



**Budapesti Műszaki és Gazdaságtudományi Egyetem**  
Villamosmérnöki és Informatikai Kar  
Villamos Energetika Tanszék

Turóczy Balázs Ágoston

# **Napelemes termelés időjárásalapú csoportosítása**

Tudományos Diákköri Konferencia dolgozat

KONZULENS

**Dr. Hartmann Bálint**

BUDAPEST, 2019

## Összefoglaló

---

A villamos energetika egyik fő fejlődési iránya napjainkban a megújuló energiaforrásokat hasznosító erőművek fejlesztése és telepítése. Ennek meghatározó oka a klímaváltozás folyamata, melyet más tényezők mellett, az üvegház-hatású gázok nagymértékű kibocsájtása okozhat. Ezen jelenség megfékezése érdekében több nemzetközi megállapodás is született. Ezek egyike bevezette az úgynevezett CO<sub>2</sub> kvóták rendszerét, mely jelentős gazdasági ösztönző, így a villamosenergia-termelésben nagymértékű dekarbonizációs folyamatok indultak meg.

Magyarországon is megfigyelhető az említett tendencia. Hazánkban az elkövetkezendő évtizedben nagy mennyiségű napenergiát hasznosító erőmű létesítését tervezik, összesen néhány ezer MW teljesítmény beépítése van kilátásban. A jelenlegi magyar villamosenergia-rendszerben ez számottevő teljesítményhányadot jelent. A napenergia azonban rendkívül volatilis energiaforrás, tehát az időjárástól függően egy percen belül is jelentősen változhat a napelemparkok által a hálózatba táplált villamosenergia mennyisége. Mindez igen nehéz feladat elé állítja a hálózati rendszerirányítókat, a napi menetrendek megtervezése jelentős nehézségekbe ütközik. Ennek okán nagyon fontos, hogy képesek legyünk egy jól használható, a valóságot elfogadható hibával közelítő becslést adni a várható, időjárásváltozásból fakadó termelésingadozásokra. A tervezéshez felhasználhatjuk az időjárás-előrejelzést vagy támaszkodhatunk historikus adatokra, melyekből tendenciaszerű becslést adhatunk. Ez utóbbihoz múltbeli termelési adatokon kívül szükségünk van egy csoportosítási módszerre is, melyek az év napjait, így az adott napon lehetséges napelemes termelést csoportosítani képes az időjárás alapján. Számos ilyen módszert fejlesztettek ki az elmúlt évtizedben, különböző országokban, különböző megfontolások alapján.

Dolgozatomban a fent említett napelemes termelési adatok időjárás alapú csoportosítására kidolgozott módszerek helyességét és adaptálhatóságát vizsgáltam meg. Munkám során 2018-as évi magyarországi adatokkal dolgoztam, melyeket a MTA Energiatudományi Kutatóközpont (EK) és a Debreceni Egyetem kihelyezett tanszéke, a Megújuló Energiapark szolgáltatott. Feladatul tűztem ki, hogy megállapítsam: vajon megfelelően használhatóak ezek a módszerek magyarországi körülmények között is? Vizsgáltam az egyes módszerek robusztusságát, valamint eredményeik korrelációját, így értékelve hatékonyságukat. Hosszútávú célom a vizsgált módszerek segítségével egy olyan osztályozási rendszer létrehozása, mely mind meteorológiai, mind villamosenergia termelési szempontból jól alkalmazható specifikusan Magyarország területén, ezzel elősegítve a hálózati rendszerirányítók jövőbeli munkáját mind az átviteli mind az elosztóhálózat szintjén.

## Abstract

---

One of the major trends in the electrical power industry today is the development and installation of power plants using renewable energy sources. The main reason for this is the climate change process, which can be caused, among other factors, by the high emission of greenhouse gases. Several international agreements have been concluded to manage this phenomenon. One of these has introduced the so-called CO<sub>2</sub> quota system, which is a major economic incentive, leading to major decarbonization processes in electric power generation.

This tendency can also be observed in Hungary. It is planned to build many, and high capacity solar power plant in Hungary in the next decade, with a total capacity of 10,000 MW. In the current Hungarian electricity system, this represents a very significant ratio in the power generation. However, solar energy is a very volatile source of energy, so depending on the weather, the amount of electric power generation supplied to the grid by solar parks can change significantly within a very short term. All of this challenge the DSO's at grid and system operation, and there are significant difficulties in predicting daily schedules. For this reason, it is very important to be able to provide a well-used estimation for the production fluctuations based on weather changes within acceptable deviation range. We can use weather forecasting or based on historical data to provide trend-based estimation. For the second solution not only historical production data, but also a grouping method is necessary, which can classify the days of the year, defined on the possible weather based solar production of the days. Many of these methods have been developed over the last decade, in different countries, based on different considerations.

In my work I have investigated the correctness and adaptability of the above-mentioned methods for clustering solar generation power data based on weather conditions. During my work I have used Hungarian data for 2018, provided by the MTA Energiatudományi Kutatóközpont (EK) and the "Renewable Energy Park" Department of the University of Debrecen. My main task was defined for establishing: Are these methods suitable to use in Hungarian conditions? I have examined the robustness of each method and the correlation of their concerning results, to evaluate their effectiveness. My long-term goal is to create a specific classification system for Hungary, both in terms of meteorology and electricity production point of view, which could facilitate and support the future daily work of system operators both at transmission and distribution network level.

# Tartalomjegyzék

---

Összefoglaló .....	i
Abstract .....	ii
Tartalomjegyzék .....	iii
1 Tiszta ég modellek .....	1
1.1 Bird-féle tiszta ég modell [1] .....	1
2 Csoportosítási módszerek .....	4
2.1 Trueblood [2] .....	4
2.1.1 Tiszta égbolt index ( $CI_{day}$ ) .....	4
2.1.2 Variabilitási index ( $VI_{day}$ ) .....	4
2.1.3 Csoportok kialakítása .....	5
2.2 Maafi [3] .....	5
2.2.1 Tiszta égbolt index ( $K(d)$ ) .....	5
2.2.2 Fraktál index ( $D(d)$ ) .....	6
2.2.3 Csoportok kialakítása .....	6
2.3 Fortuna [5] .....	7
2.3.1 Területarány index ( $A_r(d)$ ) .....	7
2.3.2 Változékonysági index ( $I(d)$ ) .....	7
2.3.3 Csoportok kialakítása .....	8
2.4 Chicco [7] .....	9
2.4.1 Chicco normált, idősoros (Chicco 1) .....	9
2.4.2 Chicco normált, rendezett (Chicco 2) .....	9
2.4.3 Csoportok kialakítása .....	9
2.5 Almohammadi [8] .....	10
2.5.1 Csoportok kialakítása .....	10
2.6 Módszerek összefoglalása .....	11
3 Csoportosítás megvalósítása .....	12



3.1	Energiatudományi Kutatóközpont (EK).....	12
3.1.1	Napelempark [9].....	12
3.1.2	Adatfeldolgozás.....	14
3.1.3	Csoportosítás.....	15
3.2	Debreceni Megújuló Energiapark.....	27
3.2.1	Adatfeldolgozás.....	28
3.2.2	Csoportosítás.....	28
3.3	Eredmények összehasonlítása.....	32
4	Módszerek értékelése és összehasonlítása.....	33
4.1	Csoportosítási egyezések elemzése.....	33
4.1.1	Tiszta égboltú napok.....	33
4.1.2	Felhős napok.....	40
4.1.3	Ellentmondásos besorolás.....	46
4.1.4	Felhőborítottság.....	50
4.2	Módszerek felbontásérzékenysége.....	53
4.3	Értékelés.....	54
4.3.1	Trueblood.....	54
4.3.2	Maafi.....	54
4.3.3	Fortuna.....	55
4.3.4	Chicco.....	55
4.3.5	Almohammadi.....	55
5	Konklúzió, javaslattétel.....	56
6	Irodalomjegyzék.....	1

# 1 Tiszta ég modellek

Tiszta ég modellek segítségével becsülhető egy adott földrajzi helyen (tiszta égboltot feltételezve), adott pillanatban a felszínt érő globál (vízszintes síkra a felette levő  $2\pi$  térszögből belépő), szórt/diffúz (a vízszintes síkra a felette levő  $2\pi$  térszögből érkező, kivéve a Nap korongjának térszögét) és direkt (a Nap korongjának térszögéből a Nap irányára merőleges felületre belépő) napsugárzás mértéke. Számos ilyen módszer létezik, melyek összetettséjüknek megfelelően, különböző légköri folyamatokat vesznek figyelembe a lehető legpontosabb becslés érdekében. Sok esetben nincs lehetőség minden adat mérésére az adott helyszínen, így azokat modellek alapján kell megbecsülni. Munkám során én a Bird-féle tiszta ég modellt alkalmaztam.

## 1.1 Bird-féle tiszta ég modell [1]

A Bird-féle tiszta ég modellt Richard E. Bird és Roland L. Hulstrom fejlesztette ki 1981-ben [1]. Ez a módszer az összetett tiszta ég modellek csoportjába tartozik, mivel számos légköri paramétert vesz figyelembe a besugárzási értékek számításakor, így bemenő adatainak száma is szükségszerűen nagy.

A Bird-féle tiszta ég modell bemenő adatai:

1. Napállandó (Solar constant): Direkt sugárzás átlagos Nap-Föld távolságnál a légkörön kívül.
2. Zenit szög (zenith angle): A függőleges egyenes és a nap által bezárt geometriai szög, mely a földrajzi hely függvényében az év minden napjának minden percére meghatározható.
3. Felszíni nyomás (surface pressure): A légnyomás az adott földrajzi területen.
4. Albedó (ground albedo): A földfelszín fényvisszaverő képessége.
5. Páratartalom (precipitable water vapor): A levegő páratartalma adott földrajzi területen.
6. Ózonréteg (total ozone): Az ózonréteg vastagsága az adott földrajzi helyen.
7. Zavarosság (turbidity at 0.5- and/or 0.38  $\mu\text{m}$  wavelength): Az aeroszol optikai mélysége függőleges vetületen (0,5- és/vagy 0,38- $\mu\text{m}$  hullámhosszon).
8. Aeroszol részecskék okozta előre szórási arány (aerosol forward scattering ratio): A teljes szórt besugárzás és az aeroszol részecskék okozta előre szórási aránya (a szakirodalomban: 0,84).

Ezen bemenő paramétereknek megfelelően a modell alapegyenletei a következők.

A direkt besugárzás meghatározására szolgáló egyenlet:

$$I_d = I_0 \cdot \cos(Z) \cdot 0,9662 \cdot T_R \cdot T_O \cdot T_{UM} \cdot T_W \cdot T_A \quad (1)$$

A diffúz besugárzás meghatározására szolgáló egyenlet:

$$I_{as} = \frac{I_0 \cdot \cos(Z) \cdot 0,79 \cdot T_O \cdot T_W \cdot T_{UM} \cdot T_{AA} \cdot [0,5 \cdot (1 - T_R) + B_a \cdot (1 - T_{AS})]}{1 - M + M^{1,02}} \quad (2)$$

A teljes, vízszintes felületre eső besugárzás meghatározására szolgáló egyenlet:

$$I_T = \frac{I_a + I_{as}}{1 - r_g r_s} \quad (3)$$

ahol

$I_0$	az atmoszférát érő besugárzás (az év során változik, a Nap-Föld távolság változása miatt)
$Z$	a zenit szög (bemeneti paraméter)
$T_R$	a Rayleigh-szóródás fényáteresztése
$T_O$	az ózonréteg fényáteresztése
$T_{UM}$	az egyenletesen kevert gázok (szén-dioxid és oxigén) fényáteresztése
$T_W$	a vízgőz fényáteresztése
$T_A$	az aeroszol abszorpciójának és szóródásának fényáteresztése
$T_{AA}$	az aeroszol fényáteresztése
$B_a$	az aeroszol részecskék okozta előre szórási arány (bemeneti paraméter)
$T_{AS}$	az aeroszol szóródásának fényáteresztése
$M$	a légtömeg
$r_g$	a környezet albedója (bemeneti paraméter)
$r_s$	az atmoszféra albedója

Látható, hogy az egyenletekben a bemeneti paramétereken kívül szerepel kilenc darab átviteli tényező is, melyek meghatározására szolgáló pontos egyenleteket is tartalmazza az [1].

$$T_R = e^{[-0.903 \cdot (M')^{0.84} \cdot (1 + M' - (M')^{1.01})]} \quad (4)$$

$$T_0 = 1 - 0,1611 \cdot X_0 \cdot (1 + 139,48 \cdot X_0)^{-0,3035} - 0,002715 \cdot X_0 \cdot (1 + 0,044 \cdot X_0 + 0,0003 \cdot X_0^2)^{-1} \quad (5)$$

$$X_0 = U_0 \cdot M \quad (6)$$

$$T_{UM} = e^{[-0,0127 \cdot (M')^{0,26}]} \quad (7)$$

$$T_W = 1 - 204959 \cdot X_W \cdot [(1 + 79,034 \cdot X_W)^{0,6828} + 6,385 \cdot X_W]^{-1} \quad (8)$$

$$X_W = U_W \cdot M \quad (9)$$

$$T_A = e^{[-\tau_A^{0,873} \cdot (1 + \tau_A - \tau_A^{0,7088}) \cdot M^{0,9108}]} \quad (10)$$

$$\tau_A = 0,2758 \cdot \tau_{A,0.38} + 0,35 \cdot \tau_{A,0.5} \quad (11)$$

$$T_{AA} = 1 - K_1 \cdot (1 - M + M^{1,06}) \cdot (1 - \tau_A) \quad (12)$$

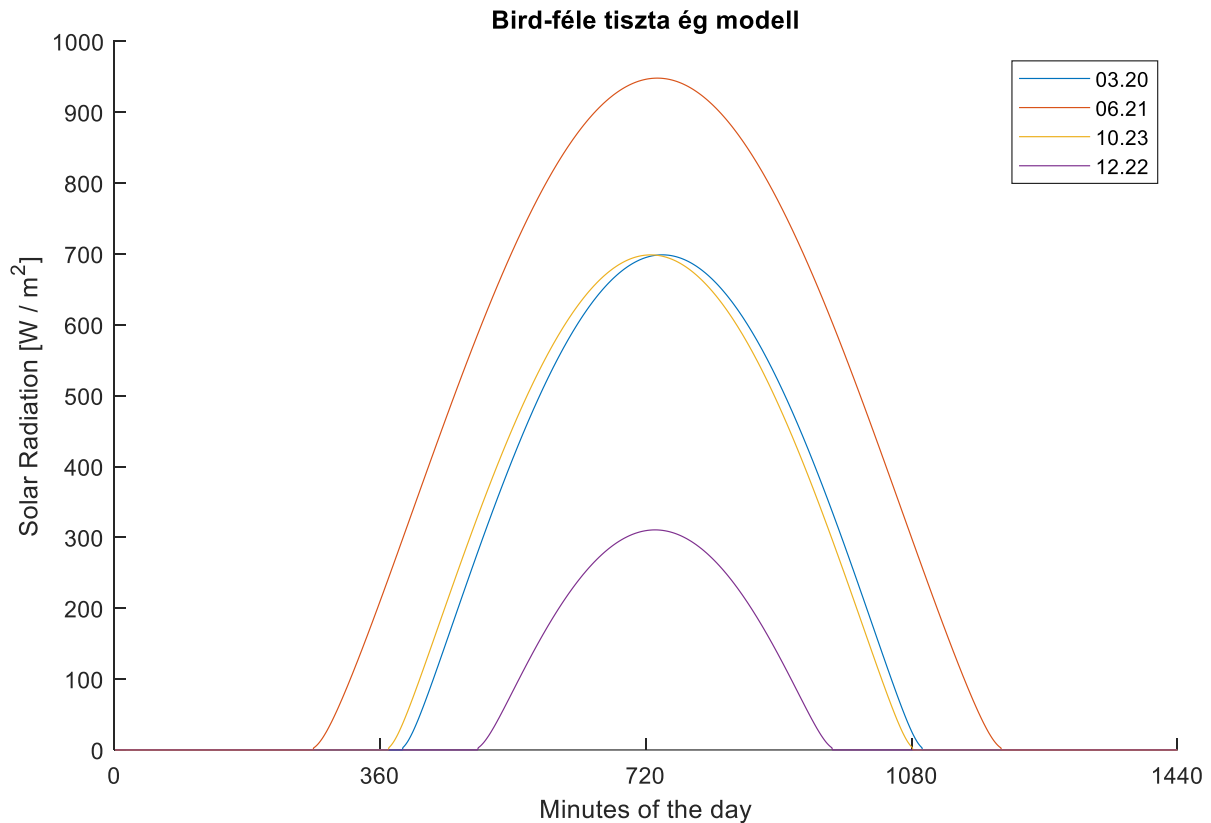
$$T_{AS} = \frac{T_A}{T_{AA}} \quad (13)$$

$$r_s = 0,0685 + (1 - B_a) \cdot (1 - T_{AS}) \quad (14)$$

$$M = [\cos(Z) + 0,15 \cdot (930885 - Z)^{1,25}]^{-1} \quad (15)$$

$$M' = \frac{M}{1013} \quad (16)$$

Ezen összefüggések alapján adott földrajzi helyre, az év minden napjára kiszámítható az ideális besugárzási görbe, a kívánt felbontással. Munkám során egy-, öt- és tízperces felbontású görbéket alkalmaztam, melyeket a Energiatudományi Kutatóközpont (EK) területére számítottam ki (1. ábra).



**1. ábra: Bird-féle tiszta ég modellből számított egyperces eredmények**

A fenti ábrán láthatók a Bird-féle tiszta ég modell által számított, egyperces felbontású besugárzási görbék a tavaszi- és őszi nap-éj egyenlőség, illetve a nyári- és téli napforduló napjaira, a fent említett földrajzi helyre.

Amint az az ábrán (1. ábra) is látható, ez a tiszta ég modell feltételezi, hogy az év minden napján azonos időpontban éri el a Nap pályájának csúcspontját (a besugárzási görbék csúcspontja egy, függőleges egyenesbe esne, ha az óráátállítást nem venné figyelembe a modell). Mivel ez a valóságban nem így van, a későbbiekben a mért besugárzási adataink korrigálásra szorulnak majd. Mivel mérési adataink nem veszik figyelembe a nyári időszámítást, ezért a tiszta ég modell besugárzási adatait is helyesbítenünk kell majd.

Munkám során a tiszta ég modellből származó besugárzási adatokat referencia adatként használtam fel az egyes napok mért besugárzási adatai alapján történő csoportosításához. Az elmúlt évtizedekben több ilyen csoportosítási módszert is kidolgoztak, melyek közül én az alábbiakat vizsgáltam meg.

## 2 Csoportosítási módszerek

Számos módszer létezik az év napjainak csoportosítására, mely csoportosítások az adott napokra vonatkozó napsugárzási görbék különböző tulajdonságait veszik alapul. Ezen módszerek számottevő része különböző tiszta égbolt és/vagy valamilyen variabilitási indexeket használ az egyes csoportok elkülönítésére. Munkám során öt, egymástól igen eltérő módszert vizsgáltam meg. Ezek elméletének ismertetése olvasható ebben a fejezetben. (Az egyes módszereket az átláthatóság kedvéért elsőszámú szerzőik alapján különböztetem meg.)

### 2.1 Trueblood [2]

Ezt a módszert az USA-ban fejlesztették ki 2013-ban egy Tennessee-ben találtató 1 MW-os napelempark egy éves, 1 másodperces felbontású mérési adatai alapján. Az eljárás kidolgozása során a kutatók eleinte ugyan másodperces felbontású adatsorral dolgoztak, azonban munkájuk során arra jutottak, hogy nem szükséges ilyen nagyfelbontású adat, 1 perces mintavételi idejű adatsorral is megfelelő pontosságú eredmények érhetők el. A tanulmány alapfeltevése, hogy a besugárzási görbék hossza, illetve a görbe alatti terület alapján a napok időjárása megállapítható. Így a napok csoportosítása elvégezhető két indexszám, a napi tisztaégbolt - és a variabilitási index alapján.

#### 2.1.1 Tiszta égbolt index ( $CI_{day}$ )

A napi tiszta égbolt index (Daily Clearness Index –  $CI_{day}$ ) az adott nap két (mért besugárzási és clear sky besugárzási) görbéje alatti terület arányával fejezhető ki [2]. Ennek számítási módja a következő.

$$CI_{day} = \frac{\text{mért besugárzási görbe alatti területe}}{\text{clear sky besugárzási görbe alatti területe}} \quad (17)$$

A  $CI_{day}$  index értékei a tanulmányban 0,0 és 1,1 között változnak. A gyakorlatban is előfordulhat 1 feletti érték (tehát a valós besugárzás mértéke meghaladja az elvi maximumot), mivel a tiszta égbolt modellek nem minden esetben adnak pontos eredményt a matematikai modell sajátosságai miatt.

#### 2.1.2 Variabilitási index ( $VI_{day}$ )

A [2] forrás az időjárás változékonyságát a tiszta égboltot feltételező besugárzásigörbe és a mért besugárzásigörbe hosszának összehasonlításával kívánja megállapítani. Az év napjainak összehasonlíthatósága érdekében a számítások elvégzéséhez olyan adatsorokat alkalmaztam, az eredeti módszertől eltérően, melyek csak a napsütéses órák mérési adatpontjait tartalmazzák, valamint az évi minden napjára azonos hosszúságúak. Ezt az eredeti adatsor újramintavételezésével valósítottam meg.

Az így számított variabilitási indexet (Daily Variability Index –  $VI_{day}$ ) meghatározó egyenlet a következő.

$$VI_{day} = \frac{\text{mért besugárzási görbe hossza}}{\text{clear sky besugárzási görbe hossza}} \quad (18)$$

Az így számított variabilitási index értékei a tanulmány szerint jellemzően 1 és 30 közötti számok.

### 2.1.3 Csoportok kialakítása

A tanulmány készítői öt jellegzetes besugárzási görbetípust különböztettek meg, melyek indexeit kiszámítva kialakították a csoporthatárokat. Felhasználva a  $CI_{day}$  és  $VI_{day}$  indexeket, a [2] forrás determinisztikusan a következő csoportokat állapítja meg:

1. táblázat: Trueblood módszer csoportjai

Csoport elnevezése	Feltétel
Derült nap (Clear)	$VI_{day} < 2$ ÉS $CI_{day} \geq 0,5$
Nagyon változékony nap (High)	$VI_{day} \geq 10$
Mérsékelten változékony nap (Moderate)	$5 \leq VI_{day} < 10$
Gyenge változékonyságot mutató nap (Mild)	$2 \leq VI_{day} < 5$
Borult nap (Overcast)	$VI_{day} < 2$ ÉS $CI_{day} < 0,5$

A fenti módszer segítségével igen egyszerűen csoportosíthatjuk a napokat, azonban éppen egyszerűségéből fakadóan a hibásan besorolt napok száma is magas lehet.

## 2.2 Maafi [3]

Ez a csoportosítási módszer egy Algériában készült tanulmányban jelent meg 2003-ban. A kutatók két sivatagban található település, Tahifet és Imehrou időjárásállomásának 10 perces felbontású, 1 évnyi napsugárzási adatai alapján dolgoztak. Tiszta égbolt és fraktál indexeket vezettek be az egyes csoportok elkülönítésére, melyek a következők.

### 2.2.1 Tiszta égbolt index ( $K(d)$ )

A tanulmány készítői egy, a 2.1.1 fejezetben ismertetett  $CI_{day}$  indexszel gyakorlatilag megegyező,  $K(d)$  tiszta égbolt indexet használtak. Ennek indoka, hogy a kutató munkájuk során tapasztalták, a fraktáldimenzió jól megmutatja, ha egy nap során nagyon változékony a besugárzás mértéke, azonban a szabályos (haraggörbe jellegű) besugárzási görbék között, melyek előfordulhatnak teljesen napos és teljesen felhős időjárás esetén is, nem tud különbséget tenni. Ezért a mutatószám fő feladata ebben az esetben a szabályos görbék egymástól való megkülönböztetése.

### 2.2.2 Fraktál index ( $D(d)$ )

Ennél a módszernél fraktál dimenzióval határozzák meg a besugárzási görbék szabálytalanságának mértékét, amit az időjárás változékonysága okoz. Szabályos görbének a haranggörbét tekinti a tanulmány, mely tiszta égboltú (és teljesen felhős) időjárás esetén fordul elő. Mivel a napi besugárzásigörbe egydimenziós, diszkrét idejű adatsor, a fraktáldimenzió számítható a Minkowski-Bouligand-dimenzió segítségével. Egy másik lehetséges módja a fraktáldimenzió számításnak, az úgynevezett dobozszámolás (box-count) módszer. Ennek során a besugárzási görbét ismert magasságú és szélességű téglalapokkal fedik le, majd kiszámolják azok összesített területét, melyből meghatározható a fraktáldimenzió. Az egyes téglalapok paramétereinek nem kell feltétlenül azonosnak lenniük, különböző méretű téglalapok használatával is végezhető az eljárás. A számítások elvégzéséhez olyan adatsorok szükségesek, melyek csak a napsütéses órák méréseit tartalmazzák, valamint az évi minden napjára azonos hosszúságúak. Ezt az eredeti adatsor újramintavételezésével lehet megvalósítani. A két számítási mód nagyon különböző, eredményeik mégis ekvivalensnek tekinthetők, ahogy az a [4] forrásban is bemutatásra kerül. Ezen módszerek segítségével minden napra megállapítható egy fraktálindex, melyek a [3] tanulmányban 1 és 1,6 közötti értékűeknek adódtak.

### 2.2.3 Csoportok kialakítása

A fent ismertetett két index segítségével már egyértelműen elkülöníthetők a különböző időjárású napok. Ezek felhasználásával az év napjait három csoportba sorolták be, melyek az alábbiak.

2. táblázat: Maafi módszer csoportjai

Csoport elnevezése	Feltétel	Jellemzés
Tiszta égboltú időjárás (Clear sky)	$1 < D(d) \leq 1,1$ ÉS $K(d) \geq 0,5$	Nagyon reguláris alakú besugárzási görbe, J alakú hisztogram
Részben felhős időjárás (Partly cloudy)	$1,1 < D(d) \leq 1,25$ ÉS $K(d) \geq 0,5$	Reguláris alakú besugárzási görbe kismértékű fluktuációval, U alakú hisztogram
Felhős időjárás (Cloudy)	$D(d) > 1,25$ VAGY $D(d) \leq 1,25$ ÉS $K(d) < 0,5$	Irreguláris alakú besugárzási görbe több fluktuációval, L alakú hisztogram

A box-count algoritmus eredményei nagy mértékben függenek a kiválasztott dobozmérettől, ezért ez a csoportosítási módszer nem nevezhető kifejezetten hibátúrónak. Nem megfelelő paraméterek (dobozszám) választása esetén a fraktáldimenzió számításához használt matematikai eljárások (box-count algoritmus) igen pontatlanok lehetnek.

## 2.3 Fortuna [5]

Ezt a csoportosítási módszert Olaszországban dolgozták ki 2016-ban. A kutatók egy teljes évnnyi (2000), 5 perces felbontású, amerikai (Aberdeen, Ohio) adattal dolgoztak. A tanulmány alapgondolata abból fakad, hogy a napenergia nagyon volatilis energiaforrás, így szélsőséges esetben előfordulhat, hogy a besugárzás nagyon rövid idő alatt nagymértékben megváltozik (trópusi éghajlaton akár 700 W/m<sup>2</sup>-es amplitúdóval is változhat rövid időn belül a földrajzi helytől függően [6]). Az ilyen nagy változásokat nehéz időtartományban érzékelni, frekvenciatartományban azonban jól detektálhatók, ezért az [5] tanulmány írói a napsugárzási görbék teljesítményspektrumát vizsgálták. A csoportosítás során azt is figyelembe vették az egyértelmű besorolás kialakítása érdekében, hogy az év során, az évszakok váltakozásának megfelelően, változik mind a besugárzási görbék hossza, mind azok burkológörbéjének magassága. Ennek kiküszöbölése érdekében, a napok összehasonlíthatóságának biztosításához úgynevezett jellemző nap (Typical Day - TD) modellt alkalmaztak. A jellemző nap modell historikus meteorológiai adatok alapján minden napra meghatároz egy jellemző időjárási profilt egy adott földrajzi helyre, mely referenciaként alkalmazható hasonló módon, mint a 2.1 és 2.2 fejezetben ismertetett módszerek esetében a tiszta ég modell.

A napok csoportosításához egy területarány (Area ratio –  $A_r$ ) és egy változékonyság (Intermittency –  $I$ ) indexet alkalmaztak, melyek segítségével négy csoportot különböztettek meg klaszterezési eljárással. Az említett indexek számítási módjai a következők.

### 2.3.1 Területarány index ( $A_r(d)$ )

A napok borultságát kifejező  $A_r(d)$  index a következőképp számítható:

$$A_r(d) = \frac{A_{pat}(d)}{A_{typ}(d)} \quad (19)$$

ahol

$A_{pat}(d)$  a napi mért besugárzási görbe alatti terület

$A_{typ}(d)$  a napi tiszta égboltú besugárzási görbe alatti terület

A (19) számítási mód alapján jól látszik az  $A_r(d)$  index és a  $Cl_{day}$  index közötti hasonlóság. A későbbiekben, mivel munkámhoz nem állt rendelkezésemre tipikus nap modell, a számítások elvégzéséhez a korábban már ismertetett  $Cl_{day}$  indexet használtam az eredeti módszerben leírt  $A_r(d)$  index helyett.

### 2.3.2 Változékonysági index ( $I(d)$ )

A változékonysági index a mért és számított besugárzási görbék spektrumának arányával fejezi ki az időjárás változékonyságát. A besugárzási adatok teljesítményspektrumának frekvenciatartománya és annak felosztása a mért adatok mintavételi frekvenciájától függ. Ötperces adatsor esetén a mintavételi



frekvencia 1/300 Hz értékű. A mért és számított görbék spektruma között a nagyfrekvenciás besugárzási teljesítményváltozásoknál van csak érdemi eltérés, hiszen a tiszta ég modellből számított görbében nincsenek hirtelen változások („haranggörbe”), míg a valós adatok igen változékonyak lehetnek. A [5] tanulmányban ezért csak a [1/30; 1/3] Hz tartományt vizsgálták. Munkám során azt tapasztaltam, hogy ezen tartomány változtatása nagymértékben befolyásolja a módszer eredményeit, így annak megválasztása alaposabb vizsgálatot igényel.

A napi időjárás változékonyságát kifejező  $I(d)$  index értéke következőképp számítható:

$$I(d) = \frac{\log(PW_{pat}(d)) - \log(PW_{typ}(d))}{\log(PW_{pat}(d)) + \log(PW_{typ}(d))} \quad (20)$$

ahol

$PW_{pat}(d)$  a napi mért besugárzás teljesítményspektrum-sűrűség görbe alatti terület adott frekvenciatartományban

$PW_{typ}(d)$  a tipikus nap modellel számított besugárzás teljesítményspektrum-sűrűség görbe alatti terület adott frekvenciatartományban

Megfigyelték, hogy az  $I(d)$  index az egy éves intervallumban 0,0 és 0,6 közötti értékeket vett fel. A görbék közötti kis eltérés esetén értelemszerűen  $I(d)$  0-hoz tart, míg, ha erős besugárzási változékonyság jellemzi az adott napot,  $I(d)$  0,6-hoz közelít. Mivel, ahogy már korábban említettem, nem állt rendelkezésemre tipikus nap modell, a változékonyság index számításához a Bird-féle tiszta ég modell eredményeit használtam fel.

### 2.3.3 Csoportok kialakítása

Az osztályozás jelen esetben két lépésben történik. Első lépésben kiszámításra kerülnek az  $A_r(d)$  és  $I(d)$  indexek, melyek alapján a napi besugárzási minták párokba rendezhetők. Ez után a második lépésben az összepárosított elemek k-közép klaszterezés segítségével kerülnek besorolásra a különböző csoportokba. A kialakított négy csoport jellemzői az alábbi táblázatban (3. táblázat) található.

3. táblázat: Fortuna módszer csoportjai

Csoport elnevezése	Jellemzés
Tiszta égboltú (Clear sky)	Tiszta égboltú nap, nagyon kevés felhővel
Enyhén felhős, nagy változékonyságú (Some clouds, high dynamic)	Felhősebb nap, magas szintű dinamikával
Enyhén felhős, közepes változékonyságú (Some clouds, medium dynamic)	Teljesen felhős nap, nagyméretű felhőkkel, melyek átvonulási sebessége kisebb, így a globális sugárzás
Felhős, alacsony változékonyságú (Cloudy, low dynamic)	Szignifikáns napsütésű nap sok kisebb felhővel, melyek átvonulási sebessége nagyobb, így a dinamika is nagyobb

Ez a csoportosítási módszer egészen más megközelítést alkalmaz a korábbiakhoz képest, mivel frekvenciatartományban vizsgálja a napsugárzási adatokat és nem determinisztikus módon, klaszterezési eljárással alakítja ki a csoportokat. Pontossága azonban nagyban függ többek között a fent említett szűkített frekvenciatartomány megválasztásától.

### 2.4 Chicco [7]

Ezt a módszert szintén Olaszországon dolgozták ki, 2014-ben. A tanulmány írói 1 perces felbontású, egy teljes évet (2012) lefedő adatsorral dolgoztak. A mérési adatok két, egymástól 70 km távolságra levő, hálózatra csatlakozó napelem parkhoz telepített, Dél-Olaszországi meteorológiai állomásból származtak.

A szerzők célja egy robusztus, hibatűrő módszer kidolgozása volt. Ennek érdekében igyekeztek olyan adatsort használni, mely a lehető leguniverzálisabb. Az eljárás közvetlenül a mért besugárzási adatok alapján végzi a csoportosítást, azonban minden nap adatsorát normálja az adott nap elméleti maximális értékével (tisztá ég modell). Ezzel gyakorlatilag függetleníti az algoritmust az évszakok váltakozásától. További egyszerűsítésként csak a napsütéses időintervallumokat vizsgálják, melyekből újramintavételezéssel azonos hosszúságú adatsorokat képeznek. Az eljárás a csoportok meghatározásához (k-közép) klaszterezést használ, melynek bemeneteként idősoros és rendezett idősoros adatok egyaránt szolgálnak.

#### 2.4.1 Chicco normált, idősoros (Chicco 1)

Az első esetben a módszer kidolgozói egyszerűen a normált idősoros adatokat használták a klaszterezési algoritmus bemeneteként. Így a napok csoportosítása gyakorlatilag a besugárzási görbék alakja alapján történik. Az algoritmus olyan csoportokat keres, melyek besugárzási görbéi hasonló jellegzetességeket mutatnak, így alakítja ki a négy különböző csoportot.

#### 2.4.2 Chicco normált, rendezett (Chicco 2)

Ebben az esetben minden nap besugárzási adatpontjait növekvő sorrendbe rendezték a besugárzás mértéke szerint, ezzel egyfajta sűrűségfüggvényt hozva létre. A normált és rendezett adatok használata a szerzők szerint az évszakok hatásának eliminálásán túl segít kiküszöbölni a napon belüli változások időbeliségét is, mely hibás klaszterbesorolást eredményezhet. A szerzők által használt adatsorok ebben az esetben tehát nem csak a besugárzás mértéke szerint normalizáltak, hanem az idő szerint is.

#### 2.4.3 Csoportok kialakítása

Mindkét változat segítségével négy csoportot különböztetnek meg. Az egyes csoportokba olyan napok kerülnek, melyek normált (és rendezett) görbéi a k-közép klaszterezés szerint egymáshoz hasonló jellegűek. Az így kialakított csoportok a következők.

4. táblázat: Chicco1 és Chicco2 módszer csoportjai

Csoport elnevezése	Jellemzés
Tiszta égboltú időjárás (Clear)	Haranggörbe jellegű, magas értékű besugárzási görbék
Közel tiszta égboltú időjárás (Quasi-clear)	Változékony, magas értékű besugárzási görbék
Közel felhős időjárás (Quasi-cloudy)	Változékony, alacsony értékű besugárzási görbék
Felhős időjárás (Cloudy)	Haranggörbe jellegű, alacsony értékű besugárzási görbék

A fenti módszer a hibatűrés növelése érdekében jelentős elhanyagolásokat és egyszerűsítéseket alkalmaz, melyek nagymértékben befolyásolhatják a csoportosítás eredményeit.

## 2.5 Almohammadi [8]

Az Almohammadi módszert 2016-ban fejlesztették ki, a Trueblood [2] módszerhez hasonlóan az USA-ban. A kutatók 10 perces felbontású, a University of California (San Diego) egyetemről származó, több, mint száz napnyi, 2014-es mérési adattal dolgoztak.

Ezt a módszert alapvetően nem csoportosításra találták ki, azonban azzal a feltételezéssel éltek, hogy kevert Gauss eloszlás (GMM – Gaussian Mixture Model) segítségével az éves besugárzási viszonyok reprodukálhatók. A szerzők normalizálták a besugárzási adatokat, ezzel függetlenítve azokat az évszakoktól, illetve a földrajzi helyszíntől. Ilyen normalizált adatsor gyakorlatilag a korábban említett  $CI$  index is. A [8] forrásban a normalizált adatokból sűrűségfüggvényeket készítettek, melyeket kevert Gauss eloszlásokkal közelítettek. A kutatók azt tapasztalták, hogy a tisztán felhős, illetve tisztán napos időszakok jól jellemezhetők egy modulusú Gauss eloszlással, ráadásul ezek várható értékei jellemzően 0,1, illetve 1 körüli értékeket vesznek fel. Azt is megfigyelték, hogy a változékony időjárású időszakokra általában két modulusú Gauss eloszlás jellemző, mely esetben a domináns modulus várható értéke a nagyobb. Ezen megfontolások alapján lehetséges csoportosítani az év napjait azok időjárása alapján. A módszer megvalósítása során a már korábban is alkalmazott  $CI$  indexek minden mintavételi időpontra számított változatára ( $CI_{sample}$ ) végeztem kevert Gauss eloszlás illesztést.

### 2.5.1 Csoportok kialakítása

A fent ismertetett Gauss eloszlások illesztésével a következő csoportok alakíthatók ki.

5. táblázat: Almohammadi módszer csoportjai

Csoport elnevezése	Jellemzés
Tiszta égboltú (Clear)	1 körüli várható értékű, egy modulusú Gauss eloszlás
Közepesen felhős (Medium)	Két modulusú Gauss eloszlás
Felhős (Cloudy)	0,1 körüli várható értékű, egy modulusú Gauss eloszlás

## 2.6 Módszerek összefoglalása

A bemutatott módszerek összefoglalását mutatja az alábbi 6. táblázat.

6. táblázat: Csoportosítási módszerek összehasonlítása

Módszer	Időfelbontás	Csoportok száma	Bemenő adatok és csoportosítás módja
Trueblood	1 sec (1 min)	5 osztály	Besugárzási görbe hossza ( $VI_{day}$ ) $Cl_{day}$ determinisztikus csoporthatárok
Maafi	10 min	3 osztály	Fraktáldimenzió ( $D_f(d)$ ) $Cl_{day}$ determinisztikus csoporthatárok
Fortuna	5 min	4 osztály	Teljesítményspektrum-sűrűség ( $I(d)$ ) $Cl_{day}$ k közép klaszterezés
Chicco 1	1 min	4 osztály	Mért besugárzási adatok (normált) k közép klaszterezés
Chicco 2	1 min	4 osztály	Mért besugárzási adatok (normált és rendezett) k közép klaszterezés
Almohammadi	10 min	3 osztály	$Cl_{sample}$ indexek k közép klaszterezés

Munkám során a 2. fejezetben ismertetett öt, egymástól jelentősen eltérő csoportosítási módszert vizsgáltam meg egy, az MTA Energiatudományi Központ (EK) által és egy, a Debreceni Egyetem kihelyezett tanszéke, a Megújuló Energiapark által szolgáltatott adatsor segítségével. Az öt megközelítés nagymértékű különbözősége miatt érdekes megfigyelni, hogy eredményeik milyen mértékben korrelálnak egymással.

## 3 Csoportosítás megvalósítása

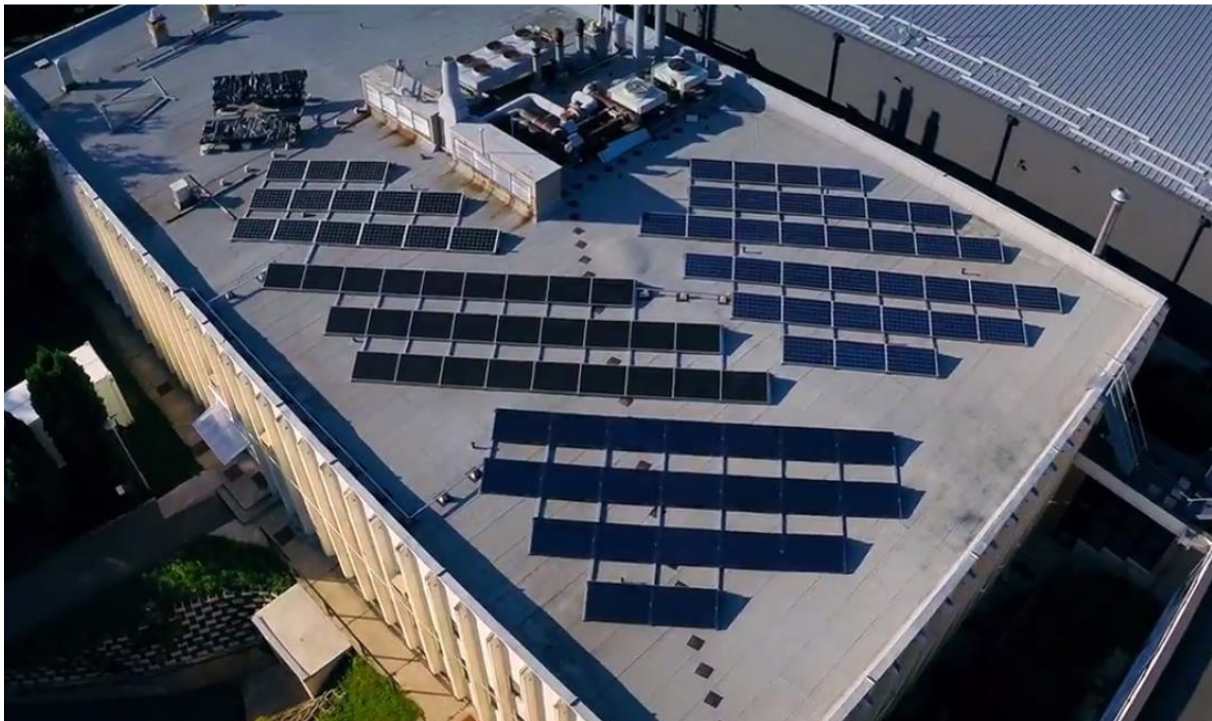
---

### 3.1 Energiatudományi Kutatóközpont (EK)

Munkám során a Magyar Tudományos Akadémia (MTA) Energiatudományi Kutatóközpontja (EK) által meghirdetett Hallgatói kutatói ösztöndíj keretében lehetőségem volt az EK Központi Fizikai Kutatóintézetben (KFKI) található napelemparkjának adataival megvizsgálni a 2. fejezetben ismertetett időjárás alapú csoportosítási eljárásokat.

#### 3.1.1 Napelempark [9]

A KFKI területén 2014-ben került sor az alábbi képen (2. ábra) látható napelemes termelőegység telepítésére.



**2. ábra: EK napelemes termelőegység [9]**

A termelőegység beépített teljesítménye 21 kWp, melyet három, különböző technológiával készült panelcsoport szolgáltat. Az egyes csoportok teljesítményei a következőképp oszlanak meg:

1. KORAX monokristályos: 3,50 kWp
2. SHARP vékonyréteg: 3,46 kWp
3. TrinaSolar polikristályos: 13,51 kWp

A napelemes rendszerhez energiatároló egységek is telepítésre kerültek, melyek kapacitása összesen 18 kWh. Az akkumulátorteleg három különböző fajta akkumulátor csoportból épül fel, melyek kapacitásmegoszlása a következő:

- 4. Hopecke ólom-savas: 8 kWh
- 5. Akasol Li-ion: 5,5 kWh
- 6. Shaft Li-ion: 5 kWh

Ez a napelemes, energiatárolóval támogatott rendszer látja el a telephelyen működő intelligens, energia-pozitív (éves szinten több energiát termel, mint fogyaszt) közvilágítási rendszert.

A termelőegységek mellett időjárás állomás helyezkedik el, melynek adatai egy központi rendszerben kerülnek eltárolásra. Az időjárás állomás méri a hőmérsékletet, légnyomást, páratartalmat, szélsőséget és a napsugárzást. Minden adatot perces és negyedórás felbontással tárol a számítógépes rendszer. Későbbi vizsgálataim során számomra a napsugárzási adatok bírtak nagy jelentőséggel az említett paraméterek közül. A napsugárzási adatok esetében globális, diffúz és e kettőből számított direkt besugárzási értékek is rendelkezésre állnak. Ezek az adatok, egész évre vonatkozóan, a 2015-ös évtől kezdve állnak rendelkezésre, ám több helyen, különböző okok miatt jelentősen hiányosak. Vizsgálataimhoz a 2018-as évi adatokat használtam fel, melyek csak elhanyagolható mértékű, könnyen korrigálható adathibával rendelkeztek. Az időjárás adatok a termelési adatokkal ekvivalensnek tekinthetők egy konstans szorzó elhanyagolásával, így a napok ezek alapján is megfelelően csoportosíthatók.

### 3.1.2 Adatfeldolgozás

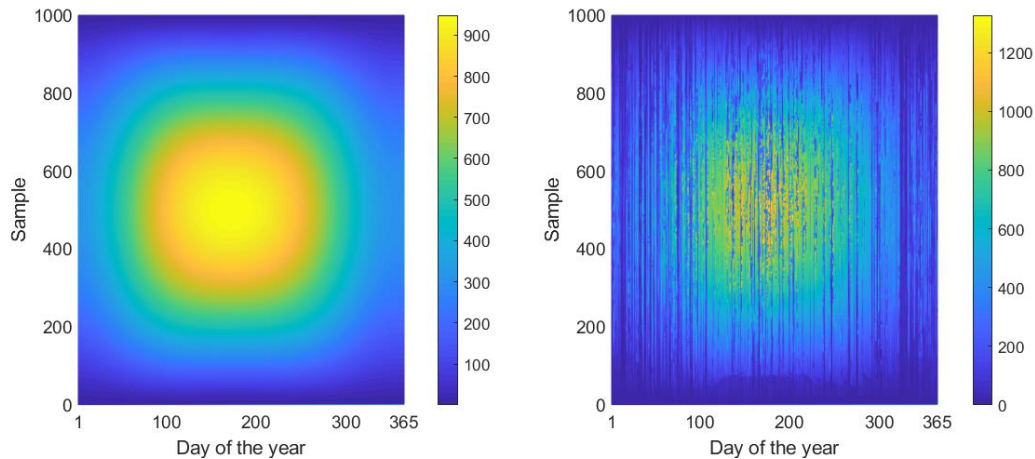
Munkám során 2018-as évi, szinte hiánytalan adatsorokkal dolgoztam. A legnagyobb adathiány mérete sem haladta meg a 3 órás időintervallumot (2018.10.20. 8:45-11:30). Ezeket a hiányokat, a számítások pontossága érdekében pótolnom kellett, amit lineáris interpolációval valósítottam meg. Adathiányokon kívül ismétlődések is előfordultak az adatsorban, valószínűleg adatrögzítési hibából adódóan. Ezen hibák korrigálás után végül egy teljes, 365 napos, 1 perces felbontású (1 nap 1440 adatpont) adatsorral tudtam tovább dolgozni, melyből a globális besugárzási értékeket vizsgáltam.

A tiszta ég görbéket a 1.1 fejezetben ismertetett Bird-féle tiszta ég modell segítségével számítottam ki, 365 napos évre, 0,2 értékű albedó és 1371 W/m<sup>2</sup> nagyságú napállandó beállításával, a megadott földrajzi lokációra (EK: 47.492, 18.954).

A mért és számított adatsor összehasonlíthatóságának érdekében szükség volt azok egymáshoz illesztésére. A mért adatsor nem tartalmazza az óráátállítást, míg a tiszta égbolt adatsor igen, így ez utóbbit a megfelelő hónapokban egy órával el kellett tolni. Amint azt az 1.1 fejezetben ismertettem, a Bird-modell feltételezi, hogy az év minden napján azonos időpontban éri el a Nap pályájának csúcát, ezért a mért adatsorokat is el kellett tolni az adott napnak megfelelő értékekkel, hogy ennek a feltételnek megfeleljenek.

Egyes számítások elvégzéséhez, ahogy azt a 2. fejezetben említettem, egységes hosszúságú adatsorokra van szükségünk, melyek csak a releváns adatpontokat tartalmazzák (az éjszakai zérus értékeket nem). Mivel az év során a nappali órák száma jelentősen változik, minden napra meg kellett határoznom a napsütötte órák időszakát. Ez a tiszta ég modellből származó adatokra egyszerűen meghatározható, azonban a mért adatsorban előfordulnak olyan értékek (pl. éjszakai órákban), amelyek biztosan nem napsugárzásból adódnak (hanem valami egyéb fényforrásból). Ezért egységesen a tiszta ég modell intervallumaival szűkítettem mind a két adathalmazt, melyet a csillagászati napkelte és napnyugta határoz meg. Így a legrövidebb adatsor 478 perces (december 21.), a leghosszabb pedig 931 perces (június 21. és 22.) lett. Ezek alapján a kialakítandó egységes adathossznak 1000 egységet választottam. Az egységes sorhosszúságok kialakítását újramintavételezéssel végeztem el, melyhez lineáris interpolációt alkalmaztam a hiányzó adatpontok meghatározásához. A két, egymáshoz igazított adatsort az alábbi 3. ábra mutatja.





**3. ábra: A tiszta égbolt (bal) és a mért (jobb) besugárzási adatok**

A fenti ábrán jól látható a két adatsor hasonlósága, megfigyelhető a tiszta égbolt adatsor szabályossága. Az adatok feldolgozását MS Excel és MATLAB programokkal végeztem, melyek segítségével létrehoztam a szükséges egy-, öt-, illetve tízperces felbontású adatsorokat.

### 3.1.3 Csoportosítás

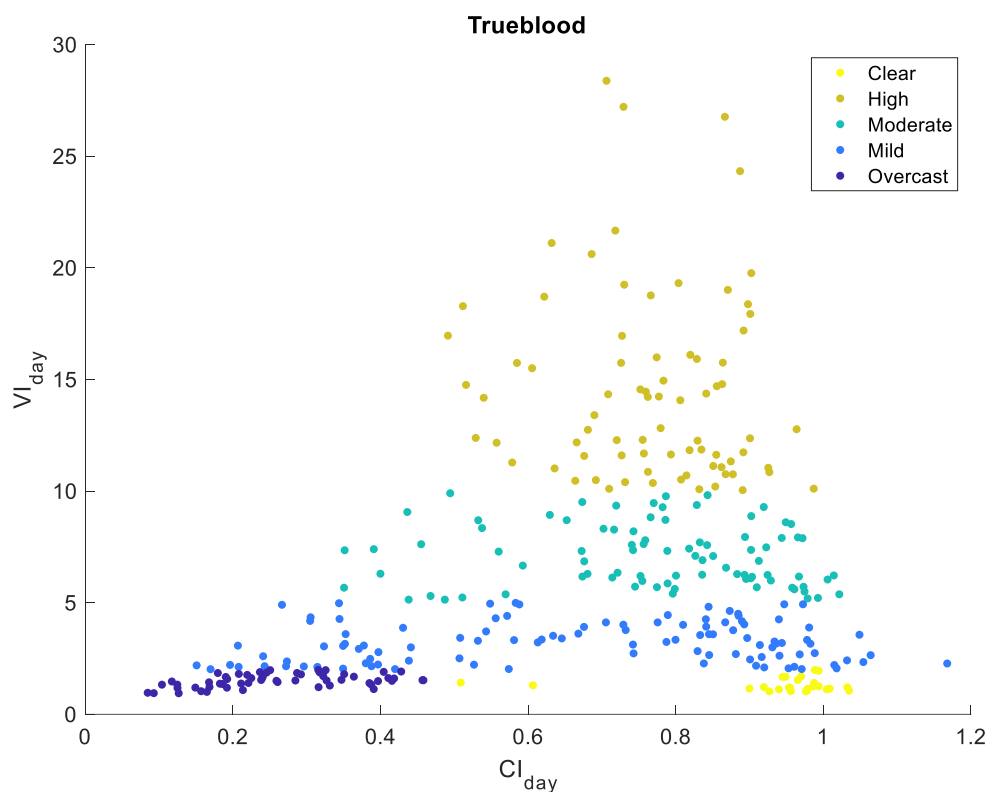
A 3.2.1 fejezetben ismertetett adatfeldolgozást követően már rendelkezésre álltak a csoportosítás elvégzéséhez szükséges adatsorok. A különböző időjárás mintázatokat mutató napok csoportosítását a 2. fejezetben ismertetett módszerek segítségével, esetleg azok egy kissé módosított változatával végeztem el. Determinisztikus módszerek esetén a tiszta ég- és variabilitási indexeket MATLAB segítségével számítottam ki, majd ezeket Excel táblázatba rendezve végeztem el a csoportosítást. A csoportosítás eredményeit felhasználva ismét MATLAB-ban elemeztem és jelenítettem meg a besorolt napokat. Nem determinisztikus módszerek esetén a teljes csoportosítási műveletsort MATLAB-ban hajtottam végre, annak klaszterezési algoritmusait (főként k-közép) felhasználva. A k-közép klaszterezés egy iteráción alapuló algoritmus, mely előre meghatározott számú (k db) középpont köré rendezi a csoportosítani kívánt adatsort. Az algoritmus első lépésben kiválaszt k db kezdő középpontot véletlenszerűen, az adatsor intervallumában, majd kiszámítja az egyes adatpontok és a középpontok közti távolságokat. Ez után minden pontot a hozzá legközelebb eső középpont csoportjába sorol. A következő lépésben az algoritmus minden csoportra meghatározza az egy csoportba tartozó pontok átlagos távolságát a középpontjuktól. Az iterációk során az algoritmus addig változtatja a középpontok helyzetét, míg a középpontok és a hozzájuk tartozó adatpontok közötti távolságok átlaga minimális nem lesz, vagy az algoritmus túl nem lépi az iterációk maximális számát. Az algoritmus akkor vezethet hibás eredményhez, ha a csoporton belüli átlagos távolságok egy lokális minimumát találja meg, ezért érdemes azt egymás után többször is (pl. a dolgozatban használt 50 ismétléssel) lefuttatni. [10]



Az egyes csoportok elnevezésére minden esetben az alkalmazott módszer eredeti csoportneveit használtam. A csoportosítás megvalósítása a következő fejezetben olvasható; a besorolások számszerűsített adatait a fejezet végén található 7. táblázat tartalmazza.

### 3.1.3.1 Trueblood

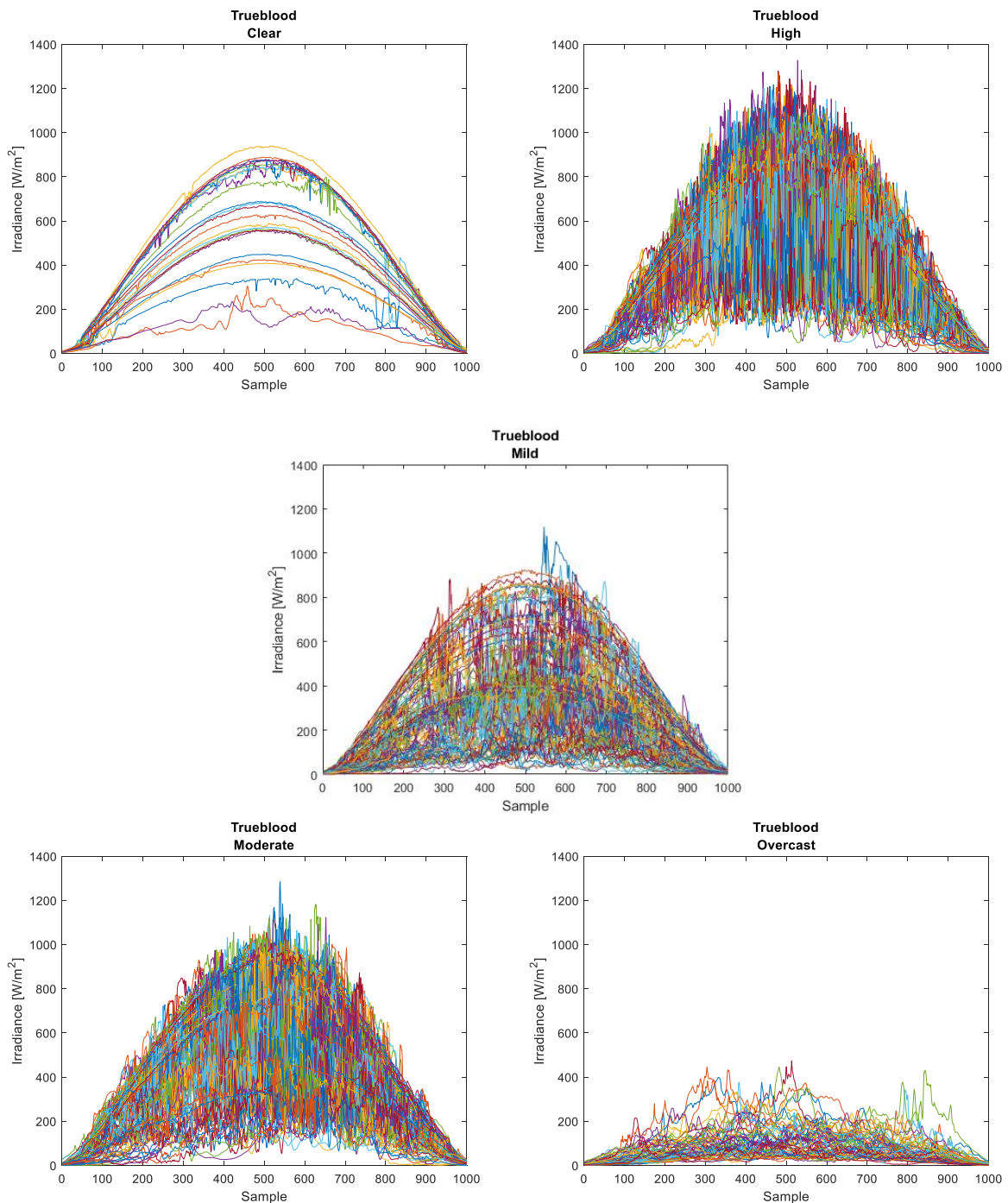
Ahogy a 2.1 fejezetben ismertetésre került, a Trueblood módszer egy tiszta égbolt- és egy variabilitási indexet használ a napok csoportosításához. Az algoritmus előírásainak megfelelően, első lépésben meghatároztam ezeket az indexeket ( $CI_{day}$  és  $VI_{day}$ ) MATLAB segítségével, egyperces felbontású, újramintavételezett adatok felhasználásával. A kiszámított indexek alapján, Excel használatával elvégeztem a csoportosítást. Ennek eredményei az alábbi ábrán (4. ábra) láthatók.



4. ábra: Trueblood módszer csoportosítási eredmények

A fenti ábrán jól megfigyelhetők a módszer által meghatározott determinisztikus csoporthatárok. Azt is megfigyelhetjük, hogy ez a módszer aránylag kevés napot sorol tiszta égboltúnak. Ezzel a későbbiekben részletesebben is foglalkozunk majd.

A mennyiségi eloszláson kívül érdemes megvizsgálni az azonos csoportba sorolt napok besugárzási görbéit is. Ezeket az alábbi, 5. ábra tartalmazza.



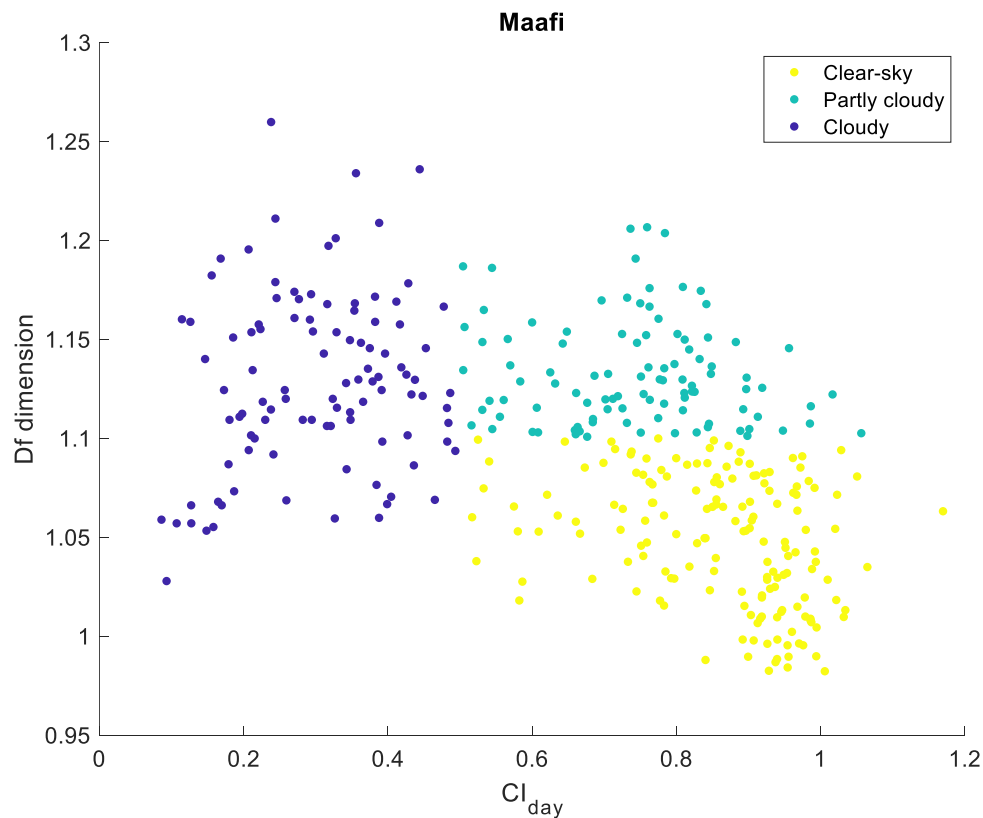
5. ábra: Trueblood módszer csoportok

A fenti ábrán jól láthatók az egyes csoportok közötti különbségek. Megfigyelhető, hogy az első (Clear) és az ötödik (Overcast) csoport markánsan különbözik egymástól, azonban a köztes csoportok (High, Mild, Moderate) idősorai között vizuálisan nem tehető egyértelmű különbség. Ez a tény felveti a kérdést, hogy vajon van-e a gyakorlatban haszna ilyen nagy csoportszám alkalmazásának.

### 3.1.3.2 Maafi

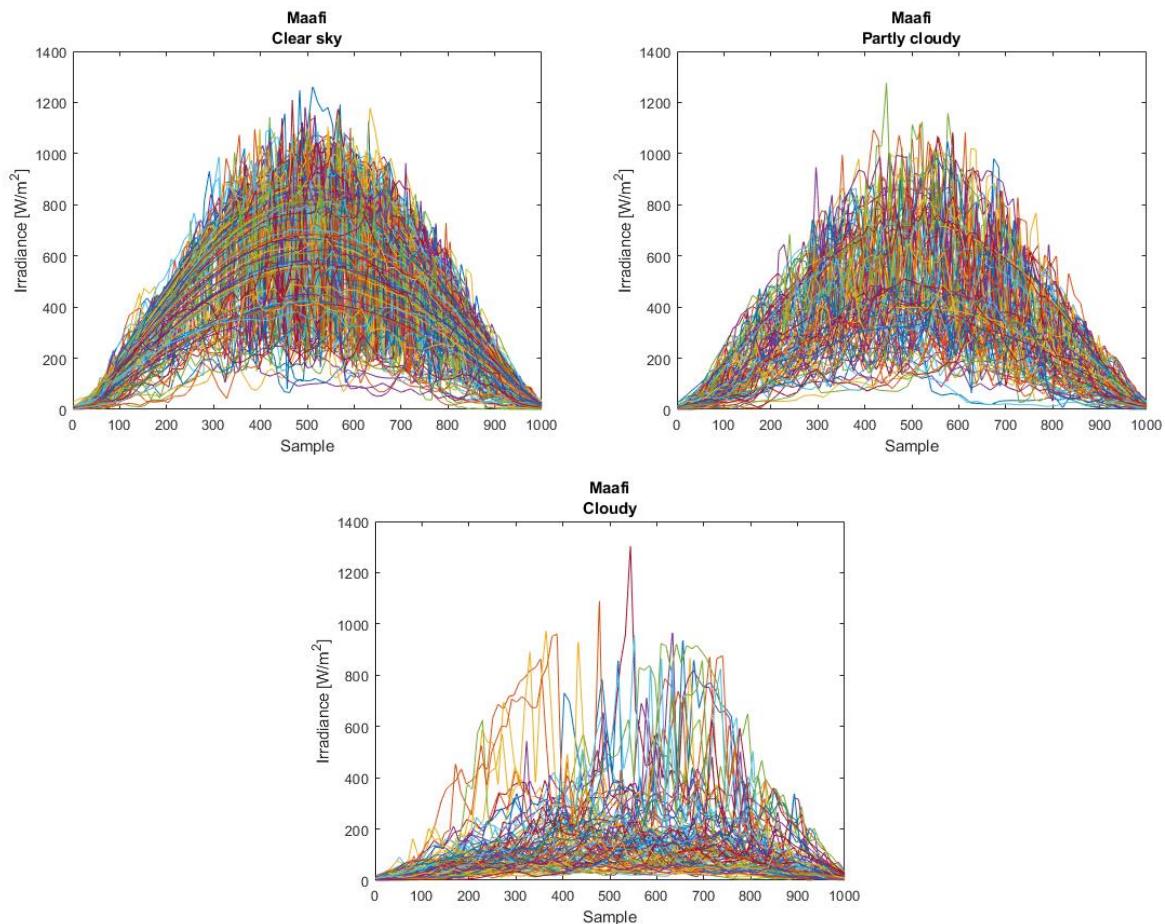
A Maafi módszer is két indexet használ, ahogy azt a 2.2 fejezetben ismertettem. A tiszta égbolt index, megegyezik az előző esetben használt  $Cl_{day}$  indexszel, a variabilitási index azonban az egyes napok

fraktáldimenziójával (Minkowski dimenziójának átlagértékével) azonos. Az algoritmus bemeneteként tízperces felbontású, újramintavételezett adatsort használtam, ahogy [3] szerzői is. A tanulmányban leírnak megfelelően, először ezeket az indexszámokat számítottam ki MATLAB segítségével, majd Excel-ben, a megfelelő csoporthatárokat beállítva, elvégeztem a csoportokba sorolást. Az így kapott eredményeket MATLAB-ban jelenítettem meg. Ez látható az alábbi ábrán (6. ábra).



6. ábra: Maafi módszer csoportosítási eredmények

A fenti ábrán egyértelműen láthatók a determinisztikus csoportosítási módszer által megszabott besorolási határok. Megfigyelhető, hogy a három csoport elemszáma hasonló nagyságrendű, de jól láthatóan tiszta égboltú (Clear sky) napból van a legtöbb. Ez ellentmond az előző módszer esetén tapasztaltakkal, a későbbiekben ennek okát alaposabban is meg fogjuk vizsgálni. Ehhez érdemes áttekinteni az egyes csoportok elemeinek besugárzási görbéit. Ezek az alábbi ábrán (7. ábra) láthatók.



7. ábra: Maafi módszer csoportok

A fenti ábra alapján az egyes csoportok között egyértelmű különbségek látszanak, bár az első és második (Clear sky és Partly cloudy) csoport görbeseregei igen hasonló összképet mutatnak. Az első csoport esetén azonban észrevehetően kirajzolódnak tisztán haranggörbe alakú görbék. A harmadik csoport (Cloudy) görbeseregén észrevehetjük, hogy a kiugró értékektől eltekintve, az alacsony besugárzású görbék dominálnak.

### 3.1.3.3 Fortuna

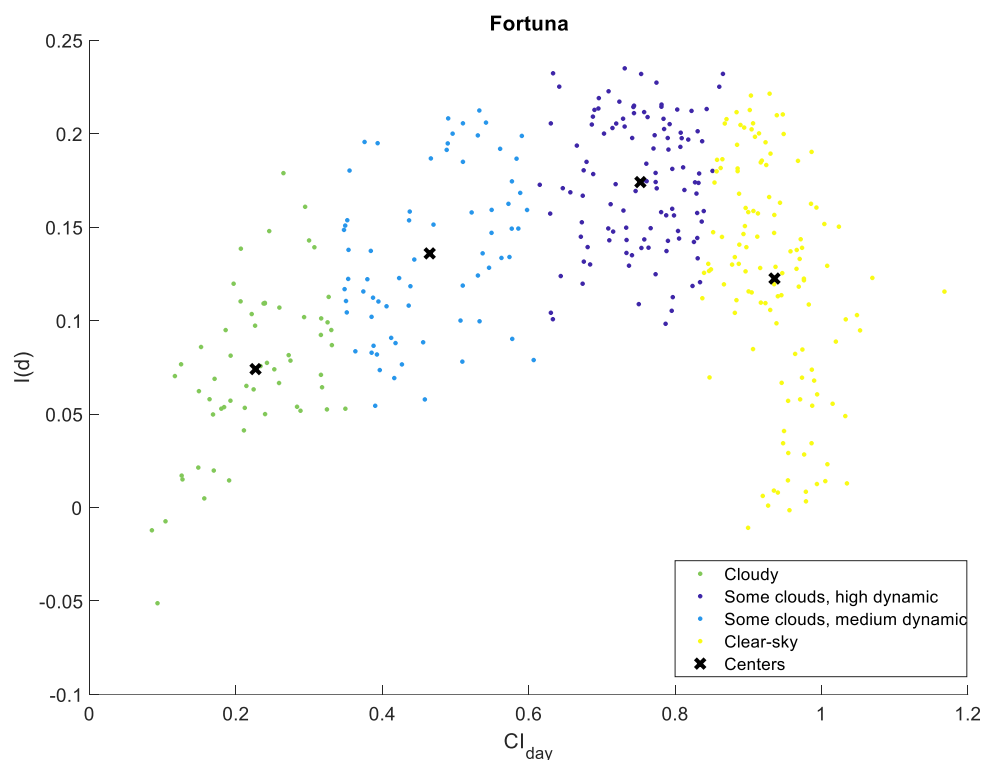
A Fortuna módszer esetén a csoportosítás alapját egy területarány - és egy változékonyság ( $I(d)$ ) index adja. A területarány index gyakorlatilag megegyezik a korábban alkalmazott  $CI_{day}$  indexszel, így ebben az esetben is ezt használtam. A változékonyság indexeket MATLAB segítségével számítottam ki. Az algoritmus bemeneteként öt perces felbontású mérési adatokat és az eredeti módszer tipikus nap adataitól eltérően, tiszta ég adatokat használtam.

Először kiszámítottam mind a mérési adatok, mind a tiszta ég adatok teljesítményspektrum sűrűségeit. Ahhoz, hogy ezek energiatartalmát össze tudjam hasonlítani, a teljes frekvenciatartomány helyett én is egy szűkebb tartományt vizsgáltam, az eredeti módszernek megfelelően. Vizsgálataimat a mintavételi frekvencia (1/300 Hz) tízszerese és százszorosa közötti, tehát az [1/30, 1/3] Hz

tartományban folytattam. Ezzel a lépéssel kiszűrhető az alacsony frekvenciaértékekhez tartozó közel azonos energiatartalom. Az így kiszámított, szűkített frekvenciatartományra értelmezett teljesítményspektrum sűrűségekből minden napra kiszámítottam a besugárzás energiatartalmát. Ezen értékek szolgálták a (20) egyenlet bemeneteként, melynek segítségével megkaptam az egész vizsgált évre minden egyes nap  $I(d)$  indexét.

Az kiszámított  $I(d)$  indexek és az öt perces felbontású tiszta ég és mérési adatokból számított  $CI_{day}$  indexek felhasználásával már könnyedén el lehetett végezni a csoportosítást. Ezen paraméterek szolgálták ugyanis a nem determinisztikus k-közép klaszterezési algoritmus bemeneteiként, melyet 50 ismétléssel futtattam le MATLAB segítségével, a hibaértékek esetleges lokális minimumainak elkerülése érdekében.

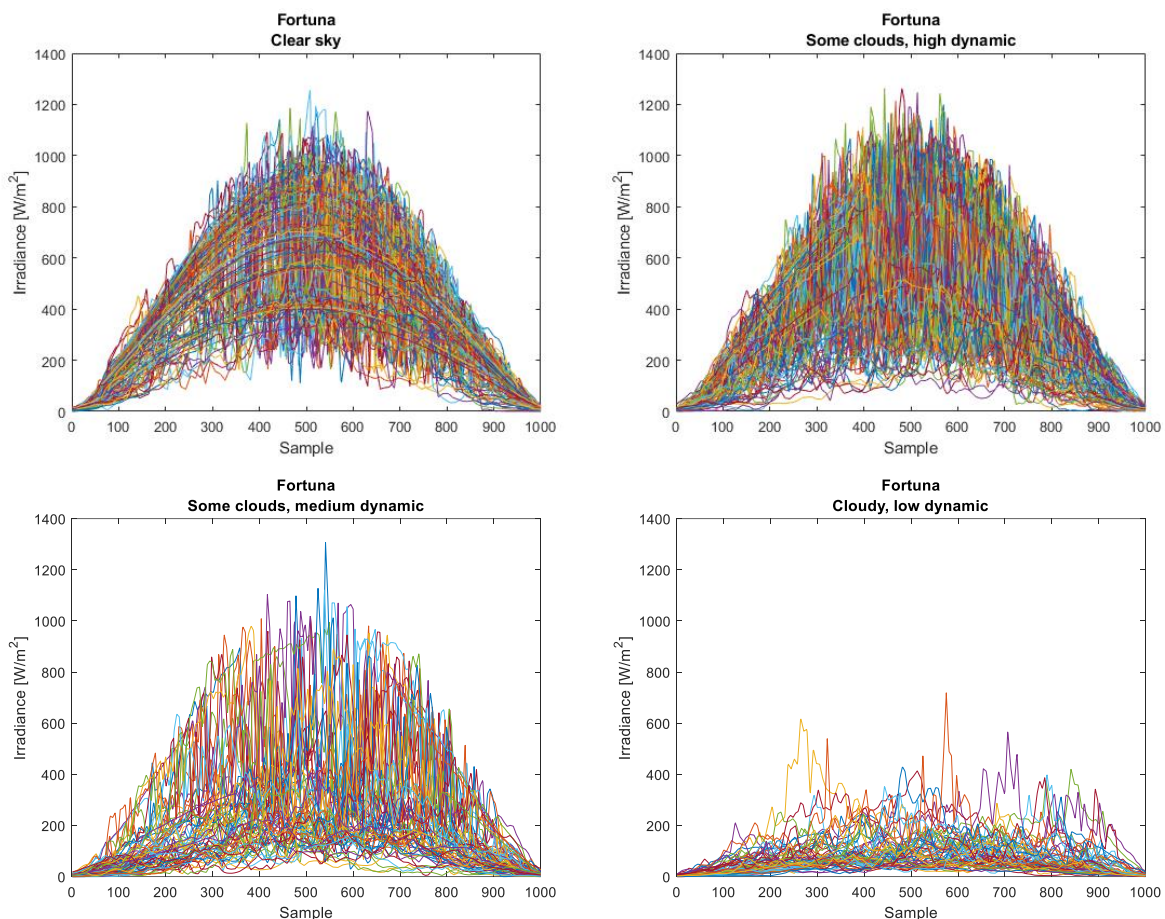
A besorolás eredményeit az alábbi 8. ábra mutatja.



**8. ábra: Fortuna módszer csoportosítási eredmények**

A fenti ábrán egyértelműen megfigyelhető a négy, egymástól elkülönülő csoport. Az érdekesség kedvéért megjelenítettem a klaszterezési algoritmus által kiválasztott, a kialakított csoportokhoz tartozó középpontokat is.

Az eredmények további szemléltetése érdekében az egyes csoportokhoz tartozó besugárzási görbéket is megjelenítettem. Ezek láthatók az alábbi ábrán (9. ábra).



9. ábra: Fortuna módszer csoportok

A fenti ábrán jól látható, hogy az első (Clear sky) és az utolsó (Cloudy, low dynamic) csoportok markánsan elkülönülnek egymástól. A középső két csoport (Some clouds, high dynamic és Some clouds, medium dynamic) azonban hasonlóságot mutat. Ez felveti a kérdést, hogy vajon hasznos e a gyakorlatban ezeket megkülönböztetni egymástól, hiszen a 8. ábra alapján az látható, hogy ezen napok változékonysági indexei közel azonos tartományba esnek, csak a tiszta égbolt indexeik tartományában van jelentősebb különbség.

### 3.1.3.4 Chicco

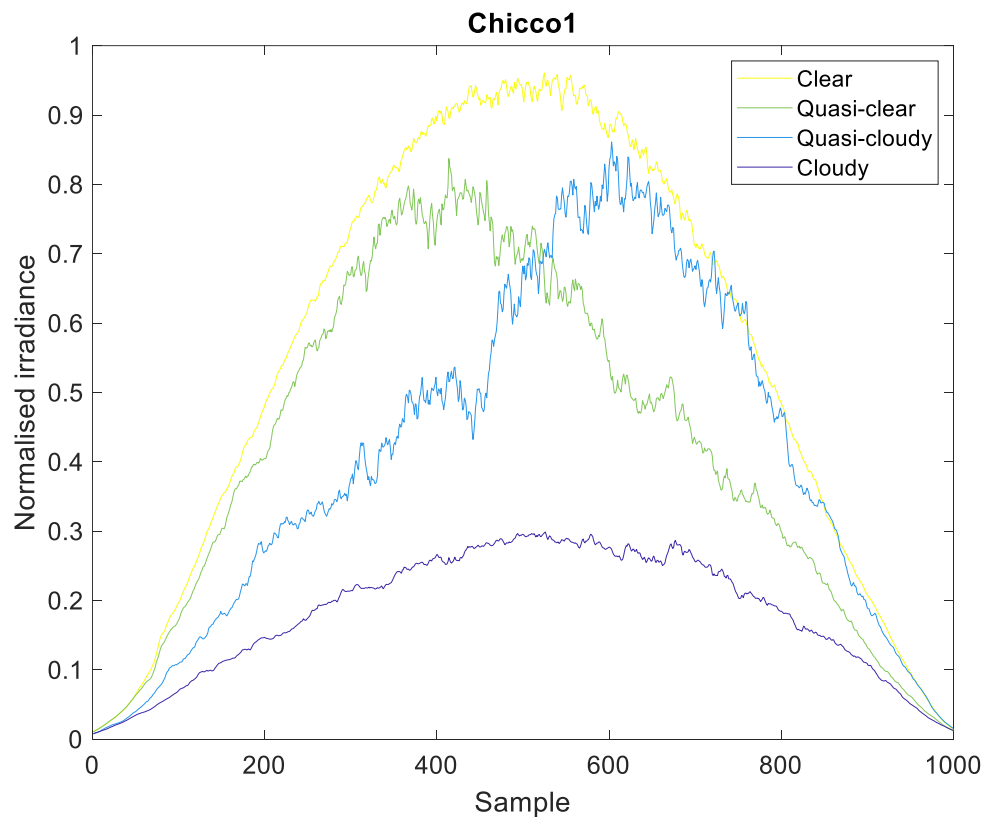
A Chicco módszer, ahogy azt a 2.4 fejezetben ismertettem, csak a besugárzási és tiszta ég adatok alapján kísérli meg csoportosítani a napokat, nem alkalmaz semmilyen minőségi jellemzőre vonatkozó indexet. Ennek megfelelően én is egy perces felbontású, újramintavételezett adatokat használtam bementeként. Ezeket az adatokat első lépésben a tiszta ég modellből származó adatok alapján normáltam MATLAB segítségével.

#### Idősoros, normált (Chicco1)

Az Chicco módszer idősoros változata esetén nincs is szükség az adatok további feldolgozására. A k-közép klaszterezést ezekkel az adatokkal hajtottam végre, szintén MATLAB használatával. A

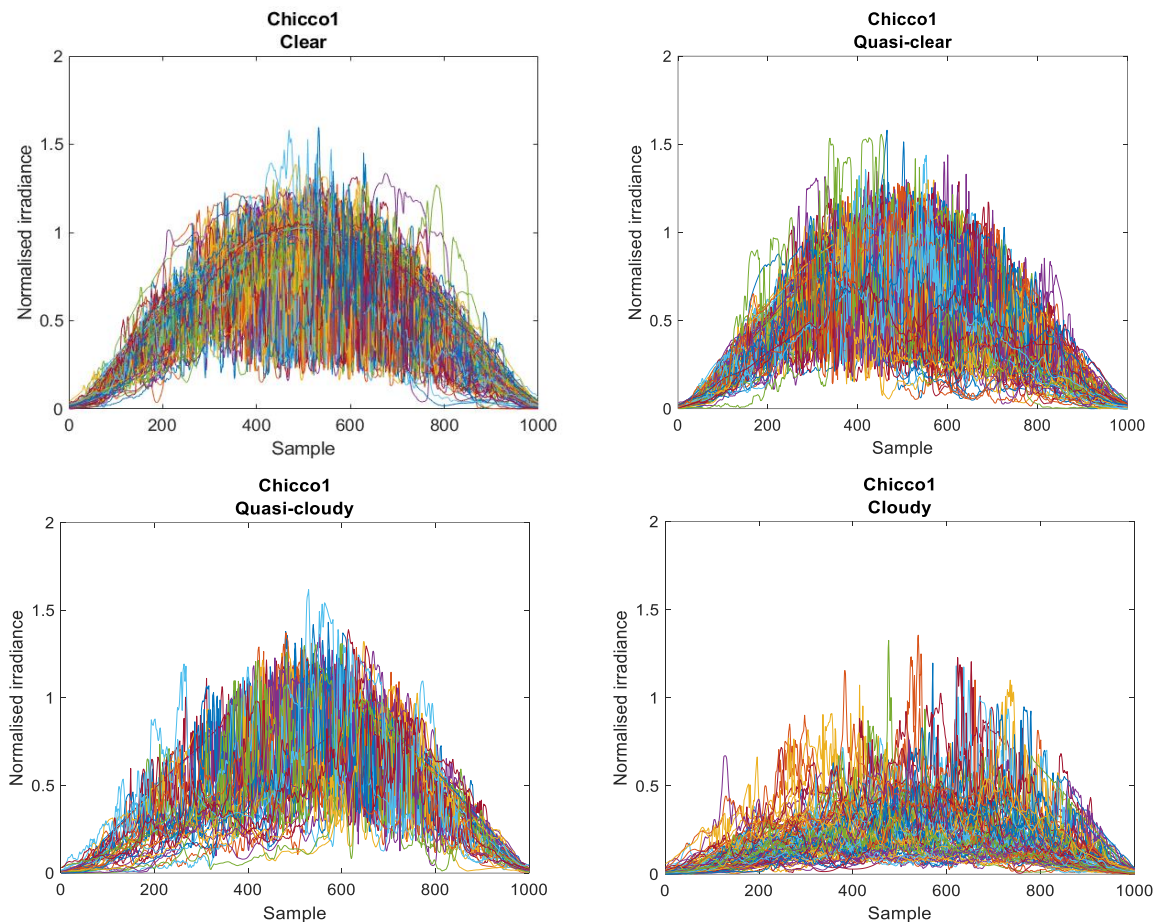


klaszterezés eredményeit úgy jelenítettem meg, hogy az egyes csoportok átlagértékeit ábrázoltam. Ez látható az alábbi ábrán (10. ábra).



**10. ábra: Chicco1 módszer csoportosítási eredmények átlagai**

A fenti ábrán megfigyelhető, hogy a négy csoport úgy helyezkedik el, ahogyan azt a csoportnevek alapján feltételeznénk. A középső két csoport (Quasi-clear és Quasi-cloudy) érdekes szimmetriát mutatnak. Úgy tűnik, mint ha az előbbi csoport tagjaira a nap második felében egy, az ideálisnál (tisztaság modell) intenzívebb besugárzáscsökkenés lenne jellemző, míg ugyan ez az utóbbi csoport tagjait a nap első felében jellemezné. A két szélső csoport (Clear és Cloudy) megfelel a „teljesen napos” és „teljesen felhős” eseteknek, kisebb fluktuációkkal, majdnem szabályos haranggörbe alakúak. Ezek alapján ez a módszer igen jól alkalmazhatónak tűnik. Érdekes lehet azonban megvizsgálni az egyes csoportok besugárzási görbéit külön-külön is, nem csak átlagolva. Ezeket a görbeseregeket tartalmazza az alábbi 11. ábra.



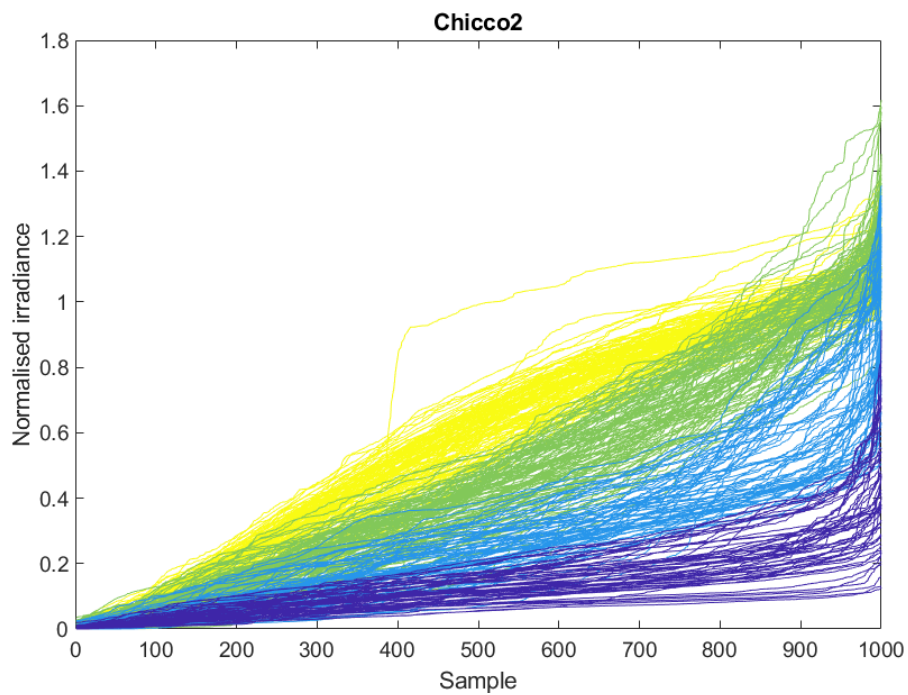
11. ábra: Chicco1 módszer csoportok

A fenti ábrán jól látható, hogy ilyen megjelenítési mód esetén a szabályosság és szimmetria nem annyira szembetűnő, mint a 10. ábra esetén. Érdekes megfigyelni, hogy az előző módszerektől eltérően ebben az esetben a „Clear” csoport görbeseregében nem rajzolódnak ki egyértelműen a szabályos, haranggörbe alakú görbék.

### Rendezett, normált (Chicco2)

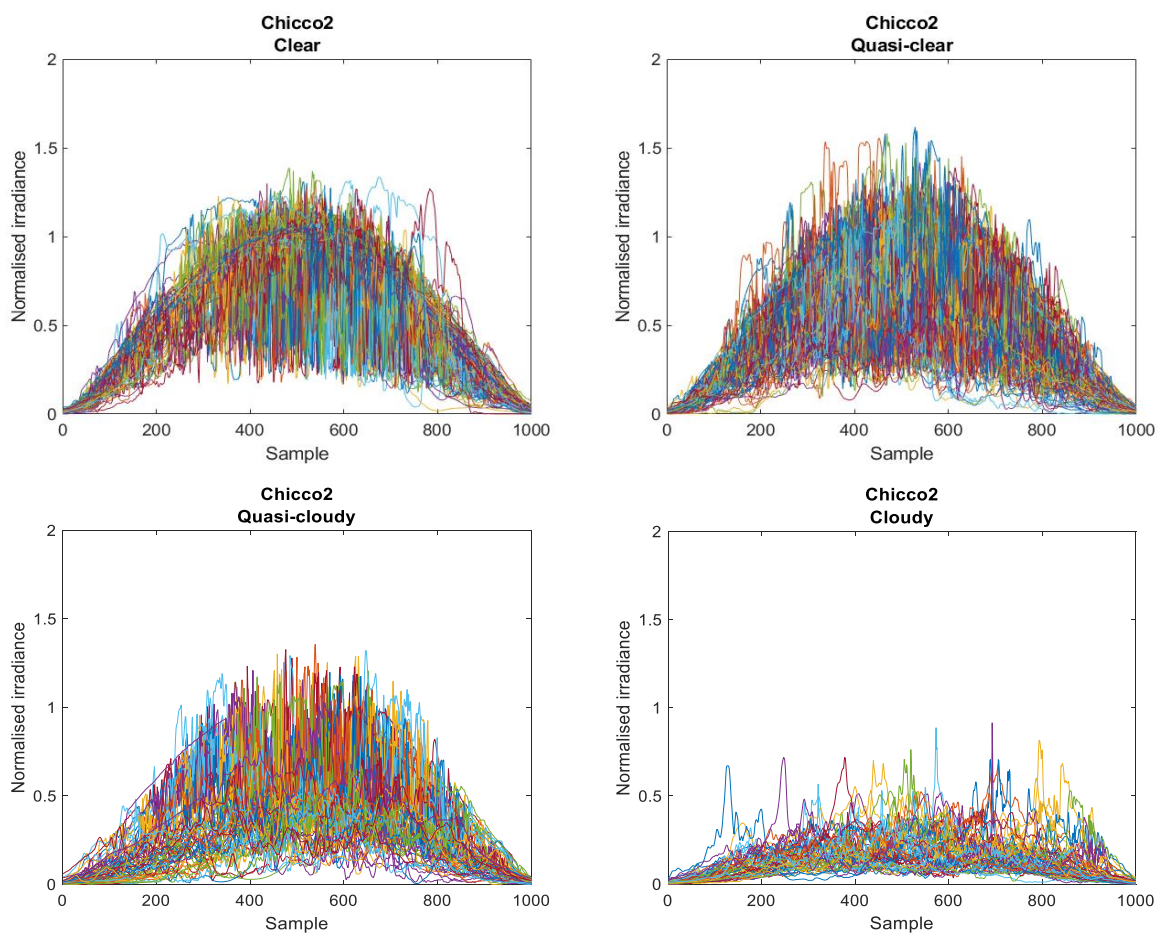
A Chicco módszer rendezett változata a tiszta ég adatokkal normált mérési adatokat gyakorlatilag eloszlásfüggvényekbe rendezi. Az adatoknak ezt a rendezését MATLAB segítségével végeztem el. Az eloszlásfüggvények a k-közép klaszterezési algoritmus bemeneteként szolgáltak. Így alakult ki az alábbi ábrán (12. ábra) látható csoportosítás, ahol sárga színnel a „Clear”, zöld színnel a „Quasi-clear”, világos kék színnel a „Quasi-cloudy”, sötétkék színnel pedig a „Cloudy” csoport görbeseregét jelöltem. Ebben az esetben is jól elkülöníthetők a csoportok, bár a megjelenítési forma (ahogy a csoportosítási alapelv is) kissé nehezen értelmezhető az előzőkhöz képest.





12. ábra: Chicco2 módszer csoportosítási eredmények

A könnyebb értelmezés érdekében ebben az esetben is megjelenítettem az egyes csoportok besugárzási görbéinek halmazát. Ezek láthatók az alábbi ábrán (13. ábra).



13. ábra: Chicco2 módszer csoportok

A fenti ábra alapján megállapítható, hogy a két módszer eredményei (Chicco1 és Chicco2) ránézésre igen hasonló, azzal a különbséggel, hogy Chicco2 módszer esetén a „Cloudy” csoportot kisebb besugárzási értékek jellemzik.

### 3.1.3.5 *Almohammadi*

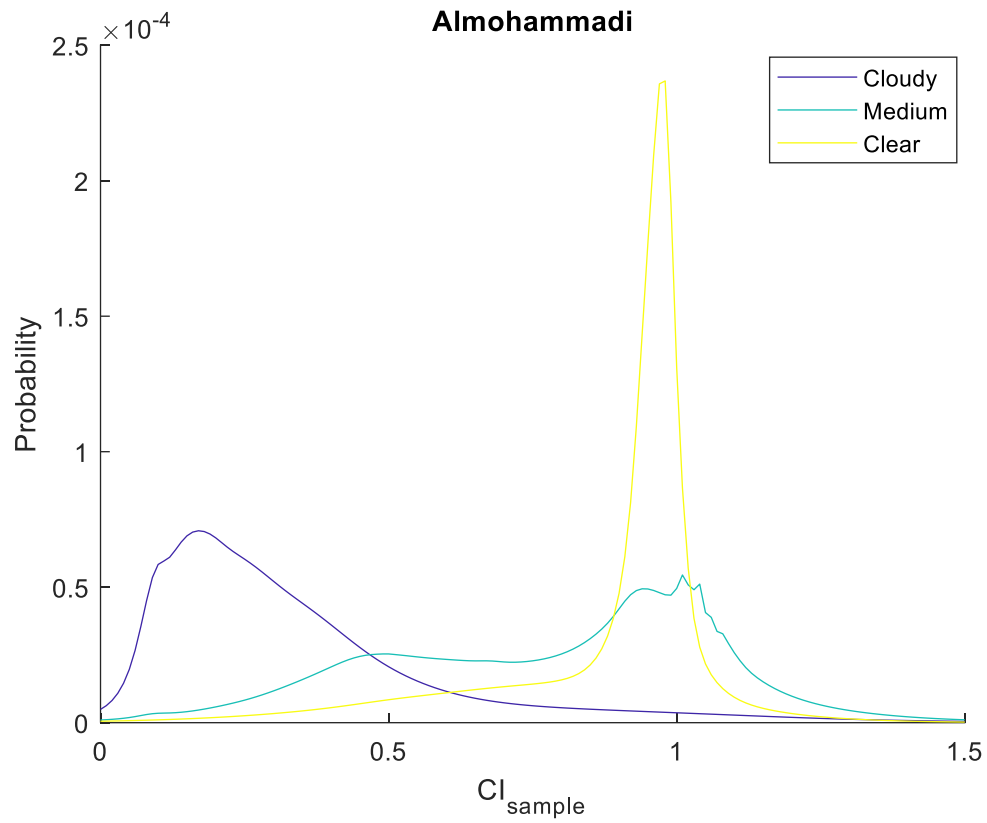
A [8] forrás, bár alapvetően nem végez csoportosítást, de feltételezi, hogy kevert Gauss eloszlás segítségével az éves besugárzási viszonyok reprodukálása lehetséges. Korábbi, jelen munkában nem részletezett vizsgálatok során bebizonyosodott, hogy valóban lehetséges ilyen módszerrel csoportosítást végezni.

Munkám során a csoportok meghatározásához a korábban ismertetett  $CI$  indexet használtam, azonban ebben az esetben az egyes napok minden egyes mintavételi időpontjára kiszámítottam MATLAB segítségével ezt az értéket ( $CI_{sample}$ ). Az index meghatározásához tízperces felbontású, újramintavételezett mérési és tiszta ég adatokat használtam. Azt tapasztaltam, hogy a napok elején és végén, viszonylag magas mért besugárzás és alacsony tisztaég értékek esetén az átlaghoz (0,6781) képest jelentősen kiugró (9,7378) értékeket mutat a számított index. Az ebből adódó esetleges hibák elkerülése érdekében megvizsgáltam, hogy milyen felső határ szabható az indexszámnak. Vizsgálataim eredményeképp arra jutottam, hogy 1,5 felső határt választva az adatoknak csak 1,3%-át kell figyelmen kívül hagynom, így ezzel a választott felső határral dolgoztam. Az hiányzó adatpontokat újramintavételezéssel, lineáris interpolációval oldottam meg.

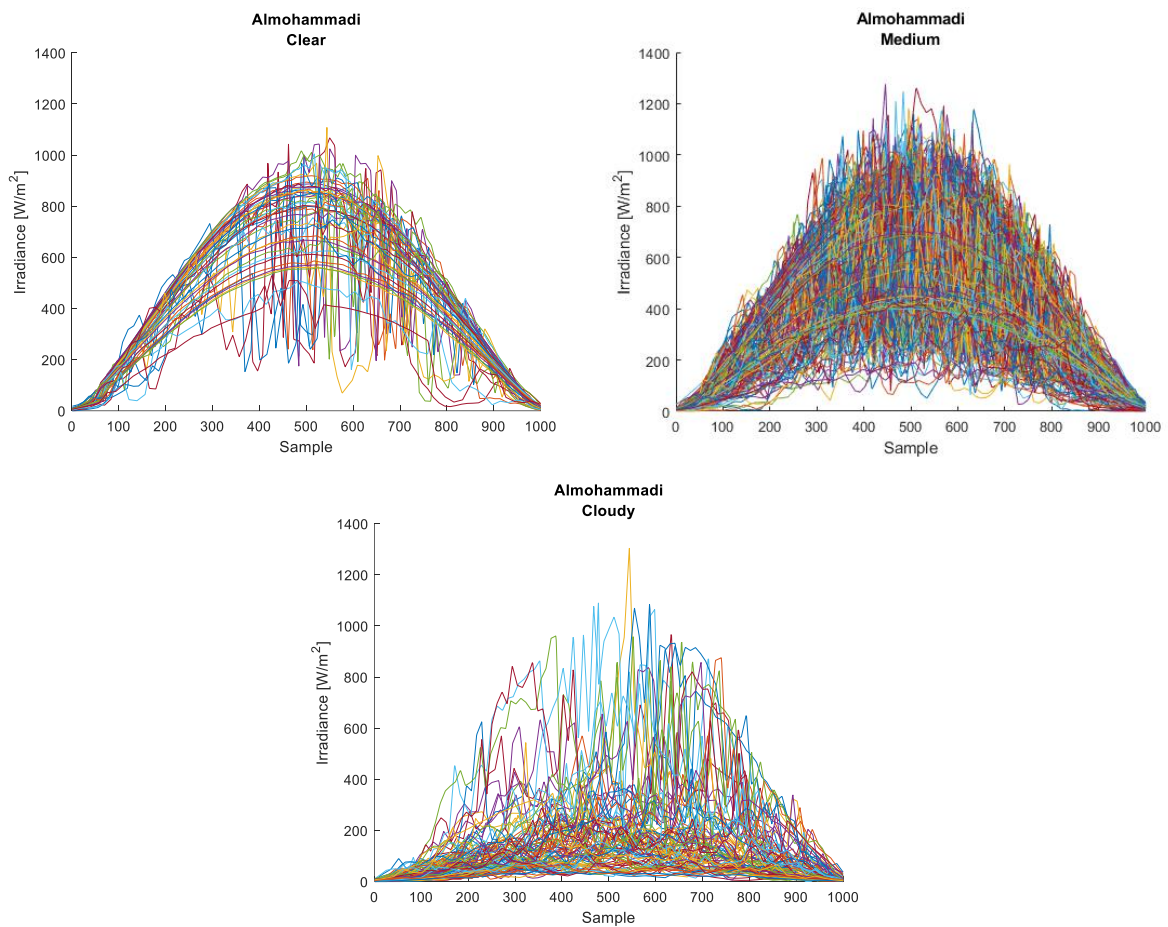
Az így számított  $CI_{sample}$  indexekre, minden napra kétmodusú Gauss eloszlást illesztettem, szintén MATLAB segítségével. Az eloszlásfüggvényeket k-közép klaszterezési algoritmus használatával rendeztem három csoportba. A csoportosítás eredményeit úgy jelenítettem meg az alábbi ábrán (14. ábra), hogy képeztem az egyes csoportok eloszlásfüggvényeinek összegét.

Az ábrán (14. ábra) jól megfigyelhető, hogy két csoport egyértelműen elkülönül. Ezek a tiszta égbolthoz (Clear) és a felhős időjárású napokhoz (Cloudy) tartozó csoportok, melyek esetében a Gauss eloszlás várható értéke 1, illetve 0,1 körüli. A harmadik csoporthoz azonban két modulus is tartozik, ezzel utalva arra, hogy az ide tartozó napok esetén felhős és napos időszakok váltják egymást. Ebben az esetben az egyes modulusokhoz tartozó várható értékek nem különülnek el érdemben egymástól.

A csoportok jellemzőinek könnyebb megfigyelhetősége érdekében az egyes csoportokhoz tartozó besugárzási görbéket is ábrázoltam (15. ábra). A csoportok az eddigiek alapján elvárható eredményeket mutatják. A „Clear” csoportban markánsan kirajzolódnak a szabályos haranggörbék, ahogy több korábbi módszer esetén is. Érdekes módon azonban a „Medium” csoport esetén is megfigyelhető ez, azonban egyértelműen a nagy fluktuációjú besugárzásokkal jellemezhető napok a dominánsnak. A „Cloudy” csoport napjaira pedig, néhány kivétellel az alacsony besugárzás jellemző.



14. ábra: Almohammadi módszer csoportosítási eredmények összegei



15. ábra: Almohammadi csoportok

### 3.1.3.6 Eredmények

A fent elvégzett csoportosítások összegzett, számszerű eredményeit tartalmazza az alábbi táblázat, a könnyebb összehasonlíthatóság érdekében százalékos értékekben (darabszám/365\*100) kifejezve.

7. táblázat: Csoportosítási eredmények - EK 2018

Módszer	Csoport elnevezése és darabszáma [%]				
<b>Trueblood</b>	Clear	High	Moderate	Mild	Overcast
	7,40	21,92	26,30	29,04	15,34
<b>Maafi</b>	Clear-sky		Partly cloudy		Cloudy
	44,66		27,40		27,95
<b>Fortuna</b>	Clear sky	Some clouds, high dynamic	Some clouds, medium dynamic	Cloudy, low dynamic	
	33,70	30,96	19,18	16,16	
<b>Chicco1</b>	Clear	Quasi-clear	Quasi-cloudy	Cloudy	
	40,55	21,64	11,51	26,30	
<b>Chicco2</b>	Clear	Quasi-clear	Quasi-cloudy	Cloudy	
	34,79	30,96	19,18	15,07	
<b>Almohammadi</b>	Clear		Medium		Cloudy
	13,42		56,44		30,14

A fenti 7. táblázat eredményei alapján jól látható, hogy a különböző módszerek által kialakított csoportosítások mennyire nem egységesek, azok végeredményeiben nagymértékű eltérések figyelhetők meg.

## 3.2 Debreceni Megújuló Energiapark

A Debreceni Egyetem kihelyezett tanszékétől, a Megújuló Energiaparktól származó 10 perces felbontású, 2018-as évi besugárzási adatok kerültek feldolgozásra, a releváns, 10 perces felbontású adatsorokat alkalmazó módszerek (Maafi és Almohammadi) segítségével.

### 3.2.1 Adatfeldolgozás

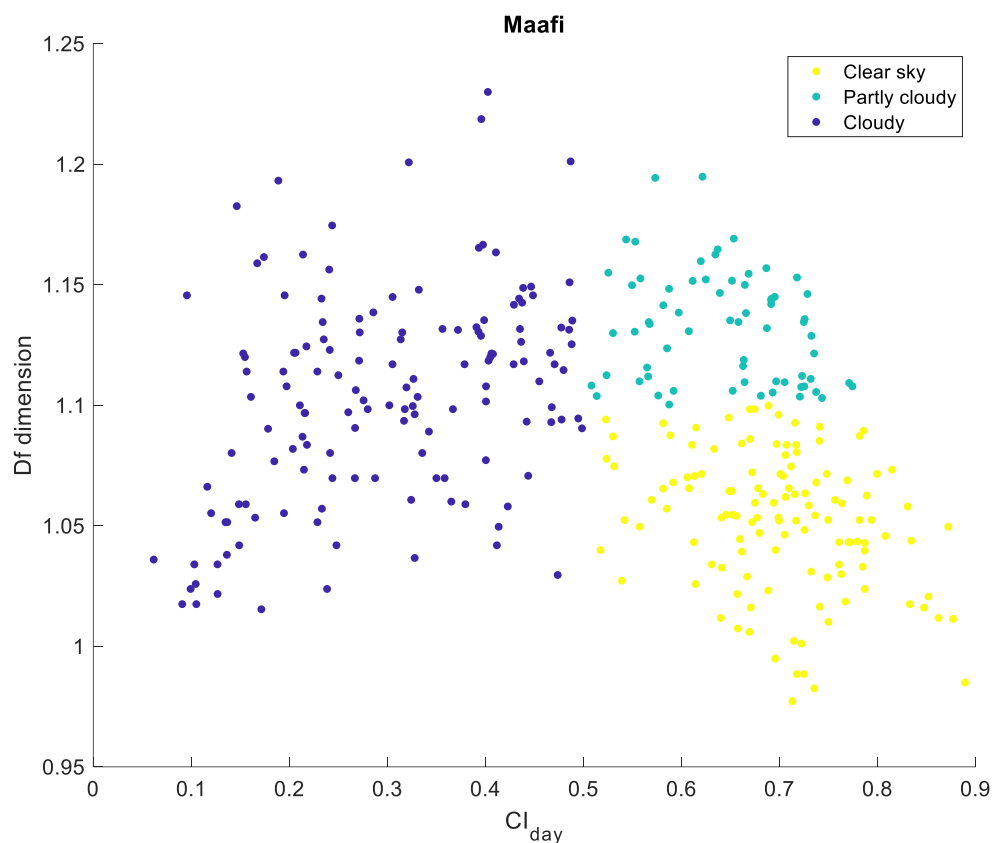
A rendelkezésre álló adatsor igen hiányos volt, így a duplikációk és hiányos napok eltávolítása, valamint a kismértékben hiányos napok kiegészítése (lineáris interpolációval) után végül 333 teljes napnyi adatmennyiséggel lehetett elvégezni a csoportosítást.

### 3.2.2 Csoportosítás

A Bird-féle tiszta ég modell segítségével legenerálásra került a tiszta ég adatsor, a mért adatsorhoz illeszkedő 333 napra, 0,2 albedó és 1371 W/m<sup>2</sup> nagyságú napállandó beállításával, a megfelelő földrajzi területre (Debreceni Megújuló Energiapark: 47.530, 21.575). Ezeket az értékeket referenciaként használva, elvégeztem a csoportosításokat a megfelelő módszerekkel a 3.1.3 fejezetben leírtakkal megegyező módon.

#### 3.2.2.1 Maafi

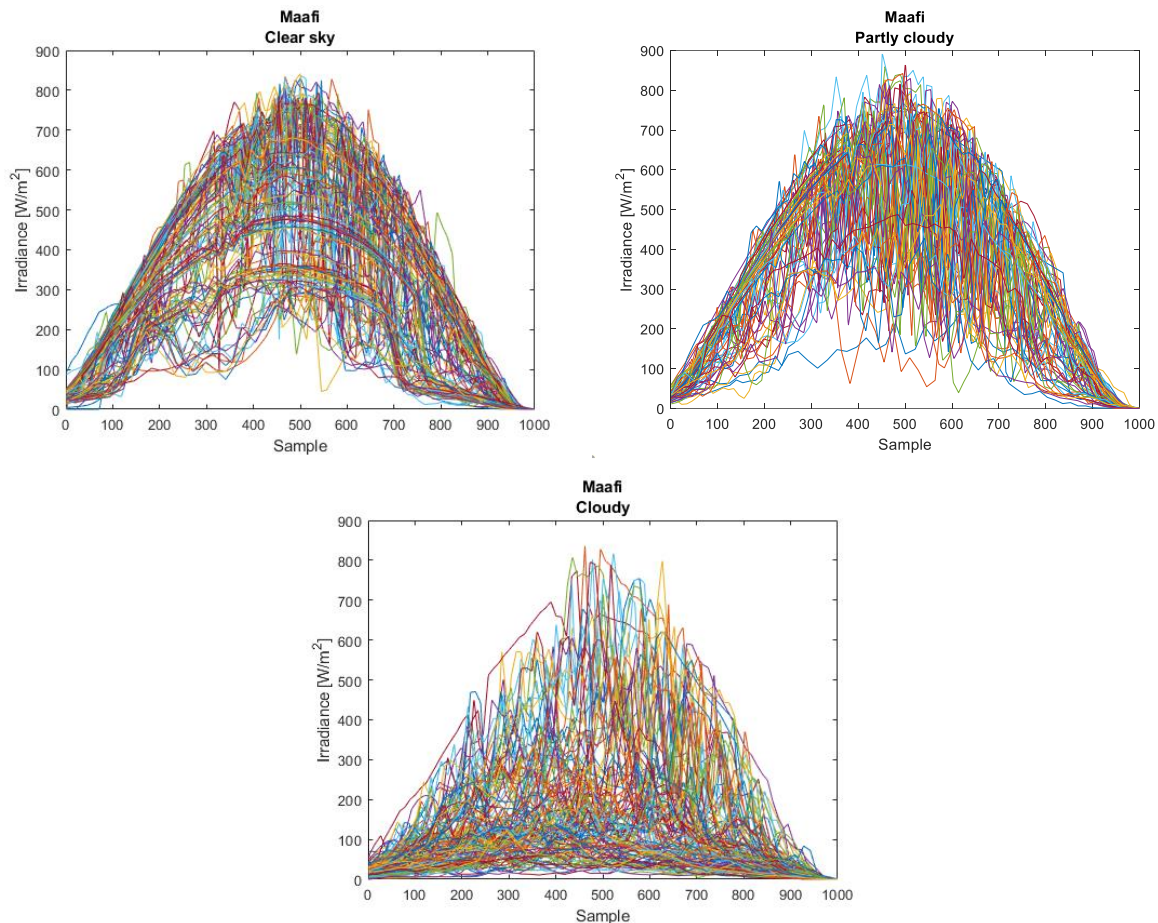
A 3.1.3.2 fejezetben leírt lépések és műveletek végrehajtása után, ebben az esetben az alábbi ábrán (16. ábra) látható eredményeket kaptam.



16. ábra: Maafi módszer csoportosítási eredmények - Debrecen

A fenti ábrán jól megfigyelhetők a determinisztikus csoporthatárok, ahogy a korábban vizsgált adatsor esetén is. Érdekes azonban, hogy ebben az esetben a felhős („Cloudy”) napok dominálnak (143 db) és csak igen kevés nap (67 db) szerepel az enyhén felhős („Partly cloudy”) csoportban. Az előző esetben

a napok megoszlása az egyes csoportok között nagyságrendileg azonos volt. Ez az arányeltolódás valószínűleg a hiányos adathalmaznak köszönhető. A további elemzéshez az egyes csoportok idősorait mutatja az alábbi 17. ábra.



17. ábra: Maafi csoportok – Debrecen

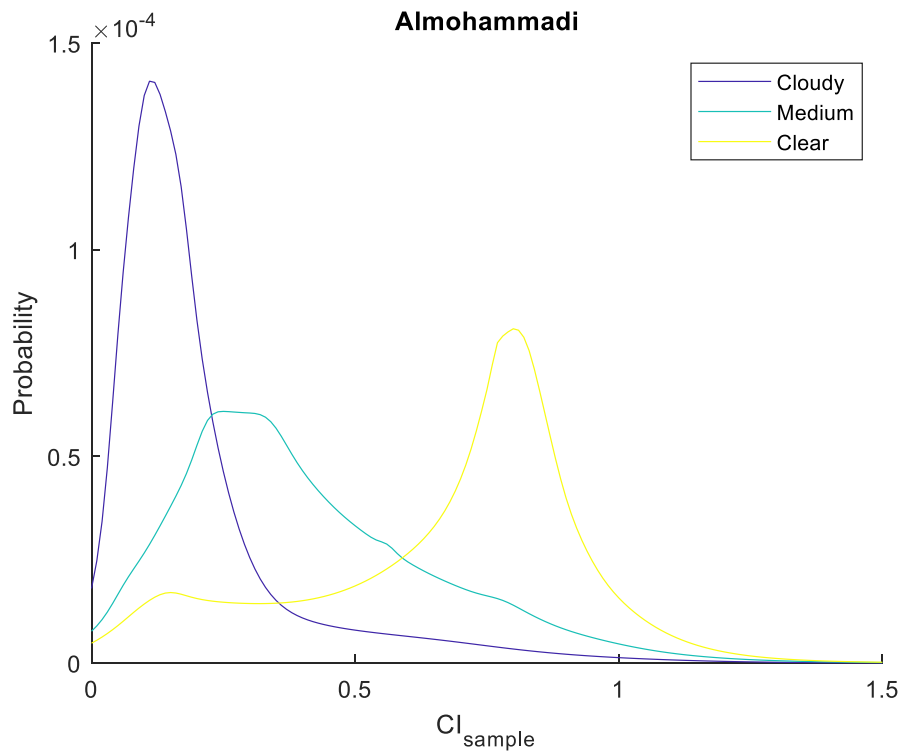
A fenti ábrán megfigyelhető, hogy akár csak a korábban vizsgálat esetében, az első csoportban („Clear sky”) markánsan kirajzolódnak a szabályos, haranggörbe alakú idősorok. Az is megfigyelhető, hogy az utolsó („Cloudy”) csoportot alapvetően az alacsony csúcsertékű besugárzásigörbék dominálják, azonban számos kiugró értékkel rendelkező idősor is szerepel ebben a csoportban. Az előző esettől eltérően, jelen esetben a második („Partly cloudy”) és harmadik („Cloudy”) csoport görbeserege csak kis mértékben hasonló, a három csoport jobban elkülönül egymástól.

### 3.2.2.2 Almohammadi

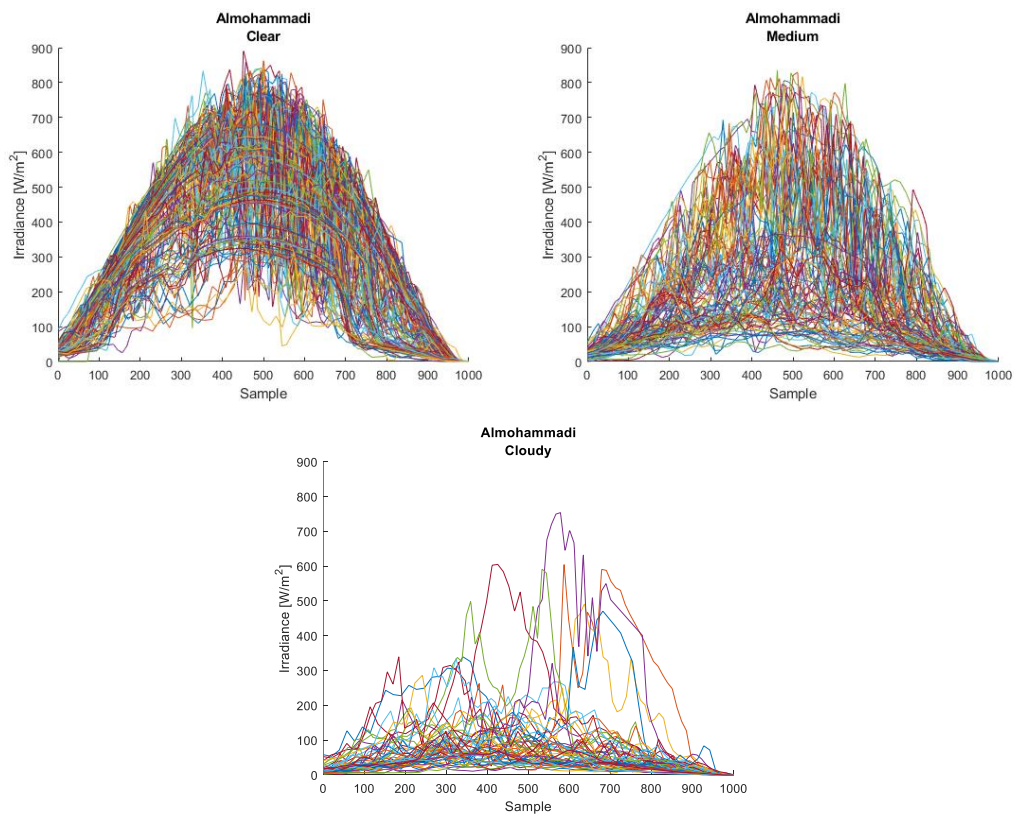
A 3.1.3.5 fejezetben leírt megfontolások alapján jelen adatsorra is elvégeztem a kétmodusú Gauss eloszlások illesztését. Az ismertetett lépéssorozat végrehajtása után kapott eredményeket az alábbi 18. ábra mutatja be. Megfigyelhető, hogy a 3.1.3.5 fejezetben tárgyalt eredményektől eltérően, ebben az esetben nem a középső („Medium”) csoport, hanem az tiszta égboltú napok csoportja („Clear”) jellemezhető kétmodusú eloszlással, és az eloszlás domináns várható értéke is jelentősen kisebb 1-nél



(0,81 értékű). Látható, hogy az EK adatsora alapján kapott eredményekhez képest jelen eloszlásfüggvények összességében balra tolódnak el. Ez a felhős napok dominanciájára utal.



18. ábra: Almohammadi módszer csoportosítási eredmények átlagai – Debrecen



19. ábra: Almohammadi módszer csoportok – Debrecen

A fenti ábrán (19. ábra) az látható, hogy a csoportok görbeseregei igen hasonló összképet mutatnak, mint a korábbi, azonos módszerrel végzett csoportosítás esetén. Megfigyelhető azonban, hogy a felhős napok csoportjának darabszáma igen alacsony. Ez ellentmond az előbb látott Maafi módszer (3.2.2.1) eredményeinek, illetve a 18. ábra kapcsán tárgyalt jobbratolódásnak is. Az ellentmondás valószínűleg annak köszönhető, hogy az adatsor alapvetően nagy számú felhős és kis mennyiségű napos napot tartalmaz. Az Almohammadi módszer által használt k-közép klaszterezés ezt az adatsort három, az algoritmus szerint jól elkülönülő csoportra osztja. Ebből adódik a fent említett jobbratolódása az eloszlásfüggvényeknek is. Megállapítható tehát, hogy ebben az esetben az Almohammadi módszer alkalmazása nem szolgáltat megfelelő eredményeket.

### 3.2.2.3 Eredmények

A fent elvégzett csoportosítások összegzett, számszerű eredményeit tartalmazza az alábbi 8. táblázat.

8. táblázat: Csoportosítási eredmények - Debrecen 2018

Módszer	Csoport elnevezése és darabszáma		
<b>Maafi</b>	Clear-sky	Partly cloudy	Cloudy
	123	67	143
<b>Almohammadi</b>	Clear	Medium	Cloudy
	164	113	56

A fenti 8. táblázat eredményeiben is egyértelműen megfigyelhető a fent leírt ellentmondás.



### 3.3 Eredmények összehasonlítása

A 3.1.3.2 és 3.2.2.1, illetve a 3.1.3.5 és 3.2.2.2 fejezetekben elvégzett csoportosítások eredményeinek összehasonlítását tartalmazza az alábbi 9. táblázat.

9. táblázat: Csoportosítási redmények összehasonlítása: EK 2018 - Debrecen 2018

Módszer	Csoport	Debrecen		EK	
		Napok száma [db]	Napok száma [%]	Napok száma [db]	Napok száma [%]
Maafi	Clear	123	36,94	163	44,66
	Partly-cloudy	67	20,12	100	27,40
	Cloudy	143	42,94	102	27,95
Almohammadi	Clear	164	49,25	49	13,42
	Medium	113	33,93	206	56,44
	Cloudy	56	16,82	110	30,14

A fenti táblázatból jól látható, hogy a két adatsor csoportosítási eredményei között jelentős eltérések vannak. Ez a főként földrajzi helyszínek eltérő időjárási körülményeinek köszönhető. Azonban, amint az előzőekben láttuk, az Almohammadi módszer esetén a módszer hibás működése is nagymértékű eltéréseket okoz.

## 4 Módszerek értékelése és összehasonlítása

---

### 4.1 Csoportosítási egyezések elemzése

A csoportosítások elvégzése után kísérletet tettem annak megállapítására, hogy a kapott eredmények elfogadhatók-e, megfelelnek-e a valóságnak. Ennek érdekében a 2018-as évi, EK-s mérési adatokra alkalmazott módszerek eredményeit összehasonlítottam egymással. (A debreceni adatokat azért hagytam figyelmen kívül, mert azok nem fedtek le egy teljes évet, illetve csak két módszer alkalmazását tették lehetővé.) Mivel az eljárások által alkalmazott csoportok száma igen eltérő, csak a szélső csoportokra (tisztá égboltú - és felhős napok) koncentráltam, melyek minden módszer esetén megtalálhatók. Így vizsgáltam a csoportosítások eredményeinek egyezését. Csak azokat az eseteket tekintettem értékelhetőnek, melyek során legalább négy módszer (a hatból) egyezést mutatott. Ezen esetek mentén igyekeztem következtetéseket levonni az egyes eljárások működésének helyességéről. A vizsgálatok eredményei az alábbiakban olvashatók.

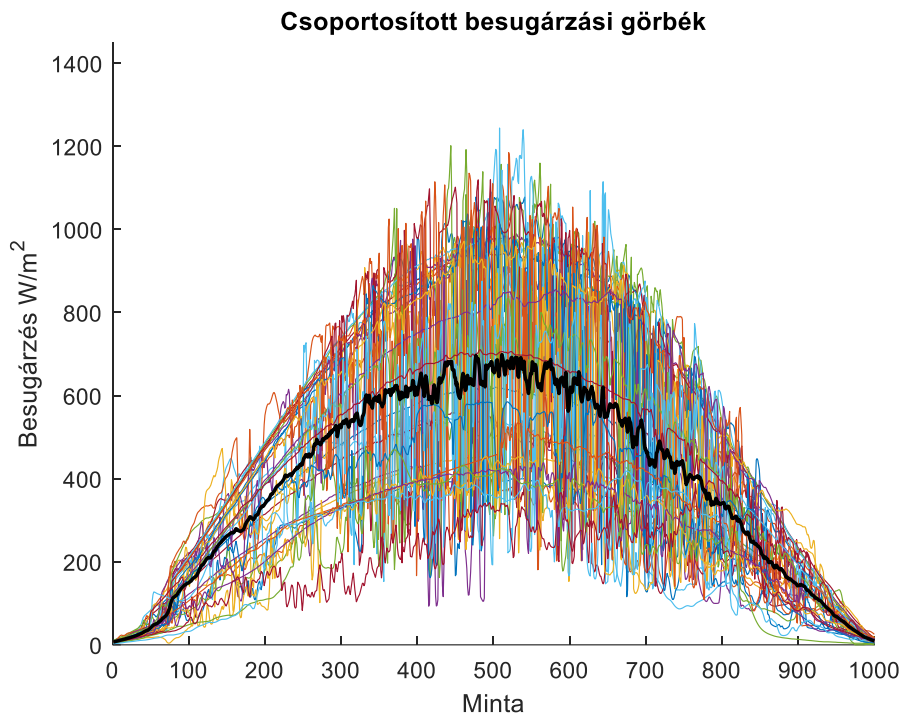
#### 4.1.1 Tiszta égboltú napok

Elsőként a tiszta égboltú csoportba sorolt napokat vizsgáltam meg. Arra kerestem a magyarázatot, hogy miért éppen ezek a módszerek értékelték azonosan az adott napok csoportját, illetve a kimaradt módszerek milyen okból ítélték másképp.

##### 4.1.1.1 4 találatot kapott napok

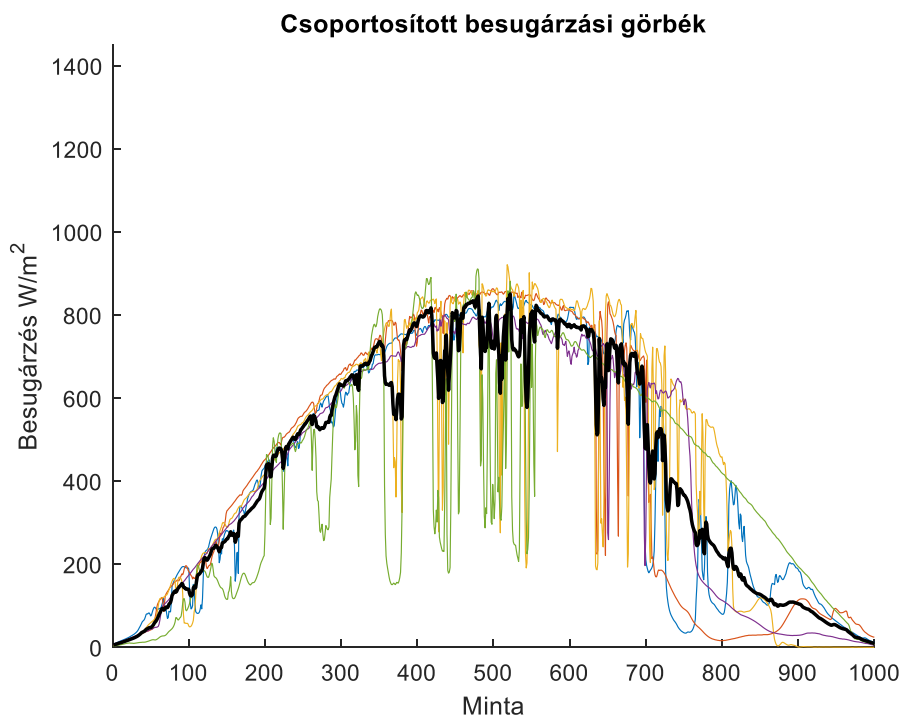
Azon napok száma, melyeket pontosan négy módszer sorolt a tiszta égboltú napok sorába, a vizsgált esetben 43 db volt. Ez alapján tehát az év 11,78%-ában napos volt az idő. Ez első benyomásra reális eredménynek tűnik.

Az említett 43 esetből 35 esetben az Almohammadi módszer nem mutatott találatot, míg a Trueblood módszer eredményei egyetlen esetben sem egyeztek meg a többi módszer eredményeivel. Ennek a 35 napnak a besugárzási görbéi, illetve az ezekből számított átlaggörbe (fekete) láthatók az alábbi ábrán (20. ábra). Megfigyelhető, hogy az esetek meghatározó részében a besugárzási görbék igen változékonyak, így ezek a napok a Trueblood módszer esetében egyértelműen nem lehetnek „Clear” besorolásúak, tehát az egyezés hiánya jogos. Ezen kívül a besugárzás mértéke ezeknél a napoknál igen alacsony értékeket is elér, így az Almohammadi módszer tiszta égboltra jellemző egymodusú eloszlásától eltérnek. Tehát ez a besorolás is helyesnek értékelhető.



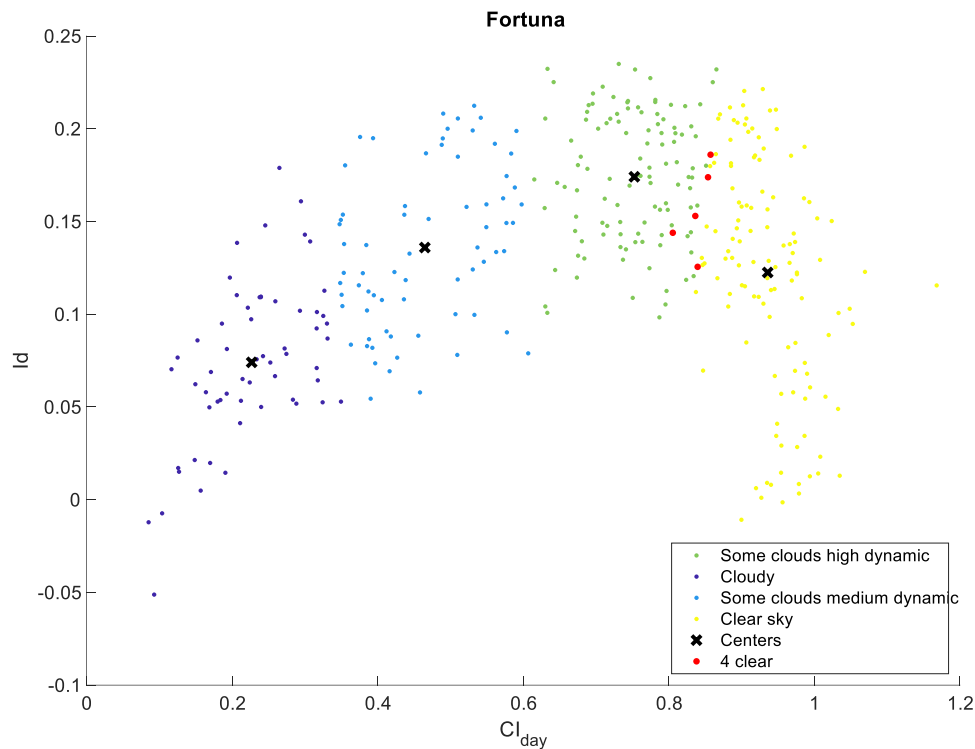
**20. ábra: Trueblood és Almohammadi módszer szerint nem tiszta égboltú napok**

A maradék 8 esetből 5 esetben a Fortuna, 3 esetben pedig a Maafi módszerek maradnak el a Trueblood módszerrel együtt. A Fortuna módszer által nem tiszta égboltúnak minősített napok az alábbi ábrán (21. ábra) láthatók. Az átlagértékből (fekete) jól látszik, hogy ezekre a napokra viszonylag magas besugárzási értékek jellemzők, valamint erős fluktuáció legalább a nap ötödében.



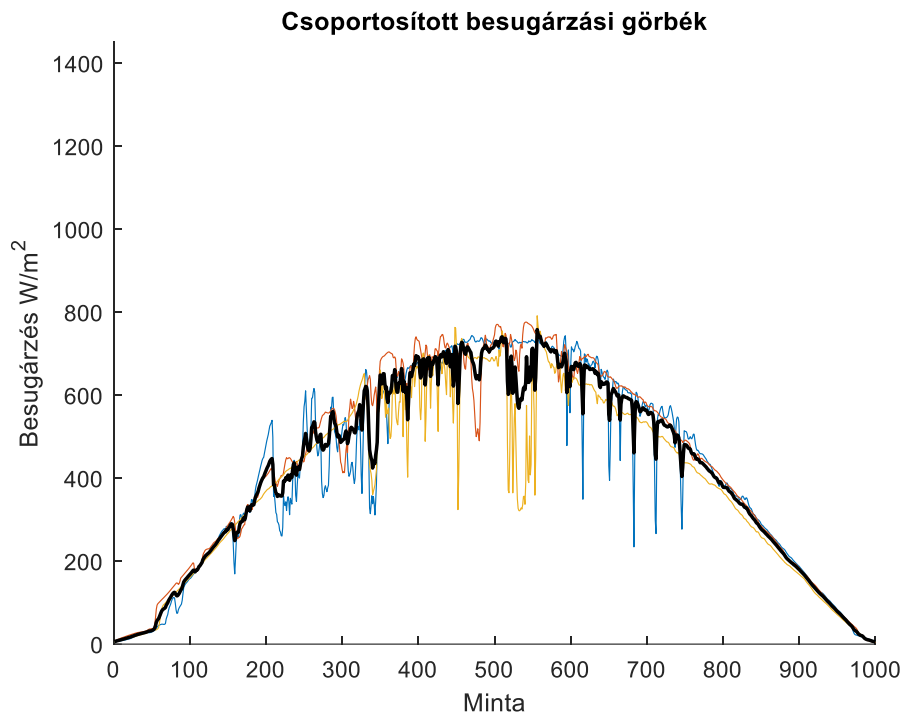
**21. ábra: Trueblood és Fortuna módszer szerint nem tiszta égboltú napok**

Ezen napokat a  $Cl_{day} - I(d)$  síkon megjelenítve (22. ábra) megfigyelhető, hogy a k-közép klaszterezés éppen a „Clear sky” és a „Some clouds, high dynamic” csoportok határára sorolta őket. Az is látszik azonban, hogy több, hasonló paraméterekkel rendelkező nap a „Clear” csoportba került, ahogy a többi eljárás esetén is. Így a Fortuna módszer besorolása nem tekinthető hibásnak.

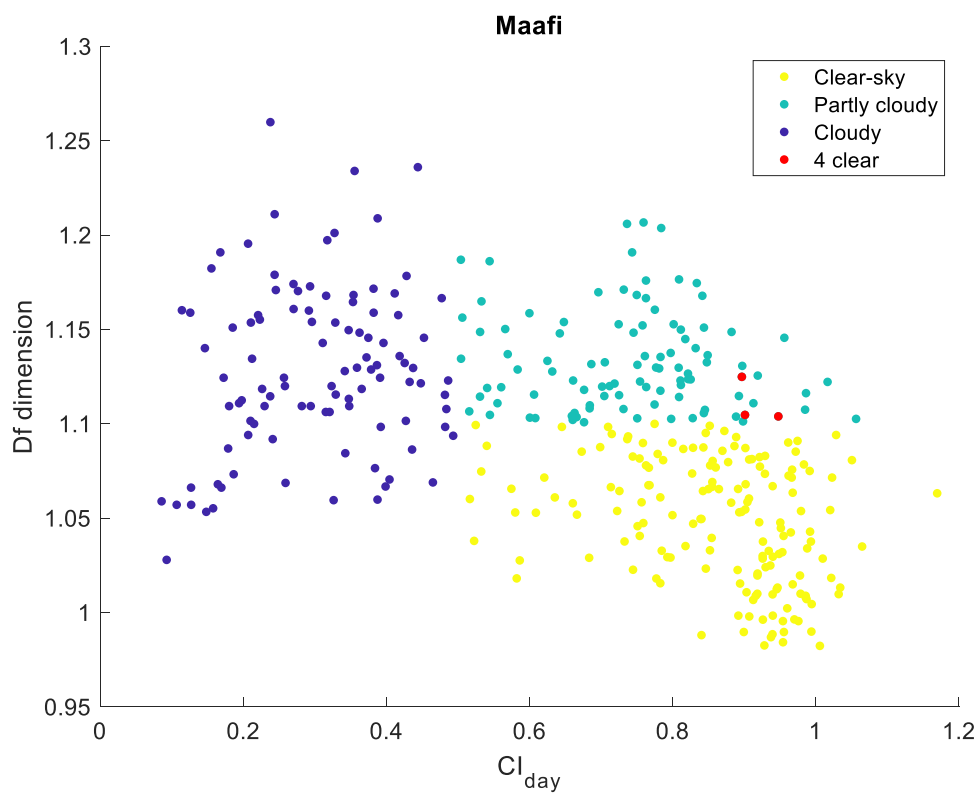


22. ábra: Fortuna módszer szerint nem tiszta égboltú, 4 találatot kapott napok kiemelése

A Maafi módszer szerint nem tiszta égboltú napok idősorait a 23. ábra mutatja. Ez a módszer tízperces felbontású adatsorokkal dolgozik. A besugárzási görbéken látható, hogy még ilyen durva felbontás esetén is könnyen észlelhetők az adott napokon bekövetkezett jelentős fluktuációk. Ezeket a napokat is megjelenítettem a  $Cl_{day} - D(d)$  diagramon (24. ábra). Ebben az esetben is azt tapasztaltam, hogy a vizsgált napok éppen csoporthatárra esnek. Azonban itt is az figyelhető meg, hogy számos, a kiemelt pontokéhoz hasonló paraméterű pont viszont „Clear” kategóriába került. Így a Maafi módszer működése is megfelelőnek mondható, annak jelentősebb utólagos korrigálására nincs szükség.



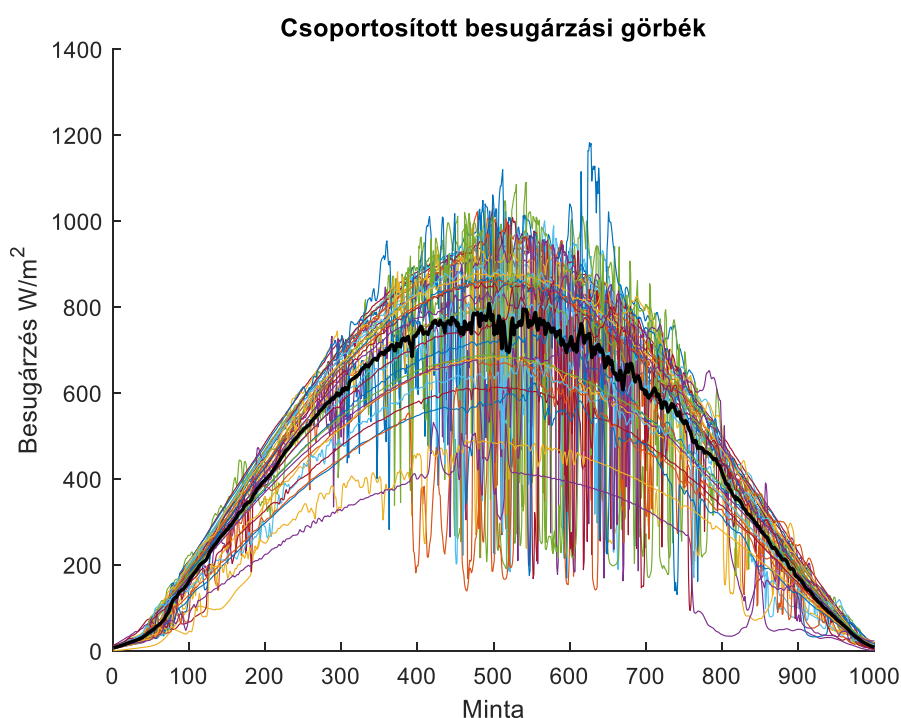
23. ábra: Trueblood és Maafi módszer szerint nem tiszta égboltú napok



24. ábra: Maafi módszer szerint nem tiszta égboltú, 4 találatot kapott napok kiemelése

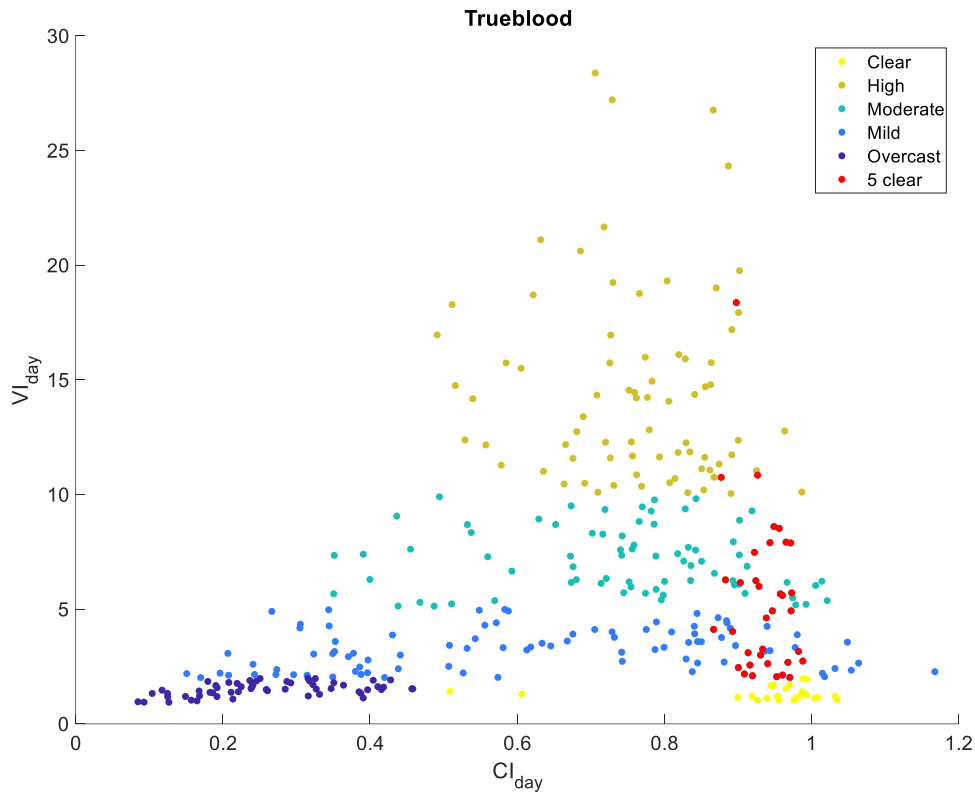
#### 4.1.1.2 5 találatot kapott napok

Az öt módszer által tiszta égboltú napnak minősített napok száma a vizsgált esetben 39 darabra (10,68%) adódott. Ezen napok besugárzási görbéi láthatók az alábbi ábrán (25. ábra). Az idősoros adatokon megfigyelhető, hogy ebben az esetben már markánsabban kirajzolódnak a haranggörbe alakú besugárzási görbék és az átlaggörbe (fekete) is szabályosabb, mint a 4 találatot kapott napok esetében. Azonban továbbra is jelentős, főként a napok középső részére koncentrálódó, fluktuációk figyelhetők meg. Jelen esetben a találatok megoszlása úgy alakult, hogy 35 esetben a Trueblood, 4 esetben pedig az Almohammadi módszer eredményei nem egyeztek meg a másik öt eljárás eredményeivel.



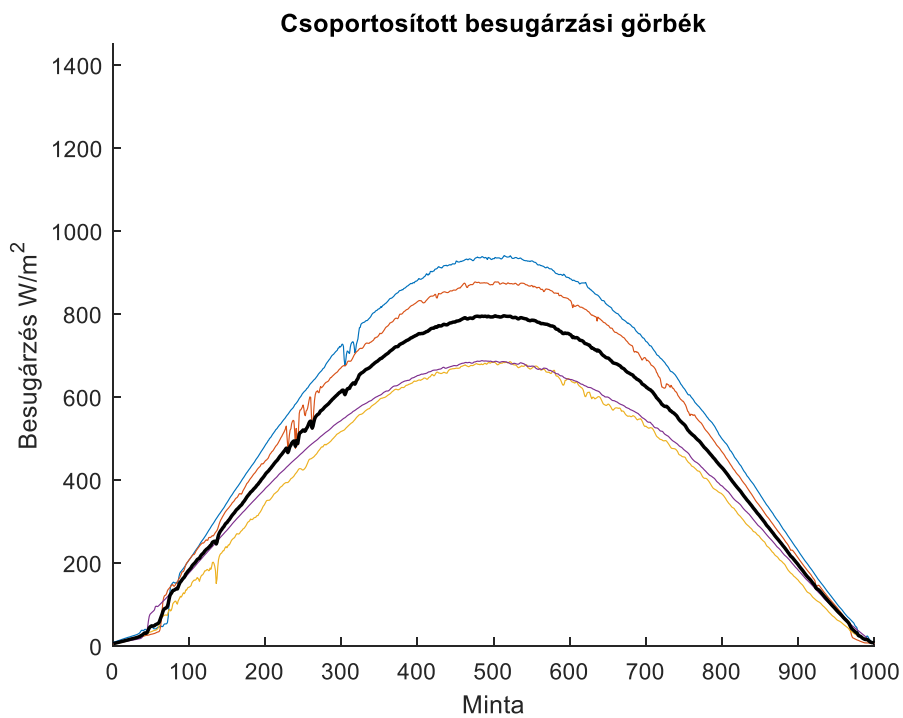
**25. ábra: 5 módszer által tiszta égboltúnak azonosított napok besugárzási viszonyai**

Mind a 25, a Trueblood módszer által nem tiszta égboltú napként azonosított napra magas összesített besugárzás jellemző, a  $CI_{day}$  indexek mind 0,86 feletti értékűek, így az eltérés kizárólag a variabilitási indexre vezethető vissza. A vizsgált napokat a  $CI_{day}-VI_{day}$  síkon kiemelve (26. ábra) megfigyelhető, hogy azok nem egy csoportba tömörülnek, hanem a szomszédos kategóriákon kívül is megtalálhatók. Három csoport (High, Moderate, Mild) között oszlanak meg. A módszer determinisztikus jellege miatt, csak úgy lehetne növelni a találatok arányát, ha a csoportok számát csökkentenénk, azaz, ha a viszonylag alacsony variabilitási indexekkel jellemezhető kategória (Mild) elemeit megosztanánk a két szomszédos (Moderate és Clear) kategóriák között.



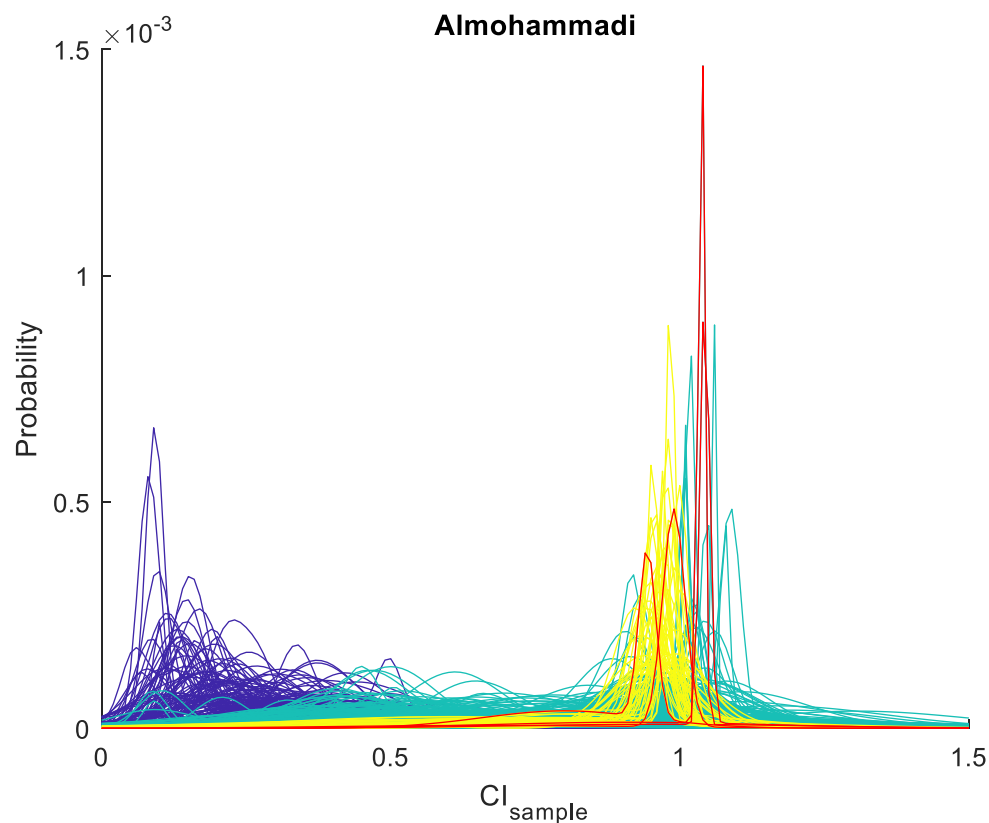
26. ábra: Trueblood módszer szerint nem tiszta égboltú, 5 találatot kapott napok kiemelése

Az Almohammadi módszer által nem tiszta égboltúnak minősített napok esetében az idősorok (27. ábra) alapján egyértelműen hibás besorolásról van szó. A hibás besorolás okainak vizsgálatához kiemeltem a besorolás során kialakított kevert Gauss modell eloszlásfüggvényeit (28. ábra).



27. ábra: Almohammadi módszer szerint nem tiszta égboltú napok sugárzási viszonyai

A 28. ábra szerint a helytelen besorolású napok várhatóértékei közel 1 értékűek (0,94 – 1,04), tehát a mért besugárzási értékek kis mértékben magasabbak, illetve alacsonyabbak voltak a tiszta égboltú viszonyokat jellemző elméleti értéknél. Az egyik, kiemelkedően magas súlyú, 1,04 várható értékű modussal bíró eloszlásfüggvény hibás besorolása a klaszterezési módszer (k-közép) jellegéből adódik.



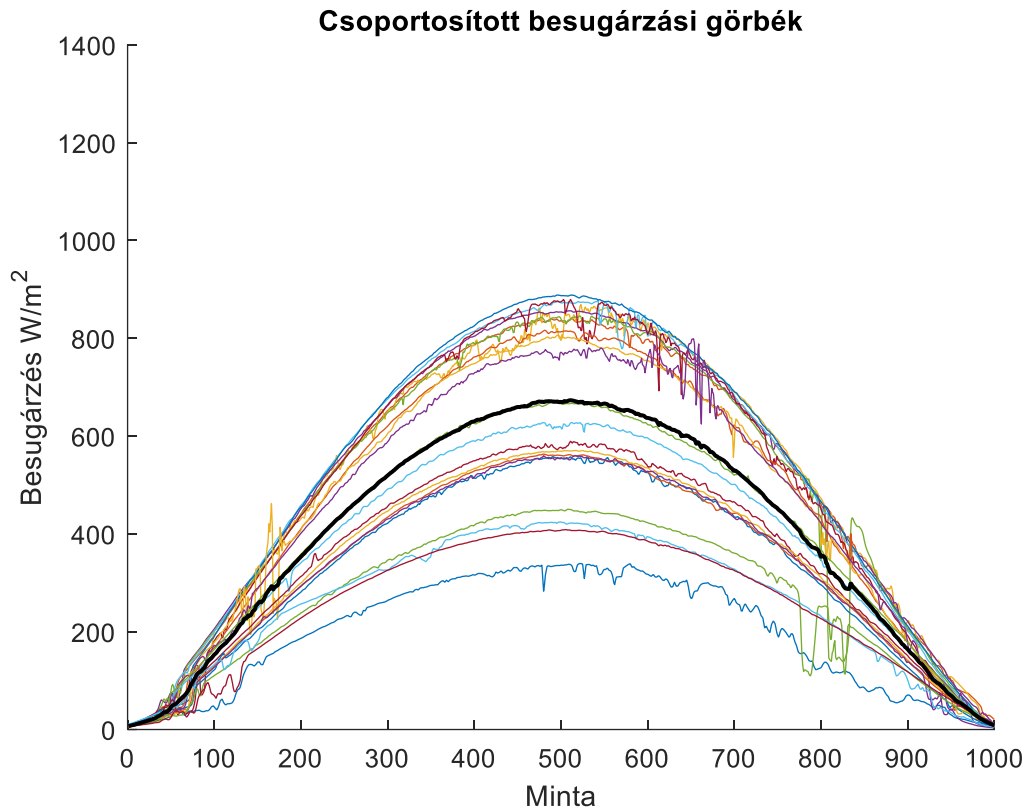
28. ábra: Almohammadi módszer szerint nem tiszta égboltú, 5 találatot kapott napok kiemelése

Nagy valószínűséggel a másik három esetben is ugyan ez a hiba oka. Tehát a módszer összességében nem nevezhető hibásnak.

#### 4.1.1.3 6 találatot kapott napok

A vizsgált év 365 napjából mindössze 21 napot (5,75%) sorolt mind a 6 módszer egyértelműen tiszta égboltúnak. Ez viszonylag alacsony arány ahhoz képest, hogy a módszerek külön-külön átlagosan 106 (29,04%) napot soroltak ebbe a kategóriába (ami viszont viszonylag magas aránynak tekinthető ahhoz képest, hogy 6 módszerből 4 legalább 4 csoportot alkalmaz). Ezen 21 nap besugárzási görbéit mutatja az alábbi 29. ábra. Jól látható, hogy a besugárzási görbék szinte kivétel nélkül haranggörbe alakúak, csak kismértékű, gyors felhőátvonulások zavarják meg a besugárzást. Megfigyelhető, hogy az átlaggörbe (fekete) majdnem tökéletes haranggörbe. Az is észrevehető azonban, hogy az átlagos besugárzás mértéke (az átlaggörbe maximuma) mindössze  $670 \text{ W/m}^2$ .





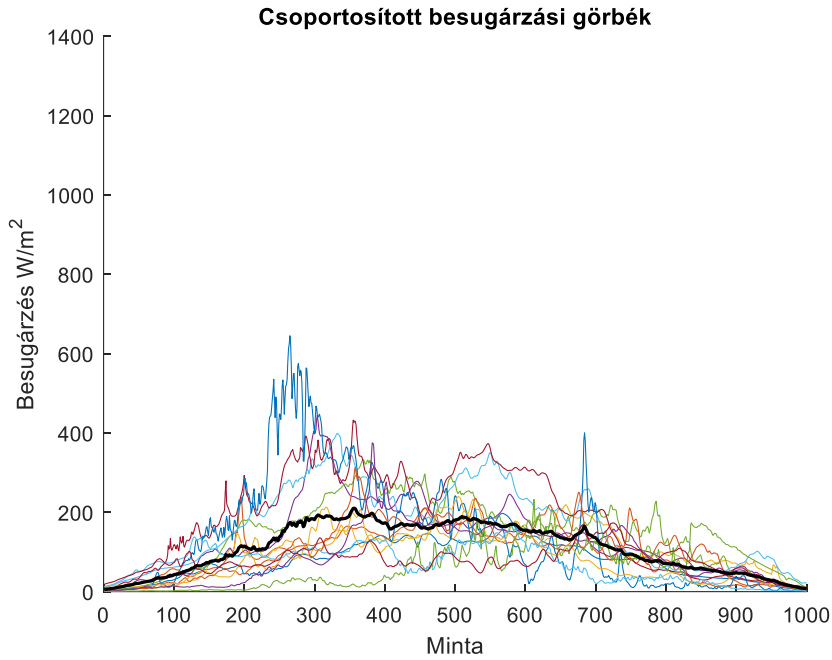
29. ábra: 6 módszer által tiszta égboltúnak azonosított napok besugárzási viszonyai

## 4.1.2 Felhős napok

A tiszta égboltú napok vizsgálata után az egyes módszerek által felhősnek ítélt napokat vizsgáltam meg, az előző esethez hasonló megfontolások alapján.

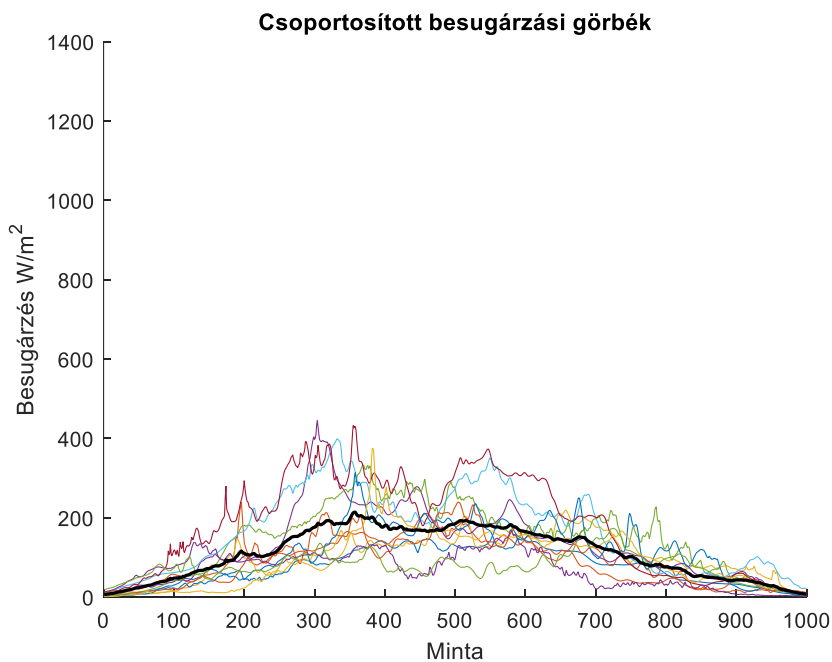
### 4.1.2.1 4 találatot kapott napok

Az általam vizsgált adatsorban 14 db (3,84%) olyan nap szerepel, melyeket pontosan négy csoportosítási módszer ítélt felhősnek. Ezek idősorai az alábbi ábrán (30. ábra) láthatók. A Chicco1 módszer ezek egyikét sem sorolta a felhős napok közé, míg a Maafi, Chicco1 és Almohammadi módszerek mindegyiket ide sorolták. Ezen napok közül a Fortuna módszer 12, a Trueblood módszer pedig 2 darabot nem ítélt felhősnek.

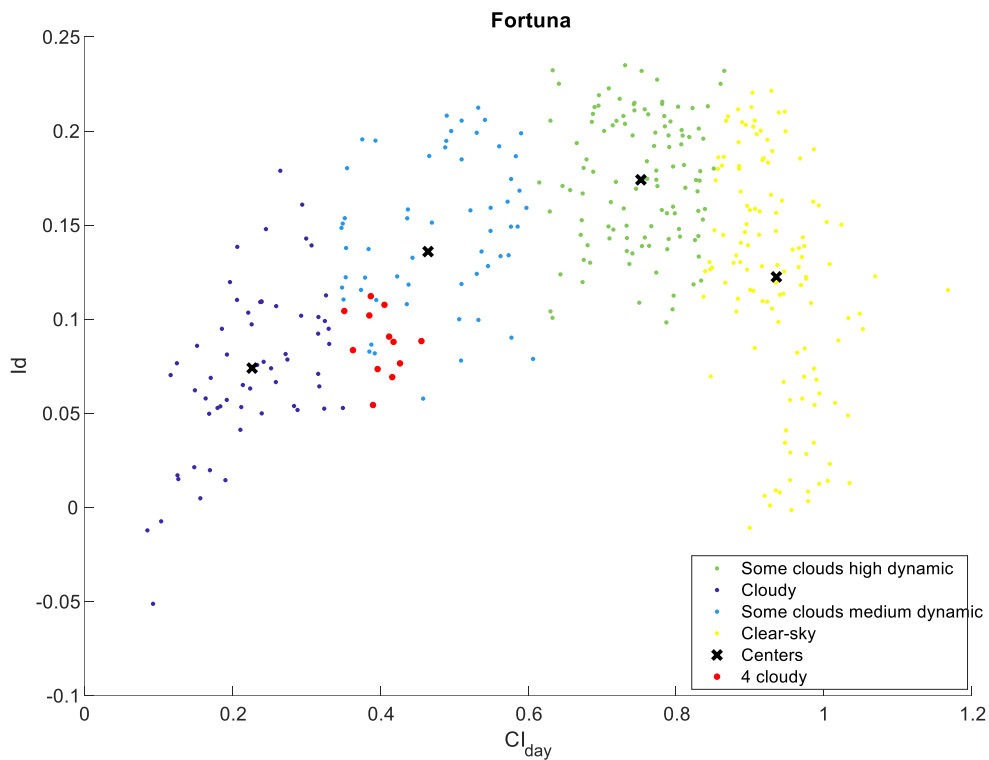


30. ábra: 4 módszer által felhősként azonosított napok besugárzási viszonyai

A Fortuna módszer által nem felhősként azonosított 12 napra (31. ábra), ahogy azt az átlaggörbe (fekete) is mutatja, viszonylag alacsony besugárzási értékek jellemzők és a variabilitás mértéke is viszonylag kicsi. Ezeket a napokat a  $Cl_{day} - I(d)$  síkon ábrázolva (32. ábra) látható, hogy a korábbi esetekhez hasonlóan, most is a csoporthatár közelében helyezkednek el a pontok. Mivel a csoportosítás nem determinisztikus módon, hanem k-közép klaszterezés segítségével történik, az eltérés nem a módszer gyengeségéből fakad, így annak működése nem tekinthető hibásnak.

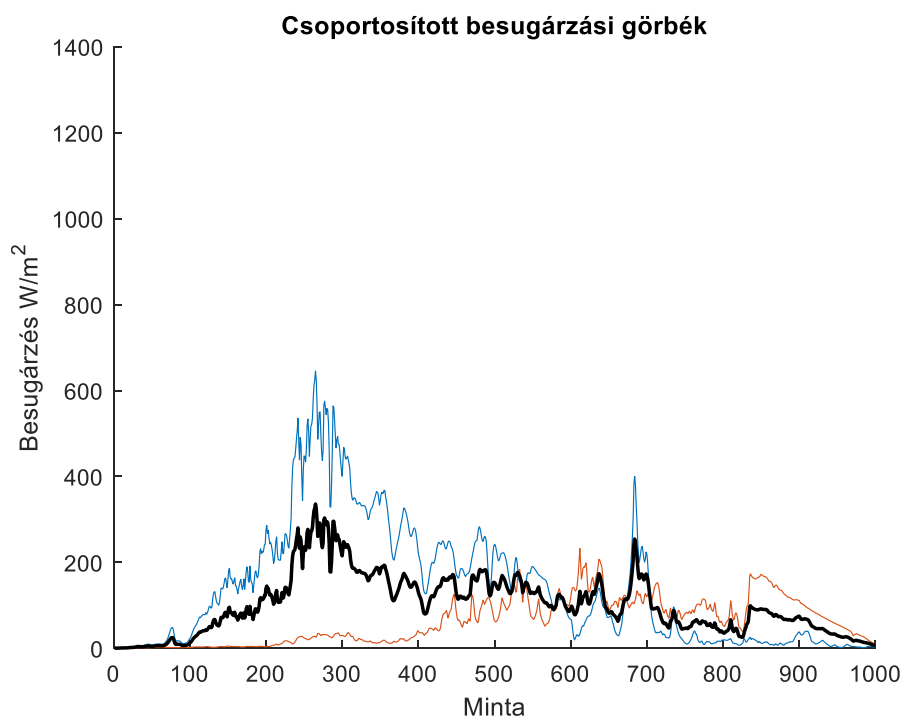


31. ábra: Fortuna módszer szerint nem felhős napok



**32. ábra: Fortuna módszer szerint nem felhős, 4 találatot kapott napok kiemelése**

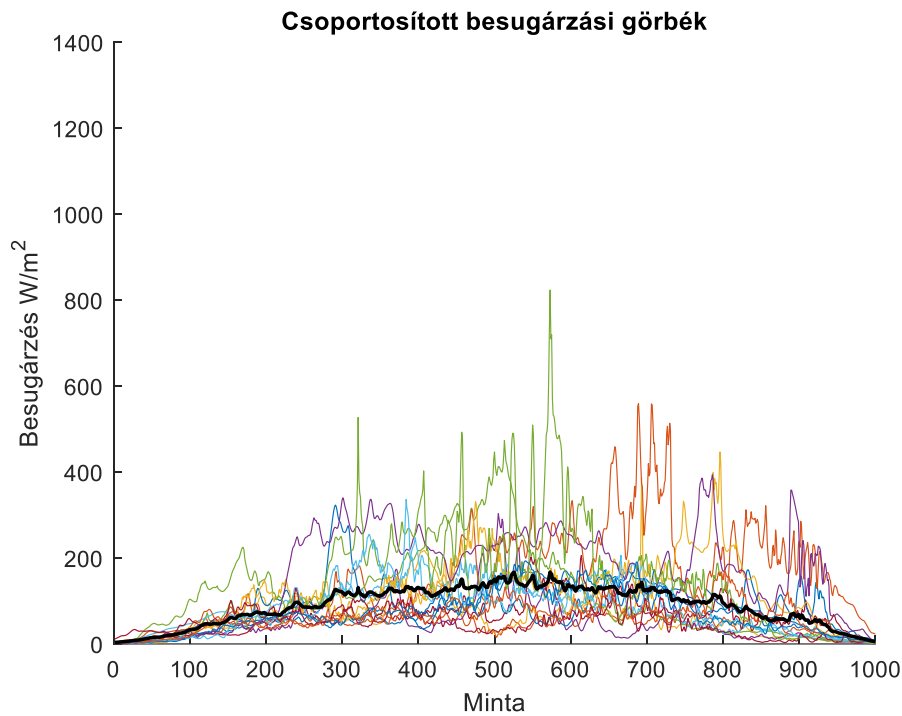
A Trueblood módszer esetén, ahogy korábban is, a magas variabilitás miatt kerültek más kategóriába (Mild) ezek a napok (33. ábra).



**33. ábra: Trueblood módszer szerint nem felhős napok**

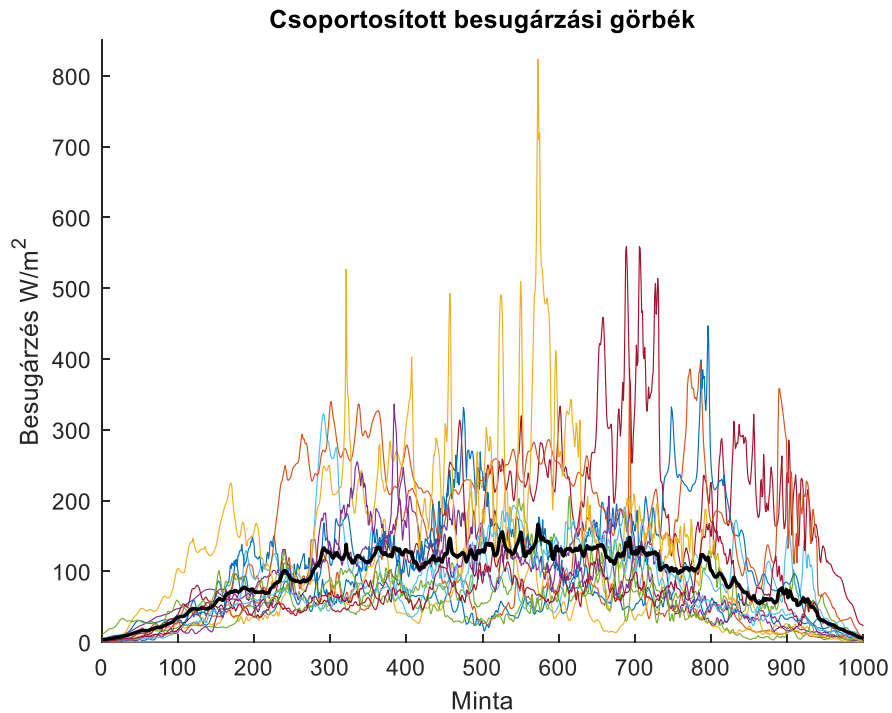
#### 4.1.2.2 5 találatot kapott napok

A pontosan öt módszer által felhősnek ítélt napok száma mindössze 16 db (4,38%) volt a vizsgált időszakban. Minden esetben vagy a Trueblood (14 eset) vagy a Chicco2 (2 eset) módszer mutatott csak eltérést. Ezen napok besugárzási viszonyait mutatja a 34. ábra.

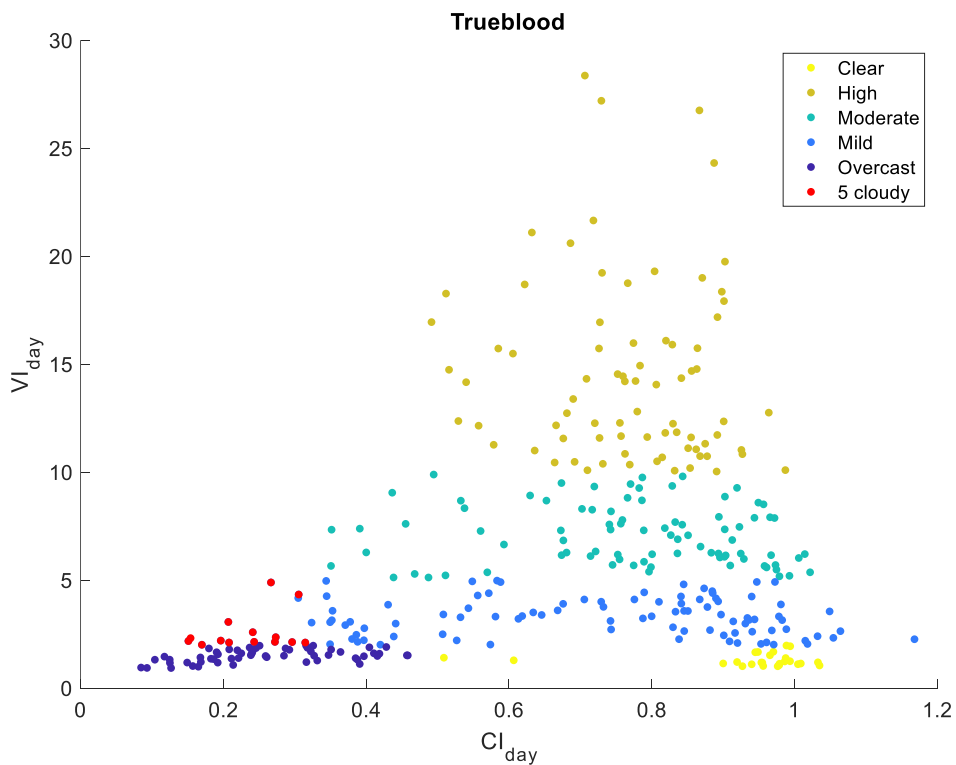


**34. ábra: 5 módszer által felhősként azonosított napok besugárzási viszonyai**

A Trueblood módszerrel nem felhősként azonosított napok besugárzásait (35. ábra) és azok pozícióit (36. ábra) megjelenítve, megvizsgálhatjuk az eltérések okát. Ahogy az összes korábbi esetben, jelen esetben is a variabilitási index alacsony határértéke miatt kerültek ezek az egyébként felhős napok „Mild” kategóriába.



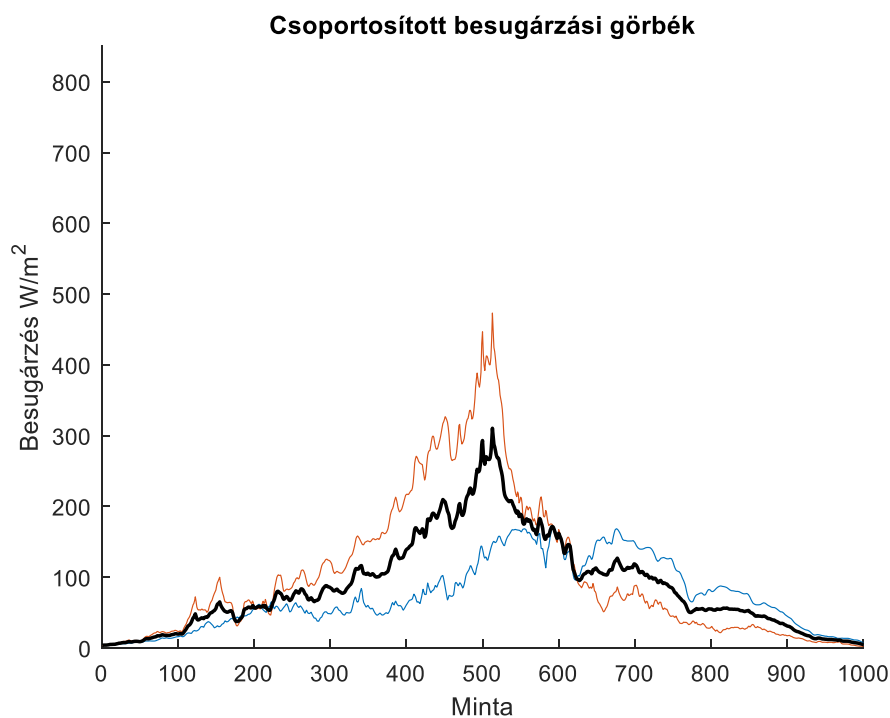
35. ábra: Trueblood módszer szerint nem felhős napok



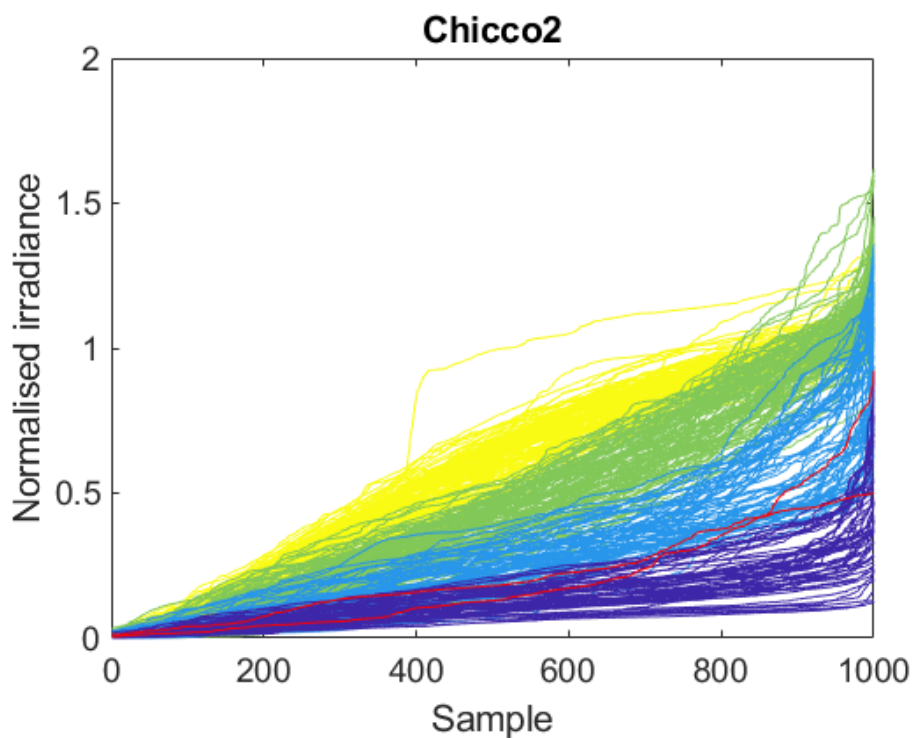
36. ábra: Trueblood módszer szerint nem felhős, 5 találatot kapott napok kiemélése

A Chicco2 módszer által nem felhősként azonosított napok besugárzási görbéit megjelenítve (37. ábra) azt láthatjuk, hogy mindkét esetben alacsony besugárzási értékekkel rendelkező napról van szó, azonban rövid idejű, kiugró értékek is előfordulnak. A Chicco2 módszer újramintavételezett, normált és eloszlásfüggvény-szerűen rendezett adatsorral dolgozik, így a tiszta égbolt értékeket gyakorlatilag

függetlenné teszi az időbeliségtől. Ez azonban éppen a fent említett kiugró értékek esetén jelentősen torzíthatja az eloszlást. Ilyen torzult eloszlás esetén pedig a nem determinisztikus (k-közép) klaszterezési algoritmus hibás besoroláshoz vezethet. Ennek alátámasztása érdekében ábrázoltam a Chicco2 módszer besorolási eredményeit (38. ábra), kiemelve a vizsgált eseteket.



37. ábra: Chicco2 módszer szerint nem felhős napok

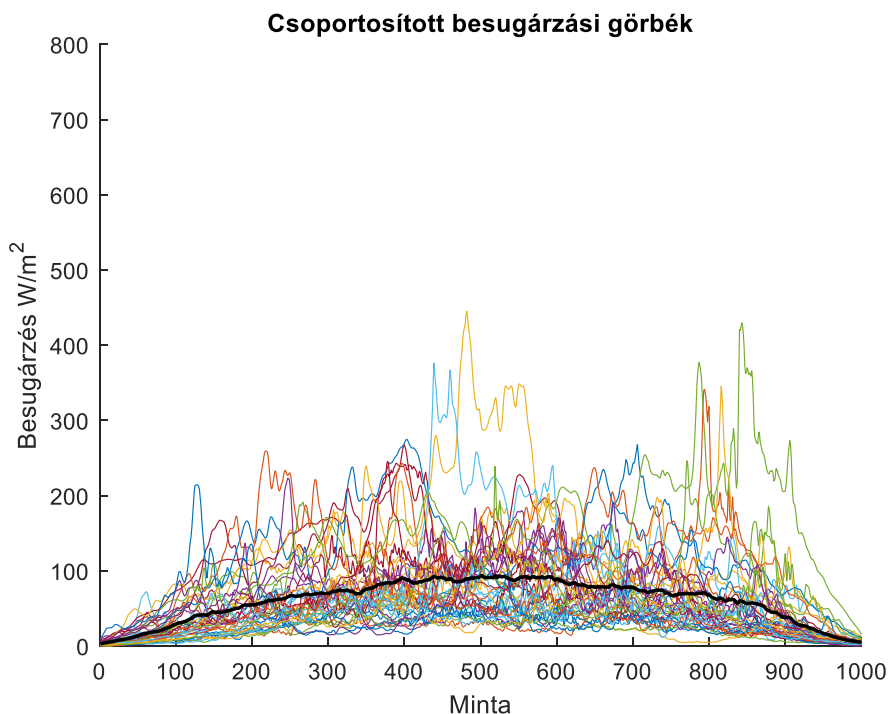


38. ábra: Chicco2 módszer szerint nem felhős, 5 találatot kapott napok kiemelése

Látható, hogy a módszer a szomszédos, „Quasi-cloudy” csoportba sorolta a napokat, tehát a hiba mértéke nem jelentős.

#### 4.1.2.3 6 találatot kapott napok

Végül azokat a napokat vizsgáltam meg, melyeket mind a hat módszer egyaránt felhősnek minősített. A vizsgált 365 napból ezek száma 41 darabra (11,23%) adódott. Ezen napok idősorai és átlaguk (fekete görbe) láthatók az alábbi ábrán (39. ábra).



39. ábra: 6 módszer által felhősnek azonosított napok besugárzási viszonyai

A fenti ábrán megfigyelhető, hogy ezekre a napokra mérsékelt variabilitás és igen alacsony átlagos besugárzási érték jellemző (az átlaggörbe maximuma  $93 \text{ W/m}^2$ ).

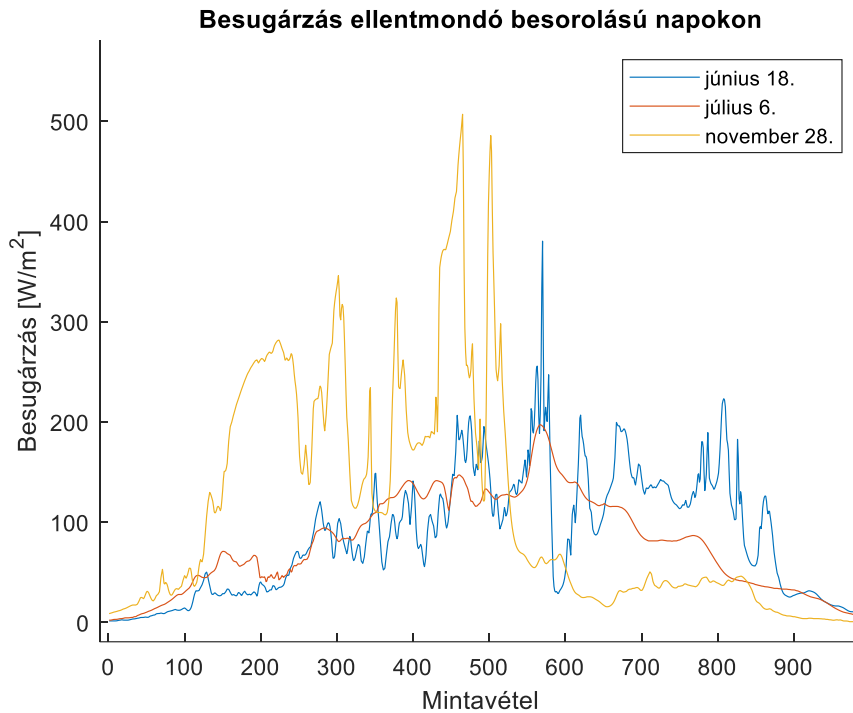
#### 4.1.3 Ellentmondásos besorolás

A napok csoportosítása során előfordult, hogy a különböző módszerek egymásnak ellentmondó kategóriába soroltak be néhány napot. Ezen esetek mentén a metódusok alkalmazhatóságának korlátjaira kaphatunk rálátást. Két ilyen lehetőséget vizsgáltam meg.

##### 4.1.3.1 Egyidejű tiszta égboltú és felhős azonosítás

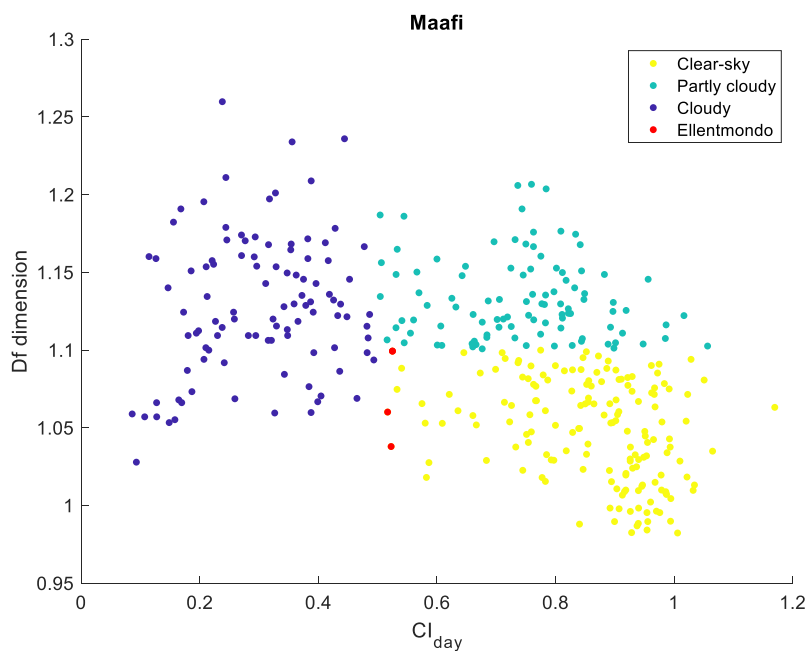
Az általam vizsgált évben 3 darab olyan nap (június 18., július 6. és november 28.) fordult elő, melyeket két módszer ellentétesen értékelt. Minden esetben a Maafi módszer sorolta tiszta égboltú kategóriába a napokat. Az első esetben a felhős csoportba soroló módszer a Chicco1 módszer volt, míg a második és harmadik esetben az Almohammadi féle módszer. Az ellentmondást okozó napok idősorait a 40. ábra mutatja.





40. ábra: Egyidejűleg tiszta égboltúként és felhősként azonosított napokk besugárzási viszonyai

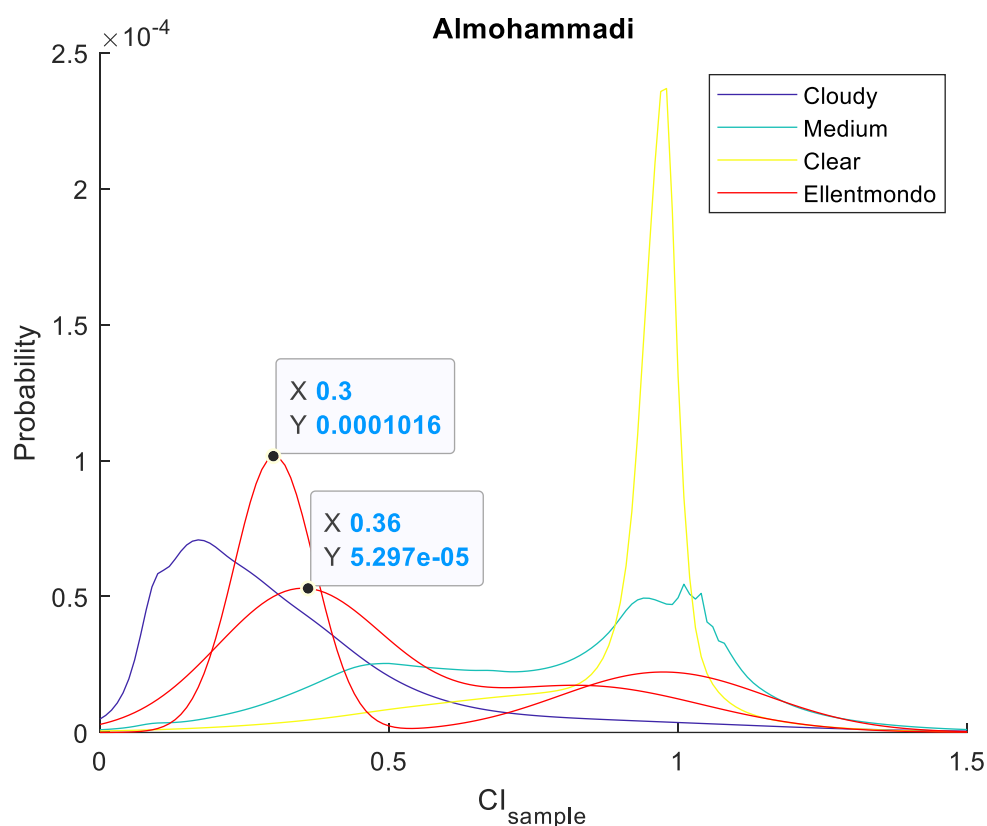
A fenti ábrán jól látható, hogy az ellentmondást okozó napokra igen alacsony besugárzási értékek jellemzők. A három nap közül kettő nyári, a nyári napfordulóhoz közeli nap, a harmadik pedig egy késő őszi, a téli napfordulóhoz közeli nap. Mind a három esetben egy determinisztikus (Maafi) és egy nem determinisztikus módszer (Chicco1 és Almohammadi) mond ellent egymásnak. Ezen napok kiemelve a  $Cl_{day}-D(d)$  síkon az alábbi ábrán (41. ábra) láthatók. Megfigyelhető, hogy mindhárom adatpont a felhős csoport határa környékén helyezkedik el.



41. ábra: Maafi módszer szerint tiszta égboltú, ellentmondásos napok kiemelése

Június 18. naphoz tartozó adatpont éppen a Maafi módszer három csoportjának közös határán helyezkedik el, így valószínűsíthető, hogy minimálisan nagyobb fluktuáció vagy kisebb tiszta égbolt index érték esetén a besorolás megváltozna. A Chicco1 módszer eredményét vizsgálva ismét a k-közép klaszterezés sajátosságával találkozhatunk.

Július 6. és november 28. napokhoz tartozó adatpontok szintén a Maafi módszer csoporthatárain helyezkednek el. Az Almohammadi módszer ezekre a napokra kétmodusú eloszlásfüggvényt illesztett (42. ábra), mely inkább a „Mild” csoportra jellemző. Ezen eloszlásfüggvények esetében azonban az alacsonyabb várhatóértékű (0,3 és 0,36) modulusok a nagyobb súlyúak. Ezek a várható értékek ráadásul éppen a felhős és átmeneti napok első modulusainak várható értékei közötti értékek.

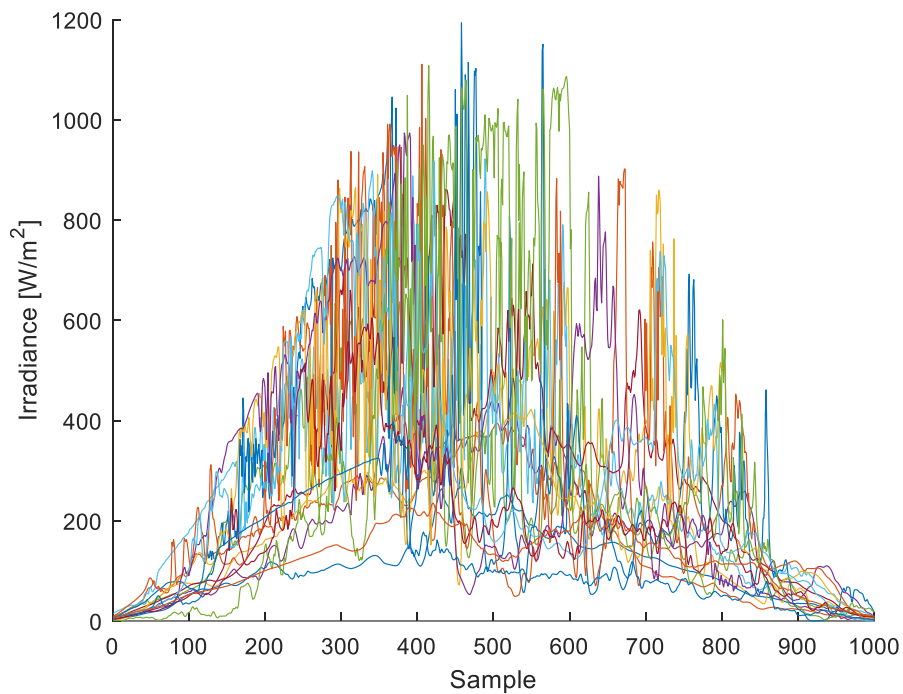


42. ábra: Almohammadi módszer szerint tiszta égboltú, ellentmondásos napok kiemelése

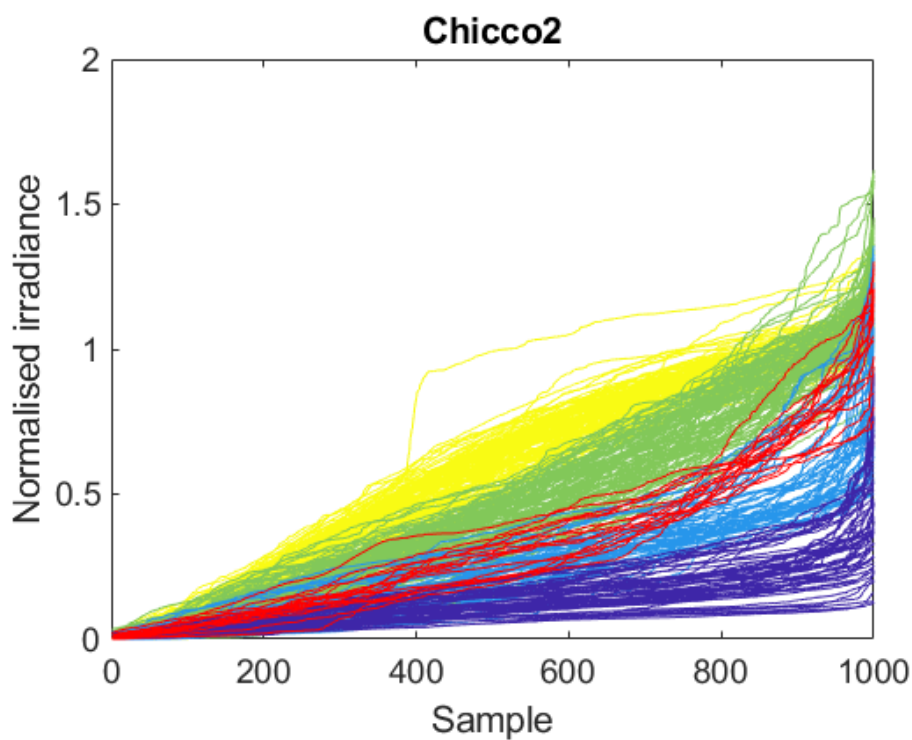
#### 4.1.3.2 Egyidejű alacsony és magas dinamikájú azonosítás

A másik olyan csoport páros, ahol az ellentmondások szembeűnők, viszonylag jól definiálhatók lehetnek, a Chicco1 és Chicco2 módszer által alkalmazott „Quasi-clear” és „Quasi-cloudy” besorolások. A vizsgált időszakban a kategóriákba sorolt elemek számában ugyan nem volt jelentős eltérés, azonban jobban megvizsgálva összesen 45 olyan elemet detektáltam, melyek eltérő besorolást kaptak. Ezen napok közül mindössze 16 db olyan volt, melyeket a Chicco1 (idősoros) ítél „Quasi-clear” -nek és a Chicco2 (rendezett) „Quasi-cloudy” -nak. Ezen napok idősorai láthatók az alábbi ábrán (43. ábra). Az idősorok alapján azt feltételezhetjük, hogy az eltérő besorolást a napok körülbelül 40-50%-án át tartó

alacsony variabilitású időszakok okozzák. Ez a jelleg inkább egy tiszta égboltú nap besugárzási görbéjére hasonlít, azonban az adatokat rendezve, egy torzabb, jobb oldal irányába eltolódott eloszlást kapunk (44. ábra), melyet a k-közép klaszterezés inkább felhősnek tekint.



43. ábra: Chicco1 által „Quasi-clear”-nek, Chicco2 (rendezett) által „Quasi-cloudy”-nak ítélt napok besugárzási viszonyai



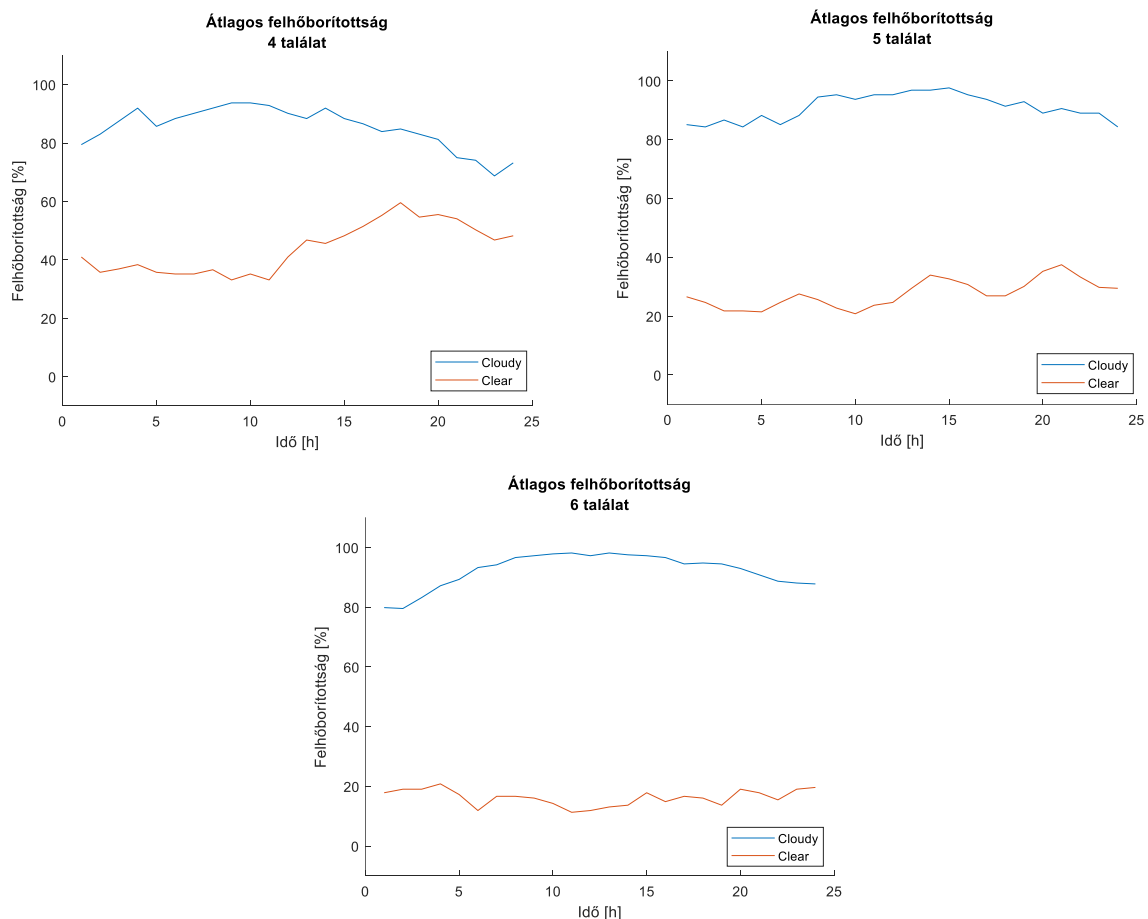
44. ábra: Chicco2 módszer szerint „Quasi-cloudy”, ellentmondásos napok kiemelése

#### 4.1.4 Felhőborítottság

A különböző módszerek alkalmazásával kapott csoportosítási eredmények összehasonlítását követően szerettem volna ezen besorolásokat egy független forrás alapján validálni. Ennek megvalósításához felhőborítottsági adatokat használtam.

Az Országos Meteorológiai Szolgálat (OMSZ) adatszolgáltatásának köszönhetően lehetőségem volt a csoportosítási módszerek eredményeit budapesti felhőborítottsági adatok elemzésének segítségével értékelni. Az OMSZ jóvoltából 2018-as évi egy órás felbontású, pestszentlőrinci felhőborítottsági adatokkal dolgozhattam.

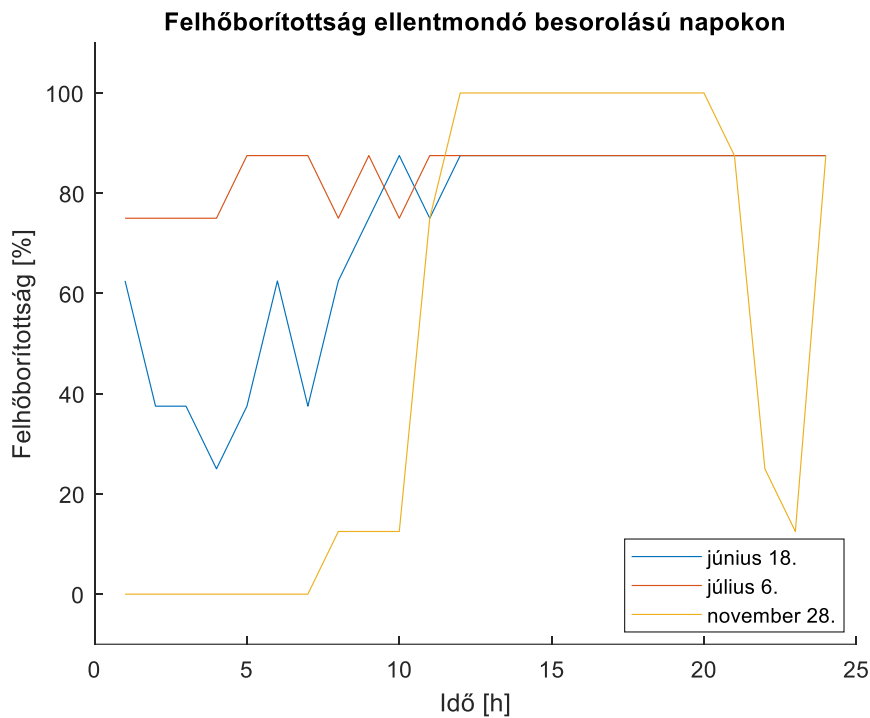
Az adatok alapján megvizsgáltam, hogy az egyes találati arányok mellett, hogyan alakul a csoport tagjainak átlagos napi felhőborítottságának mértéke. Ez látható az alábbi ábrán (45. ábra).



45. ábra: Átlagos felhőborítottság 4, 5 és 6 találatot kapott napok esetén

Látható, hogy a találati arány növekedésével az átlagértékek egyre inkább a szélső értékek felé tolódnak, felhős esetben egyre magasabb, tiszta égboltú esetben egyre alacsonyabb az átlagos, napi felhőborítottság értéke.

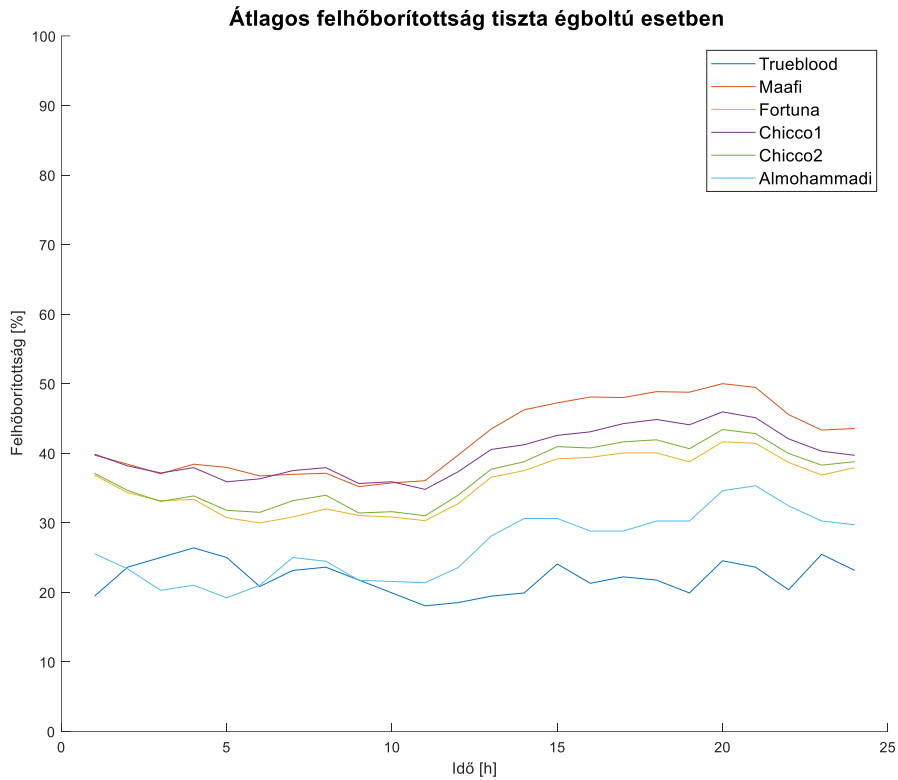
Ez után azokat az ellentmondásokat vizsgáltam meg, melyek esetén egy módszer tiszta égboltúnak, egy másik azonban felhősnek ítélte a napokat. Ezt mutatja az alábbi 46. ábra.



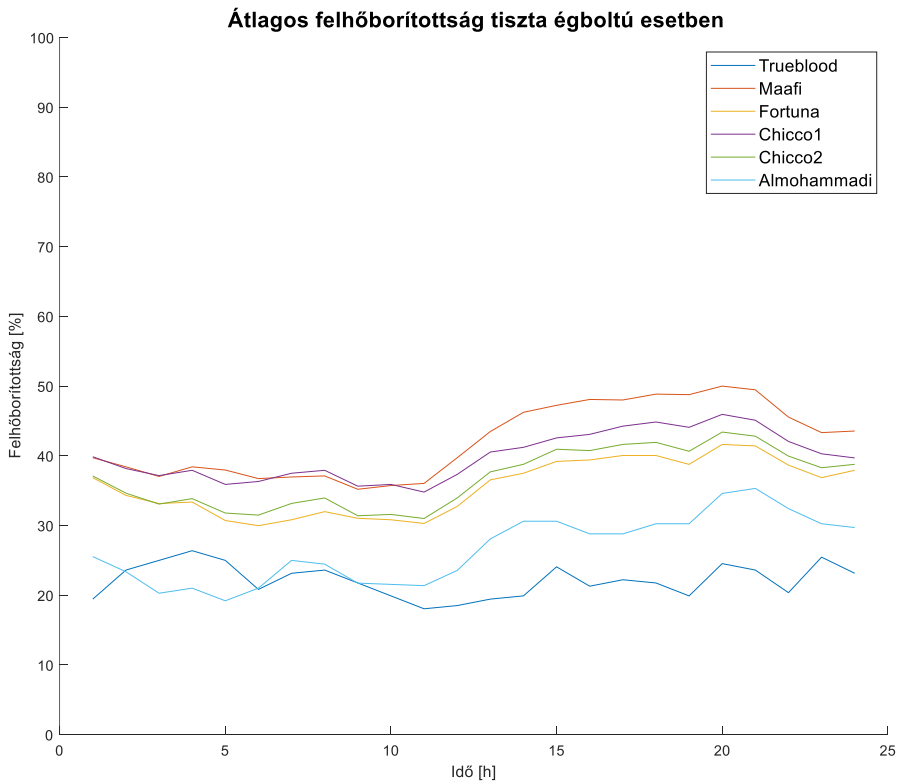
**46. ábra: Felhőborítottság ellentmondó besorolású (tiszta égboltú - felhős) napokon**

Az ábra jól tükrözi azokat a megállapításokat, amiket a fentiekben tettem. Nem meglepő, hogy a módszerek ellentétesen sorolták be ezeket a napokat, hiszen látható, hogy két esetben (június 18. és november 28.) ránézésre sem állapítható meg egyértelműen a napok időjárásának milyensége. A harmadik esetben (július 6.) megfigyelhető, hogy a felhőborítottság szintje egész nap 75% feletti értékű. Itt tehát az Almohammadi módszer egyértelműen hibásan végezte a besorolást.

Végül az egyes módszerek által tiszta égboltúnak, illetve felhősnek ítélt napok felhőborítottságainak átlagértékeit vizsgáltam meg. Ez az összehasonlítás az alábbi 47. és 48. ábrán látható.



47. ábra: Átlagos, napi felhőborítottság tiszta égboltú esetben



48. ábra: Átlagos, napi felhőborítottság felhős esetben

A fenti ábrán jól megfigyelhető, hogy a tiszta égboltúnak, illetve felhősnek ítélt napok átlagos, napi felhőborítottsága hasonló tendenciákat mutat a különböző módszerek esetén. Így megállapítható, hogy a módszerek működése ugyan más, eredményeik mégis korrelálnak.

## 4.2 Módszerek felbontásérzékenysége

A fenti módszerek vizsgálata közben felmerült a kérdés, hogy kidolgozóik vajon mi alapján választották meg a felhasznált adatsorok felbontását. Arra kerestem a választ, hogy az adatsorok felbontásának változtatása milyen mértékben befolyásolja a csoportosítási módszerek eredményeit. Az egyes eljárások a 2018-as évi EK-ból származó, különböző felbontású bemeneti adatokra az alábbi eredményeket adták.

10. táblázat: Csoportosítási módszerek felbontásérzékenysége

Cikk	Csoport elnevezése	Bemenő adatsor felbontása		
		1 min	5 min	10 min
Trueblood	Clear [db]	27	62	103
	High [db]	81	10	0
	Moderate [db]	91	74	29
	Mild [db]	110	139	145
	Overcast [db]	56	80	88
Maafi	Clear-sky [db]	144	151	163
	Partly cloudy [db]	120	113	100
	Cloudy [db]	101	101	102
Fortuna	Clear sky [db]	124	123	115
	Some clouds, high dynamic [db]	65	113	121
	Some clouds, medium dynamic [db]	107	70	70
	Cloudy [db]	69	59	59
Chicco 1	Clear [db]	148	147	138
	Quasi-clear [db]	79	79	86
	Quasi-cloudy [db]	42	44	42
	Cloudy [db]	96	95	99
Chicco 2	Clear [db]	127	126	125
	Quasi-clear [db]	113	114	113
	Quasi-cloudy [db]	70	70	69
	Cloudy [db]	55	55	58
Almohammadi	Clear [db]	104	105	107
	Medium [db]	192	56	57
	Cloudy [db]	69	204	201

A fenti táblázatban található eredményeket a későbbiekben, a módszerek értékelése során fogom felhasználni, elemzésükre akkor kerül majd sor.



## 4.3 Értékelés

A fenti csoportosítási eljárások elvégzése során szerzett benyomásaim és a csoportosítások eredményeinek kiértékeléséből leszűrt következtetéseim alapján most kísérletet teszek az egyes módszerek átfogó, objektív értékelésére.

### 4.3.1 Trueblood

A Trueblood módszer a legegyszerűbb elvű csoportosítás a vizsgált eljárások közül. A variabilitási index számítási elve könnyen érthető és szemléletes. A tiszta égbolt index, mint megfigyelhető volt, univerzálisan alkalmazható, a többi módszerhez is ezt használtam. Az algoritmus végrehajtása nem igényel különösebben nagy számítási kapacitást, a kategóriákba sorolás is könnyen elvégezhető. Az egyes csoportok határfeltételei azonban a földrajzi viszonyok függvényében esetleg eltérők lehetnek, tehát ezek felülvizsgálata szükséges. Az alkalmazott magas csoportszám (5 db) kevésbé hasznos, mint amennyivel bonyolultabbá teszi a csoportosítást. A 10. táblázat alapján megfigyelhető, hogy az adatsor felbontásának csökkentésével (10 perces adatsor) az egyik csoport (High) darabszáma 0-ra adódott. Tehát érdemes lehet a csoportok számát 4-re vagy akár 3-ra csökkenteni. Ahogy a kutatás során a szerzők is rájöttek, az egy másodperces felbontás feleslegesen nagy, bőven elegendő egy perces felbontású bemeneti adatok alkalmazása. A módszer érzékenysége a bemenő adatok felbontására egyébként a csoporthatárok megfelelő korrigálásával kiküszöbölhető. Összességében tehát a módszer, megfelelő csoportszám és csoporthatárok esetén jól alkalmazható.

### 4.3.2 Maafi

A Maafi módszer alkalmazása is viszonylag könnyűnek mondható. Mivel ugyan azt a tiszta égbolt indexet használtam, mint a Trueblood módszer esetén, ennek számítása nem okozott problémát. A variabilitási index számítása azonban számos problémát vet fel. Először is a besugárzásigörbék fraktáldimenzióinak meghatározása (box-count algoritmus) nagy számítási kapacitást igényel (egyszerű, személyi számítógépekhez mérten). Ezen kívül a kiszámított értékek nagymértékben függenek attól, hogy az algoritmus mekkora dobozméretet használ. Ennek optimalizálása komplikált feladat. Szerintem a kapott eredmények megbízhatósága nincsen arányban a kiszámítás módjának bonyolultságával. Ezen kívül az egyes csoportokat meghatározó determinisztikus értékhatárok nem feltétlenül globális érvényűek, azok optima a földrajzi lokáció alapján eltérő lehet. A módszer, a 10. táblázat eredményei alapján csak kis mértékben érzékeny a bemenő adatok felbontására, mely érzékenység a csoporthatárok megfelelő korrigálásával kiküszöbölhető. Összességében tehát a módszer jól alkalmazható a napok csoportosítására, ha rendelkezésre áll megfelelő számítási kapacitás a fraktáldimenziók és az optimális dobozméret meghatározásához, illetve ha a csoporthatárok az adott alkalmazáshoz igazíthatók.

### 4.3.3 Fortuna

A Fortuna módszer alapgondolata szerintem nagyon előremutató. A frekvenciatartományban történő vizsgálat sok olyan zavaró tényezőt kiküszöböl, melyek időtartományban problémát okoznak. Nem szükséges például bonyolult eljárásokkal kiszűrni az adatsorokból a napsütéses órák időszakát, mivel az éjszakai órák zérus amplitúdójú adatpontjai nem jelennek meg a spektrumban (0 Hz frekvenciájú, 0 amplitúdójú összetevők). Ezen kívül a módszer kevésbé lesz érzékeny a bemenő adatsor felbontására, ahogyan azt a 10. táblázat is mutatja. Az változékonyság index kiszámítása azonban, melyhez Fourier transzformáció, periodogram készítés és végül (20) egyenlet szükséges, igen komplikált. Ezen kívül módszer nem a teljes frekvenciatartományt vizsgálja, hanem annak csak egy kiválasztott tartományát. Ezen tartomány kiválasztása az eljárás gyenge pontja, mivel a csoportosítás eredményeit nagymértékben befolyásolja. A tartomány optimális kiválasztása alapos vizsgálatokat igényel, így jelentősen bonyolíthatja, a már egyébként is igen körülményes algoritmust. További negatívum, hogy a k-közép klaszterezési eljárás csak abban az esetben képes megfelelő eredményeket szolgáltatni, ha a bemeneti adatsor ezt lehetővé teszi (ld. 3.2.2.2 fejezet). Összességében a módszerről megállapítható, hogy használata nem egyszerű, de eredményei optimális feltételek mellett megbízhatók.

### 4.3.4 Chicco

A Chicco módszer legfőbb előnye, hogy független az évszakok változásától, illetve a Chicco2 esetben a napon belüli változások időbeliségétől is. Ebből adódóan gyakorlatilag érzéketlen a bemenő adatok időbeli felbontására (10. táblázat). Ugyan akkor éppen ez a tulajdonsága lehet adott esetben a legnagyobb gyengesége is, hiszen az évszakok és az idő teljes eliminálása teljesen rossz eredményekhez is vezethet. Mind ezeken túl módszer alkalmazása nagyon egyszerű, nem igényel különösebben nagy számítási kapacitást, illetve nem szükséges semmilyen komplikált index számítás. A k-közép klaszterezési algoritmus azonban ebben az esetben is okozhat problémát. Összességében a módszer jól használható, azonban eredményeinek validálására legalább egy másik módszer szükséges.

### 4.3.5 Almohammadi

Az Almohammadi módszer alapvetően nem csoportosítási módszer, ahogy azt már korábban is említettem. A módszer gyakorlatilag ugyan úgy normált adatsorral ( $C_{I_{sample}}$ ) dolgozik, mint a Chicco féle eljárás, annak minden előnyével és hátrányával együtt. Ezen kívül a kevert Gauss eloszlások illesztése igen számításigényes folyamat, valamint előzetes megfontolásokat igényel (modusok számának megválasztása). Ennél a módszernél is jelentkezik a k-közép klaszterezés működéséből adódó hibafaktor, ahogy az a 3.2.2.2 fejezetben is látható. Összességében a módszer megközelítése figyelemre méltó, érdemes lehet egy ezen alapuló csoportosítási eljárást kidolgozni.

## 5 Konklúzió, javaslatétel

---

Fent ismertetett munkám során először áttekintettem a dolgozat témájával kapcsolatos szakirodalmat. A napjainkban létező időjárás alapú csoportosítási módszerek közül kiválasztottam ötöt, melyek alapelvei igen eltérők, így a lehetőségek spektrumának nagy hányadát lefedik. A kiválasztott módszerek áttekintése az 1. és 2. fejezetben olvasható.

A szakirodalmi áttekintés után valós, mért adatsorokat kerestem, melyek felhasználásával megvizsgálhattam a megismert módszerek működését a gyakorlatban is. Két adatsorral dolgozhattam, melyeket az MTA EK-tól, illetve a Debreceni Egyetemtől kaptam. Elvégeztem a mért adatsorok letisztázását és kiegészítését, ahol arra szükség volt. Ezen folyamatok ismertetése a 3.1.2 és a 3.2.1 fejezetekben található.

Az ily módon felhasználásra alkalmassá vált adatsorokon végrehajtottam az általam kiválasztott csoportosítási algoritmusokat. Ennek leírását tartalmazza a 3.1.3 és a 3.2.2 fejezet. A két adatsor esetén igen eltérő eredményeket kaptam, ezeket a 3.3 fejezetben hasonlítottam össze egymással.

A csoportosítások megvalósítása után, megvizsgáltam, hogy az egyes módszerek eredményei milyen mértékben korrelálnak egymással, illetve a különbségek, esetleg ellentmondások, milyen okokra, módszerekre jellemző sajátosságokra vezethetők vissza. Vizsgáltam az eljárások robusztusságát az alapján, hogy eredményeik mennyiben változnak meg különböző felbontású bemenő adatok esetén. Kísérletet tettem a módszerek eredményeinek validálására is, egy független forrás alapján. Ehhez az OMSZ-től kapott felhőborítottsági adatokat használtam fel. Ezen vizsgálatok eredményei olvashatók a 4.1 és 4.2 fejezetekben.

Végül a 4.3 fejezetben a fent említett vizsgálataim és tapasztalataim alapján megkíséreltem átfogó értékelést adni az egyes módszerekre, egyaránt rávilágítva azok gyengeségeire és erősségeire is.

Összességében megállapítható, hogy az egyes módszerek külön-külön, ha kisebb módosítások vagy korrigálások (földrajzi helyszín vagy a bemenő adatsor felbontása függvényében) után is, de jól alkalmazhatók. Azonban az is látszik, hogy az eredmények különböző okokra visszavezethető pontatlanságai miatt, a napelemes termelés előrejelzésére a rövidtávú meteorológiai időjárás-előrejelzések sokkal alkalmasabbak. Az időjárás alapú csoportosítás relevanciája csak hosszútávon mutatkozik meg, több évnvi/évtizednyi eredmény alapján megadott földrajzi lokációra tendenciaszerű előrejelzések adhatók. Ehhez azonban megfelelő validálációs rendszer szükséges a módszerek eredményeit nagymértékben befolyásoló különféle hibalehetőségek miatt.

További nehézségeket okoz, ha nem egyetlen módszert szeretnénk csak alkalmazni, hanem a különböző számítási eljárások eredményeit összevetve kívánjuk kialakítani a csoportosítást. Ennek megvalósítására több lehetséges megoldás adható.

Felállíthatunk egy hierarchiát az eljárások között, például megtehetjük, hogy először a legkevesebb kategóriát használó, „durva felbontást” megvalósító besorolást végezzük el, és ennek eredményeit osztjuk tovább a többi, nagyobb számú csoportot alkalmazó eljárással.

Szintén hasznos megközelítés lehet az is, hogy az egyes módszerek eredményeit valamiféle súlyozással látjuk el, és a végső csoportosítást ezek összegzésével alakítjuk ki. Itt azonban az a probléma jelentkezhet, hogy az eltérő csoportszámok miatt az összegzés nehezen megvalósítható.

A meglévő módszerek megfontolását felhasználva és a leghatékonyabb számítási módokat kombinálva egy új csoportosítási eljárás kialakítása lehet a legmegfelelőbb megoldás. További munkám során erre fogok törekedni.

## 6 Irodalomjegyzék

---

- [1] Richard E. Bird and Roland L. Hulstrom: *A Simplified Clear Sky Model for Direct and Diffuse Insolation on Horizontal Surfaces*, Golden, Colorado, USA, February 1981.
- [2] Chris Trueblood, Steven Coley, Tom Key, Lindsey Rogers, Abraham Ellis, Cliff Hansen, Elizabeth Philpot: *PV Measures Up for Fleet Duty*, *IEEE power & energy magazine*, 2013
- [3] A. Maafi, S. Harrouni: *Preliminary results of the fractal classification of daily solar irradiances*, 2003
- [4] Samia Harrouni: *Fractal Classification of Typical Meteorological Days from Global Solar Irradiance: Application to Five Sites of Different Climates, Modeling Solar Radiation at the Earth's Surface: Recent Advances*, Chapter 2, 2008
- [5] Luigi Fortuna, Giuseppe Nunnari, Silvia Nunnari: *A new fine-grained classification strategy for solar daily radiation patterns*, 2016
- [6] T. Soubdhan, R. Emilion, R. Calif: *Classification of daily solar radiation distributions using a mixture of Dirichlet distributions*, 2009
- [7] Gianfranco Chicco, Valeria Cocina, Filippo Spertino: *Characterization of solar irradiance profiles for photovoltaic system studies through data rescaling in time and amplitude*, 2014
- [8] Shahrouz Alimohammadi, Dawei He: *Multi-Stage Algorithm For Uncertainty Analysis of Solar Power Forecasting*, 2016
- [9] Kovács András, Csáji Balázs Csanád, Váncza József: *Egy megvalósított intelligens közvilágítási rendszer: E+grid*, 2016
- [10] <https://www.mathworks.com/help/stats/kmeans.html#bueq7aj-5>