

Válasz Dr. Szeidl László bírálataira

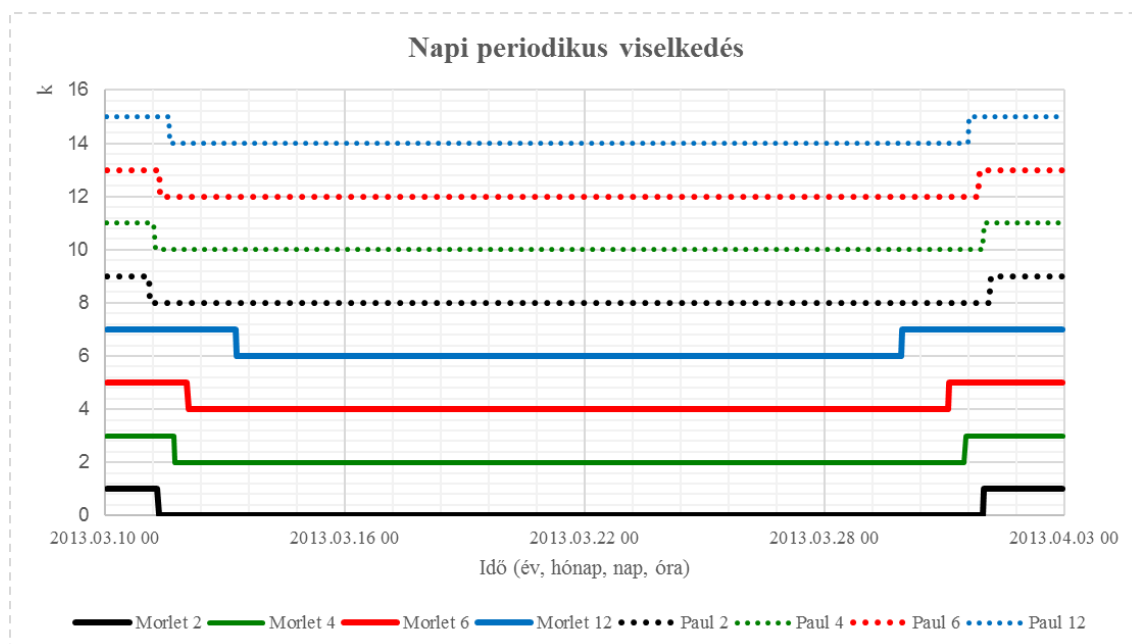
Mindenekelőtt köszönöm Dr. Szeidl Lászlónak, hogy elvállalta dolgozatom bírálatát, azt alaposan tanulmányozta és ennek kapcsán elgondolkodtató kérdéseket tett fel, melyekre ezúton kívánok válaszolni.

A bírálat első kérdése gyakorlatilag két részből áll: „A permafroszt vizsgálatánál a ψ wavelet függvény megválasztására mennyire érzékeny a spektrum becslése (42. o.)? Vannak-e olyan eredmények, amelyek a hőmérsékleti folyamatokat általánosított ARMA folyamatokkal modellezik?”

A kérdés első részének megválaszolása kapcsán mindenképpen foglalkozni kell azzal a kérdéssel, hogy miért használtam CWT-t? A wavelet transzformációk két nagy csoportjában a skála és az idő paraméter változása diszkrét, illetve folytonos. Noha csak a napi periodicitást vizsgáltam, de esetemben a periódus kimaradása időben folytonosan bármikor bekövetkezhet, ezért a folytonos wavelet transzformáció alkalmazását láttam indokoltnak. Ennek értelmezése is lényegesen egyszerűbb és közvetlenebb, mint a diszkrét transzformációé. Ez a választás jelentősen megszorítja a szóhajóvó wavelet függvények körét. A vizsgált probléma a periódus kimaradása, tehát az elsősorban zaj szűrésére használt Hermite-waveletek nem jönnek szóba. Így az ismert jelentősebb wavelet családok közül a Morlet, általánosított Morse (Lilly és Olhede, 2009), valamint a Gauss-derivált (DoG) wavelet családok egyes elemeinek választása lehet indokolt. Ezek közül a Morlet wavelet választása a legelterjedtebb. Bár így ennek választása természetesnek tűnt, de igazából a Bíráló kérdése kapcsán gondoltam bele, hogy az általam vizsgált feladat időtartományra vonatkozó kérdést vet fel. Ezzel szemben a Morlet wavelet elsősorban a frekvencia/skála meghatározásához optimális. Az irodalom az általánosított Morse waveleteket javasolja $\gamma=1$ paraméterválasztással a frekvenciakiesés detektálásához legjobban használható waveletként. Ennek elsődleges oka, hogy ez a wavelet nagyon rövid időtartományra koncentrált. Nagyon kis oszcillációs szám választással a Morlet wavelet is ilyen lenne, azonban 6-nál kisebb oszcillációs számra a Morlet wavelet nem őrzi meg az energiát, azaz nem megengedett (non-admissible). Azonban egy nagy előnye a Morlet waveletnek, hogy a skálát egyértelműen konvertálja frekvenciává, ami az általam vizsgált alkalmazásban kedvező, hiszen napi periodikus viselkedést, illetve ennek hiányát akartam vizsgálni.

A felvetett kérdés kapcsán elvégeztem egy elemzést és összehasonlítottam, hogy milyen eredményre jutnék, ha a Morlet wavelet helyett az általánosított Morse családba tartozó Paul waveletet, illetve a DoG waveletet alkalmaznám. A DoG wavelet, mint az irodalomból is ismert, nem teljesen alkalmas az adott feladatra és a kapott eredmények is pontosan ezt mutatták, mivel a napi periódust sem mutatta meg tisztán, néhány órára az rendszeresen kimaradt. Tehát a Bíráló által felvetett kérdés végül a Morlet és a Paul waveletek összehasonlítására vezetett. A kapott eredmények szerint a Paul wavelet stabilan és megbízhatóan detektálja a frekvenciakimaradás időtartamát (2013.03.11. 09 órától, 2013.03.31. 20 óráig). Ezzel szemben a Morlet wavelet növekvő oszcillációs szám mellett csökkenő időintervallumot határoz meg (2 oszcillációs szám mellett 2013.03.11. 08 órától, 2013.03.31. 23 óráig, 12 oszcillációs szám mellett 2013.03.13. 06 órától 2013.03.29. 23 óráig). Azonban, a legkisebb megengedett, azaz a 6 oszcillációs szám mellett a Paul wavelettel nagyon jól egyező eredményt kapunk, a bemutatott példában a különbség mindössze néhány óra. A periódus kimaradás elején ez 16, a végén 17 óra időkülönbséget jelent (1. ábra). Így arra a következtetésre juthatunk, hogy bár elméletileg a Paul wavelet használata az indokolt és

eredménye megbízhatóbb, mint a Morlet waveleté, az általam végzett alkalmazásban, ahol alacsony oszcillációs számmal választottam meg a Morlet waveletet, az eredmények gyakorlatilag megegyeznek. Úgy vélem, a kérdés és annak nyomán megvizsgált témakör fontos, rávilágít a helyes wavelet-megválasztás problémájára.

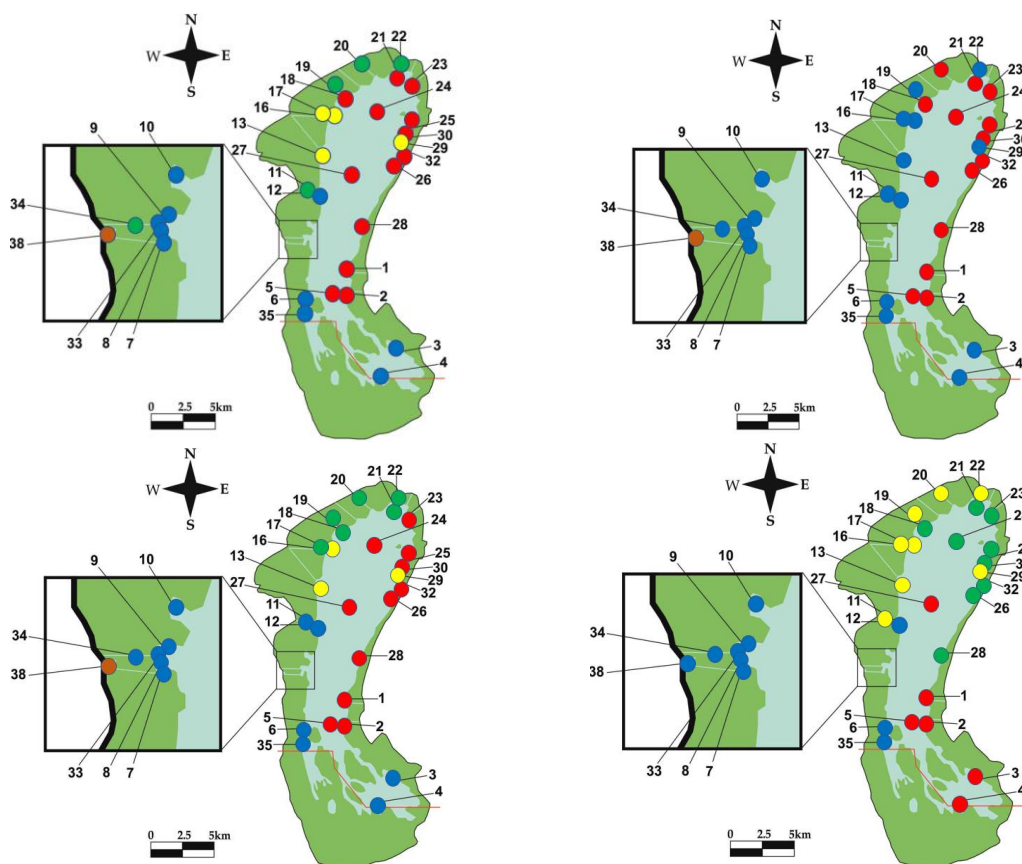


1.ábra: Morlet (folytonos vonalak) és Paul (pontosított vonalak) wavelet függvényekkel, különböző oszcillációs szám (színek) mellett végzett vizsgálatok eredményei, a Tejos megfigyelő pont -10cm mélységének hőmérséklet idősorán. A páros k értékek esetén a napi periodikus viselkedés kimarad, a páratlan $k+1$ esetén a napi periodikus viselkedés jelen van.

A kérdés második részének megválaszolása kapcsán el kell mondani, hogy hőmérséklet adatok leírására a legkorábbi modellek az autoregressziót és az Ornstein-Uhlenbeck folyamatot használták diszkrét, illetve folytonos időben. Ez természetes választás, tekintve, hogy a hővezetés egyenletét kielégítő Brown-mozgás tekinthető mindkét modell alapjának. Azonban a véletlenszerűen változó környezet perturbációkat jelent a hő terjedése során, melyet sok esetben csak nem-lineáris modellekkel lehet leírni. Az irodalomban több helyen megjelenik a GARCH típusú gerjesztés, vagy valamilyen más formában a GARCH egyenlet összekapcsolása az ARMA folyamatokkal. Egy újszerű megközelítés (Benjamin et al., 2003) az általánosított lineáris modellekkel analóg módon link-függvény segítségével kezeli a nem-linearitásokat. Bár a modell megalkotásának elsődleges célja az autoregresszió átültetése diszkrét értékű megfigyelésekre volt, de ez a modell ugyancsak jól alkalmazható bizonyos nem-linearitások kombinációjára a regresszióval. Természetesnek tűnik ezen GARMA modellek használata hőmérsékleti – így a permafrosztban mért – adatsorok elemzésére, de általam nem ismert az irodalomban ilyen megközelítés. Meteorológiai elemzésekben is csak szélsőbesség adatok elemzésére, illetve biztosítási alkalmazásban a klímaváltozás jelentette kockázat modellezésére találtam ilyen vizsgálatokat. Az alkalmazás ötlete a Bíráló részéről egy nagyon értékes felvetés, teljesen indokoltnak és megalapozottnak tűnik, további kutatások során érdemesnek tartom felhasználni.

A bíráló második kérdése: „A disszertációbeli feladatokra alkalmazott klaszterező eljárásoknál felmerült-e más, a négyzetes távolságtól eltérő távolságfogalmak használata?„ A bíráló példaként hozta a a Mahalanobis-, vagy Manhattan-távolságokat.

A válasz, hogy felmerült, azonban a négyzetes euklideszi távolság Ward algoritmussal kapott eredményei szakmailag ellentmondásmentesek voltak, publikálásra és a CCDA megírt programjába ez került. Az alábbiakban részletesen bemutatom, milyen számítási eredményekre alapul az előző megállapítás. Az R software klaszterező eljárásaiból a CCDA-ba, a bíráló nyomán a négyzetes euklideszi távolságon kívül az euklideszi, Manhattan és a Canberra távolságon alapuló számítási eljárások kerültek be, elvégezve velük a Fertő tó és a budapesti termálvizek adatainak vizsgálatát. A Fertő tó mintavételi pontjaira vonatkozó optimális csoportosítás eredményei az 2. ábrán láthatók. A dolgozatban és publikációban (Kovács et al. 2014) is megjelent négyzetes euklideszi távolsággal kapott eredményt, az öt optimális csoportot, a bal felső panel mutatja be. Az euklideszi távolsággal kapott optimális csoportszám (jobb felső panel) három, melyek közül a harmadik egyetlen mintavételi pontból áll (38-as jelű). Ez azt jelenti, ezzel a távolsággal egy nagyon leegyszerűsített kép alkotható a Fertő tó mintavételi pontjairól, holott a valóságban az ökoszisztéma ettől lényegesen színesebb, változatosabb. A Manhattan távolság használatával első ránézésre elfogadható eredményt kapunk (bal alsó panel). Néhány részlet azonban szakmailag kifogásolható. Több olyan mintavételi pontpár van, melyek közül az egyik a tó sűrű nádasában, a közlekedést lehetővé tevő csatornában, míg a másik a nyíltvízen helyezkedik el. Ilyenek a 11-12-es, a 21-22-es és a 18-19-es pontpárok. Mivel a nádasok belsejében lényegesen különböznek az ott zajló folyamatok a nyíltvíztől, ezért szakmailag jogos elvárás, hogy a nádas belsejében és a nyíltvízen elhelyezkedő mintavételi pontok a vizsgált paraméterek mért értékein keresztül, így a csoportbeosztás szerint is, különbözőségeket mutassanak. Itt említem meg, hogy a nyíltvízen elhelyezkedő 13-as, 17-es és a csatornában levő 16-os pont egy csoportba tartozik, holott elhelyezkedésük látszólag ezt nem indokolja. Az egy csoportba tartozás oka, hogy a Fertőt tápláló Wulka-patak itt folyik be, és annak vízminősége mintegy elfedi a mintavételi pontok elhelyezkedéséből adódó különbségeket. A Canberra távolság alkalmazásával sem kapunk olyan képet, ami szakmailag megfelelő lenne (jobb alsó panel). Ennek kiemelkedő példája a nyíltvízen található 27-es számú mintavételi pont, ami távol, így „elszigetelten” helyezkedik el a számítás során azonos csoportba tartozónak jelölt 1-es, 2-es és 5-as mintavételi pontoktól. Szintén jelentős problémának érzem ezt a csoportosítást a 3-as és 4-es mintavételi pontok esetében is. Ezek a magyar határ közelében helyezkednek el egy nádasokkal körbevett öbölben, ahol a Fertő tó nagy nyíltvizével csatornákon keresztül valósul meg a kapcsolat. Ennek tulajdonítható, hogy a másik három távolsággal történt csoportosítás hasonló helyzetben levő mintavételi pontokkal sorolta ezeket egy csoportba.



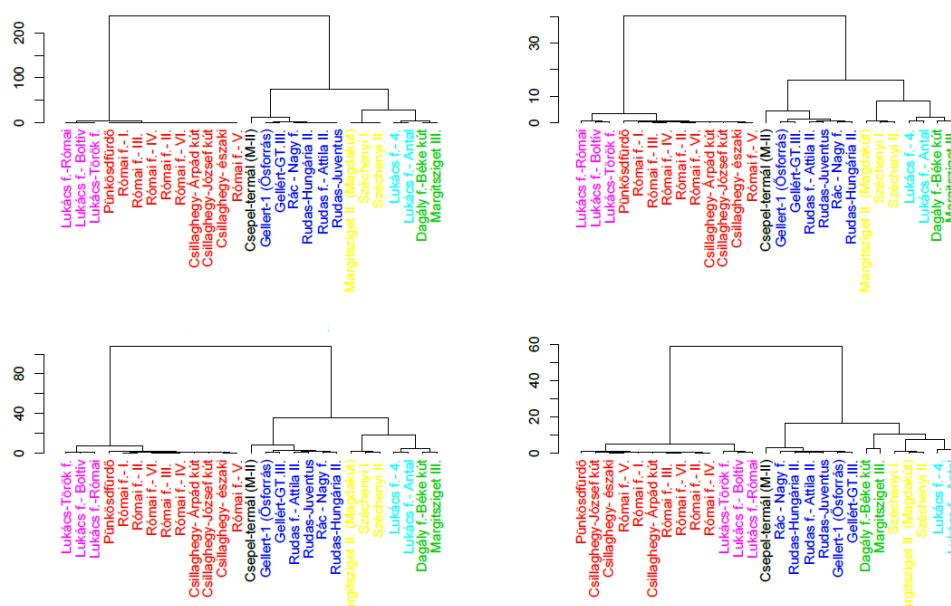
2. ábra: Mintavételi pontok optimális csoportfelosztása, Ward algoritmussal négyzetes euklideszi (bal felső panel), euklideszi (jobb felső panel), Manhattan (bal alsó panel), Canberra (jobb alsó panel) távolságokkal kapott számítások eredményeként

A homogén csoportok kialakulását befolyásolja a használt távolság és / vagy klaszterezési eljárás. Ennek legérzékletesebb példája, ha két mintavételi pont már az optimális csoportosítás meghatározásánál külön csoportba kerül, akkor azok már nem kerülhetnek azonos csoportba a teljes – a homogén csoportok meghatározását – célzó eljárásban. A következőkben tekintsük meg, milyen különbségeket eredményeztek a homogén csoportok meghatározásában a Ward eljárás alkalmazása során használt távolság fogalmak. A kapott eredményeket a négyzetes euklideszi távolság használatának eredményeivel hasonlítottam össze. A négyzetes euklideszi távolsággal 25 darab mintavételi pont került, kilenc különböző számú mintavételi pontot tartalmazó homogén csoportba. Ezeknek a homogén csoportoknak jelentős részét meg lehetett határozni, euklideszi és Canberra távolsággal 23, Manhattan távolsággal 21 mintavételi pont esetében. Amennyiben mind a 38 mintavételi pontot figyelembe vesszük, akkor a legkedvezőtlenebb esetben is azoknak kevéssel több mint 10%-a került más összetételű homogén csoportba, mint amit a négyzetes euklideszi távolság alkalmazásával kaptunk. Részleteket a többtagú homogén csoportokra vonatkozóan a 1. táblázat mutatja be.

Homogén csoportok Ward módszerrel különböző távolságokkal					
euklideszi	négyzetes euklideszi	manhattan	négyzetes euklideszi	canberra	négyzetes euklideszi
1	1	1	1	1	1
2	2	2	2	2	2
5	5	5	5	5	5
27	27	27	27	27	27
25	25	24	24	6	6
30	30	26	26	9	9
3	3	28	28	33	33
4	4	32	32	35	35
6	6	3	3	25	25
9	9	4	4	30	30
33	33	6	6	3	3
35	35	9	9	4	4
7	7	33	33	24	24
8	8	35	35	26	26
10	10	7	7	28	28
11	11	8	8	32	32
13	13	18	18	10	10
16	16	21	21	11	11
	24	23	23	7	7
26	26	25	25	8	8
28	28	30	30	18	18
32	32		13	21	21
18	18		16	23	23
21	21		10		13
	23		11		16

1.táblázat: Ward módszerrel, különböző távolságokkal meghatározott homogén csoportok a Fertő tó mintavételi pontjainak példáján. A páronként alkalmazott azonos színek, azonos csoportösszetételt jelölnek.

A budapesti termálvizek (források és kutak) adataira is elkészültek a négy távolsággal (négyzetes euklideszi, euklideszi, Manhattan, Canberra), Ward módszerével az optimális és a homogén csoportok meghatározására vonatkozó számítások. Az optimális csoportok dendrogramjait a 3. ábra mutatja be. Különbség a csoportok számában és összetételében nem látható.



3. ábra: Budapesti termálvizek (források és kutak) optimális csoportfelosztása, Ward algoritmussal négyzetes euklideszi (bal felső panel), euklideszi (jobb felső panel), Manhattan (bal alsó panel), Canberra (jobb alsó panel) távolságokkal kapott számítások eredményeként. A különböző színek egy-egy optimális csoportot jelölnek.

Homogén csoportok tekintetében alapvetően két eredmény született. A négyzetes euklideszi és a Canberra távolságokkal ugyanannyi és azonos összetételű, míg az euklideszi és a Manhattan távolságokkal az előzőekhez képest több és más összetételű homogén csoportok is létrejöttek. Az eredményeket a 2. táblázat mutatja be.

HASZNÁLT TÁVOLSÁG FOGALMAKKAL LÉTREJÖTT HOMOGEN CSOPOTOK			
négyzetes euklideszi	canberra	euklideszi	manhattan
Római f.- III.	Római f.- III.	Csillaghegy- északi	Csillaghegy- északi
Római f.- IV.	Római f.- IV.	Római f.- VI.	Csillaghegy-József kút
			Római f.- V.
			Római f.- VI.
		Csillaghegy-József kút	
		Római f.- V.	
		Római f.- I.	Római f.- I.
		Római f.- II.	Római f.- II.
		Római f.- III.	Római f.- III.
		Római f.- IV.	Római f.- IV.
Dagály f.-Béke kút	Dagály f.-Béke kút	Dagály f.-Béke kút	Dagály f.-Béke kút
Margitsziget III.	Margitsziget III.	Margitsziget III.	Margitsziget III.

2.táblázat: Ward módszerrel, különböző távolságokkal meghatározott homogén csoportok a budapesti termálkarsztos források / kutak példáján. Az azonos színek azonos csoportösszetételt jelölnek.

A bíráló harmadik kérdése így hangzik: „Vannak-e eredmények a Velencei tó vízösszetételére nézve és azok mennyire különböz(het)nek a disszertációban vizsgált tavakétól?”

A választ néhány leíró statisztika segítségével, ha nem is teljes értékűen de jelentős részben már meg lehet adni, de úgy vélem érdemes ezt a témát kicsit körbejárni és több oldalról megmutatni, nemcsak önmagában, hanem az összehasonlítást lehetővé tevő, a Fertő tó, a Balaton és a Kis-Balatonnal együtt. Továbbá mivel a CCDA a Velencei-tó esetében még nem került bemutatásra, ennek eredményeit is itt célszerűnek látom bemutatni.

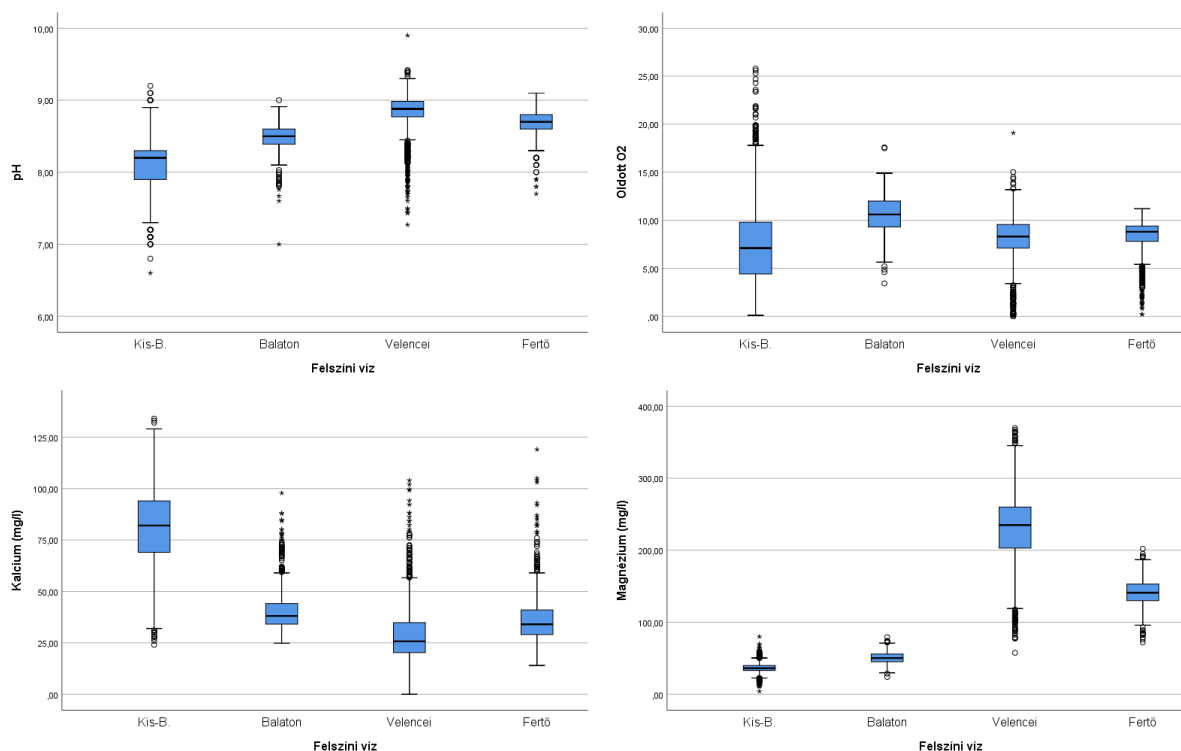
A mért paraméterek közül fontos és mindegyik vizsgált felszíni vízben mért paraméter néhány statisztikáját mutatja be a 3. táblázat. Ezek között kiemelt jelentőségű a felszíni-és a felszínalatti vizek esetében az a néhány szervetlen ion, ami jelentősebb koncentrációban van jelen, ezeket főelemeknek nevezzük. Ide tartoznak kationok: a nátrium (Na^+), a kálium (K^+), a magnézium(Mg^{2+}) és a kalcium (Ca^{2+}) és anionok: a karbonát (CO_3^{2-}), a hidrogénkarbonát (vagy bikarbonát, HCO_3^-), a szulfát (SO_4^{2-}) és a klorid (Cl^-).

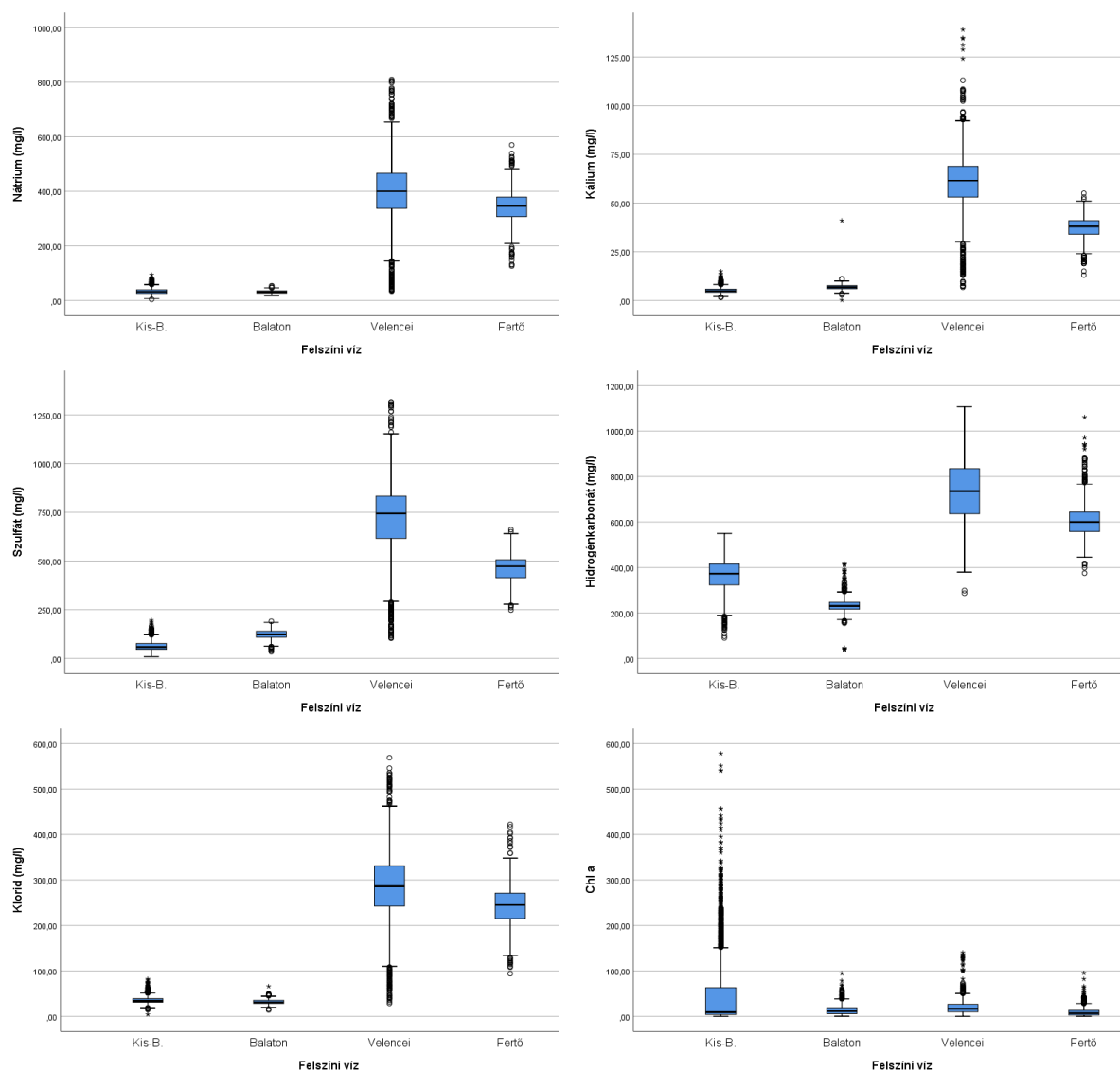
Mind a paraméterek átlagai, mind a mediánjai rámutatnak arra, hogy a Velencei-tó hasonlóan a Fertő tóhoz némely anion és kation esetében nagyon magas koncentrációval bír. A főelemek összegzett koncentrációja (ez gyakorlatilag az összes oldott anyag mennyisége) a Balaton és a Kis-Balaton esetében 500-600 mg/l, a Fertő tónál közel 2000 mg/l, a Velencei-tónál 2450 mg/l. Ez utóbbi kettő jóval meghaladja az ásványvizektől elvárt, a magyar szabványban rögzített összes oldott anyag tartalom értéket! Érdemes szólni a relatív szórásról. A főelemek közel hasonló relatív szórásúak. Az egyetlen biológiai paraméter a klorofill-a, mely messze a legmagasabb változékonysággal bír más paraméterekhez képest.

Átlag	pH	oldott O ₂	Ca	Mg	Na	K	Cl	SO ₄	HCO ₃	Chl a
Balaton	8,47	10,61	40,88	50,58	31,45	6,81	31,86	122,83	233,80	14,49
Kis-Balaton	8,11	7,36	80,31	36,35	33,43	5,09	35,57	63,09	364,65	46,90
Velencei-tó	8,80	8,03	30,19	229,50	394,84	60,20	281,00	712,58	740,62	21,00
Fertő-tó	8,68	8,36	36,40	140,95	343,55	37,44	244,01	466,14	607,86	9,56
Szórás	pH	oldott O ₂	Ca	Mg	Na	K	Cl	SO ₄	HCO ₃	Chl a
Balaton	0,18	1,95	10,47	7,65	5,87	1,71	5,13	23,69	39,62	12,51
Kis-Balaton	0,31	4,04	19,77	6,03	10,40	1,42	7,32	23,36	69,15	72,80
Velencei-tó	0,33	2,59	15,03	53,75	138,65	18,29	90,29	217,64	136,07	17,92
Fertő-tó	0,18	1,79	12,10	16,59	55,95	5,56	42,10	66,48	76,90	9,62
Rel. Szórás	pH	oldott O ₂	Ca	Mg	Na	K	Cl	SO ₄	HCO ₃	Chl a
Balaton	0,02	0,18	0,26	0,15	0,19	0,25	0,16	0,19	0,17	0,86
Kis-Balaton	0,04	0,55	0,25	0,17	0,31	0,28	0,21	0,37	0,19	1,55
Velencei-tó	0,04	0,32	0,50	0,23	0,35	0,30	0,32	0,31	0,18	0,85
Fertő-tó	0,02	0,21	0,33	0,12	0,16	0,15	0,17	0,14	0,13	1,01
Medián	pH	oldott O ₂	Ca	Mg	Na	K	Cl	SO ₄	HCO ₃	Chl a
Balaton	8,50	10,60	38,05	50,20	30,70	6,80	30,50	122,25	230,70	11,10
Kis-Balaton	8,20	7,10	82,00	36,00	32,00	4,90	34,00	58,00	373,00	9,14
Velencei-tó	8,88	8,30	25,70	235,00	400,00	61,50	286,00	744,00	735,90	16,60
Fertő-tó	8,70	8,80	34,00	141,00	347,00	38,00	245,00	473,00	600,00	7,10

3. táblázat: válogatás a Balaton, Kis-Balaton, Velencei-tó és a Fertő tó néhány paraméterének leíró statisztikáiból

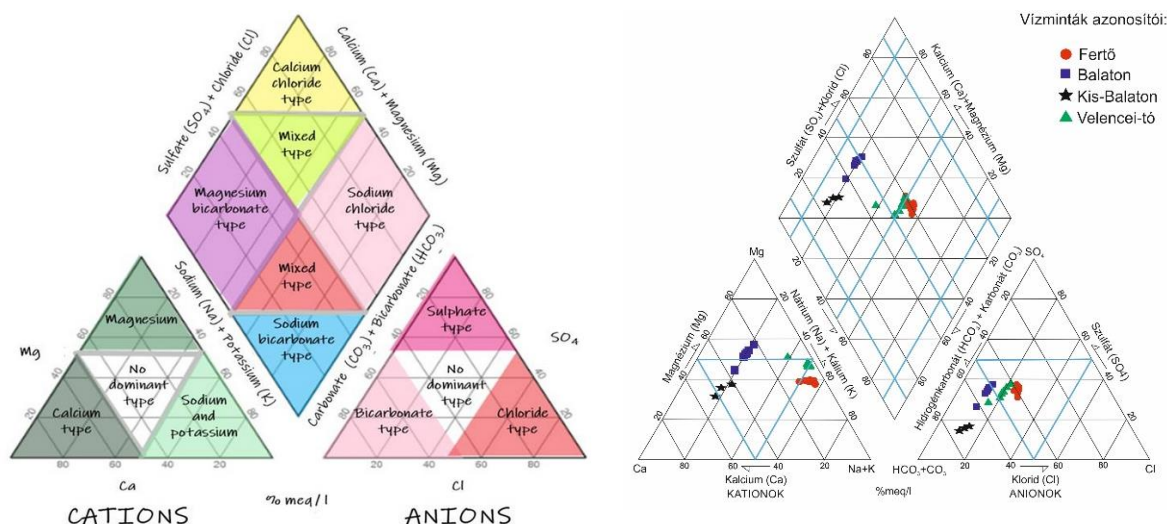
A táblázatos forma mellett a box-whisker's ábrák (4. ábra) szemléletesen mutatják be az egyes felszíni vizek közötti különbségeket. A Velencei-tó szempontjából szembetűnő a pH magas értéke. Jól látható, hogy felszíni vizeink közül a magnézium, nátrium, kálium, szulfát és hidrogénkarbonát mért értékei a Fertő és a Velencei tóban nagyon magasak.





4.ábra: Box-whisker's ábrák a Kis-Balaton, Balaton, Velencei-tó és a Fertő tó mért paramétereiről

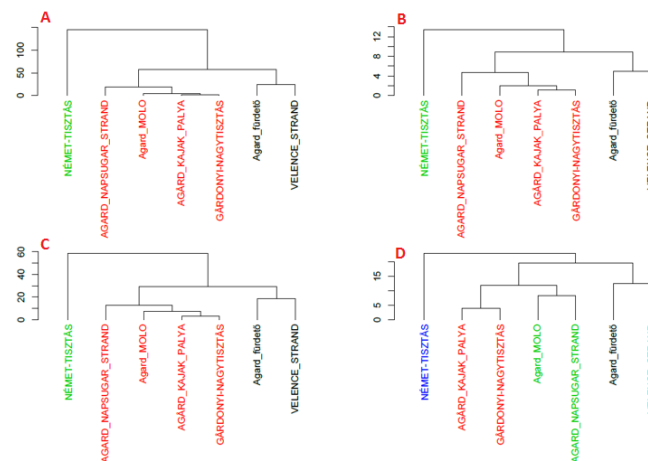
A vízkémiai elemzések eredményeit gyakran meq/l-ben adják meg, mellyel meghatározható a víz vízkémiai fáciese. Ez egy mindenki számára érthető osztályozás, melyet jól lehet szemléltetni az ún. Piper-diagramon. Attól függően, mely paramétereknek jelentős a szerepük a vizeket különböző típusokba sorolhatjuk. Ezeket az összetétel függvényében a 5. ábra bal panelje mutat be. A vizsgálat tárgyát képező felszíni vizekre jellemző pontok az 5. ábra jobb paneljén láthatók. A három zöld pont a Zala-folyóban levő („wetland része”), a pirosak a Balaton tíz mintavételi pontját jellemzik a főelemek alapján. Szembetűnő, hogy a mintavételi pontok egymás utáni sorrendje szerint a Kis-Balaton (ami a Balaton vízének utánpótlását biztosítja) kalcium-hidrogénkarbonátos vize, a Balaton észak-keleti végében már magnézium-hidrogénkarbonátos vízként van jelen. A Fertő-tó 38-as mintavételi pontjának többsége (fekete pontok) nátrium-kálium kation fáciesű, kisebb része nem domináns, míg anionok szempontjából nem domináns típusba tartoznak. A Velencei-tó ugyanakkor mind anionok, mind kationok szerint nem domináns típust mutat.



5. ábra: Vízgeokémiai fáciesek a Piper diagramon (bal panel, <https://www.hatarilabs.com> alapján), a Kis-Balaton (zöld), Balaton (piros), Fertő tó (fekete) és a Velencei-tó (kék) mintavételi pontjai vannak a jobb panelen

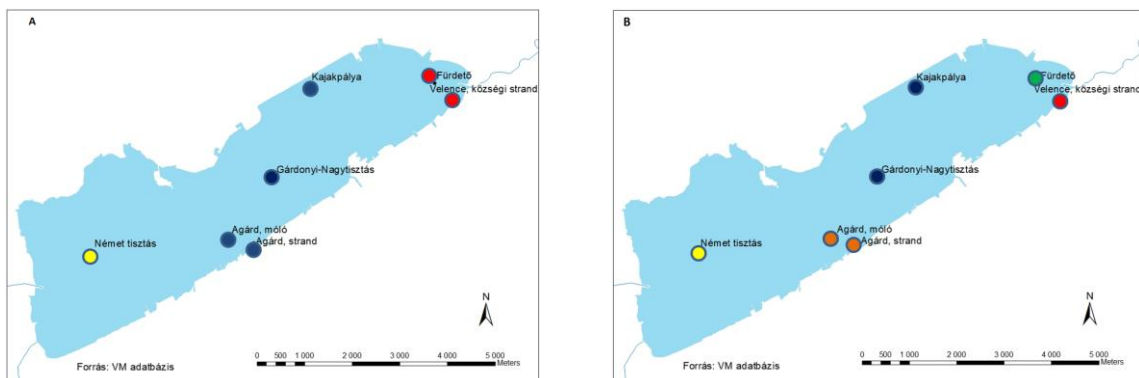
A CCDA-val megvizsgáltam, hogy a Velencei-tó hét mintavételi pontja hány optimális és hány homogén csoportot alkot. A számításokat 1991-2004 között a következő paraméterekkel végeztem: vízhőmérséklet ($^{\circ}\text{C}$), pH, vezetőképesség (μScm^{-1}), oldott oxigén (mg l^{-1}), biológiai oxigén igény (mg l^{-1}), kémiai oxigén igény (mg l^{-1}) $\text{Na}^{+}(\text{mg l}^{-1})$, $\text{K}^{+}(\text{mg l}^{-1})$, $\text{Mg}^{2+}(\text{mg l}^{-1})$, $\text{Ca}^{2+}(\text{mg l}^{-1})$, $\text{CO}_3^{2-}(\text{mg l}^{-1})$, $\text{HCO}_3^{-}(\text{mg l}^{-1})$, $\text{SO}_4^{2-}(\text{mg l}^{-1})$, $\text{Cl}^{-}(\text{mg l}^{-1})$, $\text{NH}_4\text{-N}$ (mg l^{-1}), $\text{NO}_2\text{-N}$ (mg l^{-1}), összes nitrogén (mg l^{-1}) foszfát foszfor (mg l^{-1}), összes foszfor (mg l^{-1}), chl a ($\mu\text{g l}^{-1}$).

Ward módszeréhez négyzetes euklideszi, euklideszi, Manhattan és Canberra távolságokat használtam. A kapott dendrogramokat az optimális csoportokkal (melyeket azonos színnel jelöltem) a 6. ábra mutatja be. A négyzetes euklideszi, euklideszi, Manhattan távolságokkal kapott eredmények egymástól nem különböznek. Canberra távolsággal végzett számítások eredményei azonban jelentős különbségeket mutatnak, több homogén csoport jött létre. Két egytagú csoportra vált szét a Fürdető és Velence strand, és a négy mintavételi pontot magában foglaló, a tó nagy területét lefedő csoport két kettő tagú csoportja vált szét (Agárd móló – Agárd strand és Nagytisztás – Kajakpálya)



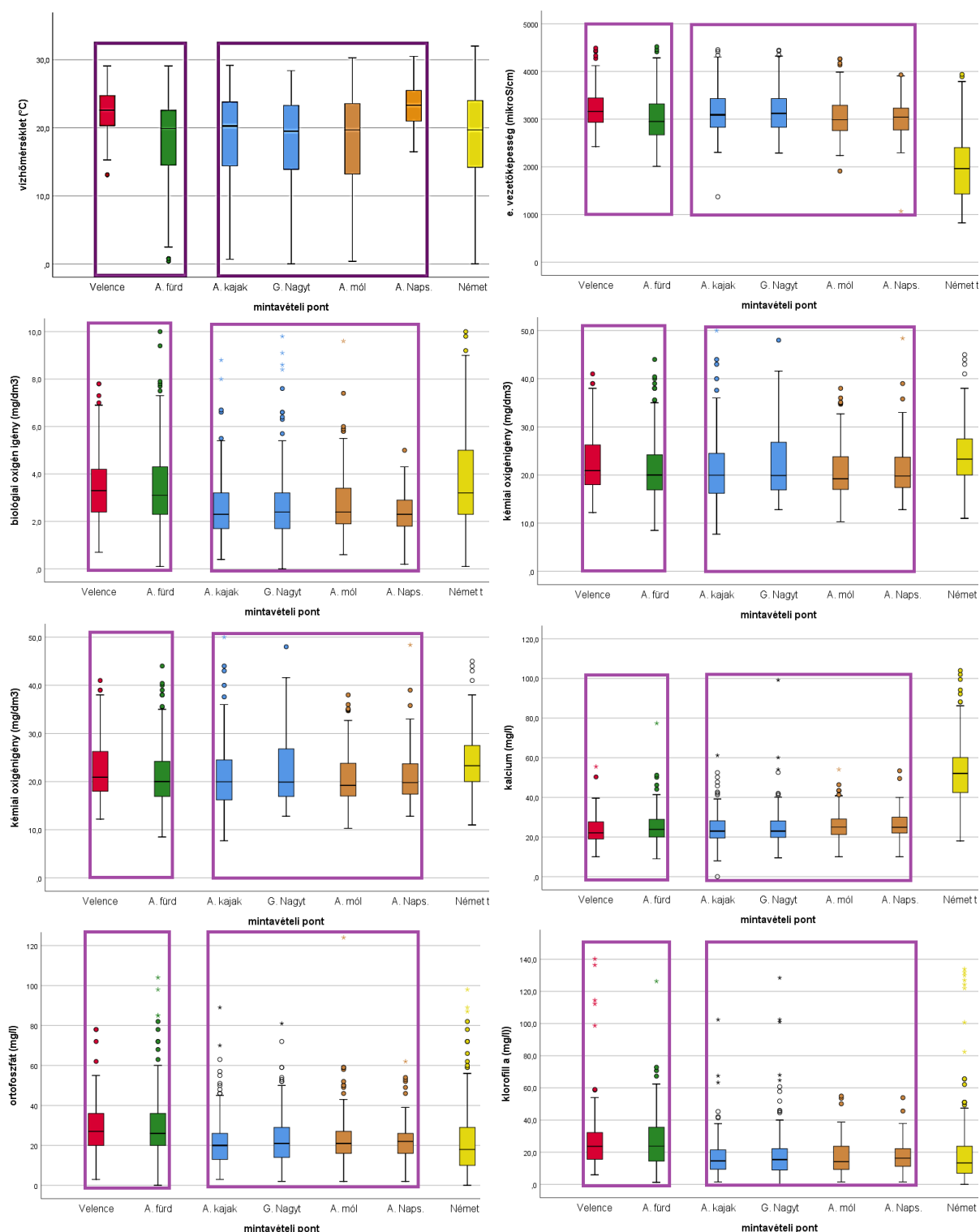
6.ábra: CCDA eredmények, optimális csoportosításra Ward módszerrel, négyzetes euklideszi (A), euklideszi (B), Manhattan (C) és Canberra (D) távolságokkal

Érdekes a csoportok elhelyezkedését térben is megtekinteni (7. ábra). És bár az ábra nem jelöli a nádas helyét és elterjedését, megállapítható a csoportok elhelyezkedése ellentmondás mentes. Az A panelen a tó keleti felében és a közepén lévő nagy csoport szétválásának valószínűsíthető oka a közöttük levő nád. A Német tisztás elnevezésű pont a tó nyugati részében levő nádas terület belsejében fekszik, mely tény különállásának okaként valószínűsíthető. A Canberra távolsággal meghatározott optimális csoportszám nagyobb, egy-egy csoport kevesebb mintavételi pontot tartalmaz.



7.ábra: CCDA-val, Ward módszerrel négyzetes euklideszi, euklideszi, Manhattan (A) és Canberra (B) távolságokkal, meghatározott optimális csoportok elhelyezkedése a Vekencei-tó területén,

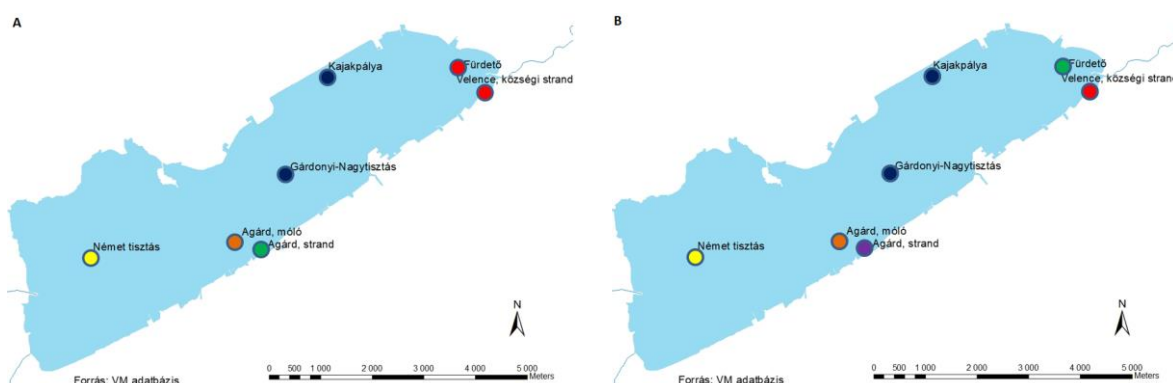
Annak megítélése, hogy melyik távolsággal készített számítás jobb avagy elfogadhatóbb paraméterenként meg lehet vizsgálni. Ehelyett itt a vizuális megjelenítést választom, box-whisker's ábrákon (8. ábra) mutatom be, hogyan alakulnak az egyes paraméterek statisztikái. Olyan paramétereket választottam, melyek a kétféle optimális csoport egyikének vagy másikának létezését támasztják alá. Nagyobb csoportszám (a csoportokat azonos színek jelölik) mellett lehetne érvelni például az elektromos vezetőképesség, vagy a kalcium paraméterekkel. Kisebb csoportszám (keretezett boxok és a Német tisztás) megfelelőbbnek tűnik biológiai oxigén igény, ortofoszfát, klorofill-a paraméterek esetében.



8.ábra: Box-whisker's ábrák néhány paraméterre, az optimális és homogén csoportok szemléltetésére a Velencei-tó mintavételi pontjain. A színezett box-ok a Canberra, lila kerettel és Német tisztás, a négyzetes euklideszi, euklideszi és Manhattan távolságokkal határozott optimális csoportokat jelölik.

Homogén csoportok tekintetében, a Canberra távolsággal meghatározott homogén csoportok száma csak annyiban különbözik a többi távolságfogalommal meghatározott csoportoktól, hogy a Fürdető és a Velencei közleki strand mintavételi pontok különváltak (9. ábra). Ezt látszik alátámasztani több paraméter (hőmérséklet, elektromos vezetőképesség, BOI, KOI) box-

whisker's ábrája (8. ábra), de vannak olyan paraméterek melyek ezt nem erősítik meg (ortofoszfát, klorofill-a). Nagytisztás és Kajakpálya mintavételi pontok homogén csoportja szinte mindegyik bemutatott paraméter esetében elfogadhatónak tűnik. Agárd móló és Agárd Napsugár strand mintavételi pontok egyenként alkotnak önálló csoportot. A különbségek a mért értékekben vagy nem látszanak, vagy nagyon kicsik, a paraméterek box-whisker's ábráin, ugyanakkor ezek a pontok a számítások során mindegyik távolság alkalmazásával szétváltak. Német tisztás megfigyelő pont szinte mindegyik mintavételi ponttól, eltérő mért értékekkel rendelkezik, nem „véletlen”, hogy már az optimális csoportok meghatározásánál is önálló külön csoportot alkotott.



9.ábra: CCD, Ward módszerrel négyzetes euklideszi, euklideszi, Manhattan (A) és Canberra (B) távolságokkal, meghatározott homogén csoportok elhelyezkedése a Vekencei-tó mintavételi pontjain,

Mivel a számítások során alkalmazott négy távolságfogalommal három esetben ugyanazt a csoportosítást kaptuk, úgy vélem elegendő ezek eredményeit figyelembe venni. A gyakorlat számára a számításaim fontos eredményének tartom, hogy véleményem szerint kettő mintavételi pont megszüntethető a vizsgált hét mintavételi pontból. Ez az eredmény természetesen a vizsgált paraméterkörre érvényes és nem szabad elfeledni hogy a számításokba bevont adatokat 1991 -2004 között mérték.

A számítások és az ábrák szerint a Velencei-tó esetében is megjelent az a tapasztalat, hogy a CCD képes minden paraméter egyidejű figyelembevételével mellett kismértékű különbségek észrevételére.

Bízva abban, hogy Dr. Szeidl László kielégítőnek tartja válaszaimat, bírálatát ismételtelen köszönöm.

Budapest, 2020.05.12

Kovács József

Felhasznált irodalom

Benjamin, M. A., Rigby, R. A., Stasinopoulos, D. M., (2003), Generalized autoregressive moving average models. Journal of the American Statistical Association, 98, 214-223.

Kovács, J ; Kovács, S ; Magyar, N ; Tanos, P ; Hatvani, I G; Anda, A: Classification into homogeneous groups using combined cluster and discriminant analysis, ENVIRONMENTAL MODELLING & SOFTWARE 57 pp. 52-59. , 8 p. (2014)

Lilly, J., Olhede, S. (2009). Higher-order properties of analytic wavelets. IEEE Transactions on Signal Processing 57: 146.160.

Internetes hivatkozás: <https://www.hatarilabs.com/ih-en/what-is-a-piper-diagram-and-how-to-create-one>