

Halak fajkompozíciója és hidromorfológiai változók közötti összefüggések a Duna magyarországi szakaszán

36. OTDK Műszaki Tudományi Szekció

Szerző:

Jandó Benedek

a Veres Pálné Gimnázium 12. osztályos tanulója

Konzulensek:

Dr. Baranya Sándor

Füstös Vivien

Ermilov Alexander Anatol

Vízépítési és Vízgazdálkodási Tanszék

Budapesti Műszaki és Gazdaságtudományi Egyetem

Budapest, 2022.

1	Bevezetés	4
1.1	Magyarország vízrajza	4
1.1.1	Vizes élőhelyek Magyarországon	4
1.1.2	A Duna szerepe	4
1.1.3	Folyószabályozások	5
1.2	Magyarország halfaunája	6
1.3	Nagy időintervallumot átölelő adatsorok	7
1.4	Machine Learning („gépi tanulás”) az ökológiában	8
1.4.1	Általános alkalmazások	8
1.4.2	Random Forest általánosságban	9
2	Célkitűzés	11
2.1	Duna halfaunisztikai adatbázisának használati lehetőségei	11
2.2	Random Forest modellezés	11
3	Anyag és módszer	12
3.1	Duna halfaunisztikai adatbázisa	12
3.1.1	Általános jellemzés	12
3.1.2	Mintavételi módszerek	12
3.1.2.1	Parti elektromos mintavételezés (PE)	12
3.1.2.2	Elektromos bentikus keretes húzóháló (EBKH)	13
3.1.3	Adatbázis tisztításának módszerei	13
3.1.4	Sokféleségi együttható	14
3.1.5	Fajpárok közötti kapcsolatok elemzése	15
3.2	Random Forest modellezés	16
3.2.1	20 leggyakoribb halfaj vizsgálata	16
3.2.2	Abiotikus változók a modellben	17
3.3	Modell paraméterei	17
4	Eredmények értékelése, diszkusszió	18
4.1	RF elemzés eredményei	18
4.2	Fajpárok elemzésének eredményei	20
4.2.1	<i>Neogobius melanostomus</i> – <i>Zingel streber</i> (invazív generalista – őshonos specialista)	20
4.2.2	<i>Neogobius melanostomus</i> – <i>Romanogobio vladykovi</i> (invazív – őshonos)	21
4.2.3	<i>Neogobius melanostomus</i> – <i>Ponticola kessleri</i> (két invazív faj)	23
4.2.4	<i>Sander lucioperca</i> – <i>Perca fluviatilis</i> (két táplálékkonkurens)	25
5	Összefoglalás, következtetések	26

1 Bevezetés

1.1 Magyarország vízrajza

1.1.1 *Vizes élőhelyek Magyarországon*

Magyarország vízrajzában meghatározó elem az ország medence fekvése. A Kárpátokban lehulló csapadék különböző felszíni és felszín alatti vízfolyások formájában hazánkban köt ki. Mi sem bizonyítja ezt jobban, minthogy az ország nyilvántartott vízfolyásai közül a hazánkban eredők a kilépő teljes vízhozam mindössze 10%-át teszik ki. A többi 90%-ot mind külföldről érkező vízfolyások adják. A Magyarországra hulló csapadék nem lefolyástalan, ugyanis déli lefolyási iránnyal idővel a Dunába, majd a Fekete-tengerbe jut. Felszíni vizeinket mintegy 9800 vízfolyás alkotja. Ezek közül fontosabbak a Duna, a Tisza, a Dráva, a Maros, a Szamos, a Bodrog, a Hármas-Körös, a Sajó, a Sió, a Hernád, a Rába, az Ipoly, a Sebes-Körös és a Zagyva. Az összes vízfolyás együttes hossza 52 355 km (A Víz keretirányelv hazai megvalósítása, 2009).

A rengeteg felszíni vízfolyás önmagában is összefüggő, hatalmas, és rendkívül jelentős vizes élőhelyrendszert alkot, amely az egész Kárpát-medencét behálózza. Ezenkívül további élőhelyek kialakításában és fenntartásában játszik jelentős szerepet a felszín alatti vízhálózat. A Duna-Tisza közén például rendkívül nagy, egykor összefüggő lápvidék található (Turjánvidék), amely az Ős-Duna egykori folyása mentén helyezkedik el és jelenleg is a Dunából átszivárgó víz, illetve a bizonyos területeit a Gödöllői-dombságból felszín alatt lefolyó csapadék táplálja (Törőcsik, Sümegi, & Sümegi, 2018). A hazánkban élő fajok és társulások jelentős része kötődik vizes élőhelyekhez, hiszen ezek a területek egykor meghatározó, összefüggő részét adták az országnak az erdők és a füves puszták mellett. Mind az erdők, mind pedig a különbözővizes élőhelyek, mint a lápok, mocsarak vagy szikes tavak, hidrológiai szempontból pufferterületeknek számítanak, amelyek lassítják a víz lefolyását hazánkban, tehát nagyszerepet játszanak a vízmegtartásban.

1.1.2 *A Duna szerepe*

A Duna Európa második leghosszabb folyója, hossza 2850 km, vízhozama Budapestnél középvízi állapotban $2300 \text{ m}^3/\text{s}$. Vízgyűjtő területe $817\,000 \text{ km}^2$, amely a teljes Kárpát-medencét, így Magyarországot is magába foglalja. Ebből adódóan hazánk vízrajza Duna-központú, az ország folyóhálózatának főengelyét ez a folyó alkotja (A Víz keretirányelv hazai megvalósítása, 2009).

A Duna hatalmas szerepet játszott hazánk jelenlegi tájszerkezetének kialakításában. Tájformáló szerepe Magyarországon is kettős. Hazai szakaszának esése mindössze 26 méter, ami komoly építő tevékenységet tesz lehetővé, főleg az alsóbb szakaszokon (például Pesti hordalékkúpsíkság) (A Víz keretirányelv hazai megvalósítása, 2009).

A magyar tájtipológia rendszer jelentős részét teszik ki folyókkal kapcsolatba hozható tájtipusok, amelyek nagy része a Duna mentén is megtalálható. A hazai szakaszon fekvő kistájakra közepes, illetve magas heterogenitás jellemző, tehát a hazai tájképi inhomogenitás kialakításában is jelentős a Duna munkája (Tózsá, 1998.).

Továbbá fontos kiemelni, hogy a Duna természetesen a felszín alatti vízhálózathoz is szervezeten kapcsolódik. Például a Duna-Tisza közének elsősorban gravitációs vezérlésű felszín alatt vízrendszerében az Ős-Duna kavicsos medre fontos víztartóként funkcionál, így itt jellemzően mocsaras, lápos területek alakulhattak ki (pl. Kolon-tó). A Duna-völgyi feláramlási zónában található területeken szikes tavak (pl. Kelemen szék) jöttek létre (Mádlné Dr. Szőnyi, 2006.).

1.1.3 *Folyószabályozások*

Az árvízmentesítés komoly gazdasági kérdés már az ókor óta. A folyók periodikus kiöntései nagyban megnehezítik a jó termőképességű hordaléktalajok megművelését. A folyók szabályozásával rengeteg ártéri terület válik könnyen megművelhetővé, ezért jelentős gazdasági érdek fűződik ehhez a tevékenységhez. Hazánkban ez hatványozottan igaz, ugyanis több mint 20 000 km²-t tesznek ki az árvízzel veszélyeztetett területek, amelyeken jelenleg közel 2,5 millió ember él, így az árvízvédelem nemcsak gazdasági, hanem lakossági kérdés is (A Víz keretirányelv hazai megvalósítása, 2009).

A folyók gátak közé szorításának másik fontos motorja a hajózási feltételek javítása. A Duna 10 országon folyik keresztül, így az áruszállításban rendkívül fontos útvonalat jelent. Ráadásul a vízi transzport sokkal nagyobb mennyiségű árucikk egyszerre történő mozgatását tette lehetővé (például a reformkorban), mint a szárazföldi útvonalak.

E céloktól vezérelve indultak meg Magyarországon a folyószabályozások a XII. században, majd vettek nagy lendületet a reformkorban (XIX. sz.) Széchenyi István, Vásárhelyi Pál, Beszédes József, Huszár Mátyás, Lányi Sámuel és Vedres István vezetésével (László, 2001). Mindezek eredményeképpen manapság 4600 km földgát,

19 km árvízvédelmi fal és 338 km magaspart övezi folyóinkat (A Víz keretirányelv hazai megvalósítása, 2009).

A modern folyószabályozások drasztikus meder- és parti átalakításokkal jártak. Jelentősen átalakult Magyarország vízrajza. A nagy kiterjedésű árvizek eltűnésével számos vizes élőhely veszítette el vízutánpótlási forrásait, művelésbe bevont területek szikesedtek el és váltak terméketlenné a folyó menti feláramlásoknak köszönhetően. A hidrológiai puffterületekként szolgáló ártéri erdők visszaszorultak, így csökkent a vízmegtartás lehetősége hazánkban. Ezenkívül eltűnt az ártéri gazdálkodás Magyarország legnagyobb részéről.

A folyómedrek leszűkítésével és kimélyítésével, illetve a litorális zóna átalakításával számos élőlény veszítette el élőhelyét. A kanyarulatok levágásával nőtt a folyóvizek esése, így a sodrási sebessége is, amely a mederanyag-összetételt is megváltoztatta számos területen. A folyók gyakorlatilag homogenizálva lettek, amely a sokszínű halközösségek visszaszorulásához vezetett. A vízi élőlények mellett a száraz parti öv fajai is megsínylelték a folyószabályozásokat. Jó példa erre a nagy fülemüle (*Luscinia luscinia*), amely a 19. században még a galériaerdők nagy részében megtalálható volt, azonban ezeknek az eltűnésével a hazai állomány is összeomlott, míg végül a 20. század végére teljesen el is tűnt a hazai költő madárfaunából (Haraszthy, 2019.).

A folyók szabályozásához sorolhatók a vízjárás befolyásolására létesített duzzasztók is, amelyek ellehetlenítik számos halfaj populációs kapcsolatait és szaporodási tevékenységét, például a vizáét (*Huso huso*), amely mára a kipusztulás szélére sodródott a túlhalászás és a folyókra „lezárása” következtében (Kottelat & Freyhof, 2007.).

1.2 Magyarország halfaunája

Magyarországon (ha az utóbbi években betelepített, mára már talán önfenntartó állománnyal rendelkező fajokat is beleszámoljuk) összesen 105 halfaj fordul elő, amelyek az alábbi rendekbe tartoznak: ingolaalakúak (*Petromyzontiformes*), tokalakúak (*Acipenseriformes*), angolnaalakúak (*Anguilliformes*), heringalakúak (*Clupeiformes*), pontyalakúak (*Cypriniiformes*), harcsaalakúak (*Siluriformes*), lazacalakúak (*Salmoniformes*), csukaalakúak (*Esociformes*), tőkehalalakúak (*Gadiformes*), fogaspontyalakúak (*Cyprinodontiformes*), pikóalakúak (*Gasterosteiformes*) és sügéralakúak (*Perciformes*) (Erős & Vörös, 2017.).

A hazánkban előforduló halfajok mintegy 45%-a (47 faj) idegenhonos. Ezen fajok nagy része direkt betelepítéssel jutott Magyarország vizeibe. Tenyésztésük elsődleges okai az

élelmiszertermelés, a rekreációs (horgászati) és az esztétikai (akvarisztikai) célú felhasználás voltak. Az akvaristák által betelepített fajok általában csak termálvizekben tudnak meghonosodni, azonban többször is előfordult már, hogy az általuk betelepített fajok a magyar vizekben is megtalálták számításukat. A betelepítések mellett és sokszor azokkal egyidejűleg azonban egyéb halfajokat is behurcoltak. Ezenkívül sok fajnál nehéz eldönteni, hogy pontosan milyen hatás vagy milyen hatások összessége játszott szerepet a betelepedésben (Erős & Vörös, 2017.).

Az idegenhonos halfajok közül számos invazív és veszélyezteti az őshonos halfaunát. Jó példa erre az amurgéb (*Perccottus glenii*), amely akvaristák által jutott el Kelet-Oroszországba, majd spontán megtelepedés vagy szintén akvaristák révén 1997-ben a Tiszán is megjelent. A faj inváziója veszélyezteti a vöröskönyves, fokozottan védett lápi póc (*Umbra krameri*) állományát (Harka & M. Bănărescu, Fish fauna of the Upper Tisa, 1999).

Az idegenhonos halfajok ökológiai hatása a betelepítés előtt nem feltétlenül feltárható, azonban a legtöbb esetben a betelepítések előtt nem is végeztek ilyen irányú vizsgálatokat. A hazánkba betelepített busa fajok (*Hypophthalmichthys spp.*) állománya például számos helyen a hagyományos halászat betiltása következtében nincs szabályozva, ugyanis a horgászok nem tudják kifogni ezeket a zooplanktonnal táplálkozó fajokat. Ugyanakkor a busák táplálékkonkurensei számos őshonos halfaj ivadékanak, például a fogassüllőének (*Sander lucioperca*) (Erős & Vörös, 2017.).

1.3 Nagy időintervallumot átölelő adatsorok

A modern tudományban egyre nagyobb jelentőséggel bírnak a hosszútávú, azaz nagy időintervallumot átölelő, standard módon gyűjtött adatsorok. Ezekkel az adatbázisokkal olyan változásokat tudunk leírni és modellezni, amely emberi léptékben mérve rendkívül lassú: egy klasszikus, néhány évet feldolgozó adatsorokkal nem igazán kimutatható. Ezenkívül a hosszútávú adatsorokban tárolt rengeteg adat apróbb, finomabb összefüggések felismerésére is lehetőséget kínál.

Jó példa erre a szürke pettyesaraszoló (*Biston betularia*) esete, amelynek az iparosodás okozta légszennyezés következtében a nagyvárosok környékén elszaporodott a melanisztikus változata a szelektív predáció miatt; majd a légszennyezés csökkenésével újra

visszaállt az eredeti arány a két színváltozat között. Ezt a folyamatot nagy területet lefedő és hosszútávú monitoringgal tudták csak bizonyítani és a természetes szelekció, illetve az adaptáció egy remek példája lett (Lees & Creed, 1975).

1.4 Machine Learning („gépi tanulás”) az ökológiában

1.4.1 *Általános alkalmazások*

A mesterséges intelligencia valósággal berobbant a tudományba. Mivel rengeteg új lehetőséget hozott szinte minden területen, így nem meglepő, hogy hamar elkezdtek használni az ökológiában is. Napjainkra gyakorlatilag a mesterséges intelligencia minden megközelítését (neurális hálózatok, genetikusan algoritmusok, random forest rendszerek, deep learning algoritmusok) alkalmazták már ökológiai kutatásokban (Sylvain, Hervet, & Lecomte, 2019).

A Machine Learning (ML) algoritmusok általánosságban arra képesek, hogy mintázatokat fedezzenek fel adatsorokban. Az ökológusok számára ez, azért rendkívül előnyös, mert így lehetővé válik a bonyolult, nem lineáris adatsorok elemzése, ami gyakran előfordul komplex ökológiai kérdéseknél (Olden, Lawler, & Poff, 2008).

A ML-nek alapvetően két fajtája van. Az egyik az „ellenőrzés” nélküli tanulás, amikor nincs meghatározott kimenet, csupán exploratív módon használják az algoritmust egy adatsorokon, hogy az csoportosítsa az adatokat. Az adatok tehát külön címkézés nélkül lesznek kategorizálva (Valletta, Torney, Kings, Thornton, & Madden, 2017). A másik módszer, amikor ellenőrzött a tanítási folyamat. Ilyenkor a kimenet meghatározott és a programnak megadott jellegek alapján kell kategorizálni az adatokat, majd a felfedezett mintázatok segítségével új adatsorokat kell elemeznie (Olden, Lawler, & Poff, 2008). Ez a hagyományos ML, amelynek gyakorlatilag továbbfejlesztése a deep learning. Ebben az esetben ugyanis, nem szükséges specifikus jellegek megadása. Például, ha képfelismerésről van szó, akkor az algoritmusnak elég annyit megadni, hogy a felismerni kívánt entitás rajta van a képen és kellő mennyiségű tanító anyag segítségével az megtalálja a felismeréshez szükséges jellegeket. Ezáltal a deep learning algoritmusok használata esetén nem szükséges a vizsgált objektum mélyreható ismerete (Sylvain, Hervet, & Lecomte, 2019).

A ML sikere a klasszikus statisztikai módszerekkel szemben a rugalmasságában keresendő, ugyanis könnyen alkalmazkodik nem lineáris adatsorokhoz is és könnyedén lehet vele párhuzamosan vizsgálni számos változót. Erre jó példa Sebastien Brosse és

munkatársainak kutatása (Brosse, Guegan, Tourenq, & Lek, 1999.), amiben 15, litorális zónában élő halfaj térbeli elterjedésének és abundanciájának becslésére használtak neurális hálózatokat és többszörös lineáris regressziót. A kutatásban 8 független abiotikus változóval vizsgáltak egy függőt, az adott halfaj denzitását, majd a 306 minta alapján felállított modellek segítségével próbálták becsülni a halfajok abundanciáját és térbeli elhelyezkedését. A két módszert összehasonlítva azt tapasztalták, hogy a neurális hálózatokon alapuló elemzés megfelelőbb a populációs becslésekhez, ugyanis képes megőrizni a tapasztalt ökológiai komplexitást, szemben a lineáris regresszióval.

A neurális hálózatokat ökológiai alkalmazásban használják még továbbá, vízminőségi index becsléséhez (Hameed, és mtsai., 2017) vagy halfajok jelenlétének becsléséhez (Mastrorillo, Lek, Dauba, & Belaud, 1997.). Ezenkívül az erdőökológiában fajok elterjedésének modellezéséhez, szénciklus modellezéséhez és veszélyértékeléshez is. Ugyanakkor ez az a terület, ahol a rendelkezésre álló adatsorok limitáltsága miatt a mesterséges intelligencia további alkalmazása is korlátozott (Liu, és mtsai., 2018.).

A deep learning algoritmusokat jellemzően valamilyen objektum felismerésére használják, például kép vagy hangfelismerés, ugyanis képes rendkívül plasztikus jeleket is, mint például a madárhangok, felismerni. A deep learning használati lehetőségeinek kibővülésében ezáltal nagy szerepet játszott a képalkotás technológiai fejlődése, például a drónokra szerelt kamerák felbontásának növekedése. Erre jó példa a tengerpartokon élő vízimadarak drónok segítségével történő monitorozása, amely sokkal pontosabb egyedszám-becslést tesz lehetővé és sokkal kevesebb zavarással jár, mint a hagyományos terepbejárás (Kellenberger, Veen, Folmer, & Tuia, 2021).

1.4.2 *Random Forest általánosságban*

A Random Forest (RF) módszer a ML megközelítések egy típusa, amely remekül alkalmazható ökológiai felhasználásban is változók prediktálására (Liu, és mtsai., 2018.). Az egész módszer az ún. „döntés fákon” (decision tree (DT)) alapul. Ezek a fák egy adott változó becslését vagy klaszterezését teszik lehetővé logikai döntések mentén, több másik változó segítségével (Olden, Lawler, & Poff, 2008).

Minden DT egy kezdeti csomópontból indul, amihez két köztes csomópont kapcsolódik. A köztes csomópontokhoz logikai döntések szerint (Igaz/Hamis) további köztes csomópontok vagy végső csomópontok kapcsolódnak. Utóbbiak adják az adott eset végkimenetelét, azaz a becsült értéket. Az egész rendszer az emberi döntések

analógiájára van felépítve. Amikor döntünk, mindig egy adott helyzetből indulunk ki és a korábbi tapasztalataink szerint veszünk számításba különböző változókat, pont, mint egy DT. Az egyes csomópontok tehát az adott változó küszöbértékét tartalmazzák numerikus változók esetén vagy egyszerűen egy állítást, logikai változók esetén. A DT használatakor a kezdeti csomópontból kiindulva haladunk végig a fán, logikai döntéseket hozva, míg végül eljutunk a prediktálni kívánt változóhoz. Ha egy DT csak numerikus változókat tartalmaz akkor „regressziós fának” (regression tree (RT)), ha csak kategóriákat, akkor pedig „osztályozó fának” (classification tree (CT)) nevezzük.

Egy DT felépítésekor mindig van egy változó, amit prediktálni vagy osztályozni szeretnénk a fa segítségével. A többi rendelkezésünkre álló változót úgy építjük bele a fába, hogy végül az adatsorunk minden értékére a lehető legpontosabban meg tudjuk adni a minket érdeklő változó egy értékét. A változók aszerint kerülnek bele, hogy mennyire tisztán vagyunk képesek döntést hozni a segítségükkel. RT-k esetén a numerikus változóknak küszöbértékeket keresünk, amelyek segítségével megfogalmazunk egy állítást, például, hogy ennek a numerikus változónak az értéke kisebb mint 80. Ha ez igaz, akkor ahhoz a csomóponthoz haladunk tovább, amelyikre ez teljesül, ha hamis akkor pedig a másik irányba. A DT előnye talán a RT-k használata mentén érhető meg a legkönnyebben: gyakran fordul elő, hogy nem lineáris változókkal van dolgunk, amelyeket regressziószámítással elemezni meglehetősen nehéz és pontatlan. Ezzel szemben egy RT segítségével a prediktálásban résztvevő numerikus változót lineáris elemekre bontva klaszterezzük és hozunk létre küszöbértékeket, amelyeket egy-egy csomóponton helyezünk el. Így a csomópontokon végighaladva a független változók értékeinek segítségével tudjuk becsülni az általunk vizsgált függő változót vagy változókat.

A DT-k előnye tehát, hogy nem lineáris kapcsolatokra is könnyedén alkalmazhatók, hátrányuk, hogy egyáltalán nem rugalmasak. Jól használhatók azon az adatsoron, amin létre lettek hozva, azonban nem lehet velük interpolálni más adatbázisokat, mert nem képesek alkalmazkodni az ott tapasztalható más küszöbértékekhez vagy a kategóriák eltérő kapcsolataihoz. Itt jön képbe a Random Forest (RF), ami ötvözi a DT-k hatékonyságát és a ML flexibilitását, bár extrapolálni a betanításra használt adatsor tartományán kívülre ez sem képes.

A RF lényege, hogy egy adott tanító adatbázis alapján random választott változók és küszöbértékek segítségével létrehoz számos DT-t. Majd az adott változó(k) új adatsorok

alapján történő prediktálásához az új adatbázis minden egyes magyarázó változójának értékét lefuttatja az összes, korábban elkészített DT-n. Regressziós RF esetében a fák által megszavazott értékek átlaga lesz a prediktált érték, míg egy klasszifikációs RF azt adja eredményül, amit a legtöbb DT javasolt. A RF tehát több száz vagy akár ezer DT összessége, amelyek mindegyikét felhasználja az új becslésekhez, ezáltal képes alkalmazkodni az új adatsokaságban található, ismeretlen esetekhez is (Olden, Lawler, & Poff, 2008).

A RF-ket általában úgy alkotják meg a tanító adatbázisból, hogy annak egy bizonyos részét (általában 70-80%-át) használják csak a tanításhoz és ezekből a véletlenszerűen kiválasztott rekordokból, ismétléses mintavétellel létrehoznak egy ún. „bootstrapped” adatbázist, ami alapján készítik a fákat, mindegyikre egyet, egymástól függetlenül. Ezután az elkészült rendszert az adatbázis, eddig fel nem használt részén tesztelik, validálják. A fák tanításához azonban nem az összes független változót használjuk fel, hanem csak egy előre meghatározható számú változó készlettel végezzük a tanítást, azt pedig, hogy mely változókkal, minden fa esetében a véletlenre bizzuk.

2 Célkitűzés

2.1 Duna halfaunisztikai adatbázisának használati lehetőségei

A Duna teljes hazai szakaszát lefedő, hosszútávú adatbázis számos ökológiai kérdés megválaszolásában segíthet minket. Ezek a kérdések lehetnek az invazív halfajok őshonos halfaunára gyakorolt hatásaival, vagy az invazív fajok terjedésének modellezésével, esetleg az egyes fajok populációs változásaival kapcsolatos kérdések. Kutatásunkban ennek az adatbázisnak az általános, leíró jellegű jellemzését tűztük ki célul. Ezenkívül példákat kerestünk a fent említett kérdések megválaszolására.

2.2 Random Forest modellezés

A tudomány részekre tagolódása, diszciplínákra való felosztása valójában csak az emberi elméhez méretezést jelenti. Ettől függetlenül a különböző területek összefüggének és sokszor célszerű összefüggéseiben vizsgálni egy-egy rendszert, mint például a dunai halfajok abiotikus igényeit. Erre az összetett feladatra hatékonyan alkalmazhatók a nagy memória- és számítási kapacitású számítógépek, a fentebb tárgyalt mesterséges intelligencia megközelítései. Kutatásunk másik részében felállítottunk egy RF modellt, amely korábban más célból kifejlesztett hidrológiai modellekből származó abiotikus változók segítségével határozza meg a 20 leggyakoribb dunai halfaj előfordulását limitáló tényezőket.

3 Anyag és módszer

3.1 Duna halfaunisztikai adatbázisa

3.1.1 *Általános jellemzés*

Konzulenseim több szálon kapcsolódnak olyan biológiai, ökológiai kutatóhelyekhez, mint a Balatoni Limnológiai Kutatóintézet, illetve az ELKH Ökológiai Kutatóközpont. E két intézmény alá tartozó kollégák elmúlt évekbeli együttműködése eredményeként létrejött egy igen részletes hosszú idejű adatsor, melyben a halászati mintavételek jegyzőkönyvét gyűjtik folytatólagosan; mely adatbázist jelen kutatáshoz rendelkezésünkre bocsátottak.

Az adatbázis általános jellemzését 4 részre osztottuk: (1) megvizsgáltuk a minták időbeli eloszlását éves léptékben 2004 és 2022 között; (2) az egyes mintavételi pontokhoz folyami kilométereket rendelve folyami kilométeres lebontásban vizsgáltuk a főági minták térbeli eloszlását; (3) a fajok egyes egyedszámadatait összesítve megvizsgáltuk a minta teljes fajkészletét; (4) megvizsgáltuk az általunk, az elemzésekben használt változók értékeinek százalékos eloszlását.

Az általunk megadott mintavételi pontokat QGIS-ben egy folyami kilométer fedvényhez viszonyítva a nearest hub (points) analízis segítségével minden főági mintához folyami kilométereket tudtunk rendelni. Az egyes mintavételi pontokra páros és páratlan folyami kilométerek esetén is elvégeztük a nearest hub (points) analízist, majd a két kapott érték közül a nagyobbbat kiválasztva adtuk meg azt a folyami kilométert, amelyben a mintavétel történt.

3.1.2 *Mintavételi módszerek*

3.1.2.1 *Parti elektromos mintavételezés (PE)*

A halfaunisztikai mintavétel két módszerrel történt. Az egyik a partközeli vagy parti elektromos mintavételezés (PE), amelyet nappal és éjszaka egyaránt végeztek. Ebben az esetben a part mellett, a folyásirány sebességével, lefelé csorgó csónakból közvetlenül a parti zónából fogták a halakat. A módszer során az áramellátást leggyakrabban pulzáló egyenárammal biztosították. Az anódot egy kézi merítőháló fém kerete adta, míg a katód egy, a csónak hátulján a vízbe lógatott, réz vezeték volt. A mintavétel hossza területenként és a mintavételt megnehezítő körülmények szerint változott. Főágon leggyakrabban 500 méter, míg általánosságban 45-800 méter hosszú mintavételezés történt ezzel a módszerrel. A halak a mintavétel végeztével lettek meghatározva lemérve és szabadon engedve.

3.1.2.2 *Elektromos bentikus keretes húzóháló (EBKH)*

A másik mintavételi módszer nyílt vízben zajlott, ez az elektromos bentikus keretes húzóhálóval történő mintavétel (EBKH), amelyet csak napközben végeztek. Ebben az esetben egy halászgépet használtak a mintavételezésre (leggyakrabban a Hans-Grassl EL65 IIGI eszközt). A halászgép egy kisebb vízilabda kapu méretű fém keretes húzóhálóból állt, amelyet nagyjából a víz sodrásának sebességével húztak a folyómederben lefelé. Itt az anód a húzóháló fém kerete volt, míg a katódot egy, a háló előtt húzott réz vezeték képezte. A halak elkábításához használt elektromosságot itt is leggyakrabban pulzáló egyenáram biztosította. A mintavétel hossza ennél a módszernél is a területtől, illetve a mintázást megnehezítő körülményektől függött, de legtöbb esetben 500 méter volt, míg általánosságban 300-870 méter hosszú mintavételezéseket végeztek. A mintázás végén felhúzták a hálót és lemérték, illetve meghatározták a benne lévő halakat, majd vissza engedték őket (Szalóky, és mtsai., 2014.).

3.1.3 *Adatbázis tisztításának módszerei*

Az dunai halfaunisztikai méréseket tartalmazó adatbázis az elemzések előtt természetesen tisztításon és szűrésen ment át.

Először a pontatlan, hiányos, vagy az általunk vizsgált területen, tehát a Duna magyarországi szakaszán túlmutató adatokat töröltük ki. Hiányosnak számított az a minta, amelyből nem volt meghatározható a mintavétel helyszíne, illetve, hogy főágban vagy mellékágban zajlott-e, továbbá nem tartalmazta a mintavétel idejét, napszakát vagy módszerét. Ezt követően ellenőriztük a mintavétel kezdő és végpontjának koordinátáit és ahol előfordult, hogy a kettő között nagyobb távolság volt, mint a mintavétel lehetséges hossza ott szintén töröltük az egész adatot vagy legalább az egyik koordinátapárt. Ezután kiszelektáltuk a redundáns vagy a további elemzések szempontjából irreleváns változókat (például a vízparti növényzet becslései). Végül a megjegyzéseket feldolgozva szűrtünk ki további rekordokat az adatbázisból.

Az elemzésekhez használt koordinátapárokat az alábbiak szerint számítottuk ki: abban az esetben, amikor mindkét koordinátapár rendelkezésre állt (minták 70%-a), ott a kettő átlagaként előálló szakaszfelezőpontot számítottuk ki. Abban az esetben viszont, ahol csak a kezdő pont állt rendelkezésünkre (minták 30%-a) ott azt a koordinátapárt vettük alapul. Az elemzésekhez használt koordináták az Országos Egységes Vetületi rendszerben (EOV) lettek számolva.

3.1.4 Sokféleségi együttható

Az ökológiai kutatásokban az abundancia mellett gyakran előkerülő érték a diverzitás-index, amelynek 3 típusát különítjük el: alfa-, béta-, és gamma-diverzitást. Emellett a diverzitás, mint fogalom az ökológiai kutatások gyakran előkerülő kérdése és feltérképezése kulcsfontosságú az élő közösségek működésének megértéséhez. Az egyes fajpárok összehasonlításához létrehoztunk egy, a diverzitással valamelyest rokon, de igazából a leggyakoribb faj mennyiségi arányára és a fajszámra fókuszáló együtthatót, amely megadja, hogy mennyire sokféle az adott minta.

Ez a sokféleségi érték (M) a minta teljes egyedszámából (N), teljes fajszámából (z), a leggyakoribb faj egyedszámából (n), illetve az összes többi faj összesített egyedszámából (k) számítható ki az alábbi képlettel:

$$M = \frac{\frac{k}{N} \times \log_N z}{\frac{n}{N}} = \frac{k \times \log_N z}{n}$$

Az első egyenletben lévő törtek nevezője természetesen kiesik, de segíti az együttható megértését. A sokféleségi együttható, tehát az összes többi faj egyedszámának a teljes minta összesített egyedszámára nézett aránya szorozva a fajszám, egyedszám alapú logaritmusával és osztva a leggyakoribb faj egyedszámának összesített egyedszámra nézett arányával. (Szükségszerűen egyes alapú logaritmus egy 1-nek kell adódjon, de mivel a számláló másik tagjának értéke ekkor mindenképpen 0, így az együttható értéke is egyértelműen 0 lesz.) A képletben résztvevő értékek 0 és 1 közé esnek, azonban az együttható értéke elméletileg bármekkora lehet, viszont a gyakorlatban egyre ritkábbak a nagy értékek. Ez az ökológia centrális nullhipotézisével magyarázható, amely azt mondja ki, hogy a populációk spontán helyezkednek el a Földön és gyakorlatilag nincs köztük kölcsönhatás. Ennek ellenhipotézise gyakorlatilag az ökológia tárgya, amelynek eredményei természetesen azt bizonyítják, hogy igen is van rendszer az egyes populációk elhelyezkedésében és igenis vannak köztük kölcsönhatások. Emiatt minél nagyobb mintát veszünk, annál kisebb a valószínűsége, hogy a mintában a fajok száma közelítsen az összesített egyedszámhoz, tehát a nagy értékek rendkívül ritkák. Ezzel szemben nagyon kis mintáknál a képletből adódóan gyakran tapasztalhatunk meglepően nagy értékeket, azonban ennek az együtthatónak nem feladata a minta reprezentativitását is figyelembe venni.

A sokféleségi érték arra használható, hogy megvizsgáljuk egy mintában mennyire uralkodik egy adott faj és a minta „maradéka” hány faj között oszlik meg. Minél inkább egyenletes a minta fajok szerinti eloszlása és minél több faj van a mintában, annál

sokfélőbb. Fontos megjegyezni azonban, hogy a maximális sokféleség nem egyenlő az diverz és jól működő ökoszisztémával, ugyanis a különböző trofikus szinteken lévő fajoknak különböző mennyiségben kell képviseltetniük magukat ahhoz, hogy egy életközösség jól tudjon működni.

3.1.5 *Fajpárok közötti kapcsolatok elemzése*

A fajok közötti kapcsolatok elemzési lehetőségeire csak néhány példát választottunk, amelyben megmutatjuk, hogy milyen populációs kapcsolatok vizsgálhatók adott fajpárok között. A fajpárok közötti kapcsolatok elemzésére alapvetően két módszert használtunk. Egyrészt vizsgáltuk a sokféleségi értékeket (1), másrészt pedig azt, hogy az egyik faj milyen hatással van a másik egyedszámára, amikor mindkettő jelen van egy adott helyen (2).

(1) A *Zingel streber* – *Neogobius melanostomus* és a *Romanogobio vladykovi* – *Neogobius melanostomus* fajpárok esetén megvizsgáltuk, hogy van-e különbség a sokféleségi értékek között, amikor az egyik faj az uralkodó vagy, amikor a másik. Ehhez kiszűrtük azokat a mintákat, amikben legalább az egyik faj jelen van, majd kiválogattuk azokat, amelyekben az egyik, majd azokat, amelyekben a másik faj az uralkodó és kiesttük az ezekhez tartozó sokféleségi értékeket. Ezután a mindkét mintából random kiválasztottunk a *Zingel streber* – *Neogobius melanostomus* fajpár esetén 70-et, a *Romanogobio vladykovi* – *Neogobius melanostomus* fajpár esetén pedig 150-et. Ezek a random minták a kisebb elemszámú eredeti minta nagyjából 70%-át tartalmazzák és függetlenek. Ezt követően a random mintákból kiszámítottuk fajonként a sokféleségi értékek átlagát és Welch-próbával (R-ben *t.test* függvény) összehasonlítottuk őket, minden esetben egyoldali ellenhipotézissel. A sokféleségi értékek fent leírt tulajdonságából következik, hogy nem normális eloszlást követnek (az eloszlás minden esetben balra ferde), azonban 30-nál nagyobb elemszámú minták esetén nem normális eloszlást követő változóknak is lehet használni közelítőlegesen a Welch-próbát (Reiczigel, Harnos, & Solymosi, 2010.).

(2) A két faj egymásra gyakorolt hatását minden vizsgált fajpárnál kiszámoltuk. Itt azokat a mintákat listáztuk ki, amelyekben mindkét faj előfordul, ugyanis abban az esetben, amikor egy faj nincs jelen egy mintában, a probléma ökológiai komplexitása miatt nehéz meghatározni, hogy ez a másik faj miatt van vagy valami

teljesen más, általunk nem is mért változó miatt. Ezt követően közös szórásdiagramon ábrázoltuk az adatpárokat. Minden szórásdiagramra illesztettünk egy $f(x)=x$ referenciaegyenest, is, amely azt a hipotetikus esetet szemlélteti, amikor a két faj egyedszáma egyenes arányosságban van egymással. Az alapján, hogy a pontok milyen mintázatot követtek, illetve az egyenes melyik oldalán sűrűsödtek felállítottuk az ellenhipotézist, majd kiszámítottuk a Pearson-féle korrelációs együtthatót a két változó kapcsolatának értékeléséhez. Azért korrelációs számítással dolgoztunk, nem pedig regressziószámítással, mert a két halfaj egyedszáma minden bizonnyal szimmetrikus kapcsolatban áll egymással. A korrelációs együttható után fajonként kiszámítottuk az egyedszám-arányokat, tehát azt, hogy a két faj együttes egyedszámának mekkora részét teszi ki az egyik és mekkora részét a másik. Az így kapott párosított minták átlagát kétmintás, páros t-próbával hasonlítottuk össze egyoldali ellenhipotézissel (R-ben `t.test(paired=T)`). Mivel az egyedszám-párok minden esetben a kis értékeknél sűrűsödtek és a nagyobbak felé haladva minden irányban erősen ritkultak, így a különbségeik sem követtek normális eloszlást, azonban a sokféleségi értékekkel szemben a különbségek tapasztalati eloszlása leginkább egyenletes volt és több csúcsú, nem pedig ferde. Viszont a nagy mintaelemszámnak ($30 \leq n$) köszönhetően közelítőleg tudtuk alkalmazni a kétmintás, páros t-próbát (Reiczigel, Harnos, & Solymosi, 2010.).

3.2 Random Forest modellezés

3.2.1 20 leggyakoribb halfaj vizsgálata

A RF modellben a 20 leggyakoribb (legnagyobb egyedszámban jelenlévő) halfaj egyedszám-adatait, a minták fajszerkezetét, illetve egyedszámát használtuk fel függő változóként, valamint 4 mintavételi körülményt leíró és 5 abiotikus változót független változóként. A 20 leggyakoribb halfajt az alábbiak alkották: szélhajtó kűsz (*Alburnus alburnus*), bagolykeszeg (*Blicca bjoerkna*), paduc (*Chondrostoma nasus*), selymes durbincs (*Gymnocephalus schraester*), balin (*Leuciscus aspius*), jászkeszeg (*Leuciscus idus*), menyhal (*Lota lota*), folyami géb (*Neogobius fluviatilis*), kerekfejű géb (*Neogobius melanostomus*), csapósügér (*Perca fluviatilis*), kessler-géb (*Ponticola kessleri*), halványfoltú küllő (*Romanogobio vladkovi*), bodorka (*Rutilus rutilus*), leánykoncér (*Rutilus virgo*), fogassüllő (*Sander lucioperca*), kőüllő (*Sander volgensis*), fejes domolykó (*Squalius cephalus*), szilvaorrú keszeg (*Vimba vimba*), német bucó (*Zingel streber*), magyar bucó (*Zingel zingel*). A 4 körülményt leíró változó a dunai halfaunisztikai adatbázisból származott.

3.2.2 *Abiotikus változók a modellben*

Az 5 abiotikus változóból egy, a vízmélység, a dunai halfaunisztikai adatbázisból származik, a többi hidrológiai modellek segítségével, illetve vízügyi adatbázisokból rendeltük az adatokhoz. Vízügyi adatsorból nyert változó a vízállás és a vízjárás, amelyek a kis-közép-nagy, illetve az árad-apad-stagnál értékek valamelyikét vették fel. A sodrási sebesség (m/s) az egyes mintavételi pontokra a vízállás és a vízjárás, továbbá az adott szakaszra jellemző esés segítségével lett megbecsülve a Duna mentén elhelyezett mércék visszamenőleges mérései alapján.

A mederanyag-összetétel 2D modellből lett a mintavételi ponthoz hozzárendelve az alábbi cikkekben használt módszerek és modellek segítségével: (Baranya, és mtsai., 2018), (Füstös, Erős, & Józsa, 2021.).

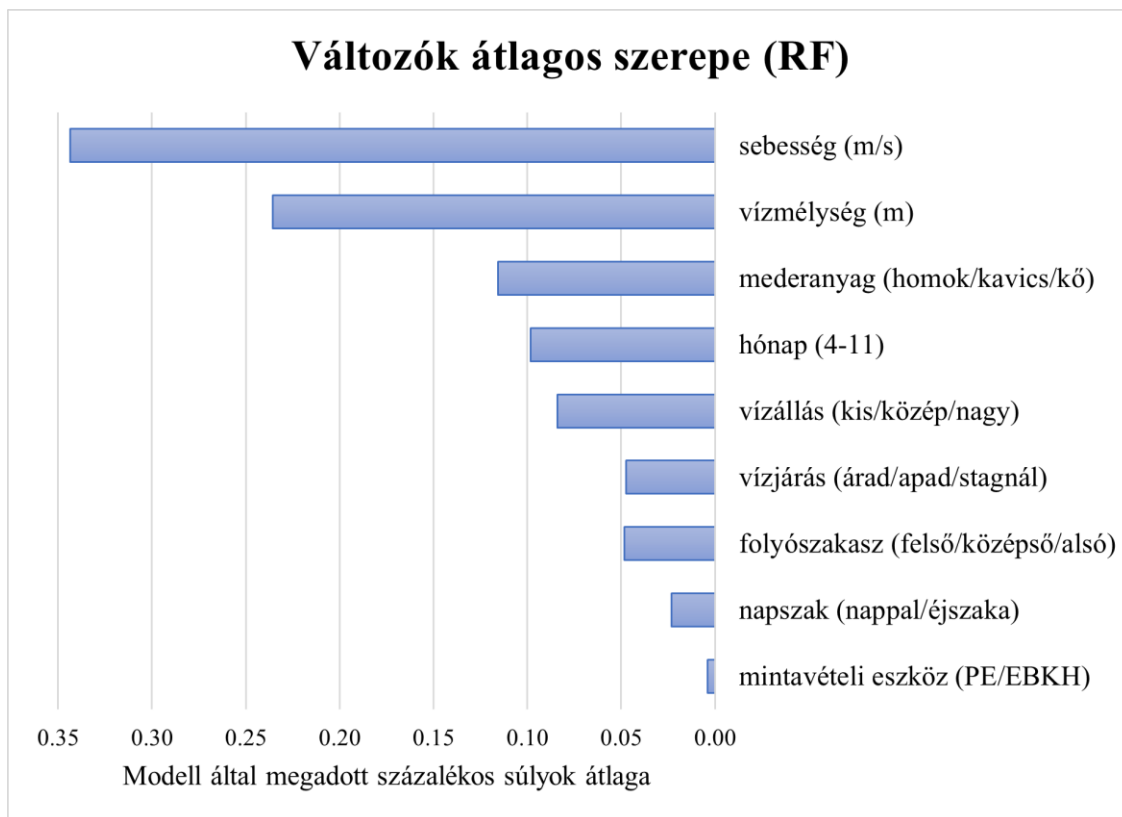
A lokális mederanyag-becslés lényege, hogy a modell egy lefektetett grid minden egyes rácspontjára a vízhozam alapján megbecsüli a nyírófeszültséget, amely alapján kikövetkeztethető a mederanyag is. Az adott területen jellemző vízhozam gyakorlatilag maximum likelihood becslési módszerrel adódik, amely alapján a kettő közötti összefüggés miatt a nyírófeszültségre is becsülhető egy érték valószínűségi úton. A korábban vett mederanyag minták és mérések alapján pedig a méret szerint csoportosított mederanyag típusok a becsült nyírófeszültséghez hozzárendelhetők minden egyes rácspontban, ami a mederanyag nagyon aprólékos leképezését adja.

3.3 Modell paraméterei

A dolgozatban egy regressziós RF modellt alkalmaztunk, melynek betanítása a rendelkezésre álló adatok 80%-ával történt meg, míg a maradék 20%-on validáltuk a betanítás sikerességét. Az RF modell 150 fát tartalmazott és maximálisan 5 szintig hatolhatott le.

4 Eredmények értékelése, diszkusszió

4.1 RF elemzés eredményei

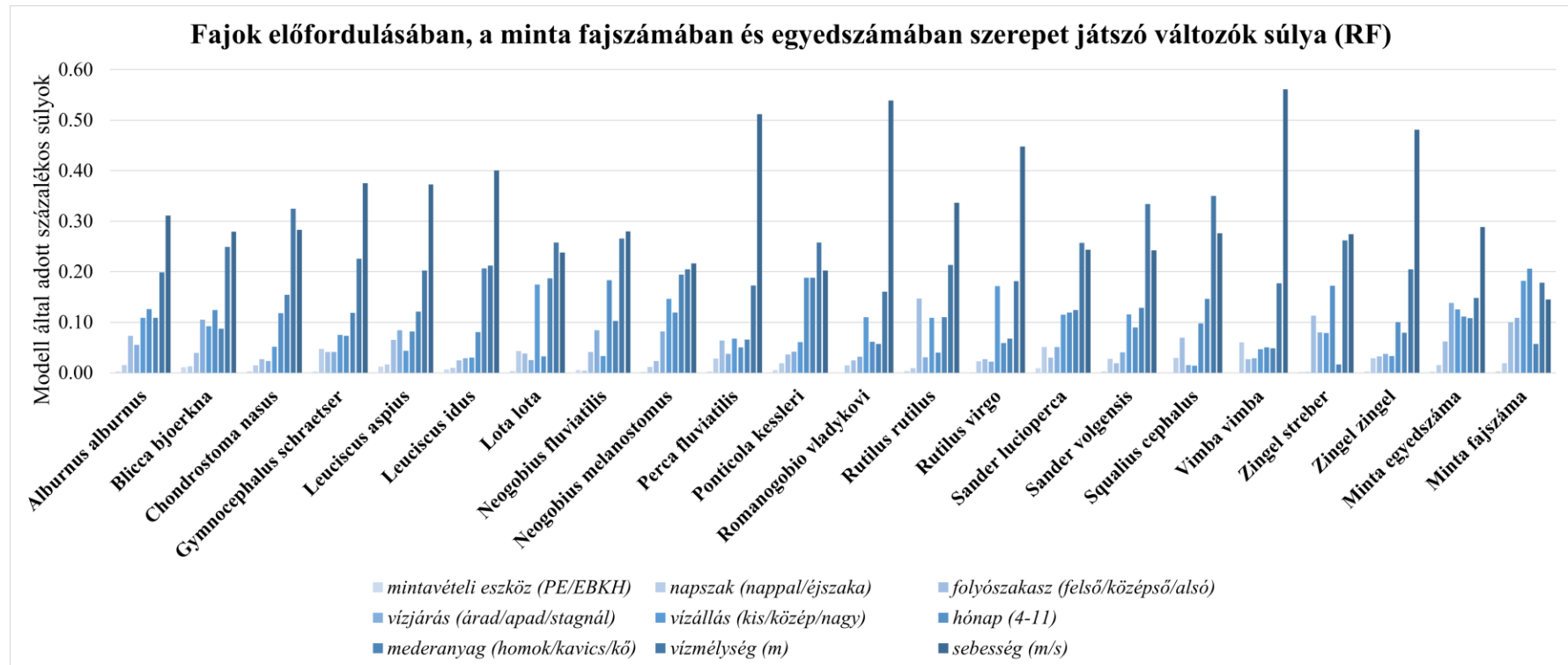


1. ábra: A kiválasztott változók átlagos szerepe a 20 leggyakoribb halfaj előfordulásában a RF modell eredményei szerint.

A RF elemzés szerint a vizsgálatra kiválasztott abiotikus változók közül kiemelt fontosságú a vízsebesség és a vízmélység; harmadik helyre pedig a mederanyag minősége került (1). Az első két paraméter ennyire kiemelt szerepe egybevág a ténnyel, hogy az élőhely- megfelelőség kutatásának kezdeti időszakában amerikai kutatók ugyanezen két paraméteren kezdték az élőhelyek minősítését (U.S. Fish and Wildlife Service, 1985), a mederanyag minőségét pedig már később vették be a halmazba (Baranya, és mtsai., 2018). A további paramétereket illetően van olyan, amit fenntartással kell kezelnünk. Saját tapasztalatainkból ismeretes, hogy a napszak és a mintavételi mód is befolyásolja a mintázott halállomány fajösszetételét, ami nem jelenik meg e paraméterek súlyánál. Fontos azonban kiemelni, hogy a vizsgálatunk függő változóiban egyedszámok és fajszámok szerepeltek, nem fajok klasszifikációja. Ezzel a megközelítéssel tehát a megváltozott összetételt nem, csak a fajszámban, egyedszámban beállt változást mutathatjuk ki (amely a modelleredmények szerint nem volt jelentős a két napszak és a két mintázási módszer között).

Az egyes fajokra külön meghatározva a 9 paraméter súlyát, szembetűnik, hogy bizonyos fajoknál az áramlási sebesség többszörösen meghatározó a többi paraméterhez képest (2).

Ismeretes, hogy a sügér (*Perca fluviatilis*) jellemzően az állóvizeket, míg a szilvaorrú keszeg (*Vimba vimba*) éppen ellenkezőleg, a nagyobb sodrású területeket kedveli (Harka & Sallai, 2004). Mindkét faj előfordulásában egyaránt fontos limitáló tényező tehát a sebesség, amelyet kimutatott a modellünk is. Egy másik megfigyelhető jelenség a kerekfejű gébhez (*Neogobius melanostomus*) kapcsolódik, melynek előfordulásában homogén módon alacsony szerepet határozott meg a RF a három legfontosabb paraméternek (2). Ez alapján adódik a következtetés, hogy a fajnak nincsenek különösebb igényei a hidrológiai, morfológiai változókat illetően. Eztalátámasztja, hogy egy invazív fajról beszélünk, amely rövid idő alatt igen nagy kiterjedésben el tudott terjedni a hazai Dunán (Harka & Sallai, Magyarország halfaunája, 2004); az inváziós siker valóban származhat ebből a fajta „igénytelenségből” is.



2. ábra: A 20 leggyakoribb faj előfordulásában szerepet játszó 9 abiotikus paraméter súlya az egyes fajokra, valamint a minta-egyedszáma, minta-fajszáma külön tekintve (RF).

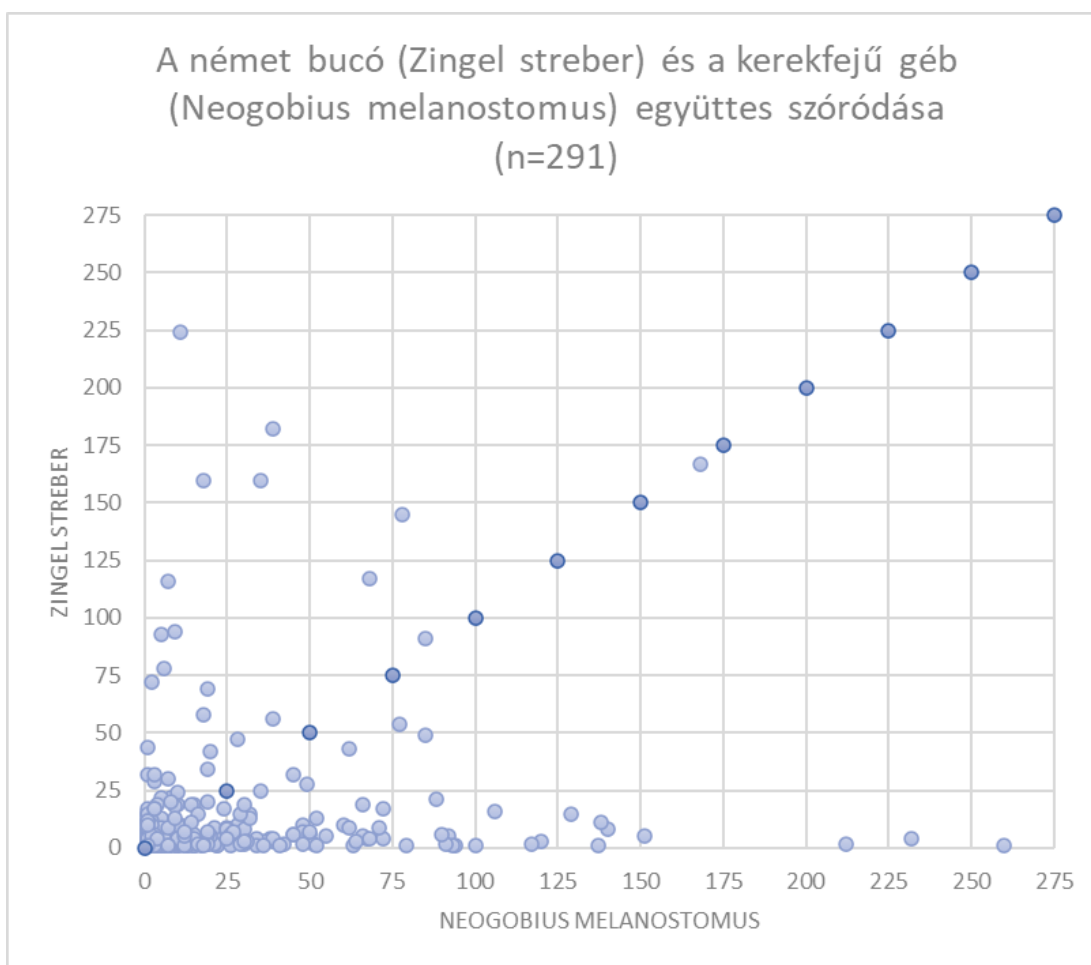
4.2 Fajpárok elemzésének eredményei

4.2.1 *Neogobius melanostomus* – *Zingel streber* (invazív generalista – őshonos specialista)

A kerekfejű géb (*Neogobius melanostomus*) rendkívül generalista faj. A túlságosan erős sodrású mély területeket nem kedveli, azonban gyakorlatilag bárhol és bármilyen aljzaton lehet vele találkozni olykor igen nagy mennyiségben. A Dunában előforduló ponto-kaszpikus gébfajok közül ez a leggyakoribb. A többi gébfajhoz hasonlóan nehéz megítélni, hogy betelepedéséhez mennyire közvetlenül járult hozzá emberi tevékenység (Erős & Vörös, 2017.). Ezzel szemben a német bucó (*Zingel streber*) a Duna vízgyűjtőjén őshonos, védett halfaj, amely elsősorban áramláskedvelő, ezáltal nagyobb egyedszámban az erősebb sodrású folyószakaszokon, a meder mélyebb részein fordul elő (Füstös, Erős, & Józsa, 2021.).

A sokféleségi érték elemzésénél a kerekfejű géb esetén az átlagos becsült sokféleség $M = 0,49$, míg a német bucónál $M = 0,36$ ($p = 0,041$ (Welch-próba)). A tapasztalt különbség magyarázható azzal, hogy a kerekfejű géb generalista faj, így sokkal több mikroélőhelyen lehet uralkodó faj (például a partmenti zónában), mint a német bucó, amely azegészen extrém, mély és nagy sodrású területeken a leggyakoribb (2) (Füstös, Erős, & Józsa, 2021.). Ezen extrém körülmények között, pedig kevesebb halfaj képes megmaradni, így nem véletlen, hogy sokféleségi érték is kisebb lett ennél a fajnál. Ettől függetlenül óvatosan lehet csak messzemenő következtetéseket levonni az eredményekből, ugyanis a szignifikancia-szint a kapott p-érték alapján elfogadható, de nem túl magas. Itt fontos megjegyezni ezzel az együttthatóval kapcsolatban, hogy átgondolt használatához elengedhetetlen az életközösségek beható ismerete, ugyanis a specialista fajok esetében kapott kisebb értékek természetesen nem azt jelentik, hogy az invazív faj jobb fajközösségekben fordul elő.

A kerekfejű géb generalizmusát támasztják alá a RF modell súlyai is, hiszen ennél a fajnál gyakorlatilag nincsen kiemelkedő, limitáló változó (2). Pont ez az a tulajdonsága a fajnak, ami miatt ilyen agresszívan terjed a Duna vízgyűjtőjében.

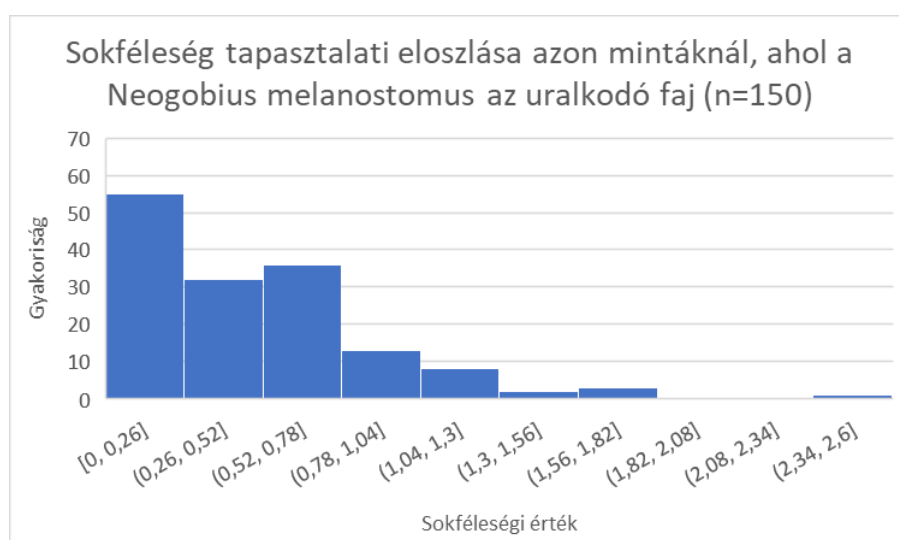
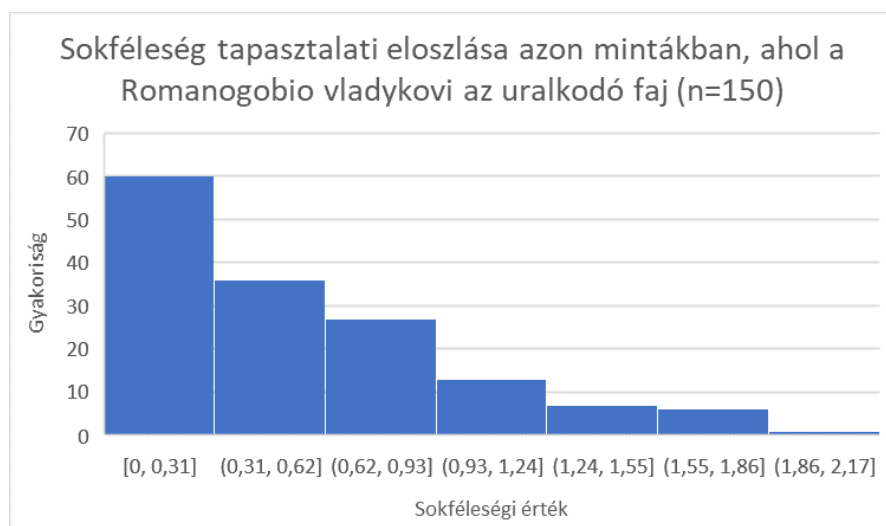


3. ábra: A német bucó (*Zingel streber*) és a kerekfejű géb együttes szórásdiagramja origón átmenő egyenessel.

A két faj mindössze 291 esetben fordult elő egy mintában. Az egyedszámok szóródási grafikonja nem mutat semmilyen határozott kapcsolatot a két változó között. Ellenben a referencia egyenes jobb oldalára, tehát a kerekfejű géb felé húznak a pontok (3), amit alátámaszt a páros t-próba eredménye is, mely szerint a rész-arány változók különbségének átlaga 0,20 ($p < 0,0001$), tehát az együttes előfordulásuk esetén átlagosan 20%-kal több kerekfejű géb van, mint német bucó. Ez komoly táplálékkonkurenciát jelenthet a két faj között és a kapott értékek aggasztóak a német bucó állományára nézve, azonban a szórásdiagramról leolvasható, hogy azokon a területeken, ahol a német bucó nagy mennyiségben fordul elő, tehát a mély és nagy sodrású helyeken, ott a kerekfejű géb csak kis egyedszámban van jelen (3), tehát a specialista faj a számára optimális élőhelyen versenytársa az agresszívan terjeszkedő invazív fajnak.

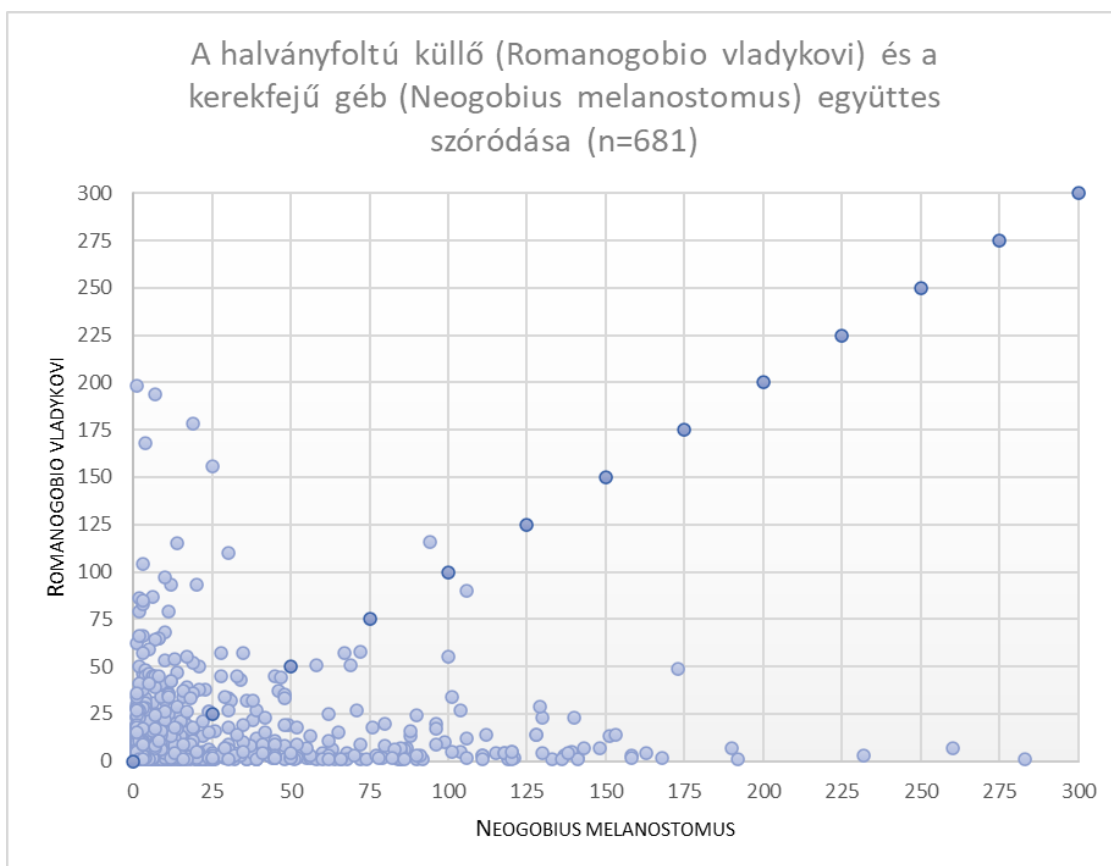
4.2.2 *Neogobius melanostomus* – *Romanogobio vladykovi* (invazív – őshonos)

Elemzéseinkben összepárosítottunk a kerekfejű gébvel egy generalistább fajt is, a halványfoltú küllőt, amely nagyon nagy egyedszámban él a Dunában. Ez a faj jóval kevésbé sodráskedvelő, mint a német bucó és a gébhez hasonlóan fenéklakó (bentikus).



4. ábra: Sokféleségi együttható tapasztalt eloszlása azon szelektált mintákban, ahol a halványfoltú küllő (*Romanogobio vladykovi*) vagy a kerekfejű géb (*Neogobius melanostomus*) az uralkodó faj.

A sokféleségi érték tekintetében a random minták hasonló tapasztalati eloszlást mutattak (4), de az őshonos fajnál átlagosan magasabb értéket feltételeztünk. A becsült átlagos értékek a sokféleségre $M = 0,52$ a halványfoltú küllőnél és $M = 0,48$ a kerekfejű gébnél, azonban a Welch-próba alapján ezek az értékek nem adódtak szignifikánsnak ($p = 0,183$).



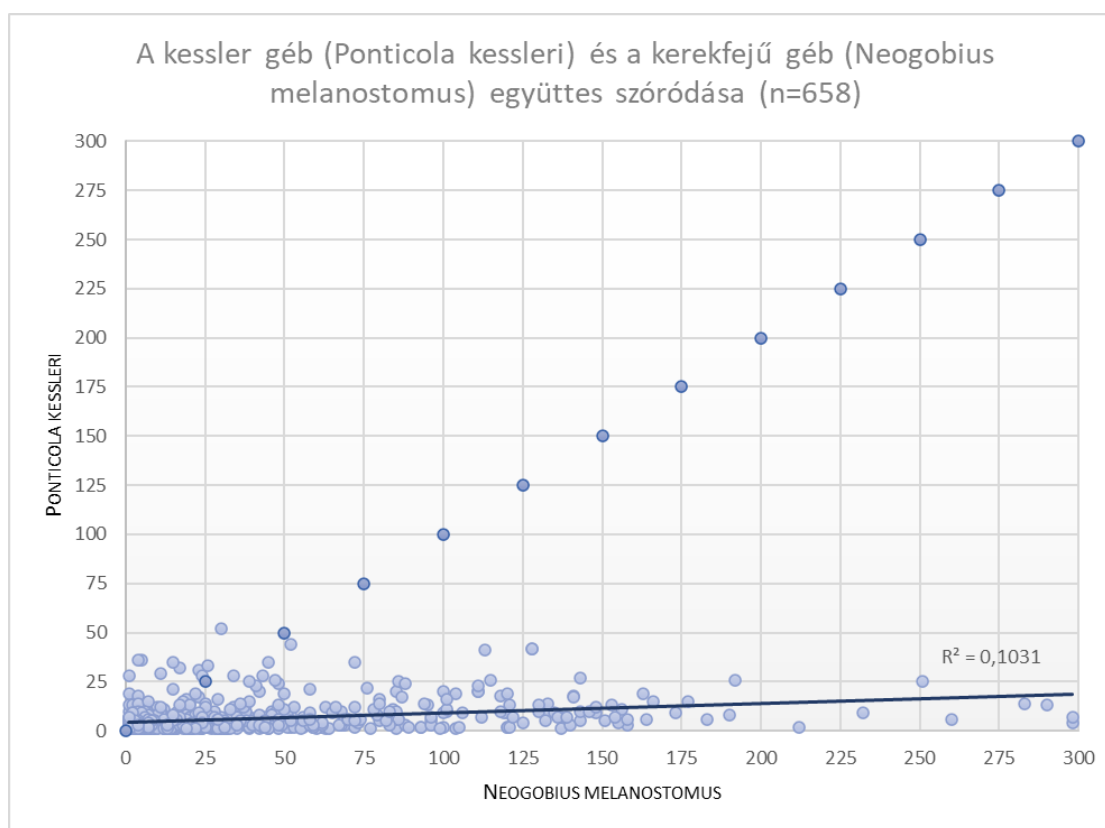
5. ábra: A halványfoltú küllő (*Romanogobio vladykovi*) és a kerekfejű géb (*Neogobius melanostomus*) együttes szórásdiagramja origón átmenő egyenessel.

A két faj kapcsolatáról a szórásdiagram alapján is elmondható, hogy az egyedszám adatok nem korrelálnak. A pontok a referencia egyenesnek inkább a jobb oldalán sűrűsödnek, de nem olyan egyértelmű a helyzet, mint a német bucó esetében (5). A rész- arány átlagok különbségének becsült értéke itt 0,14-nek adódott ($p < 0,0001$ (páros t- próba)). Ezek alapján azokban a mintákban, ahol mindkét faj előfordult átlagosan 14%-kal több a kerekfejű géb, mint a halványfoltú küllő. Ez is aggasztó adat, azonban fontos figyelembe venni, hogy ebben az elemzésben 681 rekordot használtunk fel, tehát rengeteg esetben lehetnek konkurensei ezek a fajok egymásnak azonban úgy tűnik, hogyha kis egyedszámban fordul elő az invazív faj akkor a környezethez mérten a halványfoltú küllő is képes labdába rúgni. A kapott érték értelmezésében ezenkívül érdemes megemlíteni az RF modell eredményeit (2). A halványfoltú küllőnél nagyon nagy súllyal szerepel a sodrás sebessége, mint limitáló tényező, míg a gébnél nincs igazán kiemelkedő meghatározó tényező.

4.2.3 *Neogobius melanostomus* – *Ponticola kessleri* (két invazív faj)

Az invazív fajok őshonos halállományra és fajkompozícióra gyakorolt hatásai mellett érdekes vizsgálni az invazív fajok kapcsolatát is. Esetünkben a kerekfejű géb (*Neogobius melanostomus*) és a kessler géb (*Ponticola kessleri*). Nagyjából ugyanakkor

érkezett a Duna vízgyűjtőjébe, így a hazai biotikus és abiotikus környezethez való alkalmazkodásuk gyakorlatilag egy időben történt meg. Állományaik minden bizonnyal egymásra is hatással vannak és könnyen lehetnek egymás konkurensei. Ezenkívül érdekes alanyai lehetnek azoknak a kutatásoknál, amelyek bizonyos fajok új környezethez való különböző alkalmazkodását vizsgálják. A hazai Duna szakaszon jelenleg előforduló 4 gyakori ponto-kaszpikus eredetű gébfaj közül a kessler és a kerekfejűa leggyakoribbak. Azonban az általunk vizsgált mintákban uralkodó fajként mindössze 8 esetben jelent meg a kessler géb, míg rokona több száz mintának volt a leggyakoribb faja. Ez arra az előfeltételezésre enged következtetni, hogy a kerekfejű géb volt képes a legjobban alkalmazkodni a hazai viszonyokhoz.



6. ábra: A kessler géb (*Ponticola kessleri*) és a kerekfejű géb (*Neogobius melanostomus*) együttes szórásdiagramra origón átmenő egyenessel.

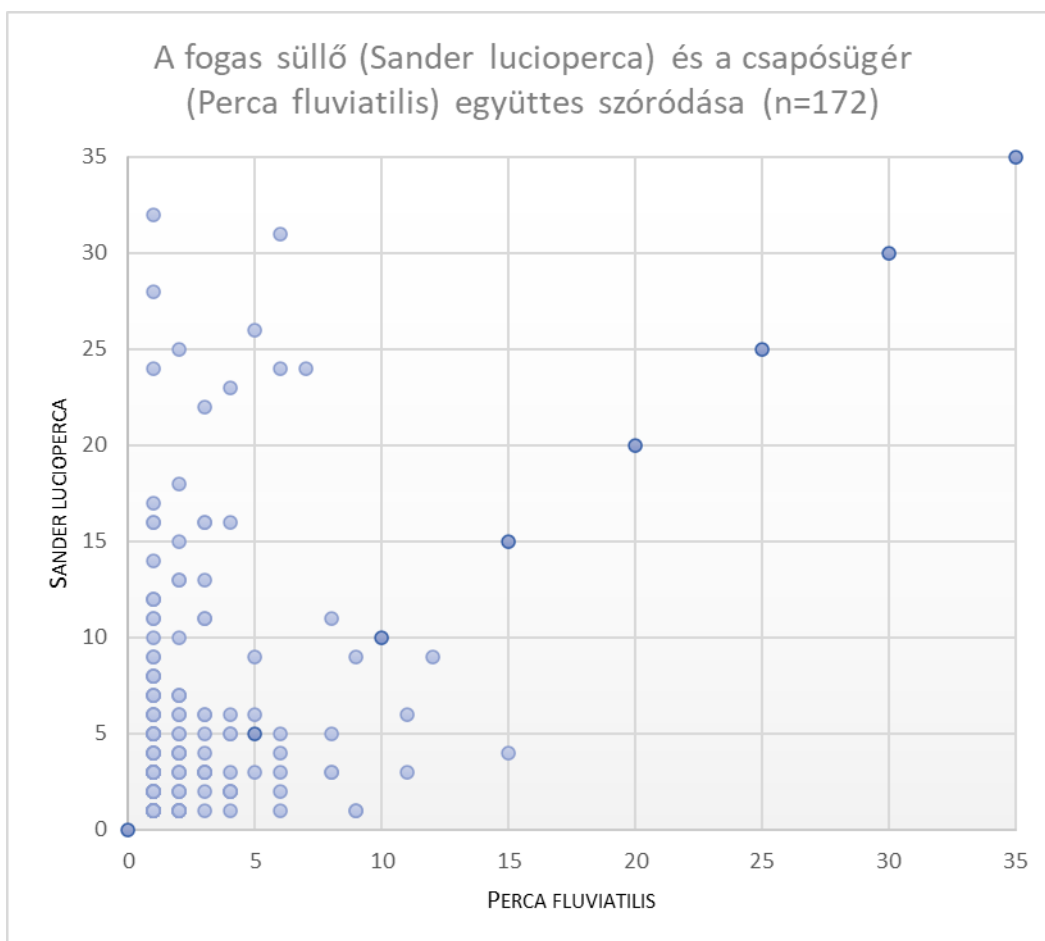
Ezt a feltételezést a kapott eredmények is alátámasztják. Az egyedszámok együttes szórásdiagramján látható, hogy a két faj egyedszáma között pozitív lineáris kapcsolat van (6) ($r = 0,32$ (pearson), $p < 0,0001$). Bár az adatok szóródása elég nagy ($R^2 = 0,10$), de úgy tűnik, hogy a kessler géb kísérfaja a kerekfejű gébnek és bizonyos fokig együtt nő az egyedszámuk is (6).

A kísérfaj jellegét alátámasztják a rész-arány változók vizsgálatának eredményei is. A kapott különbségek átlaga 0,46 ($p < 0,0001$ (páros t-próba)), tehát a két faj együttes előfordulása esetén nagyjából kétszer annyi a kerekfejű géb, mint a kessler géb. Ez

nagyon erős elnyomást jelent, ami érdekes, hiszen mindkét faj invazív, mégis nagy különbségek mutatkoznak a mennyiségükben. A RF modell alapján mindkét faj generalistának mondható, hiszen nincs kiemelkedő súlyú limitáló tényező (2), tehát elképzelhető, hogy az ennek ellenére jelentkező különbségeknek biotikus okai vannak, mégpedig az invazív gébfajok kompetíciója.

4.2.4 *Sander lucioperca* – *Perca fluviatilis* (két táplálékkonkurens)

Az invazív fajtárok mellett érdekes vizsgálni az őshonos fajokat is. A fogassüllő (*Sander lucioperca*) és a csapósügér (*Perca fluviatilis*) bizonyos korig lehetnek egymás táplálékkonkurensei, ugyanis a fiatal halak mindkét faj esetében kis halakat fogyasztanak. A Dunán folyó mintavételeknél használt módszerek, pedig pont ezt a mérettartományt tudják hatékonyan megfogni, így érdekes vizsgálni, hogy a két faj hogyan osztja fel egymás között a rendelkezésre álló élőhelyeket.



7. ábra: A fogas süllő (*Sander lucioperca*) és a csapósügér (*Perca fluviatilis*) együttes szórásdiagramja origión átmenő egyenessel.

Az egyedszámok együttes szórásdiagramján látszik, hogy a pontok szinte kizárólag a bal oldalon, azaz a fogassüllőnél sűrűsödnek (7). Ezt támasztják alá a rész-arány változókból kapott eredmények is, ugyanis a különbségek általágának becsült értéke 0,29-nek adódott ($p < 0,0001$ (páros t-próba)). A két faj együttes előfordulása esetén,

tehát átlagosan 29%-kal több a fogassüllő, mint a csapósüger. Ezenkívül a szórásdiagramról leolvasható, hogy kis egyedszámok esetén könnyen előfordulhat, hogy a két fajból ugyanannyi van (7), azonban, ahogy az eredmények is mutatják a fogassüllő uralkodóbb faj a Dunán, mint rokona. Ezt a RF modell eredményei is alátámasztják, ugyanis a csapósüger jelenlétének nagyon erős limitáló tényezője a víz sodrásának sebessége (2), tehát feltételezhető egyfajta sodrás-alapú felosztás a két faj között, amely alapján a fogassüllő a nagyobb sodrású területeken fordul elő nagy számban, míg a csapósüger inkább a lassú folyású vagy állóvizeket kedveli. Ezen hipotézis mentén érdekes lenne a két faj kapcsolatát megvizsgálni állóvizekben is.

5 Összefoglalás, következtetések

A dolgozat keretében egy hosszú időt felölelő (~15 év) mintavételi adattábla elemzését hajtottuk végre kézi és számítógépes (RF) módszerekkel. Az ilyen kiméretű, ráadásul összetett rendszerszemléletű adatelemzésre minden esetben igaz, hogy azzal elkészíteni nem, csak alkalmas időben abbahagyni lehetséges. Az eredmények, míg néhány kérdésre kielégítő választ adtak, rengeteg további lehetőséget és kérdést vetettek fel, amelyeket ebben a korlátozott időtartamban nem volt mód körüljárni.

A dolgozat egyik fő eredménye, hogy az elmúlt időszakban a hazai Duna szakaszra összegyűjtött halmintázások minden eredményét egy egységes térinformatikai adatbázisba foglaltam össze. Az adatbázis segítségével számos automatizált lekérdezés hajtható végre, ezekre a korábbiakban bemutattam példákat. Könnyen értékelhetők a mintavételi helyek, időszakok, jellemző halfajok, a mintavételi helyeken jellemző diverzitás, a jellemző egyedszámok stb. A felállított adatbázis ezen túlmenően bemeneti adatokat szolgáltatott egy Random Forest alapú adatelemző eljárásnak, amelynek segítségével a halak előfordulását, mint biotikus leírót kapcsolhattam össze az ún. hidromorfológiai paraméterekkel, a folyó áramlási sebességével, vízmélységgel, hordalékelragadó erővel és a jellemző mederanyag összetétellel, mint abiotikus leírókkal. Az abiotikus jellemzőket egy, a hazai Dunára felépített 2D számítógépes hidrodinamikai modell szolgáltatta. A dolgozat másik lényegi eredménye, hogy sikeresen alkalmaztam ezt a mesterséges intelligencia alapú módszert arra, hogy a biotikus és abiotikus jellemzők között kapcsolatot állítsak fel. Kimutattam, hogy a legtöbb halfajnál az előfordulást leginkább determináló abiotikus jellemzők a folyó áramlási sebessége és a vízmélység. Az eredmények alapján látható volt az is, hogy az őshonos halfajok előfordulásában erősebb a kapcsolat ezekkel a paraméterekkel, míg az invazív halfajoknál gyengébb, vagyis utóbbiak kevésbé érzékenyek a hidromorfológiai jellemzők megváltozására.

Az RF modell a vizsgálataim alapján megbízhatóan képes ezeket a kapcsolatokat leírni, bár a kutatás folytatásaként fontos lesz a modell megbízhatóságának számszerű jellemzése. Az RF modell lehetőséget ad, hogy a halelőfordulást determináló paraméterekre külön vizsgálatokat végezzünk, hogy kimutassuk, milyen módon függenek az adott változóktól a különböző halfajok megjelenései, tehát pl. milyen vízmélység és áramlási sebesség tartományokat preferál az adott halfaj. Az RF modell alapján továbbá lehetőség nyílik majd a halfajok előfordulásának nagyléptékű, mégis térképi kiértékelésére, mivel az eljárás a 2D hidrodinamikai modellmezőkre épül. További vizsgálatokkal tehát reményeink szerint az eddigi módszerekhez képest pontosabban vizsgálható a halelőfordulás Duna hossza menti változékonysága, a vízjárás hatása, vagy például a főág-mellékág kapcsolatok, mint az ún. keresztirányú átjárhatóságnak a szerepe.

6 Irodalomjegyzék

- A Víz keretirányelv hazai megvalósítása. (2009). Vízgyűjtő-gazdálkodási terv. Vízügyi és Környezetvédelmi Központi Igazgatóság.
- Baranya, S., Fleit, G., Józsa, J., Szalóky, Z., Tóth, B., Czeglédi, I., & Erős, T. (2018). Habitat mapping of riverine fish by means of hydromorphological tools. *Ecohydrology*.
- Brosse, S., Guegan, J.-F., Tourenq, J.-N., & Lek, S. (1999.). The use of artificial neural networks to assess fish abundance and spatial occupancy in the littoral zone of a mesotrophic lake. *Ecological Modelling*, 299-311.
- Erős, T., & Vörös, J. (2017.). Áttekintés a hazai idegenhonos inváziós halak, kételtűek és hüllők jelenlegi helyzetéről. *Magyar Tudomány*, 426-428.
- Füstös, V., Erős, T., & Józsa, J. (2021.). 2D vs. 3D Numerical Approaches for Fish Habitat Evaluation of a Large River - Is 2D Modeling Sufficient? *Periodica Polytechnica Civil Engineering*, 1114-1125.
- Hameed, M., Sharqi, S. S., Yaseen, Z. M., Afan, H. A., Hussain, A., & Elshafie, A. (2017). Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia. *Neural Computing and Applications*, 893-905.
- Haraszthy, L. (2019.). Nagy fülemüle. In L. Haraszthy, *Magyarország fészkelő madarainak költésbiológiája* (old.: 475-478.). Csákvár: Pro Vértes Nonprofit Zrt.
- Harka, Á., & M. Bănărescu, P. (1999). Fish fauna of the Upper Tisa. *Tiscia monograph series*, 439-454.
- Harka, Á., & Sallai, Z. (2004). *Magyarország halfaunája*. Szarvas: Nimfea Természetvédelmi Egyesület.
- Kellenberger, B., Veen, T., Folmer, E., & Tuia, D. (2021). Deep Learning Enhances the Detection of Breeding Birds in UAV Images. *EGU General Assembly*.
- Kottelat, M., & Freyhof, J. (2007.). *Handbook of European Freshwater Fishes*. Cornol,

Switzerland and Berlin, Germany: Maurice Kottelat and Jörg Freyhof.

László, F. (2001). *A reformkori vízimérnökök szerepe Magyarország árvízvédelmi rendszerének megalapozásában.*

Lees, D., & Creed, E. (1975). Industrial melanism in *Biston betularia*: the role of selective predation. *Journal of Animal Ecology*, 67-83.

Liu, Z., Peng, C., Work, T., Candau, J.-N., DesRochers, A., & Kneeshaw, D. (2018.). Application of machine learning methods in forest ecology: recent progress and future challenges. *Environmental Reviews*, 339-350.

Mádlné Dr. Szőnyi, J. (2006.). A Duna-Tisza köze vízföldtani típusszelvénye. *Hidrológiai tájékoztató*, 50-52.

Mastrorillo, S., Lek, S., Dauba, F., & Belaud, A. (1997.). The use of artificial neural networks to predict the presence of small-bodied fish in river. *Freshwater Biology*, 237-246.

Olden, J., Lawler, J., & Poff, N. (2008). Machine learning methods without tears: A primer for ecologists. *The Quarterly Review of Biology*, 171-193.

Reiczigel, J., Harnos, A., & Solymosi, N. (2010.). *Biostatisztika nem statisztikusoknak.* Nagykovácsi: Pars Kft.

Sylvain, C., Hervet, É., & Lecomte, N. (2019). Applications for deep learning in ecology. *Methods in Ecology and Evolution*, 1632-1644.

Szalóky, Z., György, Á. I., Tóth, B., Sevcsik, A., Specziár, A., Csányi, B., . . . Erős, T. (2014.). Application of an electrified benthic frame trawl for sampling fish in a very large European river (the Danube River) – Is offshore monitoring necessary? *Fisheries Research*, 12-19.

Tózsá, I. (1998.). Tájképi homogenitás Magyarországon. *Földrajzi Értesítő*, 432-445.

Törőcsik, T., Sümegi, B. P., & Sümegi, P. (2018). Ócsai Selyem-rét környezettörténete a jégkor végétől. *Természetvédelem és kutatás a Turjánvidék északi részén.* Budapest: Duna-Ipoly Nemzeti Park Igazgatóság.

U.S. Fish and Wildlife Service. (1985). Habitat suitability models and instream flow suitability curves: pink salmon. *Biological report*, 10.109. Forrás: <https://apps.dtic.mil/dtic/tr/fulltext/u2/a322912.pdf>

Valletta, J., Torney, C., Kings, M., Thornton, A., & Madden, J. (2017). Applications of machine learning in animal behaviour studies. *Animal Behaviour*, 203-220.