

A network diagram consisting of various sized light blue circles connected by thin white lines, set against a solid blue background. The circles vary in size and are scattered across the page, with some larger circles and some smaller ones. The lines connect these circles in a non-uniform, web-like pattern.

KWR 2021.045 | Mei 2021

**Actualisatie
regressiemodel voor
chlorideconcentraties
bij Lobith en Eijsden**

Actualisatie regressiemodel voor chlorideconcentraties bij Lobith en Eijsden

KWR 2021.045 | Mei 2021

Opdrachtnummer(s)

403334/001/001, 403772/001/001

Projectmanager

Astrid Reus, MSc.

Opdrachtgever

Rijkswaterstaat (RWS)

Auteur(s)

Dr. Tessa Pronk

Kwaliteitsborger(s)

Dr. Thomas ter Laak

Dr. Jelmer Nijp

Begeleidingscommissie

Meinte Blaas (RWS WV), Roel Burgers (RWS VWM), Gerard Stroomberg (RIWA-Rijn), Rozemarijn Neefjes (RIWA-Rijn), Vincent Beijk (WMCN LCW), Sandra Mol (DGWB), Erik Verhofstad (DGWB), Mirja Baneke (VEWIN)

Dit rapport is een herziene versie van KWR rapport 2020.088. In de huidige versie zijn de analyses voor de korte tijdsreeks (2007-2018) opnieuw gedaan met hoogfrequente (dagelijks) en recente data (2009-2020). Dit geeft een meer robuuste en actuele afleiding van de coëfficiënten voor het geactualiseerde regressiemodel. Rapport 2020.088 vervalt hierbij. Deze versie van het rapport is verstrekt aan de opdrachtgevers van het adviesproject. Delen van dit rapport mogen worden gebruikt of gepubliceerd mits er voldoende context wordt gegeven en met bronvermelding.

N.B. De opdrachtgever heeft op 15 december 2020 aan KWR kenbaar gemaakt dat dit rapport openbaar is en derhalve ook beschikbaar gemaakt mag worden via de online KWR bibliotheek.

Trefwoorden

chloride, chloridevracht, rivierafvoer, regressie, model, verwachting

Jaar van publicatie
2021

Meer informatie

Tessa Pronk
T +3130 606 96 81
E Tessa.Pronk@kwrwater.nl

PO Box 1072
3430 BB Nieuwegein
The Netherlands

T +31 (0)30 60 69 511
F +31 (0)30 60 61 165
E info@kwrwater.nl
I www.kwrwater.nl

The logo for KWR (Kwaliteitswater Rijkswaterstaat) features the letters 'KWR' in a bold, blue, sans-serif font. The 'K' and 'R' are connected at the top, and the 'W' is positioned between them.

Samenvatting

In dit rapport is een vernieuwd regressiemodel opgesteld om de chlorideconcentratie van het oppervlaktewater te voorspellen voor de locaties Eijsden en Lobith. Het regressiemodel levert een beschrijving van de historisch opgetreden variaties in chloride-vrachten en -concentraties als functie van diverse factoren waaronder rivierafvoer. Het regressiemodel is te gebruiken om operationele verwachtingen op te stellen. Ook is het te gebruiken bij scenario-berekeningen, bijvoorbeeld in het Nationaal Watermodel ten behoeve van het Deltaprogramma Zoetwater. In tegenstelling tot eerdere studies is naast de verwachtingswaarde ook de onzekerheid in deze waarde gekwantificeerd. Inzicht in de waarde en onzekerheid van de chlorideconcentratie is noodzakelijk om af te kunnen wegen of het noodzakelijk is om bij mogelijke overschrijdingen van de norm maatregelen te treffen.

Voor het afleiden van dit regressiemodel is de relatie tussen de chloridevracht en verschillende factoren onderzocht. De in deze studie gevolgde aanpak beschrijft de relatie tussen de totale chloridevracht en deze factoren in een functie. Dit is anders dan het merendeel van de gangbare relaties, die uitgaan van het concept van een vaste achtergrondconcentratie en basisvracht. Uiteindelijk is een model gekozen waarin de voorspelling zo goed mogelijk is, maar zo min mogelijk extra factoren nodig zijn die zelf ook weer voorspeld moeten worden. Voor het voorgestelde model is alleen de voorspelling van de waterafvoer en de datum van de voorspelling nodig. Op deze manier is er een goede balans tussen praktische uitvoerbaarheid en de nauwkeurigheid van de voorspelde chloridevracht en -concentratie. Dit model is gevalideerd met een historische reeks aan meetgegevens vanaf 1972 (voor Lobith) en vanaf 1977 (voor Eijsden), en uiteindelijk gefit op een meer recente, hoogfrequenter (dagelijkse) reeks (2009-2020). Het model kan voor Lobith de chlorideconcentratie gemiddeld met 90% nauwkeurigheid voorspellen; voor Eijsden is dit met 80% nauwkeurigheid.

Het nieuwe regressiemodel geeft een verbeterde voorspelling ten opzichte van andere modellen voor locaties Lobith en Eijsden. Extreem hoge en lage chlorideconcentraties blijven minder nauwkeurig te voorspellen dan meer gemiddelde waarden. Vooral hoge chlorideconcentraties lijken niet altijd afhankelijk van de waterafvoer te zijn. De toevoeging van een gekwantificeerde onzekerheid in de voorspelling geeft juist ook daarom nuttige handvatten voor beheerders om kansen op overschrijdingen te evalueren bij het opstellen van verwachtingen van concentratie op basis van operationele afvoerverwachtingen.

Inhoud

1	Inleiding	5
2	Kwaliteitseisen voor chloride	6
3	Beschrijving van chloride en waterafvoer bij Lobith en Eijsden	9
4	Bestaande modellen voor chlorideconcentratie of chloridevracht bij Eijsden en/of Lobith	13
5	Data exploratie van chloride en andere variabelen bij Lobith en Eijsden	16
6	Geactualiseerd regressiemodel voor chlorideconcentratie	21
6.1	Basismodel	21
6.2	Factoren van invloed	23
6.3	Evaluatie van het gekozen model	28
7	Verbetering regressiemodel: de voorspellingen aanpassen en spreiding bepalen	33
7.1	Voorspelling met predictie-intervallen	33
7.2	Correctiemethode 1	34
7.3	Correctiemethode 2	36
7.4	Vergelijking van Methode 1 en Methode 2	37
8	Verwachting van chlorideconcentratie en spreiding in de praktijk	38
9	Conclusies en aanbevelingen	40
9.1	Samenvatting van de resultaten van het vernieuwde regressiemodel	40
9.2	Aanbevelingen	41
10	Referenties	43
I	Bijlage I. Model Residu variatie met diverse variabelen	45
II	Bijlage II. Code voor visualisatie van chloride voorspellingen	47
III	Bijlage III. Regressiecoëfficiënten voor Vergelijking 7 en Model 7 uit Tabel 3.	49
IV	Bijlage IV. Dataverwerking van de recente, hoogfrequente data	55
V	Bijlage V. Extra figuren	61

1 Inleiding

Een verslechtering van de oppervlaktewaterkwaliteit door een te hoge chlorideconcentratie heeft nadelige gevolgen op meerdere fronten. Het heeft bijvoorbeeld een negatieve invloed op het ecologische systeem, vermindert de groei van gewassen wanneer het water wordt gebruikt voor irrigatie, leidt tot corrosie van metalen in het leidingmateriaal wanneer drinkwater wordt gemaakt en getransporteerd naar gebruikers en beïnvloed de smaak van het drinkwater. Ook kunnen kwaliteitsnormen voor drinkwater en oppervlaktewater bij een te hoge concentratie niet gehaald worden, waardoor maatregelen nodig zijn zoals het mengen met water met een lager chloridegehalte.

Het is de verwachting dat de chlorideconcentratie in oppervlaktewater in de toekomst toeneemt. Door klimaatverandering zal in de stroombekkens van de Rijn en Maas een verandering in het neerslagregime optreden (De Bruijn & Van Mazijk, 2003). Door meer piekbuien (veel regen in kortere tijd) zal er meer afspoeling zijn over de grond en minder uitspoeling door de grond. In de winter zal meer neerslag vallen en in de zomer minder. Als er in de zomer minder water is om de chloridevracht te verdunnen, neemt de chlorideconcentratie toe. Een deel van het Nederlandse drinkwater wordt gewonnen uit de Maas en de Rijn en chloride is met de huidige zuiveringstechnieken alleen tegen hoge extra kosten te verwijderen. Voor de drinkwatervoorziening is daarom een goede voorspelling van het chloride gehalte en de kans op het overschrijdingen van maximaal toegestane waarden relevant. In het laatste geval zal de waterinname namelijk tijdelijk moeten worden gestaakt (De Bruijn & Van Mazijk, 2003).

Inzicht in de variabiliteit en voorspelbaarheid van chlorideconcentratie is noodzakelijk om bij (dreigende) overschrijdingen van de norm op tijd effectieve maatregelen te kunnen treffen. Ook is een eenduidige relatie tussen verklarende variabelen en chlorideconcentratie behulpzaam bij het opstellen van klimaatscenario's. In dit rapport wordt de chlorideconcentratie gemodelleerd, met als doel een verwachting te kunnen berekenen en de onzekerheid in deze verwachting expliciet te maken. Hiermee kan de ontwikkeling van de chlorideconcentraties in het oppervlaktewater op korte en langere termijn worden voorspeld. Daarnaast kan worden ingeschat of maatregelen om de chlorideconcentratie te reduceren noodzakelijk zijn.

2 Kwaliteitseisen voor chloride

Een te hoge chlorideconcentratie heeft nadelige gevolgen op meerdere fronten. Ten eerste op de ecologische toestand van wateren. Waterplanten en dieren afhankelijk van zoetwater kunnen nadelige gevolgen ondervinden van een te hoge chlorideconcentratie. Ten tweede zijn er gevolgen voor (de productie van) drinkwater. Verhoogde chlorideconcentratie zorgt voor een verminderde levensduur van metalen pijpleidingen vanwege corrosie. Daarnaast kan een hoge chlorideconcentratie leiden tot een zoutige smaak van het drinkwater. Om deze reden zijn er standaarden vastgesteld, waaronder deze effecten niet te verwachten zijn of tot een minimum beperkt worden. Als deze standaarden voor drinkwater en oppervlaktewater worden overschreden, zijn er maatregelen nodig zoals het mengen met water met een lager chloridegehalte, of bij drinkwaterproductie overstappen op een andere bron of actief ontzilten.

Het verschil in interpretatie en toepassing van de drinkwater-gerelateerde normen in wet- en regelgeving en in de praktijk hebben in het verleden tot onduidelijkheid geleid (RIWA-Rijn Jaarrapport 2017 De Rijn, 2018). Als gevolg hiervan hebben, op voorspraak van de Beleidstafel Droogte (zie rapport: Ministerie van Infrastructuur en Waterstaat december, 2019), het Ministerie van IenW samen met drinkwaterbedrijven een handleiding chloridenormering Drinkwater(bronnen) opgesteld. Hierin is beschreven hoe de normen gehanteerd dienen te worden (zie het rapport: Handleiding Normering Chloride Drinkwater(bronnen) van Helpdesk water, 2019).

Voor de ecologie is er minder verwarring en gelden de hieronder beschreven normen (Tabel 1a). Het langdurig, chronisch maximaal toelaatbaar risico (MTR) (opgelost chloride) van 200 mg/l voor landoppervlakte-wateren is beleidsmatig vastgesteld in de Vierde Nota waterhuishouding Regeringsbeslissing (1998). De afgeleide chronische concentratie met maximaal toelaatbaar risico in water, waarbij geen nadelige effecten zijn te verwachten voor het ecosysteem (MTR_{eco}) is 94 mg/l, het niveau waarbij mogelijk ernstige effecten voor ecosystemen zijn te verwachten (ER_{eco}) is 570 mg/l (Verbruggen et al., 2008). Door Evers (2006) werd afgeleid dat voor grote rivieren, zoals de Maas en de Rijn, de ecologische toestand bij concentraties beneden de 150 mg/l zeer goed tot goed is, voor kleinere rivieren ligt deze grens voor zeer goed bij chlorideconcentraties lager dan 20-40 mg/l.

Tabel 1a. Afgeleide chloridenormen voor ecologie

Beschrijving	Waarde (mg/l)
Maximaal toelaatbaar risico oppervlaktewateren, $MTR_{opgelost}$	200
Maximaal toelaatbaar ecologisch risico in oppervlaktewater, MTR_{eco}	94
Ernstige effecten voor ecosystemen zijn te verwachten, ER_{eco}	>570
Goed tot zeer goede toestand in grote rivieren (Evers, 2006)	<150

Voor drinkwater geldt dat bij overschrijding van de jaargemiddelde norm van 150 mg/l chloride in drinkwater over één kalenderjaar een drinkwaterbedrijf maatregelen moet treffen om aan de wettelijke eisen te voldoen. Deze wettelijke eisen zijn vastgelegd in het Drinkwaterbesluit en de Drinkwaterregeling. De norm van 150 mg/l voor in te nemen oppervlaktewater, heeft eveneens betrekking op het gemiddelde in een kalenderjaar. Het probleem hierbij is echter dat bij een tijdelijke overschrijding van de waarde van 150 mg/l nog niet bekend is of het jaargemiddelde voor het hele kalenderjaar zal worden overschreden. Voor een gerapporteerde overschrijding van de jaargemiddelde chloridenorm in oppervlaktewater (voor de drinkwaterbereiding) is ontheffing van het ILT mogelijk. Dit is achteraf echter niet zinvol en kan van te voren niet worden verleend.

Mede daarom is het belangrijk om te weten wanneer de chlorideconcentraties zo hoog zijn dat het jaargemiddelde overschreden zou kunnen worden. Daarom zijn er maximale daggemiddelden vastgesteld (Tabel 1b).

Tabel 1b. Maximale drinkwater-gerelateerde daggemiddelde chlorideconcentraties en het maximale jaargemiddelde

Beschrijving	Waarde (mg/l)
Daggemiddelde chloride geleverd drinkwater	250
Daggemiddelde chloride oppervlaktewater innamepunt	200
Jaargemiddelde chloride geleverd drinkwater	150
Jaargemiddelde chloride oppervlaktewater innamepunt	150

Uitgangspunt bij de vastgestelde maximale daggemiddelden is dat voldaan wordt aan de jaargemiddelde norm van 150 mg/l die zowel voor drinkwater als in te nemen oppervlaktewater geldt. De afspraak is om in actie te komen als de maximum daggemiddelde waarde op het innamepunt de waarde van 200 mg/l overschrijdt. Deze waarde sluit qua hoogte aan bij de waarden die in het Infiltratiebesluit en Rijnverdrag worden gehanteerd. In het Rijnverdrag (1976) en protocol (1991) is een maximale (dag)waarde van 200 mg/l chloride in oppervlaktewater aan de Duits-Nederlandse landsgrens (meetstations Lobith-Bimmen) overeengekomen.

Voor drinkwater zelf wordt een maximum daggemiddelde waarde van 250 mg/l voor chloride voorgesteld, waarbij het uitgangspunt is dat het drinkwater niet agressief en corrosief mag zijn. Hiermee wordt aangesloten bij de richtwaarden en normen die internationaal worden gehanteerd. De WHO adviseert voor chloride een grenswaarde van 250 mg/l in drinkwater, gebaseerd op smaakeffecten die boven deze concentratie kunnen optreden. Beneden deze concentratie treden geen negatieve gezondheidseffecten op. Grenswaarden waarboven wél gezondheidseffecten op kunnen treden, zijn overigens niet afgeleid voor drinkwater.

Als uit de prognoses of de monitoring blijkt dat de overschrijding van 200 mg/l op het inname-punt tot gevolg heeft dat in drinkwater de waarde van 250 mg/l overschreden kan gaan worden, bepalen drinkwaterbedrijf en waterbeheerder in onderling overleg welke maatregelen worden genomen om de toename van chloride te beperken en het jaargemiddelde onder de 150 mg/l te houden. Als een overschrijding van 250 mg/l in het drinkwater langer dan 30 dagen dreigt te gaan duren, treedt het drinkwaterbedrijf in overleg met de toezichthouder en waterbeheerder over te verrichten onderzoek en te nemen maatregelen. Bij een langdurige verhoging in drinkwater boven 250 mg/l is het aannemelijk dat sprake is van nadelige gevolgen (zie rapport Helpdesk Water, 2019 en Tabel 2).

Over het algemeen zal een kleine overschrijding van de norm maximaal enkele weken beperkte effecten hebben. Als overschrijdingen groter zijn en langer duren, is er mogelijk nader onderzoek nodig. Hieronder in Tabel 2, overgenomen uit van Vossen- van den Berg et al. (2019), staan de belangrijkste potentiële risico's van tijdelijke overschrijding van de huidige norm weergegeven. In Tabel 2 wordt uitgegaan van een overschrijding van de chlorideconcentratie norm in geleverd drinkwater tot maximaal 250 mg/l voor een periode van korter dan 1 maand, tussen 1 en 2 maanden en langer dan 2 maanden.

Tabel 2. Potentiele risico's bij overschrijdingen van de huidige norm (250 mg/l) op verschillende eindpunten (overgenomen uit van Vossen- van den Berg et al. (2019))

Type	Beschrijving	<1 maand	1-2 maanden	>2 maanden
Volksgezondheid*	Effecten gezondheid nierpatiënten	Beperkt	Effect	Effect
	Gevolgen natrium beperkt dieet	Beperkt	(Mogelijk) Effect	(Mogelijk) Effect
Bedrijfstechnisch (intern en extern)	Gevolgen voor dieet zuigelingen	Onbekend	Onbekend	Onbekend
	Toename corrosie metalen	Beperkt	Beperkt	Effect
	Hoger chemicaliënverbruik	Effect	Effect	Effect
	Hoger energieverbruik	Effect	Effect	Effect
	Aanpassingen installaties	Geen	Beperkt	Effect
Organoleptisch Wettelijk kader	Bruinwater klachten	Onbekend	Onbekend	Onbekend
	Verkorte levensduur materialen	Geen	Geen	Beperkt
	Smaakklachten	Beperkt	Beperkt	Beperkt
	Overschrijding normen Drinkwaterbesluit	Effect	Effect	Effect
	Overschrijding normen infiltratiebesluit	Effect	Effect	Effect
	Overschrijding normen KRW	Effect	Effect	Effect
Overig	Overschrijding normen Besluit Kwaliteit en monitoring water	Effect	Effect	Effect
	Verziltting bekkens	Geen/ Beperkt**	Effect	Effect
	Natuurschade duingebieden	Beperkt	Effect***	Effect
	Productieverlies vee	Beperkt	Effect	Effect
	Productieverlies gewassen	Effect****	Effect****	Effect****
	Productsamenstelling industrie	Effect	Effect	Effect

*afgestemd met RIVM, **Afhankelijk van omvang bekkens en mogelijkheid om inname water te ontwijken, ***Afhankelijk van lokale ecosystemen, ****gewasafhankelijk, aanwezigheid van ontzilting en groeifase afhankelijk

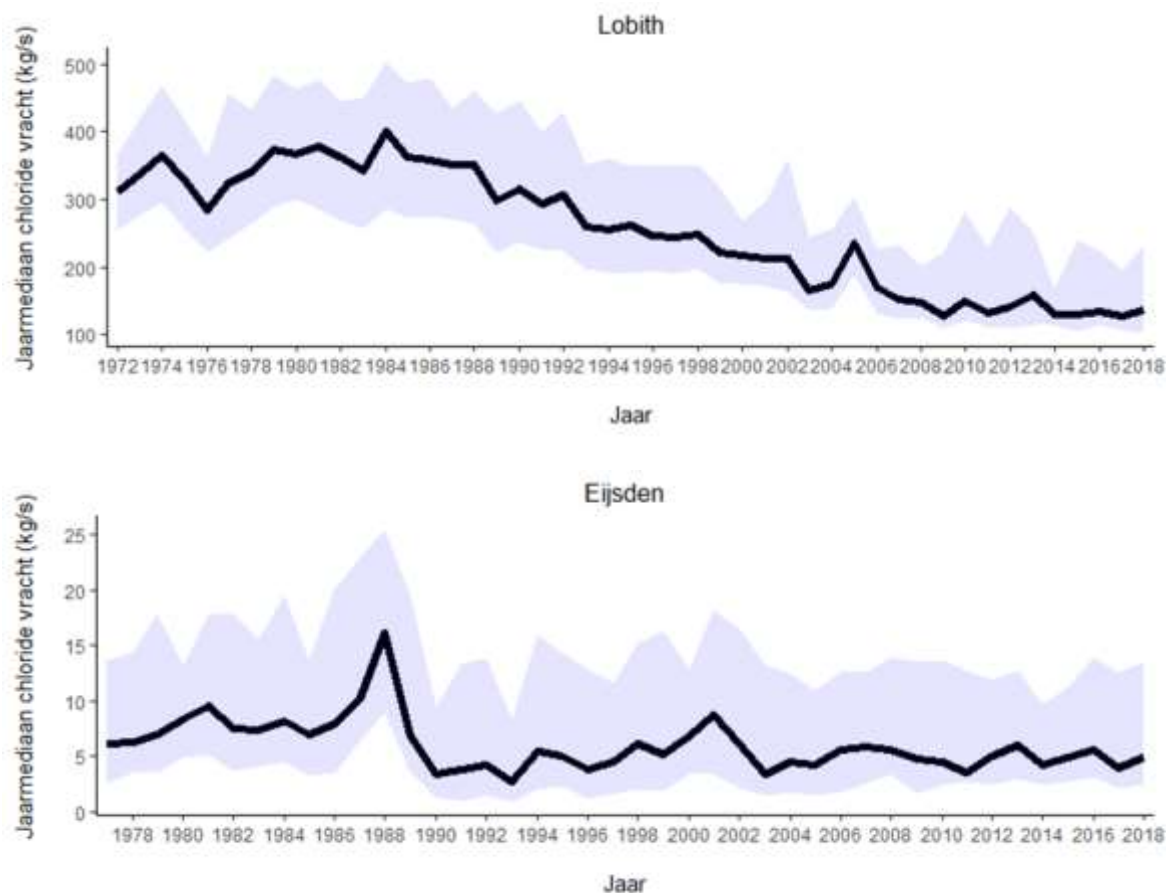
3 Beschrijving van chloride en waterafvoer bij Lobith en Eijsden

Het water van de Rijn en de Maas komt Nederland binnen bij respectievelijk de locaties Lobith en Eijsden. Deze locaties kennen een uitgebreid meetprogramma waarin ook chloride en waterafvoer worden gemeten. We bekijken de lange termijn relatie aan de hand van een historische reeks. Deze metingen gaan terug tot 1972 voor Lobith en tot 1977 voor Eijsden. Aanvankelijk wordt er elke dag gemeten, en meer recentelijk wekelijks (Eijsden) of tweewekelijks (Lobith) (bron: RIWA-Rijn database, Nieuwegein).

Chloride komt van nature in mineraalhoudende gesteenten voor, vooral in de vorm van zouten. Deze zouten zijn zeer goed oplosbaar in water en daardoor mobiel (WHO, 1996). Door afstroming van neerslag langs (afspoeling) en door het gesteente (uitspoeling) lost een hoeveelheid zout op dat vervolgens wordt afgevoerd via de rivier. Een ander deel van de natuurlijke zoutbelasting komt door atmosferische depositie. Door deze processen is er altijd chloride aanwezig in rivierwater.

Daarnaast draagt de mens bij aan de zoutbelasting door bemesting, scheepvaart, afvalwaterlozingen, gladheidsbestrijding bij vorst (voor een mogelijke invloed van deze activiteit, zie Figuur 27 in Bijlage V), mijnbouw en andere industriële activiteiten (Zwolsman & Bokhoven, 2007; Bonte, 2009; De Bruijn & Van Mazijk, 2003). Voor de Maas is de door mensen veroorzaakte chloride component waarschijnlijk laag ten opzichte van de natuurlijke vracht. Voor de Rijn heeft de industrialisatie geleid tot hogere zoutvrachten. Sinds het einde van de 19e eeuw is de natuurlijke zoutvracht voor de Rijn zelfs maar een klein deel van de totale belasting. De zoutvracht neemt in de Rijn sinds 1895 toe (Baggelaar & Van der Meulen, 2009). De Bruijn & Van Mazijk (2003) stellen dat de jaargemiddelde vracht van 50 kg/s, die vóór 1895 gold, een goede schatting is van de natuurlijke zoutvracht bij Lobith. Deze vracht van 50 kg/s levert met het langjarig gemiddelde van de afvoer van 2200 m³/s een chlorideconcentratie van 23 mg/l op. Baggelaar & Van der Meulen (2009) schatten dit als respectievelijk 40 kg/s en 19 mg/l. Sinds ongeveer 1950 lozen De Bruijnkoolmijnen bij Keulen en de kalimijnen bij Mulhouse grote hoeveelheden zout. Daarnaast is de zware industrie in het Ruhrgebied zout gaan lozen. In eerste instantie werd dit zout eenvoudig geloosd in de Rijn, maar door onder andere de Nederlandse milieu- en drinkwaterlobby is dit in het Rijnzoutverdrag verboden, waardoor de vervuilingbron van de kalimijnen is weggefallen (De Bruijn & Van Mazijk, 2003).

In Figuur 1 is de ontwikkeling van de jaargemiddelde chloridevracht uitgezet van de 70-er jaren tot en met 2018. De recente gemiddelde vracht van de Rijn bij Lobith is ongeveer 130 kg/s (Figuur 1a). Voor de Maas bij Eijsden is de chloridevracht tot 1988 iets hoger dan daarna (Figuur 1a). Dit lijkt echter toe te schrijven aan de hogere waterafvoer tussen 1977 en 1988 (Figuur 2b). De gemiddelde chloridevracht bij Eijsden is rond de 5 kg/s.

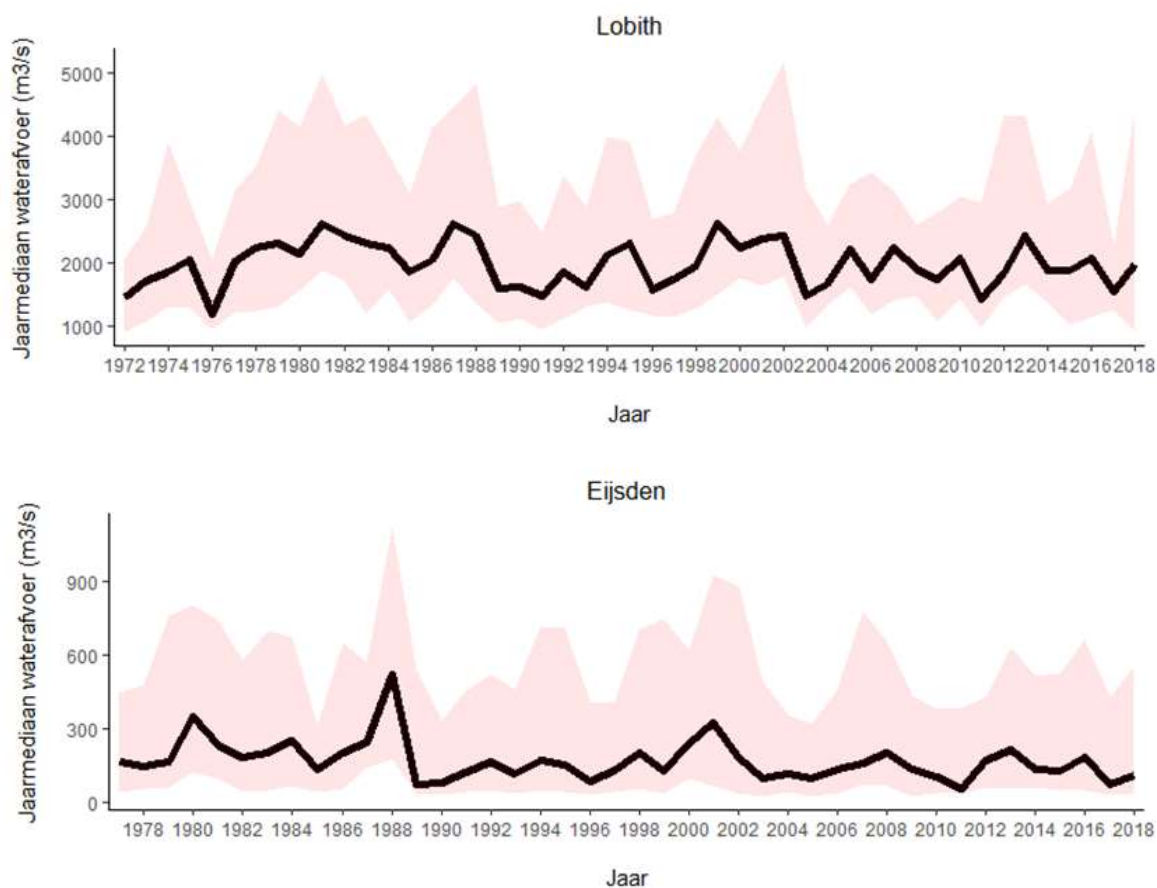


Figuur 1. Jaarmediaan chloridevracht van a. 1972 (Lobith, boven) en b. 1977 (Eijsden, onder) tot en met 2018. De blauwe band laat het 90 percentiel van alle metingen in het betreffende jaar zien en de zwarte lijn weerspiegelt de mediaan. Bron: RIWA-Rijn database, Nieuwegein.

Tot begin jaren 90 van de vorige eeuw zien we vooral bij Lobith een grotere spreiding van de chlorideconcentratie aan weerszijden van de mediaan. In de periode daarna neemt het zoutgehalte af en zijn er vooral uitschieters naar boven. Dat wijst erop dat de chlorideconcentratie recentelijk meer richting de natuurlijke belasting gaat met daarbovenop incidentele emissies.

Chloride wordt een ‘conservatieve’ stof genoemd, omdat de concentratie alleen door verdunning of emissie wordt beïnvloed en niet door biologische of fysisch-chemische processen (RIWA-Rijn Jaarrapport 2017, 2018). De concentratie hangt dus af van de hoeveelheid chloride die in de rivier komt (de vracht) door natuurlijke en menselijke invloeden, en hoe groot de verdunning daarvan is. Aangezien zoutlozingen door menselijke invloeden doorgaans voor een groot deel constant en onafhankelijk van de afvoer zijn, varieert de chlorideconcentratie met de omvang van de afvoer doordat er meer of minder verdunning van deze vracht is. De natuurlijke component van de chloridevracht is daarentegen afhankelijk van de afvoer. Bij een variërende afvoer is deze natuurlijke chlorideconcentratie juist constant doordat bij een hogere afvoer meer water over land en door de grond is afgestroomd, waarin als gevolg hiervan ook meer zout is opgelost (De Bruijn & Van Mazijk, 2003). De natuurlijke chloridevracht neemt daardoor toe bij een grotere omvang van de afvoer. Samengevat, een constante vracht levert bij een verhoogde afvoer dus een verlaging van de concentratie op, terwijl een constante concentratie bij een verhoogde afvoer een verhoging van de vracht oplevert (De Bruijn & Van Mazijk, 2003). Daarnaast wordt de chlorideconcentratie ook nog beïnvloed door incidentele lozingen of incidenten van andere aard.

In Figuur 2a en b is de ontwikkeling van de jaargemiddelde waterafvoer uitgezet tegen de tijd. De waterafvoer is een resultante van de gebeurtenissen in het stroomgebied van de rivieren. Als er veel regen valt, stroomt er veel water naar de rivier. Als er weinig regen valt, stroomt er juist weinig water naar de rivier. Als het warm is, verdamppt een deel van het water. Als het heel koud is, befrist het water deels en ligt neerslag als sneeuw op de grond, waardoor er weinig water naar de rivier afstroomt. Het smelten van sneeuw zal alleen rond de koudere maanden met sneeuwval kunnen gebeuren. Waterafvoer is dus, onder andere, een resultaat van regenval en temperatuur in het stroomgebied. Bij warm weer is er naar verwachting ook meer smeltwater van gletsjers. Door temperatuurstijging als gevolg van klimaatverandering krimpen de gletsjers in de Alpen structureel. Een deel van het water wordt door de Rijn afgevoerd, wat een verhoging van de afvoer in het algemeen oplevert. De Alpiene gletsjers zijn al zeker sinds de tweede helft van de 19e eeuw aan het krimpen. De effecten daarvan zijn nu dus al deels verdisconteerd in de rivierafvoeren (De Bruijn & Van Mazijk, 2003), alhoewel de snelheid van smelten wel toeneemt. In Stahl et al. (2016) zijn de bijdragen van smeltwater voor de Rijn bij Lobith gekwantificeerd. In de zomer is de bijdrage van gletsjerwater tot maximaal 20%. De bijdrage van smeltende sneeuw is tot enkele tientallen procenten tot maximaal 65%, vooral in januari en februari en het de bijdrage is het kleinst in de zomer. De rest van het water in de rivier bij Lobith komt van regenval. Waterafvoer kan ook variëren wanneer water actief aan de rivier onttrokken wordt of juist actief wordt aangevuld (De Bruijn & Van Mazijk, 2003).



Figuur 2. Jaarmiddelen waterafvoer van a. 1972 (Lobith, boven) en b. 1977 (Eijsden, onder) tot en met 2018. De roze band laat het 90 percentiel van alle metingen in het betreffende jaar zien en de zwarte lijn weerspiegelt de mediaan. Bron: RIWA-Rijn database, Nieuwegein.

Niet alle factoren die invloed hebben op de waterafvoer hebben ook invloed op de chlorideconcentratie. Zoals gezegd leidt meer regen tot meer afspoeling van chloride waardoor de vracht groeit maar de concentratie in het oppervlaktewater niet per sé toeneemt. Vorst zal ook de concentratie niet doen afnemen. Er is in dat geval noch afspoeling van chloride noch afspoeling van water. Onttrekken van water aan de rivier zal de concentratie niet doen

afnemen, er wordt immers zowel chloride als water onttrokken. We verwachten daardoor een min of meer constante chlorideconcentratie met waterafvoer bij meer regen, vorst of onttrekken van water.

Plotselinge hevige regenval na een lange periode van droogte kan de chlorideconcentratie wel tijdelijk verhogen, vanwege afspoeling van opgehoopte atmosferische depositie en chloride in bemesting. Meer smeltwater van gletsjers zal de chloridevracht en concentratie verlagen. In smeltwater van gletsjers zit logischerwijs minder chloride omdat deze neerslag aanvankelijk niet als water is afgestroomd via of door de grond. Smeltwater van sneeuw zal naar verwachting dezelfde weg afleggen als regen. Het is niet bekend of smeltwater meer of minder chloride bevat dan regen. Het actief toevoegen van water uit zijstromen met een andere chlorideconcentratie zal de chlorideconcentratie ook veranderen. De chloride concentratie verandert ook bij verdamping. De verdamping is weer afhankelijk van temperatuur, luchtvochtigheid, inkomende straling en windsnelheid. Doordat er minder water overblijft na verdamping terwijl de hoeveelheid zout hetzelfde blijft, zal de concentratie toenemen. Verdamping heeft invloed op zowel het water in de rivier als water dat de rivier moet bereiken. Ook neemt de concentratie toe door antropogene emissies zoals zoutlozingen uit industrie of emissies van water met een verhoogd zoutgehalte uit rioolwaterzuiveringsinstallaties (RWZI). Deze zijn niet te relateren aan de waterafvoer. In bepaalde gevallen kan de industrie wel waterafvoer-gerelateerde verhogingen van de zoutvracht veroorzaken door afspoeling van zoutrijk mijnafval en door bewust zoutrijk industrieel afvalwater te lozen bij hoge rivier afvoer. Over het algemeen verwachten we een verandering in chlorideconcentratie bij plotselinge en hevige regenval, toevoegen van water uit andere stromen, de bijdrage van gletsjerwater, verdamping en antropogene invloeden.

4 Bestaande modellen voor chlorideconcentratie of chloridevracht bij Eijsden en/of Lobith

Diverse bestaande studies belichten de relatie tussen chloride en waterafvoer specifiek voor Lobith en/of Eijsden. In dit hoofdstuk wordt een kennisoverzicht van deze studies gegeven en wordt de stand van zaken omtrent voorspelling van chlorideconcentraties beschreven. De studies gebruiken veelal de begrippen ‘achtergrondconcentratie’ en ‘basisvracht’. Rivierwater bevat van nature een minimale concentratie aan opgeloste ionen waaronder chloride afkomstig uit atmosfeer, gesteenten, bodem en afbraakproducten van organismen: de natuurlijke achtergrondconcentratie. Bovenop natuurlijke bronnen dragen antropogene bronnen (landbouw, afvalwater) structureel bij aan een verhoging van de achtergrondconcentratie tot een zgn. basisvracht.

Sjerps et al. (2016) legde relaties tussen de chlorideconcentratie C en de waterafvoer Q (een ‘QC’ relatie) vast voor Lobith en Eijsden tussen 2008 en 2012, uitgaande van een vastgestelde ‘basis’ chloridevracht die verdund wordt bij toenemende afvoer en een zogenoemde ‘achtergrondconcentratie’ (de concentratie die resulteert bij oneindig hoge afvoer). De waarden van parameters werden met lineaire regressie vastgelegd, waarbij ook met R^2 , een getal tussen de 0 (geen relatie) en 1 (perfecte verklaring), bepaald werd welk deel van de variatie van de chlorideconcentratie met het model verklaard kon worden. Voor de meetreeks bij Lobith met vastgestelde basis-chloridevracht 66,951 kg/s en achtergrond-chlorideconcentratie 42 mg/l was de R^2 0,70 na verwijdering van vijf uitbijters ($>3x$ standaardafwijking); bij Eijsden met een basisvracht van 1,955 kg/s en een achtergrondconcentratie van 21 mg/l was de R^2 0,78 (bij verwijdering van 7 uitbijters $> 3x$ standaardafwijking). De gebruikte vergelijking was

$$C_C = \frac{L_b}{Q} + C_b \quad \text{(Literatuurvergelijking type 1)}$$

Waarbij de C_C de chlorideconcentratie is, C_b de achtergrondconcentratie en L_b de vaste (basis) chloridevracht. Q is de waterafvoer.

Ook Van den Brink et al. (2019) en De Bruijn & Van Mazijk (2003) gebruiken deze relatie om de chlorideconcentratie te beschrijven, voor Lobith. Van den Brink et al. (2019) gebruikten hierbij een achtergrondconcentratie van 47 mg/l en een basisvracht van 60 kg/s. De Bruijn & Van Mazijk (2003) gebruikten een basisvracht bij Lobith van 130 kg/s, en een achtergrond chlorideconcentratie (34 mg/l). De parameters werden hier handmatig geschat, en de invloed van de Kalimijnen werd ook geschat en expliciet gemaakt. De Bruijn & Van Mazijk (2003) beschrijven dat bij lage afvoeren in dit model de fout van de concentratie ongeveer tussen de 15 en 30 mg/l ligt. De consequentie van die foutenmarges is dat de afvoer waarbij een chloride-concentratie bij Lobith van 200 mg/l optreedt binnen een marge ligt van ongeveer 690 en 900 m³/s.

Van den Boogaard & Mens (2018) beschrijven in een memo aan RWS hun model voor chlorideconcentratie te Sas van Gent. In dat model is de chlorideconcentratie een niet-lineaire relatie met waterafvoer, waarbij de exponent wordt genomen van de waterafvoer. De waterafvoer schatten ze in op basis van een gewogen gemiddelde.

$$C_C = c + a \cdot \exp\left(\sum_{s=-T}^0 w_s \cdot Q_{t+s}\right) \quad \text{(Literatuurvergelijking type 2)}$$

Hierbij is C_C de chlorideconcentratie die gemodelleerd wordt, Q de waterafvoer. De parameters c en a , en de gewichten s en w zijn via een kalibratie geschat uit de data. Met dit model kan de chlorideconcentratie op het oog goed voorspeld worden. Er is echter geen beschrijving hoe de prestatie van de voorspelling is gekwantificeerd.

Bonte (2009) maakte een model voor de verzilting van het Markermeer. Hier werd gebruik gemaakt van een 'chloridebakjesmodel' waar chloride uit verschillende locaties elkaar beïnvloeden. Vastgesteld werd dat de grootste toevoegingen van zout grotendeels overeenkomt met die aanwezig in de bronnen van water met uitzondering van directe neerslag, die een zeer geringe bijdrage levert (Bonte, 2009). Verdamping is een van de factoren die de chlorideconcentratie beïnvloedt bij de berekeningen. Voor de zoutbijdrage van de Rijn bij Lobith werd een QC relatie gebruikt (Literatuurvergelijking type 1, zie ook De Bruijn & Van Mazijk, 2003; Sjerps et al., 2016; Van den Brink et al., 2019). Voor de periode 2007-2008 leidt Bonte (2009) een basisvracht af van 60 kg/s en een achtergrondconcentratie van 52 mg/l.

In Baggelaar & Van der Meulen (2009) werd voor Lobith een chlorideconcentratie-model op maandbasis gemaakt. Dit bestaat uit een QC relatie (zie Literatuurvergelijking type 1) met een model voor de residuen, op basis van een vorige tijdstep. Het tijdreeksmodel verklaart 81,7% van de variantie (R^2 0,817) van de maandgemiddelde chlorideconcentratie.

$$C_{C,t} = C_b + \frac{L_b}{Q_t} + \frac{(1-0,19B)}{(1-0,66B) \cdot (1-0,49B_{12})} \cdot a_t \quad \text{(Literatuurvergelijking type 3)}$$

met $C_{C,t}$ de maandgemiddelde chlorideconcentratie [mg/l] voor maand t , $Q_{Lob,t}$ de maandgemiddelde afvoer (m^3/s), a_t het modelresidu (mg/l). Deze term vertegenwoordigt het stochastische deel van het model. Dit verdisconteert dat het niet door de afvoer beschreven deel van de chlorideconcentratie nog structuur vertoont in de vorm van autocorrelatie en seizoenseffecten. B is de zogenaamde tijdverschuivingsoperator ($B X_t = X_{t-1}$ en $B_{12} X_t = X_{t-12}$). Voor dit laatste is kennis nodig van de maandwaarden. De achtergrondconcentratie werd gezet op 41,3 mg/l en de basisvracht op 99,072 kg/s.

Van der Wijk (2018) bepaalde de chlorideconcentraties bij Lobith en Eijsden op basis van de formulering uit Kroon et al., '100 Jaar reeks LHM en LSM; beschrijving van de invoer'. Deze formulering wordt op het moment gebruikt om chlorideconcentraties te berekenen in het Landelijk Hydrologisch Model (LHM) en Landelijk Sobek Model (LSM) beide als onderdeel van het Nationaal Watermodel (NWM). Voor het berekenen van de chlorideconcentraties bij Lobith is de gangbare QC vergelijking conform Literatuurvergelijking type 1 gebruikt. Voor Eijsden is de volgende formulering gebruikt:

$$C_C = a \cdot Q_{Eijsden}^{-b} \quad \text{(Literatuurvergelijking type 4)}$$

Waar parameter a de waarde heeft van 350,13 en parameter b de waarde van 0,4336. De achtergrondconcentratie werd voor de QC relatie (Literatuurvergelijking type 1) voor Lobith gezet op 51,916 mg/l en de basisvracht op 63,320 kg/s. Op basis van deze regressievergelijkingen werden uitspraken gedaan over de waterafvoer waarbij overschrijding plaatsvindt. Alleen als de afvoer bij Lobith ver onder de 600 m^3/s daalt, kan de chlorideconcentratie stijgen boven de 200 mg/l. Voor Eijsden geldt volgens deze vergelijking dat zolang de afvoer boven de 10 m^3/s blijft, de achtergrondconcentratie niet boven de 150 mg/l zal stijgen. In het model voor Eijsden is niet uitgegaan van een achtergrondconcentratie, bij hele hoge verdunning gaat de chloride concentratie richting nul.

In Baggelaar & Smits (2003) is de chloridevracht van Andijk gemodelleerd, aan de hand van de vracht bij Lobith, plus de verwachte waterafvoer uit Lobith. Het chloride uit Lobith komt namelijk na enige tijd aan bij Andijk. De chloridevracht bij Lobith zelf werd gemodelleerd als een lineaire relatie met het logaritme van de waterafvoer. Twee perioden, (a) de periode januari 1976 t/m december 1986 en (b) de periode januari 1987 t/m december 2000, werden apart gemodelleerd, vanwege het geobserveerde verschil in chloridevracht. Een scherpe scheiding tussen deze perioden kan overigens niet worden aangewezen (zie Figuur 1a).

$$L_C = a + b \cdot \log(Q) \quad \text{(Literatuurvergelijking type 5)}$$

Waar L_c de chloridevracht is, a en b geschat voor de twee perioden apart, Q is de waterafvoer.

Samenvatting van het overzicht van modellen

Samengevat ligt de gehanteerde achtergrondconcentratie voor Lobith, de meest gemodelleerde locatie, tussen 34-52 mg/l en de basisvracht tussen 60-130 kg/s. Uit de literatuur blijkt dat het QC model (zie Literatuurvergelijking type 1) het meest gebruikte model is om de chlorideconcentratie te modelleren. Er is een voorkeur om chlorideconcentratie te modelleren, en daarbij de modellering van chloridevracht impliciet te laten. Een nadeel is dat er op deze manier geen rekening wordt gehouden met afvoerafhankelijke chloridevracht. De chlorideconcentratie is in QC relaties alleen afhankelijk van verdunning van een vaste chloridevracht. De chloridevracht neemt in werkelijkheid toe met toenemende waterafvoer. Baggelaar & Smits (2003) modelleren wel de chloridevracht, als functie van (log) waterafvoer (Baggelaar & Smits, 2003) (zie Literatuurvergelijking type 5). Een ander opvallend feit is dat er in geen van de studies een test wordt gedaan op data die niet zijn gebruikt om het model af te leiden ('ongeziene data') om te testen of zulke data toch goed te verklaren zijn met het model. Namelijk, als men alle data gebruikt bij het formuleren van een model kan dit model specifiek voor deze data gaan gelden waardoor andere, nieuwe data minder goed te verklaren zijn. De data zijn dan 'overfit'. Dit kan worden getoetst door het model te baseren op een deel van de beschikbare data en daarmee de resterende 'ongeziene data' te voorspellen. De ongeziene data kan een willekeurige selectie zijn van al bekende meetpunten, of bloksgewijs een periode uit de meetreeks. De modellen in de literatuur geven daarnaast geen uitgebreide beschrijving (behalve R^2) van de prestaties, bijvoorbeeld een kwantificatie van de afwijking in chloridevracht of chlorideconcentratie ten opzichte van de gemeten waarden of een andere maat voor hoe goed het model op de data past. Er wordt in het algemeen ook geen rekening gehouden met mogelijke onzekerheid (spreiding) van de voorspellingen, behalve in de studie van De Bruijn & Van Mazijk (2003).

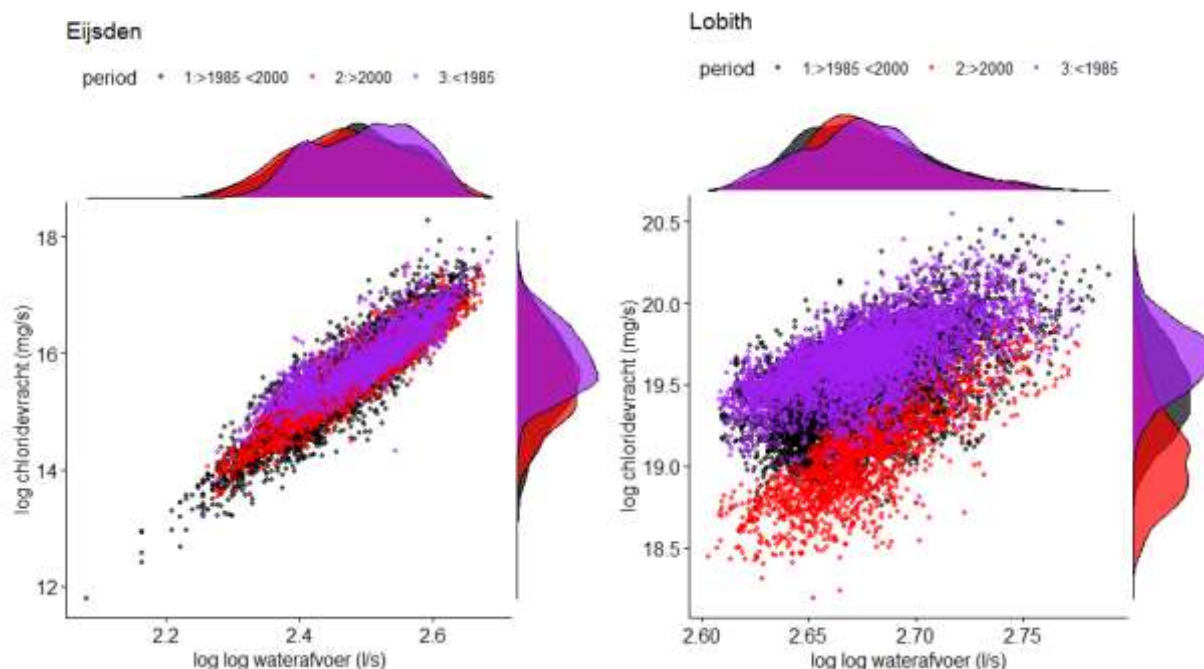
5 Data exploratie van chloride en andere variabelen bij Lobith en Eijsden

In het vorige hoofdstuk bleek dat chloride bij locaties Lobith en Eijsden tot nu toe voornamelijk is gemodelleerd aan de hand van waterafvoer als verklarende variabele. In dit hoofdstuk presenteren we naast waterafvoer mogelijk andere relevante factoren die chloride in oppervlaktewater kunnen beïnvloeden. We richten ons daarbij op het voorspellen van de chloridevracht omdat voor het voorspellen van de chlorideconcentratie beter eerst een goede voorspelling van de chloridevracht gedaan kan worden. Namelijk, in hoofdstuk 3 beargumenteren we dat deze chloridevracht zelf ook afhankelijk is van de waterafvoer. Het modelleren van de vracht biedt de mogelijkheid dit mee te nemen waarna de chlorideconcentratie op een simpele manier kan worden afgeleid door deze chloridevracht te delen door de waterafvoer.

Voor het modelleren van de chloridevracht moeten de relaties met factoren die dit mogelijk beïnvloeden in beeld gebracht worden. Relevante factoren die chloridevracht kunnen beïnvloeden zijn tijd (de vracht kan veranderen met de jaren), seizoensinvloeden, temperatuur (verdamping, smeltwater van gletsjers en sneeuw, bevroering), en direct effect van regenval (lokale afspoeling). Het meenemen van de variatie van factoren zoals industrie en socio-economische ontwikkelingen die emissies veroorzaken, valt buiten de reikwijdte van dit onderzoek maar worden impliciet meegenomen in de factoren 'jaar' en 'dag in het jaar' die trends kunnen vertonen als gevolg van deze factoren.

Variatie tussen jaren

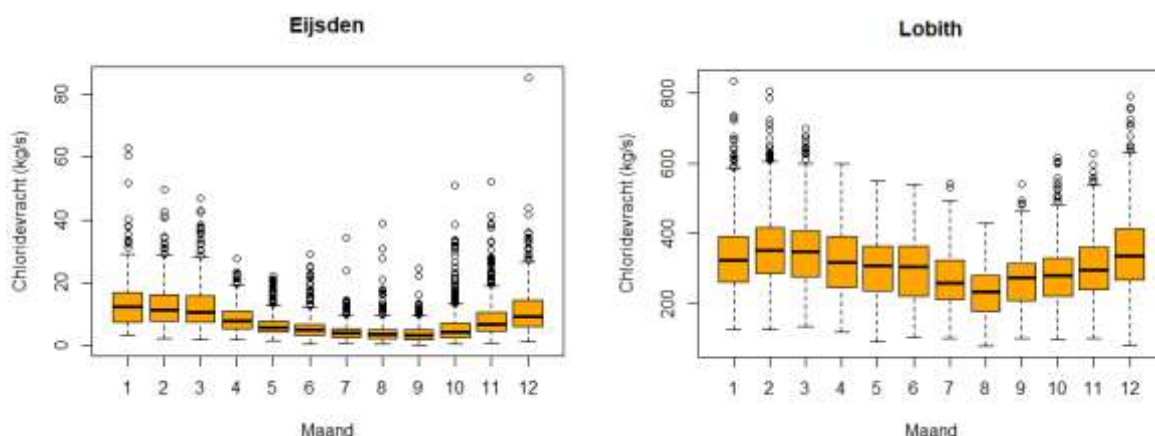
Als eerste kijken we naar de waarde van de chloridevracht door de jaren heen. In Figuur 3 hebben we de chloridevracht bij Eijsden en Lobith (bron: RIWA-Rijn database, Nieuwegein) in verschillende perioden van ongeveer gelijke lengte (1970-1984, 1985-2000, 2001-2018) uitgezet tegen de waterafvoer. We hebben hierbij chloridevracht uitgedrukt op een (natuurlijke) logschaal en waterafvoer op een dubbele (natuurlijke) logschaal, waardoor de relatie duidelijker wordt omdat deze dan visueel lineair wordt. In Figuur 3 is te zien dat er voor Eijsden geen opmerkelijk verschil is tussen de geselecteerde perioden. Voor Lobith is te zien dat vóór 1985 de chloridevracht hoger is en ook minder steil toeneemt met waterafvoer dan daarna. Dat komt waarschijnlijk doordat de chloridevracht in deze periode voornamelijk afkomstig is van emissies door industrie, die onafhankelijk zijn van de waterafvoer. Bij Eijsden is dit effect miniem.



Figuur 3. Relatie (log) chloridevracht met (log(log)) waterafvoer in drie verschillende perioden met de bijbehorende dichtheidsverdelingen. Het natuurlijke logaritme (ook wel aangegeven met 'ln') is hier genomen.

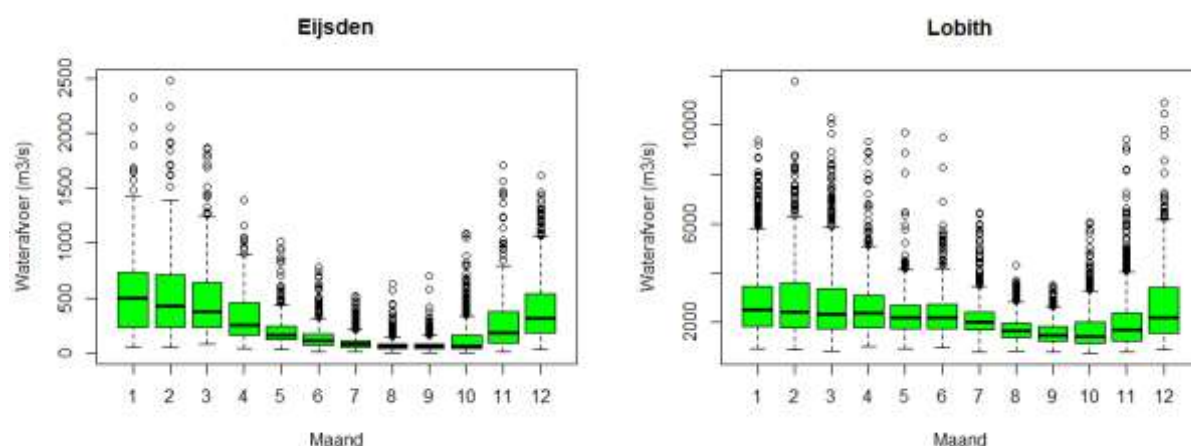
Seizoen

De tweede factor die we beschouwen is seizoen. In Figuur 4 hebben we de metingen van alle jaren weergegeven als boxplots per maand. In deze boxplots zit 50% van de meetwaarden in de box, dat zijn de meetwaarden tussen het 25 en 75 kwartiel. De lijnen die er boven- en onderuit steken bevatten alle meetwaarden die binnen maximaal 1,5 keer de lengte van de box vallen (het zogenoemde 1,5 interkwartiel-bereik). Meetwaarden die daar onder of boven liggen, zijn te zien als losse punten. In Figuur 4 is te zien dat de chloridevracht varieert met de tijd in het jaar. Dit is het meest duidelijk in Eijsden. In Lobith is de spreiding van de waarden hoger (langere boxen) doordat de vracht een grote spreiding heeft, die met de jaren geleidelijk afneemt (zie Figuur 1a).



Figuur 4. Chloridevracht hoogte en spreiding per maand in het jaar (van januari tot december). Voor dit figuur zijn alle jaren in de meetreeks per maand gecombineerd.

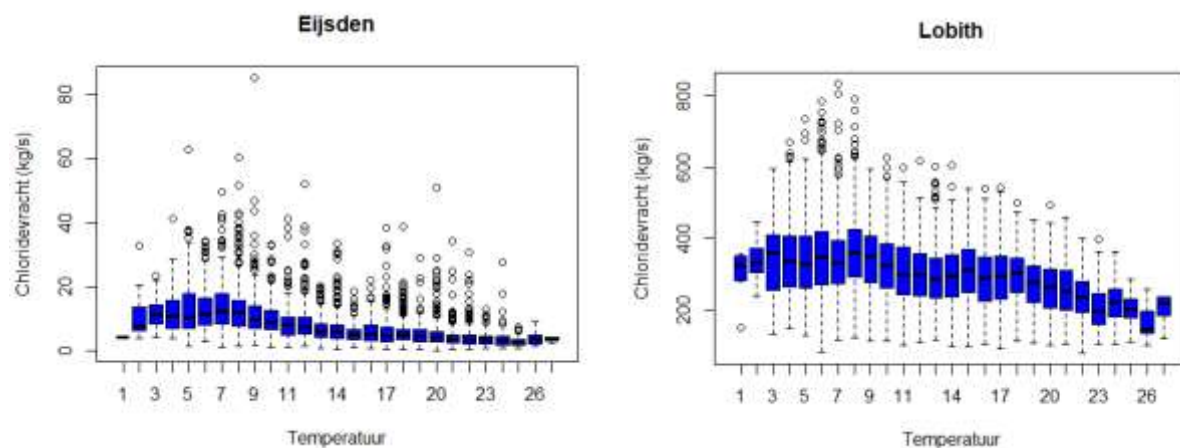
In Figuur 5 is te zien dat de waterafvoer een zelfde patroon volgt als de chloridevracht. Op het oog is de vracht dus (zoals verwacht en beargumenteerd in hoofdstuk 3) afhankelijk van de waterafvoer.



Figuur 5. Waterafvoer hoogte en spreiding per maand in het jaar (van januari tot december). Voor dit figuur zijn alle jaren in de meetreeks per maand gecombineerd.

Temperatuur

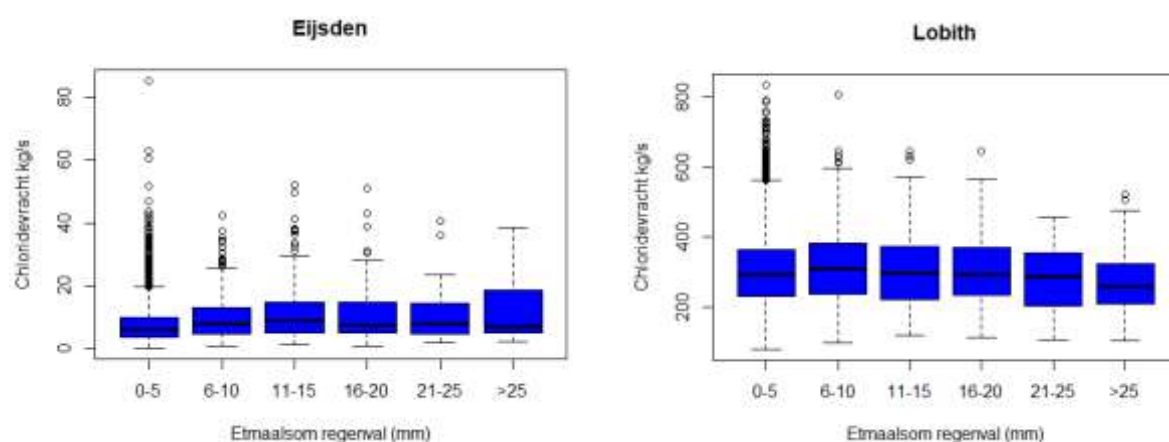
De luchttemperatuur varieert ook tussen de maanden. Vooral in Eijsden neemt de chloridevracht eerst toe, en dan af met de temperatuur (Figuur 6). Het lijkt alsof de temperatuur een structurele invloed heeft. Maar de temperatuur is het hoogst in de maanden waar ook lage waterafvoer is. Dit zijn dus overlappende factoren. Zowel temperatuur als afvoer hebben een vergelijkbaar seizoensverloop. De temperatuur heeft ook invloed op waterafvoer, dit heeft te maken met bevriezing (weinig toevoer van chloride door weinig afvoer) en met verdamping (hogere chloride concentratie, door vermindering van afvoer). In dat geval heb je met het meenemen van de waterafvoer deels ook de invloed van de temperatuur meegenomen.



Figuur 6. Chloridevracht hoogte en spreiding per temperatuurgraad (Celsius). Voor dit figuur zijn alle jaren in de meetreeks per temperatuurgraad gecombineerd.

Regenval

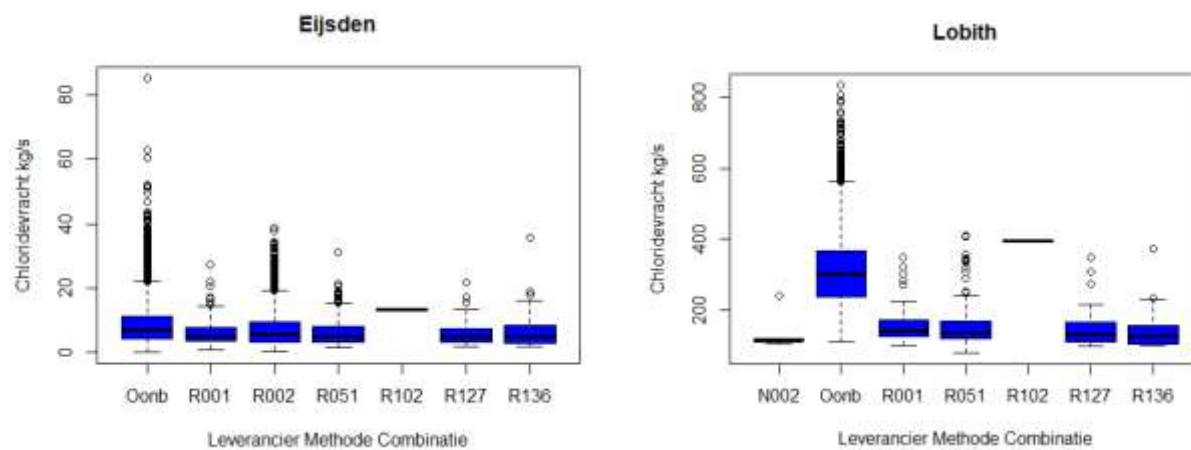
Voor het directe effect van lokale regenval op het rivierwater nemen we gegevens van het KNMI van de stations die het dichtst bij de locaties Eijsden en Lobith liggen. Dit zijn respectievelijk 'Maastricht' en 'Volkel'. Gemiddeld gezien is de chloridevracht het hoogste bij middelmatige regenval (6-15 mm) (zie Figuur 7). De chloridevracht bij lage etmaalsom (0-5 mm neerslag per dag) heeft de meeste uitschieters naar boven in zowel Lobith als Eijsden (Figuur 7). We zien deze uitschieters ook als we dezelfde plot maken voor waterafvoer (dus: etmaalsom regenval op de x-as met waterafvoer op de y-as, niet getoond). Dat duidt erop dat de incidentele hogere chloridevracht bij lage etmaalsom regenval te wijten kan zijn aan hoge waterafvoer. Mogelijk is dit het gevolg van de kans op uitschieters: meer dan 80% van de metingen ligt in de range van 0-5 mm regenval op beide locaties. Maar, stuwbeheer zou hier ook invloed kunnen hebben (Ogink & Stolker,2004).



Figuur 7. Chloridevracht hoogte en spreiding per categorie etmaalsom regen. Voor dit figuur zijn alle jaren in de meetreeks per categorie etmaalsom regen gecombineerd.

Leverancier van de data / Meetmethode

In de historische meetdata van RIWA voor Lobith en Eijsden staat ook per meting de code voor de leverancier / methode combinatie aangegeven. We analyseren deze gegevens om uit te kunnen sluiten dat resultaten beïnvloed worden door systematische meet- en analyse methoden. In Figuur 8 staan de meetwaarden chloridevracht ten opzichte van de meetmethoden. De code 'Oonb' staat voor niet geadmisteerde leverancier / methode combinatie. De andere codes staan voor unieke leverancier / methode combinaties. 'R002' en 'Oonb' bevatten de meeste metingen, en hebben daardoor ook meer kans om uitschieters op te pakken. Voor Lobith en Eijsden is naar verwachting geen effect van leverancier / methode combinatie die gebruikt kan worden voor een correctie van de chloridevracht. Omdat 'Oonb' de enige geadmisteerde code voor leverancier / methode combinatie was tot 2007 en dus vooral de periode beslaat waar de chloridevracht hoog was (zie Figuur 1a) kan de afwijking in Figuur 8 bij Lobith verklaard worden. Het zal belangrijk zijn om bij een verandering van meettechniek de resultaten te ijkten aan de resultaten van de oude techniek, zodat er geen verandering in chlorideconcentratie is bij een verandering van meettechniek. Of dit nodig was en of dit is gebeurd, is aan deze data niet af te lezen.



Figuur 8. Chloridevracht hoogte en spreiding per leverancier – methode combinatie. Codes staan voor verschillende leverancier-methode combinaties. Voor dit figuur zijn alle jaren in de meetreeks per leverancier - methode combinatie gecombineerd

6 Geactualiseerd regressiemodel voor chlorideconcentratie

6.1 Basismodel

In Figuur 9 staan de gemeten chloridevracht en de gemeten chlorideconcentratie van de historische meetreeks uitgezet tegen de waterafvoer, op een gewone (niet log-) schaal. Het is te zien dat de relatie tussen chloridevracht en waterafvoer niet lineair is. De chloridevracht neemt snel toe bij lage waterafvoer, en neemt steeds minder toe bij hogere waterafvoer. Om een vergelijking te vinden die de relatie tussen waterafvoer en chloridevracht beschrijft kan er een regressieanalyse gedaan worden. Een regressieanalyse stelt een statistisch verband vast tussen een afhankelijke en één of meer verklarende variabelen. Hierbij wordt een vergelijking opgesteld, waarvan de parameters zodanig worden aangepast dat daarbij de som van de afstand in het kwadraat van elk meetpunt met de lijn van de vergelijking het kleinst is. Een niet-lineaire regressie is lastiger in de uitvoering. Daarom willen we om praktische redenen graag een lineaire regressie doen. De exponentiële relatie tussen chloridevracht en waterafvoer is een lineaire relatie als we deze uitdrukken op een log schaal. De relatie is het rechtst bij een dubbele log van de waterafvoer (zie Figuur 3). Een dubbele log stijgt aanvankelijk nog iets sterker, wat beter bij het beeld past dat we in Figuur 9 zien bij de chloridevracht. Deze dubbele log staat in de statistiek bekend als de ‘Gumbel’ verdeling, die vaak gebruikt wordt bij het schatten van extremen. Het gebruik van de dubbele log sluit daarmee goed aan bij statistische beschrijving van tijdreeksen. In Tabel 3 is te zien dat deze relatie een iets betere fit geeft dan geen log, of een enkele log van de waterafvoer (bij vergelijk modellen, zie Tabel 3). Per jaar, en mogelijk nog bij andere variabelen, verschilt de relatie mogelijk enigszins. Daar gaan we verderop in dit hoofdstuk verder op in. Maar over het algemeen kunnen we voor beide locaties waterafvoer (Q) en chloridevracht (L_c) op log schaal beschrijven als een lineair verband:

$$\log(L_c) = a + b \cdot \log(\log(Q)) \quad \text{Vergelijking (1)}$$

L_c is de chloridevracht (mg/s), Q is de waterafvoer (l/s), a en b zijn constanten. Dit kunnen we omschrijven naar:

$$L_c = e^a \cdot \log(Q)^b \quad \text{Vergelijking (2)}$$

Dit is de vergelijking die we gebruiken als basis model. De chlorideconcentratie verkrijgen we als we de chloridevracht L_c delen door de waterafvoer. We delen daarvoor Vergelijking 2 door de waterafvoer

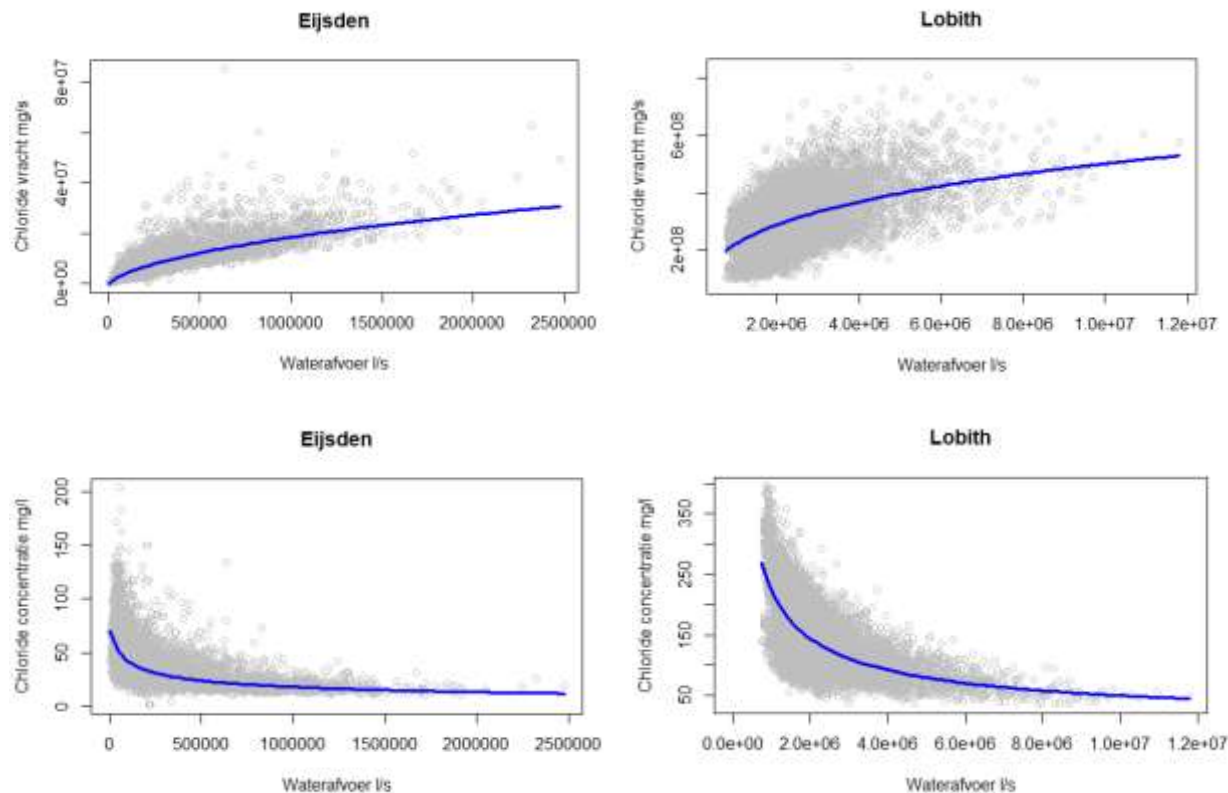
$$C_c = \frac{e^a \cdot \log(Q)^b}{Q} \quad \text{Vergelijking (3)}$$

Waarbij C_c de chlorideconcentratie is (mg/l). De chlorideconcentratie is het netto effect van twee tegengestelde effecten: de verdunning (onder de streep) en de chloridevracht die afhankelijk is van waterafvoer (boven de streep).

Om de coëfficiënten a en b voor Vergelijking 3 te schatten, doen we zoals eerder uitgelegd een lineaire regressie op de (log) chloridevracht en (log(log)) waterafvoer. Deze relatie is te zien in de figuren van Figuur 9 als de blauwe lijn.

In Figuur 9 is de relatie niet weergegeven op log schaal, maar uitgedrukt in de reguliere grootheden, na terugtransformatie.

Het is goed om te beseffen dat in dit type model de chlorideconcentratie niet oneindig blijft stijgen als de waterafvoer zeer klein wordt (wat bij het veelgebruikte model Literatuurvergelijking type 1 wel het geval is). De term onder de streep wordt voor de vergelijkingen met de geschatte waarden voor a en b voor Lobith en Eijsden voor zeer kleine afvoeren relatief groter ten opzichte van de term boven de streep (Vergelijking 3) waardoor de chlorideconcentratie vanaf een bepaald moment snel gaat dalen met nog verder afnemende waterafvoer. Dit omslagpunt ligt bij Eijsden rond 85-160 mg/l bij een waterafvoer van rond de 3000-10000 l/s (3-10 m³/s), afhankelijk van de factoren van invloed zoals jaar, maand, en andere factoren van invloed. Dit hoogste punt in chlorideconcentratie is in de praktijk mogelijk, want deze waterafvoer ligt niet ver beneden de laagste gemeten waarden van de waterafvoer bij Eijsden (zie Figuur 2b). Het potentieel voor hoge chlorideconcentraties bij locatie Eijsden is daardoor niet groot. Net zo goed kunnen incidentele emissies, en toekomstige hogere structurele emissies bij veranderde socio-economische situaties (bijvoorbeeld een nieuwe industrie die chloride gaat lozen op de Maas), toekomstige hogere structurele verdamping, en andere factoren die in hoofdstuk 5 zijn geëvalueerd zoals plotselinge regenval na droogte en seizoenseffecten, deze uiteraard wel verhogen. Bij Figuur 9 in het plot links onderaan is te zien dat de hoogste chloride concentraties inderdaad niet bij de allerlaagste waterafvoeren voorkomen, maar bij afvoeren die net iets hoger liggen. Bij Lobith ligt het omslagpunt bij een lagere waterafvoer dan bij het omslagpunt bij Eijsden. Dit omslagpunt ligt zeer ver beneden de laagst gemeten waterafvoer op deze locatie (zie Figuur 2a). Bij deze locatie is er daardoor wel een potentieel voor (zeer) hoge chlorideconcentraties, tot enkele duizenden mg/l (niet getoond) bij lagere waterstanden. Voor Lobith is het omslagpunt zelf in de praktijk onzeker omdat er historisch geen waterafvoeren zijn waargenomen die zo laag zijn (zie Figuur 2a).



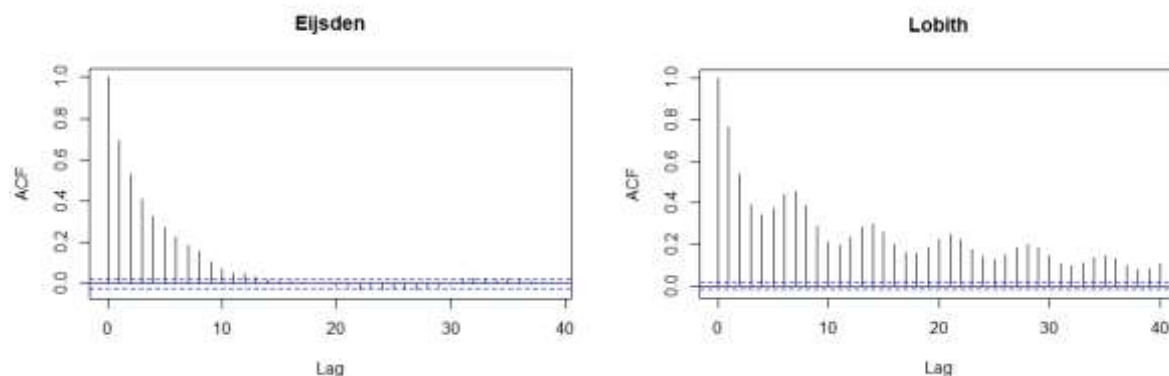
Figuur 9. Chloridevracht (bovenste plots) en chlorideconcentratie (onderste plots) uitgezet tegen waterafvoer, voor locaties Eijsden en Lobith. De blauwe lijn is het resultaat van de lineaire regressie volgens Vergelijking 3.

De chloride achtergrondconcentratie kan worden beschouwd als wat er overblijft per liter water bij hele hoge verdunning, oftewel hoge waterafvoer. Alleen de natuurlijke constante concentratie per liter blijft dan over. Om deze te bepalen voor een vergelijking met de waarden van de modellen in hoofdstuk 5 is tweemaal de maximale waterafvoer van de rivier in de periode 1972/1977 tot 2018 genomen. De chloride achtergrondconcentratie is op deze manier berekend op 27 mg/l voor Lobith en 10 mg/l bij Eijsden. De chloride achtergrondconcentratie voor Lobith ligt bij deze berekening vlak bij de schatting van 23 mg/l uit De Bruijn & Van Mazijk (2003) van vóór 1895. Het lijkt logisch dat de chloride achtergrondconcentratie, die een natuurlijke oorsprong heeft, niet veel verandert met de tijd. De waarde ligt onder de range van waarden (34-52 mg/l) voor basis-chlorideconcentratie voor Lobith in de modellen in hoofdstuk 4. Een 'basis chloridevracht' onafhankelijk van de waterafvoer, kunnen we met deze relatie niet berekenen omdat de vracht naar nul gaat als de waterafvoer ook naar nul gaat (zie Figuur 9 en 10). Als we als praktische benadering de chloridevracht waarde bij de minimale waterafvoer nemen, dan is de basis chloridevracht voor Eijsden 0 (zoals in het model van Van der Wijk, 2018) en voor Lobith ongeveer 100 kg/s (voor Lobith gebruikten we hier alleen metingen van na het jaar 2000). Dit ligt binnen de range zoals gevonden in de modellen voor Lobith in Hoofdstuk 4, de gehanteerde basisvracht voor Lobith varieert daar van 60-130 kg/s.

6.2 Factoren van invloed

Er zijn extra factoren die de regressierelatie over de gehele meetreeks beter kunnen maken. Die factoren hebben we in hoofdstuk 5 bekeken. Het ene jaar kan een hogere chloridevracht hebben dan het andere. De concentratie kan ook oplopen of afnemen binnen een jaar. Of, een industrie kan een seizoensafhankelijk productieproces hebben dat leidt tot variatie in emissies gedurende het jaar. Ook wegzout is een voorbeeld van een emissie met seizoensvariatie. Er zijn dan structurele emissies die afhankelijk zijn van de maand in het jaar. De invloed van temperatuur (verdamping en neerslagtekort) is waarschijnlijk voor een groot deel al opgevangen in de relatie in Vergelijking 3, waarbij de chloridevracht relatief hoog is bij lage afvoer (n.b. de stijging van de chloridevracht wordt minder bij toenemende waterafvoer, zie ook Figuur 9), maar kan ook onafhankelijk daarvan nog invloed hebben. Het effect van regenval in het stroomgebied is waarschijnlijk al ondervangen in de waterafvoer. Maar, lokale regenval na grote droogte zou invloed kunnen hebben door afspoeling van extra chloride van bemesting of opgehoopte atmosferische depositie.

Daarnaast kan een meting een voorspellende waarde hebben op de navolgende metingen, omdat temporele emissies of afspoeling meerdere metingen achter elkaar kunnen beïnvloeden. Dit is uiteraard afhankelijk van de meetfrequentie. Dit fenomeen, waarbij metingen met kleinere tussentijd meer op elkaar lijken, wordt temporele autocorrelatie genoemd (Hyndman & Athanasopoulos 2018). Het verband tussen opeenvolgende metingen (de temporele autocorrelatie) is weergegeven in Figuur 10. In deze 'ACF' (autocorrelation function) plot is te zien dat er na het toevoegen van de factor 'jaar' wat de variatie tussen jaren verklaart, nog autocorrelatie is, voor Eijsden tot 10 dagen terug, en bij Lobith verder terug met een cyclus van een week. Het is onbekend waardoor deze wekelijkse cycli bestaan bij Lobith, maar een mogelijke verklaring zou kunnen zijn dat in het weekend bepaalde industrieën inactief zijn. Deze cyclische autocorrelatie is het sterkst bij lage afvoeren (niet getoond).



Figuur 10. Autocorrelatie (ACF plot) voor Eijsden en Lobith, op basis van Model 3 in Tabel 3. Op de x-as staan de dagen ('Lag') waar nog autocorrelatie is. Een autocorrelatie van 0.2 wordt in het algemeen nog als bruikbaar gezien.

De factoren die we hebben geëvalueerd op hun vermogen om meer variatie in de historische chloridevrucht te verklaren staan in Tabel 3, met enkele indicatoren voor hoe goed ze de vergelijking om de chloridevrucht te verklaren verbeteren. De aangepaste (adj.) R^2 is een maat voor hoe goed de data passen op de regressielijn, gecompenseerd voor het aantal variabelen in het model. Dus door variabelen toe te voegen stijgt de adj. R^2 , mits deze bijdragen aan de verklaring van de variatie. Als de volgende variabele niet genoeg toevoegt, daalt de adj. R^2 . De adj. R^2 kan een waarde aannemen tussen de 0 en 1. Hoe hoger de adj. R^2 , hoe beter het model de variatie in de data verklaart. Akaike's Information Criterion (AIC) geeft een afweging van het aantal benodigde modelparameters en de 'fit' van het model. Bij een groot aantal voorspellende variabelen is het logisch dat de 'fit' verbetert, maar is het waarschijnlijker dat ruis in de data gefit wordt. Dit heet 'overfitten' en gaat ten koste van de kwaliteit van de voorspelling. Een lagere AIC tussen modellen die op dezelfde data worden gemaakt (ook negatief) geeft een betere balans weer tussen 'fit' en het aantal verklarende variabelen. Het AIC criterium is waardevol om 'overfitten' te voorkomen. AIC kan tussen modeltypen (bijvoorbeeld een QC relatie en de log (log) relatie) niet vergeleken worden, maar wel binnen modeltypen om de invloed van het toevoegen of weghalen van variabelen te evalueren.

Daarnaast kijken we hoe goed de chlorideconcentratie voorspeld wordt met het model als we een deel van de meetreeks buiten beschouwing laten en daarna de ontbrekende punten voorspellen. De mate waarin dit lukt, drukken we uit in MAPE, wat staat voor *mean absolute percentage error* van voorspellingen ten opzichte van de gemeten waarden. MAPE is per definitie altijd > 0 ; hoe dichterbij 0 hoe kleiner het relatieve verschil tussen voorspelling en meting.

$$MAPE = \frac{\sum_{i=1}^n \text{abs}\left(\frac{C_{ci} - C_{pi}}{C_{ci}}\right) \cdot 100}{n} \quad \text{Vergelijking (4)}$$

Hier zijn C_c de gemeten chlorideconcentratie en C_p de voorspelde chlorideconcentratie van meting i , en n is het aantal voorspellingen.

Voor het vaststellen van de MAPE doen we vijftig keer een kruis validatie (CV). Dat wil zeggen dat we vijftig keer willekeurig 10 % van de meetwaarden verwijderen en daarna deze waarden voorspellen met het model dat op de overige 90% van de data is gefit. We nemen het gemiddelde van de MAPE uit de vijftig CV's ('50 CV MAPE alles' in Tabel 3). Recenter wordt er minder vaak gemeten. Voor de waarden vanaf 2005 hanteren we dezelfde methode maar nu trekken we de te voorspellen punten alleen uit die periode ('50 CV MAPE recent' in Tabel 3). We verwijderen hiervoor 20% van deze punten in deze periode, en gebruiken de rest van de meetwaarden in de reeks voor het bepalen van de fit.

Over het algemeen zien we in Tabel 3 aan de kolommen '50 CV' dat de chlorideconcentratie bij de meer uitgebreide modellen bij Lobith beter te benaderen is dan die bij Eijsden. De factor 'jaar' (Yr) geeft vooral in Lobith de grootste verbetering aan het model. Lobith heeft dan ook de grootste lange termijn variatie in chloridevracht tussen 1972 en nu. In Lobith verandert de chloridevracht geleidelijk, ook gedurende het jaar. Daarom voegen we een interactie toe met de dagen dat een jaar heeft (Dy). Dit houdt rekening met een geleidelijke structurele verhoging of verlaging van de chloridevracht gedurende een specifiek jaar. Deze interactie is niet voor alle jaren significant, maar wel voor jaren waar ook visueel een trend is te zien binnen in het jaar (niet getoond). De factor 'maand' (Mo) geeft ook een kleine, maar wel significante verbetering van de regressie, vooral in Lobith. De factor 'weekdag' (Wd) geeft ook een verbetering, vooral bij Lobith. Alhoewel het nemen van een dubbele log niet veel verbetering geeft bij locatie Lobith of Eijsden (vergelijk Model 1 met 2) heeft het voor sommige varianten in Tabel 3 een verbetering. Het meenemen van lokale regenval (R) of temperatuur (T) geeft een verbetering. Ook de toevoeging van een factor 'tijd sinds laatste grote etmaalsom regenval (7 mm)' is getoetst (niet getoond) waarbij alle dagen met een som lager dan 7 mm waarde 'nul' kregen en dagen met waarde > 7 mm een waarde ter hoogte van het aantal nullen dat vooraf ging. Deze factor was ingebracht om afspoeling van atmosferische depositie toe te voegen door de lengte van een droge periode gevolgd door stevige neerslag mee te wegen. Dit gaf echter geen grote verbetering en is daarom niet verder meegenomen. Dit suggereert dat de afspoeling van atmosferische Cl depositie een beperkte invloed heeft óf dat de gekozen waarden en manier van uitrekenen voor de parameter niet goed gekozen zijn. Omdat factoren die zelf ook voorspeld moeten worden, zoals regenval, minder prioriteit hebben omdat ze de voorspelling compliceren, en het effect niet overduidelijk is, is dit niet verder uitgezocht.

De meting meenemen die voorafgaat aan de volgende meting (Vm) geeft ook een verbetering van het model. De voorspelling in recentere jaren (na 2005) is juist slechter (zie '50 CV MAPE recent' in Tabel 3) omdat de meetpunten hier in deze historische reeks met langere tussenpozen gedaan worden en de voorspellende waarde van opvolgende metingen verval. De meetfrequentie is dan lager dan de duur van de invloed van bepaalde factoren. Dit speelt vooral bij Lobith, waar recent slechts om de twee weken gemeten wordt. Bovendien is het meenemen van de vorige meting in de voorspelling niet in alle gevallen praktisch. Als er straks een voorspelling gedaan moet worden van bijvoorbeeld een maand vooruit, is een meting van twee weken daarvoor ook niet voorhanden. Hoewel we de temporele autocorrelatie op deze manier dus niet kunnen fitten, kunnen we de temporele autocorrelatie wel gebruiken om de voorspelling van de chlorideconcentratie te verbeteren in Hoofdstuk 7.

Tabel 3. Vergelijking van regressiemodellen. 'Goodness of fit' scores gebaseerd op concentraties. De variabele Yr staat voor jaar, Dy is dag van het jaar (1 tot en met 365), Mo is maand van het jaar (1 tot en met 12), Wd is de dag in de week (maandag tot en met zondag), T is temperatuur, Vm is de chloridevracht van de vorige meting, R is regenval categorie. Voor de Lobith modellen 00 en 01 is Yr meegefit omdat de modellen in de originele publicaties niet over de gehele periode gaan maar voor een kleinere periode gefit zijn. 'Recent' is vanaf 2005 (20% ongezien), 'alles' is de hele meetserie (10% ongezien). In groen is het uiteindelijk gekozen model aangegeven.

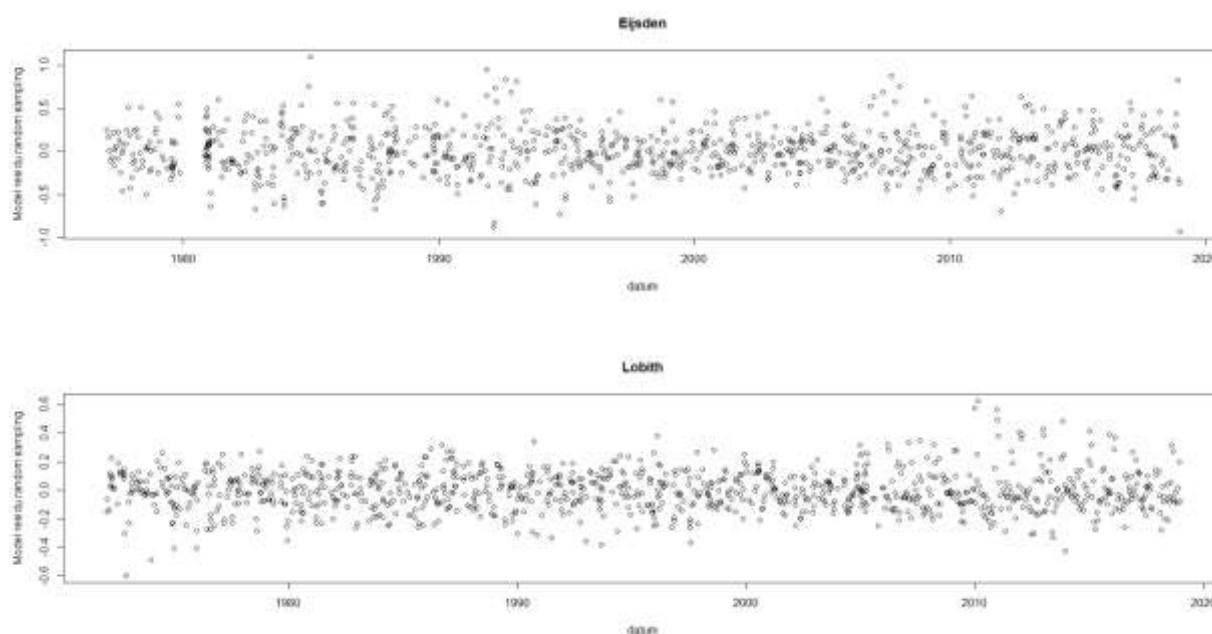
Eijsden					
Model		Adj. R ²	AIC	50 CV MAPE alles	50 CV MAPE recent
00	Lc ~ Q + Yr (Literatuurvergelijking 1, 'QC')	0,74	260873	32,9	26,9
01	Lc ~ log(Q) + Yr (Baggelaar & Smits, 2003)	0,67	260878	48,9	43,9
1	log(Lc) ~ log(Q)	0,82	4665	25,7	24,6
2	log(Lc) ~ log(log(Q))	0,82	4521	25,6	25,7
3	log(Lc) ~ log(log(Q)) + Yr	0,87	2297	21,7	21,1
4	log(Lc) ~ log(log(Q)) + Yr*Dy	0,88	1720	21,1	20,7
5	log(Lc) ~ log(log(Q)) + Yr*Dy + Mo	0,88	1345	20,5	20,3
6	log(Lc) ~ log(log(Q)) + Yr + Mo	0,87	1912	21,3	19,6
7	log(Lc) ~ log(log(Q)) + Yr*Dy + Mo + Wd	0,88	1311	20,6	19,3
8	log(Lc) ~ log(log(Q)) + Yr*Dy + T	0,88	1580	20,6	20,3
9	log(Lc) ~ log(log(Q)) + Yr*Dy + R	0,88	1495	20,6	20,5
10	log(Lc) ~ log(log(Q)) + Yr*Dy + Vm	0,88	1142	20,1	22,2
11	log(Lc) ~ log(log(Q)) + Yr*Dy + Mo + T	0,88	1250	20,4	20,5
12	log(Lc) ~ log(log(Q)) + Yr*Dy + Mo + R	0,88	1156	20,4	20,3
13	log(Lc) ~ log(log(Q)) + Yr*Dy + Mo + T + R	0,89	1067	20,2	20,1
Lobith					
Model		Adj. R ²	AIC	50 CV MAPE alles	50 CV MAPE recent
00	Lc ~ Q + Yr (Literatuurvergelijking 1, 'QC')	0,71	457223	13,1	13,7
01	Lc ~ log(Q) + Yr (Baggelaar & Smits, 2003)	0,70	456665	13,2	16,7
1	log(Lc) ~ log(Q)	0,23	4100	25,2	74,5
2	log(Lc) ~ log(log(Q))	0,23	4113	25,2	75,4
3	log(Lc) ~ log(log(Q)) + Yr	0,76	-9485	12,9	14,1
4	log(Lc) ~ log(log(Q)) + Yr*Dy	0,77	-9985	12,8	13,6
5	log(Lc) ~ log(log(Q)) + Yr*Dy + Mo	0,83	-13406	10,9	11,8
6	log(Lc) ~ log(log(Q)) + Yr + Mo	0,82	-12879	11,1	9,7
7	log(Lc) ~ log(log(Q)) + Yr*Dy + Mo + Wd	0,83	-13805	10,8	9,5
8	log(Lc) ~ log(log(Q)) + Yr*Dy + T	0,82	-12795	11,3	12,1
9	log(Lc) ~ log(log(Q)) + Yr*Dy + R	0,77	-9189	12,9	13,8
10	log(Lc) ~ log(log(Q)) + Yr*Dy + Vm	0,87	-17201	9,0	14,1
11	log(Lc) ~ log(log(Q)) + Yr*Dy + Mo + T	0,83	-13468	10,9	12,2
12	log(Lc) ~ log(log(Q)) + Yr*Dy + Mo + R	0,83	-12701	11,0	11,9
13	log(Lc) ~ log(log(Q)) + Yr*Dy + Mo + R + T	0,83	-12736	11,0	12,0

Ter vergelijking met andere bestaande typen modellen voor chloridevracht bij Lobith, de QC vergelijking (zie Literatuurvergelijking 1); en Baggelaar & Smits, 2003 (zie Hoofdstuk 4) die in Tabel 3 gefit zijn op de data, is er een lichte verbetering met betrekking tot het basis model. Dit is te zien als we in Tabel 3 voor Lobith Model 00 en 01

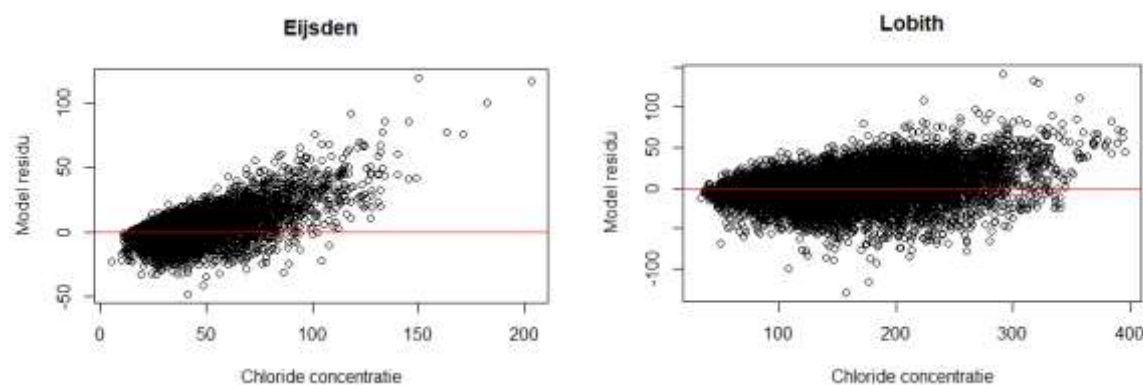
vergelijken met Model 3. Bij Eijsden is er een grote verbetering ten opzichte van deze modellen, maar hiervoor waren de modellen 00 en 01 niet ontworpen.

Model 7 in Tabel 3 lijkt voor beide locaties een relatief goede fit te hebben. Als we kijken naar de fouten die Model 7 uit Tabel 3 maakt en dit vergelijken met de gemeten waarden, willen we het liefst zien dat de fouten ('model residuen') niet meer systematisch veranderen met de verklarende variabele. Dat betekent namelijk dat alle variatie veroorzaakt door de variabele reeds in het model zit. De model residuen worden weergegeven als $\log(\text{gemeten chlorideconcentratie}) - \log(\text{voorspelde chlorideconcentratie})$. Voor temperatuur (zie Bijlage I Figuur 21), de tijd (Figuur 11), en de factor maand (zie Bijlage I Figuur 21) is dit inderdaad het geval. Omdat er recent met een lagere frequentie (7 tot 13 keer minder) wordt gemeten, zijn voor de visualisatie in Figuur 11 voor elk jaar evenveel metingen random geselecteerd, zodat voor elk jaar evenveel kans is om uitschieters te hebben. Voor de factor dag in het jaar (zie Bijlage I Figuur 21) zien we een structurele onderschatting in enkele weken van het jaar bij locatie Lobith. Mogelijk heeft dit met de vakantieperiode te maken.

Er is opvallend veel meer spreiding van chloridevracht bij een lage etmaalsom lokale regenval en lage waterafvoeren (zie Bijlage I Figuur 21). Weinig regenval is, als we dit extrapoleren naar regionale effecten, waarschijnlijk gekoppeld aan lage waterafvoeren. Een eenvoudige verklaring is dat bij een lage waterafvoer incidentele emissies een groter effect hebben vanwege de lage verdunning. Ook zou het kunnen zijn dat in de zomer, wanneer waterafvoeren laag zijn en de bodem droog is, lokale buien in het stroomgebied bijdragen aan de waterafvoer die verschilt in chlorideconcentratie door de verschillen in lokale chlorideconcentraties in of op de bodem. Het is daarnaast mogelijk dat lage waterafvoeren simpelweg zelf ook minder nauwkeurig bepaald kunnen worden. Deze waterafvoeren worden namelijk niet gemeten, maar afgeleid op basis van hydraulische condities (bv. waterstand, stuw Driel) met een 'Qf' relatie (van Vuuren et al., 1999; Ogink & Stolker, 2004), die periodiek wordt herijkt, met oog op bijvoorbeeld rivierbodemdaling.



Figuur 11. Model residuen, relatief verschil in log (gemeten chlorideconcentratie) - log (voorspelde chlorideconcentratie) van individuele metingen van de hele meetreeks. Per jaar zijn een gelijk aantal metingen random geselecteerd om de kans op uitschieters gelijk te trekken.



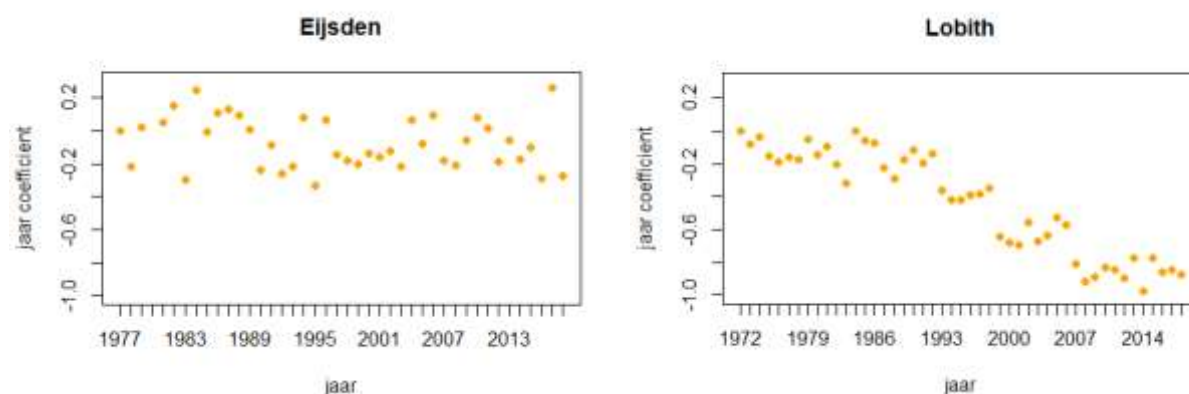
Figuur 12. Variatie in model residuen (gemeten chlorideconcentratie - voorspelde chlorideconcentratie, in mg/l) met de voorspelde chlorideconcentratie.

De model residuen (de fouten van de model waarden ten opzichte van de gemeten waarden) veranderen wel systematisch met de voorspelling van de afhankelijke variabele, de chlorideconcentratie (zie Figuur 12). Hoe hoger de chlorideconcentratie, hoe meer de residuen positief afwijken. Dit geeft aan dat het model neigt naar het onderschatten van vooral de hoge chlorideconcentraties. Dit is het duidelijkst bij locatie Eijsden. Vooral in Eijsden vindt er daarnaast juist voor lage waarden over het algemeen een overschatting plaats door het model, ten opzichte van de werkelijke gemeten waarde. Dit kan een gevolg zijn van hoe een lineair regressiemodel, dat alle meetpunten fit, aan de uiteinden van het meetdomein meer afwijkt. Deze afwijking kan een beperking opleveren ten aanzien van de hoge chlorideconcentraties, omdat juist die van belang zijn voor beheerders en gebruikers.

6.3 Evaluatie van het gekozen model

Tabel 3 geeft een overzicht van verschillende variabelen en hun verbetering aan de fit op de gehele historische meetreeks in Lobith en Eijsden. Dit zijn variabelen die het historisch verloop van de chloride concentratie mede bepaald hebben, volgens dit model. Vanuit het oogpunt van robuuste toepasbaarheid van het model willen we bij voorkeur een zo simpel mogelijk voorspellend model met zo min mogelijk variabelen die zelf ook voorspeld dienen te worden, zoals temperatuur en neerslag. Immers, als deze data niet beschikbaar is door bijvoorbeeld gebreken in meetapparatuur, en er praktische belemmeringen zijn bij het integreren van deze data in een enkel voorspellingsmodel, kan de chloride concentratie niet voorspeld worden. De waterafvoer is zo essentieel (Tabel 3) dat we deze niet buiten beschouwing kunnen laten.

De variabele 'jaar' als factor meenemen zoals in Tabel 3, is in de praktijk ook lastig. Om de coëfficiënten voor deze factor te kunnen schatten moeten er voor een deel van het jaar al metingen zijn. De coëfficiënt voor een factor vertegenwoordigt de verandering in de gemiddelde respons die verband houdt met een verandering in die factor. De jaarcoëfficiënten vertegenwoordigen het effect van dat betreffende jaar. Voor locatie Eijsden lijken jaren die dicht bij elkaar liggen niet persé méér op elkaar dan jaren die ver uiteen liggen (zie Figuur 13), zodat we de waarde van de coëfficiënt niet betrouwbaar uit het voorgaande jaar kunnen 'lenen'. In Figuur 13 zijn de jaar coëfficiënten voor Lobith en Eijsden weergegeven die zijn vastgesteld met de fit op de historische meetreeks. Of de jaar coëfficiënt voor een bepaald jaar bij Eijsden een significante parameter is (zie Bijlage III, Eijsden), lijkt een random effect. Ditzelfde geldt voor de interactieterm jaar*dag. Bij Lobith is er wel een duidelijke trend in jaar coëfficiënten over de gehele reeks (Figuur 13). Hier lijken de laatste twaalf jaar sterk op elkaar. In deze laatste twaalf jaar is er desondanks ook variatie, waarbij jaren dichtbij elkaar niet persé méér op elkaar lijken dan jaren ver uiteen. Ondanks dat de jaar*dag interactie in het regressiemodel tot een verbetering in de chloride voorspelling leiden (Tabel 3), weegt de meerwaarde van de jaar coëfficiënten en interactie met dag van het jaar niet op tegen de nadelen van de implementatie in de praktijk.

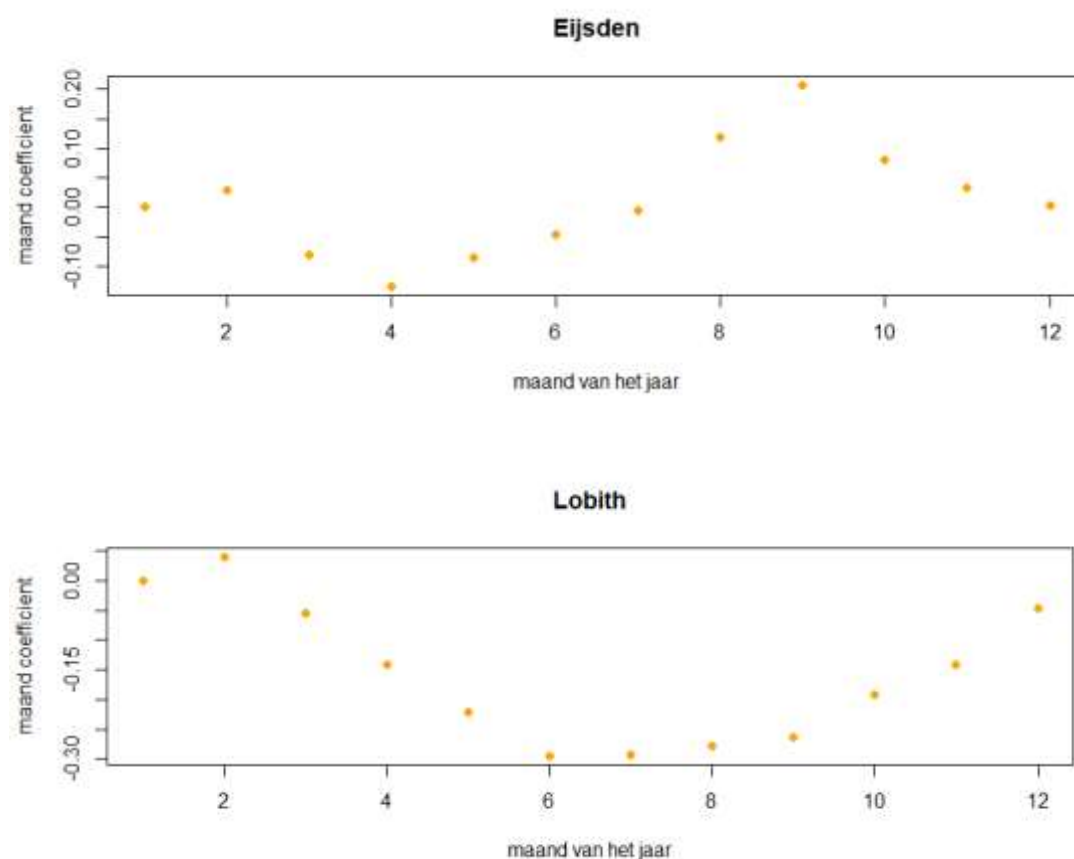


Figuur 13. De jaar coëfficiënten uit Model 7 in Tabel 3 over de gehele meetreeks van Eijsden en Lobith (zie ook Bijlage III)

Voor het voorspellingsmodel hebben we het liefst alleen variabelen die vanuit de meetgegevens zelf vast te stellen zijn. Het meest simpele model met de beste MAPE voorspellingen en relatief goede R^2 en AIC, is voor beide locaties Eijsden en Lobith Model 7 (zie Tabel 3) waar variabelen maand en weekdag in zijn meegenomen. De regressie coëfficiënten voor deze variant staan gegeven in Bijlage III. Voor de recentste 12 jaren lopen de jaarcoëfficiënten voor beide locaties niet ver uiteen (zie Figuur 13).

Vanaf hier zullen we met een andere dataset de model fit gaan bepalen. Recente (tot februari 2021) en hoogfrequente (dagelijkse) data voor Lobith en Eijsden voor debiet en chloride zijn verkregen via RWS servicedesk-data en aqualarm, met ondersteuning van Marc Philippart (RWS-WVL). Deze data zijn vrij beschikbaar via het <https://waterinfo.rws.nl> onder de knop 'meer data'. Verwerking van deze data is beschreven in Bijlage IV. De meetgegevens zijn afgekapt op eind 2020, omdat 2021 niet compleet is. De meetpunten in 2021 worden wel gebruikt als onafhankelijke toets voor de voorspellende waarde van de vernieuwde regressie relatie (zie Figuur 27 in Bijlage V). We bepalen de fit van deze recente, hoogfrequente data (2009-2020) zonder de variabelen 'jaar' en 'dag'. Door deze niet mee te nemen in de voorspelling, wordt er een model fit gedaan die over alle jaren gemiddeld optimaal is, en waarbij de mogelijkheid dat de chlorideconcentratie een geleidelijke stijging of daling heeft in een jaar niet wordt meegenomen. In Figuur 28 in Bijlage V staan de jaarcoëfficiënten voor de jaren (2009-2020) voor deze recente, hoogfrequente data. De jaarcoëfficiënten in deze meetreeks lijken wel iets meer verband te houden in de jaren die elkaar opvolgen dan in de historische meetreeks.

Uit de model fit op deze recente, hoogfrequente reeks blijkt dat de 'weekdag' variabele niet (meer) van belang is. Deze is, behalve een klein effect op zondag in Eijsden (zie Figuur 24 in Bijlage IV), niet meer significant (niet getoond). De 'maand' variabele is daarentegen wel nog steeds relevant. In Figuur 14 staan de maandcoëfficiënten weergegeven voor deze recente, hoogfrequente data. In Figuur 26 in Bijlage V zijn de verschillen in chloridevracht per maand op een andere manier weergegeven; zoals ze te zien zijn in de meetgegevens.



Figuur 14. De maandcoëfficiënten voor een model fit met de factor 'Maand' van Eijsden en Lobith voor de recente, hoogfrequente data van 2009-2020 (zie ook Bijlage III).

Om in de praktijk de variatie van maand tot maand niet te abrupt te laten verspringen, voegen we net zoals eerder bij jaarvariatie in de historische meetreeks (Model 7, Tabel 3) een interactieterm toe met 'dag in de maand'. Deze term geeft geen grote verbetering (niet getoond), maar laat de voorspellingen tussen maandgrenzen soepel verlopen.

De formule voor de basis van dit regressiemodel met waterafvoer is gegeven in Vergelijkingen 1-3. Het regressiemodel met de termen voor variabelen erbij kan geschreven worden als

$$\log(L_C) = a + b \cdot \log(\log(Q)) + bm + bmd \cdot mdag + bd \cdot mdag \quad \text{Vergelijking (5)}$$

Hierin is L_C de chloridevracht, Q het debiet, a de coëfficiënt voor de intercept, b de coëfficiënt voor het debiet, bm de coëfficiënt voor factor maand, en bmd de coëfficiënt voor de interactieterm maanddag en maand. De factor bd is de coëfficiënt voor de maanddag op zich. De factor $mdag$ is een getal dat staat voor de dag in de maand (1-28 voor februari, tot 30 of 31 voor de andere maanden). In Bijlage III staan de coëfficiënten voor deze vergelijking. Voor Lobith is de R^2 voor dit model op de laatste 12 jaar 0,82 en voor Eijsden is dit 0,86. Deze formule kunnen we omschrijven zodat we de log kwijtraken bij de vracht:

$$L_C = e^a \cdot \log(Q)^b \cdot e^{bm} \cdot e^{bmd \cdot mdag} \cdot e^{bd \cdot mdag} \quad \text{Vergelijking (6)}$$

De chloride concentratie is dan simpelweg de chloridevracht uit Vergelijking 6 gedeeld door debiet Q :

$$C_c = \frac{e^a \cdot \log(Q)^b \cdot e^{bm} \cdot e^{bmd \cdot mdag} \cdot e^{bd \cdot mdag}}{Q} \quad \text{Vergelijking (7)}$$

In deze vergelijking kunnen de model coëfficiënten (' a ' voor het snijpunt en ' b ' voor de variabelen) voor waterafvoer (b) de betreffende maand (bm) of interactieterm maanddag maand (bmd) of maanddag (bd) uit Bijlage III worden ingevuld om de chlorideconcentratie te berekenen. De variabele $mdag$ is de dag in de maand waarvoor de voorspelling wordt gedaan.

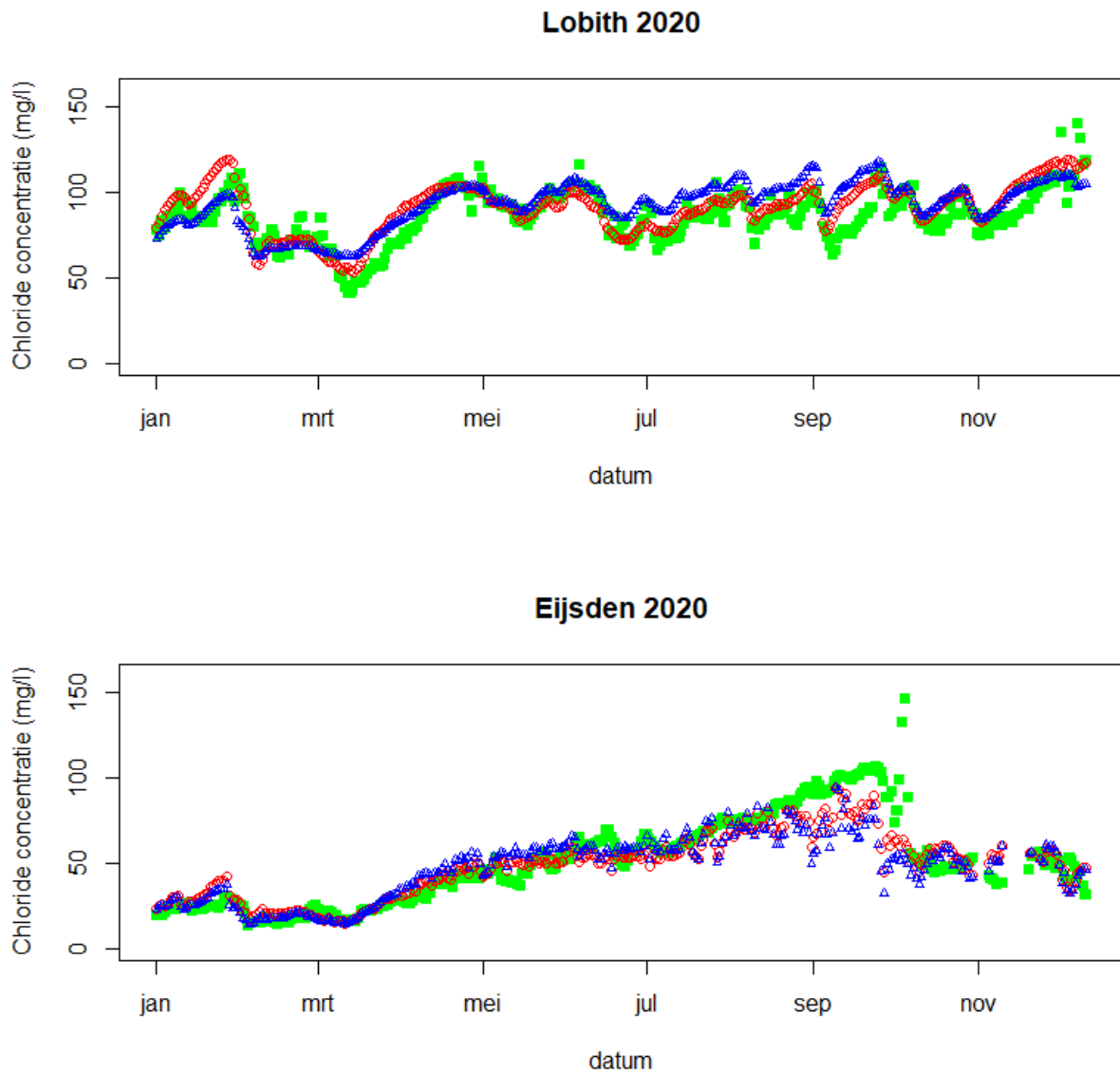
Ter referentie van de resultaten van Vergelijking 7 nemen we de vergelijkingen van Van der Wijk (2018) die momenteel veel gebruikt worden, en passen deze toe op de recente, hoogfrequente data die we gebruiken. Van deze vergelijkingen berekenen we de MAPE (Vergelijking 4) zonder kruisvalidaties. Van der Wijk (2018) past verschillende modellen toe voor Lobith (QC relatie) en Eijsden (Literatuurvergelijking type 4, Hoofdstuk 4).

De resultaten zijn te zien in Tabel 4. Als we uitgaan van MAPE, dan leidt het nieuwe regressiemodel (Vergelijking 7) tot een verbetering ten opzichte van de modellen van Van der Wijk (2018) op beide locaties. Dit beeld blijft overeind als we alleen voor recente (>2016) data de gefitte en werkelijke chlorideconcentraties vergelijken. Ook als we de MAPE alleen uitrekenen voor de hogere gemeten chloridewaarden (>80 percentiel van alle waarden tussen 2007 en 2018), blijft het vernieuwde regressiemodel iets beter.

Tabel 4. Vergelijking van de huidig gebruikte modellen uit Van der Wijk (2018) en Vergelijking 7 voor recente, dagelijkse data (2009-2020) in termen van 'goodness of fit' voor concentraties. De MAPE is berekend met Vergelijking 4. In de kolom 'MAPE > .' zijn de berekeningen alleen gedaan voor datapunten waarvan de werkelijke chloridemeting hoger dan het 80 percentiel van de 12 jarige meetreeks was. Hoe lager het MAPE getal, hoe beter het model past op de data.

Eijsden				
	MAPE (2009-2020)	MAPE recent (2019-2020)	MAPE > 58 mg/l	
Vergelijking 7	19,2	16,9	18,8	
Van der Wijk Eijsden (Hoofdstuk 4)	20,0	19,4	21,7	
Lobith				
	MAPE (2009-2020)	MAPE recent (2019-2020)	MAPE > 96 mg/l	
Vergelijking 7	9,8	10,4	9,6	
Van der Wijk Lobith (Hoofdstuk 4)	13,3	14,7	11,4	

In Figuur 15 zijn de voorspellingen van beide modellen uitgezet tegen de meetwaarden voor een recent jaar, 2020. Daar is te zien dat de waarden van de vergelijkingen van Van der Wijk (2018) over het algemeen iets hoger liggen dan die van Vergelijking 7, en hoger dan de gemeten waarden, vooral bij Lobith. Vooral bij Lobith is te zien dat de vernieuwde regressie relatie een iets groter bereik heeft (lage chloride, hoge chloride) dan de modellen van Van der Wijk (2018).



Figuur 15. Tijdreeksen van de gemeten en voorspelde chloride concentratie in 2020. De voorspellingen van Vergelijking 7 (rode cirkels) en de voorspelling met de vergelijkingen van Van der Wijk (2018) (blauwe driehoeken) ten opzichte van de meetwaarden (groene vierkanten).

In hoofdstuk 7 evalueren we hoe we de voorspelling nog kunnen verbeteren, en daarbij de kans op een goede voorspelling kunnen weergeven.

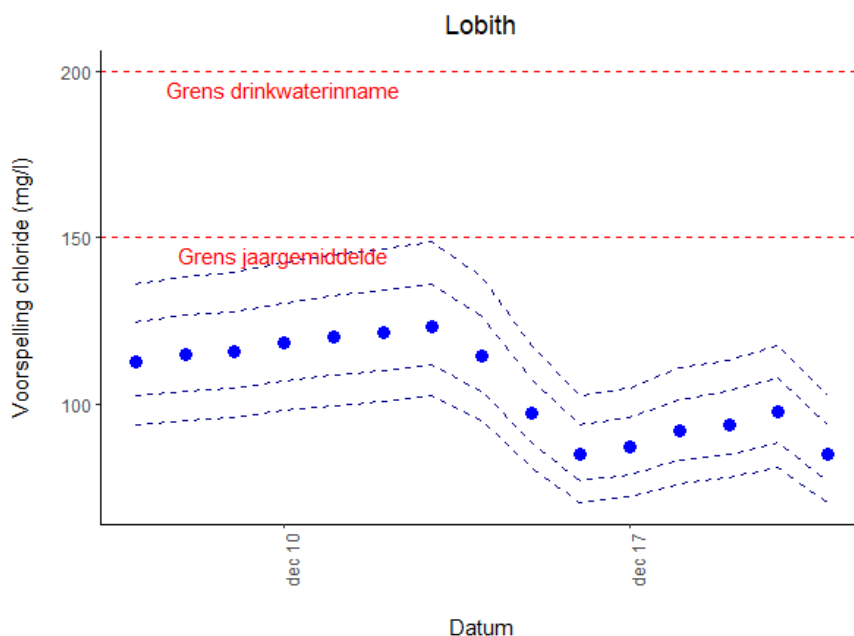
7 Verbetering regressiemodel: de voorspellingen aanpassen en spreiding bepalen

7.1 Voorspelling met predictie-intervallen

Met het model uit het vorige hoofdstuk kunnen we de chlorideconcentraties bij Lobith gemiddeld (relatief) nauwkeuriger benaderen dan die bij Eijsden (zie Tabel 3, Tabel 4, MAPE waarden).

Er blijven verhogingen (en verlagingen) van de chloridevracht die we niet goed kunnen voorspellen met het model dat beschreven is in Vergelijking 7. In Figuur 15 is te zien dat deze afwijkende voorspellingen vaak gelden voor meerdere aaneengesloten metingen. Hoge chlorideconcentraties worden vooral bij Eijsden onderschat.

De voorspellingen die gedaan worden met Vergelijking 7 zijn over het algemeen beter dan de bestaande modellen (zie Tabel 3 en Tabel 4). Maar deze voorspellingen van de chlorideconcentratie met een gegeven waterafvoer, maand en maandag geven nog niet het hele beeld. Er is voor elke voorspelling een bepaalde kans dat deze voorspelling goed is. Sommige voorspellingen zijn zekerder dan andere voorspellingen. Daarom is het inzichtelijk om deze kansen te kennen voor elke afzonderlijke voorspelling.



Figuur 16. Visualisatie van het gebruik van predictie intervallen voor dagelijkse chloridevoorspellingen voor een willekeurige periode. De buitenste lijn rond de waarden is het 80% predictie interval (10-90 percentielen), de binnenste lijn is het 50% predictie interval (25-75 percentielen). Hoe breder deze banden, hoe onzekerder de voorspelling is. Horizontale rode lijnen zijn de referentie waarden voor water dat bestemd is als bron voor drinkwater, zie Tabel 1b. Elke blauwe cirkel is een voorspelling op een dag. Het script om dit type figuur te genereren staat in Bijlage II.

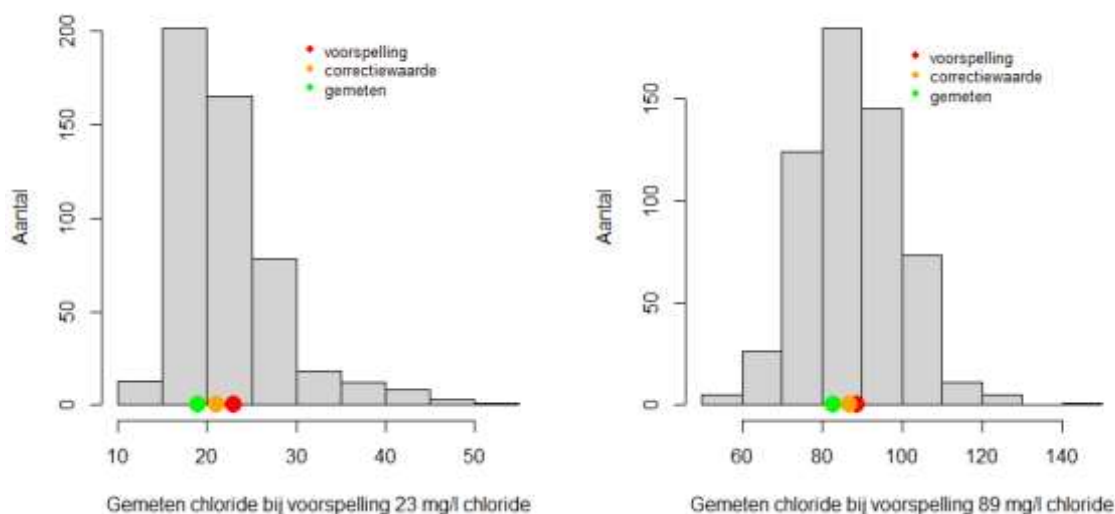
Een standaard manier om onzekerheid in de voorspelling te kwantificeren is om gebruik te maken van predictie-intervallen. Het 80% predictie interval geeft bijvoorbeeld aan dat bij een voorspelde waarde, er een 80% kans is dat de werkelijke uitkomst hierbinnen ligt. Deze predictie-intervallen kunnen in de programmeertaal R op een simpele manier worden gegenereerd. In Figuur 16 zijn de 80% en 50% predictie-intervallen voor enkele waarden te zien die voorspeld zijn voor enkele willekeurige datums en waterafvoeren. De predictie-intervallen worden in R gebaseerd op de standaard error van de predictor waarden (hier o.a. waterafvoer). Door op deze manier de voorspelling te visualiseren is het duidelijk dat elke voorspelling een bepaalde kans heeft om andere waarden te hebben. Zie <https://rpubs.com/aaronsc32/regression-confidence-prediction-intervals> voor een nadere uitleg.

Voor de spreiding kunnen we, als alternatief, ook gebruik maken van de eerder al geobserveerde verschillen met de gemeten chloridewaarden bij een bepaalde voorspelling. We beschrijven twee methoden om de voorspelde chlorideconcentraties te verbeteren, én de kans weer te geven dat een bepaalde voorspelling overeenkomt met de werkelijk gemeten waarde. De eerste methode gebruikt een correctie van de voorspelling die is gebaseerd op de eerdere werkelijke metingen in de meetreeks van een bepaalde voorspelling. De tweede methode is gebruikt een correctie van de voorspelling gebaseerd op de meting vóór de voorspelde waarde. Voor beide methoden is een langlopende en/of hoogfrequente tijdreeks nodig met veel metingen. De methoden zijn toegepast op historische data met Model 7 uit Tabel 3. De coëfficiënt voor jaarvariatie bij een voorspelling in de toekomst wordt dan toch voor een deel uit metingen van het voorgaande jaar 'geleend' zodat deze coëfficiënt op voldoende metingen is gebaseerd.

7.2 Correctiemethode 1

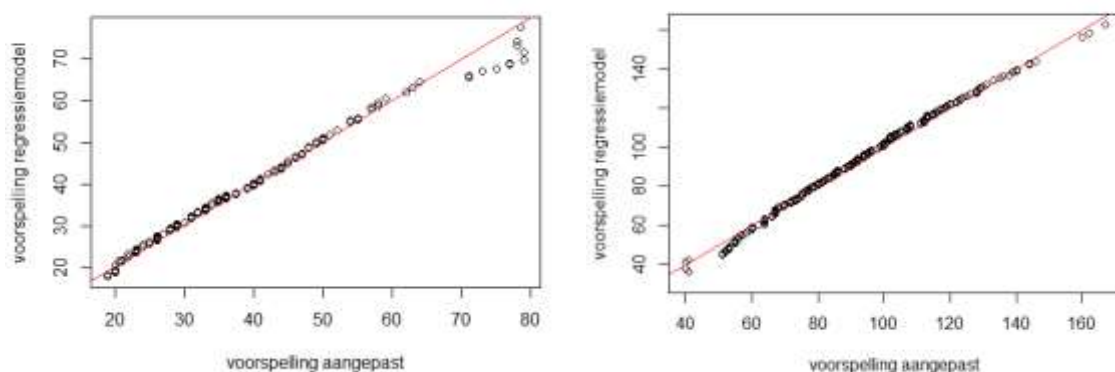
De eerste methode (Methode 1) gebruikt metingen van de gehele historische tijdreeks om een voorspelling te corrigeren. Dat gaat als volgt. Stel dat we 86 mg/l chloride voorspellen, dan kijken we in de meetreeks met chlorideconcentraties naar het verschil tussen alle punten met deze voorspelde waarde (86 mg/l) en de daadwerkelijke gemeten concentratie. Om dit te baseren op voldoende metingen, nemen we metingen in het verleden die voorspeld werden als 82-90 mg/l (+- 5%).

- Stap 1: Voorspelling van de chlorideconcentratie op basis van het regressiemodel
- Stap 2: Welke momenten in het verleden hadden ook deze voorspelling (+- 5%)
- Stap 3: Wat was de gemeten chlorideconcentratie op deze momenten
- Stap 4: De aangepaste voorspelling is de mediaan van de waarden van de metingen uit stap 3. De onzekerheid van deze aangepaste voorspelling wordt gegeven door de spreiding van de waarden in stap 3.



Figuur 17. Histogram van gemeten chloride waarden in het verleden bij een voorspelling die dicht ligt bij de voorspelling in dit plaatje. Links een voorbeeld voor Eijsden, rechts een voorbeeld voor Lobith.

In Figuur 17 is te zien dat de aangepaste voorspelling (oranje), gebaseerd op metingen in het verleden met dezelfde voorspelling inderdaad dichterbij de echte gemeten waarde (groen) van de voorspelling (stap 1) kan liggen dan het originele regressiemodel (rood). In het figuur is ook te zien dat de voorspellingen die in het verleden tussen de 82-90 mg/l lagen in werkelijkheid gemeten waarden hadden tussen de 50 en 125 mg/l. Het model en de werkelijk gemeten waarden liggen dus in sommige gevallen een eind uit elkaar. Deze aanpak corrigeert vooral de onderschatting van de hogere chloridewaarden voor locatie Eijsden (zie Figuur 18). Ook hele lage voorspellingen worden onbedoeld iets naar boven gecorrigeerd, dit komt waarschijnlijk doordat er bij deze lage voorspellingswaarden relatief veel uitschieters naar boven zijn in werkelijke gemeten chloridewaarden. In Tabel 5 staat dat de MAPE verbetering ten opzichte van de oorspronkelijke voorspelling gemiddeld ongeveer 0,2-0,5 % is. Dat lijkt niet veel, maar de verbetering is vooral bij waardevol voor het hoge concentratiebereik bij Eijsden. Juist het hoge concentratiebereik is relevant omdat deze aanleiding geven tot maatregelen.



Figuur 18. Twee voorbeelden met willekeurig getrokken recente (>2000) meetpunten om te voorspellen, met de aangepaste voorspelling met Methode 1 op de x-as ten opzichte van de voorspelling met het regressiemodel (Model 7, Tabel 3) op de y-as. Links locatie Eijsden, rechts locatie Lobith.

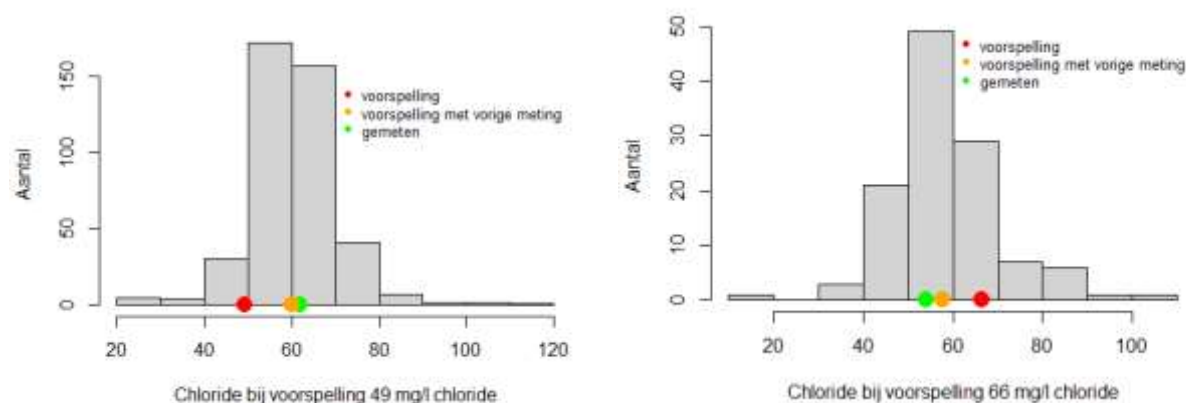
In deze methode is het aantal metingen waarop de verdeling wordt gebaseerd, afhankelijk van hoe vaak deze waarde eerder is voorspeld in de meetreeks. Daarom moet de meetreeks voldoende metingen hebben. Als de waarde niet eerder is voorspeld, kan worden teruggevallen op de originele voorspelde waarde zonder aanpassing.

7.3 Correctiemethode 2

In de tweede methode wordt de nieuwe voorspelling aangepast op basis van de meting vóór de te voorspellen waarde. De concentratie verandert niet heel snel, waardoor metingen met kleinere tussentijd meer op elkaar lijken dan metingen met grotere tussentijd. Dit fenomeen heet temporele autocorrelatie (Hyndman & Athanasopoulos 2018). We kunnen dus kijken naar het meetpunt vóór het te voorspellen punt en op basis hiervan de voorspelling aanpassen. Op basis van de ACF grafiek (Figuur 10) weten we dat hoe korter de periode tussen de metingen is, hoe groter de correlatie tussen deze punten is.

We doen opnieuw een voorspelling, maar voorspellen daarbij ook de afwijking met de gemeten waarden. Deze afwijking tellen we op bij de voorspelling, en we visualiseren de spreiding. Door niet te kijken naar de hoogte van de chloride meting zelf in eerdere metingen (we kijken alleen naar de afwijking) kunnen we voor Lobith, waar een groot verloop is in chlorideconcentraties, ook de metingen in eerdere jaren meenemen. Hieronder volgt een beschrijving van de benodigde stappen voor deze correctie:

- Stap 1. Voorspelling van chlorideconcentratie op basis van het regressiemodel
- Stap 2. Bepaal het verschil tussen gemeten waarde en voorspelde waarde door het regressiemodel voor de laatst bekende meting.
- Stap 3. Welke metingen uit het verleden hebben dezelfde afwijking als die uit stap 2 in mg/l uit stap 2 (+/- 10%)
- Stap 4. Wat was het relatieve model residu van de laatst bekende meting?
- Stap 5. Welke metingen uit het verleden hebben nog meer dit model residu uit stap 4 (+/- 15%)
- Stap 6. Wat is de overlap in metingen uit stap 3 en stap 5
- Stap 7. Wat is de periode (dagen) die ligt tussen de datum van de voorspelde waarde en de datum van de laatst bekende meting
- Stap 8. Voor de metingen uit het verleden in stap 6, welke metingen liggen precies de periode uit stap 7 later?
- Stap 9. De aangepaste voorspelling is de mediaan van de afwijking van de metingen uit stap 8 opgeteld bij de voorspelling uit stap 1. De onzekerheid van deze aangepaste voorspelling wordt gegeven door de spreiding van de waarden in stap 8 opgeteld bij de voorspelling uit stap 1.



Figuur 19. Histogram van chloride waarden spreiding op basis van de afwijking van metingen die precies even ver aflijgen van een vorig meetpunt (Methode 2). Links een voorbeeld voor Eijsden, rechts een voorbeeld voor Lobith.

Het is te zien in Figuur 19 dat de aangepaste voorspelling (oranje) nu dichterbij de werkelijke meetwaarde (groen) kan liggen dan de oorspronkelijke voorspelling (rood). De voorspelling wordt naar beneden bijgesteld als deze op basis van de vorige metingen lager moet zijn, en naar boven bijgesteld als deze hoger moet zijn. De hoge chloride voorspellingen worden niet specifiek verbeterd, zoals bij Methode 1 wel het geval is. Voor Methode 2 geldt, hoe verder in de tijd het vorige meetpunt van het te voorspellen meetpunt vandaan ligt, hoe minder voorspellende waarde de meting heeft. In Tabel 5 staat de gemiddelde MAPE nauwkeurigheid voor deze methode. De methode heeft vooral meerwaarde bij een voorspelling op de korte termijn.

7.4 Vergelijking van Methode 1 en Methode 2

In Tabel 5 staan alle MAPE waarden voor de originele voorspelling met het regressiemodel (Model 7, Tabel 3) ten opzichte van Methode 1 en Methode 2.

Tabel 5. Gemiddelde van de voorspellingsfout (MAPE) op de historische meetreeks. In groen is aangegeven wanneer Methode 1 of Methode 2 gelijk of beter is dan de originele voorspelling met het regressiemodel. In oranje is weergegeven waar deze slechter is. Tussen haakjes staat de gemiddelde standaard deviatie.

		Voorspelling Eijsden		Voorspelling Lobith	
		MAPE recent (>2000)	MAPE alles	MAPE recent (>2000)	MAPE alles
Op basis van Model 7, Tabel 3		19,6 %	21,3 %	9,5 %	10,8 %
Voorspelling aangepast Methode 1		18,8 % (11)	20,1 % (12)	9,1 % (14)	10,8 % (20)
Voorspelling aangepast Methode 2	Gemixte dagen	14,9 % (8)	13,2 % (8)	6,5 % (11)	7,9 % (15)
	1 dag	11,7 % (6)		4,8 % (10)	
	2 dagen	15,2% (8)		6,8 % (13)	
	3 dagen	16,4 % (9)		7,9% (14)	
	7 dagen	18,6 % (11)		8,6 % (14)	
	10 dagen	19,8 % (10)		9,0 % (15)	
	14 dagen	19,6 % (11)		11,9 % (14)	
	28 dagen	21,8 % (12)		12,7 % (15)	
	42 dagen	22,4 % (12)		14,5 % (16)	

In Tabel 5 is te zien dat de nauwkeurigheid van de voorspelling en ook de standaard deviatie het beste zijn met Methode 2. Maar, dit geldt alleen voor voorspellingen die kort na een al bekende meting volgen (maximaal 1 tot anderhalve week). Als we verder vooruit willen voorspellen (bijvoorbeeld om operationele verwachtingen op te stellen), presteert Methode 2 juist slechter dan Methode 1 én de originele voorspelling met het regressiemodel. Voor locatie Lobith zijn er op dit moment verwachtingen van de waterafvoer voor vijftien dagen vooruit beschikbaar, voor locatie Eijsden is dit vijf dagen (pers. comm, RWS). Omdat in de praktijk verwachtingen gewenst zullen zijn voor een periode van minimaal een week vooruit, en Methode 1 en Methode 2 ingewikkelder zijn om te implementeren, is op basis van praktische overwegingen gekozen om de voorspelling niet te corrigeren.

8 Verwachting van chlorideconcentratie en spreiding in de praktijk

In de vorige hoofdstukken hebben we een regressiemodel opgesteld en drie methoden beschreven om de spreiding in de voorspelling te berekenen, waaronder het uitrekenen van predictie-intervallen. Om de voorspelling met bijbehorende spreiding goed te kunnen interpreteren, is een geschikte weergave nodig. De applicatie is geprogrammeerd in R (voor de code, zie Bijlage II). In het operationeel waterbeheer zal in het algemeen een verwachting gewenst zijn voor meerdere momenten in de toekomst, van dichterbij (dagen) tot verder in de toekomst (weken). Om een verwachting op te stellen zijn minstens de volgende gegevens nodig als invoer:

- Een voorspelling van de waterafvoer
- De datum van de gewenste verwachting

Uit de datum van de gewenste verwachting wordt door het R script de maand gehaald die een factor bepaalt die het effect van de huidige maand weergeeft, en de interactie met dag in de maand. Voor Methode 1 en Methode 2 zullen gegevens van de historische meetreeks of de hoogfrequente, recente meetreeks nodig zijn voor de aangepaste verwachting; deze moeten ook aanwezig zijn en zullen regelmatig moeten worden aangevuld met recentere meetgegevens. Methode 1 en Methode 2 zijn lastiger in de implementatie voor gebruikers, deze zijn daarom op het moment niet geïmplementeerd.

De data (voorspelde waterafvoeren en datums) worden ingelezen in R, en het script (zie Bijlage II) wordt uitgevoerd, met als basis een regressievergelijking (Vergelijking 7). Met wat aanpassingen kan dat een informatieblad opleveren zoals in Figuur 20. Hierbij dient te worden opgemerkt dat de onzekerheid van de verwachting van de waterafvoer niet is meegenomen. Het meenemen van die onzekerheid zal de spreiding in mogelijke voorspelde waarden vergroten.



Figuur 20. Voorbeeld van het informatieblad voor de verwachting van chlorideconcentraties, voor een fictief jaar. In blauwe linten staan de percentielen van de mogelijke waarden aangegeven. In rood staan de grens voor drinkwaterinname (200 mg/l) en de grens voor het jaargemiddelde (150 mg/l) aangegeven. De tabel geeft aan hoeveel kans er is om deze waarden te overschrijden, bij deze verwachtingen. Lagere chlorideconcentraties kennen een lagere spreiding (niet getoond). De tabel helpt om de kansen op overschrijdingen bij een bepaalde verwachting expliciet te maken. De toenemende onzekerheid van de verwachting van waterafvoer zit nog niet in deze verwachting verwerkt.

De waterbeheerder kan waarden aangegeven waarboven het de moeite waard wordt om bepaalde maatregelen te nemen. Deze waarden zijn afhankelijk van de kosten die bespaard worden door niet in te grijpen wanneer het ook niet nodig is, ten opzichte van de opgetreden schade en kosten van maatregelen nadat de grens in chlorideconcentratie voor drinkwaterinname overschreden is. Deze waarde kan door de betrokkenen bijvoorbeeld gesteld worden op 10% kans om de drinkwaterinname grens (zie Tabel 1b) van 200 mg/l te overschrijden. Als deze waarde wordt bereikt, betekent dat dat in één op de tien gevallen drinkwaterinname grens overschreden wordt. De bijbehorende tabel (zie Figuur 20) maakt de kans op overschrijdingen bij een bepaalde verwachting inzichtelijk.

9 Conclusies en aanbevelingen

9.1 Samenvatting van de resultaten van het vernieuwde regressiemodel

Met het in dit rapport afgeleide regressiemodel kan voor Lobith de chloridevracht tussen 1972 en 2018 gemiddeld met 90% nauwkeurigheid voorspeld worden (en R^2 0,83) en voor Eijsden tussen 1977 en 2018 met 80% nauwkeurigheid (en R^2 0,88). Verder zijn ook drie opties ('Standaard', 'Methode 1', 'Methode 2') beschreven om de spreiding van de voorspelling te berekenen, en te visualiseren. De standaard deviatie van deze voorspellingen voor recente data (>2000) ligt rond de 11 mg/l voor Eijsden, en rond de 14 mg/l voor Lobith. De spreiding neemt toe met de hoogte van de chlorideconcentratie voorspellingen. Met de kennis over deze spreiding heeft een waterbeheerder of watergebruiker inzicht in de mogelijke afwijkingen van de gemaakte verwachtingen ten opzichte van de werkelijkheid. Bovendien kan de kans op een overschrijding van een bepaalde grens voor de chlorideconcentratie (bijvoorbeeld die van de drinkwaterinname van 200 mg/l) worden afgeleid. Vervolgens kan door betrokkenen worden afgewogen of het de moeite waard wordt om maatregelen te nemen. Bijvoorbeeld omdat bij deze kans het economisch aantrekkelijk wordt om maatregelen te nemen. De drempel om in te grijpen is dan afhankelijk zijn van de kosten die worden bespaard door niet in te grijpen, ten opzichte van de kosten die er gemaakt moeten worden om alsnog in te grijpen nadat de grens in chlorideconcentratie voor drinkwaterinname overschreden is.

Om dit regressiemodel af te leiden hebben we voor verschillende factoren van invloed eerst de historische relatie met chloridevracht gevisualiseerd. Hieruit kwam naar voren dat, zoals verwacht, waterafvoer het meest gerelateerd was aan chloridevracht. Ook maand, weekdag, jaar, lokale regenval, plotselinge regenval na droogte, temperatuur en verloop binnen een jaar hadden invloed op de chlorideconcentratie. De model fits van verschillende regressiemodellen zijn geëvalueerd, met inclusie van de verschillende factoren. Er is voor een robuuste toepassing voor een voorspellingsmodel gekozen waarin zo min mogelijk extra factoren zitten die zelf ook weer voorspeld moeten worden. Om de uiteindelijke model-fit zo robuust en actueel mogelijk te maken is er een recente, hoogfrequente (dagelijkse) meetreeks gebruikt. In het gekozen model, dat gefit is op de meest recente 12 jaar (2009-2020), zijn slechts de waterafvoer en de datum nodig voor een voorspelling, welke over het algemeen voorhanden zullen zijn. De R^2 is bij dit model voor Lobith 0,82 en voor Eijsden 0,86. Het voorspellingsmodel kan over het algemeen de chlorideconcentratie nauwkeuriger voorspellen dan bestaande modeltypen (het 'QC' model en het model van Baggelaar & Smits, 2003) en veel gebruikte modellen (Van der Wijk, 2018). Extreem hoge (en lage) chlorideconcentraties zijn vooral voor locatie Eijsden minder nauwkeurig te voorspellen dan minder extreme waarden.

Voor Eijsden wordt door Van der Wijk (2018) gesteld dat, zolang de afvoer boven de 10 m³/s blijft, de achtergrondconcentratie niet boven de 150 mg/l zal stijgen. Uit resultaten met het geactualiseerde regressiemodel blijkt dat met het huidige model onder normale omstandigheden (socio-economisch en klimatologische omstandigheden die vergelijkbaar zijn met de periode vanaf 1977 tot nu) de kritische chlorideconcentraties van 150 en 200 mg/l mogelijk helemaal niet bereikt zullen worden. Bij zeer lage waterstanden loopt de chlorideconcentratie weer terug voordat deze kritische waarden zijn bereikt. Dit komt doordat de chlorideconcentratie in het gebruikte model afhankelijk is van twee tegengestelde effecten, een toename van chloridevracht met waterafvoer boven de streep, en verdunning hiervan met waterafvoer onder de streep. Bij lage afvoer wordt de toename met waterafvoer relatief klein ten opzichte van de verdunning waardoor de chlorideconcentratie vanaf een bepaald moment snel gaat dalen met nog verder afnemende waterafvoer. Incidentele emissies, en andere factoren die zijn geëvalueerd zoals een specifieke maand in het jaar of plotselinge regenval na droogte, kunnen deze uiteraard wel verhogen. Ook toekomstige hogere structurele emissies door veranderde socio-economische situaties (bijvoorbeeld een nieuwe industrie die chloride loost op de Maas) of door

klimaatverandering (hogere structurele verdamping, heviger buien) kunnen de chlorideconcentratie verhogen. In reeks met de meest recente, hoogfrequente (dagelijkse) metingen daalt de chlorideconcentratie bij Eijsden niet bij zeer lage afvoer, maar blijft op hetzelfde maximumniveau (niet getoond). Met de fit op de historische data daalde de chlorideconcentratie bij zeer lage afvoer. Mogelijk betekent dit dat er recent een verandering gaande is ten opzichte van de historische reeks. Voor de Rijn bij Lobith blijkt uit het vernieuwde regressiemodel dat de kritische chlorideconcentraties van 150 en 200 mg/l wel gehaald kunnen worden, en zelfs in grote mate overschreden kunnen worden bij lagere afvoeren dan de huidige.

9.2 Aanbevelingen

Voor een goed passend model is het belangrijk om regelmatig te toetsen of het model nog voldoende de huidige situatie representeert. Door de verschillen in de gemeten en de voorspelde waarde van de meest recente meetwaarden (bijvoorbeeld 1-2 jaar) te bekijken, kan elk jaar de nauwkeurigheid van de voorspelling worden beoordeeld en kan eventueel de vergelijking worden aangepast. Bij voorkeur worden nieuwe meetgegevens ook direct toegevoegd aan de bestaande meetreeks. Dit is vooral belangrijk bij Methode 2, die gebruik maakt van de waarden van de vorige meting maar ook om de dataset voor Methode 1 uit te breiden. Omdat Methode 1 en Methode 2 gebruik maken van een meetreeks, is het de aanbeveling om voor het bepalen van de spreiding van voorspellingswaarden in het gekozen simpele model (Vergelijking 7) de standaard methode, die gebruik maakt van predictie intervallen, te gebruiken.

Er is nog ruimte voor verbetering. Ten eerste kan de regressie zelf verbeterd worden door extra factoren zoals lokale regenval en temperatuur, en voorspelde gebeurtenissen zoals het strooien van zout als gladheidsbestrijding (voor een mogelijke invloed van deze activiteit, zie Figuur 27 in Bijlage V) mee te fitten. Voor de effecten van lokale regenval kan in dat geval het beste een weerstation iets bovenstrooms van Lobith en Eijsden gekozen worden. Ook het effect van plotselinge regenval na droogte zou beter gedefinieerd kunnen worden met behulp van de etmaalsom en de maximale grootte van een regenbui, deze data zijn mogelijk ook beschikbaar voor bovenstroomse weerstations. Er moet voor deze factoren zelf dan wel ook een verwachting worden gedaan voordat ze meegenomen kunnen worden. Dit kan praktische bezwaren opleveren omdat er bijvoorbeeld afspraken gemaakt moeten worden met leveranciers van zulke verwachtingen, en er moet beoordeeld worden hoe nauwkeurig deze zijn.

Ten tweede kan er een andere aanpak voor het construeren van het model gekozen worden. Doordat er voor een regressiemodel is gekozen en niet voor een model waarin chloridebronnen expliciet en procesmatig worden gemodelleerd, is de methode niet geschikt om (effect van) verandering in bronnen (tijdelijke beheermaatregelen of beleidsmatige veranderingen in het stroomgebied, of veranderingen als gevolg van weersomstandigheden) te voorspellen. Wel kan deze invloed indirect worden meegenomen zodra de maatregel van kracht is en deze zich manifesteert in een verandering in de chloridevracht. Er loopt op dit moment een bronnenonderzoek door ICBR, dat meer inzicht zal geven in de invloed van verschillende bronnen op chlorideconcentratie.

Ten derde kan er verder gedacht worden over een ander, goed aansluitend twee-staps model. Het in dit rapport afgeleide regressiemodel heeft de neiging om de lage chlorideconcentraties te overschatten, en de hoge chlorideconcentraties te onderschatten, vooral voor de locatie Eijsden. Voor crisisbeheersing is vooral het accuraat voorspellen van hoge chlorideconcentraties belangrijk. Methode 1 geeft hierin al een verbetering bij Eijsden. In een nadere analyse kan er gekozen worden om vooral te richten op het goed voorspellen van de hogere chlorideconcentraties. Daarbij kan, als het moet, de nauwkeurigheid van de lagere chlorideconcentraties worden opgeofferd. Maar, het is onwaarschijnlijk dat afvoer-onafhankelijke incidentele extra emissies aan de hand van waterafvoer voorspeld kunnen worden. Het is wel goed mogelijk om bij een nadere analyse een gewogen regressie te doen waarbij uitschieters naar boven minder gewicht krijgen, en dat zo de chloridevracht die het gevolg is van waterafvoer (zonder incidentele extra emissies) beter ingeschat kan worden. Daarbovenop kan dan nog een andere

'forecasting' methode (Hyndman & Athanasopoulos, 2018) toegepast worden voor de afvoerafhankelijke incidentele emissies. Er zou bijvoorbeeld gebruik gemaakt kunnen worden van de stijging van de afgelopen meetpunten, om het volgende punt te voorspellen. Of, er kan een statistisch classificatiemodel of een 'Machine Learning tool' worden ingezet zoals een 'RandomForest' model of 'Deep Learning' met een 'Neural Network'. Dit kan dan direct toegepast worden op de gemeten waterstanden in plaats van de daarvan afgeleide waterafvoeren. Een deel van de onnauwkeurigheid van de vracht- en concentratie-voorspellingen kan namelijk veroorzaakt zijn door een onnauwkeurigheid in de bepaling van de waterafvoer. De waterafvoeren worden niet gemeten, maar geschat aan de hand van een 'Qf' (of QH) relatiemodel (zie bijv. Van Vuuren et al., 1999; Ogink & Stolker, 2004; Vermeulen & Behrens, 2005). De kracht van de toegepaste regressiebenadering ten opzichte van de 'Machine learning' benadering is dat er minder gegevens nodig zijn om tot een goed werkend model te komen. Ook is de regressievergelijking expliciet en geen 'black box'. Maar gezien er een lange tijdreeks voorhanden is en de nauwkeurigheid van chlorideconcentratie voorspelling beter zou kunnen worden is het de moeite waard de toegevoegde waarde van dergelijke technieken te verkennen.

Ten vierde kan er ook nog aandacht geschonken worden aan het vaststellen van de verwachte spreiding van de voorspelling. De spreiding wordt op dit moment met Methode 1 op basis van de eerdere voorspellingen teruggezocht in de bestaande data. Voor voorspellingen met weinig gelijke voorspellingen in het verleden, kan de spreiding op toeval gebaseerd zijn. Mochten de chlorideconcentraties in de toekomst incidenteel hoger worden dan in de meetreeks aanwezig, dan kan de spreiding met Methode 1 zelfs niet berekend worden omdat er geen referentiegegevens zijn. Het verdient de aanbeveling om, als dit de verwachting is, een methode te implementeren die de spreiding in die gevallen kan extrapoleren. Bijvoorbeeld door simpelweg de standaard berekende predictie intervallen te gebruiken. Daarbij wordt dan geen correctie gedaan van de voorspelling zelf, wat in Methode 1 wel gedaan wordt en wat een verbetering van de voorspelling geeft. De verwachte spreiding van een operationele verwachting zal in de realiteit ook afhangen van de zekerheid in de verwachting van de waterafvoer die voor de verwachting van de chlorideconcentratie wordt gebruikt. Deze onzekerheid is nu niet in de berekening van de spreiding meegenomen. Deze spreiding in waterafvoerverwachting zal toenemen hoe verder in de toekomst naar voren wordt gekeken. Het is mogelijk dat deze spreiding zelfs groter is dan de spreiding in de huidige modelvoorspellingen. In een toekomstige applicatie zal deze spreiding in de verwachting van de waterafvoer moeten worden toegevoegd aan de spreiding van de chlorideconcentratie-voorspelling zelf. Hiervoor zijn gegevens over de verdeling van de onzekerheid van de verwachting in waterafvoer nodig.

De hier ontwikkelde methode is te gebruiken voor klimaatscenario-studies waarbij trends in het optreden van lage afvoeren te koppelen zijn aan hoge chlorideconcentraties. Het hier voorgestelde regressiemodel is generieker (het geldt voor beide locaties) en nauwkeuriger dan de tot op heden toegepaste modellen, maar voor zowel de huidige als hier ontwikkelde regressie zijn er beperkingen. Het is inherent aan een regressiemodel dat processen niet expliciet gemodelleerd worden maar meegenomen in de modelfit. Hoewel het toegepaste regressiemodel goed werkt op de reeks met de ontwikkelingen in chlorideconcentratie Eijsden en Lobith van grofweg de afgelopen 50 jaar, en ook op recente data, is het geen gegeven dat relaties die de chlorideconcentratie in relatie tot debiet hebben bepaald in de (verre) toekomst niet veranderen. Een procesmodel zal in dat opzicht meer inzichtelijke scenario-studies mogelijk maken.

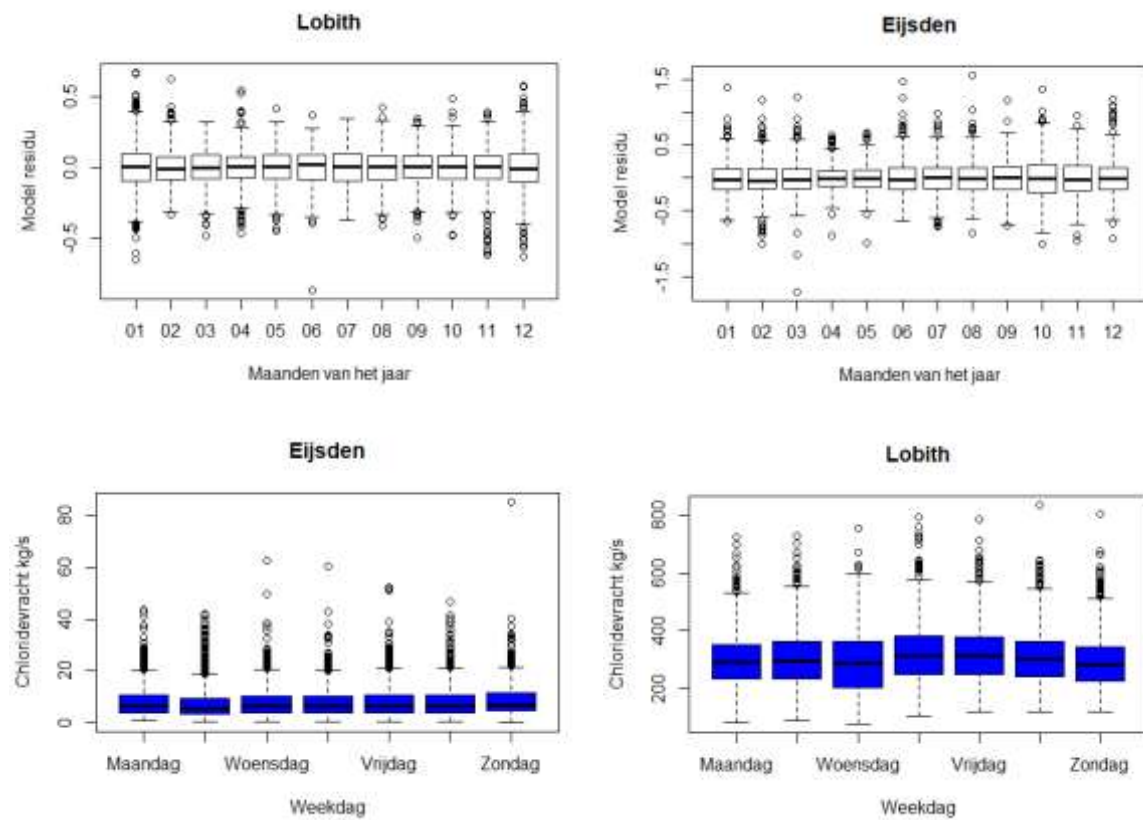
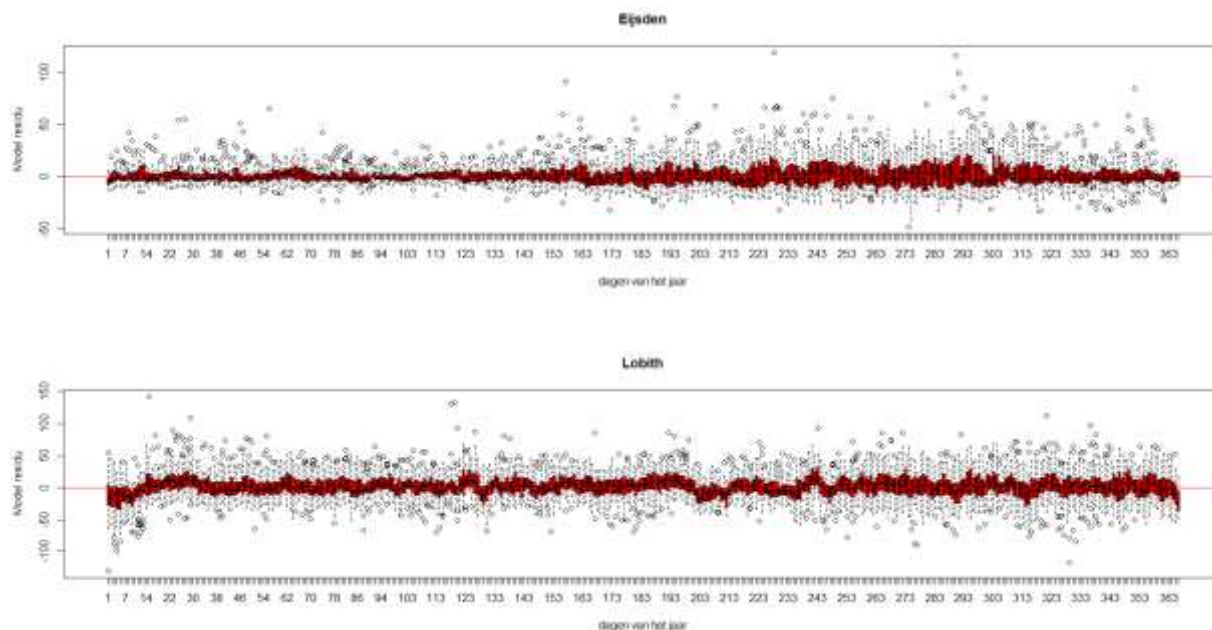
Ten slotte, als er voorspellingen gedaan worden met de regressiemodellen inclusief de onzekerheid ervan, is het ook een aanbeveling om bij het huidige R script een gebruikersinterface te maken. In deze gebruikersinterface kan data en aanvullende informatie gevraagd worden aan de gebruiker. Er kunnen dan ook interactief visualisaties worden getoond op basis van de keuze van de gebruiker. De methode is daardoor toegankelijker voor de gebruiker die chlorideconcentraties wil voorspellen.

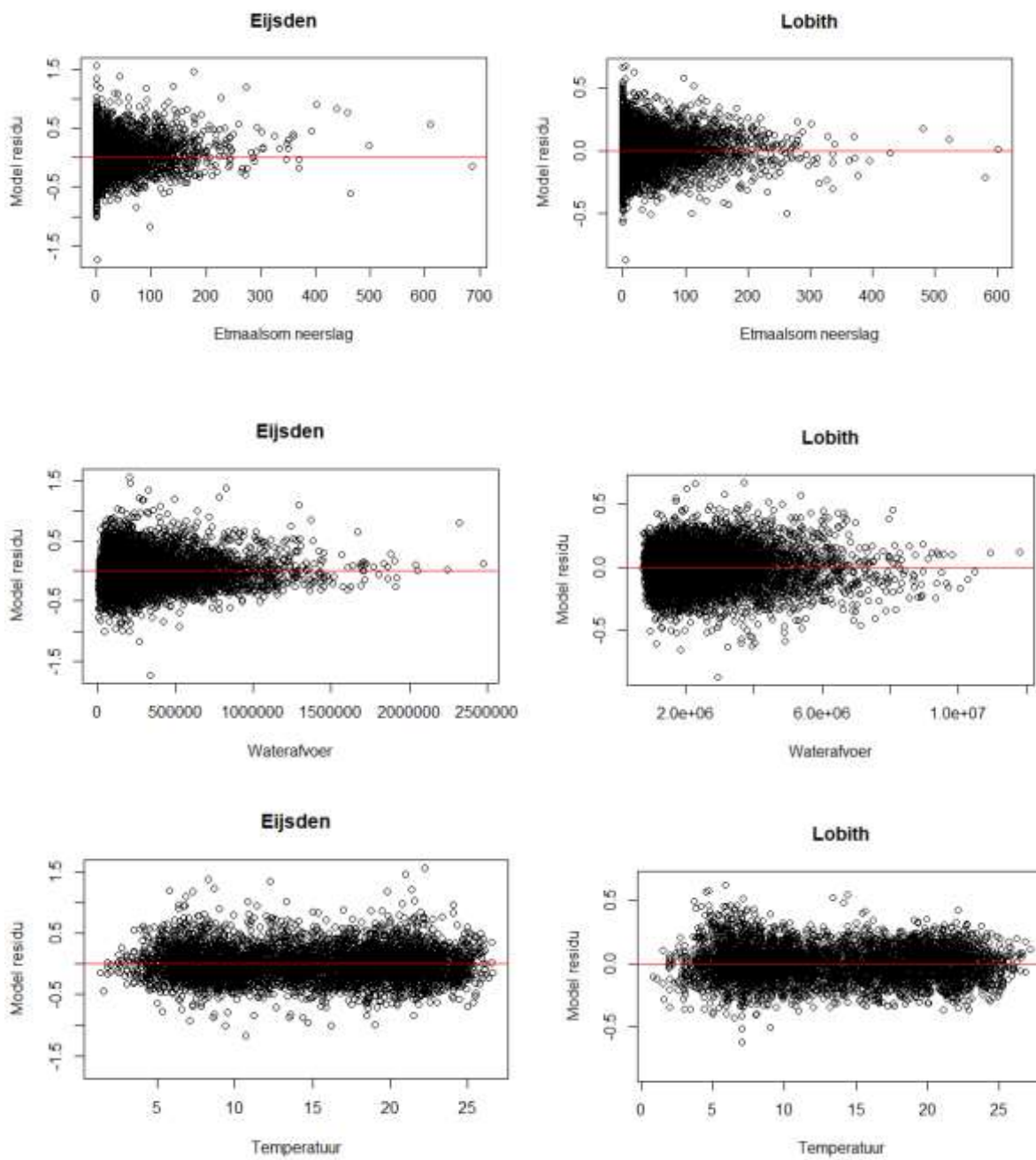
10 Referenties

- Baggelaar, P.K., Smits, A. (2003); Detecteren invloed gewijzigd spui-beheer met statistisch model chloride gehalte Andijk. RIWA-Rijn rapport.
- Baggelaar, P.K., Van der Meulen, E.C.J. (2009); Historische en toekomstige ontwikkelingen chloridebelasting in het traject Lobith tot Andijk. RIWA Rapport
- Bonte, M. (2009); Drinkwaterfunctie Markermeer en verzilting IJsselmeergebied BTO 2009.041(s)
- Bonte, M., Zwolsman, G. (2009); Klimaatverandering en verzoeting van de Rijn. Platform H2O / 20
- Van den Boogaard, H., Marjolein Mens, M. (2018); Eenvoudige zoutrelatie voor operationaal beheer van Kanaal Gent-Terneuzen. Memo.
- Van den Brink, M., Huismans, Y., Blaas, M., Zwolsman, G. (2019); Climate Change Induced Salinization of Drinking Water Inlets along a Tidal Branch of the Rhine River: Impact Assessment and an Adaptive Strategy for Water Resources Management. *Climate* 7, 49; doi:10.3390/cli7040049
- De Bruijn, F.A., Van Mazijk, A. (2003); Klimaatinvloeden op de kwaliteit van het Rijnwater, Startdocument Klimaatproject RIWA. RIWA-Rijn rapport.
- Deltaprogramma Zoetwater (2019); Agendapunt Bovenstroomse aspecten Zoetwater Agendapunt Bovenstroomse aspecten Memo
- Evers C.H.M. (2006); Getalswaarden voor de algemene fysisch-chemische kwaliteitselementen. Lelystad, The Netherlands: Rijksinstituut voor Integraal Zoetwaterbeheer en Afvalwater-behandeling (RIZA), RIZA rapport nr. 9R6513B0D0.
- Helpdesk Water (2019); Handleiding Normering Chloride Drinkwater(bronnen) Werkversie Helpdesk Water juli 2019 <https://www.helpdeskwater.nl/@214714/handleiding-chloridenormering-drinkwater-bronnen/>
- Hyndman, R.J., & Athanasopoulos, G. (2018); *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. Accessed on 2020-06-28.
- KNMI weerdata, openbaar. Beschikbaar op <https://www.knmi.nl/nederland-nu/klimatologie/daggegevens>
- Ministerie van Infrastructuur en Waterstaat (2019); Nederland beter weerbaar tegen droogte. Eindrapportage Beleidstafel Droogte. December 2019 <https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2019/12/18/nederland-beter-weerbaar-tegen-droogte/2019261596.01+Eindrapportage+Beleidstafel+Droogte.pdf>
- Ogink H.J.M., Stolker C. (2004); Verbetering Qf-relaties. Rapport Delft Hydraulics. <https://repository.tudelft.nl/islandora/object/uuid%3A16ba5745-0cdc-4f7e-bf09-a633e7fcad4d>
- Sjerps, R., Ter Laak, T., Zwolsman, G. (2016); Ontwikkeling waterkwaliteit bij innamepunten van oppervlaktewater voor de drinkwatervoorziening. KWR BTO rapport 2016.028

- Stahl, K., Weiler, M., Kohn, I., Freudiger, D., Seibert, J., Vis, M., Gerlinger, K., Böh, M. (2016) The snow and glacier melt components of streamflow of the river Rhine and its tributaries considering the influence of climate change. Report No. I-25 of the CHR International Commission for the Hydrology of the Rhine Basin. https://www.chr-khr.org/sites/default/files/chrpublications/asg-rhein_synthesis_en.pdf
- Stroomberg, G., Neefjes, R., de Jonge J., Bannink, A., van de Haar, G., Zwamborn, C. (2018); Jaarrapport 2017 De Rijn. RIWA-Rijn, Nieuwegein, The Netherlands. ISBN 978-90-6683-168-1. Beschikbaar via: <https://www.riwa-rijn.org/publicatie/jaarrapport-2017-de-rijn/> geopend op 2020-07-17.
- Verbruggen E.M.J., Moermond C.T.A., Janus J.A., Lijzen J.P.A. (2008); Afleiding van milieurisico-grenzen voor chloride in oppervlaktewater, grondwater, bodem en waterbodem. RIVM Rapport 711701075/2008
- Vewin (2017); 'Drinkwaterstatistieken 2017', een rapport in de reeks 'Drinkwaterstatistieken'
- Van Vossen- van den Berg, J., Cirkel, D.G., Nijhuis N.J., Mesman G.A.M., Huiting, H. (2019); Achtergrond chloridenormering en analyse effecten van overschrijding van de norm. KWR Rapport BTO 2019.014
- Van Vuuren, W., Moll, R., Vervoorn, H. (1999); De Q H-relatie zijn langste tijd gehad? Op weg naar het Qf-tijdperk voor de Rijntakken. H2O 25 19-24. <https://edepot.wur.nl/363506>
- WHO (1996); Guidelines for Drinking-Water Quality, 2nd edition https://www.who.int/water_sanitation_health/publications/gdwq2v1/en/index1.html
- Van der Wijk, R. (2018); Advies met betrekking tot achtergrondconcentratie Lobith en Eijsden. Memo, Deltares
- Vermeulen, C.J.M. en H.W.A. Behrens (2005); Informatiebehoefte afvoeren Rijkswaterstaat. HKV LIJN IN WATER rapport PR1049, december 2005.
- Zwolsman, J.J.G., Van Bokhoven, A.J. (2007); Impact of summer droughts on water quality of the Rhine river : a preview of climate change? Water Sci Technol (2007) 56 (4): 45–55. <https://doi.org/10.2166/wst.2007.535>

I Bijlage I. Model Residu variatie met diverse variabelen





Figuur 21. Model residuen ten opzichte van (van boven naar beneden) dag in het jaar (boxplots), maand van het jaar (boxplots), dag in de week (boxplots), etmaalsom neerslag, waterafvoer, temperatuur.

II Bijlage II. Code voor visualisatie van chloride voorspellingen

```
##### Figuur 16. Chloride voorspelling met predictie intervallen #####
## Script voor verwerking en voorspelling Chloride concentratie
## Project voor RWS
## # Eijsden en Lobith
## Tessa Pronk versie 1 op 7 april 2020
## laatste versie: 6 april 2021
## Nodig:
## Bestaande data, RIWA dataformat
## (met iig kolommen 'datum', 'paramet_naam', 'waarde')
## Bestand met datums en waterafvoer om te voorspellen
## (met kolommen "locatie", "datum", "Waterafvoer_voorspelling")
#####

library(lubridate)
library(ggplot2)
setwd("...") # de werk-directory
# gegevens van de te voorspellen chloride
VoorspDat<-read.csv2("voorspellingL1.csv",sep=";")

# Bewerk de gegevens voor
VoorspDat$waterafvoerl<-VoorspDat$Waterafvoer_voorspelling*1000
VoorspDat$Month<-format(as.Date(VoorspDat$datum,format='%d-%m-%Y'), "%m") # Voeg kolom met label maand toe
VoorspDat$Month_day <- format(as.Date(VoorspDat$datum), "%d") # Voeg kolom met label dag in de maand toe
VoorspDat$Month_day<- as.numeric(VoorspDat$Month_day)

filename<-as.character(VoorspDat$Locatie[1]) # Dit is de naam van de locatie waarvoor de voorspelling gedaan wordt

# read in DATAfiles (waarop het model wordt gefit. Alternatief: gebruik Vergelijking 7.
Dat<-read.csv2(paste("RIWABaseTM2018_",filename,".csv",sep=""))
Dat$datum <- ymd(Dat$datum) # Zet datum in datumformaat
Dat$Month <- format(as.Date(Dat$datum), "%m") # Voeg kolom met label maand toe
Dat$Month_day <- format(as.Date(Dat$datum), "%d") # Voeg kolom met label dag in de maand toe
Dat$Month_day<- as.numeric(Dat$Month_day)

# kolom 4 is datum, kolom 9 is chloride in RIWABase data
Chl<-Dat[which(Dat$paramet_naam=='chloride'),]
Chl<-Chl[-which(duplicated(Chl[,c(4,9)])==TRUE),]

Wat<-Dat[which(Dat$paramet_naam=='waterafvoer'),c(4,9)]
colnames(Wat)<-c("datum","waterafvoer")
Wat<-Wat[-which(duplicated(Wat$datum)==TRUE),]
Chor<-merge(Chl,Wat,by="datum",all.x=TRUE,all.y=FALSE)
colnames(Chor)[which(colnames(Chor)=='waarde')<-"chloride"
Chor$chloride<-round(as.numeric(Chor$chloride),0)
Chor$waterafvoerl<-Chor$waterafvoer*1000 # niet m3 per s maar l per s
Chor$ChlVr<-(Chor$chloride*Chor$waterafvoerl)

Chor<-Chor[-which(is.na(Chor$ChlVr*Chor$waterafvoerl)==TRUE),] # missende waarden weghalen

# In dataset Eijsden zitten een paar rare lage waarden. Heel mei 1991 is afwijkend. Weghalen.
# ook valt de Chloride vracht op twee data met een orde van grootte, en stijgt daarna weer.
if(filename=="Eijsden"){
```



```
Chor<-Chor[-which(format(Chor$datum,"%Y-%m")=="1991-05"),]
Chor<-Chor[-which(Chor$datum=="1992-04-11"|Chor$datum=="1992-04-12"),]
}

Chor <- Chor[Chor$datum > ymd(20090101),] # Beperk de meetreeks tot de laatste 12 jaar
Ctr<-Chor
Cte<-VoorspDat

fit1<- lm(log(Ctr$ChIVr)~log(log(Ctr$waterafvoerl))+Ctr$Month*Ctr$Month_day)

# Om te voorspellen moet het dataobject dezelfde naam hebben als het te fitten dataobject
Ctr<-Cte

# Voorspelling met predictie intervallen (alternatief: gebruik Vergelijking 7 en predictieinterval berekening)
pred.int1<-exp(predict(fit1,Ctr,interval = 'prediction',level=0.50))/Ctr$waterafvoerl # met upper lower bound
pred.int2<-exp(predict(fit1,Ctr,interval = 'prediction',level=0.75))/Ctr$waterafvoerl # met upper lower bound
pred.int3<-exp(predict(fit1,Ctr,interval = 'prediction',level=0.90))/Ctr$waterafvoerl # met upper lower bound

mydata <- cbind(Ctr, pred.int1,pred.int2,pred.int3)
colnames(mydata) <-make.unique(colnames(mydata),sep="_")

# plot de resultaten
colors <- c("50 percentile" = "blue", "75 percentile" = "green", "90 percentile" = "red")

p <- ggplot(mydata, aes(as.Date(datum,format="%d-%m-%Y"),fit)) +
  geom_point(size=2)

# Voeg predictie intervallen toe
p + geom_line(aes(y = lwr, color = "50 percentile"), linetype = "dashed") +
  geom_line(aes(y = upr, color = "50 percentile", linetype = "dashed") +
  geom_line(aes(y = lwr_1, color = "75 percentile"), linetype = "dashed") +
  geom_line(aes(y = upr_1, color = "75 percentile", linetype = "dashed") +
  geom_line(aes(y = lwr_2, color = "90 percentile"), linetype = "dashed") +
  geom_line(aes(y = upr_2, color = "90 percentile", linetype = "dashed") +

labs(x = "\nDatum",y="Voorspelling chloride (mg/l)\n",title= filename) +
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank(),plot.title = element_text(hjust = 0.5),
  panel.background = element_blank(), axis.line = element_line(colour = "black"),plot.margin = margin(1, 4, 1, 1, "cm")) +
theme(axis.text.x=element_text(angle=90, hjust=1))
```

III Bijlage III. Regressiecoëfficiënten voor Vergelijking 7 en Model 7 uit Tabel 3.

In vergelijking 7 staat het model dat op basis van de hoogfrequente, recente data optimaal (makkelijk te implementeren versus accuraat) is. Model 7 uit Tabel 3 is het model dat op de historische reeks optimaal is.

$$\log(L_c) = a + b \cdot \log(\log(Q)) + bm + bmd \cdot mdag + bd \cdot mdag \quad (\text{Vergelijking 7})$$

$$\log(L_c) = a + b \cdot \log(\log(Q)) + bm + bwd + bd \cdot dag + bj + bjd \cdot dag \quad (\text{Model 7, Tabel 3})$$

De volgende overeenkomsten tussen coëfficiënten het R-script (gegeven na deze tabel) en de notatie in de hoofdttekst zijn van toepassing:

Variabele of factor	parameter	uitleg
Chor\$ChlVr	Lc	Chloridevracht (uitvoergrootheid in mg/s)
Chor\$waterafvoerl	Q	Debiet (invoergrootheid in l/s)
Intercept	a	intercept
log(log(Chor\$waterafvoerl))	b	coëfficiënt term die bij log(log(Q)) hoort
Chor\$Month-i-	bm	coëfficiënt voor factor maand, i ∈ [1,12]
Chor\$weekdag-k-	bwd	coëfficiënt voor factor weekdag, k ∈ [1,7]
Chor\$Yr-j-	bj	coëfficiënt voor factor jaar, j ∈ [1972,2018]
Chor\$day of Chor\$Month_day	bd	coëfficiënt voor variabele dag
Chor\$Yr-j:Chor\$day	bjd	coëfficiënt voor de interactieterm jaar-dag, j ∈ [1972,2018]
Chor\$Month-i:Chor\$Month_day	bmd	coëfficiënt voor de interactieterm maand-dag, j ∈ [1,12]

Met 'log' wordt hier het natuurlijke logaritme bedoeld. De coëfficiënten zijn afgeleid op chloridevracht data in mg/s, en waterafvoer data in l/s. In de lijst hieronder staan de coëfficiëntwaarden zoals bepaald op de recente, dagelijkse reeks, en daaronder op de historische tijdreeks. De coëfficiënten voor factoren zijn altijd ten opzichte van een referentie. Bijvoorbeeld, de coëfficiënt voor Chor\$Month01 is 0. De asterisken (*) geven de mate van significantie weer (de symbolen staan hier tussen de waarden waar ze voor gelden): 0 '****' 0.001 '***' 0.01 '**' 0.05 '*' 0.1 '.' 1.

call Eijsden recente, dagelijkse metingen, met interactieterm:

```
lm(formula = log(Chor$ChlVr) ~ log(log(Chor$waterafvoerl)) + Chor$Month * Chor$Month_day)
```

Residuals:

```
      Min       1Q   Median       3Q      Max
-0.7275 -0.1640 -0.0070  0.1514  1.1160
```

Coefficients:

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -3.103758   0.175412  -17.694 < 2e-16 ***
log(log(Chor$waterafvoerl))  7.529540   0.067983  110.757 < 2e-16 ***
Chor$Month02    0.235747   0.037617   6.267 4.07e-10 ***
Chor$Month03    0.056295   0.036935   1.524  0.12754
Chor$Month04   -0.085462   0.037219  -2.296  0.02171 *
Chor$Month05    0.055460   0.037250   1.489  0.13660
Chor$Month06    0.055079   0.037763   1.459  0.14477
Chor$Month07    0.016191   0.038163   0.424  0.67139
```

Chor\$Month08	0.156880	0.039162	4.006	6.29e-05	***
Chor\$Month09	0.271916	0.040460	6.721	2.06e-11	***
Chor\$Month10	0.271321	0.038792	6.994	3.11e-12	***
Chor\$Month11	0.173900	0.039096	4.448	8.90e-06	***
Chor\$Month12	0.084759	0.038819	2.183	0.02906	*
Chor\$Month_day	0.006164	0.001419	4.343	1.44e-05	***
Chor\$Month02:Chor\$Month_day	-0.013564	0.002156	-6.291	3.48e-10	***
Chor\$Month03:Chor\$Month_day	-0.008557	0.002009	-4.260	2.09e-05	***
Chor\$Month04:Chor\$Month_day	-0.003015	0.002054	-1.468	0.14228	
Chor\$Month05:Chor\$Month_day	-0.008784	0.002002	-4.389	1.17e-05	***
Chor\$Month06:Chor\$Month_day	-0.006447	0.002060	-3.129	0.00176	**
Chor\$Month07:Chor\$Month_day	-0.001324	0.002019	-0.656	0.51205	
Chor\$Month08:Chor\$Month_day	-0.002476	0.002050	-1.208	0.22723	
Chor\$Month09:Chor\$Month_day	-0.004000	0.002156	-1.855	0.06366	.
Chor\$Month10:Chor\$Month_day	-0.012050	0.002055	-5.864	4.87e-09	***
Chor\$Month11:Chor\$Month_day	-0.008896	0.002145	-4.147	3.43e-05	***
Chor\$Month12:Chor\$Month_day	-0.005071	0.002109	-2.405	0.01622	*

 Residual standard error: 0.2423 on 4041 degrees of freedom
 Multiple R-squared: 0.8609, Adjusted R-squared: 0.8601
 F-statistic: 1042 on 24 and 4041 DF, p-value: < 2.2e-16

Call Lobith recente, dagelijkse metingen, met interactieterm:

lm(formula = log(Chor\$chlvr) ~ log(log(Chor\$waterafvoerl)) + Chor\$Month * Chor\$Month_day)

Residuals:

Min	1Q	Median	3Q	Max
-0.36571	-0.08332	-0.00462	0.07299	0.62373

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.1487554	0.2203021	-0.675	0.499565
log(log(Chor\$waterafvoerl))	7.1903370	0.0815999	88.117	< 2e-16 ***
Chor\$Month02	0.0760264	0.0197611	3.847	0.000121 ***
Chor\$Month03	0.0145175	0.0191728	0.757	0.448980
Chor\$Month04	-0.0730733	0.0194616	-3.755	0.000176 ***
Chor\$Month05	-0.1631112	0.0196609	-8.296	< 2e-16 ***
Chor\$Month06	-0.2285775	0.0194202	-11.770	< 2e-16 ***
Chor\$Month07	-0.3314265	0.0193279	-17.148	< 2e-16 ***
Chor\$Month08	-0.2915005	0.0194264	-15.005	< 2e-16 ***
Chor\$Month09	-0.3090654	0.0201021	-15.375	< 2e-16 ***
Chor\$Month10	-0.2012528	0.0200116	-10.057	< 2e-16 ***
Chor\$Month11	-0.1783171	0.0201237	-8.861	< 2e-16 ***
Chor\$Month12	-0.1213884	0.0203993	-5.951	2.89e-09 ***
Chor\$Month_day	0.0003128	0.0007464	0.419	0.675182
Chor\$Month02:Chor\$Month_day	-0.0023862	0.0011331	-2.106	0.035267 *
Chor\$Month03:Chor\$Month_day	-0.0044461	0.0010487	-4.239	2.29e-05 ***
Chor\$Month04:Chor\$Month_day	-0.0045550	0.0010873	-4.189	2.86e-05 ***
Chor\$Month05:Chor\$Month_day	-0.0037458	0.0010608	-3.531	0.000418 ***
Chor\$Month06:Chor\$Month_day	-0.0044294	0.0010792	-4.104	4.13e-05 ***
Chor\$Month07:Chor\$Month_day	0.0022487	0.0010519	2.138	0.032594 *
Chor\$Month08:Chor\$Month_day	0.0006216	0.0010603	0.586	0.557751
Chor\$Month09:Chor\$Month_day	0.0026434	0.0011049	2.392	0.016786 *
Chor\$Month10:Chor\$Month_day	0.0003502	0.0010737	0.326	0.744307
Chor\$Month11:Chor\$Month_day	0.0021992	0.0011150	1.972	0.048636 *
Chor\$Month12:Chor\$Month_day	0.0046000	0.0011138	4.130	3.70e-05 ***

 Residual standard error: 0.1271 on 4122 degrees of freedom
 Multiple R-squared: 0.8194, Adjusted R-squared: 0.8184
 F-statistic: 779.4 on 24 and 4122 DF, p-value: < 2.2e-16

Call Eijsden, historische reeks:

lm(formula = log(Chor\$chlvr) ~ log(log(Chor\$waterafvoerl)) + Chor\$Yr * Chor\$day + Chor\$Month + Chor\$weekdag)

Min 1Q Median 3Q Max
 -1.73461 -0.17062 -0.01432 0.15564 1.56926 (Residuals)

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-5.621e+00	1.639e-01	-34.299	< 2e-16	***
log(log(Chor\$waterafvoerl))	8.569e+00	5.740e-02	149.275	< 2e-16	***
Chor\$Yr1978	-2.168e-01	1.026e-01	-2.114	0.034573	*
Chor\$Yr1979	1.974e-02	7.858e-02	0.251	0.801613	
Chor\$Yr1980	3.656e+00	8.373e-01	4.366	1.28e-05	***
Chor\$Yr1981	5.318e-02	7.862e-02	0.676	0.498744	
Chor\$Yr1982	1.479e-01	7.855e-02	1.883	0.059750	.
Chor\$Yr1983	-2.968e-01	7.862e-02	-3.775	0.000161	***
Chor\$Yr1984	2.453e-01	7.858e-02	3.122	0.001804	**
Chor\$Yr1985	-9.362e-03	7.901e-02	-0.118	0.905678	
Chor\$Yr1986	1.065e-01	7.852e-02	1.356	0.175198	
Chor\$Yr1987	1.265e-01	8.276e-02	1.529	0.126274	
Chor\$Yr1988	9.578e-02	8.126e-02	1.179	0.238527	
Chor\$Yr1989	5.586e-03	1.211e-01	0.046	0.963220	
Chor\$Yr1990	-2.416e-01	8.303e-02	-2.910	0.003625	**
Chor\$Yr1991	-8.939e-02	7.865e-02	-1.137	0.255757	
Chor\$Yr1992	-2.637e-01	7.785e-02	-3.388	0.000708	***
Chor\$Yr1993	-2.179e-01	1.035e-01	-2.105	0.035324	*
Chor\$Yr1994	7.628e-02	7.756e-02	0.983	0.325432	
Chor\$Yr1995	-3.317e-01	7.851e-02	-4.225	2.41e-05	***
Chor\$Yr1996	6.154e-02	7.826e-02	0.786	0.431714	
Chor\$Yr1997	-1.428e-01	7.808e-02	-1.828	0.067533	.
Chor\$Yr1998	-1.792e-01	7.868e-02	-2.277	0.022795	*
Chor\$Yr1999	-2.042e-01	7.809e-02	-2.615	0.008929	**
Chor\$Yr2000	-1.359e-01	7.768e-02	-1.749	0.080314	.
Chor\$Yr2001	-1.573e-01	7.840e-02	-2.007	0.044819	*
Chor\$Yr2002	-1.223e-01	7.817e-02	-1.565	0.117735	
Chor\$Yr2003	-2.148e-01	1.037e-01	-2.071	0.038350	*
Chor\$Yr2004	6.177e-02	7.717e-02	0.801	0.423438	
Chor\$Yr2005	-7.772e-02	1.030e-01	-0.755	0.450370	
Chor\$Yr2006	9.510e-02	1.028e-01	0.925	0.354805	
Chor\$Yr2007	-1.821e-01	1.026e-01	-1.776	0.075813	.
Chor\$Yr2008	-2.075e-01	1.019e-01	-2.037	0.041677	*
Chor\$Yr2009	-6.014e-02	1.034e-01	-0.582	0.560720	
Chor\$Yr2010	7.836e-02	1.290e-01	0.607	0.543743	
Chor\$Yr2011	1.179e-02	1.032e-01	0.114	0.909036	
Chor\$Yr2012	-1.846e-01	1.039e-01	-1.777	0.075649	.
Chor\$Yr2013	-6.066e-02	1.023e-01	-0.593	0.553014	
Chor\$Yr2014	-1.751e-01	1.046e-01	-1.675	0.094045	.
Chor\$Yr2015	-9.974e-02	1.042e-01	-0.957	0.338649	
Chor\$Yr2016	-2.871e-01	1.036e-01	-2.771	0.005609	**
Chor\$Yr2017	2.621e-01	1.029e-01	2.547	0.010885	*
Chor\$Yr2018	-2.716e-01	1.058e-01	-2.566	0.010299	*
Chor\$day	3.124e-04	4.794e-04	0.652	0.514614	
Chor\$Month02	-3.266e-02	1.788e-02	-1.827	0.067725	.
Chor\$Month03	-3.861e-02	2.428e-02	-1.590	0.111879	
Chor\$Month04	-1.295e-01	3.319e-02	-3.903	9.59e-05	***
Chor\$Month05	-8.283e-02	4.273e-02	-1.938	0.052613	.
Chor\$Month06	-2.814e-02	5.245e-02	-0.536	0.591670	
Chor\$Month07	9.368e-03	6.235e-02	0.150	0.880575	
Chor\$Month08	1.120e-01	7.253e-02	1.544	0.122511	
Chor\$Month09	4.707e-02	8.242e-02	0.571	0.567952	
Chor\$Month10	8.946e-02	9.236e-02	0.969	0.332763	

Chor\$Month11	-4.272e-02	1.022e-01	-0.418	0.676028	
Chor\$Month12	-9.499e-02	1.121e-01	-0.847	0.396952	
Chor\$weekdagDinsdag	-4.019e-02	1.110e-02	-3.621	0.000295	***
Chor\$weekdagWoensdag	-1.445e-02	1.170e-02	-1.235	0.216951	
Chor\$weekdagDonnerdag	-5.106e-03	1.185e-02	-0.431	0.666512	
Chor\$weekdagVrijdag	5.647e-03	1.177e-02	0.480	0.631353	
Chor\$weekdagZaterdag	1.174e-02	1.178e-02	0.996	0.319221	
Chor\$weekdagZondag	2.968e-02	1.187e-02	2.500	0.012435	*
Chor\$Yr1978:Chor\$day	1.250e-03	4.950e-04	2.526	0.011563	*
Chor\$Yr1979:Chor\$day	2.716e-04	3.805e-04	0.714	0.475312	
Chor\$Yr1980:Chor\$day	-1.157e-02	2.472e-03	-4.681	2.91e-06	***
Chor\$Yr1981:Chor\$day	-8.242e-05	3.730e-04	-0.221	0.825129	
Chor\$Yr1982:Chor\$day	-3.693e-05	3.715e-04	-0.099	0.920810	
Chor\$Yr1983:Chor\$day	2.205e-03	3.734e-04	5.906	3.65e-09	***
Chor\$Yr1984:Chor\$day	-9.410e-04	3.729e-04	-2.523	0.011650	*
Chor\$Yr1985:Chor\$day	1.244e-03	3.721e-04	3.344	0.000831	***
Chor\$Yr1986:Chor\$day	1.677e-05	3.747e-04	0.045	0.964306	
Chor\$Yr1987:Chor\$day	1.971e-04	3.848e-04	0.512	0.608594	
Chor\$Yr1988:Chor\$day	4.751e-04	4.689e-04	1.013	0.310977	
Chor\$Yr1989:Chor\$day	1.746e-03	4.884e-04	3.574	0.000353	***
Chor\$Yr1990:Chor\$day	3.528e-04	3.934e-04	0.897	0.369825	
Chor\$Yr1991:Chor\$day	-6.606e-04	3.751e-04	-1.761	0.078253	.
Chor\$Yr1992:Chor\$day	-4.082e-04	3.692e-04	-1.106	0.268834	
Chor\$Yr1993:Chor\$day	-1.030e-03	4.940e-04	-2.085	0.037090	*
Chor\$Yr1994:Chor\$day	-8.493e-04	3.679e-04	-2.308	0.021000	*
Chor\$Yr1995:Chor\$day	8.534e-04	3.724e-04	2.292	0.021956	*
Chor\$Yr1996:Chor\$day	-4.582e-04	3.697e-04	-1.239	0.215261	
Chor\$Yr1997:Chor\$day	-9.635e-05	3.696e-04	-0.261	0.794348	
Chor\$Yr1998:Chor\$day	-3.186e-04	3.714e-04	-0.858	0.391117	
Chor\$Yr1999:Chor\$day	4.773e-04	3.703e-04	1.289	0.197464	
Chor\$Yr2000:Chor\$day	-3.322e-04	3.693e-04	-0.900	0.368377	
Chor\$Yr2001:Chor\$day	4.029e-04	3.907e-04	1.031	0.302542	
Chor\$Yr2002:Chor\$day	-4.629e-05	3.692e-04	-0.125	0.900249	
Chor\$Yr2003:Chor\$day	1.206e-03	4.899e-04	2.462	0.013838	*
Chor\$Yr2004:Chor\$day	-7.074e-04	3.660e-04	-1.933	0.053303	.
Chor\$Yr2005:Chor\$day	5.412e-04	4.875e-04	1.110	0.267056	
Chor\$Yr2006:Chor\$day	-9.274e-04	4.874e-04	-1.903	0.057108	.
Chor\$Yr2007:Chor\$day	-2.659e-04	4.872e-04	-0.546	0.585275	
Chor\$Yr2008:Chor\$day	2.876e-04	4.807e-04	0.598	0.549632	
Chor\$Yr2009:Chor\$day	-6.980e-05	4.875e-04	-0.143	0.886146	
Chor\$Yr2010:Chor\$day	-2.911e-04	5.741e-04	-0.507	0.612148	
Chor\$Yr2011:Chor\$day	8.299e-04	4.920e-04	1.687	0.091686	.
Chor\$Yr2012:Chor\$day	-2.946e-04	4.907e-04	-0.600	0.548263	
Chor\$Yr2013:Chor\$day	-7.376e-04	4.816e-04	-1.532	0.125681	
Chor\$Yr2014:Chor\$day	-4.317e-04	4.899e-04	-0.881	0.378222	
Chor\$Yr2015:Chor\$day	1.532e-04	4.934e-04	0.311	0.756180	
Chor\$Yr2016:Chor\$day	9.185e-04	4.934e-04	1.862	0.062670	.
Chor\$Yr2017:Chor\$day	-9.212e-04	4.878e-04	-1.888	0.059013	.
Chor\$Yr2018:Chor\$day	1.660e-03	4.966e-04	3.342	0.000835	***

 Residual standard error: 0.261 on 7858 degrees of freedom
 Multiple R-squared: 0.8836, Adjusted R-squared: 0.8821
 F-statistic: 590.4 on 101 and 7858 DF, p-value: < 2.2e-16

Call Lobith, historische reeks:

lm(formula = log(Chor\$chlvr) ~ log(log(Chor\$waterafvoerl)) +
 Chor\$Yr * Chor\$day + Chor\$Month + Chor\$weekdag)

Min 1Q Median 3Q Max
 -0.86550 -0.08538 0.00314 0.08769 0.66764 (Residuals)

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	5.954e+00	1.489e-01	39.995	< 2e-16	***
log(log(Chor\$waterafvoer1))	5.167e+00	5.622e-02	91.891	< 2e-16	***
Chor\$Yr1973	-8.020e-02	2.184e-02	-3.672	0.000242	***
Chor\$Yr1974	-3.587e-02	2.202e-02	-1.629	0.103302	
Chor\$Yr1975	-1.537e-01	2.208e-02	-6.963	3.51e-12	***
Chor\$Yr1976	-1.913e-01	2.165e-02	-8.835	< 2e-16	***
Chor\$Yr1977	-1.588e-01	2.199e-02	-7.220	5.51e-13	***
Chor\$Yr1978	-1.750e-01	2.150e-02	-8.141	4.30e-16	***
Chor\$Yr1979	-5.378e-02	2.140e-02	-2.513	0.011995	*
Chor\$Yr1980	-1.440e-01	2.128e-02	-6.766	1.39e-11	***
Chor\$Yr1981	-9.132e-02	2.126e-02	-4.295	1.76e-05	***
Chor\$Yr1982	-2.055e-01	2.143e-02	-9.592	< 2e-16	***
Chor\$Yr1983	-3.182e-01	2.170e-02	-14.665	< 2e-16	***
Chor\$Yr1984	-3.813e-03	2.131e-02	-0.179	0.857957	
Chor\$Yr1985	-5.656e-02	2.129e-02	-2.656	0.007910	**
Chor\$Yr1986	-7.286e-02	2.139e-02	-3.406	0.000662	***
Chor\$Yr1987	-2.229e-01	2.144e-02	-10.400	< 2e-16	***
Chor\$Yr1988	-2.881e-01	2.169e-02	-13.280	< 2e-16	***
Chor\$Yr1989	-1.730e-01	2.118e-02	-8.171	3.36e-16	***
Chor\$Yr1990	-1.159e-01	2.109e-02	-5.495	3.99e-08	***
Chor\$Yr1991	-1.961e-01	2.113e-02	-9.284	< 2e-16	***
Chor\$Yr1992	-1.402e-01	2.105e-02	-6.660	2.87e-11	***
Chor\$Yr1993	-3.630e-01	2.102e-02	-17.268	< 2e-16	***
Chor\$Yr1994	-4.205e-01	2.157e-02	-19.494	< 2e-16	***
Chor\$Yr1995	-4.182e-01	2.368e-02	-17.662	< 2e-16	***
Chor\$Yr1996	-3.877e-01	2.098e-02	-18.478	< 2e-16	***
Chor\$Yr1997	-3.868e-01	2.210e-02	-17.503	< 2e-16	***
Chor\$Yr1998	-3.486e-01	2.157e-02	-16.164	< 2e-16	***
Chor\$Yr1999	-6.423e-01	2.200e-02	-29.189	< 2e-16	***
Chor\$Yr2000	-6.804e-01	2.184e-02	-31.152	< 2e-16	***
Chor\$Yr2001	-6.953e-01	2.171e-02	-32.022	< 2e-16	***
Chor\$Yr2002	-5.576e-01	2.133e-02	-26.140	< 2e-16	***
Chor\$Yr2003	-6.752e-01	2.132e-02	-31.673	< 2e-16	***
Chor\$Yr2004	-6.393e-01	2.114e-02	-30.239	< 2e-16	***
Chor\$Yr2005	-5.307e-01	2.388e-02	-22.220	< 2e-16	***
Chor\$Yr2006	-5.689e-01	5.825e-02	-9.765	< 2e-16	***
Chor\$Yr2007	-8.082e-01	5.447e-02	-14.838	< 2e-16	***
Chor\$Yr2008	-9.183e-01	5.341e-02	-17.194	< 2e-16	***
Chor\$Yr2009	-8.855e-01	5.676e-02	-15.600	< 2e-16	***
Chor\$Yr2010	-8.302e-01	5.652e-02	-14.689	< 2e-16	***
Chor\$Yr2011	-8.473e-01	5.626e-02	-15.059	< 2e-16	***
Chor\$Yr2012	-8.949e-01	5.842e-02	-15.317	< 2e-16	***
Chor\$Yr2013	-7.726e-01	5.503e-02	-14.040	< 2e-16	***
Chor\$Yr2014	-9.772e-01	5.547e-02	-17.615	< 2e-16	***
Chor\$Yr2015	-7.761e-01	5.502e-02	-14.106	< 2e-16	***
Chor\$Yr2016	-8.593e-01	5.421e-02	-15.853	< 2e-16	***
Chor\$Yr2017	-8.443e-01	5.350e-02	-15.781	< 2e-16	***
Chor\$Yr2018	-8.747e-01	5.540e-02	-15.789	< 2e-16	***
Chor\$day	-3.800e-05	1.583e-04	-0.240	0.810294	
Chor\$Month02	4.256e-02	7.404e-03	5.749	9.21e-09	***
Chor\$Month03	1.088e-02	1.027e-02	1.060	0.289373	
Chor\$Month04	-7.789e-02	1.400e-02	-5.565	2.68e-08	***
Chor\$Month05	-1.132e-01	1.796e-02	-6.301	3.07e-10	***
Chor\$Month06	-1.695e-01	2.208e-02	-7.673	1.81e-14	***
Chor\$Month07	-2.431e-01	2.624e-02	-9.264	< 2e-16	***
Chor\$Month08	-3.359e-01	3.055e-02	-10.996	< 2e-16	***
Chor\$Month09	-1.931e-01	3.481e-02	-5.547	2.96e-08	***
Chor\$Month10	-1.504e-01	3.901e-02	-3.855	0.000117	***
Chor\$Month11	-1.391e-01	4.323e-02	-3.217	0.001298	**

Chor\$Month12	-1.276e-01	4.746e-02	-2.690	0.007162	**
Chor\$weekdagDinsdag	2.164e-02	4.650e-03	4.654	3.29e-06	***
Chor\$weekdagWoensdag	4.473e-02	4.627e-03	9.668	< 2e-16	***
Chor\$weekdagDonderdag	6.565e-02	4.650e-03	14.118	< 2e-16	***
Chor\$weekdagVrijdag	6.413e-02	4.658e-03	13.769	< 2e-16	***
Chor\$weekdagZaterdag	3.182e-02	4.685e-03	6.793	1.15e-11	***
Chor\$weekdagZondag	-2.634e-03	4.875e-03	-0.540	0.589006	
Chor\$Yr1973:Chor\$day	5.898e-04	1.029e-04	5.730	1.03e-08	***
Chor\$Yr1974:Chor\$day	4.735e-04	1.035e-04	4.574	4.83e-06	***
Chor\$Yr1975:Chor\$day	4.754e-04	1.032e-04	4.604	4.18e-06	***
Chor\$Yr1976:Chor\$day	7.944e-04	1.025e-04	7.749	1.00e-14	***
Chor\$Yr1977:Chor\$day	4.621e-04	1.028e-04	4.495	7.03e-06	***
Chor\$Yr1978:Chor\$day	6.166e-04	1.002e-04	6.154	7.80e-10	***
Chor\$Yr1979:Chor\$day	4.014e-04	9.967e-05	4.028	5.67e-05	***
Chor\$Yr1980:Chor\$day	8.069e-04	9.909e-05	8.143	4.24e-16	***
Chor\$Yr1981:Chor\$day	2.236e-04	9.896e-05	2.259	0.023878	*
Chor\$Yr1982:Chor\$day	7.002e-04	9.943e-05	7.042	2.00e-12	***
Chor\$Yr1983:Chor\$day	1.359e-03	1.010e-04	13.463	< 2e-16	***
Chor\$Yr1984:Chor\$day	2.820e-04	9.925e-05	2.842	0.004496	**
Chor\$Yr1985:Chor\$day	6.561e-04	9.977e-05	6.577	5.02e-11	***
Chor\$Yr1986:Chor\$day	4.028e-04	9.966e-05	4.042	5.33e-05	***
Chor\$Yr1987:Chor\$day	6.496e-04	9.953e-05	6.526	7.04e-11	***
Chor\$Yr1988:Chor\$day	1.108e-03	1.001e-04	11.073	< 2e-16	***
Chor\$Yr1989:Chor\$day	5.584e-04	9.955e-05	5.610	2.07e-08	***
Chor\$Yr1990:Chor\$day	4.825e-04	9.916e-05	4.866	1.16e-06	***
Chor\$Yr1991:Chor\$day	6.868e-04	9.949e-05	6.904	5.32e-12	***
Chor\$Yr1992:Chor\$day	2.464e-04	9.876e-05	2.494	0.012630	*
Chor\$Yr1993:Chor\$day	6.347e-04	9.888e-05	6.419	1.42e-10	***
Chor\$Yr1994:Chor\$day	4.369e-04	1.003e-04	4.356	1.34e-05	***
Chor\$Yr1995:Chor\$day	4.486e-04	1.070e-04	4.193	2.77e-05	***
Chor\$Yr1996:Chor\$day	7.384e-04	9.868e-05	7.482	7.83e-14	***
Chor\$Yr1997:Chor\$day	5.408e-04	1.030e-04	5.251	1.54e-07	***
Chor\$Yr1998:Chor\$day	1.043e-04	1.019e-04	1.024	0.305920	
Chor\$Yr1999:Chor\$day	7.824e-04	1.017e-04	7.696	1.52e-14	***
Chor\$Yr2000:Chor\$day	7.807e-04	1.022e-04	7.639	2.36e-14	***
Chor\$Yr2001:Chor\$day	6.997e-04	1.007e-04	6.951	3.82e-12	***
Chor\$Yr2002:Chor\$day	5.306e-05	9.987e-05	0.531	0.595196	
Chor\$Yr2003:Chor\$day	2.935e-04	1.005e-04	2.921	0.003493	**
Chor\$Yr2004:Chor\$day	1.771e-04	9.911e-05	1.787	0.074039	.
Chor\$Yr2005:Chor\$day	-1.037e-04	1.764e-04	-0.588	0.556603	
Chor\$Yr2006:Chor\$day	-5.130e-04	2.746e-04	-1.868	0.061741	.
Chor\$Yr2007:Chor\$day	1.497e-06	2.623e-04	0.006	0.995444	
Chor\$Yr2008:Chor\$day	4.514e-04	2.492e-04	1.811	0.070126	.
Chor\$Yr2009:Chor\$day	1.639e-04	2.624e-04	0.625	0.532035	
Chor\$Yr2010:Chor\$day	2.173e-04	2.622e-04	0.829	0.407253	
Chor\$Yr2011:Chor\$day	1.092e-04	2.623e-04	0.416	0.677228	
Chor\$Yr2012:Chor\$day	1.600e-04	2.685e-04	0.596	0.551249	
Chor\$Yr2013:Chor\$day	-4.064e-04	2.612e-04	-1.556	0.119828	
Chor\$Yr2014:Chor\$day	1.951e-04	2.623e-04	0.744	0.457036	
Chor\$Yr2015:Chor\$day	-4.962e-04	2.557e-04	-1.941	0.052305	.
Chor\$Yr2016:Chor\$day	5.063e-06	2.591e-04	0.020	0.984411	
Chor\$Yr2017:Chor\$day	-7.670e-05	2.603e-04	-0.295	0.768233	
Chor\$Yr2018:Chor\$day	1.331e-04	2.642e-04	0.504	0.614508	

Residual standard error: 0.1348 on 11791 degrees of freedom

Multiple R-squared: 0.8327, Adjusted R-squared: 0.8311

F-statistic: 528.8 on 111 and 11791 DF, p-value: < 2.2e-16

IV Bijlage IV. Dataverwerking van de recente, hoogfrequente data

Recente en hoogfrequente data voor Lobith en Eijsden voor debiet en chloride zijn verkregen via RWS servicedesk-data en Aqualarm, met ondersteuning van Marc Philippart (RWS-WVL). De chloride metingen en debiet gegevens zijn in aparte files verkregen. In deze bijlage wordt de verwerking van deze data beschreven.

Chlorideconcentratie gegevens

De chloride gegevens zijn van 2000-2021 en bestaan uit diverse meetstrategieën: “steekmonsters”, “Verzamelmonster tijdproportioneel gedurende 24 uur”, “Verzamelmonster tijdproportioneel gedurende 1 uur”, “Verzamelmonster tijdproportioneel”. Metingen met code “Zwevende stof” en “Verzamelmonster tijdproportioneel gedurende 4 weken” en metingen die hoger zijn dan 1000 mg/l of lager zijn dan 10 mg/l (Lobith) en hoger dan 300 mg/l (Eijsden) zijn *a priori* uitgesloten. Dit zijn onrealistische waarden en komen waarschijnlijk door een fout in de registratie.

Lobith

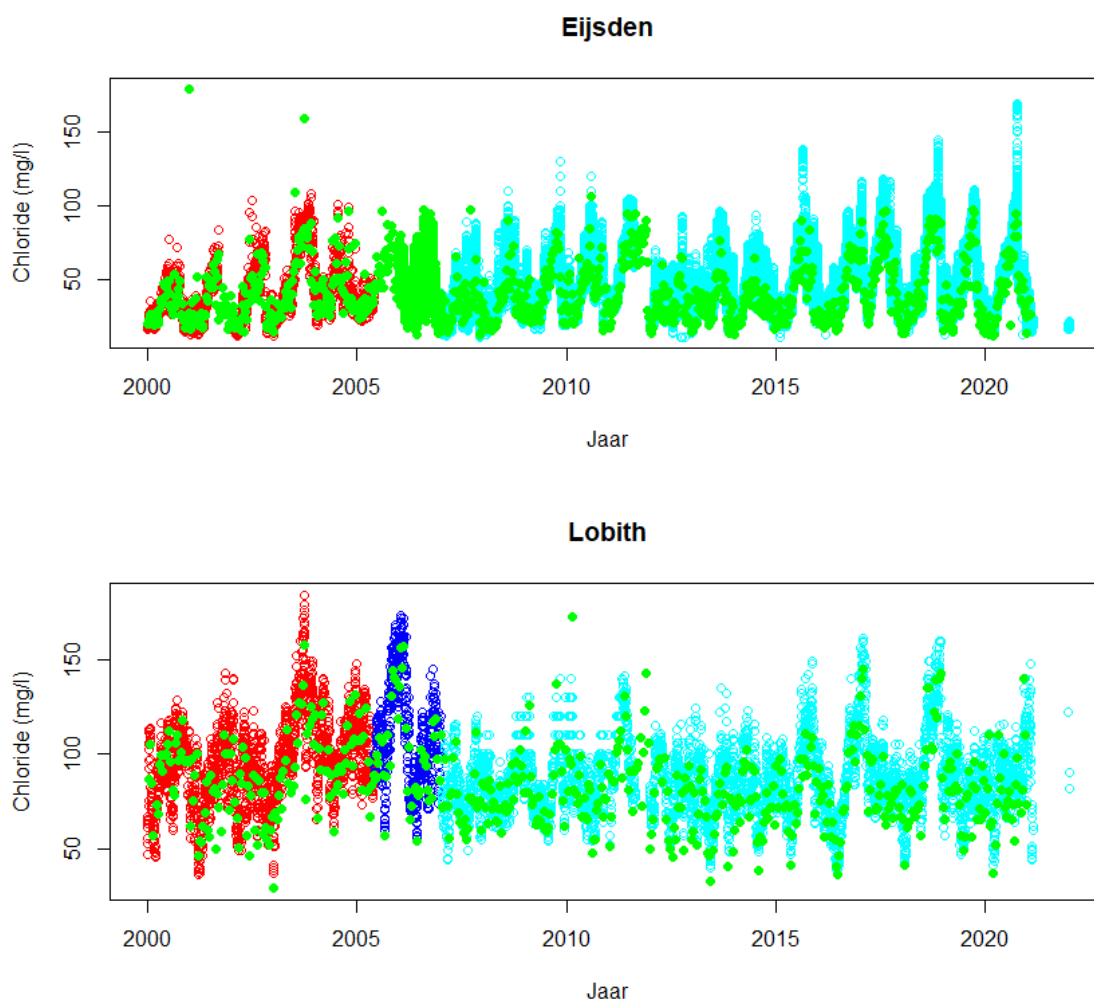
Alle verkregen metingen van meetstrategieën voor Lobith - behalve de steekmonsters - zijn (afgeleide) gemiddelden over 24 uur. Dit geldt ook de “Verzamelmonster tijdproportioneel gedurende 1 uur”. Dit is voor de zekerheid nog bevestigd door de data leverancier. Er wordt ook aangegeven in een andere kolom dat dit 24 uur gemiddelden zijn. Het is waarschijnlijk dat deze bepaald zijn via meerdere “Verzamelmonster tijdproportioneel gedurende 1 uur” metingen.

Eijsden

Voor Eijsden zijn de “Verzamelmonster tijdproportioneel” wel waarden wel per uur. Om de analyse voor Lobith en Eijsden gelijk te houden, worden de data voor Eijsden ook naar een 24 uur gemiddelde gebracht.

Meetstrategieën

De drie meetstrategieën, naast de steekmonsters, hebben elk een eigen naam: ‘Bepaling van chloride-ionselectieve meting’, ‘Chloride mbv AgNO₃ titratie’, ‘Titrimetrische bepaling Cl volgens Mohr’. Ook wordt de verwerking anders beschreven: “Verzamelmonster tijdproportioneel gedurende 1 uur”, “Verzamelmonster tijdproportioneel gedurende 24 uur”, “Verzamelmonster tijdproportioneel”. Of de manier van meten/berekenen van deze drie methoden invloed heeft op de hoogte van de chloride bepalingen, is niet te achterhalen omdat de perioden waarin deze metingen gedaan zijn niet overlappen (zie in Figuur 22 rode, lichtblauwe en donkerblauwe metingen. Groen zijn steekmonsters). Voor steekmonsters kunnen we wél bepalen of metingen via deze manier afwijken van de andere metingen, dat beschrijven we later. In Figuur 22 staan de data zoals verkregen (zonder de *a priori* uitgesloten waarden) uitgezet, elke meetstrategienaam is in een kleur aangegeven.



Figuur 22. Chlorideconcentratie data. Elke kleur is een andere meetmethode/strategie. Rood is "Verzamelmuster tijdproportioneel gedurende 1 uur", donkerblauw is "Verzamelmuster tijdproportioneel gedurende 24 uur", lichtblauw is "Verzamelmuster tijdproportioneel", groen is "Steekmonster". In deze plot zijn de uurgemiddelde waarden bij Eijsden nog niet teruggebracht tot een 24 uur gemiddelde. Bron: RWS servicedesk-data.

Steekmonsters (watermonsters die zijn geanalyseerd in het laboratorium op chloride) zijn gedurende de gehele periode (2000-februari 2021) genomen. Deze metingen zijn gemiddeld voor zowel Lobith als Eijsden lager dan de metingen met de andere methoden, en worden daarom uitgesloten van de analyse.

Dit heeft consequenties voor de resultaten wanneer deze vergeleken worden tussen de historische meetreeks (van RIWA-Rijn) en de hoogfrequente (dagelijkse) recente meetreeks van RWS. De RIWA data bevatten uitsluitend de steekmonsters en geven dus iets lagere chlorideconcentraties aan. Dat komt terug in de medianen: voor Lobith is de mediaan van de hoogfrequente, recente chlorideconcentratie data van RWS, na het terugbrengen van de data tot 24 uur gemiddelden en verwijderen van de steekmonsters, 86 mg/l en voor RIWA data is dat 79 mg/l. Voor Eijsden is de mediaan voor de hoogfrequente, recente data van RWS, na het terugbrengen van de data tot 24 uur gemiddelden en verwijderen van steekmonsters, 40,8 mg/l, voor RIWA data is dat 37 mg/l. In Figuur 23 staan de RIWA data zoals die zijn gebruikt als historische meetreeks gevisualiseerd samen met de daggemiddelde RWS data, ter illustratie. De waterafvoeren uit de eerder gebruikte RIWA data zijn hetzelfde als de 24 uur gemiddelden van de RWS data die voor de afleiding met hoogfrequente, recente data ook gebruikt worden. Voor de model fit betekent dit dat de relatie waterafvoer-chloridevracht die op de hoogfrequente recente RWS data zal worden afgeleid,

mogelijk iets steiler loopt dan de relatie op basis van historische meetgegevens. De gemeten chlorideconcentratie is immers hoger in de recente, hoogfrequente datareeks, bij dezelfde waterafvoer.

Debiet gegevens

De debiet gegevens (waterafvoer) zijn diverse series met diverse meetfrequenties; daggemiddelden, per 10 minuten, of per 60 minuten. Er zijn data voor 2000-2021. Niet alle meetfrequenties omvatten alle jaren, maar er is wel overlap. De tijdreeks per 60 minuten is voor Lobith niet compleet, en de perioden overlappen met de tijdreeks van daggemiddelde waarden. De per 60 minuten metingen worden daarom niet meegenomen in de regressie analyses. De per 10 minuten data zijn omgerekend naar gemiddelde dagwaarden.

Lobith

De chloride data bij Lobith zijn per dag. Voor dagen waar gegevens van twee meetfrequenties van debieten voorhanden waren, een daggemiddelde en het daggemiddelde van per 10 minuten, is de gemiddelde waarde van de twee meetfrequenties toegekend. De maximale afwijking tussen deze getallen is voor Lobith 2,5% (dit gaat slechts om enkele dagen), de gemiddelde afwijking is 0,02%.

Eijsden

Voor Eijsden is de maximale afwijking 0,4%, de gemiddelde afwijking is 0,01%. De daggemiddelde debiet waarden die via beide reeksen (daggemiddelden, per 10 minuten) zijn verkregen, liggen dus voor het overgrote deel niet ver uit elkaar.

Data na verwerking

Data worden afgekapt op eind 2020, omdat 2021 niet compleet is. De meetpunten in 2021 worden wel gebruikt als onafhankelijke toets voor de voorspellende waarde van de vernieuwde regressie (zie Figuur 27 in Bijlage V). Enkele datapunten in Eijsden zijn alsnog verwijderd omdat de combinatie waterafvoer-chloridevracht zeer afwijkend was. De medianen per jaar voor debiet en chloridevracht zijn weergegeven in Figuur 25 in Bijlage V. De data bevatten na verwijdering van steekmonsters en het middelen tot een 24 uur gemiddelde per dag, metingen voor bijna alle dagen in alle jaren. Daardoor zijn genoeg metingen per maand en per weekdag beschikbaar om de coëfficiënten van deze variabelen betrouwbaar vast te kunnen stellen (Tabel 6 t/m 8).

Tabel 6. Aantal metingen per jaar met 24 uur gemiddeld debiet én 24 uur gemiddelde chloride meting.

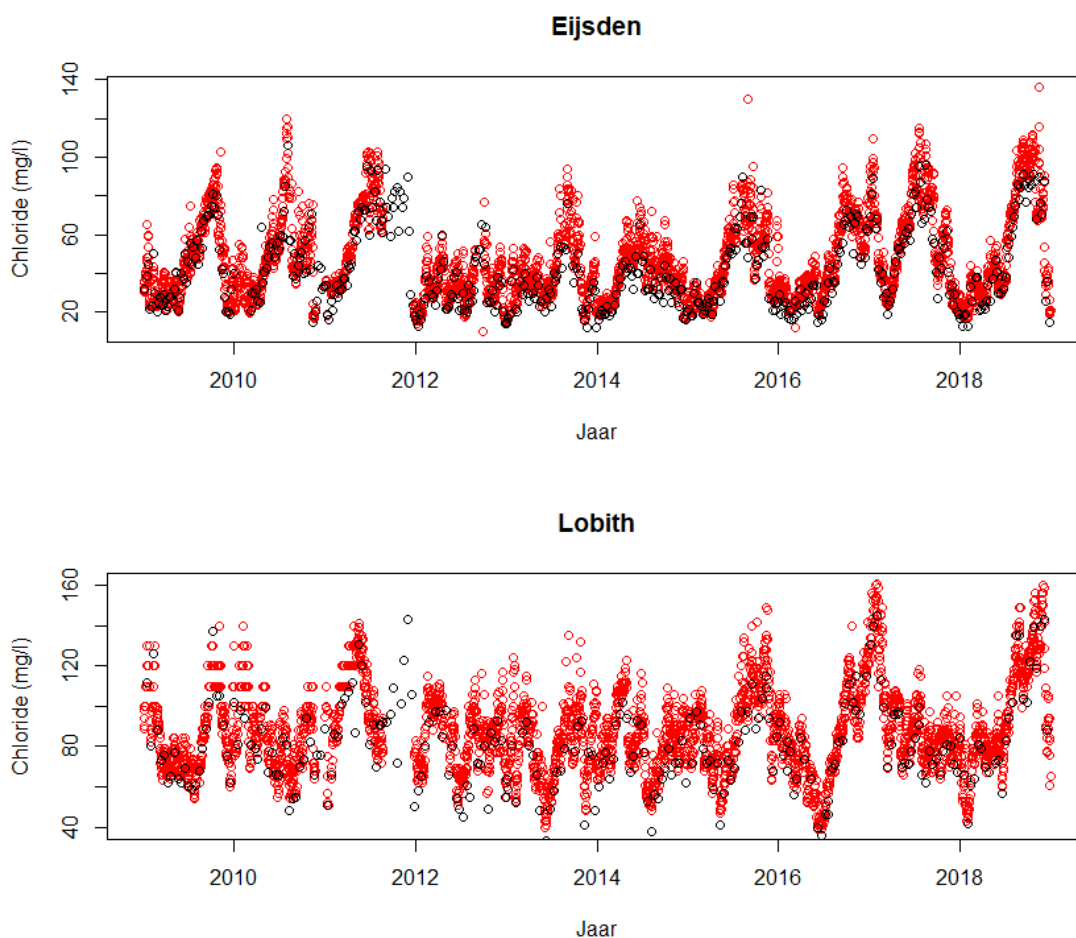
	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Eijsden	364	323	224	333	351	365	339	353	362	360	362	330
Lobith	361	323	226	347	365	363	356	366	365	365	365	346

Tabel 7. Aantal dagen met metingen per maand met 24 uur gemiddeld debiet én 24 uur gemiddelde chloride meting over de periode 2009-2020.

	jan	feb	mrt	apr	mei	jun	jul	aug	sep	okt	nov	dec
Eijsden	356	331	368	359	370	355	357	338	302	337	307	290
Lobith	359	337	372	353	357	358	369	362	326	340	321	293

Tabel 8. Aantal dagen met metingen per weekdag met 24 uur gemiddeld debiet én 24 uur gemiddelde chloride meting over de periode 2009-2020.

	maandag	dinsdag	woensdag	donderdag	vrijdag	zaterdag	zondag
Eijsden	582	586	587	583	581	577	574
Lobith	593	587	592	594	595	593	593



Figuur 23. De vergelijking tussen de RIWA-Rijn chlorideconcentratie data (zwarte cirkels) van de historische reeks en RWS chlorideconcentratie data (rode cirkels, 24 uur gemiddelden). De RWS data is hoogfrequenter (hier teruggebracht tot dagelijks) dan de RIWA data (wekelijks in Eijsden, tweewekelijks in Lobith). De RWS data hebben gemiddeld gezien een hogere waarde dan de RIWA-Rijn data (steekmonsters).

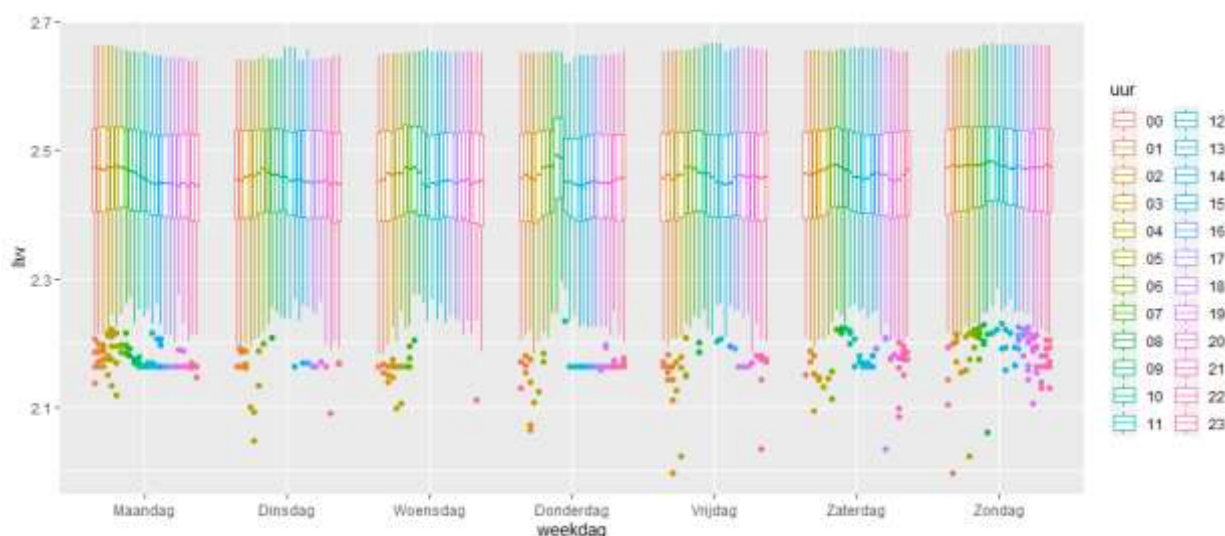
Verschil 1 uur- en 24 uurgemiddelde data Eijsden

Voor locatie Eijsden zijn er vanaf 2009, voor zowel chloride als waterafvoer, meer gedetailleerde gegevens gedurende een dag. Voor waterafvoer zijn dit 10 minuten gegevens. Voor chlorideconcentratie zijn dat 1 uur gemiddelden. Dat geeft de mogelijkheid om te kijken hoe de voorspelling op een meer gedetailleerde schaal is (1 uur gemiddelde in plaats van 24 uur gemiddelde). We middelen daarvoor de 10 minuten waterafvoer data tot een uurgemiddelde. We kijken of het model op een meer gedetailleerde schaal dezelfde fit en prestatie heeft. Voor Lobith kunnen we dit niet doen, omdat de data hiervoor ontbreken.

Weekdag variatie

Voor de uur gemiddelde data van locatie Eijsden is de invloed van de weekdays significanter dan bij 24 uur gemiddelde data (niet getoond), met name zondag. Dit kan zijn omdat er meer data zijn, waardoor een klein verschil sneller significant wordt. Het effect op de prestatie van het model is nog steeds heel klein, waardoor dit alsnog weggelaten kan worden in een model fit.

De oorzaak van de weekdag verschillen in chloridevracht in Eijsden worden veroorzaakt door de waterafvoer, die verschilt in hoogte van dag tot dag in de week en ook zelfs van uur tot uur. De chloridevracht stijgt met deze waterafvoer. De waterafvoer is het hoogste rond 08.00 uur in de ochtend. Op zondag is de waterafvoer over het algemeen hoger. Op maandag is er vervolgens een relatief groot verval (zie Figuur 24). Er is dus zeker variatie binnen 24 uur in chlorideconcentratie op basis van waterafvoer. Maar, de spreiding is ook heel groot.



Figuur 24. De variatie in $\log(\log)$ waterafvoer mg/s van uur tot uur, uitgesplitst naar weekdag voor locatie Eijsden (2009-2020). Per uur is een boxplot weergegeven met de $\log(\log)$ waterafvoer voor dat uur op die dag. Bron: RWS servicedesk-data.

De model prestatie

Bij een laat-een-jaar-weg kruisvalidatie (500 keer) heeft het model gebaseerd op uurgemiddelden voor locatie Eijsden een MAPE van 22,1. Dit is iets slechter dan het model bij de 24 uur gemiddelden (20,3).

Ten opzichte van de modellen van Van der Wijk (2018) toegepast op uurgemiddelde data blijft de vernieuwde formule ook beter: zonder kruisvalidatie is de prestatie op basis van MAPE 23,0 (Van der Wijk) ten opzichte van 20,9 (geactualiseerd). Voor recente data (2019-2020) is dit 21,4 (Van der Wijk) ten opzichte van 17,8. Voor relatief hoge chloride (>58) is dit 23,9 (Van der Wijk) ten opzichte van 20,3.

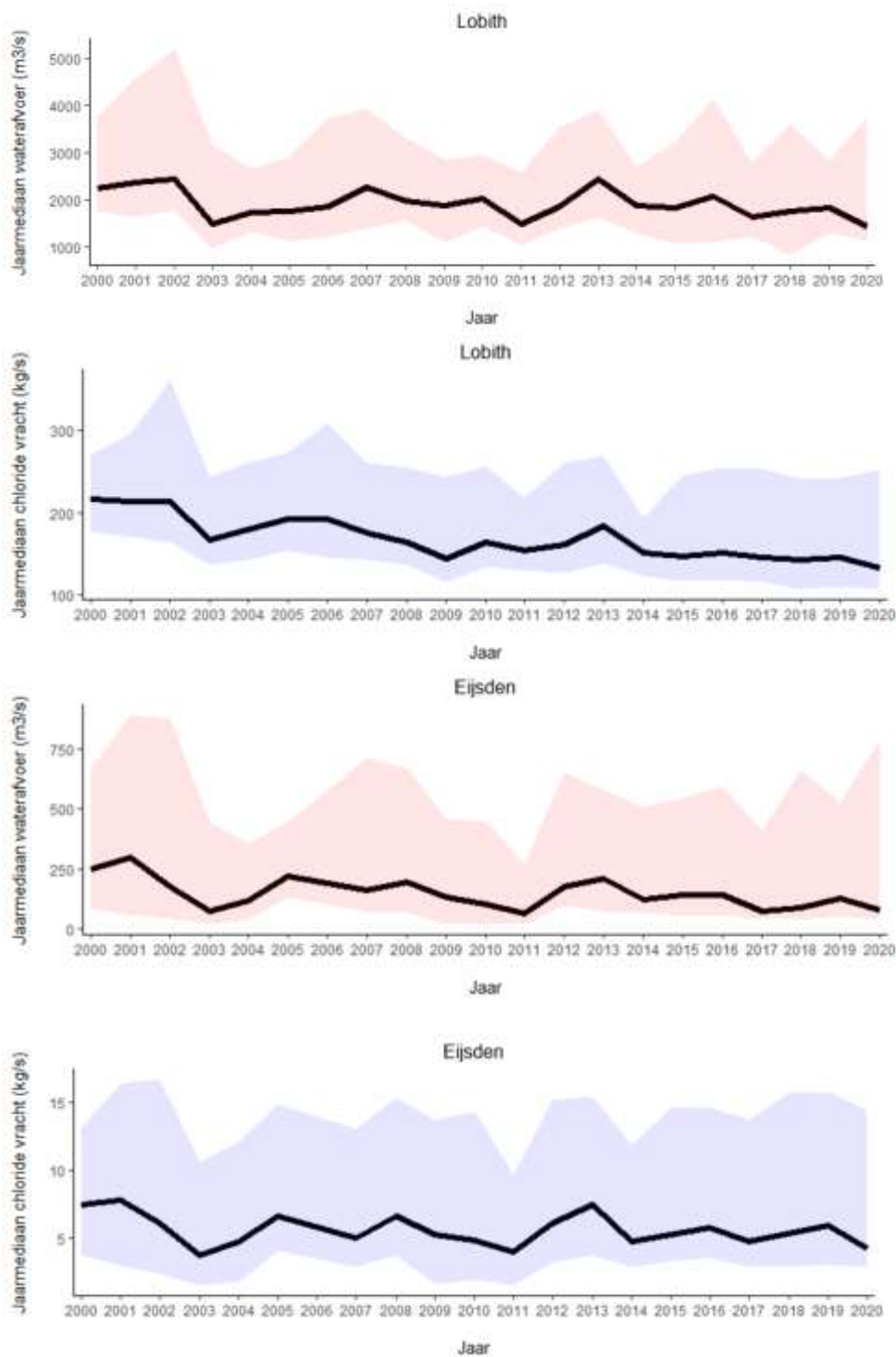
Tabel 9. Model prestaties van modellen op basis van 1 uur en 24 uur gemiddelden van chloride en waterafvoer, op data van 1 uur en 24 uur gemiddelden. Waarden zijn gegenereerd zonder kruisvalidatie.

	Model 1 uur gemiddelde MAPE	Model 24 uur gemiddelde MAPE
Data chloride 2009-2020 (1 uur gemiddelde)	20,9	21,2
Data chloride 2009-2020 (24 uur gemiddelde)	20,1	19,2

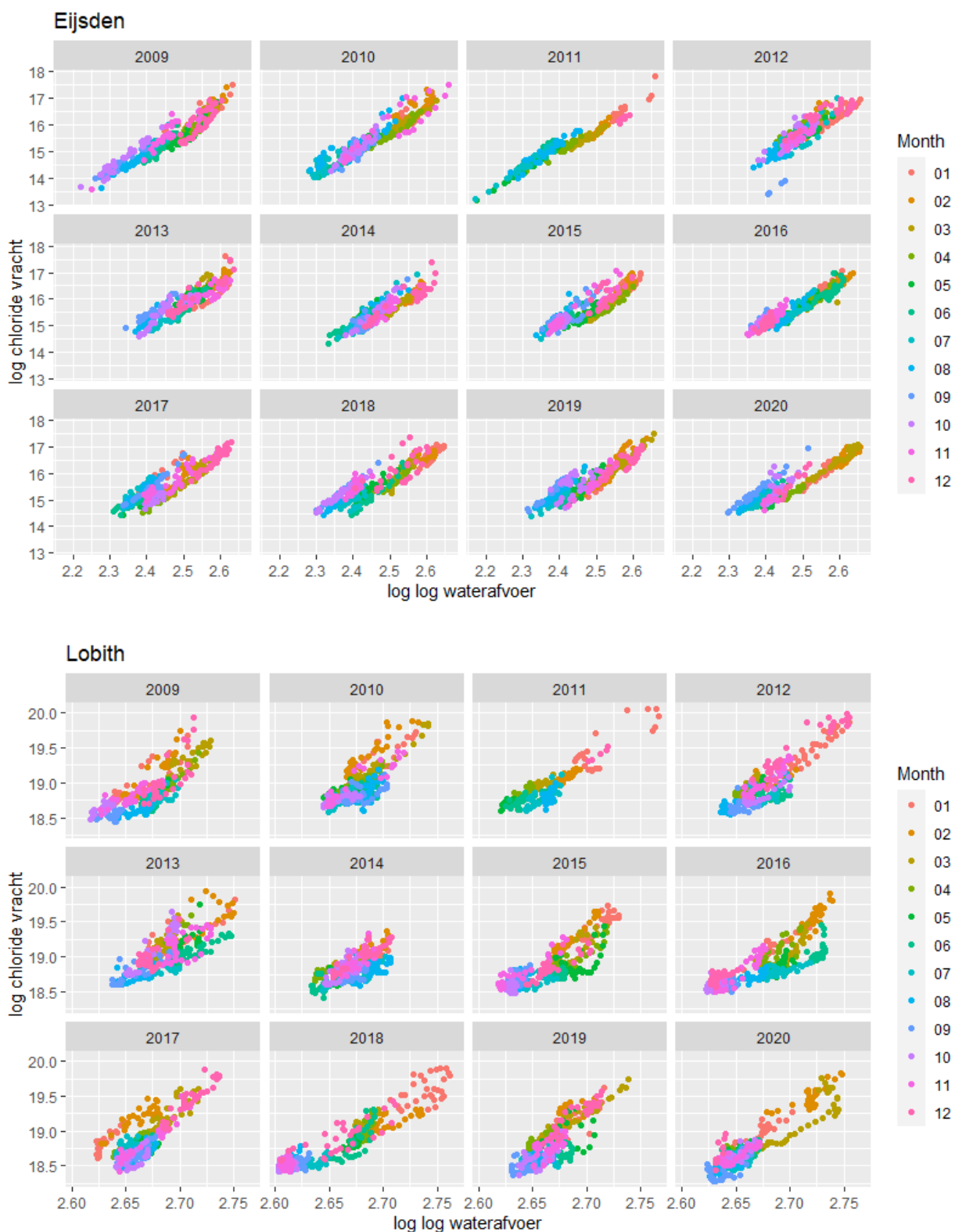
Uit Tabel 9 blijkt dat de modellen, zoals ook te verwachten zou zijn, hun eigen data het best kunnen voorspellen. Als de voorspelling van chloride per uur gedaan zou worden, liggen de prestaties van de modellen niet ver uit elkaar (bovenste rij). Als de voorspellingen voor zowel 1 uur als 24 uur gedaan gaan worden, dan is het 24 uur model gemiddeld iets beter (bij vergelijk kolommen).

De reden dat de 1 uur gemiddelde chloride over het algemeen slechter voorspeld kan worden (rij een ten opzichte van rij twee) is mogelijk veroorzaakt doordat waterafvoeren niet heel nauwkeurig vastgesteld kunnen worden, vooral als deze waterafvoer laag is. De waterafvoer wordt in sommige metingen zelfs vastgesteld als negatief (niet getoond). Deze onzekerheid wordt er iets uitgemiddeld als er uur gemiddelden worden genomen, en nog verder als er 24 uur gemiddelden worden genomen, wat de data robuuster maakt. Voor een voorspelling maakt het feit dat waterafvoer slecht gemeten kan worden niet uit, omdat de waterafvoer zelf ook voorspeld zal worden. Het maakt wel iets uit voor het predictie-interval. Het predictie interval is ongeveer 5% (hoge waarden) tot 10% (lage waarden) breder in het een uur gemiddelde model ten opzichte van het 24 uur gemiddelde model (niet getoond). Dat betekent dat de nauwkeurigheid van de voorspelling lager is. Als voorbeeld: het 90 percentiel predictie-interval is in het 24 uur gemiddelde model voor een bepaalde voorspellingswaarde 28,1 breed, in het uur gemiddelde model is dat 29,4 breed.

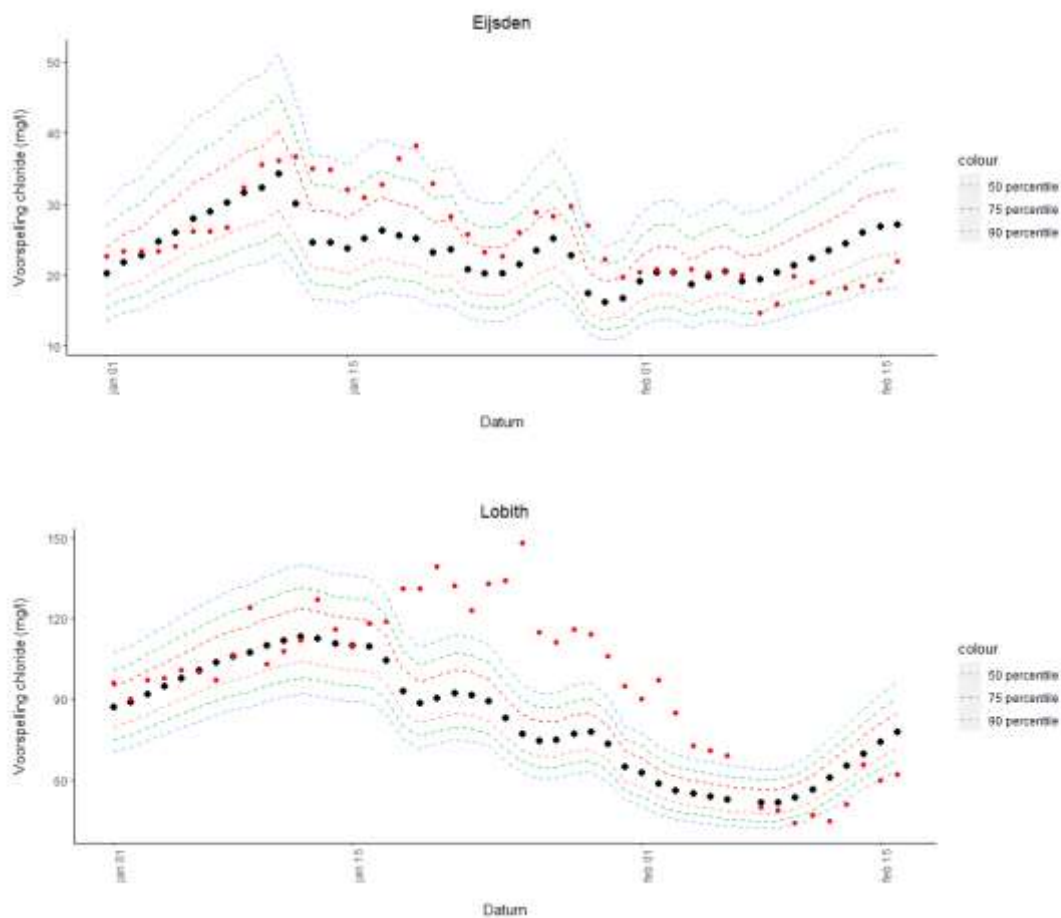
V Bijlage V. Extra figuren



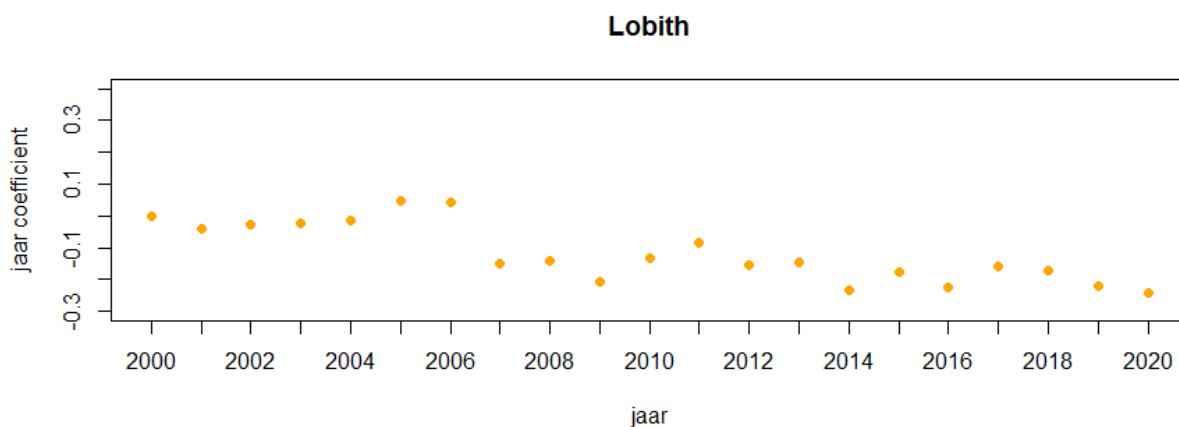
Figuur 25. Chloridevracht en waterafvoer voor Lobith en Eijsden voor hoogfrequente, recente data. De gekleurde band laat het 90 percentiel van alle metingen in het betreffende jaar zien en de zwarte lijn weerspiegelt de mediaan. Bron: RWS servicedesk-data.

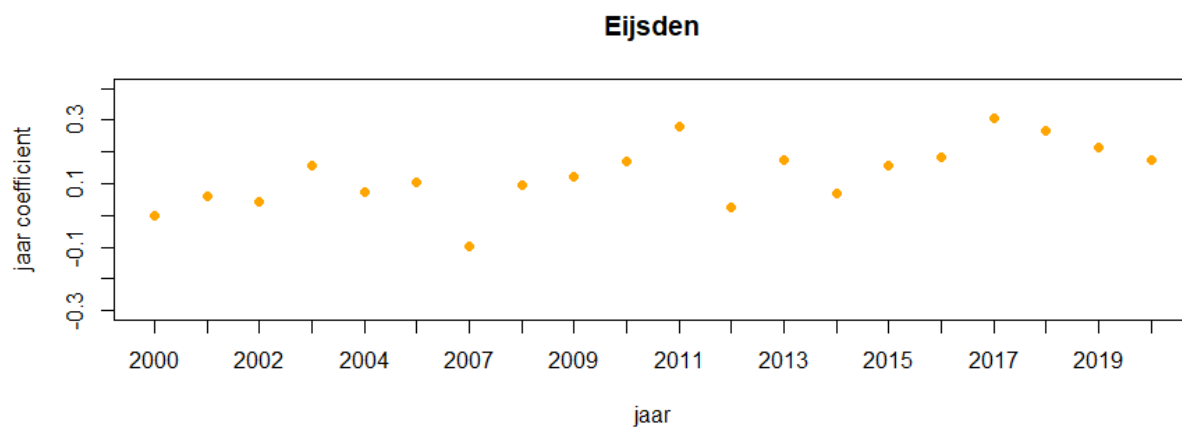


Figuur 26. De relatie log chloridevracht (mg/l) tegen de log(log) waterafvoer (l/s), per jaar voor hoogfrequente, recente data. De waarden per maand zijn aangegeven in verschillende kleuren. Bron: RWS servicedesk-data.



Figuur 27. Voorbeeld van de voorspelling van het model dat gefit is op recente, hoogfrequente data van 2009-2020 met interactieterm (zwarte cirkels) in vergelijking met werkelijke chlorideconcentratie data (rode cirkels) in 2021, met predictie-intervallen. Voor zowel Lobith als Eijsden is in de tweede helft van januari de werkelijke chlorideconcentratie hoger dan de voorspelde. Dit zou veroorzaakt kunnen zijn door het strooien van zout in combinatie met regen in die periode. De periode van 15 januari tot 8 februari had in ieder geval veel regenval (niet getoond). Daarna ging het vriezen, waardoor de neerslag is blijven liggen op het land, dit is mogelijk de oorzaak van de overschatting rond half februari. Bron: RWS servicedesk-data.





Figuur 28. Jaar coëfficiënten voor de fit jaar+ maand+ weekday, periode 2000-2020 op hoogfrequente, recente data. Zonder weekday zien de coëfficiënten er nagenoeg hetzelfde uit (niet getoond). Bron: RWS servicedesk-data.