Tijdreeksanalyse nu ook toepasbaar bij onregelmatige meetfrequenties

Marc F.P. Bierkens
Martin Knotters
Frans C. van Geer

Een belangrijke eis voor de toepassing van tijdreeksanalyse op hydrologische reeksen is dat deze zijn gemeten met een regelmatige frequentie. Deze meetfrequentie is meteen ook beperkend voor de minimale tijdstap waarmee de tijdreeksmodellen kunnen rekenen. Dit artikel beschrijft een methode om een bepaalde klasse van tijdreeksmodellen, de zogenaamde transfer–ruismodellen (TR-modellen), toe te passen op reeksen die met een te lage of onregelmatige frequentie zijn gemeten. Op deze wijze kunnen hydrologische tijdreeksen worden verdicht en verlengd en kunnen gaten in deze reeksen worden opgevuld. Bijvoorbeeld: voor een locatie waar gedurende een aan aantal jaren op onregelmatige tijdstippen, soms met gaten van maanden, de stijghoogte van het grondwater is gemeten, kan op deze wijze een langjarige reeks van dagelijkse stijghoogten worden gegenereerd. Voorwaarde is dan wel dat de belangrijkste invoervariabele van het TR-model – dit is meestal het neerslagoverschot – gedurende dezelfde langjarige periode op dagbasis is gemeten.

Inleiding

De tijdreeksanalyse zoals beschreven door Box en Jenkins (1976) is een nuttige en veel gebruikte set van methoden om hydrologische tijdreeksen te analyseren. Toepassing van tijdreeksanalyse in de hydrologie gaat meestal als volgt. Uitgangspunt is een (korte) tijdreeks met waarnemingen van een hydrologische variabele (bijvoorbeeld afvoer). Door analyse van het statistisch gedrag van deze tijdreeks volgt de keuze voor een type tijdreeksmodel. Vervolgens worden de parameters van het tijdreeksmodel geschat door dit op de tijdreeks te kalibreren. Het gekalibreerde tijdreeksmodel kan vervolgens worden gebruikt voor: a) het analyseren van mogelijke trends en de oorzaak van deze trends; b) het extrapoleren van de tijdreeks (vooruit en achteruit in de tijd); c) stochastische simulatie: het simuleren van een groot aantal evenzo waarschijnlijke lange tijdreeksen. Deze gesimuleerde reeksen kunnen vervolgens worden gebruikt voor onzekerheidanalyse (als invoer voor een hydrologische model), risicoanalyse en ontwerp.

Er bestaat een groot aantal soorten tijdreeksmodellen. Een veel gebruikt en nuttig type tijdreeksmodel is het transfer–ruismodel (hierna aangeduid met TR-model). Bij een TR-model wordt de variatie van een meetreeks van de *uitvoervariabele* (de variabele waarin men geïnteresseerd is, zoals bijv. rivierafvoer of stijghoogte van het grondwater) gerela-

Marc F.P. Bierkens en **Martin Knotters** zijn werkzaam bij DLO–Staring Centrum, Postbus 125, 6700 AC Wageningen, e-mail: bierkens@sc.dlo.nl.
Frans C. van Geer is werkzaam bij NITG–TNO, Postbus 6012, 2600 JA Delft.

teerd aan de variatie van één of meer *invoervariabelen* (ofwel verklarende variabelen, zoals bijv. het neerslagoverschot en grondwateronttrekking). De gebruikte empirische relaties tussen uitvoervariabele en de invoervariabelen worden transferfuncties genoemd. Deze kunnen zowel lineair (Box en Jenkins, 1976; Hipel en McCleod, 1994) als niet-lineair zijn (Tong, 1990; Knotters en De Gooijer, 1999). De variatie van de uitvoervariabele kan echter nooit helemaal verklaard worden door de invoervariabelen. Als we met behulp van reeksen van invoervariabelen en de bijbehorende transferfuncties een reeks van de uitvoervariabele berekenen, en deze berekende reeks aftrekken van de waargenomen tijdreeks, dan houden we een reeks van verschillen (*residuen* genaamd) over. Deze residureeks is het onverklaarde deel van de tijdreeks, en wordt in het TR-model beschreven door een, eveneens empirisch, ruismodel.

TR-modellen zijn met name geschikt om een hydrologische tijdreeks uiteen te rafelen in verschillende componenten die ieder verklaard kunnen worden door een andere invoervariabele. Zo kan bijvoorbeeld worden vastgesteld welk gedeelte van een gemeten stijghoogtereeks wordt verklaard door het neerslagoverschot, de rivierwaterstand en een naburige grondwateronttrekking. Op deze wijze kunnen dus ook de belangrijkste oorzaken van eventuele stijghoogteverlagingen worden opgespoord (Van Geer en Defize, 1987; Rolf, 1989). Voorbeelden van toepassingen van TR-modellen op hydrologische meetreeksen zijn er onder meer voor stijghoogten (Tankersley e.a., 1993; Gehrels e.a., 1994; Knotters en Van Walsum, 1997), rivierafvoeren (Hipel e.a., 1975; Chow e.a., 1983) en suspensieve sedimentgehalten (Gurnell en Fenn, 1984; Lemke, 1991).

Een nadeel van alle tijdreeksmodellen, en dus ook van TR-modellen, is dat de in- en uitvoervariabelen met een regelmatige frequentie moeten zijn gemeten. Dat is heel vaak niet het geval: stijghoogten in Nederland worden op de 14^{e} en de 28^{e} van de maand gemeten. Dit resulteert in 24 waarnemingen per jaar, maar niet met regelmatige tijdsintervallen tussen de metingen. In de praktijk wordt dit veelal opgelost door net te doen of er tussen de waarnemingen een constant tijdsinterval van 1/24ste jaar zit (tijdstap ongeveer 15,2 dagen). Vaak treft men in een gebied ook buizen aan met verschillende meetfrequenties (1 maal per week, 2 maal per maand, 1 maal per maand en 4 maal per jaar). Reeksen worden ook nogal eens afgekeurd voor analyse met een TR-model omdat er teveel gaten in zitten. Tenslotte is het vaak zo dat de frequentie waarmee een uitvoervariabele wordt gemeten niet altijd hoog genoeg is om de gewenste vragen te beantwoorden. Het optreden van natschade bij een bepaald gewas hangt bijvoorbeeld af van het aantal *achtereenvolgende* dagen dat de grondwaterstand zich in de wortelzone bevindt. Dit is niet te bepalen als de grondwaterstand maar twee keer per maand wordt gemeten.

In de meeste toepassingen van TR-modellen vormt het neerslagoverschot (neerslag minus Makkink-referentieverdamping) de belangrijkste invoervariabele. Neerslag en verdamping worden in Nederland op een groot aantal locaties op dagbasis gemeten. Als het neerslagoverschot als reeksen van dagelijkse gemiddelden bekend is, zou het dus mogelijk moeten zijn om de TR-modellen ook op dagbasis toe te passen. De uitvoervariabele *op dagbasis* wordt dan via de transfermodellen gerelateerd aan de invoerreeksen, waaronder in ieder geval het dagelijks gemiddelde neerslagoverschot. Net als voor een fysisch-deterministisch model dat op dagbasis rekent, kunnen de parameters van de transfermodellen op dagbasis eenvoudig worden gekalibreerd, namelijk via minimalisatie van de verschillen tussen de berekende en gemeten uitvoervariabele op de meettijdstippen. Een probleem blijft echter de identificatie van het ruisproces: hoe kunnen we de parameters van het ruisproces op dagbasis schatten wanneer we alleen de residuen op de meettijdstippen kennen?

In dit artikel presenteren we een methode om TR-modellen te kalibreren met dezelfde tijdstap als de (regelmatig gemeten) invoervariabele (bijv. het neerslagoverschot), zelfs als de uitvoervariabele (bijv. de stijghoogte of de afvoer) met een veel lagere of onregelmatige frequentie is gemeten. Op deze wijze kunnen hydrologische meetreeksen worden verdicht en verlengd en kunnen gaten in reeksen worden opgevuld. Voor een locatie waar gedurende aan aantal jaren op onregelmatige tijdstippen, soms met gaten van maanden, de stijghoogte is gemeten, kan op deze wijze, met behulp van het neerslagoverschot, een reeks van bijv. 30 jaar dagelijkse stijghoogten worden gegenereerd. Eerst zal de onderliggende theorie worden uitgelegd. Dat de methode werkt wordt vervolgens geïllustreerd door een toepassing op een locatie met dagelijks gemeten grondwaterstanden.

Theorie

In dit hoofdstuk wordt de theorie uitgelegd met het voorbeeld van de stijghoogte van het grondwater als uitvoervariabele en het neerslagoverschot als invoervariabele.

TR-modellen

Het startpunt is het TR-model met het neerslagoverschot als enige invoer:

$$H_t = H_t^* + N_t \tag{1}$$

$$H_t^* = \sum_{i=1}^{r} \delta_i H_{t-i}^* + \sum_{j=0}^{s} \omega_j P_{t-j-b} \tag{2}$$

$$(N_t - c) = \sum_{k=1}^{p} \phi_k (N_{t-k} - c) + \varepsilon_t + \sum_{l=1}^{q} \theta_l \varepsilon_{t-l} \tag{3}$$

met

H_t	:	de stijghoogte op tijdstip t;
H_t^*	:	de stijghoogte op tijdstip t verklaard door het neerslagoverschot;
P_t	:	het gemiddelde neerslagoverschot tussen tijdstap $t-1$ and tijdstip t. P_t wordt berekend uit de gemiddelde neerslag minus de gemiddelde Makkink- referentieverdamping;
N_t	:	het ruisproces (de residuen) op tijdstip t;
ε_t	:	een discreet witte-ruisproces met gemiddelde 0 en variantie σ_ε^2;
c	:	een constante: het gemiddelde van N_t;
δ_i	:	de autoregressieve parameters van het transfermodel tot op orde r;
ω_j	:	de *moving average* parameters van het transfermodel tot op orde s;
ϕ_k	:	de autoregressieve parameters van het ruismodel tot op orde p;
θ_l	:	de *moving average* parameters van het ruismodel tot op orde q;
b	:	de tijdsvertraging tussen invoer en uitvoer.

Vergelijking (2) is het transfermodel (deterministische gedeelte), vergelijking (3) het ruismodel (het stochastisch gedeelte) en vergelijking (1) het transfer–ruismodel (TR-model).

Om het eenvoudig te houden beschouwen we alleen een beperkte klasse van TR-modellen met $r = 1, s = 0, p = 1, q = 0$ en $b = 0$:

$$H_t = H_t^* + N_t \tag{4}$$

$$H_t^* = \delta_l H_{t-1}^* + \omega_0 P_t \tag{5}$$

$$(N_t - c) = \phi_l (N_{t-l} - c) + \varepsilon_t \tag{6}$$

Een speciaal geval van het TR-model (4) wordt verkregen als de autoregressieve parameter van het ruismodel (6) gelijk is aan die van het transfermodel (5): $\phi_l = \delta_1$. Het TR-model krijgt dan de volgende vorm:

$$(H_t - \mu) = a(H_{t-l} - \mu) + bP_t + \varepsilon_t \tag{7}$$

Vergelijking (7) staat bekend als een ARMAX(1,0)-model (Hipel en McLeod, 1994). Met de introductie van dit model lopen we vooruit op het artikel van Knotters en Bierkens (1999) waarin het ARMAX(1,0)-model een centrale rol speelt.

Matrixnotatie van het TR-model

Het TR-model kan als volgt in een matrixnotatie worden geschreven:

$$\begin{bmatrix} H_t^* \\ N_t \end{bmatrix} = \begin{bmatrix} \delta_1 & 0 \\ 0 & \phi_1 \end{bmatrix} \begin{bmatrix} H_{t-1}^* \\ N_{t-1} \end{bmatrix} + \begin{bmatrix} \omega_0 & 0 \\ 0 & c(1-\phi_1) \end{bmatrix} \begin{bmatrix} P_t \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \varepsilon_t \tag{8}$$

Als we definiëren (vetgedrukte letters zijn matrices en vectoren):

$$\mathbf{x}_t = \begin{bmatrix} H_t^* \\ N_t \end{bmatrix} \quad \mathbf{u}_t = \begin{bmatrix} P_t \\ 1 \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} \delta_1 & 0 \\ 0 & \phi_1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} \omega_0 & 0 \\ 0 & c(1-\phi) \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{9}$$

dan kan vergelijking (8) geschreven worden als

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_t + \mathbf{D}\varepsilon_t \tag{10}$$

Vergelijking (10) is de toestandsvergelijking van een stochastisch systeem, d.w.z. het geeft de toestand van $\mathbf{x}_t$ op tijdstip t aan.

Meerdere invoervariabelen

De toestandsvergelijking kan eenvoudig worden uitgebreid voor TR-modellen met meerdere invoervariabelen (Shea, 1987; Van Geer en Zuur 1997). Stel dat de stijghoogte niet alleen wordt bepaald door het neerslagoverschot maar ook door het onttrekkingsdebiet van een nabij gelegen grondwaterwinning. Het TR-model met twee inputs ziet er dan als volgt uit:

$$H_t = H_t^* + H_t^{**} + N_t \tag{11}$$

$$H_t^* = \delta_1^* H_{t-1}^* + \omega_o^* P_t \tag{12}$$

$$H_t^{**} = \delta_1^{**} H_{t-1}^{**} + \omega_o^{**} Q_t \tag{13}$$

$$(N_t - c) = \phi_1 (N_{t-1} - c) + \varepsilon_t \tag{14}$$

waarbij Q_t het gemiddelde onttrekkingsdebiet is tussen t–1 en t. De matrix-vector notatie van dit TR-model is:

$$\begin{bmatrix} H_t^* \\ H_t^{**} \\ N_t \end{bmatrix} \begin{bmatrix} \delta_1^* & 0 & 0 \\ 0 & \delta_1^{**} & 0 \\ 0 & 0 & \phi_1 \end{bmatrix} \begin{bmatrix} H_{t-1}^* \\ H_{t-1}^{**} \\ N_{t-1} \end{bmatrix} + \begin{bmatrix} \omega_o^* & 0 & 0 \\ 0 & \omega_o^{**} & 0 \\ 0 & 0 & c(1-\phi_1) \end{bmatrix} \begin{bmatrix} P_t \\ Q_t \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \varepsilon_t \tag{15}$$

Het uitbreiden naar meer dan twee invoervariabelen gaat op vergelijkbare wijze.

Kalmanfilter

Als het TR-model geschreven is in de vorm van de toestandsvergelijking (10) kan het worden ingebed in een Kalmanfilter. Zoals we later zullen zien kan het Kalmanfilter dan gebruikt worden om de parameters van het TR-model te schatten. Zoals uitgelegd in Bierkens (1998) worden met het Kalmanfilter voorspellingen gedaan met het deterministische gedeelte van het model (in dit geval $\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\mathbf{u}_t$). Hier gebeurt dat met dezelfde tijdstap als de invoerreeksen. Voor het neerslagoverschot in Nederland is dat 1 dag. Om de zoveel tijdstappen (bijv. om de 15 dagen) is er naast een voorspelling van het deterministisch model ook een waarneming beschikbaar. Op dat moment wordt de voorspelling bijgesteld met behulp van de waarneming, namelijk door een gewogen gemiddelde te nemen van de modelvoorspelling en de waarneming. Naarmate de meetfout kleiner is in vergelijking tot de grootte van de voorspelfout (o.a. afhankelijk van σ_ε^2) krijgt de waarneming meer gewicht en vice versa. De bijgestelde voorspelling van $\mathbf{x}_t$ wordt dan als beginvoorwaarde voor de volgende tijdstap gebruikt. Doordat de modelvoorspellingen regelmatig worden bijgesteld met behulp van waarnemingen, zijn voorspellingen met het Kalmanfilter gemiddeld nauwkeuriger dan die van het deterministische model alleen. Degenen die zijn geïnteresseerd in het algoritme van het Kalmanfilter worden verwezen naar Van Geer en Zuur (1997) of kunnen een document met het gebruikte Kalman-algoritme verkrijgen bij de eerste auteur.

Schatting van de parameters van het TR-model

Uit het Kalmanfilter hebben we op de meettijdstippen τ dus zowel de beschikking over modelvoorspellingen van de stijghoogte (laten we die $\overline{H}_\tau$ noemen) als waarnemingen (hier aangeduid met y_τ). Het verschil tussen de waarneming en de modelvoorspelling met het Kalmanfilter ($y_\tau - \overline{H}_\tau$) wordt *innovatie* v_τ genoemd. Naast de innovatie v_τ levert het Kal-

manfilteralgoritme op elk meettijdstip τ ook een schatting van de *innovatievariantie* $\sigma^2_{v,\tau}$. Uit de innovaties en de innovatievarianties kunnen de parameters van de toestandsvergelijking (10) worden geschat. Dit betekent dus dat de parameters van zowel het transfermodel als dat van het ruismodel kunnen worden geschat (beide werkend met een tijdstap van 1 dag), zelfs als maar op een beperkt aantal tijdstippen waarnemingen beschikbaar zijn.

Als we uitgaan van een reeks invoervariabelen met een lengte van n tijdstappen met waarnemingen van de uitvoervariabele op m tijdstippen (met $m \leq n$, dan gaat het algoritme voor de schatting van de TR-parameters als volgt (zie ook Te Stroet, 1995 en Bierkens, 1998):

1 Kies een beginschatting voor de parameters van het TR-model; dus (δ_1, ω_0, ϕ_1, c, σ_ε^2) voor het TR-model (4)–(6) of (δ_1^*, ω_0^*, δ_1^{**}, ω_0^{**}, ϕ_1, c, σ_ε^2) voor het TR-model (11)–(14);
2 Run met de invoerreeksen en de parametersschatting het Kalmanfilter voor de n tijdstappen. Dit resulteert in m innovaties v_τ en m innovatievarianties $\sigma^2_{v,\tau}$;
3 Evalueer het volgende *maximum-likelihood*-criterium (Schweppe, 1973):

$$J = m\ln(2\pi) + \sum_{\tau=1}^{m} \ln\left(\sigma^2_{v,\tau}\right) + \sum_{\tau=1}^{m}\left(\frac{v_\tau^2}{\sigma^2_{v,\tau}}\right) \tag{16}$$

4 Stel op basis van de waarde van het criterium de schatting van de parameters bij. De nieuwe parameterwaarden worden zo gekozen dat het criterium kleiner wordt (minimalisatie);
5 Ga terug naar 2 en blijf stap 2 t/m 4 herhalen totdat het criterium, binnen zekere grenzen, niet meer van waarde verandert. De parameterset die zo gevonden wordt en dus (16) minimaliseert, heet een *maximum-likelihood*-schatting van de parameters.

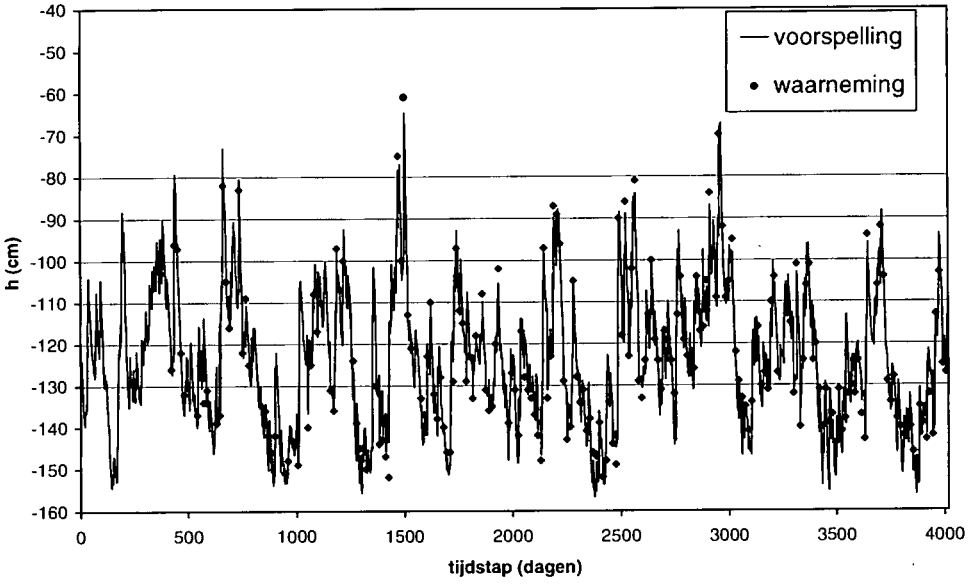
Toepassing

De dataset

Om na te gaan of de hiervoor beschreven methode werkt, is deze toegepast en gevalideerd op een tijdreeks van 15 jaar (1981–1995) met dagelijks gemeten stijghoogten. De stijghoogte is elke dag om ongeveer 9 uur in de ochtend waargenomen. Het betreft een stijghoogtebuis op het meteoveld van het KNMI in De Bilt (TNO code 32CL0034). In deze toepassing gebruiken we alleen het neerslagoverschot als invoervariabele. Het neerslagoverschot is bepaald uit de daggemiddelde neerslag en Makkink-referentieverdamping gemeten op het meteoveld. Om de methode te testen is de dataset in twee delen gesplitst. De periode 1981–1990 is gebruikt om het TR-model te kalibreren voor verschillende waarnemingsfrequenties. De periode 1991–1995 is gebruikt ter validatie van de gekalibreerde TR-modellen.

Kalibratie met verschillende waarnemingsfrequenties

Uit de reeks dagelijks waargenomen stijghoogten van de kalibratieperiode (1981–1990) hebben we alternatieve stijghoogtereeksen geselecteerd met afnemende waarnemingsfre-



Figuur 1: Voorspellingen met het TR-model binnen het Kalmanfilteralgoritme en (15-daagse) waarnemingen van de stijghoogte voor periode 1981-1990; TR-model gekalibreerd op de stijghoogtereeks met waarnemingsinterval van 15 dagen.

quenties. Dit resulteerde in reeksen met de volgende regelmatige tijdsintervallen (Δt in dagen) tussen de waarnemingen: 1 dag, 3 dagen, 7 dagen, 10 dagen, 15 dagen, 30 dagen, 60 dagen en 90 dagen. Verder hebben we een reeks van 240 waarnemingen geselecteerd op aselect (*random*) gekozen dagnummers. Dit resulteerde in een reeks waargenomen stijghoogten met gemiddeld 24 waarnemingen per jaar, maar met een zeer onregelmatige waarnemingsfrequentie.

Elk van de 9 alternatieve waarnemingsreeksen is gebruikt om het TR-model te kalibreren volgens de hierboven beschreven methode: m.b.v. het Kalmanfilter en criterium (16). Figuur 1 geeft een voorbeeld van een run met voorspellingen van het Kalmanfilter nadat dit gebruikt is om het TR-model te kalibreren op de reeks met Δt = 15 dagen. Te zien is dat doordat voorspellingen van het TR-model op gezette tijdstippen worden bijgesteld (hier elke 15 tijdstappen), de Kalmanfiltervoorspellingen erg goed zijn (standaardfout van 9,1 cm voor maximaal 15 dagen vooruit).

In tabel 1 staan de parametersets die zijn verkregen door het TR-model te kalibreren op de 9 alternatieve reeksen. Tussen haakjes staan de standaardfouten van de gekalibreerde parameters. Het 95%-betrouwbaarheidsinterval van de gekalibreerde parameter kan worden geschat als de gekalibreerde waarde plus/min twee maal de standaardfout. De standaardfouten zijn bepaald via een foutenanalyse die is beschreven door Bierkens e.a. (1999).

Tabel 1 toont dat voor de aselecte reeks en voor de reeksen met een meetinterval tot 60 dagen vergelijkbare parameters voor het transfermodel (δ_1 en ω_0) worden gevonden. De resultaten voor de parameters van het ruismodel (ϕ_1 en σ_ε^2) zijn iets minder consistent. We zien een klein verloop in σ_ε^2, en de standaardfouten maken een aanzienlijke sprong tussen

Tabel 1: Geschatte parameters van het TR-model verkregen door kalibratie van het TR-model op stijghoogtereeksen met toenemende tijdsduur tussen waarnemingen en met aselect gekozen waarnemingstijdstippen (gemiddeld twee maal per maand).

Δt (dagen)	N	δ_1		ω_0 (dagen^{-1})		ϕ_1		c (cm)		σ_ε^2 (cm^2)	
1	3180	0,9523	(0,0024)	4,705	(0,13)	0,8940	(0,0050)	-130,3	(0,67)	17,43	(0,43)
3	1060	0,9527	(0,0027)	4,947	(0,19)	0,8890	(0,0080)	-130,7	(0,65)	18,02	(0,82)
7	500	0,9480	(0,0034)	5,401	(0,24)	0,9105	(0,0086)	-130,5	(0,78)	15,08	(1,15)
10	318	0,9420	(0,0043)	5,733	(0,28)	0,9288	(0,0083)	-130,2	(0,81)	11,60	(1,23)
Aselect	240	0,9447	(0,0054)	5,486	(0,37)	0,9386	(0,0097)	-130,6	(1,15)	11,82	(1,25)
15	210	0,9427	(0,0052)	5,226	(0,38)	0,9220	(0,015)	-129,9	(0,85)	13,52	(2,12)
30	106	0,9453	(0,0068)	5,185	(0,53)	0,9205	(0,064)	-129,7	(0,89)	13,11	(9,23)
60	52	0,9537	(0,0075)	4,606	(0,61)	0,8269	(0,080)	-129,1	(1,19)	26,76	(7,26)
90	35	0,9579	(0,0098)	3,502	(0,70)	0,9102	(0,060)	-129,7	(1,57)	11,75	(5,36)

de Δt = 15 en Δt = 30 dagen, terwijl voor Δt = 60 dagen de geschatte parameters sterk gaan afwijken. Vergelijking van parameters uit de rij 'Aselect' in tabel 1 (onregelmatige meetintervallen met gemiddelde Δt = 15 dagen) met de parameters voor regelmatige Δt = 15 dagen, laat zien dat de methode ook bij zeer onregelmatige meetfrequenties vergelijkbare resultaten oplevert.

We kunnen dus concluderen dat op deze locatie de parameters van het transfermodel op dagbasis consistent kunnen worden geschat uit reeksen met waarnemingsintervallen tot 60 dagen, en dat de parameters van het ruismodel op dagbasis consistent kunnen worden geschat uit reeksen met waarnemingsintervallen tot 30 dagen. Een reeks met één waarneming per maand is dus op deze locatie voldoende om een TR-model op dagbasis te kalibreren. Verder blijkt dat deze waarnemingsintervallen niet regelmatig hoeven te zijn (zie rij 'Aselect' in tabel 1).

Het maximale waarnemingsinterval waarmee een TR-model kan worden gekalibreerd hangt af van de 'traagheid' van het stijghoogteverloop. Een reeks op een dekzandrug heeft een veel groter geheugen – opeenvolgende waarnemingen lijken veel sterker op elkaar – dan een reeks van een locatie in een beekdal. Hoe trager het stijghoogteverloop, des te groter het interval tussen de waarnemingen mag zijn. Overigens geldt ook dat hoe trager het stijghoogteverloop, des te langer moet de lengte van de reeks zijn, d.w.z. het aantal jaren dat een reeks bevat. Dus een tragere stijghoogtereeks mag minder frequent gemeten worden, maar moet voor een langere periode worden gemeten, om een goede kalibratie van een TR-model mogelijk te maken. Knotters en Van Walsum (1994) geven criteria voor de minimaal benodigde lengte van waarnemingsreeksen. Door Bierkens e.a. (1999) worden criteria gegeven voor de maximaal toegestane meetintervallen. Voor het transfergedeelte van het TR-model van de vorm (4)–(6) (parameters δ_1 en ω_0) geldt

$$\Delta t \leq \frac{-3}{\ln \delta_1} \tag{17}$$

en voor het ruisgedeelte van het TR-model (4)–(6) (parameters ϕ_1 en σ_ε^2) geldt

$$\Delta t \leq \frac{-3}{\ln \phi_1} \tag{18}$$

Hierbij geldt dat Δt wordt berekend in de eenheid van de tijdstap waarmee het TR-model rekent. Dus als het TR-model op dagbasis rekent dan is Δt in dagen, als het TR met tijdstappen van maanden rekent is Δt in maanden. Als we deze criteria toepassen op het TR-model op dagbasis voor locatie 32CL0034 (met δ_1 en ϕ_1 verkregen voor Δt = 1 dag), vinden we maximale meetintervallen van 61 en 23 dagen voor respectievelijk het transfermodel en het ruismodel. Dit komt zeer goed overeen met wat we vonden uit het kalibratie-experiment. In de praktijk kunnen criteria (17) en (18) worden gebruikt om na de kalibratie van het TR-model (dus met de hieruit geschatte parameters) na te gaan of de frequentie van de gebruikte waarnemingsreeks voldoende is geweest om de parameters van het TR-model te kunnen schatten.

Validatie

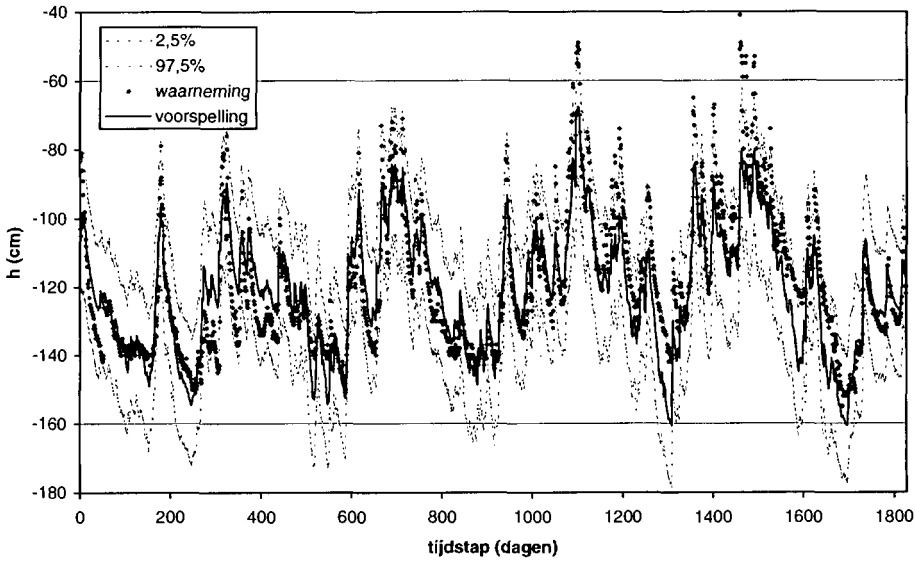
De TR-modellen die uit de 9 alternatieve stijghoogtereeksen voor de periode 1981–1990 zijn verkregen, zijn gevalideerd op de dagwaarnemingen voor de periode 1991–1995. Validatie is op twee manieren gebeurd:

1 validatie door vergelijking van voorspellingen met het transfermodel (vergelijking 5) met de waarnemingen. Figuur 2 toont de voorspellingen met het gekalibreerde TR-model op dagbasis (gekalibreerd op de reeks met Δt = 15 dagen) en de dagwaarnemingen voor de periode 1991–1995. Ook het 95%-voorspellings-interval (betrouwbaarheidsband) is aangegeven. Deze kan berekend worden uit de voorspelling met het transfermodel H_t^* en de parameters van het ruismodel ϕ_1 en σ_ε^2 als:

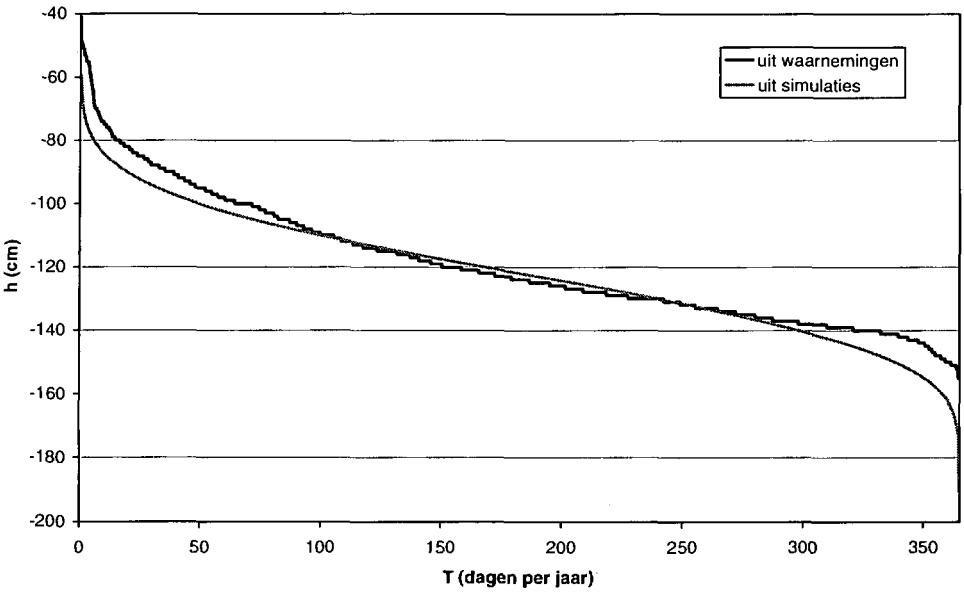
$$\left[H_t^* - 1{,}96\sqrt{\frac{\sigma_\varepsilon^2}{1-\phi_1^2}}, H_t^* + 1{,}96\sqrt{\frac{\sigma_\varepsilon^2}{1-\phi_1^2}}\right]$$

2 validatie van de statistieken van met het TR-model gesimuleerde stijghoogtereeksen met de statistieken berekend uit de dagelijkse waarnemingen voor de periode 1991–1995. Een belangrijke statistiek is bijvoorbeeld de duurlijn: het verwachte aantal dagen per jaar dat de stijghoogte een bepaald niveau overschrijdt. Figuur 3 toont de duurlijn berekend uit simulaties met het TR-model op dagbasis (gekalibreerd op de reeks met Δt = 15 dagen) met de duurlijn berekend uit de waarnemingen.

Vanwege de lengte zullen we de resultaten van de validatie hier niet geven. Deze staan vermeld in Bierkens e.a. (1999). In het kort komt het er op neer dat vergelijkbare validatieresultaten worden bereikt voor alle parametersets uit kalibratie op $\Delta t \leq 60$ dagen, alsmede voor de parameterset uit kalibratie op de aselecte meettijdstippen. We vonden voor validatiemethode 1 (vergelijking voorspellingen met waarnemingen) een gemiddelde fout van ca 2,0 cm en een standaardfout van circa 9,5 cm. Voor validatiemethode 2 (vergelijking duurlijnen) vonden we een gemiddelde fout van circa 2,3 cm en een standaardfout van circa 5,2 cm.



Figuur 2: Voorspellingen, 95%-voorspellingsintervallen en (dagelijkse) waarnemingen van de stijghoogte voor de validatieperiode 1991-1995; TR-model gekalibreerd op de stijghoogtereeks met waarnemingsinterval van 15 dagen.



Figuur 3: Duurlijnen van de stijghoogte voor de periode 1991-1995 zoals geschat uit de (dagelijkse) waarnemingen en met het TR gesimuleerde (dagelijkse) stijghoogten; TR-model gekalibreerd op de stijghoogtereeks met waarnemingsinterval van 15 dagen.

Conclusies en discussie

Met de methode die in dit artikel is gepresenteerd is het mogelijk om TR-modellen die werken met dezelfde tijdstap als de invoerreeksen, toe te passen op uitvoerreeksen die met een veel lagere of zelfs onregelmatige frequentie zijn gemeten. Op deze wijze kunnen hydrologische tijdreeksen worden verdicht en verlengd en kunnen gaten in deze reeksen worden opgevuld. Toepassing van de methode op een locatie met dagelijks waargenomen stijghoogten voor de periode 1981–1995 liet zien dat op deze locatie een gemiddelde waarnemingsfrequentie van 1 maal per maand voldoende is om het TR-model met een tijdstap van 1 dag te kalibreren.

In dit artikel zijn ook criteria gegeven waarmee achteraf kan worden nagegaan of de waarnemingsfrequentie van de gebruikte uitvoerreeks voldoende is geweest om de parameters van het TR-model te kunnen schatten. Deze criteria kunnen eventueel ook gebruikt worden voor het terugbrengen van de waarnemingsfrequentie ten behoeve van kostenbesparing. Men zou bijvoorbeeld na de installatie van een stijghoogtebuis eerst gedurende een kortere periode met een hogere frequentie (bijvoorbeeld 1 maal per week) kunnen waarnemen. Op de kort waargenomen tijdreeks kan het TR-model worden gekalibreerd. Vervolgens kan dan uit de TR-parameters worden nagegaan wat de minimaal vereiste waarnemingsfrequentie is. Vervolgens kan met deze (of een iets hogere) frequentie (bijv. 1 maal per maand) worden verder gemeten. Als het nodig is kan met het gekalibreerde TR-model de minder frequent gemeten tijdreeks altijd weer worden verdicht, mits er in de tussentijd geen grote veranderingen in het grondwatersysteem hebben plaatsgevonden.

Literatuur

Bierkens, M.F.P. (1998) Eenvoudige stochastische modellen voor grondwaterstandsfluctuaties, deel 1: Een stochastische differentiaalvergelijking; in: *Stromingen*, jrg 4, nr 2, pag 5–26.

Bierkens, M.F.P., M. Knotters en F.C. van Geer (1999) Calibration of Transfer function-noise models to sparsely or irregularly observed time series; in: *Water Resources Research* (in druk).

Box, G.E.P. en G.M. Jenkins (1976) Time Series Analysis: forecasting and control; revised edition; Holden Day, San Francisco.

Chow, K.C.A., W.A. Watt en D.G. Watts (1983) A stochastic-dynamic model for real time flood forecasting; in: *Water Resources Research*, jrg 19, nr 3, pag 746–752.

Geer F.C. van en P.R. Defize (1987) Detection of natural and artificial causes of groundwater fluctuations; in: *The influence of climate change and climatic variability on the hydrologic regime and water resources*; IAHS publication nr 168, pag 597–606.

Geer, F.C. van en A. F. Zuur (1997) An extension of Box-Jenkins transfer/noise models for spatial interpolation of groundwater head series; in: *Journal of Hydrology*, nr 192, pag 65–80.

Gehrels, J.C., F.C. van Geer en J.J. de Vries (1994) Decomposition of groundwater level fluctuations using transfer modelling in an area with shallow to deep unsaturated zones; in: *Journal of Hydrology*, nr 157, pag 105–138.

Gurnell, A.M. en C.R. Fenn (1984) Box-Jenkins transfer function models applied to suspended sediment concentration-discharge relationships in a proglacial stream; in: *Arctic and Alpine Research*, jrg 16, nr 1, pag 93–106.

Hipel, K.W., W.C. Lennox, T.E. Unny en A.I. McLeod (1975) Intervention analysis in water resources; in: *Water Resources Research*, jrg 11, nr 6, pag 855–861.

Hipel, K.W. en A.I. McLeod (1994) Time Series Modelling of Water Resources and Environmental Systems; Developments in Water Science, nr 45, Elsevier Science, Amsterdam.

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*, nr 197, pag 25–46.

Knotters, M. en J.G. de Gooijer (1999) TARSO modeling of water table depths; in: *Water Resources Research*, jrg 35, nr 3, pag 695–705.

Knotters, M. en M.F.P. Bierkens (1999) Tijdreeksmodellen voor de grondwaterstand: een kijkje in de black box. Aangeboden aan *Stromingen*.

Lemke, K.A. (1991) Transfer function models of suspended sediment concentration; in: *Water Resources Research*, jrg 27, nr 3, pag 293–395.

Rolf, H.L.M. (1989) Verlaging van grondwaterstanden in Nederland; analyse periode 1950–1986; Rapport DGV–TNO en Rijkswaterstaat.

Schweppe, F.C. (1973) Uncertain Dynamic Systems; Prentice-Hall Inc., Englewood Cliffs, New Jersey.

Shea, B.L. (1987) Estimation of multivariate time series; in: *Journal of Time Series Analysis*, jrg 8, nr 1, pag 95-109.

Stroet, C.B.M. te (1995) Calibration of Stochastic Groundwater Flow Models: Estimation of System Noise Statistics and Model Parameters; proefschrift Technische Universiteit Delft.

Tankersley, C.D., W.D. Graham en K. Hatfield (1993) Comparison of univariate and transfer function models of groundwater fluctuations; in: *Water Resources Research*, jrg 29, nr 10, pag 2517–3533.

Tong, H. (1990) Non-linear Time Series; a dynamical system approach; Clarendon Press, Oxford.