



Közzététel: 2025. január 28.

A tanulmány címe:

**Módszertani kísérlet a villamosenergia-fogyasztás előrejelzésére hagyományos eljárással és mesterséges intelligencia (MI) alkalmazásával**

Szerzők:

**OBÁDOVICS CSILLA**

PhD, a Soproni Egyetem Lámfalussy Sándor Közgazdaság-tudományi Karának egyetemi tanára

E-mail: [Obadovics.Csilla@uni-sopron.hu](mailto:Obadovics.Csilla@uni-sopron.hu)

**SZABADOS LEVENTE**

a Soproni Egyetem Széchenyi István Doktori Iskolájának PhD-hallgatója

E-mail: [L.Szabados@int.fs.de](mailto:L.Szabados@int.fs.de)

DOI: <https://doi.org/10.20311/stat2025.1.hu0021>

**Az alábbi feltételek érvényesek minden, a Központi Statisztikai Hivatal (a továbbiakban: KSH) *Statisztikai Szemle* c. folyóiratában (a továbbiakban: Folyóirat) megjelenő tanulmányra. Felhasználó a tanulmány vagy annak részei felhasználásával egyidejűleg tudomásul veszi a jelen dokumentumban foglalt felhasználási feltételeket, és azokat magára nézve kötelezőnek fogadja el. Tudomásul veszi, hogy a jelen feltételek megszegéséből eredő valamennyi kárért felelőséggel tartozik.**

1. A jogszabályi tartalom kivételével a tanulmányok a szerzői jogról szóló 1999. évi LXXVI. törvény (Szt.) szerint szerzői műnek minősülnek. A szerzői jog jogosultja a KSH.
2. A KSH földrajzi és időbeli korlátozás nélküli, nem kizárólagos, nem átadható, térítésmentes felhasználási jogot biztosít a Felhasználó részére a tanulmány vonatkozásában.
3. A felhasználási jog keretében a Felhasználó jogosult a tanulmány:
  - a) oktatási és kutatási célú felhasználására (nyilvánosságra hozatalára és továbbítására a 4. pontban foglalt kivétellel) a Folyóirat és a szerző(k) feltüntetésével;
  - b) tartalmáról összefoglaló készítésére az írott és az elektronikus médiában a Folyóirat és a szerző(k) feltüntetésével;
  - c) részletének idézésére – az átvevő mű jellege és célja által indokolt terjedelemben és az eredetihez híven – a forrás, valamint az ott megjelölt szerző(k) megnevezésével.
4. A Felhasználó nem jogosult a tanulmány továbbértékesítésére, haszonszerzési célú felhasználására. Ez a korlátozás nem érinti a tanulmány felhasználásával előállított, de az Szt. szerint önálló szerzői műnek minősülő mű ilyen célú felhasználását.
5. A tanulmány átdolgozása, újra publikálása tilos.
6. A 3. a)–c) pontban foglaltak alapján a Folyóiratot és a szerző(ke)t az alábbiak szerint kell feltüntetni:  
„*Forrás: Statisztikai Szemle* c. folyóirat 103. évfolyam 1. számában megjelent, **Obádovics Csilla – Szabados Levente** által írt, **Módszertani kísérlet a villamosenergia-fogyasztás előrejelzésére hagyományos eljárással és mesterséges intelligencia (MI) alkalmazásával** című tanulmány (link csatolása)”
7. A Folyóiratban megjelenő tanulmányok kutatói véleményeket tükröznek, amelyek nem feltétlenül esnek egybe a KSH vagy a szerzők által képviselt intézmények hivatalos álláspontjával.

Obádovics Csilla – Szabados Levente

# Módszertani kísérlet a villamosenergia-fogyasztás előrejelzésére hagyományos eljárással és mesterséges intelligencia (MI) alkalmazásával

## Methodological research for forecasting electricity consumption with traditional methods and with AI application

Obádovics Csilla PhD, a Soproni Egyetem Lámfalussy Sándor Közgazdaság-tudományi Karának egyetemi tanára  
E-mail: Obadovics.Csilla@uni-sopron.hu  
Szabados Levente, a Soproni Egyetem Széchenyi István Doktori Iskolájának PhD-hallgatója  
E-mail: L.Szabados@int.fs.de

A tanulmány egy olyan módszertani kísérletet ismertet, amelynek célja egy szabványosítható energiagazdálkodási irányítási rendszer megalkotása volt. Kiindulópontját annak vizsgálata képezte, hogy bizonyos változók ismeretében (pl. időjárási adatok, látogatószám, nettó árbevétel stb.) hogyan lehet néhány napra előre jelezni az ipari fogyasztók negyedórás felbontású villamosenergia-fogyasztását. Az ilyen terhelésbecslés gazdasági jelentősége nyilvánvaló, kiemelt szerepe van a villamos energia beszerzése területén, és segítséget nyújt a várható fogyasztás prognosztizálásához. A módszer alkalmazásával csökkenthetők az energiaköltségek. Az energia- és az éghajlati válsággal összefüggésben kulcsfontosságú, hogy a szervezetek fejlett módszereket alkalmazzanak az energiafogyasztás és az energiaköltségek visszafogására. A tanulmányban különböző modellek alkalmazását vizsgáltuk a villamosenergia-igény előrejelzésére, amely növelheti a piac hatékonyságát, és hozzájárulhat a hálózat stabilitásához. Egy magyarországi hipermarket részletes villamosenergia-fogyasztási adatsorát elemeztük, 15 perces időközönként, 48 órás előrejelzésekre fókuszálva. Statisztikai módszerekkel feltártuk a terhelési jellegzetességeket, a napi, a heti, a havi és az éves lefutások főbb jellemzőit. Az adatok alapján meghatároztuk a tipikus napokat, feltártuk és jellemeztük a tipikus eseményeket. Elemeztük az órás hőmérsékleti adatok és a villamosenergia-fogyasztás közötti összefüggéseket. A feltárt összefüggések alapján 24 és 48 órás előrejelző modellt alkottunk, amelynek segítségével adott hibahatáron belül előre becsülhetők a 15 perces átlagterhelések. Módszertanunk magában foglalja a klasszikus modelleket, például a lineáris regressziót, valamint a legmodernebb *deep learning* modelleket, így a TiDE-t és a Lag-Llama-t.

Kulcsszavak: időszerelemzés, energiafogyasztás-előrejelzés, *deep learning* modellek

The presentation describes a methodological experiment, the aim of which was to create a standardized Energy Management Management System. Its starting point was the investigation of how, given certain variables (e.g. weather data, number of visitors, net sales, etc.), it is possible to forecast the electricity consumption of industrial consumers with quarter-hour resolution for a few days. The economic importance of such load estimation is obvious and has a prominent role in the field of electricity procurement, and it provides assistance in forecasting the expected consumption. Using this method, energy costs can be reduced. In the context of the energy and climate crisis, it is crucial

that organizations adopt advanced methods to reduce energy consumption and energy costs. In the study, we examine the use of different models to forecast electricity demand, which can increase the efficiency of the market and contribute to the stability of the network. We analyze the detailed electricity consumption data set of a Hungarian hypermarket, focusing on 48-hour forecasts at 15-minute intervals. We used statistical methods to reveal the load characteristics, the main characteristics of the daily, weekly, monthly and annual runs. Based on the data, we defined typical days, explored and characterized typical events. We analyzed the correlations between hourly temperature data and electricity consumption. Based on the revealed correlations, we created a 24- and 48-hour forecasting model, which can be used to predict the 15-minute average loads within a given margin of error. Our methodology includes classic models such as linear regression, and cutting-edge “deep learning” models such as TiDE and Lag-Llama.

Keywords: time series analysis, electricity consumption prediction, “deep learning” models

A klímaválsággal és a zöldenergiára való átállással összefüggésben kiemelten fontos, hogy minden lehetséges eszközt felhasználjunk energiafogyasztásunk csökkentésére. A villamos energia nem tárolható gazdaságosan nagy mennyiségben, így a piaci hatékonyság kulcsfontosságú az alacsony árak fenntartásához. A rövid távú energiaszükségletekre vonatkozó pontosabb előrejelzések révén az energia megfizethetőbbé válik a vállalatok számára.

A jobb energiaigény-előrejelzés pozitív hatással van a hálózat stabilitására, mivel bár a megújuló energiaforrások növekvő mértékű alkalmazása előnyös a fenntarthatóság szempontjából, destabilizáló hatással lehet a hálózatra (*Biber et al., 2022*). A kereslet pontos előrejelzései segíthetnek enyhíteni ezeket a destabilizáló hatásokat azáltal, hogy az előrejelzések alapján a szolgáltatók biztosítják a kiegyensúlyozottabb és megbízhatóbb energiaellátást.

A mesterséges intelligencia (MI) technikáinak, például a *deep learning*-nek az alkalmazása a villamosenergia-igények előrejelzésére potenciálisan jelentős energiamegtakarításhoz vezethet, és bemutatja a mesterséges intelligencia jelentőségét az éghajlati kihívások kezelésében. Ez a kettős előny – gazdasági és környezeti – aláhúzza a *deep learning* technikák feltárásának szükségességét ezen a területen.

A klasszikus idősorra vonatkozó modellezési technikákat széles körben értékelik robusztusságuk és megbízhatóságuk miatt a statisztikusoktól kezdve az üzleti elemzőkig. Ezek a módszerek, köztük az ARIMA (*Box–Jenkins, 1970*) az exponenciális simítás (*Winters, 1960*) és a State-Space-modellek számos alkalmazásban hatékonyan bizonyultak. Azonban összetettebb, sokdimenziós, nemlineáris, bonyolult összefüggésekkel rendelkező adathalmaz esetében problémás az alkalmazásuk.

A technikák és modellek új hulláma van kialakulóban, amelyek közül több is az úgynevezett megalapozómodellek („foundational models”, {*Bommasani et al., 2021*}) előnyeit hivatott kihasználni. Figyelemre méltó példákat találni TimeGPT alkalmazására *Garza és Mergenthaler-Canseco (2023)*, *Woo és szerzőtársai (2024)*, *Das és szerzőtársai (2024)*, *Rasul és szerzőtársai, (2024)*, *Ansari és szerzőtársai (2024)*, valamint *Goswami és szerzőtársai (2024)* tanulmányában. Az idő-soros elemzések alapmodelljeinek szélesebb körű áttekintését lásd *Ye és szerzőtársainál (2024)*.

Ebben a cikkben a hagyományos és az innovatív technikák tesztelésére és összehasonlítására törekedtünk, továbbá meghatároztuk a módszerek hatékonyságát a villamosenergia-igény előrejelzésében, reményeink szerint ezzel is hozzájárulva a hatékonyabb energiagazdálkodáshoz és -fogyasztáshoz.

## 1. A becslési eljárás előkészítése

A becslési modell előállítását megelőzően adatbázis-szerkesztési lépések, majd az adatsor jellemzőinek teljes feltárása szükségesek. Az adatelemzés során először rendszereztük, átstrukturáltuk és kiegészítettük a szükséges kódváltozókkal a kapott adatokat. A továbbiakban megvizsgáltuk a hipermarket fogyasztási adatsorát, valamint a fogyasztás és a hőmérséklet közötti összefüggést. A feltárt összefüggéseket statisztikai és matematikai módszerekkel feldolgoztuk, amely alapján különböző modelleket dolgoztunk ki a 15 perces átlagterhelés előreszámításához.

### 1.1. Adat-előkészítés

A tanulmányban felhasznált adathalmaz egy magyarországi hipermarket villamosenergia-fogyasztásának részletes nyilvántartása, 15 perces időközönként, közel egy éven keresztül, 2007. szeptember 1. és 2008. július 31. között. A hipermarket térségére vonatkozó, órás pillanatnyi hőmérsékleti (°C) adatsort az Országos Meteorológiai Szolgálat Éghajlati Osztályától kaptuk, a 2007. augusztus 1. 1 óra és 2008. július 31. 24 óra közötti intervallumra. Az órás adatokból lineáris interpolációval számoltuk ki a negyedórás értékeket.

Az adat-előkészítési folyamat magában foglalta az adathalmaz rendszerezését és átstrukturálását, új változók létrehozását, valamint a fogyasztási adatok összefüggését a hőmérsékleti rekordokkal. Különös figyelmet fordítottunk a rendhagyó

események, például az áramkimaradások és az ünnepnapok kezelése, hogy biztosítsuk a prediktív modellek robusztusságát. Ennek az átfogó adatsornak az elemzésével egy olyan, 48 órás előrejelzési modell kidolgozását tűztük ki célul, amely képes 15 perces felbontással megbecsülni a villamosenergia-fogyasztást, ezzel is segítve az energiaköltségek csökkentését.

Mielőtt több különböző modellt alkalmaztunk volna az adathalmazra, feltáró adatelemzést végeztünk, amely erősen befolyásolta modelljeink paramétertervezési döntéseit.

### 1.1.1. Zajkezelés

Az adathalmazban volt néhány olyan szélsőséges adat, amelyeket torzító hatásuk miatt figyelmen kívül hagytunk a modellek felállítása során.

### 1.1.2. Rendhagyó események, napok kezelése

**Áramszünet előfordulása.** A megfigyelt időintervallumon belül három alkalommal fordult elő áramszünet, mindhárom alkalommal rövid ideig. Ezeket az időadatokat külön jelöltük, és az elemzés során és a becslési eljárásnál nem vettük őket figyelembe.

**Az ünnepnapok kezelése.** A munkaszüneti napok jelentős, az extrém értékekben megmutatózó hatása miatt az ünnepnapokra vonatkozóan bináris változót alkalmaztunk. Az ünnepnapok esetében a villamosenergia-fogyasztás eseti megkülönböztetést kapott, külön naptípus szerint kezeltük.

**Egyéb szélsőségek.** Az egyéb előforduló szélsőséges adatokra az aktuális elemzésnél térünk ki.

### 1.1.3 Új változók létrehozása, kódváltozók kialakítása

A megadott szempontok szerinti elemzés új kódváltozók létrehozását tette szükségessé. Ilyenek az időpontot meghatározó kódváltozók, mint az évszám, a hónap, a nap sorszáma, az óra sorszáma, a negyedóra sorszáma változók, továbbá a naptípus: hétfőtől vasárnapig a napok 1–7-es kóddal, az ünnepnapok 0-s kóddal és az áramszünet időtartamához tartozó adatok 8-as kóddal jelölve.

Az elemzés szempontjából fontos megkülönböztetni az éjjeli és a nappali időszakokat az adatsor jelentősen eltérő viselkedése miatt, így a nyitvatartási-zárvatartási időszak megkülönböztetésére is készítettünk egy kódváltozót.

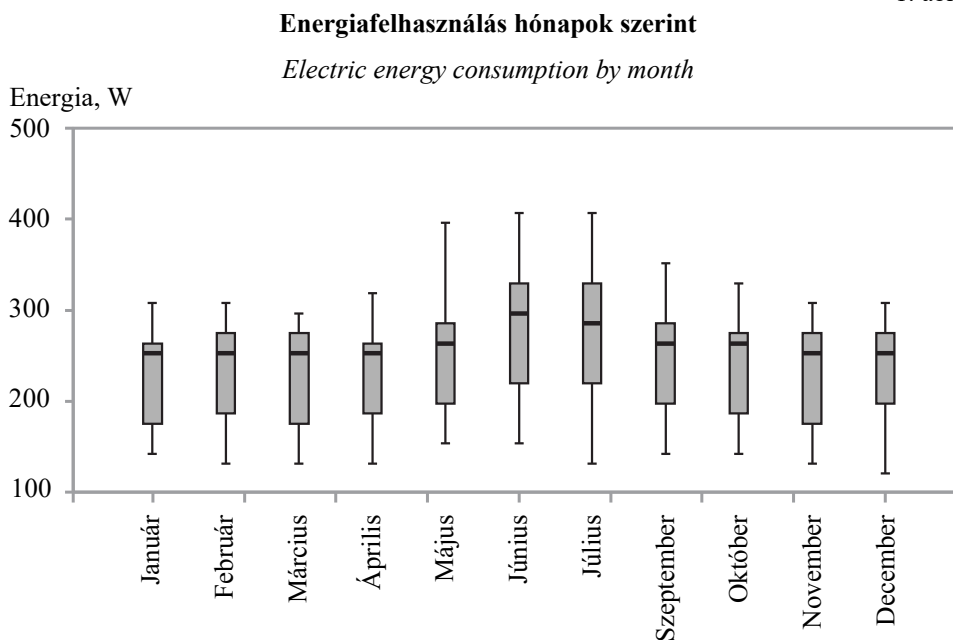
Az elemzések azt mutatták, hogy a hőmérséklet hatása az átlagterhelésre 14 °C-os napi átlaghőmérséklet felett szignifikáns és jól jellemezhető, ezért egy hőmérsékleti kódváltozót is hozzárendeltünk az adatbázishoz.

## 1.2. Terhelési adatok elemzése

A pillanatnyi villamosenergia-fogyasztás sok tényezőtől függ. Nappal több energiát használunk, mint éjszaka, nyáron a hűtés miatt fogy több áram. Ezek mellett még más tényezők – például az előzőekben jelzett típusnapok – is befolyásolhatják egy bevásárlóközpont napi villamosenergia-igényét.

A villamosenergia-fogyasztási adatok havi megoszlását tekintve (1. ábra) látható, hogy általában a nyári hónapok magasabb mediánnal rendelkeznek, így feltehető, hogy a fő energiafelhasználás a hűtésből származik.

1. ábra



Forrás: a bevásárlóközpont adatai alapján saját szerkesztés.

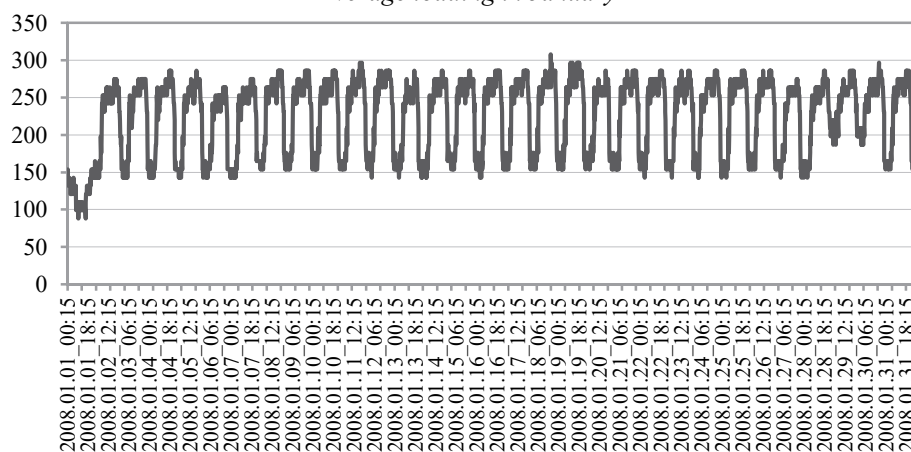
Az üzletek forgalmát és ezzel együtt az energiafogyasztást az is befolyásolja, hogy a hónap melyik időtartamára esik a nap (általában fizetési napok körül), így nap szerinti dátumot is figyelembe kell venni a modell megalkotásakor.

Példaként három különböző hónap terhelési görbét mutatjuk be (2., 3., 4. ábra).

Elsőként tekintsük a januári átlagterhelés görbét. Az ábrán jól látható a január 1-jei ünnepnap alacsony értéke, továbbá az is, hogy az egyes napok lefutása nem mutat különösebb eltérést, kivéve január 29-e és 30-a éjszakáját, amikor az éjszakai átlagértéktől magasabban maradt az elektromos terhelés (rendhagyó éjszakai műszak).

2. ábra

**A januári átlagterhelés alakulása**  
*Average loading in January*

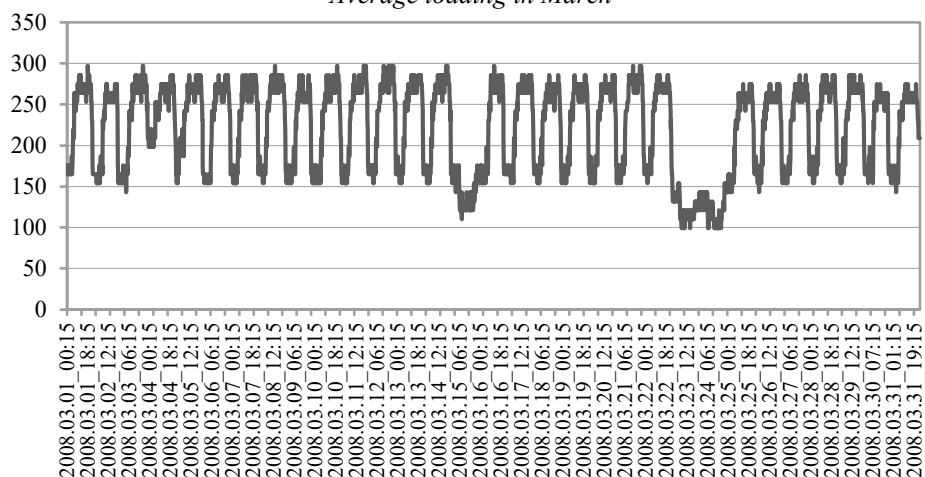


Forrás: saját szerkesztés.

A 3. ábrán a március havi terhelés görbét látjuk. Három ünnepnap volt a hónapban, március 15. és 23–24. (húsvét). Ezek terhelése az ünnepnapos típusnak felel meg. Egy rendhagyó éjszakai terhelést tapasztalhatunk, március 3-áról 4-ére a terhelés nem esett vissza a szokásos éjszakáknak megfelelő szintre. A többi nap terhelési görbéje hasonló lefutásokat mutat.

3. ábra

**A márciusi átlagterhelés alakulása**  
*Average loading in March*

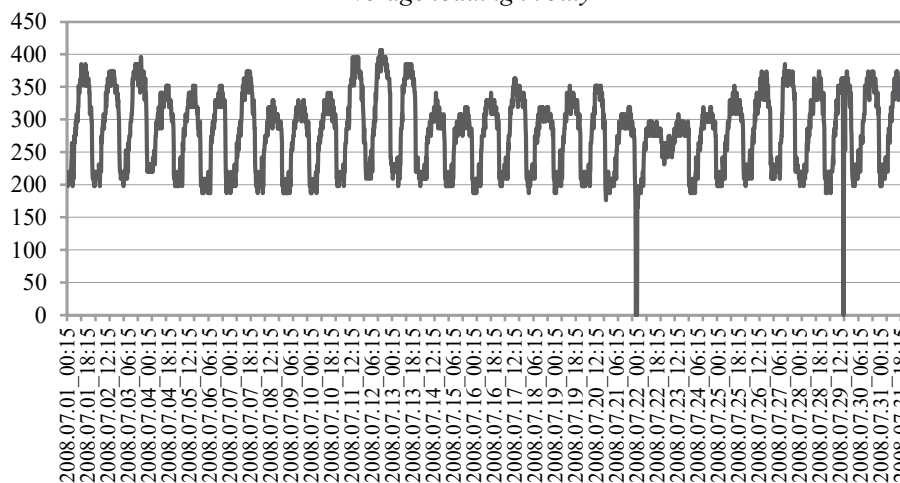


Forrás: saját szerkesztés.

A 4. ábrán a július hónap terhelési görbáját mutatjuk be. Kétszer fordult elő rövid áramszünet, amit a terhelés 0-ra való visszaesése is jelez. Összehasonlítva a téli hónapok terhelési görbéivel, szembevetendő a különbség: míg a téli hónapokban kismértékű ingadozást tapasztalunk csak az egyes napok között, alapvetően egy stagnáló egyenes körül mozognak az értékek, addig a nyári hónapban a hőmérséklet függvényében jelentős eltérések tapasztalhatók az egyes napok terhelési átlagértékében.

4. ábra

**A júliusi átlagterhelés alakulása**  
*Average loading in July*



Forrás: saját szerkesztés.

A hónapok között eltérő a napon belüli áramfogyasztás (Függelék F1. ábra), amely alapvetően a hőmérséklet függvénye lehet. Erre később részletesebben kitérünk.

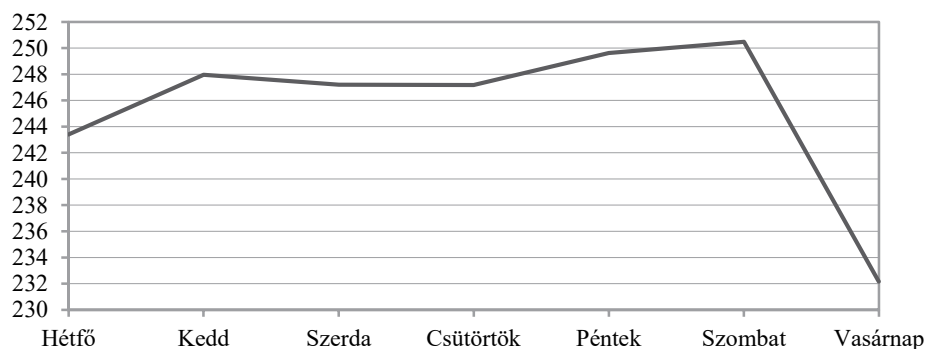
### 1.3. Típusnapok

Mint korábban jeleztük, az adott héten belül a napnak is hatása van az energiafogyasztásra. A heti eloszlás elemzése (5. ábra) erősebb energiafogyasztást mutat a pénteki és a szombati „bevásárlási napokon”, hangsúlyozva, hogy nem elég a „hétvége vs. hétköznap” jellemző, tehát a hét adott napja funkció elengedhetetlen lesz modelljeink összeállításánál.



5. ábra

**A terhelés napi átlagértéke**  
Average loading by day types



Forrás: saját szerkesztés.

Hétfőn és vasárnap szignifikánsan alacsonyabb az energiafogyasztás (ANOVA sig = 0,000), mint a hét többi napján; kedden, szerdán és csütörtökön közel azonos; míg pénteken és szombaton enyhén magasabb.

A vasárnapi alacsonyabb átlagérték a későbbi nyitásnak és a korábbi zárásnak tulajdonítható. Az adatsor részletes elemzéséből egyértelműen kitűnik, hogy vasárnap 7 órakor kezdődik a nappali időszak, ekkor veszi fel a hétköznapi 6 órai értéket az áramfogyasztás. Ugyanígy az esti időszak 22:30 helyett 20:30-kor kezdődik. A rövidebb nappali időszak miatt a napi átlagérték alacsonyabb, mint az egyéb napokon (1. táblázat).

1. táblázat

**A naphatás leíró statisztikai eredményei**  
Descriptive statistics by type of day

(W)

Napok	N	Átlag	Szórás	Standard hiba	95%-os konfidencia-intervallum		Minimum	Maximum
					alsó határ	felső határ		
Hétfő	4 368	243,42	53,416	0,808	241,83	245,00	121	385
Kedd	4 315	247,98	51,576	0,785	246,44	249,52	132	385
Szerda	4 512	247,22	51,264	0,763	245,73	248,72	132	407
Csütörtök	4 416	247,20	52,866	0,796	245,64	248,76	132	407
Péntek	4 512	249,62	53,385	0,795	248,06	251,18	143	407
Szombat	4 512	250,49	54,298	0,808	248,90	252,07	132	407
Vasárnap	4 416	232,16	55,412	0,834	230,53	233,80	132	385
Összesen	31 051	245,47	53,505	0,304	244,88	246,07	121	407

Forrás: saját számítás.

A korábbi elemzések eredményeképpen a következő típusnapokat alakítottuk ki, a további felhasználás céljából kétféle verzióban:

**1. A napok mindegyikét megkülönböztettük,** vagyis minden egyes nap külön naptípust alkot. Így hétfő, kedd, szerda, csütörtök, péntek, szombat és vasárnap különálló naptípus, az ünnepnappal kiegészülve pedig összesen 8 típust határoztunk meg.

*Előnye:* részletesebb bontás, esetleges napok közötti eltérések befolyásoló hatása mutatható ki.

*Hátránya:* az alacsony elemszám miatt a becslés bizonytalansága nagyobb.

**2. A hétköznapokat összevontuk,** hétfő, kedd, szerda, csütörtök, péntek, szombat egy naptípust alkot, ezek negyedórás átlagterhelés-alakulása szignifikánsan nem tér el egymástól. Vasárnap különálló naptípus, így az ünnepnappal kiegészítve összesen 3 naptípust definiáltunk.

*Előnye:* nagyobb az egy típusba tartozó napok száma, a becslés bizonytalansága csökken.

*Hátránya:* bizonyos differenciáló hatások összemosódhatnak.

Decemberre a vasárnapok hétköznapi besorolást kaptak, miután a nyitvatartás is hétköznapi felel meg, a terhelés alakulása is a hétköznapi lefutásokat követi.

Az ünnepnapok alakulása egyedileg is becsülhető, természetesen az adott nap egyszeri előfordulása az alapadatbázisban a becslés bizonytalanságát eredményezi. A minimum- és a maximumértékek megadásával a biztos találat maximum 77 W terjedelemben szóródik ( $\pm 38$  W), maximum  $\pm 25\%$  az átlagtól való eltérés.

Napközben téli időszámítás idején reggel 7, nyári időszámítás idején reggel 8 órától délután 6 (téli időszámításban), illetve 7 (nyári időszámításban) óráig nagyjából stagnálóan alacsony a fogyasztási érték, amely az esti, éjszakai órákban kissé emelkedik. Az őszi és a téli hónapokban az ünnepnapokon alacsonyabb a nappali átlagérték, a tavaszi ünnepnapokon – főként a három májusin, amikor magas nappali átlaghőmérséklet volt – magasabb terhelési adatokat tapasztaltunk (Függelék F2. ábra).

Az ünnepnapok átlagterhelése 127,6, szórása 18,6, a relatív szórás 14,6%.

#### 1.4. Hőmérsékleti adatok feltáró elemzése

Az adatok elemzése során szembetűnő eltérés tapasztalható a nyári és a téli, azaz a meleg és a hideg napok energiafelhasználásában. Ebben a fejezetben a hőmérséklet alakulását elemezzük a vizsgált időszakban.

A leghidegebb hónap a december volt,  $-0,3$  °C-os átlaghőmérséklettel, a második leghidegebb a január. A leghidegebb napok ( $-11$  °C-os átlaghőmérséklet) januárban és februárban voltak, január 4-én, 5-én és február 17-én. A legmelegebb

hónapok július és június voltak, 21 °C-os átlaghőmérséklettel. 14 napon volt 30 °C feletti hőmérséklet, május 28-án, 31-én, június 22–25 között, július 3-án, 7-én, 11–13 között és 29–30-án. Július 13-án volt a legmelegebb, közel 35 °C-os csúcshőmérséklettel.

A leghidegebb órák 22 és 8 óra között szóródtak, a –11 °C-ot hajnali 3 órakor mérték. A legmelegebb órák 15 és 19 óra között voltak, a 30 °C feletti hőmérsékletek 75%-át ekkor mérték. (Ebből a 4 és az 5 órás adatok 35%-ot tesznek ki.)

## 2. Becslési eljárás hagyományos módszerekkel

Első lépésben le kell szögezni, hogy az egyéves adatsor egészen pontos közelítést és előrejelzést nem tesz lehetővé. Mindamellet az egyéves idősor alapos, minden szegmensre kiterjedő részletes elemzése azt mutatja, hogy bizonyos határok között a becslés és az előrejelzés elvégezhető.

### 2.1. Az energiafogyasztás hőmérséklet-függősége

A közeli meteorológiai állomás óránkénti hőmérsékleti adatainak elemzése és az energiafogyasztás mérése szoros összefüggést mutat. Pontosabban, még ha ki is vesszük a heti és a napi szezonális mintázatokat a fogyasztásmérésből, megfigyelhető a Granger-féle ok-okozati összefüggés (*Granger, 1969*). Ez arra enged következtetni, hogy kb. 2 órás késéssel a hőmérséklet változása hatással van az energiafogyasztásra. Ezen megfigyelés alapján a többváltozós modelleknél a helyi hőmérsékletet is figyelembe vettük modelljeink készítésekor.

Ha a nappali átlaghőmérséklet meghaladja a 14 °C-ot, a terhelést már nem csak az adott időpont befolyásolja, a hőmérséklet hatása erősen meghatározóvá válik.

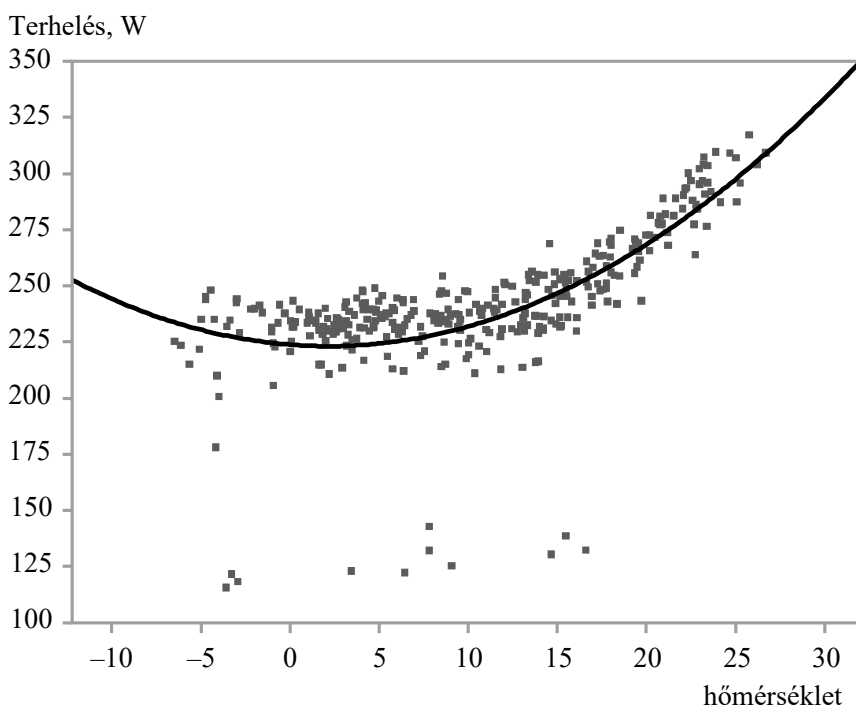
A 6. ábrán jól kivehető az ünnepnapok alacsony értéke, valamint a 14 °C-os határvonal. A terhelési és a hőmérsékleti görbe összevetéséből is látható ez az összefüggés, de a hipermarket részéről is megerősítést nyert, hogy a 14 °C-ot meghaladó nappali átlaghőmérséklet esetén a belső hőmérséklet szabályozására használatos elektromos berendezést kell üzemeltetni. Alacsonyabb hőmérséklet esetén stagnáló terhelési értékeket kapunk, míg ennél magasabb átlaghőmérséklet esetén meredeken emelkedik a terhelés (6. ábra).

A naptári év évszakos váltakozásából eredően két fő csoport kialakítására volt szükség: az ún. téli/hideg napokra és az ún. nyári/meleg napokra, amelyek nappali

átlaghőmérséklete a 14 °C-ot nem éri el, illetve meghaladja azt (Függelék F3. és F4. ábra).

6. ábra

**A hőmérséklet hatása a villamosenergia-fogyasztásra**  
*Relationship between temperature and average loading*



Forrás: saját szerkesztés.

A téli napokra a legjobb közelítést a tényadatok átlagértékével való becsléssel kaptuk, a kiugróan magas vagy a „normálistól” eltérően viselkedő adatsorok figyelmen kívül hagyása mellett. Elvégeztük a hideg napok terhelési görbéjének polinomiális közelítését is, amely igen jó eredményt adott (94,5%-os determinációs együtthatóval).

A meleg napok terhelésbecslését többféle verzióban is elkészítettük. Az első esetben a téli napok alakulása adta az alapot, a nyári napok pedig kiegészültek egy hőmérsékletfüggő hozzáadott értékkel, amelyet exponenciális közelítéssel is és lineáris közelítéssel is meghatároztunk külön minden egyes típusnapra. A második esetben a hideg napok terhelési értékeit mellőzve, a teljes terhelésen alapuló lineáris közelítéssel becsültük meg az adatokat. Ez a megoldás vezetett jobb eredményre.

Végezetül tehát olyan számolótáblát kaptunk, amelynél a hőmérséklet-előrejelzés függvényében először ki kell választani, hogy téli vagy nyári napról van-e szó, majd a típusnapnak megfelelő számolótáblából kiolvasható a negyedórás előrejelzés.

Az éjszakai órák 22–6 óra közötti, kis szórású, csak naptípusfüggő adatsorok. Néhány kivételes nap szélsőségesen magas értékkel rendelkezik, aminek oka nem ismert. Ezen napok kiválasztása megtörtént, a becslési eljárás során külön kategóriát képeznek. Vagyis az előrejelzés során, amennyiben a tervező tisztában van azzal, hogy az éjszakai terhelés az átlagnál nagyobb lesz (pl. árufeltöltés?), a kiugró értékes táblából kell az adott éjszakai órákra a becsléshez az előrejelző adatsort kiválasztania. Ezeket a szélsőségeket a becslés során az adatsor átlagérték-számításakor nem vettük figyelembe.

### 2.1.1. Becslés a hideg napokra

A hideg napok terhelésének becslését elvégezhetjük a negyedórás átlagértékek pontbecslésével (kiegészítve konfidencia-intervallum megadásával), amely az éjszakai és a nappali közelítésére egyaránt alkalmazható, ugyanis a hideg napokon a terhelés a naptípuson kívül csak az óra:perc függvénye. Ez az eljárás elvégezhető 3 naptípusra éppúgy, mint 7-re vagy 8-ra, attól függően, hogy a hétköznapokat, illetve a szombatot, vasárnapot meg akarjuk-e különböztetni, vagy összevontan kezeljük.

Készítettünk egy becselőmodellt a hétköznapok napközbeni terhelési adataira időszorelemzés segítségével.

Az idősor komponensei:

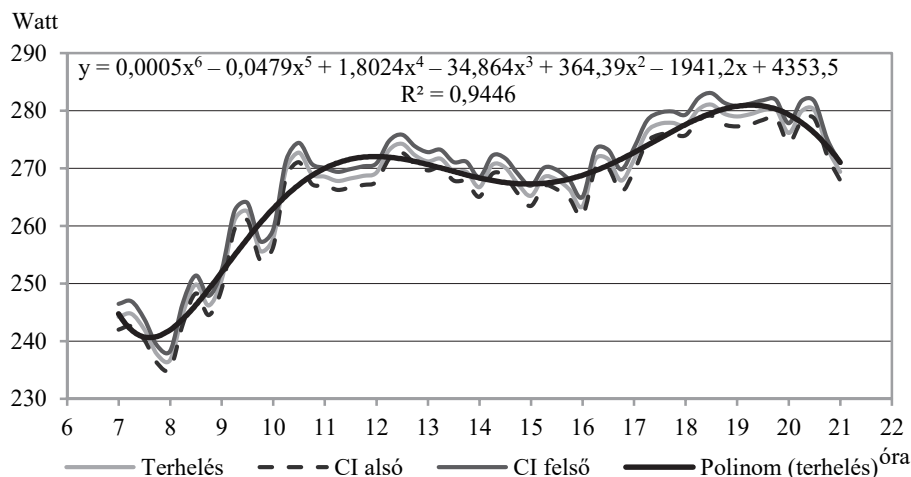
1. **Trend (alapirányzat):** az idősorban tartósan érvényesülő tendencia, a mi esetünkben ez egy állandó, az átlagterheléssel megegyező érték.
2. **Szezonális ingadozás:** szabályosan ismétlődő hullámváltozás, a negyedórás periódusok azonos szakaszában mindig azonos irányban, azonos abszolútértékkel téríti el az idősort a „trendtől”.

Megfigyelhető a havi terhelési görbén, hogy a hideg időszakban tendencia nem jellemzi a terhelés alakulását. Az viszont szembetűnő, hogy a nappali időszaknak az adott negyedórára jellemző értéke van. Az idősor összetevői közötti kapcsolat additív jelleget mutat, az alapiránytól való eltérés abszolút nagysága közel állandó.

A negyedórás időpontokra kiszámolva az addíciós értékeket, elvégeztük a hideg napos hétköznap nappali időszakának becslését. Második lépésben erre az adatsorra megkerestük a legjobban illeszkedő függvényt.

Többféle görbeillesztést kipróbálva végül a hatodfokú polinomiális görbe illeszkedett a legjobban a 7–21 óra közötti időszak negyedórás terhelési átlagadataira. Az illesztett függvény egyenletét és az illeszkedés szorosságát, valamint a pontsor konfidencia-intervallumának (CI) görbét láthatjuk a 7. ábrán.

7. ábra

**Nappali átlagterhelés konfidencia-intervallummal és közelítő függvénnyel***Daily average loading with confidence intervals and with estimate curve*

Forrás: saját szerkesztés.

A hatodfokú polinom segítségével bármely hidegnek ígérkező hétköznapra elvégezhető a terhelésbecslés. Ugyanezzel az eljárással elkészíthető a vasárnapi és az éjszakai becslőmodell is.

### 2.1.2. Becslés a meleg napokra

A meleg napok szeptemberben, májusban, júniusban és júliusban találhatók. Szeptembernek az első fele átlaghőmérséklete alapján a meleg, második fele a hideg napos csoportba tartozik. Szeptember részben hideg, részben meleg napjai a becslést bizonytalanná teszik, így a közelítőfüggvény számításakor ezt a hónapot teljes egészében figyelmen kívül hagytuk. Május első felében az átlaghőmérséklet alig éri el, illetve haladja meg a kritikus 14 °C-ot, ami az előzetes számítások szerint nagyban rontja a becslés pontosságát, ezért csak a május 13-a utáni időszak került be a számítási folyamatba.

Júniusban 15-e, júliusban 23-a szélsőségesen hűvös nyári nap volt, a közelítés bizonytalanságának csökkentése céljából ez a két nap kimaradt a számításokból. Június 9-e és 16-a este 20 óra utáni időszak kiugróan alacsony értékekkel rendelkezett, ezen időpontok adatsora törlésre került.

A 14 °C-ot meghaladó átlaghőmérsékletű napok esetében összefüggés-vizsgálatot végeztünk a hőmérséklet és az átlagterelés között úgy, hogy független változóként a hőmérséklet, függő változóként pedig a terhelés etalon (hideg napos ér-

tékek) feletti része lett figyelembe véve. A meleg, 14 °C-nál magasabb átlaghőmérsékletű napok előrejelzése tehát két részből tevődik össze. Van az ún. standard rész, amely a hőmérséklettől függetlenül jelentkező terhelés. Ehhez hozzáadódik egy közelítő függvénnyel kiszámolható, hőmérsékletfüggő rész. A becslés feltétele az ún. etalon, vagyis a hőmérséklettől függetlenül jelentkező terhelési standard előállítása. Ezt a hideg napos becslés alapján készítettük el, a zajos (szélsőséges) adatoktól és ünnepnapoktól mentes átlagértékek alapján. Itt a közelítést exponenciális illesztéssel és lineáris illesztéssel is elvégeztük. Bár az eredmények szignifikánsak, ennél jobb becslést kaptunk a másik módszer alkalmazásával.

Ebben a nappali órákra történő becsléshez függő változóként a teljes energiafelhasználást vittük be. Az eredmények jók, becslésre alkalmazhatók. Ennek alapján készítettük el a becslőtáblákat, amelyekben a nappali időszakra vonatkozó becsléshez a kapott függvénnyel becsüljük a várható energiafelhasználást. A közelítést lineáris regresszióval végeztük. A korrelációs együttható 0,902, ami szoros lineáris kapcsolatra utal. Determinációs koefficiens: 81,3%, azaz a napi átlagterhelés alakulását 81,3%-ban a napi átlaghőmérséklet határozza meg. Az ANOVA-tábla (2. táblázat) a korrelációs együtthatót teszteli. A megfigyelt szignifikanciaszint 0,000, tehát a korrelációs együttható szignifikánsan eltér 0-tól, ami erősíti a szoros lineáris kapcsolatot az átlagterhelés és az átlaghőmérséklet alakulása között.

2. táblázat

**A regressziós egyenes illesztésének eredménytáblái**  
*Results of the linear regression*

Modell	Eltérés-négyzetösszeg	Szabadság-fok	Szórás-négyzet	F	Szignifikancia
Regresszió	44 659,011	1	44 659,011	483,400	0,000
Maradék	10 254,768	111	92,385		
Teljes	54 913,779	112			

**Koefficiens**  
*Coefficients*

Modell	B	Standard hiba	t	Szignifikancia
Konstans	152,890	5,270	29,011	0,000
Hőmérséklet (X)	5,962	0,271	21,986	0,000

Megjegyzés: független változók: konstans, hőmérséklet; függő változó: energia (W).

Forrás: saját szerkesztés.

A lineáris kapcsolatot leíró egyenes paraméterei: a konstans értéke 152,89, a regressziós koefficiens, azaz a becslőegyes meredeksége 5,96.

A paraméterek értelmezése: 14 °C-os napi átlaghőmérséklet esetén 152,89 W az átlagterhelés.

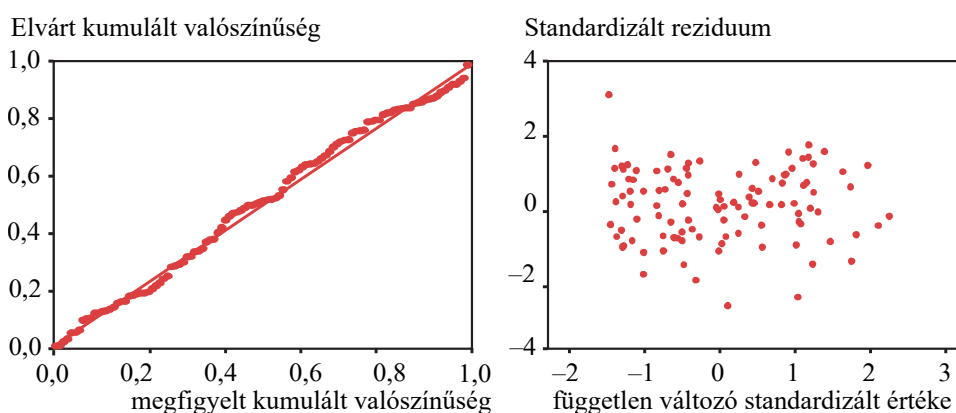
Ha a napi átlaghőmérséklet egy fokkal magasabb, az átlagterhelés 5,96 W-tal növekszik.

$$Y \text{ (becsült napi átlagterhelés)} = 152,89 + 5,96 \cdot X \text{ (napi átlaghőmérséklet)}$$

8. ábra

### A regressziós becslés alkalmazhatósági feltételeinek tesztelése

#### *Test of regression residuum*



Forrás: saját szerkesztés.

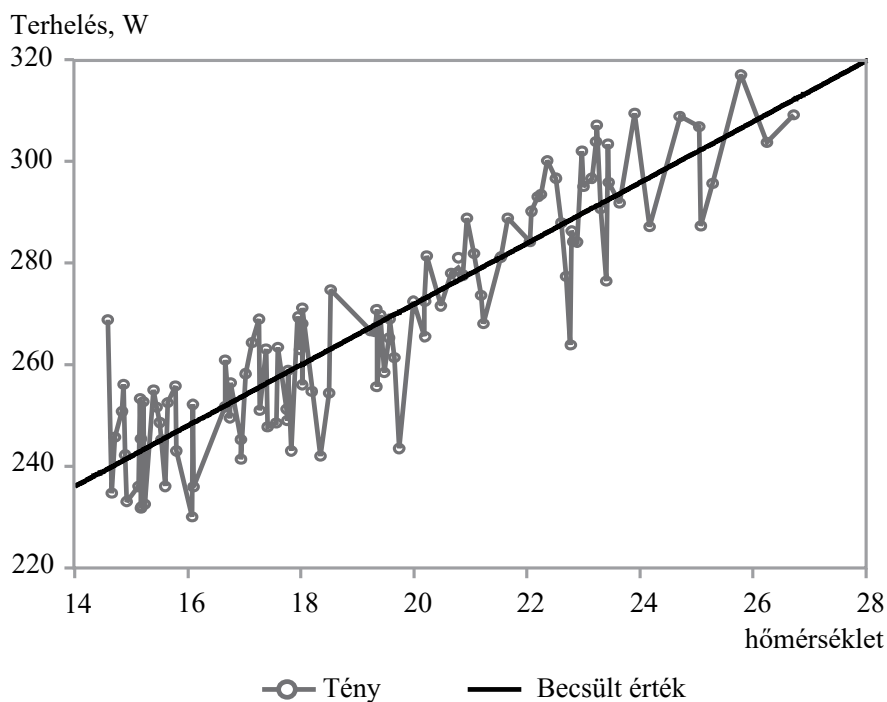
A regressziós becslés alkalmazhatósági feltételeinek tesztelésére szolgál a 8. ábra. A reziduumok normalitásvizsgálatára vonatkozik mindkettő. A bal oldali ábrán megközelítőleg az átló egyenesén kell a pontoknak elhelyezkedniük. Amennyiben az átló egyenesétől nagyfokú az eltérés, nemlineáris kapcsolatról vagy lineáris regresszióval nem közelíthető pontthalmazról van szó. Esetünkben a lineáris regressziós közelítés alkalmazható. A jobb oldali ábra a reziduumok szórását mutatja. A reziduumoknak véletlen eloszlásúnak, sűrűsödésmentesnek kell lenniük. Ez a feltétel is teljesül.

A 9. ábrán a megfigyelési értékeket és a közelítőegyenest látjuk.



9. ábra

**A hőmérséklet és a villamosenergia-fogyasztás kapcsolata 14 °C felett**  
*Relationship between temperature and average loading up to 14 °C degree*



Forrás: saját szerkesztés.

## 2.2. Negyedórás becslés naptípusonként

Egy lehetséges módszerként a napi értékből kivonjuk az etalont, és a fennmaradó rész és a hőmérséklet adatsora alapján megkeressük a legjobban közelítő függvényt. Az éjszakai időtartam hőmérséklet-független, a tényadatok átlagértékei alapján becsülhető. Ezzel a módszerrel a típusnapok közelítése nem hozta meg a várt eredményt, sok bizonytalanságot tartalmaz, és bár a determináltsága erősen szignifikáns, nem kielégítő (3. táblázat).

A három naptípusnál a korrelációs együtthatók hétfő esetében 0,778, a többi hétköznapra 0,820, vasárnapra pedig 0,792.

A két naptípusos változat esetében a hétfővel összevont hétköznapokra a korrelációs együttható 0,813, amely 66% körüli determináltságot jelent.

3. táblázat

**Összefüggés a hőmérséklet és az etalon feletti energiafelhasználás között**  
*Relationship between temperature and electricity consumption up to cool-days standard*

Napok	Korrelációs együttható	Determinációs együttható, %
Hétfő	0,7778	60,5
Kedd	0,8523	72,6
Szerda	0,8492	72,1
Csütörtök	0,7839	61,4
Péntek	0,7975	63,6
Szombat	0,8176	66,8
Vasárnap	0,7922	62,8

Forrás: saját számítás.

Másik módszerként a nappali órákra történő becsléshez függő változóként a teljes energiafelhasználást vittük be. Az eredmények jók, becslésre alkalmazhatók. Ennek alapján készítettük el a becslőtáblákat, amelyekben a nappali időszakra vonatkozó becsléshez a kapott függvénnyel becsljük a várható energiafelhasználást. A naptípusokra három különböző módon is elkészült a becslés: először kétféle bontásban (hétköznap és vasárnap szerint), másodsor minden napra (vagyis 7 naptípusra), és harmadszor háromkategóriás bontásban (hétfő különvételével, a többi hétköznap és vasárnap). A hét különálló naptípusos becslésnél jól látható, hogy hétfő és vasárnap közelítése nagyobb bizonytalansággal, a többi hétköznap pontosabb közelítéssel becsülhető. Összehasonlításhoz a hét típusnapos elemzés eredményeit mutatjuk be, a másik két csoportosításból a determinációs együtthatókat, illetve a paraméterbecslés táblázatát közöljük.

Mindkét módszernél többféle közelítő eljárást kipróbáltunk, végül a lineáris kapcsolat bizonyult a legerősebbnek, és egyben a legerősebb szignifikanciával is ez rendelkezik. A hét naptípusra készült elemzések eredményeit a 4. táblázat tartalmazza.

4. táblázat

**A hőmérséklet és energiafelhasználás közötti korrelációs együtthatók és determinációs együtthatók naptípusokon belül**

*Relationship between temperature and electricity consumption by daytypes*

Napok	Korrelációs együttható	Determinációs együttható, %
Hétfő	0,852	72,7
Kedd	0,919	84,4
Szerda	0,912	83,1
Csütörtök	0,887	78,6
Péntek	0,905	81,9
Szombat	0,895	80,2
Vasárnap	0,848	71,9

Megjegyzés: független változók: konstans, hőmérséklet; függő változó: energia (W). Forrás: saját szerkesztés.

A determinációs együtthatók erősen szignifikánsak.

A paraméterbecslések szignifikanciája 0,000. A közelítőegyenesek egyenlete:

**Hétfő:**  $Y(\text{energia}) = 171,079 + 6,13 \cdot X(\text{hőmérséklet})$

**Kedd:**  $Y(\text{energia}) = 178,75 + 5,97 \cdot X(\text{hőmérséklet})$

**Szerda:**  $Y(\text{energia}) = 164,06 + 6,64 \cdot X(\text{hőmérséklet})$

**Csütörtök:**  $Y(\text{energia}) = 154,84 + 7,01 \cdot X(\text{hőmérséklet})$

**Péntek:**  $Y(\text{energia}) = 155,889 + 7,189 \cdot X(\text{hőmérséklet})$

**Szombat:**  $Y(\text{energia}) = 161,72 + 6,834 \cdot X(\text{hőmérséklet})$

**Vasárnap:**  $Y(\text{energia}) = 164,50 + 6,265 \cdot X(\text{hőmérséklet})$

Összességében tehát azt tapasztaltuk, hogy 1 °C-os hőmérséklet-emelkedés az energiafogyasztásban körülbelül 6-7 watt közötti növekedést eredményez, függetlenül a naptípustól.

A sztenderdizált reziduumok (hibatagok) eloszlása normális eloszlású, a sztenderdizált becslült értékek és reziduumok pontfelhődiagramja véletlenszerűen szóródik, a regressziós becslés alkalmazható. A pontfelhődiagramról a szélsőséges, nem illeszkedő esetek felismerhetők, visszakérhetők.

Ezek azok a rendhagyó napok, amelyek nem illeszkednek a becslőfüggvényre, energiafelhasználásukat a hőmérsékleten kívül valamilyen más egyéb véletlen tényező is jelentősen befolyásolta pozitív vagy negatív irányban.

Amennyiben nem alkalmazunk naptípust az eljárásban, a hőmérséklet 78%-ban determinálja az energiafelhasználást (5. táblázat).

5. táblázat

#### A regressziós egyenes illesztésének eredménytáblái

*Results of the linear regression*

Modell	B	Standard hiba	t	Szignifikancia
Konstans	167,599	1,151	145,616	0,000
Hőmérséklet (X)	6,441	0,049	130,376	0,000

Megjegyzés: függő változó: energia (W).

Forrás: saját számítás.

Vagyis a közelítő egyenes egyenlete:  $Y(\text{energia}) = 167,6 + 6,4 \cdot X(\text{hőmérséklet})$ .

A végső eredmény egy interaktív modell. Az éjszakai órák becslése az átlagértékek alapján készül. A nappali időszakra vonatkozó becslés a regressziós függvény számítása alapján keletkezik. Az időponthoz tartozó hőmérsékleti előrejelzett adat módosításával automatikusan kiszámolható az energiafelhasználás becslült értéke.

### 2.3. Összegzés

1. Az ünnepnapok energiafelhasználása speciális, hasonló lefutású, a tapasztalati ünnepnapi fogyasztási értékkel becsülhető.
2. A hétköznapok energiafelhasználása naponként megkülönböztethetők, de szignifikáns különbség nem tapasztalható közöttük. A hétfő kismértékben alacsonyabb, majd a kedd magasabb értéke a jellemző, a szerda–csütörtök általában véve nem változik jellemzően, a péntek–szombat kismértékben magasabb, majd a vasárnap a legalacsonyabb értékű.
3. Megkülönböztethető az éjszaka és a nappal váltakozása. Az éjszaka reggel 6 óráig tart, a nappal 22 órakor ér véget. Vasárnap reggel 7 órától 21 óráig tart a nappali időszak.
4. Az éjszakának is van jellemző alakulása.
5. Naptípusok között éjszakai időszakban nem mutatható ki szignifikáns eltérés.
6. A nappali időszaknak rendkívül intenzív, jól jellemezhető görbéje van, hatodfokú polinommal jól közelíthető.
7. A hőmérséklet hatása csak  $+14\text{ °C}$  felett érvényesül, alacsonyabb hőmérséklet esetében nem mutatható ki kapcsolat a hőmérséklet és az energiafelhasználás között.

## 3. Elektromosenergia-igény előrejelzése mesterséges intelligencia (*deep learning*) módszerek segítségével

Az informatikai eszköztár és elméleti megismerés előrehaladtával a lineáris és nemlineáris statisztikai modellek mellett megjelentek a napjainkban „gépi tanulás” néven összefoglalt modellezési módszerek, amelyek elméleti alapjaikban a korábbiakkal osztoznak, ám kifejezőképességükben és rugalmasságukban új utakat nyitnak. A gépi tanulás kísérleti és elméleti kutatásokban való alkalmazása, az áramellátó rendszerek tervezése, az automatizált energiagazdálkodási rendszerek bevezetése széles lehetőségeket nyitott az elméleti fejlesztések gyakorlati felhasználására az adatfeldolgozás és -elemzés, valamint ezek modellezése és előrejelzése terén. Mindez lehetővé teszi annak a matematikai apparátusnak a szélesebb körű alkalmazását, amely korábban a számítási eljárások megvalósításának nehézségei, valamint az automatizált energiagazdálkodási rendszerek fejlődésének elmaradása miatt csak elméletileg fejlődhetett (*Klyuev et al., 2022*).

A gépi tanulás és ezen belül alágként a neurális hálózatok használata hatékony előrejelzési eszközök. Ez azonban nem jelenti azt, hogy az intelligens módszerek minden esetben pontosabb eredményt mutatnak az elektromos terhelés előrejelzésére.

### 3.1. Modellezési megközelítés

A „klasszikus”, tehát nem neurális tanulási módszerekkel szemben az idősor-előrejelzéshez a *deep learning* (DL-) módszerek számos előnnyel rendelkeznek. A hagyományos modellekkel ellentétben, amelyek gyakran linearitást feltételeznek, a DL-modellek – mint például a Recurrent Neural Networks (RNN-ek) és a Transformers-ek – kiválóan alkalmasak a nemlineáris idősorok modellezésére. A többváltozós előrejelzésekhez a DL-módszerek kiválóan alkalmasak több, egymástól függő idősor egyidejű kezelésére, a sorozatok közötti interakciók rögzítésére anélkül, hogy minden változóhoz független modellre lenne szükség. Végül, a DL-architektúrák előnyben részesítik az előtanítási (*pre-training*) és átviteli tanulási (*transfer learning*) képességeket, amelyek lehetővé teszik a nagy, hasonló adatkészleteken előzetesen betanított modellek finomhangolását az adott alkalmazásokhoz, csökkentve a pontos idősor-előrejelzők fejlesztéséhez szükséges adat- és időigényt. Ezek a jellemzők a DL-módszereket különösen előnyössé teszik összetett, valós idősoros feladatokhoz.

Tekintettel arra, hogy a *deep learning*<sup>1</sup> nem alkalmazható bizonyos számú adatpont alatt (és a teljesítményét kifejezetten össze akartuk hasonlítani a klasszikusabb módszerekkel), megtartottuk a 15 perces intervallumot a bemeneti adatoknál és a várható kimeneteknél egyaránt, és az üzleti követelményeknek megfelelően 48 óras (tehát  $48 \cdot 4$  intervallumú) előrejelzést készítettünk.

Az adatkészletet felosztottuk egy modellalkotási és egy validálási részre. A modellalkotáshoz a 2008. május 18. és a 2008. július 11. közötti időszakot használtuk, míg a modell teszteléséhez, validálásához a 2008. július 11. utáni időszakot alkalmaztuk.

A legegyszerűbb alapmodellünk az autoregresszív integrált mozgóátlag (ARIMA-) (*Box-Jenkins, 1970*) modell volt, konkrétan az „Auto-ARIMA” módszerrel (*Herzen et al., 2022*). Mivel ez a módszer nem adott megfelelő eredményt, a következő elemzés egy egyszerű lineáris regressziós modell volt (*Pedregosa et al., 2011*), amelybe időkódváltozót, valamint hőmérsékleti jellemzőket is beépítettünk.

<sup>1</sup> A *deep learning* (mély tanulás) alkalmazhatósági feltételeiről lásd *Alzubaidi és szerzőtársai (2021)*.

Kiinduló DL-modellként egy specifikusan idősoros adatok feldolgozására kifejlesztett mély neurális hálós modellt tanítottunk end-to-end módon, azaz szabályos véletlenszerű paraméter-inicializációval és gradiens leszállással csupán a specifikus adatunkon. Ezt a DL-architektúrát, a TiDE-<sup>2</sup> (*Das et al., 2023*) modellt kifejezetten hosszú távú előrejelzésre tervezték, így különösen alkalmas a 48 órás, 15 perces időközönkénti előrejelzésre, amely 192 előrejelzett adatot jelent. Ezt követően az előtanulási és finomhangolási paradigma (amelyet pl. *Howard és Ruder {2018}* képvisel a Natural Language Processing területen) alkalmazhatóságának teszteléséhez *Rasul és szerzőtársai (2024)* Lag-Llama-modelljét használtuk.<sup>3</sup> Utóbbi egy megalapozó modell, amelyet egyváltozós valószínűségi idősoros előrejelzésre terveztek. A modell a korábbi idősorértékekből készletetett jellemzőket dolgoz fel kihasználva az idősoros adatokban rejlő időbeli függőségeket. Lényeges módszertani eleme, hogy roppant méretű, diverz adatsor halmazon történő „előtanítása” után belső paramétereiben jelentős előrejelző erőt képviselő információ halmozódik fel, azaz „általános” idősoros modellnek tekinthető (párhuzamosan a nyelvfeldolgozási területen ismert előtanított „nagy nyelvi modellek” általános nyelvi képességeivel), így alkotói eredeti szándéka szerint alkalmas lehet specifikus adaton történő tanítás nélkül (*zero shot prediction*) vagy csupán finomhangolás (*pretraining and finetuning*) segítségével előrejelzést adni nem látott idősoros adatokra. Ez a potenciálja teszi megalapozó modellé.

A gyakorlati alkalmazhatóság érdekében a Lag-Llama implementációja a pusztán előrejelzésen túl támogatja az egyes adatkészleteken történő finomhangolást is. A finomhangolás az előre betanított modell paramétereit az új adatok jellemzőihez igazítja, jelentősen javítva az előrejelzési pontosságot. A hatékony finomhangolás képessége a Lag-Llamát az idősoros előrejelzés sokoldalú eszközévé teszi, ami rövid és hosszú távú előrejelzésekre egyaránt alkalmas, beleértve a 48 órás előrejelzési időtávunkat is. Meg kell jegyezni azonban, hogy a Lag-Llama egy egyváltozós modell, tehát bizonyos értelemben „fogyatékos”, mivel nem lehet bemenetként megadni az ünnepi vagy hőmérsékleti adatokat. A legújabb megalapozó modellek, például *Cao és szerzőtársai (2024)* TimeDiT-modellje már többváltozós bemeneteket is képes kezelni. E kutatási irány területünkre történő alkalmazása további vizsgálatot igényel.

<sup>2</sup> Time series Dense Encoder.

<sup>3</sup> *Rasul et al. (2024)* Lag-Llama modellje lásd *Garcia és munkatársai (2024)*.

### 3.2. Eredmények

A modelleredményeket szemügyre véve, a modellek eredményeit a validációs adatokon összevetve látható a TiDE-modell jó teljesítménye, így megállapítható, hogy még az általunk végzett sokjellemzős tervezés mellett is van néhány figyelemre-méltó nemlineáris mintázat az idősorban, amelyeket a TiDE-modell hasznosítani tud.

Megfigyelhetjük, hogy még a lineáris regressziós modell is képes összetett időfüggő összefüggések rögzítésére, bár a TiDE-modell határozottan dominál abban, hogy képes alkalmazkodni az energiaigény változékonyságához.

A valóban meglepő eredmény azonban az előtanított Lag-Llama-modell esetében figyelhető meg: Egyrészt *Rasul és szerzőtársai (2024)* eredeti megközelítése és a megalapozó modellek legbátrabb ígérete, miszerint általános modellként az adaton történő finomhangolás nélkül tudnak egy nem látott idősoron is használható előrejelzéseket adni, ez esetben könnyen cáfolható. A modell pusztán tanítás nélküli alkalmazása nem vezet értékelhető eredményre. Másrészt azonban egy igen ígéretes megfigyelés is tehető: a Lag-Llama-modell finomhangolást követően figyelemre-méltó és rendkívül biztató eredményt mutat. A finomhangolt modell teljesítménye kimagaslóan jó, és még alacsonyabb tanítási számítási igény mellett is határozottan felülmúlja a kifejezetten az adatunkon tanított TiDE-modellt (6. táblázat).

A kísérletek azt igazolják, hogy feladatspecifikus finomhangolásra van szükség, amely során a modell újra felhasználhatja a korábban tanult tudást a jobb predikciós eredmények érdekében, ezzel azonban a többi módszerhez képest igen jelentős teljesítménynövekedésre számíthatunk (Függelék F5. ábra).

6. táblázat

**A modelleredmények összevetése**  
*Comparing model results*

Modell	MAE	MAPE	$R^2$
Linear regression	25,90	9,05	0,698
TiDE	22,81	8,17	0,768
Lag-Llama (“zero shot”)	43,90	17,32	0,10
Lag-Llama (finomhangolt)	13,25	4,81	0,89

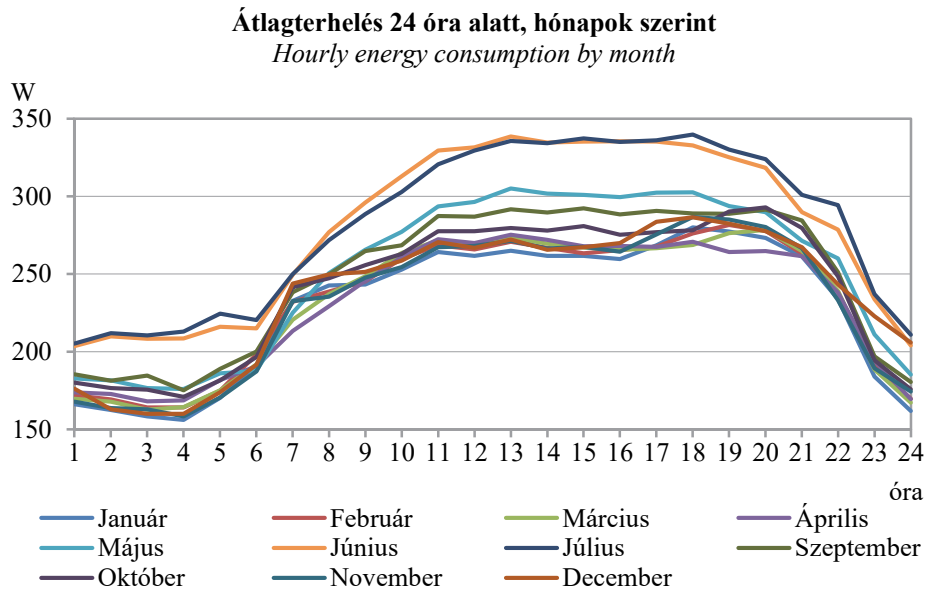
### 3.3. Következtetés és ajánlások

Bár az adathalmaz különböző részein, valamint más adatforrásokon még szélesebb körű validációra van szükség, a jelenlegi eredmények azt mutatják, hogy az idősor-modellre épülő gépi tanulás és finomhangolás alkalmazása versenyképes lehet. Az alkalmazható modellek hozzájárulhatnak a költségmegtakarításhoz és az éghajlat-tal kapcsolatos kihívások mérsékléséhez. A bemutatott eredmények aláhúzzák az idősoros előrejelzés szélesebb alkalmazásának jelentőségét a gazdasági döntésho-zatal szempontjából, különösen azokon a területeken, ahol a pontos, adatvezérelt előrejelzés elengedhetetlen. A nagy adatkészletekre előre betanított alapmodellek megjelenése, valamint a modellek kisebb, tartományspecifikus adatokon történő finomhangolása egy forradalmi lépést jelent. Ahogy azt ebben a tanulmányban be-mutattuk, az adatkészlet finomhangolása versenyképes eredményeket hozhat, ami az alapvető *deep learning* modelleket hozzáférhetővé, skálázhatóvá és rendkívül relevánssá teszi a különböző gazdasági és környezetvédelmi alkalmazásokban.



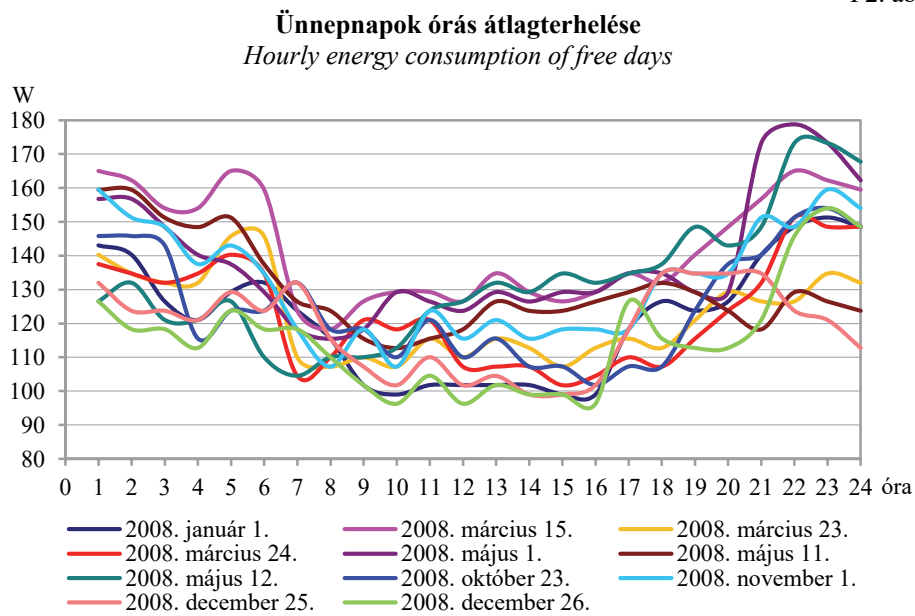
## Függelék

F1. ábra



Forrás: saját szerkesztés.

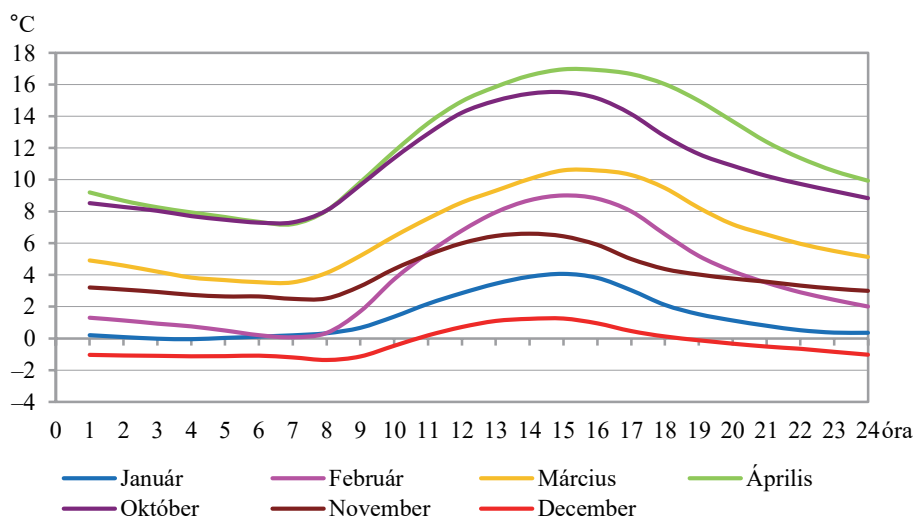
F2. ábra



Forrás: saját szerkesztés.

F3. ábra

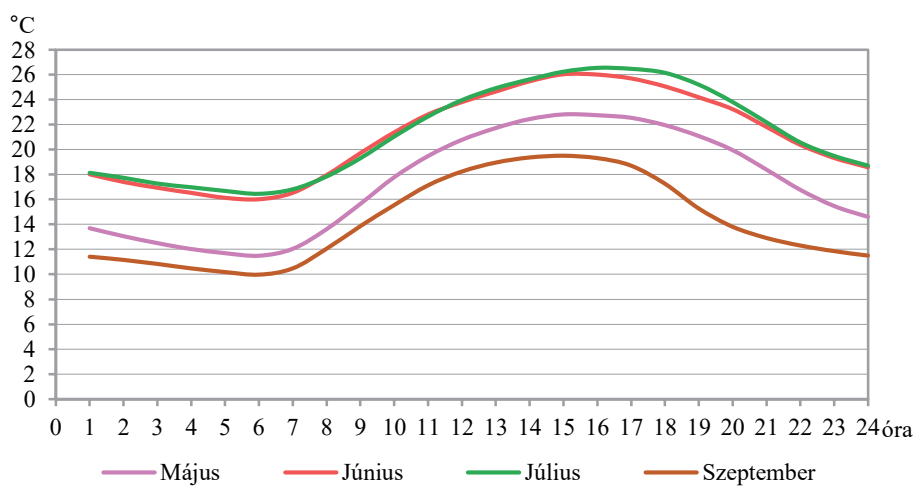
**Órás átlaghőmérséklet alakulása havonként (téli/hideg hónapok)**  
*Hourly average temperature by month (cold)*



Forrás: saját számítás alapján szerkesztés.

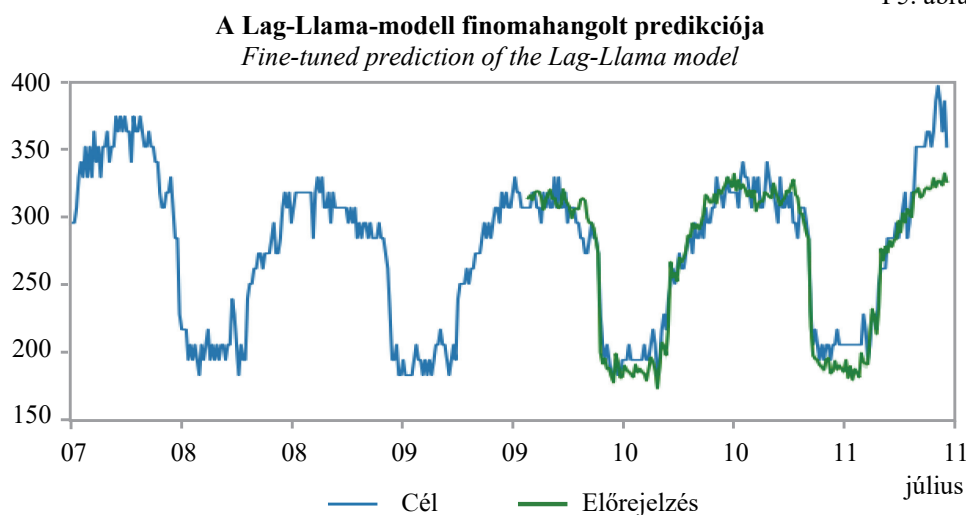
F4. ábra

**Órás átlaghőmérséklet alakulása havonként (nyári/meleg hónapok)**  
*Hourly average temperature by month (hot)*



Forrás: saját számítás alapján saját szerkesztés.

F5. ábra



## Irodalom

- Alzubaidi, L. – Zhang, J. – Humaidi, A.J. – Al-Duajili, A. – Duan, Y. – Al-Shamma, O. – Santamria, J. – Fadhel, M. A. – Al-Amidie, M. – Farhan, L. (2021): Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data* 8, 53. <https://doi.org/10.1186/s40537-021-00444-8>
- Ansari, A. F. – Stella, L. – Turkmen, C. – Zhang, X. – Mercado, P. – Shen, H. – Shchur, O. – Rangapuram, S. S. – Arango, S. P. – Kapoor, S. – Zschiegner, J. – Maddix, D. C. – Wang, H. – Mahoney, M. W. – Torkolla, K. – Gordon, A. W. – Bohlke-Schneider, M. – Wang, Y. (2024): Chronos: Learning the language of time series. *arXiv:2403.07815*. <https://doi.org/10.48550/arXiv.2403.07815>
- Biber, A. – Felder, M. – Wieland, C. – Spliethoff, H. (2022): Negative price spiral caused by renewables? electricity price prediction on the german market for 2030. *The Electricity Journal*, 35(8), 107188. <https://doi.org/10.1016/j.tej.2022.107188>
- Bommasani, R. – Hudson, D. A. – Adeli, E. – Altman, R. – Arora, S. – Arx, S. – Bernstein, M. S. – Bohg, J. – Bosselut, A. – Brunskill, E. – Brynjolfsson, E. – Buch, S. – Card, D. – Castellon, R. – Chatterji, N. S. – Chen, A. S. – Creel, K. – Davis, J. Q. – Demszky, D. – Donahue, C. – Doumbouya, M. – Durmus, E. – Ermon, S. – Etchemendy, J. – Ethayarajh, K. – Fei-Fei, L. – Finn, C. – Gale, T. – Gillespie, L. E. – Goel, K. – Goodman, N. D. – Grossman, S. – Guha, N. – Hashimoto, T. – Henderson, P. – Hewitt, J. – Ho, D. E. – Hong, J. – Hsu, K. – Huang, J. – Icard, T. – Jain, S. – Jurafsky, D. – Kalluri, P. – Karamcheti, S. – Keeling, G. – Khani, F. – Khattab, O. – Koh, P. W. – Krass, M. S. – Krishna, R. – Kuditipudi, R. – Kumar, A. – Ladhak, F. – Lee, M. – Lee, T. – Leskovec, J. – Levent, I. – Li, X. L. – Li, X. – Ma, T. – Malik, A. – Manning, C. D. – Mirchandani, S. P. – Mitchell, E. – Munyikwa, Z. – Nair, S. – Narayan, A. – Narayanan, D. – Newman, B. – Nie, A. – Niebles, J. C. – Nilforoshan, H. – Nyarko, J. F. – Ogut, G. – Orr, L. – Papadimitriou, I. – Park, J. S. – Piech, C. – Portelance, E. – Potts, C. – Raghunathan, A. – Reich, R. – Ren, H. – Rong, F. – Roohani, Y. H. – Ruiz, C. – Ryan, J. – R'e, C. – Sadigh, D. – Sagawa,

- S. – Santhanam, K. – Shih, A. – Srinivasan, K. P. – Tamkin, A. – Taori, R. – Thomas, A. W. – Tram'er, F. – Wang, R. E. – Wang, W. – Wu, B. – Wu, J. – Wu, Y. – Xie, S. M. – Yasunaga, M. – You, J. – Zaharia, M. A. – Zhang, M. – Zhang, T. – Zhang, X. – Zhang, Y. – Zheng, L. – Zhou, K. – Liang, P. (2021): On the opportunities and risks of foundation models. *CoRR*, *arXiv:2108.07258*. <https://doi.org/10.48550/arXiv.2108.07258>
- Box, G. E. P. – Jenkins, G. M. (1970): *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.
- Cao, D. – Ye, W. – Zhang, Y. – Liu, Y. (2024): TimeDiT: General-purpose diffusion transformers for time series foundation model. *arXiv*. <https://arxiv.org/abs/2409.02322>
- Das, A. – Kong, W. – Leach, A. – Mathur, S. – Sen, R. – Yu, R. (2023): Long-term forecasting with TIDE: Time-series dense encoder. <https://doi.org/10.48550/arXiv.2304.08424>
- Das, A. – Kong, W. – Sen, R. – Zhou, Y. (2024): A decoder-only foundation model for time-series forecasting. <https://doi.org/10.48550/arXiv.2310.10688>
- Garcia, C. – Yang, L. – Zhang, M. – Chen, Y. – Patel, S. – Nguyen, T. (2024): Lag-llama: Time-series foundation models. <https://arxiv.org/pdf/2310.08278> (letöltve: 2024. május)
- Garza, A., – Mergenthaler-Canseco, M. (2023): Timegpt-1. *arXiv.2310.03589*. <https://doi.org/10.48550/arXiv.2310.03589>
- Goswami, M. – Szafer, K. – Choudhry, A. – Cai, Y. – Li, S. – Dubrawski, A. (2024): Moment: A family of open time-series foundation models. *arXiv:2402.03885*. <https://doi.org/10.48550/arXiv.2402.03885>
- Granger, C. (1969): Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3), 424–438. <https://doi.org/10.2307/1912791>
- Herzen, J. – Lässig, F. – Piazzetta, S. G. – Neuer, T. – Tafti, L. – Raille, G. – Pottelbergh, T. – Pasička, M. – Skrodzki, A. – Huguenin, N. – Dumonal, M. – Kościsz, J. – Bader, D. – Gusset, F. – Benheddi, M. – Williamson, C. – Kosinski, M. – Petrik, M. – Grosch, G. (2022): Darts: User-friendly modern machine learning for time series. *Journal of Machine Learning Research*, 23(124), 1–6. [arXiv:2110.03224](https://doi.org/10.48550/arXiv.2110.03224). <https://doi.org/10.48550/arXiv.2110.03224> (letöltve: 2024. május)
- Howard, J. – Ruder, S. (2018): Universal Language Model Fine-tuning for Text Classification. In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 328–339, Melbourne, Australia. Association for Computational Linguistics. <http://doi.org/10.18653/v1/P18-1031>
- Klyuev, R. V. – Morgoev, I. D. – Morgoeva, A. D. – Gavrina, O. – Martyushev, N. – Efremenkov, E. A. – Mengxu Q. (2022): Methods of forecasting electric energy consumption: A literature review. *Energies*, 15(23), 8919.
- Pedregosa, F. – Varoquaux, G. – Gramfort, A. – Michel, V. – Thirion, B. – Grisel, O. – Blondel, M. – Prettenhofer, P. – Weiss, R. – Dubourg, V. – Vanderplas, J. – Passos, A. – Cournapeau, D. – Brucher, M. – Perrot, M. – Duchesnay, E. (2011): Scikit-learn: Machine learning in python *Journal of machine learning research*, 12(Oct), 2825–2830. <https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf> (letöltve: 2024. május)
- Winters P. R. (1960): Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3), 324–342. <https://doi.org/10.1287/mnsc.6.3.324>
- Woo, G. – Liu, Ch. – Kumar, A. – Xiong, C. – Savarese, S. – Sahoo, D. (2024): Unified training of universal time series forecasting transformers. *Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235*, <https://doi.org/10.48550/arXiv.2402.02592>, <https://doi.org/10.48550/arXiv.2402.02592>
- Ye, J. – Zhang, W. – Yi, K. – Yongzi, Y. – Zijue, L. – Jia, L. – Fugee, T. (2024): A survey of time series foundation models: Generalizing time series representation with large language model. <https://doi.org/10.48550/arXiv.2405.02358>