
Trendanalyse op maat voor een meetnet waterkwaliteit

Paul K. Baggelaar¹ en Eit C.J. van der Meulen²

Dit artikel beschrijft hoe grote aantallen meetreeksen, zoals van een meetnet waterkwaliteit, efficiënt en statistisch verantwoord kunnen worden geanalyseerd op trend. De werkwijze houdt rekening met de specifieke kenmerken van milieugegevens, zoals een scheve kansverdeling, seizoenseffecten, autocorrelatie, gecensureerde waarden en ontbrekende waarden. Ondanks de grootschalige aanpak wordt per meetreeks maatwerk geleverd, gericht op het maximaliseren van het onderscheidend vermogen bij de trenddetectie en de nauwkeurigheid bij de trendschatting. Met deze werkwijze is inmiddels al voor een aantal grootschalige meetnetten, met honderden tot duizenden meetreeksen per meetnet, objectief vastgesteld waar en voor welke parameters er statistisch significante kwaliteitsverbeteringen, dan wel -verslechtingen zijn opgetreden.

Inleiding

De meeste grootschalige meetinspanningen in het milieuonderzoek hebben als hoofddoelstellingen het beschrijven en beoordelen van: 1) de toestand en 2) de toestandsverandering. Voor het beschrijven van de toestand zijn diverse beschrijvende statistische methoden beschikbaar en voor het beoordelen van de toestand kunnen verschillende normen dienen als objectieve maatstaf. Maar het beoordelen van de toestandsverandering is minder eenvoudig, aangezien objectief onderscheid moet worden gemaakt tussen een toevallige en een structurele verandering. Dit vereist statistisch maatwerk voor elke afzonderlijke meetreeks, wat in het geval van een grootschalig meetnet - met soms tot duizenden meetreeksen - veel inspanning vergt. Dit artikel beschrijft een procedure voor trendanalyse op maat voor grote aantallen meetreeksen, zodanig dat deze ook geautomatiseerd kan worden uitgevoerd.

Trendanalyse

Trendanalyse is een statistische techniek om objectief vast te stellen of een meetreeks een trend vertoont, waarbij tevens een schatting voor de grootte van de trend wordt gegeven. Onder het begrip trend verstaan we hier een verandering van het centrum

¹ Icastat, Amstelveen, paul.baggelaar@planet.nl

² AMO, Hengelo, amo@home.nl

van de kansverdeling van de meetwaarden, over een tijdshorizon van tenminste enkele jaren. Seizoensmatige veranderingen en kortstondige calamiteiten vallen daar dus niet onder.

Een statistische trendanalyse omvat per geanalyseerde reeks twee onderdelen, namelijk:

- (1) trenddetectie - dit levert een objectief antwoord op de vraag of de reeks een trend vertoont en wordt uitgevoerd met een statistische toets op trend;
- (2) trendkwantificering - dit levert een schatting van de grootte van de trend, bijvoorbeeld uitgedrukt als verandering per jaar.

Trends kunnen verschillende vormen hebben, maar we beperken ons hier tot de monotone trend. Een monotone trend is een verandering die doorgaans dezelfde kant opgaat, dat wil zeggen opwaarts of neerwaarts, ongeacht of dit lineair, stapsgewijs of in een andere vorm is. Waar het gaat om milieuveranderingen is dit een veel voorkomende trendvorm.

Statistisch toetsen op monotone trend

Er bestaan meerdere statistische toetsen op monotone trend in meetreeksen. De bekendste is de lineaire regressietoets, gebaseerd op een lineair regressiemodel van de reeks, met de tijd als verklarende variabele (zie het tekstkader). Er zijn ook toetsen beschikbaar die uitgebreidere vormen hanteren van het lineaire regressiemodel, om rekening te kunnen houden met log-normaliteit, seizoenseffecten en/of autocorrelatie³ (zie bijvoorbeeld [Duin, 1993]). Verder zijn er diverse verdelingsvrije toetsen beschikbaar, die geen eisen stellen aan het soort kansverdeling waar de meetwaarden uit afkomstig zijn, zoals de Mann-Kendalltoets [Mann, 1945; Kendall, 1938 en 1975] en uitbreidingen daarvan om rekening te kunnen houden met seizoenseffecten [Hirsch e.a., 1982], of met zowel seizoenseffecten als autocorrelatie [Hirsch and Slack, 1984]. Andere voorbeelden zijn een toets gebaseerd op de Spearman-rangcorrelatiecoëfficiënt [Lehmann, 1975] en uitbreidingen daarvan om rekening te houden met seizoenseffecten, respectievelijk autocorrelatie [Lettenmaier, 1976; Montgomery and Reckhow, 1984] en een toets gebaseerd op lineaire regressie van rangnummers na correctie voor seizoenseffecten [Farrell, 1980; Van Belle and Hughes, 1984].

Trendtoetsen met lineaire regressie

Om met de standaarduitvoering van lineaire regressie op verantwoorde wijze te kunnen toetsen op trend, dienen de residuën van het regressiemodel onderling onafhankelijke trekkingen te vormen uit een normale kansverdeling. Maar doordat chronologisch gerangschikte waarden van milieuparameters vaak onderling afhankelijk zijn, vanwege seizoensinvloeden en autocorrelatie, is dit vaak ook bij modelresiduën het geval. Verder blijken waarden van milieuparameters doorgaans

³ Autocorrelatie is het verschijnsel dat opeenvolgende meetwaarden van een meetreeks een relatie vertonen. Veel statistische technieken zijn gebaseerd op de vooronderstelling dat meetwaarden onafhankelijk van elkaar zijn genomen.

geen normale, maar een scheve frequentieverdeling te vertonen, veelal met een langere rechterstaart (positieve scheefheid), wat dan vaak ook het geval is voor de modelresiduën. Aanvankelijk werd aanbevolen een dergelijke reeks om te zetten in een reeks jaargemiddelden, omdat dat vaak leidt tot onderling onafhankelijke waarden, afkomstig uit een normale kansverdeling, zodat dan toch op trend kan worden getoetst met lineaire regressie (zie bijvoorbeeld [Sanders e.a., 1983]). Zo'n 25 à 30 jaar geleden zijn er echter verscheidene andere statistische toetsen op trend ontwikkeld - vooral door statistici van de Amerikaanse Geologische Dienst (USGS) - die wél rekening kunnen houden met de speciale karakteristieken van meetreeksen van milieuparameters. In verschillende simulatiestudies is vastgesteld dat het toepassen van deze toetsen op de oorspronkelijke reeks doorgaans meer onderscheidend vermogen⁴ oplevert dan het toepassen van de lineaire regressietoets op de reeks van jaargemiddelden (zie bijvoorbeeld [Baggelaar en Baggelaar, 1991; Harcum e.a., 1992]).

Gezien de grote verscheidenheid aan statistische toetsen op trend, elk met specifieke mogelijkheden en onderliggende vooronderstellingen, is de nodige inspanning en expertise vereist om de meest geschikte toets te kunnen selecteren. Een statistische evaluatie van veranderingen in het milieu is dan ook wel net zo zeer een kunst als een wetenschap genoemd [Ward and Loftis, 1986]. In de praktijk is het ook niet ongebruikelijk om bij trenddetectie van een groot aantal meetreeksen slechts één en dezelfde trendtoets te hanteren, aangezien dat makkelijker is uit te voeren en ook beter is uit te leggen aan leken. Maar die voordelen wegen ons inziens niet op tegen de verkeerde conclusies die kunnen worden getrokken, aangezien dan vaker dan nodig trends over het hoofd worden gezien, of onterecht worden gedetecteerd. Gezien de grote maatschappelijke en bedrijfsmatige implicaties die ontwikkelingen van de waterkwaliteit kunnen hebben, is er behoefte aan statistisch verantwoorde en objectieve conclusies over dergelijke ontwikkelingen. Verder worden er jaarlijks dermate grote bedragen besteed aan het meten van de waterkwaliteit dat het zaak is deze gegevens zo optimaal mogelijk te verwerken tot informatie. Door te besparen op de statistische verwerking zal de duur betaalde meetinformatie onvoldoende uit de ruwe gegevens worden gefilterd.

Om de meetnetinformatie over veranderingen zo goed mogelijk te kunnen ontsluiten is op gefaseerde wijze een procedure (en ondersteunende software) ontwikkeld waarmee grote aantallen meetreeksen geautomatiseerd op trend kunnen worden geanalyseerd. De eerste aanzet daartoe is ruim twintig jaar geleden gegeven, in opdracht van de RIWA (Vereniging van Rivierwaterbedrijven), die een meetnet beheert om toestand en trends van de kwaliteit van Rijn en Maas te monitoren. Met een combinatie van Monte Carlo-simulaties en theorie is voor verschillende soorten tijdreeksen vastgesteld welke van de bovengenoemde trendtoetsen het best functioneert [Baggelaar en Baggelaar, 1989 en 1991]. De procedure is in latere fasen aangepast, mede op basis van praktijkervaringen en nieuwe inzichten [Baggelaar, 1994; Baggelaar en Van der Meulen, 2012]. De procedure is speciaal ontwikkeld voor milieuonderzoek en kan rekening houden met de specifieke eigenschappen van milieureeksen, zoals een niet-normale kansverdeling

⁴ Het onderscheidend vermogen is de kans om een werkelijk bestaande trend te detecteren. Dit vermogen neemt toe met de grootte van de trend.

(al of niet door uitschieters), autocorrelatie, seizoenseffecten, gecensureerde meetwaarden⁵ en ontbrekende meetwaarden. Hierdoor wordt maatwerk geleverd, met een toets die aansluit op de statistische karakteristieken van de meetreeks. De intentie van de procedure is om voor elke reeks zowel het onderscheidend vermogen bij de trenddetectie als de nauwkeurigheid bij het trendschatten te maximaliseren.

Op basis van een vergelijking van de mogelijkheden en de capaciteiten van de hiervoor vermelde trendtoetsen is besloten de procedure uit te rusten met de volgende twee groepen trendtoetsen: 1) de lineaire regressietoets en uitbreidingen daarvan en 2) de Mann-Kendalltoets en uitbreidingen daarvan. Een toets uit de eerste groep is van toepassing als de residuën van het lineaire regressiemodel afkomstig zijn uit een normale kansverdeling. Als dat niet het geval is, is een toets uit de tweede groep van toepassing. Voor beide groepen zijn de uitbreidingen bedoeld om ook rekening te kunnen houden met seizoenseffecten en/of autocorrelatie. De lineaire regressietoetsen behoren tot de klasse van de parametrische toetsen, terwijl de Mann-Kendalltoetsen behoren tot de klasse van de verdelingsvrije toetsen.

In het nu volgende lichtten we de details van deze trendtoetsen toe. Daarna bespreken we onze procedure voor de selectie van de meest geschikte toets voor een reeks.

Lineaire regressietoets en zijn uitbreidingen

Als de waarden van de beschouwde variabele afkomstig zijn uit een normale verdeling en er geen sprake is van seizoenseffecten of autocorrelatie, dan kan in de meeste gevallen de lineaire regressietoets worden toegepast. Het daarbij gehanteerde model luidt:

$$Z_t = b_0 + b_1 \cdot X_t + e_t \tag{1}$$

met Z de waarden van de beschouwde variabele, b_0 het geschatte intercept, b_1 de geschatte lineaire helling, X de tijd, t de tijdsindex en tenslotte e het modelresidu. De toetsingsgrootte T wordt berekend als:

$$T = \frac{b_1}{\hat{\sigma}[b_1]} \tag{2}$$

met $\hat{\sigma}[b_1]$ de standaardfout van de schatter van de lineaire helling. Als de modelresiduën ($e_1 \dots e_n$) afkomstig zijn uit dezelfde normale kansverdeling en tevens onafhankelijk van elkaar zijn, mogen we met $100\% \cdot (1-\alpha)$ betrouwbaarheid spreken van een trend, als geldt:

$$\begin{aligned} |T| > t_{(1-\alpha/2, n-2)} & \text{ bij tweezijdig toetsen} \\ T > t_{(1-\alpha, n-2)} & \text{ bij eenzijdig toetsen op een positieve trend} \\ T < t_{(\alpha, n-2)} & \text{ bij eenzijdig toetsen op een negatieve trend} \end{aligned} \tag{3}$$

met $t_{(1-\alpha/2, n-2)}$, $t_{(1-\alpha, n-2)}$ en $t_{(\alpha, n-2)}$ de Student-t-waarden met een eenzijdige onderschrijdingskans van respectievelijk $1-\alpha/2$, $1-\alpha$ en α bij $n-2$ vrijheidsgraden (n is hier het aantal waarden in de reeks).

⁵ Een gecensureerde meetwaarde is een waarde gerapporteerd als ‘kleiner dan’ een bepaalde rapportagegrens. Voorbeeld: $< 1 \mu\text{g/l}$.

Lineaire regressietoets met verdiscontering seizoenseffecten en/of autocorrelatie

Als de waarden van de beschouwde variabele afkomstig zijn uit een normale verdeling en er wél sprake is van seizoenseffecten en/of autocorrelatie, dan dient daarmee rekening te worden gehouden. Er moet dan worden getoetst op trend met een uitgebreidere vorm van het lineaire regressiemodel, zoals hieronder toegelicht. Als er zowel sprake is van seizoenseffecten als van autocorrelatie kan worden getoetst op trend met het volgende gecombineerde model:

$$Z_t = b_0 + b_1 \cdot X_t + \sum_{i=2}^s (\gamma_i \cdot I_i) + N_t \quad (4)$$
$$N_t = \phi_1 \cdot N_{t-1} + e_t$$

met Z , b_0 , b_1 , X en e als boven beschreven, s het aantal tijdstappen per jaar (voor dit soort analyses te interpreteren als het aantal seizoenen), i de seizoensindex, γ het seizoenseffect, I het seizoen (dit is 1 als X_t in seizoen i valt en anders 0) en N de ruis van het model. De autocorrelatie van de ruis wordt beschreven door het tweede deel van het model, met ϕ_1 de 1e-orde autoregressieve modelparameter [Box and Jenkins, 1976]. Het eerste seizoen is niet opgenomen in dit uitgebreide model, zodat zijn (eventuele) effect zal zijn opgenomen in de schatting van het intercept (b_0). Een relevant seizoenseffect zal daardoor tot uiting komen in een statistisch significant effect (γ) van minstens één van de andere seizoenen. De geschatte grootte van dat effect vertegenwoordigt dan het verschil tussen het seizoenseffect van seizoen i en dat van het eerste seizoen. Als er alleen sprake is van seizoenseffecten vervalt het ruisdeel van (4) en als er alleen sprake is van autocorrelatie vervalt het seizoensdeel.

Ook om statistisch te kunnen toetsen op trend aan de hand van deze uitgebreidere vorm van het lineaire regressiemodel, moeten de modelresiduën ($e_1 \dots e_n$) afkomstig zijn uit dezelfde normale kansverdeling en tevens onafhankelijk van elkaar zijn. Het toetsen gaat dan als boven beschreven voor de (gewone) lineaire regressietoets, zij het dat het aantal vrijheidsgraden niet meer gelijk is aan $n-2$, maar aan $n-k$, waarbij k het aantal geschatte modelparameters is.

Mann-Kendalltoets en zijn uitbreidingen

De Mann-Kendalltoets op trend is verdelingsvrij en stelt dus geen voorwaarden aan het soort kansverdeling waar de waarden uit afkomstig zijn. Deze toets komt daarom in aanmerking als er sprake is van niet-normaliteit. De vooronderstellingen zijn dat de waarden nóch seizoenseffecten, nóch autocorrelatie vertonen. De toetsingsgrootte S wordt berekend als [Mann, 1945, Kendall, 1938 en 1975]:

$$S = \sum_{t=1}^{n-1} \sum_{k=t+1}^n \text{sgn}[Z_k - Z_t] \quad (5)$$

met Z de waarden van de beschouwde variabele, t en k waarden van de tijdsindex, n het aantal waarden en $\text{sgn}[\]$ de signum-functie, volgens:

$$\begin{aligned} \text{sgn}[\theta] &= 1 \quad \text{als } \theta > 0 \\ \text{sgn}[\theta] &= 0 \quad \text{als } \theta = 0 \\ \text{sgn}[\theta] &= -1 \quad \text{als } \theta < 0 \end{aligned} \quad (6)$$

De nulhypothese luidt dat de waarden $Z_1 \dots Z_n$ afkomstig zijn uit n onderling onafhankelijke en identieke kansverdelingen en dat er dus geen trend bestaat. Voor $n \leq 10$ is een tabel beschikbaar met kritische waarden voor S onder de nulhypothese. Voor $n > 10$ zal de kansverdeling van S de normale verdeling naderen en kan de onderstaande benadering worden gehanteerd.

Onder de nulhypothese geldt voor respectievelijk de verwachtingswaarde en de variantie van S :

$$E[S] = 0 \text{ en } \text{Var}[S] = \frac{n \cdot (n-1) \cdot (2n+5)}{18} \tag{7}$$

Hieruit kunnen we, mede met behulp van een continuïteitscorrectie⁶, de gestandaardiseerde toetsingsgrootte T berekenen, als:

$$\begin{aligned} T &= \frac{S-1}{\sqrt{\text{Var}[S]}} \text{ als } S > 0 \\ T &= 0 \text{ als } S = 0 \\ T &= \frac{S+1}{\sqrt{\text{Var}[S]}} \text{ als } S < 0 \end{aligned} \tag{8}$$

Onder de nulhypothese ('géén trend') volgt T een standaardnormale kansverdeling. We mogen met $100\% \cdot (1-\alpha)$ betrouwbaarheid spreken van een trend, als geldt:

$$\begin{aligned} |T| > U_{(1-\alpha/2)} & \text{ bij tweezijdig toetsen} \\ T > U_{(1-\alpha)} & \text{ bij eenzijdig toetsen op een positieve trend} \\ T < U_{(\alpha)} & \text{ bij eenzijdig toetsen op een negatieve trend} \end{aligned} \tag{9}$$

met $U_{(1-\alpha/2)}$, $U_{(1-\alpha)}$ en $U_{(\alpha)}$ de waarden van de standaardnormale kansverdeling met een eenzijdige onderschrijdingskans van respectievelijk $1-\alpha/2$, $1-\alpha$ en α .

Mann-Kendalltoets met verdiscontering seizoenseffecten en/of autocorrelatie

Als er behalve niet-normaliteit tevens sprake is van seizoenseffecten en/of autocorrelatie, dan moet worden getoetst op trend met uitgebreidere vormen van de Mann-Kendalltoets, zoals hieronder toegelicht.

Als er behalve niet-normaliteit ook sprake is van seizoenseffecten, komt de seizoenale Mann-Kendalltoets in aanmerking. Dit is een multivariate uitbreiding van de Mann-Kendalltoets [Hirsch e.a., 1982]. De nulhypothese van de Mann-Kendalltoets, dat de waarden $Z_1 \dots Z_n$ afkomstig zijn uit n onderling onafhankelijke en identieke kansverdelingen, is bij het optreden van seizoenseffecten te beperkt. De verdelingen hangen dan immers af van de tijd van het jaar en zijn daarmee niet meer identiek. De multivariate uitbreiding houdt in dat voor elk afzonderlijk seizoen een toetsingsgrootte S_g wordt berekend, volgens:

$$S_g = \sum_{i=1}^{n_g-1} \sum_{k=i+1}^{n_g} \text{sgn}[Z_{kg} - Z_{ig}] \tag{10}$$

⁶ De continuïteitscorrectie dient om een discrete stochastische variabele zo goed mogelijk te benaderen met een continue stochastische variabele.

met g de index voor het seizoen, n_g het aantal jaren met een waarde in seizoen g en i en k indices voor het jaar. Vervolgens wordt door sommatie van alle S_g -waarden de toetsingsgrootte S^* verkregen:

$$S^* = \sum_{g=1}^s S_g \quad (11)$$

met s het aantal seizoenen in een jaar. De nulhypothese luidt nu dat de waarden van de reeks afkomstig zijn uit n onderling onafhankelijke kansverdelingen, die voor een bepaald seizoen tevens identiek zijn. En de alternatieve hypothese luidt dat voor één of meer seizoenen de kansverdelingen niet identiek zijn, wat neerkomt op een monotone trend in één of meer seizoenen. Onder de nulhypothese geldt voor de verwachtingswaarde van S^* :

$$E[S^*] = \sum_{g=1}^s E[S_g] = 0 \quad (12)$$

en voor de variantie van S^* :

$$\text{Var}[S^*] = \sum_{g=1}^s \text{Var}[S_g] + \sum_{g=1}^s \sum_{h \neq g}^s \text{Cov}[S_g, S_h] \quad (13)$$

Vanwege de onderlinge onafhankelijkheid geldt $\text{Cov}[S_g, S_h] = 0$, waardoor bovenstaande formule reduceert tot:

$$\text{Var}[S^*] = \sum_{g=1}^s \text{Var}[S_g] = \sum_{g=1}^s \frac{n_g \cdot (n_g - 1) \cdot (2n_g + 5)}{18} \quad (14)$$

Als het product van het aantal seizoenen (met beschikbare waarden) en het aantal jaren groot genoeg is (zeg minstens 20 à 30), zal de kansverdeling van S^* onder de nulhypothese de normale verdeling naderen. Uit S^* kan dan, mede met behulp van een continuïteitscorrectie, de gestandaardiseerde toetsingsgrootte T worden berekend, waarmee vervolgens wordt getoetst op trend, op dezelfde wijze als eerder beschreven voor de toetsingsgrootte S van de Mann-Kendalltoets (formules (8) en (9)).

Als er behalve niet-normaliteit en seizoeneffecten ook sprake is van autocorrelatie, komt de seizoenale Mann-Kendalltoets met verdiscontering van autocorrelatie in aanmerking [Hirsch and Slack, 1984]. Vanwege de autocorrelatie geldt nu niet $\text{Cov}[S_g, S_h] = 0$, zodat deze bij het schatten van de variantie van S^* wel degelijk een rol speelt (zie formule (13)). De schatter voor $\text{Cov}[S_g, S_h]$ is uitgewerkt door [Dietz and Killeen, 1981].

Voor de Mann-Kendalltoets en zijn uitbreidingen zijn ook specifieke correcties beschikbaar voor als er sprake is van ontbrekende waarden, of van gelijke waarden. Gelijke waarden kunnen bijvoorbeeld voorkomen als er meerdere gecensureerde waarden in een meetreeks zijn. Alvorens trendanalyse toe te kunnen passen moeten dergelijke semi-kwantitatieve waarden namelijk worden vervangen door kwantitatieve waarden (zie ook verder), zodat er bij het voorkomen van gecensureerde waarden gelijke waarden in de reeks ontstaan.

Wanneer welke trendtoets toepassen?

Een tijdreeks kunnen we opvatten als een realisatie van één of ander proces, dat is te kenschetsen op basis van statistische karakteristieken, zoals het soort kansverdeling en het al of niet optreden van seizoenseffecten en/of autocorrelatie. Elk van de hiervoor beschreven trendtoetsen zal het best functioneren bij tijdreeksen gegenereerd door een specifiek soort proces. We zullen dat aanduiden als het preferentieproces voor die toets. Bij toepassing van een toets op een tijdreeks die is gegenereerd door zijn preferentieproces zal de toets voldoen aan de volgende twee voorwaarden:

1. zijn empirische significantieniveau is niet hoger dan het nominale significantieniveau (zie de toelichting in het onderstaande tekstkader);
2. zijn onderscheidend vermogen ligt doorgaans hoger dan dat van andere trendtoetsen die voldoen aan voorwaarde 1.

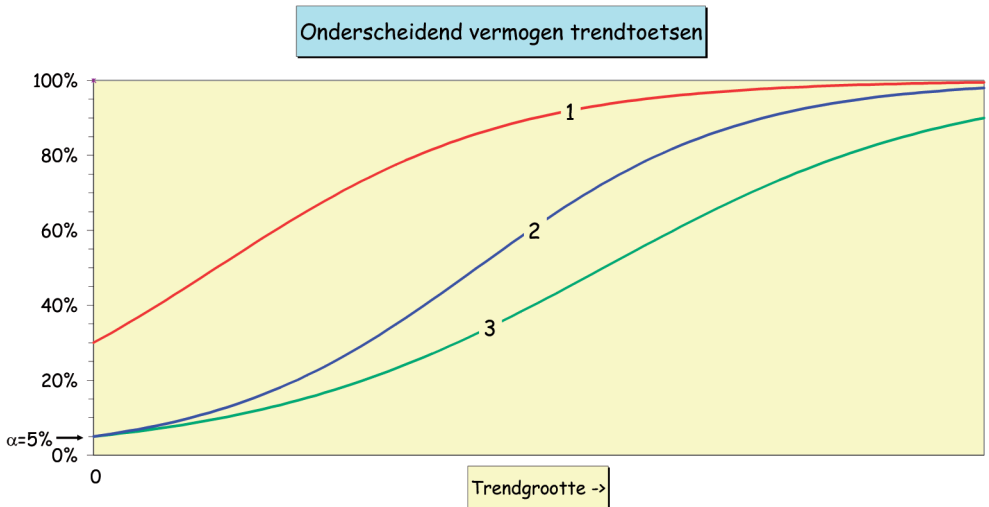
Het is voor een meetnetbeheerder van belang dat wordt voldaan aan beide voorwaarden. De eerste voorwaarde garandeert namelijk dat het percentage onterecht gedetecteerde trends binnen de vóóraf geaccepteerde limiet blijft. Dit verleent zeggingskracht en objectiviteit aan de signaleringsfunctie van het meetnet. De tweede voorwaarde bewerkstelligt dat er zo efficiënt mogelijk informatie wordt gefilterd uit de kostbare meetgegevens van het meetnet.

Het nominale versus het empirische significantieniveau van een trendtoets

Het nominale significantieniveau van een toets wordt vóóraf ingesteld door de gebruiker. In het geval van trendanalyse komt dit neer op het aangeven welk risico wordt geaccepteerd op het onterecht detecteren van een trend. Dit risico kan niet te laag worden ingesteld, want dan heeft de toets nog maar weinig onderscheidend vermogen. Doorgaans wordt het risico op 5% ingesteld, wat impliceert dat wordt getoetst met 95% betrouwbaarheid. De nulhypothese (deze luidt: “er is geen trend”) wordt verworpen als - gegeven de nulhypothese - de kans op de afgeleide toetsingsgrootte of een extremere waarde kleiner is dan het nominale significantieniveau. Die kans wordt wel aangeduid als de p-waarde van de toetsingsgrootte. Afhankelijk van het soort statistische toets blijkt uit de theorie en/of simulatiestudies dat als de toets wordt gehanteerd in gevallen waarvoor deze minder geschikt is, het empirische significantieniveau (het significantieniveau dat daadwerkelijk optreedt) kan afwijken van het nominale significantieniveau (het significantieniveau dat vóóraf is ingesteld).

Het bovenstaande principe lichten we toe met afbeelding 1. Deze toont voor een drietal (fictieve) toetsen op monotone trend de relatie tussen het onderscheidend vermogen en de trendgrootte⁷, voor een bepaald fictief proces. Daarbij geeft de Y-as-afsnijding (de waarde bij geen trend) het risico op onterechte trenddetectie (het empirische significantieniveau). Toets 1 is niet geschikt voor dit proces, omdat die niet voldoet aan voorwaarde 1. De toets geeft namelijk een groter risico op onterechte trenddetectie dan het nominale significantieniveau α (ingesteld op 5% in dit voorbeeld). De andere twee toetsen voldoen wel aan die voorwaarde. En aangezien van deze toets 2 het meeste onderscheidend vermogen heeft, is dat de meest geschikte toets voor dit proces.

⁷ In het Engels wordt een dergelijke relatie aangeduid als ‘power curve’.

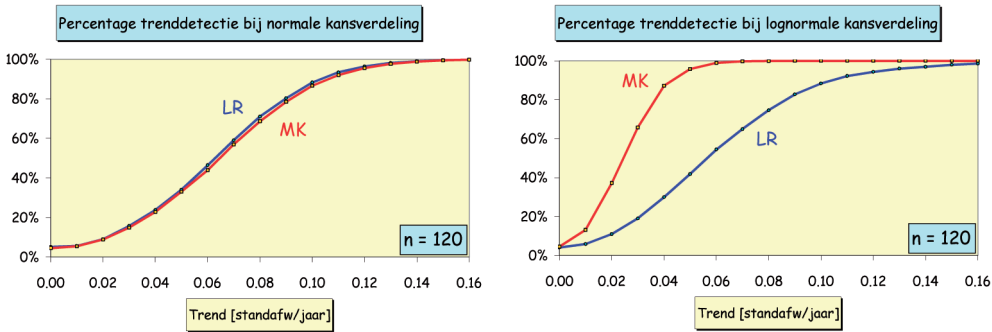


Afbeelding 1: Fictieve voorbeelden van het onderscheidend vermogen van een drietal toetsen op monotone trend, als functie van de trendgrootte. De Y-as-afsnijding (de waarde bij geen trend) is het risico op onterechte trenddetectie (het empirische significantieniveau). Het nominale significantieniveau (α) is voor elk van deze toetsen ingesteld op 5%.

Een curve als die van toets 1 in afbeelding 1 is bijvoorbeeld te verwachten als de standaarduitvoering van de lineaire regressietoets wordt toegepast op een normaal verdeeld proces met (positieve) autocorrelatie. Door de autocorrelatie zal de standaardfout van de hellingschatter worden onderschat en dit geeft volgens formule (2) een overschatting van de toetsingsgrootte en daarmee té vaak onterechte trenddetecties.

Parametrisch versus verdelingsvrij toetsen

Er leeft een hardnekkig misverstand dat een verdelingsvrije toets altijd minder onderscheidend vermogen oplevert dan een parametrische toets. Dit is namelijk alleen het geval als de meetwaarden daadwerkelijk afkomstig zijn uit het soort kansverdeling waar de parametrische toets van uitgaat. Het benutten van de kennis over de achterliggende kansverdeling levert deze toets dan immers voordeel op ten opzichte van de verdelingsvrije toets. Maar dit gaat niet meer op als de meetwaarden afkomstig zijn uit een anderssoortige kansverdeling dan verondersteld. Zo kunnen verdelingsvrije toetsen beduidend meer onderscheidend vermogen opleveren dan parametrische toetsen bij afwijkingen van normaliteit (zie bijvoorbeeld [Bradley, 1968; Helsel and Hirsch, 1988 en Helsel and Hirsch, 1991]). Dit blijkt ook uit de resultaten van een simulatie die wij hebben uitgevoerd en die zijn weergegeven in afbeelding 2. Daarin wordt het empirisch vastgestelde percentage trenddetectie (op te vatten als een raming van het onderscheidend vermogen) van de lineaire regressietoets vergeleken met dat van de Mann-Kendalltoets, voor enerzijds een normale kansverdeling (links) en anderzijds een log-normale kansverdeling (rechts). Uit deze afbeelding blijkt dat de lineaire regressietoets bij normaliteit enigszins krachtiger is dan de Mann-Kendalltoets, maar bij niet-normaliteit (dit voorbeeld betreft log-normaliteit) is de Mann-Kendalltoets veel krachtiger dan de lineaire regressietoets.



Afbeelding 2: Vergelijking van het percentage trenddetectie (op te vatten als een raming van het onderscheidend vermogen) van de lineaire regressietoets (LR) en dat van de Mann-Kendalltoets (MK), als functie van de trendgrootte (uitgedrukt als fractie van de standaardafwijking van het betreffende proces). Links voor een proces met een normale kansverdeling en rechts voor een proces met een log-normale kansverdeling. De resultaten zijn afgeleid met Monte Carlosimulatie (5.000 simulaties per trendgrootte), waarbij is getoetst met 95% betrouwbaarheid. Elke gesimuleerde reeks omvatte 120 waarden (zoals bijvoorbeeld het geval bij een reeks van 10 jaar maandwaarden). Let op dat de resultaten van de linkerfiguur niet direct kunnen worden vergeleken met die van de rechterfiguur, aangezien de standaardafwijking van het log-normale proces hier groter is dan die van het normale proces.

Wat nog minder algemeen bekend is, is dat het voordeel van verdelingsvrije toetsen zelfs al begint bij geringe afwijkingen van normaliteit. Ze presteren dan al minstens zo goed als parametrische methoden, zoals ook aangetoond in simulatiestudies (zie bijvoorbeeld [Hirsch e.a., 1991; Önöz and Bayazit, 2003]). Maar kleine afwijkingen van normaliteit zijn slechts te detecteren met grote steekproeven, die zelden beschikbaar zijn. Deze overwegingen hebben er toe geleid dat het in het milieuonderzoek - waar veel scheve kansverdelingen voorkomen - vrij gebruikelijk is geworden om verdelingsvrije statistische methoden te hanteren (zie bijvoorbeeld [Gilbert, 1987] en [Helsel and Hirsch, 1991]).

En transformeren?

Sommige onderzoekers kiezen er voor om bij statistische analyse van meetwaarden uit een scheve kansverdeling eerst een wiskundige transformatie toe te passen op de meetwaarden - zoals een log-transformatie - en vervolgens parametrische methoden toe te passen. Wij geven bij niet-normaliteit echter de voorkeur aan verdelingsvrije methoden. De belangrijkste reden hiervoor is dat het zelden lukt om bij niet-normaliteit succesvol te transformeren naar normaliteit, zodat er onderscheidend vermogen verloren gaat als dan toch een parametrische toets op de getransformeerde waarden wordt toegepast.⁸ En verder leidt transformeren bij trendanalyse tot 'kromme' trends, die zich niet makkelijk laten vergelijken met de rechte trends die in dezelfde trendstudie zijn geschat voor reeksen die niet getransformeerd zijn.

⁸ Hetzelfde argument kan worden aangevoerd tegen het toepassen (op de ongetransformeerde waarden) van statistische methoden die zijn gebaseerd op de maximum-likelihood-schatter. Dat biedt immers alleen voordeel ten opzichte van verdelingsvrije methoden als volledig wordt voldaan aan het soort kansverdeling waar de maximum-likelihood-schatter van uit gaat.

Preferentieprocessen van de trendtoetsen

Uit een combinatie van Monte Carlosimulaties en theorie is voor elk van de trendtoetsen het preferentieproces vastgesteld (zie tabel 1). Deze tabel vormt het uitgangspunt van de hieronder beschreven selectieprocedure van de toe te passen trendtoets.

Toets	Statistische kenmerken van proces							
	Normale kansverdeling			Geen normale kansverdeling				
	S	A	S+A	S	A	S+A		
LR								
LRs								
LRa								
LRsa								
MK								
MKs								
MKsa								

Tabel 1: Preferentieprocessen van de verschillende trendtoetsen.

Toelichting statistische toetsen

LR: Lineaire regressietoets | MK: Mann-Kendalltoets | s: met verdiscontering seizoens-effecten | a: met verdiscontering autocorrelatie

Toelichting statistische kenmerken van het proces dat de tijdreeks genereert

S: seizoenseffecten | A: autocorrelatie

Hoofdpijnen van onze selectieprocedure

Om voor elke reeks de best daarop aansluitende trendtoets te kunnen selecteren, is in verschillende fasen een selectieprocedure tot stand gekomen, zoals eerder toegelicht. Gezien de complexiteit van de verschillende toetsen en van de beslismomenten van de selectieprocedure is daarvoor ook programmatuur ontwikkeld. In eerste instantie betrof dit het programma Trendpak, dat is ontwikkeld in Fortran en onder MS-DOS draaide. Zijn opvolger is het programma Trendanalist, dat is ontwikkeld in Matlab. Het betreft een statistische toolbox voor trendanalyse, die zowel in handmatige als in geautomatiseerde modus kan worden gebruikt.

De selectieprocedure gaat uit van de relevante kenmerken van het proces dat de tijdreeks heeft gegenereerd, namelijk het al of niet voldoen aan de normale kansverdeling en het al of niet optreden van seizoenseffecten en/of autocorrelatie. Om deze kenmerken zo objectief mogelijk vast te kunnen stellen worden de hieronder genoemde statistische toetsen gebruikt, elk met een betrouwbaarheid van 95%.

- Toets op seizoenseffecten: Kruskal-Wallistoets. Dit is het verdelingsvrije alternatief van éénweg-variantieanalyse. De toets functioneert beter op een trendloze reeks, zodat deze wordt uitgevoerd op een hulpreeks, die ontstaat door de tijdreeks zo goed mogelijk van een eventuele trend te ontdoen, door van elke waarde de bijbehorende

waarde van de trendlijn (met de Kendall-seizoenshelling) af te trekken (zie ook verder). Aangezien de toets onafhankelijke waarden veronderstelt, wordt de hulpreeks (minstens) op kwartaalbasis gezet. Onze ervaring is dat er bij tijdsstappen van een kwartaal of groter voor veel oppervlaktewaterprocessen geen sprake meer is van een storende autocorrelatie. Voor tragere processen - zoals grondwaterprocessen - kan dat wel het geval zijn, maar daar is de meetfrequentie doorgaans al veel lager.

- Toets op normaliteit: Lilliefors-toets [Lilliefors, 1967 en 1969]. Dit is een aanpassing van de Kolmogorov-Smirnovtoets op normaliteit, voor situaties waarin het gemiddelde en de standaardafwijking van de populatie moeten worden geschat uit de steekproefgegevens.
- Toets op autocorrelatie van de residuën van een (al dan niet uitgebreid) lineair regressiemodel: Ljung-Box-Portmanteautoets [Ljung and Box, 1978], of zonodig een aanpassing daarvan die rekening houdt met ontbrekende waarden [Stoffer and Toloi, 1992]. Beide uitvoeringen baseren zich op een som van de gekwadrateerde autocorrelatiecoëfficiënten van de modelresiduën. De autocorrelatiecoëfficiënt is een maat voor de lineaire relatie tussen waarden die een bepaald tijdsinterval van elkaar liggen. De toets op autocorrelatie gaat overigens uit van de normale kansverdeling en wordt daarom pas toegepast als is gebleken dat de modelresiduën afkomstig zijn uit een normale kansverdeling.
- Toets op autocorrelatie van de tijdreekswaarden: runstoets. Dit is een verdelingssvrije toets. Aangezien het hierbij gaat om de autocorrelatie die resteert nadat de tijdreeks is ontdaan van een eventuele trend en eventuele seizoenseffecten, wordt de toets uitgevoerd op een hulpreeks, die zo goed mogelijk is ontdaan van die karakteristieken. Deze hulpreeks wordt vervaardigd door: (1) de tijdreeks eerst zo goed mogelijk van een eventuele trend te ontdoen, door van elke waarde de bijbehorende waarde van de trendlijn (met de Kendall-seizoenshelling) af te trekken; (2) indien er sprake is van seizoenseffecten (volgens de Kruskal-Wallistoets) van elke reekswaarde de bijbehorende seizoensmediaan af te trekken.

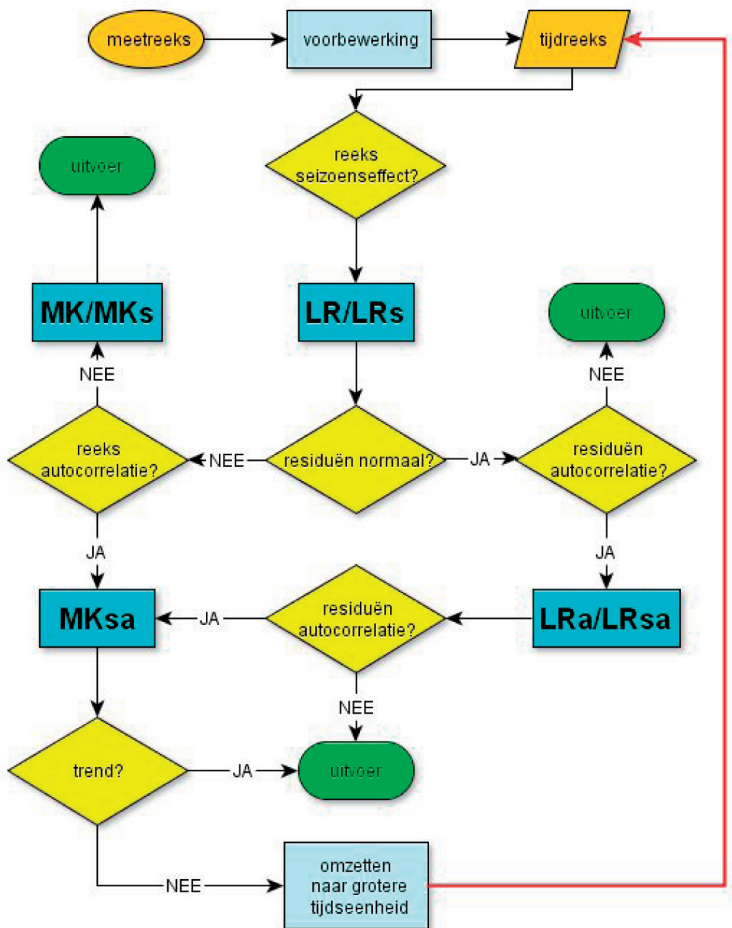
De hoofdlijnen van de selectieprocedure zijn weergegeven in afbeelding 3. Op een kleiner detailniveau wordt voor enkele toetsen nagegaan of de tijdreeks wel voldoende informatie bevat om de betreffende toets toe te kunnen passen. Anders wordt de tijdreeks omgezet naar een grotere tijdsstap en wordt de selectieprocedure opnieuw doorlopen.

Om voor trendanalyse in aanmerking te kunnen komen, hanteren we als criteria dat een meetreeks minstens 5 jaar, 5 meetwaarden en minder dan 80% gecensureerde meetwaarden moet omvatten. De nulhypothese luidt dat er geen trend is, terwijl de alternatieve hypothese luidt dat er een trend is vanaf het begin van de meetreeks. Desgewenst kan er een ander begintijdstip voor de trend worden opgegeven. Dit dient echter gebaseerd te zijn op externe informatie en níet op een visuele beoordeling van de reeks. Volgens de wetenschappelijke methode moet een hypothese immers getoetst worden met gegevens die onafhankelijk zijn van de gegevens waaruit de hypothese is afgeleid.

Eerst wordt getoetst of de reeks seizoenseffecten vertoont. Afhankelijk van het resultaat wordt de lineaire regressietoets (LR), of de lineaire regressietoets met verdiscontering van seizoenseffecten (LRs) toegepast. In beide gevallen wordt het toetsresultaat pas geldig geacht als op basis van statistische toetsing mag worden aangenomen dat de

modelresiduën afkomstig zijn uit een normale kansverdeling en tevens geen autocorrelatie vertonen. Anders wordt één van de volgende paden gevolgd:

- Modelresiduën niet normaal verdeeld en uit toetsing blijkt dat de tijdreeks geen autocorrelatie vertoont - Er wordt dan de bijpassende verdelingsvrije toets toegepast, namelijk de Mann-Kendalltoets (MK), of de Mann-Kendalltoets met verdiscontering van seizoenseffecten (MKs).
- Modelresiduën niet normaal verdeeld en uit toetsing blijkt dat de tijdreeks autocorrelatie vertoont - De Mann-Kendalltoets met verdiscontering van seizoenseffecten en autocorrelatie wordt toegepast (MKsa).
- Modelresiduën normaal verdeeld, maar ze vertonen autocorrelatie - De lineaire regressietoets met verdiscontering van autocorrelatie (LRa), of de lineaire regressietoets met verdiscontering van seizoenseffecten en autocorrelatie (LRsa) wordt toegepast. Als uit toetsing blijkt dat de modelresiduën dan nóg autocorrelatie vertonen, wordt de Mann-Kendalltoets met verdiscontering van seizoenseffecten en autocorrelatie (MKsa) toegepast. Een resterende autocorrelatie kan immers zijn veroorzaakt door een niet-lineaire trend en de MKsa-toets kan daar beter mee overweg.



Afbeelding 3: Hoofdpijnen van de procedure om de meest geschikte trendtoets voor een meetreeks te selecteren.

De verschillende onderdelen van de selectieprocedure zijn onderscheiden door vormen en kleuren:
| oranje ellips: invoer | oranje parallellogram: bewerkte invoer | lichtblauwe rechthoek: actie | gele ruit: beslismoment | blauwe rechthoek: trendtoets | groene ellips: uitvoer |
Trendtoetsen: | **LR**: lineaire regressietoets | **MK**: Mann-Kendalltoets | **s**: met verdiscontering seizoenseffecten | **a**: met verdiscontering autocorrelatie | **sa**: met verdiscontering seizoenseffecten en autocorrelatie |

Rekening houden met gering onderscheidend vermogen MKsa-toets

Voor een aantal tijdreeksen zal bij het doorlopen van de selectieprocedure de Mann-Kendalltoets met verdiscontering van seizoenseffecten en autocorrelatie (MKsa) de meest geschikte trendtoets blijken. Maar als deze toets geen statistisch significante trend detecteert, wordt de tijdreeks omgezet naar de eerstvolgende grotere tijdsstap en wordt opnieuw de selectieprocedure doorlopen. Uit simulaties is namelijk gebleken dat het onderscheidend vermogen van deze toets in bepaalde gevallen dermate gering is, dat het loont de tijdreeks om te zetten naar een grotere tijdsstap, aangezien dan wellicht een krachtigere trendtoets van toepassing wordt [Baggelaar en Baggelaar, 1991].

Omzetten meetreeks naar tijdreeks

Alvorens de selectieprocedure in te kunnen gaan, moet de meetreeks worden omgezet naar een vorm die geschikt is voor trendanalyse. Ter onderscheiding van de meetreeks noemen wij die vorm de tijdreeks. Meetreeksen van milieuv variabelen zijn namelijk zelden direct geschikt voor trendanalyse, vanwege ongelijke meetintervallen, uitschieters en ontbrekende of gecensureerde waarden.

Uitschieters

Uitschieters zijn meetwaarden die duidelijk afwijken van de andere meetwaarden. Ze kunnen zijn veroorzaakt door: 1. uitzonderlijke bemonsterings- of meetfouten en/of transcriptiefouten (fouten bij het schrijven en/of typen), of 2. een extreme situatie. Bij statistische analyses is het van belang dat meetwaarden veroorzaakt door uitzonderlijke fouten niet meegenomen worden. Het vergt echter detailkennis van alle handelingen die tot een meetwaarde hebben geleid om objectief vast te kunnen stellen of er sprake is van een foute meetwaarde of van een meetwaarde die een extreme situatie weergeeft. Bij het ontbreken van deze detailkennis - en dit zal meer regel dan uitzondering zijn - is er dan ook geen objectieve rechtvaardiging voor het verwijderen van uitschieters en moet dit achterwege blijven. Als er grote uitschieters voorkomen, zal volgens de selectieprocedure al gauw één van de verdelingsvrije Mann-Kendalltoetsen in aanmerking komen, aangezien de modelresiduen van lineaire regressie door de uitschieter(s) de toets op normaliteit niet zullen doorstaan. Dit is echter geen probleem, daar deze verdelingsvrije toetsen zeer robuust zijn tegen uitschieters.

Gecensureerde waarden

Voor de trendanalyse worden alle gecensureerde waarden van een reeks standaard op de helft van de hoogste rapportagegrens van die reeks gezet. Ook niet-gecensureerde meetwaarden die lager zijn dan de hoogste rapportagegrens worden op de helft van die grens gezet. Deze aanpak kan informatieverlies geven, maar is nodig om kunstmatige trends door veranderingen van de rapportagegrens te vermijden (zie bijvoorbeeld [Helsel and Hirsch, 1991]).

Selectie meest geschikte tijdsstap

Elke meetreeks wordt omgezet naar een tijdreeks met als mogelijke keuzen voor de tijdsstap vier weken, maand, twee maanden, kwartaal, vier maanden, zes maanden en jaar. De redenen waarom geen trendanalyse wordt uitgevoerd op tijdreeksen met een kleinere tijdsstap dan vier weken zijn: 1) eventuele seizoenseffecten zijn dan moeilijker te detecteren en 2) de autocorrelatie kan dan zulke grote vormen aannemen, dat elke trendtoets daardoor onbruikbaar wordt.

De keuze van de tijdsstap wordt bepaald door het meest voorkomende meetinterval van de meetreeks. Het omzetten naar een tijdreeks geschiedt door alle waarden van de meetreeks die binnen de tijdsstap van de tijdreeks vallen te vervangen door hun mediaan, aangezien dat een robuustere centrummaat is dan het gemiddelde, beter bestand tegen foute meetwaarden. Als de resulterende tijdreeks 30% of meer ontbrekende waarden bevat wordt een grotere tijdsstap genomen, totdat deze minder dan 30% ontbrekende waarden bevat.

Trendkwantificering

Voor wat betreft de trendkwantificering zijn de volgende drie trendschatters beschikbaar:

1. de lineaire regressiehelling in het geval van een normale kansverdeling;
2. de Theilhelling in het geval van een niet-normale kansverdeling en geen seizoenseffecten;
3. de Kendall-seizoenshelling in het geval van een niet-normale kansverdeling en seizoenseffecten.

Voor elke reeks wordt een trend geschat, ongeacht of er met de trendtoets een statistisch significante trend in de tijdreeks is gedetecteerd. Aangezien ook bij de selectie van een trendtoets wordt afgegaan op kansverdeling, seizoenseffecten en autocorrelatie, worden altijd dezelfde combinaties van trendtoets en trendschatter gehanteerd. Deze combinaties zijn vermeld in onderstaande tabel.

Trendtoets	Trendschatter
LR, LRs, LRa of LRsa	Lineaire regressiehelling
MK	Theilhelling
MKs of MKsa	Kendall-seizoenshelling

Tabel 2: De vaste combinaties van trendtoets en trendschatter die de selectieprocedure hanteert.

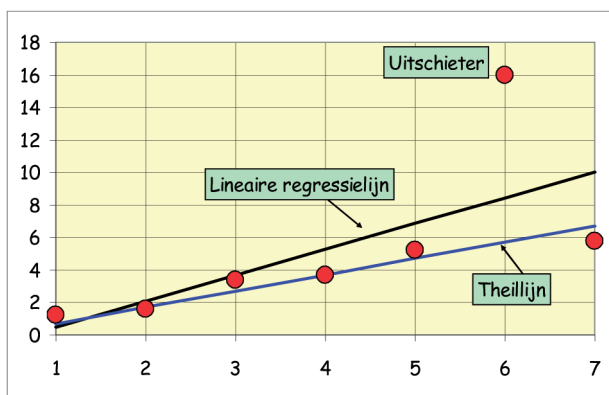
Theilhelling

In het geval van een niet-normale kansverdeling en geen seizoenseffecten wordt een trend geschat als de Theilhelling [Theil, 1950, Sen, 1968]. Dit is de mediaan van alle mogelijke hellingen tussen de afzonderlijke waarden in de tijdreeks:

$$b_{Th} = \text{mediaan} \left[\frac{Z_t - Z_k}{X_t - X_k} \text{ voor alle } 1 \leq k < t \leq n \right] \quad (15)$$

met b_{Th} de Theilhelling (meeteenheid/jaar), Z de waarden van de beschouwde variabele (meeteenheid), X de tijd (jaar, met ook de jaarfractie), t en k waarden van de tijdsindex en n het aantal waarden in de reeks. Er zijn $n \cdot (n-1) / 2$ hellingen te bepalen bij n tijdreekswaarden.

De Theilhelling is op te vatten als een verdelingsvrije schatter, die vrij resistent is tegen de invloed van extremen, omdat de mediaan van alle individuele hellingen wordt gebruikt (zie het voorbeeld in afbeelding 4).



Afbeelding 4: Illustratie van de robuustheid van de Theilhelling ten opzichte van de lineaire regressiehelling bij een uitschieter in de reeks.

Verder is deze hellingschatter zuiver - dat wil zeggen zonder systematische fout - en met een grotere nauwkeurigheid dan die van de lineaire regressiehelling bij tijdreeksen met waarden afkomstig uit een scheve kansverdeling. Dit speelt zelfs al bij geringe afwijkingen van normaliteit, die op steekproefniveau praktisch niet visueel zijn te onderkennen. Dit laatste is bijvoorbeeld met een simulatiestudie geïllustreerd door [Hirsch e.a., 1991].

Kendall-seizoenshelling

Als er sprake is van een niet-normale kansverdeling en er ook seizoenseffecten optreden, wordt een trend geschat als de Kendall-seizoenshelling [Sen, 1968; Hirsch e.a., 1982], gedefinieerd als:

$$b_{Ks} = \text{mediaan} \left[\frac{Z_{tj} - Z_{kj}}{X_t - X_k} \text{ voor alle } 1 \leq k < t \leq n_j \text{ en voor } j = 1, \dots, s \right] \quad (16)$$

met b_{K_S} de Kendall-seizoenshelling (meeteenheid/jaar), t en k waarden van de jaar-index, j de seizoensindex, n_j het aantal jaar met een tijdreekswaarde in seizoen j en s het aantal tijdreekswaarden in een jaar (bij analyse op maandschaal is s bijvoorbeeld 12). Het is dus de mediaan van alle hellingen tussen tijdreekswaarden die precies een heel aantal jaren uit elkaar liggen.

Deze schatter wordt niet beïnvloed door seizoenseffecten, omdat de hellingen slechts worden bepaald voor waarden uit hetzelfde seizoen, die één of meer jaren uit elkaar liggen. Het betreft een zuivere schatter en met een grotere nauwkeurigheid dan die van de lineaire regressiehelling bij tijdreeksen met seizoenseffecten, autocorrelatie en een scheve kansverdeling [Hirsch e.a., 1982].

Trendlijn

Om de geschatte trend van een tijdreeks grafisch te kunnen weergeven in een trendlijn, is tevens een schatting van het intercept (b_0) nodig. Dit is het snijpunt van de trendlijn met de Y-as voor het nulpunt van X (de tijd).

Als er sprake is van een normale kansverdeling, dan maakt de schatting van het intercept al deel uit van de schatting van het lineaire regressiemodel. Als er daarbij ook sprake is van een seizoenseffect en het uitgebreide lineaire regressiemodel is gebruikt, dan moet het geschatte intercept worden gecorrigeerd voor het daarin vervatte seizoenseffect.

Als er geen sprake is van een normale kansverdeling, dan wordt het intercept verdelingsvrij geschat, volgens [Conover, 1980]:

$$b_0 = \text{mediaan}[Z_t \text{ voor } 1 \leq t \leq n] - b_1 \cdot \text{mediaan}[X_t \text{ voor } 1 \leq t \leq n] \quad (17)$$

met Z de waarden van de beschouwde variabele (meeteenheid), X de tijd (jaar, met daarbij ook de jaarfractie), t de tijdsindex, n het aantal waarden in de reeks en b_1 de geschatte helling (dit is dan de Theilhelling of de Kendall-seizoenshelling).

Met het geschatte intercept en de geschatte helling kan tenslotte de trendlijn worden geconstrueerd, volgens:

$$\hat{Z}_t = b_0 + b_1 \cdot X_t \quad (18)$$

Toepassingen

De procedure wordt regelmatig toegepast op gegevens van grootschalige meetnetten waterkwaliteit, zoals van Rijkswaterstaat, de RIWA, een tiental waterschappen⁹, een drinkwaterbedrijf, de Provincie Drenthe en de Vlaamse Milieumaatschappij. Door de omvang van deze meetnetten levert elke toepassing zeer veel uitvoer op en het is dan steeds weer een uitdaging die zodanig gestructureerd te presenteren dat interpretaties worden vergemakkelijkt. Verschillende presentatievormen hebben daarbij inmiddels al hun waarde bewezen, zoals het trendpalet (een matrix van meetlocaties en parameters, met per cel het resultaat van de trendtoets aangegeven met een stoplichtkleur), boxplots van geschatte hellingen, zo mogelijk met onderscheid naar één of andere indelings-

⁹ In één geval is ook een trendanalyse uitgevoerd van de kwaliteit van het zuiveringsslib van afvalwater-zuiveringsinstallaties.

factor (zoals landgebruik, bodemtype, hydrologische situatie, etc.), gestapelde staafdiagrammen, taartdiagrammen en trendkaarten, met toestand- en trendpictogrammen. Voor illustraties hiervan kunnen we bijvoorbeeld verwijzen naar [Baggelaar e.a., 2011]. Verder zijn ook verschillende soorten toetsende statistische analyses van de resultaten mogelijk, zoals (verdelingsvrije) variantieanalyse van de geschatte hellingen, om na te gaan in hoeverre deze verschillen voor de niveaus van een bepaalde indelingsfactor. De resultaten van een trendanalyse kunnen uiteraard ook aanleiding geven tot vervolgonderzoek. Zo kan worden besloten de trendanalyse te herhalen ná correctie van de meetreeksen voor natuurlijke invloeden, zoals waterafvoer en/of –temperatuur, zodat antropogene invloeden duidelijker naar voren kunnen komen. Maar er kan ook worden besloten om een hypothese van antropogene beïnvloeding te toetsen via tijdreeksmodellering, met bijvoorbeeld een in het verleden toegepaste maatregel als verklarende variabele (al dan niet als [0,1]-variabele).

Een technische bespreking van al deze mogelijkheden voert echter te ver voor de strekking van dit artikel.

Conclusies

Het loont om bij trendanalyse maatwerk uit te voeren en per meetreeks de best daarop aansluitende trendtoets en trendschatter toe te passen. Dit levert meer onderscheidend vermogen bij het toetsen op trend en een grotere nauwkeurigheid bij het schatten van de trend. Hiertoe dient de selectie van trendtoets en trendschatter rekening te houden met de statistische karakteristieken van de betreffende reeks, in het bijzonder het soort kansverdeling en het al of niet optreden van seizoenseffecten en autocorrelatie. Bij niet-normaliteit komen verdelingsvrije methoden in aanmerking, aangezien die zelfs bij geringe afwijkingen van normaliteit al superieur zijn aan parametrische methoden en bovendien niet of nauwelijks nadelig worden beïnvloed door uitschieters. De selectieprocedure en de daarop aansluitende trendanalyse zijn zodanig geobjectiveerd, dat ze geheel geautomatiseerd kunnen worden uitgevoerd. Dit is van belang gezien het grote aantal meetreeksen dat meetnetten kunnen bevatten.

Literatuur

Baggelaar, P.K. en Baggelaar, D.H. (1989) Project Trendbepaling. Rapportage 1e fase; Kiwa-rapport SWO 89.335, augustus 1989, Nieuwegein.

Baggelaar, P.K. en Baggelaar, D.H. (1991) Project Trendbepaling. Rapportage 2e fase; Kiwa-rapport SWO 90.284, februari 1991, Nieuwegein.

Baggelaar, P.K. (1994) A protocol for trend analysis of the input of pollutants to the North Sea; Kiwa-rapport SWO 93.337, mei 1994, Nieuwegein.

Baggelaar, P.K. en Baggelaar, D.H. (1995) Trends in de oppervlaktewaterkwaliteit van Rijn en Maas; Rapport RIWA en Kiwa, Amsterdam en Nieuwegein, mei 1995.

Baggelaar, P.K., Van der Meulen, E.C.J. en Pot, R. (2011) Trendanalyse meetnetten Hoogheemraadschap van Rijnland; Icastat-AMO-Roelf Pot, mei 2011, 157 blz.

- Baggelaar, P.K. en Van der Meulen, E.C.J.** (2012) Handleiding Trendanalist; Icastat-AMO, februari 2012, 43 blz.
- Box, G.E.P. and Jenkins, J.M.** (1976) Time series analysis, forecasting and control; Holden-Day, San Francisco.
- Bradley, J.V.** (1968) Distribution-Free Statistical Tests; Prentice Hall, Englewood Cliffs.
- Conover, W.J.** (1980) Practical nonparametric statistics; John Wiley, New York.
- Dietz, E.J. and Killeen, T.J.** (1981) A Nonparametric Multivariate Test for Monotone Trend with Pharmaceutical Applications; Journ. of the Amer. Stat. Assoc., 76 (373), blz 169-174.
- Duin, R.N.M.** (1993) Trendanalyse van verschillende waterkwaliteitsvariabelen op 26 lokaties in de Noordzee over de periode 1975 t/m 1989; Dienst Getijdewateren, Rijkswaterstaat.
- Farrel, R.** (1980) Methods for classifying changes in environmental conditions; Techn. Rep. VRF-EPA7.4-FR80-1, Vector Res. Inc., Ann Arbor, Michigan.
- Gilbert, R.O.** (1987) Statistical Methods for Environmental Pollution Monitoring; Van Nostrand Reinhold, New York, 320 blz.
- Harcum, J.B., Loftis, J.C. and Ward, R.C.** (1992) Selecting trend tests for water quality series with serial correlation and missing values; Water Resources Bulletin, vol 28, no 3, blz 469-478.
- Helsel, D.R. and Hirsch, R.M.** (1988) Discussion of paper by R.H. Montgomery and J.C. Loftis: Applicability of the t-test for detecting trends in water quality variables; Water Resources Bulletin, vol 24, blz 201-204.
- Helsel, D.R. and Hirsch, R.M.** (1991) Statistical Methods in Water Resources; Studies in Environmental Science 49. Elsevier, Amsterdam, 510 blz.
- Hirsch, R.M., Slack, J.R. and Smith, R.A.** (1982) Techniques of trend analysis for monthly water quality data; Water Resources Research, vol 18, no 1, February 1982, blz 107-121.
- Hirsch, R.M. and Slack, J.R.** (1984) A nonparametric trend test for seasonal data with serial dependence; Water Resources Research, vol 20, no 6, June 1984, blz 727-732.
- Hirsch, R.M., Alexander, R.B. and Smith, R.A.** (1991) Selection of methods for the detection and estimation of trends in water quality; Water Resources Research, vol 27, no 5, May 1991, blz 803-813.
- Kendall, M.G.** (1938) A new measure of rank correlation; Biometrika, 30, 1938, blz 81-93.
- Kendall, M.G.** (1975) Rank Correlation Methods; Charles Griffin, London, 1975.
- Lehmann, E.L.** (1975) Nonparametrics: Statistical Methods Based on Ranks; Holden-Day, Inc., San Francisco.
- Lettenmaier, D.P.** (1976) Detection of trend in water quality data from records with dependent observations; Water Resources Research, vol 12, no 5, blz 1037-1046.

- Lilliefors, H.W.** (1967) On the Kolmogorov-Smirnov test for normality with mean and variance unknown; Journal of the American Statistical Association, 62, blz 399-402.
- Lilliefors, H.W.** (1969) Correction to the paper: On the Kolmogorov-Smirnov test for normality with mean and variance unknown; Journal of the American Statistical Association, 64, blz 1702.
- Ljung, G.M and Box, G.E.P** (1978) On a measure of lack of fit in time series models; Biometrika, 65, blz 297-303.
- Mann, H.B.** (1945) Non-parametric tests against trend; Econometrica 13, blz 245-259, 1945.
- Montgomery, R.H. and Reckhow, K.H.** (1984) Techniques for detecting trends in lake water quality; Water Resources Bulletin, vol 20, no 1, february 1984.
- Önöz, B. and Bayazit, M.** (2003) The power of statistical tests for trend detection; Turkish J. Eng. Env. Sci., 27, blz 247-251.
- Sanders, T.G., Ward, R.C., Loftis, J.C., Steele, T.D., Adrian, D.D. and Yevjevich, V.** (1983) Design of networks for monitoring water quality; Water Resources Publications, Littleton, USA.
- Sen, P.K.** (1968) Estimates of the regression coefficient based on Kendall's tau; Journ. Am. Statist. Assoc., 63, blz 1379-1389, 1968.
- Stoffer, D.S. and Toloi, C.M.C.** (1992) A note on the Ljung-Box-Pierce portmanteau statistic with missing data; Statistics & probability Letters 13 (1992), blz 391-396.
- Theil, H.** (1950) A rank-invariant method of linear and polynomial regression analysis, 1, 2 and 3; Ned. Akad. Wetensch. Proc., 53, blz 386-392, 521-525 en 1397-1412.
- Van Belle, G. and Hughes, J.P.** (1984) Nonparametric tests for trend in water quality; Water Resources Research, vol 20, no 1, blz 127-136.
- Ward, R.C. and Loftis, J.C.** (1986) Establishing statistical design criteria for water quality monitoring systems: review and synthesis; Water Resources Bulletin, vol 22, no 5, October 1986, blz 759-767.