
Waarom doen alsof de neerslag eens per maand valt?

Het discrete Box-Jenkins- versus het continue PIRFICT-tijdreeksmodel, in theorie

Jos von Asmuth
Kees Maas
Marc F.P. Bierkens

In twee artikelen (deel 1: theorie, deel 2: praktijk) wordt het continue PIRFICT-tijdreeksmodel gepresenteerd en vergeleken met de discrete tijdreeksmodellen van Box en Jenkins (1970). Wanneer je tijd niet beschouwt als een dimensieloze index, zoals in Box-Jenkins modellen, maar rekent met de werkelijke tijd zoals we die allemaal kennen, blijkt tijdreeksanalyse eerder eenvoudiger te worden dan ingewikkelder. Wellicht is dat laatste ook een kwestie van smaak, maar werken met continue tijd heeft in ieder geval een belangrijk praktisch pluspunt: je bent af van de vervelende consequenties die de tijdsdiscretisatie met zich meebrengt. In een continu tijdsdomein zijn grondwaterstandsreeksen met onregelmatige tijdstappen opeens geen probleem meer, mogen de frequenties van de in- en uitvoerreeksen natuurlijk verschillen en ook de modelorde en parameterwaarden zijn niet langer discretisatie-afhankelijk. Kortom, je kunt gewoon rekenen met de data die je hebt en de uitkomsten die je vindt zijn karakteristieke eigenschappen van het geanalyseerde systeem. Kies je daarbij ook nog eens voor een voorgedefinieerde impulsrespons-functie, bijvoorbeeld op basis van de fysica, dan is het PIRFICT (= Predefined Impulse Response Function In Continuous Time) tijdreeksmodel geboren en hoef je als modelleur niet meer voor elke reeks een model-identificatieprocedure te doorlopen. Op weg naar een volledig geautomatiseerde tijdreeksanalyse?

Inleiding

Transferruis(TR)-modellen, een klasse van tijdreeksmodellen die een uitvoerreeks genereren met behulp van een of meerdere invoerreeksen en een ruisproces, zijn een krachtig hulpmiddel voor het analyseren van allerhande hydrologische reeksen, waaronder grondwaterstanden. Niet verwonderlijk dus, dat ze in de praktijk veel gebruikt worden voor o.a. het voorspellen van standen, het repareren of verlengen van reeksen, het analyseren van de effecten van maatregelen of het scheiden van de invloed van verschillende factoren (Van Geer e.a., 1988; Rolf, 1989; Knotters en Van Walsum, 1997). Gemakkelijk in het gebruik zijn TR-modellen echter niet. Er zitten een aantal knelpunten aan vast die vooral voor de beginnende gebruiker een hoge drempel op kunnen werpen. Zo vragen de modellen aardig

Jos von Asmuth en **Kees Maas** zijn werkzaam bij KIWA Onderzoek en Advies, Nieuwegein en TU Delft.
Marc F.P. Bierkens is werkzaam bij Alterra, Wageningen.

wat specifieke kennis van de tijdreeks-modelleur, alleen al omdat de modelleur de zogenaamde orde van het tijdreeksmodel moet kunnen identificeren. Tijdreeksmodellen werken bovendien (in principe) alleen met reeksen die met vaste tussenpozen zijn gemeten en waaruit geen enkele waarde mag ontbreken, dit alles vanwege het verschijnsel tijdsdiscretisatie.

In tijdreeksanalyse literatuur (o.a. Box en Jenkins, 1970; Hipel en McLeod, 1994; Ljung, 1999) wordt vaak weinig aandacht besteed aan het feit dat tijd een continu verschijnsel is. Tijdreeksmodellen delen de tijd simpelweg in vaste partjes in, en slaan vervolgens aan het rekenen met data die volgens diezelfde tijdsindeling beschikbaar moet zijn. Nu is dat voor iemand die bijvoorbeeld het temperatuursverloop van een haardroger wil modelleren op zich niet zo'n punt. Diegene past gewoon de frequentie van zijn meetapparatuur aan aan de eisen van het tijdreeksmodel. De hydroloog echter heeft te maken met de weerbarstigste hydrologische praktijk, waarin hij of zij het moet stellen met de gegevens die er zijn. Grondwaterstandsgegevens werden en worden vaak door mensen verzameld en die mensen zijn helaas wel eens ziek, meten liever niet op zaterdag en zondag en produceren dus gegevens met gaten in de reeks en met onregelmatige tijdstappen. In Stromingen jrg. 5 nr. 2 dragen Bierkens e.a. een mogelijke oplossingsrichting voor dit knelpunt aan. Met behulp van het Kalmanfilter blijkt een exponentieel c.q. ARX(1,0) TR-model met dagelijkse frequentie gekalibreerd te kunnen worden op onregelmatig waargenomen grondwaterstandsreeksen, doordat het Kalmanfilter een inschatting maakt van de variantie van het ruisproces voor de tijdstippen waarop geen waarneming beschikbaar is. Alhoewel het equidistantie-knelpunt hiermee voor eenvoudige, exponentiële systemen voor een belangrijk deel opgelost is, is het met behulp van deze aanpak niet goed mogelijk om de respons te modelleren van trage systemen die niet exponentieel reageren (grote systemen met dikke onverzadigde zones, zoals die te vinden zijn in het duingebied (Van de Vliet e.a., 2000) en op de Veluwe (Gehrels, 1999). Doordat er, tot ongeveer de top van de responsfunctie, voor elke tijdstap een zogenaamde moving-average parameter nodig is, zou een grondwaterstandsreeks van een traag systeem al gauw een niet te kalibreren model met honderden parameters vragen. Daarbij vraagt dit zogenaamde 'KALMAX'-model nog steeds invoerreeksen met equidistante, maar kleine, tijdstappen.

Het identificeren van de juiste orde van TR-modellen is vaak een belangrijk struikelblok voor beginnende modellers, en onderwerp van discussie en verschil van inzicht voor hen die meer ervaren zijn. Er zijn, in navolging van Knotters (2001), globaal drie stromingen te onderscheiden, die in principe naast en door elkaar gebruikt (kunnen) worden. Een nog steeds veel gebruikte procedure is allereerst het iteratief identificeren, kalibreren en diagnostisch toetsen van verschillende modellen, zoals beschreven door Box en Jenkins. De achtergrond van zulke procedures ligt over het algemeen in de statistiek c.q. de stochastische realisatietheorie. Een belangrijk nadeel van deze aanpak is dat de modelorde vaak niet eenduidig vastgesteld kan worden (Hipel en McLeod, 1994) en dat het proces tamelijk heuristisch en kennis- en arbeidsintensief is (De Gooijer e.a., 1985). Beide nadelen worden echter ondervangen wanneer men, als tweede optie, gebruik maakt van automatische modelselectie-procedures. Hierbij worden modellen geselecteerd met behulp van informatiecriteria, zoals Aikake's 'final prediction error' (FPE) criterium, 'information criterion' (AIC) en 'Bayesian information criterion' (BIC) (Akaike, 1970; Akaike, 1974; Akaike, 1979). Zowel het AIC- als het FPE-criterium zijn echter niet consistent, maar overschatten de

ware orde van tijdreeksmodellen (Shibata, 1976). Alhoewel veel modelselectie-criteria aanlokkelijke eigenschappen hebben, is er niet één duidelijk superieure methode aan te wijzen en mist er een coherent raamwerk dat aangeeft wanneer je welk criterium moet gebruiken (De Gooijer e.a., 1985). Naast deze methoden, die beide alleen uitgaan van de data zelf, kan ten derde de modelorde ook worden geselecteerd op basis van de fysica. In de zogenaamde 'data-based mechanistic modelling' methodiek (Young en Beven, 1994; Price e.a., 2000) wordt een fysische analyse van het onderzochte systeem gecombineerd met statistische model-identificatiemethoden. De modelorde wordt in hun methodiek direct uit de data geschat, maar alleen geaccepteerd als er een plausibele fysische verklaring voor is te geven. Knotters en Bierkens (2000) gaan in feite een stap verder en gebruiken een analyse van de fysica om het aantal mogelijke TR-modellen vooraf te beperken tot één. In hun onderzoek komt het ARX(1,0)-model als meest geschikte TR-model voor grondwaterstanden uit de bus, op basis van een analyse van de waterbalans van de verzadigde zone.

In dit artikel wordt een alternatieve methode van TR-tijdreeksanalyse gepresenteerd die gebruik maakt van continue voorgedefinieerde impulsrespons(IR)-functies. De aldus verkregen klasse van tijdreeksmodellen wordt aangeduid als PIRFICT-modellen en biedt een mogelijke oplossing voor bovengenoemde knelpunten van model-identificatie en tijdsdiscretisatie. Het artikel behandelt allereerst de theorie en methodiek van de discrete Box-Jenkins-modellen, waarna hetzelfde dunnetjes wordt overgedaan in continue tijd voor het PIRFICT-model. Vervolgens lopen we de toepassing van het PIRFICT-model stap-voor-stap af, en sluiten af met een korte samenvatting van de methode. Vanwege de eenvoud beschrijven we een TR-model met één invoervariabele, het neerslagoverschot, maar de modelvergelijkingen kunnen zowel in discrete als continue tijd eenvoudig worden uitgebreid naar meerdere invoervariabelen. In het tweede artikel wordt de theorie in praktijk gebracht met behulp van 15 grondwaterstandreeksen uit het Noord-Hollands Duinreservaat.

In discrete tijd: het transferruis-model à la Box en Jenkins

In discrete tijd kan een transfer-model voor tijdreeksen van de grondwaterstand met als enige invoerreeks het neerslagoverschot als volgt geschreven worden:

$$h_t = h_t^* + n_t + d \quad (1)$$

$$h_t^* = \Theta(B)p_t = \sum_{i=0}^{\infty} \Theta_i p_{t-i} \quad (2)$$

$$n_t = \varphi(B)a_t = \sum_{i=0}^{\infty} \varphi_i a_{t-i} \quad (3)$$

met

h_t	=	waargenomen grondwaterstand op tijdstip t (met $t \in \mathbb{N}$, $[-]$) ten opzichte van NAP [m]
h_t^*	=	voorspelde grondwaterstand op tijdstip t ten opzichte van d [m]
n_t	=	de residu-reeks [m]
d	=	het niveau van h_t zonder neerslagoverschot, of met andere woorden het lokale drainageniveau ten opzichte van NAP [m]

$\Theta(B)$	=	transferfunctie [-]
B	=	zogenaamde 'backward shift operator' [-], die ervoor zorgt dat $B^k p_t = p_{t-k}$
p_t	=	het neerslagoverschot op tijdstip t [m]
$\varphi(B)$	=	ruis-transferfunctie [-]
a_t	=	een discreet witte ruis proces [m] met $E[a] = 0$ en $E[a^2] = \sigma_a^2$

De transferfunctie $\Theta(B)$ wordt in het BJ-model gedefinieerd als een breuk met een zogenaamde 'moving average' functie $\omega(B)$ als teller en een 'auto-regressive' functie $\delta(B)$ als noemer, zodat $\Theta(B) = \omega(B)/\delta(B)$. Naar de gewichten $\Theta_0, \Theta_1, \Theta_2 \dots \Theta_\infty$ van de transferfunctie wordt verwezen als impulsrespons-functie. De tijdreeksmodelleur dient tijdens het model-identificatieproces de vertragingstijd te specificeren en het aantal parameters waaruit de transferfuncties bestaan, zodanig dat:

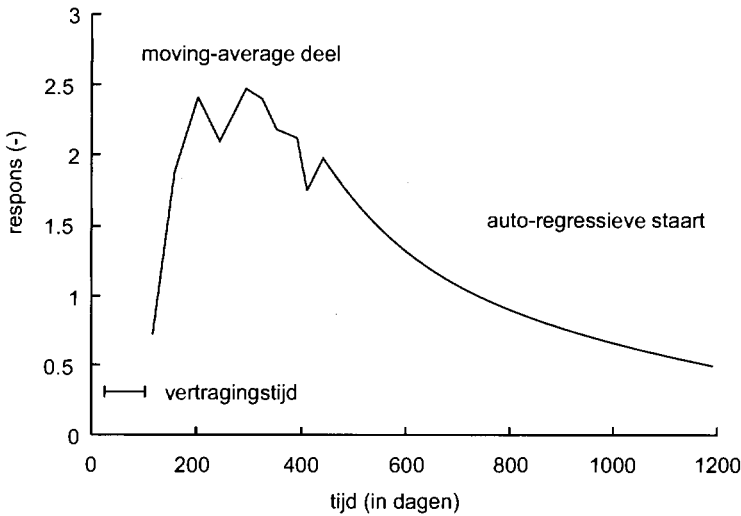
$$\Theta(B) = B^{nk} \frac{\omega(B)}{\delta(B)} = B^{nk} \frac{\omega_0 + \omega_1 B + \omega_2 B^2 + \dots + \omega_{nb-1} B^{nb-1}}{1 + \delta_1 B + \delta_2 B^2 + \dots + \delta_{nd} B^{nd}} \quad (4)$$

$$\varphi(B) = \frac{\theta(B)}{\phi(B)} = \frac{1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_{nc} B^{nc}}{1 + \phi_1 B + \phi_2 B^2 + \dots + \phi_{nf} B^{nf}}$$

met

nb	=	aantal moving-average parameters van het transfermodel
nc	=	aantal moving-average parameters van het ruismodel
nd	=	aantal auto-regressive parameters van het transfermodel
nf	=	aantal auto-regressive parameters van het ruismodel
nk	=	vertragingstijd

De orde van een Box-Jenkins-model wordt zodoende meestal aangeduid met $[nb \ nc \ nd \ nf \ nk]$. Alhoewel het vaak niet expliciet gebeurt, dient de modelleur echter ook de tijdsdiscretisaties th en tp van de grondwaterstandsreeks en het neerslagoverschot te specificeren (of wordt hij of zij daartoe gedwongen door de praktijk). Voor traditionele BJ-modellen geldt hierbij dat $th = tp$, terwijl voor een BJ[1 0 1 1 0] model met Kalmanfilter geldt dat $th = i*tp$ waarbij i een heel getal is. De volledige orde van een Box-Jenkins model geven we hier dus weer met $[nb \ nc \ nd \ nf \ nk][th = tp]$. Met het identificeren van de orde van een TR-model wordt in feite de vorm gespecificeerd die de transferfunctie aan kan nemen. De moving-average parameters zorgen er voor dat de transferfunctie aan het begin geheel vrij is qua vorm, terwijl een auto-regressive parameter ervoor zorgt dat de staart van de transferfunctie exponentieel is (zie figuur 1). De modelleur zoekt met het toevoegen van extra MA-parameters als het ware naar het punt van de transferfunctie van waaraf de staart van de respons aardig benaderd kan worden door een exponentiële functie, wat doorgaans net over de top van de curve is. Omdat het aantal parameters dat hiervoor nodig is afhangt van de waarnemingsfrequentie, moet de modelleur hierbij het aantal parameters (en dus het risico van overfittinggedrag) afwegen tegen de waarnemingsfrequentie (oftewel het aantal voor kalibratie beschikbare waarnemingen). AR-parameters benaderen bovendien bij toene-mende frequentie asymptotisch de waarde van 1, wat problemen op kan leveren door hun slechte schaling.



Figuur 1: Voorbeeld van een transferfunctie van een BJ-model met een vertragingstijd, 10 moving-average-1 auto-regressieve-parameter.

De beperkingen van discrete TR-modellen voor wat betreft de tijdsdiscretisatie ligt al in de basis van vergelijkingen (1-3). Tijd wordt hierin opgevat als een dimensieloze index $t-i$ met $i \in \mathbb{N}$, waarbij een tijdstap aldus altijd gelijk is aan 1, ongeacht de werkelijke tijdsdiscretisatie. Omdat dezelfde index wordt gebruikt voor h_t en p_t , kunnen bovendien de frequenties van de in- en uitvoerreeksen niet verschillen. De reeks met de laagste frequentie is maatgevend voor de frequentie van het model, en aldus wordt de modelleur gedwongen om de extra waarnemingen die hij of zij heeft van de andere reeksen overboord te zetten.

In continue tijd: Het PIRFICT tijdreeksmodel

Net als neerslag, verdamping en grondwaterstandsfluctuaties zijn de meeste processen die met TR-modellen worden gemodelleerd in werkelijkheid natuurlijk niet discreet maar continu. In continue tijd kan een lineaire transformatie van een tijdreeks met behulp van een tijdsafhankelijke transferfunctie geschreven worden als convolutieintegraal (Maas,1994). Vergelijkingen (1-2) kunnen aldus geschreven worden als:

$$h(t) = h^*(t) + n(t) + d \quad (5)$$

$$h^*(t) = \int_{-\infty}^t p(\tau)\theta(t-\tau)d\tau \quad (6)$$

Wanneer we vergelijking (3) transformeren naar continue tijd ontstaat (logischerwijze) een stochastische integraal, waarvan de oplossing onder voorwaarde van een exponentiële IR-functie voor het ruisproces geschreven kan worden als:

$$v(t) = n(t) - e^{-\alpha\Delta t}n(t-\Delta t) \quad (7)$$

waarbij $v(t)$ de innovatiereeks vormt. Een innovatie is daarbij gedefinieerd als het verschil tussen een waarneming en de zogenaamde 1-stap-vooruit voorspelling van het model. Het hoe en waarom hiervan is eigenlijk een verhaal apart, waarop we niet hier, maar in een aparte publicatie in willen gaan (zie ook Von Asmuth en Bierkens, in voorbereiding). Om met onregelmatige tijdstappen en discreet beschikbare gegevens in continue tijd om te kunnen gaan, definiëren we t niet als dimensieloze index maar in werkelijke tijd, en indexeren daarvoor in de plaats t zelf met i . Een discreet beschikbare tijdreeks p , met N waarnemingen kan dan geschreven worden als:

$$p_i = [p(t_0), p(t_1), p(t_2), \dots, p(t_N)] \quad (8)$$

Het lokale drainageniveau in vergelijking (5) is normaal een onbekende, maar kan op de volgende manier geëlimineerd worden uit de vergelijkingen. Wanneer we beide kanten van vergelijking (5) sommeren voor alle tijdstippen waarvoor we waarnemingen hebben, dus van $t = t_0$ tot t_N , delen door N , en aannemen dat het drainageniveau d constant is, krijgen we de volgende uitdrukking:

$$d = \frac{\sum_{i=0}^N h(t_i)}{N} - \frac{\sum_{i=0}^N h^*(t_i)}{N} - \frac{\sum_{i=0}^N n(t_i)}{N} = \bar{h} - \bar{h}^* - \bar{n} \quad (9)$$

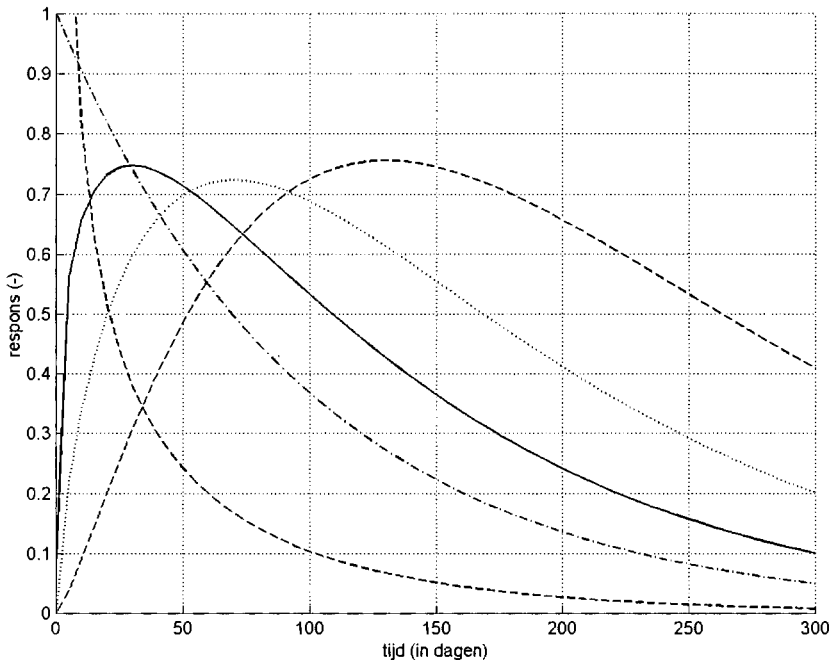
Wanneer we vergelijkingen (5) en (9) combineren krijgen we een algemene vergelijking voor een continu TFN model waarvan de componenten gecentreerd zijn rond hun temporele gemiddelde (de symbolen met een streep erboven), en waarin d bovendien niet meer voorkomt:

$$\{h(t) - \bar{h}\} = \{h^*(t) - \bar{h}^*\} + \{n(t) - \bar{n}\} \quad (10)$$

In geval van het PIRFICT-model wordt de modelorde geïdentificeerd door een continue wiskundige functie, of combinatie van functies, te kiezen als IR-functie. Vervolgens kan een schatting van de parameters van de IR-functie gemaakt worden door een bepaalde doel-functie te minimaliseren. Er zijn echter een aantal belangrijke verschillen met het discrete model-identificatieproces. Een goed gekozen continue IR-functie heeft allereerst een flexibele vorm, waardoor één continue functie de respons van een hele reeks van systemen kan beschrijven. Ten tweede is het model-identificatieproces eenvoudiger doordat het verschijnsel tijdsdiscretisatie niet interfereert met de modelidentificatie. Ten derde kan een continue IR-functie op basis van objectieve criteria gekozen worden als die functie die de fysica van het onderliggende systeem het beste weergeeft. In traditionele TR-modellen kan de modelorde weliswaar op basis van de fysica worden gekozen, maar als er MA-parameters in het model aanwezig zijn ligt de exacte vorm van de transferfunctie, en daarmee de fysische wetten, niet vast.

Voor het beschrijven van de respons van de grondwaterspiegel op neerslag en verdamping gebruiken we hier de Pearson type III verdelingsfunctie (PIII df), met een extra parameter A om de oppervlakte te kunnen aanpassen:

$$\theta(t) = A \frac{\alpha^n (t-b)^{n-1} e^{-\alpha(t-b)}}{\Gamma(n)} \quad (11)$$

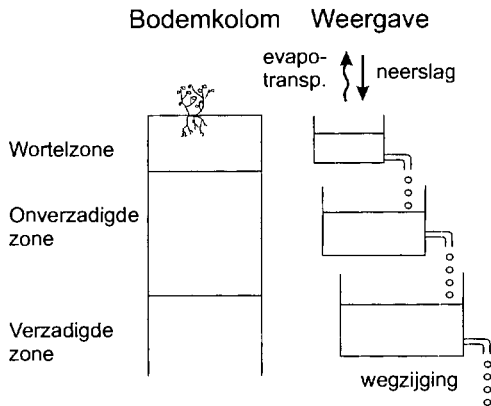


Figuur 2: Voorbeeldkrommen uit de Pearson III df familie ($A=100$, $\alpha = 0,01$, $n = [0,5 \ 1 \ 1,3 \ 1,7 \ 2,3]$)

De PIII df kan qua vorm geleidelijk overgaan van een functie die steiler is dan exponentieel, via een exponentiële functie naar een normale verdeling (zie figuur 2). In discrete termen omvat de functie daarmee het ARX(1,0) model, maar bovendien alle ARMAX-modellen waarvan de ordinaten van de transferfunctie een van de krommen uit de PIII familie vormen. De fysische basis van de PIII df ligt in het feit dat het de transfer-functie van een serie van gekoppelde lineaire reservoirs kan beschrijven, waarbij de parameter n hun aantal weergeeft. Knotters en Bierkens (2000) tonen in hun artikel over de fysische basis van tijdreeksmodellen aan dat het ARX(1,0) model een gediscretiseerde versie is van een enkel lineair reservoir, en kozen het daarom als meest geschikte model voor tijdreeksen van de grondwaterstand. De keuze voor het ARX-model is echter gebaseerd op een simpel fysisch model van een 1-dimensionale grondkolom, waarbij laterale stroming en het gedrag van de onverzadigde zone verwaarloosd worden. In die zin vormt het PIRFICT-model met de PIII df een verdere uitbreiding van hun methodiek, aangezien de PIII df bijvoorbeeld ook de gecombineerde respons van een gelaagde verzadigde en onverzadigde zone kan beschrijven (figuur 3).

Implementatie van het PIRFICT-model

Wanneer de responsfuncties geïdentificeerd zijn staat de kern van het PIRFICT-model overeind, en kunnen de modelvergelijkingen geëvalueerd worden. Om inzicht te krijgen in de werking van het model lopen we hieronder stap-voor-stap de verdere toepassing ervan af. Allereerst moeten de beschikbare invoerreeksen getransformeerd worden naar continue



Figuur 3: Schematische weergave van een freatische bodemkolom met behulp van de Pearson III df.

reeksen, aangezien de meeste reeksen op de een of andere manier gediscretiseerd zijn om op een digitaal medium opgeslagen te kunnen worden. Veel tijdreeksen, zoals het neerslagoverschot, kunnen echter beschouwd worden als een verandering van de primitieve functie van een onderliggend continu proces:

$$p_i = P(t_i) - P(t_{i-1}) = \int_{t_{i-1}}^{t_i} p(\tau) d\tau \quad (12)$$

Uit bovenstaande vergelijking valt logischerwijze de continue reeks $p(\tau)$ niet zomaar op te lossen als p_i slechts discreet beschikbaar is. We kunnen echter wel een benadering voor $p(\tau)$ vinden door aan te nemen dat het neerslagoverschot uniform verdeeld is over de periode t_{i-1} tot t_i . Vergelijking (12) kan dan geschreven worden als:

$$p(\tau) = \frac{p_i}{t_i - t_{i-1}}, \quad t_{i-1} \geq \tau > t_i \quad (13)$$

Op deze manier wordt de benadering van de continue neerslagoverschotsreeks $p(\tau)$ beter naarmate de tijdstappen kleiner worden. Omdat vergelijking (6) een neerslagreeks vanaf $t = -\infty$ bevat, moet er vervolgens een beginvoorwaarde gedefinieerd worden om met de gegevens vanaf $t = 0$ te kunnen rekenen. We kiezen als beginvoorwaarde P_{ini} voor het neerslagoverschot van $t = -\infty$ tot $t = 0$ het gemiddelde neerslagoverschot uit onze meetreeks zodat:

$$h(0) - \bar{h} = \bar{p} \int_0^{\infty} \theta(t) dt - \bar{h}^* + n(0) - \bar{n}^* \quad (14)$$

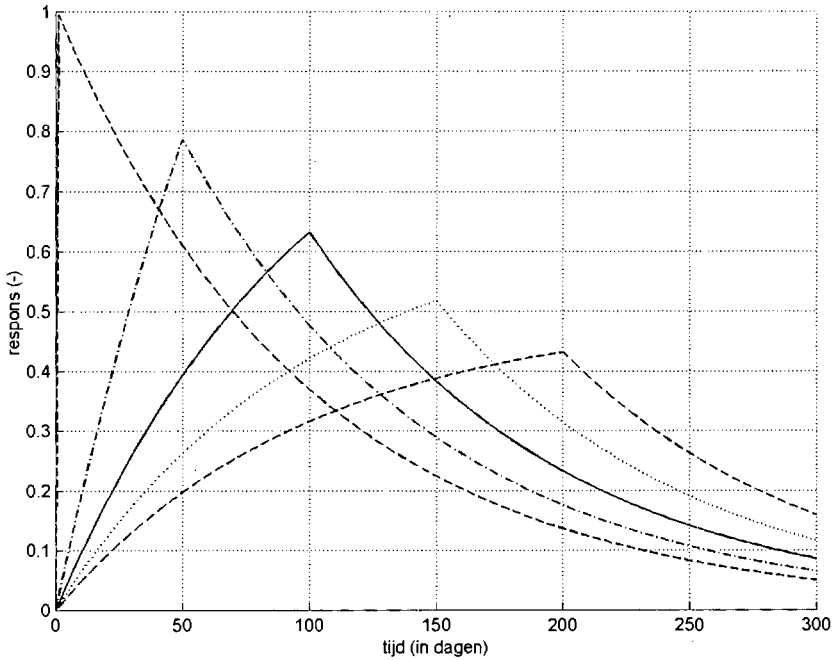
waarbij $t = 0$ het tijdstip van de eerste neerslagoverschotmeting is, wat liefst een aardige periode voor de eerste grondwaterstandsmeting dient te vallen. Een andere optie, die echter als nadeel heeft dat hij gevoeliger is voor afwijkingen in de data, is om de beginvoorwaarde zo te kiezen dat de eerste gesimuleerde waarde van het model overeenkomt met de eerste meting van de grondwaterstand:

$$P_{ini} = \frac{h(0) - \bar{h} - \bar{h}^* - \bar{n}^*}{\int_0^{\infty} \theta(t) dt} \quad (15)$$

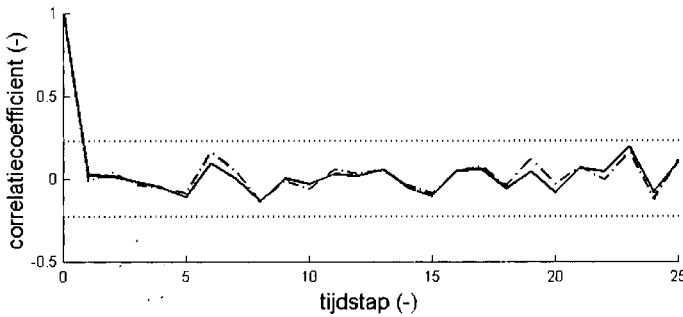
Nu een continue invoerreeks beschikbaar is kan vergelijking (6) geëvalueerd worden door gebruik te maken van de blokrespons(BR)-functie. Waar in discrete TR-modellen de gewichten van de transfer-functie beschouwd worden als respons op een impuls, is vanuit de continue optiek dit de respons op een invoerreeks in de vorm van een blok. Om verwarring te voorkomen gebruiken we de term transfer-functie voor het discrete geval en refereren we aan zijn continue equivalent als de BR-functie $\Theta(t)$. De term IR-functie kan dan gereserveerd worden voor de respons op een feitelijk instantane impuls. In tegenstelling tot de BR-functie en de transfer-functie van discrete modellen is de IR-functie van een convolutieproces geen functie van Δt , en hangt met andere woorden dus niet af van de waarnemingsfrequentie. De IR-functie beschrijft de reactie van de grondwaterspiegel op een plotselinge impuls van neerslag en is onveranderlijk in de tijd en karakteristiek voor een bepaald systeem, zolang de fysica van het systeem niet verandert (Von Asmuth en Maas, 2001). De methode is daarmee verwant aan de zogenaamde 'instantaneous unit hydrograph' die in de oppervlaktewater-hydrologie gebruikt wordt, terwijl ook het concept van gekoppelde lineaire reservoirs daar gebruikt wordt en bekend staat onder de naam 'Nash-cascade'. De BR-functie $\Theta(t)$ wordt verkregen door de IR-functie te convolueren met een 'blok' neerslagoverschot met eenheidsintensiteit over een periode Δt , wat gelijk staat aan:

$$\Theta(t) = \int_{t-\Delta t}^t \theta(\tau) d\tau \quad (16)$$

In figuur 4 worden een aantal voorbeelden van BR-functies van een enkel lineair reservoir getoond voor verschillende tijdstappen. Vanwege het superpositiebeginsel kan het verloop van de grondwaterspiegel $h^*(t)$ worden verkregen door de respons-functies van alle 'blokken' van neerslag bij elkaar op te tellen. Omdat $\Theta(t)$ een continue functie is, is $h^*(t)$ zelf ook continu gedefinieerd, en kan er vervolgens voor elke waarneming van $h(t)$ ook een waarde van de residureeks $n(t)$ worden verkregen. Wanneer uit deze residureeks met behulp van vergelijking (7) een innovatiereeks gedestilleerd wordt, zijn daarmee alle modelvergelijkingen geëvalueerd en kan het model worden ingebed in een optimalisatie-algoritme (zoals Marquardt-Levenberg) om de parameters te schatten. Als criterium wordt hierbij de Likelihood-functie van de innovaties gebruikt (Bierkens e.a., 1999), en niet een kleinste-kwadratencriterium, omdat de variantie van de innovaties niet constant is maar afhankelijk van de stapgrootte. Wanneer het model gekalibreerd is wordt, last but not least, de betrouwbaarheid en geldigheid van het model gecontroleerd met behulp van de auto- en crosscorrelatie-functies van de innovatiereeks, de covariantie-matrix van de modelparameters en de variantie van de IR-functies. De covariantiematrix van de parameters $C(p)$ wordt geschat met behulp van de Jacobiaan J en σ^2 . Uit deze covariantie matrix kunnen tevens de correlaties tussen de parameters berekend worden. De variantie van de IR-functie $\theta(t)$ wordt berekend met:



Figuur 4: Voorbeelden van blokrespons-functies van een enkel lineair reservoir voor verschillende tijdstappen ($A=100, \alpha=0,01, n=1, \Delta t = [1 \ 50 \ 100 \ 150 \ 200]$).



Figuur 5: Voorbeeld van de autocorrelatie-functies van een BJ- en een PIRFICT model van dezelfde data. De gestippelde lijnen geven het 95% betrouwbaarheids interval weer.

$$\begin{aligned} \sigma_{\theta(t)}^2 = & \left(\frac{\partial\theta(t)}{\partial A}\right)^2 \sigma_A^2 + \left(\frac{\partial\theta(t)}{\partial a}\right)^2 \sigma_a^2 + \left(\frac{\partial\theta(t)}{\partial n}\right)^2 \sigma_n^2 + 2\left(\frac{\partial\theta(t)}{\partial A}\right)\left(\frac{\partial\theta(t)}{\partial a}\right)C(A,a) + \dots \\ & + 2\left(\frac{\partial\theta(t)}{\partial A}\right)\left(\frac{\partial\theta(t)}{\partial n}\right)C(A,n) + 2\left(\frac{\partial\theta(t)}{\partial a}\right)\left(\frac{\partial\theta(t)}{\partial n}\right)C(a,n) \end{aligned} \quad (17)$$

zodat we, door aan te nemen dat $\theta(t)$ normaal verdeeld is, een betrouwbaarheidsinterval $\pm 2\sigma$ kunnen plotten voor $\theta(t)$. Net als in discrete TR-modellen komen serieuze modelproblemen tot uiting in de autocorrelatie-functie van de innovatiereeks $v(t)$ en de crosscorrelatie-functie tussen $v(t)$ en de invoerreeks $p(t)$. De autocorrelatie-functie geeft bijvoorbeeld aan of de aanname dat de innovatiereeks 'wit' is klopt of verworpen moet worden, terwijl de crosscorrelatie-functie aangeeft of er nog patronen in de ruisreeks aanwezig zijn die door de

invoerreeks verklaard hadden kunnen worden. Figuur 5 toont een voorbeeld van een crosscorrelatie-functie van een BJ- en een PIRFICT model, waarbij voor beide modellen de correlaties binnen het betrouwbaarheidsinterval vallen. De autocorrelatie- en crosscorrelatie-functies op tijdstap k worden in het continue geval op dezelfde wijze gedefinieerd als voor BJ-modellen, maar vanwege de niet-equidistante gegevens wordt een marge rond tijdstap k van $\pm 0,5 k$ toegepast.

Samenvatting methode

De toepassing van het PIRFICT-model bestaat in het kort uit de volgende stappen. Allereerst wordt voor elke invoerreeks een impulsrespons-functie gekozen, waarvoor in principe elke continue functie of sommatie van functies kan dienen. Het model is echter bij uitstek geschikt om hiervoor fysische wetten c.q. analytische formules te gebruiken. Ten tweede worden de invoerreeksen getransformeerd naar continue tijd met behulp van vergelijking (13), door aan te nemen dat ze uniform verdeeld zijn gedurende de tijdstappen. De transferconvolutieintegraal (vergelijking (6)) kan nu geëvalueerd worden met behulp van de blokrespons-functie van elke blokpuls, zodat een continue simulatie van de grondwaterstand verkregen wordt. Hiermee wordt voor elke grondwaterstandmeting ook een waarde voor de residureeks verkregen, waaruit via vergelijking (7) een innovatiereeks te destilleren valt. Het schatten van de parameters gebeurt met behulp van een Marquardt-Levenberg algoritme en de zogenaamde Likelihood-functie van de innovaties. Last but not least dient de betrouwbaarheid en geldigheid van het model gecontroleerd te worden met behulp van de auto- en crosscorrelatie-functie van de innovatiereeks, de covariantie-matrix van de modelparameters en de variantie van de IR-functies.

De winst van het PIRFICT-model voor de praktijk kan samengevat worden door te concluderen dat met deze aanpak (vooralsnog in theorie) een TR-model verkregen wordt dat eenvoudiger te identificeren, beter te standaardiseren, en beter toepasbaar is op onregelmatige meetreeksen van complexe systemen.

Literatuur

- Akaike, H. (1970)** Statistical predictor identification; in: *Ann. Inst. Statist. Math.*, jrg 22, pag 203–217.
- Akaike, H. (1974)** A new look at the statistical model identification; in: *IEEE Trans. Autom. Control*, jrg 19, pag 716–723.
- Akaike, H. (1979)** A Bayesian extension of the AIC procedure; in: *Biometrika*, jrg 66, pag 237–242.
- Asmuth, J.R. von en M.F.P. Bierkens (in voorbereiding)** Modelling irregularly observed time series as a continuous coloured stochastic process; in: conference proceedings CMWR2002.
- Asmuth, J.R. von en C. Maas (2001)** The method of impulse response moments: a new method integrating time series-, groundwater- and eco-hydrological modelling; in: Gehrels, J.C., N.E. Peters, E. Hoehn, K. Jensen, C. Leibundgut, J. Griffioen, B. Webb en

- W.J. Zaadnoordijk (red), *Impact of Human Activity on Groundwater Dynamics*; IAHS Publication 269, Wallingford, pag 51–58.
- Bierkens, M.F.P., M. Knotters en F.C. van Geer (1999)** Tijdreeksanalyse nu ook toepasbaar bij onregelmatige meetfrequenties; in: *Stromingen*, jrg 5, nr 2, pag 43–54.
- Box, G.E.P. en Jenkins, G.M. (1970)** *Time Series Analysis: Forecasting and Control*; Holden-Day, San Fransisco.
- Geer, F.C. van, P.K. Baggelaar en P.R. Defize (1988)** Toepassing van tijdreeksanalyse op meetreeksen van de stijghoogte; in: *H₂O*, jrg 21, nr 161, pag 438–442.
- Gehrels, J.C. (1999)** *Groundwater Level Fluctuations, separation of natural from anthropogenic influences and determination of groundwater recharge in the Veluwe area, the Netherlands*; Proefschrift Vrije Universiteit Amsterdam, Amsterdam.
- Gooijer, J.G. de, B. Abraham, A. Gould en L. Robinson (1985)** Methods for determining the order of an autoregressive-moving average process: A survey; in: *International Statistical Review*, jrg 53, nr 3, pag 301–329.
- Hipel, K.W. en McLeod, A.I. (1994)** *Time series modelling of water resources and environmental systems*; Elsevier, Amsterdam.
- Knotters, M. (2001)** *Regionalised time series models for water table depths*; Proefschrift Wageningen Universiteit, Wageningen.
- Knotters, M. en M.F.P. Bierkens (2000)** Physical basis of time series models for water table depths; in: *Water Resources Research*, jrg 36, nr 1, pag 181–188.
- Knotters, M. en P.E.V. van Walsum (1997)** Estimating fluctuation quantities from time series of water-table depths using models with a stochastic component; in: *Journal of Hydrology*, jrg 197, pag 25–46.
- Ljung, L. (1999)** *System Identification, Theory for the user*; Prentice Hall, Upper Saddle River.
- Maas, C. (1994)** *On Convolutional Processes and Dispersive Groundwater Flow*; Proefschrift, Delft University of Technology, Delft.
- Price, L.E., P. Goodwill, P.C. Young en J.S. Rowan (2000)** A data-based mechanistic modelling (DBM) approach to understanding dynamic sediment transmission through Wyresdale Park Reservoir, Lancashire, UK; in: *Hydrological Processes*, jrg 14, pag 63–78.
- Rolf, H.L.M. (1989)** *Verlaging van de grondwaterstanden in Nederland. Analyse periode 1950–1986*; DGV-TNO, Delft.
- Shibata, R. (1976)** Selection of the order of an autoregressive model by Akaike's information criterion; in: *Biometrika*, jrg 63, pag 117–126.
- Vliet, R.N. van de, H.L.M. Rolf en J.J. Lebbink (2000)** *Effecten van de reductie en stopzetting waterwinning Zuid-Kennemerland. Tijdreeksanalyse van gemeten grondwaterstandreeksen voor urgente gebieden*; PWN / IWACO, Rotterdam.
- Young, P. C. en K. J. Beven (1994)** Data based mechanistic (DBM) modelling and the rainfall-flow non-linearity; in: *Environmetrics*, jrg 5, pag 335–365.