

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

OTDK

Rövidtávú kiegyenlítetlenség előrejelzés gépi tanulással
proaktív mFRR szabályozás céljából

Markovics Dávid
MSc hallgató

Konzulensek:

Dr. Mayer Martin János

BME-GPK Energetikai Gépek és Rendszerek Tanszék

Bessenyei Tamás

Magyar Villamosenergia-ipari Átviteli Rendszerirányító Zrt.



Budapest, 2021

Absztrakt

Az elmúlt években az európai villamosenergia rendszer összekapcsolással párhuzamosan a kiegyenlítőenergia piacok is jelentős átalakuláson mentek keresztül, amely folyamat jelenleg is tart hazánkban, illetve az együttműködő országokban. Az átalakulás első lépése a hagyományos primer, szekunder és terciér tartalék fogalmak leváltása volt az új, egységes rendszerszintű szolgáltatási termék típusokra (FCR, aFRR, mFRR, RR). Ezek implementálása hosszabb távon a kiegyenlítőenergia szolgáltatások európai szintű harmonizálásának lehetőségét célozta meg, amelyért 2017-től kezdve aFRR esetén a PICASSO, míg az mFRR termékre a MARI projekt felelős. Ezen kereskedési platformok jelenleg előkészítés alatt állnak, mind az ENTSO-E, mind a résztvevő országok TSO-i részéről és várhatóan 2022 Q3-ban indul majd el az éles üzemük.

Jelen dolgozat fő problémafelvetése ehhez a felkészüléshez kapcsolódik; lényegében egy szemléletváltás lehetőségének vizsgálata, a jelenleg alkalmazott reaktív jellegű mFRR aktiválás átalakítása proaktívra, amely az eddigi jellemző szempontok (alacsony rendelkezésre álló forgó tartalék, erőművi kiesés) mellett gazdasági és egyéb rendszeroptimalizációs megfontolásokra ad lehetőséget. Ehhez a proaktív magatartáshoz szükség van kiegyenlítetlenség előrejelzésre, amely feltáró lehetőségelemzése képi a dokumentum tartalmi részét.

Az elemzés első részében a rendelkezésre álló adatokat vizsgáltam különböző tanuló algoritmusok alkalmazásával, a magyarázó változók lehető legjobb kombinációját keresve, illetve párhuzamosan a modellválasztást előkészítve. A dolgozat második felében egy tesztkörnyezetet hoztam létre, amelyben Gradient Boosting alapú kvantilis regresszióval meghatározott kiegyenlítetlenség előrejelzésből, az adott időpontra vonatkozó változtatható kritériummal bíró aktiválási protokollal szimuláltam a MARI platformnak megfelelő termékdefiníciókkal elérhető mFRR szolgáltatásokat a tesztidőszakban. A vizsgálat fő eredményeihez használt kimutatások a pontbecslések átlagos hibái, a konfidencia intervallum valós pontossága, a tesztidőszakban proaktívan aktivált mFRR szabályozás, annak aránya a teljes kiegyenlítetlenséghez, valamint az okozott keresztszabályozás mértéke. Ezek az eredmények különböző konfidenciaszint, illetve aktiválási kritérium mellett kerülnek bemutatásra a dolgozatban, a többszöri újratanítás hatását is kiemelve.

In recent years, in parallel with the interconnection of the European electric power system, the balancing energy markets have also undergone a significant transformation, a process that is still taking place in Hungary and in the cooperating countries. The first step in the transformation was replacing the traditional primary, secondary, and tertiary reserve concepts with new, unified system-level service product types (FCR, aFRR, mFRR, RR). In the longer term, their implementation aimed at the possibility of harmonizing balancing energy services at European level, for which PICASSO will be responsible for the aFRR, and MARI for the mFRR product. These trading platforms are under development since 2017 and are currently being prepared for the go-live scheduled for 2022 Q3, both by ENTSO-E and by TSOs of the participating countries.

The main problem of the present work is related to this preparation process; essentially examining the possibility of a change of approach, transforming the currently used reactive mFRR activation into proactive, which allows for economic and other system optimization considerations in addition to the existing activation principles (low available revolving reserve, power plant outage). This proactive behavior requires an imbalance forecast, the exploratory analysis of which forms the content of the document.

In the first part of the analysis, the available data were examined using different learning algorithms, looking for the best possible combination of explanatory variables, and in parallel

preparing the model choice. In the second half of the dissertation, test environment was created in which the possible mFRR services were simulated using the product definitions according to the MARI platform. In this part, the test was based on an imbalance prediction determined by Gradient Boosting-based quantile regression, supplemented with an activation protocol, based on a variable criteria for the given work point. The statements used for the main results of the study were the mean errors of the point estimates, the true accuracy of the confidence interval, the proactively activated mFRR control during the test period, its ratio to total imbalance, and the proportion of cross-regulation caused. These results are presented in the dissertation with different confidence levels and activation criteria, also briefly highlighting the effect of repeated retraining.

Kulcsszavak

kiegyenlítetlenség előrejelzés, gépi tanulás, proaktív mFRR szabályozás

Rövidítések

ACE	Area Control Error (területi szabályozási hiba)
aFRR	automatic Frequency Restoration Reserve (automatikus frekvencia visszaállítási tartalék)
BSP	Balancing Service Provider (kiegyenlítő termék szolgáltató)
DA	Direct Activation (direkt aktiválás)
EBGL	Electricity Balancing Guideline (villamosenergia kiegyenlítés irányelv)
FCR	Frequency Containment Reserve (frekvencia megtartó tartalék)
FRR	Frequency Restoration Reserve (frekvencia visszaállítási tartalék)
HVDC	High Voltage Direct Current (nagyfeszültségű egyenáramú átviteli hálózat)
HMKE	Háztartási méretű kiserőmű
ID	IntraDay (napon belüli)
IGCC	International Grid Control Cooperation (nemzetközi hálózatszabályozási együttműködés)
MARI	Manually Activated Reserves Initiative (manuálisan aktivált tartalékokra vonatkozó kezdeményezés)
MAVIR	Magyar Villamosenergia-ipari Átviteli Rendszerirányító
mFRR	manual Frequency Restoration Reserve (manuális frekvencia visszaállítási tartalék)
MTU	Market Time Unit (piaci időegység)
PICASSO	Platform for the International Coordination of Automated Frequency Restoration and Stable System Operation (az automatikus frekvencia visszaállítási tartalék és stabil rendszerüzemeltetés nemzetközi együttműködéséért felelős platform)
RR	Replacement Reserve (helyettesítő tartalék)
SA	Scheduled Activation (ütemezett aktiválás)
SOGL	System Operation Guideline (rendszerüzemeltetési irányelv)
TSO	Transmission System Operator (átviteli hálózati rendszerirányító)

Tartalomjegyzék

1 Bevezetés	1
1.1 Szabályozás szintjei, fogalmai	1
1.2 Proaktív mFRR szabályozás célja	3
1.3 Kiegyenlítetlenség előrejelzés szakirodalmi előfordulásai	5
1.4 Áttekintés	7
2 Módszertan.....	8
2.1 Adat előkészítés és használmagyarázó változók	8
2.2 Gépi tanulás, használt algoritmusok.....	10
2.3 Hiperparaméter hangolás.....	12
2.4 A pontbecslések értékelési szempontjai	12
2.5 Platformos rendszer tesztkörnyezetének alapjai	13
2.6 Keresztszabályozás.....	14
2.7 Újratanítás	16
3 Adatelemzés	17
3.1 Korreláció vizsgálat.....	17
3.2 IGCC rendelkezésre állás vizsgálata	19
3.3 Változó relevancia vizsgálat.....	23
4 Eredmények	25
4.1 Pontbecslések eredményei.....	25
4.2 Platformos környezet.....	27
4.3 mFRR statisztikák napi bontásban	32
4.4 mFRR – IGCC.....	33
5 Összefoglalás	35
6 Hivatkozások.....	37
7 Függelék.....	41

Köszönetnyilvánítás

Köszönettel tartozom konzulenseimnek, Dr. Mayer Martin Jánosnak és Bessenyei Tamásnak, akik segítségükkel hozzájárultak a dolgozat megalkotásához, továbbá a MAVIR Zrt.-nek és kiemelten Fábián Attilának az adatok biztosításáért, és azok értelmezésében végzett segítségükért.

A kutatás az Innovációs és Technológiai Minisztérium ÚNKP-21-2 kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült.

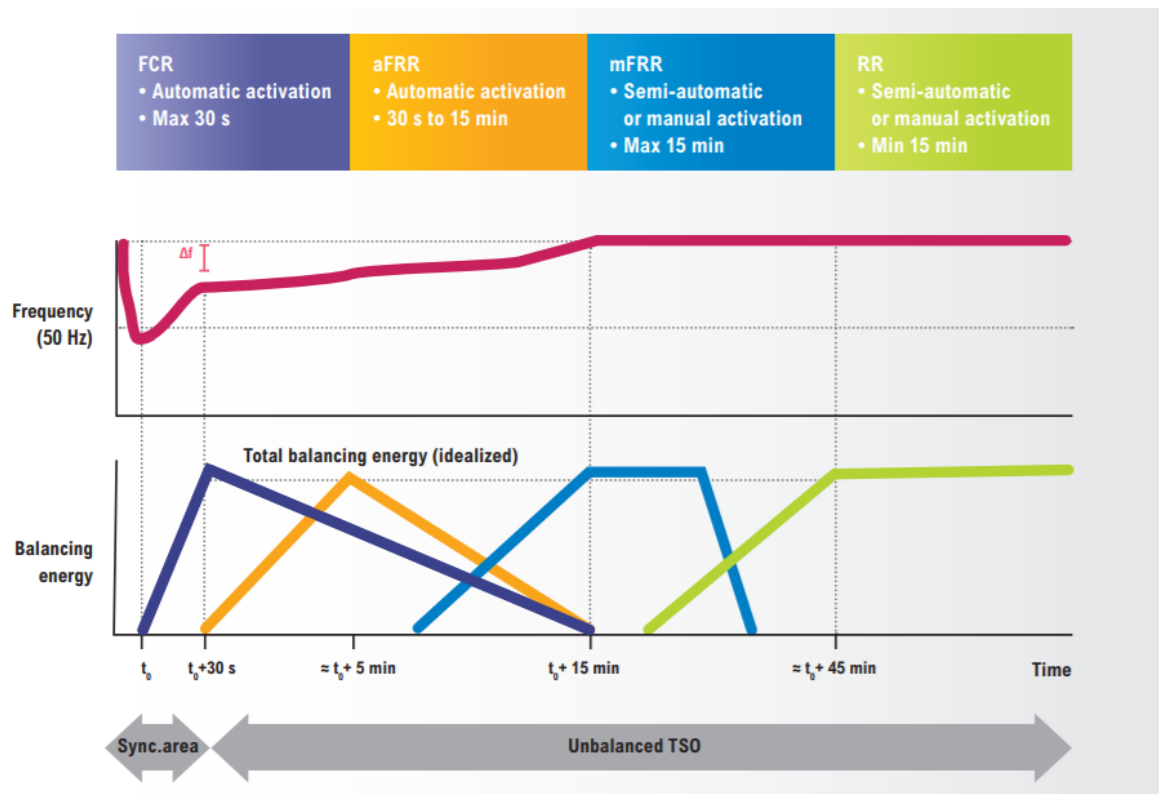
1 Bevezetés

Az előrejelzés tervezhetőséget jelent, amely nagyobb mértékű rugalmasságot és alacsonyabb üzemeltetési költségeket eredményez. Ez a gondolat szolgált alapjául jelen kutatásnak, amelynek célja egy olyan eszköz előkészítése segítő elemzésekkel, amely rövid távon jó pontossággal becsülni képes a várható rendszerszintű kiegyenlítetlenséget, a következő néhány évben bevezetendő össz-európai platformos - MARI (Manually Activated Reserves Initiative) [1] és PICASSO (Platform for the International Coordination of Automated Frequency Restoration and Stable System Operation) [2] - rendszerben (továbbiakban: platformos időszak), illetve igény szerint az átállásig. A tesztelt algoritmusok, illetve a tapasztalatok alapján javasolt módszertan a proaktív mFRR szabályozás magját képező modul szerepét töltene be, oly módon, hogy a korábbi reaktív eljárásrend helyett a becslés alapján előre tervezve lehessen aktiválni tartalékokat, lehetőség szerint csökkentve az aFRR szabályozási igényeket, illetve a keresztszabályozás mértékét.

1.1 Szabályozás szintjei, fogalmi

A villamosenergia rendszer egyensúlyának fenntartása érdekében a TSO-k különböző kiegyenlítőenergia termékeket tudnak igénybe venni [3]. Ezek közül a leggyorsabb az FCR – Frequency Containment Reserve, amely a frekvencia változását érzékelve 30 másodpercen belül automatikusan aktiválódik, a frekvenciaváltozás megszüntetése érdekében. A frekvencia megfelelő munkapontba való visszaállításáért felelősek az aFRR – automatic Frequency Restoration Reserve, illetve mFRR – manual Frequency Restoration Reserve kapacitások, amelyek között irányítástechnikai, és reakció, illetve terhelés változtatási sebességbeli különbségek vannak. Előbbi egy automatikusan vezérelt termék, amely percek alatt képes az igényelt teljesítményszintet elérni, míg utóbbi jellemzően manuálisan, telefonon keresztül diszpécseri utasításra indítható, az igényelt teljesítményt pedig maximum 15 perccel az utasítást követően érik el. A régebbi (primer, szekunder, tercier) tartalékdefiníciókkal ezek a termékek csak részben feleltethetők meg; az FCR vette át a primer, az aFRR a szekunder szabályozás szerepét, az mFRR azonban a szekunder és a tercier között található, lényegében egy gyors tercier tartaléknak minősül. A rendszer egyensúlyi helyzetének visszaállítása után annak gazdaságos üzeméért felelősek tehát az mFRR mellett az RR – Replacement Reserve tartalékok, amelyek hosszabb felfutási idővel, de kisebb egységköltséggel jellemezhetők. Utóbbi kapcsán érdemes megemlíteni a TERRE projektet [4], amely az RR szabályozás nemzetközi együttműködésért felelős, azonban ebben Magyarország csak megfigyelőként vesz részt.

Az 1. ábrán láthatók a szabályozási termékek jellemző aktiválási idejei egy frekvenciaváltozással járó esemény hatására. Maga az ábra a rendszert elvi szinten mutatja be, ugyanakkor az is elmondható, hogy az esetek csak igen kis részében szorul rá a hazai villamosenergia rendszer az mFRR, illetve különösen az RR aktiválásokra, hiszen egyéb hatások, lehetőségek is jelen vannak a rendszerben.



1. Ábra: Szabályozási termékek elvi bemutatása egy frekvenciaváltozást okozó esemény esetén[3].

Együttműködő rendszerek esetén nem pusztán a lokális menetrendi eltéréseket kell szabályozni, hanem az adott TSO felelősségébe tartozó import-export menetrendi eltéréseket is. Az ACE – Area Control Error azt mutatja meg, hogy mennyi a maradó szabályozási hiba, amelyet az import-export szaldó menetrend-tény különbségének és a frekvenciaeltérés hibajelének összegzésével kapunk meg. Amennyiben az ACE egy adott országban éppen nulla, és az import-export szaldóban eltérés van, az azt jelenti, hogy az eltérésért a másik ország (rendszerirányító) a felelős, így neki kell azt kiszabályoznia. Pozitív érték esetén fel-, negatív érték esetén pedig leszabályozás szükséges.

Az együttműködő országok kiegyenlítetlenségének „elnettósítása” az IGCC – International Grid Control Cooperation [5], amely célja a helyi szabályozási igények csökkentése a szomszédos szabályozási területek ellentétes irányú eltéréseinek kihasználásával. Ezáltal az adatsorban is szereplő pozitív, illetve negatív irányú aktivált aFRR szabályozási energia valójában az igénybe vett hazai aFRR termékek és az IGCC elnettósítással kapott, vagy adott volumenek összege. Az IGCC nettósítást alapvetően egy nemzetközi platform végzi, amely 4 másodpercenként fut le; az egyes országok kiegyenlítetlenség adatait először összerendezi, és a fennmaradó határkeresztező kapacitások mentén optimalizálja. Ezáltal, ha két országban ellentétes irányú szabályozás lenne szükséges, és közöttük rendelkezésre áll átviteli útvonala, akkor azon keresztül megtörténik az igények elnettósítása, amely mindkét ország számára kedvezőbb, mintha ellentétes irányú aFRR szabályozást vettek volna igénybe. Az IGCC platformon 2 másodperc alatt megtörténnek ezek a tranzakciók, amely során egyszerre akár több országgal is lehetséges az említett együttműködés. A kezdeti kiegyenlítetlenség és az elnettósított volumenek különbsége lesz az a mennyiség, amelyet az együttműködő (így tehát a hazai TSO is) majd a második 2 másodperces ciklusban aFRR aktiválással szabályoz.

Lényegében tehát az IGCC a rendszerszabályozás és ár szempontjából is kedvező, ugyanakkor mennyisége - és ezzel együtt rendelkezésre állása - előre nem ismert, bizonytalan.

1.2 Proaktív mFRR szabályozás célja

A proaktív mFRR szabályozás megjelenik az ENTSO-E 2018-as egységesített kiegyenlítőenergia fogalmakhoz és célokhoz tartozó magyarázó segédletében [6], illetve Haberg és társai cikkében [7], amely lényegében a rendszer operatív megvalósításáról, valamint pénzügyi értékeléséről szól. Mivel az mFRR szabályozás lassabban indítható, mint az aFRR, nem elég az aktuális rendszeradatokat látni, hanem azok értékelésén alapuló előre tekintésre, becslésre van szükség, amely alapján a rendszerirányító döntést tud hozni, hogy érdemes-e elindítani egy tartalék erőművet, vagy sem. Jelenleg a döntés előkészítési folyamata nem különül el, ez is a diszpécser feladata. Az mFRR szabályozás céljai között kiemelendők a következő szempontok:

- **Alacsony aFRR tartalék:** A hivatalos (SOGL szerinti [8]) megfogalmazás alapján: „Ha az automatikus FRR kapacitás az üzleti napon 30 percen túl az 1. szintű FRCE célparaméter teljesítéséhez szükséges tartalék mennyiség alá csökken, kézi FRR vagy RR aktiválásával gondoskodik az automatikus FRR tehermentesítéséről. Az 1. szintű FRCE célparaméter teljesítéséhez szükséges minimális tartalék mennyisége a 12. cikk (1) bekezdés (a) pontjában leírt, a VER historikus kiegyenlítetlenségeinek vizsgálata során összegyűjtött kiegyenlítetlenségek 70. percentilis értéke” [9]. Vagyis, nem megengedhető, hogy az aFRR tartalék mértéke 3 negyedórán keresztül az előírt minimum alá csökkenjen. Ilyenkor mFRR tartalékokat kell igénybe venni az aFRR szabályozói kapacitás helyreállításához.
- **Gazdasági érvek:** amennyiben a jelen rendszerben ismert mFRR és aFRR árak közül, a kockázati faktort ismerve, előbbiek megfelelően olcsóbbak, úgy pénzügyi alapon érdemes a várható kiegyenlítetlenség egy részét mFRR szabályozással kiváltani. Árnyalja a probléma egyszerűségét az IGCC, amely jellemzően olcsóbb, mint az aFRR és az mFRR tartalékok, így fontos, hogy lehetőleg proaktív mFRR szabályozással csak aFRR-t váltsunk ki. Az említett platformos rendszerben ez a szempont várhatóan még összetettebb lesz a termékek ár-bizonytalansága miatt. (A PICASSO platformnak része lesz az IGCC, illetve az utódja, az IN (Imbalance Netting)).
- **Diszpécseri döntés:** a rendszerirányító diszpécseri által észlelt komplex eseményrendszer alapján dönthetnek úgy, hogy mFRR tartalékot kell igénybe venni, olyankor is, amikor az pusztán az előbbi szempontok alapján nem egyértelmű.

Jelenleg az első és a harmadik pont lényegében hasonló, ugyanis pl. az erőművi egységek kiesése miatti mFRR aktiválás is visszavezethető az alacsony aFRR tartalék szint korrigálására. Ugyanakkor a jelen esetben érdemes elkülöníteni abból a szempontból, hogy ezek a döntések olyan ismereteken is alapulhatnak, amelyek a vizsgált előrejelző modulban nem kezelhetők.

A gazdasági vonalon érdemes megemlíteni, hogy jelenleg az mFRR ajánlatokra 112 000 Ft/MWh, míg aFRR esetén 500 000 Ft/MWh az ársapka, tehát ezek az értékek jelentik a beadható ajánlatok árának felső határát. Ez alapján, illetve abból az alapgondolatból kiindulva, hogy az aFRR egy gyorsabb, rugalmasabb termék, az következne, hogy drágább, mint az mFRR, azonban ez jellemzően nem ilyen egyszerű. Az aFRR terméket szolgáltató erőművek általában forgógépek, míg az mFRR esetén gyakran álló erőművekről van szó, így utóbbiak esetében az igénybevétel egy indítási ciklussal jár együtt, amely még a modern erőművi egységek esetén is komolyabb élettartam csökkenéssel jár, mint egy üzemben lévő forgógép pár MW-os terhelésváltoztatása, még ha az nagyobb gyakorisággal történik is. Így tehát a

MAVIR Zrt. transzparencia adatpublikációi alapján az a következtetés, hogy a gazdasági elemzéshez jelenleg irányonkénti közös aFRR-mFRR merit order szükséges, nem lehet általánosan megfogalmazni prioritást pénzügyi alapon. A következő években a platformos rendszer (csatlakozásunk előtti) kezdeti időszakának adataiból, illetve a jelenlegi rendszer fokozatos átalakulásának tapasztalataiból pedig majd következtetni kell a termékek egymáshoz képesti árazására. Emellett azt is fontos kiemelni, hogy a jelenlegi szabályzat szerint az árak (ID piac) óránként változhatnak.

A fenti gondolatok megértéséhez készült a 2. ábra, amelyen látható egy historikus adatokon alapuló mFRR, aFRR, valamint a kettőből képzett közös FRR merit order lista. Mintaként jelöltem a tevezett előrejelzésből származó adott konfidenciaszintű intervallum, illetve pontbecslést, amelyek jól szemléltetik a proaktív mFRR szabályozás céljait.



2. Ábra: Minta mFRR, aFRR és közös merit order lista, jelölve az előrejelzéssel kapott kiegyenlítetlenség intervallum és pontbecslés értékeit.

Egyrészt a vízszintes tengelyen látható teljesítmény adatokból látszik, hogy a kiegyenlítetlenség várható értéke mellett mekkora lenne a fennmaradó aFRR tartalék (az IGCC-t figyelmen kívül hagyva). Amennyiben ez az érték egy adott szint alatt van (az ábrán példaképp

100 MW), úgy az előrejelző modul figyelmeztetést adhat a diszpécsereknek, hogy számítsanak magas igénybevételre, így előre tudnak vele kalkulálni, amely remélhetőleg kevesebb keresztszabályozást eredményezne, mint az egy-két hiányos negyedóra után, a harmadik elkerülése végett történő mFRR aktiválás, amely addigra már nem feltétlenül lenne indokolt a trendek esetleges változása miatt.

Másrészt az egyedi listákhoz képest a közös merítésből láthatóan kedvezőbb árakat lehet elérni, hiszen a bemutatott példában, akár már 50 MW körüli mFRR aktiválással el lehet kerülni az ársapka közelében lévő aFRR ajánlatok lehívását. Amennyiben egy ilyen órában leszabályozási igény lett volna, úgy a legkedvezőbb mFRR ajánlatok helyett szintén többszörös áron valósult volna meg a szabályozás pusztán aFRR-ből.

1.3 Kiegyenlítetlenség előrejelzés szakirodalmi előfordulásai

A villamosenergia rendszerek tekintetében a 2020-as évek egyik meghatározó fejlesztési iránya a flexibilitás, amely szükséges az időjárásfüggő megújulók rendszerintegrációjához. A rendszer rugalmasságának igényét részben az adja, hogy nem lehet pontosan megadni a fogyasztást, illetve az utóbbi időben a beépített fotovoltaiikus, szél- és más intermittens erőművek hatására a termelés becslését is egyre nagyobb pontatlanság övezi. Erre reagálva az említett technológiák termelésének [10], [11], illetve a fogyasztásnak [12] az előrejelzése is erősen kutatott témává vált, aktívan alkalmazva a gépi tanulást, kiváltva a fizikai modelleket [13]. Ezzel szemben, a kiegyenlítetlenség előrejelzés témakörében meglepően kevés publikáció készült eddig.

Garcia és Kirschen (2004) munkájukban [14] ARIMA (Auto-Regressive Integrated Moving Average) és ANN (Artificial Neural Network) modelleket hasonlítottak össze egy hétre (négyórás intervallumokra), illetve egy hónapra előre (napi átlagra) történő előrejelzésre. Bemenő adatként alkalmaztak fogyasztás előrejelzést és annak hibáját, elfogadott licit és ajánlati mennyiségeket, határidős ügyleteket, kapuzárási egyensúlyhiány volumeneit, az ajánlatok összes és elfogadott mennyiségét, kiegyenlítőenergia árakat és a hét napjait. A publikált anyagból egyértelműen nem derül ki, hogy használtak-e historikus kiegyenlítetlenség adatokat prediktorként. Az eredmények ismertetése kapcsán megfogalmazták, hogy a neurális hálók jobban teljesítenek, mint a lineáris regresszió alapú modellek, azonban a tesztelt hálózat típusok közül nem volt olyan, amely mindegyik időjárási körülmény és a kétféle teszt módszertan szerint egyértelműen jobb lett volna a többinél.

Kratochvíl doktori disszertációjában [15] foglalkozott kiegyenlítetlenség előrejelzéssel. Kutatásában a cseh villamosenergia rendszert vizsgálta, ugyanakkor megállapította, hogy a nemzetközi összekapcsolt hálózat miatt a környező országok is komoly hatással bírnak a helyi egyensúlyra. A kiegyenlítőenergia volumenek mellett annak árát is vizsgálta, magyarázó változóként felhasználta a napon belüli (intraday) piaci árakat, teljes valós és jósolt fogyasztást, illetve ezek különbségeit, kiegyenlítőenergia piaci árakat, a német PV és szélenergia termelési adatokat, valamint a termelés és import-export menetrendekeket. Órán belüli előrejelzéseikhez nem használt magas szintű tanuló algoritmusokat, egyszerű lineáris modellezéssel készítette azokat, viszonyításként pedig ARIMA-t használt, magát a kiegyenlítetlenség becslést pedig klasszifikációs problémaként értelmezte, tehát diszkrét, ötféle kategóriába sorolt kimeneteket produkált, nem folytonosakat. Ugyanígy a bemenő adatokat is kategorizálta. Megállapította többek között, hogy a két óránál régebbi adatoknak már kevés magyarázóereje van, ugyanakkor ezt az állítást csupán egyszerű korreláció analízisre alapozta.

Contreas MSc diplomamunkájában (2016) [16] vizsgálta a területet, nem rendszerirányítás, hanem villamosenergia kereskedési stratégia szempontjából, kiemelten azzal a céllal, hogy időjárásfüggő megújulókat is menedzselő kereskedő cégek a másnapi menetrendjükből adódó pontatlanságot becsülve olyan napon belüli piaci pozíciót vegyenek fel, amellyel a kiegyenlítőenergia költségek csökkenthetők. Ehhez Random Forests (véletlen erdők) alapú órás kiegyenlítetlenség előrejelzést használt, melyet genetikus algoritmussal optimalizált. Alkalmazott magyarázó változói között szerepeltek fogyasztás és szél és szélergia gradiens előrejelzések, hőmérséklet, másnapi aukciós volumen és ár adatok, valamint kiegyenlítőenergia típusonkénti fel- és leirányú akciók eredmények, technológiai kényszerek. Ezesetben egyértelműen kiderült, hogy nem használt historikus kiegyenlítetlenség adatokat prediktorként. Elemzésében a hagyományos hibaszámítási módszerek mellett a tesztidőszakban megtakarított kiegyenlítőenergia költségeket is kiszámolta, modelljeit, melyek különböző bemenő adatkombinációkra épülnek, e mentén rangsorolta.

Tárik S. és társai cikkükben (2019) [17] röviden bemutatták a témában elkezdett munkájukat. Kutatásukban, a világon elsőként publikálva kvantilis regresszió alapú intervallum becslést vizsgáltak, amelyet a két évnnyi adatsor munkapontjait követő kétórás időtávjainak, ötperces intervallumaira végeztek el. Az előrejelzéseket a norvég energiapiacra végezték el, magyarázó változóként használták a historikus kiegyenlítetlenség adatokat, időbeli tulajdonságokat (pl. nappályas aktuális magassági szöge), energiapiaci árakat, HVDC áramlási, termelési és terv adatokat, illetve előrejelzés és mért környezeti jellemzőket a hőmérséklet, szél, valamint folyami áramlások vonatkozásában. Eredményeik között megfogalmazták, hogy a kiegyenlítetlenség meglepően erős és bizonyos mértékig előre látható mintákat mutat, illetve a tisztán historikus kiegyenlítetlenség adatok alapján készített előrejelzéseik is jobbnak bizonyultak a TSO akkori módszereinél. A kutatás kapcsán érdemes megjegyezni, hogy kifejezetten ígéretesnek tűnik, azonban sajnos 2019 óta nem jelent meg a teljes anyag az interneten, így egyelőre csak egy megkezdett munka rövid, tartalmilag szűkösen publikált kivonata, amelyből önmagában komolyabb következtetéseket a hazai rendszerre nem lehet levonni.

A felsorolt anyagokon kívül az Elia, az egyik belga TSO kapcsán lehet megemlíteni nyilvános aktivitást a témában. A honlapjuk szerint fut náluk egy projekt [18], amely mesterséges intelligencia segítségével fog rövidtávú rendszer kiegyenlítetlenséget előrejelezni a tartalékok aktiválásának javítására. Ezzel kapcsolatban 2021 nyarán az első eredményeiket ismertetve [19] hirdettek nyilvános konzultációt. A publikált jelentésből kiderül, hogy a céljuk a fejlesztés alatt álló előrejelző egységnek 2022 Q2-re a PICASSO, majd 2022 Q4-re a MARI környezetbe történő implementálása. Az anyagban alapvetően lineáris, ANN és SVM modelleket hasonlítottak össze, illetve rekurzív változó kiválasztással (Recursive Feature Extraction) vizsgálták a prediktorok mennyiségének csökkentési lehetőségeit. Egy évnnyi negyedórás adaton dolgoztak, amelyből a változókat egy napra visszamenőleg vizsgálták. Ez azt jelenti, hogy minden típusból 96 értéket vettek figyelembe, amely igen komoly számítási igényeket eredményezhetett. A változók között szerepelt a kiegyenlítetlenség, fel- és leirányú aktivált tartalékok mennyisége és ára, valós idejű és intraday terhelés előrejelzés, illetve utóbbi hibája, szél és napenergia termelésre vonatkozó adatok, IGCC, valamint nettó szabályozási volumenek. A változók kiválasztása mellett vizsgálták a tanulási időt is, amely kapcsán megállapították, hogy nagyjából 8 hétnyi adat ismeretében már jó közelítéssel eléri a határértékét a pontosság, az alatt azonban jelentősen gyengébb eredmények születtek. Magát az elemzést zárt időtartamon keresztvalidációs futtatással készítették egyébként, tehát a rendszer változásainak lekövetésére nem feltétlenül vonhatók le releváns következtetések a munkából.

Az anyag sok pozitívumot és hasznos tapasztalatot tartalmaz, azonban meg kell említeni, hogy az az eredmény, miszerint a lineáris módszer lett a legjobb, felvet kérdéseket, hiszen más alkalmazási területen nem, vagy csak elvétve található olyan eset, hogy a neurális háló, vagy az SVM ne tudná felülmúlni azt, így felmerül a kérdés, hogy vajon megfelelően lettek-e illesztve a módszerek. Nem publikálták, hogy milyen programozási nyelvet, könyvtárat alkalmaztak, így erre bővebben nem tudok reflektálni. Továbbá érdemes megemlíteni, hogy a jelentés nem tudományos célú, korábbi területi és módszertani szakirodalomra nem tesz utalást.

1.4 Áttekintés

A kutatás fő célja a platformos környezetre való felkészülés segítése a rövidtávú kiegyenlítetlenség előrejelzés proaktív mFRR szabályozás céljából történő felhasználása, amelyhez céljában és peremfeltételeiben egyaránt hasonló elemzést legjobb tudomásom szerint még nem publikáltak a nyilvánosság előtt. A 2. fejezetben az alkalmazott módszertant, azon belül a felhasznált adatokat, a gépi tanulást általánosságban, és specifikusan az elemzésben felhasznált algoritmusokat, majd a pontbecsléseket és a platformos keretrendszert, valamint az újratanítás eszközeit mutatom be. Ezt követően a 3. fejezetben a „nyers” adatok elemzésének tanulságai, a megállapított összefüggések, valamint az ezekből levont következtetések és azok felhasználásai olvashatóak. A 4. fejezetben ismertetem a kapott eredményeket a pont-, majd a kvantilis regresszió alapú intervallum becslésekre, összehasonlítva a különböző módszereket, prediktor kombinációkat, illetve utóbbi esetén az aktiválási protokollok változtatásával elért maximálisan aktiválásra javasolt mFRR mennyiségeket.

2 Módszertan

Ebben a fejezetben elsőként bemutatom a felhasznált adatokat, amelyet a gépi tanulás általános, illetve a vizsgált algoritmusok, majd a hiperparaméterek rövid áttekintése követ. Ezután a pontbecslések céljait, illetve módszertanát, majd a platformos környezet felépítését és rendszerét ismertetem. Végül az újratanítással kapcsolatos megfontolások, illetve az alkalmazott módszertan bemutatása olvasható.

2.1 Adat előkészítés és használmagyarázó változók

Az elemzésben használt adatokat a MAVIR Zrt. biztosította a 2020.01.01 és 2021.10.06 közötti időszakra különböző adatgyűjtő rendszereiből lekérdezve. Az általános módszertan az, hogy az adott időbéllyeghez az azt megelőző ötperces intervallum átlagos értékei vannak hozzárendelve. A célértékként kezelt kiegyenlítetlenségnek nevezett mennyiség alapvetően, három fő összetevő:

1. az aFRR (Automatikus) szab. (5p),
2. az mFRR és RR (Nem automatikus) szabályozás mértéke (kiegyenlítő célú),
3. illetve az ACE mért érték

nevű kimutatások irányfüggő összegzésével áll elő. Az irányfüggés azért vettem bele a képletbe, mivel még a viszonylag rövidnek számító ötperces időszámban is gyakori a domináns iránnyal ellentétes kiegyenlítőenergia igény, azonban a tartaléktervezés szempontjából a szélsőségek kezelése a fontos, így ezt a jelenséget kis volumene miatt elhanyagoltam. Az említett aFRR adatok a hazai szinten aktivált aFRR szabályozói teljesítmények és az IGCC összege. Az elemzéshez használt adatsor az előzőekben említett kiegyenlítetlenség komponensek mellett tartalmaz különböző prediktorokat, amelyeket a következőkben ismertetek.

A napenergia adatok közül felhasználásra került a MAVIR Zrt. beszállítói day-ahead és intraday előrejelzése, az üzemirányítói mérésből származó teljesítmény adatok, illetve ezek különbségéből képzett várakozás-tény eltérések. Továbbá a nemzetközi kereskedelem modellezése érdekében a vizsgált változók között megtalálhatóak az import-export menetrendek, a bruttó rendszerterhelési tervek, az ACE, illetve az IGCC adatsorok elkülönítve az aFRR adatoktól. Az említett adatok a MAVIR Zrt. adatpublikációs felületén a megfelelő kategóriát kiválasztva [20] jellemzően 15 perces bontásban elérhetőek mindenki számára.

Az elemzés során ezekből az adatokból választottam ki a prediktorokat, standard skálázást alkalmaztam rajtuk, illetve bizonyos esetekben további változókat képeztem belőlük. A magyarázó és magyarázott változókat a következő táblázatban foglaltam össze, feltüntetve az alkalmazott jelöléseket és a hozzá tartozó meghatározásokat. A gyakorlatban ezek a változók, illetve az adatok a $T = 0$ időponthoz tartozó időbéllyeghez (pl. 2021.04.07 19:45:00) rendelve szerepelnek a név oszlopban feltüntetett fejléccel, táblázatos formába rendezve.

1. Táblázat: A használt magyarázó változók bemutatása.

Név	Magyarázat
Magyarázó változók	
NoD	Megadja, hogy a nap hányadik ötperce az adott időponthoz tartozó intervallum.
Solar prog DA	Napelemparkok termelésére vonatkozó szolgáltatói negyedórás előrejelzésekből interpolált ötperces átlagok day-ahead (DA) és intraday (ID) időhorizonton, valamint
Solar prog ID	
Solar prog	

Solar prog diff	rendre a kettő átlaga, különbsége, illetve előbbi napi
Solar daily mean schedule	átlagértéke.
Solar net. measured	Napenergia termelésre vonatkozó valós idejű nettó üzemirányítói mérés.
Solar schedule error	Az üzemirányítói mérés és a szolgáltatói DA és ID előrejelzések átlagának különbsége a jelzett intervallumokra, illetve ezek átlaga.
Solar schedule error T-35	
Solar schedule error T-20	
Solar schedule error T-5	
Solar schedule error mean	
IGCC Fel	A fel- és leirányú IGCC, a kettő összege, valamint korábbi időpontokban mért értékei.
IGCC Le	
IGCC	
IGCC T-50	
IGCC T-45	
IGCC T-40	
...	
IGCC T-15	
IGCC T-10	
IGCC T-5	
Import-export	Az import-export menetrendek aktuális és tervezett/ következő órák értéke, valamint a következő és az aktuális óra menetrend szaldójának különbsége, illetve ugyanezen érték az IMPEX D esetekben súlyozva az adott munkapont órában belüli helyzetével, valamint ennek időben eltolt értékei.
Import-export plan	
Import-export diff	
Impex D	
Impex D-2	
Impex D-4	Az adott munkapontban rögzített ACE értékek.
ACE	
Bruttó terv rszt.	A bruttó terv rendszerterhelés adatsor, illetve az abból képzett gradiensek, rendre a következők szerint előállítva:
Bruttó terv rszt. grad	
Bruttó terv rszt. grad2	
Bruttó terv rszt. grad3	
Bruttó terv rszt. grad4	
Imbalance	A fejezet első bekezdésében kifejtett módon (az ACE, IGCC és hazai aFRR, valamint mFRR és RR aktiválások előjelfüggő összegzésével) előállított kiegyenlítetlenség értékek az adott munkapontban, illetve az azt megelőző 125 percben.
IB T-125	
IB T-120	
IB T-115	
...	
IB T-15	
IB T-10	
IB T-5	<i>Becsült változók</i>
IB T--5	
IB T--10	
IB T--15	
...	
IB T--50	
IB T--55	
IB T--60	

2.2 Gépi tanulás, használt algoritmusok

A gépi tanulás alapvetően egy eszköztár, amely a műszaki gyakorlatban is használt problémák megoldására használható, az általánosan alkalmazott fizikai modellezéstől eltérő adat-alapú, statisztikai módszerekkel. Egyszerű problémák esetén viszonylag könnyű általános fizikai törvényekre visszavezetett, vagy empirikus úton megállapított matematikai összefüggéseket alkalmazni, bonyolultabb rendszerek esetén azonban a sorba kapcsolt modellek egyedi pontatlansága, egy-egy mért, vagy becsült tag bizonytalansága, olyan eredő hibát tud okozni, amely nem feltétlenül elfogadható. Ilyen esetekben érdemes vizsgálni a például a felügyelt tanulás (supervised learning) [21] kategóriájába tartozó tanuló algoritmusokat, amelyek célja általánosságban bizonyos változók, az úgynevezett prediktorok, és a hozzá tartozó kimeneti mennyiségek ismeretében meghatározni az adott algoritmusra jellemző összefüggés ideális paramétereit, hogy az majd később új bemenő adatkombinációk esetén minél jobb becslést tudjon adni a keresett értékekre. Tehát az alkalmazás jellemzően két fázisból áll, a tanulás, amely során a különböző elven összefüggéseket kereső algoritmusok felparaméterezése történik, valamint a tesztelés, amely során egy elkülönített adattartományon kipróbálva a tanult algoritmust, kiderül, hogy az illeszkedés valójában mennyire pontos. A gyakorlatban ezt a két lépést ki szokás egészíteni hiperparaméter hangolással [22], amely célja az algoritmusok külső paramétereinek, egyfajta vázának manuális (valójában gépi, de nem automatikus) beállítása.

Az algoritmusok, illetve azok hiperparamétereit elsőként 2020-as adatokon kerültek tesztelésre véletlenszerűen kiválasztva az év során elszórt tanuló és teszt adatok segítségével. A véletlenszerű mintavételezéssel terv szerint a teljes időtartományra érvényes, reprezentatív megállapítások tehetők. Ugyanakkor az elemzés másik fele idősoros felosztásban készült, amely azt jelenti, hogy a teszt adatsor (2021.04.07-2021.10.04) a tanuló (2020.04.07-2021.04.06) után található, életszerűen, azzal a céllal, hogy az előrejelzés azon tulajdonságát lehessen megfigyelni, hogy mennyire képes a rendszerszintű változásokat túlíllesztési hibák nélkül kezelni. Ez kritikus fontosságú, hiszen habár a hazai primerenergia felhasználás nagyjából stagnál [23], a villamosenergia fogyasztás folyamatosan növekszik [24], szintúgy a nemzetközi piacokon kereskedett volumenek [25], illetve a napelemtermelés [24], ezért várhatóan új adatsort mutatva a tanult algoritmusnak, napról-napra lesznek az ismert tartományon kívül eső értékek.

Az előrejelzésekhez tesztelésre került 24 algoritmus, többnyire a Python *scikit-learn* könyvtárából [26], amelyekből a számítási igények, és első eredmények alapján bővebb elemzésre tíz módszer került kiválasztásra. Ezek listája megtalálható a következő táblázatban egy rövid leírással, illetve javasolt irodalommal együtt.

2. Táblázat: Összefoglaló a tesztelt algoritmusokról.

ID	Python név	Rövid leírás	Cikk
lr	LinearRegression	A magyarázott változót a prediktorok értékének lineáris függvényével közelíti, amely együtthatóit a tanítás fázisában optimalizálja a legkisebb négyzetek elvén.	
tr	TheilSenRegressor	Nem paraméteres módszer, ami az jelenti, hogy nem feltételez eloszlást az adatok mögött. Mivel a Theil-Sen egy medián alapú becselő módszer, megbízhatóbb a sérült, vagy kiugró adatokkal szemben, a közönséges legkisebb négyzetes hibák	[27]

		elven alapuló lineáris modellekhez képest. A probléma dimenzióinak növelésével ezen erősségéből vesz.	
huber	HuberRegressor	Legkisebb négyzetes hibák elven alapuló lineáris modell, amely alfa és epsilon paramétereivel csökkenti a hibás, vagy kiugró adatok befolyását, anélkül, hogy teljes mértékben ignorálná hatásukat.	[28]
knn	KNeighborsRegressor	A keresett értéket a magyarázó változók "szomszédainak" kimenetei alapján interpolálja, amely eljárás során a figyelembe vett adatpontok száma állítható a <code>n_neighbors</code> paraméterrel.	[29]
mlp	MLPRegressor	A többretegű perceptron regresszor egy előrecsatolt neurális hálózat (feedforward network), amely tanítása hiba visszaáramoltatás módszerével történik (back-propagation).	[30]
rf	RandomForestsRegressor	Az egyik legelső ensemble módszer, meghatározott számú döntési fát illeszt az adatok különböző részeire, amelyek becsléseinek átlagát adja vissza kimenetként.	[31]
et	ExtraTreesRegressor	A Random Forests-hez hasonló ensemble módszer (az almintákra illesztett randomizált döntési fák átlaga), de megnövelt véletlenszerűséggel, hogy csökkentse a szórást, az elfogultság rovására.	[32]
gbr	GradientBoostingRegressor	A GradientBoosting (gradiensnövelés) egy gépi tanulási technika regressziós és osztályozási problémák megoldására, amely előrejelzési modellt hoz létre gyenge előrejelzési modellek, jellemzően döntési fák együttese formájában. Lépésről lépésre építi fel a modellt, mint más boosting módszerek, és általánosítja őket, lehetővé téve egy tetszőleges differenciálható veszteségfüggvények optimalizálását.	[33]
xg-boost	XGBRegressor	Az előzőekben ismertetett scikit-learn algoritmusoktól elkülönülő, önálló könyvtárral rendelkező módszer, amely gradient boosting eljárás alapján alapul, célja számítási határoknak szélsőségeinek fejlesztésére a skálázhatóság, hordozhatóság és a pontosság növelése érdekében.	[34]
light-gbm	LGBMRegressor	A Microsoft által fejlesztett, szintén önálló könyvtárral rendelkező gradient boosting algoritmus, amely hisztogram alapú tanítást használ a tanítási idő és memória használat csökkentése érdekében.	[35]
mid	-	Az MLP és az XGBoost algoritmusok előrejelzéseinek átlagolásával képzett becslés.	-

Az előrejelzések több kimenetű regressziós keretrendszerrel készültek (MultiOutputRegression), amely lényegében a 12 előrejelzett intervallumra külön, párhuzamosan tanította az egyes algoritmusokat. Jelen esetben egy 12 magos, 24 logikai szál processzor volt a számításokért felelős egység, így a párhuzamos futtatás lehetséges volt,

ugyanakkor éles használat esetén ekkora számítási kapacitás mellett számítani kell komolyabb memória igényre is, amely rossz tulajdonsága, hogy rendszerint hiány esetén a futtatás egyszerűen leáll, nem pedig csupán lassabban fut le. Az említett első körös kiválasztási folyamatból ezért kerültek eltávolításra ígéretes algoritmusok, többek között a Kernel Ridge [36] és a Support Vector Machines [37], illetve az Extra Trees és a Random Forests stabil futtatása a tanítás során 16 GB RAM mellett csak ritkított adatsoron volt lehetőség.

2.3 Hiperparaméter hangolás

Az előzőekben említett célból a vizsgált algoritmusok tesztelésre kerültek különböző hiperparaméter beállításokkal. Erre az eljárásra többféle algoritmus rendelkezésre áll a Scikit-learn könyvtárban [38] például, többek között a CustomGridCV, valamint a RandomGridCV, amely szóösszetételekben az első tag utal arra, hogy a definiált paraméter hálóból milyen elven válassza ki az algoritmus a tesztelendőket, míg a végén a CV a kereszt validációra utal, hiszen fontos szempont, hogy a legjobb kombinációt ne csak egyféle tanulás-tesztelés felosztáson futtatva válasszuk meg, hanem több különböző eseten elért átlagos teljesítmény alapján. A gondolatmenet alapján tehát a hangoláshoz szükséges egy paraméter háló, a dokumentáció alapján, a futási idők ismeretének megfelelő sűrűséggel megválasztva, egy elv, amely a “Custom” esetben szimplán az összes kombinációt jelenti, míg a “Random” előtagnál véletlenszerű mintavételezést a lehetőségek közül, kiegészítve egy iterációs számmal, amely megadja, hogy hány minta legyen tesztelve. Előbbi esetén az összes kombináció, utóbbinál pedig az iterációs szám és a kereszt validációk számának szorzata lesz arányos a várható futási idővel. Jelen kutatásban a véletlenszerű („Random”) módszer került alkalmazásra, a pontbecsléseket végző algoritmusokra. A különböző időtávokra történő futtatás, a keresztvalidáció, és a kombinációk számának szorzatai a folyamat párhuzamosítása mellett is extrém számítási időket jelentős korlátokat szabtak ennek az eljárásnak. Természetesen léteznek más optimalizációs elvek [39], illetve olyan paraméterek, amelyek eddig nem kerültek tesztelésre, így ezen a területen igény esetén még lehet vizsgálatokat végezni.

A kvantilis előrejelzések esetén nem történt effektív hangolás, mivel az mFRR aktiválásra továbbvezetett célfüggvények túl komplexek. A tapasztalat azt mutatta, hogy az RMSE alapú hangolás, habár csökkentette a konfidencia intervallum átlagos szélességét, ugyanúgy csökkentette annak valós megbízhatóságát.

2.4 A pontbecslések értékelési szempontjai

Az pontbecslés célja az, hogy a valós kiegyenlítetlenséget minél pontosabban előrejelezzék, illetve annak irányát lehetőleg minden esetben eltalálják. Ezekre a célokra az alábbi hiba, illetve pontossági mutatók kerültek alkalmazásra:

$$MAE = \sum_{i=1}^N |y_i - IB_i| \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^N (y_i - IB_i)^2} \quad (2)$$

$$Skill\ score = \left(1 - \frac{RMSE}{RMSE_{perzisztencia}}\right) \cdot 100 [\%] \quad (3)$$

$$dir\% = \frac{\sum_{i=1}^N (y_i \cdot IB_i > 0) = IGAZ}{N} \cdot 100 [\%] \quad (4)$$

A fenti képletekben y_i a becslült, IB_i pedig a valós kiegyenlítetlenséget jelöli, N a tesztelt munkapontok száma, a *perzisztencia* pedig jelen esetben egy mozgóátlag alapú referencia modell, amely a következő egy órára mindig az elmúlt két óra átlagát rendeli hozzá. A hagyományos értelmezéstől ebben az esetben azért tértem el, hogy az órán belüli intenzív ingadozások hatását csökkentsem a referencia modell esetében, ezzel a középértékét célzó megoldással. A MAE alapvetően azt mutatja meg, hogy átlagosan mekkora abszolútértékű hibával dolgoznak a becslések. Az RMSE szintén irányfüggetlen, ugyanakkor a hatványozás miatt erősebben bünteti a nagyobb kilengéseket, ezáltal gazdasági szempontból reprezentatívabb mutató, a kiegyenlítőenergiát szolgáltató tartalék erőművek merit order alapú igénybevitelét figyelembe véve. A Skill score a perzisztenciához képesti relatív RMSE javulást mutatja meg, a dir% pedig azt, hogy az esetek hány százalékában sikerült a valóssal megegyező irányt becsülnie az algoritmusoknak.

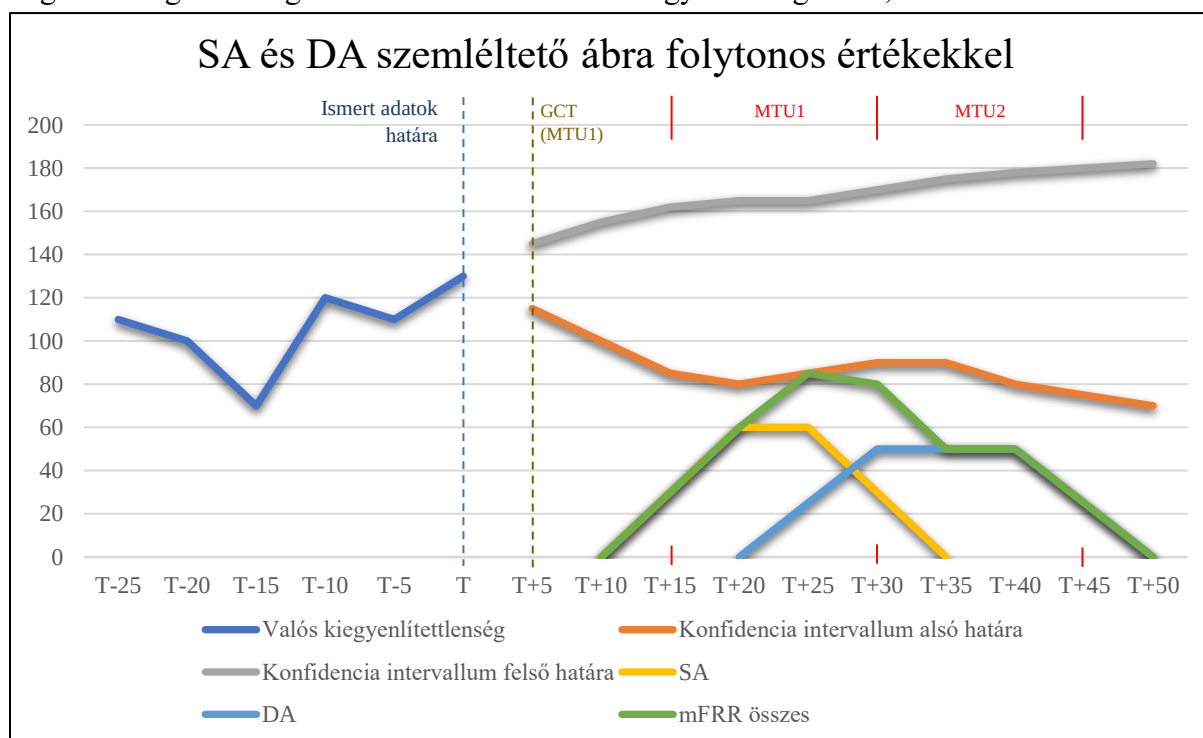
2.5 Platformos rendszer tesztkörnyezetének alapjai

A prediktorokat bemutató táblázatban látható, hogy az előrejelzések minden időpontban a következő egy óra ötperces intervallumaira kerültek elvégzésre. Ennek az időtávnak a megválasztása a diszpécser munkájának megkönnyítése érdekében történt, mivel nekik érdemes látniuk, hogy hosszabb távon mire számítsanak. A bevezetésben említettem, hogy a proaktív mFRR eszköz felépítése alapvetően a platformos rendszer igényein alapul. Ezekről a specifikációk és a célzott termékek a következőkben kerülnek ismertetésre.

A MARI [40] és PICASSO [2] platformok lényege az, hogy összekapcsolja az együttműködő országok mFRR és aFRR piacait. Jelen elemzés tárgyát előbbi képzi, amely úgy fog megvalósulni, hogy egész órától kezdve negyedórás sávokra, ún. MTU -kra (Market Time Unit) lesz lehetősége a TSO-knak szabályozási igényt és maximális árat, illetve a BSP-knek (Balancing Services Provider) ajánlatokat és hozzá tartozó minimális árat ajánlaniuk. A rendszer az MTU-kat 10 perccel megelőző GCT-ig (Gate Closure Time) beadott ajánlatokat egy erre kifejlesztett algoritmussal feldolgozza, és kirendezi, oly módon, hogy az aktivált energia marginális, azaz, a teljesült ajánlatok listájának legmagasabb árán kerül kifizetésre.

Alapvetően kétféle, időben trapéz alakú terméket lehet majd megkülönböztetni. Az első, az SA (Scheduled Activation) [40], amely az adott MTU előtt 5 perccel kezdődő, 10 perc hosszúságú felterheléssel indul, majd 5 percig az igényelt maximális teljesítményen marad, ezt követően pedig szintén 10 perc leterhelési idővel száll ki a rendszerből. Ezek alapján látható, hogy a trapéz területe, azaz, az aktivált energia megegyezik azzal az energiával, amely egy negyedóra hosszú, maximális magasságú téglalap területe lenne. Természetesen, amennyiben lefelé irányuló szabályozásról van szó, ennek ellentétes előjelű változata történne.

A másik lehetséges termék a DA (Direct Activation), amely lényege, hogy a volumen igény nincs korlátozva egy adott MTU-ra, ezért indítható valamely MTU közben is, ugyanúgy 10 perces fel- és leterhelési idővel, azzal a kikötéssel, hogy az aktiválás utáni MTU-ban érjen véget. A megértést segítő alábbi ábrán látható egy lehetséges SA, DA kombináció.



3. Ábra: A MARI platformban várható SA és DA termékek bemutatása.

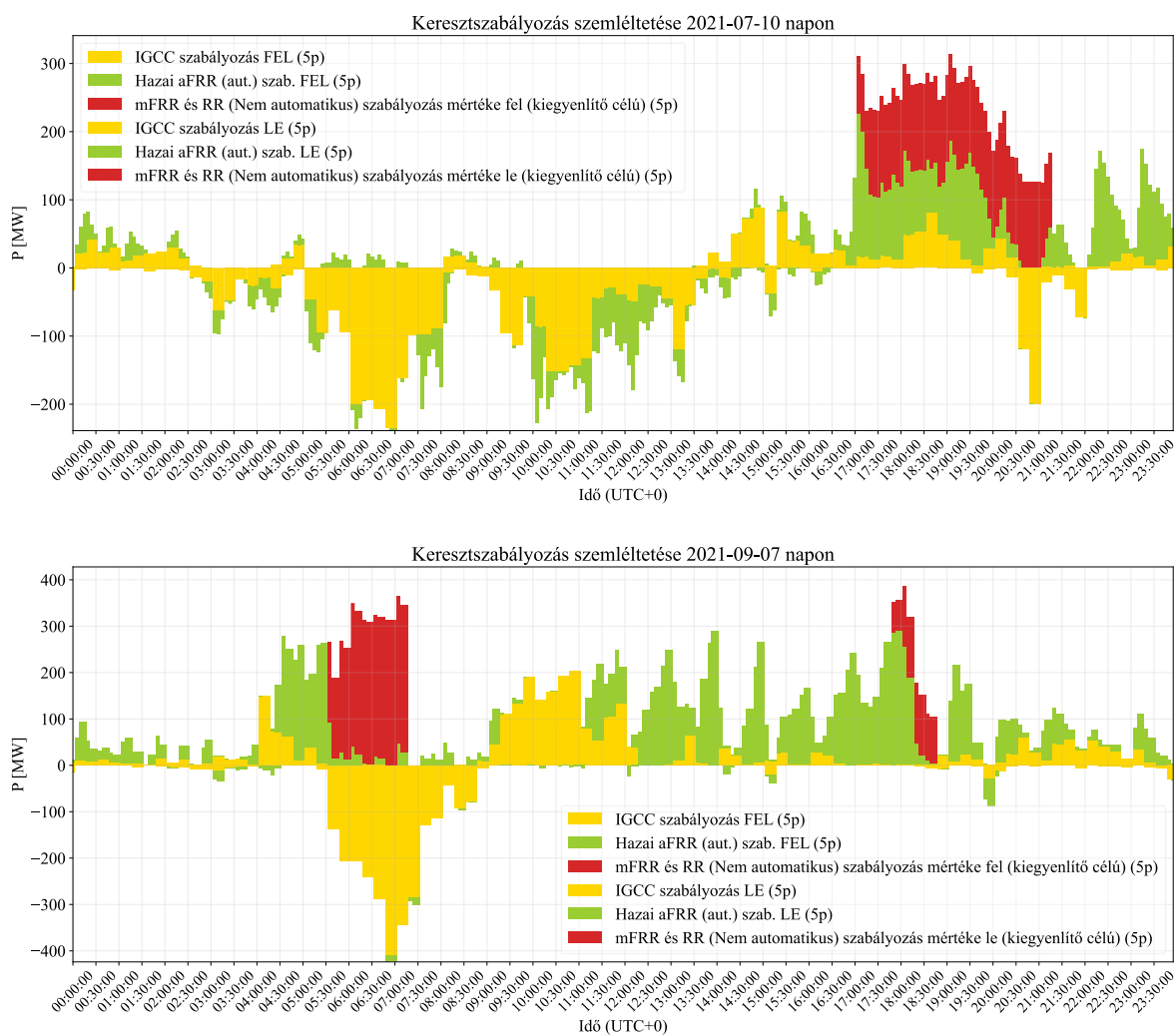
A DA tehát egy bizonyos szintű szabadságot nyújt a TSO-knak, illetve segít abban, hogy pontosabban ki lehessen rajzolni a szabályozási igényeket. Ugyanakkor, vélhetően drágább szolgáltatás lesz, elsőként mindig az SA-t kell tervezni, és amennyiben szükséges, hozzá igazítani a DA-t. Ezt azonban nehezíti az, hogy a nem-legközelebbi MTU-ra (az ábrán MTU2) igényelt SA termék T+20 időpillanatban kerül meghatározásra, az addig eltelt három intervallum adatainak ismeretében. Így tehát a DA aktiválást vélhetően olyan helyzetek fogják indukálni, amely jelen kutatás határain túl mutat, mint például a tartós teljesítmény kiesés, vagy egyéb diszpécseri döntés. Ebből a megfontolásból az elemzés fókusza a legközelebbi MTU-ra történő SA aktiválás szimulációja lett. A szimulációhoz Gradient Boosting regresszió alapú kvantilis regressziót használtam, amelynek előrejelzéseit a konfidenciaszint és az utolsó ismert adatpontra előírt minimális abszolút kiegyenlítetlenség függvényében, a következő aktiválási logikával:

- ha az adott konfidencia szintű intervallum nem metszi a 0 tengelyt a [T+15, T+35] intervallumon,
- és a T=0-ban mért kiegyenlítetlenség abszolútértéke nagyobb, az előírt minimumnál,
- akkor a konfidenciasáv a [T+15, T+35] intervallumon felvett legkisebb abszolútértékű pontja lesz a javasolt maximálisan aktiválható mFRR mennyiség.

2.6 Keresztszabályozás

Az előzőekben ismertettem az adatokat, illetve a pont-, és a konfidenciaintervallum becslés alapú proaktív mFRR aktiválás felépítését. A rendszernek az a célja, hogy a szabályozásba a tervezhetőség növelésével további rugalmasságot vigyen, azaz új erőművi beruházások nélkül,

pusztán jobb üzemeltetéssel tudjon akár magasabb szabályozási igényeket kiszolgálni, alacsonyabb költségszinten. A költségek tekintetében az egyik jelentős szempont a keresztszabályozás, amely lényegében azt jelenti, hogy egyszerre mindkét irányban történik tartalék aktiválás. Ez a jelenség az mFRR esetén viszonylag gyakori, ami annak köszönhető, hogy egy adott MTU vége és az utasítás kiadása közötti fél órában a domináns irány megváltozhat, amelyet a már üzemben lévő tartalékok teljesítmény változtatásával már nem lehet lereagálni. Szemléltetésképp a 4. ábrán láthatóak az IGCC, aFRR, valamint mFRR és RR adatok 2021 július 10-én és szeptember 7-én. Előbbinél az esti órákban aktivált mFRR egy tipikusnak mondható eset, az igények csillapodása gyorsabb volt, mint az erőművek reakcióideje, amely keresztszabályozáshoz vezetett. A szeptemberi viszont egy extrém példa, ilyen mértékű és hosszúságú hiba ritkán fordul elő, szemléltetésnek viszont megfelelő.



4. Ábra: Keresztszabályozás bemutatása valós példákon.

A keresztszabályozás tehát fontos szempont a tesztkörnyezetben kapott eredmények kiértékeléséhez, a valóságban alkalmazott módszertannal való összevetéshez. Két típusát lehet megkülönböztetni:

1. a rossz irányban aktivált,
2. vagy helyes irányban, de túl nagy volumenben aktivált szabályozás.

Ezeknek a típusoknak az összegzésével számoltam a tesztidőszakra vonatkozó keresztszabályozási hibát az (5) képlet szerint, előbbiből n , utóbbiból m db esetet feltételezve.

Fontos, hogy ebben az összevetésben az mFRR-t nem folytonos értékekkel kell figyelembe venni, hanem oly módon átalakítva, hogy az adott i, j időpillanathoz az azt megelőző öt perc átlagértéke legyen hozzárendelve.

$$Error = \left(\sum_{i=1}^n \frac{|mFRR_i|}{12}, \quad ahol \ mFRR_i \cdot IB_i < 0 \right) + \left(\sum_{j=1}^m \frac{|IB_j - mFRR_j|}{12}, \quad ahol \ (IB_j - mFRR_j) \cdot IB_j < 0 \right) [MWh] \quad (5)$$

Az adatok kapcsán korábban említettem, hogy a kiegyenlítetlenség értékeit úgy képeztem, hogy a fel- és leirányú komponensekből irányonként összegeztem, majd kiválasztottam a dominánsat és ahhoz hozzáadtam még az ACE-t. Ez a módszertan a legtöbb esetben jó, mivel a tartalékméretezéshez a szélsőségeket kell tudni kezelni, ugyanakkor, ha komolyabb keresztszabályozás következett be a valóságban a 4. ábrához hasonlóan, akkor már nem tükrözi a valóságos igényeket, hiszen egy ± 50 MW-os igény helyett 300 MW-os nagyságrendű kiegyenlítetlenségeket rögzít, ráadásul akár irányt is váltva két intervallum között. Ez egy igen ritka eset, az általam készített előrejelzések értékelésében nem okoz jelentős változást. Ugyanakkor a valóságban okozott keresztszabályozás számításában már jelentős befolyással bír, ezért ennek megállapításához a fel- és leirányú szabályozásokat intervallumomként összegeztem, majd hozzáadtam az ACE-t, és az így képzett kiegyenlítetlenséggel vettem össze a manuális aktiválásokat, ugyanúgy az (5) képlet szerint.

2.7 Újratanítás

A tanuló algoritmusok újratanítása fontos téma, melynél praktikum szempontjából alapvetően a *hogyan* és *milyen sűrűn* kérdésekre kell választ keresni. Magát az igényt az indokolja, hogy a rendszer változik. Egyrészt természetes módon lassan, a technológiák fejlődésével, fogyasztási igények folyamatos átalakulásával. Másfelől tud ugrásszerűen átalakulni egy komolyabb új erőmű belépésével, új határkeresztező vezetékek üzembe helyezésével, vagy a már meglévők kapacitásainak növelésével. Ugyanígy gyors változást tud eredményezni a kereskedési szabályok átírása, legyen szó a termékek időtartamának, vagy azok időhorizontjának csökkentéséről. Ezen gondolatmenet azt sugallja, hogy az újratanítás rendszerességét nem feltétlenül lehet általánosítani, azonban lehet kategorizálni cél szerint. Egyfelől a folyamatos változások követésére érdemes rendszeres időközönként frissíteni a betanított algoritmusokat túlillesztés elleni védelemből, másrészt amennyiben valamilyen komolyabb rendszerváltozás történik, úgy annak az előrejelzésre gyakorolt hatását érdemes vizsgálni.

Módszertani szempontból kétféle újratanítási lehetőség merült fel az elemzés során. Vannak algoritmusok, amelyek képesek az ún. inkrementális (másként: online) tanulásra [41], amely lényegében a már illesztett algoritmus belső paramétereinek az új adatok felhasználásával történő frissítésként értelmezhető. A másik alapvető megközelítés, hogy egyszerűen az új adatokkal kiegészítjük a korábban már felhasználtakat, és újra elvégezzük a tanítást.

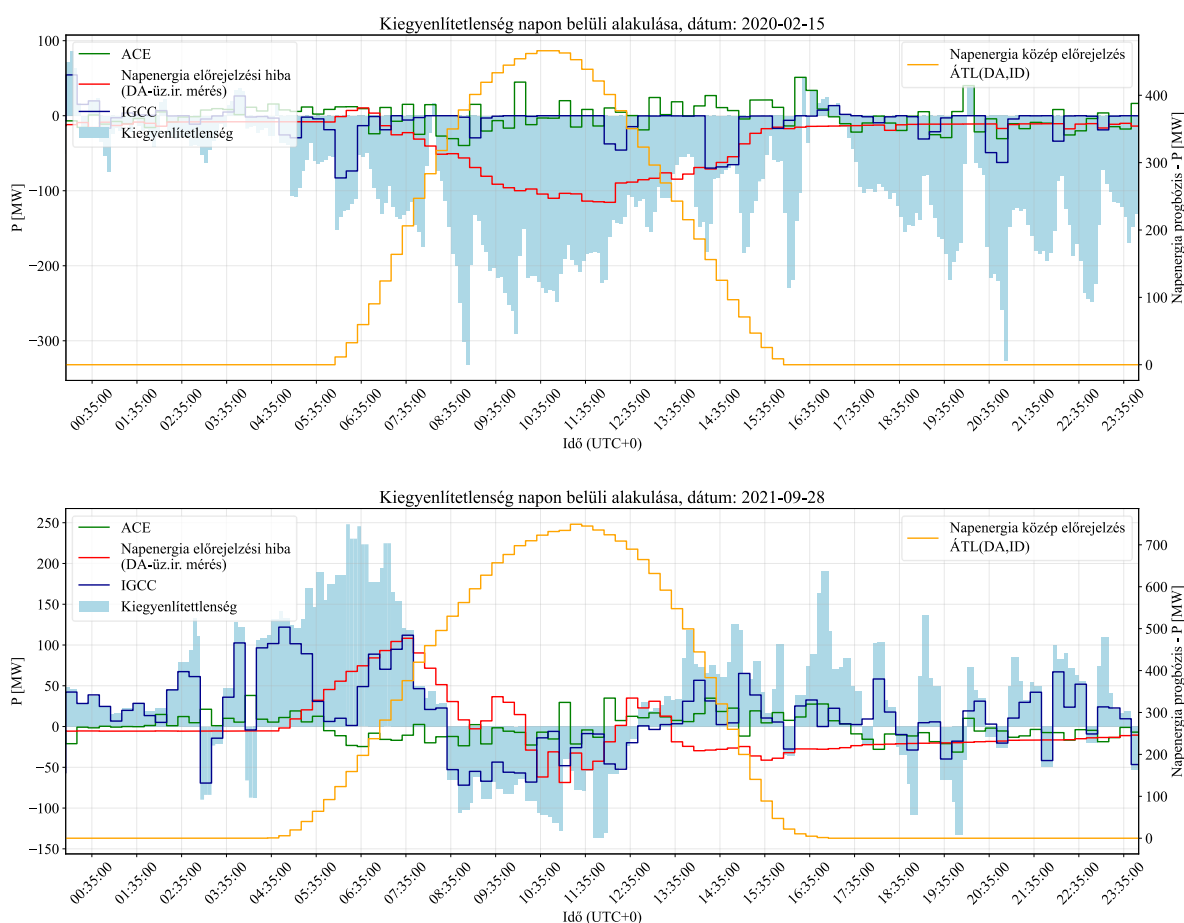
Jelen esetben az adatsor ötszöri kibővítésével történt az újratanítás az utóbbi (magasabb erőforrás igényű) módszertan mentén, mivel az algoritmusok egy része nem támogatja az inkrementális tanulást beépített függvénnyel. Az így előállított eredmények a módszerek megnevezésénél az „_exh” kiegészítéssel lesznek jelezve a későbbiekben.

3 Adatelemzés

Elsőként az ismertetett adatok szemrevételezése, illetve korreláció és más statisztikai alapú vizsgálatát végeztem el. A „nyers” adatok elemzése a szakirodalmi javaslatok mellett egy kiváló módszer a terület ok-okozati viszonyainak felderítésére, az egyidejűségek, potenciális magyarázó változók megértésére. Ennek a vizsgálatnak a tapasztalatai a következőkben olvashatók.

3.1 Korreláció vizsgálat

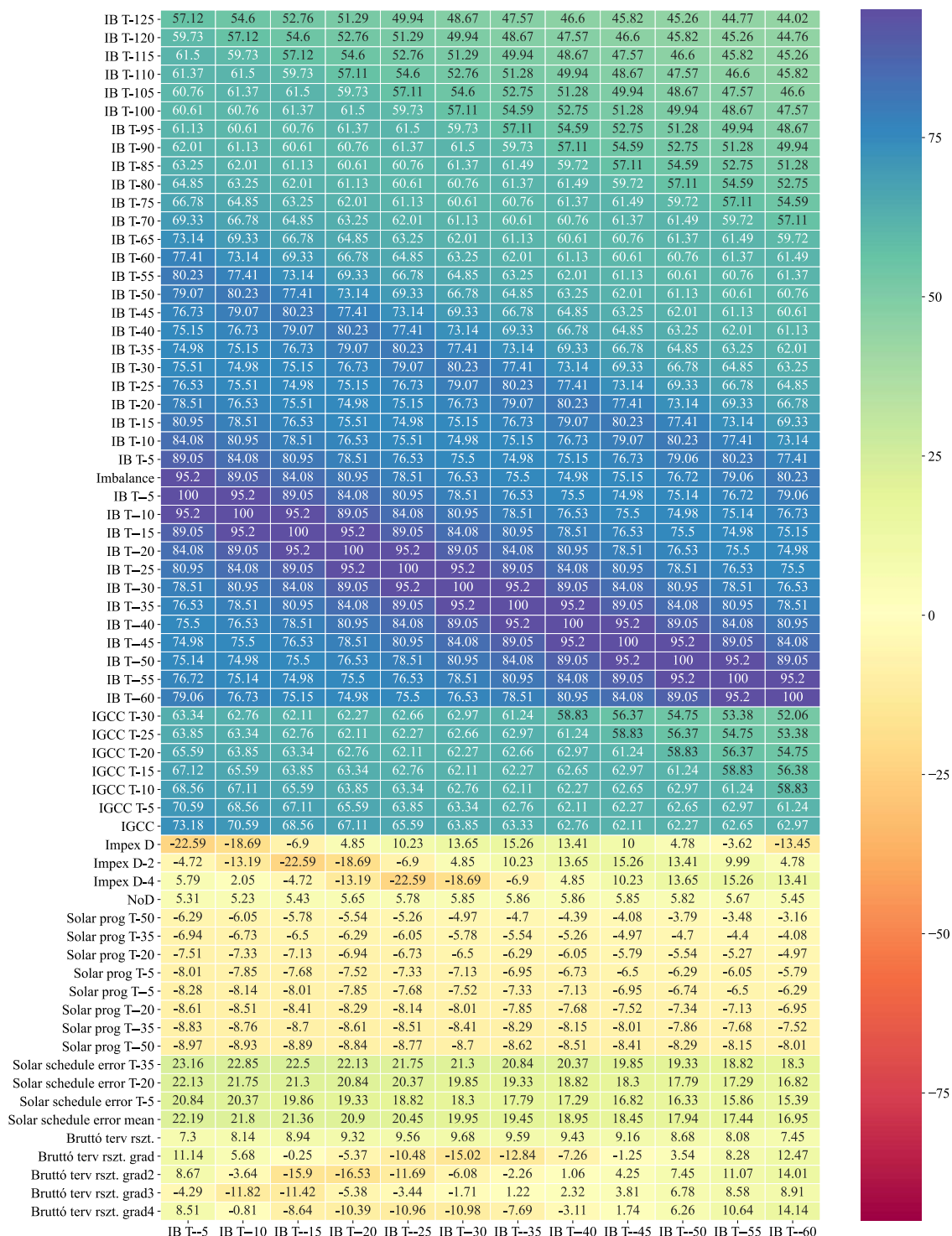
Az alábbi ábrán két, a vizsgált tartomány elején, illetve végén található nap folyamán rögzített adatok láthatóak. Elsőre szembetűnő, a kiegyenlítetlenségben van egy órás hullámszerűség, illetve az, hogy a kis méretű eltérések esetén az IGCC és az ACE rendszerint magasabb arányt képvisel. Ez a megfigyelés összhangban van a Gridradar nevű elemző oldal egyik nemrégiben közzétett anyagával [42], amelyben frekvencia adatsorokat vizsgáltak Fourier-transzformációjával és megfigyelték, hogy az órás frekvenciájú hullámszerűség az egyik legjelentősebb, hasonlóan az általam említett tapasztalatokhoz.



5. Ábra: Kiegyenlítetlenség összetevők és a napenergia termelés alakulása két vizsgált napon.

Ezen a vonalon elindulva érdemes megvizsgálni az ismertetett adatok és a valós kiegyenlítettség közti lineáris kapcsolatot, amelyhez a Pearson-féle korrelációs együtthatót használtam százalékos formában. Ennek eredményei láthatóak az alábbi hő térképen. Az iménti megállapítások közül az órás hullámszerűség egyértelműen megjelenik a korreláció esetén is. Ez volt az a tapasztalat, amely az import-export adatok felé terelte a vizsgálatot, hiszen jelenleg a

rendszerben leginkább a nemzetközi menetredek azok, amelyek órás bontásúak, két szomszéd (Ausztria, Románia) kivételével. Az Impex-D adatok generálása mögötti megfontolások innen eredeztethetők tehát.



6. Ábra: A vizsgált adatsorok közti Pearson-féle korrelációs tényezőket százalékos formában ábrázoló hő térkép.

A hő térkép alapján ennél sokkal több következtetést nem feltétlenül érdemes levonni, hiszen az algoritmusok elágazó pontjaiban, kategorizálásában nem a lineáris kapcsolat a leginkább reprezentatív jellemző, ez inkább csak előzetes képet ad arról, hogy mivel érdemes foglalkozni,

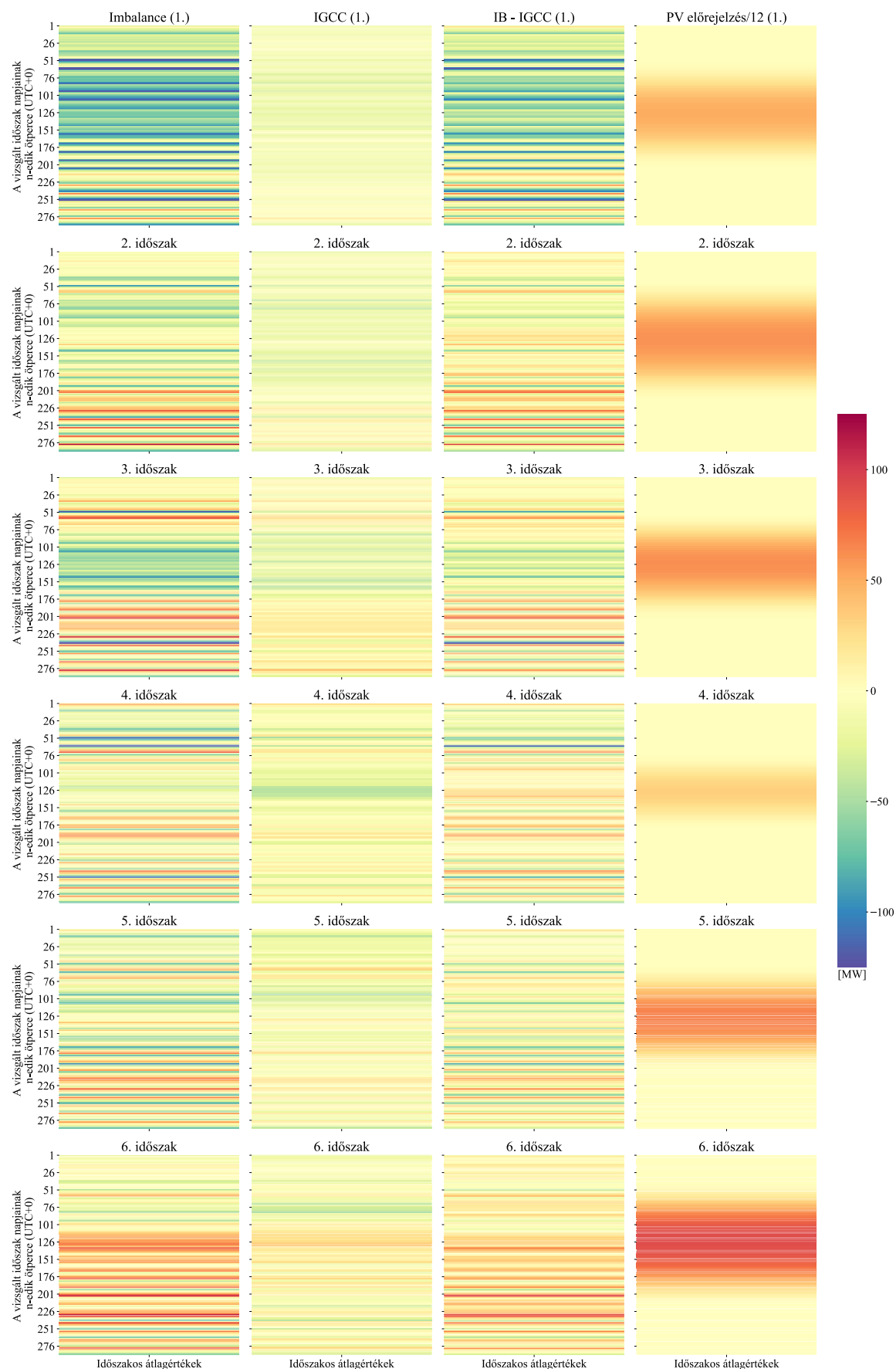
illetve melyek azok a változók, amelyek egymással is erős kapcsolatban vannak, így együttes használatuk várhatóan nem jár komoly hozzáadott értékkel. Ezen tapasztalatok alapján az adatokat további vizualizációs technikákkal ábrázoltam, amelyek célja, hogy az ok-okozati viszonyok, illetve az egyidejűséget vizsgálva az elemzéshez használt változók lehető legjobb megválasztása, valamint a kapott eredmények kiértékelésének megalapozása.

3.2 IGCC rendelkezésre állás vizsgálata

Ahogy a bevezetésben is említésre került, az IGCC egy jó alternatívája a hazai FRR aktiválásoknak, mind műszaki, mind gazdasági szempontból, hiszen jellemzően kedvezőbb árazású, illetve nem jár terhelésváltoztatással a hazai erőművek esetén. Az import-export menetrendek változása esetén lényegében azért keletkezik kiegyenlítetlenség, mivel az erőművi egységek nem tudnak ugrásszerűen terhelést változtatni, amely ugyanúgy igaz a külföldi egységekre. Ebből tehát az következne, hogy egy ilyen menetrendváltozás esetén az együttműködő országok ellentétes irányú kiegyenlítetlensége jelentős mértékben kompenzálható lenne IGCC-vel. Ugyanakkor ezt bonyolítja, hogy egy-egy szomszédos országnak a mi rendszerünkkel vett szaldója adott esetben csökkenhet úgy is, hogy egyébként a többi határon kereskedett volumenük nő, tehát akár lehet mindkét országnak azonos irányú a kiegyenlítetlensége. Az IGCC lehetőségét bizonytalanná tesz továbbá a határkeresztező kapacitások rendelkezésre állása, illetve az is, hogy nem minden környező ország vesz részt jelenleg ebben az együttműködésben [43]. Mindezek alapján érdemesnek tartottam az IGCC, a kiegyenlítetlenség, illetve a tervezett menetrendváltozás kapcsolatait a következőkben bemutatott ábrákon keresztül vizsgálni.

Az alkalmazott módszertan az, hogy a rendelkezésre álló adatokat egymást követő időszakokra osztva, az egyes ötpercekre jellemző értékek kerülnek bemutatásra. Elsőként a 7. ábrán hat időszak átlagos kiegyenlítetlenségének (Imbalance), IGCC adatainak, a kettő különbségének, illetve a napenergia termelés előrejelzésének hőtésképei láthatók. Ebből alapvetően az egyes időszakokra, illetve azon belüli ötperces intervallumokra jellemző irányokra és azok változásának trendjére lehet következtetni. Első ránézésre jól látszik a korábban említett órás változás, különösen a reggeli és esti csúcsok közelében. A tendenciák kapcsán elmondható, hogy korábban jellemzően a leszálló, míg az utóbbi időben inkább a felszálló igények dominálnak. Az jobb szélső oszlopban ábrázolt, érzékelhetően növekvő napenergia termelés időszakos alakulása, illetve a szabályozási igények iránya között egyelőre nem tapasztalható egyértelmű összefüggés, ez vélhetően az elszámolási rendszer változásának köszönhető. 2020 április 1. óta kell a KÁT, illetve METÁR rendszerben üzemelő napelemparkok termelésére menetrendet adni, illetve annak hibája után pótdíjat fizetni. Ennek a költségnek alapjául az aktuális kiegyenlítőenergia elszámolási rendszer szolgál, amelyet azóta többször módosítottak [44], ezáltal a kezdetben inkább alulmenetrendezést támogató környezet [45] mára inkább már szimmetrikus, ezáltal kevésbé alakulnak ki olyan különbségek az irányok között, mint 2020 nyarán, illetve ár extrémumok, mint 2021 elején [46]. Ez a változás lehet egy oka a reggeli és esti csúcs közötti jellemző trendek átalakulásának, amelyen még erősíthet az egyre növekvő HMKE szintű PV termelés, amely a nettó fogyasztást befolyásolja.

Lényeges megállapítás erről az ábrázolásról továbbá, hogy vannak olyan idő intervallumok, amelyekben az IGCC-vel csökkentett kiegyenlítetlenség stabil iránnyal, átlagosan akár 100 MW-os értékkel bír. Ezekben az időszakokban akár előrejelzés nélkül is lehetne mérlegelni az mFRR aktiválást, illetve másfelől – az esetlegesen ezekben az időszakokban tapasztalt alacsony aFRR rendelkezésre állás várhatóan az óraváltás után meg fog szűnni beavatkozás nélkül is.



7. Ábra: Időszakos kiegyenlítetlenség, IGCC, illetve napenergia termelés előrejelzés ötperces átlagainak hő térképe 2020.02.01-2021.10.06 között.

Míg a 7. ábrán csak átlagértékek szerepeltek, a következőkben tárgyalt 8.-on alapként az IGCC adatok boxplotja látható, amely lényege, hogy a narancs színű medián mellett megmutatja az adott intervallumra vonatkozó eloszlásának két lépcsőjét:

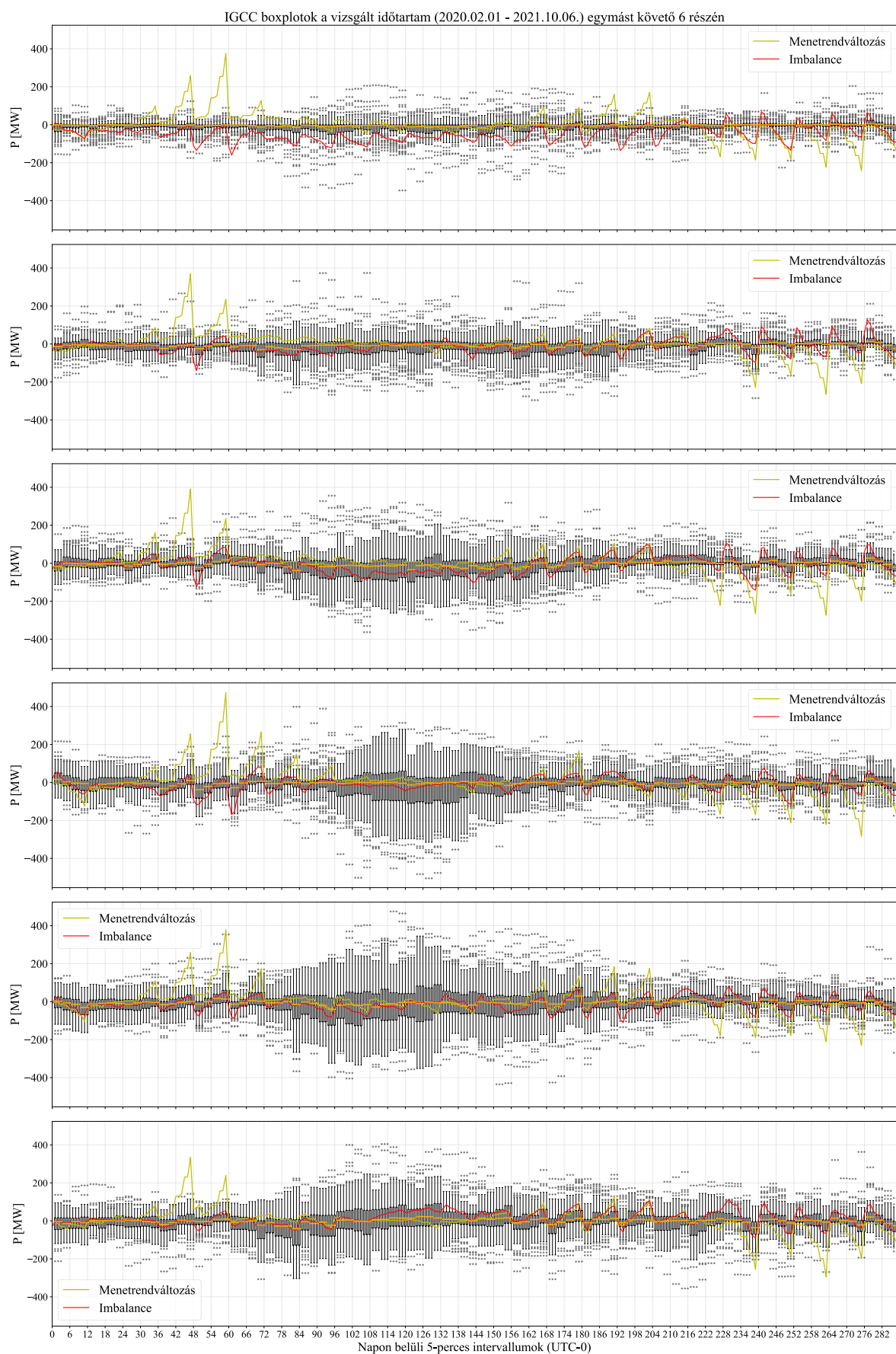
- egyrészt az interkvartilis terjedelmet, amely azt a tartományt jelöli, ahol az összes érték középső 50%-a helyezkedik el,
- másrészt az adatok 100%-os konfidenciaszint melletti várható tartományát.

Azok az adatpontok, amelyek ezeken a sávokon kívülre esnek az *outlier*-ek, azaz kiugró, vagy potenciálisan hibásnak minősíthető esetek. A boxplotot kiegészítendő, ábrázolásra került a kiegyenlítetlenség, illetve a *Bruttó terv rszt. grad4* nevű prediktor (mint képzett menetrendváltozási gradiens) átlaga.

Magyarország 2020. márciusában csatlakozott az IGCC rendszerbe, a tartományok fokozatos szélesedésével jól érzékelhető, hogy az együttműködés keretében egyre nagyobb volumenben sikerült a kiegyenlítetlenséget elnettósítani az országok között. Az is látható, hogy míg a kezdeti időszakban az IGCC jellemzően nulla körüli mediánnal bírt napszaktól függetlenül, az idő előrehaladtával reggeli és az esti csúcsok közelében egyre inkább átvette a teljes kiegyenlítetlenség trendjét, csillapított volumenekkel. Amellett tehát, hogy az rendelkezésre állás, egyre inkább biztosított ezekben az időszakokban, a jóval kisebb átlagokból látható, hogy még ilyenkor is jelentős tere van az FRR szabályozóknak.

Másfelől, a nappali intervallumokra tekintve látható, hogy a volumenek széles határok között mozognak, az esetek felében tapasztalható 60-70 MW-nál nagyobb abszolútértékű IGCC, amely jellemző irány hiányában megnehezíti a proaktív mFRR szabályozás dolgát, hiszen a cél az lenne, hogy IGCC-t lehetőségét lehetőleg nem veszítve csökkentsünk aFRR igényeket. Ez azonban jelenleg nem biztosítható, gyakran az egyik pillanatban 2-300 MW-ot is el lehet nettósítani, amely ugyanakkora igény mellett a következő ötpercben pedig akár nullára is csökkenhet.

A menetrendváltozás kapcsán a nappali órákban, a PV növekedésével egyre kevésbé ismerhetők fel a kiegyenlítetlenséggel azonos trendek, ugyanakkor a reggeli, illetve esti időszakokban jól látszik, hogy az átlagokban ugyanaz a hullámváz tapasztalható, csak előbbi esetén nagyobb amplitúdóval. Ez a megfigyelés arra következtet, hogy lényegében a kiegyenlítetlenség fő determinisztikus oka a menetrendváltozás, amely magában foglalja az import-exportot is, és ehhez adódik hozzá az időjárásfüggő energiatermelés, illetve a fogyasztás bizonytalansága, amelyek PV dominancia miatt leginkább a nappali órákban rendelkeznek erős befolyással.



8. Ábra: Időszakos IGCC boxplotok kiegészítve a kiegyenlítetttség, illetve a menetrendváltozás (*Bruttó terv rszt. grad4*) átlagos alakulásával.

3.3 Változók relevancia vizsgálata

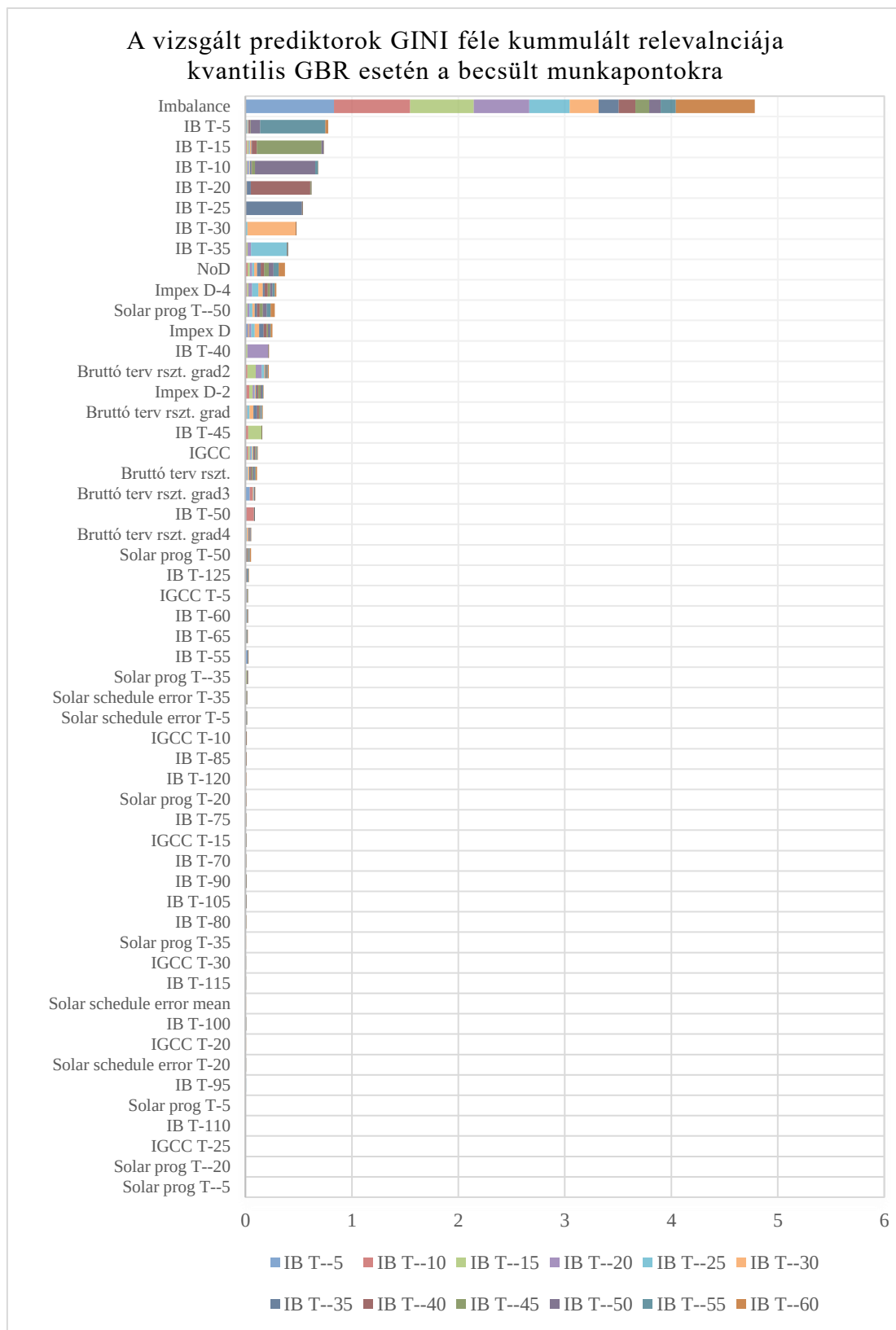
A 2. fejezetben ismertetésre kerültek azok az adatok, amelyek előzetes megfontolás alapján az elemzésben potenciálisan prediktorként alkalmazhatóak. A korreláció, illetve IGCC, menetrendváltás vizsgálata alapján elkezdett körvonalazódni, hogy melyek rendelkeznek erősebb kapcsolattal a kiegyenlítetlenséggel. A legtöbb magasabb szintű algoritmus képes arra, hogy a bemenő adatokat súlyozza relevancia szerint, ezáltal jellemzően fölösleges változók hozzáadásával nem fog romlani az előrejelzés. Ugyanakkor a futási időt jelentősen meg lehet így növelni, illetve valós alkalmazás esetén is a több bemenő adat nagyobb adatforgalmat és bonyolultabb rendszert eredményez. A következő lépés tehát az, hogy valamilyen módon számszerűsítsük a változók fontosságát, sorrendet állítsunk, és kiválasszuk azokat a változókat, amelyek használatával a pontosság elfogadható szintű csökkenése mellett lehet egyszerűsíteni a rendszert. Mivel minden egyes kombináció tesztelésére nincs lehetőség erőforrás oldalon, az értékelésre a GINI relevancia, vagy más néven átlagos szennyeződés-csökkenés (MDI) került alkalmazásra, amely más területen is megfelelőnek bizonyult erre a célra [47]. Ezt a mutatót összetett döntéshozó fa alapú rendszerekre találták ki, és úgy lehet meghatározni, hogy a csomóponti szennyeződések elérési valószínűséggel (amelyet a csomópontot elérő minták arányával lehet becsülni) súlyozott teljes csökkenését átlagoljuk az erdő fáira. A különböző algoritmusok, könyvtárak nem feltétlenül ugyanúgy számítják [48], ebben az esetben a scikit-learn implementáció került alkalmazásra, amely Breiman [49] értelmezésén alapul.

Az értékelést a Gradient Boosting Machines kvantilis regressziójával végeztem el, melynek eredményei a 9. ábrán láthatóak 12 időszakra egy összevont halmazozott oszlopdiagramon. A kapott értékek összeg szerinti csökkenő sorrendben kerültek ábrázolásra, így fentről lefelé haladva az egyes adatsorok hozzáadott értéke egyre kevesebb. Ugyanakkor fontos kiemelni, ahogyan a színekből látszik, hogy nem minden időtávra ugyanazok a prediktorok a fontosak, az IB adatok esetén jellemzően az egy órával korábbi értékek emelkednek ki, amely megfigyelés összhangban van a korreláció vizsgálatnál tapasztalt órák hullámmozásával.

Praktikusan három prediktor kombináció került tesztelésre, amelyek elnevezései és specifikációi a következők:

- **Komplex** – a korábban ismertetett, illetve a 9. ábrán látható 54 bemenet, elvi szinten a vizsgált adatokból képezhető legjobb előrejelzés céljával;
- **GINI-based** – azok a változók, amelyeknek a kumulált relevanciája elérte a 0,05-öt (23 db, a *Solar prog T-50*-ig bezárólag), azzal a céllal, hogy egyszerűbb, de lehetőleg közel azonos pontosságú előrejelzést kapjunk, mint a komplex esetben;
- **Only_IB** – az elmúlt két óra kiegyenlítetlenség adatainak felhasználásával kapott előrejelzések.

A továbbiakban az elemzések ezzel a három kombinációval párhuzamosan lesznek elvégezve a változó szelekció fontosságának meghatározására.



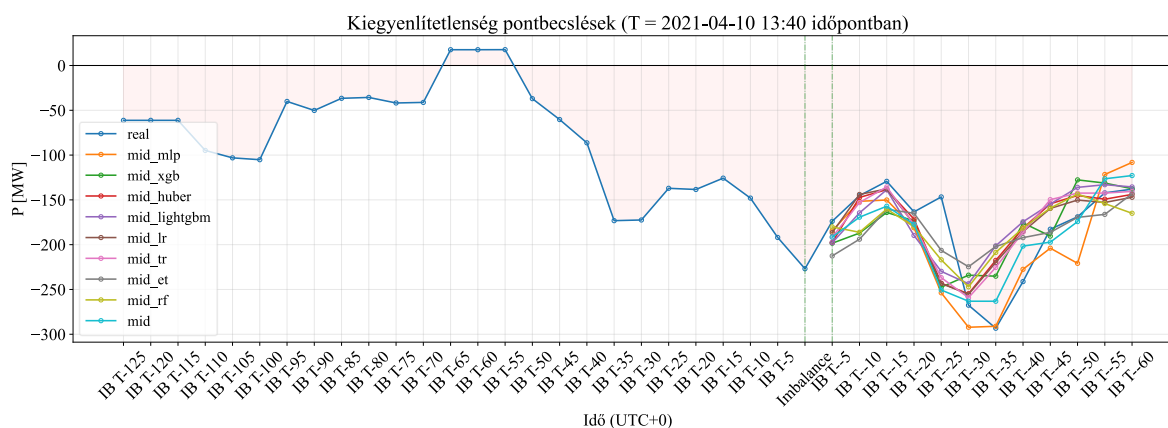
9. Ábra: GINI relevancia alapú változó ártékelés csökkenő sorrendben a 12 időszakra kumuláltan ábrázolva.

4 Eredmények

Ebben a fejezetben kerülnek bemutatásra az elemzés során kapott eredmények, elsőként a pontbecslésekre, majd a platformos tesztkörnyezetben értékelt intervallum becslésekre.

4.1 Pontbecslések eredményei

A 2. fejezetben említett módon a pontbecslések esetén a cél egyértelmű, a kiegyenlítetlenség minél pontosabb előrejelzése. A 10. ábrán látható egy munkapontra jellemző minta, lényegében ez az előrejelzés került elvégzésre a teszt időszak minden egyes ötpercében. Kemelendő, hogy ezzel az ábrázolásmóddal pillanatértékként jelennek meg a feliratokkal jelzett előrejelzések, szintúgy a valós kiegyenlítetlenség (*“real”*), amelyet az összehasonlíthatóság kedvéért előre is feltüntettem, azonban a valóságban ezek ötperes átlagértékek, az időszak végét jellemző időbéljeghez rendelve. A valósághű ábrázolás *step*, vagy *stairs* diagramokkal képezhető, amelyek a későbbiekben alkalmazásra kerültek, azonban jelen ábra úgy értelmezhetetlen lenne a túl sok görbe miatt.



10. Ábra: Pontbecslés a GINI-based esetben pillanatértékként ábrázolva.

A 3. táblázatban a pontbecslések átlagos RMSE hibái a fél éves tesztidőszakon. Az első oszlopban a *Típus* szerepelnek az alkalmazott prediktor kombinációk, míg egyvel jobbra, a *Módszer* kategóriában az algoritmusok azonosítói a 2. táblázat szerint. A módszerek nevében az „_exh” jelölés az újratánításra utal. A táblázat értékei a T+25-ös intervallum szerint lettek növekvő sorrendbe rendezve, mivel ahogyan a 3. ábrán láthattuk, ez a legfontosabb időszak a legközelebbi MTU-ra történő SA aktiválás esetén.

Az RMSE eredmények alapján az optimális előrejelzés kapcsán a vizsgált lehetőségek közötti fontossági sorrend a következő:

1. **Prediktorok megválasztása** – a csak kiegyenlítetlenség adatok figyelembevételéhez képest akár 15-20% javulás érhető el a magyarázó változók komplexebb csoportjának alkalmazásával.
2. **Modell választás** – 5-10% javulás érhető el az egyszerű lineáris modellekhez képest a komplexebb algoritmusok alkalmazásával, illetve további 3% az előrejelzések kombinálásával.
3. **Újratánítás** – jellemzően 1-2%-os javulási potenciállal bír az elemzésben alkalmazott ötszöri újratánítás.

3. Táblázat: RMSE eredmények a háromféle prediktor kombinációra, normál és újratanított esetekben.

		RMSE [MW]											
Típus	Módszer	IB T--5	IB T--10	IB T--15	IB T--20	IB T--25	IB T--30	IB T--35	IB T--40	IB T--45	IB T--50	IB T--55	IB T--60
Komplex	mid_exh	27.15	36.55	41.29	46.19	50.55	53.78	57.28	60.81	63.64	65.79	67.66	69.52
GINI-based	mid_exh	27.86	37.63	42.1	46.88	51.11	54.81	58.03	61.37	64.24	65.97	67.94	69.8
Komplex	mid	27.29	36.92	42.05	47.1	51.32	54.78	58.01	61.44	64.67	66.72	68.82	70.67
GINI-based	mid	28.28	38.03	42.79	47.41	51.94	55.55	58.6	62.5	65.07	67.02	68.95	70.91
Komplex	xgb_exh	27.3	37.77	42.37	47.41	52.01	55.61	59.01	62.3	65.14	67.25	69.55	71.51
GINI-based	xgb_exh	27.91	38.4	42.91	47.89	52.39	56.14	59.45	62.68	65.69	67.62	70.07	71.88
Komplex	mlp_exh	28.75	37.82	42.92	48.19	52.44	55.62	59.49	63.28	66.34	68.55	70.17	72.06
GINI-based	mlp_exh	29.26	39.14	43.71	48.59	52.81	56.4	59.81	63.15	66	67.86	69.54	71.38
Komplex	mlp	28.98	38.36	43.73	49.24	52.93	56.53	60.35	63.89	67.78	69.47	71.44	73.8
Komplex	lightgbm_exh	28.19	39.28	43.76	48.7	53.25	56.13	59.51	62.61	65.34	67.1	68.98	70.93
Komplex	xgb	27.56	38.17	43.17	48.6	53.27	56.65	60.06	63.37	66.4	68.91	71.13	72.48
GINI-based	xgb	28.2	38.9	43.66	48.72	53.47	56.9	60.25	63.74	67.12	69.27	71.96	73.78
GINI-based	mlp	29.85	39.63	44.46	48.93	53.59	57.35	60.52	64.46	66.93	68.74	70.18	72.35
GINI-based	lightgbm_exh	28.6	39.83	44.2	49.22	53.78	56.79	60.07	63	65.78	67.67	69.85	71.65
Komplex	lightgbm	28.46	39.65	44.19	49.3	54.02	56.94	60.45	63.56	66.42	68.81	70.49	72.36
GINI-based	lightgbm	28.92	40.11	44.65	49.59	54.62	57.42	60.58	63.84	66.82	68.66	70.57	72.63
GINI-based	et	30.02	42.34	47.25	52.84	57.33	60.84	64.23	67.32	69.33	71.18	72.85	74.45
Komplex	et	29.92	42.31	47.33	52.91	57.42	61.09	64.51	67.51	69.62	71.48	72.86	74.4
Komplex	rf	30.16	42.59	47.21	53.13	57.97	61.16	64.43	67.12	69.19	71.27	72.97	74.43
GINI-based	rf	30.16	42.74	47.4	53.16	58	61.48	64.59	67.32	69.46	71.26	72.83	74.32
Komplex	huber_exh	30.05	42.5	47.06	53.12	58.05	62.37	65.96	68.95	71.35	73.23	75.11	76.85
Komplex	huber	30.07	42.63	47.27	53.38	58.37	62.65	66.34	69.25	71.71	73.59	75.53	77.33
Komplex	tr	30.2	43.06	47.59	54.1	58.96	63.11	67.3	70.03	73.28	74.47	76.35	77.48
GINI-based	huber_exh	31.17	44.33	48.65	54.49	59.47	63.94	67.38	70.16	72.12	73.69	75.24	76.87
Komplex	lr	32.49	44.38	48.98	55.08	59.55	63.56	67.08	70.01	72.29	74.18	76	77.69
GINI-based	lr	31.21	44.47	49.01	55.05	59.74	64.04	67.47	70.26	72.23	73.79	75.3	76.84
GINI-based	huber	31.18	44.42	48.83	54.73	59.75	64.22	67.66	70.44	72.39	73.97	75.55	77.22
GINI-based	tr	31.16	44.56	49.13	55.48	59.88	65	68.01	71.16	73.03	74.17	75.61	77.05
Only_IB	mid_exh	31.94	44.95	53.05	58.22	62.29	66.05	68.47	70.77	72.66	74.08	75.45	76.97
Only_IB	mid	32.16	45.04	52.92	58.3	62.47	66.15	68.6	71.13	72.84	74.2	75.56	77.16
Only_IB	mlp_exh	32.45	45.22	53.32	58.57	62.49	66.18	68.55	70.89	72.73	74.16	75.54	77.05
Only_IB	mlp	33.12	45.45	53.11	58.39	62.63	66.21	68.55	71.3	72.96	74.19	75.7	77.33
Only_IB	xgb_exh	32.5	46.07	54.52	59.59	63.7	67.35	69.84	72.01	73.76	75.09	76.4	77.86
Only_IB	xgb	32.53	46.12	54.75	59.9	63.91	67.59	70.1	72.32	73.83	75.28	76.5	77.93
Only_IB	lightgbm_exh	32.89	46.59	55.14	60.22	64.35	67.8	70.24	72.19	73.93	75.24	76.37	77.65
Only_IB	lightgbm	32.98	46.61	55.26	60.42	64.53	68.01	70.43	72.39	74	75.4	76.53	77.72
Only_IB	lr	33.25	47.55	56.31	61.27	65.27	68.75	71.05	72.93	74.49	75.43	76.26	77.45
Only_IB	et	33.91	47.8	56.1	61.16	65.32	68.78	71.18	73.2	74.93	76.03	76.98	78.4
Only_IB	huber_exh	33.26	47.36	56.15	61.21	65.32	68.89	71.23	73.11	74.64	75.55	76.35	77.53
Only_IB	huber	33.26	47.37	56.19	61.27	65.4	68.98	71.32	73.2	74.74	75.65	76.44	77.62
Only_IB	rf	34.35	48.11	56.29	61.7	65.79	69.25	71.52	73.46	75.14	76.34	77.33	78.53
Only_IB	tr	33.42	47.91	56.47	61.75	65.99	69.32	71.38	73.59	74.97	76.43	77.33	78.1

Az algoritmusok közül az előzetes (2020-as adatsoron végzett) teszteknek megfelelően az XGBoost, MLP, illetve LightGBM bizonyult a legjobbnak, az egyébként ígértes, de a magas RAM igény miatt csak kevesebb adaton tanított módszerek, mint az ET nem tudták érdemben megközelíteni ezt a szintet. Ugyanígy a lineáris módszerek is jelentősen elmaradtak

pontosságban. Ezek a tapasztalatok egyébként ellentmondanak az Elia bevezetésben említett konzultációs anyagának [19] eredményeivel. A „mid” verzió esetén látszik, hogy az MLP, illetve XGBoost előrejelzések átlagolása effektíven csökkenti azok egyéni hibáit.

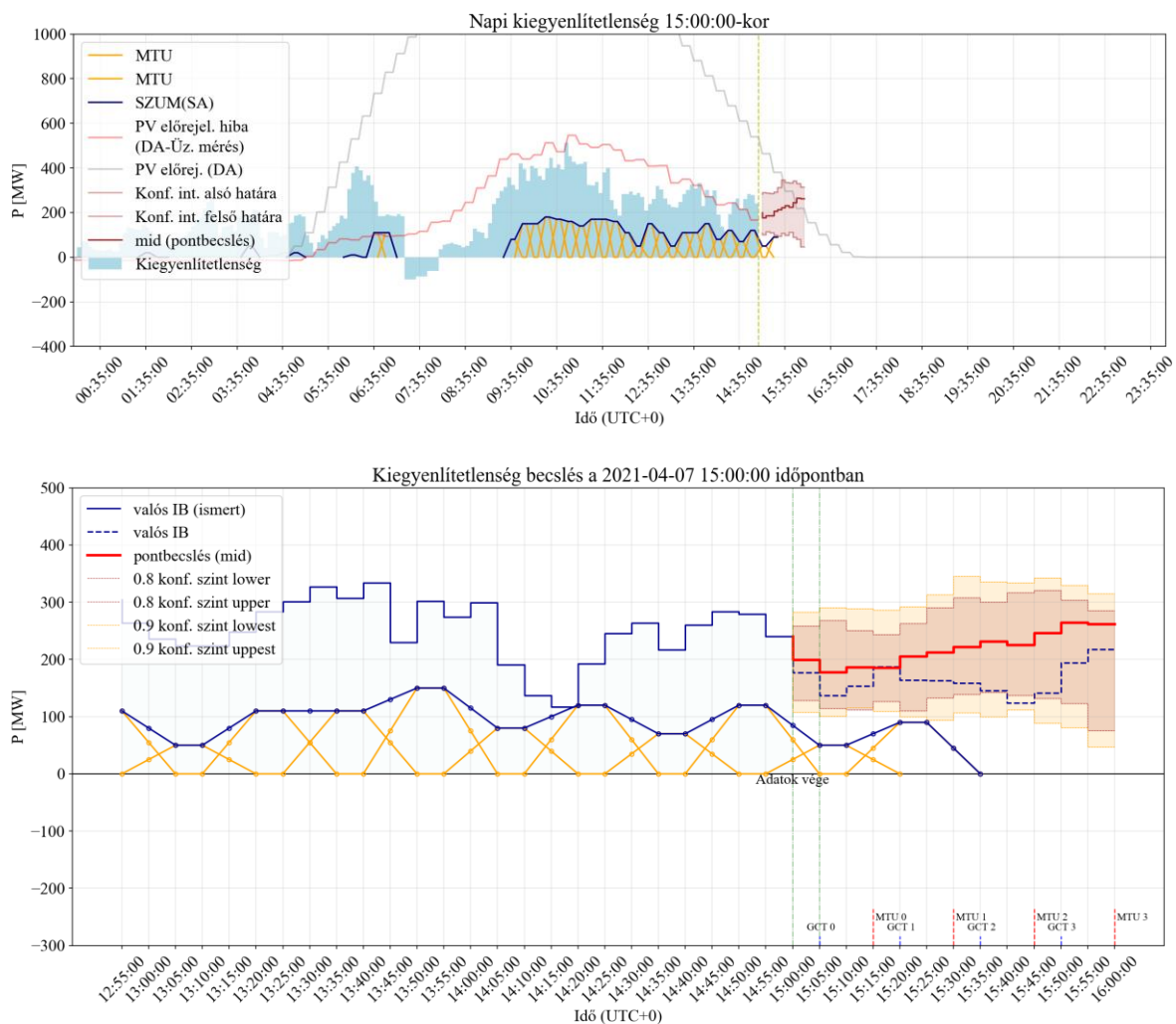
A prediktor variációk kapcsán a várt eredmények adódtak; a Komplex verzió bizonyult a legjobbnak, ugyanakkor a kiemelkedő algoritmusok esetén is csak 1%-kal teljesített jobban, mint a GINI relevancia alapú becslések, a középmezőnyben pedig ez a sorrend helyenként felcserélődött. Ez azt jelenti, hogy amennyiben erőforrás, illetve adat validációs szempontból nem kritérium a bemenő adatok kontrollálása, úgy a pontbecslésekre érdemes alkalmazni az ismertetett változók mindegyikét, amennyiben viszont igen, úgy a GINI alapú szelekció a megfelelő a komplexitás redukálására. Látható továbbá, hogy pusztán a kiegyenlítetlenség adatok alkalmazása már jelentős kompromisszumnak számít.

A további kimutatások az egyszeri tanítással kapott eredményekről a függelék 17-19. ábráin láthatók. A legjobb előrejelzések az esetek több, mint 85%-ában helyesen ítélték meg a kiegyenlítetlenség irányát a T+25-ös intervallumon. Tekintettel a gyakori 0 körüli ingadozásra, ez az érték jónak mondható. A MAE értékek ugyanezen az intervallumon 40 MW alá kerültek (38,35 MW), amely összevetve a kiegyenlítetlenség átlagos abszolút értékével (90,29 MW), 42,47%-os relatív hibának minősül.

A Skill score táblázatból kiderül, hogy a kétórás mozgóátlaghoz képest 30%-kal kisebb RMSE érhető el a vizsgált módszerekkel, amely a becslés időhorizontjának csökkentésével egyébként tovább növelhető. Az időhorizont kapcsán általánosan megfogalmazható, hogy az előrejelzések pontosságának csökkenése az első 15 percben ($RMSE_{T+20} - RMSE_{T+5}$) nagyjából megegyezik az azt követő 40 perc során tapasztaltakkal ($RMSE_{T+60} - RMSE_{T+20}$). Érdemes még megemlíteni, hogy habár az RMSE értékek a legtöbb esetben az időhorizont növelésével monoton romlanak, a Skill score esetén a gyengébb algoritmusok és prediktorok felé haladva fokozatosan tolódik át a minimum az óra végéről a T+30 felé. Ebből az látszik, hogy a mozgóátlag alapú perzisztencia becslés 30 perces időhorizonton teljesít relatíve a legjobban.

4.2 Platformos környezet eredményei és tapasztalatai

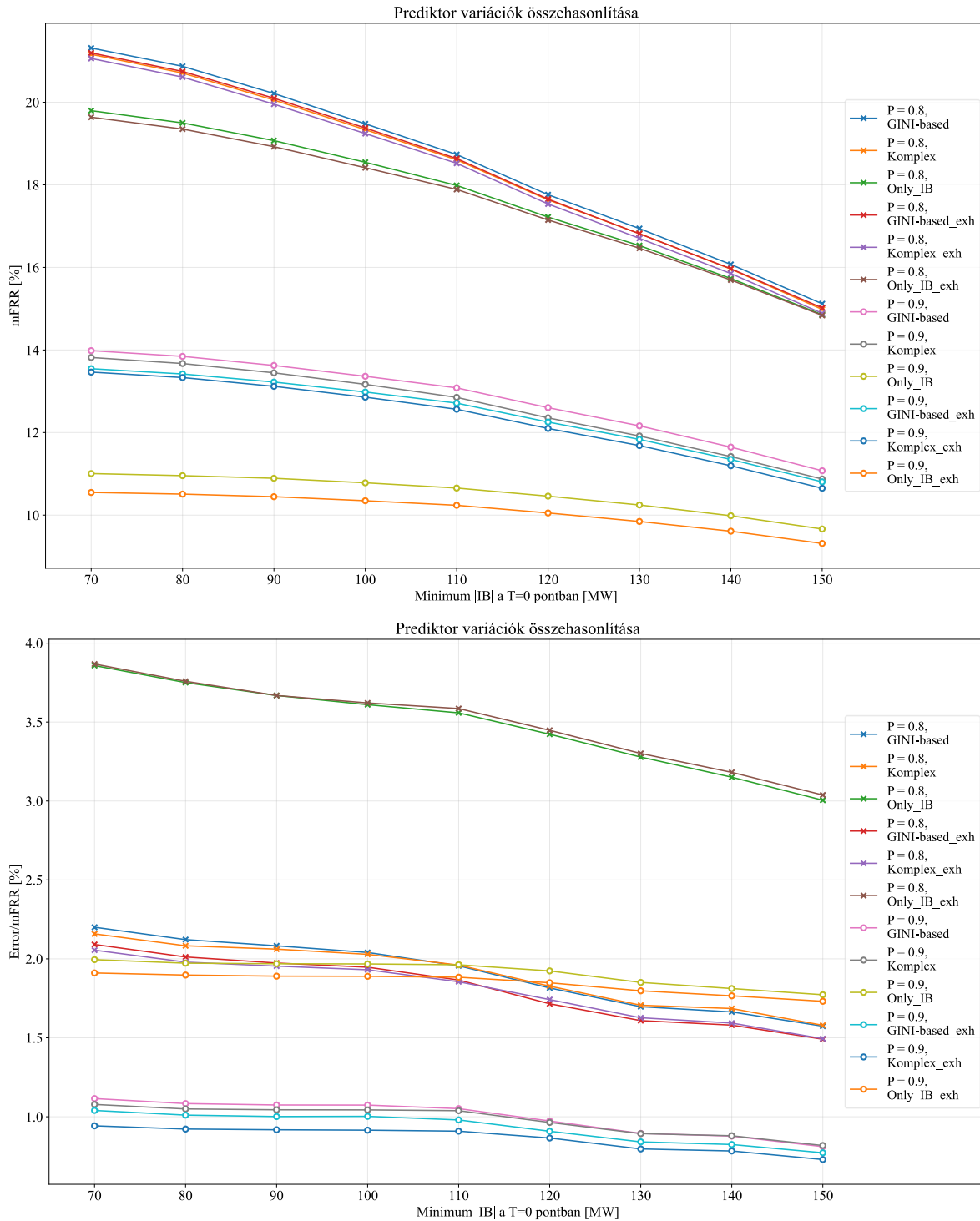
A platformos környezet tesztjei során az előbbieken tárgyalt pontbecslések közül csak a „mid” verzió került felhasználásra a különböző konfidencia intervallumok mellett. Előbbi különösebb funkciót nem kapott ebben a tesztkörnyezetben, inkább csak segítő jelleggel ábrázoltam. Az előrejelzéseket a 11. ábra segítségével mutatom be. A felső diagramon a hosszútávú trendek követése végett napi bontásban láthatók az adott munkapontig rögzített kiegyenlítetlenség adatok oszlopdigramként, illetve a napenergia előrejelzési hiba, amely a szolgáltatói becslés, és az üzemirányítói mérés különbsége, tehát nem a hitelesített menetrendezési hiba, csak egy becslés. A kék színnel ábrázolt kiegyenlítetlenség adatok folytatásaként látható a pontbecslés, és a (jellemzően, de nem kötelezően) körülötte lévő GBR alapú 90%-os konfidencia intervallum. Az alsó ábrán lényegében ugyanez az előrejelzés látható, csak szűkített vizualizációs tartományon, kiegészítve a 80%-os konfidencia intervallummal, illetve a majd később rögzítendő valós kiegyenlítetlenség értékekkel. Mindkét diagramon láthatók az adott aktiválási protokoll mellett javasolt maximális mFRR termékek valós idejű ábrázolással, sárga színnel az egyes MTU-kban elérhető SA-k, míg sötétkékek ezek összege, az átfedésben lévő időszakok megfelelő értékelésére.



11. Ábra: Platformos környezet előrejelzéseinek bemutatása egy választott munkapontban napi és rövidtávú bontásban.

A szimuláció során negyedóránkénti futtatást alkalmaztam, amely utolsó ismert adatpontja a GCT-k előtt öt perccel lévő időbélyeghez tartozik. Ezáltal csak a gyakorlatban használható legkisebb időhorizontú eseteket (17 376 db) vettem figyelembe, mivel az azokat megelőző 10 percben nincs lehetőség SA-t aktiválni. Az előrejelzések ismeretében az aktiválási protokollban szereplő minimum értéket 70 és 150 MW között tízes osztással felvéve, a különböző verziókkal, illetve konfidenciaszintekkel kapott mFRR aktiválási arányok és az egységnyi mFRR aktiválásra jutó keresztaszabályozás a 12. ábrán láthatóak. A felső diagram alapján látszik, hogy a GINI relevancia alapú esetben kapható a legnagyobb aktiválás, (a teljes kiegyenlítettséghez képest) 15-21,5% közötti értékkel a minimum függvényében. Az újratartás ebben az esetben konzervatívabb becslést eredményezett, valamelyest alacsonyabb aktiválással. Hasonlóan a pontbecslések eredményeihez, a Komplex és a GINI alapú becslések itt is közel vannak egymáshoz, az Only_IB pedig jelentősen elmarad tőlük. A minimum kritérium növelésével egyre meredekebben, de jó közelítéssel csökken az aktiválható mFRR szabályozás mennyisége és az egyes módszerek közötti különbség is egyre kevésbé releváns. A 80 és 90%-os

konfidenciaszintek közül természetesen előbbi esetén tapasztalhatóak magasabb mFRR mennyiségek, másfél-kétszeres különbséggel.



12. Ábra: Az aktiválási kritériumban szereplő minimum függvényében kapott javasolt maximális mFRR aktiválások százalékos aránya a teljes kiegyenlítetttséghez képest, illetve az egységnyi mFRR aktiválásra jutó keresztszabályozás aránya kétféle konfidenciaszinten.

Az alsó ábrára tekintve látható, hogy mennyi relatív hibával, azaz keresztszabályozással lehet számolni az egyes esetekben. Összesen 12 eset szerepel minden minimum értékhez, mindegyik különböző színnel, illetve a 80, illetve 90%-os konfidenciaszintű becslések a jelölőkben is különböznek. A várt módon az alacsonyabb konfidenciaszinthez magasabb relatív

hiba is tartozik, a tapasztalatok alapján jellemzően közel kétszeres értékkel lehet kalkulálni a kettő között. Ez alapvetően nem tűnik meglepőnek, hiszen a 80 és 90% között éppen kétszeres a várható hibák száma, ugyanakkor a keresztszabályozásra jellemző komplexebb számítási módszertan miatt az Error/mFRR értékre nem következtethető egyértelműen az azonos trend. A háromféle prediktor kombinációból az Only_IB verzió bizonyult magasan a leggyengébbnek, olyannyira, hogy a 90%-os konfidenciaszint mellett a másik két eset 80%-os eredményeit produkálta nagyságrendileg. A minimum érték növelésével a Komplex és GINI-based esetekben csak igen enyhe változás tapasztalható a 70-110 intervallumon, ami azt jelenti, hogy ezzel a szűréssel amennyi hibát megelőzünk, azzal arányos mennyiségben veszünk el mFRR aktiválási lehetőséget. Az ezeknél magasabb szűrők használatával kezd el csökkenni a hiba, ami a két konfidenciaszint esetén láthatóan nagyjából azonos trendeket mutat, csak 80% esetén meredekebben. Az újratanítás a korábban megállapított alacsonyabb mFRR aktiválás mellett érdemben, a legjobb esetekben akár több, mint (relatív) 15%-kal csökkentette az Error/mFRR arányt. Ez a tapasztalat mindenképpen azt jelzi, hogy az újratanítással érdemes foglalkozni.

A bemutatott kétféle mutató alapján komplex gazdasági értékelés szükséges a paraméterek helyes megválasztásához, ugyanis az mFRR aktiválás mennyiségének növelésével ezzel arányos költségcsökkentést lehet elérni, míg a keresztszabályozás olyan (extra) kiadás, amely okozói eltérésként nem kiszámítható a mérlegköröknek, tehát a TSO-t terheli. Ez a kockázat a jelenlegi rendszerben is fennáll, ugyanakkor a proaktív igénybevétel esetén vélhetően gyakrabban kell majd felvállalni. Ezzel kapcsolatban 4. táblázatban láthatók az aktiválási adatok 90%-os konfidenciaszint és 100 MW-os minimum előírás mellett. Ezen peremfeltételekkel a legjobb esetben 53 258 MWh energiát lehetett volna, amely 571,9 MWh keresztszabályozást vont volna maga után. A táblázatban oszloponként félkövérrel kiemelttem a legjobb eredményeket, amelyek láthatóan nem minden esetben köthetők egyféle módszerhez, tehát a preferált verzió kiválasztásához a szempontok mérlegelése szükséges.

4. Táblázat: Aktiválási statisztikák részletes bemutatása egy választott protokoll mentén, összehasonlítva a valós adatokkal

Típus	Min. [MW]	Konf. szint	mFRR [MWh]	mFRR [%]	Error [MWh]	Error [%]	Error/mFRR [%]	mFRR [MWh] +	mFRR [MWh] -	Sum IB (előjeles) [MWh]	Sum IB [MWh]
Only_IB	100	0.9	42 973	10.78	845.6	0.212	1.97	23 660	- 19 313	65 364	398 536
Only_IB_exh			41 243	10.35	779.3	0.196	1.89	22 508	- 18 735		
GINI-based			53 258	13.36	571.9	0.144	1.07	30 503	- 22 755		
GINI-based_exh			51 738	12.98	518.7	0.130	1.00	29 758	- 21 980		
Komplex			52 478	13.17	547.5	0.137	1.04	29 228	- 23 250		
Komplex_exh			51 240	12.86	468.8	0.118	0.91	28 940	- 22 300		
Valós (mFRR+RR)			17 254	4.33	1 777.2	0.446	10.30	16 270	- 983		

Az alsó sorban szerepelnek a vizsgált időszakra jellemző valós adatok, amelyekkel összehasonlítva látható, hogy közel harmadannyi mFRR aktiválás több, mint háromszoros hibát eredményezett, mint a jobb módszerek tesztjei. Itt le kell azonban szögezni, hogy egyrészt az aktiválás alakja jelenleg eltér a platformos rendszertől, másrészt az aktuális szabályok 12,5, illetve 15 perces utasítási időt írnak elő, így lényegében a korrekt összehasonlításához a tesztet az ismert adatok határának öt perccel korábbra való áthelyezésével kellett volna futtatni. Az eredmények alapján azonban arra lehet következtetni, hogy a platformos rendszerre való áttérés, és adott esetben az ismertetett proaktív mFRR szabályozás módszertanának használata jó lehetőséget nyújt mind az aFRR igények, mind a keresztszabályozás csökkentésére. Többek között ez azért is fontos, mert - új erőművi beruházások hiányában - a jelenlegi szimulációk

alapján is látható, hogy az aFRR termék-piac néhány éves távlatban nem lesz képes kiszolgálni a jelentős PV beépítés miatt prognosztizált nagyobb szabályozási igényeket (rugalmasságot). A másik szempont, amelyben láthatóan van fejlődési lehetőség, az a leszállás, amely a vizsgált időszakban a kevésbé domináns irány volt, ugyanakkor a teljes kiegyenlítetlenségre jellemző arányoktól jóval elmaradt a valós mFRR igénybevételek aránya. Ezzel szemben a tesztek során, ha nem is teljesen arányosan, de jóval nagyobb mértékben részesült a negatív irány.

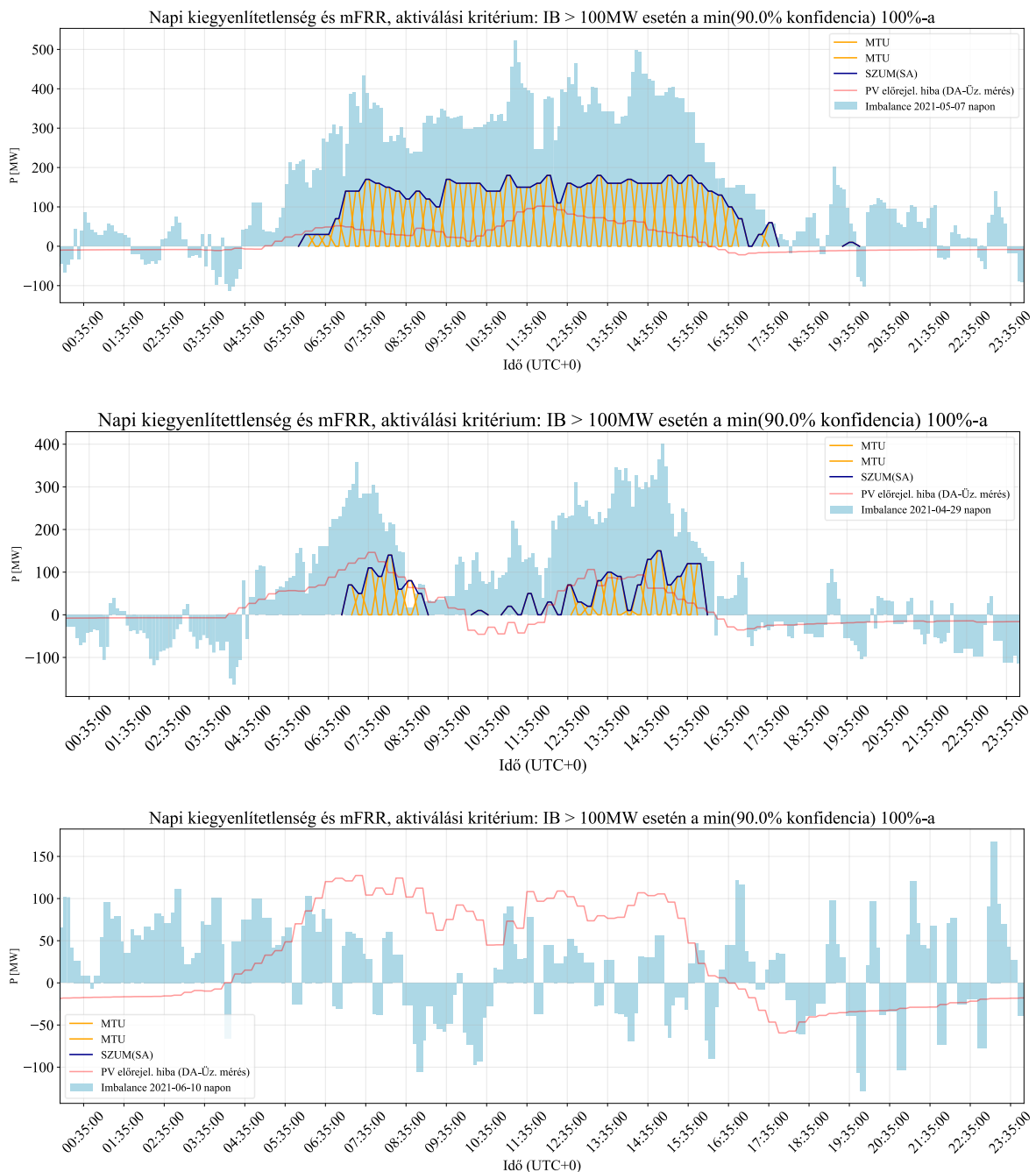
Mivel a proaktív mFRR szabályozás magját az intervallum becslés képi, a későbbiekben ennek további vizsgálata és fejlesztése szükséges. A 2. fejezetben említettem, hogy a célfüggvény viszonylag bonyolult, illetve mostanra azt is beláttuk, hogy a vizsgált néhány kimutatás alapján kapott sorrendek sincsenek feltétlenül összhangban oly módon, hogy könnyedén el lehessen dönteni, milyen paraméterek, kritériumok mellett lenne ideális a megvalósítás. Ehhez kapcsolódóan az 5. táblázatban láthatóak rendre a 90, illetve 80%-os konfidenciaintervallumok adatai a tesztidőszakon. Az első tesztek során már megállapítottam, hogy az RMSE alapú optimalizálás, habár csökkenti az átlagos szélességet, azt a pontosság rovására teszi. Ebben az esetben is az látható, hogy az intervallum hiába szűkült a prediktorok bővítésével, az azon belülre eső pontok aránya csökkent. Ugyanakkor ehhez az összehasonlításához már rendelkezésre állnak a tesztkörnyezetben kapott eredmények is, így 4. Táblázatra visszautalva kijelenthető, hogy pusztán a szélesebb intervallum és magasabb valós pontosság nem garantálja az alacsonyabb keresztszabályozást. Ettől függetlenül az eljárás gyakorlati alkalmazása esetén érdemes 2-3%-kal alacsonyabb valós megbízhatóságot feltételezni a tesztek alapján. Érdekes megfigyelés továbbá, hogy a 25-35 perccel későbbi intervallumok esetén lett a legalacsonyabb a valós pontosság, amely vélhetően szintén a korábban már többször tárgyalt órán belüli ingadozásnak köszönhető. Persze ez nem azt jelenti, hogy könnyebb lenne ennél időben távolabbra becsülni, hiszen látható, hogy a későbbi időpontokhoz, egyre szélesebb intervallum tartozik, viszont érdemes lehet olyan prediktorokkal kísérletezni, amelyek csökkenthetik ezt a visszaesést, különösen a 25 perccel későbbi időpontra, amely az egyik legfontosabb szerepet tölti majd be a platformos környezetben.

5. Táblázat: Az intervallumbecslések átlagos szélességére és valós pontosságára vonatkozó statisztikák független hőtérképpel ábrázolva.

90% konfidenciaintervallum adatai													
		IB T--5	IB T--10	IB T--15	IB T--20	IB T--25	IB T--30	IB T--35	IB T--40	IB T--45	IB T--50	IB T--55	IB T--60
Átlagos szélesség [MW]	Only_IB	102.7	147.4	172.2	186.2	195.2	204.2	209.8	214.4	216.5	218.3	219.5	222.8
	Komplex	89.8	130.2	144.4	157.4	167.9	178.6	188.4	197.0	203.5	210.5	213.2	217.7
	GINI-b.	90.0	130.0	144.9	158.5	167.9	179.3	188.2	196.4	203.8	210.5	213.5	218.7
Valós pontosság [%]	Only_IB	88.88	88.33	87.95	87.86	87.49	87.41	87.34	87.27	87.13	87	86.78	86.8
	Komplex	90.06	88.58	87.85	87.51	86.91	87.13	87.19	87.33	87.61	87.79	87.64	87.62
	GINI-b.	89.89	88.46	88.05	87.44	86.7	87.16	87.16	87.23	87.56	87.52	87.52	87.63
80% konfidenciaintervallum adatai													
		IB T--5	IB T--10	IB T--15	IB T--20	IB T--25	IB T--30	IB T--35	IB T--40	IB T--45	IB T--50	IB T--55	IB T--60
Átlagos szélesség [MW]	Only_IB	72.0	108.3	128.8	139.9	148.2	154.0	158.0	159.9	161.7	163.1	164.6	166.4
	Komplex	65.2	96.7	109.6	119.8	128.6	136.5	142.5	149.6	155.1	159.1	162.2	165.8
	GINI-b.	65.5	97.5	109.4	120.5	128.1	136.0	141.9	149.0	151.5	156.2	160.3	164.5
Valós pontosság [%]	Only_IB	80.09	78.3	77.32	77.36	77.24	77	76.92	76.55	76.28	76.04	75.89	75.93
	Komplex	81.49	78.33	77.57	77.13	76.47	76.46	76.58	77.06	77.3	77.27	77.27	77.45
	GINI-b.	81.41	78.14	77.5	76.81	76.15	76.26	76.25	76.58	76.45	76.5	76.9	77.31

4.3 mFRR statisztikák napi bontásban

A vizsgált napok közötti mFRR aktiválási statisztikákat a GINI relevancia alapú prediktorokkal tanított esetre, 90%-os konfidenciaszint és 100 MW-os minimum mellett a függelékben található 6. táblázat tartalmazza. Ezek alapján kiválasztottam három napot, amelyek a 13. ábrán láthatók. A felső májusi napon kaptam a legnagyobb aktiválható mennyiséget, amely a láthatóan hosszan fennálló felszabályozási igény stabilitása miatt igen kevés hibát eredményezett.



13. Ábra: Napi mFRR statisztikák három választott napra.

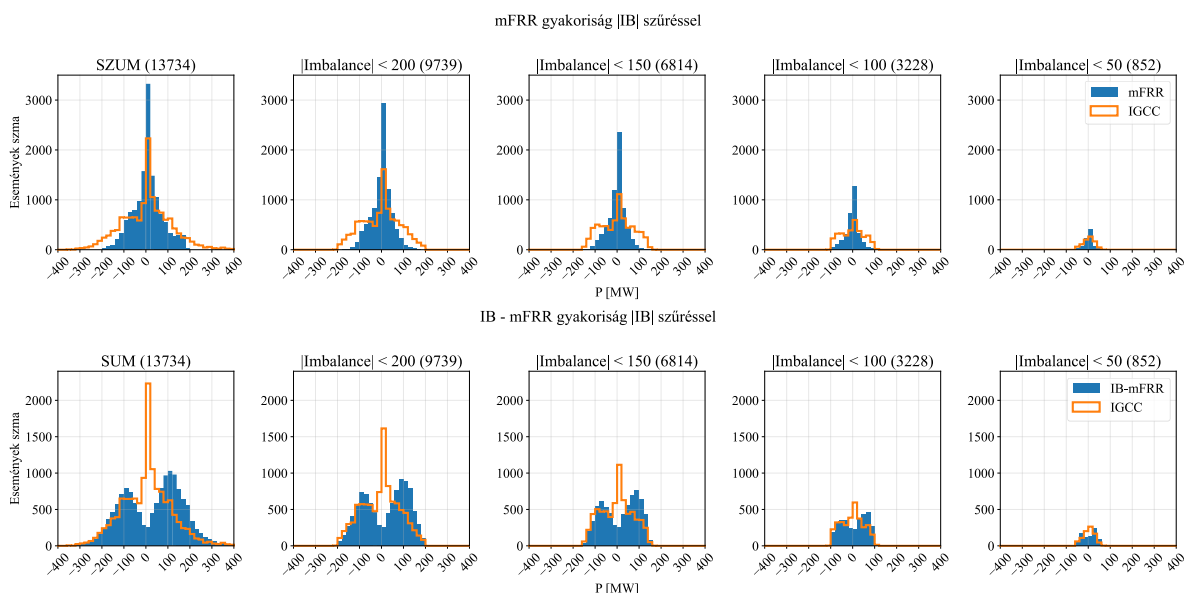
A másodikat alapvetően a magas keresztszabályozás miatt választottam, azonban azért nem a július 9.-i, ilyen szempontból leggyengébb napra esett a választásom, mivel ezen az ábrán a

hibát okozó közel 130 MW-os ugrás mellett jól látszik a napenergia előrejelzési hiba és a kiegyenlítetlenség közti kapcsolat. Habár volumenben jelentős a különbség, a trendben megfigyelhető ugyanaz a hullám mindkét változónál. A teljesség igényével a július 9-i nap megtalálható a függelékben a 20. ábrán.

A harmadik nap egy olyan esetet mutat be, amelyen egyáltalán nem történt aktiválás. Ez alapvetően sikeresnek mondható, hiszen a módszertan helyesen járt el, hogy nem próbálkozott mFRR igénybevétellel, hiszen láthatóan nem lett volna rá szükség.

4.4 mFRR és IGCC eloszlások összehasonlítása

A 3. fejezetben kifejtésre került, hogy az IGCC alapvetően egy nagyon jó lehetőség a kiegyenlítetlenség kooperatív kezelésére, ugyanakkor rendelkezésre állása bizonytalan, így tervezhetősége is gyenge, tehát nem lehet pl. alacsony aFRR esetén egyértelműen számítani rá, illetve arra sem jó, hogy a lekötött tartalékok mennyiségének csökkentse, ezáltal optimalizálva annak költségeit. A tartalékméretezés során sem lehet figyelembe venni, így hiába vállal jelentős szerepet az IGCC a szabályozásban, a minimális aFRR tartalék szint határának megállapításakor nem minősül befolyásoló tényezőnek. A 14. ábrán látható ezzel kapcsolatban 12 kimutatás, amelyeken a valós idejű kiegyenlítetlenség abszolútértéke alapján szűrve láthatóak az mFRR aktiválások, illetve a fennmaradó kiegyenlítetlenség (IB-mFRR) hisztogramjai, ábrázolva az IGCC tapasztalt alakulását. Összesen 13 734 ötperces volt, amelyben az aktiválásra javasolt standard termékek összege nullától különböző értéket vett fel. A kimutatások alapján ezek ritkábban vettek fel extrém értékeket, mint az IGCC, jellemzően -200 és 200 MW közötti sávba estek a javasolt volumenek. Az vizsgált ötpercek nagyjából 40%-a esett olyan intervallumokra, amelyben a valós kiegyenlítetlenség abszolútértéke 150 MW alatt volt, illetve csupán 5%-a 50 MW-nál kisebb esetekre. Ez alapvetően pozitív tapasztalat, hiszen kis igények esetén kevésbé indokolt az mFRR aktiválás.



14. Ábra: mFRR, illetve kiegyenlítetlenség-mFRR hisztogramok kiegyenlítetlenség abszolútérték alapján szűrve a tesztidőszak azon időszakain, amelyben a javasolt aktiválásnullától különböző volt.

Az alsó diagramokon a maradó kiegyenlítetlenség eloszlása látható. Ha azt szeretnénk, hogy az IGCC-t nem csökkentve lehessen aFRR igényeket csökkenteni, az ideális az lenne, ha a maradó kiegyenlítetlenség és az IGCC azonos eloszlást mutatna. A diagramok alapján ez

viszonylag jónak mondható, az IGCC a kis értékű, és valamelyest a szélsőségek esetében mutat nagyobb gyakoriságot. Habár ezek a kimutatások az egyidejűségről nem sok információt tartalmaznak, ugyanakkor pusztán eloszlás alapján azt sugallják, hogy közel annyi javasolt aktiválás történt, amely mellett a maradék kiegyenlítetlenség IGCC-vel fedezhető lett volna. Tehát ebből a mennyiségből, amelyet az algoritmus javasolt, lehet mérlegelni gazdasági, illetve rendelkezésre állási szempontokat figyelembe véve, hogy mekkora hányad kerüljön mFRR-ből fedezve és mennyi maradjon a hazai, vagy nemzetközi aFRR aktiválásra.

5 Összefoglalás

A kutatás célja a rendszer szintű kiegyenlítetlenség rövidtávú előrejelzése gépi tanulással, amelyet a közeljövőben élesedő PICASSO és MARI össz-európai tartalék platformokhoz való csatlakozás után (illetve opcionálisan az átállásig is – de más termékalakkal) lehet majd használni proaktív mFRR szabályozásra. A tanulmány első részében ismertettem az új szabályozási termékek (FCR, aFRR, mFRR, RR), illetve kiemelten a MARI platform jellemzői, kritériumai. A szakirodalmi áttekintés során láthattuk, hogy a kutatási terület igencsak szerény felhozattal bírt a dolgozat elkészültéig, a néhány publikált anyag jellemzően a kiegyenlítetlenség előrejelzést eltérő felbontással, módszerrel és időhorizonttal vizsgálta, valamint az alkalmazási terület sem egyezett az általam vizsgáltakkal.

A második fejezetben bemutattam a kutatás során alkalmazott módszertant, illetve többek között a felhasznált adatokról, tanuló algoritmusokról, a pontbecslések értékelési szempontjairól, illetve a platformos tesztkörnyezet felépítéséről olvashattunk. Ezután az adatok elemzése következett, különböző vizualizációs technikákkal szemléltettem az egyes változók közti kapcsolatot. Elsőként a korreláció vizsgálata alapján fogalmaztam meg, hogy a kiegyenlítetlenségben van egy jelentős órás hullámozás, amely az import-export menetrendváltozásokra vezethető vissza, amely átvezette a figyelmet az IGCC-re, annak rendelkezésre állásának vizsgálatára. Ezzel kapcsolatban a napon belüli boxplotok segítségével megállapításra került, hogy a menetrendváltásokból adódó hullámozás a reggeli és esti órákban egyértelműen dominálnak, amely trendet nap közben az időjárásfüggő megújulók torzítanak, jóval nagyobb szórást eredményezve a kialakuló kiegyenlítetlenségben. A változók GINI relevancia alapú vizsgálatával a kezdeti 54 prediktorból kiválasztásra került a 19 legjobb, e kettő (*Komplex*, *GINI-based*), illetve egy csak historikus kiegyenlítetlenség adatokból álló prediktor kombinációval teszteltem az algoritmusokat (*Only_IB*).

Minden algoritmust a különböző esetekre a következő egy óra 12 db ötperces intervallumára külön került tanításra, az adott intervallumra egyedileg illesztve. A negyedik fejezetben elsőként a pontbecslések eredményeit ismertettem, amelyek kapcsán a következő megállapításokat emelném ki:

1. **Prediktorok megválasztása** – a csak kiegyenlítetlenség adatok figyelembevételéhez képest akár 15-20% javulás érhető el a magyarázó változók komplexebb csoportjának alkalmazásával.
2. **Modell választás** – 5-10% javulás érhető el az egyszerű lineáris modellekhez képest a komplexebb algoritmusok alkalmazásával.
3. **Újratanítás** – jellemzően 1-2%-nál kisebb javulási potenciállal bír az elemzésben alkalmazott ötszöri újratanítás.

A 3. Táblázatból kiderült, hogy a legjobb módszerek az MLP, illetve XGBoost voltak, illetve az, hogy ezek becsléseinek átlagolásával a pontosság tovább növelhető. A kétórás mozgóátlaghoz képest 30%-kal jobb eredményeket lehetett elérni, a felhasználás szempontjából legfontosabbnak számító T+25-ös intervallumon 50,55 MW-os RMSE-vel, illetve 42,47%-os relatív MAE-vel.

A platformos környezetben (ugyanúgy közel 6 hónapnyi adaton) GBR alapú intervallum becsléseket végeztem, amely alapján különböző aktiválási protokollok mentén vizsgáltam a maximálisan aktiválásra javasolt mFRR szabályozás mennyiségét. Megállapítottam, hogy a relevancia alapján szelektált prediktorokkal érhető el a legmagasabb aktiválás, közel azonos egységnyi aktiválásra jutó keresztszabályozással, mint a Komplex esetben. A csak kiegyenlítetlenség adatokon alapuló előrejelzések jelentősen gyengébbnek bizonyultak,

kevesebb aktiválás mellett is több, mint másfélszeres hibát eredményeztek, mint a másik két típus esetén. A protokollban előírt minimum érték változtatásával a tapasztalatok szerint az aktiválható szabályozás a várt módon jelentősen, míg az említett hibaarány csak enyhén csökkenthető. Az eredményeket számszerűsítve, 90%-os konfidenciaszint mellett maximálisan 53 258 MWh szabályozást lehetett volna előre aktiválni, amely a teljes kiegyenlítetlenség 13,36%-át fedezte volna. Ugyanez 571,9 MWh keresztszabályozással járt volna, amely 1,07% Error/mFRR arányt jelent. A proaktív szabályozás a tapasztalatok szerint mindenképpen jár keresztszabályozással, ugyanakkor ennek mennyisége jelenlegi módszertannal okozott hiba kevesebb, mint harmadának bizonyult, közel háromszoros volumenek mellett is. A vizsgált újratanítás a maximálisan javasolt mennyiséget enyhén csökkentette, ugyanakkor 15%-kal kisebb hibaarányt eredményezett. Ebből következik, hogy az újratanítás rendszeres jelleggel javasolt, ugyanakkor a rendszerben történő komolyabb változások esetén érdemes felmérni, hogy az milyen szinten befolyásolta a szabályozást, és ennek megfelelően kell mérlegelni az újratanítást és megválasztani a továbbiakban használt historikus adatsort.

A vizsgált időszak napjai közül bemutattam néhány példát, szemléltetve, hogy milyen körülmények között lehet effektíven végezni a proaktív mFRR szabályozást, illetve mikor jár jelentős hibával. Végül pedig az aktiválás és a valóságban rendelkezésre állt IGCC eloszlását vizsgálva megállapítottam, hogy a proaktív mFRR aktiválás után maradó kiegyenlítetlenséget az egyidejűséget nem vizsgálva, jól lefedti az IGCC.

Összességében tehát az elkészített elemzés alapján a proaktív mFRR szabályozás gyakorlati megvalósítása javasolt, akár a platformos időszak előtt is. A GBR alapú intervallum becslés elfogadható pontosságú, az aktiválási kritériumokat pedig a felhasználó által javasolt módon, az elfogadható hibára kalibrálva kell megválasztani. A bemutatott módszertan a tesztek alapján alkalmas arra, hogy a kiegyenlítőenergia rendszerbe komolyabb erőművi beruházások nélkül, pusztán a tervezhetőség fejlesztésével egyszerre csökkentse a már meglévő aFRR kapacitások kihasználtságát a rendelkezésre álló manuális tartalékok aktívabb igénybevételeivel, és csökkentse az okozott keresztszabályozás mértékét.

6 Hivatkozások

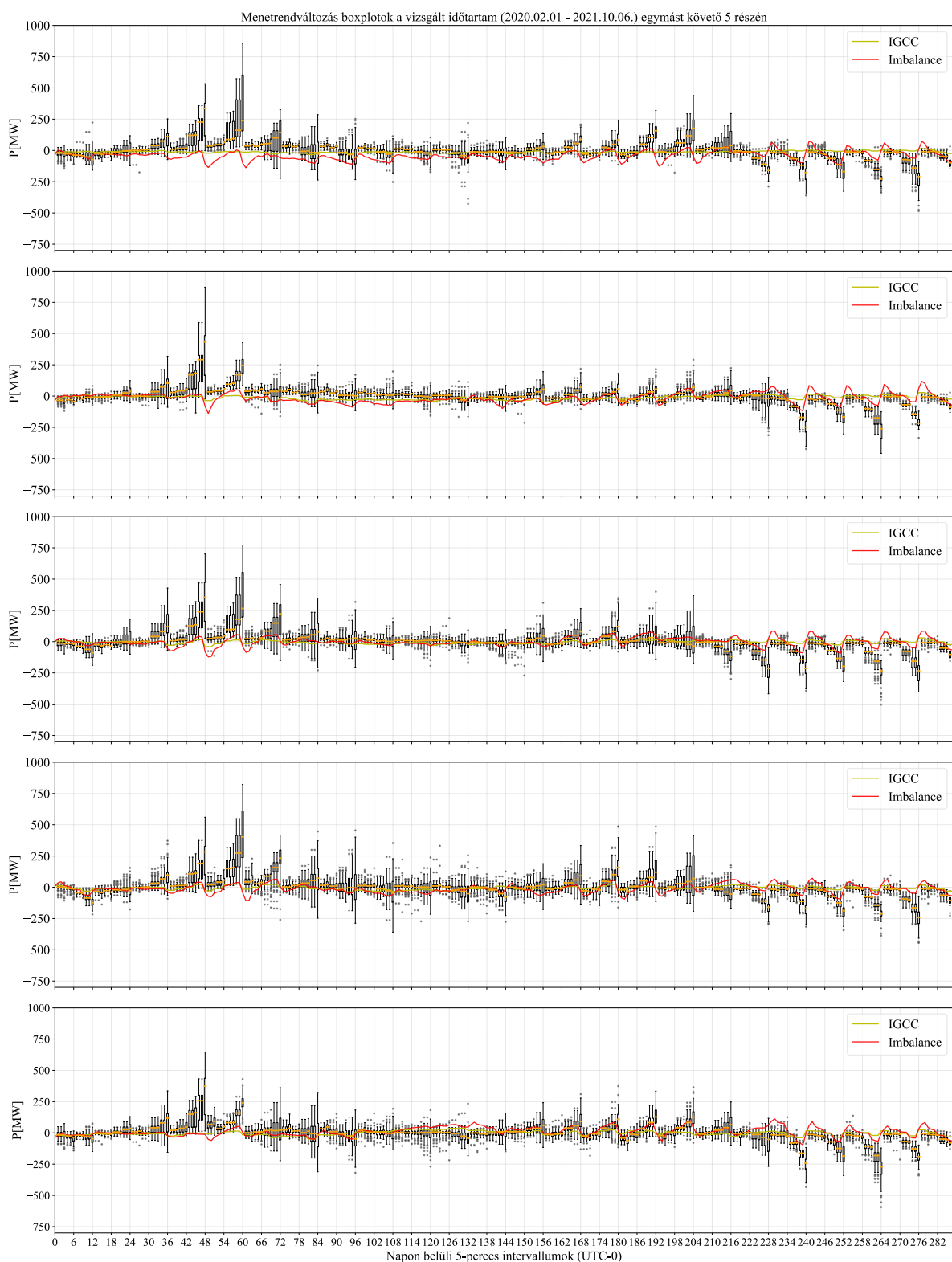
- [1] ENTSO-E, “Manually Activated Reserves Initiative.” https://www.entsoe.eu/network_codes/eb/mari/ (accessed Sep. 26, 2021).
- [2] ENTSO-E, “PICASSO.” https://www.entsoe.eu/network_codes/eb/picasso/ (accessed Sep. 26, 2021).
- [3] ENTSO-E (European Network of, Transmission System Operators, and for Electricity), “An overview of the European balancing market and electricity balancing guideline,” Nov. 2018, Accessed: Sep. 18, 2021. [Online]. Available: https://eepublicdownloads.entsoe.eu/clean-documents/Network%20codes%20documents/NC%20EB/entso-e_balancing_in%20europe_report_Nov2018_web.pdf
- [4] “TERRE.” https://www.entsoe.eu/network_codes/eb/terre/ (accessed Oct. 22, 2021).
- [5] ENTSO-E, “Imbalance Netting.” https://www.entsoe.eu/network_codes/eb/imbalance-netting/ (accessed Sep. 26, 2021).
- [6] ENTSO-E, “Explanatory document to all TSOs’ proposal for the implementation framework for a European platform for the exchange of balancing energy from frequency restoration reserves with manual activation in accordance with Article 20 of Commission Regulation (EU) 2017/2195 establishing a guideline on electricity balancing” Dec. 2018. Accessed: Sep. 26, 2021. [Online]. Available: https://eepublicdownloads.entsoe.eu/clean-documents/nc-tasks/EBGL/EBGL_A20_181218_ALL%20TSOs%20proposal_mFRRIF_explanatory_document_for%20submission.pdf?Web=0
- [7] M. Haberg and G. Doorman, “Proactive planning and activation of manual reserves in sequentially cleared balancing markets,” *2017 IEEE Electrical Power and Energy Conference, EPEC 2017*, vol. 2017-October, pp. 1–6, Feb. 2018, doi: 10.1109/EPEC.2017.8286186.
- [8] COMMISSION REGULATION - establishing a guideline on electricity transmission system operation. Official Journal of the European Union. Accessed: Oct. 22, 2021. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2017.220.01.0001.01.ENG&toc=OJ:L:2017:220:TOC#d1e9466-1-1
- [9] MAVIR Zrt., *A MAVIR Magyar Villamosenergia-ipari Átviteli Rendszerirányító Zártkörűen Működő Részvénytársaság javaslata az LFC blokkra vonatkozó üzemviteli módszertanra a villamosenergia-átviteli hálózat üzemeltetésére vonatkozó iránymutatás megalkotásáról szóló 2017/1485 Bizottság (EU) rendelet 119. cikk (1) bekezdése szerint.* Nemzetközi Üzemi és Kereskedelmi Szabályzat, 2018, pp. 9–14. Accessed: Oct. 22, 2021. [Online]. Available: https://www.mavir.hu/documents/10258/238462133/4.11+LFC_Blokk_uzemviteli_modszeran_HUN.pdf/017b3caa-1b41-65d0-924e-9b4dd6565f8e?t=1605785938004
- [10] U. K. Das *et al.*, “Forecasting of photovoltaic power generation and model optimization: A review,” *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 912–928, Jan. 2018, doi: 10.1016/J.RSER.2017.08.017.
- [11] Y. Wang, R. Zou, F. Liu, L. Zhang, and Q. Liu, “A review of wind speed and wind power forecasting with deep neural networks,” *Applied Energy*, vol. 304, p. 117766, Dec. 2021, doi: 10.1016/J.APENERGY.2021.117766.

- [12] G. Hafeez, I. Khan, S. Jan, I. A. Shah, F. A. Khan, and A. Derhab, “A novel hybrid load forecasting framework with intelligent feature engineering and optimization algorithm in smart grid,” *Applied Energy*, vol. 299, p. 117178, Oct. 2021, doi: 10.1016/J.APENERGY.2021.117178.
- [13] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, “Review of photovoltaic power forecasting,” *Solar Energy*, vol. 136, pp. 78–111, 2016, doi: 10.1016/j.solener.2016.06.069.
- [14] M. P. Garcia and D. S. Kirschen, “Forecasting system imbalance volumes in competitive electricity markets,” *2004 IEEE PES Power Systems Conference and Exposition*, vol. 3, pp. 1805–1812, 2004, doi: 10.1109/PSCE.2004.1397617.
- [15] Š. Kratochvíl, “System Imbalance Forecast,” Prague, 2016. Accessed: Oct. 15, 2021. [Online]. Available: https://dspace.cvut.cz/bitstream/handle/10467/67662/Disertace_%C5%A0t%C4%9Bp%C3%A1n_Kratochv%C3%ADl_2016.pdf?sequence=1&isAllowed=y
- [16] C. Contreras, “System imbalance forecasting and short-term bidding strategy to minimize imbalance costs of transacting in the spanish electricity market,” Madrid, 2016. Accessed: Oct. 15, 2021. [Online]. Available: <http://hdl.handle.net/11531/16621>
- [17] T. S. Salem, K. Kathuria, H. Ramampiaro, and H. Langseth, “Forecasting Intra-Hour Imbalances in Electric Power Systems”, Accessed: Oct. 15, 2021. [Online]. Available: www.aaai.org
- [18] Elia Group Innovation, “Advanced Machine Learning to support dispatching.” <https://innovation.eliagroup.eu/projects/advanced-machine-learning-to-support-dispatching/> (accessed Oct. 17, 2021).
- [19] Elia Transmission Belgium SA/NV, “System Imbalance forecast and its publication to stakeholders,” Brussels, 2021. Accessed: Oct. 17, 2021. [Online]. Available: https://www.elia.be/en/public-consultation/20210831_public-consultation-on-the-system-imbalance-forecast-and-evaluation-of-its-publication
- [20] “Rendszerterhelés - MAVIR - Magyar Villamosenergia-ipari Átviteli Rendszerirányító Zrt.” <https://www.mavir.hu/web/mavir/rendszerterheles> (accessed Oct. 15, 2021).
- [21] E. Alpaydin, *Introduction to Machine Learning*, 2nd ed. The MIT Press, 2010.
- [22] T. Agrawal, “Introduction to Hyperparameters,” *Hyperparameter Optimization in Machine Learning*, pp. 1–30, 2021, doi: 10.1007/978-1-4842-6579-6_1.
- [23] “6.2.1.1. Primary energy consumption data, by month and cumulative from the beginning of the year.” https://www.ksh.hu/stadat_files/ene/en/ene0016.html?utm_source=kshhu&utm_medium=banner&utm_campaign=theme-energy (accessed Oct. 11, 2021).
- [24] “Hungary - Countries & Regions - IEA.” <https://www.iea.org/countries/hungary> (accessed Oct. 11, 2021).
- [25] SIDC, “SIDC Stakeholder Report,” Jun. 2021.
- [26] F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [27] X. Dang, H. Peng, X. Wang, and H. Zhang, “Theil-Sen Estimators in a Multiple Linear Regression Model.” Jul. 17, 2009. Accessed: Apr. 07, 2021. [Online]. Available: <http://home.olemiss.edu/>
- [28] A. B. Owen, “A robust hybrid of lasso and ridge regression,” *Contemporary Mathematics*, vol. 443, 2007, doi: <http://dx.doi.org/10.1090/conm/443>.

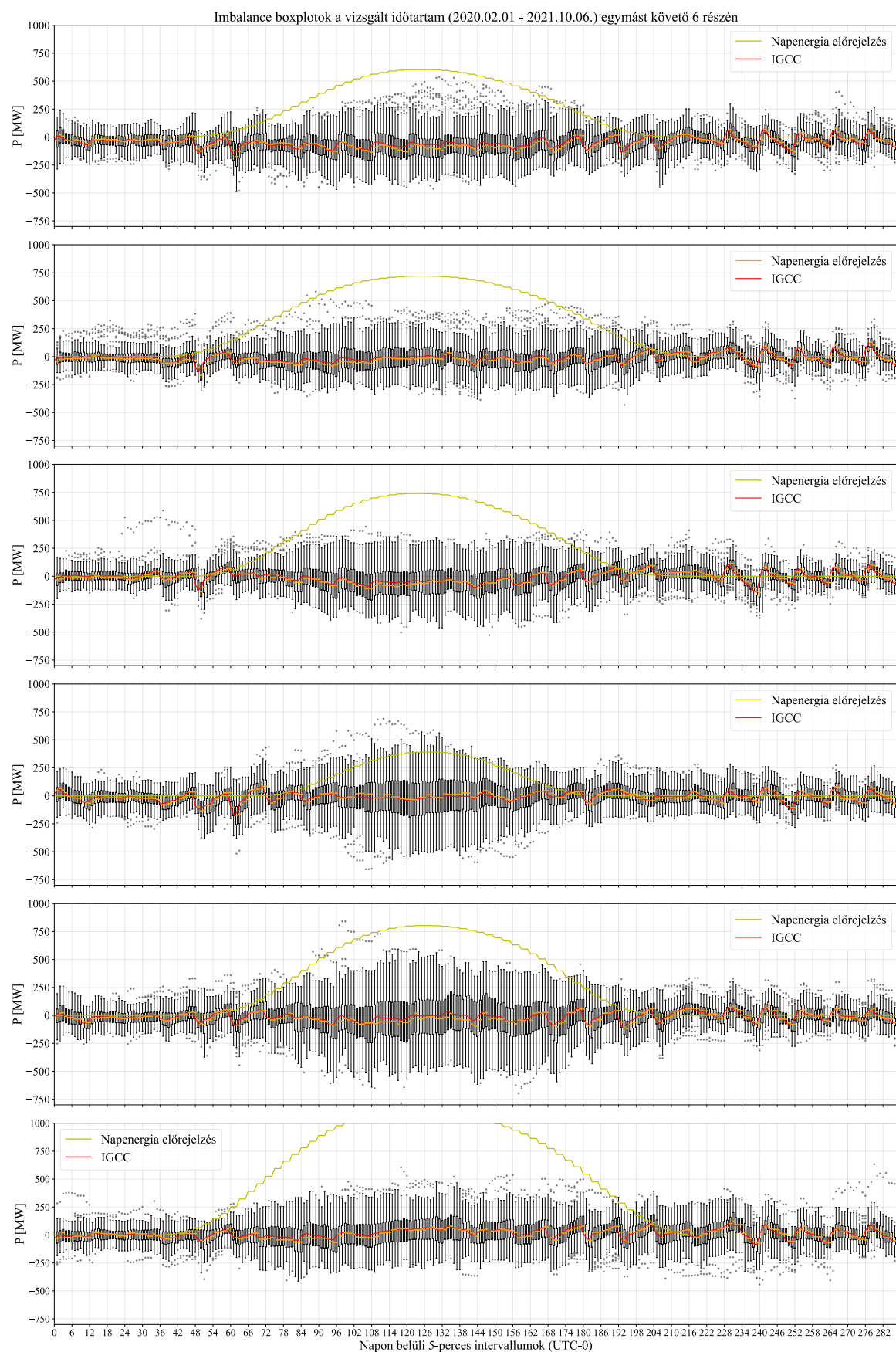
- [29] O. Kramer, “K-Nearest Neighbors,” in *Dimensionality Reduction with Unsupervised Nearest Neighbors*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 13–23. doi: 10.1007/978-3-642-38652-7_2.
- [30] D. P. Kingma and J. Lei Ba, “Adam: A method for stochastic optimization,” 2015.
- [31] L. Breiman, “Random Forests,” *Machine Learning* 2001 45:1, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
- [32] P. Geurts, D. Ernst, and L. Wehenkel, “Extremely randomized trees,” *Mach Learn*, vol. 63, pp. 3–42, 2006, doi: 10.1007/s10994-006-6226-1.
- [33] J. H. Friedman, “Greedy function approximation: A gradient boosting machine,” *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001, doi: 10.1214/aos/1013203451.
- [34] XgboostDevelopers, “Xgboost docs,” 2020. <https://xgboost.readthedocs.io/en/latest/index.html>
- [35] C. Microsoft, “LightGBM docs,” 2021. <https://lightgbm.readthedocs.io/en/latest/Features.html>
- [36] V. Vovk, “Kernel Ridge Regression,” in *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik*, B. Schölkopf, Z. Luo, and V. Vovk, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 105–116. doi: 10.1007/978-3-642-41136-6_11.
- [37] C.-C. Chang and C.-J. Lin, “LIBSVM: A Library for Support Vector Machines,” 2001. Accessed: Apr. 07, 2021. [Online]. Available: www.csie.ntu.edu.tw/
- [38] T. Agrawal, “Hyperparameter Optimization Using Scikit-Learn,” *Hyperparameter Optimization in Machine Learning*, pp. 31–51, 2021, doi: 10.1007/978-1-4842-6579-6_2.
- [39] L. Hertel, J. Collado, P. Sadowski, J. Ott, and P. Baldi, “Sherpa: Robust hyperparameter optimization for machine learning,” *SoftwareX*, vol. 12, p. 100591, Jul. 2020, doi: 10.1016/J.SOFTX.2020.100591.
- [40] Sophie Marquet, “MARI Algorithm Design Principles,” Nov. 2018. Accessed: Sep. 26, 2021. [Online]. Available: www.n-side.com
- [41] de B. Louis, “Incremental (Online) Learning with Scikit-Multiflow ,” *Towards Data Science*. <https://towardsdatascience.com/incremental-online-learning-with-scikit-multiflow-6b846913a50b> (accessed Oct. 15, 2021).
- [42] “Periodische Muster in der Netzfrequenz | Gridradar.” <https://gridradar.net/en/blog/post/analysis-periodic-patterns-network-frequency> (accessed Oct. 17, 2021).
- [43] “IGCC Merger.” https://www.entsoe.eu/network_codes/eb/imbalance-netting/ (accessed Oct. 12, 2021).
- [44] “Kiegyenlítő energia elszámolás - MAVIR - Magyar Villamosenergia-ipari Átviteli Rendszerezényítő Zrt.” <https://www.mavir.hu/web/mavir/kiegyenlito-energia-elszamolas> (accessed Oct. 12, 2021).
- [45] D. Markovics, “Különböző gépi tanulási algoritmusok optimalizálása és összehasonlítása napelemparkok statisztikai alapú termelés előrejelzésére - BME TDK Portál,” Nov. 2020. <https://tdk.bme.hu/GPK/ViewPaper/Kulonbozo-gepi-tanulasi-algoritmusok> (accessed Oct. 12, 2021).
- [46] MAVIR Zrt., “KÁT és Prémium rendszer havi monitoring jelentés,” May 2021. Accessed: Oct. 12, 2021. [Online]. Available: https://www.mavir.hu/documents/10258/0/KAT_es_premium_tamogatas_statisztikai_a_datai_202105_fin.xlsx/4f921ca2-b5f6-edac-447e-2e8c9154e96d?t=1624617544787

- [47] “A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data,” 2009, doi: 10.1186/1471-2105-10-213.
- [48] Ceshine Lee, “Feature Importance Measures for Tree Models — Part I ,” *Medium*. <https://medium.com/the-artificial-impostor/feature-importance-measures-for-tree-models-part-i-47f187c1a2c3> (accessed Oct. 12, 2021).
- [49] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, “Classification And Regression Trees,” *Classification and Regression Trees*, pp. 1–358, Oct. 2017, doi: 10.1201/9781315139470.

7 Függelék



15. Ábra: Képzett menetrendváltás boxplotok kiegészítve az átlagos kiegyenlítetlenség és IGCC értékekkel hat egymást követő időszakra



16. Ábra: Kiegyenlítetlenség boxplotok kiegészítve az átlagos napenergia előrejelzés és IGCC értékekkel hat egymást követő időszakra

		Pontbecslések eredményei - dir%											
Komplex	mid	94.44	90.79	88.6	86.96	85.77	84.52	83.69	82.78	82.05	81.57	81.11	80.58
GINI-based	mid	94.19	90.49	88.52	86.92	85.47	84.28	83.61	82.59	81.96	81.44	81.07	80.78
Komplex	mid_xgb	94.6	90.54	88.43	86.71	85.36	84.19	83.27	82.43	81.77	81.33	80.73	80.38
GINI-based	mid_xgb	94.45	90.4	88.47	86.78	85.25	84.11	83.4	82.58	81.87	81.18	80.77	80.39
Komplex	mid_mlp	93.96	90.32	88.06	86.25	85.06	84.02	82.86	81.96	81.02	80.73	80.35	79.73
Komplex	mid_lightgbm	94.15	89.95	88.09	86.43	84.87	84	82.87	82.3	81.66	81.21	80.65	80.27
GINI-based	mid_lightgbm	94.15	89.97	88.01	86.37	84.67	83.82	82.98	82.14	81.73	81.13	80.73	80.26
GINI-based	mid_mlp	93.65	89.94	87.98	86.38	84.66	83.75	82.83	81.68	81.14	80.76	80.46	80.03
Komplex	mid_et	93.85	89.28	87.13	85.16	83.88	82.62	81.78	80.89	80.37	79.9	79.62	79.38
Komplex	mid_rf	93.9	89.39	87.12	85.14	83.85	82.55	81.64	81.01	80.63	80.05	79.79	79.57
GINI-based	mid_et	93.91	89.35	87.22	85.22	83.8	82.69	81.92	81.05	80.68	79.98	79.47	79.34
GINI-based	mid_rf	93.95	89.31	87.1	85.3	83.57	82.41	81.65	81.09	80.66	80.07	79.76	79.52
Komplex	mid_lr	93.49	88.75	86.86	84.86	83.42	81.99	81.12	80.28	79.68	79.39	78.91	78.79
Komplex	mid_huber	93.59	88.99	86.93	85.01	83.42	81.98	81.13	80.28	79.65	79.41	78.92	78.74
Komplex	mid_tr	93.51	88.69	86.9	84.9	83.33	81.9	80.93	80.15	79.55	79.28	78.74	78.79
GINI-based	mid_lr	93.27	88.35	86.58	84.56	83.03	81.52	80.53	80.02	79.53	79.41	79.04	78.81
GINI-based	mid_huber	93.4	88.45	86.58	84.66	82.95	81.5	80.48	79.93	79.5	79.41	78.98	78.77
GINI-based	mid_tr	93.26	88.24	86.57	84.41	82.85	81.31	80.48	79.88	79.34	79.32	78.89	78.92
Only_IB	mid_mlp	93.13	88.38	85.14	83.36	82	80.7	80.1	79.56	79.26	78.81	78.51	78.33
Only_IB	mid	93.21	88.48	85.24	83.42	81.95	80.71	80.13	79.71	79.28	78.86	78.68	78.34
Only_IB	mid_xgb	93.06	88.22	84.66	83.01	81.47	80.33	79.84	79.42	78.9	78.58	78.4	78.03
Only_IB	mid_et	92.66	87.76	84.13	82.48	81.21	79.81	79.25	78.88	78.49	78.29	78.03	77.87
Only_IB	mid_lightgbm	92.9	87.97	84.32	82.58	81.13	80.1	79.51	79.19	78.9	78.67	78.42	78.1
Only_IB	mid_huber	93.03	87.88	83.95	82.29	80.91	79.74	79.22	78.84	78.46	78.41	78.33	78.25
Only_IB	mid_lr	93	87.75	83.99	82.31	80.87	79.68	79.18	78.85	78.4	78.38	78.3	78.15
Only_IB	mid_rf	92.36	87.76	84.12	82.33	80.85	79.8	79.26	78.81	78.51	78.28	78.04	77.79
Only_IB	mid_tr	92.96	87.62	83.99	82.23	80.69	79.54	79.08	78.79	78.39	78.36	78.09	77.97
		IB T--5	IB T--10	IB T--15	IB T--20	IB T--25	IB T--30	IB T--35	IB T--40	IB T--45	IB T--50	IB T--55	IB T--60

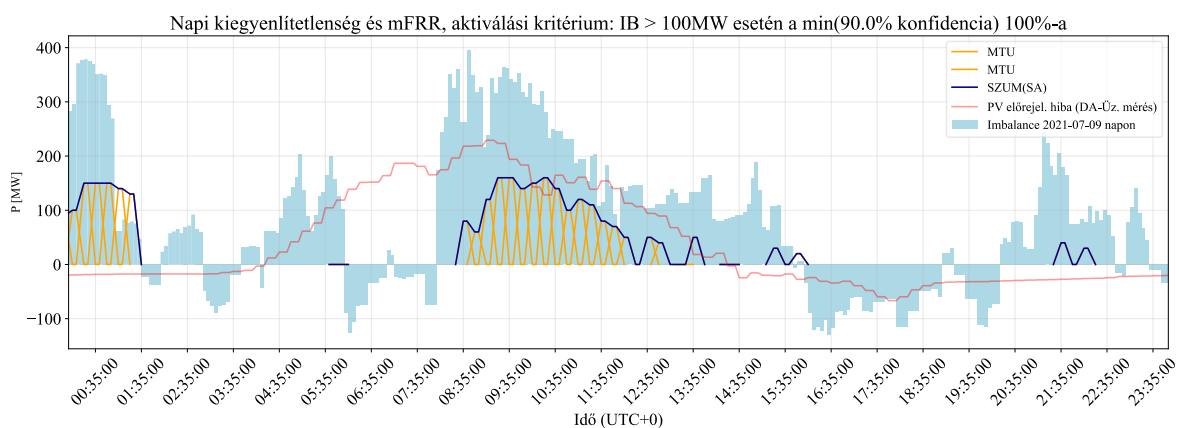
17. Ábra: Pontbecslések átlagos irány hatékonysága (dir%) eredményei egyszeri tanítással.

		Pontbecslések eredményei - MAE											
Komplex	mid	17.77	26.52	31.16	34.98	38.35	40.88	43.37	45.98	48.44	49.83	51.32	52.64
GINI-based	mid	18.53	27.37	31.66	35.23	38.84	41.63	43.89	46.92	48.77	50.17	51.54	52.84
Komplex	mid_xgb	17.69	27.32	31.83	36	39.69	42.28	44.82	47.37	49.63	51.47	53.02	53.87
Komplex	mid_mlp	19.57	27.9	32.66	36.76	39.79	42.3	45.36	47.97	50.88	51.97	53.42	55.34
GINI-based	mid_xgb	18.15	27.85	32.16	36.05	39.85	42.52	45.03	47.64	50.18	51.76	53.5	54.8
GINI-based	mid_mlp	20.26	28.82	33.15	36.64	40.31	43.15	45.55	48.75	50.43	51.7	52.77	54.25
Komplex	mid_lightgbm	18.48	28.47	32.59	36.64	40.36	42.66	45.31	47.66	49.78	51.47	52.74	53.96
GINI-based	mid_lightgbm	18.79	28.87	32.93	36.75	40.82	43.01	45.4	47.83	50.08	51.32	52.66	53.97
GINI-based	mid_et	19.3	30.29	34.97	39.46	42.91	45.54	48	50.36	51.9	53.26	54.44	55.6
Komplex	mid_et	19.26	30.24	35.04	39.54	42.98	45.7	48.24	50.53	52.13	53.52	54.5	55.57
Komplex	mid_rf	19.46	30.52	34.96	39.63	43.37	45.8	48.23	50.26	51.76	53.17	54.41	55.33
GINI-based	mid_rf	19.48	30.64	35.07	39.66	43.4	46.05	48.4	50.43	52	53.28	54.31	55.33
Komplex	mid_huber	19.09	30.76	35.12	40.04	44	47.2	49.86	52.2	53.87	55.14	56.53	57.69
Komplex	mid_lr	20.21	31.4	35.62	40.67	44.28	47.33	49.97	52.35	53.93	55.22	56.53	57.6
Komplex	mid_tr	19.92	31.44	35.42	40.7	44.45	47.56	50.68	52.78	55.07	55.88	57.1	57.72
GINI-based	mid_huber	19.96	32.21	36.41	41.19	45.13	48.56	51.01	53.14	54.4	55.41	56.5	57.58
GINI-based	mid_lr	20.89	32.55	36.62	41.48	45.15	48.44	50.89	53.03	54.33	55.34	56.36	57.35
GINI-based	mid_tr	20.72	32.6	36.67	41.91	45.32	49.31	51.46	53.89	55.11	55.81	56.72	57.52
Only_IB	mid	20.46	31.84	39.26	43.42	46.82	49.62	51.41	53.13	54.37	55.44	56.41	57.57
Only_IB	mid_mlp	21.64	32.31	39.58	43.58	47.07	49.8	51.54	53.42	54.56	55.57	56.57	57.76
Only_IB	mid_xgb	20.76	32.68	40.61	44.65	47.88	50.75	52.49	53.97	55.08	56.15	57.04	58.05
Only_IB	mid_lightgbm	21.1	33.18	41.18	45.23	48.47	51.14	52.78	54.09	55.24	56.21	57.02	57.9
Only_IB	mid_et	22.15	34.06	41.64	45.7	49.08	51.73	53.4	54.8	56.03	56.91	57.7	58.75
Only_IB	mid_lr	21.4	34	42.07	45.99	49.13	51.84	53.43	54.71	55.83	56.52	57.12	57.94
Only_IB	mid_huber	20.44	33.52	41.91	45.94	49.17	51.93	53.54	54.83	55.93	56.6	57.17	57.98
Only_IB	mid_rf	22.53	34.33	41.82	46.09	49.39	52.11	53.66	54.92	56.14	57.08	57.9	58.71
Only_IB	mid_tr	21.53	34.27	42.12	46.31	49.66	52.32	53.66	55.15	56.2	57.21	57.98	58.52
		IB T--5	IB T--10	IB T--15	IB T--20	IB T--25	IB T--30	IB T--35	IB T--40	IB T--45	IB T--50	IB T--55	IB T--60

18. Ábra: Pontbecslések MAE eredményei egyszeri tanítással.

		Skill score a pontbecslésekre											
Komplex	mid	55.3	43.66	39.41	35.2	32.19	30.18	28.42	26.45	24.78	24.43	24.11	24.18
GINI-based	mid	53.68	41.97	38.34	34.78	31.37	29.2	27.69	25.19	24.31	24.09	23.96	23.92
Komplex	mid_mlp	52.53	41.46	36.99	32.26	30.06	27.95	25.53	23.52	21.16	21.32	21.22	20.82
Komplex	mid_xgb	54.86	41.75	37.8	33.14	29.61	27.8	25.89	24.14	22.76	21.95	21.56	22.24
GINI-based	mid_xgb	53.81	40.64	37.09	32.98	29.35	27.48	25.65	23.7	21.93	21.54	20.64	20.85
GINI-based	mid_mlp	51.11	39.52	35.94	32.69	29.19	26.91	25.32	22.84	22.15	22.14	22.61	22.38
Komplex	mid_lightgbm	53.38	39.49	36.33	32.18	28.62	27.43	25.41	23.92	22.74	22.06	22.27	22.37
GINI-based	mid_lightgbm	52.63	38.79	35.66	31.78	27.83	26.82	25.25	23.58	22.28	22.23	22.18	22.08
GINI-based	mid_et	50.83	35.39	31.92	27.31	24.25	22.46	20.74	19.42	19.36	19.38	19.66	20.13
Komplex	mid_et	50.99	35.43	31.8	27.21	24.13	22.14	20.4	19.19	19.02	19.04	19.65	20.18
Komplex	mid_rf	50.6	35.01	31.97	26.91	23.4	22.05	20.5	19.66	19.52	19.28	19.53	20.15
GINI-based	mid_rf	50.6	34.78	31.7	26.87	23.36	21.64	20.3	19.42	19.2	19.29	19.68	20.27
Komplex	mid_huber	50.75	34.95	31.89	26.56	22.87	20.15	18.14	17.11	16.59	16.65	16.71	17.04
Komplex	mid_tr	50.53	34.29	31.43	25.57	22.09	19.56	16.95	16.17	14.76	15.65	15.8	16.88
Komplex	mid_lr	46.78	32.28	29.42	24.23	21.31	18.99	17.23	16.2	15.91	15.98	16.19	16.65
GINI-based	mid_lr	48.88	32.14	29.38	24.27	21.06	18.38	16.74	15.9	15.98	16.42	16.96	17.56
GINI-based	mid_huber	48.93	32.21	29.64	24.71	21.05	18.15	16.51	15.68	15.8	16.22	16.69	17.15
GINI-based	mid_tr	48.96	32	29.21	23.68	20.88	17.16	16.08	14.82	15.05	15.99	16.62	17.34
Only_IB	mid	47.32	31.27	23.75	19.8	17.46	15.69	13.55	14.86	15.27	15.96	16.67	17.22
Only_IB	mid_mlp	45.75	30.64	23.47	19.67	17.24	15.61	13.41	14.65	15.13	15.97	16.52	17.04
Only_IB	mid_xgb	46.72	29.62	21.11	17.6	15.55	13.85	13.5	13.43	14.12	14.74	15.64	16.39
Only_IB	mid_lightgbm	45.98	28.87	20.37	16.88	14.73	13.32	13.09	13.35	13.92	14.6	15.6	16.62
Only_IB	mid_lr	45.54	27.44	18.86	15.71	13.76	12.38	12.33	12.7	13.35	14.57	15.9	16.91
Only_IB	mid_et	44.46	27.06	19.16	15.86	13.69	12.34	12.17	12.38	12.84	13.89	15.11	15.89
Only_IB	mid_huber	45.52	27.71	19.03	15.71	13.58	12.08	11.99	12.38	13.06	14.32	15.7	16.73
Only_IB	mid_rf	43.73	26.58	18.89	15.12	13.07	11.74	11.75	12.07	12.6	13.53	14.72	15.75
Only_IB	mid_tr	45.26	26.89	18.63	15.05	12.8	11.65	11.92	11.91	12.8	13.43	14.72	16.21
		IB T--5	IB T--10	IB T--15	IB T--20	IB T--25	IB T--30	IB T--35	IB T--40	IB T--45	IB T--50	IB T--55	IB T--60

19. Ábra: Pontbecslések átlagos Skill score eredményei egyszeri tanítással.



20. Ábra: A legnagyobb okozott keresztszabályozású napon belül kiegyenlítetttség, napenergia előrejelzési hiba és javasolt mFRR aktiválás.

6. Táblázat: Napi mFRR statisztikák GINI-based_exh esetben (min.: 100 MW, konfidenciaszint: 90%).

	mFRR [MWh]	Error [MWh]	Error/ mFRR [%]	mFRR [MWh] +	mFRR [MWh] -	Sum IB [MWh]	Sum IB [MWh] +	Sum IB [MWh] -		mFRR [MWh]	Error [MWh]	Error/ mFRR [%]	mFRR [MWh] +	mFRR [MWh] -	Sum IB [MWh]	Sum IB [MWh] +	Sum IB [MWh] -
2021-05-07	1560.0	0.4	0.02%	1560.0	0.0	4332.3	4150.4	-182.0	2021-04-30	212.5	0.8	0.39%	0.0	-212.5	2224.5	107.1	-2117.4
2021-04-13	1137.5	1.4	0.12%	1112.5	-25.0	3414.4	2988.6	-425.8	2021-08-16	207.1	0.0	0.00%	207.1	0.0	2082.2	1660.4	-421.7
2021-06-30	1070.0	3.0	0.28%	0.0	-1070.0	3931.1	161.6	-3769.5	2021-07-16	204.6	10.0	4.91%	84.6	-120.0	2253.9	1033.2	-1220.8
2021-08-05	972.5	21.4	2.20%	632.5	-340.0	4272.0	2877.7	-1394.3	2021-06-07	202.5	0.4	0.18%	202.5	0.0	2231.0	1897.6	-333.4
2021-04-07	950.0	1.9	0.20%	950.0	0.0	4315.1	4242.5	-72.6	2021-05-17	202.5	3.3	1.65%	180.0	-22.5	1847.5	1072.8	-774.6
2021-07-19	930.0	3.1	0.34%	547.5	-382.5	3718.9	1710.5	-2008.5	2021-09-26	202.5	0.0	0.00%	202.5	0.0	2535.9	2505.0	-30.9
2021-04-08	907.5	7.9	0.87%	907.5	0.0	4023.0	3941.0	-82.0	2021-06-05	192.5	0.7	0.36%	132.5	-60.0	1940.5	1110.6	-829.9
2021-07-01	895.0	1.5	0.16%	0.0	-895.0	3334.1	100.9	-3233.2	2021-08-24	192.5	5.8	3.00%	90.0	-102.5	1841.6	1122.7	-718.9
2021-08-07	781.7	0.0	0.00%	0.0	-781.7	2985.0	297.1	-2687.9	2021-05-20	187.5	0.3	0.14%	187.5	0.0	1874.0	1537.7	-336.3
2021-05-22	780.0	0.6	0.08%	780.0	0.0	3003.8	2797.4	-206.4	2021-07-26	185.0	0.5	0.25%	185.0	0.0	1888.6	1754.2	-134.5
2021-08-17	777.9	2.3	0.30%	290.4	-487.5	3240.7	1369.1	-1871.6	2021-08-06	180.8	1.7	0.92%	0.0	-180.8	2109.7	228.0	-1881.6
2021-08-01	730.0	6.1	0.83%	517.5	-212.5	3573.0	2002.9	-1570.1	2021-04-20	180.0	0.0	0.00%	0.0	-180.0	1791.6	119.6	-1672.0
2021-07-09	698.8	25.4	3.63%	698.8	0.0	2942.1	2462.0	-480.1	2021-04-14	177.5	4.1	2.33%	22.5	-155.0	1840.0	896.5	-943.5
2021-08-18	672.5	0.0	0.00%	0.0	-672.5	2570.2	188.4	-2381.8	2021-05-05	170.0	6.7	3.92%	82.5	-87.5	2149.6	921.5	-1228.1
2021-08-25	652.5	0.0	0.00%	0.0	-652.5	2505.3	353.0	-2152.3	2021-04-22	170.0	0.4	0.23%	165.0	-5.0	1779.7	942.5	-837.3
2021-05-27	632.5	3.9	0.62%	0.0	-632.5	2454.3	76.4	-2378.0	2021-09-29	170.0	0.0	0.00%	170.0	0.0	1853.4	1631.4	-222.0
2021-09-01	630.0	11.0	1.75%	35.0	-595.0	2594.1	868.7	-1725.4	2021-08-11	170.0	0.0	0.00%	170.0	0.0	2406.3	2268.5	-137.8
2021-05-19	622.5	2.4	0.38%	210.0	-412.5	2591.3	1515.7	-1075.6	2021-08-30	167.5	5.3	3.18%	20.0	-147.5	1820.0	755.9	-1064.0
2021-07-29	620.0	0.6	0.09%	0.0	-620.0	2861.8	93.7	-2768.1	2021-09-02	167.5	5.0	2.96%	95.0	-72.5	1926.7	1079.6	-847.1
2021-07-31	605.0	5.8	0.96%	605.0	0.0	2851.9	2443.6	-408.4	2021-09-21	167.5	2.4	1.41%	27.5	-140.0	2042.1	1024.3	-1017.8
2021-07-20	580.0	4.2	0.72%	0.0	-580.0	3034.6	17.8	-3016.7	2021-07-13	162.5	0.0	0.00%	162.5	0.0	2221.6	2155.5	-66.2
2021-07-06	577.5	0.0	0.00%	577.5	0.0	3170.9	2991.8	-179.2	2021-07-03	160.0	9.9	6.21%	0.0	-160.0	1455.0	387.7	-1067.3
2021-06-25	570.0	2.7	0.48%	115.0	-455.0	3227.1	630.7	-2596.3	2021-07-18	160.0	0.1	0.06%	160.0	0.0	1879.2	1103.0	-776.2
2021-09-05	552.5	4.4	0.79%	0.0	-552.5	2097.2	243.2	-1854.1	2021-04-18	160.0	0.0	0.00%	0.0	-160.0	1605.9	492.2	-1113.7
2021-06-22	542.5	1.1	0.21%	542.5	0.0	3345.0	3334.1	-11.0	2021-08-26	157.5	5.0	3.16%	0.0	-157.5	1567.9	319.8	-1248.1
2021-05-02	538.3	13.0	2.42%	65.0	-473.3	2785.6	458.9	-2326.7	2021-08-08	157.5	0.0	0.00%	2.5	-155.0	2007.9	669.8	-1338.1
2021-09-18	532.5	0.0	0.00%	532.5	0.0	2653.5	2336.0	-317.5	2021-07-11	155.0	1.7	1.08%	155.0	0.0	1940.5	1528.5	-412.0
2021-04-12	525.0	0.0	0.00%	515.0	-10.0	2916.1	2062.7	-853.4	2021-05-01	154.2	12.5	8.10%	10.0	-144.2	2072.9	386.4	-1686.5
2021-06-23	495.0	0.0	0.00%	495.0	0.0	3616.1	3606.9	-9.1	2021-06-01	147.5	0.0	0.00%	97.5	-50.0	1738.5	940.2	-798.3
2021-08-02	492.5	5.5	1.11%	15.0	-477.5	2574.7	455.3	-2119.3	2021-07-12	145.0	0.0	0.00%	117.5	-27.5	2010.1	1622.2	-388.0
2021-08-29	492.5	13.3	2.70%	172.5	-320.0	2334.5	801.4	-1533.1	2021-09-19	145.0	0.0	0.00%	65.0	-80.0	1690.9	660.2	-1030.7
2021-04-16	490.0	14.4	2.93%	432.5	-57.5	2419.2	1777.4	-641.8	2021-08-13	135.0	0.9	0.65%	0.0	-135.0	1633.2	471.5	-1161.7
2021-04-29	487.5	13.4	2.75%	487.5	0.0	2699.5	2131.5	-567.9	2021-06-20	135.0	0.0	0.00%	135.0	0.0	2222.0	1824.3	-397.7
2021-05-30	485.0	1.7	0.35%	357.5	-127.5	2425.7	1420.9	-1004.8	2021-07-25	130.0	0.0	0.00%	67.5	-62.5	2057.2	1490.1	-567.2
2021-06-21	485.0	0.0	0.00%	485.0	0.0	2991.7	2874.5	-117.2	2021-05-18	125.0	6.7	5.33%	42.5	-82.5	1969.9	1057.8	-912.1
2021-04-28	482.5	12.0	2.48%	395.0	-87.5	2227.5	1244.5	-983.0	2021-08-04	117.5	0.5	0.45%	37.5	-80.0	1942.8	1217.3	-725.5
2021-05-03	465.0	2.8	0.60%	15.0	-450.0	2510.8	362.0	-2148.7	2021-10-04	117.5	0.0	0.00%	42.5	-75.0	2379.0	1426.1	-952.8
2021-07-10	455.0	17.3	3.80%	222.5	-232.5	2633.9	1441.4	-1192.5	2021-09-28	117.5	0.0	0.00%	117.5	0.0	1854.4	1330.2	-524.3
2021-05-24	447.5	0.5	0.10%	437.5	-10.0	2448.1	1784.9	-663.2	2021-06-06	112.5	3.3	2.97%	0.0	-112.5	1532.1	672.9	-859.2
2021-07-05	442.5	32.1	7.25%	287.5	-155.0	2853.4	2164.1	-689.2	2021-09-17	112.5	0.0	0.00%	112.5	0.0	1668.0	1168.3	-499.7
2021-05-25	437.5	0.0	0.00%	430.0	-7.5	2344.4	1725.0	-619.4	2021-05-31	107.5	0.0	0.00%	70.0	-32.5	1847.9	1140.6	-707.3
2021-05-11	432.5	4.3	0.99%	380.0	-52.5	2569.2	1799.0	-770.2	2021-05-16	102.5	0.0	0.00%	60.0	-42.5	2172.9	464.1	-1708.8
2021-08-19	432.5	4.2	0.97%	40.0	-392.5	2480.0	258.4	-2221.5	2021-08-10	102.5	0.0	0.00%	60.0	-42.5	1791.8	1069.0	-722.8
2021-06-18	427.5	0.0	0.01%	427.5	0.0	2225.5	1757.9	-467.6	2021-04-19	100.0	6.3	6.28%	47.5	-52.5	1866.9	824.6	-1042.4
2021-06-26	427.5	4.4	1.03%	0.0	-427.5	2475.8	36.5	-2439.3	2021-06-28	100.0	0.0	0.00%	0.0	-100.0	1617.8	571.5	-1046.4
2021-07-17	425.4	0.0	0.00%	425.4	0.0	2926.6	2716.2	-210.4	2021-08-20	95.0	0.0	0.00%	77.5	-17.5	1662.5	1152.7	-509.8
2021-04-10	425.0	0.8	0.18%	2.5	-422.5	2379.4	477.2	-1902.2	2021-10-03	92.5	0.0	0.00%	92.5	0.0	2304.6	1870.1	-434.5
2021-06-13	402.9	4.9	1.21%	115.4	-287.5	2331.0	670.0	-1661.0	2021-06-12	90.0	0.0	0.00%	90.0	0.0	2200.0	2145.8	-54.2
2021-04-11	402.5	0.0	0.00%	402.5	0.0	2139.6	1586.7	-552.9	2021-07-22	87.5	2.5	2.90%	0.0	-87.5	1584.3	609.1	-975.2
2021-08-27	392.5	0.0	0.00%	362.5	-30.0	2314.6	1429.8	-884.8	2021-07-30	80.0	0.0	0.00%	0.0	-80.0	1757.6	368.3	-1389.3
2021-09-22	390.0	3.3	0.85%	390.0	0.0	2710.3	2656.3	-54.0	2021-08-31	75.0	2.0	2.67%	57.5	-17.5	1913.4	1260.4	-653.0
2021-09-04	390.0	0.0	0.01%	0.0	-390.0	2052.5	159.0	-1893.5	2021-06-04	72.5	0.0	0.00%	72.5	0.0	1931.1	1517.4	-413.7
2021-07-14	385.0	0.6	0.16%	370.0	-15.0	2539.5	1753.1	-786.4	2021-06-08	70.0	0.0	0.00%	70.0	0.0	1790.1	1524.1	-266.0
2021-05-06	375.0	1.3	0.36%	365.0	-10.0	2020.8	1582.8	-438.0	2021-06-09	70.0	0.0	0.00%	70.0	0.0	1682.7	1426.6	-256.1
2021-04-21	370.0	10.3	2.79%	117.5	-252.5	2400.6	985.1	-1415.5	2021-05-08	67.5	5.4	7.99%	67.5	0.0	1767.4	1446.5	-320.9
2021-04-26	360.0	2.3	0.64%	262.5	-97.5	2233.3	1254.2	-979.1	2021-05-26	65.0	1.1	1.62%	0.0	-65.0	1545.6	360.6	-1185.0
2021-04-17	352.5	4.3	1.22%	352.5	0.0	1670.2	1315.6	-354.7	2021-08-12	65.0	0.0	0.00%	65.0	0.0	2029.2	1818.5	-210.7
2021-07-08	333.8	0.5	0.14%	333.8	0.0	3252.6	3102.9	-149.7	2021-05-12	62.5	17.5	28.00%	42.5	-20.0	1605.3	633.8	-971.5
2021-09-07	332.5	11.1	3.34%	300.0	-32.5	3226.4	2688.6	-537.8	2021-06-17	62.5	0.0	0.00%	62.5	0.0	1723.6	1622.7	-100.9
2021-09-23	320.0	5.0	1.56%	75.0	-245.0	2606.2	1425.4	-1180.8	2021-07-23	57.5	0.5	0.84%	52.5	-5.0	1321.3	854.3	-467.0
2021-09-27	317.5	0.0	0.00%	317.5	0.0	2439.3	2295.2	-144.1	2021-08-14	57.5	0.0	0.00%	57.5	0.0	2125.5	2035.6	-90.0
2021-09-15	315.0	1.7	0.53%	315.0	0.0	2797.5	2737.2	-60.3	2021-08-21	55.0	0.3	0.62%	0.0	-55.0	1455.6	418.3	-1037.3
2021-06-24	312.5	0.4	0.14%	312.5	0.0	2781.3	2697.4	-83.8	2021-08-03	55.0	1.7	3.03%	47.5	-7.5	1839.5	953.2	-886.3
2021-07-21	307.5	8.3	2.71%	0.0	-307.5	2315.2	125.2	-2190.0	2021-07-15	55.0							