Fejlesztések a jártassági vizsgálatokban

Fücsök Ferenc

MAROVISZ, ffucsok@gmail.com

Absztrakt

Mint azt a minőségirányítási alapelvek rögzítik, a Magyar Roncsolásmentes Vizsgálat Szövetség (MAROVISZ) a felhasználók érdekében fejleszti a jártassági vizsgálatok minőségét. Ezt a feladatot több kisebb, de lényeges területen folyamatosan végezzük.

A jártassági vizsgálatok értékelése közben kisebb-nagyobb értelmezési problémákkal találkozunk. Egyik ilyen volt a JV07 jelű hegesztési varrat ultrahangos vizsgálata, ahol az ROC diagram meghatározásához használt módszerünk cellamérete rejtetten olyan szigorú tűrést alkalmazott, amit kézi ultrahangos vizsgálattal nem lehetett megvalósítani. Az értékelés eredményének alapján kijelenthetjük, hogy mindegyik cellamérettel végzett értékelésnek van előnye és hátránya.

A matematikusok a vizsgáló teljesítményének jellemzésére ajánlják az ROC görbe alatti terület számítását. Az említett összehasonlító elemzésnél megvizsgáltuk a görbe alatti terület alkalmazását is, és jelenleg úgy látjuk, érzékenyebben, és kisebb hibákkal mutatja be a vizsgálók teljesítményei közti különbségeket. Az elemzés eredményeit a következő jártassági vizsgálatokban használni fogjuk.

A skaláris mérési eredmények megbízhatóságát a mérési bizonytalansággal lehet jellemezni, aminek ismertetését és használatának körülményeit is bemutatjuk.

1. Bevezetés

A Magyar Roncsolásmentes Vizsgálati Szövetség (MAROVISZ), mint a laboratóriumok többsége, rendelkezik ISO 9001 szabvány szerinti tanúsítással. A tanúsítás követelményeinek megfelelően szükség van a tevékenységek folyamatos fejlesztésére, így a jártassági vizsgálatok fejlesztésére is. Ebben a cikkben az észlelési feladatok értékelésénél használt módszerek fejlesztésével és a mérési bizonytalanság alkalmazásával kapcsolatos problémákról fogunk beszámolni.

Jelen cikkben nem foglalkozunk a szervezési problémákkal, melyek szintén fejlesztésre szorulnának. Egy példában jelezve: a jártassági vizsgálatban részt vevő laborok a megbeszélt határidőre mindig visszaküldik a vizsgált munkadarabokat, de a mérési eredmények elküldéséről gyakorta elfelejtkeznek…

2. A JV07 vizsgálat problémái

A JV07 jelű jártassági vizsgálat értékelésénél olyan problémát találtunk, ami megnehezítette a korrekt eredmény megállapítását. Ennél a vizsgálatnál egy hegesztési varrat kézi ultrahangos vizsgálatát kellett elvégezni olyan varratban, ahol hét darab szikraforgácsolással bemunkált mesterséges reflektor volt. A résztvevők vizsgálati eredményeinek értékelésénél a hegesztési varratot 1 cm hosszú cellákra osztottuk, és azt vizsgáltuk, hogy a résztvevő a cellában megtalálta-e a hiányt vagy sem. Az eredmények azt mutatták, hogy az 1 cm hosszú cellák olyan rejtett tűrési követelményt tartalmaztak, amit a kézi ultrahangos vizsgálattal nem lehetett megvalósítani. Ennek következtében az észlelési valószínűség (POD) és a téves detektálási valószínűség (PFC) számítása értékelhetetlen eredményeket adott, amit a JV07 összefoglaló jelentés nem tartalmazott. Helyette más követelmények szerinti értékelést kellett alkalmazni.

Az azóta elmúlt időben volt módunk, hogy a jellemzők értékelését különböző cella méretekkel elvégezzük, és összehasonlítsuk az eredményeket. A cellák méreteit 2,5 cm, 1 cm és 0,2 cm hosszúságúra vettük fel, és így végeztük a vizsgálati jellemzők számítását.

2.1 A vizsgálati jellemzők

Először nézzük át, hogy melyek azok a jellemzők, melyeket egy észlelési feladat értékelésénél meg szoktunk határozni. Az észlelési valószínűséget (POD) a következő képlettel számítjuk:

$$\mathrm{POD} = \frac{\mathrm{PI}}{\mathrm{N1}} = \frac{\mathrm{PI}}{(\mathrm{PI}+\mathrm{NH})}, \qquad (1)$$

ahol:

PI = pozitív igaz cellák, azaz a találatok száma.
Magyarázat: helyesen detektált hibás cellák.

NH = negatív hamis cellák, azaz a nem megtalált hibás cellák száma.
Magyarázat: tévesen nem detektált hibás cellák.

N1 = az összes hibás cellák száma.

A vizsgálók teljesítményének meghatározásához a téves detektálás valószínűségét (PFC) is ki kell számítani, amit a következő képlettel lehet elvégezni:

$$PFC = \frac{PH}{N2} = \frac{PH}{(PH + NI)}, \qquad (2)$$

ahol:

PH = pozitív hamis cellák, azaz a hamis riasztások = tévesztések száma.
Magyarázat: tévesen detektált hibátlan cellák.

NI = negatív igaz cellák, azaz a helyesen elfogadott cellák száma.
Magyarázat: helyesen nem detektált hibátlan cellák.

N2 = az összes hibátlan cellák száma.

A vizsgálók teljesítményét (K) az előbbi két jellemző összevonásával a következő képlettel határozhatjuk meg, feltéve, hogy az ROC görbe hiperbolikus (ROC = Receiver Operating Characteristic = vevő [régen a radar jelek vevője] működési jelleggörbe):

$$K = 1 - 2\sqrt{(1 - POD)PFC}. \qquad (3)$$

A K értéke 0 (rossz) és 1 (kiváló) között terjedhet. A 0 értéket a találgató vizsgáló, az 1 értéket a hibátlanul dolgozó, ideális vizsgáló érheti el.

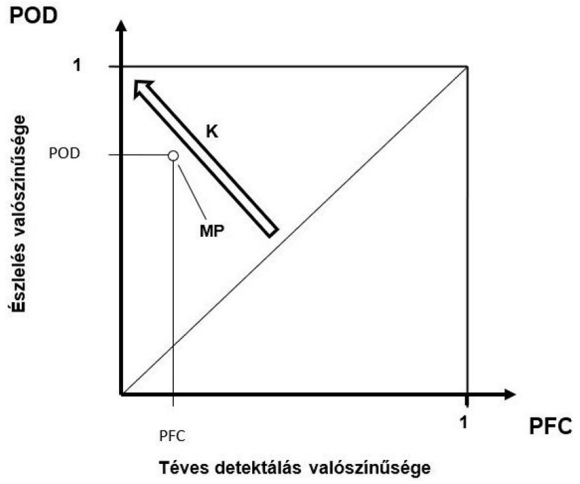
A jártassági vizsgálatokban elvégzendő feladatok statisztikailag kevés mérést tartalmaznak, melyekből az ROC görbét nem lehet meghatározni. Az eddigi jártassági vizsgálatokban a görbének csak egy pontját, a munkapontot határoztuk meg. Ebből az egy pontból pedig nem lehet megállapítani, hogy a görbe hiperbolikus-e. Statisztikusok ajánlása, hogy az ROC görbe alatti terület ennél érzékenyebb és pontosabb eredményt ad a vizsgálók teljesítményéről.

Az ROC diagram származtatását és a munkapont (MP) értelmezését az 1. ábrán mutatjuk be.

A diagramon a vízszintes x tengelyre a téves detektálást, más néven hamis riasztások valószínűségét (PFC) mérjük fel, a függőleges y tengelyre pedig az észlelés valószínűségét (POD). A két koordináta metszéspontjában a munkapont jelenik meg, az ábrán MP kódolja. Az átlóra merőleges, a (0;1) pontba mutató nyíl a K értékét jelképezi, ami nyíl irányában nő.

Ha a görbe több pontját meg tudjuk határozni, akkor a 2. ábrán látható diagramot rajzolhatjuk meg. Az esetek többségében az ROC görbe hiperbolikus, mint az ábrán is látható. Az ábrán a görbe alatti terület (AUC = Area under Curve = görbe alatti terület) értelmezését is bemutatjuk a szürke felülettel.

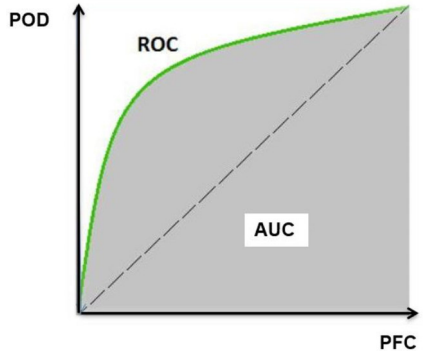
A görbe alatti terület meghatározása görbe határoló vonal esetén komplikált lehet, de ha csak egy munkapontot tudunk meghatározni, és azt kötjük össze az átló végpontjaival, akkor egy egyszerűbben kiszámítható görbe alatti területet kapunk, mint az a 3. ábrán látható.



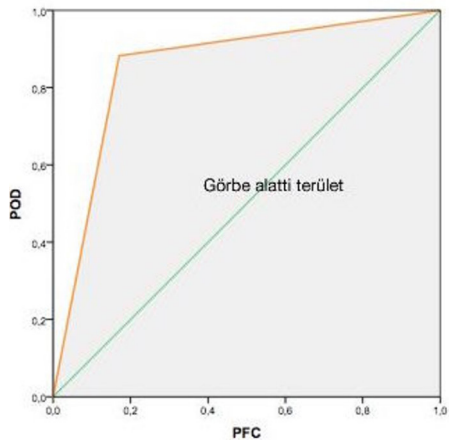
1. ábra: *A ROC görbe szerkesztése*

Az ábra alapján könnyen belátható, hogy a görbe alatti terület a (0;0) és (1;1) pontok által meghatározott négyzet felének-, és az átló felett látható háromszög területéből adódik össze.

Az ismertetett jellemzők változásait fogjuk vizsgálni a következőkben.



2. ábra: *A hiperbolikus ROC görbe, és a görbe alatti területe (AUC)*



3. ábra: *Görbe alatti terület egy munkaponttal számítva*

2.2 A számítások eredményei

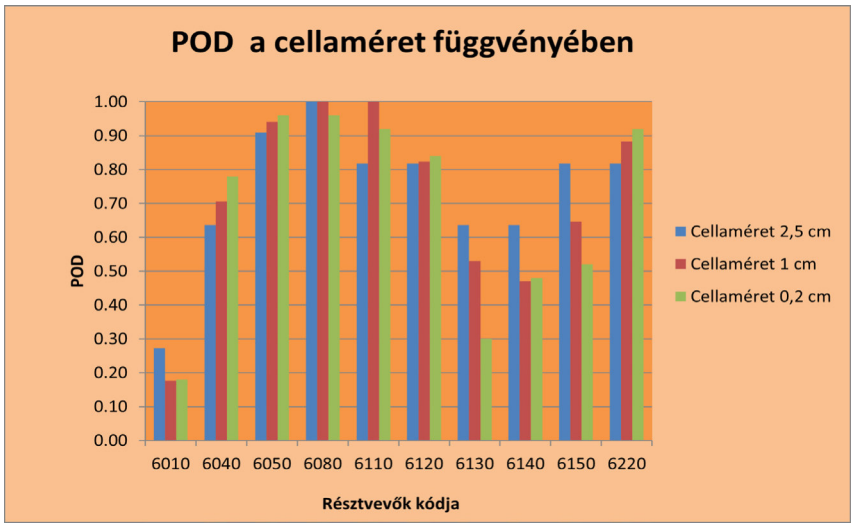
A fentiekben ismertetett jellemzők változásait értékeltük egy már lezajlott jártassági vizsgálat valódi adataival. A JV07 jelű jártassági vizsgálat 1 jelű próbatestjén elvégzett vizsgálatok eredményeit értékeltük a már említett 0,2 cm, 1 cm és 2,5 cm-es cellaméretek alkalmazásával. A következő diagramokban 10 résztvevő eredményeinek változását mutatjuk be a cellaméretek függvényében.

A 4. ábrán bemutatott adatok alapján megállapítható, hogy az értékelésnél használt cellaméretnek nincs egyértelmű hatása az észlelési valószínűség értékére. Van öt olyan résztvevő, akiknek POD értékei 2,5 cm-es cellaméretben jobbak, másik 2 résztvevőnek 1 cm-es cellaméretben jobbak, a maradék három résztvevőnél a legkisebb méret a kedvező. Tehát nincs egyértelmű korreláció a cellaméret és a POD sorrendje között.

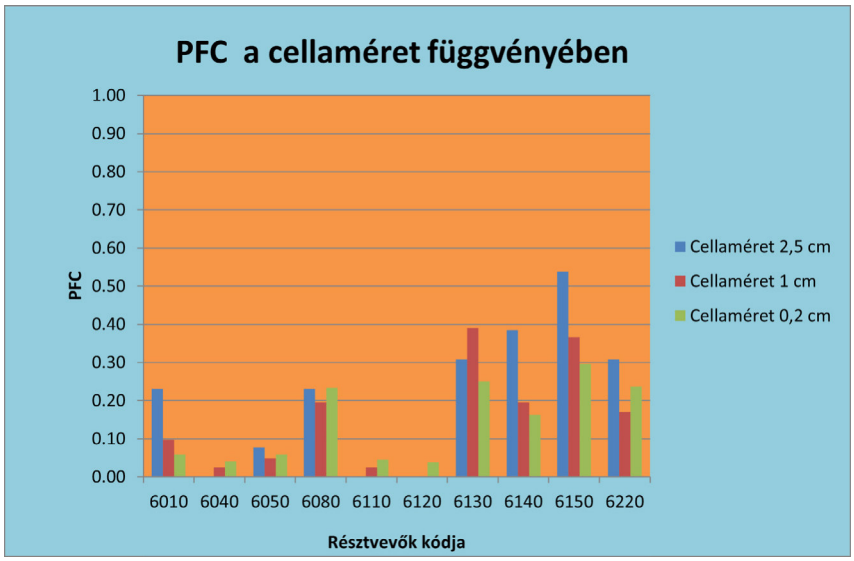
A téves riasztás valószínűségének (PFC) értékei sem kapcsolódnak a cellaméthez (5. ábra). A (kedvező) legkisebb értékeket három résztvevő érte el 2,5 cm-es cellaméret esetén, de ugyanezen résztvevők POD értékei is ekkor a legkisebbek (kedvezőtlenek). Viszont hat résztvevő PFC értéke a 2,5 cm-es cellaméretnél a legrosszabb. Összefoglalva, nincs határozott korreláció a cellaméret és a PFC sorrendje között.

Az összevont jellemző, a K érték elfedi a szélsőséges eseteket (6. ábra), például a kis PFC ellensúlyozza a rossz POD-ot. Ennek legjellemzőbb bizonyítéka, hogy négy résztvevő teljesítményének értéke elérte az elméleti maximumot 2,5 cm-es cellamérettel végzett értékelés esetén. Közülük három az 1 cm-es felosztás esetén is elérte a K = 1 értéket, annak ellenére, hogy a POD és PFC értékeik nem voltak a legjobbak. Feltételezhető, hogy esetükben az ROC görbe nem hiperbolikus. Ezt a feltételezést a kis mérési darabszám miatt nem lehet ellenőrizni.

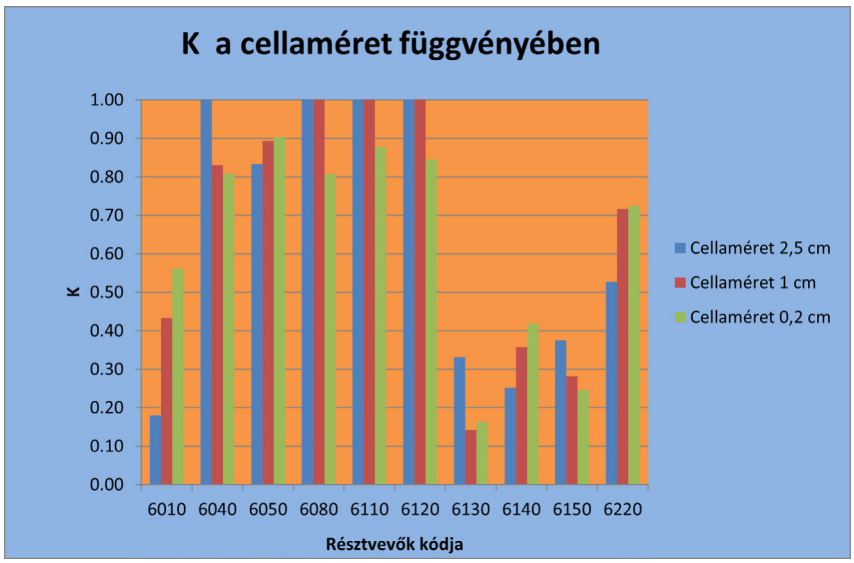
A 7. ábrán is látható, hogy nincs határozott korreláció a cellaméret és a jellemző AUC érték között.



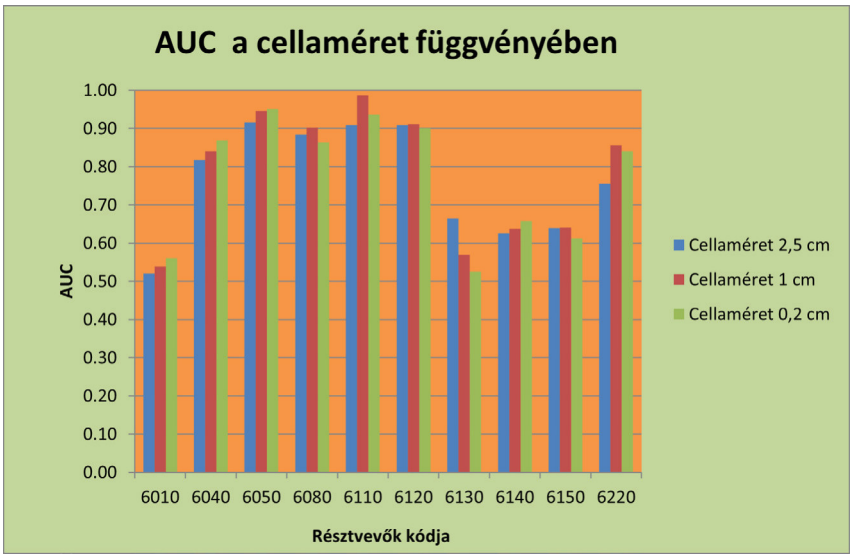
4. ábra: *Az észlelési valószínűség (POD) változása a cellaméret függvényében*



5. ábra: *A téves riasztás valószínűségének (PFC) változása a cellaméret függvényében*



6. ábra: *A vizsgáló teljesítményének (K) változása a cellaméret függvényében*



7. ábra: *A görbe alatti terület (AUC) változása a cellaméret függvényében*

Az ábrán látszik az a gyakorlati megállapítás, hogy egyik résztvevő eredménye sem éri el az elméletileg lehetséges maximális 1 értéket.

Következtetések:

- Az összevont jellemző, a K érték elfedi a szélsőséges eseteket, azaz a kis számú téves riasztás (PFC) ellensúlyozza a messze nem maximális észlelési valószínűséget (POD).
- Az észlelési valószínűség és a téves riasztás valószínűségének értéke változott a cellaméret függvényében, de nem volt tendencia a változásban. Ezért a vizsgálat tervezésénél, figyelembe véve az alkalmazott vizsgálati módszert, előre el kell dönteni az értékelés módját, beleértve az értékelésnél használatos cellaméret kiválasztását.
- A görbe alatti terület jobb jelzőszám, mint a K érték, de még nincs rá minősítési kritérium. Kipróbáljuk a következő vizsgálatoknál, addig talán találunk követelményt is.

3. A mérési bizonytalanság

A JV14 vizsgálattól kezdve minden számszerű eredményt adó mérésnél új fajta feladatot is meg kell oldani a jártassági vizsgálatok résztvevőinek. Vizsgálni fogjuk a méréseik bizonytalanságát is.

Hétköznapi szóhasználatunkban általában a mérések pontosságát emlegetjük, mint a mérések minőségi jellemzőjét. A pontosság alatt a mérési eredménynek és a mérendő mennyiség valódi értékének a közelségét értjük. A pontosság értéke azonban a valódi érték megismerhetetlensége miatt nem határozható meg. Ezért kell a mérési bizonytalanságot számítani. A feladat megoldásához először ismerjük meg a fogalmat. A Nemzetközi Metrológiai Szótár (VIM:1993) meghatározása szerint [1]:

Mérési bizonytalanság: *a mérési eredményhez társított azon paraméter, amely a mérendő mennyiségnek észszerűen tulajdonítható értékek szóródását jellemzi.*

A definíció feltételezi, hogy az adott mennyiséget többször megmérjük és a kapott egyedi értékekből számoljuk ki a mérési eredményt, majd az egyedi értékek szóródását. Mérési eredménynek rendszerint a mért egyedi értékek számtani átlagát tekintjük. A szóródást pedig az ún. standard deviációval [1], (más néven standard szórással) jellemezzük. A mérési bizonytalanságot leggyakrabban ezzel a standard szórással azonosítják, de ha nagyobb biztonságra törekszünk, akkor a standard szórás többszörösét kell vennünk.

A mérési bizonytalanság meghatározásához a Nemzetközi Súly- és Mértékügyi Hivatal először 1995-ben külön útmutatót adott ki, amit rövidítve GUM-nak nevezünk (GUM = Guide to the expression of uncertainty in measurement = Útmutató a mérési bizonytalanság kifejezéséhez). Jellemző a probléma sokoldalúságára, hogy az Útmutatót 2008-ban átdolgozták [2] és 2020-ig hat függelékkel egészítették ki. Ezeknek a kiadványoknak sincs hivatalos magyar fordítása.

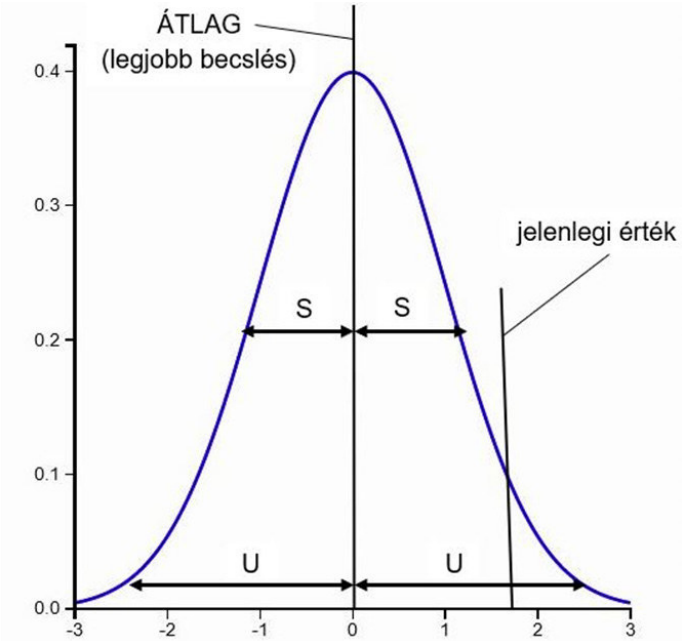
A definíció részletét, a *„mérendő mennyiségnek észszerűen tulajdonítható értékek szóródását”* különböző módon lehet értelmezni. A jártassági vizsgálataink esetében egyértelmű, hogy a mérési eredmény szórását a mérési módszer és a mért tulajdonság együttesen okozza. Ezeknek az adatoknak a szétválasztását matematikai módszerrel nem lehet korrekten megoldani, erre a GUM-ban sincs ajánlott módszer. Ezért az *„észszerűen tulajdonítható értékek szóródását”* a vizsgálatot tervezőnek, végrehajtónak és értékelőnek kell meghatároznia. Például egy merőlegesen hajlított cső, vagy belső korrózió miatt károsodott csővezeték falvastagságának mérésénél nincs értelme mérési bizonytalanságot számítani, vagy a mérések helyét pontosan ki kell jelölni. További példákat minden gyakorlati szakember tudna sorolni.

A téma iránt bővebben érdeklődőknek magyar nyelven a „Nemzeti Akkreditálási Rendszer, EA Útmutató mennyiségi vizsgálatok bizonytalanságának kifejezéséhez, NAR-EA-4/16, 1. kiadás, 2004. szeptember”

című [3] NAT áttekintő ismertetését ajánljuk.
A mérési bizonytalanságot két, alapjaiban különböző módon lehet meghatározni. Az A-típusú értékelés (Type A evaluation of measurement uncertainty): a mérési bizonytalanság egyik összetevőjének meghatározott mérési feltételek mellett elvégzett mérésekből kapott, a mért mennyiségértékek statisztikai elemzésével nyert értékelése. Az értékelés e típusánál alkalmazható az előbb említett szórásszámítás.
A mérési bizonytalanság B-típusú értékelése (Type B evaluation of measurement uncertainty): a mérési bizonytalanság összetevőjének a mérési bizonytalanság A-típusú értékelésétől eltérő módon nyert értékelése.

3.1 A mérési bizonytalanság gyakorlati értelmezése

A tapasztalatok szerint a többször megismételt mérések eredményei általában a 8. ábrán látható Gauss-eloszlást mutatják. Az eloszlást a mért átlagérték (legjobb becslés), valamint a standard deviáció (mért szórás: S) jellemzi. A mérési bizonytalanság U értelmezését a 8. ábrán mutatjuk be.
Mivel egy adott mérést nem végezhetünk el végtelen számban, a mért szórásból ki kell számítani az elvégzett mérések számától függő bizonytalanságot (U). Ez a bizonytalanság attól is függ, hogy mekkora megbízhatósági szintet (konfidenciaszint) szeretnénk elérni. A konfidenciaszint annak a valószínűsége, hogy a mérendő érték a bizonytalanság adott tartományán belül helyezkedik el. Általában megelégedünk a 95%-os konfidenciaszinttel.
Az ábrán bemutatott *jelenlegi érték* egy mérés eredményét jelképezi, aminek az értéke az S szóráson kívül esik, de a mérési bizonytalanság mértékét nem haladja meg. Fontos tulajdonsága a mérési bizonytalanságnak, hogy a mérendő mennyiségnek tulajdonított, adott mennyiségértékhez, a legjobb becslésnek nevezett átlaghoz van társítva. Ennek a mennyiségértéknek a módosítása a társított mérési bizonytalanság módosítását is eredményezi.
A VIM3 (Vocabulaire international de métrologie = Nemzetközi Mérésügyi szótár) [5] meghatározásai, melyekkel a méréseinkből meghatározhatjuk a mérési bizonytalanságot:
standard mérési bizonytalanság: *szórás formájában kifejezett mérési bizonytalanság.*
Tehát ha kiszámítjuk a méréseink szórását, akkor már egy bizonytalansági jellemzőt kapunk (S). Ha ennek az adatnak a megbízhatóságát is figyelembe szeretnénk venni, akkor kell a következő definíció a VIM3-ból:



8. ábra: *A szórás, S és a kiterjesztett mérési bizonytalanság, U viszonya*

kiterjesztett mérési bizonytalanság: *az eredő standard mérési bizonytalanság és egy 1-nél nagyobb kiterjesztési tényező szorzata.*
A szórás, S és a kiterjesztett mérési bizonytalanság, U viszonyát egy mérési sorozat eredményeinek Gauss-görbéjén a 8. ábra mutatja be.
A bizonytalanság számítása több mérési sorozat esetén (mint a jártassági vizsgálatok adatainak értékelésénél is):

$$U = \pm k\sqrt{S_1^2 + S_2^2 + ... + S_n^2}\,, \qquad (4)$$

ahol:
U = a mérések kiterjesztett bizonytalansága,
k = kiterjesztési tényező, értéke 2 és 3 között változhat, mely a megbízhatósági (konfidencia) szinttől és a mérési darabszámtól függ,
S1, S2 …. Sn = a mérési sorozatok szórásai.
A képlet szerint, ha egy munkadarabon három kijelölt helyen végzünk öt-öt mérést, és mérési helyenként kiszámítjuk a szórásokat, akkor a gyökös kifejezéssel meghatározhatjuk az eredő szórást, és ezt szorozzuk a k kiterjesztési tényezővel. A tényezőt a GUM 2008 (Table G2; 78. oldal) táblázatából [2] kereshetjük ki. A 95%-os megbízhatósági, konfidenciaszinthez (m), 15 db mérés esetén, a kiterjesztési tényező értéke k = 2,14.
Ebben az esetben a teljes mérési eredményt a következő formában kell megadni a kiterjesztett mérési bizonytalanság felhasználásával:
a teljes mérési eredmény: $x \pm U$, m% megbízhatósági szinten.

A teljes mérési eredmény, például mm-ben mért hosszúság esetén: 25 ± 0,06 mm 95%-os megbízhatósági szinten. Az eredményt úgy kell értelmezni, hogy a következő hosszúságmérések eredményei – ugyanolyan mérési körülmények esetén – 95%-os valószínűséggel a 25 mm ± 0,06 mm tartományba fognak esni.

Néhány megjegyzés a B-típusú értékeléshez, amit lehet a:

- hatóságilag közzétett mennyiségértékekhez társított,
- hiteles anyagminta mennyiségértékéhez társított,
- kalibrálási bizonyítványból kapott,
- hitelesített mérőműszer pontossági osztályából kapott, vagy a
- személyes tapasztalatok útján nyert határértékekből kapott adatokra alapozni.

A mérési bizonytalanságot úgy is meg lehet határozni, hogy számba vesszük és összegezzük a lehetséges hatásokat. Ez a B-típusú értékelés egyik formája. Az ultrahangos vastagságmérés esetében a vonatkozó szabvány C melléklete [4] összefoglalja azokat a mérési paramétereket, melyek hatással vannak a mérési bizonytalanságra. Ennek alapján is lehet mérési bizonytalanságot becsülni.

Természetesen ezt a módszert nem lehet a jártassági vizsgálatban használni, mert összehasonlíthatatlan értékeket eredményezne. Ezért a bizonytalanságot a mérési eredmények szórásából számítjuk. Ez nem korlátozza a laboratóriumokat, hogy napi vizsgálataikban a szabvány szerinti módon határozzák meg a méréseik bizonytalanságát.

3.2 A bizonytalanság kiértékelésének előnyei a vizsgáló laboratóriumok számára

A mérési bizonytalanság értékelésével kapcsolatban számos előny adódik, bár a feladat időigényes:

- A mérési bizonytalanság számszerű módon nyújt segítséget olyan fontos kérdésekben, mint a vizsgálati eredményekkel kapcsolatos kockázat és megbízhatóság.
- A mérési bizonytalanság megadása az eredmény közlése és értelmezése révén közvetlen versenyelőnyt jelenthet.
- A vizsgálati eredményeket befolyásoló egyes paraméterek hatásának (számszerű) ismerete növeli a vizsgálati eljárás megbízhatóságát.
- Eredményesebben lehet kiigazító intézkedéseket tenni, melyek ezáltal költséghatékonyabbá válnak.
- A mérési bizonytalanság kiértékelése a vizsgálati eljárás jobb megismerése révén kiindulási pontot jelent az eljárás optimalizálására.
- Az előírások teljesítésének megállapításához az ügyfelek és a termékeket minősítő testületek igénylik az eredményhez társított bizonytalanságról szóló információt.
- Csökkenthetők a kalibrációs költségek, ha a kiértékeléssel kimutathatók, hogy egyes befolyásoló mennyiségek a bizonytalanságra nincsenek jelentős hatással.

4. Összefoglalás

A MAROVISZ figyelemmel kiséri és folyamatosan fejleszti a jártassági vizsgálatok minőségét. Az észlelési vizsgálatok értékelésében mutatkozó bizonytalanságot egy valós vizsgálati eredményeket tartalmazó adathalmaz különféle módon végzett értékeléseivel hasonlítottuk össze. Megállapítottuk, hogy nincs minden fajta vizsgálat értékelésére alkalmas módszer. Az eredmények összehasonlíthatóságához szükséges a vizsgálati technológiához illesztett kiértékelési módszer.

A mérési bizonytalanság számításának bevezetése a skaláris eredmények feldolgozásában egy új szempont lesz a következő vizsgálati feladatoknál. Reméljük, a fejlesztések növelni fogják a laboratóriumok felkészültségét és előnyeit a jó eredményekkel elvégzett jártassági vizsgálatok után.

Hivatkozások:

[1] International vocabulary of basic and general terms in metrology, second edition, 1993, International Organization for Standardization (Geneva, Switzerland)

[2] JCGM 100:2008 Evaluation of measurement data – Guide to the expression of uncertainty in measurement, GUM 2008. https://www.bipm.org/en/publications/guides

[3] Nemzeti Akkreditálási Rendszer EA Útmutató mennyiségi vizsgálatok bizonytalanságának kifejezéséhez NAR-EA-4/16 1. kiadás 2004. szeptember https://docplayer.hu/216016-Nemzeti-testulet-nemzeti-akkreditalasi-rendszer-ea-utmutato-menynyisegi-vizsgalatok-bizonytalansaganak-kifejezesehez-nar-ea-4-16-1.html

[4] MSZ EN ISO16809:2019 Roncsolásmentes vizsgálat. Ultrahangos vastagságmérés (ISO 16809:2017)

[5] International vocabulary of metrology – Basic and general concepts and associated terms 3rd edition, VIM3: www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf