

Nogmaals: monsternamerequentie van rivierwater

1. Inleiding

Dit is een reactie op het artikel 'Enkele berekeningen inzake monsternamerequentie van rivier- en afvalwater' van de auteurs J. C. Schippers, H. Hofman en J. Roelands, dat in *H₂O* (10) 1977, nr. 3, is verschenen (blz. 71 t/m 74). De auteur van dit artikel, werkzaam bij de afdeling Onderzoek en Planning van DWL Rotterdam, is secretaris van de Subgroep Statistiek van de Werkgroep Waterkwaliteit van de Rijncommissie Waterleidingbedrijven (RIWA). De Subgroep houdt zich ondermeer met de problematiek van de monsternamerequentie van rivierwater bezig en vindt de aanpak van de heren Schippers, Hofman en Roelands in hun artikel wat erg eenvoudig. In deze reactie wordt hierop nader ingegaan. De in het artikel gegeven theorie gaat alleen op wanneer geldt, dat:

- de monsters op aselechte wijze zijn genomen;
- de monsters onderling onafhankelijk zijn (dat wil zeggen er bestaat geen enkele samenhang tussen de verschillende monsters);
- de gemeten concentraties van een stof in de monsters alle geacht kunnen worden afkomstig te zijn uit een normale verdeling. De auteurs noemen deze zelfde voorwaarden wel in hun artikel, doch gaan alleen in de discussie hier iets dieper op in. Hier zal getracht worden de konsekwenties van deze voorwaarden aan te geven.

2. Aselechte trekkingen

Het aselechte nemen van een monster zou inhouden dat elk tijdstip in de loop van het jaar een gelijke kans moet krijgen om te worden uitgekozen als moment voor een monsterneming. Dit is momenteel niet het geval. De in het artikel als voorbeeld genomen monsters zijn op een vast tijdstip om de week genomen. Daar de weekdag en het jaargetijde van invloed zijn op de gemeten grootheden kan niet van aselechte trekking worden gesproken. Wel kan men zich afvragen of het nuttig is om volledig aselechte monsters te nemen. Door precies elke week één monster te nemen ontstaat een soort gelaagde steekproef. Op deze wijze wordt dan voorkomen dat door volledig aselechte trekken er bijv. een onevenredig groot aantal in één seizoen valt. Dit is belangrijk omdat de kansverdeling van de gemeten grootheden in de loop van het jaar verandert, bijv. onder invloed van de temperatuur. In een goede gelaagde steekproef moet echter wel in de week een aselechte tijdstip worden gekozen. Wordt dit niet gedaan dan wordt de spreiding tengevolge van de weekdag geëlimineerd en dus mogelijk een kleinere spreiding gevonden dan die in

werkelijkheid bestaat. Ook bestaat hierdoor de kans op een systematische afwijking. Het voordeel van op gelijke tijdsintervallen betrokken monsters is wel dat eenvoudiger met tijdreeksanalyse kan worden gewerkt.

3. Onderlinge onafhankelijkheid

Erg belangrijk voor het vaststellen van een monsternamerequentie is het al dan niet onderling onafhankelijk zijn van 2 opeenvolgende waarnemingen. Wanneer twee opeenvolgende waarnemingen een grote overeenkomst (correlatie) vertonen voegt de tweede weinig nieuwe informatie toe. Het lijkt dan ook, ten behoeve van de berekening van gemiddelden en spreidingen, weinig zinvol de meetfrequentie op te voeren wanneer de overeenkomst tussen direkt opeenvolgende waarnemingen erg hoog is. De mate van overeenkomst wordt numeriek weergegeven als een autocorrelatiecoëfficiënt. In tabel I zijn een aantal autocorrelatiecoëfficiënten berekend van bij Kop van 't Land wekelijks gemeten grootheden in 1974. Naarmate de samenhang groter is, liggen de in de tabel gegeven waarden dichter bij de 100 %. Veel gemeten waarden blijken een grote mate van samenhang te vertonen.

TABEL I - Autocorrelatiecoëfficiënt van wekelijkse waarnemingen van het Rijnwater bij de Kop van 't Land in 1974.

	1 week verschil	2 weken verschil	3 weken verschil
kleur	0,63	0,62	0,56
Gel. vermogen	0,79	0,57	0,47
pH	0,11	—	—
Cl ⁻	0,77	0,55	0,46
NO ₂ ⁻	0,87	0,78	0,71
SO ₄ ²⁻	0,70	0,47	0,29
NH ₄ ⁺	0,76	0,64	0,54
O ₂	0,84	0,75	0,71
PO ₄ ³⁻ (totaal)	0,13	—	—

Het vaststellen van de maximale meetfrequentie kan gebeuren met de autocorrelatiefunctie. Dit is de functie die het verband weergeeft tussen de autocorrelatiecoëfficiënt en de tijdsafstand tussen twee waarnemingen waarvoor deze coëfficiënt geldt. De maximale meetfrequentie wordt bijv. zodanig gekozen, dat de autocorrelatiecoëfficiënt tussen de monsters 90 % is. Bij de Drinkwaterleiding Rotterdam wordt gedacht over toepassing van deze methode voor het nemen van monsters uit het zuiveringsproces ten behoeve van de verslaggeving. Dat de auteurs in de discussie van hun artikel schrijven, dat er geen methode bestaat om de autocorrelatie bij de berekeningen te betrekken is, gezien het bovenstaande, dus niet juist.

4. Het normaal verdeeld zijn

Het vaststellen of een gemeten grootheid

al dan niet afkomstig is uit een normale verdeling is bij een aantal waarnemingen van circa 50 per jaar veelal onmogelijk. Aan de hand van een lijn op waarschijnlijkheidspapier is geen kwantitatieve uitspraak te doen. Getracht is voor de grootheden uit de vorige tabel een verwerping te krijgen van de hypothese, dat ze genomen kunnen worden gedacht uit een normale verdeling. Dit is gebeurd met de toets van Kolmogorov-Smirnov. Het 95 % betrouwbaarheidsinterval werd door het geringe aantal waarnemingen echter zo groot dat geen enkele keer tot verwerping kon worden overgegaan. Dit wil echter niet zeggen dat de nulhypothese van het normaal verdeeld zijn waar is. De eis van normaal verdeeld zijn is echter minder belangrijk bij vergelijking van de gemiddelden van bijv. twee verschillende jaren. Veel belangrijker is zij echter voor de spreiding.

5. Kritiek op de gebruikte methode voor bepaling van de monsternamerequentie

Gezien het bovenstaande moet de juistheid van de toepassing van de door de auteurs gebruikte theorie, betreffende de bepaling van de monsternamerequentie in rivierwater, in twijfel worden getrokken. De berekende standaardafwijking is bij niet onderling onafhankelijk aselechte waarnemingen onjuist. Omdat veel stoffen in de loop van het jaar een verloop vertonen is het de vraag of sommige spreidingen niet kunnen worden gesplitst in een spreiding ten gevolge van bijv. een seizoensinvloed en een zekere restspreiding. Dit kan gebeuren door het jaar in stukken te verdelen. De kans bestaat dat de op deze wijze per seizoen berekende monsternamerequentie veel kleiner wordt dan $\frac{1}{4}$ van de berekende benodigde jaarfrequentie. Ook lijkt het baseren van een monsternamerequentie op basis van de in het voorgaande jaar opgetreden spreiding zeer twijfelachtig. Spreidingen van gemeten grootheden kunnen in verschillende jaren sterk variëren (zie tabel IV van het besproken artikel bijv. Na⁺) ten gevolge van bijv. debietsinvloed. Hier kan ook nog even gewezen worden op een fout in hoofdstuk 3. Stel dat in 1973 een berekend gemiddelde \bar{m}_1 en in 1974 \bar{m}_2 is gevonden met een nauwkeurigheid van resp. Δc_1 en Δc_2 elk met een betrouwbaarheid van $x\%$. De betrouwbaarheid van de uitspraak, dat de werkelijke gemiddelden van beide jaren gelijk zijn, wanneer $\bar{m}_2 - \Delta c_2 = \bar{m}_1 + \Delta c_1$ is niet $\frac{x}{2}$ gelijk $x\%$ doch ongeveer $\frac{x}{100}$ %. De beide

gemiddelden van geval II in afbeelding 1 verschillen dus significant (met kans x).

6. Verhoging van de nauwkeurigheid van het gemiddelde

De auteurs noemen in hun artikel 3 methoden voor het verhogen van de nauwkeurigheid van het gemiddelde:

- a. vergroting van het aantal geanalyseerde monsters;
- b. mengmonsters;
- c. verbetering analysemethode.

ad a. Voor wat betreft de vergroting van het aantal monsters moet weer gewezen worden op de invloed van de autocorrelatie in de waarnemingen. Wanneer twee op elkaar volgende waarnemingen een grote correlatie vertonen, voegt de tweede waarneming weinig informatie toe en kan dus achterwege blijven. Daar veel grootheden, zoals uit tabel I blijkt, een grote samenhang bezitten lijkt het opvoeren van het aantal monsters weinig zinvol.

ad b. Voor wat betreft de mengmonsters kan men zich afvragen wat hiermee is bedoeld. Aangenomen wordt dat hier bedoeld wordt het monster dat ontstaat wanneer monsters op verschillende tijdstippen genomen bij elkaar worden gedaan (verzamelmonsters). Wanneer de analysemethode 100 % nauwkeurig is, zal het effect hiervan echter op de nauwkeurigheid nihil zijn. Het nu gevonden gemiddelde is gelijk aan het gemiddelde van de afzonderlijke monsters, dus is de nauwkeurigheid gelijk gebleven. Wanneer de analysemethode zelf ook een spreiding vertoont, wordt de spreiding zelfs groter dan die van de afzonderlijke monsters en worden minder waarnemingen gedaan, waardoor elke analysefout belangrijker invloed krijgt. Door het werken met verzamelmonsters wordt echter een relevante informatie over het verloop van de grootte gedurende het jaar verwijderd. Hierdoor lijkt het werken met verzamelmonsters minder nuttig. Wanneer het verzamelmonster wordt gebruikt om het aantal monsters op te voeren zonder verhoging van het aantal analyses wordt het in de vorige alinea gestelde van belang. In dit geval kunnen verzamelmonsters natuurlijk wel nuttig zijn.

ad c. De door de auteurs geplaatste opmerking met betrekking tot de analysemethode is juist. Aan het onderzoek naar de verbetering van de nauwkeurigheid en vergelijkbaarheid van analyseresultaten wordt onder andere in RIWA-verband gewerkt.

7. Conclusie

Het artikel van de heren Schippers, Hofman en Roelands is gebaseerd op een theorie,

die voor het beschreven soort problemen niet toereikend is. Aan de uitgangspunten van de theorie wordt namelijk niet voldaan. Het genoemde artikel lijkt dan ook niet praktisch bruikbaar voor het vaststellen van de monsternamingsfrequentie van rivierwater. Om te komen tot een andere aanpak zal kunnen worden uitgegaan van de autocorrelatiefunctie van de te meten grootheden.

Ing. A. Wedemeier

Sekretaris Subgroep Statistiek
Rijncommissie Waterleidingbedrijven

