



Ed van der Mark, Dunea

Ferry Roman, Ropec

Luc Zandvliet, Het Waterlaboratorium

Toine Ramaker, Dunea

# Ruim tien jaar wateronderzoek: een statistische goudmijn

**Sinds de jaren '90 slaat Dunea de resultaten van drinkwateronderzoek op in de LIMS-databank. Inmiddels is het aantal waterkwaliteitsmetingen van het distributienet de 400.000 gepasseerd. Op deze data is statistisch onderzoek uitgevoerd in het kader van DisConTO\*. Resultaten hiervan met betrekking tot *Aeromonas* spp. laat zien dat een wijziging in de meetfrequentie een trendbreuk oplevert in de onderzochte periode 1999 tot en met 2010. Ook wordt duidelijk dat een significante relatie bestaat tussen het aantal *Aeromonas* en de drinkwatertemperatuur. Verder laat het onderzoek zien dat nagroei van *Aeromonas* regionaal verschillend is.**

**D**rinkwaterbedrijven voeren wateronderzoek uit van bron tot kraan. De metingen zijn deels verplicht gesteld in het Drinkwaterbesluit en deels gebaseerd op bedrijfsspecifieke eisen, zoals voor kwaliteitsbewaking, bedrijfsvoering en onderzoek. Dunea beschikt met LIMS (het Laboratorium Informatie Management-Systeem) over ruim twaalf jaar aan analyse-resultaten. Dit zijn 400.000 waarnemingen in het voorzieningsgebied van Dunea (van Hillegom tot Den Haag en van Den Haag tot Nieuwerkerk aan den IJssel). Sinds de jaren '90 is de toegankelijkheid van de data sterk verbeterd door gebruik van databanken. Hét moment dus om statistisch onderzoek uit te voeren.

Statistisch onderzoek kan inzichten opleveren over de ontwikkeling van de waterkwaliteit in het distributienet over een langere periode. Maar ook eventuele kwaliteitsverschillen per regio kunnen in beeld worden gebracht. Daarnaast kan worden getoetst of resultaten van onderzoeksparameters onderlinge relaties hebben of in verband gebracht kunnen worden met externe factoren. Bevindingen uit statistisch onderzoek kunnen dus aanleiding vormen voor aanbevelingen om de waterkwaliteit in het distributienet verder te verbeteren, meetnetten te optimaliseren of beslissingen op het gebied van assetmanagement te voeren. Hoe kwamen we hiertoe?

## Cijfers kunnen bedriegen

Voordat de statistische analyses daadwerkelijk kunnen beginnen, zijn de opgeslagen data kritisch bekeken. Er is namelijk een aantal valkuilen:

- wijziging van de systematiek van het getal dat wordt opgeslagen als het gevonden gehalte kleiner is dan de onderste analyse-grens;
- wijziging(en) in de analysemethode(n);
- metingen ter plaatse versus metingen in het laboratorium, van bijvoorbeeld zuurgraad en troebelheid;
- wijziging van de meetfrequentie van de parameter tijdens het onderzoek;
- monsterneming direct na openen van de kraan of nadat de temperatuur van het drinkwater constant is geworden.

Deze wijzigingen leveren andere meetresultaten op en geven trendbreuken in grafieken. Dit kan de lengte van de periode waarover statistisch onderzoek kan worden uitgevoerd, sterk beïnvloeden.

## 'De koe bij de horens vatten'

De analyse begon met de parameter *Aeromonas*: een groep van bacteriën die over het algemeen in lage hoeveelheden voorkomt in drinkwater. Het is een bedrijfskundige parameter, dus niet schadelijk voor de gezondheid maar wel beeldvormend voor de waterkwaliteit ter plaatse. Het geeft een indicatie voor de mate van nagroei en daarmee de biologische stabiliteit van het drinkwater. *Aeromonas* is in staat om zich boven een temperatuur van 12 tot 14°C te vermeerderen bij lage gehalten aan makkelijk afbreekbare verbindingen. Deze parameter leent zich goed voor analyse, omdat hiervan stabiele meetdata over een langere periode (twaalf jaar) beschikbaar zijn. In het Waterleidingbesluit is voor *Aeromonas* een norm opgenomen van < 1.000 kve/100

ml. Dunea past een bedrijfsnorm toe van < 200 kve/100 ml.

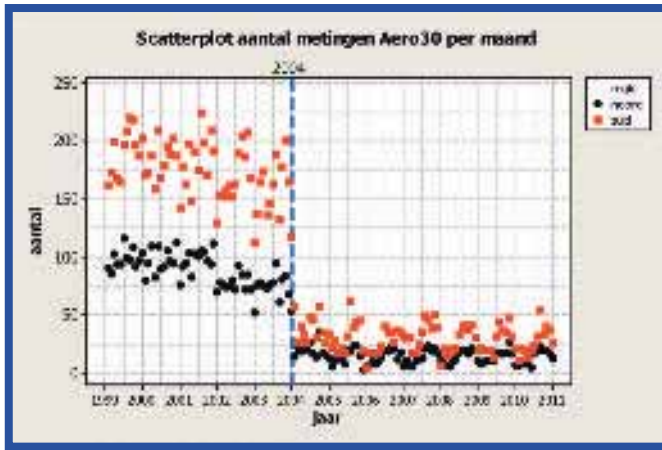
## Toetsen van drie relaties

Bij de analyse van *Aeromonas* gaat de interesse uit naar drie mogelijke relaties:

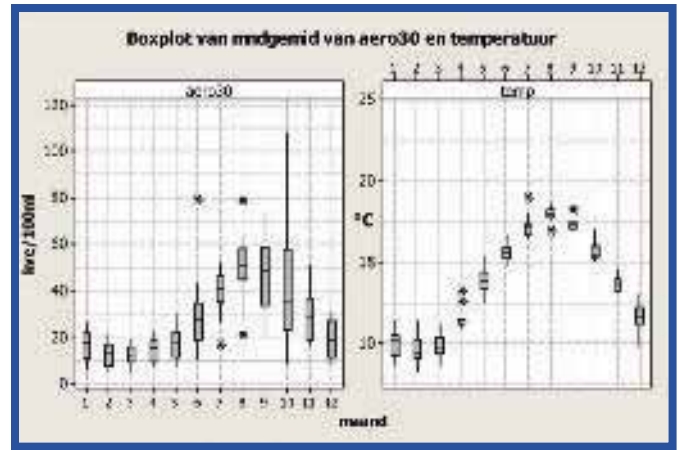
- nagaan of de reductie in de meetfrequentie in 2004 tot een trendbreuk heeft geleid;
- toetsen van de hypothese dat een relatie bestaat tussen het aantal *Aeromonas* en de temperatuur van het drinkwater;
- toetsen of een verschil bestaat in het aantal *Aeromonas* in regio Noord en regio Zuid. Regio Noord krijgt drinkwater van productielocatie Katwijk en koopt in bij Waternet, in een verhouding van ongeveer 9,5:0,5. Regio Zuid wordt voornamelijk voorzien van drinkwater van productielocaties Scheveningen en Monster en koopt in bij Evides globaal in een verhouding van 8,4:1,5:0,1.

## Trendbreuk

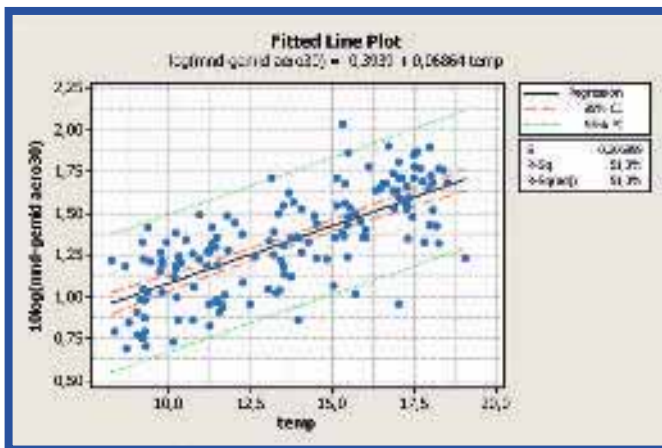
In totaal zijn in de jaren 1999 tot en met 2010 19.479 watermonsters onderzocht op *Aeromonas*. De bedrijfsnorm (< 200 kve/100 ml) is in deze periode 481 keer en de Waterleidingbesluit-norm (< 1000 kve/100 ml) 28 keer overschreden. Dit betekent dat in 97,5 procent van de gevallen de watermonsters voldeden aan de bedrijfsnorm en in 99,8 procent van de gevallen aan het Waterleidingbesluit. Voor de regio Noord bedroeg het aantal overschrijdingen van de bedrijfsnorm en het Waterleidingbesluit respectievelijk 26 en 1 en voor de regio Zuid respectievelijk 455 en 27. Belangrijke vermelding hierbij is dat regio Zuid ongeveer



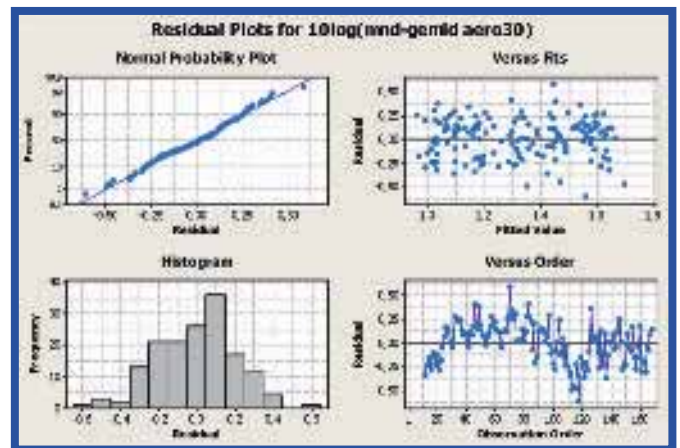
Afb. 1: Aantal metingen Aeromonas per maand per regio.



Afb. 2: De gemeten aantallen Aeromonas laten eenzelfde patroon zien als de temperatuur van het drinkwater (periode: 1999 tot en met 2010).



Afb. 3: De relatie van 10log(maandgemiddelde aero30) versus de maandgemiddelden van de temperatuur.



Afb. 4: Grafische residu-analyse voor de 10log-maandgemiddelden van Aeromonas.

tweemaal zoveel aansluitingen heeft als de regio Noord.

Uit een eerste beschouwing van de data van *Aeromonas* blijkt dat in de meetfrequentie in de afgelopen twaalf jaar (1999 tot en met 2010) in 2004 flink veranderde (zie afbeelding 1).

Het aantal metingen per maand nam in beide regio's af. Deze frequentieverlaging roept de vraag op of de resultaten van voor 2004 nog te vergelijken zijn met de periode na 2004. Uit de cijfers blijkt dat de trefkans van een overschrijding bij een lagere meetfrequentie in ieder geval kleiner is (zie tabel 1).

Met een zogeheten t-toets is na te gaan of de maandgemiddelden in beide perioden significant van elkaar afwijken. Een t-toets is alleen toepasbaar als de waarnemingen normaal zijn verdeeld. Met behulp van een test is geconcludeerd dat de waarnemingen alleen normaal zijn verdeeld als de data getransformeerd worden naar de 10log van de maandgemiddelden ( $p = 0,182$ ). Dit maakt gebruik van de t-toets mogelijk na transformeren van de data in 10log-waarden. Uit de t-toets blijkt dat de gemiddelden over beide perioden van 1999 tot en met 2003 en van 2004 tot en met 2010 significant verschillend zijn ( $p = 0,000$ ). Dit betekent dat de reductie in onderzoeksfrequentie een negatieve invloed heeft op de representativiteit van *Aeromonas* die na 2004 in het

voorzieningsgebied wordt aangetroffen. In de periode 1999 tot en met 2010 treedt dus een trendbreuk op in 2004.

**Relatie *Aeromonas* en de temperatuur**

Aan de hand van boxplots per maand van de gemiddelden van 1999 tot en met 2010 is eenvoudig te zien dat zowel *Aeromonas* als de drinkwatertemperatuur een seizoens-effect laten zien (zie afbeelding 2).

Voor de hand ligt dat een relatie bestaat tussen de groei van *Aeromonas* en de temperatuur van het drinkwater. In tabel 1 was reeds vermeld dat in de twaalfjarige periode 481 overschrijdingen van de bedrijfsnorm en 28 overschrijdingen van de norm in het Waterleidingbesluit hebben plaatsgevonden. In de maanden juli tot en met oktober treffen we de meeste normover-

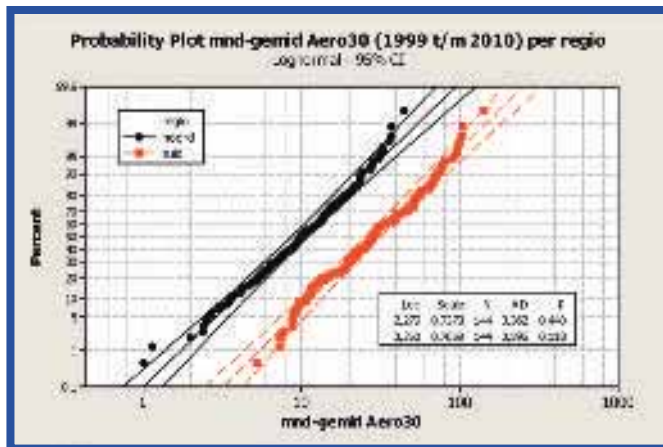
schrijdingen aan. Dit is tevens de periode dat de drinkwatertemperatuur haar maximum bereikt.

Met behulp van regressierekenen kan worden bepaald of er een lineair verband bestaat tussen het aantal *Aeromonas*-bacteriën en de drinkwatertemperatuur. Omdat bekend is dat de maandgemiddelden lognormaal verdeeld zijn, kan ook hier de t-toets worden uitgevoerd.

In afbeelding 3 is de logaritme van de aantallen kve van *Aeromonas*, log(maandgemiddelde aero30) uitgezet tegen de maandgemiddelden van de temperatuur. De berekende regressielijn is  $\log(\text{maandgemiddelde aero30}) = 0,3939 + 0,06864 \cdot \text{temp}$  met de volgende parameters:  $S = 0,2026959$ ,  $R\text{-Sq}(\text{adj}) = 51,0\%$ ,  $p = 0,000$ .

Tabel 1. Aantal overschrijdingen van de bedrijfsnorm- (< 200 kve / 100 ml) en de norm uit het Waterleidingbesluit (< 1000 kve / 100 ml) per regio per periode.

regio	periode	bedrijfsnorm	%	norm uit Waterleidingbesluit	%	aantal metingen
noord	1999 t/m 2003	24	0,5	1	0,02	5.274
	2004 t/m 2010	2	0,2	0	0	1.189
zuid	1999 t/m 2003	389	3,7	25	0,24	10.525
	2004 t/m 2010	66	2,6	2	0,08	2.491
<b>totaal</b>	<b>1999 t/m 2010</b>	<b>481</b>	<b>2,5</b>	<b>28</b>	<b>0,14</b>	<b>19.479</b>



Afb. 5: Toets op lognormale verdeling van de data van beide regio's.

De p-waarde is kleiner dan 0,05 en dus significant te noemen.

Afbeelding 4 geeft een residu grafiek weer. Bij deze analyse van verschillen in modelwaarden en geobserveerde waarden kan de kwaliteit van de uitgevoerde lineaire regressie eenvoudig grafisch worden beoordeeld. De residuen moeten dan wel een normale verdeling volgen, rondom de nullijn zijn verdeeld en het histogram moet een klok vorm hebben.

Uit bovenstaande kunnen we concluderen dat een positieve, significante relatie bestaat tussen de drinkwatertemperatuur en het aantal *Aeromonas*-bacteriën in het distributienet. Dat de temperatuur van het drinkwater af productielocatie varieert van ongeveer 8°C tot 16°C, geeft aan dat de temperatuur in het distributienet nog sterk wordt beïnvloed door de diepteligging van de leiding en straatbekleding erboven. Stijging van de drinkwatertemperatuur als gevolg van klimaatverandering zou dus in de toekomst voor meer nagroei van *Aeromonas* in het voorzieningsgebied kunnen zorgen. In het bedrijfstakonderzoek van de waterleidingbedrijven wordt hier aandacht aan besteed.

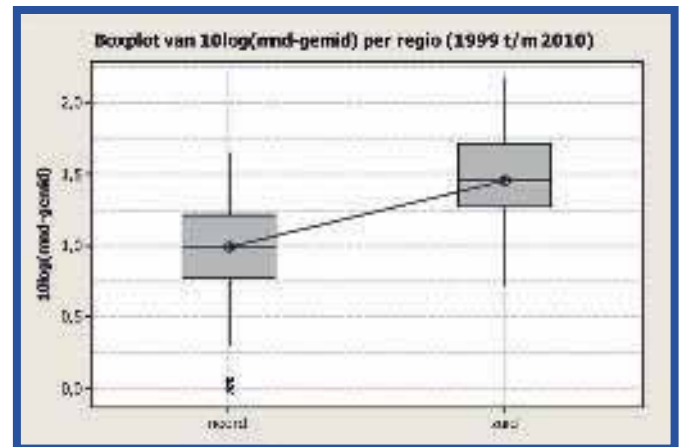
### Regioverschillen

Omdat regio Noord voornamelijk wordt voorzien van drinkwater van productielocatie Katwijk en regio Zuid door Scheveningen,

is het interessant om na te gaan of een significant regionaal verschil bestaat voor *Aeromonas*. De datasets van beide regio's zijn lognormaal verdeeld (zie afbeelding 5); de t-toets kan dus worden toegepast. De toets levert een p-waarde op van 0,000. Dus de gemiddelden van regio Noord en regio Zuid over de jaren 1999 t/m 2010 zijn significant verschillend. De grafiek met de boxplots per regio in afbeelding 6 laat dit ook zien. De langjarige gemiddelden zijn respectievelijk 10 en 29 kve/100ml. Met andere woorden: de nagroei van *Aeromonas* in het drinkwater van de noordelijke regio is significant minder dan die in de zuidelijke regio. Dit verschil kan worden veroorzaakt door diverse factoren, zoals de kwaliteit van het geproduceerde drinkwater, verschil in leidingmateriaal, verblijftijd van het drinkwater, mate van vermaasheid van het net en meng-gebieden. Om dit vast te stellen, is vervolgonderzoek noodzakelijk.

### Conclusie

Gedegen statistisch onderzoek naar waterkwaliteitsgegevens is een belangrijk gereedschap om trends en relaties in de waterkwaliteit vast te leggen. Tijdens dit onderzoek kwam aan het licht dat de verlaging van de onderzoeksfrequentie voor *Aeromonas* in 2004 leidde tot een trendbreuk in de data. Ook de trefkans van normoverschrijdingen is in deze periode afgenomen. Met deze informatie kan het meetnet en de frequentie van de metingen kritisch worden



Afb. 6: Boxplot log maandgemiddelden per regio over de periode 1999 t/m 2010.

bekeken en geëvalueerd. Tevens kon via dit statistisch onderzoek de voor de hand liggende hypothese dat groei van *Aeromonas* temperatuurafhankelijk is, worden vastgesteld. Dit ondanks de doorgaans zeer lage waarden *Aeromonas* die in het voorzieningsgebied van Dunea worden aangetroffen. Verder bleek dat de gemiddelde waarden voor *Aeromonas* in de regio Noord significant lager zijn dan in de regio Zuid. Hier kunnen verschillende oorzaken aan ten grondslag liggen, die nader onderzoek vragen.

Dunea gaat verder met de analyse van de waterkwaliteitsgegevens. De resultaten van de analyse op *Aeromonas* heeft belangrijke informatie opgeleverd om de ontwikkeling van de waterkwaliteit in het distributienet beter te begrijpen, het meetnet nog eens kritisch te beschouwen en mogelijk later voeding te geven aan investeringsbeslissingen.

### NOTEN

\* DisConTO is een samenwerkingsverband van Vitens, PWN, Brabant Water, RIVM, TU Delft, DHV, softwarebedrijf UReason en Dunea, dat onderzoek verricht naar de ontwikkeling van de drinkwaterkwaliteit in het distributienet onder normale omstandigheden en bij calamiteiten. Het onderzoek wordt financieel ondersteund door AgentschapNL.