

DOKTORI (PhD) ÉRTEKEZÉS

NOVOTNI ADRIENN



**SOPRONI
EGYETEM**

FAIPARI MÉRNÖKI ÉS
KREATÍVIPARI
KAR

Sopron

2023

Doktori (PhD) értekezés

Soproni Egyetem

Faipari Mérnöki és Kreatívipari Kar

Cziráki József Faanyagtudomány és Technológiák Doktori Iskola

Vezető: Prof. Dr. Németh Róbert

Doktori program: Menedzsment a faiparban

Programvezető: Pakainé Dr. Kováts Judit

Tudományág: Anyagtudomány és technológiák

CÍM: A HETEDIK ÉS A NYOLCADIK KERETPROGRAMBAN MEGVALÓSULT FAIPARI PROJEKTEK STRUKTÚRÁJÁNAK, FINANSZÍROZÁSÁNAK, HÁLÓZATAINAK ÉS EREDMÉNYEINEK ELEMZÉSE

Készítette: Novotni Adrienn

Témavezető: Dr. habil Pásztory Zoltán

Sopron

2023

**A HETEDIK ÉS A NYOLCADIK KERETPROGRAMBAN MEGVALÓSULT FAIPARI PROJEKTEK
STRUKTÚRÁJÁNAK, FINANSZÍROZÁSÁNAK, HÁLÓZATAINAK ÉS EREDMÉNYEINEK ELEMZÉSE**

Értekezés doktori (PhD) fokozat elnyerése érdekében
a Soproni Egyetem Cziráki József Faanyagtudomány és Technológiák
Doktori Iskolája

Menedzsment a faiparban programja

Írta:
Novotni Adrienn

Témavezető: Dr. habil Pásztory Zoltán
Elfogadásra javaslom (igen / nem)

(aláírás)

A jelölt a doktori szigorlaton % -ot ért el,

Sopron,
a Szigorlati Bizottság elnöke

Az értekezést bírálóként elfogadásra javaslom (igen /nem)

Első bíráló (Dr.) igen /nem
(aláírás)

Második bíráló (Dr.) igen /nem
(aláírás)

(Esetleg harmadik bíráló (Dr.) igen /nem
(aláírás)

A jelölt az értekezés nyilvános vitáján.....% - ot ért el

Sopron,
.....
a Bírálóbizottság elnöke

A doktori (PhD) oklevél minősítése.....

.....
Az EDHT elnöke

Tartalom

1.	Kivonat	1
2.	Abstract.....	2
3.	Bevezetés	3
4.	Szakirodalmi összefoglaló	5
4.1.	A faipari kutatások jelentősége a fenntartható fejlődés szemszögéből	5
4.1.1.	A fenntartható faipar koncepciója	6
4.2.	Faipari projektek a szakirodalomban	8
4.3.	Az SNA-elemzés szakirodalmi háttere	10
4.4.	Az FP7 és a H2020 projektek SNA-módszerekkel történő elemzése a szakirodalomban.....	10
5.	Legfontosabb szakmai hipotézisek	12
5.1.	1. szakmai hipotézis.....	12
5.2.	2. szakmai hipotézis	12
5.3.	3. szakmai hipotézis.....	12
5.4.	4. szakmai hipotézis.....	12
5.5.	5. szakmai hipotézis.....	12
5.6.	6. szakmai hipotézis.....	12
5.7.	7. szakmai hipotézis.....	12
5.8.	8. szakmai hipotézis.....	12
5.9.	9. szakmai hipotézis.....	12
5.10.	10. szakmai hipotézis	12
5.11.	11. szakmai hipotézis	13
5.12.	12. szakmai hipotézis	13
5.13.	13. szakmai hipotézis	13
5.14.	14. szakmai hipotézis	13
5.15.	15. szakmai hipotézis	13
5.16.	16. szakmai hipotézis	13
5.17.	17. szakmai hipotézis	13
5.18.	18. szakmai hipotézis	13
5.19.	19. szakmai hipotézis	13
6.	Az empirikus kutatás módszerei	14
6.1.	A szakirodalom szövegbányászati és tudománymetriai elemzésének módszerei	14

6.2.	A fenntartható faipar súlya a tudományos kutatásban vizsgálatának módszerei	15
6.3.	Az FP7 és a H2020 faipari projektjeinek statisztikai elemzése	15
6.3.1.	Projektrészvételek	16
6.3.2.	Projekt támogatások	18
6.4.	Az SNA-elemzés módszerei	19
6.4.1.	A faipari projektháló elemzésének módszerei	19
6.4.1.1.	A kapcsolatháló felrajzolása	19
6.4.1.2.	Hálózati mutatók számítása	20
6.4.1.3.	Hálózati klaszterelemzés	28
6.5.	A faipari projektek outputjai	28
7.	Eredmények	30
7.1.	A szakirodalom szövegbányászati és tudományometriai elemzése	30
7.2.	A fenntartható faipar súlya a tudományos kutatásban – eredmények	37
7.2.1.	Az eredmények értékelése	41
7.3.	Az FP7 és a H2020 faipari projektjeinek statisztikai elemzésének eredménye	42
7.3.1.	A projektek általános jellemzői	42
7.3.2.	Az egyes országok programonkénti részvételének elemzése	51
7.3.3.	Az egyedi és országonkénti részvételek változásai az egyes programok között	57
7.3.4.	Tevékenység és programonkénti részvétel összehasonlítása	59
7.3.5.	A faipari projektrészvevők programonkénti támogatása	60
7.3.6.	Az egyes országok által kapott közösségi támogatás a két programban	66
7.3.7.	Tevékenység szerinti egyéni támogatás összehasonlítása az egyes programokban	74
7.3.8.	Országonkénti támogatás és az általános faipari teljesítmény kapcsolata	76
7.4.	A faipari projektháló elemzésének eredményei	83
7.4.1.	A kapcsolatháló felrajzolása	83
7.4.2.	A hálózati mutatók számított értékei	86
7.4.3.	„Nevezetes” hálózatelméleti eloszlások vizsgálata	89
7.4.3.1.	Poisson-eloszlás	89
7.4.3.2.	Hatványfüggvény-eloszlás	91
7.4.3.3.	Egyéb eloszlások	95
7.4.4.	Központi résztvevők	101
7.4.5.	Hálózati klaszterek	104
7.5.	Projekteredmények hatékonyság- és tartalomelemzése	107

7.5.1. Az európai uniós támogatási programok faipari projektjeinek projekteredményei	107
7.5.2. A faipari projektek legfontosabb kulcsszavai	111
7.5.3. Faipari szabadalmak elemzése	114
8. Összefoglalás.....	119
9. Felhasznált szakirodalom	121
10. Tézisek	135
10.1. Saját publikációk bibliográfiai adatai	137

1. Kivonat

A fenntartható faiparhoz, a faipari projektekhez, általában a kapcsolatháló elemzéshez és konkrétan az európai uniós projektek kapcsolatháló-elemzéséhez kapcsolódó szakirodalomról nyújtott kép alapján beigazolódott, hogy az általunk használt módszerek a Hetedik és Nyolcadik Keretprogram (a továbbiakban: FP7 és H2020) faipari projektjeinek elemzése során újszerűek. A faipari projektek szövegbányászati és tudományometriai elemzése alapján a folyóiratokra vonatkozó Bradford-eloszlás nem érvényes, a szerzői produktivitásra vonatkozó Lotka-eloszlás viszont igen. A publikációk dimenziócsökkentési eljárásokkal kulcsszavaik szerint 4-5 klaszterre oszthatók. A fenntartható faiparhoz kapcsolódó publikációk idősorainak elemzése alapján a faipar szerepének felismerése a fenntarthatóságban egyelőre nem egyértelmű. Az FP7-hez képest a H2020 faipari projektekben a támogatás kevésbé volt egyenletes. A koordinátorok átlagos támogatása mindkét támogatási programban magasabb, mint a résztvevőké. A két programban való országonkénti részvétel és támogatás között közepes erősségű korreláció és regresszió tapasztalható. A tevékenység alapján mért részvétel különbséget mutat a két program között. A tevékenység alapján a két program között a támogatás súlypontja látványosan átalakult, de ez nem az egyes csoportok támogatásának különbségére, hanem a belső szórásnégyzet alapján a csoportokon belüli különbségekre vezethető vissza. Az egy résztvevőre eső közösségi támogatás mediánja – különösen reálértékben – csökkent a két időszak között. A többes projektrészvétel magasabb támogatással jár. Az FP7 faipari projektjeiben résztvevő országok közösségi támogatása a faipari kibocsátással közepes rangkorrelációt és közepesnél gyengébb lineáris regressziót mutat. A faipari projektháló töredezettségét a tartalomvizsgálat segítségével részben faiparinak minősített projektek részvevőivel való kiegészítés szignifikánsan nem csökkentette. A vizsgált kapcsolatháló hatványfüggvény-eloszlást követtek egy bizonyos minimumérték felett, az értékeket a bootstrapping-eljárások érdekében nem módosították. Egyes jelek alapján a tényleges együttműködések a minimumérték alatt is hatványfüggvény-eloszlás felé mutathatnak. A faipari projektek foksám és közöttség alapján kimutatható központi aktorai a vizsgált csoportokban különböző mértékű átfedést mutattak. A támogatásokhoz képest a faipari projektek eredményeinek száma nem kiemagasló, tartalmi szempontból szórványos átfedést mutatnak a faipari szabadalmak absztraktjaival.

2. Abstract

The insight provided by the literature on the sustainable wood industry, on wood projects, on network analysis in general and specifically on network analysis of EU projects has confirmed that the methods used in our analysis of wood projects in Seventh and Eighth Framework Programme (hereafter: FP7 and H2020) are novel. Based on the text mining and science metrics analysis of the wood projects, the Bradford distribution for journals does not prove valid, but the Lotka distribution for author productivity does. Publications are divided into 4-5 clusters according to their keywords using dimension reduction procedures. The time series analysis of publications related to the sustainable wood industry suggests that the recognition of the role of the wood industry in sustainability is not yet evident. Compared to FP7, support for H2020 projects in the wood sector has been less steady. Average support for coordinators is higher than for participants in both funding programmes. There is a medium-strength correlation and regression between participation and support per country in the two programmes. Activity-based participation shows a difference between the two programmes. By activity, the centre of gravity of support appears to have shifted significantly between the two programmes, but this is not due to differences in support between groups, but to differences within groups based on the internal variance coefficient. Median support per participant, especially in real terms, decreased between the two periods. More project participation is associated with higher support. Community support to countries participating in FP7 wood projects shows a medium rank correlation with wood output and a weaker than medium linear regression. The fragmentation of wood project nets was not significantly reduced by the addition of participants in projects classified as partly related to the industry using the content analysis. The tested link nets followed a power function distribution above a certain minimum value, the values were not significantly modified by bootstrapping procedures. Some indications suggest that the actual interactions may also tend towards a power distribution below the minimum values. The central actors in wood industry projects, identified by degree number and betweenness, showed varying degrees of overlap in the groups studied. Compared to funding, the number of outputs of wood industry projects is not outstanding, with a minimal degree of overlap in terms of content with the abstracts of wood patents.

3. Bevezetés

Az értekezés keretében – a kapcsolódó szakirodalom elemzése és egyéb kapcsolódó vizsgálatok mellett – a Hetedik és a Nyolcadik Keretprogram faipari projektjeit és azok eredményeit (mint döntően, de nem kizárólag nyílt innovációs eredményeket) elemezzük szövegbányászati, statisztikai és hálózatkutatási eszközökkel, s emellett a faipari szabadalmakat (mint zárt innovációs eredményeket) is az elemzés körébe vonjuk. A két keretprogramra a továbbiakban FP7 és H2020 rövidítéssel hivatkozunk az egyszerűbb megkülönböztetés érdekében. Az FP7 a *Framework Program 7*, a H2020 pedig a *Horizon 2020*-ra utal. Bár kitérünk a kutatások költségeire és eredményeire, azaz a hagyományos értelemben vett hatékonyságelemzésre, a hatékonyság fogalmát jelentősen kiterjesztjük. Hatékonysági kérdésként fogjuk tehát fel azt is, hogy a faipari kutatási projektek résztvevői hatékony kutatási hálózatokba tömörülnek-e, illetve, hogy a hálózati szempontból jól működőnek tűnő részstruktúráknak melyek a jellemzői. Emellett szintén hatékonysági kérdésként megvizsgáltuk azt is, hogy mi lehet a hozzájuk való eredményes csatlakozási stratégia kulcsa, ha ez egyáltalán egyértelműen megállapítható. A megfelelő kutatási partnerek és együttműködési formák mellett a jól kiválasztott kutatás-fejlesztési irányok is fontosak lehetnek, ez felértékeli a „kurrens” kutatási témák, illetve a szabadalmak szövegbányászati eszközökkel végzett kutatását.

Ha a kutatás-fejlesztési folyamatokat tisztán az üzleti folyamatok részének tekintenénk, a fenti kérdések akkor is fontosak lennének. Azonban a kutatás-fejlesztés során nemcsak az outputból származó üzleti eredmény, hanem a döntően közösségi (pl. európai uniós) forrás elnyerése is fontos hatékonysági tényező, ezért a kutatás-fejlesztési folyamatok tisztán üzleti logika mentén már csak ezért sem szemlélhetők. Természetesen ettől még a kevés és üzleti szempontból kevésbé izgalmas, kizárólag a közösségi forrásokat felhasználó outputok esetén alapvető gazdaságossági kérdések merülhetnek fel. Úgy gondoljuk tehát, hogy az üzleti-gazdaságossági szemlélet fontos, de nem kizárólagos eleme a hatékonysági folyamatok vizsgálatának.

A dolgozat alapvetően a fő hipotézisek → pontosított statisztikai hipotézisek és számítások → statisztikai hipotézisek nem elutasítása/elutasítása és a számított eredmények értékelése → fő hipotézisek értékelése → összefoglaló tézisek utat követi.

A megfogalmazott legfontosabb hipotéziseket a későbbi fejezetekben szükség szerint statisztikai szempontból még pontosítottuk és kiegészítettük.

A dolgozat jóval több statisztikai hipotézist és számítást tartalmaz vagy feltételez, mint amennyit a fő hipotézisek között vázoltunk, mivel összetett számítási módszerek köztes statisztikai próbáinak statisztikai hipotéziseit gyakran a dolgozat egésze szempontjából másodlagosnak ítéltük meg.

A későbbi fejezetekben leírt statisztikai hipotézisek számát megpróbáltuk kordában tartani, de nem mehettünk el amellett a lassan félévszázados tudományos „trend” mellett (Hartley & Sielken, 1975), amely szinte az összes statisztikai elemzést a következtetési statisztikai vizsgálatok alá rendel, s még nyilvánvalóan deskriptív statisztikák esetén is hipotéziseket vizsgál, így lényegében szinte minden adatsort „mintának” tekint, illetve minden minta – a nem véletlen minta is – visszavezethető egy elméleti, végtelen szupersokaságra. Az ezzel kapcsolatos módszertani viták (Greenland, et al., 2016) túlmutatnak a dolgozat keretein. Annyit érdemes

megjegyezni, hogy a „szupersokaságok” létét általánosító, legradikálisabb megközelítés szerint minden véges számú sokaság egy elméletileg létező „szupersokaság” egyetlen, véletlen mintája, s ez a visszavezetés az iskola hívei szerint módszertani – pl. a robusztusságot biztosítandó – okokból szükséges. (Binder, 1983) A „szupersokaságok” feltételezését azonban explicit vagy implicit módon számos statisztikus elutasítja. (Bárdits, et al., 2016) A leggyakoribb, „köztes” szemlélet szerint „amennyiben egy teljes sokaság megfigyelésére kerül sor, a tesztek használata per definitionem értelmetlen, ám bizonyos esetekben a teljes sokaság egy nagyobb ún. szuperpopuláció részeként is felfogható, s akár mintaként is kezelhető.” (Hunyadi & Vita, 2016, p. 436)

A p érték általános félreértelmezése, a nagy mintáknál vitatható érvényességű próbák, a túl könnyen megugorható szignifikanciaszint, a próbák érvényességnek figyelmen kívül hagyása és még számos elméleti és gyakorlati probléma miatt a statisztikai próbák használatát az utóbbi években nagyon sok kritika éri. Az egész vita marginálisnak tűnhet, de mégis odavezetett, hogy miközben a tudományos folyóiratok többsége a következtetési statisztikai eljárások használatát várja el szinte minden statisztikai elemzésben, egyes folyóiratok egyenesen tiltják azok használatát. Ez utóbbiak közül elhíresült példa, hogy a *Basic and Applied Social Psychology* folyóirat 2015-ben a nullhipotézisre épülő szignifikanciateszteket – lényegében a statisztikai eljárások döntő többségét – érvénytelennek minősítette, használatukat nem fogadja el a kutatásokban.

Mi – mivel az ellentmondásos helyzetet megoldani nem tudjuk – a szakirodalomban leggyakrabban követett megközelítést és alkalmazott módszereket követtük, de emellett pl. statisztikai próbák p értékéből óvatosan vontunk le következtetéseket.

Szinte minden lehetséges esetben kiszámoltuk p (empirikus szignifikanciaszint, leegyszerűsítve: annak valószínűsége, hogy a próbastatisztika a mintából számított értéket követi) értékét, de szükség szerint jeleztük, hogy a populációs adatsor mintaként való kezelése és mögöttes sokaság (esetleg „szupersokaság”) feltételezése az olvasó számára mehökkentő lehet.

A sokaságnak tűnő adathalmazok azonban nemcsak a tudományos trendekhez való igazodás miatt kezelhetők mintaként. Pl. egy támogatási program bizonyos kategóriájához (esetünkben faipari kutatásokhoz) tartozó projektek tekinthetők a támogatási programok adott kategóriáján belüli szegmensének, hiszen a támogatási program nem fed le az összes formális vagy informális projektegység működését az adott kategórián belül. Ha pedig az elemzésben az időbeliség is szerepet játszik, a terjedőben lévő – a klasszikus statisztikai alapokat „kreatívan” használó – adattudományi megközelítés eleve feltételezi, hogy a múltbeli és jövőbeli értékek nélkül a rendelkezésre álló adatok nem tekinthetők teljes sokaságnak, s rutinszerűen használ mintáknál általános eljárásokat sokaságnak tűnő adathalmazoknál is. Persze a mintavétel jellege vagy a torzítottság kérdése ilyen esetekben is súlyos aggályokat vethet fel, de nem nagyobbakat, mint a tudományos kutatásokban általános, kérdőíves adatokra épülő elemzések, miközben a legtöbb kérdőíves felmérés nem teljesíti a tudományosan elfogadható mintavételek előírásait.

Még egyértelműbb a helyzet, amikor azt vizsgáljuk, hogy az adatok eloszlása egy nevezetes eloszlásból származik-e, hiszen ilyenkor az adatsor mintaként kezelése a nevezetes eloszlás felételezett „adatsorához” képest kevésbé kifogásolható.

A módszertani bizonytalanságokat – a statisztikai szoftverek és a statisztikai programozási nyelvek lehetőségeire építve – az adattudományban gyakran egy viszonylag egyszerű „trükkkel” oldják meg. A sokaságnak tűnő adatokból egyszeri vagy többszöri (bootstrapping) visszatevés nélküli vagy visszatevéses mintavételi eljárásokkal mintákat képeznek, s ezeket elemzik. Az eljárás segít annak a problémának a kezelésében is, hogy a nagy minták esetén a próbák szisztematikusan a nullhipotézis elutasítása felé mutatnak (Hunyadi & Vita, 2016). A dolgozatban erre is látunk példát.

A probléma kezelését elősegíthetik pl. az „effect size” módszerek, a nemparaméteres eljárások és a nagyobb minták használata, vagy a deskriptív statisztikák továbbfejlesztése. Ezek többnyire szintén túlmutatnak a dolgozat keretein, de lehetőség szerint próbáltunk meríteni a módszerekből.

Az elemzés során Excel, SPSS és JASP szoftvereket és az R statisztikai programozási nyelv egyes csomagjait használtuk.

4. Szakirodalmi összefoglaló

A szakirodalmi összefoglalóban az empirikus kutatás módszertanához kapcsolódó szakirodalmat nem érintettük, ezeket a műveket az empirikus kutatás módszereinél tüntettük fel.

4.1. A faipari kutatások jelentősége a fenntartható fejlődés szemszögéből

A faipari kutatások egyik legfontosabb jellemzője a fenntartható gazdálkodás felé való elmozdulás igényének egyre hangsúlyosabb megjelenése az elemzésekben. Az elmozdulás a faipari kutatási hálózatok kialakításának egyik fő mozgatórugójának tűnik. A *Sustainable Wood for a Sustainable World* (SW4SW, Fenntartható Fa egy Fenntartható Világért) kezdeményezést a FAO 2018-ban fogadta el, megvalósításában számos nemzetközi szervezet (ACFSI, CIFOR, ITTO, WWF) részt vesz. (FAO, 2018b)

A faipar és a tőle nehezen elválasztható egyes erdészeti tevékenységek fenntarthatóságának koncepciója fontos tényező az Agenda 2030-ban, az ENSZ integrált fenntartható fejlődési és fejlesztési keretrendszerében is (United Nations, 2019), de a 2015-ös Párizsi Egyezményben is megjelent. (United Nations, 2015) A faipari termékek teljes életciklusának fenntarthatósága viszonylag új elem a környezetvédelemben, a korábbi megközelítések az erdők megőrzését alapvetően az ipari tevékenység mellőzésétől várták. Egyre világosabbá vált azonban, hogy az erdők csak akkor őrizhetők meg, ha elfogadjuk, hogy az erdő számos egyéb funkciója mellett a faipar nyersanyagforrása. Különösen a fejlődő országokban elképzelhetetlen az erdőterületek megőrzése a gazdasági realitások figyelembevétele nélkül. (Nambiar, 2019)

Az SW4SW szembenéz egy másik alapproblémával is. A faipar évtizedek óta nagyon súlyos támadásoknak van kitéve. Ez részben a nem, vagy korlátozottan fenntartható technológiáknak köszönhető. Azonban számos fából készült vagy faalapú termék helyettesítő termékei jóval

nagyobb környezeti károkat okoznak, különösen, ha a teljes termékéletről vizsgáljuk. (Klarić & Obučina, 2020)

A faipar elleni kritikák egyes termékek keresleti oldalán visszaesést okoztak a fejlett országokban, s a folyamat korántsem zárult le. (Paluš & Slašťanová, 2019) Ez részben fenntartható gazdálkodásra való törekvést eredményezett (Casey, et al., 2017), részben még fenntarthatatlannabb gazdálkodáshoz vezetett.

A braziliai erdőtüzek felhívták a figyelmet arra, hogy amennyiben az erdei nyersanyagok (főleg a fa) és termékek nem elég jövedelmezőek, vagy felhasználásuk esetleg adminisztratív akadályokba ütközik, akkor az erdő helyén gyakran szarvasmarha-legelőket vagy szőjababföldeket alakítanak ki. (DeFries, et al., 2013) Ugyanilyen súlyos problémát jelentenek az olajpálmaligetek Délkelet-Ázsiában az egykori „természetes” erdők helyén. Eközben a fa mint energiahordozó és döntően alacsony hozzáadott értékű feldolgozásra váró alapanyag iránti igény a fejlődő és felzárkózó országokban szintén hatalmas erdőpusztításokhoz vezet, amelyeket a legkedvezőbb esetben is csak monokultúras ültetvényekkel pótolnak. (Henders, et al., 2015)

A magas szinten feldolgozott faipari termékek környezetvédelmi szempontból reális szemlélése, a jobb marketing, a teljes termékéletről fenntarthatóbbá tétele, s az „erdő” és a „fa” jövedelmezőségének javítása a helyi közösségek szintjén kulcsfontosságú tényező lehet az erdők megőrzése és a fenntarthatósági célok teljesítése érdekében. A fa növekvő high-tech felhasználása a kedvező folyamatokat erősítheti. (FAO, 2018b)

4.1.1. A fenntartható faipar koncepciója

A különböző erdőtípusok közül környezetvédelmi szempontból az örökzöld lombos erdők (EBF: evergreen broad-leaved forest, örökzöld lomboserdő) megőrzése kulcsfontosságú, mivel karbonmegkötési képességük kiemelkedően magas ($0,388 \text{ kgC m}^{-2}\cdot\text{yr}^{-1}$), egységnyi területre számítva minden más erdőtípus karbonmegőrzési képességének a többszöröse (3,5-9-szerese). (Zhao, et al., 2019) A többi erdőtípus – ECF (evergreen coniferous forest – örökzöld tűlevelű erdő), DBF (broad-leaved forest, lomboserdő), DCF (deciduous coniferous forest – lombhullató tűlevelű erdő) és ECDBMF (evergreen coniferous deciduous broad-leaved mixed forest – örökzöld-lombhullató vegyes erdő) – megőrzése is fontos, hiszen ezek karbonmegkötési képessége is jóval magasabb, mint az adott éghajlaton helyükre telepíthető különböző fásszárú ültetvényeké, szántóföldeké és legelőké. A különböző természetes erdőtípusok egymás helyén nagyon korlátozottan alakulhatnak ki vagy telepíthetők, így a kevésbé kedvező karbonmegkötési képességű erdőtípusok „leváltása” korlátozott hatékonyságú. A szén-dioxid kibocsátás és a szén-dioxid megkötés mérlege a természetes erdőtakarók esetén szintén jóval kedvezőbb. (Pang, et al., 2019)

A természetes, örökzöld lombos erdők azonban súlyos veszélyeknek vannak kitéve, s megőrzésük elképzelhetetlen fenntartható fakitermelés és faipar nélkül.

A Sustainable Wood for a Sustainable World a *Fenntartható Fejlődési Célok* (SDGs, Sustainable Development Goals) közül 11-nek a megvalósításhoz járulhat hozzá az alábbi négy területen megvalósított lépések segítségével (FAO, 2018b):

1. A fenntartható faipari értékláncok és termékek előnyeit nemzetközi szinten kell elismerni, s be kell építeni a fenntarthatósági és klímavédelmi stratégiák, ill. a nemzeti politikai keretek közé.
2. Növelni kell az erdőgazdálkodás kapacitásait a fenntartható fakitermelés érdekében, a fenntartható fakitermelést pedig a fenntartható tájgazdálkodás különböző szintjeinek megfelelően kell szemlélni.
3. Szélesíteni kell a fenntartható faipari értékláncok és termékek piaci és finanszírozási lehetőségeit.
4. A fenntartható faipari értékláncok és a fenntartható faipari termékek biogazdálkodáshoz és újrahaznosításhoz való hozzájárulását mind a fejlett, mind a fejlődő országokban erősíteni kell.

A fenti célok természetesen támogatandók. Azonban már az alapvető célokat megfogalmazó 2017-es, római SW4SW-világkonferencián kirajzolódtak a megvalósítás előtt álló fő akadályok.

A faipar fenntarthatósága ugyanis igen összetett műszaki, gazdasági és politikai problémák megoldását igényli. A megoldás módszerei alapvetően ismertek a szakemberek előtt, s egyes országok Kameruntól Kínáig igen komoly eredményeket értek el az utóbbi években. Néhány probléma azonban jelenleg globális szinten megoldhatatlannak tűnik:

- A fejlődő országok a fejlett országokéhoz képest szerény, de folyamatosan növekvő fogyasztói igényeit a fejlett országok közvéleménye és politikai vezetői nem fogadják el, s emellett rendkívül szigorú normákat fogalmaznak meg számukra. (Sommer, et al., 2017)
- A SW4SW a probléma súlyához képest igen alacsony költségvetéssel tervez (19 millió USD), de egyelőre még ennek a forrásoldalát is csak részben sikerült biztosítani, főként az ENSZ utóbbi időben tapasztalt alulfinanszírozottsága miatt. (FAO, 2018a) Az elmúlt években a környezetvédelmi támogatások helyett egyes nyugati kormányok a szigorú környezetvédelmi szabályozások politikáját követik, miközben a problémák megoldása hatalmas anyagi forrásokat kívánna. (Adhikari, 2019)
- A faipar és az erdőgazdálkodás a világ GDP-jének alig 1%-át adja, főleg a fejlett országok kormányai számára marginális területnek számít. (FAO, 2018b)
- Az élelmiszer, az ipari és az energianövények, ill. az ipari területek iránti igény folyamatosan nő, miközben az alacsony feldolgozottságú fatermékek (pl. cellulóz) iránti igény 2030-ra megnégyszereződhet. Mindez nem kedvez a fenntartható gazdálkodásnak, különösen az alapanyag-kitermelésben. (FAO, 2018b)
- A faalapú high-tech termékek előállítására főleg a fejlett országokban és Kínában történik, így a magas hozzáadott értékből fakadó többletjövedelmek fejlesztési forrásként elsősorban nem a fejlődő országokban jelennek meg. (Klarić & Obučina, 2020)

A problémák rendkívül gazdag kutatási területet jelentenek, s a FAO is külön kiemelte a tudományos közvéleményt, mint az SW4SW egyik célcsoportját.

Érdekes kérdésnek tűnik tehát annak megvizsgálása, hogy növekedés tapasztalható-e a fenntartható faipari tudományos kutatások (publikációk) területén a környezetvédelmi kutatásokhoz és azon belül a fenntartható nyersanyagokhoz képest.

A dolgozat egészének szempontjából a kérdés azért fontos, mert a későbbiekben vizsgált faipari projektek lényegében mindegyike kiemelte a fenntarthatóságot, s úgy tűnik, hogy a fenntarthatóság hangsúlyozása – és egyben biztosítása – nélkül a faipari projektek nem számíthatnak európai uniós forrásokra. A forrásbevonás során tehát a fenntartható jelleg első számú sikertényező.

4.2. Faipari projektek a szakirodalomban

A faipari projektekkel foglalkozó cikkek tartalomelemzése a dolgozat szempontjából kiemelt fontosságú. A cikkek tartalmi elemzését visszamenőleg 2016-ig végeztük el. A régebbi cikkek aktualitása kérdéses.

A faipari K+F projektek keretében készült vagy közvetlenül a projekteket leíró cikkek jelentős része a dolgozatban korábban érintett fenntartható fejlődés témakörébe esik. Ez abból a szempontból is érdekes, hogy a keresőszavak között a fenntarthatóság egyáltalán nem jelent meg.

Fontos témaként tűnik fel a fenntartható fa mint alapanyag és mint építőanyag felhasználásának növelése: a fémből és betonból készült, építőipari szerkezetek faelemekkel való kiváltása (Riggio, et al., 2022) (Kitriniaris, 2018), az illegális fakitermelésből származó fa helyettesítése fenntartható gazdálkodásból származó faanyaggal a fejlődő világban (Wongnaa, et al., 2022), a hagyományos üzemanyagok faalapú biodízzel való felváltása (Boglioli, et al., 2022). A fenntarthatóság technológiai szempontból egyik fontos eleme a fafeldolgozás során keletkező hulladék arányának csökkentése. (Larchenko, et al., 2022) A fenntarthatóság általános, holisztikus megközelítése faipari példák alapján némileg szokatlannak tűnik a szűkebb szegmenseket érintő faipari elemzések között, de egyáltalán nem példa nélküli. (Gibson & Warren, 2020) (Santana-Sosa & Fadai, 2019) (Chambers, 2019) (Marsh & Nelson, 2018) (Schramm & Litschel, 2017) (Lallemand & Guérin, 2017) (Davies, et al., 2017)

A fenntarthatóság és a hatékonyság együttes biztosítása szintén fontos cél. (Martins, et al., 2022). A fejlődő országokban a túlzott, pazarló kitermelés jelent gondot, s a gazdaságossági és környezetvédelmi szempontok összehangolása tűnik égetően fontos kérdésnek. (Mbatu, 2020) (Prescott, et al., 2017) (Salame, et al., 2016) Meglepően sok szerző elemez oroszországi projekteket. A világ legnagyobb erdőállományával rendelkező Oroszországban a ritkán lakott területek munkaerőhiánya miatt jelentős technikai (automatizálási) fejlesztések zajlanak, a pazarlás csökkentése inkább gazdaságossági szempontok miatt kerül előtérbe. (Gordeev, et al., 2021) (Butko, et al., 2021) A közhiedelemmel ellentétben viszonylag szigorú és folyamatosan szigorodó orosz törvények (Kolesnik & Sinyatullina, 2017) és szabályozók (Ivantsova, 2021) ellenére fenntarthatósági szempontból (is) kiszámíthatatlan, részben közvetett globális hatások valószínűsíthetők, ha a technológiai és szervezeti fejlesztések (Ivanova, et al., 2018), illetve a termelés regionális összehangolása és átstrukturálása (Ulyanova, et al., 2019) (Titova, et al., 2019) (Perfiliev & Zadrauskaite, 2019) miatt orosz feldolgozott faipari termékek tömege árasztja el az ázsiai és részben az európai piacokat (Burov, et al., 2021).

Természetesen a faipar jövedelmezőségének javítása – mint minden gazdasági tevékenységé – a fenntarthatósági szempontok nélkül is hangsúlyos a faipari projekteket érintő publikációkban. A magasabb hozzáadott érték biztosítása a jövedelemőség javításának kulcsa (Dieste, et al., 2019), amely számos részterületen keresztül biztosítható. A kereskedelmi tevékenység modernizálása (Pinto, et al., 2022) (Barriault, et al., 2017), a termőhelyek alapanyagtermelési szempontokat figyelembe vevő kialakítása (Vaca, et al., 2022), a regionális faipari klaszterek kialakítása (Polyanin, et al., 2020), az energiafelhasználás optimalizálása (Dobretsov, et al., 2020), az innovatív projektek aktívabb finanszírozása (Golovina & Dykusova, 2020), a faipari üzemszervezés és logisztika modernizációja (Wieruszewski, et al., 2020) (Trojanowska, et al., 2016) (Koppelhuber, et al., 2016) egyaránt hatékonyságjavulást ígér. Versenyképességi és technológiai problémák egyébként nemcsak a második és a harmadik (Vrabcová, et al., 2019) világban lehetnek jelen, tapasztalhatunk ilyet a fejlett országokban (Lüthi, et al., 2019) (Kay, 2017) (Derikvand, et al., 2017) (Coombs & Forster, 2017) (Adamowicz & Noga, 2017) is.

A faipar szigorúan vett – ha egyáltalán ilyen lehatárolás lehetséges – technológiai kérdései (törvényszerűen) lényegében minden projektben központi szerepet játszottak. Az építőipar technológiai kérdései különösen hangsúlyosak. A hosszú fesztávú fafödémek felé támasztott elvárások (Nesheim, et al., 2021), a rétegelt lemezek jellemzői (Parsons, et al., 2021), a faanyag tárolási körülményeinek hatása a felhasználhatóságra (Vieira, et al., 2021), a szabálytalan alakú fűrészáru feldolgozhatósága (Aagaard & Larsen, 2020), a hőszigetelő, faalapú panelek jellemzői (Grohmann, et al., 2020), az axiálisan terhelt önmetsző csavarok felhasználhatósága (Brandner, et al., 2019), az alapanyag védelme a betegségek ellen (Gazis, et al., 2018) (Czaplicki, et al., 2016), a magas teherbírású fa-üveg kompozitok termelése (Buyuktaskin, et al., 2017), a fapadlók technológiai problémái (Kujawińska, et al., 2017), a faszervezeti módosítások (Németh, et al., 2016), a faanyag biomasszaként történő hasznosítása (Goryunov, et al., 2016) mind-mind fontos technológiai kérdések, de nyilván gazdaságossági és fenntarthatósági vonatkozásaik is vannak.

A faipari projektek nem szakmai, hanem projektmenedzsment-szempontok szerinti elemzése is megjelenik a cikkekben. Fiataloknak szánt, részben pedagógiai projektek elemzésére számos példát találtunk. (Kingdon, 2021) (Kekkonen & Isoherranen, 2021) (Isa, et al., 2020) A projektek faipari-szakmai szempontok szerinti, összehasonlító elemzésére kevesen vállalkoztak. (Ellefson, 2019) Általános elemzések (Shegelman & Vasilev, 2020) mellett egészen speciális elemzésekkel is találkozhatunk. Például a göccsel foglalkozó szakirodalom átfogó elemzése (Qu, et al., 2020) mindenképpen ide sorolható.

Az új, információs-kommunikációs és üzleti technológiákban rejlő lehetőségek elemzése (Henriksson & Johansen, 2018) (Salim & Johansson, 2018) (Aberger, et al., 2018) (Santana-Sosa & Riola-Parada, 2018) (Hofmair, et al., 2017) (Le Roux, et al., 2016) (Kropivšek, et al., 2016) mellett az új technológiák hátulütőinek bemutatására is akad példa (Özgün, et al., 2021).

Mindenképpen érdekes, hogy az „open innovation” és a „closed innovation” kulcsszavak nem jelentek meg a cikkekben.

A szakirodalmi eredmények fontossága miatt az általános bemutatást szövegbányászati elemzéssel is kiegészítettük a későbbi fejezetekben.

4.3. Az SNA-elemzés szakirodalmi háttere

Az értekezésben közölt hálózatkutatási elemzés az SNA-ra (*Social Network Analysis – társadalmi kapcsolatháló elemzése*), a hálózatkutatás egy részterületére ad példát.¹

A társadalmi kapcsolatháló elemzése számos előzmény után a '90-es évek második felétől került a tudományos érdeklődés középpontjába, Magyarországon főleg Barabási Albert-László munkásságának köszönhetően. (Barabási, 2003) Barabási hozzájárulása a tudományterülethez – különösen a tudományszervezés és a már meglévő elméletek informatikai megvalósítása terén – világszerte is jelentős. (Barabási, 2018) Azonban a hálózatkutatás már évtizedekkel korábban is fontos tudományterületnek számított, kb. az '50-es évek második felétől számítjuk a „nagykorúvá válását”. A dolgozatban alkalmazott hálózatkutatási módszerek az elmúlt évtizedek eredményeire épülnek, azaz a véletlen hálózatokat (gráfokat) leíró modellre (Erdős & Rényi, 1960); a rögzített fokszámeloszlással rendelkező, de amúgy teljesen véletlenszerű kapcsolatokból felépülő hálózatok modellezésére alkalmas konfigurációs modellre (Bollobás, 1980) (Molloy & Reed, 1995) (Newman, 2010); a „bárkitől legfeljebb csak hatlépésnyire vagyunk?” kérdés köré szerveződő kisvilág modellre (Watts & Strogatz, 1998); a skálafüggetlen hálózatok modelljére (Barabási & Albert, 1999); a skálafüggetlen hálózatok kialakulását leíró Barabási–Albert-modellre (Albert & Barabási, 2002).

4.4. Az FP7 és a H2020 projektek SNA-módszerekkel történő elemzése a szakirodalomban

A kutatók SNA-módszereket egyaránt felhasználtak a projekteredmények, illetve a projektek különböző jellemzőinek vizsgálata során.

Az SNA-módszerek egyik fő alkalmazási területe a projekteken belül létrejövő innovációk hatékonyságának elemzése. Az érintett témák rendkívül szerteágazók. Az energiarendszer innovációi (Calvo-Gallardo, et al., 2022) (Klitkou, et al., 2022) (Balest, et al., 2022) (Klöckner, 2019), a mezőgazdasági-rurális innovációk (Guerrero-Ocampo, et al., 2022) (Fernandez de Arroyabe, et al., 2021) (Lombardi, et al., 2020) (Yang, et al., 2020) (Favilli, et al., 2015), az információs-kommunikációs innovációk (Vicente, et al., 2020), a pénzügyi innovációk (Barucca, et al., 2020), a rendészeti innovációk (Burcher, 2020), a technológiai diverzitás ösztönzése (Muscio, et al., 2022), a termelési eljárások modernizációja (Yang, et al., 2019) (Capone, 2014) (Lemonnier & Millet, 2012), a településvédelmi innovációk (Gourbesville, 2012) vagy a magán-gazdaság és az egyetemek együttműködéséből fakadó innovációk vizsgálata (Pinheiro, et al., 2015) egyaránt fontos témának tűnik.

Az SNA élettudományi-egészségügyi innováció-vizsgálatokban való alkalmazása különösen jelentős. (Esteve-Codina, et al., 2021) (Freeze, et al., 2020) (Vimalajeewa, et al., 2019) (Choudhury, et al., 2017) (Millard, et al., 2015) (Pasquale, et al., 2013) A COVID-19 járvány felértékelte az ilyen jellegű kutatásokat. Bár a projektek kezdése, ill. gyakran érdemi befejezése megelőzte a járvány kitörését, találtunk néhány publikációt a témában. (Hung, et al., 2020) (Ahmed, et al., 2020) (Pascual-Ferrá, et al., 2020) Ez arra utalhat, hogy az európai

¹ A *social network*-öt magyarra többnyire „közösségi hálóként” fordítják, de a „társadalmi kapcsolatháló” pontosabb kifejezés.

finanszírozású programok „menet közben” is mutatnak némi rugalmasságot a gazdasági-társadalmi környezet változásaira. Az innovációk vizsgálatán belül az SNA és a szövegbányászat (Ludovico, et al., 2020) (Cha & Cho, 2012), a gépi tanulás (Vu & Parker, 2015) (Jankowski, et al., 2015) (Tosatto & van Zee, 2014), a hálózati tanulás (Filvà, et al., 2014) (Lim, 2009) vagy a játékelmélet (Yadati & Narayanam, 2011) összekapcsolása módszertani szempontból tűnik érdekesnek.

Hálózati jellegéből fakadóan a „közösségi média” a támogatási programok SNA-módszerekkel folytatott kutatásainak kedvelt terepe. A tweetek/posztok/üzenetek jellege és terjedése (Karamouzas, et al., 2022) (Ahmed, et al., 2020) (Scharl, et al., 2013) (Zielinski, et al., 2013) (Gau, et al., 2011), az átalakuló médiamanipuláció (Guarino, et al., 2020), az új médiafelületek biztonsági kérdései (Naqvi, et al., 2019) (Ciapetti, et al., 2019) (Soliman, et al., 2016) (Aliprandi, et al., 2014), a kurrens módszertanok „tesztelése” (Hao, et al., 2017) (Fischer-Hübner, 2012) a kutatások legfontosabb témái. A különböző „obskúrus” együttműködések vizsgálata is a témához sorolható. (Slobozhan, et al., 2022)

A támogatási projektek menedzsmentkérdéseinek SNA-eszközökkel történő vizsgálata szintén kedvelt kutatási terület (Akimova, 2021) (Marinova, et al., 2019) (Nunes & Abreu, 2020) (Wang, et al., 2020) (Zaabi & Bashir, 2020) (Herrera, et al., 2020) (Wanzenböck, et al., 2020) (Herrera, et al., 2020) (Pollack & Matouš, 2019) (Marinova, et al., 2019) (Gu, et al., 2019) (Lee, et al., 2018), többnyire kevés módszertani újdonságot vagy érdekes eredményt felmutatva. Kivételt képez talán a társadalmi tőke szerepének vizsgálata. (Pinheiro, et al., 2016) A területen belül a nyitott innováció kérdésének vizsgálata alárendelt. (Nunes & Abreu, 2020) Többben vizsgálták az európai programok különböző iparági vagy regionális szegmenseit mint hálózatot (Balland, et al., 2019) (Enger, 2018) (Schröpfer, et al., 2017) (Pirker, et al., 2017) (Duarte, 2013) (Derntl & Klamma, 2012), bár gyakran – a nyilvánvaló technikai akadályok miatt – inkább csak egy-egy projektre koncentrálnak (Hall, et al., 2018) (Pasquale, et al., 2013) (Peppoloni, et al., 2016) (Hampson, et al., 2012). Szakmai szempontból a technikai akadályok megoldását célzó módszertani ajánlások emelkednek ki. (Aslam, et al., 2018)

A térbeli mozgás leíró – különböző szállítási és logisztika hálózatok, illetve a turizmus – vizsgálata a terület hálózati jellege miatt szintén kedvelt elemzési terület. (Park & Stamato, 2021) (Valeri & Baggio, 2020) (Acampa, et al., 2020) (Yi, et al., 2020) Az elektromos autók projektjeinek SNA-módszerekkel történő vizsgálata valószínűleg a jövőben lesz kedvelt kutatási terület. (Li & Lu, 2020)

A számunkra kiemelkedő fontosságú faipari területen kevés publikáció született, de találunk példát a faipari termékek kereskedelmi hálózatainak SNA-eszközökkel történő elemzésére. (Lovrić, et al., 2018)

Az eddigi eredmények alapján megállapítható, hogy a 4.2. fejezetben vizsgált faipari projektek és a projektek SNA-módszerekkel történő elemzése között elég kicsi az átfedés a szakirodalomban, ezért a faipari projektek SNA- és egyéb statisztikai módszerekkel történő vizsgálata új eredményeket ígér.

5. Legfontosabb szakmai hipotézisek

5.1. 1. szakmai hipotézis

A faipari projektekkel foglalkozó szakirodalom idézőinek száma hatványfüggvény-eloszlást követ egy bizonyos x_{min} érték után. Azaz a legtöbb cikk nagyon kevés, ellenben néhány cikk nagyon sok hivatkozást kap.

5.2. 2. szakmai hipotézis

A faipari projektekkel foglalkozó cikkek és idézők évenkénti számának alakulása között korrelációt feltételezünk.

5.3. 3. szakmai hipotézis

A Scopusból elérhető, faipari projekteket leíró publikációk nem tekinthetők a témában megszületett publikációk valódi reprezentációjának. Tehát a Bradford-tétel nem teljesül.

5.4. 4. szakmai hipotézis

A szerzői produktivitásra vonatkozó Lotka-eloszlás érvényes a faipari projektekkel foglalkozó publikációs mintára.

5.5. 5. szakmai hipotézis

A projektek támogatása és teljes költségvetése között szoros összefüggés mutatkozik.

5.6. 6. szakmai hipotézis

Az FP7-ben és a H2020-ban az országokénti részvétel gyakorisága között van összefüggés.

5.7. 7. szakmai hipotézis

Az FP7 és a H2020 faipari projektjeiben az egyedi projektrészvételek eloszlása a két programban és az országokénti projektrészvételek eloszlása a két programban ugyanolyan eloszlásból származik.

5.8. 8. szakmai hipotézis

Van összefüggés a tevékenység és az FP7 és a H2020 programokban való részvétel gyakorisága között.

5.9. 9. szakmai hipotézis

A faipari projektek közösségi támogatásának teljes sokasággal összevethető középértéke (átlag vagy medián) nem tér el a teljes sokaságétól.

5.10. 10. szakmai hipotézis

A faipari projektek közösségi támogatása a két támogatási időszak között lényegesen nem változott.

5.11. 11. szakmai hipotézis

Az FP7-ben és a H2020-ban az országokénti közösségi támogatás mértéke között van összefüggés.

5.12. 12. szakmai hipotézis

Az FP7 és a H2020 projekttámogatások országokénti eloszlása ugyanolyan eloszlásból származik.

5.13. 13. szakmai hipotézis

A tevékenység, illetve az FP7 és a H2020 programokban kapott támogatás közötti kapcsolat jelentős, azonban a kategóriák támogatása közötti különbség nem az ismerv szerinti heterogenitásra, hanem a kategóriakon belüli heterogenitásra vezethető vissza. A faipari projekt-résztevők két program közötti átlagos támogatásnövekedése alapvetően a részátlagok növekedésére vezethető vissza.

5.14. 14. szakmai hipotézis

Az FP7 és H2020 programokból kapott országokénti támogatás és a faipari kibocsátás között összefüggés van.

5.15. 15. szakmai hipotézis

A projektrésztevők kapcsolathálóiban mért fokszámok eloszlásai skálafüggetlenek, azaz kevés résztevőnek van sok és sok résztevőnek kevés kapcsolata.

5.16. 16. szakmai hipotézis

A faipari projekthálóok jelentős hálózati klaszterei hierarchikus felépítésűek.

5.17. 17. szakmai hipotézis

A faipari projekthálóokban nem érvényesül a kisvilág jelensége. Azaz a résztevők nincsenek egymáshoz elég „közel” a projektegyüttműködések hálózatában.

5.18. 18. szakmai hipotézis

Az FP7 és a H2020 támogatási program faipari projektjeinek eredményei a felhasznált közösségi erőforrásokhoz képest korlátozott nyitott és zárt innovációs eredményeket hoztak. A támogatás és az eredmények között számszerűleg korlátozott az összefüggés.

5.19. 19. szakmai hipotézis

Az FP7 és a H2020 támogatási program faipari projektjeinek absztraktjai és a nemzetközi szinten elérhető faipari szabadalmak absztraktjai között mérsékelt átfedés lehet.

A hipotézisek nem fedik le az összes vizsgálatunkat, de alapvetően meghatározták a kutatási irányokat.

6. Az empirikus kutatás módszerei

6.1. A szakirodalom szövegbányászati és tudományometriai elemzésének módszerei

A szövegbányászati elemzés során az R *base* alapsomagon kívül jelentős mértékben az R *bibliometrix* (Aria & Cuccurullo, 2017), *dplyr*, *tidyr* (Wickham & Francois, 2015), *knitr* (Xie, 2016) és *ggplot2* (Valero-Mora, 2010) csomagjait használtuk fel.

A Scopusból letöltött, faipari projekteket leíró BibTeX-fájlt bibliográfiai data frame-be töltöttük, majd deskriptív, bibliográfiai statisztikát készítettünk.

Megvizsgáltuk, hogy a publikációk idézőinek száma hatványfüggvény-eloszlást követ-e egy bizonyos értéktől. (1. szakmai hipotézis) Ehhez a 6.4. fejezet módszereit használtuk.

Korrelációs teszttel vizsgáltuk a 2. szakmai hipotézis teljesülését.

Ez a szakmai hipotézis lényegében ekvivalens azzal, hogy a

$$H_0: \rho = 0 \tag{1}$$

statisztikai hipotézis elvetését feltételezzük.

A 3. szakmai hipotézist a talán kevésbé gyakran használt Bradford-törvény segítségével teszteltük. A Bradford-törvény folyóiratokra és cikkekre értelmezett változata alapján (Naranan, 1970), ha csökkenő sorrendbe rendezzük a folyóiratokat aszerint, hogy hány cikk jelent meg bennük, majd három olyan harmadot alakítunk ki – szintén a csökkenő sorrendet követve –, ahol az egyes harmadokba eső cikkek száma azonos, akkor az egyes harmadokba eső folyóiratok számának aránya a felső harmadból kiindulva $1: n: n^2$, ahol n a Bradford-érték (vagy Bradford-állandó).

Az eredeti eljárás azonban nagy könyvtárak különböző szempontok szerint csoportosított állományainak hivatkozásainál volt érvényes, illetve érvényes a mai napig, megnyugtató tudományos magyarázat nélkül. Kisebb mintáknál elterjedt módszer a Bradford-törvényre épülő, az eredeti összefüggést nem sértő Leimkuhler-modell használata. (Wardikar, 2013)

Három zóna esetén, Bradford-eloszlást feltételezve, az egyszerűsített képlet (Egghe, 1986)

$$k = (1,781y_m)^{\frac{1}{3}} \tag{2}$$

, ahol k a Bradford-állandó, y_m a legtöbb cikkel rendelkező folyóirathoz tartozó cikkek száma.

A Bradford-állandó együtthatója pedig a

$$r_0 = \frac{T(k-1)}{k^3-1} \tag{3}$$

képlettel számolható ki, ahol T a folyóiratok száma.

Ha a Bradford-eloszlástól a tapasztalati eloszlás eltér, akkor élhetünk azzal a feltételezéssel, hogy a vizsgált minta valamilyen okból torzított. A Bradford-eloszlásnak széles körben használt, egzakt statisztikai-bebecslési eljárása nincs.

Megvizsgáltuk a 4. szakmai hipotézissel (Lotka-eloszlás) elvivalens statisztikai hipotézist, miszerint:

$$H_0: \text{az } x \text{ darab cikket írt szerzők száma } 1/x^n \text{-nel arányos, ahol } n(\beta) \approx 2. \quad (4)$$

A tapasztalati eloszlásra és az elméleti, Lotka-eloszlásra kétmintás Kolmogorov-Szmirnov tesztet végeztünk.

6.2. A fenntartható faipar súlya a tudományos kutatásban vizsgálatának módszerei

A fenntartható faipar súlyának időszaki változását a Scopus adatbázisának évenkénti találatértékeivel mértük fel.

A vizsgálat során a „sustainable wood” (fenntartható fa), az „environmental protection” (környezetvédelem) és a „sustainable material” (fenntartható nyersanyag) kulcsszavak évenkénti gyakorisága jelentette a számítások alapját (a továbbiakban rendre: SW, EP, SM).

Az adatok vizuális alakulásuk alapján polinomiális (harmadfokú) vagy exponenciális trenddel tűntek leírhatónak. Az összehasonlíthatóság kedvéért az 1975 és 2018 közötti időszakra írtuk fel ezeket. A trendegyenletek tapasztalati értékekre való illeszkedését az R^2 mutatóval (determinációs együtthatóval) vizsgáltuk.

Az outlierok kiszűrésénél ebben a fejezetben elfogadtuk a statisztikai programok javaslatait, külön számítást nem végeztünk rájuk.

Statisztikai hipotéziseket ebben a fejezetben nem állítottunk fel, mivel semmilyen trend meglétét vagy hiányát nem feltételeztük. Ennek ellenére csak az amúgy szignifikáns eredményeket közöltük.

6.3. Az FP7 és a H2020 faipari projektjeinek statisztikai elemzése

A vizsgálat során először a CORDIS-ból letölthető hatalmas adatállományt kezelhető, soktáblás relációs adatbázissá alakítottuk. Az adatbázisból lehetőség nyílik az összes FP7 és H2020 projekt adatainak vizsgálatára.

A faipari projektek kiszűrésére két módszert alkalmaztunk.

Mivel a vizsgálat kezdetekor a SciVoc-kódok – amelyek egyértelműen meghatározzák egy projekt tematikus besorolását – (és több más oszlop eredménye) még nem állt rendelkezésre, a faipari projektek előszűrését keresőszavak segítségével végeztük el, majd a kiszűrt állományt egyenként tartalomvizsgálatnak vetettük alá, s így az állomány tovább csökkent.

2022 nyarán azonban elérhetővé vált az FP7 és a H2020 program lényegében teljes adatbázisa, ahol a megfelelő SciVoc-kódok is megjelentek. Úgy döntöttünk, hogy az eredeti adatállományunkat – a továbbiakban H2020(fa2) néven összehasonlítási alapként – nem vetjük el teljesen, s ahol lehet, összevetjük az FP7 és a H2020 „woodworking” SciVoc-kóddal (is) rendelkező projektjeivel (ez utóbbiakat a továbbiakban FP7(fa) és H2020(fa) néven).

Ezután a teljesség kedvéért keresőszavak és tartalomvizsgálati eszközökkel az FP7(fa2) állományt is összeállítottuk.

Az FP7(fa)/H2020(fa) és az FP7(fa2)/H2020(fa2) adatállomány eltérései az általunk alkalmazott módszer tökéletlenségei mellett abból is fakadnak, hogy mi „minden”, a faiparhoz akár csak kis mértékben kapcsolódó projektet szerepeltetünk az FP7(fa2)/H2020(fa2) állományokban, míg az FP7(fa)/H2020(fa) SciVoc-kódjai erős, előzetesen definiált faipari jelleget tükröznek. Az FP7(fa)/H2020(fa) ezért jóval szűkebb adatállomány, mint az FP7(fa2)/H2020(fa2).

Az FP7(fa2)/H2020(fa2) előzőekben vázolt összevetését a másik két adatállománnyal azonban a „klasszikus” statisztikai számítások terén nem tartjuk helyes módszernek, viszont hálózat-elemzési szempontból igen, ezért FP7(fa2)/H2020(fa2) állományoknak a 6.4. fejezetben vázolt alapokon a 7.4 fejezetben lesz szerepe.

A statisztikai számítások során tehát csak az FP7(fa) és a H2020(fa) adatállománnyal dolgoztunk.

6.3.1. Projektrésztvételek

A 6. szakmai hipotézis ellentettjének statisztikai megfelelője (Spearman-féle rangkorreláció próbájával vizsgáljuk):

$$H_0: \rho = 0 \quad (5)$$

A Spearman-féle rangkorreláció elvégezhető diszkrét adatok esetén is. (Bár a nagyszámú diszkrét érték esetén felvetődik a „folytonosként kezelt diszkrét változó” lehetősége, ahogy azt 6.4.1.2 fejezetben elemeztük.) A próba elvégzése a bevezetőben jelzett módszertani problémákat veti fel, így a kapott ρ érték mellett p értékelhetőségét az olvasóra bízunk.

A 7. szakmai hipotézis statisztikai megfelelője (folytonos vagy folytonosként kezelt diszkrét változó esetén Kolmogorov-Szmirnov próbát használva):

$$H_0: F_X(x) = F_Y(y) \quad (6)$$

vagy:

$$H_0: F_X(x) = F_0(x) \cdot F_Y(y) = F_0(y) \cdot F_0(x) = F_0(y)$$

Diszkrét változó esetén – az értékeket azonos számú kategóriába rendezve – a 7. szakmai hipotézis a 8. szakmai hipotézis módszerét követve is vizsgálható, ha az előfeltételek teljesülnek. (Ezt teszteltük.)

A Pearson-féle χ^2 mellett/helyett (ez utóbbinál a próba előfeltételeinek esetleges nem teljesülése esetén) Kullback-Leibler és Jensen–Shannon divergenciát is mértünk.

Az országok összehasonlítása során csak azokat az országokat vettük figyelembe, amelyek mindkét programban részt vettek, s ez vélelmezhetően jelentős adatvesztéssel járt.

Megvizsgáltuk a Pearson-féle korrelációs együttható előfeltételeinek teljesülését is (ha a Shapiro-Wilk és/vagy a Kolmogorov-Szmirnov próba alapján a normalitás kizárható volt, a parametrikusság további feltételeit nem vizsgáltuk), illetve kiszűrtük az esetleges outliereket.

A normalitás „teljesülését” feltételeztük volna folytonosként kezelt diszkrét változó esetén, ha a próba ezt megerősítette volna.

A 8. szakmai hipotézis ellentettjének statisztikai megfelelője (Pearson-féle χ^2 próbával vizsgáltuk):

$$H_0: \{\pi_{ij}\} \quad (7)$$

A potenciális outlierok alsó és felső határát három módszerrel vizsgáltuk. A próba elvégzése a bevezetőben említett módszertani problémát veti fel, ezért az eredmény értelmezése mindenképpen óvatosságra int.

Az outlierok kiszűrésére a statisztikában és a statisztikai szoftverek által gyakran használt módszer az interkvartilis terjedelemre (IQR) és az alsó-felső kvartilisekre épül. (8) Az I tartományon kívül eső értékeket általában jó eséllyel outliernek tekinthetjük.

$$I = [q_{0,25} - 1,5 \times IQR; q_{0,75} + 1,5 \times IQR] \quad (8)$$

Alternatív módszer a Hampel-szűrő (Hampel, 1985):

$$I = [Me - 3 \times MAD; Me + 3 \times MAD] \quad (9)$$

(8) helyett széles körben használják a megengedőbb Tukey-határokat (Tukey, 1977):

$$I = [q_{0,25} - 2,2 \times IQR; q_{0,75} + 2,2 \times IQR] \quad (10)$$

A korreláció és regressziószámítás előtt megvizsgáltuk, hogy egyelőre linearitást feltételezve többdimenziós (esetünkben: kétdimenziós) outlierok vannak-e a vizsgált adatsorokban.

A Mahalanobis-távolság (pontosabban: távolságnégyzet) általánosan használt a többdimenziós outlierok kiszűrésére (Münnich, et al., 2006):

$$D_{ij}^2 = \sum_{r=1}^k \sum_{s=1}^k (\mu_{ri} - \mu_{rj}) v^{rs} (\mu_{si} - \mu_{sj}) \quad (11)$$

A képletben v^{rs} a k változóból álló kovarianciamátrix inverzében az r . sor és s . oszlop aktuális eleme. Ha a vizsgált populáció normális eloszlású, akkor a D^2 értékei egy p szabadságfokú χ^2 -eloszlást követnek. Szignifikánsan nagy D^2 érték – két változó mellett ($df = 2$) a küszöbérték 13,82 – esetén az értéket outliernek tekinthetjük. Vagy p értékkel kifejezve, ha p érték kisebb, mint 0,001, akkor outlierre gyanakodhatunk.

A többdimenziós outlierok kiszűrésénél érdemes figyelembe venni azt is, hogy az outlierok milyen hatással vannak a változók közötti összefüggésre, esetleg nem torzítják-e az adatsort. Ezt leggyakrabban Cook-távolsággal jellemzik. (A Cook-távolság sem tartozik a gyakran számolt statisztikai mutatók közé, így képletét közölnünk illene, de a számítás bonyolultsága miatt ezt mégsem tesszük. A számításra az R base `cooks.distance` függvényét használtuk a megfelelő lineáris modell definiálása után.)

A Cook-távolság 1-nél nagyobb érték esetén gyakorol jelentős hatást a korrelációra.

A pontfelhőre legkisebb négyzetek módszerével egy egyenest rajzoltunk fel, majd meghatároztuk a lineáris regresszió paramétereit.

Student-féle t -próbával és F -próbával megvizsgáltuk, hogy lehet-e szignifikáns lineáris kapcsolat a két változó között. (A statisztikai próbák nullhipotézise szerint nincs.)

$$H_0: \beta_1 = 0 \quad (12)$$

A próba elvégzése itt kapcsolódik a bevezetőben érintett módszertani problémához, így az eredmény óvatosan értékelendő.

A lineáris regressziós kapcsolat előfeltételei közül az egyik legáltalánosabban vizsgált a reziduálisok eloszlásának normalitása (mint a homoszkedaszticitás előfeltétele), amelyet legkönnyebben a lineáris modell felírása után tudunk ellenőrizni. A normalitás ellenőrzéséhez Shapiro-Wilk, Kolmogorov-Szmirnov, Cramér-von Mises és Anderson-Darling próbákat használtunk.

$$H_0: \text{A reziduálisok normális eloszlást követnek.} \quad (13)$$

A reziduálisok eloszlásának normalitása mellett a homoszkedaszticitás (a hibatagok szórásnégyzetének állandósága) szintén fontos támpontot ad a lineáris regresszió megbízhatóságáról, ezért mi is teszteljük studentizált Breusch–Pagan próbával és White-próbával:

$$H_0: \text{A homoszkedaszticitás érvényesül.} \quad (14)$$

A hipotézisek vizsgálata mellett természetesen deskriptív statisztikai adatokat is közöltünk.

6.3.2. Projekt támogatások

A két támogatási időszak közötti időszak vizsgálatát részben leíró jellegű összehasonlításra építettük. A 9. szakmai hipotézisnél először az egymintás parametrikus próba (a faipari projektek projektrésztvételekénti támogatási átlaga megegyezik a támogatási program összes projektrésztvételekénti tapasztalt átlaggal) előfeltételei közül a normalitást vizsgáltuk, majd annak esetleges nem teljesülése esetén nem parametrikus, kétmintás Wilcoxon-próbákat végeztünk a mediánra:

$$H_0: m = m_0 \quad (15)$$

A próba elvégzésének szükségessége itt is statisztikai megközelítés kérdése, ahogy azt a bevezetőben érintettük. Az eredményt ezért megfelelő körültekintéssel kell értékelnünk.

A (15) vizsgálatát FP7(fa)-FP7, H2020(fa)-H2020, FP7(fa)-H2020(fa), FP7-H2020 viszonylatokban vizsgáltuk nominális és reálértéken is. Nem minden viszonylat tekinthető minta-sokaság párosításnak, de megfelelő értelmezés mellett a próba véleményünk szerint elvégezhető.

A 10. szakmai hipotézist is (15) segítségével vizsgáljuk.

A 11. szakmai hipotézist – statisztikai szempontból a hipotézis az állítás ellentéte – (5), ill. (8)-(14) statisztikai hipotézisei és mutatói alapján vizsgáltuk.

A 12. szakmai hipotézist (6) alapján vizsgáljuk.

A 13. szakmai hipotézis vizsgálata során a vegyes kapcsolatból kiindulva főátlagokat, részátlagokat számoltunk. Ezután a teljes szórásnégyzetet összetevőkre bontottuk (külső és belső szórásnégyzet), majd szórásnégyzet-hányados típusú mutatószámokat számoltunk. A két

támogatási program közötti változás kimutatására a főátlagokat hasonlítottuk össze standardizálással.

Az 5. szakmai hipotézist és a 14. szakmai hipotézist a 11. szakmai hipotézishez hasonlóan teszteltük. Az 5. szakmai hipotézis esetén azonban a feltételezés meglehetősen trivialisága miatt csak korrelációt számoltunk.

6.4. Az SNA-elemzés módszerei

A hálózatelemzésnek viszonylag jól leírható módszertana alakult ki, amelynek egyes módszerei kapcsolódnak a különböző elméleti irányzatokhoz, de elméleti szempontból inkább egy sajátos logikájú „patternnek” tekinthetők.

Az elemzés során az 6.3. fejezetben vázolt módon FP7(fa) és H2020(fa) adatállomány mellett az FP7(fa2) és a H2020(fa2) adatait is felhasználtuk. Az R alapcsomagján kívül elsősorban az *RSQLite*, *dplyr*, *igraph*, *vcd*, *powerLaw* és *fitdistrplus* csomagokra építettünk.

6.4.1. A faipari projektháló elemzésének módszerei

6.4.1.1. A kapcsolatháló felrajzolása

A faipari projektek elemzése során tehát négy adatsort hasonlítottunk össze.

A faipari projekthálókat (a projektrésztvevők hálóját) irányítatlan hálózatként kezeltük, azaz nem vettük figyelembe a csomópontok közötti kapcsolat irányát. A tudományos együttműködések ugyanis az eltérő státusztól függetlenül irányítatlan kapcsolatként szokás értelmezni. (Barabási, 2017) Tehát nem vettük figyelembe, hogy az adott résztvevő pl. egyben projektkoordinátor-e.

Különböző, statisztikai programozási átalakítások után felírtuk a *szomszédsági* (vagy *adjacencia-*) *mátrixot*. A szomszédsági mátrix a hálózat kutatásban fontos elem. Az éleket nem tekintettük irányítottak, most csak a kapcsolatrendszert – a hierarchiát nem – vizsgáltuk.

Egy N pontból álló irányítatlan hálózat szomszédsági mátrixának N sora és N oszlopa van. A mátrix csak 0 és 1 értékeket vehet fel.

Ha i és j között van kapcsolat, akkor:

$$A_{ij} = A_{ji} = 1 \quad (16)$$

Ha nincs, akkor:

$$A_{ij} = A_{ji} = 0 \quad (17)$$

Egyszerű példa irányítatlan szomszédsági mátrixra:

$$A_{ij} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad (18)$$

A szomszédsági mátrixból ezután különböző, csak technikai szempontból érdekes lépések során felrajzolhatjuk a hálót, amely esetünkben lényegében a szomszédsági mátrix grafikus reprezentációja.

6.4.1.2. Hálózati mutatók számítása

A rendelkezésre álló adatok alapján kiszámítottuk az elemzés szempontjából releváns hálózati mutatókat.

Először kiszámoltuk a *sűrűséget*, tehát azt, hogy az összes lehetséges kapcsolatnak milyen hányada jött létre.

Egy N elemű hálóban a felírható kapcsolatok száma:

$$PE = \frac{N * (N - 1)}{2} \quad (19)$$

A sűrűség irányítatlan hálóban az alábbi módon írható fel:

$$D = \frac{2E}{N(N - 1)} \quad (20)$$

ahol E a tapasztalati élek száma.

Ha minden lehetséges kapcsolat létezik, tehát mindenki kapcsolatban áll mindenkivel, akkor a sűrűség értéke 1. 0 sűrűségérték esetén senki sem áll kapcsolatban senkivel. A sűrűség értéke tehát 0 és 1 közötti szám, amelynek magasabb értékei nagyobb hálózati sűrűséget jeleznek. (Molnár, 2020) Értékelése véleményünk szerint csak nagyon hasonló méretű hálózatok összehasonlításánál egyértelmű. FP7(fa) és H2020(fa) esetén talán ezzel a helyzettel állunk szemben.

A *tranzitivitás* azt mutatja meg, hogy milyen átlagos esélye van annak, hogy ha egy csomópont kapcsolatban van egy másik csomóponttal, s ez a csomópont egy harmadikkal, akkor a kiinduló csomópontunk is kapcsolatban van a harmadik csomóponttal. (Kisfalusi, 2018)

A tranzitivitást *átlagos klaszterezettségi együtthatónak* (Barabási, 2017) is nevezik, amely a klaszterezettségi együtthatóból – egy adott csomópont tranzitivitásából – vezethető le.

A k_i fokszámú i -edik csomópont klaszterezettségi együtthatója:

$$C_i = \frac{2L_i}{k_i(k_i - 1)} \quad (21)$$

ahol L_i az i -edik pont k_i számú szomszédja közötti kapcsolatok száma. Értéke mindig 0 és 1 közé esik.

A teljes hálózatra vonatkozó átlagos klaszterezettségi együttható:

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^N C_i \quad (22)$$

Fontos mutató a csomópontok fokszáma, tehát a csomópont és a hálózat többi csúcsa közötti kapcsolatok száma.

Írányítatlan hálózatban a csomópontok (N) és a fokszámok (k) segítségével kiszámítható az összes kapcsolat száma:

$$L = \frac{1}{2} \sum_{i=1}^N k_i \quad (23)$$

A hálózat fontos jellemzője az átlagos fokszám, ennek képlete irányítatlan hálózatban:

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i = \frac{2L}{N} \quad (24)$$

A hálózatoknak szintén alapvető jellemzője, hogy az adott hálózat skálafüggetlennek minősíthető-e.

Egy hálózatot skálafüggetlennek nevezünk, ha fokszámeloszlása hatványfüggvénnyel írható le. (Barabási, 2017)

A fokszámeloszlás hatványfüggvény-eloszlása:

$$p_k \sim k^{-\gamma} \quad (25)$$

ahol γ a fokszámkitevő. Mivel az eloszlások próbáinál γ helyett gyakran α is használatos, ahol a szakirodalomban α -ra hivatkoznak, γ mellett α -t is használtunk. (Gillespie, 2020)

Skálafüggetlen hálózatoknál a „legnagyobb” középpont fokszáma:

$$k_{max} = k_{min} + N^{\frac{1}{\gamma-1}} \quad (26)$$

A skálafüggetlenség lényegét a fokszámeloszlás momentumai fejezik ki.

A fokszámeloszlás n -edik momentuma (nemcsak skálafüggetlen hálózatoknál):

$$\langle k^n \rangle = \sum_{k_{min}}^{\infty} k^n p_k \approx \int_{k_{min}}^{\infty} k^n p(k) dk \quad (27)$$

Az alacsonyabb momentumok jelentése:

- $n = 1$: első momentum, átlagos fokszám, $\langle k \rangle$
- $n = 2$: második momentum, $\langle k^2 \rangle$ a szórásnégyzet (a fokszámok szórásnégyzete) számításánál használható, mivel $\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2$.
- $n = 3$: harmadik momentum, $\langle k^3 \rangle$, az eloszlás ferdesége, megmutatja, hogy p_k mennyire szimmetrikus $\langle k \rangle$ körül.

Skálafüggetlen hálózatokban a fokszámeloszlás n -edik momentuma:

$$\langle k^n \rangle = \int_{k_{min}}^{k_{max}} k^n p(k) dk = C \frac{k_{max}^{n-\gamma+1} - k_{min}^{n-\gamma+1}}{n - \gamma + 1} \quad (28)$$

Ha a fokszámok normális eloszlást követnének, akkor egy véletlenszerűen kiválasztott csomópont fokszáma

$$k = \langle k \rangle \pm \sigma_k \quad (29)$$

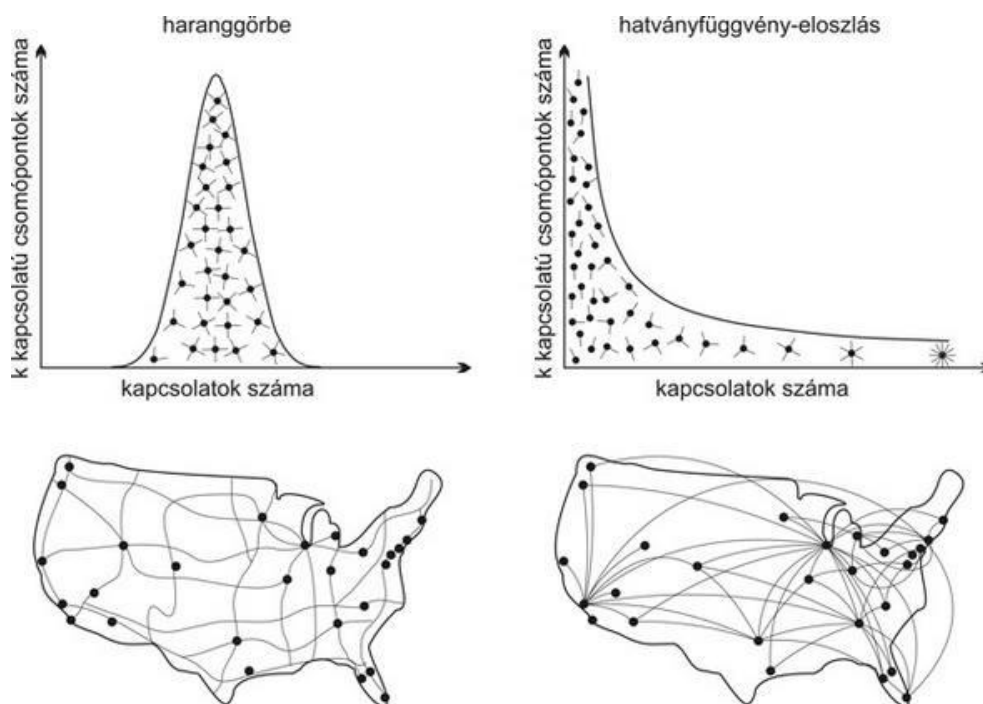
A $\langle k \rangle$ és a σ_k (szórás) azonban más nagyságrendű a véletlen és a skálafüggetlen hálózatokban.

Poisson-eloszlást követő fokszámeloszlású véletlen hálózatokban a hálózati csomópontok fokszáma a

$$k = \langle k \rangle \pm \langle k \rangle^{1/2} \quad (30)$$

tartományba esik. A véletlen hálózatok fokszáma tehát hasonló, a $\langle k \rangle$ átlagos fokszám a véletlen hálózat „méretskálája”. Ellenben a $\gamma < 3$ fokszámeloszlású hatványfüggvények esetén az átlag körüli ingadozás tetszőleges lehet, tehát nincs értelmezhető belső skála. (Barabási, 2017)

A nem skálafüggetlen és a skálafüggetlen hálózatokra ad példát az USA autóúthálózatát és belföldi légitársasági-hálózatát összehasonlító 1. ábra.



1. ábra Nem skálafüggetlen és skálafüggetlen hálózat (Barabási, 2003)

A valódi skálafüggetlen és véletlen hálózatok meglehetősen ritkák, a legtöbb létező hálózat kialakulásában és fejlődésében számos hatás szerepet játszik.

A hálózat kutatásban ezért fontos kérdés, hogy a véletlen (Poisson- vagy normális), illetve a hatványfüggvény-eloszlás mellett esetleg más eloszlással jellemezhetjük-e a fokszámok gyakoriságait.

A statisztikában ez a kérdés az illeszkedésvizsgálat témakörébe tartozik.

A fokszámok kapcsán azonban tisztáznunk kell két látszólag triviális kérdést.

Első probléma, hogy a csomópontonkénti fokszámot vajon diszkrét változóként vagy folytonos változóként kezeljük-e.

A fokszám nyilván diszkrét változó, azonban értéke elvben jóval magasabb is lehetne, illetve még a mostaninál is többféle értéket vehetne fel nagyobb hálózat, vagy intenzívebb együttműködés esetén, ráadásul arányskálán mozog. Ilyen esetekben az elterjedt statisztikai szoftverek bevezetik a „folytonosként kezelt diszkrét változó” fogalmát, s a diszkrét változóra folytonos elemzési módszerek használatát javasolják. (Acock, 2018) (IBM Corp., 2020) Az eloszlások illeszkedésvizsgálatára használt *fitdistrplus* csomag megalkotói, s *vignette*-jének szerzői² is folytonosként kezelnek nagy elemszámú diszkrét változókat. (Delignette-Muller & Dutang, 2020) Gyakoribb eloszlások illeszkedésvizsgálati módszereinél azonban nagyon változatos a folytonos módszerek használhatósága diszkrét változókra. (Clauset, et al., 2009)

Hasonló problémát jelentett, hogy az elemzés során sokasági vagy mintavételi statisztikai eljárásokat használjunk-e. A faipari projektek egyes jellemzőinek sokaságként történő lehatárolásának problémái, a nem feltétlenül teljes projektháló, az adatgyűjtés törvényszerűen tökéletlen volta miatt indokolt a bizonytalanságokat és tökéletlenségeket is „kezelő” – esetünkben pl. bootstrapping módszerek – választása, amelyek megfelelő körültekintéssel többféle adatsor esetén bevethetők. Az általunk vizsgált projekthálót az összes lehetséges projektháló vagy a faipari projektek egy lehetséges reprezentációjaként, mintájaként értelmezve sokkal súlyosabb mintavételi kérdésekkel szembesülünk, mintha a H2020-ban regisztrált faipari projekteket sokaságként vizsgáljuk. Emiatt az egész elemzés során inkább az utóbbi felé hajlottunk, de a szakirodalmi tanácsokat követve esetenként elvégeztük a szükséges próbákat.

A szokványos illeszkedésvizsgálati kérdés ezért esetünkben módosulhat, hiszen nem a minta alapján akarunk becslést adni arra, hogy a sokaság megfelel-e az adott eloszlásnak, hanem (módszertani fenntartásokkal és óvatossággal) arra keressük a választ, hogy maga a sokaság megfelel-e annak.

Ha nem keresünk pontos paramétereket, hanem csak az illeszkedésre vonatkozó hipotézisünket akarjuk megerősíteni vagy elvetni, ez valójában nem jelent kardinális problémát. (Clauset, et al., 2009) (Delignette-Muller & Dutang, 2020) (Gillespie, 2020)

Természetesen minden ilyen vitatható módszertani kérdést nagyon óatosan kell megközelíteni. Törekedtünk a „bármilyen” adatsornál megbízható eredményt adó módszerek választására, illetve a nagy eltéréseket adó módszerek kizárására.

Először – diszkrét változóra érvényes eljárást választva – megvizsgáltuk, hogy az adatsor Poisson-eloszlást követ-e. A Poisson-eloszlás tesztelésére χ^2 próbát és maximum likelihood módszert is használhatunk. A tesztelés mintára és sokaságnak tekintett csoportra is használható.

² A francia szerzők elismert nemzetközi szaktekintélyek a statisztikai programozásban, az *Université de Lyon*, illetve az *Université de Strasbourg* oktatói.

Az χ^2 próba használatának feltétele, hogy az elemszám nagyobb legyen, mint az összes kategória számának ötszöröse. Emellett 20%-nál nem lehet több az 5-nél kisebb elemszámok aránya. (Larntz, 1978) Mivel ez utóbbi nem teljesült a maximum likelihood módszert választottuk. A 20%-os feltételt gyakran a maximum likelihood módszernél is említik, de jelentősége itt kisebb. A maximum likelihood módszer kisebb ($n > 40$) elemszámnál is viszonylag megbízható eredményt ad Poisson-eloszlás illeszkedésvizsgálatára. (Pervin, et al., 2020) Más kutatók szerint maximum likelihood módszer esetén a megbízhatóság feltétele, hogy a kategóriákhoz tartozó elemszámok átlaga ne legyen kisebb ötnél. (Friendly & Meyer, 2015)

Poisson-eloszlás esetén az eloszlás egyetlen paramétere a λ , vagy becslés esetén a $\hat{\lambda}$.^{3 4}

A 15. szakmai hipotézishez kapcsolódóan – egyfajta ellenpróbaként – megvizsgáltuk, hogy az eloszlás Poisson-eloszlásból származik-e.

$$H_0: \text{Az eloszlás Poisson-eloszlásból származik, } \lambda > 0 \text{ értékkel.} \quad (31)$$

$$\hat{p}_k = \frac{\hat{\lambda}^k}{k!} e^{-\lambda} \quad (32)$$

ahol $k = (1, 2, 3, \dots)$ esetünkben a fokszámokat jelöli.

Az egyes kategóriákhoz a Poisson-függvény által becsült értékek:

$$\hat{m}_i = \hat{p}_k \sum n \quad (33)$$

ahol n a tapasztalati gyakoriság értékeit, \hat{m} pedig a becsült Poisson-eloszlású függvényünk értékeit jelöli az egyes kategóriáknál.

A becsült paraméterek száma d , ezért a szabadságfok:

$$df = r - d - 1 \quad (34)$$

ahol r a kategóriák száma, de látunk majd példát arra, hogy az adattudományban jellemző felfogást követve a $df = r$ képletből indulunk ki. A két végeredmény között ritkán van jelentős különbség.

A próbastatisztika ebben az esetben:

³ Bár a korábbi okfejtés alapján használhatnánk λ -t is, az eljárást becslésnek tekintve $\hat{\lambda}$ -t számoljuk ki.

⁴ A számítás menetére a magyar nyelvű források közül egyedül Backhausz (2020) levezetését találtuk követhetőnek, számítását megerősítettük, azonban az általa alkalmazott kategóriaösszevonás – amely a szakirodalomban az alacsony elemszámú kategóriáknál gyakori (bár vitatott) módszer – véleményünk szerint a nullhipotézist a „talán” nem elutasítandóból a biztosan nem elutasítandóba tolja át. Ezért a χ^2 próba helyett ilyen esetben is a maximum likelihood módszer használata indokolt úgy, hogy G^2 számítása során a végtelenhez közelítő logaritmus-értékeket nem vesszük figyelembe. A szabadságfokot ilyen esetben a „kihagyott” értékek számával csökkenteni kell. (Friendly & Meyer, 2015)

$$G^2 = 2 \sum_i^N n_i \ln \left(\frac{n_i}{\hat{m}_i} \right) \quad (35)$$

Ha a próbastatisztika értéke kisebb, mint a táblázatban a szabadságfokhoz és a szignifikanciaszinthez tartozó χ^2 -érték (c_{krit}), akkor a tapasztalati értékeink Poisson eloszlásúak.

Diszkrét változók esetén az alábbi valószínűségi eloszlás α és ζ paramétereit keressük (Clauset, et al., 2009):

$$p(x) = \frac{x^{-\alpha}}{\zeta(\alpha, x_{min})} \quad (36)$$

ahol n tapasztalati értékre:

$$\zeta(\alpha, x_{min}) = \sum_{n=0}^{\infty} (n + x_{min})^{-\alpha} \quad (37)$$

Tehát keresünk egy α paraméterű hatványfüggvényt, amely $x_{min} > 0$ értéktől leírja a tapasztalati értékeinket is. Az 15. szakmai hipotézis statisztikai szempontból azonos a

$$H_0: \text{Az eloszlás hatványfüggvény-eloszlásból származik.} \quad (38)$$

hipotézissel.

A hipotézisvizsgálat a Kolmogorov–Smirnov tesztre épül. (Clauset, et al., 2009)

A módszer lényege esetünkben, hogy felírjuk a tapasztalati adatok kumulatív eloszlásfüggvényének ($S(x)$) és a legjobban illeszkedő hatványfüggvény-eloszlás kumulatív eloszlásfüggvényének ($P(x)$) értékei közötti távolságot:

$$D = \max_{x \geq x_{min}} |S(x) - P(x)| \quad (39)$$

Az általunk (a számítógépes algoritmus által) becsült \hat{x}_{min} a D -t minimalizáló x_{min} értéke.

A szakirodalomban a $p \leq 0,05$ vagy a $p \leq 0,1$ (konzervatív felfogás) érték esetén elutasítják H_0 -t. Azonban ennél magasabb p értéknél sem hozunk feltétlenül más döntést, például magas x_{min} esetén.

A számítás során felhasznált algoritmusok (függvények) elvben lehetőséget adnak tisztán diszkrét változót feltételező becslésre, de jobb közelítést adnak akkor, ha a folytonos algoritmust használjuk úgy, hogy közben jelezzük, hogy diszkrét változóról van szó. (Nepusz & Csardi, 2023)

A fenti számítás tökéletlensége miatt érdemes a nagyobb gépigényű – a vizsgált adatsorból sok visszatevéses mintavételt és becslést végrehajtó, majd a kapott értékeket kumuláló – bootstrapping-módszereket használnunk. Az előző bekezdésekben tárgyalt értékekre – hasonló matematikai statisztikai összefüggések alapján – ezért bootstrapping-módszerrel is adtunk becslést.

A használt algoritmussal előbb még megerősítő számítást végeztünk az előző számításra, majd alkalmaztuk a bootstrapping módszert 1000 iterációval. (Jellemzően 500 – 5000 iteráció használata tekinthető általánosnak.) Ezután a projektrészvételek eloszlását is teszteltük.

Bár a használt csomagok függvényeit mintavételi eljárásokra tervezték, azonban az megfelelő paraméterezéssel sokasági változókra is megfelelő információkat nyújtanak. Ráadásul mi „csak” az eloszlás jellegét akarjuk megbecsülni, illetve a korábban kapott értékeinket tesztelni.

A következő lépésben felrajzoltuk a foksámok sűrűség- és a kumulatív eloszlásfüggvényét.

Tehát a diszkrét változónkat itt a kapcsolódó dokumentációban támogatott módon folytonosként kezeltük, de vizuálisan diszkrét változóként kezelve is apró lett volna a különbség.

A vizuálisan levonható következtetések után a deskriptív statisztikából ismert harmadik (*ferdeség – skewness*) és negyedik (*kurtózis – kurtosis*) momentumot használtuk a foksámeloszlást leíró eloszlás kiválasztására. A nem negatív ferdeség a szimmetria hiányára, a kurtózis 3-tól eltérő (tehát normál eloszlástól eltérő) értéke lapitottságra vagy csúcsosságra utalhat.

A ferdeség és a kurtózis ebben az esetben (Casella & Berger, 2002):

Független és azonos eloszlású véletlenszerű változókat feltételezve, $(x_i)_i$ megfigyelés esetén $(X_i)_i \sim X$.

$$sk(X) = \frac{E[(X - E(X))^3]}{Var(X)^{\frac{3}{2}}} \quad (40)$$

$$\widehat{sk} = \frac{\sqrt{n(n-1)}}{n-2} \times \frac{m_3}{m_2^{\frac{3}{2}}} \quad (41)$$

$$kr(X) = \frac{E[(X - E(X))^4]}{Var(X)^2} \quad (42)$$

$$\widehat{kr} = \frac{n-1}{(n-2)(n-3)} \left((n+1) \times \frac{m_4}{m_2^2} - 3(n-1) \right) + 3 \quad (43)$$

ahol m_2, m_3, m_4 a tapasztalati momentumokat írja le a

$$m_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k \quad (44)$$

képlettel, ahol x_i x n darab tapasztalati értéke, \bar{x} pedig ezek átlaga.

A használt függvény a klasszikus deskriptív statisztika mellett ferdeséget és kurtózist is becsül alapértelmezés szerint mintavételi torzítások nélkül.

Mivel itt sokaságként kezeljük a csoportot, nem módosítottuk az eljárást mintavételre. A két eredmény közötti különbség egyébként minimális lenne.

Az eljárás ráadásul előállítja a magyar szakirodalomban ritkán említett *Cullen-Frey diagramot* (*Kurtózis-ferdeség diagramot*), amely feltünteti a leggyakoribb eloszlásokhoz tartozó lehetséges értékeket. (Cullen & Frey, 1999) Mivel a ferdeség és a kurtózis nem robusztus (kis paraméterváltozásokkal nagy kilengéseket mutathat) nem paraméteres (nem normalitásra épülő) *bootsrapping* eljárást (Efron & Tibshirani, 1994) választottunk $boot = 1000$ értékkel. Ez szinte bizonyosan megbízható és vizuálisan jól kirajzolódó eredményt ad. A diagramot diszkrét változóra és folytonosan kezelt diszkrét változóra is előállítottuk.

Ezután kiszámoltuk az *átmérőt*. Az átmérő a hálózat „pályahossza”: a legnagyobb számú lépés, amely ahhoz szükséges, hogy egy csomóponttól eljussunk bármely más csomóponthoz a lehető legrövidebb utat bejárva. A kis átmérőjű hálózatokat „kisvilág”-nak nevezik. (Barabási, 2006)

Átlagos fokszámú véletlen hálózatból kiindulva, a kezdőponttól d távolságnál messzebb lévő csomópontok száma:

$$N(d) \approx 1 + \langle k \rangle + \langle k \rangle^2 + \dots + \langle k \rangle^d = \frac{\langle k \rangle^{d+1} - 1}{\langle k \rangle - 1} \quad (45)$$

ahol $\langle k \rangle$ az átlagos fokszám.

$N(d)$ nem lehet nagyobb N -nél (a csomópontok teljes számánál), ezért a távolság tetszőleges értéket nem vehet fel. A maximális távolságra (d_{max}), illetve a hálózat átmérőjének elemszámára igaz, hogy:

$$N(d_{max}) \approx N \quad (46)$$

Ha $\langle k \rangle \gg 1$, akkor (45) egyenlet számlálójából és nevezőjéből egyaránt elhagyhatjuk a „ -1 ”-et:

$$\langle k \rangle^{d_{max}} \approx N \quad (47)$$

A hálózat átmérője tehát:

$$d_{max} \approx \frac{\ln N}{\ln \langle k \rangle} \quad (48)$$

Az (48) egyenlet a kisvilág jelenséget írja le. (Barabási, 2017)

Mivel a legtöbb hálózatban az (48) egyenlet jobb közelítést ad két véletlenül kiválasztott csomópont átlagos távolságára ($\langle d \rangle$ -re), mint d_{max} -ra, ezért a kisvilág jelenséget gyakran a következő képlettel jellemzik:

$$\langle d \rangle \approx \frac{\ln N}{\ln \langle k \rangle} \quad (49)$$

Kisvilág esetén tehát a „kis”-en azt értjük, hogy az átlagos úthossz vagy az átmérő logaritmusan függ a hálózat hosszától. A 17. szakmai hipotézist a fenti értékek kiszámításával teszteltük.

A *közöttiség* azt vizsgálja, hogy az adott szereplő hálózati elhelyezkedése mennyire kritikus a hálózati együttműködés és információáramlás szempontjából. Ha egy csomópont sok olyan útvonalon fekszik, amely minimális útvonal két másik szereplő között, akkor valószínűleg fontos szerepet tölt be a hálózatban. (Kürtösi, 2011) (Freeman, 1977)

A közöttiség v csomópontra:

$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (50)$$

ahol σ_{st} az s és a t csomópontok közötti legrövidebb utak száma, $\sigma_{st}(v)$ pedig az ezek közül v -n áthaladó utak száma.

Gyakran használják a normalizált formát, ahol a (50) kifejezést – irányítatlan gráfok esetén – elosztják $(N - 1)(N - 2)/2$ -vel.

Normalizált formaként gyakran használják az alábbi kifejezést is:

$$normal(g(v)) = \frac{g(v) - \min(v)}{\max(v) - \min(v)} \quad (51)$$

Az érték mindkét esetben a $[0,1]$ tartományba esik.

A közöttiség és a fokszám felhasználásával kiszűrtük a legfontosabb aktorokat.

A szakirodalom alapján még számos mutató számítható, mi azonban csak az elemzési céljaink szempontjából legfontosabbakkal foglalkoztunk.

6.4.1.3. Hálózati klaszterelemzés

A hálózatban lévő klaszterek felírása az előző két alfejezethez hasonlóan szinte önálló tudományterület, mi itt is csak az elemzés szempontjából leglényegesebb módszerekre szorítkoztunk.

Hálózati klasztereket írtunk fel, felrajzoltuk azokat egy hálózati klaszterdiagram formájában, majd elemeztük a legnagyobb elemszámú klasztert.

A használt algoritmus egyre kisebb és kisebb részekre osztja a hálót, amíg meg nem találja azokat az elemeket, amelyek az egyes csoportok között hídként szolgálnak, mivel náluk nagy a közöttiség (50) értéke. Bár a technikai dokumentáció szerzői (The igraph core team, 2003-2020) egyértelműen utalnak az algoritmus matematikai-statisztikai alapjaira (Newman & Girvan, 2004), a hivatkozott mű alapján (50)(51) egyenletek szolgálnak elsősorban a klaszterek elhatárolására az összetett algoritmuson belül.

A 16. szakmai hipotézist az egyes programokon belüli faipari projekthálóok legnagyobb klasztereinek közöttiségalapú elemzésével teszteltük.

6.5. A faipari projektek outputjai

A faipari projektek eredményeit a nyílt és a zárt innovációs modellek tükrében, főleg azok kimenetei alapján mértük fel.

Regressziót az innovációs folyamatok összetettsége, illetve az eredmények és a támogatás egyéb tényezőik (pl. a pályázati kiírásokban foglalt minimumkövetelmények) általi meghatározottsága miatt nem számoltunk, ezért inkább korrelációt mértünk többféle viszonylatban. A 18. szakmai hipotézist tehát az 5. szakmai hipotézishez hasonlóan vizsgáltuk.

Az európai uniós projektek outputjai mellett fontosnak tartottuk, hogy a zárt innovációk legfontosabb eredményét, a faiparhoz kapcsolódó szabadalmakat is elemezzük. Az elemzéshez legfontosabb adatforrásként az Európai Szabadalmi Hivatal (EPO) Espacenet szolgáltatása és a Szellemi Tulajdon Világszervezetének Patentscope szolgáltatása szolgált.

A letöltött adatokon egyszerű megoszlási viszonyszámok mellett szövegbányászati elemzést végeztünk. A szövegbányászati elemzés szorosan kapcsolódik a hatékonyság tágabb értelmezéséhez, hiszen a szabadalmak összefoglalóiból esetleg kinyert kurrens kulcsszavak segíthetnek a hatékony kutatási és forrásbevonási stratégiák kialakításához.

Az európai uniós projektekben azonban alacsony számuk miatt szövegbányászati eszközökkel nem a projektekben keletkezett szabadalmakat, hanem a projektek absztraktjait vizsgáltuk. Ez az összehasonlítás érvényességét némileg csökkenti. Kérdéses, hogy az EU-projektek kis mintája és a szabadalmak hatalmas mintája – tartalmi szempontból – mutathat-e átfedést.

Az általunk használt közepes erőforrásigényű algoritmusok igen hosszú futási idővel korlátozott eredményeket ígértek, ennél bonyolultabb eszközök használata nem tűnt célszerűnek, illetve a reálisan elérhető számítógépek számítási és memóriakapacitásait is meghaladják. Véleményünk szerint a szövegbányászat jelentősebb erőforrások bevonása mellett kijelölheti a jövőbeli kutatási irányokat.

A különböző szófajok gyakoriságaira vonatkozó számításokat is végeztünk, illetve az egyes szófajok leggyakoribb előfordulásait is kiszűrtük. Ezeket azonban csekély információtartalmuk miatt nem közöltük, hanem a gyakori jelzős és egyéb szerkezetekre koncentráltunk. Az algoritmusok minden esetben a természetes nyelvet leíró szótáraadatokkal való összevetéssel és kigyűjtéssel kezdődnek, ez igen hosszadalmassá teszi az elemzéseket.

A korábban már említett R-csomagok mellett a természetes nyelvi szövegek elemzésére széles körben használt *UDPipe* csomagot használtuk különböző algoritmusokkal a faipari projektek és a szabadalmak absztraktjaira.

Az algoritmusok általános (de nyelvfüggő) elemeit, tehát a tokenizálást (a dokumentumok szavak sorozatára bontását), az egységes írásmódra alakítást, a lemmatizálást (a szavak szótőre, szótári alakra való visszavezetését), a felesleges írásjelek eltávolítását, a jelentés nélküli stopszavak negligálását, a szavak szófajának meghatározását (POS – part-of-speech, szófajok szerinti – tagging) nem érintjük.

A *RAKE-algoritmus* azon a megfigyelésen alapul, hogy a kulcsszavak gyakran tartalmaznak több szót a szokásos írásjelekkel vagy szünetjelekkel, kötőszavak és névelők után, minimális lexikai jelentéssel. Az ilyen stopszavakat jellemzően a legtöbb információs rendszerben elhagyják, és a különböző szövegelemzésekben sem veszik figyelembe, mivel értelmetlennek tekintik őket. Azonban ezek a szavak gyakran mégis tartalomhordozónak bizonyulnak. A bonyolult összehasonlítások és szűrések után előállt kifejezések leginkább az adathalmazban

elfoglalt sorrendet kijelölő – önálló, általános értékkel nem nagyon bíró – RAKE-értéket kapnak. A magas RAKE-értékű kifejezés „fontosabb”, mint az alacsony RAKE-értékű. (Wijffels, 2023)

A POS-szekvenciákra épülő algoritmus az adott nyelven jellemző szerkezeteket (pl. jelzős szerkezeteket) szűri ki. A PMI-algoritmus (Pointwise Mutual Information) hasonló célt követ, de jobban épít az asszociációs kapcsolatokra.

7. Eredmények

7.1. A szakirodalom szövegbányászati és tudományometriai elemzése

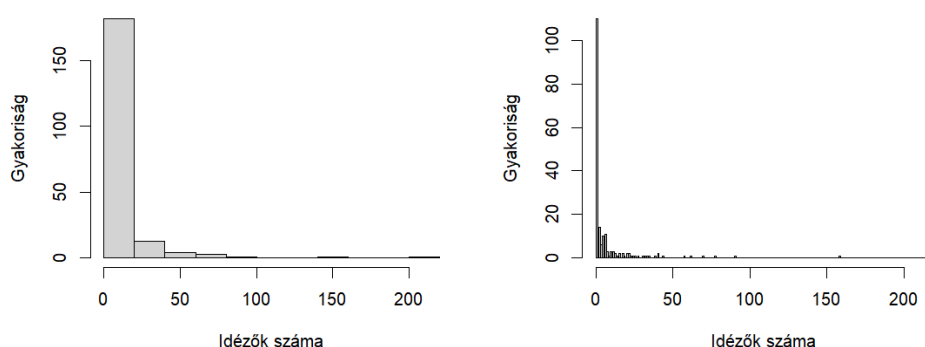
A faipari projekteket elemző szakirodalom szövegbányászati elemzése során fokozott óvatossággal kellett eljárunk, mivel a pontatlan BibTeX-fájlokból az R algoritmusai gyakran megtévesztő statisztikákat állítottak elő. Ez elsősorban az első szerzők országait érintette. (Egyes országok nem egységes formában/rövidítéssel szerepeltek.) Manuális módosítás/kiegészítés után azonban megfelelő eredményt kaptunk.

Az 1. táblázat a vizsgált publikációk leíró statisztikájának legfontosabb adatait tartalmazza.

1. táblázat A faipari projektekhez kapcsolódó publikációk legfontosabb adatai

Mutató	Érték
Évek	1984-2022
Dokumentumok száma	205 darab
Átlagos éves növekedés	6,48%
Idézők átlagos száma	8,37
Éves átlagban	0,74
Hivatkozott források száma	4.198 darab
A publikációk átlagos kora (2022.10.19.)	11,9 év

A 2. ábra első diagramja (kevés osztópont az x-tengelyen) vizuálisan megerősíteni látszik a hatványfüggvény-eloszlást, a második diagram (sok osztópont az x-tengelyen) kevésbé valószínűsíti ezt.

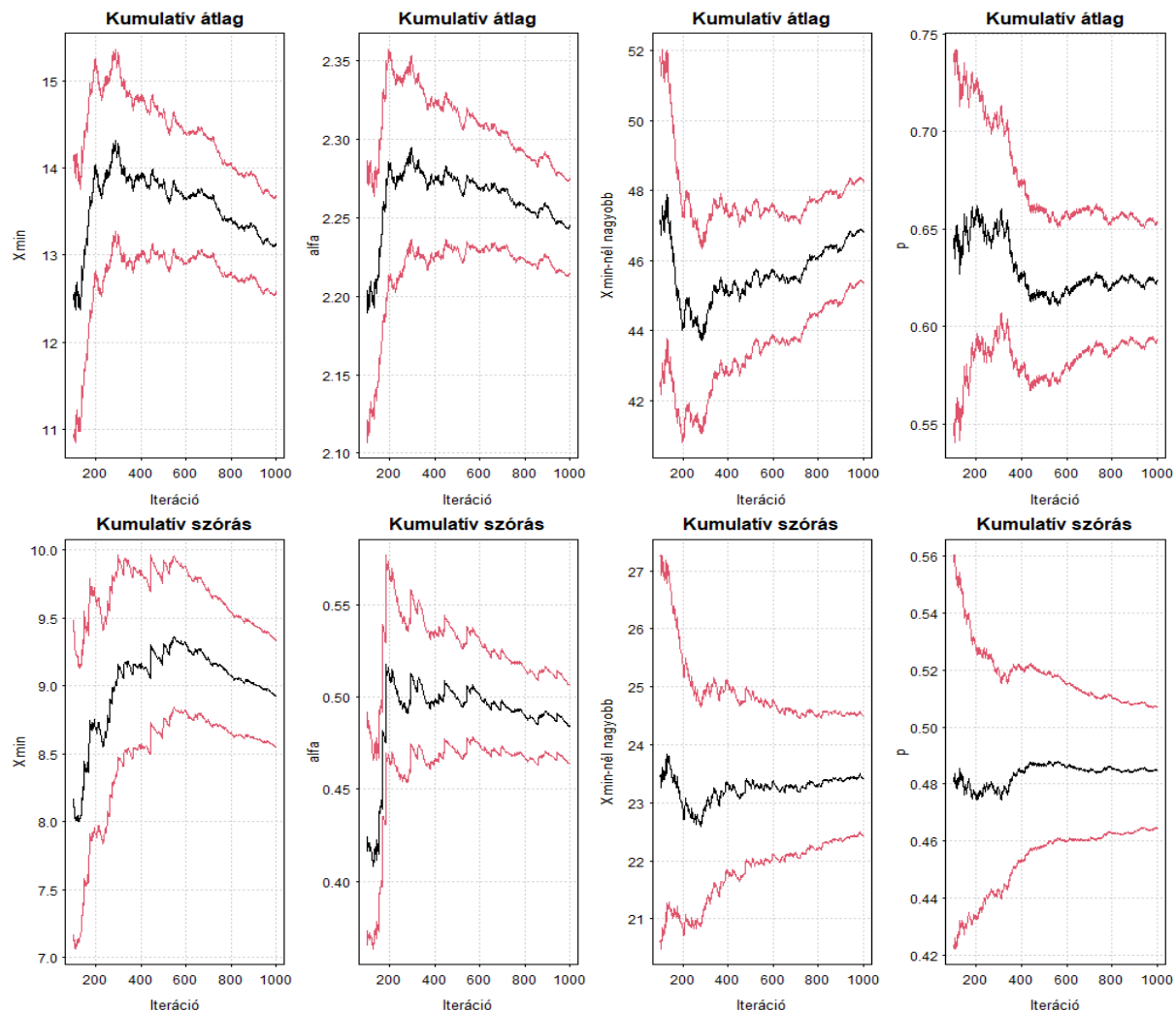


2. ábra A publikációk idézőinek száma eltérő számú osztópont mellett (x-tengely)

1000 iterációra épülő, p becslésére és az eloszlás tesztelésére irányuló bootstrapping-eljárás mellett $x_{min} = 13$, $p = 0,623$, α (vagy γ) = 2,25. Mivel p értéke jóval nagyobb, mint a

módszer elméleti alapjainak kidolgozói (Clauset, et al., 2009) által javasolt $0,1$, emellett $2 < \alpha(\gamma) < 3$, a hatványfüggvény-eloszlásra vonatkozó $H_0(38)$ hipotézist nem utasítjuk el. (3. ábra)

Ez egyben azt jelenti, hogy az 1. szakmai hipotézist, miszerint a faipari projektekkel foglalkozó szakirodalom idézőinek száma hatványfüggvény-eloszlást követ egy bizonyos x_{min} érték után, nem vetjük el. Azaz a legtöbb cikk nagyon kevés, ellenben néhány cikk nagyon sok hivatkozást kap.



3. ábra Idézők hatványfüggvényeloszlás-próbájának értékei bootstrapping(p) módszerrel

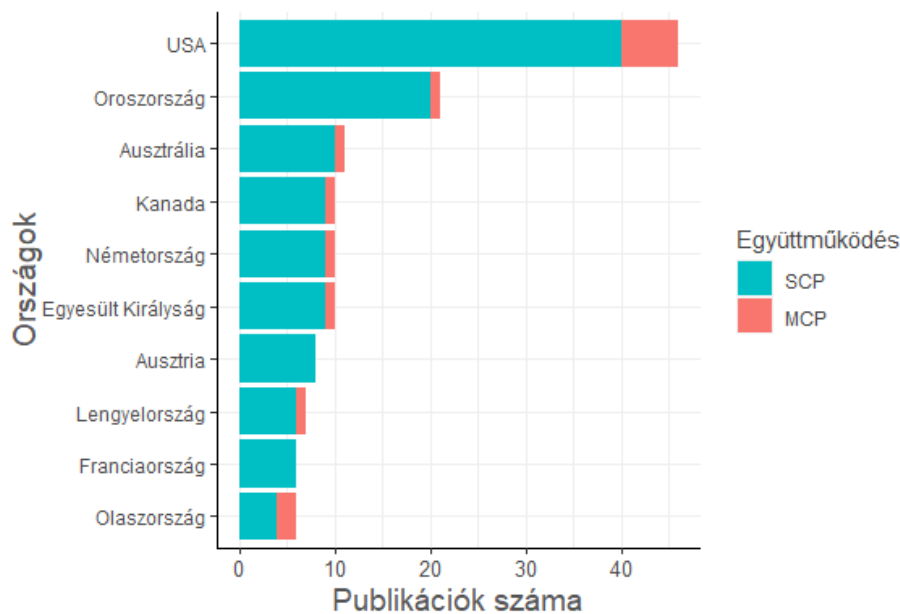
Hangsúlyozni kell azonban azt is, hogy x_{min} -nél kisebb értékekre a hatványfüggvény-eloszlás nem érvényes. A 205 kapcsolódó cikkből átlagosan csak 47 esik az x_{min} -nél nagyobb tartományba, a cikkek többségének pedig egyetlen idézője sincs.

A kevésbé idézett – gyakran régi, aktualitását veszített vagy a centrumországok tudományos közvéleménye számára különböző okokból kevésbé érdekes cikkek – pontos idézettsége azonban nem feltétlenül tárható fel a Scopus alapján, elképzelhető, hogy ezek idézettsége a Scopusban nem szereplő cikkeket is figyelembe véve magasabb. Ez azonban nagyon nehezen igazolható. (Az x_{min} becslésére épülő bootstrapping-eljárás kissé más eredményt ad, mint a

hatványfüggvény-eloszlás tesztelését szolgáló bootstrapping(p), de az eltérés nagysága lényegtelen, oka pedig túlmutat a dolgozat keretein.)



4. ábra Legtermékenyebb szerzők

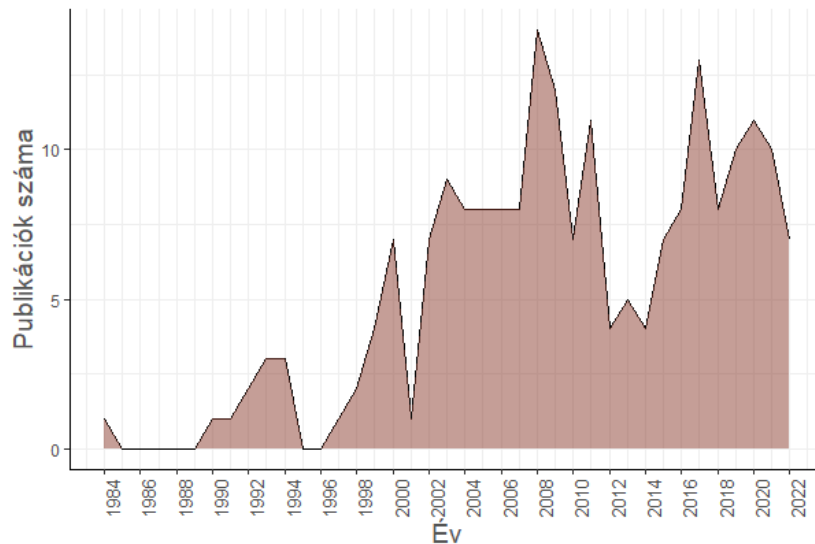


SCP: Single Country Publications (egy ország), MCP: Multiple Country Publications (több ország)

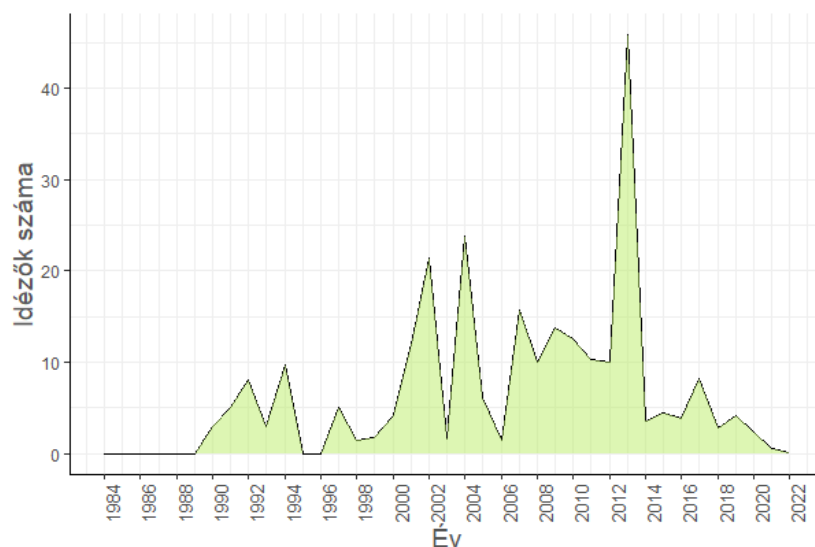
5. ábra A legtöbb cikket kibocsátó országok

A témában publikáló szerzők között a publikációk száma alapján nincs igazán kiemelkedő szerző. A legtöbb cikk az Egyesült Államokban született, Oroszország második helye – figyelembe véve, hogy az oroszországi szerzők a legtöbb országhoz képest nagyobb mértékben publikálnak nem angol nyelvű (orosz), a Scopusban szórványosan megtalálható kiadványokba (Macháček, 2019) – némileg meglepő, bár korábbi tartalmi elemzésünk már utalt rá. A cikkek és a cikkek idézettségének évenkénti alakulásának mintázata mutat némi hasonlóságot. (4. ábra - 7. ábra)

A hasonlóság tesztelését Spearman-féle rangkorreláció-számítással végeztük el. A normalitás hiánya külön próba nélkül is nyilvánvaló és a sok 0 érték miatt a más elemzésekben gyakoribb, de itt ritkábban használt Pearson-féle korrelációs együttható esetünkben nem értelmezhető. A szignifikancia-szintet (α) 0,05-nek választottuk.



6. ábra Publikációk száma évenként



7. ábra Idézők száma évenként

A 2. táblázat alapján a 2. szakmai hipotézis, tehát, hogy a faipari projektekkel foglalkozó cikkek és idézők évenkénti számának alakulása között korreláció van – azaz a függetlenségre vonatkozó $H_0(1)$ statisztikai hipotézis elutasítása –, nemcsak az adott évi, hanem a publikációkhoz képest egy vagy két évvel későbbi citációk esetén is teljesül, mivel $p < \alpha$. A publikációk és a citációk száma között tehát a 0 – 2. években közepes korreláció tapasztalható 95%-os konfidencia-intervallum mellett. A későbbi években is tapasztaltunk esetenként a második éves értéknél magasabb ρ -t, de p már minden esetben nagyobb volt, mint α , így H_0 -t egyértelműen nem utasíthattuk el. A publikációk és a citációk éveinek talán nem szokványos lehatárolását egyfelől az esetleges időbeli elcsúsztatás biztosításának igénye, másfelől a két vizsgált idősor korrelációanalízishez szükséges egyenlő elemszáma indokolta.

2. táblázat Spearman-féle rangkorrelációs teszt a publikációk és a citációk évenkénti száma között ($\alpha = 0,05$)

Publikációk évei	Citációk évei	ρ	p	Döntés
1984-2022	1984-2022	0,52	0,0007	H_0 -t elvetjük.
1984-2021	1985-2022	0,42	0,008	H_0 -t elvetjük.
1984-2020	1986-2022	0,45	0,005	H_0 -t elvetjük.
1984-2019	1987-2022	0,25	0,14	H_0 -t nem vetjük el.

Az 8. ábra a keresőkifejezésünknek megfelelő cikkek kulcsszavai közötti kapcsolatot mutatja. Az alkalmazott algoritmus nemcsak a kulcsszavak közötti kapcsolatokat tárta fel, hanem három témaklasztert is elkülönített. A késsel jelölt klaszter faipari, a pirossal jelölt faipari-erdészeti, míg a zöld inkább faipari menedzsmenthez kapcsolódó cikkekre-projektekre utal.

A Bradford-zónákhoz tartozó értékeket az 3. táblázat mutatja.

3. táblázat A Bradford-zónák értékei

Zóna	Folyóiratok száma	Megoszlás	Cikkek száma
1.	24	15,4%	69
2.	65	41,7%	69
3.	67	42,9%	67
Σ	156	100%	205

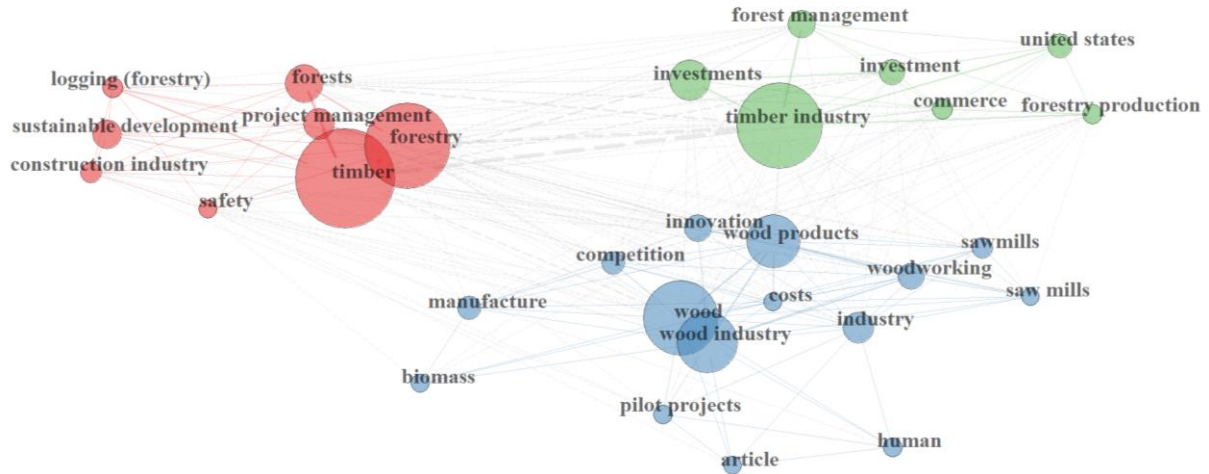
Az egyes zónákba eső folyóiratok számának aránya – 24: 65: 67 – nem közelítenek a Bradford 1: n : n^2 tapasztalati értékeihez.

(2) alapján $k = 2,42$, (3) szerint $r_0 = 16,77$. Az Egghe-féle módosított képlettel számított Bradford-eloszlás ekkor a három harmadban 16,77: 40,66: 98,57.

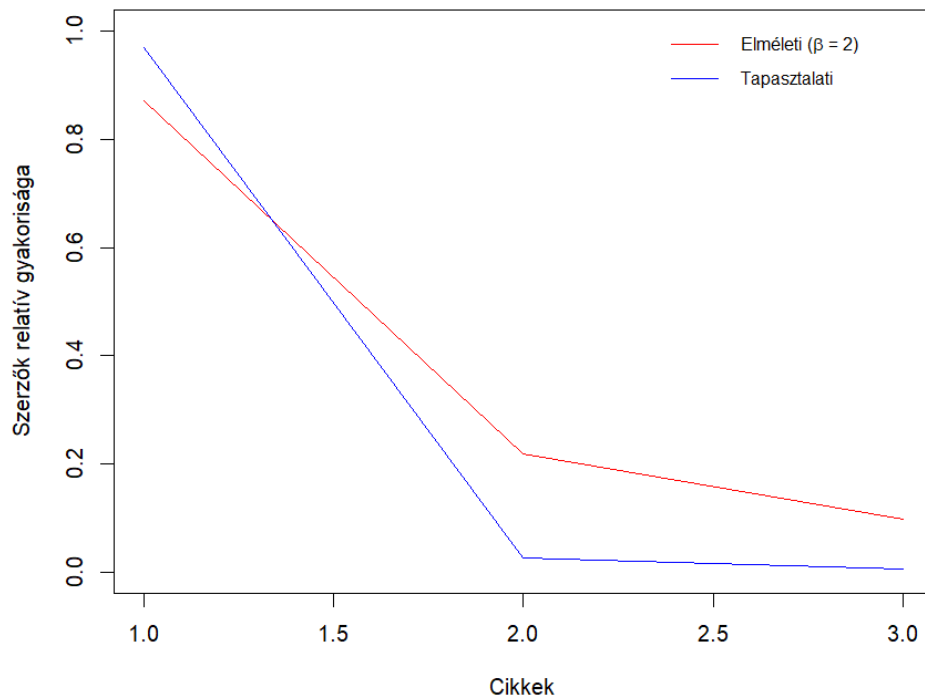
Látható, hogy a tapasztalati értékeink nem követnek Bradford-eloszlást. Tehát a 3. szakmai hipotézist, miszerint a Scopusból elérhető, faipari projekteket leíró publikációk nem tekinthetők a témában megszületett publikációk valódi reprezentációjának, nem vetjük el.

A Lotka-eloszlás kétmintás, Kolmogorov-Szmirnov próbájának β értéke (4,72) jelentős mértékben eltér az Alfred J. Lotka által becsült általános értéktől (2). Ez azonban nem feltétlenül jelenti azt, hogy a szerzői produktivitás gyakorisági eloszlása ne követne Lotka-eloszlást.

A p értéke (0,52) a szignifikanciaszintnél ($\alpha = 0,05$) jelentősen magasabb. Tehát $H_0(4)$ -et nem vetjük el, ezért nem vetjük el azt a feltételezést (4. szakmai hipotézis), hogy a szerzői produktivitás gyakorisága Lotka-elosztást követ. Ezt valamelyest vizuálisan az 9. ábra is megerősíti.

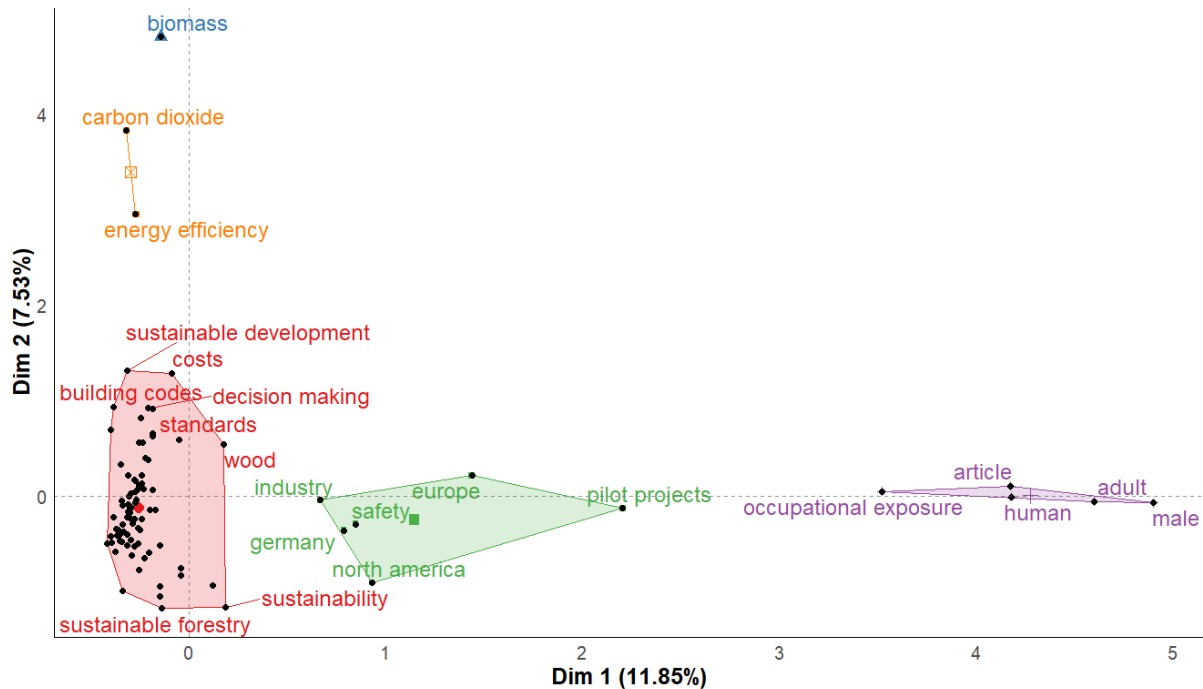


8. ábra Kulcsszavak többsé előfordulásának kapcsolathálója
(a leggyakoribb 30 kulcsszó esetén)



9. ábra Szerzői produktívás tapasztalati és Lotka-eloszlása

A többsé előfordulású kulcsszavak közötti kapcsolat ábrázolása önmagában ritkán ad valós információkat, bár esetünkben ez talán nincs így.



10. ábra A többes előfordulású kulcsszavak konceptuális struktúratérképe

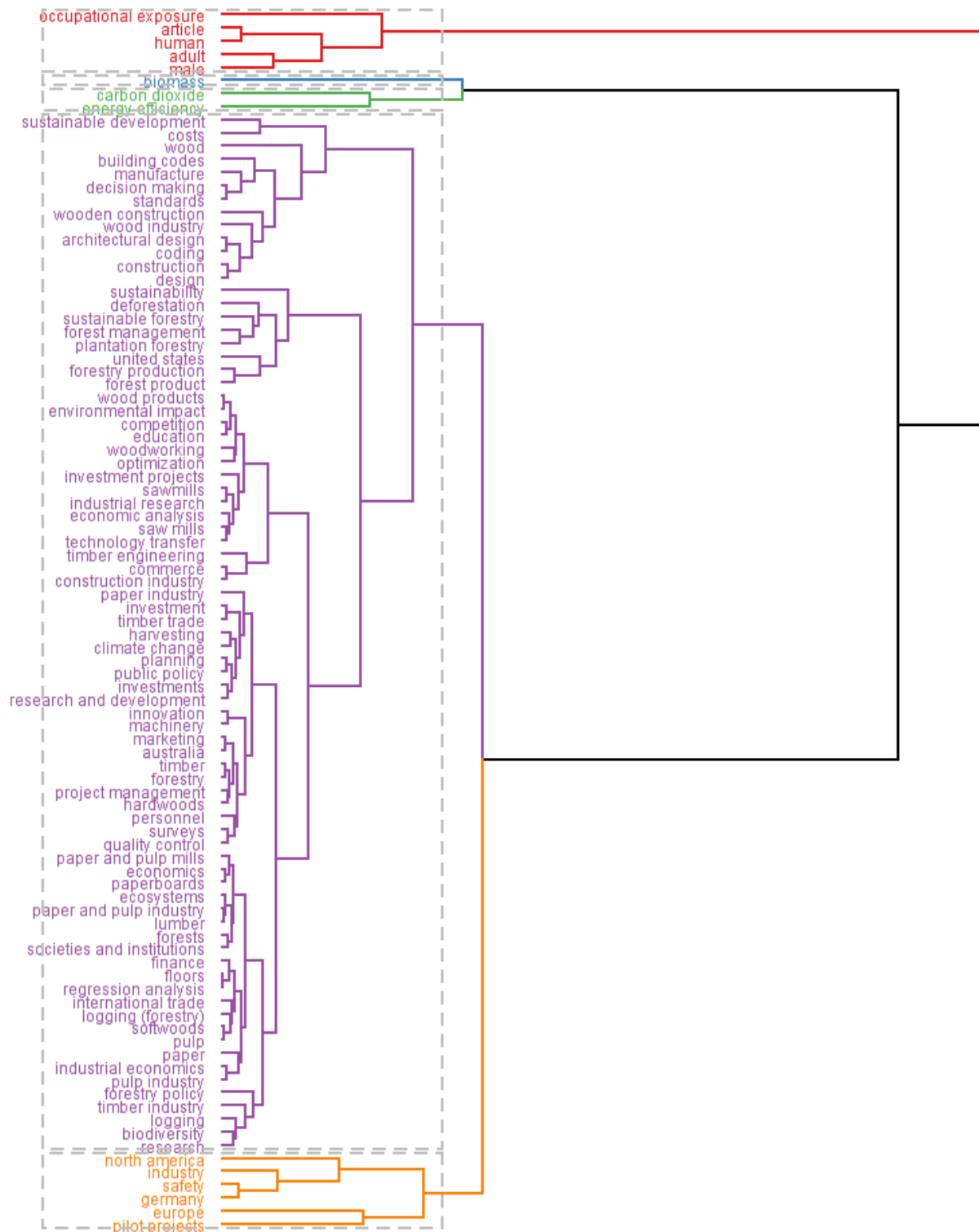
Dimenziócsökkentési eljárásokkal általában relevánsabb információk nyerhetők ki.

A CA (Correspondence Analysis) és az MCA (Multiple Correspondence Analysis) dimenziócsökkentési eljárás, illetve az adattudományban általánosan használt K-közép klaszterezési eljárás együttes eredményét a 10. ábra mutatja.

Az eredmény megerősíti korábbi benyomásunkat, miszerint a faipari projektekben a fenntarthatóság kulcsszerepet játszik. A diagramon 4-5 klaszter rajzolódik ki.

A kutatási témák dendrogramja a többes előfordulású kulcsszavak struktúráját írja le. Az algoritmus emellett 4-5 klaszterre osztja a kulcsszavakat (11. ábra).

A struktúratérkép és a dendrogram egymással lényegében ekvivalens.



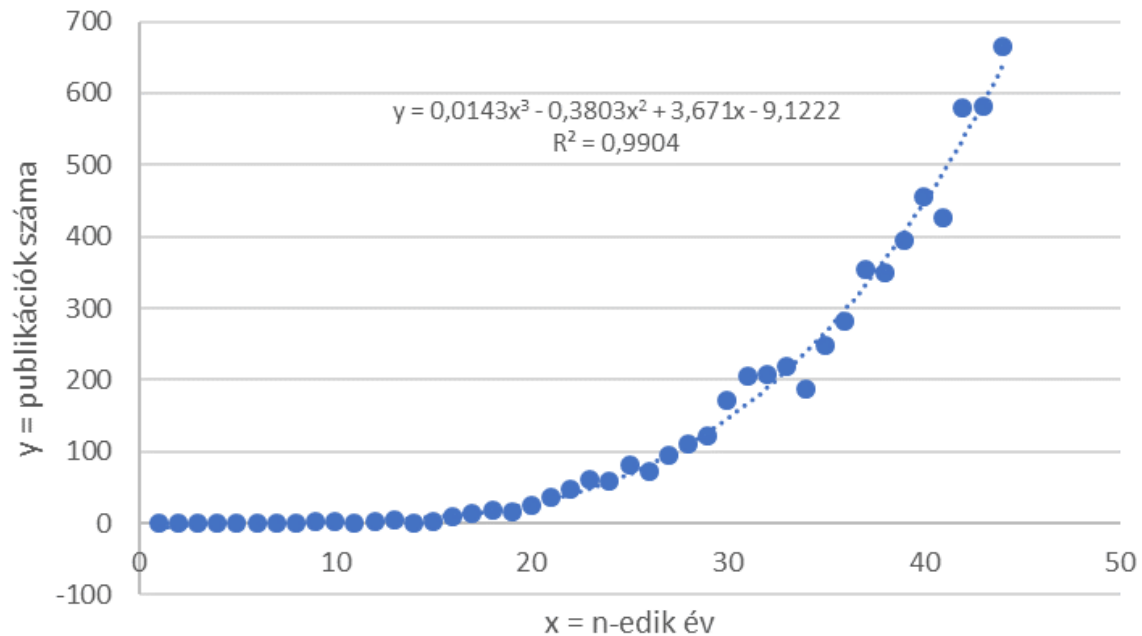
11. ábra A kulcsszavak dendrogramja

7.2. A fenntartható faipar súlya a tudományos kutatásban – eredmények

Az SW-hez tartozó publikációkat 1975 és 2020 között években közzétették az adatbázis szerint. Az SW gyakorisága 2018-ban vette fel a maximális értékét (664-es értékkel). (2019 és 2020 az adatfelvétel idején nem tekinthető teljes évek.)

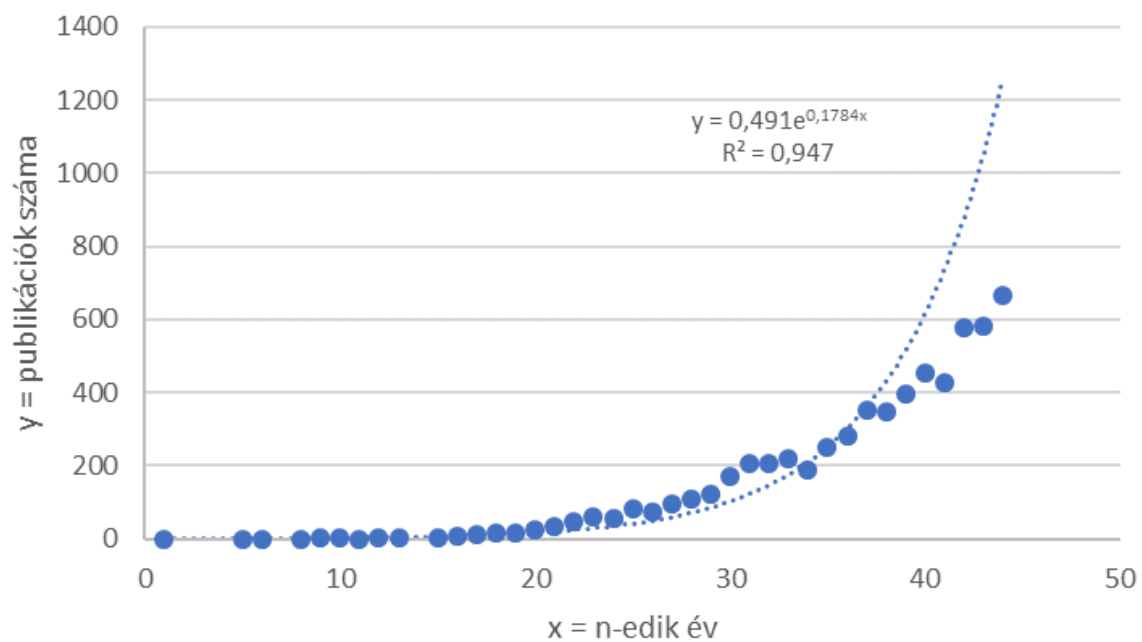
Ugyanezek az adatok EP és SM esetén rendre: 1922 – 2021 között vizsgált évek, a maximális érték (2018-ban): 12.850; 1968 – 2020 között vizsgált évek, a maximális érték (2018-ban): 7.112.

A polinomiális (SW) trend az R^2 alapján igen pontosan leírja a tapasztalati értékeink alakulását (12. ábra), azonban a trend exponenciálisnak is tekinthető.



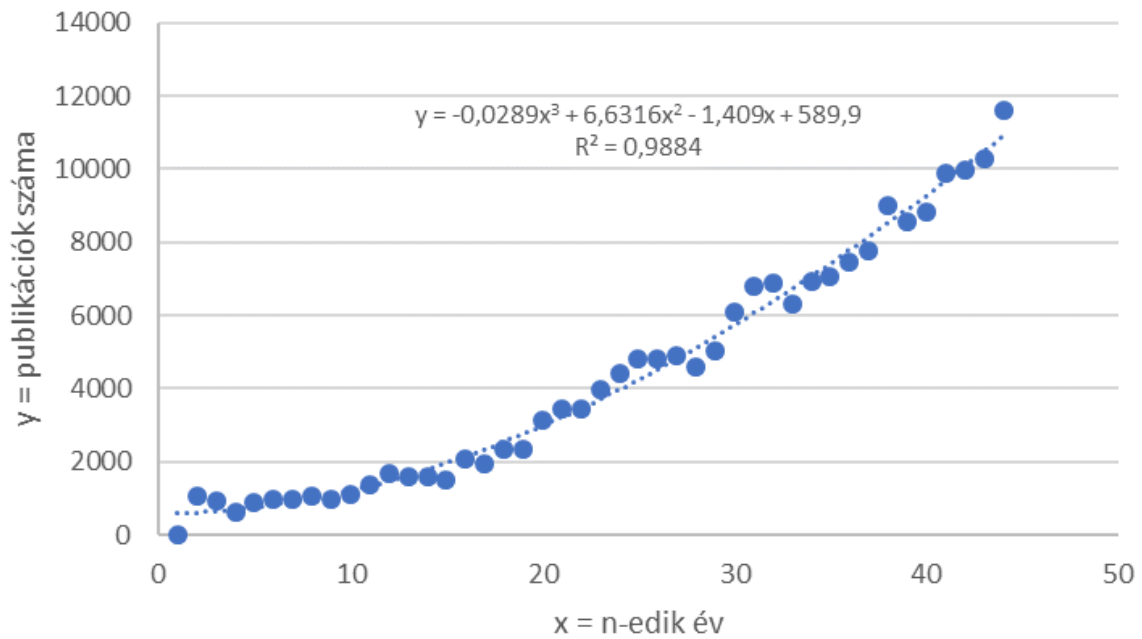
12. ábra Polinomiális trend (SW)

Az exponenciális trend (SW) R^2 értéke ugyan kisebb, azonban olyan közel van a maximális értékhez (1), hogy a trendet minden további nélkül tekinthetjük exponenciálisnak. (13. ábra)



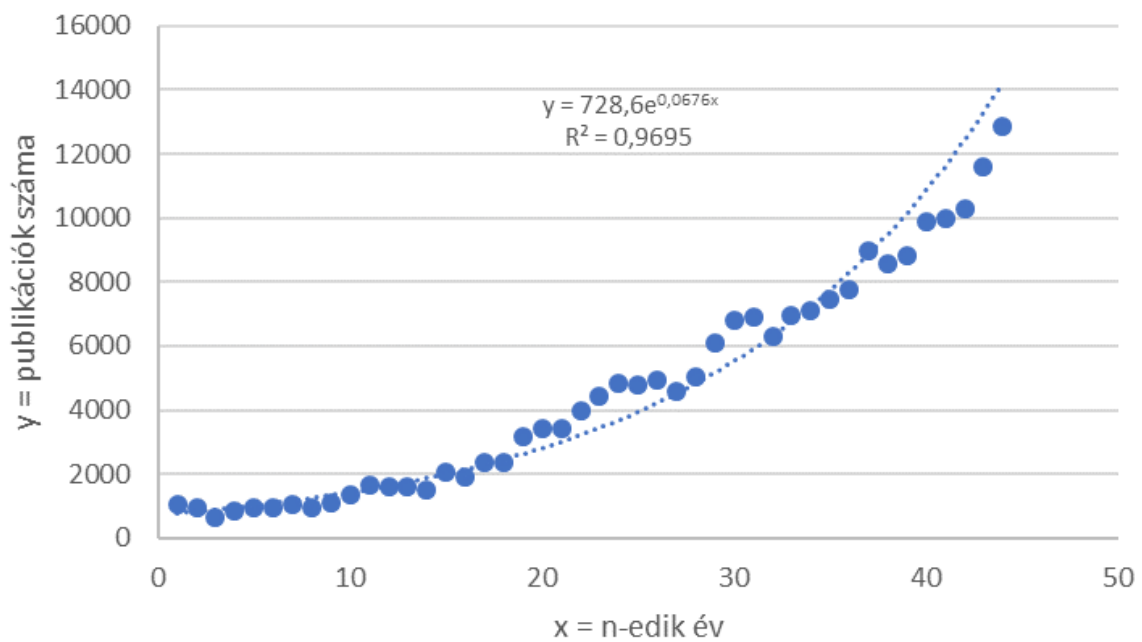
13. ábra Exponenciális trend (SW)

Az előző összefüggéshez hasonlóan alakul az EP polinomiális trendje is. (14. ábra)



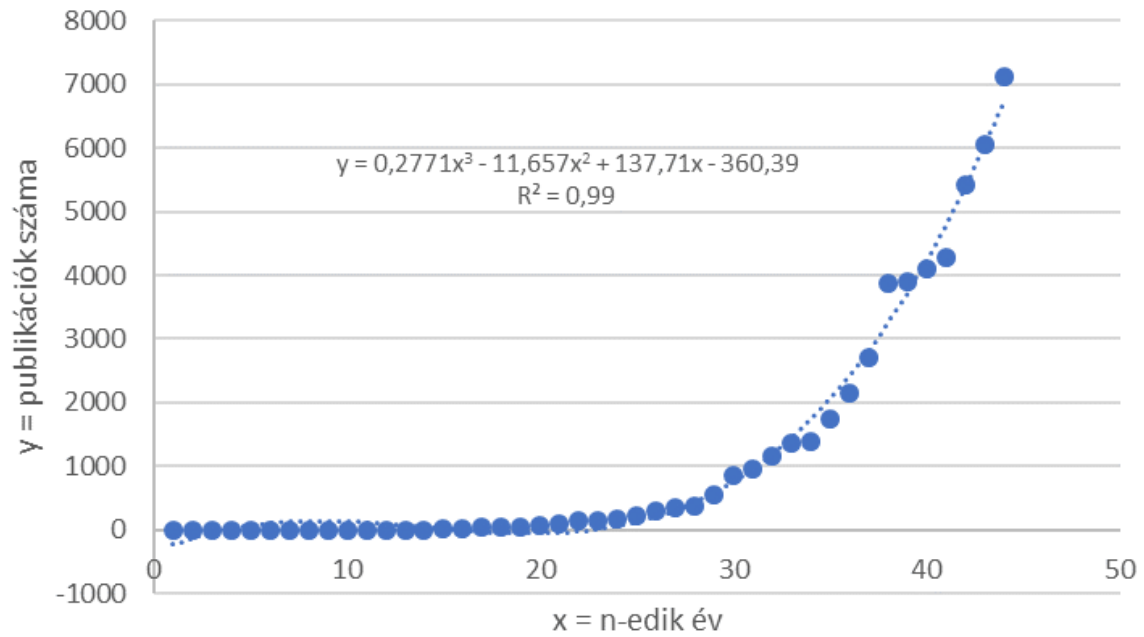
14. ábra Polinomiális trend (EP)

Az EP exponenciális trendje még az SW exponenciális trendjénél is pontosabban írja a tapasztalati idősor alakulását. (15. ábra)



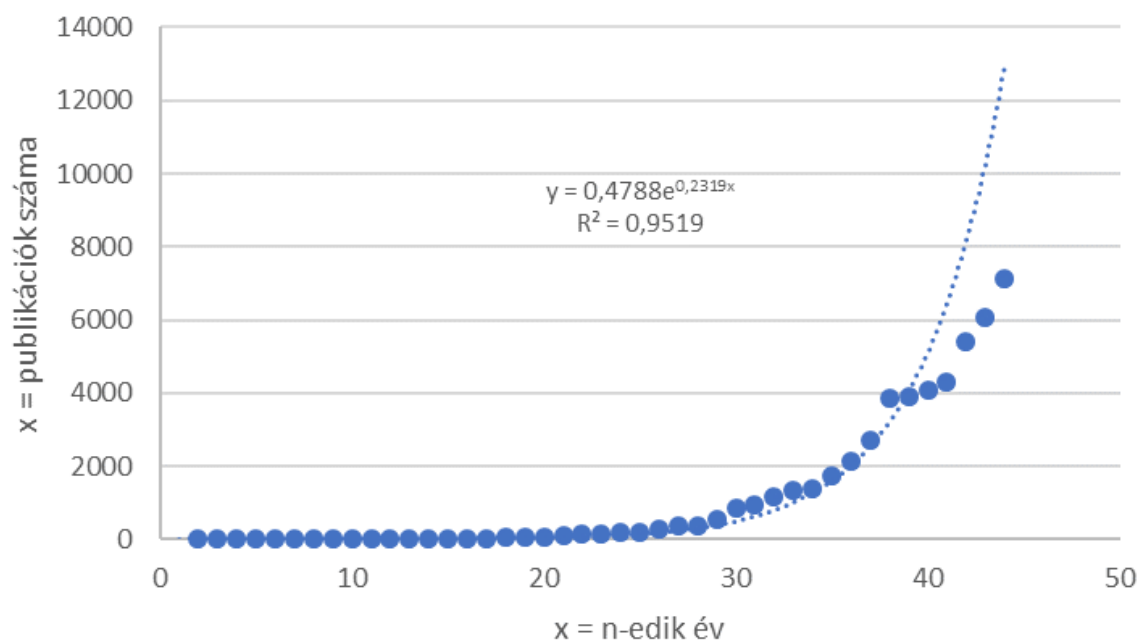
15. ábra Exponenciális trend (EP)

Nincs változás az SM polinomiális trendjénél sem. (16. ábra)



16. ábra Polinomiális trend (SM)

Ahogy az SM exponenciális trendje sem mutat lényeges eltérést az előzőkhez képest. (17. ábra)



17. ábra Exponenciális trend (SM)

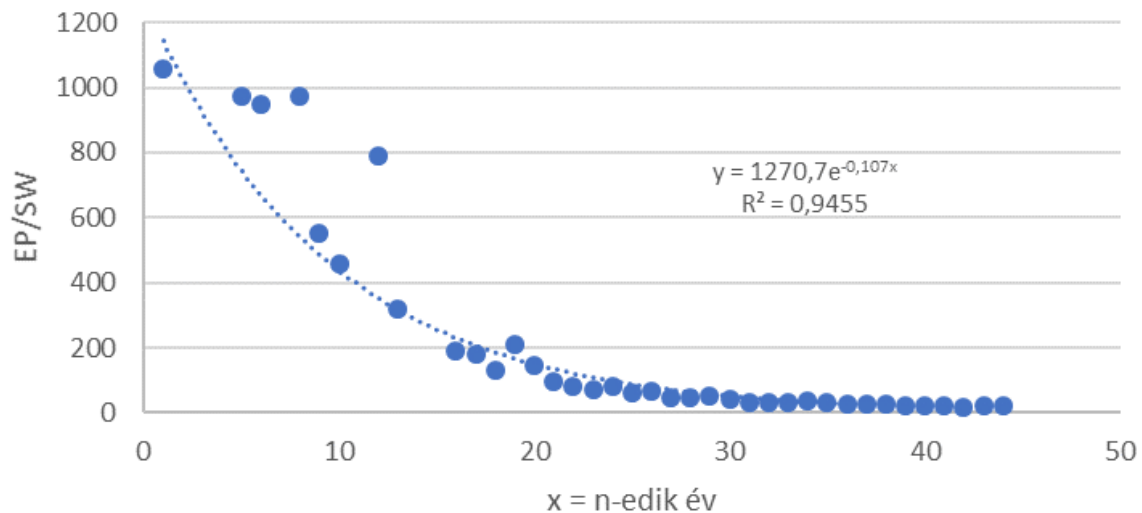
Három, nagyon pontosan a tapasztalati idősorra illeszkedő, lényegében azonos típusú trend esetén, csak az idősorok egymáshoz való viszonya adhat pontosabb képet pl. a fenntartható faipar helyzetéről.

7.2.1. Az eredmények értékelése

Bár formálisan a harmadfokú polinomiális egyenletek illeszkedtek jobban a tapasztalati értékeinkre, az adatok jellege alapján úgy véljük, hogy helyesebb, ha mindhárom esetben exponenciálisan növekvő trendről beszélünk. (Az R^2 mutató értéke exponenciális trend esetén is meglehetősen magas.)

Annak ellenére, hogy a kapott egyenletek összehasonlíthatók, s ábrázolásuk után különösen feltűnő a trendek közötti különbség, úgy találtuk, hogy a helyzetet az olvasó számára leginkább az jellemzi, ha évenként képezzük EP és SW, illetve SM és SW gyakoriságának hányadosait.

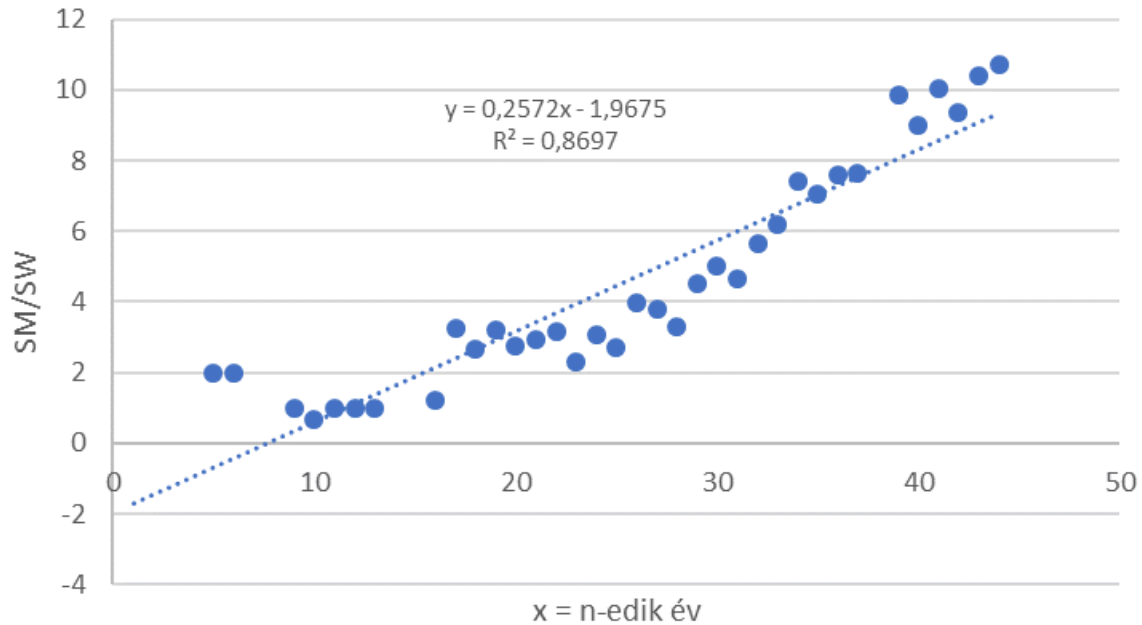
A zéró értékek és az outlierek – ebben az esetben, a lényeges adatvesztés minimális esélye miatt automatizmusokra bízott – kiszűrése után azt tapasztaltuk, hogy az első esetben (18. ábra) csökkenő exponenciális trend, a második esetben (19. ábra) növekvő lineáris trend tapasztalható. R^2 mutató értéke mindkét esetben magas.



18. ábra EP/SW arány zéró értékek és outlierek nélkül

Megállapítható, hogy a fenntartható faiparhoz kapcsolódó publikációk száma exponenciálisan nőtt, de ez a másik két kulcsszóval ellátott – s még nagyon sok egyéb területet érintő – publikációkra is igaz.

A környezetvédelmi publikációk „főlénye” a fenntartható faiparhoz kapcsolódó cikkek gyakoriságához képest csökkent (azaz a fenntartható faiparhoz kapcsolódó cikkek aránya nőtt), de a fenntartható nyersanyagok területén évről évre több publikáció születik, mint a fenntartható faiparban. Mivel a fenntartható faipar ma már szinte az összes faipari cikk általános hívószava, a felemás folyamat a faiparra *en bloc* is érvényes.



19. ábra SM/SW arány zéró értékek és outlierok nélkül

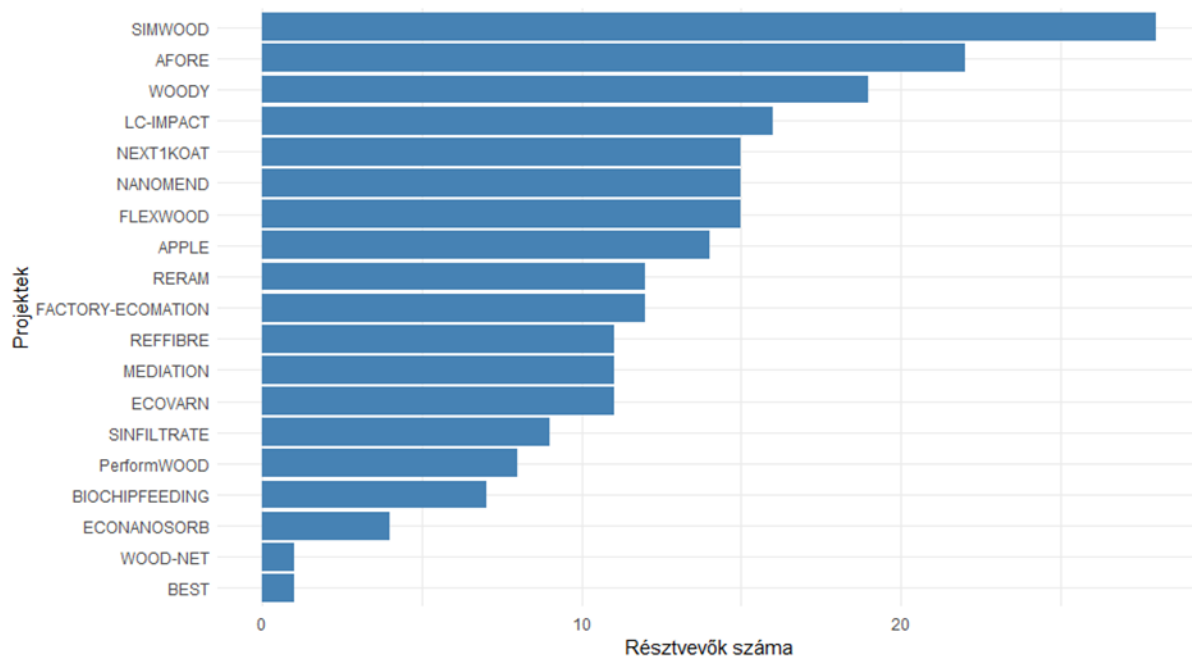
Úgy tűnik, hogy a fenntarthatósági szempontból kulcsfontosságú terület felfedezése egyelőre várat magára. Ebben minden bizonnyal szerepet játszanak azok a tényezők, amelyek az SW4SW stratégiai céljainak megvalósulását is akadályozzák.

7.3. Az FP7 és a H2020 faipari projektjeinek statisztikai elemzésének eredménye

7.3.1. A projektek általános jellemzői

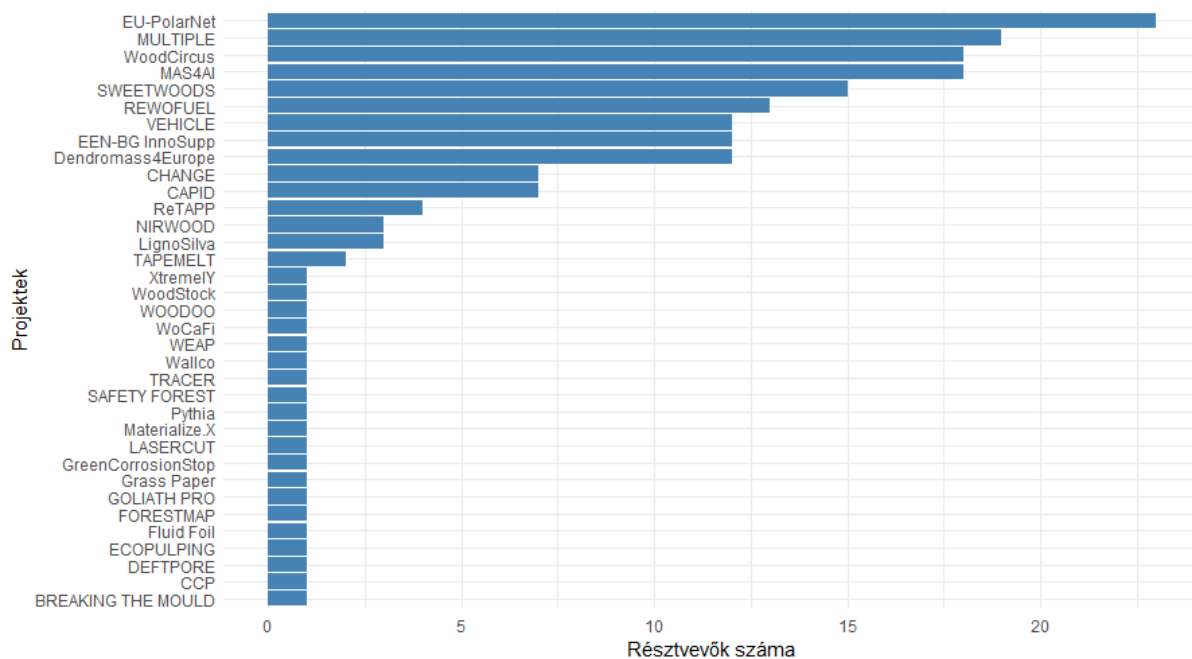
A 19 FP7(fa) projektben a SIMWOOD projekt résztvevőinek száma 28 volt, ezzel ez a projekt bizonyult a „legnépesebbnek”. Szorosan követi az AFORE és a WOODY 22, illetve 19 résztvevővel. A projektek között csak két egyrésztvevős található, tehát az együttműködésre épülő projektek voltak a jellemzők. (A résztvevők alatt természetesen résztvevő intézményeket, cégeket, szervezeteket értünk.)

A 36 H2020(fa) projekt közül az EU-PolarNet, a MULTIPLE és a MAS4AI projektekben vettek részt a legtöbben, rendre 23, 19 és 18 résztvevővel. Ezekben a projektekben sokkal gyakoribb a nem konzorciumként működő projekt. 20 olyan projektet láthatunk, amelyben csak egyetlen projektrésztvevő van. Kérdéses, hogy az együttműködés csökkenése a két projektidőszak között járt-e pozitív hatásokkal.



20. ábra Az FP7(fa) projektek résztvevőinek száma

A 21. ábra 35 projektet tartalmaz, mivel egy projekt (NIRWOOD) azonos névvel, de más projektazonosítóval a támogatási időszak alatt újraindult, ezért itt a két projektet összevontuk.



21. ábra A H2020(fa) projektek résztvevőinek száma

Az FP7(fa) és a H2020(fa) projektek költségeinek jellemzőit a 4. táblázat mutatja.

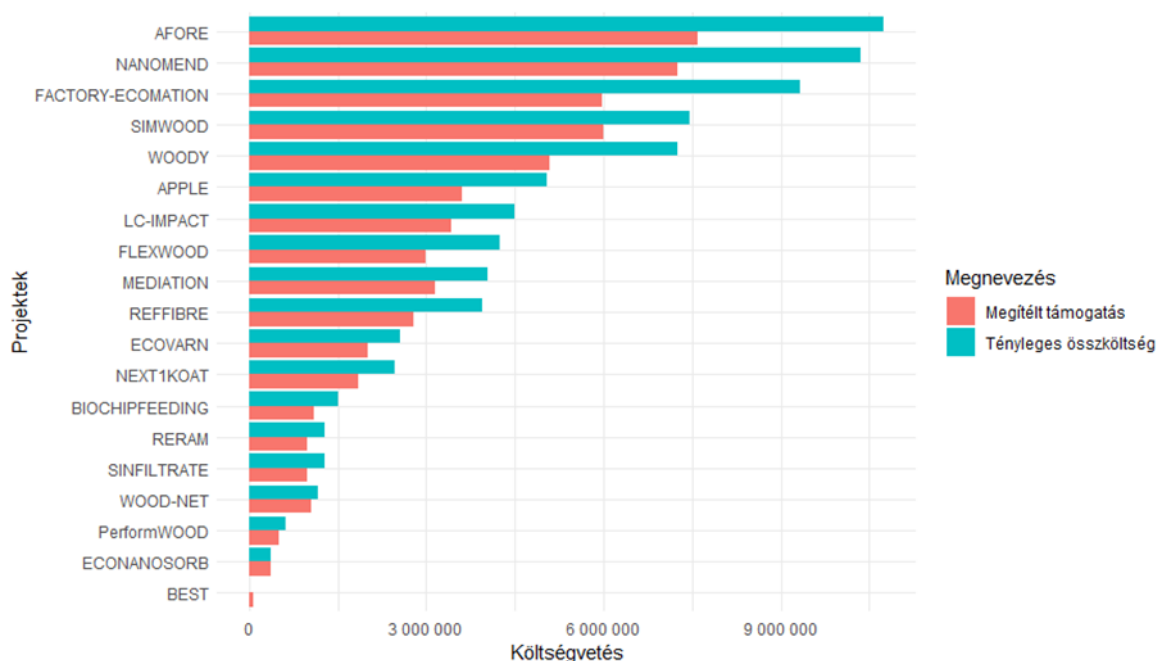
Miközben a faipari programok száma, összköltségvetése és az összes támogatás jelentősen nőtt, a közösségi támogatás és a teljes költségvetés projektenkénti átlaga jelentősen, mediánja drámaian csökkent.

4. táblázat A faipari projektek támogatása (euróban)

Mutató (támogatás)	FP7(fa)	H2020(fa)	Változás
Összes megítélt közösségi támogatás	56.807.620	83.568.704	+47,11%
Összes költség	78.159.801	128.100.494	+63,90%
Közösségi támogatás átlaga	2.989.875	2.321.353	-22,36%
Közösségi támogatás mediánja	2.785.782	331.446	-88,10%
Teljes költségvetés átlaga	4.113.674	3.558.347	-13,50%
Teljes költségvetés mediánja	3.937.100	331.446	-91,58%

A medián abszolút értékben és az átlaghoz képest is drámai csökkenésének okát jól mutatja a 22. ábra és a 23. ábra. A H2020 faipari projektjeiben ugyanis nagyon sok egyszerreplős projekt alacsony támogatást kapott, míg egyes projektek a korábbiakhoz képest magas támogatásban részesültek.

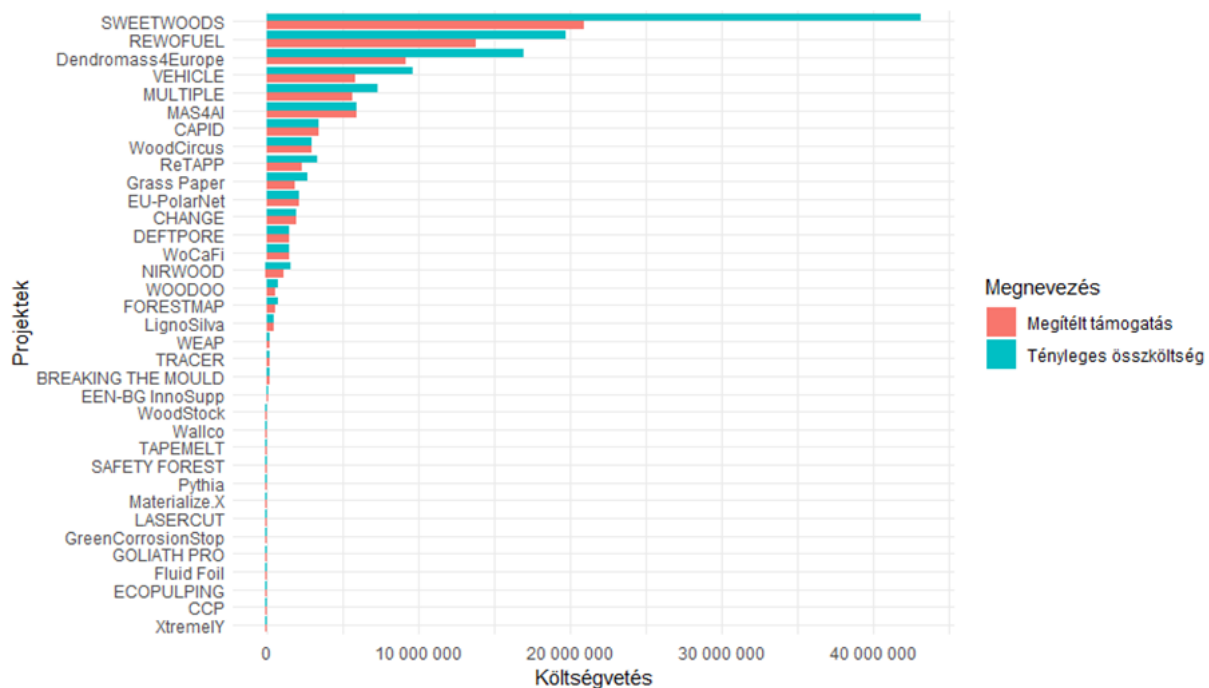
Az FP7(fa) projektjei közül a legtöbb támogatást az AFORE (7.600.000 euró), a NANOMEND (7.250.000 euró) és a SIMWOOD (5.990.311 euró) projektnek ítélték meg, s a legkevesebb megítélt támogatás 79.200 euró volt. (Ez utóbbit egyébként a projekt megghiúsulása miatt a projekt végül nem vette igénybe.) Ezt a projektet leszámítva a legalacsonyabb tíz támogatási összeg a 378.100 – 2.993.300 eurós sávba esett. Az összköltség terén hasonló a sorrend. A legmagasabb költségvetésű projekt az AFORE (10.746.919 euró) volt, a második helyen pedig a NANOMEND (10.357.517 euró) szerepel. A SIMWOOD (7.465.592 euró) azonban itt már a negyedik helyre szorult a FACTORY-ECOMATION (9.337.750 euró) mögött. A korábban említett, végül 0 euró összköltségű projekttől eltekintve a tíz legszerényebb költségvetésű projekt összköltsége a 378.100 – 4.050.579 eurós sávba esett.



22. ábra Az FP7(fa) projekteknek megítélt közösségi támogatás és a tényleges összköltség (euróban)

A H2020(fa) projektjei közül a legnagyobb EU-támogatást a SWEETWOODS (20.959.745 euró), a REWOFUEL (13.856.302 euró) és a Dendromass4Europe (9.166.453 euró) projekteknek

ítélték meg. A projektek teljes költsége alapján ugyanez a sorrend, rendre 43.233.731, 19.791.557 és 16.971.581 euró összköltséggel.



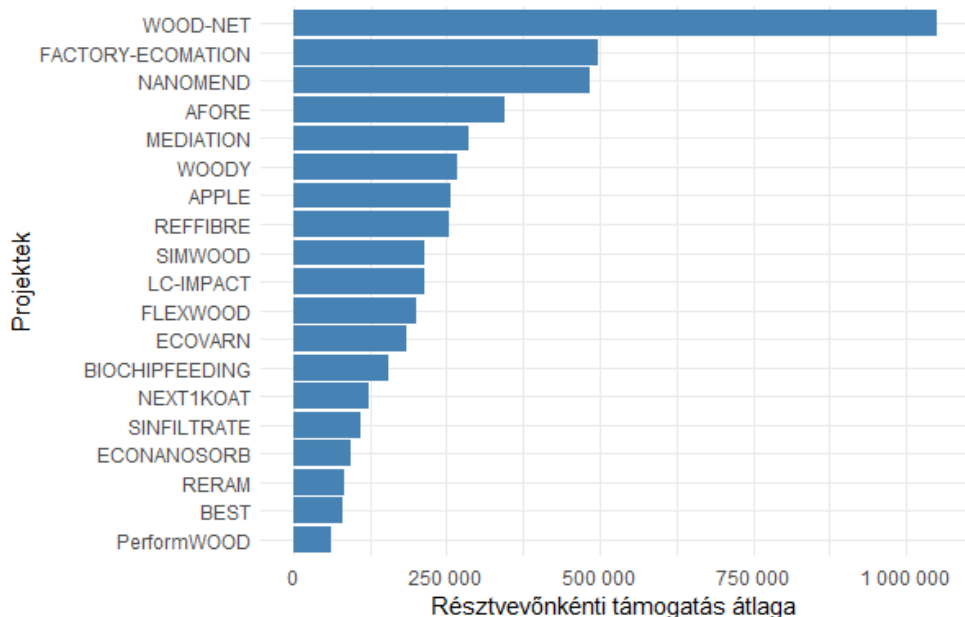
23. ábra A H2020(fa) projekteknek megítélt közösségi támogatás és a tényleges összköltség (euróban)

A legkisebb közösségi támogatást kapó tíz projektnek 50.000 euró támogatást ítélték meg, összköltségeik pedig egy esetet leszámítva – ahol ez 50.472 euró volt – egyenként 71.429 eurót tettek ki.

Néhány magas költségvetésű (magas összeggel támogatott) projekt miatt az átlagos támogatás tehát jóval kisebb mértékben csökkent, mint a medián.

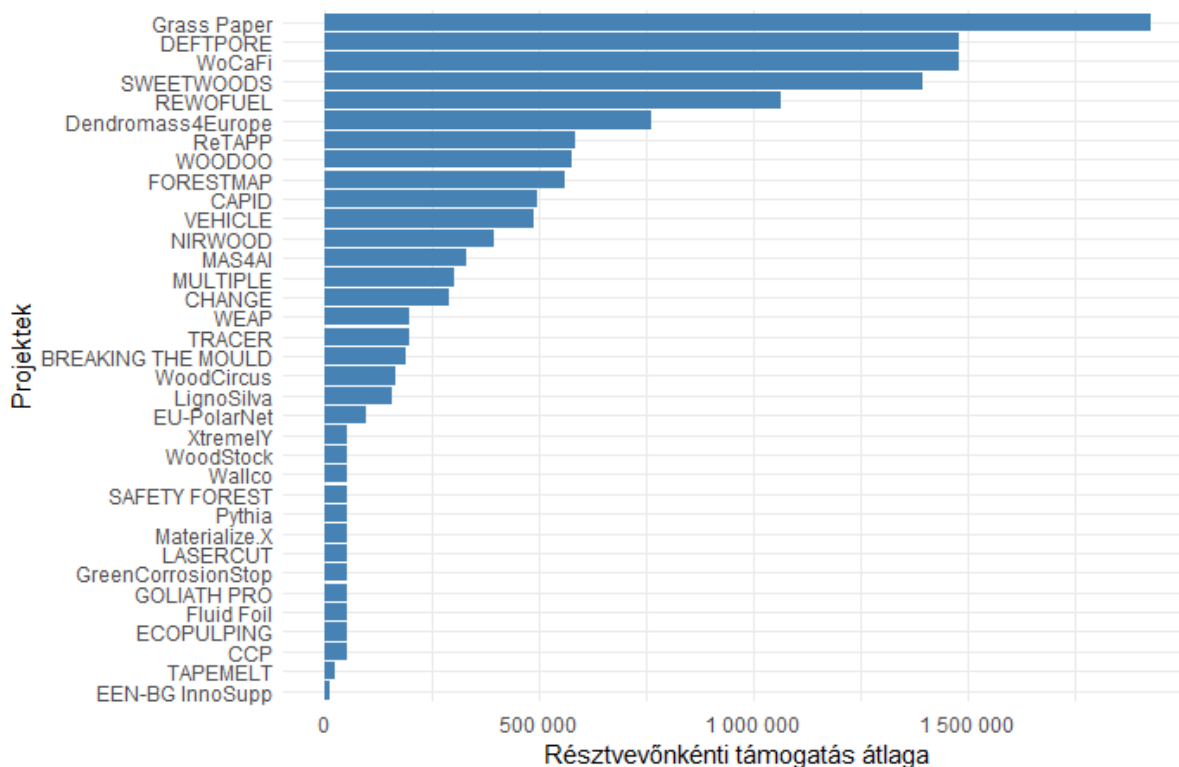
Az FP7(fa) és a H2020(fa) projektek résztvevőinek átlagtámogatása – a résztvevőnkénti teljes költségről az adatközlés erősen hiányos, ezért ezt nem tudtuk érdemben vizsgálni – jelentős különbségeket mutat.

Az FP7(fa) adatsoron belül átlagosan a legnagyobb támogatást a WOOD-NET (1.049.320 euró), a FACTORY-ECOMATION (497.646 euró) és a NANOMEND (483.333 euró) résztvevői kapták. Az átlagtámogatás négy projektben nem érte el a 100.000 eurót, a legalacsonyabb átlagos támogatás 61.977 eurót tett ki. (24. ábra)



24. ábra A résztvevők átlagtámogatása az FP7(fa) projektekben (euróban)

A H2020(fa) projektjei közül az átlagos legnagyobb támogatásban a Grass Paper (1.930.250 euró), a DEFTPORE (1.482.862 euró) és a WoCaFi (1.481.008 euró) résztvevői részesültek. 15 projektben az átlagos támogatás nem érte el a 100.000 eurót, a legkisebb átlagos támogatási összeg 9.827 euró volt. (25. ábra)



25. ábra A résztvevők átlagtámogatása az H2020(fa) projektekben (euróban)

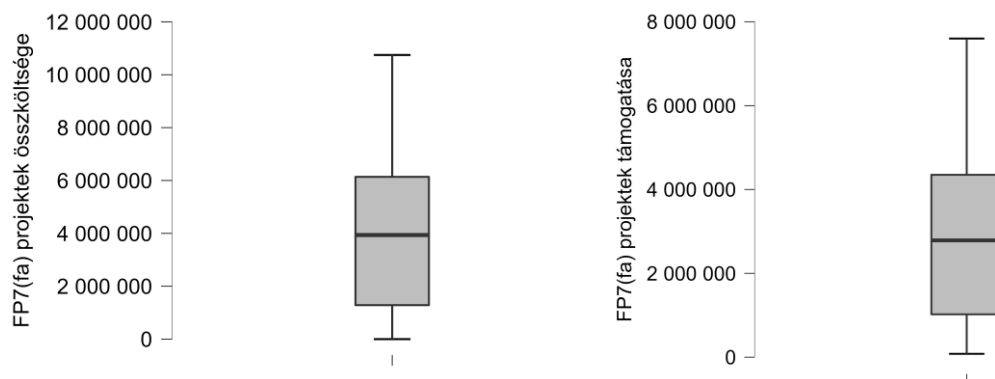
Az átlagos projektenkénti támogatás adatait a két programban az 5. táblázat mutatja. Jól látható, hogy az FP7(fa) adatsorhoz képest a H2020(fa) adatsorban a projektenkénti

átlagtámogatásban is magasabb átlagot, de kisebb mediánt, s emellett nagyobb szórást mutat. Tehát a támogatottak átlagos támogatása alapján a projektek támogatása kevésbé volt egyenletes, mint korábban.

5. táblázat A projektenkénti átlagtámogatás fő jellemzői

Mutató	FP7(fa)	H2020(fa)
Projektek száma	19	35(36)
Módusz	61.976,87	50.000,00
Medián	213.371,87	187.866,00
Átlag	261.034,63	393.547,06
Szórás	227.745,52	499.554,43
MAD	89.305,21	137.866,00
IQR	159.784,36	477.576,31
Ferdeség	2,54	1,73
Csúcsosság	7,89	2,27
Minimum	61.976,87	9.827,42
Maximum	$1,049 \times 10^6$	$1,930 \times 10^6$

A 26. ábra alapján az FP7(fa) projektek költségvetésében és közösségi támogatási összegében nincs outlier.



26. ábra Az FP7(fa) projektek összköltségének és közösségi támogatásának boxdiagramja (euróban)

FP7(fa) projektjeinek összköltsége esetén $I(8)(9)(10)$ rendre $[-5.994.594; 13.417.701]$, $[-7.883.534; 15.757.734]$, $[-9.391.745; 16.814.852]$.

FP7(fa) projektjeinek közösségi támogatása esetén $I(8)(9)(10)$ rendre $[-3.970.221; 9.341.523]$, $[-5.183.501; 10.755.065]$, $[-6.299.776; 11.671.078]$.

Mindkét adatsor értékei az elfogadási tartományba esnek mindhárom számítási mód mellett, ezért megállapíthatjuk, hogy egyik adatsorban sincsenek outlierok.

Az elvégzett Shapiro-Wilk próba p értéke az 5%-os szignifikanciaszint körül mozgott, de mivel a Kolmogorov-Szmirnov normalitási próba értékei 0-hoz közeledek, s kisebb mintáknál ez utóbbi próba gyakran megbízhatóbb, a normális eloszlásra vonatkozó feltételezést mindkét

adatsornál elvetjük, ezért Pearson-féle korrelációt nem számolunk, de Spearman-féle rangkorreláció számolható. (6. táblázat)

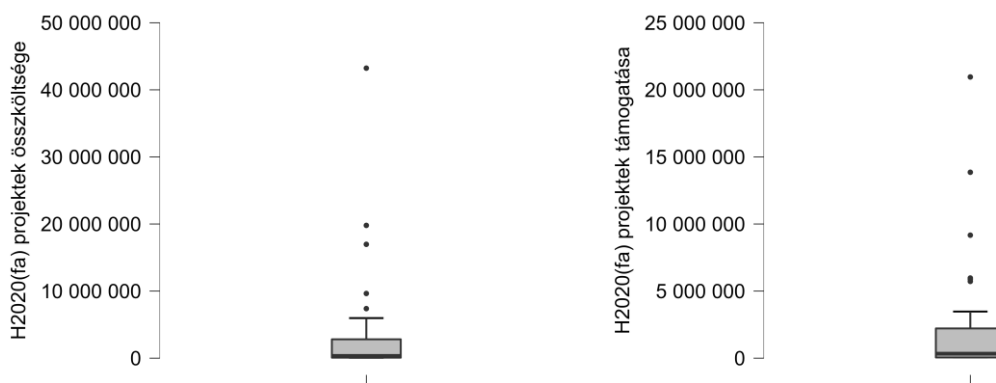
6. táblázat Az FP7(fa) projektek összköltségének és közösségi támogatásának normalitáspróbája

Mutató	FP7(fa)	
	Összköltség	Támogatás
Shapiro-Wilk (W)	0,899	0,907
P (Shapiro-Wilk)	0,046	0,065
Kolmogorov-Szmirnov (D)	0,947	1
P (Kolmogorov-Szmirnov)	$2,22 \times 10^{-16}$	$2,22 \times 10^{-16}$

A Mahalanobis-távolság minden esetben kisebb, mint 13,82. (A maximális érték 9,70.) Tehát a Mahalanobis-távolság alapján nincs többdimenziós outlier. A Cook-távolság egyetlen egy esetben sem nagyobb, mint egy. (A legnagyobb érték 0,87.) Többdimenziós outlier tehát nem akadályozza a korrelációs számítás elvégzését.

A Spearman-féle rangkorrelációs együttható értéke ($\rho = 0,99$) rendkívül szoros korrelációt jelez. Bár alapvetően sokasági statisztikáról van szó, ha valamilyen sokaságot feltételezünk még a vizsgált adatok mögött, akkor p értéke ($8,418 \times 10^{-6}$) alapján a függetlenségre vonatkozó hipotézist elvethetjük.

A boxdiagram alapján a H2020(fa) projektek összköltségei és közösségi támogatása jelentős számú outliert hordoz magában (27. ábra).



27. ábra A H2020(fa) projektek összköltségének és közösségi támogatásának boxdiagramja (euróban)

A H2020(fa) projektjeinek összköltsége esetén $I(8)(9)(10)$ rendre $[-4.038.848; 6.921.890]$, $[-871.664; 1.534.557]$, $[-5.956.977; 8.840.019]$.

A H2020(fa) projektjeinek közösségi támogatása esetén $I(8)(9)(10)$ rendre $[-3.198.230; 5.463.717]$, $[-920.371; 1.583.264]$, $[-4.714.071; 11.671.078]$.

A H2020(fa) projektjeinek összköltségének adatai között (8) alapján öt, (9) alapján 13, (10) alapján négy outlier található. A H2020(fa) projektjeinek közösségi támogatási adatai között (8) alapján hat, (9) alapján 12, (10) alapján három outlier található. A nagyszámú outlier

kizárása esetünkben felesleges adatvesztéssel járna, mivel kialakulásuk egyértelműen a számos „nagy” projekt mellett megjelenő kis, egyéni projektekre vezethető vissza. Tehát az anomáliát nem valamilyen hiba vagy extremitás okozza, hanem a nem eléggé konzisztens adatok. Viszont a nagyszámú, outlierként viselkedő adat miatt a további számítások eredményét óvatosan kell fogadnunk.

A 7. táblázat alapján egyik adatsor sem normális eloszlású, ezért Pearson-féle korrelációs együttható nem számítható.

7. táblázat A H2020(fa) projektek összköltségének és közösségi támogatásának normalitáspróbája

Mutató	H2020(fa)	
	Összköltség	Támogatás
Shapiro-Wilk (W)	0,479	0,582
P (Shapiro-Wilk)	$3,697 \times 10^{-10}$	$6,096 \times 10^{-9}$
Kolmogorov-Szmirnov (D)	1	1
P (Kolmogorov-Szmirnov)	$< 2,2 \times 10^{-16}$	$< 2,2 \times 10^{-16}$

A Mahalanobis-távolság két esetben nagyobb, mint 13,82. (15,26 és 29,41) Tehát a Mahalanobis-távolság alapján két adatszám többdimenziós outlier. A Cook-távolság ugyanebben a két esetben nagyobb, mint egy (1,05 és 22,34). Mivel a két adatszám mindkét adata külön-külön is outlier volt, ezt a két adatszámot nem vettük figyelembe a számítás során. Talán ennyi adat elhagyása nem okoz jelentős információvesztést.

Az outlierektől megtisztított adatok között a Spearman-féle korrelációs együttható értéke igen magas ($\rho = 0,97$). Bár az adatokat itt is célszerűbb sokaságnak tekinteni, p értékét itt is kiszámoltuk ($< 2,2 \times 10^{-16}$). A nullához közeli érték alapján a függetlenségre vonatkozó nullhipotézis elvethető.

Az egyes projektek összköltsége és közösségi támogatása között tehát az FP7(fa) és a H2020(fa) projektekben igen szoros kapcsolat van, ezért az 5. szakmai hipotézist, miszerint a projektek támogatása és teljes költségvetése között szoros összefüggés mutatkozik, nem vesszük el.

A 19 FP7(fa) projektben minden koordinátor különböző. A H2020(fa) projektekben 32 koordinátor egy projektet vezet, míg két koordinátor kettőt. Egyetlen olyan koordinátor van a vizsgált faipari projektekben, amelyik mindkét támogatási programban koordinátor volt. A központi pozíciók „halmazása” tehát egyáltalán nem jellemző.

A koordinátorok és a résztvevők átlagos közösségi támogatása között jelentős különbség mutatkozik, de ez minden bizonnyal a feladatkörök és a felelősség különbségeire vezethető vissza. Feltűnő, hogy a H2020(fa) projektekben milyen nagy különbség tapasztalható a koordinátorok támogatásának átlaga és mediánja között, ez néhány kiugró támogatási értékre vezethető vissza. (8. táblázat)

8. táblázat Koordinátorok és résztvevők által kapott közösségi támogatás (euró)

Mutató	FP7(fa)	H2020(fa)
Koordinátorok átlagos támogatása	609.953	1.034.169
Koordinátorok támogatásának mediánja	572.978	195.455
Résztvevők átlagos támogatása	213.295	343.249
Résztvevők támogatásának mediánja	172.185	217.556

A 9. táblázat alapján megállapítható, hogy nem jellemző, hogy akár a koordinátorok, akár a résztvevők több projektben vettek volna részt.

9. táblázat A résztvevőkre jutó projektrésztvételek száma

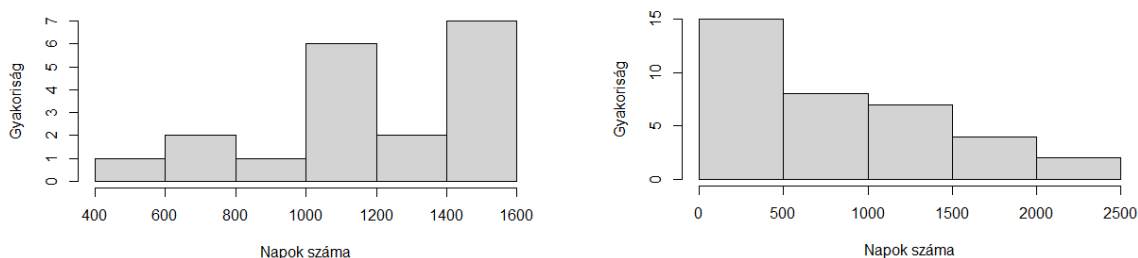
Mutató	FP7(fa)	H2020(fa)
Átlagos projektrésztétel	1,095	1,119
<i>Koordinátorok átlagos projektrésztétele</i>		
Koordinátorként	1	1,059
Összesen	1,421	1,265
Koordinátor és résztvevői szerepkört betöltők	5	5

A projektek megvalósítási ideje a két támogatási időszak alatt jelentősen csökkent, nominálisan az egy napra jutó támogatás nőtt, reálértéken lényegében szinten maradt (10. táblázat).

10. táblázat A projektek megvalósítási ideje

Mutató	FP7(fa)	H2020(fa)
Fizikai megvalósítás átlaga (nap)	1.167	791
Fizikai megvalósítás átlaga (medián)	1.095	729
Egy napra eső támogatás projektenként (euró)	2.562	2.934

A megvalósítási idők gyakoriságának eloszlása jelentős különbséget mutat a két időszak között. Míg az FP7(fa) projektekben az átlagosnál hosszabb, a H2020(fa) projektekben inkább az átlagosnál rövidebb ideig tartó projektek domináltak (28. ábra).



28. ábra Az FP7(fa) és a H2020(fa) projektek fizikai megvalósítási idejének eloszlása

Az önrész átlagos mértéke a két időszak között valamelyest nőtt (27,32%-ról 34,76%-ra), viszont az önrészt is tartalmazó projektek aránya csökkent (89,47%-ról 69,44%-ra).

Röviden ejtsünk szót a magyar részvételről is! Bár tartalmilag több program is kapcsolódott a faiparhoz, a „woodworking” SciVoc-kód mindkét támogatási programban 1 – 1 magyarországi résztvevő projektjénél található meg. Az FP7(fa) projektekben a szentendrei Regional

Environmental Center for Central and Eastern Europe vett részt. Ez a részvételek 0,43%-át, a közösségi támogatásoknak pedig (158.785 euróval) 0,28%-át teszi ki. A H2020(fa) projektekben a soproni Ökoforestino Kft. vett részt, amely a részvételek 0,53%-ának, (428.925 euróval) a támogatások 0,51%-nak felel meg.

7.3.2. Az egyes országok programonkénti részvételének elemzése

A faipari projektekben nemcsak európai uniós országok vettek részt. Az FP7(fa) résztvevők 91,77%-a volt EU-07 tag, s ők a támogatások 93,34%-át kapták, a H2020(fa) résztvevők között az EU-13 tagok 97,34%-ot tettek ki, s a közösségi támogatások 98,51%-ában részesültek.

A 11. táblázat értékei alapján az országonkénti részvétel gyakorisága a két támogatási időszakban igen sok hasonlóságot mutat.

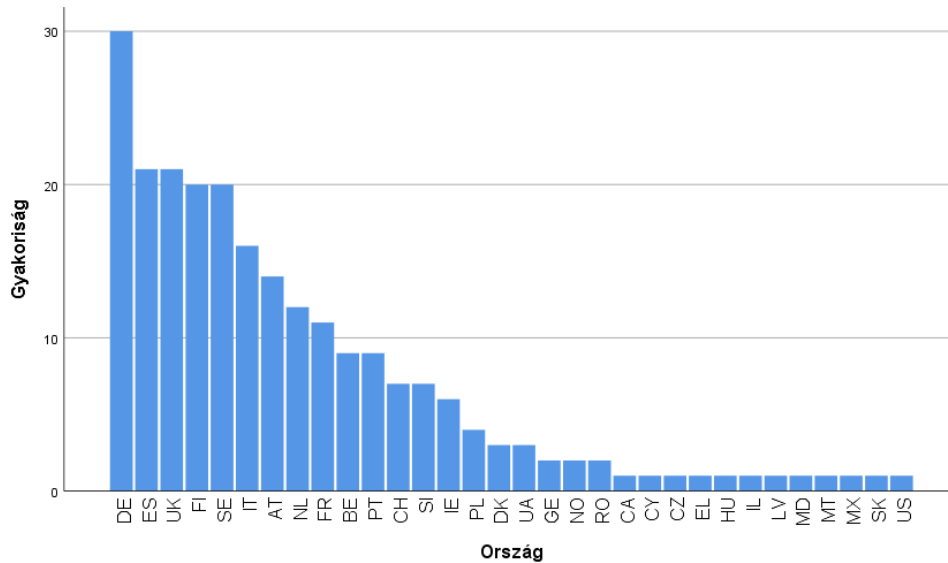
11. táblázat Országonkénti részvételi gyakoriság leíró statisztikája

Mutató	FP7(fa)	H2020(fa)
Országok száma	32	26
Módusz	1	1
Medián	3	5
Átlag	7,22	7,23
Szórás	8,00	7,07
MAD	2	4
IQR	10,25	11
Ferdeség	1,27	1,44
Csúcsosság	0,74	2,20
Minimum	1	0
Maximum	30	19
Összes részvétel	231	188

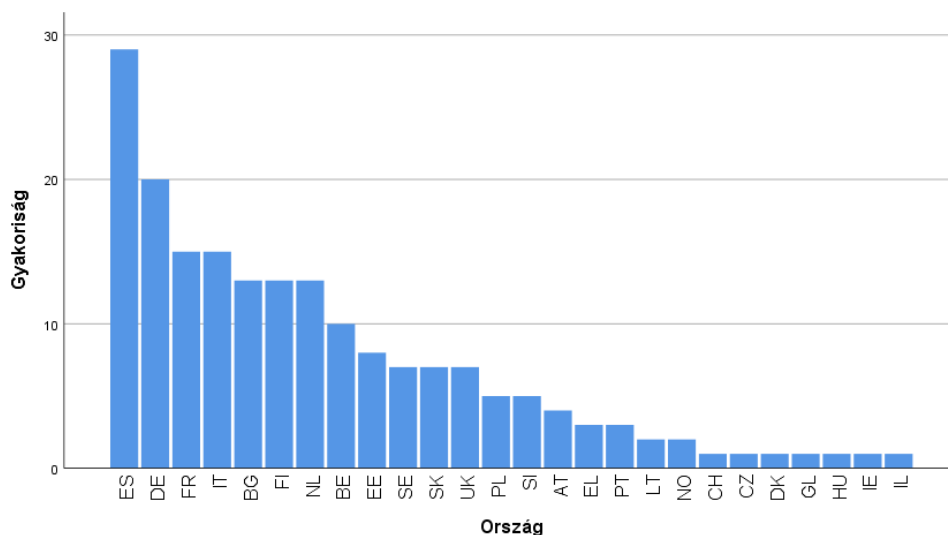
Az FP7 program faipari projektjeiben még részt vett országok közül a H2020 faipari projektjeiben 10 már nem vett részt, de a H2020-ban megjelent négy olyan ország, amely az FP7 faipari projektjeiben még nem volt résztvevő. A két országcsoportot az összehasonlító vizsgálatok során kivettük az elemzett adatsorokból. Az adatvesztés miatt a következtetéseket óvatosan kell kezelni.

0-hoz közelítő ferdeség és csúcsosság esetén felmerülhet a normalitás, de itt nem ez a helyzet egyik program esetén sem.

Az 29. és a 30. ábra alapján nagyon nehéz bármit kijelenteni arról, hogy a két támogatási program faipari projektjeiben a részvétel alapján történt-e áttrendeződés az egyes országok között.



29. ábra Projektrészvételek gyakorisága országonként az FP7(fa) adattáblában



30. ábra Projektrészvételek gyakorisága országonként a H2020(fa) adattáblában

Az összehasonlító vizsgálatok során csak a mindkét programban részt vett országokat vizsgáltuk. Az adatvesztés miatt a következtetéseket óvatosan kell kezelni.

A 31. ábra alapján a H2020(fa) részvételi gyakoriságai között egy érték outliernek tűnik.

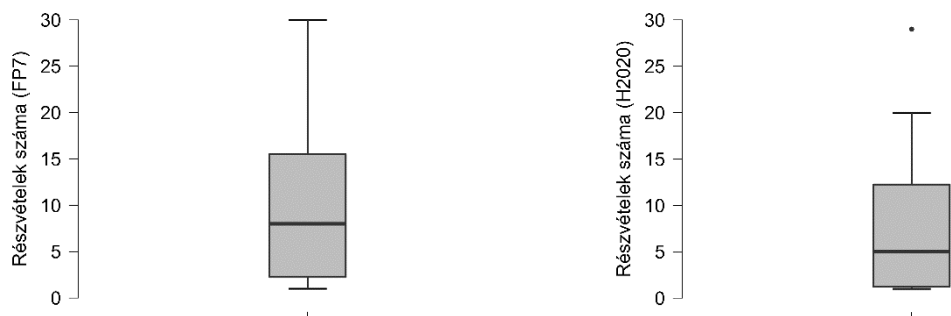
FP7(fa) esetén $I(8)(9)(10)$ rendre $[-17,6; 35,4]$, $[-20,9; 36,9]$, $[-26,9; 44,7]$.

H2020(fa) esetén $I(8)(9)(10)$ rendre $[-15,2; 28,8]$, $[-12,8; 22,8]$, $[-23; 36,5]$.

FP7(fa) adatai minden esetben I tartományba esnek. H2020(fa) esetén viszont Spanyolország adata a három módszerből kettő szerint outlier, így Spanyolországot kivettük a mintából.

A Cramér-együttható értéke alapján ($C = 0,31$) az egyes országokhoz tartozó, a két programban tapasztalt projektrészvételek között közepesnél gyengébb, sztochasztikus kapcsolatot

számszerűsítettünk. Az outliereket jellemzően nem Cramér-mutató számítása előtt negligálják az adatok közül. Ha a leírt outliert töröljük, a Cramér-mutató értéke lényegében nem változik.



31. ábra Országokénti részvétel boxdiagramjai az FP7 és a H2020 programok faipari projektjeiben

A vizsgált adatok a Shapiro-Wilk és az egymintás Kolmogorov-Szmirnov normalitási próba alapján nem normális eloszlásúak (12. táblázat), mivel $p < 0,05$, ezt a 32. ábra is megerősíti. A normális eloszlást feltételező H_0 -t tehát elvetjük.

12. táblázat Országokénti részvételek normalitási próbái

Mutató	Részvételek (országoként)	
	FP7(fa)	H2020(fa)
Shapiro-Wilk (W)	0,89	0,86
P (Shapiro-Wilk)	0,02	0,01
Kolmogorov-Szmirnov (D)	0,84	0,84
P (Kolmogorov-Szmirnov)	0,00	0,00

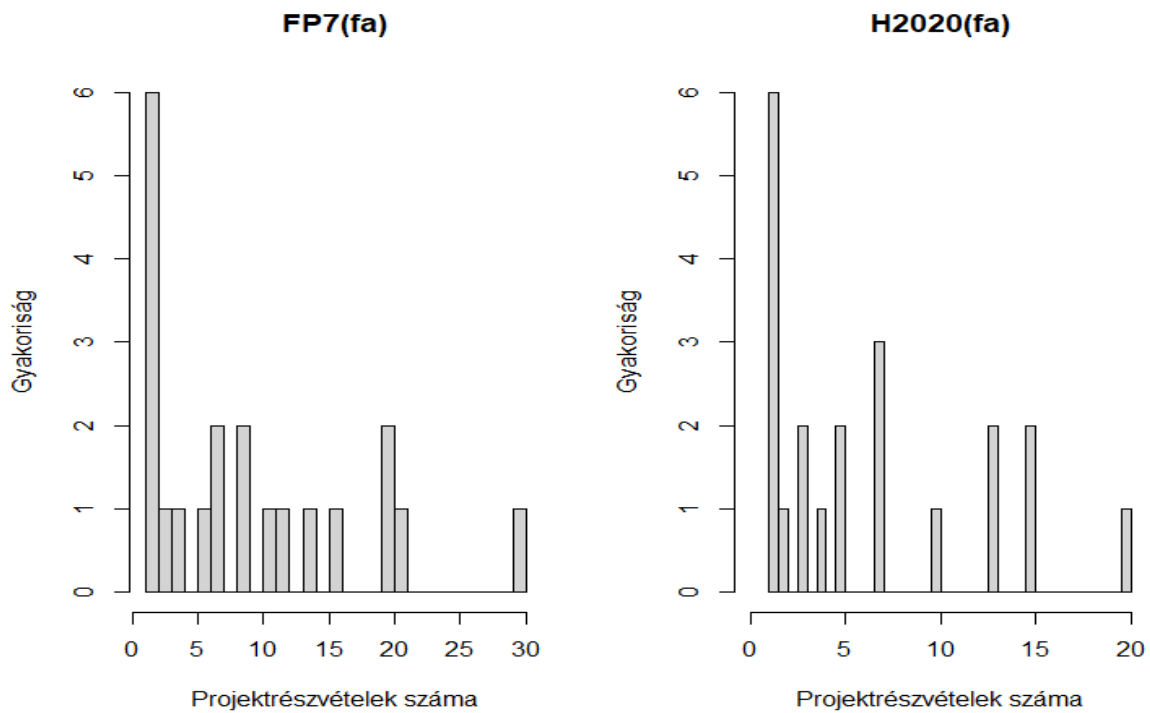
A normalitási próbák igen hasonló eredményei előrevetítik, hogy a normalitás nem teljesülése ellenére a két gyakorisági sor azonos eloszlást mutat (29. ábra és 30. ábra), illetve a két változó közötti összefüggés is valószínűsíthető (33. ábra).

A Mahalanobis-távolság minden esetben kisebb, mint 13,82. (A maximális érték is csak 6,82.) Tehát a Mahalanobis-távolság alapján nincs többdimenziós outlier. A Cook-távolság egyetlen esetben sem nagyobb, mint egy. (A legnagyobb érték 0,23.) Többdimenziós outlier tehát nem akadályozza a korrelációs számítás és a regresszió számítás elvégzését.

A Spearman-féle rangkorrelációs számítással és próbával viszonylag erős (közepesnél erősebb), pozitív korrelációt mértünk FP7(fa) és H2020(fa) országoként mért projektrészvételi gyakorisága között ($\rho = 0,719$). A $p = 0,00023 < \alpha$ érték alapján a függetlenségre vonatkozó $H_0(5)$ -t elutasítjuk. Ezért a 6. szakmai hipotézist, miszerint az FP7-ben és a H2020-ban az országokénti részvétel gyakorisága között van összefüggés, nem vetjük el.

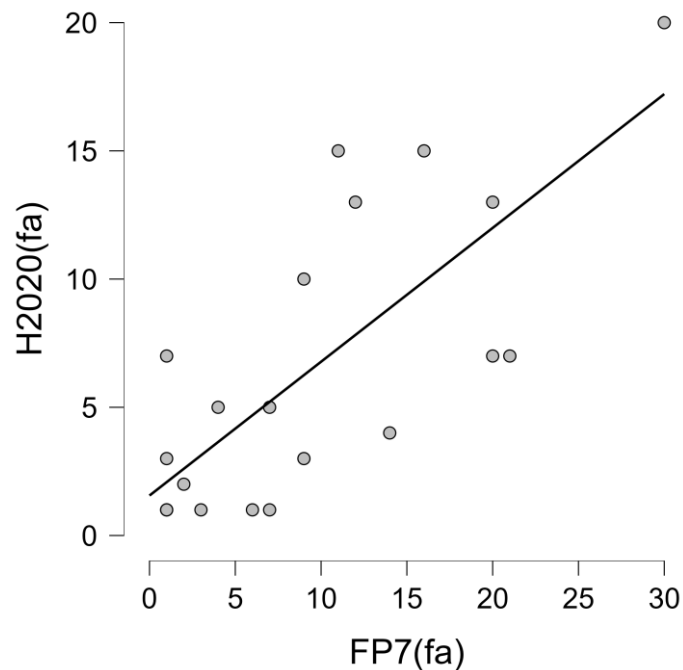
Ebben az esetben is felmerül, hogy érdemes-e a két adatsort mintának tekintenünk és egy nagyobb sokaságot vagy esetleg szupersokaságot feltételeznünk. Valószínűleg, ha csak nem feltételezzük ilyen sokaság(ok) létezését, p értékét nem feltétlenül kellett volna kiszámolnunk. Természetesen más a helyzet, ha pl. feltételezzük, hogy a projektrészvételek az adott programban a szakmai jellegű projektek és projektrészvételek csak egy részét reprezentálják, hiszen az FP7 és a H2020 nyilván nem fedi le az összes együttműködést. Statisztikai

megközelítéstől függ, hogy a kérdést korrelációs számítással vagy korrelációs teszttel vizsgáljuk-e, de a próba értékét nagyon óvatosan kell kezelni.



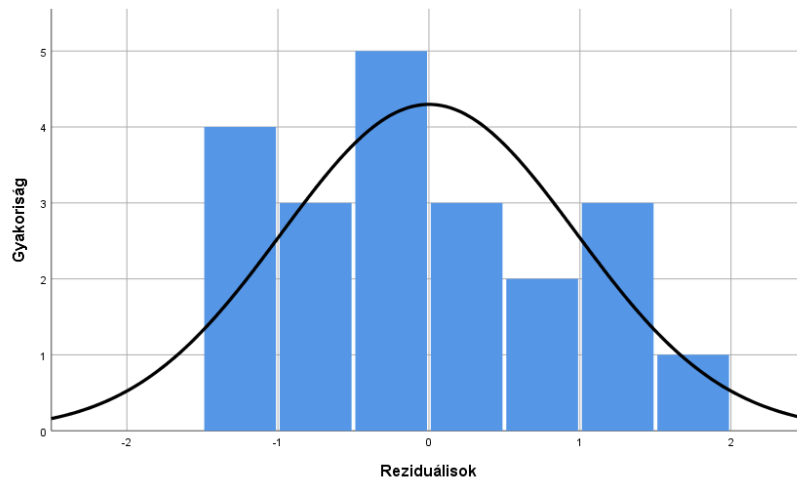
32. ábra Országokénti projektrészvételek gyakorisága

A regressziószámítás során az FP7(fa) országokénti részvételi gyakoriságait tekintjük független, a H2020(fa) országokénti részvételi gyakoriságait tekintjük függő változónak.



33. ábra Országokénti projektrészvételek számának lehetséges kapcsolata

A lineáris regresszió egyenes, ill. a becslőfüggvény paraméterei: $\hat{\beta}_0 = 1,558$, $\hat{\beta}_1 = 0,522$, $s = 3,915$, $s^2 = 15,324$, $s_{\hat{\beta}_0} = 1,309$, $s_{\hat{\beta}_1} = 0,106$.



34. ábra Az országok programonkénti projektrésztvételeinek száma közötti lineáris regresszió standardizált reziduálisainak eloszlása

Student-féle t -próba értékei: $t = 4,914$, $t_{0,975}(df = 22) = 2,074$. F -próba p értéke: 0,00. Bár a szakirodalomban a szabadságfokok számát a sorok számához képest gyakran csökkentik, mi az adattudományban általános módszer szerint a sorok számával azonosítjuk itt és a későbbiekben is (Gao & Jojic, 2016). (A végeredményben ez apró különbséget jelent.) Mivel a számított t érték a kritikus tartományban van, illetve $p < \alpha$, szignifikáns lineáris összefüggést feltételezünk a változók között. A lineáris kapcsolatot kizáró $H_0(12)$ -t ezért elutasítjuk.

A minta versus sokaság dichotómia a regresszió vizsgálatnál is felmerül. Az utolsó két próbára nem feltétlenül lett volna szükségünk. (Ha nem feltételezzük sokaság/szupersokaság létezését.)

A determinációs együttható (R^2) értéke 0,560 (prediktorok számához igazított értéke 0,536). Az eredmény közepes erősségű kapcsolatra utal.

A 34. ábra alapján nehéz eldönteni, hogy az eloszlás megfelel-e a normalitás statisztikai követelményeinek. Az 13. táblázat alapján viszont a helyzet egyértelműnek tűnik. A négyből három próba esetén $p > \alpha$, ezért a normalitásra vonatkozó $H_0(13)$ -at nem utasítjuk el. Kis elemszámú minták esetén általában a Kolmogorov-Szmirnov próba a legmegbízhatóbb, s p értéke itt egészen magas.

13. táblázat Reziduálisok normalitásvizsgálata

Próba	Próbastatisztika	p
Shapiro-Wilk	0,9553	0,4269
Kolmogorov-Szmirnov	0,0935	0,9930
Cramér-von Mises	1,7988	0,0000
Anderson-Darling	0,3069	0,5341

A homoszkedaszticitásra vonatkozó Breusch-Pagan próba p értéke (0,1047) magasabb, mint α , ezért $H_0(14)$ -t nem vetjük el. A White-próba esetén viszont ($p = 0,0281 < \alpha$) eltérő

eredményt kaptunk. A homoszkedaszticitásra vonatkozó $H_0(14)$ -et ugyan egyértelműen nem utasíthatjuk el, de kétségeink vannak az érvényességéről.

Az előfeltételek részleges teljesülése melletti bizonytalanság mellett az FP7 és a H2020 faipari projektjeiben mért országokénti projektrésztvételek száma között közepes erősségű lineáris regressziót állapítottunk meg.

A lineáris modell ennek ellenére nem tűnik teljesen meggyőzőnek, különösen akkor, ha megvizsgáljuk néhány egyéb regressziós modell szignifikáns illeszkedési próbáinak eredményét.

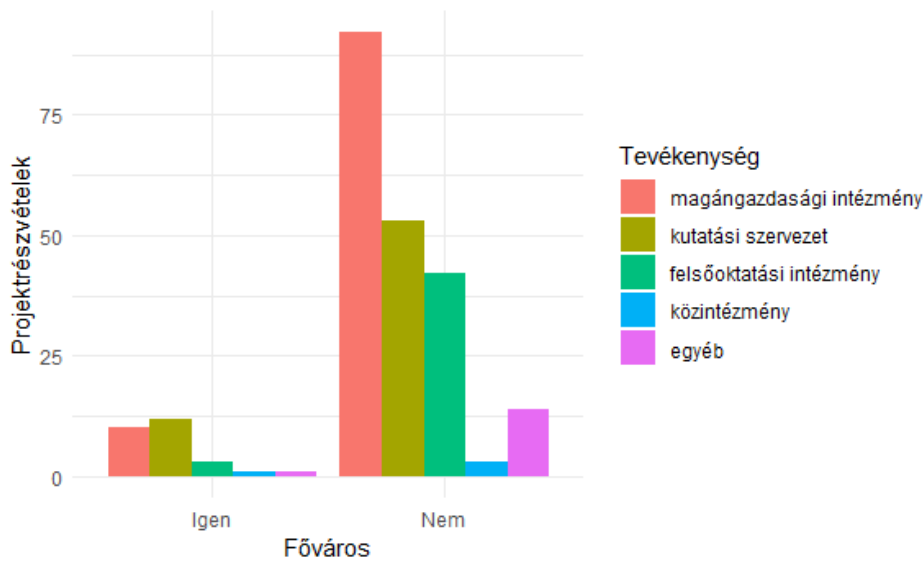
14. táblázat FP7(fa) és H2020(fa) programokban mért országokénti projektrésztvételek regresszióinak determinációs együtthatói

Regressziós modell	R ²	Igazított R ²
Logaritmikus	0,439	0,410
Inverz	0,259	0,220
Parabolikus	0,560	0,511
Harmadfokú	0,575	0,501
Vegyes	0,499	0,473
Hatványkitevős	0,453	0,425
Szigmoid	0,293	0,255
Növekedési	0,499	0,473
Exponenciális	0,499	0,473
Logisztikus	0,499	0,473

Látható, hogy elsősorban a parabolikus és a harmadfokú regresszió a lineáris modell alternatívájaként merült fel.

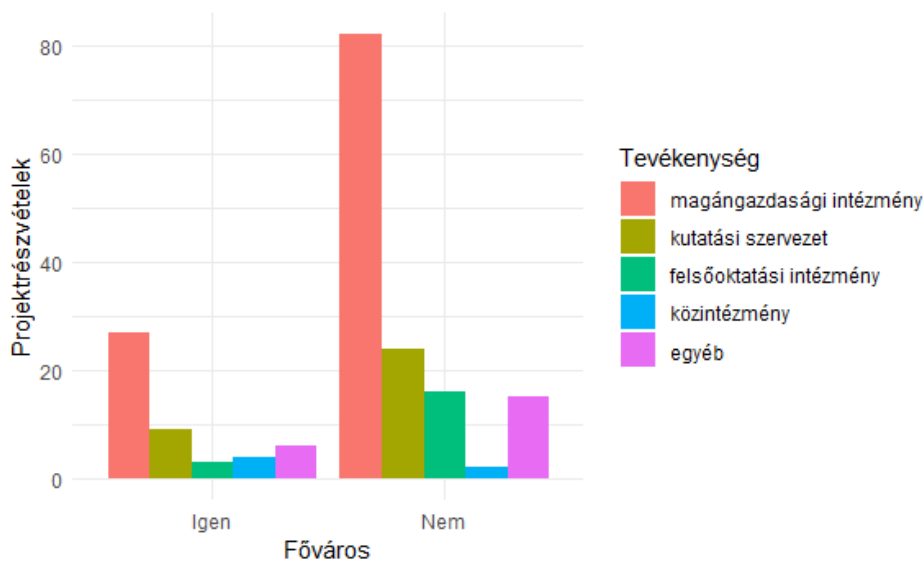
Talán kisebb jelentőségű, ezért csak leíró jelleggel vizsgáltuk meg, hogy a fővárosi (tehát a nemzeti szinten döntési központokban elhelyezkedő) résztvevők és a nem fővárosi projekt-résztvevők száma hogyan alakult.

Az FP7(fa) projektjeiben a fővárosok dominanciája (10,5%) nem igazán érhető tetten. A kis minta és a rész minta miatt a különböző tevékenységekhez tartozó szervezetek megoszlása is kevés releváns információt hordoz.



35. ábra Az egyes országok fővárosainak és egyéb településeinek résztvevői az FP7(fa) projektekben

A H2020(fa) projektekben az egyes országok fővárosaiban bejegyzett résztvevők aránya jelentősen nőtt (26,1%).



36. ábra Az egyes országok fővárosainak és egyéb településeinek résztvevői a H2020(fa) projektekben

7.3.3. Az egyedi és országokénti részvételek változásai az egyes programok között

A kétmintás Kolmogorov-Szmirnov próba eredményei alapján ($D = 0,04$, $p = 0,99$) az azonos eloszlásból származásra vonatkozó $H_0(6)$ -t az egyedi részvétel gyakoriságának eloszlása esetén nem vetjük el. Tehát a 7. szakmai hipotézis első felét, miszerint az FP7 és a H2020 faipari projektjeiben az egyedi projektrészvételek eloszlása a két programban ugyanolyan eloszlásból származik, nem vetjük el.

Problémát okoz azonban az, hogy nagy elemszámú diszkrét változókat ugyan elterjedt gyakorlat folytonosként kezelni bizonyos számítások során, de az egyedi résztvevők esetén egy-öt,

ill. egy-három darab részvételt tapasztaltunk, amely aligha tekinthető folytonosnak. Így a fenti próba nem feltétlenül alkalmazható.

Az alternatív számításhoz a 15. táblázatban a projektrészvételek számát három kategóriába rendeztük. A Pearson-féle χ^2 próba szigorú előfeltétele a 15. táblázatban – az elvárt gyakoriság minden egyes cellában minimum 5 – nem teljesül. A megengedőbb feltétel, azaz, hogy az összes cella maximum 20%-ában lehet az elvárt gyakoriság száma kevesebb, mint 5, sem teljesül (két cellában, a cellák 33%-ban nem teljesül a szigorú feltétel). Ezért úgy ítéljük meg, hogy a Pearson-féle χ^2 próba nem végezhető el.

Cramér-mutató viszont számítható. Értéke ($C = 0,11$) alapján nagyon gyenge sztochasztikus kapcsolat számszerűsíthető. Az érték önmagában azonban inkább a sztochasztikus kapcsolat erősségére utal, az eloszlások azonosságának nem jó mérőszáma.

15. táblázat Projektrésztevőkénti projektrészvételek és a programok kontingenciátáblája

Projektrészvételek száma	Gyakoriság	Program		Összesen
		FP7	H2020	
1	megfigyelt gyakoriság	198	150	348
	elvárt gyakoriság	193,7	154,3	348
2	megfigyelt gyakoriság	9	16	25
	elvárt gyakoriság	13,9	11,1	25
3 vagy több	megfigyelt gyakoriság	4	2	6
	elvárt gyakoriság	3,3	2,7	6
Összesen	megfigyelt gyakoriság	211	168	379
	elvárt gyakoriság	211	168	379

A Kullback-Leibler divergencia (0,03) és a Jensen-Shannon divergencia (0,008) nullához közeli, alacsony értékei alapján megállapíthatjuk, hogy az egyedi projektrészvételek számának eloszlása azonos. Ezért az azonos eloszlásra vonatkozó $H_0(6)$ -t nem utasítjuk el. Tehát ez a számítás is megerősíti, hogy a 7. szakmai hipotézis első felét, miszerint az FP7 és a H2020 faipari projektjeiben az egyedi projektrészvételek eloszlása a két programban ugyanolyan eloszlásból származik, nem vetjük el.

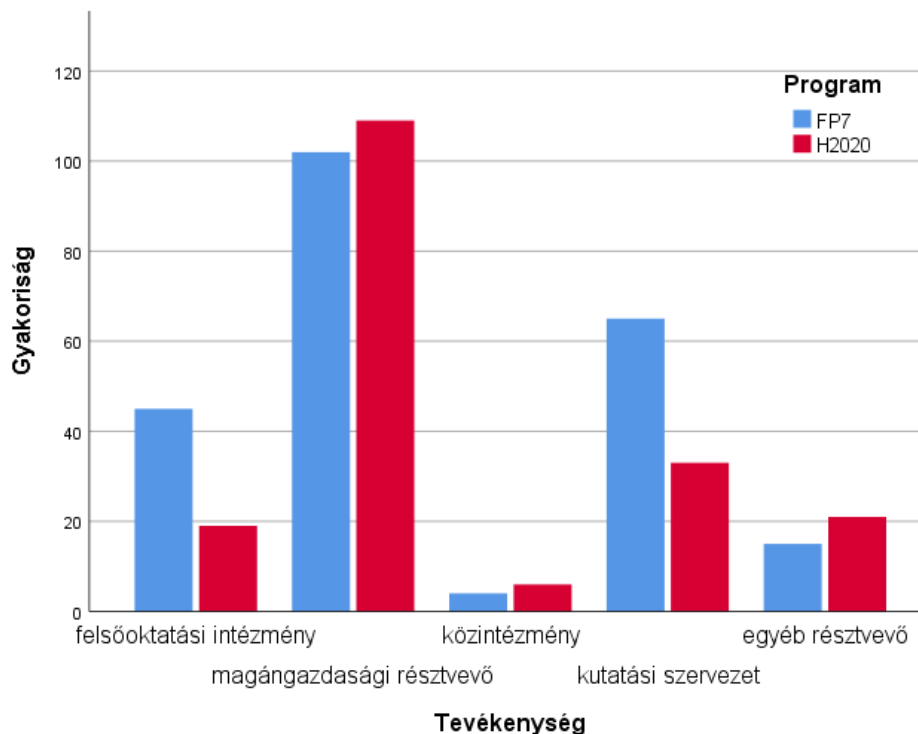
A kétmintás Kolmogorov-Szmirnov próba eredményei alapján ($D = 0,12$, $p = 0,83$), ezért az azonos eloszlásból származásra vonatkozó $H_0(6)$ -t az országonkénti részvétel gyakorisági eloszlása esetén nem vetjük el. Tehát a 7. szakmai hipotézis második felét, miszerint az FP7 és a H2020 faipari projektjeiben az országonkénti projektrészvételek eloszlása a két programban ugyanolyan eloszlásból származik, sem vetjük el.

A diszkrét-folytonos dichotómia ebben a második esetben kevésbé érvényesül, s a nagyszámú kategória miatt diszkrét változókra érvényes módszereket itt nem vizsgáltunk, a kategóriák esetleges összevonását pedig túlságosan nagy adatvesztésként értékeltük. (A diszkrét változót tehát folytonosként kezeltük).

A hasonló eloszlások általános, a dolgozat keretein valószínűleg túlmutató törvényszerűsége utalnak, de a módszertani problémák miatt érdemes óvatosan kezelni az eredményeket.

7.3.4. Tevékenység és programonkénti részvétel összehasonlítása

A 37. ábra alapján úgy tűnik, hogy a különböző tevékenységi formák szerinti részvétel gyakoriságában a két támogatási időszak között történt némi átrendeződés.



37. ábra A faipari projektek résztvevőinek tevékenység szerinti megoszlása

A Pearson-féle χ^2 próba szigorú előfeltétele a 16. táblázatban (ha hajszálnyival is, de) nem teljesül. Azonban a megengedőbb feltétel teljesül (egy cellában, a cellák csak 10%-ban nem teljesül a szigorú feltétel). Ezért úgy ítéljük meg, hogy a Pearson-féle χ^2 próba elvégezhető.

16. táblázat Tevékenység és program kontingenciatáblája

Tevékenység	Gyakoriság	Program		Összesen
		FP7	H2020	
felsőoktatási intézmény	megfigyelt gyakoriság	45	19	64
	elvárt gyakoriság	35,3	28,7	64
egyéb résztvevő	megfigyelt gyakoriság	15	21	36
	elvárt gyakoriság	19,8	16,2	36
magángazdasági résztvevő	megfigyelt gyakoriság	102	109	211
	elvárt gyakoriság	116,3	94,7	211
közintézmény	megfigyelt gyakoriság	4	6	10
	elvárt gyakoriság	5,5	4,5	10
kutatási szervezet	megfigyelt gyakoriság	65	33	98
	elvárt gyakoriság	54	44	98
Összesen	megfigyelt gyakoriság	231	188	419
	elvárt gyakoriság	231	188	419

Mivel a tábla nagyobb 2x2-nél több dimenziót tartalmaz, Cramér-féle V mutató számítható, értéke 0,21. Mivel $df = \min(\text{sorok száma} - 1, \text{oszlopok száma} - 1) = 1$, V értéke kis-

közepes hatást tükröz. 5%-os szignifikanciaszint mellett a Pearson-féle χ^2 értéke 18,425, $p = 0,001$, ezért az eredmény szignifikáns. $H_0(7)$ -t elvetjük, tehát a két program faipari projektjeiben való részvétel gyakorisága a tevékenység alapján különbséget mutat. A felsőoktatási intézmények és a kutatási szervezetek részvétele visszaesett a magángazdasági intézmények javára. Tehát a 8. szakmai hipotézist, miszerint van összefüggés a tevékenység, illetve az FP7 és a H2020 programokban való részvétel gyakorisága között, nem vetjük el.

A vizsgálat Cramér-féle V mutatón kívüli, következtetési statisztikai részét sokaság/szupersokaság feltételezése nélkül nem feltétlenül volt szükséges elvégezni.

7.3.5. A faipari projektrésztevők programonkénti támogatása

A két támogatási időszak között a faipari projektek összes támogatása nominálisan 47,11%-kal nőtt. A növekedés elmarad a két támogatási időszak összes támogatásában mért 52%-os növekménytől. Az euróövezet 2007 és 2014 (az FP7, ill. a H2020 kezdete) közötti időszakban 16,75%-os, a teljes 2007-2020 közötti időszakban pedig 22,7%-os inflációt szenvedett el (The World Bank, 2022), de a növekmény reálértéken még így is impozánsnak tűnhet.

A faipari projektekben tapasztalt egy projektrészvételre jutó támogatás két támogatási időszak közötti 80,75%-os növekedése igen jelentős forrásbővülésre utal az egyes résztvevők szintjén. A nemcsak faipari projekteket tartalmazó, tehát az összes projekt szintjén az egy projektrészvételre jutó támogatás azonban csak 22,04%-kal nőtt, ez lényegében az inflációhoz közeli növekmény.

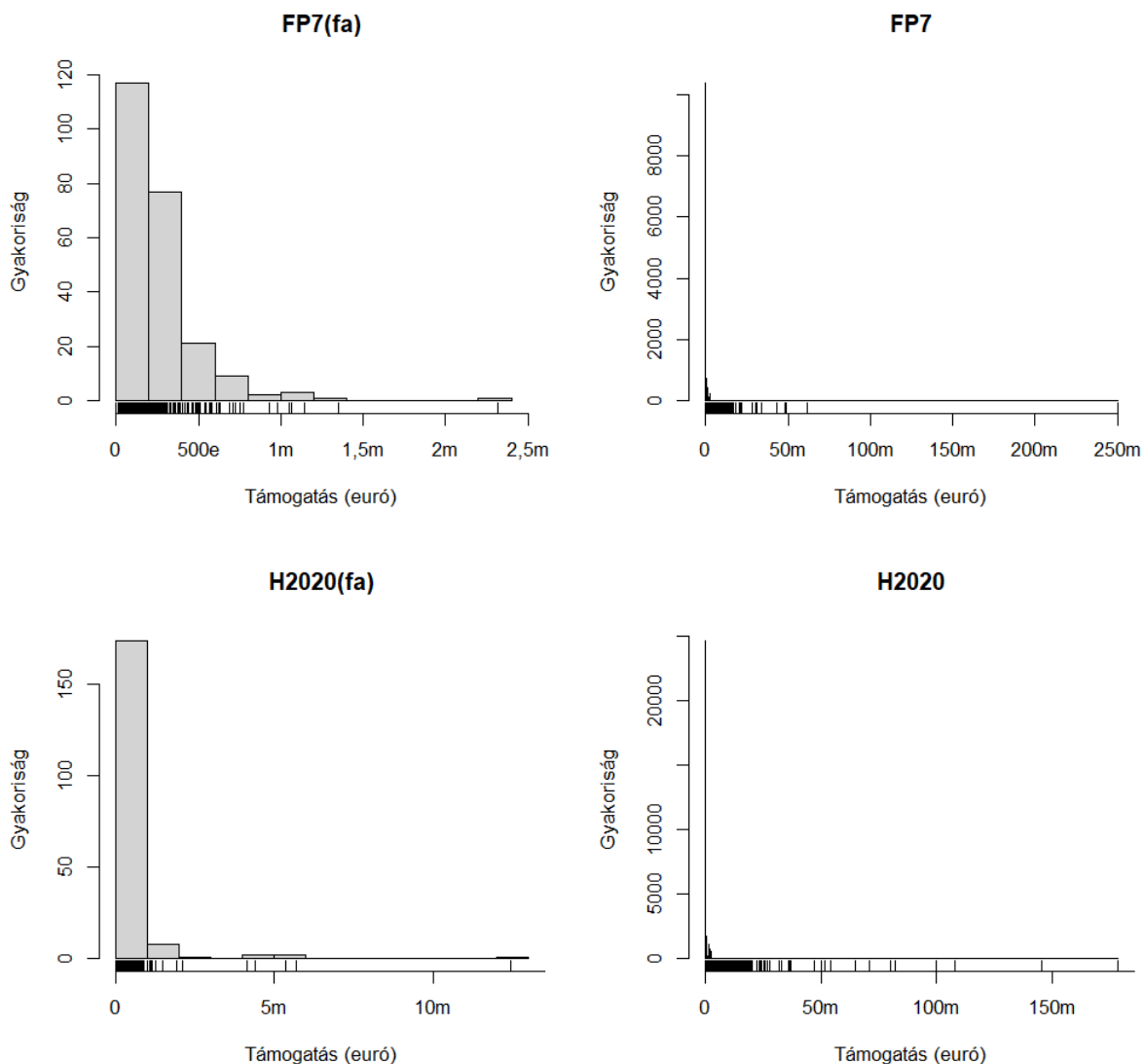
Tehát a faipari projektrészvételekre átlagosan jóval magasabb támogatás jutott a H2020 programban az FP7-hez képest.

17. táblázat Projektrészvételekre jutó közösségi támogatás deskriptív statisztikája

Mutató	FP7(fa)	FP7	H2020(fa)	H2020
Elemsszám	231	140.055	188	174.438
Módusz	0	0	0	0
Medián	196.075	189.664	182.187	206.000
Átlag	245.920	320.609	444.514	391.256
Szórás	261.878	907.762	1,165×10 ⁶	1,060×10 ⁶
MAD	121.275	135.379	151.844	156.000
IQR	243.849	295.331	385.200	362.942
Ferdeség	3,2	157,1	7,3	67,0
Csúcsosság	18,3	41.244,3	65,1	8.383,3
Kolmogorov-Szmirnov D	0,948	0,928	0,851	0,860
Kolmogorov-Szmirnov p	0	0	0	0
Minimum	0	0	0	0
Maximum	2,315×10 ⁶	2,500×10 ⁸	1,248×10 ⁷	1,782×10 ⁸
Összesen	5,681×10 ⁷	4,490×10 ¹⁰	8,357×10 ⁷	6,825×10 ¹⁰

Az átlagos értékek azonban megtévesztőek, mivel két csoport (minta, sokaság) viszonyáról csak normális eloszlás mellett adnak megbízható információt. A 17. táblázat (Kolmogorov-Szmirnov $p = 0$) és a 38. ábra alapján azonban a normális eloszlásra vonatkozó felvetés nyilván magabiztosan elvethető. (Bár a normalitás vizsgálatára a Shapiro-Wilk próba gyakoribb

választás, 3000 feletti elemszámnál ezt nem lehet kiszámolni, ezért – az összehasonlíthatóság kedvéért – választottuk a Kolmogorov-Szmirnov próbát. Nagy elemszámnál a Kolmogorov-Szmirnov próba is adhat pontatlan eredményt, de egyértelmű eredmény esetén használható.)



38. ábra A projektrészvételekre jutó támogatás eloszlása

A $H_0(15)$ vizsgálata során a reálértékre való diszkontálás (H2020) a korábban számított 16,75%-os inflációs értékkel történt, tehát 2007-es árakon számoltunk 2014-es reálértéket.

A 18. táblázat alapján két esetben nem vethetjük el a két medián azonosságára vonatkozó $H_0(15)$ hipotézist, a mintabeli medián eltérését μ (az ismertnek tekintett medián) értéktől a faipari projektek esetében nem sikerült igazolni.

Ez a konkrét esetben azt jelenti, hogy nincs szignifikáns különbség az egyes csoportok mediánja között (amennyiben azt a minták mögötti – talán csak elméleti – sokaságokra vonatkoztatjuk). Azaz a résztvételekenkénti támogatás azonosságát egyik faipari viszonylatban sem sikerült cáfolnunk.

18. táblázat Wilcoxon-próba a projektrésztvelekre jutó támogatásokra

Medián	Minta	Érték	μ	μ forrása	V	p
196.075	FP7(fa)	nominális/reál	189.664	FP7	14.667	0,21
182.187	H2020(fa)	nominális/reál	206.000	H2020	9.578	0,35
182.187	H2020(fa)	nominális	196.075	FP7(fa)	9.909	0,17
156.078	H2020(fa)	reál	196.075	FP7(fa)	8.968	0,91
206.000	H2020	nominális	189.664	FP7	$9,5 \times 10^9$	0
176.446	H2020	reál	189.664	FP7	$8,5 \times 10^9$	0

A Wilcoxon-próba nem tartozik az igazán „erős” próbák közé, így az amúgy is ellentmondásos eredményt nagyon óvatosan kell kezelnünk. A próba egészen extrém μ értékeknél sem feltétlenül ad $p < \alpha$ eredményt. A nagy elemszámok és extrém eltéréseket mutató adatsorok miatt a próba használhatósága kétséges.

Ráadásul ebben az esetben nem feltétlenül kell elszakadnunk a deskriptív statisztika szintjétől, így következtetési statisztikai eljárás nélkül (mögöttes sokaság feltételezése nélkül, tehát az adatokat sokaságnak tekintve) is tehetünk olyan megállapítást, hogy a faipari projektrésztvelekre jutó támogatás mediánja csökkent az FP7 és a H2020 között. Ez tűnik reális következtetésnek.

A fentiek alapján a 9. szakmai hipotézist, miszerint a faipari projektek közösségi támogatásának teljes sokasággal összevethető középértéke (átlag vagy medián) nem tér el a teljes sokaságtól, kétségesnek tartjuk. A 10. szakmai hipotézist pedig, miszerint a faipari projektek közösségi támogatása a két támogatási időszak között lényegesen nem változott, szintén nem tudjuk egyértelműen elvetni/nem elvetni. Ezért ezzel a két hipotézissel kapcsolatban tézist nem fogalmazunk meg.

A 38. ábra alapján az FP7-ben és a H2020-ban a faipari projektrésztvelekre jutó támogatás között valamilyen összefüggés sejlik fel.

A vizsgálat során csak azokhoz a projektrésztvevőkhöz tartozó projektrésztveleket vehettük figyelembe, amelyek mindkét támogatási programban részt vettek. Azonban mindösszesen 13 ilyen projektrésztvevő akadt. A résztvevőik során kapott támogatások közötti korreláció és regresszió nagyon kevés valós információt hordozott volna, ezért ezeket a számításokat nem végeztük el.

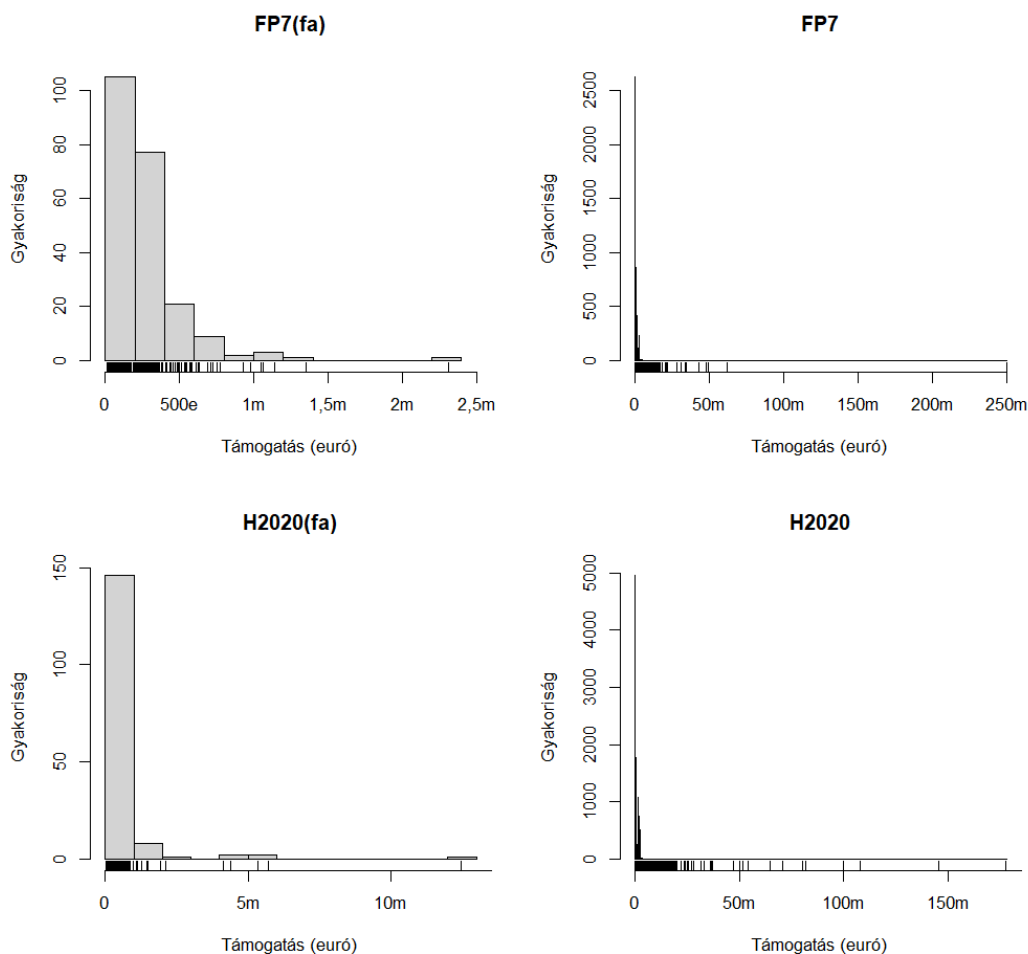
A fenti eredmények alapján azonban felmerül a kérdés, hogy helyes-e a támogatást nem kapó projektrésztveleket a vizsgálat tárgyának tekintenünk, ezért a fenti számításokat e résztvevőket elhagyva is elvégeztük.

Számos projektrésztvevő nem kapott közösségi támogatást, részvételét saját forrásból finanszírozta. A nullértékek nélküli 19. táblázat nem mutat drámai változást. (Az FP7-ben és a H2020-ban kapott támogatások minimumértékével – egy eurócent – rendelkező projekteket is elhagyhatnánk, komolyabb változást az sem jelentene.)

19. táblázat Projektrésztvelekre jutó közösségi támogatás
deszkriptív statisztikája (nullértékek nélkül)

Mutató	FP7(fa)	FP7	H2020(fa)	H2020
Elemszám	219	130.030	160	149.988
Módusz	9.185	100.000	50.000	50.000
Medián	215.602	207.631	233.725	250.905
Átlag	259.396	345.327	522.304	455.036
Szórás	262.381	937.565	1,247×10 ⁶	1,130×10 ⁶
MAD	119.362	132.558	183.725	150.708
IQR	235.637	292.162	452.324	357.464
Ferdeség	3,3	153,5	6,8	64,1
Csúcsosság	18,6	39.023,1	56,6	7.525,9
Kolmogorov-Szmirnov D	1	1	1	1
Kolmogorov-Szmirnov p	0	0	0	0
Minimum	9.185,3	0,01	6.267,4	0,01
Maximum	2,315×10 ⁶	2,500×10 ⁸	1,248×10 ⁷	1,782×10 ⁸
Összesen	5,681×10 ⁷	4,490×10 ¹⁰	8,357×10 ⁷	6,825×10 ¹⁰

A 39. ábra sem jelez jelentős változást az eloszlásban.



39. ábra A projektrésztvelekre jutó támogatás eloszlása (nullértékek nélkül)

Az előzőektől eltérően a 20. táblázat adatai némi változást mutatnak, hiszen $H_0(15)$ hipotézist egy újabb esetben vethetjük el. Tehát elvetjük azt a hipotézist, miszerint a faipari projektrészvételekre jutó támogatás a két támogatási időszak között ne változott volna.

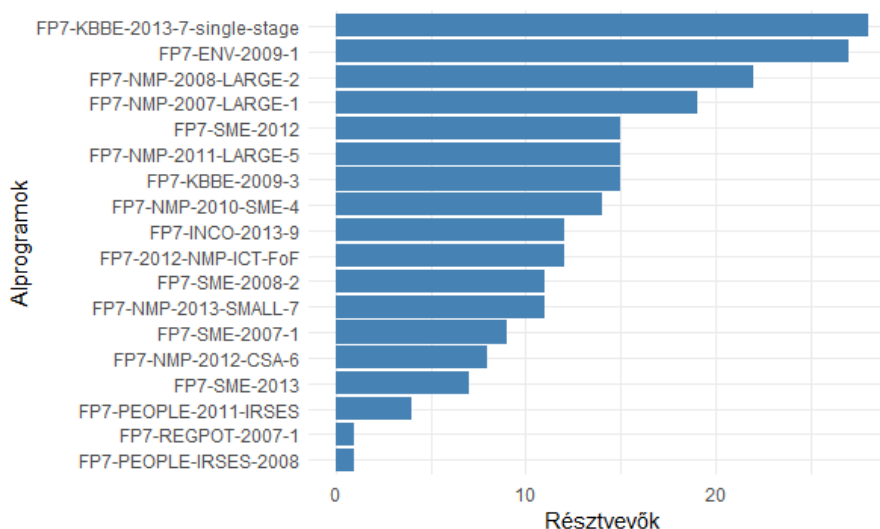
20. táblázat Wilcoxon-próba a projektrészvételekre jutó támogatásokra (nullértékek nélkül)

Medián	Minta	Érték	μ	μ forrása	V	p
215.602	FP7(fa)	nominális/reál	207.631	FP7	12.747	0,45
233.725	H2020(fa)	nominális/reál	250.905	H2020	7.272	0,16
233.725	H2020(fa)	nominális	215.602	FP7(fa)	8.005	0,01
200.192	H2020(fa)	reál	215.602	FP7(fa)	7.253	0,17
250.905	H2020	nominális	207.631	FP7	$7,8 \times 10^9$	0
214.908	H2020	reál	207.631	FP7	$6,9 \times 10^9$	0

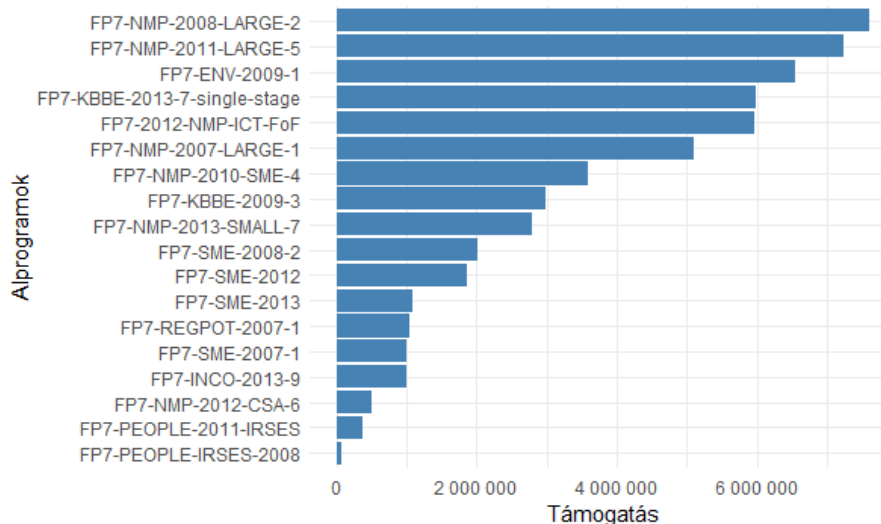
A mögöttes sokaság feltételezését az előzőekhez hasonlóan itt sem tartjuk fontosnak, s a próba használhatósága is kétséges (bár ebben a második esetben az eredmény talán megbízhatóbb). Ha a mintát sokaságnak tekintjük, akkor a változások a 19. táblázat alapján egyértelműen megfogalmazhatóak. Tehát, ha csak az FP7-re és a H2020-ra koncentrálunk, akkor nominálisan a faipari projektrészvételekre jutó támogatások mediánja az FP7-ben magasabb, a H2020-ban alacsonyabb, mint a teljes programban. Nominálisan a faipari projektek támogatása nőtt a két programban, reálértékben azonban nem. (Ez az állítás azonban a programokon kívül eső támogatásokra nem általánosítható.)

A magasabb támogatásban részesült projektrészvételek esetén a projektrészvevő gyakran több projektben is részt vett. Az FP7(fa) projektekben a több projektben is részt vevők átlagosan 37%-kal, a H2020(fa) projektekben 144%-kal kaptak több támogatást projektrészvételeként. Bár ez nyilván számos tényező függvénye, de a hatékony partnerkeresési stratégia része lehet a többes projektrészvétel megcélzása.

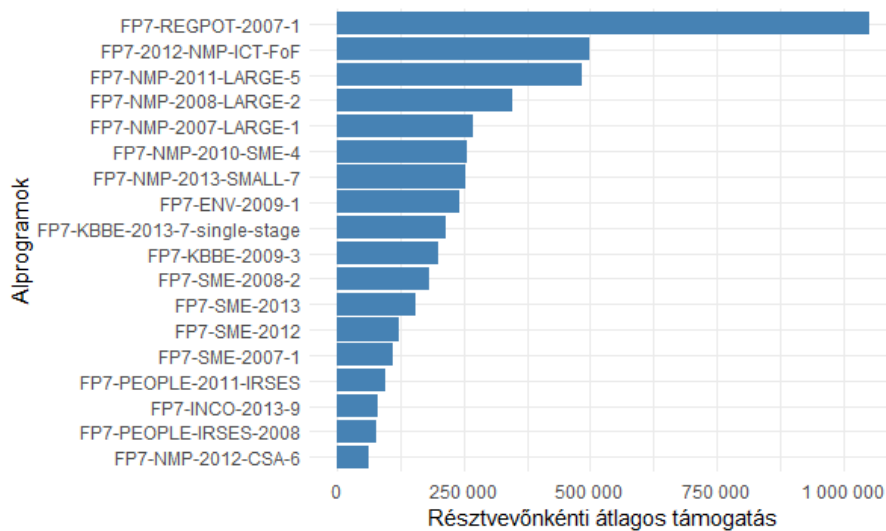
Az egyes alprogramokból kapott támogatás mind az FP7(fa), mint a H2020(fa) projektekben igen nagy eltéréseket mutat. (40. ábra - 45. ábra)



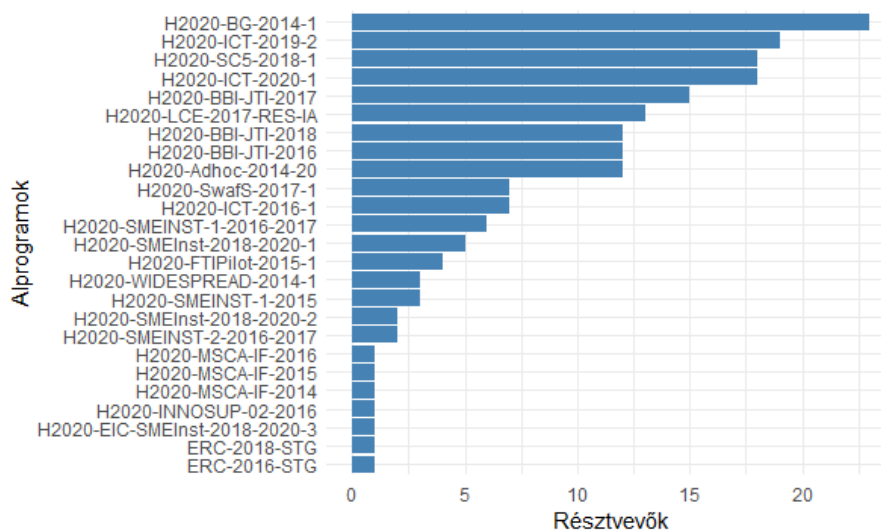
40. ábra Projektrészrtvevők száma az FP7(fa) projektekben FP7-alprogramonként



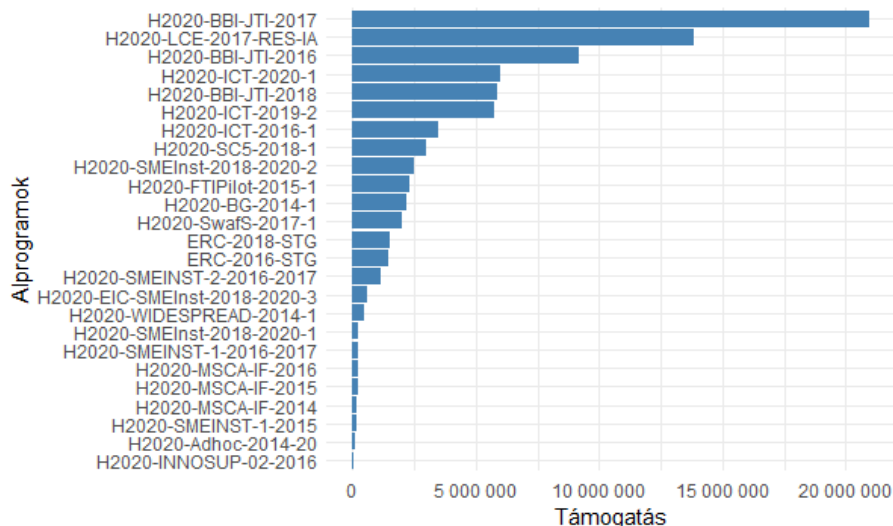
41. ábra Közösségi támogatás az FP7(fa) projektekben FP7-alprogramonként (euróban)



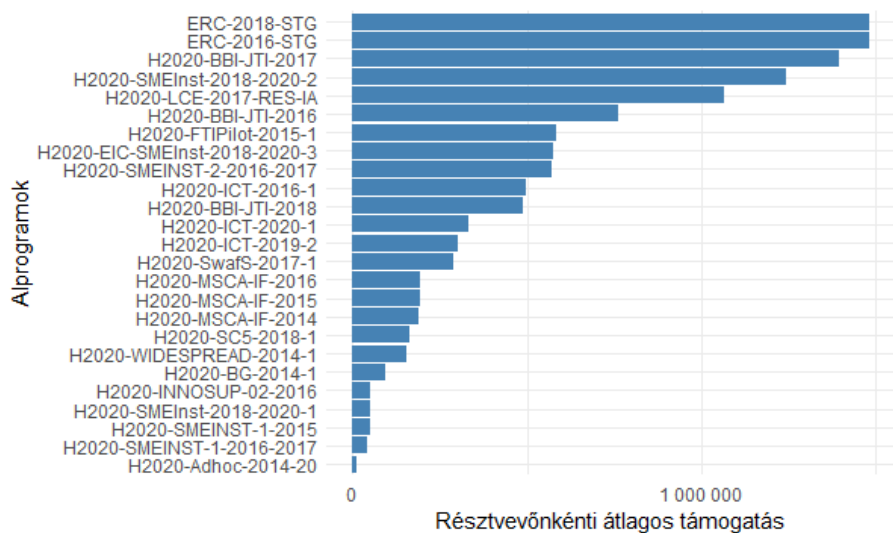
42. ábra Résztvevőnkénti közösségi támogatás az FP7(fa) projektekben FP7-alprogramonként (euróban)



43. ábra Projekt résztvevők száma a H2020(fa) projektekben H2020-alprogramonként



44. ábra Közösségi támogatás a H2020(fa) projektekben H2020-alprogramonként (euróban)



45. ábra Résztevőnkénti közösségi támogatás a H2020(fa) projektekben H2020-alprogramonként (euróban)

7.3.6. Az egyes országok által kapott közösségi támogatás a két programban

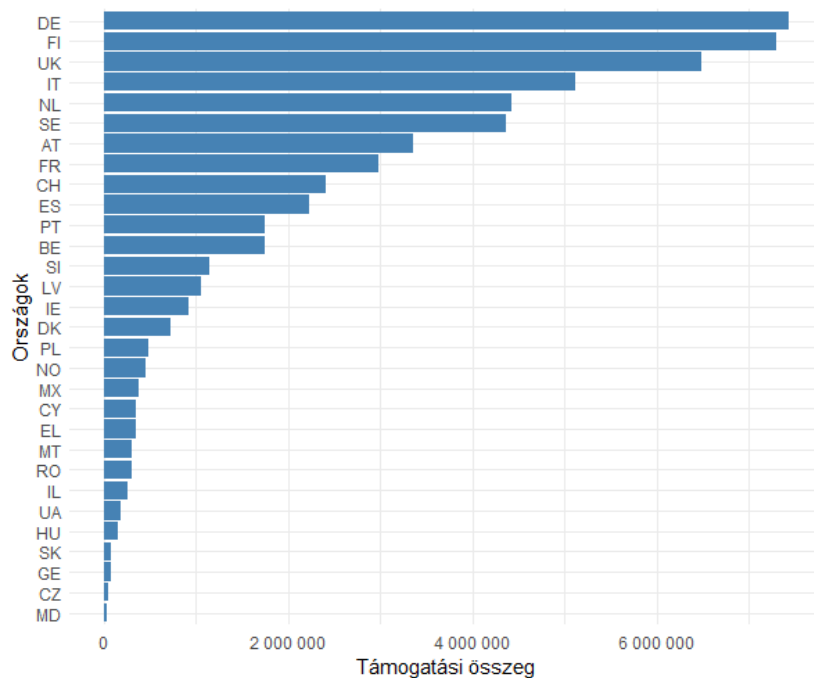
A 21. táblázat értékei alapján a két adatsorban inkább az eltérések dominálnak, bár pl. a ferdeség értéke meglehetősen hasonló.

A ferdeség és a csúcosság itt sem utal normális eloszlásra.

21. táblázat Az egyes országok közösségi támogatásának leíró statisztikája

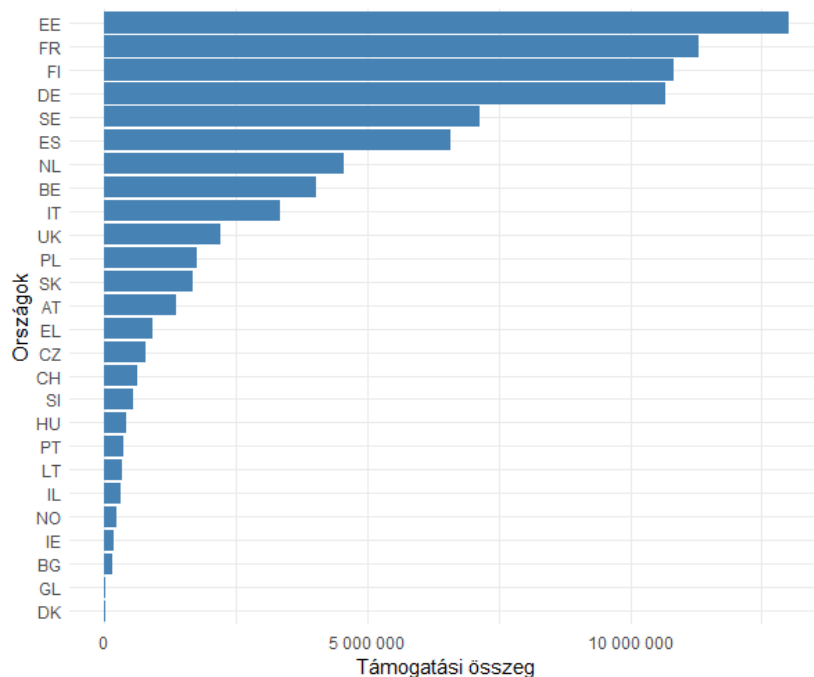
Mutató	FP7(fa)	H2020(fa)
Országok száma	32	26
Módusz	0	31.250
Medián	605.828,40	1,151×10 ⁶
Átlag	1,775×10 ⁶	3,214×10 ⁶
Szórás	2,243×10 ⁶	4,088×10 ⁶
MAD	550.981,10	982.876,87
IQR	2,309×10 ⁶	4,074×10 ⁶
Ferdeség	1,42	1,34
Csúcsosság	0,96	0,46
Minimum	0	31.250
Maximum	7,422×10 ⁶	1,300×10 ⁷
Összes támogatás	5,681×10 ⁷	8,357×10 ⁷

Az FP7(fa) adatsoron belül az össztámogatást tekintve Németország (7.422.483 euró), Finnország (7.288.517 euró), Egyesült Királyság (6.480.051 euró), Olaszország (5.116.302 euró), Hollandia (4.425.509 euró) volt az első öt ország sorrendje. Az FP7(fa) adatsoron belül két olyan ország – Kanada és az USA – található, amely nem kapott támogatást. (46. ábra)



46. ábra Az FP7 faipari projektjeiből az egyes országokra jutó támogatás (euróban)

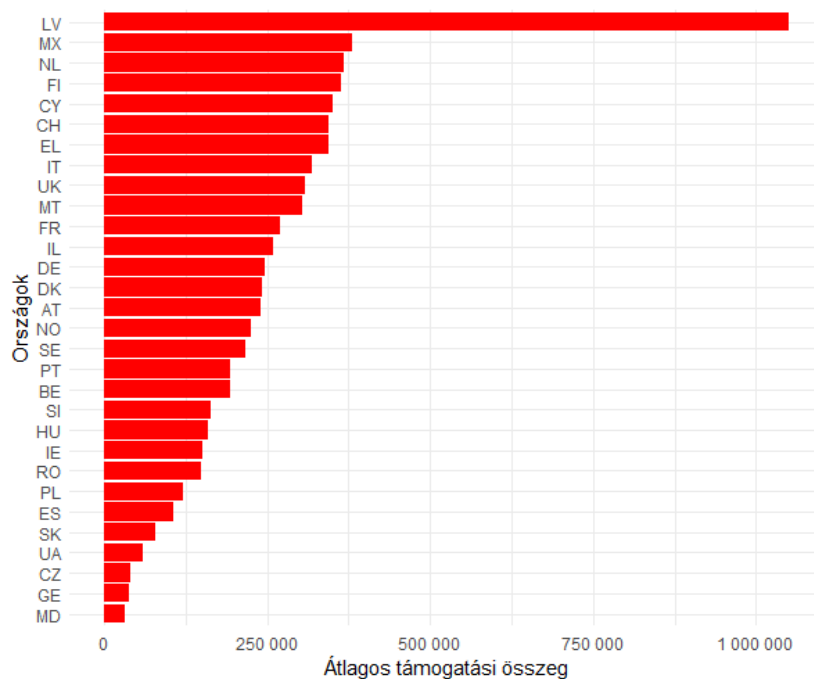
A H2020(fa) adatsoron belül viszont a legtöbb össztámogatást talán meglepő módon Észtország kapta (13.002.351 euró), majd Franciaország (11.300.647 euró), Finnország (10.830.420 euró), Németország (10.682.206 euró) és Svédország (7.143.528 euró) a sorrend. (47. ábra)



47. ábra A H2020 faipari projektjeiből az egyes országokra jutó támogatás (euróban)

A kiváló észt szereplés okainak feltárása a dolgozat keretein túlmutató elemzést igényelne.

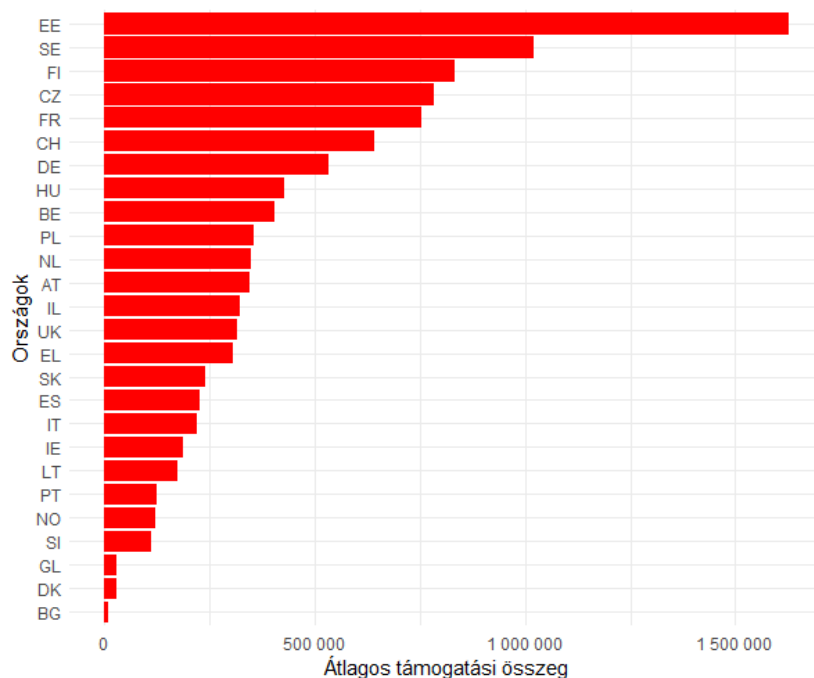
A résztvevők átlagos támogatása az FP7(fa) projektekben országonként csoportosítva igen jelentős különbségeket mutat. Az átlagos támogatás a litvániai résztvevőknél volt a legnagyobb (1.049.320 euró), majd Mexikó (380.000 euró), Hollandia (368.792 euró), Finnország (364.426 euró) és Ciprus (351.915 euró) következett. (48. ábra)



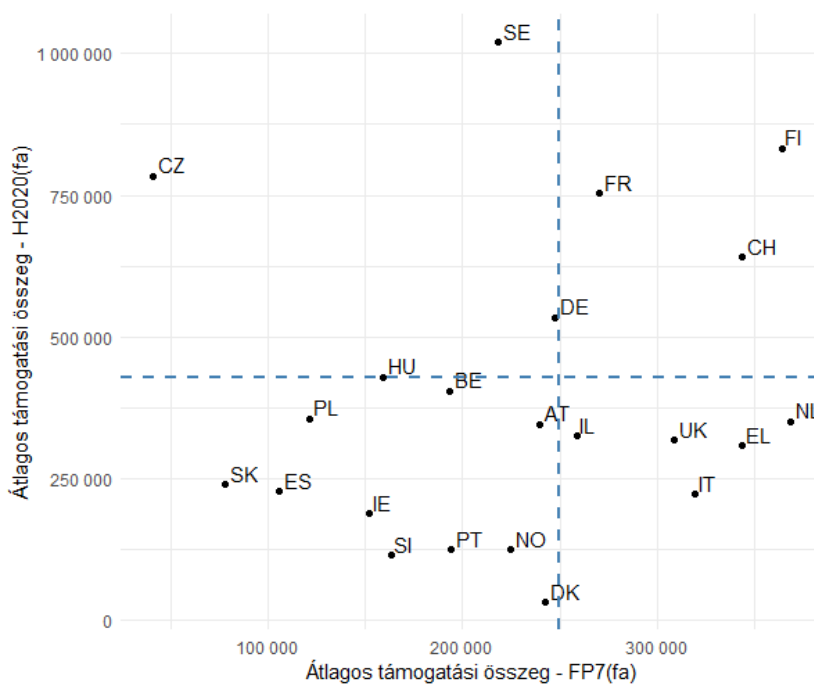
48. ábra Projekt résztvevőnkénti átlagos támogatás országonként csoportosítva – FP7(fa)

A H2020(fa) projektekben az átlagos támogatási összeg alapján Észtország (1.625.294 euró), Svédország (1.020.504 euró), Finnország (833.109 euró), Csehország (784.278 euró), Franciaország (753.376 euró) volt az első öt ország. (49. ábra)

A 50. ábra a mindkét támogatási programban részt vett országok aktorainak átlagos támogatását mutatja a két programban. Az egyes országok pozíciója alapján nehéz pontos képet (klasztereket) alkotni.

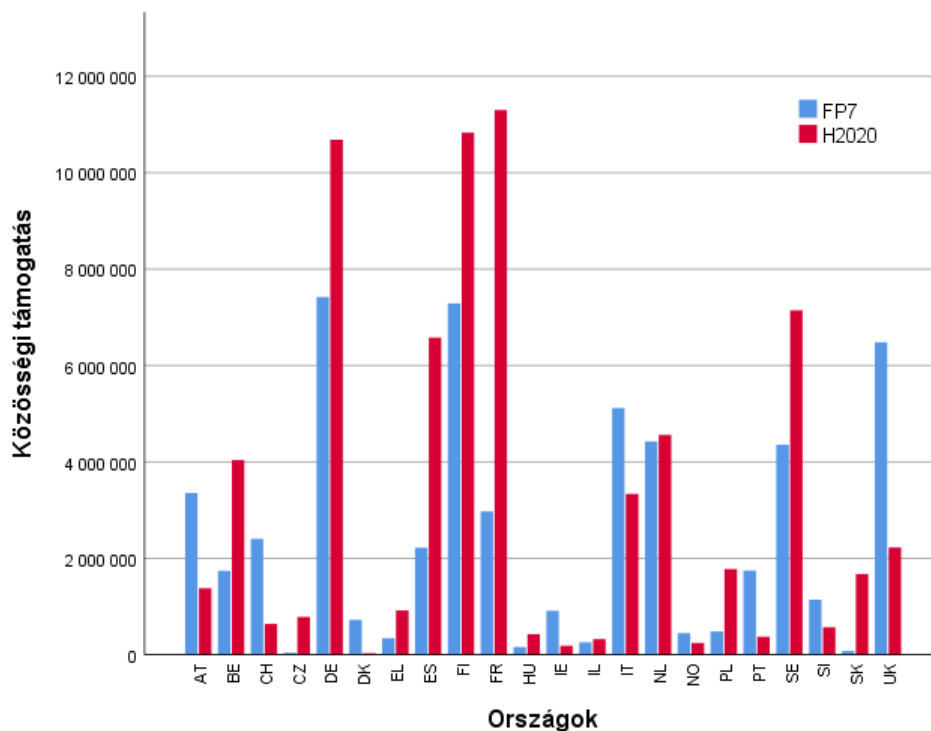


49. ábra Projekt résztvevőnkénti átlagos támogatás országonként csoportosítva – H2020(fa)



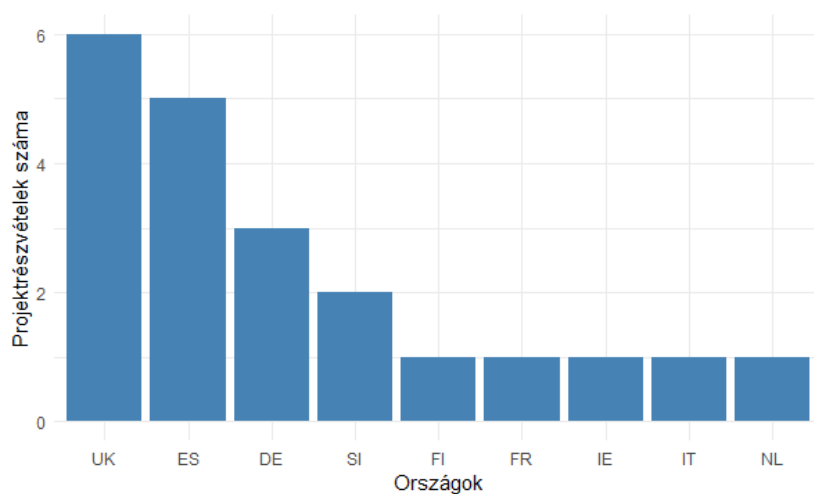
50. ábra Projekt résztvevőnkénti átlagos támogatás a két programban az egyes országokban

A 22, mindkét programban részt vett ország közül mindegyik kapott közösségi támogatást a két programból. (51. ábra) (Természetesen ország alatt, az országhoz tartozó intézményre, cégre stb. gondolunk.) A fejezet további részében csak ezt a 22 országot elemezzük.



51. ábra A faipari projektek támogatása országonként az egyes programokban a mindkét programban támogatást kapó résztvevőknél (euróban)

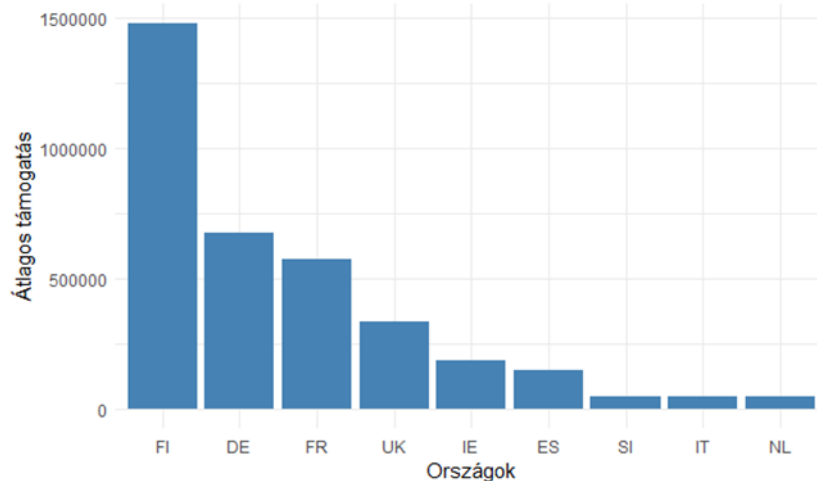
A faipari együttműködésekbe kevésbé integrált országok potenciális résztvevői számára érdekes lehet, hogy az egyrésztvevős pályázatokon mely országokhoz tartozó résztvevők értek el közösségi támogatásban tetten érhető sikert, illetve milyen mértékben tették ezt. Az FP7(fa) projektekben elég kevés ilyen projektet találunk. Egyet Lettországból 1.049.320 euró, egyet pedig Olaszországból 79.200 euró támogatással. A H2020(fa) projektekben sokkal több egyszereplős projektet találunk, szám szerint 21-et (52. ábra).



52. ábra Egyszereplős projektek száma országonként a H2020(fa) projektekben

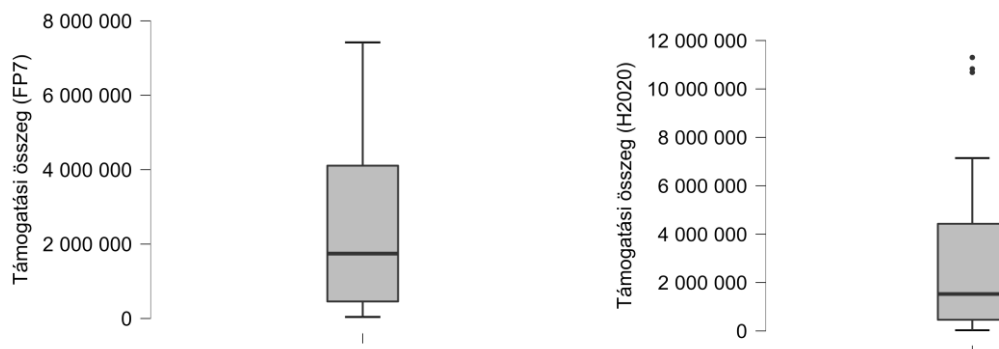
Leginkább a britek (6), a spanyolok (5) és a németek (3) vettek részt ilyen projektekben.

Az egyszereplős projektek átlagbevétele az FP7(fa) projektekben 564.260 euróra, a H2020(fa) projektekben 345.572 euróra rúgott. A H2020(fa) projektekben az egyszereplős projektek közösségi támogatása elég nagy szórást mutat (53. ábra).



53. ábra Egyszereplős projektek által kapott közösségi támogatás átlaga országonként a H2020(fa) projektekben

A 54. ábra alapján a H2020(fa) adattábla országonkénti közösségi támogatási összegei között outliereket feltételezhetünk.



54. ábra A közösségi támogatás boxdiagramjai az FP7 és a H2020 programok faipari projektjeiben országonként (euróban)

FP7(fa) esetén $I(8)(9)(10)$ rendre $[-5.017.106; 9.583.538]$, $[-4.673.959; 8.162.135]$, $[-7.572.218; 12.138.650]$.

H2020(fa) esetén $I(8)(9)(10)$ rendre $[-5.484.023; 10.376.824]$, $[-3.996.812; 7.053.077]$, $[-8.259.671; 13.152.473]$.

Az FP7(fa) esetén egyik eljárás sem mutat outliert. H2020(fa) adatai közül azonban az első eljárás két, a második eljárás 4, a harmadik eljárás egy outliert sem szűrt ki. Nincs két azonos eredmény a három módszer alapján, így úgy döntöttünk, hogy egyik résztvevő adatait sem tekintjük outliernek.

A vizsgált adatok a Shapiro-Wilk és az egymintás Kolmogorov-Szmirnov normalitási próba alapján sem normális eloszlásúak (22. táblázat), mivel $p < 0,05$. A normalitást feltételező H_0 -t tehát elvetjük.

22. táblázat Az egyes országok programonkénti támogatásának normalitáspróbái

Mutató	Támogatás (országonként)	
	FP7(fa)	H2020(fa)
Shapiro-Wilk (W)	0,86	0,78
P (Shapiro-Wilk)	0,01	0,00
Kolmogorov-Szmirnov (D)	1	1
P (Kolmogorov-Szmirnov)	0,00	0,00

A normalitási próbák igen hasonló eredményei előrevetítik, hogy a normalitás nem teljesülése ellenére a két gyakorisági sor hasonló eloszlást mutat (22. táblázat), illetve a két változó közötti összefüggés is valószínűsíthető.

A Mahalanobis-távolság minden esetben kisebb, mint 13,82. (A maximális érték is csak 8,11.) Tehát a Mahalanobis-távolság alapján nincs többdimenziós outlier. A Cook-távolság egyetlen egy esetben sem nagyobb, mint egy. (A legnagyobb érték 0,51.) Többdimenziós outlier tehát nem akadályozza a korrelációs számítás és a regresszió számítás elvégzését.

A Spearman-féle rangkorrelációs számítással és próbával viszonylag erős (közepesnél erősebb), pozitív korrelációt mértünk FP7(fa) és H2020(fa) országonként mért közösségi támogatás között ($\rho = 0,65$). A $p = 0,00023 < \alpha$ érték alapján a függetlenségre vonatkozó $H_0(5)$ -t elutasítjuk. Tehát a 11. szakmai hipotézist, miszerint Az FP7-ben és a H2020-ban az országonkénti közösségi támogatás mértéke között van összefüggés, nem vetjük el.

Természetesen a p értéket (illetve a próbát) csak akkor kell figyelembe vennünk, ha mögöttes sokaságot (vagy szupersokaságot) feltételezünk.

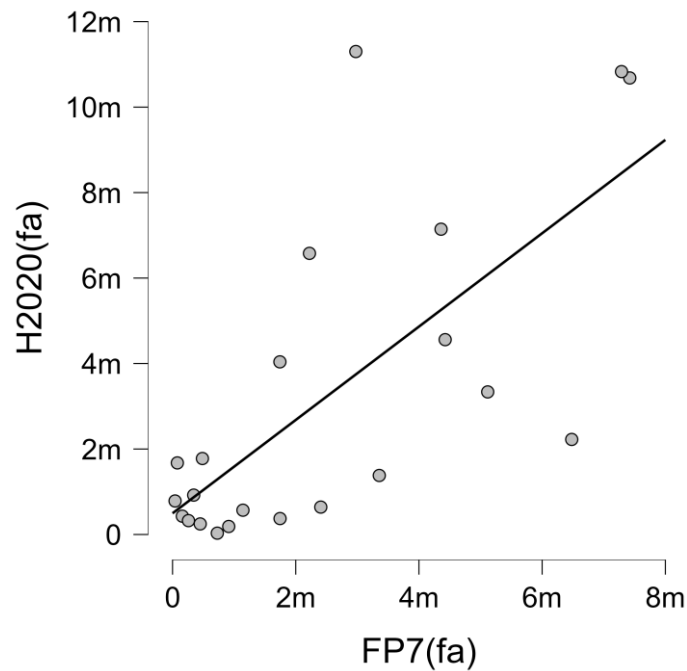
A 0 támogatási összeget kapó résztvevőket mellőző adatsorokra elvégzett, kétmintás Kolmogorov-Szmirnov próba eredményei alapján ($D = 0,16$, $p = 0,77$) az országok szerinti támogatás szintjén az azonos eloszlásra vonatkozó $H_0(6)$ -t nem utasítjuk el. Tehát a 12. szakmai hipotézist, miszerint az FP7 és a H2020 projekt támogatások országonkénti eloszlása ugyanolyan eloszlásból származik, nem vetjük el.

A regresszió számítás során az FP7(fa) országonkénti közösségi támogatási összegét tekintjük független, a H2020(fa) országonkénti támogatási összegét tekintjük függő változónak.

A lineáris regresszió egyenes, ill. a becslőfüggvény paraméterei: $\hat{\beta}_0 = 4,96 \times 10^5$, $\hat{\beta}_1 = 1,09$, $s = 2,73 \times 10^6$, $s^2 = 7,44 \times 10^{12}$, $s_{\hat{\beta}_0} = 8,41 \times 10^5$, $s_{\hat{\beta}_1} = 0,25$.

A Student-féle t -próba értékei: $t = 4,42$, $t_{0,975}(df = 22) = 2,073$. F -próba p értéke: 0,00. Mivel a számított t érték a kritikus tartományban van, illetve $p < \alpha$, szignifikáns lineáris összefüggést feltételezünk a változók között. A linearitást kizáró $H_0(12)$ -t ezért elutasítjuk.

A fenti próbák értelmezése – a korábban említett esetekhez hasonlóan – sokaság vagy szuper-sokaság feltételezése mellett értelmesek, de a lineáris regresszió becsült paraméterei enélkül is értelmezhetők.



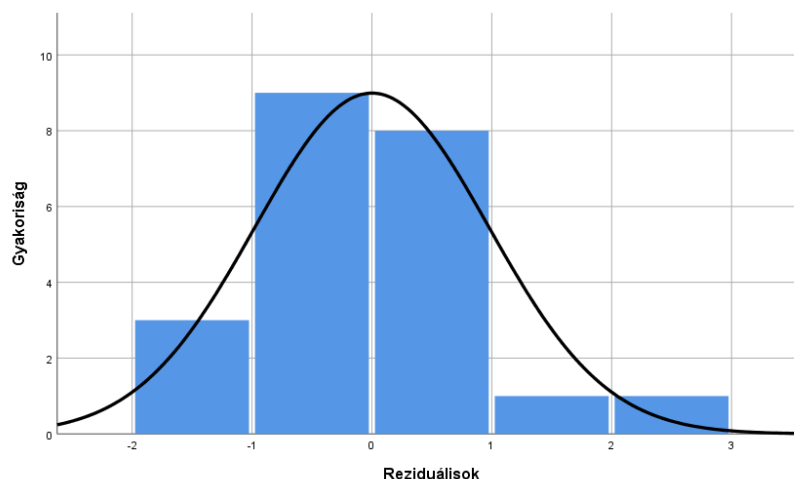
55. ábra Országokénti közösségi projekt támogatások lehetséges kapcsolata

A determinációs együttható (R^2) értéke 0,495 (a prediktorok számához igazított értéke 0,469). Az eredmény közepes erősségű kapcsolatra utal.

A 56. ábra alapján nehéz eldönteni, hogy az eloszlás megfelel-e a normalitás statisztikai követelményeinek. A 23. táblázat alapján viszont a helyzet egyértelműnek tűnik. A négyből három próba esetén $p > \alpha$, ezért a normalitásra vonatkozó $H_0(13)$ -at nem utasítjuk el. Kis elemszámú minták esetén általában a Kolmogorov-Szmirnov próba a legmegbízhatóbb, s p értéke itt egészen magas.

23. táblázat A reziduálisok normalitásvizsgálata

Próba	Próbastatisztika	p
Shapiro-Wilk	0,9514	0,3373
Kolmogorov-Szmirnov	0,1026	0,9563
Cramér-von Mises	1,8788	0,0000
Anderson-Darling	0,3638	0,4082



56. ábra Az országok programonkénti projektrésztvételeinek száma közötti lineáris regresszió standardizált reziduálisainak eloszlása

A homoszkedaszticitásra vonatkozó Breusch-Pagan próba p értéke (0,1383) magasabb, mint α , ezért a homoszkedaszticitásra vonatkozó $H_0(14)$ -t nem vetjük el. A White-próba esetén hasonló a helyzet ($p = 0,110 > \alpha$). A homoszkedaszticitásra vonatkozó $H_0(14)$ -et ezért nem utasítjuk el.

Az előfeltételek teljesülése mellett tehát az FP7 és a H2020 faipari projektjeiben kapott országok szerinti támogatás között közepes erősségű lineáris regressziót állapítottunk meg.

A lineáris modell ennek ellenére nem tűnik teljesen meggyőzőnek, különösen akkor, ha megvizsgáljuk néhány egyéb regressziós modell szignifikáns illeszkedési próbáinak eredményét. (A nem szignifikánsakat vagy más megközelítésben az igen alacsony R^2 -tel rendelkezőket nem tüntettük fel.)

24. táblázat FP7(fa) és H2020(fa) programokban mért országonkénti projektrésztvételek regresszióinak determinációs együtthatói

Regressziós modell	R^2	Igazított R^2
Logaritmikus	0,353	0,320
Parabolikus	0,495	0,441
Harmadfokú	0,533	0,455
Vegyes	0,446	0,419
Hatványkitevős	0,285	0,250
Növekedési	0,446	0,419
Exponenciális	0,446	0,419
Logisztikus	0,446	0,419

Látható, hogy elsősorban a parabolikus és a harmadfokú regresszió a lineáris modell alternatívájaként merül fel.

7.3.7. Tevékenység szerinti egyéni támogatás összehasonlítása az egyes programokban

A 25. táblázat adataiból leginkább annyit lehet kiolvasni, hogy a H2020 faipari projektjeiben a tevékenységtípusok szerinti egyéni támogatásátlag nagyobb különbséget mutatott, mint az

FP7-ben. A szórás a két támogatási program egyéni faipari támogatásai között nőtt, de az egyes csoportokon belül a kép vegyes, három csoporton belül a szórás jelentősen csökkent, két csoporton belül extrém módon nőtt.

25. táblázat A támogatások tevékenységtípus szerinti átlaga és szórása az egyes programokban

Kategória	FP7(fa)			H2020(fa)		
	n_j	\bar{x}_j	σ_{x_j}	n_j	\bar{x}_j	σ_{x_j}
felsőoktatási intézmény	45	246.171	238.622	19	541.116	1.000.496
magángazdaság	102	189.512	162.858	109	550.716	1.442.842
közintézmény	4	198.037	174.493	6	81.711	44.481
kutatási szervezet	65	347.842	370.782	33	354.246	253.035
egyéb résztvevő	15	199.856	164.749	21	51.385	76.961
Összesen	231	245.920	261.311	188	444.514	1.162.111

A belső szórásnégyzet aránya a teljes szórásnégyzethez képest (H_B^2) mindkét programban jóval magasabb volt, mint a külső szórásnégyzet aránya a teljes szórásnégyzethez képest (H_K^2). A részsokaságok (a tevékenység alapján képzett csoportok) tehát korántsem homogének. A támogatást tehát nem elsősorban a tevékenységtípus, hanem valamilyen egyéb tényező határozta meg mindkét támogatási programban. (26. táblázat)

26. táblázat A támogatások szórásai és szórásnégyzet-mutatói

Mutató	FP7(fa)	H2020(fa)
σ_{xB}	252.543	1.148.967
σ_{xK}	67.123	174.291
σ_x	261.311	1.162.111
H_B^2	0,934	0,978
H_K^2	0,066	0,022

Bár nem éreztük úgy, hogy ebben az esetben érdemes lehet outliereket keresni, mivel épp a fő összefüggést – a csoportokon belüli eltéréseket – tompította volna szükségtelen mértékben, indokolt lehet a projektekben részt vevő, de támogatást azokból nem kapó résztvevők adatait figyelmen kívül hagyni a számítások során. A 27. táblázat és a 28. táblázat azonban jól mutatja, hogy a fő összefüggést a megváltozott adatsorokkal végzett számítások sem módosítják.

27. táblázat A támogatások tevékenységtípus szerinti átlaga és szórása az egyes programokban (nullértékek nélkül)

Kategória	FP7(fa)			H2020(fa)		
	n_j	\bar{x}_j	σ_{x_j}	n_j	\bar{x}_j	σ_{x_j}
felsőoktatási intézmény	42	263.755	237.423	19	541.116	1.000.496
magángazdaság	98	197.247	161.491	89	674.472	1.570.395
közintézmény	4	198.037	174.493	6	81.711	44.481
kutatási szervezet	60	376.829	371.500	33	354.246	253.035
egyéb résztvevő	15	199.856	164.749	13	83.007	83.326
Összesen	219	259.396	261.782	160	522.304	1.243.467

Nem az egyes tevékenységekhez tartozás jelenti a tevékenységtípusok közötti támogatási különbségek fő okát, hanem az egyes csoportokon belüli különbségek kiugróan nagyok. A belső szórásnégyzet a teljes szórásnégyzet 91,5%-a, illetve 97,3%-a, míg a külső szórásnégyzet csupán a teljes szórásnégyzet 8,5%-át és 2,7%-át teszi ki.

28. táblázat A támogatások szórásai és szórásnégyzet-mutatói
(nullértékek nélkül)

Mutató	FP7(fa)	H2020(fa)
σ_{xB}	250.415	1.226.582
σ_{xK}	76.301	204.222
σ_x	261.782	1.243.467
H_B^2	0,915	0,973
H_K^2	0,085	0,027

Látható tehát, hogy a támogatási összegre a látszat ellenére minimális hatása van a csoporthoz tartozásnak, ha az egyes támogatási programokat külön-külön szemléljük.

A két támogatási program közötti eltolódás vizsgálata során a főátlagok standardizálását a nullértékek nélküli csoportokra végeztük el.

A főátlag különbsége ($K = 262.909$) és hányadosa ($I = 2,014$) a faipari projektrésztvevők átlagos támogatásában meglehetősen nagy növekedésre utal. Ez a jelentős átlagos támogatásbővülés alapvetően a részátlagok átlagos növekedéséből ($K' = 279.879$), illetve a részátlag-index jelentős változásából ($I' = 2,154$) fakad. Olyannyira, hogy az összetétel-változás ($K'' = -16.970$), illetve az összetételhatás-index ($I'' = 0,935$) egyenesen csökkenést mutat.

Tehát a H2020(fa) programban a faipari projektek támogatásának átlagos növekedése az FP7(fa)-hoz képest alapvetően egyes csoportok megugrott támogatására vezethető vissza (amelyről korábban láttuk, hogy a csoporton belül igen egyenetlenül oszlik meg). A csoportonkénti támogatott résztvevők számának átalakulása pedig mérsékelte az átlagos támogatás növekedését, mivel az alacsony támogatási összeggel rendelkező tevékenységek támogatott projektrésztvevőinek aránya nőtt a többi tevékenységhez tartozó résztvevőkhöz képest.

A fentiek alapján tehát a 13. szakmai hipotézist, miszerint a tevékenység, illetve az FP7(fa) és a H2020(fa) programokban kapott támogatás közötti kapcsolat jelentős, azonban a kategóriák támogatása közötti különbség nem az ismérv szerinti heterogenitásra, hanem a kategóriákon belüli heterogenitásra vezethető vissza, nem vetjük el. A faipari projektrésztvevők két program közötti átlagos támogatásnövekedése alapvetően a részátlagok növekedésére vezethető vissza.

7.3.8. Országokénti támogatás és az általános faipari teljesítmény kapcsolata

Az egyes országok faipari teljesítményét (kibocsátását) a négy legfontosabb faipari termékcsoporthoz tartozó termelés alapján becsültük meg. Ez természetesen nem azonos az összkibocsátással, de az egyes országok faipari kibocsátásának összemérhetőségét biztosítja (UNECE/FAO, 2022).

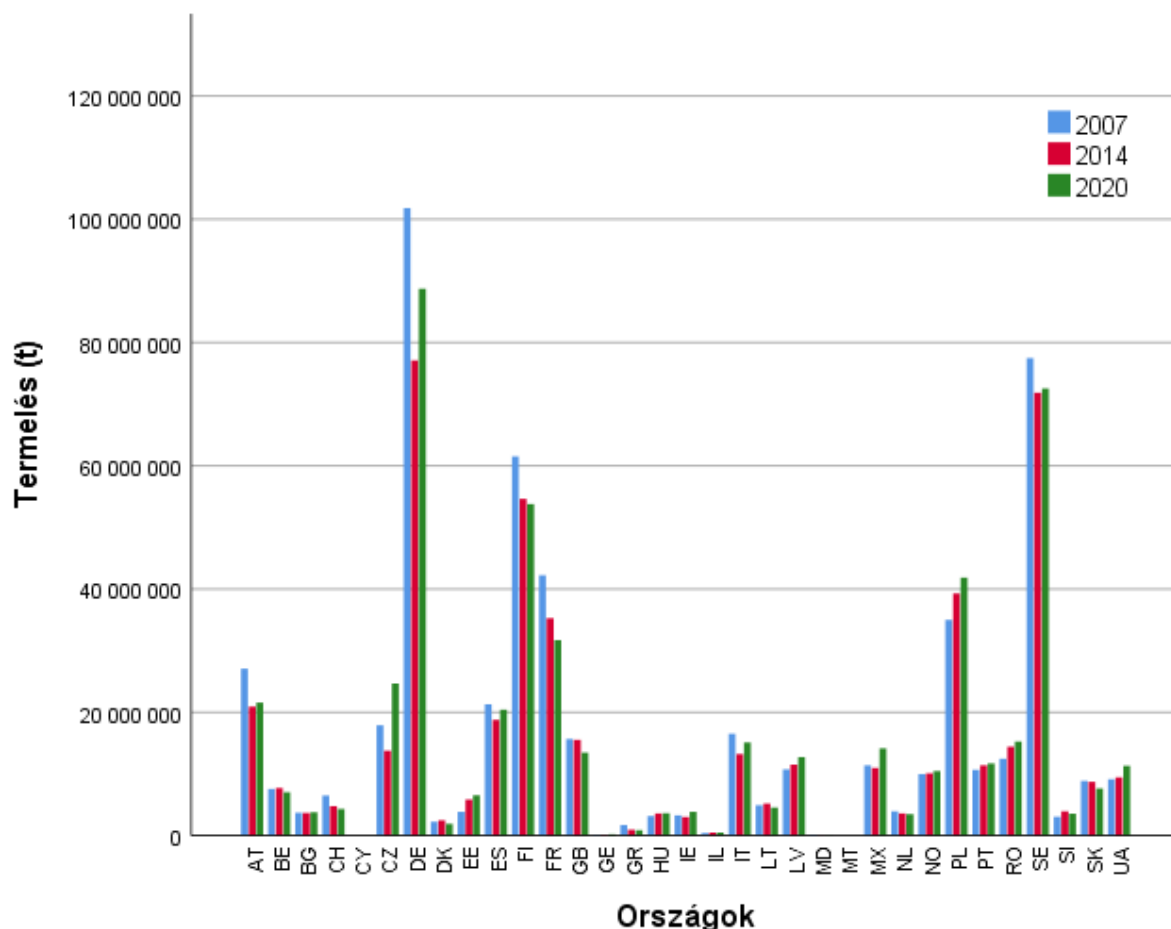
Nehézséget okozott, hogy a rendelkezésre álló hivatalos adatok három termékcsoporthoz tartozó – ipari hengeresfa, fűrészáru és faalapú panelek – csak köbméterben, egy termékcsoporthoz tartozó pedig – papír/papírkarton – csak tonnában álltak rendelkezésre (FAOSTAT, 2023). Az

átváltást célszerűbbnek tűnt volna egy esetben – a papír/papírkarton értékét köbméterre átváltva – elvégezni, azonban megbízható átváltási arány a másik három esetben állt rendelkezésre. Az átváltási arányt a puhafa-keményfa szorzó átlagolásával számoltuk ki. (Úgy ítéltük meg, hogy ennél alaposabb, súlyozott módszer itt most felesleges.)

Felmerülhet a kérdés, hogy miért nem a faipari termelés gazdasági adataiból, pl. a GDP-ből való faipari részesedés alapján tettünk összehasonlítást.

Ezt egyfelől a nem egységes gazdaságstatisztikai rendszerek és csoportosítások (a statisztikai kimutatások a faiparhoz gyakran az erdészetet és egyéb ágazatokat is „hozzácsapnak”), másfelől a nemzetgazdasági vásárlóerő-paritás nem minden ágazatban egyformán érvényesülő jellege indokolta.

Az ipari hengeresfa egy köbméterét 0,75 tonnának vettük, fűrészáru esetén ugyanez az érték 0,62 tonna, faalapú panelek esetén pedig 0,60 tonna (Forest Research, 2023). A számítások eredményét a 57. ábra mutatja.

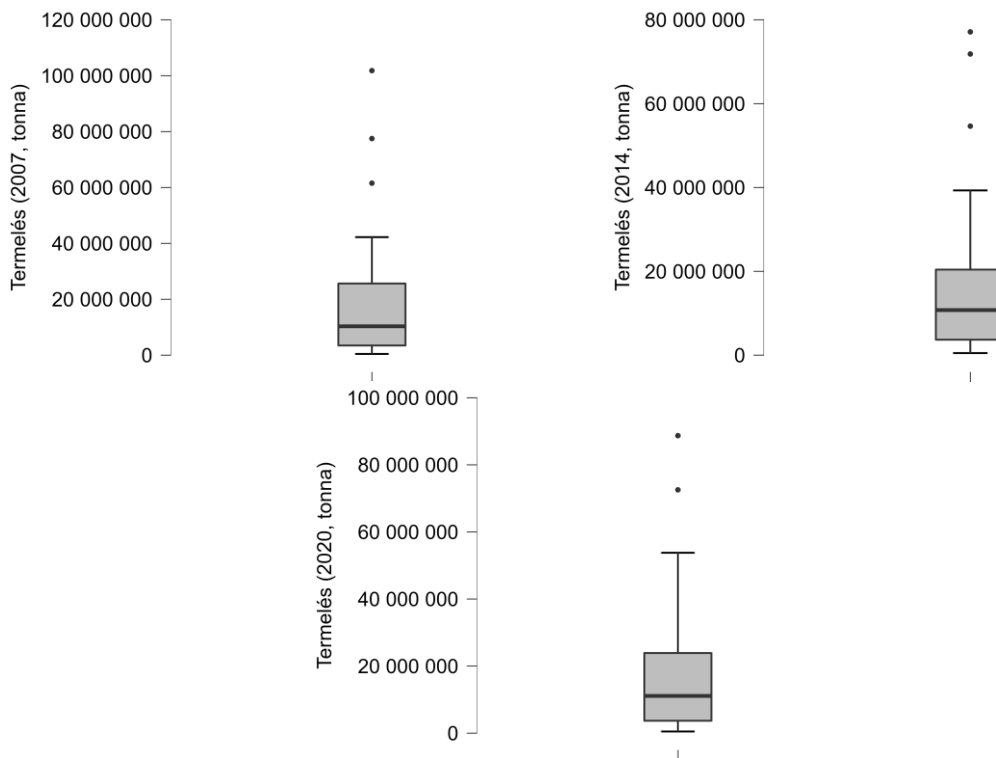


57. ábra Az FP7 és a H2020 programokból támogatást kapott faipari projektek résztvevőikhez tartozó országok becsült össztermelése négy kiemelt termékcsoportban

A termelésben mutatkozó hatalmas különbségek előrevetítik, hogy az adatsorban outlierok lehetnek. Az outlierok vizsgálatánál – mivel az elemzés az FP7 és H2020 programokban kapott támogatással való összevetést szolgálja – már eleve nem vettük figyelembe azokat az országokat, amelyek valamelyik programból nem kaptak támogatást, ezért a vizsgált adatok körét

22 országra csökkentettük. Ha támogatási programonként, külön-külön vizsgáltuk volna a pozitív támogatást, akkor kissé más összetételű országcsoportokat hasonlítottunk volna össze, s ez nem lett volna szerencsés.

A 58. ábra alapján úgy tűnik, hogy mindhárom adatsorban vannak outlierok.



58. ábra A becsült faipari termelés boxdiagramjai

Az egyes országok támogatásbeli különbségeinek outlierjeit a korábbiak alapján nem töröljük.

A 2007-es termelés esetén $I(8)(9)(10)$ rendre $[-29.801.542; 58.901.043]$, $[-22.621.810; 43.268.852]$, $[-45.324.494; 74.423.996]$.

A 2014-es termelés esetén $I(8)(9)(10)$ rendre $[-21.391.903; 45.485.770]$, $[-22.319.265; 43.819.633]$, $[-33.095.496; 57.189.363]$.

A 2020-as termelés esetén $I(8)(9)(10)$ rendre $[-26.604.984; 54.195.170]$, $[-26.277.290; 48.482.561]$, $[-40.745.010; 68.335.197]$.

A 2007-es termelési adatok alapján (8) szerint három, (9) szerint szintén három, (10) szerint két outlier szűrhető ki. A 2014-es termelési adatok alapján mindhárom módszerrel ugyanezeket az eredményeket kaptuk. A 2020-as termelés alapján rendre két, három és két outlier lehet kiszűrni. Mivel becsült termelés alapján minden évben Németország, Svédország, Finnország a sorrend, tehát az outlierok törlése a legnagyobb termelők törlését jelentené. Ez meglehetősen nagy adatvesztésnek tűnik, így az outlierok esetleges törléséről a kétdimenziós outlierok alapján döntünk. Az outlierok törlése ellen szól, hogy mindhárom adatsornál ugyanazok, ugyanolyan sorrendben adják a legnagyobb kibocsátást, aligha tekinthetjük ezeket az országokat valamilyen anomáliának.

A vizsgált adatok a Shapiro-Wilk és az egymintás Kolmogorov-Szmirnov normalitási próba alapján sem normális eloszlásúak (29. táblázat), mivel $p < 0,05$. A normalitást feltételező H_0 -t tehát elvetjük.

A normalitási próbák eredményei alapján arra a feltételezésre jutottunk, hogy a három adatsor azonos eloszlást követ, azaz azonos eloszlásból származik.

29. táblázat Az országonkénti becsült faipari termelés normalitáspróbái

Mutató	Faipari termelés		
	2007	2014	2020
Shapiro-Wilk (W)	0,750	0,760	0,769
P (Shapiro-Wilk)	$9,006 \times 10^{-5}$	0,0001	0,0002
Kolmogorov-Szmirnov (D)	1	1	1
P (Kolmogorov-Szmirnov)	$3,331 \times 10^{-16}$	$3,331 \times 10^{-16}$	$3,331 \times 10^{-16}$

A 30. táblázat alapján megállapítható, hogy mindhárom adatsor azonos eloszlást követ, a magas p érték miatt legalábbis az erre vonatkozó $H_0(6)$ hipotézist nem vetjük el.

30. táblázat Faipari termelési adatsorok azonos eloszlására vonatkozó, kétmintás Kolmogorov-Szmirnov próba eredményei

Adatsorok	Kolmogorov-Szmirnov D	Kolmogorov-Szmirnov p
Termelés (2007, 2014)	0,136	0,990
Termelés (2007, 2020)	0,136	0,990
Termelés (2014, 2020)	0,091	1,000

Mivel a normalitás nem teljesül – bár nem tartozik fő kutatási céljainak közé – érdemes megvizsgálunk, hogy a három adatsor között a Spearman-féle rangkorrelációs együttható milyen értéket vesz fel.

A 31. táblázat szerint minden termelési adatsor között igen szoros rangkorreláció mutatkozik.

31. táblázat Spearman-féle rangkorrelációs együttható a faipari termelés adatsorai között

Adatsorok	Spearman-féle ρ	Spearman-féle p
Termelés (2007, 2014)	0,985	$2,956 \times 10^{-6}$
Termelés (2007, 2020)	0,989	$2,838 \times 10^{-6}$
Termelés (2014, 2020)	0,983	$3,033 \times 10^{-6}$

A későbbiekben a faipari termelést össze kívánjuk hasonlítani az egyes programokban kapott támogatással, s ez mindkét program esetén két-két összehasonlítást igényelne. Az azonos eloszlás és a magas korreláció miatt azonban a 2007-es és a 2014-es, illetve a 2014-es és 2020-as termelés minden további nélkül – jelentősebb információvesztés nélkül – átlagolható, ezért az említett átlagokból létrehoztuk a Termelés(A) és a Termelés(B) adatsorokat. Így feleannyi számításra van szükségünk.

A későbbi számítások során az FP7(fa) projektek országonkénti közösségi támogatását a Termelés(A), a H2020(fa) projektek országonkénti közösségi támogatását a Termelés(B) adataival

hasonlítjuk össze korreláció és regressziószámítással. A regressziószámítás során a termelési adatok lesznek a független, a támogatási adatok a függő változók.

A Mahalanobis-távolság minden esetben kisebb, mint 13,82. (FP7(fa) támogatás – Termelés(A): 7,91, H2020(fa) támogatás – Termelés(B): 7,41.) Tehát a Mahalanobis-távolság alapján nincs többdimenziós outlier. A Cook-távolság maximális értéke egyik oszloppárban sem nagyobb, mint egy. (FP7(fa) támogatás – Termelés(A): 0,19, H2020(fa) támogatás – Termelés(B): 0,28) Többdimenziós outlier tehát nem akadályozza a korrelációszámítás és a regressziószámítás elvégzését egyik esetben sem.

A Spearman-féle rangkorrelációs együttható (ρ) FP7(fa) támogatás – Termelés(A) viszonylatban 0,57, H2020(fa) támogatás – Termelés(B) viszonylatban pedig 0,70. Tehát első esetben közepes, második esetben közepesnél erősebb korrelációról beszélhetünk. Ha a sokasági statisztikán túl szeretnénk lépni, akkor p értéke alapján - rendre 0,007 és 0,0004 – elvetjük a függetlenségre vonatkozó nullhipotézisünket, mivel a kapott értékek kisebbek az 5%-os szignifikanciaszintnél.

A 14. szakmai hipotézist tehát nem vetjük el, az FP7 és H2020 programokból kapott országokénti támogatás és a faipari kibocsátás között összefüggés van.

FP7(fa) támogatás – Termelés(A) viszonylatban a lineáris regresszióegyenes, ill. a becslőfüggvény paraméterei: $\hat{\beta}_0 = 1,15 \times 10^6$, $\hat{\beta}_1 = 6,42 \times 10^{-2}$, $s = 1,86 \times 10^6$, $s^2 = 3,47 \times 10^{12}$, $s_{\hat{\beta}_0} = 5,21 \times 10^5$, $s_{\hat{\beta}_1} = 0,016$.

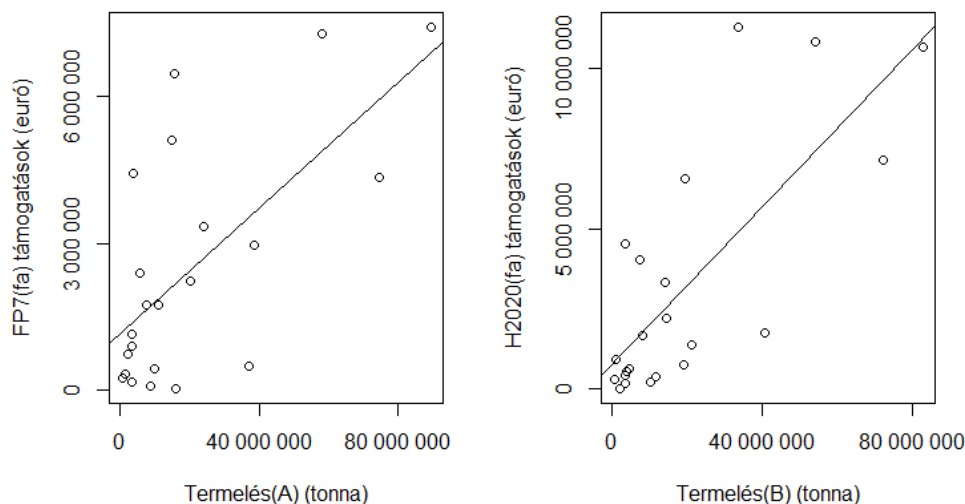
A Student-féle t -próba értékei: $t = 3,897$, $t_{0,975}(df = 22) = 2,074$. F -próba p értéke: 0,00. Mivel a számított t érték a kritikus tartományban van, illetve $p < \alpha$, szignifikáns lineáris összefüggést feltételezünk a változók között. A függetlenséget feltételező $H_0(12)$ -t ezért elutasítjuk.

H2020(fa) támogatás – Termelés(B) viszonylatban a lineáris regresszióegyenes, ill. a becslőfüggvény paraméterei: $\hat{\beta}_0 = 7,79 \times 10^5$, $\hat{\beta}_1 = 0,12$, $s = 2,49 \times 10^6$, $s^2 = 6,19 \times 10^{12}$, $s_{\hat{\beta}_0} = 7,01 \times 10^5$, $s_{\hat{\beta}_1} = 0,023$.

A Student-féle t -próba értékei: $t = 5,245$, $t_{0,975}(df = 22) = 2,074$. Az F -próba p értéke: 0,00. Mivel a számított t érték a kritikus tartományban van, illetve $p < \alpha$, szignifikáns lineáris összefüggést feltételezünk a változók között. A függetlenséget feltételező $H_0(12)$ -t ezért elutasítjuk.

A fenti próbák értelmezése – a korábban említett esetekhez hasonlóan – sokaság vagy szuper-sokaság feltételezése mellett végezhető el, de a lineáris regresszió becslt paraméterei enélkül is érvényesek.

A 59. ábra alapján a lineáris regresszió nem tűnik meggyőzőnek egyik esetben sem.



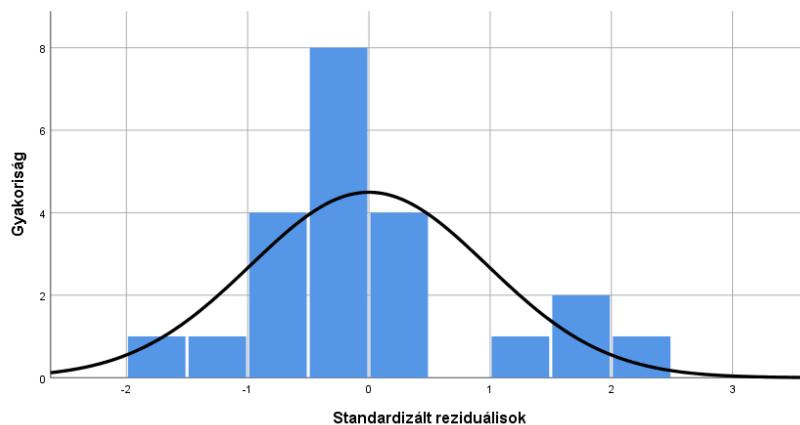
59. ábra Termelés és közösségi támogatás kapcsolata

A determinációs együttható (R^2) értéke FP7(fa) támogatás – Termelés(A) viszonylatban 0,432 (a prediktorok számához igazított értéke 0,403). Az eredmény közepesnél gyengébb erősségű kapcsolatra utal. H2020(fa) támogatás – Termelés(B) viszonylatban ugyanezek az értékek 0,579, illetve 0,558.

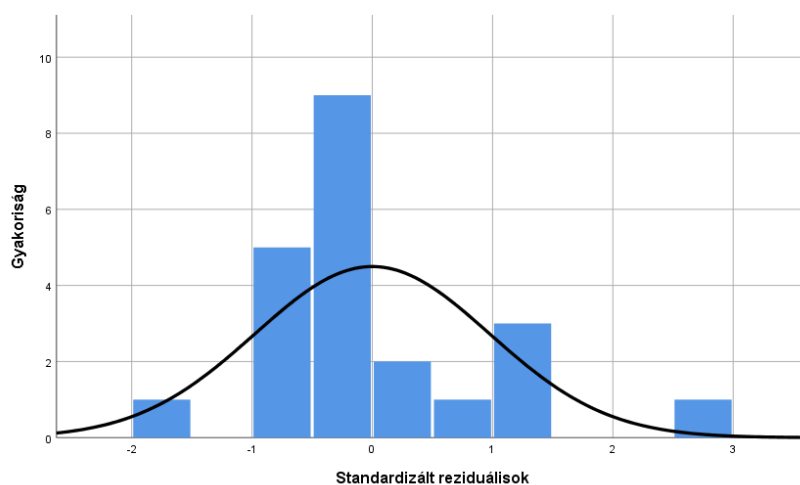
A 60. ábra és az 61. ábra alapján nehéz eldönteni, hogy az eloszlás megfelel-e a normalitás statisztikai követelményeinek, de mintha a normális eloszlás nem lenne teljesen kizárható. Az FP7(fa) támogatás – Termelés(A) viszonylatban négyből két próba esetén, H2020(fa) támogatás – Termelés(B) viszonylatban egy próba esetén $p > \alpha$. A rezualisok normalitására vonatkozó $H_0(13)$ -at azért nem utasítjuk el, mert kis elemszámú minták esetén általában a Kolmogorov-Szmirnov próba a legmegbízhatóbb, s p értéke itt mindkét esetben magasabb, mint a szignifikanciaszint. (32. táblázat)

32. táblázat A reziduálisok normalitásvizsgálata

Próba	FP7(fa) tám. – Termelés(A)		H2020(fa) tám. – Termelés(B)	
	Próbastatisztika	p	Próbastatisztika	p
Shapiro-Wilk	0,930	0,122	0,908	0,044
Kolmogorov-Szmirnov	0,161	0,564	0,222	0,195
Cramér-von Mises	2,242	0,000	2,561	0,000
Anderson-Darling	0,688	0,062	0,900	0,017



60. ábra Az FP7(fa) támogatás – Termelés(A) lineáris regresszió reziduálisainak normalitásvizsgálata



61. ábra A H2020(fa) támogatás – Termelés(B) lineáris regresszió reziduálisainak normalitásvizsgálata

A homoszkedaszticitásra vonatkozó Breusch-Pagan és White-próba értékei alapján FP7(fa) támogatás – Termelés(A) viszonylatban nem vetjük el a homoszkedaszticitásra vonatkozó $H_0(14)$ hipotézist. A próbák alapján nem egyértelmű, hogy H2020(fa) támogatás – Termelés(B) viszonylatban is így járjunk-e el. A próbák közé a fentiek miatt, ill. a két népszerű próba kis mintánál nem kielégítő megbízhatósága miatt a korábbiakon túl a kis mintánál is megbízható, de összetettebb paraméterezést igénylő Goldfeld-Quandt próbát is felvettük. (33. táblázat)

33. táblázat A homoszkedaszticitásra vonatkozó próbák

Próba	FP7(fa) tám. – Termelés(A)		H2020(fa) tám. – Termelés(B)	
	Próbastatisztika	p	Próbastatisztika	p
Breusch-Pagan (st.)	0,033	0,855	1,721	0,189
White	2,420	0,298	9,200	0,010
Goldfeld-Quandt	3,602	0,056	5,675	0,017

A kapott p értékek alapján FP7(fa) támogatás – Termelés(A) viszonylatban a homoszkedaszticitásra vonatkozó $H_0(14)$ -t nem utasítjuk el, ellenben H2020(fa) támogatás – Termelés(B) viszonylatban a hipotézist nem fogadjuk el, így – az érdekes módon – magasabb determinációs

együtthatóval rendelkező kapcsolat linearitása erősen kérdéses. Mindenképpen érdemes tehát a többi lehetséges regresszió paramétereit is megvizsgálni. Az 34. táblázatban csak a szignifikáns eredmények láthatóak.

34. táblázat Támogatások és termelés közötti regressziós kapcsolat

Regressziós modell	FP7(fa) tám. – Termelés(A)		H2020(fa) tám. – Termelés(B)	
	R ²	Igazított R ²	R ²	Igazított R ²
Logaritmikus	0,370	0,339	0,446	0,419
Parabolikus	0,432	0,372	0,592	0,542
Harmadfokú	0,442	0,349	0,594	0,526
Vegyes	0,232	0,194	0,413	0,384
Hatványkitevős	0,226	0,187	0,441	0,414
Növekedési	0,232	0,194	0,413	0,384
Exponenciális	0,232	0,194	0,413	0,384
Logisztikus	0,232	0,194	0,413	0,384

FP7(fa) támogatás – Termelés(A) viszonylatban semmi sem indokolja, hogy a közepesnél gyengébb lineáris regresszió helyett más regresszió érvényességét vizsgáljuk. H2020(fa) támogatás – Termelés(B) esetén a közepesnél erősebb paraméterekkel számolt lineáris regressziót nem tekintettük érvényesnek, talán a parabolikus vagy a harmadfokú lehet a lineáris regresszió alternatívája, azonban sokkal valószínűbb, hogy számos egyéb tényező (pl. szakmai szempontok, gazdasági fejlettség, lobbierdekek) határozzák meg a támogatások országokénti szintjét.

A korrelációnál tett megállapításokat tehát annyiban módosítjuk, hogy a 14. szakmai hipotézis – tehát, hogy az FP7 és H2020 programokból kapott országokénti támogatás és a faipari kibocsátás között összefüggés van – elutasításának hiánya ellenére, regresszióról a korrelációtól eltérően nem beszélhetünk.

7.4. A faipari projektháló elemzésének eredményei

7.4.1. A kapcsolatháló felrajzolása

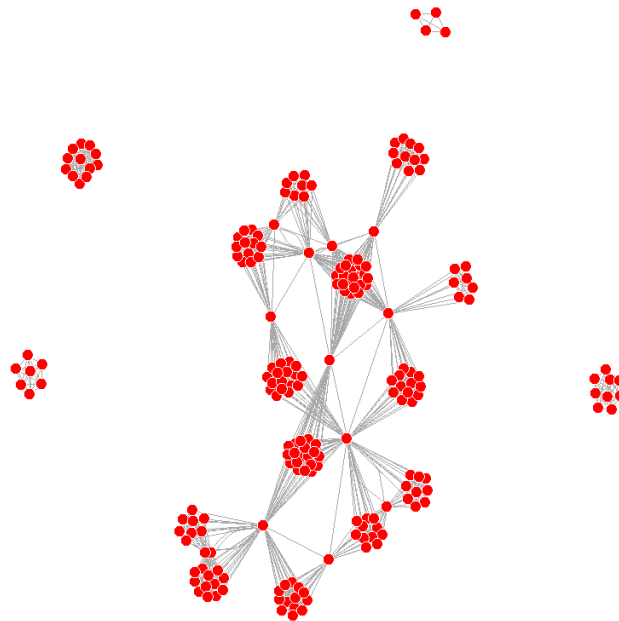
A 6.4.1.1. fejezet alapján kiszámolt értékek és a felrajzolt ábrák a következők:

A hálózatkutató célra használt, tisztított, csak a hálózati kapcsolatokkal rendelkező csomópontokat tartalmazó adatállomány legfontosabb adatait a 35. táblázat mutatja. A szűkebben és tágabban értelmezett faipari projektekre is igaz, hogy az FP7-hez képest több projektre – bár a többszereplős projektek száma alig változott – kevesebb projektrésztvevő jutott. A csomópontok száma az egyszereplős projektek negligálása és a gyakran egy projektrésztvevőre eső több projektrésztvétel miatt természetesen alacsonyabb, mint a projektrésztvételek összesített értéke.

A szomszédsági mátrixból kiindulva a projektrésztvevők között kapcsolathálókat a 62. ábra-65. ábra mutatja.

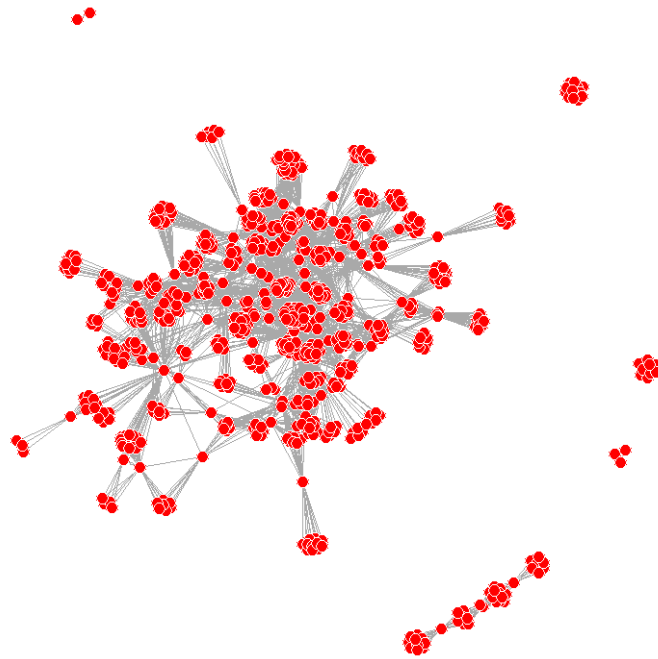
35. táblázat *Fapari projektháló leíró adatai*

Adatsor	Projektek száma	Többszereplős projektek száma	Csomópontok száma
FP7(fa)	19	17	209
FP7(fa2)	102	84	830
H2020(fa)	36	15	149
H2020(fa2)	197	80	747



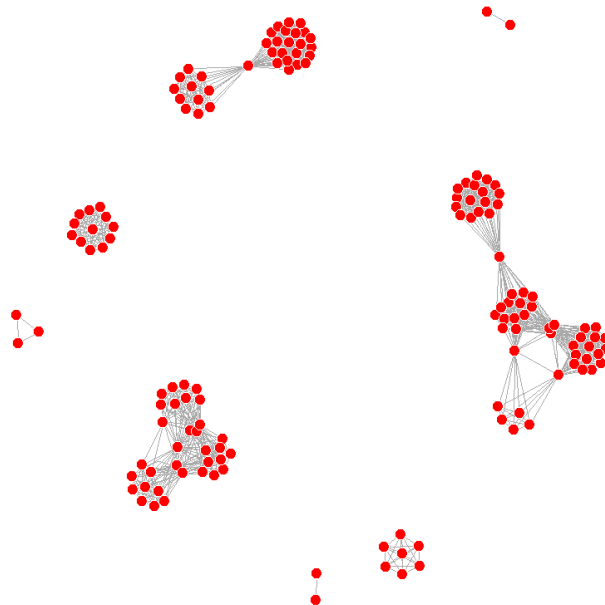
62. ábra A FP7(fa) projektrésztevők kapcsolathálója

A 62. ábra alapján látható, hogy az FP7-ben a faipari projektek résztvevői között már a faipari projektek által közvetített kapcsolatokon keresztül viszonylag összefüggő háló alakult ki. A csak részben faipari projektek kapcsolatainak hálóba illesztése után nehéz eldönteni, hogy az „összekapcsoltság” nőtt vagy csökkent (63. ábra).



63. ábra A FP7(fa2) projektrésztvevők kapcsolathálója

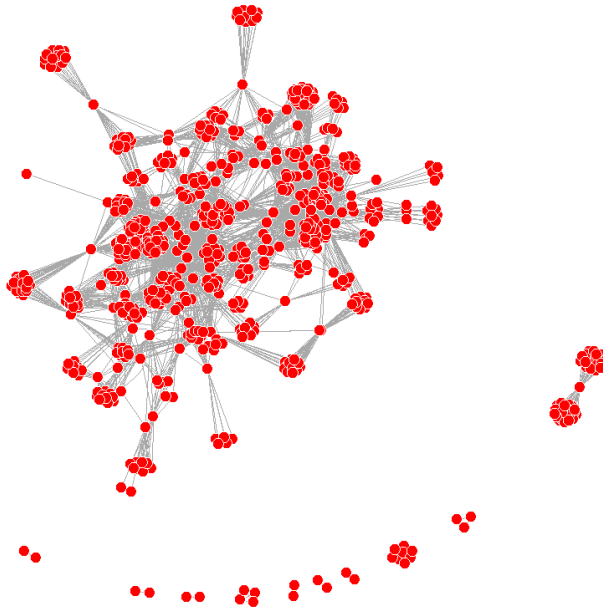
A H2020(fa) hálózatban a háló jóval szétterjedtebbnek tűnik, mint az FP7(fa) hálóban. (64. ábra)



64. ábra A H2020(fa) projektrésztvevők kapcsolathálója

A 65. ábra alapján úgy tűnik, hogy a csak részben faipari jellegű projektek – gyenge kapcsolat-ként – az eredeti besorolás alapján szétterjedt hálót jóval „összekapcsoltabbá” tették.

A fenti megállapítások azonban csak vizuális benyomásokat tükröznek, kérdés, hogy a különböző hálózati mutatók a megállapításainkat megerősítik-e.



65. ábra A H2020(fa2) projektrésztevők kapcsolathálója

A felrajzolt kapcsolatháló tehát önmagában keveset árul el a hálózat jellegéről. Annyi azért még megállapítható, hogy mindegyik mintában vannak periférikus csoportok, résztvevők is. Ez az egyes mutatók értékelésénél fontos lehet.

7.4.2. A hálózati mutatók számított értékei

A 6.4.1.2. fejezetben részletezett hálózati mutatók értékeit a 36. táblázat mutatja.

36. táblázat A hálózati mutatók értékei

Adatsor	Sűrűség	Tranzitivitás	Átmérő
FP7(fa)	0,08	0,80	4
FP7(fa2)	0,02	0,59	6
H2020(fa)	0,11	0,88	3
H2020(fa2)	0,28	0,64	6

Az (19)(20) egyenletre épülő sűrűség értékei alapján annyit állíthatunk, hogy a H2020(fa) projekthálóban több kapcsolat jött létre a lehetséges kapcsolatokból, mint az FP7(fa) hálóban. Ugyanez érvényes H2020(fa2) hálóra is az FP7(fa2) hálójához képest. A csak részben faipari jellegű projektek adataival kiegészítve a sűrűség érték H2020 esetében nőtt, ez arra utal, hogy ezeknek a projekteknek az elemzésbe vonása – legalábbis hálózatkutatási szempontból – teljesebb képet adhat a valós együttműködésről. Ez előrelépés az FP7-hez képest, ahol a részben faipari projektekkel kiegészített projektháló – FP7(fa2) – sűrűsége csökkent FP7(fa) projekthálójához képest.

A (21)(22) egyenlet alapján számolt tranzitivitás értéke meglehetősen alacsony. Ha csak a kapott eredményre koncentrálunk, akkor úgy tűnhet, hogy a „barátom barátja a barátom” jelenség a faipari projekthálók közül leginkább a H2020(fa) hálóban érvényesül. Ez azonban erősen kétséges, mivel a háló erősen széttörözött. Tehát ez a megállapítás esetünkben alapvetően az

egy részhálókra igaz, mivel az alkalmazott algoritmus nem tudja kezelni, hogy az egyes csomópontokból nem érünk el minden csomópontot.

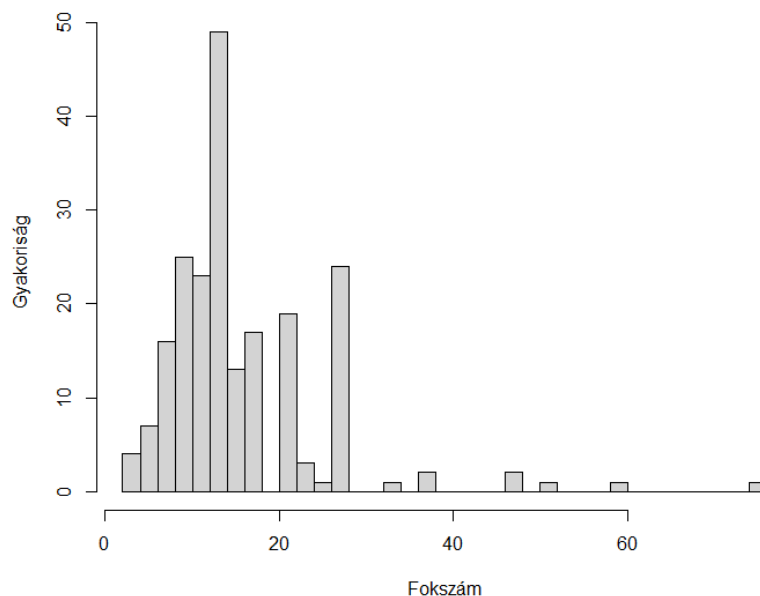
Megoldást jelentene, ha az egyes, nem kapcsolódó részhálókat különválasztanánk, de a teljes hálóra ilyen széttöredezett háló vagy hálók esetén nagyon nehéz bármilyen egzakt következtetést levonnunk.

Hasonlóképpen óvatosan kell fogadnunk az átmérőre vonatkozó következtetést, hiszen az átmérő a legtöbb esetben lényegében végtelenül nagy. Az algoritmus azonban itt is csak az elérhető csomópontokat veszi figyelembe. Tehát az érték az összefüggő hálózatrészek maximális átmérőjének háló szintű maximuma.

Eljutottunk a szakirodalmi elemzésnél már említett problémához. Ha egzakt hálózat kutatási eredményeket szeretnénk kapni, akkor csak egy vagy több projekt összekapcsolt csomópontjait kutathatjuk korlátozott érvényű eredménnyel. Ha szélesebb körben vizsgálódunk, az eredményeink a háló széttöredezettségével arányosan egyre kevésbé lesznek megbízhatóak. Kompromisszumos választásnak tűnik a hálózatok legfontosabb klasztereinek elemzése, erre a későbbiekben látunk még példát.

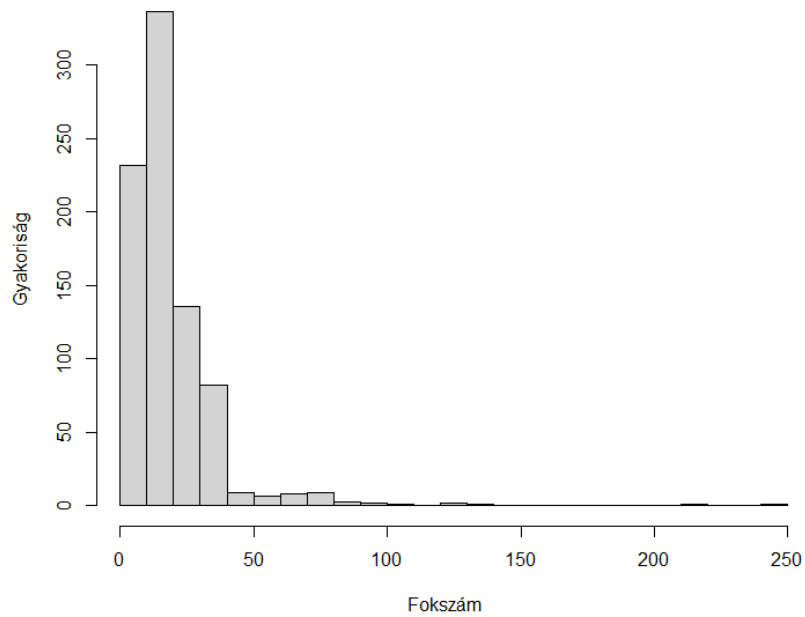
A (23) egyenlet alapján számított fokszámokat a 66-69. ábra mutatja.

A 66. ábra alapján a hálózat kutatásban nevezetes eloszlások nem vélelmezhetőek, a fokszámok eloszlása leginkább lognormális eloszlásnak tűnik.



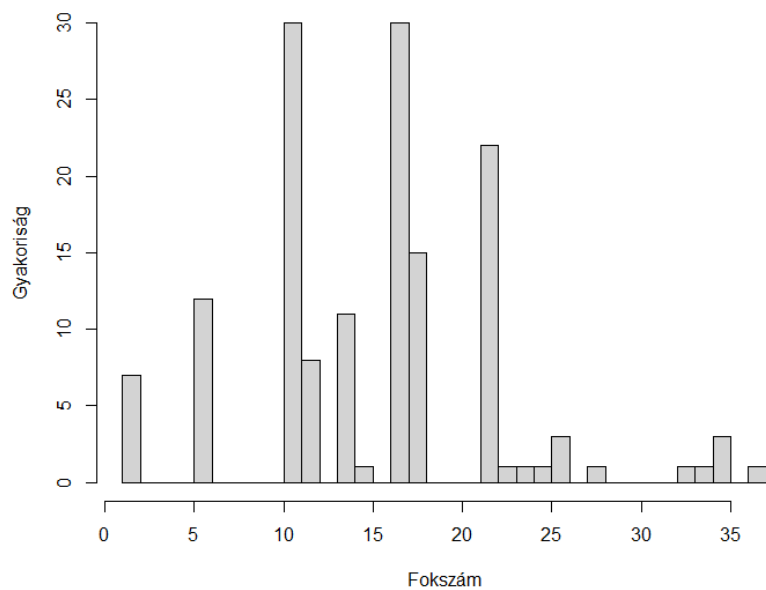
66. ábra FP7(fa) hálójának fokszámai

A 67. ábra által mutatott eloszlás hatványfüggvény-eloszlást sugall.



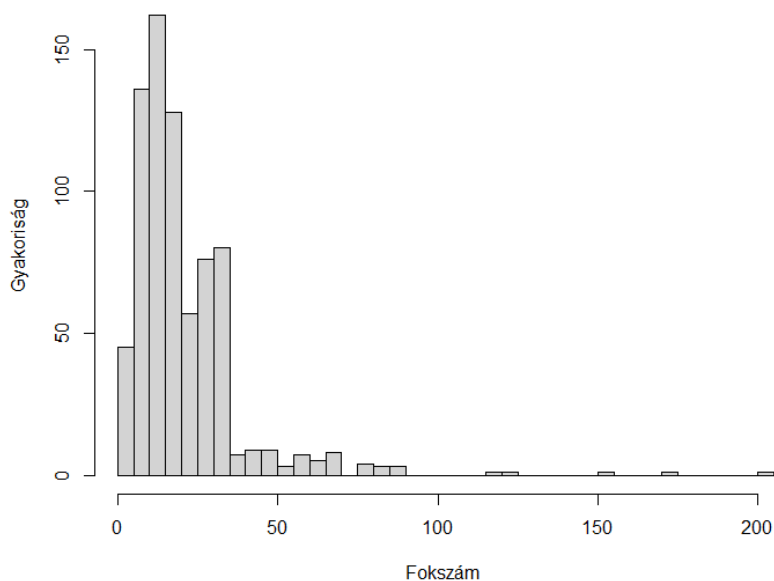
67. ábra FP7(fa2) hálójának fokszámai

A 68. ábra alapján lényegében lehetetlen bármilyen egyértelmű következtetést levonni a fokszámok eloszlására, esetleg a Poisson- (vagy a normális) eloszlás lehet óvatos feltételezés.



68. ábra H2020(fa) hálójának fokszámai

Talán érdekes, hogy az FP7(fa2) és a H2020(fa2) fokszámeloszlása némi hasonlóságot mutat az FP7(fa) fokszámeloszlásával. (69. ábra)



69. ábra H2020(fa2) hálójának fokszámai

Természetesen a tényleges eloszlások, illetve az eloszlások azonosságai csak próbákkal ellenőrizhetők. Kétmintás Kolmogorov-Szmirnov próbával azonban minden esetben megcáfoltuk, hogy a foksámeloszlások azonos eloszlásból származnának (p egyetlen esetben sem érte el a 0,001-et).

7.4.3. „Nevezetes” hálózatelméleti eloszlások vizsgálata

7.4.3.1. Poisson-eloszlás

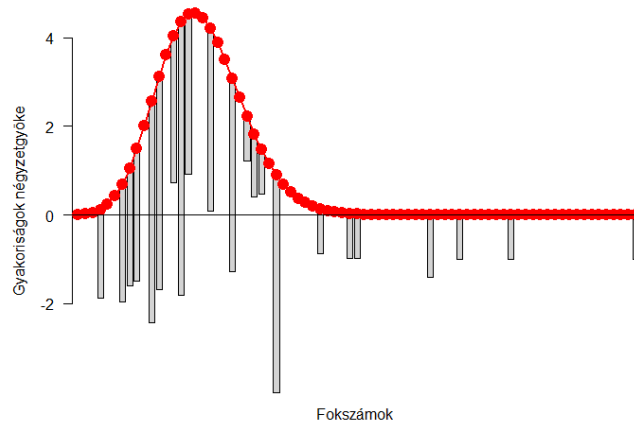
A (32)-(35) alapján elvégzett maximum likelihood becslés eredményét a 37. táblázat mutatja.

37. táblázat Fokszámok Poisson-eloszlására végzett maximum likelihood becslés eredményei

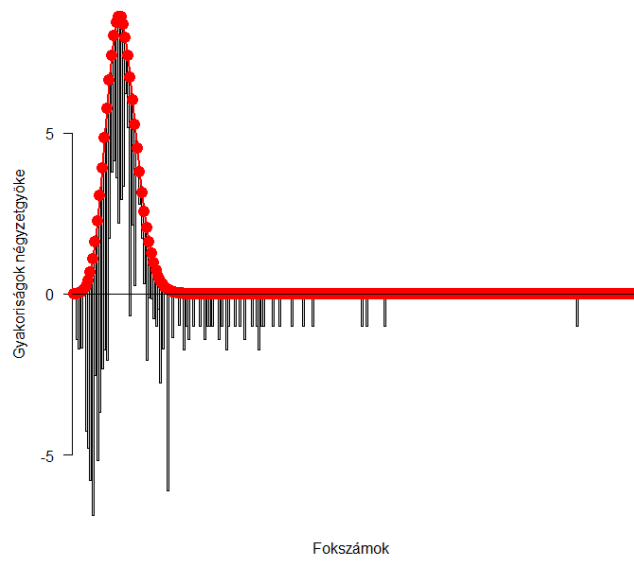
Adatsor	G^2	df	p	λ	α	c_{krit}
FP7(fa)	783,49	20	0	16,21	0,05	31,41
FP7(fa2)	7.161,77	68	0	19,89	0,05	88,25
H2020(fa)	476,03	17	0	15,62	0,05	27,59
H2020(fa2)	6.373,53	65	0	21,25	0,05	84,82

$G^2 > c_{krit}$ és $p < \alpha$ mind a négy esetben, ezért Poisson-eloszlásra vonatkozó $H_0(31)$ -t elvetjük.

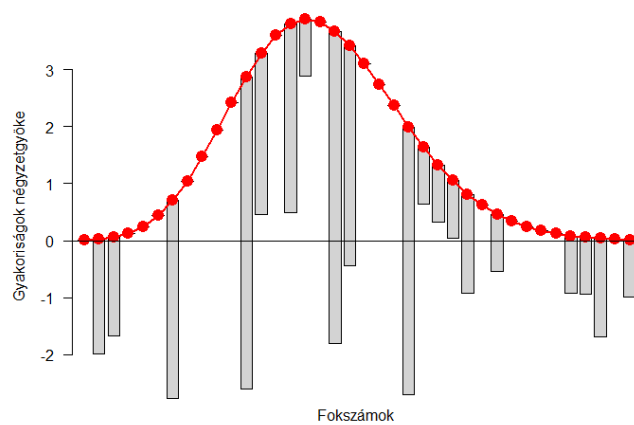
A foksámok gyakorisági eloszlásai tehát egyik esetben sem Poisson-eloszlásúak. Eredményeinket a rootogramok is megerősítik, amelyek megmutatják, hogy a tapasztalati értékeinket mennyivel kellene eltolnunk ahhoz, hogy a kívánt eloszlást kapjuk. (70. ábra – 73. ábra)



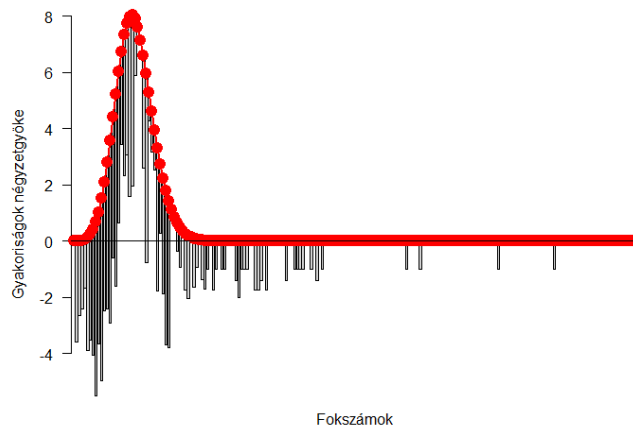
70. ábra FP7(fa) háló fokszámainak rootogramja Poisson-eloszlásra



71. ábra FP7(fa2) háló fokszámainak rootogramja Poisson-eloszlásra



72. ábra H2020(fa) háló fokszámainak rootogramja Poisson-eloszlásra



73. ábra H2020(fa2) háló fokszámainak rootogramja Poisson-eloszlásra

7.4.3.2. Hatványfüggvény-eloszlás

A (36)-(39) egyenletekkel vázolt illeszkedésvizsgálat (hatványfüggvény-eloszlás) során kapott értékeket az 38. táblázat mutatja.

38. táblázat Hatványfüggvény-eloszlás próbájának értékei

Adatsor	$\alpha(\gamma)$	x_{min}	p
FP7(fa)	3,45	13	0,07
FP7(fa2)	3,95	62	1
H2020(fa)	6,03	17	0,02
H2020(fa2)	4,40	58	0,99

FP7(fa) hálójának fokszámai esetén, 5%-os szignifikanciaszint és $x \geq 13$ mellett nem vetjük el a hatványfüggvény-eloszlásra vonatkozó $H_0(38)$ -t, mivel p kis mértékben meghaladja a szignifikanciaszintet. Azaz ebben a tartományban a minta hatványfüggvény-eloszlásból való származását nem utasítjuk el, bár az eredmény nem különösebben meggyőző.

FP7(fa) hálójától eltérően H2020(fa) hálójának fokszámai még a próbához tartozó algoritmus által generált x_{min} -nél nagyobb vagy egyenlő értékeket tartalmazó tartományban sem követnek hatványfüggvény-eloszlást, $H_0(38)$ tehát elvethető.

FP7(fa2) és H2020(fa2) hálójának fokszámai viszont magas x_{min} mellett magas p értéket vesznek fel.

FP7(fa2) és H2020(fa2) esetén tehát kijelenthetjük, hogy a hálót a skálafüggetlenség felé mutató hatások érik.

Tehát az eddigiek alapján a 15. szakmai hipotézis, miszerint a projektrésztvevők kapcsolathálóiban mért fokszámok eloszlásai skálafüggetlenek, azaz kevés résztvevőnek van sok és sok résztvevőnek kevés kapcsolata, nem utasítható el egyértelműen. Ez egyelőre kevés a viszonylag egyértelmű tézis megfogalmazásához.

A fenti értékeket a 6.4.1.2. fejezetben jelzett alternatív eljárással is megerősítettük.

Az 1000 iterációs bootstrapping(p) módszerrel kapott értékeket a 39. táblázat mutatja. (Ha az *R powerLaw* csomag x_{min} becslésére irányuló bootstrapping-módszerével dolgoznánk, magas

iterációszám mellett általában visszakapnánk az előző próba eredményeit.) A sorozatos mintavételek elemszámainak meghatározását az R *powerLaw* csomagjának algoritmusára bíztuk. A 74. ábra – 77. ábra a nagy ingadozást mutató első 100 iteráció eredménye nélkül mutatja az eredményhez vezető folyamatot.

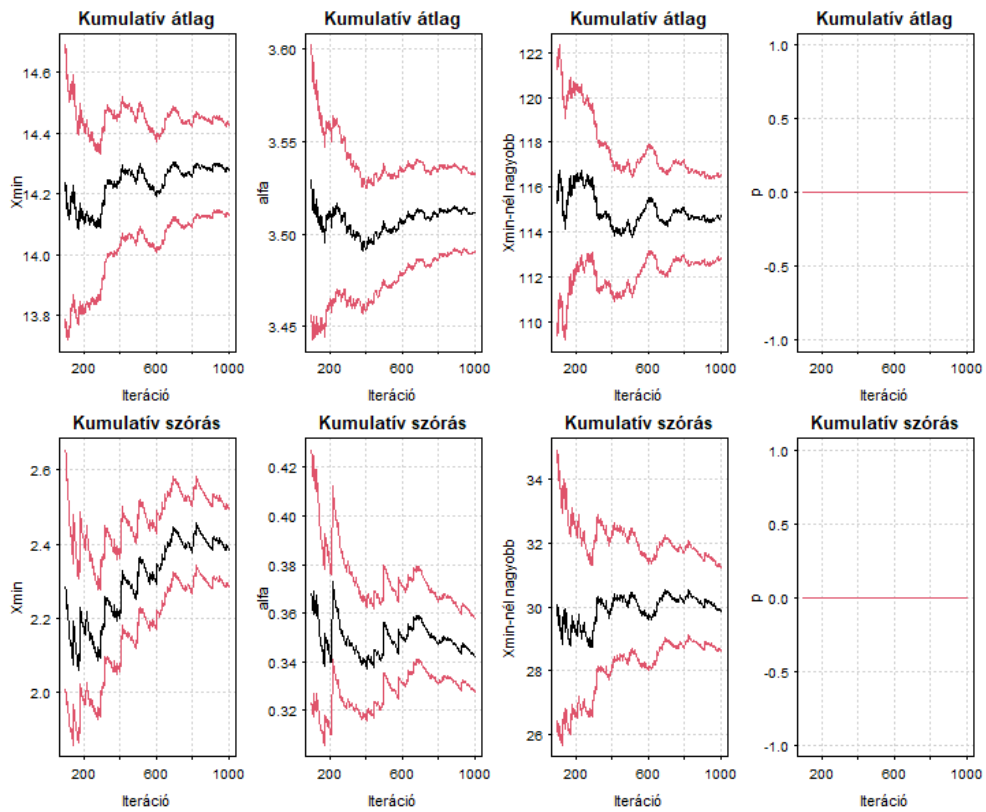
Bár a bootstrapping(p) módszer jellemzően kisebb p és kisebb x_{min} értéket eredményez, mint a hagyományos próbák, FP7(fa) és H2020(fa) esetében – valószínűleg a hatványfüggvény-eloszlásra nyomaikban is alig emlékeztető eloszlások miatt – nem teljesült. Viszont az FP7(fa2) és H2020(fa2) adatsorokra igen.

Tehát egy bizonyos x_{min} érték felett a szigorúan vett faipari projekteket kiegészítő – valószínűleg a gyenge kapcsolatokat megtestesítő – élek mindkét módszer esetén biztosították, hogy a hatékony működés szinonimájaként is felfogható hatványfüggvény-eloszlás érvényesüljön.

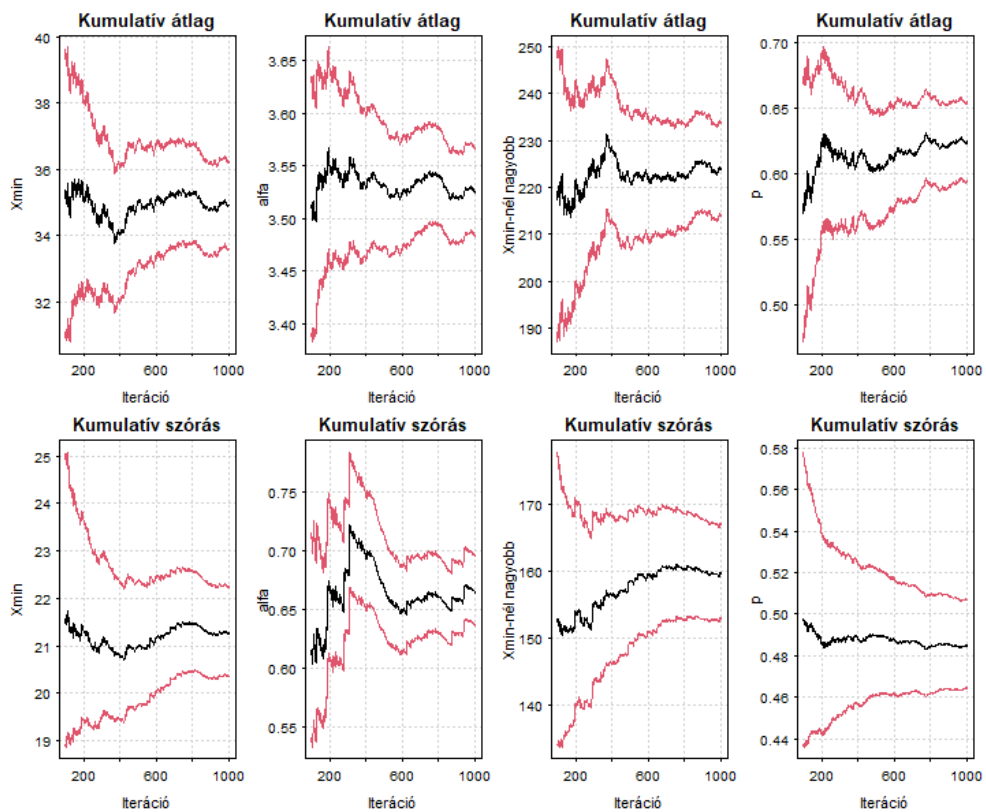
39. táblázat Fokszámok hatványfüggvényeloszlás-próbájának értékei bootstrapping(p)-módszerrel

Adatsor	$\alpha(\gamma)$	x_{min}	p
FP7(fa)	3,51	14,28	0
FP7(fa2)	3,52	34,90	0,62
H2020(fa)	6,24	17,95	0
H2020(fa2)	4,31	51,59	0,36

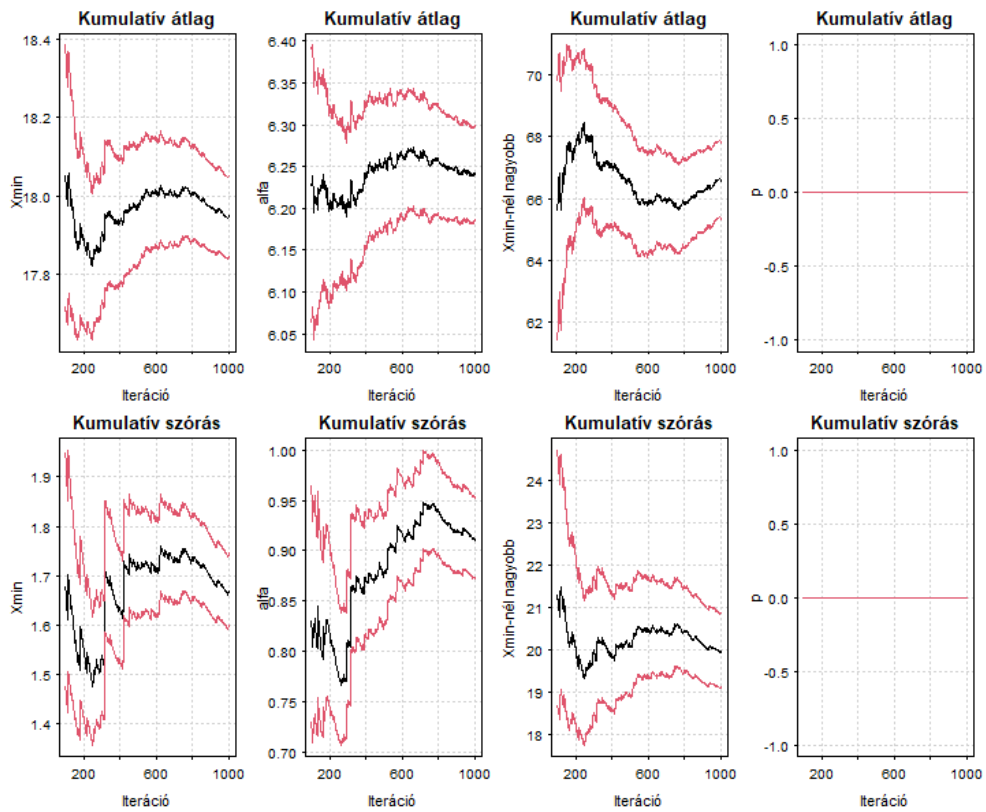
Az x_{min} alatti tartományra az összefüggés nem érvényesül. Kérdés azonban – s erre még vizsgatérünk –, hogy a kapott eredmény nem látszólagos-e, nincsenek-e olyan tényezők, amelyek ebben a tartományban is a hatványfüggvény-eloszlást valószínűsítik.



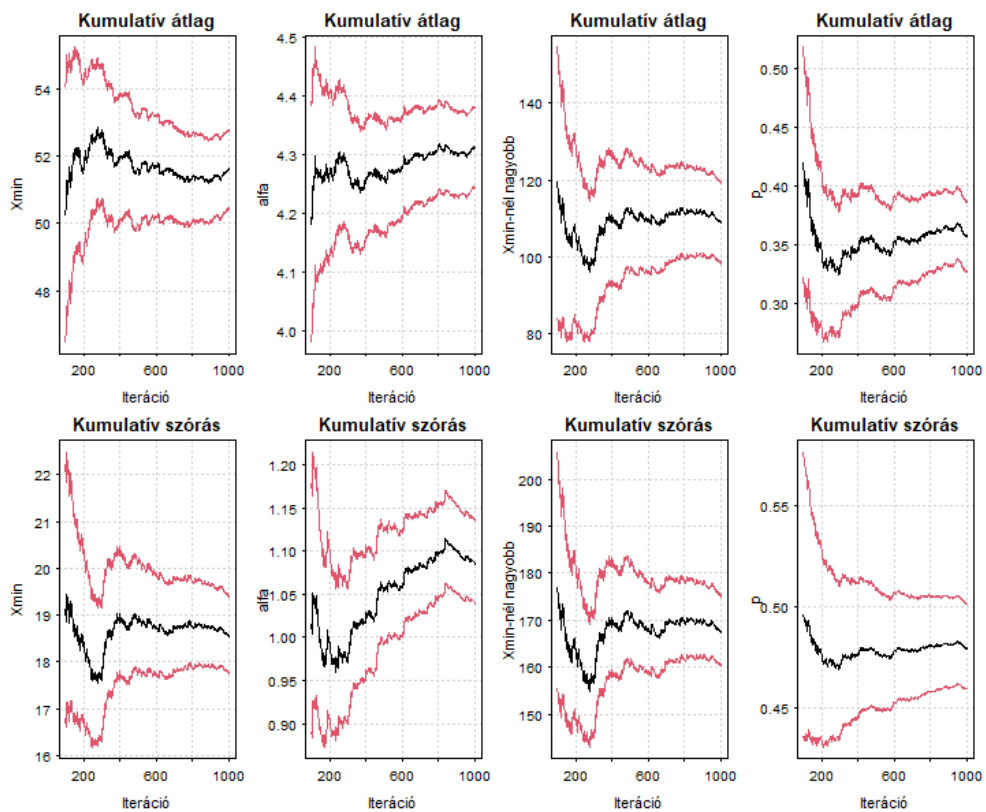
74. ábra FP7(fa) fokszámok hatványfüggvényeloszlás-próbájának folyamata bootstrapping(p) módszerrel



75. ábra FP7(fa2) fokszámok hatványfüggvényeloszlás-próbájának folyamata bootstrapping(p) módszerrel



76. ábra H2020(fa) foksámok hatványfüggvényeloszlás-próbájának folyamata bootstrapping(p) módszerrel

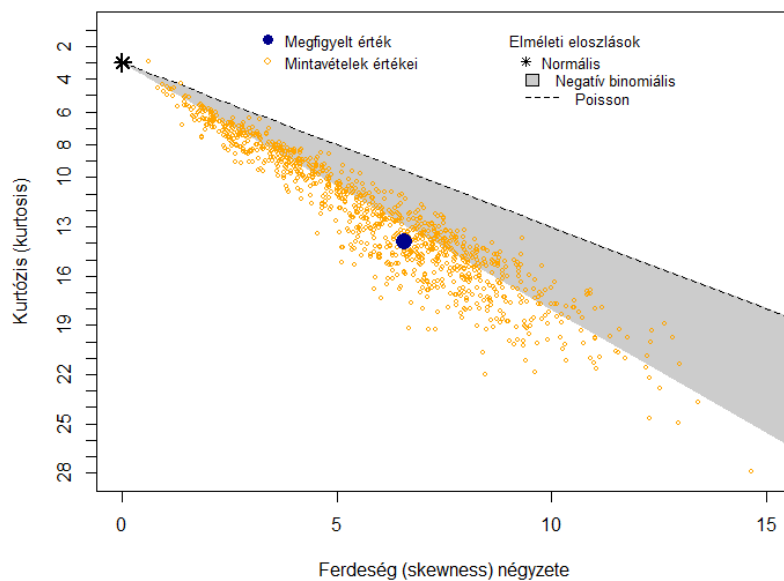


77. ábra H2020(fa2) foksámok hatványfüggvényeloszlás-próbájának folyamata bootstrapping(p) módszerrel

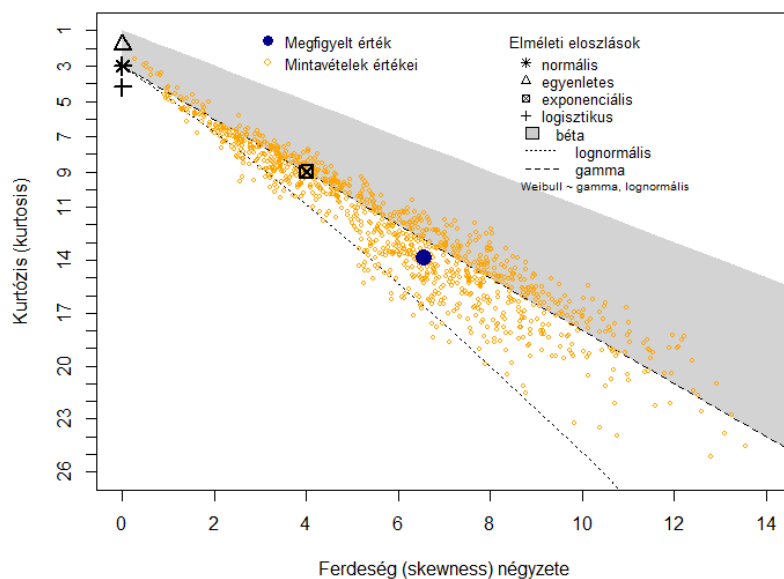
7.4.3.3. Egyéb eloszlások

A (40)-(44) egyenletek összefüggései alapján felrajzolt Cullen-Frey diagramok segítségével szélesebb körben is tesztelhetjük a fokszámok lehetséges eloszlását (78-85. ábra).

Az FP7(fa) diszkrét változóként kezelt fokszámaira felrajzolt Cullen-Frey diagram megerősíti a korábbi számításainkat, hogy a változó nem Poisson-eloszlású. Emellett a negatív binomiális eloszlás sávjába sem esik bele. (78. ábra) A fokszámok eloszlása a gamma- és a lognormális eloszlás közé tehető, s a béta-eloszlás tartományán is kívül esik. A kurtózis 3-nál sokkal nagyobb értéke magas csúcsosságra utal. Ilyen esetben a Weibull-eloszlás illeszkedése is korlátozott a másik három „ferde” eloszláshoz hasonlóan. (79. ábra)

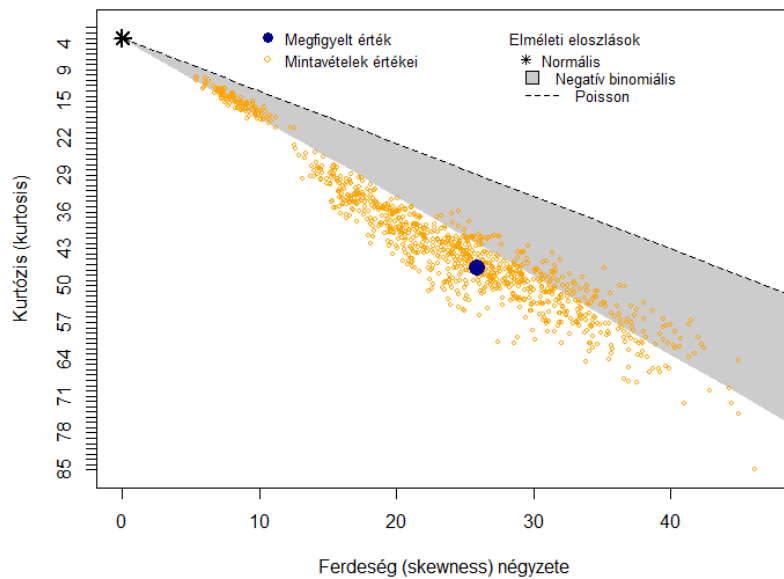


78. ábra FP7(fa) fokszámainak Cullen-Frey diagramja (diszkrétként kezelt változó)

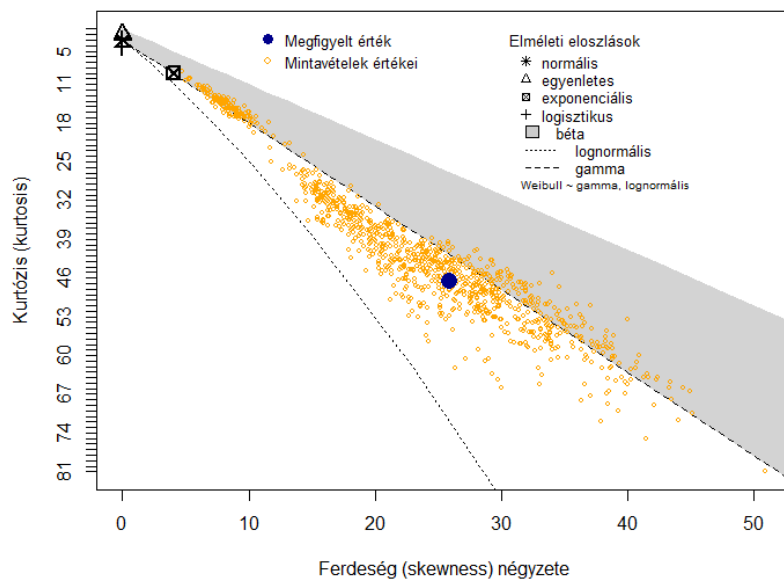


79. ábra FP7(fa) fokszámainak Cullen-Frey diagramja (folytonosként kezelt változó)

A Cullen-Frey diagram az FP7(fa2) fokszámaint diszkrét (80. ábra) és folytonos (81. ábra) változóként kezelve is nagyon hasonló mintázatot mutat, mint az előző esetben, tehát itt sem valószínűsíthető semmilyen, a korábbiakban nem érintett, „gyakori” eloszlás.

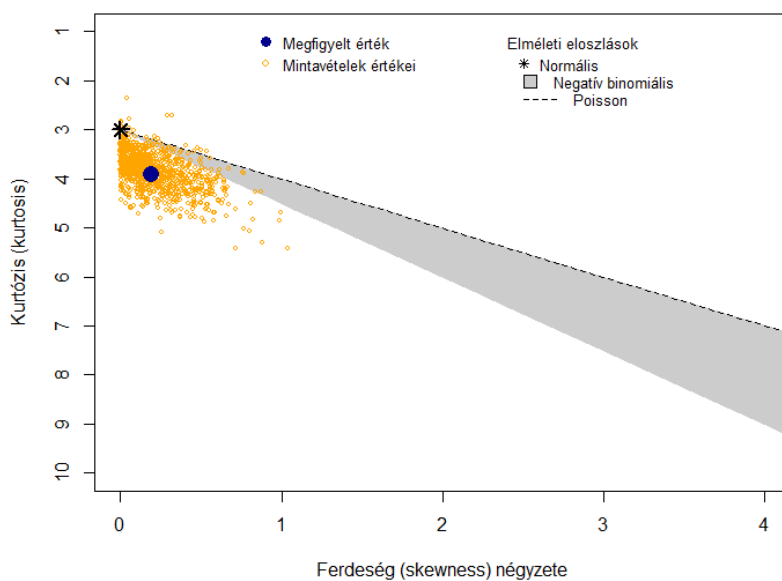


80. ábra FP7(fa2) fokszámainak Cullen-Frey diagramja (diszkrétként kezelt változó)

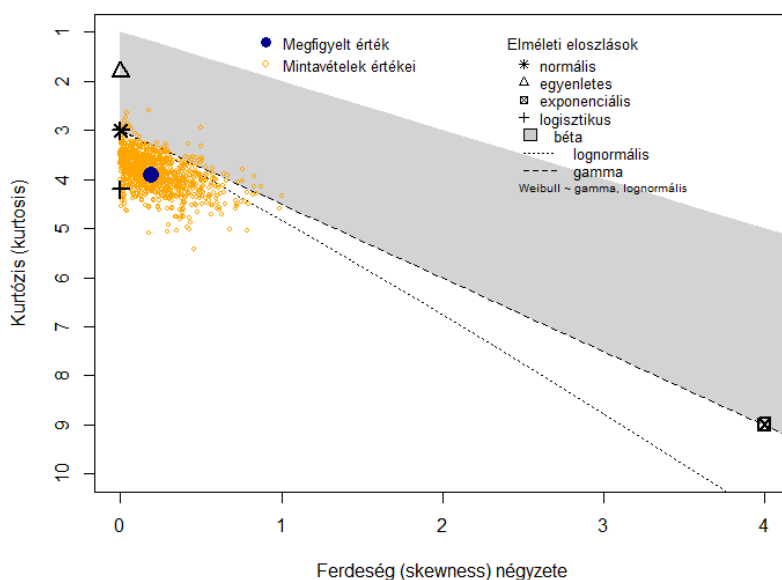


81. ábra FP7(fa2) fokszámainak Cullen-Frey diagramja (folytonosként kezelt változó)

Némileg más a helyzet a H2020(fa) fokszámainál. Az adatsort diszkrét változóként kezelve az némileg közelebb esik néhány gyakori diszkrét eloszláshoz (82. ábra), mint a korábbi adatrok, illetve a folytonos eloszlásként kezelve is ezt tapasztaljuk (83. ábra).



82. ábra H2020(fa) foksámainak Cullen-Frey diagramja (diszkrétként kezelt változó)

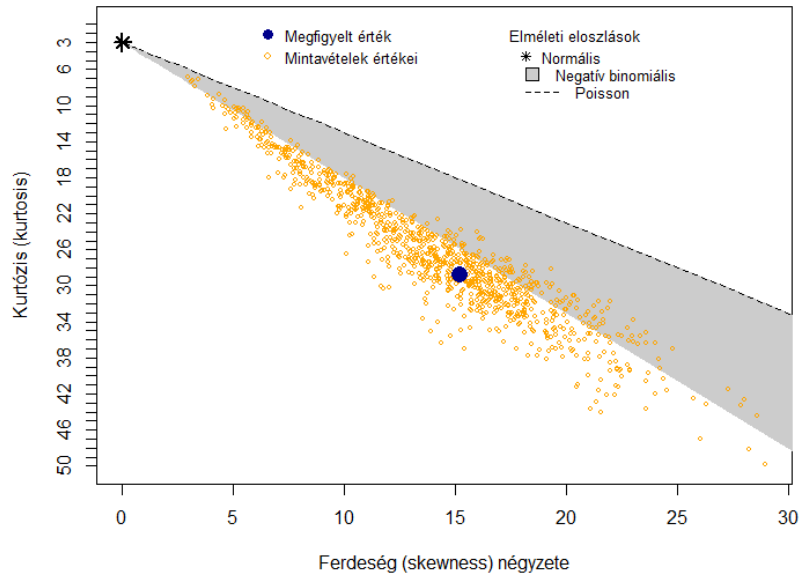


83. ábra H2020(fa) foksámainak Cullen-Frey diagramja (folytonosként kezelt változó)

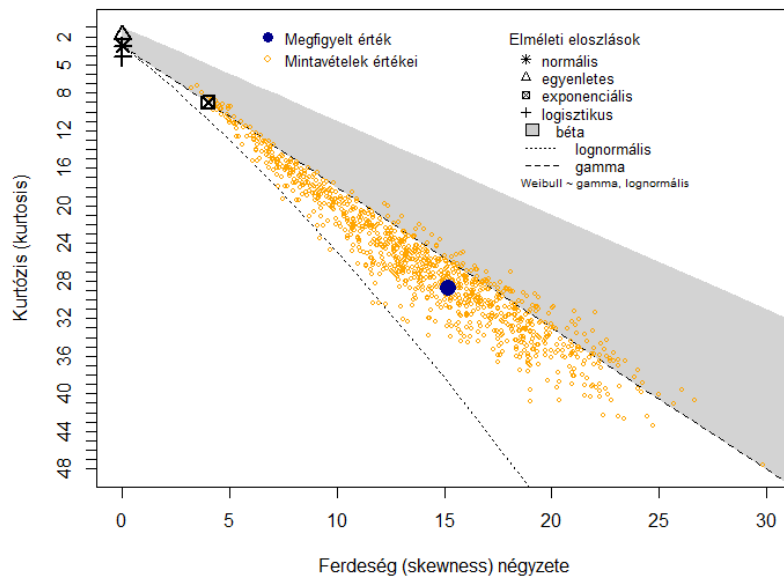
A H2020(fa2) foksámai viszont mind diszkrét változóként kezelve (84. ábra), mind folytonos változóként kezelve (85. ábra) inkább az első két adatsor foksámaira emlékeztetnek.

Matematikai módszerekkel természetesen kereshetünk foksámokat egészen pontosan leíró függvényt és (adatelemzésben ritkábban használt) eloszlást. De célszerűbb inkább a gazdasági-társadalmi jelenségeknél általános, s egyben óvatosabb összefüggések keresésére törekednünk. Ezért az eloszlások további keresése ebben az esetben nem indokolt.

A 40. táblázat a kapcsolatháló foksámainak leíró statisztikáját mutatja. Ha a tapasztalati foksámainkat és magát a hálózatot sokaságnak, s nem mintának tekintjük, akkor a „becsült” jelző feleslegesnek tűnik. Viszont a ferdeség és a kurtózis jellege, az alkalmazott bootstrapping módszer miatt indokolt a jelző használata.



84. ábra H2020(fa2) fokszámainak Cullen-Frey diagramja (diszkrétként kezelt változó)



85. ábra H2020(fa2) fokszámainak Cullen-Frey diagramja (folytonosként kezelt változó)

40. táblázat A kapcsolatháló fokszámainak deskriptív statisztikája

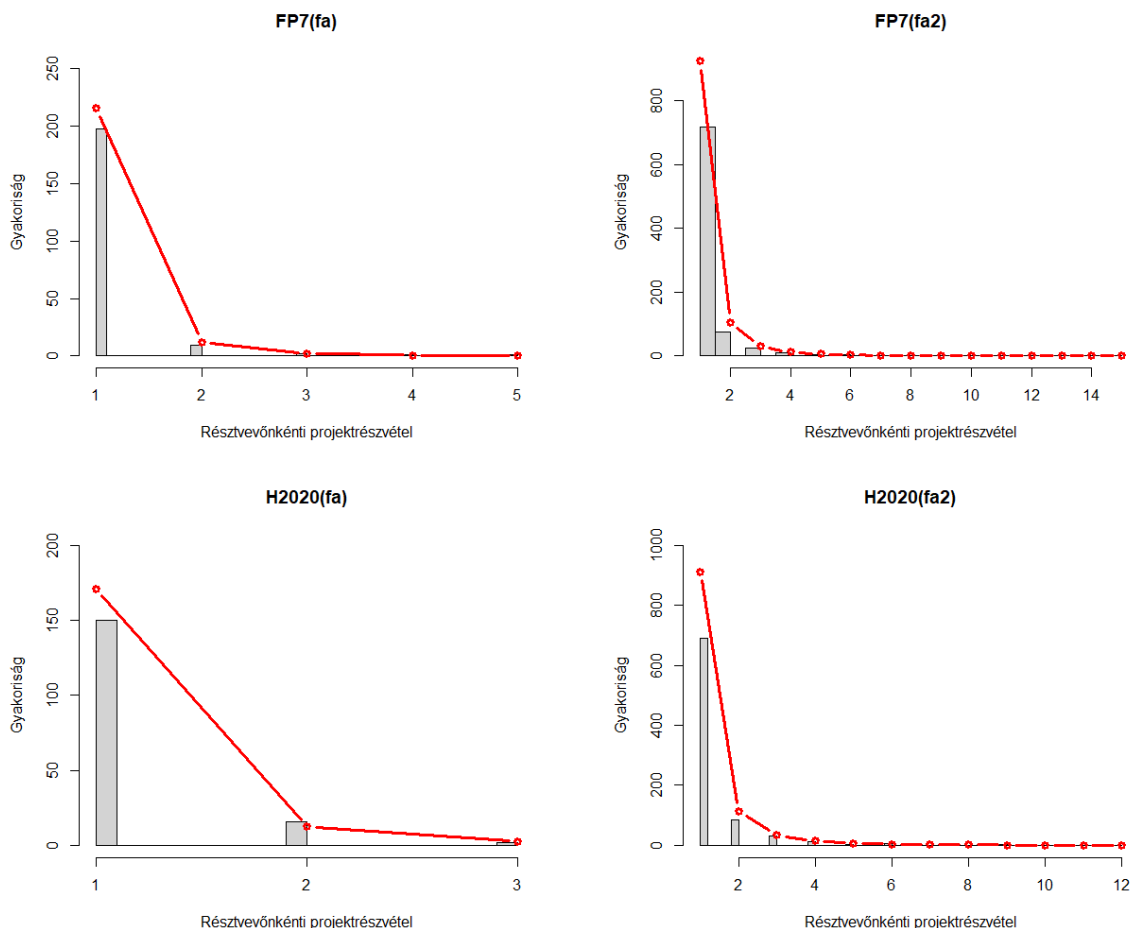
Mutató	FP7(fa)	FP7(fa2)	H2020(fa)	H2020(fa2)
Minimum	3	1	1	1
Maximum	76	242	37	202
Medián	14	14	17	17
Átlag	16,21	19,89	15,62	21,25
(Becsült) szórás	9,29	18,61	7	18,14
(Becsült) ferdeség	2,56	5,09	0,44	3,90
(Becsült) kurtózis	13,89	46,62	3,91	28,81

Talán érdekes, hogy FP7(fa) és FP7(fa2), illetve H2020(fa) és H2020(fa2) fokszámainak mediánja megegyezik, miközben a kibővített faipari kutatási hálóknak résztvevőinek fokszáma magasabb átlagot, szórást, ferdeséget, kurtózist mutat.

A kapott eredmények alapján azt feltételezhetjük (ez a legvalószínűbb), hogy a „balra ferde” eloszlást véletlen tényezők, illetve a skálafüggetlenség felé ható tényezők együtt befolyásolták, ill. alakították ki.

A kapott eredmény levezethető a projektekhez kapcsolódó partnerkeresés jellegéből is.

A projektek potenciális résztvevői a hálózat növekedése során legszívesebben a már elismert, vagy már a benyújtási szakaszban sok hálózati kapcsolatot felmutató csomópontokhoz kapcsolódtak (preferenciális kapcsolódás), de ezáltal kapcsolatba kerültek a többi projektrésztvevővel is, így a legalacsonyabb fokszámok gyakorisága törvényszerűen alacsonyabb, mint a skálafüggetlen hálózatokban. A támogatási rendszer ráadásul a többszereplős projekteket preferálja.



86. ábra A résztvevőkre jutó projektrészvételek gyakorisága

Az is elképzelhető, hogy a vizsgált adatok az együttműködéseknek csak egy kis – nem feltétlenül reprezentatív – mintáját jelentik, s a teljes együttműködő hálózat jóval több alacsony fokszámot tartalmazna.

Ezt az összefüggést látszik alátámasztani az 86. ábra, amely azt mutatja, hogy milyen gyakorisággal bír a projektekben való részvételek száma résztvevőnként.

A hatványfüggvény-eloszlásra végzett Kolmogorov-Szmirnov próba minden esetben maximális p értéket mutat.

41. táblázat A résztvevőkre jutó projektrészvételek számának Kolmogorov-Szmirnov próbája hatványfüggvény-eloszlásra

Hálózat	$\alpha(\gamma)$	x_{min}	p
FP7(fa)	4,19	1	1
FP7(fa2)	3,14	1	1
H2020(fa)	3,76	1	1
H2020(fa2)	3,02	1	1

A hatványfüggvény-eloszlásra vonatkozó $H_0(38)$ -t tehát egyik esetben sem vetjük el. Mivel γ értéke mind a négy esetben csak kis mértékben haladja meg a 3-at, joggal feltételezhetjük, hogy a skálafüggetlenség valamilyen szinten érvényesül.

Tehát míg a foksámok eloszlása nem követett hatványfüggvény-eloszlást, a projektrészvételek gyakorisága szinte tökéletes illeszkedést mutatott már a kezdőértéktől. Azaz a projektrészvételek gyakoriságára valószínűleg érvényes a gazdasági-társadalmi folyamatokban fontos jelenségnek tartott skálafüggetlenség.

Ezek alapján a 15. szakmai hipotézist alátámasztó érvek, miszerint a projektrészvevők kapcsolathálóiiban mért foksámok eloszlásai skálafüggetlenek, azaz kevés résztvevőnek van sok és sok résztvevőnek kevés kapcsolata, tovább erősödnek. Az eredmények azonban túl sok bizonytalanságot hordoznak, ezért ezzel a hipotézissel kapcsolatban tézist nem fogalmazunk meg.

A kisvilágjelleg a hálózati mutatók korábban vázolt megbízhatatlansága ellenére megvizsgáltuk, ez eredményeket azonban a hálók töredezettsége miatt óvatosan kell kezelnünk (42. táblázat).

42. táblázat Kisvilágjelleg adatai

Hálózat	$\ln N / \ln \langle k \rangle$	d_{max}	$\langle d \rangle$
FP7(fa)	1,92	4	2,53
FP7(fa2)	2,25	6	3,00
H2020(fa)	1,82	3	1,71
H2020(fa2)	2,16	6	3,04

A (48) egyenletben leírt összefüggés láthatóan egyik esetben sem teljesül. Látszólag H2020(fa) átmérője (maximális átmérője) esik legközelebb az optimális kisvilág átmérőjének értékéhez, de ha az elkülönülő részhálókat összekapcsolnánk, d (d_{max}) értéke minden bizonnyal – amennyiben néhány új kapcsolatról van szó – nőne. Az átlagos átmérő értéke a kisvilág átmérőjének értékéhez több esetben közel esik (49), ezért óvatosan feltételezhetjük a kisvilágjellegget, de a hálózatok struktúrája miatt egyértelműen nem jelenthetjük ki a kisvilágjelleg érvényesülését. Inkább több lazán vagy egyáltalán nem kapcsolódó kisvilágot feltételezhetünk.

A fenti értékek alapján a faipari projektek hálózata nem vagy nagyon korlátozottan tekinthető kisvilágnak. Emiatt a 17. szakmai hipotézist, miszerint a faipari projekthálókból nem érvényesül a kisvilág jelensége, nem utasítjuk el egyértelműen. Azonban ez a megállapítás nem elegendő egy tézis megfogalmazásához.

7.4.4. Központi résztvevők

A foksám és a közöttiség (50) alapján a legfontosabb tíz szereplő az egyes csoportokon – FP7(fa), FP7(fa2), H2020(fa) és H2020(fa2) – belül a 43. táblázattól a 50. táblázatig láthatóak.

Bár az egyes táblák csak többé-kevésbé hasonlítanak egymásra, már az első, 43. táblázat alapján kibontakozik néhány közös tendencia.

Egyetlen volt szocialista ország sem játszik központi szerepet a faipari projektekben. A természettudományi-műszaki, vagy kimondottan faipari, nemzeti vagy nemzetközi szinten kialakított K+F klaszterek vagy ernyőszervezetek központi szerepre tudtak szert tenni a faipari projektekben.

43. táblázat Az FP7(fa) projektek központi résztvevői a foksám alapján

Név	Ország	Foksám
Teknologian tutkimuskeskus VTT	Finnország	76
Fraunhofer-Gesellschaft	Németország	59
JRC – European Commission	Belgium	52
Institut Technologique FCBA	Franciaország	48
Instituto Superior de Agronomia	Portugália	48
Internationale Institut für Wald und Holz NRW	Németország	38
Stichting Wageningen Research	Hollandia	37
Danmarks Tekniske Universitet	Dánia	33
Forest Enterprises Ltd	Írország	27
Wexford Wood Producers Ltd	Írország	27

Iparági súlyuk miatt a finn, német, francia szervezetek központi szerepe teljesen természetes, de szintén minden listában feltűnnek erdőikkel alig, faipari kibocsátással korlátozottan rendelkező országok. Úgy tűnik, hogy a központi szerepnek vannak olyan tényezői, amelyek statisztikai eszközökkel nehezen feltárhatóak.

44. táblázat Az FP7(fa) projektek központi résztvevői a közöttiség alapján

Név	Ország	Közöttiség
Teknologian tutkimuskeskus VTT	Finnország	6.285,50
Fraunhofer-Gesellschaft	Németország	4.728,65
Instituto Superior de Agronomia	Portugália	3.738,53
Institut Technologique FCBA	Franciaország	2.546,73
JRC – European Commission	Belgium	1.965,83
Internationale Institut für Wald und Holz NRW	Németország	1.815,00
Danmarks Tekniske Universitet	Dánia	1.478,90
TNO	Hollandia	669,77
Stichting Wageningen Research	Hollandia	540,68
IIASA	Ausztria	162,40

45. táblázat Az FP7(fa2) projektek központi résztvevői a foksám alapján

Név	Ország	Foksám
Teknologian tutkimuskeskus VTT	Finnország	242
Fraunhofer-Gesellschaft	Németország	215
Fundación Tecnalia R&I	Spanyolország	133
Institut Technologique FCBA	Franciaország	125
Teknologian tutkimuskeskus VTT Oy*	Finnország	123
Universität für Bodenkultur Wien	Ausztria	102
RISE Research Institutes of Sweden AB	Svédország	98
INRAE	Franciaország	93
European Forest Institute	Finnország	88
Teknologisk Institut	Dánia	85

*A cég 2015-ben jött létre a Teknologian tutkimuskeskus VTT egy másik szervezettel való egyesülés után. A Teknologian tutkimuskeskus VTT és a Teknologian tutkimuskeskus VTT Oy külön szerepeltetése eltérő rcn-kódjuk miatt indokolt.

46. táblázat Az FP7(fa2) projektek központi résztvevői a közöttiség alapján

Név	Ország	Közöttiség
Teknologian tutkimuskeskus VTT	Finnország	78.304,64
Fraunhofer-Gesellschaft	Németország	57.465,94
RISE Research Institutes of Sweden AB	Svédország	49.417,75
Fundación Tecnalia R&I	Spanyolország	29.618,45
Institut Technologique FCBA	Franciaország	28.721,23
Tecnologias Avanzadas Inspiralía SI	Spanyolország	13.583,28
Teknologisk Institut	Dánia	13.444,34
Teknologian tutkimuskeskus VTT Oy	Finnország	13.413,96
BIOS Bioenergiesysteme GmbH	Ausztria	13.170,53
Universiteit Gent	Belgium	11.800,11

Egészen meglepő, hogy a foksám alapján egy ész szervezet lehetett a H2020(fa) projektek első számú aktora (47. táblázat), s talán az is meglepő, hogy az FP7(fa) projektek központi szereplői alig jelennek meg a H2020(fa) központi szereplői között.

47. táblázat A H2020(fa) projektek központi résztvevői a foksám alapján

Név	Ország	Foksám
FIBENOL OÜ	Észtország	37
CSR Consorzio Studi Ricerche S.r.l.	Olaszország	35
SCM Group S.p.A	Olaszország	35
AIMEN	Spanyolország	35
Fundación Tecnalia R&I	Spanyolország	34
Consiglio Nazionale delle Ricerche	Olaszország	33
Metgen Oy	Finnország	28
Global Bioenergies GmbH	Németország	26
Global Bioenergies	Franciaország	26
AS Graanul Invest	Észtország	26

48. táblázat A H2020(fa) projektek központi résztvevői a közöttiség alapján

Név	Ország	Közöttiség
Fundación Tecnalia R&I	Spanyolország	646,00
Consiglio Nazionale delle Ricerche	Olaszország	242,00
TNO	Hollandia	170,25
CSR Consorzio Studi Ricerche S.r.l.	Olaszország	166,50
SCM Group S.p.A	Olaszország	166,50
AIMEN	Spanyolország	166,50
FIBENOL OÜ	Észtország	129,64
IMEC	Belgium	86,25
Metgen Oy	Finnország	41,02
Spinverse Oy	Finnország	31,75

Talán még meglepőbb, hogy a H2020(fa2) adatok alapján összeállított listák több átfedést mutatnak az FP7(fa) listáival, mint a H2020(fa) listái.

49. táblázat A H2020(fa2) projektek központi résztvevői a foksám alapján

Név	Ország	Foksám
Teknologian tutkimuskeskus VTT Oy	Finnország	202
Luonnonvarakeskus	Finnország	172
Fundación Tecnalia R&I	Spanyolország	152
Fraunhofer-Gesellschaft	Németország	124
RINA Consulting S.p.A	Olaszország	119
Consejo Superior de Investigaciones Científicas	Spanyolország	89
Consiglio Nazionale delle Ricerche	Olaszország	87
Instituto Superior de Agronomia	Portugália	87
European Forest Institute	Finnország	86
Soprema	Franciaország	82

50. táblázat A H2020(fa2) projektek központi résztvevői a közöttiség alapján

Név	Ország	Közöttiség
Teknologian tutkimuskeskus VTT Oy	Finnország	51.635,02
Luonnonvarakeskus	Finnország	34.868,73
Fraunhofer-Gesellschaft	Németország	28.950,85
Fundación Tecnalia R&I	Spanyolország	15.103,15
University of Aveiro	Portugália	13.132,68
IRIS Technology Solutions S.L.	Spanyolország	11.361,76
L'ACR+	Belgium	11.037,01
Metgen Oy	Franciaország	10.917,89
Consejo Superior de Investigaciones Científicas	Spanyolország	10.680,90
Luleå tekniska universitet	Svédország	10.576,00

Összesen három olyan projektrésztvevő volt, amely a 8 listából 4-nél többre felkerült, egyenként 6-6 részvétellel: Teknologian tutkimuskeskus VTT (vagy Teknologian tutkimuskeskus VTT Oy), Fundación Tecnalia R&I és a Fraunhofer-Gesellschaft.

Ők tekinthetők az utolsó két európai keretprogram legfontosabb faipari résztvevőinek. Talán nem elhamarkodott a feltételezés, hogy a velük kialakított K+F együttműködés kulcsfontosságú szerepet játszhat az európai források bevonása során.

A 51. táblázat az egyes központszereplő-listák közös elemeinek számát mutatja. A Technologicalian kutimuskus VTT-t és a Technologicalian kutimuskus VTT Oy-t a számítás előtt összevontuk Technologicalian kutimuskus VTT-vé. Mivel az FP7(fa2) első tíz fokszámát és a közöttségét jellemző adatsorok elemszáma ezzel kilencre csökkent, így az eredetileg 11. helyezetteket bevontuk az adatok közé. Ezután előállítottuk a különböző mutatók és adatsorokon belüli első tíz helyezettek azonos elemeinek számát mutató mátrixot.

51. táblázat A különböző szempontok szerinti tízes listák közös elemeinek száma

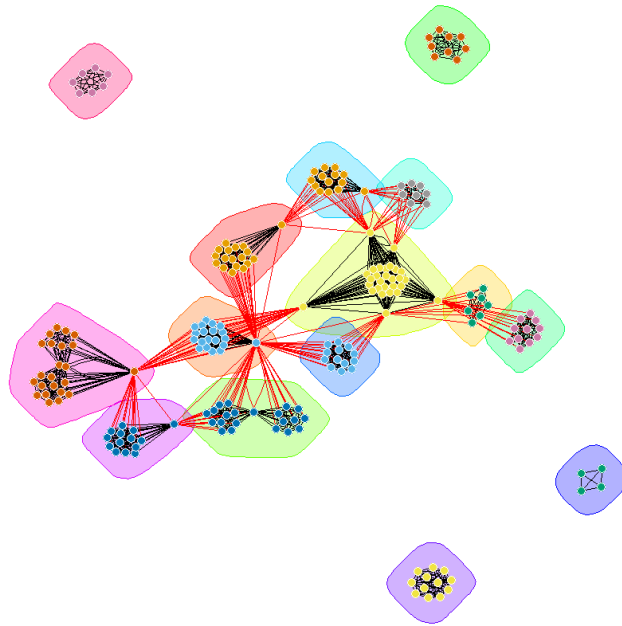
	FP7(fa) fokszám	FP7(fa) közöttség	FP7(fa2) fokszám	FP7(fa2) közöttség	H2020(fa) fokszám	H2020(fa) közöttség	H2020(fa2) fokszám	H2020(fa2) közöttség
FP7(fa) fokszám	10							
FP7(fa) közöttség	8	10						
FP7(fa2) fokszám	3	3	10					
FP7(fa2) közöttség	3	3	6	10				
H2020(fa) fokszám	0	0	1	1	10			
H2020(fa) közöttség	0	1	1	1	7	10		
H2020(fa2) fokszám	3	3	4	3	2	2	10	
H2020(fa2) közöttség	2	2	3	3	2	2	5	10

A közös elemek számát mutató mátrix megerősíti a korábban tapasztaltakat. A szűken vett faipari projektek tartalmi alapú kiegészítése a részben faipari projektekkel és résztvevőkkel a H2020 fő központi aktorainak listáját az FP7(fa) projektjeivel több közös elemet tartalmazó listákká tette a szűkebben értelmezett faipari projekteket tartalmazó H2020(fa) listákhoz képest.

A fokszámok közötti apró különbségek és a számos azonosság miatt a közöttség jobb mutatónak tűnik a központi aktorok kiválasztására.

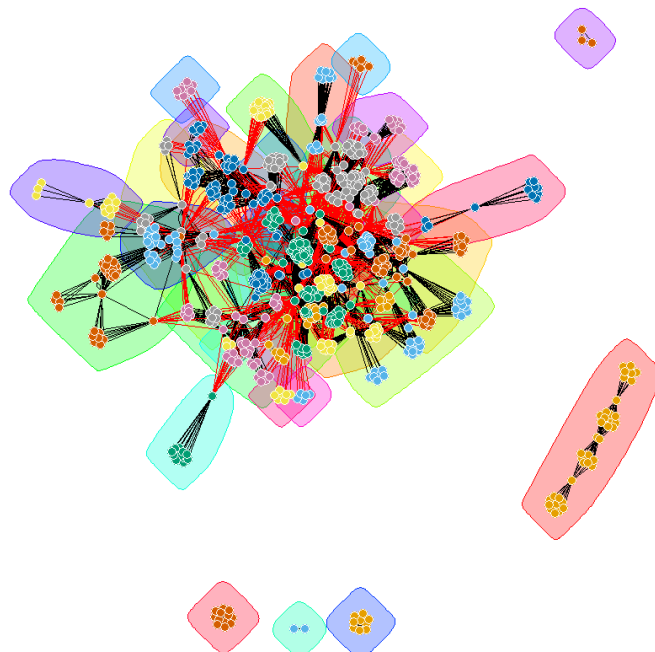
7.4.5. Hálózati klaszterek

Az FP7(fa) projekthálózában A 6.4.1.3. fejezetben vázoltak alapján a 17 projektből 15 klaszter rajzolható fel, tehát az algoritmus alig vont össze projekteket, illetve csak kis mértékben, egy-egy elem erejéig rajzolta át a projekthatárokat. (87. ábra) A legnagyobb klaszternek 28 résztvevője van, amelyben ötnek van olyan magas közöttség értéke, amellyel a teljes lista első tíz helyére is felkerült. A többi résztvevő közöttségértéke viszont 0.



87. ábra Az FP7(fa) hálózatának klaszterei

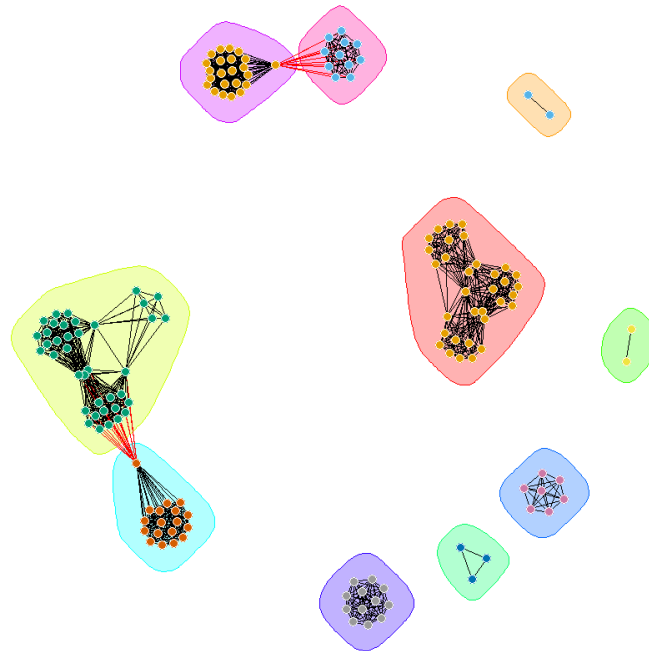
Az FP7(fa2) projekthálózában az eredetileg 84 projektet számoló hálózat méretét az algoritmus az előzőnél jobban, 38 klaszterre csökkentette le. (88. ábra) A legnagyobb klaszterben 80 résztvevő található, a közöttiség értéke három résztvevőnél magas, tíznél mérhető, 67-nél pedig 0.



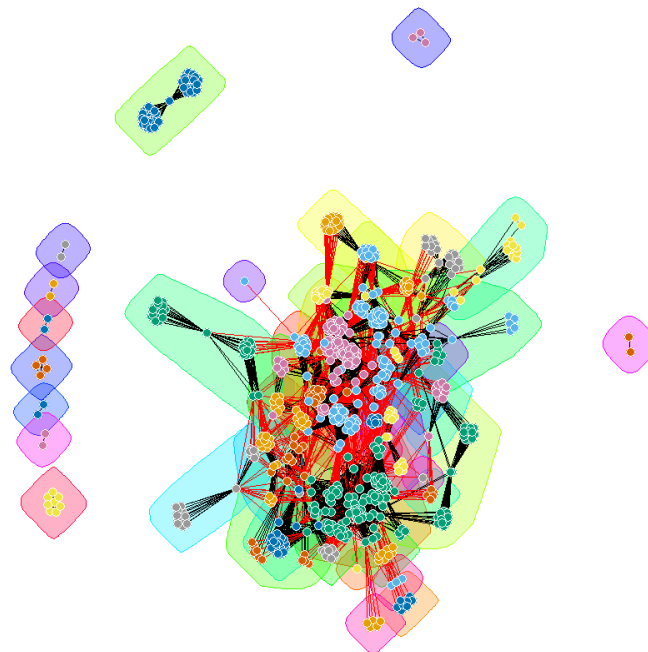
88. ábra Az FP7(fa2) hálózatának klaszterei

A H2020(fa) hálózatban a 15 projektből 10 klaszter jött létre. (89. ábra) A legnagyobb klaszter 38 résztvevőt számlál, közülük öt résztvevő közöttisége magas, 33-é pedig 0.

A H2020(fa2) projekthálóban a 80 projektre 45 klaszter épül. A legnagyobb klaszter elemszáma 101. (90. ábra) Közülük egy magas, 37 mérhető, 63 pedig 0 közöttiségértékkel rendelkezik.



89. ábra A H2020(fa) hálózatának klaszterei



90. ábra A H2020(fa2) hálózatának klaszterei

A legnagyobb klaszterekben tehát a hálózat szempontjából is központinak számító vagy legalább jelentős résztvevő mellett szinte minden esetben találunk a hálózat szempontjából jelentéktelen, periférikus résztvevőket is. Ez felvillantja a hatékony partnerkeresés stratégiáját. Mivel a jelentős faipari projektekben megfelelő szakmai kvalitások mellett periférikus – vagy addig periférikus – résztvevőknek is lehet esélye a becsatlakozásra, érdemes lehet közvetlenül

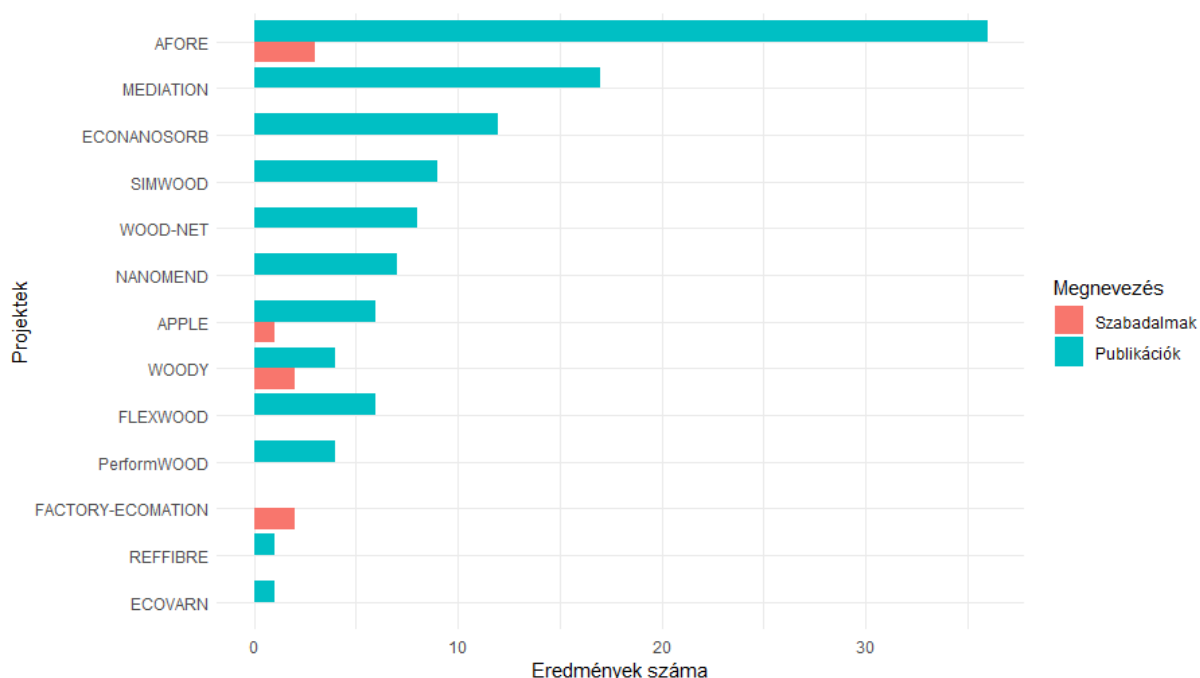
a hálózatok központi résztvevőivel szakmai kapcsolatokat kialakítani, hiszen a projektegyüttműködések belső hierarchiájában a kevésbé „fontos” résztvevőkre is szükség van.

Az eddigiek alapján nem utasítjuk el a 16. szakmai hipotézist, miszerint a faipari projekthálóok jelentős hálózati klaszterei hierarchikus felépítésűek.

7.5. Projekteredmények hatékonyság- és tartalomelemzése

7.5.1. Az európai uniós támogatási programok faipari projektjeinek projekteredményei

Az elérhető adatok alapján az FP7(fa) projektekben 8 szabadalom és 111 publikáció született. A legtöbb projekteredményt az AFORE (39), a MEDIATION (17), ECONANOSORB (12) mutatta fel. (91. ábra)

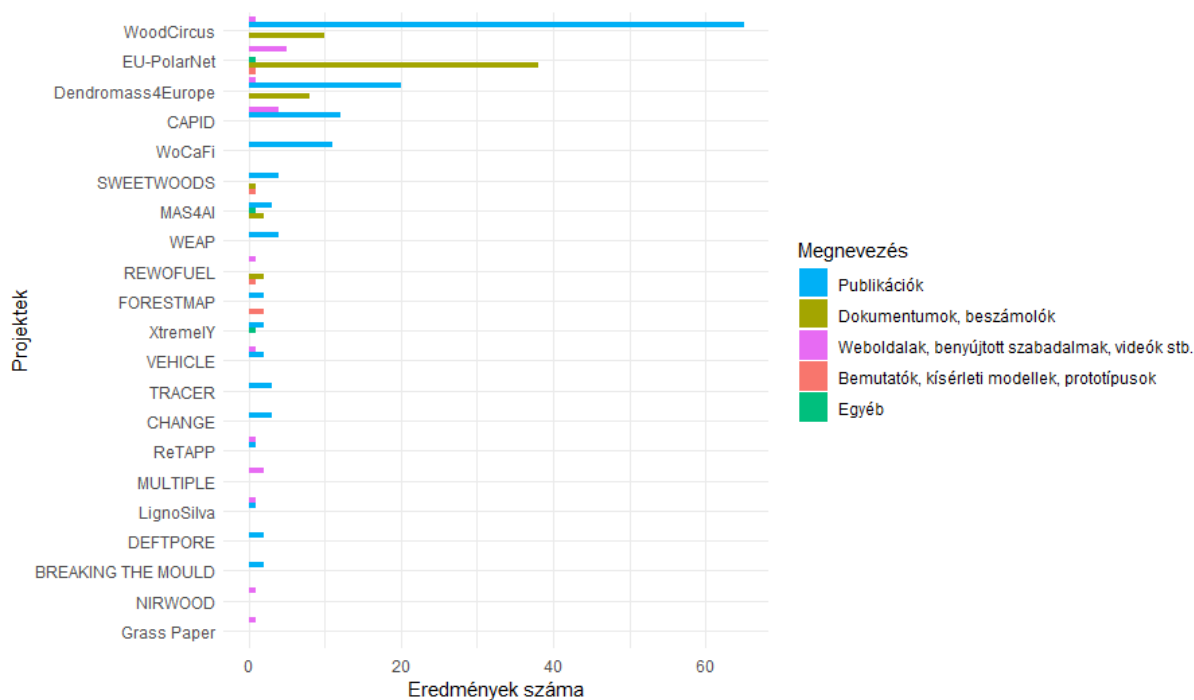


91. ábra FP7(fa) projektek eredményei

A H2020(fa) projekteknel – a H2020 kibővített adatközlése miatt – lényegesen több és részletesebb adatot találunk: 137 publikáció, 61 dokumentáció és beszámoló, 19 weboldal, benyújtott szabadalom és videó (ez utóbbi három egy kategória), öt bemutató, kísérleti modell és prototípus (ez a három is egy kategória) és három egyéb output jött létre. A legtöbb eredményt a WoodCircus (76), az EU-PolarNet (45) és a Dendromass4Europe (29) mutatták fel. (92. ábra)

Az összesen 225 projekteredmény szakmai megítélése nyilván nem egyszerű, s nem is teszünk erre kísérletet pl. az eredmények meglehetősen heterogenitása miatt. Természetesen azt sem könnyű megítélni, hogy az egy projekteredményre jutó közösségi támogatás (371.416,5 euró) sok vagy kevés, de az eredmények összetétele és az általános magángazdasági/kutatóintézeti/felsőoktatási (pl. publikációs) elvárások alapján az outputokhoz képest a közösségi támogatás meglehetősen bőségesnek tűnik.

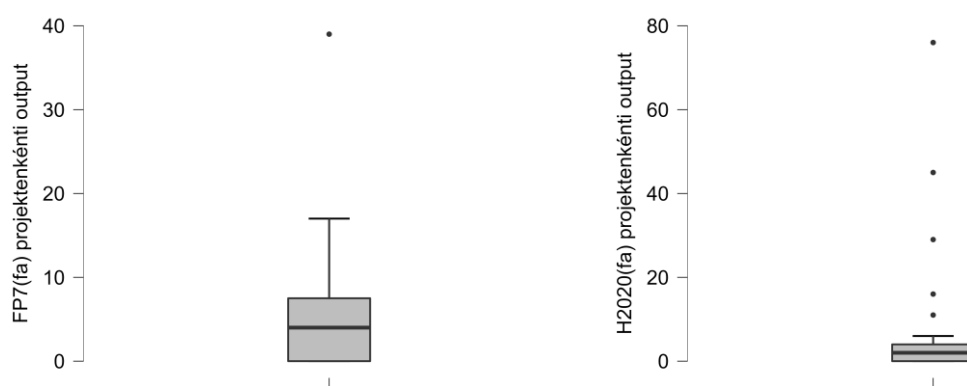
Az egyes projektek sikerességének megítélése a projekteredmények alapján az eredmények szakmai megítélése és a projektekkel szemben támasztott elvárások ismerete nélkül szintén lehetetlen.



92. ábra H2020(fa) projektek eredményei

Korábban megállapítottuk, hogy az FP7(fa) projektek támogatása nem normális eloszlású, illetve nincsenek benne outlierok. A H2020(fa) projektek támogatásáról szintén megállapítottuk, hogy nem normális eloszlású, viszont az itt talált nagyszámú outliert végül nem minősítettük outliernek, így nem zártuk ki őket az elemzésből.

A 93. ábra alapján mindkét adatsorban outlierok valószínűsíthetők.



93. ábra A projektenkénti outputok boxdiagramjai

Az FP7(fa) projektenkénti outputok esetén $I(8)(9)(10)$ rendre $[-11,25; 18,75]$, $[-13,79; 21,79]$, $[-16,5; 24]$ értéket vesz fel. Minden esetben egy érték tekinthető outliernek. Megvizsgáltuk, hogy semmilyen projekteredményt fel nem mutató projektek figyelmen kívül hagyása milyen változásokkal jár, de az outlier ugyanúgy megmaradt. Ezt az értéket ezért a hozzá tartozó sorral (AFORE projekt) töröljük az adatok közül.

A H2020(fa) projektenkénti outputok esetén $I(8)(9)(10)$ rendre $[-6; 10]$, $[-6,9; 10,9]$, $[-8,8; 12,8]$ értéket vesz fel. Az első két esetben öt, az utolsóban négy érték tekinthető outliernek. Itt is megvizsgáltuk, hogy a semmilyen projekteredményt fel nem mutató projektek figyelmen kívül hagyása milyen változásokkal jár, s így négy darab outlier maradt. Úgy döntöttünk ezért, hogy a négy legmagasabb értéket – WoodCircus (76), EU-PolarNet (45), Dendromass4Europe (29), CAPID (16) – töröljük az adatok közül.

A 52. táblázat alapján egyik adatsor sem tekinthető normális eloszlásúnak.

52. táblázat A projektenkénti outputok normalitásvizsgálata (outlierek nélkül)

Mutató	Output (projektenként)	
	FP7(fa)	H2020(fa)
Shapiro-Wilk (W)	0,85	0,75
P (Shapiro-Wilk)	0,01	0,00
Kolmogorov-Szmirnov (D)	0,53	0,50
P (Kolmogorov-Szmirnov)	0,00	0,00

A nagyszámú (hat, ill. 15) nullérték (közzétett eredmény nélküli projektek) miatt, ezek elhagyása után is megvizsgáltuk az eloszlásokat (53. táblázat). A Shapiro-Wilk próba alapján az FP7(fa) projektek outputjainak eloszlása ebben az esetben követhet normális eloszlást, de a kis mintáknál megbízhatóbb Komogorov-Szmirnov próba itt is kizárja a normalitást feltételező nullhipotézist.

53. táblázat A projektenkénti outputok normalitásvizsgálata (outlierek és nullértékek nélkül)

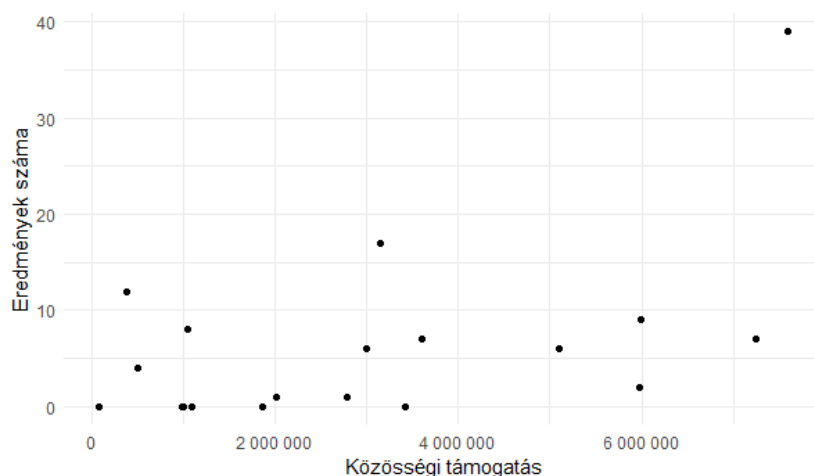
Mutató	Output (projektenként)	
	FP7(fa)	H2020(fa)
Shapiro-Wilk (W)	0,93	0,79
P (Shapiro-Wilk)	0,35	0,00
Kolmogorov-Szmirnov (D)	0,84	0,85
P (Kolmogorov-Szmirnov)	0,00	0,00

Pearson-féle korrelációs együttható ezért nem számítható egyik esetben sem, Spearman-féle rangkorrelációs együttható viszont igen.

A Mahalanobis-távolság az FP7(fa), outlierek nélküli projektek esetében kisebb, mint 13,82. (A maximális érték 6,67.) Tehát a Mahalanobis-távolság alapján nincs többdimenziós outlier. A Cook-távolság egyetlen egy esetben sem nagyobb, mint egy. (A legnagyobb érték 0,27.) Többdimenziós outlier tehát nem akadályozza a korrelációs számítás elvégzését. A H2020(fa) projektek esetén viszont a Mahalanobis távolság két projekt esetén, nagyobb, mint 13,82 (18,25 és 18,04 értékkel), viszont a Cook-távolság maximuma csak 0,3, tehát az amúgy is elég rövid adatsort felesleges tovább csökkenteni a kétdimenziósnak tűnő outlierekkel, mivel igen kis hatást gyakorolnak a korrelációs együtthatóra.

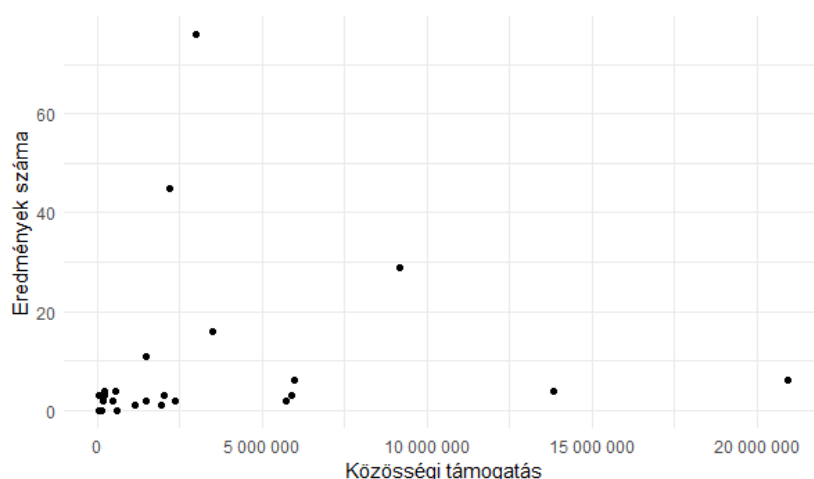
A Spearman-féle rangkorrelációs számításal közepesenél gyengébb, pozitív korrelációt mértünk az FP7(fa) projektenkénti output és a projektenkénti közösségi támogatás között ($\rho = 0,35$). Mivel $p = 0,15 > \alpha$, a függetlenségre vonatkozó $H_0(5)$ -t nem utasítjuk el. Megállapításunkat erősíti, hogy ha nemcsak az outliert, hanem a nullértékeket és a hozzájuk kapcsolódó támogatásokat is elhagyjuk, akkor $\rho = 0,06$, $p = 0,85$. A 18. szakmai hipotézis második fele tehát teljesülni látszik, azaz nincs összefüggés a támogatás és az eredmények száma között, ez a 94. ábra alapján is leolvasható.

Ettől természetesen még az eredmények minősége és tudományos hatása összefügghet a támogatás mértékével, de ezt nehezen tudjuk feltárni.



94. ábra Közösségi támogatás (euróban) és az eredmények száma az egyes FP7(fa) projektekben

A H2020(fa) projektek viszont látszólag rácafolnak az előző megállapításunkra. A projektenkénti output és a projektenkénti közösségi támogatás között meglehetősen erős kapcsolat bontakozik ki ($\rho = 0,81$, $p = 0,00$), amelyet egyébként vizuális benyomásainak nem nagyon erősítenek meg. (95. ábra)



95. ábra Közösségi támogatás (euróban) és az eredmények száma az egyes H2020(fa) projektekben

Viszont, ha az outlierok mellett a nullértékeket is elhagyjuk, akkor a függetlenségre vonatkozó $H_0(5)$ elutasíthatósága kérdésessé válik ($\rho = 0,24$, $p = 0,35$).

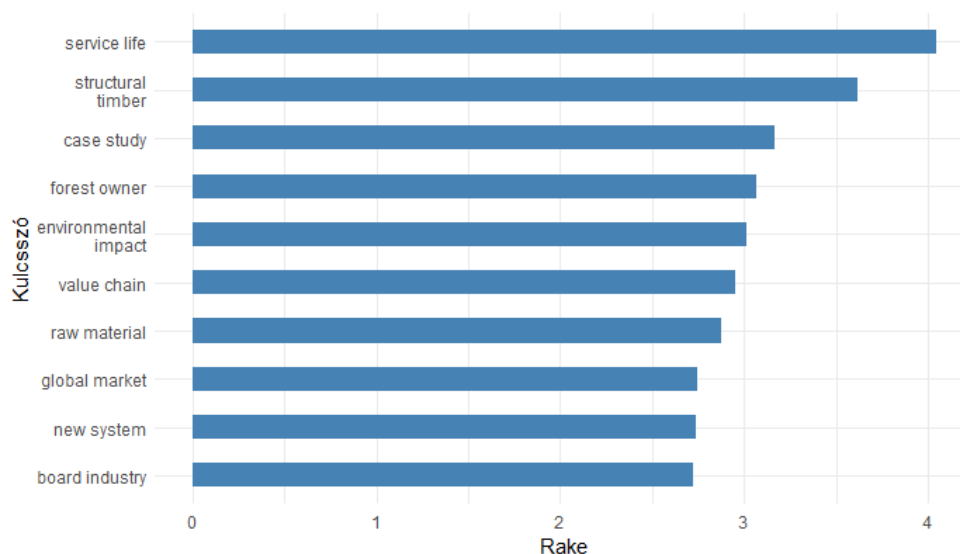
A kis minták és az ellentmondásos eredmények miatt nem utasítjuk el a 18. szakmai hipotézis második felét.

Természetesen a p értéket (illetve a próbát) itt is csak akkor kell figyelembe vennünk, ha mögöttes sokaságot (vagy szupersokaságot) feltételezünk, s a megállapításaink arra vonatkoznak.

Összességében a 18. szakmai hipotézist, miszerint az FP7 és a H2020 támogatási program faipari projektjeinek eredményei a felhasznált közösségi erőforrásokhoz képest korlátozott nyitott és zárt innovációs eredményeket hoztak, és a támogatás és az eredmények között szám szerint korlátozott az összefüggés, nem utasítjuk el.

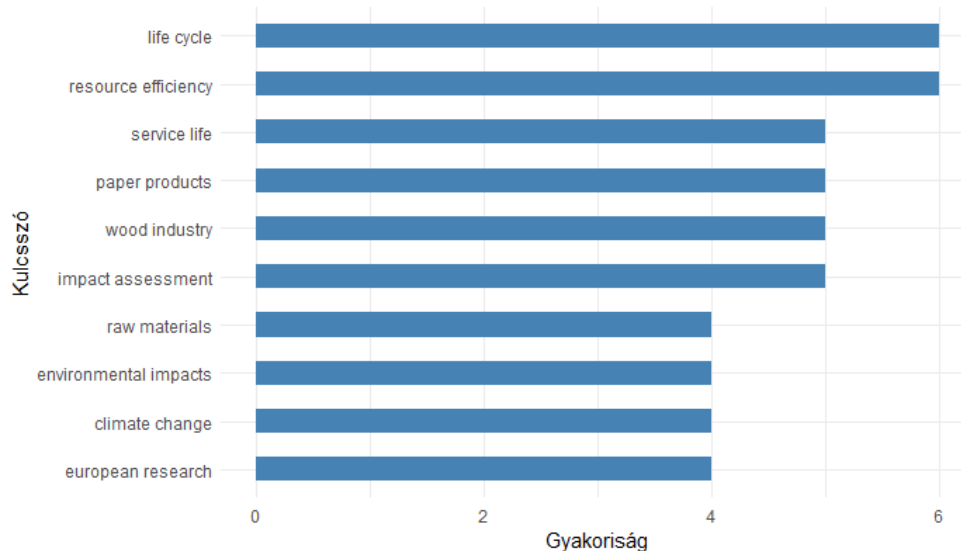
7.5.2. A faipari projektek legfontosabb kulcsszavai

Az FP7(fa) projektekben a legmagasabb RAKE-értékkel bíró szószervezetek – „élettartam”, „szerkezeti faanyag”, „esettanulmány”, „erdőtulajdonos”, „környezeti hatás”, „értéklánc”, „nyersanyag”, „globális piac”, „új rendszer”, „kartongyártás” eredeti megfelelői – meglehetősen általános kifejezéseket tartalmaznak, s nem nagyon jelölnek ki új technológiai irányokat. (96. ábra) Ellenben a forrásbevonás során hatékony hívószavakat jól kijelölhetik.



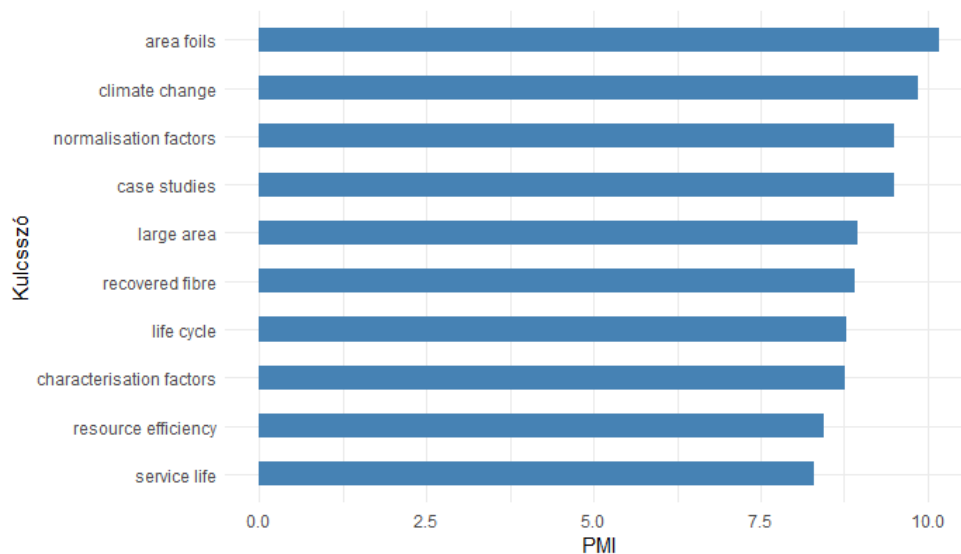
96. ábra Az FP7(fa) projektjeinek absztraktjaiból kinyert, legmagasabb RAKE-értékkel bíró kifejezések

A POS-szekvencia eljárással kapott legfontosabb kifejezések az FP7(fa) projektekben – „életciklus”, „erőforrás-hatékonyság”, „élettartam”, „papírtermékek”, „faipar”, „hatásvizsgálat”, „nyersanyagok”, „környezeti hatások”, „éghajlatváltozás”, „európai kutatás” eredeti megfelelői – is meglehetősen általánosnak tűnnek. Mutatnak némi tartalmi (nem feltétlenül szó szerinti) átfedést az előző eredményekkel, s szintén hasznosak lehetnek a forrásbevonáshoz kapcsolódó pályázatok összeállítása során. Bár az eltelt idő miatt inkább ex post, mint ex ante elemzésre nyílik lehetőségünk. (97. ábra)



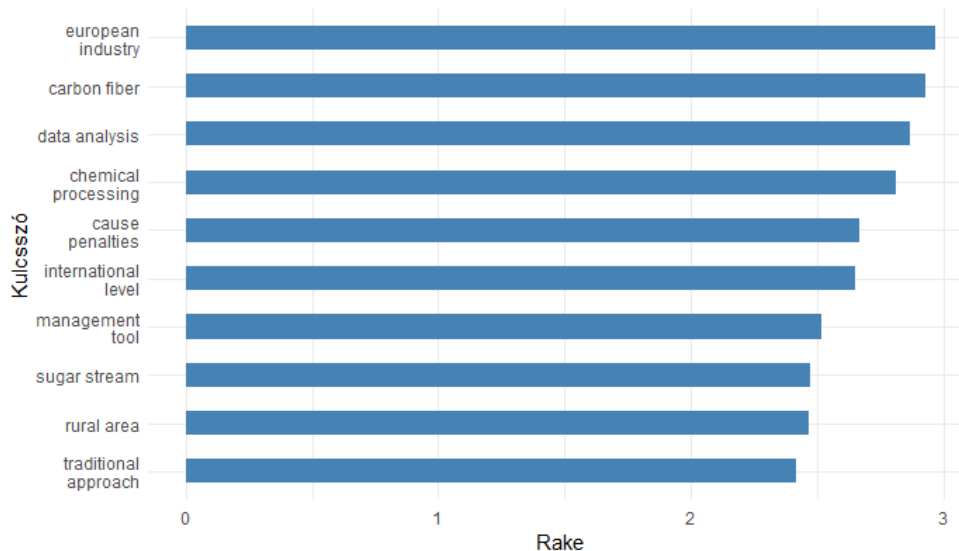
97. ábra A POS-szekvencia eljárással az FP7(fa) absztraktokból kiszűrt legfontosabb kulcsszavak

A PMI-algoritmussal az FP7(fa) projektek leírásaiból kigyűjtött kifejezések – „fedőfóliák”, „klímaváltozás”, „normalizációs tényezők”, „esettanulmányok”, „nagy felület”, „visszanyert szál”, „életciklus”, „jellemző tényezők”, „erőforrás-hatékonyság”, „élettartam” eredeti megfelelői – esetén is csak ismételni tudjuk magunkat; vannak átfedések az előzőekkel, s szintén elég általánosak. (98. ábra)



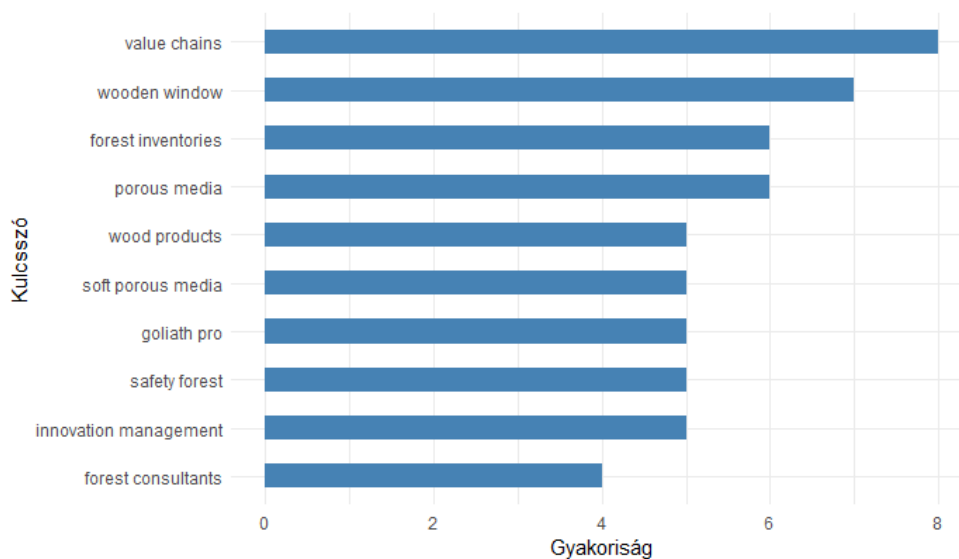
98. ábra A PMI-algoritmussal az FP7(fa) absztraktokból kiszűrt legfontosabb kulcsszavak

A RAKE-algoritmussal a H2020(fa) projektek absztraktjaiból kiszűrt legfontosabb kulcsszavak – „európai ipar”, „szénszál”, „adatelemzés”, „kémiai feldolgozás”, „büntetések generálása”, „nemzetközi szint”, „menedzsment eszköz”, „faalapú cukorszál”, „vidéki terület”, „hagyományos megközelítés” eredeti megfelelői – az általánosságok mellett technikai elemeket is felvil-lantanak. (99. ábra)



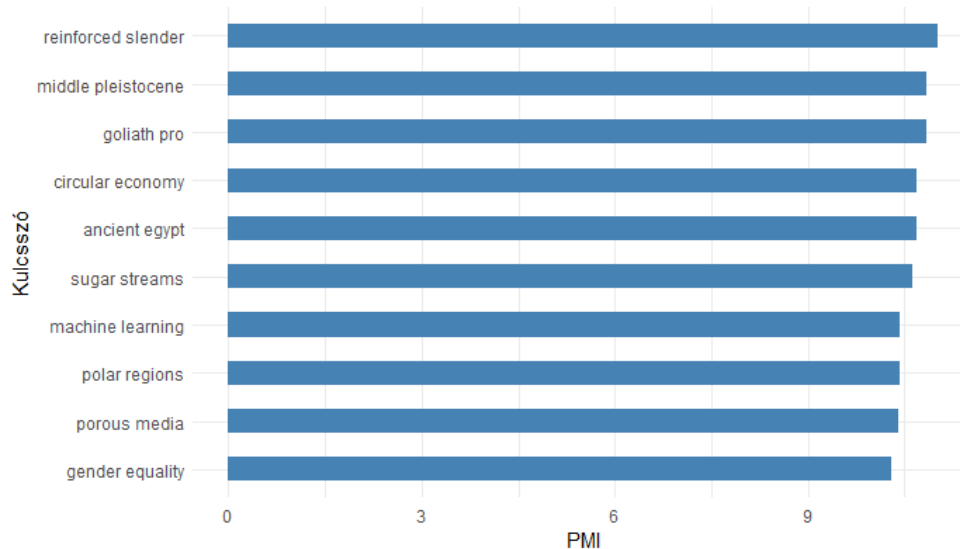
99. ábra A H2020(fa) projektjeinek absztraktjaiból kinyert, legmagasabb RAKE-értékkel bíró kifejezések

A POS-szekvencia eljárással a H2020(fa) projektek absztraktjaiból kinyert legfontosabb kifejezések – „értékláncok”, „faablak”, „erdészeti leltárak”, „porózus közegek”, „fatermékek”, „puha porózus közegek”, „Goliath Pro” (CNC-marógép típus), „biztonsági erdő”, „innovációmenedzsment”, „erdészeti tanácsadók” eredeti megfelelői – általános és technikai elemeket is tartalmaznak, némileg hasonló eredményekkel, mint az előző eljárásnál. (100. ábra)



100. ábra A POS-szekvencia eljárással a H2020(fa) absztraktokból kiszűrt legfontosabb kulcsszavak

A H2020(fa) projektek absztraktjaiból a PMI-algoritmussal kinyert kifejezések – „megerősített vékony lemez”, „középső pleisztocén”, „Goliath Pro”, „körforgásos gazdaság”, „ókori Egyiptom”, „cukorszálak”, „gépi tanulás”, „sarkvidékek”, „porózus közegek”, „nemek közötti egyenlőség” eredeti megfelelői – általános és technikai elemek mellett tartalmaznak néhány meglehetősen meglepő, de valószínűleg a sikeres pályázatot elősegítő kifejezést is (pl. a faipari kutatásokban a „nemek közötti egyenlőség” nem szokványos kifejezés). (101. ábra)



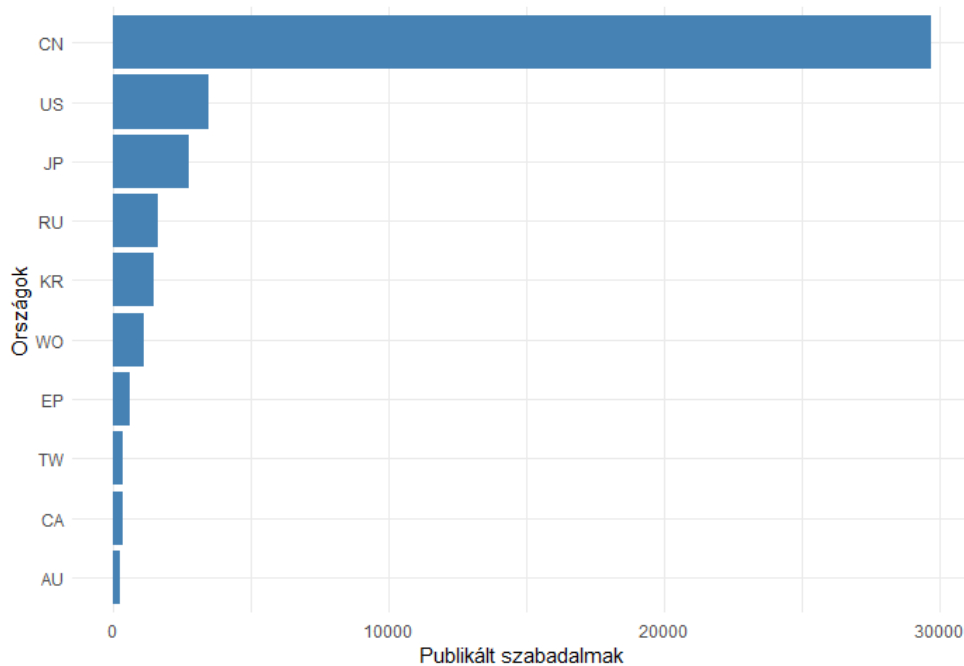
101. ábra A PMI-algoritmussal a H2020(fa) absztraktokból kiszűrt legfontosabb kulcsszavak

Ha a szó szerinti (vagy nagyon közeli) egyezést vizsgáljuk, akkor a három alkalmazott algoritmus az FP7(fa) projektek absztraktjaiból az „élettartam” kifejezést szűrte csak ki minden esetben, a H2020(fa) projektekben viszont nem volt ilyen kifejezés. Legalább két előfordulást tapasztaltunk az FP7(fa) projektek absztraktjainak „életciklus”, „erőforrás-hatékonyság”, „klímaváltozás” kifejezéseinél, illetve a H2020(fa) projektek absztraktjainak „porózus közegek” és „Goliath Pro” kifejezéseinél.

7.5.3. Faipari szabadalmak elemzése

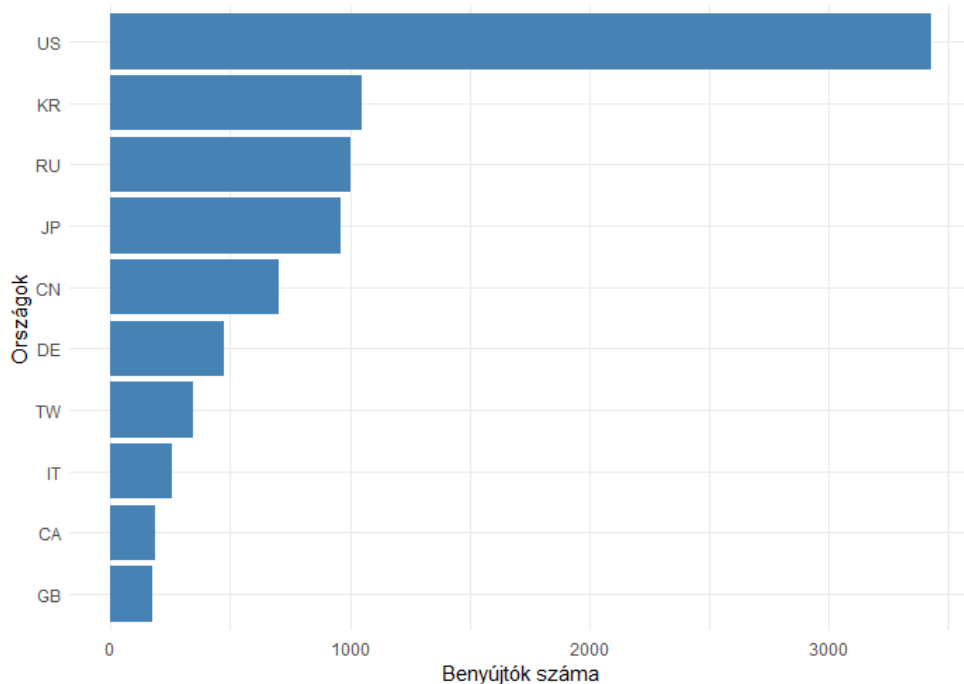
Az Európai Szabadalmi Hivatal Espacenet adatbázisban a „woodworking” keresőszót a 2008-2021-es időszakban, a 2022. december 15-i állapotok szerint 34.481 rekord (szabadalom) tartalmazta. Úgy ítéltük meg, hogy ezek a rekordok döntően faiparinak tekinthetőek, még ha nem is fogják át az összes faiparhoz kapcsolódó szabadalmat.

A vizsgált szabadalmakat 42.356 esetben tették közzé összesen 43 nemzeti vagy nemzetközi szabadalmi hivatalban, tehát egyes szabadalmakat többször is. A legtöbb szabadalmat (29.714) Kínában tették közzé, majd messze leszakadva az USA (3.443), Japán (2.723), Oroszország (1.623) és Dél-Korea (1.445) következik. Az első tizenben egyetlen európai uniós ország sem szerepel, bár a közvetlenül a WIPO és az EPO által közzétett szabadalmak száma (az ábrán WO rövidítéssel: 1.105, ill. EP rövidítéssel: 598) számos európai szabadalmat tartalmaz. (102. ábra) Ennek ellenére az adatok alapján nem úgy tűnik, hogy az Európai Unió lenne a faipari szabadalmak közzétételének fő helyszíne. Közvetlenül a magyar Szellemi Tulajdon Nemzeti Hivatalában (illetve elődjében, a Magyar Szabadalmi Hivatalban) nem nyújtottak be a keresőszónak megfelelő faipari szabadalmat.



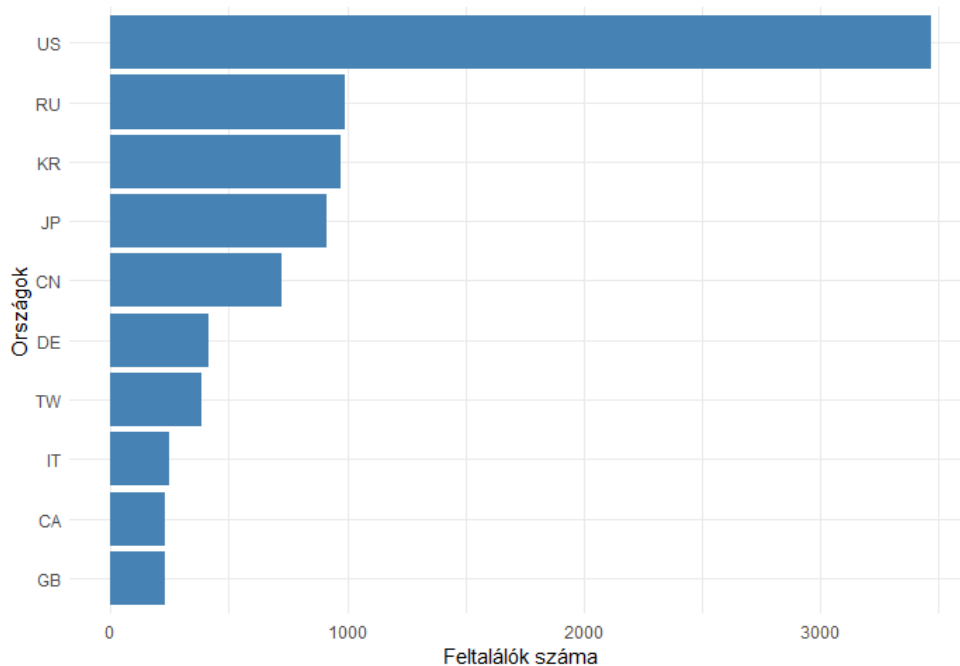
102. ábra A faipari szabadalmak publikálásának tíz legfontosabb helyszíne

A szabadalmi bejelentést benyújtók száma szerint valamelyest más kép bontakozik ki. Összesen 68 országból 9.692-en nyújtottak be faipari szabadalmat. Legtöbben az Egyesült Államokból (3.431), Dél-Koreából (1.048), Oroszországból (1.004), Japánból (959) és Kínából (702). Itt már az európai uniós faipari „nagyhatalmak” sokkal jobban szerepeltek. Németország (473), Olaszország (254) és az időszak nagy részében EU-tag Nagy-Britannia (176) is bekerült az első tízbe. Magyarország a 41. helyen áll nyolc szabadalmi kérelemben való részvétellel.



103. ábra A legtöbb szabadalmi kérelem benyújtásában érintett ország

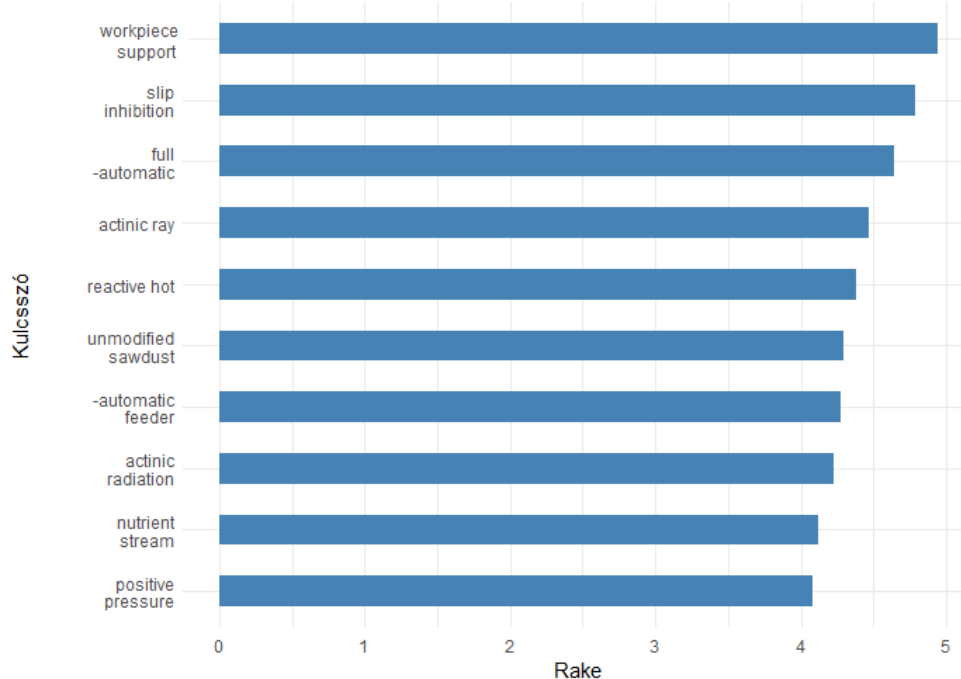
A szabadalmakhoz tartozó 9.778 feltaláló 69 országhoz tartozik. A szabadalmakban megjelölt feltalálók száma szerint kialakuló országsorrend (az első tíz alapján) alig tér el az előzőtől. Az Egyesült Államok (3.472) után Oroszország (993), Dél-Korea (975), Japán (911) és Kína (722) következik. Az európai uniós országok közül megint csak Németország (414), Olaszország (247) és az időszak nagy részében EU-tag Nagy-Britannia (230) fért be az első tízbe. 13 feltalálóval Magyarország Portugáliával holtversenyben 33. a listán.



104. ábra A leggyakoribb, szabadalomban megjelölt feltalálóhoz tartozó országok

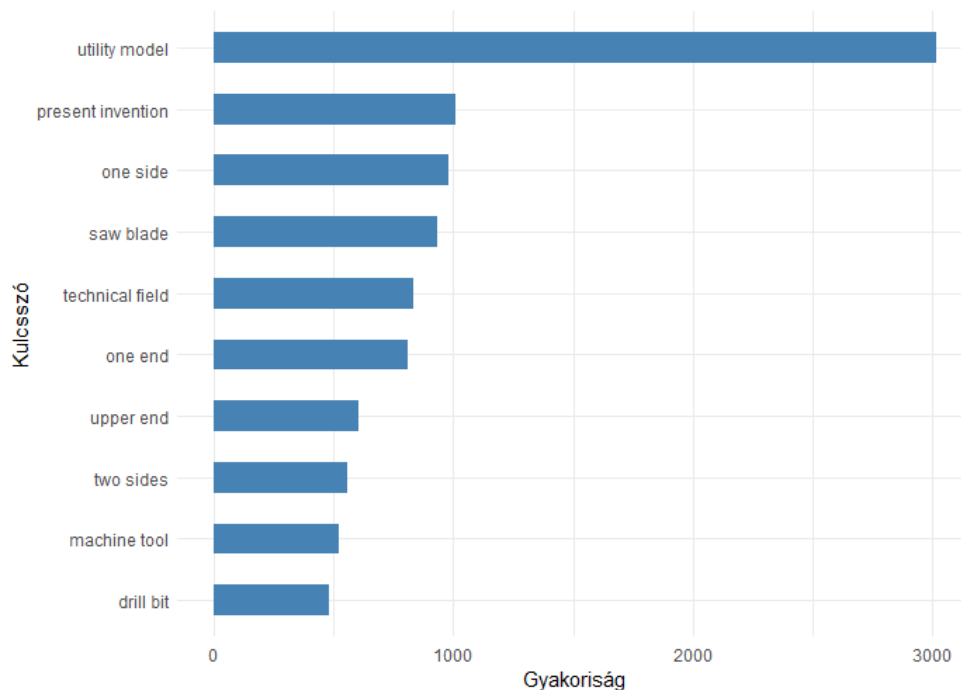
Az adatok mélyebb elemzést igényelnének, de az alapképlet viszonylag egyszerűen felrajzolható. A főként fejlett országokban zajló faipari K+F tevékenység eredményeit igen nagy számban Kínában jegyzik be (vagy Kínában is bejegyzik), ami valószínűsíti, hogy a termelés nagy része is Kínában vagy kínai érdekeltségű cégekben zajlik majd.

A „munkadarab alátámasztása”, „csúsztatógátlás”, „teljesen automatikus”, „aktinikus sugárzás” (azaz kémiai reakciókat generáló elektromágneses sugárzás, két kifejezéssel is), „reaktív forróság”, „nem módosított fűrészpor”, „automata adagoló” (utótagként), „tápanyagáramlás”, „túlnyomás” kulcsszavak (természetesen az eredeti megfelelőik) RAKE-algoritmussal megállapított kiemelkedő helyezése a faipari szabadalmakban igen változatos asszociációkra ad lehetőséget, s kurrens témaként való esetleges kezelésük a hatékony kutatás-fejlesztési stratégiák kidolgozása során további – nagyon komoly számítási igénnyel járó – finomítást igényel. (105. ábra)



105. ábra A szabadalmak absztraktjaiból kinyert, legmagasabb RAKE-értékkel bíró kifejezések

A faipari szabadalmak absztraktjainak POS-szekvenciákra épülő vizsgálata (106. ábra) általánosabb, az iparági fejlesztések általános irányaira jobban utaló kifejezések gyakoriságát szűrte ki, de általános kifejezéseket is olvashatunk a találatok között („használati minta”, „jelen találmány”, „egyik oldal”, „fűrészlap”, „műszaki terület”, „egyik vége”, „felső vége”, „két oldal”, „szerszámgép”, „fűrészszám”, „fűrőfej” eredeti megfelelői).



106. ábra A POS-szekvencia eljárással az absztraktokból kiszűrt leggyakoribb kulcsszavak

A PMI-algoritmus több esetben a legmagasabb értéket állította elő, ezért ezeket értékükkel nem ábrázoltuk, illetve töröltük közülük a dokumentációhoz kapcsolódó, nem szakmai hivatkozásokat, az így előálló szószerkezetek talán érdekesek lehetnek: „glideway crowning” (siklókoronázó), „interactively customize” (interaktívan testreszabható), „posterior endothelial” (hátsó endoteliális – azaz behámsejtek), „Populus euphratica” (*lat.*, eufráteszi nyárfa), „through the envelope” (a burkolaton keresztüli, pl. hőveszteség), „herbal medicines” (növényi gyógyszer), „combating microbes” (mikrobák elleni küzdelem), „boating mooring” (csónakkötő), „chronic obstructive” (krónikusan obstruktív, káros, pl. füstgázok esetén), „obstructive pulmonary” (obstruktív tüdőbetegség), „ischial tubularities” (ischiális tubularitások).

A három különböző algoritmus többféle adatsorra való alkalmazása után úgy tűnik, hogy az egyes projektek legfontosabb kulcsszavai leginkább a sikeres pályázathoz, a faipari szabadalmakból kinyerhető kulcsszavak pedig az ezt és más forrásbevonási vagy gazdasági sikereket megalapozó kutatás-fejlesztési tevékenységhez nyújthatnak támpontot.

A fentiek szerint a 19. szakmai hipotézist, miszerint az FP7 és a H2020 támogatási program faipari projektjeinek absztraktjai és a nemzetközi szinten elérhető faipari szabadalmak absztraktjai között mérsékelt átfedés lehet, egyértelműen nem utasítjuk el. Azonban az eredmények nem eléggé megalapozottak ahhoz, hogy tézist is megfogalmazzunk.

8. Összefoglalás

Az értekezés keretében szakirodalmi vizsgálatok mellett szövegbányászati, leíró és következtetési statisztikai, illetve hálózatkutatási eszközökkel főként az FP7 és a H2020 program faipari projektjeit, projektrésztvevőit, a projektrésztvevők kapcsolatait és az egyes projekteredményeket (a faipari szabadalmakkal való tartalmi összevetéssel) elemeztük. Az elemzés során fontosnak tartottuk a hatékonyság vizsgálatát, azonban a hatékonyságot nemcsak a közvetlen outputból kinyerhető gazdasági eredmény, hanem a hatékony forrásbevonási stratégia szempontjából is vizsgáltuk, tehát a dolgozat a közvetlen leíró jelleg mellett egyben jövőorientált. A szakirodalom és kutatási projektek általánosan megfogalmazott céljaihoz igazodva fontosnak éreztük a faipar és a fenntarthatóság összeegyeztethetőségének hangsúlyozását, a fenntartható faipar vizsgálatát is.

A szakirodalmi kutatás kiterjedt a fenntartható faiparra, a faipari projektekre, illetve általában a kapcsolatháló elemzésre és konkrétan az európai uniós projektek kapcsolatháló-elemzésére vonatkozó szakirodalomra.

A dolgozat empirikus kutatásai során több adatforrást használtunk fel. A faipari projekteket vizsgáló cikkek szövegbányászati és tudományometriai eszközökkel történő elemzését a Scopusból kiszűrt és letöltött BibTeX-állományra építettük. A fenntartható faipar tudományos súlyának vizsgálati adatforrása szintén egy Scopusból kiszűrt és letöltött adatállomány volt. Az FP7 és a H2020 faipari projektjeinek elemzéséhez a CORDIS-ban elérhető adattáblákat letöltöttük, majd relációs adatbázissá alakítottuk, végül kiszűrtük a szükséges sorokat. Az FP7 és a H2020 faipari projektek hálózat- és klaszterelemzését tartalmi vizsgálatok és szűrések után az előzőleg felhasznált adattáblákat kiegészítve végeztük el. A faipari projektek outputjait az említett, CORDIS-adatállományra felépített adatbázisból szűrtük ki. A tartalomelemzéshez az Európai Szabadalmi Hivatal Espacenet és a Szellemi Tulajdon Világszervezetének Patentscope szolgáltatásából letöltött adatokat használtuk fel.

Az elemzések során számos egy- és többváltozós módszert felhasználtunk. A többváltozós (esetünkben kétváltozós) módszerek – korreláció, lineáris és nem lineáris regresszió, χ^2 próba használata előtt törekedtünk azok előfeltételeinek vizsgálatára vagy biztosítására. Különböző módszerekkel egy- és kétdimenziós outliereket kerestünk, az adatok normalitása mellett a reziduálisok (reziduuumok) normalitását és a homoszkedaszticitást is vizsgáltuk. Az SNA-elemzés során számos központi mutatót számoltunk, kapcsolathálót és klaszterhálót rajzoltunk, a hálózatelemzés során nevezetes eloszlásokat vizsgáltunk. A vizsgálatok során reagáltunk a következtetési statisztikával kapcsolatos szakirodalmi vitákra. Ahol a legkisebb értelme/jelentősége volt, következtetési statisztikai módszereket is használtunk, de jeleztük, hogy az eredmény szakmai megközelítéstől függően esetleg megkérdőjelezhető. A gyakran használt próbák (pl. Kolmogorov-Szmirnov, Shapiro-Wilk) mellett a szakirodalmi példákat követve ritkábban használt próbákat is alkalmaztunk (pl. Anderson-Darling, Breusch-Pagan, White, Goldfeld-Quandt). A talán túlcsondulónak tűnő próbahasználatot esetenként manapság kevésbé divatos módszerekkel (pl. a szórásnégyzet összetevőkre bontása, divergenciák) ellensúlyoztuk. A szövegbányászati elemzések során leginkább RAKE-, POS-, PMI- algoritmusokat használtunk, de adtunk példát dimenziócsökkentési és klaszterezési eljárásra is.

A kutatás során Microsoft Excel, JASP, IBM SPSS statisztikai szoftvereket és GNU R statisztikai programozási nyelvet használtunk. A kapcsolódó számítások, fájlok, kódok többségét a számítások reprodukálhatóságának kedvéért közzétettük a zenodo.org oldalon:

https://zenodo.org/communities/novotni_adrienn_phd/

Eredményeink szerteágazóak. A szakirodalomról nyújtott kép alapján beigazolódott, hogy az általunk használt módszerek az FP7 és a H2020 faipari projektjeinek elemzése során több szempontból újszerűnek tekinthetők. A felhasznált módszerek külön-külön szórványosan megjelennek a szakirodalomban, de az általunk vizsgált adatok körében és az általunk használt módszertani „patternre” nem találtunk példát. Megállapítottuk, hogy a faipari projektek szövegbányászati és tudományometriai elemzése alapján a folyóiratokra vonatkozó Bradford-eloszlás nem érvényes, a szerzői produktivitásra vonatkozó Lotka-eloszlás viszont igen. A publikációk dimenziócsökkentési eljárásokkal, kulcsszavaik szerint 4-5 klaszterre oszthatók. A fenntartható faiparhoz kapcsolódó publikációk időszorelemzése alapján megállapítható, hogy a faipar szerepének fenntarthatóságban játszott szerepének felismerése egyelőre nem egyértelmű. Az FP7-hez képest a H2020 faipari projektekben a közösségi támogatás kevésbé volt egyenletes. A koordinátorok átlagos támogatása mindkét támogatási programban magasabb, mint a résztvevőké. A két programban való országonkénti részvétel és támogatás között közepes erősségű korreláció és regresszió tapasztalható. A tevékenység alapján mért részvétel különbséget mutat a két program között. A tevékenység alapján a két program között a támogatás súlypontja látszólag jelentősen átalakult, de ez nem az egyes csoportok támogatásának különbségére, hanem a belső szórásnégyzet alapján a csoportokon belüli különbségekre vezethető vissza. Az egy résztvevőre eső közösségi támogatás mediánja – különösen reálértékben – csökkent a két időszak között. A többes projektrészvétel magasabb támogatással jár. Az FP7 faipari projektjeiben résztvevő országok közösségi támogatása a faipari kibocsátással közepes rangkorrelációt és közepesnél gyengébb lineáris regressziót mutat. A faipari projektháló töredezettségét a tartalomvizsgálat segítségével részben faiparinak minősített projektek részvevőivel való kiegészítés szignifikánsan nem csökkentette. A vizsgált kapcsolatháló hatványfüggvény-eloszlást követtek egy bizonyos minimumérték felett, az értékeket a bootstrapping-eljárások érdemben nem módosították. Egyes jelek alapján a tényleges együttműködések a minimumértékek alatt is hatványfüggvény-eloszlás felé mutathatnak. A faipari projektek foksám és közöttség alapján kimutatható központi aktorai a vizsgált csoportokban különböző mértékű átfedést mutattak. A támogatásokhoz képest a faipari projektek eredményeinek száma nem kimagasló, tartalmi szempontból szórványos átfedést mutatnak a faipari szabadalmak absztraktjaival.

9. Felhasznált szakirodalom

- Aagaard, A. K. & Larsen, N. M., 2020. Developing a fabrication workflow for irregular sawlogs. *International Journal of Architectural Computing*, Volume 18, pp. 270-283.
- Aberger, E., Koppelhuber, J. & Heck, D., 2018. *Building information modeling in timber construction - A solution for planning process, design phases and the unification of scope of works*. s.l., World Conference on Timber Engineering (WCTE).
- Acampa, G., Grasso, M., Marino, G. & Parisi, C., 2020. Tourist Flow Management: Social Impact Evaluation through Social Network Analysis.
- Acock, A. C., 2018. *A Gentle Introduction to Stata*. 6 szerk. Lakeway Drive: Stata Press.
- Adamowicz, K. & Noga, T., 2017. Assessment applicability of selected models of multiple discriminant analyses to forecast financial situation of Polish wood sector enterprises. *Folia Forestalia Polonica, Series A*, Volume 59, pp. 59-67.
- Adhikari, B., 2019. United Nations general assembly voting and foreign aid bypass. *International Politics*, 56(4), pp. 514-535.
- Ahmed, W., Seguí, F. L., Vidal-Alaball, J. & Katz, M., 2020. COVID-19 and the “Film Your Hospital” Conspiracy Theory: Social Network Analysis of Twitter Data.
- Ahmed, W., Vidal-Alaball, J., Seguí, F. L. & Moreno-Sanchez, P. A., 2020. A Social Network Analysis of Tweets Related to Masks during the COVID-19 Pandemic.
- Akimova, E., 2021. Limitations in the Use of Social Network Analysis Methods in Marketing Project Management.
- Albert, R. & Barabási, A.-L., 2002. Statistical mechanics of complex networks. *Reviews of Modern Physics*, Issue 74, pp. 47-97.
- Aliprandi, C. et al., 2014. *CAPER: Crawling and analysing Facebook for intelligence purposes*. s.l., Institute of Electrical and Electronics Engineers Inc., pp. 665-669.
- Aria, M. & Cuccurullo, C., 2017. bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, , 11(4), pp. 959-975.
- Aslam, W., Butt, W. H. & Anwar, M. W., 2018. *A Systematic Review on Social Network Analysis: Tools, Algorithms and Frameworks*. New York, NY, USA, Association for Computing Machinery, p. 92–97.
- Backhausz, Á., 2020. *Nemparaméteres próbák: Kétnégyzet próbák*, Budapest: ELTE.
- Balest, J., Pezzutto, S., Giacobelli, G. & Wilczynski, E., 2022. Engaging Stakeholders for Designing a FAIR Energy Data Management Tool: The Horizon 2020 EnerMaps Project.
- Balland, P.-A., Boschma, R. & Ravet, J., 2019. Network dynamics in collaborative research in the EU, 2003–2017. *European Planning Studies*, Volume 27, pp. 1811-1837.
- Barabási, A.-L., 2003. *Behálózva – a hálózatok új tudománya*. Budapest: Magyar Könyvklub.

- Barabási, A.-L., 2006. A hálózatok tudománya: a társadalomtól a webig. *Magyar Tudomány*, Issue 11, p. 1298.
- Barabási, A.-L., 2017. *A hálózatok tudománya*. Budapest: Libri.
- Barabási, A.-L., 2018. *Network science*. Cambridge: Cambridge University Press.
- Barabási, A.-L. & Albert, R., 1999. Emergence of scaling in random networks. *Science*, Issue 286, pp. 509-512.
- Bárdits, A., Németh, R. & Terplán, G., 2016. Egy régi probléma újra előtérben: a nullhipotézis szignifikanciateszt téves gyakorlata. *Statisztikai Szemle*, 94(1), pp. 52-75.
- Barriault, F. et al., 2017. The softwood lumber dispute: Is a solution possible. *Forestry Chronicle*, Volume 93, pp. 9-16.
- Barucca, P. és mtsai., 2020. Network valuation in financial systems. 30. kötet, pp. 1181-1204.
- Binder, D., 1983. On the variances of asymptotically normal estimators from complex surveys. *International Statistical Review*, Issue 51, pp. 279-292.
- Boglioli, M. et al., 2022. Searching for Culture in “Cultural Capital”: The Case for a Mixed Methods Approach to Production Facility Siting. *Frontiers in Energy Research*, Volume 9.
- Bollobás, B., 1980. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4), pp. 311-316.
- Brandner, R., Ringhofer, A. & Reichinger, T., 2019. Performance of axially-loaded self-tapping screws in hardwood: Properties and design. *Engineering Structures*, Volume 188, pp. 677-699.
- Burcher, M., 2020. *Social Network Analysis and Crime Intelligence*. New York: Springer International.
- Burov, M. P., Vershinin, V. V. & Kovaleva, T. N., 2021. *Geospatial factors of the organization of sustainable forest management*. s.l., IOP Publishing Ltd.
- Butko, G. et al., 2021. *Assessment of the innovative potential of the timber enterprises in the Russian Ural region*. s.l., IOP Publishing Ltd.
- Buyuktasgin, H. A. A., Aslankaya, G. & Dilmaghani, M., 2017. Load bearing timber-glass composites: New opportunities for Turkey. *Baltic Forestry*, Volume 23, pp. 698-705.
- Calvo-Gallardo, E., Arranz, N. & de Arroyabe, J. C. F., 2022. Innovation systems' response to changes in the institutional impulse: Analysis of the evolution of the European energy innovation system from FP7 to H2020. *Journal of Cleaner Production*, Volume 340.
- Capone, F., 2014. Technological clusters and innovation trajectories in shipbuilding in Europe: An analysis of FP6 and FP7 European projects. *Australian Journal of Maritime and Ocean Affairs*, Volume 6, pp. 89-105.
- Casella, G. & Berger, R., 2002. *Statistical Inference*. 2 szerk. Duxbury: Thomson Learning.

- Casey, S. és mtsai., 2017. *Developing a Sustainable Roadmap for Ebony Production in Cameroon*. Los Angeles, CA: UCLA Environmental Science Senior Practicum 2016-17.
- Chambers, T., 2019. 'Performed Conviviality': Space, bordering, and silence in the city. *Modern Asian Studies*, Volume 53, pp. 776-799.
- Cha, Y. & Cho, J., 2012. *Social-network analysis using topic models*. New York, NY, USA, Association for Computing Machinery, p. 565–574.
- Choudhury, A., Kaushik, S. & Dutt, V., 2017. *Social-Network Analysis for Pain Medications: Influential physicians may not be high-volume prescribers*. New York, NY, USA, Association for Computing Machinery, p. 881–885.
- Ciapetti, A., Ruggiero, G. & Toti, D., 2019. A semantic knowledge discovery framework for detecting online terrorist networks. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Volume 11296 LNCS, pp. 120-131.
- Clauset, A., Shalizi, C. R. & Newman, M. E. J., 2009. Power-law distributions in empirical data. *SIAM Review*, 51(4), pp. 661-703.
- Coombs, S. & Forster, W., 2017. *An additive approach to the building envelope using Welsh-grown timber*. s.l., NCEUB 2017 - Network for Comfort and Energy Use in Buildings, pp. 1487-1494.
- Cullen, A. & Frey, H., 1999. *Probabilistic Techniques in Exposure Assessment*. London: Plenum Publishing Co.
- Czaplicki, L. M. et al., 2016. A New Perspective on Sustainable Soil Remediation-Case Study Suggests Novel Fungal Genera Could Facilitate in situ Biodegradation of Hazardous Contaminants. *Remediation*, Volume 26, pp. 59-72.
- Davies, D. W., Johnson, L. L., Corigliano, T. P. & Young, M. P., 2017. *Carbon choices: Bringing CO2 into life-cycle decision-making*. s.l., CRC Press/Balkema, pp. 3-20.
- DeFries, R. és mtsai., 2013. Export-oriented deforestation in Mato Grosso: Harbinger or exception for other tropical forests?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1619), pp. 1-8.
- Delignette-Muller, M. L. & Dutang, C., 2020. *fitdistrplus: An R Package for Fitting Distributions*, Vienna: R Consortium.
- Derikvand, M., Nolan, G., Jiao, H. & Kotlarewski, N., 2017. What to Do with Structurally Low-Grade Wood from Australia's Plantation Eucalyptus; Building Application?. *BioResources*, Volume 12, pp. 4-7.
- Derntl, M. & Klamma, R., 2012. *Social network analysis of European project consortia to reveal impact of technology-enhanced learning projects*. Rome, s.n., pp. 746-747.

Dieste, A., Cabrera, M. N., Clavijo, L. & Cassella, N., 2019. Analysis of wood products from an added value perspective: The Uruguayan forestry case. *Maderas: Ciencia y Tecnologia*, Volume 21, pp. 305-316.

Dobretsov, R. Y. et al., 2020. *Power distribution mechanism for the transmission of forest tracked and wheeled vehicles*. s.l., IOP Publishing Ltd.

Duarte, F., 2013. *From framework programme 7 to Horizon 2020*. New York, NY, USA, Association for Computing Machinery, p. 142.

Efron, B. & Tibshirani, R., 1994. *An Introduction to the Bootstrap*. London: Chapman & Hall.

Egghe, L., 1986. The Dual of Bradford's Law. *Journal of American Society for Information Science*, Issue 37., pp. 173-89.

Ellefson, P. V., 2019. *Problem orientation and investments in research programs*. s.l.:Taylor and Francis.

Enger, S. G., 2018. Closed clubs: Network centrality and participation in Horizon 2020. *Science and Public Policy*, Volume 45, pp. 884-896.

Erdős, P. & Rényi, A., 1960. On The Evolution of Random Graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, Issue 5, pp. 17-61.

Esteve-Codina, A. és mtsai., 2021. Gender specific airway gene expression in COPD sub-phenotypes supports a role of mitochondria and of different types of leukocytes. 11. kötet.

FAO, 2018a. *Concept note for a JOINT INITIATIVE of the Collaborative Partnership on Forests*. Rome: FAO Headquarters.

FAO, 2018b. *Sustainable Wood for a Sustainable World, Global Meeting Report*, Rome: FAO Headquarters.

FAOSTAT, 2023. *Forestry Production and Trade*. [Online]
Hozzáférhető: <https://www.fao.org/faostat/en/#data/FO>
[Hozzáférés dátuma: 20 01 2023].

Favilli, E., Rossi, A. & Brunori, G., 2015. Food networks: collective action and local development. The role of organic farming as boundary object.

Fernandez de Arroyabe, J. C., Schumann, M., Sena, V. & Lucas, P., 2021. Understanding the network structure of agri-food FP7 projects: An approach to the effectiveness of innovation systems. *Technological Forecasting and Social Change*, Volume 162.

Filvà, D. A., García-Peñalvo, F. J. & Forment, M. A., 2014. *Social network analysis approaches for social learning support*. New York, NY, USA, Association for Computing Machinery, p. 269–274.

Fischer-Hübner, S., szerk., 2012. *Trust, privacy and security in digital business*. Heidelberg [u.a.], Springer.

Forest Research, 2023. *Conversion factors*. [Online]

Hozzáférhető: <https://www.forestresearch.gov.uk/tools-and-resources/statistics/forestry-statistics/forestry-statistics-2018/sources/timber-2/conversion-factors/>
[Hozzáférés dátuma: 20 01 2023].

Freeman, L. C., 1977. A Set of Measures of Centrality Based on Betweenness. *Sociometry*, pp. 35-41.

Freeze, T. A. és mtsai., 2020. *Improving the health behaviours of COPD patients: Is health literacy the answer?*. Vermont: Vermont UP.

Friendly, M. & Meyer, D., 2015. *Discrete Data Analysis with R: Visualization and Modeling Techniques for Categorical and Count Data*. hely nélkül.:Chapman and Hall/CRC.

Gao, T. & Jojic, V., 2016. *Degrees of Freedom in Deep Neural Networks*, Ithaca NY: Cornell University.

Gau, R.-H., Hsieh, T.-C., Tsai, S.-W. & Cheng, C.-P., 2011. *An implementation framework of mapreduce email social network analysis*. New York, NY, USA, Association for Computing Machinery, p. 67–70.

Gazis, R. et al., 2018. Mycobiota associated with insect galleries in walnut with thousand cankers disease reveals a potential natural enemy against *Geosmithia morbida*. *Fungal Biology*, Volume 122, pp. 241-253.

Gibson, C. & Warren, A., 2020. Keeping time with trees: Climate change, forest resources, and experimental relations with the future. *Geoforum*, Volume 108, pp. 325-337.

Gillespie, C. S., 2020. *The poweRlaw package: Examples*, Vienna: R Consortium.

Golovina, E. Y. & Dykusova, A. G., 2020. *Financing of Innovative Projects in the Forest Industry*. s.l., Institute of Physics Publishing.

Gordeev, R. V., Pyzhev, A. I. & Yagolnitsers, M. A., 2021. Drivers of spatial heterogeneity in the russian forest sector: A multiple factor analysis. *Forests*, Volume 12.

Goryunov, A. G., Goryunova, N. N., Ogunlana, A. O. & Manenti, F., 2016. Production of energy from biomass: Near or distant future prospects?. *Chemical Engineering Transactions*, Volume 52, pp. 1219-1224.

Gourbesville, P., 2012. Urban flooding and Resilience: concepts and needs. *Geophysical Research Abstracts*, April, 14. kötet, p. 14282.

Greenland, S. és mtsai., 2016. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol.*, Issue 31, p. 337–350.

Grohmann, D., Prospero, F. & Menconi, M. E., 2020. *Tilia sp.'s pruning residues wood panels for thermal insulation*. s.l.:Elsevier.

Guarino, S., Trino, N., Chessa, A. & Riotta, G., 2020. Beyond Fact-Checking: Network Analysis Tools for Monitoring Disinformation in Social Media. *Studies in Computational Intelligence*, Volume 881 SCI, pp. 436-447.

Guerrero-Ocampo, S. B., Díaz-Puente, J. M. & Espinoza, J. F. N., 2022. Multi-Actor Partnerships for Agricultural Interactive Innovation: Findings from 17 Case Studies in Europe.

Gu, W., Luo, J.-d. & Liu, J., 2019. Exploring small-world network with an elite-clique: Bringing embeddedness theory into the dynamic evolution of a venture capital network. 57. kötet, pp. 70-81.

Hall, H., Cruickshank, P. & Ryan, B., 2018. Long-term community development within a researcher network: A social network analysis of the DREaM project cadre.

Hampel, F. R., 1985. The Breakdown Points of the Mean Combined with Some Rejection Rules. *Technometrics*, 27(2), pp. 95-107.

Hampson, C. et al., 2012. The CULTURA project: Supporting next generation interaction with digital cultural heritage collections. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Volume 7616 LNCS, pp. 668-675.

Hao, F. és mtsai., 2017. K-Clique Community Detection in Social Networks Based on Formal Concept Analysis. *IEEE Systems Journal*, 11(1), p. 250–259.

Hartley, H. & Sielken, R., 1975. A "Super-Population Viewpoint" for Finite Population Sampling. *Biometrics*, 31(2), pp. 411-422.

Henders, S., Persson, U. M. & Kastner, T., 2015. Trading forests: Land-use change and carbon emissions embodied in production and exports of forest-risk commodities. *Environmental Research Letters*, 10(12), p. Article number 125012.

Henriksson, F. & Johansen, K., 2018. *Integrated Product and Production Research on Introducing Internet of Things in Swedish Wood Industry Products*. s.l., Elsevier B.V., pp. 10-16.

Herrera, R. F., Matus, J., Santelices, C. & Atencio, E., 2020. Interaction between project management processes: a social network analysis. 12. kötet, p. 133.

Hofmair, M. W. et al., 2017. Patching process optimization in an agent-controlled timber mill. *Journal of Intelligent Manufacturing*, Volume 28, pp. 69-84.

Hung, M. és mtsai., 2020. Social Network Analysis of COVID-19 Sentiments: Application of Artificial Intelligence.

Hunyadi, L. & Vita, L., 2016. Számúzott szignifikanciatesztek. *Statisztikai Szemle*, 94(4), pp. 436-444.

IBM Corp., 2020. *IBM SPSS Advanced Statistics V27*, Armonk: IBM.

- Isa, M. U. et al., 2020. Enhancing employability skills of woodwork technology education students through project based learning at colleges of education (technical) in Nigeria. *Universal Journal of Educational Research*, Volume 8, pp. 31-40.
- Ivanova, A. V., Rafailov, M. K., Matveev, S. M. & Sibiryatkina, A. V., 2018. *Project management in the forestry complex of Russia*. s.l., International Business Information Management Association, IBIMA, pp. 3771-3777.
- Ivantsova, E. D., 2021. *Success Factors of the Implementation of State Measures of Investments Stimulation in Russian Timber Industry*. s.l., IOP Publishing Ltd.
- Jankowski, K., Knutas, A., Ikonen, J. & Porras, J., 2015. *Automated social network analysis of online student collaboration activity*. New York, NY, USA, Association for Computing Machinery, p. 326–333.
- Karamouzas, D., Mademlis, I. & Pitas, I., 2022. Public opinion monitoring through collective semantic analysis of tweets. 12. kötet.
- Kay, K., 2017. Rural Rentierism and the Financial Enclosure of Maine's Open Lands Tradition. *Annals of the American Association of Geographers*, Volume 107, pp. 1407-1423.
- Kekkonen, M. & Isoherranen, V., 2021. *Mechanical engineering students project-based learning in OUAS*. s.l., University of Minho, pp. 347-352.
- Kingdon, P., 2021. Correction to: The cosmopolitan engineering student: an analysis of a recruitment campaign for KTH Royal Institute of Technology in Stockholm (*International Journal of Technology and Design Education*, (2018), 28, 3, (787-802), 10.1007/s10798-017-9405-4). *International Journal of Technology and Design Education*, Volume 31, pp. 199-200.
- Kisfalusi, D., 2018. Az exponenciális random gráf modellek bemutatása: Egy iskolai osztály baráti hálózatának modellezése. *Szociológiai Szemle*, 28(2), pp. 75-88.
- Kitrinariis, A., 2018. *Cradle to cradle regenerative design: From circular economy to sustainable construction*. s.l., WITPress, pp. 3-14.
- Klarić, S. & Obučina, M., 2020. New Trends in Engineering Wood Technologies. *Lecture Notes in Networks and Systems*, 76. kötet, pp. 712-727.
- Klitkou, A., Fevolden, A. M. & Andersen, A. D., 2022. EU R&D Funding for Electricity Grid Technologies and the Energy Transition: Centralised versus Decentralised Transition Pathways. *Energies*, Volume 15.
- Klöckner, C. A., 2019. *Understanding the social dynamics of consumer energy choices - Some lessons learned from two H2020 projects (ECHOES, SMARTTEES)*. s.l., European Council for an Energy Efficient Economy, pp. 3-12.
- Kolesnik, V. G. & Sinyatullina, L. K., 2017. State management of the forestry complex: Current situation and main challenges. *Public Administration Issues*, pp. 129-148.

- Koppelhuber, J., Hintersteiner, K. & Heck, D., 2016. *Industrialized timber construction: Construction management aspects and influences in modular timber building systems*. s.l., ISEC Press, pp. 399-404.
- Kropivšek, J., Oblak, L., Zupančič, A. & Jošt, M., 2016. *Impact of organizational learning on the development of competencies: Case of Slovenian wood-industry*. s.l., WoodEMA, i.a., pp. 93-99.
- Kujawińska, A., Rogalewicz, M., Diering, M. & Hamrol, A., 2017. Statistical approach to making decisions in manufacturing process of floorboard. *Advances in Intelligent Systems and Computing*, Volume 571, pp. 499-508.
- Kürtösi, Z., 2011. A társadalmi kapcsolatháló-elemzés módszertani alapjai. In: T. Károly, szerk. *Társadalmi kapcsolathálózatok elemzése*. Budapest: BCE Szociológia és Társadalompolitika Intézet, pp. 19-31.
- Lallemand, F. & Guérin, A.-J., 2017. What is the future of european forests confronted with climate change and the carbon neutrality objective? [Quel avenir pour la forêt européenne face au changement climatique et à l'objectif de neutralité carbone?]. *Revue Forestiere Francaise*, Volume 69, pp. 259-271.
- Larchenko, Y. G., Shusharina, G. A. & Tretyakov, A. V., 2022. State and Prospects of Logging and Sawmilling Waste Management in Khabarovsk Territory. *Advances in Science, Technology and Innovation*, pp. 887-892.
- Larntz, K., 1978. Small-Sample Comparisons of Exact Levels for Chi-Squared Goodness-of-Fit Statistics. *Journal of the American Statistical Association*, 73(362), pp. 253-263.
- Le Roux, S., Bannier, F., Bossanne, E. & Stieglmeier, M., 2016. *Investigating the interaction of building information modelling and lean construction in the timber industry*. s.l., Vienna University of Technology.
- Lee, C.-Y., Chong, H., Liao, P.-C. & Wang, X., 2018. Critical Review of Social Network Analysis Applications in Complex Project Management.
- Lemonnier, F. & Millet, P., 2012. *FlexTiles: self adaptive heterogeneous manycore based on flexible tiles (FP7 project)*. New York, NY, USA, Association for Computing Machinery, p. 37–38.
- Lim, H. L., 2009. *Social network analysis: measuring symmetry of information flow in virtual learning groups*. New York, NY, USA, Association for Computing Machinery, p. 40–44.
- Li, S. & Lu, J., 2020. Business model innovation of China's New Energy Vehicle industry based on Social Network Analysis: Demonstration city cases.
- Lombardi, M. et al., 2020. NETWORK IMPACT OF SOCIAL INNOVATION INITIATIVES IN MARGINALISED RURAL COMMUNITIES. *Social Networks*, Volume 63, pp. 11-20.
- Lovrić, M. és mtsai., 2018. Social network analysis as a tool for the analysis of international trade of wood and non-wood forest products. 86. kötet, pp. 45-66.

Ludovico, N., Dessi, F. & Bonaiuto, M., 2020. Stakeholders Mapping for Sustainable Biofuels: An Innovative Procedure Based on Computational Text Analysis and Social Network Analysis.

Lüthi, T., Gautschi, M. & Lädach, T., 2019. How can forestry and the timber industry participate more in the timber construction boom? (Essay) [Comment la forêt et l'industrie du bois peuvent-elles participer davantage au succès de la construction en bois? (Essai)] [Wie können der Wald und die Holzindustrie vermehrt am «Holzbauboom» teilhaben? (Essay)]. *Schweizerische Zeitschrift für Forstwesen*, Volume 170, pp. 176-181.

Macháček, V., 2019. *Globalization of science: Evidence from authors in academic journals by country of origin*. Prague, Charles University Prague, Faculty of Social Sciences, Institute of Economic Studies, p. 339–350.

Marinova, Z. és mtsai., 2019. WeVerify: Wider and Enhanced Verification for You - Project Overview and Tool Demonstration.

Marsh, M. & Nelson, J., 2018. *Structural resiliency through sustainability*. s.l., Earthquake Engineering Research Institute, pp. 5629-5640.

Martins, R. et al., 2022. Technical and Economic Analysis of the Implementation of a Self-Sustainable Briquetting Process for Electric Generation. *Energies*, Volume 15.

Mbatu, R. S., 2020. Discourses of FLEGT and REDD + regimes in cameroon: A nongovernmental organization and international development agency perspectives. *Forests*, Volume 11.

Millard, C., Brhlikova, P. & Pollock, A., 2015. Social networks and health policy: The case of misoprostol and the WHO model essential medicine list. *Social Science and Medicine*, Volume 132, pp. 190-196.

Molloy, M. & Reed, B. A., 1995. Critical point for random graphs with a given degree sequence. *Random Structures & Algorithms*, 6(2–3), p. 161–180.

Molnár, L., 2020. A hálózatelemzés alapfogalmai – gráfok, centralitás, szomszédosság, hidak és a kis világ. In: *Rendszerelmélet*. Budapest: Dialóg Campus, pp. 123-140.

Muscio, A., Ciffolilli, A. & Lopolito, A., 2022. Technological diversity in collaborative projects: insights into European research policy. *Journal of Economic Policy Reform*, Volume 25, pp. 322-343.

Münnich, Á., Nagy, Á. & Abari, K., 2006. *Többváltozós statisztika pszichológus hallgatók számára*. Debrecen: Bölcsész Konzorcium.

Nambiar, E., 2019. Re-imagining forestry and wood business: Pathways to rural development, poverty alleviation and climate change mitigation in the tropics. *Forest Ecology and Management*, 448. kötet, pp. 160-173.

Naqvi, S. et al., 2019. *Towards Fully Integrated Real-Time Detection Framework for Online Contents Analysis-RED-Alert Approach*. s.l., Institute of Electrical and Electronics Engineers Inc., pp. 257-263.

- Naranan, S., 1970. Bradford's Law of Bibliography of Science: an Interpretation. *Nature*, , 227(5258), pp. 631-632.
- Németh, R. et al., 2016. *Wood modification related research at the University of West Hungary*. s.l., Vienna University of Technology.
- Nepusz, T. & Csardi, G., 2023. *Fitting a power-law distribution function to discrete data*, Vienna: R Consortium.
- Nesheim, S., Malo, K. A. & Labonnote, N., 2021. Effects of interconnections between timber floor elements: dynamic and static evaluations of structural scale tests. *European Journal of Wood and Wood Products*, Volume 79, pp. 1163-1182.
- Newman, M. E. J., 2010. *Networks: An Introduction*. Oxford: Oxford University Press.
- Newman, M. E. J. & Girvan, M., 2004. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69(026113).
- Nunes, M. & Abreu, A., 2020. Applying Social Network Analysis to Identify Project Critical Success Factors.
- Nunes, M. & Abreu, A., 2020. Managing Open Innovation Project Risks Based on a Social Network Analysis Perspective.
- Özgün, K., Aklan, S. C., Tekin, A. T. & Çebi, F., 2021. Malfunction Detection on Production Line Using Machine Learning: Case Study in Wood Industry. *Advances in Intelligent Systems and Computing*, Volume 1197 AISC, pp. 1116-1124.
- Paluš, H. & Slašťanová, N., 2019. Requirements for green public procurement of wood-based products in the EU. *Digitalisation and Circular Economy: Forestry and Forestry Based Industry Implications - Proceedings of Scientific Papers*, pp. 309-314.
- Pang, J., Li, H., Tang, X. & Geng, J., 2019. Carbon dynamics and environmental controls of a hilly tea plantation in Southeast China. *Ecology and Evolution*, 9(17), pp. 9723-9735.
- Park, A. J. & Stamato, S. Z., 2021. *Social network analysis of global transshipment: a framework for discovering illegal fishing networks*. Virtual, IEEE Press, p. 776–783.
- Parsons, B., Ferrell, K. & Brashaw, B., 2021. *A practical approach to developing and sharing advanced mass timber solutions*. s.l., World Conference on Timber Engineering (WCTE).
- Pascual-Ferrá, P., Alperstein, N. & Barnett, D., 2020. Social Network Analysis of COVID-19 Public Discourse on Twitter: Implications for Risk Communication.
- Pasquale, D. D. és mtsai., 2013. *Advancements in ICT for healthcare and wellbeing: towards Horizon 2020*. New York, NY, USA, Association for Computing Machinery, p. 353–358.
- Peppoloni, S., Di Capua, G. & Haslinger, F., 2016. *ENVRI PLUS project: Developing an ethical framework for Environmental and Earth System Research Infrastructures*. Vienna, EGU - General Assembly 2016, pp. EPSC2016-6973.

- Perfiliev, P. & Zadrauskaite, N., 2019. *Research of structure units designs for woodflow on the rivers of north Russia*. s.l., International Multidisciplinary Scientific Geoconference, pp. 797-804.
- Pervin, M., Rahman, S., Parvin, S. & Akter, S., 2020. Minimum Sample Sizes for Likelihood Ratio Tests: A Simulation Study. *JnUJSci*, 7(1), pp. 63-75.
- Pinheiro, M. L., Lucas, C. & Pinho, J. C., 2015. Social network analysis as a new methodological tool to understand university-industry cooperation. *International Journal of Innovation Management*, Volume 19.
- Pinheiro, M. L., Serôdio, P., Pinho, J. C. & Lucas, C., 2016. The role of social capital towards resource sharing in collaborative R&D projects: Evidences from the 7th Framework Programme. *International Journal of Project Management*, Volume 34, pp. 1519-1536.
- Pinto, A. S., Costa, E., Guimarães, L. & Passos, R., 2022. A Digital Business Model for the Wood Industry. *Smart Innovation, Systems and Technologies*, Volume 280, pp. 623-634.
- Pirker, J., Khosmood, F. & Gütl, C., 2017. *Social network analysis of the global game jam network*. New York, NY, USA, Association for Computing Machinery, p. 10–14.
- Pollack, J. & Matouš, P., 2019. Testing the impact of targeted team building on project team communication using social network analysis.
- Polyanin, A., Pronyaeva, L., Pavlova, A. & Stepanova, Y., 2020. *Integration development processes in the wood industry based on clusterization*. s.l., IOP Publishing Ltd.
- Prescott, G. W. et al., 2017. Political transition and emergent forest-conservation issues in Myanmar. *Conservation Biology*, Volume 31, pp. 1257-1270.
- Qu, H., Lyu, J. & Chen, M., 2020. *Visual analysis of wood defect knot research based on CiteSpace*. s.l., Institute of Physics Publishing.
- Riggio, M. et al., 2022. Leveraging Structural Health Monitoring Data Through Avatars to Extend the Service Life of Mass Timber Buildings. *Frontiers in Built Environment*, Volume 8.
- Salame, C. W. et al., 2016. Use of spatial regression models in the analysis of burnings and deforestation occurrences in forest region, Amazon, Brazil. *Environmental Earth Sciences*, Volume 75, pp. 1-12.
- Salim, R. & Johansson, J., 2018. *Automation decisions in investment projects: A study in the Swedish wood products industry*. s.l., Elsevier B.V., pp. 255-262.
- Santana-Sosa, A. & Fadai, A., 2019. *A holistic approach for industrializing timber construction*. s.l., Institute of Physics Publishing.
- Santana-Sosa, A. & Riola-Parada, F., 2018. *A theoretical approach towards resource efficiency in multi-story timber buildings through BIM and lean*. s.l., World Conference on Timber Engineering (WCTE).
- Scharl, A. és mtsai., 2013. DELIVERABLE SUBMISSION SHEET.

Schramm, E. & Litschel, J., 2017. Heterogeneous Actors in a Dialog: Insights and Perceptions concerning climate-caused Risks in forestry [Heterogene Akteure im Dialog: Einsichten und Erkenntnisse zu klimabedingten Risiken in der Waldbewirtschaftung]. *Allgemeine Forst- und Jagdzeitung*, Volume 188, pp. 73-84.

Schröpfer, V., Tah, J. & Kurul, E., 2017. Mapping the knowledge flow in sustainable construction project teams using social network analysis.

Shegelman, I. R. & Vasilev, A. S., 2020. The analysis of the research areas in Russian timber industry using the knowledge base on the scientific publications of Russian researchers. *Journal of Advanced Research in Dynamical and Control Systems*, Volume 12, pp. 62-71.

Slobozhan, I., Ormosi, P. & Sharma, R., 2022. Detecting shadow lobbying. 12. kötet.

Soliman, A. és mtsai., 2016. CADIVA: cooperative and adaptive decentralized identity validation model for social networks. 6. kötet.

Sommer, J., Shandra, J. & Restivo, M., 2017. The World Bank, contradictory lending, and forests: A cross-national analysis of organized hypocrisy. *International Sociology*, 32(6), pp. 707-730.

The igraph core team, 2003-2020. *R igraph manual pages*, Vienna: R Consortium.

The World Bank, 2022. *DataBank*. [Online]

Hozzáférhető: <https://data.worldbank.org/indicator/FP.CPI.TOTL.ZG?locations=XC>
[Hozzáférés dátuma: 22. 12. 2022].

Titova, E. V., Subhonberdiev, A. S. H., Malitskaya, V. B. & Safonova, N. M., 2019. *Strategy for the sustainable development of the timber industry as subsystems of the regional economy*. s.l., International Business Information Management Association, IBIMA, pp. 9119-9125.

Tosatto, S. C. & van Zee, M., 2014. *Social network analysis for judgment aggregation*. Richland, International Foundation for Autonomous Agents and Multiagent Systems, p. 1529–1530.

Trojanowska, J., Żywicki, K., Varela, M. L. R. & Machado, J., 2016. Improving Production Flexibility in an Industrial Company by Shortening Changeover Time: A Triple Helix Collaborative Project. *Innovation, Technology and Knowledge Management*, pp. 133-146.

Tukey, J. W., 1977. *Exploratory data analysis*. Reading, Massachusetts: Addison-Wesley.

Ulyanova, S. B., Anosova, N. E. & Fisheva, A. A., 2019. *The Arctic region in the economic and geographical concept of Vladimir Dehn*. s.l., Institute of Physics Publishing.

UNECE/FAO, 2022. *Forest Products Annual Market Review, 2021-2022, Statistical Annex*. Genf: UNECE/FAO.

United Nations, 2015. *Paris Agreement*, hely nélk.: United Nations, Climate Change.

United Nations, 2019. *The Sustainable Development Goals Report 2019*, New York: United Nations Publications, Department of Economic and Social Affairs.

- Vaca, R. A. et al., 2022. Site Quality for Mahogany (*Swietenia macrophylla* King) in Natural Forests in Quintana Roo. *Journal of Sustainable Forestry*, Volume 41, pp. 33-50.
- Valeri, M. & Baggio, R., 2020. Social network analysis: organizational implications in tourism management.
- Valero-Mora, P. M., 2010. ggplot2: Elegant Graphics for Data Analysis. *Journal of Statistical Software*, , 35(1), pp. 1-3.
- Vicente, M. R., Garcíá-Muñiz, A. S. & Billón, M., 2020. Exploring European-funded project-based networks in ICT and their links with regional total factor productivity: The FP7 and the CIP. *Research Evaluation*, Volume 29, pp. 441-455.
- Vieira, C. T., da Silva Rosado, S. C. & Pereira, F. J., 2021. Morphology and viability of pollen grains of *Toona ciliata* M. Roemer (Meliaceae) at different flower stages and storage times [Morfológia e viabilidade de grãos de pólen de *Toona ciliata* M. Roemer (Meliaceae) em diferentes estádios florais e tempos de armazenamento]. *Scientia Forestalis/Forest Sciences*, Volume 48.
- Vimalajeewa, D. és mtsai., 2019. Leveraging Social Network Analysis for Characterizing Cohesion of Human-Managed Animals. 6. kötet, pp. 323-337.
- Vrabcová, P., Nikodemus, A. & Hájek, M., 2019. Utilization of forest resources and socio-economic development in Uukolonkadhi community forest of Namibia. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, Volume 67, pp. 197-206.
- Vu, T. & Parker, D. S., 2015. *Node Embeddings in Social Network Analysis*. New York, NY, USA, Association for Computing Machinery, p. 326–329.
- Wang, Y. és mtsai., 2020. Collaborative relationship discovery in BIM project delivery: A social network analysis approach.
- Wanzenböck, I., Lata, R. & Ince, D., 2020. Proposal success in Horizon 2020: A study of the influence of consortium characteristics.
- Wardikar, V. V. G., 2013. Application of Bradford's Law of Scattering to the Literature of Library & Information Science: A Study of Doctoral Theses Citations Submitted to the Universities of Maharashtra. *Library Philosophy and Practice*, Issue 12.
- Watts, D. J. & Strogatz, S. H., 1998. Collective dynamics of 'small-world' networks. *Nature*, Issue 393, p. 440–442.
- Wickham, H. & Francois, R., 2015. *A Grammar of Data Manipulation*. [Online] Hozzáférhető: <https://cran.r-project.org/web/packages/dplyr/index.html> [Hozzáférés dátuma: 19 10 2022].
- Wieruszewski, M., Trociński, A., Kawalerczyk, J. & Mydlarz, K., 2020. Logistics of Scots pine wood supply [Logistyka dostaw surowca sosnowego]. *Sylwan*, Volume 164, pp. 637-642.
- Wijffels, J., 2023. *Package 'udpipe'*, Bécs: R Foundation.

Wongnaa, C. A., Nyarko, M., Addison, M. & Awunyo-Vitor, D., 2022. Viability of artisanal timber milling: cases from Ghana. *Journal of Enterprising Communities*, Volume 16, pp. 278-298.

Xie, Y., 2016. *A General-Purpose Package for Dynamic Report Generation in R*. [Online] Hozzáférhető: <https://cran.r-project.org/web/packages/knitr/index.html> [Hozzáférés dátuma: 19 10 2022].

Yadati, N. & Narayanam, R., 2011. *Game theoretic models for social network analysis*. New York, NY, USA, Association for Computing Machinery, p. 291–292.

Yang, M., Chen, H. & Xu, Y., 2020. Stakeholder-Associated Risks and Their Interactions in PPP Projects: Social Network Analysis of a Water Purification and Sewage Treatment Project in China.

Yang, Q. és mtsai., 2019. Clustering Product Development Project Organization From the Perspective of Social Network Analysis.

Yi, J., Ryan, C. & Wang, D.-C., 2020. China's Village Tourism Committees: A Social Network Analysis.

Zaabi, H. A. & Bashir, H. A., 2020. Modeling and analyzing project interdependencies in project portfolios using an integrated social network analysis-fuzzy TOPSIS MICMAC approach.

Zhao, J., Ma, J. & Zhu, Y., 2019. Evaluating impacts of climate change on net ecosystem productivity (NEP) of global different forest types based on an individual tree-based model FORCCHN and remote sensing. *Global and Planetary Change*, 182. kötet, p. Article number 103010.

Zielinski, A., Tokarchuk, L., Middleton, S. & Chaves, F., 2013. A Framework for Achieving Situational Awareness during Crisis based on Twitter Analysis. *Geophysical Research Abstracts*, April.pp. EGU2013-6961.

10. Tézisek

A dolgozatban felállított 19 szakmai hipotézist számos statisztikai hipotézissel, kapcsolódó próbával vagy egyéb módszerrel vizsgáltuk. A szakmai hipotézisek közül 14 esetben az elemzés eredményei egyértelműen a nem elutasítás felé, öt esetben pedig inkább az elutasítás felé mutatnak. A 19 szakmai hipotézisünk alapján 14 tézist állítottunk fel.

1. tézis

A faipari projektekkel foglalkozó szakirodalom idézőinek száma hatványfüggvény-eloszlást követ egy bizonyos x_{min} érték után. Tehát a legtöbb cikk nagyon kevés, ellenben néhány cikk nagyon sok hivatkozást kap.

Kapcsolódó publikációk: [3][10][15][28][31]

2. tézis

A faipari projektekkel foglalkozó cikkek és idézők évenkénti számának alakulása között korreláció tapasztalható nemcsak az adott évi, hanem a publikációkhoz képest egy vagy két évvel későbbi citációk esetén is.

Kapcsolódó publikációk: [3][10][15][28][31]

3. tézis

A Scopusból elérhető, faipari projekteket leíró publikációk nem követnek Bradford-eloszlást, ezért valószínűleg nem tekinthetők a témában megszületett publikációk valódi szakmai reprezentációjának.

Kapcsolódó publikációk: [3][10][15][28][31]

4. tézis

A Scopusban elérhető, faipari projekteket leíró publikációkban a szerzői produktivitás gyakorisága Lotka-eloszlást követ.

Kapcsolódó publikációk: [3][10][15][28][31]

5. tézis

Az egyes projektek összköltsége és közösségi támogatása között az FP7 és a H2020 faipari projektjeiben igen szoros kapcsolat (erős Spearman-féle rangkorreláció) mutatkozik.

Kapcsolódó publikációk: [1][4][5][6][11]

6. tézis

Az FP7 és a H2020 faipari projektjeiben az országonkénti részvétel gyakorisága között összefüggés (közepesnél erősebb Spearman-féle rangkorreláció) tapasztalható.

Kapcsolódó publikációk: [2][18][22]

7. tézis

Az FP7 és a H2020 faipari projektjeiben az egyedi projektrésztvételek eloszlása a két programban és az országonkénti projektrésztvételek eloszlása a két programban ugyanolyan eloszlásból származik, tehát azonos eloszlásúnak tekinthető.

Kapcsolódó publikációk: [2][6][10][22]

8. tézis

A tevékenység, illetve az FP7 és a H2020 faipari projektjeiben való részvétel gyakorisága között (a Cramér-féle V alapján közepes erősségű) összefüggés tapasztalható.

Kapcsolódó publikációk: [2][3][10]

9. tézis

Az FP7-ben és a H2020 faipari projektjeiben az országonkénti közösségi támogatás mértéke között van összefüggés. Közepesnél erősebb Spearman-féle rangkorreláció tapasztalható, és közepesnél erősebb lineáris regresszió írható fel.

Kapcsolódó publikációk: [1][2][3][10][11][22]

10. tézis

Az FP7 és a H2020 faipari projektjeiben a projekttámogatások országonkénti eloszlása ugyanolyan eloszlásból származik.

Kapcsolódó publikációk: [1][2][3][10][22]

11. tézis

A tevékenység, illetve az FP7 és a H2020 programokban kapott támogatás közötti kapcsolat jelentős. A kategóriák támogatása közötti különbség nem az ismérv szerinti heterogenitásra, hanem a kategóriákon belüli heterogenitásra vezethető vissza.

Kapcsolódó publikációk: [1][2][3][10][22]

12. tézis

Az FP7 és a H2020 programokból kapott országonkénti támogatás és a faipari kibocsátás között összefüggés (közepesnél erősebb Spearman-féle rangkorreláció) tapasztalható.

Kapcsolódó publikációk: [1][2][3][10][22]

13. tézis

A faipari projekthálóok jelentős hálózati klaszterei hierarchikus felépítésűek.

Kapcsolódó publikációk: [2][3][6][22][24][26]

14. tézis

Az FP7 és a H2020 faipari projektjeinek eredményei a felhasznált közösségi erőforrásokhoz képest korlátozott nyitott és zárt innovációs eredményeket hoztak, és a támogatás és az eredmények között számszakilag közepesnél gyengébb az összefüggés.

Kapcsolódó publikációk: [2][3][4][5][6][13][14][16][24][26]

10.1. Saját publikációk bibliográfiai adatai

- [1] Báder, M. et al., 2023. The effect of agroforestry farming on wood quality and timber industry and its supportation by Horizon 2020. *Agroforestry Systems*, 2023 (Közölve: 2023. január 24.), pp. 1–17.
- [2] Novotni, A., Pásztory, Z. & Tóth, Z., 2022. Social Network Analysis in Wood Industry Projects. *Acta Silvatica et Lignaria Hungarica: An International Journal in Forest, Wood and Environmental Sciences*, 18(2), pp. 89–101.
- [3] Novotni, A. & Tóth, Z., 2021. Analyzing projects related to sustainable wood. In *International Conference of Economics PhD Students and Researchers in Komarno*. pp. 117–123.
- [4] Novotni, A., 2021. A Horizont 2020 által támogatott agrárerdészeti projekt. In „Termeljünk együtt a természettel! – Az agrárerdészet, mint új kitörési lehetőség”. pp. 213–219.
- [5] Novotni, A., 2021. A Horizont 2020 által támogatott agrárerdészeti projektek. In „Termeljünk együtt a természettel! – Az agrárerdészet, mint új kitörési lehetőség”. pp. 208–212.
- [6] Novotni, A. & Tóth, Z., 2021. Két projekttámogatási eloszlás következményeinek vizsgálata faipari projektekben. In *Közgazdász Doktoranduszok és Kutatók VII. Nemzetközi Téli Konferenciája Konferenciakötet*. pp. 240–248.
- [7] Novotni, A. & Tóth, Z., 2021. Két projekttámogatási eloszlás következményeinek vizsgálata faipari projektekben. In *Doktoranduszok Országos Szövetsége Közgazdaságtudományi osztályának VII. Téli Konferenciája Absztraktkötet*. pp. 39–39.
- [8] Tóth, Z. & Novotni, A., 2021. Analysis of Wood Industry Project Networks with R. In *III. International Conference of Economics PhD Students and Researchers in Komarno : Book of Abstracts*. pp. 30–30.
- [9] Tóth, Z. & Novotni, A., 2021. Students' level of informedness about smart city activities by field of study. In *International Conference of Economics PhD Students and Researchers in Komarno*. pp. 152–159.
- [10] Tóth, Z. & Novotni, A., 2021. Analysis of Wood Industry Project Networks with R. In *III. International Conference of Economics PhD Students and Researchers in Komarno*. pp. 183–189.

- [11] Tóth, Z. & Novotni, A., 2021. Demographic Reasons for High Birth Rates in Vienna in March 1946. *RussianStudies.hu*, 3(2), pp. 1–22.
- [12] Novotni, A. & Tóth, Z., 2020. Analysing projects related to sustainable wood. In *PhD Students and Researcher Conferences in Central Europe I*. pp. 18–18.
- [13] Novotni, A. & Tóth, Z., 2020. Retro on the furniture market. In *VI. International Winter Conference of Economics PhD Students and Researchers: Conference Proceedings*. pp. 181–187.
- [14] Novotni, A. & Tóth, Z., 2020. IKEA and global marketing trends. In *VI. International Winter Conference of Economics PhD Students and Researchers: Conference Proceedings*. pp. 172–180.
- [15] Novotni, A. & Tóth, Z., 2020. Innovation models as a basis of efficiency analysis. In *VI. International Winter Conference of Economics PhD Students and Researchers: Conference Proceedings*. pp. 165–171.
- [16] Novotni, A., 2020. A fenntarthatóság és az IKEA. In *IX. Interdiszciplináris Doktorandusz Konferencia 2020 [9th Interdisciplinary Doctoral Conference 2020]*. pp. 375–386.
- [17] Novotni, A., 2020. A fenntarthatóság és az IKEA. In *IX. Interdiszciplináris Doktorandusz Konferencia 2020 Absztraktkötet*. p. 360.
- [18] Novotni, A., 2020. The role of the research of wood industry waste and by-products in H2020 projects. In *Proceedings of the Miskolc IPW- IV. Sustainable raw materials international project week*. pp. 1–12.
- [19] Novotni, A. & Tóth, Z., 2020. Innovation models and efficiency analysis. In *VI. Winter Conference of Economics PhD Students and Researchers*. p. 94.
- [20] Novotni, A. & Tóth, Z., 2020. Retro on the furniture market. In *VI. Winter Conference of Economics PhD Students and Researchers*. p. 93.
- [21] Novotni, A. & Tóth, Z., 2020. IKEA and global marketing trends. In *VI. Winter Conference of Economics PhD Students and Researchers*. p. 92.
- [22] Tóth, Z. & Novotni, A., 2020. A Horizont 2020 által támogatott fenntartható faipari projektek. In *Válság és kilábalás: Innovatív megoldások Nemzetközi Tudományos Konferencia Sopron, 2020. november 5. – Konferenciakötet / Crisis and Recovery: Innovative Solutions International Scientific Conference 5 November 2020, Sopron – Conference Proceedings*. pp. 545–551.
- [23] Tóth, Z. & Novotni, A., 2020. A Horizont 2020 által támogatott fenntartható faipari projektek - Sustainable Wood Projects Supported by Horizon 2020. In *Válság és kilábalás: Innovatív megoldások: Programfüzet és előadáskivonatok = Crisis and Recovery : Innovative Solutions : Schedule and Book of Abstracts*. pp. 85–85.
- [24] Pásztor, Z. & Novotni, A., 2020. The Utilization of Tree Bark as Thermal Insulation Panels and Formaldehyde Absorber. *Geosciences and Engineering: A Publication of the University of Miskolc*, 8(12), pp. 205–216.
- [25] Tóth, Z. & Novotni, A., 2020. Students' level of informedness about smart city activities by field of study. In *PhD Students and Researcher Conferences in Central Europe I*. pp. 17–17.
- [26] Novotni, A., 2019. Info-communication Trends in Marketing Research. In *Modern gazdaság, okos fejlődés Nemzetközi Tudományos Konferencia. Sopron, 2019. november 7. –*

- Konferenciakötet / Modern Economy, Smart Development International Scientific Conference. Sopron, 7 November 2019. – Publications. pp. 641–653.
- [27] Novotni, A., 2019. Infokommunikációs trendek a marketingben - Info-communication Trends in Marketing Research. In Modern gazdaság, okos fejlődés = Modern economy, smart development. pp. 96–96.
- [28] Novotni, A., 2019. A science metric analysis of publications about sustainable wood. In III. Ring – Fenntartható nyersanyag-gazdálkodás - III. Sustainable Raw Materials Konferenciakötet - Proceedings. pp. 262–271.
- [29] Novotni, A., 2019. A közösségi marketing és az online trendek. In Tavaszi Szél 2019 Konferencia. Nemzetközi Multidiszciplináris Konferencia. pp. 378–378.
- [30] Novotni, A., 2019. Fenntarthatóság a vállalati menedzsmentben az IKEA példáján keresztül. In Tavaszi Szél 2019 Konferencia. Nemzetközi Multidiszciplináris Konferencia. pp. 377–377.
- [31] Tóth, Z. & Novotni, A., 2019. A fenntarthatósághoz kapcsolódó faipari kutatások tudományometriai vizsgálata. In Logisztika-Informatika-Menedzsment Nemzetközi Konferencia 2019. pp. 67–73.
- [32] Tóth, Z. & Novotni, A., 2019. Az olvasáskutatási eredmények felhasználása az e-learning módszertanában. In Logisztika-Informatika-Menedzsment Nemzetközi Konferencia 2019. pp. 51–58.
- [33] Pásztor, Z. et al., 2011. Mecseki Erdészeti Zrt. tájegységein lévő tölgy, cser, gyertyán faanyagok sűsűségének tudományos szintű vizsgálata, valamint a tüzelési célú átvételi hibák feltárása.
- [34] Alpár, T., Szabadhegyi, V., et al., 2010. Faanyagok modifikálása, tulajdonságainak javítása, segédanyagok fejlesztése (1. rész).
- [35] Alpár, T., Selmeczi, É., et al., 2010. Faanyagok modifikálása, tulajdonságainak javítása, segédanyagok fejlesztése.
- [36] Novotni, A., 2010. Online vállalati teljesítményértékelő eszköze lesz a klasztertagoknak. pp. 27–27.
- [37] Novotni, A., 2010. Benchmarking tanácsadó füzet 2., Sopron: NYME Erdő- és Fahasznosítási Regionális Egyetemi Tudásközpont (ERFARET).
- [38] Varga, M. et al., 2010. Energiatakarékos, technológia-rugalmas por-forgács elszívó rendszerek kifejlesztése, valamint szerszámfejlesztés az elszívási hatékonyság és az élettartam növelése céljából.
- [39] Bartha, D. et al., 2009. A Dalos-hegyen lévő átalakító üzemmódú erdőtömb állapotfelvétele, monitoringjának kialakítása és a megfelelő erdőkezelési módszerek kidolgozása.
- [40] Molnár, S. et al., 2009. Kocsánytalan tölgy, cser és bükk (álgesztes és fehér) fafajok faanyag tulajdonságainak esztétikai és fizikai javítása, modifikálása új módszerekkel.
- [41] Novotni, A., 2009. Benchmarking adatlap, Sopron: NYME Erdő- és Fahasznosítási Regionális Egyetemi Tudásközpont (ERFARET).
- [42] Novotni, A., 2009. Benchmarking tanácsadó füzet, Sopron: NYME Erdő- és Fahasznosítási Regionális Egyetemi Tudásközpont (ERFARET).
- [43] Novotni, A., 2009. Faipari online benchmarking adatbázis.

- [44] Veperdi, G. et al., 2009. A szálaló és átalakító üzemmódú faállományok lőfakészletének meghatározása koncentrikus körös mintavétellel, egyváltozós fatérfogat-függvény alkalmazásával.