

onderwerp Validatie van (nieuwe) parameters en/of matrices	code KHRO-6.0	bladzijde 1 van 8	Datum 2003-08-15
	Revisie 7	vastgesteld door Hoofd LMC	paraaf

6.0 Validatie van (nieuwe) parameters en/of matrices

VERANTWOORDELIJKHEID

De ringonderzoekcoördinator is verantwoordelijk voor de opzet van het validatie-onderzoek. De labhoofden zijn verantwoordelijk voor de bereiding, meting van monsters en de rapportage aan de ringonderzoekcoördinator van gegevens ten behoeve van de validatie van een ringonderzoek.

Ook de verantwoordelijkheid van uitbesteding van de monsters ten behoeve van een validatie-onderzoek ligt bij de labhoofden. De ringonderzoekcoördinator is verantwoordelijk voor de uitvoering van de benodigde berekeningen en raadpleegt zonodig de statisticus. Het hoofd LMC keurt het geheel uiteindelijk goed. Uitzondering op deze regel is de validatie van een microbiologisch ringonderzoek. Deze ringonderzoeken en bepalingsmethoden verschillen zodanig van de chemische dat er een aparte validatieprocedure voor is ontwikkeld. Deze is opgenomen in KHRO-4.3.5.. Het hoofd van het Laboratorium voor Microbiologie (LMB) is verantwoordelijk voor de opzet en uitvoering van het validatieonderzoek voor microbiologische parameters.

BETREFT

Een beschrijving van de procedure voor het uitvoeren van een validatie-onderzoek voor chemische parameters in een ringonderzoek. Hierin worden de homogeniteit en de stabiliteit van de bereide monsters onderzocht.

DOEL

Het waarborgen van onderling identieke en voldoende stabiele ringonderzoekmonsters.

Met betrekking tot stabiliteit: het kan voorkomen dat een monster zeer kort of niet stabiel is (bijvoorbeeld CO₂). In deze gevallen wordt geadviseerd de monsters 24 uur na aankomst bij de deelnemer, maar bij voorkeur op dezelfde dag, te meten. Als deze instructie wordt opgevolgd levert een verlopende stabiliteit een relatief geringe bijdrage aan de spreiding in het groepsresultaat.

UITVOERING

Aanleiding

De behoefte van de deelnemers aan ringonderzoeken ontwikkelt zich. Dit omdat er in toenemende mate vraag is naar passende ringonderzoeken en omdat het analysepakket van de deelnemers zich wijzigt. Deelnemers aan de Kiwa-ringonderzoeken kunnen hun wensen met betrekking tot nieuwe parameters en/of nieuwe matrices (bijvoorbeeld andere watertypen) kenbaar maken bij de inventarisatie voor het nieuwe jaarprogramma (zie KHRO-4.1), dit inbrengen bij besprekingen van de rapportages (zie KHRO-4.3.10) of de help-desk (assistent ringonderzoeken) of de ringonderzoekcoördinator of het hoofd van LMC hierop attenderen. Ook kan het zijn dat Kiwa, op basis van haar inzicht in de markt, zelf van mening is dat het zinvol is haar pakket uit te breiden.

Ongeacht de aanleiding volgt eventueel een (beperkt) marktonderzoek. Kiwa moet de overtuiging hebben dat er in elk geval minimaal acht belangstellende laboratoria zijn, alvorens een onderzoek naar de praktische uitvoerbaarheid gestart wordt. Indien dat het geval is, wordt dat onderzoek uitgevoerd.

Validatieonderzoek

De ringonderzoekcoördinator overlegt met het betreffende laboratoriumhoofd over de mogelijkheden. Zij besteden onder andere aandacht aan de wijze van bereiden van de monsters (o.a. homogeniteit), de conserveringswijze (stabiliteit), de controle van de monsters (door Kiwa te doen, of uit te besteden; toetsingscriteria). Zo nodig verricht het laboratoriumhoofd literatuuronderzoek, of verzamelt op andere wijze relevante gegevens. Tevens gaat hij de mogelijkheden na om gebruik te maken van gecertificeerde referentiematerialen. In het laboratorium wordt zonodig oriënterend praktisch werk gedaan. Vervolgens stelt het laboratoriumhoofd in samenwerking met de ringonderzoekcoördinator een validatieplan op. In het validatieplan wordt in elk geval vastgelegd hoe de homogeniteit en stabiliteit getest gaat worden. Uiteraard kunnen voor de validatie eerder verzamelde gegevens

onderwerp
**Validatie van (nieuwe) parameters
en/of matrices**

code
KHRO-6.0

bladzijde
2 van 8

Datum
2003-08-15

benut worden. Voor het onderzoek naar de stabiliteit worden monsters bereid, die direct na bereiding, verdeeld over twee aansluitende dagen worden geanalyseerd. Deze analyses worden, opnieuw verdeeld over twee aansluitende dagen, herhaald aan het einde van de periode waarin de monsters stabiel geacht worden te zijn. Hiermee worden effecten van de binnen-laboratorium reproduceerbaarheid ondervangen. Voor het onderzoek naar de homogeniteit wordt een vergelijkbaar aantal monsters geanalyseerd, maar nu in een zo kort mogelijk tijdsbestek. Uiteraard kunnen de onderzoeken naar de homogeniteit en stabiliteit deels met elkaar gecombineerd worden. De spreiding in deze resultaten wordt vergeleken met de standaardafwijking van de binnen-laboratorium reproduceerbaarheid en indien bekend, met de standaardafwijking van de tussen-laboratorium reproduceerbaarheid (Thompson en Wood, 1993). Over het validatieplan kan overleg gepleegd worden met het hoofd van LMC en/of met de statisticus. Het plan wordt door het hoofd LMC vastgesteld.

Algemeen:

De validatie wordt doorgaans uitgevoerd in de moeilijkste matrix. Tot op heden is dat oppervlaktewater. De monsters ten behoeve van het validatie-onderzoek worden in principe altijd bij Kiwa bereid. Alleen de analyse ervan kan worden uitbesteed. Een uitgebreide uitleg over de gehanteerde berekeningen en formules staan in de syllabus toegepaste statistiek voor drinkwaterlaboratoria, SWE 96.001.

Homogeniteitstest:

Aan de hand van de s_R van de binnenlaboratoriumreproduceerbaarheid (relatieve standaarddeviatie in %, RSD) in de bepaling bij Kiwa van de additiemonsters of die van het laboratorium waaraan het onderzoek wordt uitbesteed, wordt het aantal monsters bepaald waarmee de homogeniteit wordt getest. De concentratie van deze monsters moet in dezelfde orde van grootte zijn als de aangeboden concentraties in ringonderzoeken.

Om het aantal monsters voor de homogeniteitstest te bepalen, wordt de volgende regel gehanteerd:

- $RSD \leq 5\%$, 6 monsters voor homogeniteit,
- $5 > RSD \leq 10\%$, 8 monsters voor homogeniteit,
- $RSD > 10\%$, 10 monsters voor homogeniteit.

Als er geen gegevens bekend zijn (b.v. nieuwe parameter) dient er een reële schatting gemaakt te worden op basis van soortgelijke parameters of gegevens uit de literatuur.

Fase 1:

Deze monsters worden allen op 1 dag, in 1 meetserie door 1 analist gemeten. Uit de waarden wordt het gemiddelde en de spreiding berekend. De ringonderzoekcoördinator onderzoekt de resultaten en besluit naar aanleiding daarvan tot verwijdering van uitschieters. Hierbij kan een beroep gedaan worden op de expertise van de statisticus. Wanneer tot statistisch gefundeerde uitschieterverwijdering wordt overgegaan, wordt dat met behulp van de Veglia-toets uitgevoerd. Deze toets is niet afhankelijk van het type kansverdeling van de resultaten. Deze is namelijk moeilijk vast te stellen omdat het hier altijd om een beperkt aantal resultaten gaat (maximaal 10). Vervolgens wordt de gemiddelde spreiding van het groepsresultaat per Youdenpaar berekend uit minstens 2 ringonderzoeken (s_R). Deze kentallen zijn te vinden in de betreffende rapporten onder het hoofdstuk 'rapportcijfers'. Deze gemiddelde spreiding wordt gedeeld door de spreiding berekend uit de homogeniteitstest (s_r). Hieruit moet een factor 3,3 of hoger komen. De monsters mogen nu als homogeen beschouwd worden (Zie Thompson and Wood, 1993). De voorwaarde om monsters voor één bepaalde parameter homogeen te verklaren staan nogmaals weergegeven in de onderstaande formule.

$$\frac{s_R}{s_r} \geq 3,3$$

Bij veel organische bepalingen omvat één analyse meerdere soortgelijke componenten. Zo'n bepaling wordt een multicomponentbepaling genoemd. Bij de bereiding van een ringonderzoek worden deze componenten

onderwerp

Validatie van (nieuwe) parameters en/of matrices

code

KHRO-6.0

bladzijde

3 van 8

Datum

2003-08-15

gezamenlijk aan het monster toegevoegd. Omdat alle componenten uit één multicomponentbepaling in één ringonderzoek exact dezelfde behandeling ondergaan, wordt voor het testen van de homogeniteit een aangepast criterium gebruikt. Van de parameters uit de multicomponentbepaling die in één specifiek ringonderzoek worden toegevoegd aan de monsters, wordt de factor $s_r:s_r$ bepaald. Hiervan wordt vervolgens een gemiddelde factor berekend. Deze gemiddelde factor moet $>3,3$ zijn én 50% van de afzonderlijk berekende factoren dienen $>3,3$ te zijn.

Wanneer niet aan bovenstaande criteria wordt voldaan volgt nader onderzoek naar de homogeniteit zoals beschreven onder fase 2.

Fase 2;

In deze fase wordt een absoluut criterium aangehouden, gebaseerd op criteria voor de precisie voor analysemethoden zoals opgenomen in het Waterleidingbesluit. In het Waterleidingbesluit zijn echter alleen eisen opgenomen voor de matrix drinkwater, terwijl Kiwa de homogeniteitstesten doorgaans uitvoert in oppervlaktewater (de moeilijkste matrix). Omdat dit de eisen juist scherper zal stellen, kan van deze richtlijn uit worden gegaan. Er wordt onderscheid gemaakt tussen anorganische en organische parameters.

- Anorganische parameters

Voor anorganische parameters is het criterium voor de precisie van een analysemethode in het Waterleidingbesluit doorgaans 5 tot 20%. Dit betekent dat de standaardafwijking van de herhaalbaarheid 2.5 tot 10% mag zijn. Op basis van het rapport SWE-98.003 "Validatie van bepalingmethoden voor anorganische parameters", betekent dit dat de standaardafwijking van de reproduceerbaarheid 2 tot 3 maal de standaardafwijking van de herhaalbaarheid mag bedragen, in dit geval 5 tot 30%. Op basis van praktijkervaring wordt het criterium voor s_R op 5% gesteld.

Dit houdt in dat wanneer het gestelde criterium uit fase 1 wordt overschreden maar $s_R \leq 5\%$ dat het niet relevant is om homogeniteit voor anorganische parameters aan te tonen. Echter, wanneer $s_R > 5\%$ dan dient wederom nader onderzoek uitgevoerd te worden zoals omschreven onder fase 3.

- Organische parameters

Voor organische parameters is het criterium voor de precisie van een analysemethode in het Waterleidingbesluit doorgaans 50%. Dit betekent dit dat de standaardafwijking van de herhaalbaarheid 25% mag zijn. Op basis van het rapport SWE-96.016 "Validatie van bepalingmethoden voor organische microverontreinigingen", betekent dit dat de standaardafwijking van de reproduceerbaarheid 2 tot 3 maal de standaardafwijking van de herhaalbaarheid mag bedragen, in dit geval 25 tot 75%. Op basis van praktijkervaring wordt het criterium voor s_R op 10% gesteld.

Dit houdt in dat wanneer het gestelde criterium uit fase 1 wordt overschreden maar $s_R \leq 10\%$ dat het niet relevant is om homogeniteit voor organische parameters aan te tonen. Echter, wanneer $s_R > 10\%$ dan dient wederom nader onderzoek uitgevoerd te worden zoals omschreven onder fase 3.

Fase 3

Er wordt 1 monster bereid dat in meervoud gemeten wordt. Het aantal metingen is weer gerelateerd aan de RSD. De spreiding die dan gemeten wordt, is de spreiding uit de analyse (s_r) en zou vergelijkbaar moeten zijn aan de spreiding die veroorzaakt werd door de analyse in de homogeniteitstoets. Als de beide spreidingen niet significant van elkaar afwijken worden de monsters alsnog homogeen verklaard. Dit wordt berekend met de f-toets.

De nulhypothese (H_0) luidt:

$$H_0 : s_h = s_r$$

onderwerp
**Validatie van (nieuwe) parameters
en/of matrices**

code
KHRO-6.0

bladzijde
4 van 8

Datum
2003-08-15

De alternatieve hypothese luidt:

$$H_a: s_h > s_r \quad (3)$$

De spreiding om de f-toets uit te voeren wordt berekend volgens:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (4)$$

Waarin n het aantal waarnemingen, x_i de meetwaarde en \bar{x} het gemiddelde.

De toetsingsgrootte f wordt berekend volgens:

$$f = \frac{s_h^2}{s_r^2} \quad (5)$$

Er wordt in dit geval eenzijdig getoetst. Als de nulhypothese wordt aangenomen zijn de monsters homogeen. We verwerpen de nulhypothese als geldt:

$$f > F_{(\alpha=5\%, vgt, vgn)} \quad (6)$$

met $F_{(\alpha=5\%, vgt, vgn)}$ de waarde van de F-verdeling met een eenzijdige overschrijdingskans van 5%, bij vgt vrijheidsgraden voor de teller en vgn vrijheidsgraden voor de noemer.

In dit geval zijn de spreidingen dus niet vergelijkbaar en de monsters niet homogeen.

Wanneer na het doorlopen van de drie fasen van het homogeniteitsonderzoek niet kan worden aangetoond dat een monster homogeen is, zal de parameter niet meer aangeboden worden tot de oorzaak van de inhomogeniteit bekend en opgelost is.

Houdbaarheid:

Ook hier wordt het aantal te meten monsters bepaald door de spreiding (relatieve standaard deviatie in %, RSD) in de betreffende analyse bij Kiwa of het laboratorium waar het validatie-onderzoek wordt uitbesteed.

Fase 1

De regel is:

- $RSD \leq 5\%$, 6 monsters voor het houdbaarheidsexperiment, verdeeld over twee opeenvolgende dagen,
- $5 > RSD \leq 10\%$, 8 monsters voor het houdbaarheidsexperiment, verdeeld over twee opeenvolgende dagen,
- $RSD > 10\%$, 10 monsters voor het houdbaarheidsexperiment, verdeeld over twee opeenvolgende dagen.

Gegevens over de RSD worden verkregen uit de Shewartkaartgegevens voor de additiemonsters (binnenlaboratoriumspreiding).

Dit experiment wordt aan het begin en aan het eind van de periode uitgevoerd, waarin de monsters houdbaar

onderwerp

**Validatie van (nieuwe) parameters
en/of matrices**

code

KHRO-6.0

bladzijde

5 van 8

Datum

2003-08-15

dienen te zijn. Hierin wordt rekening gehouden met de bereidingstijd op Kiwa, de analysetijd die deelnemers nodig hebben en de tijd die het vervoeren van de monsters in beslag neemt. Ook kan informatie verkregen worden uit literatuur (div. normen, de praktijkrichtlijn monsterneming Kiwa).

Het concentratieniveau moet representatief zijn voor ringonderzoekmonsters.

Door de monsters aan het begin en het eind van de testperiode te verdelen over twee opeenvolgende dagen te meten wordt de binnenlabspreiding verdisconteerd. Hierna wordt bezien of de spreidingen vergelijkbaar zijn door de *t*-toets toe te passen. Wanneer dit zo is wordt de *t*-toets gebruikt om de gemiddelden te vergelijken, zo niet wordt de aangepaste *t*-toets gebruikt.

De gemiddelden die berekend worden, worden middels één van beide *t*-toetsen met elkaar vergeleken om te bezien of ze al dan niet significant van elkaar verschillen. Hierbij moet rekening gehouden worden met het feit of er één of tweezijdig getoetst dient te worden. Als bij een parameter alleen een afname of alleen een toename wordt verwacht betekent dat dat er éénzijdig getoetst wordt. Als het niet bekend is of het om een toe- of afname gaat (of beide is mogelijk) dient er tweezijdig getoetst te worden.

De nulhypothese voor het vergelijken van de spreiding (H_0) luidt:

$$H_0 : s_b = s_e \quad (7)$$

Waarin s_b de spreiding van de waarden gemeten in de beginperiode, en s_e de spreiding van de waarden gemeten in de eindperiode.

De alternatieve hypothese voor het vergelijken van de spreiding luidt:

$$H_a : s_b \neq s_e \quad (8)$$

De spreiding om de *f*-toets uit te voeren wordt berekend volgens:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (9)$$

De toetsingsgrootte *f* wordt berekend volgens:

$$f = \frac{s_b^2}{s_e^2} \text{ of } \frac{s_e^2}{s_b^2} \quad (f \geq 1) \quad (10)$$

Er wordt in dit geval tweezijdig getoetst.

We verwerpen de nulhypothese als geldt:

$$f > F_{(a2=5\%, vgt, vgn)} \quad (11)$$

met $F_{(a2=5\%, vgt, vgn)}$ de waarde van de *F*-verdeling met een tweezijdige overschijdingskans van 5%, bij *vgt* vrijheidsgraden voor de teller en *vgn* vrijheidsgraden voor de noemer.

Als de nulhypothese wordt aangenomen wordt de *t*-toets gebruikt. Als de alternatieve hypothese geldt, dan wordt de aangepaste *t*-toets gebruikt.

onderwerp
**Validatie van (nieuwe) parameters
en/of matrices**

code
KHRO-6.0

bladzijde
6 van 8

Datum
2003-08-15

De nulhypothese voor het vergelijken van de gemiddelden (H_0) luidt:

$$H_0: \bar{x}_b = \bar{x}_e \quad (12)$$

De alternatieve hypothese voor het vergelijken van de gemiddelden kan luiden:

$$H_a: \bar{x}_b \neq \bar{x}_e \quad (13)$$

Of, bij eenzijdig toetsen

$$\bar{x}_b > \bar{x}_e \text{ of } \bar{x}_e > \bar{x}_b \quad (14)$$

De toetsingsgrootheid (T) is als volgt gedefinieerd:

$$T = \frac{\bar{x}_b - \bar{x}_e}{s_c \sqrt{\frac{1}{n_b} + \frac{1}{n_e}}} \quad (15)$$

met s_c de gecombineerde schatting van de standaardafwijking van de waarnemingen (we mochten er immers van uitgaan dat de standaardafwijkingen van beide populaties gelijk zijn), berekend als:

$$s_c = \sqrt{\frac{(n_b - 1) \cdot s_b^2 + (n_e - 1) \cdot s_e^2}{n_b + n_e - 2}} \quad (16)$$

Als uit de f-toets is gebleken dat de standaardafwijkingen van beide uitkomsten-populaties *niet* gelijk zijn, moet - in afwijking van de t-toets - de toetsingsgrootheid (T) nu als volgt worden berekend:

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad (17)$$

(er is nu immers geen grond om de twee geschatte standaardafwijkingen te combineren tot één schatting). Onder de nulhypothese volgt T een Student-t verdeling waarvan het gemiddelde nul bedraagt. We verwerpen de nulhypothese bij tweezijdig toetsen als:

$$|T| > t_{(a/2=5\%,vg)} \quad (18)$$

met $t_{(a/2=5\%,vg)}$ de Student-t waarde met een tweezijdige overschrijdingskans van 5%, die geldt bij vg vrijheidsgraden. Bij enkelzijdig toetsen wordt de nulhypothese verworpen als:

$$|T| > t_{(a=5\%,vg)} \quad (19)$$

met $t_{(a=5\%,vg)}$ de Student-t waarde met een eenzijdige overschrijdingskans van 5%, die geldt bij vg vrijheidsgraden. Als de nulhypothese wordt aangenomen zijn de monsters stabiel, als de alternatieve hypothese wordt aangenomen zijn de monsters niet stabiel gedurende de geteste periode.

onderwerp
**Validatie van (nieuwe) parameters
en/of matrices**

code
KHRO-6.0

bladzijde
7 van 8

Datum
2003-08-15

Het aantal vrijheidsgraden is nu te berekenen als:

$$vg = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} \quad (20)$$

Wanneer niet aan bovenstaande criteria wordt voldaan, volgt nader onderzoek naar de stabiliteit zoals beschreven onder fase 2.

Fase 2

In deze fase wordt een criterium aangehouden gebaseerd op de reproduceerbaarheid van de bepaling. Het gemiddelde van de opbrengst van dag 0 en dag 1 wordt vergeleken met het gemiddelde van de opbrengst van de geachte stabiele periode. Als de opbrengst binnen de 1S-grens van de shewhartkaart valt van de desbetreffende parameter dan mogen we aannemen dat de monstres stabiel zijn over de geteste periode.

Als na het doorlopen van de stabiliteitstoets blijkt dat een parameter niet stabiel is over de geteste periode dan zal in eerste instantie een nieuw stabiliteitsexperiment over een kortere periode worden uitgevoerd. Mocht de stabiliteit van een parameter nog niet aangetoond zijn voordat een ringonderzoek wordt uitgegeven, dan zal geen theoretische waarde worden afgegeven en zal aan de deelnemers worden gevraagd deze parameter binnen 24 uur te analyseren.

De validatieonderzoeken voor stabiliteit worden steekproefsgewijs herhaald. De keuze van de ringonderzoeken zullen zo veel mogelijk representatief zijn voor het gehele jaarprogramma. De ringonderzoeken die daarvoor in aanmerking komen worden per jaar door de ringonderzoekcoördinator vastgesteld.

Opmerking:

Statistiek is de berekening van een kans. Dus zelfs als uit de toegepaste statistiek blijkt dat de monsters niet stabiel en/of homogeen zijn, moeten de bevindingen altijd aan de praktijkervaring getoetst worden.

Validatie van een nieuwe matrix

Uit het schema wordt een aantal representatieve en eventueel problematische parameters gekozen. Hiervoor wordt een validatieprocedure uitgevoerd zoals hierboven beschreven is. De concentratieniveau's worden zoveel mogelijk gelijk gehouden als bij het validatieonderzoek in oppervlaktewater (of de op dat moment moeilijkste matrix waarin onderzoek is uitgevoerd). Uit de gegevens wordt middels de t-toets bezien of de gemiddelden aan het begin en het eind van de geteste periode in beide matrices van elkaar afwijken. Wanneer aangetoond kan worden dat een bepaalde matrix de moeilijkste matrix is, men mag aannemen dat de overige parameters in de nieuwe matrices homogeen en houdbaar zijn als aangetoond deze moeilijkste matrix. Wanneer het een geheel nieuwe matrix betreft, wordt er altijd onderzoek gedaan. Dit kan een literatuuronderzoek, validatie zoals boven beschreven of een test in een ringonderzoek zijn. Bij de laatste optie wordt het ringonderzoek kosteloos of tegen een lage prijs aangeboden.

Opname in het reguliere programma

De resultaten van de proef worden intern geëvalueerd. Indien de resultaten bevredigend waren, wordt het voorlopige validatierapport met deze resultaten en evaluatie uitgebreid en kan de betreffende parameter en/of matrix in het reguliere programma worden opgenomen. De werkwijze wordt beschreven in het zogenaamde kookboek, waarin voor elke parameter en matrix de bereiding is vastgelegd. Deze beschrijving wordt opgenomen als bijlage in het validatierapport. Het definitieve rapport wordt door het hoofd van LMC vastgesteld. De labhoofden dragen zorg voor archivering in het laboratoriumarchief.

Indien de resultaten niet bevredigend zijn, worden afspraken gemaakt over nader onderzoek.

onderwerp

**Validatie van (nieuwe) parameters
en/of matrices**

code

KHRO-6.0

bladzijde

8 van 8

Datum

2003-08-15

CONTROLE EN BIJSTELLING

De ringonderzoekcoördinator houdt toezicht op de juiste uitvoering van deze instructie en neemt zondig corrigerende maatregelen.