DIMENZIÓK
Matematikai Közlemények
VII. kötet, 2019

doi:10.20312/dim.2019.05

Kettős Gauss függvény alkalmazása

Csanády Viktória
SOE Matematikai Intézet
csanady.viktoria@uni-sopron.hu

ÖSSZEFOGLALÓ. Az időbeli folyamatok jellemzése esetén gyakran a folyamat tendenciáját vizsgálva szükségünk lehet bonyolultabb matematikai modellek alkalmazására. A számítógépes statisztikai programok által felkínált alap modellek viszont erre nem adnak kielégítő választ, a felhasználónak kell a speciális modellt megtalálni, majd a gép adta lehetőség révén alkalmazni. Az alábbiakban két speciális kettős Gauss függvény kerül bemutatásra, annak rugalmasságát kiemelve, gyakorlati példákon illusztrálva.

ABSTRACT. To describe tendencies in time series processes we often need to apply complex mathematical models. The basic models offered by statistical computer programmes are usually not applicable, so the user must find the special model, apply it and exploit the computers' potentials. We present two special double-Gaussian functions below, stress their flexibilities through practical examples.

1. Bevezetés

Az egyes időfüggő folyamatok jellemzése során, amikor azok, nem mutatnak határozott tendenciát menetükben (monoton növekedés ill. csökkenés), gyakran csak a mérési pontok összekötésére hagyatkoznak ami persze egy meglehetősen primitív megoldás, nem beszélve arról, hogy a főbb matematikai jellemzők ezáltal nem vállnak ismerté. Ilyen esetekben, ha mégis modellillesztésre kerül sor, akkor az alkalmasnak tűnő modell általában - amit használnak - egy magasabb fokszámú polinom függvény. Ez a függvény persze számos előírható és elvárható kritériumnak nem tesz eleget, továbbá az illesztésnél szükséges kezdőértékek megadása problémás, olykor sok időt és energiát felemésztő feladat. Az említett kritériumok közül a legkézenfekvőbb a végtelenekben történő viselkedése a polinom függvénynek, ami a vizsgált időbeli folyamatokra egyáltalán nem jellemző, így az alkalmasság megkérdőjelezhető. Az említett folyamatok esetében egyszerűbb a helyzet akkor, amikor lokális szélsőértékkel nem, vagy csak egy ilyen jellemző ponttal bír a leírandó pontsorozat. Ezekben az esetekben általában jól alkalmazhatók az un. telítési függvények un. életfüggvények és ezek lineáris kombinációi, bonyolultabb esetekben szuperponált modelljei. Ennek fő indokai a függvények kedvező tulajdonságaiból fakadnak, így a zérushelyből történő indíthatóságból, a függvény korlátosságából illetve aszimptotikus tulajdonságából. Azokban az esetekben viszont, amikor a folyamatot jellemző adatsor több szélsőértékkel bír, inflexiós pontok megléte feltételezhető, akkor az említett modellek már nem alkalmasak a folyamat jellemzésére. Ez az illesztési probléma ösztönözte a gondolatot újabb összetett modellek létrehozására, melyek több szélsőérték esetén is működőképes rugalmas modellek, melyekből igény esetén számíthatók a matematikai jellemzők. Az alábbiakban bemutatásra kerülnek az összetett modellek illetve azok alkalmazása két erdészeti vonatkozásban, a felhasznált adatok a Központi Statisztikai Hivatal adatbázisából származnak.

A vizsgálat tárgya:

- az 1996-2017 éves időszakban évenként kitermelt tölgy, nyár és akác fa mennyiségének időbeli változása. 1. táblázat: Kitermelt faanyag.
- az 1995/1996-2016/2017 időszak évenkénti fásítás-erdőtelepítés mértékének időbeli változása. 2. táblázat: Erdőtelepítés.

A vizsgált adathalmazok és az alkalmazott modellek:

A lehetséges reprodukálhatóság végett megadásra kerülnek a vizsgált adatsorok.

	FAKITERMELÉS 1000 m^3/év			
	Év (VAR1)	TÖLGY (VAR2)	NYÁR (VAR3)	AKÁC (VAR4)
1	1996	980	1252	1370
2	1997	1009	1188	1331
3	1998	1033	1207	1266
4	1999	1104	1170	1344
5	2000	1184	1069	1496
6	2001	1158	1090	1425
7	2002	1161	976	1480
8	2003	1133	921	1527
9	2004	1119	933	1462
10	2005	1188	883	1401
11	2006	1157	898	1351
12	2007	1189	875	1206
13	2008	1166	921	1427
14	2009	1052	942	1480
15	2010	1102	980	1628
16	2011	1104	1212	1857
17	2012	1055	1113	1793
18	2013	1039	1115	1745
19	2014	941	1100	1541
20	2015	941	1072	1488
21	2016	938	1071	1429
22	2017	898	1020	1488

1. táblázat. Kitermelt faanyag

	Erdőtelepítés hektár/év			
	TÉNY IDŐSZAK	ÉV (VAR1)	ELSŐ TELEPÍTÉS (VAR2)	ÖSSZ TELEPÍTÉS (VAR3)
1	1995/96	1995	6610	7804
2	1996/97	1996	8289	9742
3	1997/98	1997	8183	9795
4	1998/99	1998	8661	9858
5	1999/00	1999	9790	10842
6	2000/01	2000	13150	15516
7	2001/02	2001	14830	17169
8	2002/03	2002	12015	15028
9	2003/04	2003	7574	11581
10	2004/05	2004	7657	9439
11	2005/06	2005	13989	15008
12	2006/07	2006	18948	20289
13	2007/08	2007	7332	9441
14	2008/09	2008	5168	6303
15	2009/10	2009	5096	5960
16	2010/11	2010	2803	3461
17	2011/12	2011	4537	5009
18	2012/13	2012	2530	3250
19	2013/14	2013	1287	1599
20	2014/15	2014	318	452
21	2015/16	2015	158	300
22	2016/17	2016	626	694

2. táblázat. Erdőtelepítés

Az alkalmazott regressziós modellek:

- Két Gauss függvény kompozíciója (GAUÉGAU)

- matematikai alakja:

$$y = \frac{b_6}{e^{\left(b_5(x-b_4)\right)^2}} + \frac{b_3}{e^{\left(b_2(x-b_1)\right)^2}} + b_0$$

- számítógépes alak:

var2=b6/exp((b5*(var1-1*b4))^2)+b3/exp((b2*(var1-1*b1))^2)+b0.

Kezdőérték választás általános (2 maximum vagy minimum) esetben az adatsor értékei alapján:

$b6=var2_{\text{első max.}}-var2_{\text{min.}}$ vagy $b6=var2_{\text{első min.}}-var2_{\text{max.}}$

$b3=var2_{\text{másod. max.}}-var2_{\text{min.}}$ vagy $b3=var2_{\text{másod. min.}}-var2_{\text{max.}}$

$b4=var1_{\text{első max.}}$ vagy $var1_{\text{első min.}}$

$b1=var1_{\text{másod. max.}}$ vagy $var1_{\text{másod. min.}}$

$b5 \sim 0{,}05$
$b2 \sim 0{,}05$
$b0 = var2_{min.}$

- Két abszolútértékes módosított Gauss függvény kompozíciója (ABGAUÉABGAU)

- matematikai alakja:

$$y = \frac{b_8}{e^{(b_7|x-b_6|)^{b_5}}} + \frac{b_4}{e^{(b_3|x-b_2|)^{b_1}}} + b_0$$

- számítógépes alak:

var2=b8/exp((b7*(abs(var1-1*b6)))^b5)+b4/exp((b3*(abs(var1-1*b2)))^b1)+b0.

Kezdőérték választás általános (2 maximum vagy minimum) esetben az adatsor értékei alapján:

$b8=var2_{\text{első max.}}-var2_{min.}$ vagy $b8=var2_{\text{első min.}}-var2_{max.}$
$b4=var2_{\text{másod. max.}}-var2_{min.}$ vagy $b4=var2_{\text{másod. min.}}-var2_{max.}$
$b6=var1_{\text{első max.}}$ vagy $var1_{\text{első min.}}$
$b2=var1_{\text{másod. max.}}$ vagy $var1_{\text{másod. min.}}$
$b7 \sim 0{,}1$
$b3 \sim 0{,}1$
$b5 \sim 3$
$b1 \sim 3$
$b0 = var2_{min.}$

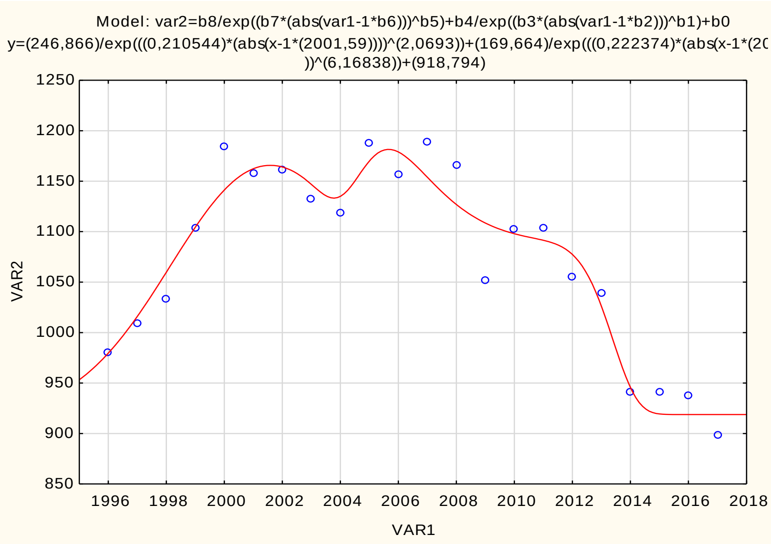
A modellek rövid jellemzése:

Az alkalmazott GAUÉGAU két transzformált Gauss függvény összegéből áll egy függőleges eltolással, ez adja meg a két szélsőérték létezésének feltételét továbbá a lehetséges asszimetriát. Az ABGAUÉABGAU függvény esetében a kitevő kitevője nem kettő, tetszőleges érték lehet, ami magával vonja az abszolútérték szükségességét. Az említett tetszőleges kitevő kitevője a modell rugalmaságát lényegesen növeli, akár törési pont is előfordulhat a szélsőértékek mellett a modellben.
Első esetben 7, második esetben 9 paraméter befolyásolja a függvény alakját, ezen értékek kezdőértékeinek megadása viszont a már említett módokon könnyűszerrel megadható a pontsorozat ismeretében.

2. Számított eredmények, kiértékelés

2.1. A fakitermelési adatsorok regressziós eredményei

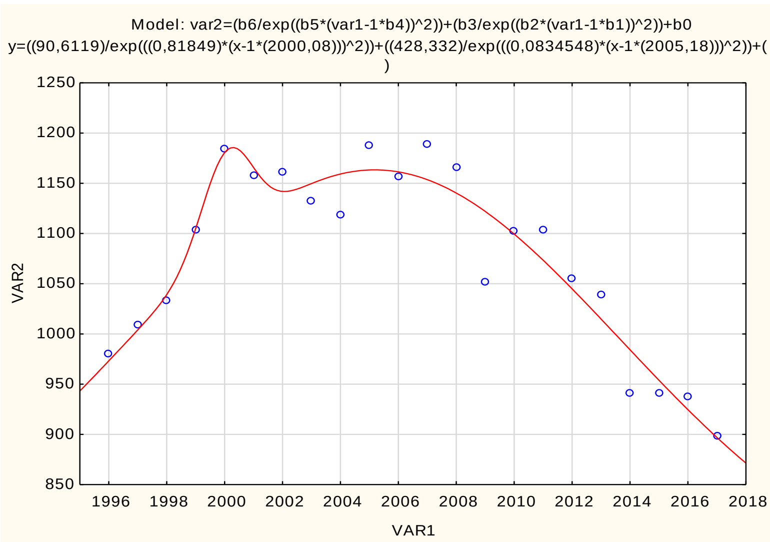
Mind a három fa faj esetén elsőként az ABGAUÉABGAU alkalmazás eredményeit tüntetjük fel majd közvetlen utána a GAUÉGAU eredményeit illetve azok grafikus repezentációját.



1. ábra. Tölgy ABGAUÉABGAU

	Model: var2=b8/exp((b7*(abs(var1-1*b6)))^b5)+b4/exp((b3*(a... (Tölgy) Dep. var: VAR2 Loss: (OBS-PRED)**2 Final loss: 11954,830685 R= ,96584 Variance explained: 93,285%								
N=22	b8	b7	b6	b5	b4	b3	b2	b1	b0
Estimate	246,8658	0,210544	2001,592	2,069304	169,6643	0,222374	2009,035	6,168376	918,7937

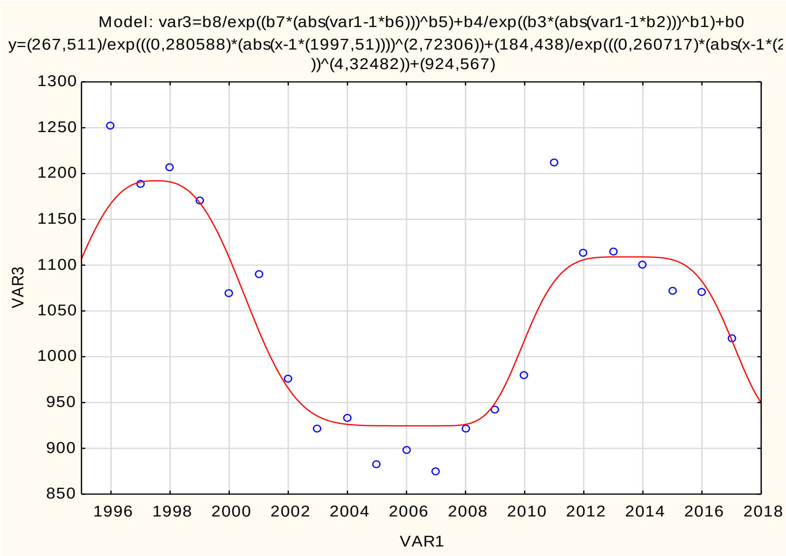
3. táblázat. Tölgy ABGAUÉABGAU



2. ábra. Tölgy GAUÉGAU

	Model: var2=(b6/exp((b5*(var1-1*b4))^2))+(b3/exp((b2*(var1... (Tölgy) Dep. var: VAR2 Loss: (OBS-PRED)**2 Final loss: 13732,520756 R= ,96066 Variance explained: 92,286%						
N=22	b6	b5	b4	b3	b2	b1	b0
Estimate	90,6119	0,81849	2000,07	428,332	0,08345	2005,18	735,043

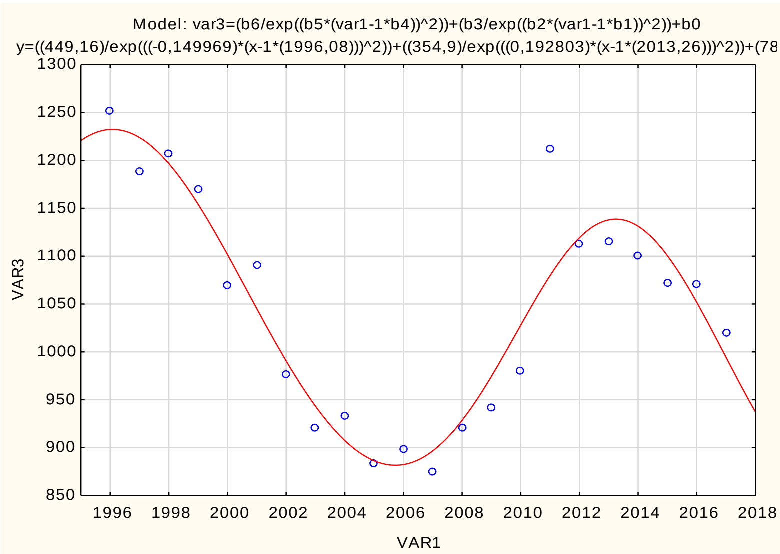
4. táblázat. Tölgy GAUÉGAU



3. ábra. Nyár ABGAUÉABGAU

	Model: var3=b8/exp((b7*(abs(var1-1*b6)))^b5)+b4/exp((b3*(...(Nyár) Dep. var: VAR3 Loss: (OBS-PRED)**2 Final loss: 37826,017227 R= ,93247 Variance explained: 86,950%								
N=22	b8	b7	b6	b5	b4	b3	b2	b1	b0
Estimate	267,511(	0,28058{	1997,50{	2,72306:	184,438:	0,26071:	2013,49{	4,32482{	924,567:

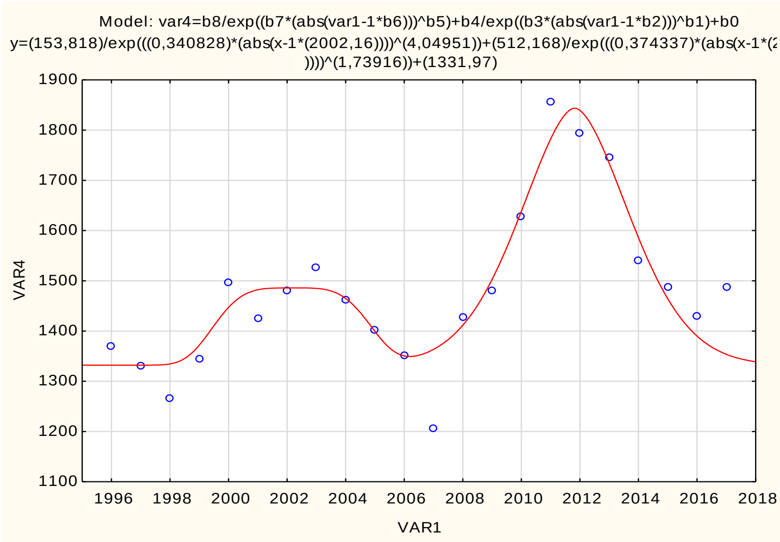
5. táblázat. Nyár ABGAUÉABGAU



4. ábra. Nyár GAUÉGAU

	Model: var3=(b6/exp((b5*(var1-1*b4))^2))+(b3/exp((b2*(var1... (Nyár) Dep. var: VAR3 Loss: (OBS-PRED)**2 Final loss: 31501,971245 R= ,94410 Variance explained: 89,132%						
N=22	b6	b5	b4	b3	b2	b1	b0
Estimate	449,159{	-0,14996{	1996,08:	354,899{	0,19280:	2013,25{	783,235⁄

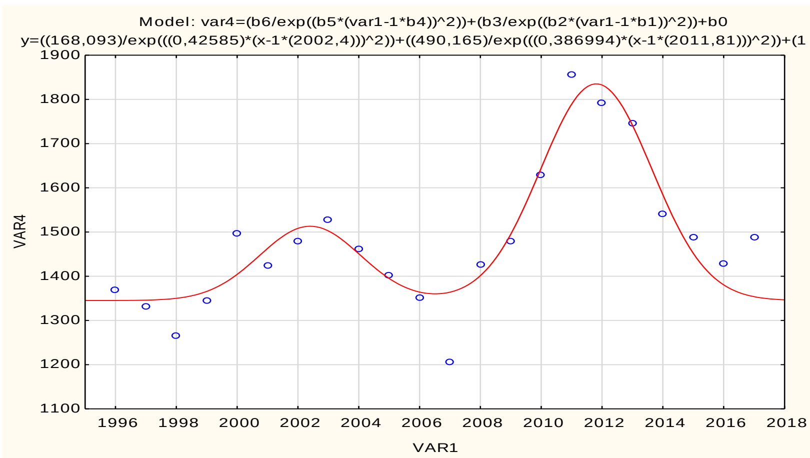
6. táblázat. Nyár GAUÉGAU



5. ábra. Akác ABGAUÉABGAU

	Model: var4=b8/exp((b7*(abs(var1-1*b6))))^b5)+b4/exp((b3*(...(Akác) Dep. var: VAR4 Loss: (OBS-PRED)**2 Final loss: 70028,264512 R= ,93340 Variance explained: 87,123%								
N=22	b8	b7	b6	b5	b4	b3	b2	b1	b0
Estimate	153,8184	0,340828	2002,158	4,049508	512,1679	0,374337	2011,821	1,739160	1331,966

7. táblázat. Akác ABGAUÉABGAU



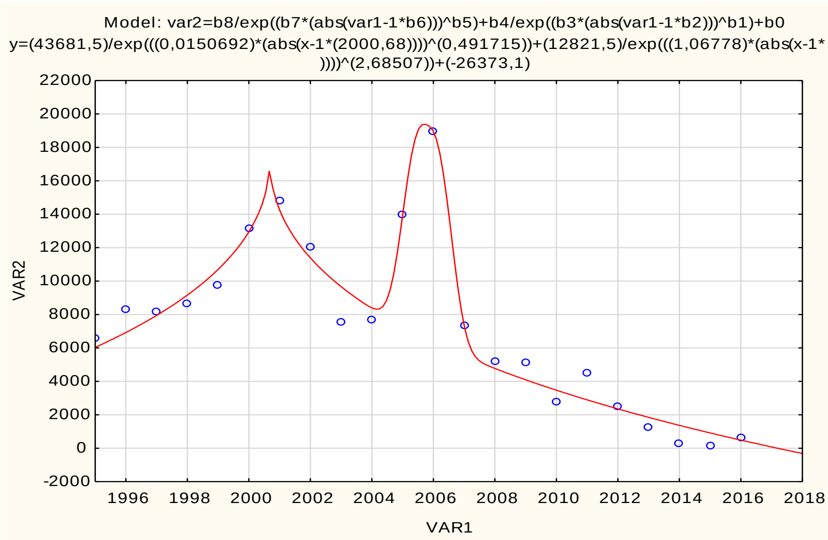
6. ábra. Akác GAUÉGAU

	Model: var4=(b6/exp((b5*(var1-1*b4))^2))+(b3/exp((b2*(var1... (Akác) Dep. var: VAR4 Loss: (OBS-PRED)**2 Final loss: 75839,028602 R= ,92766 Variance explained: 86,054%						
N=22	b6	b5	b4	b3	b2	b1	b0
Estimate	168,092	0,42585	2002,40	490,164	0,38699	2011,80	1345,10

8. táblázat. Akác GAUÉGAU

2.2. Az erdőtelepítési adatsorok regressziós eredményei

Mind a két erdőtelepítési adatsor esetén az illesztési sorrend változatlan (2.1)

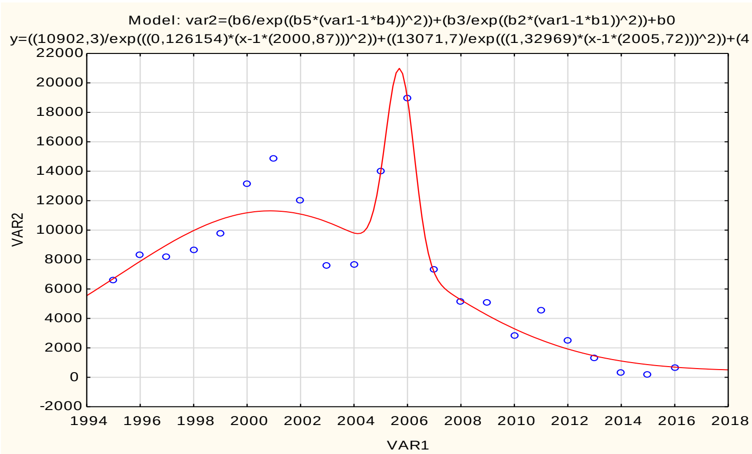


7. ábra. Első telepítés ABGAUÉABGAU

Model: var2=b8/exp((b7*(abs(var1-1*b6)))^b5)+b4/exp((b3*(a... (Fásítás)
Dep. var: VAR2 Loss: (OBS-PRED)**2
Final loss: 15386487,604 R= ,98577 Variance explained: 97,174%

N=22	b8	b7	b6	b5	b4	b3	b2	b1	b0
Estimate	43681,49	0,015069	2000,683	0,491715	12821,50	1,067781	2005,798	2,685068	-26373,1

9. táblázat. Első telepítés ABGAUÉABGAU

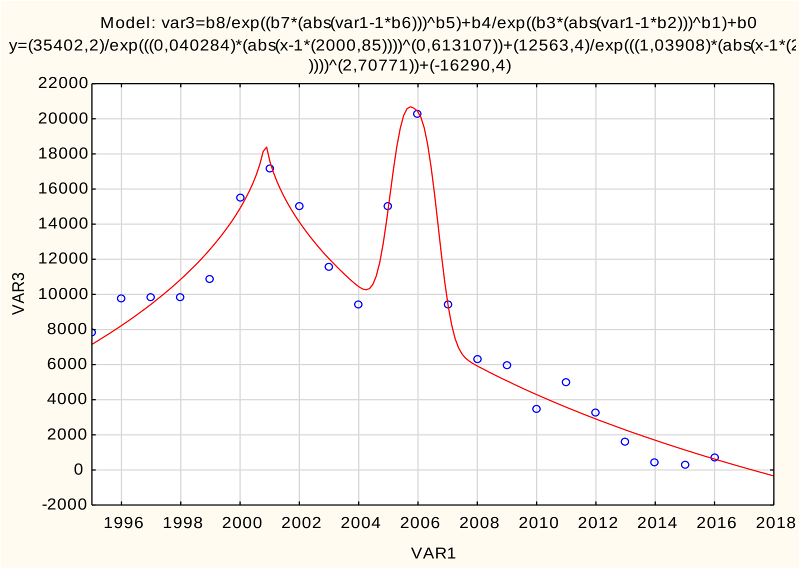


8. ábra. Első telepítés GAUÉGAU

Model: var2=(b6/exp((b5*(var1-1*b4))^2))+(b3/exp((b2*(var1... (Fásítás)
Dep. var: VAR2 Loss: (OBS-PRED)**2
Final loss: 40639234,125 R= ,96196 Variance explained: 92,536%

N=22	b6	b5	b4	b3	b2	b1	b0
Estimate	10902,3	0,12615	2000,86	13071,7	1,32969	2005,72	405,569

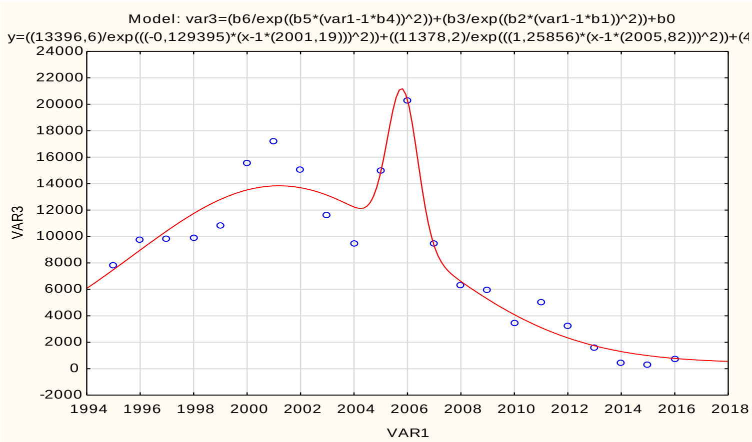
10. táblázat. Első telepítés GAUÉGAU



9. ábra. Összes telepítés ABGAUÉABGAU

	Model: var3=b8/exp((b7*(abs(var1-1*b6)))^b5)+b4/exp((b3*(a... (Fásítás)								
	Dep. var: VAR3 Loss: (OBS-PRED)**2								
	Final loss: 16064995,504 R= ,98825 Variance explained: 97,663%								
N=22	b8	b7	b6	b5	b4	b3	b2	b1	b0
Estimate	35402,21	0,040284	2000,848	0,613107	12563,43	1,039083	2005,872	2,707708	-16290,4

11. táblázat. Összes telepítés ABGAUÉABGAU



10. ábra. Összes telepítés GAUÉGAU

	Model: var3=(b6/exp((b5*(var1-1*b4))^2))+(b3/exp((b2*(var1... (Fásítás)						
	Dep. var: VAR3 Loss: (OBS-PRED)**2						
	Final loss: 42254557,076 R= ,96878 Variance explained: 93,854%						
N=22	b6	b5	b4	b3	b2	b1	b0
Estimate	13396,60	-0,129395	2001,191	11378,23	1,258563	2005,819	438,3610

12. táblázat. Összes telepítés GAUÉGAU

2.3. Elemzés, értékelés

***A fakitermelési adatok regressziós vizsgálatának elemzése fafajonként*:**

- *Tölgy*

A kapott R értékek alapján az ABGAUÉABGAU pontosabb illesztést eredményez, amit jól igazolnak az illesztési ábrák. Bár a szélsőértékek száma a két modell esetén megegyező, a pontsor követése az első esetben sokkal szorosabb.

- *Nyár*

A kapott R értékek alapján itt a GAUÉGAU mutat század nagyságrendben nagyobb értéket. Az ábrák vizsgálatából az tűnik ki, hogy az említett modell határozottabb lefutású a szélsőértékekre nézve, ennek oka a pontsorozat jellegéből adódik.

- *Akác*

A R értéke ennél a fafajnál az ABGAUÉABGAU esetében nagyobb. Bár hasonló jelleget mutat a két modell grafikonja a jobb illesztés észlelhető. A kettős Gauss esetében az egyes szélsőértékekhez tartozó hullámokat szimmetria jelzi míg a módosított kitevő esetén az asszimetria is fellelhető, ami indokolja a már említett nagyobb R értéket.

Mind a három fafaj esetében az egyes illesztések eredményeiből meghatározhatók a növekvő, csökkenő időszakok intervallumai mind a független mind a függő változóra, meghatározhatók a szélsőérték helyek és értékek. Az ábra alapján közelítőleg a modell alapján pontosan megadhatók az inflexiós helyek és értékek, melyek a változási sebesség meghatározó pontjai, valamint megadják a modell konvex illetve konkáv intervallum határait.

Értékelés:

A felsoroltak igazolják, hogy az alkalmazott modellek hasonló jellegű adatsorok vizsgálatára alkalmasak, könnyen használhatók és részletes tájékoztatók.

Az erdőtelepítési adatok regressziós vizsgálatának elemzése telepítési típusonként:

- *Első telepítés*

Az adatsor értékeinek jelentős ingadozását itt az ABGAUÉABGAU modell követi kellő pontossággal, amit a R=0,9857 érték is fémjelez, míg ugyan magas R értéket jelez a GAUÉGAU modell R=0,9619 a pontsorozat követése ennél azonban lényegesen pontatlanabb.

- *Összes telepítés*

Az adatsor hasonló jellegű az első telepítés adatsorához. Az eredmény itt is azt mutatja, hogy az ABGAUÉABGAU modell rugalmasan követi a pontsorozatot, amit alátámaszt az R érték, ami itt eléri a 0,9882-t. A kettős Gauss követési pontossága lényegesen gyengébb még az R=0,9687 magas értéke ellenére is.

Értékelés:

A fenti adatsorok igazolják a két modell közötti jelentős rugalmassági tulajdonsági eltérést. Bizonyos esetekben lehet elegendő a kettős Gauss alkalmazása is, ha eltekintünk az adatsor erős szóródásától, viszont ekkor a folyamat leírása pontatlanná válik. Ezzel szemben a módosított kitevő révén kapott modell kellő pontossággal írja le a folyamatot

3. Összefoglaló

A gyakorlati életből, az erdészet területéről választott adatsorok nem hasonlíthatók egy klasszikus fizikai, kémiai mérési sorozathoz, ahol a tendencia várható, a mért értékek szélsőséges ingadozása nem jellemző. Az itt vizsgált adathalmazok változékonyak, több maximum vagy minimum értékkel bírhatnak, növekedési vagy csökkenési sebesség változása mellett. Jellemzésükre nem alkalmasak a hagyományos egyszerű modellek. A bemutatott két modell, a kettős Gauss GAUÉGAU valamint a kitevőkben módosított ABGAUÉABGAU rendelkezik azzal a rugalmassággal melyre az illesztésnél szükség lehet.

Az eredmények igazolják, hogy különösen az utóbbi függvény alkalmas ehhez hasonló szélsőséges adatsorok jellemzésére megfelelő pontossággal. A függvény jól mutatja az adathalmaz szakaszainak menetét, szélsőértékeit, konvexitását, inflexiós pontjait, ugyanekkor magas korrelációs értéke jelzi az illesztés pontosságát. Mindez indokolja a függvény alkalmazását szélsőséges adatsorok esetén.

Irodalomjegyzék

[1] **Csanády V., Horváth–Szováti E., Szalay L.**, Alkalmazott statisztika, Sopron, Nyugat-Magyarországi Egyetem Kiadó (2013), 175p.

[2] Központi Statisztikai Hivatal. https://www.ksh.hu/stadat .