

Beoordeling Tijdreeksmodellen

D.A. Brakenhoff, O.N. Ebbens

January 20, 2021

1 Beoordeling Tijdreeksmodellen

Auteurs: O.N. Ebbens (Artesia), D.A. Brakenhoff (Artesia)

Een veelgestelde vraag is hoe je een geoptimaliseerd tijdreeksmodel moet beoordelen. In dit notebook worden methodes voor modelbeoordeling gepresenteerd en geïllustreerd met voorbeelden uit de praktijk.

1.1 Inhoudsopgave

- Section 1.2
- Section 1.3
- Section 1.4
- Section 1.5
- Section 1.6
- Section 1.7
- Section 1.8
- Section 1.9
- Section 1.10

1.2 Introductie

Voor het gebruiken of afkeuren van een tijdreeksmodel is het belangrijk om vast te stellen wat het doel van het tijdreeksmodel is. Op basis van dit doel kunnen de criteria worden gesteld voor het beoordelen van het model. In dit notebook worden verschillende methodes toegelicht om tijdreeksmodellen te beoordelen.

Dit notebook geeft een overzicht van de meest voorkomende methodes om tijdreeksmodellen te beoordelen. Per onderwerp wordt de belangrijkste achtergrondinformatie gepresenteerd, maar het is zeker geen volledig overzicht van al het onderzoek en werk wat op dat vlak is gedaan. Voor de lezers die daarin geïnteresseerd zijn hebben we referenties opgenomen waarmee men dieper in het onderwerp kan duiken.

<IPython.core.display.HTML object>

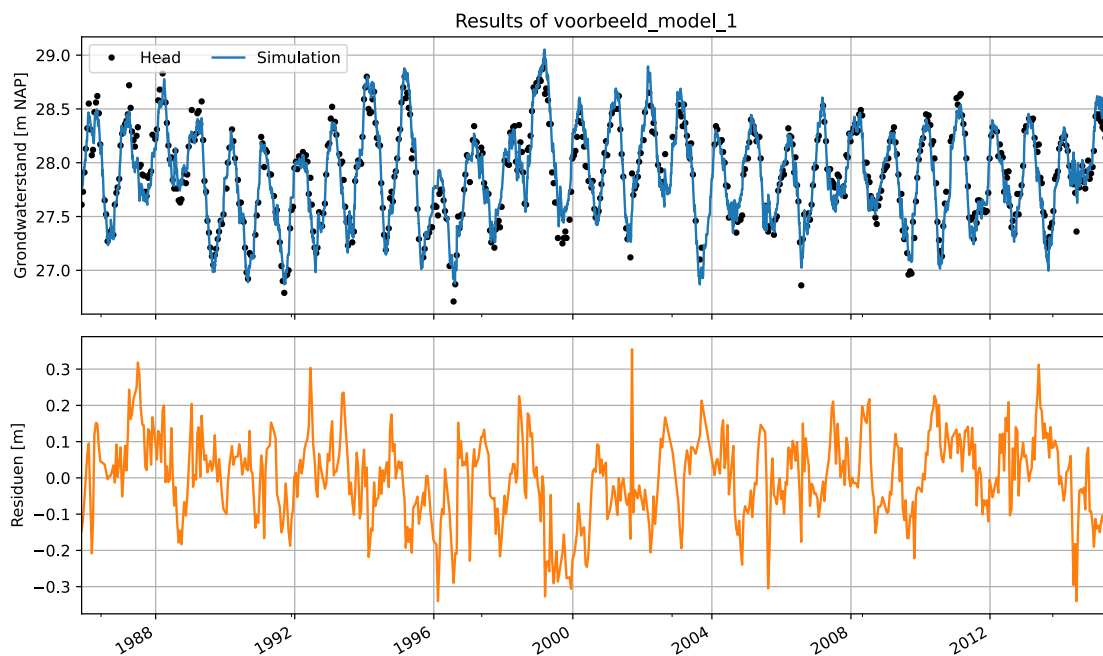
1.3 Visuele beoordeling

Een eerste, eenvoudige manier om een tijdreeksmodel te beoordelen is het visualiseren van de resultaten. De modelsimulatie wordt samen met de metingen in een grafiek weergegeven. Vaak worden

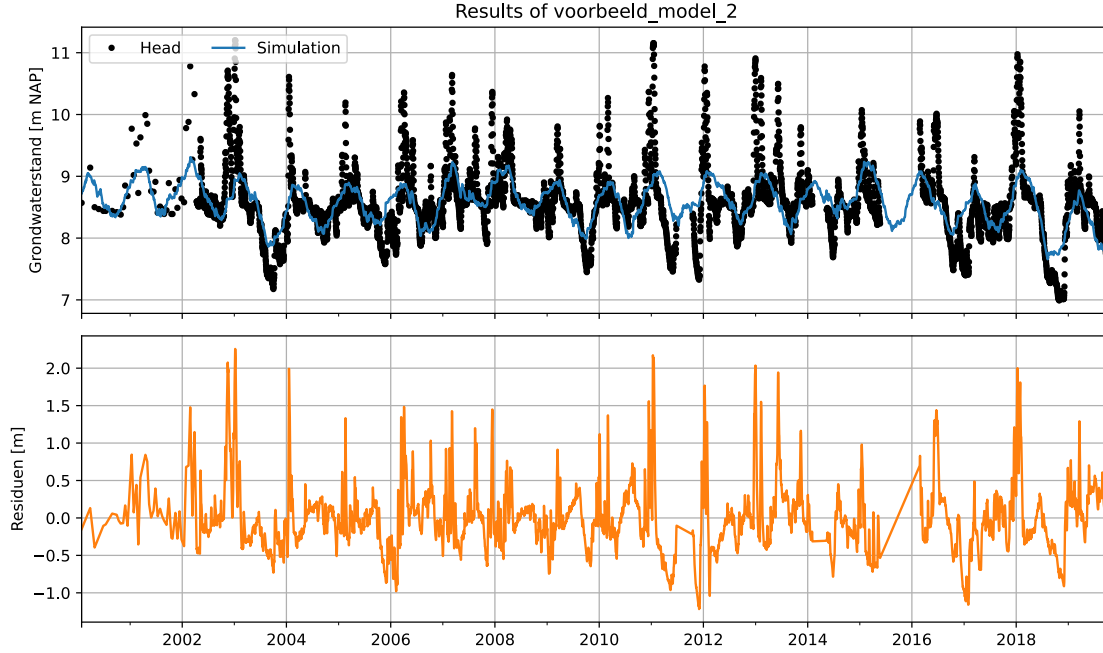
de modelresiduën ook gepresenteerd. Aan de hand van beide grafieken kan visueel beoordeeld worden of de simulatie door de metingen loopt, of de metingen helemaal niet goed benadert. Ook kan beoordeeld worden in welke periodes het model het goed doet, en in welke periodes niet. Is het model bijvoorbeeld goed in staat om de pieken en dalen te simuleren?

Voorbeeld visuele beoordeling 1 Als voorbeeld van een visuele beoordeling is een tijdreeksmodel geselecteerd waarin de grondwaterstand wordt verklaard aan de hand van neerslag en verdamping. De modelsimulatie (blauwe lijn) en de metingen (zwarte punten) staan in de bovenste grafiek. De residuen (oranje lijn) zijn in de onderste grafiek weergegeven.

In de figuur is te zien dat de simulatie en de metingen behoorlijk goed overeen komen. In de residuen is te zien dat het verschil tussen model en meting het grootste is bij de extremen: de dalen in de zomer en de pieken in de winter.



Voorbeeld visuele beoordeling 2 Onderstaand voorbeeld laat een tijdreeksmodel zien waar de modelsimulatie (blauwe lijn) slecht overeenkomt met de metingen (zwarte punten). De simulatie wijkt in sommige gevallen wel 2 meter af van de gemeten grondwaterstand. Met dit model is het uiteraard niet mogelijk om betrouwbare simulaties te maken.



1.4 Fit statistieken

Naast de visuele beoordeling kunnen de modelresultaten ook worden beoordeeld met behulp van statistische parameters. Er zijn verschillende statistische parameters om de fit tussen gemodelleerde en gemeten grondwaterstanden in getallen uit te drukken. De statistieken kunnen opgedeeld worden in statistieken die de modelfout (de residuen) uitdrukken als een getal, en statistieken die proberen de “goodness-of-fit” uit te drukken als een getal. Section 1.10 behandelt een aantal fit statistieken van tijdreeksmodellen. Section 1.10 geeft een overzicht van statistieken die gebruikt worden voor hydrologische modellen. In onderstaande tabel zijn een aantal veelgebruikte statistieken opgenomen.

afkorting	statistische parameter	formule *	waarde bij optimale fit	minimum	maximum	type
SSE	som kwadratische fout	$\sum_{i=1}^N (y_i - \hat{y}_i)^2$	0	0	∞	residuën
MAE	gemiddelde absolute fout	$\sum_{i=1}^N \frac{\ y_i - \hat{y}_i\ }{N}$	0	0	∞	residuën

afkorting	statistische parameter	formule *	waarde bij optimale fit	minimum	maximum	type
RMSE	wortel kwadratisch gemiddelde fout	$\sqrt{\sum_{i=1}^N \frac{(y_i - \hat{y}_i)^2}{N}}$	0	0	∞	residuën
EVP	verklaarde variantie	$\frac{\sigma_y^2 - \sigma_r^2}{\sigma_y^2} * 100$	100%	0	100	goodness- of-fit
R^2	determinatiecoëfficiënt	$1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$	1	0	1	goodness- of-fit
NSE	Nash-Sutcliffe model efficiëntie	$1 - \frac{\sum_{i=1}^N (y_i - f_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$	1	$-\infty$	1	prognosefout

* N is het aantal metingen, y_i en \hat{y}_i zijn de gemeten en met het model gefitte grondwaterstand op tijdstip i , respectievelijk. f_i is de met het model voorspelde grondwaterstand op tijdstip i en σ_y^2 en σ_r^2 zijn de variantie in de gemeten grondwaterstanden en de variantie van de residuën, respectievelijk.

1.4.1 Verschillen R^2 , NSE en EVP

De R^2 en NSE worden regelmatig door elkaar gebruikt omdat ze grote gelijkenissen vertonen. Het verschil zit hem in de toepassing. De R^2 wordt gebruikt om te kijken hoe goed het lukt om een model te fitten op metingen, de NSE om te kijken hoe goed een modelvoorspelling past bij metingen. De NSE kan negatief worden, omdat de gekwadrateerde fout van de modelvoorspelling groter kan zijn dan de totale gekwadrateerde som. In het geval van de R^2 is het slechtst mogelijke model in theorie het gemiddelde van de metingen waardoor de waarde van R^2 nooit lager kan worden dan nul.

Ook de EVP en de R^2 tonen veel gelijkenissen. In principe verschilt de EVP in waarde van de R^2 tenzij het gemiddelde van de modelresiduën gelijk is aan nul. In dat geval is de waarde van de R^2 , mits uitgedrukt als percentage, gelijk aan de EVP. Bij modellen die gefit zijn op de metingen is het gemiddelde van de modelresiduën over het algemeen nagenoeg gelijk aan nul.

Voorbeeld fit statistiek 1 De fit statistieken van het hierboven weergegeven model. De gemiddelde absolute modellfout bedraagt ca. 9 cm. De fit-statistieken geven aan dat het model goed overeenkomt met de metingen; de verklaarde variantie (of EVP) is 92,9%.

	waarde	eenheid
Mean absolute error (mae)	0.090	m
Root mean squared error (rmse)	0.114	m
Sum of the squares of the error (sse)	8.431	m ²
Explained variance percentage (evp)	92.926	%
R-squared (rsq)	0.929	-

Voorbeeld fit statistiek 2 De fit statistieken van het tweede voorbeeldmodel zijn een stuk slechter. Zoals ook al visueel aangegeven komt het model niet overeen met de metingen. De verklaarde variantie (EVP) waarde is slechts 39,6% en de gemiddelde afwijking maar liefst 33 cm.

	waarde	eenheid
Mean absolute error (mae)	0.326	m
Root mean squared error (rmse)	0.464	m
Sum of the squares of the error (sse)	1281.269	m ²
Explained variance percentage (evp)	38.597	%
R-squared (rsq)	0.386	-

1.4.2 Overfitting

Als er meer verklarende variabelen (en daarmee modelparameters) aan een tijdreeksmodel worden toegevoegd wordt de fit over het algemeen beter, zelfs als de verklarende variabele helemaal geen invloed heeft op de grondwaterstand. Het model kan deze extra vrijheidsgraden (parameters) namelijk gebruiken om het model beter op de metingen te fitten. Het toevoegen van een niet-relevante verklarende variabele zorgt niet voor een verbetering in de voorspellende waarde van het model. Het fenomeen waarbij meer parameters aan een model worden toegevoegd dan gerechtvaardigd is op basis van de data heet overfitting. Overfitting kan vaak herkend worden aan de correlaties tussen parameters (zie Section 1.7) en de standaardfouten van de parameters (zie Section 1.6) (Section 1.10).

Dit is belangrijk omdat het dus niet altijd mogelijk is om op basis van de fit statistieken aan te geven of het model verbetert na het toevoegen van een verklarende variabele. Er bestaan diverse methodes om te bepalen of de voorspellende waarde van een model verbetert na het toevoegen van verklarende variabelen. Section 1.10 laat een aantal methodes zien om modellen op dit criteria te beoordelen.

1.5 Cross validatie

Cross validatie is een methode om tijdreeksmodellen te beoordelen. In plaats van de term cross validatie wordt ook vaak gesproken over modelvalidatie. Hierbij wordt een tijdreeks opgesplitst in twee delen, een trainings-deel en een test-deel. Het deel van de metingen dat gebruikt wordt om het model te kalibreren heet de trainings-set. Vervolgens wordt het gekalibreerde model gebruikt om een simulatie te berekenen voor de tweede periode. Dit deel wordt de test-set genoemd. De simulatie wordt vergeleken met de metingen om te beoordelen of het model ook goede resultaten oplevert voor het deel waar het niet op gekalibreerd is. Een tijdreeks moet voldoende lang zijn en voldoende metingen bevatten om deze op te kunnen splitsen trainings-deel en een test-deel.

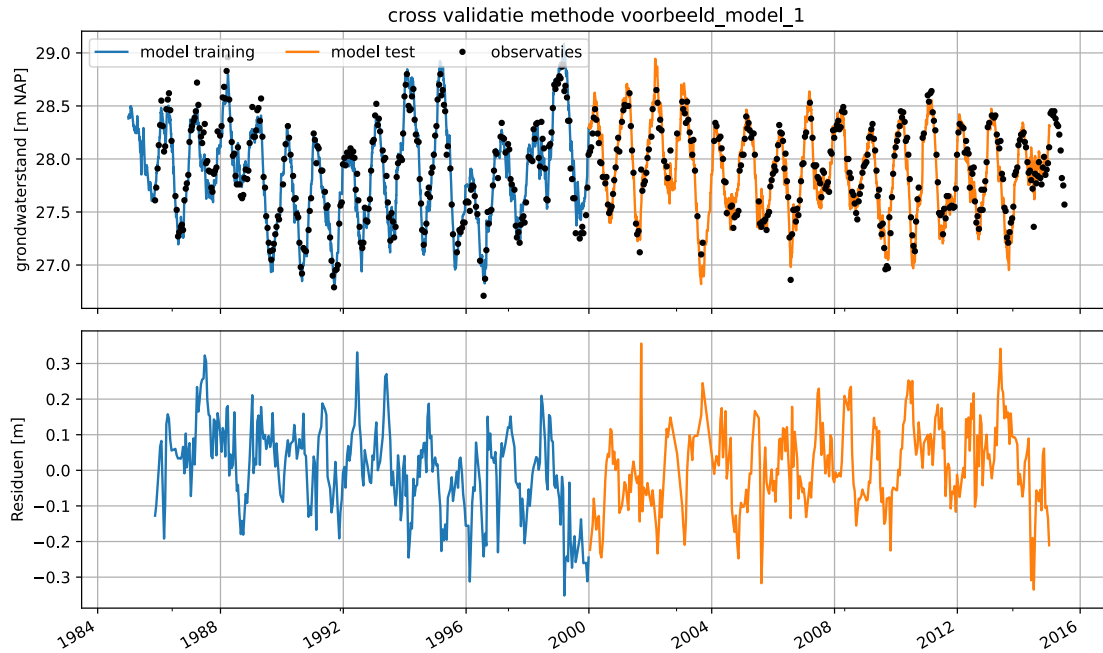
Als het model goed presteert in de test-periode dan is het model dat is afgeleid op basis van de trainings-periode goed genoeg om de grondwaterstanden te voorspellen in die periode. Dat geeft vertrouwen dat het model goed presteert, maar dat betekent niet per definitie dat het model ook in de toekomst goede voorspellingen zal geven. Als er in de toekomst iets structureel veranderd in het systeem (b.v. een onttrekking die start met pompen buiten beide geselecteerde periodes), dan zal het model dat uiteraard niet goed voorspellen.

Als het model een goede fit heeft in de trainings-periode maar niet in de test-periode, dan is het model geen goede voorspeller van de grondwaterstand. Dit kan verschillende oorzaken hebben. Oorzaken zijn bijvoorbeeld veranderingen in het systeem die tijdens de test-periode hebben plaatsgevonden zoals de start van een nieuwe onttrekking of het dempen van een watergang, of het meenemen van een invloed in het model die tijdens de trainingsperiode nog niet actief was.

De hierboven besproken Section 1.4 kunnen ook gebruikt worden om de voorspelling in de testperiode te toetsen. In dit geval is in plaats van de R^2 de NSE toegepast omdat het om een voorspelling met het model gaat. Voor een uitgebreide analyse van een set aan tijdreeksmodellen met deze methode, inclusief een beschouwing van bijbehorende fit-statistieken, zie Section 1.10.

Voorbeeld cross validatie 1 De cross validatie is uitgevoerd voor het voorbeeldmodel hierboven. De dataset is in twee periodes opgesplitst. Het model is getraind op de periode 1985 - 2000 en getest op de periode 2000 - 2015. De resultaten staan in de plots hieronder evenals de berekende fit statistieken beide periodes.

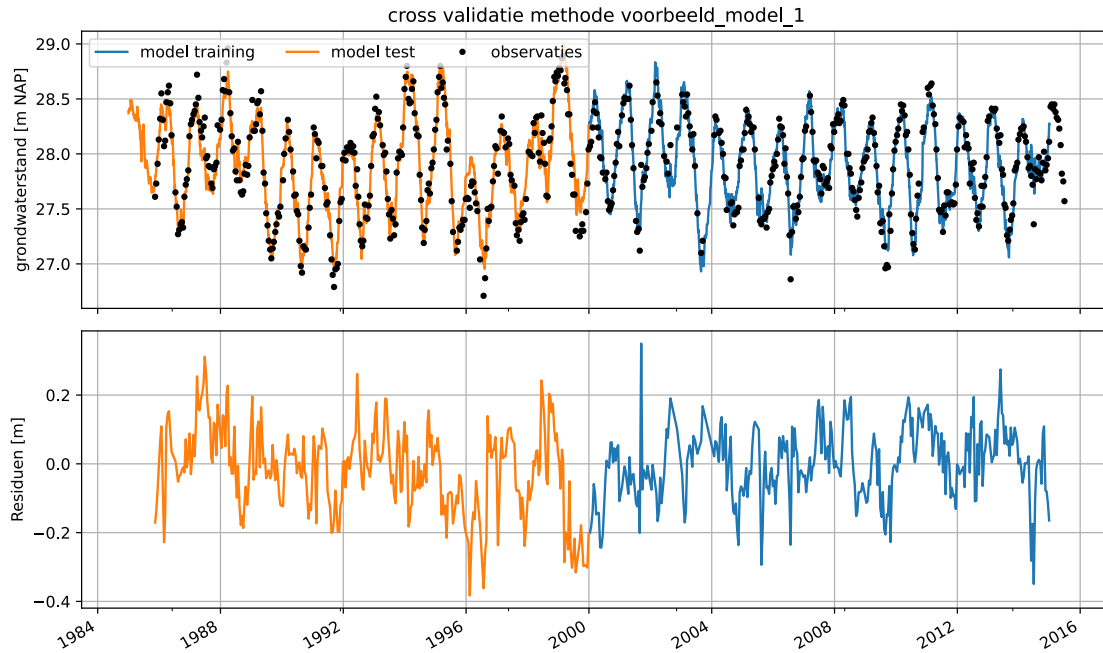
Het model presteert vrij goed in de test periode wat vertrouwen geeft in de voorspellende waarde van het model. De EVP is in de testperiode weliswaar 3 procentpunt lager dan in de trainingsperiode maar met 90,5% nog steeds relatief hoog. Aan de fit statistieken is duidelijk te zien dat het model vergelijkbaar presteert in de testperiode vergeleken met de trainingsperiode.



	training: 1985-2000	test: 2000-2015	eenheid
Mean absolute error (mae)	0.097	0.094	m
Root mean squared error (rmse)	0.123	0.117	m
Sum of squares of the error (sse)	4.727	4.353	m^2
Explained variance percentage (evp)	93.441	90.524	%
Nash-Sutcliffe model efficiency (nse)	0.934	0.905	-

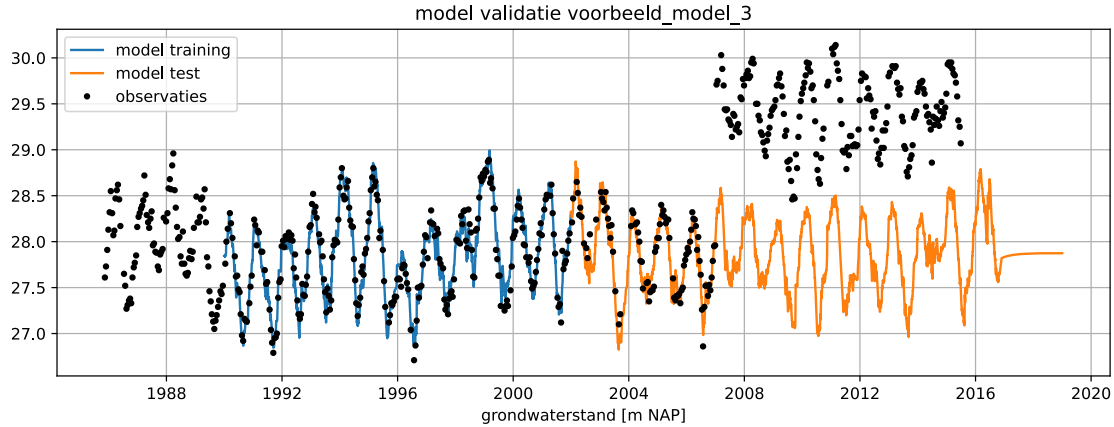
Voorbeeld cross-validatie 2 Bij de cross-validatie methode wordt vaak een trainingsperiode in het verleden gekozen en een testperiode die daarna ligt. Deze kunnen echter ook worden omgedraaid. In onderstaand voorbeeld is hetzelfde model gebruikt als in het vorige voorbeeld maar zijn de trainings- en testperiode omgedraaid.

De resultaten laten eenzelfde beeld zien als in het eerste voorbeeld. De EVP in de periode 2000-2015 is verbeterd met 2 procentpunten. Deze toename is verklaarbaar omdat het model nu gefit is op deze periode. Daarnaast valt op dat het model nu in de trainingsperiode een slechtere EVP en NSE waarde heeft dan in de testperiode. Dit wordt waarschijnlijk veroorzaakt door slecht verklaarbare metingen in de trainingsperiode.



	training: 2000-2015	test: 1985-2000	eenheid
Mean absolute error (mae)	0.080	0.096	m
Root mean squared error (rmse)	0.101	0.124	m
Sum of squares of the error (sse)	3.299	4.855	m^2
Explained variance percentage (evp)	92.776	93.393	%
Nash-Sutcliffe model efficiency (nse)	0.928	0.933	-

Voorbeeld cross-validatie 3 Hieronder een voorbeeld van een model met een goede fit in de trainingsperiode en een slechte fit in de testperiode. De slechte fit wordt veroorzaakt door een sprong in de grondwaterstanden in 2007. Deze sprong is veroorzaakt door een externe invloed die niet is meegenomen bij het trainen van het model. Het zal daarom niet mogelijk zijn om met het model een goede fit te krijgen voor de testperiode.



	training: 1990-2003	test: 2003-2019	eenheid
Mean absolute error (mae)	0.091	1.132	m
Root mean squared error (rmse)	0.115	1.315	m
Sum of squares of the error (sse)	3.637	474.127	m^2
Explained variance percentage (evp)	93.994	23.061	%
Nash-Sutcliffe model efficiency (nse)	0.940	-1.696	-

1.6 Betrouwbaarheid modelparameters

%% PM: Referentie naar Pastas notebook model diagnostiek? Rapport Paul & Eit? %%

In tijdreeksmodellering volgens de PIRFICT methode worden externe invloeden worden gemodelleerd met responsfuncties die ieder één of meerdere parameters bevatten Section 1.10. Daarnaast kunnen in tijdreeksmodellen ook nog een constante en een ruismodel worden toegevoegd. Tijdreeksmodellen kunnen beoordeeld worden aan de hand van de berekende onzekerheden van deze modelparameters. Het is goed voor te stellen dat een model waarvan de parameters lastig te bepalen zijn (de parameters zijn onzeker) minder betrouwbaar is dan een model waarbij de parameters goed vast te stellen zijn (met een kleine onzekerheid).

Voor het beschouwen van de onzekerheden van de parameters moet de modelruis (de ruis is het resultaat van het toepassen van een ruismodel op de residuen) voldoen aan bepaalde statistische voorwaarden. Vaak wordt een ruismodel toegepast om ervoor te zorgen dat de verschillen tussen het model en de metingen aan deze statistische eisen voldoet. De voorwaarden voor de ruis zijn dat het:

- zich gedraagt als witte ruis met een gemiddelde van nul. Er is sprake van witte ruis als:
 - er geen significante autocorrelatie in de ruis aanwezig is
 - de ruis homoscedastisch is
 - de ruis voldoet aan een, bij voorkeur normale, statistische verdeling.
- niet gecorreleerd is aan een verklarende tijdreeks.

Er wordt in dit notebook geen uitgebreide uitleg gegeven over deze voorwaarden. Er wordt wel gepresenteerd hoe deze aspecten visueel beoordeeld kunnen worden. Naast een visuele beoordeling bestaan er allerlei diagnostische toetsen die statistisch beoordelen of wordt voldaan aan deze voorwaarden. Het voordeel van deze toetsen is dat ze meer objectief zijn dan een visuele controle. Het nadeel is dat ze in veel gevallen alleen werken voor equidistante tijdstappen, en dus formeel niet toegepast mogen worden op reeksen met een variabele tijdstap. Zie bijvoorbeeld [deze notebook](#) voor implementatie en een uitgebreidere beschrijving van dergelijke diagnostische toetsen Section 1.10.

Samenvattend mogen de berekende onzekerheden van de parameters gebruikt worden in verdere analyse als voldaan wordt aan de hierboven genoemde statistische voorwaarden.

Voorbeeld betrouwbaarheid modelparameters De onzekerheid in de modelparameters is berekend voor het voorbeeld model en weergegeven in onderstaande tabel. De standaardafwijking van de geschatte parameters is relatief klein ten opzichte van de absolute waarde van de parameters. Het lijkt er dus op dat de parameters relatief nauwkeurig ingeschat kunnen worden in de optimalisatie. Maar we weten nog niet of we deze berekende standaardafwijking (of onzekerheid) kunnen vertrouwen, omdat we nog niet hebben beschouwd of het model doet aan de hierboven genoemde statistische voorwaarden. Dus in de volgende voorbeelden wordt onderzocht of het model aan die voorwaarden voldoet.

parameter	optimale waarde	standaardafwijking	standaardafwijking (%)
recharge_A	686.247	36.404	$\pm 5.30\%$
recharge_a	159.386	7.991	$\pm 5.01\%$
recharge_f	-1.305	0.053	$\pm 4.04\%$
constant_d	27.920	0.059	$\pm 0.21\%$
noise_alpha	49.912	5.919	$\pm 11.86\%$

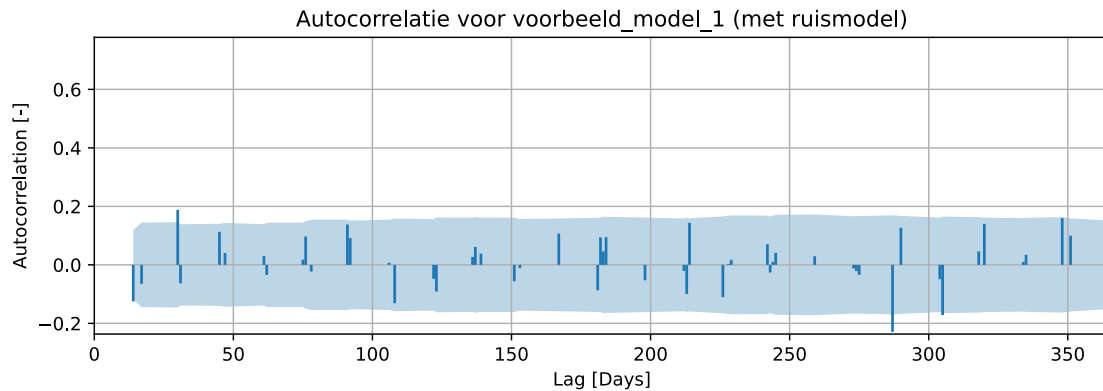
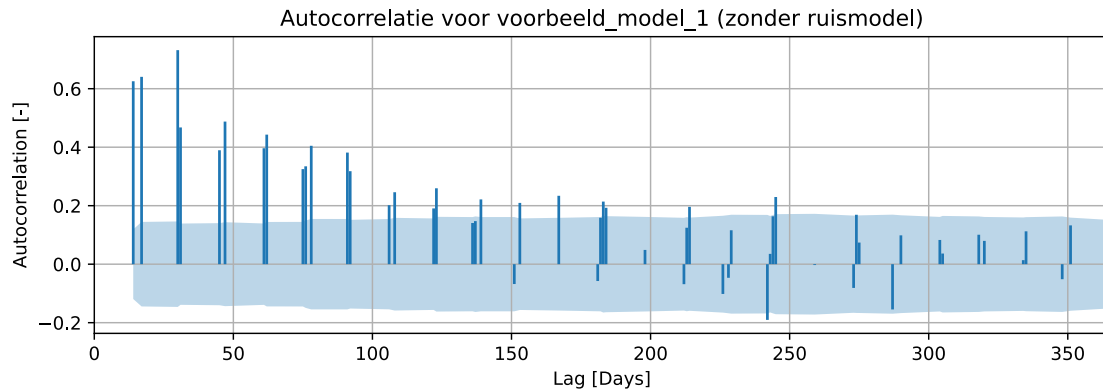
1.6.1 Autocorrelatie

Autocorrelatie is de correlatie van een tijdreeks met een in de tijd verschoven versie van zichzelf. De betrouwbaarheden van de schatting van de modelparameters wordt te laag ingeschat als er nog significante autocorrelatie aanwezig is in de ruis. Deze vorm van correlatie kan worden weergegeven in een autocorrelatiediagram. Daarin is op de x-as de verschuiving in de tijd weergegeven (lag) en op de y-as de berekende correlatie.

Voorbeeld autocorrelatiediagram De autocorrelatie is voor het voorbeeld model berekend met en zonder een ruismodel. De resultaten zijn weergegeven in onderstaande figuren. De maximale verschuiving is 365 dagen. Er is niet voor elke verschuiving een correlatie berekend omdat de tijdreeks die we beschouwen metingen om de ± 14 dagen bevat.

De blauwe band in de grafiek geeft het 95% betrouwbaarheidsinterval weer. Als 95% van de berekende autocorrelaties binnen dit interval liggen kunnen we concluderen dat er geen significante autocorrelatie aanwezig is in de reeks. Het is duidelijk te zien dat de residuen van het model zonder ruismodel niet voldoen aan de eis voor de autocorrelatie. De berekende onzekerheden in

de parameters zal dus niet goed ingeschat zijn. In het model met ruismodel zien we dat er geen significante autocorrelatie aanwezig is in de ruis.

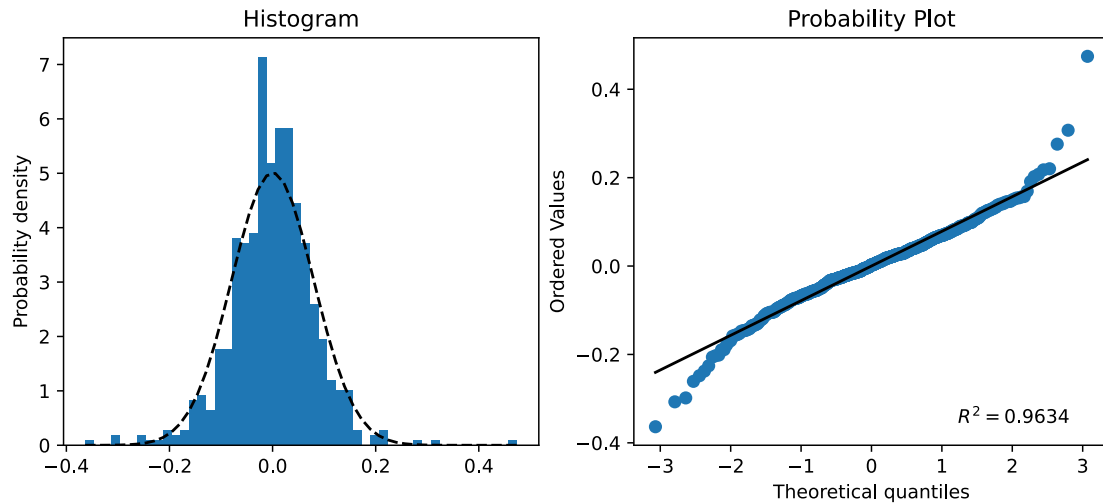


1.6.2 Normale verdeling ruis

Een andere voorwaarde voor witte ruis is dat deze voldoet aan een, bij voorkeur normale, statistische verdeling. De verdeling van de ruis kan weergegeven worden met een histogram in combinatie met een normale kansverdeling. Visueel kan dit beoordeeld worden door een histogram te maken van de residuen/ruis en deze te vergelijken met een normale verdeling met μ gelijk aan het gemiddelde van de ruis, en σ gelijk aan de standaardafwijking van de ruis.

Voorbeeld visualisatie normale verdeling ruis Voor het voorbeeld model is een histogram geplott van de ruis en wordt deze vergeleken met de normale verdeling. Dit is in de linker grafiek weergegeven. In de rechter grafiek is eigenlijk dezelfde analyse op een andere manier weergegeven, een zogenaamde “probability plot”. Als de blauwe punten op de zwarte rechte lijn liggen is de ruis normaal verdeeld.

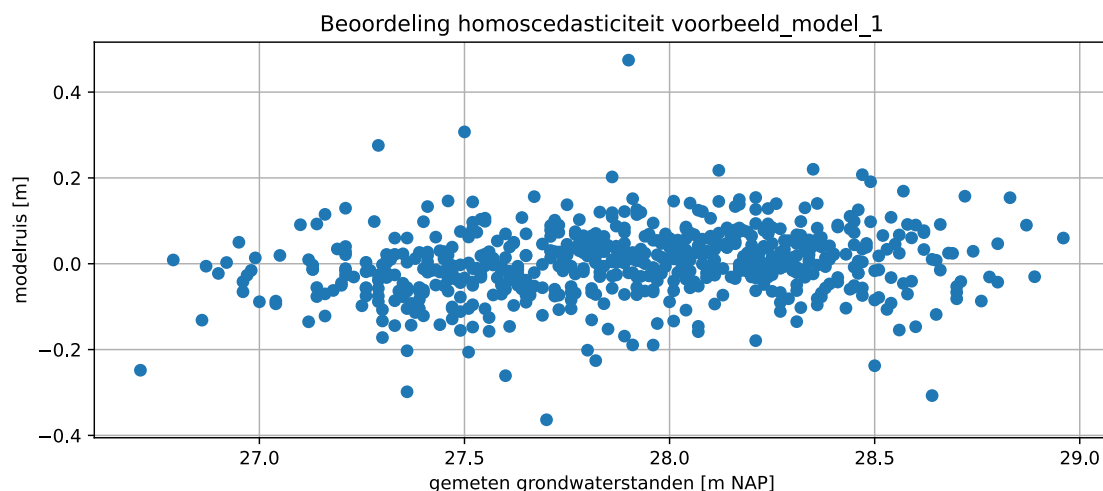
Op basis van deze resultaten kan visueel wordt ingeschat dat de verdeling van de ruis de normale verdeling voldoende goed benaderd om het model op dit aspect goed te keuren.



1.6.3 Homoscedasticiteit

De derde statistische voorwaarde voor witte ruis is dat de ruis homoscedastisch moet zijn. Een tijdreeks is homoscedastisch als de variantie onafhankelijk is van de absolute waarde van de reeks. Voor een tijdreeksmodel voor het simuleren van stijghoogtes betekent dit dat de variantie van de ruis onafhankelijk is van de gemeten stijghoogtes. Als de ruis niet homoscedastisch is (= heteroscedastisch) dan is de variantie van de ruis afhankelijk van de grondwaterstand. Dit kan visueel beoordeeld worden door de ruis uit te zetten tegen de stijghoogtemetingen. Als de spreiding van de ruis ongeveer gelijk is over het gehele bereik van de stijghoogtemetingen dan is de ruis homoscedastisch.

Voorbeeld visualisatie homoscedasticiteit Voor het voorbeeld model is de ruis uitgezet tegen de metingen. De spreiding in de ruis lijkt redelijk constant te zijn over het bereik van de stijghoogtemetingen. De ruis lijkt dus homoscedastisch.



1.6.4 Correlatie verklarende tijdreeks

%% PM %%

1.7 Correlatie modelparameters

Tijdens het optimaliseren van een tijdreeksmodel wordt ook de correlatie tussen de geschatte parameters berekend. Als twee parameters sterk gecorreleerd zijn betekent dat deze in de optimalisatie moeilijk van elkaar te onderscheiden zijn. Dit heeft meestal geen invloed op de voorspellende kracht van het tijdreeksmodel in zijn geheel, maar mogelijk wel op het inschatten van individuele bijdrages van de verklarende variabelen.

Voorbeeld parameter correlaties 1 Je maakt een grondwatermodel met als verklarende reeksen recharge en oppervlaktewaterpeilen. Het model presteert goed maar de parameters van beide verklarende variabelen zijn sterk gecorreleerd. Het model is daardoor niet geschikt om de invloed van recharge en oppervlaktewater van elkaar te scheiden. Dus kan het model niet gebruikt worden om het effect van een stijging in het oppervlaktewaterpeil te berekenen. Het is immers niet duidelijk of de stijging van de grondwaterstand wordt veroorzaakt door toegenomen recharge of door een stijging van het oppervlaktewaterpeil.

Het model kan nog wel bruikbaar zijn voor andere doeleinden. Als er bijvoorbeeld een grondwateronttrekking is gestart in de omgeving kan het effect daarvan mogelijk wel met het model bepaald worden. Voor het bepalen van dat effect is het immers niet van belang om de invloeden van recharge en oppervlaktewater van elkaar te scheiden.

%% PM: invullen pastas model met rivier? voorbeeld lijkt niet bovenstaande uit te beelden...%%

Voorbeeld parameter correlaties 2 %% PM: voorbeeld met gecorreleerde onttrekkingsreeksen %%

1.8 Hydrologische beoordeling

Tijdreeksmodellen worden veel gebruikt om systeem-identificatie uit te voeren, waarbij de bijdrages van verschillende invloeden aan veranderingen in de stijghoogte worden ingeschat. Vanuit vraagstukken uit de praktijk wordt tijdreeksanalyse vaak toegepast op een heel meetnet, of op meerdere peilbuizen binnen een interessegebied. De hydrologische beoordeling van individuele, maar ook hele sets van tijdreeksmodellen is interessant om uit te voeren.

1.8.1 Hydrologie

Het beoordelen van modellen op basis van hydrologische aspecten kan allerlei vormen aannemen. De vraag daarbij is eigenlijk “is hetgeen dat ik bereken hydrologisch plausibel?”. Daarbij zijn modellen die wel aan de verwachtingen voldoen net zo interessant als de modellen die daar niet aan voldoen.

Door de modellen op hydrologische aspecten te toetsen wordt het inzicht in het systeem vergoot.

Voorbeeld hydrologische beoordeling 1 %% PM: echte voorbeelden met winning invoegen %%

Als voorbeeld van een hydrologische beoordeling van een tijdreeksmodel wordt de respons van een onttrekking beschouwd. In onderstaande figuur zijn twee verschillende stapresponsen weergegeven met een onzekerheidsbandbreedte. De bandbreedte van 2x de standaardafwijking is in dit geval berekend aan de hand van de onzekerheid in de gain parameter (de parameter die het stationaire effect van de onttrekking bepaald). In de linker grafiek is de onzekerheid van de staprespons klein, met andere woorden, de parameters kunnen relatief nauwkeurig bepaald worden in de optimalisatie. In de rechter grafiek is deze onzekerheid veel groter, de invloed van de onttrekking kan dus niet nauwkeurig worden ingeschat in de optimalisatie. In dit geval betekent een positieve staprespons dat de onttrekking een verlaging van de stijghoogte veroorzaakt.

Stationaire invloed van een onttrekking met een kleine onzekerheid (links) en een grote onzekerheid (rechts).

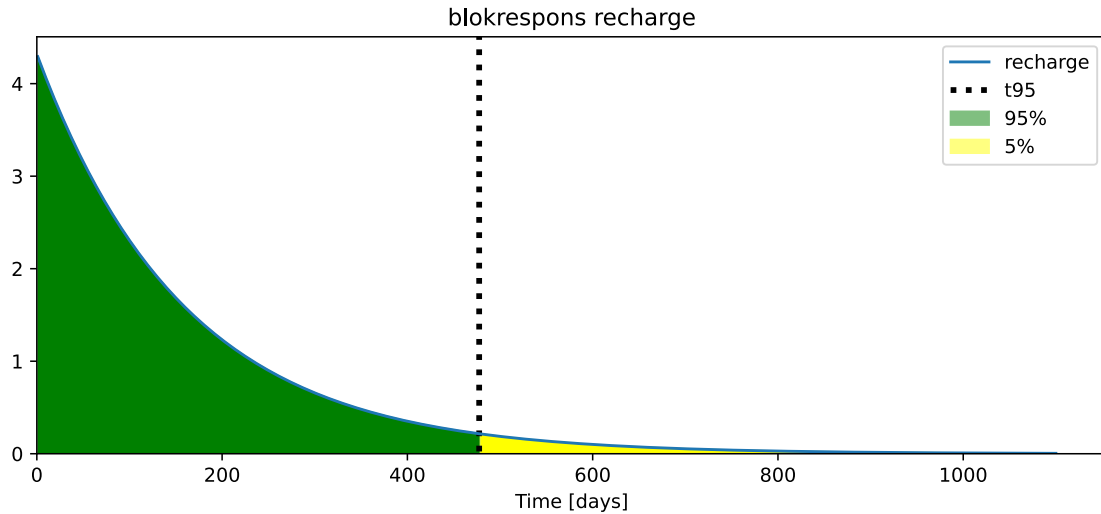
Vanuit de hydrologie moet de onttrekking voor een daling van de grondwaterstand zorgen, maar in het geval van het onzekere model (rechts) is er een meer dan 5%-kans dat de onttrekking een verhoging van de grondwaterstand veroorzaakt. De bandbreedte is dusdanig groot dat er een kans is dat de staprespons ook negatief kan zijn. In dat geval kunnen we stellen dat dat hydrologisch gezien niet plausibel is, en dat dit model waarschijnlijk niet goed is. De onttrekking heeft mogelijk geen significante invloed op de stijghoogte.

1.8.2 Geheugen van de responsfunctie

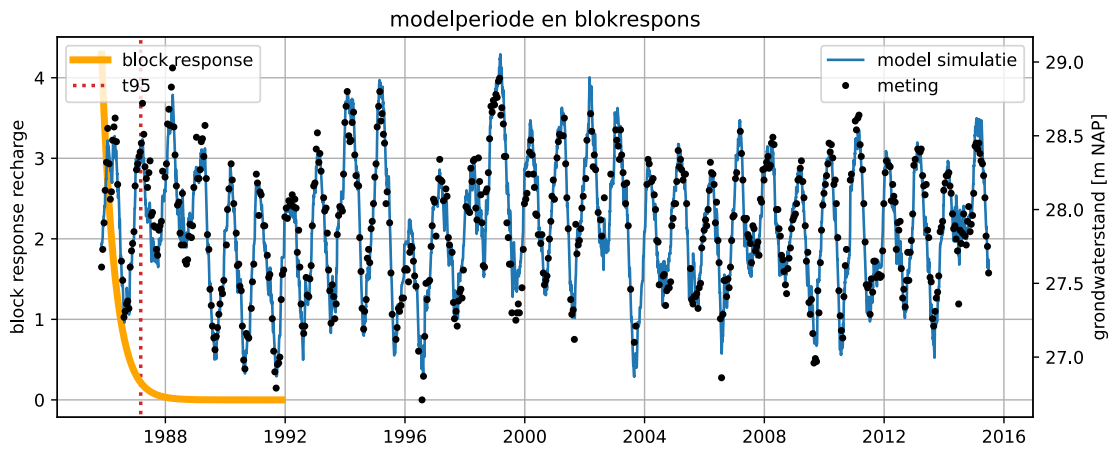
%% PM %%

Voorbeeld bepalen geheugen van de responsfunctie Het bepalen van het geheugen van de responsfunctie is niet altijd even vanzelfsprekend. Regelmatig worden exponentiële responsfuncties gebruikt met een horizontale asymptoot op $y=0$, de respons zal dus nooit gelijk zijn aan 0. Het is daarom niet mogelijk om het geheugen te bepalen door te kijken wanneer de respons 0 is. Om deze reden wordt als maat voor het geheugen vaak de tijd genomen die het duurt voordat 95% van de respons is uitgewerkt. Hieronder is dit visueel weergegeven.

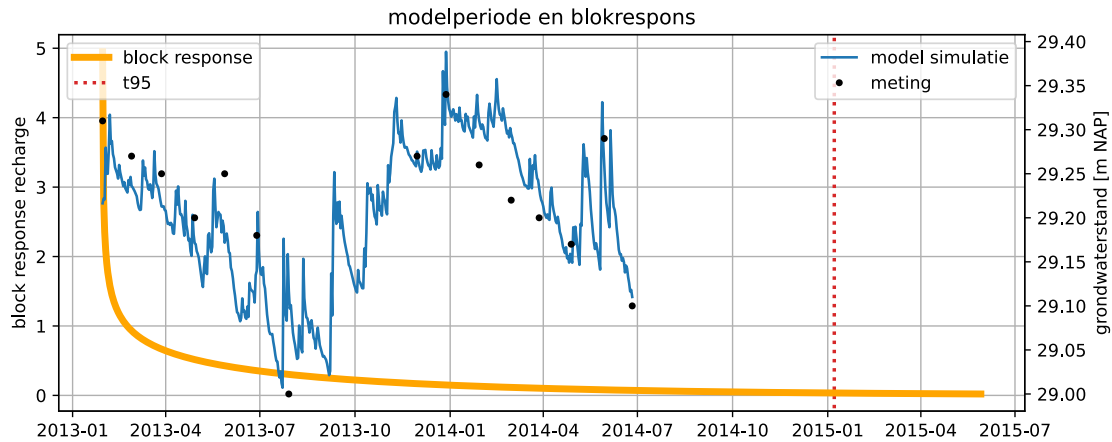
Geheugen (t_{95}) = 477 dagen



Voorbeeld geheugen 1 Hieronder zijn van het voorbeeldmodel de blokrespons en de modelresultaten in één figuur geplot. Bij dit model is de lengte van het tijdreeksmodel een aantal maal groter dan het geheugen van de responsfunctie.



Voorbeeld geheugen 2 Voor een ander voorbeeldmodel is ook de blokrespons en de modelresultaten in één figuur geplot. Bij dit model is de lengte van het tijdreeksmodel kleiner dan het geheugen van de responsfunctie.



1.8.3 Ruimtelijke beoordeling

In veel studies wordt tijdreeksanalyse toegepast op meerdere stijghoogtemeetpunten. Deze studies hebben bijvoorbeeld als doel om uit te zoeken waar de stijghoogtes goed verklaard kunnen worden aan de hand van een grondwateraanvulling, of willen graag uitzoeken wat de invloed is van een drinkwaterwinning op de grondwaterstanden, en zo zijn er tal andere voorbeelden te bedenken. In dit soort studies is het nuttig om een ruimtelijke beoordeling te doen van de resultaten. Het is onmogelijk om alle mogelijke soorten ruimtelijke beoordelingen te noemen dus wordt dit aspect aan de hand van het volgende voorbeeld toegelicht.

Voorbeeld hydrologische ruimtelijke beoordeling Een voorbeeld van ruimtelijke beoordeling van tijdreeksmodellen wordt gegeven aan de hand van een voorbeeld uit Section 1.10. In die studie zijn tijdreeksmodellen afgeleid voor peilbuizen binnen een straal van ± 5 km rondom een drinkwater winning van Brabant Water bij Seppe. In deze tijdreeksmodellen zijn neerslag, verdamping en het onttrekkingsdebiet van winning Seppe als verklarende variabelen meegenomen. Vervolgens zijn de tijdreeksmodellen geoptimaliseerd. In onderstaande figuur zijn de berekende stationaire invloeden (ook wel bekend als de ‘gain’) van de winning Seppe uitgezet tegen de afstand tussen de peilbuizen en het de winlocatie. Daarbij is onderscheid gemaakt tussen peilbuizen in het eerste en tweede watervoerende pakket. De verticale balken geven de onzekerheid aan, en komen overeen met 2 maal de standaardafwijking van de parameter. De stationaire invloed is berekend voor een onttrekkingsdebiet van 1 miljoen m^3/jaar . In de figuur is ook de stationaire invloed van de onttrekking als functie van de afstand opgenomen. Deze is berekend met een 2-laags analytisch elementen grondwatermodel Section 1.10 en dient als extra controle op de geschatte invloed volgens de tijdreeksmodellen.

Berekende stationaire invloed winning Seppe als functie van de afstand bij een onttrekkingsdebiet van 1 Mm^3/jaar (bron: Collenteur et al. (2019))

Elk tijdreeksmodel is onafhankelijk ten opzichte van de andere modellen; er wordt per model op basis van de gemeten stijghoogtes ingeschat wat de bijdrages van neerslag, verdamping en onttrekking zijn op de verandering in de grondwaterstand. Het berekende patroon van een grotere verlaging nabij de winning past goed op de theoretische invloed van een onttrekking. Ook het verschil tussen het eerste en tweede watervoerende pakket is duidelijk op te maken uit de resultaten van de

tijdreeksmodellen. De resultaten komen ook aardig overeen met een berekening met een eenvoudig grondwatermodel.

Deze onafhankelijke modellen lijken goed in staat om de invloed van de drinkwaterwinning op de stijghoogtes in te schatten, dus op basis van deze ruimtelijke beoordeling wordt geconcludeerd dat de tijdreeksmodellen toegepast kunnen worden verdere analyses te doen over de invloed van de winning op de stijghoogtes. Een ruimtelijke hydrologische beschouwing van de resultaten van meerdere tijdreeksmodellen kan dus waardevol zijn in de beoordeling van die tijdreeksmodellen.

1.9 Discussie

%% PM: wat komt hier? %%

1.10 Referenties

%% PM: invullen referenties %%

- Asmuth 2002
- Collenteur, R.A., M. Bakker, R. Caljé, S.A. Klop, F. Schaars. 2019. Pastas: Open Source Software for the Analysis of Groundwater Time Series. Groundwater, Vol. 57, No. 6, 877–885
- Bakker, M., and O.D.L. Strack. 2003. Analytic Elements for Multiaquifer Flow. Journal of Hydrology, 271(1-4), 119-129.
- Jackson et al. (2019) Introductory overview: Error metrics for hydrological modelling - A review of common practices and an open source library to facilitate use and adoption. Environmental Modelling & Software 119, 32-48.
- van Geer, F. (2012) Tijdreeksanalyse: Introductie en aandachtspunten. Stromingen 18, nummer 2.
- Knotters, M. (2012) Validatie van tijdreeksmodellen voor de grondwaterstand. Stromingen 18, nummer 2.
- Poeter, E. & Anderson, D. (2005) Multimodel Ranking and Inference in Ground Water Modeling. Ground Water 43, 597-605