



## **6-2 háttéranyag**

### **Víztestek biológiát támogató fizikai-kémiai minősítése**

**Dr. Kardos Máté Krisztián**

# Tartalom

Bevezetés .....	3
Minőségi osztály meghatározása .....	<b>Hiba! A könyvjelző nem létezik.</b>
“a” – “c” – Mérési adatok alapján .....	5
Súlyszám.....	5
“m” – Modellel – kizárólag 3. és 6. típusú vízfolyás víztestek esetén	<b>Hiba! A könyvjelző nem létezik.</b>
3. típus – Dombvidéki kisvízfolyások .....	10
6. típus – Síkvidéki kisvízfolyások.....	10
“s” – Szakértői becsléssel – kizárólag állóvíz víztestek esetén	<b>Hiba! A könyvjelző nem létezik.</b>
A minősítés megbízhatóságának számítása .....	12
Néhány észrevétel .....	<b>Hiba! A könyvjelző nem létezik.</b>
Hivatkozások.....	14

## Bevezetés

A víztestek fiziko-kémiai elemek szerinti állapotának értékelése során a VGT2-ben kidolgozott módszertanból indultunk ki, azt fejlesztettük tovább (Clement és Szilágyi, 2015). Az ott meghatározott típuspecifikus határértékeken nem változtattunk. A VGT2 során az osztályzatok komponenscsoportonként<sup>1</sup> átlagolásra kerültek, és a leggyengébb komponenscsoport osztályzata adta a víztest fiziko-kémiai osztályzatát. Ezt az elvet megtartottuk.

A VGT2 során egy víztest egy monitoring pont alapján került minősítésre, ettől most eltértünk. Egy víztesthez egy vagy több monitoring pont adatait vettük figyelembe, eltérő súllyal. Továbbá előfordult, hogy egy monitoring pont adatait egynél több víztestnél is figyelembe vettük, szintén eltérő súllyal.

A monitoring adatok három forrásból származtak. A legjelentősebb adatbázis az Országos Környezetvédelmi Információs Rendszer (OKIR) Felszíni Vizek Szakterületi Rendszermoduljában (FEVISz) szereplő 2013 – 2018. évi adatok. A második adatforrás a KEHOP kiegészítő monitoringprojekt<sup>2</sup> K1, K2 és K15 projektelemei keretében 2019. évben végzett mérések. E két adatforrást közös adatbázisba dolgozták össze és egységes kritériumok szerinti minőségellenőrzésen estek át. Az összedolgozott adatbázis alapján (2020. július 31-i dátummal) elkészítettük az állapotértékelés első változatát, melyben azonban a vártnál magasabb volt a nem minősíthető, ún. „szürke” víztestek száma. Ezért a második (2020. október 22-i) változatban a Vízügyi Igazgatóságok véleményével összhangban néhány víztest esetében figyelembe vettük a Vízügyi Igazgatóságok saját, 2013 – 2018. évi mérési eredményeit.

Azokat a víztesteket, melyekre nem állt rendelkezésre mérési adat, statisztikai modellek eredményei alapján vagy szakértői becsléssel minősítettük.

---

<sup>1</sup> komponenscsoportok: savasság, sótartalom, oxigénháztartás, tápanyagok

<sup>2</sup> A projekt címe: „A Víz Keretirányelv előírásai szerinti monitoring vizsgálatok és az ahhoz szükséges fejlesztések végrehajtása, továbbá a Víz Keretirányelv végrehajtásához kapcsolódó monitoring állomások kiépítése, fejlesztése”; azonosítószáma: KEHOP-1.1.0-15-2016-00002.

## Állapotértékelés

Az állapot értékelése ötosztályú skálán (1 – kiváló, 2 – jó, 3 – mérsékelt, 4 – gyenge, 5 – rossz), döntően vízminőségi monitoring adatok alapján történt. Vízfolyás víztestek esetén – a VGT2-vel egyezően – négy komponenscsoportban tizennégy komponens elemet vettük figyelembe:

- savasodás (pH)
- sótartalom (elektromos vezetőképesség, kloridion-koncentráció)
- oxigénháztartás (oldott oxigén, oxigén telítettség, biokémiai oxigénigény, dikromátos kémiai oxigénfogyasztás, összes szerves szén, ammónium-ion)
- tápanyagok (összes szerves nitrogén, összes nitrogén, ortofoszfát ion, összes foszfor).

Állóvíz víztestek esetén – szintén a VGT2-ben megteremtett hagyományokat követve – ugyanebben a négy komponenscsoportban a következő komponenseket vettük figyelembe:

- savasodás (pH)
- sótartalom (elektromos vezetőképesség)
- oxigénháztartás (biokémiai oxigénigény, dikromátos oxigénfogyasztás, összes szerves szén)
- tápanyagok (ammónium ion, nitrit-ion, nitrát-ion, összes nitrogén, ortofoszfát ion, összes foszfor).

A határértékek tekintetében a vízfolyás víztesteket tíz típusba (1-10), az állóvíz víztesteket 11 típusba (1-4, 5A, 5B, 5C, 6-8) soroltuk. A típus-specifikus határértékeket a VGT2 6-3. mellékletéből vettük át. A határértékek tehát függetlenek voltak a víztest kategóriájától (természetes, erősen módosított, mesterséges). Komponensenként elvégeztük az osztályba sorolást, tehát minden víztest kapott minden komponens tekintetében egy osztályzatot (1 azaz kiváló és 5 azaz rossz között). Az osztályzatokat komponenscsoportonként átlagoltuk és a leggyengébb csoport-osztályzat vált a víztest fiziko-kémiai osztályzatává.

A monitoring adatok első sorban magukból a víztestből vett minták mérési eredményei, másodsorban a víztesttel hidrológiai kapcsolatban álló (felvív vagy alvív irányban hozzá csatlakozó) víztestből vett minták, vagy e kettő súlyozott átlaga alapján történt.

## Az állapot értékelése mérési adatok alapján

A minősítés alapja a Fevisz2020v2b.accdb tábla, amely a 2013-18. évi OKIR-FEVISz adatokat, valamint a KEHOP projekt 2019. évi adatait tartalmazza. Néhány víztest esetében figyelembe vettük a területileg illetékes Vízügyi Igazgatóság vízminőségmérési eredményeit „Adathianyos\_viztestek\_VIZIG-ek.xlsx”, („d”). Ilyen esetekben a víztest minősítésének megbízhatósága csak alacsony lehet, hiszen az értékelés nem a Kormányhivatalok Laboratóriumai által vett és mért minták eredményei alapján történt.

“a” – a minősítés a víztesten található egy vagy több monitoring pont adatai alapján történt

“b” – a minősítés kettő vagy több monitoring pont adatai alapján történt, melyek közül legalább egy a víztesten, legalább egy pedig azon kívül, a víztesttel felvízi vagy alvízi irányban hidrológiai kapcsolatban álló szomszédos víztesten helyezkedett el.

“c” – a minősítés egy vagy több monitoring pont adatai alapján történt, melyek mindegyike a víztesten kívül, a víztesttel felvízi vagy alvízi irányban hidrológiai kapcsolatban álló szomszédos víztesten helyezkedett el. A “b” – “c” eseteket “kapcsolásnak” nevezzük: valamely víztest

„d” – a minősítés a Vízügyi Igazgatóság kiegészítő mérései alapján történt minősítés, a méréseket közvetlenül a víztestből vett mintákon végezték (tehát ezen adatforrás esetén nem végeztünk „kapcsolás”-t).

### **Monitoring állomások súlyszáma**

A monitoring állomás súlyszáma azt fejezi ki, hogy annak adatai mennyire jellemzők az adott víztestre. Két részből tevődik össze: egyrészt a monitoring hely és víztest földrajzi viszonyából, másrészt a monitoring hely adatellátottságából. Tehát a súlyszám csak az “a” – “c” esetekben értelmezett, és csak akkor, ha egynél több monitoring állomás játszik szerepet az adott víztest minősítési osztályzatának kialakításában.

Matematikailag egy-egy monitoring állomás súlya két szám szorzataként adódik.

$$s = s_1 \cdot s_2 \quad (1)$$

Az  $s_1$  súlyszám szakértői becsléssel kerül meghatározásra és azt hivatott tükrözni, hogy az adott pont mennyire reprezentatív a víztest egészének vízminősége szempontjából. Az  $s_1$  súlyszám meghatározásakor az alábbi szempontokat mérlegeltük.

- a monitoring állomáshoz és a víztest kifolyási pontjához tartozó vízgyűjtő aránya
- a monitoring pont és a víztest kifolyási pontja között található mellékágak, terhelések
- a vízminőség változékonysága (gyors folyású nagyvizek vs. pangó csatornák)
- összetett víztestek esetén: a monitoring pont a főágon, a mellékágon, vagy már a „közös szakaszon” található-e?

Az  $s_2$  súlyszám objektív adat: az adott állomáson adott komponens tekintetében rendelkezésre álló mérések száma.

### Állapotértékelés statisztikai modellekkel

A statisztikai modellek egy vagy több független változó és egy függő változó közötti kapcsolatot leírására alkalmasak. Mind a független, mind a függő változók lehetnek akár folytonos, akár kategorikus (azon belül akár ordinális, akár nominális) skálán mért változók. Egy statisztikai modell felállításának első lépése a tanítás vagy kalibráció (beállítás). Ennek során az egyenletek paramétereit úgy állítjuk be, hogy azok minél pontosabban írják le a tanító adathalmaz eseteit, azaz a tanító adathalmaz független változóiból a lehető legpontosabban számolják a tanító adathalmaz függő változóját. A folyamatot illesztésnek is nevezzük. A kalibrációt követi a validáció, melynek egyik fajtája a keresztvalidáció. A sikeres kalibrációt, validációt követi a modell alkalmazása független adatbázison, azaz a becslés vagy predikció (Kardos, 2020; Kardos és Clement, 2020).

A legismertebb statisztikai modelles család a regresszió, mely hagyományosan folytonos független és folytonos függő változókkal dolgozik. A független változók száma szerint megkülönböztetünk egy- vagy többváltozós, a leíró függvény szerint pedig lineáris vagy nemlineáris regressziót. A bináris logisztikus regresszió kétosztályú kategorikus, az ún. „arányos esélyek” logisztikus regresszió (angolul: proportional odds logistic regression – POLR) ordinális kettő- vagy többosztályú kategorikus változó becslésére használható. Mindkettő folytonos független változókkal dolgozik. A diszkriminancia elemzés lehet lineáris

vagy négyzetes, nominális skálán mért két- vagy többosztályú függő változó értékét becsüli elsősorban szintén folytonos független változókból (Kardos, 2020, 2019).

Az illesztés sikerességét, az illeszkedés mértékét több módon is vizsgálhatjuk. Érdekes felírni a megfelelési mátrixot, mely a víztestek becsült állapotát mutatja a tényleges állapot függvényében. Kiolvasható tehát belőle, hogy a modell hány víztest állapotát „találta el”, illetve mennyiét „lőtte mellé”. A téves osztályba sorolt eseteket a megfelelési mátrix alapján külön is vizsgáltuk. A hibás besorolásnak két fajtája van. Az egyik, ha egy jó állapotú víztestet nem jó állapotúnak mondunk. Ez tévedés ugyan, de pusztán annyi vonzata lesz, hogy intézkedéseket hozunk a víztest állapotának javítására, úgymond „feleslegesen”; az intézkedéseknek azonban így is pozitív hatásuk lesz a vízminőségre. Ez tehát a kevésbé súlyos tévedés, az elsőfajú hiba. A durvább tévedés, ha egy nem jó állapotú víztestet jónak mondunk, ilyenkor ugyanis elmarad egy szükséges intézkedés: ezt másodfajú hibának hívjuk. A becsülő változók „hasznosságát” mutatja a variancia inflációs faktor (VIF), értéke azt mutatja meg, hogy az újonnan bevont becsülő változó a korábban bevont változókkal vagy a függő változóval korrelál-e erősebben. Maximális értéke szakirodalmi forrásonként változik; a megengedőbbek 10, a szigorúbbak az 5-ös variancia infláció alatti értékeket tartják elfogadhatónak (Hair és mtsai., 2014; O'Brien, 2007). A modellek egymással való összehasonlítására alkalmasak a különböző információs kritériumok, például az Akaike információs kritérium (AIC) (Akaike, 1969).

Esetünkben az alapprobléma, hogy csupán a víztestek egy részére áll rendelkezésre helyszíni mérési adat és a fiziko-kémiai állapot. A viszonylag kis vízgyűjtőterületű vízfolyások ( $A < 1000 \text{ km}^2$ , víztest típus: 1-3 és 5-6) esetében széleskörű információval rendelkezünk azonban a vízgyűjtőről (domborzat, területhasználat, pontszerű terhelések). Ezért a következőképpen jártunk el. Statisztikai modellt illesztünk azokra a víztestekre, melyekre rendelkezésünkre áll a fiziko-kémiai állapot. A modell független változói a víztest kifolyási pontjához tartozó vízgyűjtő domborzati, területhasználati és pontszerű terhelési jellemzői, függő változó pedig a víztest fiziko-kémiai állapota (1 -5 skálán). A modellillesztést külön-külön végeztük a 3. illetve 6. típusba tartozó kisvízfolyás víztestekre.

**1. táblázat. A tanító adathalmazban szereplő víztestek kifolyási pontjához rendelt vízgyűjtőterületnek a modellillesztésbe bevont jellemzői. Zárójel előtt az átlagos, zárójelben a legkisebb – legnagyobb értékek. A víztest osztálya: kiváló: 1 ... rossz: 5.**

Jel	Jellemző	3. típus	6. típus
<b>Domborzat és hidrológia</b>			
Ter.	Vízgyűjtőterület (km <sup>2</sup> )	240 (25 - 1000)	360 (4 - 1000)
Mag.	Átlagos tengerszint feletti magasság (m tsz. f.)	280 (150 - 540)	150 (85 - 270)
Lejt.	Átlagos lejtés (%)	8,8 (1,7 - 16)	3,1 (0,32 - 8,8)
KÖQ	Sokéves középvízhozam (m <sup>3</sup> /s)	0,75 (0,053 - 4)	0,75 (0,008 - 3,5)
<b>Területhasználati arányok</b>			
Mest.	Mesterséges burkolatú területek (%)	7 (2,4 - 33)	6,6 (1,6 - 22)
Mg.	Mezőgazdasági területek (%)	47 (5,6 - 88)	62 (26 - 90)
Erdő	Erdő és természetes vegetációk (%)	46 (4,3 - 86)	30 (1,1 - 71)
Víz	Nyíltvizes és vizenyős területek (%)	0,27 (0 - 2)	1,4 (0 - 12)
<b>Pontszerű terhelések</b>			
szvQ	Szennyvízarány a KÖQ-hoz képest (%)	5,3 (0 - 63)	26 (0 - 350)
szvBOI	BOI terhelés a KÖQ-hoz képest (g/s) / (m <sup>3</sup> /s)	0,61 (0 - 5)	3,3 (0 - 53)
szvKOI	KOI terhelés a KÖQ-hoz képest (g/s) / (m <sup>3</sup> /s)	3,3 (0 - 56)	14 (0 - 230)
szvN	Összes N terhelés a KÖQ-hoz képest (g/s) / (m <sup>3</sup> /s)	0,85 (0 - 5)	4,7 (0 - 63)
szvP	Összes P terhelés a KÖQ-hoz képest (g/s) / (m <sup>3</sup> /s)	0,074 (0 - 0,63)	0,74 (0 - 11)
<b>Az állapotértékelés eredménye</b>			
FKO	Víztest fiziko-kémiai osztálya	2,44 (1 - 5)	2,47 (1 - 5)

Jelen munka során a Kardos & Clement (2020) cikkben illetve Kardos (2020)-ban ismertetett gondolatmenetet alkalmaztuk. A víztestek összevont vízgyűjtőire képzett háttérváltozók a hivatkozott publikációkban leírtakkal megegyeznek. Különbség két ponton van. Egyrészt a hivatkozásokban ismertetett BLR és LDA modellekkel szemben jelen munka során a POLR modellt alkalmaztuk. Másrészt a modellekbe bevont becsülő változókban: a hivatkozásokhoz képest más becsülő változók kerültek kiválasztásra. A modellezésbe vont becsülő változók kiválasztását itt is korrelációs mátrix segítette (ld. 2. táblázat **Hiba! A hivatkozási forrás nem található.**).

Dombvidéki modellekhez a szvP, Erdő, szvKOI változókat használtuk prediktorként, melyeket ebben a sorrendben fokozatosan vontunk be. Síkvidéki változók: szvP, Erdő, Mest., szvKOI, Ter., Víz, melyeket szintén fokozatosan vontunk be, ebben a sorrendben.



**7. táblázat. Korrelációs mátrix. Jobbra fent: 3. típusú (dombvidéki) vízfolyás víztestek. Balra lent: 6. típusú (síkvidéki) vízfolyás víztestek. Szignifikanciaszintek: \*\*\* < 0,001 < \*\* < 0,01 < \* < 0,05 < <sup>x</sup> < 0,1.**

	Ter.	Mag.	Lejt.	Mest.	Mg.	Erdő	Víz	szvQ	szvBOI	szvKOI	szvN	szvP	FKO
Ter.		0.26***	0.07	0.1 <sup>x</sup>	-0.01	-0.01	-0.08	0.10 <sup>x</sup>	0.13*	0.11*	0.14**	0.15**	0.05
Mag.	0.03		0.75***	0.1 <sup>x</sup>	-0.6***	0.58***	-0.34***	0.09 <sup>x</sup>	0.00	0.01	0.06	0.03	-0.28***
Lejt.	0.01	0.87***		0.01	-0.63***	0.62***	-0.27***	-0.02	-0.01	0.00	-0.02	-0.04	-0.29***
Mest.	0.00	0.05	0.04		-0.10 <sup>x</sup>	-0.16**	-0.02	0.48***	0.58***	0.55***	0.6***	0.58***	0.32***
Mg.	0.01	-0.52***	-0.51***	-0.12*		-0.97***	0.31***	-0.10 <sup>x</sup>	-0.04	-0.07	-0.07	-0.05	0.36***
Erdő	0.00	0.6***	0.57***	-0.11 <sup>x</sup>	-0.87***		-0.35***	-0.02	-0.1 <sup>x</sup>	-0.07	-0.08	-0.1 <sup>x</sup>	-0.43***
Víz	-0.02	-0.16**	-0.1 <sup>x</sup>	-0.12*	-0.29***	-0.11 <sup>x</sup>		-0.04	0.00	-0.02	-0.03	-0.03	0.03
szvQ	0.04	0.02	0.00	0.33***	-0.08	0.01	-0.05		0.69***	0.77***	0.81***	0.79***	0.4***
szvBOI	0.02	-0.03	-0.08	0.23***	-0.10 <sup>x</sup>	0.06	-0.06	0.73***		0.89***	0.85***	0.83***	0.39***
szvKOI	0.03	-0.04	-0.08	0.28***	-0.07	0.02	-0.05	0.9***	0.92***		0.84***	0.84***	0.39***
szvN	0.03	-0.01	-0.03	0.34***	-0.07	0.00	-0.05	0.96***	0.79***	0.94***		0.94***	0.43***
szvP	0.02	-0.04	-0.06	0.26***	-0.02	-0.04	-0.05	0.92***	0.75***	0.93***	0.93***		0.44***
FKO	0.11*	-0.07	-0.09	0.19**	0.21***	-0.2***	-0.17**	0.29***	0.17**	0.26***	0.31***	0.28***	

### 3. típusú vízfolyás víztestekre (dombvidéki kisvízfolyások) illesztett modellek táblázatai

A 3. táblázat a dombvidéki kisvízfolyás víztestekre illesztett modelleket foglalja össze. Látható, hogy a legjobb találati arány a 3.3 modellre jellemző (62%), ugyanakkor a VIF értéke 5 alatt marad. A 3.3 modell becslő változói az erdő területarány, a szv. KOI terhelés valamint az szv. foszfor terhelés.

#### 3. táblázat. A vizsgált modellek összehasonlító táblázata

modell	találati arány	VIF erdő	VIF szvKOI	VIF szvÖP	AIC
3.1	58%				770
3.2	61%	1.88		3.51	696
3.3	62%	4.64	4.94	3.53	696

4. táblázat. A kiválasztott modell megfelelési mátrixa. A mátrix értékei az adott osztályú víztestek számát jelölik. Piros: másodfajú hibák.. <sup>x</sup>kétosztálynyi tévedés, <sup>\*</sup>három osztálynyi tévedés. A nullákat nem tüntettem föl.

mért	össz.	3.4 modell					találat
		kiváló	jó	mérs.	gyenge	rossz	
kiváló	23		21	2			0%
jó	181	1	<b>151</b>	29			83%
mérsékelt	111		<b>54</b>	<b>51</b>	6		46%
gyenge	30		<b>6<sup>x</sup></b>	16	<b>8</b>		27%
rossz	1			1 <sup>x</sup>			0%
<i>összesen</i>	<i>346</i>	<i>1</i>	<i>232</i>	<i>99</i>	<i>14</i>	<i>0</i>	<i>62%</i>

### 6. típusú vízfolyás víztestekre (síkvidéki kisvízfolyások) illesztett modellek táblázatai

#### 5. táblázat. A vizsgált modellek összehasonlító táblázata

mod	találati arány	VIF Ter.	VIF Mest.	VIF Erdő	VIF Viz	VIF szvKOI	VIF szvÖP	AIC
6.1	40%							973
6.2	40%			1.15			2.73	964

mod	találati arány	VIF Ter.	VIF Mest.	VIF Erdő	VIF Viz	VIF szvKOI	VIF szvÖP	AIC
6.3	40%		2.99	1.19			3.43	963
6.4	40%		3.00	3.34		3.07	3.44	962
6.5	41%	2.31	3.00	3.52		3.24	3.69	959
6.6	40%	2.31	3.06	1.19	3.43	3.14	3.93	947

Síkvidéken a választás a 6.5 számú modellre esett, ennek becslő változói a vízgyűjtőterület, a burkolt területek aránya, az erdők aránya, a szvKOI és a szvÖP terhelések.

**6. táblázat. A vizsgált modellek összehasonlító táblázata. A mátrix értékei az adott osztályú víztestek számát jelölik. Piros: másodfajú hibák.. <sup>x</sup>kétosztálynyi tévedés, <sup>\*</sup>három osztálynyi tévedés. A nullákat nem tüntettem föl.**

	mért	6.5 modell					találat
		össz	kiváló	jó	mérsékelt	gyenge	
kiváló	53	4	45	4 <sup>x</sup>			8%
jó	136	2	129	5			95%
mérsékelt	100	2 <sup>x</sup>	87	9	2		9%
gyenge	51		40 <sup>x</sup>	7	3	1	6%
rossz	12		7 <sup>*</sup>	2 <sup>x</sup>	2	1	8%
<i>összesen</i>	352	8	308	27	7	2	41%

### Állapotértékelés szakértői becsléssel

Csak kivételes esetekben értékeltük az állapotot szakértői becsléssel. Ezekben az esetekben a terhelés hiánya, történelmi adatok, földrajzilag közel elhelyezkedő, azonos típusba tartozó víztestek adatai, illetve területileg illetékes Vízügyi Igazgatóságok helyismerete alapján adtunk osztályzatot. A konkrét értékeket és azok indoklását a 7. táblázat tartalmazza.

### **7. táblázat. Szakértői becsléssel megállapított osztályzatok és azok indoklása**

VT VOR	VT NÉV	Típus	FKO	indoklás
ANS517	Korcánypusztai-halastavak	8	2	nincs terhelés

VT VOR	VT NÉV	Típus	FKO	indoklás
AIH084	Kardoskúti-Fehér-tó	3	1	VGT2-es osztályzat
AIG930	Biharugrai-halastavak	8	2	AIG926 Begécsi-halastavak
AIP534	Öregállás V. tó	6	5	AIP532 Öregállás II., AIP533 Öregállás III., AIP539 Tripolisz
AIH042	Zalaszentmihályi-horgásztó	5	3	AIH009 Pölöskei K-i tározó
AIH092	Kolon-tó	6	3	ADUVIZIG észrevétele alapján
AEP525	Gögő-Szenke főcsatorna	6	3	FETIVIZIG észrevétele alapján

## A minősítés megbízhatóságának számítása

Alacsonynál jobb megbízhatóság csak OKIR-FEVISz illetve KEHOP monitoring adatok alapján történő minősítés esetén adható („a” – „c” esetek). Ebben az esetben az adott tartományba esés  $p$  valószínűségét a centrális határeloszlás tétele alapján a következő módon számoljuk.

A mintából az (2) – (3) képletekkel becsülhetjük a sokaság várható értéket, szórását:

$$\hat{\mu} = \frac{\sum_{i=1}^n x_i}{n} \quad (2)$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (x_i - \hat{\mu})^2}{n - 1}} \quad (3)$$

A becsült átlag szórását az alábbi (4) képlettel számolhatjuk.

$$\sigma(\hat{\mu}) = \frac{\hat{\sigma}}{\sqrt{n}} =: s \quad (4)$$

Az osztályhatároktól ( $x_1, x_2$ ) mért távolságot a sűrűségfüggvény segítségével számoljuk, melynek standard ( $\mu=0, \sigma=1$ ) alakját mutatja a (4) képlet:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{t=-\infty}^x e^{-\frac{t^2}{2}} dt \quad (5)$$

Ha a várható érték  $\mu = \hat{\mu}$ , a szórás pedig  $\sigma = s$ , akkor a sűrűségfüggvény a (5) szerinti alakban írható föl.

$$F(x) = \frac{1}{s\sqrt{2\pi}} \int_{t=-\infty}^x e^{-\frac{(t-\hat{\mu})^2}{2s^2}} dt \quad (6)$$

Tehát  $\hat{\mu}$  várható értékű,  $s$  szórású valószínűségi változó ( $x_1, x_2$ ) tartományba esésének valószínűsége a (6) képlet szerinti.

$$p = F(x_2) - F(x_1) = \frac{1}{s\sqrt{2\pi}} \left( \int_{t=-\infty}^{x_2} e^{-\frac{(t-\hat{\mu})^2}{2s^2}} dt - \int_{t=-\infty}^{x_1} e^{-\frac{(t-\hat{\mu})^2}{2s^2}} dt \right) \quad (7)$$

Egyszerűsítve:

$$p(x_1 < x < x_2) = \frac{1}{s\sqrt{2\pi}} \int_{t=x_1}^{x_2} e^{-\frac{(t-\hat{\mu})^2}{2s^2}} dt \quad (8)$$

A fiziko-kémiai osztályzat megbízhatóságát annak a komponenscsoportnak a megbízhatósága adja, amelyik az osztályzatot is adta (tehát a legrosszabb osztályzatú csoport). A megbízhatósági számérték helyett a víztesteket – a minősítésük megbízhatósága alapján – három osztályba soroljuk

- magas a megbízhatóság, ha  $p \geq 0,9$ ,
- közepes, ha  $0,7 \leq p < 0,9$ ,
- és alacsony, ha  $p < 0,7$ .

## Hivatkozások

Akaike, H., 1969. Fitting autoregressive models for prediction. *Ann. Inst. Stat. Math.* 21, 243–247.

Clement, A., Szilágyi, F., 2015. Felszíni víztestek fizikai kémiai állapotértékelési rendszere - OVGT 6-2 háttéranyag). Budapest.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., 2014. *Multivariate Data Analysis (MVDA)*. Pearson. <https://doi.org/10.1002/9781118895238.ch8>

Kardos, M.K., 2020. Felszíni vízminőségi monitoring rendszer hatékonyságának növelése statisztikai modellek alkalmazásával. THESIS PhD. Budapesti Műszaki és G. Egyetem.

Kardos, M.K., 2019. Folyóhálózati monitoringhelyek csoportosítására szolgáló algoritmus kidolgozása. BME VKKT 1–28.

Kardos, M.K., Clement, A., 2020. Predicting small water courses' physico-chemical status from watershed characteristics with two multivariate statistical methods. *Open Geosci.* 12, 71–84. <https://doi.org/10.1515/geo-2020-0006>

O'Brien, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. *Qual. Quant.* 41, 673–690. <https://doi.org/10.1007/s11135-006-9018-6>

## Melléklet

FEV\_AVK\_minosites\_v090\_out20201210.xlsx