model_structuur

December 7, 2020

1 Model structuur

Auteurs: R.A. Collenteur (Universiteit Graz)

Een belangrijke stap in tijdreeksanalyse is het bepalen van de modelstructuur van het tijdreeksmodel. Dit is het onderwerp van dit Notebook. We beperken ons hier tot het bepalen van het deterministische deel van het model. Het bepalen van het ruismodel is onderdeel van het notebook over modelkalibratie. In dit notebook wordt een overzicht gegeven van de verschillende opties die mogelijk zijn voor de modelstructuur.

1.0.1 Inhoudsopgave

1.1 Algemeen stappenplan

Voor het kiezen van de gewenste modelstructuur is het belangrijk om vast te stellen wat het doel van het tijdreeksmodel is. Wanneer het doel van de modellering is vastgesteld, kunnen de volgende stappen worden doorlopen:

1. **Systeem analyse**: bepaal mogelijke hydrologische variabelen,
2. **Model bouw**: Bepaal hoe deze variabelen om te zetten 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 en derde stap. Eerst wordt een overzicht gegeven van de bouwstenen die beschikbaar zijn om een model te bouwen. Daarna kijken we naar methoden om te bepalen welke bouwstenen nodig zijn in een model.

1.2 Basis model structuur

Het tijdreeksmodel kan in algemene vorm als volgt worden geschreven:

$$h(t) = \sum_{m=0}^{M} h_m(t) + d + r(t)$$

Waar $h(t)$ de gemeten stijghoogte is, h_m de bijdrage van invloed m, d het basisniveau van het model en $r(t)$ de residuen. Er kunnen meerdere invloeden bijdragen aan de fluctuaties in de stijghoogte. In dit notebooks gaat het voornamelijk over het bepalen van deze bijdragen, en of een invloed moet worden meegenomen in het model.

1.2.1 Het basisniveau van het model

Vrijwel elk tijdreeksmodel bevat een constante d waarmee het model op de juiste hoogte wordt gebracht. Hoewel verleidelijk, is het niet altijd juist een physische 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 zou convergeren wanneer alle invloeden nul zijn. Of deze interpretatie juist is hangt echter af van het exacte model. Beter is daarom om eenvoudig naar d te verwijzen als het basisniveau van het model.

Voor het bepalen van de exacte waarde van de constante zijn verschillende methoden beschikbaar. De eerste optie is om na een simulatie het gemiddelde van de residuen te nemen als waarde voor d (e.g., von Asmuth et al., 2002). Een andere optie is om d eenvoudigweg mee te schatten als een parameter (e.g., Collenteur et al. 2019). Dit heeft als voordeel dat er ook iets gezegd kan worden over de onzekerheid van deze parameter. Meer onderzoek is nodig om vast te stellen wat het effect is van het verschil tussen deze twee methoden om de uiteindelijke resultaten.

1.2.2 Bijdragen

Stijghoogte fluctuaties kunnen door vele verschillende oorzaken ontstaan. Vaak hebben we een idee welke oorzaken belangrijk zijn, maar net zo vaak is dit juist het doel van de modellering. Wat het exacte effect van een bepaalde invloed kan afhangen van de gekozen vorm van de modellering, oftewel hoe de invloed vertaald wordt in een bijdrage aan de stijghoogte fluctuaties.

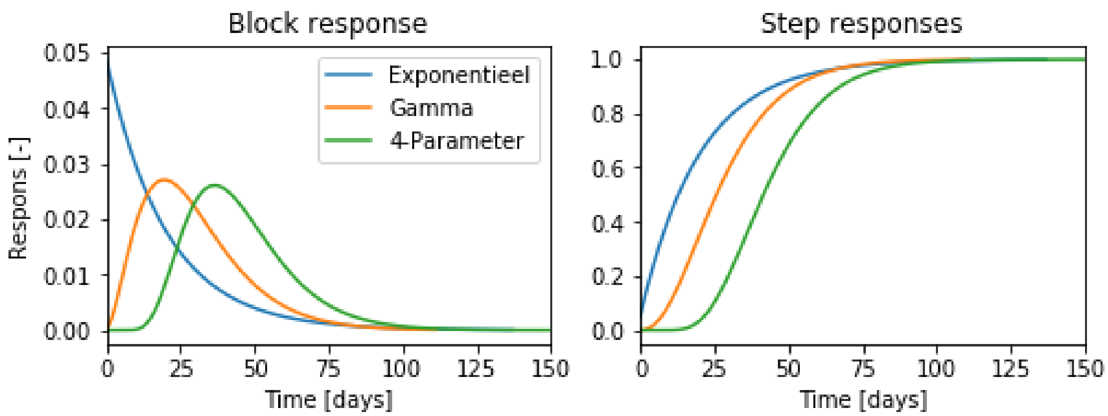
We kunnen twee typen invloeden onderscheiden: gemeten en niet gemeten invloeden. Voor gemeten invloeden zijn verklarende tijdreeksen beschikbaar om de bijdrage van een invloed aan de stijghoogte fluctuaties te bepalen. Dit zijn bijvoorbeeld: neerslag, potentiele verdamping, ontrekkingen, oppervlakte waterstanden en luchtdruk metingen. Dit type invloeden kan worden vertaald in een bijdrage met behulp van een impuls respons functie.

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, ingreep in de waterhuishouding (bv. opzet Ijsselmeer). Dit type bijdrage kan worden meegenomen door een trend mee te nemen in de modellering. De op te geven informatie is in dit geval een periode of tijdstap wanneer een verandering heeft plaats gevonden.

1.3 Respons functies

De vertaling van een verklarende tijdreeks naar een bijdrage aan de stijghoogte fluctuaties kan worden gedaan met behulp van een impuls respons functies. De impuls respons functie beschrijft in dit geval hoe de stijghoogte reageert op een impuls van de verklarende tijdreeks. Er zijn tal van respons functies die kunnen worden gekozen om een verklarende tijdreeks om te zetten in een bijdrage aan de stijghoogte sfluctuaties. De keuze voor een bepaalde respons functie kan bijvoorbeeld afhangen van het type verklarende reeks (bv. neerslag of een rivier), maar ook van het systeem (bv. een dikke of dunne overzadigde zone).

Vaak wordt een 4-parameter gamma functie of een vereenvoudigde vorm daarvan gebruik als impuls respons functie. Afhankelijk van het systeem kunnen bepaalde parameters worden vastgezet, wat het aantal te schatten parameters verlaagt. In de volgende figuur worden drie impuls respons functie getoond die elk een vereenvoudigde vorm van de 4-parameter gamma functie zijn.



De exponentiele respons heeft 2 parameters, de Gamma heeft 3 parameters, en de 4-parameter gamma functie heeft 4 parameters. Het kan echter goed voorkomen dat de parameters van de 4-parameter functie na kalibratie eigenlijk een Gamma of zelfs exponentiele functie simuleren. Het is dan mogelijk parameters vast te zetten en op die manier het aantal te schatten parameters te reduceren. Het dient over het algemeen aanbeveling verschillende respons functies te testen. Per software pakket dient te worden bepaalt welke respons functies beschikbaar zijn.

1.4 Effect van neerslag en verdamping

Vrijwel alle grondwaterstanden in Nederland worden in meer of mindere mate beinvloed door neerslag en verdamping. Deze twee verklarende tijdsreeksen worden daarom bijna altijd in een tijdreeksmodel opgenomen. De manier waarop deze tijdreeksen worden meegenomen in het model kunnen sterk verschillen. We kunnen grofweg drie methoden onderscheiden:

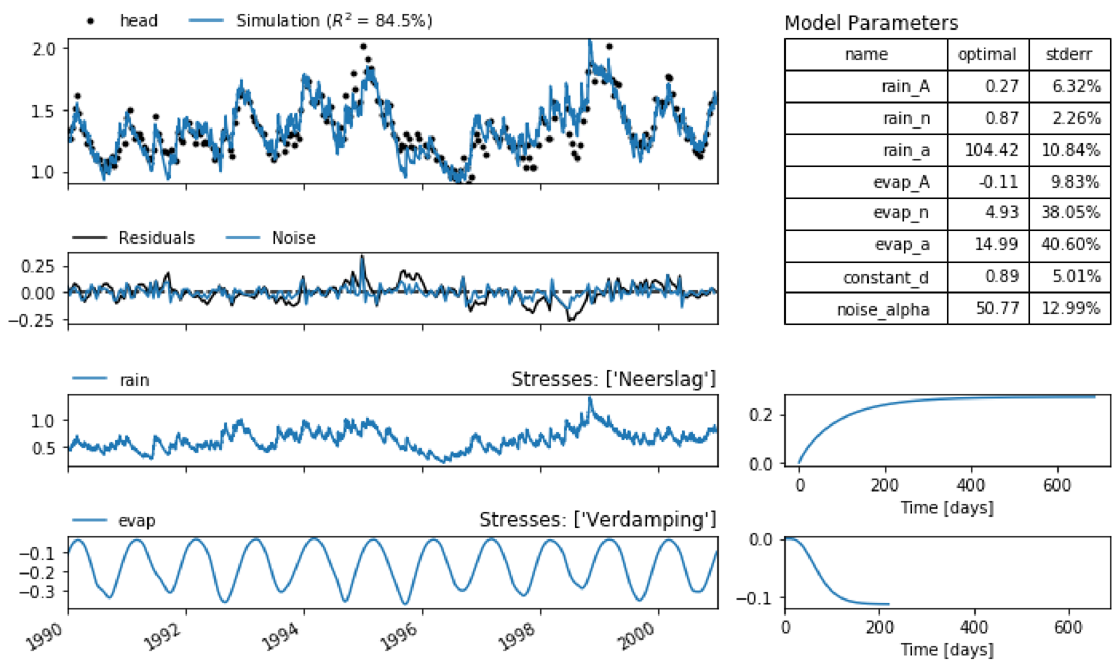
- Neerslag en verdamping als aparte invloeden
- Netto neerslagoverschot ($N = P - fE$) en 1 respons functie
- Niet-lineaire grondwateraanvulling en 1 respons functie

1.4.1 Voorbeeld data

In de volgende drie secties wordt een model gemaakt volgens deze drie methoden. De stijghoogte tijdreeks die in dit voorbeeld is gebruikt is gemeten in een monitorings put B32C0639 (filter 1) nabij de Bilt. Neerslag en potentiele Makkink verdamping is beschikbaar van het KNMI meteorologisch meetstation de Bilt.

1.4.2 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 (en wederom onafhankelijke) responsfuncties worden de tijdreeksen vertaald in bijdragen aan de stijghoogte fluctuaties. Hieronder is een voorbeeld gegeven van het resultaat van een dergelijk model voor de voorbeeld data.

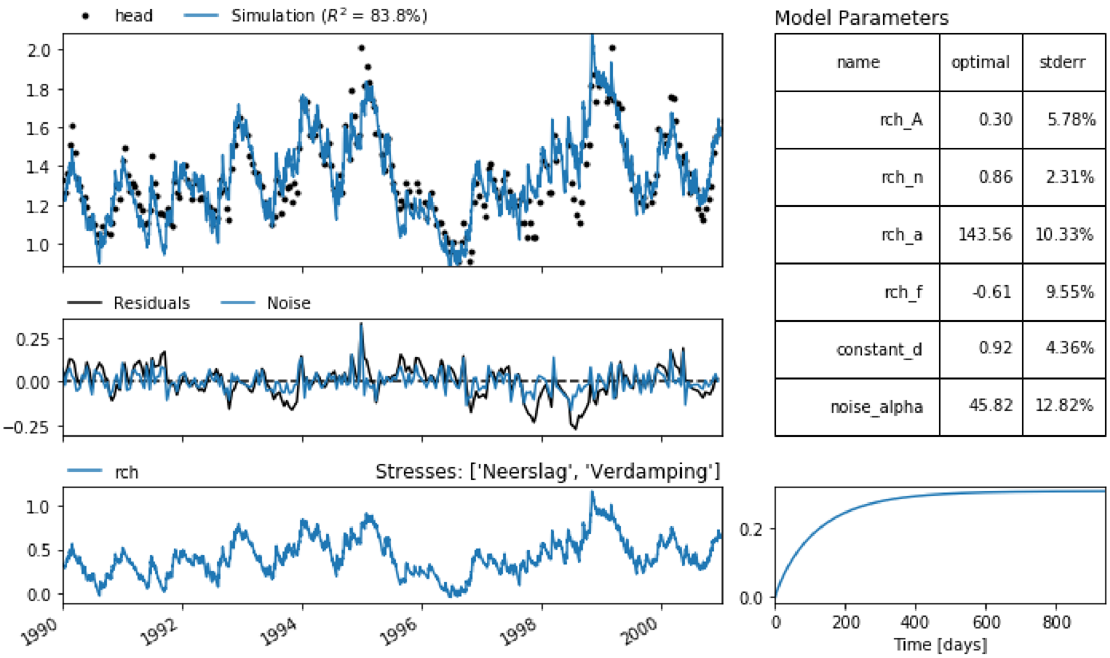


De onderste twee linker subplots tonen de geschatte invloed van de neerslag en de verdamping. De neerslag heeft een positief (stijgend) effect op de stijghoogte en de verdamping een dalend effect. Voor beide respons functies (zie de onderste twee rechter subplots) moeten de parameters worden geschat. Het totaal aantal geschatte parameters van dit model is 8.

1.4.3 Lineair neerslagoverschot

Een veelgemaakte aanname is dat de respons van de stijghoogte op neerslag en verdamping vergelijkbaar is, maar in de tegenovergestelde richting. 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 ($N = P - E$) en deze flux te vertalen in een bijdrage aan de stijghoogtefluctuaties met een enkele respons functie.

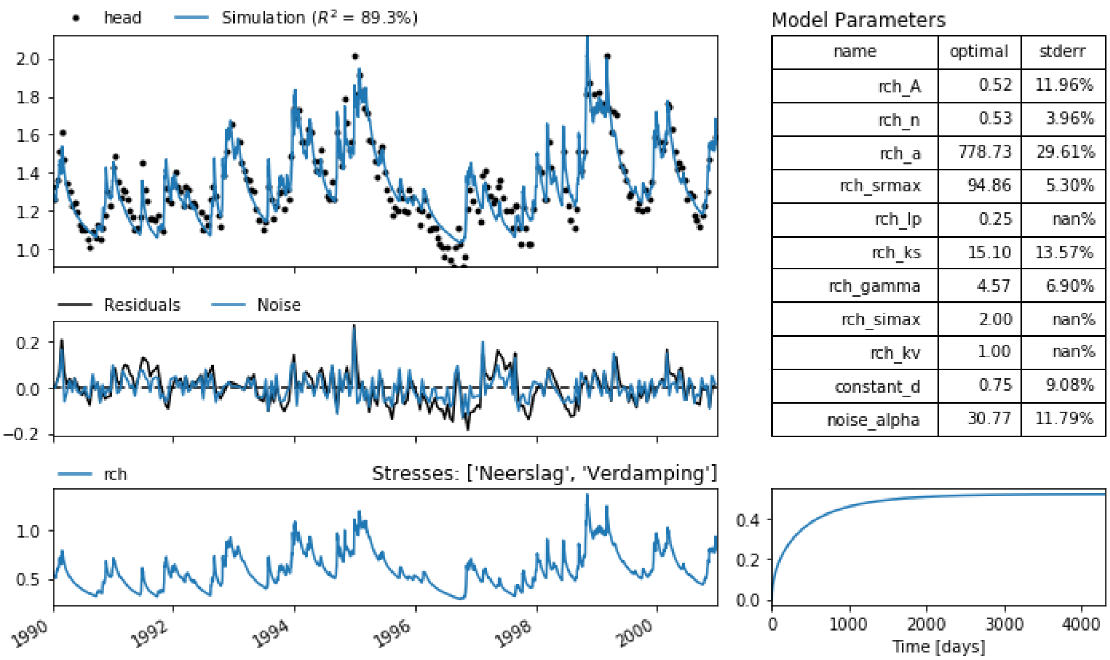
Omdat vaak gewerkt wordt met de potentiele verdamping (bijvoorbeeld Makkink of Penman-Monteith) wordt een verdampingsfactor f gebruikt. Het neerslag overschot wordt dan berekend als : $N = P - f * E$, waar f een kalibratie parameter is. Een belangrijk voordeel van dit model ten op zichte van het voorgaande model is het kleinere aantal parameters dat geschat moet worden.



In de figuur hierboven is het resultaat van dit model voor de voorbeeld data getoond. De geschatte verdampingsfactor is $f = -0.61$. Terwijl dit model twee kalibratie parameters minder heeft is de model fit vrijwel gelijk aan het eerste model.

1.4.4 Niet-lineaire grondwateraanvulling

Bij de derde methode wordt een niet-lineair model gebruikt om grondwateraanvulling te berekenen, die vervolgens ook geconvolueerd wordt met een respons functie. Verschillende niet-lineaire modellen zijn beschikbaar, zie bijvoorbeeld Berendrecht et al. (2005), Peterson en Western (2014) en Collenteur et al. (2020). Een voordeel van niet-lineaire modellen is dat er rekening wordt gehouden met processen in de onverzadigde zone, waardoor de grondwaterstand niet meer lineair reageert op neerslag en verdamping. Dit gebeurt bijvoorbeeld wanneer de verdamping wordt beperkt door de de beschikbare bodemvocht ten tijd van droogte, of door de berging van neerslag in de onverzadigde zone.



name	optimal	stderr
rch_A	0.52	11.96%
rch_n	0.53	3.96%
rch_a	778.73	29.61%
rch_srmax	94.86	5.30%
rch_lp	0.25	nan%
rch_ks	15.10	13.57%
rch_gamma	4.57	6.90%
rch_simax	2.00	nan%
rch_kv	1.00	nan%
constant_d	0.75	9.08%
noise_alpha	30.77	11.79%

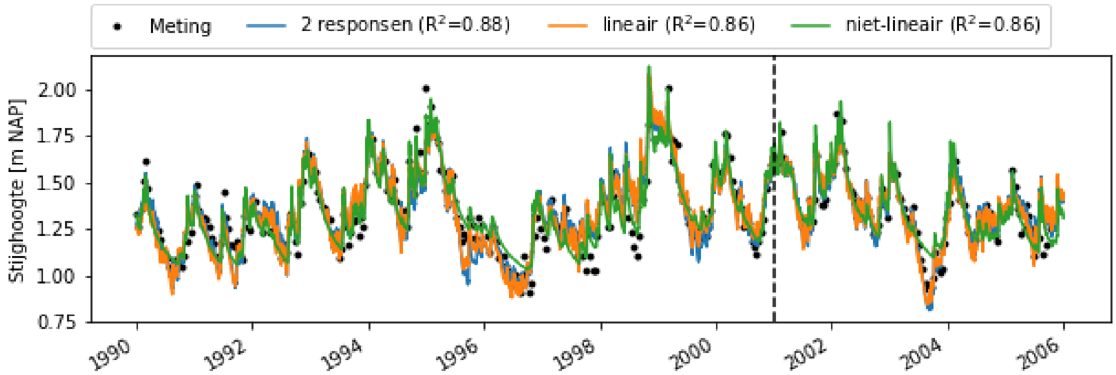
Mogelijke nadelen van niet-lineaire modellen kunnen een groter aantal vrije parameters zijn of de rekentijd van het model. Het is echter vaak mogelijk een aantal parameters vast te zetten en een response functie te gebruiken met minder parameters, waardoor het aantal vrije parameters sterk is te reduceren. Door ontwikkelingen in de software is de rekentijd eigenlijk geen argument meer om niet-lineaire modellen niet te gebruiken.

1.4.5 Welk model is het beste?

We hebben nu drie modellen voor een stijghoogte reeks die goed verklaart kan worden door neerslag en verdamping. Het model met twee aparte responsen heeft een $R^2 = 0.85$, het lineaire model $R^2 = 0.84$ en het niet-lineaire model $R^2 = 0.89$. De modellen ontlopen elkaar dus eigenlijk niet zo veel wat betreft deze fit statistiek. Het is aan de modelleur om nu 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. In de figuur hieronder is met elk model de stijghoogte in voor een periode na de kalibratieperiode gesimuleerd (2001-2006). In de legenda is de determinatie coefficient weergegeven. Nu is het model met de twee responsen het beste model volgens de R^2 statistiek ($R^2 = 0.88$), en hebben het lineaire en niet-lineaire model een slechtere fit ($R^2 = 0.86$).

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 ($n = 6$), en de andere twee modellen hebben beide meer parameters hebben (beide $n = 8$).

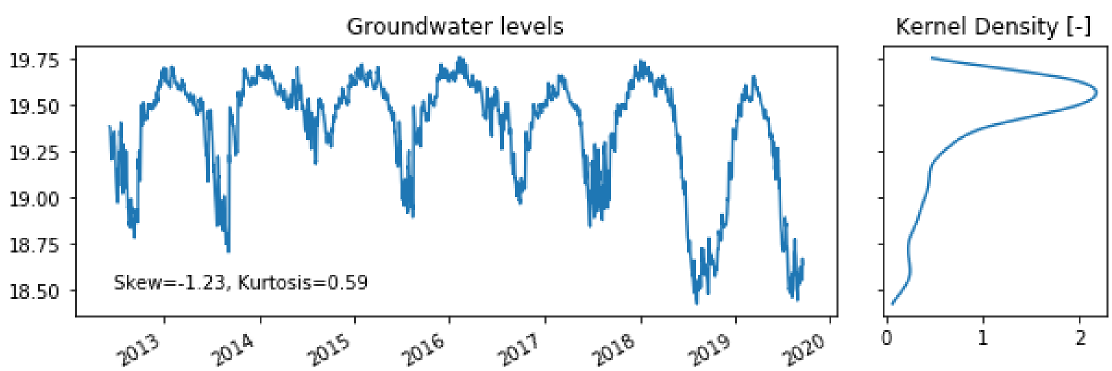


1.5 Drempel niet-lineariteit

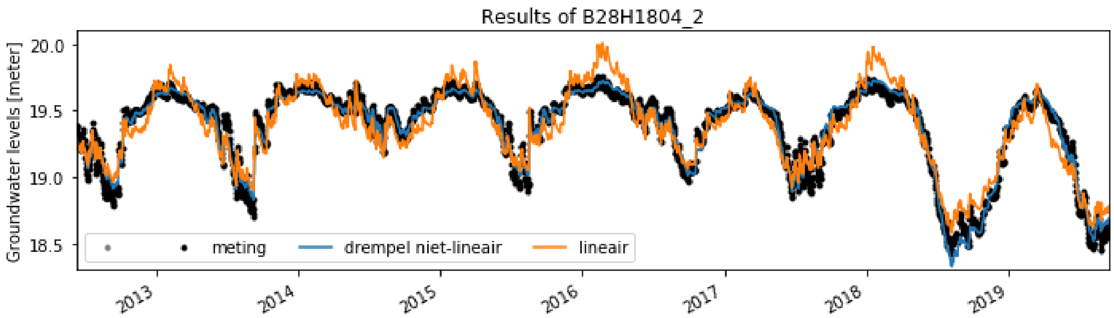
Een veelvoorkomend fenomeen in Nederland is drempel niet-lineariteit. Daarbij reageert een grondwatersysteem anders op een impuls (bv. neerslag) nadat een bepaalde grenswaarde in de stijghoogte is overschreden. Knotters et al. (1999) hebben hiervoor een model ontwikkeld, waarbij de respons op neerslagoverschot afhangt van de stijghoogte boven of onder een bepaalde drempelwaarde.

Of er sprake is van drempel niet-lineariteit kan worden vastgesteld door:

1. een visuele inspectie of de stijghoogte tijdreeks afgevlakt is,
2. het plotten van een histogram van de stijghoogte metingen, of
3. het testen van verschillende modelstructuren.



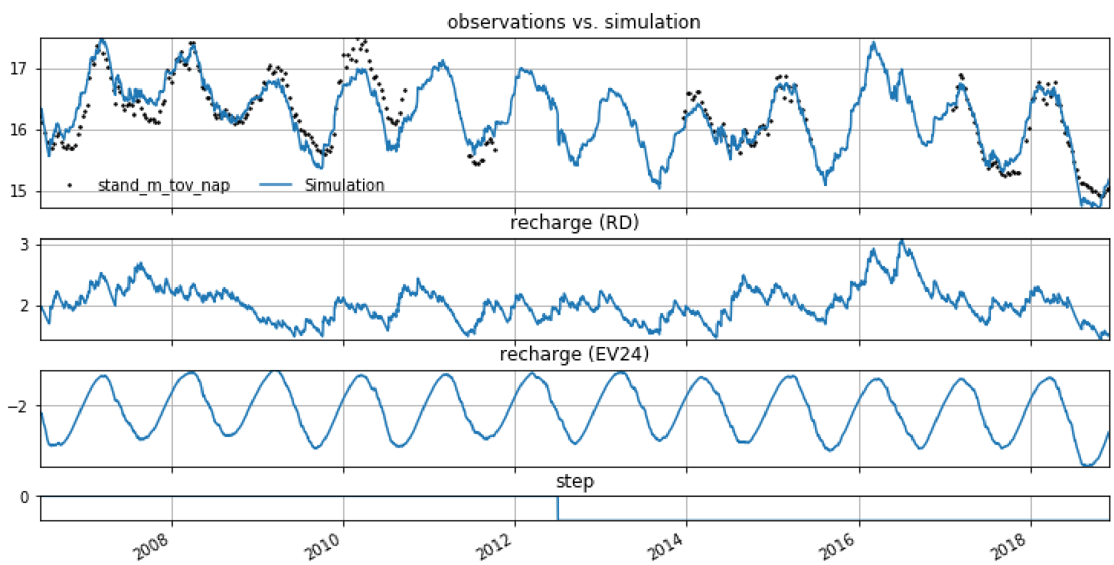
Uit bovenstaande plots is af te lezen dat er mogelijk sprake is van drempel niet-lineariteit. We maken daarom twee modellen: de eerste met een simpel neerslag overschot en een tweede met drempel niet-lineariteit. Onderstaande figuur geeft de resultaten van de twee modellen weer. Het lineaire model geeft een structurele overschatting van de hoge grondwaterstanden, waar het drempel niet-lineaire model netjes de gemeten grondwaterstand volgt.



1.6 Veranderende Systemen

Systemen kunnen niet allen veranderen als gevolg van de grondwaterstand (bv. drempel niet-lineariteit), maar ook als gevolg van systeemsveranderingen door de tijd. - Stap response - Lineaire trend - Veranderende respons functies (Obergfjell et al. (2019)).

1.6.1 Stap response



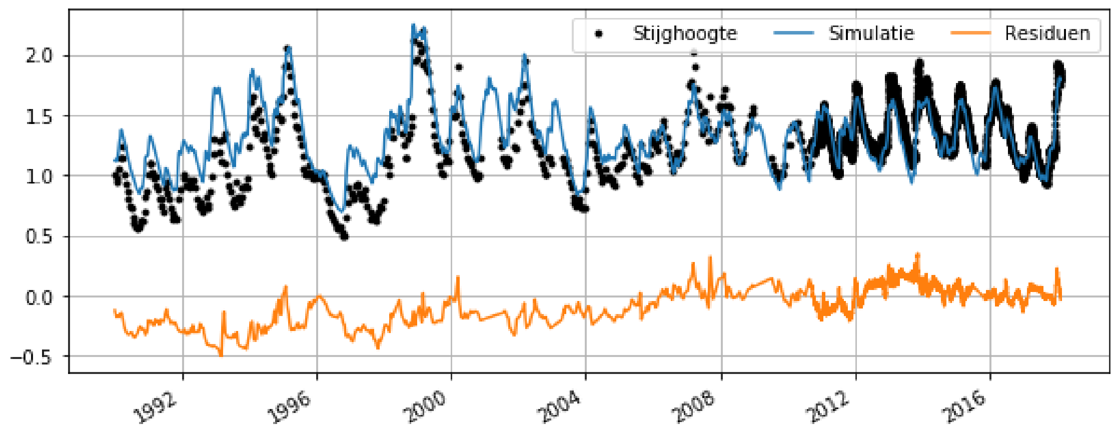
1.6.2 Voorbeeld van een model met lineaire trend

Het gebeurt vaak dat een stijghoogte tijdreeks een opwaarste of neerwaartse trend vertoond, die niet door neerslag of verdamping verklaart kan worden. We beschouwen hier een model voor het

1e filter in de peilbuis “B32C0609” nabij de Bilt. Als eerste stap maken we een model met neerslag en verdamping als verklarende variabelen.

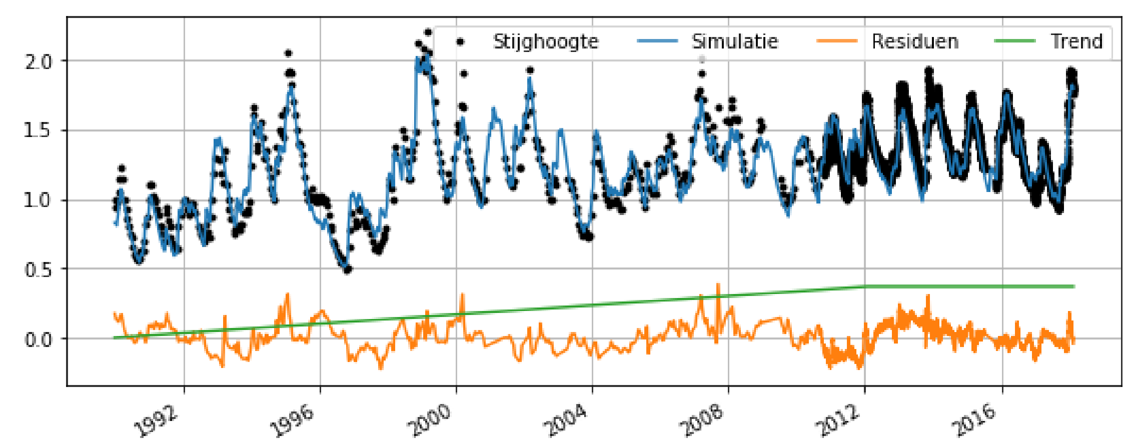
```
Fit report Stijghoogte              Fit Statistics
====================================================
nfev     48                     EVP            76.08
nobs     3016                   R2              0.76
noise    False                  RMSE            0.12
tmin     1990-01-01 00:00:00    AIC              nan
tmax     2018-02-14 00:00:00    BIC              nan
freq     D                      Obj            21.98
warmup   3650 days 00:00:00     ___
solver   LeastSquares           Interpolated      No

Parameters (6 were optimized)
====================================================
                optimal    stderr     initial   vary
rch_A          0.640193    ±2.21%    0.772130   True
rch_a        161.991751    ±2.77%   10.000000   True
rch_srmax    502.071955   ±10.37%  250.000000   True
rch_lp         0.250000     ±nan%    0.250000  False
rch_ks      9999.735234  ±103.36%  100.000000   True
rch_gamma      5.497705   ±11.25%    4.000000   True
rch_simax      2.000000     ±nan%    2.000000  False
rch_kv         1.000000     ±nan%    1.000000  False
constant_d     0.501271    ±4.35%    1.322862   True

Parameter correlations |rho| > 0.5
====================================================
rch_A     rch_a       0.83
          constant_d -0.91
rch_a     constant_d -0.63
rch_srmax rch_ks      0.81
          rch_gamma   0.73
rch_ks    rch_gamma   0.99
```

```
<matplotlib.legend.Legend at 0x7f9be865fb50>
```



In de figuur hierboven is het resultaat van de modellering te weergegeven. We zien dat het model de fluctuaties in de gemeten stijghoogte in het algemeen goed kan verklaren, maar de residuen van het model laten een duidelijke opwaarste trend zien. Het model overschat de stijghoogte in het begin van de kalibratie periode en onderschat de stijghoogte aan het einde van deze periode.

```
<matplotlib.legend.Legend at 0x7f9be8b88a10>
```



1.7 Verklarende variabelen

Welke verklarende variabelen moeten worden meegenomen

- Toetsen van hydrologische plausibiliteit van evenwichtsrelatie (bv. well_A > 0), diagnostisch toetsen noodzakelijk
- Lengte van de respons functie
- Residuen analyse

Einstein, "everything should be as simple as possible, but not simpler." Occums razor

1.8 Discussie

1.9 Referenties

- Knotters et al. (1999)
- von Asmuth et al. (2002)
- Berendrecht et al. (2006)
- Peterson and Western (2014)
- Collenteur et al. (2020)

1.10 Beschikbaarheid van opties

In de volgende tabel wordt een overzicht gegeven welke opties in verschillende software pakketten beschikbaar zijn. **Dit overzicht komt niet in dit Notebook**

Optie	Pastas	Menyanthes	Hydrosight
Drempel niet-lineariteit	Ja	Ja	Nee
niet-lineaire grondwateraanvulling	Ja	Nee	Ja
Stap respons	Ja	Ja	Onbekend
Lineaire trend	Ja	Ja	Onbekend
Veranderende respons	Nee	Nee	Nee