



**inbo**



Instituut voor  
Natuur- en Bosonderzoek

## **Opstellen van een distributiekaart van zware metalen in sigmagebieden langsheen Zeeschelde, Rupel en Nete**

*Pieter Verschelde, Thierry Onkelinx, Suzanna Lettens en Bruno De Vos*

**Auteurs:**

Pieter Verschelde, Thierry Onkelinx, Suzanna Lettens en Bruno De Vos  
*Instituut voor Natuur- en Bosonderzoek*

Het Instituut voor Natuur- en Bosonderzoek (INBO) is het Vlaams onderzoeks- en kenniscentrum voor natuur en het duurzame beheer en gebruik ervan. Het INBO verricht onderzoek en levert kennis aan al wie het beleid voorbereidt, uitvoert of erin geïnteresseerd is.

**Vestiging:**

INBO Geraardsbergen  
Gaverstraat 4, 9500 Geraardsbergen  
www.inbo.be

**e-mail:**

suzanna.lettens@inbo.be

**Wijze van citeren:**

Verschelde P., Onkelinx T., Lettens S., De Vos B. (2013). Opstellen van een distributiekaart van zware metalen in sigmagebieden langsheen Zeeschelde, Rupel en Nete. Rapporten van het Instituut voor Natuur- en Bosonderzoek 2013 (11). Instituut voor Natuur- en Bosonderzoek, Brussel.

**Revisor:**

Maarten Jans

**D/2013/3241/048**

**INBO.R.2013.11**

**ISSN: 1782-9054**

**Verantwoordelijke uitgever:**

Jurgen Tack

**Druk:**

Managementondersteunende Diensten van de Vlaamse overheid

**Foto cover:**

Paardeweide, Berlare  
Y. Adams/Vildaphoto.net

**Dit onderzoek werd uitgevoerd in opdracht van:**

Waterwegen & Zeekanaal NV, Afdeling Zeeschelde  
Anna Bijnsgebouw - Lange Kievitstraat 111/113 bus 44, 2018 Antwerpen  
www.wenz.be



Waterwegen en Zeekanaal NV  
**weg van water**

# **Opstellen van een distributiekaart van zware metalen in sigmagebieden langsheen Zeeschelde, Rupel en Nete**

**Pieter Verschelde, Thierry Onkelinx, Suzanna Lettens &  
Bruno De Vos**



## Samenvatting

Voorliggend rapport bekijkt hoe de overstromingsgebieden uit het geactualiseerd Sigmaplan (Beslissing Vlaamse Regering 2005 en 2006) optimaal bemonsterd kunnen worden voor het vastleggen van de referentietoestand en nadien voor het opvolgen van de bodemkwaliteit in de tijd. INBO bemonsterde in het verleden 11 sigmagebieden, waarvan een deel reeds onder invloed stond van een aantal recente overstromingen (en bijhorende alluviale afzettingen) en andere gebieden die de laatste decennia niet overstroomd zijn. De staalnames, die gebaseerd waren op een regelmatig grid met 1 bodemstaal per hectare, dienden als basis om de ruimtelijke variatie van zware metalen en eigenschappen van de bodem binnen en tussen de gebieden te onderzoeken.

In eerste instantie gebeurt een algemene screening van de data. De fysico-chemische variabelen die in het labo geanalyseerd werden, worden grafisch voorgesteld. Hier blijkt dat er grote verschillen bestaan tussen gebieden, zowel qua elementconcentraties als qua andere bodemeigenschappen (textuur, organisch materiaal, pH, ...). Ook de spreiding verschilt sterk van gebied tot gebied. Sigmagebieden die momenteel reeds in gebruik zijn als overstromingsgebied bevatten meestal meer klei en organisch materiaal evenals hogere concentraties zware metalen.

In een tweede fase wordt gekozen om de modelbouw en interpolatie uit te voeren voor het zware metaal zink. Alle zware metalen zijn sterk gecorreleerd en zink heeft een brede range en geen enkele waarde beneden de bepaalbaarheidsgrens. De conclusies voor zink zijn gezien de sterke samenhang tussen alle zware metalen ook grotendeels geldig voor deze andere zware metalen. Er vindt een selectie plaats van de mogelijke verklarende variabelen, gebaseerd op verschillende criteria, met name een maximale relevantie van de variabele (bv. milieufactoren of bodemeigenschappen die nodig zijn voor de berekening van Vlarebo normen) en een minimale correlatie met de andere verklarende variabelen. De volgende verklarende variabelen worden weerhouden: kleigehalte, pH, LOI, gehalten aan sulfaat, kalium, fosfor en plantbeschikbare fosfor, hoogteligging, afstand tot de rivier en X,Y-coördinaten. Per gebied wordt een model opgesteld en nagekeken welke van deze variabelen significant bijdragen om de responsvariabele zink te verklaren. Hieruit blijkt dat klei en/of organisch materiaal (die onderling vaak positief gecorreleerd zijn) bijna altijd positief gecorreleerd zijn met het gehalte aan zink. Hetzelfde geldt voor de pH. Vaak is er ook een negatieve correlatie met de hoogteligging (dit is de relatieve hoogte, waarbij het laagste punt per gebied gelijk wordt gesteld aan hoogte 0). In een aantal gebieden komt tenslotte een positieve correlatie voor tussen nutriëntgehalte (fosfor, plantbeschikbare fosfor, kalium) en zinkgehalte.

Behalve een statistisch model wordt per gebied ook een ruimtelijke interpolatie van het zinkgehalte uitgevoerd op basis van een variogram. De ruimtelijke afhankelijkheid van het zinkgehalte in de bodem verschilt sterk tussen gebieden, zowel wat betreft de korte-afstandscorrelatie als hoe deze autocorrelatie afneemt met toenemende afstand. De range (maximale afstand waarop punten geautocorreleerd zijn) varieert van minder dan 300 m in Anderstadt-Lier tot meer dan 900 m in Wijmeers 1. Ook Bovenzanden en De Bunt hebben een grote range. De nugget (de variantie die verwacht wordt tussen 2 punten die zeer dicht bij elkaar liggen) en de sill (de maximale spatiale variantie) variëren sterk van gebied tot gebied. Tielrodebroek heeft een zeer lage nugget en zeer lage sill. Anderstadt-Lier daarentegen heeft een veel hogere nugget en Wijmeers I een veel hogere sill. De overstroomde gebieden hebben over het algemeen hogere concentraties zink en vertonen een hogere spatiale variabiliteit.

Er werd per gebied ook een kaart opgesteld met de kans op overschrijding van de Vlarebo bodemsaneringsnorm type 1 en 2. Deze kaart werd opgesteld aan de hand van indicator kriging, wat wil zeggen dat de kans op overschrijding gemodelleerd wordt aan de hand van een 1-0 grid gebaseerd op het werkelijke voorkomen van normoverschrijdingen in de meetpunten. Een variogram interpoleert dit grid naar alle tussenliggende punten.

Er werd ook getracht een algemeen variogram op te stellen, over alle gebieden heen, maar dit bleek moeilijk. De voorspelde waarden verschillen voor een aantal gebieden sterk van deze afgeleid uit de lokale variogrammen, in het bijzonder voor De Bunt, Anderstadt-Lier en Paardeweide. Ook het onderscheid tussen overstromde en niet-overstromde gebieden bemoeilijkt het opstellen van één allesomvattend spatiaal model.

Daarnaast wordt ook een algemeen model opgesteld, dat alle gebieden samen bekijkt. Om het effect van het gebied zelf buiten beschouwing te laten, wordt van alle variabelen inclusief de respons het gemiddelde per gebied afgetrokken. Heel wat van de variabelen blijken duidelijk significant te zijn in het algemeen model. De enige effecten die wegvallen zijn alle variabelen die de afstand tot de rivier gebruiken en de interactie tussen klei en het al dan niet overstromd zijn. Bij klei is ook organisch materiaal (LOI) mee gevarieerd, omdat deze vrij sterk gecorreleerd zijn met elkaar. Klei+LOI en logP hebben het sterkste effect, pH en hoogte hebben een kleiner effect. Het effect van pH, hoogte en logP is sterker uitgesproken voor overstromde gebieden. Het effect van klei+LOI is sterker voor niet-overstromde gebieden.

Een reductie in het aantal te analyseren parameters betekent een besparing voor de inventarisatie. Zo zou bijvoorbeeld de meting van zware metalen vervangen kunnen worden door een proxy die een goede benadering van het zware metaal gehalte oplevert. Of een alternatieve, goedkopere variabele zou kunnen opgemeten worden. De mogelijkheden van een dergelijke winst in efficiëntie zijn echter beperkt. Ten eerste blijkt uit de analyse dat de staalnamekost een groot deel uitmaakt van de totale kost (verplaatsingskosten en werkuren) zodat de labokost een relatief klein aandeel van de totale kost vertegenwoordigt. Ten tweede zijn er geen variabelen die voldoende sterk gecorreleerd zijn met de concentratie zware metalen om dienst te doen als volwaardige proxy. Echter, een variabele die met hogere dichtheid en lage kost wordt opgemeten en goed gecorreleerd is met het zware metaalgehalte, zoals electro-magnetische inductie metingen met bv de EM-38, zou toch nuttig kunnen zijn in deze context, maar is niet onderzocht in deze studie.

Verschillende manieren van steekproeftrekking worden met elkaar vergeleken. Hieruit blijkt dat 100 staalnamepunten per gebied een absoluut minimum is voor het vastleggen van de referentietoestand, bij voorkeur gelegen volgens een patroon dat een regelmatig grid combineert met een deel punten die zeer dicht bij elkaar liggen, om zo ook inzicht te krijgen in de variatiepatronen op korte afstand (nugget).

Tenslotte worden concrete aanbevelingen geformuleerd. De referentietoestand van gebieden dient steeds vastgelegd te worden vóór ingebruikname als overstromingsgebied. Als er beslist wordt om de gebieden in de tijd op te volgen, is het van groot belang dat de meetpunten niet verplaatsen in de tijd. Om de concentraties van de zware metalen in kaart te brengen, blijft de meest aangewezen methode om gewoon de concentratie van het metaal zelf op te meten, en deze dan te interpoleren over het gebied. Omdat de bodemsaneringsnormen gebaseerd zijn op formules die de pH, het kleigehalte en de hoeveelheid organische materiaal bevatten, moeten deze drie variabelen steeds opgemeten worden.

# Inhoud

<b>Samenvatting</b> .....	<b>5</b>
<b>1 Inleiding</b> .....	<b>9</b>
1.1 Situering .....	9
1.2 Doelstelling.....	10
<b>2 Data beschrijving</b> .....	<b>12</b>
2.1 FloodGis databank.....	12
2.2 GIS lagen .....	14
2.3 Algemene screening van de data .....	14
2.3.1 Chemische variabelen per gebied .....	14
2.3.2 Gebieden.....	19
2.3.3 Bepalen van de responsvariabele.....	20
<b>3 Methodologie</b> .....	<b>22</b>
3.1 Data manipulaties .....	22
3.2 Basisverkenning .....	22
3.3 Gebiedspecifieke analyse.....	23
3.3.1 Model.....	23
3.3.2 Kaarten .....	26
3.4 Algemene analyse over de gebieden heen .....	28
3.5 Kosten .....	28
3.5.1 Sampling strategie .....	28
3.5.2 Cokriging analyse .....	29
3.5.3 Kostenraming .....	29
<b>4 Gebiedspecifieke analyse</b> .....	<b>30</b>
4.1 Basisverkenning .....	30
4.2 Nog niet overstroomde gebieden .....	31
4.2.1 Wijmeers I.....	31
4.2.2 De Bunt .....	34
4.2.3 Kruibeke - Bazel - Rupelmonde .....	38
4.2.4 Vlassenbroekse polder .....	42
4.2.5 Grote Wal, Kleine Wal en Zwijn .....	44

4.3	Reeds overstroomde gebieden .....	47
4.3.1	Anderstadt I en II en de polder van Lier .....	47
4.3.2	Bovenzanden .....	50
4.3.3	Paardeweide .....	53
4.3.4	Tielrodebroek.....	56
4.4	Algemene verschillen en gelijkenissen tussen gebieden .....	59
4.5	Alternatieve analysemethodes.....	65
4.5.1	Cokriging.....	65
4.5.2	Stochastische methode voor kans op overschrijding norm .....	66
<b>5</b>	<b>Algemene analyse .....</b>	<b>69</b>
5.1	Resultaten van het veralgemeend model .....	69
5.2	De impact op de predicties .....	71
<b>6</b>	<b>Kostenramingen .....</b>	<b>74</b>
6.1	Staalname strategie.....	74
6.2	Cokriging.....	76
6.3	Kostenraming .....	78
<b>7</b>	<b>Conclusies en aanbevelingen .....</b>	<b>81</b>
	<b>Lijst van figuren.....</b>	<b>82</b>
	<b>Lijst van tabellen .....</b>	<b>85</b>
	<b>Literatuurlijst.....</b>	<b>86</b>



# 1 Inleiding

## 1.1 Situering

In opdracht van de Afdeling Zeeschelde van Waterwegen en Zeekanaal NV (W&Z) onderzoekt INBO de aanwezigheid van verontreiniging in Sigmagebieden langs de bevaarbare waterlopen. Er worden bodemstalen genomen van de bovenste laag en in sommige gevallen ook van diepere lagen. Deze stalen worden geanalyseerd op zware metalen. Indien mogelijk worden de gebieden bemonsterd vóór ingebruikname als overstromingsgebied. Bedoeling is in eerste instantie om de referentietoestand ( $T_0$ ) vast te leggen en in tweede instantie om na te gaan hoe de bodemkwaliteit evolueert onder invloed van de afzetting van riviersedimenten tijdens overstromingen.

Het Sigmaplan werd opgemaakt in 1977 en bevat 13 gecontroleerde overstromingsgebieden (GOG). Deze gebieden overstroomden door middel van een overlooppolder bij abnormale hoogwaterstanden, zoals bij stormvloed. Van zodra de waterstand in de rivier daalt, stroomt het GOG via uitwateringssluizen geleidelijk aan terug leeg. Het doel van deze gebieden is het aftoppen van een stormtjgolf. Het Sigmaplan werd geactualiseerd in 2006 (zoals goedgekeurd door de beslissingen van de Vlaamse Regering van 22/7/2005 en 28/4/2006). De 13 oorspronkelijke gebieden werden aangevuld met 55 nieuwe en (delen van) een aantal gebieden kregen een nieuwe invulling, bijvoorbeeld als gecontroleerd gereduceerd getijdengebied (GGG) of wetland. Een GGG is een gebied waar het getij dagelijks doordringt, zij het met verminderde amplitude, dus "gedempt" via een inwateringssluis. Deze inrichting moet instaan voor de ontwikkeling van estuariene natuur. Een wetland is een natuurinrichting met specifiek aandacht voor waterrijke systemen. De ontwikkeling van een wetland gebeurt vaak, maar niet noodzakelijk, binnen een GOG.

De nieuwe sigmagebieden uit het geactualiseerd Sigmaplan (Beslissing Vlaamse Regering 2005 en 2006) functioneren momenteel nog niet als overstromingsgebied. De inrichtingswerken zijn in sommige gebieden wel al begonnen. Er wordt bij de opdeling van sigmagebieden een onderscheid gemaakt naargelang de voorziene startdatum van de inrichtingswerken. Zo zouden de inrichtingswerken van de 2010 sigmagebieden moeten gestart zijn in 2010. De gebieden uit het Sigmaplan (1977) functioneren wel reeds als overstromingsgebied (met uitzondering van KBR). De inrichting van deze gebieden zal binnen het geactualiseerd Sigmaplan echter vaak nog aangepast worden. Deze gebieden worden bijgevolg opgedeeld in 2010, 2015, 2020,... sigmagebieden naar gelang de aanvangsdatum van de inrichtingswerken.

INBO bemonsterde in opdracht van W&Z in eerste instantie de 2010 sigmagebieden die als GOG en GGG zullen worden ingericht. Later werd ook gestart met de 2015 sigmagebieden met GOG of GGG inrichting. Er werd gekozen voor een bemonstering volgens een regelmatig grid van 100 x 100 m<sup>2</sup>, ofwel 1 staalnamepunt per hectare. Op deze punten werd een staal genomen van de bovenste 10 cm bodem door vier keer te steken met een guts binnen een cirkel met een straal van 1 m. In sommige gebieden werden ook diepere lagen (20-30 cm, 50-60 cm en 100-110 cm) bemonsterd op een beperkte subset van de punten. Met deze staalname werd de referentietoestand van de overstromingsgebieden vastgelegd, en dit vóór het in werking treden van de plannen zoals ze vastgelegd zijn in het geactualiseerd Sigmaplan (Beslissing Vlaamse Regering 2005 en 2006). Het is ook de bedoeling om de afzetting van nieuwe sedimenten en mogelijke veranderingen in de verontreinigingsgraad van deze gebieden op te volgen na ingebruikname als GOG of GGG. Daarnaast wordt ook een ecologische risico-evaluatie van de verontreinigde gebieden uitgevoerd, waarbij de

bodemconcentraties worden gerelateerd aan de bladconcentratie van specifieke indicatorplanten om zo een globale uitspraak te kunnen doen over het ecologisch risico van de bodemverontreiniging. Bepaalde bodemeigenschappen en milieufactoren kunnen immers leiden tot het meer of minder biobeschikbaar worden van de polluenten voor opname in planten en bodemorganismen.

Voor het vaststellen van verontreiniging voorziet de OVAM het Compendium voor Monsterneming en Analyse (CMA), dit is een bundel van monsternemings- en analysemethoden. Om een rechtsgeldige uitspraak te kunnen doen over de kwaliteit van bodem ten opzichte van de toepasselijke wetgeving moeten de analyses uitgevoerd worden door een erkend laboratorium en overeenkomstig de methoden opgenomen in het CMA. De manier van bemonsteren die vooropgesteld wordt in het CMA is duidelijk verschillend van de aanpak van INBO. De OVAM werkt bijvoorbeeld steeds met diepe boringen (tot minimaal 120 cm) waarbij van de verdachte bodemlagen een vast aantal stalen per oppervlakte-eenheid geanalyseerd moet worden. Er zijn ook veel gelijkenissen tussen beide bemonsteringsstrategieën, zo wordt in het geval van depositie (waaronder alluviale afzettingen) de bovenste 10 cm vanaf het maaiveld afzonderlijk bemonsterd en wordt een representatief aantal stalen van deze toplaag geanalyseerd. En hoewel de INBO analyses niet uitgevoerd werden door een erkend laboratorium, werden wel grotendeels de in het CMA voorgeschreven analysemethodes gebruikt. Enkel voor de bepaling van organisch materiaal (door INBO gemeten via *loss on ignition* en niet op basis van infrarood spectrometrie in een TOC analyser) en van kleigehalte (door INBO gemeten via laserdiffractometrie en niet met de pipetmethode) werd afgeweken van de standaard methodes. Voor het bepalen van het gehalte aan elementen werd de standaard analysemethode (ICP-OES) gebruikt, maar de ontsluiting gebeurde met een verschillend solvent (namelijk koningswater in plaats van een mengsel van zoutzuur, waterstoffluoride en salpeterzuur zoals het CMA voorschrijft). Een belangrijke kanttekening hierbij is dat het doel van INBO bij de bodembemonstering is om de referentietoestand vast te leggen en trends van de bodemverontreiniging op te volgen. Dit in tegenstelling tot de OVAM die de aanwezigheid van polluenten in de bodem wil vaststellen en de verspreiding ervan ruimtelijk aflijnen om, indien van toepassing en gebaseerd op de geldende wetgeving, bodemsanering op te leggen.

## 1.2 Doelstelling

Het doel van voorliggende studie is om de bemonsteringsstrategie die tot nu toe door INBO toegepast werd, te evalueren en indien nodig bij te sturen. Er wordt gekeken hoeveel bodemstalen nodig zijn, en in welke configuratie, om de doelstelling van de bemonstering te bereiken. Er wordt ook nagegaan of het aantal stalen of labo-analyses beperkt kan worden door gebruik te maken van andere datasets of van variabelen die de concentraties zware metalen kunnen vervangen. Er zal ook gekeken worden welke bemonstering optimaal is voor de monitoring van de gebieden na ingebruikname als overstromingsgebied.

De gebruikte datasets worden eerst beschreven, met name de chemische variabelen met bijhorende analysemethode en nauwkeurigheid, evenals de GIS datalagen (hoofdstuk 2). Vervolgens worden in hetzelfde hoofdstuk ook eventuele transformaties of andere aanpassingen van de data beschreven. Nadien illustreert een eerste screening van de data de belangrijkste verschillen tussen en binnen gebieden. In hoofdstuk 3 wordt de statistische verwerking van de data in detail beschreven. Hoofdstuk 4 brengt de concentraties zware metalen in kaart per overstromingsgebied, waarbij de ruimtelijke patronen en variatie bekeken wordt, maar ook de relatie tussen de geanalyseerde parameters. Daarna wordt getracht om tot een algemeen model te komen om de ruimtelijke patronen van zware metalen in de bodem te voorspellen over de gebieden heen (hoofdstuk 5). In hoofdstuk 6 worden de kosten berekend van mogelijke alternatieve staalnameprocedures. Dit moet

leiden tot een het formuleren van aanbevelingen voor een efficiënte(re) staalname en gegevensinzameling (Hoofdstuk 7).

## 2 Data beschrijving

### 2.1 FloodGis databank

De FloodGis databank bevat alle veldgegevens en labo-analyses die de onderzoeksgroep Milieu en Klimaat van INBO verzamelde in de sigmagebieden. Op het moment van staalname werd voor elk staal de datum en de diepte van staalname genoteerd, alsook de huidige bodembedekking. Er werden elf categorieën onderscheiden, namelijk bos, populier, wilg, grasland, akker, mais, moeras, riet, braak, veen en tuin. De gebieden werden bezocht tussen november 2007 en april 2010, met uitzondering van Paardeweide, dat reeds in januari 2005 bemonsterd werd. Stalen werden in het labo geanalyseerd op de variabelen die worden weergegeven in Tabel 2-1. Voor elke variabele wordt weergegeven indien er in de databank waarden onder de bepaalbaarheidsgrens lagen en hoe hoog de bepaalbaarheidsgrens was. Voor sommige variabelen ligt de bepaalbaarheidsgrens niet vast maar varieert ze afhankelijk van de instellingen van het labo.

De bodemstalen werden eerst gedurende twee weken gedroogd in droogovens op 40°C. Vervolgens werden ze gemalen in een mechanische molen voorzien van een 2 mm zeef (Retsch, Haan, Duitsland). De hoeveelheid organisch materiaal werd bijna overal gemeten via *loss on ignition* (LOI) na verassing. Verassing werd uitgevoerd in 2 stappen, eerst werd het staal gedurende 12u verwarmd tot 105°C om alle water te laten verdampen en vervolgens werd het verast in een moffeloven op 550°C gedurende 4u. In 2 gebieden werd de totale concentratie aan organische koolstof ook gemeten met een TOC analyser, uitgerust met vaste monstermodule, bij 900°C (Shimadzu 5050A Solid Sample Module Analyser, Shimadzu, Kyoto, Japan). Bodemtextuur werd bepaald via laser diffractometrie (Coulter LS, Miami, US), waarbij de kleifractie gedefinieerd wordt als de 0-6 µm fractie. Voor elk bodemstaal werd de  $pH_{H_2O}$ ,  $pH_{CaCl_2}$  en elektrische conductiviteit (EC) bepaald in een 1:5 (m/m) bodem-water-suspensie op massabasis.  $CaCO_3$  werd gemeten door terugtitratie van een overmaat  $H_2SO_4$ , toegevoegd aan 1 g luchtdroog sediment, met 0,5 M NaOH. De totale N concentratie werd gemeten met de Kjeldahl-methode. De concentratie sulfaat in waterextract werd bepaald met behulp van een ionenchromatograaf (Dionex IC 3000CD). De absorbantie van de bariumsuspensie werd gemeten met een spectrofotometer. De hoeveelheid plantbeschikbare P werd colorimetrisch bepaald na extractie met natriumbicarbonaat, volgens de methode van Olsen (Olsen et al., 1954). De pseudo-totale *aqua-regia*-extraheerbare concentratie van de elementen Ca, K, Mg, Na, P, S, Al, As, Cd, Co, Cu, Cr, Fe, Mn, Ni, Pb, Se en Zn werden gemeten met de ICP-OES (Varian VISTA-MPX, Varian, Palo Alto, CA) na microgolf ontsluiting. De concentratie Hg tenslotte werd gemeten met een kwikanalyser (Mercury Analyzer AMA-254). Het staal werd opgewarmd tot 750°C en het gevormde gas gezuiverd met zuurstofgas. Het Hg werd verzameld uit deze gassen en verwarmd tot 900°C. Vervolgens werd de hoeveelheid Hg bepaald via atomaire spectrometrie.

Tabel 2-1. Variabelen die in het labo werden opgemeten op de bodemstalen van de FloodGis databank. De gebruikte eenheid wordt getoond, evenals de bepaalbaarheidsgrens (BG) die voorkomt in de databank en een courante bepaalbaarheidsgrens (BG) waarvan de helft van deze waarde werd gebruikt bij de verwerking van de data (BG/2).

Variabele	Eenheid	BG in dataset	BG standaardwaarde	BG/2
Klei	(%)	-	-	-
Leem	(%)	-	-	-
Zand	(%)	-	-	-
Droge stof	(%)	-	0,1	0,05
Loss on ignition (LOI)	(%)	-	0,5	0,25
Totale organische koolstof (TOC)	(%)	-	0,1	0,05
pH <sub>H2O</sub>		-	-	-
pH <sub>CaCl2</sub>		-	-	-
Elektrische conductiviteit (EC)	( $\mu$ S/cm)	-	-	-
Kjeldahl N	(%)	-	0,05	0,025
plantbeschikbare P	(mg/kg)	1	1	0,5
Sulfaat (SO4)	(mg/kg)	5, 10	5	2,5
CaCO <sub>3</sub>	(%)	0.1, 0.5, 1.0	1	0,5
Ca	(mg/kg)	-	50	25
K	(mg/kg)	-	50	25
Mg	(mg/kg)	-	50	25
Na	(mg/kg)	50	50	25
P	(mg/kg)	-	50	25
S	(mg/kg)	-	50	25
Al	(mg/kg)	-	10	5
Fe	(mg/kg)	-	10	5
Mn	(mg/kg)	-	5	2,5
As	(mg/kg)	2	2	1
Cd	(mg/kg)	0.1, 0.15	0,15	0,075
Co	(mg/kg)	1	1	0,5
Cr	(mg/kg)	-	0,5	0,25
Cu	(mg/kg)	5	5	2,5
Hg	(mg/kg)	-	0,05	0,025
Ni	(mg/kg)	-	0,5	0,25
Pb	(mg/kg)	-	0,5	0,25
Se	(mg/kg)	0.5, 2.0	2	1
Zn	(mg/kg)	-	1	0,5

Op het moment van voorliggende studie bevatte de FloodGis databank gegevens voor 1004 punten gelegen in elf overstromingsgebieden (Tabel 2-2). Hoewel soms ook diepere bodemlagen bemonsterd werden, worden hier enkel de resultaten voor de bovenste 10 cm bekeken. Een deel van de bezochte gebieden werd voorzien als overstromingsgebied binnen het Sigmaplan (1977) (Tabel 2-2) en met uitzondering van Kruikeke-Bazel-Rupelmonde zijn deze gebieden reeds in gebruik als gecontroleerd overstromingsgebied. De overige gebieden werden pas afgelijnd binnen het geactualiseerd Sigmaplan (Beslissing Vlaamse Regering 2005 en 2006) (Tabel 2-2). Dit zijn allemaal 2010 sigmagebieden, dit wil zeggen dat de inrichtingswerken van start zouden moeten gaan ten laatste in 2010. Op het moment van staalname was geen enkel van deze gebieden reeds in gebruik genomen als overstromingsgebied.

Tabel 2-2. Lijst van de onderzochte sigmagebieden met hun oppervlakte en aantal staalnamepunten. Tevens wordt weergegeven of het een gebied uit het Sigmaplan (1977) betreft ("1977") of uit het geactualiseerd Sigmaplan (Beslissing Vlaamse Regering 2005 en 2006) ("2010" slaat op het voorziene startjaar voor de aanleg van deze gebieden) en of het reeds in gebruik is als overstromingsgebied.

Gebiedcode	Gebiednaam	Oppervlakte (ha)	Sigmaplan	Sigmaplan 2006 startdatum	Overstroomd?	Aantal punten weerhouden in de dataset
1WIJ	Wijmeers I	159	2006	2010	Nee	144
AS1L	Anderstadt I	12	1977	2015	Ja	14
AS2L	Anderstadt II	15	1977	2015	Ja	10
BOZW	Bovenzanden	34	1977	2020	Ja	38
BUNT	De Bunt	88	2006	2010	Nee	85
KBR_	Kruikeke-Bazel-Rupelmonde	660	1977	2010	Nee	130
LIER	Polder van Lier	26	1977	2015	Ja	31
PAWE	Paardeweide	85	1977	2010	Ja	177
TIEL	Tielrodebroek	97	1977	2025	Ja	94
VLAS	Vlassenbroekse Polder	239	2006	2010	Nee	138
WALZ	Grote Wal, Kleine Wal, Zwijn	149	2006	2010	Nee	143

## 2.2 GIS lagen

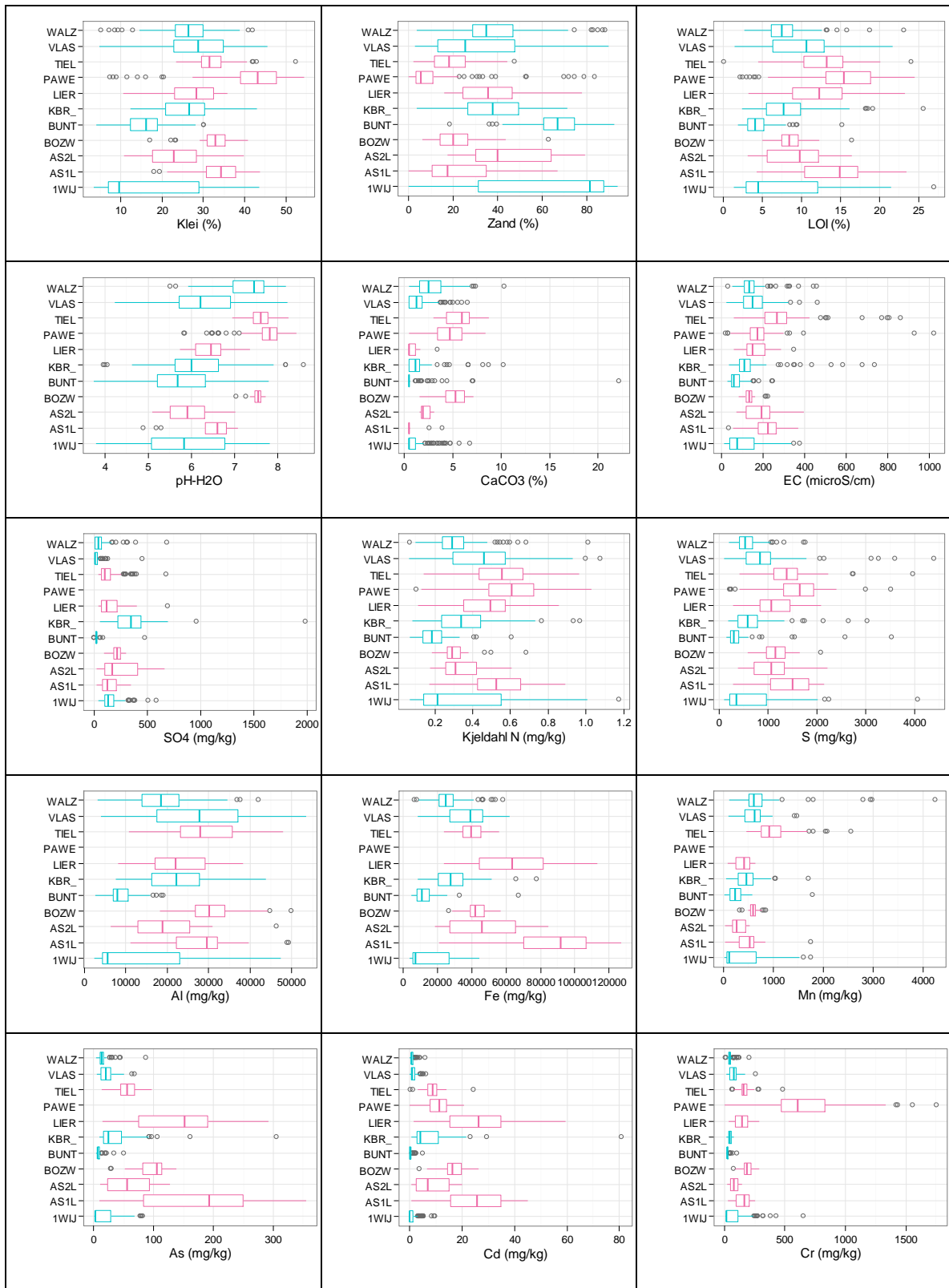
De belangrijkste bijkomende bron van data is het Digitaal Hoogtemodel Vlaanderen (DHM). Deze kaart werd uitgegeven door het Agentschap voor Geografische Informatie Vlaanderen (AGIV) en is gebaseerd op laseraltimetrie, een techniek waarbij het aardoppervlak wordt gescand d.m.v. een lasersysteem dat aan boord van een vliegtuig of helikopter wordt gemonteerd. Dit levert een verzameling onregelmatig verspreide punten op maaiveldhoogte met een gemiddelde dichtheid van 1 punt per 20m<sup>2</sup>. Om dit hoogtemodel bruikbaar te maken werd via ArcGIS een uitsnijding gemaakt van dit hoogteraster op de gebiedspolygonen en via R werd een overlay gemaakt tussen de meetpunten van het FloodGIS meetnet en dit hoogteraster.

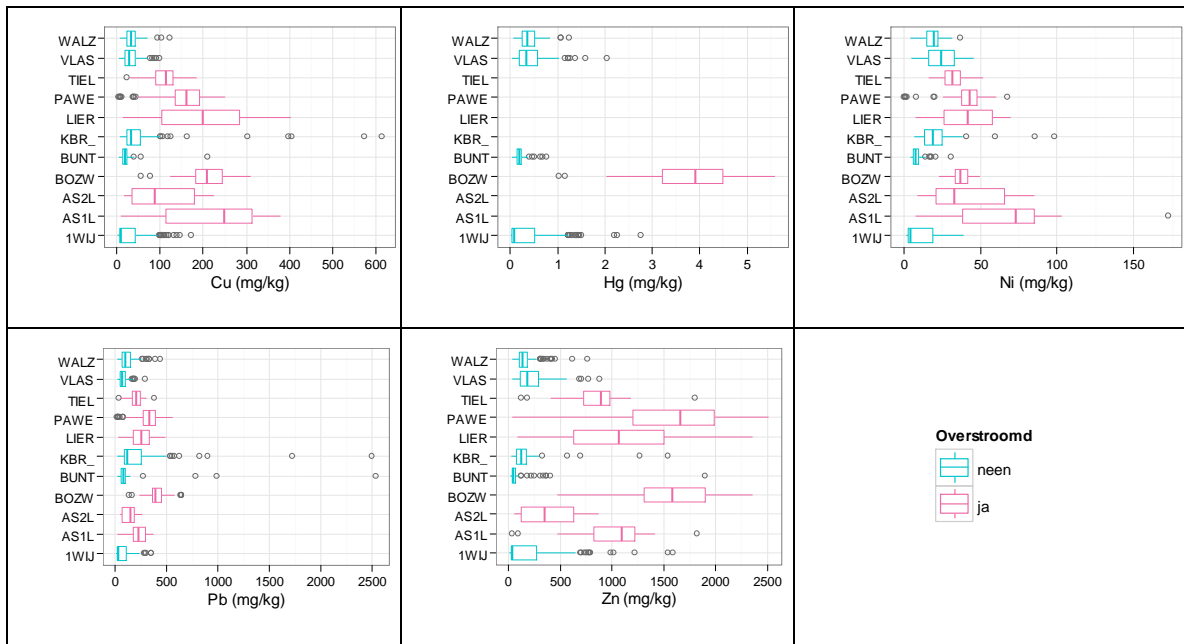
Er werd ook gebruik gemaakt van de Vlaamse Hydrografische Atlas (VHA), die eveneens ter beschikking gesteld wordt door AGIV. Deze vectoriële kaart geeft de centrale assen van bevaarbare waterlopen (waterwegen), onbevaarbare-geklasseerde en een aantal niet-geklasseerde waterlopen weer en wordt continu geactualiseerd waarbij de Afdeling Water van de VMM als centrale beheerder optreedt.

## 2.3 Algemene screening van de data

### 2.3.1 Chemische variabelen per gebied

Figuur 2-1 toont per gebied een aantal bodemvariabelen die in het labo werden opgemeten, namelijk enerzijds basiseigenschappen (textuur, organisch materiaal, pH) en anderzijds elementconcentraties (zware metalen en nutriënten). Gebieden die reeds in gebruik zijn als gecontroleerd overstromingsgebied vertonen hogere gehalten As, Cd, Cr, Cu, Ni, Pb, Zn en Fe. Ook het kleigehalte en organisch materiaal gehalte lijken gemiddeld iets hoger in reeds overstroomde gebieden. AS1L, LIER en BOZW bevatten hoge concentraties As, Cd, Cu en Zn en PAWE bevat uitzonderlijk veel Cr. Ook Zn-concentraties zijn er hoog. PAWE bevat ook veel klei, organisch materiaal en CaCO<sub>3</sub>, wat wijst op een hoge buffercapaciteit. BOZW bevat veel klei en CaCO<sub>3</sub> en AS1L veel klei en organisch materiaal. Hg werd pas opgemeten vanaf midden 2010. Er zijn dus niet voor alle gebieden resultaten voor Hg.

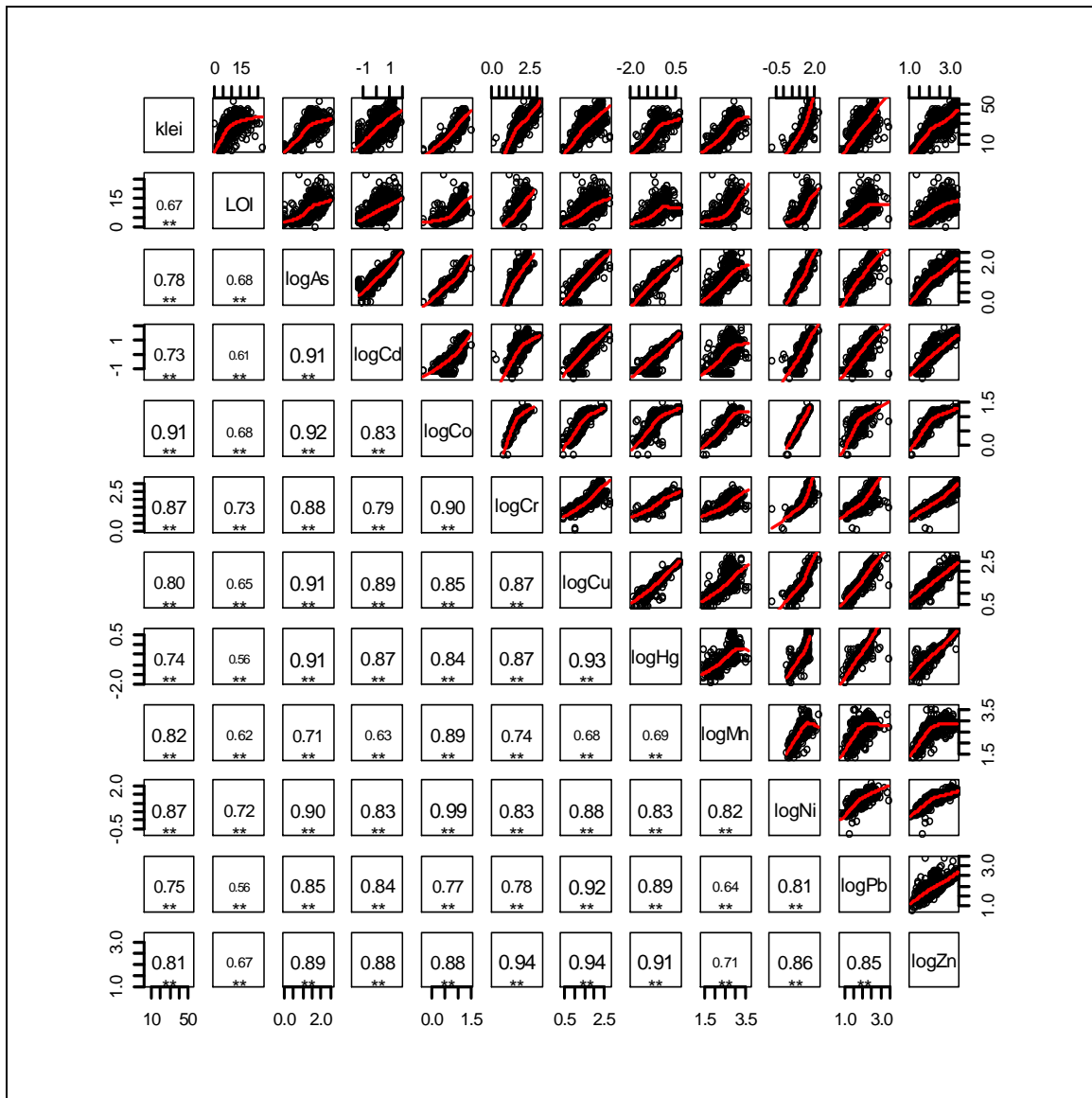




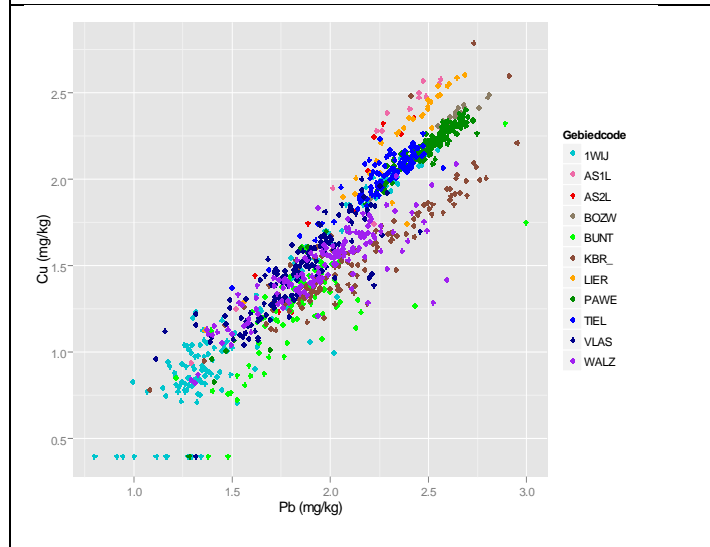
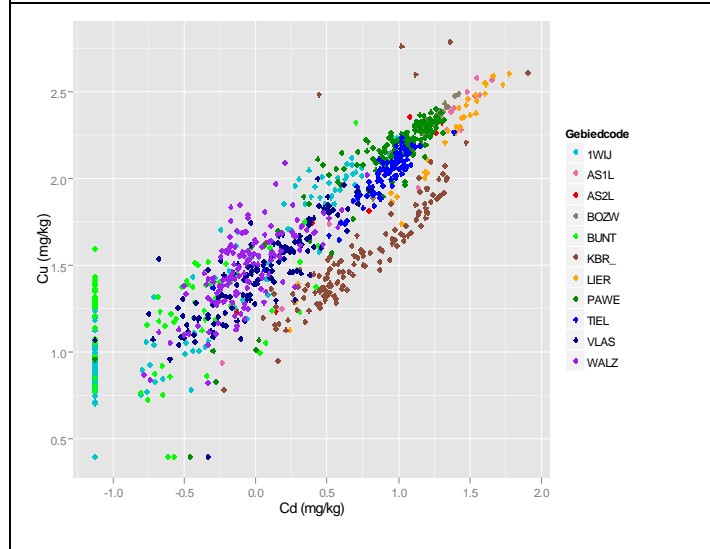
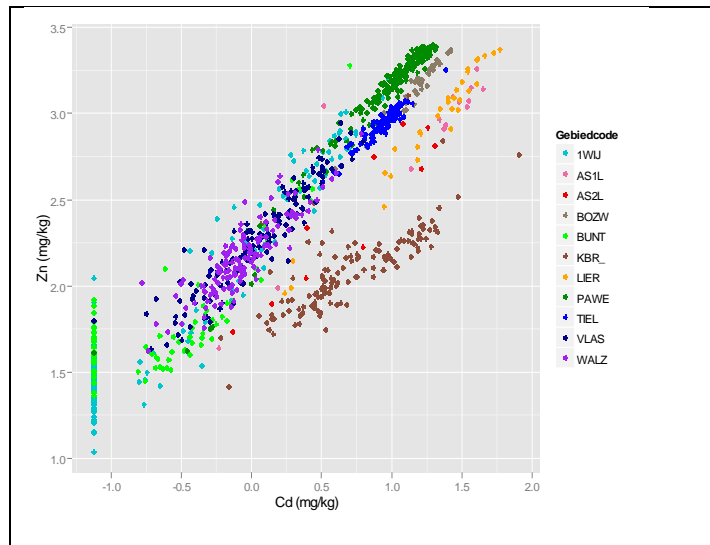
Figuur 2-1 bodemfysische en -chemische variabelen per sigmagebied.

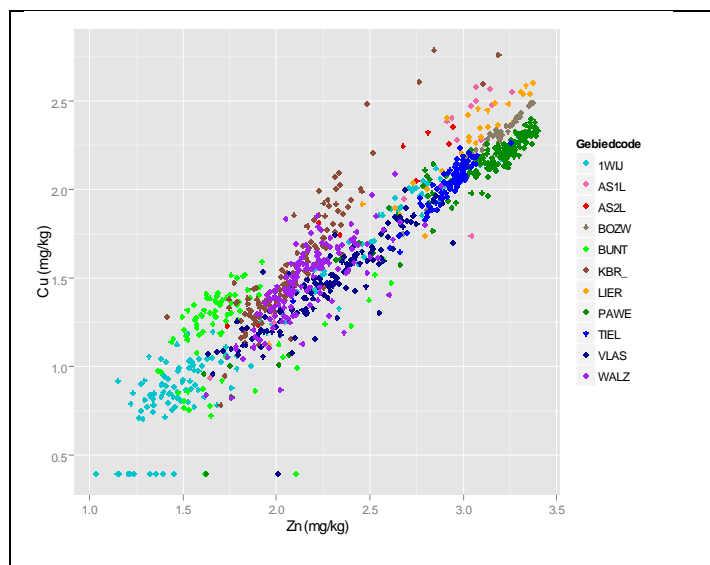
Figuur 2-2 toont de correlaties tussen zware metalen. De metalen zijn duidelijk gecorreleerd maar voor Cd, Zn, Cr en Cu zijn er meerdere correlatielijnen te zien. De puntenwolken voor Pb worden sterk beïnvloed door een beperkt aantal outliers. Bekijken we de correlatie tussen Cd en Zn meer in detail (Figuur 2-3), dan valt op dat KBR en AS1L-AS2L-LIER (drie aaneengesloten gebieden) sterk verschillen van de overige gebieden. Het is onduidelijk hoe deze verschillen verklaard kunnen worden. Beide gebieden liggen geografisch vrij ver verwijderd van de rest van de overstromingsgebieden, dus er zou een invloed kunnen zijn van de rivier. Ook valt op uit Figuur 2-1 dat het gehalte aan  $\text{CaCO}_3$  zeer laag is in KBR en AS1L-AS2L-LIER, er zouden dus ook bodemchemische factoren kunnen spelen. Voor KBR zou ook de atmosferische depositie van nabijgelegen industrie de vaste verhouding tussen paren van zware metalen kunnen verstoren. De correlatie tussen Cd - Cu, Pb - Cu en Zn - Cu is eveneens afwijkend voor dezelfde gebieden, zij het in mindere mate.





Figuur 2-2 Pearson correlaties tussen zware metalen onderling (concentraties zijn log10 getransformeerd) en tussen zware metalen, klei en organisch materiaal. De rode curve is een locally weighted regression (Lowess) lijn.

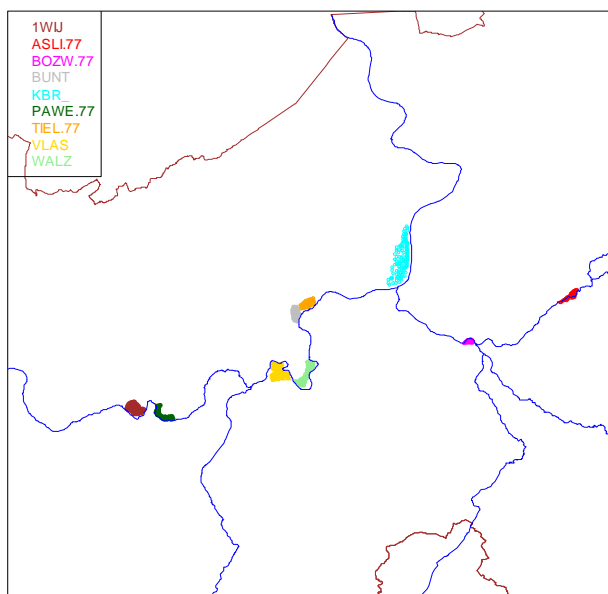




Figuur 2-3 Correlaties tussen gelogaritmeerde ( $\text{Log}_{10}$ ) concentraties van zware metalen met aanduiding van de gebiedcode. Voor Pb werden enkel de concentraties kleiner dan 1000 ppm bewaard, 3 outliers werden verwijderd.

### 2.3.2 Gebieden

Doordat er te weinig meetpunten waren in Anderstadt I en II en de polder van Lier voor een ruimtelijke analyse, maar deze toch vlak bij elkaar liggen, zijn deze samengevoegd tot het gebied ASLI. Zo zijn er nog negen gebieden in de dataset, waarvan er vier al eens overstroomd zijn in 1977 (Tabel 2-2). Deze worden verder aangeduid met het suffix '.77'. Het gebied KBR is anders gesampled dan de andere gebieden. KBR maakt gebruik van een onregelmatig sampleschema, terwijl de andere gebieden bemonsterd zijn op een grid van  $100 \times 100 \text{ m}^2$ . In 1WIJ zijn daar enkele afwijkingen op om de meetpunten niet op maar rond de vijver te leggen.

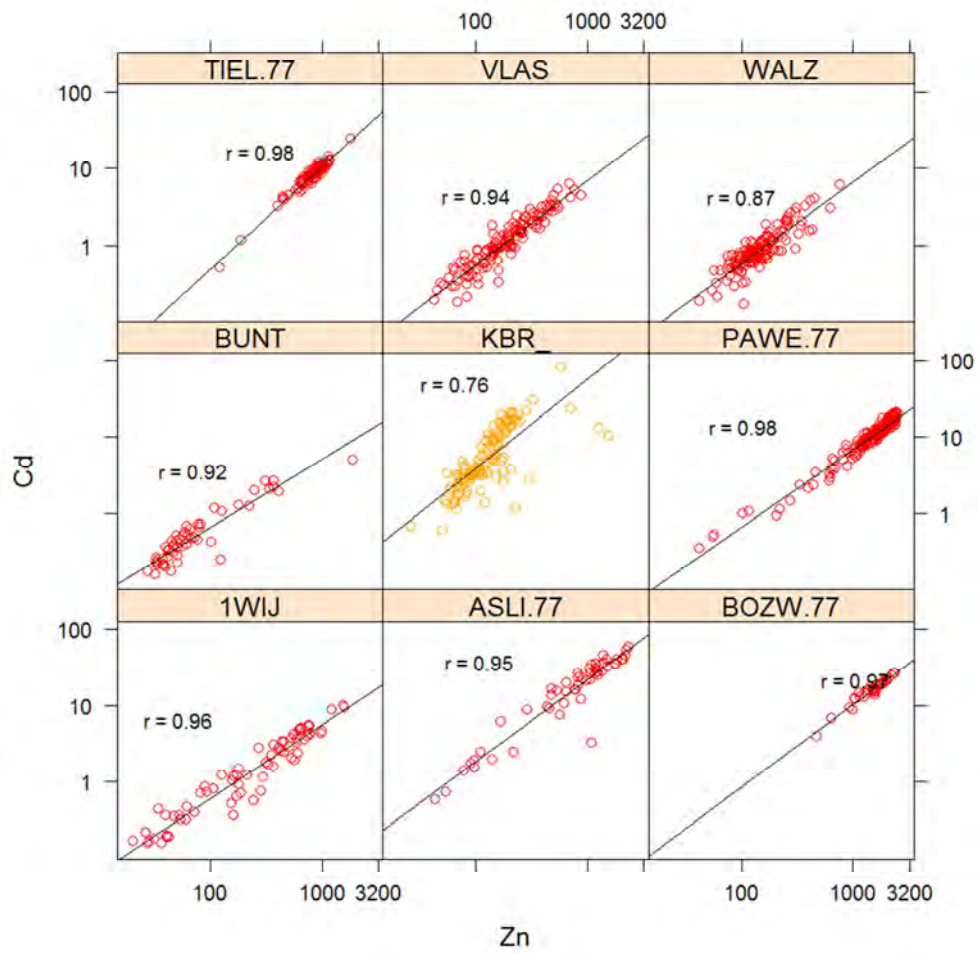


Figuur 2-4 De ligging van de overstromingsgebieden. AS1L, AS2L en LIER zijn gecombineerd in ASLI.

### 2.3.3 Bepalen van de responsvariabele

In de dataset zijn er acht verschillende zware metalen aanwezig (As, Cd, Cr, Cu, Hg, Ni, Pb en Zn). Alle zware metalen zijn sterk samenhangend met correlaties gaande van 0.75 tot zelfs meer dan 0.90. Hierdoor is het niet nodig de analyse te herhalen voor ieder zwaar metaal apart, maar kan een analyse op 1 metaal volstaan. Cadmium heeft ecotoxicologisch gezien een grote relevantie aangezien dit element relatief vlot opneembaar is voor planten en zeer schadelijk is voor zowel planten als organismen die voor hun voeding afhankelijk zijn van planten. Nadeel is echter dat dit element vaak gemeten wordt in concentraties die onder de bepaalbaarheidsgrens van het labo liggen, wat hem als variabele minder geschikt maakt voor statistische analyse. Omwille hiervan werd niet cadmium maar zink gekozen als responsvariabele. Dit element is net als cadmium vlot opneembaar door planten maar het is enkel toxisch in hoge concentraties. Zink is sterk gecorreleerd met cadmium, dus bevat dezelfde informatie, maar hier stelt zich geen probleem met de detectielimiet. Een ander voordeel is dat zink zelfs een nog grotere samenhang heeft met de andere metalen dan cadmium.

In de gebieden waarbij er veel cadmiumwaarden onder de detectielimiet gemeten werden (BUNT en 1WIJ) is de correlatie slechts rond de 0.75. Indien echter enkel cadmiumwaarden boven de detectielimiet beschouwd worden, is de correlatie ook voor deze gebieden boven de 0.9. Toch blijft er nog 1 gebied over waarbij de correlatie niet zo goed is en dat is in KBR, waarvoor de correlatie slechts rond 0.75 ligt. Dit komt waarschijnlijk doordat KBR zwaar beïnvloed wordt door een grote vervuilingbron ten oosten van dit gebied, die de verhouding tussen aanwezig zink en cadmium verstoort.



Figuur 2-5 Samenhang tussen zink en cadmium per gebied voor cadmiumwaarden boven de detectielimiet. De concentraties zijn in de log<sub>10</sub>-schaal uitgedrukt omdat dit beter correspondeert met de werkelijke samenhang.

## 3 Methodologie

### 3.1 Data manipulaties

Omdat data “as is” niet rechtstreeks bruikbaar is voor statistische analyse, zijn eerst enkele voorbereidende stappen nodig:

- de schaal waarin wij meten is vaak niet de natuurlijke schaal van de meetgrootte. Het belangrijkste voorbeeld hiervan zijn concentraties. Deze zijn vaak van nature uit lognormaal verdeeld, maar wij meten deze als een overvloed aan lage concentraties en een klein aandeel hoge concentraties, terwijl na een logaritmische transformatie, deze eigenlijk mooi normaal en symmetrisch verdeeld zijn. In deze studie worden alle zware metalen, maar ook aluminium, calcium, fosfor, ...in de  $\log_{10}$ -schaal uitgedrukt. Ook andere grootheden zoals KjN gedragen zich lognormaal en zullen dus als dusdanig getransformeerd worden.
- Een ander soort transformatie gebeurt voor het landgebruik. In de verschillende gebieden werden 11 verschillende soorten landgebruik genoteerd (zie 2.1), maar voor een betekenisvolle analyse zijn voldoende observaties nodig per categorie. Daarom wordt gekozen om alles op te delen in vier categorieën, namelijk akkergebied (akker, mais), weidegebied, bosgebied (bos, populieren, wilg) en een restcategorie (moeras, riet, veen, braakliggend, tuin).
- Tenslotte worden ook enkele extra hulpvariabelen aangemaakt in de dataset:
  - Overstroomd: een dummy variabele die aangeeft of een gebied al dan niet reeds in gebruik is als overstromingsgebied (binnen het Sigmaplan (1977)).
  - Hoogte, HoogteBoven, HoogteMin en AfstandRivier. De hoogte wordt gehaald uit het 5x5-m hoogtemodel van Vlaanderen. Hieruit wordt de minimale hoogte van het gebied bepaald (HoogteMin) en de hoogte van ieder meetpunt boven deze minimumhoogte (HoogteBoven). Verder wordt uit GIS ook de afstand van een meetpunt tot de rivier bepaald, met de functie “gDistance” in R (referentie) en wordt deze uitgedrukt in de kilometerschaal.

### 3.2 Basisverkenning

Vooraleer de gebieden van meer nabij worden onderzocht, moet nog op twee punten gelet worden:

1. Variabelen die niet voldoende opgemeten zijn, dus veel missende waarden hebben, zijn niet bruikbaar voor statistische analyse.
2. Veel variabelen bevatten intrinsiek heel wat dezelfde informatie. Bij de analyse, en vooral bij de modellering is het niet aangewezen deze samen in het model op te nemen, omdat deze elkaar gaan beconcurreren en hierdoor het model foutieve aanwijzingen kan geven.

Er is als volgt tewerk gegaan om de variabelen te selecteren:

- De verklarende variabelen worden gegroepeerd als ze een zelfde soort informatie bevatten, zoals bijvoorbeeld bodemfysische eigenschappen, mineralen, ...
- Alle variabelen worden per gebied apart getest op hun onderlinge correlatie. Indien deze voor alle gebieden of op 1 gebied na zeer hoog is ( $> 0.75$ ), wordt een selectie gemaakt uit de verschillende gecorreleerde variabelen, omdat dit bij modelbouw problemen kan geven. De gekozen variabele moet in de analyse dan gezien worden als een proxy voor een bepaald soort milieufactor in plaats van het effect van de variabele zelf. Immers als verschillende variabelen sterk gecorreleerd zijn, is het niet mogelijk te achterhalen welke variabele nu juist de grootste invloed heeft.

- Als een keuze moet gemaakt worden tussen variabelen krijgen gemakkelijk te meten of in de wettelijke normen gebruikte variabelen voorrang. Dit geldt bijvoorbeeld voor LOI (TOC) of pH. Ook wordt er getracht variabelen over te houden van zoveel mogelijk uiteenlopende variabelengroepen zoals (organisch, mineralen, ...)
- De overgebleven variabelen worden dan verder gebruikt in de analyse.

Deze opeenvolging van stappen maakt het mogelijk om met een beperktere variabelenset te werken dan deze die beschikbaar is in de volledige dataset.

De focus zal vooral liggen op de verbanden tussen zink en de andere variabelen. Zink doet ook dienst als voorspeller voor de andere zware metalen. Een scatter matrix zal duidelijk maken of deze veronderstelling goed opgaat. Deze veronderstelling is correct als de metalen sterk met elkaar gecorreleerd zijn voor ieder gebied apart. Bovendien vertoont Zn geen left-censored values (waarden onder LOQ) zoals Cd, Cu, As.

### 3.3 Gebiedspecifieke analyse

#### 3.3.1 Model

Eens de gebieden en verklarende variabelen afgebakend zijn, wordt de invloed op de respons per variabele apart onderzocht. Hierbij wordt geen onderscheid gemaakt tussen variabelen die sturend kunnen werken zoals bijvoorbeeld de hoogte of de afstand tot de rivier en variabelen zonder duidelijke sturende werking, zoals bijvoorbeeld de kaliumconcentratie van het sediment. Er wordt gewerkt met een veralgemeend kleinste kwadratenmodel (gls) (Pinheiro & Bates, 2000; Venables & Ripley, 1997). Als responsvariabele wordt  $\log_{10}(\text{Zn})$  (mg/kg) gebruikt en als verklarende variabelen pH-CaCl<sub>2</sub>, LOI (% "loss on ignition"), kleigehalte (%), logP (mg/kg), logPIP (mg/kg, "plantbeschikbare fosfor"), logSO<sub>4</sub> (mg/kg), afstand tot de rivier (km), hoogte boven het laagste punt in een gebied (m), landgebruikcategorie (bos, weide, akker, varia) en logK (mg/kg). De andere variabelen in de dataset komen te vervallen, omdat ze te zwaar gecorreleerd zijn met een van de bovenstaande variabelen. Daarnaast wordt in het model ook nog een ruimtelijke correlatiestructuur geschat. De parameters hiervan worden niet verder geïnterpreteerd, omdat in deze verkennende modellen het belangrijker is dat er met een ruimtelijke correlatiestructuur rekening gehouden wordt, eerder dan de exacte structuur ervan te bepalen.

Vervolgens wordt het model vereenvoudigd, door variabelen die weinig aan het model toevoegen uit het model te halen en wordt het model gevalideerd met enkele simpele validatieplots. De overgebleven berekende parameters worden telkens in een tabel weergegeven per gebied.

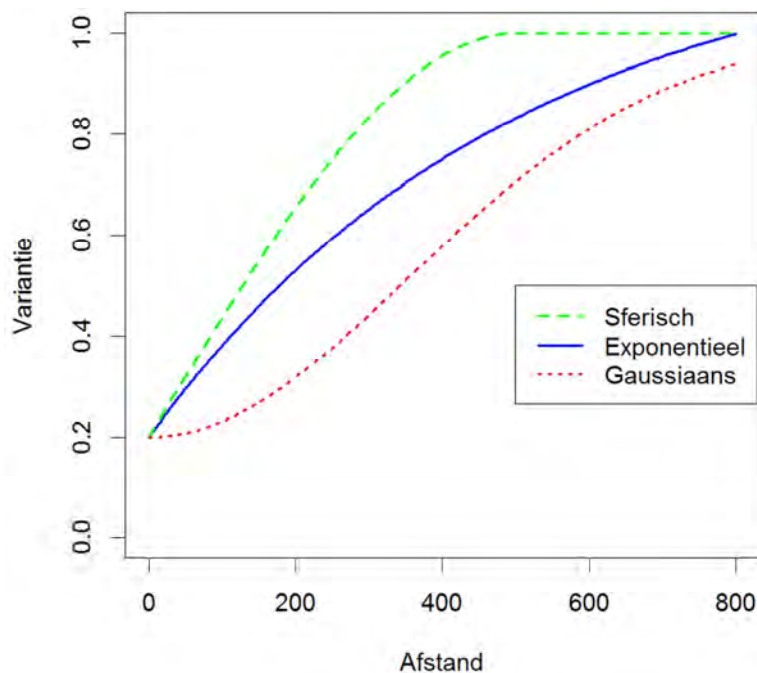
Deze modellen dienen vooral om te illustreren welke variabelen samenhangen met de concentratie van de zware metalen voor ieder gebied apart. Naar interpretatie toe moeten sterk gecorreleerde variabelen steeds als een proxy voor een variabelengroep geïnterpreteerd worden. In de samenvattende tabellen per gebied worden telkens alle mogelijke verklarende variabelen opgelijst (10 in totaal). Indien ze gebruikt zijn in het model wordt de schatting en het 95%-betrouwbaarheidsinterval gegeven met een indicatie voor p-waarden. Indien ze niet weerhouden zijn in het model wordt aangegeven of ze niet significant zijn of met welke andere variabelen uit het model ze te sterk gecorreleerd zijn. Verder wordt ook het aantal meetpunten gegeven. De ruimtelijke parameters worden niet getoond, omdat deze anders overgeïnterpreteerd worden, aangezien deze extreem variabel zijn en totaal anders kunnen zijn in een ander sample van meetpunten in hetzelfde gebied.

## Variogram

Een variogram is een manier om te achterhalen hoe sterk dichtbij gelegen meetpunten met elkaar samenhangen. Je kan verwachten dat, zelfs bij een relatief constante concentratie over een gebied, dat de gemeten concentraties toch telkens iets andere waarden zullen aannemen. Als de punten ver genoeg uit elkaar liggen heb je een bepaalde variabiliteit op de metingen, die wordt in een variogram de "sill" genoemd. Eigenlijk is dit de variabiliteit die je kan verwachten als 2 meetpunten onafhankelijk van elkaar zijn. Indien je echter een meting in de buurt van een andere meting uitvoert, verwacht je dat de gemeten concentratie niet zoveel zal afwijken, dit is wat men noemt ruimtelijke autocorrelatie. Een variogram wordt gebruikt om deze in een wiskundige vorm te gieten. Een variogram heeft 4 basisparameters:

- de range: dit is de afstand (m) vanaf waar aangenomen kan worden dat 2 meetpunten niet meer met elkaar gecorreleerd zijn, dus de metaalconcentraties mogen vanaf deze afstand als ruimtelijk onafhankelijk van elkaar aanzien worden. Als punten verder van elkaar gemeten worden - op voorwaarde dat je nog altijd dezelfde gemiddelde concentratie verwacht - zal deze variantie ook niet of nauwelijks hoger zijn dan de variantie op een afstand van de range. De waarde van deze variantie wordt de "sill" genoemd.
- de nugget: dit is de variantie die verwacht wordt tussen 2 punten die zeer dicht bij elkaar liggen. Deze zal heel wat lager zijn dan de variantie tussen punten die verder van elkaar liggen.. Dit is dus de variabiliteit die voornamelijk veroorzaakt wordt door fouten in staalname en analyse.
- de psill (partiële sill): dit is de variantie die bovenop de variantie van punten dichtbij elkaar (nugget) komt eens we op een afstand gelijk aan de "range" zijn.
- het type model: hoe verder we meetpunten uit elkaar leggen, hoe minder deze elkaar beïnvloeden, en eens voorbij de "range" verwachten we geen invloed meer tussen de punten. De invloed tussen punten vlakbij elkaar en punten op een afstand van de "range" neemt gradueel af. Hoe deze afname gebeurt, wordt bepaald door het type model. In dit verslag wordt vooral "sferisch", "exponentieel" en "gaussiaans" gebruikt. Bij sferisch en exponentieel neemt de invloed zeer snel af wanneer de punten verder uit elkaar liggen, maar deze afname wordt minder en minder uitgesproken als we nog verder gaan; bij gaussiaans gaat dit eerst traag, daarna snel, en daarna weer trager, een soort S-curve. Bij gaussiaans en sferisch bereiken we op de "range" het minimum aan autocorrelatie, bij het exponentiële model, blijft er nog een kleine invloed voorbij de range, maar hoe verder hoe minder. Dit staat grafisch afgebeeld in Figuur 3-1. Voor dit rapport is telkens gezocht naar een zo goed mogelijke benadering van het variogram dat inherent is aan de data, met vooral de nadruk op de korte afstanden. Meestal is het moeilijk om te onderscheiden welke vorm nu het beste is, omdat de meetpunten op een grid liggen, en er dus geen informatie is over hoe punten zeer dicht bij elkaar een gelijkaardig meetresultaat geven





Figuur 3-1 De drie voornaamste basisvormen voor variogrammen. De variantie tussen meetpunten wordt in functie van hun onderlinge afstand uitgedrukt.

### Concrete werkwijze

1. Eerst wordt de correlatie tussen de verschillende verklarende variabelen onderzocht. Variabelen met een correlatie groter dan circa 0.60 worden niet samen in het model opgenomen en de keuze wordt gemaakt op basis van 2 criteria: de basisvariabelen voor de norm (LOI, pH en kleigehalte) krijgen voorrang en er wordt gekeken naar de samenhang met andere verklarende variabelen (hoe lager hoe beter). Verder worden ook de punten van een zeldzaam landgebruik uit de data verwijderd, omdat deze de conclusies zouden kunnen scheeftrekken.
2. Er wordt een variogram gemaakt op het metaal zonder rekening te houden met verklarende variabelen. Ondanks meestal een 100 x 100 grid gebruikt is, wordt de basisbreedte toch op 50m gezet, omdat de afstand tussen 2 diagonaal naast elkaar liggend punten 141 dus ca. 150m bedraagt, en dus een interval van 50m wel relevant is. Verder wordt als basiscutoff 1500m (dus 30 x 50m) gebruikt om het variogram te berekenen, zodat het variogram geschat wordt op een vrij groot aantal punten, en met een startrange van 800m, wat wil zeggen dat we als startwaarde er vanuit gaan dat punten op 800m van elkaar niet meer met elkaar gecorreleerd zijn. De startwaarde van de nugget wordt op 0.02 gezet. Naargelang hoe het variogram fit wordt er wel wat gespeeld met de basisafstand en de cutoff.
3. Dit variogram wordt als startsituatie gebruikt in een GLS model. Niettemin berekent het GLS algoritme vaak een totaal verschillend variogram, omdat dit op basis van de residu's in het model is. Omdat dit variogram ook verder niet gebruikt wordt voor

predicties zijn de exacte vorm en parameters niet zo belangrijk, zolang de autocorrelatie maar verdwijnt uit de residu's.

4. De niet beduidende termen (keuze  $\alpha = 0.10$  als drempel) in het model worden uit het model gehaald, behalve LOI, pH en kleigehalte, die worden langer in het model weerhouden, tenzij het niet anders kan omdat ze onderling teveel correleren met elkaar. Verder zijn deze modellen puur illustratief.
5. Het variogram dat in de modellen wordt berekend, wordt verder helemaal niet meer gebruikt. De interpolatie van de zware metalen voor de kaarten gebeurt op de gemeten waarde zelf, dus er worden opnieuw variogrammen gemaakt die los staan van het variogram uit de modellen. Anders geldt echter de beperking dat de verklarende variabelen op dezelfde manier ruimtelijk moeten variëren als het zware metaal, en dit is hier niet gegarandeerd.

### 3.3.2 Kaarten

Per gebied zal de concentratie van zink en cadmium op een kaartje worden weergegeven. De kleurenschaal op de kaart is niet lineair, maar is gebaseerd op een constant aantal meetwaarden per kleur op de kaart, de kwantielen dus. De kaarten zijn steeds gebaseerd op een variogram die berekend worden op de gemeten concentraties van de metalen zelf.

De interpolatie gebeurt in de log-schaal, waardoor er geen problemen zijn met uitbijters die de berekening kunnen negatief beïnvloeden. Op de kaarten zelf echter wordt er teruggetransformeerd naar de originele schaal. Wiskundig kan dit echter niet gewoon met een antilog-transformatie, maar dient er een extra correctiefactor op basis van de kriging variantie in rekening gebracht worden. Deze transformatie is als volgt:

$$10^{(a + b)}$$

$$\text{Met: } a = \log \text{predictie}, b = \frac{1}{2} \sigma_{krige}^2$$

#### Concentratie

Eens de variogrammen gemodelleerd zijn, kan op basis van kriging de ruimte tussen de meetpunten per gebied apart geïnterpoleerd worden op een 5x5m grid. Op de meetpunten zelf zal deze voorspelling hetzelfde zijn als de waarde in de meetdata, tenzij er meerdere metingen binnen hetzelfde griddeel belanden. Tussen de meetpunten worden dan de concentraties voorspeld op basis van het variogram en de omliggende meetpunten. Hoe verder een voorspeld punt en een meetpunt van elkaar liggen hoe groter de variantie op de voorspelling is. Deze variantie wordt de kriging-variantie genoemd.

#### Kans om boven de norm te zitten

Naast de voorspelling van de metaalconcentraties zelf, is ook de kans op overschrijding van de Vlarebo bodemsaneringsnorm type I,II gemodelleerd (Bijlage 4 van het Besluit van de Vlaamse Regering houdende vaststelling van het Vlaams Reglement betreffende de bodemsanering en de bodembescherming van 14/12/2007).

**Belangrijk om hierbij te vermelden is dat de bemonstering van de bodem in het kader van dit onderzoek niet gebeurde volgens de code van goede praktijk en de analyses niet werden uitgevoerd door een erkend laboratorium. Op basis van de bekomen meetresultaten kan dan ook niet geoordeeld worden of de bodem verontreinigd is overeenkomstig het decreet betreffende de bodemsanering en bodembescherming en het Vlaams reglement betreffende de bodemsanering en bodembescherming. Hiervoor dient de bemonstering immers te gebeuren volgens de code van goede praktijk en dienen de analyses te worden uitgevoerd door een**

## **erkend laboratorium. Vergelijkingen met de bodemsaneringsnormen kunnen bijgevolg niet als representatief of rechtsgeldig worden beschouwd.**

Dit werd op 2 verschillende manieren gedaan. Enkel de tweede manier is weerhouden in dit verslag.

- De relevante variabelen voor het berekenen van de norm (klei, TOC, pH) worden net als het zware metaal geïnterpoleerd. Op basis hiervan wordt de norm berekend alsook de statistische kans dat de norm overschreden wordt. Meer informatie rond deze techniek is te vinden in de rubriek rond een alternatieve bepaling voor de overschrijding van de norm.
- Er wordt gebruik gemaakt van indicator kriging. Dus gewoon het effectief al dan niet boven de norm liggen van de meetpunten wordt omgezet naar een 1-0 variabele. Dit grid van 1 en 0 wordt dan via een variogram met kriging geïnterpoleerd en zo krijg je in de tussenliggende gebieden een continue waarde tussen 0 en 1 die overeenkomt met de kans dat dit punt boven de norm ligt.

### **Vergelijking tussen gebieden**

Wanneer de analyses op alle gebieden afgerond zijn, zullen de gebieden onderling met elkaar vergeleken worden, met een duidelijke focus op het verschil tussen het al dan niet overstroemd zijn. Dit gebeurt grafisch. Verder wordt ook een overzicht van de variogrammen in alle gebieden gegeven.

### **Cokriging**

Voor een willekeurig gebied wordt het model in plaats van op de gewone modelleringsmanier gefit via cokriging. Bij cokriging worden alle variabelen simultaan geïnterpoleerd over het gebied en zou de extra informatie die voortkomt uit de covariantie tussen de variabelen tot een betere predictie moeten leiden. Deze techniek heeft wel twee belangrijke beperkingen, namelijk dat de ranges van de variogrammen van iedere variabele gelijk moeten zijn en dat de rekenintensiteit enorm hoog is.

Daarnaast wordt nagekeken of het digitale hoogtemodel kan gebruikt worden om de concentratie van zware metalen mee te helpen voorspellen. Als dit mogelijk is zal de extra informatie die we in het hoogtemodel hebben over het hele gebied meehelpen aan de voorspelling van de zware metalen, dus hebben we plotseling heel wat meer meetpunten die kunnen gebruikt worden voor de voorspelling. Hiervoor is het wel nodig dat de relatie tussen de hoogte en de concentratie aan zware metalen sterk genoeg is.

### **Alternatieve bepaling voor overschrijding van de norm**

In plaats van gewoon de overschrijding van de norm als een binaire variabele te interpoleren, wordt hier een alternatieve methode voorgesteld. Dit alternatief houdt ook rekening met hoeveel de concentratie boven de norm ligt.

Om dit te bewerkstelligen, worden alle variabelen nodig om de norm te berekenen (klei, TOC, pH en zink) geïnterpoleerd op basis van een apart variogram voor elke variabele. Dan wordt voor het hele grid berekend wat de norm is. Om deze om te zetten naar een kans wordt gewerkt met een t-statistiek. Dit houdt in dat de geïnterpoleerde metaalconcentratie wordt verminderd met de norm, en gedeeld door de standaardafwijking (de vierkantswortel van de variantie) op de krigivoorspelling.

$$t = \frac{(\log Z_n - \log(\text{Norm}(\text{klei, TOC, pH})))}{\sqrt{\text{var}_{\text{krige}}}}$$

Dit komt overeen met een standaard t-statistiek, waarvan in de software bepaald kan worden met welke kans deze groter is dan 0 (dus de concentratie groter dan de norm).

Deze methode heeft het voordeel dat er meer informatie gebruikt wordt, namelijk de grootte van de overschrijding, maar heeft als nadeel dat er veel interpolaties moeten gebeuren, waaraan telkens een onzekerheid gekoppeld is, met het gevolg dat de t-statistiek ook door andere variabelen beïnvloed kan worden. Dit wil echter niet zeggen dat deze methode geen goede benadering kan zijn van de werkelijke overschrijdingskans.

### 3.4 Algemene analyse over de gebieden heen

In plaats van de analyse per gebied apart uit te voeren, wordt deze ook eens uitgevoerd tegelijk over alle gebieden heen. Hierdoor kan informatie van alle gebieden samen gebruikt worden om tot een goed model te komen, maar deze methode kan problemen geven als de gebieden anders reageren op verschillende variabelen.

Hier is gekozen om van alle variabelen inclusief de respons het gemiddelde per gebied af te trekken. Dus in het model wordt geen uitspraak meer gedaan in hoeverre de aanwezigheid van een bepaalde stof het zinkgehalte bepaalt, maar hoe de relatieve waarde van de variabele binnen een gebied invloed heeft op de zinkconcentratie. De absolute cijfers van de coëfficiënten zullen nog moeilijk te interpreteren zijn, maar een positief verband met zink zal nu betekenen dat binnen een gebied er meer zink is op de plaatsen in het gebied waar die bepaalde stof of eigenschap de hoogste concentratie heeft. Omdat de concentraties van zink in het ene gebied van 100 tot 200 mg/kg en in een ander gebied van 100 tot 2000 mg/kg kunnen uiteenlopen, wordt er een wegingsfactor per gebied aan het model toegevoegd, zodat de homogeniteit van de zinkconcentraties tussen de gebieden - wat een noodzakelijke voorwaarde is voor een model - verkregen wordt en zo de gebieden met de grootste spreiding het model niet zullen domineren.

Nu kunnen we uit het model halen hoe de verschillende verklarende variabelen samenhangen met de zinkconcentratie, zonder er rekening mee te moeten houden dat de concentraties in ieder gebied anders zijn. Het gebied zelf toevoegen aan het model heeft in deze opzet geen zin, aangezien het gebiedseffect al in rekening gebracht is toen we de variabelen herschaalden door het gemiddelde per gebied in mindering te brengen. Dit lost wel het probleem op dat het gebied zeer bepalend is voor de grootte van de metingen van de verklarende variabelen. Het gebied zal dan het effect van de andere variabelen maskeren.

De partiële predicties van het model worden in een figuur samengebracht. Doordat bij deze figuren rekening gehouden is met de range van de verschillende verklarende variabelen, en het gemiddelde effect per definitie 0 is, kunnen de subplots met elkaar vergeleken worden naar effectgrootte.

## 3.5 Kosten

### 3.5.1 Sampling strategie

Verschillende manieren van steekproeftrekking worden naast elkaar gelegd, gebruik makend van een gesimuleerde dataset op basis van een sferisch model. Enerzijds worden de meetpunten op een grid gelegd (200x200m, 200x100m, 100x100m, 50x50m), anderzijds wordt voor een steekproefpuntenaantal ongeveer gelijk aan dat van een 100x100m grid de meetpunten ofwel allemaal aselekt gekozen, ofwel wordt de helft op een 200x100m grid gelegd en de andere helft aselekt gekozen. Op basis van de gesimuleerde dataset worden

een 200-tal willekeurige steekproefcases gemaakt en wordt de variabiliteit tussen de cases voor nugget, psill en range van het bij dat horende variogram onderzocht. Zo wordt wat inzicht verkregen in welke impact de steekproef kan hebben.

### 3.5.2 Cokriging analyse

We willen weten of we op basis van een geringer aantal analyses op zware metalen in combinatie met metingen op goedkopere variabelen die wel intensief gesampled zijn, toch een betere voorspelling kunnen krijgen. Hiervoor houden we slechts een fractie van de metaalmetingen van de originele data over. Terwijl we voor de afhankelijke gecorreleerde variabele (klei in dit geval) wel alle beschikbare punten gebruiken. De predictie op het klein aantal metingen, de predictie op basis van de cokriging en de effectieve gemeten waarden worden dan met elkaar vergeleken om het nut van de cokriging af te wegen.

### 3.5.3 Kostenraming

Hier wordt enkel een eerste aanzet gegeven tot de berekening van de kosten voor scenario's, waarbij beslist wordt om in plaats van telkens het metaal zelf een andere, veel goedkopere variabele op te meten, dus op basis van cokriging. De relatieve kost voor verschillende scenario's wordt grafisch voorgesteld in functie van de kost van het veldwerk zelf zonder analyses.

## 4 Gebiedspecifieke analyse

### 4.1 Basisverkenning

Van de 1019 observaties over alle gebieden heen in de oorspronkelijke dataset zijn er 1004 behouden. Deze zijn verspreid over negen gebieden, vijf niet-overstroomde en vier wel reeds overstroomde sigmagebieden (Figuur 2-4).

Volgende variabelengroepen werden gebruikt als eerste selectiecriteria:

- Bodemtextuur: klei, zand en leemfractie. Deze zijn per definitie alle drie sterk gecorreleerd (som is 100%). Omdat in de Vlarebo norm het kleigehalte wordt gebruikt, zal enkel deze variabele weerhouden worden
- Bodemkarakteristieken:  $\text{pHCaCl}_2$ ,  $\text{pHH}_2\text{O}$ ,  $\log\text{EC}$ ,  $\log\text{CaCO}_3$ . De beide pH-waarden zijn zo goed als perfect gecorreleerd (gemiddeld verschil van 0.6 eenheden), en de keuze valt op de pH op basis van calciumchloride.  $\text{CaCO}_3$  heeft zeer veel waarden onder de detectielimiet en daardoor minder geschikt. Als de lage waarden buiten beschouwing gelaten worden is er wel een vrij sterk verband met de pH, waardoor de keuze om deze variabele te laten vallen gerechtvaardigd is.
- Mineralen:  $\log\text{Ca}$ ,  $\log\text{K}$ ,  $\log\text{Mg}$ ,  $\log\text{Na}$ . Magnesium en natrium zijn sterk gecorreleerd met kalium, dus dient één van de drie gekozen te worden. Nu is gekozen voor kalium, maar de andere twee zullen zich (statistisch) zo goed als identiek gedragen. Calcium is wel duidelijk verschillend en wordt nog meegenomen.
- Mineralen 2e groep:  $\log\text{P}$ ,  $\log\text{S}$ ,  $\log\text{Al}$ ,  $\log\text{Fe}$ ,  $\log\text{KjN}$ ,  $\log\text{PIP}$ . Kjeldahl-stikstof en zwavel zijn sterk gecorreleerd, dus is gekozen om enkel KjN over te houden. Ook aluminium en ijzer zijn sterk gecorreleerd, dus hier blijft alleen aluminium over.
- Organisch en andere:  $\text{LOI}$ ,  $\text{TOC}$ ,  $\log\text{SO}_4$ .  $\text{TOC}$  is maar gedefinieerd in één gebied en wordt weggelaten, maar kan wel voorspeld worden aan de hand van  $\text{LOI}$ . Hierdoor kunnen we de hoeveelheid organisch materiaal voor ieder gebied afzonderlijk in beschouwing nemen.  $\text{LOI}$  en  $\log\text{SO}_4$  worden geselecteerd.
- Beschrijvende variabelen die hoe dan ook meegenomen worden in de analyse: minimale hoogte van het gebied, de hoogte van een punt boven deze minimum hoogte en de afstand tot de rivier.
- Noodzakelijke variabelen: de gebiedcode en de XY-coördinaten.
- Daarna werden alle groepen samengebracht en blijken nog steeds enkele duidelijke correlaties aanwezig te zijn in de dataset.
- $\text{LOI}$ , stikstof en zwavel zijn sterk gecorreleerd, dus wordt enkel  $\text{LOI}$  weerhouden, omdat deze van belang is voor de berekening van de Vlarebo bodemsaneringsnorm.
- Aluminium en kalium zijn sterk gecorreleerd met elkaar en in grote maar iets mindere mate ook met ijzer. Hier wordt voor kalium gekozen.
- Calcium en pH hangen logischerwijs ook sterk samen, en enkel pH wordt behouden, omdat deze makkelijker meetbaar is.
- De elektrische conductiviteit en  $\text{LOI}$  hangen ook vrij sterk samen, doch in iets mindere mate (over het algemeen ligt correlatie ( $\rho$ ) tussen 0.65 en 0.85 en voor Tielrodebroek slechts rond 0.50).

Samenvattend gebruiken we:

- als responsvariabele  $\log\text{Zn}$ ;
- als verklarende variabelen: klei,  $\text{pHCaCl}_2$ ,  $\text{LOI}$ ,  $\log\text{SO}_4$ ,  $\log\text{K}$ ,  $\log\text{P}$ ,  $\log\text{PIP}$ ;
- andere relevante variabelen: HoogteBoven, AfstandRivier, XY-coördinaten.

## 4.2 Nog niet overstroomde gebieden

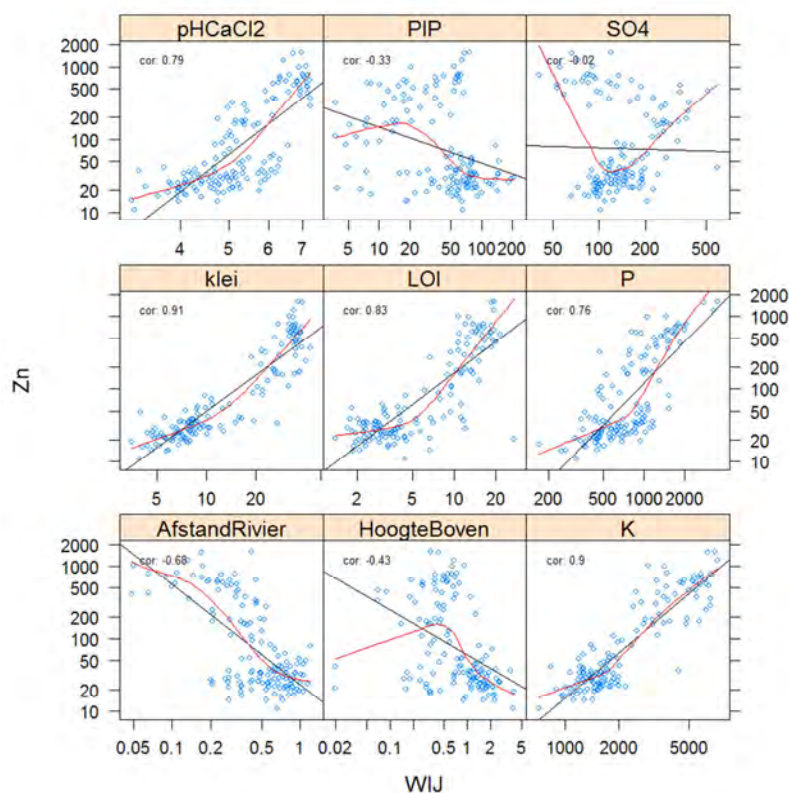
### 4.2.1 Wijmeers I

#### Model

Twee waarnemingen zijn uit deze analyse gelaten, omdat dit de enige 2 waren met het landgebruik "varia". Zo blijven 142 waarnemingen over in de Wijmeers I dataset.

Grafisch lijkt zink vooral sterk samen te hangen met fosfor, pH, kalium, klei en LOI. In het model kunnen we echter enkel het effect van klei inschatten, omdat de andere sterk met zink samenhangende variabelen zwaar gecorreleerd zijn met klei, waardoor hun effecten dus niet onafhankelijk geschat kunnen worden.

Het model duidt op een negatieve samenhang tussen zink en de afstand van de rivier en de hoogte in het gebied, terwijl er meer zink is bij hogere kleigehaltes of plantbeschikbare fosfor. Er dient opgelet te worden bij het interpreteren van deze 2 coëfficiënten, want klei en plantbeschikbare fosfor zijn in dit gebied negatief gecorreleerd ( $\rho = -0.44$ ). Verder is er volgens de modelcoëfficiënt ook minder zink in de meer sulfatrijke meetpunten. Dit laatste verband blijkt echter niet uit Figuur 4-1, waarschijnlijk is dit te wijten aan een andere variabele die reeds het zinkgehalte verklaart, maar niet volledig onafhankelijk van sulfaat is.



Figuur 4-1 Verbanden tussen zink en verklarende variabelen voor het gebied 1WIJ. De punten geven de ruwe data aan en linksboven wordt de correlatiecoëfficiënt weergegeven. Verder wordt ook nog de lineaire trendlijn aangegeven als een donkergroen-zwarte lijn en wordt er een lowess-smoother (lokaal gewogen schattingen) als een rode lijn weergegeven.

Tabel 4-1. Modelschattingen voor het gebied 1WIJ. N is het aantal waarnemingen. Per variabele wordt de geschatte modelcoëfficiënt weergegeven, samen met het bijhorende 95%-betrouwbaarheidsinterval. De significantie wordt aangegeven met . :  $0.05 \leq p < 0.10$ ; \* :  $0.01 \leq p < 0.05$ ; \*\* :  $0.001 \leq p < 0.01$ ; \*\*\* :  $p < 0.0001$ . De variabelen die niet opgenomen zijn in het model worden met schatting X aangegeven en de reden waarom ze niet opgenomen zijn is te vinden waar anders het betrouwbaarheidsinterval staat. Indien ze oorspronkelijk wel in het model zijn opgenomen, maar helemaal niet significant leken, worden ze met schatting 0 aangegeven.

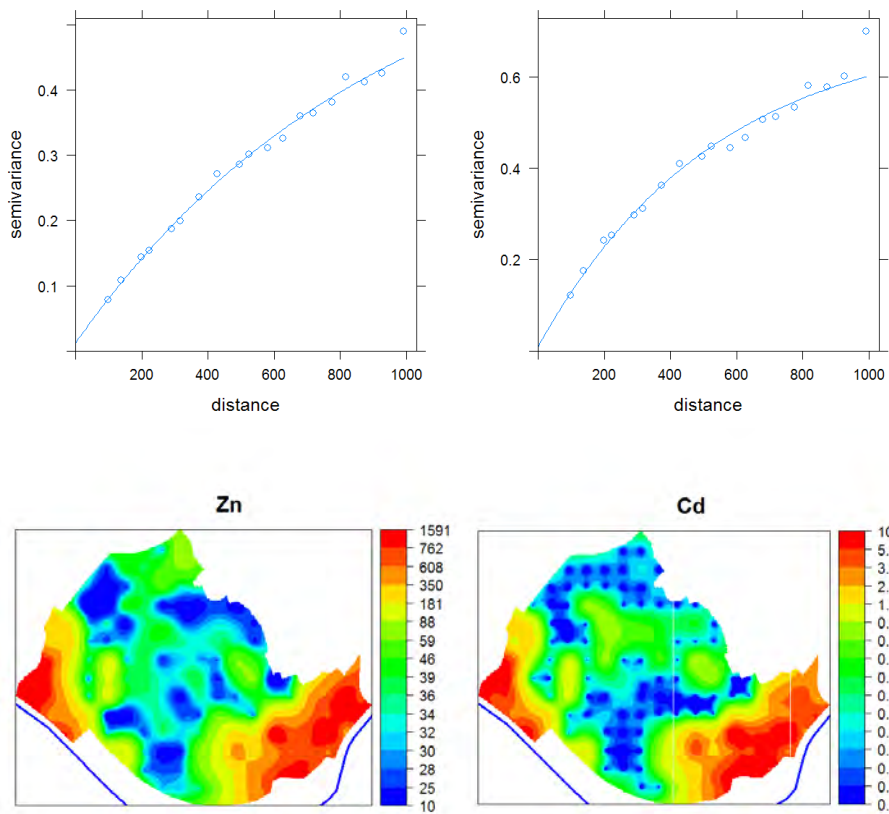
Var	Schatting	Interval95
N	142	NA
(Intercept)	1.87	[1.349, 2.386] ***
klei	0.04	[0.032, 0.042] ***
logSO4	-0.24	[-0.427, -0.05] *
logPIP	0.13	[0.035, 0.227] **
AfstandRivier	-0.29	[-0.635, 0.057]
HoogteBoven	-0.09	[-0.15, -0.034] **
pHCaCl2	X	hoge correlatie met klei
LOI	X	hoge correlatie met klei
logK	X	hoge correlatie met klei
logP	X	hoge correlatie met klei
Landgebruik	0	niet- significant

## Kaarten

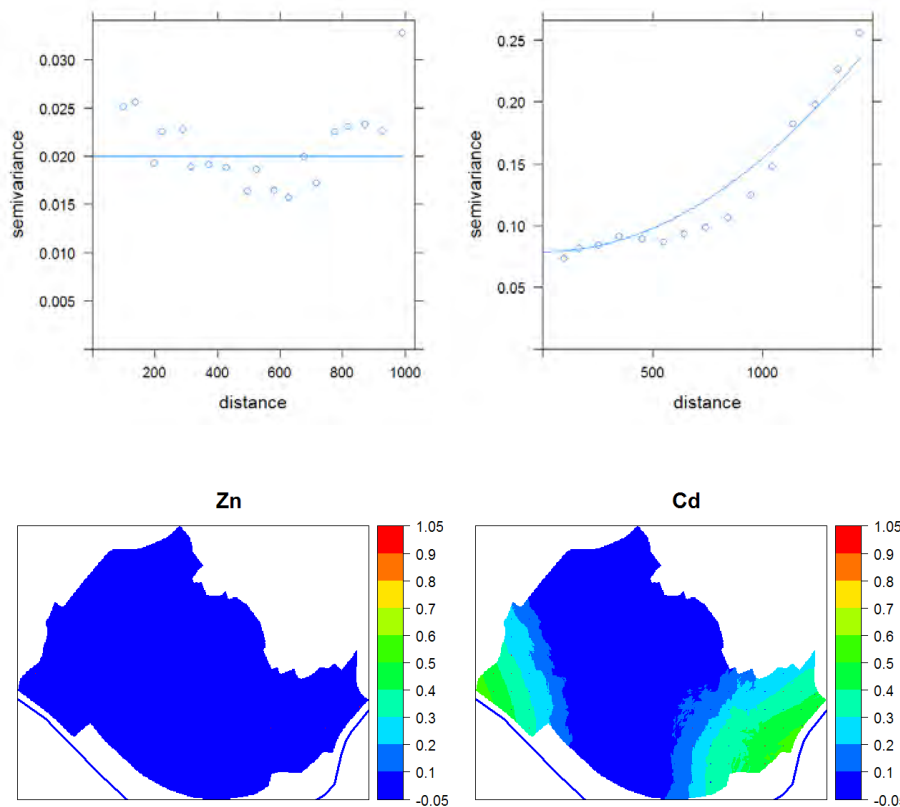
Bij de interpolatie is het duidelijk dat de concentratie van zink hoger is dicht bij de rivier.

De kans op overschrijding van de norm voor zink kan moeilijk op kaart voorgesteld worden, omdat er van alle meetpunten maar een 4-tal boven de norm uitkomen. Zo was het wiskundig niet mogelijk om tot een stabiel variogram te komen. Voor cadmium is er aan de oost- en westgrens wel een gebied waar de kans iets groter is, maar deze blijft over het algemeen beperkt.





Figuur 4-2 Predicties Zn en Cd in 1WIJ (mg/kg).



Figuur 4-3 Kans op overschrijding van de norm in 1WIJ.

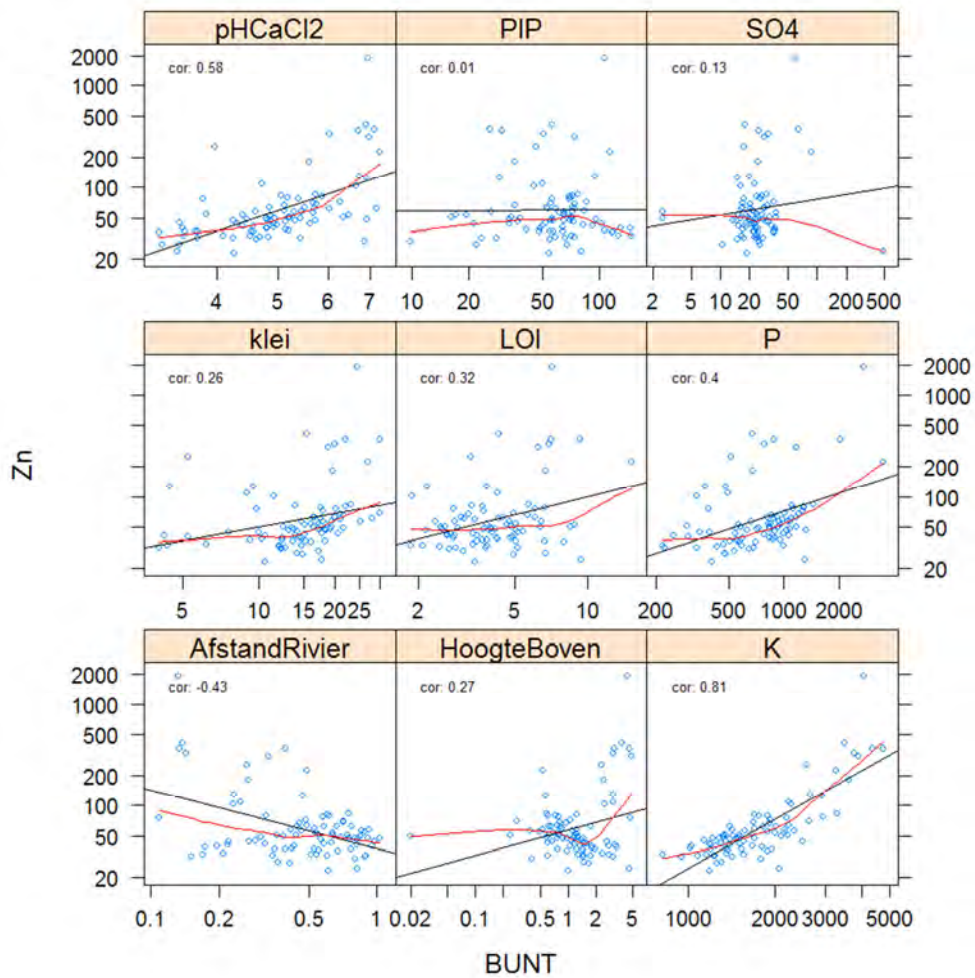
## 4.2.2 De Bunt

### Model

Er werd 1 waarneming weggelaten omdat dit de enige was met het landgebruik "varia". Zo blijven 84 waarnemingen over.

Het fosforgehalte kan niet in het model worden ondergebracht omdat deze sterk gecorreleerd is met klei ( $\rho = 0.65$ ). Verder valt ook de afstand tot de rivier en het landgebruik er a priori uit wegens hun onderlinge correlatie en de correlatie met hoogte. Gebieden dichtbij de rivier liggen hoger dan deze verder er vandaan en er komen vooral bossen voor dicht bij de rivier.

In dit model lijken hoger gelegen gebieden (dus dicht bij de rivier) meer zink te bevatten. Ook kalium, plantbeschikbare fosfor en de pH lijken overeen te komen met meer zink.



Figuur 4-4 Verbanden met Zn voor het gebied BUNT.

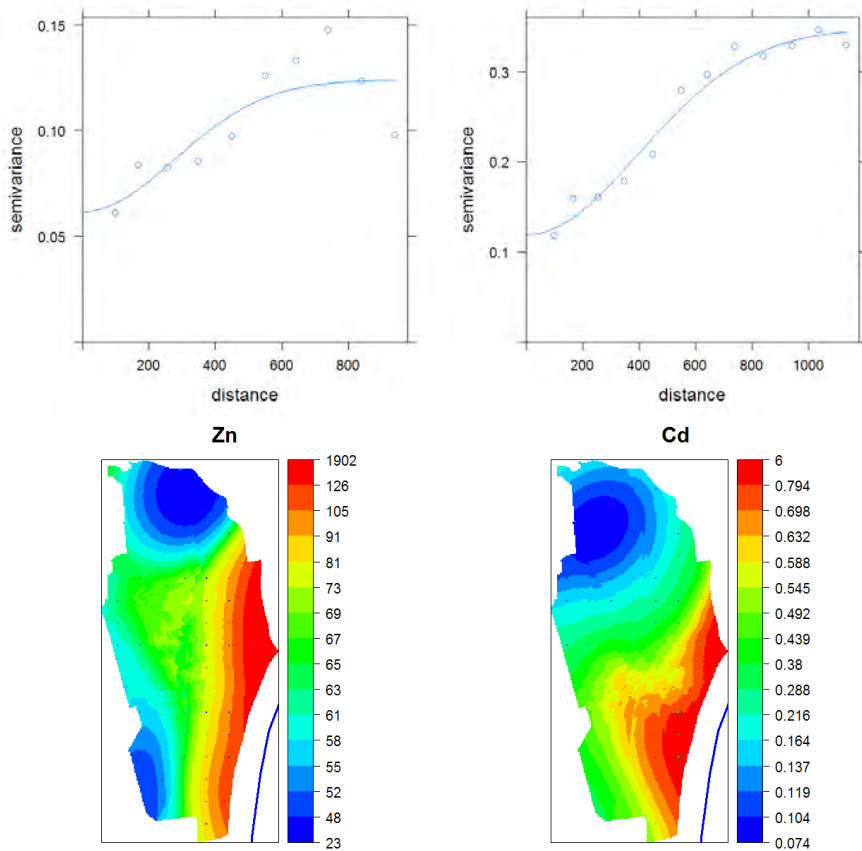
Tabel 4-2 Modelcoëfficiënten voor het gebied BUNT.

Var	Schatting	Interval95
N	84	NA
(Intercept)	-2.89	[-3.762, -2.026] ***
pH <sub>CaCl2</sub>	0.11	[0.072, 0.15] ***
logSO <sub>4</sub>	-0.11	[-0.238, 0.026]
logK	1.17	[0.889, 1.441] ***
logPIP	0.21	[0.054, 0.375] **
HoogteBoven	0.06	[0.023, 0.098] **
klei	0	niet-significant
LOI	0	niet-significant
logP	X	correlatie met klei
AfstandRivier	X	correlatie met hoogte
Landgebruik	0	net niet significant

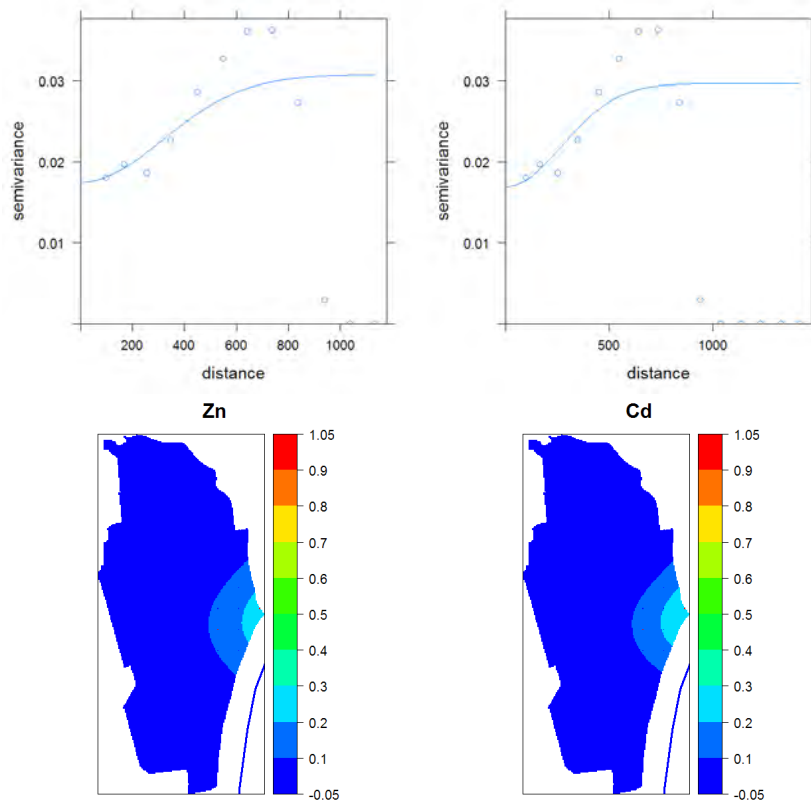
## Kaarten

Er zit een duidelijke oost-west gradiënt in de verdeling van zink over dit gebied. Deze gradiënt loopt sterk samen de afstand tot de rivier (of de hoogte). Dichter bij de rivier bevinden zich de hoogste concentraties, toch is het verschil in voorspelde concentratie niet zo groot en varieert deze tussen de 44 en 227 mg/kg. In de data zelf is het meest oostelijke punt een punt met een heel wat hogere concentratie dan de rest. Uit de stafkaart blijkt duidelijk dat dit punt beïnvloed wordt door twee waterlopen, Durme én Schelde.

Enkel het meest oostelijke meetpunt is boven de norm gemeten, dus interpolatiemogelijkheden voor de normoverschrijdingskaart is beperkt. In het kaartje met de normoverschrijding is er weinig basis om de spatiale evolutie van de kans op overschrijding als waar aan te nemen. Een ander grid zou een heel andere kansevolutie kunnen geven.



Figuur 4-5 Predicties Zn en Cd in BUNT (mg/kg).



Figuur 4-6 Kans op overschrijding van de norm in BUNT.

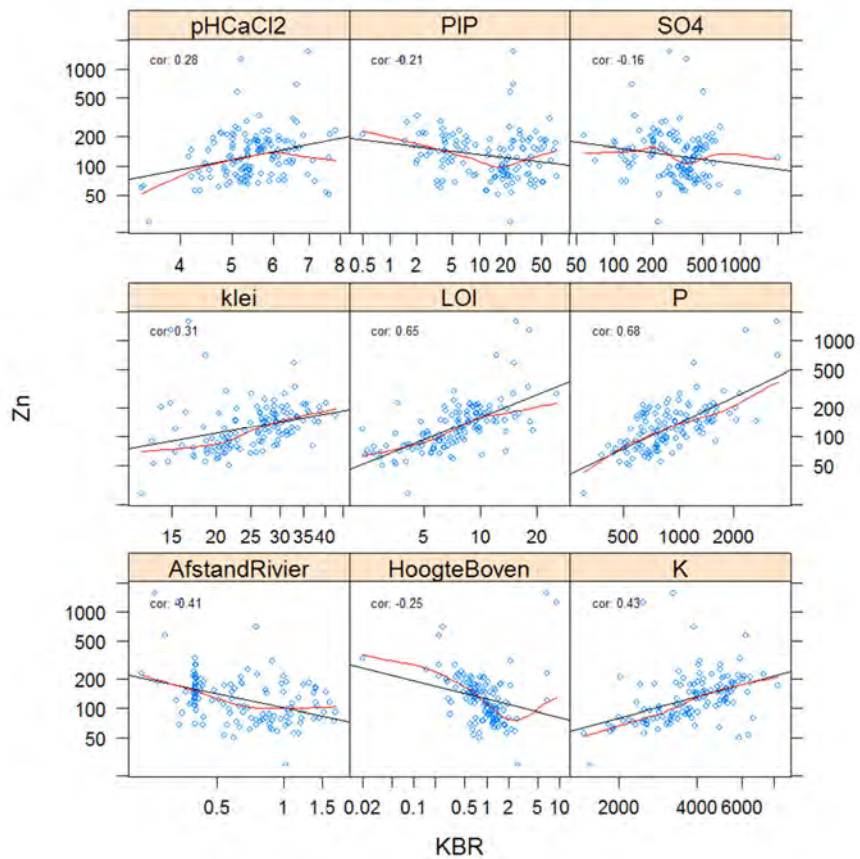
### 4.2.3 Kruibeke - Bazel - Rupelmonde

#### Model

Net als in de vorige gebieden is er ook hier één zeldzaam landgebruik aanwezig: op slechts één punt is er akkerland. Deze wordt dus om dezelfde redenen als voorheen uit de analyse geweerd. Zo blijven toch nog altijd 129 observaties over. Er is echter wel een essentieel verschil met de vorige gebieden: hier worden de meetpunten niet volgens een vast grid, maar met meer ruimtelijke variatie gekozen. Verder is er in dit gebied ook een serie waarnemingen aanwezig - een scherm als het ware - aan de oostkant, grenzend aan de rivier. Omdat deze waarschijnlijk met een bepaald doel gekozen zijn, vermoedelijk een vervuilingbron dichtbij, worden deze via een extra variabele "Scherm" onderscheiden van de rest.

Kalium wordt uit de dataset weggehaald omdat deze sterk ( $\rho = 0.68$ ) gecorreleerd is met klei. Verder blijkt ook de schermvariabele niet echt een invloed te hebben op de metaalconcentraties en kan deze opgevangen worden door de afstand tot de rivier, sulfaten en plantbeschikbare fosfor. Volgens het model vallen deze drie, net zoals klei en pH weg wegens niet significant. LOI van het scherm van meetpunten is over het hele scherm vrij gelijkwaardig, maar deze valt niet op omdat deze LOI zeer gemiddeld is voor het gebied. De enige variabelen die nog overblijven in het model zijn LOI, fosfor en hoogte van het

meetpunt ten opzichte van het gebied. Elk van deze drie hangt positief samen met het zinkgehalte, althans na correctie voor de andere verklarende variabelen, want in de figuur wordt net het omgekeerde verband getoond.



Figuur 4-7 Verbanden met Zn voor het gebied KBR.

Tabel 4-3 Modelcoëfficiënten voor het gebied KBR.

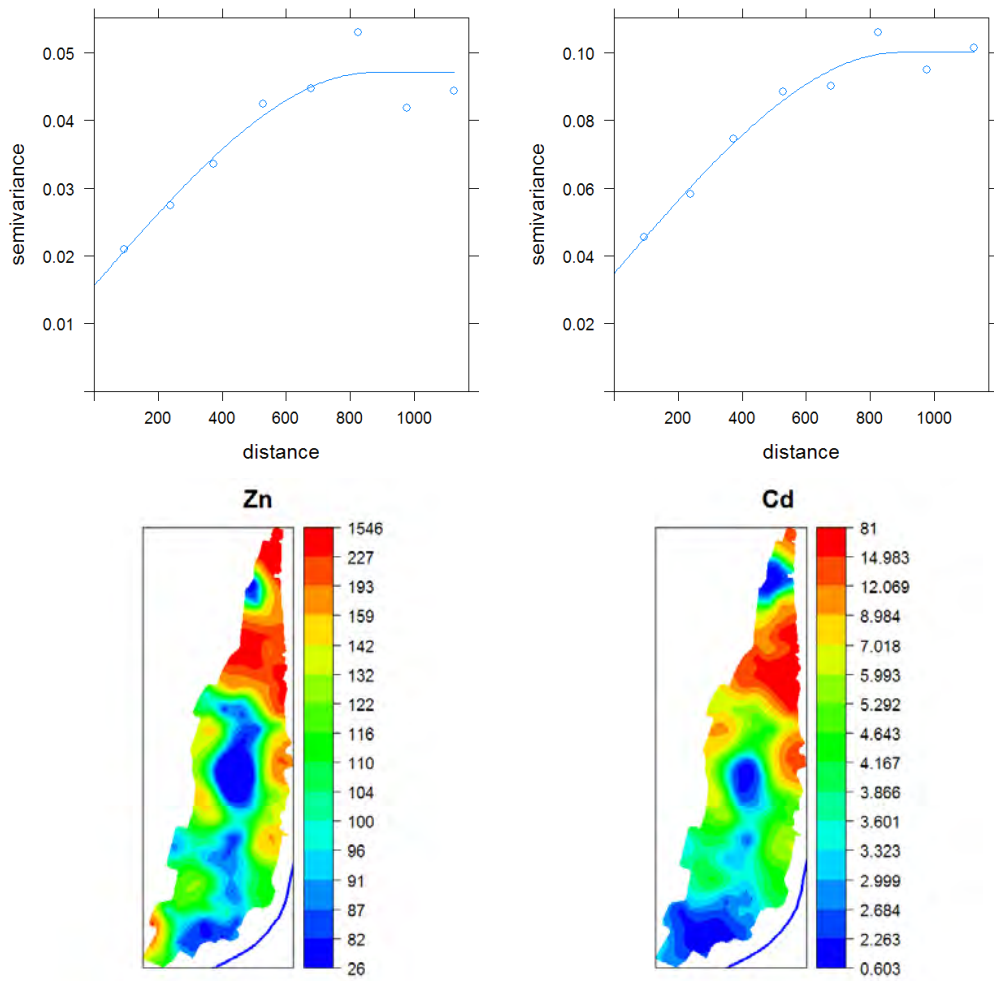
Var	Schatting	Interval95
N	129	NA
(Intercept)	0.31	[-0.134, 0.76]
LOI	0.03	[0.02, 0.033] ***
logP	0.53	[0.398, 0.661] ***
HoogteBoven	0.04	[0.012, 0.072] **
klei	0	niet-significant
pHCaCl2	0	niet-significant
logSO4	0	niet-significant
logK	X	correlatie met klei
logPIP	0	niet-significant
AfstandRivier	0	niet-significant
Landgebruik	X	correlatie met scherm en P
Scherm	X	correlatie met P, Hoogte, LOI

### Kaarten

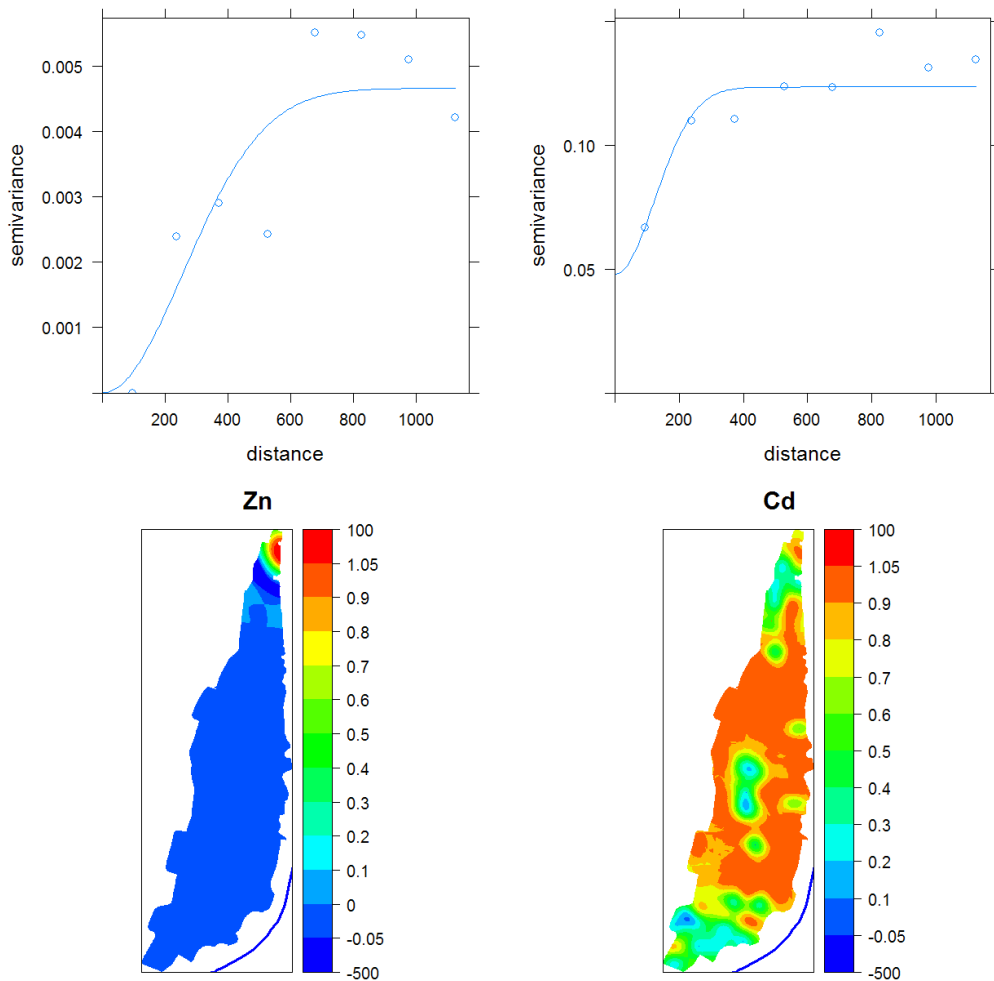
Vooraf in het noordoosten is de zinkwaarde zeer hoog, in de rest van het gebied is er niet enorm veel verschil in concentratie en liggen zowat alle waarden tussen de 80 en 320 mg/kg, met nog iets lagere waarden in het centrum van het gebied.

Net als in de vorige nog-niet-overstroomde gebieden zijn er ook hier nauwelijks overschrijdingen van de norm voor zink. Hier zijn er 2, de punten in het meest noordelijke deel van het gebied. Uit de concentratiekaart bleek ook al dat hier de concentratie uitzonderlijk hoog is. Cadmium daarentegen heeft wel een gelijkaardige spreiding als zink, maar hier is zowat het hele centrale gebied boven de norm.





Figuur 4-8 Predicties Cd en Zn voor KBR (mg/kg).



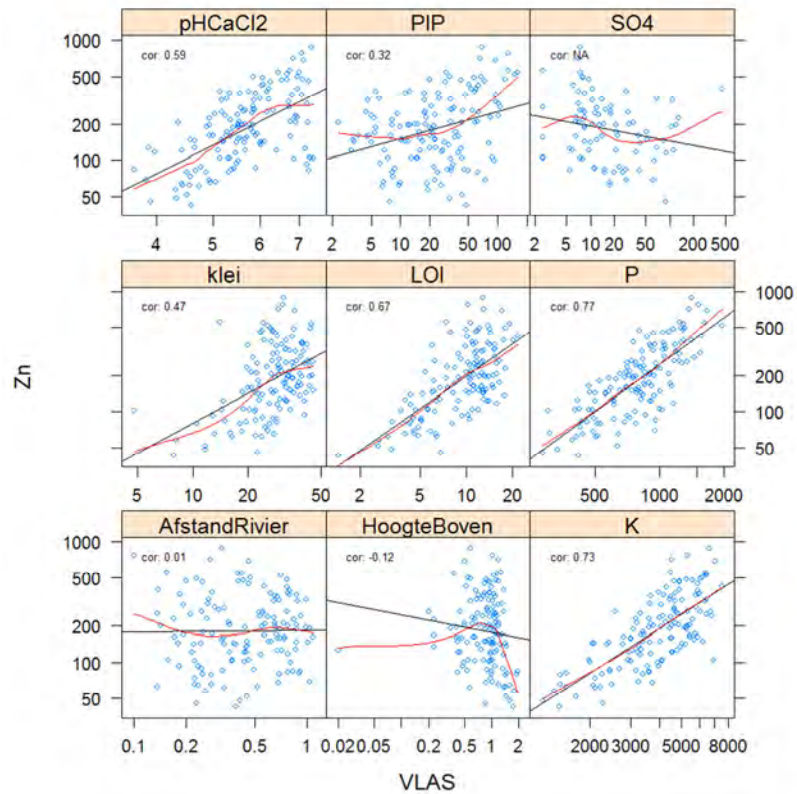
Figuur 4-9 Kans op overschrijding van de norm in KBR.

#### 4.2.4 Vlassenbroekse polder

##### Modellen

Oorspronkelijk zijn er 138 meetpunten en ook hier valt er eentje weg omdat het als landgebruik "varia" heeft. Akkerland komt ook relatief weinig voor, maar omdat er toch nog een aantal waarnemingen zijn blijven deze in de dataset.

Heel wat verklarende variabelen zijn gecorreleerd in deze dataset. Zo vallen LOI en K weg omdat deze teveel samenhangen met klei en fosfor. Ook de sulfaten verdwijnen a priori omdat deze missende waarden bevatten, maar tegelijk wel gecorreleerd zijn met pH, die dan het sulfaateffect zal opnemen.. De variabelen landgebruik, hoogte en afstand tot de rivier bleken niet significant. Verder beïnvloeden de modelparameters voor klei en fosfor elkaar (positieve correlatie) en krijgt plantbeschikbare fosfor een onlogische negatieve coëfficiënt. Er wordt besloten om plantbeschikbare P weg te laten uit het model. Uiteindelijk levert dit een model op met enkel klei, pH en fosfor die allemaal positief samenhangen met het zinkgehalte.



Figuur 4-10 Verbanden met Zn voor het gebied VLAS.

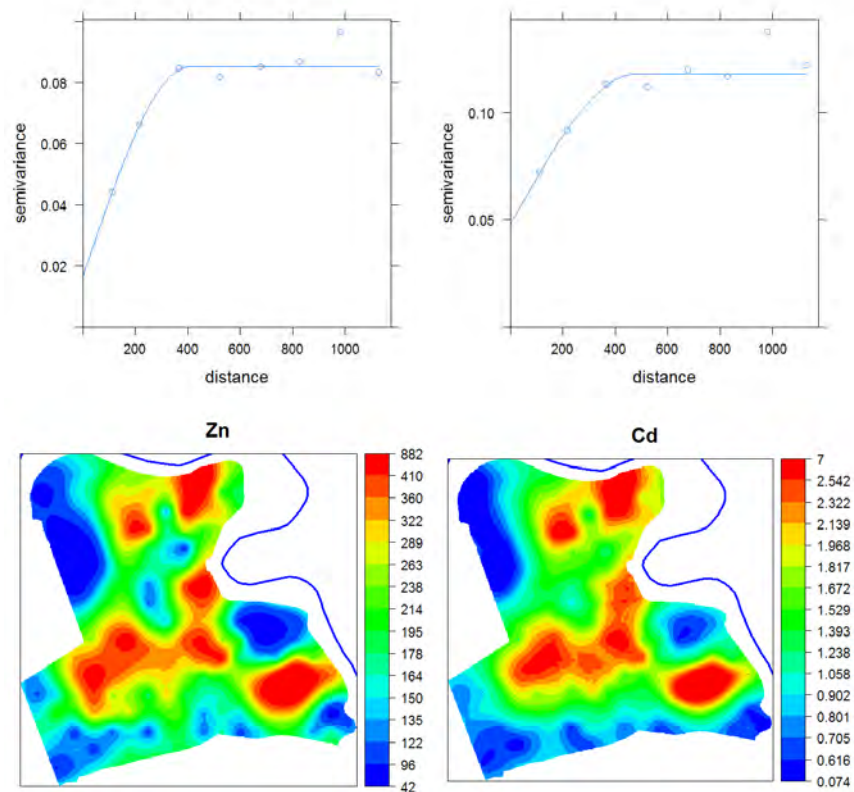
Tabel 4-4 Modelcoëfficiënten voor het gebied VLAS.

Var	Schatting	Interval95
N	137	NA
(Intercept)	-0.89	[-1.319, -0.468] ***
klei	0.01	[0.002, 0.009] **
pHCaCl2	0.12	[0.084, 0.149] ***
logP	0.8	[0.625, 0.976] ***
LOI	X	correlatie met klei en P
logSO4	X	correlatie met pH en missings
logK	X	correlatie met klei en P
logPIP	X	Onlogisch verband en maskeert klei
HoogteBoven	0	niet-significant
AfstandRivier	0	niet-significant
Landgebruik	0	niet-significant

## Kaarten

De zinkconcentratie is gespreid over het hele gebied, met enkele stukken met hogere en enkele met lagere concentraties. Er zit geen patroon in de verdeling. De cadmiumconcentratie is gelijkaardig gespreid als zink.

Aangezien er verder ook geen enkele concentratie boven de norm gemeten is voor zink, kan de kans op overschrijding van de norm niet gemodelleerd worden. Voor cadmium zijn er enkele meetpunten die de norm overschrijden, maar niet voldoende om tot een stabiel variogram te komen.

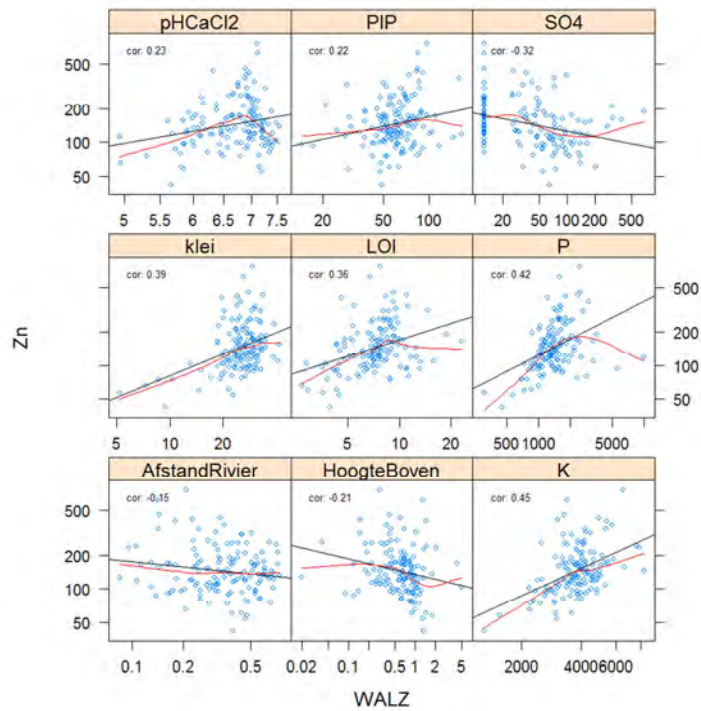


Figuur 4-11 Predicties Zn en Cd in VLAS (mg/kg).

### 4.2.5 Grote Wal, Kleine Wal en Zwijn

#### Model

Dit gebied bestaat uit 143 waarnemingen, waarvan er 4 worden weggehaald omdat ze het landgebruik "varia" kregen. Ook nu zijn er enkele stevige correlaties binnen de verklarende variabelen te vinden, namelijk tussen kalium en klei en tussen plantbeschikbare fosfor en het totale fosforgehalte. Klei en plantbeschikbare fosfor worden behouden in de *a priori* fase. Uiteindelijk blijken enkel klei, LOI en plantbeschikbare fosfor significant over te blijven in het model.



Figuur 4-12 Verbanden met Zn voor het gebied WALZ.

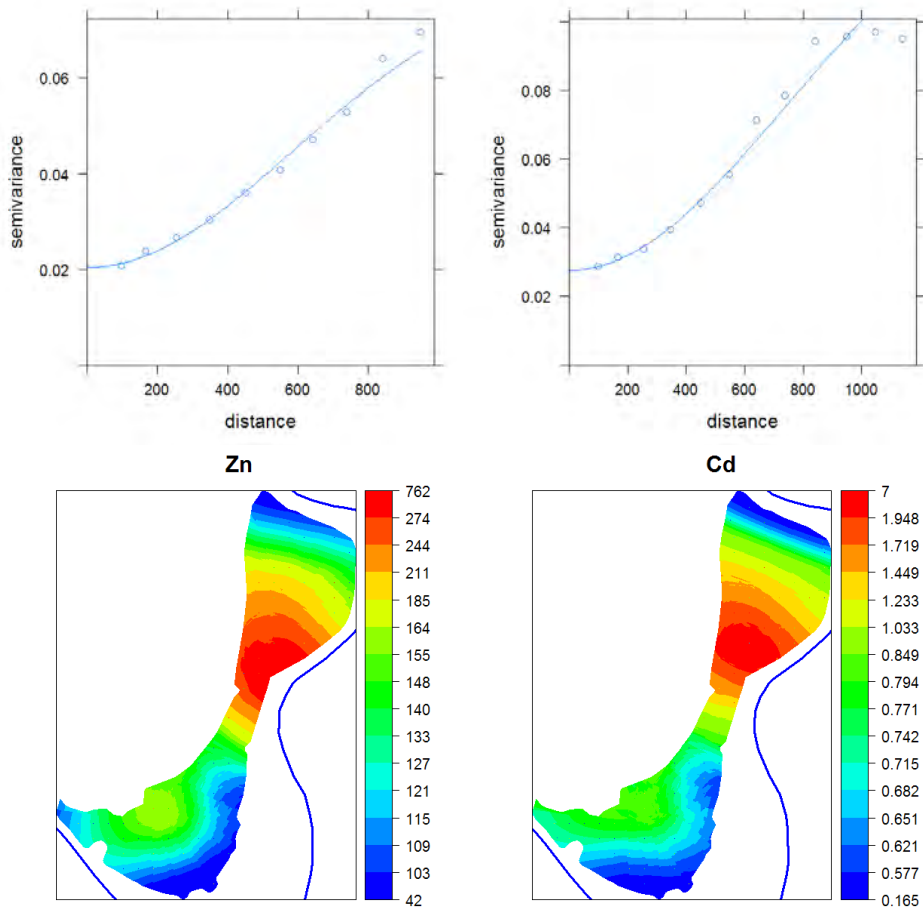
Tabel 4-5 Modelcoëfficiënten voor het gebied WALZ.

Var	Schatting	Interval95
N	139	NA
(Intercept)	1.24	[0.255, 2.229] *
klei	0.01	[0.001, 0.01] *
LOI	0.02	[0.007, 0.026] ***
logPIP	0.2	[0.032, 0.373] *
pHCaCl2	0	niet-significant
logSO4	0	niet-significant(p=0.15)
logK	X	correlatie met klei
logP	X	correlatie met PIP
HoogteBoven	0	niet-significant
AfstandRivier	0	niet-significant
Landgebruik	0	niet-significant

## Kaarten

De gemeten zinkwaarden zijn vrij gemiddeld in vergelijking met de andere niet-overstroomde gebieden. Er zijn enkele uitschieters in de voorspelde concentratie en de hoogste waarden worden gemeten in het middenstuk van het gebied dat dichtbij de rivier ligt. Voor cadmium loopt het verhaal bijna identiek.

Er werd geen enkele zinkwaarde gemeten boven de norm. Voor Cd was er slechts een overschrijding van de norm. De kans op overschrijding van de norm kon dus niet gemodelleerd worden.



Figuur 4-13 Predicties Zn en Cd in WALZ (mg/kg).

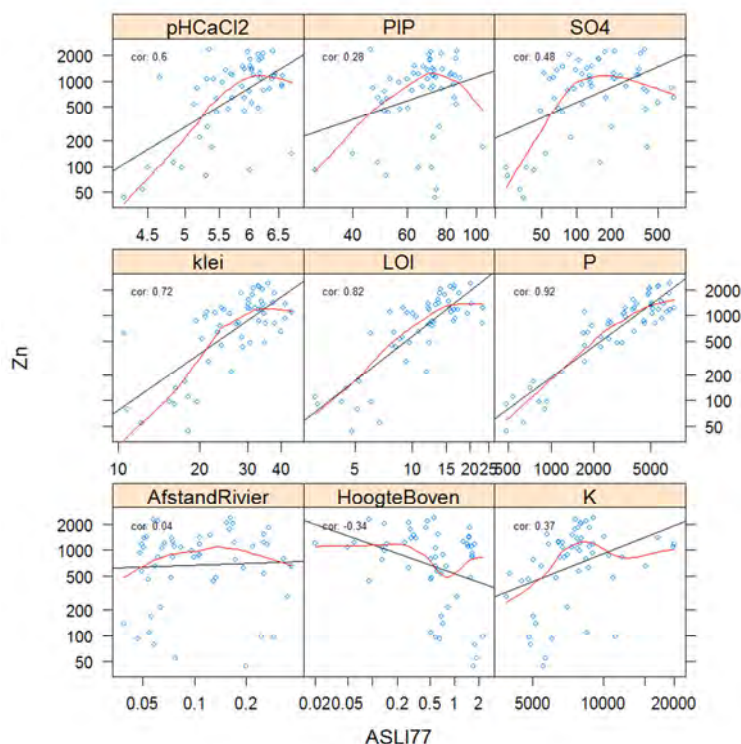
## 4.3 Reeds overstroomde gebieden

### 4.3.1 Anderstadt I en II en de polder van Lier

#### Model

Dit gebied is een samenstelling van 3 kleine gebiedjes, en blijft met 55 meetpunten nog altijd een gebied met een zeer beperkt aantal meetpunten. Het landgebruik hangt sterk samen met het deelgebied, maar tegelijk ook met de hoogte boven het laagste punt van het gebied. De deelgebieden liggen dan ook op sterk verschillende hoogte. Verder zijn klei, LOI en fosfor onderling zeer sterk gecorreleerd ( $> 0.75$ ).

Omdat LOI en klei beiden basisvariabelen zijn, worden ze apart getest in het model, en werd LOI behouden wegens iets betere statistische eigenschappen. In het uiteindelijke model blijven enkel pH, kalium, LOI en de hoogte over als verklarende variabelen. pH, kalium en LOI hangen positief samen met de zinkconcentratie, en lager gelegen delen bevatten meer zink. Dit laatste kan erop wijzen dat lager gelegen delen in het verleden vaker overstroomd werden maar kan net zo goed betekenen dat er onder grasland meer zink in de bodem zit.



Figuur 4-14 Verbanden met Zn voor het gebied ASLI.

Tabel 4-6 Modelcoëfficiënten voor het gebied ASLI.

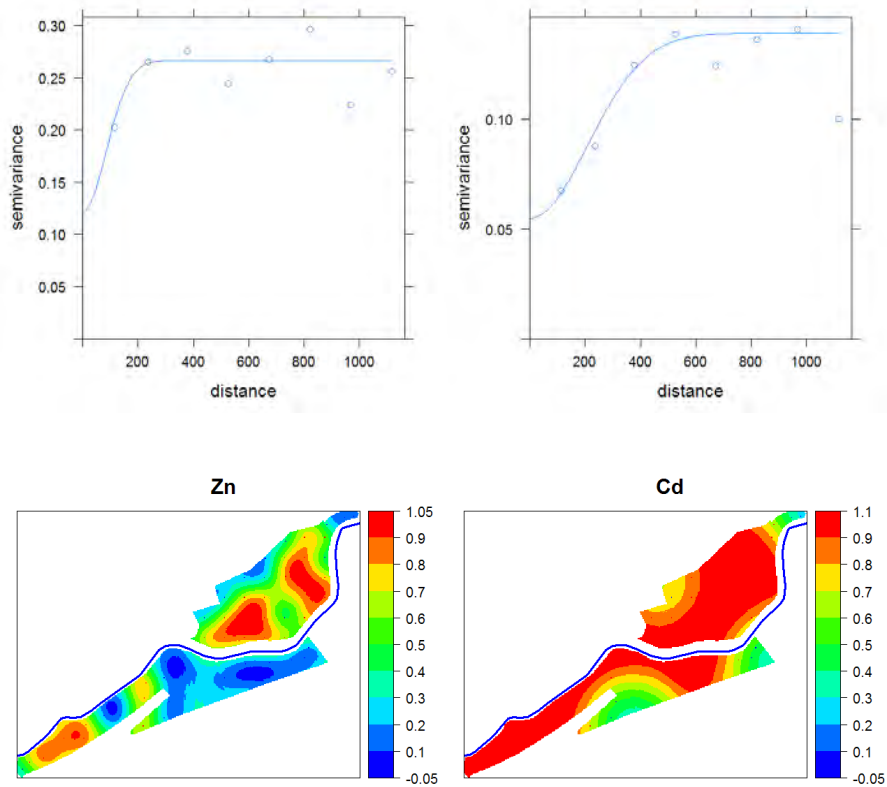
Var	Schatting	Interval95
N	55	NA
(Intercept)	-0.94	[-2.582, 0.708]
pH <sub>CaCl2</sub>	0.18	[0.054, 0.316] **
logK	0.61	[0.181, 1.036] **
LOI	0.05	[0.03, 0.06] ***
HoogteBoven	-0.28	[-0.451, -0.103] **
klei	X	correlatie met LOI en K
logSO <sub>4</sub>	0	niet-significant
logP	X	correlatie met LOI en klei
logPIP	0	niet-significant
AfstandRivier	0	niet-significant
Landgebruik	X	correlatie met hoogte

### Kaarten

De zinkconcentraties liggen in dit gebied heel wat hoger dan in de niet-overstroomde gebieden en dan vooral in dichtbij de rivier of in laaggelegen delen. Zoals aangegeven op de topokaart (Figuur 4-16) is Anderstadt II gedeeltelijk vijver. Hier werden geen bodemstalen genomen dus de predicties zijn hier minder betrouwbaar.

In tegenstelling tot de niet-overstroomde gebieden wordt de norm voor zink hier wel vaak overschreden. De kans dat de norm overschreden wordt, is vooral hoog in de polder van Lier en Anderstadt I. Voor cadmium ligt zowat het hele gebied boven de norm.

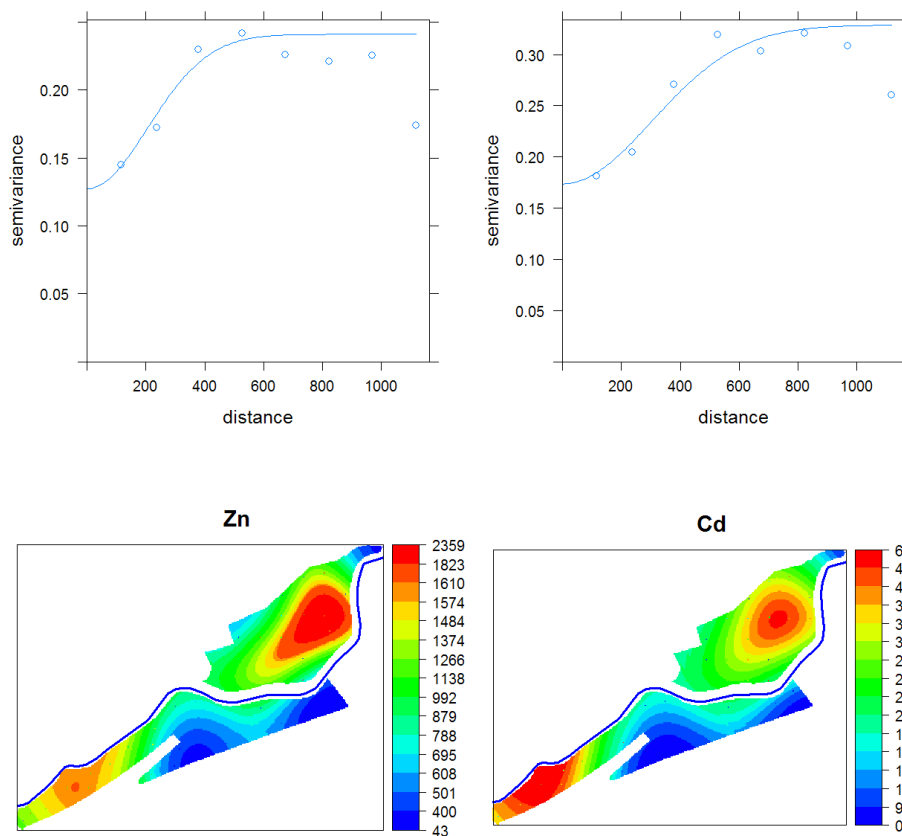




Figuur 4-15 Predicties Zn en Cd in ASLI (mg/kg).



Figuur 4-16 Topokaart van Anderstadt I en II en Polder van Lier.

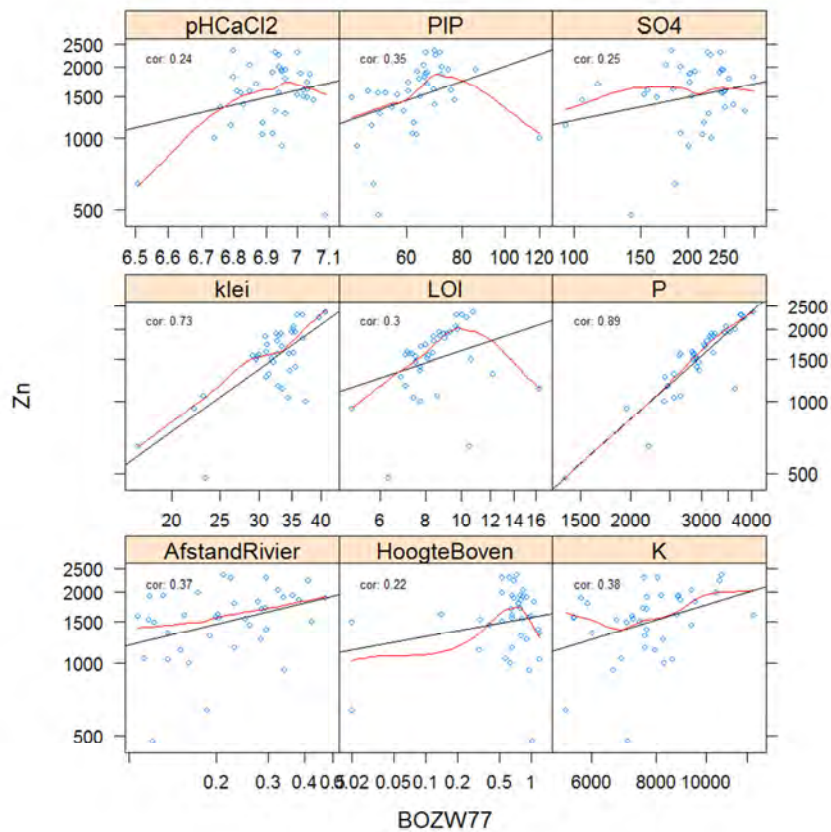


Figuur 4-17 Kans op overschrijding van de norm in ASLI.

### 4.3.2 Bovenzanden

#### Model

Bovenzanden is een klein gebied met slechts 38 meetpunten dat in het westen bestaat uit bosgebied en voor de rest uit weidegebied. Dit landgebruik is ook gekoppeld aan de hoogte binnen dit gebied, aangezien de weides voornamelijk voorkomen op de hoger gelegen meetpunten. Naast landgebruik wordt ook fosfor uit het model geweerd wegens een te grote correlatie met klei. Omdat de dataset zeer klein is, is het niet mogelijk een model te maken met alle verklarende variabelen er tegelijk in. Er valt zelfs nog een extra meetpunt weg omdat die coördinaten op dezelfde plaats heeft als een ander meetpunt. Wegens de kleine omvang zullen de termen sequentieel toegevoegd worden. Eens klei in het model zit valt kalium weg omdat deze vrij sterk gecorreleerd zijn ( $\rho = 0.56$ ). Hetzelfde geldt voor plantbeschikbare fosfor. De overgebleven variabelen zijn klei en afstand tot de rivier die positief samenhangen met het zinkgehalte. Voor de afstand tot de rivier is dat wel wat verassend. Eens er gecorrigeerd is voor deze variabelen blijkt ook dat er kans is dat de hoogte negatief samenhangt met het zinkgehalte.



Figuur 4-18 Verbanden met Zn voor het gebied BOZW.

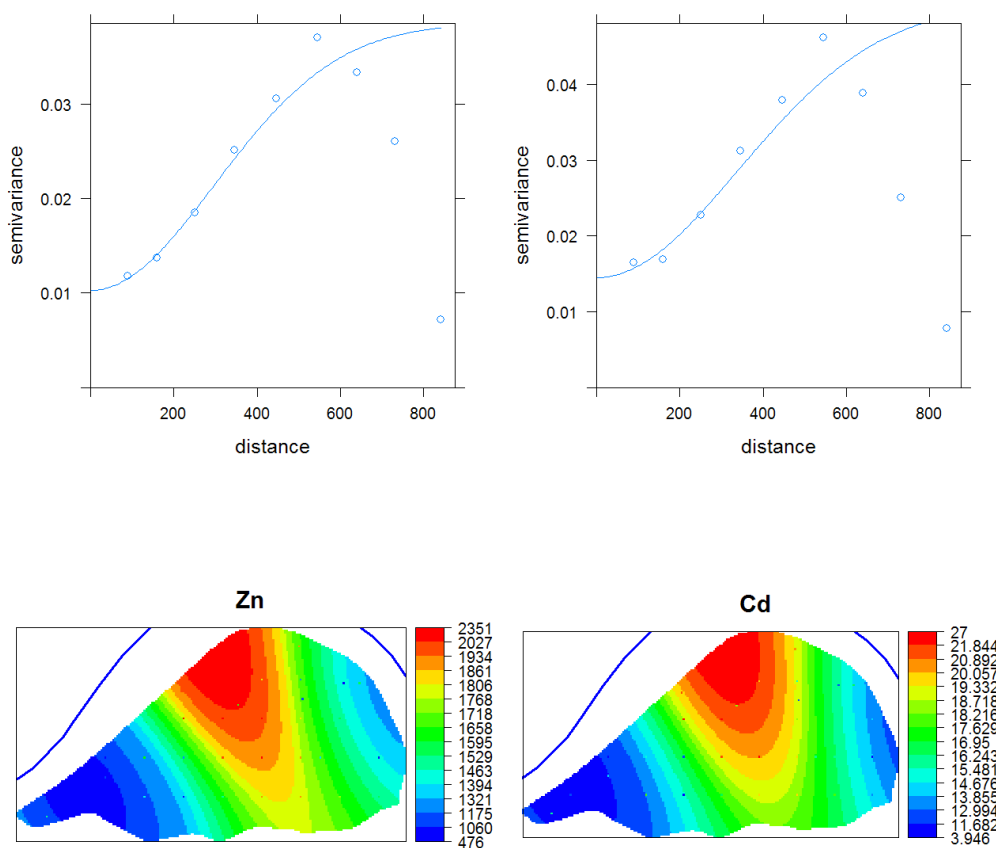
Tabel 4-7 Modelcoëfficiënten voor het gebied BOZW.

Var	Schatting	Interval95
N	37	NA
(Intercept)	2.62	[2.334, 2.906] ***
klei	0.02	[0.01, 0.026] ***
AfstandRivier	0.35	[-0.184, 0.891]
HoogteBoven	-0.15	[-0.279, -0.024] *
pHCaCl2	0	niet-significant
LOI	0	niet-significant
logSO4	0	niet-significant
logK	0	niet-significant
logP	X	correlatie met klei
logPIP	X	correlatie met klei
Landgebruik	X	correlatie met hoogte

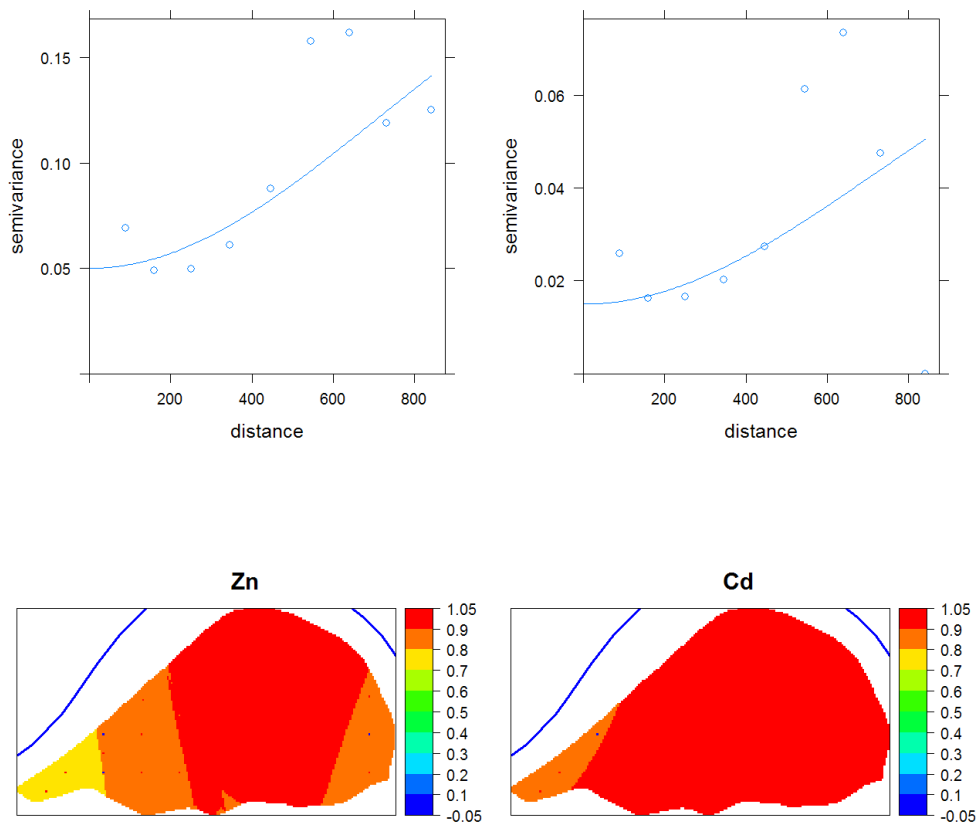
## Kaarten

Over het hele gebied is de zinkconcentratie hoog, maar in het westelijke punt (het bosgebied, waar de concentratie klei lager is) is dit iets lager.

De norm wordt in het grootste deel van het gebied overschreden voor zowel zink als cadmium.



Figuur 4-19 Predicties Zn en Cd in BOZW (mg/kg).



Figuur 4-20 Kans op overschrijding van de norm in BOZW.

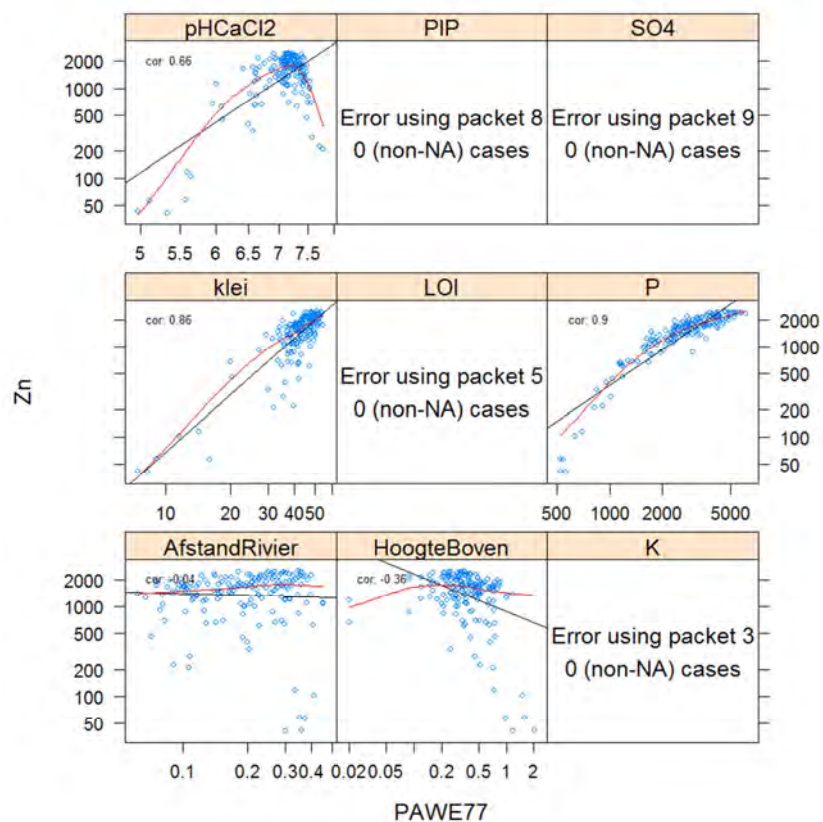
### 4.3.3 Paardeweide

#### Model

In Paardeweide zijn er 177 meetpunten, waarvan en 173 opgenomen zijn voor de analyse. De 3 punten in bosgebied zijn weggelaten, alsook een meting die dezelfde coördinaten kreeg als een andere meting. Verder zijn in Paardeweide ook niet altijd dezelfde variabelen opgemeten als in de andere gebieden. Kalium, sulfaat en plantbeschikbare fosfor kunnen dus niet gebruikt worden in de analyse. LOI is ook niet aanwezig, maar deze kan berekend worden uit het totaal organische koolstof volgens de formule  $LOI = 1.724 * TOC$ . Aangezien klei met de andere variabelen, behalve de afstand tot de rivier, sterk gecorreleerd is, wordt deze uit de analyse geweerd.

Bij verfijning van het model blijven enkel nog pH, LOI, en de hoogte boven de rivier over. Niettemin is de nodige voorzichtigheid geboden bij de interpretatie van de geschatte coëfficiënten en hun betrouwbaarheid, omdat alle variabelen wel op een of andere manier

met elkaar gecorreleerd zijn. Dus een model met klei alleen zou even zeer goed werken en een niet zoveel lagere verklaringskracht hebben.



Figuur 4-21 Verbanden met Zn voor het gebied PAWE. De foutmeldingen komen omdat deze variabelen niet opgemeten zijn voor dit gebied.

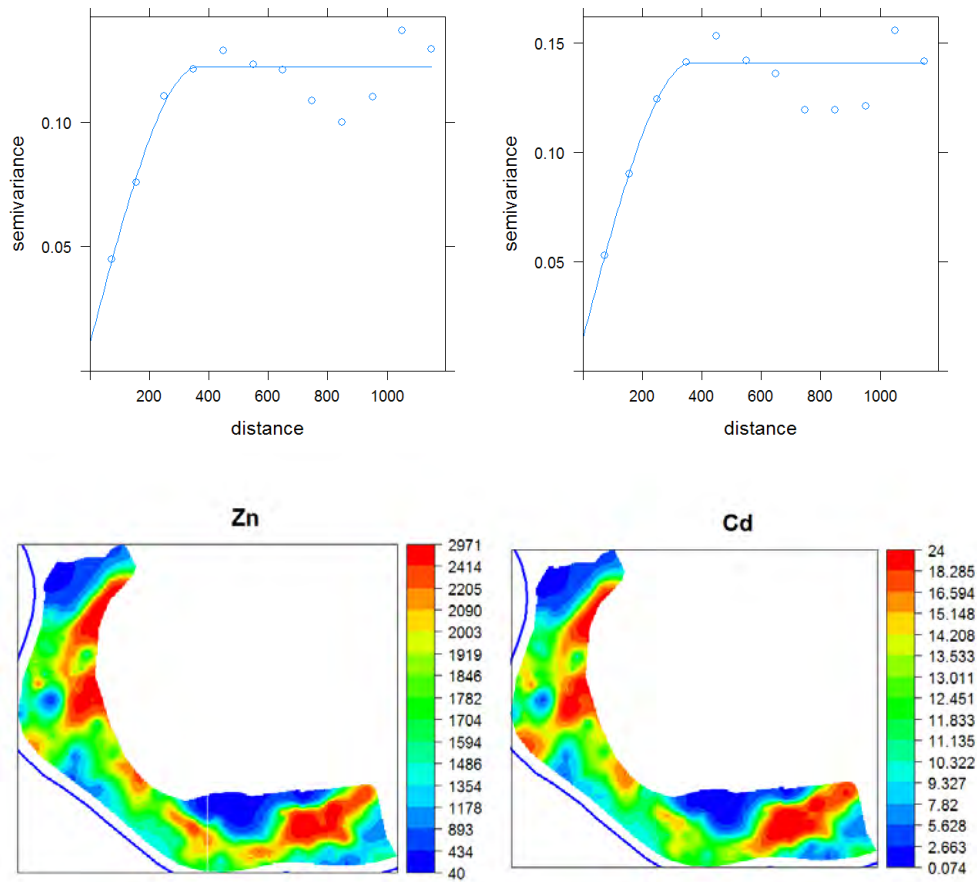
Tabel 4-8 Modelcoëfficiënten voor het gebied PAWE.

Var	Schatting	Interval95
N	173	NA
(Intercept)	-1.28	[-1.553, -0.998] ***
pHCaCl2	0.18	[0.147, 0.214] ***
LOI	0.02	[0.018, 0.024] ***
logP	0.83	[0.754, 0.916] ***
HoogteBoven	-0.13	[-0.188, -0.068] ***
klei	X	correlatie met de rest
afstand tot de rivier	0	niet-significant

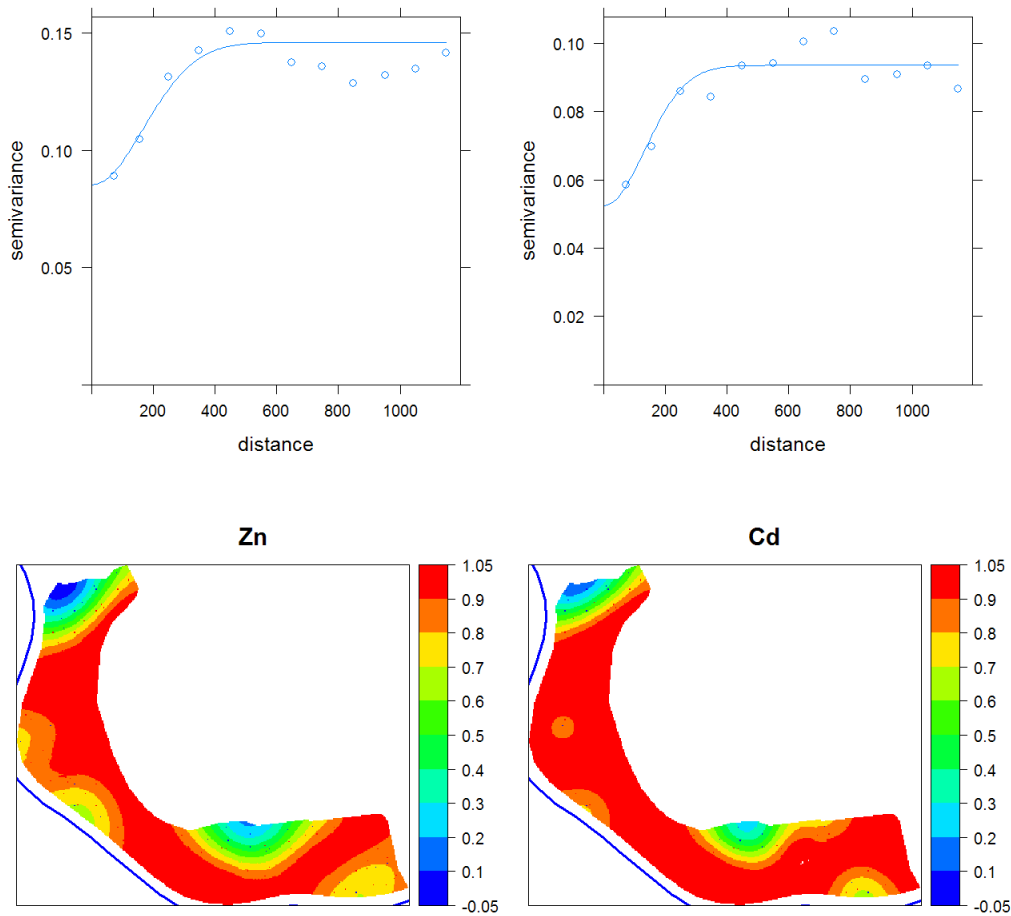
## Kaarten

De zinkconcentratie is hoog in dit gebied. Er zijn 2 clusters waar de concentratie duidelijk lager ligt. Voor cadmium is er iets gelijkaardigs gaande - niet onlogisch door de correlatie tussen beiden.

De kans dat de norm overschreden wordt is zeer hoog voor het grootste deel van het gebied. In de clusters met een lagere voorspelde concentratie is de kans dat de norm overschreden wordt veel lager.



Figuur 4-22 Predicties Zn en Cd in PAWE (mg/kg).



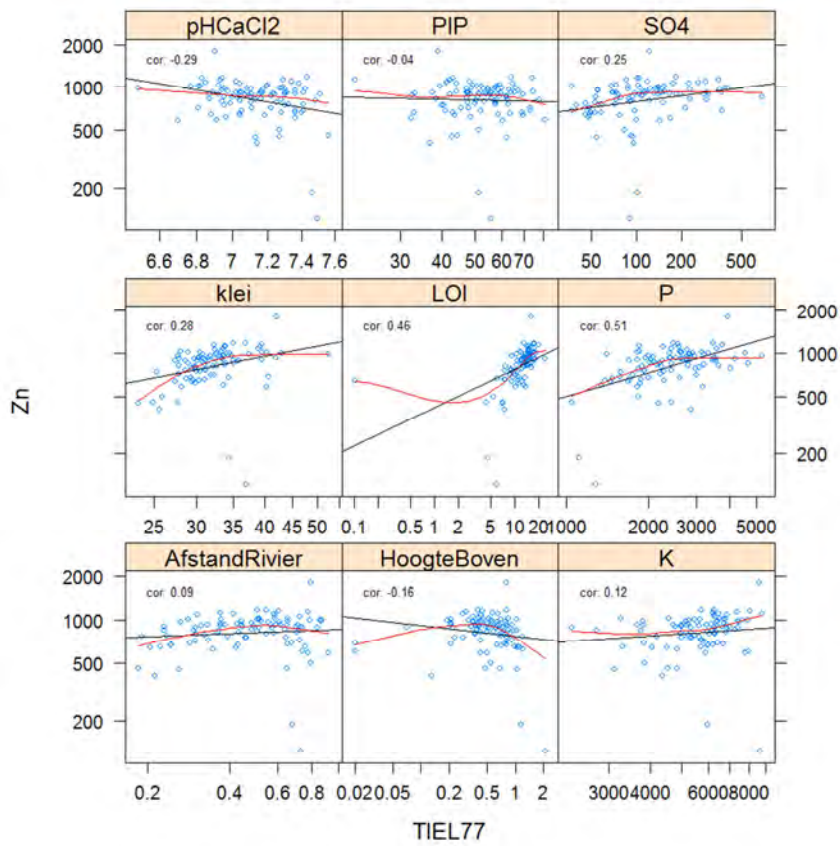
Figuur 4-23 Kans op overschrijding van de norm in PAWE.

#### 4.3.4 Tielrodebroek

##### Model

Dit gebied bestaat uit 94 punten. Eén punt in bosgebied wordt weggelaten, zodat er nog 93 meetpunten overblijven. Omdat de akkers over het algemeen dicht bij de rivier voorkomen dan de andere types landgebruik, zal deze variabele geweerd worden en opgevangen worden door de afstand tot de rivier. Verder zijn ook LOI en fosfor met ongeveer 0.6 gecorreleerd en als ze samen in het model voorkomen beïnvloeden ze elkaar duidelijk, dus wordt fosfor geweerd uit het model. Ook LOI en klei samen in het model geeft problemen, ondanks dat hun correlatie niet zo hoog - maar toch aanzienlijk - is. Als finaal model is er een positief verband met LOI en een negatief verband met de hoogte en de afstand tot de rivier.





Figuur 4-24 Verbanden met Zn voor het gebied TIEL.

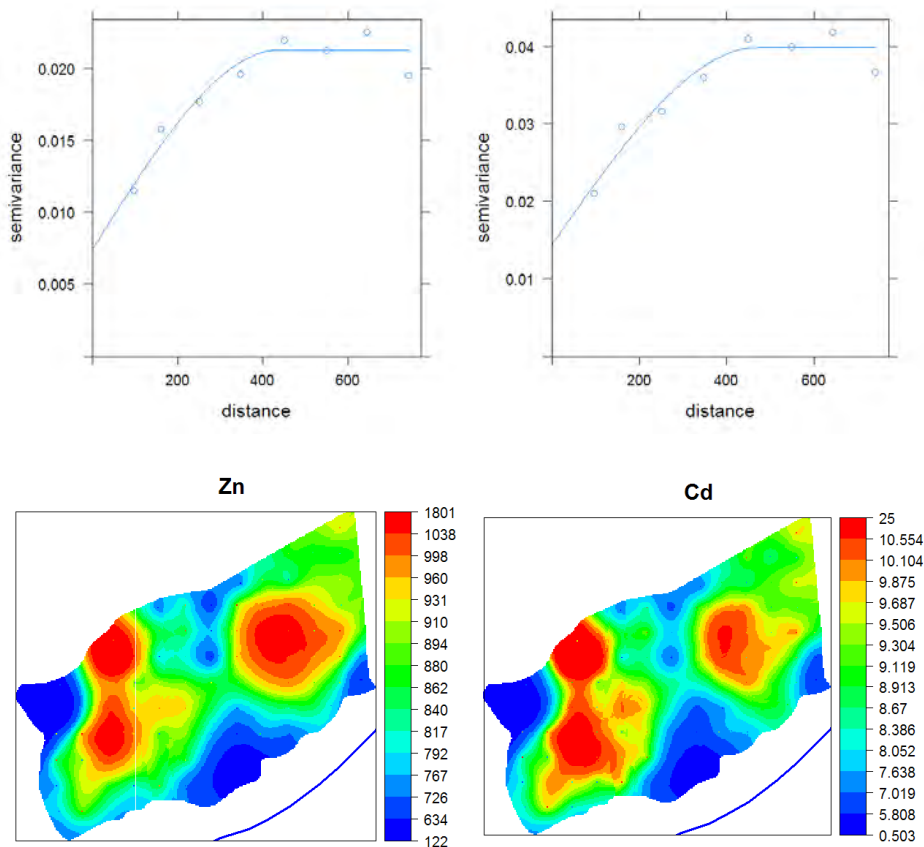
Tabel 4-9 Modelcoëfficiënten voor het gebied TIEL.

Var	Schatting	Interval95
N	93	NA
(Intercept)	2.79	[2.655, 2.924] ***
LOI	0.02	[0.018, 0.03] ***
HoogteBoven	-0.11	[-0.193, -0.032] **
AfstandRivier	-0.24	[-0.434, -0.05] *
klei	X	correlatie met LOI
pHCaCl2	X	correlatie met LOI
logSO4	X	correlatie met LOI
logK	X	correlatie met LOI
logP	X	correlatie met LOI
logPIP	0	niet-significant
Landgebruik	X	correlatie met afstand tot rivier

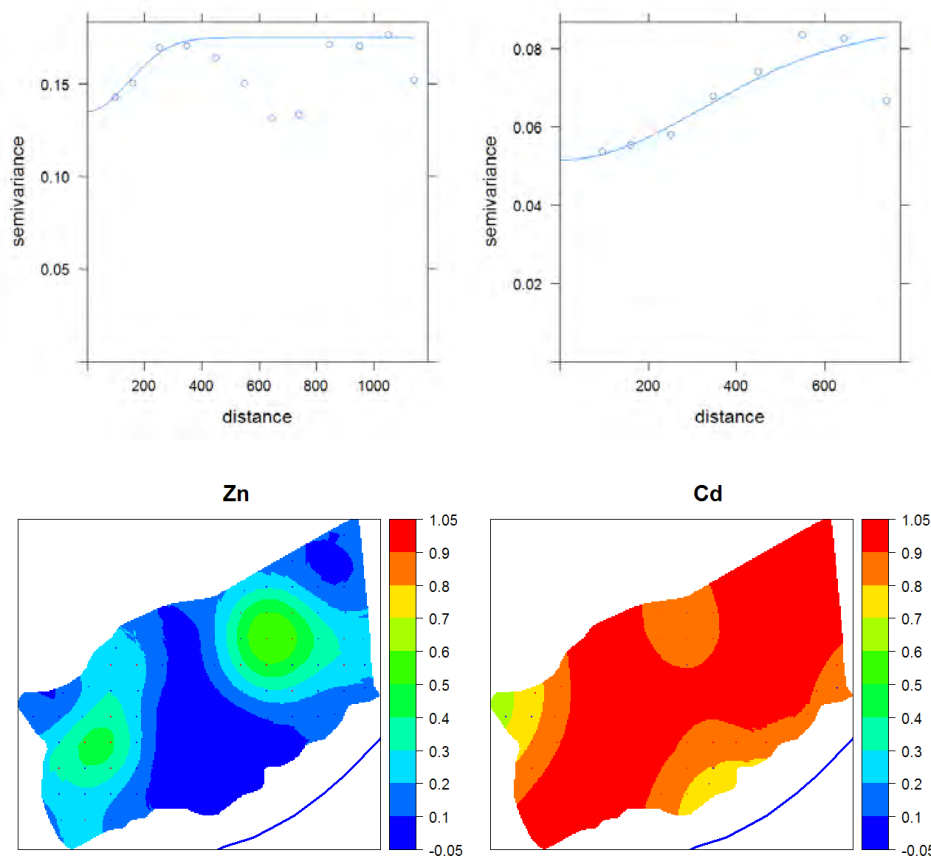
## Kaarten

Net zoals bij de andere overstromde gebieden ligt de concentratie zink hier vrij hoog. In het westen is er een uitschieter naar beneden, maar dit komt waarschijnlijk omdat hier één uitzonderlijk hoog gelegen meetpunt aanwezig is.

De kans op overschrijding van de norm voor zink is hier kleiner dan bij de andere overstromde gebieden. Voor cadmium wordt de norm wel duidelijk overschreden.



Figuur 4-25 Predicties Zn en Cd in TIEL (mg/kg).



Figuur 4-26 Kans op overschrijding van de norm in TIEL.

#### 4.4 Algemene verschillen en gelijkenissen tussen gebieden

Er zijn enkele zeer duidelijke verschillen en gelijkenissen tussen de gebieden. De verschillen bestaan er vooral in hoe sterk de gelijkenis van de metaalconcentraties op korte afstand tussen meetpunten is en evolueert bij toenemende afstand. Dit kan gebeuren door de nugget (snijpunt van de curve met de y-as in Figuur 4-27) met de sill (de asymptotische waarde) te vergelijken. Deze verhouding verschilt sterk van gebied tot gebied. Ook de range kan van gebied tot gebied sterk verschillen. Dit heeft deels te maken met het verschil in grootte en vorm van de gebieden, en dat er geen rekening gehouden is met anisotropie (=correlatie tussen punten daalt sneller in de ene richting dan de andere richting). Het algemene model (Hoofdstuk 5) heeft een nog grotere range, wat waarschijnlijk deels komt omdat de ruimtelijke structuur van zink voor ieder gebied verschillend is. De grotere range is dan mogelijk een correctie voor het feit dat er verschillende gebieden samengenomen zijn.

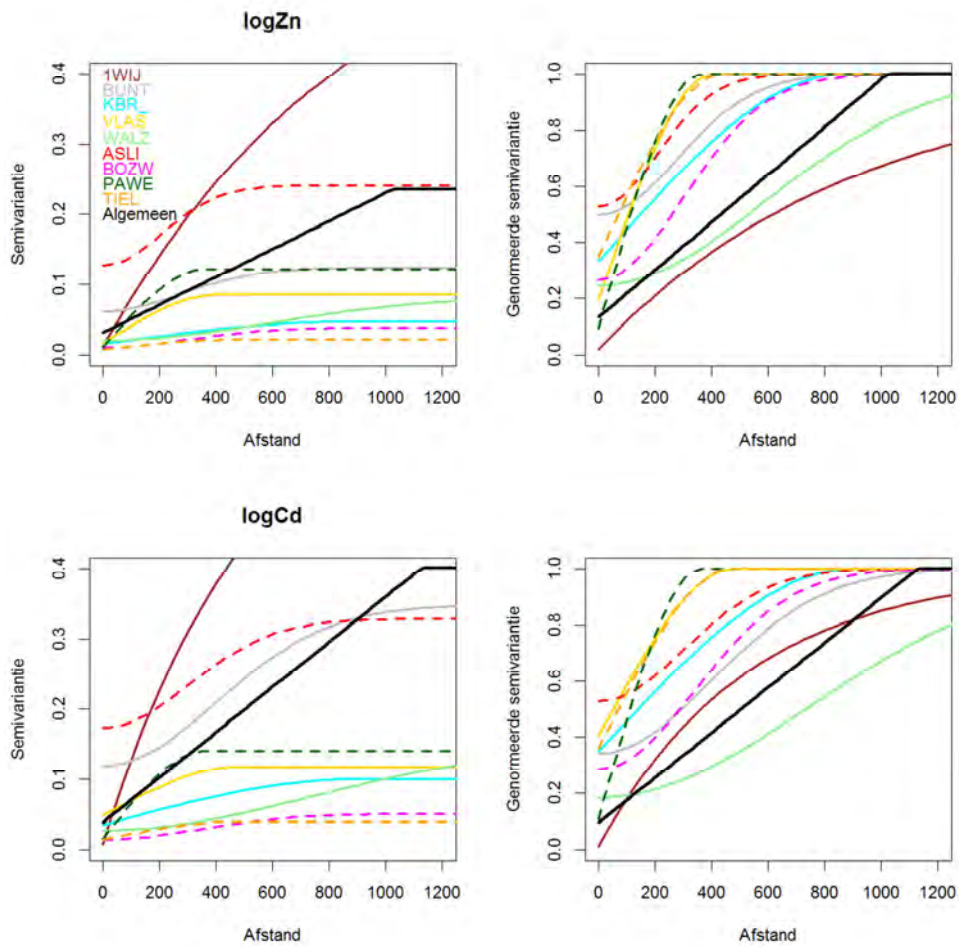
Een ander verschil doet zich voor tussen de overstroomde en niet-overstroomde gebieden. De overstroomde gebieden hebben over het algemeen veel hogere concentraties van zware metalen. Deze metaalconcentraties zijn gelijkmatiger verspreid over het gebied in tegenstelling tot de niet-overstroomde gebieden waar de concentraties zeer afhankelijk zijn van waar precies in het gebied gemeten wordt. Bij ieder overstroomd gebied wordt de norm vaak tot bijna overal overschreden, terwijl dit bij de nog-niet-overstroomde gebieden veel

minder gebeurt. De cadmiumnorm wordt in tegenstelling tot de zinknorm ook al regelmatig overschreden in de niet-overstroomde gebieden. De range lijkt in de overstroomde gebieden ook iets kleiner dan bij de niet-overstroomde gebieden, maar om dit hard te maken hebben we meer gebieden nodig om te testen. Als dit patroon algemeen voorkomt dan is dat waarschijnlijk omdat de concentraties meer verdeeld worden over het gebied en dus de zeer lokale effecten uitgevlakt worden.

Over het algemeen liggen de zinkconcentraties een stuk lager in de niet-overstroomde gebieden. De kaart kleurt voornamelijk blauw en groen, wat duidt op een concentratie van minder dan ca. 1000 mg/kg. Bij de wel reeds overstroomde gebieden valt op dat de kaart veel roder kleurt en dat de meeste concentraties boven de 1000 mg/kg liggen. Mogelijk kunnen gebieden bij overstroming meer verontreinigd te worden. Omdat er geen nultoestand gekend is van de gebieden die reeds overstroomd zijn is het echter mogelijk dat deze een hogere uitgangskonzentratie hadden. De huidige sedimentkwaliteit kan bovendien ook verschillen van deze van de voorbije drie decennia.

### **Kriging**

De krige variantie, wat een maat is voor de betrouwbaarheid van de zinkconcentratievoorspellingen is in ieder gebied verschillend. Wijmeers wijkt het meest af van allemaal, omdat hier een exponentieel variogrammodel het beste leek. De range bedraagt hier ongeveer 900 m wat betekent dat na 900m ongeveer 95% van de autocorrelatie is weggevallen. Aangezien het gebied maar ongeveer 1500m overspant zijn de meeste waarnemingen met elkaar gecorreleerd. Ook Bovenzanden en De Bunt hebben een zeer grote range in vergelijking met de gebiedsgrootte.



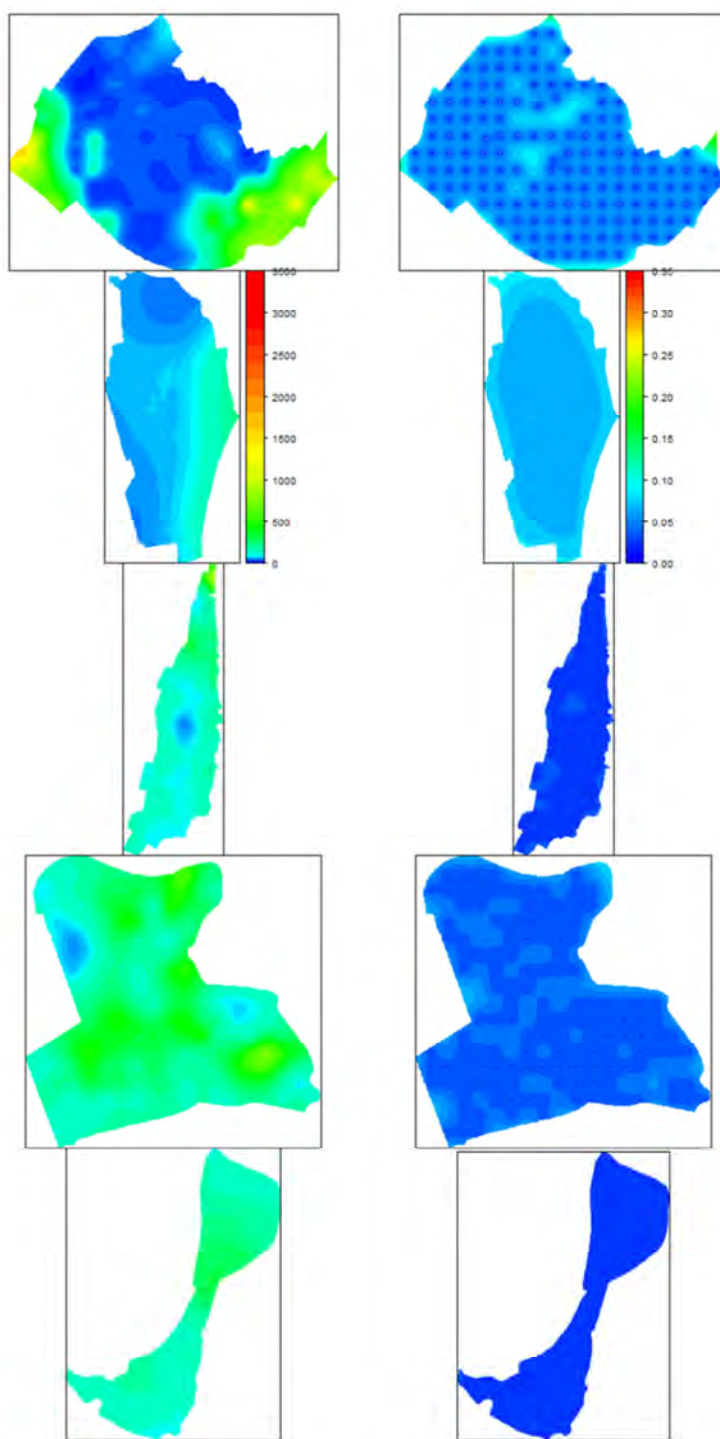
Figuur 4-27 Vergelijking van de semivariogrammen over de verschillende gebieden. Bovenaan staan de semivariogrammen voor logZink van alle gebieden en ook van het algemene model van het volgend hoofdstuk. Onderaan staan deze voor cadmium. In de rechter figuur worden semivariogrammen genormeerd.

Tabel 4-10 Vergelijking variogrammen tussen de gebieden voor zink. Het type model, range (maximale afstand waarop punten gecorreleerd zijn), nugget (ruisvariantie), var inf (variantie tussen punten ver van elkaar), Cor 0m (1 - signaal/ruisverhouding) worden weergegeven.

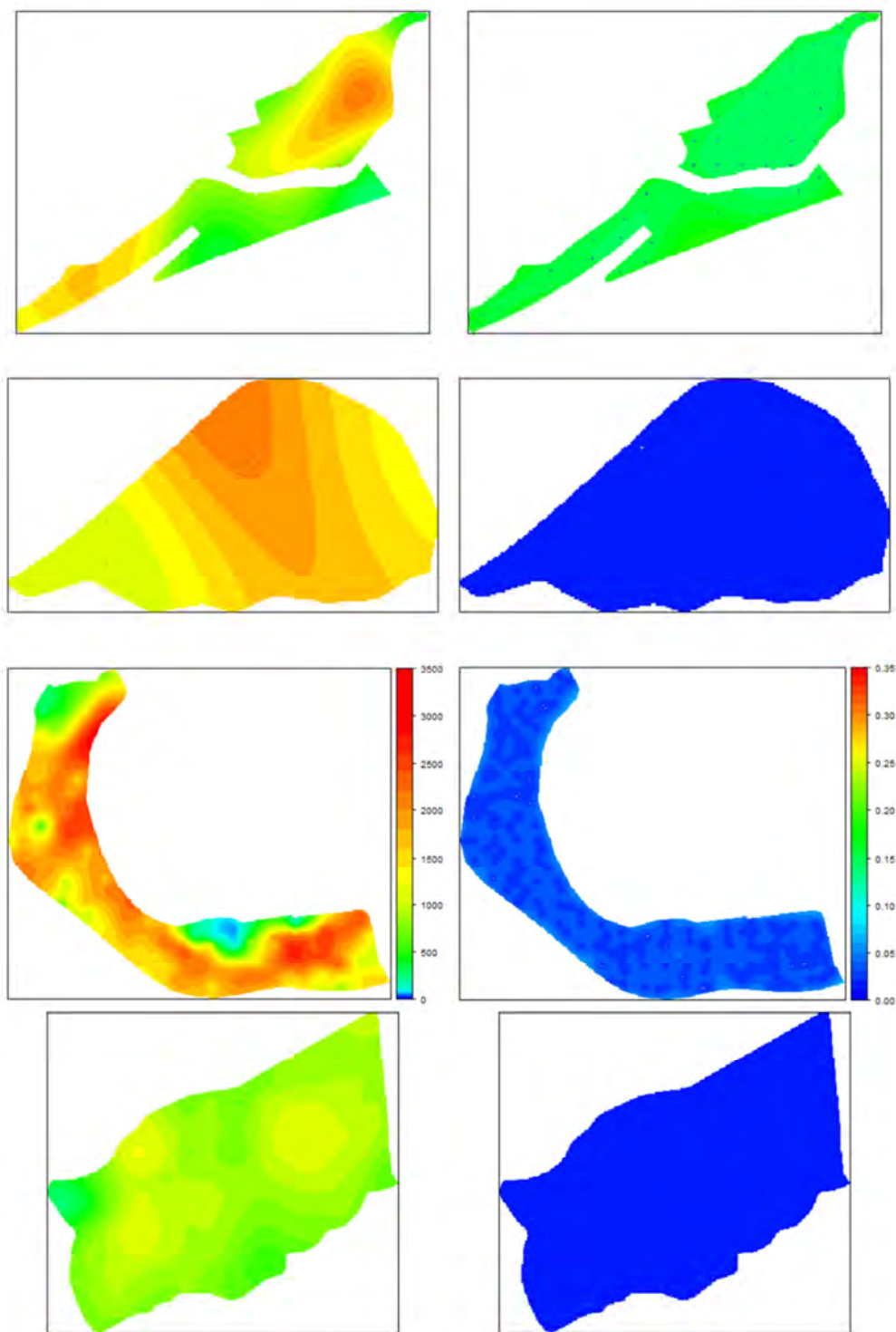
Gebied	Model	Nugget	Psill	Var_inf	Cor_0m	Range
1WIJ	Exp	0.012	0.662	0.674	0.982	917
BUNT	Gau	0.062	0.062	0.124	0.503	390
KBR	Sph	0.016	0.031	0.047	0.668	872
VLAS	Sph	0.017	0.069	0.085	0.804	407
WALZ	Gau	0.02	0.062	0.082	0.753	828
ASLI	Gau	0.127	0.114	0.241	0.472	291
BOZW	Gau	0.01	0.028	0.039	0.734	420
PAWE	Sph	0.011	0.111	0.122	0.908	365
TIEL	Sph	0.007	0.014	0.021	0.651	440
Algemeen	Lin	0.031	0.203	0.235	0.867	1022

Tabel 4-11 Vergelijking variogrammen tussen de gebieden voor cadmium. Het type model, range (maximale afstand waarop punten gecorreleerd zijn), nugget (ruisvariantie), var inf (variantie tussen punten ver van elkaar), Cor 0m (1 - signaal/ruisverhouding) worden weergegeven.

Gebied	Model	Nugget	Psill	Var_inf	Cor_0m	Range
1WIJ	Exp	0.008	0.701	0.709	0.989	531
BUNT	Gau	0.119	0.229	0.348	0.658	564
KBR	Sph	0.035	0.065	0.1	0.652	894
VLAS	Sph	0.048	0.07	0.118	0.595	480
WALZ	Gau	0.028	0.122	0.149	0.815	1047
ASLI	Gau	0.174	0.155	0.329	0.472	428
BOZW	Gau	0.014	0.036	0.051	0.714	482
PAWE	Sph	0.015	0.126	0.141	0.892	365
TIEL	Sph	0.014	0.025	0.04	0.639	472
Algemeen	Lin	0.038	0.363	0.402	0.904	1127



Figuur 4-28 Alle predicties en krige-varianties voor zink onder 1 noemer voor de niet-overstroomde gebieden.



Figuur 4-29 Alle predicties en krigé-varianties voor zink onder 1 noemer voor de overstroomde gebieden.



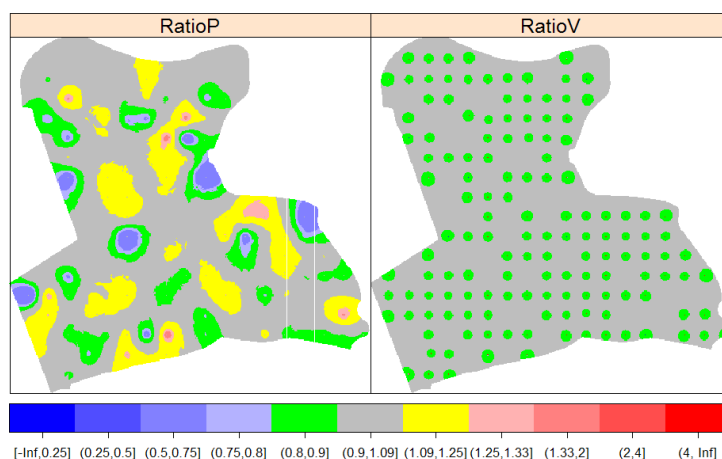
## 4.5 Alternatieve analysemethoden

### 4.5.1 Cokriging

#### Extra informatie op de meetpunten zelf

Cokriging is een multivariaat alternatief om metingen te interpoleren over een gebied. Bij gewone kriging gebruikt men enkel de metingen van de interessevariabele zelf, terwijl bij cokriging de voorspelling van de interessevariabele wordt geschat gebruik makend van verschillende variabelen. Natuurlijk heeft dit enkel zin als deze andere variabelen genoeg verband hebben met de interessevariabele. De voorspelling van de interessevariabele zal nu gebruik maken van zowel haar eigen meting, als de meting van andere stoffen op dezelfde locatie. De belangrijkste informatie zit vervat in hoe de verschillende variabelen met elkaar variëren (covariëren). Door van deze informatie gebruik te maken, krijgen we meer inzicht hoe de variabelen samenhangen, waardoor ook voorspellingen in de tussenliggende punten van deze informatie gebruik kunnen maken, meer dus dan van de metingen van de individuele stoffen alleen.

Hier is de cokriging toegepast op het gebied VLAS. Uit Figuur 4-30 kunnen we afleiden dat de impact op de predicties niet zo enorm groot is. Ook bij de variantie is het verschil niet groot. Het gebied net rond het meetpunt wordt wel iets beter voorspeld. De best correlerende variabele met het zinkgehalte was fosfor met een correlatie van 0.77. De andere variabelen waren voor ongeveer 0.50 gecorreleerd. Vermoedelijk is een correlatie van 0.77 nog niet genoeg om echt een meerwaarde te bieden als op het punt ook de zinkconcentratie gemeten werd.



Figuur 4-30 Relatief verschil in predictie (links) en variantie (rechts) uitgedrukt als een verhouding tussen de gewone krigingvoorspelling en de cokrigingvoorspelling van zink. De verhouding van de varianties is de verhouding in de log-schaal.

## Gebruik maken van het digitale hoogtemodel

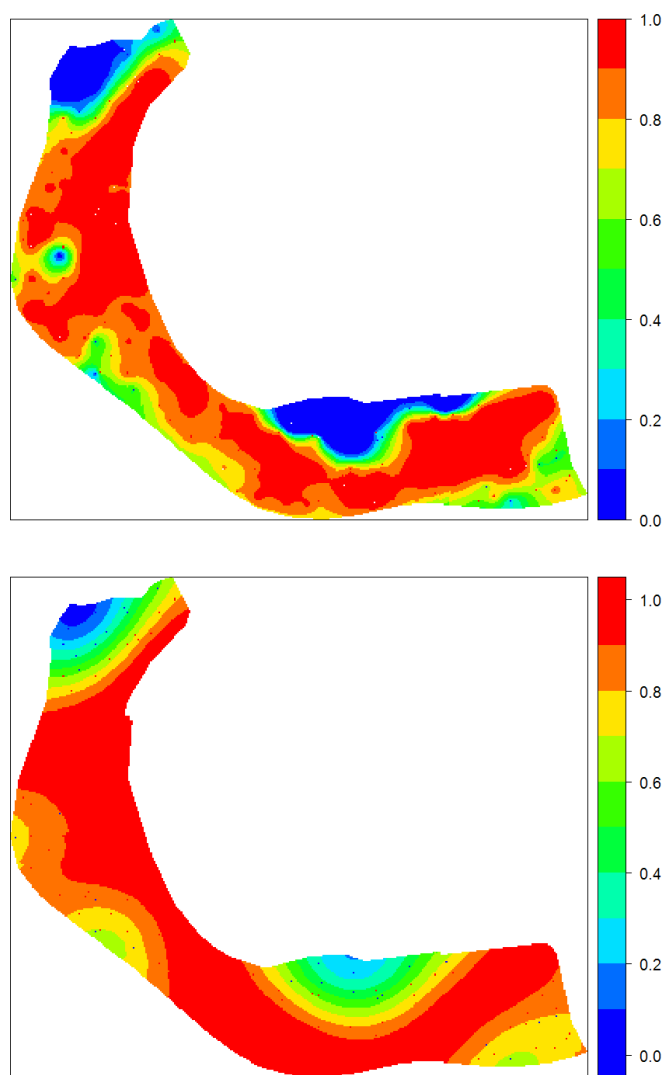
Omdat voor het gebied VLAS de hoogte zeer zwak gecorreleerd is met de zinkconcentratie wordt de cokriging hier hernomen voor het gebied 1WIJ. Hier is er een negatieve correlatiewaarde van -0.5, maar ook een sterke positieve correlatie tussen zink en het kleigehalte van ongeveer 0.9. De kriging met zink alleen wordt vergeleken met deze met zink en klei en deze met daaraan nog eens het hoogtemodel gekoppeld.

Een meerwaarde had deze methode echter niet, en de meeste schattingen bleven binnen de plus of min 5% van de originele schatting. Een belangrijke reden hiervoor is dat er niet kan gewerkt worden met een kleinere afstand tussen de punten dan het grid toelaat, aangezien cokriging vereist dat de ranges en de minimale variogramafstand gelijk moet zijn voor alle variabelen.

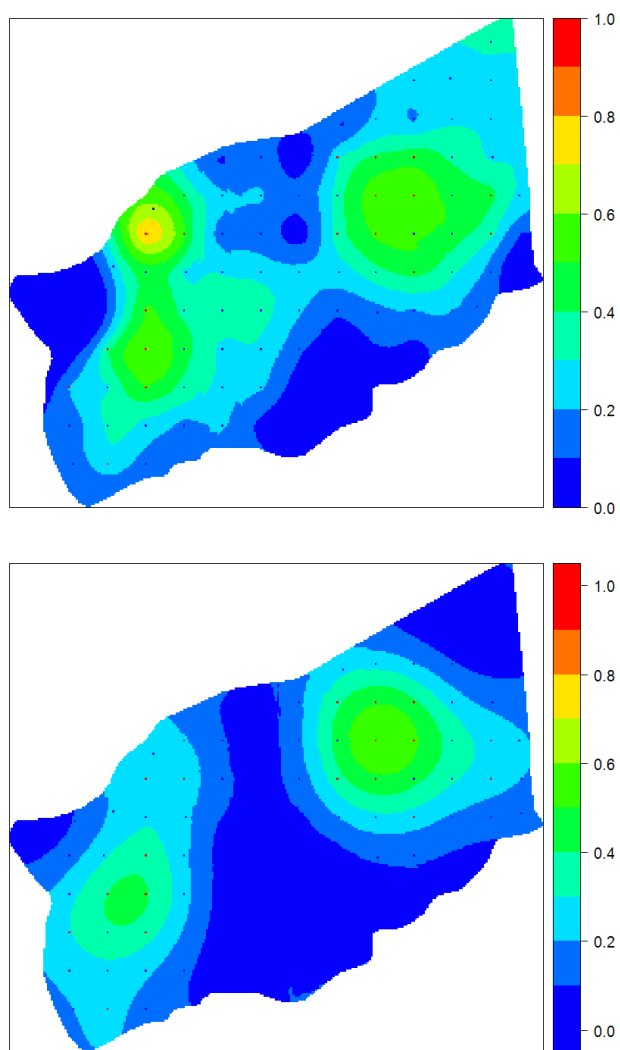
### 4.5.2 Stochastische methode voor kans op overschrijding norm

Om de kans op overschrijding van de norm te bepalen werd gebruik gemaakt van indicator kriging, waarbij het al dan niet boven de norm liggen werd uitgedrukt als een binaire variabele, die binnen het gebied geïnterpoleerd werd. Dit heeft als nadeel dat er geen rekening gehouden wordt met hoeveel een punt boven de norm ligt. Hoe deze stochastische methode te werk gaat is reeds in de methodologie beschreven. In de methodologie werd ook een tweede methode beschreven die alle variabelen nodig voor de berekening van de norm interpoleert voor het ganse gebied. Hier volgen twee voorbeelden om het verschil tussen beide methodes te illustreren. De gekozen gebieden zijn PAWE en TIEL en de gekozen meetvariabele is zink.

In grote lijnen blijft de voorspelling van de overschrijdingskans gelijk, maar de stochastische methode kan deze iets verfijnen (Figuur 4-31). Dat het verschil tussen de gemeten concentratie en de norm ook zijn rol speelt in het bepalen van de overschrijdingskans is een goede zaak, maar heeft wel als minpunt dat een zeer grote overschrijding lokaal zwaar zal doorwegen op de voorspellingen rond zo een punt. Voor Tielrodebroek is ook duidelijk dat het verschil niet zo groot is tussen de indicator kriging en de stochastische benadering.



Figuur 4-31 Indicator Kriging (onder) en een eigen vorm van stochastische voorspelling (boven) voor de overschrijdingskans van de zinknorm voor het gebied PAWE.



Figuur 4-32 Indicator Kriging (onder) en een eigen vorm van stochastische voorspelling(boven) voor de overschrijdingskans van de zinknorm voor het gebied TIEL.

## 5 Algemene analyse

### 5.1 Resultaten van het veralgemeend model

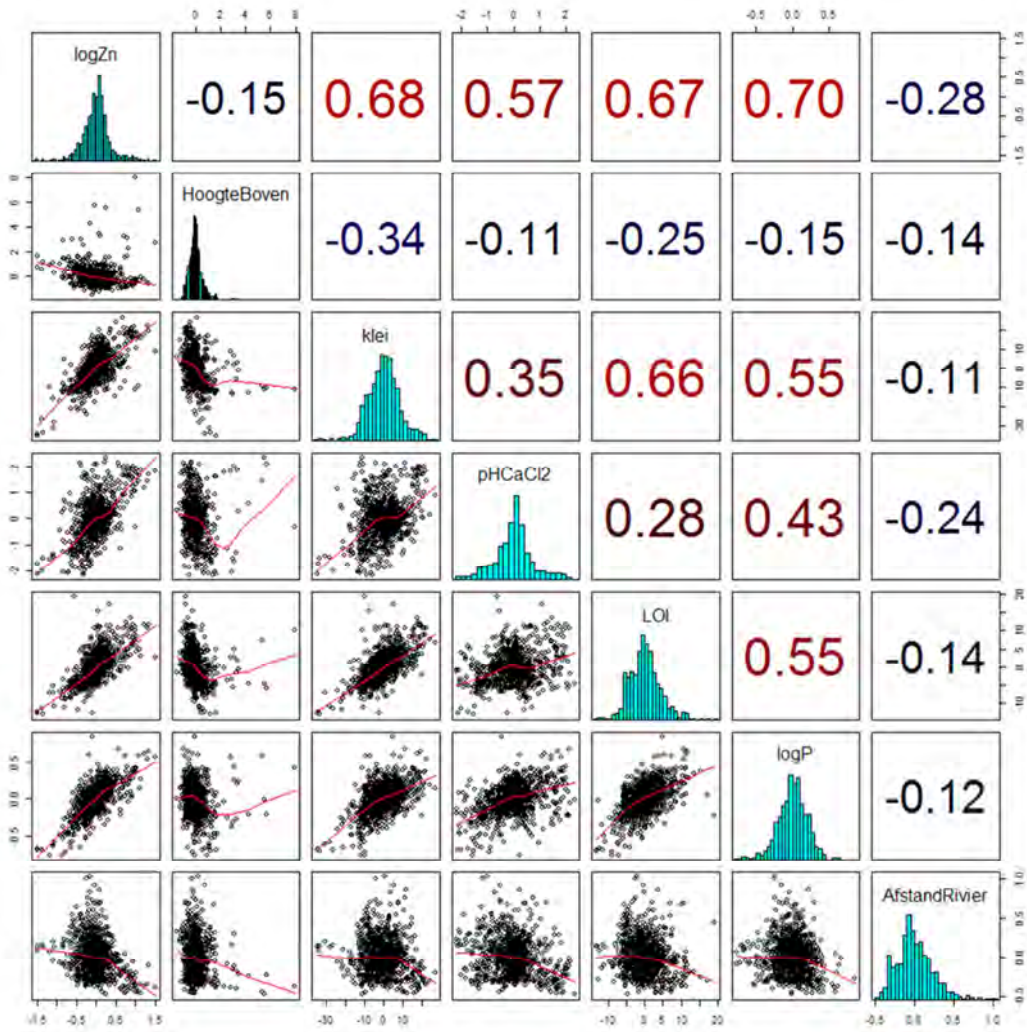
Omdat we ruimtelijke correlatie verwachten tussen de meetpunten binnen een gebied, moet hiervoor in het model gecorrigeerd worden. Zowel een sferisch als een exponentieel model zijn uitgetest en het exponentiële model kwam iets beter uit de verf. Naast de wegingsfactor voor de variantie per gebied die we vastgelegd hebben, bevat het algemene model de basisvariabelen die in ieder gebied opgemeten zijn (hoogte, afstand tot de rivier, klei, pH, LOI, logP) en hun interactie met het al dan niet overstroomd zijn van een gebied. Verder wordt ook de interactie tussen afstand en hoogte meegenomen.

Heel wat van de variabelen blijken duidelijk significant te zijn in het algemene model. De enige effecten die wegvallen zijn alle variabelen die de afstand tot de rivier gebruiken en de interactie tussen klei en het al dan niet overstroomd zijn.

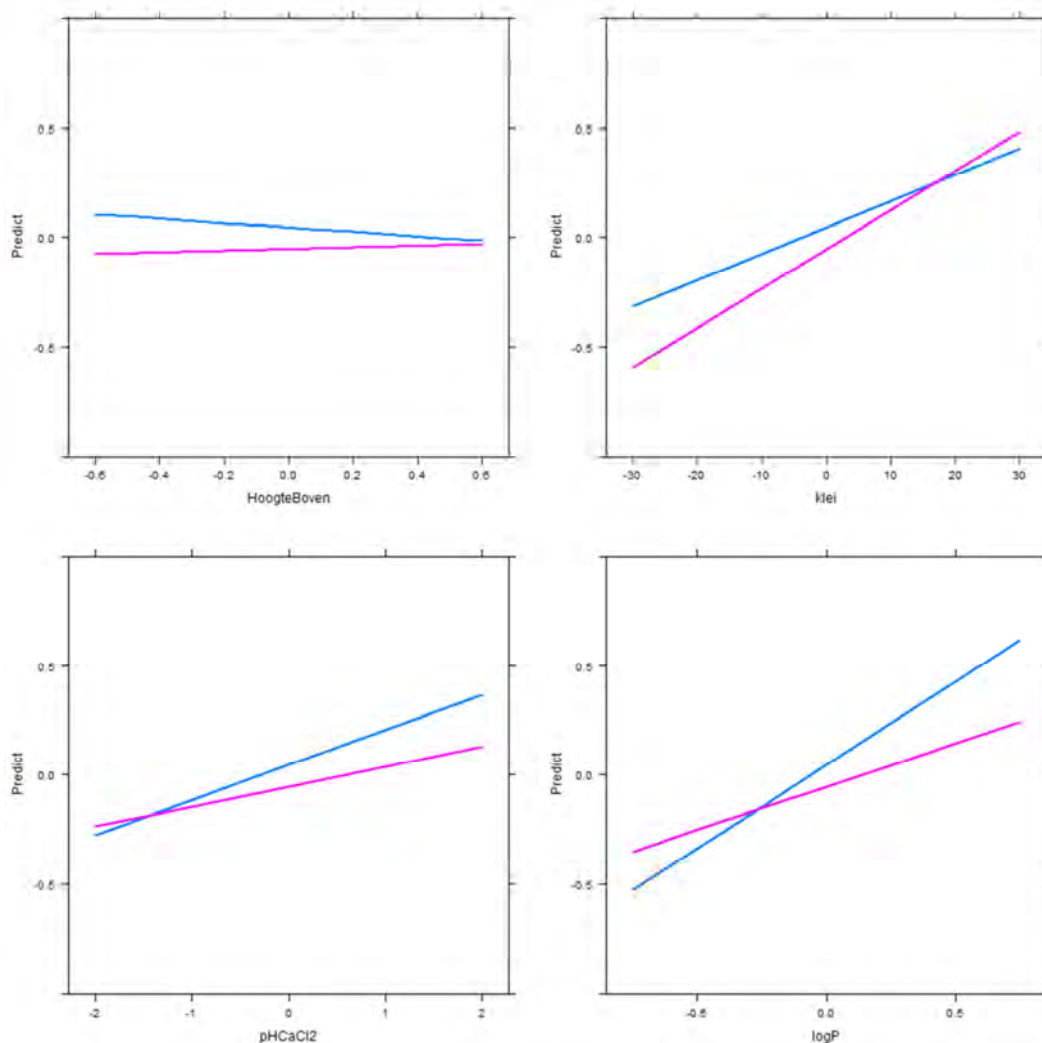
In Figuur 5-2 is het verband te vinden tussen de zinkconcentratie en de verklarende variabelen. Er is telkens één variabele die gevarieerd wordt, terwijl de rest op het gemiddelde, wat hier 0 is door het centreren van de data, gefixeerd wordt. Bij klei is ook LOI mee gevarieerd, omdat deze vrij sterk gecorreleerd zijn met elkaar. Aangezien bij de figuren rekening is gehouden met het bereik van de data, kan worden geconcludeerd dat hoogte en pH een kleiner effect hebben dan klei+LOI en logP. De interactie met het al dan niet overstroomd zijn beïnvloedt deze trend. Voor pH en fosfor is de samenhang met zink sterker in de overstroomde gebieden, terwijl de samenhang met klei en dus ook LOI sterker is in niet-overstroomde gebieden. Hier moeten we voorzichtig zijn met uitspraken, omdat er maar een heel beperkt aantal gebieden in het model opgenomen zijn. Enkel aan de hand van een dataset met meer overstroomde en niet-overstroomde gebieden, of met gebieden met een gekende toestand voor en na overstromen, kan met zekerheid aangetoond worden of dit komt door de dataset of dat dit een algemeen fenomeen is.

Tabel 5-1 Coëfficiënten van het algemene model.

Var	Schatting	Interval95
N	1002	
(Intercept)	0.05	[-0.007, 0.099] .
HoogteBoven	-0.1	[-0.144, -0.057] ***
klei	0.01	[0.004, 0.007] ***
pHCaCl2	0.16	[0.127, 0.19] ***
LOI	0.01	[0.009, 0.016] ***
logP	0.76	[0.681, 0.832] ***
Overstroomdneen	-0.1	[-0.171, -0.027] **
LOI:Overstroomdneen	0.01	[0.007, 0.017] ***
pHCaCl2:Overstroomdneen	-0.07	[-0.103, -0.032] ***
logP:Overstroomdneen	-0.36	[-0.471, -0.254] ***
HoogteBoven:Overstroomdneen	0.14	[0.091, 0.186] ***



Figuur 5-1 Verbanden algemene gebiedsoverschrijdende analyse.



Figuur 5-2 Predicties voor de algemene analyse, telkens voor 1 variabele die varieert terwijl de rest op het gemiddelde (0) is gefixeerd en er rekening is gehouden met de range van de variabelen om onderlinge vergelijkingen mogelijk te maken. LOI covarieert met klei wegens hun correlatie. De blauwe lijn is voor overstromde gebieden, de magenta lijn voor niet overstromde gebieden.

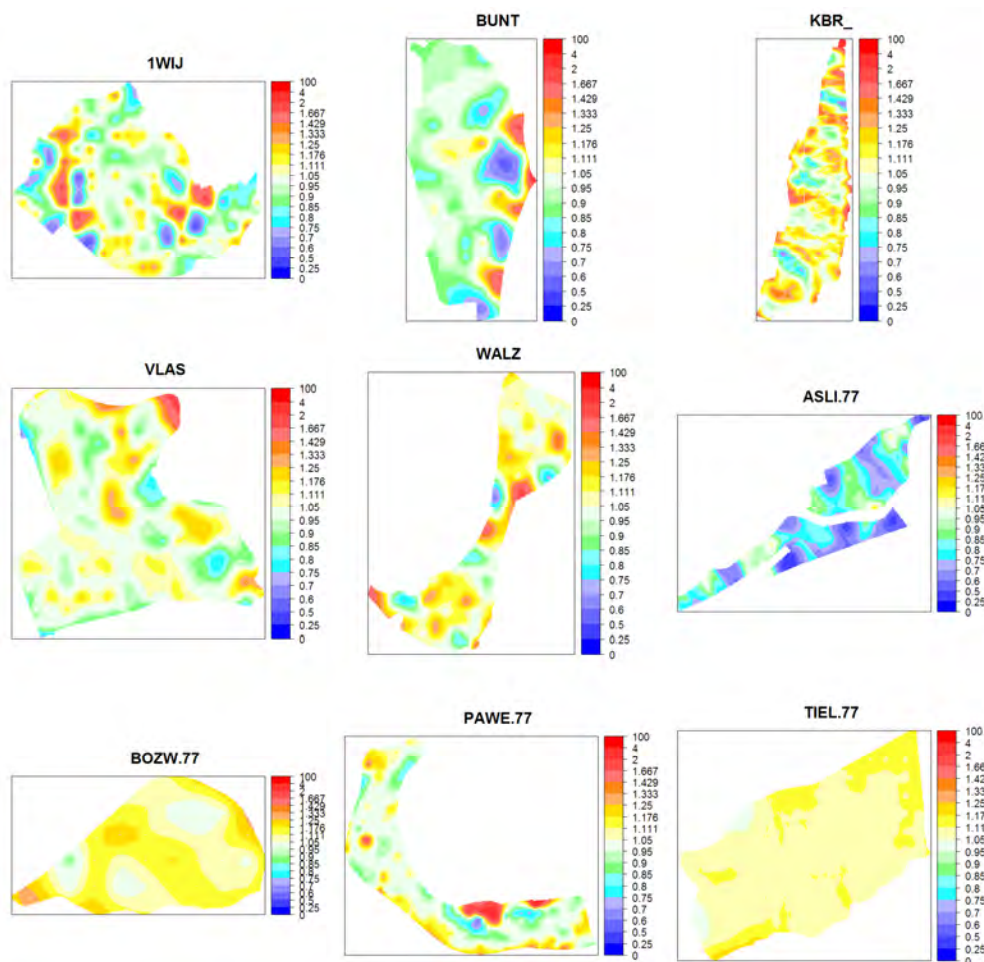
## 5.2 De impact op de predicties

We hebben een variogram gefit voor de volledige dataset. Dit is mogelijk door de cutoff waarde klein genoeg te zetten, waardoor de software het variogram enkel berekent op punten in hetzelfde gebied. Dit algemene variogram wordt dan gebruikt om predicties te doen voor ieder gebied apart.

Figuur 5-3 geeft aan hoe verschillend de predicties zijn. Zolang de verschillen geel of groen zijn, is het verschil beperkt tot ongeveer 25%. Wordt het rood of blauw dan zijn de verschillen groter. Voor bepaalde gebieden is de voorspelling vrij gelijklopend (VLAS, BOZW, TIEL), voor andere is de predictie helemaal verschillend (BUNT, ASLI, PAWE) en voor de

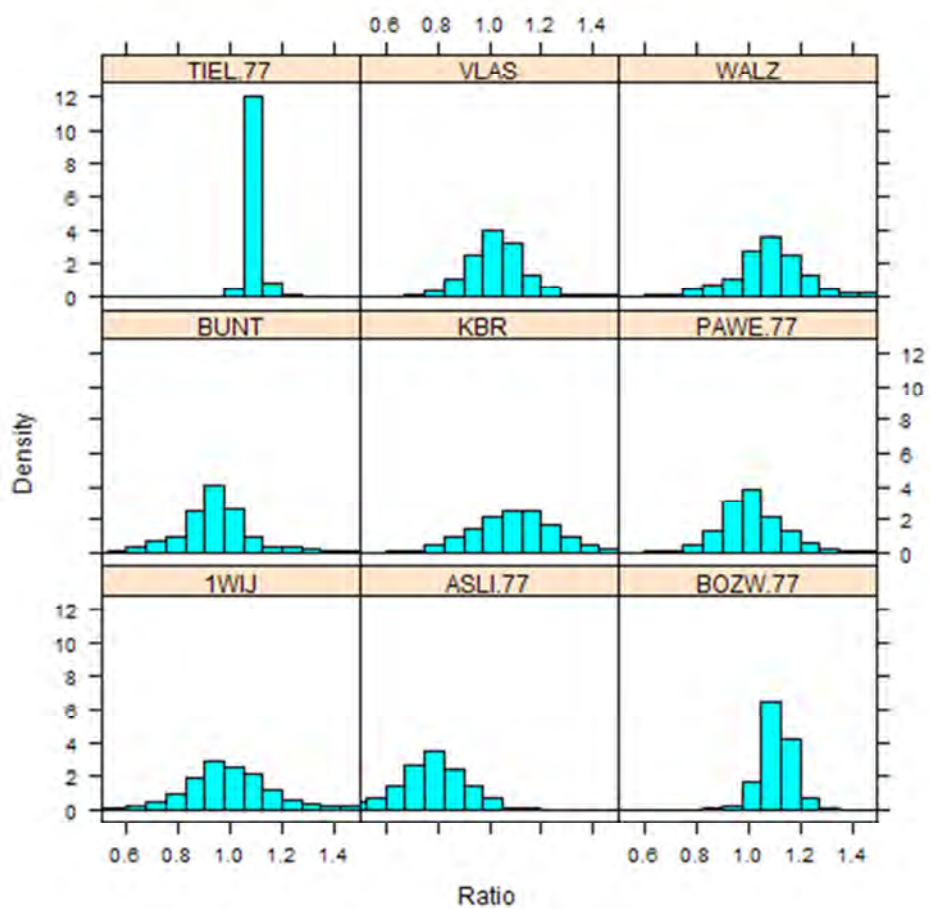
overige gebieden zijn er nog steeds vrij duidelijke verschillen. Voor BOZW geldt dan weer dat de gemiddelde schatting hoger is dan voorheen.

Niettemin kunnen we er van uitgaan dat het moeilijk zal worden om met 1 variogrammodel alle gebieden te bestrijken, omdat er al een groot verschil is tussen overstromde en niet-overstromde gebieden en de verschillen met de lokale predictie toch vaak zeer groot zijn.



Figuur 5-3 Verhouding van de voorspelling met het algemene model en deze met de lokale modellen uitgedrukt op de kaarten van de gebieden.





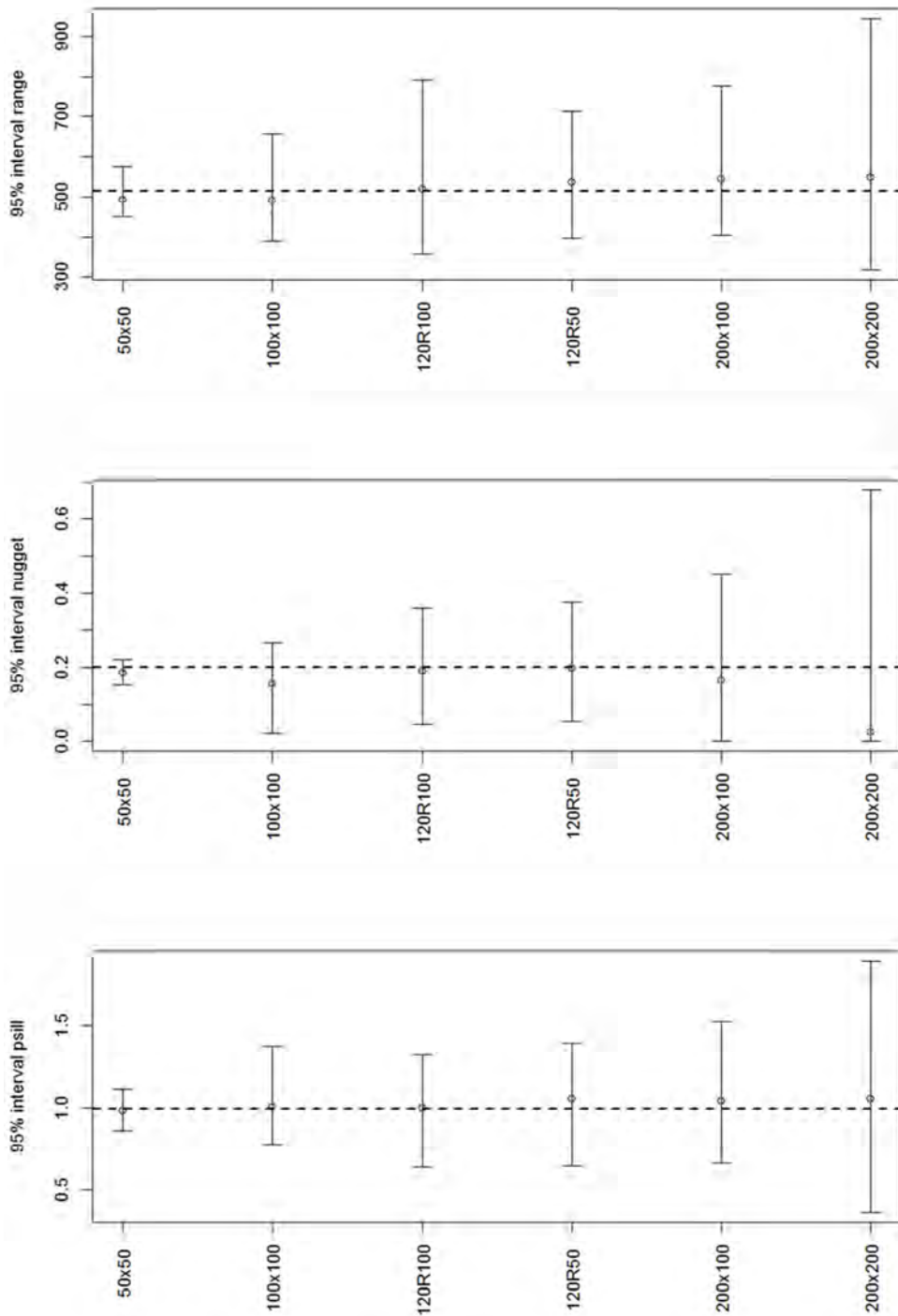
Figuur 5-4 Verhouding van de voorspelling met het algemene model en deze met de lokale modellen, uitgedrukt in een histogram.

## 6 Kostenramingen

### 6.1 Staalname strategie

Figuur 6-1 onderzoekt een aantal mogelijke configuraties om de bemonstering van een gebied uit te voeren, gebruik makend van een gesimuleerde dataset op basis van een sferisch model. Het grofste grid komt overeen met ongeveer 30 meetpunten, het 100x100m en de twee gedeeltelijk aselechte grids bevatten 120 punten, het 200x100m grid logischerwijs een 60-tal punten en het 50x50m grid een kleine 500 meetpunten. Het is duidelijk dat de grootte van de steekproef de grootste impact heeft op de nauwkeurigheid van de parameterbepalingen. Of er al dan niet een aantal punten willekeurig op een verschillende afstand dan het grid worden gekozen, maakt niet zoveel uit naar de parameterschatting. Het kan wel een veel beter inzicht geven hoe de correlatie evolueert op zeer korte afstand, wat zeker voor voorspellingen een grote impact kan hebben.

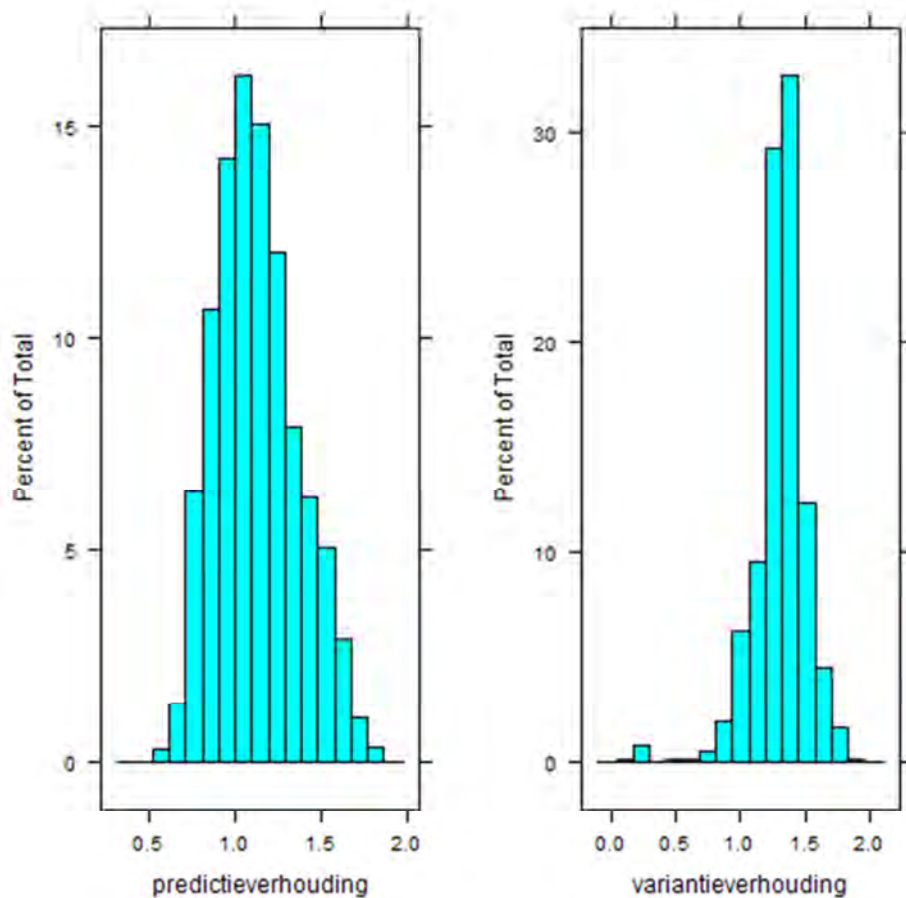
Verder viel op dat doordat het 200x200 grid uit slechts 30 punten bestaat, de variogrammen vaak zeer onstabiel zijn en zelfs niet gefit kunnen worden. Op Figuur 6-1 is aangegeven waarbinnen 95% van de predicties vallen. Maar zelfs voor een steekproef van een goeie 100 punten, kunnen de schattingen nog altijd zeer ver uit elkaar lopen.



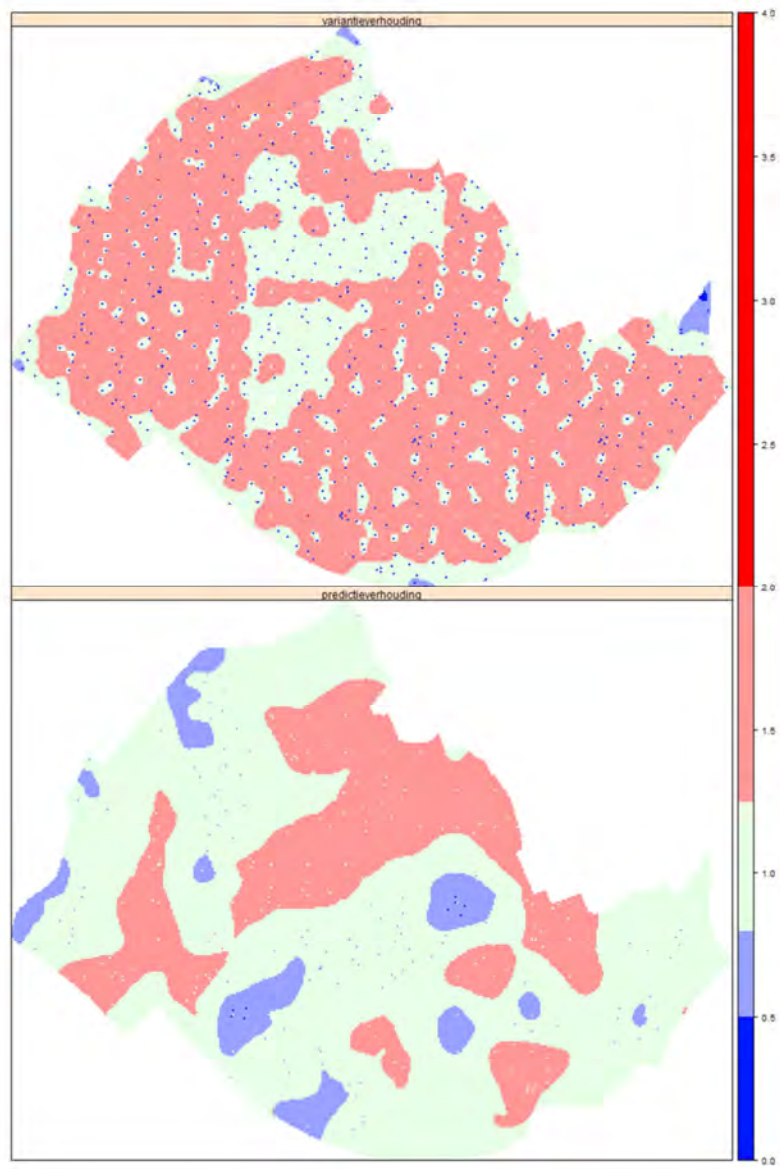
Figuur 6-1 Impact van de keuze van de steekproef op de range (boven), sill (midden) en nugget (onder).

## 6.2 Cokriging

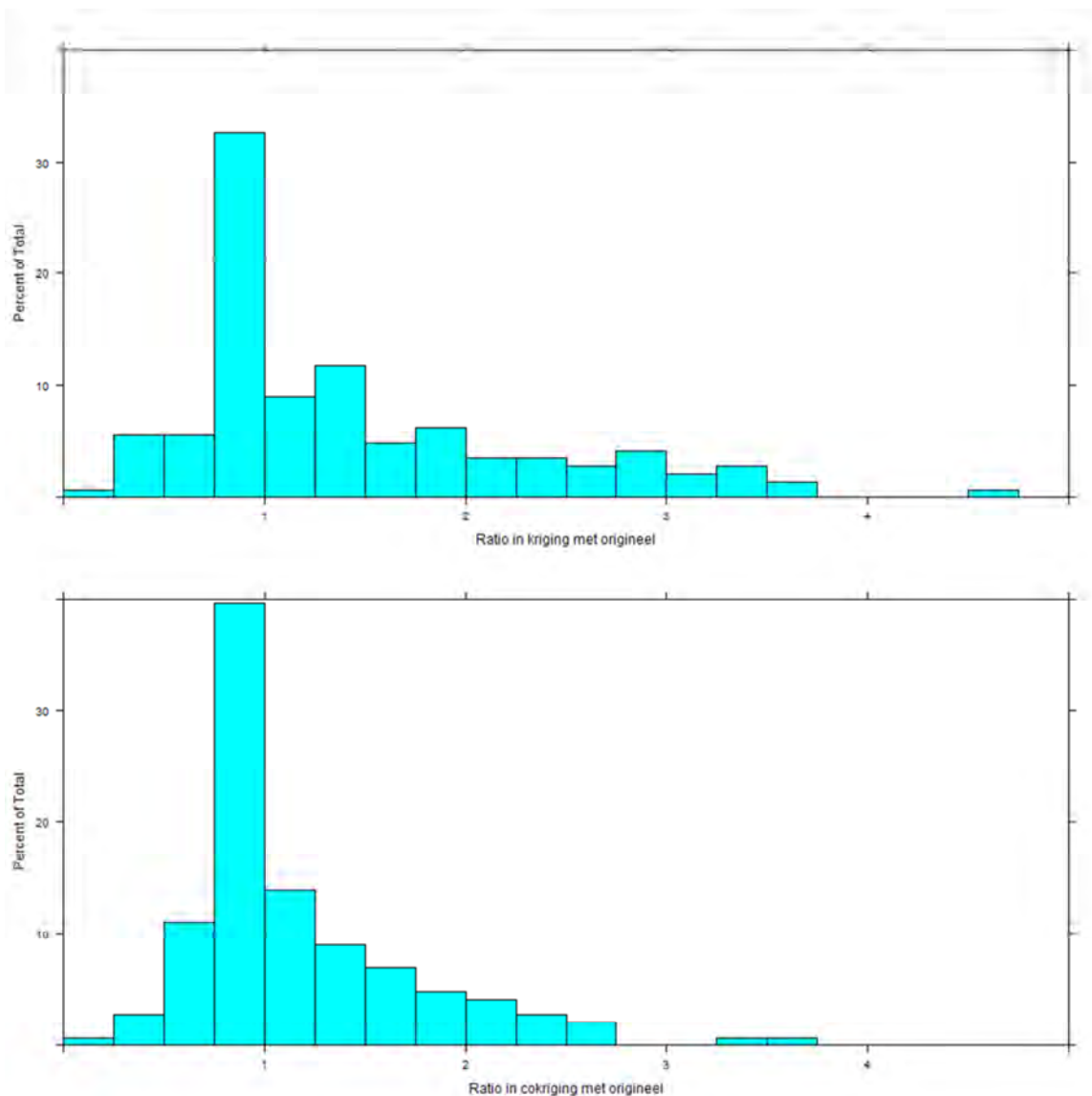
In het geval van deze dataset heeft cokriging wel duidelijk zijn nut indien we in bijvoorbeeld maar een kwart van de punten van het gebied 1WIJ de metaalconcentratie meten, en in alle meetpunten de kleiconcentratie kunnen meten. De kleiconcentratie is met meer dan 0.9 gecorreleerd met het zware metaal. Het grootste deel van de voorspellingen wijkt maximaal een factor 2 af tussen beide methodes (Figuur 6-2) en meestal is deze afwijking zeer beperkt (dus ratio dicht bij 1). Op Figuur 6-4 is duidelijk zichtbaar dat de gemeten waarde van zink een stuk dichter benaderd wordt dan op basis van de meetpunten van zink alleen.



Figuur 6-2 Spreiding van de verhouding van predictie (links) en variantie (rechts) bij normale kriging en via cokriging.



Figuur 6-3 Verhouding van kriging predictie en variantie op kaart voor het gebied Wijmeers I. Rood betekent dat de waarde van cokriging hoger ligt dan deze van de gewone kriging.



Figuur 6-4 Spreiding van de verhouding van predictie tussen de originele waarde enerzijds en de voorspelling van kriging op een beperkte set alsook de voorspelling met behulp van cokriging anderzijds.

### 6.3 Kostenraming

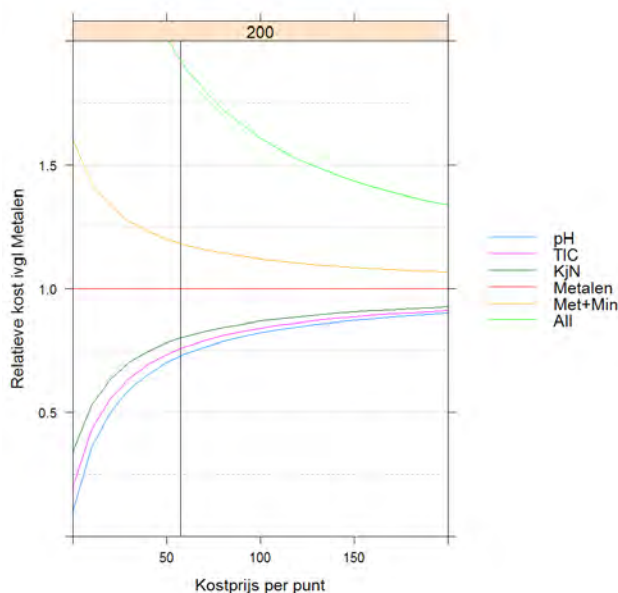
Tabel 6-1 geeft weer in welke mate er een besparing gerealiseerd wordt wanneer niet altijd de volledige standaardset van bodemvariabelen gemeten wordt. Wel dient het nut daarvan niet overschat te worden, want hier wordt vanuit gegaan dat alle opgemeten variabelen een perfecte proxy zijn voor de zware metalen. In de realiteit is dit echter niet het geval, aangezien de voorlopig gekende variabelen onvoldoende met de zware metalen gecorreleerd zijn. Dit is dus een puur theoretische oefening.

Tabel 6-1 Totale kostprijs per meetpunt in verschillende scenario's, waarbij de kostprijs van een staalname (zonder labo-analyse) op 57 euro is ingesteld.

Kost	Labokost	Totale kost	Relatieve kost
Alle laboanalyses	101	158	1.93
Metalen+Mineralen	40	97	1.18
Metalen	25	82	1
Mineralen	20	77	0.94
KjN	8.5	65.5	0.8
DS+LOI+TIC	5	62	0.76
pH+EC	2.5	59.5	0.73
Niets	0	57	0.7

Indien er geen staalnamekost is, dan is het puur de labo-analyse die de doorslag geeft, en is de relatieve kost gelijk aan de verhouding in labokosten. Hoe groter de staalnamekost, hoe minder profijt gedaan wordt met het opmeten van alternatieve goedkopere variabelen, behalve natuurlijk deze, waarvoor een stuk van de staalnamekost verdwijnt, omdat ze gewoon op het veld zonder bodemmonster gemeten kunnen worden.

Volgens de gegevens in de tabel blijkt dat de gemiddelde kost per staal ongeveer 57 euro bedraagt. In dit geval is de winst om pH te meten in plaats van de metalen, nog slechts 30%. Dus indien pH zwak gecorreleerd is met de metalen, wat hier ook het geval is, bespaar je 30 % kosten, maar heb je een zeer slechte proxy voor de metaalconcentratie. Echter alles voor ieder meetpunt meten, lijkt niet aangewezen, omdat - als de kostprijs voor staalname 57 euro bedraagt - dit ongeveer dubbel zo duur is dan in het geval dat we alleen de zware metalen meten, wat hier in deze studie onze enige interessevariabele is.



Figuur 6-5 Relatieve kostprijs ten opzichte van een meting van enkel zware metalen voor 200 meetpunten, uitgerekend voor 6 scenario's van wat er standaard gemeten wordt.

We kunnen besluiten dat kiezen voor een veel goedkopere variabele in labo-analysekost, niet altijd zoveel winst oplevert, en wel om volgende redenen:

- Een benaderende variabele is alleen nuttig als deze een sterke samenhang met de interessevariabele vertoont.
- Niet alleen de labokost is van belang, een relatief grote staalnamekost heeft tot gevolg dat er niet zoveel winst gedaan wordt met een goedkope labo-analyse.



## 7 Conclusies en aanbevelingen

Er zijn enkele alternatieve sporen uitgeprobeerd om tot een goede interpolatie van de zware metalen in overstromingsgebieden te komen. Het bleek dat de concentratie van zware metalen wel samenhangt met verschillende verklarende variabelen, maar deze samenhang is niet heel sterk. Dit heeft als gevolg dat het op basis van de huidige variabelenset niet mogelijk is om de meting van de zware metalen te vervangen door goedkopere alternatieven. Verder zijn alle zware metalen zo zwaar gecorreleerd dat de resultaten van een spatiale analyse op meerdere metalen niet nodig is, omdat de conclusies voor de andere zware metalen zeer gelijkaardig zullen zijn. Hier is gekozen voor zink als het te analyseren metaal.

Om de concentraties van de zware metalen in kaart te brengen, blijft de meest aangewezen methode gewoon de concentratie van het metaal zelf op te meten, en deze dan te interpoleren over het gebied. Aangezien de gebieden wel een duidelijk verschillende ruimtelijke samenhang hebben - zo zijn de zware metalen in reeds overstroomde gebieden, veel meer ad random verspreid dan in de niet-overstroomde gebieden - zal een apart interpolatiemodel per gebied moeten worden gebruikt. Dit heeft echter wel de implicatie dat kleinere gebieden veel intensiever bemonsterd moeten worden, omdat er zeer veel meetpunten nodig zijn om tot een goed interpolatiemodel te komen. Hoeveel punten precies bemonsterd moet worden, moet blijken uit een steekproefstudie, maar het is wel mogelijk enkele algemene aanbevelingen te doen.

Aanbevelingen voor de toekomst zijn:

- Zorgen dat er een referentietoestand opgemeten wordt alvorens het gebied kan overstromen.
- Het opmaken van een apart variogram voor ieder gebied. Dit heeft als gevolg dat de kleinere gebieden intensiever gesampled zullen moeten worden. Minstens 100 meetpunten per gebied zijn zeker geen overbodige luxe. Om een beter inzicht te krijgen in de correlatie-evolutie op zeer korte afstand, is het best een significant deel van de punten te reserveren voor meetpunten zeer dicht (< 50 m) bij elkaar. Dat kan met een bepaalde systematiek gebeuren, maar mag ook aselekt. Om een mooie ruimtelijke gebalanceerde steekproef te hebben bestaan er intussen enkele algoritmes gebaseerd op "tesselation" zoals het grts-algoritme (Stevens & Olsen, 2004).
- Als er beslist wordt om de gebieden in de tijd op te volgen is het van groot belang dat de meetpunten niet verplaatsen in de tijd. Op 100 waarnemingen zijn de variogrammen nog altijd niet echt stabiel, en dus een verplaatsing van het grid, zal leiden tot een verschillend variogram, waardoor ook de interpolaties niet meer te vertrouwen zijn.
- Omdat de overschrijdingsnormen gebaseerd zijn op formules die de pH, het kleigehalte en de hoeveelheid organisch materiaal bevatten, is het belangrijk dat deze variabelen systematisch mee opgemeten worden.

## Lijst van figuren

Figuur 2-1	bodemfysische en -chemische variabelen per sigmagebied. ....	16
Figuur 2-2	Pearson correlaties tussen zware metalen onderling (concentraties zijn log <sub>10</sub> getransformeerd) en tussen zware metalen, klei en organisch materiaal. De rode curve is een locally weighted regression (Lowess) lijn. ....	17
Figuur 2-3	Correlaties tussen gelogarithmeerde (Log <sub>10</sub> ) concentraties van zware metalen met aanduiding van de gebiedcode. Voor Pb werden enkel de concentraties kleiner dan 1000 ppm bewaard, 3 outliers werden verwijderd. ....	19
Figuur 2-4	De ligging van de overstromingsgebieden. AS1L, AS2L en LIER zijn gecombineerd in ASLI. ....	20
Figuur 2-5	Samenhang tussen zink en cadmium per gebied voor cadmiumwaarden boven de detectielimiet. De concentraties zijn in de log <sub>10</sub> -schaal uitgedrukt omdat dit beter correspondeert met de werkelijke samenhang. ....	21
Figuur 3-1	De drie voornaamste basisvormen voor variogrammen. De variantie tussen meetpunten wordt in functie van hun onderlinge afstand uitgedrukt. ....	25
Figuur 4-1	Verbanden tussen zink en verklarende variabelen voor het gebied 1WIJ. De punten geven de ruwe data aan en linksboven wordt de correlatiecoëfficiënt weergegeven. Verder wordt ook nog de lineaire trendlijn aangegeven als een donkergroen-zwarte lijn en wordt er een lowess-smoother (lokaal gewogen schattingen) als een rode lijn weergegeven. ....	31
Figuur 4-2	Predicties Zn en Cd in 1WIJ (mg/kg). ....	33
Figuur 4-3	Kans op overschrijding van de norm in 1WIJ. ....	34
Figuur 4-4	Verbanden met Zn voor het gebied BUNT. ....	35
Figuur 4-5	Predicties Zn en Cd in BUNT (mg/kg). ....	37
Figuur 4-6	Kans op overschrijding van de norm in BUNT. ....	38
Figuur 4-7	Verbanden met Zn voor het gebied KBR. ....	39
Figuur 4-8	Predicties Cd en Zn voor KBR (mg/kg). ....	41
Figuur 4-9	Kans op overschrijding van de norm in KBR. ....	42
Figuur 4-10	Verbanden met Zn voor het gebied VLAS. ....	43
Figuur 4-11	Predicties Zn en Cd in VLAS (mg/kg). ....	44
Figuur 4-12	Verbanden met Zn voor het gebied WALZ. ....	45
Figuur 4-13	Predicties Zn en Cd in WALZ (mg/kg). ....	46
Figuur 4-14	Verbanden met Zn voor het gebied ASLI. ....	47
Figuur 4-15	Predicties Zn en Cd in ASLI (mg/kg). ....	49
Figuur 4-16	Topokaart van Anderstadt I en II en Polder van Lier. ....	49
Figuur 4-17	Kans op overschrijding van de norm in ASLI. ....	50
Figuur 4-18	Verbanden met Zn voor het gebied BOZW. ....	51
Figuur 4-19	Predicties Zn en Cd in BOZW (mg/kg). ....	52
Figuur 4-20	Kans op overschrijding van de norm in BOZW. ....	53

Figuur 4-21 Verbanden met Zn voor het gebied PAWE. De foutmeldingen komen omdat deze variabelen niet opgemeten zijn voor dit gebied.....	54
Figuur 4-22 Predicties Zn en Cd in PAWE (mg/kg).....	55
Figuur 4-23 Kans op overschrijding van de norm in PAWE.....	56
Figuur 4-24 Verbanden met Zn voor het gebied TIEL.....	57
Figuur 4-25 Predicties Zn en Cd in TIEL (mg/kg).....	58
Figuur 4-26 Kans op overschrijding van de norm in TIEL.....	59
Figuur 4-27 Vergelijking van de semivariogrammen over de verschillende gebieden. Bovenaan staan de semivariogrammen voor logZink van alle gebieden en ook van het algemene model van het volgend hoofdstuk. Onderaan staan deze voor cadmium. In de rechter figuur worden semivariogrammen genormeerd.....	61
Figuur 4-28 Alle predicties en krige-varianties voor zink onder 1 noemer voor de niet-overstroomde gebieden.....	63
Figuur 4-29 Alle predicties en krige-varianties voor zink onder 1 noemer voor de overstroomde gebieden.....	64
Figuur 4-30 Relatief verschil in predictie (links) en variantie (rechts) uitgedrukt als een verhouding tussen de gewone krigingvoorspelling en de cokrigingvoorspelling van zink. De verhouding van de varianties is de verhouding in de log-schaal. ....	65
Figuur 4-31 Indicator Kriging (onder) en een eigen vorm van stochastische voorspelling(boven) voor de overschrijdingskans van de zinknorm voor het gebied PAWE. ....	67
Figuur 4-32 Indicator Kriging (onder) en een eigen vorm van stochastische voorspelling(boven) voor de overschrijdingskans van de zinknorm voor het gebied TIEL. ....	68
Figuur 5-1 Verbanden algemene gebiedsoverschrijdende analyse.....	70
Figuur 5-2 Predicties voor de algemene analyse, telkens voor 1 variabele die varieert terwijl de rest op het gemiddelde (0) is gefixeerd en er rekening is gehouden met de range van de variabelen om onderlinge vergelijkingen mogelijk te maken. LOI covarieert met klei wegens hun correlatie. De blauwe lijn is voor overstroomde gebieden, de magenta lijn voor niet overstroomde gebieden.....	71
Figuur 5-3 Verhouding van de voorspelling met het algemene model en deze met de lokale modellen uitgedrukt op de kaarten van de gebieden. ....	72
Figuur 5-4 Verhouding van de voorspelling met het algemene model en deze met de lokale modellen, uitgedrukt in een histogram. ....	73
Figuur 6-1 Impact van de keuze van de steekproef op de range (boven), sill (midden) en nugget (onder).....	75
Figuur 6-2 Spreiding van de verhouding van predictie (links) en variantie (rechts) bij normale kriging en via cokriging. ....	76
Figuur 6-3 Verhouding van kriging predictie en variantie op kaart voor het gebied Wijmeers I. Rood betekent dat de waarde van cokriging hoger ligt dan deze van de gewone kriging. ....	77
Figuur 6-4 Spreiding van de verhouding van predictie tussen de originele waarde enerzijds en de voorspelling van kriging op een beperkte set alsook de voorspelling met behulp van cokriging anderzijds. ....	78

Figuur 6-5 Relatieve kostprijs ten opzichte van een meting van enkel zware metalen voor 200 meetpunten, uitgerekend voor 6 scenario's van wat er standaard gemeten wordt..... 79

## Lijst van tabellen

Tabel 2-1. Variabelen die in het labo werden opgemeten op de bodemstalen van de FloodGis databank. De gebruikte eenheid wordt getoond, evenals de bepaalbaarheidsgrens (BG) die voorkomt in de databank en een courante bepaalbaarheidsgrens (BG) waarvan de helft van deze waarde werd gebruikt bij de verwerking van de data (BG/2).....	13
Tabel 2-2. Lijst van de onderzochte sigmagebieden met hun oppervlakte en aantal staalnamepunten. Tevens wordt weergegeven of het een gebied uit het Sigmaplan (1977) betreft ("1977") of uit het geactualiseerd Sigmaplan (Beslissing Vlaamse Regering 2005 en 2006) ("2010" slaat op het voorziene startjaar voor de aanleg van deze gebieden) en of het reeds in gebruik is als overstromingsgebied.....	14
Tabel 4-1. Modelschattingen voor het gebied 1WIJ. N is het aantal waarnemingen. Per variabele wordt de geschatte modelcoëfficiënt weergegeven, samen met het bijhorende 95%-betrouwbaarheidsinterval. De significantie wordt aangegeven met . : $0.05 \leq p < 0.10$ ; * : $0.01 \leq p < 0.05$ ; ** : $0.001 \leq p < 0.01$ ; *** : $p < 0.0001$ . De variabelen die niet opgenomen zijn in het model worden met schatting X aangegeven en de reden waarom ze niet opgenomen zijn is te vinden waar anders het betrouwbaarheidsinterval staat. Indien ze oorspronkelijk wel in het model zijn opgenomen, maar helemaal niet significant leken, worden ze met schatting 0 aangegeven.....	32
Tabel 4-2 Modelcoëfficiënten voor het gebied BUNT.....	36
Tabel 4-3 Modelcoëfficiënten voor het gebied KBR.....	40
Tabel 4-4 Modelcoëfficiënten voor het gebied VLAS. ....	43
Tabel 4-5 Modelcoëfficiënten voor het gebied WALZ. ....	45
Tabel 4-6 Modelcoëfficiënten voor het gebied ASLI.....	48
Tabel 4-7 Modelcoëfficiënten voor het gebied BOZW.....	51
Tabel 4-8 Modelcoëfficiënten voor het gebied PAWE. ....	54
Tabel 4-9 Modelcoëfficiënten voor het gebied TIEL. ....	57
Tabel 4-10 Vergelijking variogrammen tussen de gebieden voor zink. Het type model, range (maximale afstand waarop punten gecorreleerd zijn), nugget (ruisvariantie), var inf (variantie tussen punten ver van elkaar), Cor 0m (1 - signaal/ruisverhouding) worden weergegeven.....	62
Tabel 4-11 Vergelijking variogrammen tussen de gebieden voor cadmium. Het type model, range (maximale afstand waarop punten gecorreleerd zijn), nugget (ruisvariantie), var inf (variantie tussen punten ver van elkaar), Cor 0m (1 - signaal/ruisverhouding) worden weergegeven.....	62
Tabel 5-1 Coëfficiënten van het algemene model. ....	69
Tabel 6-1 Totale kostprijs per meetpunt in verschillende scenario's, waarbij de kostprijs van een staalname (zonder labo-analyse) op 57 euro is ingesteld.....	79

## Literatuurlijst

Pinheiro, J.C., and Bates, D.M. (2000) *Mixed-Effects Models in S and S-PLUS*, Springer, esp. pp. 100, 461.

Venables, W.N. and Ripley, B.D. (1997) *Modern Applied Statistics with S-PLUS*, 2nd Edition, Springer-Verlag.

Stevens, D.L., Jr., and A.R. Olsen. (2004) Spatially-balanced sampling of natural resources. *Journal of the American Statistical Association* 99, 262-278.

