

Hoe doen we tijdreeksanalyse van grondwaterstanden in Nederland?

Verlag van een workshop

Op 5 december 2001 was er te Oosterhout een workshop over tijdreeksanalyse van grondwaterstanden (en stijghoogten), georganiseerd door Brabant Water en Icastat. Deelname was op uitnodiging en er waren 31 personen aanwezig, grotendeels ervaren of aankomende tijdreeksanalisten, aangevuld met enkele beoordelaars van resultaten van tijdreeksanalyse. De workshop was bedoeld als eerste aanzet tot een structurele vorm van directe uitwisseling van ervaringen en ideeën, om zo de mogelijkheden én beperkingen van de tijdreeksanalyse beter in kaart te kunnen brengen. Een andere doelstelling was om – waar mogelijk – tot meer consensus in aanpak te komen. De resultaten van een tijdreeksanalyse kunnen namelijk sterk afhangen van de gevolgde aanpak, maar welke is het meest geschikt, gezien de situatie en de vraagstelling? Er is daarover nog geen consensus onder de tijdreeksanalisten, vermoedelijk ook doordat zij nog te weinig ervaringen met elkaar hebben uitgewisseld. Dit werkt uiteraard verwarrend voor opdrachtgevers en voor beoordelaars van tijdreeksanalyse en kan uiteindelijk zelfs het aanzien van deze techniek ondergraven.

Er zijn verschillende toepassingsmogelijkheden van tijdreeksanalyse, zoals voorspellen, of simuleren (scenario-analyse), maar deze eerste workshop richtte zich vooral op het vaststellen van causale relaties met tijdreeksanalyse.

Enige historie

In Nederland zijn sinds ongeveer 1970 verschillende vormen van tijdreeksanalyse toe-

gepast om te reconstrueren hoe grondwaterstanden en stijghoogten beïnvloed zijn door factoren zoals de meteorologische omstandigheden, grondwateronttrekkingen en waterhuishoudkundige ingrepen. Het leek een verfrissend andere benadering om nu eens de gegevens 'vrijuit' te laten spreken en ze niet slechts te gebruiken als ijkmetaal voor een op de theorie gebaseerd model. De eerste 15 jaar betrof het nog slechts rudimentaire vormen van tijdreeksanalyse, maar vanaf circa 1985 is geleidelijk overgegaan op geavanceerdere vormen, zoals de Box–Jenkins-analyse, die beter aansluiten op de karakteristieken van de onderzochte tijdreeksen, met name seizoenseffecten en seriële afhankelijkheid.

Inmiddels hebben we in Nederland een selecte groep beoefenaren van deze 'kunst', die we niet alleen aantreffen bij onderzoeksinstituten, zoals Kiwa, TNO-NITG en Alterra, maar ook bij verschillende ingenieursbureaus en drinkwaterbedrijven. Vermoedelijk hebben we in Nederland vergeleken bij andere landen een relatief groot aantal hydrologische tijdreeksanalisten. Dit zal te maken hebben met volgende twee factoren: 1) de ruimtedruk, waardoor er een grote belangenverstrengeling rond het grondwater optreedt en de beïnvloeding van dat grondwater voortdurend scherp in de gaten wordt gehouden; 2) het feit dat er in Nederland zoveel geschikte tijdreeksen van grondwaterstanden en stijghoogten beschikbaar zijn.

De workshop

Om een beeld te kunnen krijgen van de huidige diversiteit in aanpak, was de deelnemers verzocht om vóóraf hun antwoorden op een aantal vragen – voornamelijk van technische aard – in te sturen. De resultaten van deze mini-enquête werden tijdens de workshop door mij gepresenteerd. Daarbij zijn van elke vraag eerst kort de achter-

gronden toegelicht en geplaatst in het kader van de probleemstelling, zo mogelijk met een praktijkvoorbeeld, waarna het overzicht van de antwoorden werd gepresenteerd. Deze vorm bleek gelukkig goede uitgangspunten voor de groepsdiscussie op te leveren, die tevens werd bevorderd door de knusse locatie: een klein soort amfitheater op de productielocatie Oosterhout van Waterleidingmaatschappij Noord-West Brabant¹. Ander voordeel van de voorbereiding was dat deze eerste workshop daardoor kort (circa 3 uur) en krachtig kon blijven.

Gehanteerde invoervariabelen

Uit de antwoorden bleek dat er door de aanwezigen al een groot aantal verschillende invoervariabelen is gehanteerd bij het statistisch modelleren van grondwaterstands- of stijghoogtereeksen. Het potentieel neerslagoverschot bleek daarbij het meest gehanteerd, op de voet gevolgd door een (naburige) grondwaterwinning. Andere veel gebruikte invoervariabelen zijn de grondwaterstand of stijghoogte in een ander peilfilter en de oppervlaktewaterafvoer. Verder zijn gehanteerd het bewerkte potentieel neerslagoverschot (bijvoorbeeld de uitvoer van een model van de onverzadigde zone), de neerslag en de verdamping afzonderlijk (dus niet samengevoegd tot potentieel neerslagoverschot), de afstand tot het zwaartepunt van de winning (die kan namelijk veranderen als pompputten worden bijgeplaatst), een ruilverkaveling, een anti-verdrogingsmaatregel, het stuwbeheer en zelfs de temperatuur. Als een invloedsfactor moeilijk of niet kwantificeerbaar is, zoals een ruilverkaveling of een anti-verdrogingsmaatregel, dan wordt deze doorgaans geschematiseerd tot een [0,1]-variabele (ook

wel een dummy-, aan/uit- of indicator-variabele genoemd).

Potentieel neerslagoverschot of stambuis?

Waar het ging om het verdisconteren van de meteorologische invloed op de te modelleren grondwaterstand, had circa 60% aangegeven daar bij voorkeur het potentieel neerslagoverschot voor te hanteren. De resterende 40% had daarentegen een voorkeur voor de grondwaterstand in een stambuis, dat wil zeggen een grondwaterstand die representatief mag worden geacht voor de "ongestoorde" situatie in het gebied. De discussie over deze opties is overigens al heel oud, maar wordt nog steeds levendig gevoerd. Het potentieel neerslagoverschot heeft als voordeel een redelijk onverdachte vertegenwoordiger te zijn van de meteorologische omstandigheden, maar heeft als groot nadeel dat het slechts een theoretische benadering vormt van de grootte die écht invloed heeft op de grondwaterstand, namelijk de grondwateraanvulling. Een ander nadeel is dat de relatie tussen het neerslagoverschot en de grondwaterstand soms vrij complex, zelfs niet-lineair kan zijn. In een dergelijk geval kan het alleszins de moeite lonen om toch een geschikte stambuis te zoeken, zoals later in de workshop ook werd geïllustreerd met een voorbeeld (zie verder).

Gehanteerde tijdseenheden

De halve-maand en de maand blijken veel gebruikt als tijdseenheid bij de tijdreeksanalyse van grondwaterstanden en stijghoogten. Maar een groot deel van de aanwezigen had aangegeven de tijdseenheid in principe te laten afhangen van de omstandigheden, zoals de doelstelling van de analyse, de te verwachten procesdynamiek

¹ Inmiddels samengegaan met Waterleidingmaatschappij Oost-Brabant tot Brabant Water.

(responstijden²), de kwaliteit van de gegevens en modelkarakteristieken. Criteria die daarbij een rol kunnen spelen zijn het aantal benodigde modelparameters, de standaardfouten van geschatte evenwichtsrelaties en de pasvorm van het model (uitgedrukt in de standaardafwijking van de modelresiduën).

Modelidentificatie

Een tijdreeksmodel wordt ontwikkeld volgens de cyclus: (1) identificeren model, (2) schatten modelparameters en (3) verifiëren of wordt voldaan aan de eisen om statistisch verantwoorde uitspraken op basis van het model te kunnen doen. Belangrijke eisen hiervoor zijn dat de modelresiduën afkomstig zijn uit dezelfde normale kansverdeling met gemiddelde nul en dat ze onafhankelijk van elkaar zijn. De wijze waarop het tijdreeksmodel wordt geïdentificeerd kan een grote invloed hebben op de resultaten, maar er is nog onvoldoende consensus over wat daarvoor nou de te verkiezen insteek is. Waar het bijvoorbeeld gaat om het vaststellen van de invloed van een grondwaterwinning op een grondwaterstand, levert het kruiscorrelogram van deze twee variabelen – een grafische weergave van hun statistische relatie als functie van de onderlinge tijdsverschuiving – vaak nog allerlei ruimte tot verschillende interpretaties. Ter illustratie werd een kruiscorrelogram getoond dat kon duiden op een responstijd die ergens tussen 2 en 5 maanden ligt. Maar uit de geohydrologische situatie kon met een analytische formule worden afgeleid dat de responstijd wel 12 maanden of meer bedroeg. En doordat deze verschillende responstijden tot duidelijke verschillen leiden in de uit-

eindelijk geschatte evenwichtsrelatie tussen de grondwaterwinning en de grondwaterstand, mag de keuze van de in het model te hanteren responstijd cruciaal worden genoemd.

Volgens de mini-enquête gaat circa 10% van de analisten bij de modelidentificatie bij voorkeur louter uit van het kruiscorrelogram (“wat ik zie”), terwijl circa 40% bij voorkeur louter uitgaat van proceskennis (“wat ik weet”). De resterende 50% van de analisten combineert deze insteken. Dit onderdeel kent dus duidelijke verschillen in aanpak, die ook tot verschillen in resultaten kunnen leiden. Ik ben inmiddels van mening dat die verschillen in aanpak geen probleem hoeven op te leveren, als we ons maar realiseren dat ze verschillende doelen dienen en onze aanpak ook duidelijk bij de resultaten vermelden. Want als we het tijdreeksmodel bijvoorbeeld identificeren op basis van wat de beschikbare gegevens ons in het kruiscorrelogram laten “zien” over de onderlinge statistische relatie, dan zijn we in feite met een *explorerende* analyse bezig (hypothese-vormend). Op zich is daar niets mis mee, maar het zou bij een dergelijke aanpak misleidend zijn om vervolgens op basis van het geschatte model ook nog keiharde uitspraken te doen over de statistische significantie van een relatie, want dat zou er dan op neer komen dat het vormen én het toetsen van de hypothese met dezelfde gegevens zijn uitgevoerd, wat neerkomt op het uitvoeren van een schijntoets. Als we daarentegen het tijdreeksmodel identificeren op basis van onze proceskennis, verkregen met andere gegevens (zoals resultaten van geologische boringen en pompproeven), dan voldoet onze aanpak wél aan de wetenschappelijke methode. We toetsen onze hypothese over de statistische relatie dan immers met andere gegevens dan die waar we de hypothese op hebben gebaseerd. We kunnen dit aanduiden als *bevestigende* analyse. Uitspraken over de

² De responstijd is de tijd die het duurt voordat de invloed van een verandering van een invloedsfactor op de grondwaterstand of stijghoogte is uitgewerkt.

statistische significantie van een relatie hebben bij een bevestigende analyse uiteraard veel meer zeggingskracht dan bij een explorerende analyse. Een ander voordeel van de bevestigende analyse is dat deze meer bescherming biedt tegen het beschrijven van een niet-causale relatie met het model. Een opdrachtgever of beoordelaar zal er dan ook zeker mee gediend worden als we de door ons gevolgde insteek duidelijk in de rapportage vermelden.

De discussie leverde ook nog een goede tip op over een manier om een niet-lineaire relatie te kunnen identificeren: zet de meetwaarden van de uitvoervariabele en een invoervariabele tegen elkaar uit in een spreidingsdiagram en doe dat voor verschillende onderlinge tijdsverschuivingen. Met de gebruikelijke manier – inspectie van het kruiscorrelogram – kunnen alleen lineaire relaties worden geïdentificeerd.

De koffiepauze werd verrijkt met een toepasselijke Sinterklaas-verrassing van Martin Knotters (Alterra), die trakteerde op heerlijke banketletters, nog ter viering van zijn promotie, een maand eerder. Een bespreking van zijn proefschrift – over gerealiseerde tijdreeksmodellen voor grondwaterstanden - vindt u overigens in het vorige nummer van stromingen (jaargang 7, nummer 4, blz. 59–63).

Intermezzo: hydrologische tijdreeks-analyse van een continue meetreeks

Na de pauze gaf Jos von Asmuth (Kiwa Water Research) een korte toelichting op zijn promotie-onderzoek naar een alternatieve vorm van tijdreeksanalyse, die vooral toegesneden is op het beschrijven van hydrologische processen en die uitgaat van continue in plaats van discrete tijd. Doordat de relaties worden beschreven met analytische functies die aansluiten op de te verwachten hydrologische processen, past deze

aanpak bij een *bevestigende* analyse, zoals gedefinieerd in de voorgaande paragraaf. En door uit te gaan van continue tijd worden de beschikbare gegevens maximaal benut en worden problemen door ontbrekende meetwaarden en niet-overeenstemmende meetfrequenties omzeild. In het vorige en in dit nummer van Stromingen treft u twee artikelen aan van Jos, Kees Maas (Kiwa) en Marc Bierkens (Alterra) over deze vorm van tijdreeksanalyse.

Modelselectie

Alle analisten hadden aangegeven het belangrijk te vinden dat een model robuust is, oftewel dat een kleine verandering in de modelvorm, zoals in de responstijd, geen duidelijk ander resultaat oplevert. En als er meerdere modellen geschikt lijken, dan wordt door circa 33% van de analisten bij de modelselectie vooral afgegaan op kwantitatieve criteria, zoals de pasvorm (uitgedrukt in de standaardafwijking van de modelresiduen of het percentage verklaarde variantie), de statistische significantie van de modelparameters, de correlaties van de modelparameters, een informatiecriterium, zoals dat van Akaike of Bayes, de robuustheid van het model of validatieresultaten. Door circa 30% wordt bij de modelselectie vooral afgegaan op hydrologisch inzicht en door 33% op een combinatie van kwantitatieve criteria en hydrologisch inzicht. Verder werd ook nog als mogelijkheid genoemd het principe van equifinaliteit te hanteren. Want waarom zou je slechts één model kiezen, als er meerdere geschikt zijn? Je gooit dan immers bruikbare informatie weg. Volgens deze redenering zouden uitspraken ook niet alleen moeten worden gebaseerd op de binnen-model betrouwbaarheid van één model, maar zou daar ook de tussen-model betrouwbaarheid van alle in principe geschikte modellen bij moeten worden betrokken.

Alle geënuquëterden hadden instemmend geantwoord op de vraag of het zinvol zou zijn als de analist bij de rapportage op één of andere wijze melding maakt van de zeggingskracht die elk van de modellen - en daarmee ook de op die modellen gebaseerde uitspraken en de bijbehorende gekwantificeerde betrouwbaarheden - volgens hem/haar heeft. Dit zou dan neerkomen op een oordeel gebaseerd op alle indrukken die zijn opgedaan bij het ontwikkelen van de modellen en het aansluitend controleren van de resultaten, zoals op robuustheid en op aannemelijkheid van causaliteit (zie ook verder). Maar het vermelden van dit oordeel is helaas nog weinig gebruikelijk, wellicht doordat harde criteria ontbreken en men nog terugschrikt voor het overbrengen van een subjectief oordeel, ook al kan een opdrachtgever of beoordelaar hier enorm mee geholpen zijn.

Causaliteit

Een inspirerend discussiepunt betrof de vraag hoe we de analyse kunnen toespitsen op causale relaties en die onderscheiden van de louter statistische, oftewel niet-causale relaties. Tijdreeksanalyse is namelijk sterk gericht op *statistische* relaties. Maar waar het gaat om het vaststellen van *causale* relaties kan die sterke gerichtheid op statistische relaties juist roet in het eten gooien. Het komt dan ook regelmatig voor dat er sterk moet worden getwijfeld aan de zeggingskracht van resultaten, wellicht omdat bij nadere studie blijkt dat de uitkomsten bizar, of zelfs fysisch onmogelijk zijn, óf omdat blijkt dat een gering gewijzigde aanpak, die evengoed verdedigbaar is, tot duidelijk andere uitkomsten leidt (gebrek aan robuustheid).

Om het causaliteits-dilemma te illustreren werd eerst het voorbeeld gepresenteerd van de relatie tussen een stijghoogte en een grondwaterwinning, waarbij de peilbuis en

de winplaats zich op 117 km afstand bevinden. Die relatie bleek toch statistisch significant, ondanks het feit dat een causale relatie door de afstand is uitgesloten. Een dergelijk voorbeeld doet het altijd goed om het welhaast mystieke aura weg te nemen dat beginnende analisten de techniek in hun enthousiasme nog toedichten. Vervolgens werd het voorbeeld gepresenteerd van een groot verlagingsonderzoek in de omgeving van de winplaats Seppe (Noord-Brabant), waarbij een 50-tal grondwaterstands- en stijghoogtereeksen is geanalyseerd om de invloed van de winning vast te stellen. In eerste instantie zijn bij het modelleren het potentieel neerslagoverschot en de winning Seppe als invoervariabelen gehanteerd. Het resulterende verlagingsbeeld bleek echter niet alleen onlogisch - vooral in het freatische watervoerende pakket ontbrak er een logisch verband tussen de geschatte evenwichtsrelatie met de winning en de afstand tot de winning -, maar in veel gevallen zelfs fysisch onmogelijk. De gereconstrueerde onbeïnvloede grondwaterstanden, dat wil zeggen zónder de geschatte invloed van de winning, bleken namelijk vaak vér boven maaiveld uit te komen, terwijl het ondenkbaar is dat dit vroeger het geval was. De meest waarschijnlijke verklaring voor deze onbruikbare resultaten is een ongelukkige combinatie van omstandigheden, namelijk: 1) een niet-lineaire reactie van de grondwaterstand op het neerslagoverschot en 2) de correlatie tussen het neerslagoverschot en de winning Seppe. Als we onder dergelijke omstandigheden als enige invoervariabele het potentieel neerslagoverschot zouden hanteren, zal het (lineaire) model slechts een soort gemiddelde reactie van de grondwaterstand op het neerslagoverschot kunnen beschrijven en dus geen goede pasvorm hebben, maar als we vervolgens tevens de winning Seppe toevoegen als invoervariabele, verkrijgt het model ineens veel meer flexibiliteit om de niet-lineaire relatie tussen de grondwaterstand en het

neerslagoverschot te beschrijven: de winning Seppe lijkt immers op dat neerslagoverschot. In een dergelijke situatie is het dus vrijwel onvermijdelijk dat de winning een deel van de relatie van de grondwaterstand met het neerslagoverschot gaat beschrijven, ook al heeft de winning geen enkele invloed.

Daarom zijn in tweede instantie bij het modelleren de grondwaterstand in een stambuis en de winning Seppe als invoervariabelen gehanteerd. Het resulterende verlagingenbeeld bleek nu wél logisch en ook fysisch aannemelijk. Blijkbaar was de complexe, niet-lineaire relatie tussen het neerslagoverschot en de te modelleren grondwaterstand nu voldoende verdisconteerd door de grondwaterstand in de stambuis als invoervariabele in het model op te nemen. De schatting van de relatie met de winning Seppe raakte daardoor niet meer sterk vertekend. Dit is dus een duidelijk voorbeeld van een situatie waar het modelleren met de grondwaterstand in een stambuis te verkiezen is boven het modelleren met het potentieel neerslagoverschot.

Bovenstaande voorbeelden illustreren hoe sterk tijdreeksanalyse is gericht op *statistische* relaties en ons eigenlijk slechts in staat stelt dáárover objectieve uitspraken te doen. We moeten ons daarom steeds blijven realiseren dat elke uitspraak die we vervolgens doen over causaliteit subjectief is en geheel voor onze rekening blijft!

Uit de mini-enquête bleek dat circa 55% van de analisten al wel een soort controle uitvoert op de resultaten van de tijdreeksanalyse alvorens tot uitspraken over causaliteit te komen. Een veelzeggend hulpmiddel dat hierbij werd genoemd was het 'boerenverstand'. Zo wordt er bijvoorbeeld op gelet dat de correlaties tussen de invoervariabelen, of tussen de geschatte evenwichtsrelaties niet té hoog zijn, of alle mogelijk relevante invloedsfactoren wel bij de modellering zijn betrokken, of het geschetste ruim-

telijke beeld fysisch voldoende plausibel is, ook gelet op de gereconstrueerde, onbeïnvloede grondwaterstanden (zoals is gedaan in bovenstaand voorbeeld) en of het geschetste beeld wel aansluit bij de resultaten van andere methoden.

Resumerend

Deze eerste workshop was goed bezocht en is – voorzover dat peilbaar is – als zeer nuttig ervaren. Een consensus over de aanpak van tijdreeksanalyse is niet bereikt, dat zou uiteraard ook teveel gevraagd zijn, zeker voor een eerste bijeenkomst. Maar het belang van de inzet van hydrologische proceskennis bij tijdreeksanalyse blijkt al wel breed gedragen te worden, zeker waar het er om gaat tot uitspraken over causaliteit te komen. Ook kwam duidelijk naar voren dat modelkeuzen goed moeten worden toegelicht en verantwoord.

De aankomende analisten en de beoordeelaars zullen door de discussies ongetwijfeld veel hebben kunnen opsteken over de 'zachte', oftewel subjectieve elementen van deze techniek, die in eerste instantie zo objectief lijkt, en ook over zijn verschillende valkuilen. Gelukkig bleek er een groot animo voor vervolgsessies, met voorkeur voor een interactieve vorm, zoals toegepast bij deze workshop. Deze uitwisseling tussen hydrologische tijdreeksanalisten krijgt dan ook ergens in de herfst van dit jaar een vervolg, ditmaal met een open uitnodiging, ondermeer te verspreiden via de LASSA³. De inspirator en gastheer van deze eerste workshop, Eric Broers van Brabant Water, verdient tenslotte een compliment voor zijn stimulerende aanmoediging en voor de prettige ambiance.

Paul Baggelaar
Icastat

³ Landelijke Studiegroep Statistiek in de Aardwetenschappen.