



Methoden voor trendanalyse bij de Evaluatie Meststoffenwet

Pim Ouwehand
Centraal Bureau voor de Statistiek (CBS)

Werkgroep Statistiek Evaluatie Meststoffenwet

Met medewerking van:

Arthur Denneman (Centraal Bureau voor de Statistiek, CBS)
Dico Fraters (Rijksinstituut voor Volksgezondheid en Milieu, RIVM)
Lan Ge (Wageningen Economic Research, WEER)
Joachim Rozemeijer (Deltares)
Hans Visser (Planbureau voor de Leefomgeving, PBL)
Dennis Walvoort (Wageningen Environmental Research, WENR)

Gefinancierd door het Ministerie van Landbouw, Natuur en Voedselkwaliteit (LNV),
in samenwerking met het Ministerie van Infrastructuur en Waterstaat (IenW).

CBS Den Haag
Henri Faasdreef 312
2492 JP Den Haag
Postbus 24500
2490 HA Den Haag
+31 70 337 38 00
www.cbs.nl

projectnummer 304689

2 september 2019

Samenvatting

Dit rapport is het resultaat van de Werkgroep Statistiek en geeft een overzicht van trendmethoden voor de Evaluatie Meststoffenwet (EMW). Het rapport richt zich specifiek op onderzoekers die met de trendmethoden werken. Daarnaast is er een werkwijzer rapport, waarin beschreven wordt hoe om te gaan met de trendmethoden bij de EMW.

Dit rapport beschrijft eerst de problemen die optraden rond de berekening van trends bij de meest recente EMW, te weten de diversiteit aan trends en hun presentatie, en de discrepanties tussen rapporten. Vervolgens gaan we in op de definitie van een trend en de significantie ervan. Naast de keuze voor trendmethode hebben de verschillende voor- en nabewerkingen die uitgevoerd worden op de data een belangrijke invloed hierbij. We behandelen daarom de bewerkingen die zoal mogelijk zijn, en geven aan waarop gelet moet worden. Daarna gaan we in op de verschillende trendmethoden en beschrijven hun voor- en nadelen.

De belangrijkste conclusies zijn:

- Verschillende trendmethoden kunnen leiden tot verschillende conclusies over het stijgen, dalen of gelijk blijven van een reeks. Een mogelijkheid is om altijd naar meerdere methoden te kijken bij het beoordelen van trends in een dataset.
- Een trend kent altijd onzekerheid, ook al levert de methode daar geen informatie over. Methoden die onzekerheidsmarges leveren kunnen verkeerde conclusies voorkomen.

Inhoudsopgave

1. Inleiding	5
1.1 Evaluatie Meststoffenwet	5
1.2 Werkgroep statistiek	6
2. Huidige werkwijze	7
2.1 Meetnetten	7
2.2 Trends in de EMW 2016	15
3. Afbakening	23
3.1 Doel van dit rapport	23
3.2 Doel van de trendmethoden	23
3.3 Data	23
4. Selectie van een trendmethode	25
4.1 Wat is een trend?	25
4.2 Keuze tussen methoden	30
4.3 Statistische vs. niet-statistisch methoden	31
5. Voor- en nabewerkingen	32
5.1 Aggregatie	32
5.2 Zomer- en wintergemiddelden	33
5.3 Transformaties	33
5.4 Corrigeren voor verstorende effecten	33
5.5 Veranderingen in meetnetten	34
6. Overzicht van methoden	35
6.1 Methodes	35
6.2 Eigenschappen van de methodes	35
6.3 Voortschrijdende gemiddelden	36
6.4 Mann-Kendall (MK) en Seasonal Mann-Kendall (SMK)	37
6.5 Theil-Sen	37
6.6 Lineaire regressie	37
6.7 LOWESS	38
6.8 Structurele tijdreeksmodellen (STM)	38
6.9 ARIMA-modellen	41
6.10 Twee deelperioden vergelijken	41
7. Vergelijking tussen methodes	42
8. Conclusies	44
Bronnen en achtergronddocumenten	46
Lijst van afkortingen	48

1. Inleiding

1.1 Evaluatie Meststoffenwet

De Meststoffenwet (MW) is de Nederlandse uitwerking van de Europese Nitraatrichtlijn met als belangrijkste doel de vervuiling van het water door de landbouw te verminderen. De veehouderij zorgt voor een overschot aan restanten van deze productieketen, met name mestgerelateerde nutriënten (fosfaat, stikstof (=nitraat) en ammoniak). Een gedeelte daarvan wordt over de eigen landbouwgrond verspreid, de rest wordt onder andere geëxporteerd. De Meststoffenwet reguleert de bemesting van landbouwgrond, opdat de belasting van de bodem met stikstof, fosfaat en verontreinigende stoffen niet leidt tot overschrijding van de doelen voor de kwaliteit van grond- en oppervlaktewater. In de MW is opgenomen dat de minister ten minste eens in de 5 jaar verslag uitbrengt aan de Tweede Kamer over de doeltreffendheid en effecten van deze wet in de praktijk. Met het rapport Evaluatie Meststoffenwet (EMW) van het Planbureau voor de Leefomgeving (PBL) wordt invulling gegeven aan deze verplichting. Deze evaluatie wordt gedaan in opdracht van de Ministeries van LNV en IenW, en uitgevoerd door verschillende onderzoeksinstituten, waaronder PBL, Deltares, RIVM, WUR, en CBS. Het PBL is verantwoordelijk voor het opstellen van de synthese van de evaluatie. Hiervoor dienen achtergrondrapporten van verschillende onderzoeksinstituten als basis. Het meest recente syntheserapport betreft de EMW 2016 (PBL, 2017). De voorbereidingen voor de evaluatie van 2020 zijn inmiddels begonnen.

In de Evaluatie Meststoffenwet wordt, terugkijkend (de ex post evaluatie), de werking van de Meststoffenwet tot nu toe beoordeeld en, vooruitkijkend (de ex ante evaluatie), een aantal beleidsopties voor toekomstige regelgeving bekeken. In de ex post evaluatie wordt gekeken of de doelen van het mestbeleid gehaald worden en in hoeverre dat door het gevoerde beleid wordt bereikt. Centraal daarbij staan de doelen voor het grondwater en het oppervlaktewater (nl. het terugdringen van eutrofiëring, i.e. een overmaat aan voedingsstoffen (nutriënten) waardoor ecologische processen ontregeld raken).

De ex post-evaluatie is vooral gebaseerd op metingen en landbouwstatistieken van bemesting, gewasopbrengsten, nutriëntenoverschotten en waterkwaliteit. Er wordt gekeken naar trends op basis van waarnemingen in bemesting en milieukwaliteit vanaf 2000, met een focus op de periode 2006-2014 bij de meest recente EMW. Aangezien er grote verbeteringen zijn geweest in de jaren negentig, is er voor gekozen alleen te kijken naar recentere gegevens.

De meetgegevens om de ontwikkelingen te analyseren zijn voornamelijk afkomstig van meetnetwerken, zoals het Landelijk Meetnet effecten Mestbeleid (LMM), Meetnet Nutriënten Landbouw Specifiek Oppervlaktewater (MNLSO), Landelijk Meetnet Grondwaterkwaliteit (LMG) en Provinciale Meetnetten Grondwaterkwaliteit (PMG). De meetnetten verschillen in omvang, doelstelling, en hoeveelheid historische gegevens die inmiddels beschikbaar zijn. In hoofdstuk 2 geven we een beschrijving van deze meetnetten.

In de EMW worden uit de gegevens van de meetnetten conclusies getrokken of bepaalde indicatoren stijgend, dalend of ongeveer gelijk gebleven zijn. Het gaat dan niet om de waarden in opvolgende jaren, maar ook om de meer structurele, onderliggende ontwikkeling in de tijd. Deze ontwikkeling noemen we de trendmatige ontwikkeling. Er wordt gekeken naar de ontwikkeling van de stikstof- en fosfaatoverschotten (milieudruk) en van de kwaliteit van het ondiepe grondwater en het regionale oppervlaktewater. De focus bij de EMW ligt op het oppervlaktewater.

1.2 Werkgroep statistiek

Een van de onderliggende rapporten die als basis heeft gediend voor de synthese van de Evaluatie Meststoffenwet (PBL, 2017) is Velthof et al. (2017). Hierin wordt een zestigtal vragen beantwoord die de toenmalig ministeries van Economische Zaken (EZ) en Infrastructuur en Milieu (IenM) gesteld hebben in het kader van de EMW 2016, en die aspecten betreffen van het mestbeleid en de gevolgen daarvan op landbouw en milieu. Hiervoor is gebruik gemaakt van data van o.a. CBS en verschillende landelijke meetnetwerken voor waterkwaliteit en landbouw. In de aanbevelingen voor volgende evaluaties wordt een aantal aandachtspunten benoemd. Meerdere van deze aandachtspunten richten zich op de data en gebruikte methoden. Zo bleek bij de beantwoording van de vragen dat er soms op basis van dezelfde of vergelijkbare data andere conclusies getrokken werden. In Hoofdstuk 2 zullen we in meer detail ingaan op de situaties waarin verschillen tussen rapportages optraden.

Om dergelijke verschillen in de toekomst te voorkomen bij de EMW heeft het Ministerie van LNV de wens geuit om een werkgroep statistiek zich te laten buigen over het bereiken van meer consistentie in trendmethodes voor de volgende EMW. Naast het CBS hebben vijf instituten deelgenomen aan de activiteiten van de werkgroep. Het gaat om: Deltares, PBL, RIVM, WUR/WEER en WUR/WENR. Het doel van de werkgroep was om in een vroeg stadium overeen te komen welke statistische methode er wordt ingezet bij een volgende Evaluatie Meststoffenwet. In de opdracht aan de werkgroep wordt het doel beschreven als *een eenduidige, objectieve werkwijze opstellen waarmee trends en de significantie van trends vastgesteld kan worden van meetdata/statistieken die relevant zijn voor Evaluatie Meststoffenwet ex-post rapportages*.

Om tot zo'n werkwijze te komen is eerst een methoderapport opgesteld, waarin de verschillen tussen de methodieken beschreven worden. Dit betreft het huidige rapport, waarin de door de deelnemende instituten gehanteerde methodieken waarmee de trends en de significantie van trends vastgesteld worden op een rij worden gezet. Op basis hiervan heeft de werkgroep een discussie gevoerd over voorkeursmethoden. De deelnemende instituten hebben elementen benoemd die onderdeel dienen te zijn van de eenduidige, objectieve werkwijze voor de vaststelling van trends en de significantie van trends. Deze elementen hebben hun plek gekregen in een door de werkgroep statistiek opgesteld werkwijzerapport, waarin de te gebruiken werkwijze tijdens de komende Evaluatie Meststoffenwet aan bod komt.

2. Huidige werkwijze

In dit hoofdstuk gaan we nader in op de werkwijze bij EMW 2016, in het bijzonder op de problemen die ervaren werden en die aanleiding waren voor het oprichten van de werkgroep statistiek. Aangezien de gegevens uit een aantal meetnetten de basis vormen van de EMW, geven we eerst een beschrijving van de verschillende meetnetten. Elk meetnet is met een bepaald doel voor ogen opgezet en kent zijn eigen werkwijze. Deze werkwijze beschrijven we na de meetnetten zelf.

2.1 Meetnetten

De meetnetten die gebruikt worden voor de EMW zijn: het Landelijk Meetnet effecten Mestbeleid (LMM), Meetnet Nutriënten Landbouw Specifiek Oppervlaktewater (MNLISO), Landelijk Meetnet Grondwaterkwaliteit (LMG) en Provinciale Meetnetten Grondwaterkwaliteit (PMG). De meetnetten kennen verschillende beheerders: Het LMM (www.rivm.nl/lmm) wordt beheerd door RIVM/WEER, het MNLISO (www.deltares.nl/nl/projecten/meetnet-nutriënten-landbouw-specifiek-oppeervlaktewater-mnliso/) door Deltares, de LMG (www.rivm.nl/lmg) door het RIVM, en de PMG door de provincies.

2.1.1 Landelijk Meetnet effecten Mestbeleid (LMM)

Het LMM bestaat sinds 1987 en volgt de kwaliteit van grond- en oppervlaktewater op landbouwbedrijven, gerelateerd aan de bedrijfsvoering op deze bedrijven. Tot 2006 bestond het LMM uit steeds wisselende meetlocaties. Door de inrichting van het Derogatiemeetnet en de bepalingen van de EU die daarvoor gelden, worden sinds 2006 jaarlijks steeds dezelfde bedrijven bemonsterd en vindt er alleen wisseling plaats als bedrijven niet meer deel willen nemen.

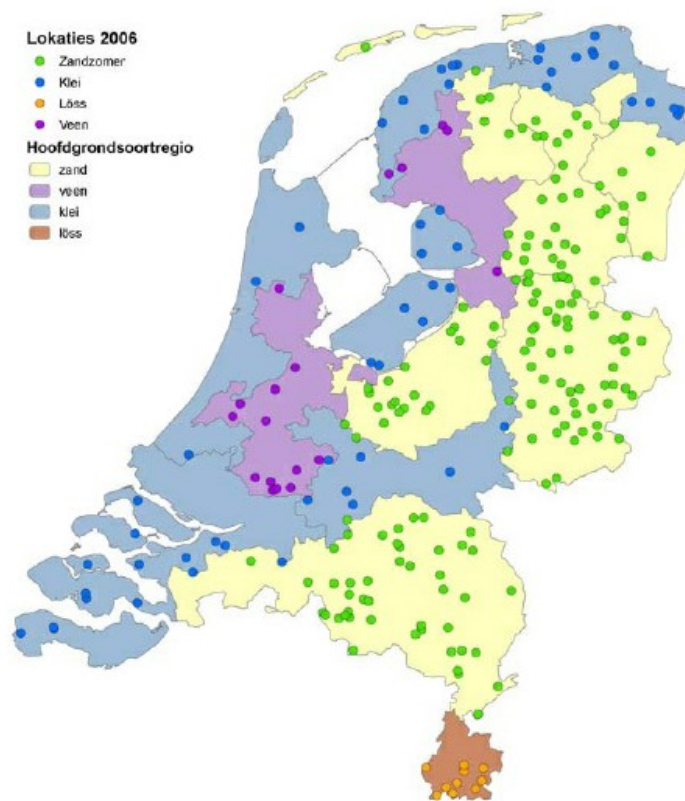
Bij de analyse van de gegevens van het LMM moet rekening gehouden worden met de complexe opbouw en ontwikkeling van het meetnet, alsook de wijzingen die zijn doorgevoerd om aanvullende beleidsvragen te kunnen beantwoorden. Daarnaast hebben natuurlijke omstandigheden (weer- en seizoenseffecten, verdeling bodemtypen en grondwatertrappen) een grote invloed op de gemeten concentratie.

Voor de EMW worden de nitraatconcentraties uit het LMM gerapporteerd. Het detailniveau verschilt per rapport (groter in onderliggende rapporten dan in het syntheserapport).

Opbouw van het LMM

Twee meetnetten

Het LMM omvat twee meetnetten, een basismmeetnet en een derogatiemeetnet. Het basismmeetnet beoogt de ontwikkeling in de waterkwaliteit op landbouwbedrijven in Nederland te volgen in relatie tot de landbouwpraktijk. Het derogatiemeetnet beperkt zich tot bedrijven die zich hebben aangemeld voor derogatie. Bedrijven uit het basismmeetnet met derogatie zijn ook opgenomen in het derogatiemeetnet. De verplichte omvang van het derogatiemeetnet (300 bedrijven) en het type bedrijven dat van derogatie gebruik maakt is zodanig dat extra bedrijven geworven zijn voor het derogatiemeetnet. Daarnaast zijn er regelmatig verkennende programma's geweest. In **Figuur 1** is een kaart met de meetpunten van het LMM weergegeven. Dit is de stand van 2006, en geeft dus een enigszins verouderd beeld. Inmiddels is de waarneming in met name de Lössregio en de regio Zand Zuid geïntensiveerd (zie Tabellen 2 en 3 verderop), en is de gebiedsindeling veranderd. Desalniettemin geeft dit een beeld van de spreiding van meetpunten binnen het LMM.



Figuur 1: De vier hoofdgrondsoortregio's waarvoor deelprogramma's van het LMM zijn opgezet en de ligging van de voor de jaarrapportage 2006 gemonitorde bedrijven waarvan zowel waterkwaliteitsgegevens als gegevens over bedrijfsvoering beschikbaar zijn.

Steekproef

Beide meetnetten zijn gebaseerd op gestratificeerde aselechte steekproeven welke in principe getrokken zijn uit het Bedrijfsinformatienet (BIN) van Wageningen Economic Research. Indien in het BIN onvoldoende bedrijven aanwezig waren, is rechtstreeks uit de landbouwtellingbedrijven getrokken. De bedrijven zijn vervolgens wel aanvullend opgenomen in het BIN, maar als aparte groep.

Steekproefstrata

Het LMM onderscheidt vier hoofdgrondsoortregio's in Nederland, de Zand-, Löss-, Klei- en Veenregio (zie **Figuur 2**). Binnen elke regio zijn één of meerdere gebieden onderscheiden. Daarnaast wordt bij de trekking voor het basismetnet onderscheid gemaakt tussen 4 bedrijfscategorieën (zie **Tabel 1**): melkveehouderij, akkerbouw, hokdierbedrijven en overige dierbedrijven. In elke regio worden alleen die categorieën meegenomen die een relevant areaal vertegenwoordigen. In elke regio is minstens 80% van het landbouwareaal vertegenwoordigd. Bij de trekking wordt verder rekening gehouden met een evenredige verdeling over bedrijfseconomische grootteklassen.



Figuur 2: Regio's met grondtype in het LMM.

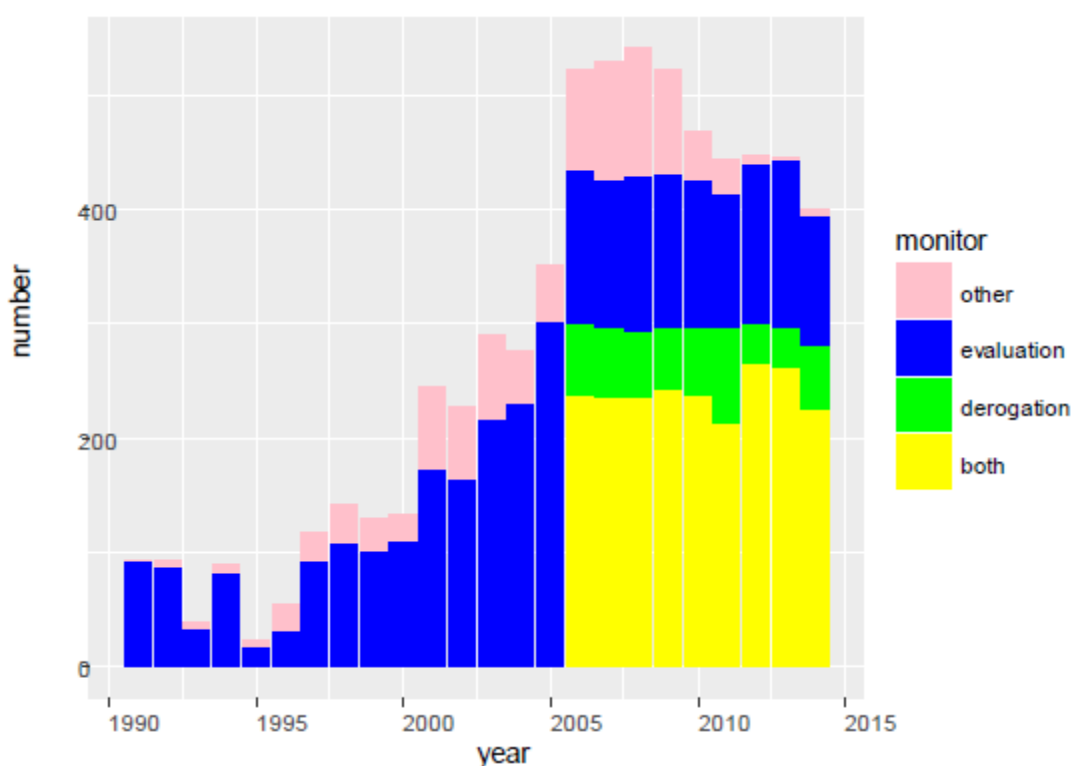
Tabel 1: Indeling basismetnet

Regio	Zandregio	Lössregio	Kleiregio	Veenregio
Melkveehouderij	3 gebieden onderscheiden – elk gebied hetzelfde aantal	1 gebied	5 gebieden onderscheiden - areaalevenredig	2 gebieden onderscheiden – elk gebied zelfde aantal
Akkerbouw	3 gebieden onderscheiden – Noord en Midden samen hetzelfde aantal als Zuid. Verdeling N-M areaal evenredig	1 gebied	5 gebieden onderscheiden - areaalevenredig	niet relevant
Hokdier	1 gebied	niet relevant	niet relevant	niet relevant
Overig dier	1 gebied	1 gebied	5 gebieden onderscheiden - areaalevenredig	niet relevant

Aantal bedrijven

Het aantal bedrijven varieert in de tijd vanwege de genoemde opbouw van het programma (zie **Figuur 3**). De som van het aantal bedrijven per meetnet (zie **Tabel 2**) is hoger dan het aantal dat daadwerkelijk wordt bemonsterd (zie **Tabel 3**). Dit komt doordat bedrijven uit het basismeetnet die gebruik maken van derogatie ook mogen meetellen voor het derogatiemeetnet. Er is dus sprake van overlap. Ook de bedrijven die deelnemen aan het verkennende programma zijn tevens deelnemer aan het derogatiemeetnet.

Per bedrijf wordt op 16 punten uitspoeling gemeten, dit kan zijn door bemonstering van drainwater, grondwater en/of bodemvocht. Daarnaast wordt er, indien van toepassing, op maximaal 8 punten metingen in het slootwater gedaan op een bedrijf. Per meetpunt vindt een beperkt aantal (individuele) metingen plaats. Per bedrijf per ronde worden één of meerdere mengmonsters gemaakt. Deze mengmonsters worden op een groot aantal parameters geanalyseerd.



Figuur 3: Ontwikkeling in het aantal bedrijven in het LMM per programma.

Tabel 2: Aantal geplande bedrijven per deelprogramma in de periode 2011-2014

	Evaluation Monitoring	Derogation Monitoring	Exploratory Monitoring
	EM	DM	VM
Clay	60	60	4
Loess	50	20	1
Peat	24	60	3
Sand	109	160	8
Total	243	300	16
	EM	DM	VM
Dairy farms	109	261	16
Non-dairy farms	134	39	0
Total	243	300	16

Tabel 3: Aantal bedrijven per regio en deelprogramma

	Only EM	Only DM	Both EM and DM	Other	Total
2010*	129	60	238	41	468
Clay	40	13	62	7	122
Loess	22	3	17	9	51
Peat	7	8	44	0	59
Sand	60	36	115	25	236
2011	118	84	220	30	452
Clay	40	15	56	6	117
Loess	25	1	18	6	50
Peat	5	23	31	0	59
Sand	48	45	115	18	226
2012	143	34	283	6	466
Clay	45	3	74	0	122
Loess	31	1	19	0	51
Peat	6	8	51	0	65
Sand	61	22	139	6	228
2013	147	35	263	1	446
Clay	46	3	57	0	106
Loess	31	1	20	0	52
Peat	5	8	51	0	64
Sand	65	23	135	1	224

* The years are the FADN years (see Sections 2.1.4 and 4.3.1). The groundwater sampling was performed between one-half and one-and-a-half years later, depending on the region.

Meetfrequentie

De meetfrequentie was voor 2006 afhankelijk van de regio. In de Zand- en Veenregio werd een bedrijf drie keer in een periode van zeven jaar (de deelname periode) bemonsterd. In de Kleiregio werden de bedrijven wel jaarlijks bemonsterd in de periode van deelname. Vanaf 2006 worden alle LMM-bedrijven jaarlijks bemonsterd. Het aantal bemonsteringen per jaar varieert van 1 tot 7; dit is afhankelijk van het watertype dat wordt bemonsterd (grondwater, drainwater, bodemvocht, slootwater). Drain- en slootwater wordt in de regel vier keer bemonsterd gedurende de winter. Daarnaast worden er nog drie slootwaterbemonsteringen in de zomer uitgevoerd. De grondwater- en bodemvochtbemonstering wordt één keer per jaar uitgevoerd, met uitzondering van de bedrijven in de Kleiregio en de natte delen van de Zandregio, waar twee keer per jaar wordt bemonsterd.

Gegevensanalyse

Steekproefveranderingen

Er zijn meerdere soorten van veranderingen die relevant zijn:

1. Wijzigingen op een deelnemend bedrijf (verandering van categorie, aankoop/verkoop of huur/verhuur grond);
2. Wijzigingen in de groep van bedrijven (stoppers, nieuw geworven bedrijven);
3. Aanpassing van strata n.a.v. nieuwe beleidsvragen, bijv. onderscheid maken bij akkerbouw in de Zandregio tussen Zand Zuid enerzijds en Zand Noord en Zand midden anderzijds in 2012. Hierdoor zijn sinds dat jaar de akkerbouwbedrijven in Zand Zuid oververtegenwoordigd in de steekproef.

Grondsoort en grondwatertrap

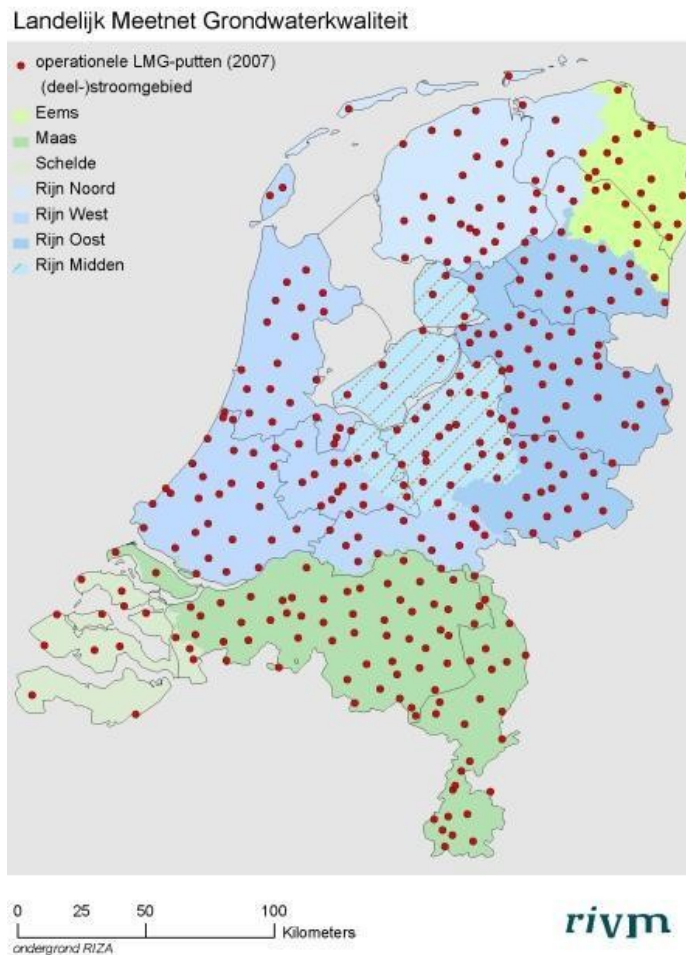
De reden dat steekproefveranderingen een grote rol spelen, is o.a. de grote invloed van grondsoort en grondwatertrap op de gemeten concentratie, zeker bij nitraat. Ook binnen een grondsoortregio, dus bijvoorbeeld de Zandregio, komen lokaal andere grondsoorten (klei, veen, löss) voor. Een bedrijf dat een zandig perceel met een diepe grondwaterstand verkoopt en een veenhoudend of nat perceel koopt, zal de gemiddelde nitraatconcentratie zien dalen, ook als er in de landbouwpraktijk zelf niets veranderd.

Weers- en seizoensinvloed

Het weer, vooral de hoeveelheid neerslag en verdamping, heeft invloed op de gemeten concentratie. Daarnaast is er ook nog een invloed van het moment van meten in een jaar op de hoogte van de gemeten concentratie.

2.1.2 Landelijk Meetnet Grondwaterkwaliteit (LMG)

In dit meetnet wordt de kwaliteit van het ondiep en middeldiep grondwater in Nederland vastgesteld (zie o.a. Fraters et al., 2016 en www.rivm.nl/lmg). Het meetnet is opgebouwd tussen 1979 en 1984 en bestaat uit ongeveer 350 meetlocaties die zijn verspreid over heel Nederland (zie **Figuur 4**).



Figuur 4: meetpunten LMG

Het gaat om permanente putten die speciaal voor monitoringdoeleinden zijn aangelegd. De belangrijkste criteria voor de selectie van de locaties waren de grondsoort, het landgebruik en de hydrogeologische toestand. Op elke locatie worden grondwatermonsters genomen op diepten van 5-15 m

en 15-30 m onder het grondoppervlak. Er wordt onderscheid gemaakt naar typen landgebruik en grondsoorten. In **Tabel 4** is het aantal putten per categorie weergegeven.

Tabel 4: Aantal putten waarvoor complete gegevensreeksen beschikbaar zijn voor de periode 1984-2014 voor alle grondsoorten, typen landgebruik en monsternemingsdiepten

Landgebruik	Diepte (m)	Zand	Klei	Veen	Overig
Landbouw	5-15	121	61	32	5
	15-30	120	60	32	4
Natuur	5-15	55	4	4	3
	15-30	52	4	4	2
Overig	5-15	37	19	2	8
	15-30	37	18	2	12

Tussen 1984 en 1998 zijn er jaarlijks steekproeven genomen op de locaties. Na een evaluatie in 1998 is het aantal putten verminderd en werd de bemonsteringsfrequentie gereduceerd voor bepaalde combinaties van grondsoorten en diepten. Zo is het aantal putten waaruit jaarlijks een monster wordt genomen teruggebracht van 756 naar ongeveer 350. De meetfrequentie is omlaag gegaan voor put-filters waar weinig veranderingen werden verwacht en gezien. De ondiepe filters in de zandgebieden worden 1 maal per jaar bemonsterd; de ondiepe filters in de overige gebieden eens per 2 jaar; alle diepe filters 1 maal per 4 jaar; ondiepe filters die tot nu toe alle jaren een chlorideconcentratie groter dan 1000 g/m³ te zien geven worden 1 maal per 4 jaar bemonsterd, ongeacht het bodemtype. Het Rijksinstituut voor Volksgezondheid en Milieu beheert het netwerk en is verantwoordelijk voor de interpretatie van en rapportage over de gegevens.

Gegevensanalyse

Vanwege de opzet van het LMG zijn er locaties die niet jaarlijks worden bemonsterd. Om onjuiste trends die voortvloeien uit de opzet van het meetnet te vermijden, wordt er een schatting gemaakt van alle ontbrekende gegevens. Deze schatting wordt berekend door de beschikbare gegevens te interpoleren. Voor gegevens die ontbreken aan het begin of het eind van een reeks wordt respectievelijk de eerste of de laatste beschikbare waarde gebruikt om de ontbrekende informatie te schatten. De jaarlijkse gemiddelde concentraties worden simpelweg berekend aan de hand van de gemeten concentraties. De indeling in grondsoortregio's en gebieden is hetzelfde als die bij het LMM (Zie **Figuur 2**). Voor de LMG-putten in de Zandregio is afzonderlijk het gemiddelde bepaald voor de gebieden Zand noord, Zand midden en Zand zuid.

2.1.3 Provinciale Meetnetten Grondwaterkwaliteit (PMG).

Samen met de provincies zijn na de inrichting van het LMG ook provinciale meetnetten grondwaterkwaliteit (PMG) ingericht, zie IenM (2013). De PMG's komen wat betreft doelstelling, inrichting en gebruik overeen met het LMG, maar hebben vaak ook nog aanvullende doelstellingen, zoals de bewaking van de grondwaterkwaliteit in bijzondere gebieden.

Alle provinciale meetnetten zijn in de negentiger jaren geoptimaliseerd. De ligging van de meetpunten in de verschillende provinciale meetnetten is goed afgestemd op het landelijk meetnet. Maar voor de keuze in bemonsteringsdiepte, analysepakket, laboratorium, en indeling in homogene deelgebieden is in elke provincie net weer een andere opzet gekozen. Gemiddeld komt elke provincie ongeveer op een gelijke dekking uit van 2 tot 3 peilbuizen per 100 km². De Provinciale Meetnetten (PMG) bevatten bij elkaar zo'n 600 meetpunten (IenM, 2013).

2.1.4 Meetnet Nutriënten Landbouw Specifiek Oppervlaktewater (MNLISO)

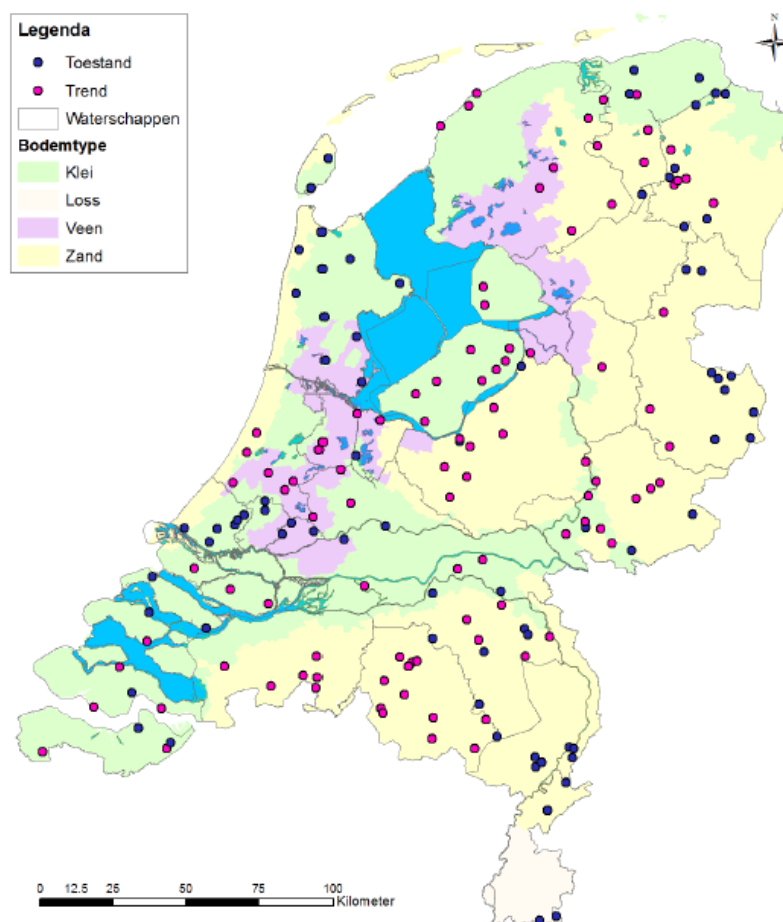
Het MNLISO (zie UVW, 2015) is in 2010-2012 door de waterschappen en Deltares opgezet om te onderzoeken hoe het staat met de nutriënten (meststoffen) in landbouw specifiek oppervlaktewater. Voor het meetnet zijn bestaande meetlocaties van alle waterschappen geselecteerd, die landbouw als enige humane bron van nutriënten hebben. Met de gegevens uit het meetnet zijn door Deltares toestand- en trendanalyses uitgevoerd om te kunnen vaststellen of:

- Er neerwaartse of opwaartse trends in nutriëntenconcentraties zijn;
- De doelen met betrekking tot nutriënten worden gehaald (toestand).

Het MNLISO bestaat uit 173 meetlocaties die landbouw als enige bron van nutriënten hebben (zie **Figuur 5**). De waterschappen en Deltares hebben daarvoor gezocht naar bestaande meetlocaties in landbouwgebieden die niet onder invloed staan van lozingen vanuit rioolwaterzuiveringen of industrie. Van de 173 meetlocaties hebben 99 momenteel een meetreeks langer dan 10 jaar, waardoor het mogelijk is naar trends in de nutriëntenconcentraties te kijken.

Gegevensanalyse

De trends in de MNLISO-gegevens zijn met meerdere statistische methoden vastgesteld, ook voor de zomer- en winterconcentraties afzonderlijk, voor de deelgebieden zand, klei en veen en voor verschillende meetperiodes. De conclusies (nl. dat de nutriëntenconcentraties dalen) zijn dus niet afhankelijk van de gekozen statistische methode, meetperiode of deelgebied.



Figuur 5: Toestand- en trendmeetlocaties behorend tot het MNLISO geprojecteerd op de kaart van de bodemtypes zoals gehanteerd in het LMM. De trendmeetlocaties zijn onderdeel van de toestandsmetlocaties.

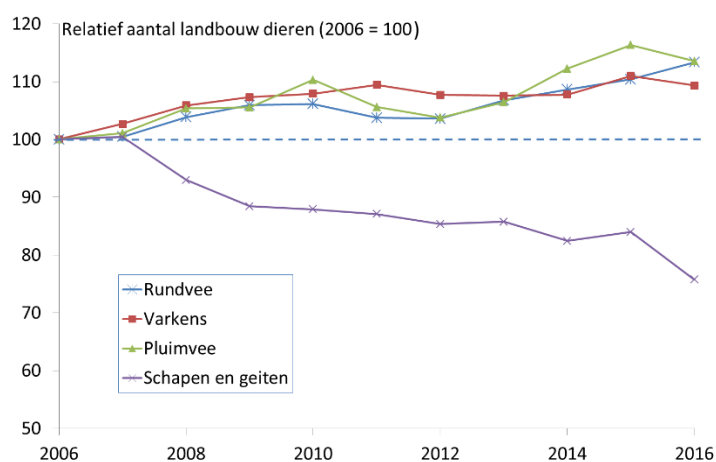
2.2 Trends in de EMW 2016

In deze paragraaf geven we een overzicht van trendanalyses in de EMW 2016. Hiermee willen we twee problemen laten zien: de diversiteit aan trends en discrepanties tussen verschillende analyses.

In 2.2.1 geven we een aantal voorbeelden van de verschillende soorten data, methoden, figuren en tabellen waarmee over trends gerapporteerd wordt in het ex-postrapport (Velthof et al., 2017) en het syntheserapport (PBL, 2017) van de EMW, en geven aan welke bijbehorende conclusies er in deze rapporten getrokken worden. Dit om een beeld te krijgen van de diversiteit aan trends die onderzocht worden en de manier waarop deze weergegeven worden. In Velthof et al. (2017) zijn de figuren direct overgenomen uit de onderliggende rapporten. Elk is met een eigen doel gemaakt en heeft een eigen layout. Zodra ze samengebracht worden in één rapport vallen de verschillen op. In PBL (2017) is gepoogd zoveel mogelijk te uniformeren. Dit leidt soms echter weer tot discrepanties met de onderliggende rapporten. In 2.2.2. illustreren we dit.

2.2.1 Trends in de EMW

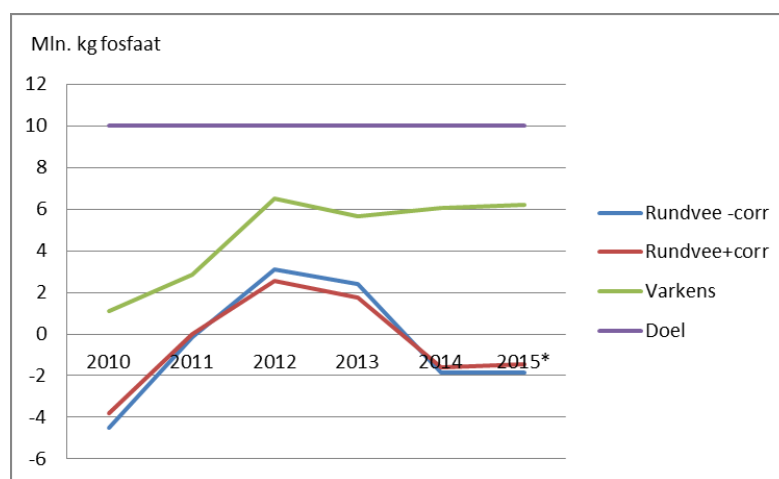
Om reeksen te kunnen vergelijken worden soms alle reeksen gepresenteerd in de vorm van een ratio. Door alle waarden van de reeks te delen door de waarde in een basisjaar starten alle reeksen in het basisjaar op 100. De keuze van het basisjaar beïnvloedt daarom het beeld. Daarnaast heeft een stijging in de ene reeks minder consequenties dan in een andere reeks. Dat wordt echter in een dergelijke presentatie niet zichtbaar, omdat de onderliggende aantallen niet zichtbaar zijn. Een voorbeeld is **Figuur 6** (Figuur 1 in Velthof et al, 2017). Deze geeft aan dat de laatste jaren het aantal dieren stijgt voor pluimvee (12% hoger in 2015 dan in 2012), melkvee (7% hoger in 2015 dan in 2012) en varkens (3% hoger in 2015 dan in 2012). Het aantal schapen en geiten is iets afgenomen in de periode 2012-2015 (2%). Volgens voorlopige cijfers in 2016 stijgt het aantal runderen verder en dalen de aantallen varkens en pluimvee iets. Om deze reeksen te kunnen vergelijken worden alle reeksen op 100 gezet in het startjaar 2006. Aangezien het om integrale tellingen gaat wordt aangenomen dat de cijfers geen marge hebben. Dit is een voorbeeld waarvoor opeenvolgende jaarwaarden verbonden worden met lijnstukken maar waar geen trend geschat wordt. Dat zou mogelijk zijn als alle lijnstukken in de reeks dalen (zoals voor schapen en geiten tot 2014 het geval is), maar als lijnstukken soms dalen en soms stijgen, kun je geen eenduidige uitspraak doen. Dit is bijvoorbeeld het geval voor de reeks Varkens en Pluimvee.



Figuur 6: Relatief aantal landbouwdieren (2006=100). Cijfers voor 2016 zijn voorlopig. Bron: CBS-Statline.

Uit **Figuur 7** (Figuur 3 in Velthof et al, 2017) blijkt, dat de varkenshouderij in 2012 60% van het voerspoordoel (dus 6 miljoen kg fosfaat) heeft gerealiseerd en dat sindsdien dit niveau is gestabiliseerd. De rundveehouderij had in 2012 zo'n 30% van de 10 miljoen kg fosfaat gerealiseerd, maar vervolgens nam de beoogde reductie weer af en was in 2014 en 2015 de fosfaatuitscheiding via het voer zelfs hoger dan in het uitgangsjaar. De doelstelling van 10 miljoen kg reductie in fosfaatproductie voor de melkveesector is dus niet gerealiseerd, mede doordat het convenant zich beperkte tot het fosfaatgehalte in krachtvoerders.

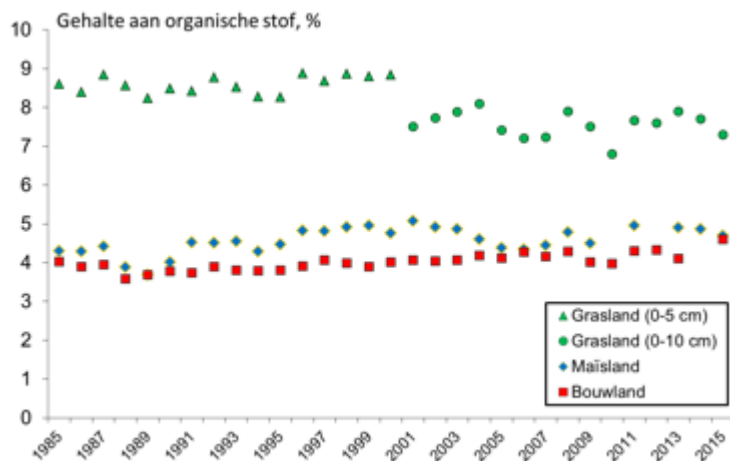
In deze figuur zijn de cijfers niet als ratio gepresenteerd maar als een absoluut verschil, namelijk de waarde van 2009 minus de betreffende waarde (het omgekeerde zou meer voor de hand liggen). Net als in de vorige figuur is dit een voorbeeld waar geen trend wordt weergegeven. Een conclusie trekken over de trendmatige ontwikkeling in bijvoorbeeld de twee reeksen over rundvee is erg lastig: de reeksen stijgen, dalen en stabiliseren. Ook wordt hier geen informatie over de onzekerheid gegeven.



*) 2015 is voorlopig

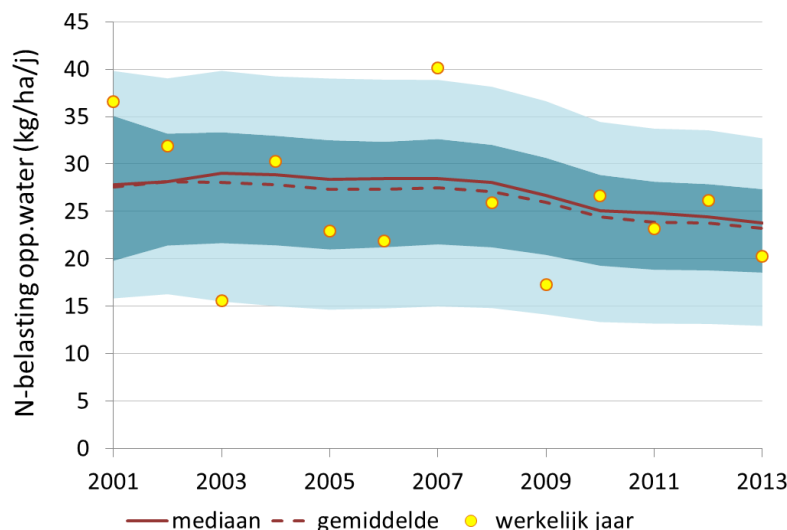
Figuur 7: Voortgang van de realisatie voerspoor in kader van convenant tussen Nevedi en LTO in de rundvee- en varkenshouderij voor de jaren 2010-2015. Op de Y-as is de afname van de fosfaatproductie weergegeven ten opzichte van de productie in het referentiejaar 2009. Een negatief getal betekent dat de fosfaatproductie is toegenomen ten opzichte van 2009. Rundvee+corr: correctie naar melkproductie van 2009 (De Koeijer et al., 2016).

Figuur 8 (Figuur 16 in Velthof et al, 2017) toont dat het gehalte (de y-as geeft percentages weer) aan organische stof in bodemmonsters een stabiele tot licht stijgende trend laat zien voor grasland, bouwland (akkerbouw) en maïsland. Ook hier wordt geen trend getoond, en wordt een conclusie getrokken op basis van een visuele inspectie van de ontwikkeling. Aangezien dit om meetgegevens gaat mogen we aannemen dat hier wel enige onzekerheid is ten gevolge van meetfouten. Bij het trekken van een conclusie over de ontwikkeling wordt hier echter geen rekening mee gehouden.



Figuur 8: Gemiddelde trends van het gehalte aan organische stof in bodemmonsters van grasland (0- 5 cm in periode 1995 -2001 en 0-10 cm in periode 2002-2015), maïsland (0-25 cm) en akkerbouw (0-20 cm) in de periode 1985 -2015. Bron: Eurofins Agro.

Figuur 9 (Figuur 41 in Velthof et al, 2017) laat de dalende trends in concentraties van stikstof (N-totaal) zien voor de oppervlaktewater sinds 2001. De y-as is hier in kg/ha, dus weer anders dan bij de vorige grafieken. De stikstofbelasting is voor weer gecorrigeerd en laat een dalende trend zien. Interessant is hier dat voor jaarwaarden zowel mediaan als rekenkundig gemiddelde wordt getoond. De gele punten geven berekende nitraatconcentraties aan bij het werkelijke weerjaar. De donkerblauwe geven de range tussen het 25- en 75-percentiel waarde en de lichtblauwe band tussen de 10 en 90-percentielwaarde. Uit de figuur wordt niet duidelijk waarvan de percentielen genomen zijn en waarom de gele stippen buiten de gearceerde banen liggen.

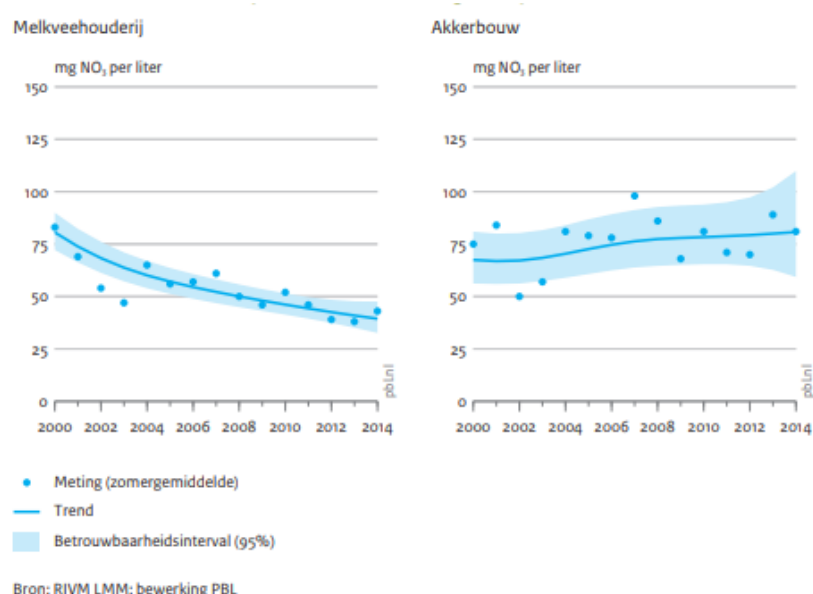


Figuur 9. Verloop van gemiddelde stikstofbelasting van het oppervlaktewater berekend met het model STONE, uitgaande van het werkelijke weersgegevens per jaar en een methode waarbij gecorrigeerd is voor weers effecten.

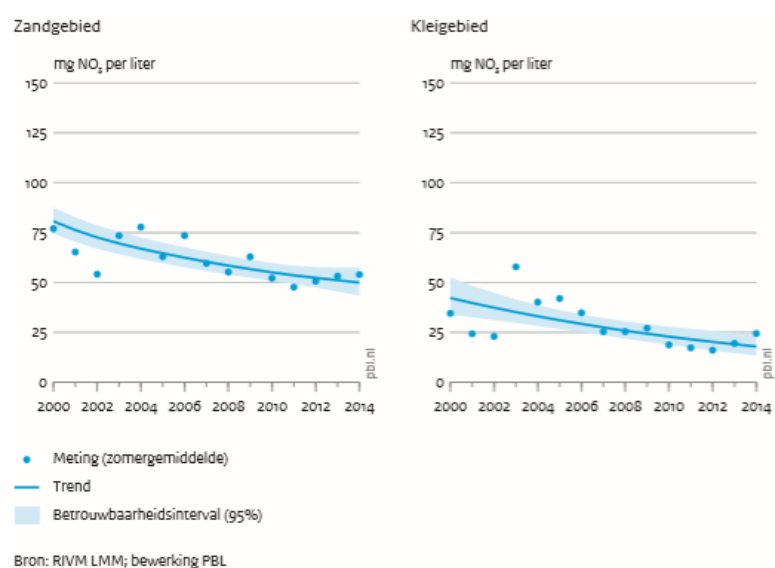
In het syntheserapport (PBL, 2017) is wat meer uniformiteit in de weergave van trends. Zo komen we onder andere de volgende trendanalyses tegen.

In **Figuur 10a en Figuur 10b** (Figuren 4.4 en 4.5 in PBL, 2017) worden de nitraatconcentraties getoond (op de y-as). Er wordt o.b.v. o.a. deze figuur geconcludeerd dat in de periode 2000-2015 de nitraatconcentraties in het zandgebied significant en relevant dalen, behalve onder akkerbouw en in het centrale zandgebied. De dalingen in het zandgebied stagneren en zijn na 2005 niet meer significant in de deelgebieden. In het kleigebied is de daling relevant en significant van 2006 tot en met 2014. Met 'relevantie' van de ontwikkeling wordt bedoeld op de betekenis die de daling of stijging heeft. Soms is een ontwikkeling statistisch niet significant, maar heeft deze wel een inhoudelijke betekenis. Dit staat ook wel bekend als 'practical significance'.

Dit is de eerste figuur waar een statistische trendmethode gebruikt is, in alle voorgaande figuren nog niet. De trendlijn is berekend met TrendSpotter, waarbij als model een Integrated Random Walk is toegepast. Hiermee is het ook mogelijk om de significanties van een verschil [trendwaarde van 2014 – trendwaarde van 2000] te bepalen.



Figuur 10a: Nitraatconcentratie in uitspoelend water in zandgebied per sector



Figuur 10b: Nitraatconcentratie in uitspoelend water per grondsoortgebied

2.2.2 Discrepanties tussen rapporten

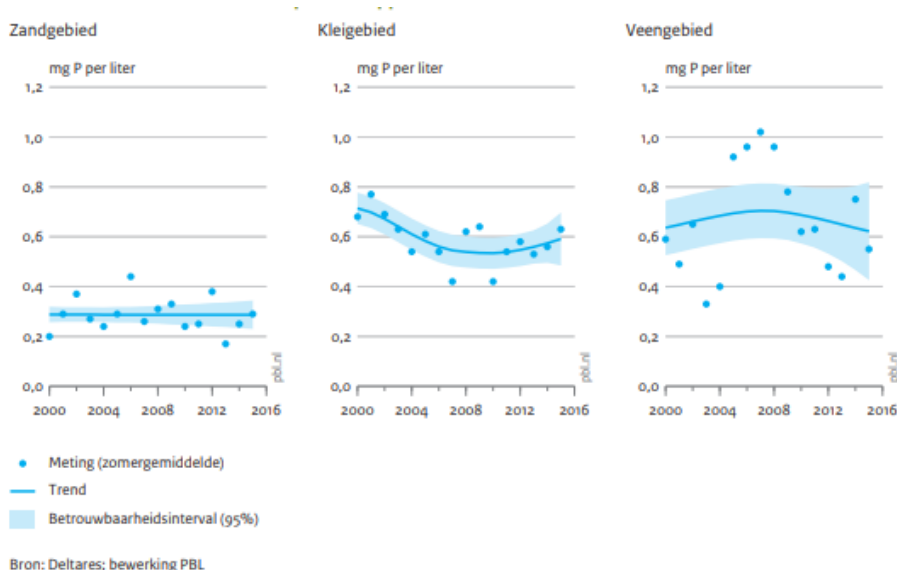
De werkgroep statistiek is opgericht om discrepanties tussen verschillende rapportages te voorkomen. Dit betrof in het bijzonder verschillen in trends die gevonden werden in de verschillende databronnen (o.a. waterkwaliteit en landbouwgegevens LMM, waterkwaliteit MNLSO, gewasopbrengsten en bodemvruchtbaarheid). Dit resulteerde er bijvoorbeeld in dat een instituut op basis van een dataset concludeerde dat de waterkwaliteit verbeterd was, terwijl een ander instituut, op basis van (nagenoeg) dezelfde data, concludeert dat het stabiel was gebleven of zelfs verslechterd was. Een belangrijke reden hiervoor ligt in de verschillende methoden die voor de trendanalyses gehanteerd worden, waaronder de selectie en bewerking van de data en de gehanteerde trendmethode. Die verschillen zijn deels ingegeven door de verschillen in doelen tussen rapportages.

In het algemeen zijn de volgende verschillen opgetreden:

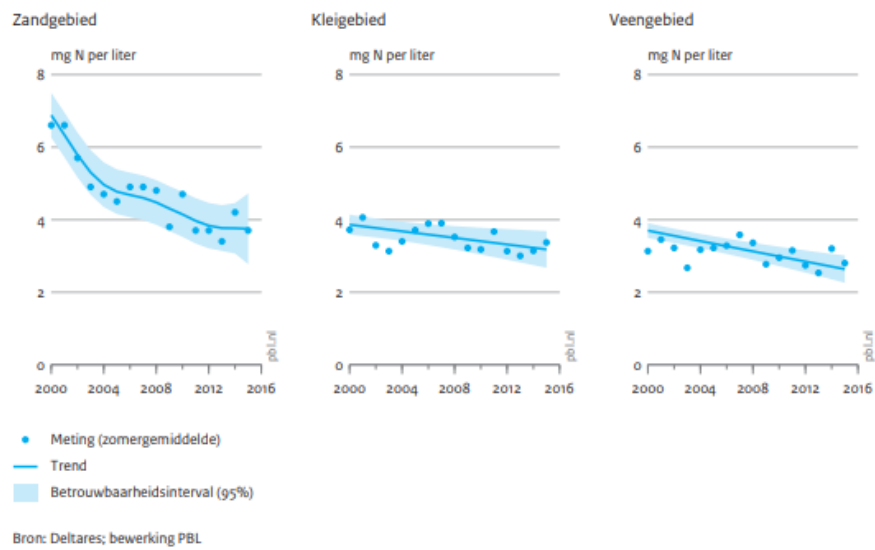
- Twee rapporten gebruiken ongeveer dezelfde variabelen en komen toch tot andere conclusies.
- Twee rapporten zijn op dezelfde brondata gebaseerd, maar in beide gevallen is een andere bewerking gedaan (zoals aggregatie), waardoor de conclusies toch verschillen.
- Bij een van de onderliggende rapporten die als basis voor het syntheserapport door PBL gebruikt is, is de data verder bewerkt (er is een nieuwe trend bepaald), waardoor de uitkomsten niet in lijn zijn met elkaar.

Hieronder beschrijven we van dit laatste geval een voorbeeld waarin een dergelijk verschil tussen rapportages optrad.

Figuur 11 en Figuur 12 (Figuren 4.10 en 4.11 in PBL, 2017) richten zich op de stikstof- en fosforconcentratie in landbouwspecifiek oppervlaktewater. Er wordt hier uitgegaan van zomergemiddeldes van de metingen. In het zandgebied dalen de concentraties stikstof al jaren, maar vlakt de trend in de laatste jaren af. Deze figuur is gebaseerd op data uit het MNLSO. Hierbij heeft Deltares de gemiddelde concentraties per grondsoort aangeleverd (de blauwe stippen), en heeft PBL met TrendSpotter een analyse gedaan om de trend en marge daarop te berekenen (met een Integrated Random Walk (IRW) model).

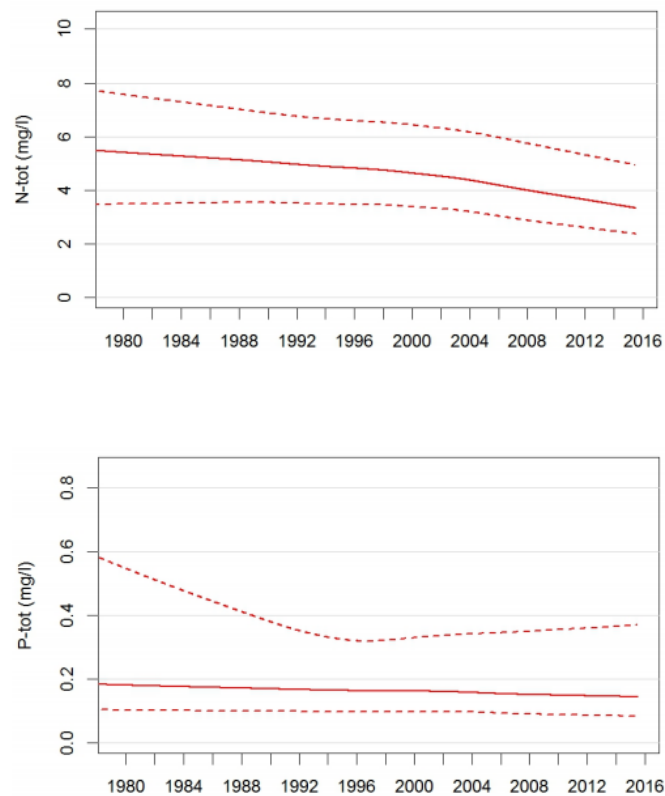


Figuur 11: Fosforconcentratie in landbouwspecifiek oppervlaktewater



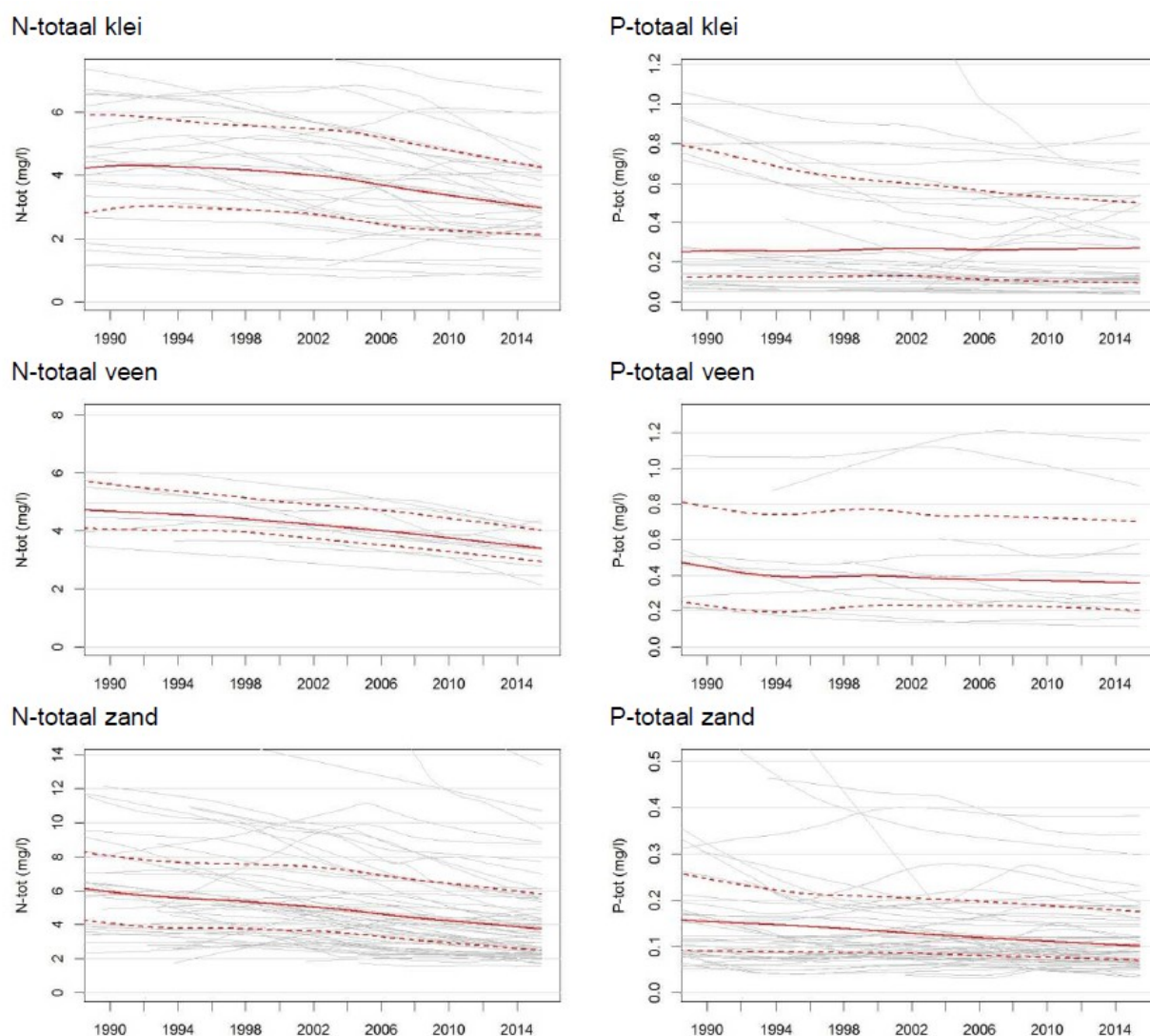
Figuur 12: Stikstofconcentratie in landbouwspecifiek oppervlaktewater

Figuur 13 komt uit het ex-post rapport (Velthof et al., 2017). In deze figuur is niet uitgesplitst naar bodemtype. Ook hier zien we dalende trends, maar geen afvlakking, zowel voor Stikstof (N) als Fosfor (P).



Figuur 13: Berekende trend (LOWESS) van de concentratie N-totaal (bovenste grafiek) en P-totaal (onderste grafiek) voor de landbouwspecifieke water; mediaan (doorgetrokken lijn) en 25- en 75-percentiel.

Figuren die wel uitgesplitst zijn per bodemtype zijn opgenomen in het achterliggende MNLSO-rapport (Klein & Rozemeijer, 2016). In **Figuur 14** is geen afvlakking van de trend in Stikstof in het zandgebied te zien. Voor Fosfor in klei is een lichte stijging (in tegenstelling tot de ombuigende trend in Figuur 11) en in zand is een duidelijke daling te zien (in tegenstelling tot de stijging in Figuur 11). In deze Figuren zijn alle individuele MNLSO-metingen (12 per jaar, minimaal 10 jaar lang) meegenomen: Hierbij is eerst de methode LOWESS toegepast per meetpunt (grijze lijnen), vervolgens een geaggregeerde LOWESS hierdoorheen. De stippellijnen zijn LOWESS-trends door de bovenste en de onderste helft van de gegevens.



Figuur 14: Geaggregeerde LOWESS-trendlijn en 25- en 75-percentiel LOWESS-trendlijnen (gestippeld) voor N-totaal en P-totaal (zomer- en winterconcentraties) voor de deelgebieden klei, veen en zand.

Het verschil tussen de trends tussen de rapportages komt door een viertal oorzaken. In de eerste plaats worden verschillende modellen gebruikt: Integrated random walk (IRW) en LOWESS. Ten tweede gebruikt het IRW-model zomergemiddelden en geen medianen. Dat kan veel uitmaken en wellicht al het verschil in trends verklaren. Ten derde zijn de onzekerheden anders bepaald. Een vierde verschil is dat bij MNLSO de trends op alle reeksen geschat zijn en vervolgens geaggregeerd. Bij het PBL zijn eerst alle reeksen gemiddeld en is daarna een trend bepaald. Verder is nog de presentatie verschillend, te zien aan de verhoudingen op de x-as en y-as.

Conclusie

Met name in het ex post rapport (Velthof et al., 2017) is een diversiteit aan trends en weergaves daarvan te zien. In het syntheserapport (PBL, 2017) is dit uniformer. Er komen in de EMW veel onderwerpen aan bod, waarvoor met verschillende methodieken de trends zijn berekend. Vervolgens kan nog de weergave van de trends verschillen. Zoals getoond kunnen verschillende aanpakken ook tot andere conclusies leiden over trends. Om die reden is het belangrijk de trendberekening te uniformeren waar mogelijk. Er moeten veel keuzes gemaakt worden om een trend te berekenen. Dit begint bij de voorbereiding van de data, gevolgd door de keuze voor de trendmethode, en uiteindelijk nog bij de nabewerking en visualisatie. Aan deze opeenvolgende stappen besteden we daarom aandacht in de komende hoofdstukken.

3. Afbakening

N.B. Hieronder wordt onderscheid gemaakt tussen het huidige methodenrapport, waarin de trendmethoden vergeleken worden, en het daarna opgestelde werkwijzerrapport, waarin beschreven wordt welke trendmethode(n) de voorkeur verdienen voor de EMW.

3.1 Doel van dit rapport

Doel van de werkgroep statistiek EMW is een eenduidige, objectieve werkwijze opstellen waarmee trends en de significantie van trends vastgesteld kan worden van meetdata/statistieken die relevant zijn voor Evaluatie Meststoffenwet ex-post rapportages. Onder de 'significantie van trends' verstaan we het aangeven of een trend significant stijgend of dalend is, dan wel dat deze gelijk gebleven is. In dit rapport geven alleen een beschrijving van de methoden, en zullen we de eigenschappen zoveel mogelijk zonder waardeoordeel op een rij zetten. We zullen geen uitgebreide experimenten doen met bijv. gesimuleerde data. In de op dit rapport volgende werkwijzerrapport zal een oordeel en gewicht toegekend worden aan de eigenschappen van de verschillende methoden, om zo tot een geschikte werkwijze voor de EMW te komen.

3.2 Doel van de trendmethoden

Hoewel we reeds hebben aangegeven dat we ons richten op trendmethoden voor data uit o.a. meetnetten, omschrijven we hier iets preciezer waarover de trendmethoden een uitspraak kunnen doen. De in de EMW te hanteren trendmethode(n) moeten worden toegepast voor het volgende:

- Er moet een trend voor een dataset uit een meetnet of voor een indicatorreeks geschat kunnen worden en in een grafiek of tabel weergegeven kunnen worden
- Er moet aangegeven kunnen worden of de trend over de laatste x jaar significant dalend/stijgend is.

Evaluatie van beleidseffecten valt buiten de scope (dus of we de verandering in trend kunnen verklaren uit een maatregel). We richten ons dus alleen op het beschrijven van de trend in de doelvariabele en niet op het leggen van verbanden met andere variabelen. Ook richten we ons op een specifiek type trend, namelijk temporele trends. Naast temporele trends zijn er immers ook nog ruimtelijke trends en spatio-temporele trends. Verder beperken we ons tot methoden voor een univariate outputvariabele. Als er bijvoorbeeld trends voor 12 provincies moeten komen wordt dat per provincie berekend, en niet in een multivariaat model waarbij correlaties tussen trends van verschillende provincies gemodelleerd kunnen worden. Als laatste beperking geldt dat we ons richten op de ex post rapportage EMW. We kijken dus niet naar methoden die expliciet bedoeld zijn om toekomstvoorspellingen mee te maken.

3.3 Data

Over welke data moeten uitspraken gedaan worden, ofwel op welke data wordt de trendmethode toegepast? Dit valt uiteen in de databronnen die als input dienen voor de EMW, en de typen data die we daarin aantreffen.

We gaan er vanuit dat er trends berekend moeten worden voor data afkomstig uit verschillende onderwerpen, afkomstig uit verschillende databronnen:

- Qua onderwerp leggen we de focus op trends in waterkwaliteit. Hier kwamen in het verleden verschillende resultaten met dezelfde data. Bij andere onderwerpen was dat minder het geval. We zullen wel de insteek nemen dat de methoden ook toegepast kunnen worden op de mineralenbalans (stikstof- en fosfaatoverschotten) en gewasopbrengsten en bodemvruchtbaarheid.

- Er zal gebruik gemaakt worden van de data afkomstig uit de volgende bronnen: LMM, MNLSO, LMG en PMG

De datatypen die we daarin tegenkomen, en waarover trends berekend moeten worden, zijn meestal metingen van meetpunten uit een meetnetwerk. Het gaat om meerdere metingen, soms met ongelijke intervallen ertussen, gedurende meerdere jaren. Ook gaat het meetpunten verspreid over het land. Verderop gaan we in op het aggregeren van deze meetpunten tot landelijke cijfers. Soms zijn de metingen al geaggregeerd tot bijvoorbeeld een landelijk of regionaal cijfer en hebben we te maken met een reeks van waarnemingen op jaarbasis waarin een trend geschat moet worden.

De data die we aantreffen kan ook nog verschillen wat betreft datakarakteristieken. Zo kunnen de data de volgende eigenschappen hebben

- Frequentie: jaarlijks, maar soms ook hogere frequentie
- Seizoenpatronen
- Hoeveelheid ruis t.o.v. het signaal
- Lengte van reeksen
- Missende waarden
- Uitbijters
- Etc.

In de beschrijving van de verschillende trendmethoden zullen we aangeven of de trendmethoden rekening kunnen houden met deze eigenschappen.

4. Selectie van een trendmethode

In dit hoofdstuk gaan we in op algemene aspecten van de trendmethoden. We beschrijven wat een trend eigenlijk is, en enkele belangrijke aspecten die voor alle trendmethoden gelden. Verder behandelen we een aantal zaken ter voorbereiding op de vergelijking van de verschillende methoden in het volgende hoofdstuk.

4.1 Wat is een trend?

Om inzicht te krijgen in het verloop van allerlei fenomenen over langere tijd worden in een publicatie vaak tijdreeksen gepubliceerd. Een probleem met het presenteren van tijdreeksen is echter de volatilititeit (de mate waarin de reeks op- en neergaat) die een reeks meestal kent. Hierdoor is het moeilijker zicht te krijgen op de onderliggende ontwikkeling. Met name aan het einde van een reeks is dit van belang, omdat we willen weten of de meest recente waarnemingen een daadwerkelijke ontwikkeling betekenen, of dat het eenmalig afwijkingen van een meer structureel patroon zijn. Om een beter beeld te krijgen van de onderliggende ontwikkeling is het vaak nuttig om een tijdreeks glad te strijken en zo de trend zichtbaar te maken. We definiëren de trend daarom als de onderliggende langetermijnontwikkeling van de reeks, los van terugkerende verstoringen zoals seizoenpatronen of eenmalige verstoringen die we ruis noemen, en uitbijters en breuken.

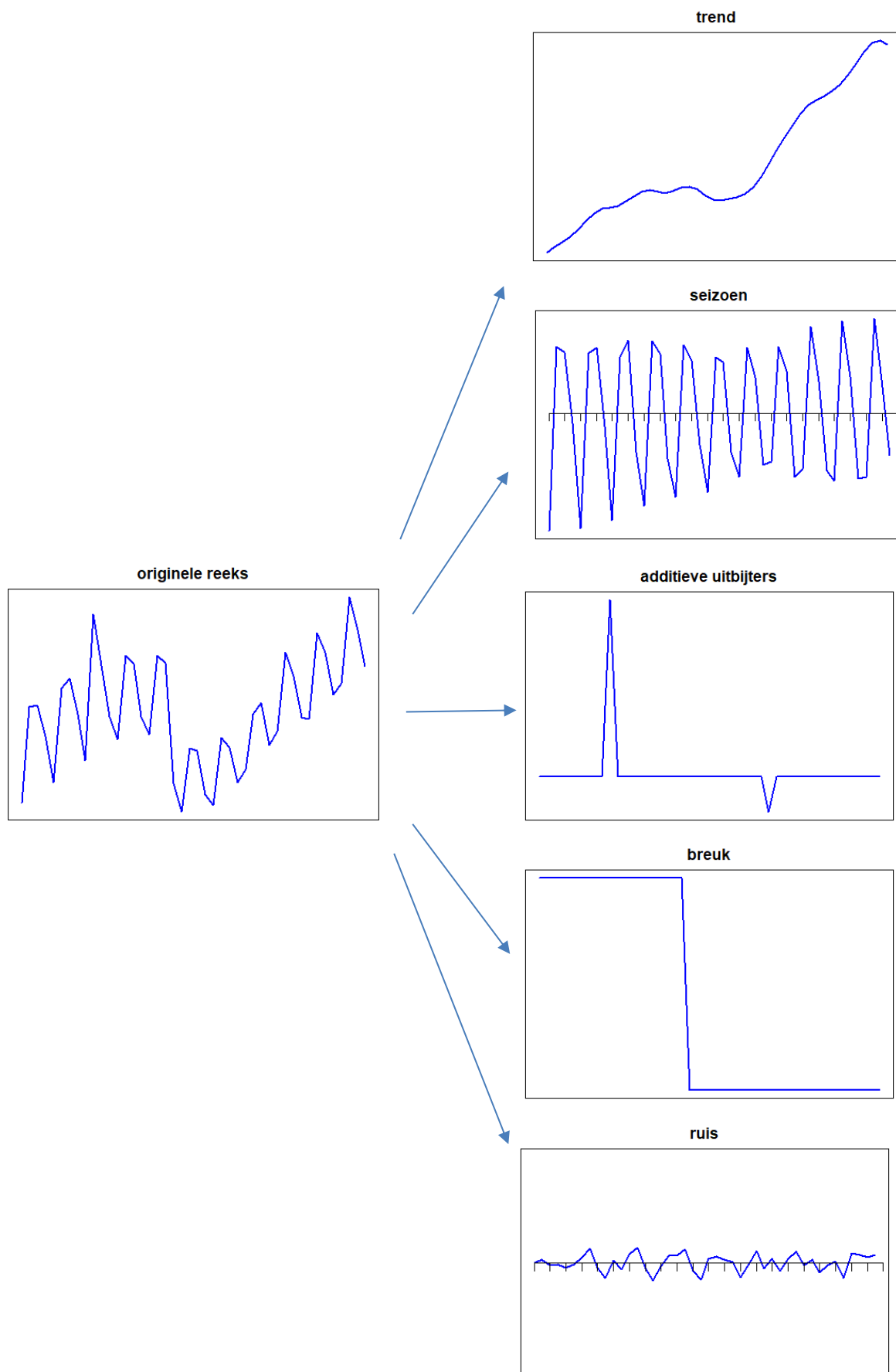
Er zijn veel manieren om deze trend te schatten. Een aanname die vaak gemaakt wordt bij tijdreeksmodellen is dat elke tijdreeks te scheiden is in een onderliggend patroon (het signaal) en eenmalige verstoringen daarvan. Deze relatie geldt voor elke periode t van de reeks:

$$\text{Reeks}(t) = \text{Signaal}(t) + \text{Ruis}(t)$$

Een volgende aanname daarbij is dat het signaal uit een aantal componenten bestaat:

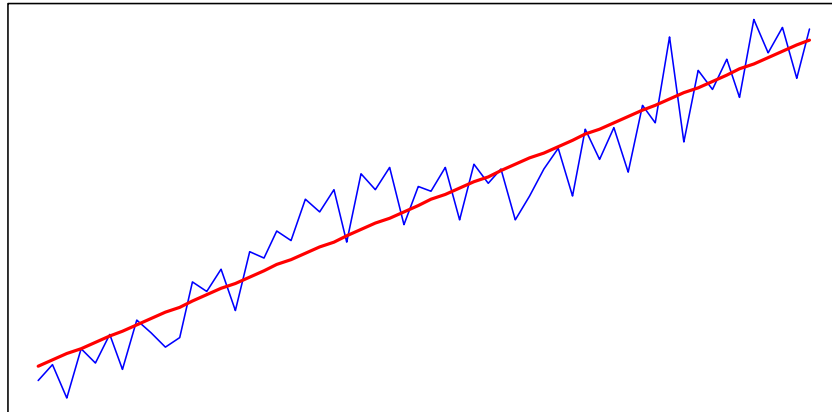
$$\text{Signaal}(t) = \text{Trend}(t) + \text{Seizoen}(t) + \text{Uitbijters}(t) + \text{Breuken}(t)$$

De reeks valt dus uiteen in een aantal deelreeksen, zoals geïllustreerd in **Figuur 15**. Met de juiste schattingsmethoden kunnen we de reeks op deze manier ontleden in de genoemde tijdreekscomponenten. Afhankelijk van het doel krijgen we dan als onderliggende trendmatige ontwikkeling de deelreeks $\text{Trend}(t)$ of bijvoorbeeld $\text{Trend}(t) + \text{Uitbijters}(t)$. Afhankelijk van of uitbijters als reële effecten gezien worden kunnen we ze dus wel of niet opnemen in de uiteindelijk trendlijn.

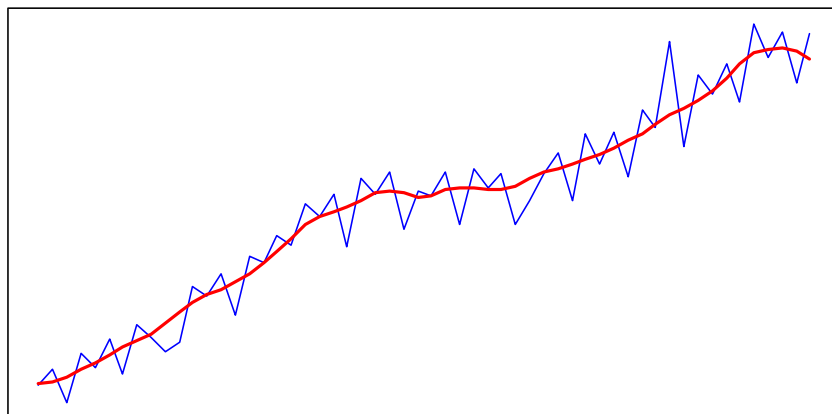


Figuur 15: decompositie van een tijdreeks

Hoewel we de trend in bovenstaande figuur hebben gepresenteerd als een flexibele lijn, zijn er meerdere manieren om naar een trend te kijken. Een trend kan globaal zijn, maar ook lokaal. Dit wordt geïllustreerd in **Figuur 16** en **Figuur 17**.



Figuur 16: voorbeeld van een globale trend



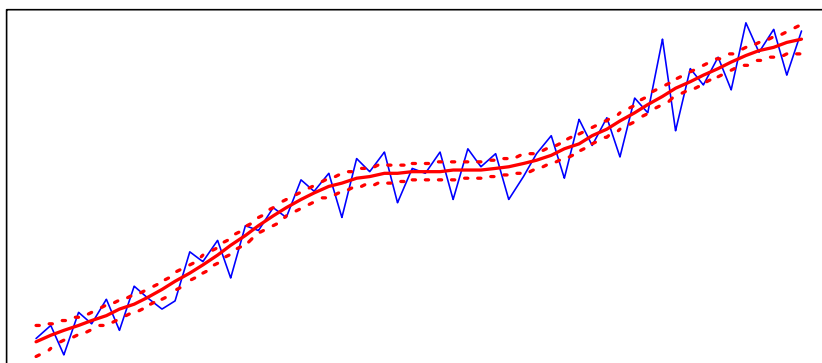
Figuur 17: Voorbeeld van een lokale trend

In het geval van een globale trend is dit bijvoorbeeld een rechte lijn met een intercept en vaste hellingshoek, maar ook een kromme is mogelijk, zolang de eigenschappen maar voor de hele reeks hetzelfde zijn. In het geval van de rechte lijn zal de onderliggende ontwikkelingen van periode tot periode dus voor de hele tijdreeks dezelfde zijn. De gevonden hellingshoek is daarmee sterk afhankelijk van het interval wat we nemen om de trend te schatten. Door de trendschatting een aantal perioden later te laten beginnen kan een andere conclusie getrokken worden over de meest actuele ontwikkelingen aan het einde van de reeks. Ook de fluctuaties in de hele reeks, zoals seizoenpatronen of op- en neergaande bewegingen over langere tijd, hebben hun weerslag op de trendschatting van de meest actuele periode. Verder is een globale trend mogelijk gevoelig voor revisies van de data. Als de reeks met een of meerder periode wordt uitgebreid, verandert de trendschatting van eerdere perioden ook. Er zijn echter ook voordelen te bedenken van globale trends. Als we een kleine hoeveelheid data hebben, ontbrekende waarnemingen, of data met veel ruis ten opzichte van het signaal, kan dit het moeilijk maken om een flexibele trend (ofwel een complexer model) betrouwbaar te schatten. In dat geval kan het beter zijn een methode te kiezen die een globale trend schat.

Een variant op de globale lineaire trend is de globale niet-lineaire trend. Dit is een trend die niet afwisselend kan stijgen en dalen, maar ook niet een rechte lijn is. Een voorbeeld hiervan is een exponentieel stijgende trend.

Zoals te zien in de figuur 16 kan een globale trend een omslagpunt aan het einde van een reeks missen. Vaak wordt daarom een flexibele of lokale trend beschouwd. In Figuur 17 is een lokale trend weergegeven voor dezelfde reeks als in Figuur 16. Hier krijgen we een duidelijker zicht op de ontwikkeling aan het einde van de reeks. Hoewel afhankelijk van de precieze methode die gekozen wordt zal deze trend in tegenstelling tot een globale trend vaak minder gevoelig zijn voor het interval wat gekozen wordt en minder last hebben van bijstellingen. Elke tijdreeksmethode blijft echter gevoelig voor het 'eindpuntprobleem'. Dit wil zeggen dat het moeilijker is om eigenschappen van een tijdreeks te schatten aan het einde van een reeks, omdat we nog niet weten hoe deze verder loopt. Hetzelfde geldt voor het beginpunt van de reeks. Het is dan ook van belang om naast de trend zelf ook altijd de betrouwbaarheid ervan te kwantificeren (betrouwbaarheidsinterval). Bij zowel de globale als lokale aanpak moet een aantal keuzes gemaakt worden door de gebruiker. De mate van gladheid van een lokale trend of de hellingshoek van de globale trend is immers afhankelijk van de methode en zelfs de parametrisering daarbinnen. Bij voorkeur wordt dus een goodness-of-fit-criterium gebuikt om de juiste modellering te kiezen, zoals de Likelihood, AIC (Akaike Information Criterion, die rekening houdt met de complexiteit van het model) of R^2 . Voor sommige methoden kunnen deze criteria niet uitgerekend worden of zijn ze niet geschikt (bijvoorbeeld bij moving averages, zie het overzicht met de methoden). In dit geval kan ook gekeken worden naar voorspelfouten of cross validatie.

Zelfs als we de beste methode kiezen volgens een objectief criterium, blijft er een mate van onzekerheid over de gevonden trend. Zoals aangegeven is een trend slechts een schatting van een onderliggende ontwikkeling. Hoewel we de trend tot nu toe presenteren als iets deterministisch, is deze dus wel degelijk met onzekerheid omgeven. Dit wordt geïllustreerd in **Figuur 18**. In de figuur is ook het eindpuntprobleem zichtbaar, aan het einde van de reeks zijn de marges op de trend groter dan midden in de reeks. Helaas is niet voor alle methoden een schatting van de onzekerheid beschikbaar, maar is het via een omweg wel mogelijk om hier iets over te zeggen. Hier komen we later op terug.

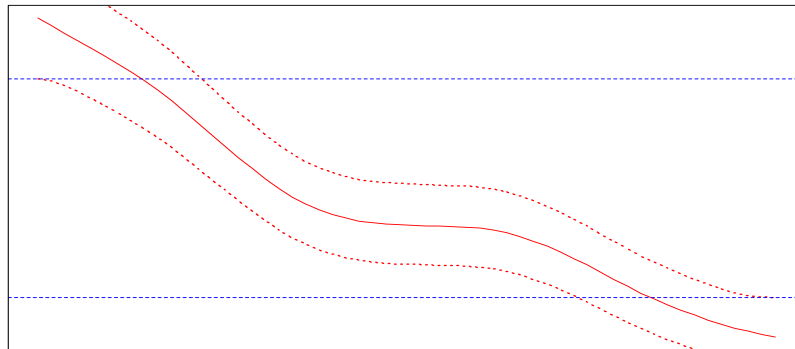


Figuur 18: Voorbeeld van een lokale trend met marges

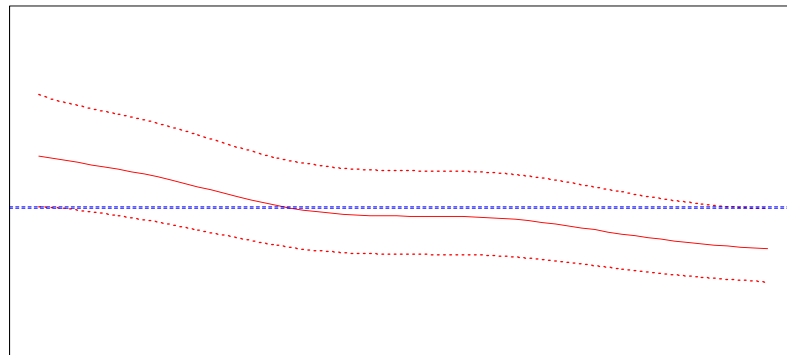
Vanwege deze onzekerheid is het niet altijd makkelijk een uitspraak te doen over het stijgen of dalen van een trend. Bij het beoordelen of een trend significant stijgt of daalt, moeten we de meest recente periode vergelijken met een basisperiode. In **Figuur 19** is een dalende trendlijn weergegeven

met onzekerheidsmarge. We zien dat de bovenmarge aan het einde van de reeks strikt lager ligt dan de ondermarge aan het begin van de reeks, en kunnen dus stellen dat de trend significant dalend is.

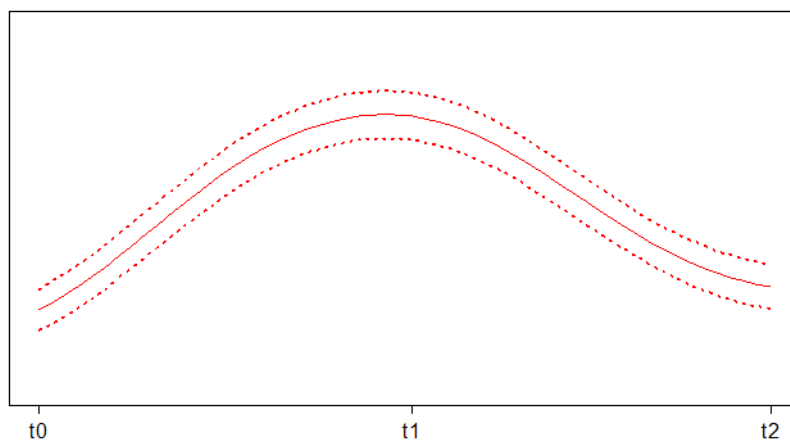
In **Figuur 20** daarentegen ligt de bovenmarge aan het einde van de reeks op gelijke hoogte met de ondermarge aan het begin van de reeks. De marges aan begin en eind zijn dus niet strikt gescheiden van elkaar, en we kunnen dus niet concluderen dat de trend aan het einde van de reeks significant lager ligt dan aan het begin. Bij een lokale trend kan deze tussen einde van de tijdreeks en de basisperiode in de tussentijd gestegen en gedaald zijn of vice versa. In **Figuur 21** is een trendlijn weergegeven waar dit speelt. Als we perioden t_0 en t_1 , of t_2 en t_1 met elkaar vergelijken kunnen we met veel zekerheid zeggen dan de trend stijgend dan wel dalend is. Als we t_0 en t_2 met elkaar vergelijken kunnen we geen conclusie trekken.



Figuur 19: Voorbeeld van een lokale trend met marges



Figuur 20: Voorbeeld van een lokale trend met marges



Figuur 21: Voorbeeld van een lokale trend met marges

Bij deze illustraties geldt wel een belangrijke kanttekening. Het is in de regel niet zonder meer mogelijk om twee punten van een trendlijn met elkaar te vergelijken en op basis van de marges rond die punten conclusies te trekken over het significant zijn van de stijging of daling. De waarnemingen van een tijdreeks zijn niet onafhankelijk, en dat geldt ook voor de waarden van de trendlijn. Bij het bepalen van de marges moet rekening gehouden zijn met de correlaties tussen de opeenvolgende waarden. In de methodiek moet daar rekening mee gehouden worden. In dat geval is een analyse zoals bovenstaand geoorloofd. Als de methodiek daar geen rekening mee houdt zijn de marges puntsgewijze onzekerheden en ontbreekt de informatie over de onderlinge correlaties daartussen. In dat geval is een aanvullende analyse nodig om conclusies over de significantie te trekken. Bijvoorbeeld in TrendSpotter wordt deze output geleverd.

4.2 Keuze tussen methoden

Er zijn verschillende afwegingen te maken om een trendmethode te kiezen. Niet elke trendmethode is geschikt voor elk type data. De data kan erg volatiel zijn, of juist zeer stabiel. Er kan een lange of korte historische reeks zijn. Daarnaast kan de frequentie (maandelijks, per kwartaal of jaarlijks) verschillen en kan er sprake zijn van cyclische of seizoenpatronen, en verschillende soorten verstoringen zoals eenmalige uitbijters of permanente niveaushuivingen (breuken). Verder wordt niet elke reeks even regelmatig waargenomen of kunnen er ontbrekende waarnemingen zijn. Niet alle methoden kunnen daar goed mee omgaan of zijn robuust genoeg. Ook is niet elke methode even geschikt voor het omgaan met gecensoreerde waarnemingen (denk aan nutriëntconcentraties beneden de detectielimiet).

Daarnaast is de keuze voor een methode afhankelijk van het doel. Soms is het belangrijk om te weten of er over een periode gemiddeld genomen een stijging of daling is. Dan kan een methode voor een globale trend geschikt zijn. Als er echter een zo goed mogelijke weergave van de dynamiek over de afgelopen jaren gewenst is, is een gedetailleerdere modellering nodig, al dan niet met onzekerheidsmarges.

Verder heeft elke methode zijn eigen (soms impliciete) aannames. Die hebben vaak betrekking op de eigenschappen van de data, in het bijzonder de stochastische eigenschappen van de onderliggende datagenererende processen. Veel voorkomende aannames richten zich op de storingstermen, en gaan er van uit dat deze (in volgorde van belangrijkheid):

- Onafhankelijk van elkaar zijn, ofwel geen autocorrelatie vertonen.
- Homoskedastisch zijn: constante variantie door de reeks heen
- Normaal verdeeld zijn:

Bij het toepassen van een model zouden deze getoetst moeten worden. Bij voorkeur wordt aan alle drie de voorwaarden voldaan, maar dat lukt niet altijd. Als er meerdere methoden zijn waarbij aan de aannames voldaan wordt, moet er een keuze gemaakt kunnen worden welke methode het meest geschikt is. Dit kan voor sommige modellen via selectiecriteria, zoals (Adjusted) R^2 , AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), Likelihood ratio toets of F-toets. Deze criteria houden niet alleen rekening met de 'fit' van het model, maar ook met de complexiteit ervan. In theorie kan een perfecte fit verkregen worden door een complexer model te kiezen. Het risico daarvan is echter 'overfitting', waarbij we een model hebben wat de toevallige fluctuaties (ruis) in de reeks modelleert, in plaats van het onderliggende patroon. Overfitten treedt op als we een te complex model toepassen in verhouding tot de hoeveelheid data, bijvoorbeeld een model met teveel verklarende componenten. Hiermee wordt in-sample een goede fit verkregen, maar zal out-of-sample de voorspellende waarde van het model klein zijn.

Keuze van instellingen

Hoewel er dus al verschillende trends gevonden kunnen worden afhankelijk van de gekozen methode, kunnen ook nog binnen een methode de uitkomsten behoorlijk beïnvloed worden. Ten eerste kan dit via de parametrisering. Bij sommige methoden (zoals ARIMA) zijn er vaste procedures die gevolgd kunnen worden om de parameters in te stellen. Bij andere methoden zijn er minder mogelijkheden of is dit subjectiever. Zoals eerder genoemd zijn de uitkomst ook vaak gevoelig voor de eigenschappen van de data. Als er dus bijvoorbeeld een kortere of langere periode genomen wordt om de trend te schatten kan dit invloed hebben op de trend aan het einde van de reeks. Sommige methoden kunnen daarnaast nog hulpinformatie meenemen bij het schatten van de trend. Dit gebeurt dan via regressoren. Door een regressor op te nemen in het model kan een deel van de variaties in de reeks verklaard worden. Dit betekent dan wel dat er een andere trend uit het model zal komen. Bij het al dan niet opnemen van regressoren wordt gekeken of deze significant zijn. Aangezien het in dit om een goede modellering van de trend gaat, en de opgenomen regressor ook de trend zal beïnvloeden, moet daarbij inhoudelijk afgewogen worden of deze regressor inderdaad een deel van de variaties in de reeks verklaart en er niet slechts sprake is van een meer toevallige correlatie.

4.3 Statistische vs. niet-statistisch methoden

Zoals uit de hoofdstuk 2 blijkt is het een randvoorwaarde dat een trendmethode ook marges levert. Eerder is toegelicht dat een trend een schatting van de onderliggende ontwikkeling van een reeks is en daarom dus met onzekerheid omgeven. In het volgende hoofdstuk zullen we zien dat er methoden zijn die de trendschatting als een deterministische uitkomst beschouwen, en geen uitspraak over de onzekerheid doen. Desondanks is het toch mogelijk om bij een niet-statistische methode een schatting van de onzekerheid te geven. Dit kan met bijvoorbeeld Monte Carlo simulatie. Voor meetnetten kan mogelijk ook Bootstrap geschikt zijn, dit is ook een vorm van Monte Carlo, maar dan met teruglegging. Als we via een dergelijke route marges kunnen afleiden, hebben we een niet-statistische methode toch in zekere zin statistisch gemaakt.

De Monte Carlo procedure werkt als volgt voor een tijdreeks met N waarnemingen. Stap 1: schat de trend met methode x op de data en daarmee ook de residuen, dat wil zeggen data minus trend. Nu zijn er twee mogelijkheden:

- (i) schat op de residuen een kansverdeling (bijv. de normale verdeling of log-normale verdeling) en doe daaruit N trekkingen, of
- (ii) maak op basis van de residuen een bootstrap sample ter grootte N (met teruglegging).

Tel nu deze gesimuleerde residu-reeks op bij de geschatte trend. Dat geeft een tijdreeks met gesimuleerde data. Schat nu op deze reeks met de methode x opnieuw een trend. Herhaal dit schema 1000 maal. Dit levert voor methode x een aantal van 1000 trends. Op die gegenereerde data kan dan elk gewenste statistiek toegepast worden, zoals bijvoorbeeld de onzekerheid in een willekeurig trendverschil $[\mu_t - \mu_s]$.

5. Voor- en nabewerkingen

Alvorens de data geanalyseerd kan worden met een trendmethode, zijn er vaak meerdere bewerkingen nodig. Sommige van deze bewerkingen kunnen ook achteraf gedaan worden. Hieronder beschrijven we deze bewerkingen, en de keuzes die daarbij gemaakt kunnen worden. Hieronder zullen we de voorbewerkingen als losstaande bewerkingen beschrijven, in de praktijk zijn ze soms in de trendmethode verwerkt. De verschillende voorbewerking hebben een belangrijke invloed op de uiteindelijke conclusies die getrokken worden. Zoals in hoofdstuk 2 is geïllustreerd met enkele voorbeelden kunnen er verschillen optreden tussen verschillende rapporten. Dit komt vaak niet in de laatste plaats door de voorbewerkingsstappen. Er is een veelheid aan mogelijkheden en de keuzes zijn afhankelijk van de specifieke eigenschappen van het meetnet en het doel van de analyse. Dit vergt dus altijd inhoudelijke afwegingen. We zullen hier een overzicht bieden van wat er zoal mogelijk is.

5.1 Aggregatie

Aangezien we naar meetgegevens uit meetnetten kijken krijgen we te maken met de vraag hoe uit deze detailinformatie een trend op aggregaatsniveau bepaald kan worden. Uit de gegevens van de lokale meetstations moet bijvoorbeeld een trend berekend worden op provinciaal of op landelijk niveau. Hier zijn meerdere aanpakken voor. Feitelijk gaat het om een weging van de verschillende meetpunten. Samenvattend moeten er de volgende methodologische keuzes gemaakt worden voor het volgende:

- Stratificatie
- Welke meetpunten aggregeren we?
- Aggregatiefunctie
- Volgorde

We zullen deze toelichten aan de hand van een aantal bijkomende aspecten waarmee rekening gehouden moet worden bij de keuze voor een geschikte methode. Bij de stratificatie kunnen we denken aan de volgende problemen:

- Een meetnet is niet altijd uniform verdeeld over het land. Soms worden er extra meetpunten geplaatst in gebieden waar problemen zijn, en die men dus extra in de gaten wil houden.
- Ook kunnen de meetpunten van karakter verschillen. Zo wordt bij het LMM onderscheid gemaakt naar grondsoortregio's: de Zand-, Klei-, Veen- en Lössregio en binnen deze regio's één of meerdere gebieden onderscheiden (Zie Figuur 2).

Dit betekent dat we de meetpunten niet zomaar kunnen aggregeren tot een landelijk totaal. Ofwel we moeten de meetpunten eerst samenvoegen per deelgebied (bijvoorbeeld een provincie of een gebied van 10x10 km), of we moeten alle punten een bepaalde weging meegeven zodanig dat bepaalde type punten niet oververtegenwoordigd zijn. In de ruimtelijke statistiek worden hier methoden voor gebruikt zoals polygoondeclustering en celdeclustering (Isaaks & Srivastava, 1989, Hfst 10). Een meer gesofisticeerde methode is om hiervoor de covariantiematrix te gebruiken. Ook kunnen we besluiten om niet (volledig) te aggregeren, maar gestratificeerd te publiceren, zoals bij het LMM naar grondsoortregio.

Bij het aggregeren hoeven niet alle punten meegenomen te worden. Soms vertekenen bepaalde metingen ongewenst het beeld. Van sommige meetpunten is bijvoorbeeld een korte meetreeks beschikbaar, of is deze zeer onregelmatig door ontbrekende waarden. We kunnen dan bijvoorbeeld per periode alle meetpunten aggregeren die er zijn (methode 'all'), of alleen de meetpunten meenemen die in minimaal (bijv.) 80% van de perioden van de meetreeks voorkomen (methode 'beauties').

Bij de aggregatiefunctie gaat het erom hoe we meetpunten bij elkaar optellen. We kunnen van alle meetwaarden bijvoorbeeld een gemiddelde nemen, een geometrisch gemiddelde, of een mediaan. De mediaan heeft als voordeel dat deze relatief ongevoelig is voor extreme waarnemingen die het beeld kunnen verstoren. Tegelijkertijd kan ook het gewone gemiddelde juist voordelen bieden:

- Bij de analyse van een aantal meetstations met Fosfor- en Stikstof-metingen bleek dat de trends nogal verschilden. Dat kwam door een klein aantal vieze beken met hele hoge waarden aan het begin, die allemaal sterk verbeterden. Dat gaf een sterk dalende trend in de tijd. Maar het geometrisch gemiddelde zorgde voor een vrijwel horizontale trend over de hele meetperiode.

Als laatste kan nog gekozen worden in welke volgorde geaggregeerd wordt. Zo zouden eerst alle meetpunten geaggregeerd kunnen worden, en daarna een trendlijn bepaald kunnen worden, maar het kan ook andersom. Voordeel van eerst de ruwe gegevens aggregeren is dat er een heleboel ruis wegvalt op aggregaatsniveau, waarna het berekenen van de trend nauwkeuriger kan gebeuren. Een nadeel is dat je informatie verliest en dat een geaggregeerde gemiddelden (of andere statistiek) onzeker is door de grote ruimtelijke verschillen in concentratieniveaus. De ruimtelijke verschillen in trend zijn mogelijk veel kleiner dan de ruimtelijk verschillen in concentraties. Voordeel van een trendberekening voor de aggregatiestap kan zijn dat de invloed van uitbijters beperkt wordt, en dat op het laagste niveau de onderliggende ontwikkeling bepaald wordt.

5.2 Zomer- en wintergemiddelden

Gerelateerd aan het aggregatievraagstuk over de meetpunten is de vraag hoe er binnen een jaar geaggregeerd moet worden. Bij de analyse van trends wordt vaak naar jaarcijfers gekeken. Dat kan een eenvoudig gemiddelde over de 12 maanden zijn, maar dat is bepaald niet triviaal omdat binnen het jaar vaak veel missend is, alleen in het zomerhalfjaar wat metingen zijn, etc. Ook hier gelden dezelfde argumenten om te kiezen tussen bijv. mediaan of rekenkundig gemiddelde.

5.3 Transformaties

Veel statische modellen gaan uit van een aantal aannames. Bijvoorbeeld dat er een lineair verband in de data aanwezig is of dat de ruisterm een normale verdeling volgt. In de praktijk wordt hier niet altijd aan voldaan, en is het moeilijker een geschikt model te vinden voor een dataset. Een oplossing is dan om een transformatie op de data toe te passen voordat het model geschat wordt. In de praktijk wordt vaak een logaritmische transformatie toegepast. Veel van de genoemde problemen, of problemen door uitbijters en negatief-wordende trends, zijn te voorkomen met een transformatie vooraf.

5.4 Corrigeren voor versturende effecten

Soms is het wenselijk het verloop van trends te corrigeren voor externe invloeden. Door deze invloeden te verwijderen ontstaat een helderder beeld op de onderliggende trend. Voor de concentraties van nutriënten in het oppervlaktewater bijvoorbeeld zijn neerslag en temperatuur belangrijk. Vooral de uit- en afspoeling van landbouwgronden worden sterk beïnvloed door neerslag. We zijn in dat geval geïnteresseerd in de trend zonder dat deze vertekend wordt door het neerslagpatroon. Andere factoren waarvoor gecorrigeerd kan worden zijn verschillen in reistijd (hoe oud is het water op bepaalde diepte) en temperatuur.

Daarnaast zijn er nog uitbijters en breuken. Deze verstoren de berekening van een trend. Aan de ene kant is het dus vaak gebruikelijk om deze voorafgaand aan de trendberekening te verwijderen uit de data. Aan de andere kant bevatten ze ook nuttige informatie, die je ofwel wil laten zien in de uiteindelijke trend, of die je zelfs mee wil nemen in de trendberekening. In het eerste geval kan je ze eerste uit de reeks verwijderen, en na berekening van de trend weer toevoegen aan de reeks. Een andere optie is om een methode te kiezen die robuust is ten aanzien van uitbijters.

Een breuk is een bijzonder soort uitbijter, deze is namelijk permanent van aard. Een breuk kan veroorzaakt worden doordat er daadwerkelijk iets bijzonders gebeurt in de meetreeks, maar ook door een definitieverandering of methodewijziging. In het kader van meetnetwerken valt bijvoorbeeld te denken aan detectiegrenzen (de minimale concentratie die waargenomen kan worden) die wijzigen door technologische vooruitgang, waardoor concentraties gedetailleerder gemeten kunnen worden.

5.5 Veranderingen in meetnetten

Een belangrijk punt waarmee rekening gehouden moet worden is de samenstelling van de meetnetten zelf. De verzameling van meetpunten kan veranderen, er kunnen namelijk meetpunten bijkomen of afvallen, om verschillende redenen. Als het gaat om bedrijven kunnen deze stoppen met meewerken aan de steekproef. In andere gevallen kunnen meetpunten toegevoegd worden omdat er een specifiek doel is (zoals het extra meten van een bepaalde regio). Daarnaast kunnen de meeteenheden zelf nog veranderen qua eigenschappen. Denk aan een bedrijf wat een stuk zandgrond verkoopt en kleigrond aankoopt.

6. Overzicht van methoden

6.1 Methoden

Er zijn in de literatuur veel methoden te vinden voor het berekenen van Trends. Een uitgebreid overzicht van 30 methoden is te vinden in Visser et al. (2015). Daarin worden de volgende categorieën onderscheiden:

1. Exploratieve methoden
2. Parametrisch methoden
3. Nonparametrische methoden
4. Stochastische modellen
5. Overig

Voor de EMW zullen we in dit hoofdstuk de eigenschappen van de volgende methoden beschrijven (met tussen haakjes de categorie volgens Visser et al, 2015):

1. Voortschrijdende gemiddelden (cat. 1)
2. Seasonal Mann-Kendall (cat. 1)
3. Theil-Sen (cat. 1)
4. Lineaire regressie (OLS) (cat. 2)
5. Locally Weighted Scatterplot Smoothing (LOWESS) (cat. 3)
6. Structurele tijdreeksmodellen (cat. 4)
7. Auto Regressive Moving Average (ARIMA)-modellen (cat. 4)
8. Vergelijken van twee deelperioden

6.2 Eigenschappen van de methoden

We beschrijven de methoden adhv een aantal kenmerken die ze kunnen hebben. We doen dit zo uniform mogelijk zodat het makkelijk vergelijken is tussen de methoden. We gaan in op de volgende aspecten:

- Datakenmerken.
 - i) Geschiktheid en robuustheid van de methode: in hoeverre kan de methode überhaupt omgaan met de volgende aspecten die zich kunnen voordoen in de data, en zijn de uitkomsten van de methode gevoelig voor het optreden van deze aspecten in de data? Hier is het mogelijk nuttig om enkele gestileerde voorbeeldreeksen te simuleren om aan te tonen dat een methode daar al dan niet mee om kan gaan.
 - o allerhande typen trends
 - o veel ruis (signaal-ruisverhouding)
 - o missende data
 - o uitbijters (die niet verwijderd zijn in de voorbewerking)
 - o Omslagpunten (=trendomkering)
 - o Korte reeksen
- Schatten en parametrisering
 - o Hoe wordt de trend geschat door de methode?
 - o Is het een parameter-vrije methode of niet? Is het een heuristische methode of veronderstelt het een onderliggend model?
 - o Hoe wordt de methode geparametriseerd?
- Onzekerheid
 - o Is de methode een wiskundige methode (dus deterministisch) of een statistische methode (dus stochastisch, m.a.w. er wordt ook ruis verondersteld) en kan hij dus een marge leveren?
- Aanvullende informatie meenemen

- Bijvoorbeeld regressoren om de trend beter te kunnen bepalen
- Robuustheid
 - Ten aanzien van modelveronderstellingen, in het bijzonder degene die moeilijk te verifiëren zijn.
 - Bijstellingen bij uitbreiden reeksen (hoe veranderen de vorige waarden als een waarneming aan de reeks wordt toegevoegd)

6.3 Voortschrijdende gemiddelden

Werking van de methode

Er zijn meerdere manieren om reeksen glad te strijken. Een van de eenvoudigste is om voortschrijdende gemiddelden te nemen. Bij een voortschrijdend gemiddelde of Moving Average (MA) wordt een gemiddelde van een aantal tijdreekswaarden genomen. Vervolgens wordt dit interval één periode opgeschoven en wordt weer een gemiddelde genomen. Op deze manier ontstaat een gladde, flexibele trend. In de regel worden de gemiddelden symmetrisch rondom de periode waarvoor we de trend willen berekenen genomen. Bij maand- of kwartaaldata is dit niet goed mogelijk, waardoor vaak gekozen wordt voor een gemiddelde van meerdere voortschrijdende gemiddelden. Een nadeel van voortschrijdende gemiddelden is dat deze altijd enigszins achter lopen bij de ontwikkelingen in de originele reeks. Aangezien de gemiddelden symmetrisch worden genomen reageren ze vertraagd op veranderingen. Om dezelfde reden kunnen ze niet goed toegepast worden aan de uiteinden van de reeks (tenzij we de reeks eerst extrapoleren). Verder zijn ze vaak te eenvoudig om met de complexe dynamiek in een reeks rekening te houden.

Varianten

Als variant op voortschrijdende gemiddelden kan ook een voortschrijdende mediaan genomen worden. Dit kan bij bepaalde toepassingen (zoals meetnetten) mogelijk robuustere resultaten opleveren. Bij kwartaal- of maanddata is het niet mogelijk een MA symmetrisch rondom een periode te nemen. Met complexere moving averages lukt dit wel. Er zijn veel van dergelijke filters te bedenken, bijvoorbeeld een 2x4-filter of 2x8-filter.

Het verder uitbreiden van de methode met bijvoorbeeld hulpinformatie is niet mogelijk. Hetzelfde geldt voor uitbijters en trendbreuken; deze zouden via voorbehandeling eerst verwijderd moeten worden.

Parametrisering

De lengte van de filters kan gekozen worden door de gebruiker. Lange filters halen meer ruis weg, maar zijn trager en kunnen bovendien resulteren in te veel uitvlakking (smoothing). Korte filters volgen nauwer de originele reeks, maar zijn dus gevoeliger voor toevallige fluctuaties.

Lineaire filters

Standaard voortschrijdende gemiddelden hebben als nadeel dat ze een bepaalde vertraging kennen. In Perrucci en Pijpers (2017) wordt gekeken naar een variant op de voortschrijdende gemiddelden, waar de filtergewichten zorgvuldig op specifieke wijze gekozen zijn. In dat geval kan de vertraging beperkt gehouden worden. Met deze methode is het ook mogelijk om te extrapoleren (zie Pijpers, 2018). Deze lineaire methoden hebben bovendien nog voordelen als je multivariate reeksen wil behandelen die een gemeenschappelijke trend hebben, of als er restricties zijn op lineaire combinaties van de reeksen waar je exact aan wilt (blijven) voldoen na de trendschatting (zie Pijpers, 2018).

6.4 Mann-Kendall (MK) en Seasonal Mann-Kendall (SMK)

Werking van de methode

De methode is een toets en geeft dus geen trend. De SMK-toets is speciaal bedoeld voor cijfers met een seizoenpatroon. De toets is vooral nuttig als men geïnteresseerd is in de significantie van een trend. De methode is niet-parametrisch en geeft de significantie en richting van de trend. De significantie hangt af van de mate van stijging of daling en van het aantal waarnemingen. Bij een korte reeks of als er veel ontbrekende waarden zijn kan een op het oog significante trend toch niet als significant bestempeld worden.

Voor- en nadelen

SMK is robuust in de zin dat deze ongevoelig is voor uitbijters en kan omgaan met ontbrekende waarnemingen. De toets zoekt een monotoon stijgende of dalende trend, bij complexere patronen kan de toets dus trends onterecht als insignificant aanmerken.

6.5 Theil-Sen

Werking van de methode

Theil-Sen is een methode om op robuuste wijze een lijn te schatten in een puntenwolk door de mediane helling te berekenen van lijnen die tussen elk puntenpaar getrokken kunnen worden. De uitkomst is een globale trend. Aangezien het een wiskundige methode is kent deze geen onzekerheidsmarges.

Voor- en nadelen

Een belangrijk voordeel is dat deze methode weinig last heeft van uitbijters of ontbrekende datapunten, in tegenstelling tot OLS.

6.6 Lineaire regressie

Werking van de methode

De methode schat via de kleinste kwadratenmethode een rechte lijn door een puntenwolk, en legt daarmee een verband tussen een verklarende en een te verklaren variabele. De methode levert dus een globale trend. Door als verklarende variabele de tijd te nemen is deze methode ook voor tijdreeksen toepasbaar.

Varianten

Er zijn veel uitbreidingen zoals multiële regressie met meerdere verklarende variabelen, kwadratische OLS, piecewise linear OLS.

Voor- en nadelen

OLS is gevoelig voor uitbijters. Als bekend is in welke periode deze optreden kunnen ze echter gemodelleerd worden. Daarnaast gelden een aantal aannamen ten aanzien van de residuen, die eerst gecontroleerd moeten worden.

6.7 LOWESS

Werking van de methode

LOWESS (Locally Weighted Scatterplot Smoothing) is een generalisatie van moving averages en regressietechnieken. Voor elk datapunt wordt een polynoom geschat met de (gewogen) kleinste kwadratenmethode, op basis van een deel van de gehele reeks. De punten dichtbij krijgen daarbij een hoger gewicht en de punten verder weg een lager gewicht. De methode is non-parametrisch.

6.8 Structurele tijdreeksmodellen (STM)

Werking van de methode

De klasse van structurele tijdreeksmodellen (en met name de hierbij gebruikte schattingstechniek) modelleert elk van de componenten (zoals de trend of het seizoenpatroon) van deze reeks expliciet. Dit heeft als voordeel dat het inzichtelijk gemaakt wordt hoe de dynamiek van de reeks in verschillende onderdelen uiteenvalt en zorgt dat eenvoudig een of meerdere van deze componenten bekeken kan worden.

Een STM kent meerdere onbekenden: i) de waarden van de toestandsvariabelen en ii) de varianties van de storingstermen, de hyperparameters. Het schatten van een structureel tijdreeksmodel bestaat daarom uit twee stappen:

- Via een Maximum Likelihood (ML)-procedure worden de hyperparameters geschat. Hiertoe maximaliseren we de likelihoodfunctie via een numerieke optimalisatieprocedure.
- Vervolgens wordt middels de Kalmanfilter recursief door de reeks heengelopen om de waarden van de toestandsvariabelen voor elke periode van de reeks te bepalen. Zodra alle opeenvolgende waarnemingen behandeld zijn, wordt nogmaals door de reeks heengelopen, en worden de schattingen aangepast via een 'smoothing'-algoritme. Hiermee worden de componenten van de tijdreeks geschat. De Kalmanfilter levert (o.a.) 'gesmoothde' schattingen van de tijdreekscomponenten op. Dit is waar we hier in geïnteresseerd zijn. Deze laatste stap levert ook marges voor de gesmoothde reeks op.

Bij het smoothen zijn we vaak geïnteresseerd in de trendlijnen. Deze worden berekend als het gesmoothde signaal, minus de seizoeneffecten, maar inclusief de eventuele uitbijters en andere verklarende effecten.

In een STM wordt de reeks waarnemingen y_t uitgedrukt als een functie van een trend μ_t (plus eventueel andere componenten) en een onregelmatige term ε_t . De trend van de volgende periode μ_{t+1} is gelijk aan de trend van de huidige periode μ_t plus een richtingscoëfficiënt v_t , die ook weer stochastisch is. Een speciaal geval van STM is het smooth trend model ofwel integrated random walk (IRW). Dit is een variant waarbij de slope-component (richtingscoëfficiënt van de trend) stochastisch is, door de storingstermen ζ_t , maar de storingsterm η_t in de niveaucomponent ontbreekt:

$$y_t = \mu_t + \varepsilon_t \quad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2)$$

$$\mu_{t+1} = \mu_t + v_t$$

$$v_{t+1} = v_t + \zeta_t \quad \zeta_t \sim NID(0, \sigma_\zeta^2)$$

Hoewel het niveau geen storingsterm heeft, is deze component wel stochastisch aangezien de slope stochastisch is. De veranderingen in de slope-component, de storingen ζ_t , komen op deze manier

echter niet direct in de meetvergelijking terecht, maar werken indirect via de niveau-component door. Dit zorgt voor een gladder verloop van het signaal μ_t in de eerste vergelijking.

Varianten en flexibiliteit

Onder andere OLS en ARIMA-modellen zijn speciale gevallen van STM. STM hebben verder de mogelijkheid om regressoren op te nemen. Deze kunnen tijdsafhankelijk gemodelleerd worden (waarbij de coëfficiënt door de tijd heen varieert). Verder kan de Kalmanfilter omgaan met ontbrekende data, waardoor de methode geschikt is voor meetreeksen die niet elk jaar worden waargenomen.

STM zijn zeer flexibel en kunnen veel tijdreeksdata goed modelleren. De modellen kunnen bijvoorbeeld een combinatie van (flexibele of globale) trends, (deterministische of stochastische) seizoenpatronen en cyclische bewegingen modelleren, aangevuld met regressoren, uitbijters en trendbreuken. Er zijn zowel univariate als multivariate formuleringen.

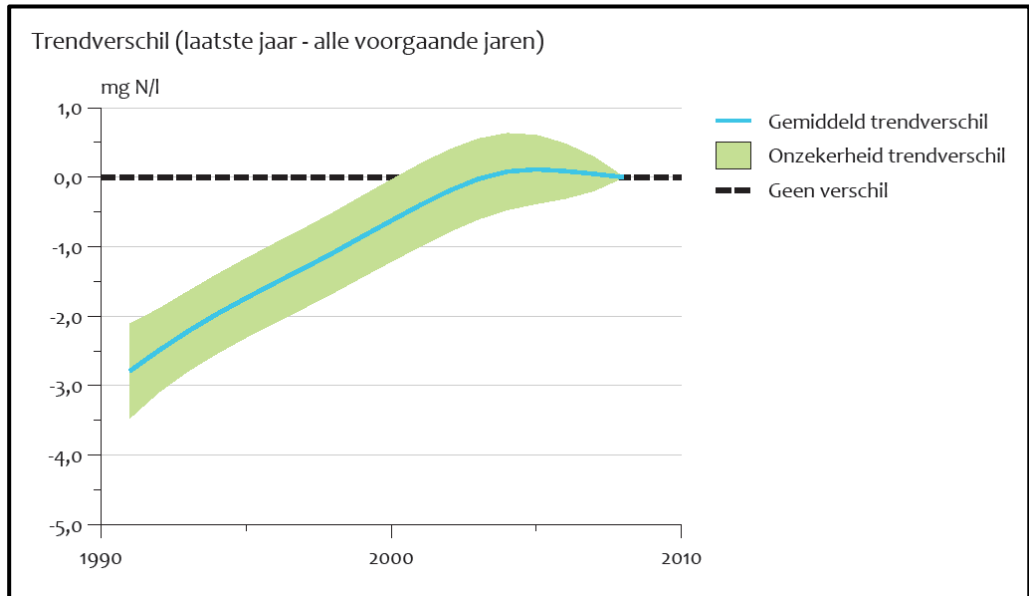
In tegenstelling tot ARIMA-modellen is er geen automatische identificatieprocedure beschikbaar om de juiste modelspecificatie te kiezen. Dit vergt handmatig beoordelen van de kwaliteitsmaten zoals AIC, residuen en significantie van hyperparameters.

Parametrisering

De ML-procedure is afhankelijk van startwaarden voor de hyperparameters. De keuze voor de startwaarden bepaalt of de ML-procedure wel met een goede oplossing kan komen, en heeft dus invloed op de parameterschattingen. De likelihoodfunctie kan namelijk meerdere lokale optima bevatten, en afhankelijk van de gekozen startwaarden kan deze in een lokaal optimum terecht komen. Om hier achter te komen moeten eenvoudigweg meerdere startwaarden geprobeerd worden, er is helaas geen optimalisatiemethode voor beschikbaar. De uiteindelijke startwaarden worden gekozen o.b.v. de beste modelfit.

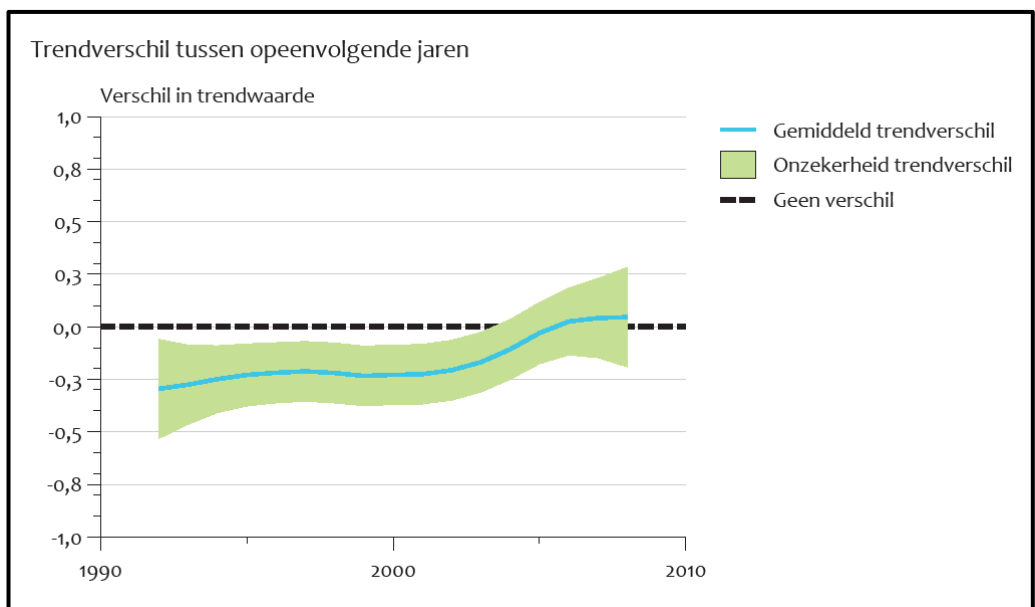
TrendSpotter

Bij het PBL is een softwarepakket ontwikkeld wat gebaseerd is op structurele tijdreeksmodellen. Het pakket wordt veel gebruikt in de context van de meststoffenwet, in het bijzonder in PBL (2017). Naast de berekening van trends en marges, en de visualisatie hiervan, levert TrendSpotter ook de trendverschillen. Dit is de trendwaarde van een bepaald basisjaar t.o.v. de actuele trendwaarde. Hiermee kan vastgesteld worden of een trend significant gestegen of gedaald is. In **Figuur 22** is hier een voorbeeld van te zien. Deze geeft de trendverschillen $[\mu_{2008} - \mu_t]$ met 95% onzekerheden in die verschillen. Omdat de bovenste onzekerheidslijn in het jaar 2000 de nullijn kruist, geldt dat trendverschillen tussen 2008 en de jaren voor 2000 significant verschillend zijn. Dus over de periode 2001-2008 is de trend in het eindjaar niet significant ten opzichte van alle andere jaren in deze periode.



Figuur 22: Trendverschillen ten opzichte van het eindjaar 2008. Ook de 95% betrouwbaarheidsbanden voor deze verschillen zijn gegeven.

Figuur 23 geeft het trendverschil $[\mu_t - \mu_{t-1}]$, wederom met 95% onzekerheidsbanden. Dit trendverschil kan gezien worden als de eerste afgeleide van de trend μ_t . De figuur laat zien dat trendverschillen over de periode 1991-2003 statistisch significant negatief zijn. Dit duidt op een sterke daling in deze periode, hetgeen natuurlijk klopt met het verloop getoond in de bovenste grafiek.



Figuur 23: Eerste gedifferentieerde van de trend μ_t met 95% betrouwbaarheidsintervallen. Negatieve waarden duiden op daling van de trend zoals getoond in Figuur 22.

6.9 ARIMA-modellen

Werking van de methode

Ook wel bekend als de Box-Jenkins-methodologie. Dit is net als STM een klasse van modellen. De modelvorm wordt gespecificeerd door 3 parameters, (p,d,q) , die de afhankelijkheid van de waarnemingen van eerdere waarnemingen specificeren. Op die manier is elke waarneming een som van autoregressieve termen en stochastische storingsprocessen. Eventueel is dit uit te breiden naar seizoensmodellen, we krijgen dan een SARIMA(p,d,q)(P,D,Q)-model. In tegenstelling tot de klasse van STM zijn hier de tijdreekscomponenten niet expliciet gegeven.

Varianten en flexibiliteit

Een ARIMA-formulering bestaat uit slechts 1 vergelijking. Deze kent de vorm van een regressie-vergelijking, waardoor eenvoudig regressoren, uitbijters en breuken toegevoegd kunnen worden. Omgaan met missende data is minder eenvoudig.

Parametrisering

Wordt vaak gedaan aan de hand van de (partiële) autocorrelatiefunctie. De structuur hiervan is een indicatie voor de orde van de parameters p en q . Als de data niet stationair is, wordt ook parameter d ongelijk aan nul gekozen. Deze laatste stap is belangrijk, omdat een aanname onder deze klasse van modellen is dat een reeks stationair is (of dus eerst gemaakt moet worden). Bij STM geldt de eis van stationariteit niet. In de praktijk zijn reeksen zelden perfect stationair, zelfs niet na deze stap. Onder andere in de seizoencorrectiemethode X-13-ARIMA is een automatische modelidentificatie-procedure opgenomen. Als de orde (p,d,q) van het model gekozen is, moeten vervolgens alle parameters nog geschat worden.

6.10 Twee deelperioden vergelijken

Een eenvoudige trendmethode is die waarbij een tijdreeks wordt opgedeeld in twee deelperioden en vervolgens de gemiddeldes (of een ander aggregaat) van beide perioden vergeleken worden. Dat kan bijvoorbeeld door de meetperiode in twee gelijke delen te splitsen of bijvoorbeeld de waarde in het eindjaar te vergelijken met het gemiddelde van de hele voorgaande periode.

Deze methode is erg gevoelig is voor de keuze van de deelperioden. Verder is het incorrect om de deelperiode-gemiddelden te testen op statistische significantie met een Student t-toets of de toets van Wilcoxon omdat men dan geen rekening houdt met de aanwezige autocorrelatie in de tijdreeks.

7. Vergelijking tussen methoden

Hier geven we een overzicht van de eigenschappen van de verschillende methoden.

Algemene kenmerken

	Methode	Parametrisch of non-parametrisch	Wordt onzekerheid gegeven?	Lokale of globale trend	Wordt de trend expliciet gegeven?	Geeft de trendmethode ook een waarden voor het begin en eind van de reeks?	Predicties onderdeel van de methode?
1	Voortschrijdende gemiddelden	Non-param.	Nee	Lokaal	Ja	Nee 1)	Nee
2	Lineaire regressie	Param.	Ja	Globaal 2)	Ja	Ja	Ja
3	Theil-Sen	Non-param.	Nee	Globaal	Ja	Ja	Nee
4	Mann-Kendall (Seasonal)	Non-param	Ja	Globaal	Nee	Nvt	Nee
5	Lowess	Non-param	Nee 3)	Beide	Ja	Ja	Nee
6	STM	Param.	Ja	Beide	Ja	Ja	Ja
7	ARIMA	Param.	ja	Beide	Nee	Ja	Ja

- 1) Bij symmetrische voortschrijdende gemiddelden niet. Met aangepaste filters, bijvoorbeeld door speciale eindgewichten te nemen, is dit eventueel wel mogelijk.
- 2) Met piecewise linear regression is ook een wat meer lokale trend mogelijk.
- 3) Dit kan bijvoorbeeld wel met de "loess" implementatie in R.

Bij de volgende kolommen gelden nog enkele opmerkingen:

- Bij de kolom 'lokale of globale trend' wordt bedoeld wat voor een trend de methode schat als we in één keer alle datapunten gebruiken. Door bijvoorbeeld een reeks in stukken op te delen zou met een globale methode toch een min of meer lokale trend benaderd kunnen worden.
- Bij de kolom 'predicties onderdeel van de methode' bedoelen we dat de methode standaard op eenvoudige wijze predicties kan maken, en iets zegt over de kwaliteit hiervan. Met enige moeite kan dat met bijna elke methode, maar dan zijn vaak aanvullende methoden nodig. Bijvoorbeeld bij voortschrijdende gemiddelden zou de gladgestreken reeks weer geëxtrapoleerd kunnen worden met een ARIMA-model.

Parametrisering

	Methode	Is er een methode om de parameter voor flexibiliteit te schatten?	Model-identificatie mogelijk	Uitbreiding met regressoren mogelijk?
1	Voortschrijdende gemiddelden	Nee (maar wel beïnvloedbaar)	Nee	Nee
2	Lineaire regressie	Ja 1)	Ja	Ja
3	Theil-Sen	n.v.t.	Nee	Nee
4	Seasonal Mann-Kendall	n.v.t.	n.v.t.	Nee
5	Lowess	Ja 2)	Nee	Ja 1)
6	STM	Ja	Ja	Ja
7	ARIMA	Ja (impliciet)	Ja	Ja

- 1) Je kunt de graad van de polynoom (veelterm) die je fit met lineaire regressie schatten met kruisvalidatie. De graad van de polynoom is een maat voor de flexibiliteit
- 2) Via kruisvalidatie
- 3) In de loess-implementatie in R

Eigenschappen t.a.v. data

	Methode	Kan de methode omgaan met...				
		ontbrekende waarden	Volatiele data	Seizoenpatronen	uitbijters / niveauverschuivingen	korte reeksen
1	Voortschrijdende gemiddelden	Nee	Ja	Ja	Nee	Afhankelijk van gewenste window
2	Lineaire regressie	Ja	ja	Nee	Ja	Ja
3	Theil-Sen	Ja	ja	nee	Ja	Ja
4	Seasonal Mann-Kendall	Ja		Ja	Ja	Ja
5	Lowess	ja		Nee	Ja*	Ja
6	STM	Ja	Ja	Ja	Ja	Afh.v complexiteit model
7	ARIMA	nee	Ja	Ja	Ja	Afh.v complexiteit model

* Alleen door een transformatie vooraf, zoals een logaritmische transformatie.

8. Conclusies

Verschillende trendmethoden kunnen verschillende conclusies opleveren. Om dit bij toekomstige EMW-rapporten te voorkomen is het een goed idee om een uniforme aanpak te hanteren. Doel van dit CBS-rapport is een vergelijking van de methoden te geven. Aanvullend hierop is een werkwijze-rapport geschreven, waarin een aanpak staat waarmee binnen de EMW gewerkt kan worden.

Er zijn een aantal algemene overwegingen te maken om tot een werkwijze voor de EMW te komen:

- Verschillende trendmethoden kunnen leiden tot verschillende conclusies over het stijgen, dalen of gelijk blijven van een reeks. Een mogelijkheid is om altijd naar meerdere methoden te kijken bij het beoordelen van trends in een dataset.
- Omslagpunten aan het eind van de reeks worden gemist als alleen gekeken wordt naar de globale trend. Het is dan ook belangrijk om de betrouwbaarheid van de trend aan het einde van de reeks te kwantificeren om de trend op waarde te kunnen schatten.
- Een trend kent altijd onzekerheid, ook al levert de methode daar geen informatie over. Methoden die onzekerheidsmarges leveren (al dan niet via een omweg, zoals Monte Carlo simulatie of Bootstrap) kunnen verkeerde conclusies voorkomen.
- Een trendmethode moet geschikt zijn voor de te onderzoeken data. De al dan niet impliciete aannames over de (stochastische) eigenschappen van de data die een methode veronderstelt moeten dus aansluiten bij die dataset. Bij het toepassen van een trendmethode zouden die aannames gecontroleerd moeten worden. Een keuze tussen verschillende methoden kan gemaakt worden via modelselectiecriteria aangevuld met proceskennis om de plausibiliteit van het model te waarborgen. Op deze manier kan een objectieve keuze gemaakt worden, en deze onderbouwd worden naar gebruikers.
- Er kan per situatie of reeks een optimale methode gekozen worden. Dit vergt veel inspanning per situatie. Om tot een aanpak te komen die voor de EMW als geheel geschikt is, is het te overwegen om de voorkeur te geven aan een uniforme aanpak, die wellicht niet overal optimaal is, maar wel overal hetzelfde is en ook duidelijk uit te leggen is naar gebruikers. Een overweging daarbij is dat een statistisch gezien optimale trend niet per se een makkelijk te interpreteren trend hoeft te zijn. Een voorkeur kan zijn dat trends een bepaalde mate van gladheid vertonen, omdat dit aansluit bij wat gebruikers mogelijk verwachten van een trend.
- De reden achter verschillende uitspraken voor dezelfde data kan liggen aan de precieze dataselectie en datavoorbewerking die bij de instituten anders ingevuld worden. Zo'n uitgebreidere beschrijving kan leiden tot een soort van sjabloon waarmee tijdens de volgende EMW voor elke relevante meetreeks vastgelegd kan worden hoe de dataselectie en datavoorbewerking gedaan is door desbetreffend instituut. Dit betekent dat de meta-data horend bij de gebruikte meetreeksen beschreven dient te worden. Dit bevat o.a. informatie over: (i) welke dataset gebruikt is, (ii) wat de kenmerken van de meetreeksen zijn, (iii) welke datavoorbewerking gedaan is, (iv) welke statistische methode gebruikt is (en waarom), en (v) welke nabewerking gedaan is.
- De complexiteit en de te gebruiken software van de methode kan ook nog een rol spelen. Sommige methoden zijn moeilijker te implementeren of vereisen specifieke kennis.
- Hoe visualiseren we de uitkomsten van de methode in grafieken en tabellen?
- Welke onzekerheidsmarge presenteren we? Kiezen we bijvoorbeeld voor 95%-betrouwbaarheidsintervallen? De gekozen onzekerheidsmarge hangt af van het doel en van de beschikbare data. Hoe minder data, hoe lager dit percentage moet zijn.

Als suggestie voor verder onderzoek noemen we nog dat het nuttig kan zijn om eens verschillende methoden toe te passen op een aantal voorbeelddatasets, al dan niet synthetische data, om zo de verschillen tussen de methoden te kunnen onderzoeken. Ook kan het nuttig zijn om na te gaan welke methoden internationaal gebruikt worden voor dezelfde typen gegevens.

Bronnen en achtergronddocumenten

Evaluatie Meststoffenwet

PBL (2017), *Evaluatie Meststoffenwet 2016: Syntheserapport*, Planbureau voor de Leefomgeving, Den Haag.

Velthof, G.L., T. Koeijer, J.J. Schröder, M. Timmerman, A. Hooijboer, J. Rozemeijer, C. van Bruggen, A. Bleeker en P. Groenendijk (2017), *Effecten van het mestbeleid op landbouw en milieu, Beantwoording van de ex post vragen in het kader van de evaluatie van de Meststoffenwet*, Wageningen Environmental Research, Wageningen.

Methoden

PBL

Puijtenbroek, P.J.T.M van, P. Cleij, H. Visser (2010), *Nutriënten in het Nederlandse zoete oppervlaktewater: toestand en trends*, Planbureau voor de Leefomgeving (PBL), Den Haag/Bilthoven.

Visser, H. (2004), *Estimation and detection of flexible trends*, Atmospheric Environment 38: 4135-4145.

Visser, H., S. Dangendorf en A.C. Petersen (2015), *A review of trend models applied to sea level data with reference to the "acceleration-deceleration debate"*, J. Geophys. Res. Oceans, 120, 3873–3895, doi:10.1002/2015JC010716.

Visser, H., S. Dangendorf, D.P. van Vuuren, B. Bregman en A.C. Petersen (2018), *Signal detection in global mean temperatures after "Paris": an uncertainty and sensitivity analysis*, Clim. Past, 14, 139–155, 2018 <https://doi.org/10.5194/cp-14-139-2018>.

RIVM

Boumans, L., D. Fraters (2017), *Actualisering van de trendmodellering van gemeten nitraatconcentraties bij landbouwbedrijven; Landelijk Meetnet effecten Mestbeleid*, RIVM rapport 2016-0211.

Claessens, J. (2018), *Notitie voor Landelijke Werkgroep Grondwater- Trendanalyse KRW*, RIVM, 15 november 2018.

Lukács, S., S. Wuijts, H.F.R. Reijnders, J.A. Ferreira, H.H.J. Dik en L.J.M. Boumans (2014), *Trendanalyse van kwaliteit grondwater in drinkwaterwinningsgebieden (2000-2009)*, RIVM Briefrapport 607402011/2014

Claessens, J., H.F.R. Reijnders, J.A. Ferreira, H.H.J. Dik (2014), *Trendanalyse van kwaliteit van grondwater in drinkwaterwinningsgebieden (2000-2012)*, RIVM Briefrapport 607402012/2014

Fraters, B., A.E.J. Hooijboer, A. Vrijhoef, J. Claessens, M.C. Kotte, G.B.J. Rijs, A.I.M. Denneman, C. van Bruggen, C.H.G. Daatselaar, H.A.L. Begeman, J.N. Bosma (2016), *Landbouwpraktijk en waterkwaliteit in Nederland; toestand (2012-2014) en trend (1992-2014)*, Resultaten van de monitoring voor de Nitraatrichtlijn, RIVM Rapport 2016-0076.

Deltares

Klein, J. J. Rozemeijer, H.P. Broers, B. van der Grift (2012), *Meetnet Nutriënten Landbouw Specifiek Oppervlaktewater Deelrapport B: Toestand en trends*, Bijdrage aan de Evaluatie Meststoffenwet 2012, Deltares

Klein, J. & J. Rozemeijer (2016), Update LOWESS-grafieken MNLSO-locaties t/m 2015. Deltares Memo 1 augustus, 2016.

Rozemeijer, J.C., J. Klein, H.P. Broers, T.P. van Tol-Leenders, B. van der Grift (2014), *Water quality status and trends in agriculture-dominated headwaters; a national monitoring network for assessing the effectiveness of national and European manure legislation in the Netherlands*, Environ Monit Assess 186:8981-8995.

Unie van Waterschappen (2015) Factsheet meetnet nutriënten landbouw specifiek oppervlaktewater (MNLSO), gedownload van <https://www.uvw.nl/wp-content/uploads/2015/10/Meetnet-nutri%C3%ABnten-landbouw-15-10-factsheet.pdf>

WUR

Knotters, M., D.J.J. Walvoort, T.P. Leenders (2007), *Een statistische analyse van de invloed van het mestbeleid op de oppervlaktewaterkwaliteit*, Alterra-rapport 540, Alterra, Wageningen. <https://library.wur.nl/WebQuery/wurpubs/fulltext/18753>

Walvoort, D.J.J. en W.M.G.M. van Loon (2017), *EDATA: Exploratory Data Analysis and non-parametric Trend Analysis*; Wageningen Environmental Research, Rijkswaterstaat Ministerie van Infrastructuur en Milieu (IenM). <http://publicaties.minienm.nl/documenten/edata-exploratory-data-analysis-and-non-parametric-trend-analysis>

Overig

Durbin, J. en S. J. Koopman (2012), *Time series analysis by state space methods*, Oxford University Press.

Chandler, R.E. and E.M. Scott (2011), *Statistical methods for trend detection and analysis*, Wiley & Sons Statistics in Practice

IenM, Ministerie van (2013), *Draaiboek monitoring grondwater KRW*.

Isaaks E.H. en R.M. Srivastava, 1989. *An introduction to applied geostatistics*. New York. 561 pag.

Pijpers, F.P. (2018), *Nowcasting using linear time series filters*, CBS discussion paper

Pijpers, F.P. en E. Perruci (2017), *Filtering in the Fourier domain: a new set of filters for seasonal adjustment of time series and its evaluation*, CBS discussion paper.

Van der Laan, D.J. , E. de Jonge en J. Solcer (2015), *Effect of displaying uncertainty in line and bar charts, Presentation and interpretation*, in: Proceedings of the 6th International Conference on Information Visualization Theory and Applications (IVAPP-2015), pages 225-232

Lijst van afkortingen

AIC	Akaike Information Criterion
ARIMA	Autoregressive Integrated Moving Average
BIC	Bayesian Information Criterion
BIN	Bedrijfsinformatienet
CBS	Centraal Bureau voor de Statistiek
EMW	Evaluatie Meststoffenwet
EZ	Economische Zaken
IenW	Infrastructuur en Waterstaat
IenM	Infrastructuur en Milieu
IRW	Integrated Random Walk
LMG	Landelijk Meetnet Grondwaterkwaliteit
LMM	Landelijk Meetnet effecten Mestbeleid
LNV	Landbouw, Natuur en Voedselkwaliteit
LOWESS	Locally Weighted Scatterplot Smoothing
MA	Moving Average
MK	Mann-Kendall
ML	Maximum Likelihood
MNLSO	Meetnet Nutriënten Landbouw Specifiek Oppervlaktewater
MW	Meststoffenwet
N	Stikstof
OLS	Ordinary least squares
P	Fosfor
PBL	Planbureau voor de leefomgeving
PMG	Provinciale Meetnetten Grondwaterkwaliteit
RIVM	Rijksinstituut voor Volksgezondheid en Milieu
SARIMA	Seasonal Autoregressive Integrated Moving Average
SMK	Seasonal Mann-Kendall
STM	Structural time series model
WENR	Wageningen Environmental Research
WEER	Wageningen Economic Research
WUR	Wageningen University & Research