

Leeswijzer meetonzekerheid geaccrediteerde verrichtingen

Het Interlaboratorium Vergelijkend Onderzoek.

Tussen de geaccrediteerde laboratoria binnen de sector bouw worden regelmatig ringonderzoeken gehouden, het zogenoemde InterLaboratorium Vergelijkend Onderzoek (ILVO). Met behulp van de ringonderzoeken worden de standaardafwijking van de herhaalbaarheid (S_r) en de reproduceerbaarheid (S_R) bepaald. Met behulp van deze grootheden is het dan weer mogelijk om waarden voor de herhaalbaarheid (r) en de reproduceerbaarheid (R) te bepalen:

- Herhaalbaarheid (r) = $2\sqrt{2}$ x de standaardafwijking van r ($2\sqrt{2} \cdot S_r$);
- Reproduceerbaarheid (R) = $2\sqrt{2}$ x de standaardafwijking van R ($2\sqrt{2} \cdot S_R$).

In sommige normen worden eisen gesteld aan de herhaalbaarheid en de [reproduceerbaarheid](#). De leeswijzer geeft geen informatie omtrent de herhaalbaarheid. De reproduceerbaarheid is vergeleken met de waarden uit de vigerende norm. Bij het niet voldoen aan de norm worden vanuit de sectorcommissie bouw acties uitgezet. De reproduceerbaarheden zijn vastgesteld in december 2015 en blijven gehandhaafd zolang zij binnen het 95% betrouwbaarheidsinterval liggen. De in deze leeswijzer gepresenteerde reproduceerbaarheden voor de verrichtingen zijn berekend uit de resultaten van de gedurende de laatste 12 à 13 ringonderzoeken verzamelde data.

De meetonzekerheid wordt berekend als:

- Meetonzekerheid = 2 x de standaardafwijking van S_R ($2 \cdot S_R$).

De geaccrediteerde laboratoria hebben besloten om per verrichting een gemeenschappelijke meetonzekerheid op te geven. [De meetonzekerheden en marges](#) zijn vastgesteld op 12 april 2007 en blijven gehandhaafd zolang zij binnen het 95% betrouwbaarheidsinterval liggen. De in deze leeswijzer gepresenteerde meetonzekerheden voor de verrichtingen zijn berekend uit de resultaten van de gedurende de laatste 12 à 13 ringonderzoeken verzamelde data. De meetonzekerheid wordt bepaald onder zogenoemde "reproduceerbaarheidomstandigheden", hetgeen betekent dat de genoemde onzekerheid ontstaat wanneer meerdere laboratoria hetzelfde monster onderzoeken. Vanwege het destructieve karakter van de meeste proeven is het niet mogelijk om hetzelfde monster te onderzoeken, waardoor er strenge eisen aan de samenstelling van de monsters worden gesteld. De meetonzekerheden gelden uitsluitend voor verrichtingen die volgens de vermelde norm zijn uitgevoerd door voor deze verrichting geaccrediteerde laboratoria. De meetonzekerheden worden niet standaard bij de resultaten gerapporteerd, maar zijn op aanvraag door het betreffende laboratorium leverbaar.

Waarvoor gebruik je de meetonzekerheid?

Een meting geeft een schatting van de werkelijkheid. De meetonzekerheid geeft aan wat de betrouwbaarheid van deze geschatte werkelijkheid is. Indien een willekeurig geaccrediteerd laboratorium een schatting van de werkelijkheid geeft, zal deze met een betrouwbaarheid van 95% liggen in het interval "meetresultaat \pm meetonzekerheid".

De meetonzekerheid geeft dus aan binnen welke grenzen de werkelijke waarde zal kunnen liggen. Deze grenzen kunnen enger gemaakt worden door het betreffende laboratorium meer metingen op het monster te laten verrichten.

Waarvoor mag je de meetonzekerheid niet gebruiken?

De meetonzekerheid mag niet toegepast worden als een marge op de eisen, zoals die bijvoorbeeld in de "Standaard RAW Bepalingen" zijn opgenomen. Deze eisen zijn empirisch bepaald en gelden voor de meetwaarde zelf.

Leeswijzer meetonzekerheid geaccrediteerde verrichtingen

Opzet ILVO:

Per jaar wordt 2 tot 3 maal een pakket monsters verstuurd naar momenteel 16 laboratoria. Voor elk pakket moet een aantal variabelen worden gemeten. Slechts één van alle laboratoria voert alle metingen uit. Daarom zijn de variabelen ingedeeld in clusters, waarin in principe steeds dezelfde laboratoria meedoen. Een deelnemend laboratorium hoeft echter niet aan alle verrichtingen binnen een cluster mee te doen. De frequentie kan per cluster verschillen en ook binnen een cluster kan de frequentie verschillen. Een en ander is afhankelijk van de resultaten binnen het ringonderzoek.

Cluster	Aantal maal per jaar	Aantal Laboratoria
1	3	5 tot 7
5	2 tot 3	6
8	3	3 tot 8
9	3	2 tot 8
10	3	4 tot 6
11	2	2 tot 4
12	3	4 tot 5
13	3	2 tot 3
14	3	2 tot 3
15	3	2 tot 11

De pakketten bestaan uit monsters die zeer zorgvuldig zijn samengesteld. Zo zorgvuldig dat de berekende varianties kunnen worden toegeschreven aan de uitvoering van de methode en niet aan materiaalspreiding (ILVO 192). Voor elke variabele wordt één meting uitgevoerd. Per pakket en per laboratorium is dus slechts één meetresultaat beschikbaar. Bovendien doen ook binnen een cluster de laboratoria niet aan alle verrichtingen mee. Dit heeft tot gevolg dat:

- Het niet mogelijk is om per laboratorium afzonderlijk de standaardonzekerheid te berekenen;
- Het schema van metingen ontbrekende waarnemingen bevat;
- De standaardonzekerheden per laboratorium alleen kunnen worden berekend met behulp van alle meetresultaten uit het ringonderzoek. Voor deze berekening is gebruik gemaakt van een variantieanalyse-methode die rekening houdt met de ongelijke aantallen metingen per laboratorium ("restricted maximum likelihood"-methode (REML)).

ILVO-rapportage:

Zoals gezegd wordt de meetonzekerheid bepaald aan de hand van de ringonderzoeken. Aan de hand van de laatste 13 ringonderzoeken in [cluster 3](#) zal de gehanteerde ILVO-rapportage worden uitgelegd.

De tabel "metingen" op pagina 2 geeft van elk deelnemend laboratorium per ringonderzoek de meetresultaten weer. In de laatste kolom wordt het gemiddelde resultaat per pakket weergegeven. In de laatste rij wordt het gemiddelde resultaat per laboratorium weergegeven.

De tabel "verschillen t.o.v. de pakketgemiddelden" geeft van elk deelnemend laboratorium per ringonderzoek het verschil ten opzichte van het gemiddelde pakketresultaat (laatste kolom tabel "metingen") weer. In de laatste kolom wordt per pakket de gemiddelde afwijking ten opzichte van het pakketgemiddelde weergegeven. In de laatste rij wordt per laboratorium de gemiddelde afwijking ten opzichte van het pakketgemiddelde weergegeven.

De laatste tabel geeft informatie over resultaten die niet in de analyse zijn betrokken.

De derde pagina bestaat uit twee delen, namelijk een analyse van de meetmethode (bovenste gedeelte) en een analyse van de laboratoria (grafieken).

Op pagina 3 wordt in de eerste regel onder de kop vermeld dat bij de analyse 6 laboratoria zijn betrokken, dat er 13 verschillende pakketten zijn onderzocht en dat er in totaal 64 metingen zijn geanalyseerd.

Leeswijzer meetonzekerheid geaccrediteerde verrichtingen

Analyse van de meetmethode.

Om te beginnen wordt de methode geanalyseerd onder herhaalbaarheidsomstandigheden, dat wil zeggen dat de genoemde onzekerheid geldt binnen een laboratorium, in een klein tijdsbestek, bij één laborant en één apparaat:

- S_r = standaardafwijking van de herhaalbaarheid. Deze waarde wordt ook wel de standaardonzekerheid (u) genoemd.
- $2S_r$ = uitgebreide onzekerheid (U).
Met een betrouwbaarheid van 95% kan worden gesteld dat de werkelijke waarde zal liggen in het interval "meetresultaat $\pm U$ ". Voor deze waarde wordt tevens het 95% betrouwbaarheidsinterval vermeld.
- $r = 2\sqrt{2} \times S_r$ = herhaalbaarheid, ook wel herhaalbaarheidsgrens genoemd.
Met een betrouwbaarheid van 95% kan worden gesteld dat het absolute verschil van twee meetresultaten kleiner zal zijn dan r .

Verder wordt de methode geanalyseerd onder reproduceerbaarheidsomstandigheden, dat wil zeggen dat de genoemde onzekerheid ontstaat wanneer meerdere laboratoria hetzelfde monster onderzoeken.

- S_R = standaardafwijking van de reproduceerbaarheid. Deze waarde wordt ook wel de standaardonzekerheid (u_c) genoemd.
- $2S_R$ = uitgebreide onzekerheid (U) of [meetonzekerheid](#).
Met een betrouwbaarheid van 95% kan worden gesteld dat de werkelijke waarde zal liggen in het interval "meetresultaat $\pm U$ ". Voor deze waarde wordt tevens het 95% betrouwbaarheidsinterval vermeld.
- $R = 2\sqrt{2} \times S_R$ = reproduceerbaarheid, ook wel reproduceerbaarheidsgrens genoemd.
Met een betrouwbaarheid van 95% kan worden gesteld dat het absolute verschil van twee meetresultaten kleiner zal zijn dan R . Voor deze waarde wordt tevens het 95% betrouwbaarheidsinterval vermeld.

Daarnaast worden de volgende grootheden onderscheiden:

- $\gamma = S_R/S_r$
Deze grootheid geeft de verhouding tussen de standaardafwijking van de Reproduceerbaarheid en de standaardafwijking van de Herhaalbaarheid weer.
Hoe dichter deze waarde in de buurt van 1 ligt, hoe kleiner de systematische verschillen tussen de laboratoria zijn. [\[3\]](#)
- Marge (M) = $1,645S_R$
De marge wordt gebruikt om te toetsen aan een grens. Met 95% betrouwbaarheid kan worden gesteld dat de werkelijke waarde kleiner zal zijn dan "meetresultaat + M " of dat de werkelijke waarde groter zal zijn dan "meetresultaat - M ".

Analyse per laboratorium

Per laboratorium worden de grootheden k en h van de heer dr. J. Mandel berekend, waarmee de prestaties van een laboratorium in termen van de herhaalbaarheidsstandaardafwijking en gemiddelde worden vergeleken met de andere laboratoria met betrekking tot de binnen-lab spreiding en het lab-niveau. [\[3, 4\]](#)
Ook in de norm ISO 5725-2 (1994) wordt voor deze toetsingen de methode van Mandel's k en Mandel's h beschreven.

Mandel's k is de verhouding van de standaarddeviatie van de binnen-lab spreiding voor een specifiek lab ten opzichte van de overall S_r .

In de rapportage van de ILVO kalibratieronden wordt Mandel's k per lab in een staafdiagram weergegeven. Daarnaast wordt getoetst of deze waarde statistisch significant groter is dan 1, wat aangeeft of een lab een relatief grote binnen-lab spreiding heeft. Dit komt in de grafiek tot uiting door het overschrijden van de grenswaarden die met een gestippelde lijn zijn aangegeven.

Bij de berekening van Mandel's k wordt voor een specifiek lab i de voor dat lab geldende binnen-lab standaardafwijking geschat met een waarde s_i . Dit gebeurt op basis van de afwijkingen van het betreffende lab ten opzichte van het

Leeswijzer meetonzekerheid geaccrediteerde verrichtingen

pakketgemiddelde. De waarde s_i wordt dan gedeeld door de overall s_r zodat dus voor dat lab de waarde van Mandel's k is: $k_i = s_i/s_r$.

De grenswaarde wordt aangegeven met krit- k (zie de tabel onder het staafdiagram). Deze waarde kan per lab verschillend zijn, omdat deze wordt berekend op basis van het aantal vrijheidsgraden die bij de betreffende binnen-lab spreiding s_i horen. Bij meer vrijheidsgraden neemt de zekerheid toe en de kritieke waarde krit- k dus af. Het aantal vrijheidsgraden neemt toe naarmate er meer pakketten door het betreffende lab zijn gemeten en ook naarmate de s_i groter is.

Opmerking

Soms is de binnen-lab spreiding voor een lab zo klein dat deze in de statistische analyse als een negatief getal wordt geschat. In dat geval wordt een * afgedrukt in de rapportage. Dit moet dus worden gelezen als "zeer klein" en dus in elk geval kleiner dan de grenswaarde.

Mandel's h is het verschil tussen het gemiddelde meetniveau van een specifiek lab ten opzichte van het overall gemiddelde.

In de rapportage van de ILVO kalibratieronden wordt Mandel's h per lab in een staafdiagram weergegeven. Daarnaast wordt getoetst of deze waarde statistisch significant afwijkt van 0, wat aangeeft of een lab systematisch te laag dan wel te hoog meet. Dit komt in de grafiek tot uiting door het overschrijden van de grenswaarden die met twee gestippelde lijnen zijn aangegeven.

De berekening van Mandel's h voor een specifiek lab i gaat als volgt. Uit de statistische analyse wordt voor het betreffende lab zijn overall gemiddelde M_i geschat. Als alle labs alle pakketten hebben gemeten, zou dit simpelweg het gemiddelde van alle meetwaarden van het betreffende lab zijn. In de praktijk zijn er vaak missende waarden in de tabel (soms door het weglaten van uitbijters). Dan wordt daar in de statistische analyse voor gecorrigeerd. Van de M_i 's per lab wordt het gemiddelde MM en de standaardafwijking s_M berekend.

Vervolgens wordt de waarde van Mandel's h voor dat lab berekend via
$$h_i = \frac{M_i - MM}{s_M} .$$

De grenswaarde wordt aangegeven met krit- h (zie de tabel onder het staafdiagram). Deze grenswaarde is voor ieder lab gelijk, omdat deze wordt berekend op basis van het aantal vrijheidsgraden die bij de schatting van s_M horen en dit hangt alleen af van het aantal labs dat de betreffende variabele gemeten heeft. Bij meer deelnemende labs neemt de zekerheid toe en de kritieke waarde krit- h dus af.

Opmerking

In het algemeen is de ware waarde van de gemeten variabele onbekend. Het is gebruikelijk om dan het gemiddelde MM over alle labs als de waarheid te benoemen. Dat betekent dat het gemiddelde van de h_i 's over de labs gelijk is aan 0. In een beperkt aantal situaties, zoals bijvoorbeeld variabele 1 van cluster 1 (bitumengehalte), is de ware waarde wel bekend. In dat geval kan als grootheid de afwijking van de meetwaarde ten opzichte van de ware waarde worden

$$h_i = \frac{M_i}{s_M} .$$

geanalyseerd. Voor deze grootheid wordt de waarde van Mandel's h berekend als
De grenswaarde krit- h blijft daarbij ongewijzigd.

Voor gedetailleerde informatie over de toegepaste statistiek wordt verwezen naar het document "Toelichting statistiek ILVO ringonderzoek".

http://www.fenelab.nl/_wpx/wp-content/uploads/2018/11/Toelichting-statistiek-ILVO-ringonderzoek.pdf