

## TOEGEPASTE STATISTIEK IN HET WATERBEHEER (4)

# Waarden beneden de detectiegrens

Analyses van een watermonster leveren niet altijd bruikbare getallen. Soms zijn de gehalten in het water zo laag dat ze in het laboratorium niet nauwkeurig te meten zijn. De gehalten liggen dan onder de detectiegrens. Niet vast te stellen is dan of de stof wel of niet aanwezig is. Binnen één meetreeks (serie metingen van een stof, op een locatie gedurende een jaar) komt het voor dat een aantal keer de stof wel wordt aangetroffen (gehalten boven de detectiegrens) en een aantal keer niet (gehalten onder de detectiegrens). Hoe bereken je dan een jaargemiddelde voor deze stof?

Tot voor kort bestonden drie gangbare methoden voor het omgaan met de waarden onder de detectiegrens (dg). Bij het berekenen van een jaargemiddelde konden dan de waarden onder de detectiegrens op een tijdstip (i) worden vervangen door een vaste waarde ( $x_i$ ).

Hierbij wordt de detectiegrens vermenigvuldigd met een bepaalde factor

- door nul:  $x_i = 0 * dg$   
Het berekende gemiddelde is dan een onderschatting van het werkelijke gemiddelde.
- door de helft van de detectiegrens:  $x_i = 0,5 * dg$   
Het berekende gemiddelde zal soms een overschatting en soms een onderschatting zijn van het werkelijke gemiddelde.
- door de detectiegrens:  $x_i = 1 * dg$   
Het berekende gemiddelde is dan een overschatting van het werkelijk gemiddelde.

In een statistisch onderzoek met diverse stoffen in zoet en zout water is uitgezocht wat de beste methode is voor het omgaan met waarden onder de detectiegrens bij de berekening van de kengetallen, zoals het gemiddelde en de mediaan. Naast de drie gangbare is ook een aantal nieuwe methodes in het onderzoek meegenomen. Voor de beoordeling van de methoden is eerst de schatten parameter (gemiddelde of mediaan) bepaald zonder waarden onder de detectiegrens. Vervolgens zijn verschillende fracties detectiegrenzen aan de reeks toegevoegd.

In de methode die uiteindelijk het beste resultaat gaf, de DG90-methode, worden alle waarden onder de detectiegrens vervangen door een berekende waarde. In tegenstelling tot de huidige methoden, is de berekende

waarde afhankelijk van de verdeling van de meetgegevens van de beschouwde stof. De berekende waarde is namelijk gebaseerd op de fractie van de meetwaarden onder de detectiegrens (f) en de breedte van de verdeling.

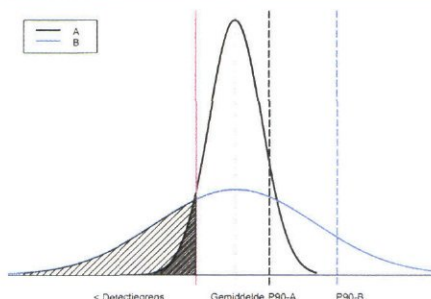
Deze breedte wordt bepaald door afstand tussen de waarde van de detectiegrens en het 90 percentiel (p90) van de meetreeks. De factor die dan gebruikt wordt is:

$$x_i = \left( \frac{dg}{p90} \right)^{c * f} dg,$$

waarin c een constante is waarvan de grootte afhangt van de fractie waarden onder de detectiegrens. De factor kan een waarde tussen 0 en 1 aannemen.


Uit de formule valt af te leiden dat hoe breder de verdeling is, des te kleiner de verhouding tussen de detectiegrens en het 90

Afb. 1a: Twee verdelingen met dezelfde detectiegrens (rode verticale lijn). Bij een verdeling die 'breed' (blauw, B) is, ligt de detectiegrens verder van het 90-percentiel (P90). In dit geval zal de verhouding detectiegrens/90-percentiel kleiner zijn dan bij een 'smalle' verdeling (zwart, A). Het effect van de fractie en de verhouding op de factor is te zien in afbeelding 1b.



percentiel is en dat de berekende waarde dichter bij nul zal liggen. Als de meetwaarden daarentegen dicht bij elkaar liggen (afstand tussen detectiegrens en het 90 percentiel is klein) zal de vervangende waarde dicht bij de detectiegrens liggen (zie afbeelding 1).

Met deze nieuwe methode, omgedoopt als de DG90-methode, wordt een reëler getal verkregen voor de berekende kengetallen gemiddelde en mediaan. De nieuwe methode heeft in tegenstelling tot de gangbare methoden een theoretische basis en is door middel van simulaties met gegevens uit het monitoringmeetnet van het zoete en zoute water getest. Tevens is ze uit te breiden voor meervoudige detectiegrenzen, waarbij de rekenregels niet complexer worden dan bovengenoemde formule.

Interessant is het om de discussie aan te gaan over het gebruik van de werkelijk gemeten waarden onder de detectiegrens. Deze waarden zijn in het laboratorium beschikbaar, maar worden in de rapportage vervangen door de detectiegrens. Bij gebruik van de werkelijk gemeten waarden worden de genoemde substitutiemethoden overbodig. Ongeacht welke methode gekozen wordt, is het voor de gebruikers van de kengetallen te allen tijde van belang te weten welke methode gehanteerd is bij de berekening van het getal. 

Voor meer informatie:

r.n.m.duin@rikz.rws.minvenw.nl  
deVries@cqm.nl

Afb. 1b: Het effect van de verhouding detectiegrens/90-percentiel en de fractie waarden onder de detectiegrens (kleiner dan dg) op de factor. Drie lijnen zijn getekend voor respectievelijk de fracties onder de detectiegrens van 10, 30 en 50. Duidelijk is te zien dat de factor een grotere range krijgt als de fractie onder de detectiegrens groter wordt.

