeldolgozó neurális háló tanítása több napot vett igénybe. A kisebb erősforrásigény a módszer korlátos erőforrású rendszerekbe történő integrálását is lehetővé teszi.

[1]Processzor: Intel Core i7-6700 @ 3,40 GHz, 4 mag, 16 GRAM

4. fejezet

Összefoglalás

Jelen dolgozatom fő célkitűzése az általam kidolgozott mesterséges neurális háló alapú sugárzásbecslő módszer bemutatása, mely a hagyományos algoritmusok és az adatalapú megközelítések előnyeit kombinálja. A teljes égbolt képek használata helyett az égbolt felhőzetét egy, a hagyományos képfeldolgozó algoritmusok segítségével előállított paramétervektorral jellemzem. Ez lecsökkenti a feldolgozott adatmennyiséget és igen hatékonnyá teszi a módszert, miközben továbbra is figyelembe veszi az égboltképekből származó információkat.

A módszer átfogó értékeléséhez több meteorológiai adat-kombinációt állítottunk össze, illetve a kapott modelleket pontosság és erőforrásigény szempontjából összehasonlítottam a kutatócsoportunk által kidolgozott, teljes kameraképeket használó neurális hálókkal is.

Az eredmények azt mutatják, hogy míg az időjárási adatösszeállításnak nincs jelentős hatása a pontosságra, addig a kevesebb meteorológiai adatot használó, de komplexebb háló bonyolult időjárási összefüggések felismerésére képes. Összességében megállapítható, hogy a becslés pontossága szempontjából nincs nagy jelentősége annak, hogy a neurális háló égboltképeket vagy képekből kinyert paramétereket kap bemeneti adatként, azonban a háló tanításához szükséges erőforrások szempontjából jelentős a különbség. Emellett elmondható, hogy a neurális háló nem kezeli elég pontosan a hirtelen felhőzetváltozásból fakadó napsugárzás változásokat, így a becslés javítása érdekében további tanítás szükséges.

4.1. Kitekintés

Az EK Környezetfizikai Laboratóriumában zajló kutatás fő célja a napelemek termelését közvetlenül befolyásoló globálsugárzás előrejelzése. A kutatás jelenlegi fázisában a minél pontosabb globálsugárzás becslés a cél, melynek érdekében a hálót tovább fogjuk fejleszteni, illetve további összehasonlításokat fogunk végezni az égboltparaméterekkel és a kamerakképekkel végzett becslések között.

Emellett terveink közé tartozik egy előrejelző modell kidolgozása a hálók visszacsatolásával, valamint a becslési bizonytalanság meghatározása különböző időjárási körülmények között, hogy teljesebb képet kapjunk a hálók előrejelzési képességeiről.

5. fejezet

Köszönetnyilvánítás

Köszönöm Barancsuk Lillának és Oláhné Groma Veronikának az értékes szakmai iránymutatást és emberi támogatást, amit az elmúlt több, mint másfél évben kaptam Tőlük! Támogatásuk nélkül ez a TDK dolgozat nem jöhetett volna létre!

Irodalomjegyzék

[1] „Magyarország nemzeti energia- és klímaterve 2023. évben felülvizsgált változat," https://kormany.hu/dokumentumtar/magyarorszag-nemzeti-energia-es-klimaterve-felulvizsgalt-valtozat-tars-egy, 2023, Letöltve: 2023. 10. 30.

[2] P. Bórawski, R. Wyszomierski, A. Bełdycka-Bórawska, B. Mickiewicz, B. Kalinowska, J. W. Dunn, and T. Rokicki, „Development of renewable energy sources in the european union in the context of sustainable development policy," *Energies*, vol. 15, no. 4, p. 1545, 2022.

[3] MAVIR, „Naptermelés becslés és tény adatok," Elérhető: https://mavir.hu/web/mavir/naptermeles-becsles-es-teny-adatok, 2023, Letöltve: 2023. 10. 30.

[4] B. Turóczi, B. Hartmann, and A. Kazsoki, „Felhőátvonulások hatása az elosztóhálózat feszültségére"," *Elektrotechnika*, vol. 114, no. 03, pp. 14–16, 2021.

[5] A. Mellit, A. Massi Pavan, E. Ogliari, S. Leva, and V. Lughi, „Advanced methods for photovoltaic output power forecasting: A review," *Applied Sciences*, vol. 10, no. 2, p. 487, 2020.

[6] Z. Li, K. Wang, C. Li, M. Zhao, and J. Cao, „Multimodal deep learning for solar irradiance prediction," 07 2019, pp. 784–792.

[7] K. Hu, L. Wang, W. Li, S. Cao, and Y. Shen, „Forecasting of solar radiation in photovoltaic power station based on ground-based cloud images and bp neural network," *IET Generation, Transmission & Distribution*, vol. 16, no. 2, pp. 333–350, 2022.

[8] R. Ahmed, V. Sreeram, Y. Mishra, and M. Arif, „A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization," *Renewable and Sustainable Energy Reviews*, vol. 124, p. 109792, 2020.

[9] A. Holtslag and A. Van Ulden, „A simple scheme for daytime estimates of the surface fluxes from routine weather data," *Journal of Applied Meteorology and Climatology*, vol. 22, no. 4, pp. 517–529, 1983.

[10] F. Barbieri, S. Rajakaruna, and A. Ghosh, „Very short-term photovoltaic power forecasting with cloud modeling: A review," *Renewable and Sustainable Energy Reviews*, vol. 75, pp. 242–263, 2017.

[11] Reuniwatt, „24/7 cloud observations with sky insight™ all-sky imager," Elérhető: https://reuniwatt.com/en/247-all-sky-observation-sky-insight, Letöltve: 2023. 10. 30.

[12] Solcast, „How solcast generates irradiance and weather data," Elérhető: https://solcast.com/irradiance-data-methodology, Letöltve: 2023. 10. 30.

[13] B. J. Martins, A. Cerentini, S. L. Mantelli, T. Z. L. Chaves, N. M. Branco, A. von Wangenheim, R. Rüther, and J. M. Arrais, „Systematic review of nowcasting approaches for solar energy production based upon ground-based cloud imaging," *Solar Energy Advances*, vol. 2, p. 100019, 2022.

[14] P. Kumari and D. Toshniwal, „Deep learning models for solar irradiance forecasting: A comprehensive review," *Journal of Cleaner Production*, vol. 318, p. 128566, 2021.

[15] J. Xu, S. Yoo, D. Yu, D. Huang, J. Heiser, and P. Kalb, „Solar irradiance forecasting using multi-layer cloud tracking and numerical weather prediction," in *Proceedings of the 30th Annual ACM Symposium on Applied Computing*, 2015, pp. 2225–2230.

[16] Y. Chu, H. T. Pedro, M. Li, and C. F. Coimbra, „Real-time forecasting of solar irradiance ramps with smart image processing," *Solar Energy*, vol. 114, pp. 91–104, 2015.

[17] J. Zhang, R. Verschae, S. Nobuhara, and J.-F. Lalonde, „Deep photovoltaic nowcasting," *Solar Energy*, vol. 176, pp. 267–276, 2018.

[18] Y. Lin, D. Duan, X. Hong, X. Han, X. Cheng, L. Yang, and S. Cui, „Transfer learning on the feature extractions of sky images for solar power production," in *2019 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2019, pp. 1–5.

[19] W.-C. Tsai, C.-S. Tu, C.-M. Hong, and W.-M. Lin, „A review of state-of-the-art and short-term forecasting models for solar PV power generation," 2023.

[20] F. Wang, J. Li, Z. Zhen, C. Wang, H. Ren, H. Ma, W. Zhang, and L. Huang, „Cloud feature extraction and fluctuation pattern recognition based ultrashort-term regional pv power forecasting," *IEEE Transactions on Industry Applications*, vol. 58, no. 5, pp. 6752–6767, 2022.

[21] R. Samu, M. Calais, G. Shafiullah, M. Moghbel, M. A. Shoeb, B. Nouri, and N. Blum, „Applications for solar irradiance nowcasting in the control of microgrids: A review," *Renewable and Sustainable Energy Reviews*, vol. 147, p. 111187, 2021.

[22] D. Peters, R. Völker, T. Kilper, M. Calais, T. Schmidt, C. Carter, and C. von Maydell, „Model-based design and simulation of control strategies to maximize the PV hosting capacity in isolated diesel networks- using solar short-term forecasts for predictive control of diesel generation," in *32nd European Photovoltaic Solar Energy Conference and Exhibition*, 2016.

[23] R. Harris, „Karratha Airport Solar Project, 1 MWp solar PV system," Elérhető: https://arena.gov.au/assets/2017/02/karratha-solar-farm-public-impact-report.pdf, 2018, Letöltve: 2023. 10. 30.

[24] R. Samu, S. G. Bhujun, M. Calais, G. Shafiullah, M. Moghbel, M. A. Shoeb, and B. Nouri, „Solar irradiance nowcasting system trial and evaluation for islanded microgrid control purposes," *Energies*, vol. 15, no. 17, 2022.

[25] D. M. Günter, „Felhőborítottság becslése képfeldolgozás segítségével napelem termelésbecslés támogatására," Tudományos Diákköri Konferencia dolgozat, Budapesti Műszaki és Gazdaságtudományi Egyetem, 2022.

[26] D. M. Günter, L. Barancsuk, V. Groma, and B. Sinkovics, „Az égboltkamerák szerepe a fotovillamos termelés rövidtávú előrejelzésében," *Elektrotechnika*, vol. 116, pp. 7–10, 2023.

[27] L. Barancsuk, D. M. Günter, V. Groma, and B. Hartmann, „Fotovillamos termelésbecslés támogatása multimodális neurális háló segítségével égboltkameraképek és különböző konfigurációjú időjárási adatok alapján," in *Mechwart András Ifjúsági Találkozó Absztraktkötet*, 2023, Közlésre elfogadva.

[28] M. Horváth, „Napelem termelésbecslésének támogatása égboltkamera képek mesterséges intelligencia alapú elemzésével," Szakdolgozat, Budapesti Műszaki és Gazdaságtudományi Egyetem, 2022.

[29] T. Mertens, J. Kautz, and F. Van Reeth, „Exposure fusion," in *15th Pacific Conference on Computer Graphics and Applications (PG'07)*, 2007, pp. 382–390.

[30] G. Bradski, „The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[31] M. Lefevre, A. Oumbe, P. Blanc, B. Espinar, B. Gschwind, Z. Qu, L. Wald, M. Schroedter-Homscheidt, C. Hoyer-Klick, A. Arola *et al.*, „Mcclear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions," *Atmospheric Measurement Techniques*, vol. 6, no. 9, pp. 2403–2418, 2013.

[32] Wikipedia, „Solar zenith angle," Elérhető: https://en.wikipedia.org/wiki/Solar_zenith_angle, Letöltve: 2023. 10. 31.

[33] H. Gábor, E. Ádám, B. Miklós, B. András, B. Pál, H. Ákos, K. Bumke, and A. Macke, „Felhôzöttségmérés, optikai felhôfelismerô algoritmusok összehasonlítása–1. rész," *Fizikai Szemle*, vol. 7, p. 227, 2015.

[34] Wikipedia, „Accuracy and precision," Elérhető: https://en.wikipedia.org/wiki/Accuracy_and_precision, Letöltve: 2023. 10. 30.

[35] Wikipedia, „Precision and recall," Elérhető: https://en.wikipedia.org/wiki/Precision_and_recall, Letöltve: 2023. 10. 30.

[36] B. Nouri, S. Wilbert, L. Segura, P. Kuhn, N. Hanrieder, A. Kazantzidis, T. Schmidt, L. Zarzalejo, P. Blanc, and R. Pitz-Paal, „Determination of cloud transmittance for all sky imager based solar nowcasting," *Solar Energy*, vol. 181, pp. 251–263, 2019.

[37] P. Tzoumanikas, E. Nikitidou, A. Bais, and A. Kazantzidis, „The effect of clouds on surface solar irradiance, based on data from an all-sky imaging system," *Renewable energy*, vol. 95, pp. 314–322, 2016.

[38] B. Holtslag and A. Ulden, „A simple scheme for daytime estimates of the surface fluxes from routine weather data," *Journal of Applied Meteorology*, vol. 22, pp. 517–529, 04 1983.

[39] C. Balafas, M. Athanassopoulou, T. Argyropoulos, P. Skafidas, and C. Dervos, „Effect of the diffuse solar radiation on photovoltaic inverter output," in *Melecon 2010 - 2010 15th IEEE Mediterranean Electrotechnical Conference*, 2010, pp. 58–63.

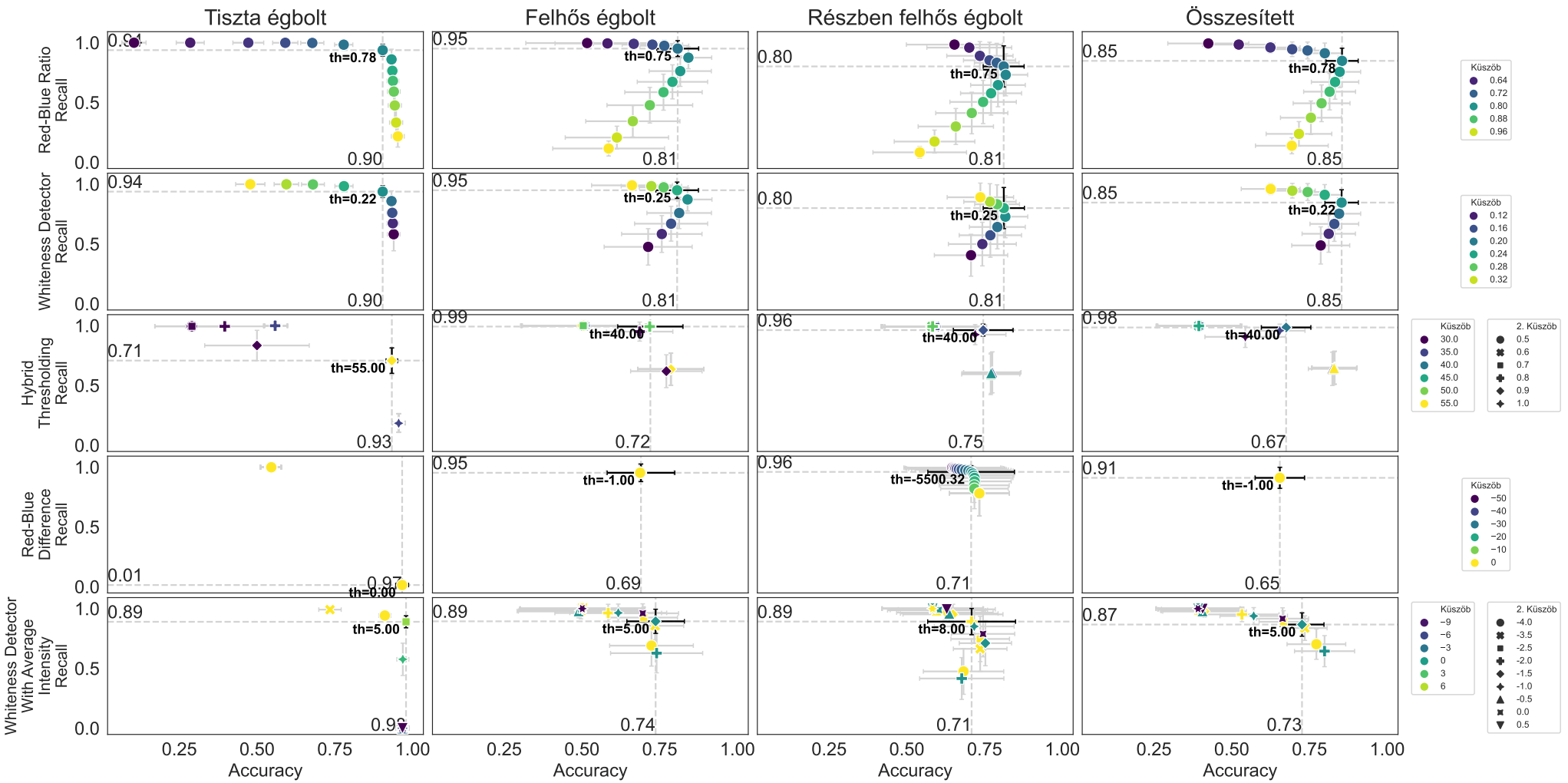
A. függelék

Képfeldolgozó algoritmusok optimális küszöbértékeinek beállítása

A felhőparaméterek megállapításához különböző pixelszintű küszöbözésen alapuló algoritmusokat alkalmaztunk, amelyek az égboltképek képpontjait felhő illetve tiszta égbolt kategóriákba sorolják. Az algoritmusok mindegyike küszöbparamétereket használ az égbolt és a felhők megkülönböztetésére. Ahhoz, hogy megtaláljuk az algoritmusokhoz tartozó legjobb küszöbértéket, összeállítottunk egy 120 égboltképből álló, a magyarországi felhőtípusok tekintetében reprezentatív adatbázist, amely minden képéhez elkészítettünk egy referencia felhőmaszkot. Az adatbázis alapján az algoritmusok teljesítménye számszerűsíthető.

Először empirikusan megállapítottunk egy tartományt minden algoritmushoz, amelyen belül a megfelelő küszöbértékek elhelyezkedhetnek, majd a tartományon belül több különböző paraméterbeállítás mellett futtattuk az algoritmusokat, és mindegyikhez meghatároztuk a recall és accuracy értékeket. A kapott értékek közül (A.1. ábra) kiválasztottuk azt a küszöbértéket, amely a recall mind az accuracy értékek négyzetes középértékét maximalizálja.

A referencia adatbázis képeit osztályoztuk aszerint is, hogy milyen az égbolt állapota: tiszta égbolt, felhős égbolt és részben felhős égbolt időjárási helyzeteket definiáltunk, és az algoritmusok pontosságát időjárási helyzet függvényében is értékeltük. Az A.1. ábráról leolvasható, hogy míg a Hybrid Thresholding algoritmus igen jól teljesít tiszta égbolt esetén, felhős égbolt mellett pontossága jelentősen csökken. További vizsgálódásainkhoz ezért a robusztus Red–Blue Ratio algoritmust választottuk, amely többféle időjárási helyzet mellet is nagy pontosságot mutat.



A.1. ábra. Algoritmusok optimális paraméterbeállítása.