

stowa

HANDLEIDING TIJDREEKSANALYSE



RAPPORT

2021
32

HANDLEIDING TIJDREEKSANALYSE

RAPPORT

2021

32

ISBN 978.90.5773.939.2



COLOFON

UITGAVE Stichting Toegepast Onderzoek Waterbeheer
Postbus 2180
3800 CD Amersfoort

AUTEURS J. Von Asmuth (Trefoil Hydrology): Hoofdstuk 2 en 13
P. Baggelaar (PB-Icastat): Hoofdstuk 3, 4, 6, 7, 8, 11 en 12
M. Bakker (TU Delft): Hoofdstuk 9
D. Brakenhoff (Artesia): Hoofdstuk 9
R. Collenteur (Universität Graz): Hoofdstuk 9
O. Ebbens (Artesia): Hoofdstuk 9
H. Mondeel (Witteveen+Bos): Hoofdstuk 1, 2 en algehele redactie
S. Klop (Witteveen+Bos): Hoofdstuk 1, 5 en 6, 10
F. Schaars (Artesia): Hoofdstuk 9

BEGELEIDINGSCOMMISSIE

Jos Beemster (Waternet)
Robert Broekhuis (Waterschap Vechtstromen)
Francis de Graaf (Waterschap Drents Overijsselse Delta)
Jan Hoogendoorn (Vitens)
Ewa van Kooten (Waterschap Vallei en Veluwe)
Thomas de Meij (Staatsbosbeheer)
Jacob Oosterwijk (Brabant Water)
Pim van Santen (Waterschap Aa en Maas)
Michelle Talsma (STOWA)
Elena Uibel (Waterschap Hunze en Aa)

DANKWOORD

In de handleiding is gebruik gemaakt van materiaal en illustraties uit het Engelstalige Handboek Tijdreeksanalyse (in prep.) dat in opdracht van de drinkwaterbedrijven Vitens en Brabant Water wordt opgesteld door KWR / Jos von Asmuth en Kees Maas

DRUK Kruyt Grafisch Adviesbureau bv
STOWA STOWA 2021-32
ISBN 978.90.5773.939.2

Copyright Teksten en figuren uit dit rapport mogen alleen worden overgenomen met bronvermelding.
Disclaimer Deze uitgave is met de grootst mogelijke zorg samengesteld. Niettemin aanvaarden de auteurs en de uitgever geen enkele aansprakelijkheid voor mogelijke onjuistheden of eventuele gevolgen door toepassing van de inhoud van dit rapport.

TEN GELEIDE

HANDLEIDING VOOR HET UITVOEREN VAN TIJDREEKSANALYSES

Hoe bepaal je het effect van een drinkwaterwinning op de grondwaterstand? Wat voor effect heeft een peilaanpassing op de omgeving? Waterbeheerders en andere belanghebbers willen graag weten wat het effect van dergelijke ingrepen in het watersysteem is. Om daar achter te komen wordt vaak gebruik gemaakt van meetgegevens in de omgeving en wordt hiermee een zogenoemde tijdreeksanalyse uitgevoerd. Een praktische handleiding voor het maken van een tijdreeksanalyse ontbrak. STOWA heeft op verzoek van de waterbeheerders het initiatief genomen om deze handleiding op te laten opstellen. De handleiding helpt de gebruikers stap voor stap om een tijdreeksanalyse op een verantwoorde wijze uit te voeren.

Tijdreeksanalyse is een techniek die het verband legt tussen een ingreep in het hydrologische systeem zoals het plaatsen van een stuw en de uitbreiding van een drinkwaterwinning en de reactie hierop van het (grond)watersysteem. Het kan gaan om een daling of stijging van grondwaterstand of een afname van de beekafvoer.

Er zijn twee methoden beschikbaar voor het maken van een tijdreeksanalyse. Het gaat om de 'statistische' methode (Box-Jenkins Transferruis, BJTR) en de 'fysische' methode (Predefined Impulse Response Function In Continuous Time: PIRFICT). De fysische methode wordt ook wel de 'impuls-respons' methode genoemd. De gepresenteerde handleiding beschrijft beide methoden.

De handleiding geeft inzicht in de te doorlopen stappen, randvoorwaarden en mogelijkheden van de tijdreeksmodellering. In de handleiding zijn casussen opgenomen waarmee u zelf aan de slag kunt. Links naar websites met de benodigde software zijn in het handboek opgenomen. Het handboek is opgesteld voor beginnende tijdreeksanalisten en geeft ervaren analisten verdiepende kennis en informatie.

Dit handboek is mede mogelijk gemaakt door een financiële bijdrage van diverse waterschappen.

Joost Buntsma
Directeur STOWA

DE STOWA IN HET KORT

STOWA is het kenniscentrum van de regionale waterbeheerders (veelal de waterschappen) in Nederland. STOWA ontwikkelt, vergaart, verspreidt en implementeert toegepaste kennis die de waterbeheerders nodig hebben om de opgaven waar zij in hun werk voor staan, goed uit te voeren. Deze kennis kan liggen op toegepast technisch, natuurwetenschappelijk, bestuurlijk-juridisch of sociaalwetenschappelijk gebied.

STOWA werkt in hoge mate vraaggestuurd. We inventariseren nauwgezet welke kennisvragen waterschappen hebben en zetten die vragen uit bij de juiste kennisleveranciers. Het initiatief daarvoor ligt veelal bij de kennisvragende waterbeheerders, maar soms ook bij kennisinstellingen en het bedrijfsleven. Dit tweerichtingsverkeer stimuleert vernieuwing en innovatie.

Vraaggestuurd werken betekent ook dat we zelf voortdurend op zoek zijn naar de 'kennisvragen van morgen' – de vragen die we graag op de agenda zetten nog voordat iemand ze gesteld heeft – om optimaal voorbereid te zijn op de toekomst.

STOWA ontzorgt de waterbeheerders. Wij nemen de aanbesteding en begeleiding van de gezamenlijke kennisprojecten op ons. Wij zorgen ervoor dat waterbeheerders verbonden blijven met deze projecten en er ook 'eigenaar' van zijn. Dit om te waarborgen dat de juiste kennisvragen worden beantwoord. De projecten worden begeleid door commissies waar regionale waterbeheerders zelf deel van uitmaken. De grote onderzoeklijnen worden per werkveld uitgezet en verantwoord door speciale programmacommissies. Ook hierin hebben de regionale waterbeheerders zitting.

STOWA verbindt niet alleen kennisvragers en kennisleveranciers, maar ook de regionale waterbeheerders onderling. Door de samenwerking van de waterbeheerders binnen STOWA zijn zij samen verantwoordelijk voor de programmering, zetten zij gezamenlijk de koers uit, worden meerdere waterschappen bij één en het zelfde onderzoek betrokken en komen de resultaten sneller ten goede aan alle waterschappen.

De grondbeginselen van STOWA zijn verwoord in onze missie:

Het samen met regionale waterbeheerders definiëren van hun kennisbehoeften op het gebied van het waterbeheer en het voor én met deze beheerders (laten) ontwikkelen, bijeenbrengen, beschikbaar maken, delen, verankeren en implementeren van de benodigde kennis.

INHOUDSOPGAVE

1	INLEIDING	5
1.1	Inleiding	5
1.2	Doel en doelgroep	5
1.3	Vershil in terminologie tussen beide methoden	5
1.4	Structuur en leeswijzer van deze Handleiding	6
	1.4.1 Leeswijzer	6
	1.4.2 Structuur	6
2	TIJDREEKSMETHODEN EN -MODELLEN	8
2.1	Tijdreeksmodellen, de statistische en fysische benadering	8
	2.1.1 Inleiding in de statistische en fysische methoden	8
	2.1.2 Overeenkomsten tussen beide methoden	9
2.2	Over regressie en ARMA-functies	10
2.3	Het transfer- en ruismodel	11
2.4	Verschillende transfermodellen	13
	2.4.1 Gemeenschappelijke basis	13
	2.4.2 Box-Jenkins transferruismodel (BJTR)	15
	2.4.3 Het PIRFICT-model	18
2.5	Verschillende ruismodellen	20
	2.5.1 Het ruismodel	20
	2.5.2 Verschillen in ruismodellen en terminologie	20
2.6	TRA met simultane invloeden en reeksen	20
	2.6.1 Inleiding	20
	2.6.2 Simultane invloeden - scheiden van effecten	20
	2.6.3 Simultane analyse en identificatie	24
2.7	Vergelijking PIRFICT- en BJTR-model (overzicht)	26
	2.7.1 Wiskundige verschillen	26
	2.7.2 Vergelijking van resultaten	27
	2.7.3 Voor- en nadelen van de BJTR en PIRFICT methode	28
2.8	Software	29
	2.8.1 Menyanthes/Hydromonitor	29
	2.8.2 Tijdreeksanalist	30
	2.8.3 Pastas	30

3	VOORBEREIDEN TIJDREEKSANALYSE	31
3.1	Inleiding	31
3.2	Specificeren soort gewenste resultaat	32
3.3	Verkennen systeem en data	33
	3.3.1 Verkennen systeem	33
	3.3.2 Verkennen en beoordelen data	34
3.4	Opstellen modelleerstrategie	36
	3.4.1 Keuze soort tijdreeksmodel	37
	3.4.2 Keuze tijdseenheid	38
	3.4.3 Keuze modelperiode	39
	3.4.4 Hoe meenemen neerslag en verdamping als invoervariabelen	40
	3.4.5 Omgaan met verandering van het systeem	41
	3.4.6 Omgaan met sterk gecorreleerde invoerreeksen	41
	3.4.7 Omgaan met onverklaarbaar patroon van de uitvoerreeks	42
3.5	Vorbewerken meetreeksen	42
	3.5.1 Vorbewerken bij een BJTR-model	43
	3.5.2 Vorbewerken bij een PIRFICT-model	44
4	ONTWIKKELEN BJTR-MODEL	45
4.1	Inleiding	45
4.2	Identificeren modelvorm	46
	4.2.1 Empirisch identificeren vorm transferfunctie	46
	4.2.2 Theoretische aanpak bij identificeren vorm transferfunctie	53
	4.2.3 Instellen vorm ruismodel	53
4.3	Schatten modelparameters BJTR-methode	53
4.4	Verifiëren model en eventueel aanpassen model en/of strategie	54
	4.4.1 Beoordelen verloop en resultaten schattingsproces	55
	4.4.2 Beoordelen kenmerken modelresidu a_t	58
	4.4.3 Beoordelen kruiscorrelaties modelresidu/modelruis en invoerreeksen	68
	4.4.4 Valideren model met onafhankelijke meetwaarden	69
5	ONTWIKKELEN PIRFICT-MODEL	72
5.1	Inleiding	72
5.2	Identificeren modelvorm	73
	5.2.1 Welke invloeden zijn van belang	73
	5.2.2 Hoe het effect van neerslag en verdamping kan worden meegenomen	73
	5.2.3 keuze van de responsfunctie	73
	5.2.4 Hoe om te gaan met veranderende hydrologische systemen	75
	5.2.5 Toepassen ruismodel	75
5.3	Schatten modelparameters	76
5.4	Verifiëren model	77

5.4.1	Beoordelen ruis	78
5.4.2	Beoordelen op kruiscorrelaties ruis en invoerreeksen	81
6	TOEPASSEN MODEL EN PRESENTEREN RESULTATEN	82
6.1	Inleiding	82
6.2	Toepassen model	82
6.2.1	Houd rekening met kalibratiegebied van tijdreeksmodel	83
6.2.2	Voorspellen met een BJTR-model	84
6.2.3	Voorspellen met een PIRFICT-model	85
6.3	Presenteren resultaten	86
6.3.1	Interpretatie en patroonanalyse	86
6.3.2	De valkuil van te vergaande onzekerheidsreductie	88
7	CASUS BJTR-METHODE: INVLOED GRONDWATERWINNING NIJ BEETS	91
7.1	Specificeren soort gewenste resultaat	91
7.2	Verkennen systeem en data	91
7.2.1	Topografische, geografische en hydrogeologische kenmerken van de omgeving	91
7.2.2	Verkennen data mogelijke invloedsfactoren grondwaterstand	93
7.3	Opstellen modelleerstrategie	95
7.3.1	Keuze soort tijdreeksmodel	95
7.3.2	Keuze tijdseenheid	96
7.3.3	Keuze modelperiode	96
7.4	Vorbewerken meetreeksen	96
7.4.1	Opschonen van fouten	96
7.4.2	Kalibreren loggermetingen	97
7.4.3	Omzetten meetreeksen naar maandbasis	97
7.4.4	Selectie te modelleren grondwaterstandreeksen	98
7.4.5	Iteratieve modelontwikkeling	98
7.5	Identificeren vorm tijdreeksmodel	98
7.5.1	Modelleren dynamische relatie met neerslagoverschot	98
7.5.2	Modelleren dynamische relatie met grondwaterwinning Nij Beets	99
7.5.3	Modelleren dynamische relaties met oppervlaktewaterstanden	100
7.5.4	Formuleren ruismodel	100
7.6	Schatten modelparameters	100
7.7	Verifiëren en zonodig bijstellen tijdreeksmodel	101
7.8	Presenteren resultaten tijdreeksanalyse	102
8	VOORBEELDEN TIJDREEKSANALYSE MET BJTR-METHODE	108
8.1	Modelleren grondwaterstand met PNO als invoer	109

8.2	Deterministisch voorspellen grondwaterstand	116
8.3	Stochastisch voorspellen grondwaterstand	119
9	VOORBEELDEN EN ACHTERGROND PIRFICT-METHODE	122
9.1	Vorbewerking van reeksen voor tijdreeksanalyse	122
9.1.1	Opschonen stijghoogte reeks	122
9.1.2	Vereenvoudigen stijghoogte reeks	124
9.1.3	Equidistant maken verklarende reeks	127
9.1.4	Voorgeschiedenis verklarende reeks	127
9.1.5	Eenheden afstemmen	129
9.2	Model structuur	130
9.2.1	Algemeen stappenplan	130
9.2.2	Basis modelstructuur	131
9.2.3	Responsfuncties	132
9.2.4	Effect van neerslag en verdamping	132
9.2.5	Dikkere onverzadigde zone	137
9.2.6	Drempel niet-lineariteit	138
9.2.7	Andere verklarende variabelen	139
9.2.8	Veranderingen in het systeem	144
9.2.9	Discussie	146
9.3	Kalibratie van een tijdreeksmodel	147
9.3.1	Introductie	147
9.3.2	Effect van de lengte van de tijdreeks (of kalibratieperiode)	148
9.3.3	Het ruismodel	150
9.3.4	Effect van een ruismodel	151
9.3.5	Effect van het ruismodel op de schatting van de waarden en betrouwbaarheid van de parameters	154
9.3.6	Effect van een fout in de gemeten grondwateraanvulling	156
9.3.7	Effect van de meetfrequentie op de werking van het AR1 ruismodel	157
9.4	Beoordeling Tijdreeksmodellen	160
9.4.1	Introductie	160
9.4.2	Algemeen stappenplan	160
9.4.3	Visuele beoordeling	161
9.4.4	Fit statistieken	162
9.4.5	Testen model	165
9.4.6	Modelonzekerheid	167
9.4.7	Hydrologische beoordeling	171
10	CASUSSEN PIRFICT-METHODE	179
10.1	Casus: Invloeden bepalen	179
10.1.1	Deel I: Invloed peilopzet	179
10.1.2	Deel II: Invloed waterloop	185
10.2	Casus: Karakteristiek bepalen	190
10.2.1	Deel I: Inschatten GXG voor korte reeks	190
10.2.2	Deel II: Invullen missend deel van meetreeks	195

10.3	Casus: Systeemanalyse	197
	10.3.1 Verlaging natuurgebied	197
10.4	Casus: Voorspellen	202
	10.4.1 Deel I: Uitzetten onttrekking	202
	10.4.2 Deel II: Herstel grondwaterstand in april	212
11	INTRODUCTIE STATISTIEK VOOR TIJDREEKSANALYSE	218
11.1	Inleiding	218
11.2	Onzekerheid hoort bij modelleren	218
11.3	Het begrip autocorrelatie	222
11.4	Statistisch toetsen	223
11.5	Schade bij tijdreeksanalyse als modelresiduen geen witte ruis vormen	227
12	INTRODUCTIE BOX-JENKINS TIJDREEKSMODELLEN	231
12.1	Uitleg aan de hand van systeemtheorie	231
12.2	Statistisch modelleren van een tijdreeks	232
12.3	Ontwikkelen van een BJTR-model	233
12.4	Formuleren van een BJTR-model	234
13	RUIS, RESIDUEN EN INNOVATIES: WAT IS WAT?	237
13.1	Beknopte toelichting op de achterliggende concepten van ruis, residu en innovaties	237
13.2	Gebruik binnen de tijdreeksanalyse	238
13.3	Gebruik buiten de tijdreeksanalyse	239
	LITERATUUR	241

BEGRIPPENLIJST

Autocorrelatie – Er is sprake van autocorrelatie als opeenvolgende waarden van een tijdreeks niet onafhankelijk van elkaar zijn, maar aan elkaar gerelateerd zijn. Te onderscheiden zijn:

- **positieve** autocorrelatie: opeenvolgende reekswaarden verschillen doorgaans minder van elkaar dan reekswaarden met een groter tijdsverschil. Als een natuurlijk proces met een hoge frequentie wordt waargenomen treedt doorgaans positieve autocorrelatie op;
- **negatieve** autocorrelatie: opeenvolgende reekswaarden verschillen doorgaans meer van elkaar dan reekswaarden met een groter tijdsverschil.

De gepresenteerde onzekerheden van geschatte modelparameters en voorspellingen zullen onjuist zijn als bij de modellering van een tijdreeks geen rekening met autocorrelatie wordt gehouden. Zo zullen bij niet of onvoldoende verdisconteren van positieve autocorrelatie zowel de standaardfout als het betrouwbaarheidsinterval van een geschatte evenwichtsrelatie worden onderschat. Bij hoge positieve autocorrelatie kunnen dat zelfs forse onderschattingen zijn.

Autocorrelatiecoëfficiënt - Een maat voor de lineaire samenhang van reekswaarden die zijn gescheiden door een bepaald aantal tijdseenheden.

Autocorrelatiefunctie (ACF, of autocorrelogram) - Een functie die de autocorrelatiecoëfficiënt koppelt aan het aantal tijdseenheden tussen de waarden. De grafische weergave van de ACF wordt ook wel aangeduid als *autocorrelogram*.

ARMA – Een modelformulering met zowel autoregressieve termen (AR-termen), als moving average-termen (MA-termen). In de BJTR-methode is de transferfunctie voor de relatie tussen invoerreeks en uitvoerreeks geformuleerd als een ARMA-model, evenals het ruismodel (N_t). Als het ruismodel ook een differentie omvat wordt het aangeduid als een ARIMA-model en als het tevens seizoensparameters omvat wordt het aangeduid als een SARIMA-model.

Betrouwbaarheid bij statistisch toetsen – De kans op het terecht niet verwerpen van de nulhypothese (H_0). Deze kans wordt ook wel aangeduid als $1 - \alpha$, waarbij α het vóóraf ingestelde significantieniveau (of ook wel fout van de 1^e soort) is, zijnde het geaccepteerde risico op het onterecht verwerpen van de nulhypothese. Te onderscheiden zijn:

- **nominale** betrouwbaarheid – deze is vóóraf ingesteld door de analist (meestal op 95 %);
- **empirische** betrouwbaarheid – deze wordt in de praktijk gerealiseerd en kan afwijken van de nominale betrouwbaarheid als niet wordt voldaan aan bepaalde randvoorwaarden (zoals het ontbreken van autocorrelatie).

Betrouwbaarheidsinterval van een schatting – Een maat voor de precisie van een schatting. Als wordt voldaan aan de daarvoor geldende randvoorwaarden, zal een 95 %-betrouwbaarheidsinterval van een schatting met 95 % betrouwbaarheid de werkelijke waarde bevatten. Engels: *confidence interval of an estimate*.

BJTR-methode - De Box-Jenkins-methode voor het uitvoeren van tijdreeksanalyse.

Deterministische deel van een tijdreeksmodel: Het modeldeel dat de relaties met de beschikbare invoerreeksen (zoals neerslag, verdamping en grondwaterwinning) beschrijft.

Evenwichtsrelatie (ook wel aangeduid als A, of als M0) - De uiteindelijke stationaire verandering van de uitvoerreeks van een systeem als de invoerreeks een constante niveauperandering van één eenheid heeft. Engels: *steady-state gain*.

GHG of GLG - Gemiddeld hoogste, of laagste grondwaterstand. De GHG (GLG) is het over een bepaalde periode bepaalde gemiddelde van alle jaargemiddelde hoogste (laagste) grondwaterstanden, de laatste bepaald over de drie hoogste (laagste) grondwaterstanden in dat jaar, uitgaande van halfmaandelijke metingen (op de 14^e en 28^e van elke maand) van april tot april. Vroeger werd daarvoor meestal een periode van minstens acht jaar gehanteerd, maar inmiddels wordt ook wel geadviseerd daarvoor een rekenperiode van 30 jaar te hanteren, om de invloed van extreme jaren te verminderen.

Homoscedasticiteit - Een reeks vertoont homoscedasticiteit als de variantie van de reekswaarden niet verandert met de tijd en/of het niveau. Engels: *homoscedasticity*.

Impulsresponsfunctie of transferfunctie - Functie die de relatie beschrijft tussen een enkelvoudige impuls en de respons van het systeem.

Informatie criterium - Een maat waarmee de kwaliteit van verschillende modellen voor dezelfde dataset (in ons geval tijdreeks) kan worden vergeleken. Daarbij wordt zowel de modelpasvorm meegewogen (de mate waarin het model aansluit op de reeks), als de complexiteit van het model (uitgedrukt in het aantal modelparameters). Een dergelijke maat kan objectieve onderbouwing bieden bij modelselectie. Bekende informatiecriteria zijn het Akaike-InformatieCriterium (AIC) en het Bayes-InformatieCriterium (BIC).

Kruiscorrelatiecoëfficiënt - Een maat voor de lineaire samenhang van reekswaarden van twee variabelen die zijn gescheiden door een bepaald aantal tijdseenheden.

Kruiscorrelatiefunctie (KCF of kruiscorrelogram) - Een functie die de kruiscorrelatiecoëfficiënt koppelt aan het aantal tijdseenheden tussen de waarden. De grafische weergave van de KCF wordt ook wel aangeduid als *kruiscorrelogram*.

p-waarde (of significantie) - Bij statistisch toetsen is dit de kans op een extremere toetsingsgrootte dan degene die is gerealiseerd, bij geldigheid van de nulhypothese. Als de p-waarde kleiner is dan de vooraf ingestelde fout van de 1^e soort α (ook wel aangeduid als significantieniveau, zoals 5% of 1%), kan de nulhypothese worden verworpen.

Partiële autocorrelatiecoëfficiënt - Een maat voor de lineaire samenhang van reekswaarden die zijn gescheiden door een bepaald aantal tijdseenheden, ná correctie voor de doorwerking van de lineaire samenhang van reekswaarden die zijn gescheiden door minder tijdseenheden.

Partiële autocorrelatiefunctie (PACF, of partieel autocorrelogram) - Een functie die de partiële autocorrelatiecoëfficiënt koppelt aan het aantal tijdseenheden tussen de waarden. De grafische weergave van de PACF wordt ook wel aangeduid als *partieel autocorrelogram*. Een ACF en PACF samen geven een completer beeld van de achterliggende correlatiestructuur dan alleen een ACF en stellen zo in staat het ruismodel (N_t) beter te identificeren.

PIRFICT-methode - Fysisch-hydrologisch georiënteerde methode voor het uitvoeren van tijdreeksanalyse. Het acroniem staat voor *Predefined Impulse Response Function In Continuous Time*.

Ruis en residuen - Benamingen van twee stochastische reeksen die tevoorschijn komen bij het ontwikkelen van een tijdreeksmodel van een tijdreeks (daarom worden ze ook wel aangeduid als latente variabelen). Dit betreft:

- N_t : het stochastische deel van de uitvoerreeks, dat niet kan worden beschreven door het deterministische deel van het tijdreeksmodel. In de BJTR-methode wordt deze aangeduid als de *ruisreeks* en in de PIRFICT-methode als de *residuereeks*. De component van residuereeks die niet tot de systeemfout behoort wordt in de PIRFICT methode als *ruis* aangeduid;
- a_t : een stochastische reeks die de invoer vormt van het ruismodel dat de N_t -reeks beschrijft. In de BJTR-methode wordt deze aangeduid als de *residuereeks* en in de PIRFICT-methode als de *ruisreeks*.

Responstijd (van twee variabelen) – De tijd die het duurt voordat de invloed van een verandering van de eerste variabele (zoals het neerslagoverschot) op de tweede variabele (zoals de grondwaterstand) is uitgewerkt.

Schatter - Een methode om de waarde van een bepaalde eigenschap te schatten. Maar wordt ook gebruikt om de kansverdeling van de schattingen aan te duiden. Engels: *estimator*.

Schattingsfout – Verschil tussen geschatte en werkelijke waarde van een bepaalde eigenschap.

Significantieniveau – Het vóóraf ingestelde – en daarmee nog geaccepteerde - risico op het onterecht verwerpen van de nulhypothese. Deze wordt aangeduid als α , of ook wel als het risico van de 1^e soort. We onderscheiden:

- **nominaal** significantieniveau - deze is vóóraf ingesteld (meestal op 5 %) en geeft aan welk risico wordt geaccepteerd;
- **empirisch** significantieniveau - deze wordt in de praktijk gerealiseerd en kan afwijken van het nominale significantieniveau als niet wordt voldaan aan bepaalde randvoorwaarden.

Standaardfout van een schatting – Standaardafwijking van de schattingsfout en daarmee ook van de schatter. Op te vatten als een maat voor de onzekerheid van een schatting. Engels: *standard error of an estimation*.

Systematische fout – Een afwijking tussen de verwachtingswaarde van een schatter en de werkelijke waarde van het schattingsdoel, oftewel $E[\hat{\theta}] \neq \theta$, waarin $E[.]$ de verwachtingswaarde van de schatter, $\hat{\theta}$ de schatting en θ de werkelijke waarde. Engels: *bias*.

Tijdreeks - Een chronologisch gerangschikte reeks waarden van een bepaalde variabele, zoals de grondwaterstand, of de neerslag op een bepaalde locatie. Engels: *time series*.

Verklaarde variantie – Een maat voor de pasvorm van het tijdreeksmodel. Het betreft het deel van de variantie van de uitvoerreeks dat verklaard wordt door het deterministische deel van het tijdreeksmodel, met waarden tussen 0 en 100 %. Wordt ook wel aangeduid als R^2 , of als EVP (*explained variance percentage*).

Zuivere schatter - Een schatter zonder systematische fout. De verwachtingswaarde van de schatter is dan gelijk aan de werkelijke waarde van het schattingsdoel, oftewel $E[\hat{\theta}] = \theta$, waarin $E[.]$ de verwachtingswaarde van de schatter, $\hat{\theta}$ de schatting en θ de werkelijke waarde. Engels: *unbiased estimator*.

Deel I. Handleiding

1

INLEIDING

1.1 Inleiding

Tijdreeksanalyse is een techniek om relaties tussen verschillende systeem-variabelen te leggen, zowel binnen als buiten het hydrologisch domein. De tijdreeksanalyse bestaat uit statistische algoritmes die het verband herleiden tussen een verandering of input in het hydrologisch systeem (bijvoorbeeld een peilverandering, grondwateronttrekking, neerslag en verdamping) en de respons of output hierop ((grond)waterstand of afvoer). De tijdreeksanalyse is een nuttig instrument in de hydrologie en wordt steeds meer toegepast.

Bij het toepassen van tijdreeksanalyse is het nodig dat de gebruiker zich bewust is van de (on)mogelijkheden en randvoorwaarden van de verschillende tijdreeksmodellen. Deze handleiding dient om gebruikers van tijdreeksanalyse, op een overzichtelijke manier, inzichten te geven in de verschillende stappen in de tijdreeksmodellering, verduidelijkt met voorbeeld casussen. Daarnaast beschrijft de handleiding ook de theoretische achtergrond van de tijdreeksanalyse.

In de tijdreeksanalyse wordt in de hydrologie onderscheid gemaakt tussen de 'statistische' methode (Box-Jenkins Transferruis, BJTR) en de 'fysische' methode (Predefined Impulse Response Function In Continuous Time' (PIRFICT) transferruismodellen, ook wel impuls-respons genoemd. Deze handleiding bevat beide methoden.

1.2 Doel en doelgroep

Dit document is een praktisch toepasbare handleiding voor de gebruikers van tijdreeksmodellen. Dit geldt voor zowel gebruikers die bekend zijn met tijdreeksanalyse als gebruikers die geen eerdere ervaring hebben met tijdreeksanalyse. Voor de verschillende gebruikers is in de leeswijzer (zie paragraaf 1.4.1) beschreven hoe de handleiding gebruikt kan worden. De handleiding bevat daarnaast ook een deel theorie en achtergrond, welke nodig is om tijdreeksanalyse toe te passen in de praktijk. Kennis en/of affiniteit met statistiek en wiskunde wordt aanbevolen om voor een goed begrip van deze handleiding én het toepassen van de tijdreeksanalyse.

1.3 Verschil in terminologie tussen beide methoden

In de statistische methode (BJTR) en de 'fysische' methode (PIRFICT) worden de termen ruis en residu verschillend gebruikt.

In de BJTR-methode wordt met ruis de stochastische component van de uitvoerreeks bedoeld, zijnde het deel dat niet kan worden beschreven door het deterministische deel van het tijdreeksmodel. En met (model)residu wordt bedoeld een stochastische reeks die de invoer vormt van het ruismodel dat de ruis beschrijft (zie Afbeelding 12.2). Bij de PIRFICT-methode is dit omgekeerd en betreft de 'ruis' de stochastische reeks die de invoer vormt van het ruismodel en 'residu' de stochastische component van de uitvoerreeks.

Hoewel dit uiteraard verwarrend is houden we de in de handleiding de terminologie zoals gehanteerd in de verschillende methoden aan. Dit betekent bijvoorbeeld dat ruis en residu in hoofdstuk 4 een andere betekenis heeft dan in hoofdstuk 5.

Aan het begin van het hoofdstuk wordt aangegeven welke termen gehanteerd worden. Wel hanteren we een consistent gebruik van de symbolen, waarbij a_t de ruis/modelresidu van de invoerreeksen betreft en N_t de ruis/modelresidu van de uitvoerreeksen. De symbolen geven we tussen haakjes achter de termen weer om zodat duidelijk is wat met de term bedoeld wordt.

De termen worden in de handleiding verschillend gebruikt conform de terminologie zoals deze in de methoden toegepast wordt. Hierdoor sluit de handleiding aan bij de literatuur van de betreffende methoden.

1.4 Structuur en leeswijzer van deze Handleiding

1.4.1 Leeswijzer

Deze handleiding beschrijft de werkstappen voor het uitvoeren van tijdreeksanalyse, waarbij ook casussen en voorbeelden zijn opgenomen. Daarnaast is theoretische achtergrondinformatie opgenomen welke nodig is voor het goed uitvoeren van tijdreeksanalyse. In de handleiding wordt vanuit de beschrijving van de werkstappen (hoofdstuk 3 tot en met 6) verwezen naar relevante voorbeelden en theorie.

Voor *beginners* wordt aanbevolen om te starten met de beschrijving van de werkstappen van de twee beschreven methoden (statistische en impuls-respons) en vooral de bijbehorende casussen te lezen en/of zelf uit te voeren. Paragraaf 2.1.1 kan de lezer helpen om inzicht te krijgen in de twee methoden. Het wordt aanbevolen bij het uitwerken van de casussen, de theoretische hoofdstukken te gebruiken als ondersteuning en naslagwerk. Na het lezen en uitvoeren van de werkstappen en casussen zal hoofdstuk 2 en de theoretische achtergrondinformatie eenvoudiger te lezen zijn.

Voor *ervaren gebruikers* van tijdreeksmodellering biedt hoofdstuk 2 de basis theorie van beide methoden en de overeenkomsten tussen de methoden, daarnaast wordt in deze hoofdstukken verwezen naar de relevante achtergrond literatuur op het gebied van tijdreeksanalyse. Ook voor deze gebruikers bevat de beschrijving van de werkstappen, en haar aandachtspunten en valkuilen, nuttige informatie.

Voor alle lezers geldt dat de handleiding bedoeld is als ondersteuning bij het uitvoeren van tijdreeksanalyse. Hierbij hoeft de handleiding niet volgordegelezen worden maar kunnen de hoofdstukken ook individueel worden geraadpleegd indien nodig (voor de inhoud van de individuele hoofdstukken zie paragraaf 1.4.2). Om deze reden worden enkele onderdelen meermaals in de handleiding behandeld en kan sprake zijn van enige herhaling, dit is gedaan om de leesbaarheid van de individuele hoofdstukken en paragrafen te bevorderen.

1.4.2 Structuur

Dit document bevat 3 delen:

- Deel I - Handleiding (hoofdstuk 2 t/m 6),
- Deel II - Casussen en uitgewerkte voorbeelden (hoofdstuk 7 t/m 10),
- Bijlage - Verdiepende theoretische achtergrond van de tijdreeksanalyse (hoofdstuk 11, 12 en 13).

Deze 3 delen beschrijven de werkstappen die uitgevoerd worden bij de tijdreeksmodellering, zoals in afbeelding 1.1 weergegeven:

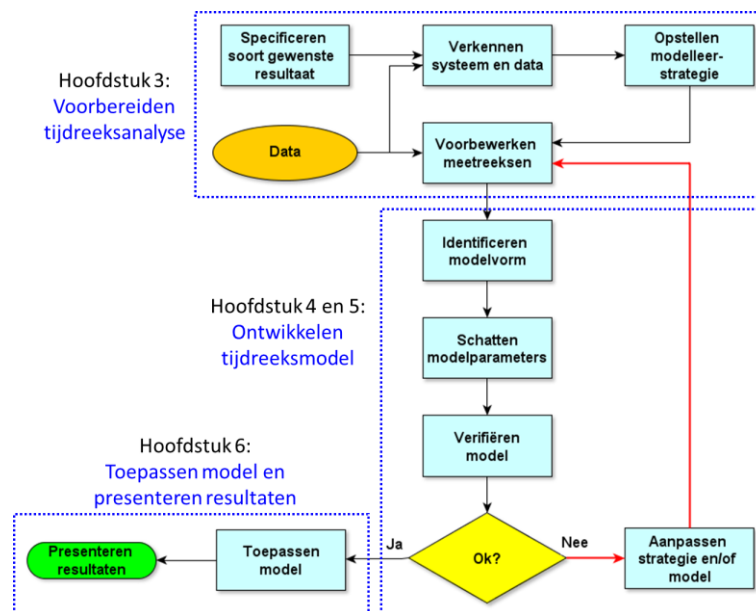
- voorbereiden tijdreeksanalyse;
- ontwikkelen, kalibratie en validatie van het tijdreeksmodel;
- toepassen model en presenteren resultaten.

De handleiding beschrijft de werkstappen, aandachtspunten en valkuilen. Deze werkstappen worden verduidelijkt met casussen en voorbeelden in deel II. In de handleiding wordt verwezen naar paragrafen die dieper ingaan op de theorie (bijlage).

In het document worden de 'statistische' (BJTR) methode en de 'fysische' (PIRFICT) methode naast elkaar behandeld. Hierbij is een indeling gemaakt zodat de lezer voor beide methoden de handleiding kan volgen. De indeling van hoofdstukken voor beide methoden is als volgt:

- algemeen (voor beide methoden):
 - hoofdstuk 2: Korte toelichting op tijdreeksanalyse en het onderscheid in de statistische en fysische methode. Dit hoofdstuk bevat ook een vergelijking van en afwegingskader voor deze twee methoden;
 - paragraaf 3.1 - 3.4: Voorbereiding tijdreeksanalyse;
 - hoofdstuk 6: Toepassen model en presenteren resultaten;
 - hoofdstuk 11: Introductie statistiek voor tijdreeksanalyse;
 - hoofdstuk 13: Ruis, residuen en innovatie (bijlage);
- statistische methode (BJTR):
 - paragraaf 3.5.1: Voorbewerken meetreeksen bij statistische methode (deel I);
 - hoofdstuk 4: Ontwikkelen BJTR -model (deel I);
 - hoofdstuk 7 en 8: Casussen en voorbeelden BJTR -methode (deel II);
 - hoofdstuk 12: Introductie BJTR modellen (bijlage);
- fysische methode (PIRFICT):
 - paragraaf 3.5.2: Voorbewerken meetreeksen bij fysische methode (deel I);
 - hoofdstuk 5: Ontwikkelen PIRFICT-model (deel I);
 - hoofdstuk 9 en 10: Voorbeelden en casussen PIRFICT-methode (deel II).

Afbeelding 1.1 Stroomschema voor het uitvoeren van tijdreeksanalyse, met verwijzing naar de vier hoofdstukken die de fasen van de uitvoering beschrijven. Bij meerdere onderdelen kan blijken dat verantwoorde tijdreeksanalyse niet mogelijk is, maar die uitgangen van het stroomschema zijn niet apart weergegeven, om de overzichtelijkheid te bewaren



2

TIJDREEKSMETHODEN EN -MODELLEN

2.1 Tijdreeksmodellen, de statistische en fysische benadering

2.1.1 Inleiding in de statistische en fysische methoden

Tijdreeksanalyse is een techniek om relaties tussen verschillende variabelen te leggen. De theorie en methoden voor het analyseren en modelleren van tijdreeksen hebben een veel groter toepassingsgebied dan de hydrologie alleen: tijdreeksen komen aan de orde binnen een breed scala aan wetenschappelijke disciplines. Tijdreeksen ontstaan dan ook zodra er verspreid over de tijd vergelijkbare metingen gedaan worden aan een bepaald systeem en/of toestand daarvan.

Binnen het hele toepassingsgebied van 'tijdreeksmethoden' kun je van nature een onderscheid maken tussen meer statistische en meer fysische (hydrologische) benaderingen. In algemene termen is er een onderscheid tussen:

- **statistische methoden** – die gebruik maken van de (statistische) eigenschappen van de gegevens en tijdreeksen die verzameld worden;
- **fysische methoden** – deze zijn in het algemeen gericht op het in kaart brengen en modelleren van de specifieke fysische wetten en eigenschappen die gelden voor het bemeten systeem.

Beide methoden en benaderingen sluiten elkaar niet uit maar kunnen elkaar juist aanvullen. De statistische en fysische benaderingen hebben elk bepaalde voor- en nadelen (zie paragraaf 2.7), in sommige situaties zal een statistische methode beter toegepast worden, in andere gevallen de meer fysische benadering. Het is om te beginnen de kracht van de statistiek dat (gedetailleerde) kennis van de onderliggende fysica in principe niet nodig is voor toepassing ervan. 'Statistiek' is daarmee generiek, en overstijgt en is toepasbaar binnen verschillende fysische disciplines. Deze algemene toepasbaarheid biedt naast praktische ook principiële voordelen, omdat de fysische eigenschappen van natuurlijke systemen in het algemeen niet alleen niet volledig bekend zijn (zie Afbeelding 2.1). Het is ook niet mogelijk om alle fysische eigenschappen volledig en exact in kaart te brengen, omdat daarvoor om te beginnen feitelijk het gehele grondwatersysteem afgegraven zou moeten worden (Konikow en Bredehoeft, 1992; Oreskes e.a., 1994).

Het generieke karakter van statistiek kan naast een voordeel ook een nadeel zijn, omdat statistiek abstract en daarmee lastig te begrijpen kan zijn voor hydrologen en andere niet-statistici. Het ontbreken van gebruiksvriendelijke software helpt daar niet bij. Wellicht daarom, én vanwege het ontbreken van gebruiksvriendelijke software, was het toepassen van (statistische) tijdreeksmodellen in de jaren negentig voorbehouden aan een beperkt aantal experts, zoals Van Geer e.a. (1988), Baggelaar (1988) en Rolf (1989).

Afbeelding 2.1 De ondergrond is heterogeen. De fysisch-hydrologische eigenschappen ervan zijn in het algemeen niet volledig bekend en ook niet volledig kenbaar zonder bijv. de ondergrond af te graven (Konikow en Bredehoeft, 1992; Oreskes e.a., 1994, bron illustratie: www.nationalgeographic.es)



Dit hoofdstuk bevat verder een toelichting op de methoden en modellen die algemeen gebruikt worden voor het analyseren van tijdreeksen van de stijghoogte of grondwaterstand, en beperken ons daarbij in technische zin tot de zogeheten transferruismodellen (zie paragraaf 2.3). We hanteren hierbij eenzelfde indeling als hiervoor beschreven, namelijk tussen:

- statistisch georiënteerde ARMA-transferruismodellen van Box en Jenkins, (zie paragraaf 2.2), hierna genoemd Box-Jenkins Transferruis (**BJTR**) modellen zoals gebruikt binnen Tijdreeksanalist en Menyanthes (Hydromonitor); en
- fysisch-hydrologisch georiënteerde 'Predefined Impulse Response Function In Continuous Time' (**PIRFICT**) transferruismodellen (zie paragraaf 2.4.3), die min of meer een integratie vormen tussen statistiek en hydrologie, zoals gebruikt binnen Pastas, Menyanthes (HydroMonitor) en Metran.

Opgemerkt wordt dat het tweede type model ook wel als impulsrespons-model wordt aangeduid, maar ook de statistische methode kent impulsresponsfuncties.

2.1.2 Overeenkomsten tussen beide methoden

De dwarsverbanden tussen statistiek en hydrologie worden onvoldoende herkend, waardoor de verschillen tussen de tijdreeksmethoden doorgaans eerder uitvergroot dan overbrugd worden. In deze handleiding wordt vooral ingegaan op de (technisch-inhoudelijke) overeenkomsten en dwarsverbanden tussen het statistisch georiënteerde BJTR-model en het hydrologisch georiënteerde PIRFICT-model. Met het inzicht in de overeenkomsten, worden de verschillen ook helder en wordt het mogelijk om waar nuttig of nodig een gefundeerde keuze te maken tussen beide methoden, zie ook paragraaf 2.6.

Zoals hierboven beschreven gaat het bij beide methoden om een transferruismodel. Dergelijke modellen bevatten een transferfunctie om een impuls (een verandering in het (hydrologisch) systeem, zoals neerslag) te vertalen naar een respons (verandering grondwaterstand). Tegelijkertijd is en blijft het zo dat de ogen van een hydroloog in stijghoogtereeksen heel andere dingen zien en opvallen dan die van een statisticus, en ook de doelen en toepassingen zijn deels anders. Daarnaast vergen beide typen methoden een andere aanpak. Daarom scheiden we beide typen methoden wel in het 'praktische stappenplan' dat te vinden is in de hoofdstukken hierna (hoofdstukken 3 t/m 5).

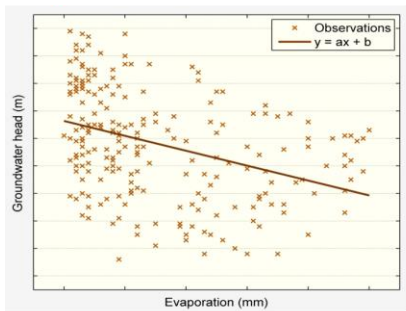
2.2 Over regressie en ARMA-functies

De oorsprong van tijdreeks- of transferruismodellen, zoals die het onderwerp vormen van deze handleiding, ligt in de statistiek en het daar ontwikkelde en toegepaste ARMA model (dat staat voor 'Auto Regressive Moving Average', zie Box en Jenkins (1970) en/of paragraaf 12.1). Het ARMA-model lichten we hieronder toe, waarbij we lineaire regressie als startpunt nemen. Van daaruit voegen we stap-voor-stap een stukje complexiteit toe, gaande dus van eenvoudige lineaire regressie tot aan een ARMA-transfermodel (zie Afbeelding 2.2). De achterliggende vergelijkingen zijn te vinden in het theorie-onderdeel van deze handleiding (zie Bijlage van de handleiding, paragraaf 12.1 en 12.2).

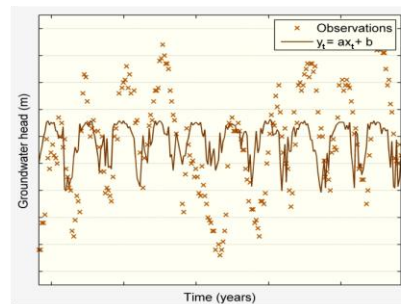
- 1 **Lineaire regressie** – is wellicht de meest bekende en eenvoudige methode voor het leggen van (statistische) verbanden tussen zogeheten 'verklaarde' en 'verklarende' variabelen. Eenvoudig gezegd komt lineaire regressie neer op het 'fitten van een lijn op een puntenwolk'. In Afbeelding 2.2 is een niet alledaags, maar wel alledaags ogend, voorbeeld te zien van het resultaat van lineaire regressie, waarbij verdamping als verklarende en de stijghoogte als verklaarde variabele is gebruikt.
- 2 **Dynamische regressie** – is in essentie niet anders dan lineaire regressie, maar alle meetwaarden van de verklarende en de verklaarde variabele hebben in dat geval een tijdsindex waarmee ze tijdsafhankelijk of dynamisch worden. In Afbeelding 2.2 is wederom een voorbeeld te zien van het resultaat van dynamische regressie. Het zijn dezelfde data en resultaten als die van de lineaire regressie, maar in deze vorm zien die er al uit als die van een tijdreeksmodel (echter met 'slechte fit').
- 3 **Regressie met verleden** – is de volgende stap, met als essentie dat de verklarende variabele (verdamping) naast een instantaan ook een mogelijk persistent effect heeft. Vertaald naar ons voorbeeld betekent dit dat de verdamping van gisteren ($t - 1$) ook vandaag (t) nog niet zomaar aangevuld is. Dit principe kan eenvoudigweg in het regressiemodel verdisconteert worden, door de verdamping op ($t - 1$) ook mee te nemen in de regressievergelijking. In Afbeelding 2.2 is een illustratie te zien van hoe parameters in de regressievergelijking in dat geval vertaald kunnen worden in een zogeheten responsfunctie.
- 4 **ARMA-functie** – Het zetten van deze laatste stap is nodig vanwege de beperkte toepasbaarheid van de vorige (regressie met verleden) voor trage systemen. Hydrologische effecten kunnen, bijvoorbeeld in het geval van een groot grondwatersysteem als het Veluwemassief, tot één of meer decennialang na-ijlen. Een tijdreeksmodel met een dagelijkse frequentie zou in de hiervoor geschetste aanpak duizenden modelparameters vereisen, die niet meer onafhankelijk te schatten zijn met een dergelijke beperkte dataset. Een oplossing hiervoor is het combineren van termen van **verklarende** variabelen uit het verleden (zogeheten MA (Moving Average) termen) met termen uit het verleden van de **verklaarde** variabele zelf (zogeheten AR (AutoRegressive) termen). Vertaald in de vorm van een responsfunctie zorgt één enkele AR-parameter voor een tot in het oneindige exponentieel afnemend effect of 'staart' van de respons.

Afbeelding 2.2 Stapsgewijs antwoord op de vraag 'wat is een tijdreeksmodel' vanuit statistisch oogpunt, gaande van eenvoudige lineaire regressie tot aan een ARMA-transferfunctie. Zie voor toelichting de tekst hierboven

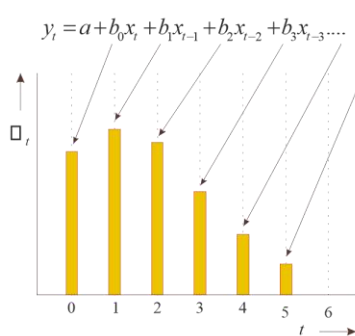
1) Lineaire regressie



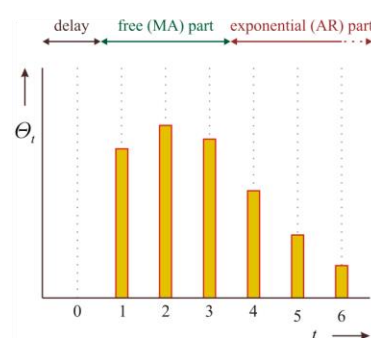
2) Dynamische regressie



3) Effecten verleden



4) ARMA-functie



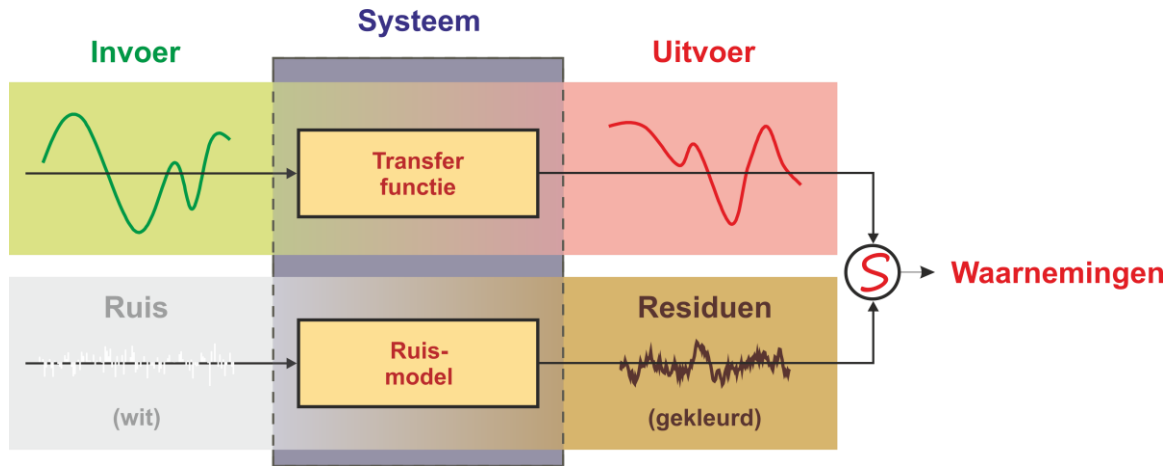
2.3 Het transfer- en ruismodel

Naast het in de vorige paragrafen gehanteerde onderscheid tussen statistiek en fysica, speelt ook het onderscheid tussen determinisme en stochastiek een belangrijke rol bij tijdreeksmodellen, omdat dit samenhangt met het onderscheid tussen het zogeheten transfergedeelte van het model, en het ruismodel. Dit onderscheid is als volgt (zie Afbeelding 2.3):

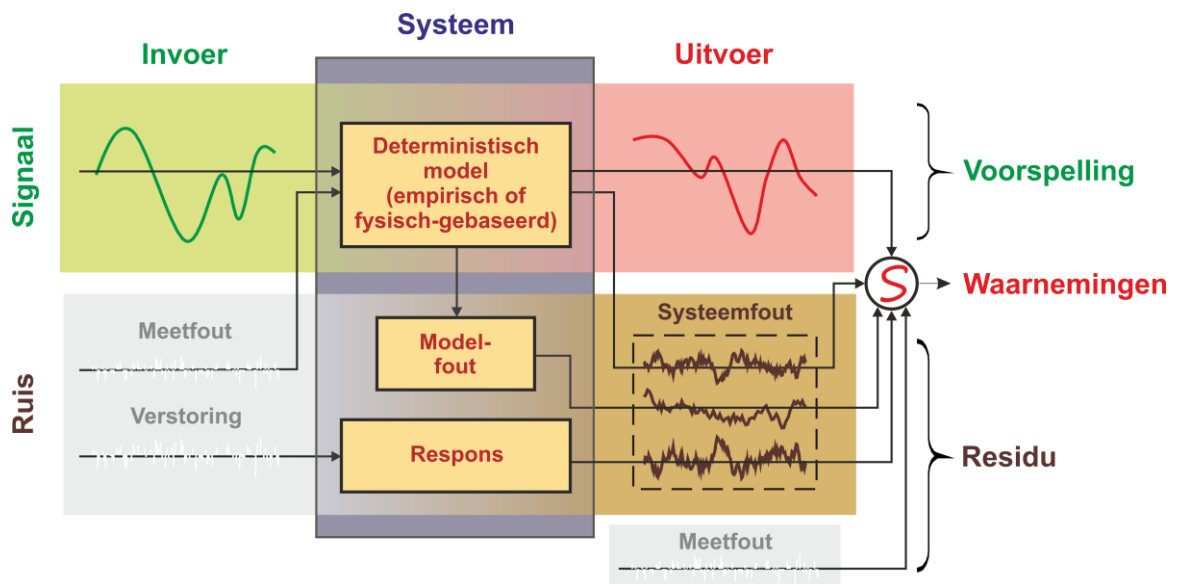
- 1 **Transfermodel (deterministisch)** – deterministische modellen of vergelijkingen leveren bij een gegeven situatie en/of een gegeven invoer slechts één enkele uitkomst of toestand op. Indien de situatie en/of invoer van tevoren bekend zijn, is ook de toekomstige toestand van tevoren te voorspellen. Het transfermodel verzorgt bij tijdreeksmodellen de conversie van de (deterministische) invoerreeksen naar hun bijdrage aan de uitvoer;
- 2 **Ruismodel (stochastisch)** – de uitkomsten van een stochastisch model of vergelijking zijn toevallig en niet van tevoren kenbaar. Er is geen causaal verband tussen situatie en uitkomst, of sprake van sturing. De toestand kan wel afhangen van die uit het verleden, indien er sprake is van systemen met geheugen. Het ruismodel verzorgt bij tijdreeksmodellen de modellering van fouten of residuen van het transfermodel. Voor berekeningen waarbij de onzekerheid aangegeven wordt, dient de invoer van het ruismodel (conceptueel gezien) stochastische, zogeheten witte ruis te zijn (Afbeelding 2.3). Dat wil zeggen dat de invoer geen zogeheten autocorrelatie vertoont (zie paragraaf 3.4.6, 4.4.3 en 5.4.3) en gemiddeld 0 is.

Het onderscheid tussen determinisme en stochastiek moet niet verward worden met dat tussen statistiek en fysica: een statistisch (of beter gezegd empirisch) model kan deterministisch zijn, en ook een fysisch model stochastisch. Het onderscheid tussen statistiek en fysica dat ook verderop in deze handleiding gehanteerd wordt slaat met name op het deterministische deel, ofwel de transferfuncties van tijdreeksmodellen. Het transfermodel verzorgt in feite ook de kern van de voorspelling (als de tijdreeks tenminste afdoende verklaard kan worden a.d.h.v. de invoer).

Afbeelding 2.3 Schema van een gecombineerd *single input* transferruismodel, met deterministisch transfergedeelte en stochastisch ruis-model (von Asmuth, 2012). De transferfunctie zelf kan zowel statistisch als fysisch-hydrologisch van aard zijn, of iets daar tussen in. NB zoals in paragraaf 1.3 beschreven is ruis en residu bij BJTR omgekeerd



Afbeelding 2.4 Overzicht van bronnen en effecten van ruis en verstoring in een lineair, open systeem (von Asmuth, 2012). Hierin wordt voor de ruis als invoer onderscheid gemaakt tussen de meetfout en verstoringen die niet in het transfermodel meegenomen worden (bijvoorbeeld een tijdelijke bemaling). In de uitvoer zijde bestaat het residu uit ruis van het transfermodel, modelfout, verstoringen en de meetfout



2.4 Verschillende transfermodellen

2.4.1 Gemeenschappelijke basis

De gemeenschappelijke basis van de BJTR en PIRFICT modellen wordt onder andere gevormd door de impulsresponsfunctie. In het PIRFICT-model wordt hierbij de volgende vergelijking gehanteerd, in vereenvoudigde vorm:

$$h(t) - d = (\theta * p)(t) \quad (1)$$

Dit betekent dat de toestand van het systeem (h) ten opzicht van de voorheen stabiele toestand d (bijvoorbeeld drainagebasis, deze parameter wordt doorgaans meegeeft met het tijdreeksmodel) bepaald wordt door de impulsresponsfunctie θ en invoer p (bijvoorbeeld neerslag). Het bovenstaande impliceert om te beginnen dat je de dynamiek van een willekeurige variabele op een willekeurige locatie in een willekeurig fysisch systeem kunt berekenen. Die berekening zou zelfs exact zijn, als de responsfunctie θ en verder de neerslag p en de drainagebasis d exact bekend zouden zijn. Net zoals de eigenschappen van de ondergrond zelf is uiteraard ook θ niet zomaar exact bekend. Met kan worden formules aangetoond dat zowel het BJTR-, het PIRFICT- als (lineaire) grondwatermodellen in de basis gelijk en in elkaar om te schrijven zijn.

De overeenkomst tussen het BJTR- en PIRFICT-model is dat het beiden methoden zijn om de systeemrespons te modelleren en benaderen, het verschil zit hem in de manier waarop ze dat doen. In de volgende paragrafen zal dit verder uit de doeken gedaan worden.

Gemeenschappelijke basis in de vergelijkingen

Het discrete en statistische BJTR-tijdreeksmodel kan het best vergeleken worden met het continue en fysisch-hydrologisch georiënteerde PIRFICT-model, door uit te gaan van impulsresponsfuncties en convolutie (een wiskundige bewerking die 2 functies combineert tot een nieuwe functie) als gemeenschappelijke basis.

Het opstellen van een fysisch gebaseerd model begint in het algemeen met het afleiden van een differentiaalvergelijking die op het desbetreffende systeem of variabele van toepassing is. Differentiaalvergelijkingen worden daarbij op hun beurt afgeleid van twee soorten vergelijkingen (zie bijvoorbeeld Bear (1972):

- continuïteitsvergelijkingen;
- constitutieve vergelijkingen.

De continuïteitsvergelijking (die uitgaat van het principe van massabehoud) is bij hydrologen beter bekend onder de naam waterbalans. Als constitutieve vergelijking (die de eigenschappen van het medium specificeert) geldt voor grondwater de wet van Darcy. Uit de combinatie van beide kan de differentiaalvergelijking voor grondwaterstroming afgeleid worden, die je vervolgens op kunt lossen voor een impuls of Dirac-deltafunctie δ als invoer (Dirac, 1947). De Dirac-deltafunctie is feitelijk een wiskundige abstractie, met als definitie :

$$\begin{cases} \delta(t) = 0, & t \neq 0 \\ \int_{-\infty}^{\infty} \delta(t) dt = 1 \end{cases} \quad (2)$$

Uitgedrukt in woorden komt de deltafunctie neer op een instantane impuls met eenheidsoppervlak. Toegepast op grondwater kun je hem zien als (oneindig) korte regenbui met eenheidshoogte (bijv. 1 mm). De manier waarop het grondwatersysteem vervolgens reageert op een dergelijke impuls van neerslag, wordt beschreven door de zogeheten impulsresponsfunctie θ . Ook θ is (logischerwijze) een wiskundige abstractie, met als definitie:

$$\begin{cases} \theta(t) = h(t) - d \\ h(t) = d, & t < 0 \\ p(t) = \delta(t) \end{cases} \quad (3)$$

De impulsresponsfunctie θ kun je in woorden definiëren als het effect dat invoer p heeft op de toestand h van het systeem, waarbij θ de afwijking in de tijd beschrijft van de voorheen stabiele toestand d . Toegepast op grondwater kun je θ zien als reactie van de stijghoogte h op een zeer korte regenbui p met eenheidshoogte (bijv. 1mm), terwijl de stijghoogte afgezien daarvan verder constant is en gelijk aan de lokale drainagebasis d . Door uit te gaan van lineariteit ofwel het superpositiebeginsel (zie paragraaf x) kun je vervolgens de reacties op verschillende gebeurtenissen of regenbuien eenvoudigweg bij elkaar op tellen. Bij een onveranderlijk systeem is bovendien ook de impulsresponsfunctie θ onveranderlijk. Met behulp van deze uitgangspunten kun je de stijghoogte h al met al berekenen voor een willekeurige neerslagreeks met behulp van de volgende formule, volgens het principe van Duhamel (Duhamel, 1833):

$$h(t) - d = \int_{-\infty}^t \theta(t - \tau) p(\tau) d\tau \equiv \int_0^{\infty} \theta(t) p(t - \tau) d\tau \equiv (\theta * p)(t) \quad (4)$$

De integraal in deze formule staat bekend als convolutie-integraal, de berekening kortweg als convolutie. De laatste vorm in vergelijking (4) achter het ' \equiv ' teken is feitelijk slechts een compacte manier om een convolutieproduct weer te geven. Het symbool τ in de formule is de integratievariabele.

2.4.2 Box-Jenkins transferruismodel (BJTR)

Deze paragraaf beschrijft het klassieke BJTR-model, zoals beschreven in Box en Jenkins (1970). In hoofdstuk 3 en 4 wordt beschreven hoe dit model wordt opgesteld. Daarnaast zijn voor deze methoden voorbeelden en casussen opgenomen in hoofdstuk 7 en 8. Het gaat uit van equidistante tijdreeksen, dat wil zeggen met een constant tijdsinterval tussen opeenvolgende reekswaarden. De invoerreeksen mogen daarbij een kleiner meetinterval hebben dan de uitvoerreeksen, maar ze mogen geen missende waarden hebben.

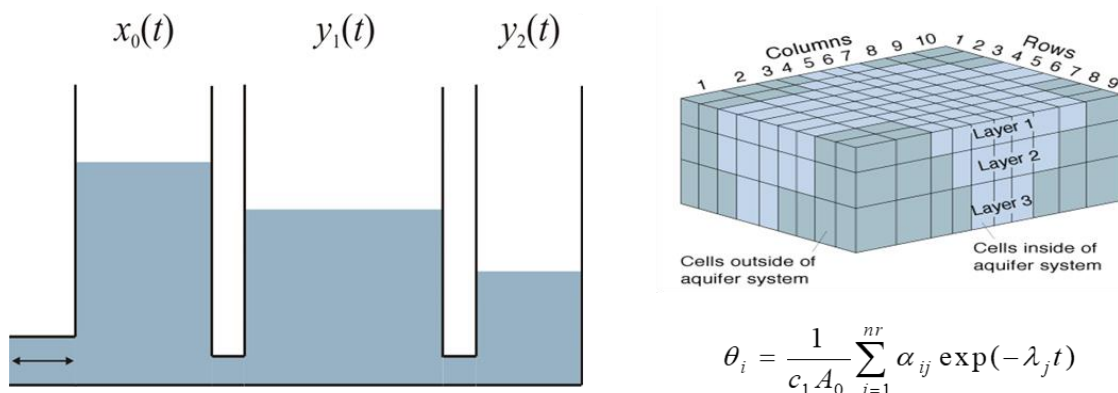
Het model hanteert transferfuncties in ARMA-vorm, waarmee een breed scala aan dynamische relaties kan worden beschreven. In paragraaf 2.2 is het ARMA-model al geïntroduceerd en toegelicht vanuit het gezichtspunt van de statistiek, en stond de link met hydrologie en fysica op de achtergrond.

Onderstaand zijn de vergelijkingen weergegeven van het ARMA model. Het resultaat van deze vergelijkingen betreft onderstaande formule, die tevens inzicht geeft in het AR (Auto Regressive) en Moving Average gedeelte.

$$h_t - d - \underbrace{\delta_1(h_{t-1} - d) - \dots - \delta_{nr}(h_{t-nr} - d)}_{\text{Auto Regressive}} = \underbrace{\omega_0 P_t + \omega_1 P_{t-1} + \dots + \omega_{ns} P_{t-ns}}_{\text{Moving Average}} \quad (5)$$

Het rechter- of MA-deel van vergelijking komt neer op dynamische regressie met verdiscontering van de effecten van de **verklarende** variabele uit het verleden (bijvoorbeeld de neerslag), het linkerdeel is het AR-deel dat neerkomt op dynamische regressie op basis van de toestand van de **verklaarde** variabele in het verleden zelf (bijvoorbeeld de grondwaterstand). De vraag hoever het model hierbij terug kijkt in de tijd, wordt bepaald door de symbolen nr en ns in de vergelijking, die zowel het aantal tijdstappen als het aantal AR- en MA-parameters definiëren. Bij ARMA-transferruismodellen moet het aantal AR- en MA-parameters vooraf door de modelleur gekozen worden op basis van de zogeheten modelidentificatieprocedure, en dat van zowel het transfer- als het ruismodel (zie paragraaf x). De gekozen parameteraantallen vormen daarbij de zogeheten modelorde van het tijdreeksmodel. Dit wordt tussen haakjes aangeduid voor het AR en MA deel, bijvoorbeeld ARMA(1,0) of ARMA (1,2). Vaak wordt voor de orde van een AR-model één parameter gekozen (AR(1)).

Afbeelding 2.5 Fysische equivalenten van AR-modellen. Links: een reeks gekoppelde lineaire reservoirs volgens Box en Jenkins (1970). Hierbij is er een respons in de 2 rechter reservoirs op de impuls in het linker reservoir. Rechts: schematische weergave van een 3-dimensionaal model: een lineair, numeriek grondwatermodel met nr cellen, met som van exponentiele functies of AR(nr)-model als oplossing (Sahuquillo, 1983)



ARMA-modellen worden op allerlei soortige systemen en variabelen toegepast, en worden daarbij meestal gezien als 'black box'-modellen. Van veel systemen in de alfa- en gammawetenschappen (bijv. de econometrie) is het ook niet zomaar mogelijk om een fysisch-deterministisch model of verklaring te geven. Wiskundig gezien zorgt AR-parameter δ_1 voor exponentiele af- of toename van de verklaarde variabele (zie ook paragraaf 2.2), vanuit hydrologisch perspectief komt dit neer op een wegzijgingsterm (in het geval van een enkele AR-parameter is dit een lineair waterreservoir en/of eenvoudige bodemkolom (Box en Jenkins, 1970; Knotters en Bierkens, 2000)). MA-parameters kunnen daarentegen gezien worden als aanvulling die vertraagd het grondwater bereikt, bijv. vanwege de werking van de onverzadigde zone (Von Asmuth, 2012).

Vergelijking ARMA

In het ARMA-model wordt vergelijking (3) geformuleerd in discrete tijd als discrete convolutie:

$$h_t - d = \sum_{i=-\infty}^t \Theta_{t-i} P_i \equiv \sum_{i=0}^{\infty} \Theta_i P_{t-i} \equiv \Theta(B)P_t \equiv (\Theta * P)_t \quad (6)$$

De convolutie-integraal in vergelijking (3) wordt hier dus een sommatie. In deze vergelijking is t een discrete tijdsindex ($t \in \mathbb{N}$), bij P gaat het om neerslaghoeveelheden, Θ is de transfer- of blokresponsfunctie, en B is de zogeheten *backward shift* operator, die gedefinieerd wordt door:

$$B^b P_t = P_{t-b} \quad (7)$$

De transfer- of blokresponsfunctie Θ is afhankelijk van de modelfrequentie, wordt hier benaderd met een ARMA-functie en daarbij symbolisch geschreven als (Box and Jenkins, 1970; Von Asmuth 2012):

$$\Theta(B) = \delta^{-1}(B)\omega(B) \quad (8)$$

In deze vergelijking is $\delta(B)$ de zogeheten AutoRegressive (AR) operator (niet te verwarren met de Dirac delta functie), en $\omega(B)$ is de zogeheten Moving Average (MA) operator. De berekening van h_t gebeurt vervolgens recursief met behulp van de volgende vergelijking:

$$h_t - d - \delta_1(h_{t-1} - d) - \dots - \delta_{nr}(h_{t-nr} - d) = \omega_0 P_t + \omega_1 P_{t-1} + \dots + \omega_{ns} P_{t-ns} \quad (9)$$

Het rechter- of MA-deel van vergelijking (7) na het '='-teken komt neer op dynamische regressie met verdiscontering van de effecten van de **verklarende** variabele uit het verleden, het linkerdeel is het AR-deel dat neerkomt op dynamische regressie op basis van de toestand van de **verklaarde** variabele in het verleden zelf. De vraag hoever het model hierbij terug kijkt in de tijd, wordt bepaald door nr en ns in vergelijking (7), die zowel het aantal tijdstappen als het aantal AR- en MA-parameters definiëren. De gekozen parameteraantallen vormen daarbij zoals gezegd de zogeheten modelorde van het tijdreeksmodel.

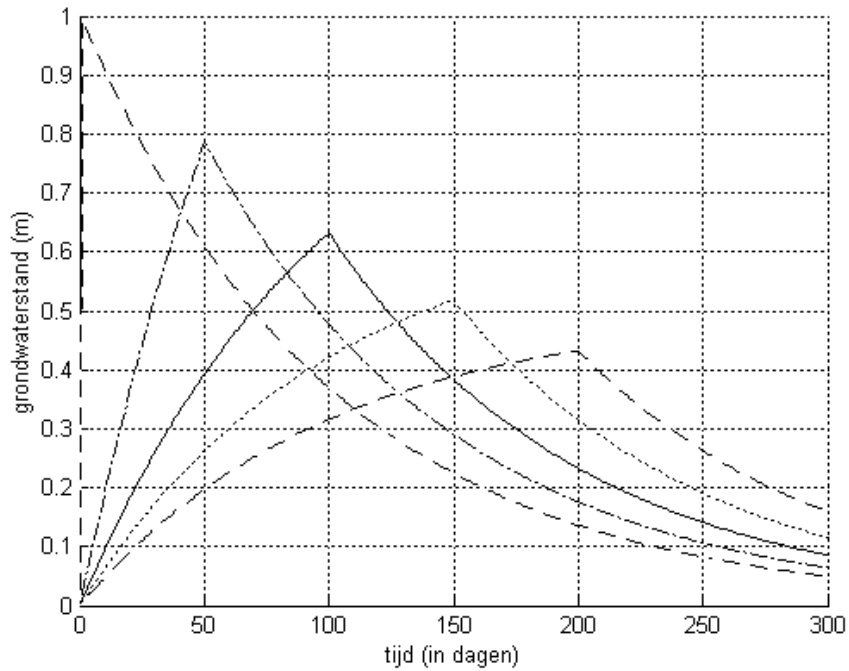
Vanwege de recursieve formulering van vergelijking (7) is een tijdreeksmodel in deze vorm rekenkundig efficiënter en minder veeleisend dan geformuleerd als convolutie (vergelijking (3) of (4)). In het verleden was dit waarschijnlijk een belangrijk voordeel (ARMA-modellen bestonden immers al in de begintijd van de computers), maar tegenwoordig is rekenkracht geen beperking meer. Bij een tijdreeksmodel dat geschreven is in convolutievorm is het ook mogelijk om voor de impulsresponsfunctie θ en/of transferfunctie Θ fysisch-hydrologische of andere benaderingen te gebruiken, wat in de volgende paragrafen verder aan de orde komt.

Discrete tijd in ARMA-modellen

Bij ARMA-modellen wordt een discrete tijd gehanteerd. Dit betekent dat de transferfunctie vanuit continue tijd gezien een *blok*responsfunctie betreft (ofwel de respons op invoer die de vorm heeft van een blok, bijvoorbeeld 1 mm regen in 1 dag).

Onderstaande afbeelding geeft blokresponsfuncties weer voor 4 verschillende modelfrequenties of tijdstappen Δt , uitgaande van dezelfde, exponentiële impulsrespons. De gehanteerde tijdstap of frequentie is dus van invloed op de respons en resultaten.

Afbeelding 2.6 Blokresponsfuncties van een systeem met een exponentiële impulsrespons voor verschillende tijdstappen ($A=100$, $a=0.01$, $n=1$, $\Delta t = [1 \ 50 \ 100 \ 150 \ 200]$) (bron: von Asmuth, 2002)



Discrete naar continue waarneming

Een discrete waarneming, bijvoorbeeld een neerslaghoeveelheid P die op een bepaald moment verzameld is in een regenmeter, kun je vanuit continu gezichtspunt zien als integraal van de onderliggend continue neerslagintensiteit p over de meetperiode. In formule vorm:

$$P_i = \int_{t_{i-1}}^{t_i} p(\tau) d\tau \quad (10)$$

Door uit te gaan van constante, gemiddelde neerslag kan de neerslagintensiteit gedurende de meetperiode eenvoudigweg afgeleid worden uit de discreet waargenomen neerslaghoeveelheden via:

$$p(\tau) = \frac{P_i}{t_i - t_{i-1}}, \quad t_{i-1} \geq \tau > t_i \quad (11)$$

Grafisch weergegeven neemt het verloop van de neerslag gedurende de meetperiode in dat geval de vorm aan van een blok, vandaar dus de naam blokresponsfunctie. De resultaten van ARMA- en PIRFICT-modellen zijn vervolgens met elkaar te vergelijken door de impulsresponsfunctie van die laatste om te rekenen naar een blokresponsfunctie Θ , door θ te convolueren met een blok met eenheidsintensiteit over meetperiode $\Delta t = t_i - t_{i-1}$. Wederom in formule vorm:

$$\Theta(t) = \int_{t-\Delta t}^t \theta(\tau) d\tau \quad (12)$$

Uit vergelijking (12) volgt o.a. dat Θ inderdaad afhankelijk is van de modelfrequentie of tijdstap Δt . In Afbeelding 2.6 zijn ter illustratie verschillende blokresponsfuncties voor verschillende tijdstappen weergegeven voor het geval dat θ een exponentiële functie is (of in hydrologische termen een eenvoudig lineair reservoir). Omdat Θ en p continue functies zijn, is de stijghoogte h in vergelijking (4) zelf ook continu gedefinieerd, en kan er voor elke waarneming van de stijghoogte ook een afwijking of waarde van de residureeks worden berekend.

2.4.3 Het PIRFICT-model

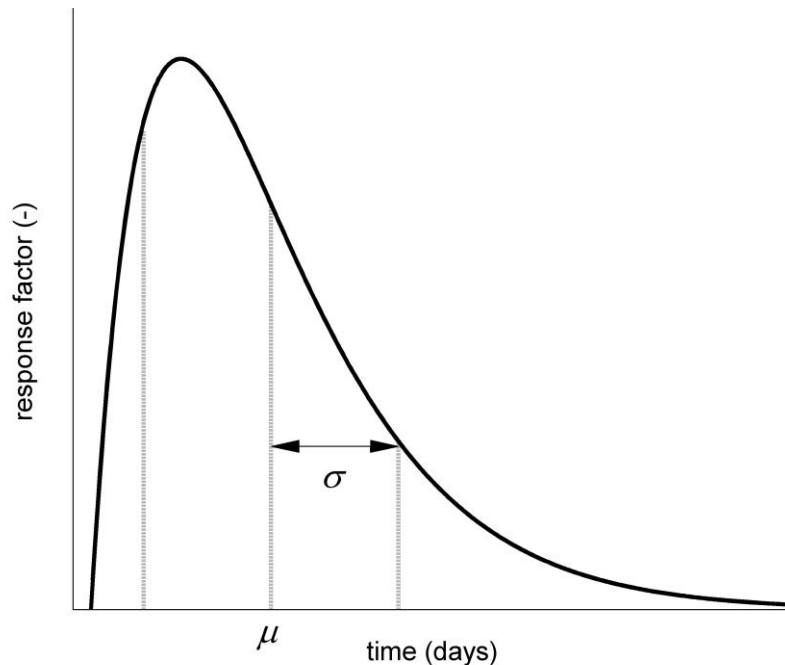
Het PIRFICT-model (dat staat voor Predefined Impulse Response Function In Continuous Time) is door Von Asmuth e.a. (2002b) ontwikkeld en beschreven als nieuw type tijdreeksmodel, met als belangrijk kenmerk dat het **continu is in de tijd**. In deze paragraaf wordt het principe van de PIRFICT methode uiteengezet. In hoofdstuk 3 en 5 wordt de opzet van een PIRFICT tijdreeksmodel beschreven. In hoofdstuk 10 en 11 zijn uitgewerkte voorbeelden en casussen opgenomen waarbij deze methode wordt gehanteerd.

Vanwege deze formulering kun je binnen het PIRFICT-model in principe een willekeurige continue impulsresponsfunctie θ kiezen als transfermodel, zolang deze een bekende integraal heeft en een analytische oplossing heeft. Vaak worden algemene oplossingen, zoals de **gammaverdeling** gebruikt als impulsresponsfunctie van neerslag naar grondwater. De gammaverdeling is een scheef-normale verdeling (Afbeelding 2.7). Deze verdeling is door Nash (1958), toegepast als *Unit Hydrograph* om de effecten van neerslag op afvoer van rivier- en beeksystemen te modelleren. Voor andersoortige hydrologische invloeden (zoals onttrekkingen) worden in praktijk echter ook andersoortige responsfuncties gebruikt (zie paragraaf 5.2.3).

Het is echter ook mogelijk om als impulsresponsfunctie een volledig **fysisch-deterministische**, analytische oplossing toe te passen. Belangrijk nadeel van een fysisch-deterministische oplossing is dat om te beginnen het kiezen of uitwerken van analytische oplossingen tijd en hydrologische kennis vergt.

Een andere belangrijke beperking is dat niet van elke specifieke situatie en systeem goede analytische oplossingen voorhanden zijn. Analytische oplossingen (en hydrologische modellen in het algemeen) zijn daarnaast gebaseerd op vereenvoudigingen en aannamen bij de modelschematisatie, die in de regel onzeker en soms ook onterecht zijn.

Afbeelding 2.7 Scheef-normale impulsresponsfunctie met gemiddelde μ en variantie σ^2 (von Asmuth, 2012)



Gamma functie als impulsresponsfunctie in PIRFICT?

De vraag of, wanneer en waarom de impulsresponsfunctie θ goed benaderd kan worden met een scheef-normale verdeling is helaas niet eenvoudig te beantwoorden (zie echter Maas (1994); Von Asmuth (2012) voor een gedeeltelijke verklaring m.b.v. de centrale limietstelling uit de statistiek). Dát verdelingsfuncties bruikbaar zijn is wellicht ook minder goed intuïtief invoelbaar, omdat in dit geval de wet van behoud van massa en de waterbalans niet opgaan (een **hoeveelheid** water vertaalt zich hier in een **waterstand**, en niet in een verdeling van waterhoeveelheden over de tijd). Om die reden heeft θ hier in de regel ook geen eenheidsoppervlakte. Bij toepassing als impulsresponsfunctie van de stijghoogte moet de gammaverdeling daarom vermenigvuldigd en geschaald worden tot geschaalde gammaverdeling, die beschreven wordt door:

$$\begin{cases} \theta(t) = A \frac{a^n t^{n-1} \exp(-at)}{\Gamma(n)}, & t \geq 0; \\ \theta(t) = 0 & , \quad t < 0. \end{cases} \quad (13)$$

met A , a en n als constanten. Een manier om te onderzoeken of en in hoeverre de geschaalde gammaverdeling bruikbaar is voor dit doel, is om de resultaten en vorm ervan te vergelijken met de resultaten van een BJTR-model dat toegepast wordt op dezelfde data (zie volgende paragraaf). Een andere manier is om te onderzoeken of en in hoeverre analytische oplossingen voor de respons van verschillende fysisch-hydrologische systemen inderdaad een scheef-normale verdeling volgen (von Asmuth, 2002). Hoe het ook zij, feit is dat de geschaalde gammaverdeling inmiddels in verschillende softwarepakketten en in talrijke studies succesvol is toegepast voor het analyseren van tijdstijghoogtereeksen en modelleren van de respons op neerslag en verdamping daarbij.

2.5 Verschillende ruismodellen

2.5.1 Het ruismodel

Het ruismodel beschrijft zoals gezegd dus de fouten of afwijkingen en onzekerheden van het transfermodel. Die fouten of residuen zijn daarbij gedefinieerd als het verschil tussen de (door het transfermodel) geschatte waarden en de waarnemingen. In formulevorm is dat te schrijven als de volgende eenvoudige optelsom:

$$h(t) = \hat{h}(t) + n(t) \quad (14)$$

waarbij h de waarnemingen zijn (de grondwaterstand of stijghoogte), \hat{h} de schatting daarvan en n het residu. Het ruismodel modelleert dus kortgezegd de fouten (afwijkingen) en onzekerheden van het transfermodel. Alhoewel ook deze afwijkingen (gedeeltelijk) een fysische oorsprong kunnen hebben, gaat het daarbij doorgaans om een verzameling van fouten en verstoringen van verschillende aard, inclusief fouten in de metingen of meetruis zelf. Omdat deze fouten per definitie niet door het transfermodel worden beschreven, vallen ze ook per definitie onder het ruismodel dat dus stochastisch van aard is. Voor een verdere uitleg over de toepassing van het ruismodel zie paragraaf 5.2.5 en de voorbeelden in paragraaf 9.3.

2.5.2 Verschillen in ruismodellen en terminologie

Voor het ruismodel geldt, net als voor het transfermodel, dat er zowel een gemeenschappelijk basis als verschillen spelen tussen het Box-Jenkins- en het PIRFICT-transferruismodel. Ook hierbij speelt als belangrijkste verschil dat het ruismodel van **BJTR discreet** is, terwijl dat van het **PIRFICT-model continu** is in de tijd. BJTR past voor het ruismodel ARMA-functies toe (Box en Jenkins, 1970), die in staat zijn om complexe autocorrelatiestructuren van de ruis te beschrijven. Dit biedt veel flexibiliteit om het modelresidu terug te brengen tot witte ruis. Het PIRFICT-model gebruikt een continue (exponentiele) functie (Von Asmuth en Bierkens, 2005). Hierdoor is het niet geschikt om complexe ruismodellen op te stellen.

Zoals in hoofdstuk 13 beschreven is, is er ook een verschillend gebruik van terminologie met betrekking tot het ruismodel ontstaan. Dat levert in de praktijk verwarring op. Het gaat hierbij met name om de termen **ruis**, **residuen** en **innovaties** en de precieze betekenis daarvan in verschillende contexten.

2.6 TRA met simultane invloeden en reeksen

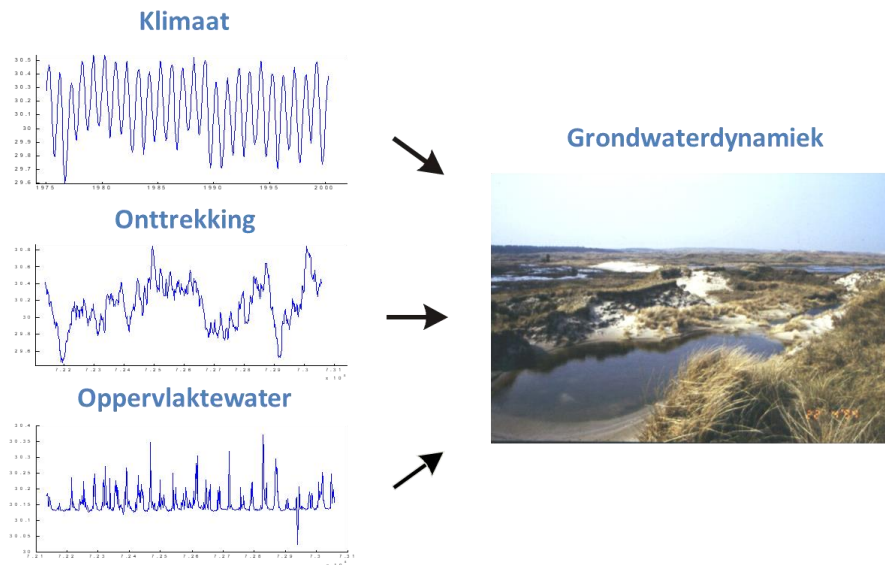
2.6.1 Inleiding

In de voorgaande paragrafen zijn we voor het gemak uitgegaan van de meest eenvoudige situatie waarin één enkele verklarende variabele of invloedsfactor de dynamiek van één verklaarde variabele bepaalt. In deze paragraaf gaan we in op de praktijk die vaak complexer is, met meerdere invloeden/variabelen met ook nog eens meerdere meetreeksen (verklaarde variabelen). In deel II van de handleiding zijn voorbeelden en casussen opgenomen waarin tijdreeksanalyse wordt uitgevoerd met zowel één als meer verklaarde variabelen worden gebruikt.

2.6.2 Simultane invloeden - scheiden van effecten

In de voorgaande paragrafen zijn we voor het gemak uitgegaan van de meest eenvoudige situatie waarin één enkele verklarende variabele of invloedsfactor de dynamiek van de verklaarde variabele bepaalt. In dat geval bestaat het transfergedeelte van het tijdreeksmodel feitelijk uit één enkele transfer- of responsfunctie. De grondwaterstand wordt in de regel echter door verschillende factoren tegelijk beïnvloedt, zeker in een drukbevolkt en sterk gereguleerde land als Nederland (Afbeelding 2.8).

Afbeelding 2.8 De grondwaterdynamiek op een bepaalde locatie wordt in de regel bepaald door de simultane invloed die verschillende factoren tegelijk daarop uitoefenen



Simultane invloeden bij BJTR, sommatie van invloeden

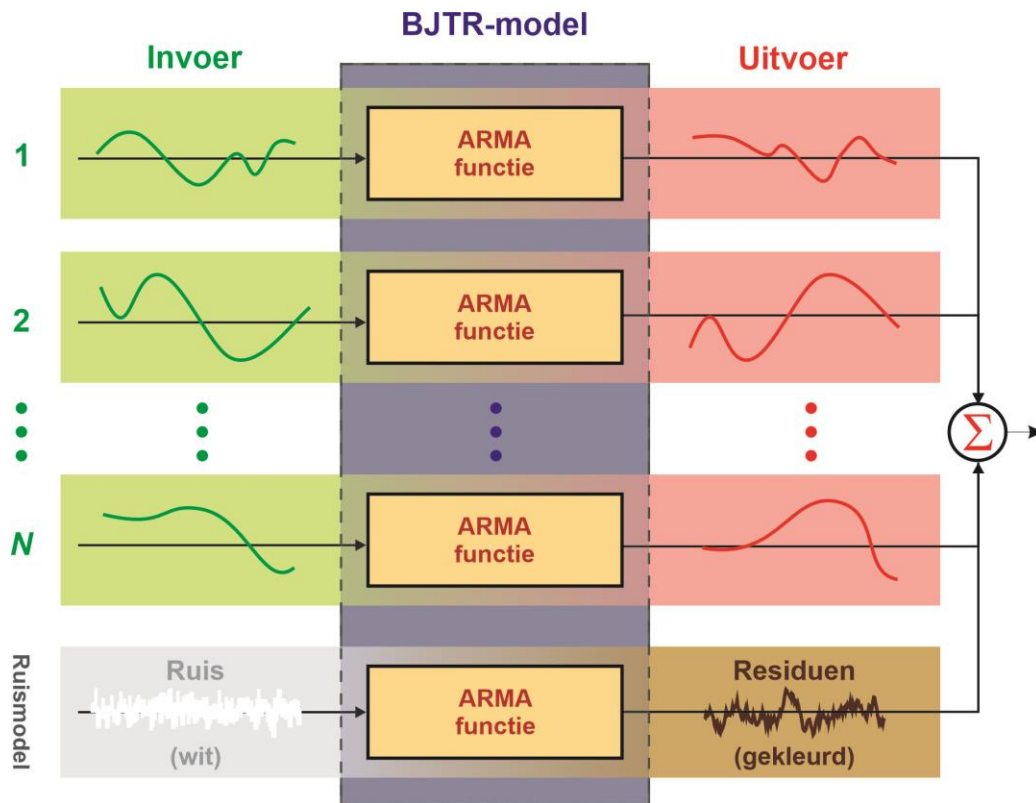
Bij BJTR analyse worden de effecten van simultane factoren (*multiple input*) in het algemeen verdisconteert door de bijdragen of effecten van die factoren bij elkaar op te tellen. Hierbij heeft elke invoer haar eigen ARMA-functie. In formulevorm betekent dit dat we de eerdere basisvergelijking uit paragraaf 2.5.1 uitbreiden tot:

$$h(t) = \sum_{i=1}^N \hat{h}_i(t) + d + n(t) \tag{15}$$

In deze formule doet *i* eenvoudigweg dienst als index of nummering van de verschillende (*N*) factoren die invloed uitoefenen in een bepaalde situatie.

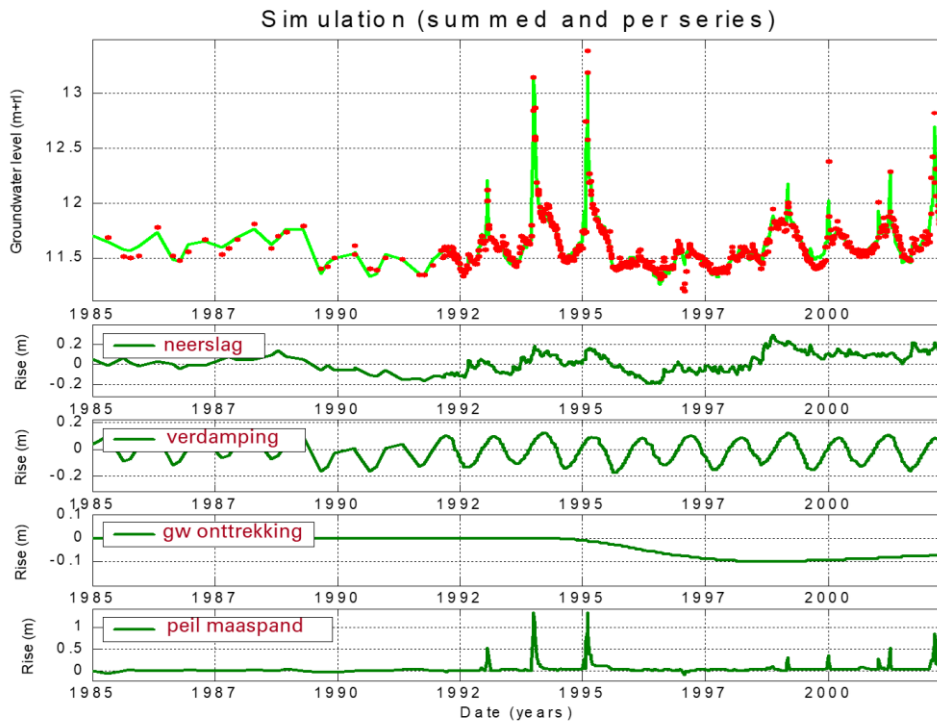
Vanuit statistisch oogpunt zijn alle invoerreeksen gelijkwaardig, er wordt in principe geen onderscheid gemaakt in de aard of het type van een bepaalde invloed met betrekking tot de impulsresponsfunctie. De verschillende factoren worden daarmee parallel geschakeld, waarbij elke invloed een eigen ARMA-functie toegewezen krijgt. De orde of vorm van de ARMA-functies wordt gekozen op basis van statistische of empirische criteria (zie paragraaf 4.2.1), de effecten van alle factoren worden vervolgens eenvoudigweg gesommeerd om tot een schatting van het hele transfermodel te komen (zie Afbeelding 2.9).

Afbeelding 2.9 Schema van een (klassiek) *multiple input* BJ-transferruismodel. De verschillende (N) invloedsfactoren of invoerreeksen worden parallel geschakeld en hun uitvoer of bijdrage gesommeerd



Het feit dat de effecten van alle factoren gesommeerd worden in een standaard transferruismodel betekent dat uitgegaan wordt van het zogeheten lineariteitsbeginsel. Superpositie kan immers alleen toegepast worden op lineaire systemen. Door uit te gaan van lineariteit zijn tijdreeksmodellen in staat om een schatting te geven van de individuele effecten van verschillende factoren (Afbeelding 2.10) die in het model dus losstaan van elkaar. Dit maakt het mogelijk om tijdreeksmodellen toe te passen in effectenstudies en schaderegelingen. Hydrologische systemen zijn echter welbeschouwd allemaal in meer of mindere mate niet-lineair. Superpositie is daarom alleen mogelijk als de niet-lineaire processen verwaarloosbaar zijn.

Afbeelding 2.10 Resultaten van een *multiple input* transferruismodel, weergegeven als match tussen waarnemingen (rode stippen) en simulatie van het transfermodel (lichtgroene lijn) in de bovenste grafiek en de bijdragen of effecten van de verschillende invloedsfactoren daaronder (donkergroene lijnen, Von Asmuth e.a. (2008))

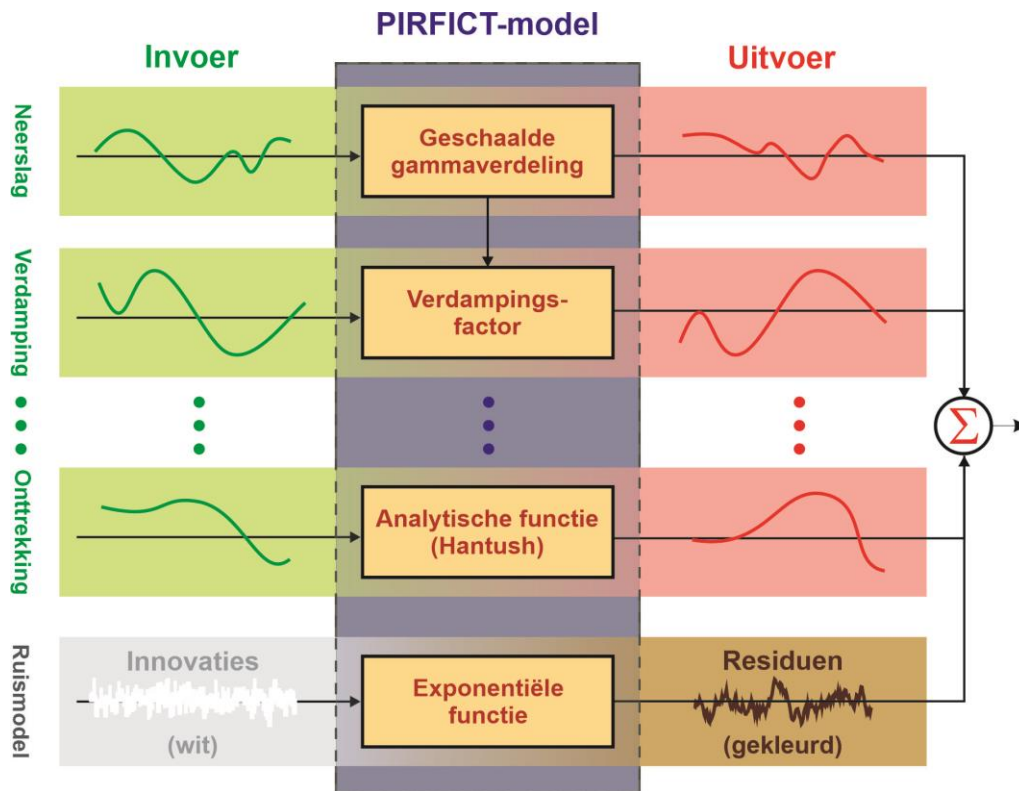


Simultane invloeden bij PIRFICT, aard van invloed

De verschillende invoerreeksen hebben verschillende hydrologische invloeden (Von Asmuth e.a., 2008): neerslag is een ander type invloed (of randvoorwaarde in de termen van een grondwatermodel) dan grondwateronttrekking, en zelfs de effecten van neerslag zijn niet overal en altijd hetzelfde. Een hevige stortbui in een winterperiode met hoge grondwaterstanden kan bijvoorbeeld leiden tot snelle en oppervlakkige afvoer, terwijl aan de andere kant motregen na een droge periode in de zomer in zijn geheel in de onverzadigde zone wordt geborgen, of als interceptie in het bladerdak van bomen en struiken blijft hangen. Voor de verschillende hydrologische invloeden worden dan ook verschillende responsfuncties toegepast.

Binnen het PIRFICT-model worden verschillende factoren in principe, net als bij BJTR, parallel geschakeld, maar de manier waarop deze behandeld worden en de responsfuncties die ze vervolgens toegewezen krijgen kan verschillen (Afbeelding 2.11).

Afbeelding 2.11 Schema van een *multiple input* PIRFICT-transferruismodel. De verschillende (N) invloedsfactoren of invoerreeksen worden daarbij parallel geschakeld en hun uitvoer of bijdrage eenvoudigweg gesommeerd



2.6.3 Simultane analyse en identificatie

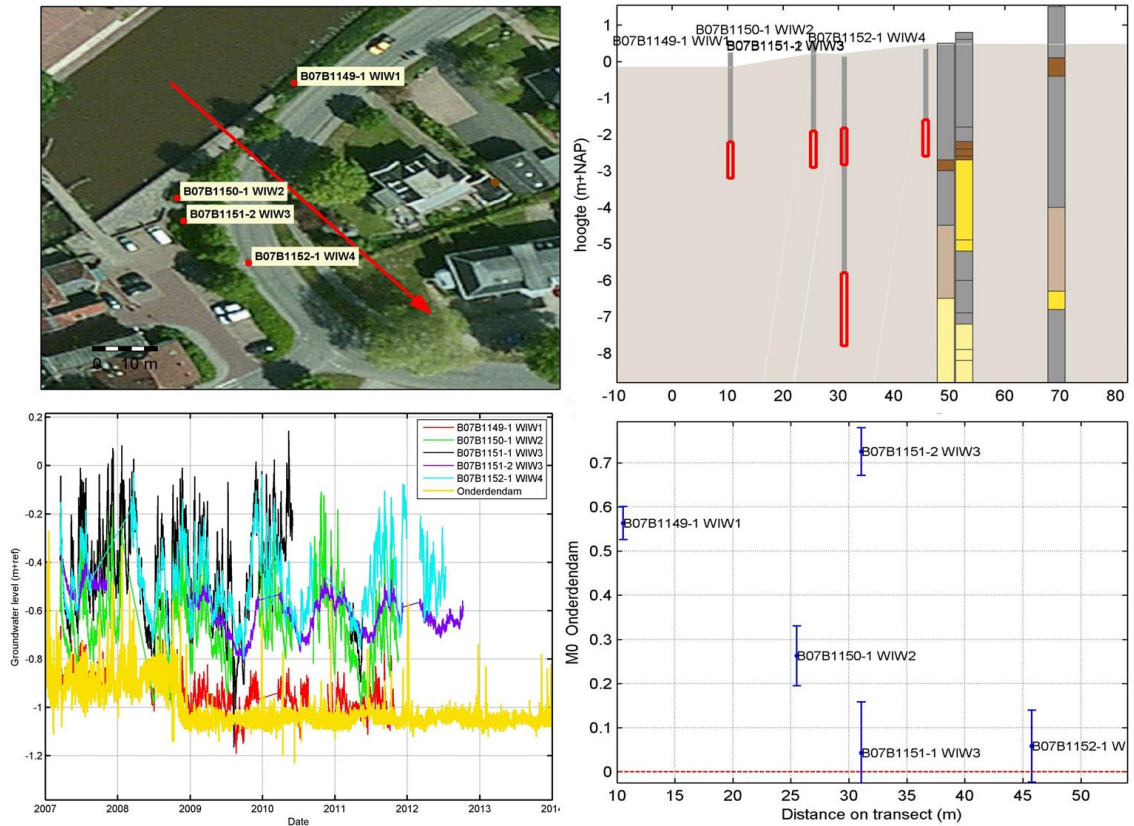
Naast dat er vaak meerdere invloeden zijn op de meetreeks, zijn er in de meeste gevallen ook meerdere meetreeksen (verklaarde variabelen). Hydrologisch gezien speelt het onderliggende geohydrologische systeem uiteraard een belangrijke rol en bepaalt dat systeem ook het de mate van bepaalde invloeden op de meetreeksen. De consequenties hiervan zijn (zie bijv. Afbeelding 2.12, (Von Asmuth en Leunk, 2014)):

- naast de eigenschappen van de meetreeks zijn ook de geohydrologische eigenschappen, bodemopbouw en ligging van de peilbuizen en hun filters van wezenlijk belang;
- het identificeren en beantwoorden van de vraag wélke factoren invloed uitoefenen op de grondwaterdynamiek hoeft niet noodzakelijkerwijs per factor en per meetreeks te gebeuren (zie ook. Dát bijvoorbeeld neerslag en verdamping invloed uitoefenen staat bij voorbaat vast;
- ook de effecten van een bepaalde factor (bijv. het boezempeil) vertonen in het algemeen ruimtelijke samenhang. De ruimtelijke samenhang voegt daarbij een dimensie toe (of letterlijk drie dimensies) aan de tijdreeksanalyseresultaten en de beoordeling daarvan.

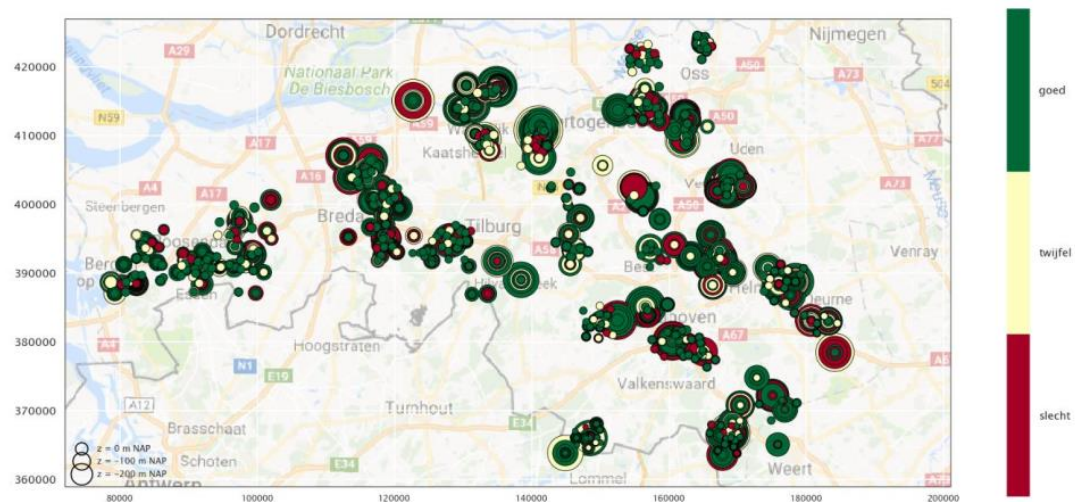
Bij toepassing van het PIRFICT-model is het mogelijk om veel grotere aantallen (honderden of zelfs duizenden) stijghoogtereeksen tegelijk te analyseren (zie Afbeelding 2.13, en bijv. (Von Asmuth en Leunk, 2012; Collenteur e.a., 2017)). Dit komt omdat de nadruk meer ligt op de *a posteriori* (naderhand) analyse, beoordeling en toetsing te liggen, dan op *a priori* (vooraf) empirisch onderzoek.

De ruimtelijke samenhang en situering van verschillende meetpunten binnen eenzelfde systeem, komt niet aan de orde in het klassieke tijdreeksanalysehandboek van Box en Jenkins (1970). Elke meetreeks wordt apart geanalyseerd en is daarmee bewerkelijk voor grote aantallen meetreeksen. Het analyseren van meerdere meetreeksen in samenhang komt wel aan de orde in het op (geo)hydrologische toepassingen gerichte handboek van Hipel en McLeod (1994), in de vorm van zogeheten *contemporaneous ARMA models*, ofwel CARMA-modellen.

Afbeelding 2.12 Voorbeeld van analyse van meerdere meetreeksen (linksonder), ligging van meetpunten (linksboven), geohydrologische situatie (rechtsboven) en geschatte doorwerking van het boezempeil (rechtsonder). Doel van de analyse is het schatten van de effecten van boezempeilverlaging (Von Asmuth en Leunk, 2014)

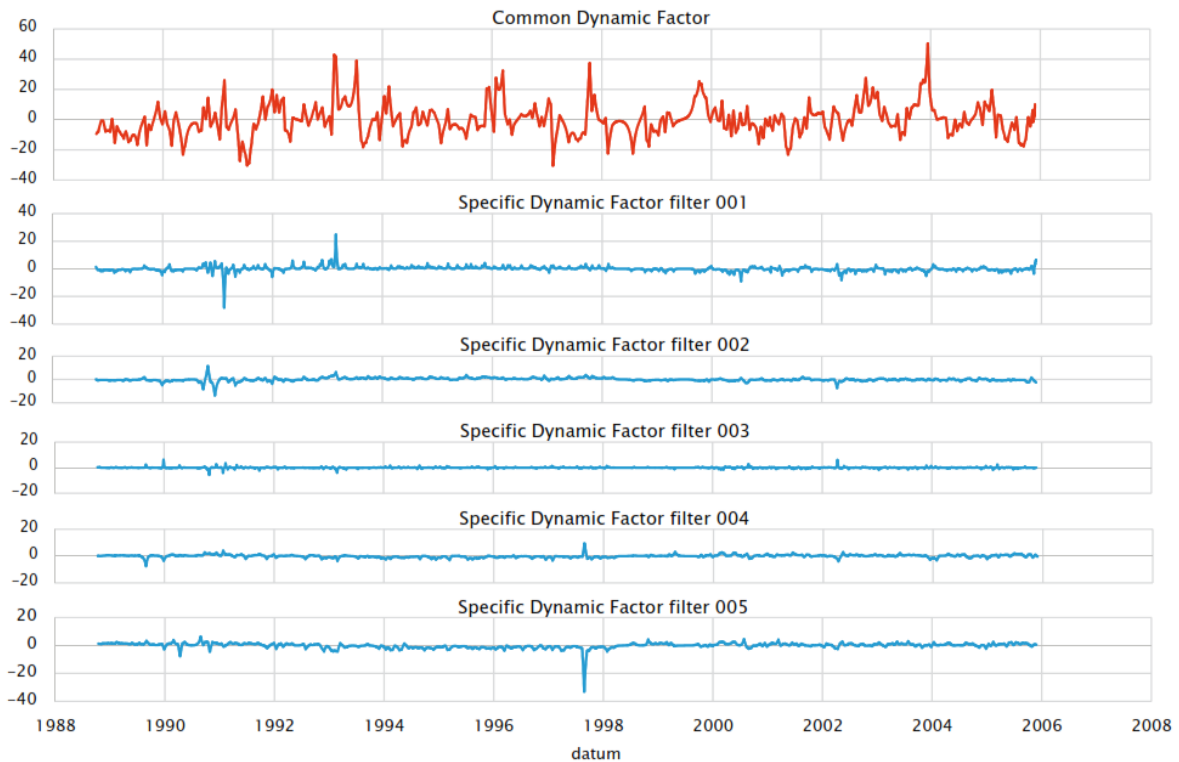


Afbeelding 2.13 Voorbeeld van analyse van grote aantal tijdreeksen tegelijk (boordeling op basis van de verklaarde variantie. 0=slecht, 1=twijfelachtig en 2=goed, bron: Colleteur e.a., 2017)



Een ander voorbeeld van simultane analyse is het gebruik van een zogeheten meervoudig tijdreeksmodel, c.q. het METRAN-model (Van Geer en Berendrecht, 2015; Berendrecht en van Geer, 2016). In een dergelijk model worden een aantal stijghoogtereeksen ook in letterlijke zin tegelijk geanalyseerd en gemodelleerd. In die benadering worden alle stijghoogtereeksen allereerst gemodelleerd volgens de PIRFICT-benadering met neerslag en verdamping als invoerreeksen, waarna de residuereeksen met een Dynamic Factor Model (Geweke, 1977; Stock en Watson, 2011) gemodelleerd en daarbij opgesplitst worden in een gezamenlijke Common Dynamic Factor en afzonderlijke Specific Dynamic Factors als resultaat (zie Afbeelding 2.14).

Afbeelding 2.14 Resultaat van toepassing van een meervoudig tijdreeksmodel, met opsplitsing van de residuereeksen in een Common Dynamic Factor en afzonderlijke Specific Dynamic Factors (bron: Van Geer en Berendrecht, 2015)



2.7 Vergelijking PIRFICT- en BJTR-model (overzicht)

2.7.1 Wiskundige verschillen

Zoals toegelicht en uiteengezet in de vorige paragrafen zijn het BJTR- en PIRFICT-model in de basis gelijk, en te vergelijken en in elkaar om te schrijven via de convolutievorm. Door de modellen te beschrijven vanuit hun gemeenschappelijke basis worden ook de verschillen tussen beide duidelijk, waarvan de belangrijkste zijn:

- 1 **gebruik van discrete versus continue tijd:** In discrete tijd is t in principe een eenvoudige index ($t \in \mathbb{N}$) die geen dimensie of eenheid heeft. Er bestaat in discrete tijd niet zoiets als halve of variërende tijdstap. Alle meetgegevens (van zowel de verklaarde als verklarende variabelen) moeten daarom dezelfde, vaste meetfrequentie hebben, of in een voorbereidingsstap daarnaartoe omgewerkt worden (zie paragraaf 3.4.2). De meetreeksen mogen daarbij in principe ook geen gaten of ontbrekende waarden vertonen, vanwege beperkingen in het ruismodel (die deels op te heffen zijn via inzet van een Kalman-filter, zie bijv. Bierkens e.a. (1999)). Het feit dat een modelleur de werkelijke tijdstap verschillend kan kiezen, zorgt er daarbij voor dat de modelparameters en de modelvorm frequentieafhankelijk zijn.

In continue tijd is de tijd t de DatumTijd van de meetgegevens in de database zelf ($t \in \mathbb{R}$). Metingen zijn dus gewoon bruikbaar zoals ze zijn, ook als ze onregelmatig zijn waargenomen, en hoeven niet voorberekt te worden (zie paragraaf 3.4.2). Er bestaat in beginsel ook niet zoiets als een 'gat' in een stijghoogtereeks, ook niet voor een ruismodel dat continu is (Von Asmuth en Bierkens, 2005). De wiskundige formulering daarvan zorgt echter wel voor een beperking van de modelvorm tot exponentiele functie of AR(1)-model. In continue tijd komt zijn de modelvergelijkingen en parameters daarnaast in principe onafhankelijk van de frequentie, omdat er geen frequentie voorkomt in de vergelijkingen;

- 2 **formulering van respons- of transferfuncties** – deze bestaan in het BJTR-model dus uit een combinatie van AR- en MA-parameters (zie paragraaf 4.2), terwijl het in het PIRFICT-model om continue wiskundige functies gaat (zie paragraaf 5.2.3). Bij die laatste kan het om volledig fysisch-deterministische, analytische functies gaan, in de meeste gevallen worden empirische functies zoals de gammaverdeling toegepast.

Hieronder wordt in meer detail ingegaan op de vergelijking en verschilpunten tussen beide modellen, als facetten of consequenties van bovengenoemde hoofdpunten.

2.7.2 Vergelijking van resultaten

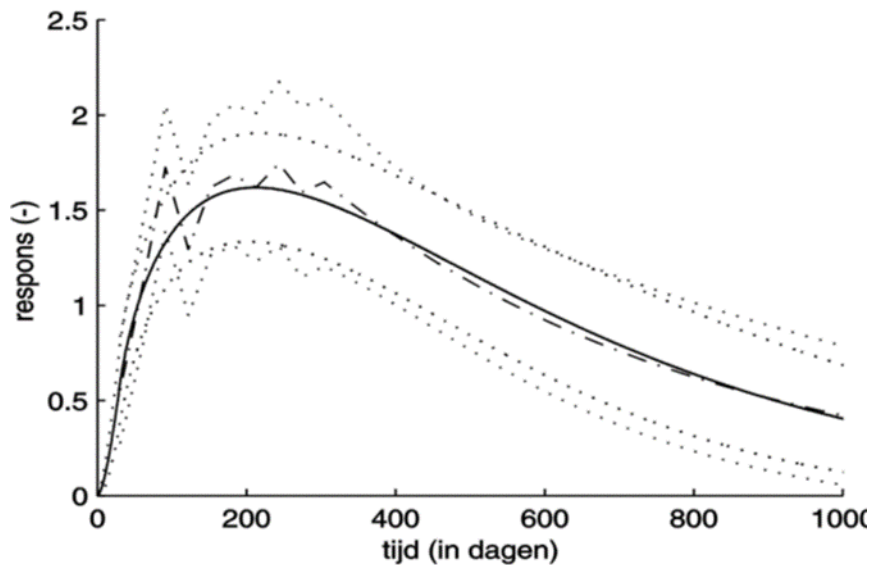
Een pragmatische manier om verschillende methoden te vergelijken is om de resultaten daarvan met elkaar te vergelijken. De kern van een continu tijdreeksmodel wordt daarbij gevormd door de impulsresponsfunctie. De impulsresponsfunctie beschrijft zoals gezegd de reactie van de grondwaterspiegel of stijghoogte op bijvoorbeeld een plotselinge impuls van neerslag. De impulsresponsfunctie is onveranderlijk in de tijd en karakteristiek voor een bepaald systeem en de locatie daarbinnen, zolang het systeem zelf niet verandert.

Bij het ontwikkelen en testen van het PIRFICT-model is een dergelijke vergelijking gemaakt voor 15 tijdstijghoogtereeksen uit het Noord-Hollands Duinreservaat (Von Asmuth e.a., 2002b). Afbeelding 2.15 toont daarvan de geschatte transferfuncties van buis 19AZW246_1, waarbij opvalt dat beide functies grotendeels overeenkomen, met als belangrijkste verschil dat de BJTR-functie deels onregelmatig is terwijl de PIRFICT-respons glad is. Ook de performance of fit van beide modellen verschilde niet veel van elkaar (qua RMSE, RMSI en bij kruisvalidatie), en dat voor alle onderzochte tijdstijghoogtereeksen. De verschillen die er wel waren waren daarbij in de richting van *overfitting* of toevalligheid van de onregelmatige vorm van de BJTR-functie.

Uit deze vergelijking van resultaten in praktijk komt dus niet zozeer een beste methode uit de bus. Een lastig punt bij het vergelijken van modellen is dat niet onomstotelijk en objectief vast te stellen is welke van beide de werkelijkheid het best benadert, zolang de werkelijkheid zelf niet bekend is. Het vergelijken van beide modellen o.b.v. een gesimuleerde werkelijkheid biedt in dit geval ook geen goede oplossing, omdat van een modelsimulatie niet aangetoond kan worden dat ze de werkelijkheid correct beschrijft. Een model dat een gesimuleerde reeks goed beschrijft, hoeft nog geen goede beschrijving van de werkelijkheid te geven. Het geeft alleen aan dat kalibratie goed is uitgevoerd. Conclusies die uit deze vergelijking wél te trekken zijn, zijn:

- 1 als beide methoden niet of nauwelijks verschillen qua resultaat, dan zijn andere verschilpunten leidend bij eventuele afweging van een keuze tussen beide (zie hoofdstuk 3);
- 2 het vergelijken van de resultaten zelf is hoe dan ook nuttig en levert synergie. Dit is ook als standaardpraktijk toepasbaar, bijv. als onderdeel van zogeheten *ensemble forecasts*. Verschillen die daarbij optreden wijzen op onzekerheden en gevoeligheden van de uitkomsten voor de modelkeuze.

Afbeelding 2.15 Vergelijking van transferfuncties voor buis 19AZW246_1, geschat m.b.v. dezelfde meetreeksen met maandelijkse frequentie. De doorgetrokken lijn is de blokrespons van het PIRFICT-model, de gestreepte de BJTR-functie, de gestippelde lijn hun 95 % betrouwbaarheidsinterval (bron: von Asmuth, 2002)



2.7.3 Voor- en nadelen van de BJTR en PIRFICT methode

BJTR-gebruikers noemen de volgende voor- en nadelen van het BJTR-model:

- 1 **minder aannamen** - Het model laat zoveel mogelijk de data spreken, zonder deze in een of ander conceptueel model van het systeem te plaatsen. Dit voorkomt dat de resultaten beïnvloed worden door hypothesen of voorstelling van het conceptueel model;
- 2 **minder hydrologische kennis nodig** - Het vorige punt betekent ook dat er geen of beperkt hydrologische kennis nodig is van het (geo)hydrologisch systeem om een model te kunnen maken. Dit maakt deze methode geschikt voor systemen waar minder systeemkennis van is;
- 3 **flexibiliteit** - Het model hanteert transferfuncties in ARMA-vorm, waarmee een breed scala aan dynamische relaties kan worden beschreven en die voldoende flexibiliteit bevatten om ook hydrologische relaties te kunnen beschrijven;
- 4 **discrete tijd** - BJTR-modellen zijn discreet en kunnen in principe alleen overweg met equidistante meetreeksen, maar de uitvoerreeks mag daarbij ook ontbrekende waarden omvatten. In geval van ontbrekende waarden wordt het BJTR-model overigens veelal ingebed in een Kalmanfilter (zie paragraaf 2.5 of bijv. [Bierkens et al., 1999 en Berendrecht et al., 2003]);
- 5 **uitgebreid ruismodel** - Het model heeft een uitgebreid ruismodel, dat wil zeggen dat het complexe autocorrelatiestructuren van de ruis kan beschrijven, dankzij de ARIMA- en SARIMA-functionaliteiten van het ruismodel (zie paragraaf 4.2.3). Dit biedt veel flexibiliteit om het modelresidu terug te brengen tot witte ruis.

Het nadeel van de BJTR-methode is dat het bewerkelijk is voor grote aantallen meetreeksen, omdat de meetreeksen niet in samenhang beschouwd worden.

PIRFICT-gebruikers noemen de volgende voor- en nadelen van het PIRFICT-model:

- 1 **voorbewerking van de stijghoogte reeks is niet nodig** - Stijghoogte meetreeksen, die als invoer en basis nodig zijn voor elk tijdreeksmodel, zijn in continue tijd in principe gewoon bruikbaar zoals ze zijn (na datakwaliteitscontrole (QC)). Voorbewerking kan daarbij gepaard gaan met data- en informatieverlies (zie paragraaf 3.5);
- 2 **het model is eenvoudiger te identificeren en automatiseren** - Het feit dat model en impulsresponsfunctie frequentie-onafhankelijk zijn (i.t.t. de transferfunctie, zie vergelijking (12)), beperkt de complexiteit ervan. Er hoeft geen frequentie-afhankelijke modelorde gekozen te worden zoals bij BJTR-modellen, waarmee het PIRFICT-model eenvoudiger te identificeren en automatiseren is;
- 3 **het model is beter hydrologisch te interpreteren** - Naast het feit dat frequentie-onafhankelijkheid ook bij de interpretatie een stukje complexiteit wegneemt, is het PIRFICT-model ook beter hydrologisch te interpreteren. Het fysisch-hydrologische gedrag van grondwatersystemen kan daarbij vooraf opgelegd worden aan het model bij het kiezen van de responsfunctie of de toegestane parameterwaarden daarvan (*predefined*, bijvoorbeeld als analytische oplossing), maar ook achteraf geïnterpreteerd worden bij het analyseren van de resultaten. Het vooraf opleggen van een bepaald gedrag, kan echter ook een nadeel zijn (zie hoofdstuk 6);
- 4 **het model is inzetbaar voor 'massaproductie'** – De combinatie van bovenstaande eigenschappen maakt het mogelijk om (zoals in paragraaf 2.7 aan de orde kwam) grote aantal meetreeksen simultaan te modelleren en interpreteren;
- 5 **beperkt ruismodel** – In continue tijd is het ruismodel beperkt tot exponentiele afname (equivalent met een AR(1)-model in discrete tijd (Von Asmuth en Bierkens, 2005). Een goede, wiskundige oplossing voor meer complexe, continue ruismodellen is er nog niet. Deze beperking kan een nadeel zijn.

De meerwaarde van het feit dat de stijghoogte invoer van het PIRFICT-model niet voorberekt hóéft te worden, is afhankelijk van het systeem en bijvoorbeeld in het geval van trage systemen niet zo groot. Ook bij toepassing van het PIRFICT-model wordt de invoer in praktijk vaak voorberekt, om de rekentijd te beperken en de berekening te vereenvoudigen. De voorbereking wordt daarbij echter doorgaans volledig geautomatiseerd, en kan zich daarmee grotendeels aan de aandacht van de modelleur onttrekken.

2.8 Software

In deze paragraaf wordt een kort overzicht gegeven van de veel gebruikte software voor het uitvoeren van tijdreeksanalyse. Voor een uitgebreide beschrijving van de beschikbare software pakketten wordt verwezen naar de individuele documentatie van de software. De voorbeelden en casussen in deze handleiding maken gebruik van deze software pakketten.

2.8.1 Menyanthes/Hydromonitor

Menyanthes is een softwarepakket voor het toepassen van tijdreeksanalyse ontwikkeld door KWR (Asmuth, 2002)¹, en is in ontwikkeling sinds 2000. Binnen Menyanthes kan zowel gebruikt gemaakt worden van de PIRFICT-methode en de Box-Jenkins methode voor het uitvoeren van tijdreeksanalyse. Het software pakket bevat methoden om datareeksen uit verschillende bronnen direct in te kunnen lezen, onder andere KNMI, DINOloket en InSitu gegevens. Sinds begin 2020 gaat Menyanthes door onder de naam Hydromonitor².

¹ <https://www.kwrwater.nl/tools-producten/menyantes/>

² <https://www.kwrwater.nl/tools-producten/hydromonitor/>

2.8.2 Tijdreeksanalist

Tijdreeksanalist is een Windows-applicatie om tijdreeksen te modelleren en analyseren¹. Hierin kunnen de verschillende stappen van het opstellen van een tijdreeksmodel worden doorlopen. Tijdreeksanalist is ontwikkeld door Eit van der Meulen en Paul Baggelaar, om het programma te gebruiken is een licentie nodig. In de kader van deze handleiding is een oefenversie beschikbaar gesteld van de software. Het programma bevat mogelijkheden om een tijdreeksmodel op te zetten met zowel de BJTR-methode als de PIRFICT methode. Het doel hierbij is om een transparante presentatie te geven (dus geen black box).

2.8.3 Pastas

Pastas is een software pakket ontwikkeld in python3 door de Technische Universiteit Delft, Artesia en de Universiteit Graz (Collenteur et al. 2019). De software is gratis beschikbaar en kan binnen elke python omgeving worden gebruikt. De software kan worden gecombineerd met andere python toepassingen voor de analyses. De code voor Pastas is openbaar op Github (<https://github.com/pastas/pastas>) en gebruikers kunnen bijdrage aan de ontwikkeling. Pastas maakt gebruik van de PIRFICT-methode voor het opstellen van tijdreeksanalyse. Hierbij heeft de gebruiker inzicht in elke stap door de modulaire opzet van de software. Er zijn verschillende voorbeelden beschikbaar om met Pastas aan de slag te gaan. De modellering in Pastas wordt gedaan met behulp van scripts en tijdreeksanalyse studies zijn daardoor beter reproduceerbaar.

¹ <https://www.amo-nl.com/software/tijdreeksanalist-trg/>

3

VOORBEREIDEN TIJDREEKSANALYSE

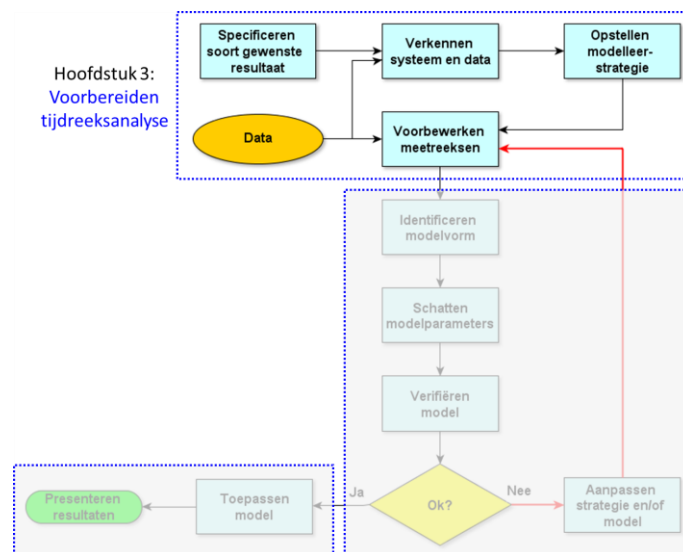
3.1 Inleiding

Dit hoofdstuk beschrijft de fase van het voorbereiden van de tijdreeksanalyse. Deze omvat vier stappen:

- specificeren soort gewenste resultaat (paragraaf 3.2);
- verkennen systeem en data (paragraaf 3.3);
- opstellen modellerstrategie (paragraaf 3.4);
- voorbereiden meetreeksen (paragraaf 3.5).

De eerste drie stappen zijn voor beide methoden van toepassing. Het bespreken van de laatste stap (paragraaf 3.5) maakt onderscheid tussen de BJTR-methode en de PIRFICT-methode.

Afbeelding 3.1 Stroomschema voor het uitvoeren van tijdreeksanalyse, met uitgelicht het voorbereiden van de tijdreeksanalyse

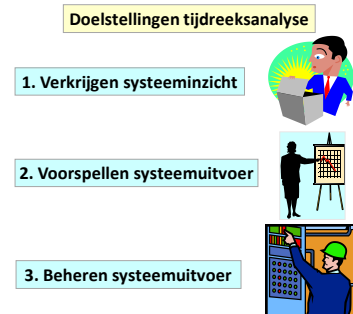


3.2 Specificeren soort gewenste resultaat

Start de voorbereiding van de tijdreeksanalyse met te specificeren welk soort resultaat de analyse op moet leveren. Dat is namelijk nodig om in de tweede stap richting te geven aan het verkennen van het systeem en de data en om in de derde stap een effectieve modelleerstrategie op te kunnen stellen.

Het soort gewenste resultaat hangt uiteraard af van de doelstelling van de tijdreeksanalyse. Daar zijn op hoofdlijnen drie categorieën van te onderscheiden:

1. verkrijgen van inzicht in de werking van een systeem, zoals het vaststellen van de relatie tussen de grondwaterstand en het neerslagoverschot,
2. voorspellen van de uitvoer van een systeem, zoals het voorspellen van de grondwaterstand bij een bepaald theoretisch verloop van het neerslagoverschot,
3. beheren van een systeem, zoals het zodanig begrenzen van een grondwateronttrekking dat de grondwaterstand niet te laag wordt.



Trendanalyse

Trendanalyse is ook een vorm van tijdreeksanalyse en valt onder doelstelling 1 (Verrijgen van inzicht in de werking van een systeem). Daarbij wordt echter geen gebruik gemaakt van een model dat relaties met andere hydrologische variabelen beschrijft, maar van een model dat de relatie met de tijd beschrijft. Een overzicht van trendanalysemethoden en een bijpassend keuzeschema worden gegeven in [Baggelaar en van der Meulen, 2012].

Voor wat betreft het soort gewenste resultaat is het ook belangrijk of de onzekerheid van de resultaten moet worden meegegeven. Dat kan namelijk een andere modelleerstrategie en ook een uitgebreidere modelverificatie vergen dan wanneer dat niet het geval is (zie paragraaf 3.4.1 en paragraaf 4.4). Het is overigens sterk aan te bevelen om altijd de onzekerheid van de resultaten mee te geven, aangezien daarmee hun zeggingskracht direct inzichtelijk wordt. Het zal in staat stellen tot een beter onderbouwde risicoafweging en besluitvorming te komen op basis van de resultaten.¹

¹ Bij hydrologisch modelleren wordt niet altijd de onzekerheid van de resultaten meegegeven. Daar worden door modelleers verschillende argumenten voor aangehaald. [Pappenberger and Beven, 2006] konden bij een kritisch onderzoek hiernaar zeven categorieën van dergelijke argumenten onderscheiden.

Voorbeelden van soort gewenste resultaat

Enkele voorbeelden van het soort gewenste resultaat van een tijdreeksanalyse zijn:

- A. een schatting van de evenwichtsrelatie van de grondwaterstand met de naburige grondwaterwinning¹;
- B. idem, maar dan aangevuld met het 95%-betrouwbaarheidsinterval van die schatting;
- C. het resultaat van een statistische toets op een relatie tussen die grondwaterstand en de naburige grondwaterwinning;
- D. een voorspelling van het verloop van de grondwaterstand bij een bepaald scenario van het verloop van neerslag en verdamping;
- E. idem, maar dan aangevuld met het 95%-voorspelinterval van het verloop;
- F. een schatting van de evenwichtsrelatie van de grondwaterstand met een bepaalde antiverdrogingsmaatregel;
- G. idem, maar dan aangevuld met het 95%-betrouwbaarheidsinterval van die schatting;
- H. het resultaat van een statistische toets op een relatie tussen die grondwaterstand en de antiverdrogingsmaatregel.

Merk op dat de voorbeelden met **blauwe tekst**, aangegeven bij B, C, E, G en H, ook een kwantificering vergen van de onzekerheid van het resultaat, uitgedrukt in het 95 %-betrouwbaarheidsinterval (B en G), het 95 %-voorspelinterval (E), of het resultaat van een statistische toets (C en H). In deze Handleiding is op diverse plaatsen toegelicht hoe die onzekerheden kunnen worden gekwantificeerd. Voor de betreffende tijdreeksmodellen gelden dan meer verificatiecriteria en de modelleerstrategie zal daar dan ook specifiek op afgestemd moeten worden.

3.3 Verkennen systeem en data

Verken na het specificeren van het soort gewenste resultaat, het te modelleren systeem en de daarvan beschikbare data, zodanig dat op basis daarvan een effectieve modelleerstrategie kan worden opgesteld. Met 'systeem' bedoelen we hier een afgebakend deel van de werkelijkheid dat de dynamische relatie bepaalt tussen zijn invoersignalen en het uitvoersignaal.

3.3.1 Verkennen systeem

Richt de systeemverkenning in eerste instantie op de eigenschappen van het systeem en op het identificeren, verzamelen en beoordelen van de variabelen waar de te modelleren tijdreeks aan gerelateerd kan zijn. Die laatste komen namelijk in aanmerking als invoervariabele bij het modelleren, mits daar geschikte meetreeksen van beschikbaar zijn, of beredeneerd kunnen worden samengesteld (zie paragraaf 3.4.5).

Voorbeelden van invoervariabelen en te verzamelen gegevens

Als bijvoorbeeld een grondwaterstandreeks moet worden gemodelleerd, komen als invoervariabelen in aanmerking neerslag, verdamping, grondwaterwinning of -infiltratie, oppervlaktewaterafvoer en -peil, peilbeheer en waterhuishoudkundige maatregelen. Metadata en meetdata van dergelijke signalen kunnen bijvoorbeeld worden opgevraagd bij het KNMI, het drinkwaterbedrijf, de provincie en het waterschap. De meteorologische variabelen zullen vrijwel altijd in aanmerking komen als invoervariabelen. Om te kunnen bepalen welke van de niet-meteorologische variabelen als invoervariabelen in aanmerking komen, moeten hun invloedstralen worden geraamd, op basis van hydrologisch inzicht en de variatie van de betreffende variabele in de beoogde modelperiode. Het is raadzaam daarbij een ruime veiligheidsmarge aan te houden. Waar het grote winningen betreft, zoals voor de openbare drinkwatervoorziening, of voor industriële grootgebruikers, is er vaak sprake van meerdere pompputten. Om na te kunnen gaan of en hoe daar bij de tijdreeksanalyse rekening mee moet worden gehouden, is het aan te bevelen de metadata en meetdata van alle pompputten op te vragen. Zo kan bijvoorbeeld worden vastgesteld of er sprake is geweest van een verplaatsing van het zwaartepunt van de winning (deze kan de grondwaterstand beïnvloeden), of dat er wellicht pompputten zijn die min of meer hetzelfde pomppregiem vertonen (hun onttrekkingen kunnen dan wellicht geclusterd worden tot één invoervariabele).

¹ De evenwichtsrelatie van een grondwaterstand met een winning is de uiteindelijke stationaire verandering van de grondwaterstand die zal optreden als de winning een stapverandering van één eenheid ondergaat (bijvoorbeeld uitgedrukt als cm/miljoen m³/jaar). Dit kengetal van de systeemdynamiek wordt in het Engels aangeduid als *steady state gain*.

Ook de statische eigenschappen van het systeem zijn van belang

Richt de systeemverkenning ook op de statische eigenschappen die de vorm en omvang van de dynamische relatie tussen de invoervariabelen en de uitvoervariabele bepalen. In een hydrologisch systeem kunnen die eigenschappen van velerlei aard zijn, zoals topografisch, geografisch, geologisch, hydrologisch en waterhuishoudkundig. Met deze kennis kan al vóóraf voor elke invoervariabele een eerste theoretische inschatting worden gemaakt van zijn dynamische relatie met de uitvoervariabele. Dat betreft dan een inschatting van de vorm van de relatie (uitgedrukt in een transferfunctie of impulsresponsfunctie) en een inschatting van de omvang van de relatie (uitgedrukt in de evenwichtsrelatie). Die inschatting komt in ieder geval van pas als de vorm van de relatie moet worden bepaald op basis van theoretische kennis (zie paragraaf 4.4.2). Zo kan bijvoorbeeld voor een systeem met een diepe grondwaterspiegel worden verwacht dat het effect van de neerslag op de grondwaterstand sterk uitgesmeerd zal zijn over de tijd. Deze kennis kan van pas komen bij het kiezen van de mate van aggregatie of uitdunning van de te modelleren grondwaterstandreeks.

Ga bij de systeemverkenning ook na of er sprake kan zijn van (drempel)niet-lineariteit¹, zoals bijvoorbeeld door 's zomers droogvallende waterlopen (zie ook paragraaf 9.2 voor voorbeelden en uitwerking hiervan). Dan kan daarmee rekening worden gehouden bij het kiezen van het soort tijdreeksmodel.

3.3.2 Verkennen en beoordelen data

Beoordeel na de systeemverkenning de data die beschikbaar zijn van de te modelleren uitvoervariabele(n) en de bijbehorende invoervariabele(n). Het gaat dan vooral om het beoordelen van de kwaliteit van de tijdreeksen. Dit kan worden ondersteund met:

- 1 visueel beoordelen tijdreeksen;
- 2 beoordelen loggermetingen.

Ad 1

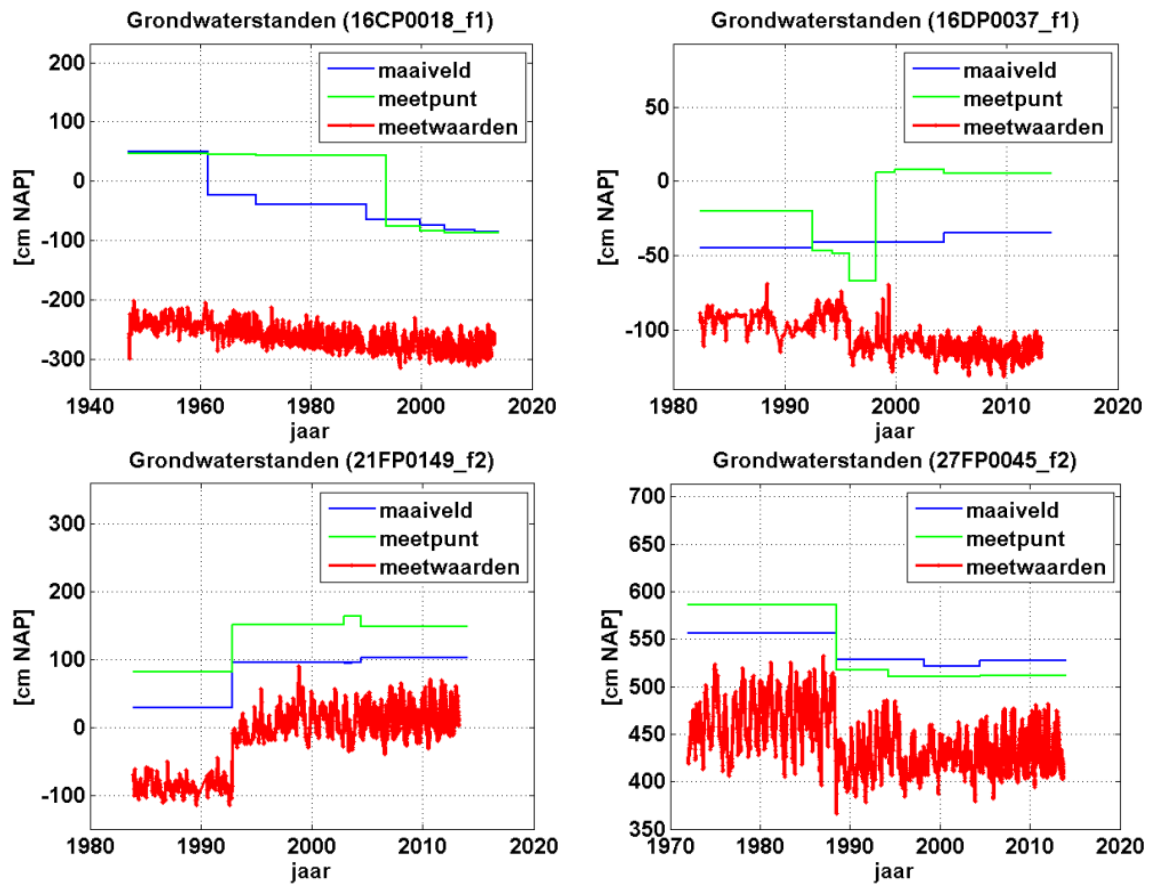
Richt het visueel beoordelen van de tijdreeksen op het detecteren van meet- en/of transcriptiefouten. Deze kunnen naar voren komen als uitbijters –dit zijn meetresultaten die niet in het normale beeld passen- of een vreemd verloop van de tijdreeks. Gebruik daarbij hydrologische kennis en ervaring en ga na of de anomalie kan zijn veroorzaakt door een lokale hydrologische ingreep. Verwijder een anomalie alleen als het voldoende aannemelijk is dat deze een meet- en/of transcriptiefout betreft. Zo kan er bij handmatig meten bijvoorbeeld sprake zijn van een zogenaamde meterfout², die zondig kan worden gecorrigeerd (zie ook paragraaf 9.1 voor voorbeelden met uitwerking).

Zet in de plot ook het verloop van relevante metadata uit. Bij een grondwaterstandreeks zijn dit bijvoorbeeld maaiveldhoogte, meetpunthoogte, diepte bovenkant en onderkant filter en X- en Y-coördinaten. Als een sprong van de reeks gepaard gaat met een sprong van één of meer van deze metadata is dat een sterke aanwijzing dat de reeks twee onvergelykbare situaties beschrijft, of dat een wijziging van maaiveld- of meetpunthoogte niet is verwerkt. Zo is in de voorbeelden van onderstaande Afbeelding 3.2 waarschijnlijk sprake van invloed van niet verwerkte wijzigingen van maaiveld- en/of meetpunthoogte op de gemeten grondwaterstand.

¹ Een systeem is niet-lineair als de reactie van de uitvoervariabele op een invoervariabele geen lineaire functie is van de instelling van die laatste.

² Een meterfout is een afwijking van exact een meter of een veelvoud daarvan.

Afbeelding 3.2 Voorbeelden van grondwaterstandreeksen die waarschijnlijk zijn beïnvloed door niet of slecht verwerkte wijzigingen van metadata (hier maaiveld- en/of meetpunthoogte)



Ad 2

In Nederland is vanaf circa 2004 in toenemende mate het handmatig meten van de grondwaterstand vervangen door het automatisch meten met loggers. Deze meetinstrumenten zijn echter niet zonder problemen gebleken en kunnen na verloop van tijd minder nauwkeurig worden en/of drift¹ gaan vertonen. Zonder geschikte controle- en correctieprocedure kunnen deze fouten dan ook relevante vertekeningen geven van het verloop van de grondwaterstand.

Een geschikte controle vergt een combinatie van:

- 1 frequente (zoals maandelijkse of zelfs tweewekelijkse) handmatige controlemetingen (in het veld) van de grondwaterstand en;
- 2 kennis van de kansverdeling van de meetfout van een goed functionerende logger, zoals bijvoorbeeld afgeleid uit de technische fabrieksspecificaties van de betreffende logger.

Alleen met deze combinatie kan met methoden uit de Statistische Procesbeheersing (zoals de Shewart-controlekaart) objectief worden vastgesteld wanneer het meetproces niet meer beheerst is. Ook Alterra adviseerde in dit kader al eerder voldoende frequent - zelfs tweewekelijks - handmatige controlemetingen in het veld uit te voeren om de fouten binnen de perken te kunnen houden [Ritzema et al., 2012]. Maar doordat er steeds meer sprake is van volledig automatisch meten, waarbij elke meting via telemetrie wordt doorgestuurd, is er momenteel juist een tegenovergestelde tendens naar minder veldbezoeken. Een voordeel van de telemetrie is nog wel dat het presteren van de logger online kan worden gevolgd en er dan direct een herstellende actie kan worden ondernomen als blijkt dat de logger niet of slecht functioneert. Het slecht functioneren kan bijvoorbeeld ook blijken uit niet-realistische metingen van hulpvariabelen zoals watertemperatuur, luchtdruk en waterdruk.

¹ Onder drift verstaan we hier een geleidelijk opbouwende systematische fout van de logger bij het bepalen van de grondwaterstand. Het komt neer op het verlopen van het nulpunt.

Een geleidelijk opbouwende en niet-gedetectedeerde systematische fout van een logger kan er bij tijdreeksanalyse toe leiden dat:

- de ruis onnodig wordt vergroot, óf dat;
- de schatting van de relatie van de grondwaterstand met een invoerreeks – zoals een winning of een ingreep – vertekend wordt, resulterend in onterecht wel/niet statistisch significant detecteren van een relatie.

De ervaring heeft geleerd dat dergelijke vertekeningen makkelijk kunnen optreden en daarom is grote zorgvuldigheid geboden. Tijdreeksanalyse kan een gevoelig instrument zijn, maar daardoor kan het ook makkelijker de plank mislaan bij een slecht gemeten grondwaterstandreeks.

Als eerste controle is aan te bevelen om in de plot van een tijdreeks met loggermetingen tevens de handmatige controlemetingen uit te zetten, die zijn gedaan bij het uitlezen van de logger. Dit stelt in staat de kwaliteit van de loggermetingen visueel te beoordelen. Ga verder na of er een procedure voor controle en correctie is gehanteerd, die voldoende solide is om de gevalideerde meetreeks geschikt te achten voor tijdreeksanalyse. Als bijvoorbeeld een met een logger verkregen tijdreeks van de grondwaterstand alleen is gecontroleerd en gecorrigeerd door deze te vergelijken met zijn tijdreeksmodel, met als invoerreeksen neerslag en verdamping, dan ontbreekt een onafhankelijke en objectieve beoordeling (zoals ten opzichte van handmetingen). Er is dan een reëel risico dat de gevalideerde tijdreeks slechts een conceptueel model van de werkelijkheid nabootst.

Als uit het beoordelen van de kwaliteit van de loggermetingen blijkt dat deze ongeschikt zijn voor tijdreeksanalyse, ga ze dan zo mogelijk zelf controleren en corrigeren. Daarvoor zijn verschillende methoden en tools beschikbaar. Zo ontwikkelde Alterra in samenwerking met Eijkelpark Soil & Water een validatietool voor hoogfrequente grondwaterstandreeksen. Deze stelt in staat grote hoeveelheden metingen automatisch en efficiënt te controleren en ook overzichtelijk te rapporteren over de kwaliteit van de waarnemingen [Walvoort et al, 2013].¹ De tool is gebaseerd op drie uitgangspunten:

- 1 metingen van slechte of matige kwaliteit moeten als zodanig worden herkend en beoordeeld, maar daarbij mogen geen metingen van goede kwaliteit verloren gaan;
- 2 de controles moeten toegepast kunnen worden op reeksen met uiteenlopende eigenschappen. Daarom worden robuuste statistische methoden gebruikt;
- 3 de resultaten van de controles moeten op een intuïtief niveau kunnen worden geïnterpreteerd. Daartoe wordt gebruik gemaakt van *fuzzy-logic*.

De tool maakt geen gebruik van tijdreeksmodellen voor de relatie tussen neerslagoverschot en grondwaterstand, om te voorkomen dat het resultaat van de kwaliteitscontrole gaat afhangen van veronderstellingen over die relatie.

3.4 Opstellen modelleerstrategie

Het opstellen van de modelleerstrategie start met het specificeren van de uitvoervariabele (de te modelleren variabele) en de invoervariabele(n). De uitvoervariabele zal doorgaans al vastliggen door de vraagstelling. Als invoervariabelen komen in aanmerking de variabelen waar de uitvoervariabele volgens de systeemverkenning aan gerelateerd kan zijn.

¹ Deze maakt deel uit van het Smart Sensing-systeem van Eijkelpark.

3.4.1 Keuze soort tijdreeksmodel

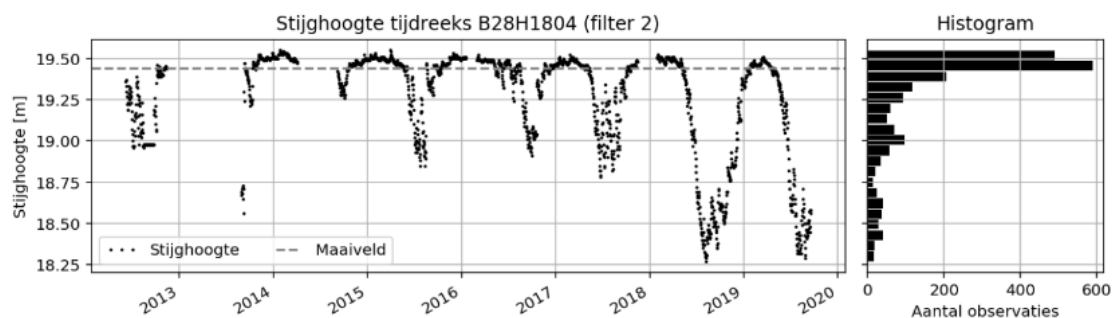
Er zijn voor tijdreeksanalyse meerdere soorten modellen beschikbaar, elk met specifieke kenmerken en met sterke en zwakke punten. Voor wat betreft hydrologische toepassingen zijn de bekendste het BJTR-model en het PIRFICT-model. Voor elk van deze zijn de sterke punten beschreven in paragraaf 2.7. Het soort tijdreeksmodel dat het meest geschikt is wordt bepaald door onder andere het soort gewenste resultaat, de meetfrequentie, het al of niet voldoen aan equidistantie van de beschikbare tijdreeksen en de beschikbare kennis van het te modelleren systeem.

Rekening houden met niet-lineaire relaties

Als een grondwatersysteem niet lineair reageert op de invoerreeksen vraagt dit van de tijdreeksmodelleur een andere aanpak. Het systeem zal immers afhankelijk van de toestand van het systeem anders reageren op invloeden. Voorbeelden zijn dat als de grondwaterstand aan maaiveld komt deze niet verder kan stijgen bij neerslag, of dat een zeer lage grondwaterstand minder sterk reageert op verdamping.

Het is aan de tijdreeksmodelleur om deze niet-lineariteit te herkennen en de modelstrategie hierop aan te passen. Afbeelding 3.3 toont een voorbeeld van een reeks die duidelijk niet-lineair gedrag vertoont (zie ook paragraaf 9.2.6), in dit geval drempel niet-lineariteit. Dit houdt in dat de grondwaterstand anders reageert op bijvoorbeeld neerslag en verdamping boven een bepaald grenswaarde (of drempel). Niet voor alle meetreeksen geldt dat niet-lineariteit vóóraf duidelijk vast te stellen is. Het testen met verschillende modelstructuren kan hierin soms beter inzicht geven dan visuele inspectie van de uitvoerreeks. Als een lineair tijdreeksmodel relevant minder presteert dan een niet-lineair model wijst dat op niet-lineariteit.

Afbeelding 3.3 Voorbeeld van drempel niet-lineariteit (zie paragraaf 9.2)



Het klassieke BJTR- of PIRFICT-model kan alleen lineaire relaties beschrijven. Mogelijkheden om drempel-niet-lineariteit te verdisconteren worden geboden door:

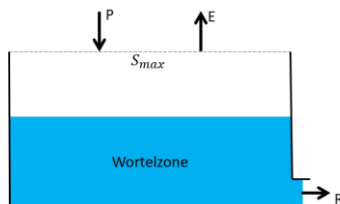
- een aangepaste vorm van het ARMA-model, aangeduid als het TARSO-model¹ [Knotters and De Gooijer, 1999];
- een niet-lineair Kalmanfilter [Berendrecht et al., 2004];
- een niet-lineaire transferfunctie [Collenteur et al., 2020].

Bij het TARSO-model verandert de respons op neerslag en verdamping wanneer de grondwaterstand boven of onder een bepaalde drempelwaarde komt. Dit kan bijvoorbeeld optreden als de grondwaterstand dicht nabij maaiveld staat, of bijvoorbeeld bij sloten die alleen bij een hoge grondwaterstand draineren. Bij het opstellen van een tijdreeksmodel moet worden gekeken of de prestaties verbeteren met het toevoegen van drempel niet-lineariteit.

¹ Dit model hanteert drempelwaarden van de grondwaterstand waar de reactie op het neerslagoverschot verandert, bijvoorbeeld door bodemverschillen of door verschillen tussen drainageniveaus.

Andere voorbeelden zijn uitwerkingen van het zogenaamde bakjes-model, dat veel in de neerslag-afvoermodellering wordt gebruikt (voor een voorbeeld met uitwerking zie paragraaf 9.2.5). Afbeelding 3.4 toont een voorbeeld van een bakjes-model met één bak voor de wortelzone. Bij dit voorbeeld wordt de grondwateraanvulling R niet alleen bepaald door de neerslag (P) en verdamping (E) maar ook door de berging in de wortelzone. Dit model kan bijvoorbeeld worden uitgebreid met een interceptiebak. Door het toevoegen van dit bakjes-model worden extra parameters aan het tijdreeksmodel toegevoegd, die vervolgens bij het schatten van de modelparameters worden geoptimaliseerd. Bij het voorbeeld van Afbeelding 3.4 worden twee parameters (berging in wortelzone en afvoerfactor) aan het tijdreeksmodel toegevoegd, om de massabalans van het bakjes-model te simuleren.

Afbeelding 3.4 Een bakjes-model voor de wortelzone



3.4.2 Keuze tijdseenheid

In het verleden werden grondwaterstanden meestal handmatig uitgelezen, op of rond de 14^e en de 28^e van elke maand, wat neerkwam op ongeveer halfmaandelijks meten. Maar door de introductie van automatische loggers vanaf ongeveer 2004 zijn er in toenemende mate dagelijkse, uurlijkse, of soms zelfs nog hoogfrequentere metingen beschikbaar. Afhankelijk van de modelperiode kan een grondwaterstandreeks dan verschillende meetintervallen hebben.

BJTR-methode

Voor het BJTR-model moeten alle betrokken reeksen (uitvoer en invoer) equidistante zijn, of worden gemaakt. De invoerreeksen mogen een kleiner meetinterval hebben dan de uitvoerreeks, maar ze mogen geen ontbrekende waarden hebben. Voor de keuze van de tijdseenheid bij modelleren met een BJTR-model is onder meer van belang de minimale tijdseenheid (zoals dag, week, maand, et cetera) waarvoor een equidistante tijdreeks van de uitvoervariabele beschikbaar is, die voldoende lang is en ook de beoogde modelperiode omvat. Verder is het zaak zo goed mogelijk aan te sluiten op de dynamiek van de verschillende processen die de uitvoervariabele beïnvloeden. Daarbij is met name van belang welke responstijden optreden, oftewel hoe lang het duurt voordat de invloed van een impuls van een invloedvariabele op de uitvoervariabele is uitgewerkt. De responstijden kunnen bijvoorbeeld empirisch worden ingeschat met kruiscorrelogrammen (zie paragraaf 4.2.1). Naarmate het te modelleren proces trager is, past een grotere tijdseenheid voor het modelleren. Het is dan aan te bevelen om elke reeks per tijdseenheid te aggregeren tot een gemiddelde of som, omdat er anders veel modelparameters nodig kunnen zijn.

In de praktijk worden grondwaterstandreeksen met het BJTR-model vaak gemodelleerd met zowel de uitvoerreeks als de invoerreeks(en) op week-, tweeweeks-, of maandbasis. Daar zijn verschillende argumenten voor:

- het sluit vaak beter aan op de lange responstijd van hydrologische systemen dan modelleren op dagbasis, wat het eenvoudiger maakt om transferfuncties empirisch te identificeren (zie paragraaf 4.2.1);

- bij modelleren op dag- of uurbasis is doorgaans sprake van een dermate sterke autocorrelatie¹ van de ruis (N_t) dat het moeilijk of onmogelijk is die zodanig te beschrijven in het ruismodel dat het modelresidu (a_t) witte ruis vormt. Er kan dan niet worden voldaan aan een belangrijke randvoorwaarde om tevens de onzekerheden van modelresultaten mee te kunnen geven (zie daarvoor paragraaf 7.6);
- waar het gaat om de klassieke BJTR-methode is een aanvullend argument dat bij verhogen van de tijdsbasis doorgaans kan worden volstaan met minder modelparameters. Dit sluit aan op het spaarzaamheidsprincipe². Maar door een aanpassing kan de BJTR-methode inmiddels ook voor gedefinieerde functies toepassen - zoals de Betafunctie, de Gammafunctie en de exponentiële functie - om zo het aantal benodigde modelparameters te reduceren [Van der Meulen en Baggelaar, 2019].

Als het gaat om het modelleren van snellere processen, zoals de doorwerking van het getijde, ligt het meer voor de hand een kortere tijdsbasis te hanteren, zoals de uurbasis. Maar het kan dan niet mogelijk blijken het modelresidu (a_t) terug te brengen tot witte ruis, wat de toepassingsmogelijkheden van het model zal beperken.

PIRFICT-methode

Voor PIRFICT geldt dit niet, bij deze methode zijn equidistante tijdstappen geen vereiste. Het is daarom niet nodig om invoerreeksen met ongelijke tijdstappen om die reden voor te bewerken. Hierdoor is het ook mogelijk om reeksen te gebruiken met missende data. Bij hoogfrequente data wel kan worden gekeken of de gegevens passen bij het doel van de tijdreeksanalyse, indien gewenst kan de data worden geaggregeerd. Een voorbeeld waarbij hoogfrequente data gewenst kan zijn, is wanneer getijde een rol speelt. Bij bijvoorbeeld zeer trage grondwatersystemen kan data worden geaggregeerd tot een lagere frequentie. Hierbij is ook sprake van minder sterke autocorrelatie.

3.4.3 Keuze modelperiode

De modelperiode dient bij voorkeur zodanig te worden gekozen dat de kans op het verkrijgen van het soort gewenste resultaat wordt gemaximaliseerd. Verder dient de modelperiode in ieder geval minstens zo lang te zijn als de maximale responstijd van het systeem op relevante invloedsvariabelen (zie ook paragraaf 9.4) Afhankelijk van de bodemkundige en hydrologische situatie kan de responstijd van de grondwaterstand op het neerslagoverschot ergens tussen enkele weken en meer dan vijf jaar bedragen [Knotters en Bierkens, 1999].

Als het gaat om het vaststellen van de relatie van de grondwaterstand met het neerslagoverschot, dan is het in het algemeen raadzaam om een modelperiode van minimaal 3 à 4 jaar te hanteren, maar als het een diepe freatische grondwaterstand betreft kan het zelfs nodig zijn minimaal 5 à 6 jaar te hanteren.

Modelperiode voor bepalen grondwaterstandkarakteristieken

Voor het bepalen van grondwaterstandkarakteristieken (zoals GxG of grondwatertrap) is formeel geen tijdreeksmodellering nodig als de grondwaterstandreeks voldoende lang is en halfmaandelijks of frequenter is bemeaten. Die karakteristieken kunnen dan immers direct uit de metingen worden bepaald. Er zijn echter verschillende criteria in omloop voor wat 'voldoende lang' is, zoals 8 jaar, 10 jaar en 30 jaar. [Ritzema et al., 2012] adviseren de GxG te berekenen over een periode van 30 jaar, omdat de praktijk heeft uitgewezen dat bij kortere perioden de invloed van extra droge of natte jaren sterk kan doorwerken, wat een instabiel beeld verschaft van de GxG. Dat maakt die informatie minder bruikbaar bij planvorming. Dit speelt vooral als de fluctuatie van de grondwaterstand hoofdzakelijk wordt veroorzaakt door fluctuaties in het neerslagoverschot. Als de fluctuatie vooral wordt veroorzaakt door ontwatering kunnen ook kortere reeksen bruikbaar zijn.

¹ Er is sprake van autocorrelatie als opeenvolgende waarden van een tijdreeks niet onafhankelijk van elkaar zijn, maar aan elkaar gerelateerd zijn.

² Bij het ontwikkelen van een model moet bij voorkeur het spaarzaamheidsprincipe (*principle of parsimony*) in acht worden genomen. Dit houdt in dat moet worden gestreefd de tijdreeks voldoende goed te beschrijven met een model dat zo min mogelijk modelparameters heeft. Als er teveel modelparameters zijn, zullen die slecht zijn geschat, doordat de informatie aanwezig in de data dan over teveel parameters is verdeeld.

Als een grondwaterstandreeks volgens bovenstaande criteria te kort is, zal het nodig zijn deze eerst te verlengen op basis van de in een tijdreeksmodel vastgelegde relatie met het neerslagoverschot. Voor het afleiden van dat model dient dan wel een modelperiode van minstens 3 à 4 jaar te worden gehanteerd. Maar als het een diepe freatische grondwaterstand betreft kan het ook nodig zijn daarvoor minimaal 5 à 6 jaar te hanteren.

Let op dat in alle gevallen, dus ook bij het modelmatig verlengen van een grondwaterstandreeks, de GxG alleen moet worden bepaald uit de halfmaandelijke meetwaarden op of rond de 14^e en 28^e van elke maand, dus ook als het een reeks met hoogfrequenter metingen betreft. Deze restrictie komt door de historische definitie van de GxG. Een uitgewerkt voorbeeld voor het bepalen van deze grondwaterkarakteristieken is te zien in paragraaf 10.2.

Modelperiode voor vaststellen van dynamische relatie met winning of ingreep

Als het gaat om het vaststellen van een relatie tussen de grondwaterstand en een grondwaterwinning of waterhuishoudkundige ingreep, moet de modelperiode bij voorkeur zo worden gekozen dat in die periode:

- de grondwaterwinning of waterhuishoudkundige situatie een zo groot en abrupt mogelijke verandering ondergaat, met voldoende lang te achten perioden vóór en ná die verandering. Wat hier *voldoende* is te achten zal afhangen van het systeem en van de omvang van de verandering. In het algemeen is het raadzaam om de modelperiode zowel minstens 2 à 3 jaar vóór de verandering als ná de verandering te laten uitstrekken;
- de uitvoervariabele zo min mogelijk wordt beïnvloed door andere variabelen;
- de invoervariabelen onderling zo min mogelijk zijn gecorreleerd. Afhankelijk van de patronen van gecorreleerde invoervariabelen kunnen er perioden zijn waarbinnen ze minder gecorreleerd zijn.

Zie verder hoofdstuk 7 en paragraaf 9.2.8 voor uitgewerkte voorbeelden van casussen met deze toepassing van tijdreeksanalyse, voor zowel de BJTR-methode als de PIRFICT-methode.

Verdisconteer een opstartperiode

Als op basis van de systeemverkenning kan worden aangenomen dat er sprake is van een lange responstijd van de te modelleren reeks op één of meer van de invloedsvariabelen, streef er dan naar de invoerreeksen zodanig eerder te laten starten dat van elk de volledige responstijd is gepasseerd vóórdát de te modelleren reeks start. Anders gaat het opstarten ten koste van de lengte van de modelperiode.

3.4.4 Hoe meenemen neerslag en verdamping als invoervariabelen

Bij het modelleren van de grondwaterstand komen in eerste instantie neerslag en verdamping in aanmerking als invoervariabelen. Deze kunnen daarbij op verschillende manieren in het tijdreeksmodel worden opgenomen:

- 1 als potentieel neerslagoverschot, bepaald als $\text{neerslag} - f \cdot \text{verdamping}$, waarin f de verdampingsfactor is. Voor de Makkink-verdamping wordt deze factor doorgaans op 1 gesteld, maar het is mogelijk dat deze een seizoensmatige variatie vertoont;
- 2 apart, als twee invoerreeksen, met afzonderlijke transferfuncties. De verdampingsfactor wordt dan in feite meegeschat als een modelparameter;
- 3 semi-apart, als twee invoerreeksen, maar met exact dezelfde transferfunctie (qua vorm en modelparameters), waarbij tevens de verdampingsfactor f als een modelparameter wordt geschat.

De eerste manier, met als invoerreeks het potentieel neerslagoverschot is van oudsher toegepast bij modelleren met de BJTR-methode. De tweede manier, waarbij neerslag en verdamping aparte invoerreeksen vormen wordt niet vaak toegepast, omdat deze kan resulteren in afwijkende vormen van de transferfuncties voor neerslag en verdamping, wat het model dan minder aannemelijk maakt. De derde manier wordt veel toegepast bij modelleren met de PIRFICT-methode, maar kan ook worden toegepast bij de BJTR-methode. Deze leidt tot dezelfde transferfunctie voor neerslag en verdamping - die wordt immers opgelegd - en hanteert een verdampingsfactor die wordt meegeschat en daardoor in principe realistischer kan zijn dan een vóóraf aangenomen verdampingsfactor. Maar er kan desondanks soms toch een weinig realistische verdampingsfactor worden geschat, zoals met een waarde kleiner dan 0,5 of een waarde groter dan 2.

Een verdampingsfactor wordt nog realistisch geacht als deze tussen 0,5 en 2,0 ligt. Een uitgewerkte vergelijking van de verschillende methoden is te zien paragraaf 9.2 en de beoordeling is uitgewerkt in 9.4.

3.4.5 Omgaan met verandering van het systeem

Als het systeem in de beoogde modelperiode is veranderd van karakteristieken moet worden getracht het model daar op af te stemmen. Zo kan een hydrologisch systeem veranderen door bijvoorbeeld beekherstel, een antiverdrogingsmaatregel, of een verandering van het oppervlaktewaterpeil. Van dergelijke ingrepen zullen doorgaans geen tijdreeksen beschikbaar zijn, zodat moet worden gepoogd die beredeneerd samen te stellen. Daarvoor moet echter minstens bekend zijn wanneer de ingreep actief werd en moet theoretisch kunnen worden beredeneerd in welke vorm zijn relatie met de uitvoerreeks zich zal manifesteren, zoals een niveauperandering via een stap, lineaire helling, parabool of hyperbool. Een antiverdrogingsmaatregel met een duidelijke startdatum en een verondersteld direct effect kan dan worden vertegenwoordigd door een reeks met waarden 0 tot de startdatum en waarden 1 daarna. Dit wordt aangeduid als een interventiereeks. Mocht de daarvoor benodigde informatie of theoretische kennis ontbreken, ga dan niet over tot het zelf invullen van start en/of vorm van de relatie op basis van een visuele verkenning van de te modelleren uitvoerreeks. Er ontstaat dan namelijk een substantieel risico dat de interventiereeks wordt gekneed naar een toevallige verandering van de uitvoerreeks, wat neerkomt op curvefitting. Het afstemmen van een invoerreeks op visuele verkenning van de uitvoerreeks introduceert een sterk subjectieve lading en de modelresultaten hebben dan weinig zeggingskracht (zie ook paragraaf 5.2.4).

Als de verandering van het systeem niet-lineariteit in zijn reacties op invloedsfactoren kan hebben geïntroduceerd, moet het model daar ook rekening mee kunnen houden.

3.4.6 Omgaan met sterk gecorreleerde invoerreeksen

Als twee invoerreeksen onderling relevant positief of negatief zijn gecorreleerd (vuistregel: absolute correlatiecoëfficiënt van 0,3 of meer), is het niet mogelijk hun relaties met de uitvoerreeks onafhankelijk van elkaar te schatten. Zo kan het potentieel neerslagoverschot sterk gecorreleerd zijn aan een drinkwaterwinning die niet of nauwelijks een trendmatige verandering vertoont. In veel voorzieningsgebieden is namelijk het drinkwatergebruik in de zomer groter dan in de winter, waardoor de drinkwaterwinning een seizoenpatroon vertoont, net als het potentieel neerslagoverschot. Als in zo'n geval met tijdreeksanalyse de relatie van de grondwaterstand met het potentieel neerslagoverschot en/of de drinkwaterwinning wordt vastgesteld, moet er terdege rekening mee worden gehouden dat de resultaten vertekend kunnen zijn.

Maar als de tijdreeksanalyse alleen dient om de relatie van de grondwaterstand met een derde variabele vast te stellen (zoals een waterhuishoudkundige maatregel), dan zijn er de volgende mogelijkheden:

- 1 gebruik bij het modelleren slechts één van de gecorreleerde invoerreeksen, of;
- 2 gebruik bij het modelleren beide gecorreleerde invoerreeksen, zij het één daarvan na correctie voor de relatie met de eerste¹. Een dergelijke correctie is bijvoorbeeld uit te voeren met lineaire regressie. Schat hiertoe het volgende lineaire regressiemodel voor de relatie tussen de twee invoerreeksen (X_1 en X_2):

$$X_{2t} = b_0 + b_1 X_{1t} + e_t$$

waarin b_0 de schatting van het intercept, b_1 de schatting van de regressieparameter, e de modelfout en t de index. De voor de relatie met X_1 gecorrigeerde reeks van X_2 volgt dan uit:

$$X_{2t}^* = X_{2t} - (b_0 + b_1 X_{1t}) = e_t$$

waarin X_{2t}^* de voor de relatie met X_1 gecorrigeerde reeks van X_2 .

¹ Deze voorbewerking wordt in de statistiek aangeduid als het *orthogonaliseren* van twee tijdreeksen.

Voorbeeld van modelleerstrategie

Een modelleerstrategie dient te worden afgestemd op de vraagstelling en de betreffende hydrologische situatie, zoals hier met een voorbeeld toegelicht. Stel dat moet worden nagegaan in welke mate een grondwateronttrekking, die is verspreid over een groot aantal wingroepen, de grondwaterstand in de omgeving heeft beïnvloed. De meeste van deze winningen zijn onderling echter zeer sterk gecorreleerd, waardoor het niet mogelijk is om van elk afzonderlijk de invloed op de grondwaterstand te reconstrueren. En het sommeren van alle winningsreeksen tot één winningsreeks zal tot een sterke vertekening leiden, omdat het model dan geen rekening meer kan houden met het feit dat de winningen op verschillende afstanden liggen van het betreffende peilfilter.

Een oplossing van dit probleem is het toepassen van geohydrologische weging, door voor elke te analyseren grondwaterstandreeks alle winningen gewogen te sommeren tot één winnings-reeks, die vervolgens als een invoerreeks wordt gehanteerd bij het statistisch modelleren van die grondwaterstandreeks. Als gewicht voor een afzonderlijke winning kan bijvoorbeeld de met een grondwatermodel berekende stationaire verlaging van de betreffende grondwaterstand door (een eenheidshoeveelheid van) die winning worden gehanteerd. Het gewicht kan ook worden afgeleid met de formule van Hantush–Jacob voor een gedeeltelijk afgesloten watervoerend pakket. Deze strategie is bijvoorbeeld toegepast door [Rolf en Lebbink, 1998] en [Van Geer en Lourens, 2001] en leverde fysisch-hydrologisch plausibele en bruikbare resultaten. Als kanttkening past wel dat deze strategie enigszins gekneed is naar het conceptuele beeld van het beschouwde geohydrologische systeem.

3.4.7 Omgaan met onverklaarbaar patroon van de uitvoerreeks

Soms blijkt een te modelleren uitvoerreeks een niet-stationair¹ patroon (zoals een trend) te vertonen die bij geen van de beschikbare invoerreeksen valt te bespeuren en waar ook geen directe verklaring voor te vinden is. Dit is een aanwijzing - eigenlijk op te vatten als waarschuwing - dat er onvoldoende bekend is over het te modelleren systeem. Ga hier niet over tot het introduceren van een nieuwe invoerreeks met de vorm van het betreffende niet-stationaire patroon. Dat zou het modelleren van het systeem reduceren tot curvefitting. Net zoals hiervoor al vermeld bij paragraaf 3.4.5 geeft dat het model een sterk subjectieve lading en de resultaten hebben dan weinig zeggingskracht.

Een objectievere manier om een niet-stationair patroon mee te kunnen nemen in het model is het te verdisconteren in het ruismodel, middels een differentie (zie daarvoor paragraaf 4.4.2). Het moge echter evident zijn dat het tijdreeksmodel dan niet veel voorspelkracht kan hebben, omdat de dynamiek van het betreffende systeem onvoldoende wordt beschreven met het deterministische deel van het model.

Een niet-stationair patroon kan het overigens ook lastig of onmogelijk maken de vorm van de transferfunctie met een invoerreeks empirisch af te leiden, in welk geval een vorm zal moeten worden aangenomen op basis van theoretische kennis (zie paragraaf 4.2.2).

3.5 Vorbewerken meetreeksen

Bij het opstellen van de modelleerstrategie zijn de bij de tijdreeksanalyse te betrekken variabelen geselecteerd (uitvoervariabele en invoervariabelen). Nu moeten hun meetreeksen worden voorberekt om ze gereed te maken voor de tijdreeksanalyse. De wijze van voorbereken kan afhangen van de gekozen methode van tijdreeksanalyse, zodat dit onderdeel in twee subparagrafen wordt toegelicht, namelijk:

- 1 voorbereken bij een BJTR-model (zie paragraaf 7.4);
- 2 voorbereken bij een PIRFICT-model (zie paragraaf 9.1).

¹ We noemen een reeks niet-stationair als het proces dat deze reeks genereert niet-stationair is, in de zin dat zijn gemiddelde en/of variantie tijdsafhankelijk zijn.

3.5.1 Voorbewerken bij een BJTR-model

Het BJTR-model gaat uit van equidistante tijdreeksen en wordt vaak toegepast op week-, tweeweeks- of maandbasis, omdat daarmee doorgaans beter wordt aangesloten op de tijdschaal van de dynamiek van de verschillende processen die spelen in het te modelleren systeem. Het toepassen op dagbasis is minder gebruikelijk, niet alleen omdat dit bij niet snel reagerende systemen veel modelparameters kan vergen, maar ook omdat het op dagbasis lastig of onmogelijk kan zijn te voldoen aan de randvoorwaarde dat het modelresidu (a_t) witte ruis vormt.

Om te kunnen aansluiten op de dynamiek van de systeemprocessen kan het nodig zijn tijdreeksen te aggregeren, of uit te dunnen. Als het een bijvoorbeeld een reeks (of deel daarvan) betreft met halfmaandelijke metingen - doorgaans op of rond de 14^e en de 28^e - kan die worden omgezet tot een reeks maandgemiddelde grondwaterstanden, door per maand de meetwaarden te middelen. Het verlies aan informatie over het verloop van de grondwaterstand door het middelen van halfmaandelijke meetwaarden naar maandwaarden zal meestal gering zijn, gezien de grote autocorrelatie (op te vatten als een maat voor de overlap in informatie, zie daarvoor paragraaf 10.3) van halfmaandelijke meetwaarden van de grondwaterstand. Als dezelfde reeks ook hoogfrequente loggermetingen bevat, zoals dagelijks of uurlijks, moeten ook die worden omgezet tot reeksen maandgemiddelden. Maar daarvoor moeten dan ook slechts twee meetwaarden worden gemiddeld, namelijk degene die het dichtst liggen bij het destijds voor die reeks gebruikelijke tijdstip van het handmatig meten (zoals de 14^e of de 28^e). Als er per maand meer loggerwaarden worden gemiddeld dan handmetingen ontstaat een reeks met niet-constante precisie van de waarden (aangeduid als heteroscedasticiteit), wat ongewenst is bij tijdreeksanalyse. Het residu (a_t) vormt dan namelijk geen witte ruis. Bij modelleren op maandbasis kunnen invoerreeksen bestaande uit somwaarden - zoals potentieel neerslagoverschot en grondwaterwinning - worden omgezet tot reeksen van maandgemiddelden, door per maand de maandsom te delen door het aantal dagen in die maand. Op die wijze wordt de kunstmatige variatie door verschil in maandlengte genivelleerd.

Het is verder aan te bevelen om bij de voorbewerking de invoerreeksen en uitvoerreeksen zodanig te schalen dat er geen geschatte modelparameters resulteren met uiterst kleine waarden. Zo zal het doorgaans handiger zijn om grotere grondwaterwinningen uit te drukken in 1.000 of 10.000 m³/d, in plaats van als m³/d.

Omgaan met niet-parallel beschikbare reekswaarden

Voor een BJTR-model, dat uitgaat van discrete tijd, moeten de waarden van alle reeksen parallel beschikbaar zijn, dus voor dezelfde tijdstippen. Dat zal echter niet altijd het geval zijn en dan moet een voorbewerking worden toegepast. Hieronder een voorbeeld.

De neerslag gemeten op een neerslagstation betreft de 24-uursom van 08:00 uur voorgaande dag tot 08:00 uur van de bij de meetwaarde opgegeven datum. Maar de neerslag en verdamping die zijn gemeten op een weerstation betreffen de etmaalsom van 0:00 uur tot 24:00 uur van de bij de meetwaarde opgegeven datum. Als die neerslag en verdampingsreeksen parallel moeten worden gemaakt voor tijdreeksanalyse kan dit met verschillende methoden, zoals:

1. omrekenen 24-uursom van de verdamping om 24:00 uur ($V_{24,t}$), naar die om 08:00 uur van de betreffende dag ($V_{8,t}$), volgens:

$$V_{8,t} = \frac{16}{24} V_{24,t-1} + \frac{8}{24} V_{24,t}$$

2. idem, maar dan volgens:

$$V_{8,t} = V_{24,t-1}$$

3. omrekenen 24-uursom van de neerslag om 08:00 uur ($R_{8,t}$) naar die om 24:00 uur ($R_{24,t}$), volgens:

$$R_{24,t} = \frac{8}{24} R_{8,t} + \frac{16}{24} R_{8,t+1}$$

De eerste twee methoden zijn bijvoorbeeld geschikt als de grondwaterstandreeks bestaat uit metingen van 08:00 uur en de derde als die bestaat uit metingen van 24:00 uur, of uit etmaalgemiddelden. Als bij modelleren op grotere tijdsbasis, zoals week of maand, wordt uitgegaan van geaggregeerde waarden zullen de verschillen tussen de drie methoden min of meer vervagen.

Schatten van ontbrekende waarden

Als waarden van de te modelleren reeks ontbreken op één of meer van de equidistante tijdstippen die mee moeten doen bij het ontwikkelen van een BJTR-model, zal dit uiteraard een nadelig effect hebben op het model. Bij oudere software gebaseerd op de klassieke BJTR-aanpak kon in zo'n geval zelfs niet worden gemodelleerd, omdat er anders een onbekend gebied zou worden betreden, waardoor de resultaten minder betrouwbaar zijn te achten.

Zoals eerder aangegeven, mogen de invoerreeksen een kleiner meetinterval hebben dan de uitvoerreeks, maar in principe mogen deze ook geen ontbrekende waarden hebben.

Het is aan de analist hoe om te gaan met ontbrekende waarden. In de praktijk worden deze vaak geïnterpoleerd, maar het moge duidelijk zijn dat dit geen ideale situatie is. Enigszins pragmatisch kan op basis van praktijkervaringen worden aanbevolen om niet te modelleren als de uitvoerreeks en/of één of meer van de invoerreeksen meer dan 10 % ontbrekende waarden heeft. Of, als in zo'n geval tóch wordt gemodelleerd, voldoende terughoudend te zijn met het trekken van conclusies naar aanleiding van het model.

3.5.2 Vorbewerken bij een PIRFICT-model

Zoals eerder genoemd is het voor het opzetten van een PIRFICT-tijdreeksmodel niet noodzakelijk om de stijghoogtereeksen om te zetten naar equidistante reeksen. Om deze redenen kan de PIRFICT-methode ook omgaan met eventuele gaten in de meetreeksen. Vorbewerking om deze redenen is daarom niet nodig.

Op het moment dat een PIRFICT-tijdreeksmodel wordt opgesteld met een relatief korte tijdseenheid, op bijvoorbeeld uurbasis, wordt aanbevolen om ook invoerreeksen met deze of kleinere tijdsstap te gebruiken. Een voorbeeld hiervan is wanneer een tijdreeksmodel op uurbasis gewenst is (bijvoorbeeld om de invloed van het getij op de stijghoogte te bepalen), maar een of meerdere van de invoerreeksen alleen op dagbasis beschikbaar zijn. Voor deze situatie wordt aanbevolen om de verklarende reeksen terug te brengen naar een kleinere frequentie (uur). Dit kan bijvoorbeeld gedaan worden op basis van het gemiddelde (bijvoorbeeld in het geval van een onttrekking) of, in het geval van bijvoorbeeld verdamping, op basis van een verdeling over de dag. Het is aan de modelleur om hier een keuze in te maken, hierbij moet gekeken worden naar de betreffende hydrologische processen.

4

ONTWIKKELEN BJTR-MODEL

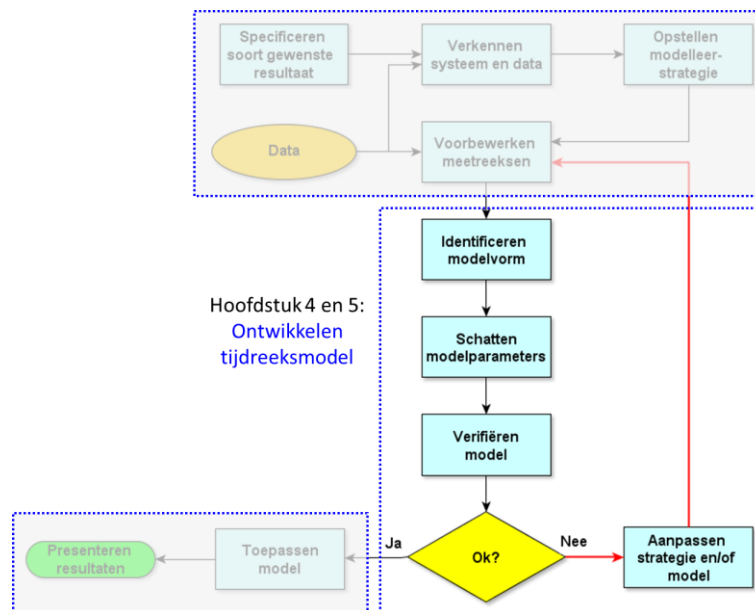
4.1 Inleiding

Dit hoofdstuk beschrijft de fase van het ontwikkelen van een BJTR-model. Deze bestaat uit drie stappen:

- identificeren modelvorm (paragraaf 4.2);
- schatten modelparameters (paragraaf 4.3);
- verifiëren model en eventueel aanpassen strategie en/of model (paragraaf 4.4).

De drie stappen worden zo nodig herhaald doorlopen, tot het model voldoet aan bepaalde randvoorwaarden.

Afbeelding 4.1 Stroomschema voor het uitvoeren van tijdreeksanalyse, met uitgelicht het ontwikkelen van het tijdreeksmodel



Ruis en residu

In dit hoofdstuk (BJTR-methode) heeft de term 'ruis' (N_t) betrekking op de stochastische component van de uitvoerreeks en de term 'residu' (a_t) op de stochastische invoerreeks.

4.2 Identificeren modelvorm

De eerste stap bij het ontwikkelen van het tijdreeksmodel is het identificeren van zijn vorm. Het gaat dan vooral om het deterministische deel van het model, dat bestaat uit de transferfuncties van de afzonderlijke invoerreeksen. Een transferfunctie (of ook wel impulsresponsfunctie) beschrijft de relatie van een invoerreeks met de uitvoerreeks. Op hoofdlijnen zijn er twee soorten aanpak voor het identificeren van de vorm van een transferfunctie, namelijk:

- 1 de empirische aanpak, met als adagium 'laat de meetdata zo vrijuit mogelijk spreken' en
- 2 de theoretische aanpak, waarbij de vorm van de relatie wordt opgelegd op basis van het conceptuele model van de werkelijkheid dat de analist heeft.

Bij het ontwikkelen van een BJTR-model wordt voor het identificeren van de vorm van een transferfunctie meestal gestart met de empirische aanpak, maar als die te weinig oplevert moet worden overgegaan op de theoretische aanpak. De empirische aanpak is namelijk minder effectief als de relatie tussen invoerreeks en uitvoerreeks onvoldoende sterk is om een helder beeld van zijn vorm te kunnen krijgen, of als er sprake is van meerdere invloedsfactoren van de uitvoervariabele, vooral als die ook onderling gecorreleerd zijn. Onderstaande paragraaf 4.2.1 licht toe hoe de vorm van een transferfunctie kan worden geïdentificeerd met de empirische aanpak en paragraaf 4.2.2 doet dit voor de theoretische aanpak. Deze stap is uitgewerkt voor de grondwaterwinning Nij Beets in paragraaf 7.5.

4.2.1 Empirisch identificeren vorm transferfunctie

Een sterk punt van de BJTR-methode is dat deze in staat stelt de vorm van de transferfunctie tussen een invoervariabele en de uitvoervariabele zo goed mogelijk te laten aansluiten op wat de meetdata daarover aan informatie verschaffen. Die vorm kan namelijk empirisch worden geïdentificeerd en vervolgens worden vertaald naar een wiskundige formulering van de transferfunctie. De details daarvan worden in deze subparagraaf beschreven.

Het praktisch uitvoeren van het empirisch identificeren is overigens eenvoudig, omdat de lastigste onderdelen van deze stap met geschikte software automatisch kunnen worden uitgevoerd. De analist hoeft dan alleen nog de empirisch vastgestelde vorm van de transferfunctie visueel te vergelijken met een aantal standaardvormen van respons op een enkelvoudige impuls, om op basis daarvan de best passende wiskundige formulering van de transferfunctie te kunnen kiezen. Illustraties van deze vereenvoudiging door softwarematige ondersteuning zijn te vinden in de beschrijving van de casus Nij Beets (paragraaf 7.5) en in het oefenvoorbeeld van paragraaf 8.1.

In het nu volgende worden de volgende aspecten van het empirisch identificeren besproken:

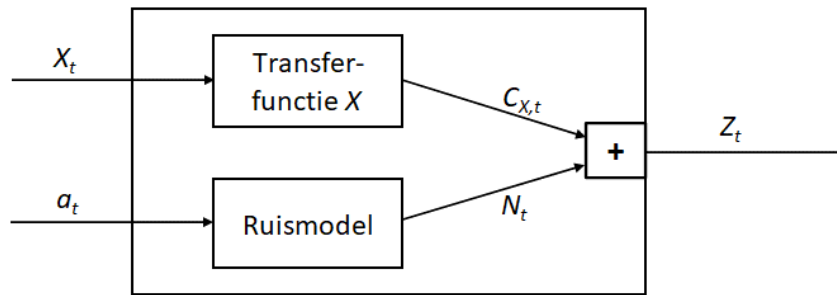
- ideale situatie om vorm transferfunctie te identificeren;
- het kruiscorrelogram kan vorm helpen identificeren, maar vertroebelt beeld bij autocorrelatie;
- de oplossing: het gewitte kruiscorrelogram;
- automatisch witten kruiscorrelogram;
- enkele voorbeelden van empirisch identificeren;
- empirisch identificeren kan ook meerwaarde geven bij toepassen PIRFICT-methode.

De toelichting wordt gegeven aan de hand van een systeem met één uitvoerreeks (Z_t) en één invoerreeks (X_t). Het BJTR-model van dit systeem heeft dan de volgende algemene vorm:

$$Z_t = C_{X,t} + N_t$$

waarin $C_{X,t}$ de component van Z_t die gerelateerd is aan X_t en N_t de ruis (zie Afbeelding 4.2 hieronder).

Afbeelding 4.2 Structuur van het BJTR-model met één invoerreeks (X_t), waarbij a_t het modelresidu is



Autocorrelatiestructuur ruis

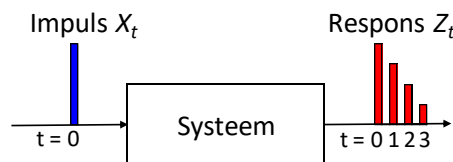
De ruis N_t kunnen we opvatten als de component van de uitvoerreeks die wordt veroorzaakt door: i) alle invloedsvariabelen die niet als invoerreeks zijn meegenomen, ii) onvolkomenheden van het model en iii) onvolkomenheden van de data. De ruisreeks N_t zal doorgaans nog autocorrelatiestructuur vertonen, zodat het zal lonen die structuur ook in het model te beschrijven. De uitvoerreeks Z_t kan dan immers nóg beter worden beschreven, wat bijvoorbeeld betere voorspellingen zal opleveren. Daarom wordt er ook een ruismodel toegevoegd aan het BJTR-model, waarmee de autocorrelatiestructuur van de ruis N_t wordt beschreven als lineaire functie van de residureeks a_t , die bestaat uit volledig willekeurige waarden, afkomstig uit dezelfde kansverdeling en zonder autocorrelatiestructuur. De residureeks a_t bevat dan geen informatie meer waarmee Z_t nóg beter zou kunnen worden beschreven.

Ideale situatie om vorm transferfunctie te identificeren

Om van de bovenvermelde algemene vorm van het BJTR-model tot een specifiekere vorm te komen, moet de vorm van de transferfunctie tussen X_t en $C_{X,t}$ (de bijbehorende component van Z_t) worden geïdentificeerd. Empirisch identificeren is het eenvoudigst als:

- 1 X_t een onderscheidende vorm vertoont, zoals een stap, een impuls (zie Afbeelding 4.3) of een sinus;
- 2 er geen sprake is van andere variabelen die Z_t beïnvloeden.

Afbeelding 4.3 De vorm van een transferfunctie tussen invoerreeks X_t en uitvoerreeks Z_t is eenvoudig vast te stellen als de invoerreeks een duidelijk onderscheidbare vorm heeft, zoals in dit voorbeeld een enkelvoudige impuls



In dit voorbeeld van Afbeelding 4.3 gaat de enkelvoudige impuls van X_t op tijdstip $t = 0$ gepaard met een onmiddellijke respons van Z_t (op tijdstip $t = 0$), gevolgd door een geleidelijk afnemende respons tot en met tijdstip $t = 3$. Uit de vorm van de respons kan de wiskundige formulering van de betreffende transferfunctie worden gekozen, door deze te vergelijken met standaardvormen van respons op een enkelvoudige impuls (zoals in Afbeelding 4.4). Als we de vorm van de respons van Z_t in Afbeelding 4.3 vergelijken met de standaardvormen in Afbeelding 4.4, komen er twee ARMA-transferfuncties in aanmerking voor het beschrijven van de betreffende relatie, namelijk de ARMA(1,0)-transferfunctie¹ en de ARMA(0,3)-transferfunctie.

¹ Een ARMA-transferfunctie is gekenmerkt door zijn ordes r en s en wordt uitgeschreven als ARMA(r,s). Hierbij is r het aantal AR-parameters ($\delta_1, \delta_2, \dots, \delta_r$) en is $s+1$ het aantal MA-parameters ($\omega_0, \omega_1, \omega_2, \dots, \omega_s$). Één AR-parameter kan een reeks exponentieel afnemende MA-parameters vervangen en zorgt zo voor reductie van het aantal modelparameters (dit is wiskundig uitgewerkt in paragraaf 12.4). Bij hydrologische processen blijkt in de praktijk meestal een r van 0 of 1 te volstaan.

De twee in aanmerking komende BJTR-modellen voor het betreffende systeem zijn dan:

- met de ARMA(1,0)-transferfunctie:

$$Z_t = \frac{\omega_0}{(1 - \delta_1 B)} X_t + N_t$$

- met de ARMA(0,3)-transferfunctie:

$$Z_t = \omega_0 X_t - \omega_1 X_{t-1} - \omega_2 X_{t-2} - \omega_3 X_{t-3} + N_t$$

waarin δ_1 een AR(1)-parameter en ω_i een MA(i)-parameter (i is hier respectievelijk 0, 1, 2 en 3). Zie hoofdstuk 11 voor een toelichting op het wiskundig formuleren van een BJTR-model.

Afbeelding 4.4 Standaardvormen van respons (positief of negatief) op een enkelvoudige impuls, voor enkele veel voorkomende ARMA-transferfuncties. Deze zijn bedoeld als hulpmiddel bij het empirisch identificeren van de vorm van een transferfunctie, op basis van visuele beoordeling van een gewit kruiscorrelogram. Zie de aanwijzingen onder deze Afbeelding

Transferfunctie	Vorm van respons	Transferfunctie	Vorm van respons
ARMA(0,0) $\omega_0 X_t$		ARMA(1,0) $\frac{\omega_0 X_t}{(1 - \delta_1 B)}$	
ARMA(0,1) $\omega_0 X_t - \omega_1 X_{t-1}$		ARMA(1,1) $\frac{\omega_0 X_t - \omega_1 X_{t-1}}{(1 - \delta_1 B)}$	
ARMA(0,2) $\omega_0 X_t - \omega_1 X_{t-1} - \omega_2 X_{t-2}$		ARMA(1,2) $\frac{\omega_0 X_t - \omega_1 X_{t-1} - \omega_2 X_{t-2}}{(1 - \delta_1 B)}$	

Toelichting:

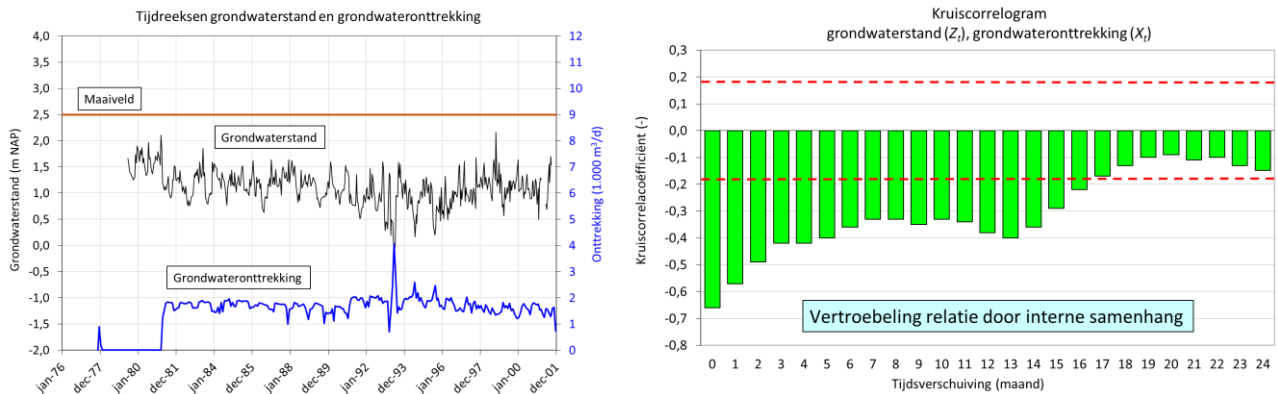
Bij de notatie ARMA(r,s) is r het aantal AR-parameters δ en $s+1$ het aantal MA-parameters ω (zie ook paragraaf 12.4). Uit de voorbeelden van Afbeelding 4.4 komen ondermeer de volgende punten naar voren:

- als de transferfunctie alleen een MA-structuur heeft, is er na één of meer duidelijk positieve of negatieve coëfficiënten een abrupte terugkeer richting nul;
- als de transferfunctie een AR-structuur heeft, is er sprake van een min of meer geleidelijke terugkeer richting nul;
- als de transferfunctie zowel een MA als een AR-structuur heeft, is er na één of meer duidelijk positieve of negatieve coëfficiënten een min of meer geleidelijke terugkeer richting nul.

Het kruiscorrelogram kan vorm helpen identificeren, maar vertroebelt beeld bij autocorrelatie

Een praktisch probleem met het empirisch identificeren van de vorm van een transferfunctie is dat ideale omstandigheden voor identificatie, zoals met een enkelvoudige impuls van X_t , normaliter niet optreden in hydrologische systemen. Er zal dan moeten worden uitgegaan van wat er wél beschikbaar is. Een hulpmiddel daarbij kan het kruiscorrelogram zijn, een plot van de kruiscorrelatiefunctie (zie Afbeelding 4.5).

Afbeelding 4.5 Voorbeeld van een kruiscorrelogram (rechts), ter identificatie van de vorm van de dynamische relatie tussen een grondwateronttrekking (X_t) en de grondwaterstand (Z_t) op 120 meter afstand (links). De twee tijdreeksen bestaan uit maandgemiddelden van halfmaandelijke meetwaarden. Door de autocorrelatie van de reeksen geeft het kruiscorrelogram een vertekend beeld van de vorm van hun dynamische relatie



Weergave theoretisch betrouwbaarheidsinterval kruiscorrelatie dat geldt bij geen relatie tussen X_t en Z_t

Om te bepalen of een geschatte kruiscorrelatiecoëfficiënt statistisch significant is, is in het kruiscorrelogram ook het theoretisch 95 %-betrouwbaarheidsinterval van de schatting weergegeven (met de horizontale rode stippellijnen), dat geldt onder de nulhypothese dat de twee tijdreeksen X_t en Z_t geen relatie vertonen. Dit interval wordt benaderd met ondergrens $-2/\sqrt{n}$ en bovengrens $+2/\sqrt{n}$, waarin n het aantal meetwaarden van een reeks is. ¹

Afbeelding 4.5 toont de geschatte kruiscorrelatiecoëfficiënt² ($r_{c,k}$), een maat voor de samenhang tussen de waarde van de invoerreeks X_t en de uitvoerreeks Z_t k tijdseenheden later, als functie van k (de tijdsverschuiving, in het Engels aangeduid als *lag*). Als deze reeksen equidistant zijn en even lang, dan kan deze coëfficiënt worden geschat als:

$$r_{c,k} = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(z_{t+k} - \bar{z})}{\sqrt{\sum_{t=1}^n (x_t - \bar{x})^2 \sum_{t=1}^n (z_t - \bar{z})^2}}$$

waarin x_t en z_t de waarden van invoerreeks X en uitvoerreeks Z op tijdstip t , \bar{x} en \bar{z} de geschatte gemiddelden van X en Z en n het aantal waarden. Deze formule komt er in het kort op neer dat de kruiscorrelatiefunctie wordt bepaald door de equidistante tijdreeksen van X en Z in twee kolommen naast elkaar te zetten en vervolgens de kolom van X telkens één rij naar beneden te schuiven en de correlatiecoëfficiënt van de twee kolommen te bepalen (met weglating van de rijen die slechts één meetwaarde bevatten).

Veel hydrologische tijdreeksen vertonen echter positieve autocorrelatie, wat inhoudt dat opeenvolgende waarden doorgaans meer op elkaar lijken dan waarden die veel eerder of later zijn genomen. Dit komt door de traagheid van het systeem en zorgt voor een veel minder ideale situatie voor identificatie dan de situatie met een enkelvoudige impuls van X_t , omdat er door de autocorrelatie niet meer goed kan worden vastgesteld door welke impuls van X_t een bepaalde respons van Z_t wordt veroorzaakt en evenmin hoe een respons is opgebouwd uit responsen van voorgaande impulsen van X_t . Het kruiscorrelogram kan dan slechts een vertoebeeld beeld geven van de dynamische relatie tussen twee tijdreeksen. Dit is ook te zien in Afbeelding 4.5 waar het kruiscorrelogram een vertekend beeld geeft van de vorm van de transferfunctie, gelet op de onrealistisch lange responstijd (> 2 jaar).

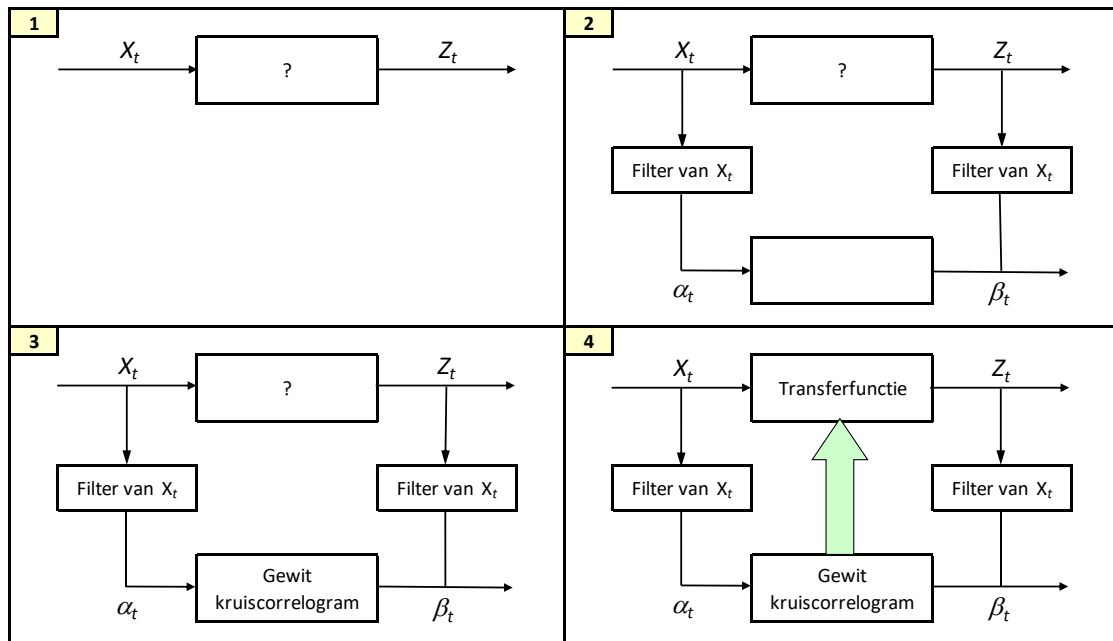
¹ Deze benadering verdisconteert nog niet dat de breedte van het betrouwbaarheidsinterval toeneemt met het aantal tijdsverschuivingen. Dit komt doordat een geschatte kruiscorrelatiecoëfficiënt op minder waarden is gebaseerd naarmate de tijdsverschuiving groter is.

² De kruiscorrelatiecoëfficiënt wordt ook wel aangeduid als de tijdsafhankelijke correlatiecoëfficiënt.

De oplossing: het gewitte kruiscorrelogram

De vertroebeling van een kruiscorrelogram door autocorrelatie kan sterk worden verminderd door beide reeksen vóór het schatten van de kruiscorrelatiefunctie door hetzelfde lineaire filter te halen. Deze remedie tegen vertroebeling van een relatie door autocorrelatie wordt aangeduid als het *witten* (Engels: *prewhitening*) van het kruiscorrelogram [Box and Jenkins, 1970]. Het witten is in een vierluik toegelicht in Afbeelding 4.6.

Afbeelding 4.6 Vierluik ter toelichting van het *witten* van het kruiscorrelogram. Het witten dient om een beter beeld te krijgen van de vorm van de dynamische relatie tussen invoerreeks (X_t) en uitvoerreeks (Z_t) van een systeem. Zo kan tot een empirisch onderbouwde formulering van de vorm van hun transferfunctie worden gekomen



De stappen van het in bovenstaande vierluik weergegeven witten van het kruiscorrelogram zijn dan:

1. de vraagstelling is de vorm te identificeren van de dynamische relatie tussen invoerreeks X_t en uitvoerreeks Z_t van een bepaald systeem, zodat deze kan worden vertaald naar de wiskundige formulering van hun transferfunctie. Die vorm is echter niet af te lezen uit hun kruiscorrelogram, doordat er vertroebeling optreedt door autocorrelatie van een reeks of van beide reeksen;
2. leidt daarom een lineair filter van de invoerreeks (X_t) af, dat deze terugbrengt tot witte ruis en filter daarmee zowel X_t als Z_t , tot respectievelijk reeks α_t en reeks β_t . De reeks α_t zal dan witte ruis vormen, maar dit zal doorgaans niet gelden voor de reeks β_t (tenzij de reeks Z_t veel lijkt op X_t). Dit afleiden van een lineair filter is in principe de enige lastige stap van het witten, maar deze is eenvoudig te automatiseren (zie alinea hierna);
3. stel visueel vast welke vorm het kruiscorrelogram van α_t en β_t (aangeduid als het *gewitte* kruiscorrelogram van X_t en Z_t) heeft. Doordat de reeks α_t witte ruis vormt, zonder autocorrelatie, is duidelijker te zien in welke vorm een impuls van α_t tevoorschijn komt in β_t . De dynamische relatie tussen α_t en β_t zal gelijk zijn aan die tussen X_t en Z_t , doordat α_t en X_t dezelfde lineaire relatie hebben als β_t en Z_t (beschreven door het lineaire filter van X_t);
4. vergelijk de in het gewitte kruiscorrelogram weergegeven vorm van de dynamische relatie tussen α_t en β_t met standaardvormen (zoals van Afbeelding 4.4) en kies op basis daarvan de wiskundige formulering van de transferfunctie voor de relatie tussen X_t en Z_t .

Automatisch witten kruiscorrelogram

Het in stap 2 van bovenstaand vierluik vermelde afleiden van een lineair filter van X_t kan een nogal bewerkelijke exercitie zijn, die bovendien de nodige expertise vergt. Maar een pragmatische oplossing, die doorgaans voldoende werkt, is het volgende eenvoudige lineaire AR-filter van X_t af te leiden:

$$\alpha_t = X_t - (\phi_1 X_{t-1} + \phi_2 X_{t-2} + \phi_3 X_{t-3} - \dots + \phi_u X_{t-u})$$

waarin α_t , de gefilterde vorm van de invoerreeks X_t en ϕ_i de AR-parameter van orde i ($i = 1 \dots u$), waarbij u voldoende groot is, zoals 20, 30, of meer.¹ De parameters van dit filter worden geschat als de parameterset die de som van de kwadraten van alle α_t minimaliseert.

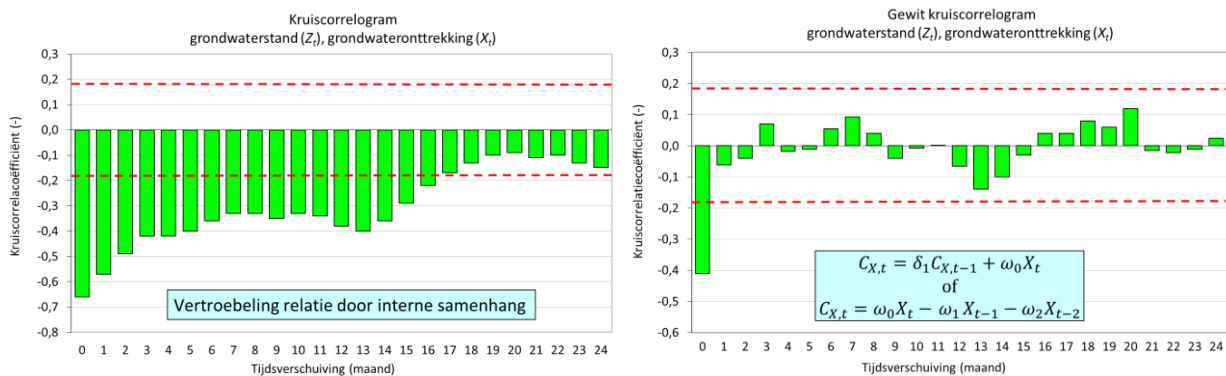
Ditzelfde AR-filter, met de voor X_t geschatte parameters $\phi_1, \phi_2, \dots, \phi_u$, moet vervolgens ook worden toegepast op de uitvoerreeks Z_t , om zo β_t te herleiden.

Deze pragmatische wijze van witten is geheel automatisch uit te voeren met geschikte software, zodat de analist alleen nog in stap 4 het gewitte kruiscorrelogram visueel hoeft te vergelijken met standaardvormen, om zo de wiskundige formulering van de transferfunctie te vinden. Dit wordt bijvoorbeeld geïllustreerd in de casus Nij Beets (paragraaf 7.5) en ook in het oefenvoorbeeld van paragraaf 8.1.

Enkele voorbeelden van empirisch identificeren

Als voorbeeld toont Afbeelding 4.7 zowel het kruiscorrelogram van Afbeelding 4.5 (links) als de gewitte versie daarvan (rechts). De vorm van de transferfunctie van de grondwaterwinning en de grondwaterstand is dankzij het witten duidelijker en ook veel realistischer geworden. Als we deze vorm vergelijken met de standaardvormen van Afbeelding 4.4 kan het een ARMA(1,0)- of een ARMA(0,2)-transferfunctie betreffen.

Afbeelding 4.7 Voorbeeld van kruiscorrelogram vóór het witten (links) en ná het witten (rechts). Door het witten is de vorm van de transferfunctie duidelijker en ook realistischer geworden



Het BJTR-model met de ARMA(1,0)-transferfunctie is uit te schrijven als:

$$Z_t = \frac{\omega_0}{(1 - \delta_1 B)} X_t + N_t$$

En met de ARMA(0,2)-transferfunctie is het uit te schrijven als:

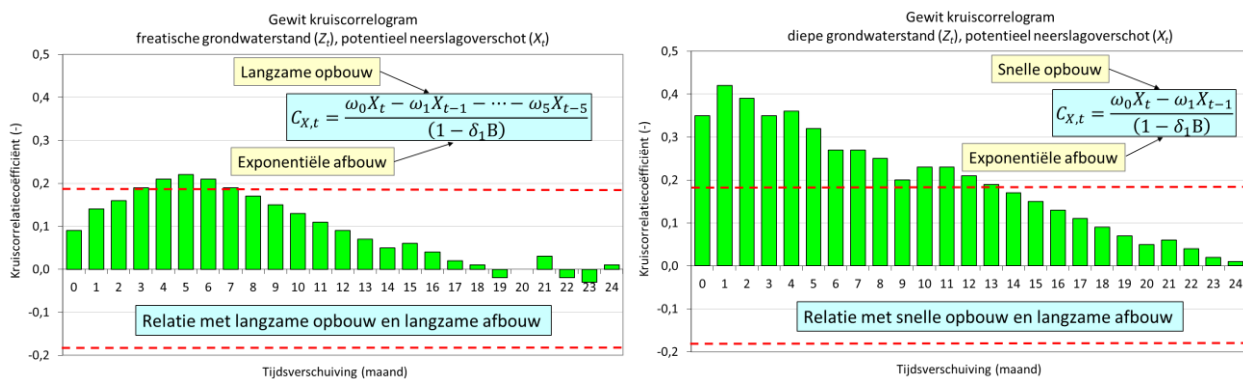
$$Z_t = \omega_0 X_t - \omega_1 X_{t-1} - \omega_2 X_{t-2} + N_t$$

Desgewenst kan de geschikteste van deze modellen worden gekozen op basis van één of ander informatiecriterium (zoals AIC, of BIC, zie paragraaf 4.4.1).

¹ Het betreffende filter is meestal té uitgebreid om als efficiënt model voor X_t bij praktijktoepassingen te kunnen fungeren, maar in veel gevallen wel voldoende in staat om de autocorrelatiestructuur zodanig te verwijderen dat het gewitte kruiscorrelogram bruikbare informatie kan verschaffen over de relatie tussen X_t en Z_t .

Afbeelding 4.8 toont twee andere voorbeelden van een gewit kruiscorrelogram, ter identificatie van de vorm van de dynamische relatie tussen het potentieel neerslagoverschot en de grondwaterstand. Het eerste voorbeeld betreft een freatische grondwaterstand, waar de maximale respons van de grondwaterstand op een impuls van het neerslagoverschot pas na vijf maanden optreedt. Het tweede voorbeeld betreft een stijghoogte in een gedeeltelijk afgesloten watervoerend pakket, waar de maximale respons veel sneller optreedt. In beide voorbeelden is er na het optreden van het maximum sprake van een min of meer geleidelijke terugkeer richting nul, welk deel efficiënt beschreven kan worden door een AR-parameter.

Afbeelding 4.8 Twee voorbeelden van een gewit kruiscorrelogram, ter identificatie van de vorm van de dynamische relatie tussen het potentieel neerslagoverschot en de grondwaterstand. Links betreft het een freatische grondwaterstand, waar de maximale respons van de grondwaterstand op een impuls van het neerslagoverschot pas na vijf maanden optreedt. Rechts betreft het een stijghoogte in een gedeeltelijk afgesloten watervoerend pakket, waar de maximale respons veel sneller optreedt. In beide gevallen is er na het optreden van het maximum sprake van een min of meer geleidelijke terugkeer richting nul. Deze afbouw kan worden beschreven door het AR-deel (met parameter δ_1)



Het empirisch identificeren zal doeltreffender zijn naarmate er sprake is van een sterkere relatie tussen de invoervariabele en de uitvoervariabele. Als meerdere tijdreeksen van hetzelfde gebied moeten worden gemodelleerd is het dan ook aan te bevelen te starten met de tijdreeks waarvan verwacht mag worden dat die de sterkste relatie vertoont met de voor het onderzoek relevantste invoervariabele. Bij een studie naar het effect van een ingreep is dat bijvoorbeeld de reeks gemeten in de peilbuis die zich het dichtst bij de ingreep bevindt.

Als er sprake is van twee of meer invoerreeksen, waarvan er één met een zeer sterk bepalende invloed op de uitvoerreeks, zal het lastig zijn de vormen van transferfuncties van andere invoerreeks(en) empirisch te identificeren. Een oplossing kan zijn de uitvoerreeks voorafgaand aan de identificatie te corrigeren voor zijn sterke relatie met de invloedrijke invoerreeks. Daarvoor moet dan eerst een BJTR-model worden afgeleid dat de relatie tussen die invoerreeks en de uitvoerreeks beschrijft. De ruisreeks van dat model kan vervolgens dienen om de transferfunctie(s) van de uitvoerreeks met de andere invoerreeks(en) empirisch te identificeren.

Empirisch identificeren kan ook meerwaarde geven bij toepassen PIRFICT-methode

Het boven beschreven identificeren op basis van empirie kan ook een geschikt hulpmiddel zijn bij het selecteren van een voorgedefinieerde impulsresponsfunctie voor een PIRFICT-model. Het kan namelijk dienen ter staving en soms zelfs ter vervanging van de selectie op basis van theoretische kennis. Maar evenzo kan bij toepassen van de BJTR-methode het identificeren op basis van theoretische kennis een geschikt substituuat zijn als de empirische aanpak te weinig informatie oplevert (zie paragraaf 4.2.2 hieronder).

4.2.2 Theoretische aanpak bij identificeren vorm transferfunctie

Als de empirische aanpak bij het identificeren van de vorm van de transferfunctie te weinig houvast geeft, moet worden overgegaan op de theoretische aanpak. Dit kan bijvoorbeeld het geval zijn als een invoerreeks te weinig variatie heeft vertoond over de modelperiode, of als het systeem een veelvoud aan invoerreeksen heeft. De vorm en zo mogelijk ook de responstijd van de relatie tussen invoerreeks en uitvoerreeks moeten dan worden aangenomen op basis van kennis van de hydrologische kenmerken van het beschouwde systeem. Zo kunnen bij gebrek aan empirische identificatiemogelijkheden zowel vorm als responstijd van de relatie tussen de grondwaterstand en een grondwateronttrekking worden geraamd door afstand, onttrekkingsdebiet en bodemconstanten in te voeren in een hydrologische formule voor de instationaire verlaging. Zo komt bijvoorbeeld de formule van Hantush in aanmerking bij een onttrekking van grondwater onder een deklaag met watervoerende sloten met een vast waterpeil.

4.2.3 Instellen vorm ruismodel

De vorm van het stochastische deel van het BJTR-model – aangeduid als het ruismodel - kan bij aanvang van de modelontwikkeling nog niet worden geïdentificeerd, omdat de ruis een latente variabele is, wat inhoudt dat zijn eigenschappen pas kunnen worden vastgesteld na het schatten van de parameters van het BJTR-model (voor een uitwerking voor een casus zie 7.5.4). Voor de eerste modelronde kan echter worden volstaan het ruismodel in te stellen als een AR(1)-model, volgens:

$$N_t - c = \phi_1(N_{t-1} - c) + a_t$$

waarin ϕ_1 de AR(1)-parameter en c een constante. Als bij de modelverificatie blijkt dat de modelresiduen nog geen witte ruis (onafhankelijke trekkingen uit dezelfde kansverdeling) vormen kan het ruismodel worden uitgebreid, bijvoorbeeld door een of meerdere MA-parameters (θ_1, θ_2 , etc.) en/of een AR-seizoensparameter (ϕ_1) toe te voegen, waarna het tijdreeksmodel opnieuw kan worden geschat. Een voorbeeld van een dergelijk uitgebreid ruismodel is¹:

$$(1 - \phi_1 B)(1 - \phi_1 B^S)(N_t - c) = (1 - \theta_1 B)a_t$$

waarin B de backshift operator ($BN_t = N_{t-1}$) en B^S de seizoengebonden backshift-operator ($B^S N_t = N_{t-S}$). Bijvoorbeeld bij modelleren op maandbasis geldt dat S 12 is en wordt dus een relatie gelegd met de N_t van een jaar geleden.

Als het bij modelleren van een grondwaterstandreeks nodig blijkt het ruismodel uit te breiden om een seizoen effect te beschrijven, terwijl het neerslagoverschot als invoerreeks meedoet, is de betreffende transferfunctie blijkbaar onvoldoende in staat het seizoenpatroon te beschrijven. Dit kan wijzen op een niet-lineaire relatie tussen neerslagoverschot en grondwaterstand, die onvoldoende wordt beschreven door de (lineaire) transferfunctie.

4.3 Schatten modelparameters BJTR-methode

Na het identificeren en formuleren van de vormen van de transferfunctie(s) en het instellen van de vorm van het ruismodel worden de modelparameters geschat, evenals hun standaardfouten (zie ook paragraaf 7.6). Het schattingsproces komt neer op het iteratief zoeken naar de combinatie van parameterwaarden die voldoet aan een bepaald criterium voor de mate waarin het model de uitvoerreeks beschrijft. Vaak wordt daarbij het criterium gehanteerd dat de combinatie van parameterwaarden de laagste kwadraatsom van de modelresiduen (a_t) moet opleveren. Dit principe wordt aangeduid als de kleinste kwadratenmethode (*Least Squares Method*).

¹ In paragraaf 12.4 is de algemene vorm van het ruismodel beschreven. Die algemene vorm wordt aangeduid als het ARIMA-model.

Per iteratiestap wordt de reeks van het modelresidu (a_t) bepaald als het verschil tussen de meetreeks van de uitvoervariabele en de modelreeks, de laatste zoals berekend met de dan geldende schattingen van alle modelparameters, dus zowel van het deterministische modeldeel, als van het ruisdeel (N_t). Daarbij wordt een formulering van het BJTR-model gebruikt die de N_t -reeks beschrijft als een naar de ruisparameters (ϕ, θ) gewogen som van de modelresiduen ($a_t, a_{t-1}, a_{t-2}, \text{etc.}$). De N_t -reeks is daarmee bij de BJTR-methode een latente variabele, die pas tevoorschijn komt na het schatten van de modelparameters, net als de a_t -reeks en de componenten van de invoerreeksen. Deze directe manier van parameterschatten is eenduidig en optimaal en is mogelijk doordat in discrete tijd wordt gemodelleerd met equidistante tijdstappen. Bij de PIRFICT-methode is dit niet mogelijk, omdat ervoor is gekozen dat deze methode ook overweg moet kunnen met niet-equidistante tijdstappen. Daar wordt eerst de realisatie van de N_t -reeks bepaald als het verschil van de uitvoerreeks en het deterministische deel van het tijdreeksmodel, waarna pas in een tweede stap de realisatie van de a_t -reeks wordt bepaald uit de N_t -reeks en de ruisparameter (α). Dit komt dan neer op een indirecte manier van parameterschatten.

Het schattingsproces stopt als het ingestelde maximum aantal iteraties is bereikt, of als wiskundige convergentie wordt bereikt voor de vector met modelparameters. Dit laatste is het geval als de kwadraatsom van de verschillen van parameters van opeenvolgende iteraties minder is dan een ingestelde tolerantie (dit heet convergentie volgens de Euclidische norm).

Als het model teveel parameters bevat en deze daardoor onderling sterk gecorreleerd zijn, kan het moeilijk zijn convergentie te bereiken. Daarom is het aan te raden de lijst met uitvoer van alle iteraties van het schattingsproces door te nemen, om zo te kunnen beoordelen of met hervatten van de iteraties convergentie kan worden bereikt. Als er geen convergentie kan worden bereikt is aan te bevelen de oorzaak te achterhalen en model en/of aanpak daar op aan te passen (zie daarvoor paragraaf 4.4.1).

Voor het schatten van het BJTR-model met de kleinste kwadratenmethode wordt vaak gebruik gemaakt van een niet-lineair schattingsalgoritme, zoals dat van Levenberg-Marquardt. Bij een tijdreeksmodel kan er namelijk - afhankelijk van de structuur - sprake zijn van een niet-lineair optimalisatieprobleem. Bij complexere modellen kan het ook nodig zijn geschikte beginschattingen van de parameters te verstrekken. Dit kan het aantal benodigde iteraties beperken, evenals het risico dat het schattingsproces in de meerdimensionale parameter ruimte terecht komt in een lokaal minimum van de kwadraatsom van de modelresiduen.

4.4 Verifiëren model en eventueel aanpassen model en/of strategie

Het verifiëren van het BJTR-model dient om onvolkomenheden te detecteren. Als deze worden gevonden, kan getracht worden ze te herstellen en als er geen onvolkomenheden worden gevonden geeft dat zowel een onderbouwd vertrouwen in geschatte evenwichtsrelaties (en hun standaardfouten), als in de met het model gegenereerde voorspellingen.

De verificatiefase bestaat vooral uit diagnostische checks, om te beoordelen of het model de uitvoerreeks van het systeem goed beschrijft, of het voldoet aan een aantal relevante vooronderstellingen en ook of het aansluit op fysisch-hydrologische logica. De verschillende diagnostieken van de verificatiefase zijn vermeld in onderstaande Tabel 4.1 en ze worden hierna toegelicht. Zie ook paragraaf 7.7, waar deze stappen zijn uitgewerkt voor de casus Nij Beets.

Tabel 4.1 Overzicht van diagnostieken om een geschat BJTR-model te kunnen beoordelen

Diagnostiek	Kan aanwijzingen geven over	In
Verloop en resultaten schattingsproces		
Wel of geen convergentie schattingsproces	Model dat suboptimaal is, té uitgebreid, instabiel, of onrealistisch	§ 4.4.1
Correlatiematrix van parameterschattingen		
Parameterschattingen en hun significanties		
Fysisch-hydrologische logica		
Informatie criterium (AIC, BIC)		
Kenmerken modelresidu a_t		
Reeks van modelresidu	Anomale reekswaarden, gebrek aan voorspelcapaciteit, modelresidu geen normaal verdeelde witte ruis	§ 4.4.2
Gemiddelde en standaardafwijking van modelresidu		
Histogram van modelresidu		
ACF en PACF van modelresidu		
Variatie modelresidu versus tijd en modelwaarde		
Relatie modelresidu a_t en invoerreeks X_t		
KCF's van modelresidu/modelruis en invoerreeksen	Transferfunctie die relatie invoerreeks en uitvoerreeks onvoldoende beschrijft	§ 4.4.3
Modelvalidatie met onafhankelijke meetwaarden		
Kenmerken verschil onafhankelijke meetwaarde en modeluitkomst	Kalibratieperiode met onvoldoende informatie over dynamische relaties	§ 4.4.4

4.4.1 Beoordelen verloop en resultaten schattingsproces

Wel of geen convergentie schattingsproces

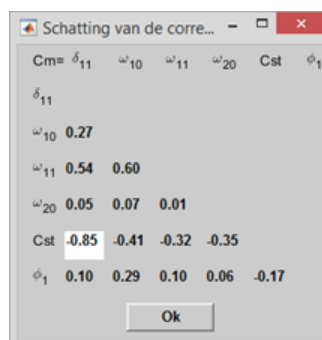
Als er geen convergentie kan worden bereikt is aan te bevelen de oorzaak te achterhalen en model en/of aanpak daar op aan te passen. Als mogelijke oorzaken komen in aanmerking:

- het schattingsproces is in een lokaal optimum blijven steken. Het is dan aan te bevelen om van een aantal bepalend geachte modelparameters beginschattingen en zo nodig ook grenzen van het schattingsgebied op te geven aan de software die het schattingsproces uitvoert;
- de schattingen van bepaalde modelparameters zijn onderling sterk gecorreleerd. Dit kan worden beoordeeld aan de hand van de correlatiematrix van de parameterschattingen (zie hieronder).

Correlatiematrix van parameterschattingen

De correlatiematrix van parameterschattingen vermeldt voor elke combinatie van modelparameters de correlatiecoëfficiënt van hun schattingen (zie het voorbeeld van Afbeelding 4.9).

Afbeelding 4.9 Voorbeeld van correlatiematrix van de geschatte parameters van een BJTR-model. Bij sterke correlatie ($|r| > 0,60$) is de cel wit gekleurd



Als parameterschattingen sterk positief of negatief gecorreleerd zijn (zoals $|r| > 0,60$), kan dit er op wijzen dat:

- het model teveel parameters bevat. Dit kan verder worden beoordeeld met de statistische significanties van de parameterschattingen (zie hieronder);
- bepaalde invoerreeksen sterk gecorreleerd zijn. In paragraaf 3.4.6 zijn suggesties te vinden hoe hiermee om te gaan.

Parameterschattingen en hun statistische significanties

Het schattingsproces levert als uitvoer van het BJTR-model voor elke modelparameter (zoals $\delta_1, \omega_0, \phi_1, \theta_1, c, g$, etc., zie daarvoor hoofdstuk 12) een schatting en bijbehorende standaardfout. Een parameter kwalificeert als statistisch significant als het - gegeven de modelonzekerheden - voldoende zeker is te achten dat zijn werkelijke waarde afwijkt van nul. Voor het bepalen van die statistische significantie kan de t-toets voor één steekproef worden gebruikt, op voorwaarde dat de modelresiduen normaal verdeelde witte ruis vormen (zie paragraaf 4.4.2). Als bijvoorbeeld de statistische significantie van de geschatte evenwichtsrelatie met een invoerreeks moet worden bepaald, volgt de toetsingsgrootte van die toets uit:

$$T = \frac{|\hat{g}|}{s_{\hat{g}}}$$

waarin T de toetsingsgrootte, \hat{g} de schatting van de evenwichtsrelatie en $s_{\hat{g}}$ zijn standaardfout. De te toetsen nulhypothese luidt dat de evenwichtsrelatie nul is ($H_0: g = 0$). Het is bij deze toepassing aan te raden als alternatieve hypothese te hanteren dat de evenwichtsrelatie ongelijk is aan nul ($H_a: g \neq 0$), ongeacht of zijn teken wel of niet aansluit op fysisch-hydrologische logica. Want als een parameter statistisch significant blijkt, maar met een teken (positief of negatief) dat niet aansluit op fysisch-hydrologische logica, geeft dat een zeer bruikbaar signaal dat er iets niet in orde is met de data, het model, of onze kennis van het betreffende systeem. Gezien deze tweezijdige alternatieve hypothese moet tweezijdig worden getoetst.

Het toetsen kan op twee manieren (zie de uitgebreidere toelichting in paragraaf 11.4):

- door de toetsingsgrootte te vergelijken met de kritische waarde van de Student-t-verdeling die volgt uit het gewenste significantieniveau. Het significantieniveau (α , ook wel aangeduid als fout van de 1^e soort) is het vóóraf ingestelde - en daarmee nog geaccepteerde - risico op het onterecht verwerpen van de nulhypothese (zoals 5 % of 1 %);
- door de significantie (ook wel aangeduid als p -waarde) van de toetsingsgrootte te bepalen en die te vergelijken met het significantieniveau. De significantie is de kans op een extremere toetsingsgrootte dan degene die is gerealiseerd, bij geldigheid van de nulhypothese. Als de significantie kleiner is dan het vooraf ingestelde significantieniveau, kan de nulhypothese worden verworpen.

Beide manieren zijn uitgewerkt in onderstaand voorbeeld.

Voorbeeld: Bepalen statistische significantie modelparameter

Stel dat een tijdreeks Z_t van 8 jaar lengte is gemodelleerd op maandbasis (96 waarden) en dat het BJTR-model 7 modelparameters heeft. De MA-parameter ω_0 voor de transferfunctie van invoerreeks X_{1t} is daarbij geschat op 9,44 met standaardfout 1,83 (beide met de dimensie eenheid Z_t / eenheid X_{1t}). Het aantal vrijheidsgraden is 89, bepaald als het aantal modelresiduen dat bij de kalibratie was betrokken (96) minus het aantal geschatte parameters (7). Stel dat we een toetsrisico α van 10 % accepteren dat onterecht de nulhypothese wordt verworpen. Dit komt er op neer dat tweezijdig wordt getoetst met 90 % betrouwbaarheid ($1 - \alpha$). Er geldt dan:

$$T = \frac{|9,44|}{1,83} = 5,158 > t_{95\%;89} = 1,662$$

waarin T de toetsingsgrootte en $t_{95\%;89}$ het 95-percentiel¹ van de Student-t-verdeling met 89 vrijheidsgraden, aangeduid als de kritische waarde voor de toets. Aangezien de toetsingsgrootte (5,158) groter is dan de kritische waarde (1,662) wordt de nulhypothese verworpen. De conclusie is dat er met 90 % betrouwbaarheid sprake is van een statistisch significant geschatte parameter ω_0 .

De tweede manier om te toetsen is de significantie (p -waarde) van de toetsingsgrootte te bepalen en die te vergelijken met het significantieniveau. De p -waarde volgt uit²:

$$p = 2 \cdot \text{Kans}[t_{89} > 5,158] = 0,00015\%$$

Aangezien deze significantie kleiner is dan het significantieniveau (10 %) wordt de nulhypothese verworpen.

Als een modelparameter statistisch niet significant blijkt, kan deze in principe worden verwijderd voor de volgende modelschatting. Dit hoeft echter niet per se en kan zelfs af te raden zijn in de volgende gevallen:

- als de statistisch niet-significante parameter de eerste is, of bij de eerste hoort van een groep parameters die logischerwijs bij elkaar horen, zoals een statistisch niet-significante θ_1 en/of θ_2 in een groep van meerdere MA-ruisparameters θ , waarbij de andere wél statistisch significant zijn;
- als de parameters samen een transferfunctie met een statistisch significante evenwichtsrelatie beschrijven. Zo ligt het niet voor de hand één of meer statistisch niet-significante MA-parameters van een transferfunctie ($\omega_0, \omega_1, \dots$) te verwijderen als de daaruit volgende evenwichtsrelatie tussen de betreffende invoerreeks en uitvoerreeks wél statistisch significant is.

Er moet ook altijd worden nagegaan of de 1^e-orde AR-parameter van het ruismodel (ϕ_1) niet te hoog is geschat. Als voor deze parameter namelijk een waarde van circa 0,95 of meer is geschat, kan dat er op duiden dat de ruis een trend tracht te beschrijven en zo wellicht onterecht een externe invloed verdisconteert. Als de schatting van ϕ_1 bijna 1 of meer bedraagt, probeert het ruismodel daadwerkelijk een trend te beschrijven en is te adviseren daarvoor een differentie in het ruismodel op te nemen (zie § 4.4.2).

Fysisch-hydrologische logica

Als onderdeel van het verifiëren van een tijdreeksmodel moet ook worden nagegaan of het model aansluit op fysisch-hydrologische logica. In bijvoorbeeld de volgende gevallen is daar duidelijk geen sprake van:

- een statistisch significante negatieve evenwichtsrelatie tussen neerslagoverschot en grondwaterstand;
- een statistisch significante positieve evenwichtsrelatie tussen grondwaterwinning en grondwaterstand;
- een voorspelde grondwaterstand die volgens het model bij uitzetten van de grondwaterwinning statistisch significant boven maaiveld uitkomt³, terwijl het een van oudsher droog gebied betreft.

In dergelijke gevallen moet het model als onrealistisch en daarmee ongeschikt worden beoordeeld.

Het is van belang hierbij tevens de modelonzekerheid in acht te nemen, zodat deze beoordeling alleen mogelijk is als de modelresiduen voldoen aan de randvoorwaarde dat ze normaal verdeelde witte ruis vormen. Anders kan de statistische significantie niet of verkeerd worden bepaald en is het niet mogelijk objectief vast te stellen of het model aansluit op fysisch-hydrologische logica.

¹ Er wordt tweezijdig getoetst met 90 % betrouwbaarheid, zodat de kritische waarde het 95-percentiel is van de kansverdeling van de toetsingsgrootte onder de nulhypothese. Hier is dat de Student-t-verdeling met 89 vrijheidsgraden. In Excel volgt die kritische waarde uit: = T.INV(0,95;89).

² In Excel volgt deze p -waarde uit: = 2 * T.VERD.RECHTS(5,158;89) of uit = T.VERD.2T(5,158;89).

³ Er is sprake van een voorspelde grondwaterstand die met 95 % betrouwbaarheid statistisch significant boven maaiveld uitkomt, als de ondergrens van zijn 95 %-voorspelinterval boven maaiveld uitkomt.

Alleen als er sprake blijkt van statistisch significant niet-aansluiten op fysisch-hydrologische logica is er een objectieve grond het model daarop af te keuren. Het niet aansluiten op fysische logica is immers irrelevant als het statistisch niet-significant is.

Door sommige analisten wordt een model ook wel gekwalificeerd als *plausibel* of *niet-plausibel*, afhankelijk van het wel of niet aansluiten op fysisch-hydrologische logica. De term *plausibel* kan echter een té positief oordeel over het model uitstralen, aangezien deze een combinatie van de volgende eigenschappen suggereert: *aannemelijk, aanvaardbaar, redelijk, waarschijnlijk, begrijpelijk, logisch, overtuigend, geloofwaardig of steekhoudend*. Dergelijke positieve kwalificaties gaan veel verder dan te rechtvaardigen is uit een vaststelling dat het model niet in tegenspraak is met fysisch-hydrologische logica.

Informatie criterium

Een informatie criterium is een maat waarmee de kwaliteit van verschillende modellen voor dezelfde dataset (in ons geval tijdreeks) kan worden vergeleken. Het kwantificeert de informatie die verloren gaat als een bepaald model wordt gebruikt om het proces te beschrijven dat de reeks genereerde. Daarbij wordt zowel de modelpasvorm meegewogen (de mate waarin het model aansluit op de reeks), als de complexiteit van het model (uitgedrukt in het aantal modelparameters). Een dergelijke maat kan objectieve onderbouwing bieden bij modelselectie.

Bekende informatiecriteria zijn het Akaike-InformatieCriterium (AIC) en het Bayes-InformatieCriterium (BIC). Als de modelresiduen normaal verdeelde witte ruis vormen wordt het AIC bepaald als:

$$\text{AIC} = -2\log(L) + 2p$$

waarin L de gemaximaliseerde *likelihood* (maat voor de pasvorm van het model) en p het aantal modelparameters.

En het BIC wordt dan bepaald als:

$$\text{BIC} = -2\log(L) + \log(n)p$$

waarin n het aantal waarden van de gemodelleerde reeks.

Beide informatiecriteria komen dus grofweg neer op het sommeren van een (inverse) maat voor de pasvorm en een maat voor de complexiteit van het model.

Voor beide criteria geldt dat het model met de laagste waarde relatief het best is, bij vergelijking van verschillende modellen voor dezelfde reeks. Maar let wel dat deze informatiecriteria slechts relatieve maten zijn om verschillende modellen voor dezelfde reeks te kunnen vergelijken. Ze leveren namelijk geen informatie over de kwaliteit van het model en kunnen in principe dus ook voor een aantal slechte modellen de minst slechte aangeven.

AIC kan tot overparametrisatie leiden en BIC tot onderparametrisatie. De laatste is dus wat strenger, maar het is aan te bevelen beide criteria te gebruiken. Meestal zullen ze namelijk hetzelfde model als beste aanwijzen en als ze dat niet doen, is het nog steeds informatief om dat dan zo te vermelden.

Maar houd wel voor ogen dat automatische modelselectie een analist nooit van de plicht kan ontslaan te verifiëren of aan de modelveronderstellingen is voldaan.

4.4.2 Beoordelen kenmerken modelresidu a_t

Bij het verifiëren moeten meerdere kenmerken van het modelresidu worden beoordeeld.

Reeks van modelresidu

Ga na of de tijdreeks van het modelresidu anomalieën vertoont. Als dat het geval is kan het tot vertekende schattingen van modelparameters hebben geleid. Hiervoor kan een visuele beoordeling van de geplote residuereeks volstaan.

Gemiddelde en standaardafwijking van modelresidu

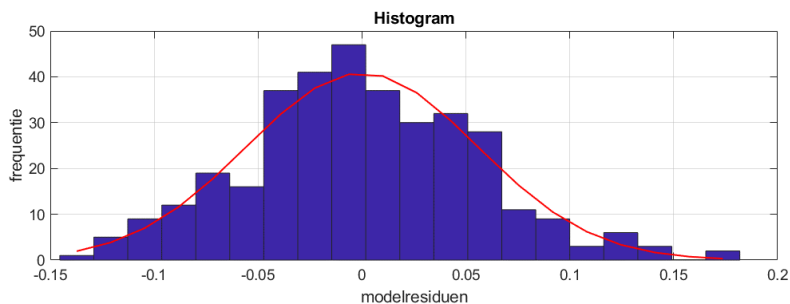
Als het gemiddelde van het modelresidu statistisch significant afwijkt van nul zullen de modelvoorspellingen onzuiver zijn, wat betekent dat ze een systematische fout zullen vertonen. Het duidt op een onvolkomenheid bij het identificeren van het model. Als het model nog geen constante omvat, kan dit worden opgelost door een constante in het model op te nemen. Anders moet worden getracht één of meer transferfuncties anders te formuleren.

De standaardafwijking van het modelresidu wordt geschat met de reeks modelresiduen en is op te vatten als een maat voor het deel van de variatie van de te modelleren reeks dat niet beschreven kan worden door het BJTR-model (dat wil zeggen het deterministische deel plus het stochastische deel). Deze standaardafwijking is ook een maat voor de nauwkeurigheid waarmee het model één stap vooruit kan voorspellen. Want als een voorspelling één stap vooruit wordt gedaan, dan zal gemiddeld 95 % van deze voorspellingen een fout hebben die minder is dan tweemaal de standaardafwijking (er van uitgaande dat het modelresidu normaal verdeelde witte ruis vormt). En samen met de modelstructuur bepaalt deze standaardafwijking ook de nauwkeurigheid waarmee het model twee of meer stappen vooruit kan voorspellen.

Histogram van modelresidu

Om te beoordelen of de modelresiduen afkomstig zijn uit een normale kansverdeling kan in principe worden volstaan met een visuele beoordeling van een histogram (zie Afbeelding 4.10) van het modelresidu, of andere grafische weergaven van de verdeling (zoals de Q-Q-plot), mits de analist daar enige ervaring mee heeft opgebouwd.

Afbeelding 4.10 Voorbeeld van een histogram van de modelresiduen. De rode lijn geeft aan wat de vorm zou zijn van een normale verdeling met het gemiddelde en de standaardafwijking van de modelresiduen. Op basis van een visuele beoordeling van deze afbeelding kan worden aangenomen dat de modelresiduen afkomstig zijn uit een kansverdeling die niet sterk afwijkt van de normale kansverdeling



Er zijn uiteraard ook diverse statistische toetsen beschikbaar om te beoordelen op normaliteit, zoals D'Agostino-K2 [D'Agostino, 1986], Jarque-Bera [1980, 1981 en 1987] en Shapiro-Wilk [1965]. Deze hebben echter de volgende tekortkomingen:

- bij kleine steekproeven (< 20) hebben ze slechts weinig onderscheidend vermogen¹ en kunnen ze alleen grote afwijkingen van normaliteit detecteren;
- bij grote steekproeven (> 250) hebben ze juist veel onderscheidend vermogen en kunnen dan afwijkingen van normaliteit detecteren die praktisch niet relevant zijn. Kleine afwijkingen zijn immers niet relevant omdat toepassingen die uitgaan van de randvoorwaarde van normaliteit in het geval van een grote steekproef vrij robuust zijn tegen afwijkingen van normaliteit. Er is alleen geen objectieve methode om vast te stellen tot welke afwijking van normaliteit die robuustheid nog opgaat, zodat dit vaak subjectief moet worden ingeschat, zoals op basis van visuele beoordeling van het histogram van de gegevens.

¹ Onderscheidend vermogen is de kans om terecht de nulhypothese te verwerpen. Dit vermogen is mede een functie van de mate waarin de werkelijkheid afwijkt van de nulhypothese.

Als ervoor wordt gekozen te toetsen op normaliteit, is het raadzaam dit uit te voeren met een hoge betrouwbaarheid (zoals 99 %), omdat zelfs een geringe en praktisch niet-relevante afwijking van normaliteit al snel statistisch significant kan zijn door het grote aantal reekswaarden bij tijdreeksanalyse. Maar zelfs dan vergt het beoordelen op normaliteit nog pragmatisme en mag aan het oordeel op basis van visuele beoordeling van het histogram meer gewicht worden toegekend dan aan het resultaat van de statistische toets.

Als uit de beoordeling blijkt dat er sprake is van een relevant te achten afwijking van normaliteit van het modelresidu zijn er twee mogelijkheden dat probleem op te lossen, te weten:

- aggregeren van de tijdreeksen naar een grotere tijdseenheid (zoals van een enkele meting naar een week-, tweeweeks, of maandgemiddelde van de metingen);
- transformeren van de uitvoervariabele. Bij een verdeling die sterk scheef naar rechts is kan bijvoorbeeld de logaritmische transformatie uitkomst bieden:

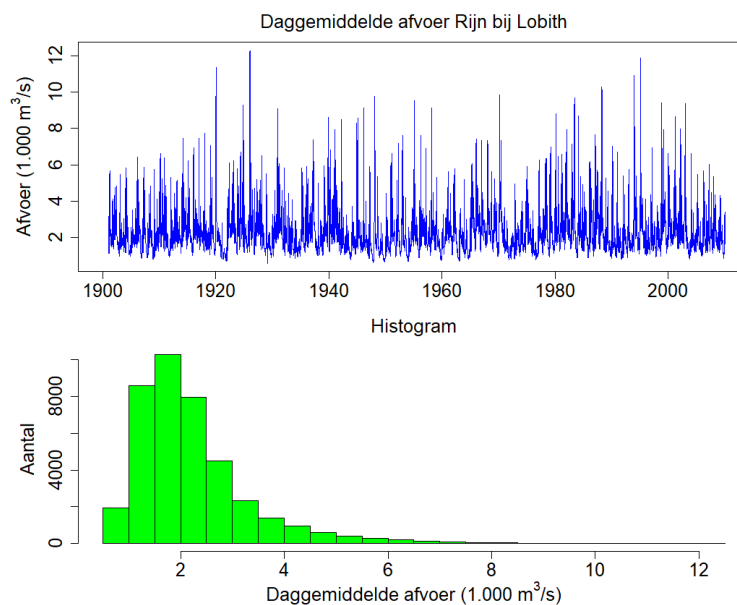
$$Z_t^* = \log_{10}(Z_t)$$

Bij het modelleren van grondwaterstandreeksen zal meestal wel worden voldaan aan normaliteit van het modelresidu, maar bij een snel reagerend grondwatersysteem kan er sprake zijn van een positief scheve kansverdeling van de grondwaterstand, vooral als wordt gemodelleerd op uur- of dagbasis. De neerslag is immers sterk scheef verdeeld en in een snel reagerend systeem is er een min of meer instantane reactie op de neerslag, met weinig uitsmering over de tijd, zodat die scheefheid ook naar voren zal komen in de grondwaterstand.

Meetreeksen van de afvoer vertonen doorgaans wél een sterke scheefheid (zie het voorbeeld in Afbeelding 4.11), zodat daar rekening mee moet worden gehouden. Bij tijdreeksanalyses is het dan ook niet ongebruikelijk om de logaritme van de afvoer te modelleren.

Afbeelding 4.11 Meetreeks van de daggemiddelde afvoer van de Rijn bij Lobith, bepaald over de periode 1901 t/m 2009.

Daaronder het histogram van de meetwaarden. Het betreft een verdeling die zeer sterk scheef naar rechts is (dat wil zeggen met uitschieters naar boven, wat in het histogram naar rechts is)



Autocorrelatiestructuur modelresidu a_t

Als de reeks van het modelresidu a_t nog autocorrelatiestructuur vertoont, biedt dat een goede gelegenheid het model te verbeteren. Een autocorrelatiestructuur kan namelijk in veel gevallen met het stochastische deel (het ruismodel) van het BJTR-model worden beschreven, waardoor het model beter zal aansluiten op de tijdreeks en zo ook betere voorspellingen kan opleveren. Verder is het ontbreken van autocorrelatie van het modelresidu zelfs een vereiste als het model dient om uitspraken met gekwantificeerde onzekerheden te kunnen doen (zie daarvoor paragraaf 6.2.2). De modelresiduen dienen dan namelijk te voldoen aan de volgende randvoorwaarden:

- 1 ze zijn gemiddeld nul;
- 2 ze zijn onafhankelijk van elkaar (er is dan geen sprake van een autocorrelatiestructuur);
- 3 ze hebben dezelfde variantie¹ (dit wordt aangeduid als *homoscedasticiteit*);
- 4 ze zijn afkomstig uit de normale kansverdeling (of een andere kansverdeling waarvan alle eigenschappen bekend zijn).

Als een reeks voldoet aan de eerste drie voorwaarden spreken we van *witte ruis*. Als deze ook voldoet aan normaliteit spreken we van *normaal verdeelde witte ruis*. Aanvullende randvoorwaarden, die echter vaak over het hoofd worden gezien en daardoor meestal impliciet worden aangenomen, zijn dat de formulering van het deterministische deel van het tijdreeksmodel overeen stemt met de werkelijke vorm (ook wel aangeduid als het *werkelijke systeem*) en dat de schatter van de evenwichtsrelatie tussen een invoerreeks en de uitvoerreeks zuiver is, dat wil zeggen zonder systematische fout.

Het checken op autocorrelatie van het modelresidu kan worden uitgevoerd door visuele beoordeling van de combinatie van twee diagnostieken van de reeks van het modelresidu a_t , namelijk de autocorrelatiefunctie (ACF) en de partiële autocorrelatiefunctie (PACF):

- de autocorrelatiefunctie (ACF) toont de schatting van de autocorrelatiecoëfficiënt (r_k), een maat voor de samenhang tussen het modelresidu op tijdstip t en het modelresidu k tijdseenheden later, als functie van k (de tijdsverschuiving). Deze coëfficiënt wordt per tijdsverschuiving bepaald als:

$$r_k = \frac{\sum_{t=1}^{n-k} (a_t - \bar{a})(a_{t+k} - \bar{a})}{\sum_{t=1}^n (a_t - \bar{a})^2}$$

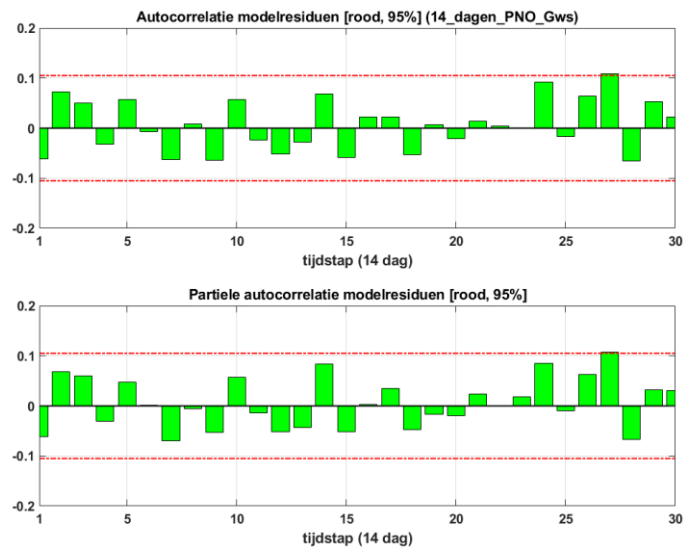
waarin a_t de waarde van het modelresidu op tijdstip t , \bar{a} het geschatte gemiddelde van het modelresidu en n het aantal modelresiduen. Merk op dat r_k voor $k = 0$ altijd precies 1 is.

- de partiële autocorrelatiefunctie toont de schatting van de partiële autocorrelatiecoëfficiënt (pr_k) als functie van k (de tijdsverschuiving). Deze wordt per tijdsverschuiving berekend als de autocorrelatiecoëfficiënt, zij het ná correctie voor de autocorrelaties bij kleinere tijdsverschuivingen ($k-1$, $k-2$, et cetera, tot het ingestelde maximum van de X-as van de PACF).

Afbeelding 4.12 toont een combinatie van ACF en PACF voor modelresiduen zonder autocorrelatie.

¹ Dit impliceert dat de residuen kunnen worden gezien als trekkingen uit dezelfde kansverdeling.

Afbeelding 4.12 Voorbeeld van combinatie van ACF en PACF als de modelresiduen geen autocorrelatie vertonen. De horizontale rode stippellijnen begrenzen het 95 %-betrouwbaarheidsinterval dat voor normaal verdeelde modelresiduen geldt bij geen autocorrelatie. De tijdstap waarmee is gemodelleerd bedraagt hier 14 dagen. Hieronder is de visuele beoordeling van een combinatie van ACF en PACF visueel toegelicht



Hoe een combinatie ACF en PACF visueel te beoordelen

Ga na of er in ACF en/of PACF aanwijzingen zijn voor autocorrelatie, zoals:

- binnen ongeveer de eerste 5 tijdsverschuivingen wijken één of meer coëfficiënten duidelijk af van de rest en/of vertonen een duidelijk samenhangend patroon;
- er is sprake van duidelijk afwijkende coëfficiënten op tijdsverschuivingen die kunnen samenhangen met periodiciteit. Zo kan bij modelleren met tijdseenheid 14 dagen een seizoenseffect tot uiting komen door afwijkende coëfficiënten op 26, 52, 78, ... tijdsverschuivingen;
- ruim meer dan 5 % van de coëfficiënten valt buiten het 95 %-betrouwbaarheidsinterval.

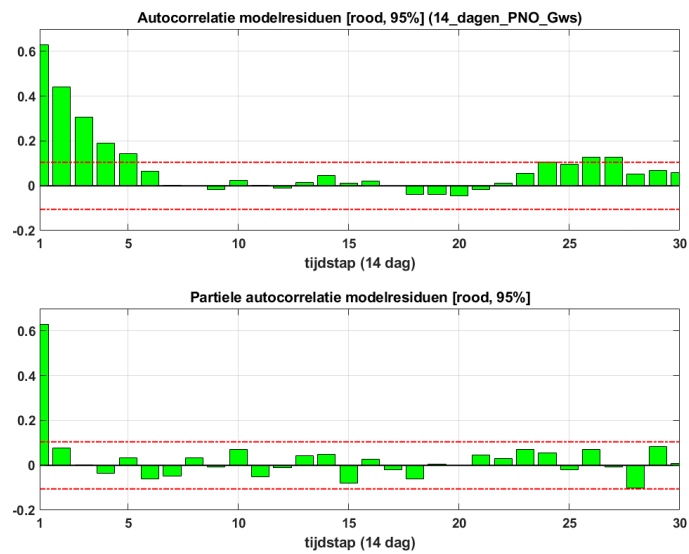
Als dit één of meer aanwijzingen voor autocorrelatie oplevert, identificeer dan aan de hand van de combinatie van ACF en PACF van de ruis welke vorm van het ruismodel nodig is om het modelresidu zonder autocorrelatie te krijgen. Als hulpmiddel daarvoor toont

Afbeelding 4.14 voorbeelden van combinaties van ACF en PACF voor verschillende autocorrelatiestructuren.

Een visuele beoordeling van de combinatie van ACF en PACF van Afbeelding 4.12 levert volgens de hierboven vermelde criteria geen aanwijzingen voor autocorrelatie.

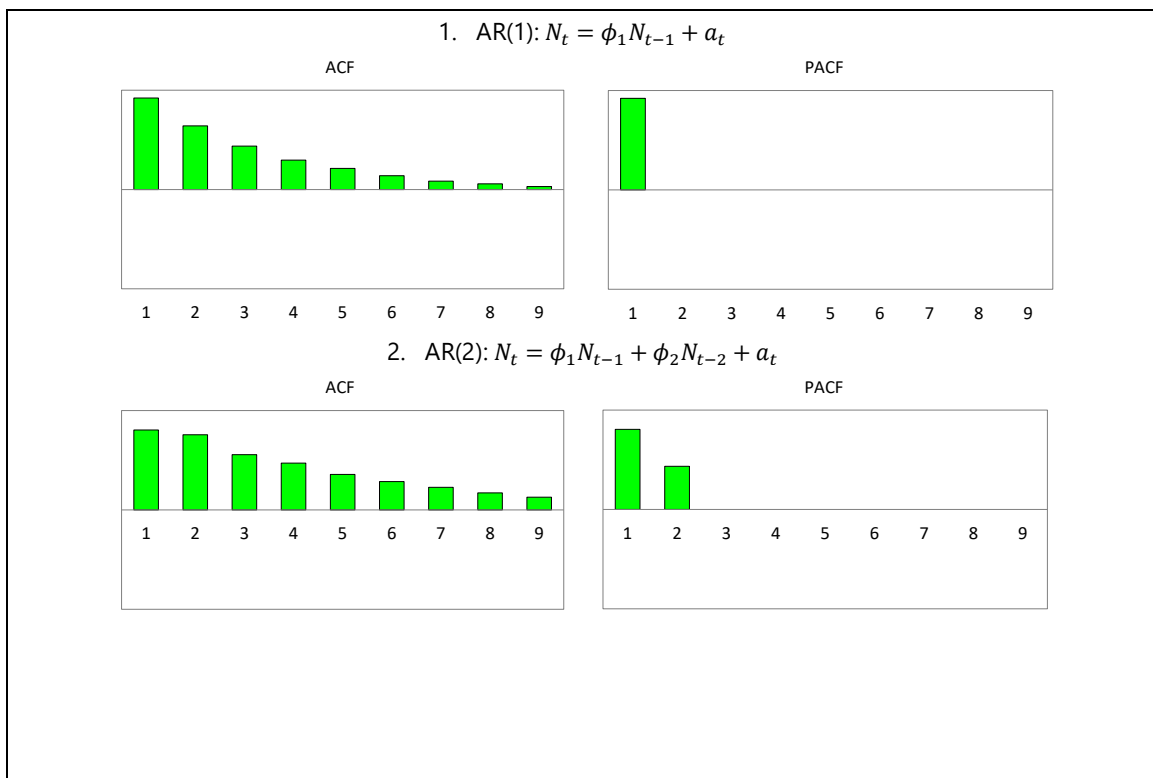
En Afbeelding 4.13 toont een voorbeeld van hoe de combinatie van ACF en PACF er uit kan zien als de modelresiduen wél autocorrelatie vertonen. De ACF toont zowel één of meer duidelijk afwijkende coëfficiënten als een duidelijk samenhangend patroon van coëfficiënten binnen de eerste 5 tijdsverschuivingen. Ook toont de ACF nog een kleine aanwijzing voor autocorrelatie op jaarbasis, gezien het samenhangend patroon en enigszins verhoogde waarden van de coëfficiënten rond tijdsverschuiving 26 (bij tijdstap 14 dagen komt dat neer op 364 dagen). En de PACF toont een duidelijk afwijkende coëfficiënt voor tijdsverschuiving 1.

Afbeelding 4.13 Voorbeeld van combinatie van ACF en PACF als de modelresiduen autocorrelatie vertonen

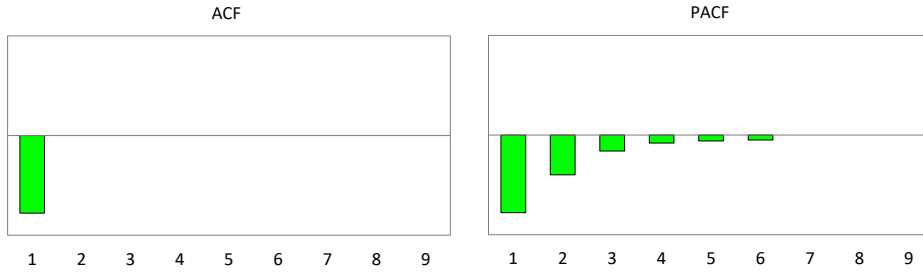


Als hulpmiddel voor het identificeren van de vorm van het ruismodel toont Afbeelding 4.14 voorbeelden van combinaties van ACF en PACF voor verschillende autocorrelatiestructuren. Dit kan van pas komen om het ruismodel te herformuleren als na de eerste keer schatten van het BJTR-model blijkt dat de modelresiduen nog autocorrelatie vertonen. Als we bijvoorbeeld de combinatie van ACF en PACF van bovenstaande Afbeelding 4.13 vergelijken met de voorbeelden van Afbeelding 4.14 lijkt er sprake van een AR(1)-proces.

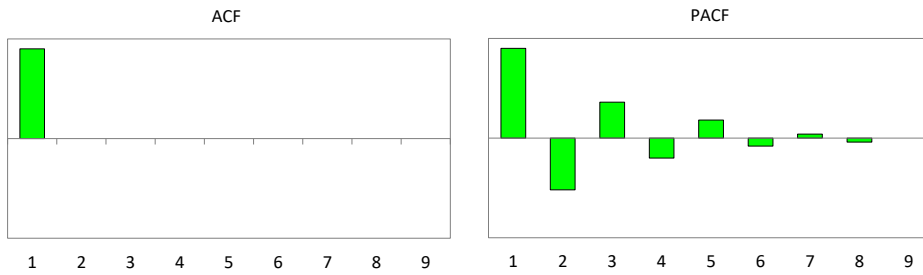
Afbeelding 4.14 Combinaties van ACF en PACF voor enkele AR-, MA- en ARMA-processen. Deze zijn bedoeld als hulpmiddel bij het identificeren van de vorm van het ruismodel. Dat identificeren is echter pas mogelijk na de eerste schatting van het BJTR-model, omdat de ruis niet bemeaten is, maar een latente variabele is, die pas zichtbaar kan worden na het schatten van het BJTR-model



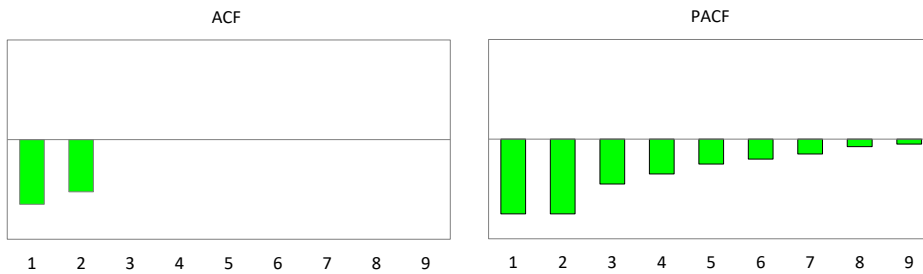
3. MA(1): $N_t = a_t - \theta_1 a_{t-1}$ (bij positieve θ_1)



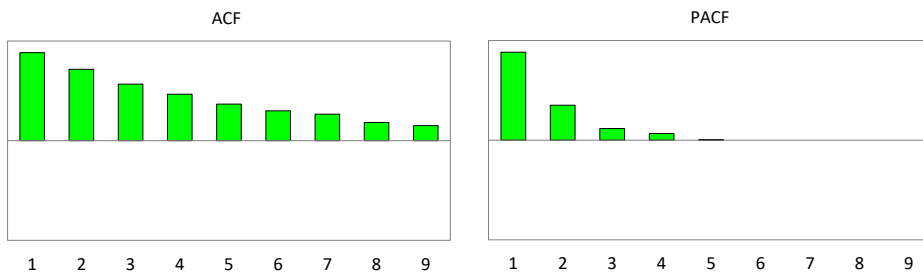
4. MA(1): $N_t = a_t - \theta_1 a_{t-1}$ (bij negatieve θ_1)



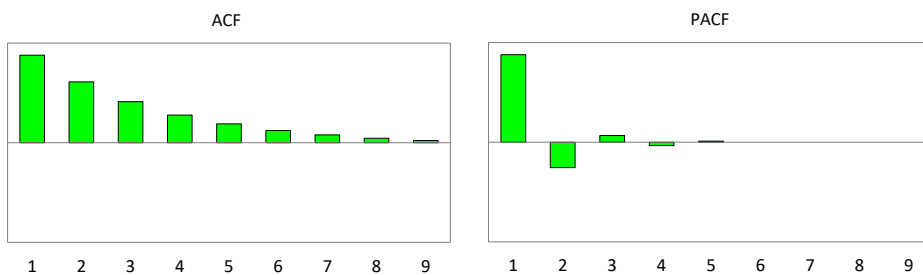
5. MA(2): $N_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$



6. ARMA(1,1): $N_t = \phi_1 N_{t-1} + a_t - \theta_1 a_{t-1}$ (bij positieve θ_1)



7. ARMA(1,1): $N_t = \phi_1 N_{t-1} + a_t - \theta_1 a_{t-1}$ (bij negatieve θ_1)

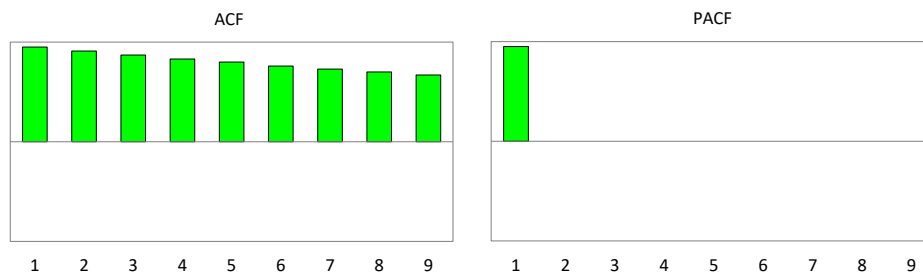


Uit de voorbeelden van Afbeelding 4.14 komen onder meer de volgende punten naar voren:

- als er sprake is van autocorrelatie komt dit al tot uiting bij de coëfficiënten van de eerste paar tijdsverschuivingen. Tenzij er alleen sprake is van autocorrelatie op jaarbasis;
- als we alleen de ACF beoordelen geeft dat nog onvoldoende uitsluitsel over de vorm van de autocorrelatiestructuur. De ACF's van vier van deze voorbeelden lijken namelijk sterk op elkaar (voorbeelden 1, 2, 6 en 7), terwijl uit hun PACF's blijkt dat het verschillende processen betreft. Dit illustreert dat voor het identificeren van de autocorrelatiestructuur de ACF en de PACF altijd in samenhang moeten worden beoordeeld;
- als de ACF een min of meer geleidelijke afname richting nul vertoont is er sprake van een AR-proces of een ARMA-proces. Een nadere specificatie daarvan kan worden afgeleid uit de PACF;
- als in de ACF de coëfficiënten van de eerste tijdsverschuivingen duidelijk afwijken van de andere en er daarna een abrupte terugval is naar niet-afwijkende waarden, is er sprake van een MA-proces.

Er is sprake van een bijzonder geval als de ACF van de ruis slechts zeer langzaam afneemt richting nul (zoals in Afbeelding 4.15). Het kan dan zijn dat de ruisreeks niet-stationair is (dat wil zeggen een trend vertoont), in welk geval een differentie aan het ruismodel moet worden toegevoegd (dit wordt hieronder verder uitgewerkt). Maar het kan ook een indicatie zijn om op grotere tijdsbasis te modelleren.

Afbeelding 4.15 Voorbeeld van combinatie van ACF en PACF die kan duiden op een niet-stationaire ruisreeks



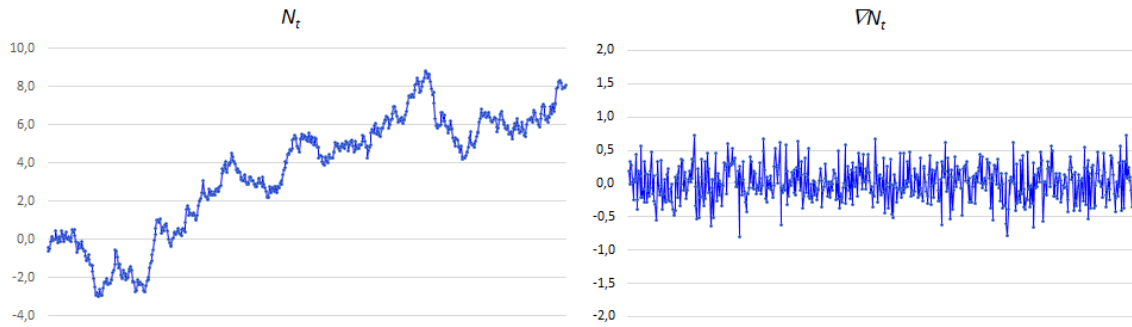
Soms blijkt een te modelleren uitvoerreeks een trend te vertonen die door geen van de beschikbare invoerreeksen kan worden verklaard. Dit is een aanwijzing dat er onvoldoende bekend is over het te modelleren systeem. De trend zal zich dan ook manifesteren in een combinatie van ACF en PACF van de ruis zoals weergegeven in Afbeelding 4.15.

Een objectieve manier om een dergelijke onverklaarbare trend modelmatig te verdisconteren is een differentie (∇) op te nemen in het ruismodel, volgens:

$$\nabla N_t = N_t - N_{t-1}$$

De trend zal dan vaak verdwijnen, zodat een stationaire reeks resteert. Het voorbeeld Afbeelding 4.16 hieronder illustreert hoe een niet-stationaire ruisreeks door een differentie stationair kan worden gemaakt.

Afbeelding 4.16 Voorbeeld van het resultaat van het verdisconteren van een trend in de ruis door een differentie



Na toevoegen van een differentie aan het ruismodel zal het bij de daaropvolgende modelverificatie vaak ook nog nodig blijken de MA-structuur aan te passen, door toevoegen van één of meer MA-parameters θ . Dit om te kunnen bewerkstelligen dat de residuen witte ruis vormen, zoals:

$$\nabla N_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-1} \Rightarrow N_t = N_{t-1} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-1}$$

Bij sterk dynamische trends, met willekeurige veranderingen van niveau en helling, kan het zelfs nodig zijn een dubbele differentie toe te passen om stationariteit te bewerkstelligen, volgens:

$$\nabla^2 N_t = \nabla(N_t - N_{t-1}) = N_t - 2N_{t-1} + N_{t-2}$$

Het opnemen van een differentie in het ruismodel is overigens een noodzaak - en dus niet slechts een mogelijkheid - als een te modelleren uitvoerreeks een trend vertoont die bij geen van de beschikbare invoerreeksen voorkomt. De ruis dient namelijk te voldoen aan stationariteit, wat inhoudt dat zijn gemiddelde en variantie niet veranderen met de tijd.

Het moge overigens wel duidelijk zijn dat als een te modelleren uitvoerreeks een trend vertoont die door geen van de beschikbare invoerreeksen kan worden verklaard, het dan niet mogelijk zal zijn om met het afgeleide tijdreeksmodel betrouwbare voorspellingen te kunnen doen, ongeacht hoe die trend in het model is beschreven. De kwaliteit van die voorspellingen hangt immers sterk af van ons inzicht in het toekomstige verloop van de trend en dat inzicht ontbreekt in zo'n geval blijkbaar.

Een trend kan het ook lastig of zelfs onmogelijk maken de vorm van de transferfunctie met een invoerreeks langs empirische weg te identificeren, in welk geval deze vorm op basis van theoretische kennis zal moeten worden aangenomen (zie paragraaf 4.2.2).

Bij hoogfrequent modelleren is het raadzaam om bij de visuele beoordeling van de autocorrelatiestructuur van het modelresidu steeds twee versies van de ACF en PACF te beschouwen, namelijk:

- 1 met een beperkt doorlopende X-as, tot een tijdsverschuiving van hooguit 20 of 30 tijdseenheden. Want als er sprake is van een autocorrelatiestructuur zal deze vooral tot uiting komen voor de kleinste tijdsverschuivingen. Dit is beter visueel te beoordelen als de X-as niet teveel tijdsverschuivingen omvat;
- 2 met een sterk verlengde X-as, tot een tijdsverschuiving van enigszins meer dan een jaar. Zo kan namelijk worden nagegaan of er ook autocorrelatie op jaarbasis optreedt. Dit kan wijzen op een niet verdisconteerd seizoenseffect of op niet-lineariteit.

Statistisch toetsen op autocorrelatie modelresidu

De visuele beoordeling van de combinatie van ACF en PACF zal doorgaans al voldoende uitsluitel geven over een eventuele autocorrelatiestructuur van het modelresidu. Maar daarnaast kan daar desgewenst ook nog formeel op worden getoetst. Daarvoor zijn bijvoorbeeld de volgende statistische toetsen beschikbaar:

- de Ljung-Box-toets [Ljung and Box, 1978], ook wel aangeduid als de Portmanteau-toets. Deze gaat er van uit dat het modelresidu een normale kansverdeling volgt en dat zijn tijdreeks equidistant is, zonder ontbrekende waarden. De toetsingsgrootte is een som van de gekwadrateerde autocorrelatiecoëfficiënten van de modelresiduen. Als de modelresiduen afkomstig zijn uit een normale verdeling zal de toetsingsgrootte afkomstig zijn uit een χ^2 -verdeling¹;
- een aangepaste versie daarvan, die overweg kan met ontbrekende waarden in de reeks van het modelresidu [Stoffer and Tolo, 1992];
- de runstoets, die een geschikt alternatief is voor de Ljung-Box-toets als de modelresiduen niet afkomstig zijn uit een normale verdeling. Het is een verdelingsvrije toets, met als toetsingsgrootte het aantal runs. Een run is daarbij een aaneengesloten groep waarden die alle aan dezelfde kant van de reeksmediaan liggen. Deze toets kan ook overweg met ontbrekende waarden van het modelresidu.

Deze statistische toetsen worden doorgaans uitgevoerd met 95 % betrouwbaarheid (α , het risico op onterecht verwerpen van de nulhypothese, bedraagt dan 5 %).

Als het modelresidu geen witte ruis vormt en er ook geen verbetering optreedt met toevoegen van invoerreeksen en/of gebruik van andere transferfuncties, dient een uitgebreider ruismodel te worden toegepast, of moet worden gemodelleerd op een grotere tijdsbasis. Dit laatste kan op de volgende manieren:

- door elke reeks te aggregeren tot een reeks van gemiddelden, dus bijvoorbeeld een reeks meetwaarden die om de 24 uur zijn genomen te aggregeren tot een reeks week-, tweeweeks-, of maandgemiddelden;
- door de uitvoerreeks uit te dunnen, dus bijvoorbeeld een reeks meetwaarden die om de 24 uur zijn genomen uit te dunnen tot een reeks meetwaarden die om de 14 dagen zijn genomen.

In tijdreeksmodel verdisconteren van niet-lineaire relatie

Als een tijdreeksmodel het neerslagoverschot als invoerreeks heeft, zou daarmee het seizoenseffect verdisconteerd moeten zijn. Als er dan toch nog autocorrelatie blijkt te zijn voor modelresiduen die een jaar of een veelvoud daarvan uit elkaar liggen, kan het wijzen op een niet-lineaire relatie tussen het neerslagoverschot en de grondwaterstand, die niet afdoende kan worden beschreven met het model.

Variantie modelresidu versus tijd en modelwaarde

Om te kunnen voldoen aan de randvoorwaarde van witte ruis dienen de modelresiduen ook homoscedastisch te zijn, wat inhoudt dat hun variantie niet verandert met de tijd of met de modelwaarde. Als dat wel het geval is spreken we van heteroscedasticiteit.

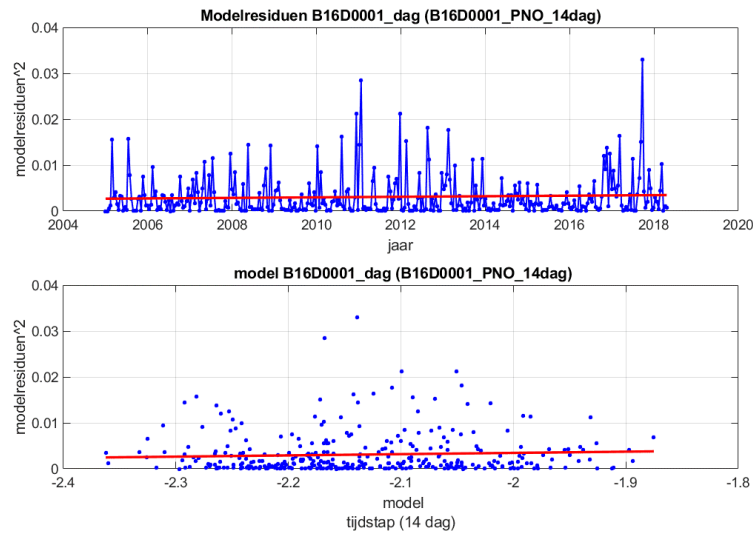
Er kan op heteroscedasticiteit worden getoetst, zoals met de Engle-toets of de Breusch-Pagan-toets:

- de Engle-toets [Engle, 1982] is gebaseerd op de eigenschap dat als de modelresiduen a_t heteroscedastisch zijn, de reeks gekwadrateerde modelresiduen (a_t^2) autocorrelatie zal vertonen. De toets komt neer op nagaan of de helling van een lineaire regressielijn door de reeks gekwadrateerde modelresiduen al of niet statistisch significant is;
- de Breusch-Pagan-toets [Breusch and Pagan, 1979] gaat na of de variantie van het modelresidu a_t gerelateerd is aan de modelwaarden.

In de twee plots van Afbeelding 4.17 zijn de gekwadrateerde modelresiduen uitgezet. In de eerste plot zijn deze uitgezet tegen de tijd en in de tweede tegen de bijbehorende waarden van het BJTR-model. In beide plots is ook de lineaire regressielijn weergegeven (rode lijn), die het gekwadrateerde modelresidu beschrijft als functie van de tijd (bovenste plot), respectievelijk de modelwaarde (onderste plot).

¹ Uit te spreken als Chikwadrat-verdeling.

Afbeelding 4.17 Plot van de gekwadrateerde modelresiduen als functie van de tijd (boven) en als functie van de bijbehorende waarde van het BJTR-model (onder). De rode lijn in de bovenste plot is de lineaire regressielijn van gekwadrateerd modelresidu en tijd en in de onderste plot van gekwadrateerd modelresidu en modelwaarde. De bovenste plot hoort bij de Engle-toets en de onderste hoort bij de Breusch–Pagan-toets, allebei toetsen op heteroscedasticiteit

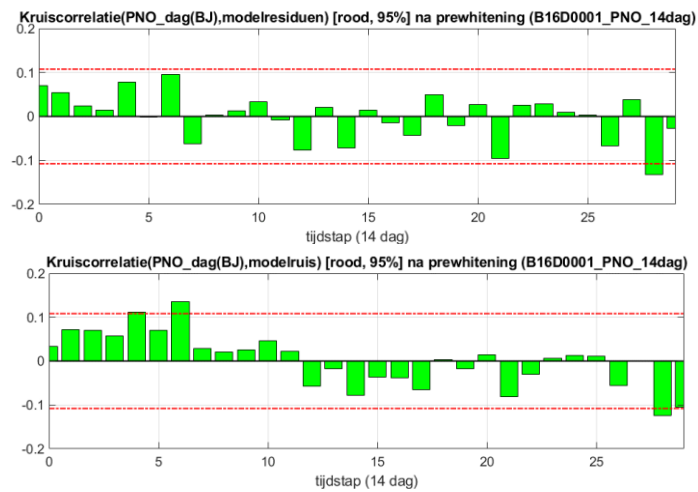


4.4.3 Beoordelen kruiscorrelaties modelresidu/modelruis en invoerreeksen

Als er sprake blijkt van een kruiscorrelatie tussen het modelresidu (a_t) en een invoerreeks, of tussen de modelruis (N_t) en een invoerreeks, is dat een aanwijzing dat het model nog onvoldoende de dynamische relatie met de betreffende invoervariabele beschrijft. Ga in zo'n geval na of dit kan worden verbeterd door de vorm van de transferfunctie aan te passen. Anders wordt niet voldaan aan één van de randvoorwaarden om aan de hand van het model uitspraken met gekwantificeerde onzekerheden te kunnen doen, namelijk de randvoorwaarde dat de formulering van het deterministische deel van het tijdreeksmodel overeen moet stemmen met de werkelijke vorm. Verder zal het model bij verbetering ook geschikter worden om er bijvoorbeeld voorspellingen mee uit te voeren.

Afbeelding 4.18 toont voorbeelden van de kruiscorrelatie tussen het potentieel neerslagoverschot en het modelresidu en tussen het potentieel neerslagoverschot en de modelruis. Het betreft de modellering van een grondwaterstandreeks. Een visuele beoordeling van deze kruiscorrelatiefuncties geeft geen duidelijke aanleiding om de betreffende transferfunctie aan te passen.

Afbeelding 4.18 Voorbeelden van de kruiscorrelatie tussen het potentieel neerslagoverschot en het modelresidu (boven) en van de kruiscorrelatie tussen het potentieel neerslagoverschot en de modelruis (onder). Het betreft de modellering van een grondwaterstandreeks

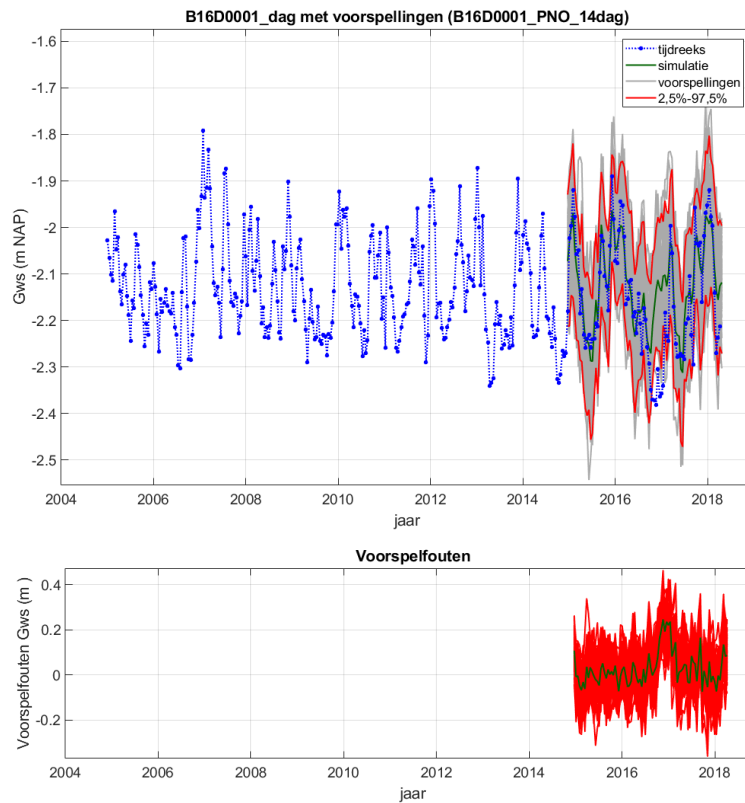


4.4.4 Valideren model met onafhankelijke meetwaarden

Valideren is het nagaan of het model voldoet aan zijn gebruiksdoel. Daartoe worden modeluitkomsten vergeleken met waarden die niet betrokken waren bij het kalibreren van het model. Deze laatste kunnen worden verkregen door een deel van de uitvoerreeks apart te houden, of door aselekt een aantal meetwaarden van de uitvoerreeks te trekken en die apart te houden.

Hieronder een voorbeeld van valideren, waarbij een deel van de uitvoerreeks apart is gehouden (Afbeelding 4.19). Het betreft een model dat de grondwaterstand beschrijft als functie van het potentieel neerslagoverschot en dat moet dienen om toekomstige grondwaterstanden te voorspellen bij een bepaald ontwikkelingsscenario van het potentieel neerslagoverschot. Beide reeksen zijn beschikbaar op dagbasis, waarbij de grondwaterstandreeks loopt van 1 januari 2005 t/m 18 april 2018. Het potentieel neerslagoverschot start eerder (1 januari 1995) en eindigt later (10 januari 2021). Het te valideren model is gekalibreerd over de tienjaarsperiode 2005 t/m 2014. De modellering is op 14-daagse basis, waartoe de reeksen door middeling zijn omgezet van dagbasis naar tweewekenbasis. Ter validering van het model is vervolgens het verdere verloop van de grondwaterstand voorspeld van 2015 tot het eind van de beschikbare grondwaterstandreeks (18 april 2018) en vergeleken met de over die periode beschikbare metingen van de grondwaterstand (Afbeelding 4.19 en Afbeelding 4.20).

Afbeelding 4.19. Voorbeeld van het valideren van een tijdreeksmodel met een apart gehouden deelreeks. Bovenste subplot: tijdreeks (blauw), deterministische reeksvoorspelling (groen), 100 stochastische reeksvoorspellingen (grijs) en grenzen van 95 %-voorsplinterval (rood). Onderste subplot: reeks van deterministische voorspelfouten (groen) en 100 reeksen van stochastische voorspelfouten (rood)



Afbeelding 4.19 toont in de bovenste subplot:

- blauw - de tijdreeks van de grondwaterstand van januari 2005 t/m april 2018;
- groen - de deterministische reeksvoorspelling van de grondwaterstand vanaf 2015 (aangeduid als 'simulatie'), gebaseerd op de voor die periode beschikbare waarden van PNO;
- grijs - een bundel van 100 stochastische reeksvoorspellingen van de grondwaterstand¹;
- rood - de grenzen van het 95 %-voorspelinterval, zoals per tijdstip bepaald uit het 2,5-percentiel en het 97,5-percentiel van de daarvoor voorspelde waarden.

Uit deze subplot blijkt dat de gerealiseerde grondwaterstand in de periode midden 2016 tot midden 2017 voor een relevant deel onderin of zelfs buiten het 95 %-voorspelinterval ligt van het model dat is gekalibreerd over de periode 2005 t/m 2015.

En Afbeelding 4.19 toont in de onderste subplot:

- groen - de reeks van de deterministische voorspelfout van de grondwaterstand vanaf 2015;
- rood - de bundel van 100 reeksen stochastische voorspelfouten van de grondwaterstand.

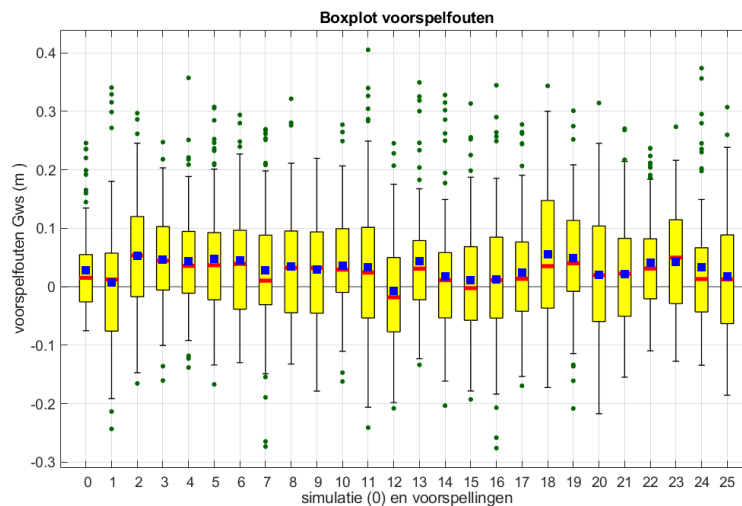
Uit deze subplot blijkt dat de deterministische voorspelfout in de periode midden 2016 tot midden 2017 kan oplopen tot circa 25 cm.

¹ Bij stochastisch voorspellen worden alle modeleigenschappen verdisconteerd die voorspellingen onzeker maken. Dit betreft: i) de onzekerheden van de geschatte modelparameters, ii) de onzekerheid van de modelruis en iii) het gecorreleerd zijn van de parameterschattingen. Er zijn dan zeer veel toekomstige verlopen mogelijk. Daarom wordt een voorspelbundel gegenereerd, bestaande uit een groot aantal van die mogelijke toekomstige verlopen. Dit is nader toegelicht in § 6.2.2.

Op basis van deze validatie moet worden uitgezocht wat de grote voorspelfouten veroorzaakt, alvorens te besluiten de modellering aan te passen of het model als onbruikbaar te kwalificeren. Zo kan er bijvoorbeeld sprake zijn van nog een andere invloedsfactor, of van disfunctioneren van de logger.

Afbeelding 4.20 toont per voorspelling de boxplot van de voorspelfout (zie hieronder). De eerste boxplot (met 0 op de X-as) betreft de voorspelfout van de deterministische voorspelling. De overige 25 boxplots betreffen de voorspelfouten van 25 willekeurige stochastische reeksvoorspellingen uit de bundel van 100. De meeste van deze boxplots duiden op een scheve verdeling van de voorspelfout, met veel positieve uitschieters. Verder is bij de meeste zowel het gemiddelde (blauw vierkant) als de mediaan (rode streep) van de voorspelfout groter dan nul. Ook dit wijst op het slechter functioneren van het BJTR-model over deze periode.

Afbeelding 4.20. Boxplots van de voorspelfout. De eerste boxplot (met 0 op de X-as) betreft de voorspelfout van de deterministische voorspelling. De overige 25 boxplots betreffen de voorspelfouten van 25 stochastisch voorspelde reeksen, die willekeuring zijn getrokken uit de bundel van 100 stochastisch voorspelde reeksen



[Knotters, 2012] adviseert om de belangrijkste validatiecriteria af te stemmen op het gebruiksdoel van het model. Waar het gaat om een tijdreeksmodel van de grondwaterstand komen bijvoorbeeld de volgende gebruiksdoelen in aanmerking:

- het kwantificeren van een bepaalde invloed op de grondwaterstand;
- het voorspellen van toekomstige grondwaterstanden
- het bepalen van overschrijdingskansen en/of GxG's middels het simuleren van grondwaterstandreeksen bij een scenario voor de ontwikkeling van het neerslagoverschot.

Voor het voorspellen komen als validatiecriteria in aanmerking grenswaarden voor kengetallen van de verzameling verschillen tussen onafhankelijke meetwaarden en de bijbehorende modelwaarden (zoals gemiddelde en standaardafwijking). En voor het bepalen van overschrijdingskansen middels simuleren komen in aanmerking maten voor de kwaliteit van geschatte overschrijdingsfrequenties.

5

ONTWIKKELEN PIRFICT-MODEL

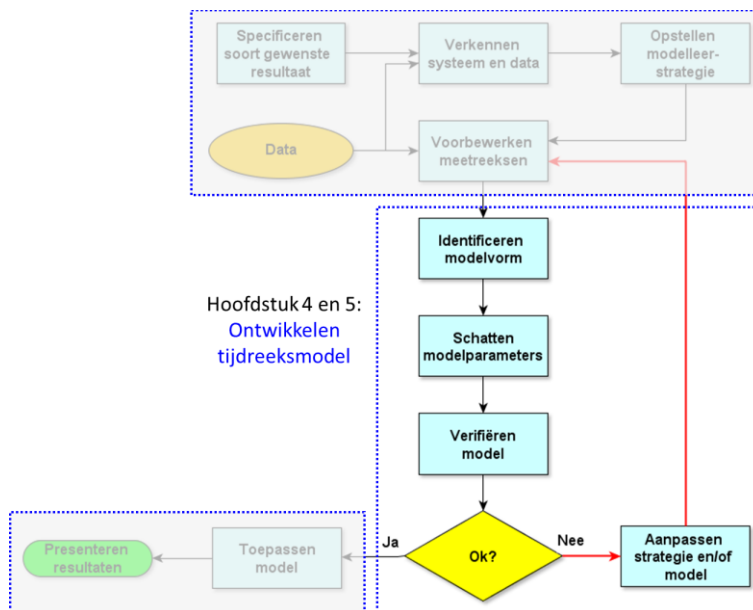
5.1 Inleiding

Dit hoofdstuk beschrijft de fase van het ontwikkelen van een PIRFICT-model. Deze bestaat uit de stappen:

- identificeren modelvorm;
- schatten modelparameters;
- verifiëren model.

Deze stappen worden zo nodig herhaald doorlopen, tot het model voldoet aan bepaalde randvoorwaarden.

Afbeelding 5.1 Stroomschema voor het uitvoeren van tijdreeksanalyse, met uitgelicht het ontwikkelen van het tijdreeksmodel



Ruis en residu

In dit hoofdstuk (PIRFICT-methode) heeft de term 'residu' $(N(t))$ betrekking op de stochastische component van de uitvoerreeks en de term 'ruis' $(a(t))$, op de stochastische invoerreeks en het deel van het residu wat niet tot de systeemfout behoort (idealiter betreft dit de 'witte ruis'), zie hiervoor ook afbeelding 2.4.

5.2 Identificeren modelvorm

5.2.1 Welke invloeden zijn van belang

Voordat een tijdreeksmodel wordt opgesteld is het aan de gebruiker om de vorm van het tijdreeksmodel te bepalen aan de hand van de beschikbare gegevens en de hydrologische kennis van het systeem (zie ook paragraaf 9.2). Hierbij hoort ook de keuze van de verklarende reeksen die worden meegenomen in het tijdreeksmodel. Meestal wordt in elk geval gebruik gemaakt van de neerslag en verdamping als verklarende reeksen (zie ook paragraaf 3.4.4). Afhankelijk van het systeem kunnen hier extra reeksen aan worden toegevoegd zoals bijvoorbeeld een onttrekking of oppervlaktewaterstand. Welke reeksen moeten worden meegenomen is niet altijd op voorhand vast te stellen en kan dan slechts achteraf worden bepaald. Zo is het niet altijd vooraf te bepalen of een onttrekking op 5 km afstand wel of niet van invloed is op de tijdreeks en dus wel of niet dient te worden meegenomen in het tijdreeksmodel.

Om de invloed van een bepaalde reeks te kunnen bepalen in de tijdreeksanalyse is het van belang dat de invoerreeds voldoende varieert. Bij een lage variatie van de invloed is het niet mogelijk om met behulp van tijdreeksanalyse het effect hiervan te bepalen. Als bijvoorbeeld een onttrekking een constant debiet heeft gedurende de meetperiode van de grondwaterstand kan het effect hiervan op de grondwaterstand niet worden gesimuleerd.

5.2.2 Hoe het effect van neerslag en verdamping kan worden meegenomen

Het effect van neerslag en verdamping wordt bij een groot deel van de tijdreeksmodellen meegenomen als invloedsreeks op de stijghoogte. Hierbij is dit voor een groot deel van de tijdreeksmodellen de dominante invloedsreeks. De neerslag en verdamping kunnen op meerdere manieren worden meegenomen in het tijdreeksmodel zoals beschreven in paragraaf 3.4.4.

Voor een PIRFICT-model kunnen de neerslag en verdamping worden meegenomen met een niet-lineair model (zie ook paragraaf 3.4.1). Hierbij kunnen niet-lineaire interacties die plaatsvinden tussen de gemeten grondwaterstand en de grondwateraanvulling worden meegenomen, bijvoorbeeld processen in de wortelzone ten tijde van droogte. Er zijn verschillende niet-lineaire modellen beschikbaar, zie bijvoorbeeld Berendrecht et al. (2005), Peterson en Western (2014) en Collenteur et al. (2020).

5.2.3 keuze van de responsfunctie

Zoals beschreven in paragraaf 2.4.3 maakt het PIRFICT-model gebruik van een responsfunctie om de relatie tussen een invoerreeds en de gemeten grondwaterstand te simuleren. Hierbij wordt de vorm van de respons ten gevolge van bijvoorbeeld neerslag en verdamping vooraf door de opsteller van het tijdreeksmodel vastgelegd. Voor elke invoerreeds (bijvoorbeeld neerslag, verdamping, een onttrekking) is het mogelijk om een andere responsfunctie te gebruiken.

Er wordt vaak gebruik gemaakt van de geschaalde Gammafunctie (zie paragraaf 2.4.3 en 2.6.2). Naast deze Gamma-responsfunctie zijn er nog enkele responsfuncties die binnen het tijdreeksmodelleren ook vaak worden gebruikt. In Tabel 5.1 is een overzicht weergegeven van de verschillende invloeden met de veelgebruikte responsfuncties voor deze invloeden. Dit zijn slechts suggesties en geen wetmatigheden. Per tijdreeksmodel moet gekeken worden of deze relaties passen bij de beschikbare gegevens.

Tabel 5.1 Overzicht van invloedtypen die onderscheiden worden binnen de PIRFICT-benadering, met hun eigenschappen. Bij de bijbehorende responsfuncties gaat het om een standaardkeuze zoals die bijvoorbeeld in het programma Menyanthes gehanteerd wordt (von Asmuth e.a., 2012), en geen wetmatigheid

Invloed	Eenheid	Dimensie	Type	Effect	Responsfunctie
neerslag	[LT ⁻¹]	2D	Flux	+	Gamma respons
verdamping	[LT ⁻¹]	2D	Flux	-	Gamma respons/verdampingsfactor
onttrekking	[L ³ T ⁻¹]	1/2D	Flux	-	Putformule Hantush
oppervlaktewater	[L]	1/2D	Stand	+	Polderfunctie Bruggeman
luchtdruk	[FL ⁻²]	2D	Druk	+	Barometrische efficiëntie
ingreep	[Div.]	Diverse	Diverse	+/-	Diverse

Naast deze veelgebruikte responsfuncties zijn er meer opties beschikbaar voor het opstellen van het PIRFICT-model, per software zal de beschikbaarheid van deze responsfuncties mogelijk verschillen. De Exponentiele responsfunctie is een versimpelde versie van de Gamma-respons, waarbij de parameter n gelijk is aan 1 (zie paragraaf 2.4.3). Het voordeel van deze responsfunctie is dat het aantal parameters van de respons kleiner is (2 i.p.v. 3), waardoor minder parameters moeten worden geoptimaliseerd door het tijdreeksmodel wat kan zorgen voor een lagere onzekerheid. Een nadeel van deze respons kan zijn dat door de versimpelde vorm de relatie tussen de grondwaterstand en de invoerreeks niet juist kan worden gesimuleerd. Daarnaast bestaat ook een 4-Parameter responsfunctie. Hierbij wordt een extra term toegevoegd aan de exponent van de Gammaverdeling, de responsfunctie krijgt dan 4 in plaats van 3 parameters. Deze respons kan door de extra vrijheidsgraad andere vormen aannemen dan de Gamma-respons (en dus een andere relatie simuleren), maar heeft als nadeel dat het toevoegen van extra parameters kan leiden tot een hogere onzekerheid in de modelparameters.

Naast deze standaard-responsfuncties is het mogelijk om een casus specifieke responsfunctie te gebruiken, bijvoorbeeld een responsfunctie die is afgeleid op basis van de hydrologische kennis van een gebied of op basis van eerder uitgevoerde onderzoeken (denk hierbij bijvoorbeeld aan een pompproef).

De keuze van de responsfunctie ligt bij de opsteller van het tijdreeksmodel. Voor verschillende casussen zullen verschillende responsfuncties verschillende resultaten als gevolg hebben. Op basis van hydrologische kennis kan voor het opstellen van een tijdreeksmodel een inschatting worden gemaakt van de responsfunctie die past bij de invoerreeks en de gemeten grondwaterstand.

Als uit de modelresultaten blijkt dat de inschatting voor de respons niet de juiste is, kan deze eventueel worden bijgesteld. Een voorbeeld hiervan kan zijn dat voor een neerslag en verdampingsreeks de Gamma-functie gebruikt is als respons voor deze reeksen. Als uit de modelresultaten blijkt dat de geoptimaliseerde responsfunctie een waarde voor de n parameter heeft van 1, suggereert dit dat de respons lijkt op de Exponentiele responsfunctie. Als het model met deze responsfunctie vergelijkbare resultaten geeft is deze wellicht beter geschikt omdat het aantal modelparameters voor de exponentiele responsfunctie lager is. Een ander voorbeeld is, wanneer initieel gekozen is voor de Exponentiele responsfunctie, dat bij beoordeling van de modelresultaten blijkt dat het tijdreeksmodel met piekneerslag sneller reageert dan bij de metingen is waargenomen, gekozen kan worden voor de Gamma-responsfunctie. Hierbij moet wel worden opgemerkt dat het doel hierbij is om het hydrologische systeem zo goed mogelijk te simuleren en niet om de beste modelfit te verkrijgen.

5.2.4 Hoe om te gaan met veranderende hydrologische systemen

Het is mogelijk dat gedurende de meetperiode van een of meerdere peilbuizen het hydrologische systeem verandert (zie ook paragraaf 3.4.5). Voorbeelden hiervan zijn: het baggeren van een kanaal, het verhogen van het oppervlaktewaterpeil en het dempen van sloten. Hier zal bij het opstellen van een tijdreeksmodel rekening mee moeten worden gehouden. Er zijn drie methoden om dit mee te nemen in een PIRFICT-tijdreeksmodel:

- het toevoegen van een staptrend;
- het toevoegen van een lineaire trend;
- het veranderen van een responsfunctie met de tijd.

Het toevoegen van een staptrend wordt doorgaans gebruikt wanneer in een relatief korte periode een verandering heeft plaatsgevonden. Denk hierbij bijvoorbeeld aan een peilopzet. De staptrend kan in het model worden gesimuleerd met verschillende responsfuncties net als elke andere invloed, dit kan bijvoorbeeld zijn een instantane of een exponentiele respons. Per situatie zal de prestatie van deze responsfunctie verschillen, op basis van de kennis van het hydrologische systeem en de ingreep kan deze worden gekozen. Als er bijvoorbeeld sprake is van een peilopzet in een waterloop kan deze als instantane invloed worden gemodelleerd. Als over een langere periode werkzaamheden zijn geweest, zoals het dempen van sloten, past een andere responsfunctie waarschijnlijk beter, bijvoorbeeld een exponentiele respons. Voor een uitgewerkt voorbeeld zie paragraaf 9.2 en 10.4. Naast de responsfunctie kan met behulp van het tijdreeksmodel ook het moment van optreden van de staptrend worden geoptimaliseerd, dit kan nuttig zijn als van werkzaamheden niet meer exact bekend is wanneer deze zijn uitgevoerd. Hierbij moet worden opgemerkt dat dit nog een extra parameter aan het model toevoegt, waardoor de resultaten van het model onzekerder kunnen worden. Ook is er een substantieel risico dat de interventiereeks wordt gekneed naar een toevallige verandering van de uitvoerreeks, wat neerkomt op curvefitting. Het afstemmen van een invoerreeks op visuele verkenning van de uitvoerreeks introduceert een sterk subjectieve lading en de modelresultaten hebben dan minder zeggingskracht.

De mogelijkheid bestaan ook om een lineaire trend mee te simuleren in een tijdreeksmodel (zie ook paragraaf 3.4.7), zoals bij het veranderen van vegetatietypes. Een lineaire trend kan worden geoptimaliseerd door het tijdreeksmodel, hierbij wordt de gradiënt van de trend meegenomen als modelparameter. Het toevoegen van een trend moet alleen worden toegepast als hiervoor een fysische/hydrologische verklaring voor is en mag niet worden gebruikt om het model te 'fitten'. In zo'n geval moet onderzocht worden welke invloed mogelijk mist in het tijdreeksmodel.

Daarnaast bestaat de mogelijkheid dat de respons van de grondwaterstand op externe invloeden verandert. Dit kan bijvoorbeeld komen door het dichtslibben van een watergang (Oberghell, 2019). Het is mogelijk om een PIRFICT-model op te stellen met een responsfunctie die in de tijd verandert.

De voorbeelden van veranderingen in het systeem zoals hierboven beschreven zijn niet altijd direct zichtbaar in de meetreeksen of bekend bij de gebruiker van tijdreeksanalyse. Het is daarom noodzakelijk dat bij het beoordelen rekening wordt gehouden met de aanwezigheid van deze veranderingen. Door de residuen van het model te observeren kan een mogelijk verandering worden geïdentificeerd en de modelstrategie hierop worden aangepast (indien hier ook een hydrologische verklaring voor bestaat). Hierbij is het streven dat de ruis ($a(t)$) geen informatie meer bevat en witte ruis vormt (zie ook paragraaf 2.5.1).

5.2.5 Toepassen ruismodel

Het ruismodel is een onderdeel van het PIRFICT-tijdreeksmodel. Het ruismodel kan worden gebruikt om statistische uitspraken te doen met het model. Bij het residu (verschil tussen de meetreeks en het tijdreeksmodel) van het tijdreeksmodel is meestal sprake van correlatie. Concreet betekent dit dat het residu van afhangt van het residu van de vorige dag. Bijvoorbeeld, als het op de locatie van de peilbuis gisteren heeft geregend terwijl op de locatie van het regenstation geen regen is gemeten, wordt in het tijdreeksmodel een te hoge stijghoogte gesimuleerd. Het duurt een aantal dagen (of zelfs weken) voordat dit effect is uitgedempt, dus zal het model waarschijnlijk vandaag ook een te hoge stijghoogte simuleren.

Op deze manier ontstaat deze correlatie in de residuen. Met autocorrelatie in deze reeks kunnen geen statistische uitspraken worden gedaan met het tijdreeksmodel.

Een ruismodel kan toegevoegd worden aan een tijdreeksmodel om te trachten de gecorreleerde residuen te transformeren naar ruis die ongecorrleerd is. Het eenvoudigst mogelijke ruismodel is een zogenaamd autoregressief model van orde 1, of kortweg een AR1-model. In dit model is het residu $\varepsilon(t)$ op tijdstip t gelijk aan een factor ρ keer het residu op tijdstip $t - \Delta t$ plus een random (ongecorreleerde) fout, $a(t)$, of in formule vorm:

$$\varepsilon(t) = \rho \varepsilon(t - \Delta t) + a(t)$$

Als de tijd Δt tussen twee residuën varieert, dan kan de factor ρ een functie zijn van Δt die afneemt naarmate de tijdstap Δt tussen twee metingen groter wordt. De formule is dan:

$$\varepsilon(t) = e^{-\Delta t / \alpha} \cdot \varepsilon(t - \Delta t) + a(t)$$

waarbij α een parameter is die aangeeft hoe snel de correlatie tussen residuën afneemt als Δt toeneemt. De correlatie is verwaarloosbaar als $\Delta t > 3 \alpha$ (want $e^{-3} \approx 0.05$). De parameter α wordt geoptimaliseerd bij het optimaliseren van het tijdreeksmodel. De rest term, $a(t)$, van de functie beschrijft de ruis van het model. Op basis van de resulterende ruis kan worden gekeken of het model gebruikt kan worden om statistische uitspraken te doen, zie paragraaf 5.4.2. Hieruit blijkt ook dat bij een grotere tijdstap de autocorrelatie minder is (zie paragraaf 3.4.2).

Het is ook mogelijk om het model op te stellen zonder het gebruik van een ruismodel. Het is dan niet mogelijk om statistische uitspraken te doen over de onzekerheden van parameters en voorspellingen.

5.3 Schatten modelparameters

Na het kiezen van de responsfuncties voor de verschillende invoerreeksen en het instellen van het ruismodel worden de modelparameters geschat, evenals hun standaardfouten. Het schattingsproces zoekt iteratief naar de combinatie van parameterwaarden die voldoet aan een bepaald criterium voor de mate waarin het model de uitvoerreeks beschrijft.

Naast de invloeden wordt over het algemeen ook een constante geoptimaliseerd in het tijdreeksmodel (zie ook paragraaf 2.4.1, waar deze constante deel uitmaakt van de algemene uitdrukking voor het tijdreeksmodel). Er kan gekozen worden om deze constante niet mee te fitten in het model, dit kan bijvoorbeeld van toepassing zijn als alle reeksen zijn genormaliseerd. Deze constante wordt soms aangeduid als de drainage basis, dit is echter niet per definitie correct. De parameter is het basis niveau van het model. Voor het bepalen van de waarde van het basisniveau zijn verschillende methoden. Een eerste optie is om na een simulatie het gemiddelde van de residuën te nemen als waarde voor de constante. Dit betekent dat de parameter niet geschat hoeft te worden tijdens de kalibratie. Een andere optie is om de constante mee te schatten als een parameter (Collenteur et al. 2019). Dit heeft als voordeel dat er ook iets gezegd kan worden over de onzekerheid van deze parameter. Zie hiervoor ook paragraaf 9.2.2.

De meest gebruikte methode om de optimale parameters te vinden is de minimalisatie van de som van de kwadraten van de verschillen tussen de gemeten stijghoogtes en de gemodelleerde stijghoogtes, ook wel de kleinste kwadraten methode genoemd. De som van de kwadraten is een niet-lineaire functie van de modelparameters. Om de optimale modelparameters te vinden wordt gebruikt gemaakt van een zoekmethode (zie paragraaf 9.3.1).

Er bestaan een aantal verschillende zoekmethoden die gebruikt kunnen worden om de optimale modelparameters te vinden. Iedere zoekmethode is gebaseerd op een ander zoek algoritme, maar het idee is altijd hetzelfde: het model wordt een aantal keer gedraaid met verschillende waarden van de modelparameters. Gebaseerd op de verschillende modeluitkomsten (en de daaruit voortvloeiende som van de kwadraten) bepaald het zoekalgoritme een volgende mogelijke set modelparameters die een kleinere

som van de kwadraten zou kunnen opleveren. Het zoekalgoritme gaat door totdat het besluit dat de beste parameters gevonden zijn.

De efficiëntie van het zoekalgoritme wordt bepaald door het aantal keer dat het model gedraaid moet worden om de optimale modelparameters te vinden. Het is vanzelfsprekend aan te bevelen om de initiële modelparameters zo dicht mogelijk bij de optimale modelparameters te kiezen. Daartoe heeft de meeste software geprobeerd om slimme keuzes te maken voor de initiële waarden van de modelparameters en de initiële parameters deels afhankelijk te maken van de waarden van de invoerreeksen. Soms heeft de modelleur specifieke kennis van het systeem waardoor betere initiële waarden gekozen kunnen worden.

De meest gebruikte zoekalgoritmes zoeken de optimale modelparameters door de modelparameters steeds een beetje aan te passen zodat de som van de kwadraten steeds een beetje kleiner wordt. Op deze manier leidt er een direct pad van de initiële waarden naar de optimale waarden en bij elk stapje langs het pad is de som van de kwadraten kleiner. Dit kan tot gevolg hebben dat het zoekalgoritme in een lokaal minimum belandt. Het is namelijk goed mogelijk dat langs het pad van de initiële parameter waarden naar de optimale parameter waarden de kwadraten som een tijdje toeneemt voordat het weer afneemt. Er bestaan geavanceerde zoekalgoritmes die het model ook draaien met modelparameters die juist een grotere kwadraten som geven met als doel om uiteindelijk het globale minimum te vinden. Dit kost vanzelfsprekend (veel) meer rekentijd.

Bij het optimaliseren is het nodig de startwaardes voor de modelparameters moeten zo worden gekozen dat convergentie wordt bereikt. Hiervoor wordt door de veelgebruikte software vaak automatisch een waarde voor aangenomen, bijvoorbeeld op basis van de gemiddelde waarde van de invoerreeks. Bij het opstellen van een tijdreeksmodel is het aan de opsteller van het model om de startwaardes voor de optimalisatie te bepalen (of controleren indien deze automatisch worden gekozen). Als convergentie niet wordt bereikt kan dit erop wijzen dat de startwaardes voor de kalibratie niet goed gekozen zijn.

Op basis van de resultaten van de kalibratie van de modelparameters kan het nodig zijn om de modelvorm, zoals deze is opgesteld in de voorgaande fase van de tijdreeksanalyse, te herzien. Op basis van de modelparameters kan bijvoorbeeld besloten worden om de gekozen responsfuncties aan te passen.

5.4 Verifiëren model

Na het optimaliseren van de modelparameters kan een fysische verificatie worden uitgevoerd op het model en de resulterende simulatie. Een eerste verificatie op de modelresultaten is om te kijken of de gesimuleerde grondwaterstand realistisch is. Hierbij kan de gesimuleerde grondwaterstand vergeleken worden met het maaiveldniveau nabij de peilbuis. Als de gesimuleerde grondwaterstand boven deze waarde uitkomt is de simulatie niet realistisch. Mogelijk is dan een aanpassing van de modelvorm nodig, zoals het toevoegen van een drempel-niet-lineariteit.

Daarnaast kan een tijdreeksmodel ook worden gevalideerd door middel van veldobservaties. Hierbij kan de gebiedskennis worden gebruikt om het tijdreeksmodel te toetsen. Als bijvoorbeeld in een zomer door het tijdreeksmodel een grondwaterstand is gesimuleerd die vergelijkbaar is met andere zomers maar in het veld extreme droogte is waargenomen, kan dit aanleiding geven om het tijdreeksmodel aan te passen. Mogelijk is er een sloot drooggevallen of reageert de grondwaterstand bij droogte niet meer lineair op verdamping. Het model dient dan mogelijk te worden aangepast.

Daarnaast kan naar de geoptimaliseerde modelparameters worden gekeken en of deze passen bij het hydrologische beeld. Een voorbeeld hiervan zou zijn dat de hydrologische verwachting is dat een peilbuis op grotere afstand tot een onttrekking een lagere invloed van deze onttrekking ondervindt dan een peilbuis die dichterbij staat, bij een uniforme ondergrond (hierbij kan gekeken worden naar de gain parameter van de respons, voor de Gamma functie is dit de parameter A, zie paragraaf 2.4.3). Of dat de grondwaterstand positief reageert op neerslag en negatief op verdamping. Als dit niet het geval is moet de modelvorm mogelijk worden aangepast, mogelijk zijn niet alle invloeden meegenomen in de tijdreeksanalyse of zijn de respons functies niet passend bij de invloedreeks. Een aantal van deze verificaties worden ook op voorhand

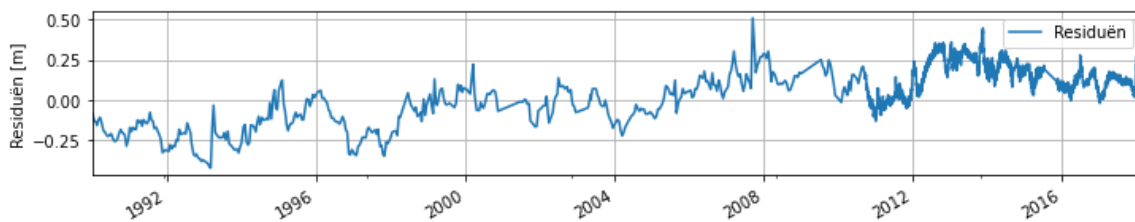
al gemaakt, bijvoorbeeld door het vastleggen van een positieve bijdrage van neerslag bij het opstellen van het tijdreeksmodel, dan is het niet nodig om dit te evalueren na het optimaliseren van het model.

5.4.1 Beoordelen modelresidu

Ga na of de tijdreeks van het modelresidu anomalieën vertoont. Als dat het geval is kan het tot vertekende schattingen van modelparameters hebben geleid. Hiervoor kan een visuele beoordeling van de geplote residueerks volstaan.

Soms blijkt een te modelleren uitvoerreeks een trend te vertonen die door geen van de beschikbare invoerreeksen kan worden verklaard. Dit is een aanwijzing dat er onvoldoende bekend is over het te modelleren systeem (zie ook paragraaf 5.2.4).

Afbeelding 5.2 Model met een trend in de residueerks (zie voorbeeld in paragraaf 9.2.8)



5.4.2 Beoordelen ruis

Bij het verifiëren moeten meerdere kenmerken van worden beoordeeld. Dit is nodig om te beoordelen of het gaat om witte ruis en bepaald of het model de parameter-onzekerheden goed heeft ingeschat. Voor een voorbeeld zie paragraaf 9.3 en 10.4 waarin deze beoordeling is uitgevoerd voor een PIRFICT-tijdreeksmodel.

Gemiddelde en standaardafwijking van modelruis

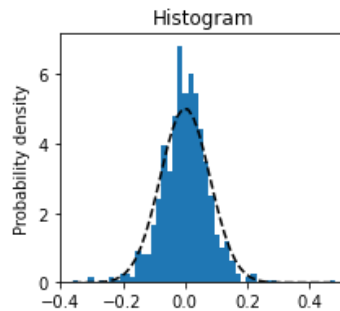
Als het gemiddelde van het modelruis statistisch significant afwijkt van nul zullen de modelvoorspellingen onzuiver zijn, wat betekent dat ze een systematische fout zullen vertonen. Het duidt op een onvolkomenheid bij het identificeren van het model. Als het model nog geen constante omvat, kan dit mogelijk worden opgelost door een constante in het model op te nemen (zie ook paragraaf 9.2.2). Anders moet worden getracht één of meer responsfuncties anders te kiezen.

De standaardafwijking van het modelruis wordt geschat en is op te vatten als een maat voor het deel van de variatie van de te modelleren reeks dat niet beschreven kan worden door het PIRFICT-model. Deze standaardafwijking is ook een maat voor de nauwkeurigheid waarmee het model één stap vooruit kan voorspellen. Want als een voorspelling één stap vooruit wordt gedaan, dan zal gemiddeld 95 % van deze voorspellingen een fout hebben die minder is dan tweemaal de standaardafwijking (er van uitgaande dat de ruis normaal verdeelde witte ruis vormt). En samen met de modelstructuur bepaalt deze standaardafwijking ook de nauwkeurigheid waarmee het model twee of meer stappen vooruit kan voorspellen.

Histogram van modelruis

Om te beoordelen of de modelruis afkomstig zijn uit een normale kansverdeling kan in principe worden volstaan met een visuele beoordeling van een histogram (zie Afbeelding 5.3) van de ruis, of andere grafische weergaven van de verdeling (zoals de Q-Q-plot), mits de analist daar enige ervaring mee heeft opgebouwd.

Afbeelding 5.3 Voorbeeld van een histogram van de ruis $a(t)$ van een PIRFICT-model (zie ook paragraaf 9.3). De zwarte lijn geeft aan wat de vorm zou zijn van een normale verdeling met het gemiddelde en de standaardafwijking van de ruis. Op basis van een visuele beoordeling van deze afbeelding kan worden aangenomen dat de ruis afkomstig is uit een kansverdeling die niet sterk afwijkt van de normale kansverdeling



Er zijn ook diverse statistische toetsen beschikbaar om te beoordelen op normaliteit, zoals D'Agostino-K2 (D'Agostino, 1986), Jarque-Bera (1980, 1981 en 1987) en Shapiro-Wilk (1965). Deze hebben echter een aantal tekortkomingen:

- bij kleine steekproeven (< 20) hebben ze slechts weinig onderscheidend vermogen¹ en kunnen ze alleen grote afwijkingen van normaliteit detecteren;
- bij grote steekproeven (> 250) hebben ze juist veel onderscheidend vermogen en kunnen dan afwijkingen van normaliteit detecteren die praktisch niet relevant zijn. Kleine afwijkingen zijn immers niet relevant omdat toepassingen die uitgaan van de randvoorwaarde van normaliteit in het geval van een grote steekproef vrij robuust zijn tegen afwijkingen van normaliteit. Er is alleen geen objectieve methode om vast te stellen tot welke afwijking van normaliteit die robuustheid nog opgaat, zodat dit vaak subjectief moet worden ingeschat, zoals op basis van visuele beoordeling van het histogram van de gegevens.

Als ervoor wordt gekozen te toetsen op normaliteit, is het raadzaam dit uit te voeren met een hoge betrouwbaarheid (zoals 99 %), omdat zelfs een geringe en praktisch niet-relevante afwijking van normaliteit al snel statistisch significant kan zijn door het grote aantal reekswaarden bij tijdreeksanalyse. Maar zelfs dan vergt het beoordelen op normaliteit nog pragmatisme en mag aan het oordeel op basis van visuele beoordeling van het histogram meer gewicht worden toegekend dan aan het resultaat van de statistische toets.

Als uit de beoordeling blijkt dat er sprake is van een relevant te achten afwijking van normaliteit van de ruis zijn er twee mogelijkheden dat probleem op te lossen, te weten:

- aggregeren van de tijdreeksen naar een grotere tijdseenheid (zoals van een enkele meting naar een week-, tweeweeks, of maandgemiddelde van de metingen);
- het aanpassen van de vorm van het tijdreeksmodel, bijvoorbeeld door de keuze van de invoerreeksen en de keuze van de responsfuncties.

¹ Onderscheidend vermogen is de kans om terecht de nulhypothese te verwerpen. Dit vermogen is mede een functie van de mate waarin de werkelijkheid afwijkt van de nulhypothese.

Autocorrelatiestructuur ruis a_t

Als de reeks van de ruis nog autocorrelatiestructuur vertoont, biedt dat een goede gelegenheid het model te verbeteren. Verder is het ontbreken van autocorrelatie van de ruis zelfs een vereiste als het model dient om uitspraken met gekwantificeerde onzekerheden te kunnen doen (zie daarvoor paragraaf 6.2.3). De ruis dient dan namelijk te voldoen aan de volgende randvoorwaarden:

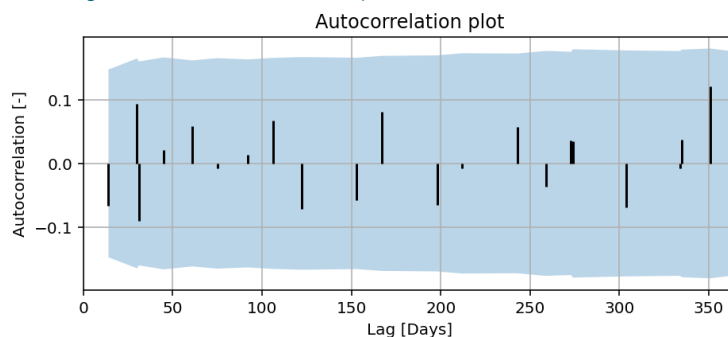
- 1 ze zijn gemiddeld nul;
- 2 ze zijn onafhankelijk van elkaar (er is dan geen sprake van een autocorrelatiestructuur);
- 3 ze hebben dezelfde variantie¹ (dit wordt aangeduid als *homoscedasticiteit*);
- 4 ze zijn afkomstig uit de normale kansverdeling (of een andere kansverdeling waarvan alle eigenschappen bekend zijn).

Als een reeks voldoet aan de eerste drie voorwaarden spreken we van *witte ruis*. Als deze ook voldoet aan normaliteit spreken we van *normaal verdeelde witte ruis*. Aanvullende randvoorwaarden, die echter vaak over het hoofd worden gezien en daardoor meestal impliciet worden aangenomen, zijn dat de formulering van het deterministische deel van het tijdreeksmodel overeen stemt met de werkelijke vorm (ook wel aangeduid als het *werkelijke systeem*) en dat de schatter van de evenwichtsrelatie tussen een invoerreeks en de uitvoerreeks zuiver is, dat wil zeggen zonder systematische fout.

Autocorrelatie is de correlatie van een tijdreeks met een in de tijd verschoven versie van zichzelf. Als er nog significante autocorrelatie aanwezig is in de ruis, dan wordt het betrouwbaarheidsinterval van de parameter te klein geschat (zie paragraaf 9.3 voor een voorbeeld). Deze vorm van correlatie kan worden weergegeven in een autocorrelatiediagram. Daarin is op de horizontale as de verschuiving in de tijd weergegeven (in het Engels: 'lag') en op de verticale as de berekende correlatie (zie Afbeelding 5.4).

De verschuivingen in de tijd (lags) moeten overeenkomen met de tijdstap van het model. Als er op dagbasis gerekend wordt, moet de reeks ook verschoven worden met (veelvouden van) die tijdstap voor het bepalen en beoordelen van de autocorrelatie (Collenteur, 2018).

Afbeelding 5.4 Voorbeeld autocorrelatie plot



Statistisch toetsen op autocorrelatie modelruis

De visuele beoordeling van de ACF zal doorgaans al voldoende uitsluitsel geven over een eventuele autocorrelatiestructuur van de ruis. Maar daarnaast kan daar desgewenst ook nog formeel op worden getoetst. Daarvoor zijn bijvoorbeeld de volgende statistische toetsen beschikbaar:

- de Ljung-Box-toets [Ljung and Box, 1978], ook wel aangeduid als de Portmanteau-toets. Deze gaat er van uit dat de ruis een normale kansverdeling volgt en dat zijn tijdreeks equidistant is, zonder ontbrekende waarden. De toetsingsgrootte is een som van de gekwadrateerde autocorrelatiecoëfficiënten van de ruis. Als de ruis afkomstig zijn uit een normale verdeling zal de toetsingsgrootte afkomstig zijn uit een χ^2 -verdeling²;
- een aangepaste versie daarvan, die overweg kan met ontbrekende waarden in de reeks van de ruis [Stoffer and Toloï, 1992];

¹ Dit impliceert dat de residuen kunnen worden gezien als trekkingen uit dezelfde kansverdeling.

² Uit te spreken als Chikwadrat-verdeling.

- de runstoets, die een geschikt alternatief is voor de Ljung-Box-toets als de ruis niet afkomstig is uit een normale verdeling. Het is een verdelingsvrije toets, met als toetsingsgrootte het aantal runs. Een run is daarbij een aaneengesloten groep waarden die alle aan dezelfde kant van de reeksmediaan liggen. Deze toets kan ook overweg met ontbrekende waarden van de ruis.

Deze statistische toetsen worden doorgaans uitgevoerd met 95 % betrouwbaarheid (α , het risico op onterecht verwerpen van de nulhypothese, bedraagt dan 5 %).

Als de ruis geen witte ruis vormt en er ook geen verbetering optreedt met toevoegen van invoerreeksen en/of gebruik van andere responsfuncties, dient een uitgebreider ruismodel te worden toegepast, of moet worden gemodelleerd op een grotere tijdsbasis. Dit laatste kan op de volgende manieren:

- door elke reeks te aggregeren tot een reeks van gemiddelden, dus bijvoorbeeld een reeks meetwaarden die om de 24 uur zijn genomen te aggregeren tot een reeks week-, tweeweeks-, of maandgemiddelden;
- door de uitvoerreeks uit te dunnen, dus bijvoorbeeld een reeks meetwaarden die om de 24 uur zijn genomen uit te dunnen tot een reeks meetwaarden die om de 14 dagen zijn genomen.

Variantie ruis versus tijd en modelwaarde

Om te kunnen voldoen aan de randvoorwaarde van witte ruis dient de ruis ook homoscedastisch te zijn, wat inhoudt dat hun variantie niet verandert met de tijd of met de modelwaarde. Als dat wel het geval is spreken we van heteroscedasticiteit.

Er kan op heteroscedasticiteit worden getoetst, zoals met de Engle-toets of de Breusch-Pagan-toets:

- de Engle-toets [Engle, 1982] is gebaseerd op de eigenschap dat als de modelruis a_t heteroscedastisch zijn, de reeks gekwadraterde ruis (a_t^2) autocorrelatie zal vertonen. De toets komt neer op nagaan of de helling van een lineaire regressielijn door de reeks gekwadraterde ruis al of niet statistisch significant is;
- de Breusch-Pagan-toets [Breusch and Pagan, 1979] gaat na of de variantie van de ruis gerelateerd is aan de modelwaarden.

5.4.3 Beoordelen op kruiscorrelaties ruis en invoerreeksen

Als er sprake blijkt van een kruiscorrelatie tussen de ruis en een invoerreeks is dat een aanwijzing dat het model nog onvoldoende de dynamische relatie met de betreffende invoervariabele beschrijft. Het is dan aan te bevelen om te kijken of alle invloeden zijn meegenomen in het tijdreeksmodel en of de gekozen responsfuncties passen bij de verschillende invoerreeksen.

6

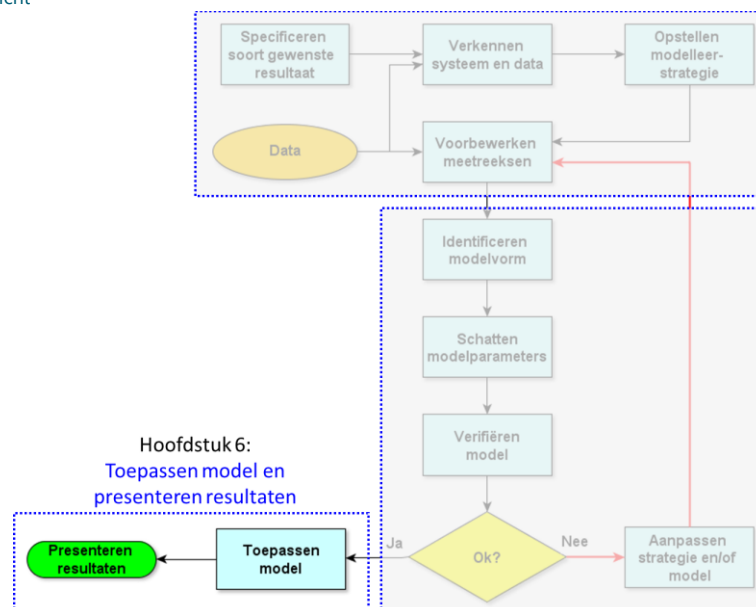
TOEPASSEN MODEL EN PRESENTEREN RESULTATEN

6.1 Inleiding

Dit hoofdstuk beschrijft de fase van het toepassen van het model en het presenteren van de resultaten:

- toepassen model (paragraaf 6.2);
- presenteren resultaten (paragraaf 6.3).

Afbeelding 6.1 Stroomschema voor het uitvoeren van tijdreeksanalyse, met toepassen model en presenteren van de resultaten uitgelicht



6.2 Toepassen model

Als het doel van de tijdreeksanalyse is om inzicht in de werking van het systeem te verkrijgen (categorie 1 uit paragraaf 3.2), en er geen extrapolaties naar andere omstandigheden nodig zijn, zullen de geschatte modelparameters en zo nodig ook hun gekwantificeerde onzekerheden daar vaak al voldoende informatie over bevatten en kan direct worden overgegaan tot het presenteren van de resultaten (zie daarvoor paragraaf 6.3).

Als er echter wél extrapolaties naar andere omstandigheden nodig zijn, of als het doel van de tijdreeksanalyse is om de uitvoer van het systeem te voorspellen of te beheren (categorieën 2 en 3 uit paragraaf 3.2), moeten er uitkomsten worden gegenereerd die verder gaan dan de resultaten van het kalibreren van het model. Daarbij moet echter wel rekening worden gehouden met het kalibratiegebied van het model, zoals toegelicht in paragraaf 6.2.1. En in paragraaf 6.2.2 wordt toegelicht hoe met een tijdreeksmodel voorspellingen kunnen worden gegenereerd.

6.2.1 Houd rekening met kalibratiegebied van tijdreeksmodel

Een tijdreeksmodel dat de verificatie goed doorkomt kan toepasbaar worden geacht voor de omstandigheden die opgingen in de periode waarover het model is gekalibreerd. Dat stelsel van omstandigheden kunnen we aanduiden als het *kalibratiegebied* van het model. Als het model slechts één invoervariabele heeft, is het kalibratiegebied het bereik van de betreffende variabele in de kalibratieperiode. En als het model n invoervariabelen heeft is het kalibratiegebied een n -dimensionale ruimte die de punten omvat die zijn gedefinieerd door de combinaties van waarden van de betreffende variabelen in de kalibratieperiode.

Een tijdreeksmodel kan (lineaire) relaties beschrijven die gelden in het kalibratiegebied. Maar er is geen garantie dat die relaties ook opgaan buiten het gebied waarvoor het model is gekalibreerd. Zo kan het gedrag van een hydrologisch systeem drempel-niet-lineariteiten vertonen, bijvoorbeeld door het droogvallen van waterlopen, of grondwaterstanden die het maaiveld bereiken. Het is daarom nodig om bij toepassing van een model eerst na te gaan of daarbij wellicht buiten het kalibratiegebied wordt gekomen.

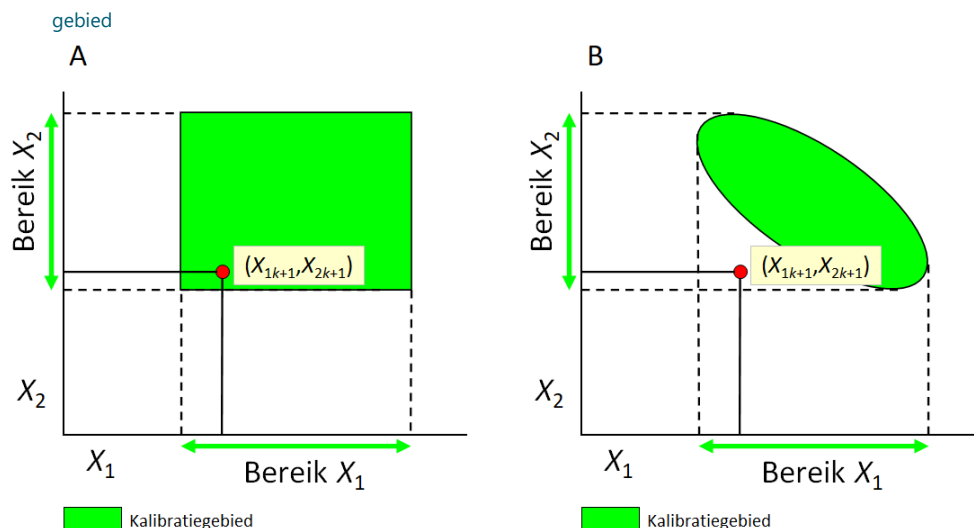
Hieronder wordt dit geïllustreerd voor het volgende meervoudig lineair regressiemodel met twee invoervariabelen en zonder intercept (het model gaat door de oorsprong):

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

waarin Y_i de waardenset van de uitvoervariabele, i de index ($i = 1 \dots k$), X_{1i} en X_{2i} de waardensets van de invoervariabelen, β_1 en β_2 regressieparameters en ε_i de waardenset van de modelfout.

Onderstaande Afbeelding 6.2 toont twee voorbeelden van een 2-dimensionaal kalibratiegebied, dat k punten (X_{1i}, X_{2i}) omvat die zijn gedefinieerd door de combinaties van waarden van X_1 en X_2 in de kalibratieset. In het voorbeeld links (A) ligt het (nieuwe) punt (X_{1k+1}, X_{2k+1}) in dat kalibratiegebied, zodat het model kan worden toegepast voor dat punt. Maar in het voorbeeld rechts (B, waar overigens sprake blijkt van enige negatieve correlatie tussen X_1 en X_2) ligt dat punt niet in het kalibratiegebied, ook al valt zowel zijn waarde voor X_1 als die voor X_2 binnen het bereik dat deze variabelen hebben in de kalibratieset. In principe is het model dan niet toepasbaar te achten voor dat punt.

Afbeelding 6.2 Twee voorbeelden van een 2-dimensionaal kalibratiegebied en de ligging van een nieuw punt ten opzichte van dat gebied



Het kan dus lastig zijn vast te stellen of een punt, gedefinieerd door een bepaalde nieuwe combinatie van waarden van de invoervariabelen, binnen het kalibratiegebied ligt. Want zelfs als van elke invoervariabele de waarde binnen het bereik van de kalibratieperiode valt, kan het punt gedefinieerd door al die waarden toch buiten het kalibratiegebied liggen.

In principe moet dus terughoudend omgegaan worden met het toepassen van een model voor omstandigheden die niet opgingen gedurende de kalibratieperiode. Maar soms is er geen andere keuze, met als voorbeelden:

- wanneer we met een tijdreeksmodel het verloop van de grondwaterstand willen voorspellen bij een bepaald extreem verloop van weersomstandigheden, of;
- als we met een tijdreeksmodel de evenwichtsrelatie tussen een grondwaterstand en een grondwaterwinning hebben vastgesteld en die lineair willen vertalen naar de verlaging van de grondwaterstand bij een verdubbeling van de winningshoeveelheid.

In beide gevallen zal bij het presenteren van de resultaten minstens aannemelijk moeten worden gemaakt dat het betreffende systeem geen (drempel-)niet-lineariteiten bevat. Dat zal dan moeten worden onderbouwd op basis van gebiedskennis.

6.2.2 Voorspellen met een BJTR-model

Om voorspellingen te kunnen genereren dient de invoer van het model zich niet alleen uit te strekken over de kalibratieperiode, maar ook over de voorspelperiode. Het gekalibreerde tijdreeksmodel moet dan worden gevoed met het toekomstige verloop van de invoer. Uiteraard is dat niet bekend en moet worden uitgegaan van één of meer scenario's voor het verloop. Zo zijn er voor lange termijn-voorspellingen van de grondwaterstand reeksen van het potentieel neerslagoverschot nodig voor verschillende klimaatscenario's.

Deterministisch voorspellen met BJTR-model

Er zijn verschillende mogelijkheden om voorspellingen te genereren met een BJTR-model, die verschillen in de mate waarin ze de modelonzekerheden verdisconteren. De eenvoudigste vorm is het deterministisch voorspellen, dat geen modelonzekerheid verdisconteert, omdat het alleen is gebaseerd op het deterministische deel van het BJTR-model. Een voorspelling komt hierbij voor elk tijdstip van de voorspelperiode tot stand door als invoer aan het gekalibreerde model te verstrekken de waarde van elke invoerreeks voor het betreffende tijdstip en zo nodig ook voorgaande tijdstippen (dit laatste als één of meer transferfuncties een MA(2)-structuur of hoger hebben). Het deterministisch voorspellen levert voor elk tijdstip dus één voorspelling op. Het is eenvoudig uit te voeren, maar het resultaat zal voor veel doeleinden minder geschikt zijn, doordat de variatie van de voorspelde reeks kleiner zal zijn dan wat in werkelijkheid te verwachten is (de ruis N_t is immers niet verdisconteerd in de voorspelling). Er kan dan bijvoorbeeld geen betrouwbaar te achten GHG of GLG uit worden bepaald.

Stochastisch voorspellen met BJTR-model

Een completere en realistischere vorm van voorspellen is het stochastisch voorspellen, aangezien dat alle modelonzekerheden verdisconteert. Gezien die onzekerheden wordt er een voorspelbundel gegenereerd, bestaande uit een groot aantal mogelijke toekomstige verlopen. Deze voorspelbundel komt tot stand met de volgende handelingen:

- 1 genereer een reeks van het modelresidu over de voorspelperiode, door voor elk tijdstip een aselechte trekking te doen uit de kansverdeling van het modelresidu (zie onderstaand tekstkader);
- 2 doe voor elke modelparameter een trekking uit zijn kansverdeling, maar zodanig dat de trekkingen van de verschillende modelparameters dezelfde onderlinge correlaties vertonen als hun schatters (deze onderlinge correlaties volgen uit de covariantiematrix van de geschatte parameters);
- 3 bereken met het model voor elk tijdstip een voorspelling van de uitvoerreeks door het model als invoer te verstrekken:
 - de waarde van elke invoerreeks voor het betreffende tijdstip en zo nodig ook voorgaande tijdstippen (hangt af van soort transferfunctie);
 - het in stap 1 voor dat tijdstip getrokken modelresidu en zo nodig ook van voorgaande tijdstippen (hangt af van soort ruismodel);
 - de in stap 2 getrokken modelparameters;
- 4 doorloop de voorgaande stappen 1 t/m 3 een groot aantal keer (zoals 500, 5.000 of meer), zodat er een bundel met veel voorspelde verlopen ontstaat.

Het stochastisch voorspellen levert voor elk tijdstip dus een groot aantal voorspellingen op, zodanig dat er een voorspelband ontstaat over de voorspelperiode. Uit die band kan bijvoorbeeld het 95 %-voorspelinterval worden afgeleid, door voor de ondergrens het 2,5-percentiel van de voorspellingen te nemen en voor de bovengrens het 97,5-percentiel.

Modelresidu als invoer bij het voorspellen

Het modelresidu vormt ook een invoerreeks van het tijdreeksmodel, zij het dat die niet bestaat uit gemeten waarden, maar uit waarden die pas ontstaan bij het modelleren, als verschil van de gemeten uitvoerreeks en de modelreeks daarvan. Zijn toekomstig verloop is uiteraard niet bekend, maar we kennen wel zijn statistische karakteristieken. De tijdreeksmodellering was er immers op gericht de modelresiduen terug te brengen tot willekeurige trekkingen uit een normale kansverdeling met een zo klein mogelijke standaardafwijking. Daar elke trekking uit die kansverdeling een mogelijke toekomstige realisatie van een modelresidu is, kan door vele trekkingen achter elkaar te doen een mogelijk toekomstig verloop van het modelresidu worden gesimuleerd en door middel van het ruismodel ook een mogelijk toekomstig verloop van de ruiscomponent van de uitvoerreeks. Het is overigens ook mogelijk die trekkingen te doen uit de verzameling modelresiduen van de modelkalibratie, om zo goed mogelijk aan te sluiten op hun empirische kansverdeling.

Omzetten tijdsbasis na voorspellen

Stel dat met tijdreeksmodellering van een grondwaterstandreeks de GHG en GLG moeten worden voorspeld bij een bepaald verloop van het neerslagoverschot en dat een BJTR-model is afgeleid op maandbasis, bijvoorbeeld vanwege de traagheid van het proces en/of om te kunnen voldoen aan de voorwaarde dat het modelresidu normaal verdeelde witte ruis vormt. Het model kan dan alleen de maandgemiddelde grondwaterstand voorspellen, terwijl voor het ramen van de GHG en GLG meetwaarden van de grondwaterstand nodig zijn voor de 14^e en 28^e van elke maand.

Dit kan worden bewerkstelligd door elke met het model voorspelde toekomstige maandgemiddelde grondwaterstand te vervangen door twee grondwaterstanden, namelijk één geldend voor de 14^e van die maand en één geldend voor de 28^e van die maand. Daarvoor moeten per halfmaandelijks meettijdstip trekkingen worden gedaan uit de historische verzameling afwijkingen tussen de gemeten grondwaterstanden op dat meettijdstip en de bijbehorende maandgemiddelden, zoals berekend over de modelkalibratieperiode. Zo kan bijvoorbeeld de voorspelde maandgemiddelde grondwaterstand van februari 2030 worden omgezet naar een voorspelde grondwaterstand op 14 februari 2030 door een willekeurige trekking te doen uit de over de modelkalibratieperiode berekende afwijkingen van de gemeten grondwaterstanden op 14 februari ten opzichte van de gemiddelde grondwaterstand in dezelfde februarimaand. De voorspelde grondwaterstand van 28 februari 2030 ligt daarmee ook meteen vast: die bevindt zich namelijk precies aan de andere kant van de voorspelde maandgemiddelde grondwaterstand van februari 2030 (een gemiddelde van twee waarden ligt immers exact in hun midden).

6.2.3 Voorspellen met een PIRFICT-model

Voor het voorspellen met een PIRFICT-model wordt gebruik gemaakt van de responsfuncties van het tijdreeksmodel (zie hiervoor ook het voorbeeld in paragraaf 10.4). Met behulp van deze responses kan de uitvoerreeks worden gegenereerd. Hiervoor zijn voorspellingen nodig voor de invoerreeksen van het tijdreeksmodel. Bij neerslag en verdamping kan hiervoor bijvoorbeeld gebruikt gemaakt worden van klimaatscenario's of weersvoorspellingen. Met behulp van deze invoerreeksen en de bijbehorende respons kan een voorspelling worden gemaakt. Omdat de responsfuncties continu zijn kan de uitvoer voor elke tijdstap worden voorspeld.

Bij het voorspellen met een tijdreeksmodel moet altijd worden gekeken naar het voorspelinterval. Hierbij moet worden gekeken naar de zekerheid van de afgeleide responsfuncties. Deze kan worden ingeschat met de onzekerheid van de modelparameters zoals deze zijn bepaald met het PIRFICT-model (de onzekerheden kunnen enkel goed worden ingeschat als het tijdreeksmodel voldoet aan de voorwaarden zoals beschreven in paragraaf 5.4.2). Het zekerheidsinterval van de respons kan worden bepaald door middel van Monte Carlo Sampling van de parameters van de responsfuncties. Hierbij worden willekeurige trekkingen gedaan uit de modelparameters, inclusief de bijbehorende onzekerheid, om een parameterset te genereren.

Met elk van deze parametersets wordt een responsfunctie opgesteld (op basis van de gekozen respons in het tijdreeksmodel, bijvoorbeeld de Gamma-respons), waaruit een voorspelinterval wordt afgeleid.

Met deze set van parameters en responsfuncties kunnen voorspellingen worden opgesteld met de voorspellingen voor de invoerreeks(en). Dit genereert een bandbreedte voor de voorspelling van de uitvoerreeks. Uit deze band kan bijvoorbeeld het 95 %-zekerheidsinterval worden afgeleid, door voor de ondergrens het 2,5-percentiel van de voorspellingen te nemen en voor de bovengrens het 97,5-percentiel. De procedure zoals hierboven beschreven kan indien gewenst herhaald worden voor verschillende voorspellingen voor de invoerreeksen (bijvoorbeeld verschillende klimaatscenario's).

6.3 Presenteren resultaten

6.3.1 Interpretatie en patroonanalyse

Ruimtelijke samenhang van de resultaten

Door de verbeterde mogelijkheden van geautomatiseerde verwerking zijn de inspanning en kosten die nodig zijn voor tijdreeksanalyse niet (PIRFICT-methode) of slechts weinig (BJTR-methode) afhankelijk van het aantal meetreeksen dat wordt geanalyseerd.

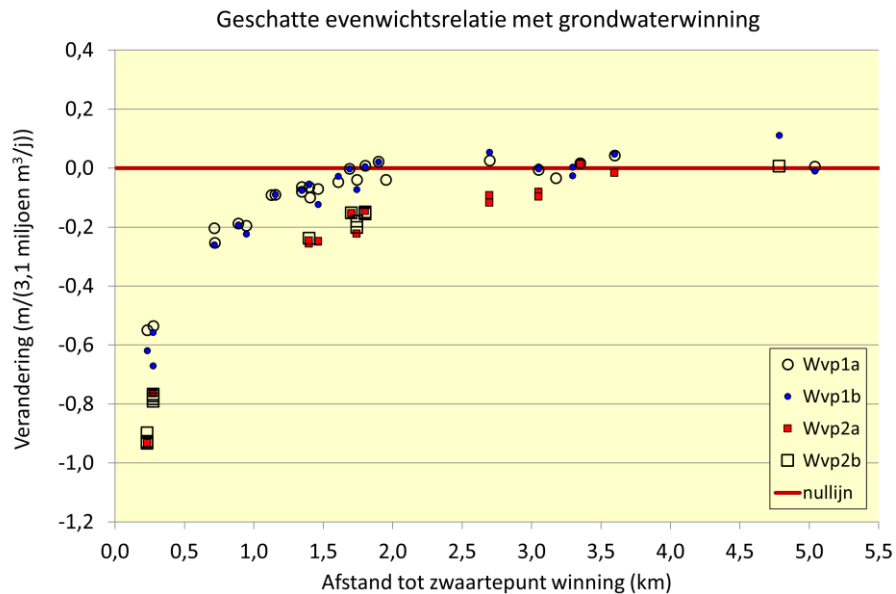
Hoewel tijdreeksmodellen in principe 'puntmodellen' zijn, leveren dergelijke resultaten bij het dichte netwerk van peilbuizen dat we in Nederland kennen toch een ruimtelijk beeld. Het ruimtelijke beeld geeft vaak een beter beeld op van de hydrologische betekenis van de resultaten dan een individuele reeks: terwijl de resultaten van een individuele reeks op zich best plausibel kunnen lijken, kunnen de resultaten in vergelijking met die in de nabije omgeving er toch uitspringen als uitbijter.

Het stationaire effect (M_0 of gain) van de responsfunctie voor grondwateronttrekkingen is bijvoorbeeld op te vatten als de stationaire verlaging van de grondwaterstand in meters per eenheid onttrokken debiet (in m per m³/d). De M_0 van een grondwaterwinning is afhankelijk van de positie van het peilbuisfilter dat geanalyseerd wordt, en van de geohydrologische eigenschappen van het grondwatersysteem. Enkele factoren die hierop betrekking hebben zijn:

- het al dan niet freatisch zijn van het grondwater;
- de afstand van de winning tot de peilbuis;
- de afstand van de peilbuis tot een voedende grens;
- de weerstand van het watervoerende pakket waaruit onttrokken wordt;
- de weerstand van eventuele afsluitende lagen tussen winning en peilbuisfilter.

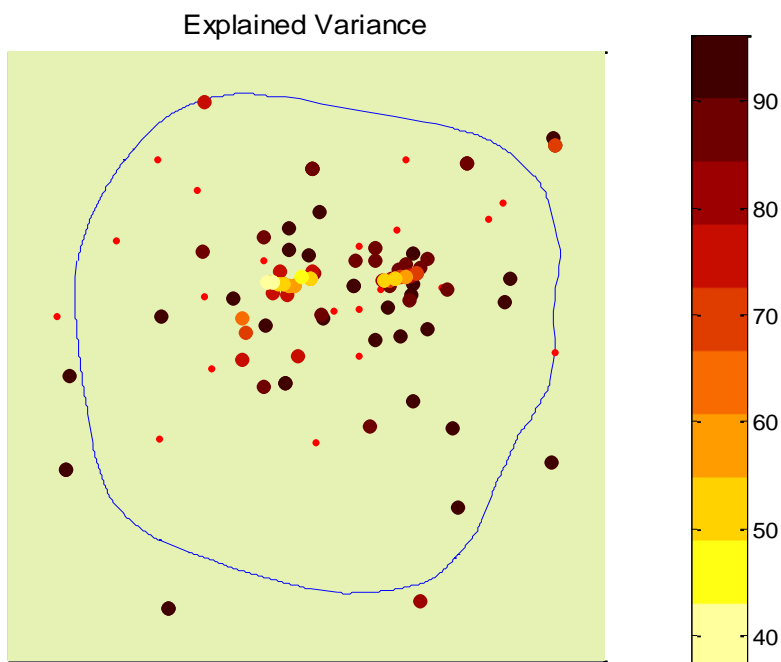
Aan de ene kant geven de resultaten van een tijdreeksanalyse via de M_0 en de andere parameters dus informatie over de eigenschappen van het grondwatersysteem. Aan de andere kant kan informatie over het grondwatersysteem helpen om plausibele van niet-plausibele relaties te scheiden. Wanneer men bijvoorbeeld de geschatte relatie met een winning uitzet tegen de afstand tot de winning, zou een onttrekkingskegel zichtbaar moeten worden als de omgeving geohydrologisch min of meer homogeen is (Afbeelding 6.3).

Afbeelding 6.3 Bij plotten van de geschatte relatie met de winning tegen de afstand tot het zwaartepunt van de winning zal er een onttrekkingskegel zichtbaar worden als de omgeving geohydrologisch min of meer homogeen is. Hier is er sprake van kegels in verschillende watervoerende pakketten



Een andere optie die veel inzicht op kan leveren is het plotten van de resultaten (parameters, validatiematen (typen afwijking), et cetera) op de locatie van de geanalyseerde filters. De plot in Afbeelding 6.4 geeft een ruimtelijk overzicht van de verklaarde varianties op de plaats van de geanalyseerde reeksen van een waterwingebied in Duitsland (Von Asmuth e.a., 2008). Duidelijk zichtbaar zijn twee 'strengen' met minder goede resultaten. In dit geval was de verklaring daarvan snel gevonden: als verklarende variabele was het verloop van de totale onttrekking meegenomen, terwijl de afzonderlijke winputten, op de plek van de twee strengen, individueel worden aangestuurd.

Afbeelding 6.4 Plot van de verklaarde variantie op de plaats van de geanalyseerde reeksen. Duidelijk zichtbaar zijn twee 'strengen' met minder goede resultaten



6.3.2 De valkuil van te vergaande onzekerheidsreductie

De betrouwbaarste resultaten van tijdreeksanalyse van de uitvoer van een bepaald systeem zijn te verwachten als de invoervariabelen van het systeem over de modelperiode voldoende onderscheidend zijn veranderd en niet aan elkaar gecorreleerd zijn. Dit kan bij bepaalde niet-natuurlijke systemen worden bewerkstelligd door een geschikte proefopzet, waarbij het verloop van invoervariabelen vooraf zodanig wordt ingesteld dat de tijdreeksanalyse maximale kans van slagen krijgt. Maar bij natuurlijke systemen is dat doorgaans niet of hooguit beperkt mogelijk en zal men het moeten doen met de beschikbare tijdreeksen. Men kan dan slechts hopen dat de tijdreeksanalyse toch bruikbare resultaten oplevert.

Soms zal ook blijken dat een tijdreeks met verschillende modellen kan worden beschreven, waarbij deze niet relevant verschillen in pasvorm. Dit verschijnsel wordt aangeduid als *equifinaliteit*. Het duidt er op dat de tijdreeks te weinig informatie bevat ten aanzien van de vraagstelling. Er zijn twee mogelijkheden hiermee om te gaan: i) een keuze maken uit de resultaten, of ii) alle resultaten presenteren. De tweede mogelijkheid is daarbij meer verantwoord.

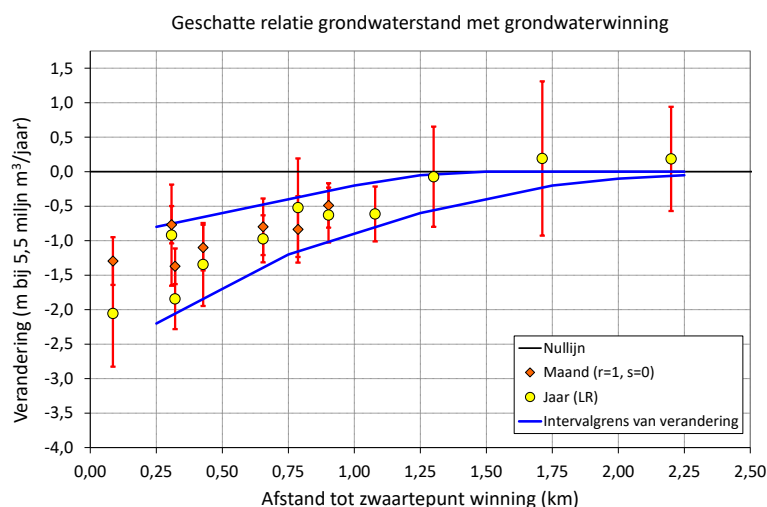
Wees alert op dit soort situaties en vermijd de valkuil van te vergaande onzekerheidsreductie. Probeer dus altijd op basis van fysisch-hydrologisch inzicht te beoordelen of er sprake kan zijn van equifinaliteit en laat dan niet uitsluitend één of ander kwantitatief informatiecriterium (zoals AIC, of BIC) de doorslag geven, tenzij de kandidaat-modellen volgens dat criterium sterk verschillen.

Houd in gedachte dat de onzekerheid:

- een maat is voor het gebrek aan informatie inhoud van de combinatie model + data;
- niet door de modelleur wordt veroorzaakt en geen onkunde is;
- inspiratie voor en richting kan geven aan verbetering;
- een verschijnsel is waar managers bij hun besluitvorming rekening mee kunnen houden.

Dus als het vermoeden bestaat dat er sprake is van equifinaliteit, presenteer dan de resultaten van alle geschikt geachte modellen en breng de samenvoeging van binnen-model-onzekerheid en tussen-model-onzekerheid in beeld. In het voorbeeld van Afbeelding 6.5 is de samengevoegde onzekerheid van twee modelsoorten in beeld gebracht middels een interval dat de onzekerheid aangeeft van de met twee modelsoorten geschatte verlagingskegel door de grondwaterwinning.

Afbeelding 6.5 De samengevoegde onzekerheid van twee modelsoorten (op maandbasis en op jaarbasis) is in beeld gebracht middels een interval dat de onzekerheid aangeeft van de met twee modelsoorten geschatte verlagingskegel door de grondwaterwinning. Een verticale rode lijn is het 95 %-betrouwbaarheidsinterval van de geschatte evenwichtsrelatie



Verder kan in het algemeen worden aangenomen dat de aannemelijkheid van modelresultaten toeneemt naarmate meer van de volgende voorwaarden opgaan:

- alle mogelijk relevante invoervariabelen zijn meegenomen als invoerreeksen;
- de geschatte evenwichtsrelaties zijn fysisch-hydrologisch verklaarbaar;
- de geschatte evenwichtsrelaties zijn onderling niet sterk gecorreleerd;
- de invoervariabelen zijn voldoende veranderd over de modelperiode;
- het ruimtelijk beeld van de geschatte invloeden is aannemelijk.

Deel II: Casussen en uitgewerkte voorbeelden

7

CASUS BJTR-METHODE: INVLOED GRONDWATERWINNING NIJ BEETS

Deze casus beschrijft de studie Nij Beets, waar met tijdreeksanalyse de ruimtelijke invloed van de gelijknamige grondwaterwinning is bepaald. De studie is uitgevoerd in opdracht van Vitens [Baggelaar en Van der Meulen, 2010]. Daarbij zijn van 67 grondwaterstandreeksen tijdreeksmodellen ontwikkeld volgens de BJTR-methode. De reeksen zijn afkomstig uit 44 peilbuizen, gelegen binnen een straal van 4 km van het zwaartepunt van de winning Nij Beets. De uitvoering van deze studie is conform de in de hoofdstukken 3, 4 en 6 van deze handleiding beschreven stappen.

7.1 Specificeren soort gewenste resultaat

Deze studie dient de invloed van de grondwaterwinning op de grondwaterstand in de omgeving empirisch te bepalen. Daartoe dient met tijdreeksanalyse van een groot aantal grondwaterstandreeksen de evenwichtsrelatie met de winning te worden geschat. Gezien het belang van de bevindingen voor het drinkwaterbedrijf en voor agrariërs, natuurorganisaties en omwonenden, ligt het voor de hand ook de onzekerheid van de resultaten mee te geven. Dat zal in staat stellen tot een beter onderbouwde risicoafweging en besluitvorming te komen op basis van de resultaten.

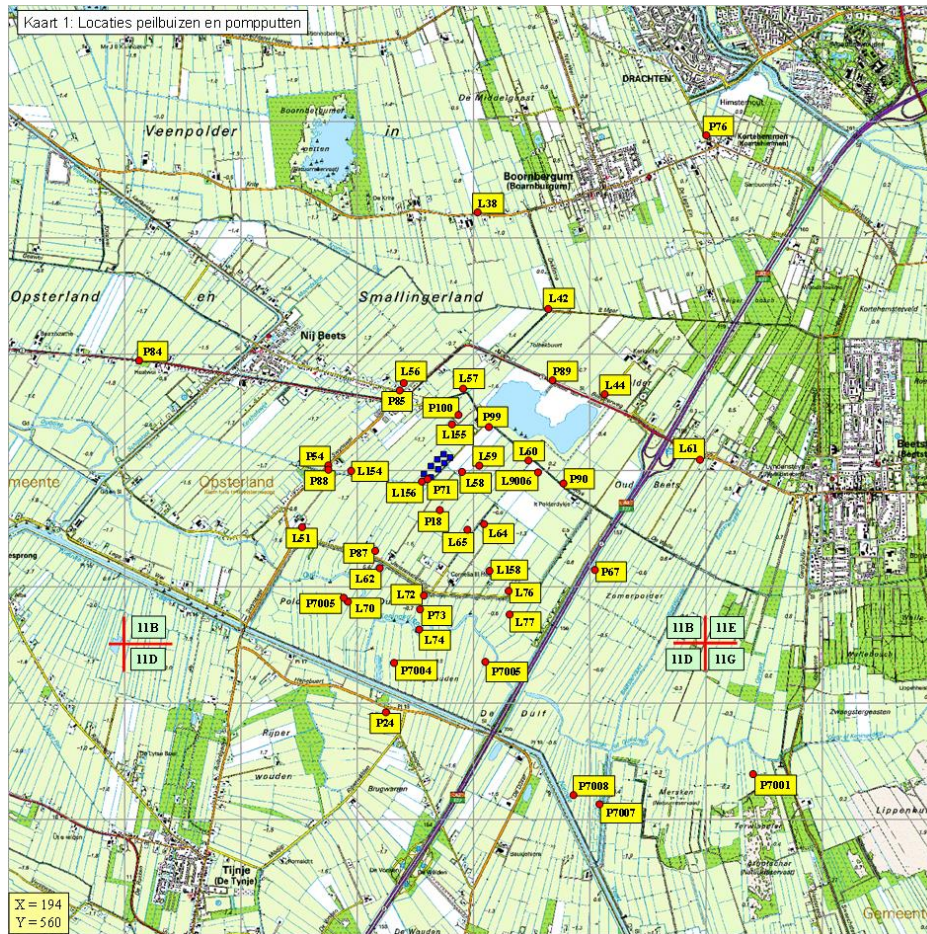
7.2 Verkennen systeem en data

7.2.1 Topografische, geografische en hydrogeologische kenmerken van de omgeving

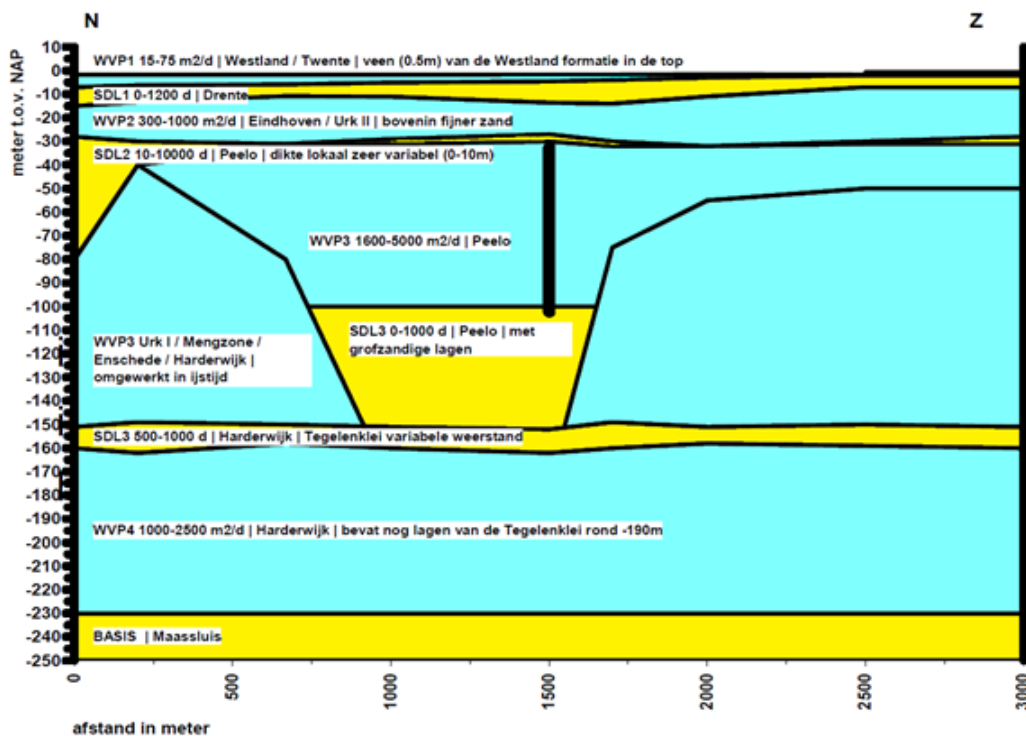
De omgeving is grotendeels te kenmerken als ontgonnen laagveengebied, met veel sloten en uitgestrekte weilanden. De maaiveldhoogte bevindt zich overwegend tussen NAP -0,5 en -1,5 m. Er zijn diverse natuurgebieden in de omgeving, waaronder een natuurreservaat dat bestaat uit een laagveenmoeras. Verder is er een Natura 2000-gebied van 846 ha. In Afbeelding 7.1 is het projectgebied weergegeven.

In de ondergrond van het onderzoeksgebied komen drie à vier watervoerende pakketten voor, gescheiden door slecht doorlatende pakketten. De winning Nij Beets onttrekt grondwater aan het 3^e watervoerende pakket (Afbeelding 7.2).

Afbeelding 7.1 Kaart van het studiegebied, met peilbuizen (rood) en de pompputten van de winning Nij Beets (blauw)

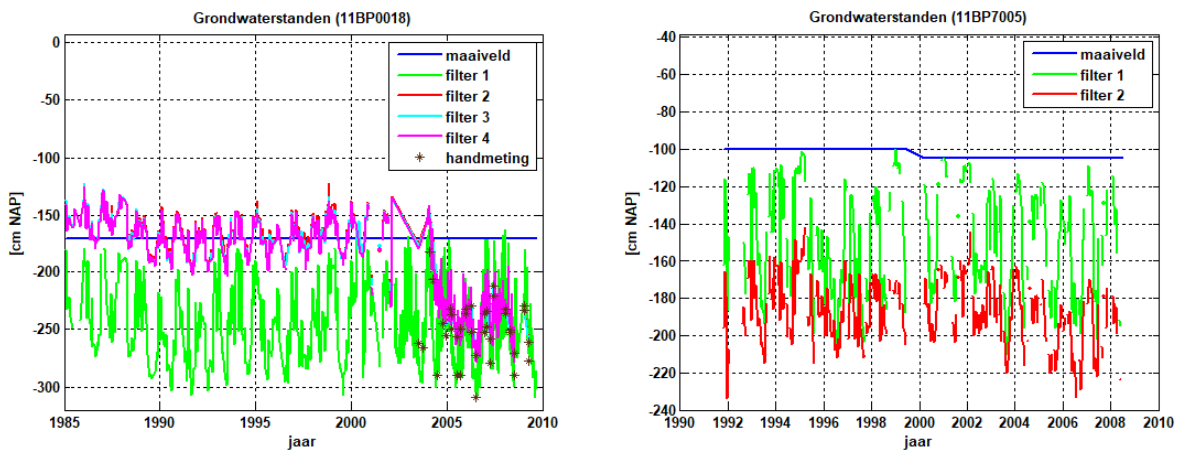


Afbeelding 7.2 Geohydrologisch schema van de ondergrond in de omgeving van de winplaats Nij Beets. Ook de positie van de pomppilters van de winning Nij Beets is aangegeven. Afkomstig uit [Vitens, 2007]



In het laaggelegen centrale deel van het gebied komt veel kwel voor. Deze wordt via de sloten en een zandwinplas door bemaling afgevoerd naar de Friese boezem. Verder wordt ook een deel van de kwel afgevangen door de winning Nij Beets. De kwel blijkt bijvoorbeeld uit een vergelijking van de grondwaterstands- en stijghoogtereeksen van peilbuis 11BP0018 (zie Afbeelding 7.3 - links). Aangezien deze peilbuis dicht bij de winning Nij Beets ligt, zijn de diepe stijghoogten sinds de start van deze winning in maart 2004 sterk gedaald.

Afbeelding 7.3 Links: Freatische grondwaterstandreeks (filter 1) en drie diepere stijghoogtereeksen (filter 2-4) van peilbuis 11BP0018, gelegen 370 m ten zuiden van het zwaartepunt van de winning Nij Beets. De handmetingen zijn controlemetingen bij het uitlezen van de loggers (geplaatst in april 2002). Rechts: Freatische grondwaterstandreeks (filter 1) en diepere stijghoogtereeks (2^e wvp, filter 2) van peilbuis 11BP7005, gelegen 1,4 km ten zuid-zuidwesten van het zwaartepunt van de winning Nij Beets, in polder De Dulf



De kwel is afkomstig van de hoger gelegen gebieden in het noordoosten en oosten en de natuurgebieden in het zuidoosten en zuiden. In Afbeelding 7.3 - rechts is bijvoorbeeld te zien dat er infiltratie optreedt in polder De Dulf, die circa 1 km ten zuiden van de winplaats ligt. Het maaiveld ligt in deze polder ongeveer 1 meter hoger dan rond de winplaats (NAP -0,7 m versus NAP -1,7 m).

7.2.2 Verkennen data mogelijke invloedsfactoren grondwaterstand

Meteorologische grondwateraanvulling

De belangrijkste invloedsfactor van de grondwaterstand is doorgaans de grondwateraanvulling, maar daar zijn geen meetgegevens over beschikbaar. Als theoretische benadering van deze factor kan echter het potentieel neerslagoverschot dienen. Dat wordt berekend uit de neerslag en de verdamping, volgens:

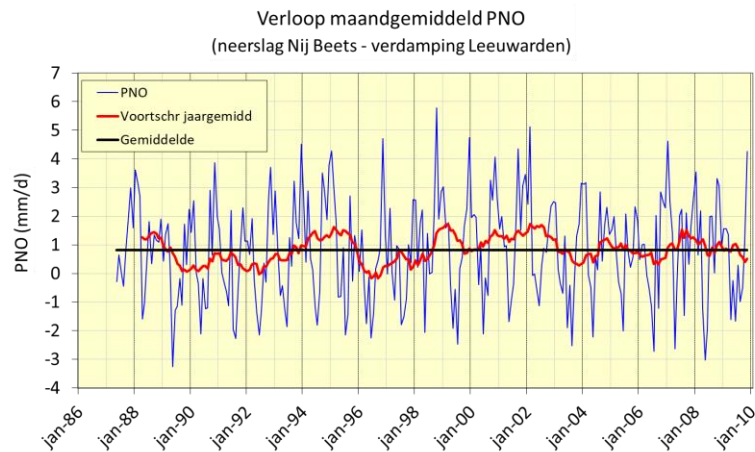
$$PNO_t = R_t - V_t$$

waarin PNO_t het potentieel neerslagoverschot, R_t de neerslag en V_t de Makkink-verdamping¹, alle in mm per tijdseenheid en t de index van de tijdseenheid.

¹ Deze wordt ook wel aangeduid als de referentiegewasverdamping.

Voor de reconstructie van het potentieel neerslagoverschot is gebruik gemaakt van de neerslagcijfers van het nabijgelegen neerslagstation Nij Beets. Verder zijn verdampingscijfers gebruikt van het weerstation Vliegbasis Leeuwarden, circa 24 kilometer ten noordwesten van de winning Nij Beets. Aangezien de verdampingscijfers doorgaans betrekking hebben op een grote regio, nemen we aan dat die ook voldoende representatief zijn voor het onderzoeksgebied. Afbeelding 7.4 toont het uit deze cijfers gereconstrueerde verloop van het maandgemiddelde van het potentieel neerslagoverschot over de periode mei 1987 t/m november 2009.

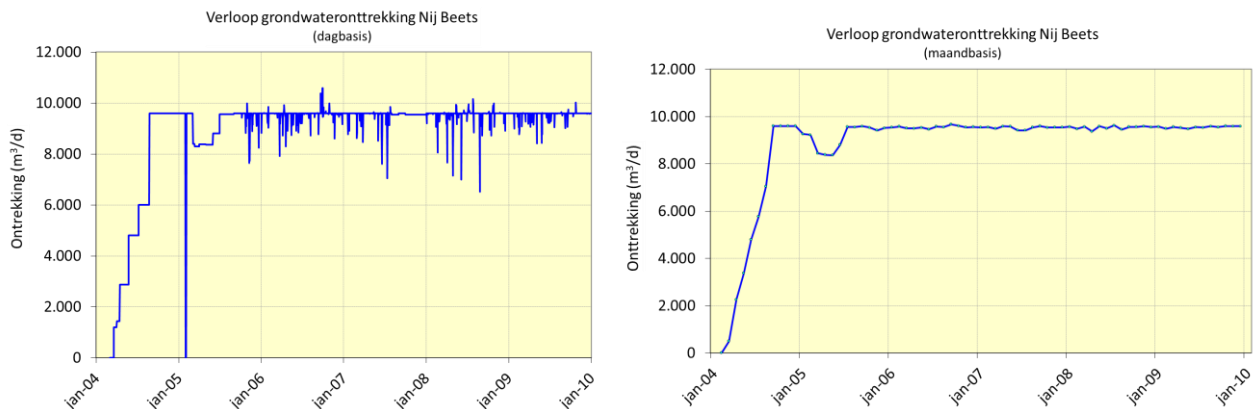
Afbeelding 7.4 Verloop van het maandgemiddeld potentieel neerslagoverschot, zoals afgeleid uit de neerslag van het neerslagstation Nij Beets en de Makkink-verdamping op het weerstation Leeuwarden, beide op dagbasis. Tevens zijn weergegeven het gemiddelde en het voortschrijdend jaargemiddelde van het potentieel neerslagoverschot



Grondwaterwinning Nij Beets

De winning Nij Beets bestaat uit acht pompputten, die maximaal op 260 m afstand van elkaar liggen (zie Afbeelding 7.1). Ze onttrekken grondwater uit het 3^e watervoerende pakket, met de pomppfilters tussen 30 en 100 m onder maaiveld. Deze pompputten worden roulerend ingezet. De totale winning bedraagt bijna 3,5 miljoen m³/jaar. Afbeelding 7.5 toont het verloop van de winning, zowel op dagbasis (links), als op maandbasis (rechts). De winning heeft een grote variatie vertoond, aangezien deze vanaf de start op 19 maart 2004 in slechts vier maanden is toegenomen naar een vrijwel constant niveau van rond de 3,5 miljoen m³/jaar. In de eerste helft van 2005 is echter minder onttrokken dan deze vergunde hoeveelheid, door problemen met de zuivering.

Afbeelding 7.5 Verloop van de grondwaterwinning Nij Beets, zowel op dagbasis (links), als op maandbasis (rechts)



Door de grote en vrij abrupte sprong van de winning zal deze over een periode die de sprong omvat niet storend gecorreleerd zijn aan het neerslagoverschot, zodat er geen verstrengeling is te verwachten van de geschatte statistische relaties van deze twee factoren met de grondwaterstand. We mogen daarom veronderstellen dat het reconstrueren van de invloed van deze winning middels tijdreeksanalyse een redelijke kans van slagen heeft.

Verloop zwaartepunt winning Nij Beets

Als het zwaartepunt van een winning sterk verandert, zal dat invloed hebben op de grondwaterstanden in de omgeving en moet daar dus rekening mee worden gehouden bij de tijdmodellering van de grondwaterstandreeksen die zijn gemeten in peilbuizen dicht bij de winning. Hier zal dat echter niet nodig zijn, omdat alle acht pompputten vanaf de start van de winning hebben meegedraaid. Daarbij is weliswaar sprake van een roulerende inzet, maar dit leidt er alleen toe dat het zwaartepunt van de winning enigszins rond een vast punt danst en niet dat het een stapsgewijze verandering vertoont. Waar we in het vervolg refereren aan afstanden van peilbuizen tot de winning Nij Beets, zijn deze berekend ten opzichte van het zwaartepunt van de winning, berekend als gemiddelde van de coördinaten van de acht pompputten.

Overige grondwaterwinningen

Binnen een straal van acht km rond de winning Nij Beets bevinden zich sinds 1986 84 bestaande, of inmiddels gestopte grondwateronttrekkingen, voor een groot deel bronbemalingen (36) en verder onttrekkingen voor sanering (21), beregening (14), koude-warmte-opslag (7), koelwater (4), proceswater (1) en een proefonttrekking (1). Uit een beoordeling van de afstanden van deze grondwateronttrekkingen tot de in dit onderzoek beschouwde peilbuizen en de mate waarin een onttrekking heeft gevarieerd over de modelperioden van de verschillende grondwaterstandreeksen, kon worden aangenomen dat geen van deze hoefde te worden verdisconteerd bij het modelleren van grondwaterstandreeksen.

Oppervlaktewaterstanden

De oppervlaktewaterstanden in de onderzoekomgeving worden beheerd door Wetterskip Fryslân, met afhankelijk van het peilvlak een vast peil, of afwisselend een zomerpeil en een winterpeil. Er zijn meerdere peilwijzigingen in dit gebied geweest, waarvan de meest vergaande in de tweede helft van de jaren tachtig, in het kader van de ruilverkaveling Midden-Opsterland. Verder zijn rond 1995 in het kader van deze ruilverkaveling oppervlaktewaterpeilen verlaagd ten zuiden van De Mersken tussen Lippenhuizen/Terwispel. Aan de zuidkant van de Beetsterweg, een paar honderd meter ten noordoosten van de winplaats Nij Beets bevindt zich een zandwinplas van ongeveer 30 hectare. In 2006 is het peil van deze plas circa 12 cm verhoogd. Met deze waterhuishoudkundige ingrepen is rekening gehouden bij de modellering van de grondwaterstandreeksen, bijvoorbeeld door de modelperiode later te starten, of door de ingreep mee te nemen als invoerreeks, in de vorm van een $[0,1]$ -variabele (dit wordt ook wel aangeduid als een *interventie-variabele*).

7.3 Opstellen modelleerstrategie

Voor elke grondwaterstandreeks is een model ontwikkeld dat deze beschrijft als functie van een aantal invoerreeksen. Hier zijn dat in ieder geval de meteorologische grondwateraanvulling (vertegenwoordigd door het potentieel neerslagoverschot) en de grondwaterwinning Nij Beets, voor de peilbuizen nabij de zandwinplas aangevuld met de $[0,1]$ -reeks voor de peilverandering van de zandwinplas.

7.3.1 Keuze soort tijdreeksmodel

De tijdreeksanalyses zijn uitgevoerd met het BJTR-model, het klassieke transfer-ruismodel, zoals beschreven in [Box and Jenkins, 1970]. Er zijn verschillende redenen voor deze keuze, namelijk:

- daar zijn bij soortgelijke studies al goede ervaringen mee opgedaan;
- de onzekerheid van de resultaten moet ook worden meegegeven. Het BJTR-model bevat een uitgebreid ruismodel dat doorgaans goed in staat is de modelresiduen terug te brengen tot witte ruis (dit is een randvoorwaarde om die onzekerheden mee te kunnen geven).

7.3.2 Keuze tijdseenheid

De grondwaterstandreeksen zijn in dit onderzoek gemodelleerd met als tijdseenheid de maand, voor zowel de uitvoerreeks als de invoerreeksen (reeksen dienen equidistant te zijn). Dit sluit aan op de tijdschaal van de dynamiek van de verschillende processen die de grondwaterstand beïnvloeden. Uit verkennende modelleer-exercities bleek namelijk dat de responstijd van de grondwaterstand op het neerslagoverschot hier kan oplopen tot circa een jaar, terwijl die op de winning Nij Beets kan oplopen tot circa een half jaar.¹

Modelleren met kleinere tijdseenheden zal dan niet alleen veel modelparameters vergen, wat minder gewenst is, maar kan het bovendien moeilijk of onmogelijk maken om het modelresidu terug te brengen tot witte ruis, door de grote autocorrelatie bij kleine tijdseenheden. Het overgaan op maandwaarden geeft hier ook nauwelijks een uitvlakking van de variatie van de winning Nij Beets, zodat het onderscheidend vermogen bij de tijdreeksanalyse niet sterk wordt geschaad. Want over de periode 1 januari 2000 t/m 31 december 2009 bedraagt de standaardafwijking van de winning op dagbasis 4.632 m³/d, terwijl deze op maandbasis 4.620 m³/d bedraagt, een afname van slechts 0,2 %. Dit zeer geringe verschil komt vooral door de snelle en grote winningstoename bij de start in 2004.

7.3.3 Keuze modelperiode

De (maximale) modelperiode die voor de tijdreeksanalyse is gehanteerd loopt van 1990 t/m 2009 (studie is in 2010 uitgevoerd). Deze periode omvat circa 14 jaar vóór de start van de winning - en de snelle en grote toename daarvan - en circa vijf jaar na de start van de winning. Dit alles om zoveel mogelijk onderscheidend vermogen te krijgen om een relatie van de grondwaterstand met de winning te kunnen detecteren met tijdreeksanalyse.

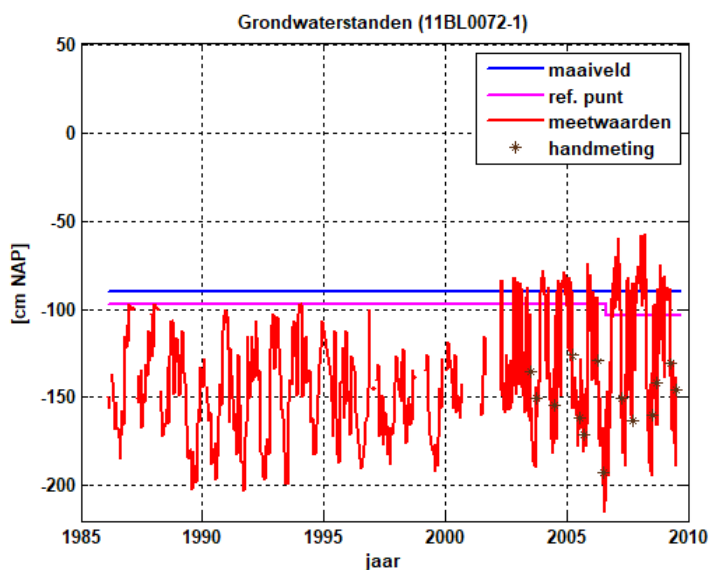
7.4 Voorbewerken meetreeksen

7.4.1 Opschonen van fouten

De verzamelde grondwaterstandreeksen zijn eerst onderzocht op het voorkomen van meet- en/of transcriptiefouten. Deze kunnen naar voren komen als uitbijters - dit zijn meetresultaten die niet in het normale beeld passen - of een vreemd verloop van de tijdreeks. Hiertoe zijn de grafieken van alle grondwaterstandreeksen visueel beoordeeld op eventuele anomalieën, gebruik makende van hydrologische kennis en ervaring met betrekking tot grondwaterstandreeksen. Daarbij is ook nagegaan of de anomalie kon zijn veroorzaakt door een lokale hydrologische ingreep. Indien voldoende aannemelijk was dat de anomalie is veroorzaakt door meet- en/of transcriptiefouten, is deze uit de reeks verwijderd, tenzij het hoogstwaarschijnlijk een zogenaamde meterfout betrof, in welk geval de anomalie handmatig is gecorrigeerd.

¹ De responstijd van een variabele op een andere variabele is de tijd die het duurt voordat de invloed van een impuls van de eerste variabele geen uitwerking meer heeft op de tweede variabele.

Afbeelding 7.6 Voorbeeld van een grondwaterstandsreeks waarbij het loggerdeel niet goed lijkt aan te sluiten op het voorgaande, handmatig gemeten deel. De logger is geïnstalleerd in april 2002. De met sterren aangegeven handmetingen zijn de handmatige controlemetingen die zijn gedaan bij het uitlezen van de loggers



7.4.2 Kalibreren loggermetingen

De loggermetingen waren wel gecorrigeerd voor luchtdrukschommelingen, maar nog niet gekalibreerd op de handmatige controlemetingen die (meestal) bij het uitlezen van de loggerdata zijn genomen. Daarom zijn de loggerreeksen gekalibreerd aan de handmatige metingen.

7.4.3 Omzetten meetreeksen naar maandbasis

De meeste grondwaterstandreeksen bevatten doorgaans twee meetwaarden per maand, gemeten op of rond de 14^e en de 28^e, maar enkele reeksen bevatten ook dagelijkse meetwaarden, die zijn gemeten met loggers. Elke reeks is omgezet tot een reeks van de maandgemiddelde grondwaterstand, door per maand de meetwaarden te middelen. Het verlies aan informatie over het verloop van de grondwaterstand door het middelen naar maandwaarden zal beperkt zijn, gezien de grote positieve autocorrelatie (op te vatten als een maat voor de overlap in informatie) van halfmaandelijke meetwaarden van de grondwaterstand en de zeer grote positieve autocorrelatie van dagelijkse meetwaarden van de grondwaterstand.

Van de 67 voor tijdreeksanalyse geselecteerde grondwaterstandreeksen (zie verder) heeft een aantal het begindeel met halfmaandelijke grondwaterstanden en het eindeel (vanaf mei 2002 of later) met dagelijkse grondwaterstanden (het deel met de loggerdata). Bij middelen op maandbasis ontstaat er dan een reeks maandgemiddelden die niet alle dezelfde precisie hebben, wat dan zal ook zal opgaan voor de residuen van het tijdreeksmodel. Dat druist echter in tegen een belangrijke randvoorwaarde bij de tijdreeksanalyse, namelijk dat de residuen geen heteroscedasticiteit vertonen. Daarom zijn bij een reeks met verschil in meetfrequentie alleen de grondwaterstanden van de 14^e en de 28^e gemiddeld tot maandgemiddelde voor het deel met de loggerdata. Dit maakt de precisie van de maandgemiddelden over de gehele reeks vergelijkbaar. Als een grondwaterstand van de 14^e of de 28^e ontbrak, is de dichtstbijzijnde genomen, binnen een venster van 7 dagen aan weerskanten van de betreffende datum.

De invoerreeksen die beschikbaar zijn als reeksen van maandsommen (het potentieel neerslagoverschot en de grondwateronttrekking), zijn omgezet tot reeksen van maandgemiddelden, door per maand de maandsom te delen door het aantal dagen in die maand. Zonder die omzetting zouden de variaties van deze invoerreeksen kunstmatig zijn vergroot, doordat niet alle maanden even lang zijn.

7.4.4 Selectie te modelleren grondwaterstandreeksen

De maximale modelperiode die voor de tijdreeksanalyse is gehanteerd loopt van 1990 t/m 2009. Per grondwaterstandreeks is de modelperiode gekozen op basis van de beschikbaarheid van metingen en een visuele beoordeling op eventuele discontinuïteiten. Als een reeks duidelijk een invloed van één of andere niet nader bekende, of onvoldoende gedetailleerd bekende peilverandering weerspiegelde, is de modelperiode later gestart of eerder geëindigd. Verder is van een aantal peilfilters die in 2002 of later voorzien zijn van een logger, als start het begin van de loggerreeks genomen, als de twee reeksdelen (handmatig gemeten deel en loggerdeel) niet goed op elkaar leken aan te sluiten. Geen van de modelperioden start echter later dan juli 2003, of eindigt eerder dan mei 2006, zodat elke modelperiode de start van de grondwaterwinning Nij Beets in maart 2004 omvat.

Een grondwaterstandreeks is geselecteerd voor de tijdreeksanalyse, als deze voldoet aan elk van de volgende criteria:

- 1 de reeks start medio 2003 of eerder, dus circa een jaar voor het starten van de winning Nij Beets;
- 2 de reeks eindigt niet eerder dan medio 2006, dus circa 2 jaar na het starten van de winning Nij Beets;
- 3 de reeks is over de periode medio 2003 tot medio 2006 grotendeels minimaal maandelijks bemeten en bevat niet te veel of te grote hiaten (dit is beoordeeld op basis van hydrologische expertise);
- 4 als de reeks of een deel daarvan is bemeten met een logger, dan is deze niet té onbetrouwbaar gebleken bij het kalibreren aan de handmatige controlemetingen (dit is beoordeeld op basis van hydrologische expertise).

Er bleken 67 van de 94 grondwaterstandreeksen, gemeten in 44 peilbuizen, te voldoen aan deze criteria (zie Afbeelding 7.1 voor hun ligging). Het betreft 35 grondwaterstandreeksen uit het 1^e, 18 uit het 2^e, 7 uit het 3^e (het winpakket) en 7 uit het 4^e watervoerende pakket.

7.4.5 Iteratieve modelontwikkeling

Alle grondwaterstandreeksen zijn dus gemodelleerd met als invoerreeksen: (1) het potentieel neerslagoverschot en (2) de grondwaterwinning Nij Beets. Zoals toegelicht in het voorgaande is voor enkele reeksen nabij de zandwinplas tevens een derde invoerreeks gehanteerd, representatief voor de peilverandering van de zandwinplas in 2006.

In de eerste modelronde is elke grondwaterstandreeks gemodelleerd met al zijn kandidaat-invloedsfactoren als invoerreeksen. Vervolgens zijn de geschatte modellen geverifieerd. Indien nodig is het tijdreeksmodel daarop handmatig aangepast, opnieuw geschat en weer geverifieerd, totdat het zo goed mogelijk voldeed aan de vooronderstellingen die aan het gehanteerde tijdreeksmodel ten grondslag liggen (zie paragraaf 7.5).

7.5 Identificeren vorm tijdreeksmodel

7.5.1 Modelleren dynamische relatie met neerslagoverschot

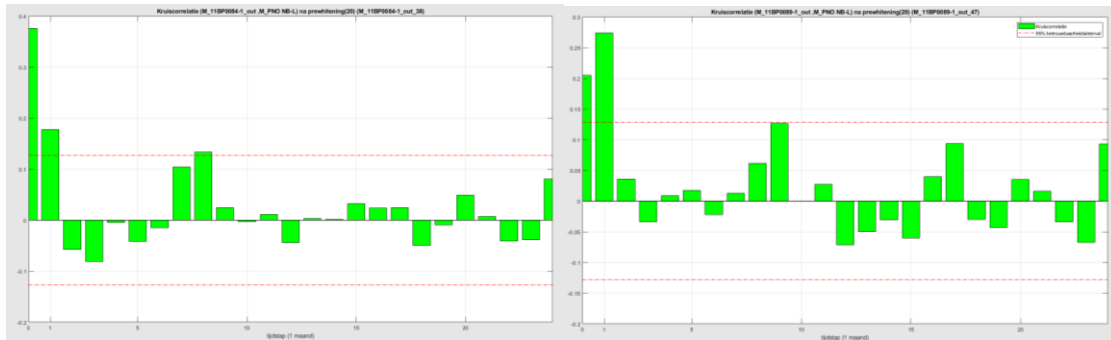
Om de relatie tussen het neerslagoverschot en een grondwaterstandreeks te kunnen modelleren, is het potentieel neerslagoverschot als invoerreeks gehanteerd. Om voor de eerste modelronde tot een geschikte formulering van de relatie tussen het neerslagoverschot en de grondwaterstand te kunnen komen, is voor elke grondwaterstandreeks afzonderlijk het gewitte kruiscorrelogram van het potentieel neerslagoverschot en de grondwaterstand visueel beoordeeld. Een dergelijk kruiscorrelogram zal namelijk min of meer de vorm hebben van de impulsresponsfunctie en daarmee een indruk geven van de wijze waarop de grondwaterstand reageert op het neerslagoverschot (zie paragraaf 4.2.1). Voor de meeste grondwaterstandreeksen wees het gewitte kruiscorrelogram op een vrijwel directe (in dezelfde maand) maximale reactie van de grondwaterstand op het neerslagoverschot, gevolgd door een stapterugkeer richting nul (zie Afbeelding 7.77 - links).

Deze relatie is in het tijdreeksmodel geformuleerd met één AR-parameter (δ_1) en één MA-parameter (ω_0) als parameters van de betreffende transferfunctie, volgens:

$$C_{PNO,t} = \delta_1 C_{PNO,t-1} + \omega_0 PNO_t$$

waarin $C_{PNO,t}$ de component van Z_t (grondwaterstand) die is gerelateerd aan PNO , het potentieel neerslagoverschot (zie paragraaf 7.2.2).

Afbeelding 7.7 Twee voorbeelden van een bij deze studie beoordeeld gewit kruiscorrelogram van het potentieel neerslagoverschot en de grondwaterstand. Beide betreffen peilfilters in het 2^e watervoerend pakket. Links een snelle maximale reactie en rechts een enigszins vertraagde maximale reactie



Voor een aantal grondwaterstandreeksen wees het gewitte kruiscorrelogram echter op een enigszins vertraagde maximale reactie van de grondwaterstand op het neerslagoverschot, gevolgd door een stapterugkeer richting nul (zie Afbeelding 7.77 - rechts). Deze relatie is in het model geformuleerd met één AR-parameter (δ_1) en twee MA-parameters (ω_0 en ω_1) als parameters van de betreffende transferfunctie, volgens:

$$C_{PNO,t} = \delta_1 C_{PNO,t-1} + \omega_0 PNO_t - \omega_1 PNO_{t-1}$$

7.5.2 Modelleren dynamische relatie met grondwaterwinning Nij Beets

Om een beeld te krijgen van de responstijd van de grondwaterstand op de winning Nij Beets, zijn eerst enkele stijghoogtereeksen gemodelleerd uit het winpakket, gemeten in peilbuizen nabij de winning. Meestal zijn dergelijke reeksen, die (semi)-spanningswater vertegenwoordigen namelijk vrij eenduidig te modelleren, doordat ze van nature minder fluctueren en snel reageren op grondwaterwinning. Uit deze modelleer exercitie bleek dat de diepe stijghoogten vrijwel instantaan reageren op de grondwaterwinning met een verlaging, om vervolgens min of meer via een 1^e-orde exponentieel verloop terug te keren naar de evenwichtssituatie. Op basis hiervan is voor alle diepe stijghoogtereeksen gemeten binnen 2,5 km van het zwaartepunt van de winning de relatie met de winning Nij Beets voor de eerste modelronde geformuleerd met één AR-parameter (d_1) en één MA-parameter (w_0) als parameters van de betreffende transferfunctie:

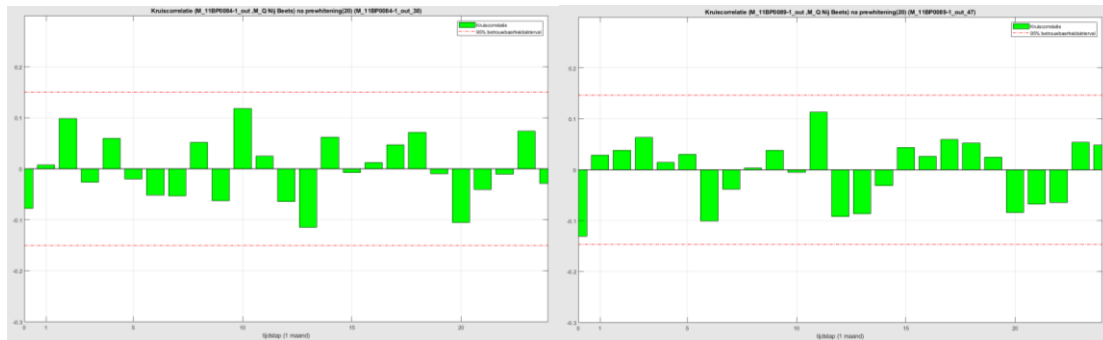
$$C_{Q,t} = \delta_1 C_{Q,t-1} + \omega_0 Q_t$$

Uit deze modelleer-exercitie kwam ook naar voren dat de responstijd van de stijghoogte op de winning Nij Beets circa 6 maanden bedroeg. Daarom is in de eerste modelronde voor elke freatische grondwaterstandreeks en voor de stijghoogtereeksen gemeten op meer dan 2,5 km van het zwaartepunt van de winning uitgegaan van zeven omega's (de tijdseenheid is immers een maand) als parameters van de transferfunctie met de winning, volgens:

$$C_{Q,t} = \omega_0 Q_t - \omega_1 Q_{t-1} - \dots - \omega_7 Q_{t-7}$$

waarin $C_{Q,t}$ de component van Z_t die is gerelateerd aan de winning Nij Beets (zie paragraaf 7.2.2). Deze formulering van de transferfunctie houdt rekening met een responstijd van 6 à 7 maanden.

Afbeelding 7.8 Twee voorbeelden van een bij deze studie beoordeeld gewit kruiscorrelogram van de winning Nij Beets en de grondwaterstand. Beide betreffen peilfilters in het 2^e watervoerend pakket. Deze kruiscorrelogrammen gaven nog geen duidelijke visuele informatie over de vorm van de relatie, maar bij het modelleren werd dat beeld duidelijker



7.5.3 Modelleren dynamische relaties met oppervlaktewaterstanden

Voor de tijdreeksmodellering van grondwaterstandreeksen in peilbuizen nabij de zandwinplas is de peilverhoging van de plas in 2006 meegenomen als invoerreeks. Er van uitgaande dat de peilwijziging effectief is vanaf augustus 2006, is deze vertegenwoordigd door een interventiereeks die de waarde 0 heeft tot augustus 2006 en de waarde 1 daarna.

De totstandkoming van de plas in de noordelijke helft van de Bouwespolder in 1999/2000 hoefde niet te worden verdisconteerd bij de tijdreeksmodellering, aangezien de tijdreeksmodelleringperiodes van de in een peilbuis naast die plas gemeten grondwaterstandreeksen 11DP7007-1 en 11DP7007-2 pas starten in 2000.

Voor de tijdreeksmodellering van grondwaterstandreeksen gemeten in of nabij peilvlakken waar afzonderlijke zomer- en winterpeilen worden nagestreefd is geen aparte invoerreeks meegenomen ter verdiscontering van die seizoensmatige peilwijzigingen. Deze seizoensvariatie zal deels worden beschreven door de relatie van de grondwaterstand met het potentieel neerslagoverschot.

7.5.4 Formuleren ruismodel

Voor elke gemodelleerde reeks is het ruismodel voor de eerste modelronde geformuleerd met alleen een AR(1)-parameter (ϕ_1) en een constante (c), volgens:

$$N_t - c = \phi_1(N_{t-1} - c) + a_t$$

waarin N_t de ruis en a_t het residu, beide in de maand met index t . Bij de modelverificatie is vervolgens nagegaan of dit afdoende was om te kunnen voldoen aan de randvoorwaarden. Zo nodig is het ruismodel hier op aangepast (zie paragraaf 4.2.3).

7.6 Schatten modelparameters

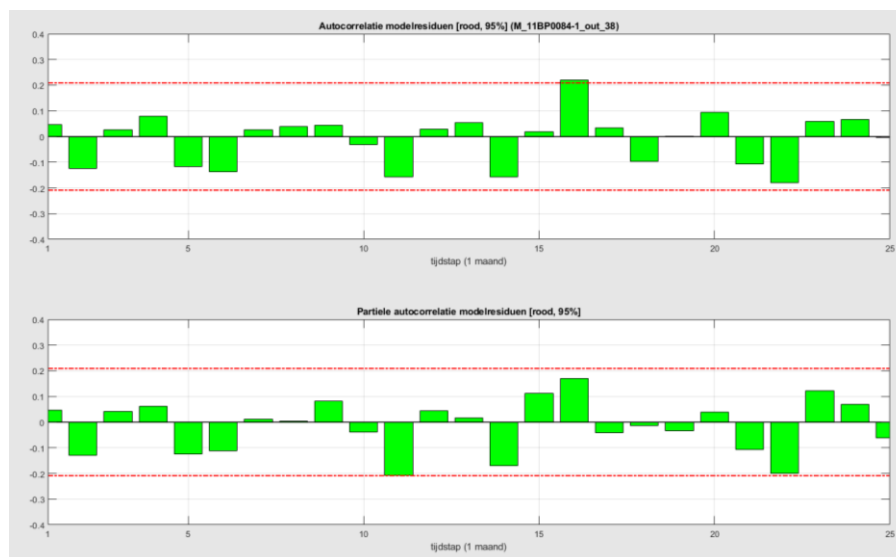
De tijdreeksmodellen zijn geschat met het programma Tijdreeksanalist [Van der Meulen en Baggelaar, 2019]. Het betreft een toolbox met allerlei subroutines, die via een menustructuur kunnen worden aangeroepen. Het programma is gebaseerd op de System Identification Toolbox van Matlab. In paragraaf 4.3 is het schattingsproces van de BJTR-methode beschreven.

7.7 Verifiëren en zonodig bijstellen tijdreeksmodel

Na het schatten zijn alle modellen geverifieerd op kenmerken van de modelresiduen en het ruismodel, zoals hieronder beschreven.

Het ruismodel is voor de eerste modelronde geformuleerd met alleen een 1^e-orde AR-parameter (ϕ_1) en een constante (c). Bij de modelverificatie is vervolgens voor elk tijdreeksmodel nagegaan of met dit ruismodel voldoende kon worden bewerkstelligd dat de modelresiduen normaal verdeelde *witte ruis* vormen, dat wil zeggen dat ze afkomstig zijn uit dezelfde normale kansverdeling, met gemiddelde nul en dat ze geen autocorrelatie vertonen (zie paragraaf 11.4). Als dat niet het geval bleek, is het ruismodel handmatig aangepast, bijvoorbeeld door één of meerdere MA-parameters ($\theta_1, \theta_2, \text{etc.}$) en/of een seizoenale AR-parameter (ϕ_7) toe te voegen, waarna het tijdreeksmodel opnieuw is geschat. Voor deze verificatie is onder meer afgegaan op visuele beoordeling van ACF en PACF van de modelresiduen (zie Afbeelding 7.9). Verder is ook afgegaan op de Ljung-Box-toets, die is gebaseerd op de som van de gekwadrateerde autocorrelatiecoëfficiënten van de modelresiduen (zie paragraaf 4.4.3).

Afbeelding 7.9. Voorbeeld van ACF (boven) en PACF (onder) van de residuen van de gemodelleerde grondwaterstandreeks 11BP0084-1. Deze diagnostieken wijzen niet op een autocorrelatiestructuur van de residuen

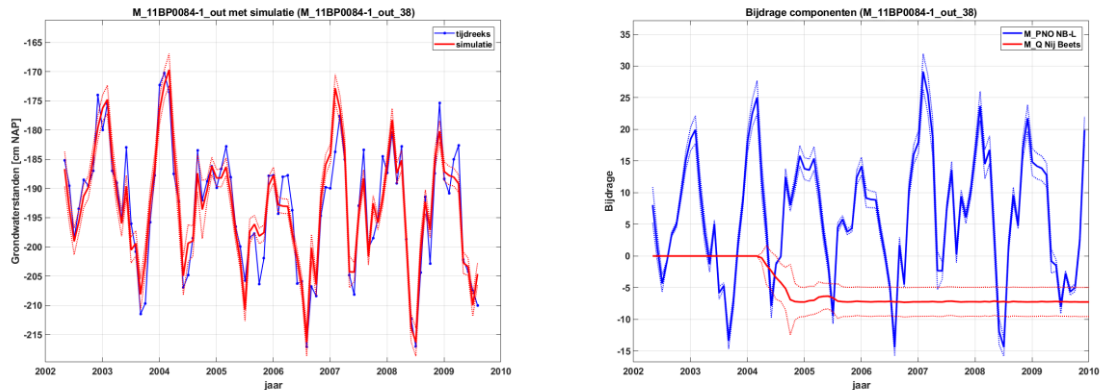


Bij de modelverificatie is ook nagegaan of de 1^e-orde AR-parameter van het ruismodel (ϕ_1) niet te hoog was geschat. Als voor deze parameter namelijk een waarde van circa 0,95 of meer is geschat, kan dat er op duiden dat de ruis een trend tracht te beschrijven en zo wellicht onterecht een deel van de invloed van de winning verdisconteert. Dit bleek echter voor geen van de modellen het geval.

Voorbeeld van gemodelleerde grondwaterstandreeks

Ter illustratie van de aanpak, hier als voorbeeld de modellering van grondwaterstandreeks 11BP0084-1, waarvan hierboven in Afbeelding 7.9 al het autocorrelogram en het partieel autocorrelogram van de modelresiduen weergegeven is. Het betreft een peilbuis met filter 1 in het 2^e watervoerende pakket, op 2,7 km ten noordwesten van de winning Nij Beets. Bij de modellering is de grondwaterstandreeks in feite ontleed in verschillende componenten, die de bijdragen van de verschillende invoerreeksen weergeven. Afbeelding 7.10 toont zowel het verloop van de maandgemiddelde grondwaterstand als het bij de modellering geschatte deterministische deel, dat hier bestaat uit de som van twee componenten, namelijk die van het potentieel neerslagoverschot en die van de winning Nij Beets.

Afbeelding 7.10 Links: De blauwe lijn toont het verloop van de maandgemiddelde grondwaterstand in peilbuis 11BP0084-filter 1. De rode lijn is het bij de modellering geschatte deterministische deel (aangeduid als 'simulatie'), dat hier bestaat uit de som van twee componenten, namelijk het potentieel neerslagoverschot en die van de winning Nij Beets. De rode stippellijnen begrenzen het 95 %-betrouwbaarheidsinterval van het geschatte deterministische deel. Rechts: Aparte weergave van de twee geschatte componenten en hun 95 %-betrouwbaarheidsinterval



Voor dit filter is de evenwichtsrelatie met de winning Nij Beets geschat op 7 cm verlaging van de grondwaterstand bij een winning van 3,5 miljoen m³/jaar. Gezien de relatief geringe standaardfout van deze schatting (1 cm) betreft het bij 90 % betrouwbaarheid een statistisch significante evenwichtsrelatie.

7.8 Presenteren resultaten tijdreeksanalyse

De tijdreeksanalyse leverde voor elke bij de modellering betrokken invoerreeks een schatting op van zijn statistische evenwichtsrelatie met de betreffende grondwaterstand, evenals de bijbehorende standaardfout, als maat voor de precisie van de schatting. Aan de hand van deze resultaten zijn de onder- en bovengrens van het 90 %-betrouwbaarheidsinterval van de geschatte evenwichtsrelatie van een grondwaterstand met de winning Nij Beets als volgt bepaald:

$$bi_{90\%}(\hat{g}) = \hat{g} \pm t_{95\%,n-k} \cdot s_{\hat{g}}$$

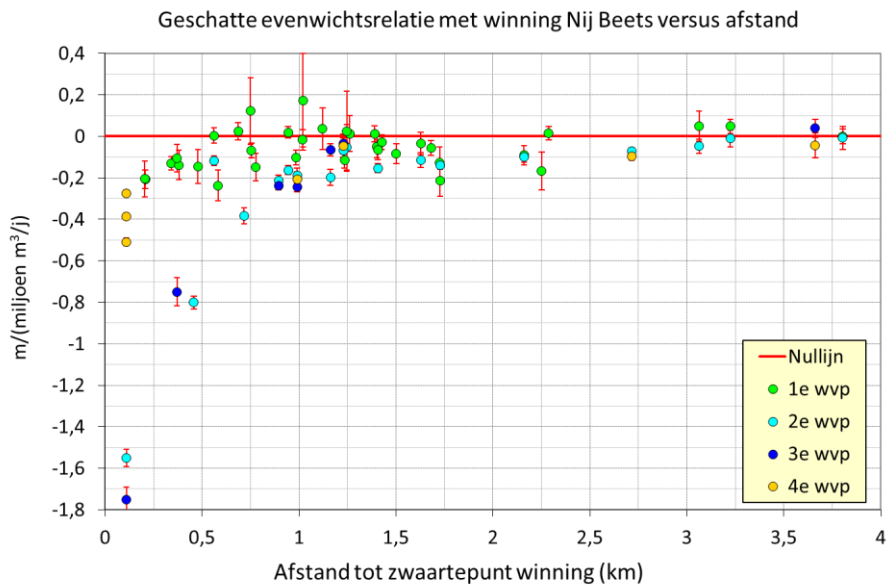
waarin \hat{g} de geschatte evenwichtsrelatie van de betreffende grondwaterstand met de winning Nij Beets (m/(3,5 miljoen m³/jaar)), $s_{\hat{g}}$ zijn standaardfout (m/(3,5 miljoen m³/jaar)) en $t_{95\%,n-k}$ het 95-percentiel van de Student-t-verdeling met aantal vrijheidsgraden gelijk aan $n-k$: het aantal reekswaarden (n) min het aantal modelparameters (k).

Als kan worden aangenomen dat wordt voldaan aan de daarvoor geldende randvoorwaarden (zie paragraaf 11.5) mag met 90 % betrouwbaarheid worden aangenomen dat de werkelijke evenwichtsrelatie zich ergens in dat 90 %-betrouwbaarheidsinterval bevindt. Dit impliceert dat als er ook vele andere realisaties van hetzelfde proces over dezelfde periode beschikbaar zouden zijn en we daarvan het tijdreeksmodel en bijbehorend betrouwbaarheidsinterval van de geschatte evenwichtsrelatie zouden afleiden, dat de werkelijke evenwichtsrelatie zich dan in 90 % van de gevallen in het aldus afgeleide interval zal bevinden.

Als het 90 %-betrouwbaarheidsinterval niet de waarde 0 omvat, is er met 90 % betrouwbaarheid sprake van een statistisch significante relatie tussen de grondwaterstand en de winning Nij Beets. Dit geldt overigens ook als het gehele interval boven de nullijn ligt (wat wijst op een stijging van de grondwaterstand), maar een causale relatie is dan uiteraard niet logisch.

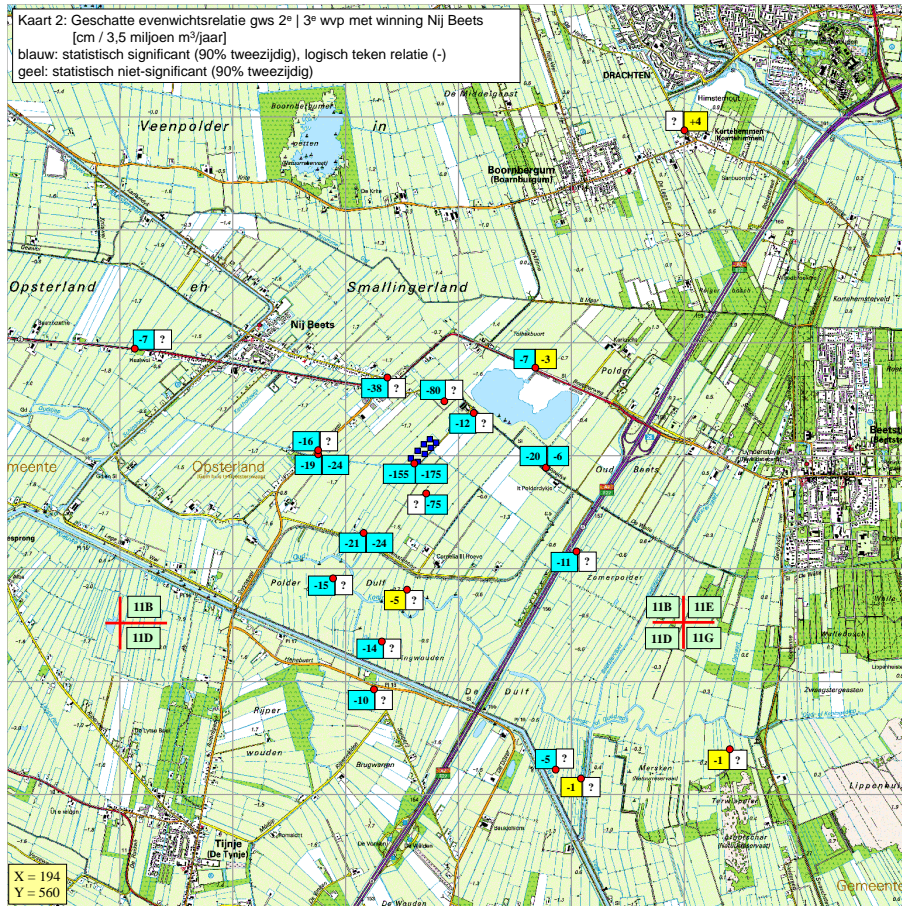
In Afbeelding 7.11 zijn de geschatte evenwichtsrelaties weergegeven versus de afstand tot het zwaartepunt van de winning, met onderscheid naar het watervoerend pakket. Bij elke geschatte evenwichtsrelatie is tevens het 90 %-betrouwbaarheidsinterval van die schatting weergegeven. Uit deze afbeelding blijkt dat in elk watervoerend pakket de geschatte verlaging doorgaans groter is naarmate de peilbuis dichterbij de winning ligt. En de verlagingen zijn in het 3^e watervoerende pakket - het winpakket - doorgaans groter dan in de andere watervoerende pakketten, waarbij het verschil toeneemt naarmate de peilbuis dichterbij de winning ligt.

Afbeelding 7.11 Geschatte evenwichtsrelatie van de grondwaterstand met de winning Nij Beets versus de afstand van de betreffende peilbuis tot het zwaartepunt van de winning, met ook onderscheid naar watervoerend pakket. Bij elke schatting is tevens het 90 %-betrouwbaarheidsinterval weergegeven (verticale rode lijn)



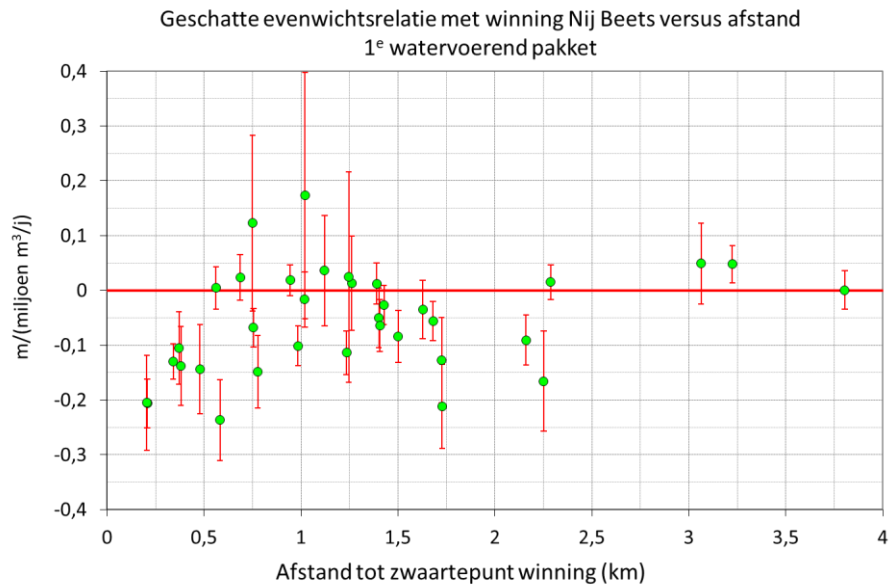
Zoals te zien in de kaart van Afbeelding 7.12 zijn naast de zandwinplas de verlagingen in het 2^e en 3^e watervoerende pakket relatief gering, namelijk 12 cm voor 11BP0099-2 (2^e wvp, zuidkant plas, op 0,56 km van de winning) en 7 en 3 cm voor 11BP0089-1 en -3 (2^e, respectievelijk 3^e wvp, noordkant plas, op 1,23 km van de winning). Mogelijke verklaringen zijn dat de verlagingen daar worden getemperd door voeding vanuit de plas en/of dat er plaatselijke obstructies voor grondwaterstroming zijn door potklei boven het winpakket.

Afbeelding 7.12 Kaart met de geschatte evenwichtsrelaties voor het 2^e en 3^e watervoerende pakket (bij de vraagtekens is er geen filter in het betreffende pakket)

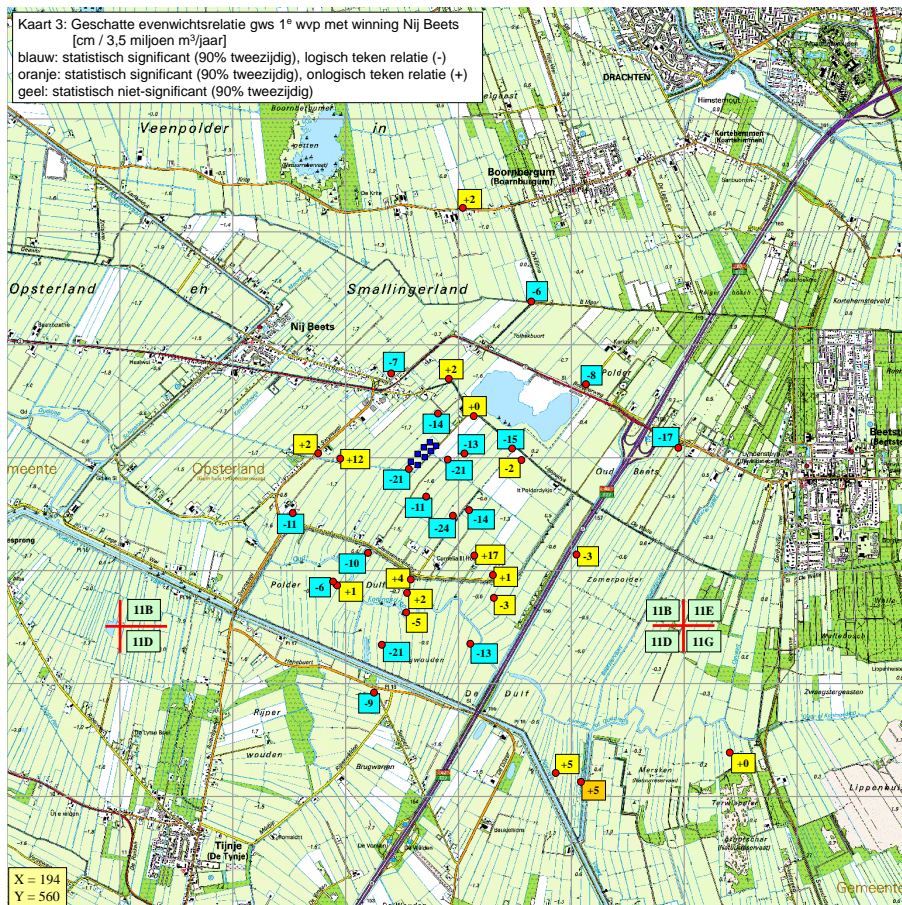


De resultaten voor het 1^e watervoerende pakket zijn afzonderlijk weergegeven in Afbeelding 7.13. Bij elke geschatte evenwichtsrelatie is tevens het 90 %-betrouwbaarheidsinterval van die schatting weergegeven. Verder zijn ze ruimtelijk weergegeven in de kaart van Afbeelding 7.14.

Afbeelding 7.13 Geschatte evenwichtsrelatie van de grondwaterstand met de winning Nij Beets versus de afstand van de betreffende peilbuis tot het zwaartepunt van de winning, voor het 1^e watervoerende pakket. Bij elke schatting is tevens het 90 %-betrouwbaarheidsinterval weergegeven (rode lijn)



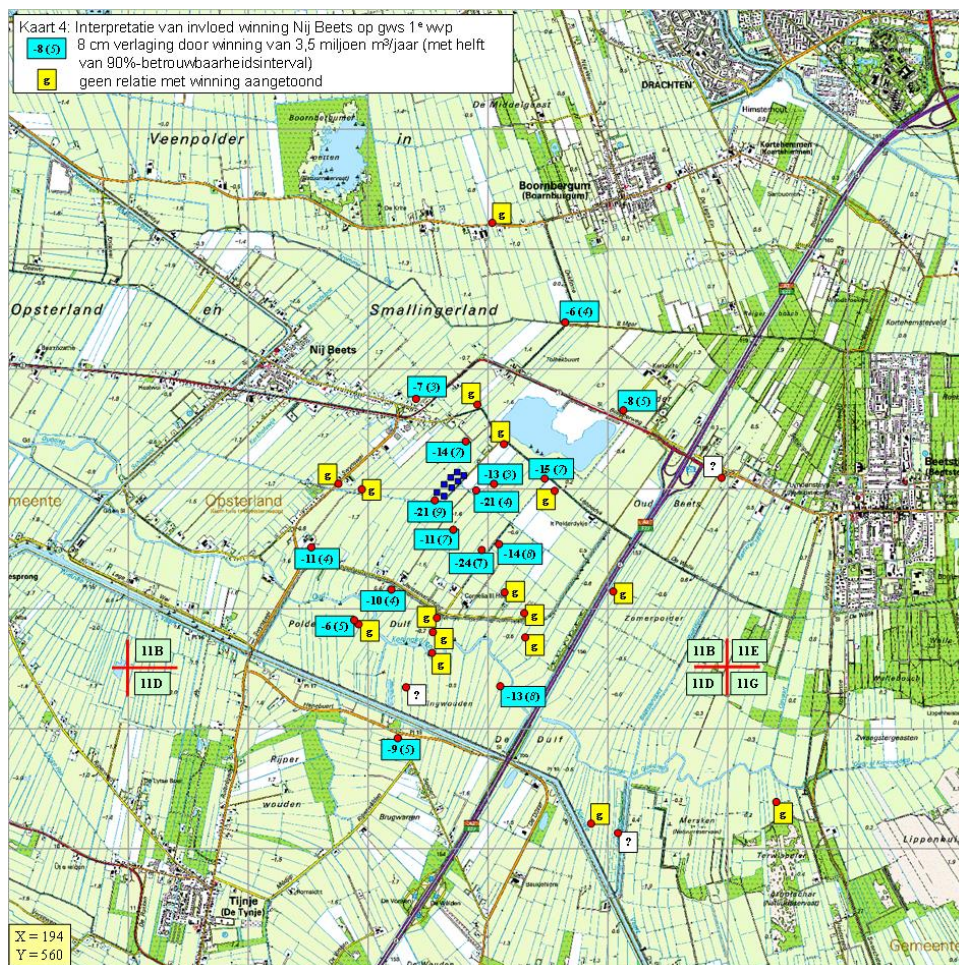
Afbeelding 7.14 Kaart met de geschatte evenwichtsrelaties voor het 1^e watervoerende pakket



Uit Afbeelding 7.13 blijkt dat het beeld van de geschatte verlagingen in het 1^e watervoerende pakket vrij heterogeen is, met op korte afstand van elkaar soms een verlaging en geen verlaging. Vermoedelijk speelt de afstand van de betreffende peilbuis tot een waterloop hierbij een rol. De grootste geschatte verlagingen betreffen peilbuizen binnen een straal van 600 meter van het zwaartepunt van de winning, maar deze bedragen niet meer dan 25 cm. Buiten deze straal blijft de geschatte verlaging doorgaans beperkt tot circa 10 cm. Er zijn drie afwijkingen op dat algemene beeld, maar er zijn voor die gevallen gegronde redenen om te twijfelen aan de causaliteit van de geschatte relaties. Deze afwijkende resultaten worden daarom te discutabel geacht om verder mee te nemen.

Afbeelding 7.15 geeft het ruimtelijk beeld van onze interpretatie van de invloed van de winning Nij Beets op de grondwaterstand in het 1^e watervoerend pakket.

Afbeelding 7.15. Kaart met onze interpretatie van de invloed van de winning Nij Beets op de grondwaterstand in het 1^e watervoerende pakket



Bij de tijdreeksmodellering is er in 45 van de 67 gevallen een statistisch significante relatie met de winning Nij Beets vastgesteld, bij 90 % betrouwbaarheid. Een statistisch significante relatie is echter nog geen bewijs dat het ook een causale relatie betreft.

De grootste geschatte verlagingen treden op in het 2^e en 3^e watervoerende pakket op de winplaats, van respectievelijk 155 en 175 cm bij 3,5 miljoen m³/jaar. In elk watervoerend pakket is de geschatte verlaging doorgaans groter naarmate de peilbuis dichterbij de winning ligt. En de verlagingen zijn in het 3^e

watervoerende pakket – het winpakket - doorgaans groter dan in de andere watervoerende pakketten, waarbij het verschil toeneemt naarmate de peilbuis dichterbij de winning ligt.

Het is aan de onderzoeker om te concluderen óf en zo ja, in hoeverre een geschatte evenwichtsrelatie een causale relatie vertegenwoordigt. De tijdreeksanalyse alleen levert daarvoor namelijk geen bewijs. Er zijn geen aanwijzingen dat de resultaten vertekend zijn en daarom mogen we veronderstellen dat de statistisch significant geschatte evenwichtsrelaties met de winning Nij Beets ook causale relaties vertegenwoordigen.

8

VOORBEELDEN TIJDREEKSANALYSE MET BJTR-METHODE

Dit hoofdstuk geeft enkele voorbeelden van het toepassen van de BJTR-methode. De modelleringen zijn uitgevoerd met het programma Tijdreeksanalist en kunnen daarmee desgewenst ook zelf worden uitgevoerd. Daarvoor is een oefenversie van dat programma beschikbaar, die de volledige functionaliteit heeft, maar alleen werkt op de meegeleverde databestanden. Bij de hier beschreven voorbeelden wordt zoveel mogelijk de uitvoer van het programma weergegeven, zodat alles eventueel ook zonder gebruik van Tijdreeksanalist is te volgen. Toch is het aan te raden de voorbeelden ook zelf uit te werken met het programma, omdat dan de materie beter zal landen. Bovendien kan dan ervaring worden opgedaan met de gevoeligheid van resultaten voor enigszins andere keuzen bij het modelleren.

De beschreven praktijkvoorbeelden zijn de volgende:

- 1 modelleren grondwaterstand met PNO als invoer;
- 2 deterministisch voorspellen grondwaterstand;
- 3 stochastisch voorspellen grondwaterstand.

Installeren oefenversie

De oefenversie van Tijdreeksanalist kan als volgt kan worden geïnstalleerd:

- gebruik deze link: <https://www.amo-nl.com/software/tijdreeksanalist-trg#Licenties>
- ga naar het eind van de webpagina en klik daar op **wachtwoord**
- er verschijnt een formulier dat moet worden ingevuld en verzonden (knop Verzenden)
- u ontvangt een automatische e-mail met daarin de link naar de downloadpagina



Over de opmaak van de voorbeelden

De voorbeelden maken gebruik van Tijdreeksanalist. Dit programma heeft 11 hoofdmenu's, die zijn te vinden in het hoofdscherm (zie afbeelding hierboven).

Bij de beschrijving van de voorbeelden zijn aan te klikken hoofdmenu's in **blauw** aangegeven als: Menu **Sessie**, Menu **Metingen**, etc.

Submenu's of invulvakken zijn aangegeven tussen verticale strepen, zoals: | Grafiek in- en output |

Nadere uitleg is in **rood** aangegeven, zoals: ... **het BJTR-model kan worden beoordeeld** ...

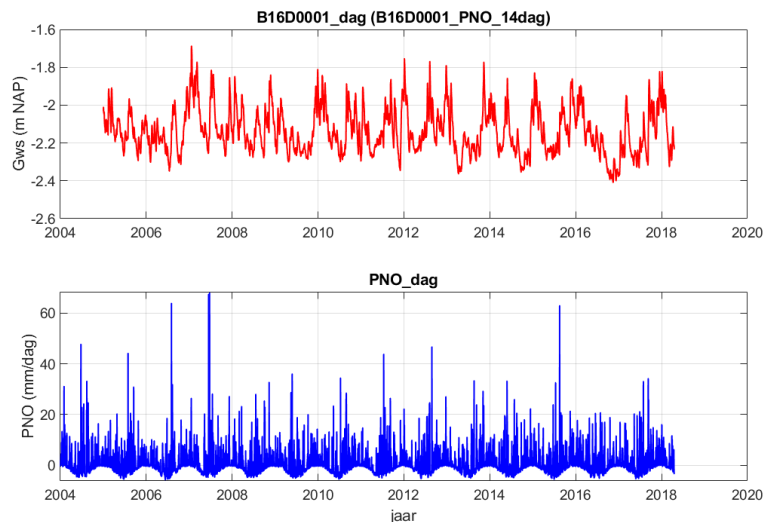
8.1 Modelleren grondwaterstand met PNO als invoer

Inlezen, verkennen en zondig voorbereiden data

Menu [Sessie](#) | Nieuw | Sessienaam - Vul in: B16D0001_PNO_14dag (of geef deze een eigen naam) |
| Output - Kies: B16D0001_dag.trg | Input - Kies: PNO_dag.trg | Ok |

Menu [Metingen](#) | Grafiek in- en output |

De ingelezen reeksen op dagbasis (grondwaterstand en PNO, zónder middeling) worden parallel geplot (zie hieronder). Zo kan visueel worden gecontroleerd op eventuele problemen. Maar hier blijkt geen sprake van duidelijke problemen.

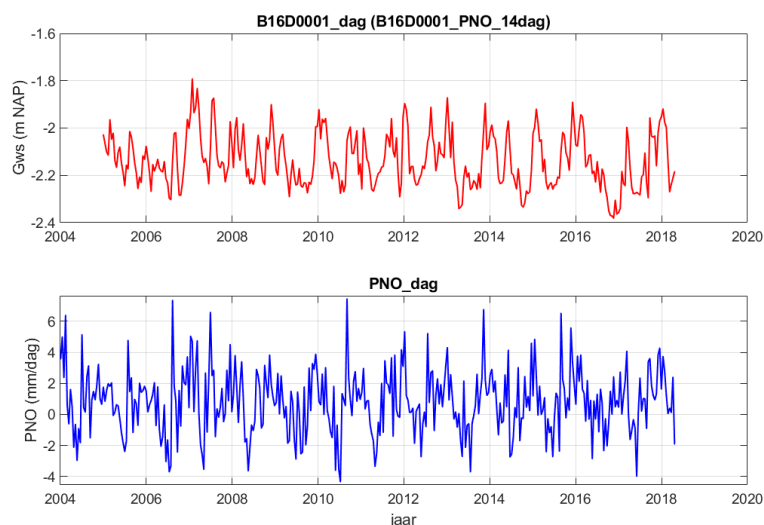


Menu [Bewerking](#) | Middelingstijd - Vul in: 14 | Ok |

Elke reeks wordt hierdoor geaggregeerd tot een reeks 14-daagse gemiddelden. Het middelen dient om een grotere garantie te verkrijgen dat het modelresidu normaal verdeelde witte ruis vormt en tevens dat de transferfunctie van PNO en grondwaterstand weinig parameters nodig heeft. De ervaring heeft geleerd dat een dergelijke garantie bij modelleren van de grondwaterstand op dagbasis vaak ontbreekt.

Menu [Metingen](#) | Grafiek in- en output |

De reeksen van 14-daagse gemiddelden (zie hieronder) zijn minder grillig en gepiekt dan de reeksen op dagbasis.



Menu [Opties](#) | Correlatieparameters | Vul in bij **Maximale tijdsvertraging**: 30 | Vul in bij **Orde AR-model voor prewhitening**: 30 | Ok |

Maximale tijdsvertraging stelt de maximale tijdsverschuiving in van de horizontale as van de ACF-plot en die van de PACF-plot. Zorg dat deze ruim een jaar wordt (zoals 30 bij middelingstijd 14 dagen, omdat dit 420 dagen omvat). Zo kunnen ACF en PACF ook worden beoordeeld op een eventuele autocorrelatie op jaarbasis (door een seizoenseffect).

Orde AR-model voor prewhitening stelt de orde in van het lineaire filter van de invoerreeks, dat gebruikt wordt bij het automatisch witten van het kruiscorrelogram van invoerreeks en uitvoerreeks. Dit is toegelicht in paragraaf 4.2.1, bij het onderdeel *Automatisch witten kruiscorrelogram*.

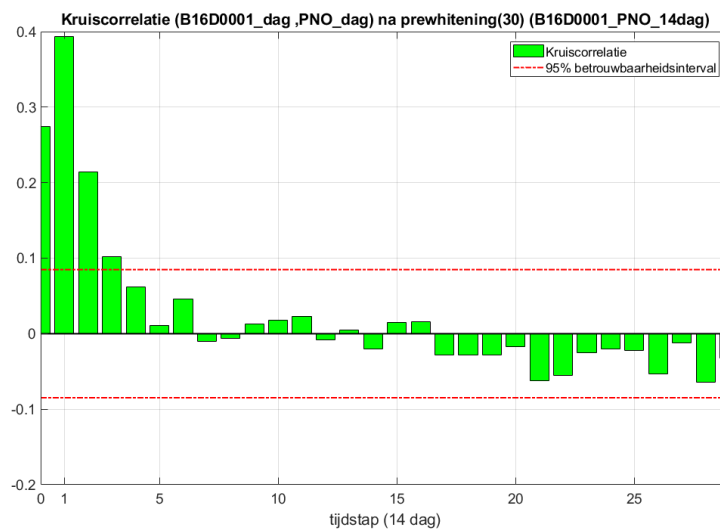
Menu [Opties](#) | Bewaar Opties |

Hiermee worden de instellingen van **Maximale tijdsvertraging** en **Orde AR-model voor prewhitening** bewaard als nieuwe standaardinstellingen.

Identificeren transferfunctie

Menu [TR-Modellering](#) | Kruiscorrelatiefunctie na prewhitening |

Het gewitte kruiscorrelogram (zie hieronder) suggereert een transferfunctie van PNO en grondwaterstand met twee MA-parameters (ω_0 en ω_1) en één AR-parameter (δ_1). Zie daarvoor de voorbeelden in Afbeelding 4.4. Het witten is automatisch uitgevoerd met het eenvoudige lineaire filter van PNO, dat is afgeleid zoals beschreven in paragraaf 4.2.1, bij het onderdeel *Automatisch witten kruiscorrelogram*.



| Geef aantal omega's | Vul in: 2 | Geef aantal delta's | Vul in: 1 | Ok |

De getoonde kruiscorrelatiefunctie betreft een gewitte versie. De standaardinstelling van Tijdreeksanalist voor het aantal omega's (ω) is 2 en voor het aantal delta's (δ) 1. Louter toevallig zijn dat hier voor PNO ook de te adviseren instellingen, gegeven de vorm van de gewitte kruiscorrelatiefunctie. Zie paragraaf 4.2.1 voor hoe een transferfunctie te identificeren.

Formuleren rest van BJTR-model (ruismodel)

Menu [TR-Modellering](#) | Kies ordes | Vul in bij **AR-deel**: 1 | Ok |

De aantallen parameters (1 δ en 2 ω) voor de transferfunctie van PNO waren al ingevuld bij de voorgaande handeling. Na het invullen van 1 bij **AR-deel** van het ruismodel ziet het scherm **Kies ordes** er uit als hieronder.

Bij Ruismodel is het eerste getal bij AR-deel het aantal AR-parameters van het ruismodel (aantal ϕ). Bij de 1^e modelronde kan dit op 1 worden gezet. En het eerste getal bij het MA-deel is het aantal MA-parameters van het ruismodel (aantal θ). Bij de 1^e modelronde kan dit op nul worden gezet. Dit initiële ruismodel omvat alleen een AR-parameter en is daarmee geformuleerd als: $N_t - c = \phi_1(N_{t-1} - c) + a_t$, waarin c een constante. Pas later, na het schatten van het BJTR-model, kan worden beoordeeld of het ruismodel moet worden aangepast. De ruis is immers een latente variabele, die pas tevoorschijn komt na de 1^e model-schatting.

Schatten BJTR-model en verkennen resultaten

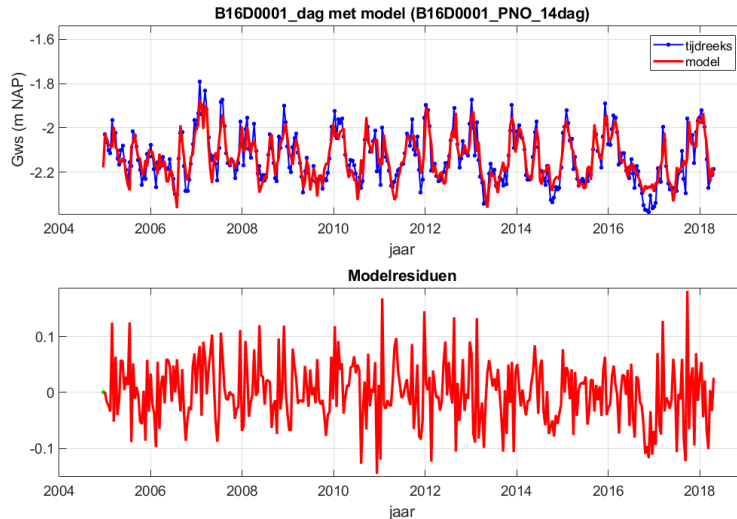
Menu **TR-Modelling** | Schat parameters |

Tijdens het schatten verschijnt er een schermje dat kwantitatieve kenmerken van het schattingsproces toont. Bij voorkeur moet de kolom **verschil** in dat schermje uiteindelijk alleen waarden bevatten die nul of vrijwel nul zijn. Als dat na het stoppen van het schatten niet het geval is – het ingestelde maximum aantal iteraties¹ is dan blijkbaar bereikt zonder voldoende convergentie -, kan het optimalisatieproces hervat worden. Dit kan met de knop **Verder** (met inachtneming van het maximaal iteraties), of met de knop **Stap** (er wordt dan slechts één iteratie uitgevoerd). Dit alles om zonnodig uit een lokaal minimum te kunnen komen. Als de kolom **verschil** alleen waarden bevat die nul of vrijwel nul zijn, of als dat bij herhaald proberen uiteindelijk onhaalbaar blijkt, dan kunnen de resultaten worden verkend en geverifieerd.

Parameters	beste	vorige	verschil
Optimalisatiedoel	0.0031064	0.0030974	0.0000090
Doelfunctie	0.0026330	0.0026330	0.0000000
Verandering	0.0000000	0.0000000	0.0000000
δ_{11}	0.4653212	0.4653212	0.0000000
ω_{10}	0.0194834	0.0194834	0.0000000
ω_{11}	-0.0178671	-0.0178671	0.0000000
Cst	-2.1915454	-2.1915454	0.0000000
ϕ_1	0.6443184	0.6443184	0.0000000

¹ Dit aantal kan worden aangepast met **Menu Opties** | Optimalisatiecriteria | Vul in bij **Maximaal aantal iteraties**

Na herhaald proberen blijft alleen de eerste waarde in de kolom **verschil** ongelijk nul, maar het betreft een kleine waarde. De resultaten kunnen worden verkend en geverifieerd. De AR(1)-parameter van het ruismodel (ϕ_1) is niet groter dan 0,95, zodat we mogen aannemen dat de ruis stationair is. Er worden gedurende het schatten ook twee plots getoond, die laten zien hoe het BJTR-model (rood) de tijdreeks (blauw) nadert. De plot daaronder toont de bijbehorende veranderingen in de reeks van het model-residu a_t . Hieronder de twee plots zoals die eruit zien ná het schatten.



Verifiëren BJTR-model

Menu [Verificatie TR-model](#) | [Parameters en verklaarde variantie](#) | [Resultaten](#) |

Er verschijnt een schermje met de aanduiding *De evenwichtsrelatie en verklaarde varianties*. Het bovenste deel bevat geschatte evenwichtsrelatie(s) en bijbehorende standaardfout(en). Het eerste middendeel bevat varianties van ruis en residu en per invoerreeks het percentage variantie dat zijn transferfunctie verklaart van de uitvoerreeks. Het tweede middendeel vermeldt gemiddelden van ruis en residu. Het onderste deel geeft de resultaten van het toetsen van de residuen op normaliteit, geen autocorrelatie en homoscedasticiteit (geen heteroscedasticiteit), met uitkomst (**Ja**, bij voldoen en een rood gekleurde **Nee** bij niet voldoen) en p-waarde of toetsingsgrootheid. Hier vormen de residuen volgens deze toetsen normaal verdeelde witte ruis.

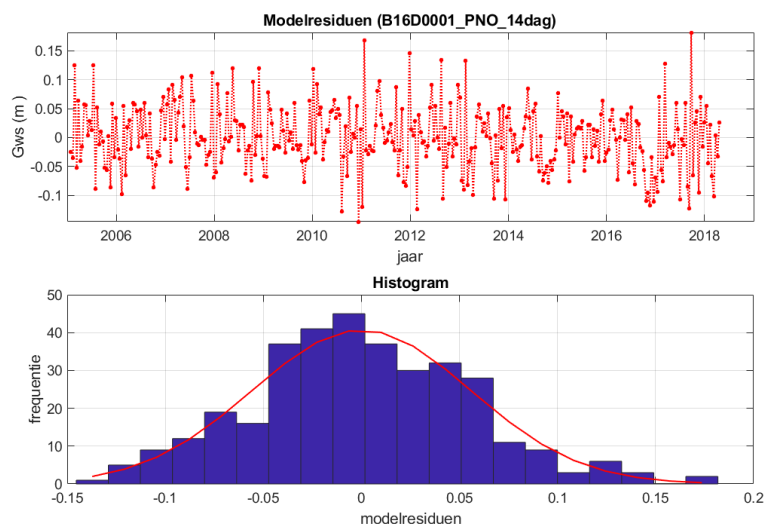
Transfermodel	omschrijving/eenheid	evenwichtsrela	standaardfout
PNO_dag(BJ)	PNO (mm/dag)	0.06985590	0.00766664
Naam	omschrijving/eenheid	variantie	percentage
B16D0001_dag	Gws (m NAP)	0.01275098	
Transfermodel PNO_dag(BJ)	PNO (mm/dag)	0.00872261	62.10 %
Modelruis	verschil tussen meting en simulatie	0.00532411	37.90 %
Modelresiduen	verschil tussen meting en fit	0.00310640	
Akaike FPE	modelresiduen gewogen naar het aantal parameters	0.00319723	
Simulatie	gegenereerd met de invoervariabelen	0.00872261	
Naam	omschrijving/eenheid	gemiddelde	
Modelresiduen	B16D0001_dag - fit	0.00001586	
Modelruis	B16D0001_dag - simulatie	0.00088812	
Toets modelresiduen	toetst	uitkomst	detail
Lilliefors-toets	normaal verdeeld (99%)	Ja	p = 0.44
Jacque-Bera-toets	normaal verdeeld (99%)	Ja	p = 0.33
Portmanteau-toets	geen autocorrelatie (95%)	Ja	14.20 < 23.68 (14)
Runs-toets	geen autocorrelatie (95%)	Ja	p = 0.87
Engle-toets	geen heteroscedasticiteit residuen (95%)	Ja	p = 0.35
Breusch-Pagan-toets	geen heteroscedasticiteit residuen-fit (95%)	Ja	p = 0.62

Volgens deze resultaten is bijvoorbeeld de evenwichtsrelatie van de 14-daagse gemiddelden van PNO en grondwaterstand geschat op 7,0 cm per mm/d, met standaardfout 0,8 cm per mm/d.

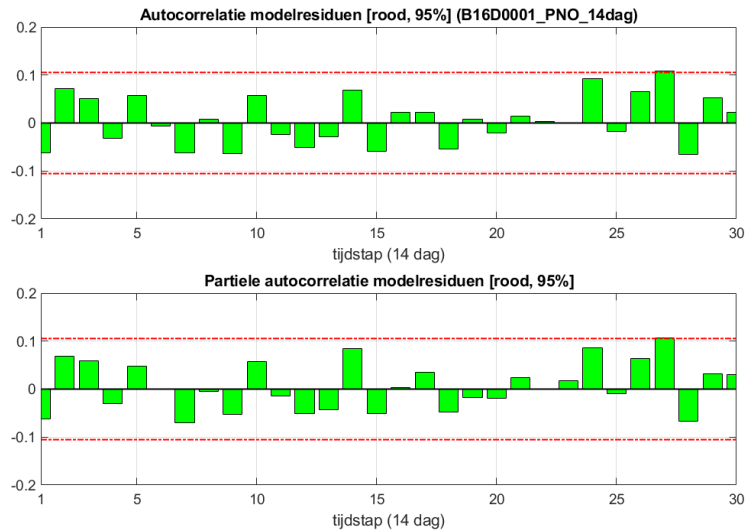
Menu [Verificatie TR-model](#) | Modelresiduen |

Beoordeel of de plot van de modelresiduen (1^e plot hieronder) geen vreemd verloop of grote anomalieën vertoont. Maar die zijn hier niet te zien.

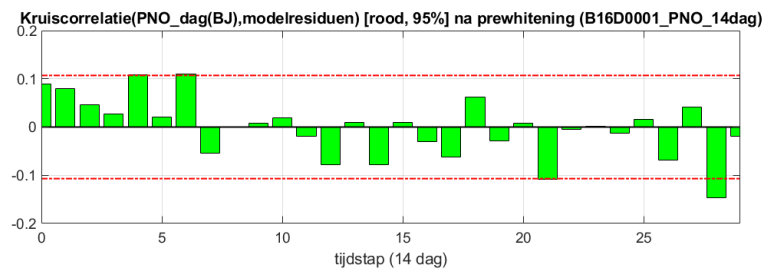
En beoordeel het histogram van de modelresiduen (2^e plot hieronder) visueel op het niet teveel afwijken van normaliteit. De rode lijn in het histogram toont welke vorm een normale kansverdeling met gemiddelde en standaardafwijking van de residuen zou hebben. Er zijn geen aanwijzingen voor grote afwijking van normaliteit.



Beoordeel de combinatie van ACF en PACF van de modelresiduen (zie hieronder) op aanwijzingen van autocorrelatie. Anders moet het ruismodel hier op worden aangepast. Zie daarvoor de adviezen in paragraaf 4.4.2. Deze combinatie van ACF en PACF bevat echter geen aanwijzingen van autocorrelatie, ook niet op jaarbasis (26 tijdstappen).

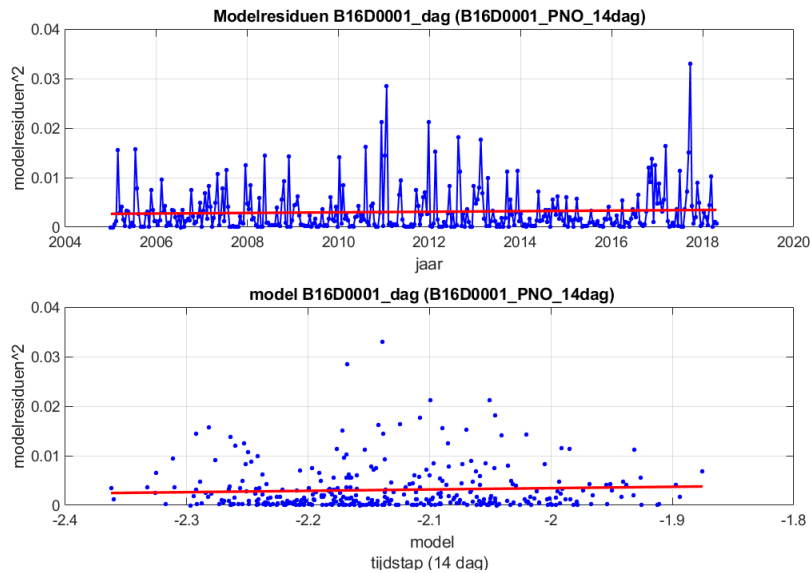


Beoordeel het kruiscorrelogram van PNO en modelresiduen (zie hieronder) op een mogelijk nog resterende relatie. Als daar duidelijk sprake van is, pas dan de transferfunctie van PNO aan. Er zijn hier echter geen duidelijke aanwijzingen voor een nog resterende relatie.



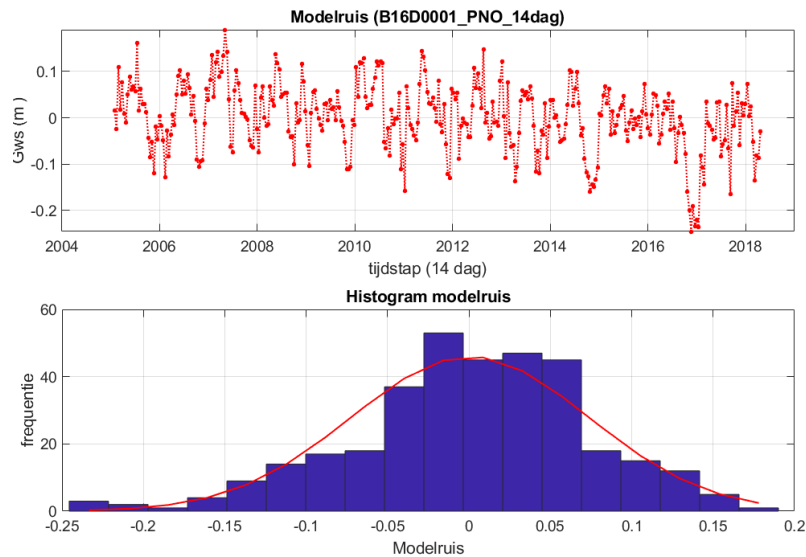
Menu [Verificatie TR-model](#) | Heteroscedasticiteit |

Beoordeel de twee plots op een duidelijke relatie tussen het gekwadrateerde modelresidu (\hat{a}_t^2) en tijd (bovenste plot) en tussen \hat{a}_t^2 en de uitkomst van het BJTR-model (onderste plot). In geen van beide plots is echter een duidelijke relatie te zien. Dit stemt overeen met de uitkomsten van respectievelijk de Engle-toets en de Breusch-Pagan-toets (zie ook het eerder besproken een schermje met de aanduiding *De evenwichtsrelatie en verklaarde varianties*).

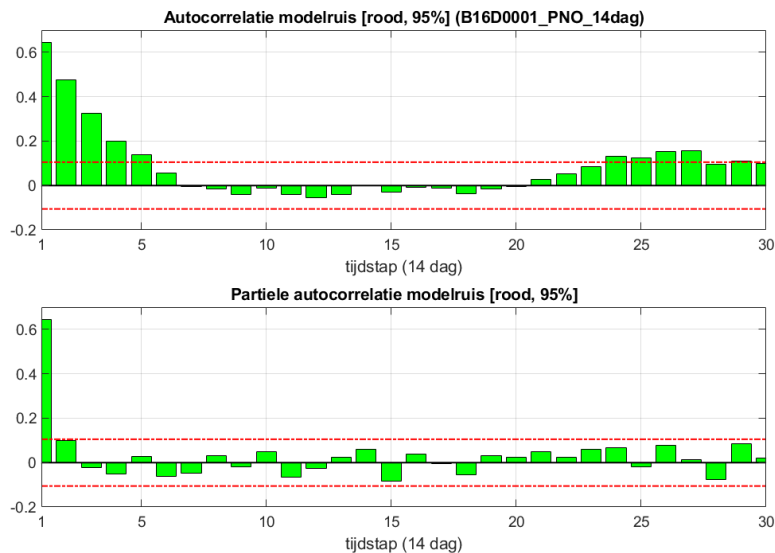


Menu [Verificatie TR-model](#) | Modelruis |

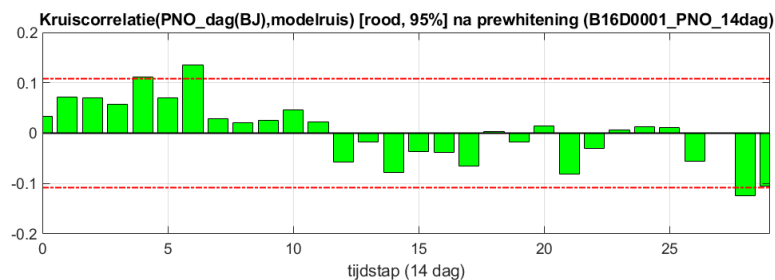
Beoordeel plot en histogram van de modelruis (zie hieronder). De plot bevat een aanwijzing dat het deterministische deel van het model minder goed de grondwaterstand van eind 2016 / begin 2017 beschrijft. Dit zullen we bij het valideren van dit BJTR-model verder onderzoeken (zie daarvoor Voorbeelden 2 en 3).



Mocht bij het hiervoor uitgevoerde beoordelen van de modelresiduen naar voren zijn gekomen dat die geen witte ruis vormen, beoordeel dan de combinatie van ACF en PACF van de modelruis om een beter ruismodel te kunnen identificeren. Maar dat is hier niet nodig gebleken. Deze combinatie wijst hier overigens op een ARMA(1,0)-proces van de ruis (zie hieronder) en zo was het ruismodel ook geformuleerd.



Beoordeel het kruiscorrelagram van PNO en modelruis (zie hieronder) op een mogelijk nog resterende relatie. Als daar duidelijk sprake van is, pas dan de transferfunctie van PNO aan. Er zijn hier echter geen sterke aanwijzingen voor een nog resterende relatie.



Als het model niet goed door de verificatie is gekomen, pas het dan aan.
Ga daarvoor terug naar [Identificeren transferfunctie](#).

Sla de sessie dan op. Deze wordt gebruikt bij Voorbeeld 2.

Menu [Sessie](#) | Bewaar als – Vul in: pad en naam (uitgang wordt .sess) |

8.2 Deterministisch voorspellen grondwaterstand

Zie de technische toelichting op deterministisch voorspellen in paragraaf 6.2.2.

Inlezen data

Menu [Sessie](#) | Open | Selecteer de sessie die bij Voorbeeld 1 is vervaardigd (met middelingstijd 14 dagen en gekalibreerd over de periode 2005 t/m medio april 2018) | Ok |

Menu [Bewerking](#) | Deelreeks output | Vul in bij **Einddatum**: 31/12/2014 |

Bij het afleiden van het BJTR-model zal het programma nu alleen nog de deelreeks van de grondwaterstand t/m 2014 hanteren. Zo wordt het model gekalibreerd over de periode 2005 t/m 2014.

Menu [Opties](#) | Correlatieparameters |

Check of **Maximale tijdsvertraging** staat op: 30 en check of orde **AR-model voor prewhitening** staat op: 30. Zo niet, vul dan die waarden in. | Ok |

Als er waarden moesten worden ingevuld, bewaar ze dan met:

Menu [Opties](#) | Bewaar Opties |

Hiermee worden de aanpassingen bewaard als nieuwe standaardinstelling.

Identificeren transferfunctie, formuleren ruisdeel en schatten en verifiëren BJTR-model

Doorloop alle stappen die hiervoor bij Voorbeeld 1 zijn beschreven vanaf [Identificeren transferfunctie](#). Zo wordt het BJTR-model afgeleid van het deel van de betreffende grondwaterstandreeks dat loopt van 2005 t/m 2014. Daarbij zal blijken dat de vorm van het model hetzelfde is als bij Voorbeeld 1 en ook dat de modelparameters weinig verschillen met die van Voorbeeld 1.

Sla dit model voor de periode 2005 t/m 2014 op, want het is straks ook nodig voor Voorbeeld 3.

Menu [Sessie](#) | Bewaar als – Vul in: pad en naam (uitgang wordt .sess) |

Menu [Opties](#) | Simulatie- en voorspelparameters |

Check of zowel **Fractie van de modelresiduenvariantie** als **Fractie van de parametervariantie** op 1 staat en zo niet, zet beide dan op 1.

| Vul in bij **Aantal voorspellingen**: 0 | Ok |

Bij deze instelling wordt voor het voorspellen alleen het deterministische deel van het BJTR-model gebruikt. Dat deel bestaat hier uit de transferfunctie van PNO en grondwaterstand. Die voorspellingen houden nog geen rekening met het stochastische deel van het BJTR-model (het ruismodel) en ook niet met de onzekerheden van de geschatte parameters van de betreffende transferfunctie. In Voorbeeld 3 wordt daar wel rekening mee gehouden.

| Vul in bij **Begindatum simulaties/voorspellingen**: 01/01/2015 |

| Vul in bij **Einddatum simulaties/voorspellingen**: 01/01/2030 |

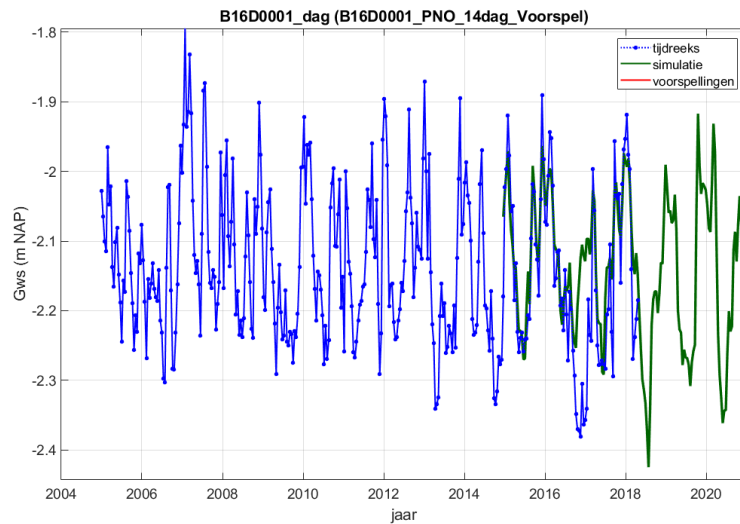
Menu [Opties](#) | Bewaar Opties |

Hiermee worden de aanpassingen bewaard als nieuwe standaardinstelling.

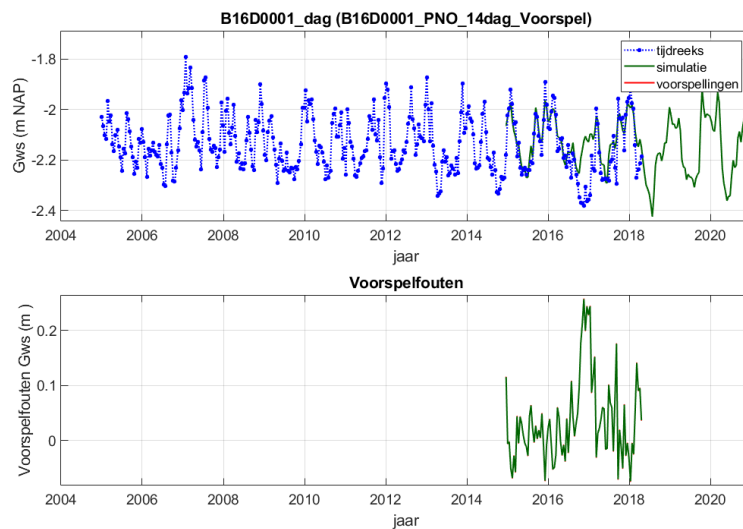
Menu [Verificatie TR-model](#) | Voorspellingen |

Er verschijnen dan vier uitvoerschermen met betrekking tot de voorspellingen, waarbij het model is gekalibreerd over de 10-jaarsperiode 2005 t/m 2014.

Het eerste uitvoerscherm (zie hieronder) toont de plot van de tijdreeks van de grondwaterstand van januari 2005 t/m april 2018, met vanaf 2015 de deterministische voorspelling, gebaseerd op de voor die periode beschikbare waarden van PNO. De deterministische voorspelling (groen, aangeduid als 'simulatie') loopt van 2015 tot de laatst beschikbare PNO-waarde (eind 2020). Van circa midden 2016 tot midden 2017 voorspelt het model tussen 10 en 25 centimeter hogere grondwaterstanden dan gerealiseerd. Dit kan er op wijzen dat er in die periode ook sprake was van een andere invloedsfactor dan PNO.

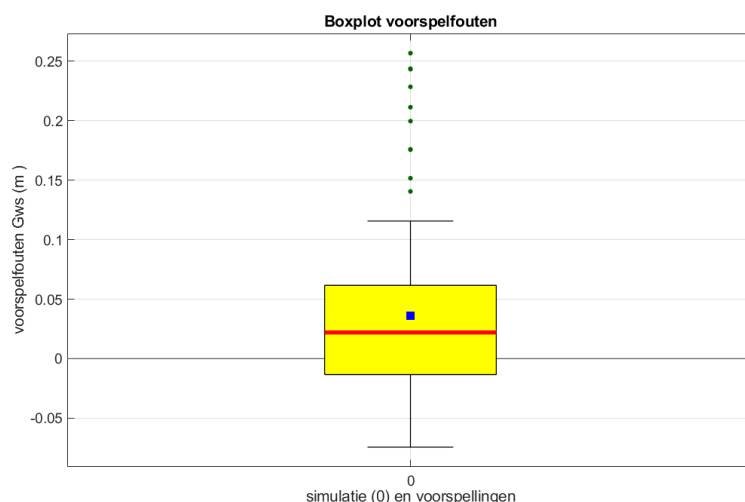


Het tweede uitvoerscherm (zie hieronder) toont boven dezelfde plot als het eerste uitvoerscherm en daaronder de reeks voorspelfouten vanaf 2015. De anomalie van midden 2016 tot midden 2017 is daarin nog duidelijker te zien.



Het derde uitvoerscherm toont vrijwel dezelfde plot als het eerste uitvoerscherm en wordt daarom hier niet weergegeven. Die plot wordt pas onderscheidend bij stochastisch voorspellen (zie Voorbeeld 3).

Het vierde uitvoerscherm toont de boxplot van de voorspelfout (zie hieronder). Deze duidt op een scheve verdeling van de voorspelfout, met veel positieve uitschieters. Verder is de mediaan van de voorspelfout (rode lijn in de box) groter dan nul. Ook dit wijst op het slechter functioneren van het deterministische model over deze periode.



Dit voorbeeld betrof niet alleen het deterministisch voorspellen van de grondwaterstand, maar ook het valideren van het deterministische deel van het BJTR-model met onafhankelijke meetwaarden. Dat was hier mogelijk doordat de grondwaterstanden van een relevant deel van de voorspelperiode beschikbaar zijn.

De voorspellingen van de grondwaterstand betreffen hier 14-daagse gemiddelden. Als er behoefte is aan puntvoorspellingen van de grondwaterstand (dat wil zeggen meetwaarden zonder aggregatie) moet elk 14-daags gemiddelde worden uitgewerkt tot 14 puntwaarden op dagbasis.

8.3 Stochastisch voorspellen grondwaterstand

Zie de technische toelichting op stochastisch voorspellen in paragraaf 6.2.2.

Inlezen data

Menu [Sessie](#) | Open | Selecteer de sessie die bij Voorbeeld 2 is opgeslagen (met middelingstijd 14 dagen en gekalibreerd over de periode 2005 t/m 2014) | Ok |

Menu [Opties](#) | Simulatie- en voorspelparameters | Vul in bij **Aantal voorspellingen**: 100 | Ok |

Dit houdt in dat als met het model wordt voorspeld, daarvoor het volledige BJTR-model wordt gebruikt, dus zowel het deterministische deel als het stochastische deel. Daarbij wordt rekening gehouden met:

- i. de onzekerheden van de geschatte parameters van de betreffende transferfunctie.
- ii. de mate waarin de geschatte parameters gecorreleerd zijn (dit is samengevat in de covariantie-matrix van de geschatte parameters);
- iii. de onzekerheid van de ruis. Daartoe worden met het ruismodel 100 ruisreeksen (N_t) gegenereerd met behulp van trekkingen uit de kansverdeling van het modelresidu (a_t), die vervolgens worden gesuperponeerd op het deterministische deel ($Z_t = \text{deterministisch deel} + N_t$).

Er kunnen ook meer of minder dan 100 ruisreeksen worden gegenereerd, door bij **Aantal voorspellingen** een ander aantal in te vullen. Maar omdat tevens het 95 %-voorspelinterval uit de bundel voorspelde reeksen moet worden bepaald, is aan te bevelen **Aantal voorspellingen** op minstens 100 te zetten. Bij hoge aantallen kan het rekenen overigens wel lang gaan duren.

Controleer of in het betreffende scherm **Begindatum simulaties/voorspellingen** op 01/01/2015 staat en **Einddatum simulaties/voorspellingen** op 01/01/2030. Als dat niet zo is, vul dan die datums in.

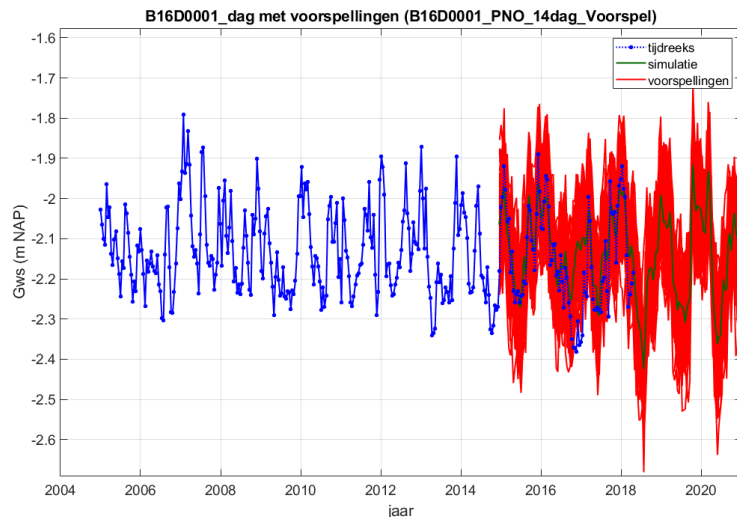
Menu [Opties](#) | Bewaar Opties |

Hiermee worden de aanpassingen bewaard als nieuwe standaardinstelling.

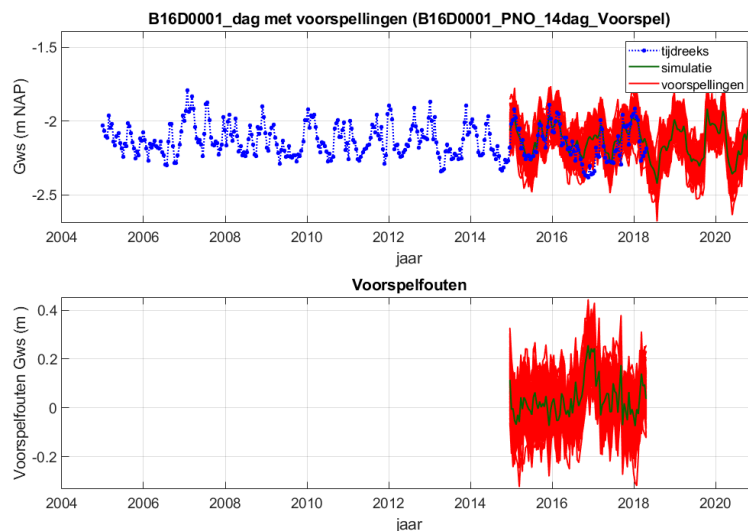
Menu [Verificatie TR-model](#) | Voorspellingen |

Er verschijnen dan vier uitvoerschermen met betrekking tot de stochastische voorspellingen, waarbij het model is gekalibreerd over de 10-jaarsperiode 2005 t/m 2014.

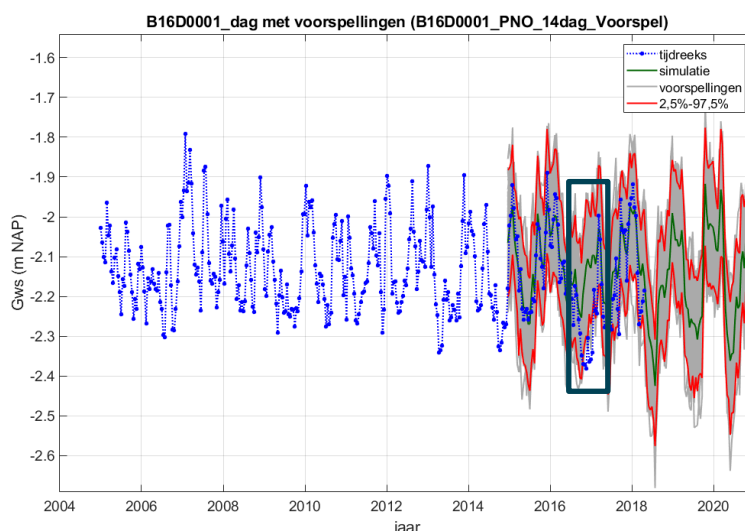
Het eerste uitvoerscherm (zie hieronder) toont de plot van de tijdreeks van de grondwaterstand van januari 2005 t/m april 2018, met vanaf 2015 zowel de deterministische voorspelling (groen, aangeduid als 'simulatie', gebaseerd op de voor die periode beschikbare waarden van PNO), als de bundel van 100 stochastische voorspellingen (rood). Zoals we ook al constateerden bij Voorbeeld 2 lijkt er van circa midden 2016 tot midden 2017 ook sprake van een andere invloedsfactor dan alleen PNO.



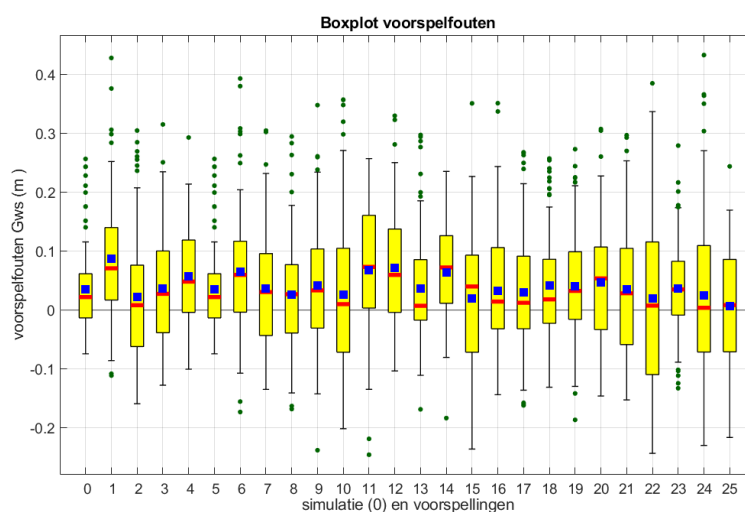
Het tweede uitvoerscherm (zie hieronder) toont boven dezelfde plot als het eerste uitvoerscherm en daaronder de reeks voorspelfouten vanaf 2015. De anomalie van midden 2016 tot midden 2017 is daarin nog duidelijker te zien.



In het derde uitvoerscherm is de voorspelbundel nu lichtgrijs in plaats van rood weergegeven, waarbij ook de grenzen van het 95 %-voorspelinterval zijn weergegeven (rode lijnen), zoals per tijdstip bepaald uit het 2,5-percentiel en het 97,5-percentiel van de daarvoor voorspelde waarden. Het is nu ook duidelijk zichtbaar dat de gerealiseerde grondwaterstand in de periode midden 2016 tot midden 2017 voor een relevant deel onderin of zelfs buiten het 95 %-voorspelinterval ligt van het model dat is gekalibreerd over de periode 2005 t/m 2015.



Het vierde uitvoerscherm toont de boxplot van de voorspelfout (zie hieronder). De eerste boxplot (met 0 op de X-as) is die van de deterministische voorspelling, zoals we die ook al zagen bij Voorbeeld 2. De overige 25 boxplots zijn die van 25 willekeurige reeksen uit de bundel van de 100 gegenereerde grondwaterstandreeksen. De meeste van deze boxplots duiden op een scheve verdeling van de voorspelfout, met veel positieve uitschieters. Verder is bij de meeste de mediaan van de voorspelfout groter dan nul. Ook dit wijst op het slechter functioneren van het BJTR-model over deze periode.



Dit voorbeeld betrof niet alleen het stochastisch voorspellen van de grondwaterstand, maar ook het valideren van een BJTR-model met onafhankelijke meetwaarden. Dat was hier mogelijk doordat de grondwaterstanden van een relevant deel van de voorspelperiode beschikbaar zijn.

De voorspellingen van de grondwaterstand betreffen hier 14-daagse gemiddelden. Als er behoefte is aan puntvoorspellingen van de grondwaterstand (dat wil zeggen meetwaarden zonder aggregatie) moet elk 14-daags gemiddelde worden uitgewerkt tot 14 puntwaarden op dagbasis.

9

VOORBEELDEN EN ACHTERGROND PIRFICT-METHODE

In dit hoofdstuk worden enkele uitgewerkte voorbeelden behandeld voor het uitvoeren van tijdreeksanalyse voor de PIRFICT-methode. Het hoofdstuk bestaat uit 4 delen:

- voorbereken;
- modelstructuur;
- kalibratie;
- beoordeling.

In deze delen worden verschillende voorbeelden gegeven met verschillende tijdreeksmodellen. Het doel hiervan is de lezer een overzicht geven van de verschillende mogelijkheden en keuzes bij het opstellen van de modellen. Deze delen zijn opgesteld om ook op zichzelf leesbaar te zijn, hierdoor kan enige herhaling tussen de delen voorkomen.

De delen van dit hoofdstuk zijn daarnaast ook beschikbaar in de vorm van Jupyter Notebooks op "[LINK invoegen](#)", in deze notebooks kan de gebruiker zelf aan de slag met de voorbeelden en kan de achterliggende modelcode worden ingezien.

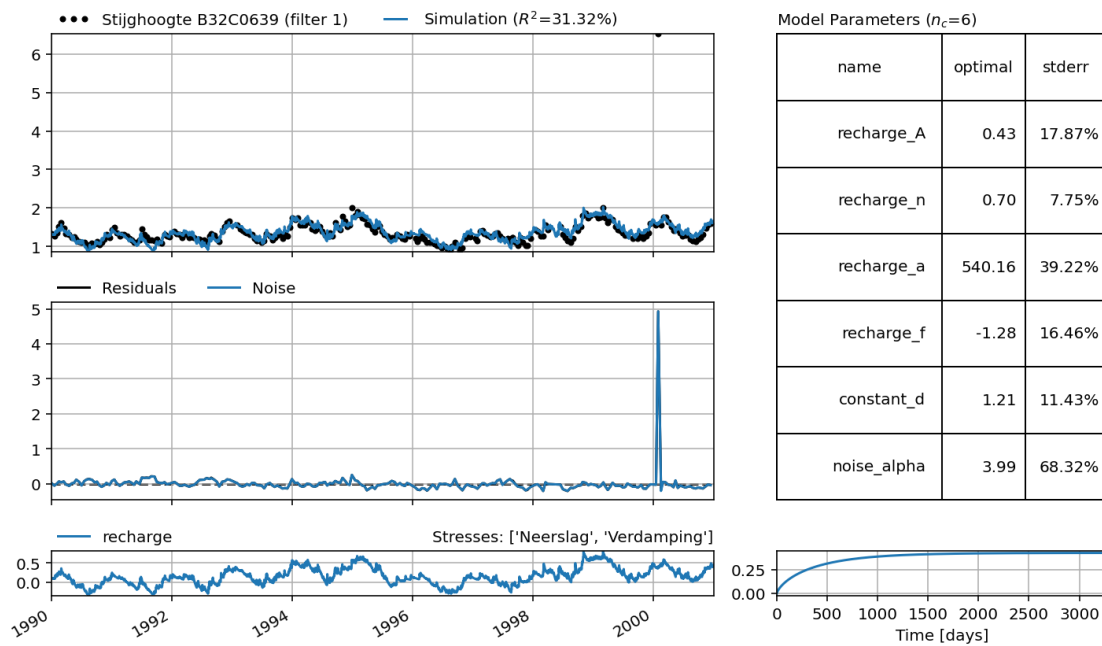
9.1 Voorbewerking van reeksen voor tijdreeksanalyse

Voor tijdreeksanalyse zijn grondwaterstanden en verklarende reeksen nodig. Beide komen in verschillende lengtes, frequenties en kwaliteit. In veel gevallen is een voorbereking van de reeksen nodig. In dit voorbeeld wordt geïllustreerd welke bewerkingen je kan doen en wanneer deze nodig zijn.

9.1.1 Opschonen stijghoogte reeks

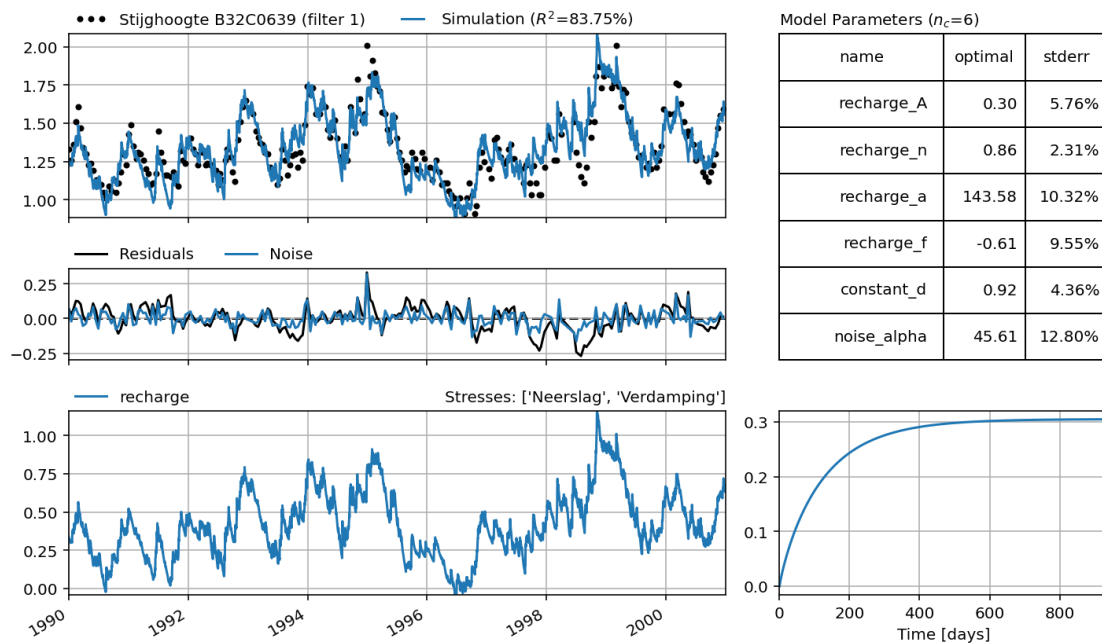
Het doel van tijdreeksanalyse is een grondwaterstands- of stijghoogte reeks te verklaren uit invoerreeksen. Door middel van automatische optimalisatie probeert de software de reeks zo dicht mogelijk te benaderen. Wanneer er fouten in de reeks aanwezig zijn kan dat grote gevolgen hebben voor de kwaliteit van het model. Onderstaand voorbeeld geeft een model dat is geoptimaliseerd zonder een duidelijke uitbijter te verwijderen.

Afbeelding 9.1 Tijdreeksmodel met uitbijter



De resultaten van dit model zijn niet erg bemoedigend. De verklaarde variantie bedraagt slechts 31 %. Dit komt mede door de grote afwijking bij de uitbijter, maar ook de standaardfouten van de parameters zijn relatief groot en de responstijd van de neerslag en verdamping lijkt voor Nederlandse begrippen aan de hoge kant. We gaan nu kijken wat de resultaten worden als we de uitbijter verwijderen.

Afbeelding 9.2 Tijdreeksmodel zonder uitbijter

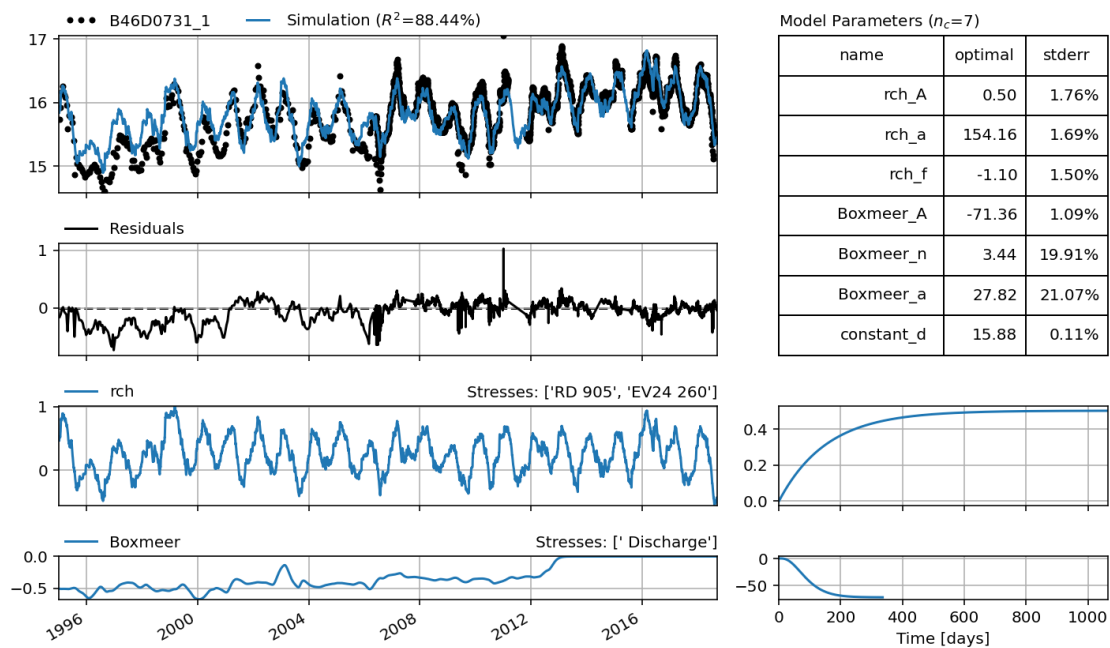


In bovenstaande afbeelding is te zien dat het model veel beter presteert met de opgeschoonde reeks. Het blijkt dat de uitbijter (ook als is het er maar één) een grote invloed heeft op het resultaat. Naast een betere fit (84 %), zijn ook de standaardfouten kleiner en is de respons korter geworden.

9.1.2 Vereenvoudigen stijghoogte reeks

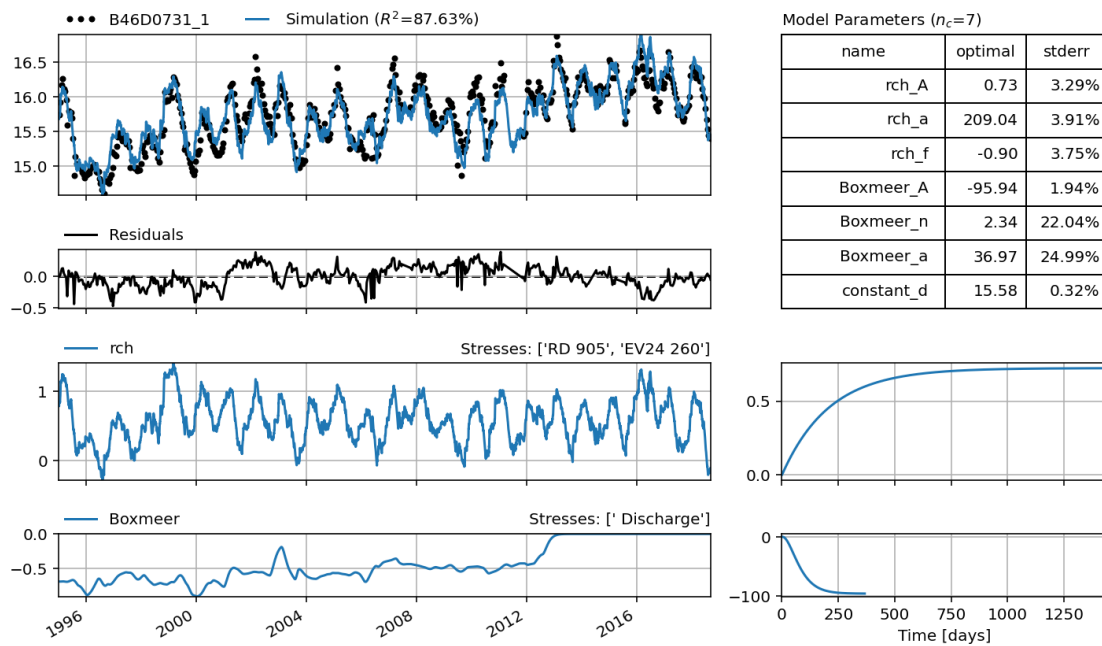
Sommige reeksen bevatten zeer hoogfrequente data, met bijvoorbeeld een tijdstap van een dag of nog korter. Afhankelijk van het doel van de tijdreeksanalyse en de beschikbare invoer, kan het gewenst zijn om de reeksen te aggregeren of uit te dunnen. Het eerste voorbeeld betreft een veelvoorkomende situatie: er is gedurende lange tijd ongeveer tweewekelijks gemeten en sinds enkele jaren op een hoge frequentie. Wanneer we alle metingen gebruiken is het mogelijk dat de nadruk bij de optimalisatie komt te liggen op de periode waar hoogfrequent is gemeten, simpelweg omdat er meer metingen zijn. In het onderstaande voorbeeld is dat het geval.

Afbeelding 9.3 Tijdreeksmodel met hoogfrequente data



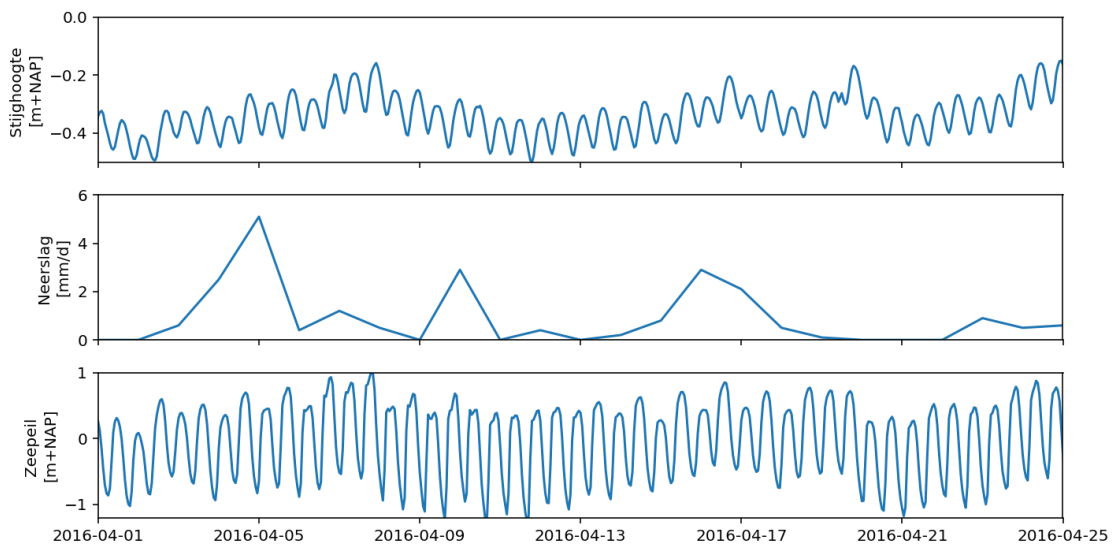
Vooraf in en begin van de reeks zijn de gesimuleerde waarden structureel hoger dan de metingen, terwijl in de tweede helft van de reeks de simulatie gemiddeld beter klopt. Het onderstaande voorbeeld gaat uit van dezelfde invoer, alleen is nu de stijghoogtereeks geresampled zodat er gedurende de hele periode tweewekelijks metingen zijn.

Afbeelding 9.4 Tijdreeksmodel waarbij data is geresampled



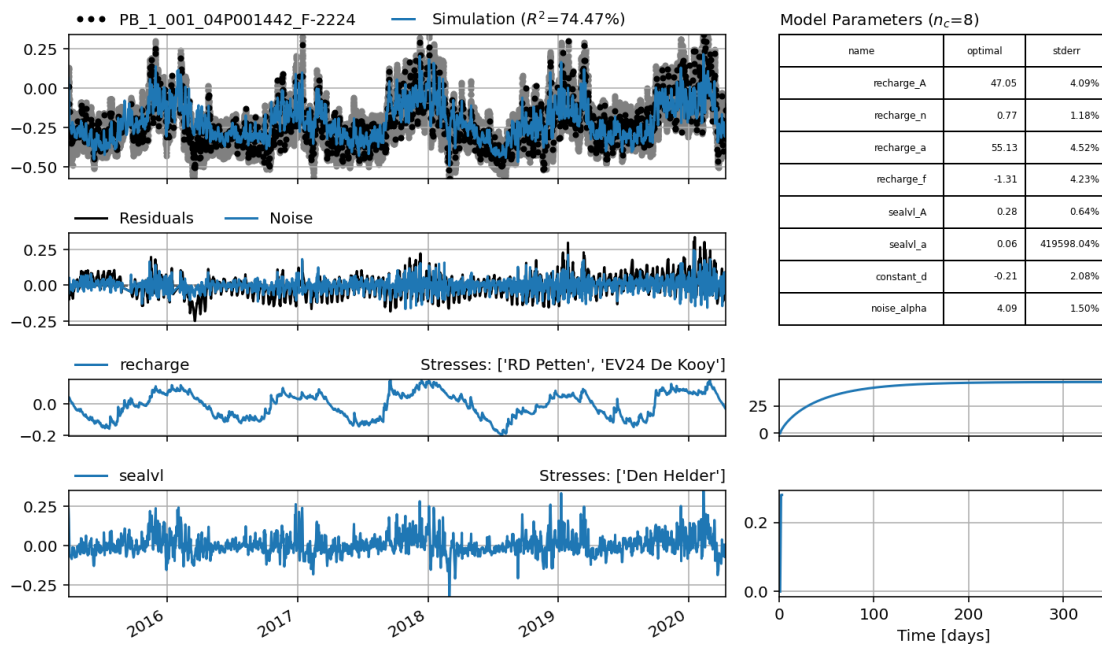
Na het filteren van de hoogfrequente data is te zien dat de residuen nu evenwichter zijn verdeeld over de reeks. De fit in de beginperiode is beter geworden ten koste van de fit aan het eind. Dit heeft ook geleid tot een ander model; ook de responsen van de bijdragen zijn gewijzigd. Er zijn ook situaties waarin het juist belangrijk is om een hoge frequentie te handhaven. In het volgende voorbeeld wordt dit geïllustreerd. Het gaat om een peilfilter dat onder invloed staat van het getij. In onderstaande figuur is goed te zien dat de stijghoogte wordt beïnvloed door het getij.

Afbeelding 9.5 Stijghoogte beïnvloed door getij



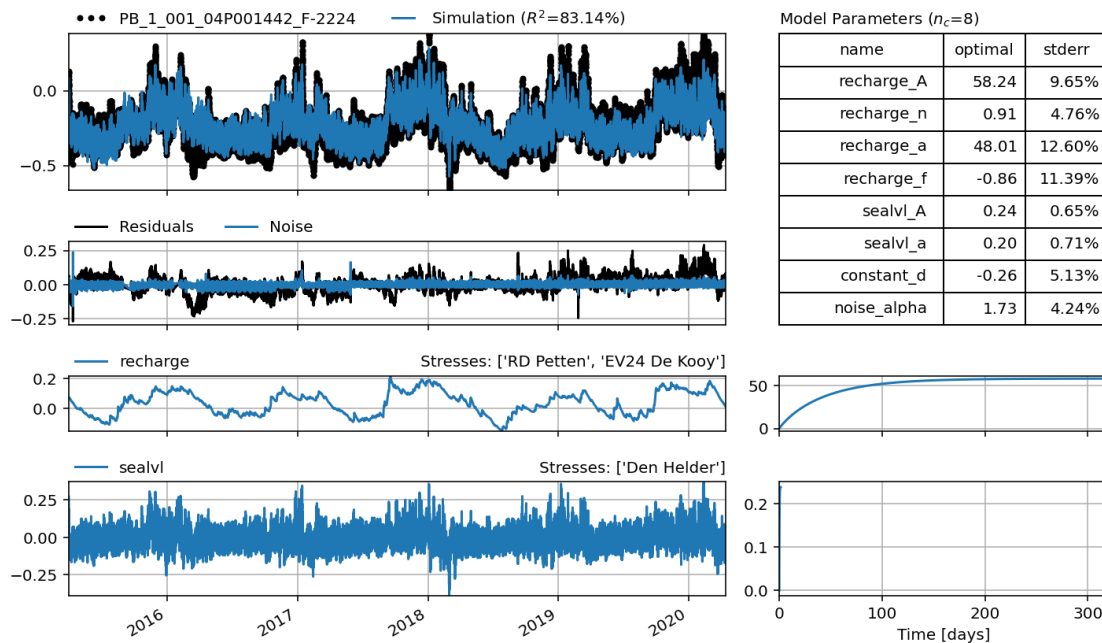
Als eerste poging maken we een model met een tijdsinterval van een dag. Onderstaande afbeelding toont het resultaat:

Afbeelding 9.6 Tijdreeksmodel met tijdsinterval van 1 dag met getij



Het is duidelijk te zien dat de dagelijkse fluctuatie niet wordt meegenomen. Er is ook een duidelijk patroon in de residuen te zien. We proberen nu een simulatie op uurwaarde.

Afbeelding 9.7 Tijdreeksmodel met tijdsinterval van 1 uur met getij

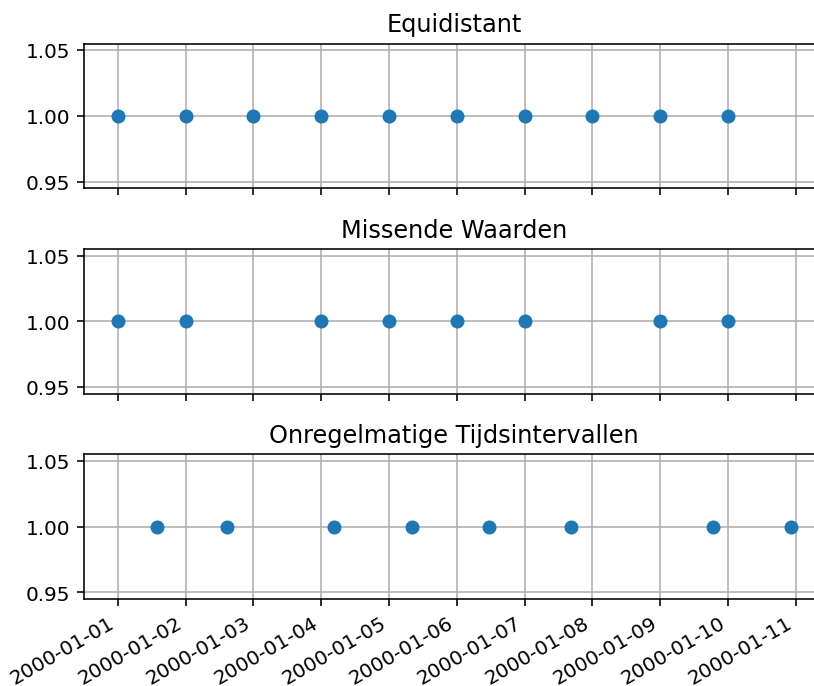


Dit model presteert veel beter, omdat de informatie die in het zeepil aanwezig is beter gebruikt kan worden om de grondwaterstand te verklaren.

9.1.3 Equidistant maken verklarende reeks

Voor de meeste tijdreeksanalyseprogramma's is het nodig dat de verklarende reeksen gelijke tijdstappen hebben, ofwel equidistant zijn (zie ook paragraaf 2.7.1 voor de verschillen tussen BJTR- en PIRFICT-modellen, en paragraaf 3.4.2 en 3.5.2 voor uitleg over het opzetten van een tijdreeksmodel). In sommige gevallen worden ze zo aangeleverd. Als dat niet zo is, is het nodig om de reeksen equidistant te maken. De manier waarop dit gebeurt, hangt af van het type verklarende reeks. Rivierpeilen kunnen bijvoorbeeld geïnterpoleerd worden afhankelijk van de grootte van de tijdstap en de dynamiek van de reeks. Bij onttrekkingsreeksen moet men zich afvragen of een missende waarde betekent dat de onttrekking uit stond, of dat er gegevens ontbreken omdat er geen waarneming heeft plaatsgevonden. Het opvullen van de data introduceert onzekerheden, die niet automatisch worden meegenomen in de analyse. Men kan er ook voor kiezen de tijdstap van de tijdreeksanalyse te vergroten, wanneer dat beter bij de resolutie van de beschikbare gegevens past. In onderstaande afbeelding worden de verschillende type reeksen geïllustreerd:

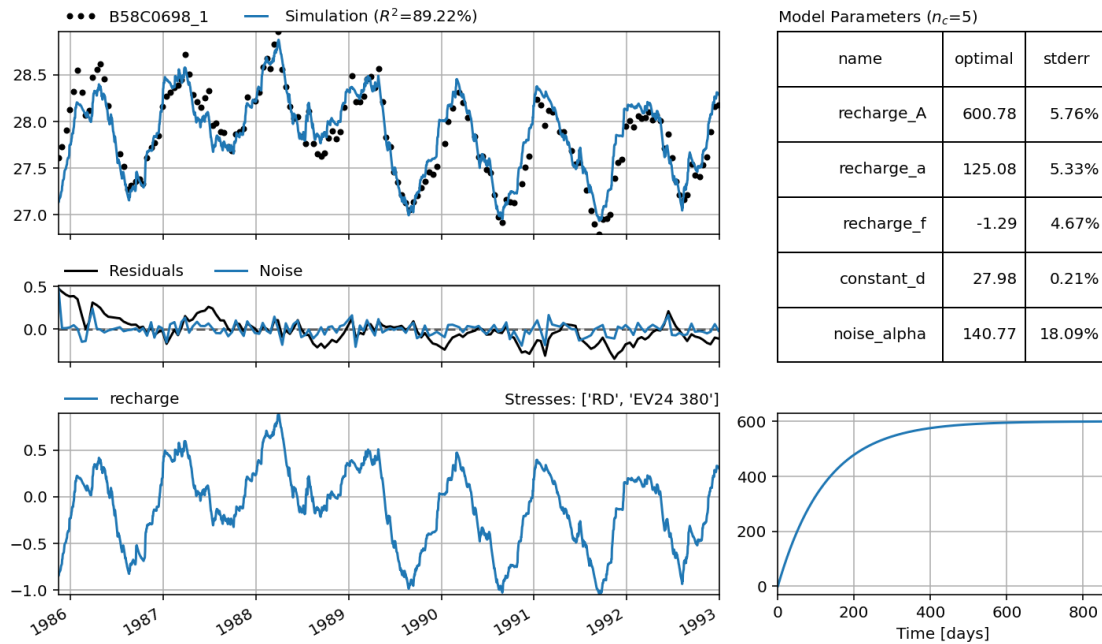
Afbeelding 9.8 Verschillende type reeksen



9.1.4 Voorgeschiedenis verklarende reeks

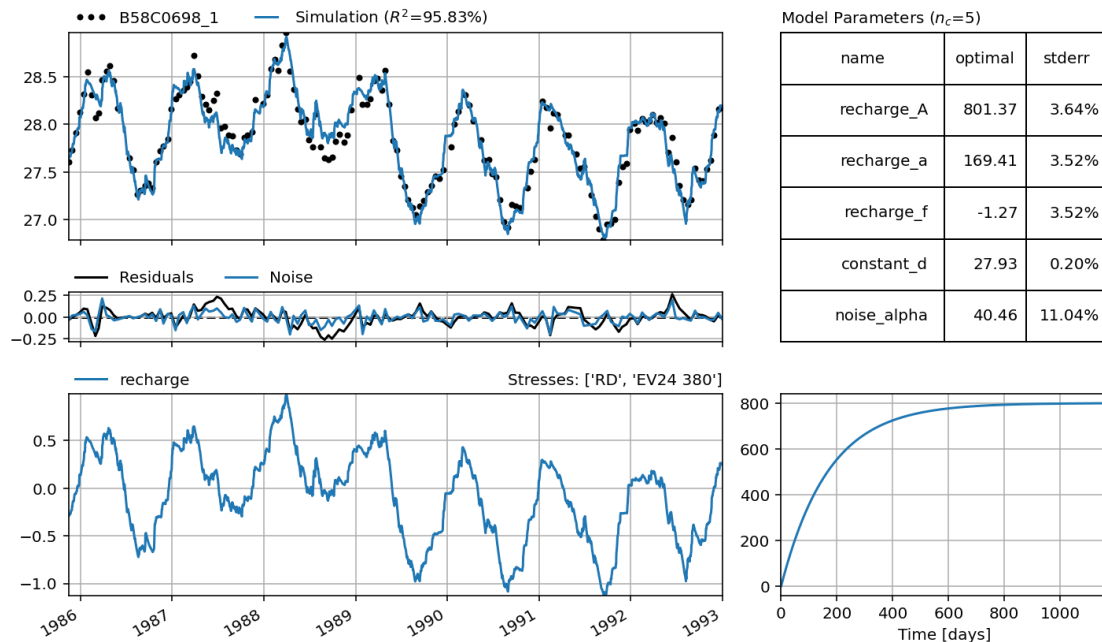
Elk hydrologisch systeem heeft een zeker geheugen. Dit betekent dat de stijghoogte afhangt van wat er een bepaalde periode ervoor is gebeurd. Daarom kent elk model een zekere inlooperperiode. Bij het fitten van een grondwaterstand is het belangrijk dat er voldoende informatie over verklarende reeksen aanwezig is vóór de eerste waarneming waarop gekalibreerd wordt. Wanneer dat niet het geval is, zal er een schatting gemaakt moeten worden, óf de betreffende waarnemingen zullen niet in de optimalisatie meegenomen moeten worden. In onderstaand voorbeeld is de neerslag bekend gedurende een periode van 100 dagen vóór de eerste stijghoogtemeting. Daarvóór wordt aangenomen dat de neerslag 0 is.

Afbeelding 9.9 Tijdreeksmodel zonder inlooperperiode



Het is duidelijk te zien dat simulatie in het begin van de reeks lager ligt dan de metingen. Dit wordt gecompenseerd door een te hoge ligging aan het eind van de reeks. In onderstaand voorbeeld nemen we hetzelfde model, maar nu gaan we uit van de gemiddelde neerslag tijdens de inlooperperiode.

Afbeelding 9.10 Tijdreeksmodel met inlooperperiode met gemiddelde neerslag

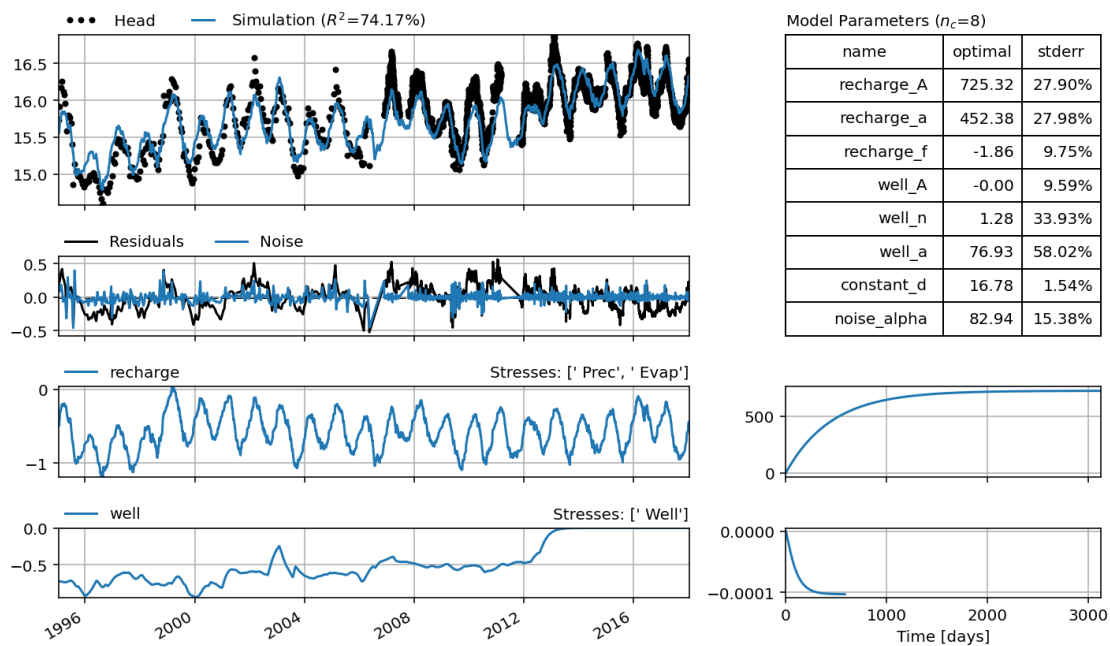


Nu het gemiddelde wordt genomen vóór dat er neerslagdata aanwezig is, is er geen trend in de residuen meer aanwezig. De modelfit is ook verbeterd, en de respons van de neerslag ziet er ook significant anders uit. De gevonden responstijd van de neerslag is veel groter dan 100 dagen. Daarom is het van belang neerslag voldoende neerslaggegevens te hebben of een realistische schatting, indien deze niet aanwezig zijn.

9.1.5 Eenheden afstemmen

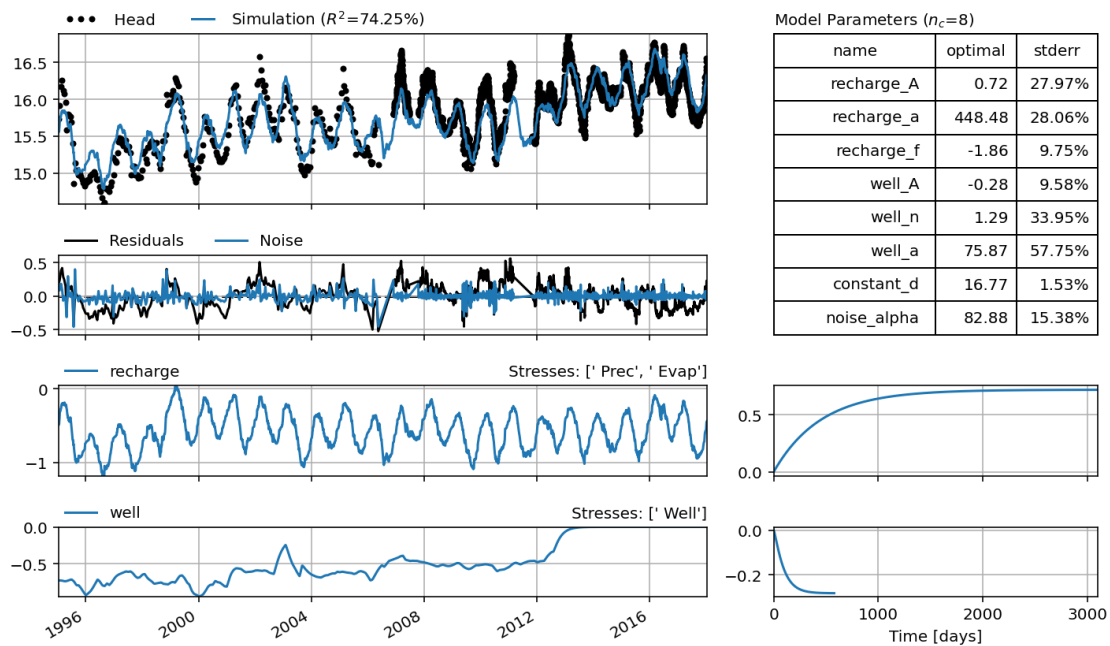
Voor de interpretatie van de resultaten (en soms ook voor het succes van de kalibratie) is het gewenst om de eenheden van de verschillende reeksen goed te kiezen en eventueel op elkaar af te stemmen. In onderstaand voorbeeld proberen we de invloed van een onttrekking vast te stellen. De invoer is aangeleverd op dagbasis in standaardeenheden. Neerslag en verdamping worden geleverd in meters per dag, en de onttrekking in kubieke meters per dag. We zijn geïnteresseerd in het totaal-effect van de onttrekking ter hoogte van de waarnemingsput. Eerst doen we een tijdreeksanalyse met de eenheden zoals we die zijn aangeleverd:

Afbeelding 9.11 Tijdreeksmodel met onttrekking



Het effect van de put (de 'gain') is gelijk aan de parameter well_A in de tabel. Deze heeft de eenheid van stijghoogte gedeeld door de eenheid van de onttrekking. In dit geval is dat dus een verlaging (in meter) bij een onttrekking van één kubieke meter per dag. Vanzelfsprekend is dat een heel klein getal, en in de tabel wordt dit zelfs afgerond naar 0.00. De gain van de recharge stelt de stijghoogte-verhoging voor die zou ontstaan als het continu één meter per dag zou regenen (737 meter). Evenzo is de verlaging door één meter verdamping per dag gelijk aan $1.86 \cdot 737 = 1370$ meter. Deze eenheden zijn niet erg praktisch voor de interpretatie. Daarom doen we deze tijdreeksanalyse nog eens opnieuw, maar dan veranderen we de eenheden van verdamping en neerslag naar mm/dag, en de onttrekking in miljoen kubieke meters per jaar.

Afbeelding 9.12 Tijdreeksmodel met geschaalde onttrekking



Bovenstaande afbeelding toont de resultaten met de gewijzigde eenheden. Wat opvalt is dat ook na omrekening de getallen iets verschillen. Dit zijn afwijkingen die ontstaan bij de optimalisatie van de parameters, die nu iets anders verloopt omdat de verhoudingen tussen de invoerreeksen verschillen. De gain bij één millimeter neerslag per dag is nu 0.72 meter. De verlaging ten gevolge van een onttrekking van één miljoen kuub per jaar is 0.28 meter. Om de eindverhoging bij stopzetting van de winning te schatten moet deze waarde vermenigvuldigd worden met de werkelijke wincapaciteit. (in miljoen kuub per jaar). Wanneer het een winning van 2 miljoen kuub per jaar betreft, schatten we de verhoging op 56 centimeter.

9.2 Model structuur

Een belangrijke stap in tijdreeksanalyse is het bepalen van de modelstructuur van het tijdreeksmodel. Dit is het onderwerp van deze paragraaf. We beperken ons hier tot het bepalen van het deterministische deel van het model. Het bepalen van het ruismodel is onderdeel van paragraaf 9.3. In deze paragraaf wordt een overzicht gegeven van de verschillende opties die mogelijk zijn voor de modelstructuur.

9.2.1 Algemeen stappenplan

Voor het kiezen van de gewenste modelstructuur is het belangrijk om vast te stellen wat het doel van de tijdreeksmodellering is. Wanneer het doel van de modellering is vastgesteld, kunnen de volgende stappen worden doorlopen:

- 1 **systemanalyse:** bepaal mogelijke hydrologische variabelen;
- 2 **modelbouw:** bepaal hoe deze variabelen worden omgezet in fluctuaties;
- 3 **controle:** bepaal welke variabelen daadwerkelijk invloed hebben.

We gaan ervan uit dat de eerste stap al is uitgevoerd en bespreken de tweede stap. In deze paragraaf wordt een overzicht gegeven van de bouwstenen die beschikbaar zijn om een model te bouwen.

9.2.2 Basis modelstructuur

Het tijdreeksmodel kan in zijn meest algemene vorm als volgt worden beschreven:

$$h(t) = \sum_{m=1}^M h_m(t) + d + r(t)$$

waar $h(t)$ de gemeten stijghoogte is, $h_m(t)$ de bijdrage van invloed m , d het basisniveau van het model, en $r(t)$ zijn de modelresiduën. Het totaal aantal invloeden dat bijdraagt aan de stijghoogte fluctuaties is M . Deze paragraaf gaat voornamelijk over het bepalen van deze bijdragen, maar eerst volgt een korte uitleg over het basisniveau d .

Het basisniveau van het model

Vrijwel elk tijdreeksmodel bevat een constante (hier parameter d) waarmee het model op de juiste hoogte wordt gebracht. Hoewel verleidelijk, is het niet altijd juist een fysische betekenis aan deze constante toe te kennen. In de literatuur wordt soms verwezen naar de drainagebasis, het niveau waar het grondwater na verloop van tijd naar toe gaat wanneer alle invloeden nul zijn. Of deze interpretatie juist is hangt echter af van het model. Veiliger is daarom om naar d te verwijzen als het basisniveau van het model.

Voor het bepalen van de waarde van het basisniveau zijn verschillende methoden beschikbaar. Een eerste optie is om na een simulatie het gemiddelde van de residuën te nemen als waarde voor d . Dit betekent dat de parameter d niet geschat hoeft te worden tijdens de kalibratie. Een andere optie is om d mee te schatten als een parameter (Collenteur et al. 2019). Dit heeft als voordeel dat er ook iets gezegd kan worden over de onzekerheid van deze parameter. Het is nog onduidelijk wat het effect van deze twee methoden is op de uiteindelijke resultaten.

Voor de voorbeelden in deze paragraaf maakt het wel of niet meefitten van de constante vrijwel niet uit voor de resultaten. In de rest van deze paragraaf wordt de waarde van het basisniveau d geschat tijdens kalibratie.

Bijdragen

Fluctuaties in de stijghoogte kunnen door verschillende oorzaken ontstaan. Vaak hebben we een idee welke variabelen belangrijk zijn, maar net zo vaak is juist het achterhalen van deze oorzaken het doel van de modellering. Het berekende effect van een bepaalde invloed zal in zekere mate afhangen van de gekozen vorm van de modellering, of, in andere woorden, de manier waarop een invloed wordt vertaald in een bijdrage aan de stijghoogtefluctuaties. De keuze voor een bepaalde modelstructuur kan dus van invloed zijn op de modelresultaten (zie ook hoofdstuk 5).

We kunnen grofweg twee typen invloeden onderscheiden: gemeten en niet gemeten invloeden. Voor gemeten invloeden zijn verklarende tijdreeksen beschikbaar om de bijdrage van een invloed aan de stijghoogte fluctuaties te bepalen. Dit zijn bijvoorbeeld: neerslag, potentiële verdamping, onttrekkingen, oppervlakte waterstanden, en luchtdruk metingen. Dit type invloeden kan worden vertaald in een bijdrage aan de stijghoogtefluctuaties met behulp van een impuls responsfunctie.

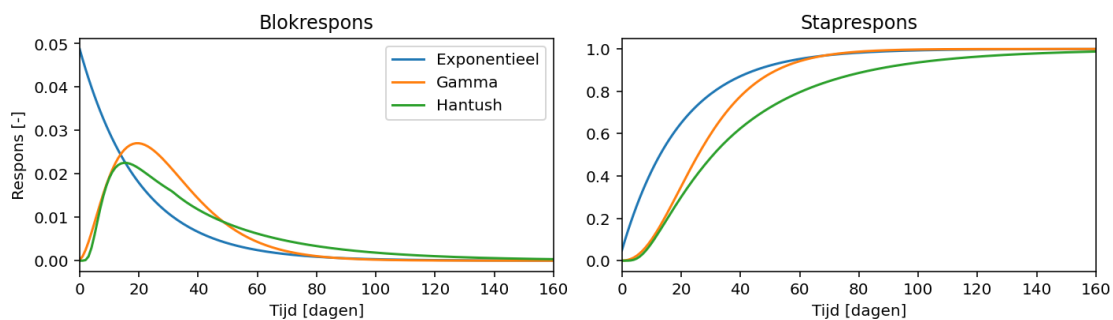
Voor niet gemeten invloeden zijn (vaak) geen verklarende tijdreeksen beschikbaar en het effect van dit type invloeden is lastiger vast te stellen. Voorbeelden zijn veranderend landgebruik en een ingreep in de waterhuishouding (bijv. het dempen van een sloot of het uitbaggeren van een kanaal). Sommige van deze bijdragen kunnen worden meegenomen door een trend mee te nemen in de modellering. De informatie die de modelleur moet opgeven is dan een tijdsperiode of tijdstip waarop een verandering heeft plaatsgevonden.

9.2.3 Responsfuncties

De vertaling van een verklarende tijdreeks naar een bijdrage aan de stijghoogtefluctuaties kan worden gedaan met behulp van een impuls responsfunctie (zie ook von Asmuth et al., 2002). De impuls responsfunctie beschrijft in dit geval hoe de stijghoogte reageert op een impuls in de verklarende variabele. Er zijn tal van responsfuncties die kunnen worden gekozen om een verklarende tijdreeks om te zetten in een bijdrage aan de stijghoogtefluctuaties (zie paragraaf 5.2.3). De keuze voor een specifieke responsfunctie kan bijvoorbeeld afhangen van het type verklarende reeks (bv. neerslag of de waterstand in een rivier), maar ook van het systeem (bijv. een dikke of dunne onverzadigde zone).

Vaak wordt een algemene responsfunctie met 4 parameters of een vereenvoudigde vorm daarvan gebruikt als impuls responsfunctie. Afhankelijk van het systeem kunnen bepaalde parameters worden vastgezet, wat het aantal te schatten parameters verlaagt. In de volgende figuur worden drie impuls responsfuncties getoond die elk een vereenvoudigde vorm van de 4-parameter functie zijn: Exponentieel, Gamma, en Hantush.

Afbeelding 9.13 Voorbeelden verschillende responsfuncties



De Exponentieel respons heeft 2 parameters, en de Gamma en de Hantush respons hebben elk 3 parameters. Zoals blijkt uit de figuur hierboven helpt de extra parameter van de Gamma en de Hantush om een grotere vertraging in de response van de stijghoogte te simuleren. Het kan voorkomen dat de parameters van de Gamma responsfunctie na kalibratie eigenlijk een Exponentieel responsfunctie simuleren. Het is dan mogelijk parameters vast te zetten of de responsfunctie te veranderen en op die manier het aantal te schatten parameters te reduceren. Over het algemeen heeft het de voorkeur de responsfunctie met minder parameters te kiezen bij een gelijke modelfit.

Hoewel de naam 'Hantush' suggereert dat de parameters fysisch geïnterpreteerd kunnen worden, hangt de validiteit van een dergelijke interpretatie af van de lokale bodemopbouw in relatie tot de aannames van de Hantush functie. Verder dient het aanbeveling verschillende responsfuncties te testen. Dit zal in deze paragraaf dan ook meermaals worden gedaan. Per software pakket dient te worden bepaald welke responsfuncties beschikbaar zijn.

9.2.4 Effect van neerslag en verdamping

Vrijwel alle grondwaterstanden in Nederland worden in meer of mindere mate beïnvloed door neerslag en verdamping. Deze twee verklarende tijdsreeksen worden daarom bijna altijd in een tijdreeksmodel opgenomen. De manier waarop deze tijdsreeksen worden meegenomen in het model kunnen sterk verschillen. We kunnen grofweg vier methoden onderscheiden (zie hiervoor ook paragraaf 3.4.4 en 5.2.2):

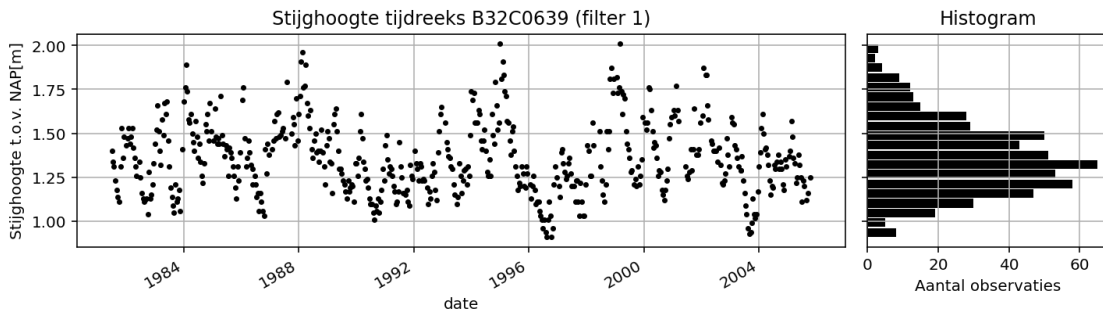
- neerslag en verdamping als aparte invloeden met aparte responsefuncties;
- netto neerslagoverschot ($N = P - fE$) en 1 responsfunctie;
- niet-lineaire grondwateraanvulling en 1 responsfunctie;
- direct opgeven van de grondwateraanvulling en 1 responsfunctie.

In de volgende drie secties worden drie modellen gemaakt volgens de eerste drie methoden voor het modelleren van het effect van neerslag en verdamping. De laatste methode wordt niet getoond maar kan eenvoudig worden toegepast. Bij deze methode wordt buiten de tijdreeksmodellering om de grondwateraanvulling of neerslagoverschot berekend (bijv. met een SWAP of Hydrus-1D model), waarbij deze tijdreeks dan als invoer wordt gebruikt en vertaald met een responsfunctie.

Voorbeeld data

De stijghoogte tijdreeks die in dit voorbeeld wordt gebruikt is gemeten in peilbuis B32C0639 (filter 1) nabij de Bilt. Neerslag en potentiële Makkink verdamping zijn beschikbaar van het meteorologisch meetstation de Bilt van het KNMI. De afbeelding hieronder laat de gemeten stijghoogte tijdreeks zien (links) en een histogram van de observaties (rechts).

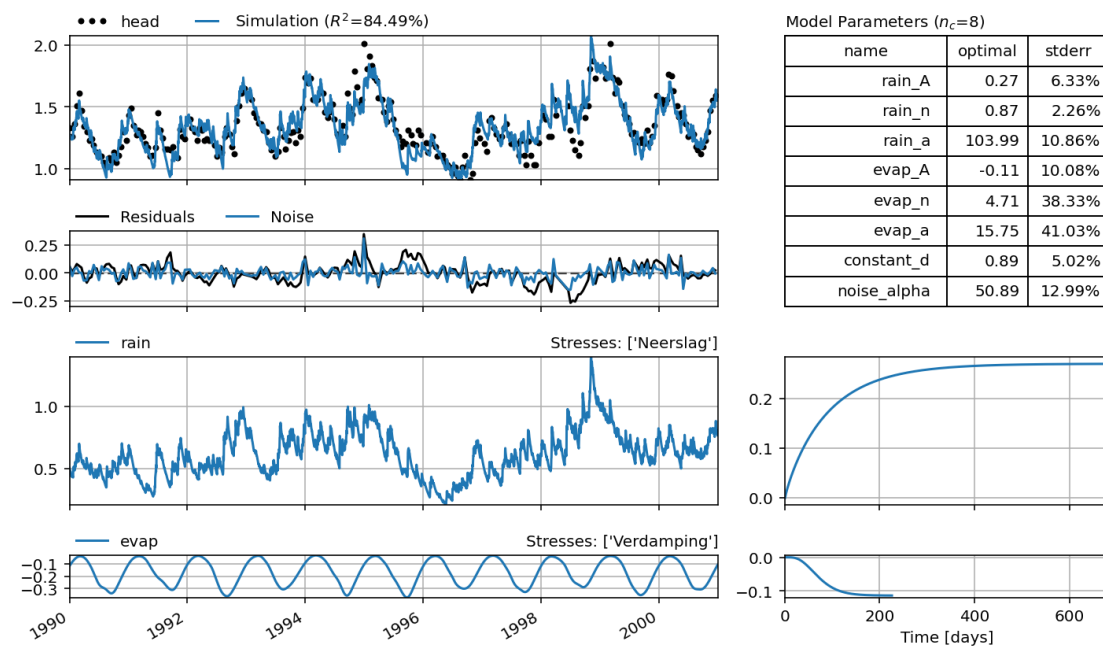
Afbeelding 9.14 Gemeten stijghoogte met histogram



Neerslag en verdamping als aparte invloeden

Bij deze methode worden de twee tijdreeksen van neerslag en verdamping onafhankelijk van elkaar meegenomen in het tijdreeksmodel. Door middel van twee aparte responsfuncties worden de tijdreeksen vertaald in bijdragen aan de stijghoogtefluctuaties. Hieronder is een voorbeeld gegeven van het resultaat van een dergelijk model voor de stijghoogte tijdreeks gemeten in peilbuis B32C0639 (filter 1).

Afbeelding 9.15 Resultaat Tijdreeksmodel B32C0639



De onderste twee linker subplots tonen de geschatte invloed van de neerslag en de verdamping. De neerslag heeft een positief (stijgend) effect op de stijghoogte en de verdamping een dalend effect. Voor beide responsfuncties (zie de onderste twee rechter subplots) moeten de parameters worden geschat. Het totaal aantal geschatte parameters van dit model is 8.

Enkele kanttekeningen over de fysische interpretatie van deze modelvorm zijn van toepassing. De responsen van de stijghoogte op neerslag en verdamping verschillen in dit voorbeeld sterk van elkaar. Vanuit een grondwater perspectief is het de vraag of het logisch is dat de responsen op neerslag en verdamping zo sterk van elkaar verschillen. Aan de andere kant is de responsfunctie ook verantwoordelijk voor de vertaling van de onverzadigde zone processen, en omdat het patroon van neerslag en verdamping anders zijn verdeeld in de tijd zou het ook zo kunnen zijn dat de respons anders is.

Lineair neerslagoverschot

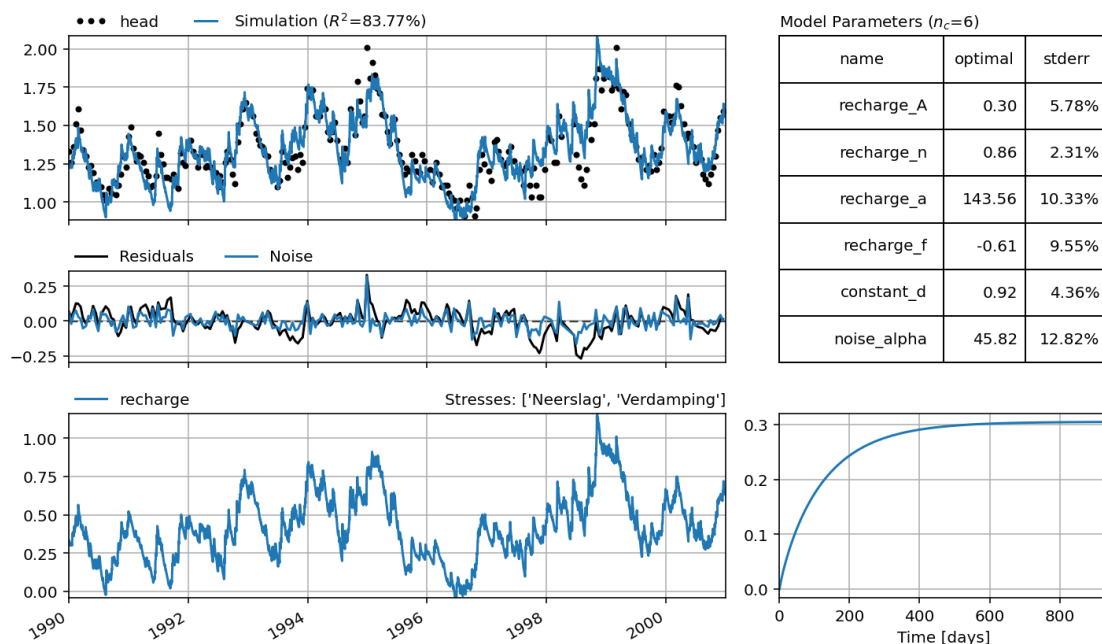
Een veelgemaakte aanname is dat de respons van de stijghoogte op neerslag en verdamping vergelijkbaar is, maar in de tegenovergestelde richting (als het regent gaat de stijghoogte omhoog, als het verdampt gaat de stijghoogte omlaag). Hoewel uit bijvoorbeeld bovenstaande figuur blijkt dit niet altijd het geval is, is in de praktijk gebleken dat deze aanname toch vaak goede modellen oplevert. Onder deze aanname is het mogelijk eerst het neerslagoverschot te berekenen en deze flux te vertalen in een bijdrage aan de stijghoogtefluctuaties met een enkele responsfunctie.

Omdat vaak gewerkt wordt met de potentiële of referentie verdamping (bijvoorbeeld Makkink of Penman-Monteith) wordt een verdampingsfactor f gebruikt. Het neerslag overschot N wordt dan berekend als (zie bv. Berendrecht et al., 2003):

$$N(t) = P(t) - f * E(t)$$

waar $P(t)$ de neerslag is, $E(t)$ de potentiële verdamping, en f een kalibratie parameter. Een belangrijk voordeel van dit model ten opzichte van het voorgaande model is het kleinere aantal parameters dat geschat moet worden.

Afbeelding 9.16 Resultaat tijdreeksmodel met verdampingsfactor

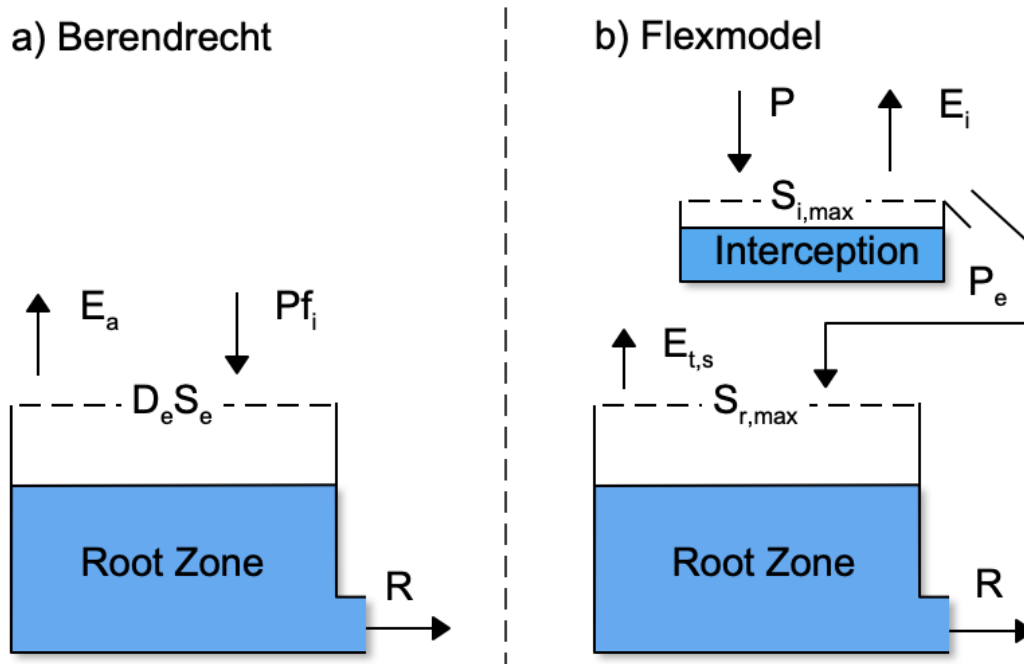


In de afbeelding hierboven is het resultaat van dit model voor de voorbeeld data getoond. Het totaal aantal geschatte parameters in dit model is 6. De geschatte verdampingsfactor is $f = -0.61$. De verdamping is in dit voorbeeld dus 61 % van de Makkink verdamping. Terwijl dit model twee kalibratie parameters minder heeft is de model fit (gemeten als R^2) vrijwel gelijk aan het eerste model.

Niet-lineaire grondwateraanvulling

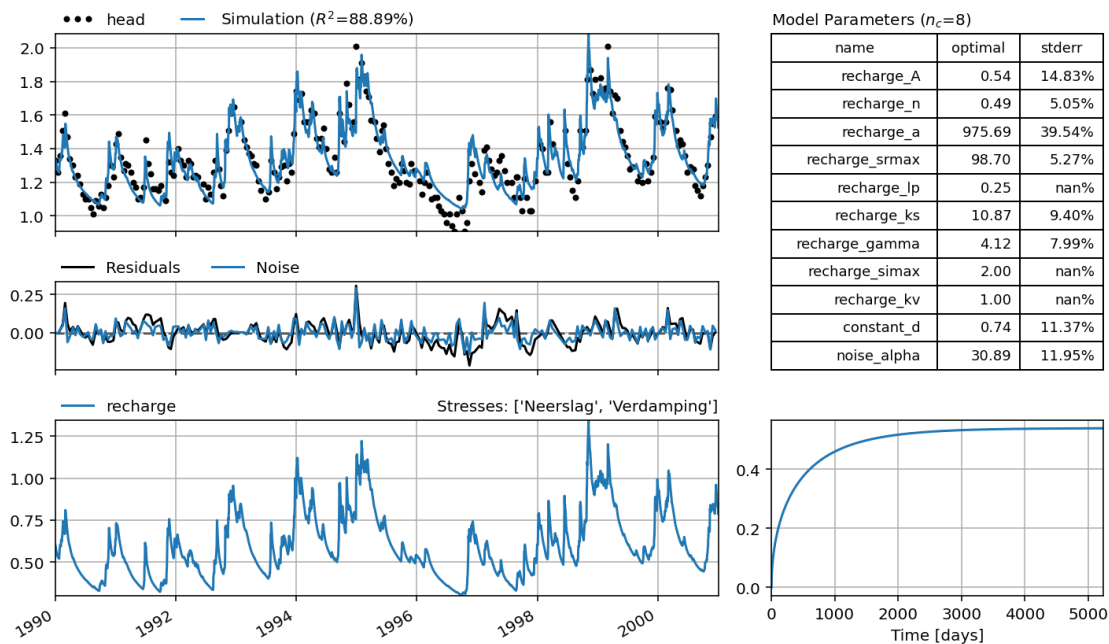
Bij de derde methode wordt een niet-lineair model gebruikt om de grondwateraanvulling te berekenen. Deze flux wordt vervolgens vertaald in een bijdrage aan de stijghoogtefluctuaties met behulp van een responsfunctie. Verschillende niet-lineaire modellen zijn beschikbaar, zie bijvoorbeeld Berendrecht et al. (2005), Peterson en Western (2014) en Collenteur et al. (2020). Al deze modellen hebben de vorm van een bakje, waar neerslag (P) en verdamping (E) in en uit gaat aan de bovenkant, en grondwateraanvulling (R) uitkomt aan de onderkant. Ter illustratie zijn de conceptuele modellen van Berendrecht et al. (2005) en Collenteur et al. (2020) hieronder visueel weergegeven. Voor de interpretatie van de verschillende modelparameters van de niet-lineaire modellen wordt verwezen naar bovengenoemde publicaties.

Afbeelding 9.17 Conceptueel model niet lineaire grondwateraanvulling



Een voordeel van niet-lineaire modellen is dat er rekening wordt gehouden met (niet-lineaire) processen in de wortelzone, waardoor de stijghoogte niet meer lineair reageert op neerslag en verdamping (dat wil zeggen, twee keer zo veel neerslag resulteert niet in twee keer zo grote stijghoogte toename). Dit gebeurt bijvoorbeeld wanneer de verdamping wordt beperkt door het beschikbare bodemvocht ten tijde van droogte, of door de berging van extreme neerslag in de wortelzone. In onderstaande afbeelding is het resultaat van dit model getoond. Het totaal aantal geschatte parameters van dit model is 8, omdat 3 parameters vast gezet zijn op basis van expert kennis.

Abbeelding 9.18 Resultaten tijdreeksmodel met niet-lineaire wortelzone



Een mogelijk nadeel van niet-lineaire modellen kan een groter aantal vrije parameters zijn. Het is echter vaak mogelijk een aantal parameters vast te zetten en een response functie te gebruiken met minder parameters, waardoor het aantal vrije parameters is te reduceren. De rekentijd van niet-lineaire modellen is tegenwoordig vergelijkbaar met de rekentijd van lineaire modellen.

Welk model is het beste?

We hebben nu drie modellen voor een stijghoogte tijdreeks die goed verklaard kan worden door neerslag en potentiële verdamping. Het model met twee aparte responsen heeft een $R^2 = 84.5\%$, het lineaire model $R^2 = 83.8\%$ en het niet-lineaire model $R^2 = 89.3\%$, elk voor de periode 1990-2001. De modellen ontlopen elkaar dus eigenlijk niet zo veel wat betreft deze fit statistiek. Het is nu aan de modelleur om een van deze modellen te kiezen of deze allemaal te gebruiken.

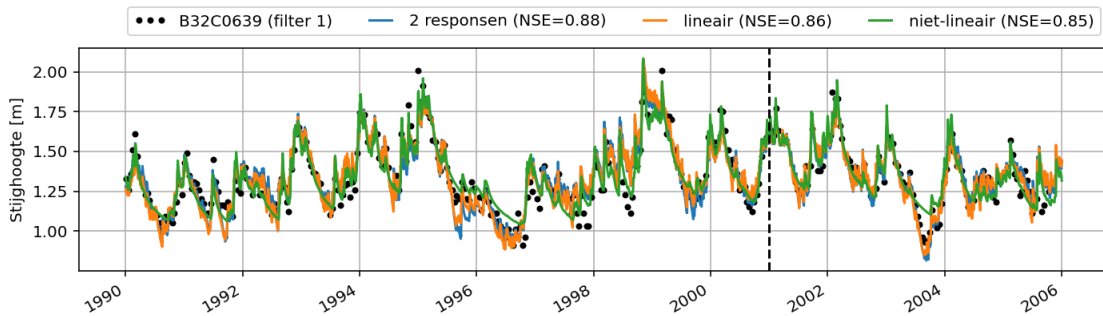
Een veelgebruikte methode om het 'beste' model te selecteren is door de fit van de modellen buiten de kalibratieperiode te berekenen (zie ook de paragraaf over 'Beoordeling', paragraaf 9.4). In de afbeelding hieronder is met elk model de stijghoogte voor een periode na de kalibratieperiode gesimuleerd (2001-2006), de zogeheten validatieperiode. In de legenda is de Nash-Sutcliffe statistiek weergegeven voor elk model. Nu is het model met de twee responsen het 'beste' model volgens de NSE statistiek (NSE=0.88), en hebben het lineaire en niet-lineaire model een iets mindere fit (NSE=0.85).

De model fits zijn in dit voorbeeld zeer vergelijkbaar. De modelleur kan er in dit geval bijvoorbeeld voor kiezen het model met het laagste aantal parameters te kiezen. Dit aantal is het laagste voor het lineaire model ($n = 6$); de andere twee modellen hebben beide meer parameters (beide $n = 8$). Ook de de AIC statistiek geeft hier een duidelijke voorkeur voor het lineaire model (AIC=10.75). Een ander type motivatie zou kunnen afhangen van het doel van de modellering. Wanneer we bijvoorbeeld vooral geïnteresseerd zijn in lage grondwaterstanden lijken de lineaire modellen beter toepasbaar, terwijl hoge grondwaterstanden juist beter met het niet-lineaire model worden gesimuleerd. Hier is duidelijk dat de keuze voor een specifiek model afhangt van het doel van de modellering.

Tabel 9.1 AIC verschillende modellen

Model	AIC
2 responsen	-36.23
lineair	-18.18
niet-lineair	-49.63

Afbeelding 9.19 Vergelijking resultaten verschillende tijdreeksmodellen



9.2.5 Dikkere onverzadigde zone

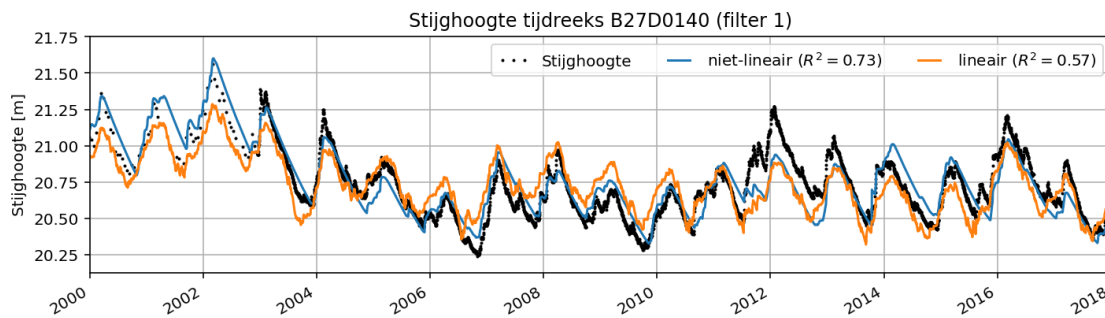
Uit het voorgaande voorbeeld bleek dat een lineair model voor deze peilbuis en stijghoogte tijdreeks beter in staat is de lage stijghoogtes te simuleren. De grondwaterspiegel in dit voorbeeld ligt dichtbij het maaiveld, wat betekent dat de wortelzone (of eigenlijk de vegetatie) water op kan nemen uit het grondwater. Het lineaire model kan een negatieve flux voor het neerslagoverschot simuleren, terwijl het niet-lineaire model dat in het voorbeeld hierboven is gebruikt dit niet kan. Peterson et al (2014) hebben een niet-lineair model uitgewerkt voor diepe grondwaterstanden en diepe wortelsystemen in Australië, waarbij ook water uit het grondwater wordt onttrokken en kan verdampen.

In Nederland zullen dergelijke diepe wortelsystemen echter minder vaak voorkomen, en beweegt bodemwater in systemen met dikke onverzadigde zones van de wortelzone naar het grondwater. Wanneer er geen verdamping vanuit het grondwater plaatsvindt kan een niet-lineair model helpen om de stijghoogte te modelleren. Of het toepassen van een niet-lineair model de simulatie van een stijghoogte tijdreeks verbetert, is onderdeel van onderzoek en zal per peilbuis moeten worden vastgesteld. Ook hier is de aanbeveling om deze modelvorm eenvoudigweg te testen en te variëren zoals bij de responsfuncties.

Voorbeeld peilbuis B27D0140

In onderstaand voorbeeld is een stijghoogte tijdreeks gemodelleerd met een dikkere onverzadigde zone. De tijdreeks is gemeten in peilbuis B27D0140 (filter 1) nabij Oldebroek ten noorden van de Veluwe. In dit voorbeeld is de fit met de metingen beter door het gebruik van een niet-lineair model voor de berekening van de grondwateraanvulling.

Afbeelding 9.20 Vergelijking tijdreeksmodellen B27D0140



9.2.6 Drempel niet-lineariteit

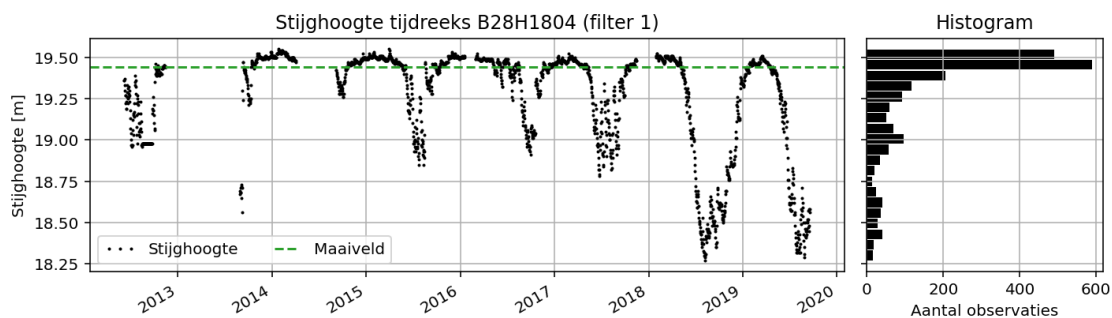
Een veelvoorkomend fenomeen in Nederland is drempel niet-lineariteit (zie bv. Berendrecht et al. (2004) en von Asmuth et al. (2012)). Daarbij reageert een grondwatersysteem anders op neerslag nadat een bepaalde grenswaarde in de stijghoogte is overschreden. Onder andere Knotters et al. (1999) hebben hiervoor een model ontwikkeld, het zogeheten TARSO model, waarbij de respons op neerslagoverschot afhangt van de stijghoogte boven of onder een bepaalde drempelwaarde. Wanneer de stijghoogte boven deze drempelwaarde komt, kan het bijvoorbeeld zo zijn dat er afstroming naar het oppervlaktewater plaatsvindt (bv. een sloot die anders droog staat). Of er sprake is van drempel niet-lineariteit kan worden vastgesteld door:

- een visuele inspectie van de stijghoogte tijdreeks;
- het plotten van een histogram van de stijghoogte metingen;
- het testen van verschillende modelstructuren.

Voorbeeld peilbuis B28H1804

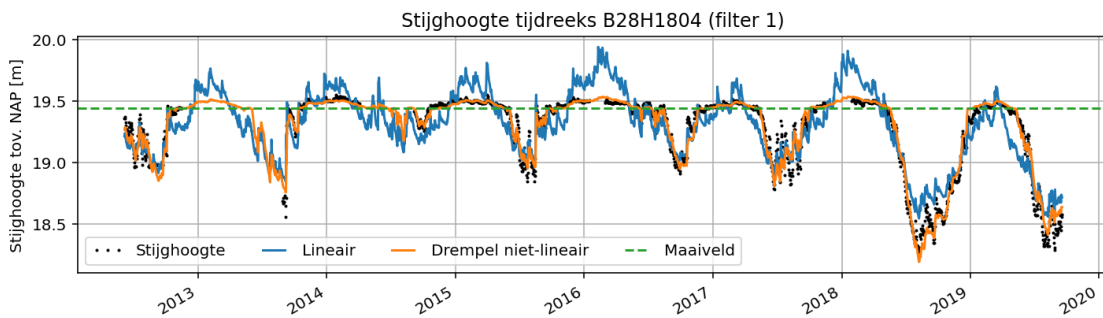
Als voorbeeld kijken we naar de stijghoogte tijdreeks gemeten in peilbuis B28H1804 (filter 1) in de buurt van Weerselo. De gegevens in Dinoloket laten zien dat het maaiveld bij dit punt op 19.44 meter t.o.v. NAP ligt. In onderstaande afbeelding is de stijghoogte tijdreeks en een histogram van de reeks afgebeeld voor een visuele inspectie (stappen 1 & 2).

Afbeelding 9.21 Stijghoogte reeks B28H1804



Uit bovenstaande plots is af te leiden dat er mogelijk sprake is van drempel niet-lineariteit. De stijghoogte reeks is afgevlakt rond 19.50-19.75 en het histogram laat duidelijk een scheve verdeling zien. Op basis van deze visuele inspectie kunnen we besluiten om twee modellen te proberen: de eerste met een simpel neerslagoverschot ($N = P - fE$) en een tweede met drempel niet-lineariteit (hier, het TARSO model). Vervolgens kan dan worden beoordeeld welk model beter is.

Afbeelding 9.22 Resultaat tijdreeksmodellen B28H1804



Bovenstaande afbeelding geeft de resultaten van de twee modellen weer. Het lineaire model geeft een structurele overschatting van de hoge stijghoogtes, waar het model met drempel niet-lineariteit netjes de gemeten grondwaterstand volgt. Het geobserveerde hydrologische gedrag wordt hier beter gesimuleerd met het niet-lineaire drempel model. Het is in het TARSO model ook mogelijk de afvoertermen weer te geven.

Afbeelding 9.23 Afvoertermen TARSO model



9.2.7 Andere verklarende variabelen

Tot op dit punt hebben we gekeken naar stijghoogte tijdreeksen die zich goed laten verklaren door neerslag en verdamping. In veel gevallen zijn echter andere verklarende tijdreeksen nodig om de stijghoogte goed te kunnen modelleren. Twee invloeden die vaak belangrijk zijn, zijn oppervlaktewaterpeilen en grondwateronttrekkingen. In de volgende twee secties wordt beschreven hoe de effecten van deze invloeden kunnen worden gemodelleerd.

Oppervlaktepeil als verklarende tijdreeks

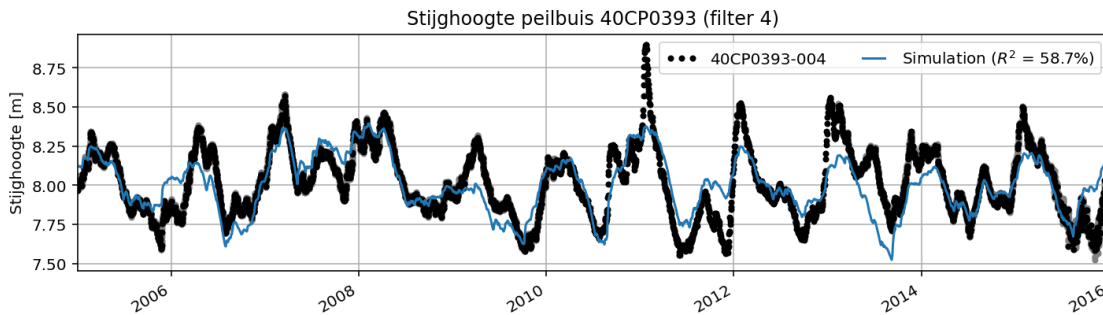
In grote delen van Nederland worden de stijghoogten beïnvloed door oppervlaktewaterpeilen. Dit kunnen grote rivieren en meren als de Rijn en het IJsselmeer zijn, maar ook kleinere vennen en slootjes in de polder. Het is niet ongebruikelijk dat het nodig is een tijdreeks van een nabijgelegen oppervlaktewaterpeil toe te voegen aan het tijdreeksmodel om de stijghoogte goed te kunnen modelleren.

We kunnen dit wederom doen door de tijdreeks van het gemeten oppervlaktewaterpeil te vertalen naar een bijdrage aan de stijghoogtefluctuaties met behulp van een responsfunctie. De responsfunctie beschrijft in dit geval de reactie van de stijghoogte op fluctuaties in het oppervlaktewaterpeil.

Voorbeeld peilbuis 40CP0393

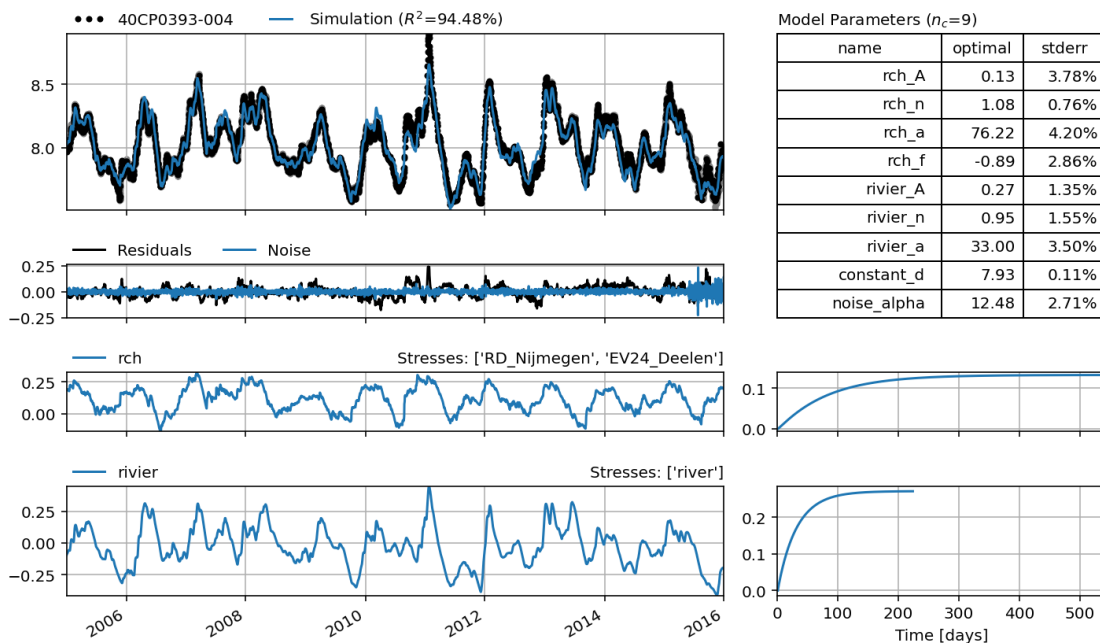
Als voorbeeld modelleren we de stijghoogte tijdreeks gemeten in peilbuis 40CP0393 (filter 4). Eerst maken we een model met enkel neerslag en verdamping als verklarende variabelen. Deze worden vertaald in stijghoogte fluctuaties met een lineair neerslagoverschot model en een Gamma responsfunctie. Dit model levert de volgende simulatie van de stijghoogte op.

Afbeelding 9.24 Resultaat simulatie met tijdreeksmodel 40CP0393



Uit bovenstaande afbeelding blijkt duidelijk dat het neerslagoverschot niet genoeg is om de stijghoogte fluctuaties te verklaren. Uit de hydrologische systeem analyse van het studiegebied is gebleken dat de peilbuis dichtbij een rivier ligt. We voegen daarom de rivierwaterstand als verklarende tijdreeks aan het model toe en kijken of dit de simulatie van de stijghoogte verbetert. We kiezen hier voor de Gamma responsfunctie met drie parameters om het oppervlaktewaterpeil te vertalen in stijghoogtefluctuaties. Deze respons functie is vaak een goede start om oppervlaktewaterpeilen te vertalen.

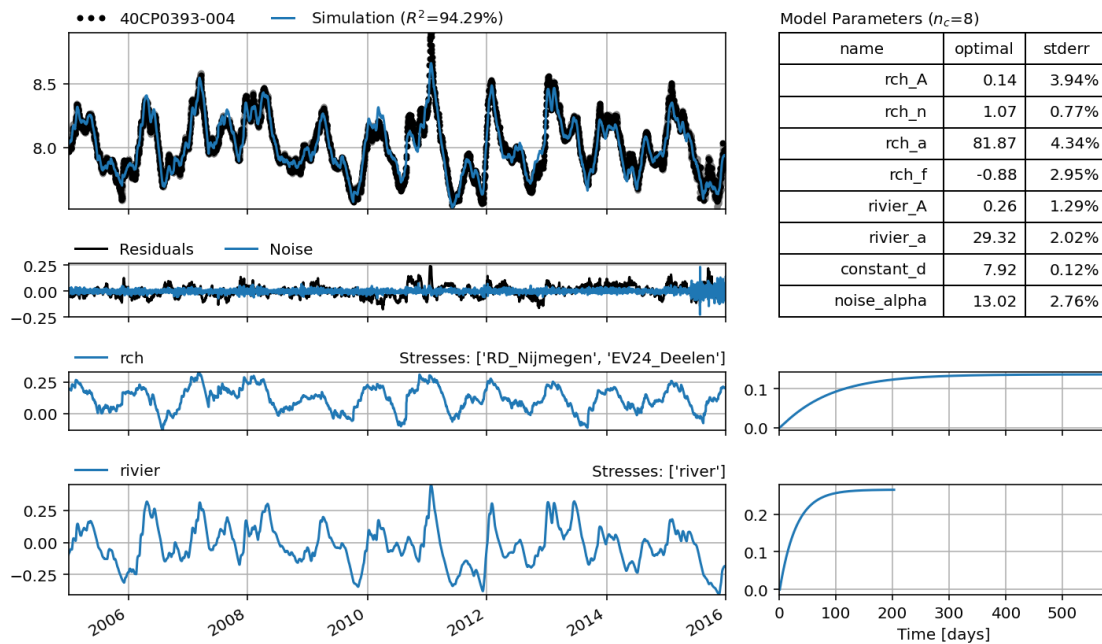
Afbeelding 9.25 Resultaat tijdreeksmodel met oppervlaktewater



Het model laat een duidelijke verbetering zien in de fit tussen de gesimuleerde en de gemeten stijghoogte. De fit, hier uitgedrukt als R^2 , is gestegen van $R^2 = 59\%$ naar $R^2 = 94\%$. Het toevoegen van het oppervlaktewaterpeil levert hier dus een duidelijke verbetering op.

De gekalibreerde responsfunctie voor het effect van de rivier lijkt erg op een exponentiële responsfunctie, te zien aan de blok respons en de waarde van parameter rivier_n die bijna 1 is. Dit betekent dat de stijghoogte vrijwel instantaan reageert op de rivierwaterstand. Wanneer dit het geval is, verdient het aanbeveling de Gamma responsfunctie te vervangen door de exponentiële responsfunctie. Dit scheelt 1 kalibratie parameter in het model, en levert hetzelfde model op. Wanneer de fit gelijk is, heeft een model met minder parameters de voorkeur over een model met meer parameters. Onderstaande figuur toont het model waar een exponentiële responsfunctie gebruikt is in plaats van de Gamma functie. Het resultaat laat zien dat dit een vergelijkbaar model oplevert, maar met 1 parameter minder.

Afbeelding 9.26 Resultaat tijdreeksmodel



Zowel de grondwateraanvulling als het oppervlaktewaterpeil is afhankelijk van de neerslag (als het hard regent is vaak het rivierpeil ook hoog). Dit kan er voor zorgen dat de parameters van deze modelonderdelen sterk gecorreleerd zijn en de bijdragen aan de stijghoogte fluctuaties lastig van elkaar te onderscheiden zijn. In het voorgaande model is dit visueel al zichtbaar; de correlatie tussen de twee bijdragen is $r = 0.44$. Des te sterker de correlatie tussen deze twee bijdragen, des te lastiger is het om een deel van de stijghoogtefluctuatie toe te kennen aan 1 van deze invloeden. Dit kan de mogelijkheid beperken om precieze uitspraken te doen over het effect van een bepaalde invloed op de stijghoogte.

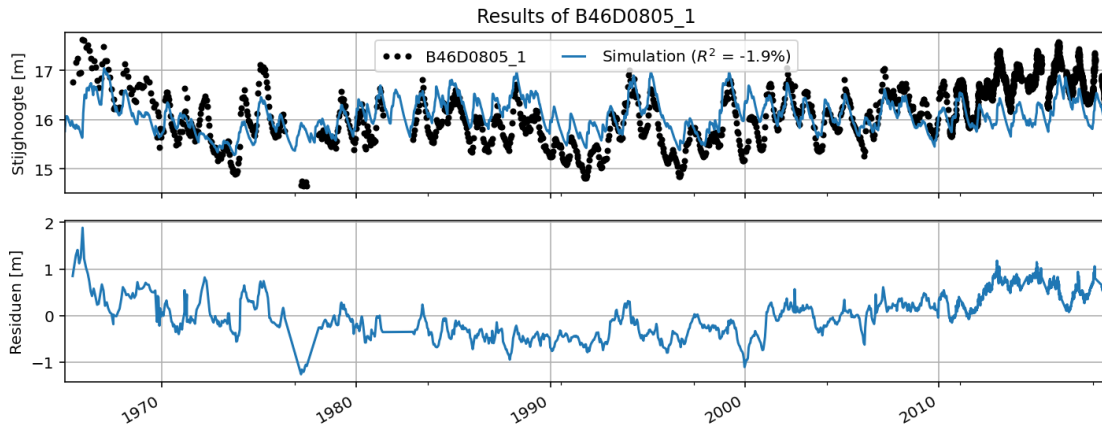
De invloed van een grondwateronttrekking bepalen

Grondwater wordt in vrijwel heel Nederland gebruikt voor de productie van drinkwater en als bron van irrigatiewater. Voor het modelleren van stijghoogte tijdreeksen in gebieden waar (in het verleden of nu) grondwater wordt onttrokken is het vaak nodig om een tijdreeks van deze onttrekking mee te nemen. Enige voorbewerking van deze tijdreeksen is over het algemeen nodig, bijvoorbeeld om tijdreeksen met maandelijkse debietstotalen terug te brengen naar een tijdreeks met een gelijke tijdstap (bv. dagsommen). Ook kan het helpen de debietreeksen uit te drukken in Miljoen kuub per tijdstap, om te voorkomen dat extreem kleine parameter waarden worden gesimuleerd. Meer over deze voorbewerking en kalibratie is te vinden in de paragrafen 9.1 en 9.3.

Voorbeeld peilbuis B46D0805 met neerslag en verdamping

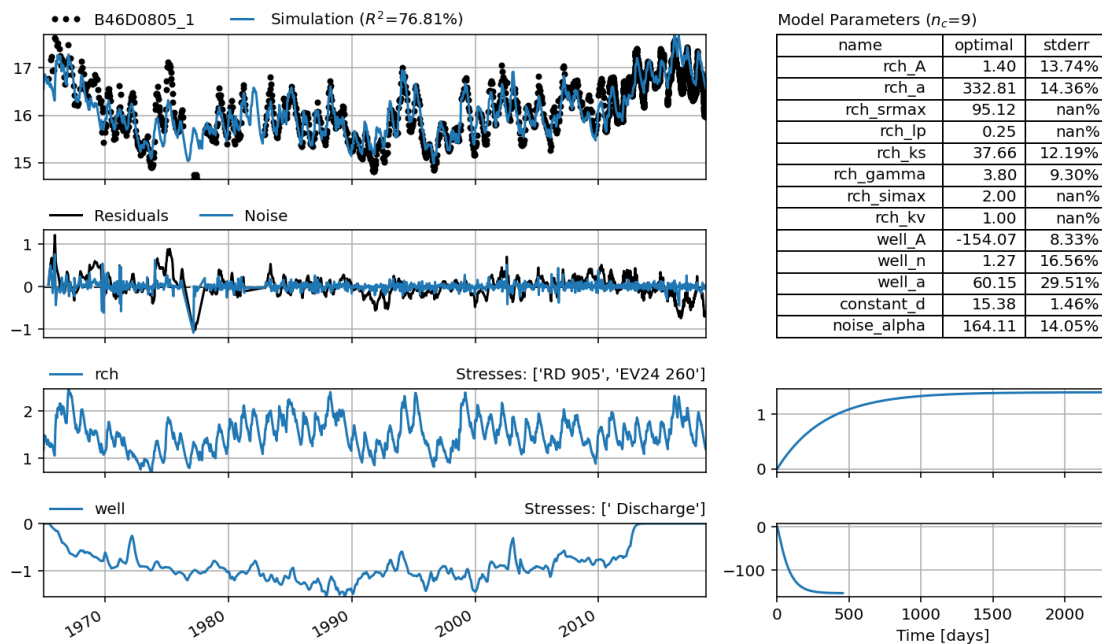
Als voorbeeld kijken we naar een peilbuis B46D0805 (filter 1) in de omgeving van een grondwater pompstation in Vierlingsbeek. Op basis van de hydrologische systeem analyse is te verwachten dat neerslag, verdamping, en mogelijk de grondwateronttrekking als verklarende variabelen in het model moeten worden meegenomen. We kiezen er hier voor met een simpel model te beginnen op basis van neerslag en verdamping. Wanneer dit genoeg is om de stijghoogte tijdreeks te modelleren zijn we klaar. Wanneer dit niet het geval is, voegen we de debietreeks van het pompstation toe als verklarende variabele.

Afbeelding 9.27 Resultaat tijdreeksmodel B46D0805_1



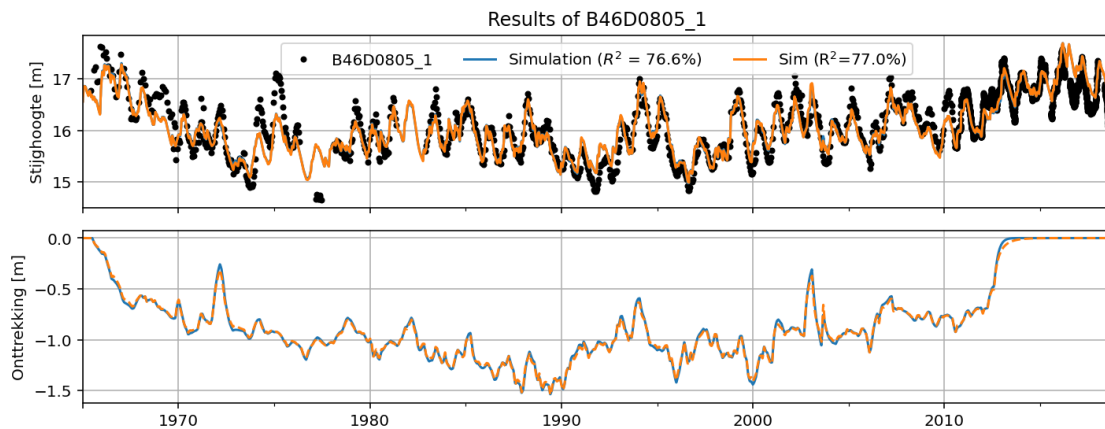
Bovenstaande afbeelding toont de gemeten en gesimuleerde stijghoogtetijdreeks in de bovenste plot, en de modelresiduen in de onderste plot. In de modelresiduen is een duidelijke trend over de tijd te zien, wat wijst op het mogelijk ontbreken van een invloed. We voegen daarom de tijdreeks van het onttrekkingsdebiet toe aan het model. Deze wordt vertaald naar een bijdrage aan de stijghoogtefluctuaties met behulp van een Gamma response functie. Dit levert het volgende model op.

Afbeelding 9.28 Resultaat tijdreeksmodel B46D0805_1 met onttrekking



Bovenstaande afbeelding laat duidelijk zien dat de fit van het model wordt verbeterd door het toevoegen de grondwateronttrekking aan het model. Dit geeft ook de mogelijkheid om de invloed van het onttrekken van grondwater op de stijghoogte in te schatten. De stap respons kan een indicatie geven van de te verwachten verlaging (verhoging) door het extra (minder) onttrekken van een hoeveelheid grondwater. In dit model is gebruik gemaakt van een Gamma responsfunctie met drie parameters. Een andere responsfunctie die gebruikt kan worden is bijvoorbeeld de Hantush functie. De volgende afbeelding toont het modelresultaat wanneer een Hantush responsfunctie wordt gebruikt voor de onttrekking.

Afbeelding 9.29 Resultaat rijdreksmodel met verschillende responsfuncties



In dit voorbeeld maakt het weinig verschil welke van deze twee responsfuncties wordt gebruikt om het effect van de onttrekking op de stijghoogte te modelleren.

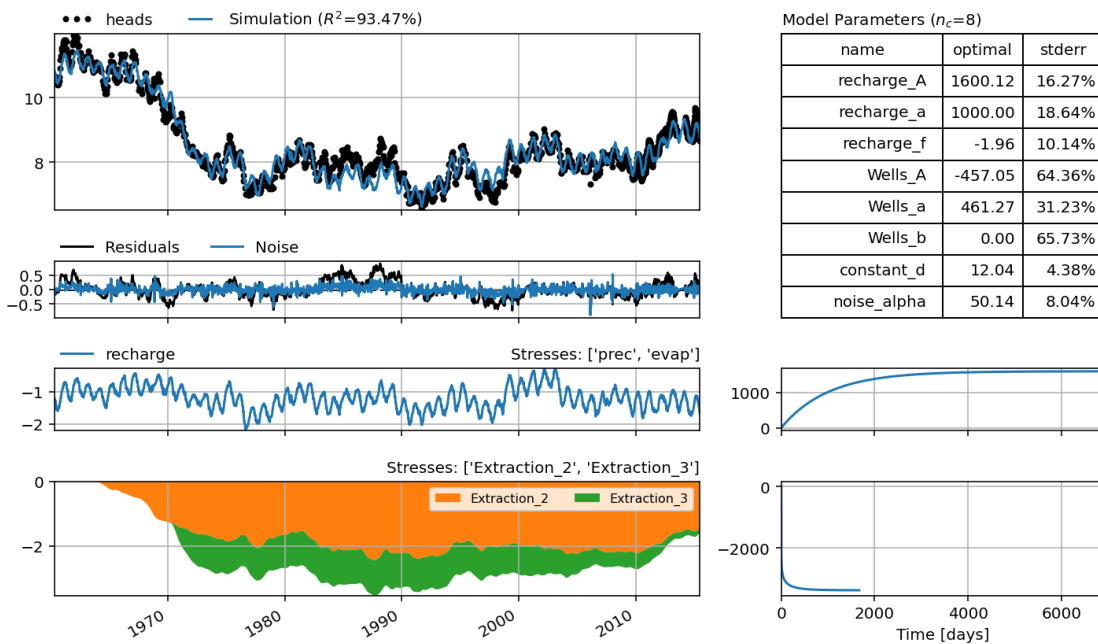
Tot slot een opmerking over het modelleren van onttrekkingen en de variatie in de debietsreeks. Om de invloed van het onttrekken van grondwater op de stijghoogte vast te stellen is het belangrijk dat er enige variatie aanwezig is in de debiet tijdreeks. Er bestaat helaas (nog) geen vuistregel om te bepalen hoeveel variatie in een verklarende tijdreeks genoeg is om de invloed goed vast te kunnen stellen. Of de debietsreeks een periode zonder onttrekken moet bevatten voor het vaststellen van de stijging in de stijghoogtes als gevolg van het stoppen van pompen is bijvoorbeeld onbekend. Ook kan het voorkomen dat de debietsreeksen een vergelijkbaar patroon vertonen als andere tijdreeksen (bv. het neerslagoverschot en verhoogde drinkwatervraag in de zomer), waardoor het lastig is deze invloeden van elkaar te onderscheiden. Zie voor voorbeelden hiervan bijvoorbeeld Maas, 2011 en Baggelaar et al., 2018.

Meerdere grondwateronttrekkingen

Wanneer meerdere pompstations in een gebied aanwezig zijn is het goed mogelijk dat de debietsreeksen sterk aan elkaar gecorreleerd zijn. Het is dan lastig om de invloed van deze reeksen van elkaar te onderscheiden. In dit geval is het mogelijk om de onttrekkingen te vertalen in stijghoogtefluctuaties met een enkele responsfunctie die geschaald is met de afstand van de pompput tot de peilbuis. Een dergelijk model en toepassing is bijvoorbeeld ook beschreven in Berendrecht et al. (2006). Dit scheelt een aantal parameters, terwijl vaak de fit van het model zeer vergelijkbaar is. Wanneer met meerdere onttrekkingen wordt gewerkt die water oppompen uit hetzelfde watervoerende pakket is het dus aan te bevelen een dergelijke model structuur te gebruiken en zo het aantal (te schatten) parameters laag te houden. In het voorbeeld hieronder is dit uitgewerkt.

De peilbuis ligt op 2282 meter afstand van pompput 1 en op 2784 meter afstand van pompput 2. De onttrekkingsdebieten zijn in enige mate gecorreleerd ($r=0.65$). Beide onttrekkingen vinden plaats in hetzelfde watervoerende pakket. De responsen van de stijghoogte op beide onttrekkingen zouden dus aan elkaar gerelateerd moeten zijn. Hieronder is het hierboven beschreven model toegepast op dit voorbeeld.

Afbeelding 9.30 Resultaat tijdreeksmodel met meerdere onttrekkingen



Zoals blijkt uit de afbeelding hierboven is het model goed in staat de stijghoogte te simuleren, terwijl het aantal geschatte parameters ($n=8$) laag blijft. Wanneer meerdere onttrekkingen in een gebied aanwezig zijn dient het aanbeveling een dergelijk model toe te passen zodat het aantal te schatten parameters beperkt blijft.

9.2.8 Veranderingen in het systeem

Veranderingen in de stijghoogte fluctuaties dienen niet altijd gemodelleerd te worden met behulp van een verklarende tijdreeks. Er zijn tal van invloeden te bedenken waarbij simpelweg geen tijdreeks van een verandering beschikbaar is, maar wel een kwalitatieve beschrijving van de verandering of het moment van de verandering. Dit kan bijvoorbeeld een peilopzet zijn in een polder waar het oppervlaktewaterpeil niet wordt gemeten, een verandering van het type vegetatie in de loop van de observatie periode, of het langzaam dichtslibben van de rivierbodem waardoor de respons op fluctuaties in het oppervlaktewaterpeil verandert in de tijd. Grofweg drie methoden zijn te onderscheiden om rekening te houden met dit type veranderingen:

- het toevoegen van een lineaire trend over een bepaalde periode;
- het toevoegen van een staptrend voor een instantane verandering;
- een responsfunctie die door de tijd verandert.

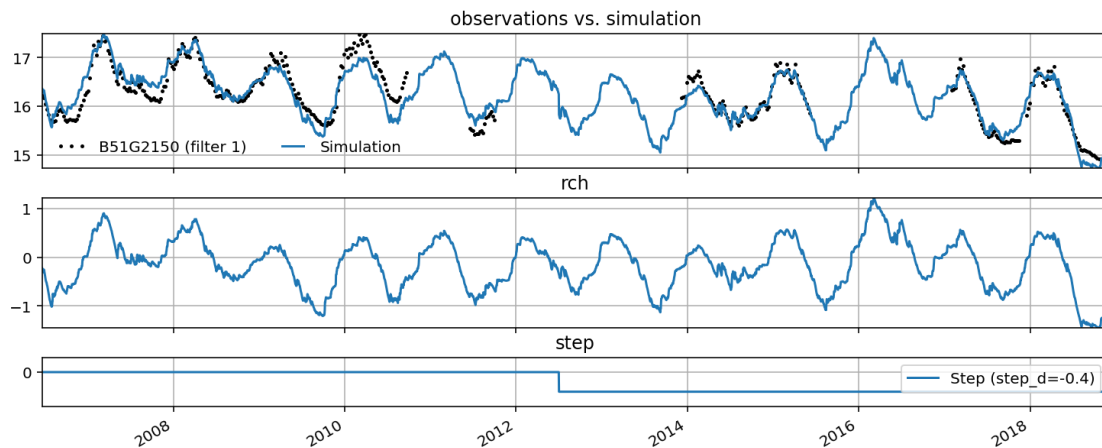
Staptrend

Een eenmalige ingreep in de waterhuishouding kan worden gemodelleerd met behulp van een staptrend (zie bv. von Asmuth et al., 2012). Een staptrend kan gebruikt worden wanneer een peilverandering is doorgevoerd die de stijghoogte beïnvloedt. Er kan hierbij worden gekozen voor een abrupte stijging (wanneer de stijghoogte snel reageert), of een meer geleidelijke stijging (wanneer het grondwatersysteem langzamer reageert). Er kan gebruik gemaakt worden van verschillende responsfuncties die beschrijven hoe de stijghoogte op een verandering reageert.

Voorbeeld peilbuis B51G215

Een voorbeeld van een dergelijk model is gegeven voor filter 1 in waarnemingsput B51G2150. Hierbij is de staptrend gesimuleerd als een instantaan (de responsfunctie is instantaan zonder vertraging) en de hoogte van de staptrend is een kalibratie parameter. In dit voorbeeld is de geschatte stap trend (parameter step_d) - 0.40 meter, met een standaard fout van ± 0.10 meter. Voor de interpretatie en het gebruik van de berekende staptrend voor verdere analyse verwijzen we naar Knotters et al. (2020).

Afbeelding 9.31 Resultaat tijdreeksmodel met stap respons



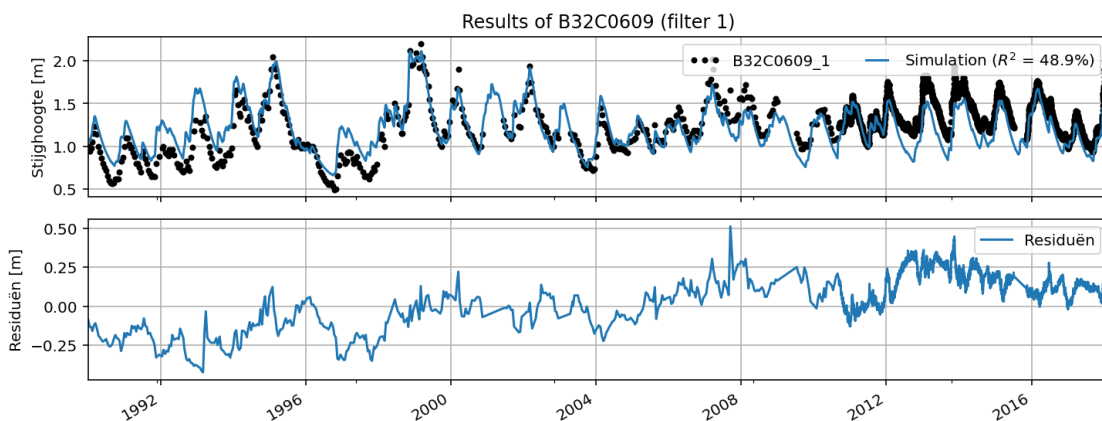
Voorbeeld van een model met lineaire trend

Het gebeurt vaak dat een stijghoogte tijdreeks een opwaartse of neerwaartse trend vertoont, die niet door neerslag of verdamping verklaard kan worden. Vaak is het lastig een specifieke oorzaak aan te wijzen van een daling of stijging. Zo kan het bijvoorbeeld goed dat het vegetatietype en daardoor de verdamping en grondwateraanvulling is veranderd in de loop van de tijd. Het zal echter lastig zijn een dergelijk effect te kwantificeren en mee te nemen in het model. Vast staat echter dat zonder de trend mee te nemen in het model het lastig is de stijghoogte goed te modelleren.

Voorbeeld peilbuis B32C0609

We beschouwen hier een model voor filter 1 in de waarnemingsput B32C0609 nabij de Bilt. Als eerste stap maken we een model met neerslag en verdamping als verklarende variabelen. Dit levert het volgende model op.

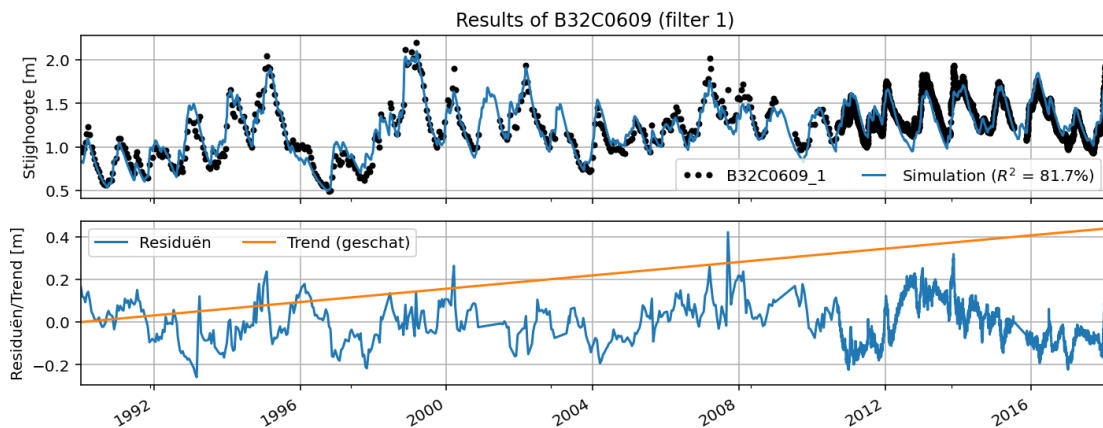
Afbeelding 9.32 Resultaat tijdreeksmodel B32C0609



In de afbeelding hierboven is het resultaat van de modellering weergegeven. We zien dat het model de fluctuaties in de gemeten stijghoogte in het algemeen goed kan verklaren, maar de residuën van het model laten een duidelijke opwaartse trend zien. Het model overschat de stijghoogte in het begin van de kalibratie periode en onderschat de stijghoogte aan het einde van deze periode.

Uit de hydrologische systeemanalyse is gebleken dat er waarschijnlijk geen grondwateronttrekking in de nabijheid van de peilbuis plaatsvindt. Een exacte oorzaak van de trend in de residuën kan dus niet worden aangewezen. In dit geval kan het zinvol zijn een lineaire trend aan het model toe te voegen om de simulatie van de stijghoogte te verbeteren. Het resultaat hiervan is hieronder weergegeven.

Afbeelding 9.33 Resultaat tijdreeksmodel B32C0609 met trend



De afbeelding hierboven laat een duidelijke verbetering in de simulatie van de stijghoogte zien door het toevoegen van een trend. Het is aan de modelleur een start- en einddatum te kiezen, of een trend toe te voegen voor de gehele observatieperiode (zoals hier gedaan). Dit hangt af van het doel van de modellering en de kennis van het systeem en de onderliggende oorzaak van de trend.

Verandering van de responsfunctie in de loop van de tijd

Tot op dit punt in deze paragraaf is verondersteld dat de vorm van de responsfunctie niet verandert in de tijd. Hoewel deze aanname vaak blijkt te werken, zijn er ook situaties waar dit niet het geval is. Obergfell et al. (2019) heeft bijvoorbeeld gekeken naar de situatie waarbij de rivierbodem is uitgebaggerd, waardoor de respons van de stijghoogte op fluctuaties in de rivier waterstand verandert in de tijd. Zij hebben daarvoor een model ontwikkeld waarbij de vorm van de respons verandert over de tijd.

9.2.9 Discussie

In deze paragraaf is getracht een overzicht te geven van de verschillende modelstructuren die in Nederland gangbaar zijn. Het is aan te raden meerdere modelstructuren te testen om de stijghoogte te modelleren, voordat een definitief model wordt gekozen. Het is over het algemeen makkelijk om modelstructuren te variëren om bijvoorbeeld het aantal parameters te verminderen of de fit te verbeteren. Bovendien is het natuurlijk ook mogelijk meerdere modellen te gebruiken. Een dergelijke multi-model aanpak kan een betere inschatting geven van de modelonzekerheid als gevolg van de model structuur (zie bv. Singh et al., 2010 en Ferré, 2017). Waar dit voor numerieke grondwatermodellen een significante tijdsinspanning zal opleveren, is het met tijdreeksmodellen relatief makkelijk meerdere modellen te gebruiken.

Hoewel is geprobeerd hier een zo compleet mogelijk overzicht te geven van de mogelijke modelstructuren, moet worden benadrukt dat het relatief makkelijk is nieuwe modelstructuren te ontwikkelen die zijn toegespitst op specifieke situaties. In deze paragraaf is gebruik gemaakt van tijdreeksen met verschillende meetfrequenties. Een meetreeks met een hogere meetfrequentie kan meer informatie bevatten over processen die met een lagere meetfrequentie niet zichtbaar zijn (zie bv. Herwijnen, 2001 en Berendrecht et al. 2003). Anderzijds reageren veel grondwatersystemen dermate traag dat laag frequente metingen voldoende informatie bevatten. De relatie tussen de meetfrequentie en de informatie die aanwezig is in een tijdreeks is onderdeel van actief onderzoek en zal in zekere mate van belang zijn bij het kiezen van de modelstructuur.

9.3 Kalibratie van een tijdreeksmodel

9.3.1 Introductie

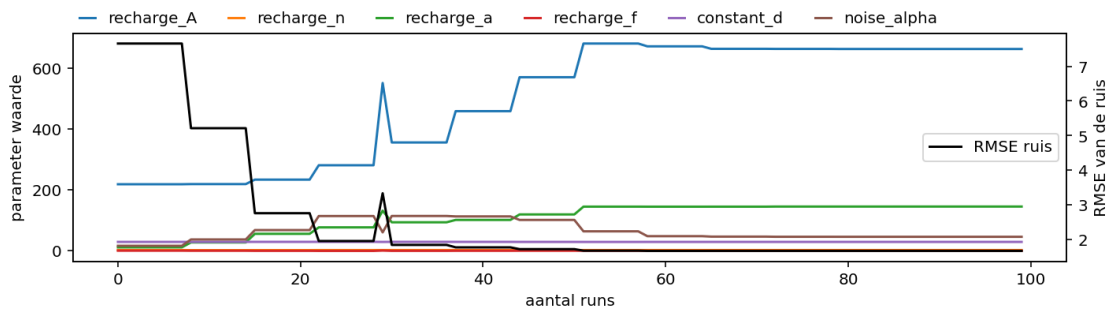
Onder kalibratie van een tijdreeksmodel verstaan we het vinden van de modelparameters zodat de (met het tijdreeksmodel) gesimuleerde stijghoogtes zo goed mogelijk overeenkomen met de gemeten stijghoogtes. De gevonden parameters worden de optimale parameters genoemd. De meest gebruikte methode om de optimale parameters te vinden is de minimalisatie van de som van de kwadraten van de verschillen tussen de gemeten stijghoogtes en de gemodelleerde stijghoogtes, ook wel de kleinste kwadraten methode genoemd. De som van de kwadraten is een niet-lineaire functie van de modelparameters. Om de optimale modelparameters te vinden wordt gebruikt gemaakt van een zoekmethode.

Er bestaan een aantal verschillende zoekmethoden die gebruikt kunnen worden om de optimale modelparameters te vinden. Iedere zoekmethode is gebaseerd op een ander zoek algoritme, maar het idee is altijd hetzelfde: het model wordt een aantal keer gedraaid met verschillende waarden van de modelparameters. Gebaseerd op de verschillende modeluitkomsten (en de daaruit voortvloeiende som van de kwadraten) bepaald het zoekalgoritme een volgende mogelijke set modelparameters die een kleinere som van de kwadraten zou kunnen opleveren. Het zoekalgoritme gaat door totdat het besluit dat de beste parameters gevonden zijn. De efficiëntie van het zoekalgoritme wordt bepaald door het aantal keer dat het model gedraaid moet worden om de optimale modelparameters te vinden. Het is vanzelfsprekend aan te bevelen om de initiële modelparameters zo dicht mogelijk bij de optimale modelparameters te kiezen. Daartoe heeft de meeste software geprobeerd om slimme keuzes te maken voor de initiële waarden van de modelparameters en de initiële parameters deels afhankelijk te maken van de waarden van de invoerreeksen. Soms heeft de modelleur specifieke kennis van het systeem waardoor betere initiële waarden gekozen kunnen worden.

De meest gebruikte zoekalgoritmes zoeken de optimale modelparameters door de modelparameters steeds een beetje aan te passen zodat de som van de kwadraten steeds een beetje kleiner wordt. Op deze manier leidt er een direct pad van de initiële waarden naar de optimale waarden en bij elk stapje langs het pad is de som van de kwadraten kleiner. Dit kan tot gevolg hebben dat het zoekalgoritme in een lokaal minimum belandt. Het is namelijk goed mogelijk dat langs het pad van de initiële parameter waarden naar de optimale parameter waarden de kwadraten som een tijdje toeneemt voordat het weer afneemt. Er bestaan geavanceerde zoekalgoritmes die het model ook draaien met modelparameters die juist een grotere kwadraten som geven met als doel om uiteindelijk het globale minimum te vinden. Dit kost vanzelfsprekend (veel) meer rekentijd.

Ter illustratie wordt hieronder het verloop van de zes parameters en de doelfunctie (in dit geval de 'root mean squared noise') weergegeven voor elke run tijdens het zoekproces naar de optimale parameters. Dezelfde buis wordt in het volgende voorbeeld gebruikt, waar de buis verder beschreven wordt. Voor elke iteratie in het zoekproces wordt het model een aantal keer gedraaid, tenminste 1 keer per parameter. Vandaar dat parameters stapsgewijs omhoog of omlaag gaan. Het model is bijna 100 keer gedraaid om de optimale parameters te vinden.

Afbeelding 9.34 visualisatie van parameter optimalisatie

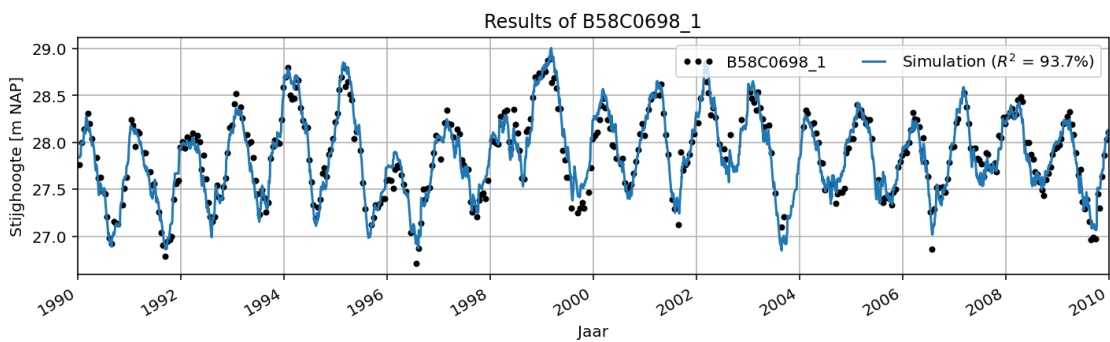


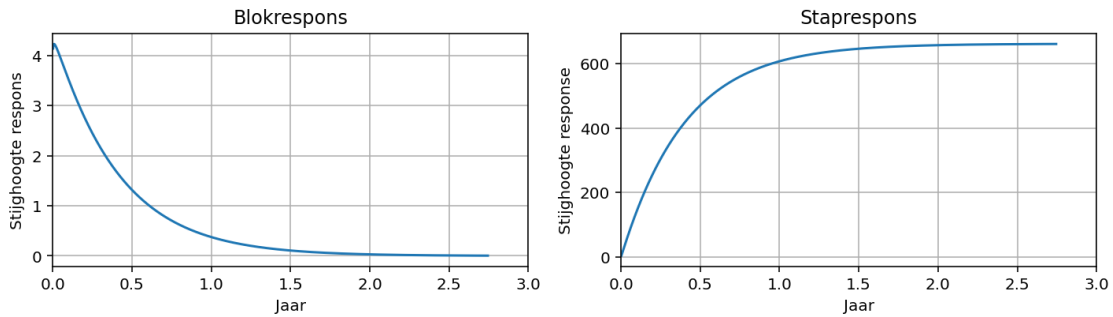
9.3.2 Effect van de lengte van de tijdreeks (of kalibratieperiode)

Het effect van de lengte van de kalibratieperiode op de uitkomsten van een tijdreeksmodel wordt besproken aan de hand van de metingen in buis B58C0698 in de plaats Swartbroek nabij Weert; de data voor de periode 1990-2010 worden beschouwd, waarin ongeveer twee keer per maand gemeten is. De regenreeks komt van regenstation Heibloem en de potentiële verdamping van weerstation Maastricht.

Er is een tijdreeksmodel gemaakt waarbij de regen en potentiële verdamping de verklarende reeksen zijn en de geschaalde Gamma functie gebruikt wordt als responsfunctie. De hele reeks wordt gebruikt om het model te kalibreren. Het model fit de data goed, zoals in onderstaande figuur te zien is. De blokrespons en staprespons functies worden ook getoond. Het geheugen van de blokrespons functie is iets meer dan 2.5 jaar. Na een regenbui op de eerste dag gaat de stijghoogte vrij snel omhoog, maar het duurt ongeveer 2.5 jaar voordat de stijghoogte weer terug is op het niveau voor de regenbui. De piek van de blokrespons functie is iets boven de 4, hetgeen betekent dat als het in een dag 1 mm regent, dat dan de stijghoogte uiteindelijk 4 mm omhoog gaat. Dat komt overeen met een effectieve porositeit (freatische berging) van $1 / 4 = 0.25$, wat een redelijke waarde is. De staprespons gaat uiteindelijk naar een waarde iets boven de 650. Dat betekent dat als het vanaf de eerste dag continu 1 mm/d regent, dat dan de stijghoogte uiteindelijk iets meer dan 650 mm omhoog gaat. Dit is gelijk aan de parameter A van de Gamma respons functie (zie paragraaf 9.2).

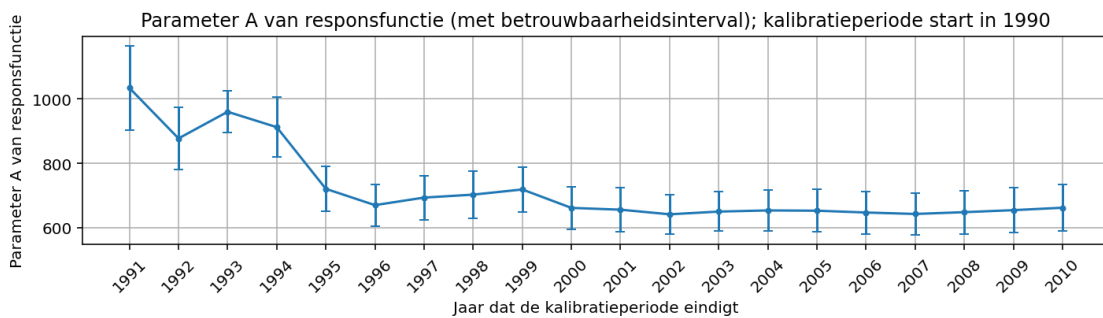
Afbeelding 9.35 Resultaten van de simulatie, met de bijbehorende stap- en blokrespons





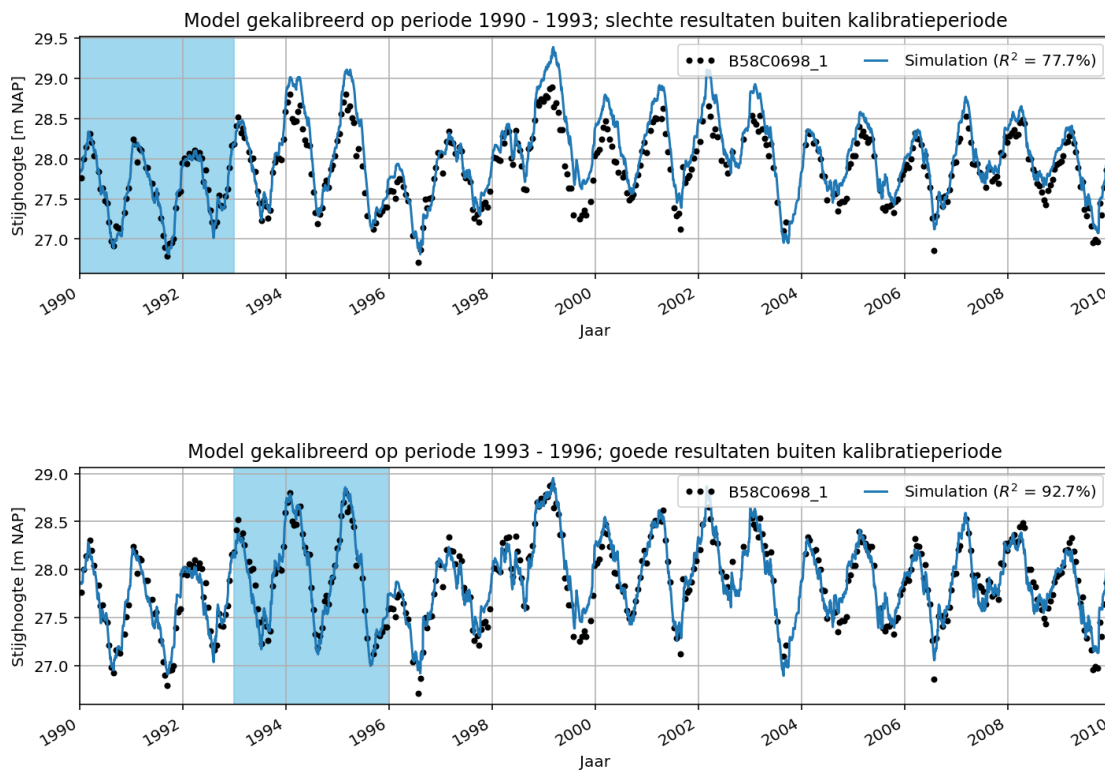
In het model hierboven wordt 20 jaar gebruikt om het model te kalibreren. Wat nu als er een kortere reeks gebruikt wordt? Wordt dan dezelfde responsfunctie gevonden? Om dit te onderzoeken wordt het model gekalibreerd op 1 jaar data (1990-1991), twee jaar data (1990-1992), et cetera tot 20 jaar data (1990-2010). Voor elke kalibratie wordt de waarde van de parameter A opgeslagen (de uiteindelijke hoogte van de staprespons). In onderstaande figuur wordt de gekalibreerde waarde van A, inclusief 95 % betrouwbaarheidsinterval, geplot tegen de lengte van de kalibratieperiode. Uit de figuur blijkt dat als de kalibratieperiode lang genoeg is, er steeds ongeveer hetzelfde uitkomt. Voor dit geval is een kalibratieperiode van 5 jaar (1990-1995) genoeg. Dat is, voor dit geval, ongeveer twee keer het geheugen van de blokrespons.

Afbeelding 9.36 Parameter A van de responsfunctie



Bij een kortere kalibratieperiode wordt, voor dit voorbeeld, een te hoge waarde van A geschat en is ook het betrouwbaarheidsinterval te klein. Immers, de waarde van A die gevonden wordt voor een langere kalibratieperiode ($A \approx 650$) ligt ver buiten het betrouwbaarheidsinterval van de waarde van A die gevonden wordt bij een kalibratieperiode van 1 of een paar jaar. De waarde die gevonden wordt met, bijvoorbeeld, een kalibratieperiode van 3 jaar geeft dus wel een goede fit voor de drie jaar van de kalibratieperiode, maar daarna zal de gesimuleerde stijghoogte vooral boven de gemeten stijghoogte liggen, zoals hieronder weergegeven is. Als een andere periode van 3 jaar gekozen wordt als kalibratieperiode, dan kan het model wel redelijk kloppen, zoals hieronder is weergegeven voor de periode 1993-1996, maar dat is alleen te bepalen als er ook data beschikbaar is voor de validatieperiode.

Afbeelding 9.37 Tijdreeksmodellen gekalibreerd op verschillende periodes



Intuïtief lijkt het te kloppen dat de kalibratieperiode ten minste een aantal keer het geheugen van het systeem moet zijn. Het is echter moeilijk om een vaste regel te geven voor de benodigde lengte van de kalibratieperiode. Als voorbeeld is hierboven geïllustreerd dat 3 jaar genoeg kan zijn voor deze reeks, maar dat het wel uitmaakt welke 3 jaar gekozen worden en dat werkt natuurlijk niet prettig in de praktijk.

Van der Spek en Bakker (2017) hebben onderzocht wat de minimale lengte van een kalibratieperiode is waarna de onzekerheid van voorspelde stijghoogtes niet verder afneemt. Zij onderzochten 18 reeksen met ongeveer twee metingen per maand en responstijden die varieerden van 60 tot 1200 dagen, maar konden geen duidelijke relatie vinden tussen het geheugen van het systeem (de responstijd) en de lengte van de kalibratieperiode. Bij een kalibratieperiode van 20 jaar, vonden zij dat het gemiddelde 95 % voorspellingsinterval van de stijghoogte ongeveer 50 % van de totale stijghoogte variatie was (het verschil tussen de hoogst en laagst gemeten stijghoogte), terwijl het betrouwbaarheidsinterval van het model ongeveer 10 % van de totale stijghoogte variatie was. Bij een kalibratieperiode van slechts 10 jaar hingen de resultaten af van welke periode van 10 jaar gekozen werd: bij sommige periodes was het betrouwbaarheidsinterval veel groter dan bij andere periodes. Bij een kalibratieperiode van slechts 5 jaar was het betrouwbaarheidsinterval vrijwel altijd groot. De conclusie van der Spek en Bakker (2017) is dat de betrouwbaarheid van het model toeneemt als de kalibratieperiode langer is, maar dat de betrouwbaarheid niet verder toeneemt als de kalibratieperiode langer is dan 20 jaar.

9.3.3 Het ruismodel

Een ruismodel kan gebruikt worden om te proberen er voor te zorgen dat de overgebleven verschillen tussen het model en de metingen aan een aantal statistische toetsen voldoen zodat er statistische uitspraken gedaan kunnen worden met het model (bijvoorbeeld: wat is de kans dat de stijghoogte boven een bepaalde grenswaarde uitkomt). In het vorige model was er ook een ruismodel gebruikt, hoewel daar niet over uitgeweid is, zodat de ruis bij benadering voldeed aan de statistische toetsen en er een betrouwbaarheidsinterval van de parameters geschat kon worden.

De verschillen tussen de gemeten stijghoogte en de gemodelleerde stijghoogte (de residuën) zijn vrijwel altijd gecorreleerd in de tijd. Simpel gezegd: als het model vandaag iets hoger is dan de gemeten stijghoogte, dan is de kans groot dat het model volgende week ook iets hoger is dan de gemeten stijghoogte. Dat is op zich prima te verklaren. Het kan bijvoorbeeld zo zijn dat er een regen reeks gebruikt wordt van een weerstation dat 20 km verderop ligt. Als het nu bij het regenstation wel geregend heeft maar bij de peilbuis niet, dan zal het model waarschijnlijk een te hoge stijghoogte simuleren. En aangezien het een tijdje duurt voordat de stijghoogte weer gezakt is na een regenbui, zal het model waarschijnlijk voor een aantal dagen/weken/maanden te hoge waarden geven. Gelukkig is de kans groot dat het een tijdje later bij de peilbuis juist wel regent terwijl het bij het regenstation niet regent, zodat de effecten elkaar weer een beetje opheffen. Een 'fout' in de regen is slechts één van de mogelijke oorzaken van een verschil tussen model en meting. Andere oorzaken zijn bijvoorbeeld meetfouten, niet gemeten veranderingen en niet-lineariteit van het systeem.

Autocorrelatie in de ruis vormt een probleem als het model gebruikt gaat worden om statistische uitspraken te doen. Als er statistische uitspraken gedaan moeten worden, dan moeten de residuën aan bepaalde statistische voorwaarden voldoen. Eén van de belangrijkste voorwaarden is dat de residuën niet gecorreleerd zijn. Een ruismodel kan toegevoegd worden aan een tijdreeksmodel om te trachten de gecorreleerde residuën te transformeren naar ruis die ongecorrleerd is. Het eenvoudigst mogelijke ruismodel is een zogenaamd AutoRegressief model van orde 1, of kortweg een AR1 model. In dit model is het residu $\varepsilon(t)$ op tijdstip t gelijk aan een factor ρ keer het residu op tijdstip $t - \Delta t$ plus een random (ongecorreleerde) fout $a(t)$, of in formule vorm:

$$\varepsilon(t) = \rho \varepsilon(t - \Delta t) + a(t)$$

Als de tijd Δt tussen twee residuën varieert, dan kan de factor ρ een functie zijn van Δt die afneemt naarmate de tijdstap Δt tussen twee metingen groter wordt. De formule is dan:

$$\varepsilon(t) = e^{-\Delta t/\alpha} \varepsilon(t - \Delta t) + a(t)$$

waarbij α een parameter is die aangeeft hoe snel de correlatie tussen residuën afneemt als Δt toeneemt. De correlatie is verwaarloosbaar als $\Delta t > 3\alpha$ (want $e^{-3} \approx 0.05$).

In onderstaande voorbeeld worden synthetische reeksen gegenereerd en geanalyseerd om te laten zien dat bij tijdreeksmodellen zonder ruismodel de parameters wel goed geschat worden, maar de betrouwbaarheidsintervallen onderschat worden.

9.3.4 Effect van een ruismodel

Voor het maken van de synthetische reeks wordt de dagelijkse grondwateraanvulling gelijk gesteld aan de gemeten regen min de gemeten potentiële verdamping van het vorige voorbeeld (dat is dus echte regen- en verdampingsdata). De responsfunctie van de grondwateraanvulling is een exponentiële functie met $A=600$ en $a=150$ d. Het basisniveau is $d=25$ m. Een synthetische stijghoogte reeks wordt gesimuleerd op dezelfde tijdstippen waarop gemeten was in het vorige voorbeeld. Deze gegenereerde synthetische stijghoogtereeks bevat geen fouten. Een tijdreeksmodel van deze synthetische stijghoogtereeks met een exponentiële responsfunctie kan de parameters A , a en d vrijwel exact terugvinden, zoals hieronder te zien is:

Fit report Simulation		Fit Statistics	
nfev	10	EVP	100.00
nobs	423	R2	1.00
noise	1	RMSE	0.00
tmin	1990-01-14 00:00:00	AIC	-19131.64
tmax	2009-12-28 00:00:00	BIC	-19115.45
freq	D	Obj	0.00
warmup	3650 days 00:00:00	—	
solver	LeastSquares	Interp.	No

Parameters (4 optimized)

	optimal	stderr	initial	vary
recharge_A	600.000000	±0.00%	217.448496	True
recharge_a	150.000000	±0.00%	10.000000	True
constant_d	25.000000	±0.00%	25.271506	True
noise_alpha	178.473347	±14.56%	15.000000	True

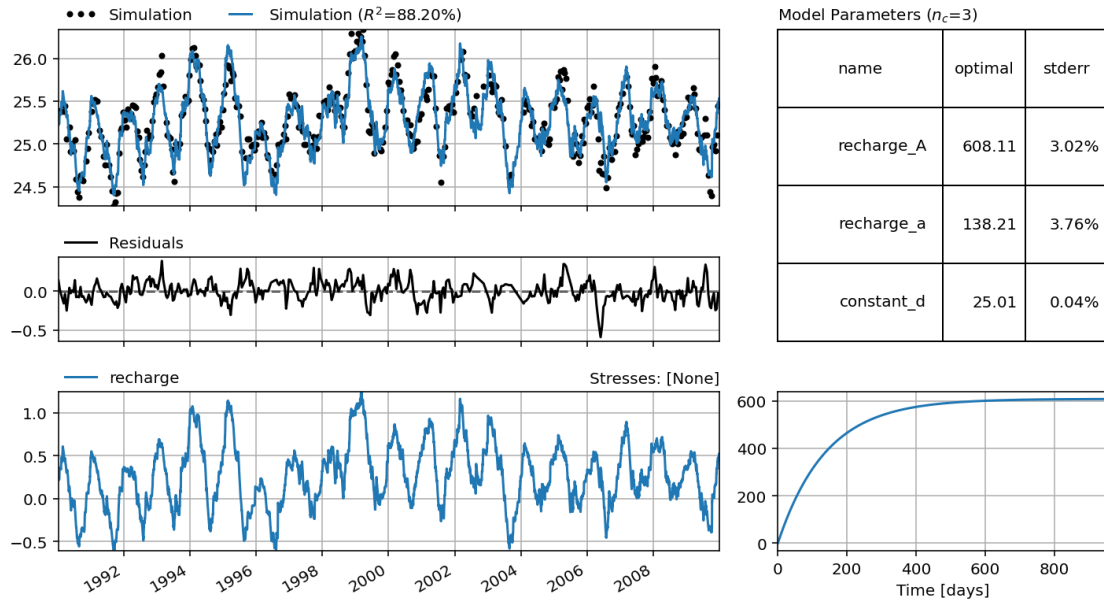
Om de werking van het ruismodel te illustreren, wordt er een residu (een fout) opgeteld bij de synthetische stijghoogte reeks. Het residu op tijdstip t , $\varepsilon(t)$, is gecorreleerd met het residu bij de vorige meting op tijdstip $t - \Delta t$ volgens de formule die hierboven al gegeven was. De ruis $a(t)$ wordt getrokken uit een normale verdeling met gemiddelde nul en standaard afwijking σ_a . Hoe groter de correlatie tussen de residuën is, hoe groter de standaardafwijking van de residuën. De standaard afwijking σ_r van de residuën kan berekend worden uit de standaardafwijking σ_a van de ruis als:

$$\sigma_r = \sigma_a / \sqrt{1 - e^{-2\Delta t/\alpha}}$$

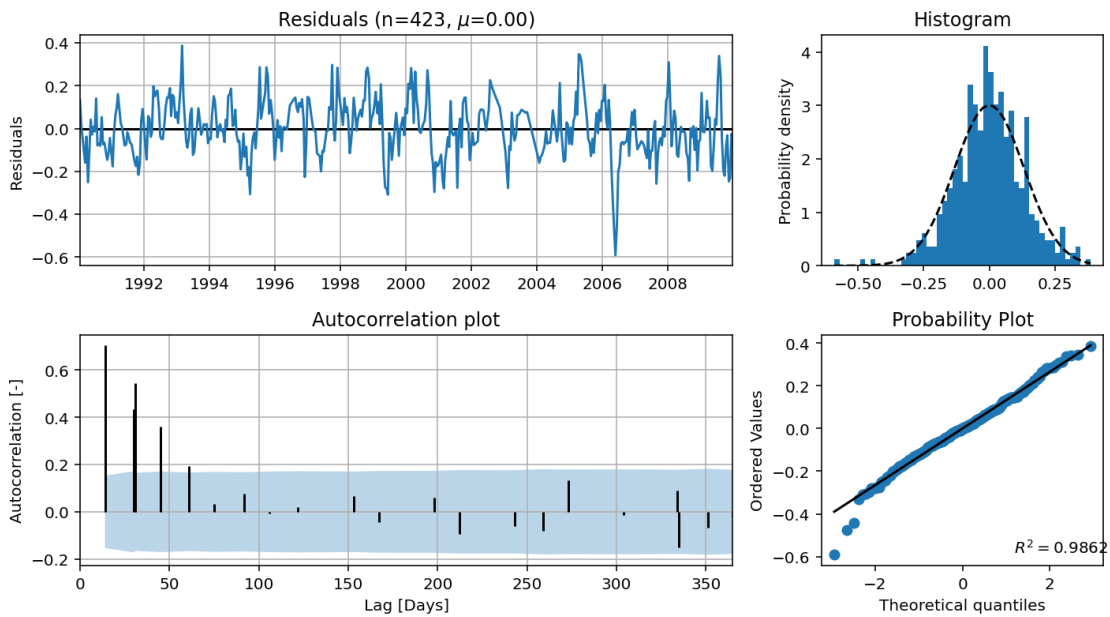
De residuën worden gegenereerd met $\sigma_a=0.1$ m, $\alpha=50$ d. De standaardafwijking van de residuën, σ_r , voor een tijdstap van 14 dagen is dan gelijk aan 0.15 m

De synthetische stijghoogte reeks bevat nu een gecorreleerde fout, net als in de werkelijkheid. We gaan nu eerst een tijdreeksmodel maken van de synthetische reeks zonder ruismodel. Het tijdreeksmodel blijkt een goede fit te geven en de geschatte parameters komen redelijk goed overeen met de parameters waarmee de synthetische reeks gemaakt was. Merk echter wel op dat de waarde van a die gebruikt is om de synthetische reeks te maken ($a=150$ d), niet binnen het geschatte betrouwbaarheidsinterval van de geschatte waarde van a ligt. De diagnostische test laat echter zien dat er autocorrelatie in de residuën zit, dus formeel klopt het geschatte betrouwbaarheidsinterval niet. Dit gaan we verder onderzoeken. Het geschatte 95 % betrouwbaarheidsinterval van a is 128.03 - 148.40 dagen.

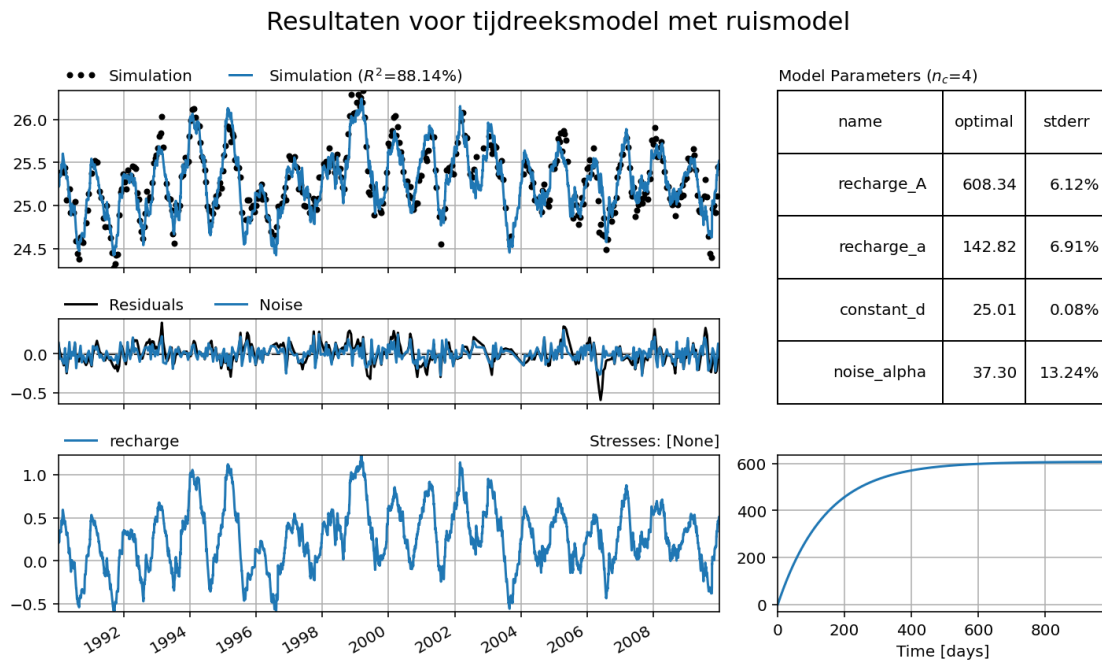
Resultaten voor tijdreeksmodel zonder ruismodel



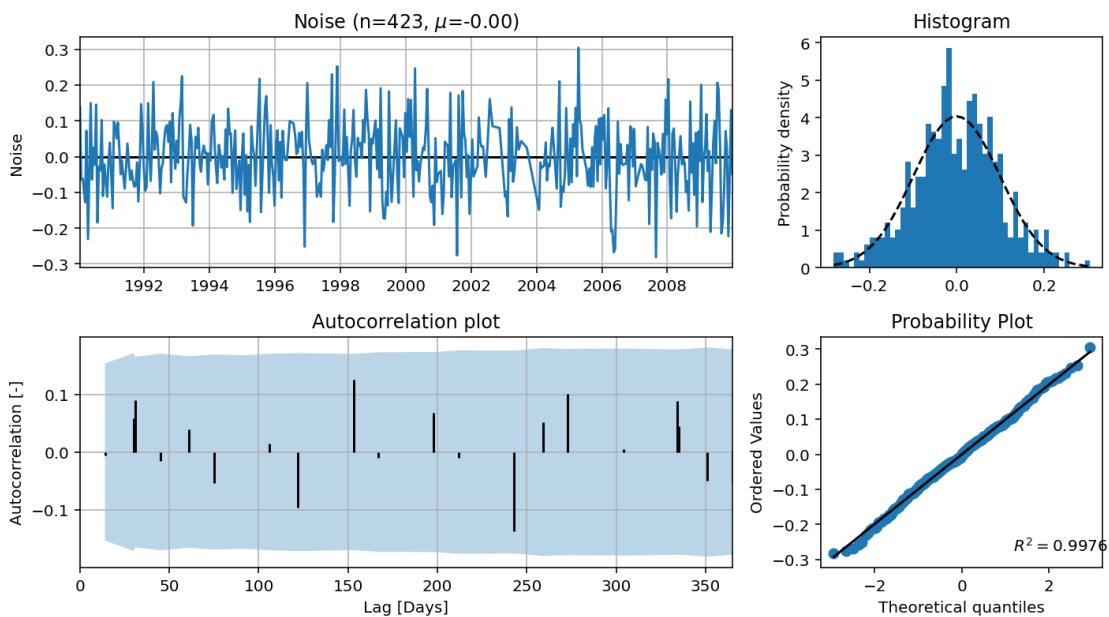
Diagnostische checks voor tijdreeksmodel zonder ruismodel



Als we nu de wederom een tijdreeksmodel maken van de synthetische stijghoogtereeks, maar dit keer wel een ruismodel toevoegen, dan is de fit wederom goed en komen de geschatte parameters wederom goed overeen met de opgegeven parameters, maar nu is de autocorrelatie van de ruis binnen het geaccepteerde niveau. Het ruismodel heeft zijn werk goed gedaan. De fit van het tijdreeksmodel met een ruismodel is altijd lager dan de fit van een tijdreeksmodel zonder ruismodel. Het verschil is hier echter klein. Het geschatte 95 % betrouwbaarheidsinterval van a is 123.46 - 162.17 dagen.



Diagnostische checks voor tijdreeksmodel met ruismodel



9.3.5 Effect van het ruismodel op de schatting van de waarden en betrouwbaarheid van de parameters

Als het ruismodel niet toegepast wordt, dan valt de echte waarde van a dus niet binnen het geschatte betrouwbaarheidsinterval, terwijl als het ruismodel wel toegepast wordt, valt de echte waarde van a wel binnen het geschatte betrouwbaarheidsinterval. Dit kan natuurlijk toeval zijn (het is immers een 95 % betrouwbaarheidsinterval, dus er is een 5 % kans dat de echte waarde er buiten valt). Daarom gaan we een experiment doen. In het experiment creëren we 100 keer een synthetische reeks, met steeds andere (gecorrleerde) foutjes. Van alle 100 synthetische reeksen maken we een tijdreeksmodel.

Daar de synthetische reeks elke keer andere (gecorrleerde) foutjes heeft, zijn de optimale parameters elke keer een beetje anders. Het experiment is twee keer gedaan. De eerste keer is er 100 keer een tijdreeksmodel gemaakt zonder een ruismodel en vervolgens is er 100 keer een tijdreeksmodel gemaakt met een ruismodel.

Tabel 9.2 Vergelijking parameters tijdreeksmodellen met en zonder ruismodel

Parameter	Zonder ruismodel	Met ruismodel
A	603.36	600.78
a	151.43	150.43
d	25.00	25.00
α	-	45.86
EVP	85.12	85.07

Het blijkt dat zowel het tijdreeksmodel zonder het ruismodel als het tijdreeksmodel met het ruismodel gemiddeld een goede schatting geeft van de parameters. Dus waarom moeten we dan een ruismodel gebruiken? De reden is dat de geschatte onzekerheid van de parameters te klein is als er geen ruismodel gebruikt wordt. De statistische manier waarop de onzekerheid geschat wordt gaat er namelijk van uit dat de overgebleven ruis niet gecorrleerd is. Als de ruis echter toch gecorrleerd is, dan is de geschatte onzekerheid te klein. Dit kan gemakkelijk nagegaan worden met het hierboven uitgevoerde experiment. Tijdens het experiment zijn de parameters 100 keer geschat inclusief een schatting van de standaard fout van de parameters. Onder enkele statistische aannames (inclusief de aannames van normaliteit en ongecorrleerde ruis), is het 95 % betrouwbaarheidsinterval tussen min twee keer de standaard fout tot plus twee keer de standaard fout (eigenlijk is het 1.96 in plaats van 2, maar laten we dat hier maar afronden). Er kan dus eenvoudig nagegaan worden hoe vaak de echte waarden van de parameters (waarmee de synthetische reeks gegenereerd is) binnen het 95 % betrouwbaarheidsinterval van de geschatte parameters liggen. Aangezien we het experiment 100 keer gedaan hebben, zou dit ongeveer 95 keer het geval moeten zijn. Zoals te zien is hieronder, geeft het tijdreeksmodel zonder ruismodel slechts zo'n 60-65 keer een betrouwbaarheidsinterval dat de echte waarde bevat. Het tijdreeksmodel met een ruismodel doet het veel beter: meer dan 95 keer liggen de echte parameters van de responsfuncties binnen het 95 % betrouwbaarheidsinterval. Samenvattend kan gezegd worden dat zonder ruismodel de parameters gemiddeld wel goed geschat worden, maar dat de betrouwbaarheidsintervallen van de parameters te klein geschat worden.

Tabel 9.3 Parameters binnen betrouwbaarheidsinterval met en zonder ruismodel

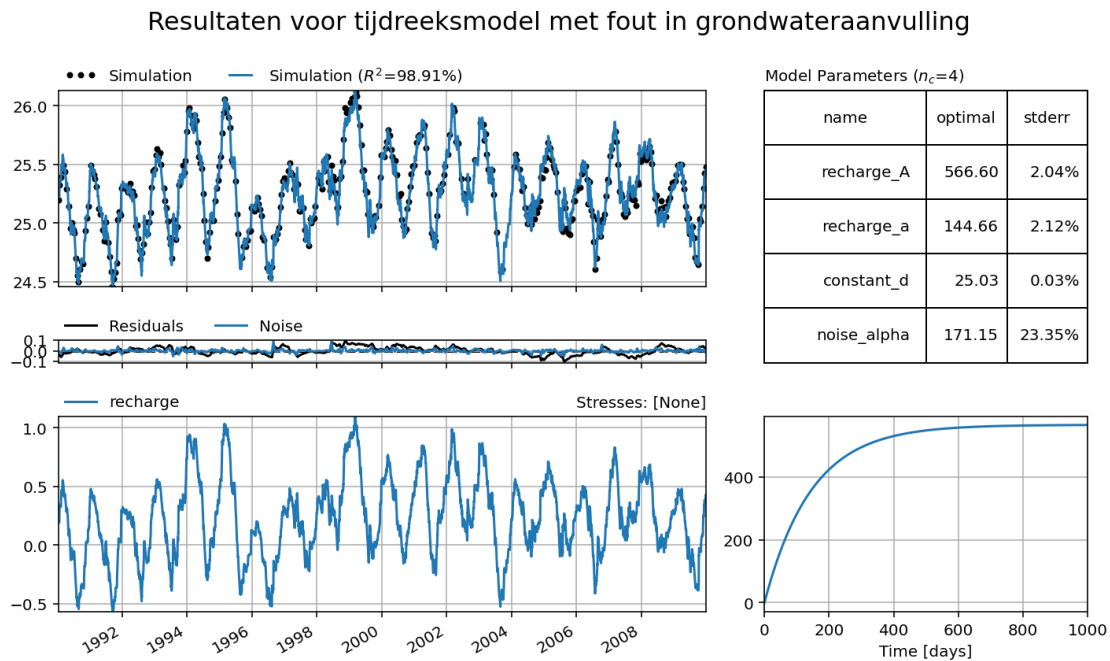
Parameter	Zonder ruismodel	Met ruismodel
A	66	96
a	72	95
d	53	98
α	-	86

9.3.6 Effect van een fout in de gemeten grondwateraanvulling

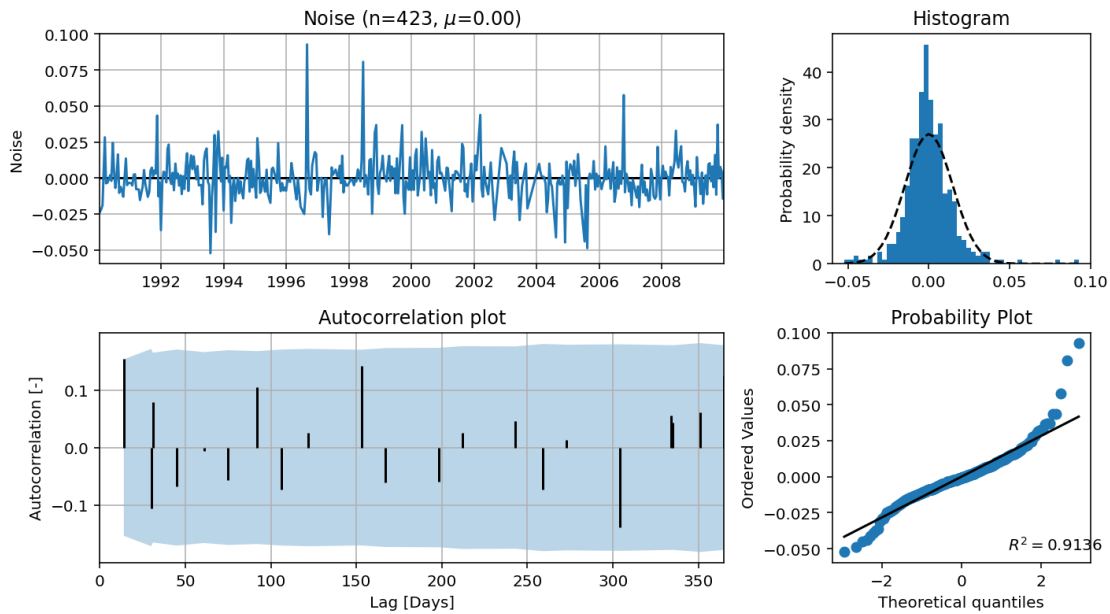
In bovenstaand experiment is uitgelegd wat het voordeel van het gebruik van een ruismodel is. Als er geen ruismodel gebruikt wordt, dan worden de parameters goed geschat, maar de betrouwbaarheidsintervallen van de parameters zijn te klein. Als het ruismodel wel meegenomen wordt, dan wordt het betrouwbaarheidsinterval veel beter geschat. In bovenstaand voorbeeld was de fout in de stijghoogte gegenereerd met een model dat precies hetzelfde is als het ruismodel dat in het tijdreeksmodel gebruikt werd. Zoals uitgelegd is de correlatie in de residuën in de werkelijkheid het gevolg van fysische processen, bijvoorbeeld een verschil tussen de gemeten regenval bij het weerstation en de gevallen regenval bij de peilbuis. In de volgende test wordt het bovenstaande experiment herhaald, maar deze keer wordt er geen gecorreleerde fout opgeteld bij de stijghoogte, maar wordt er een grondwater aanvullingsreeks (regen min verdamping) gebruikt die een fout bevat. Er zijn vele manieren waarop er een realistische fout toegevoegd kan worden aan regen of verdampingsreeksen. Hier is er voor gekozen om de gemeten grondwateraanvulling te vermenigvuldigen met een factor $(1 + \varepsilon_g)$, waarbij ε_g normaal verdeeld is met een gemiddelde nul en standaardafwijking σ_g . Zodoende is de grondwateraanvulling elke dag dus iets meer of iets minder dan gemeten. De gekozen waarde van σ_g is 0.2.

Hieronder wordt een tijdreeksmodel gemaakt, waarbij wederom een AR1 ruismodel meegeeft wordt. Zoals te zien is, is de fit zeer goed en is er vrijwel geen autocorrelatie in de ruis, maar de ruis is niet erg normaal verdeeld. Dit baart zorgen, want in het bepalen van de betrouwbaarheidsintervallen van de parameters wordt er vanuit gegaan de ruis normaal verdeeld is.

Afbeelding 9.40 Resultaten tijdreeksmodel met fout in grondwateraanvulling



Diagnostische checks voor tijdreeksmodel met fout in grondwateraanvulling



We gaan wederom het hierboven beschreven experiment uitvoeren. De tijdreeksmodellen zonder noise model blijken de parameters wederom goed te schatten, maar de betrouwbaarheidsintervallen van de parameters zijn veel te klein: bij slechts 30-40 % van de modellen liggen de echte parameters binnen de betrouwbaarheidsintervallen van de geschatte parameters. Ook de tijdreeksmodellen met ruismodel geven goede resultaten voor de geschatte parameters. Bij 80-85 % van de modellen liggen de echte parameters binnen de betrouwbaarheidsintervallen van de geschatte parameters. Dat is minder dan de verwachte 95 % en komt hoogst waarschijnlijk omdat de ruis niet normaal verdeeld is. Om dit beter te krijgen zou een tijdreeksmodel gebruikt moeten worden gebaseerd op, bijvoorbeeld, een maximum likelihood schatting van de parameters, waarbij opgegeven kan worden wat de verwachte verdeling is van de ruis (zie bijvoorbeeld van der Spek en Bakker, 2017). In de meeste huidige versies van software voor tijdreeksanalyse is het echter niet mogelijk om een maximum likelihood functie te definiëren.

Tabel 9.4 Vergelijking modellen met en zonder ruismodel

Parameter	Parameterwaarde zonder ruismodel	Parameter binnen betrouwbaarheidsinterval	Parameterwaarde met ruismodel	Parameter binnen betrouwbaarheidsinterval
A	591.77	38	593.64	80
a	149.31	37	150.54	84
d	25.00	30	25.00	86
α			142.90	
EVP	99.16		99.15	

9.3.7 Effect van de meetfrequentie op de werking van het AR1 ruismodel

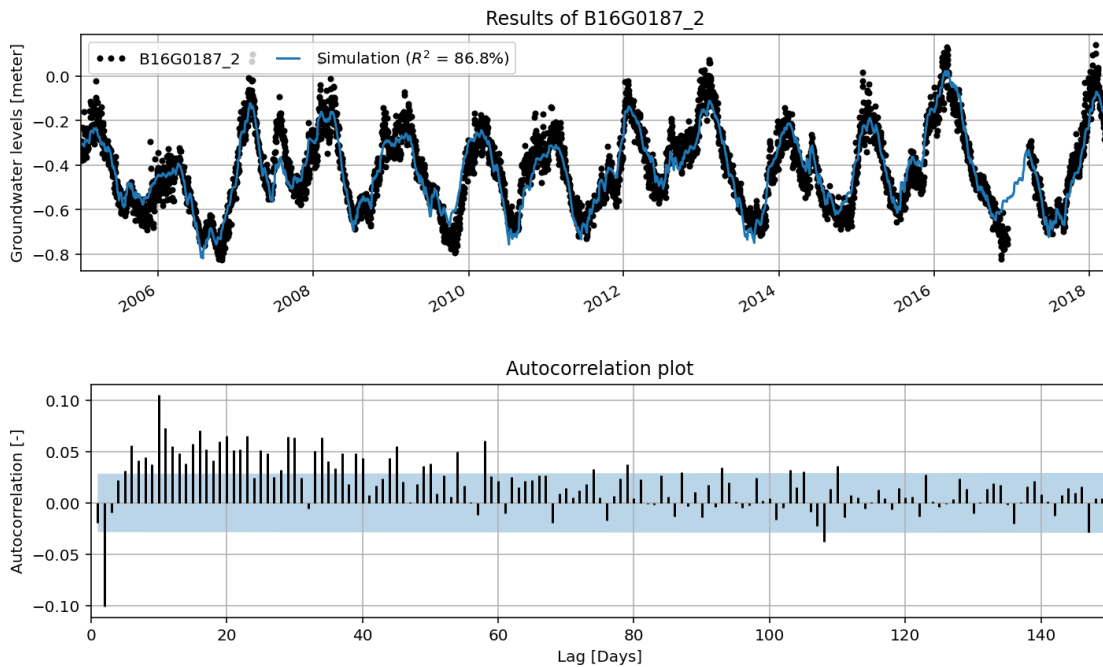
Het is vaak moeilijk, met de huidige algoritmes, om de residuën te transformeren naar ruis met verwaarloosbare autocorrelatie voor tijdreeksmodellen van stijghoogtes die dagelijks gemeten zijn. Een praktische oplossing kan dan zijn om de frequentie van de metingen te verlagen, bijvoorbeeld om alleen een meting om de week of om de twee weken te gebruiken.

Beschouw bijvoorbeeld peilbuis B16G0187_2 in het plaatsje Witte Paarden in Steenwijkerland. De stijghoogte is dagelijks gemeten tussen 2005 en 2018. Een tijdreeksmodel met regen en potentiële verdamping geeft een goede fit (zie hieronder), maar er is nog een duidelijke autocorrelatie in de residuën.

Fit report B16G0187_2		Fit Statistics	
nfev	30	EVP	86.78
nobs	4744	R2	0.87
noise	1	RMSE	0.07
tmin	2005-01-01 00:00:00	AIC	9681.41
tmax	2018-04-04 00:00:00	BIC	9720.20
freq	D	Obj	3.84
warmup	3650 days 00:00:00	—	
solver	LeastSquares	Interp.	No

Parameters (6 optimized)				
	optimal	stderr	initial	vary
recharge_A	324.803629	±3.99%	202.753032	True
recharge_n	1.161756	±2.24%	1.000000	True
recharge_a	128.228562	±6.00%	10.000000	True
recharge_f	-0.986536	±3.40%	-1.000000	True
constant_d	-0.666379	±3.52%	-0.419248	True
noise_alpha	4.275741	±4.82%	1.000000	True

Afbeelding 9.41 resultaten tijdreeksmodel met autocorrelatie van de residuen

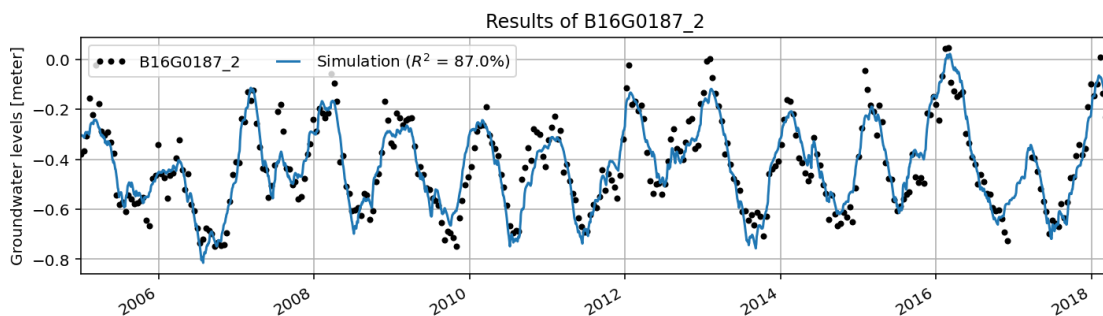


Vervolgens is het tijdreeksmodel nogmaals gefit, maar nu alleen op 14-daagse metingen. Het aantal observaties waarop gefit wordt is daardoor teruggebracht van meer dan 4500 naar minder dan 400. De fit is ongeveer even goed en alle modelparameters zijn vergelijkbaar, behalve de parameter α van het ruismodel die nu 5 keer zo groot is. Maar, en daar ging het om, de autocorrelatie in de ruis is nu verwaarloosbaar, zodat de standaardfouten (stderr) van de parameters gebruikt mogen worden om schattingen te doen van de betrouwbaarheidsintervallen van de parameters (als ook aan de andere statistische voorwaarden voldaan is). Merk ook op dat voor deze buis, de standaardfout van de parameters in het model dat gefit is op 14-daagse metingen voor de meeste parameters ongeveer twee keer zo groot is als voor de parameters in het model dat gefit is op dagelijkse metingen.

Fit report B16G0187_2		Fit Statistics	
nfev	14	EVP	86.99
nobs	339	R2	0.87
noise	1	RMSE	0.07
tmin	2005-01-01 00:00:00	AIC	25.45
tmax	2018-03-24 00:00:00	BIC	48.41
freq	D	Obj	0.52
warmup	3650 days 00:00:00	—	
solver	LeastSquares	Interp.	No

Parameters (6 optimized)				
	optimal	stderr	initial	vary
recharge_A	311.113217	±8.37%	202.753032	True
recharge_n	1.282130	±4.82%	1.000000	True
recharge_a	104.497496	±11.70%	10.000000	True
recharge_f	-0.947278	±7.20%	-1.000000	True
constant_d	-0.678718	±7.12%	-0.420647	True
noise_alpha	22.393984	±13.98%	14.000000	True

Afbeelding 9.42 resultaten tijdreeksmodel



9.4 Beoordeling Tijdreeksmodellen

Een veel gestelde vraag is hoe je een geoptimaliseerd tijdreeksmodel moet beoordelen. In deze paragraaf worden methodes voor modelbeoordeling gepresenteerd en geïllustreerd met voorbeelden uit de praktijk.

9.4.1 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. Dit voorbeeld geeft een overzicht van veel voorkomende methodes om tijdreeksmodellen te beoordelen. Per onderwerp wordt de belangrijkste achtergrondinformatie gepresenteerd. Deze paragraaf bevat 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.

9.4.2 Algemeen stappenplan

We presenteren hier een aantal stappen die vaak worden toegepast om tijdreeksmodellen te beoordelen. Er is niet zoiets als een definitief stappenplan voor modelbeoordeling, dus er bestaan ook andere methodes voor modelbeoordeling die niet in deze paragraaf zijn opgenomen. Ook de volgorde van de stappen staat niet vast, maar is logisch gestructureerd naar inzicht van de auteurs:

- **visuele beoordeling:** een eerste goede stap is vaak een visuele beoordeling van het tijdreeksmodel. Het gedrag van het model wordt visueel geïnspecteerd waarbij veel waardevolle inzichten worden verkregen over het model. Het nadeel van deze stap is dat deze relatief veel tijd kost als het om veel modellen gaat;
- **fit-statistiek:** de tweede stap is een beschouwing van de overeenkomst tussen model en metingen (de fit) op basis van fit-statistieken. Het voordeel van deze stap is dat er een objectieve maat wordt gegeven die iets zegt over overeenkomst tussen model en metingen. Het nadeel van deze stap is dat er veel waardevolle informatie verloren gaat bij het terugbrengen van de prestatie van het model tot enkele getallen;
- **testen model:** de derde stap is het testen van het model (dit wordt ook vaak validatie genoemd), door bijvoorbeeld een simulatie te doen met het model op een periode met metingen die niet zijn gebruikt in de kalibratie. Hiermee wordt het model getest op informatie die niet is gebruikt bij het optimaliseren van de modelparameters;
- **modelonzekerheid:** de vierde stap is het beschouwen van de modelonzekerheid. Daarbij stellen we eigenlijk de vraag: kunnen de modelparameters goed bepaald worden? Voordat we die onzekerheid kunnen beoordelen, moet het model aan enkele statistische eisen voldoen. Als het model niet aan die eisen voldoet, kunnen we de onzekerheid van de modelparameters niet goed inschatten, en dus niet gebruiken bij onze beoordeling. De parameters worden goed geschat, maar de onzekerheden dus niet;
- **hydrologische beoordeling:** de laatste stap die we hier presenteren is de hydrologische beoordeling waarbij de resultaten van de tijdreeksmodellen naast de hydrologische kennis over het systeem worden gelegd. Dit kan zowel op individuele tijdreeksmodellen worden toegepast als ruimtelijk op de resultaten van meerdere tijdreeksmodellen. Hier wordt bijvoorbeeld de kennis over het hydrologische systeem gebruikt om de modellen te beoordelen, maar worden ook de resultaten van de modellen gebruikt om het hydrologisch inzicht in het systeem uit te breiden.

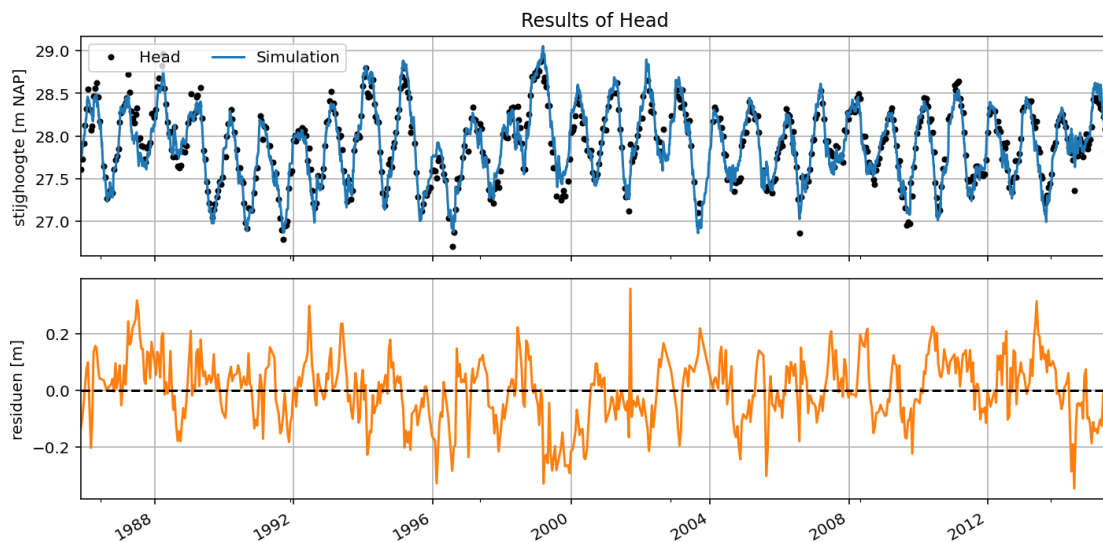
9.4.3 Visuele beoordeling

Een 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 modelresiduen 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 benaderd. 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 stijghoogte 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 afbeelding is te zien dat de simulatie en de metingen goed overeenkomen. 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. De residuen laten echter geen duidelijk patroon zien en zijn netjes verdeeld om de nul-lijn, dus er is geen sprake van een structurele fout in het model.

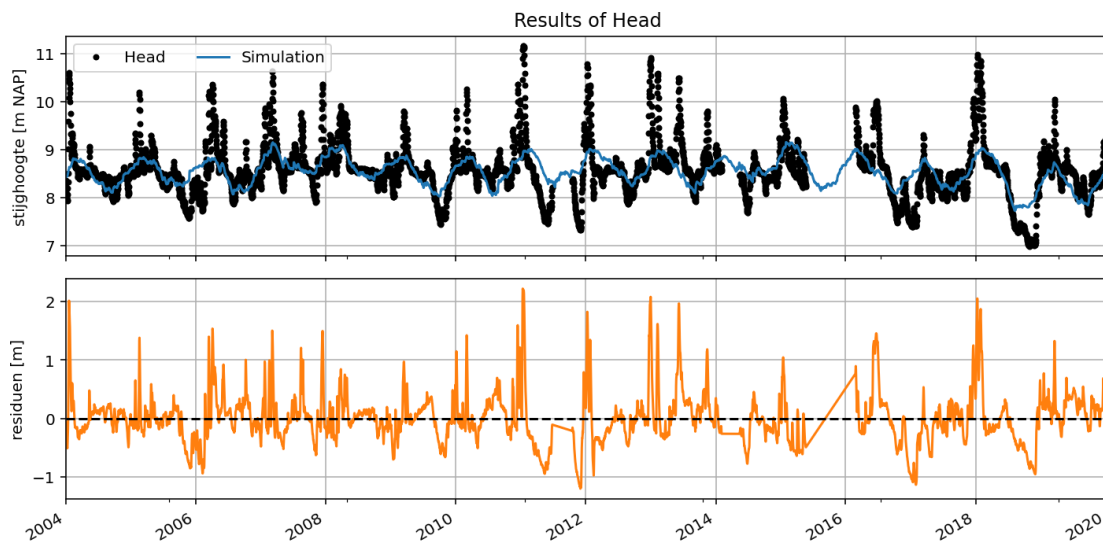
Afbeelding 9.43 Resultaten tijdreeksanalyse met residuen



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 tot wel 2 meter af van de gemeten stijghoogtes. Met dit model is het niet mogelijk om betrouwbare simulaties te maken.

Afbeelding 9.44 Resultaten tijdreeksanalyse



9.4.4 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 stijgthoogtes 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. In onderstaande tabel zijn een aantal veelgebruikte statistieken opgenomen. Voor een overzicht van verschillende fit-statistieken die vaak worden toegepast in de hydrologie zie Jackson et al. (2019).

Het wordt aanbevolen meerdere fit-statistieken tegelijk te beschouwen voor de beoordeling van tijdreeksmodellen. Elke statistiek heeft namelijk zijn eigen specifieke toepassing en geen enkel getal kan alle complexiteit met betrekking tot de fit van een model in een enkel getal uitdrukken.

Afbeelding 9.45 Overzicht fit-statistieken

afkorting	naam	formule	optimale waarde	min	max	type
SSE	som kwadratische fout	$\sum_{i=1}^N (y_i - \hat{y}_i)^2$	0	0	∞	residuen
MAE	gemiddelde absolute fout	$\sum_{i=1}^N \frac{ y_i - \hat{y}_i }{N}$	0	0	∞	residuen
RMSE	wortel kwadratisch gemiddelde fout	$\sqrt{\sum_{i=1}^N \frac{(y_i - \hat{y}_i)^2}{N}}$	0	0	∞	residuen
EVI**	verklaarde variantie	$\frac{\sigma_y^2 - \sigma_r^2}{\sigma_y^2} \cdot 100$	100%	$-\infty^*$	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	$-\infty$	1	goodness-of-fit

- * In verschillende tijdreeksanalyseprogramma's, waaronder Menyanthes en Pastas, wordt de minimale waarde van de EVP begrenst op 0.
- ** Von Asmuth (2012)
- *** De determinatiecoëfficiënt wordt in hydrologische modellen ook vaak de Nash Sutcliffe model efficiency (NSE) coëfficiënt genoemd.

In de formules hierboven is N het aantal metingen, y_i en \hat{y}_i zijn de gemeten en met het model gefitte stijghoogte op tijdstip i, respectievelijk. \bar{y}_i is het gemiddelde van de metingen. σ_y^2 en σ_r^2 zijn de variantie in de gemeten stijghoogtes en de variantie van de residuën r, waarbij $\sigma_y^2 = \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2$ en $\sigma_r^2 = \frac{1}{N} \sum_{i=1}^N (r_i - \bar{r})^2$

Verskil MAE en RMSE

De 'Mean Absolute Error' (MAE) (in het Nederlands: gemiddelde absolute fout) is de het gemiddelde van de absolute waarden van alle fouten. Hiermee wordt elke fout even zwaar gewogen. De 'Root Mean Squared Error' (RMSE) (in het Nederlands: wortel kwadratisch gemiddelde fout) worden de fouten gekwadrateerd waardoor grote fouten een relatief zwaardere weging krijgen.

Verskil R² en EVP

De verklaarde variantie (EVP) en de determinatiecoëfficiënt (R²) zijn subtiel verschillend. De waarde van de EVP is gelijk aan R² als het gemiddelde van de residuen gelijk is aan nul. De optimalisatie van een tijdreeksmodel probeert het model zo de fitten dat het gemiddelde van de residuen gelijk is aan nul, dus vaak hebben deze statistieken dezelfde waarde.

Voorbeeld fit statistiek 1

De fit statistieken van het eerste, hierboven weergegeven model staan in de tabel hieronder. De gemiddelde absolute modelfout bedraagt circa 9 cm. De fit-statistieken geven aan dat het model goed overeenkomt met de metingen; de verklaarde variantie (of EVP) is 92.9 %.

Tabel 9.5 Fit-statistieken tijdreeksmodel

	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)	8.442	m ²
Explained variance percentage (evp)	92.905	%
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 goed overeen met de metingen. De verklaarde variantie (EVP) waarde is slechts 39,6 % en de gemiddelde, absolute afwijking (MAE: mean absolute error) is maar liefst 31 cm.

Tabel 9.6 Fit-statistieken tweede tijdreeksmodel

	Waarde	Eenheid
Mean absolute error (mae)	0.314	m
Root mean squared error (rmse)	0.452	m
Sum of the squares of the error (sse)	1097.701	m ²
Explained variance percentage (evp)	36.210	%
R-squared (rsq)	0.362	-

Overfitten

Als er meer verklarende variabelen (en daarmee modelparameters) aan een tijdreeksmodel worden toegevoegd, wordt de fit in principe beter, zelfs als de verklarende variabele weinig of geen invloed heeft op de stijghoogte. 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 standaardfouten van de parameters (zie Modelonzekerheid) (Van Geer, 2012). Er zijn statistieken die een relatieve score berekenen op basis van de fit van het model waarbij gestraft wordt voor het aantal parameters. Hiermee kan worden ingeschat of het toevoegen van parameters tot een significante verbetering heeft geleid van het model. De Akaike Information Criterion (AIC) en de Bayesian Information Criterion (BIC) zijn twee voorbeelden daarvan. Zowel de AIC als de BIC zijn een maat voor de relatieve hoeveelheid informatie die verloren is gegaan in het model (zogenaamde 'penalized likelihood criteria'). Het zijn dus relatieve waardes en worden gebruikt om verschillende modellen voor dezelfde meetreeks met elkaar te vergelijken. Als de AIC en BIC lager worden na het toevoegen van een verklarende variabele betekent dit dat het model verbetert (er gaat relatief minder informatie verloren). Als de AIC en BIC hoger worden dan heeft het toevoegen van de parameters niet voldoende bijgedragen aan een verbetering van het model. Voor deze paragraaf is het alleen van belang te weten dat het verschil tussen de AIC en BIC zit in de manier waarop er gestraft wordt voor het aantal modelparameters. De BIC straft over het algemeen iets zwaarder voor het toevoegen van parameters.

Samenvattend kunnen deze statistieken gebruikt worden om verschillende tijdreeksmodellen voor hetzelfde meetpunt (met andere modelstructuren) onderling te vergelijken. Deze statistieken geven een relatieve score voor de fit (hoe goed het model op de metingen past) waarbij er gestraft wordt voor het aantal modelparameters. Hoe lager de statistiek, hoe minder informatie in het model verloren is gegaan, en dus is het model mogelijk beter.

Voorbeeld AIC/BIC

Voor dit voorbeeld maken we een tijdreeksmodel voor een meetpunt (40CP0393 (filter 4)) waarbij neerslag en verdamping lineair gecombineerd worden als schatting voor de grondwateraanvulling. We optimaliseren het model en berekenen de EVP, AIC en BIC. Vervolgens voegen we de invloed van de rivierwaterstand toe aan het tijdreeksmodel. Eerst kiezen we daarvoor de exponentiële responsfunctie (2 parameters), vervolgens doen we dat nog eens maar dan met de Gamma responsfunctie (3 parameters). Voor beide modellen berekenen we weer de EVP, AIC en BIC. Uit de resulterende tabel is op te maken dat het toevoegen van de rivier een enorme verbetering oplevert. De EVP wordt aanzienlijk hoger, en de AIC en BIC aanzienlijk lager. Nu willen we weten welke van de twee modellen met rivier beter is.

Op basis van de EVP zouden we concluderen dat het model met rivier en de Gamma-responsfunctie het beste is. De AIC is ook het laagste voor dit model, dus blijktbaar neemt volgens deze statistiek de kwaliteit van de fit voldoende toe om de toevoeging van een extra parameter te rechtvaardigen. Op basis van de AIC zouden we wederom dit model als beste selecteren. De BIC (die iets zwaarder straft voor extra parameters) is het laagste voor het model met rivier en de exponentiële respons, dus volgens deze statistiek is dat het beste model. Helaas geven de statistieken in dit geval geen eenduidig antwoord, dus is het aan de gebruiker om een keuze te maken. Gezien het geringe verschil in zowel de AIC en BIC tussen de tijdreeksmodellen met rivier zijn beide modellen voldoende goed om verdere analyses mee te doen.

Tabel 9.7 Overzicht statistieken tijdreeksmodel

	EVP	AIC	BIC
model zonder rivier	50.475	926.498	964.277
model met rivier (Exponential)	93.783	445.418	495.790
model met rivier (Gamma)	94.162	444.027	500.696

9.4.5 Testen model

Het doel van deze stap is om te controleren of het model geschikt is voor het beoogde doel. Deze stap wordt ook vaak model-validatie genoemd (voor een betoog over de juiste terminologie, zie bijvoorbeeld Konikow et al. (1992)). Cross-validatie (in het Engels: 'split-sample testing') is een veel toegepaste methode in de neerslag-afvoer hydrologie om modellen te testen. Het is een methode om de voorspellende waarde van een model te beoordelen. Bij deze methode splitsen we een tijdreeks in twee delen, een trainings-deel en een test-deel. Het deel van de metingen dat we gebruiken om het model te kalibreren heet de trainings-set. Vervolgens gebruiken we het gekalibreerde model om een simulatie te berekenen voor de tweede periode. Dit deel noemen we de test-set. De test set vergelijken we 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 in een trainings-deel en een test-deel.

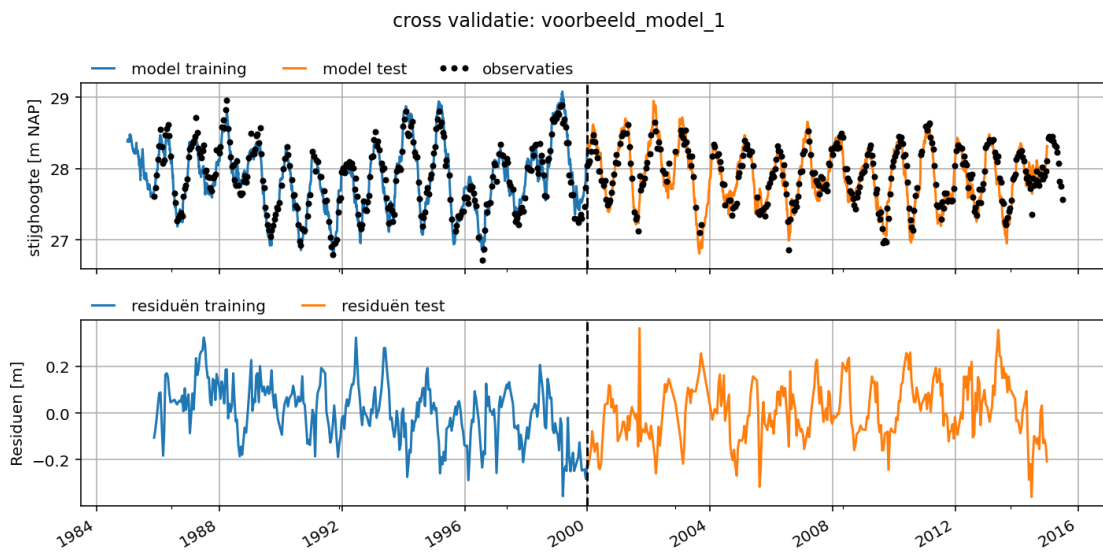
De hierboven besproken fit statistieken kunnen ook gebruikt worden om de voorspelling in een testperiode te toetsen. Voor een uitgebreide analyse van een set aan tijdreeksmodellen met deze methode, inclusief een beschouwing van bijbehorende fit-statistieken, zie Knotters (2012).

Voorbeeld model test 1

Het voorbeeld model van hierboven is getest door de dataset op te splitsen in twee periodes. 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 voor 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.

Afbeelding 9.46 Cross validatie tijdreeksmodel 1



Tabel 9.8 Fit-statistieken cross validatie

	Training: 1985-2000	Test: 2000-2015	Eenheid
Mean absolute error (mae)	0.098	0.098	m
Root mean squared error (rmse)	0.123	0.122	m
Sum of squares of the error (sse)	4.781	4.744	m^2
Explained variance percentage (evp)	93.345	89.498	%
R-squared (rsq)	0.933	0.895	-

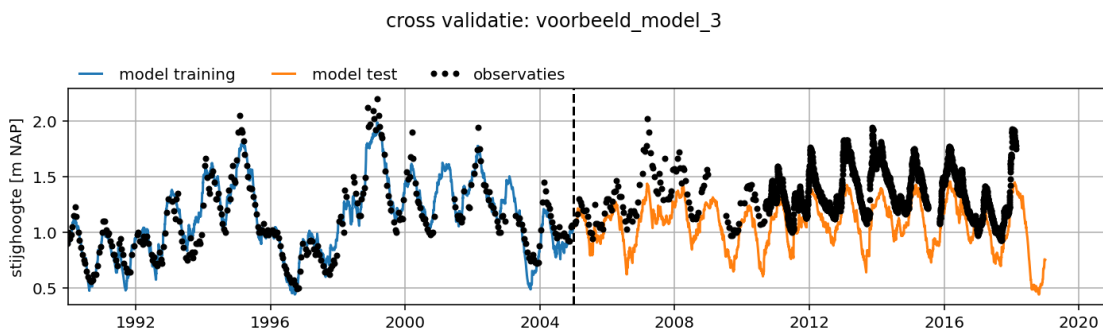
Als het model goed presteert in de test-periode dan is het model dat is afgeleid op basis van de trainingsperiode goed genoeg om de stijghoogtes te voorspellen. Dat geeft vertrouwen dat het model goed presteert, maar dat betekent niet per definitie dat het model ook in de toekomst goede voorspellingen zal geven. Als er in de toekomst iets structureel verandert 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 trainingsperiode maar niet in de test-periode, dan is het model geen goede voorspeller van de stijghoogte. 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, of overfitten.

Voorbeeld model test 2

Hieronder een voorbeeld van een model met een goede fit in de trainingsperiode en een slechte (minder goede) fit in de testperiode. De slechte fit wordt veroorzaakt door een tot nu toe onverklaarde trend in de stijghoogte. Deze trend wordt mogelijk veroorzaakt door een externe invloed die niet is meegenomen bij het trainen van het model. Het is daarom niet mogelijk om met het model een goede fit te krijgen voor de testperiode.

Afbeelding 9.47 Cross validatie tijdreeksmodel 2



Tabel 9.9 Fit-statistieken tijdreeksmodel

	Training: 1990-2005	Test: 2005-2019	Eenheid
Mean absolute error (mae)	0.083	0.245	m
Root mean squared error (rmse)	0.105	0.266	m
Sum of squares of the error (sse)	3.545	190.171	m^2
Explained variance percentage (evp)	90.738	76.882	%
R-squared (rsq)	0.907	-0.506	-

9.4.6 Modelonzekerheid

Elk tijdreeksmodel kent een bepaalde onzekerheid. Deze onzekerheid kan opgesplitst worden in onzekerheid over de modelstructuur en onzekerheid in de modelparameters. In deze paragraaf wordt alleen de onzekerheid van de modelparameters besproken, maar er is een relatie tussen deze twee vormen van modelonzekerheid. Een grote onzekerheid in de parameters kan duiden op een suboptimale modelstructuur, en andersom zal een suboptimale modelstructuur zich vaak uiten in een grotere onzekerheid in de modelparameters. Voor een toelichting op de modelstructuur van tijdreeksmodellen wordt verwezen naar paragraaf 9.2.

In tijdreeksmodellering volgens de PIRFICT methode worden externe invloeden gemodelleerd met responsfuncties die ieder één of meerdere parameters bevatten (Asmuth, 2002). 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 reeks die geminimaliseerd wordt (de residuen of de ruis) voldoen aan bepaalde statistische voorwaarden. Vaak wordt een ruismodel toegepast om ervoor te zorgen dat het model aan deze statistische eisen voldoet. De voorwaarden voor de residuen/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, wat betekent dat de standaardafwijking van de ruis constant is;
 - de ruis voldoet aan een, bij voorkeur, normale statistische verdeling.
- niet gecorreleerd is aan een verklarende tijdreeks.

Er wordt in deze paragraaf niet uitgelegd wat het gevolg is als niet aan deze voorwaarden wordt voldaan; zie daarvoor paragraaf 9.3. Er wordt wel gepresenteerd hoe deze voorwaarden visueel beoordeeld kunnen worden. Naast een visuele beoordeling bestaan er verschillende 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. 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 standaardafwijking van de modelparameters is berekend voor het eerste voorbeeldmodel 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 geschat kunnen worden in de optimalisatie. Maar we weten nog niet of we deze berekende standaardafwijking voor de verschillende parameters kunnen gebruiken om iets te zeggen over de modelonzekerheid. We hebben namelijk nog niet beschouwd of het model voldoet aan de hierboven genoemde statistische voorwaarden.

Tabel 9.10 Betrouwbaarheid modelparameters

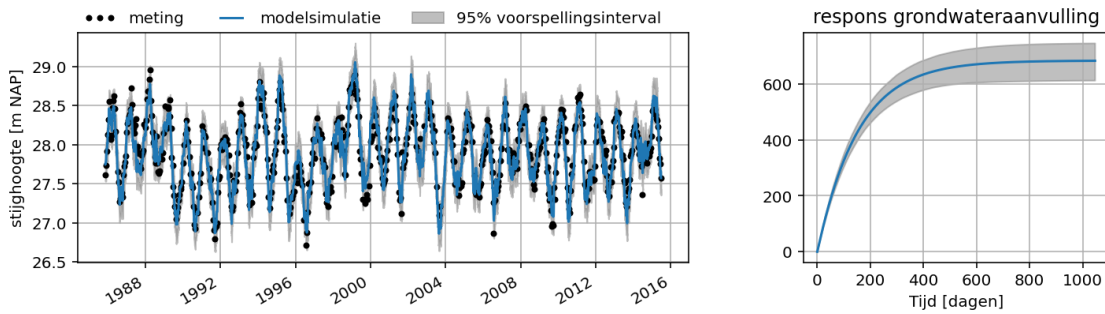
Parameter	Optimale waarde	Standaardafwijking	Standaardafwijking (%)
recharge_A	682.473	35.735	±5.24 %
recharge_n	1.018	0.018	±1.78 %
recharge_a	150.382	11.252	±7.48 %
recharge_f	-1.271	0.061	±4.77 %
constant_d	27.882	0.068	±0.24 %
noise_alpha	50.095	5.960	±11.90 %

De berekende standaardafwijking in de parameters kan gebruikt worden om een betrouwbaarheidsinterval te berekenen. Dit interval zegt iets over de mogelijke model uitkomsten gegeven de onzekerheid in de parameters. Dit interval wordt bijvoorbeeld berekend door N trekkingen te doen uit de berekende parameter-verdelingen en daarmee de stijghoogte met het model te simuleren. Het betrouwbaarheidsinterval wordt op basis van die N simulaties voor een bepaald niveau (α) berekend (bijvoorbeeld $\alpha=0.05$, het 95 %-betrouwbaarheidsinterval).

In onderstaande figuur is in de rechter grafiek het 95 %-betrouwbaarheidsinterval berekend voor de respons op grondwateraanvulling voor voorbeeldmodel 1. In de linker grafiek is het voorspellingsinterval weergegeven met een grijze band rondom de modelsimulatie van de stijghoogte. Het voorspellingsinterval is de som van het betrouwbaarheidsinterval (de onzekerheid als gevolg van de onzekerheid in de parameters) plus de standaardafwijking van de residuen (de onzekerheid als gevolg van de modelfout).

Deze bandbreedtes kunnen dus waardevol zijn in de beoordeling van het model, of zelfs in een toepassing van het model, maar daarvoor moet de modelruis wel aan de eerder genoemde statistische voorwaarden voldoen. In de volgende paragrafen wordt onderzocht of het model aan die voorwaarden voldoet.

Afbeelding 9.48 modelsimulatie met voorspellingsinterval



Voorwaarde 1: Autocorrelatie

Autocorrelatie is de correlatie van een tijdreeks met een in de tijd verschoven versie van zichzelf. Als er nog significante autocorrelatie aanwezig is in de ruis, dan wordt het betrouwbaarheidsinterval van de parameter te klein geschat (zie de paragraaf 9.3 voor een voorbeeld). Deze vorm van correlatie kan worden weergegeven in een autocorrelatiediagram. Daarin is op de horizontale as de verschuiving in de tijd weergegeven (in het Engels: 'lag') en op de verticale as de berekende correlatie.

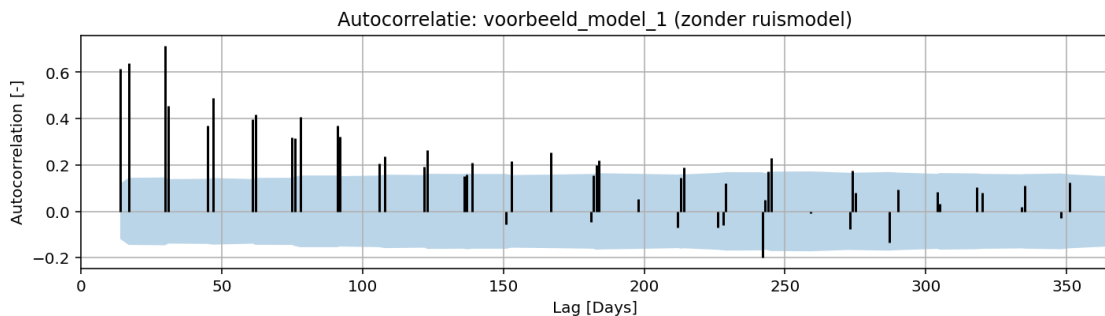
De verschuivingen in de tijd (lags) moeten overeenkomen met de tijdstap van het model. Als er op dagbasis gerekend wordt, moet de reeks ook verschoven worden met (veelvouden van) die tijdstap voor het bepalen en beoordelen van de autocorrelatie (Collenteur, 2018).

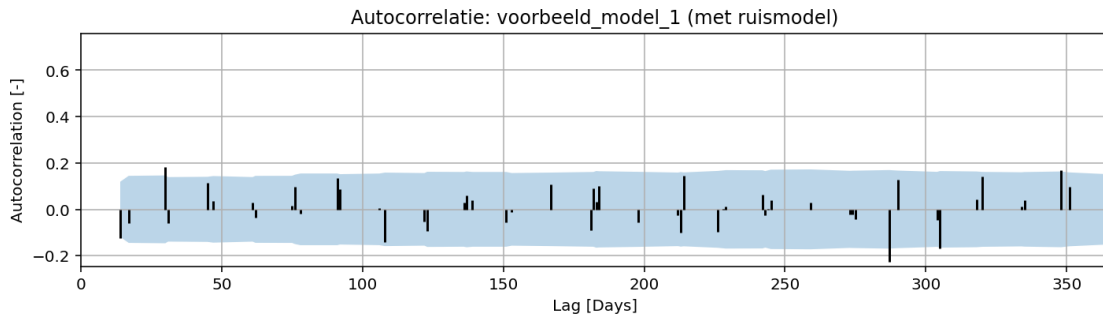
Voorbeeld autocorrelatiediagram

De autocorrelatie is voor het voorbeeld model berekend met en zonder een ruismodel. De resultaten zijn weergegeven in onderstaande afbeeldingen. De maximale, berekende 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 van de berekende autocorrelatie. Als de autocorrelatie van een bepaalde lag binnen het 95% betrouwbaarheidsinterval valt, dan kan met 95% zekerheid gezegd worden dat de autocorrelatie voor die lag nul is. Het is duidelijk te zien dat de residuen van het model zonder ruismodel niet voldoen aan de eis voor de autocorrelatie. Het berekende betrouwbaarheidsinterval voor de parameters zal dus niet goed geschat zijn. In het model met ruismodel zien we dat er geen significante autocorrelatie aanwezig is in de ruis.

Afbeelding 9.49 Autocorrelatie met en zonder ruismodel





Voorwaarde 2: Normale verdeling ruis

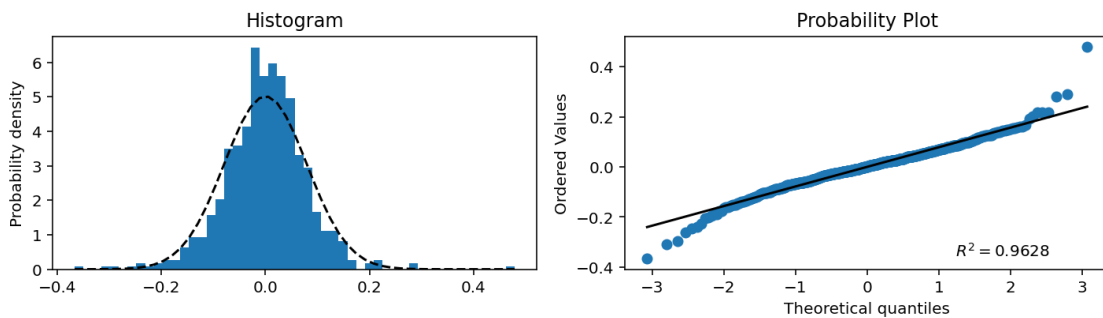
Een andere voorwaarde voor witte ruis is dat deze voldoet aan een, bij voorkeur, normale statistische verdeling. De verdeling van de ruis kan weergegeven worden met een histogram in combinatie met een normale kansverdeling. Visueel kan dit beoordeeld worden door een histogram te maken van de residuen/ruis 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 geplot van de ruis en wordt deze vergeleken met de normale verdeling. Dit is in de linker grafiek weergegeven. In de rechter grafiek is 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 wordt ingeschat dat de verdeling van de ruis de normale verdeling voldoende goed benaderd om het model op dit aspect goed te keuren.

Afbeelding 9.50 Visualisatie normale verdeling ruis



Voorwaarde 3: Homoscedasticiteit

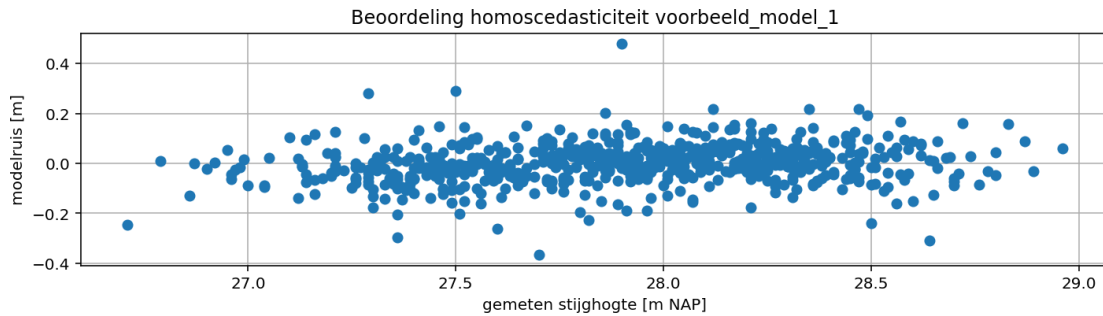
De derde statistische voorwaarde voor witte ruis is dat de ruis homoscedastisch moet zijn. Een tijdreeks is homoscedastisch als de variantie onafhankelijk is. Eén van de mogelijkheden om hier naar te kijken is via de absolute waarde van de reeks. Als de ruis niet homoscedastisch is (oftewel heteroscedastisch) dan is de variantie van de ruis afhankelijk van de stijghoogte. Dit kan visueel beoordeeld worden door de ruis uit te zetten tegen de stijghoogtemetingen.

Als de spreiding van de ruis niet gelijk is over het gehele bereik van de stijghoogtemetingen dan is de ruis heteroscedastisch. Dit betekent nog niet zonder meer dat de ruis homoscedastisch is. De variantie van de ruis kan nog op andere manieren gecorreleerd zijn, bijvoorbeeld in de tijd.

Voorbeeld visualisatie homoscedasticiteit

Voor voorbeeldmodel 1 is de ruis uitgezet tegen de metingen. De spreiding in de ruis lijkt redelijk constant te zijn over het bereik van de stijghoogtemetingen. Voor de absolute waarde lijkt er dus sprake te zijn van homoscedastische ruis.

Afbeelding 9.51 Beoordeling homoscedasticiteit



Voorwaarde 4: Correlatie ruis met verklarende tijdreeks

De laatste voorwaarde ten aanzien van de ruis is dat deze niet significant gecorreleerd is met een van de verklarende reeksen in het tijdreeksmodel. Als dit wel het geval is, is de gehanteerde responsfunctie blijkbaar niet in staat om het gedrag van de stijghoogte op basis van de verklarende reeks volledig te beschrijven. In dat geval kan de onzekerheid van de parameters niet nauwkeurig worden bepaald en moet de modelstructuur aangepast worden.

9.4.7 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 om inzicht in het systeem te vergroten.

Hydrologie

Het beoordelen van modellen op basis van hydrologische aspecten kan allerlei vormen aannemen. De vragen daarbij zijn eigenlijk 'werkt het systeem zoals ik denk dat het werkt' en 'is wat ik bereken hydrologisch plausibel?'. Daarbij zijn modellen die wel aan de verwachtingen voldoen net zo interessant als de modellen die daar niet aan voldoen. Denk bijvoorbeeld aan een model met grondwateraanvulling en een onttrekkingsdebiet als verklarende variabelen. Een eerste eenvoudige hydrologische beoordeling is de vraag of de stijghoogte ook toeneemt als het regent, en of deze daalt als er grondwater onttrokken wordt. Is dat niet het geval, dan is de kans groot dat het model niet betrouwbaar is.

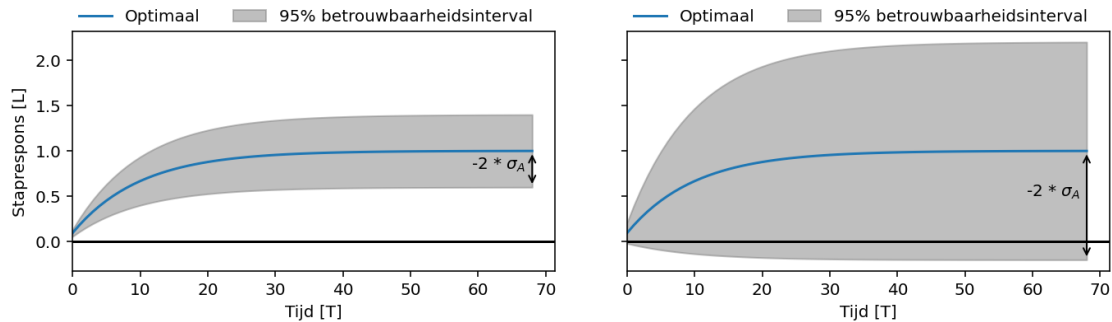
Op basis van de hydrologie kunnen tijdreeksmodellen beoordeeld worden, maar andersom kunnen de modelresultaten ook het hydrologisch inzicht in het systeem vergroten.

Voorbeeld hydrologische beoordeling

Als voorbeeld van een hydrologische beoordeling van een tijdreeksmodel wordt de respons van een onttrekking beschouwd. In onderstaande afbeelding 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. Overigens betekent in dit voorbeeld een positieve staprespons dat de onttrekking een verlaging van de stijghoogte veroorzaakt.

Afbeelding 9.52 Respons met betrouwbaarheidsinterval met een kleine onzekerheid (links) en een grote onzekerheid (rechts)



Vanuit de hydrologie moet de onttrekking voor een daling van de stijghoogte zorgen, maar in het geval van het onzekere model (rechts) is er een meer dan 5 %-kans dat de onttrekking een verhoging van de stijghoogte 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. De onttrekking heeft mogelijk geen significante invloed op de stijghoogte.

Geheugen van de respons

Een responsfunctie zegt onder andere iets over de traagheid waarmee de stijghoogte zich aanpast aan een verandering in een verklarende variabele. Een simpel voorbeeld is bijvoorbeeld hoe lang een verhoging van de stijghoogte aanhoudt na een bui. De tijd tot het effect van een verandering in een verklarende variabele op de stijghoogte verwaarloosbaar klein is wordt vaak het geheugen van de respons genoemd.

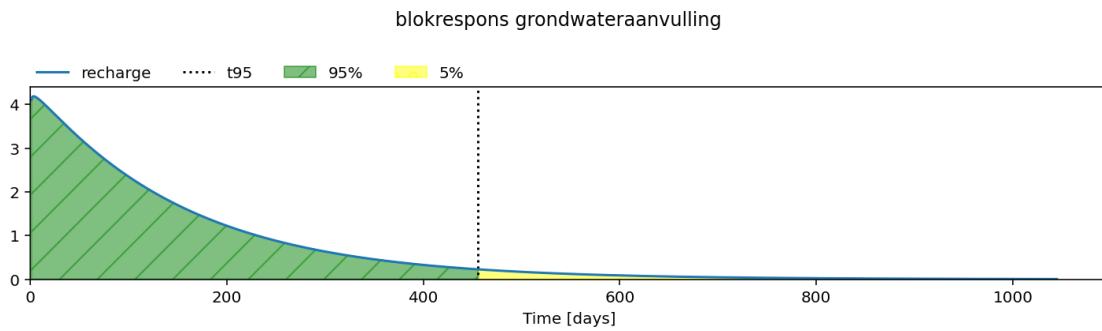
Als het geheugen van de respons langer is dan de lengte van de tijdreeks waarop het tijdreeksmodel gekalibreerd wordt, dan is de tijdreeks eigenlijk niet voldoende lang om een dergelijke trage respons nauwkeurig in te schatten. Dus in dat geval zou het tijdreeksmodel daarop afgekeurd kunnen worden.

Een andere manier om het geheugen van de respons te beoordelen is vanuit de hydrologie. Vanuit de hydrologische systeemkennis is het wellicht mogelijk om grenzen af te leiden ten aanzien van het geheugen van de respons, of door meerdere modellen te vergelijken kan beoordeeld worden of het geheugen van de respons plausibel is. Dit laatste voorbeeld is een vorm van ruimtelijke beoordeling, dat verder op in deze paragraaf nog aan bod komt.

Voorbeeld bepalen geheugen van de responsfunctie

Het geheugen van de responsfunctie kan niet bepaald worden door te kijken waar de respons gelijk is aan nul. De meeste responsfuncties zijn nooit gelijk aan nul maar benaderen nul als $t \rightarrow \infty$. 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.

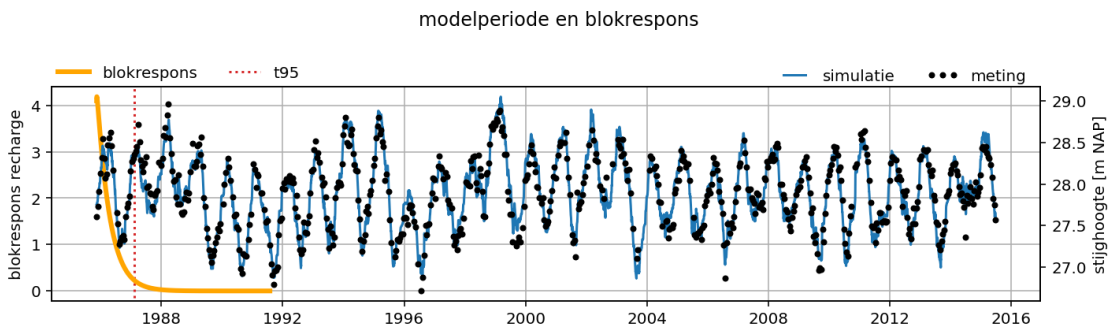
Afbeelding 9.53 Blokrespons met $t_{95\%}=456$ dagen



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. Daarmee kunnen we stellen dat de tijdreeks voldoende lang is om de respons te kunnen inschatten.

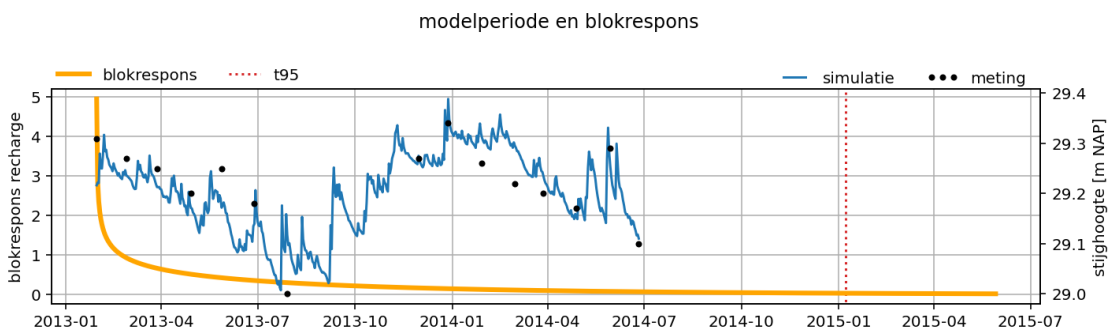
Afbeelding 9.54 Modellsimulatie met blokrespons



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. Als de respons langer is dan de meetreeks, kan op basis van de gegevens niet ingeschat worden of de respons terecht zo lang is. In dit geval kan gesteld worden dat de tijdreeks niet voldoende lang is om een tijdreeksmodel voor af te leiden.

Afbeelding 9.55 modellsimulatie met lange blokrespons



Neerslagoverschot

De invloed van neerslag en verdamping wordt vaak meegenomen als het neerslagoverschot, berekend met de volgende formule:

$$R = P + f \cdot E$$

waarin R het neerslagoverschot is, P de neerslag en E de verdamping. Voor de verdamping wordt in Nederland vaak de referentie-gewasverdamping gebruikt. Dit is een waarde voor de potentiële verdamping van een bepaald gewas. Hoeveel er in werkelijkheid verdampt, de actuele verdamping, is meestal niet bekend. In bovenstaande formule wordt de actuele verdamping berekend door de potentiële verdamping te vermenigvuldigen met de verdampingsfactor f . De verdampingsfactor wordt in een tijdreeksmodel vaak geoptimaliseerd.

De geoptimaliseerde waarde van de verdampingsfactor kan vanuit hydrologisch oogpunt worden beoordeeld. Daarbij kan de volgende betekenis worden gegeven aan de geoptimaliseerde waarde:

- $f < -2$, het neerslagoverschot wordt berekend met een actuele verdamping die meer dan twee maal zo groot is als de referentie-gewasverdamping;
- $-2 < f < 0$, het neerslagoverschot wordt berekend waarbij de verdamping met een bepaalde factor wordt geschaald;
- $f = 0$, de neerslag is voldoende om de stijghoogte te verklaren. De verdamping wordt niet meegenomen in het tijdreeksmodel;
- $f > 0$, de verdampingsfactor is positief, dit zou betekenen dat de verdamping zorgt voor een stijging in de stijghoogte. Dit is hydrologisch niet te verklaren.

Naast het beoordelen van de geoptimaliseerde waarde is het ook mogelijk om grenswaarden op te geven waarbinnen een bepaalde parameter geoptimaliseerd mag worden. Daarmee kunnen fysische randvoorwaarden worden opgelegd aan het model. Voor de verdampingsfactor worden vaak de volgende grenzen gehanteerd:

- $f < 0$, de verdamping moet voor een verlaging van de stijghoogte zorgen. De verdampingsfactor moet dus kleiner zijn dan nul;
- $f > -2$, de actuele verdamping die wordt gebruikt voor het neerslagoverschot mag niet groter zijn dan twee maal de referentie-gewasverdamping. De referentie-gewasverdamping gaat uit van een gewas dat goed is voorzien van water en nutriënten. Het is dus niet logisch dat de actuele verdamping nog eens twee maal zo groot is.

Verdampingsfactor en gewasfactor

In veel grondwatermodellen wordt een gewasfactor gebruikt om de referentiegewasverdamping om te rekenen naar de actuele verdamping. Deze factor wordt vaak vergeleken met de verdampingsfactor in tijdreeksmodellen. Hoewel er zeker overeenkomsten zijn tussen de factoren zijn er ook een paar belangrijke verschillen:

- de verdampingsfactor wordt in een tijdreeksmodel geoptimaliseerd samen met andere parameters. De verdampingsfactor kan gecorreleerd zijn aan andere modelparameters. Zo kan het gebeuren dat een iets te hoge drainagebasis wordt gecompenseerd met een lagere verdampingsfactor. Hierdoor heeft de geoptimaliseerde waarde van de verdampingsfactor alleen betekenis in combinatie met andere modelparameters en is het daarmee niet direct te vergelijken met een fysische gewasfactor;
- de werkelijke verdamping (of transpiratie) van een gewas varieert sterk gedurende het jaar afhankelijk van het seizoen en de (groei-)eigenschappen van het gewas. De gewasfactor is dus niet constant. In veel tijdreeksmodellen wordt de verdampingsfactor wel als constante aangenomen.

Voorbeeld verdampingsfactor

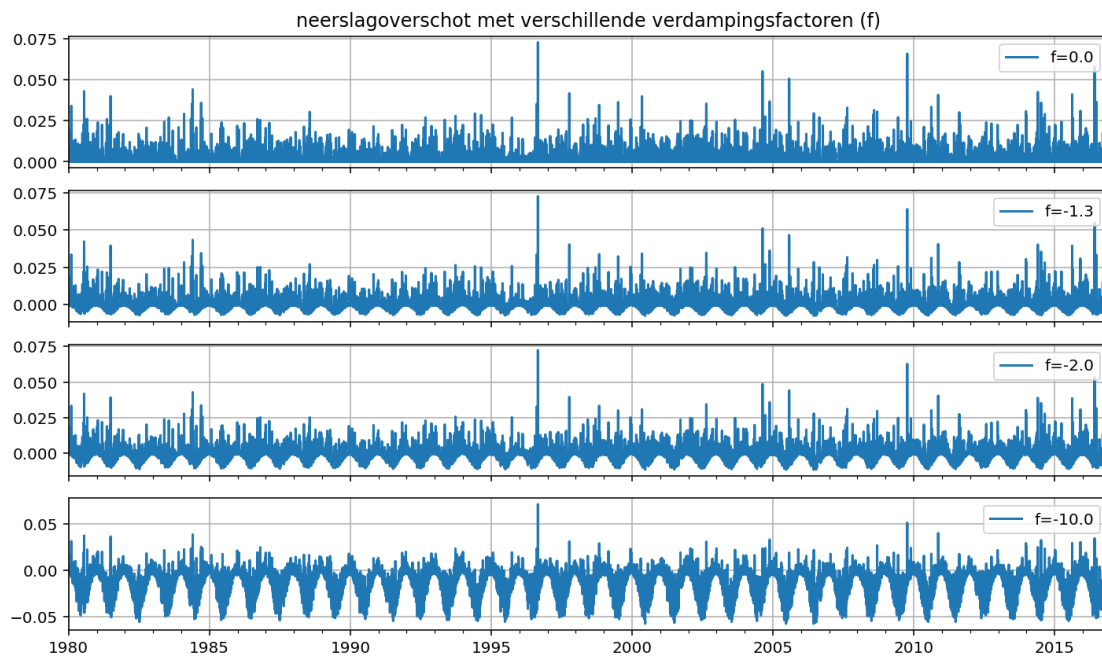
De tijdreeksen uit het basismodel zijn gebruikt om het effect van de verdampingsfactor op de modelresultaten weer te geven. Eerst is het neerslagoverschot weergegeven met verschillende verdampingsfactoren. Daaronder zijn de modelresultaten weergegeven van steeds hetzelfde model welke telkens geoptimaliseerd is met een andere vaste waarde voor de verdampingsfactor.

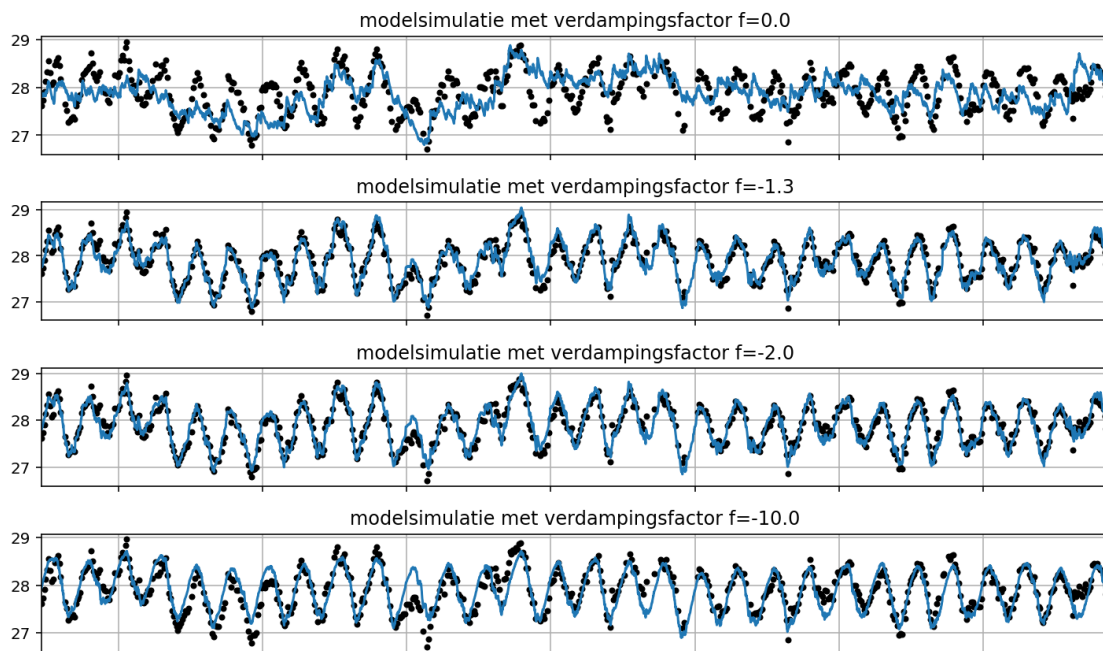
In de grafieken vallen een aantal zaken op:

- bij een verdampingsfactor van $f = 0$ wordt verdamping niet meegenomen, het model wordt alleen geoptimaliseerd o.b.v. de neerslag. De fit tussen gemodelleerde en gesimuleerde stijghoogtes is slecht;
- bij een verdampingsfactor van $f = -1,3$ wordt de beste fit verkregen. Dit is de optimale waarde van verdampingsfactor uit het oorspronkelijke model;
- bij een verdampingsfactor van $f = -2,0$ weegt de verdamping 2x zwaarder dan de neerslag en is de fit van de modelsimulatie met de metingen behoorlijk goed;
- bij een onrealistische verdampingsfactor van $f = -10,0$ weegt de verdamping 10 keer zwaarder dan de neerslag. Omdat de actuele verdamping extreem wordt overschat is het effect van de neerslag bijna niet meer terug te zien in de gesimuleerde reeks.

Tot slot is voor alle modellen ook de geoptimaliseerde drainagebasis getoond. Hierin is de correlatie te zien met de verdampingsfactor. Als in de optimalisatie de verdampingsfactor wordt vastgezet op een lage waarde wordt dit gecompenseerd met een hogere drainagebasis.

Afbeelding 9.56 Neerslagoverschot met verschillende verdampingsfactoren





Parameter correlatie

Bij de optimalisatie van tijdreeksmodellen kan ook worden gekeken naar de correlaties tussen parameters. Wanneer een model parameters bevat die 100 % gecorreleerd zijn dan kan een aanpassing in de ene parameter gecompenseerd worden met een aanpassing aan de gecorreleerde parameter terwijl de fit van het model nagenoeg hetzelfde blijft. Een hoge parameter correlatie is geen probleem voor de betrouwbaarheid van het tijdreeksmodel. In het model wordt immers telkens de combinatie van deze parameters gebruikt om een simulatie uit te voeren. Hoge correlaties zorgen er wel voor dat de absolute waarde van de geoptimaliseerde parameter niet los beschouwd mag worden van de absolute waarde van de gecorreleerde parameter.

Voorbeeld parameter correlaties

De verdampingsfactor en de drainagebasis zijn parameters die vaak sterk gecorreleerd zijn. In het model uit het vorige voorbeeld is de correlatie tussen de verdampingsfactor en de drainagebasis -0.98 . In het vorige voorbeeld zijn de modellen telkens doorgerekend met een andere verdampingsfactor, in de resultaten van dit model kunnen we mooi het effect van de hoge correlatie met de drainagebasis zien. In onderstaande tabel zijn de waarden van de verdampingsfactor, geoptimaliseerde drainagebasis en de EVP weergegeven. Het is duidelijk te zien dat bij een hogere verdampingsfactor de geoptimaliseerde drainagebasis lager ligt. Ook is zichtbaar dat de EVP slechts 1,5 procentpunt lager wordt ten opzichte van het optimale model wanneer de verdampingsfactor op -2.0 wordt gezet. De waarde van de drainagebasis wordt hoger om te compenseren voor deze toename in verdamping. In dit geval kunnen de absolute waarde van de drainagebasis en verdampingsfactor dus niet los van elkaar worden geïnterpreteerd.

Tabel 9.11 Verdampingsfactor vergeleken met drainagebasis

Verdampingsfactor	Drainagebasis (m NAP)	EVP (%)
0.0	25.07	15.76
-1.3	27.88	92.91
-2.0	28.55	91.80
-10.0	30.36	76.76

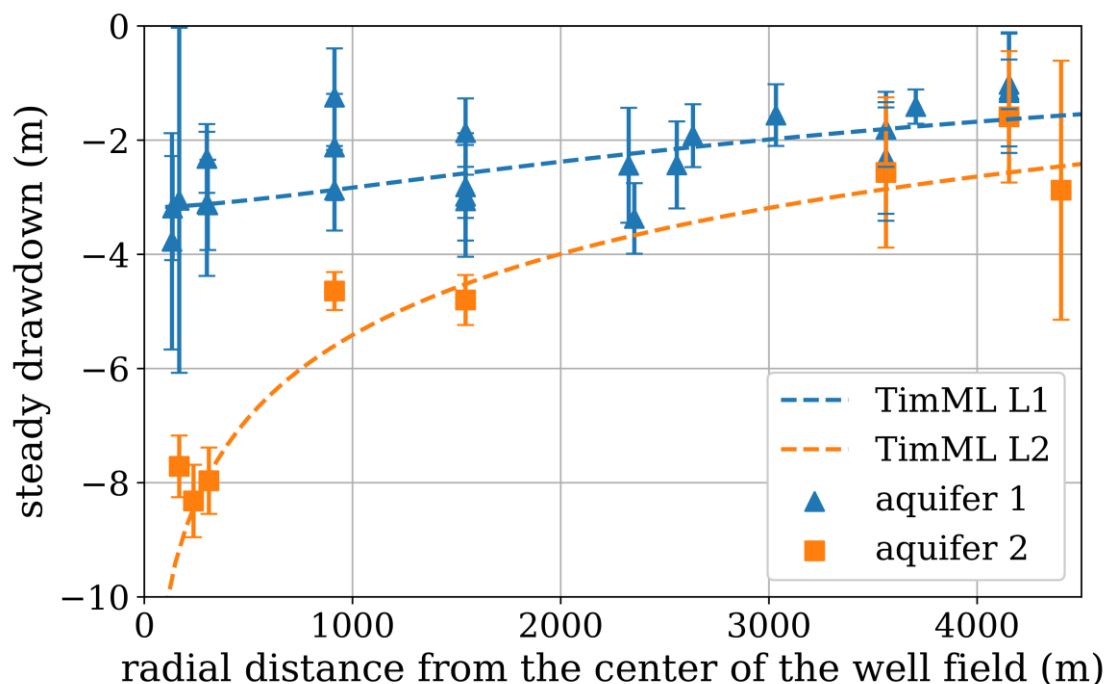
Ruimtelijke beoordeling

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 stijghoogte, en zo zijn er tal van andere voorbeelden te bedenken. In dit soort studies is het nuttig om een ruimtelijke beoordeling te doen van de resultaten. Voor voorbeelden van ruimtelijke analyses, zie bijvoorbeeld Baggelaar (1988), Asmuth (2012), en Berendrecht (2016). Dit aspect wordt 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 Collenteur et al. (2019). In die studie zijn tijdreeksmodellen afgeleid voor peilbuizen binnen een straal van ± 5 km rondom een drinkwaterwinning 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 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 afbeelding is ook de stationaire invloed van de onttrekking als functie van de afstand opgenomen. Deze is berekend met een 2-laags analytisch elementen grondwatermodel (Bakker et al., 2003) en dient als extra controle op de geschatte invloed volgens de tijdreeksmodellen.

Afbeelding 9.57 Berekende stationaire invloed winning Seppe als functie van de afstand bij een onttrekkingsdebiet van $1 \text{ Mm}^3/\text{jaar}$ (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 stijghoogte. 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. Op basis van deze ruimtelijke beoordeling wordt geconcludeerd dat de tijdreeksmodellen toegepast kunnen worden om verdere analyses te doen over de invloed van de winning. Een ruimtelijke hydrologische beschouwing van de resultaten van meerdere tijdreeksmodellen kan waardevol zijn in de beoordeling.

10

CASUSSEN PIRFICT-METHODE

In dit hoofdstuk worden 4 casussen behandeld. Deze casussen zijn voorbeelden van de toepassing van tijdreeksanalyse en weergeven de stappen voor het opstellen van het model zoals beschreven in de hoofdstukken 3, 5 en 6. Daarnaast bouwen deze casussen voor op de voorbeelden en uitwerken zoals te zien in hoofdstuk 9.

Afhankelijk van het doel van de beschreven casus worden de stappen van de tijdreeksanalyse anders doorlopen. De casussen beschrijven realistische voorbeelden van gegevens en situaties die gebruikers tegen kunnen komen, hierbij moet worden opgemerkt dat het hier gaat om voorbeelden, deze kunnen niet zondermeer worden overgenomen.

Om de leesbaarheid van de individuele casussen te bevorderen, bevatten de casussen enige herhaling. De casussen zijn als Jupyter Notebooks beschikbaar op '[LINK casus notebooks](#)'. In deze notebooks kan de gebruiker de achterliggende modelcode zien en zelf uitvoeren.

10.1 Casus: Invloeden bepalen

Deze casus laat zien hoe tijdreeksanalyse gebruikt kan worden om verschillende invloeden op de stijghoogte te bepalen.

Hierbij wordt gekeken naar het bepalen of een externe invloed, zoals een onttrekking, een significante invloed heeft en hoe groot deze invloed is. Deze casus is opgedeeld in twee delen. In het eerste deel wordt een voorbeeld behandeld met een onttrekking. In deel II van de casus wordt gekeken naar de invloed van een waterloop op de stijghoogte.

10.1.1 Deel I: Invloed peilopzet

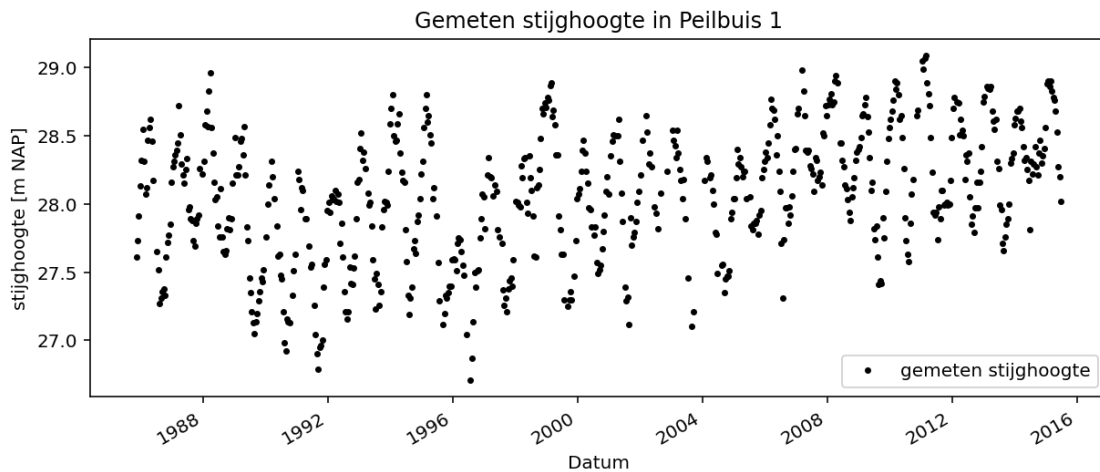
Het waterschap heeft de stijghoogte gemeten in Peilbuis 1. De peilbuis bevindt zich in een gebied waar het waterschap in de zomer van 2005 maatregelen heeft genomen om de stijghoogte in het gebied te verhogen om de natuurwaarden te verbeteren. Hiervoor zijn onder andere enkele stuwen omhoog gezet en enkele watergangen gedempt.

10 jaar na de ingrepen wil het waterschap evalueren of de ingrepen het gewenste effect hebben gehad op de stijghoogte. Hiervoor wordt de stijghoogte in Peilbuis 1 geanalyseerd met behulp van tijdreeksanalyse.

Beschikbare data voorbeeld 1

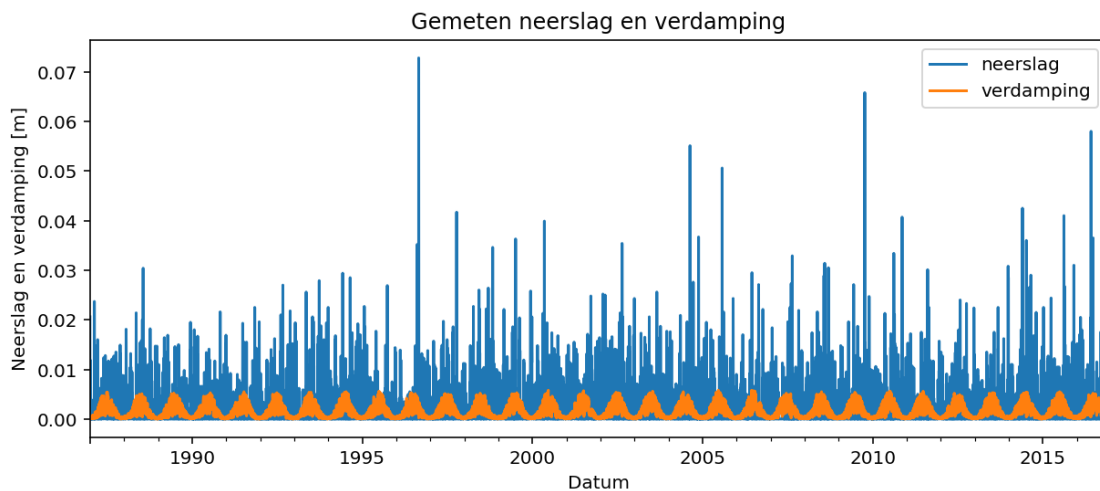
In de onderstaande afbeelding is de gemeten stijghoogte in Peilbuis 1 te zien. De stijghoogte is gemeten voor de periode van 1985 tot en met 2015. De stijghoogte is voor en na de ingreep gemeten. Op het eerste gezicht lijkt het alsof de stijghoogte sinds de ingrepen in de zomer van 2005 hoger is geworden.

Afbeelding 10.1 Gemeten stijghoogte peilbuis 1



De neerslag en verdamping ter plaatse van Peilbuis 1 worden gebruikt als verklarende reeksen. Hiervoor zijn de gegevens gebruikt van het dichtstbijzijnde weerstation van het KNMI. De neerslag en verdamping zijn te zien in de onderstaande afbeelding.

Afbeelding 10.2 Gemeten neerslag en verdamping



Opstellen tijdreeksmodel voorbeeld 1

Er wordt een model opgesteld voor de stijghoogte observaties. Hiervoor wordt de volledige reeks gebruikt. In de reeks zijn geen uitschieters gevonden, er is daarom geen reden om de meetreeks voor te bewerken voordat deze in het model kan worden toegepast.

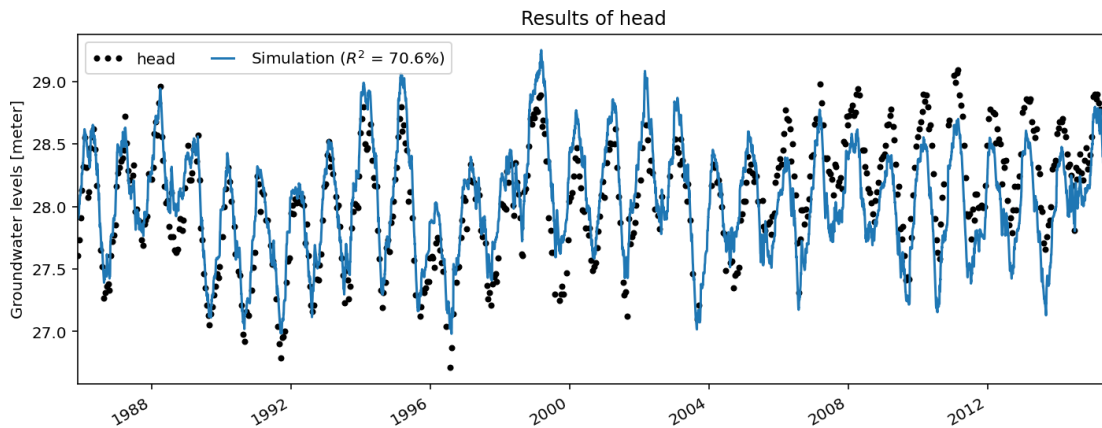
De neerslag en verdamping worden gebruikt als verklarende reeksen. Voor de verklarende reeksen wordt een respons gekozen. De respons functie beschrijft de reactie van de stijghoogte op een externe invloed. Deze moet worden vastgesteld voor elke verklarende reeks, waarbij de gebruiker het type respons functie kiest en waarbij de parameters van deze functie worden geoptimaliseerd. Voor het tijdreeksmodel wordt gekozen voor de Gamma respons functie voor de neerslag en verdamping.

In het tijdreeksmodel wordt gekozen om dezelfde response functie te gebruiken voor de neerslag en de verdamping. Hierbij wordt het verband tussen de neerslag en verdamping beschreven met de formule $R = P - f \cdot E$ (zie ook paragraaf 9.2), waarbij R de grondwateraanvulling is [m], P de neerslag [m], f de verdampingsfactor [-] en E de verdamping [m]. De verdampingsfactor wordt door het tijdreeksmodel geoptimaliseerd. Naast de verklarende reeksen wordt ook een constante meegefit in het model. Na het kiezen van de respons functies voor de verklarende reeksen kan het tijdreeksmodel worden geoptimaliseerd.

Het model heeft een R^2 van 0,71. Met behulp van het geoptimaliseerde tijdreeksmodel kan de stijghoogte worden gesimuleerd. De simulatie van de stijghoogte is te zien in de onderstaande afbeelding.

In de afbeelding zijn de metingen en de simulatie visueel weergegeven. Daaronder zijn de model residuen en ruis weergegeven. Voor de verschillende verklarende reeksen zijn de bijdrages te zien samen met de respons functies van deze reeksen. Dit zijn de Gamma response functies zoals geoptimaliseerd door het tijdreeksmodel. Daarnaast is aan de afbeelding een tabel toegevoegd met daarin de geoptimaliseerde parameters met de bijbehorende standaarddeviaties. Hierin zijn bijvoorbeeld ook de geoptimaliseerde constante en verdampingsfactor weergegeven.

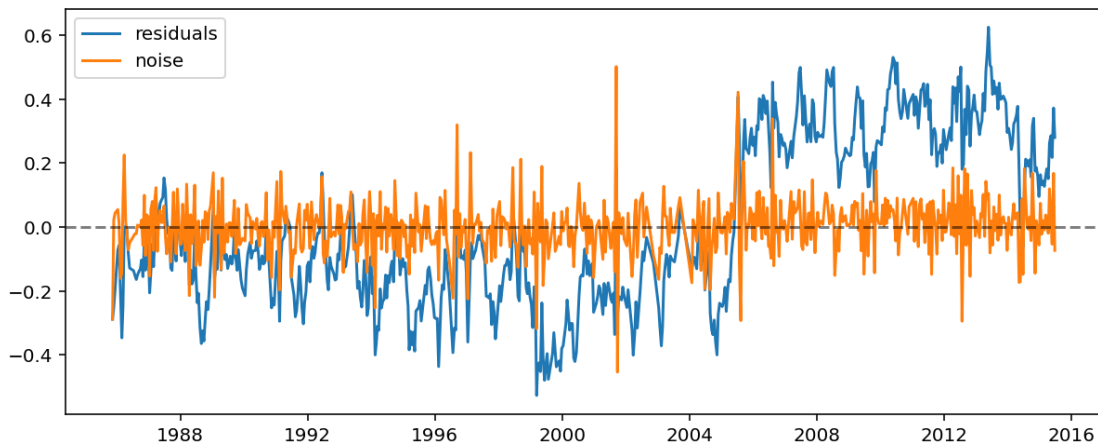
Afbeelding 10.3 Resultaten simulatie



In de simulatie is te zien dat het tijdreeksmodel de gemeten stijghoogte redelijk simuleert. Op basis van een visuele inspectie van deze simulatie is te zien dat voor de periode na 2005 de stijghoogte onderschat wordt door het model.

Dit is beter zichtbaar in de residuen van de reeks. Deze zijn weergegeven in de onderstaande afbeelding. In de periode voor 2005 overschat het model de gemeten stijghoogte (negatieve residuen). In de periode na 2005 maakt het model een onderschatting (positieve residuen). Dit suggereert dat rond 2005 een verandering heeft plaatsgevonden.

Afbeelding 10.4 Residuen en ruis tijdreeksmodel



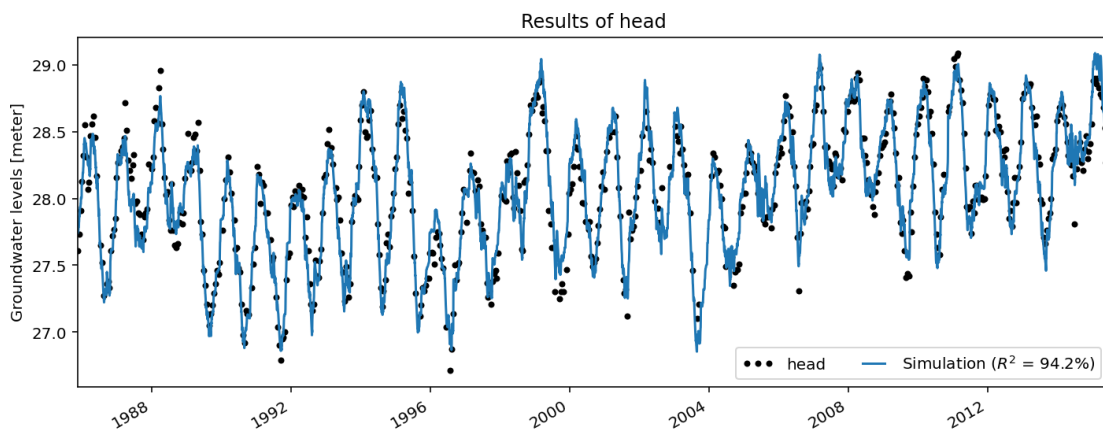
Invloed van de ingreep

Het tijdreeksmodel wordt nogmaals opgesteld, met als doel de invloed van de ingreep te bepalen. Ten eerste wordt geprobeerd het model op te stellen met de toevoeging van een stap respons, dit is een instantane sprong van de stijghoogte. Dit kan worden gezien als een verklarende reeks die gelijk is aan 0 (geen effect), voor 1 juni 2005 en gelijk is aan 1 (wel een effect), na 1 juni 2005.

Deze verklarende reeks wordt gebruikt bij het opstellen van het tijdreeksmodel. De stap respons (of sprong) wordt opgelegd op het moment waarop de ingrepen in het gebied zijn ingevoerd, 1 juni 2005.

Het nieuwe model heeft een R^2 van 0.94, dit is significant hoger dan het oude model zonder stap trend met een R^2 van 0.71. De simulatie van het nieuwe tijdreeksmodel zijn te zien in de onderstaande afbeelding.

Afbeelding 10.5 Simulatie tijdreeksmodel met staptrend



Voor het nieuwe model is wenselijk om iets te zeggen over de hoogte van de stap en de onzekerheid daarvan. Om dit te doen moet worden gecontroleerd of de resultaten van het tijdreeksmodel voldoende zijn om de onzekerheid goed in te kunnen schatten. Hiervoor moet het model aan voorwaarden.

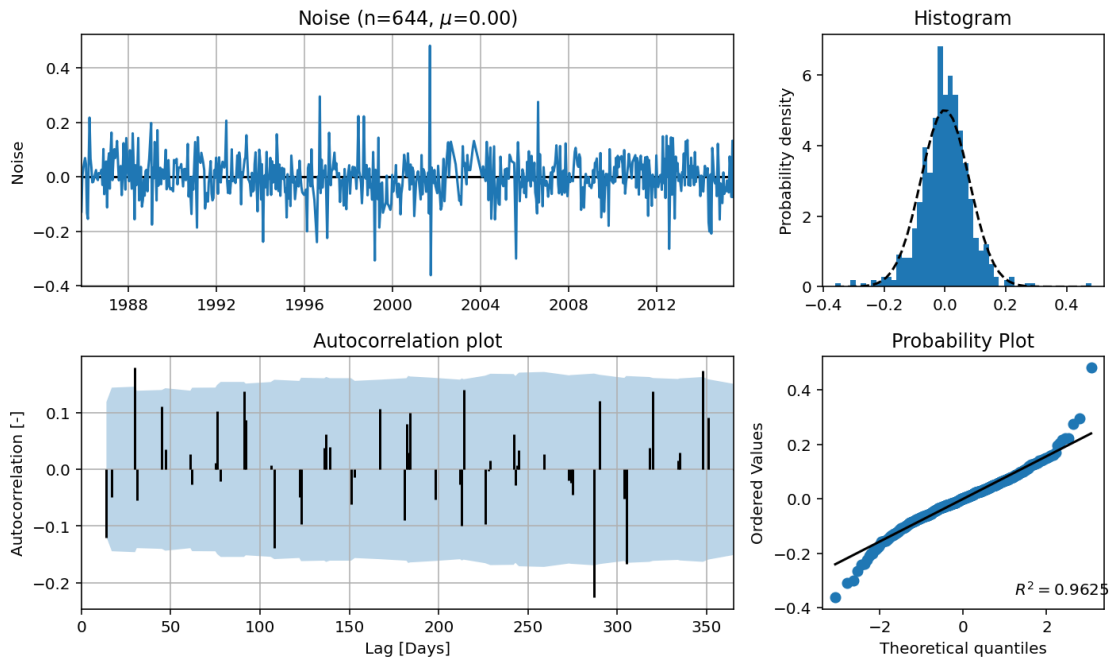
Bij het bepalen van de onzekerheid van de modelparameters is de aanname dat de ruis van het model zogenaamde 'witte ruis' is. Hiervoor worden de volgende punten gecontroleerd:

- het gemiddelde van de ruis is nul;
- de waarden zijn onafhankelijk van elkaar;

- de ruis is normaal verdeeld;
- de ruis heeft een constante variantie.

Als de ruis aan deze toetsen voldoet mag aangenomen worden dat de ruis 'witte ruis' betreft en dat de standaard deviatie van de modelparameters juist is ingeschat.

Afbeelding 10.6 Diagnostische toetsen tijdreeksmodel



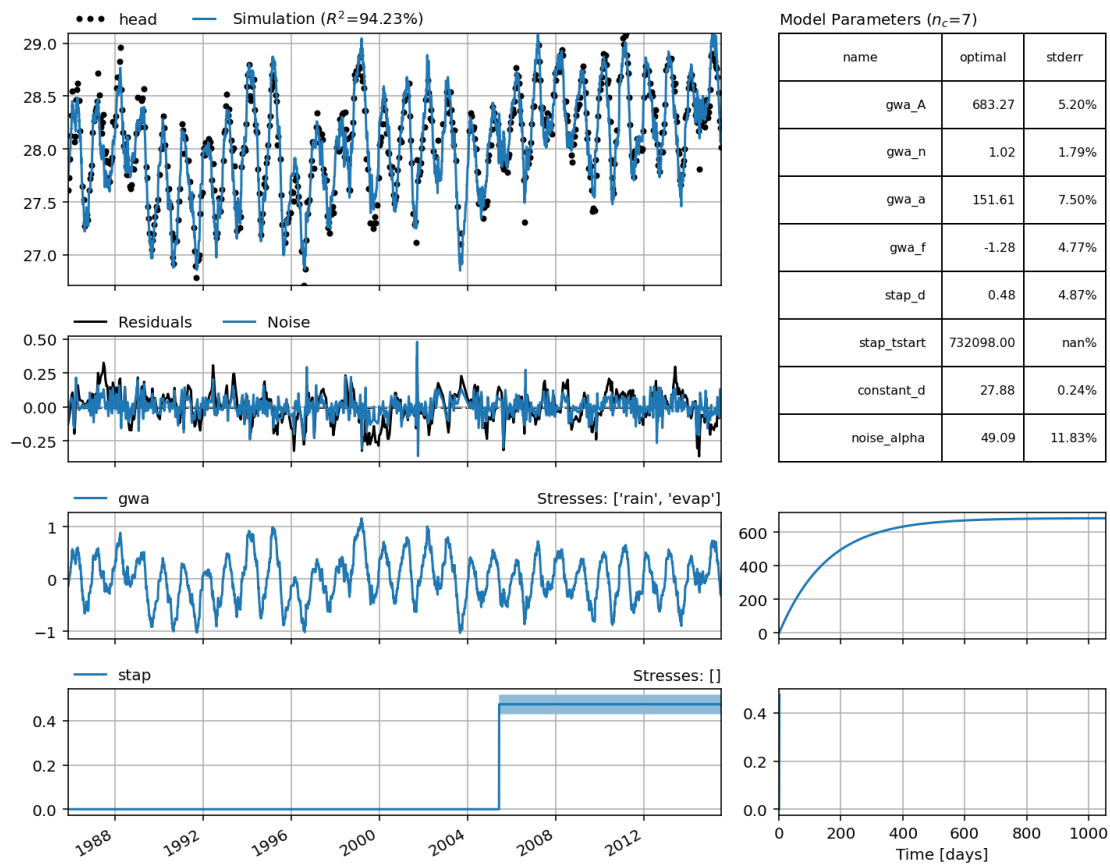
In de afbeelding linksboven is de ruis weergegeven. Te zien is dat er geen duidelijke trend zichtbaar is en dat de gemiddelde waarde (μ) gelijk is aan 0,00. In de afbeelding linksonder is de autocorrelatie van de lag weergegeven met het bijbehorende 95 % zekerheidsinterval. Hiervoor geldt dat 95 % van de autocorrelatie waarden binnen dit interval moeten liggen. In dit geval voldoet de reeks aan deze toets.

De afbeelding rechtsboven weergeeft de ruis als verdeling, met een bijbehorende normale verdeling, gefit op de ruis. Deze afbeelding kan worden gebruikt om te kijken of de ruis aan een normale verdeling voldoet. De figuur rechtsonder kan ook worden gebruikt om de normaliteit te toetsen. In dit geval voldoet de ruis goed aan een normale verdeling.

Op basis hiervan wordt aangenomen dat het gaat om 'witte' ruis en dat de onzekerheden van de modelparameters juist zijn ingeschat. Daarom is het mogelijk om de onzekerheid van de stap respons in te schatten. In de onderstaande afbeelding zijn de uitgebreide resultaten van het tijdreeksmodel weergegeven.

In de afbeelding zijn de residuen van het nieuwe tijdreeksmodel te zien. Hierin is de sprong in de residuen, die zichtbaar was bij het eerste tijdreeksmodel, niet in meer te zien bij het nieuwe tijdreeksmodel. De stap die door het tijdreeksmodel wordt berekend is gelijk aan 0.48 meter. Hierbij is het 95 % zekerheidsinterval weergegeven van de hoogte van de stap. Dit interval is tussen 0.43 en 0.52 m. Dit interval is weergegeven in de resultaten.

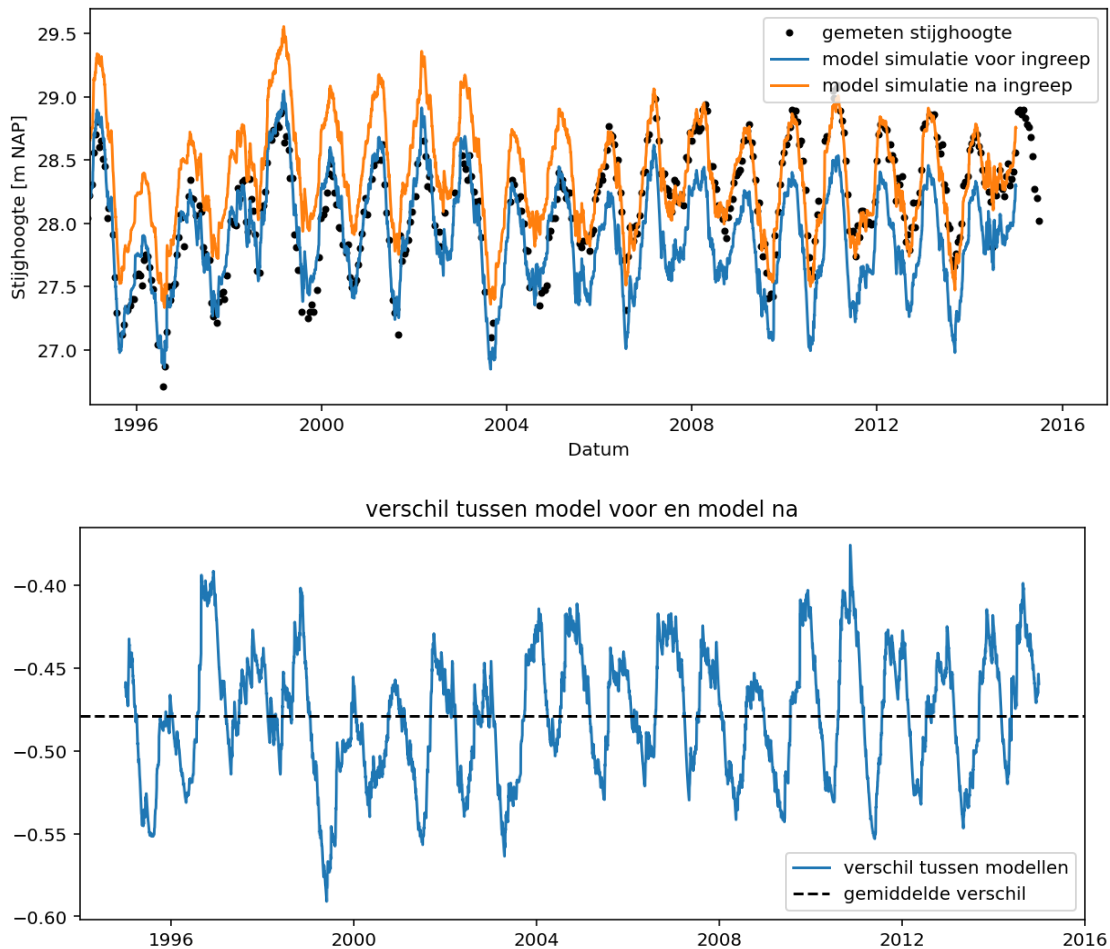
Afbeelding 10.7 Resultaten tijdreeksmodel met stap en onzekerheid



Een tweede methode wordt gebruikt om dezelfde analyse uit te voeren om de effecten van de ingreep te bepalen. Hiervoor worden 2 tijdreeksmodellen opgesteld; een model voor de periode voor dat de ingreep heeft plaatsgevonden (stijghoogtereeks tot mei 2005) en een model voor de periode na de ingrepen (stijghoogtereeks vanaf september 2005). Hierbij worden de neerslag en verdamping als verklarende reeksen gebruikt. Net als bij het vorige tijdreeksmodel wordt de Gamma respons functie gebruikt voor de neerslag en verdamping. Er wordt geen staptrend aan de modellen toegevoegd.

Met behulp van beide modellen kan de stijghoogte worden gesimuleerd voor een bepaalde periode. Met beide modellen wordt de periode van 1995 tot en met 2015 gesimuleerd, 10 jaar voor en na de ingrepen. In de onderstaande afbeelding zijn beide simulaties weergegeven, daarnaast is ook het verschil tussen beide simulaties berekend voor deze periode.

Afbeelding 10.8 Simulatie met tijdreeksmodel opgesteld voor de periode voor en na de ingreep met het verschil tussen de modellen



Zoals te zien is het verschil tussen beide modellen tussen de 0.4 en 0.6 meter. Het gemiddelde verschil is 0.48 meter. Dit komt overeen met de bevindingen van het tijdreeksmodel met de stap respons functie, hiermee kan deze model conclusie worden ondersteund. Hierbij moet wel worden opgemerkt dat bij de tweede analyse geen rekening is gehouden met de onzekerheden in het model.

10.1.2 Deel II: Invloed waterloop

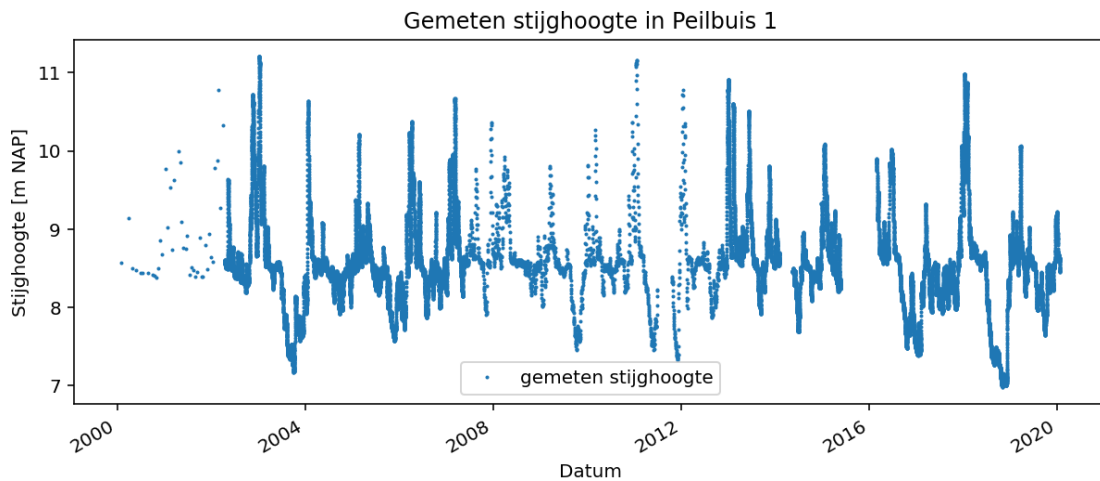
Het waterschap heeft nabij een grote waterloop een peilbuis staan om de stijghoogte in het gebied te monitoren. Het is bij de beheerder van het gebied bekend dat de waterstand in de waterloop sterk aan elkaar zijn gecorreleerd.

In de zomer van 2018 is de stijghoogte door de droogte sterk gedaald. Om dit in de toekomst te voorkomen wordt gekeken of het verhogen van de waterstanden in de waterloop hier een oplossing voor kan zijn. De hydroloog van het waterschap wil bepalen hoe de stijghoogte in het gebied reageert op veranderingen van de waterstand in de waterloop.

Beschikbare data voorbeeld 2

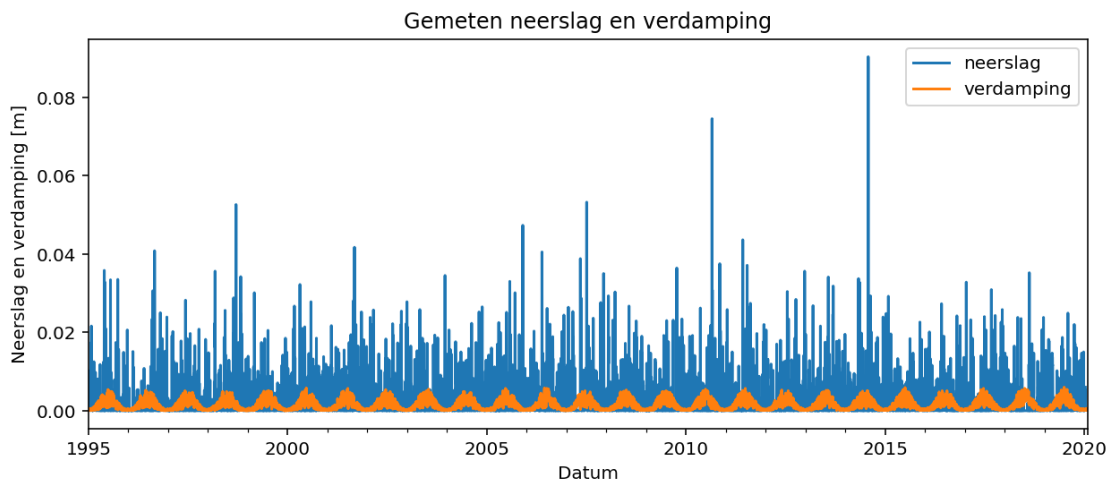
De stijghoogte is gemeten voor de periode van 2000 tot 2020. De gemeten stijghoogte is weergegeven in de onderstaande afbeelding. Te zien is dat in de zomer van 2018 de stijghoogte sterk is gedaald.

Afbeelding 10.9 Gemeten stijghoogte peilbuis 1



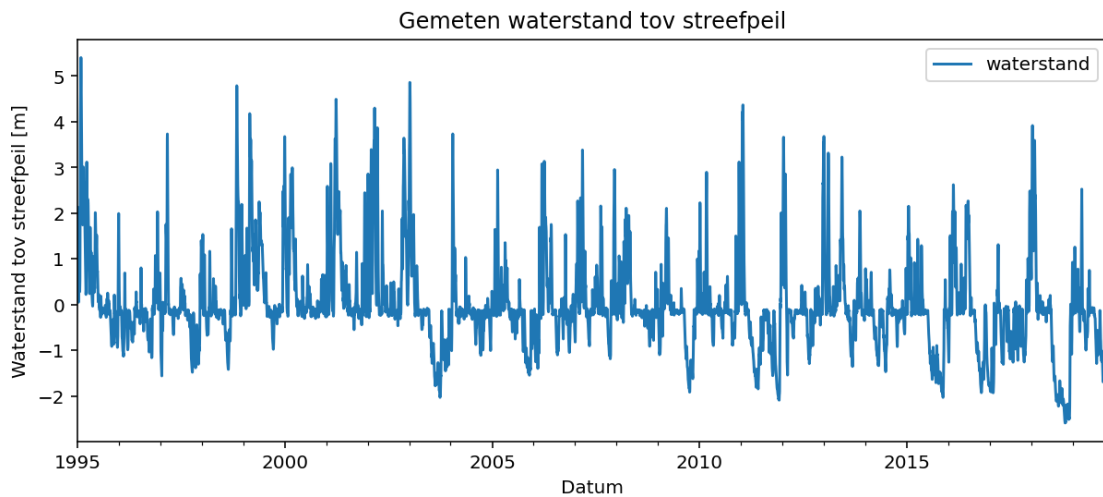
Nabij het gebied zijn de neerslag en verdampingsgegevens opgehaald van het dichtstbijzijnde KNMI weerstation. De neerslag en verdamping zijn weergegeven in de onderstaande afbeelding.

Afbeelding 10.10 Gemeten neerslag en verdamping



Naast de neerslag en verdamping wordt voor het tijdreeksmodel de waterstand in de waterloop gebruikt als verklarende reeks. In de waterloop is door het waterschap de waterstand gemeten voor de periode 1995 tot 2020. De variatie in waterstand ten opzichte van het streefpeil is te zien in de onderstaande afbeelding. In de meetreeks is te zien dat de waterstand in de zomer van 2018 erg laag was, circa 2 meter onder het streefpeil.

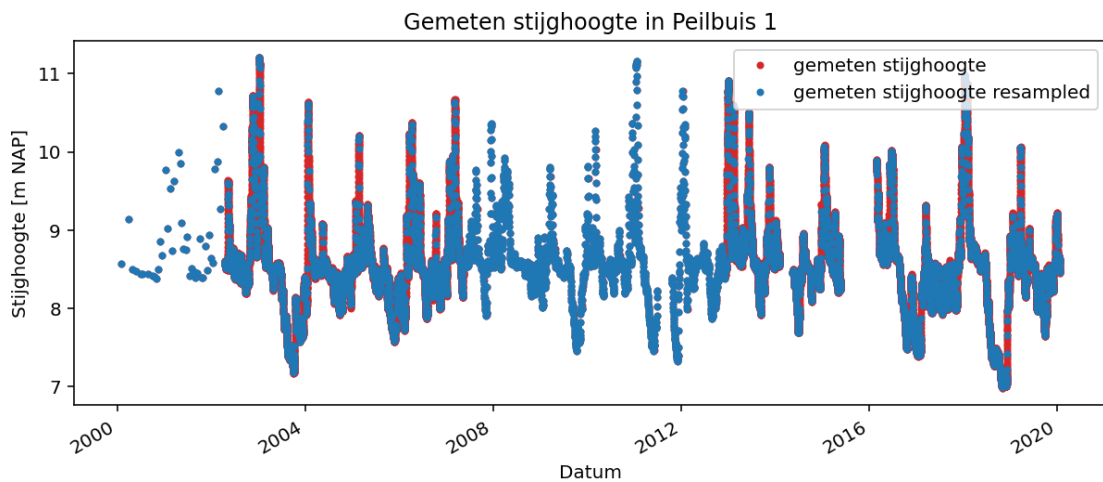
Afbeelding 10.11 Gemeten waterstand tov streefpeil



Vorbewerken reeksen voorbeeld 2

De gemeten stijghoogte heeft een meetfrequentie van eens per 3 uur. De meetreeks van het waterniveau in de waterloop heeft een meetfrequentie van eens per dag, net als de neerslag en de verdamping. Hierom wordt de meetreeks van de stijghoogte geresampled naar een frequentie van eens per dag. Hiervoor wordt de gemiddelde stijghoogte per dag genomen bij het resampelen van de meetreeks.

Afbeelding 10.12 Geresamplede stijghoogtereeks

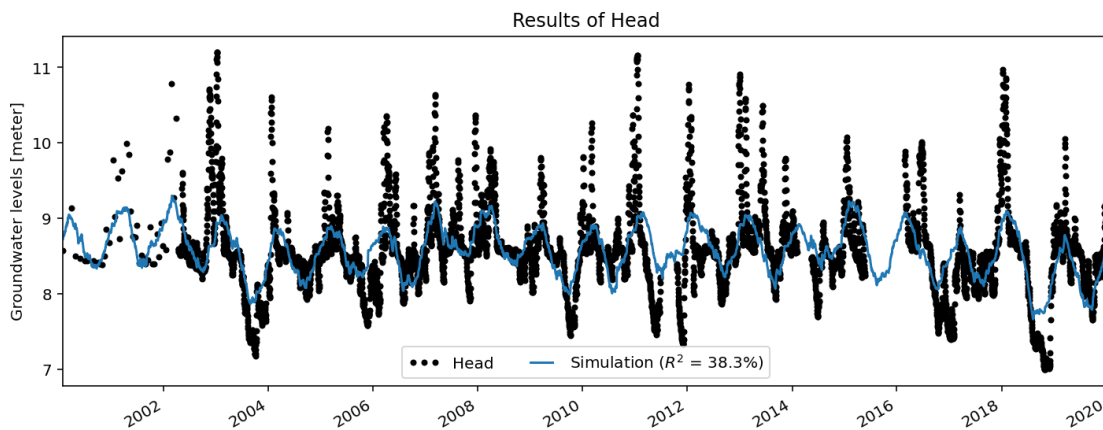


Opzetten model voorbeeld 2

Met de bewerkte meetreeks wordt een tijdreeksmodel opgezet voor de periode vanaf 2000. Eerst is een model opgesteld met enkel neerslag en verdamping als verklarende reeksen. Voor deze reeksen is de Exponentiële responsfunctie gekozen voor zowel de neerslag en verdamping.

In de onderstaande afbeelding is de gesimuleerde stijghoogte weergegeven. Te zien is dat het tijdreeksmodel geen goede verklaring geeft voor de gemeten stijghoogte. Het model is niet in staat om de variatie in stijghoogte te simuleren.

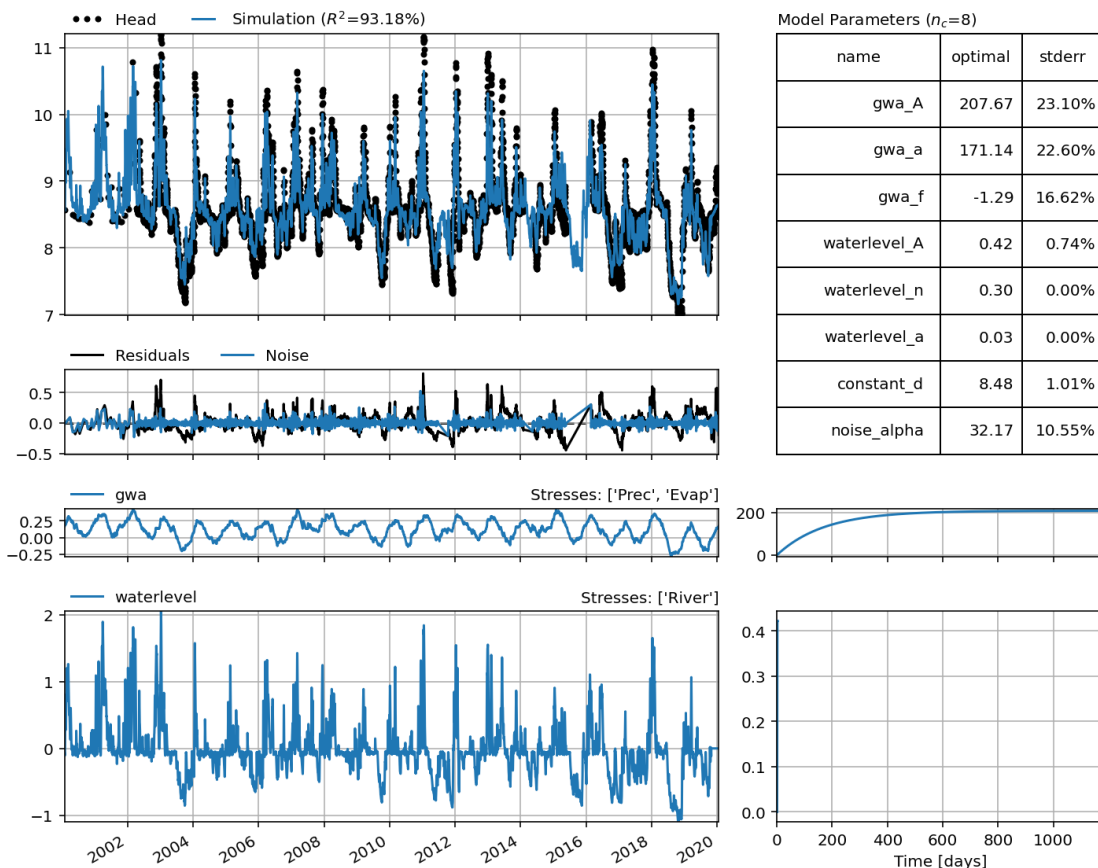
Afbeelding 10.13 Simulatie met het tijdreeksmodel zonder oppervlaktewater



Er wordt nogmaals een tijdreeksmodel opgesteld, nu met de toevoeging van de gemeten waterstand in de waterloop als verklarende reeks. Voor deze verklarende reeks worden de Gamma respons gekozen als responsfunctie. De variatie van het waterpeil is circa 6 à 7 meter, de variatie van de stijghoogte is ca. 4 meter. Omdat de stijghoogte sterk afhankelijk lijkt te zijn van de waterstand in de waterloop kan op basis van deze visuele inspectie een inschatting te maken van de reactie van de stijghoogte op de waterloop. Naar verwachting zal deze rond de 57 % tot 66 % zijn (tussen de 4/6 en 4/7).

In de onderstaande afbeelding is de simulatie van het tijdreeksmodel te zien. Hierbij is te zien dat het tijdreeksmodel de gemeten stijghoogte significant beter simuleert dan het tijdreeksmodel zonder het waterniveau in de waterloop. De R^2 van dit tijdreeksmodel is 0.93 ten opzichte van 0.38 van het eerste model. Hierin is ook te zien dat het tijdreeksmodel de stijghoogte in de zomer van 2018 goed inschat.

Afbeelding 10.14 Resultaten tijdreeksmodel met oppervlaktewater



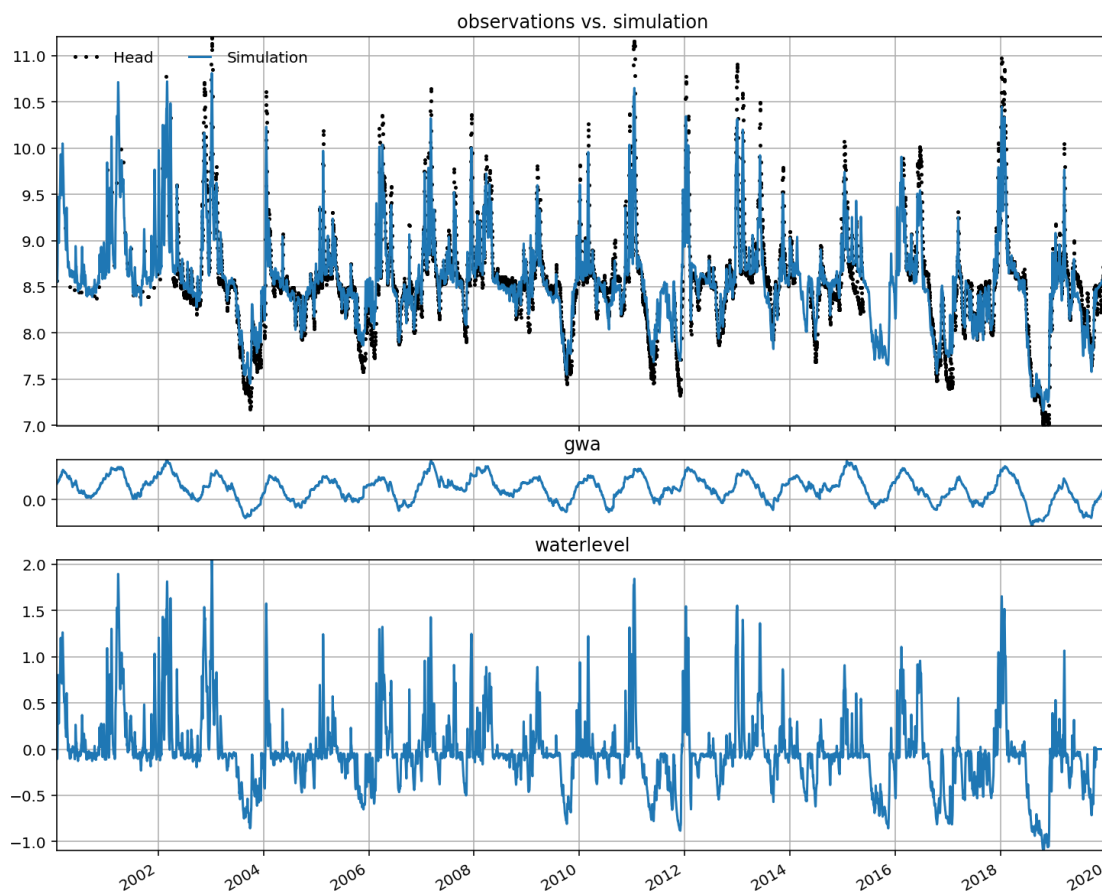
Bepalen invloed van waterloop

In de onderstaande afbeelding zijn de bijdrage van de verklarende reeksen aan de gemeten stijghoogte weergegeven. Hierbij is te zien dat de variatie in waterstand zorgt voor een bijdrage van circa -2 tot 2 m op de gemeten stijghoogte. De waterloop levert de grootste bijdrage aan de variatie van de stijghoogte. De neerslag en verdamping hebben een geringe invloed op de gemeten stijghoogte.

De stationaire waarde van de respons van de onttrekking is gelijk aan 0.57 meter, dit betekent dat de stijghoogte 0.57 meter stijgt bij een waterstandsverhoging van 1 m in de waterloop. Dit komt goed overeen met het beeld wat vooraf werd afgeleid uit de visuele inspectie van de reeks. Hierbij moet wel worden opgemerkt dat in dit geval geen onzekerheid van de hoogte van de respons is bepaald.

Met deze informatie kan de hydroloog het Waterschap adviseren. Uit het tijdreeksmodel blijkt dat de waterloop een significante invloed heeft op de stijghoogte in het gebied. Daarnaast kan met behulp van het tijdreeksmodel een inschatting worden gemaakt van de toelaatbare verlaging van de waterstand in de waterloop voordat de stijghoogte onder een niveau komt waarbij droogteschade ontstaat in het gebied waar de peilbuis zich bevindt.

Afbeelding 10.15 Bijdrage verschillende invloeden op stijghoogte



10.2 Casus: Karakteristiek bepalen

In deze casus wordt behandeld hoe tijdreeksanalyse (TRA) kan worden gebruikt voor het bepalen van de GXG waarden van een grondwaterstand meetreeks.

10.2.1 Deel I: Inschatten GXG voor korte reeks

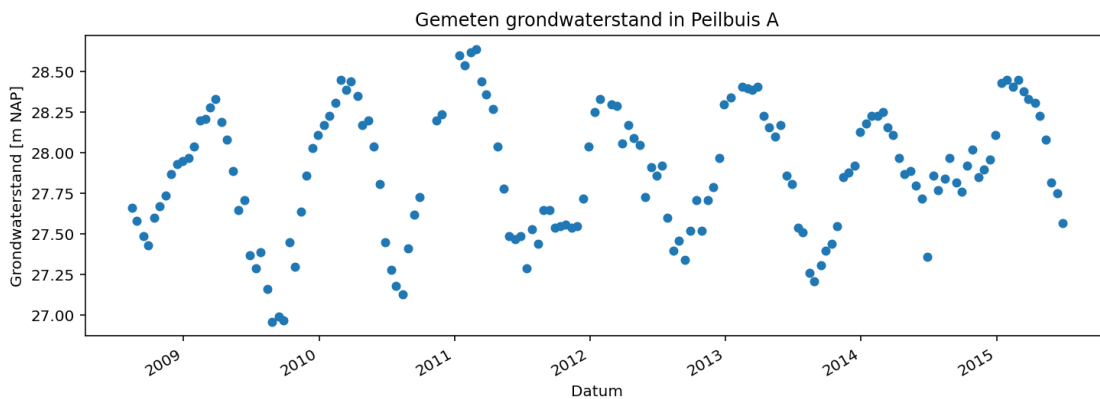
Het waterschap heeft de grondwaterstand in Peilbuis A gemeten vanaf 1 augustus 2008 tot 28 juli 2015, een totale periode van 7 jaar. Dit is de enige peilbuis in het gebied. Het Waterschap wil graag een inschatting maken van de Gemiddelde Laagste Grondwaterstand (GLG) en Gemiddelde Hoogste Grondwaterstand (GHG) binnen het gebied.

Voor de GLG worden jaarlijks de 3 laagste grondwaterstanden gemiddeld over de periode van 1 april tot en met 31 maart, over een periode van minstens 8 jaar wordt deze jaarlijkse waarde gemiddeld om de GLG te bepalen. Voor de GHG worden de 3 hoogste waarden per hydrologisch jaar gebruikt. De stijghoogte is in peilbuis A te kort gemeten voor deze berekening (<8 jaar). Met behulp van tijdreeksanalyse kan worden geprobeerd om toch een inschatting te maken van de GHG en GLG in het gebied, door de grondwaterstand terug in de tijd te extrapoleren.

Beschikbare gegevens

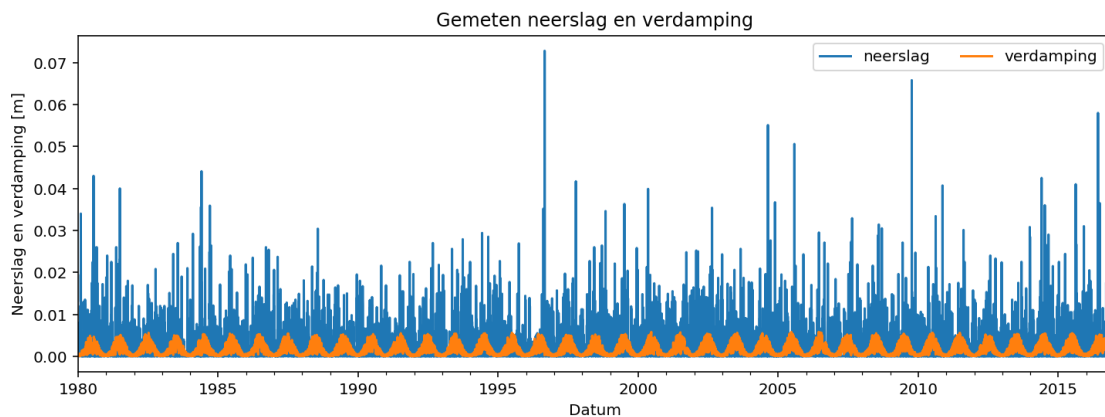
De gemeten grondwaterstand is te zien in de onderstaande afbeelding. De gemeten grondwaterstand varieert tussen de 27.2 en 28.7 m NAP.

Afbeelding 10.16 Gemeten grondwaterstand in peilbuis A



Om een tijdreeksmodel op te stellen voor grondwatermeetreeks wordt gebruik gemaakt van de neerslag en verdamping ter plaatse van Peilbuis A. Hiervoor zijn de gegevens gebruikt van het dichtstbijzijnde weerstation van het KNMI. De neerslag en verdamping zijn te zien in de onderstaande afbeelding. Te zien is dat deze meetreeksen teruggaan tot 1980, en dus langer zijn dan de gemeten grondwaterreeksen.

Afbeelding 10.17 Gemeten neerslag en verdamping



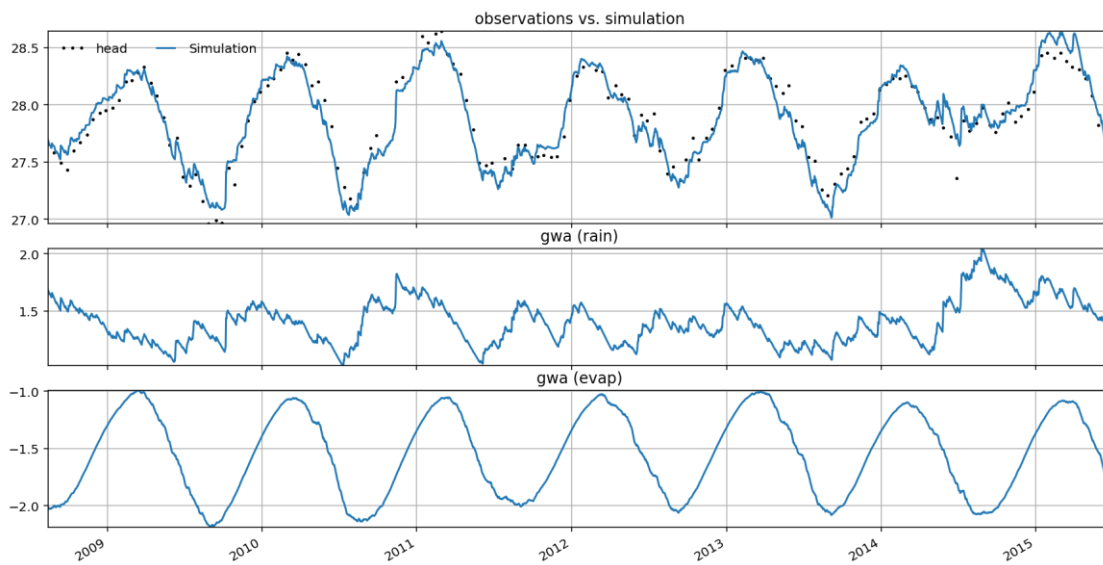
Opstellen van het tijdreeksmodel

Er wordt een model opgesteld om de grondwaterstand te simuleren. Hiervoor wordt de volledige tijdreeks gebruikt. In de tijdreeks zijn geen uitbijters gevonden. Er is daarom geen reden om de meetreeks voor te bewerken voordat deze kan worden gemodelleerd.

De neerslag en potentiële verdamping worden gebruikt als verklarende reeksen. Voor de verklarende reeksen wordt een respons gekozen. De responsfunctie beschrijft de reactie van het grondwater op een externe invloed. Deze moet worden vastgesteld voor elke verklarende reeks, waarbij de gebruiker het type responsfunctie kiest en waarbij de parameters van deze functie worden geoptimaliseerd. Hier wordt gekozen voor de Gamma responsfunctie voor de neerslag en verdamping.

In het tijdreeksmodel wordt gekozen om dezelfde responsefunctie te gebruiken voor de neerslag en de verdamping. Hierbij wordt het verband tussen de neerslag en verdamping beschreven met de formule $R = P - f \cdot E$, waarbij R de grondwateraanvulling is [m], P de neerslag [m], f de verdampingsfactor [-] en E de verdamping [m]. De verdampingsfactor wordt door gekalibreerd. Naast de verklarende reeksen wordt ook een constante gefit. Na het kiezen van de modelstructuur kan het tijdreeksmodel worden geoptimaliseerd. Het tijdreeksmodel heeft een R^2 van 0.93. In de onderstaande afbeelding is het resultaat van de simulatie van de grondwaterstand door het tijdreeksmodel weergegeven. De bijdrages van de neerslag en verdamping zijn apart weergegeven. Voor de verdamping is duidelijk het seizoensgebonden effect te zien, in de zomerperiode wordt de negatieve bijdrage van de verdamping groter.

Afbeelding 10.18 Bijdragen neerslag en verdamping op de grondwaterstand

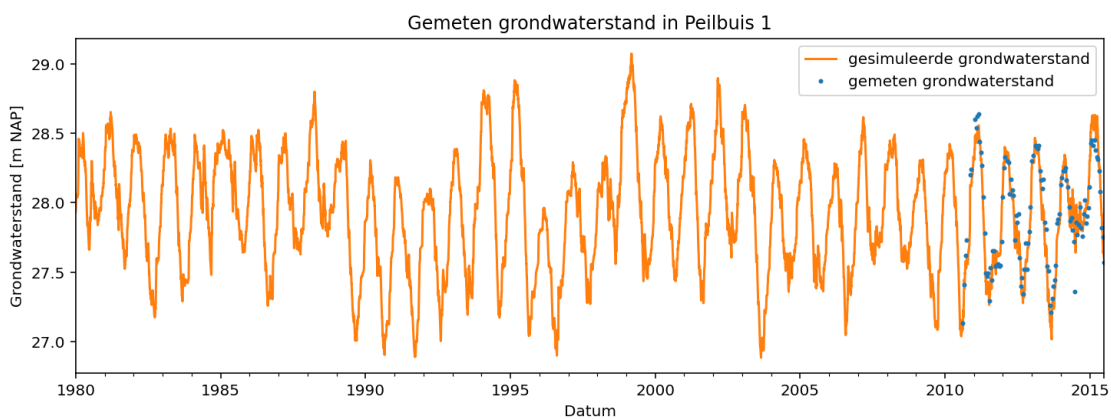


Bepalen GLG en GHG

De meetreeks is te kort om de GXG op basis van de gemeten waarden te bepalen. Met behulp van het tijdreeksmodel kan de grondwaterstand voor een langere periode worden gesimuleerd. De grondwaterstand kan worden gesimuleerd tot aan de start van de neerslag en verdampingsmeetreeksen. Hierbij wordt de aanname gedaan dat het hydrologische systeem in deze periode niet is veranderd en dat de gevonden relatie (of respons) gelijk is gebleven.

In de onderstaande afbeelding is de gesimuleerde reeks te zien vanaf 1980 (het moment waarop de neerslag en verdampingsmeetreeksen starten). Op basis hiervan kunnen de GLG en GHG worden afgeleid. De beperking hiervan is dat er op die manier geen enkele rekening wordt gehouden met de modelonzekerheid.

Afbeelding 10.19 Historische simulatie grondwaterstand



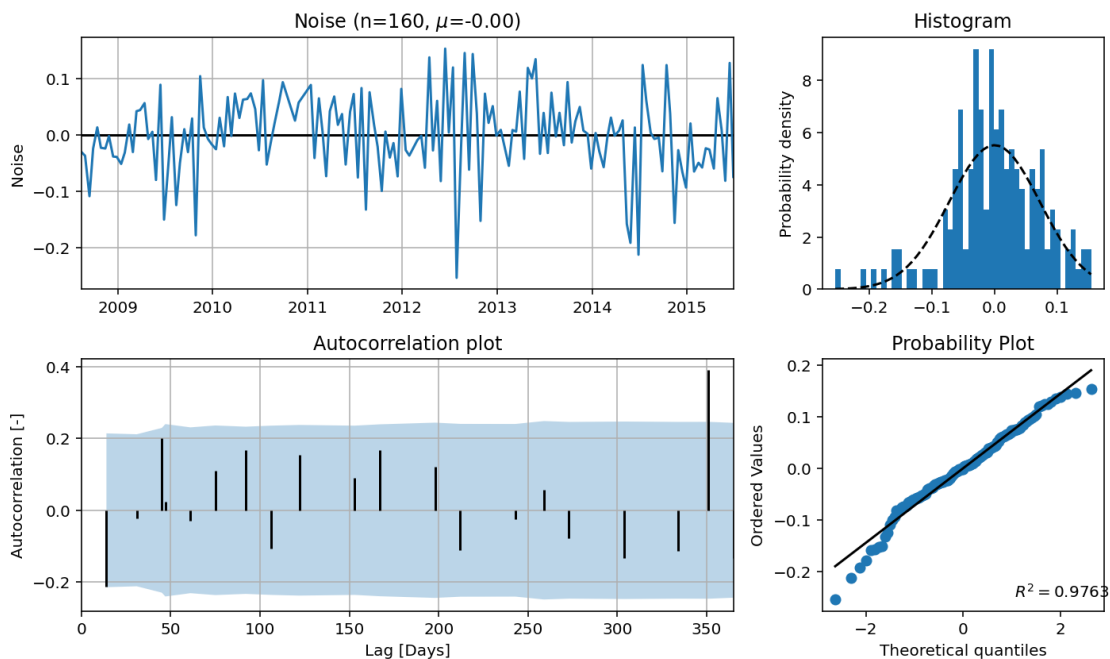
Om de onzekerheden van de GLG en GHG te bepalen is het nodig iets te zeggen over de onzekerheid van het tijdreeksmodel, uitgedrukt in de standaarddeviatie van de modelparameters. Om dit te doen moet worden gecontroleerd of de resultaten van het tijdreeksmodel voldoende zijn om de onzekerheid van de parameters goed in te kunnen schatten.

Bij het bepalen van de onzekerheid van de modelparameters is de aanname dat de ruis van het model zogenaamde 'witte ruis' is. Hiervoor worden de volgende punten gecontroleerd:

- het gemiddelde van de ruis is nul;
- de waarden zijn onafhankelijk van elkaar;
- de ruis is normaal verdeeld;
- de ruis heeft een constante variantie.

Als de ruis aan deze toetsen voldoet mag aangenomen worden dat de ruis 'witte ruis' betreft en dat de standaard deviatie van de modelparameters juist is ingeschat.

Afbeelding 10.20 Diagnostische toetsen tijdreeksmodel



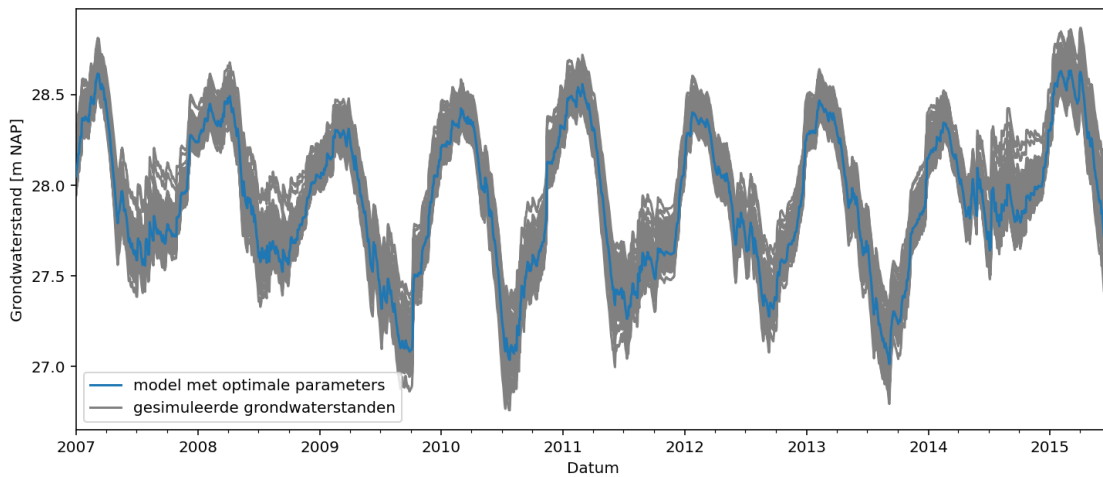
In de afbeelding linksboven is de ruis weergegeven. Te zien is dat er geen duidelijke trend zichtbaar is en dat de gemiddelde waarde (μ) gelijk is aan 0,00. In de afbeelding linksonder is de autocorrelatie van de lag weergegeven met het bijbehorende 95 % zekerheidsinterval. Hiervoor geldt dat 95 % van de autocorrelatie waarden binnen dit interval moeten liggen. Er lijkt geen sprake te zijn van een duidelijk autocorrelatie. De figuur rechtsboven weergeeft de ruis als verdeling, met een bijbehorende normale verdeling, gefit op de ruis. Deze figuur kan worden gebruikt om te kijken of de ruis aan een normale verdeling voldoet. De figuur rechtsonder kan ook worden gebruikt om de normaliteit te toetsen. De ruis lijkt voor deze reeks redelijk normaal verdeelt te zijn. Op basis van deze analyse van de ruis wordt aangenomen dat de standaard fouten van de parameters juist zijn ingeschat.

Om rekening te houden met de onzekerheid van de modelparameters bij het bepalen van de GXG worden de volgende stappen doorlopen:

- 1 met behulp van de covariantie matrix van het geoptimaliseerde model worden willekeurig 1.000 verschillende parameters sets getrokken uit een multivariate normaal verdeling;
- 2 met elk van deze 1000 parameter sets wordt de grondwaterstand gesimuleerd vanaf 2007, deze reeksen zijn te zien in de onderstaande afbeelding;
- 3 voor elk van deze gesimuleerde reeksen wordt vervolgens de GHG en GLG bepaald.

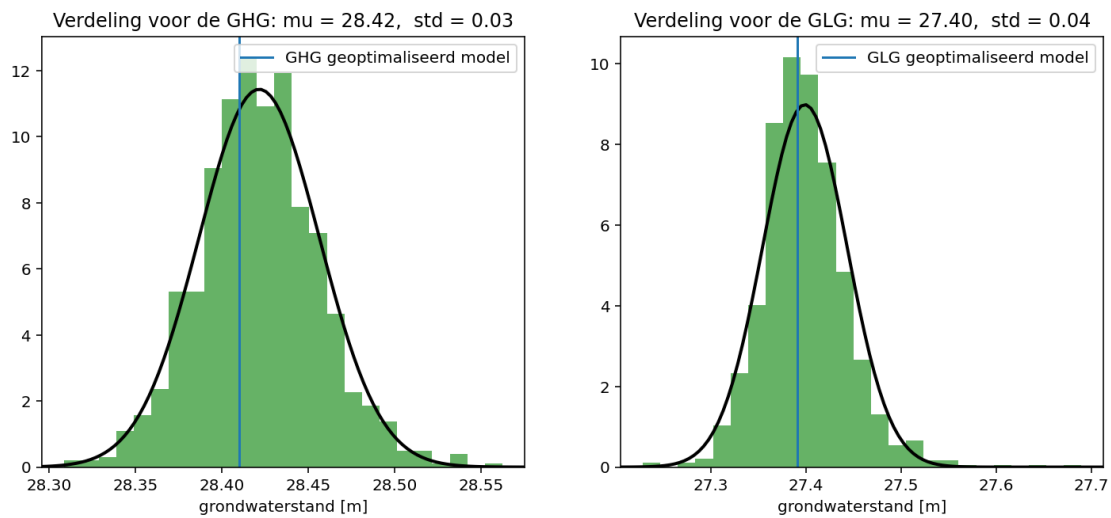
Het resultaat is niet een deterministische GXG, maar een GXG met een bandbreedte die de onzekerheid van de modelparameters in acht neemt.

Afbeelding 10.21 Resultaat modelsimulaties met verschillende parameter combinaties



In de onderstaande afbeeldingen zijn de berekende GHG en GLG waarden te zien. Over deze waarden is een normale verdeling gefit waarbij het gemiddelde en de standaarddeviatie zijn bepaald. In de afbeelding is met een blauwe lijn de GHG of GLG aangegeven welke horen bij het geoptimaliseerde tijdreeksmodel.

Afbeelding 10.22 Berekende GXG waarden



Na het uitvoeren van de analyse is toch de oudere meetreeks van Peilbuis A teruggevonden, deze meetreeks start in 1985. Het waterschap besluit om de eerdere analyse de controleren met de nieuw gevonden data. Voor de periode vanaf 2007 zijn de GXG waarden bepaald, deze zijn weergegeven in de onderstaande tabel.

De GXG op basis van de nieuwe metingen vallen binnen de bandbreedte die uit bepaald is in de vorige analyse.

Tabel 10.1 Resultaten GXG simulaties

	GLG	GHG
tijdreeksmodel vanaf 2007	27.39	28.41
controle nieuwe metingen	27.37	28.41

10.2.2 Deel II: Invullen missend deel van meetreeks

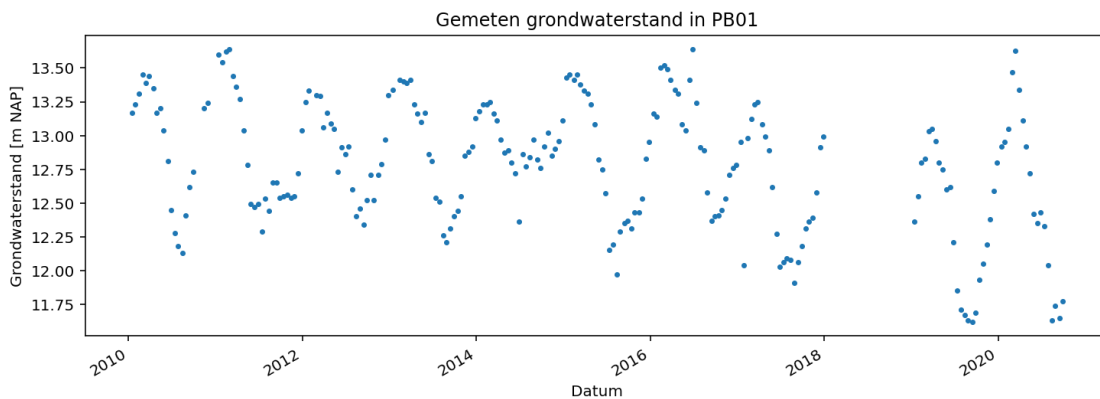
In peilbuis PB05 van het waterschap is de grondwaterstand gemeten. Het waterschap wil inzicht krijgen in de GXG in het gebied. Voor peilbuis PB05 zijn helaas geen meetgegevens beschikbaar voor het jaar 2018. In dit jaar is de grondwaterstand in het gebied zeer laag geweest. De GXG waarden op basis van de afgelopen 8 jaar worden beïnvloed door deze lage grondwaterstanden in 2018. de verwachting is dat dit voornamelijk in de GLG te zien zal zijn.

Daarom wordt gebruik gemaakt van tijdreeksanalyse om de meetreeks op te vullen. Op basis van deze reeks worden de GXG bepaald. Dan kan worden ingeschat wat het effect is van het jaar 2018 op de GXG.

Beschikbare data

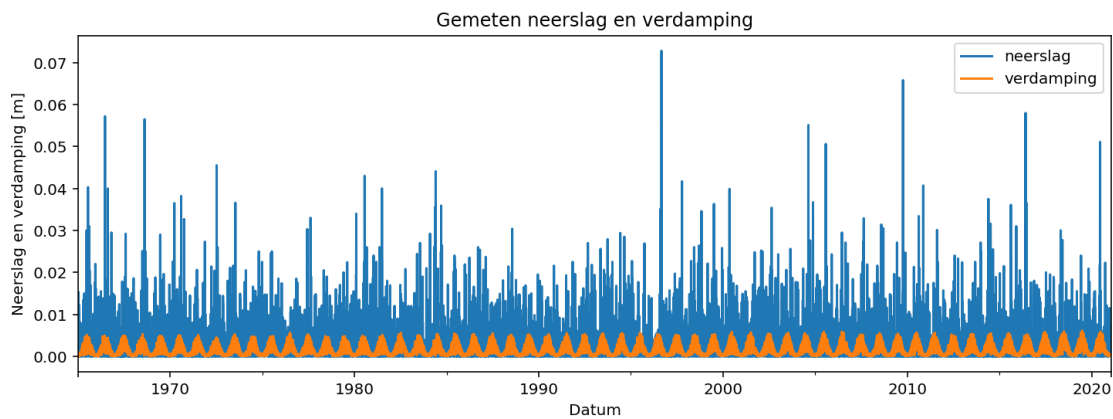
De gemeten grondwaterstand is te zien in de onderstaande afbeelding voor de periode van 2010 tot en met 2020. De meetperiode van de peilbuis is 1985 tot 2020, waarbij in het jaar 2018 geen grondwaterstanden zijn gemeten.

Afbeelding 10.23 Gemeten grondwaterstand PB01



Om een tijdreeksmodel op te stellen voor grondwatermeetreeks wordt gebruik gemaakt van de neerslag en verdamping ter plaatse van de peilbuis. Hiervoor zijn de gegevens gebruikt van het dichtstbijzijnde weerstation van het KNMI. De neerslag en verdamping zijn te zien in de onderstaande afbeelding.

Afbeelding 10.24 Gemeten neerslag en verdamping

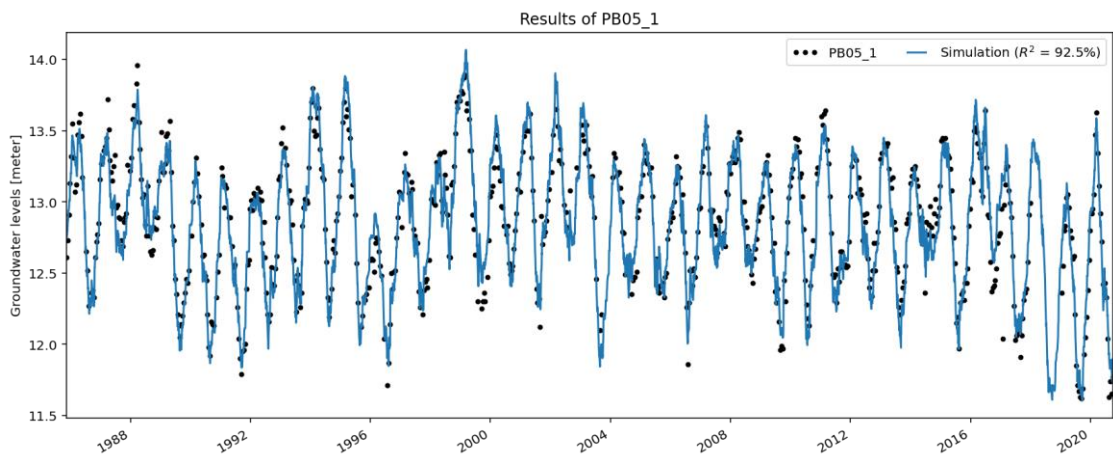


Opzetten tijdreeksmodel

Er wordt een tijdreeksmodel opgesteld op basis van de meetreeks in PB05. Hiervoor worden de neerslag en verdamping gebruikt als verklarende reeksen. Voor de neerslag en verdamping wordt de Exponentiele respons functie gekozen voor de tijdreeksanalyse.

Het tijdreeksmodel heeft een R^2 van 0.92. De gesimuleerde grondwaterstand is te zien in de onderstaande afbeelding. Hierin is te zien dat het tijdreeksmodel de grondwaterstand goed simuleert. Wel is hierbij te zien dat de data van 2018 ontbreekt. Het tijdreeksmodel simuleert voor de zomer van 2018 een lage grondwaterstand zoals verwacht.

Afbeelding 10.25 Simulatie met tijdreeksmodel



In de onderstaande tabel zijn de GXG waarden weergegeven die zijn berekend met behulp van de gesimuleerde tijdreeks. Hierbij is het droge jaar 2018 meegenomen ook al zijn voor dit jaar geen metingen beschikbaar. Voor deze analyse moet worden opgemerkt dat de aanname is gemaakt dat het grondwatersysteem in het jaar 2018 hetzelfde reageert (de respons is gelijk) op de neerslag en verdamping als in andere jaren, ondanks de extreme droogte. Daarnaast is in deze analyse niet gekeken naar de onzekerheden van de berekende GXG.

Tabel 10.2 Berekende GXG

	GLG	GHG
tijdreeksmodel	12.16	13.38

10.3 Casus: Systemanalyse

In deze casus wordt behandeld hoe met behulp van tijdreeksanalyse een hydrologische systeem kan worden geanalyseerd. Hierbij wordt gekeken naar een netwerk van meerdere peilbuizen. Door het opstellen en analyseren van tijdreeksmodellen voor al deze peilbuizen kan informatie over het systeem worden verkregen.

10.3.1 Verlaging natuurgebied

In de omgeving een natuurgebied bevindt zich een onttrekking voor een groot industrieel complex. Bij het waterschap is het vermoeden dat deze onttrekking mogelijk leidt tot een verlaging van de stijghoogte in het gebied, terwijl dit op basis van de vergunning niet toegestaan is. Het waterschap besluit daarom te kijken of op basis van peilgegevens kan worden aangetoond of de onttrekking voor verlagingen van de stijghoogte heeft gezorgd. Dit wordt onderzocht met behulp van tijdreeksanalyse.

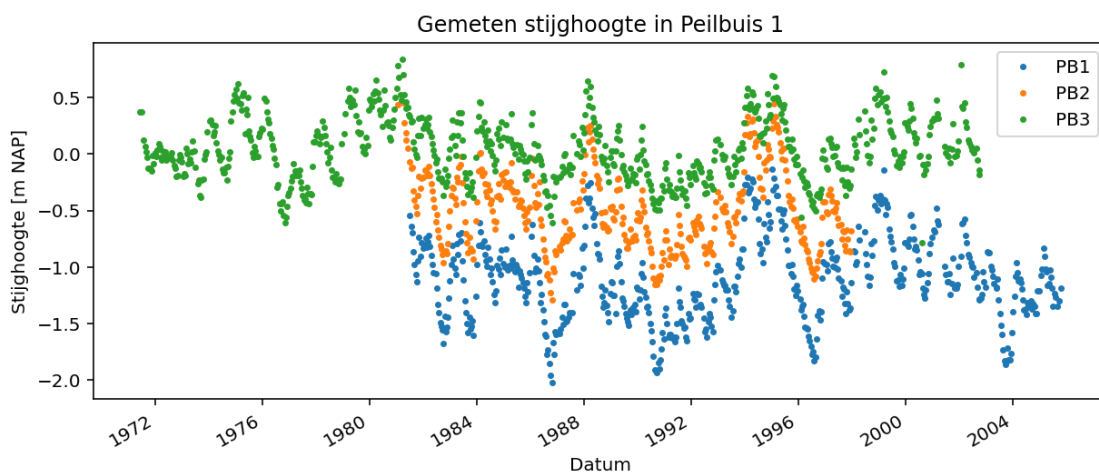
De onttrekking bevindt zich op een afstand van 500 m van het natuurgebied. Hierbij wordt onttrokken uit het eerste watervoerende pakket. Dit is ook het pakket waarin de peilbuizen zich bevinden.

Beschikbare data

Tabel 10.3 Afstand peilbuizen tot onttrekking

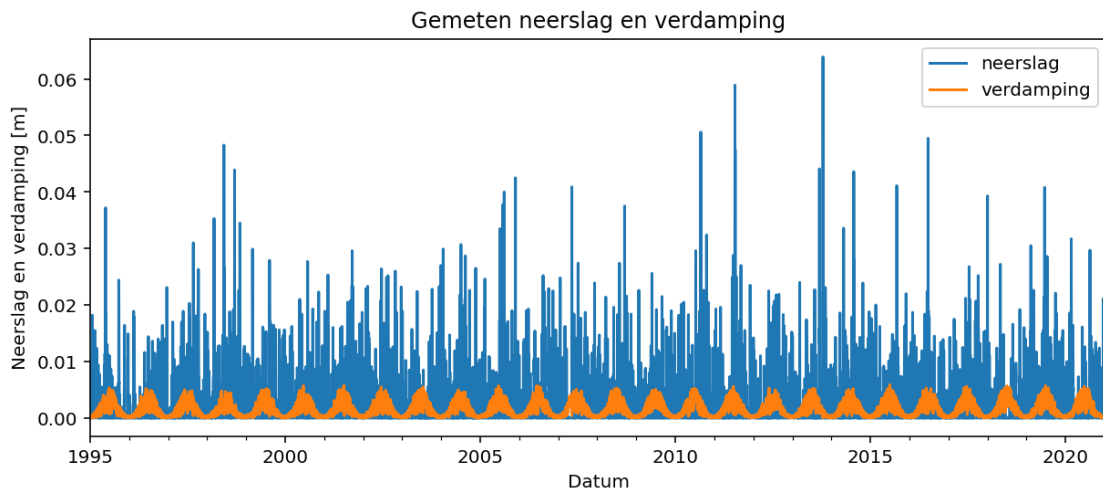
Peilbuis	Afstand [m]
PB1	112
PB2	180
PB3	321

Afbeelding 10.26 Gemeten stijghoogte in verschillende peilbuizen



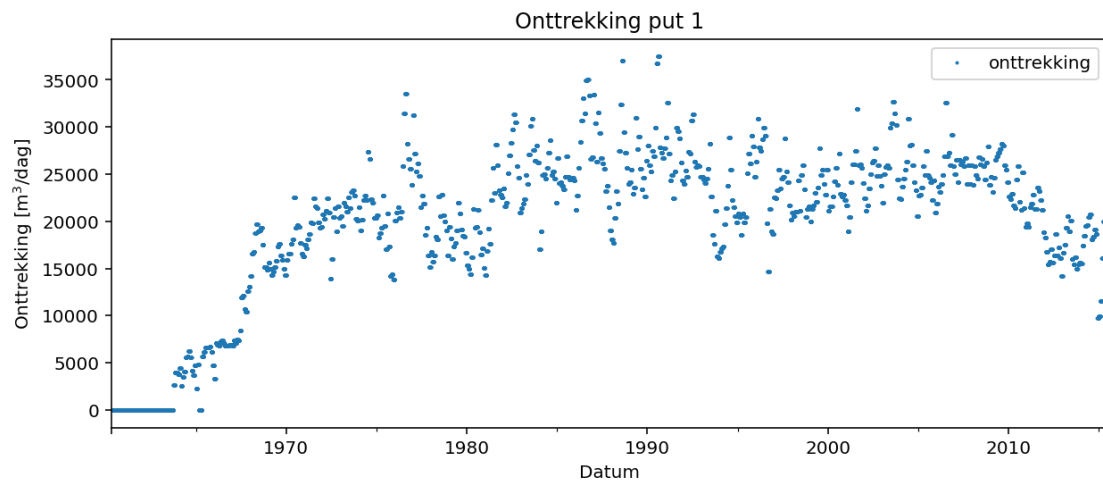
Om een tijdreeksmodel op te stellen voor stijghoogtemeetreeks wordt gebruik gemaakt van de neerslag en verdamping. Hiervoor zijn de gegevens gebruikt van het dichtstbijzijnde weerstation van het KNMI. De neerslag en verdamping zijn te zien in de onderstaande afbeelding. Voor alle peilbuizen is dit weerstation het dichtstbijzijnde.

Afbeelding 10.27 Gemeten neerslag en verdamping



Om uit te zoeken wat het effect is van de onttrekking op de gemeten stijghoogte is bij de eigenaar van de onttrekking de meetreeks voor de onttrekking opgehaald. De onttrekking is opgestart in de jaren 60. De onttrekking varieert door de jaren heen en is sinds 2010 iets afgebouwd.

Afbeelding 10.28 Onttrekking put 1



Opstellen van het tijdreeksmodel

Er wordt een model opgesteld voor de stijghoogte observaties, voor elke peilbuis wordt een apart model opgezet. Hierbij wordt de volledige reeks gebruikt voor alle peilbuizen. In de reeksen zijn geen uitschieters gevonden, er is daarom geen reden om de meetreeksen voor te bewerken.

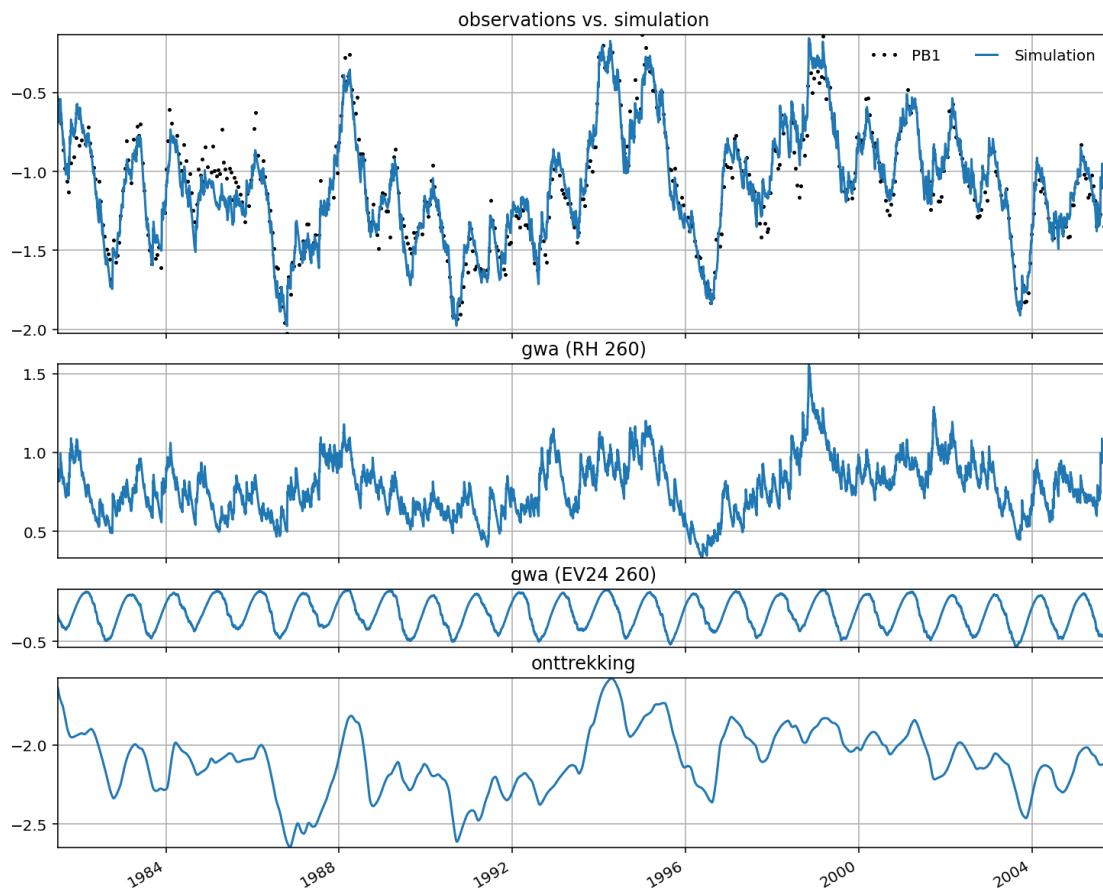
De neerslag en verdamping worden gebruikt als verklarende reeksen. Voor de verklarende reeksen wordt een respons gekozen. De respons functie beschrijft de reactie van de stijghoogte op een externe invloed. Voor het tijdreeksmodellen wordt gekozen voor de Gamma respons functie voor de neerslag en verdamping. Voor elk tijdreeksmodel wordt deze respons functie individueel geoptimaliseerd.

In het tijdreeksmodel wordt gekozen om dezelfde response functie te gebruiken voor de neerslag en de verdamping. Hierbij wordt het verband tussen de neerslag en verdamping beschreven met de formule $R = P - f \cdot E$, waarbij R de grondwateraanvulling is [m], P de neerslag [m], f de verdampingsfactor [-] en E de verdamping [m]. De verdampingsfactor wordt door het tijdreeksmodel geoptimaliseerd.

Daarnaast wordt ook de onttrekking als verklarende reeks gebruikt. Voor deze verklarende reeks wordt de Hantush responsfunctie gekozen. Naast de verklarende reeksen wordt ook de constante (in het tijdreeksmodel d) geoptimaliseerd in het model.

Met deze verklarende reeksen en de gekozen responsfuncties worden de tijdreeksmodellen geoptimaliseerd. Het tijdreeksmodel voor PB1 heeft een R^2 van 0.93. Met het tijdreeksmodel kan de gemeten stijghoogte goed worden gesimuleerd. De simulatie van het tijdreeksmodel voor PB1 is te zien in de onderstaande afbeelding. Hier is het effect van elke invoerreeds te zien als aparte bijdrage. Hierbij is te zien dat ten gevolge van de onttrekking de stijghoogte wordt verlaagt tussen de circa -2.5 en -1.5 m.

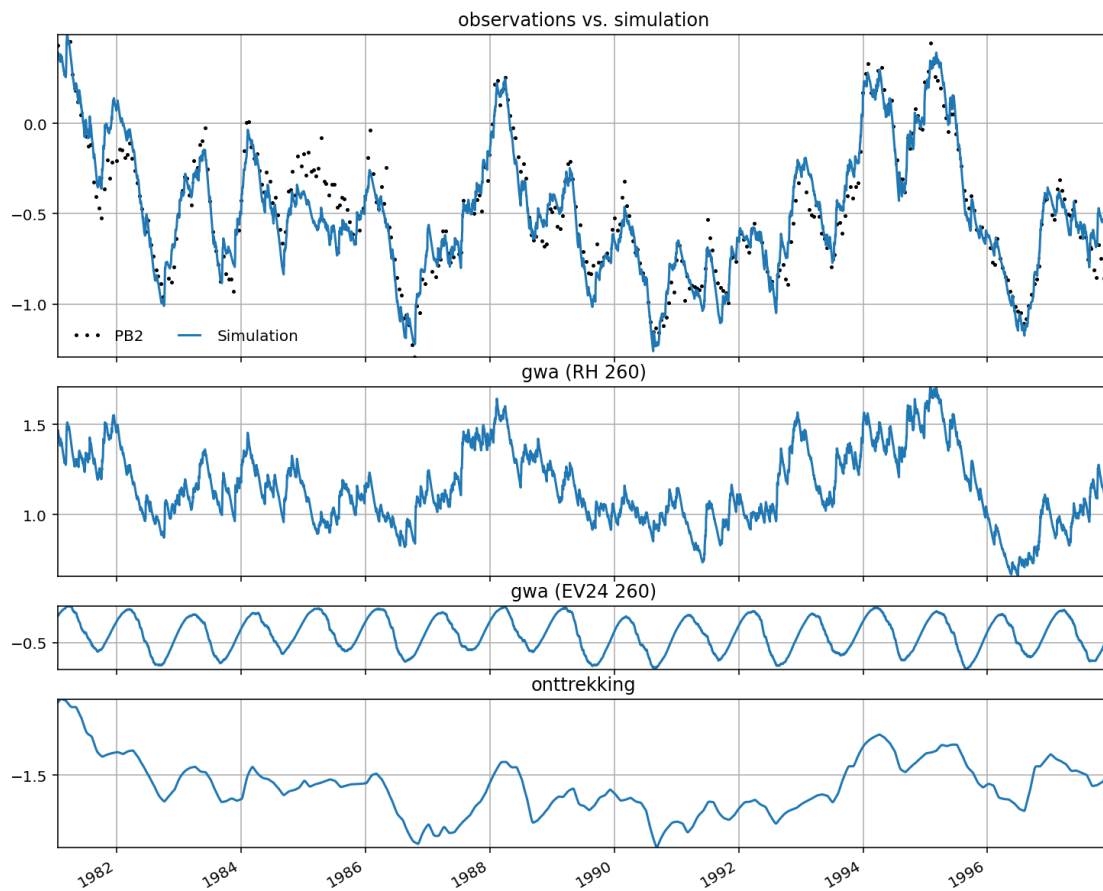
Afbeelding 10.29 Resultaten tijdreeksmodel PB1 met bijdragen van verschillende invloeden



Voor peilbuis PB2 is een tijdreeksmodel opgezet met dezelfde verklarende reeksen.

Te zien is dat het tijdreeksmodel een R^2 heeft van 0.91. In de onderstaande afbeelding is de simulatie van de stijghoogte weergegeven. Daarnaast is per invloed de bijdrage te zien. Hierbij is te zien dat ten gevolge van de onttrekking de stijghoogte wordt verlaagt tussen de circa -1 en -2 m. Dit is lager dan de verlaging die bepaald is ter plaatse van PB1. Dit komt overeen met de hydrologische verwachting, PB2 staat immers verder van de onttrekking dan PB1.

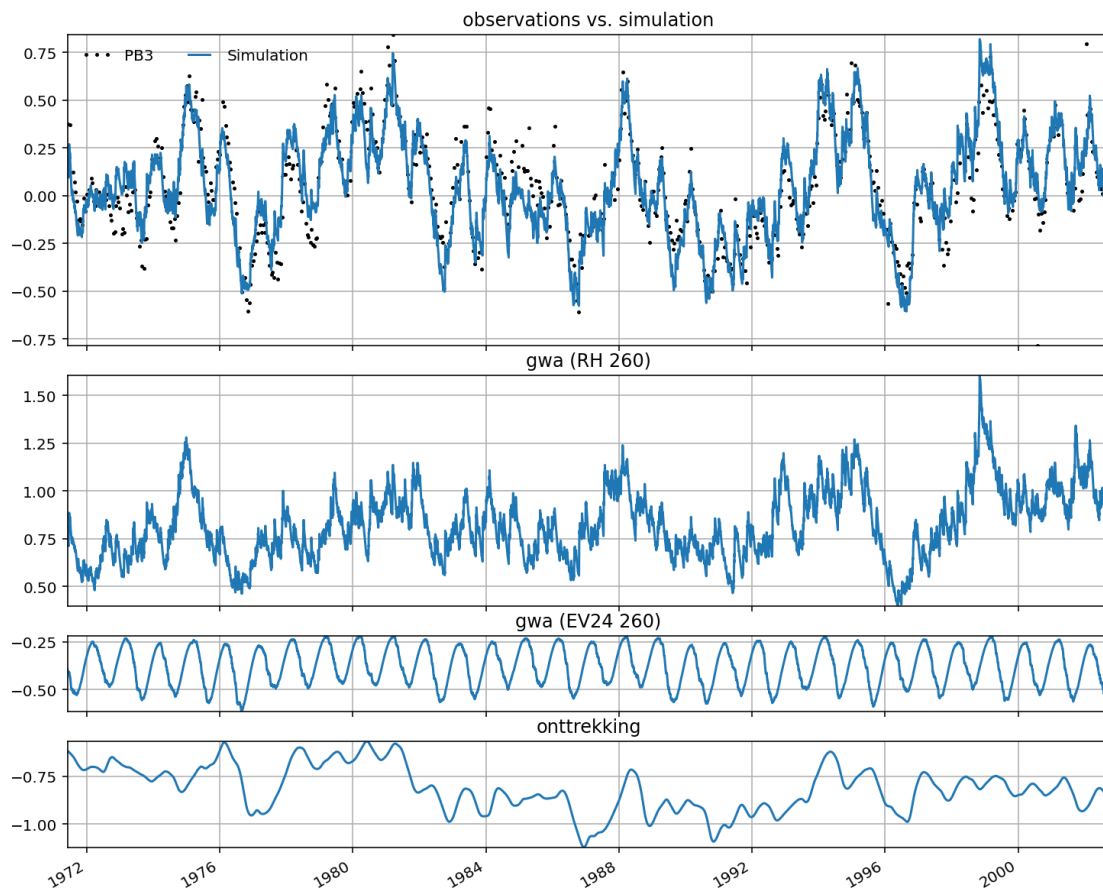
Afbeelding 10.30 Resultaten tijdreeksmodel PB2 met bijdragen van verschillende invloeden



Als laatste is ook voor peilbuis PB3 een tijdreeksmodel opgesteld.

Het tijdreeksmodel heeft een R^2 van 0.84, in de onderstaande afbeelding is de gesimuleerde reeks te zien. Per invloed is de bijdrage weergegeven. Hier is het effect van elke invoerreeks te zien volgens de modellering. Hierbij is te zien dat de verlaging ten gevolge van de onttrekking tussen de circa -0.5 en -1 m is. Op deze afstand is dus nog steeds een effect van de onttrekking te zien in de stijghoogtemeetreeks.

Afbeelding 10.31 Resultaten tijdreeksmodel PB3 met bijdragen van verschillende invloeden



Verlaging ten gevolge van onttrekking

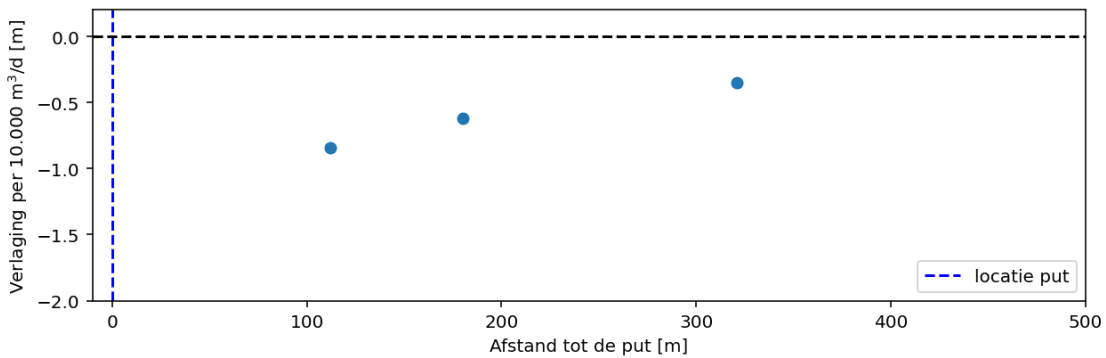
In de onderstaande tabel zijn de gemiddelde verlagingen opgenomen zoals deze met de tijdreeksanalyse zijn bepaald. Omdat de meetreeksen van de peilbuizen verschillende meetperiodes hebben wordt gekeken naar de gemiddelde onttrekking voor deze periode en wordt de gemiddelde verlaging per 10.000 m³/d bepaald.

Tabel 10.4 Overzicht resultaten tijdreeksanalyse

Peilbuis	Afstand [m]	Gemiddelde verlaging [m]	Gemiddelde onttrekking [m ³ /d]	Gemiddelde verlaging per 10.000 m ³ /d [m]
PB1	112	-2.12	24985	-0.85
PB2	180	-1.57	25129	-0.63
PB3	321	-0.81	23448	-0.35

In de onderstaande grafiek zijn de gemiddeld bepaalde verlagingen per 10.000 m³/d uitgezet tegen de afstand tot de onttrekkingsput. Het watervoerendpakket in de omgeving is uniform, daarom is de verwachting dat op een afstand van 500 m, de afstand tussen het natuurgebied en de onttrekking, een significante verlaging optreed ten gevolge van de onttrekking bij het huidige debiet (circa 20.000 m³/d).

Afbeelding 10.32 Verlagen versus afstand tot de put



10.4 Casus: Voorspellen

In deze casus laten we de verschillende manieren zien hoe een tijdreeksmodel kan worden gebruikt om te voorspellen. De casus is opgedeeld in twee delen. In het eerste deel wordt gekeken hoe een tijdreeksmodel kan worden gebruikt om te onderzoeken wat de grondwaterstand mogelijk doet na het uitzetten van een pompstation. Het tweede deel beschrijft een voorspelling voor de langere termijn.

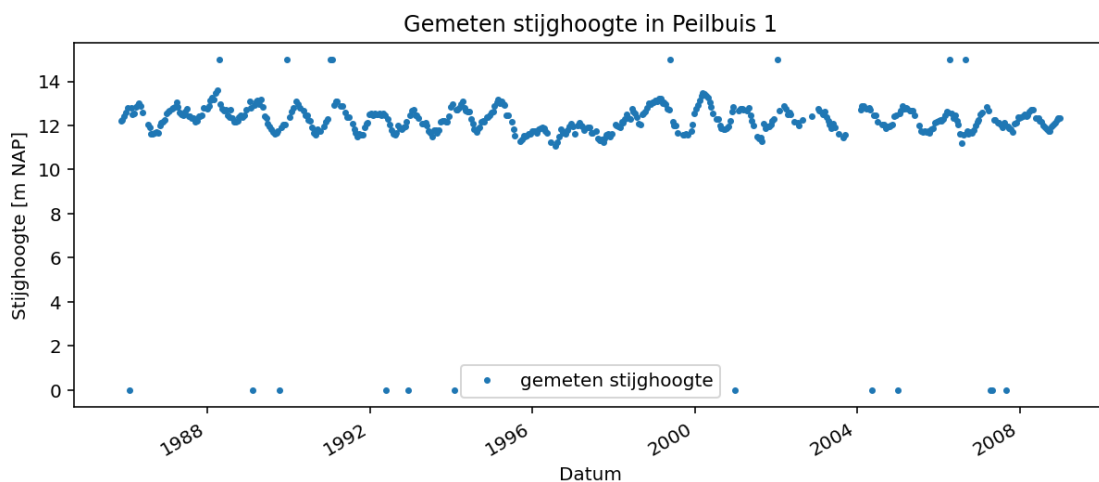
10.4.1 Deel I: Uitzetten onttrekking

Het waterschap heeft de stijghoogte gemeten in Peilbuis 1. In de omgeving van de peilbuis bevindt zich een onttrekking. Vanaf 5 november 2009 de onttrekking uitgezet en is het onttrokken debiet teruggeschroefd naar 0 m³/d. Aan een adviesbureau is gevraagd om te bepalen wat de verwachte stijging is van de stijghoogte ter plaatse van peilbuis 1 als gevolg van het uitzetten van deze onttrekking. Hierbij zijn gegevens beschikbaar tot het moment waarop de onttrekking is uitgezet.

Beschikbare data voorbeeld 1

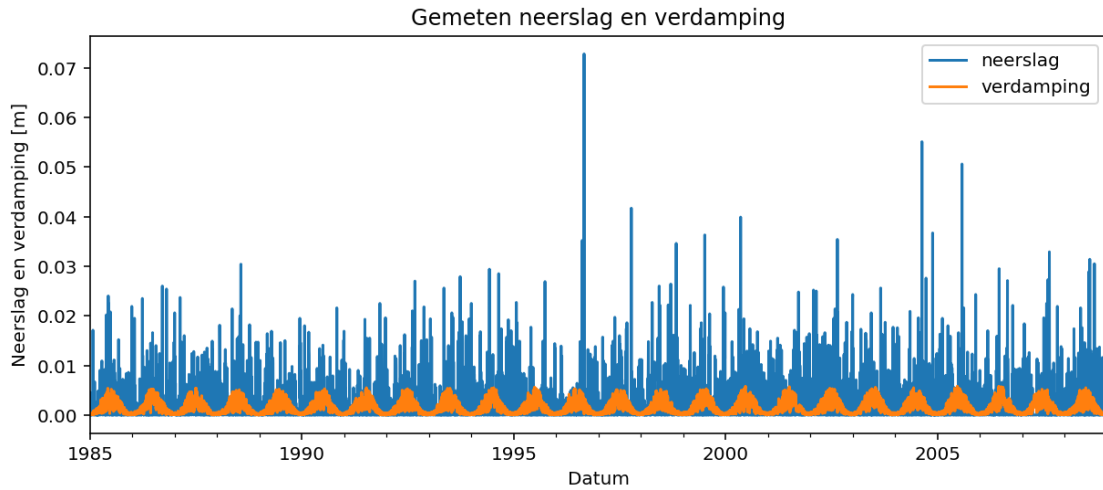
De gemeten stijghoogte is te zien in de onderstaande afbeelding. De meetperiode is van 1985 tot en met 2009. De meetreeks bevat enkele waarden die niet in de reeks passen. Daarom is het nodig om de meetreeks voor te bewerken voordat het in het tijdreeksmodel kan worden toegepast.

Afbeelding 10.33 Gemeten stijghoogte in peilbuis 1



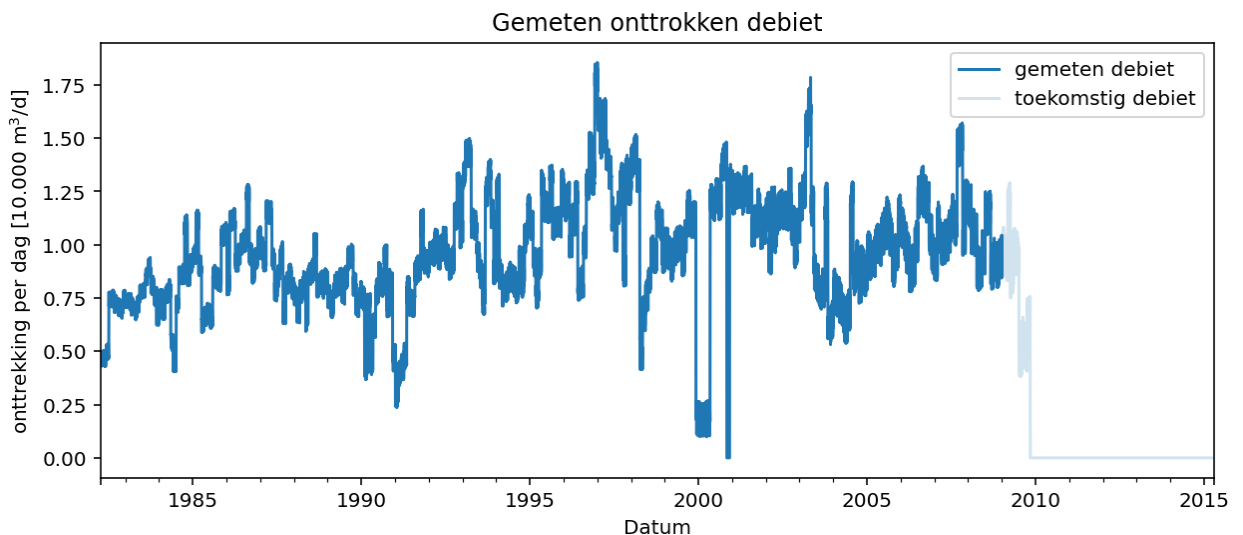
Om een tijdreeksmodel op te stellen voor stijghoogte meetreeks wordt gebruik gemaakt van de neerslag en verdamping 1. Hiervoor zijn de gegevens gebruikt van het dichtstbijzijnde weerstation van het KNMI. De gemeten neerslag en verdamping zijn te zien in de onderstaande afbeelding.

Afbeelding 10.34 Gemeten neerslag en verdamping



Naast de neerslag en verdamping is ook de data van de onttrekking gebruikt als verklarende reeks in het tijdreeksmodel. Hiervoor is bij de beheerder van de onttrekking het dagelijkse debiet opgevraagd. Deze is te zien in de afbeelding hieronder. Hierin is ook te zien dat de onttrekking aan het eind van 2009 volledig is afgeschaald. In de periode hiervoor varieerde het debiet van de onttrekking tussen de circa 5.000-18.000 m³/d.

Afbeelding 10.35 Gemeten onttrokken debiet en toekomstig debiet (vanaf 2009)

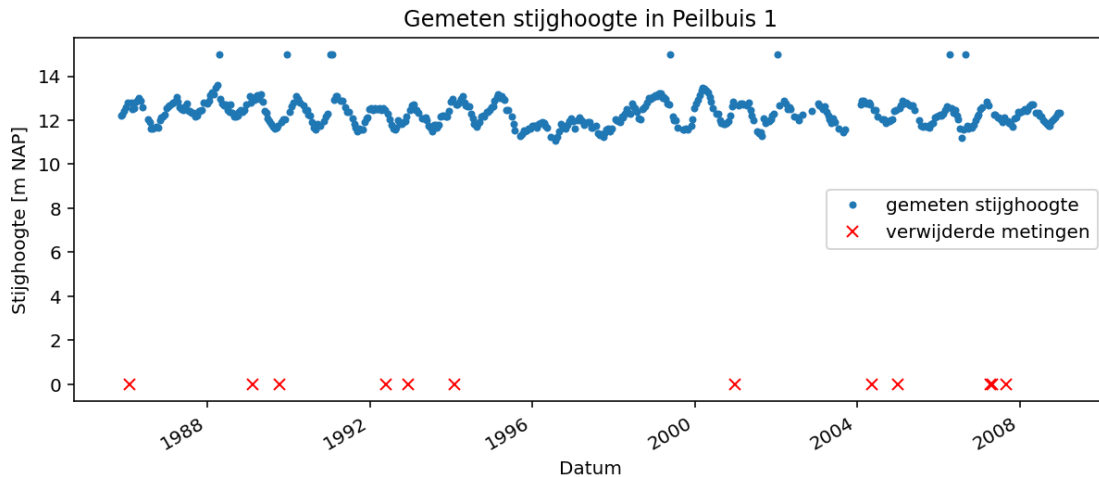


Vorbewerken reeksen voorbeeld 1

Zoals te zien in de stijghoogtemeetreeks bevat de meetreeks een aantal 0 waarden. Deze waarden zijn mogelijk het gevolg van een meetfout van de sensor of een fout in het uitlezen van de sensor. Om de reeks bruikbaar te maken voor tijdreeksanalyse moet de meetreeks worden voorberekt.

De metingen met een waarde van 0 worden verwijderd uit de reeks om deze te gebruiken in het tijdreeksmodel. Hierbij worden de gaten in de reeks niet opgevuld.

Afbeelding 10.36 Gemeten stijghoogte zonder 0 metingen



Opstellen van het tijdreeksmodel

De eerste stap voor het opstellen van het model is het toevoegen van de stijghoogte observaties aan het model. Voor een eerste berekening wordt de stijghoogtereeks gebruikt tot aan de huidige datum. De neerslag en verdamping worden gebruikt als verklarende reeksen. Voor de verklarende reeksen wordt een respons gekozen. De respons functie beschrijft de reactie van de stijghoogte op een externe invloed. Deze moet worden vastgesteld voor elke verklarende reeks, waarbij de gebruiker het type respons functie kiest en waarbij de parameters van deze functie worden geoptimaliseerd. Voor het tijdreeksmodel wordt gekozen voor de Gamma respons functie voor de neerslag en verdamping.

In het tijdreeksmodel wordt dezelfde responsfunctie gebruikt voor de neerslag en de verdamping. Hierbij wordt het verband tussen de neerslag en verdamping beschreven met de formule $R = P - f \cdot E$, waarbij R de grondwateraanvulling is [m], P de neerslag [m], f de verdampingsfactor [-] en E de verdamping [m]. De verdampingsfactor wordt door het tijdreeksmodel geoptimaliseerd samen met de drie parameters van de Gamma responsfunctie.

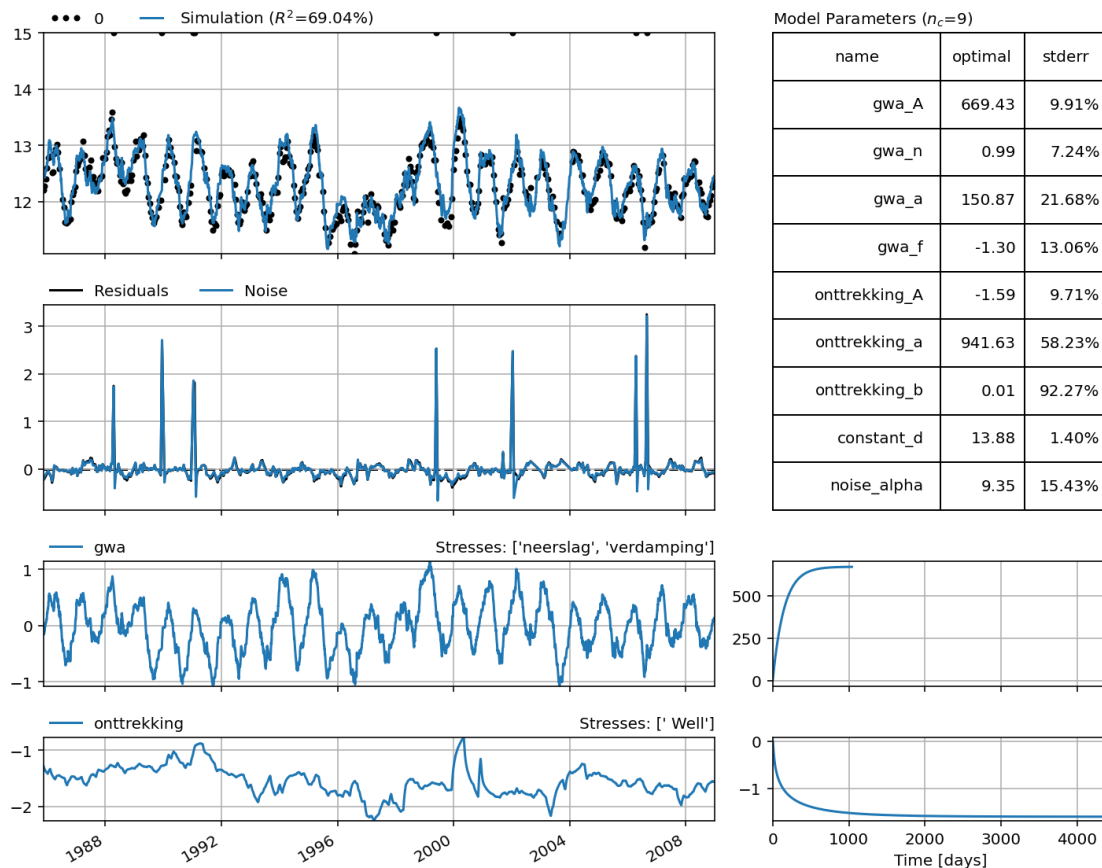
De onttrekking wordt als verklarende reeks gebruikt in het tijdreeksmodel. Voor deze verklarende reeks wordt de Hantush responsfunctie gekozen. In de opzet van het tijdreeksmodel is de onttrekking geschaald naar 10.000 m³/d, zie ook paragraaf 9.1. Naast de verklarende reeksen wordt ook de constante (in het tijdreeksmodel d) geoptimaliseerd in het model.

Na het toevoegen van de observaties en de verklarende reeksen kan het tijdreeksmodel worden geoptimaliseerd. De resultaten van het tijdreeksmodel zijn te zien in de onderstaande tabel.

Als het tijdreeksmodel is geoptimaliseerd kan de gesimuleerde reeks worden geplott. Dit is te zien in de onderstaande afbeelding samen met een overzicht van de modelresultaten.

In de afbeelding zijn de metingen en de simulatie visueel weergegeven. Daaronder zijn de model residuen en ruis weergegeven. Voor de verschillende verklarende reeksen zijn de bijdrages te zien samen met de respons functies van deze reeksen. Dit zijn de Gamma response functies zoals geoptimaliseerd door het tijdreeksmodel. Daarnaast is aan de afbeelding een tabel toegevoegd met daarin de geoptimaliseerde parameters met de bijbehorende standaarddeviaties. Hierin zijn bijvoorbeeld ook de geoptimaliseerde constante en verdampingsfactor weergegeven.

Afbeelding 10.37 Resultaten tijdreeksmodel

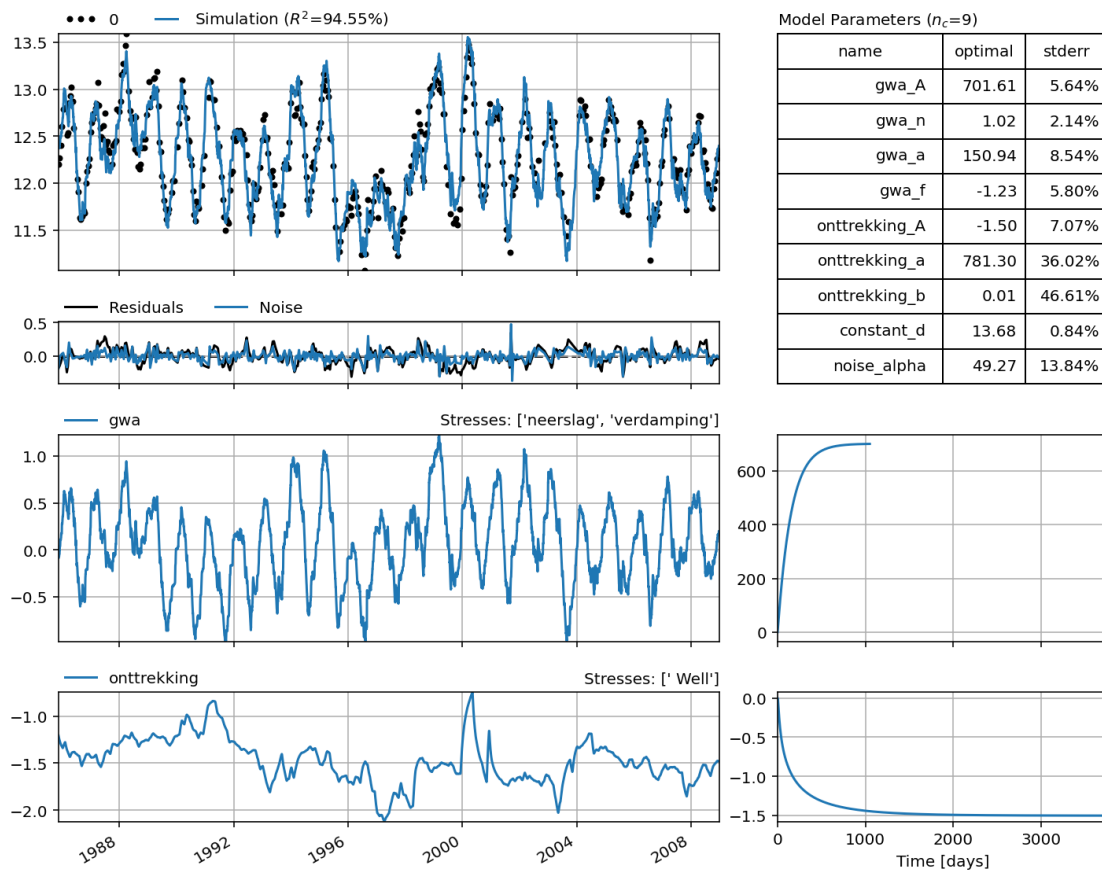


In de resultaten van het model valt op dat het tijdreeksmodel niet in staat is om de uitschieters in de meetreeks correct te simuleren. Na overleg met de beheerder van het gebied waar Peilbuis 1 in staat blijkt dat de metingen boven de +14.5 m NAP onrealistisch zijn, deze komen niet overeen met de observaties in het veld. Mogelijk is de peilbuis gedurende deze dagen volgelopen met water waardoor hogere waarden zijn geregistreerd.

Deze uitschieters kunnen het tijdreeksmodel mogelijk negatief beïnvloeden omdat deze metingen niet verklaard kunnen worden met behulp van de verklarende reeksen die worden gebruikt. Daarom wordt het tijdreeksmodel opnieuw gemaakt maar worden nu enkel stijghoogte meegenomen van < +14.5 m NAP.

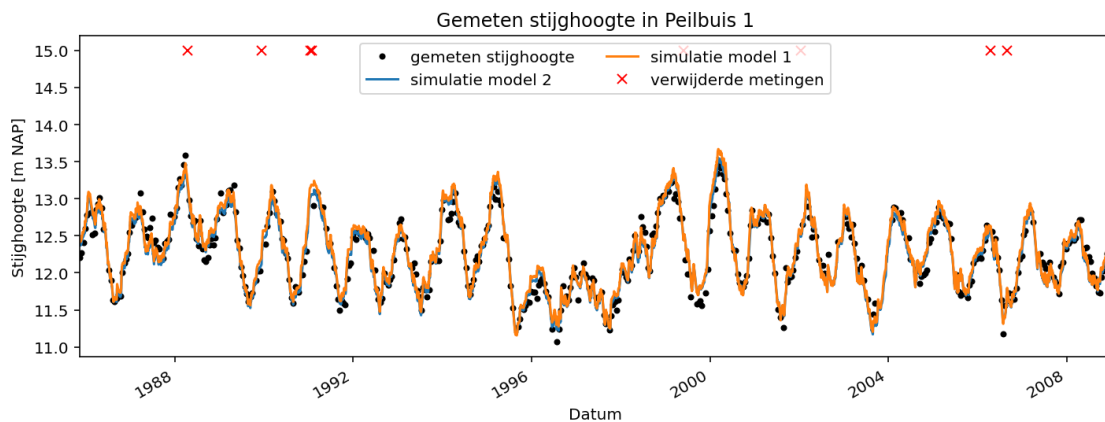
De resultaten van het nieuwe tijdreeksmodel zijn te zien in de onderstaande afbeelding.

Afbeelding 10.38 Resultaten tijdreeksmodel zonder uitbijters



Het tweede tijdreeksmodel, waarbij de metingen van > 14.5 m NAP niet zijn meegenomen, heeft een hogere verklaarde variantie en R^2 , en een lagere root mean square error (RMSE). In de onderstaande afbeelding is de gesimuleerde stijghoogtereeks van het nieuwe tijdreeksmodel (zonder uitbijters en onrealistische metingen) vergeleken met het eerste tijdreeksmodel. Hierbij is te zien dat op basis van een visuele inspectie niet te zien is welk tijdreeksmodel beter presteert, op het oog is het een vergelijkbare prestatie.

Afbeelding 10.39 Vergelijking simulatie met en zonder uitbijters



In eerste instantie zijn voor de verklarende reeksen de Gamma en Hantush respons functies gekozen. Het is aan de modellieur om de keuze voor de respons functie te maken. Daarom wordt het model nogmaals opgesteld met verschillende respons functies om deze te vergelijken. Hierbij wordt voor beide verklarende reeksen de volgende respons functies gebruikt: Exponentieel, Gamma en de Hantush respons.

In de onderstaande tabellen zijn de R^2 en RMSE van de modellen weergegeven. Langs de rijen van de tabellen staat de respons functie van de neerslag en verdamping en langs de kolommen de respons functie van de onttrekking.

Tabel 10.5 Vergelijking tijdreeksmodellen met verschillende responsfuncties

Model R^2	Respons onttrekking		
	Exponential	Gamma	Hantush
Respons grondwateraanvulling			
Exponential	0.94	0.94	0.94
Gamma	0.94	0.94	0.95
FourParam	0.94	0.93	0.94
Model RMSE	Respons onttrekking		
	Exponential	Gamma	Hantush
Respons grondwateraanvulling			
Exponential	0.12	0.12	0.11
Gamma	0.12	0.12	0.11
FourParam	0.12	0.13	0.12

Zoals te zien in de tabellen zijn er verschillende combinaties voor de respons functies die gelijkwaardige model prestaties geven. Te zien is dat het model de hoogste R^2 en laagste RMSE heeft wanneer voor de respons van de grondwateraanvulling de Gamma respons wordt gekozen. De keuze van de respons voor de onttrekking is minder van invloed op de modelresultaten.

Er wordt gekozen om de eerdere keuze voor de Gamma respons voor de grondwateraanvulling en de Hantush respons voor de

Toetsen van het tijdreeksmodel

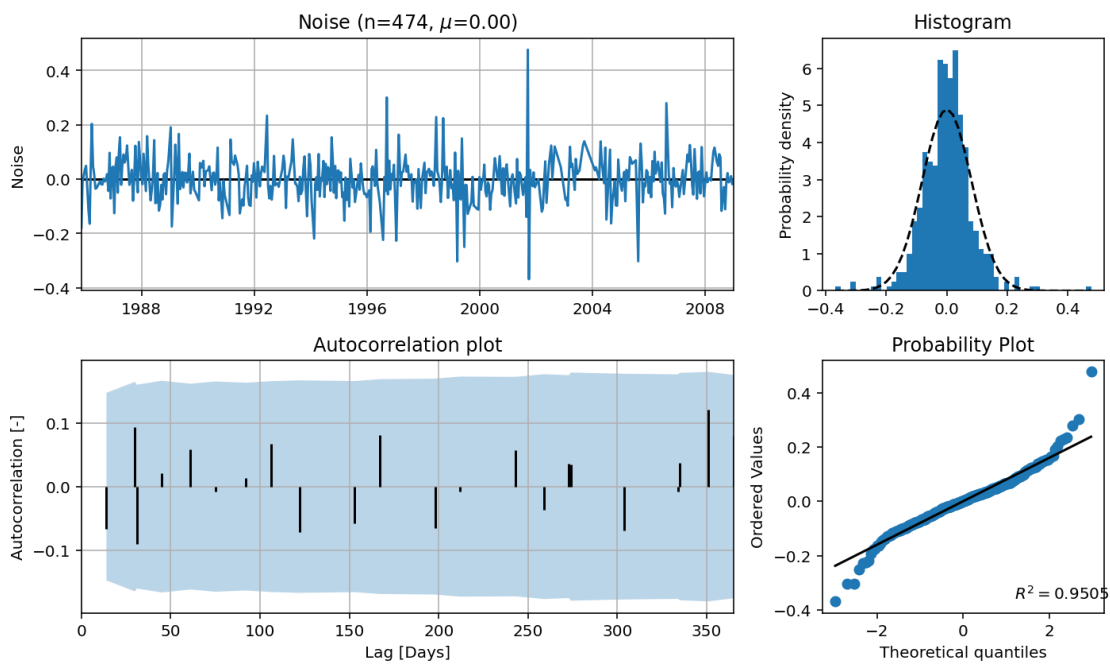
Voor het maken van een voorspelling is het nodig iets te zeggen over de onzekerheid van het tijdreeksmodel, uitgedrukt in de standaarddeviatie van de modelparameters. Om dit te doen moet worden gecontroleerd of de resultaten van het tijdreeksmodel voldoende zijn om de onzekerheid goed in te kunnen schatten.

Bij het bepalen van de onzekerheid van de modelparameters is de aanname dat de ruis van het model zogenaamde 'witte ruis' is. Hiervoor worden de volgende punten gecontroleerd:

- het gemiddelde van de ruis is nul;
- de waardes zijn onafhankelijk van elkaar;
- de ruis is normaal verdeeld;
- de ruis heeft een constante variantie.

Als de ruis aan deze toetsen voldoet mag aangenomen worden dat de ruis 'witte ruis' betreft en dat de standaard deviatie van de modelparameters juist is ingeschat.

Afbeelding 10.40 Diagnostische toetsen tijdreeksmodel



In de afbeelding linksboven is de ruis weergegeven. Te zien is dat er geen duidelijke trend zichtbaar is en dat de gemiddelde waarde (μ) gelijk is aan 0,00. In de figuur linksonder is de autocorrelatie van de lag weergegeven met het bijbehorende 95 % zekerheidsinterval. Hiervoor geldt dat 95 % van de autocorrelatie waarden binnen dit interval moeten liggen. Zoals te zien in deze plot geldt dit voor de huidige reeks met ruis.

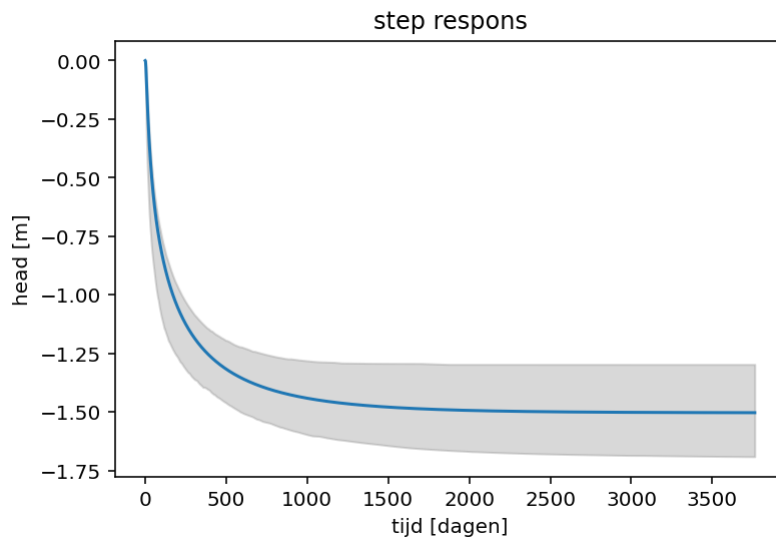
De afbeelding rechtsboven weergeeft de ruis als verdeling, met een bijbehorende normale verdeling, gefit op de ruis. Deze afbeelding kan worden gebruikt om te kijken of de ruis aan een normale verdeling voldoet. De afbeelding rechtsonder kan ook worden gebruikt om de normaliteit te toetsen. In dit geval voldoet de ruis goed aan de eis van de normale verdeling.

Voorspellen van de stijghoogte

Nu het tijdreeksmodel is opgesteld kan het gebruikt worden om het effect van het afbouwen van de onttrekking te voorspellen.

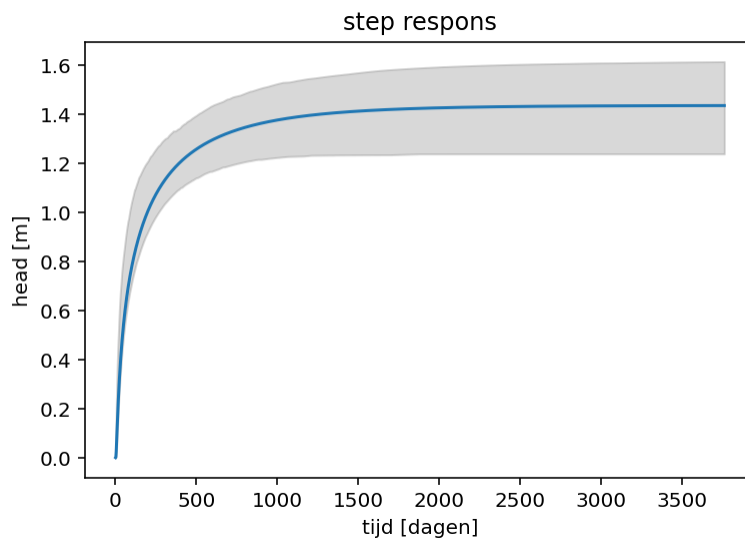
Hiervoor wordt gekeken naar de respons functie van de onttrekking. Deze is weergegeven in de onderstaande afbeelding. Hierbij is ook de bandbreedte van de respons functie te zien met een 95 % zekerheidsinterval. Deze is bepaald met behulp van een Monte Carlo Sampling op basis van de resultaten van de tijdreeksanalyse. De respons functie stabiliseert na circa 3.600 dagen. De 'gain' (niveau waarop de respons functie stabiliseert) van de respons functie is gelijk aan circa -1.6 m, de verlaging die optreedt ten gevolge van permanente een onttrekking van 10.000 m³/d. Voor een constante onttrekking van bijvoorbeeld 1.000 m³/d zou de stijghoogte naar verwachting 0.16 meter dalen.

Afbeelding 10.41 Staprespons met 95% betrouwbaarheidsinterval



Met behulp van deze responsfunctie kan de verwachte stijging van de stijghoogte ten gevolge van het uitzetten van de onttrekking worden voorspeld. De gemiddelde onttrekking voor de periode tot 2009-11-05 bedraagt $9544 \text{ m}^3/\text{d}$. De respons ten gevolge van de verandering van gemiddeld $9544 \text{ m}^3/\text{d}$ naar $0 \text{ m}^3/\text{d}$ is te zien in de onderstaande afbeelding, hierbij is de respons, zoals deze is bepaald door middel van het tijdreeksmodel, geschaald naar de gemiddelde onttrekking. De gemiddelde verwachte stijging is tussen de 1.25 en 1.63 meter, waarbij het effect van de onttrekking na 3.600 dagen volledig is uitgedempt (circa 10 jaar).

Afbeelding 10.42 Verwachte stijging tgv gemiddelde onttrekking

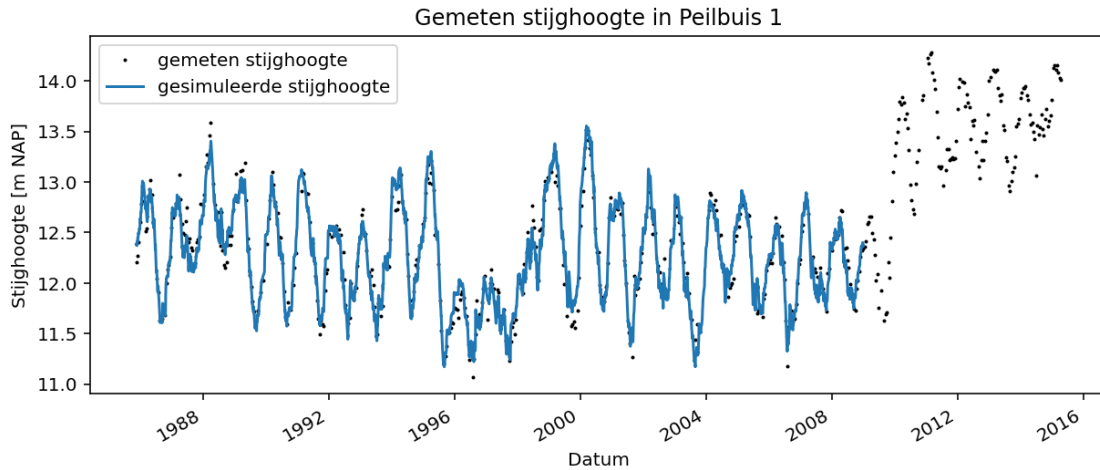


Controle voorspelling

Een aantal jaar later besluit het waterschap te kijken of de voorspelling van het adviesbureau voor de stijghoogte is uitgekomen. Hierbij worden de meeste recente metingen van de grondwaterstand gebruikt. De volledige reeks is te zien in de onderstaande afbeelding. Hierin is te zien dat de grondwaterstand sinds 2009 sterk is toegenomen.

In de afbeelding is daarnaast de simulatie te zien waarbij de meeste recente data is gebruikt voor de neerslag en verdamping. Hiervoor is hetzelfde tijdreeksmodel gebruikt zoals deze is opgesteld voor het uitvoeren van de voorspelling.

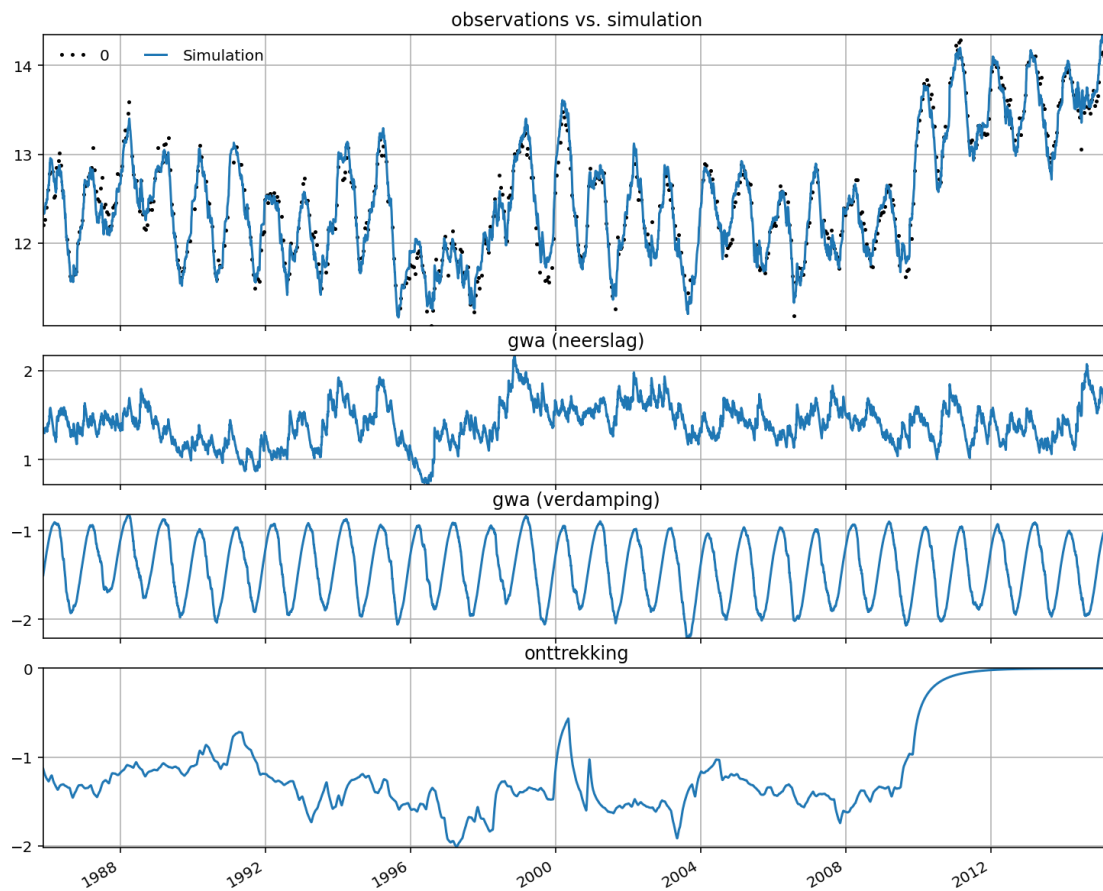
Afbeelding 10.43 Gemeten stijghoogte met periode na 2009



Er wordt een nieuw tijdreeksmodel opgesteld met de meest recente data. De meetreeksen voor de neerslag en verdamping en de onttrekking zijn aangevuld. Voor de verklarende reeksen worden dezelfde respons functies gekozen zoals in de vorige analyse.

De simulatie van het tijdreeksmodel met de volledige meetreeks is te zien in de onderstaande afbeelding. Daarnaast zijn de bijdrages van de verschillende verklarende reeksen te zien. Voor de onttrekking is te zien dat de bijdrage van deze verklarende reeks afneemt na 2009 (wanneer de onttrekking wordt uitgezet). Hieruit is ook af te leiden dat het effect hiervan langzaam uitdempt en niet instantaan optreedt.

Afbeelding 10.44 Resultaten tijdreeksmodel met bijdrage van invloeden

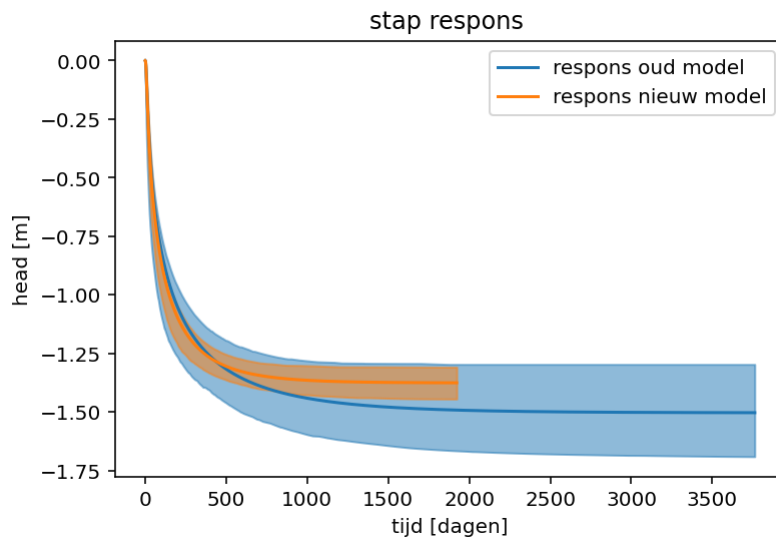


De step responses van beide modellen zijn weergegeven in de onderstaande afbeelding. Hierin is te zien dat de respons functies verschillen. Met de nieuwe respons functie ligt de bandbreedte van het gemiddelde effect van het uitzetten van de onttrekking tussen de 1.25 en 1.37 meter tegenover de bandbreedte tussen de 1.24 en 1.61 meter van het oude tijdreeksmodel. Daarnaast zal volgens het nieuwe tijdreeksmodel het effect van de onttrekking ook sneller uitdempen.

In de onderstaande afbeelding is te zien dat de respons zoals bepaald met het nieuwe tijdreeksmodel binnen de bepaalde bandbreedte van het oude tijdreeksmodel valt. De bandbreedte van de nieuwe respons valt ook binnen de bandbreedte van het oude tijdreeksmodel.

Uit deze resultaten blijkt dat de voorspelling van de verhoging ten gevolge van het uitzetten van de onttrekking past bij de huidige metingen. Met het nieuwe tijdreeksmodel is de onzekerheid van de respons verkleind.

Afbeelding 10.45 Vergelijking verwachte invloed gemiddelde onttrekking



10.4.2 Deel II: Herstel grondwaterstand in april

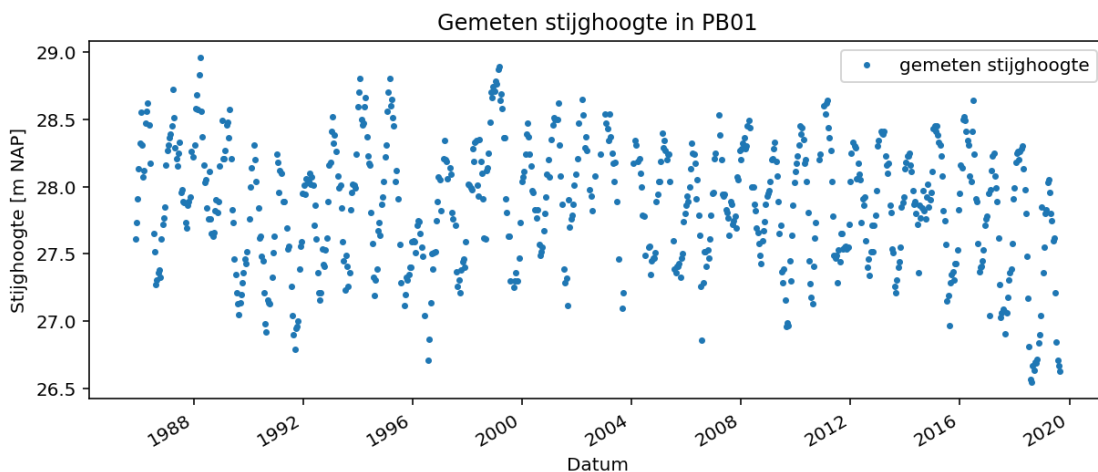
In een natuurgebied wordt de grondwaterstand al jaren gemeten in PB01. In de afgelopen jaren (2018 en 2019) merkt het waterschap dat de grondwaterstand in de zomerperiode lager ligt dan gemiddeld. Ook bij de laatste metingen in augustus 2019 was dit het geval. Aan de hydroloog van het waterschap is gevraagd of een inschatting kan worden gegeven voor de grondwaterstand bij de start van het nieuwe groeiseizoen in april volgend jaar (april 2020).

Op basis van een tijdreeksanalyse wordt gekeken naar de verwachte grondwaterstand ten gevolge van een 'natte', 'droge' en een 'gemiddelde' winterperiode (van september tot april). Zo hoop het waterschap inzicht te krijgen in het herstel van de stijghoogte onder verschillende omstandigheden.

Beschikbare data voorbeeld 2

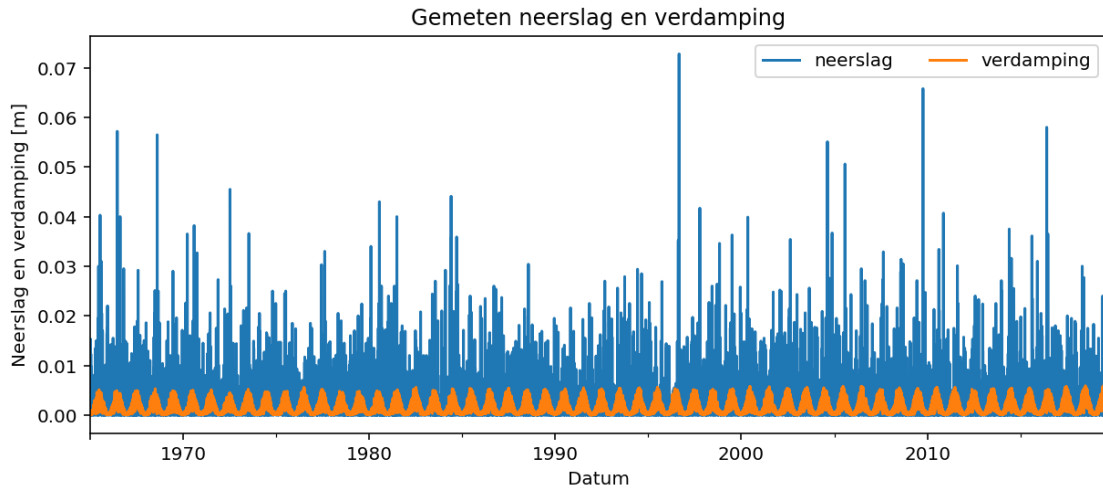
In de onderstaande afbeelding is de meetreeks van peilbuis PB01 te zien. De meetreeks loopt van 1985 tot september 2019. De variatie in de grondwaterstand is ca. 2 meter. In de reeks is te zien dat de stijghoogte in de zomer van 2018 en 2019 relatief laag is geweest vanwege de relatief droge periodes.

Afbeelding 10.46 Gemeten stijghoogte PB01



De data voor de neerslag en verdamping nabij PB01 zijn verzameld van het dichtstbijzijnde KNMI weerstation. De neerslag en verdamping zijn weergegeven in de onderstaande grafiek. Het weerstation bevat data vanaf 1965 tot en met september 2019.

Afbeelding 10.47 Gemeten neerslag en verdamping

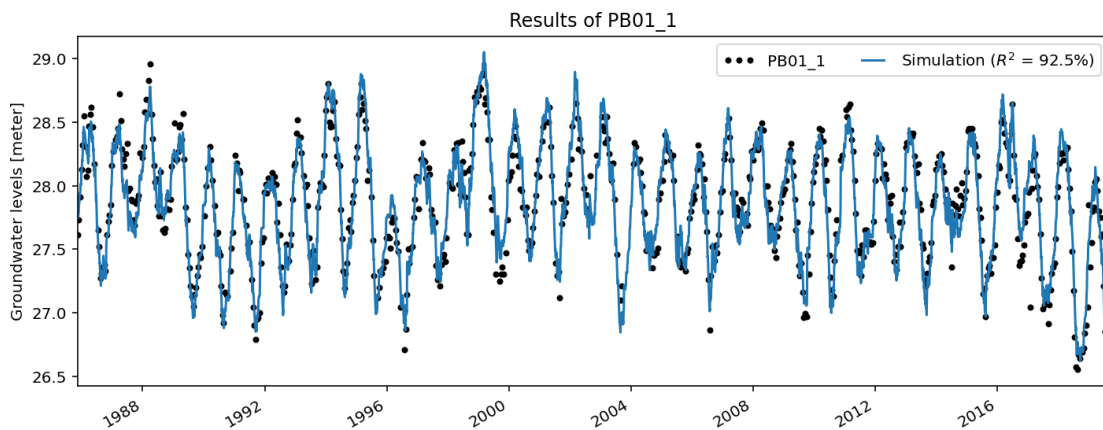


Opzetten tijdreeksmodel

Er wordt een tijdreeksmodel opgesteld op basis van de meetreeks in PB01. Hiervoor worden de neerslag en verdamping gebruikt als verklarende reeks. Voor de neerslag en verdamping wordt de Gamma respons functie gekozen in de tijdreeksanalyse.

Het tijdreeksmodel heeft een R^2 van 0.92. In de onderstaande afbeelding is de simulatie van het tijdreeksmodel te zien samen met de stijghoogte metingen. De variatie in stijghoogte lijkt goed verklaard te kunnen worden met behulp van de neerslag en verdamping. Het tijdreeksmodel is gesimuleerd tot aan 1 september 2019.

Afbeelding 10.48 Simulatie met tijdreeksmodel



Voorspelling neerslagoverschot

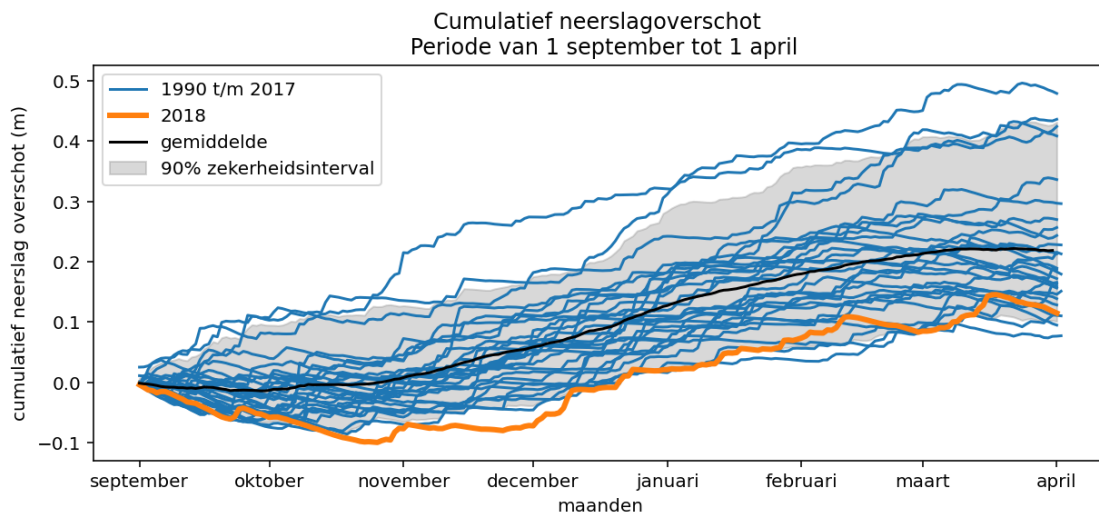
Om een inschatting te maken van de stijghoogte ten gevolge van de 'natte', 'droge' en 'gemiddelde' winterperiode wordt gekeken naar het neerslagoverschot van de afgelopen 30 jaar.

Het neerslagoverschot wordt berekend op basis van de neerslag, verdamping en de verdampingsfactor. Bij het gebruik van de verdampingsfactor wordt de onzekerheid van deze parameters niet meegenomen in de analyse, de onzekerheid van het bepaalde neerslagoverschot wordt daarom ook niet meegenomen. In de onderstaande afbeeldingen zijn de neerslagoverschotten van de afgelopen 30 jaar weergegeven, waarbij de jaren 2018 en 2019 zijn gemarkeerd. De aanname bij het bepalen van deze neerslagoverschotten voor de afgelopen 30 jaar is dat de relatie tussen neerslag, verdamping en grondwateraanvulling niet is veranderd en dat deze relatie ook in droge of natte jaren gelijk is gebleven.

In het overzicht valt op dat, voor de periode september-april, het neerslagoverschot in 2018 tot een van de laagste van de afgelopen 30 jaar behoort.

Voor het 'natte' jaar wordt de 95 % bovengrens van het neerslagoverschot aangehouden. Voor het droge jaar de 5 % ondergrens. Het 'gemiddelde' jaar wordt op basis van het gemiddelde neerslagoverschot berekend. Deze scenario's zijn berekend op basis van de dagelijkse neerslagoverschotten, per dag is de 95 %, 5 % of gemiddelde waarde voor het neerslagoverschot bepaald. De bandbreedte die dit oplevert is te zien in de onderstaande afbeelding.

Afbeelding 10.49 Cumulatief neerslagoverschot voor de periode van 1 september tot 1 april voor 1990 - 2018

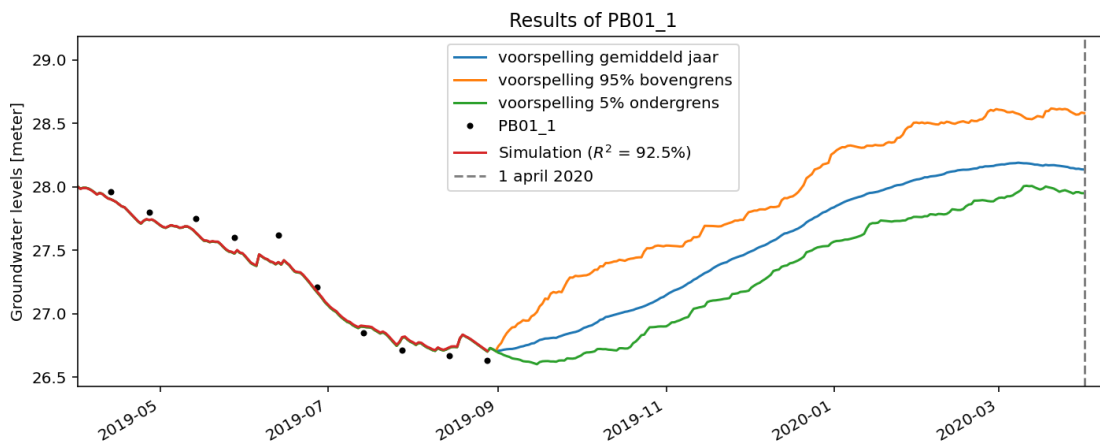


Voorspellen stijghoogte

Voor de 3 scenario's wordt met behulp van de resultaten van het tijdreeksmodel de stijghoogte bepaald voor 1 april 2020. Hierbij wordt er vanuit gegaan dat het systeem niet verandert, de respons gelijk blijft en dat er geen extra invloeden op de stijghoogte bijkomen in de simulatieperiode.

Met behulp van de responsfunctie uit het tijdreeksmodel wordt het verloop van de stijghoogte voor de periode 1 september 2019 tot 1 april 2020 berekend. In de onderstaande afbeelding zijn voor de 3 scenario's de voorspellingen van de stijghoogte te zien. Deze geven 3 schattingen van de grondwaterstand voor de periode tot 1 april 2020. Met de voorspelling kan worden gekeken of de stijghoogte bij een 'droog', 'nat' of 'gemiddeld' winterseizoen hersteld na een droge periode.

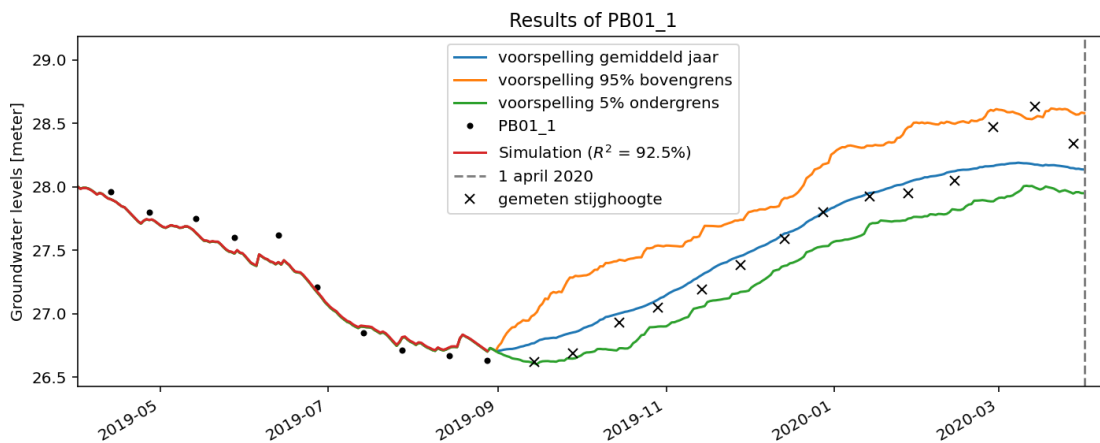
Afbeelding 10.50 Voorspelling stijghoogte



Na 1 april 2020 wordt door het waterschap de meest recente data van peilbuis PB01 verzameld. Deze kunnen worden vergeleken met de voorspelling gemaakt in september. In de onderstaande afbeelding zijn de daadwerkelijk gemeten stijghoogtes samen met de voorspelling weergegeven.

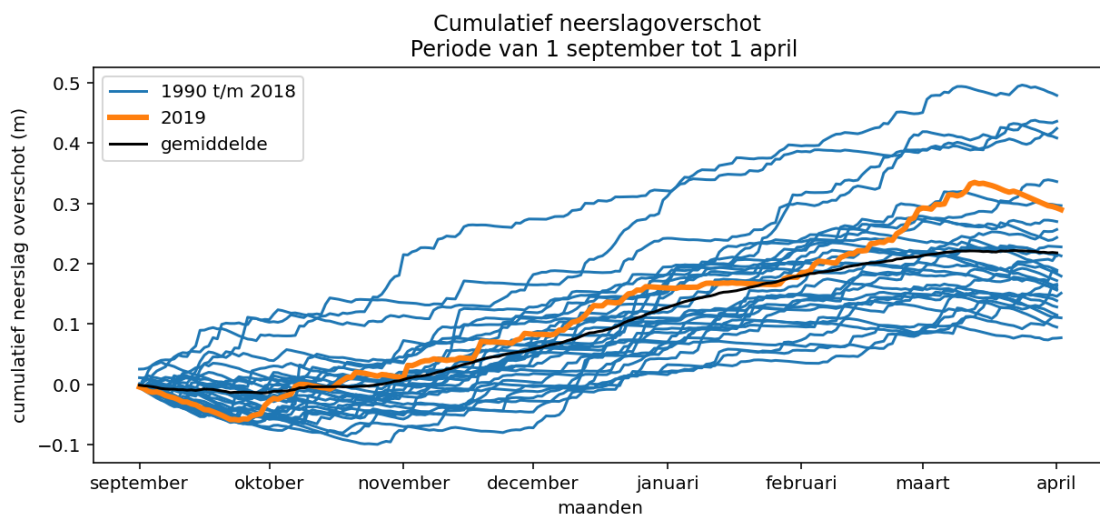
De daadwerkelijk gemeten stijghoogte in de periode september-oktober ligt lager dan het gemiddelde scenario. In de periode vanaf maart 2020 is de gemeten stijghoogte hoger dan het gemiddelde scenario. In het algemeen vallen de metingen binnen de breedte van de 3 scenario's.

Afbeelding 10.51 Voorspelling stijghoogte in vergelijking met metingen



In de onderstaande afbeelding is voor de periode september 2019 - april 2020 het neerslagoverschot vergeleken met de neerslagoverschotten van de overige jaren. Dit neerslag overschot is op dezelfde manier bepaald als voor de overige jaren. Hierbij is te zien dat voor dit jaar in september-oktober het overschot lager is dan gemiddeld en in de periode maart-april hoger. Dit beeld komt overeen met de vergelijking tussen de metingen en de 3 scenario's.

Afbeelding 10.52 Cumulatief neerslag overschot 1 september tot 1 april inclusief 2019



Bijlage: Theoretische achtergrond tijdreeksanalyse

11

INTRODUCTIE STATISTIEK VOOR TIJDREEKSANALYSE

11.1 Inleiding

Statistiek kan een belangrijk hulpinstrument zijn bij tijdreeksanalyse, omdat het niet alleen kan helpen het tijdreeksmodel te identificeren, maar ook in staat stelt de modelonzekerheid in beeld te brengen en uitspraken over het beschouwde systeem te doen met gekwantificeerde onzekerheden. Dit hoofdstuk geeft een introductie in de statistische methoden die meerwaarde kunnen verlenen aan tijdreeksanalyse.

11.2 Onzekerheid hoort bij modelleren

Het is vrijwel onmogelijk om een tijdreeks met waarden van een bepaald hydrologisch proces exact te modelleren, dat wil zeggen zonder modelafwijkingen. Dit komt vooral door data- en modelbeperkingen. Zo zal het zelden mogelijk zijn alle invloedsfactoren van het beschouwde systeem in het model op te nemen, omdat ze niet alle worden gemeten. Een modelbeperking kan ontstaan als een ruimtelijk proces wordt vereenvoudigd tot een puntproces. Zo moet in een tijdreeksmodel de ruimtelijk verdeelde neerslag op het hydrologische systeem worden vereenvoudigd tot de geregistreerde neerslag op een nabijgelegen neerslagstation. En een winplaats met verspreid liggende pompputten, die niet synchroon pompen, wordt vereenvoudigd tot een fictieve puntonttrekking, bestaande uit één pompput. Een andere modelbeperking kan zijn dat alleen lineaire relaties kunnen worden verdisconteerd, terwijl er vaak ook sprake kan zijn van niet-lineaire relaties. Door dergelijke beperkingen zijn modelafwijkingen onvermijdelijk en ontstaat modelonzekerheid. Tijdreeksanalyse maakt het echter mogelijk om deze onzekerheid te kwantificeren en mee te geven bij het presenteren van de resultaten. Zo stellen we de klant of gebruiker in staat beter onderbouwde besluiten te nemen dan alleen op basis van de resultaten. Het verschaft meerwaarde aan de resultaten van tijdreeksanalyse.

Steekproef versus populatie

De meetwaarden die verkregen wordt vormen doorgaans slechts een steekproef uit de verzameling van alle mogelijke meetwaarden die van het beschouwde verschijnsel (zoals de grondwaterstand) verkregen kan worden. Onze belangstelling zal echter vooral uitgaan naar die laatste verzameling, die ook wel wordt aangeduid als de *populatie*. Kennis van de populatie betekent immers een breder en algemener inzicht dan kennis van een steekproef uit die populatie. Beschrijvende statistiek kan ons helpen de informatie over de populatie van alle mogelijke meetwaarden op efficiënte wijze te onthullen en samen te vatten, op basis van de meetwaarden van de steekproef. Dit is dan uiteraard kostenbesparend, maar het is onvermijdelijk dat hierdoor enige onzekerheid blijft bestaan over de populatie.

Het belang van de kansverdeling

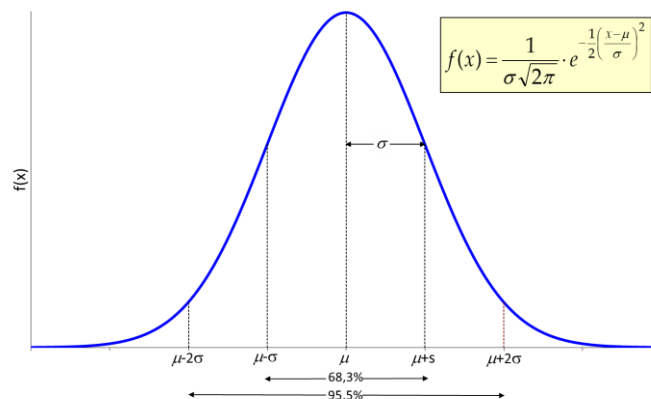
De beste manier om een populatie meetwaarden in zijn totaliteit te kenschetsen is door zijn kansverdeling. Als metingen gedaan worden van een bepaald verschijnsel zal iedere meetwaarde een bepaalde kans hebben om gerealiseerd te worden. De functie die voor elke mogelijke meetwaarde de kans op realisatie geeft noemen we de kansverdeling.

Als we de kansverdeling van de meetwaarden van een bepaald experiment al vóóraf kennen, zoals op basis van de theorie, hoeven we eigenlijk geen metingen meer te doen, we kennen dan immers al alle eigenschappen van de hypothetische verzameling meetwaarden. In de praktijk is dit echter zelden het geval - experimenten zoals met de dobbelsteen of met de munt daargelaten - en zullen we de eigenschappen van de meetwaarden moeten afleiden op basis van een steekproef van meetwaarden. Maar zodra we daarmee voldoende inzicht in de kansverdeling van de meetwaarden hebben verkregen, zijn verdere experimenten in principe overbodig geworden. Er is dan namelijk al genoeg kennis om allerlei soorten uitspraken te kunnen doen over het door ons bestudeerde deel van de werkelijkheid, zij het met bijbehorende onzekerheidsbanden.

Alhoewel er in theorie natuurlijk oneindig veel soorten kansverdelingen kunnen voorkomen, blijken enkele zeer vaak voor te komen. Zo blijken veel continue variabelen een klokvormige kansverdeling te hebben, de zogenaamde normale kansverdeling, ook wel aangeduid als de Gauss-verdeling. Er is inmiddels veel theoretische kennis beschikbaar over de eigenschappen van deze kansverdeling en daar zijn dan ook veel bekende statistische methoden op gebaseerd, zoals de t-toets, lineaire regressieanalyse en variantie-analyse.

De normale kansverdeling is symmetrisch van vorm en kan gezien zijn formule al volledig worden gekarakteriseerd zodra we zijn gemiddelde en standaardafwijking¹ kennen (zie Afbeelding 11.1). Dus als we er van uit mogen gaan dat meetwaarden uit een normale kansverdeling komen, kunnen we volstaan met het schatten van het gemiddelde en de standaardafwijking, om vervolgens allerlei soorten uitspraken te kunnen doen over de verzameling van alle mogelijke meetwaarden (de populatie).

Afbeelding 11.1 De normale kansverdeling is volledig te karakteriseren door zijn gemiddelde (μ) en standaardafwijking (σ). De horizontale as betreft de waarden van de verdeling (x) en de verticale as betreft de kansdichtheid ($f(x)$). Het interval $\mu \pm \sigma$ omvat 68,3 % van de mogelijke waarden en het interval $\mu \pm 2\sigma$ omvat 95,5 % van de mogelijke waarden (het interval $\mu \pm 1,96\sigma$ omvat precies 95 % van de mogelijke waarden)



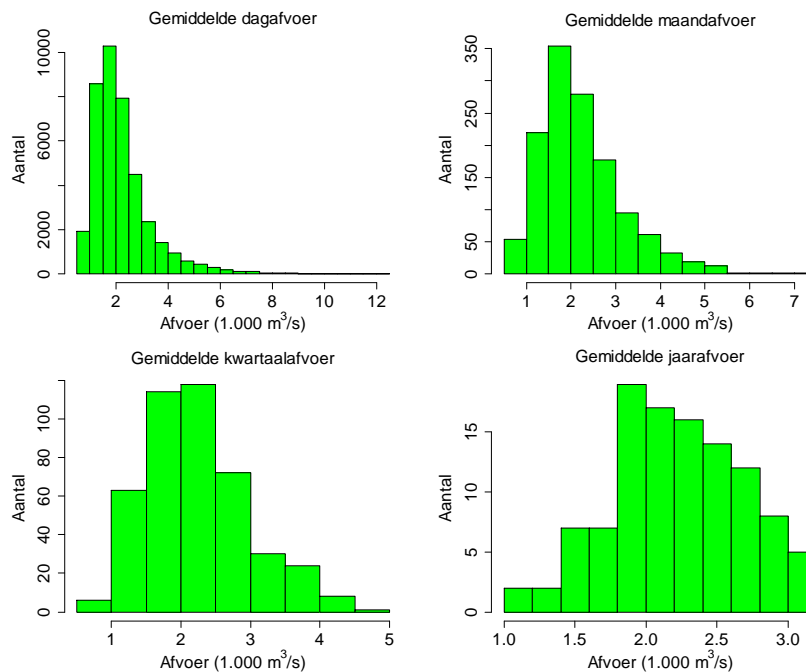
Er zijn diverse redenen waarom de normale kansverdeling vaak voor het voetlicht komt:

- veel variabelen zijn normaal verdeeld, of kunnen door een wiskundige transformatie normaal verdeeld worden gemaakt (zoals de logaritmische transformatie);
- als variabelen worden gesommeerd of gemiddeld zal de verdeling van de som of het gemiddelde meer naderen tot de normale kansverdeling naarmate er meer waarden worden gesommeerd of gemiddeld, ongeacht de kansverdeling van de waarden. Deze belangrijke bevinding uit de kansrekening wordt aangeduid als de Centrale Limietstelling. Zo wordt de verdeling van de afvoer van de Rijn bij Lobith duidelijk minder scheef naarmate over grotere tijdeenheden wordt gemiddeld (zie Afbeelding 11.2);

¹ De standaardafwijking is een kwantitatieve maat voor de spreiding van een groep waarden. Als de waarden aselechte trekkingen vormen uit een populatie kan daaruit de standaardafwijking van de populatie worden geschat. In Excel kan dit met de formulering: = STDEV.S(cellenbereik). De formulering = STDEV.P(cellenbereik) is alleen van toepassing als alle waarden van de populatie beschikbaar zijn. Bij hydrologisch modelleren zal dat laatste niet opgaan.

- als de meetwaarden afkomstig zijn uit een zeer scheve kansverdeling moet er nog wel worden gemiddeld over veel meetwaarden (75 à 100) om dit te bewerkstelligen, maar naarmate de meetwaarden afkomstig zijn uit een meer symmetrische kansverdeling reduceert het hiervoor benodigde aantal sterk;
- er is inmiddels veel theoretische kennis over deze verdeling beschikbaar.

Afbeelding 11.2 De verdeling van de gemiddelde afvoer van de Rijn bij Lobith wordt minder scheef naarmate over grotere tijdeenheden wordt gemiddeld. Dit is conform de Centrale Limietstelling, die stelt dat de verdeling van een gemiddelde meer zal naderen tot de normale kansverdeling naarmate meer waarden worden gemiddeld, ongeacht de kansverdeling van de waarden



Onzekerheidsmaten van een schatting: standaardfout en betrouwbaarheidsinterval

Een kengetal – zoals gemiddelde, of mediaan - dat we schatten uit de meetwaarden van een steekproef is op te vatten als een trekking uit de denkbeeldige populatie van geschatte kengetallen. Afhankelijk van de toevallige samenstelling van meetwaarden in de steekproef zal er namelijk een andere schatting van het kengetal resulteren. Dus als we oneindig vaak een steekproef van n meetwaarden zouden nemen uit de populatie van meetwaarden en daaruit iedere keer het gemiddelde zouden schatten (\bar{x}), volgens:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

ontstaat een populatie van gemiddelden geschat uit n meetwaarden. Een dergelijke populatie van schattingen – die ook wel wordt aangeduid als de *schatting* - kan worden gekenschetst door zijn gemiddelde en zijn standaardafwijking. In het geval van een schatter wordt die laatste aangeduid als standaardfout¹. Deze is op te vatten als een maat voor de onzekerheid die inherent is aan de schatting. Vanuit de theorie volgt dat de standaardfout van een geschat gemiddelde ($\sigma_{\bar{x}}$) omgekeerd evenredig is aan de grootte (n) van de steekproef waaruit het gemiddelde is berekend, mits de steekproef bestaat uit aselechte trekkingen uit de populatie:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

¹ Waar het gaat om spreiding spreken we in het geval van meetwaarden van *standaardafwijking* en in het geval van een uit de meetwaarden geschat populatiekenmerk van *standaardfout*. In dit tweede geval is de afwijking immers per definitie op te vatten als een fout, maar in het eerste geval niet.

waarin σ de standaardafwijking van de populatie. Maar in de praktijk komen we doorgaans niet verder dan een schatting van de standaardafwijking van de populatie, zodat we ook de standaardfout van een geschat gemiddelde slechts kunnen schatten, volgens:

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

waarin s de geschatte standaardafwijking van de populatie.

Stel bijvoorbeeld dat we de gemiddelde grondwaterdiepte van een bepaald gebied op een bepaald tijdstip willen bepalen en daartoe in dat gebied op 20 aselekt gekozen punten met boorgaten de grondwaterdiepte vaststellen. De gemiddelde diepte blijkt 63 cm en de standaardafwijking van de diepte blijkt 17 cm. De standaardfout van dat gemiddelde volgt dan uit:

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{17}{\sqrt{20}} = 3,8 \text{ cm}$$

Als de onderzochte populatie een normale kansverdeling volgt (en de meetwaarden daaruit aselekt zijn getrokken), kunnen we vervolgens ook het 95%-betrouwbaarheidsinterval van het geschatte gemiddelde schatten, met als grenzen:

$$\bar{x} \pm t_{(97,5\%;n-1)} \cdot s_{\bar{x}}$$

waarin $t_{(97,5\%;n-1)}$ het 97,5-percentiel¹ van de Student-t-verdeling met een onderschrijdingskans van 97,5 % bij $n-1$ vrijheidsgraden². Voor het boven beschreven voorbeeld van de gemiddelde grondwaterdiepte is dat:

$$\bar{x} \pm t_{(97,5\%;n-1)} \cdot s_{\bar{x}} = 63 \pm t_{(97,5\%;19)} \cdot 3,8 = 63 \pm 8,0 \text{ cm}$$

Het 95 %-betrouwbaarheidsinterval is op te vatten als het interval waarbinnen het werkelijk gemiddelde van de populatie (μ) in 95 van de 100 gevallen zal liggen, mochten we de steekproef en het daaruit schatten van gemiddelde, standaardfout en betrouwbaarheidsinterval van het geschatte gemiddelde oneindig vaak kunnen herhalen. Dit kunnen we ook wel formuleren als het interval waarbinnen dat werkelijk gemiddelde *bijna altijd* zal liggen.

Beoordelen op het soort kansverdeling

We kunnen op verschillende manieren nagaan uit wat voor soort kansverdeling onze meetwaarden afkomstig zijn. Een geschikte mogelijkheid om *visueel* te beoordelen uit wat voor soort kansverdeling de meetwaarden van een steekproef afkomstig zijn, wordt geboden door het histogram. Dit toont het aantal meetwaarden dat in een bepaalde grootteklasse valt, als functie van de grootteklasse. Bij het beoordelen van histogrammen van steekproeven dienen we ons echter wel te realiseren dat hun vorm sterk kan variëren met de steekproefgrootte.

Een aanvullend uitsluitsel over het al of niet voldoen aan een normale kansverdeling wordt verkregen door hierop statistisch te toetsen. Daar zijn meerdere toetsen voor beschikbaar, zoals de D'Agostino-K2-toets, de Jarque-Bera-toets, of de Shapiro-Wilk-toets. Maar er kleven praktische bezwaren aan deze toetsen. Zo is het onderscheidend vermogen³ gering als er weinig meetwaarden zijn, zoals minder dan circa 50. Afwijkingen van de getoetste kansverdeling kunnen dan dus niet snel worden gedetecteerd. Als er daarentegen veel meetwaarden zijn, is het onderscheidend vermogen dermate groot, dat ook al een praktisch gezien verwaarloosbare afwijking van de normale kansverdeling wordt gedetecteerd.

¹ Dit is de waarde met een onderschrijdingskans van 97,5 %.

² In Excel (Nederlandse uitvoering) kan deze worden berekend als $t_{(0,975;n-1)} = \text{T.INV}(0,975;n-1)$.

³ Het onderscheidend vermogen is de kans op het terecht verwerpen van de nulhypothese bij het toetsen. Deze kans neemt toe naarmate de werkelijkheid sterker afwijkt van de nulhypothese. Zie ook paragraaf 11.4.

11.3 Het begrip autocorrelatie

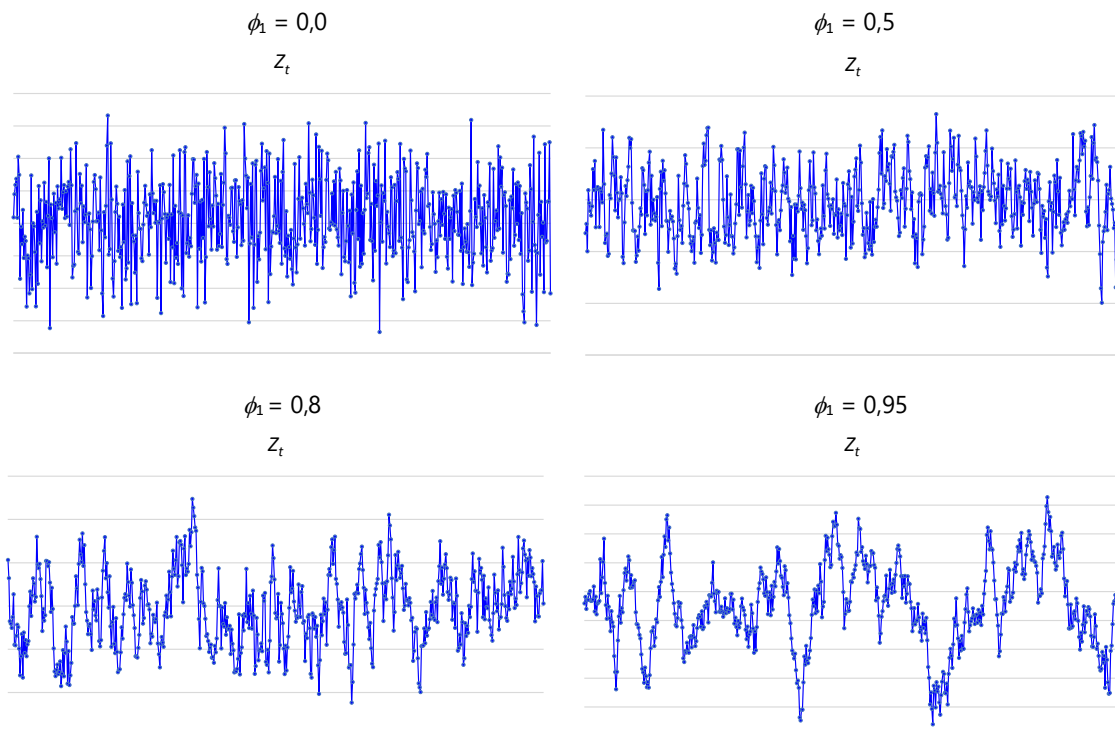
Er is sprake van autocorrelatie als opeenvolgende waarden van een tijdreeks niet onafhankelijk van elkaar zijn, maar aan elkaar gerelateerd zijn. Bij natuurlijke processen heeft autocorrelatie meestal de vorm dat opeenvolgende reekswaarden minder van elkaar verschillen dan reekswaarden met een groter tijdsverschil. Dit verschijnsel wordt aangeduid als positieve *autocorrelatie*.¹

Afbeelding 11.3 toont voorbeelden van tijdreeksen die verschillen in de mate van (positieve) autocorrelatie. De reeksen zijn gegenereerd volgens het volgende AR(1)-model:

$$Z_t = \phi_1 Z_{t-1} + a_t$$

waarin Z de uitvoerreeks, t de discrete tijdsindex, ϕ_1 de AR(1)-parameter en a het modelresidu, dat afkomstig is uit een normale kansverdeling, met gemiddelde nul. Uit deze afbeelding blijkt dat naarmate ϕ_1 en daarmee de autocorrelatie toeneemt, de reeks meer persistentie vertoont en de stochastische dynamiek vertraagt. Het tijdsverloop wordt als het ware stroperiger.

Afbeelding 11.3 Voorbeelden van tijdreeksen die verschillen in de mate van (positieve) autocorrelatie, aangebracht door de AR(1)-parameter ϕ_1 (zie formulering hierboven)



De autocorrelatie kan duidelijk in beeld worden gebracht middels de autocorrelatiefunctie (ACF), die de autocorrelatiecoëfficiënt (r) weergeeft als functie van het tijdsinterval. De autocorrelatiecoëfficiënt voor tijdsinterval k (r_k) wordt geschat als:

$$r_k = \frac{\sum_{t=1}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^n (z_t - \bar{z})^2}$$

¹ Bij negatieve autocorrelatie zijn opeenvolgende waarden doorgaans meer verschillend van elkaar dan waarden die willekeurig uit de tijdreeks zijn getrokken. Maar dit komt bij natuurlijke processen niet of nauwelijks voor, zodat *positieve autocorrelatie* in deze Handleiding meestal wordt aangeduid met de algemene term *autocorrelatie*.

waarin z_t de waarde van Z op tijdstip t en \bar{z} het geschatte gemiddelde van Z .¹ Meerdere voorbeelden hiervan zijn te vinden in paragraaf 4.4.2.

11.4 Statistisch toetsen

Een statistische toets is een procedure om de geldigheid van een hypothese objectief te verifiëren aan de hand van meetwaarden. In het nu volgende lichten we het principe van statistisch toetsen toe.

Het uitvoeren van een statistische toets vergt een aantal stappen. De stappen zijn hieronder beschreven voor het toetsen of de werkelijke waarde van een geschatte modelparameter afwijkt van nul.

- 1 Formuleer op basis van de probleemstelling de te toetsen nulhypothese (H_0) en eventueel ook de alternatieve hypothese (H_a). Als we willen toetsen of de werkelijke waarde van een geschatte modelparameter (θ) afwijkt van nul, ongeacht of deze positief of negatief is, luiden deze hypothesen:

$$H_0: \theta = 0 \quad \text{en} \quad H_a: \theta \neq 0$$

- 2 Stel vast welk toetsrisico op het onterecht verwerpen van de nulhypothese we nog accepteren. Dit wordt aangeduid als α , het significantieniveau, of ook wel het risico van de 1^e soort. Doorgaans wordt α gesteld op 5 % en soms op 1 %, onder meer afhankelijk van het effect dat het onterecht verwerpen van de nulhypothese kan hebben.
- 3 Selecteer de statistische toets die aansluit bij de kenmerken van de steekproef waar de schatting van de modelparameter op gebaseerd is. Voor het toetsen of een werkelijke modelparameter afwijkt van nul komt de t-toets voor één steekproef in aanmerking, op voorwaarde dat de modelresiduen (a_i) normaal verdeelde witte ruis vormen en ook dat residuen en ruis onafhankelijk zijn van de invoerreeksen.
- 4 Bereken de toetsingsgrootheid T . Bij tweezijdig toetsen met de t-toets voor één steekproef is dat de ratio van de absolute waarde van de geschatte modelparameter en zijn standaardfout, volgens:

$$T = \frac{|\hat{\theta}|}{s_{\hat{\theta}}}$$

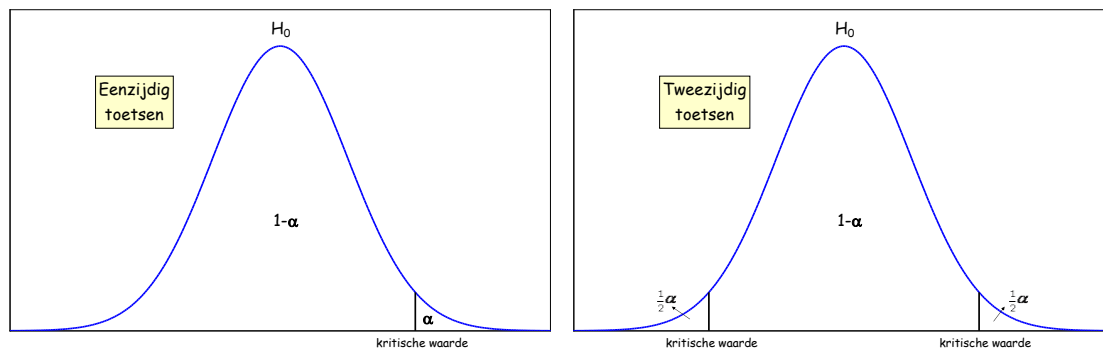
Gezien de tweezijdigheid van de alternatieve hypothese moet tweezijdig worden getoetst (zie de toelichting onder Afbeelding 11.4).

- 5 Stel vast of de toetsingsgrootheid al of niet past bij de nulhypothese, gegeven de kansverdeling van de toetsingsgrootheid onder de nulhypothese en het ingestelde significantieniveau. Voor ons voorbeeld zal de toetsingsgrootheid onder de nulhypothese afkomstig zijn uit de Student-t-kansverdeling.² Het vaststellen of de toetsingsgrootheid al of niet past bij de nulhypothese kan op twee manieren:
 - door de toetsingsgrootheid te vergelijken met de kritische waarde van de Student-t-verdeling die volgt uit het gewenste significantieniveau;
 - door de significantie (ook wel aangeduid als p -waarde) van de toetsingsgrootheid te bepalen en die te vergelijken met het significantieniveau (α). De significantie is de kans op een extremere toetsingsgrootheid dan degene die is gerealiseerd, bij geldigheid van de nulhypothese. Als de significantie kleiner is dan het vooraf ingestelde significantieniveau, kan de nulhypothese worden verworpen.Voorbeelden van deze twee manieren zijn uitgewerkt in paragraaf 4.4.1.
- 6 Concludeer op basis van het resultaat of de nulhypothese wel of niet wordt verworpen.

¹ Deze formulering gaat er van uit dat het proces dat Z_t genereert stationair is, in de zin dat zijn gemiddelde en variantie tijdsafhankelijk zijn.

² Statistische handboeken bevatten tabellen waarin voor de kansverdelingen van de gebruikelijkste toetsingsgrootheden, zoals de standaardnormale-, de Student-t-, de χ^2 - of de F-verdeling, de kritische waarden zijn vermeld die gelden bij verschillende significantieniveaus. Maar deze zijn inmiddels ook in Excel te herleiden.

Afbeelding 11.4 Voorbeeld van kansverdeling van de toetsingsgrootte onder de nulhypothese en de kritische waarde(n), respectievelijk bij eenzijdig toetsen (links) en bij tweezijdig toetsen (rechts)



Toelichting op Afbeelding 11.4

Afhankelijk van het soort alternatieve hypothese kan onderscheid worden gemaakt tussen eenzijdig en tweezijdig toetsen. Als de alternatieve hypothese ook het teken van een eventuele afwijking specificceert past eenzijdig toetsen en anders past tweezijdig toetsen. Bij eenzijdig toetsen is er één kritische waarde. Deze bevindt zich links of rechts, afhankelijk van hoe de alternatieve hypothese is geformuleerd. De nulhypothese wordt verworpen als de berekende toetsingsgrootte extremer is dan deze kritische waarde. Bij tweezijdig toetsen zijn er twee kritische waarden en de nulhypothese wordt dan verworpen als de berekende toetsingsgrootte extremer is dan hetzij de kritische waarde links, óf de kritische waarde rechts. Bij de combinatie van tweezijdig toetsen en een symmetrische kansverdeling van de toetsingsgrootte is het gebruikelijk de nulhypothese te verwerpen als de absolute waarde van de berekende toetsingsgrootte extremer is dan de kritische waarde rechts.

De *significantie* van de toetsingsgrootte is bij eenzijdig toetsen - afhankelijk van de alternatieve hypothese - de overschrijdings- of de onderschrijdingskans van de berekende toetsingsgrootte, onder geldigheid van de nulhypothese. Bij de combinatie van tweezijdig toetsen en een symmetrische kansverdeling van de toetsingsgrootte is de significantie tweemaal de overschrijdingskans van de absolute waarde van de berekende toetsingsgrootte.

De grootte van de significantie van de toetsingsgrootte ten opzichte van het *significantieniveau* α , het geaccepteerde risico op onterecht verwerpen van de nulhypothese, bepaalt of de nulhypothese wel of niet wordt verworpen. Als bijvoorbeeld de significantie van de toetsingsgrootte kleiner is dan α , dan wordt de nulhypothese verworpen. De kans op die, of een extremere waarde, wordt dan namelijk té klein geacht onder geldigheid van de nulhypothese.

Statistisch toetsen om beslissing te onderbouwen

Als het statistisch toetsen dient om een belangrijke beslissing te onderbouwen, waarbij twee partijen met tegengestelde belangen betrokken zijn, is het zinvol de onzekerheden voor beide partijen in beeld te brengen en zo mogelijk ook eerlijk te verdelen. Dit vergt een andere vorm van toetsen dan hierboven beschreven. Hieronder is dat toegelicht met een voorbeeld. Een uitgebreidere toelichting is te vinden in [Knotters et al, 2020].

Stel dat in een natuurgebied een antiverdrogingsmaatregel is genomen voor natuurherstel. Nu is de vraag of deze maatregel ook in de naaste omgeving heeft geleid tot hogere grondwaterstanden, want in dat geval kunnen grondgebruikers in aanmerking komen voor schadevergoeding. Er is slechts één tijdreeks beschikbaar van de grondwaterstand in de omgeving, die zowel enkele jaren vóór als enkele jaren ná de maatregel omvat. Om de bovengenoemde vraag objectief te kunnen beantwoorden is een tijdreeksmodel afgeleid, dat die grondwaterstandreeks beschrijft als functie van neerslag, verdamping en een interventievariabele, die de waarde 0 heeft tot de start van de maatregel en de waarde 1 daarna. Met dat model is de evenwichtsrelatie tussen de maatregel en de grondwaterstand geschat op 5,3 cm, met standaardfout 2,4 cm. Bij de modelverificatie bleek het model te voldoen aan de randvoorwaarden die gelden om er uitspraken mee te kunnen doen met gekwantificeerde onzekerheid (vermeld in paragraaf 11.5).

Om dan objectief de vraag te kunnen beantwoorden of de maatregel en de grondwaterstand gerelateerd zijn, voeren we een statistische toets uit. Er zijn meerdere combinaties van nulhypothese en alternatieve hypothese mogelijk, maar laten we als voorbeeld uitgaan van de nulhypothese (H_0) dat de werkelijke evenwichtsrelatie (g) nul (of minder) bedraagt en als alternatieve hypothese dat de werkelijke evenwichtsrelatie (g) groter dan nul is.

Dit wordt formeel als volgt uitgedrukt:

$$H_0: g \leq 0 \text{ en } H_a: g > 0$$

Als de modelresiduen normaal verdeelde witte ruis vormen (zie paragraaf 4.4.2) kan hiervoor de t-toets voor één steekproef worden gebruikt. Bij eenzijdig toetsen met 95 % betrouwbaarheid kan dan de nulhypothese worden verworpen als geldt:

$$T = \frac{\hat{g}}{s_{\hat{g}}} > t_{95\%,vg}$$

waarin T de toetsingsgrootte, \hat{g} de geschatte evenwichtsrelatie, $s_{\hat{g}}$ zijn standaardfout en $t_{95\%,vg}$ het 95-percentiel van de Student-t-verdeling met vg vrijheidsgraden, waarbij vg gelijk is aan het aantal modelresiduen dat bij de modelkalibratie was betrokken minus het aantal geschatte modelparameters. Het betreffende percentiel is de *kritische waarde* voor de toets. Het aantal vrijheidsgraden voor bovenstaand voorbeeld is 110, aangezien het een reeks betrof van 10 jaar lengte, die op maandbasis is gemodelleerd (120 waarden), met 10 modelparameters. Voor bovenstaand voorbeeld geldt dan:

$$T = \frac{5,3}{2,4} = 2,21 > t_{95\%,110} = 1,66$$

Aangezien in dit voorbeeld de toetsingsgrootte (2,21) groter is dan de kritische waarde (1,66) wordt de nulhypothese van geen positieve relatie verworpen. De conclusie is dat er met 95 % betrouwbaarheid sprake is van een statistisch significante positieve relatie tussen de ingreep en de betreffende grondwaterstandreeks.

Het risico op het trekken van een verkeerde conclusie bij het toetsen valt slechts uit te sluiten als er geen sprake is van onzekerheid over het te toetsen kengetal (zoals hier de evenwichtsrelatie). Dat is echter zelden of nooit het geval, zeker bij hydrologisch onderzoek, zodat we rekening moeten houden met toetsrisico's. Er zijn twee toetsrisico's te onderscheiden, namelijk:

- 1 het onterecht verwerpen van de nulhypothese (α , de fout van de 1^e soort). Voor de combinatie van hypothesen van bovenstaande toetscasus is dat het risico dat ten onrechte wordt geconcludeerd dat er sprake is van een positieve relatie en daarmee natschade;
- 2 het onterecht niet verwerpen van de nulhypothese (dit noemt men β , de fout van de 2^e soort). Voor de combinatie van hypothesen van bovenstaande toetscasus is dat het risico dat ten onrechte wordt geconcludeerd dat er *geen* sprake is van natschade.

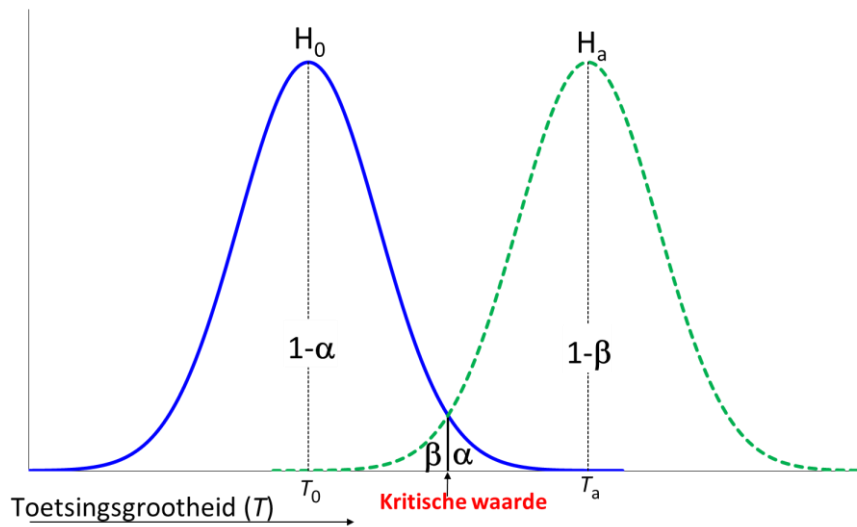
Tabel 11.1 en Afbeelding 11.5 illustreren de vier mogelijke uitkomsten bij het toetsen van een hypothese.

Tabel 11.1. De aanduidingen van de vier mogelijke uitkomsten bij het toetsen van een hypothese

Toetsuitkomst	De werkelijkheid	
	H_0 is waar	H_a is waar
Verwerp H_0 niet	Betrouwbaarheid ($1-\alpha$)	Fout 2 ^e soort (β)
Verwerp H_0	Fout 1 ^e soort (α)	Onderscheidend vermogen ($1-\beta$)

Bij het toetsen wordt α , het risico op onterecht verwerpen van de nulhypothese, doorgaans klein gehouden, zoals 5 % of 1 %. Slechts in uitzonderlijke situaties zal een nóg kleinere α worden gehanteerd. Er moet dan sprake zijn van zeer nadelige consequenties bij het onterecht verwerpen van de nulhypothese en beperkt nadelige consequenties bij het onterecht niet verwerpen van de nulhypothese.

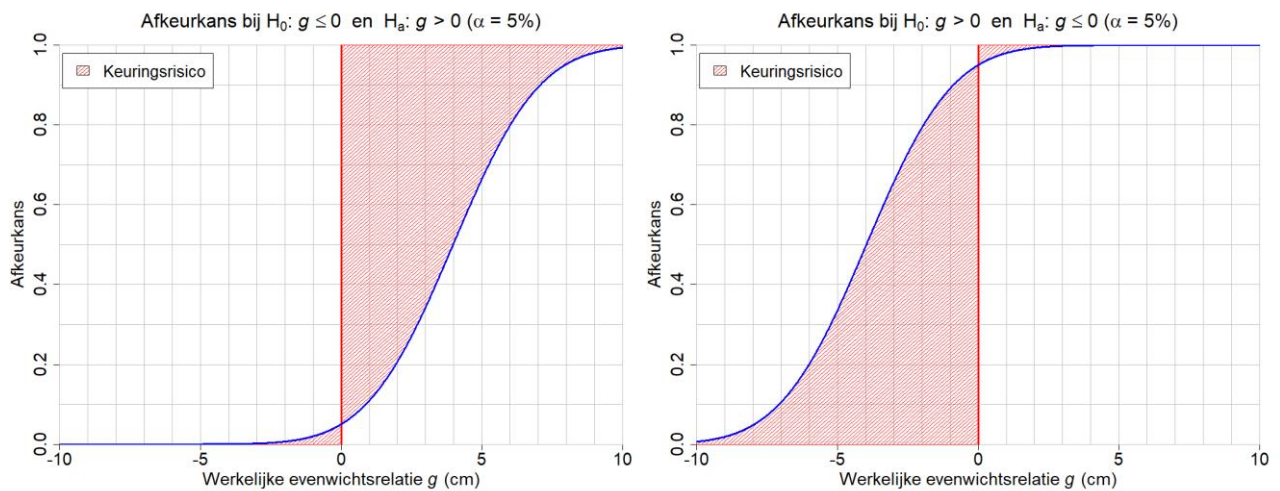
Afbeelding 11.5 Illustratie van de vier mogelijke uitkomsten bij het toetsen van een nulhypothese. De blauwe kansverdeling is die van de toetsingsgrootheid (T) onder de nulhypothese (H_0) en de groene is die van de toetsingsgrootheid onder één van de vele uitvoeringen van de alternatieve hypothese (H_a). De kansen α en β en hun complementen zijn toegelicht in tabel 11.1 hierboven



Zoals af te leiden uit Afbeelding 11.5, zal β toenemen als α lager wordt ingesteld (de kritische waarde verschuift dan naar rechts). De twee toetsrisico's kunnen dus niet simultaan worden verlaagd.

De toetsrisico's voor de twee partijen bij een toetscasus kunnen inzichtelijk worden gemaakt als we de casus beschouwen als een keuringsprobleem. We keuren de situatie af als de ingreep de grondwaterstand heeft verhoogd (de grondgebruikers hebben dan recht op een vergoeding) en we keuren goed als dat niet het geval is. Beide partijen lopen keuringsrisico's: de grondgebruikers op onterecht goedkeuren van de situatie (het concluderen dat er geen verhoging is, terwijl die er wel is) en de natuurbeheerder op onterecht afkeuren van de situatie (het concluderen dat er verhoging is, terwijl die er niet is). De keuringsrisico's voor de twee partijen kunnen inzichtelijk worden gemaakt met een keuringskarakteristiek. Afbeelding 11.6 toont twee voorbeelden voor de hier besproken casus, die slechts verschillen in de combinatie van hypothesen. In beide gevallen is er sprake van een zeer onevenwichtige verdeling van de keuringsrisico's. In het voorbeeld links ligt het risico vooral bij de grondgebruikers (onterecht goedkeuren situatie) en in het voorbeeld rechts vooral bij de natuurbeheerder (onterecht afkeuren).

Afbeelding 11.6 Twee voorbeelden van een keuringskarakteristiek voor deze toetscasus. Elk toont de afkeurkans (verticale as) als functie van de werkelijke evenwichtsrelatie van de grondwaterstand met de maatregel (horizontale as). Tevens zijn de keuringsrisico's van de twee betrokken partijen aangegeven (rood gearceerde oppervlakten aan weerskanten van $g = 0$)



Door α aan te passen, of een kleinste relevant geachte verhoging (δ) te verdisconteren kan een evenwichtiger verdeling van keuringsrisico's worden bewerkstelligd (zie daarvoor (Knotters et al, 2020)). Uit bovenstaande blijkt dat er bij het onderbouwen van een besluit met statistisch toetsen op moet worden toegezien dat de toetsrisico's voldoende evenwichtig zijn verdeeld. Dit kan ook samenspraak met de twee betrokken partijen vereisen.

11.5 Schade bij tijdreeksanalyse als modelresiduen geen witte ruis vormen

Als het tijdreeksmodel (mede) dient om uitspraken met gekwantificeerde onzekerheden te kunnen doen, moeten de modelresiduen voldoen aan de volgende randvoorwaarden:

- 1 ze zijn gemiddeld nul;
- 2 ze zijn onafhankelijk van elkaar (er is dan geen sprake van autocorrelatie);
- 3 ze hebben dezelfde variantie (homoscedasticiteit);
- 4 ze zijn afkomstig uit de normale kansverdeling (of een andere kansverdeling waarvan alle eigenschappen bekend zijn).

Als een reeks voldoet aan de eerste drie voorwaarden spreken we van *witte ruis*. Als deze ook voldoet aan normaliteit spreken we van *normaal verdeelde witte ruis*. Aanvullende randvoorwaarden, die echter vaak over het hoofd worden gezien en daardoor meestal impliciet worden aangenomen, zijn:

- 1 de formulering van het deterministische deel van het tijdreeksmodel stemt overeen met de werkelijke vorm (ook wel aangeduid als het *werkelijke systeem*) en
- 2 de schatter van de evenwichtsrelatie tussen een invoerreeks en de uitvoerreeks is zuiver, dat wil zeggen zonder systematische fout.

Met een eenvoudig simulatie-experiment is in beeld gebracht welke vertekeningen er kunnen optreden als we uitspraken met gekwantificeerde onzekerheden doen op basis van een tijdreeksmodel dat niet voldoet aan de randvoorwaarden [Baggelaar en Van der Meulen, 2020]. Dit is toegespitst op uitspraken over betrouwbaarheidsinterval en statistische significantie van de geschatte evenwichtsrelatie tussen een uitvoerreeks en een invoerreeks, waarbij het experiment zich beperkt tot de randvoorwaarde dat de modelfouten normaal verdeelde witte ruis vormen. De hoofdconclusies van het experiment zijn hieronder vermeld. Voor de geïnteresseerden zijn daaronder beknopt de relevante details van uitvoering en resultaten beschreven.

Conclusies over schade als modelresiduen geen witte ruis vormen

Als bij tijdreeksmodellering de modelresiduen (a_t) geen witte ruis vormen, kan dit leiden tot sterke vertekeningen van zowel het betrouwbaarheidsinterval als de statistische significantie van een geschatte evenwichtsrelatie. Voor wat betreft het betrouwbaarheidsinterval is de empirische dekkinggraad dan ongelijk aan de ingestelde nominale dekkinggraad. En voor wat betreft de statistische significantie is het empirische significantieniveau dan ongelijk aan het ingestelde nominale significantieniveau. Het risico op statistisch significant detecteren van een niet-bestaande relatie kan dan aanzienlijk worden.

Dus als de modelresiduen geen witte ruis vormen is het doen van uitspraken met gekwantificeerde onzekerheden niet meer te verantwoorden. De zuiverheid van een geschatte relatie is daarentegen onafhankelijk van of de modelresiduen wel of geen witte ruis vormen.

Dit leidt tot de belangrijkste adviezen:

- als het tijdreeksmodel (mede) dient om uitspraken met gekwantificeerde onzekerheden te doen, onderbouw dan dat voldaan wordt aan de daarvoor geldende randvoorwaarden (normaal verdeelde witte ruis);
- zie af van, of wees minstens zeer terughoudend met uitspraken over gekwantificeerde onzekerheden als niet wordt voldaan aan de randvoorwaarden.

Enkele details uitvoering en resultaten van het simulatie-experiment (voor geïnteresseerden)

Voor de eenvoud gaat het simulatie-experiment uit van een systeem waarbij een eventuele relatie tussen een invoervariabele X en de uitvoervariabele Z lineair en instantaan zal zijn, maar zonder na-ijling. Een dergelijke relatie kan worden beschreven met het enkelvoudig lineair regressiemodel, volgens:

$$Z_t = \beta_0 + \beta_1 X_t + \varepsilon_t$$

waarin β_0 het intercept, β_1 de regressieparameter, ε_t de modelfout en t de (discrete) tijdsindex. De eerste twee termen van het rechterlid van deze formule vormen het deterministische deel van het regressiemodel en de modelfout vormt het stochastische deel. Aangezien het een relatie tussen tijdreeksen betreft, kunnen we dit regressiemodel opvatten als een eenvoudig tijdreeksmodel. En de parameter β_1 kunnen we dan aanduiden als de evenwichtsrelatie van Z en X . Stel nu dat we objectief willen vaststellen of Z en X gerelateerd zijn en dat het antwoord belangrijke implicaties kan hebben, bijvoorbeeld op maatschappelijk, beleidsmatig, juridisch of technisch vlak. We zullen dan niet alleen de grootte van de relatie willen schatten, maar tevens zijn 95%-betrouwbaarheidsinterval en eventueel statistische significantie willen vaststellen. Als we niet alleen de grootte van β_1 willen schatten, maar ook zijn 95%-betrouwbaarheidsinterval en eventueel statistische significantie, dient volgens de statistische theorie de modelfout witte ruis te vormen en afkomstig te zijn uit een kansverdeling waarvan alle eigenschappen bekend zijn (zoals de normale kansverdeling). Om te kunnen vaststellen in welke mate die statistische kenmerken vertekend kunnen zijn als de modelfout correlatie vertoont, voegen we bij onze simulaties geen witte ruis, maar 'gekleurde' ruis toe aan de lineaire relatie tussen deze twee variabelen, volgens:

$$Z_t = \beta_0 + \beta_1 X_t + N_t$$

waarin N_t de ruis. Voor onze simulatie is gekleurde ruis gegenereerd volgens een ARMA(1,1)-ruismodel:

$$N_t = \phi_1 N_{t-1} + a_t - \theta_1 a_{t-1}$$

waarin a_t een trekking uit een normale kansverdeling met gemiddelde nul. Dit ruismodel is uitgebreider dan het AR(1)-model, omdat bij hydrologische tijdreeksmodellering het AR(1)-ruismodel zelden afdoende blijkt om te kunnen voldoen aan de randvoorwaarde van witte ruis, vooral als het een hoogfrequente tijdreeks betreft.

We hanteren bij de simulatie 190 uitvoeringen van dit ruismodel, namelijk voor alle combinaties van ϕ_1 en θ_1 , waarbij ϕ_1 waarden aanneemt tussen 0 en 0,9, met stapgrootte 0,1 en θ_1 waarden aanneemt tussen -0,9 en 0,9, eveneens met stapgrootte 0,1. Voor elk van deze 190 combinaties van ϕ_1 en θ_1 wordt 5.000 maal een trekking gedaan van a_t (uit een normale kansverdeling). De resulterende ruisreeks wordt toegevoegd aan het deterministische deel van het tijdreeksmodel. De standaardafwijking van de verdeling van a_t (σ_a) is per combinatie zo ingesteld in dat bij elk van de 190 ruismodellen de theoretische variantie van de ruis (σ_N) dezelfde, vooraf beoogde grootte heeft.

Verdere technische specificaties van de opzet

- er zijn twee simulatie-exercities uitgevoerd (A en B), elk met vaste instellingen voor β_0 (intercept), β_1 (regressieparameter) en σ_N (standaardafwijking van de ruis). De exercities verschillen alleen in de instelling van β_1 . Dit betreft:
 - A. $\beta_0 = 10$, $\beta_1 = 1$ en $\sigma_N = 2$ (Z en X zijn hierbij dus gerelateerd via β_1);
 - B. $\beta_0 = 10$, $\beta_1 = 0$ en $\sigma_N = 2$ (Z en X zijn hierbij dus niet gerelateerd);
- elk van de tijdreeksen Z, X en N bevat 3.652 dagwaarden ($k = 3.652$) en beslaat de tienjarige periode 2009 t/m 2018;
- de tijdreeks van de variabele X is eenmalig samengesteld (handmatig). Deze omvat elf niveauperanderingen, zodat een relatie tussen Z en X makkelijker kan worden gedetecteerd;
- op elk van de 950.000 (= 190 x 5.000) gesimuleerde tijdreeksen $Z_t = \beta_0 + \beta_1 X_t + N_t$ zijn de volgende statistische bewerkingen uitgevoerd:
 - schatten van de modelparameters β_0 en β_1 als b_0 en b_1 , leidend tot het volgende uit de meetwaarden geschatte tijdreeksmodel:

$$z_t = b_0 + b_1 \cdot x_t + e_t$$

waarin e_t de modelfout;

- bepalen of de ingestelde waarde van β_1 wel of niet binnen het 95%-betrouwbaarheidsinterval van zijn schatting b_1 ligt, dat volgt uit:

$$95\% \text{ b.i. } [b_1] = b_1 \pm t_{97,5\%;k-2} \cdot s_{b_1}$$

waarin $t_{97,5\%;k-2}$ het 97,5-percentiel van de Student-t-verdeling met $k-2$ vrijheidsgraden, k het aantal waarden, s_{b_1} de standaardfout van b_1 en de andere symbolen als boven toegeelicht.

Resultaten simulatie-exercitie

Bij simulatie-exercitie A is een relatie aangebracht tussen variabele Z en variabele X ($\beta_1 = 1$). Voor alle 190 combinaties van ϕ_1 en θ_1 is vastgesteld in hoeveel van de 5.000 simulaties de ingestelde waarde van β_1 binnen het 95 %-betrouwbaarheidsinterval van zijn schatting b_1 ligt. Het zo bepaalde percentage duiden we aan als de *empirische dekkingsgraad* van het betrouwbaarheidsinterval. Als het model voldoet aan de eerder vermelde randvoorwaarden zal volgens de statistische theorie de empirische dekkingsgraad gelijk zijn aan de *nominale dekkingsgraad* (dit is de ingestelde dekkingsgraad, hier 95 %). Tabel 11.2 vermeldt de empirische dekkingsgraad van β_1 voor elk van de 190 combinaties van ϕ_1 en θ_1 , bij een nominale dekkingsgraad van 95 %.

Tabel 11.2 Empirische dekkingsgraad van het betrouwbaarheidsinterval van de geschatte modelparameter β_1 voor elk van de 190 combinaties van ϕ_1 en θ_1 van een simulatie-experiment, bij een nominale dekkingsgraad van 95 %

$\beta_1 = 1$		θ_1																		
		-0,9	-0,8	-0,7	-0,6	-0,5	-0,4	-0,3	-0,2	-0,1	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
ϕ_1	0	84%	84%	84%	84%	85%	87%	88%	90%	93%	95%	97%	99%	100%	100%	100%	100%	100%	100%	100%
	0,1	82%	81%	82%	82%	83%	84%	85%	87%	90%	92%	95%	97%	99%	100%	100%	100%	100%	100%	100%
	0,2	78%	78%	79%	80%	81%	81%	82%	84%	85%	88%	92%	95%	97%	99%	100%	100%	100%	100%	100%
	0,3	76%	75%	77%	76%	78%	76%	79%	80%	83%	85%	87%	92%	95%	98%	99%	100%	100%	100%	100%
	0,4	72%	72%	73%	73%	73%	74%	74%	76%	78%	81%	84%	87%	91%	95%	99%	100%	100%	100%	100%
	0,5	68%	67%	67%	68%	68%	69%	69%	71%	73%	74%	77%	81%	85%	91%	95%	99%	100%	100%	100%
	0,6	62%	61%	64%	61%	64%	63%	63%	63%	65%	66%	70%	71%	77%	82%	89%	95%	99%	100%	100%
	0,7	56%	56%	55%	55%	55%	56%	57%	57%	58%	58%	61%	63%	67%	72%	78%	87%	95%	100%	100%
	0,8	46%	46%	45%	47%	46%	47%	47%	48%	48%	48%	50%	50%	55%	57%	62%	71%	82%	95%	100%
	0,9	36%	34%	33%	35%	35%	35%	35%	35%	34%	36%	36%	36%	37%	39%	41%	46%	53%	69%	95%

Empirische dg	Kleur
96% t/m 100%	100%
94% t/m 96%	96%
90% t/m 94%	94%
75% t/m 90%	90%
50% t/m 75%	75%
≤50%	50%

Op basis van Tabel 11.2 kan het volgende geconcludeerd worden:

1. als er sprake is van witte ruis ($\phi_1 = 0$ en $\theta_1 = 0$), is de empirische dekkinggraad van het betrouwbaarheidsinterval van b_1 gelijk aan de nominale dekkinggraad (hier 95 %). Dit is in overeenstemming met de statistische theorie;
2. de empirische dekkinggraad blijkt ook gelijk aan de nominale dekkinggraad in de negen gevallen dat $\phi_1 = \theta_1$. Dit kan worden verklaard uit de eigenschappen van het ARMA(1,1)-model. Zie daarvoor [Baggelaar en Van der Meulen, 2019];
3. in alle andere gevallen is de empirische dekkinggraad ongelijk aan de nominale dekkinggraad. Dit zijn situaties waar het doen van uitspraken met gekwantificeerde betrouwbaarheden niet meer te verantwoorden is;
4. voor de gevallen dat $\theta_1 < \phi_1$ is de empirische dekkinggraad kleiner dan de nominale dekkinggraad. Het 95%-betrouwbaarheidsinterval is dan te smal en het risico op onterecht detecteren van statistische significantie is dan groter dan ingesteld. De afwijking neemt toe met toenemende ϕ_1 en met kleiner of negatiever wordende θ_1 , zij het dat deze vanaf θ_1 ergens tussen -0,7 en -0,8 weer licht afneemt;
5. voor de gevallen dat $\theta_1 > \phi_1$ is de empirische dekkinggraad groter dan de nominale dekkinggraad. Het 95%-betrouwbaarheidsinterval is dan te breed, wat neerkomt op verlies aan onderscheidend vermogen¹. De afwijking neemt toe met toenemende θ_1 en met kleiner wordende ϕ_1 , maar kan (uiteraard) niet meer bedragen dan 5 % (het maximum van de empirische dekkinggraad is immers 100 %);
6. de afwijkingen lopen in deze simulatie op tot maximaal 62 %, aangezien de geconstateerde minimale dekkinggraad 33 % bedraagt, namelijk bij de combinatie $\phi_1 = 0,9$ en $\theta_1 = -0,7$.

We kunnen uit deze simulatie-exercitie concluderen dat het niet modelmatig verdisconteren van gekleurde ruis het betrouwbaarheidsinterval van een geschatte relatie sterk kan vertekenen. De toetsrisico's (op onterecht verwerpen respectievelijk onterecht niet verwerpen van de nulhypothese) zijn dan ongelijk aan degene waar de analist rekening mee houdt.

Resultaten simulatie-exercitie B

Net als bij simulatie-exercitie A is voor elk van de 190 combinaties van ϕ_1 en θ_1 de empirische dekkinggraad vastgesteld van het betrouwbaarheidsinterval van β_1 , bij een nominale dekkinggraad van 95 %. Het blijkt dat dit een vrijwel identiek beeld oplevert als de resultaten van simulatie-exercitie A, die zijn vermeld in Tabel 11.2. We mogen aannemen dat de kleine verschillen komen door stochastische variatie, aangezien er sprake is van een eindige steekproefgrootte. De conclusie luidt dan dat de onder Tabel 11.2 vermelde bevindingen onafhankelijk zijn van het wel of niet bestaan van een relatie tussen de uitvoerreeks (Z) en de invoerreeks (X).

Omdat bij deze simulatie-exercitie B de regressiecoëfficiënt β_1 is ingesteld op nul, is er geen sprake van een relatie tussen Z en X . Dit stelt in staat na te gaan in welke mate het niet verdisconteren van gekleurde ruis het risico kan beïnvloeden op onterecht concluderen dat er sprake is van een statistisch significante relatie. Daarbij bleek dat het niet modelmatig verdisconteren van gekleurde ruis de statistische significantie van een geschatte relatie sterk kan vertekenen. Het risico op statistisch significant detecteren van een niet-bestaande relatie kan dan aanzienlijk worden.

¹ Het onderscheidend vermogen is de kans om een werkelijk optredend verschil van een bepaalde grootte statistisch significant te detecteren.

12

INTRODUCTIE BOX-JENKINS TIJDREEKSMODELLEN

Aan de hand van tijdreeksanalyse volgens de Box-Jenkins-methode, kan de tijdreeks van een variabele beschreven worden met een discreet lineair dynamisch stochastisch model. In het nu volgende worden de daarvoor beschikbare modellen geïntroduceerd. Voor een diepgaande theoretische verhandeling wordt verwezen naar het standaardwerk (Box and Jenkins, 1970).

Deze introductie hanteert de klassieke Box-Jenkins-notatie en terminologie, omdat zo makkelijker de literatuur hierover kan worden geraadpleegd voor verdere verdieping.

12.1 Uitleg aan de hand van systeemtheorie

Het principe van de Box-Jenkins benadering kan goed worden toegelicht aan de hand van begrippen uit de systeemtheorie. Een systeem zal daartoe worden omschreven als een deel van de werkelijkheid, dat via in- en uitgangssignalen wisselwerkingen vertoont met z'n omgeving. De grondwaterstand in een peilbuis zullen we dan opvatten als het uitgangssignaal (of de uitvoerreeks) van een systeem, dat wordt aangedreven door meerdere ingangssignalen (of invoerreeksen). In het geval van een grondwaterstand zullen deze laatste meestal van meteorologische aard zijn, maar daarnaast kan het ook waterhuishoudkundige ingrepen of grondwaterwinningen betreffen.

Een dergelijk systeem kan mathematisch worden uitgeschreven als:

$$Z(t) = g(G, X, t)$$

met Z de grondwaterstand, $g(\cdot)$ één of andere functie, G een verzameling modelparameters, X de verzameling invoerreeksen en t de continue tijd.

Als we de grondwaterstand empirisch willen modelleren kan het voordelen bieden deze te beschouwen als de uitvoerreeks van een discreet lineair dynamisch stochastisch systeem:

- discreet: omdat relevante gegevens over de grondwaterstand en de invoerreeksen meestal slechts beschikbaar zijn in de vorm van discrete tijdreeksen;
- lineair: omdat dit de minste theoretische en praktische problemen oplevert, bovendien is gebleken dat veel systemen redelijk beschreven kunnen worden met lineaire modellen;
- dynamisch: omdat de tijd een essentiële rol speelt in het gedrag van de uitvoerreeks;
- stochastisch: omdat zo'n systeem nooit exact (deterministisch) kan worden beschreven, onder meer vanwege: i) de benadering die is ingevoerd door de discretisatie en linearisatie, ii) het feit dat het onmogelijk is alle invoerreeksen bij de beschrijving mee te nemen en iii) meetfouten.

Een discreet lineair dynamisch stochastisch systeem kan mathematisch worden uitgeschreven als:

$$Z_t = f(F, X, t)$$

met Z de grondwaterstand, $f(\cdot)$ een lineaire functie, F een verzameling modelparameters, X de verzameling invoerreeksen, N de ruis en t de discrete tijdsindex. Omdat de algemene vorm van de Box-Jenkins-modellen ook kan worden uitgeschreven als bovenstaande formule, is de Box-Jenkins-methode geschikt om zo'n systeem langs empirische weg te beschrijven.

En omdat de methode er op is gericht het stochastische element terug te brengen tot een verschijnsel dat minimaal is en dat bekende waarschijnlijkheidswetten volgt, kunnen er uitspraken omtrent het systeem worden gedaan met minimale en kwantificeerbare onzekerheden.

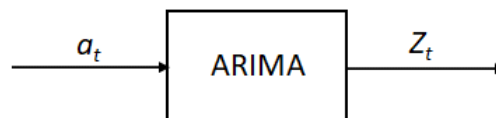
12.2 Statistisch modelleren van een tijdreeks

De klassieke statistische methodologie is zelden geschikt voor een statistische modellering van tijdreeksen, omdat deze gebaseerd is op onafhankelijke waarnemingen, terwijl de waarnemingen in een tijdreeks vaak niet onafhankelijk van elkaar zijn. De belangrijkste bijdragen van Box en Jenkins tot de oplossing van dit probleem bestonden uit de formulering van een familie van statistische modellen om tijdreeksen te beschrijven en uit aanwijzingen tot een verantwoorde modelbouw (Box and Jenkins, 1970). Voor de modellering van een tijdreeks ontwikkelden ze twee benaderingen, namelijk univariate modellering aan de hand van een ARIMA-model en transfer-ruismodellering aan de hand van een transfer-ruismodel (in deze Handleiding aangeduid als BJTR-model). Beide zullen in het nu volgende kort en voornamelijk intuïtief worden omschreven.

Univariaat modelleren

Bij univariaat modelleren wordt uitsluitend uitgegaan van de te modelleren tijdreeks, die wordt beschouwd als de uitvoerreeks van een discreet lineair dynamisch stochastisch systeem dat wordt aangedreven door een stochastische invoerreeks. Deze laatste is onbekend en ontstaat eigenlijk pas in de modelleerfase als de residueeks a_t (zie Afbeelding 12.1) De modellering is er dan ook op gericht die terug te brengen tot een verschijnsel dat bekende waarschijnlijkheidswetten volgt (normaal verdeelde witte ruis) en een minimale variantie heeft. Die variantie mag in dit verband als een soort maat voor de onzekerheid van het model worden beschouwd. Witte ruis kan worden opgevat als een volledig toevallig signaal, bestaande uit een opeenvolging van onafhankelijke trekkingen uit een kansverdeling, met gemiddelde nul en variantie σ_a^2 .

Afbeelding 12.1 Principe van het univariate ARIMA-model



Het zal duidelijk zijn dat deze univariate benadering weinig inzicht kan leveren in het beschouwde systeem. Maar het wordt wel veel toegepast om voorspellingen van een bepaalde variabele te genereren. Deze zijn dan uitsluitend gebaseerd op een statistische analyse van het gedrag van deze variabele in het verleden. Deze toepassing heeft dan ook vooral opgang gemaakt in de economie, met zijn gecompliceerde en minder inzichtelijke dynamische systemen.

Verder is het ARIMA-model door zijn grote flexibiliteit zeer geschikt om te fungeren als ruismodel bij het transfer-ruismodelleren (zie daarvoor verder), omdat het daarmee vaker zal lukken het modelresidu a_t terug te brengen tot witte ruis. Als het ARIMA-model geen differentie bevat (zie paragraaf 12.4), wordt het doorgaans aangeduid als ARMA-model.

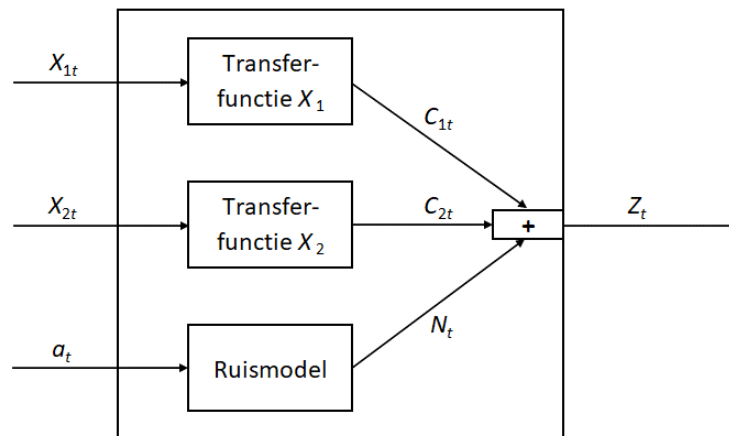
Transfer-ruismodelleren

Bij de transfer-ruismodelleren wordt de te modelleren tijdreeks beschouwd als de uitvoerreeks van een discreet lineair dynamisch stochastisch systeem dat wordt aangedreven door meerdere invoerreeksen, waaronder ook witte ruis. Elke invoerreeks (X_i) levert hierbij volgens een afzonderlijke lineaire transferfunctie een component (C_i) van de uitvoerreeks (Z). De resterende ruis (N) vertegenwoordigt de component die wordt veroorzaakt door: i) alle invloedsfactoren die bij het modelleren niet meededen als invoerreeksen, ii) onvolkomenheden van het model en iii) meetfouten. Deze ruis kan op zijn beurt worden weergegeven als de uitvoerreeks van een deelsysteem, dat volgens een ARIMA-model (in feite ook een soort transferfunctie) wordt aangedreven door witte ruis. De verschillende componenten geven door middel van lineaire superpositie de uiteindelijk waargenomen uitvoerreeks, volgens:

$$Z_t = C_{1t} + \dots + C_{mt} + N_t$$

Afbeelding 12.2 toont het principe van het transfer-ruismodel, hier met twee invoerreeksen, zoals wanneer de grondwaterstand wordt gemodelleerd als functie van het potentieel neerslagoverschot en een grondwaterwinning.

Afbeelding 12.2 Principe van het transfer-ruismodel, hier weergegeven voor een systeem met twee invoervariabelen



De kracht van het transfer-ruismodel is dat daarmee een groot aantal dynamische situaties kunnen worden beschreven met een beperkt aantal modelparameters.

12.3 Ontwikkelen van een BJTR-model

Het ontwikkelen van een BJTR-model is - net als bij de meeste modellen - een iteratief proces. Per ronde worden drie fasen doorlopen:

- 1 identificeren van de modelvorm;
- 2 schatten van de modelparameters en;
- 3 verifiëren van het model.

Het nu volgende licht de drie fasen toe, met als toepassing de ontwikkeling van een transfer-ruismodel.

Identificeren van de modelvorm

Aan de hand van diagnostieken van de relaties tussen de uitvoerreeks en de invoerreeksen en/of theoretisch inzicht in deze relaties, wordt de vorm van het model geformuleerd. De diagnostieken worden geleverd door de kruiscorrelatiefuncties van de uitvoerreeks en elke afzonderlijke invoerreeks. Zo'n functie geeft de kruiscorrelatiecoëfficiënt, een maat voor de samenhang tussen de waarde van de invoerreeks en de uitvoerreeks k tijdseenheden later, als functie van k (de tijdsverschuiving, of *lag*). Een kruiscorrelatiefunctie wordt vaak geplot weergegeven, omdat dit de diagnose aanzienlijk vergemakkelijkt.

Omdat het beeld van de relatie vertroebeld kan worden door de autocorrelaties binnen de reeksen, worden beide vóór het berekenen van de kruiscorrelatiefunctie door hetzelfde lineaire filter gehaald (zie daarvoor paragraaf 4.2.1). Deze aanpak is in [Box and Jenkins, 1976] aangeduid als *prewhitening* (witten).

In de identificatiefase dient tevens de vorm van het ruismodel te worden geformuleerd. Hiervoor zijn verschillende benaderingen mogelijk. Als er met grote waarschijnlijkheid van kan worden uitgegaan dat alle relevante invoerreeksen in het transfer-ruismodel zijn opgenomen, kan het ruismodel worden geformuleerd als een stationair model. Zo'n model beschrijft een tijdreeks die voldoet aan een bepaald statistisch evenwicht en in feite met een constante spreiding rond een constante waarde schommelt. Als er daarentegen (vermoedelijk) relevante invoerreeksen ontbreken, kan het ruismodel in eerste instantie worden geformuleerd als het ARIMA-model voor de uitvoerreeks. In alle gevallen zal een juistheid of onjuistheid van de formulering van het ruismodel vanzelf blijken in de verificatiefase, waarna het model kan worden aangepast.

Schatten van de modelparameters

De parameters van het geïdentificeerde model worden geschat volgens een optimalisatieprocedure, waarbij de som van de kwadraten van de modelresiduen - een functie van de modelparameters - wordt geminimaliseerd. Dit noemt men Kleinste Kwadratenmethode (zie hiervoor paragraaf 4.3).

Verifiëren van het model

In de verificatiefase wordt nagegaan of het model de tijdreeks adequaat beschrijft en of voldaan wordt aan alle vooronderstellingen die aan het model ten grondslag liggen. (modelresiduen normaal verdeeld, onafhankelijk van elkaar en zonder relatie met de afzonderlijke invoerreeksen). Hierbij wordt gebruik gemaakt van diagnostieken van de modelresiduen, zoals een plot, een histogram en enkele correlatiefuncties, alsmede de schatting van de variantie van de modelresiduen, de statistische significantie van de modelparameters en de correlatiematrix van de parameterschatters. Bij een onbevredigende diagnose kan een betere modelvorm geformuleerd worden aan de hand van de geconstateerde discrepanties. Het ontwikkelproces vervolgt dan weer met de schatting van de modelparameters.

Betrouwbaarheidsinterval van een geschatte parameter

De schatting van de parameter van een transfer-ruismodel kan opgevat worden als een waarde van een schatter, een stochastische (of toevals-)variabele met een bepaalde kansverdeling. Een schatting gaat gepaard met een maat voor de spreiding van deze kansverdeling: de standaardafwijking (de wortel uit de variantie). Deze standaardafwijking van een schatter wordt ook wel als standaardfout aangeduid. Als voldaan wordt aan de vooronderstellingen die ten grondslag liggen aan het transfermodel, heeft een schatter een normale kansverdeling waarvan het gemiddelde overeenkomt met de werkelijke waarde van de te schatten parameter. Aan de hand van de schatting en de standaardfout kan dan het interval worden aangegeven waarbinnen de echte waarde van de parameter zal liggen met een bepaalde betrouwbaarheid (meestal wordt hiervoor 95 % genomen). Dit noemt men het (95 %-)betrouwbaarheidsinterval. Als dit interval niet de waarde 0 bevat, is de modelparameter statistisch significant.

12.4 Formuleren van een BJTR-model

De Box-Jenkins formulering van de dynamische relatie tussen een invoerreeks en een uitvoerreeks van een systeem is afgeleid van de algemene weergave van een discreet lineair dynamisch systeem met een lineaire differentievergelijking. Voor een systeem met één invoerreeks geldt bijvoorbeeld deze differentievergelijking:

$$(1 + C_1 \nabla + \dots + C_r \nabla^r) Z_t = g(1 + D_1 \nabla + \dots + D_s \nabla^s) X_t$$

met Z de afwijking van de uitvoerreeks ten opzichte van zijn gemiddelde, X de afwijking van de invoerreeks ten opzichte van zijn gemiddelde, t de discrete tijdsindex, $C_1 \dots C_r$ en $D_1 \dots D_s$ modelparameters, ∇ de differentie-operator ($\nabla Z_t = Z_t - Z_{t-1}$) en g de evenwichtsrelatie tussen Z en X . Deze evenwichtsrelatie is een belangrijke karakteristiek van een dynamische relatie. Het is de uiteindelijke stationaire verandering van de uitvoerreeks als de invoerreeks een constante niveauperandering van één eenheid heeft. Of, in het geval van de modellering van de grondwaterstand aan de hand van een grondwaterwinning, de uiteindelijke stationaire verlaging van de grondwaterstand als de grondwaterwinning een constante niveauperandering van één eenheid heeft (er van uitgaande dat de relatie causaal is en niet louter statistisch).

De oplossing van bovenstaande differentievergelijking kan worden uitgeschreven als:

$$Z_t = \sum_{u=0}^{\infty} v_u X_{t-u} = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots$$

$$= (v_0 + v_1 B + v_2 B^2 + \dots) X_t$$

met v_0, v_1, v_2, \dots de impulsresponsgewichten (die samen de discrete impulsresponsfunctie vormen) en B de backshift-operator ($BX_t = X_{t-1}$), die slechts dient om de notatie te vereenvoudigen. Deze impulsresponsfunctie vertegenwoordigt de dynamica van het systeem en beschrijft de reactie van de uitvoerreeks Z op een enkelvoudige impuls van de invoerreeks X. En als X bestaat uit een opeenvolging van verschillende impulsen, wordt Z beschreven als de som van door X geschaalde impulsresponsgewichten.

In Box-Jenkins-notatie wordt een differentievergelijking met één invoerreeks (en nog zonder ruis) uitgeschreven als:

$$(1 - \delta_1 B - \dots - \delta_r B^r) Z_t = (\omega_0 - \omega_1 B - \dots - \omega_s B^s) X_t$$

met $\delta_1, \dots, \delta_r$ de AR-parameters, $\omega_0, \dots, \omega_s$ de MA-parameters en B de backshift-operator (hier gebruikt als substituuut voor ∇ , aangezien $B = 1 - \nabla$). Dit noemt men de transferfunctie van de orde (r,s). In zeer compacte vorm wordt dit ook wel als volgt uitgeschreven:

$$Z_t = \frac{\omega(B)}{\delta(B)} X_t$$

met $\omega(B)$ de MA-operator voor X, volgens:

$$\omega(B) X_t = \omega_0 X_t - \omega_1 X_{t-1} - \dots - \omega_s X_{t-s}$$

en $\delta(B)$ de AR-operator, die inwerkt op Z, volgens:

$$\delta(B) Z_t = Z_t - \delta_1 Z_{t-1} - \dots - \delta_r Z_{t-r}$$

Door het AR-deel kan een transferfunctie die wordt gekenmerkt door een serie exponentieel afnemende MA-parameters door veel minder parameters worden beschreven. Dit wordt duidelijker als we de onderstaande formule voor een systeem met één invoerreeks (en geen ruis) verder uitwerken:

$$Z_t = \frac{\omega_0}{(1 - \delta_1 B)} X_t$$

$$(1 - \delta_1 B) Z_t = \omega_0 X_t$$

$$Z_t = \delta_1 Z_{t-1} + \omega_0 X_t$$

$$Z_t = \delta_1 (\delta_1 Z_{t-2} + \omega_0 X_{t-1}) + \omega_0 X_t$$

$$Z_t = \delta_1 (\delta_1 (\delta_1 Z_{t-3} + \omega_0 X_{t-2}) + \omega_0 X_{t-1}) + \omega_0 X_t$$

$$Z_t = \delta_1^3 Z_{t-3} + \delta_1^2 \omega_0 X_{t-2} + \delta_1 \omega_0 X_{t-1} + \omega_0 X_t$$

$$Z_t = \delta_1^m Z_{t-m} + \delta_1^{m-1} \omega_0 X_{t-(m-1)} + \delta_1^{m-2} \omega_0 X_{t-(m-2)} + \dots + \omega_0 X_t$$

Aangezien moet opgaan dat $|\delta_1| < 1$ (anders kan er geen sprake zijn van een stationair model), geldt:

$$\delta_1^m \rightarrow 0, \text{ als } m \rightarrow \infty$$

De formule voor het systeem reduceert daardoor tot:

$$Z_t = \omega_0 X_t + \delta_1 \omega_0 X_{t-1} + \delta_1^2 \omega_0 X_{t-2} + \delta_1^3 \omega_0 X_{t-3} + \delta_1^4 \omega_0 X_{t-4} + \dots$$

Een MA-transferfunctie met exponentieel afnemende MA-parameters is dus uit te schrijven als een transferfunctie met één MA-parameter (ω_0) en één AR-parameter (δ_1), wat een grote reductie geeft van het aantal benodigde modelparameters. Dit is handig, omdat veel impulsresponsrelaties gekenmerkt worden door een geleidelijke terugkeer naar nul.

Het BJTR-model is een model om een tijdreeks te beschrijven als een som van transferfuncties van de relevante invoerreeksen (tezamen het deterministische deel van het model), aangevuld met een ruisterm (het stochastische deel van het model). In deze zin vormt het een discreet lineair dynamisch stochastisch model. In zeer compacte vorm kan het BJTR-model als volgt worden uitgeschreven:

$$Z_t = \frac{\omega_1(B)}{\delta_1(B)} X_{1t} + \dots + \frac{\omega_u(B)}{\delta_u(B)} X_{ut} + N_t$$

met X_1, \dots, X_u de invoerreeksen, $\omega_i(B)$ de MA-operator voor X_i , $\delta_i(B)$ de AR-operator voor X_i en N de ruis. Meestal bevat zo'n ruisreeks nog regelmatigigheden die beschreven kunnen worden met een ARIMA-model (zie hieronder), waardoor het BJTR-model nog beter de uitvoerreeks kan beschrijven en daarmee ook betere voorspellingen kan opleveren.

Formuleren van het stochastische deel van het BJTR-model (het ruismodel)

Het ruisdeel van het BJTR-model wordt geformuleerd als een ARIMA-model. Als dat geen differentie bevat wordt het ook wel aangeduid als ARMA-model. Het beschrijft de ruis N_t als lineaire functie van voorgaande ruiswaarden en modelresiduen. De algemene vorm van een ARIMA-ruismodel is:

$$\phi(B) \cdot \Phi(B^S) \cdot (\nabla^d \nabla_S^D N_t^\lambda - c) = \theta(B) \cdot \Theta(B^S) \cdot a_t$$

met N de ruis, a het modelresidu, t de tijdsindex, ∇ de differentie-operator ($\nabla N_t = N_t - N_{t-1}$), d het aantal differenties, S de seizoensperiode (met maandelijke waarnemingen en een jaarcyclus is deze bijvoorbeeld 12), ∇_S de seizoensdifferentie-operator ($\nabla_S N_t = N_t - N_{t-S}$), D het aantal seizoensdifferenties, λ de transformatieparameter (deze is bijvoorbeeld 1 als N normaal verdeeld is) en c een constante. De transformatieparameter maakt het bij scheve kansverdelingen vaak mogelijk om het modelresidu normaal verdeeld te krijgen.

Verder bevat deze formulering de volgende operatoren:

- $\phi(B)$ de AR-operator, volgens: $\phi(B)N_t = N_t - \phi_1 N_{t-1} - \dots - \phi_p N_{t-p}$
waarin $\phi_1 \dots \phi_p$ de AR-parameters;
- $\Phi(B^S)$ de AR-seizoensoperator, volgens: $\Phi(B^S)N_t = N_t - \Phi_1 N_{t-S} - \dots - \Phi_P N_{t-PS}$
waarin $\Phi_1 \dots \Phi_P$ de AR-seizoenparameters;
- $\theta(B)$ de MA-operator, volgens: $\theta(B)a_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$
waarin $\theta_1 \dots \theta_q$ de MA-parameters;
- en $\Theta(B^S)$ de MA-seizoensoperator, volgens: $\Theta(B^S)a_t = a_t - \Theta_1 a_{t-S} - \dots - \Theta_Q a_{t-QS}$
waarin $\Theta_1 \dots \Theta_Q$ de MA-seizoenparameters.

13

RUIS, RESIDUEN EN INNOVATIES: WAT IS WAT?

In dit hoofdstuk wordt toelichting gegeven op de termen ruis, residuen en innovaties en hoe deze termen gebruikt worden binnen en buiten de tijdreeksanalyteliteratuur.

13.1 Beknopte toelichting op de achterliggende concepten van ruis, residu en innovaties

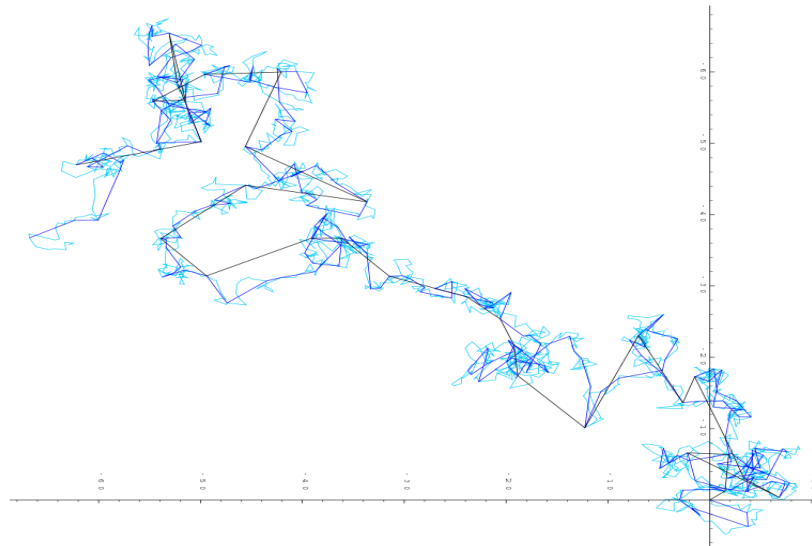
Alhoewel tijdreeksanalyse tegenwoordig snel en eenvoudig toepasbaar is geworden door de komst van gebruiksvriendelijke software, blijven de achterliggende methoden en theorie ingewikkeld. Zolang de tijdreeksanalysetoepassing standaard en de resultaten van de analyse eenduidig zijn hoeft dat op zich geen probleem te zijn. Zodra die resultaten echter om nadere interpretatie en duiding vragen, is het ook belangrijk om de achterliggende theorie, concepten en bijbehorende terminologie te kennen. Omgekeerd is het daarmee óók belangrijk om de juiste termen voor de juiste concepten te gebruiken bij het rapporteren of communiceren over tijdreeksanalyseresultaten, omdat anders al snel (spraak)verwarring en misverstanden ontstaan.

Binnen het toepassingsgebied van de tijdreeksanalyse bestaat al enige tijd spraakverwarring waar het gaat om het gebruik van de termen 'ruis', 'residuen' en 'innovaties'. We gaan hier allereerst kort in op de achterliggende concepten, omdat basisbegrip daarvan nodig is om ook de specifieke betekenis van een bepaalde term te doorgronden (net zoals begrip van kansverdelingen nodig is om het verschil in betekenis tussen de Engelstalige termen *mean* en *average* in die context te begrijpen). We gebruiken ook hier een voorbeeld uit de fysica daarbij, namelijk de zogenoemde Brownse beweging. Bij Brownse of Browniaanse beweging gaat het in beginsel om een natuurkundig verschijnsel, namelijk de autonome beweging van kleine deeltjes die de Schotse botanicus Robert Brown waarnam bij een suspensie van stuifmeelkorrels in water (zie Afbeelding 13.1 [https://nl.wikipedia.org/wiki/Brownse_beweging_\(natuurkunde\)](https://nl.wikipedia.org/wiki/Brownse_beweging_(natuurkunde))). Het was nota bene Einstein die de Brownse beweging ook wist te beschrijven in wiskundige formules (Einstein, 1905), formules die daarna hun weg vonden in de definitie van een stochastisch proces dat eveneens 'Brownse beweging' of ook Wienerproces wordt genoemd (zie bijv. Wiener (1949)), en in de beschrijving van stochastische processen en onderwerpen als filtertheorie in het algemeen.

De verklaring van Brownse beweging is dat kleine deeltjes zoals stuifmeelkorrels onderhevig zijn aan botsingen met de moleculen van de vloeistof waarin ze zich bevinden, wat een autonome, chaotische beweging veroorzaakt. Brownse beweging is dus een **indirect** bewijs voor het bestaan van moleculen en botsingen daarmee als drijvende kracht van die beweging. De moleculen, botsingen en drijvende kracht zelf zijn **niet direct waarneembaar**. Van belang in deze context is ook dat de beweging leidt tot verplaatsing over een afstand, waarvan het gemiddelde evenredig is met de wortel van de tijdstap (Afbeelding 13.1). We komen hiermee op het verschil in concept tussen 'ruis' of 'random shocks' als drijvende kracht van verstoringen in discrete tijd en in een discrete (statistische) context, en de 'ruis' als drijvende kracht in continue tijd en een continue context (die al dan niet fysisch is). Dergelijke discrete, random of 'witte' ruis heeft een vaste variantie en kan als zodanig **wel** geschat of afgeleid worden op basis van meetreeksen, die in dat geval een vaste meetfrequentie dienen te hebben.

Het concept van ruis of verstoringen die continu plaatsvinden is dat van de Brownse beweging, waarvan alleen de **resultante** c.q. het effect waarneembaar of afgeleid kan worden op basis van meetreeksen (die dan gaten of een willekeurige meetfrequentie mogen hebben). De resultante c.q. de geschatte effecten worden in deze context ook niet 'ruis' maar 'innovaties' genoemd, de variantie van deze innovaties is niet vast maar afhankelijk van de tijdstap.

Afbeelding 13.1 Computersimulatie van drie sporen van Brownse beweging in een vlak, die steeds gedetailleerder berekend wordt: met 32 (blauw), 256 (lichtblauw) en 2048 stapjes (bron: www.nationalgeographic.es)

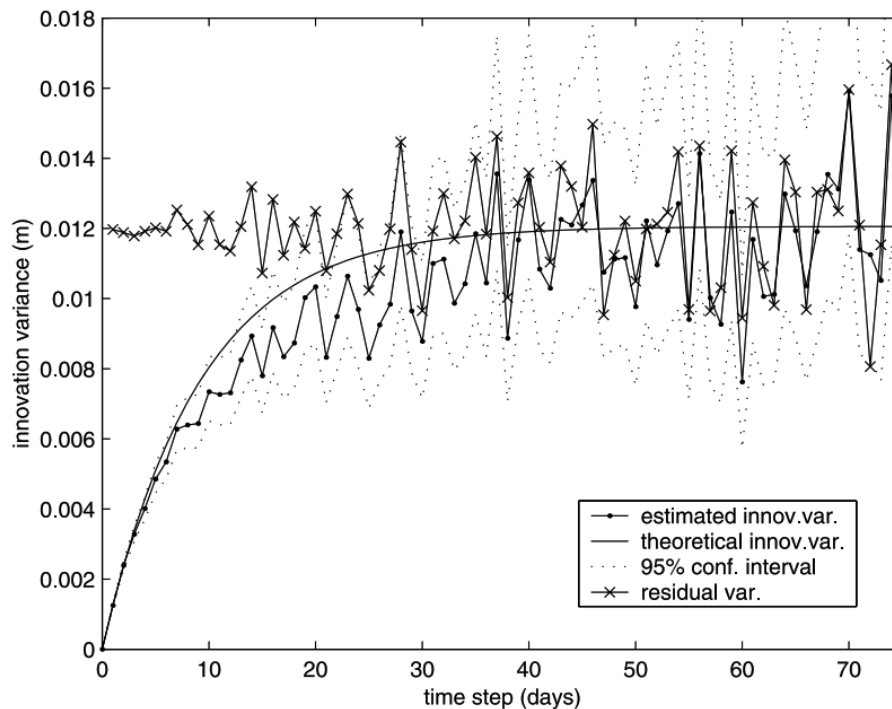


13.2 Gebruik binnen de tijdreeksanalyse

Binnen de (internationale) tijdreeksanalysetheorie en literatuur hebben de termen ruis, residuen en innovaties in principe ieder hun eigen en specifieke betekenis. Omdat er echter ook systematische verschillen bestaan tussen de gebruikte concepten in verschillende modellen of methoden (bijv. discrete tijd in het BJTR-model en continue tijd in het PIRFICT-model), kunnen er ook systematische verschillen voorkomen in de terminologie die gehanteerd wordt door verschillende auteurs in een verschillende context. Een dergelijk systematisch verschil in gebruik van terminologie geldt ook voor deze handleiding, waarbij eveneens opgemerkt dient te worden dat de termen 'ruis' en 'residuen' ook buiten de context van de tijdreeksanalyse gebruikt worden (zie volgende paragraaf).

Als basis voor de manier waarop de termen 'ruis', 'residuen' en 'innovaties' gebruikt worden binnen de tijdreeksanalyse gebruiken we de boeken van Box en Jenkins (Box en Jenkins, 1970; Box e.a., 2016), die gelden als grondleggers van de theorie van tijdreeks- en transferruismodellen. In de oorspronkelijke publicatie uit 1970 worden de termen 'white noise' en 'residuals' daarbij naast elkaar gebruikt als aanduiding voor of toelichting op symbool a_t . Voor symbool N_t wordt de term 'noise process' gebruikt. De nieuwe, vijfde editie uit 2016 bevat ook een passage over het gebruik van ARMA-modellen in zogeheten state-space vorm, voor toepassing op reeksen met missende waarnemingen. In die context wordt symbool a_t (of meer specifiek $a_{t|t-1}$) aangeduid als 'innovatie'. De term 'innovatie' komt daarbij zoals gezegd uit de filtertheorie, waarvan de oorsprong dus ligt bij de Brownse beweging. De variantie van deze innovaties is niet vast maar afhankelijk van de tijdstap, wat een belangrijk conceptueel verschil is met discrete witte ruis. Op dit aspect wordt o.a. nader ingegaan in het artikel waarin het ruismodel van het PIRFICT-model wordt beschreven (Von Asmuth en Bierkens, 2005). De innovatievariantie nadert daarbij naar nul bij het toenemen van de meetfrequentie (zie Afbeelding 13.2), wat complicaties oplevert voor het schattingsproces.

Afbeelding 13.2 Zogeheten innovatievariantiefunctie (IVF) of schattingen van de variantie van de residuen en innovaties als functie van de tijdstap, vergeleken met die van het achterliggende model (Von Asmuth en Bierkens, 2005)



13.3 Gebruik buiten de tijdreeksanalyse

De termen ruis, residu en innovatie worden ook buiten de specifieke context van de tijdreeksanalyse gebruikt, in breder en/of meer algemeen verband:

- de term **ruis** vindt zijn oorsprong in de elektronica en signaalverwerking. De effecten ervan werden voor iedereen hoorbaar bij de komst van de eerste radiotoestellen, en zichtbaar op de eerste televisies. Met ruis wordt in het algemeen bedoeld op ongewenste verstoringen van elektrische, akoestische of andersoortige signalen. De kleur van de ruis duidt daarbij op de spectrale eigenschappen van het signaal, waarbij 'witte ruis' een 'plat' of egaal frequentiespectrum heeft (de gemiddelde amplitude is voor iedere frequentie gelijk). Bij de PIRFICT-methode betreft de ruis de verstoringen van de invoerreeksen. Bij BJTR betreft de ruis de verstoringen van de uitvoerreeksen;
- de term **residuen** betekent volgens van Dale 'overblijfsels, m.n. na een chemisch proces'. De term wordt gebruikt als aanduiding van de 'voorspelfout', bijvoorbeeld bij regressieanalyse of fysisch-deterministische modellen, analoog aan de manier waarop Box en Jenkins hem ook gebruiken. Residu verwijst in het algemeen naar het deel van de metingen dat niet verklaard wordt door het model. Dit kan ofwel synoniem zijn aan de schattingen van a_t als 'one step ahead prediction error' van het ruismodel indien naar de 'verklaring' van het gecombineerde transferruismodel gekeken wordt, maar ook synoniem aan het deel dat niet deterministisch verklaard wordt (zoals bij een grondwatermodel). Bij PIRFICT betreft het residu de verstoringen van de uitvoerreeksen. Bij de BJTR methode betreft residu de verstoringen van de invoerreeksen. Verwarring kan ontstaan doordat een voorwaarde bij BJTR is dat de invoer van het ruismodel witte ruis vormt. De term **witte ruis** heeft echter betrekking op de eigenschappen van een reeks (kleur van de ruis) en niet zozeer op de naam;

- de term **innovatie** betekent in het algemeen spraakgebruik iets als 'vernieuwing'. Dit sluit in zoverre aan bij het gebruik binnen de tijdreeksanalyse, dat het daar om de nieuwe, onafhankelijke informatie gaat die besloten is in een nieuwe meting (het deel dat dus niet op basis van andere informatie voorspeld kan worden). Innovatie verwijst (binnen de tijdreeksanalyse) specifiek naar de reeks van 'one step ahead prediction errors' bij meetreeksen met gaten of onregelmatige frequenties, en/of bij gebruik van continue tijd.

Alhoewel er dus spraakverwarring bestaat waar het gaat om het gebruik van de termen 'ruis', 'residuen' en 'innovaties' binnen de tijdreeksanalyse, is er niet zozeer sprake van dat de ene term beter is dan de andere. De termen worden ook in de (inter)nationale literatuur naast elkaar gebruikt, omdat ze op verschillende aspecten van model, reeks of signaal, of de toepassing daarvan betrekking hebben.

LITERATUUR

- Baggelaar, P.K. (1988), Tijdreeksanalyse bij verlagingsonderzoek: principe en voorbeeld. H2O, vol. 21, nr. 16, blz. 443-450.
- Baggelaar, P.K. en E.C.J. Van der Meulen (2010), Statistische reconstructie invloed grondwaterwinning Nij Beets op grondwaterstanden. Icastat, november 2010, 95 blz.
- Baggelaar, P.K. en E.C.J. van der Meulen (2012), Trendanalyse op maat voor een meetnet waterkwaliteit. Stromingen 18 (2012), nummer 2, blz. 77 - 96.
- Baggelaar, P.K. en E.C.J. van der Meulen (2020), Naar betere tijdreeksmodellering met Pastas. AMO en PB Icastat, 28 april 2020, 70 blz.
- Baggelaar, P.K., Niet, A.C., Jaarsma, N., Ek, R., van Vught, A.C. (2018), Evaluatie reallocatie Mander. Witteveen+Bos, Amsterdam.
- Bakker, M., en O.D.L. Strack (2003), Analytic Elements for Multiaquifer Flow. Journal of Hydrology, 271(1-4), 119-129.
- Bear, J. (1972), American Elsevier Publishing Company, New York.
- Berendrecht, W. L., Heemink, A. W., Van Geer, F. C., en Gehrels, J. C. (2006), A non-linear state space approach to model groundwater fluctuations. Advances in water resources, 29(7), 959-973.
- Berendrecht, W., H. Gehrels, F. van Geer en A. Heemink (2003), Tijdreeksanalyse kan veel beter door kleiner meetinterval. Stromingen 9 (2003), nummer 1, blz. 5 - 22.
- Berendrecht, W., H. Gehrels, F. van Geer en A. Heemink (2004), Een niet-lineair tijdreeksmodel voor grondwaterstandsfluctuaties in gedraineerde gebieden. Stromingen 10 (2004), nummer 2, blz. 21 - 35.
- Berendrecht, W.L., A.W. Heemink, F.C. Van Geer en J.C. Gehrels (2003), Decoupling of modeling and measuring interval in groundwater time series analysis based on response characteristics; in: Journal of Hydrology, vol 278, pag 1-16.
- Berendrecht, W.L., en F.C. Van Geer (2016), A dynamic factor modeling framework for analyzing multiple groundwater head series simultaneously. Journal of Hydrology, Vol. 536, 50-60.
- Bidwell, V.J. (2005), Realistic forecasting of groundwater level, based on the eigenstructure of aquifer dynamics. Mathematics and Computers in Simulation, vol 69, pag 12-20.
- Bierkens, M.F.P., M. Knotters en F.C. van Geer (1999), Tijdreeksanalyse nu ook toepasbaar bij onregelmatige meetfrequenties. Stromingen 5 (1999), nummer 2, blz. 43 - 54.
- 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. Water Resources Research, vol 35, no 6, pag 1741-1750.
- Box, G.E.P. en G.M. Jenkins (1970), Time Series Analysis: Forecasting and Control. Holden-Day, San Fransisco.
- Box, G.E.P., G.M. Jenkins, G.C. Reinsel en G.M. Ljung (2016), Time Series Analysis: Forecasting and Control - Fifth Edition; John Wiley & Sons, Inc., Hoboken, New Jersey.
- Breusch, T. S. en A.R. Pagan (1979), A Simple Test for Heteroskedasticity and Random Coefficient Variation. Econometrica. 47 (5), blz. 1287 - 1294.
- Collenteur, R., Bakker, M., Klammler, G., en Birk, S. (2020), Estimating groundwater recharge from groundwater levels using non-linear transfer function noise models and comparison to lysimeter data. Hydrology and Earth System Sciences Discussions, 1-30.
- Collenteur, R.A. (2018), Over autocorrelatie van tijdreeksmodellen met niet-equidistante tijdstappen, Artesia, Schoonhoven, Nederland.
- 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

- Collenteur, R.A., M. van Baar en F. Schaars (2017), Evaluatie Hydrologisch Meetnet Brabant WaterFase 2: Hydrologische evaluatie van het meetnet.; Projectnummer 16.35.16, Artesia Water, Schoonhoven.
- D'Agostino, R.B. (1986), Tests for Normal Distribution. In Goodness-Of-Fit Techniques, edited by R.B. D'Agostino and M.A. Stephens, Marcel Dekker, 1986.
- Dirac, P.A.M. (1947), The Principles of Quantum Mechanics; Clarendon Press, Oxford.
- Dooge, J.C.I. (1959), A General Theory of the Unit Hydrograph. *J. Geophys. Res.*, vol 64, no 2, pag 241–256, doi:10.1029/JZ064i002p00241.
- Duhamel, J.M.C. (1833), Memoire sur la methode generale relative au mouvement de la chaleur dans les corps solides plonges dans les milieux dont la temperature varie avec le temps. *Journal d' Ecole polytechnique de Paris*, vol 14, no 22, pag 20.
- Einstein, A. (1905), Uber die von der molekular theoretischen Theorie der Warme geforderte Bewegung von in ruhenden Flussigkeiten suspendierten Teilchen; in: *Annalen der Physik*, vol 17, no 549, pag.
- Engle, R.F. (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50 (4), blz. 987 - 1007.
- Ferré, T.P. (2017), Revisiting the Relationship Between Data, Models, and Decision-Making. *Groundwater*, 55: 604-614.
- Geweke, J. (1977), The dynamic factor analysis of economic time series; in: *Latent variables in socio-economic models*, ed. D. J. Aigner en A. S. Goldberger, 365-383, Amsterdam.
- Hipel, K.W. en A.I. McLeod (1994), *Time series modelling of water resources and environmental systems*; Elsevier, Amsterdam.
- 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.
- Jarque, C.M. en A.K. Bera (1980), Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*. 6 (3), blz. 255 – 259.
- Jarque, C.M. en A.K. Bera (1981), Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence. *Economics Letters*. 7 (4), blz. 313 – 318.
- Jarque, C.M. en A.K. Bera (1987), A test for normality of observations and regression residuals. *International Statistical Review*. 55 (2), blz. 163 – 172.
- Jury, W.A. (1982), Simulation of solute transport using a transfer function model. *Water Resources Research*, vol 18, pag 363-368.
- Jury, W.A. en K. Roth (1990), *Transfer functions and solute movement through soil: theory and applications*; Birkhäuser Verlag, Basel, Switzerland.
- Kalman, R.E. (1960), A new approach to linear filtering and prediction problems; in: *J. bas. engng. trans. ASME D*, vol 82, pag 35-45.
- Klemes, V. (1978), Physically based stochastic hydrologic analysis. *Advances in Hydrosience*, vol 11, pag 285-356.
- Knotters, M. (2012), Validatie van tijdreeksmodellen voor de grondwaterstand. *Stromingen* 18 (2012), nummer 2, blz. 31 t/m 42.
- Knotters, M. en De Gooijer, J.G. (1999), TARSO modeling of water table depths. *Water Resources Research*, Vol. 35, No. 3, pp. 695-705, March 1999.
- Knotters, M. en M.F.P. Bierkens (1999), Hoe lang moet je de grondwaterstand meten om iets over de dynamiek te meten? *Stromingen* 5 (1999), nummer 4, blz. 5 – 12.
- Knotters, M. en M.F.P. Bierkens (2000), Physical basis of time series models for water table depths. *Water Resources Research*, vol 36, no 1, pag 181-188.

- Knotters, M., P. Baggelaar en E. Van der Meulen (2020), Hoezo significant? Over het effect van een ingreep op de grondwaterstand. *Stromingen* 26 (2020), nummer 3, blz. 37 t/m 48.
- Konikow, L.F. en J.D. Bredehoeft (1992), Groundwater models cannot be validated. *Advances in Water Resources*, Vol. 15, No. 1, 75-83
- Kruihof, A.J. (2001), De impuls-responsfunctie van de grondwaterspiegel, scheiden van de invloed van de onverzadigde en de verzadigde zone; Master Thesis, TU Delft, BTO no. 2001.163(s) Kiwa Water Research, Delft, Nieuwegein.
- KWR (2014), Hydrologisch meetnet Electraboezem 2e schil, effecten van boezempeilverlaging en bodemdaling.; Rapportnr KWR 2014.039, KWR Watercycle Research Institute, Nieuwegein.
- Ljung, G.M en G.E.P. Box (1978), On a measure of lack of fit in time series models. *Biometrika*, 65, blz. 297 – 303.
- Ljung, L. (1999), *System Identification, Theory for the user*; Prentice Hall, Upper Saddle River.
- Maas, C. (1994), *On Convolutional Processes and Dispersive Groundwater Flow*; PhD Thesis, Delft University of Technology, Delft.
- Maas, C. (2011), *Grondwatermodellen versus Tjddreksanalyse*, KWR, Nieuwegein.
- McLeod, G. (1983), *Box Jenkins in Practice*. Time Series Library, Gwilym Jenkins & Partners Ltd., Lancaster.
- Nash, J.E. (1958), Determining runoff from rainfall. *Proc. Inst. Civ. Eng.*, vol 10, pag 163-184.
- Obergfell, C., Bakker, M., en Maas, K. (2019), Identification and explanation of a change in the groundwater regime using time series analysis. *Groundwater*, 57(6), 886-894.
- Obergfell, C., M. Bakker en K. Maas (2016), A time-series analysis framework for the flood-wave method to estimate groundwater model parameters. *Hydrogeology Journal*, vol, pag 1-13.
- Oreskes, N., K. Schrader-Frechette en K. Belitz (1994), Verification, validation and confirmation of numerical models in the earth sciences. *Science*, vol 263, pag 641–646.
- Pappenberger, F., en K. J. Beven (2006), Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resour. Res.*, 42, W05302.
- Parlange, M.B., G.G. Katul, R.H. Cuenca, M.L. Kavvas, D.R. Nielsen en M. Mata (1992), Physical basis for a time series model of soil water content. *Water Resources Research*, vol 28, no 9, pag 2437-2446.
- Peterson, T. J., en Western, A. W. (2014), Nonlinear time-series modeling of unconfined groundwater head. *Water Resources Research*, 50(10), 8330-8355.
- Ritzema, H.P., G.B.M. Heuvelink, M. Heinen, P.W. Bogaart, F.J.E. van der Bolt, M.J.D. Hack-ten Broeke, T. Hooglen, M. Knotters, H.T.L. Massop en H.R.J. Vroon (2012), Meten en interpreteren van grondwaterstanden – Analyse van methodieken en nauwkeurigheid. *Alterra-rapport 2345*, Wageningen, 122 blz.
- Rolf, H.L.M. (1989), *Verlaging van de grondwaterstanden in Nederland. Analyse periode 1950-1986 (Lowering of the water table in the Netherlands. Analysis of the period 1950-1986)* (in Dutch); DGV-TNO, Delft.
- Rolf, H.L.M. en J.J. Lebbink (1998), *Hydrologische effecten als gevolg van de reductie duinwaterwinning in het Noord-Hollands Duinreservaat*. NV PWN Waterleidingbedrijf Noord-Holland, 1998.
- Sahuquillo, A. (1983), an eigenvalue numerical technique for solving unsteady linear groundwater models continuously in time. *Water Resources Research*, vol 19, no 1, pag 87-93, doi:10.1029/WR019i001.
- Shapiro, S.S. en M.B. Wilk (1965), An analysis of variance test for normality (complete samples). *Biometrika*. 52
- Singh, A., Mishra, S. en Ruskauff, G. (2010), Model Averaging Techniques for Quantifying Conceptual Model Uncertainty. *Groundwater*, 48: 701-715.

- Stock, J.H. en M. Watson (2011), Dynamic factor models; in: Oxford Handbook on Economic Forecasting, ed. M. P. Clements en D. F. Hendry. Oxford University Press, Oxford
- Stoffer, D.S. en C.M.C. Toloï (1992), A note on the Ljung-Box-Pierce portmanteau statistic with missing data. *Statistics & probability Letters* 13 (1992), blz. 391 – 396.
- Van der Meulen, E.C.J. en P.K. Baggelaar (2019), Tjiddreeksanalist (TRGnet en TRG) - Tjiddreeksanalyse met uitgebreide Box-Jenkins modellering – Versie 6. AMO, Hengelo, januari 2019, 114 blz.
- van der Spek, J.E. en Bakker, M. (2017), The influence of the length of the calibration period and observation frequency on predictive uncertainty in time series modeling of groundwater dynamics. *Water Resources Research*, 53(3), pp.2294-2311.
- van Geer, F. (2012), Tjiddreeksanalyse: Introductie en aandachtspunten. *Stromingen* 18, nummer 2.
- Van Geer, F.C. en A. Lourens (2001), Trends in de stijghoogte en relatie tussen verschillende dieptes. TNO-rapport NITG 01-125-B, juli 2001, NITG-TNO, Delft.
- Van Geer, F.C. en W.L. Berendrecht (2015), Meervoudige tijdreeksmodellen en de samenhang in stijghoogtereeksen; in: *Stromingen*, vol 23, no 3, pag 25-36.
- Van Geer, F.C., P.K. Baggelaar en P.R. Defize (1988), Toepassing van tijdreeksanalyse op meetreeksen van de stijghoogte. *H20*, vol. 21, no .161, blz. 438-442.
- Van Herwijnen, T. (2002), Tjiddreeksanalyse van grondwaterpeilen: de invloed van het meetinterval. *Stromingen* 8, nummer 4, blz. 19 - 30.
- Vitens (2007), Factsheet Vitens Visie Duurzame Woningen (fase 2). Winning Nij Beets. Versie 29 oktober 2007.
- von Asmuth, J. R., Bierkens, M. F., en Maas, K. (2002), Transfer function-noise modeling in continuous time using predefined impulse response functions. *Water Resources Research*, 38(12), 23-1.
- Von Asmuth, J. R., Maas, K., Knotters, M., Bierkens, M. F., Bakker, M., Olsthoorn, T. N., Cirkel, D.G., Leunk, I., en Von Asmuth, D. C. (2012), Software for hydrogeologic time series analysis, interfacing data with physical insight. *Environmental Modelling & Software*, 38, 178-190.
- von Asmuth, J., van der Schaaf, S., Grootjans, A., en Maas, K. (2012), Vennen en veentjes: (niet-)ideale systemen voor niet-lineaire tijdreeksmodellen, *Stromingen* 18, Nummer 2.
- Von Asmuth, J.R. (2012), Groundwater System Identification, through Time Series Analysis; Ph.D. thesis, Delft University of Technology, Delft.
- Von Asmuth, J.R. en I. Leunk (2012), Winterpeilaanpassing Veerse Meer. (Niet-lineaire) hydrologische situatie en effecten op watervoerend pakket en deklaag; Rapportnr KWR 2012.091, KWR Watercycle Research Institute, Nieuwegein.
- Von Asmuth, J.R. en M.F.P. Bierkens (2005), Modeling irregularly spaced residual series as a continuous stochastic process. *Water Resources Research*, vol 41, no 12, pag W12404, doi:10.1029/2004WR003726.
- Von Asmuth, J.R., C. Maas, M. Knotters, M.F.P. Bierkens, M. Bakker, T.N. Olsthoorn, D.G. Cirkel, I. Leunk, F. Schaars en D.C. Von Asmuth (2012), Software for hydrogeologic time series analysis, interfacing data with physical insight. *Environmental Modelling & Software*, vol 38, pag 178-190, <http://dx.doi.org/10.1016/j.envsoft.2012.06.003>.
- Von Asmuth, J.R., K. Maas, M. Bakker en J. Petersen (2008), Modeling time series of groundwater head fluctuations subjected to multiple stresses. *Ground Water*, vol 46, doi: 10.1111/j.1745-6584.2007.00382.x, no 1, pag 30-40.
- Von Asmuth, J.R., M.F.P. Bierkens en C. Maas (2002a), Soms is weten beter dan meten (tenzij je verkeerd zit natuurlijk). Het discrete Box-Jenkins- versus het continue PIRFICT-tjiddreeksmodel, in praktijk; in: *Stromingen*, vol 8, no 1, pag 5-14.
- Von Asmuth, J.R., M.F.P. Bierkens en K. Maas (2002b), Transfer function noise modeling in continuous time using predefined impulse response functions; in: *Water Resources Research*, vol 38, no 12, pag 1287-1299, doi:10.1029/2001WR001136.
- Walvoort, D., M. Knotters en T. Hooglen (2013), Een tool voor controle van hoogfrequente grondwaterstandsreeksen. *Stromingen* 19, nummer 3 & 4, blz. 49 – 62.

Wiener, N. (1949), *Extrapolation, Interpolation and Smoothing of Stationary Time Series*; MIT press, 56-103, Cambridge, Massachusetts.

Zadeh, L.A. (1956), On the identification problem. *IRE Transactions on Circuit Theory*, vol 3, pag 277-281.

Ziemer, R.E., W.H. Tranter en D.R. Fannin (1998), *Signals and systems: continuous and discrete*; Prentice-Hall, Upper Saddle River.