

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR

Varga Melinda

**A CO_2 KIBOCSÁTÁS MODELLEZÉSE ÉS
ELŐREJELZÉSE**

Szakkdolgozat

Matematika BSc, matematikai elemző szakirány

Témavezető:

Zempléni András egyetemi docens

Valószínűségelméleti és Statisztika Tanszék



Budapest, 2020

NYILATKOZAT

Név: Varga Melinda

ELTE Természettudományi Kar, szak: Matematika Bsc.

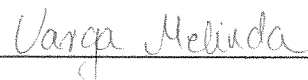
NEPTUN azonosító: VIIQCY

Szakdolgozat címe:

A CO₂ kibocsátás modellezése és előrejelzése

A **szakdolgozat** szerzőjeként fegyelmi felelősségem tudatában kijelentem, hogy a dolgozatom önálló szellemi alkotásom, abban a hivatkozások és idézések standard szabályait következetesen alkalmaztam, mások által írt részeket a megfelelő idézés nélkül nem használtam fel.

Budapest, 2020. május 29.



a hallgató aláírása

Köszönetnyilvánítás

Szeretném megköszönni témavezetőmnek, Zempléni Andrásnak a téma felvetését, megannyi segítségét, útmutatását és idejét a dolgozatom megírásához.

Továbbá köszönettel tartozom családomnak, akik a kezdetektől fogva mellettem álltak; barátaimnak, akikkel egymás támogattuk az évek során; és eddigi tanáraimnak, akiknek tudása és segítőkészsége segített eljutni idáig.

Tartalomjegyzék

1. Előszó	11
2. A kibocsátási ágazatok és adatok áttekintése	12
2.1. Kibocsátási ágazatok	12
2.2. Országok emissziója	15
2.2.1. Kibocsátás világviszonylatban	15
2.2.2. Az Európai Unió emissziós értékei és országonkénti eloszlása	18
2.2.3. A kibocsátási értékek időbeli alakulása	22
3. Lehetséges elemzési módszerek	27
3.1. Artificial Neural Network (Mesterséges neurális háló)	27
3.2. Lineáris regresszió és többszörös lineáris regresszió	28
4. Idősorok	29
4.1. Elméleti áttekintés és alapvető fogalmak	29
4.2. Az ARMA modell felépítése	31
4.3. Az ARIMA modell felépítése	32
4.4. Becslési módszerek	33
4.5. Illeszkedésvizsgálatok	34
4.5.1. Box-Pierce teszt	34
4.5.2. Ljung-Box teszt	34

5. Modellezés és előrejelzés	35
5.1. Az összes kibocsátás elemzése és előrejelzése	35
5.1.1. Az Európai Unió	35
5.1.2. Magyarország	40
5.2. Magyarország egy főre jutó kibocsátásának elemzése és előrejelzése	44
6. Konklúzió	48
Irodalomjegyzék	49
Függelék	52

Ábrák jegyzéke

2.1. A kiválasztott országok 1940-es és 2017-es kibocsátásának összehasonlítása	16
2.2. Az EU tagországok 1940-es és 2017-es kibocsátásainak összehasonlítása	21
2.3. Az országok összes kibocsátásának változása	22
2.4. Az országok egy főre jutó kibocsátásának változása	23
2.5. Az EU néhány tagországának éves CO_2 kibocsátásának változása	23
2.6. Néhány tagország egy főre jutó kibocsátásának változása	24
2.7. A világ és az EU kibocsátásának változása	24
2.8. Magyarország kibocsátásainak alakulása	25
5.1. Az EU adatainak autokovariancia vizsgálata	36
5.2. Az EU kibocsátására illesztett kvadratikus regresszió	37
5.3. Az EU kibocsátásának előrejelzése ARIMA(2,0,1) modellel 99%-os konfidenciaszinten	39
5.4. Az EU kibocsátásának előrejelzése ARIMA(0,2,1) modellel 99%-os konfidenciaszinten	39
5.5. Magyarország adatainak autokovariancia vizsgálata	40
5.6. Magyarország kibocsátására illesztett kvadratikus regresszió	41
5.7. Hazánk kibocsátásának előrejelzése ARIMA(3,0,1) modellel (fekete szakasz) 99%-os konfidenciaintervallumon	43
5.8. Hazánk kibocsátásának előrejelzése ARIMA(0,2,1) modellel 99%-os konfidenciaintervallumon	43
5.9. Hazánk egy főre eső kibocsátásának autokovariancia vizsgálata	44
5.10. Egy főre jutó kibocsátására illesztett kvadratikus regresszió	45

5.11. Egy főre jutó kibocsátás előrejelzése ARIMA(3,0,1) modellel 99%-os konfidenciaszinten	47
5.12. Egy főre jutó kibocsátás előrejelzése ARIMA(0,2,1) modellel 99%-os konfidenciaszinten	47

Táblázatok jegyzéke

2.1. Magyarország gazdasági ágazatai CO_2 kibocsátásuk szerint rendezve , a kibocsátás millió tonnában megadva (2017)	13
2.2. Magyarország ipari ágazatai CO_2 kibocsátásuk szerint rendezve millió tonnában (2017)	14
2.3. 1940-es kibocsátás pár adata	15
2.4. 2017-es kibocsátás legfőbb résztvevői	16
2.5. Az Európai Unió tagországok ismert kibocsátásai (1940)	18
2.6. Az "Európai Unió" 2017-es kibocsátása	20
5.1. Az EU kvadratikus regressziójának együtthatói	36
5.2. Az EU_ARIMAmodell modell együtthatói és standard hibáik . . .	36
5.3. Az EU_ARIMAmodell2 modell együtthatói és standard hibáik . .	37
5.4. ARIMA (2,0,1) és ARIMA (0,2,1) összehasonlítása	37
5.5. Az EU kibocsátásának regresszió alapuló becslése és korlátai . .	38
5.6. Az EU kibocsátásának idősoron alapuló becslése és korlátai . . .	38
5.7. Magyarország kvadratikus regresszióvizsgálatának együtthatói . .	40
5.8. Az EU kvadratikus regressziójának együtthatói	41
5.9. Az ARIMAmodell modell együtthatói és standard hibáik	41
5.10. Az ARIMAmodell2 modell együtthatói és standard hibáik	41
5.11. ARIMA (3,0,1) és ARIMA (0,2,1) összehasonlítása	42
5.12. Hazánk regresszió alapuló kibocsátásának becslése és korlátai . .	42
5.13. Hazánk idősoron alapuló kibocsátásának becslése és korlátai . . .	42
5.14. Egy főre jutó kibocsátás kvadratikus regresszióvizsgálatának együtthatói	44

5.15. Az egyfo_ARIMAmodell modell együtthatói és standard hibáik .	45
5.16. Az egyfo_ARIMAmodell2 modell együtthatói és standard hibáik .	45
5.17. Egy főre jutó kibocsátás ARIMA (3,0,1) és ARIMA (0,2,1) összehasonlítása	45
5.18. Egy főre jutó kibocsátás regresszió alapuló becslése és korlátai .	46
5.19. Egy főre jutó kibocsátás időszoron alapuló becslése és korlátai . . .	46

1. fejezet

Előszó

Az üvegházhatású gázok (ÜHG) olyan gázok, melyek az üvegházhatások előidézéséért és így a globális felmelegedésért felelősek. A légkörben felhalmozódva átengedik a Napból érkező rövid hullámhosszúságú elektromágneses sugarakat, így azok eléri és felmelegítik a földfelszínt. A felszín az így keletkező energiatöbbletet hosszú hullámhosszúságú sugárzás formájában veri vissza, ám ezt az üvegházhatású gázok már elnyelik, és így a légkör felmelegszi.

A üvegházhatású gázok a természetben is előfordulnak, ám a problémát az antropogén eredetű, azaz emberi tevékenység hatására nagy mennyiségben légkörbe kerülő gázok okozzák. Ilyen gázok például a vízgőz, a szén-dioxid (CO_2), a metán (CH_4), a dinitrogén-oxid (N_2O), illetve a mesterséges fluorozott szénhidrogének (HFC).

Az egyik legjelentősebb üvegházhatású, mennyisége miatt a legnagyobb környezeti problémát okozó gáz a szén-dioxid, aminek kibocsátás-csökkentése elvezethet a problémák megoldásához. Dolgozatomban a CO_2 kibocsátás modellezését és lehetséges előrejelzését fogom bemutatni, főként az Európai Unióra és Magyarországra koncentrálva.

2. fejezet

A kibocsátási ágazatok és adatok áttekintése

2.1. Kibocsátási ágazatok

Az ipari forradalom hatására fokozatosan nőtt az üvegházhatású gázok mennyisége, köszönhetően a fosszilis energiahordozók (szén, kőolaj, földgáz) égetésének. Ez még a mai napig előkelő helyet foglal el (világviszonylatban és országonként is) a CO_2 kibocsátási rangsorban a villamosenergia- és hőtermelési folyamatok miatt.

Az ipari kibocsátás több ágazatból tevődik össze, ugyanakkor külön-külön és összesen is nagy kárt okoznak. Ilyen például az energiafogyasztáshoz kapcsolódó kémiai és kohászati folyamatokból származó kibocsátás, a feldolgozóipari és az épületipari kibocsátás.

Ugyancsak CO_2 -t és más üvegházhatású gázt termelnek a mezőgazdasági folyamatok (növénytermesztés, állattenyésztés), a halászat, a bányászat, de még a vízellátás, a szennyvíz gyűjtése és kezelése, és a hulladékkezelés is.

Megjegyzés 2.1.1 *Növénytermesztés és üvegházhatás: a növények a fotoszintézis során úgymond "lekötik" a CO_2 -t, és így kvázi csökkentik az üvegházhatást. Mindazonáltal a termesztés során használt műtárgya megemeli a légköri nitrogén-oxid szintjét, ami például savas esőként visszakerül a talajra, vizekbe, növényekre. Ezenkívül a mezőgazdasági szállítmányozás is növeli a CO_2 kibocsátást. Így a növénytermesztés közvetlen és közvetett módon is növeli (növelheti) az üvegházhatást.*

Nem meglepő módon a szállítási folyamatok, a kereskedelem, illetve a közlekedés során is rengeteg gáz szabadul fel: közúti, vasúti, légi és tengeri szállítás

során felhasznált fosszilis tüzelőanyagok égetése miatt. A világ közlekedési energiafelhasználásának szinte teljes egésze a fosszilis üzemanyagokból származik, habár a technológiai fejlődés a jövőben ezt visszaszoríthatja (elektromos autók, buszok). Szállítmányozási folyamatok során nem csak maga a szállítás, de a raktározás is komoly kibocsátásokat eredményezhet.

Ezekon kívül vannak olyan területek, amikről először nem is gondolnánk, hogy emissziós rátájuk fontos tényező. Ilyenek az adminisztratív és szolgáltatást támogató tevékenységek, az információs és kommunikációs ágazatok, a humán egészségügy, de még az oktatás is. Úgyszintén fontos megemlíteni a háztartások kibocsátásait, melyek talán családonként nem magasak, de összességükben már komoly gondot okoznak, az országok és régiók fejlettségétől függően.

A jobb átláthatóság érdekében készítettem egy összesítő táblázatot, mely a 2017-es magyarországi adatokat (a Központi Statisztikai Hivatal adatbázisa alapján [11]) és a teljes kibocsátáshoz képesti arányukat tartalmazza. A kibocsátási értékek "millió tonna" mértékegységben vannak.

Kibocsátási faktorkok	2017-es kibocsátás	Részarány (%)
Ipari kibocsátás	28,49	43,94
Szállítás, raktározás	4,38	4,3
Mezőgazdaság, erdőgazdálkodás, halászat, bányászat	2,79	4,31
Kereskedelem, gépjárműjavítás	1,96	3,02
Szolgáltató szektor, oktatás, kommunikáció, pénzügy	1,67	2,58
Humán egészségügy	0,27	0,41
Egyéb nemzetgazdasági területek	1,78	2,74
Háztartások	23,29	35,92
Összes	64,83	100

2.1. táblázat. Magyarország gazdasági ágazatai CO_2 kibocsátásuk szerint rendezve , a kibocsátás millió tonnában megadva (2017)

A táblázatból jól kivehető, hogy az összes kibocsátás több, mint 40 %-át az ipari kibocsátások, 35 %-át a háztartások adják. A többi szektor emissziója szinte eltörpül ezek mellett, ám korántsem elhanyagolhatóak összességükben.

Fontos megemlíteni, hogy az ipari kibocsátás a fenti táblázatban gyűjtőfogalomként van jelen, több ágazatból tevődik össze, melyek kibocsátási értékei és azok aránya a lenti táblázatban láthatóak.

Ipari ágazat	Kibocsátás	Részarány (%)
Villamosenergia-, gáz-, gőzellátás, légkondicionálás	15,35	53,88
Feldolgozóipar	12,06	42,32
Építőipar	1,08	3,8
Összesen	28,49	100

2.2. táblázat. Magyarország ipari ágazatai CO_2 kibocsátásuk szerint rendezve millió tonnában (2017)

Jól látszik, hogy a villamosenergia-, gáz- és gőzellátás, valamint a légkondicionálás az ipari kibocsátás több, mint felét adja. Ugyancsak magas rátával rendelkezik a feldolgozóipar. Ez az országok fejlettségi szintjével áll összefüggésben, és így a fejlett országok kibocsátása magasabb, mint a fejlődő országoké. Jó példa erre Oroszország és Jemen. Mindkettő olajkitermelő-és finomító ország, de míg Oroszország a G20-ak tagja, addig Jemen az arab világ egyik legfejletlenebb régiója. Az "ourworldindata.org" [9] adatbázisa alapján Oroszország 2017-es egy főre jutó CO_2 kibocsátása 11,76 tonna/fő, Jemen egy főre jutó kibocsátása pedig 0,67 tonna/fő volt. A következőkben pár ország, köztük Oroszország kibocsátásainak időbeli alakulásáról is szó lesz.

2.2. Országok emissziója

Az emissziós faktorokat és a konkrét (éves) kibocsátásokat megvizsgálva könnyen felállíthatunk egy sorrendet, mely a világ országait sorolja fel a kibocsátott CO_2 mennyiség alapján. Az adatok vizsgálatát leszűkítettem az 1940. és 2017. közti adatokra, mivel a későbbiekben ezen évek adatainak segítségével szeretném majd a jövőbeli előrejelzéseket levezetni. Így először ezt a két végpontot vettem jobbra szemügyre. A kibocsátási adatokat az "ourworldindata.org" [9] honlapról töltöttem le (csak 2017-ig vannak adataik), a százalékos eloszlást azonban én számoltam ki, a jobb átláthatóság érdekében.

2.2.1. Kibocsátás világviszonylatban

Először is az 1940-es és a 2017-es összes kibocsátást, valamint 2017. adatai alapján a 10 legnagyobb emisszióval rendelkező országokat és államokat mutatom be szemléltetésképpen, illetve az egy főre eső értékeket is, melyeket könnyen megkapunk, ha elosztjuk az összes kibocsátást az adott ország népességével. A kibocsátási értékeket millió tonna/év mértékegységben, az egy főre vetített értékeket pedig tonna/fő mértékegységben adom meg, az egész dolgozatban. A sorrend az összes kibocsátás szerinti csökkenő sorrendben van. Vannak országok, ahol az 1940. évi kibocsátás nem ismert, ott NA (Not Available) szerepel.

Országok	Hivatalos rövidítés	1940-es kibocsátás (millió tonna)	Részarány (%)	Egy főre jutó kibocsátás (tonna/fő)
Világ összesen	-	4 904	-	2,13
Egyesült Államok	USA	1 871, 49	38,16	13, 84
Japán	JPN	153, 26	3,13	2,09
Kanada	CAN	108, 63	2,22	9,32
Kína	CHN	86, 10	1,76	0,17
India	IND	57,70	1,18	0,18
Irán	IRN	27,13	0,55	1,80
Mexikó	MEX	23,21	0,47	1,12
Szaúd-Arábia	SAU	2, 15	0,04	0,72
Dél-Korea	KOR	NA	-	-
Oroszország	RUS	NA	-	-

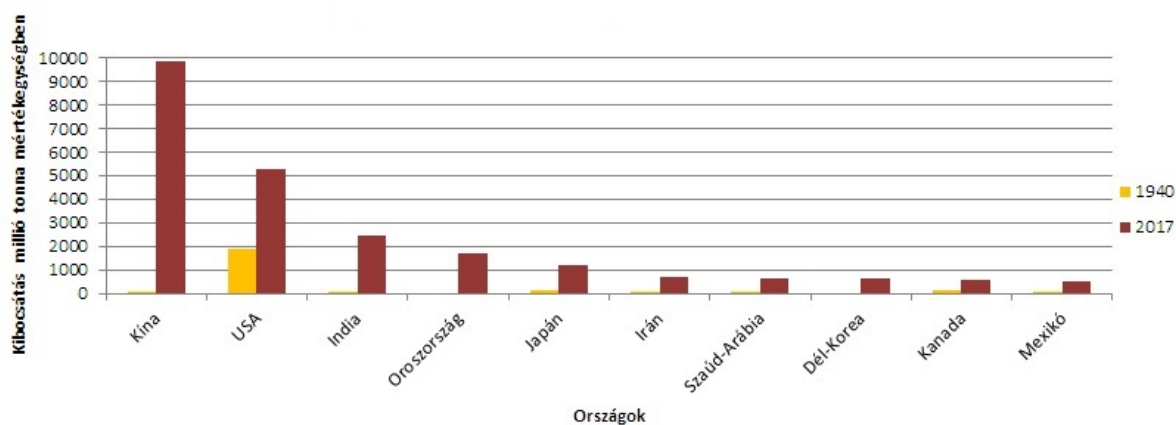
2.3. táblázat. 1940-es kibocsátás pár adata

Dél-Korea és Oroszország adatai nem ismertek erről az évről, ám manapság már nagy befolyással vannak az összes kibocsátásra.

A következő táblázat ugyanezen országok 2017-es kibocsátását tartalmazza.

Országok	Hivatalos rövidítés	2017-es kibocsátás (millió tonna)	Részarány (%)	Egy főre jutó kibocsátás (tonna/fő)
Világ összesen	-	36 153,26	-	4,79
Kína	CHN	9 838,75	27,21	6,98
Egyesült Államok	USA	5 269,53	14,58	16,24
India	IND	2 466,77	6,82	1,84
Oroszország	RUS	1 692,79	4,68	11,76
Japán	JPN	1 205,06	3,33	9,45
Irán	IRN	672,31	1,86	8,28
Szaúd-Arábia	SAU	635,01	1,76	19,28
Dél-Korea	KOR	616,09	1,70	12,08
Kanada	CAN	572,78	1,58	15,64
Mexikó	MEX	490,29	1,36	3,80

2.4. táblázat. 2017-es kibocsátás legfőbb résztvevői



2.1. ábra. A kiválasztott országok 1940-es és 2017-es kibocsátásának összehasonlítása

Az első észrevétel, hogy ezek az országok az összes CO_2 emisszió 64,89 %-át adják. Ezenkívül Kína kibocsátása körülbelül 9752 millió tonnával, az USA kibocsátása 3398 millió tonnával, a világszintű emisszió pedig közel hatszorosára növekedett. Értelemszerűen az egy főre jutó kibocsátások is megnőttek: világszinten a duplájára; a legnagyobb növekedés a vizsgált államok közül Szaúd-Arábiában, Dél-Koreában és Oroszországban történt.

Az előzőek következménye, hogy például Kína részaránya nagyjából 25 %-kal nőtt, de az USA-é 23%-kal csökkent.

2.2.2. Az Európai Unió emissziós értékei és országonkénti eloszlása

Most vizsgáljuk meg az Európai Unió emissziós értékeit. Az alábbi táblázat a tagországok 1940-es és 2017-es évi összes kibocsátását, azok arányát az EU összes kibocsátásához képest, illetve az egy főre levetített mennyiséget tartalmazza. Mivel az Egyesült Királyság 2020. január 31.-én lépett ki, ezért őket is "tagállamként" kezeltem, de dőlt betűvel jelzem, hogy már nem tagok. Vannak országok, ahol az 1940.-beli kibocsátás nem ismert, ott NA (Not Available) szerepel.

Országok	Hivatalos rövidítés	1940-es kibocsátás (millió tonna)	Részarány (%)	Egy főre jutó kibocsátás (tonna/fő)
EU 28	-	1 703, 21	96,39	-
Németország	DEU	654,37	38,42	9,16
<i>Egyesült Királyság</i>	<i>GBR</i>	<i>475,14</i>	<i>27,90</i>	<i>9,85</i>
Lengyelország	POL	167,25	9,82	5,57
Csehszlovákia	CZE	76,24	4,48	-
Belgium	BEL	62,10	3,65	7,45
Olaszország	ITA	49,05	2,88	1,11
Hollandia	NLD	33,44	1,96	3,77
Románia	ROU	23,95	1,41	1,52
Spanyolország	ESP	21,62	1,27	0,84
Svédország	SWE	18,48	1,08	2,91
Magyarország	HUN	15,82	0,93	1,71
Dánia	DNK	14,82	0,87	3,87
Írország	IRL	11,52	0,68	3,92
Ausztria	AUT	7,35	0,43	1,10
Bulgária	BGR	4,10	0,24	0,61
Portugália	PRT	3,90	0,23	0,51
Finnország	FIN	2,32	0,14	0,63
Görögország	GRC	0,24	0,01	0,03

2.5. táblázat. Az Európai Unió tagországok ismert kibocsátásai (1940)

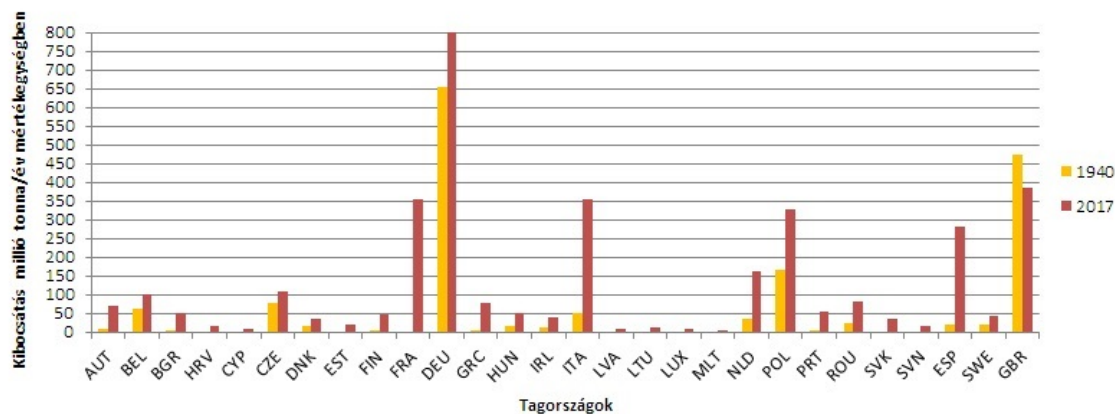
A százalékos összesítés csak 96,39 %-ra jött ki, így a nem ismert kibocsátású országok még 3 % körüli értéket produkáltak. Jól elkülönül Németország és az Egyesült Királyság, ketten együtt több mint, 65 %-ban részesedtek a jövőbeli EU

28-ak összes kibocsátásából. Az egy főre jutó értékek is náluk a legnagyobbak, hiába magas a népességük, sajnos a kibocsátásuk is. Ezzel szemben ott van Olaszország és Magyarország példája: habár Olaszország emissziója magasabb, a népessége is nagyobb, mint hazánké, így az egy főre jutó kibocsátási érték alacsonyabb, mint a magyaroké.

Most nézzük a 2017-es adatokat.

Országok	Hivatalos rövidítés	2017-es kibocsátás (millió tonna)	Részarány (%)	Egy főre jutó kibocsátás (tonna/fő)
EU 28	-	3 543,68	100	-
Németország	DEU	799,37	22,56	9,73
<i>Egyesült Királyság</i>	<i>GBR</i>	<i>384,71</i>	<i>10,86</i>	<i>5,81</i>
Franciaország	FRA	356,30	10,05	5,48
Olaszország	ITA	355,45	10,03	5,99
Lengyelország	POL	326,60	9,22	9,63
Spanyolország	ESP	281,42	7,94	6,07
Hollandia	NLD	164,05	4,63	3,35
Csehország	CZE	107,90	3,04	10,16
Belgium	BEL	100,12	2,83	8,76
Románia	ROU	79,99	2,26	5,31
Görögország	GRC	76	2,14	6,81
Ausztria	AUT	69,94	1,97	8,01
Portugália	PRT	54,86	1,55	8,56
Magyarország	HUN	50,34	1,42	5,18
Bulgária	BGR	49,07	1,38	6,93
Finnország	FIN	45,96	1,30	8,32
Svédország	SWE	41,50	1,17	4,19
Írország	IRL	39,74	1,12	8,35
Szlovákia	SVK	35,39	1,00	6,50
Dánia	DNK	34,55	0,97	6,03
Észtország	EST	19,81	0,56	15,13
Horvátország	HRV	17,18	0,48	4,10
Szlovénia	SVN	14,61	0,41	7,02
Litvánia	LTU	13,39	0,38	4,63
Luxemburg	LUX	9,29	0,26	15,93
Ciprus	CYP	7,51	0,21	6,37
Lettország	LVA	7,17	0,20	3,68
Málta	MLT	1,44	0,04	3,35

2.6. táblázat. Az "Európai Unió" 2017-es kibocsátása



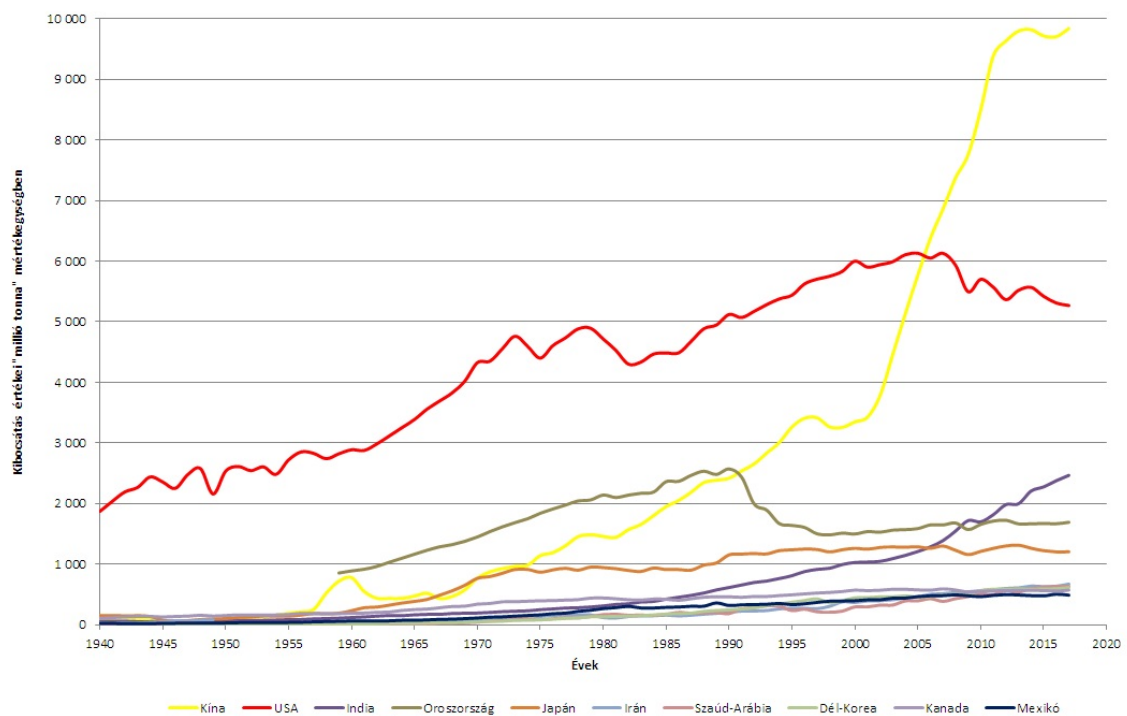
2.2. ábra. Az EU tagországok 1940-es és 2017-es kibocsátásainak összehasonlítása

Az első, ami feltűnik, hogy az EU 28-ak összes kibocsátás a majdnem háromszorosára nőtt a két időpont között. A 2.5-ös és 2.6-os táblázat hasonlósága, hogy mindkét évben Németország és az Egyesült Királyság foglalta el az első két helyet, illetve a tagországok sorrendje sem változott sokat, ez a százalékos eloszlásokból is megfigyelhető. Mindenhol kisebb-nagyobb mértékben megnövekedett a kibocsátás, de kicsit érdekesebb az egy főre jutó kibocsátás alakulása. Az első helyüket tartó németeknél alig növekedett az érték (9,16-ról 9,73-ra), addig például hazánkban ez eléggé megugrott (1,71-ről 5,18-ra). Ez a növekedés a legtöbb országra sajnos elmondható, bár az Egyesült Királyságban szinte drasztikusan csökkent a mennyiség (9,85-ről 5,81-re). Ugyanakkor figyelembe kell venni, hogy ez az érték az adott ország összes kibocsátásának és népességének hányadosa, így precízebb elemzés esetén meg kell vizsgálni a népesség számbeli változását, legyen az csökkenés, vagy növekedés (például munkahely miatti be- és kivándorlás alakulása).

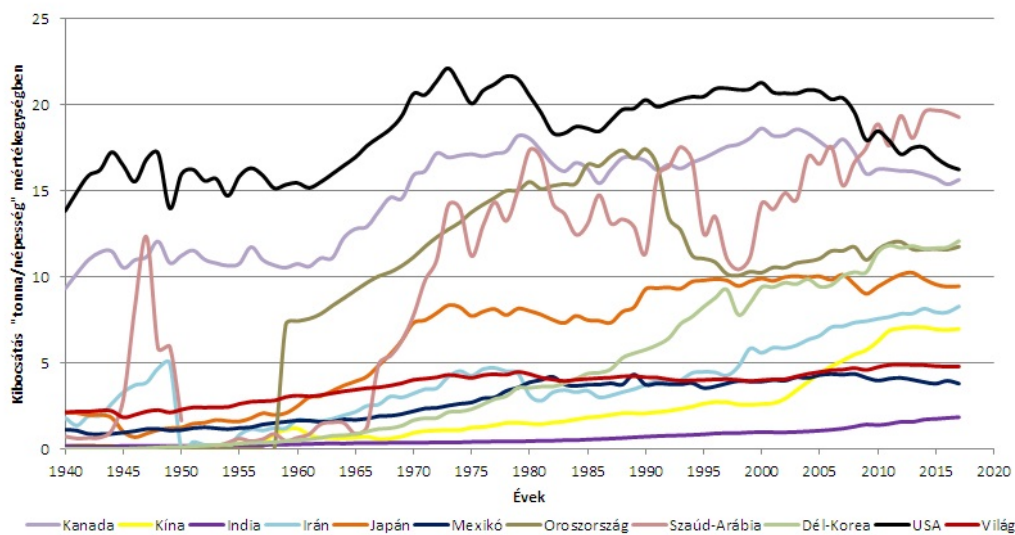
2.2.3. A kibocsátási értékek időbeli alakulása

Az eddigi táblázatok és ábrák átfogó képet mutattak a CO_2 emisszió alakulásáról, várható volt, hogy az évek alatt növekedett a kibocsátás. Ennek ellenére a konkrét időbeli alakulás nem látszik, így ebben a részben pontosabb képet szeretnék rajzolni erről.

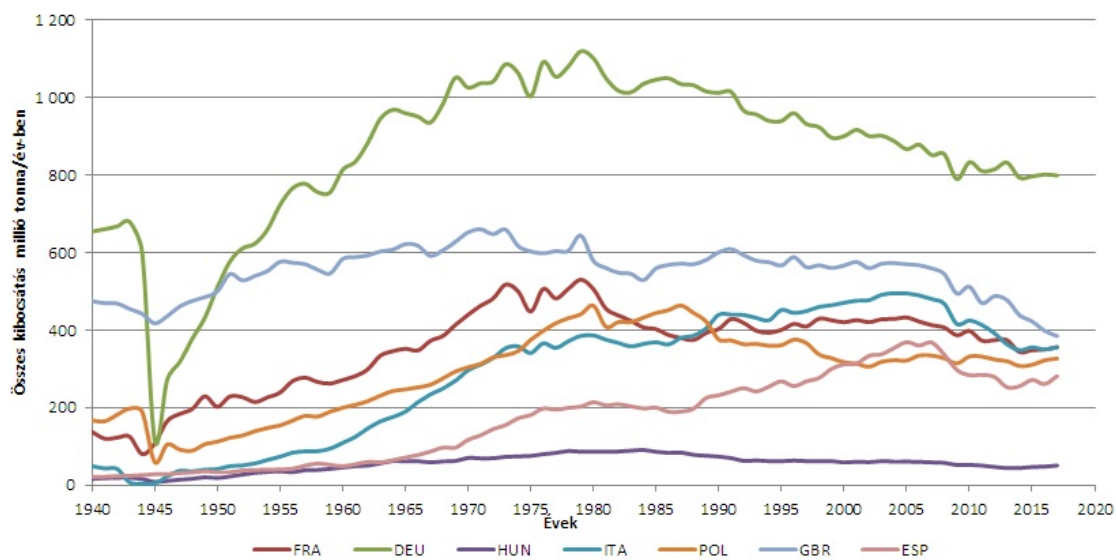
Az alábbi vonaldiagramok szemléltetik a kibocsátások időbeli alakulásait 1940. és 2017. között: a kiválogatott 10 ország összes és egy főre jutó kibocsátásának változása a "világ" emissziójához képest; az Európai Unió tagországainak összes és egy főre jutó kibocsátása; és Magyarország kibocsátásainak alakulása.



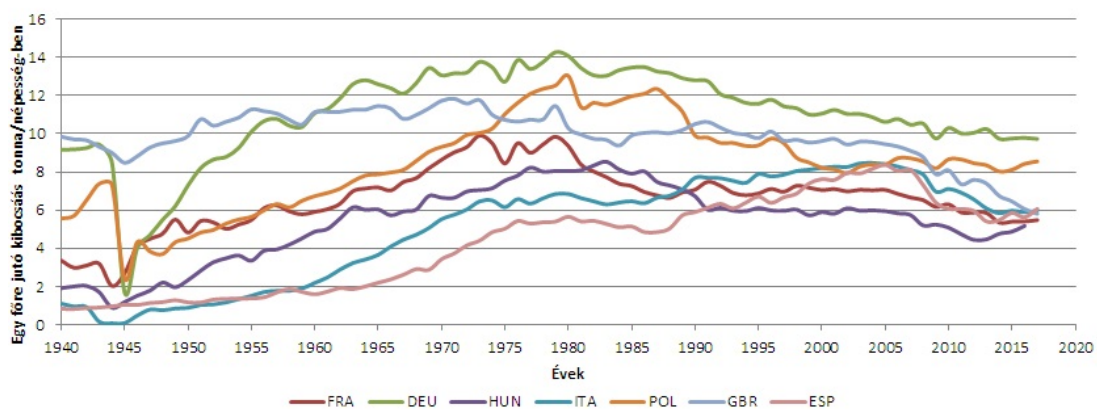
2.3. ábra. Az országok összes kibocsátásának változása



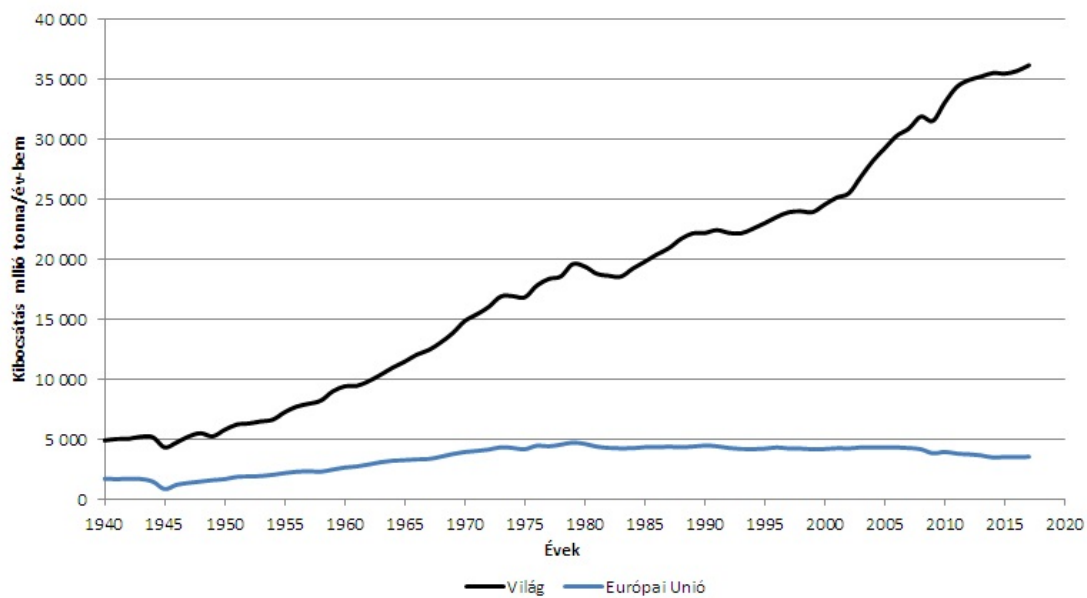
2.4. ábra. Az országok egy főre jutó kibocsátásának változása



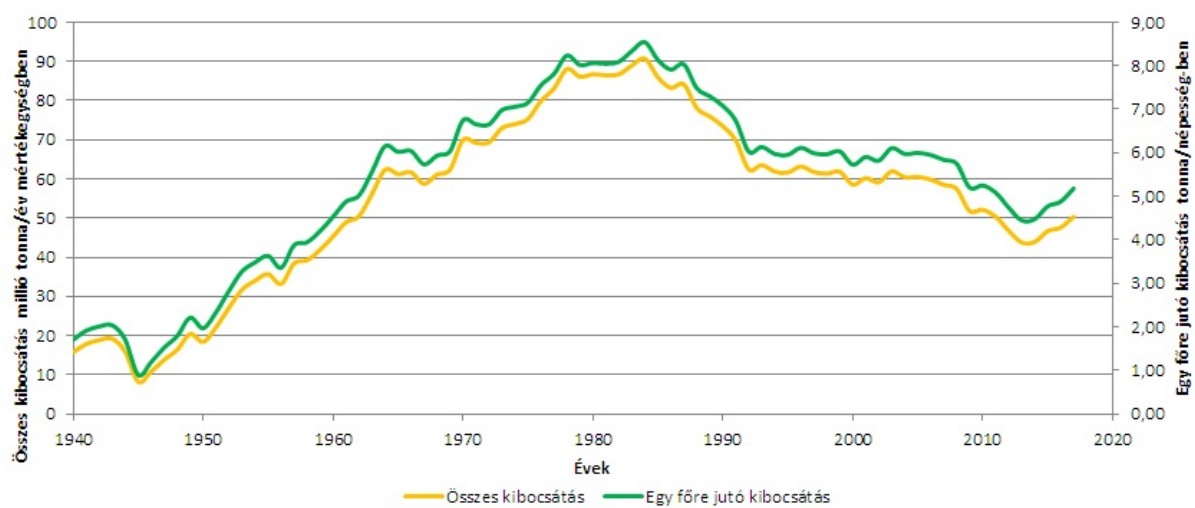
2.5. ábra. Az EU néhány tagországának éves CO_2 kibocsátásának változása



2.6. ábra. Néhány tagország egy főre jutó kibocsátásának változása



2.7. ábra. A világ és az Eu kibocsátásának változása



2.8. ábra. Magyarország kibocsátásainak alakulása

A 2.8-as ábrán jól látszik, hogy az 1940-es évek közepétől elkezdett nagyobb ütemben nőni a kibocsátás, egészen az 1980-as évekig. A csökkenés oka az 1979. november 13-án aláírt *Genfi Egyezmény* [8], mely a nagy távolságra jutó, országhatáron áterjedő légszennyezés csökkentéséről szól. Az Egyezményt hazánk 1980. szeptember 22-én ratifikálta. A következő ilyen megállapodást 1992.-ben írta alá Magyarország. Az *Egyesült Nemzetek Éghajlatváltozásról Szóló Keret-egyezménye* [21], melyet aláírva a résztvevő országok elismerték, hogy az üvegházhatású gázok felborítják az éghajlati rendszert és hogy a klímaváltozás valós veszély. Célkitűzésük az ÜHG-k kibocsátásának olyan szinten tartása, melyen megelőzhető az emberiség általi, visszafordíthatatlan "éghajlati beavatkozás". Ezt 1997.-ben egészítették ki, az ún. *Kiotói jegyzőkönyvvel* [7, 21]. A kiegészítés kötelezően teljesítendő, üvegházhatású gázok kibocsátásának csökkentését írja elő, melyek arányát az országok maguknak írhatják elő. Hazánk 6%-os csökkentést határozott el az 1985-1987-es évek átlagos kibocsátásához képest.

3. fejezet

Lehetséges elemzési módszerek

Már régóta foglalkoztatja a tudósokat, hogy milyen modellel lehetne előrejelezni a kibocsátás mennyiségét, értelemszerűen a legalacsonyabb hibarátával. Ebben a fejezetben azokat a leggyakoribb módszereket járom körül, melyeket kutatók és elemzők használtak az elmúlt években.[12]

3.1. Artificial Neural Network (Mesterséges neurális háló)

A modellek egy része a mesterséges intelligencián (AI) alapul, míg a többi matematikai modellként értelmezhető inkább. Ilyen sikeres AI modell az "ANN", melynek létrehozását az emberi agy és az idegrendszer működése inspirálta. Főként a gépi tanulás területein alkalmazzák, ezen belül is gyakoriak az osztályozási problémákra épített modellek. Felépítése az agyhoz hasonló, "csomópontokból" és a köztük lévő irányított kapcsolatokból áll. Általában több rétegűek, van bemeneti, rejtett és kimeneti rétegük, ezek a rétegek tartalmazzák a csomópontokat. A bemeneti réteg "fogadja" az információt és továbbítja a rejtett rétegnek vagy rétegeknek, melyek átalakíthatják, kódolhatják az adatokat. A kimeneti réteg mutatja a modell úgymond "megoldását", például osztályozási feladatoknál a megfelelő csoportokba sorolja az adatokat.

Ilyen modellt használtak például *Radojevic* és munkatársai Szerbiában. Vizsgálatuk során 1999. és 2007. közti adatokat, a megújuló energiaforrások értékeit, a GDP-t, a bruttó energiafogyasztást és energiaszéntintést használták, az Európai Unió országaira, azon belül pedig Bulgáriára vetítve. (Bulgária nagyban hasonlít Szerbiára, gazdasági fejlettségükben, ipari szerkezetükben, éghajlatukban és ener-

giaintenzitásukban.) Eredményeik szerint az ANN elfogadható modellezést adhat a kibocsátások vizsgálatára.

Hasonlóan alkalmazták az ANN-t más kutatók is, például *Liu* Kínában 1990-2010 közti adatok alapján. A neurális hálózatot a GDP, az exportok, a fogyasztó árindex, az állóeszközökbe való befektetések és a népesség alapján építették fel, a CO_2 kibocsátást pedig a fosszilis üzemanyagok égetéséből és a cementgyártásból vizsgálták. A következő 10 évre készítettek el az előrejelzést, melynél a relatív hiba (a valós és a becsült kibocsátások közti százalékos eltérés) +1% és -1% között voltak.

3.2. Lineáris regresszió és többszörös lineáris regresszió

A lineáris regresszió lényege, hogy ha van két folytonos változó, melyek lineáris kapcsolatban állnak egymással, akkor az egyik változóval előre jelezhetjük a másikat és minél erősebb a két változó közti összefüggés, annál kisebb a hiba.

Murad és társai ezt a tulajdonságot használták ki, mikor a malajziai adatokra a legkisebb négyzetek módszerét alkalmazták. Céljuk az éghajlatváltozás és a mezőgazdasági növekedés közti kapcsolat azonosítása és elemzése, azon belül az egy főre eső CO_2 kibocsátás és a mezőgazdasági termelési index közti kapcsolat vizsgálata volt. Elemzésükben észrevették, hogy a kapcsolat a hőmérséklet és a termelési index között fordított, valamint az egy főre jutó CO_2 kibocsátás és a termelési index között jelentős és pozitív irányú.

Ezzel szemben *Hajji* és *Lewis* többszörös lineáris regressziót alkalmaztak, a becsléshez a termelési mennyiséget és időtartamot, a teljes üzemanyag felhasználást és a földmunkák során keletkező összes szennyezőanyag kibocsátást használták. Modelljük pontosnak bizonyult, az üzemanyag felhasználása és kibocsátásának becslésére (földmunkát végző tehergépjárművek esetében) remekül működött.

Ezekén felül persze sok más módszer is rendelkezésre már kibocsátási modellezésre és előrejelzésre, hasonlóan jó eredményekkel az 1990-es és 2000-es évek közti adatok alapján.

4. fejezet

Idősorok

Az előzőekben bemutatott AI-n alapuló módszerek is tudnak előrejelzést adni, azonban más módszerekkel is képesek vagyunk modellezni. Szeretnék statisztikai módszerek segítségével egy megbízható modellt építeni, ezért ebben a fejezetben ebből az irányból közelítem meg a feladatot.

Az egyik legjobban használható, általános idősoros modell az ARIMA(p,d,q). Ezek nagyban hasonlítanak a regressziós modellezéshez, ám itt már nemlineáris összefüggéseket is figyelembe lehet venni.

Az alábbi fejezet megírásában *Hulyák Katalin: Idősorok rövidtávú előrejelzése ARIMA modellekkel* [10] című könyve, *Márkus LászlóI "Idősorok" jegyzete*[13], *Prof. Dr. Závoti József:Matematikai statisztikai elemzések 3. könyve*[16] és *Zempléni András jegyzetei*[23] [24] segítettek.

4.1. Elméleti áttekintés és alapvető fogalmak

Az egyszerű statisztikai elemzéseknél adataink függetlenségét követeljük meg, addig a valós élethelyzetek, mint például a gazdasági, társadalmi vagy akár környezeti jelenségek korántsem függetlenek, sőt, éppen az egymás utáni megfigyelések összefüggése ad támpontot az idősor elemzéséhez.

Az elemzéseknél nagy segítség, hogy a feladatot két részre oszthatjuk:

- először leválasztjuk az időbeni szisztematikus változásokat, mint például a trend vagy a periódus
- majd az így megmaradt -remélhetőleg már- stacionárius folyamatok összefüggőségi struktúráit vizsgálhatjuk.

Definíció 4.1.1 Az $X_1, X_2, \dots, X_t, \dots$ valószínűségi változók sorozatát idősornak nevezzük, ha az indexparaméter időként is értelmezhető.

Definíció 4.1.2 Az $X(t)$ idősor gyengén stacionárius, ha $E(X_t) = \text{konstans}$ és $\text{cov}(X_t, X_s)$ csak a $t - s$ időkülönbség függvénye, azaz eltolásinvariáns.

Megjegyzés 4.1.3 Az eltolásinvarianciából következik, hogy a gyengén stacionárius idősor változóinak szórása azonosak és végesek.

Definíció 4.1.4 Az $X(t)$ idősor erősen stacionárius, ha a véges dimenziós eloszlásai eltolásinvariánsak, azaz minden n és t_1, \dots, t_n esetén $(X_{t_1}, \dots, X_{t_n}) \simeq (X_{t_1+h}, \dots, X_{t_n+h})$.

Mivel egy $X(t)$ idősor időbeli fejlődést ír le, ezért az egymás utáni állapotai összefüggnek. Ezt az összefüggést az autokorreláció függvényvel jellemezhetjük.

Definíció 4.1.5 A gyengén stacionárius idősor $\text{corr}(X(t), X(s)) = r(t - s)$ korrelációt autokorreláció függvénynek nevezzük.

Definíció 4.1.6 Az $\epsilon(t)$ folyamat fehér zaj, ha ezek az $\epsilon(t)$ -k azonos eloszlásúak minden t -re és korrelálatlanok (de nem feltétlenül), valamint a várható értékük 0.

Definíció 4.1.7 Az $\epsilon(t)$ folyamat független értékű zaj, ha $\epsilon(t)$ -k független, azonos eloszlású valószínűségi változók, 0 várható értékkel.

Definíció 4.1.8 Legyen B az eltolás vagy visszaléptetés operátora, azaz:

$$BX(t) = X(t - 1), B^2X(t) = X(t - 2), \dots \quad (4.1)$$

4.2. Az ARMA modell felépítése

Az ARMA(p,q) modell az úgynevezett *p-ed rendű autoregressziós folyamatból* - későbbiekben AR(p)- és *q-ad rendű mozgóátlag folyamatból* - MA(q)- tevődik össze, amit stacionárius idősorok modellezésére alkalmaznak.

Az AR(p) modellnél az adott idősor egy t -pontbeli értéket állítjuk elő az előző, $t-1, t-2, \dots$ pontok értékeiből:

$$X(t) = \alpha_1 X(t-1) + \alpha_2 X(t-2) + \dots + \alpha_p X(t-p) + \sigma_\epsilon \epsilon(t) \quad (4.2)$$

ahol $\epsilon(t)$ független értékű zaj, 1 szórással és $\alpha_1 \dots \alpha_p \in \mathbb{R}$. Az egyenletet átrendezhetjük:

$$X(t) - \sum_{i=1}^p \alpha_i X(t-i) = \sum_{i=0}^p \tilde{\alpha}_i X(t-i) = \sigma_\epsilon \epsilon(t) \quad (4.3)$$

Ennek karakterisztikus polinomja

$$P(x) = \sum_{i=0}^p \tilde{\alpha}_i x^i(t-i) \quad (4.4)$$

ahol $\tilde{\alpha}_0 = 1$ és $\tilde{\alpha}_i = -\alpha_i$.

Tétel 4.2.1 Az AR(p) egyenletnek pontosan akkor létezik egyértelmű, gyengén stacionárius megoldása, ha a $P(x)$ karakterisztikus polinom gyökei az egységkörön belül vannak.

Az MA(q) modell esetében az $\epsilon(t)$ független értékű zaj, $\beta_i \in \mathbb{R}$ konstansok. Ekkor:

$$X(t) = \beta_0 \epsilon(t) + \beta_1 \epsilon(t-1) + \dots + \beta_q \epsilon(t-q). \quad (4.5)$$

Állítás 4.2.2 Az MA(q) folyamatok mindig erősen stacionáriusak.

Az idősorok valószerűbb közelítése érdekében kifejlesztették a vegyes ARMA(p,q) modelleket, melyek egyaránt tartalmazzák az autoregresszív és a mozgóátlagolású tagokat, így már kétféle dimenzióval, p -vel és q -val jellemezhetőek:

$$\alpha_1 X(t-1) + \dots + \alpha_p X(t-p) + \sigma_\epsilon \epsilon(t) = \beta_0 \epsilon(t) + \beta_1 \epsilon(t-1) + \dots + \beta_q \epsilon(t-q) \quad (4.6)$$

Az ARMA(p,q) folyamat stacionárius, ha az AR(p) része stacionárius.

4.3. Az ARIMA modell felépítése

Az ARIMA modellek nagy előnye, hogy ezek már képesek trendet is modellezni, nem kell különválasztani a feladatokat.

Az ARIMA modell háromféle dimenzióval jellemezhető:

- p az autoregresszió rendje
- d a differenciaképzés foka
- q a mozgóátlagolás rendje

Definíció 4.3.1 Az $X(t)$ folyamatot $ARIMA(p,1,q)$ folyamatnak nevezzük, ha az egyszeres differenciáltja

$$Y(t) = X(t) - X(t-1) = (1-B)X(t) \quad (4.7)$$

ARMA folyamat.

Definíció 4.3.2 Az $X(t)$ folyamat $ARIMA(p,d,q)$, ha a d -szeres differenciáltja

$$(1-B)^d X(t) \quad (4.8)$$

ARMA folyamat.

Ugyan az $AR(p)$, $MA(q)$ és $ARMA(p,q)$ modellek stacionárius esetben használhatóak, a valós jelenségeket (például gazdasági fejlődés) leíró idősorok általában nem stacionáriusak. Ezek a sorok általában d -ed rendű differenciálással már stacionáriusak lesznek. Ha ezekre a differenciákra írunk fel $ARMA(p,q)$ modellt, akkor abból $ARIMA(p,d,q)$ modellt kapunk. Az $ARIMA(p,d,q)$ modellek speciális eseteinek tekinthetők az $AR(p,0,0)$, a $MA(0,0,q)$ és az $ARMA(p,0,q)$ modellek.

4.4. Becslési módszerek

Statisztikai modellezések során gyakran használunk becslési eljárásokat. Az adott modellünk szabad paramétereit szeretnénk meghatározni, ám a konkrét értékét nem, csak egy közelítést tudunk adni. Az egyik legáltalánosabb módszer a *maximum-likelihood (ML)*, mellyel a keresett paraméter olyan becslését adja meg, ami a legnagyobb valószínűséggel fog bekövetkezni.

Definíció 4.4.1 Legyenek X_1, X_2, \dots, X_n független, azonos eloszlású valószínűségi változók és legyen v az ismeretlen paraméter. Ekkor a likelihood függvény a következő:

$$L(v, x) = f_v(x) = \prod_{t=1}^n f_v(x_t), \quad (4.9)$$

ha az eloszlás folytonos, és

$$L(v, x) = P_v(X = x) = \prod_{t=1}^n P_v(X_t = x_t), \quad (4.10)$$

ha az eloszlás diszkrét.

Definíció 4.4.2 Maximum-likelihood becslés esetében a $\max_v L(v, x)$ -et keressük.

Megjegyzés 4.4.3 Az ML becslés alkalmazható nem független esetekben is, ám akkor már bonyolultabbak a formulák.

A fent bevezetett pontbecslésnél célszerűbb az intervallumbecslés, melynek leggyakoribb megvalósítási módja a konfidenciaintervallum becslése. A konfidenciaintervallum megadja a becsült változó alsó és felső korlátját egy adott szignifikanciaszinten. A becsült érték várhatóan az esetek $100 \cdot (1 - \alpha)$ százalékban fog az intervallumba esni. Az elvárt megbízhatóság függvényében mi szabhatjuk meg, hogy mennyivel legyen egyenlő a α . (A szignifikanciaszintet így %-ban szoktuk általában megadni.) Ez sokkal több információt ad, mintha csak egy konkrét értéket adna vissza a modell. Több intervallumot is adhatunk, ha változtatjuk a szignifikanciaszintet.

4.5. Illeszkedésvizsgálatok

Egy idősor modellezési feladatánál fontos megvizsgálni a modell illeszkedését. Erre adnak módszert az ún. Portmanteau próbák. Ezek olyan statisztikai hipotézisvizsgálati módszerek, melyeknél a nullhipotézist pontosan megadják, ám az alternatív hipotézis viszont általános. Az idősorok elemzésénél két ismert verziót használunk, amelyeket a modellünk reziduálisaira fogunk alkalmazni.

4.5.1. Box-Pierce teszt

A Box-Pierce teszt, vagy más néven a klasszikus Portmanteau próba esetében egy T elemű idősor megfigyelése mellett a nullhipotézis az, hogy a folyamat fehér zaj. Ekkor, ha az első h "lag" autokorrelációjának négyzetösszegét vesszük és beszorozzuk a T elemszámmal:

$$Q = T \cdot \sum_{\tau=1}^h r^2(\tau) \quad (4.11)$$

akkor ennek úgy kell viselkednie, mint h db független, standard normális eloszlás négyzetösszege, és így a statisztika χ_h^2 eloszlású. Ha a $Q > \chi_{h,q}^2$ a χ_h^2 eloszlás q kvantilisével, akkor $1-q$ szignifikancia szinten elutasítjuk a nullhipotézist, azaz a folyamat nem fehér zaj. Mivel ez egy elég gyenge próba, finomítottak rajta.

4.5.2. Ljung-Box teszt

A Ljung-Box teszt egy módosított Box-Pierce teszt. A különbség az előző (4.9) képlethez képest az, hogy a nagyobb "lag"-ekhez tartozó autokorrelációk súlyát megnöveljük.

$$Q_{LB} = T \cdot (T+2) \sum_{\tau=1}^h \frac{r^2(\tau)}{T-\tau} \quad (4.12)$$

Hasonlóan a Box-Pierce teszthez, ez is χ_h^2 eloszlású. Ha a $Q > \chi_{1-q,h}^2$, akkor $1-q$ szignifikancia szinten elutasítjuk a nullhipotézist.

5. fejezet

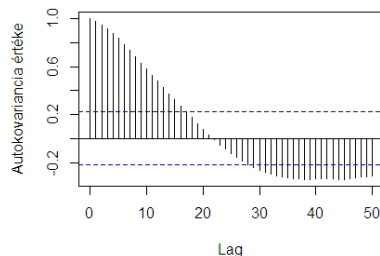
Modellezés és előrejelzés

Elemzéseim során elsősorban az Európai Unió, azon belül Magyarország emisszióját helyeztem előtérbe. Adataimat a már említett "ourworldindata.org" [9] adatbázisaiból vettem. Az adatok vizsgálatához az RStudio [18] statisztikai programot használtam. Az Európai Unió kibocsátásának elemzéséhez és előrejelzéséhez használt kódot a Függelékben mellékelem.

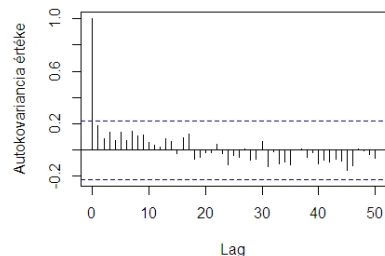
5.1. Az összes kibocsátás elemzése és előrejelzése

5.1.1. Az Európai Unió

Első lépésként a megfelelő adatbázisból kiszűrtem az 1940. és 2017. közti, az Európai Unió tagországaira vonatkozó összesített adatokat és idősorra alakítottam őket, illetve korrigáltam, hogy az elemzés során is millió tonna mértékegységben legyenek. Ezután megvizsgáltam az autokovarianciát az eredeti idősorra és a stacionáriusnak gondolt változatra.(51.a) és b) ábra)



(a) Eredeti idősor



(b) Stacionárius idősor

5.1. ábra. Az EU adatainak autokovariancia vizsgálata

Az autokorrelációk mintázata alapján (nagyon lassú, nem exponenciális lecsengés, nem reális negatív korrelációk a nagy lagekre) az eredeti idősor esetében elvethető a stacionaritás.

Az idősorra kvadratikus regressziót is illesztettem, melynél az idő és annak négyzete volt a magyarázó változó. Ezzel azt vizsgáljuk, hogy a trend, ami láthatóan nem lineáris (2.7 ábra), vajon mennyire jól közelíthető az idő legfeljebb másodfokú polinomjával.

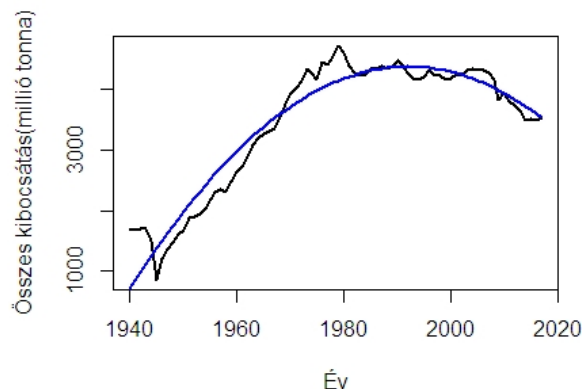
Jelen esetben a t és a t^2 is szignifikánsak voltak, a p-érték nagyon kicsi (ez leolvasható a függelékben látható outputból, de abból is látszik, hogy a standard hiba kicsi a becsült paraméter-értékhez képest), ezért valószínűtlen, hogy bármelyik együttható 0 lenne.

	Együttható	Standard hiba
Intercept	559,88	106,36
t	144,07	6,21
t ²	-1,36	0,08

5.1. táblázat. Az EU kvadratikus regressziójának együtthatói

Ezután ARIMA modellt illesztettem a regresszió után fennmaradó reziduálisokra (EU_ARIMAmodell) és az eredeti idősorra (EU_ARIMAmodell2) is.

	ARIMA (2,0,1)	Standard error
AR1	1,90	0,04
AR2	-0,94	0,04
MA1	-0,97	0,05



5.2. ábra. Az EU kibocsátására illesztett kvadratikus regresszió

5.2. táblázat. Az EU_ARIMAmodell modell együtthatói és standard hibáik

	ARIMA (0,2,1)	Standard error
MA1	-0,87	0,07

5.3. táblázat. Az EU_ARIMAmodell2 modell együtthatói és standard hibáik

	ARIMA (2,0,1)	ARIMA (0,2,1)
AIC	988,51	970,65
BIC	997,94	975,31
sum(res^2)	1 252 166	1 458 488

5.4. táblázat. ARIMA (2,0,1) és ARIMA (0,2,1) összehasonlítása

A regresszió reziduálisaira illesztett ARIMA modell paraméterei a (2,0,1), azaz 2 rendű AR, 1 rendű MA modellt illesztett a program és nem kellett differenciálni, azaz stacionáriusnak volt tekinthető a reziduálisok adatsora.

Ezzel szemben az idősorra illesztett modell egy ARIMA(0,2,1) lesz, azaz egy 1 rendű MA modell kétszeres differenciálással, ami éppen a kvadratikus hatások kiszűrésének felel meg.

A modelleket általában az AIC, BIC és $\text{sum}(res^2)$ szerint szoktuk összehasonlítani (minden mérőszámra igaz, hogy "minél kisebb, annál jobb"), jelen esetben AIC és BIC szerint az idősorra illesztett ARIMA, $\text{sum}(res^2)$ szerint pedig a regresszió reziduálisaira illesztett ARIMA modell a jobb.

Meg kell nézni, hogy a modellek illesztése helyes-e, erre a Ljung-Box tesztet használhatjuk. Az első ARIMA modellre a p-érték 0,6535; a másodikra 0,6799.

Mindkét ARIMA modellel prediktálva a következő előrejelzéseket kapjuk a következő 3 évre. A következő táblázat a 95%-os és 99%-os konfidenciaintervallumok alsó és felső korlátját is megadja. Ennél hosszabb időtávra a sok váratlan, a matematikai modellek által nem megfogható esemény (járványok, bozóttüzek stb) miatt nem lenne értelme becslést adni.

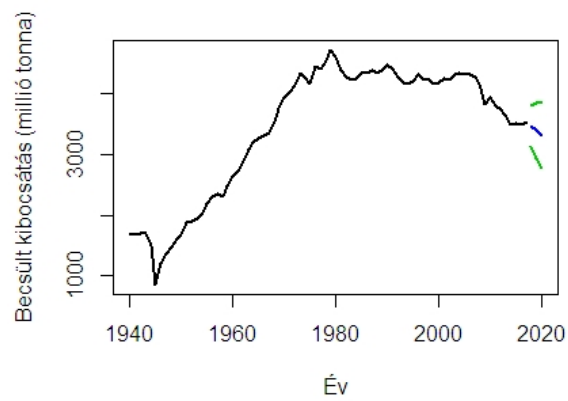
	ARIMA(2,0,1)	95 %	99 %
2018	3472	[3219; 3725]	[3139; 3805]
2019	3398	[3051; 3745]	[2942; 3854]
2020	3320	[2913; 3729]	[2784; 3857]

5.5. táblázat. Az EU kibocsátásának regresszió alapuló becslése és korlátai

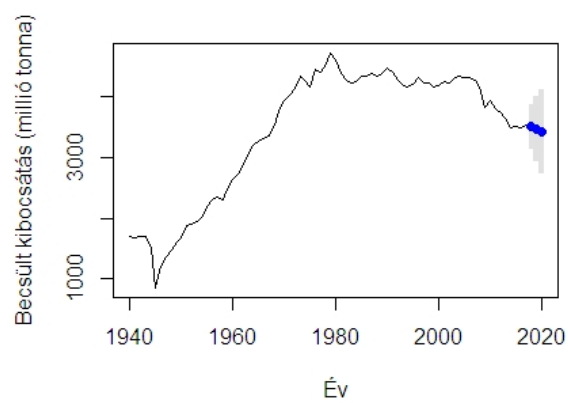
	ARIMA(0,2,1)	95 %	99 %
2018	3504	[3231; 3777]	[3145; 3863]
2019	3465	[3054; 3876]	[2924; 4006]
2020	3426	[2891; 3961]	[2722; 4129]

5.6. táblázat. Az EU kibocsátásának idősoron alapuló becslése és korlátai

Az EU_ARIMAmoell2, azaz a közvetlenül az idősorra illesztett modell magasabb értékű előrejelzést adott, mint a regresszió reziduálisaira illesztett, hiszen ebben végül a becsült regresszió is megjelenik, ami csökkenő trendet prognosztizál. A konfidenciaintervallumok mindkét esetben a 2018-as előrejelzésnél szűkebbek; míg a 2020-as már nagyobb intervallum. Ez teljesen érhető, hiszen 2019. és 2020. előrejelzéséhez a modelleknek már becsült értéket (értékeket) is kellett használniuk, emiatt nőtt a hibaráta is.



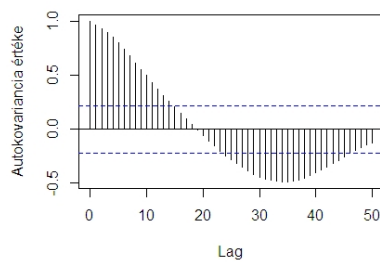
5.3. ábra. Az EU kibocsátásának előrejelzése ARIMA(2,0,1) modellel 99%-os konfidenciaszinten



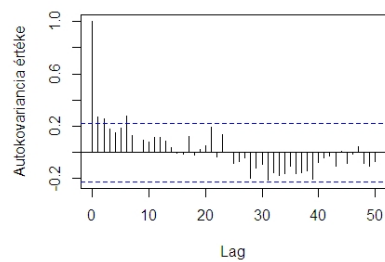
5.4. ábra. Az EU kibocsátásának előrejelzése ARIMA(0,2,1) modellel 99%-os konfidenciaszinten

5.1.2. Magyarország

Hasonlóan az Európai Unióhoz, itt is kiszűrtem az 1940. és 2017. közti magyarországi adatokat és idősorra konvertáltam őket, illetve átváltottam az értékeket, így már millió tonna mértékegységben vannak. Az elemzést ezután ugyanúgy végeztem el, mint az előző részben.



(a) Eredeti idősor



(b) Stacionárius idősor

5.5. ábra. Magyarország adatainak autokovariancia vizsgálata

Most is igaz, hogy az autokorrelációk mintázata alapján az eredeti idősor esetében szintén elvethető a stacionaritás.

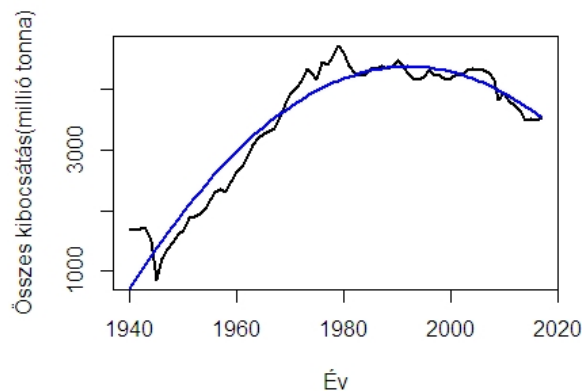
Az idősorra itt is kvadratikus regressziót illesztettem, és ugyancsak az idő és annak négyzete volt a magyarázó változó. Az a t és a t^2 is szignifikánsak voltak, a p -érték nagyon kicsi, ezért valószínűtlen, hogy van nulla együttható.

	Együttható	Standard hiba
Intercept	-4,969	2,73
t	3,481	0,159
t^2	-0,038	0,002

5.7. táblázat. Magyarország kvadratikus regresszióvizsgálatának együtthatói

Eredményei hasonlóak az EU regressziós értékeihez, a szignifikancia szintek itt is a legjobbak, a p értéke szintén nagyon kicsi, így itt valószínűtlen, hogy lenne nulla együttható.

A következő az ARIMA modellek illesztése. A regressziós modellre az ARIMAmodell, az idősorra pedig az ARIMAModell2 vonatkozik.



5.6. ábra. Magyarország kibocsátására illesztett kvadratikus regresszió

	Együttható	Standard hiba
Intercept	559,88	106,36
t	144,07	6,21
t2	-1,36	0,08

5.8. táblázat. Az EU kvadratikus regressziójának együtthatói

Ezután ARIMA modellt illesztettem a regressziós eredményekre (EU_ARIMAmodell) és az idősorra(EU_ARIMAmodell2) is.

	ARIMA (3,0,1)	Standard error
AR1	1,98	0,12
AR2	-1,07	0,24
AR3	0,06	0,12
MA1	-0,95	0,04

5.9. táblázat. Az ARIMAmodell modell együtthatói és standard hibáik

	ARIMA (0,2,1)	Standard error
MA1	-0,81	0,09

5.10. táblázat. Az ARIMAmodell2 modell együtthatói és standard hibáik

	ARIMA (3,0,1)	ARIMA (0,2,1)
AIC	388,43	384,37
BIC	400,22	389,03
sum(res^2)	556,85	654,23

5.11. táblázat. ARIMA (3,0,1) és ARIMA (0,2,1) összehasonlítása

A regresszió reziduálisaira illesztett ARIMA modell paraméterei a (3,0,1), azaz az AR rendje 2, az MA rendje 1; az idősorra illesztett ARIMA modell paraméterei (0,2,1), azaz egy 1 rendű MA-t illesztett kétszeres differenciával. A modelleket összehasonlítva, AIC és BIC szerint az idősorra illesztett, sum(res^2) szerint a regresszió reziduálisaira illesztett ARIMA modell a jobb.

Ezekre a modellekre is elvégezhetjük a Ljung-Box tesztet: az elsőre $p = 0,6736$; a másodikra $p = 0,7468$.

Az ARIMA modellekkel elkészíthetjük az előrejelzéseket, szintén 3 évre, ugyanazokkal a szignifikanciaszintekkel.

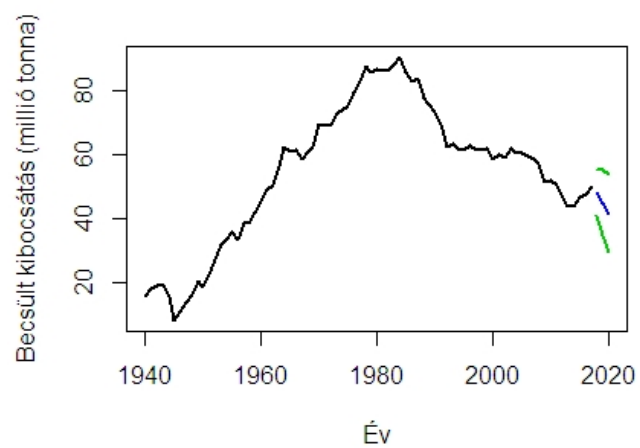
	ARIMA(3,0,1)	95 %	99 %
2018	48	[42; 53]	[41; 55]
2019	45	[37; 53]	[35; 55]
2020	42	[32; 51]	[29; 54]

5.12. táblázat. Hazánk regresszió alapuló kibocsátásának becslése és korlátai

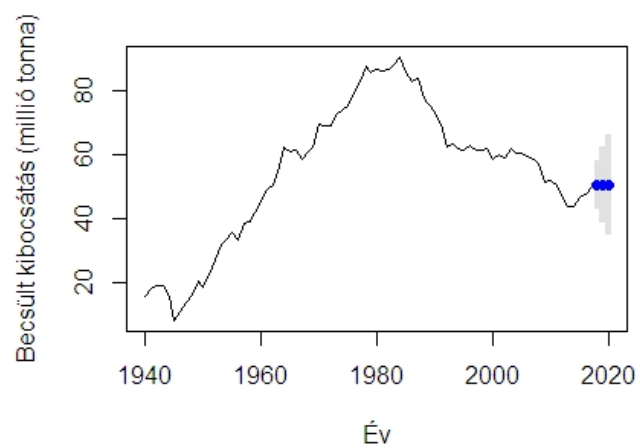
	ARIMA(0,2,1)	95 %	99 %
2018	50	[44; 56]	[43; 58]
2019	51	[42; 59]	[39; 62]
2020	51	[39; 63]	[35; 67]

5.13. táblázat. Hazánk idősoron alapuló kibocsátásának becslése és korlátai

Az ARIMAmoell2 (idősorra illesztett) előrejelzései magasabb értéket adtak vissza, mint az ARIMAmoell (regresszió reziduálisaira illesztett). Ugyanazon okok miatt, mint az EU előrejelzéseinél, a konfidenciaintervallumok itt is egyre nagyobbak az idő előrehaladtával.



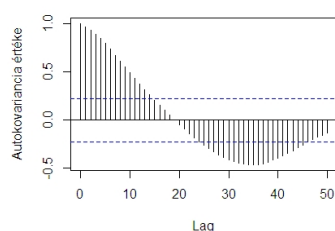
5.7. ábra. Hazánk kibocsátásának előrejelzése ARIMA(3,0,1) modellel (fekete szakasz) 99%-os konfidenciaintervallumon



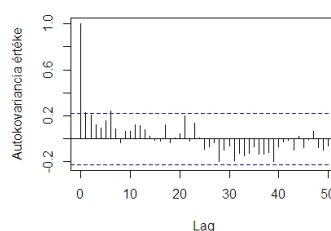
5.8. ábra. Hazánk kibocsátásának előrejelzése ARIMA(0,2,1) modellel 99%-os konfidenciaintervallumon

5.2. Magyarország egy főre jutó kibocsátásának elemzése és előrejelzése

Érdemes megnézni a modellek működését az egy főre eső adatbázis alapján is, ugyancsak 1940. és 2017. közti időszakra szűkítve. Itt a kibocsátások tonna/főben vannak, így nincs szükség átváltásra. Az adatsort itt is idősorra kellett konvertálni, majd megnézni az autokovarianciákat.



(a) Eredeti idősor



(b) Stacionárius idősor

5.9. ábra. Hazánk egy főre eső kibocsátásának autokovariancia vizsgálata

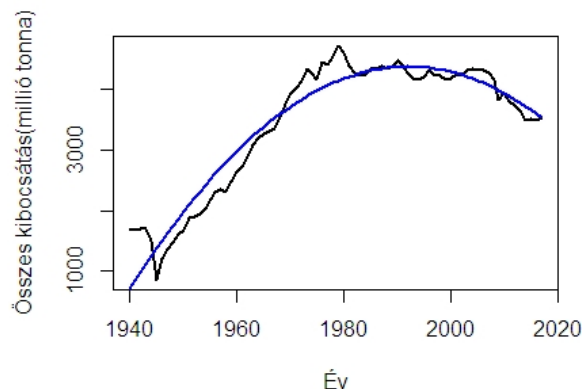
Az eredeti idősor autokorreláció vizsgálata alapján jelen esetben is elvethető a stacionaritás.

Az elemzést ugyanúgy végeztem, mint az előző esetekben, kvadratikus regresszióval kezdtem. A magyarázó változók (idő és annak négyzete) szignifikánsa a lehető legjobb; a p-érték nagyon kicsi, így valószínűtlen a nulla együttható.

	Együttható	Standard hiba
Intercept	-0,1113	0,2409
t	0,3141	0,0141
t ²	-0,003	0,0001

5.14. táblázat. Egy főre jutó kibocsátás kvadratikus regresszióvizsgálatának együtthatói

A következő lépés az ARIMA modellek illesztése. A regresszió reziduálisaira az egyfo_ARIMAmoell-t, az idősorra az egyfo_ARIMAmoell2-t használtam.



5.10. ábra. Egy főre jutó kibocsátására illesztett kvadratikus regresszió

	ARIMA (3,0,1)	Standard error
AR1	1,97	0,12
AR2	-1,06	0,23
AR3	0,06	0,12
MA1	-0,96	0,04

5.15. táblázat. Az egyfo_ARIMAmodell modell együtthatói és standard hibáik

	ARIMA (0,2,1)	Standard error
MA1	-0,85	0,08

5.16. táblázat. Az egyfo_ARIMAmodell2 modell együtthatói és standard hibáik

	ARIMA (3,0,1)	ARIMA (0,2,1)
AIC	26,82	32,62
BIC	38,6	37,28
sum(res ²)	5,42	6,38

5.17. táblázat. Egy főre jutó kibocsátás ARIMA (3,0,1) és ARIMA (0,2,1) összehasonlítása

A kapott ARIMA modellek paraméterei megegyeznek a magyarországi teljes kibocsátást leíró ARIMA modellek paramétereivel: a regresszió reziduálisaira illesztett modell egy ARIMA(3,0,1) modell és az idősorra illesztett ARIMA paraméterei (0,2,1), azaz egy 1 rendű MA kétszeres differenciálással.

A modelleket összehasonlítva, AIC és $\text{sum}(\text{residuals}^2)$ szerint a regresszió reziduálisaira illesztett, BIC szerint az idősorra illesztett ARIMA modell a jobb.

A Ljung-Box tesztet elvégezve a következő p-értékek jönnek ki: az elsőre $p = 0,7039$; a másodikra $p = 0,6487$.

Az ARIMA modellekkel adhatunk egy előrejelzést a következő 3 évre, ugyanazokkal a szignifikanciaszintekkel, mint a teljes kibocsátás előrejelzésénél.

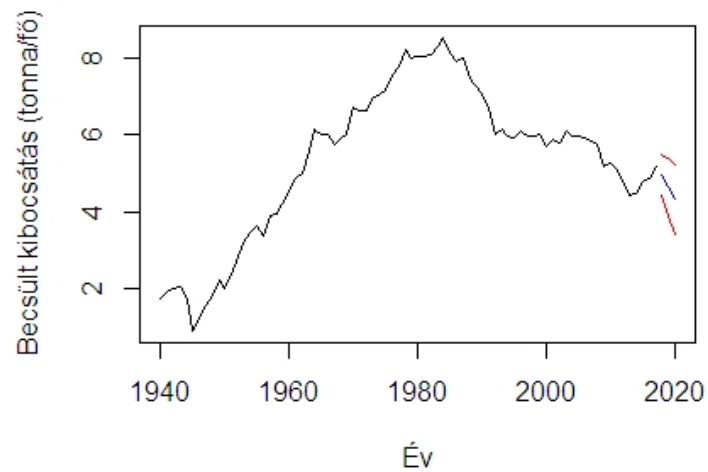
	ARIMA(3,0,1)	95 %	99 %
2018	5	[4; 5]	[4; 6]
2019	5	[4; 5]	[4; 6]
2020	4	[3; 5]	[3; 5]

5.18. táblázat. Egy főre jutó kibocsátás regresszió alapuló becslése és korlátai

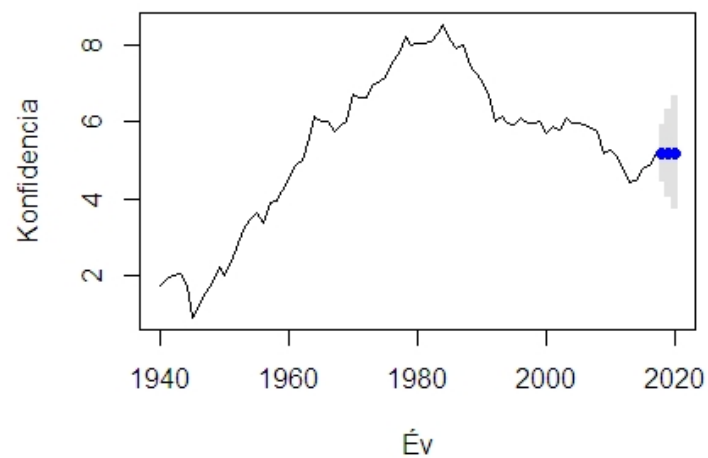
	ARIMA(0,2,1)	95 %	99 %
2018	5	[5,6]	[4,6]
2019	5	[4,6]	[4,6]
2020	5	[4,6]	[4,7]

5.19. táblázat. Egy főre jutó kibocsátás idősoron alapuló becslése és korlátai

Az egy főre jutó előrejelzés esetében az ARIMA modellek szinte teljesen egyező konfidenciaszinteket adnak vissza az adott szignifikanciaszinteken (a teljes kibocsátás előrejelzéseivel ellentétben).



5.11. ábra. Egy főre jutó kibocsátás előrejelzése ARIMA(3,0,1) modellel 99%-os konfidenciaszinten



5.12. ábra. Egy főre jutó kibocsátás előrejelzése ARIMA(0,2,1) modellel 99%-os konfidenciaszinten

6. fejezet

Konklúzió

Dolgozatomban megpróbáltam egy reális elemzést és lehetséges előrejelzést bemutatni a CO_2 kibocsátásra. Két módszert is bemutatam, amelyek a legtöbb esetben hasonló eredményre vezettek. Az ARIMA módszert használva konfidenciaintervallumokat is meghatározva nagyobb megbízhatóságot kapunk, hiszen a CO_2 kibocsátás rengeteg tényezőtől függ, amit a modell nem tud kezelni. Pontosabb előrejelzéshez külön kellene vizsgálni az egyes kibocsátási ágazatok eddigi emisszióit, majd ezek alapján egy összesített előrejelzést adni. Sajnos így sem lehet "tökéletes" az előrejelzés, mert vannak olyan változók, amiket nem lehet figyelembe venni. Ilyen tényezőre jó példa a mostani COVID-19 világjárvány is.[3] A kijárási korlátozások miatt kevesebben szállnak gépjárművekbe, visszaesett a kereslet és így a termelés is, ezek miatt kevesebb az antropogén eredetű CO_2 kibocsátás. A világjárvány visszahúzódtól függően még változhatnak az emissziós értékek, ám a végső kibocsátásra már lehetetlen egy ARIMA modell felépítése.

Irodalomjegyzék

- [1] Abdel Karim Baareh, *CO₂ Solving the Carbon Dioxide Emission Estimation Problem: An Artificial Neural Network Model*,
Journal of Software Engineering and Applications Vol.6 No.7, 2013.,p.338-342
- [2] Abdur Rahman and Md Mahmudul Hasan, *CO₂ Modeling and Forecasing of Carbon Dioxide Emissions in Bangladesh Using Autoregressive Intergrated Moving Average (ARIMA) Models*,
Open Journal of Statistics Vol.7 No.4, 2017., p. 560-566
- [3] BBC News, *Climate change and coronavirus: Five charts about the biggest carbon crash*,
<https://www.bbc.com/news/science-environment-52485712>
- [4] Dr. Ron Tibben-Lembke, *Confidence Intervals for Forecasting*, 2006.
- [5] Environmental Protection Agency, *Global Greenhouse Gas Emissions Data*,
<https://www.epa.gov/ghgemissions/global-greenhouse-gas-emissions-data>
- [6] Európai Parlament, *Üvegházhatású gázok kibocsátása az EU-ban (infografika)*,
<https://www.europarl.europa.eu/news/hu/headlines/society/20180301STO98928/uveghazhatasu-gazok-kibocsatasu-az-eu-ban-infografika>
- [7] Faragó T.,Folányi Zs., Pálvölgyi T., Poós M., *Az üvegházhatású gázok kibocsátásának csökkentése: kiotói jegyzőkönyv az ENSZ Éghajlatváltozási Keretegyezményhez és a hazai feladatok*, 1998.
http://real.mtak.hu/65839/1/Kiotoi_Jegyzokonyv.pdf
- [8] Földművelésügyi Minisztérium, *Vegyianyagok szabályozása Genfi Egyezmény*,
<https://vegyszeranyag.kormany.hu/lrtap>

- [9] Hannah Ritchie and Max Roser, *CO₂ and Greenhouse Gas Emissions*,
[https://ourworldindata.org/co2-and-other-greenhouse-\
\gas-emissions](https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions)
- [10] Hulyák Katalin, *Idősorok rövidtávú előrejelzése ARIMA modellekkel*,
Sigma Matematikai-Közgazdasági Folyóirat, 1977.
- [11] Központi Statisztikai Hivatal, *Nemzetgazdasági ágak és háztartások szén-dioxid (CO₂) kibocsátása (1985–)*, https://www.ksh.hu/docs/hun/xstadat/xstadat_eves/i_ua026d.html
- [12] Lazim Abdullah and Herrini Mohd Pauzi, *CO₂ Methods in Forecasting Carbon Dioxide Emissions: A Decade Review*,
UTM Jurnal Teknologi Vol.75 No.1, 2015.,p. 67-82
- [13] Márkus László, *Idősorok előadásjegyzet*,
[https://web.cs.elte.hu/probability/markus/
AlkmatTS1/](https://web.cs.elte.hu/probability/markus/AlkmatTS1/)
- [14] *Nemzetközi Egyezmények*,
[https://www.parlament.hu/biz37/korb/torveny/
nemzetkozi.htm](https://www.parlament.hu/biz37/korb/torveny/nemzetkozi.htm)
- [15] Szaktudás Kiadó Ház ZRt., *"A" Tételű modul - Környezetvédelem*,
[https://regi.tankonyvtar.hu/hu/tartalom/tamop425/
0032_kornyezetvedelem/ch14s02.html](https://regi.tankonyvtar.hu/hu/tartalom/tamop425/0032_kornyezetvedelem/ch14s02.html)
- [16] Prof. Dr. Závoti József, *Matematikai statisztikai elemzések 3., Becslésselmélet: alapfogalmak, nevezetes statisztikák, intervallum-becslés*,
Nyugat-magyarországi Egyetem, 2010.
- [17] *Quick Guide: Interpreting Simple Linear Model Output in R*
[https://feliperego.github.io/blog/2015/10/23/
Interpreting-Model-Output-In-R](https://feliperego.github.io/blog/2015/10/23/Interpreting-Model-Output-In-R)
- [18] *RDocumentation* <https://www.rdocumentation.org/>
- [19] Ruslana Dalinina, *Introduction to Forecasting with ARIMA in R*,
[https://blogs.oracle.com/datascience/
introduction-to-forecasting-with-arima-in-r](https://blogs.oracle.com/datascience/introduction-to-forecasting-with-arima-in-r)
- [20] Tihanyi László, Horánszky Beáta, *Európai trendek a szén-dioxid-kibocsátás területén*,
Műszaki Földtudományi Közlemények 83.kötet 1.szám, 2012., p. 249–257.

- [21] United Nations Information Service, *Az Egyesült Nemzetek Éghajlatváltozásról szóló Keretegyezménye*
http://www.unis.unvienna.org/unis/hu/thematic_info_climate_change_unfccc.html
- [22] *Üvegházhatású gázok*,
<https://www.metnet.hu/kislexikon/uveghazhatasu-gazok>
- [23] Zempléni András, *Idősorok és többdimenziós statisztika előadásjegyzet*
- [24] Zempléni András, *Leíró és matematikai statisztika előadásjegyzet*

Függelék

Az Európai Unió kibocsátásának elemzése és előrejelzése, R programkódban. A Magyarországi elemzések ehhez hasonlóak.

```
# CSV fájl betöltése
annual_co_emissions_by_region=read_csv("annual-co-emissions-by-region.csv")

# Új tábla, mely csak az 1940-es évektől tartalmazza
# a kibocsátásokat és nem szerepel a "Code" oszlop
tabla=subset(annual_co_emissions_by_region,Year>=1940,
  select = -c(Code))

# Új tábla, mely csak az EU-28-ak kibocsátásait tartalmazza
# és nem szerepelnek az "Entity" és "Year" oszlopok
EUösszesített=subset(tabla, Entity=="EU-28",
  select = -c(Entity, Year))
EUösszesített

# Idősor-e?
is.ts(EUösszesített)

# Adatok idősorra alakítása, és az értékek leosztása
# a mértékegység miatt
EUadatok=(ts(EUösszesített,frequency=1,start=c(1940,1),
  end=c(2017,1)))/1000000
plot(EUadatok,xlab="Év",ylab="CO2_kibocsátás(millió_tonna)")

# Stacionaritás és autokorreláció vizsgálata
EU_stacionarius = diff(EUadatok, differences=1)
plot(EU_stacionarius,xlab="Év", ylab="CO2_kibocsátás(millió_tonna)")
acf(EUadatok,lag.max=50,main="_",xlab="Lag",
  ylab="Autokovariancia_értéke")
acf(EU_stacionarius,lag.max=50, main="_",xlab="Lag", y
```

```

lab="Autokovariancia_erteke")

# Regresszio
adf.test(EUadatok)
t=c(1:78)
t2=t^2
EU_lr=lm(EUadatok~t+t2)
summary(EU_lr)
plot(EUadatok,type="l",lwd=2.5, xlab="Ev",
ylab="Osszes_kibocsatas(millio_tonna) ")
lines(c(1940:2017),eu_lr$fitted.values,lwd=2.5,col=4)
AIC(EU_lr)
BIC(EU_lr)
sum(EU_lr$residuals^2)

```

```

Call:
lm(formula = EUadatok ~ t + t2)

Residuals:
    Min       1Q   Median       3Q      Max
-522.65 -196.33  -43.82  160.17 1000.62

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  559.88070   106.35842    5.264  1.3e-06 ***
t            144.07146    6.21348   23.187  < 2e-16 ***
t2           -1.35830    0.07622  -17.822  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 305.1 on 75 degrees of freedom
Multiple R-squared:  0.9224,    Adjusted R-squared:  0.9203
F-statistic: 445.9 on 2 and 75 DF,  p-value: < 2.2e-16

```

Magyarázat a lineáris regressziós módszer output-hoz:

A "Residuals" megadja az előrejelzés és a tényleges értékek közti eltérést. A "Coefficients", azaz az együtthatók azok a becslések, melyek minimalizálják a hibák négyzetének összegét: az "Estimate" a becslült érték; az "Std. Error" a standard hiba; a "t-value" a becslés értéke elosztva a standard hibával; a "Pr(>|t|)" pedig a t-eloszlás táblázatból nézhető meg. A "Signif.codes" a szignifikancia szintjeit jelöli. A "residuals standard error" a reziduálisok szórása a regressziós feltételezések alapján, ez mutatja, hogy a pontok milyen közel vannak az illesztéshez. A "Multiple R-squared" a magyarázott változó varianciájának százaléka a modell hibájának kivonása után. Az "Adjusted R-squared" ugyanaz, mint az előző, de ez már figyelembe veszi a felhasznált paraméterek számát is. Az "F-statistic"

egy globális teszt, mely azt mondja meg, hogy a modellnek van-e legalább egy szignifikáns, azaz nem nulla változója.

```
# ARIMA modell a regresszióra, illeszkedés megvizsgálása
EU_ARIMAmodell=auto.arima(EU_lr$residuals)
EU_ARIMAmodell
EU_ARIMAmodell$LjungBox
EU_teszt=Box.test(EU_ARIMAmodell,type="Ljung-Box")
EU_teszt
sum(EU_ARIMAmodell$residuals^2)
```

```
> EU_ARIMAmodell
Series: EU_lr$residuals
ARIMA(2,0,1) with zero mean

Coefficients:
          ar1          ar2          ma1
      1.9083   -0.9402   -0.9724
s.e.  0.0399   0.0395   0.0471

sigma^2 estimated as 16696:  log likelihood=-490.26
AIC=988.51   AICc=989.06   BIC=997.94
> sum(EU_ARIMAmodell$residuals^2)
[1] 1252166
```

```
# Elorejelzes, n=3 evre
EU_regelorejelzes=predict(EU_ARIMAmodell,n.ahead = 3)
EU_regelorejelzes
plot(EU_regelorejelzes$pred, xlab=" v ",
     ylab="Becsult_kibocsatasok_(millio_tonna)")
matr=cbind(c(1,1,1),c(79:81),c(79:81)^2)
predlm=eu_lr$coef%*%t(matr)
plot(c(2018:2020),t(EU_regelorejelzes$pred[1:3]+predlm),
     xlab="Ev", ylab="Becsult_kibocsatasok_(millio_tonna)")

# Konfidenciaintervallum: 95% s 99%
EU_regkonfidencia95= forecast(EU_ARIMAmodell,h=3, level=95)
EU_regkonfidencia95
EU_regkonfidencia95m=as.vector(EU_regkonfidencia95$mean)+predlm
EU_regkonfidencia95l=as.vector(EU_regkonfidencia95$lower)+predlm
EU_regkonfidencia95u=as.vector(EU_regkonfidencia95$upper)+predlm
EU_regkonfidencia95c=cbind(EU_regkonfidencia95m,EU_regkonfidencia95l,
EU_regkonfidencia95u)
```

```

plot(c(1940:2017), EUadatok[, 1], xlim=c(1940, 2020), main="_",
     xlab="Ev", ylab="Becsult_kibocsatas_(millio_tonna)", type="l")
lines(c(2018:2020), t(EU_regkonfidencia95m), col=4)
lines(c(2018:2020), t(EU_regkonfidencia95l), col=2)
lines(c(2018:2020), t(EU_regkonfidencia95u), col=2)

EU_regkonfidencia99= forecast(EU_ARIMAmode11, h=3, level=99)
EU_regkonfidencia99
EU_regkonfidencia99m=as.vector(EU_regkonfidencia99$mean)+predlm
EU_regkonfidencia99l=as.vector(EU_regkonfidencia99$lower)+predlm
EU_regkonfidencia99u=as.vector(EU_regkonfidencia99$upper)+predlm
EU_regkonfidencia99c=cbind(EU_regkonfidencia99m, EU_regkonfidencia99l,
EU_regkonfidencia99u)

plot(c(1940:2017), EUadatok[, 1], xlim=c(1940, 2020), main="_",
     xlab="Ev", ylab="Becsult_kibocsatas_(millio_tonna)", type="l", lwd=2.5)
lines(c(2018:2020), t(EU_regkonfidencia99m), lwd=2.5, col=4)
lines(c(2018:2020), t(EU_regkonfidencia99l), lwd=2.5, col=3)
lines(c(2018:2020), t(EU_regkonfidencia99u), lwd=2.5, col=3)

# ARIMA modell az idosorra
EU_ARIMAmode112= auto.arima(EUadatok, seasonal = FALSE)
EU_ARIMAmode112
EU_ARIMAmode112$LjungBox
EU_teszt2=Box.test(EU_ARIMAmode112$residuals, type="Ljung-Box")
EU_teszt2
sum(EU_ARIMAmode112$residuals^2)

> EU_ARIMAmode112
Series: EUadatok
ARIMA(0,2,1)

Coefficients:
      ma1
    -0.874
s.e.    0.065

sigma^2 estimated as 19447: log likelihood=-483.32
AIC=970.65   AICc=970.81   BIC=975.31
> sum(EU_ARIMAmode112$residuals^2)
[1] 1458488

```

```

# Elorejelzes, n=3 evre
EU_elorejelzes=predict(EU_ARIMAmoell2,n.ahead = 3)
EU_elorejelzes
plot(EU_elorejelzes$pred, xlab="Ev",
ylab="Becsult_kibocsatasok_(millio_tonna)")

# Konfidenciaintervallum: 95% s 99%
EU_konfidencia95= forecast(EU_ARIMAmoell2,h=3, level=95)
EU_konfidencia95
plot(EU_konfidencia95,main="_", xlab="Ev",
ylab="Becsult_kibocsatas_(milli_tonna)", fcol=3)

EU_konfidencia99= forecast(EU_ARIMAmoell2,h=3, level=99)
EU_konfidencia99
plot(EU_konfidencia99, main="_", xlab=" v ",
ylab="Becsult_kibocsatas_(millio_tonna)",fcol=4)

```