

Hoe de kartering van de grondwaterstand (nog) beter kan

Hoe de kartering van de grondwaterstand (nog) beter kan

T. Hoogland
D.J. Brus
G.B.M. Heuvelink
M. Knotters

Alterra-rapport 1032

Alterra, Wageningen, 2004

REFERAAT

Hoogland, T., D.J. Brus, G.B.M. Heuvelink en M. Knotters, 2004. *Hoe de kartering van de grondwaterstand (nog) beter kan*. Wageningen, Alterra, Alterra-rapport 1032. 88 blz. . 19 fig; 24 tab.; 26 ref.

De afgelopen jaren ontwikkelde Alterra een methode voor kartering van grondwaterstands-karakteristieken (de Gd-karteringsmethode). Deze methode, die met behulp van tijdreeksanalyse, lineaire regressie en geostatistische interpolatie ruimtelijke beelden van grondwaterstands-karakteristieken produceert, is de ontwikkelfase inmiddels ontstegen en al een aantal keren in de praktijk toegepast. Op onderdelen is aanpassing, verfijning, correctie en verbetering van de methode mogelijk. Daarom is in dit project de Gd-methode grondig onder de loep genomen en is een uitgebreide lijst gemaakt van mogelijke verbeterpunten. Op basis hiervan zijn belangrijke en kansrijke verbeterpunten onderzocht in drie deelonderzoeken. Het eerste deelonderzoek, naar tijdreeksmodellering van grondwaterstandsreeksen voor de predictie van klimaatsrepresentatieve GxG (GHG, GVG en GLG), gaf geen aanleiding de Gd-methode aan te passen. Op het gebied van zowel stambuisregressie voor de omzetting van gemeten grondwaterstanden naar klimaats-representatieve GxG als van vlakdekkende predictie van klimaatsrepresentatieve GxG zijn echter wel mogelijke verbeteringen gevonden. De voorstellen betreffen vooral kleine veranderingen die relatief eenvoudig te implementeren zijn en die de uitgangspunten en globale werkwijze van de Gd-methode intact laten.

Trefwoorden: Grondwaterstandsdynamiek, Gd, Grondwatertrap, Gt, steekproef, waarnemings-dichtheid, stambuis, regressie, gebiedsdekkende predictie, kriging

ISSN 1566-7197

Dit rapport kunt u bestellen door €20,- over te maken op banknummer 36 70 54 612 ten name van Alterra, Wageningen, onder vermelding van Alterra-rapport 1032. Dit bedrag is inclusief BTW en verzendkosten.

© 2004 Alterra
Postbus 47; 6700 AA Wageningen; Nederland
Tel.: (0317) 474700; fax: (0317) 419000; e-mail: info.alterra@wur.nl

Niets uit deze uitgave mag worden veelevoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van Alterra.

Alterra aanvaardt geen aansprakelijkheid voor eventuele schade voortvloeiend uit het gebruik van de resultaten van dit onderzoek of de toepassing van de adviezen.

Inhoud

Woord vooraf	7
Samenvatting	9
1 Inleiding	15
1.1 Probleemstelling en achtergrond	15
1.2 Projectdoelstelling	16
1.3 Leeswijzer	16
2 Werkwijze en methoden bij kartering van de grondwaterstandsdynamiek	17
2.1 Inleiding	17
2.2 Indeling van het gebied in homogene deelgebieden	18
2.3 De GxG op peilbuislocaties	19
2.3.1 Selectie van grondwaterstandsreeksen	19
2.3.2 Overige grondwaterstandsreeksen	20
2.3.3 Berekening van de klimaatrepresentatieve GxG	20
2.4 De GxG op gerichte opnamelocaties	22
2.4.1 Berekening van de GxG voor gerichte-opnamelocaties	22
2.5 Gebiedsdekkende hulpinformatie over topografie en waterhuishouding	23
2.5.1 Afleiding van gebiedsdekkende hulpinformatie	24
2.6 Gebiedsdekkende predictie van de GxG	28
2.6.1 Regressieanalyse	28
2.6.2 Ruimtelijke structuur van de residuen	29
2.6.3 Universal kriging van GxG	29
3 Mogelijke verbeterpunten in de Gd-methode	31
3.1 Gebiedsindeling	31
3.2 Gegevensverzameling	31
3.3 Tijdreeksmodellering	32
3.4 GxG-voorspellingen voor gerichte opnames	32
3.5 Afleiden hulpgegevens	33
3.6 Gebiedsdekkende predictie van GxG	33
3.7 Overig	34
3.8 Nader te onderzoeken mogelijkheden voor verbetering	34
4 Vergelijking van een lineair en niet-lineair tijdreeksmodel voor berekening van de GxG op stambuislocaties	37
4.1 Inleiding	37
4.2 Beschrijving van de stambuizenet en het onderzoeksgebied	38
4.3 Beknopte beschrijving lineaire en niet-lineaire tijdreeksmodel	40
4.3.1 Het lineaire model	41
4.3.2 Het niet-lineaire model	41
4.4 Kalibratie van de tijdreeksmodellen op de stambuizenet	43
4.4.1 Kalibratie van het lineaire model	43

4.4.2	Kalibratie van het niet-lineaire model	44
4.5	Vergelijking van modelvoorspellingen van de GxG	47
4.6	Vergelijking van modelvoorspellingen met de werkelijkheid	50
4.7	Invloed van modelkeuze op resultaten van de stambuisregressie	53
4.8	Discussie	57
4.9	Conclusies	59
5	Effect van het meettijdstip en de locatie van stambuizen op de kwaliteit van met stambuisregressie voorspelde GxG	61
5.1	Inleiding	61
5.2	Gebruikte data	61
5.3	Onderzoeksopzet	61
5.3.1	Meettijdstippen	61
5.3.2	Steekproefopzetten	61
5.3.3	Fitten van het regressiemodel	63
5.4	Validatie	64
5.4.1	Leave-one-out	64
5.4.2	Kwaliteitsmaten	64
5.5	Resultaten	66
5.6	Discussie	71
5.7	Conclusies	71
6	Het effect van gerichte opnamelocaties op de nauwkeurigheid van gebiedsdekkende GxG-voorspellingen	73
6.1	Inleiding	73
6.2	Studiegebied en dataverzameling	73
6.3	Opzet van het experiment	74
6.3.1	Steekproefopzet	74
6.3.2	Toepassing van de Gd-methode bij verschillende steekproefopzetten	75
6.3.3	Validatie	77
6.4	Resultaten	78
6.4.1	Systematische fout	78
6.4.2	Toevallige fout	81
6.4.3	Modellschattingen van de voorspelnauwkeurigheid	81
6.4.4	Correctie voor de meetfout in GxG	82
6.4.5	Proportionele of preferente allocatie per Gt-klasse	83
6.5	Discussie	83
6.6	Conclusies	84
7	Conclusies en aanbevelingen	85
	Literatuur	87

Woord vooraf

Dit project is gefinancierd vanuit het DWK-programma 395, basis- en kerngegevens bovengrond, waaruit ontwikkeling, onderhoud en verbetering van de kartering van grondwaterstandsdynamiek (Gd) grotendeels gefinancierd worden.

Om een Gd-kartering te laten aansluiten bij de beoogde toepassing en de nauwkeurigheid te realiseren die daarvoor vereist is, is het van belang inzicht te krijgen in de belangrijkste oorzaken van fouten in de Gd-kaarten de grootte van deze fouten bij bepaalde keuzes aangaande de dataverzameling. Om dit inzicht te verkrijgen is in 2002 door de toenmalige projectleider Peter Finke dit project gestart. Het veldwerk voor dit project heeft plaatsgevonden in de zomer van 2002. De overige data zijn afkomstig uit lopende Gd-projecten, uitgevoerd in het kader van de aanwijzing van uitspoelingsgevoelige gronden. De inschatting van de relatieve bijdrage van verschillende foutenbronnen helpt bij de prioriteitstelling van mogelijke verbeteringen. Deze mogelijkheden voor verbetering betreffen zowel de kwaliteit van de gebruikte gegevens als de gebruikte methoden voor de verwerking van de gegevens.

We danken het waterschap Regge en Dinkel voor het ter beschikking stellen van digitale gegevens over het watersysteem en de gehanteerde streefpeilen. Ook willen we Jaco van der Gaast en Jaap de Gruijter danken voor het meedenken over de keuze van relevante en haalbare mogelijkheden voor verbetering die in dit onderzoek zijn beschouwd.

Samenvatting

Bij Alterra is een methode ontwikkeld om de karakteristieken van de grondwaterstandsdynamiek (Gd) gebiedsdekkend te voorspellen en de nauwkeurigheid van deze voorspellingen te kwantificeren. Deze karakteristieken zijn onder meer de gemiddeld hoogste en laagste grondwaterstand (GHG en GLG, samengevat als GxG). Om de Gd-methode toe te kunnen passen voor doeleinden die hoge eisen stellen aan de nauwkeurigheid van de voorspellingen moet onderzocht worden hoe de nauwkeurigheid van de Gd-kartering te verhogen is. Het **doel** van deze studie is om de belangrijkste en meest belovende mogelijkheden voor verbetering van de Gd-methode op te sporen, en aan de hand van cases nader te onderzoeken.

De Gd-methode start met de indeling in hydrologisch homogene deelgebieden (stratificatie). Binnen deze strata worden de locaties van de gerichte opnames geloot. Voor de locaties waar tijdreeksen van grondwaterstanden zijn waargenomen kan de GxG worden geschat met behulp van een tijdreeksmodel. Met behulp van deze schattingen, gerichte opnames en regressieanalyse (stambuisregressie) wordt de GxG geschat voor de locaties van de gerichte opnames. Vervolgens wordt de GxG gebiedsdekkend geïnterpoleerd, waarbij gebruik wordt gemaakt van zowel de indeling in strata als van regressiemodellen die de samenhang beschrijven tussen de GxG en de diverse bronnen van gebiedsdekkende hulpinformatie over topografie en waterhuishouding.

De Gd-methode zou op de volgende onderdelen kunnen worden verbeterd:

- stratificatie: minimalisatie van voorspelfouten van GxG's binnen de strata;
- gerichte opname: meer controles op meetfouten, betere ruimtelijke spreiding van gerichte opnamelocaties, beter rekening te houden met de onnauwkeurigheid van gecensureerde of 'dieper dan'-waarnemingen;
- tijdreeksmodellering: baseren op meer representatieve buizen, onzekerheid van tijdreeksmodelparameters verdisconteren, invoerreeksen beter screenen, niet-lineaire tijdreeksmodellen toepassen, beter verifiëren of aan modelveronderstellingen zoals stationariteit en normaliteit is voldaan;
- stambuisregressie: een ander, meer fysisch gebaseerd criterium voor modelselectie hanteren, niet-lineaire regressiemodellen gebruiken, het tijdstip van de gerichte opname optimaliseren, de selectie van stambuizen verbeteren, onnauwkeurigheid van GxG-schattingen voor stambuislocaties verdisconteren, en bij extrapolatie meer onnauwkeurigheid in rekening brengen;
- geografische hulpinformatie: de kwaliteit van predictoren zoals de hoogtecijfers uit het Actueel Hoogtebestand Nederland (AHN) verhogen, meer hydrologische kennis gebruiken, en nieuwe digitale informatie die recent beschikbaar is gekomen aanwenden;
- regressieanalyse GxG-hulpinformatie: betere selectie van predictorvariabelen, variabelen voorafgaand aan regressieanalyse transformeren, gebruik maken van *generalised linear models*, rekening houden met correlatie van residuen tussen strata, het net van gerichte opnames verdichten;

- overig: rekening houden met positiefouten, een hogere resolutie dan de huidige (25×25 m) gebruiken, meer fysieke proceskennis benutten en Gd-kaarten valideren, d.w.z. objectief toetsen aan onafhankelijke waarnemingen.

Uit de bovenstaande mogelijkheden voor verbetering werden de meest belovende geselecteerd, wat tot de volgende drie deelonderzoeken leidde:

1. onderzoek naar de gevolgen van de veronderstelling van lineariteit in de respons van grondwaterstand op neerslagoverschot;
2. onderzoek naar de invloed van het meettijdstip van de gerichte opname en de omvang en samenstelling van de stambuizenset op de nauwkeurigheid van de GxG die bij stambuisregressie wordt voorspeld;
3. onderzoek naar het effect van dichtheid en keuze van locaties van de gerichte opname op de nauwkeurigheid van GxG-voorspellingen en nauwkeurigheid van gekwantificeerde onzekerheid.

Het **eerste deelonderzoek** is uitgevoerd met tijdreeksen van grondwaterstanden voor 19 stambuizen nabij Nijverdal. De grondwaterstand is halfmaandelijks gemeten gedurende ca. acht jaar. Op elke reeks is zowel een lineair transfer-ruismodel als een niet-lineaire stochastische differentiaalvergelijking gekalibreerd. Met de gekalibreerde modellen en meteorologische invoer zijn vervolgens grondwaterstandsreeksen van 30 jaar gesimuleerd, waaruit de GxG is berekend.

Het niet-lineaire model blijkt moeilijker te kalibreren dan het lineaire model, mogelijk doordat de parameters drainageweerstand en verticale flux elkaars invloed grotendeels kunnen compenseren. Bovendien bleek informatie over bodemfysica en hydrologie deels onvolledig en onbetrouwbaar te zijn. Het lineaire model bleek tot realistischer voorspellingen van de GxG te leiden dan het niet-lineaire model. De GxG die is geschat met het lineaire model bleek beter bij de waarnemingen van de gerichte opname te passen dan de GxG die met het niet-lineaire model is geschat.

Bij het **tweede deelonderzoek** is gebruik gemaakt van grondwaterstandsreeksen van drie tot tien jaar lang, die zijn waargenomen in 102 stambuizen in het waterschap Regge en Dinkel. Met een lineair tijdreeksmodel zijn de GxG's voor deze locaties berekend. Deze zijn beschouwd als de werkelijke GxG's. De analyse van het effect van het meettijdstip op de kwaliteit van de GxG-voorspellingen met het stambuisregressiemodel is gebaseerd op twaalf meettijdstippen in 1997: zes in het voorjaar en zes in het najaar. Omdat we GxG's van 102 stambuislocaties hebben terwijl in de praktijk de GxG's van ca 20 stambuizen worden gebruikt om het stambuisregressiemodel te kalibreren, kunnen we het effect van de ligging van de stambuizen in de geografische ruimte en in de eigenschappenruimte te onderzoeken. Hiervoor zijn een groot aantal keren 20 stambuizen geselecteerd uit de 102 beschikbare buizen. Deze selecties zijn op verschillende manieren uitgevoerd. Op het hoogste niveau kan onderscheid gemaakt worden tussen selectie door middel van loting (kanssteekproefname), en gerichte selectie. Kanssteekproeven zijn in dit onderzoek volgens twee verschillende opzetten uitgevoerd, te weten enkelvoudig aselect zonder teruglegging (SI), en gestratificeerd, enkelvoudig aselect zonder teruglegging (STSI). De volgende varianten van stratificatie zijn onderzocht:

1. de grondwatertrap volgens de Bodemkaart van Nederland 1 : 50 000 (STSI(Gt));
2. de gemiddelde grondwaterstand in stambuizen (STSI(h)), en
3. de geografische ligging (STSI(x)).

Het aantal strata is bij alle varianten vier. Het aantal gelote steekproefbuizen per stratum is òf evenredig met het totaal aantal buizen in het stratum (proportionele allocatie) òf gelijk voor alle strata (vijf in elk, niet-proportionele allocatie). Om goed zicht te krijgen op het effect van het type allocatie, is de loting 100 keer herhaald. Volgens de theorie van *experimental design* is bij lineaire regressie met 20 waarnemingen gerichte selectie van de tien buizen met een minimale waarde voor de predictor en de tien buizen met een maximale waarde van de predictor optimaal. Bij stambuisregressie is de grondwaterstand tijdens de gerichte opname de predictor. Bij de selectie van de stambuizen is nog niet bekend wanneer de gerichte opnames zullen worden verricht. Daarom is de gemiddelde grondwaterstand als selectiecriteria gebruikt, resulterend in tien stambuizen met de kleinste gemiddelde grondwaterstand en de tien met de grootste (Gericht (h)).

Voor alle steekproeven van 20 stambuizen zijn de modellen gefit met de *Ordinary Least Squares*-methode. Bij de kanssteekproeven zijn de gewichten die bij het steekproef-*design* behoren niet gebruikt bij het fitten van het model.

De kwaliteit van de GxG-voorspellingen is bepaald door middel van kruisvalidatie met *leave-one-out*: elk van de 102 stambuizen wordt eenmaal apart gebruikt voor validatie en 101 maal gebruikt bij de voorspelling. De validatieresultaten laten zien dat het tijdstip van grondwaterstandsmeting een veel grotere invloed heeft op de kwaliteit van de voorspellingen van de stambuisregressie dan de steekproefopzet voor selectie van stambuizen. Het gemiddelde absolute of gekwadrateerde verschil tussen de GHG en de grondwaterstand op een bepaalde dag blijkt geen geschikt criterium te zijn voor het kiezen van een geschikte dag voor gerichte opname van de grondwaterstand. Daarom dient voor de GHG het model met de kleinste restvariantie te worden geselecteerd. Voor de GLG valt de dag met de kleinste restvariantie wel samen met de dag met het kleinste gemiddelde absolute of gekwadrateerde verschil tussen GLG en grondwaterstand op een bepaalde dag.

Kanssteekproeven presteren iets beter dan gerichte selectie, omdat bij de laatste systematische fouten in de voorspelde GxG optreden. Stratificatie op basis van de gemiddelde grondwaterstand is nauwelijks beter dan stratificatie op basis van de Gt.

De steekproefopzetten verschillen ten aanzien van de nauwkeurigheid waarmee de Gd-methode de onzekerheid over de werkelijke GxG kwantificeert. Dit is gevalideerd door de berekende modelvariantie te vergelijken met de gemiddelde gekwadrateerde voorspelfout. Bij SI en STSI met proportionele allocatie is de gemiddelde modelvariantie een goede maat voor de gemiddelde gekwadrateerde fout, maar bij Gericht(h), en in mindere mate voor STSI met niet-proportionele allocatie wordt de gemiddelde gekwadrateerde fout onderschat. Dit komt mogelijk door heteroscedasticiteit (niet-constante variantie): bij Gericht(h) en STSI met niet-evenredige allocatie zijn de insluitkansen van de stambuizen ongelijk, waardoor in geval van heteroscedasticiteit de schatter van de restvariantie onzuiver wordt. Door te zorgen voor gelijke insluitkansen wordt de methode minder gevoelig voor de veronderstelling van homoscedasticiteit (constante variantie). Voor alle steekproefopzetten is de onzekerheid over de modelvariantie erg groot: de berekende modelvariantie verschilt sterk tussen de 100 steekproeven.

In het **derde deelonderzoek** is het effect van de dichtheid en de ruimtelijke verdeling van de locaties met gerichte grondwaterstandsmetingen op de kwaliteit van ruimtelijke voorspellingen van de GLG onderzocht. Hierbij zijn acht steekproefopzetten geanalyseerd, met verschillende waarnemingsdichtheden en verdelingen van meetlocaties over Gt-klassen zoals weergegeven op de Bodemkaart van Nederland 1 : 50 000. Het effect van de steekproefopzet op de voorspelde GLG is vastgesteld door middel van een validatiestudie in een gebied van 2 000 ha nabij Markelo in Overijssel.

Op 530 locaties is tussen juni en september 2002 eenmalig de grondwaterstand ten opzichte van maaiveld gemeten (gerichte opnames). De locaties zijn geselecteerd met een gestratificeerde enkelvoudig aselechte steekproef, met als strata de Gt-klassen zoals weergegeven op de Bodemkaart van Nederland 1 : 50 000. De GLG voor deze meetlocaties is geschat met stambuisregressie. Uit de 530 meetlocaties zijn, gestratificeerd enkelvoudig aselekt 79 meetlocaties geselecteerd als validatieset, eveneens met de Gt-klassen als strata. Het aantal validatielocaties is proportioneel aan het oppervlak van het stratum. Uit de resterende 451 meetlocaties zijn gestratificeerd enkelvoudig aselekt 20, 40, 80 en 160 locaties als kalibratieset geselecteerd. Dit komt overeen met een meetnetdichtheid van respectievelijk één, twee, vier en acht waarnemingslocaties per km². De allocatie is ofwel proportioneel met het oppervlak van de Gt-klassen, ofwel niet-proportioneel met een oververtegenwoordiging van locaties in natte Gt-klassen. Dit laatste komt overeen met de huidige Gd-karteringen, waarbij meer belang wordt gehecht aan betrouwbare uitspraken in natte Gt's dan in droge. In totaal zijn dus acht steekproefopzetten onderzocht. Met elke steekproefopzet zijn tien steekproeven gegenereerd. De Gd-methode is 80 keer uitgevoerd, waarbij de regressie tussen GxG en hulpinformatie op elk van de 80 steekproeven is gebaseerd. De kwaliteit van de gebiedsdekkende GLG-voorspellingen is gekwantificeerd met 1) de systematische fout in de voorspelde GLG, 2) de toevallige fout in de voorspelde GLG, en 3) de fout in de schatting van de variantie van de voorspelfout in de GLG.

Bij alle steekproefopzetten blijken de GLG-voorspellingen gemiddeld te droog te zijn. Bij lage waarnemingsdichtheden blijkt vaak een significante ($\alpha=0.05$) systematische fout op te treden, die kan oplopen tot 19 cm. Bij één waarneming per km² treedt een significante systematische fout vier tot vijf van de tien keer op, terwijl dit bij acht waarnemingen per km² slechts tussen nul en twee van de tien keer het geval is. Ook de spreiding in systematische fouten neemt af bij een toenemende waarnemingsdichtheid. Verhoging van de waarnemingsdichtheid vermindert dus het risico van systematische fouten.

De precisie van de voorspellingen blijkt sterk afhankelijk te zijn van de configuratie van de gerichte opnamelocaties: tussen steekproeven met een waarnemingsdichtheid van één per km² varieert de variantie van de fouten van 500 tot 1 500 cm². Bij hogere waarnemingsdichtheden neemt de variantie van de fout af tot 500 à 800 cm² bij een waarnemingsdichtheid van acht per km². Grotere waarnemingsdichtheden leiden dus tot preciezere voorspellingen.

Uit de validatieresultaten blijkt dat bij een steekproef met proportionele allocatie over de Gt's het risico van een systematische fout groter is dan bij een steekproef met preferente allocatie in natte Gt's. Dit kan vooralsnog niet worden verklaard.

Bij steekproefopzetten met één, twee en acht waarnemingen per km² blijkt de nauwkeurigheid van de GLG-voorspellingen significant ($\alpha=0.05$) te worden

overschat. Bij één waarneming per km² wordt de nauwkeurigheid zes tot acht van de tien keer te groot ingeschat, gemiddeld zelfs met een factor twee. Een mogelijke verklaring is dat het aantal vrijheidsgraden van het gekozen regressiemodel wordt overschat. Bij stapsgewijze regressie moet het aantal vrijheidsgraden namelijk worden verlaagd, omdat de predictoren niet vaststaan maar op basis van de beschikbare gegevens worden geselecteerd. Hoe het juiste aantal vrijheidsgraden kan worden bepaald bij groepsgewijze modelselectie van maximaal vijf uit de 15 beschikbare predictoren moet nader worden onderzocht.

Als wordt gecorrigeerd voor 'pseudomeetfouten' in de geschatte GLG op gerichte opnamelocaties (de voorspelfouten ten gevolge van de stambuisregressie) is de modelvariantie hoger dan variantie die door validatie is bepaald. Als de correctie achterwege blijft treedt echter het omgekeerde op. De correctie voor 'pseudo-meetfouten' leidt dus tot overcompensatie. De oorzaak hiervan is dat wordt verondersteld dat de pseudomeetfouten ruimtelijk ongecorreleerd zijn. Bij lage waarnemingsdichtheden wordt deze overcompensatie deels opgeheven door het effect van het te grote aantal gehanteerde vrijheidsgraden, waardoor de variantie uit validatie en de modelvariantie bij de laagste waarnemingsdichtheden redelijk met elkaar overeenkomen.

De conclusie van het eerste deelonderzoek is dat het gebruik van een niet-lineair in plaats van een lineair tijdreeksmodel niet tot nauwkeuriger GxG-voorspellingen leidt. Uit het tweede deelonderzoek blijkt dat de keuze van het tijdstip voor de gerichte opname van de grondwaterstand veel meer invloed op de kwaliteit van de GxG-voorspellingen heeft dan de selectie van stambuizen. De restvariantie van het gefitte stambuisregressiemodel is een goed criterium voor het kiezen van een geschikte dag voor gerichte opname. Belangrijk is dat er in de dagen voorafgaand aan de gerichte opname geen neerslag is gevallen. Kanssteekproeven voor de selectie van stambuizen zijn aan te bevelen boven gerichte selectie. In het bijzonder gestratificeerde steekproeven met proportionele allocatie zijn aan te bevelen, omdat deze zuiver schattingen opleveren van regressiecoëfficiënten, GxG-voorspellingen en de variantie van de voorspelfout. Een groter aantal dan de gangbare 20 stambuizen wordt aanbevolen als nauwkeurige kwantificering van de onzekerheid belangrijk is. Uit het derde deelonderzoek blijkt dat bij de gebruikelijke dichtheid van één gerichte opnamelocatie per km² aanzienlijke systematische en toevallige fouten in de GxG-voorspellingen kunnen optreden, en de onzekerheid onnauwkeurig wordt gekwantificeerd. Aanbevolen wordt de dichtheid van de gerichte-opnamelocaties te verhogen. De kwantificering van de onzekerheid kan verder worden verbeterd door 'pseudo-meetfouten' te verdisconteren, en door het juiste aantal vrijheidsgraden bij de selectie van het regressiemodel voor de relatie tussen GxG en hulpinformatie in rekening te brengen. Naar dit laatste is nader onderzoek noodzakelijk.

Er zijn geen fundamentele fouten in de Gd-methode aan het licht gekomen. Grondige herziening van de methode is dus niet nodig. Met relatief kleine aanpassingen kan de GxG nauwkeuriger worden voorspeld en de onzekerheid over de werkelijke GxG beter worden gekwantificeerd.

1 Inleiding

1.1 Probleemstelling en achtergrond

De nauwkeurigheid van de gebiedsdekkende voorspelling van de GxG (verzamelterm voor GHG, GVG en GLG) die de Gd-karteringsmethode (Finke et al., 2002) levert, is in veel gevallen acceptabel. Voor specifieke doeleinden kan echter een hogere voorspelnauwkeurigheid nodig zijn dan tot nu toe werd gerealiseerd. De gebruikelijke toepassing van de methode op regionale schaal levert voorspellingen van de GHG met standaardafwijkingen van gemiddeld 30 cm (Hoogland et al., 2003). De Gd-methode levert overigens kwantitatieve informatie over de voorspelnauwkeurigheid, op van kaarten van de voorspelfout. Om de Gd-methode toe te kunnen passen voor doeleinden die hoge eisen stellen aan de nauwkeurigheid van de GxG-voorspellingen moet onderzocht worden hoe de nauwkeurigheid van de Gd-methode te verhogen is. De nauwkeurigheid van de GxG-voorspellingen wordt door een groot aantal foutenbronnen beïnvloed. Het is onduidelijk hoe groot de relatieve bijdragen van de diverse foutenbronnen is. Daarom is het van belang te weten wat de belangrijkste foutenbronnen zijn. Het inzicht in de bijdrage van de verschillende foutenbronnen kan worden gebruikt om prioriteiten aan te geven bij methodische verbeteringen ten behoeve van nieuwe karteringen, en om beter te kunnen sturen op de verhouding tussen kosten en kwaliteit van Gd-karteringen.

De aanleiding voor dit onderzoek is een casus waarin de opdrachtgever een kwaliteit van de GxG-kaarten verwachtte die met de huidige Gd-methode en tegen het beschikbare investeringsniveau niet kon worden geleverd. Hieruit bleek dat het belangrijk is om vóóraf duidelijk aan te geven welke kaartkwaliteit bereikt kan worden bij een bepaald investeringsniveau. Daarom zouden in de aanbestedingsfase van een project vragen moeten worden gesteld en beantwoord zoals:

- Op welke schaal worden de kaarten toegepast;
- wat is de gewenste kwaliteit (nauwkeurigheid), en
- welke waarnemingsdichtheid (en dus investering) hoort daar bij?

De beantwoording van deze vragen vereist inzicht in de belangrijkste oorzaken van fouten in de GxG-kaarten en in de grootte van deze fouten bij verschillende keuzes aangaande de dataverzameling. Kennis over de relatieve bijdrage van verschillende foutenbronnen helpt bij de prioriteitstelling van mogelijke verbeteringen. Deze verbeteringen betreffen zowel de kwaliteit van de gegevens als de methoden voor de verwerking van de gegevens.

Uitgangspunt van dit project is te zoeken naar verbeterpunten in de bestaande Gd-methode. Omdat de Gd-methode als uitgangspunt is gekozen zijn de bevindingen en aanbevelingen van dit onderzoek dan ook sterk gekoppeld aan deze methode.

1.2 Projectdoelstelling

Het doel van deze studie is om de belangrijkste en meest belovende mogelijkheden voor verbetering van de Gd-methode op te sporen, en deze aan de hand van concrete cases nader te onderzoeken. Eerst wordt een zo compleet mogelijke inventarisatie gemaakt van alle mogelijke punten waarop de Gd-methode verbeterd zou kunnen worden. Op basis hiervan selecteren wij de belangrijkste en meest veelbelovende mogelijkheden voor verbetering. Dit resulteert in drie deelonderzoeken:

1. onderzoek naar de gevolgen van veronderstelde lineariteit in de respons van grondwaterstanden op het neerslagoverschot. Onderzocht wordt of voorspellingen van GxG verbeteren indien de veronderstelde lineariteit in de gebruikte tijdreeksmodellen wordt vervangen door niet-lineaire relaties tussen grondwaterstand en neerslagoverschot die fysisch beter te verklaren zijn;
2. onderzoek naar de invloed van meettijdstip van gerichte grondwaterstandsopname en de omvang en samenstelling van de stambuizenset op de kwaliteit van de met stambuisregressie voorspelde GxG. De bestaande Gd-methode gaat uit van een minimum aantal van 20 stambuizen voor stambuisregressie. Het tijdstip van gerichte opname wordt bij voorkeur zo gekozen dat de grondwaterstand zich rond de GxG bevindt. In dit deelonderzoek worden deze keuzes onder de loep genomen en wordt bepaald of het maken van andere keuzes tot verbetering van de Gd-methode leidt;
3. onderzoek naar het effect van de waarnemingsdichtheid en locatiekeuze van gerichte opnames op de kwaliteit van de GxG-voorspellingen en gekwantificeerde voorspelfouten. Dit deel van het project wordt ingevuld door in een reeds gekarteerd gebied de waarnemingsdichtheid te vergroten met aanvullend veldwerk en de modelmatige kartering voor verschillende waarnemingsdichtheden te herhalen.

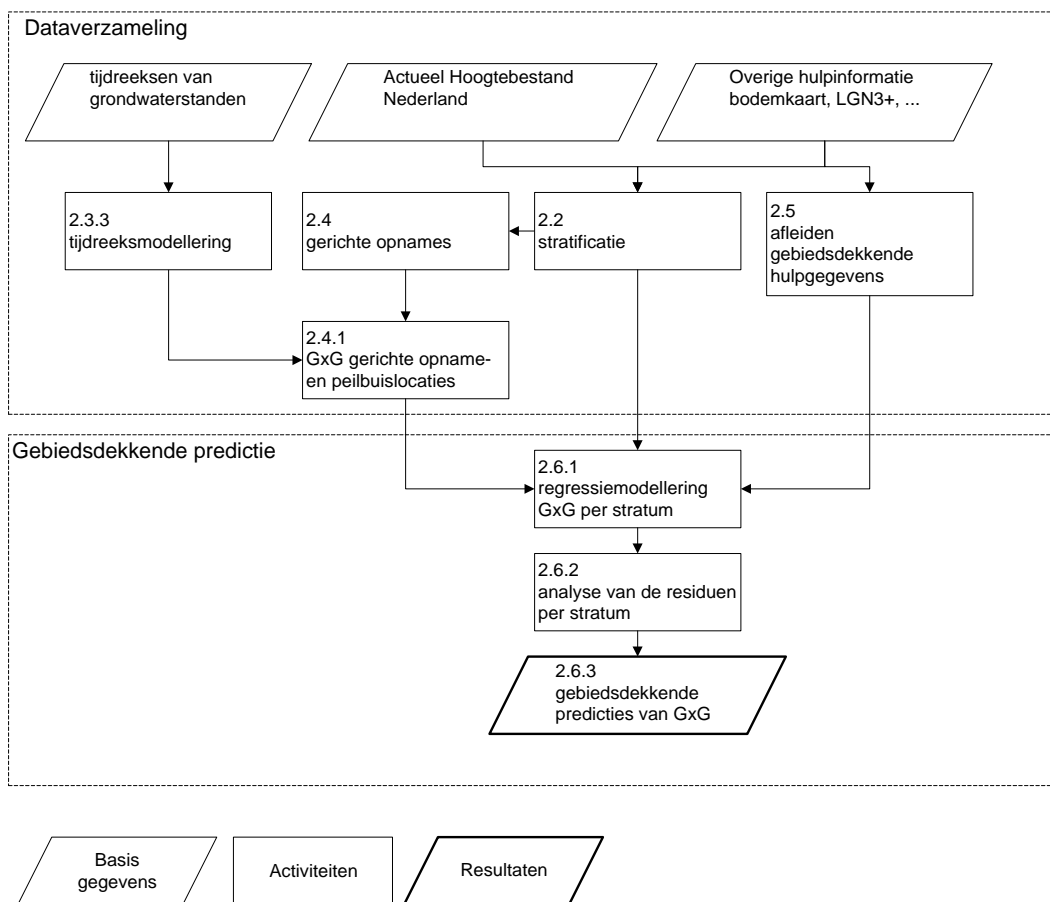
1.3 Leeswijzer

In hoofdstuk 2 wordt de Gd methode besproken zoals die in gangbare regionale karteringen wordt toegepast. Hoofdstuk 3 geeft een overzicht van de mogelijkheden tot verbetering en beschrijft welke verbeteringen nader worden onderzocht. In hoofdstuk 4, 5 en 6 worden van drie experimenten voor verbetering van de Gd-methode de onderzoeksopzet en de resultaten besproken. Hoofdstuk 4 behandelt het experiment waarin het effect van het gebruik van een niet-lineair in plaats van een lineair tijdreeksmodel op de voorspelnaauwkeurigheid van de GxG wordt onderzocht. Hoofdstuk 5 beschrijft het experiment waarin de samenstelling van de stambuizenset en de keuze van het meetmoment worden geoptimaliseerd. In hoofdstuk 6 wordt het effect van de steekproefopzet en de waarnemingsdichtheid bij de gerichte opnames geanalyseerd. Hoofdstuk 7 sluit af met aanbevelingen en conclusies over de mogelijke verbeteringen van de Gd-methode.

2 Werkwijze en methoden bij kartering van de grondwaterstandsdynamiek

2.1 Inleiding

Dit hoofdstuk beschrijft de wijze waarop uit verschillende informatiebronnen gegevens worden afgeleid die als basis dienen voor de gebiedsdekkende voorspelling (kartering) van de GxG. De globale werkwijze bij de gebiedsdekkende voorspelling van GxG is weergegeven in Figuur 1. De nummers in de figuur verwijzen naar de paragrafen in dit hoofdstuk, waarin een gedetailleerde beschrijving van de methoden wordt gegeven zoals die in een regionale Gd-kartering wordt toegepast. Voor elke Gd-kartering kunnen door de opdrachtgever extra gegevens worden aangeleverd of kan om een andere waarnemingsdichtheid worden gevraagd.



Figuur 1 Stroomschema van de gevolde werkwijze bij een Gd-kartering

Een belangrijk onderdeel van de Gd-methode is de indeling van het studiegebied in homogene deelgebieden. Deze zogenaamde stratificatie speelt al een rol bij de verzameling van grondwaterstandswaarnemingen. Daarom volgt eerst een beschrijving van de stratificatie in § 2.2. Vervolgens komt in § 2.3 de selectie van

peilbuizen aan de orde, en de berekening van de GxG uit de tijdreeksen die in deze peilbuizen zijn waargenomen. § 2.4 gaat in op de selectie van de locaties waar de zogenaamde gerichte opnames worden verricht, en geeft een beschrijving van de wijze waarop voor deze locaties GxG's worden berekend. In § 2.5 volgt tenslotte een beschrijving van de gebiedsdekkende hulpinformatie over de topografie en de waterhuishouding. § 2.6 behandelt de manier waarop gebiedsdekkende voorspellingen van de GxG tot stand komen.

2.2 Indeling van het gebied in homogene deelgebieden

Het onderscheiden van homogene deelgebieden wordt stratificatie genoemd. Het doel van de stratificatie is om de onzekerheid over de gebiedsdekkende voorspellingen van de GxG te reduceren. Er mag namelijk worden verondersteld dat binnen homogene deelgebieden de samenhang tussen de GxG en allerlei hulpinformatie sterker is dan in het gebied als geheel. Er worden deelgebieden (strata) onderscheiden op basis van eigenschappen die gerelateerd zijn aan hydrologie en maaiveldshoogte. Voor elk van de strata wordt de samenhang tussen de GxG en gebiedsdekkende hulpinformatie beschreven met een regressiemodel. Dit regressiemodel wordt gebruikt bij de gebiedsdekkende predictie van de GxG. De regressiemodellen voor de verschillende strata kunnen onderling aanmerkelijk verschillen.

Voor de stratificatie wordt de volgende gebiedsdekkende informatie gebruikt:

- de geologische kaart van Nederland, schaal 1 : 50 000 (TNO-NITG);
- de geomorfologische kaart van Nederland, schaal 1 : 50 000 (Alterra);
- het Actueel Hoogtebestand Nederland, AHN 25 × 25 meter (Adviesdienst Geoinformatie en ICT);
- de ligging van waterlopen volgens de topografische kaart van Nederland, Top10-Vector (Topografische Dienst);
- de Landelijke Grondgebruikskaart Nederland, LGN3+, 25 × 25 meter (Alterra);
- de bodem- en Gt-kaart van Nederland, schaal 1 : 50 000 (Alterra).

Verder worden er gegevens gebruikt die niet in alle gevallen gebiedsdekkend beschikbaar zijn:

- de indeling in stroomgebieden en peilgebieden volgens de waterschappen;
- gedetailleerde bodemkaarten, schaal 1 : 10 000 of schaal 1 : 25 000, die zijn vervaardigd voor bijvoorbeeld landinrichtingsprojecten (Alterra);
- de locaties van droogvallende waterlopen;
- de ligging van gebieden met vernattingsprojecten.

De bodem- en Gt-kaart van Nederland, schaal 1 : 50 000, vormt de basis voor de stratificatie. Indien nodig worden de kaartvlakken aan de hand van additionele informatie gesplitst. De stratificatie verloopt in drie stappen:

1. Er worden geohydrologische hoofdeenheden onderscheiden op basis van dagzomende geologische formaties en een aantal gedigitaliseerde breuklijnen die op de geologische kaart zijn weergegeven;
2. De geohydrologische hoofdeenheden worden onderverdeeld in bodemkundig-hydrologische eenheden, op basis van de bodem- en Gt-kaart, en gegevens over het afwateringspatroon en de maaiveldshoogte. Hierbij wordt gelet op:
 - een indeling in Gt;
 - de aanwezigheid van leemlagen in de ondergrond;
 - de textuur van de bodem;
 - de aanwezigheid van grof zand in de ondergrond;
 - de intensiteit van afwateringspatroon.Voorbeelden van deze bodemkundig-hydrologische eenheden zijn:
 - beekdalen;
 - droge zandgronden;
 - lemige gronden met stagnatie van grondwater door leemlagen.
4. In deze stap worden de uiteindelijke strata gevormd. Vlakken van bodemkundig-hydrologische eenheden worden samengevoegd tot strata, om tot meer aaneengesloten strata te komen en versnippering tegen te gaan. De strata zijn gemiddeld 3000 hectare groot (Finke et al., 2002). In gebieden met grote bodemkundig-hydrologische verschillen kunnen strata echter kleiner zijn, terwijl in homogene gebieden de strata juist groter kunnen zijn.

2.3 De GxG op peilbuislocaties

2.3.1 Selectie van grondwaterstandsreeksen

De grondwaterstandsreeksen komen voor het overgrote deel uit de OLGA-databank van TNO-NITG (Van Bracht, 1988). Vanaf 1 januari 2001 heet deze databank DINO. Daarnaast worden korte meetreeksen van grondwaterstanden aangeleverd door de waterschappen en de provincies. Hierna wordt beschreven hoe grondwaterstandsreeksen worden geselecteerd die geschikt zijn voor de berekening van de GxG. De selectiecriteria zijn:

1. de reeksen moeten zijn waargenomen in peilbuizen met een filter dat maximaal vijf meter beneden maaiveld begint;
2. de grondwaterstand t.o.v. maaiveld moet bekend zijn;
3. de plaatscoördinaten van de buislocatie moeten bekend zijn;
4. er moeten minimaal negen waarnemingen per hydrologisch halfjaar zijn verricht;
5. de meetreeks moet minimaal drie jaar lang zijn;
6. de meetreeks moet doorlopen tot minimaal een half jaar voor het begin van de kartering;
7. de afstand van de peilbuis tot een rivier of kanaal moet tenminste 40 m zijn;
8. de afstand van de peilbuis tot waterlopen en beken moet tenminste 20 m zijn;

9. de afstand van de peilbuis tot een watervoerende perceelssloot moet tenminste acht m zijn;
10. het meetpunt mag niet worden beïnvloed door afstromend oppervlaktewater (bijvoorbeeld nabij verharding), zich bevinden in een kuil, op een dijk, een oprit, en dergelijke.

De criteria 7 tot en met 10 worden getoetst in het veld.

2.3.2 Overige grondwaterstandsreeksen

Waterschappen en provincies exploiteren een aantal grondwaterstandsmeetnetten. Over het algemeen hebben deze meetnetten een tijdelijk karakter en zijn ingericht voor een specifiek doel. De gegevens die door waterschappen en provincies worden aangeleverd moeten tenminste voldoen aan criteria 1, 2 en 3 hierboven. Afhankelijk van de waarnemingsfrequentie (criterium 4), de reekslengte (criterium 5) en de waarnemingsperiode (criterium 6) wordt besloten op welke wijze uit de gegevens de GxG wordt berekend. Als de gegevens voldoen aan de criteria die hierboven zijn gesteld dan wordt de GxG berekend met behulp van tijdreeksmodellering (zie § 2.3.3). Als het aantal waarnemingen klein is, dan zullen de gegevens worden beschouwd als gerichte opnames van de grondwaterstand en zal de methode te bespreken in § 2.4 worden gevolgd. Als de reeksen te kort zijn voor tijdreeksmodellering, dan kan de methode van regressiemodellering worden gevolgd die is beschreven door Oude Voshaar en Stolp (1997). In tegenstelling tot de peilbuizen uit het OLGA/DINO-bestand worden de buizen uit de lokale meetnetten niet in het veld bezocht. Dit is een gevolg van de korte doorlooptijd tussen het beschikbaar komen van de gegevens en de kartering.

2.3.3 Berekening van de klimaatrepresentatieve GxG

Voor de berekening van een klimaatrepresentatieve GxG op peilbuislocaties worden tijdreeksmodellen gebruikt die de samenhang tussen het neerslagoverschot en de grondwaterstand beschrijven. Dit is nodig omdat maar zelden gedurende dertig jaar grondwaterstanden worden gemeten op een locatie, zonder dat het hydrologische regime in die periode wijzigt door ingrepen in de waterhuishouding. Deze modellen worden gebruikt om op basis van reeksen van neerslagoverschotten die door het KNMI gedurende tenminste dertig jaar zijn verzameld grondwaterstandsreeksen te simuleren. Uit deze gesimuleerde grondwaterstandsreeksen van dertig jaar lang kunnen klimaatrepresentatieve parameters zoals de GxG en Gt worden berekend, en kunnen duurlijnen en regimecurves worden geconstrueerd (Knotters en van Walsum, 1994).

Er is gebleken dat op basis van uitsluitend korte reeksen een vertekend (systematisch te nat of te droog) beeld van de grondwatersituatie kan ontstaan ten opzichte van de klimaatperiode van dertig jaar. Uit een analyse van Knotters en Bierkens (1999) blijkt dat reeksen van vier tot acht jaar meestal lang genoeg zijn om de samenhang tussen

het neerslagoverschot en de grondwaterstand te kunnen modelleren, en dat ook reeksen met lengte van drie jaar in veel situaties nog zullen voldoen.

In het tijdreeksmodel dat de dynamische relatie tussen neerslagoverschot p en de grondwaterstand b beschrijft bestaat de grondwaterstand op tijdstip t uit een som van twee componenten: een deterministische of transfercomponent, b_t^* , en een ruiscomponent, n_t ; $b_t = b_t^* + n_t$. De relatie wordt transfer-ruismodel genoemd.

De deterministische component b_t^* is dat deel van de grondwaterstand dat kan worden verklaard uit een lineaire samenhang met het neerslagoverschot. De deterministische component b_t^* wordt geschat met: de vorige grondwaterstandmeting uit de tijdreeks en het neerslagoverschot tussen de huidige en de vorige meting. De ruiscomponent n_t bevat de resterende invloeden. Knotters (2001) toonde aan dat dit model eenvoudig fysisch is te verklaren en in veel situaties een goede beschrijving van de grondwaterstandsdynamiek geeft. Het transfer-ruismodel is ingebed in een Kalman-filter waardoor kalibratie op de minder frequent en onregelmatig waargenomen grondwaterstandsreeksen mogelijk is (Bierkens et al., 1999). Het gehanteerde tijdreeksmodel wordt in §4.1.1. verder uitgeschreven.

Het gebruik van alleen de deterministische component bij de simulaties zou leiden tot een onderschatting van de temporele variatie. Omdat zowel de GHG als de GLG extreme grondwaterstanden beschrijven, is het van belang dat de temporele variatie juist wordt gesimuleerd. Daarom wordt gebruik gemaakt van stochastische simulatie, (zie Knotters, 2001, hoofdstuk 2). Omdat de grondwaterstandsreeksen worden gesimuleerd door middel van stochastische simulatie, kan ook de onnauwkeurigheid als gevolg van het veronderstelde verband tussen neerslagoverschot en grondwaterstand worden gekwantificeerd.

De coëfficiënten van het transfer-ruismodel worden gekalibreerd met het programma KALTFN. De neerslaggegevens zijn afkomstig van het KNMI-neerslagstation dat zich het dichtst bij de peilbuis bevindt. De gegevens betreffende de referentie-gewasverdamming zijn eveneens afkomstig van het dichtstbijzijnde weerstation.

Het dagelijkse potentiële neerslagoverschot p_t wordt berekend uit het verschil tussen dagneerslagsom (p_N) en de etmaalverdamming voor een referentiegewas volgens Makkink (e_M), zoals deze door het KNMI is gegeven (De Bruin, 1987):

$$p_t = p_{N,t} - e_{M,t} \quad (2.1)$$

Met het gekalibreerde tijdreeksmodel en de neerslagoverschotreeks van de dertig jaar worden voor elke peilbuislocatie 100 grondwaterstandsreeksen gesimuleerd van dertig jaar lang. Uit deze 100 gesimuleerde reeksen wordt de GxG en zijn standaardfout berekend. De berekening van de GxG uit de gesimuleerde reeksen van dertig jaar lang verloopt in twee stappen:

1. per hydrologisch jaar wordt het gemiddelde van de drie hoogste en laagste grondwaterstanden genomen (respectievelijk HG3 en LG3);
2. de dertig HG3's en LG3's worden gemiddeld tot een GHG en een GLG.

2.4 De GxG op gerichte opnamelocaties

Tijdens de kartering worden grondwaterstanden gemeten in een groot aantal boorgaten, op vooraf vastgestelde locaties en tijdstippen. Omdat zowel het tijdstip als de locatie van de opnames vooraf worden bepaald, wordt er gesproken van twee 'gerichte opnames'. Het doel van deze gerichte opnames is om het net van GxG's te verdichten, zodat er uiteindelijk voldoende GxG waarnemingen zijn om de statistische relatie met maaiveldhoogten en daarvan afgeleide eigenschappen te kunnen analyseren (§ 2.6). Hiertoe is het nodig dat in elk stratum voor tenminste twintig locaties GxG-waarden bekend zijn. Een aanvullende eis is dat deze locaties ruimtelijk gelijkmatig zijn verspreid over het stratum en over de droge en natte terreindelen die in het stratum voorkomen. De locaties worden als volgt gekozen:

1. voor elk stratum wordt bepaald hoeveel waarnemingen er worden gedaan (minimaal twintig);
2. met het Top10-Vectorbestand en het AHN wordt een kaart van de drooglegging gemaakt met voor elke 25×25 m-pixel een drooglegging (zie § 2.5.1);
3. de droogleggingen, of indien deze niet beschikbaar zijn de Gt-klassen, worden per stratum gesorteerd van nat (ondiep) naar droog (diep);
4. de gesorteerde droogleggingen of Gt's worden in een aantal klassen opgedeeld gelijk aan het aantal waarnemingen in het stratum. Deze klassen zijn kleiner in 'natte' dan in 'droge' terreindelen, waardoor relatief meer opnames in natte terreindelen plaatsvinden;
5. per klasse wordt één waarnemingslocatie geloot en twee reservelocaties op voldoende afstand daarvan verwijderd.

Er wordt tweemaal gemeten: één maal in de winter, als de grondwaterstand zich rond het GHG-niveau bevindt, en één maal in de zomer, als de grondwaterstand zich rond het GLG-niveau bevindt. Tijdens neerslagperiodes wordt niet gemeten, omdat de diepte van het freatisch vlak dan te veel varieert in de tijd. Tijdens het veldwerk worden boorgaten gemaakt tot tien cm onder het grondwatervniveau (maar nooit dieper dan 2,50 m). Na een instelperiode (1-2 dagen) wordt de grondwaterstand gemeten.

2.4.1 Berekening van de GxG voor gerichte-opnamelocaties

Op de tijdstippen van de gerichte opnames worden ook grondwaterstanden gemeten in de peilbuizen waar de GxG met een tijdreeksmodel is berekend (§ 2.3.3). Deze peilbuizen liggen idealiter in de directe omgeving van de meetpunten. In de praktijk worden er op één meetdag in uitgestrekte gebieden metingen verricht, en liggen de

peilbuizen over een groot gebied verspreid. Het aantal peilbuizen waarin per meetdag wordt gemeten varieert tussen de vijftien en de vijfentwintig. Dit is voldoende om voor elke meetdag een regressiemodel op te stellen dat de relatie beschrijft tussen de waargenomen grondwaterstanden en de berekende GxG. Het lineaire regressiemodel dat de relatie beschrijft tussen de GxG en een gemeten grondwaterstand noemen we het stambuisregressiemodel. Dit model heeft de volgende vorm:

$$GxG = \beta_0 + \beta_1 x_1 + \epsilon, \quad (2.2)$$

waarin x_1 de grondwaterstand, β_0 en β_1 regressiecoëfficiënten en ϵ de foutenterm;

Met het stambuisregressiemodel wordt de GxG voorspeld uit grondwaterstandswaarnemingen die tijdens de gerichte opname zijn verricht en worden de standaardfouten van deze voorspellingen berekend. De standaardfouten van de GxG-voorspellingen op puntniveau worden ook wel aangeduid als pseudomeetfouten en voor deze meetfouten kan in de verdere verwerking worden gecorrigeerd (zie § 2.6.1).

Soms wordt de grondwaterstand niet waargenomen binnen de maximale boordiepte van 2,50 meter. Dit noemen we ‘een gecensureerde waarneming’: de precieze waarde is onbekend, maar wel is bekend dat de waarneming ‘dieper dan’ een grenswaarde is. In deze situatie kiezen wij voor de volgende benadering (Cohen, 1991): eerst wordt een *maximum likelihood*-schatting gemaakt van de grondwaterstand. Daarbij wordt de code ‘> 2,50’ vervangen door de meest waarschijnlijke diepte groter dan 2,50 m. Dit getal wordt vervolgens ingevoerd in de regressievergelijking. De check op gecensureerde waarnemingen en de vervanging met het meest waarschijnlijke getal is standaard ingebouwd in het computerprogramma waarmee de GxG wordt voorspeld.

2.5 Gebiedsdekkende hulpinformatie over topografie en waterhuishouding

Het Actueel Hoogtebestand Nederland wordt eerst gecontroleerd op ‘onnatuurlijke’ hoogten zoals veroorzaakt door schuren, huizen, wegen en viaducten. Deze onnatuurlijke hoogten zijn verwijderd. Vervolgens zijn alle cellen die behoren tot de LGN4-klassen ‘zoet water’, ‘zout water’, ‘stedelijk bebouwd gebied’ en ‘hoofdwegen en spoorwegen’ verwijderd. De standaardafwijkingen van de hoogten is berekend voor een schuivend vensters van de drie × drie omliggende cellen. Cellen zijn verwijderd als deze standaardafwijking groter is dan 100 cm, en de cellen behoren tot één van de volgende klassen van het LGN4-bestand: ‘bebouwing in agrarisch en buitengebied’, ‘loof- en naaldbos in bebouwd gebied’, ‘bos met dichte bebouwing’, ‘gras in bebouwd gebied’, ‘kale grond in bebouwd buitengebied’. Als de standaardafwijking groter is dan 100 cm en de cellen liggen minder dan 50 m verwijderd van ‘hoofdwegen en spoorwegen’ zijn deze ook verwijderd.

2.5.1 Afleiding van gebiedsdekkende hulpinformatie

Van het AHN-bestand, de 1 : 50 000 Gt-kaart en het Top10- Vectorbestand zijn vijf groepen hulpinformatie afgeleid. Elk van deze groepen hulpinformatie bevat één of meer kaarten met daarop parameters die hydrologisch relevant zijn. Parameters die min of meer uitwisselbaar zijn, zijn ondergebracht in dezelfde groep. De inhoud en de afleiding van de gebiedsdekkende hulpinformatie wordt hieronder voor elk van de vijf groepen beschreven. Het AHN-bestand tezamen met de hulpbestanden wordt in het vervolg het 'AHN+-bestand' genoemd.

Groep 1: relatieve maaiveldhoogten

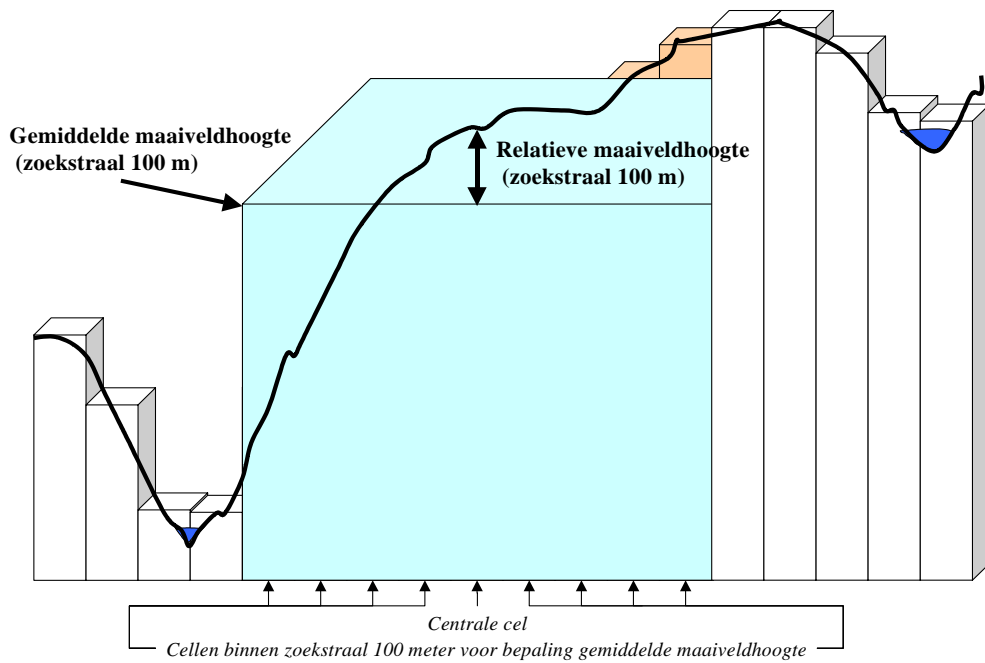
Uit onderzoek van Te Riele en Brus (1992) en Te Riele et al. (1995) is gebleken dat grondwaterstanden kunnen samenhangen met de maaiveldhoogte ten opzichte van NAP en met de relatieve maaiveldhoogte. Met relatieve maaiveldhoogte wordt de hoogte van een punt ten opzichte van de gemiddelde hoogte in een gebied binnen een bepaalde straal rond dat punt bedoeld (Figuur 2). Voor elk punt in het AHN is voor omgevingen met een straal van 100, 200, 300, 400 en 500 meter de relatieve maaiveldhoogte bepaald.

Groep 2: drainagedichtheid

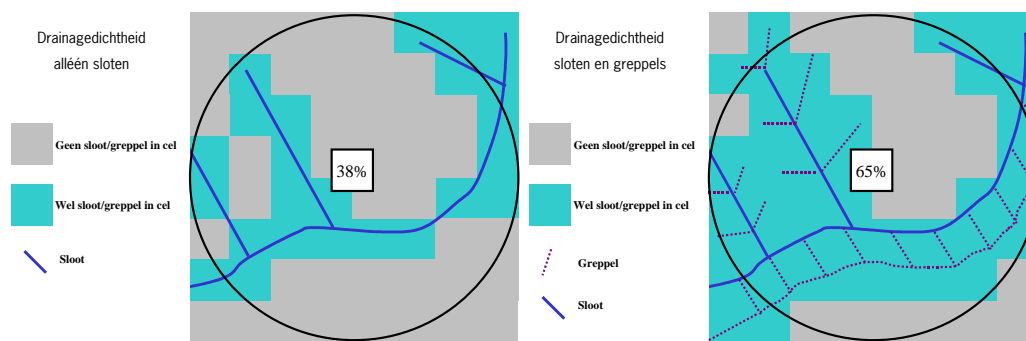
De dichtheid waarmee een gebied is ontwaterd beïnvloedt de grondwaterstand (met name de GHG, maar indien er sprake is van waterinlaat ook de GLG). Daarom wordt de drainagedichtheid gebiedsdekkend geschat en gebruikt als hulpinformatie. Uit het Top10-Vectorbestand zijn alle watergangen geselecteerd. Hiervan zijn twee bestanden gemaakt: een bestand met alle watergangen ('sloot en greppel') en een bestand waar de detailontwatering uit is verwijderd ('alleen sloot'). Voor elke 25 × 25 meter cel is bepaald in hoeveel naburige cellen binnen een zoekstraal van 300 en 400 meter drainage voorkomt (Figuur 3). Hieruit volgt een indicatie van de drainagedichtheid voor 'alleen sloot' en 'sloot en greppel'.

Groep 3: drooglegging ten opzichte van maaiveld

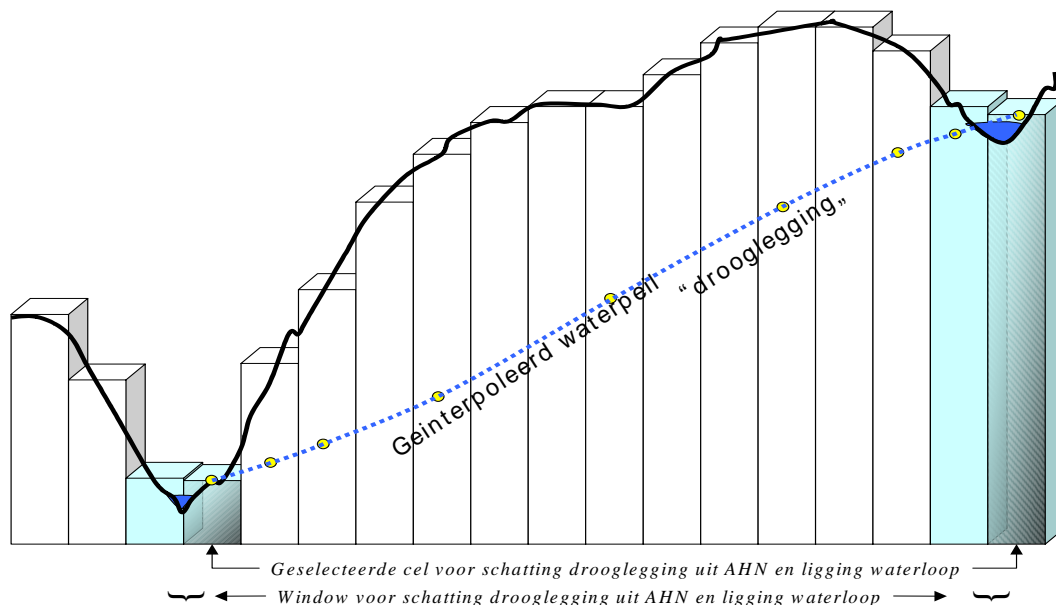
De drooglegging is het hoogteverschil tussen de waterspiegel in een waterloop en het grondoppervlak. Het kan worden geïnterpreteerd als het verwachte effect van het peilbeheer en de waterlopeninfrastructuur op de grondwaterstand. Een gebiedsdekkende schatting van de drooglegging zal daarom naar verwachting nuttige hulpinformatie opleveren bij een grondwaterstandskartering.



Figuur 2 Bepaling relatieve maaiveldhoogte uit het AHN



Figuur 3 Bepaling van de drainagedichtheid



Figuur 4 Schatting van de drooglegging

Uit een combinatie van de gedigitaliseerde waterlopen uit het Top10-Vectorbestand en het AHN is een bestand afgeleid waar per 25×25 meter pixel de drooglegging is bepaald (Figuur 4). Dit is gebeurd in drie stappen:

1. in elk segment van een watergang is een peil ten opzichte van NAP geschat met de waarde van het laagst gelegen punt van het AHN in de directe omgeving;
2. dit peil ten opzichte van NAP is gebiedsdekkend geïnterpoleerd, gewogen naar de inverse afstand tot de waterloop;
3. door deze geïnterpoleerde peilen af te trekken van de maaiveldhoogte uit het AHN is de drooglegging ten opzichte van het maaiveld gebiedsdekkend berekend.

Waarschijnlijk geeft de drooglegging die op bovenstaande wijze is berekend een systematisch te nat beeld van de drooglegging, omdat het peil geschat in stap 1 deels op hoogtecijfers van de oevers zal zijn gebaseerd, en niet op het waterpeil zelf. Met behulp van metingen van het slootpeil ten opzichte van de lokale maaiveldhoogte kan hiervoor worden gecorrigeerd. Zowel de gecorrigeerde droogleggingkaart als de kaart met oorspronkelijke schattingen van de drooglegging zijn als hulpinformatie gebruikt in het vervolg van het onderzoek.

Tabel 1 Onder- en bovengrenzen van GHG en GLG per Gt-klasse

Gt op kaart	GHGboven	GHGonder	GLGboven	GLGonder
I	-17	40	26	50
II	-9	40	50	80
II*	25	50	50	80
III	0	40	80	120
III*	25	40	80	120
IV	40	120	80	120
V	0	40	120	150
V*	25	40	120	160
VI	40	80	120	190
VII	80	120	160	260
VII* / VIII	140	220	160	400

Groep 4: maaiveld ten opzichte van NAP

Dit is de absolute maaiveldhoogte volgens het 25 × 25 meter AHN-bestand.

Groep 5: de GHG en GLG volgens de Gt-kaart 1 : 50 000 en de geschatte berging

De Gt-kaart 1 : 50 000 is weliswaar verouderd wat betreft de absolute niveaus van de GHG en GLG, maar geeft de ruimtelijke variatie mogelijk nog wel goed weer. Om die reden is de 1 : 50 000 Gt-kaart omgezet in kaarten van de GHGoud en GLGoud, waarbij gebruik is gemaakt van de karakterisatie van Gt's door Van der Sluijs (1982, 1990), zie Tabel 1. De Gt-klassen van de 1 : 50 000 Gt-kaart en het AHN-maaiveldhoogtebestand worden gebruikt om een neergeschaalde kaart van de GHG en GLG te maken. Voor elk Gt-vlak van de 1 : 50 000 kaart wordt het 15-de en 85-ste percentiel van de AHN-hoogten in dat vlak bepaald. Voor de hoogten tussen het 15 -de en 85-ste percentiel wordt verondersteld dat deze lineair samenhangen met de GHG's en GLG's tussen de onder- en bovengrens voor de betreffende Gt-klasse. Vervolgens worden de hoogten uit het AHN-bestand met deze lineaire relatie getransformeerd in GHG's en GLG's. Voor de locaties waarvan de hoogte niet in het interval van het 15-de tot het 85-ste percentiel ligt, wordt een GHG en GLG geïnterpoleerd t.o.v. NAP, vanuit de overige punten die wel in het interval vallen. Vervolgens zijn deze GHG's en GLG's teruggerekend naar diepten t.o.v. maaiveld, met behulp van de hoogten uit het AHN-bestand.

Met behulp van de bodemfysische vertaling van de bodemkaart (Wösten et al., 1988), het programma CAPSEV (Wesseling, 1991) en de grondwaterstanden in Tabel 1 is de bergingscapaciteit bij GHGoud en GLGoud berekend. Deze bergingscapaciteiten zijn als gebiedsdekkende hulpinformatie gebruikt.

2.6 Gebiedsdekkende predictie van de GxG

De gebiedsdekkende predictie van de GxG vindt plaats in een aantal stappen, die weergegeven zijn in Figuur 1. Allereerst wordt met behulp van regressieanalyse de samenhang tussen de GxG op puntlocaties en de vijf groepen met gebiedsdekkende hulpinformatie onderzocht. De regressieanalyse wordt beschreven in § 2.6.1. en wordt per stratum uitgevoerd. Dit levert voor elk stratum een selectie van predictoren op, een gekalibreerd regressiemodel en een aantal residuen (verschilwaarden tussen de GxG op puntlocaties en op basis van regressie voorspelde waarden). Vervolgens wordt de ruimtelijke structuur van de residuen geanalyseerd. Dit gebeurt niet per stratum, maar voor alle strata tegelijk, teneinde over voldoende metingen te beschikken om de ruimtelijke structuur te kunnen modelleren. Voorafgaand hieraan worden de residuen eerst per stratum gestandaardiseerd, zodat de variantie van de residuen in elk stratum gelijk is aan één. Op basis van de gestandaardiseerde residuen wordt een model voor de ruimtelijke structuur, een variogram, geschat (zie § 2.6.2.). Dit variogram wordt vervolgens gestandaardiseerd, hetgeen wil zeggen dat het wordt geschaald naar de oorspronkelijke varianties van de residuen in de strata. Dit levert voor elk stratum een variogram op. De laatste stap betreft de predictie van GxG en wordt beschreven in § 2.6.3. Voor elke cel uit het 25×25 -metergrid wordt de GxG voorspeld op basis van de, met regressie, geselecteerde predictoren, de variogrammen en de waarden van de GxG op meetlocaties.

2.6.1 Regressieanalyse

Voor elk stratum is een model geselecteerd dat de samenhang beschrijft tussen de GxG op meetlocaties en de vijf groepen met gebiedsdekkende hulpinformatie beschreven in § 2.5.1. Omdat verklarende variabelen (gebiedsdekkende hulpinformatie) die tot dezelfde groep behoren naar verwachting uitwisselbaar zijn in de regressieanalyse, is uit elke groep slechts één variabele in het regressiemodel opgenomen. Aldus wordt een set kandidaatmodellen samengesteld, met elk een regressieconstante en met maximaal vijf predictorvariabelen. Voor elk stratum is het beste model geselecteerd op basis van het Mallows' Cp-criterium (Mallows, 1966). Om de nauwkeurigheid van de GxG-voorspellingen voor peilbuis- en gerichte opnamelocaties te kunnen verdisconteren is er een gewogen regressie uitgevoerd, waarbij de nauwkeurigheid als gewicht diende: hoe nauwkeuriger de GxG-predictie, hoe meer gewicht in de regressieanalyse. In formulevorm is het gewicht gelijk aan $f/(1 + s_i^2)$, waarin s_i^2 de variantie is van de fout in de berekende GxG op locatie i , $i = 1 \dots n$, en f een correctiefactor die ervoor zorgt dat de gewichten sommeren tot het aantal waarnemingen n gebruikt in de regressieanalyse.

2.6.2 Ruimtelijke structuur van de residuen

Het verschil tussen de GxG-waarde op een peilbuislocatie of de locatie van een gerichte opname (GxG_{obs}), en de waarde die is berekend met het geselecteerde regressiemodel (GxG_{mod}), noemen we residu en wordt als volgt berekend:

$$\varepsilon_k(x) = GxG_{\text{obs},k}(x) - GxG_{\text{mod},k}(x), \quad (2.3)$$

waarin k aangeeft in welk stratum het residu ligt. Het gemiddelde van de residuen per stratum is ongeveer nul en ze hebben een variantie s_{k}^2 . Waarschijnlijk vertonen de residuen een ruimtelijke correlatie of structuur. Het regressiemodel behoeft immers niet alle ruimtelijke structuur in de GxG verklaard te hebben uit de gebiedsdekkende hulpinformatie. Voor de predictie van de GxG en de berekening van de voorspelfout, is het van belang rekening te houden met de ruimtelijke structuur van de residuen. De ruimtelijke structuur wordt gemodelleerd met een (semi-)variogram (Davis, 2002, blz. 254-264). Het variogram geeft de afhankelijkheid weer tussen een waarneming op locatie x en een waarneming op locatie $x + h$, waarbij h een vector is die zowel afstand als richting aangeeft.

Het variogram wordt gemodelleerd voor GHG, GVG en GLG afzonderlijk. Eerst worden de residuen per stratum gestandaardiseerd door ze te delen door de variantie van de residuen per stratum s_{k}^2 . Na standaardisatie worden de residuen uit alle strata gebruikt bij de modellering van één en hetzelfde, gestandaardiseerd variogram. De variogrammen worden vervolgens per stratum gedestandaardiseerd op basis van s_{k}^2 . Dit stratum-specifieke variogram wordt nog gecorrigeerd voor de pseudomeetfout (eigenlijk predictiefout) van de GxG-opnamen door vermindering van de *nugget*-parameter van het variogram met de gemiddelde variantie van de pseudomeetfout in de GxG-opnamen per stratum. Indien de *nugget*-parameter van het variogram kleiner is dan de gemiddelde variantie van de pseudomeetfout wordt ze gelijk gesteld aan nul.

2.6.3 Universal kriging van GxG

De gebiedsdekkende predictie van GxG vindt voor elk stratum afzonderlijk plaats. Binnen een stratum wordt bij een standaard kartering rekening gehouden met de correlatie tussen GHG, GVG en GLG door cross-variogrammen te gebruiken bij predictie met Gstat (Pebesma en Wesseling, 1998). In deze studie, waar alleen GLG gebiedsdekkend wordt voorspeld, is geen rekening gehouden met de onderlinge correlatie tussen GHG, GVG en GLG. De predictie wordt hier met universal kriging uitgevoerd in het geval van simultane voorspelling van GHG, GVG en GLG met co-kriging. Als driftvariabelen bij kriging wordt gebruik gemaakt van de, met regressie, geselecteerde predictoren (gebiedsdekkende hulpinformatie).

Het resultaat is een gebiedsdekkende krigingpredictie van GxG en een gebiedsdekkende krigingvariantie van de GxG per stratum. De GxG-voorspellingen en varianties per stratum worden in een GIS samengevoegd tot kaarten voor het totale gebied. De resolutie en gebiedsbegrenzing is gelijk aan de gebiedsdekkende hulpinformatie waaraan de gebruikte predictoren zijn ontleend.

3 Mogelijke verbeterpunten in de Gd-methode

In het vorige hoofdstuk is in grote lijnen de Gd-methode beschreven. De Gd-methode is het product van vele jaren van onderzoek en heeft haar toepasbaarheid inmiddels bewezen. De Gd-methode is een waardevol instrument gebleken dat bestaansrecht heeft naast andere aanpakken, zoals het gebruik van hydrologische procesmodellen of traditionele, veldbodemkundige kartering van grondwaterstandskarakteristieken. De grote lijn van de Gd-methode willen we hier dan ook niet ter discussie stellen. Op onderdelen van de methode is aanpassing, verfijning, correctie en verbetering echter nodig om aan specifieke eisen die gebruikers stellen te kunnen beantwoorden. In dit hoofdstuk inventariseren we zo volledig mogelijk de punten waarop de methode verbeterd zou kunnen worden. Sommige punten op de lijst betreffen keuzes tussen twee of meer aanpakken; de vraag is dan welke aanpak de voorkeur verdient. Hieronder presenteren we de lijst van punten door stap voor stap door de methode te lopen.

3.1 Gebiedsindeling

De indeling in strata gebeurt idealiter zodanig, dat het regressiemodel dat later per stratum wordt opgesteld identificeerbaar is en een zo hoog mogelijke voorspelkracht heeft. Het eerste verbeterpunt is:

1. Optimalisatie van de gebiedsindeling. Hoe kies je deelgebieden zodanig dat het residu van de maaiveldregressie een zo klein mogelijke variantie heeft?

3.2 Gegevensverzameling

Waarnemingen van de grondwaterstanden in boorgaten en peilbuizen kunnen fouten bevatten, die grote invloed kunnen hebben op de uiteindelijke GxG-voorspellingen. Ook ten aanzien van de selectie van gerichte-opnamelocaties en de behandeling van de zogenaamde gecensureerde waarnemingen zijn verbeteringen mogelijk. De volgende drie verbeterpunten zijn daarom:

2. Mogelijk kunnen meetfouten in gerichte opnames en stambuistijdreeksen worden gereduceerd. Hoe kan dit gerealiseerd worden en welke controles op het maken van fouten zijn mogelijk bij het transport van gegevens?
3. Is een betere keuze van het aantal gerichte opnames (aantal per deelgebied) en de locaties van gerichte opnames mogelijk? De ruimtelijke spreiding is niet optimaal en de spreiding over de predictoren ruimte houdt alleen rekening met de predictor drooglegging of Gt volgens de bodemkaart.
4. Op sommige gerichte-opnamelocaties is een grondwaterstand dieper dan 2.5 meter waargenomen. De grondwaterstand wordt dan geschat met *maximum likelihood*. Hoe kan rekening worden gehouden met gegeven dat de aldus resulterende ‘meting’ een relatief grote meetfout heeft?

3.3 Tijdreeksmodellering

Selectie van stambuizen gebeurt op basis van technische gegevens over filterdiepte, meetfrequentie en de lengte van de waarnemingsperiode, gevolgd door controle van de ligging van stambuizen in het veld. Uit de gekalibreerde tijdreeksmodellen worden de GxG en de standaardfout van de GxG per stambuislocatie afgeleid. Hier komt een simulatie (meestal 100 runs) met een gecombineerd transfer-ruismodel aan te pas om tijdreeksen met een lengte van 30 jaar te genereren, waaruit vervolgens de GxG wordt berekend. Ten aanzien van de selectie van grondwaterstandsreeksen en de berekening van klimaatsrepresentatieve GxG's uit deze reeksen is verbetering op vijf punten mogelijk:

5. De ligging van OLGA-buizen is mogelijk niet representatief voor het onderzoeksgebied. Vaak ontbreken buizen in bouwland en gedraineerde percelen en er is een oververtegenwoordiging van buizen in wegbermen en huispercelen.
6. In de berekende voorspelfout (sd_{GxG}) is de onzekerheid over de parameters van het tijdreeksmodel niet verdisconteerd. Daardoor wordt de onzekerheid over de werkelijke GxG onderschat met de sd_{GxG} .
7. De meetreeksen worden niet gescreend op invoerfouten hoewel deze fouten grote gevolgen voor GxG voorspellingen kunnen hebben.
8. Er wordt een lineair tijdreeksmodel gebruikt terwijl op basis van proceskennis niet-lineaire relaties verwacht worden. Wellicht is een niet-lineair tijdreeksmodel beter in staat de GxG te voorspellen.
9. Ook andere aannames in de tijdreeksmodellering zouden kritisch onder de loep kunnen worden genomen. Is de aanname dat de ruis stationair is en normaal verdeeld bijvoorbeeld wel realistisch?

3.4 GxG-voorspellingen voor gerichte opnames

Er wordt een relatie gelegd tussen de GxG die berekend is op stambuislocaties en de gemeten grondwaterstand ten tijde van de gerichte opnames. Het verband wordt beschreven met een lineair regressiemodel (de zogenaamde stambuisregressie), met als kandidaatpredictoren zowel zomer- als winterstanden. De predictoren wordt nu geselecteerd op basis van Mallows' C_p -criterium. Ten aanzien van de GxG-voorspellingen op gerichte-opnamelocaties is op de volgende punten verbetering mogelijk:

10. Is een ander statistisch, of wellicht een fysisch criterium voor selectie van predictoren beter?
11. Is het gebruik van een niet-lineair regressiemodel voor de stambuisregressie beter?
12. Hoe kan de keuze van het moment van de gerichte opname worden geoptimaliseerd?
13. Hoe belangrijk is de samenstelling van een gebiedsspecifieke stambuizen set en hoe kan deze stambuizen set het beste worden gekozen?

14. De stambuisregressie houdt geen rekening met onnauwkeurigheid van de berekende GxG op stambuislocaties. Levert het verdisconteren van deze onnauwkeurigheid relevante verbetering van de GxG-voorspellingen op?
15. Is extrapolatie bij voorspellingen met het stambuisregressiemodel acceptabel, of moet bij extrapolatie de variantie van voorspelfout extra worden verhoogd?

3.5 Afleiden hulpgegevens

Met behulp van het AHN worden diverse soorten gebiedsdekkende hulpinformatie afgeleid, zoals de relatieve maaiveldshoogte, de drooglegging en de neergeschaalde GHG en GLG uit de Gt-kaart 1: 50.000. Deze attributen fungeren in het vervolg van de methode als mogelijke predictoren bij de maaiveldregressie. De kwaliteit van de hulpinformatie en de wijze waarop deze wordt samengesteld kan worden verbeterd op de volgende punten:

16. Vaststellen en indien mogelijk verbeteren van de kwaliteit van het AHN.
17. Verbeteren van de kwaliteit van de andere predictoren.
18. Toevoegen van andere predictoren toevoegen (op basis van hydrologisch inzicht en recentelijk beschikbaar gekomen digitale informatie).

3.6 Gebiedsdekkende predictie van GxG

De gebiedsdekkende predictie van de GxG kan op de volgende punten worden verbeterd:

19. Wordt gezien de stapsgewijze (en groepsgewijze) selectie van predictoren wel op de juiste wijze rekening gehouden met vrijheidsgraden, en is de verhouding tussen het aantal waarnemingen en het aantal predictoren wel goed?
20. Is transformatie van predictoren dan wel GxG voorafgaand aan de maaiveldregressie zinvol?
21. Is verbetering mogelijk door vervanging van de lineaire maaiveldregressie door een *Generalised Linear Model* (niet-normaal verdeeld residu, gebruik van *link functions*)?
22. Is Mallows' Cp het juiste criterium bij de maaiveldregressie of is een ander statistisch criterium te prefereren? Zou selectie van de predictoren ook fysieke kennis kunnen worden gebruikt?, Zou met de predictoren die voor aangrenzende deelgebieden zijn geselecteerd rekening gehouden moeten worden?
23. Moet de mogelijkheid open worden gehouden dat residuen gecorreleerd zijn over deelgebiedsgrenzen heen?
24. Hoe groot is het effect van een hogere meetnetdichtheid van gerichte opnames op de nauwkeurigheid van de gebiedsdekkende GxG-voorspellingen?
25. Kan worden aangenomen dat het gestandaardiseerde variogram gelijk is voor alle deelgebieden of moet deze veronderstelling worden afgezwakt?

3.7 Overig

Hieronder een restgroep van verbeterpunten die niet eenvoudig bij één van de bewerkingsstappen hierboven zijn onder te brengen.

26. Rekening houden met of reduceren van positionele fouten in de locatiebepaling van gerichte opnames en stambuislocaties.
27. De rasterkaarten hebben een resolutie van 25×25 meter. Een hogere resolutie met als gevolg meer ruimtelijk detail kan de resultaten wellicht verbeteren.
28. Meer fysieke proceskennis onderbrengen in de kwantificering van de relatie tussen de predictoren en de GxG.
29. Zorgdragen voor een degelijke validatie van (de uitkomsten van) de Gd-karteringsmethode.

3.8 Nader te onderzoeken mogelijkheden voor verbetering

Selectie van nader te onderzoeken onderdelen heeft plaatsgevonden op basis van het veronderstelde belang van het betreffende onderdeel voor de kwaliteit van de resulterende GxG-kaarten en de haalbaarheid van onderzoek aan deze onderdelen. Hierbij zijn de onderdelen 3, 5, 8, 12, 13, 16, 17, 18, 19, 24 en 28 als belangrijkste foutenbronnen uit de bus gekomen. Vanwege het grote aantal als belangrijk ervaren onderdelen was het niet mogelijk om deze onderdelen in dit project allemaal op hun belang te onderzoeken. Verdere inperking is dus noodzakelijk.

Onderdeel 5 over de representativiteit van beschikbare OLGA-buizen voor het te karteren gebied werd als belangrijk ingeschat. Met het experiment zoals beschreven in hoofdstuk 5, over de steekproefopzet voor de stambuizenset, kan hierop maar gedeeltelijk antwoord worden gegeven omdat alleen onderzocht kon worden hoe uit beschikbare OLGA-buizen het best een set stambuizen kan worden samengesteld. Het gegeven dat nauwelijks OLGA-buizen in akkerbouwpercelen en gedraineerde percelen beschikbaar zijn terwijl voorspellingen uit stambuisregressie wel voor deze gebieden worden uitgevoerd, maakt beantwoording van de complete onderzoeksvraag zoals genoemd bij punt 5 in dit project niet mogelijk.

Onderdelen 16 en 17 over de nauwkeurigheid van het AHN en de daaruit afgeleide gebiedsdekkende hulpinformatie die wordt gebruikt voor gebiedsdekkende predictie van de GxG werden ook als belangrijk ingeschat. Het bleek echter lastig om een haalbare onderzoeksopzet voor dit onderdeel te formuleren omdat daarvoor informatie over de werkelijke maaiveldshoogten (het AHN zonder fouten) nodig is. Deze onderdelen zijn, hoewel als belangrijk ingeschat, daarom in dit project niet verder onderzocht.

Na selectie op haalbaarheid van een onderzoeksopzet zijn de onderdelen 3, 8, 12, 13, 19 en 24 over gebleven. Een aantal van deze onderdelen kon soms in één experiment worden onderzocht. De volgende experimenten zijn daarom opgezet en behandelen de tussen haakjes aangeduide onderdelen uit de lijst:

1. Onderzoek naar het effect van de veronderstelde lineariteit in de response van grondwaterstanden op het neerslagoverschot. Onderzocht is of voorspellingen van de GxG verbeteren door het lineaire tijdreeksmodel te vervangen door een niet-lineair tijdreeksmodel (8).
2. Onderzoek naar het effect van het tijdstip van meting van de grondwaterstand van de gerichte opname en de steekproef van stambuizen op de kwaliteit van de met stambuisregressie voorspelde GxG op punten met een gerichte grondwaterstandsmeting (12 en 13).
3. Onderzoek naar het effect van de waarnemingsdichtheid en de steekproefopzet voor de gerichte opnames op de voorspelnaauwkeurigheden en de nauwkeurigheid van de geschatte voorspelfouten (3, 19 en 24).

Deze drie deelonderzoeken (experimenten) zijn in deze studie uitgewerkt en worden gerapporteerd in de hoofdstukken 4, 5 en 6.

4 Vergelijking van een lineair en niet-lineair tijdreeksmodel voor berekening van de GxG op stambuislocaties

4.1 Inleiding

De Gd-methode maakt gebruik van een lineair tijdreeksmodel om de Gd-karakteristieken op stambuislocaties uit tijdreeksen van grondwaterstanden te berekenen. In het voorgaande hoofdstuk is opgemerkt dat de Gd-methode mogelijk zou kunnen worden verbeterd door het lineaire tijdreeksmodel te vervangen door een niet-lineair model. Een niet-lineair model de werkelijke (niet-lineaire) fysische processen beter beschrijft dan het lineaire model, mits voldoende informatie over het beschouwde hydrologische systeem beschikbaar is voor de parametrisatie van de fysische processen.

In dit hoofdstuk onderzoeken we aan de hand van een concrete toepassing of vervanging van het lineaire door een niet-lineair model daadwerkelijk tot een substantiële verbetering leidt. Hiertoe vergelijken we de lineaire en niet-lineaire aanpak op een aantal aspecten. Hoewel de bevindingen van dit hoofdstuk strikt genomen alleen betrekking hebben op het beschouwde voorbeeld, zullen we aan het eind van het hoofdstuk toch proberen enkele meer algemeen geldende conclusies te trekken.

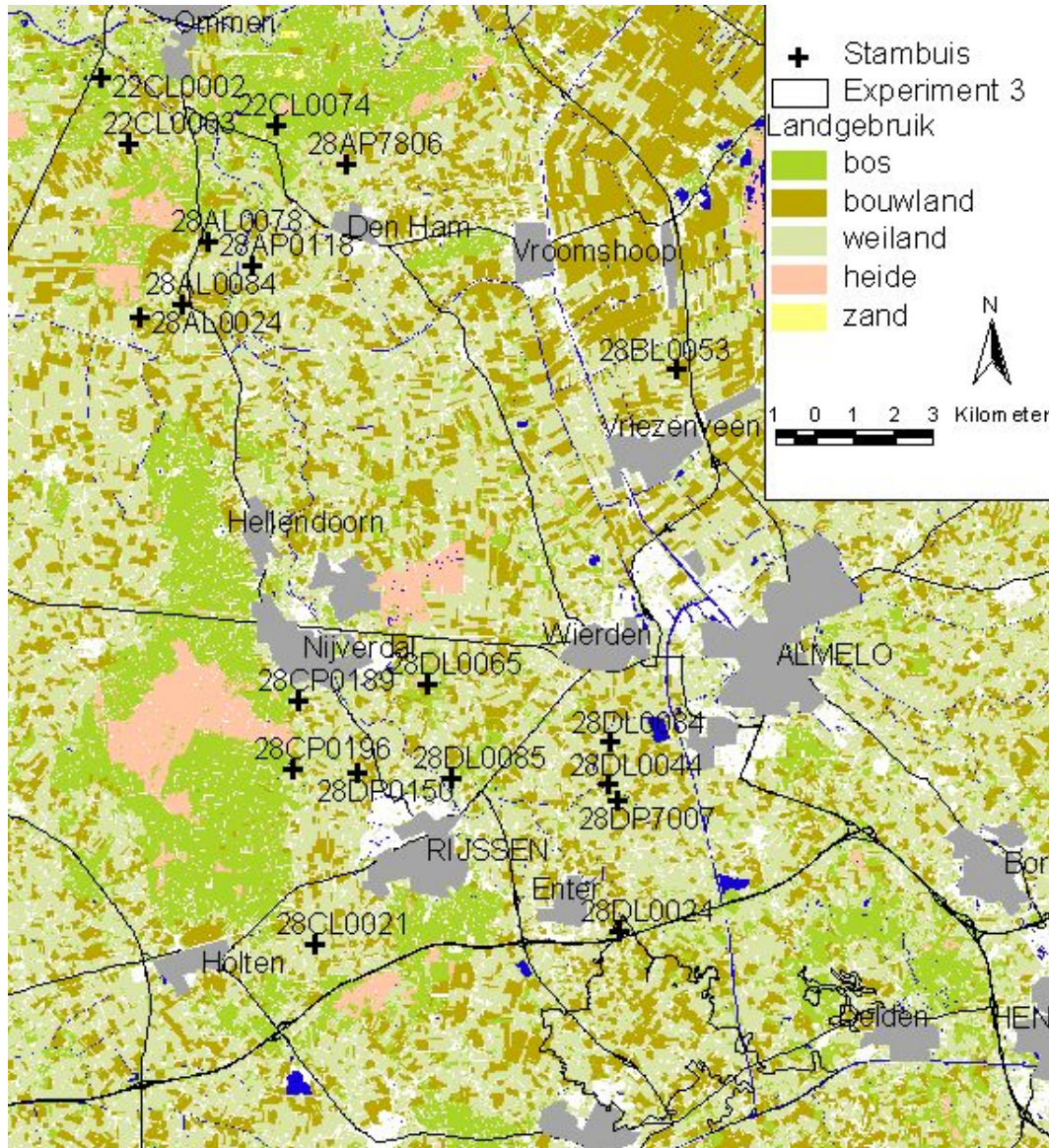
We willen in dit hoofdstuk de volgende deelvragen beantwoorden:

1. Hoe verschilt de berekening van de GxG in de niet-lineaire situatie van die in de lineaire situatie? Welke extra informatie is nodig en hoe (soepel) verloopt de procedure? (paragrafen 4.3 en 4.4)
2. Hoe groot zijn de verschillen in de voorspelde GxG op stambuislocaties bij gebruikmaking van het lineaire en niet-lineaire model? Kunnen we de waargenomen verschillen karakteriseren en verklaren? (§ 4.5)
3. Indien de lineaire en niet-lineaire modellen substantiële verschillen in voorspelde GxG te zien geven, welke van de twee modellen zit dan dichterbij de werkelijke GxG? (§ 4.6)
4. Leidt gebruikmaking van het niet-lineaire model tot betrouwbaardere voorspellingen van de regressiecoëfficiënten en tot een lagere restvariantie van de stambuisregressie? Met andere woorden, is bij het niet-lineaire model sprake van een sterker verband tussen de met het model verkregen GxG en de via gerichte opnames verkregen gerichte opnames van de grondwaterstand? (§ 4.7)

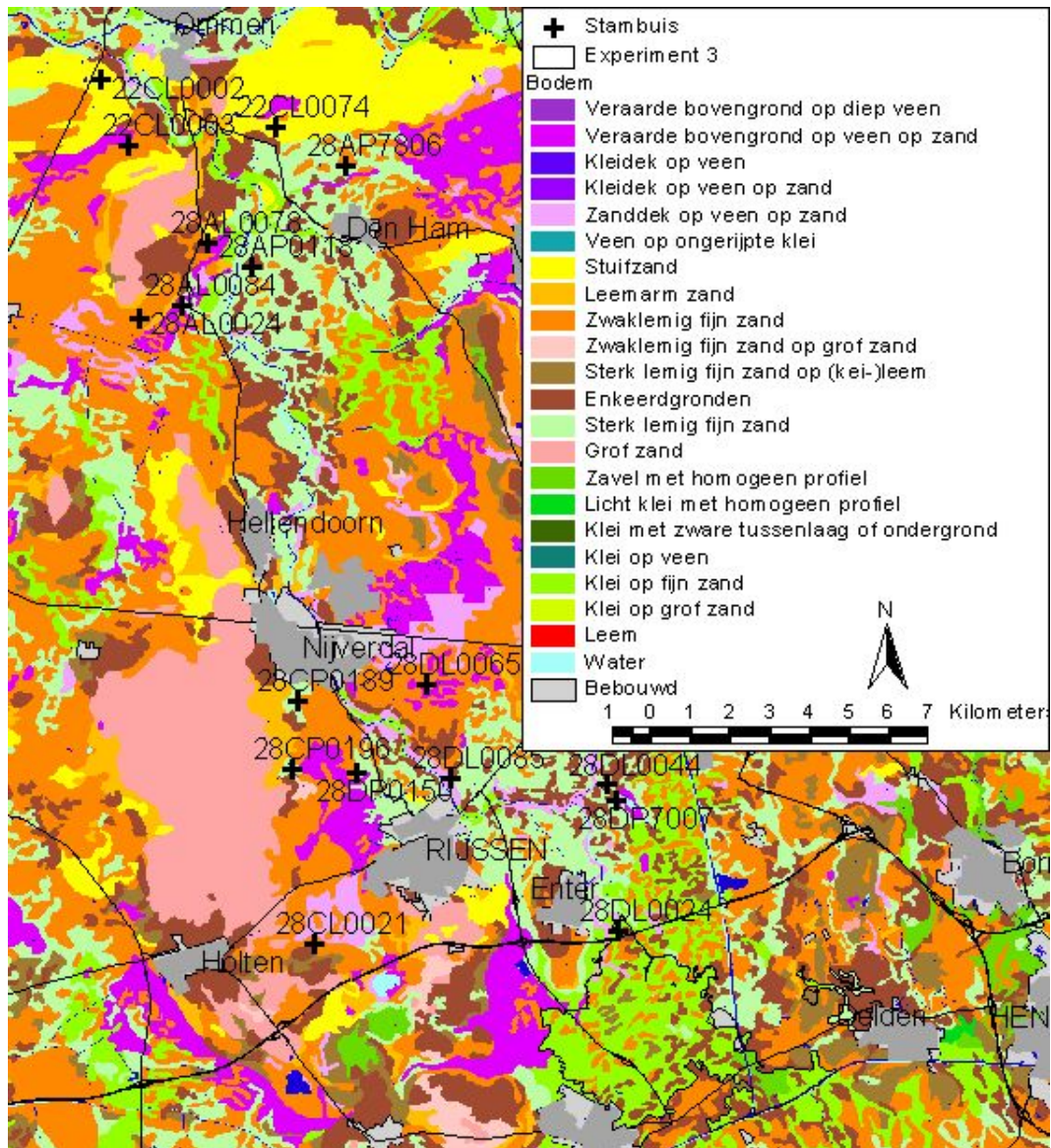
Voordat we aan beantwoording van de vier deelvragen toekomen geven we in § 4.2 een beschrijving van de data en het studiegebied, en in §4.3 een korte beschrijving van het lineaire en niet-lineaire tijdreeksmodel. We sluiten het hoofdstuk af met een discussie (§ 4.8) en trekken een aantal conclusies (§ 4.9).

4.2 Beschrijving van de stambuizenset en het onderzoeksgebied

De vergelijking van het lineaire en niet-lineaire model gebeurt aan de hand van een dataset bestaande uit 19 stambuizen uit het waterschap Regge en Dinkel. Deze stambuizen liggen rond Nijverdal, zie Figuur 5. Figuur 6 geeft de bodemopbouw van het gebied. Tabel 2 geeft een aantal kenmerken van de stambuizen en de meetreeksen.



Figuur 5 Ligging van de 19 stambuizen in het onderzoeksgebied nabij Nijverdal



Figuur 6 Bodemfysische schematisatie van de bodemkaart voor het onderzoeksgebied

Tabel 2 Kenmerken van de 19 stambuizen en meetreeksen

buisnummer	begin- datum tijdreeks	eind- datum tijdreeks	gemiddelde grondwater- stand [cm]	st.afwijking grondwater- stand [cm]	diepte slootbodem t.o.v. maaiveldhoogte op stambuislocatie [cm]
22CL0002	15-01-90	15-12-00	115	33	194
22CL0003	14-01-92	15-12-00	99	28	96
22CL0074	15-01-90	29-12-00	158	46	145
28AL0024	14-01-89	28-12-99	144	29	86
28AL0078	14-02-89	28-12-99	207	36	254
28AL0084	28-06-92	14-12-00	96	29	85
28AP0118	13-01-89	28-12-99	128	38	212
28AP7806	14-01-89	14-02-99	47	36	118
28BL0053	14-01-92	14-04-95	92	30	70
28CL0021	13-02-90	15-12-00	166	36	236
28CP0189	14-01-92	14-11-00	245	41	397
28CP0196	13-02-90	15-12-00	137	37	137
28DL0024	13-01-89	14-12-99	132	21	136
28DL0044	13-01-89	28-07-99	189	27	168
28DL0065	17-01-89	28-12-99	138	62	166
28DL0084	17-01-99	28-12-99	193	29	187
28DL0085	17-01-89	28-12-99	114	23	180
28DP0150	13-02-90	15-12-00	107	43	141
28DP7007	30-01-92	29-12-00	12	14	15

4.3 Beknopte beschrijving lineaire en niet-lineaire tijdreeksmodel

De Gd-methode maakt gebruik van een lineair stochastisch tijdreeksmodel voor de relatie tussen het neerslagoverschot en de grondwaterstand om de GxG op stambuislocaties te berekenen uit gesimuleerde grondwaterstandsreeksen van 30 jaar lang. Dit is nodig omdat de beschikbare tijdreeksen op de meeste locaties te kort zijn om een GxG te berekenen die representatief is voor de heersende klimatologische omstandigheden. Toevallige weersomstandigheden (een erg droog of nat jaar) hebben een te grote invloed op de GxG als deze rechtstreeks uit korte grondwaterstandsreeksen wordt berekend. In de Gd-methode wordt dit probleem verkleind door op basis van de (korte) tijdreeks de parameters van een tijdreeksmodel voor de relatie tussen neerslagoverschot en grondwaterstand te schatten, waarna met het tijdreeksmodel reeksen grondwaterstanden van 30 jaar lang wordt gesimuleerd, met als invoer de neerslag- en verdampingsgegevens over die 30 jaar. Er worden een groot aantal (100) mogelijke grondwaterstandsreeksen gesimuleerd (stochastische simulatie). De stochastische component van het tijdreeksmodel karakteriseert de modelfout. Door trekkingen uit deze modelfout te doen en steeds de bijbehorende GxG te berekenen wordt via de aldus toegepaste Monte-Carlomethode de kansverdeling van de werkelijke GxG gegenereerd (zie ook § 2.2.3). Het gemiddelde van de gesimuleerde GxG's gebruiken we vervolgens als predictie van de GxG op de stambuislocatie. De standaardafwijking van de gesimuleerde GxG's is een maat voor de onzekerheid over de werkelijke GxG ten gevolge van de modelfout.

In deze paragraaf geven we een beschrijving van het lineaire en niet-lineaire model. Beide modellen staan uitgebreider beschreven in Bierkens et al. (2002).

4.3.1 Het lineaire model

We beschouwen het volgende model:

$$h(t) = h^*(t) + n(t) \quad t = 1, \dots, T \quad (4.1a)$$

$$h^*(t) = \delta \cdot h^*(t-1) + \omega \cdot P^*(t) \quad t = 1, \dots, T \quad (4.1b)$$

$$n(t) - c = \phi(n(t-1) - c) + a(t) \quad t = 1, \dots, T \quad (4.1c)$$

waarbij t [dagen] de tijd, h [cm] de grondwaterstand ten opzichte van een zeker referentieniveau, h^* [cm] de deterministische component van de grondwaterstand, P^* [cm] het neerslagoverschot, n [cm] een autoregressieve ruiscomponent, a [cm] een normaal verdeeld witte ruisproces met gemiddelde nul [cm] en variantie σ_a^2 [cm²]. c [cm] is de gemiddelde grondwaterstand. De coëfficiënten δ [-] en ω [-] zijn gewichten die aan respectievelijk de vorige grondwaterstand en het huidige neerslagoverschot worden toegekend. De coëfficiënt ϕ [-] is de autoregressieve parameter van het ruismodel.

Om model (4.1) toe te kunnen passen zijn als invoer nodig de begintoestanden h_0^* en n_0 en een tijdreeks van P_t^* ($t=1, \dots, T$). Verder moeten de vijf parameters c , δ , ω , ϕ en σ_a^2 bekend zijn. Deze parameters worden geschat uit een gemeten tijdreeks van de grondwaterstand.

4.3.2 Het niet-lineaire model

De discrete-tijdvorm van het niet-lineaire model wordt gegeven door

$$h(t) = h(t-1) + f(h(t-1), t) + m(t) \quad t = 1, \dots, T \quad (4.2)$$

met

$$f(h,t) = \frac{1}{G(h)} [P(t) - E_a(S(h),t) + q_v(t) - q_d(h,t)] \quad (4.3a)$$

$$G(h) = \varepsilon_0 + (\theta_s - \theta_r) \cdot \left[1 - (1 + (\alpha(z_s - h))^{-n})^{-\frac{1}{n}} \right] \quad (4.3b)$$

$$E_a(S(h),t) = F_c \cdot E_r(t) \cdot (S(h))^c \quad (4.3c)$$

$$S(h) = \frac{1}{\alpha(z_s - h)} \cdot (1 + (\alpha(z_s - h))^{-n})^{-\frac{1}{n}} \quad (4.3d)$$

$$q_d(h,t) = \sum_{i=1}^{m_d} q_{d,i}(h,t) \quad (4.3e)$$

$$q_{d,i}(h,t) = \delta_{H_{s,i}(t) < h(t)} \cdot \frac{h(t) - H_{s,i}(t)}{\gamma_i} \quad (4.3f)$$

waarbij

P	=	netto neerslag, d.w.z. neerslag minus interceptieverdamping
q_v	=	verticale flux naar dieper grondwater
q_d	=	drainage naar oppervlaktewater
F_c	=	gewasfactor
E_a	=	actuele evapotranspiratie
E_r	=	Makkink referentieverdamping
α, n	=	van Genuchtenparameters
θ_r	=	residueel vochtgehalte
θ_s	=	verzadigd vochtgehalte
ε_0	=	residuele grondwaterberging
z_s	=	maaiveldhoogte ten opzichte van een gekozen referentieniveau
m_d	=	aantal oppervlaktewatersystemen (sloot of drainage)
$H_{s,i}$	=	hoogte van de bodem van het i -de watersysteem t.o.v. het referentieniveau
γ_i	=	drainageweerstand van het i -de watersysteem
$\delta_{conditie}$	=	1 als conditie waar, anders 0
m	=	een normaal verdeeld witte ruisproces met gemiddelde nul en variantie σ^2

Het niet-lineaire model is, onder zekere vereenvoudigende voorwaarden, afgeleid uit fysische beschrijvingen van het infiltratie- en drainageproces. Voor een gedetailleerde

uitleg verwijzen we naar Bierkens et al. (2002). De niet-lineaire vergelijking (4.2) houdt rekening met de grondwaterstandsafhankelijke waterberging in de onverzadigde zone en met grondwaterstandsafhankelijke laterale drainage.

Het niet-lineaire model is complexer dan het lineaire model en bevat meer parameters, waarvan de meeste een fysische betekenis hebben en in principe meetbaar zijn in het veld of op enigerlei wijze zijn af te leiden uit bodemkenmerken. Hoe in de praktijk aan de parameters van het niet-lineaire model te komen bespreken we in de volgende paragraaf.

4.4 Kalibratie van de tijdreeksmodellen op de stambuizenset

De programmaschil Vidente (Bierkens et al., 2002) bevat implementaties van de modellen die in de vorige paragraaf werden gepresenteerd. Dit zijn het programma KALTFN voor het lineaire en het programma SSD voor het niet-lineaire model. Beide programma's bestaan uit drie hoofdonderdelen, te weten een kalibratiedeel, een predictiedeel en een simulatiedeel. Daarnaast is er een module die een nabewerking op de uitvoer van de simulaties uitvoert en uit deze uitvoer samenvattende statistieken als gemiddelde en standaardafwijking van de GxG berekent. Vidente is een gebruikersvriendelijke, menugestuurde schil die de diverse instellingen van de programma's bewaart en grafische uitvoer genereert. De ervaren gebruiker kan ook direct de *stand-alone*-versies KALTFN en SSD aanroepen om zodoende meer controle en zicht te hebben over de uitgevoerde berekeningen.

4.4.1 Kalibratie van het lineaire model

De kalibratie van het lineaire model betreft de vijf parameters c , ω , δ , ϕ en σ_a^2 , die worden geschat op basis van de waargenomen tijdreeks van grondwaterstanden en meteorologische gegevens (neerslag en verdamping). De grondwaterstandsreeks is in dit geval afkomstig uit de DINO-databank, de meteorologische gegevens zijn afkomstig van het KNMI-weerstation dat het dichtst bij de betreffende buis ligt. Als optimalisatiecriterium bij de kalibratie wordt het criterium van Schweppe gebruikt, dat er naar streeft de (gekwadrateerde) verschillen tussen de waargenomen en voorspelde grondwaterstand zo klein mogelijk te maken. Kalibratie met KALTFN verloopt voor alle 19 buizen naar tevredenheid, hoewel er geen garantie is dat het numerieke optimalisatiealgoritme (de zogeheten amoebe-methode, zie Bierkens et al., 2002) niet in een lokaal minimum blijft steken. Ook predictie en simulatie met KALTFN verloopt naar wens en geeft geen bijzondere problemen. Ter illustratie zijn in Figuren 7 en 8 de waarnemingen en voorspellingen van de grondwaterstand voor twee van de 19 stambuizen als functie van de tijd uitgezet. De gekalibreerde parameterwaarden van het lineaire tijdreeksmodel zijn voor alle 19 buizen gegeven in Tabel 3.

Tabel 3 Gekalibreerde parameters van het lineaire model voor de 19 stambuizen.

Buisnummer	c [cm]	ω [-]	δ [-]	ϕ [-]	σ_a^2 [cm ²]
22CL0002	138.1	0.362	0.9899	0.9548	17.6
22CL0003	111.1	0.405	0.9783	0.9735	13.7
22CL0074	199.9	0.372	0.9942	0.9928	7.5
28AL0024	159.1	0.235	0.9887	0.9897	10.4
28AL0078	229.7	0.372	0.9888	0.9895	10.3
28AL0084	114.4	0.231	0.9915	0.9884	10.9
28AP0118	145.0	0.660	0.9728	0.9273	65.3
28AP7806	62.0	0.620	0.9722	0.9812	17.7
28BL0053	103.3	0.613	0.9651	0.9662	20.6
28CL0021	194.6	0.370	0.9906	0.9899	7.1
28CP0189	279.4	0.292	0.9942	0.9892	18.0
28CP0196	164.0	0.329	0.9912	0.9962	4.9
28DL0024	138.8	0.776	0.9148	0.9358	20.2
28DL0044	199.7	0.636	0.9557	0.9412	29.9
28DL0065	175.4	0.634	0.9882	0.9915	27.6
28DL0084	203.8	0.818	0.9430	0.9638	21.1
28DL0085	121.3	0.481	0.9571	0.9745	12.6
28DP0150	127.7	0.673	0.9773	0.9603	46.5
28DP7007	17.7	0.295	0.9580	0.9706	4.8

4.4.2 Kalibratie van het niet-lineaire model

Toepassing van het niet-lineaire tijdreeksmodel vereist, naast de meteorologische gegevens en een initiële grondwaterstand, specificatie van elf parameters. De maaiveldhoogte z_0 is afgeleid uit het AHN, terwijl de gewasfactor F_c voor alle 19 locaties op één is gesteld. Uit de profielbeschrijving bij elke buis worden uit het bodemtype met behulp van de Staringreeks de bodemfysische parameters θ_s , θ_v , α en n afgeleid. Uit gegevens van het waterschap over het oppervlaktewatersysteem wordt het drainageniveau H afgeleid (zie Tabel 2). Het drainageniveau wordt constant in de tijd verondersteld en gelijk gesteld aan de hoogte van het peil in de waterloop (bij een grondwaterstand lager dan het drainageniveau vindt geen drainage plaats, zie vergelijking 4.3f). De keuze van het oppervlaktewatersysteem waarop wordt gedraineerd bleek niet eenvoudig omdat een keuze gemaakt moest worden tussen ofwel nabijliggende greppels of verder weg liggende sloten (tertiaire waterlopen). Uiteindelijk is voor de tertiaire waterlopen gekozen, mede omdat de tijdreeksen van de grondwaterstanden doen vermoeden dat er hoegenaamd geen drainerende werking van de greppels uitgaat. De resterende parameters ε_0 , σ^2 , γ en q_v zijn alle gekalibreerd (zie Tabel 4), waarbij net als voor het lineaire model het criterium van Scheppe en de amoebemethode wordt gebruikt.

Hoewel kalibratie voor het niet-lineaire model een kleiner aantal parameters betreft bleek convergentie van het optimalisatiecriterium veel lastiger dan bij het lineaire model. Het succes van de kalibratie bleek sterk afhankelijk van de keuze van de initiële parameterwaarden. Voor de meeste buizen waren meerdere pogingen nodig om tot een bevredigende parameterset te komen. In twee gevallen (buisen 28AL0084 en 28DL0024) werd op basis van een visuele inspectie van de waargenomen en

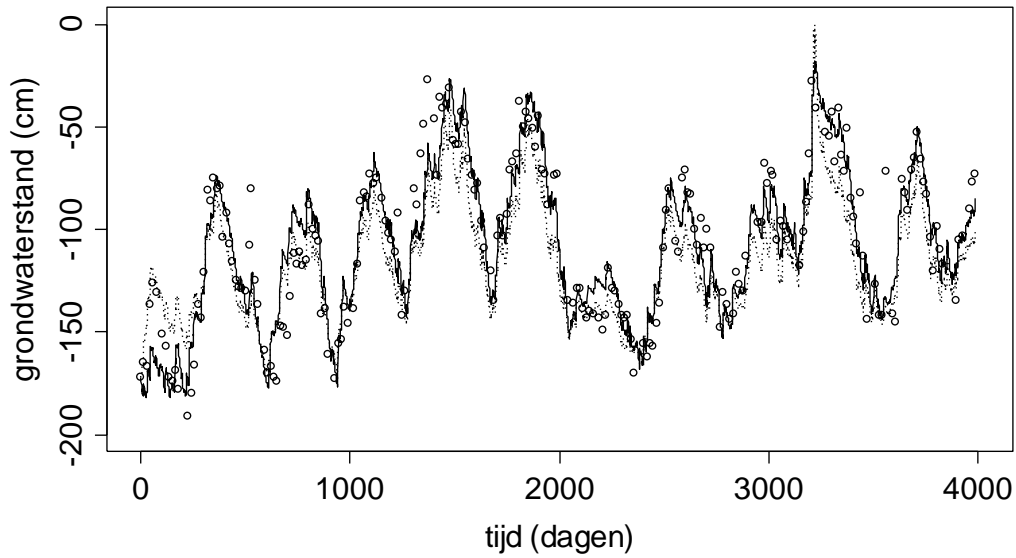
voorspelde grondwaterstanden de geoptimaliseerde parametercombinatie alsnog verworpen. Deze twee buizen zijn daarom in het vervolg van deze studie buiten beschouwing gelaten voor zowel het lineaire als niet-lineaire model. Ter illustratie zijn in Figuren 7 en 8 de voorspelde grondwaterstand voor twee locaties gegeven.

Een verklaring voor de moeizame kalibratie voor het niet-lineaire model kan zijn dat de twee parameters waar het systeem het gevoeligst voor is, te weten de verticale drainage q_v en de (laterale) drainageweerstand γ , elkaars invloed deels compenseren. De effecten van een verhoogde drainageweerstand kunnen grotendeels teniet gedaan worden door een grotere neerwaartse verticale flux te veronderstellen. Verlaging van de drainageweerstand geeft een grotere drainage naar het oppervlaktewater en deze kan gecompenseerd worden door de neerwaartse flux te laten afnemen of de opwaartse flux (kwel) te laten toenemen. Met andere woorden, uit tijdreeksen van de grondwaterstand alleen is het moeilijk de drainageweerstand en de verticale flux beide te kalibreren.

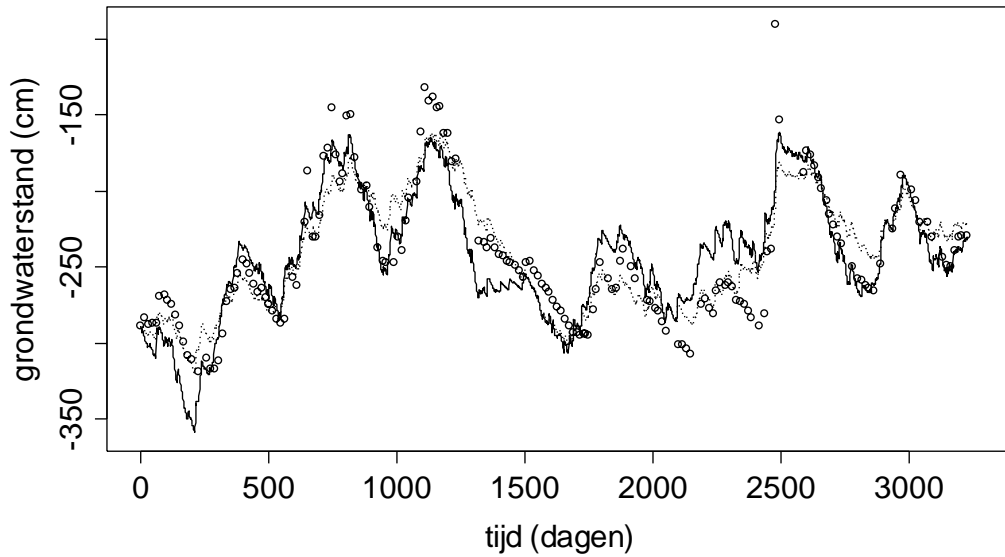
Samenvattend kan gesteld worden dat vastlegging van de diverse parameters van het niet-lineaire model meer kennis vraagt over de hydrologische en bodemkundige eigenschappen van de stambuislocatie en het gebied eromheen. Ook de parameterkalibratie verloopt veel minder soepel.

Tabel 4 Gekalibreerde parameterwaarden van het niet-lineaire model voor de 17 stambuizen.

Buisnummer	γ [dagen]	ε_0 [dm ³ /dm ³]	q_v [cm/dag]	σ^2 [cm ²]
22CL0002	340	0.057	1.06	85
22CL0003	118	0.007	-0.30	48
22CL0074	637	0.050	-0.90	45
28AL0024	79	0.100	-1.05	112
28AL0078	415	0.003	-0.20	74
28AP0118	400	0.000	1.00	188
28AP7806	283	0.090	1.52	61
28BL0053	43	0.084	-0.60	52
28CL0021	539	0.002	-0.02	40
28CP0189	1300	0.090	-0.50	184
28CP0196	399	0.090	-0.99	51
28DL0044	59	0.000	-0.71	167
28DL0065	500	0.010	-0.50	88
28DL0084	96	0.000	-0.65	156
28DL0085	223	0.003	1.83	22
28DP0150	300	0.002	0.00	127
28DP7007	40	0.031	1.19	54



Figuur 7 Waargenomen en voorspelde grondwaterstand voor stambuis 22CL0002. Cirkels zijn waarnemingen, de doorgetrokken lijn is de predictie met het lineaire model, de stippellijn is predictie met het niet-lineaire model. Dag 1 is 15 januari 1990



Figuur 8 Waargenomen en voorspelde grondwaterstand voor stambuis 28CP0189. Cirkels zijn waarnemingen, de doorgetrokken lijn is de predictie met het lineaire model, de stippellijn is predictie met het niet-lineaire model. Dag 1 is 14 januari 1992

4.5 Vergelijking van modelvoorspellingen van de GxG

Tabel 5 en 6 geven de gemiddelden en standaardafwijkingen van de GHG en GLG voor de 17 stambuizen zoals verkregen door simulatie, met het lineaire en het niet-lineaire model. De voorspelde GxG's (de gemiddelden van de simulaties) zijn ook tegen elkaar uitgezet in Figuur 9. Figuur 10 geeft de histogrammen van de verschillen. Uit de tabellen en figuren blijkt het volgende:

1. Het lineaire en niet-lineaire model voorspellen GxG's die behoorlijk van elkaar kunnen verschillen. Vaak zijn de verschillen meer dan 20 cm groot, in sommige gevallen zelfs meer dan 40 cm, met uitschieters tot zelfs 100 cm (alleen voor de GLG). Blijkbaar maakt het wel degelijk uit of het lineaire dan wel het niet-lineaire model wordt gebruikt om de GxG op stambuislocaties te voorspellen. Verwacht mag worden dat dit ook invloed zal hebben op de kaarten van de GxG die met de Gd-methode worden berekend.
2. Er is een duidelijk systematisch verschil tussen voorspellingen van de GxG met het lineaire en niet-lineaire model. Het niet-lineaire model geeft een GxG die gemiddeld zo'n 20 tot 30 cm lager ligt. De grootte van dit effect is voor GHG en GLG ongeveer even groot. Het treedt zowel op bij buizen die in natte als in droge delen van het onderzoeksgebied liggen, hoewel het effect voor natte gebieden minder geprononceerd is. Een duidelijke verklaring voor de systematische verschillen is niet gemakkelijk te geven, maar zou kunnen liggen in het feit dat de stochastische component van het niet-lineaire model een grotere variantie heeft. Dit leidt tot grilliger simulaties (hetgeen overigens niet zichtbaar is in de voorspelde GxG in Figuren 7 en 8), waardoor vaker extremen worden gesimuleerd. Dit kan goed verklaren waarom de GLG voor het niet-lineaire model systematisch lager is, maar voor de GHG is dit geen verklaring. Nadere analyse van gesimuleerde tijdreeksen kan wellicht uitsluitsel geven. Merk overigens op dat systematische verschillen een grotere invloed op de GxG kaarten zullen hebben dan toevallige verschillen. Dit is omdat systematische verschillen niet worden uitgemiddeld in de krigingstap van de Gd-methode.
3. Voor sommige buizen zijn de verschillen tussen de voorspelde GxG voor de twee modellen klein, terwijl voor andere buizen de verschillen juist groot zijn. Het lijkt er op dat de verschillen vooral daar groot zijn waar de variantie van het residu van het niet-lineaire model groot is. In die gevallen geeft het niet-lineaire model beduidend grotere voorspellingen. Dit versterkt het onder het punt 2 geuite vermoeden dat systematische verschillen mogelijk veroorzaakt zijn door de toename van het aantal extremen bij het niet-lineaire model, hoewel je voor de GLG juist een tegenovergesteld effect zou verwachten.
4. De waargenomen verschillen tussen het lineaire en niet-lineaire model lijken grotendeels verklaard te kunnen worden door de modelonzekerheid van de twee modellen, zoals vervat in de standaardafwijkingen gepresenteerd in Tabellen 5 en 6. In de meeste gevallen liggen de verschillen in de orde van grootte van één à twee standaardafwijkingen. Hierop zijn wel een aantal uitzonderingen (neem als voorbeeld de GHG van buis 28DP0150, waar het verschil 26.1 cm bedraagt en de standaardafwijkingen 2.8 en 4.3 cm), maar

dergelijke uitschieters mogen wellicht verwacht worden bij het maken van $2 \cdot 17=34$ vergelijkingen.

5. De standaardafwijkingen voor het niet-lineaire model zijn aanzienlijk groter dan die voor het lineaire model. Daarnaast zijn de verschillen tussen de buizen onderling ook groter. Dit is een belangrijke constatering, want de standaardafwijking geeft weer hoe goed het model de werkelijkheid beschrijft (hoe groter de standaardafwijking, hoe slechter de beschrijving). Verwacht mocht worden dat het niet-lineaire model de werkelijkheid beter zou beschrijven omdat het meer rekening houdt met de fysische processen. Dit blijkt echter niet op te treden.

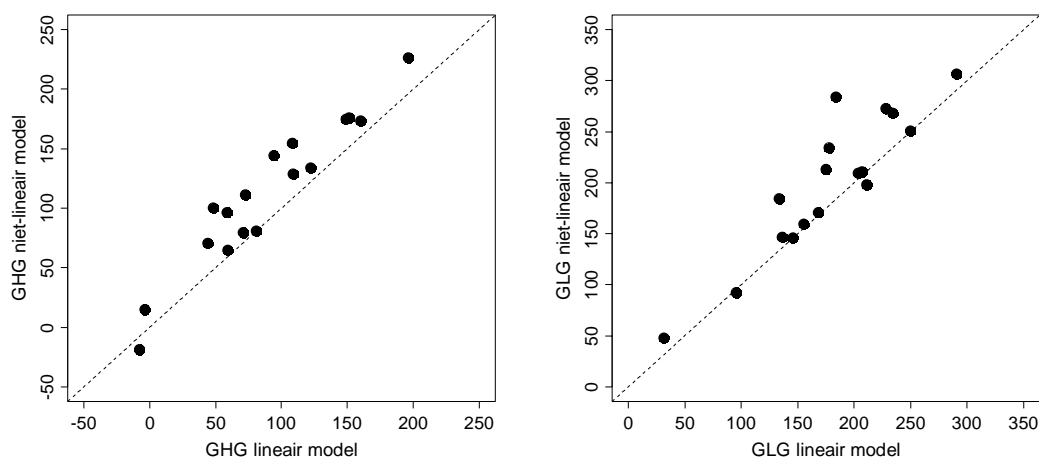
Het bovenstaande doet ernstige twijfels rijzen over de kwaliteit van het niet-lineaire model in deze toepassing. Deze twijfels zijn echter gebaseerd op de 'berekende' nauwkeurigheid in de modeluitvoer, welke niet noodzakelijkerwijs gelijk zijn aan de waargenomen nauwkeurigheid in de modeluitvoer. Immers, het lineaire of niet-lineaire model zou haar voorspelnauwkeurigheid kunnen onder- of overschatten. Een vergelijking van de berekende GxG's met de werkelijke GxG kan hier uitsluitsel over geven. We gaan hier in de volgende paragraaf dieper op in.

Tabel 5 Gemiddelde en standaardafwijking van de gesimuleerde GHG voor de 17 stambuizen

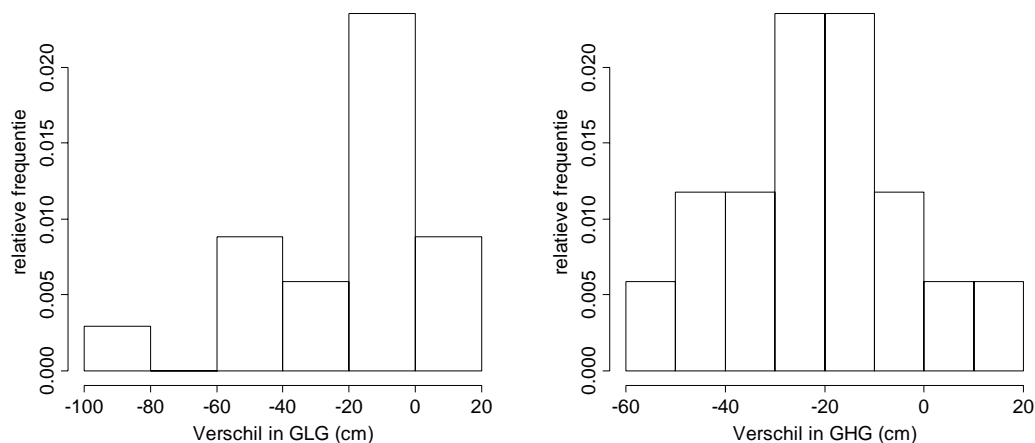
Buisnummer	lineaire model		niet-lineaire model	
	gemiddelde	standaard-afwijking	Gemiddelde	standaard-afwijking
22CL0002	71.3	1.6	79.4	4.0
22CL0003	59.7	2.0	64.5	2.9
22CL0074	109.0	3.8	128.4	16.2
28AL0024	108.2	3.3	154.6	33.5
28AL0078	160.1	3.3	173.1	4.5
28AP0118	72.8	2.2	110.8	10.0
28AP7806	-3.2	2.8	14.8	1.5
28BL0053	48.2	2.0	99.7	19.3
28CL0021	122.2	2.9	133.4	4.0
28CP0189	196.3	4.2	226.2	15.1
28CP0196	94.4	5.4	143.7	18.0
28DL0044	149.4	1.6	174.1	14.9
28DL0065	59.1	6.4	95.9	8.5
28DL0084	151.4	2.0	175.8	11.0
28DL0085	81.3	1.9	80.5	1.7
28DP0150	44.1	2.8	70.2	4.3
28DP7007	-7.7	1.0	-18.6	0.3

Tabel 6 Gemiddelde en standaardafwijking van de gesimuleerde GLG voor de 17 stambuizen

buisnummer	lineaire model		niet-lineaire model	
	gemiddelde	standaard-afwijking	gemiddelde	standaard-afwijking
22CL0002	155.4	1.7	159.0	4.9
22CL0003	136.5	2.0	146.4	3.8
22CL0074	204.1	4.0	209.1	14.6
28AL0024	178.2	3.5	234.0	32.2
28AL0078	250.1	3.5	250.3	4.8
28AP0118	184.0	2.2	283.5	20.9
28AP7806	95.7	3.0	92.1	2.4
28BL0053	134.2	2.1	183.7	18.1
28CL0021	207.6	3.1	210.7	3.5
28CP0189	291.1	4.4	306.5	15.0
28CP0196	175.5	5.4	212.6	17.1
28DL0044	228.6	1.8	272.4	18.2
28DL0065	211.1	6.7	197.5	8.7
28DL0084	234.7	2.1	267.7	14.4
28DL0085	146.0	2.0	145.9	1.3
28DP0150	168.7	2.9	170.6	6.4
28DP7007	32.0	1.1	47.3	3.7



Figuur 9 Scatter plots van de GHG (links) en GLG (rechts) voor de 17 stambuizen. Voorspellingen volgens het lineaire model uitgezet op de x-as, voorspellingen volgens het niet-lineaire model op de y-as



Figuur 10 Histogrammen van de verschillen in voorspelde GHG (links) en GLG (rechts). Verschillen zijn gedefinieerd als predictie volgens het lineaire model minus predictie volgens het niet-lineaire model

4.6 Vergelijking van modelvoorspellingen met de werkelijkheid

De belangrijkste vraag die we in dit hoofdstuk willen beantwoorden is welk model de werkelijke GxG het best beschrijft. In dit geval verkeren we in de gunstige omstandigheid dat de gemeten tijdreeksen van de grondwaterstanden gemiddeld zo'n acht jaar beslaan, hetgeen volgens de definitie van de GxG voldoende lang is om ze te kunnen berekenen. Uitzondering hierop is overigens buis 28DP7007, die wel over acht jaar gemeten is maar waarvoor berekening van de GxG niet mogelijk is vanwege de vele hiaten in de tijdreeks. Met andere woorden, we 'kennen' de GxG voor 16 stambuizen in het gebied. Hierbij dient wel opgemerkt dat de GxG bij voorkeur over een langere periode berekend dient te worden, om het effect van toevallige weersinvloeden te reduceren. Vanwege die invloeden spreken we in dit hoofdstuk daarom over de 'waargenomen' GxG in plaats van de 'werkelijke' GxG. Daarnaast is het zo dat het lineaire en niet-lineaire model beide gekalibreerd zijn op dezelfde tijdreeks als waarop de waargenomen GxG is gebaseerd. Dit maakt dat de voorspelde GxG waarschijnlijk dichter bij de waargenomen GxG zit dan wanneer de modellen op een andere tijdreeks gekalibreerd zouden zijn. Beide modellen maken gebruik van dit voordeel. Tenslotte moet vermeld dat we hier over slechts 16 buizen beschikken, zodat niet uitgesloten mag worden dat eventuele waargenomen verschillen op toeval berusten en niet statistisch significant zijn. Ondanks deze kanttekeningen is het toch de moeite waard te bepalen hoe de twee modellen de vergelijking met de waargenomen GxG op de 16 locaties doorstaan.

Tabellen 7 en 8 geven de waargenomen GxG naast de voorspelde GxG volgens de twee modellen en de verschillen tussen de waargenomen en voorspelde GxG. Figuren 11 en 12 geven *scatter plots* van voorspelde GxG uitgezet tegen waargenomen GxG.

De uitkomsten laten zien dat het lineaire model veel dichter bij de waargenomen GxG zit dan het niet-lineaire model. Voor de GHG heeft het niet-lineaire model een

duidelijke positieve systematische afwijking. Voor de GLG hebben beide modellen een positieve systematische afwijking waarbij de afwijking van het lineaire model veel geringer is dan die van het niet-lineaire model. Naast systematische verschillen zijn ook de toevallige verschillen voor het niet-lineaire model groter, met name voor de GLG.

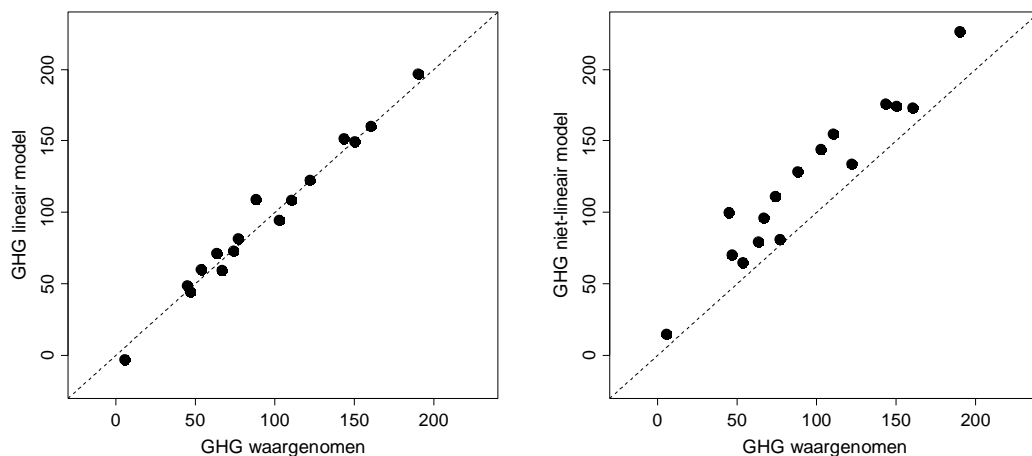
Een mogelijke verklaring voor de systematische verschillen zou kunnen zijn dat de voorspelde GxG is berekend over een periode van 30 jaar en de waargenomen GxG over een periode van acht jaar. De achtjarige periode zou gemiddeld genomen droger kunnen zijn waardoor de waargenomen GxG groter is dan voorspeld. Uit de meteorologische gegevens blijkt echter dat de periode van acht jaar (potentiële neerslagoverschot gemiddeld 0.91 mm/dag) juist iets natter was dan de periode van 30 jaar (potentiële neerslagoverschot gemiddeld 0.74 mm/dag), waarvan het effect wordt ingeschat als een toename in de grondwaterstand in de orde grootte van twee tot tien cm.

Tabel 7 *Vergelijking van waargenomen GHG en voorspelde GHG volgens het lineaire en niet-lineaire model*

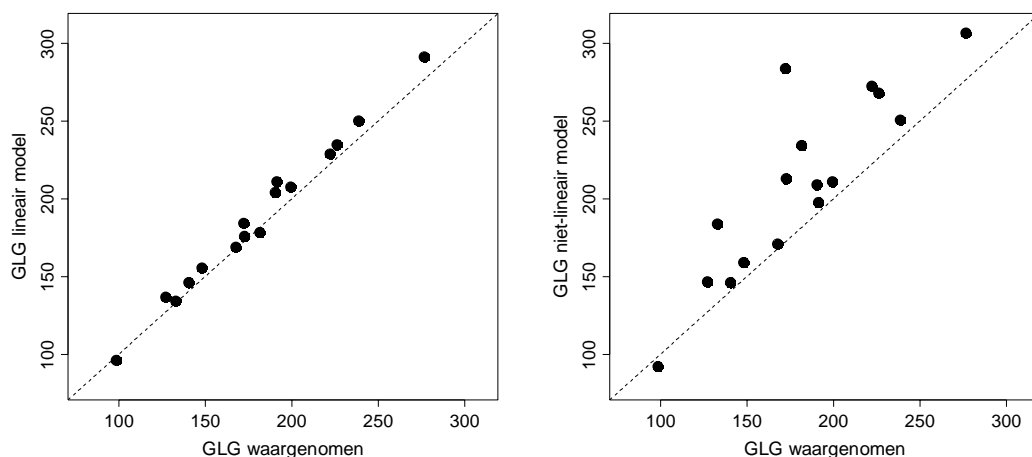
Buisnummer	GHG waargenomen	GHG lineair	GHG niet-lineair	Vershil lineair	Vershil niet-lineair
22CL0002	63.5	71.3	79.4	-7.8	-15.9
22CL0003	53.7	59.7	64.5	-6.0	-10.8
22CL0074	88.1	109.0	128.4	-20.9	-40.3
28AL0024	110.8	108.2	154.6	2.6	-43.8
28AL0078	160.5	160.1	173.1	0.4	-12.6
28AP0118	74.0	72.8	110.8	1.2	-36.8
28AP7806	5.9	-3.2	14.8	9.1	-8.9
28BL0053	44.9	48.2	99.7	-3.3	-54.8
28CL0021	122.2	122.2	133.4	0.0	-11.2
28CP0189	190.1	196.3	226.2	-6.2	-36.1
28CP0196	102.8	94.4	143.7	8.4	-40.9
28DL0044	150.2	149.4	174.1	0.8	-23.9
28DL0065	66.9	59.1	95.9	7.8	-29.0
28DL0084	143.5	151.4	175.8	-7.9	-32.3
28DL0085	77.2	81.3	80.5	-4.1	-3.3
28DP0150	47.1	44.1	70.2	3.0	-23.1

Tabel 8 *Vergelijking van waargenomen GLG en voorspelde GLG volgens het lineaire en niet-lineaire model*

Buisnummer	GLG waargenomen	GLG lineair	GLG niet-lineair	Vershil lineair	Vershil niet-lineair
22CL0002	148.4	155.4	159.0	-7.0	-10.6
22CL0003	127.5	136.5	146.4	-9.0	-18.9
22CL0074	190.9	204.1	209.1	-13.2	-18.2
28AL0024	181.8	178.2	234.0	3.6	-52.2
28AL0078	239.0	250.1	250.3	-11.1	-11.3
28AP0118	172.5	184.0	283.5	-11.5	-111.0
28AP7806	98.6	95.7	92.1	2.9	6.5
28BL0053	133.2	134.2	183.7	-1.0	-50.5
28CL0021	199.9	207.6	210.7	-7.7	-10.8
28CP0189	276.8	291.1	306.5	-14.3	-29.7
28CP0196	172.7	175.5	212.6	-2.8	-39.9
28DL0044	222.3	228.6	272.4	-6.3	-50.1
28DL0065	191.5	211.1	197.5	-19.6	-6.0
28DL0084	226.5	234.7	267.7	-8.2	-41.2
28DL0085	140.9	146.0	145.9	-5.1	-5.0
28DP0150	168.0	168.7	170.6	-0.7	-2.6



Figuur 11 Scatter plots van de waargenomen en voorspelde GHG voor 16 stambuizen. Links lineair model, rechts niet-lineair model.



Figuur 12 Scatter plots van de waargenomen en voorspelde GLG voor 16 stambuizen. Links lineair model, rechts niet-lineair model

4.7 Invloed van modelkeuze op resultaten van de stambuisregressie

Tot nu toe is in dit experiment onderzocht hoe goed het lineaire en niet-lineaire tijdreeksmodel in staat zijn de GxG op stambuislocaties te voorspellen. In de Gd-methode volgen echter nog diverse stappen om vanuit GxG's op stambuislocaties tot een gebiedsdekkende GxG te komen. De belangrijkste stappen hierbij zijn de stambuisregressie, de maaiveldregressie en de krigingstap. In dit hoofdstuk doorlopen we niet dit hele traject maar bekijken alleen hoe de stambuisregressie beïnvloed wordt door de keuze tussen het lineaire en niet-lineaire model.

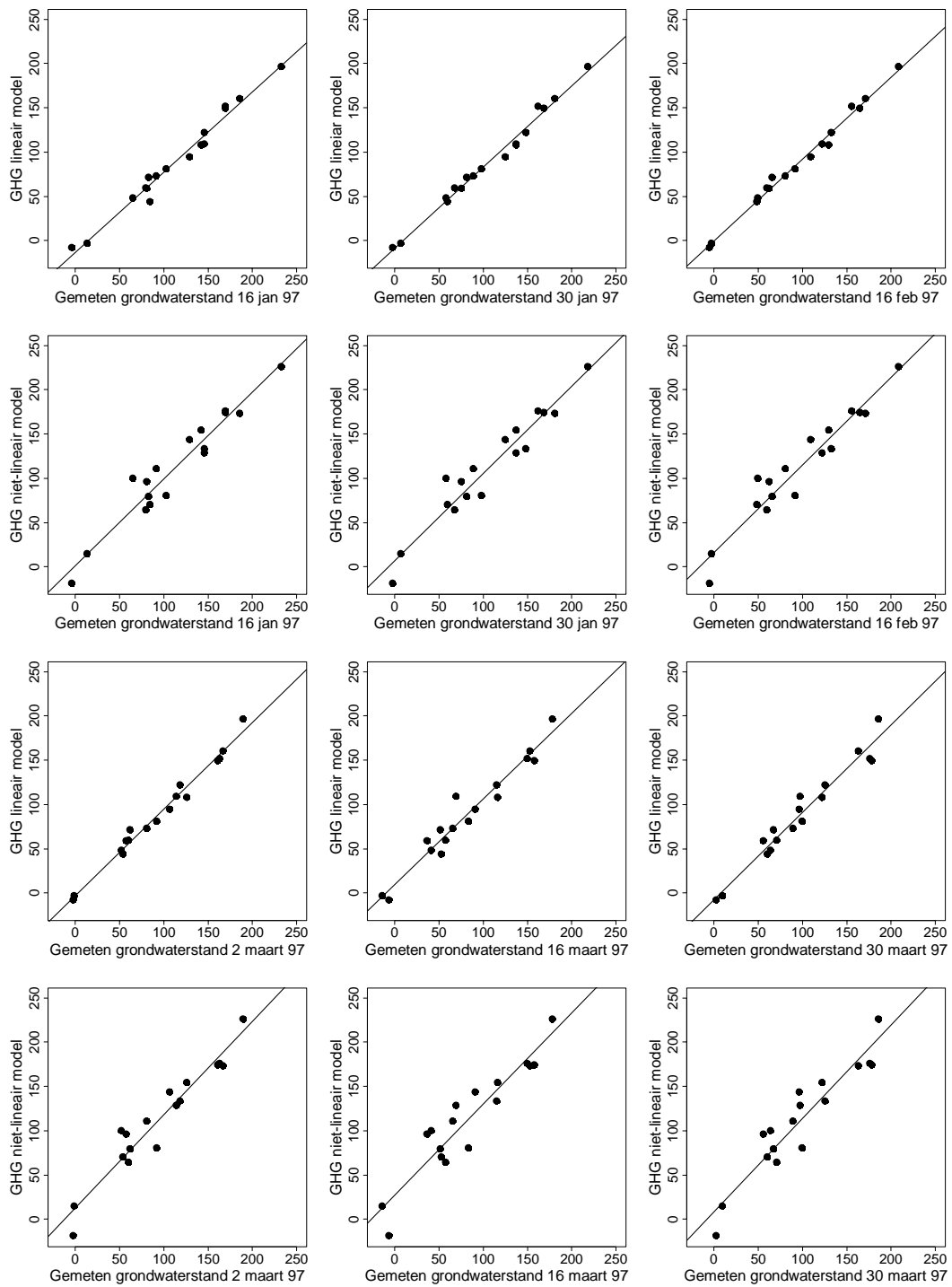
De stambuisregressie voorspelt de GxG op basis van een gerichte opname van de grondwaterstand (zie § 2.4). Het regressiemodel wordt gefit op de stambuisgegevens en toegepast op de locaties waar gerichte opnames zijn gedaan. De stambuislocaties

zijn immers de enige locaties waar de GxG bekend is, namelijk door voorspelling met het lineaire of niet-lineaire tijdreeksmodel. In § 4.4 is duidelijk geworden dat de twee modellen verschillende GxG's opleveren. Dit betekent dat ook de gefitte regressielijn en de voorspelkracht van de stambuisregressie waarschijnlijk zullen verschillen voor de twee modellen. In deze paragraaf kwantificeren we deze verschillen.

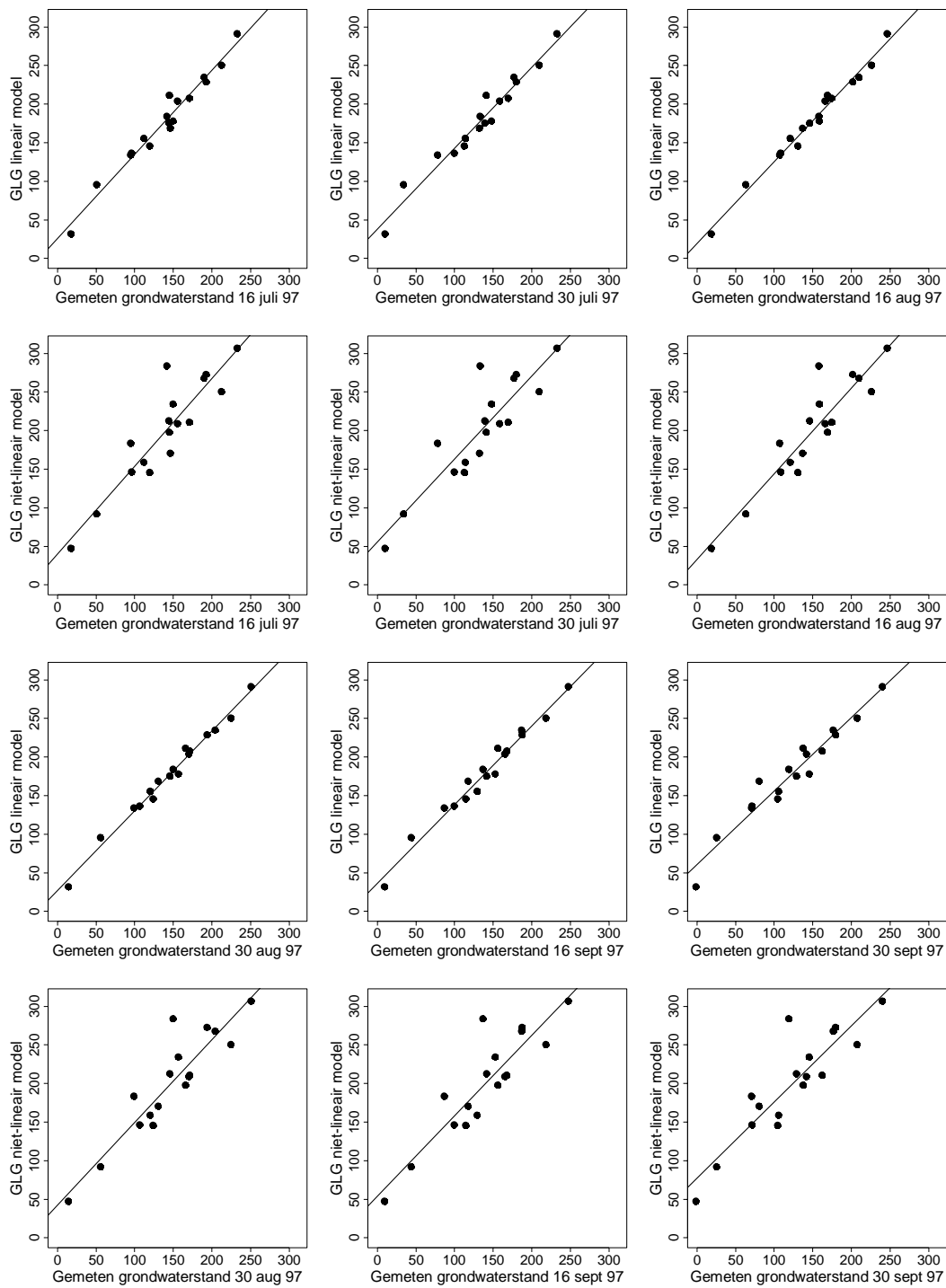
Figuren 13 en 14 geven de gefitte regressielijnen. Figuur 15 geeft de *scatter plots* van de geschatte regressiecoëfficiënten, inclusief standaardfouten. Figuur 16 geeft *box plots* van de *goodness-of-fit* van de stambuisregressies. Uit de figuren is het volgende af te leiden:

1. De regressies horend bij het niet-lineaire model vertonen een grotere spreiding rondom de regressielijn. De grotere spreiding rondom de lijn wordt bevestigd door de lagere *goodness-of-fit* en leidt tot grotere standaardafwijkingen van de geschatte regressiecoëfficiënten. De *goodness-of-fit* is voor het niet-lineaire model fors lager dan voor het lineaire model;
2. De gemeten grondwaterstand is voor alle zes zomerdagen lager dan de voorspelde GLG, zowel voor het lineaire als niet-lineaire model. Dit vertaalt zich in positieve *intercepts*. Ook de gemeten grondwaterstand op de zes winterdagen is steeds lager dan de voorspelde GHG, hoewel het effect minder groot is voor het lineaire model dan voor het niet-lineaire model. Deze systematische verschillen hebben overigens geen nadelige invloed op de kwaliteit van de regressie;
3. Er zijn ook verschillen in de regressies tussen de meetdagen (zie § 5.4). Dit is met name zichtbaar in het intercept van de regressies, die een behoorlijk grote spreiding te zien geven (Figuur 15). Ook de richtingscoëfficiënt is verschillend voor verschillende dagen, met waarden variërend van 0.90 tot 1.05 voor de GHG en van 0.95 tot 1.15 voor de GLG. De regressiecoëfficiënten behorend bij het lineaire dan wel niet-lineaire model zijn in Figuur 15 als twee enigszins overlappende groepen herkenbaar. Dit betekent dat naast dagkeuze ook modelkeuze invloed heeft op de geschatte regressiecoëfficiënten. Verschillen in regressiecoëfficiënten hebben een directe invloed op de voorspelde GxG op locaties waar gerichte opnames zijn gedaan, en als gevolg daarvan ook op de GxG-kaarten zoals berekend met de Gd-methode;
4. De invloed van de dagkeuze op de *goodness-of-fit* is veel minder groot dan de invloed van modelkeuze. Dit blijkt uit Figuur 16, die zowel voor de GHG als de GLG een veel grotere *goodness-of-fit* geeft voor het lineaire model. De GHG-regressies hebben een iets hogere R^2 dan de GLG-regressies. Deze figuur geeft aan dat de variantie van het residu van de stambuisregressie voor het niet-lineaire model fors groter is dan voor het lineaire model. Berekeningen laten zien dat deze gemiddeld genomen acht keer zo groot is.

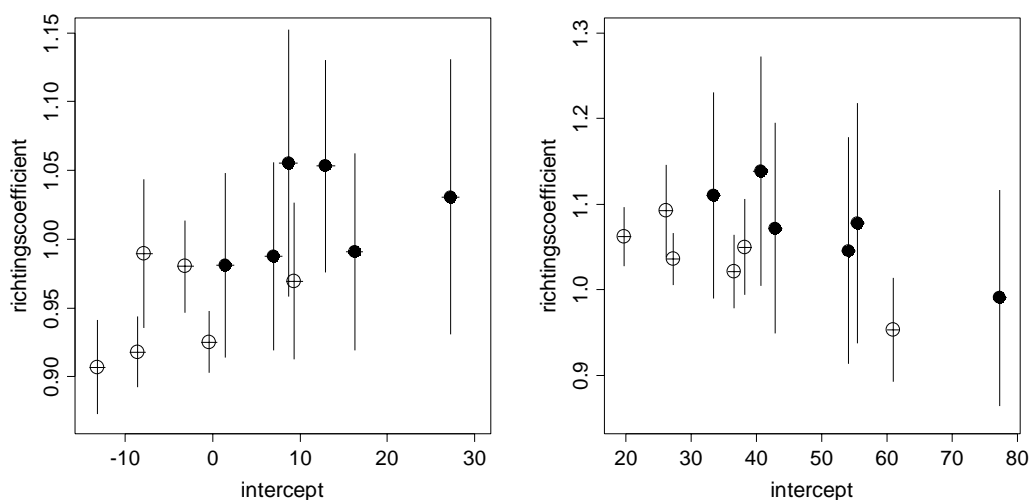
De resultaten van de stambuisregressie bevestigen de uitkomsten van de voorgaande paragrafen: de resultaten verkregen met het niet-lineaire model zijn beduidend slechter dan die verkregen met het lineaire model.



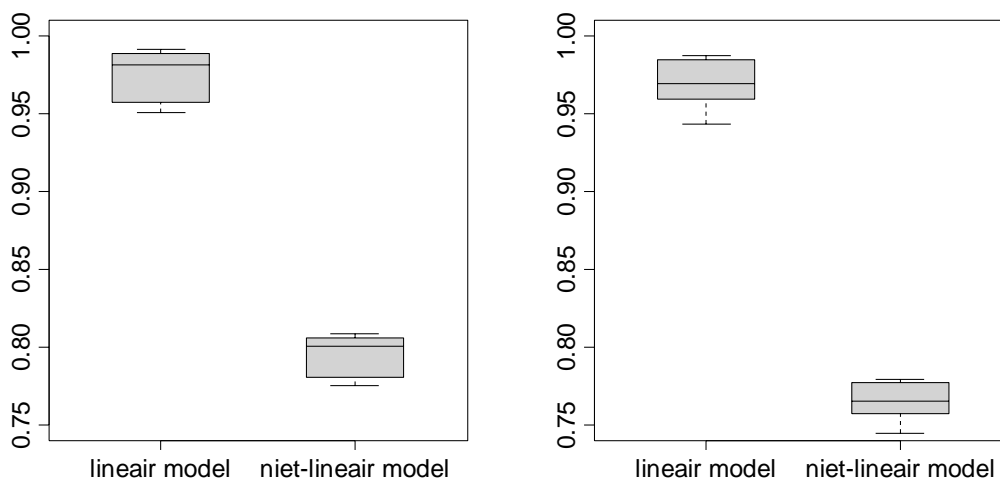
Figuur 13 Stambuisregressies die GHG voorspellen op basis van gemeten grondwaterstand op zes winterdagen in 1997. De eerste en derde rij betreffen het lineaire model, de tweede en vierde rij het niet-lineaire model. Dichte cirkels zijn de waarnemingen



Figuur 14 Stambuisregressies die GLG voorspellen op basis van gemeten grondwaterstand op zes zomerdagen in 1997. De eerste en derde rij betreffen het lineaire model, de tweede en vierde rij het niet-lineaire model. Dichte cirkels zijn de waarnemingen



Figuur 15 Scatter plots van de geschatte regressiecoëfficiënten voor de zes dagen waarop het regressiemodel is geschat. GHG links, GLG rechts. Open cirkels betreffen het lineaire tijdreeksmodel, gesloten cirkels het niet-lineaire tijdreeksmodel. De verticale en horizontale lijnen representeren \pm de standaard fout in de geschatte regressiecoëfficiënt



Figuur 16 Box plots van de goodness-of-fit (R^2) van de stambuisregressie verkregen met het lineaire en niet-lineaire tijdreeksmodel. GHG links, GLG rechts

4.8 Discussie

Dit hoofdstuk heeft aan de hand van een toepassing in het waterschap Regge en Dinkel laten zien dat het behoorlijk uitmaakt of je het lineaire dan wel het niet-lineaire tijdreeksmodel gebruikt om de GxG op stambuislocaties te voorspellen. De verschillen in voorspelde GxG tussen de twee modellen zijn groot en zeker niet verwaarloosbaar. De verschillen werken ook door naar de uitkomsten van de stambuisregressie. Op haar beurt zullen deze verschillen doorwerken in de GxG-kaarten die met de Gd-methode worden geproduceerd.

De toepassing in dit hoofdstuk laat zien dat het lineaire model het op alle fronten een stuk beter doet dan het niet-lineaire model. Deze bevinding staat haaks op de vooraf geuite verwachtingen. De verwachting was immers dat het niet-lineaire model beter zou presteren omdat het in staat is de werkelijke fysische processen beter te representeren. Een verklaring voor dit onverwachte resultaat is niet gemakkelijk te geven, maar aannemelijk lijkt dat de parametrisatie van het niet-lineaire model hier grotendeels debet aan is. Het niet-lineaire model vereist meer kennis over de (bodem)fysische en hydrologische eigenschappen van het beschouwde systeem. Deze kennis was voor het onderzoeksgebied onvoldoende aanwezig. Het was bijvoorbeeld onduidelijk welke waterloop te kiezen als belangrijkste drainerend systeem. Ook de kalibratie van het niet-lineaire model bleek een stuk lastiger en onbevredigender te verlopen dan bij het lineaire model. Bij gebrek aan voldoende informatie over de belangrijkste fysische processen heeft het blijkbaar weinig zin een model toe te passen dat die fysische processen realistischer meeneemt. Het niet-lineaire model presteert dan zelfs slechter dan het lineaire model. Verkeerd gekozen of verkeerd gekalibreerde parameters kunnen veel schade aanrichten en deze schade wordt onvoldoende gecompenseerd door de beter bij de werkelijkheid aansluitende modelstructuur.

Een andere oorzaak voor de slechtere resultaten van het niet-lineaire model zou kunnen liggen in het simpele feit dat de implementatie van het model niet geheel foutloos is. In de eerste plaats is het niet-lineaire model complexer dan het lineaire model en dit doet het risico op fouten in de programmatuur toenemen. Daarnaast is de implementatie van het lineaire model veel uitgebreider getest dan die van het niet-lineaire model. Het lineaire model is een vast onderdeel van de bestaande Gd-methode en is al talloze malen toegepast en de uitkomsten ervan kritisch onderzocht. Voor het niet-lineaire model geldt dit in veel mindere mate en dus zijn er minder mogelijkheden geweest om eventuele fouten in de programmatuur te verwijderen.

De slechtere prestatie van het niet-lineaire model kan al in een vroeg stadium worden onderkend. De fout in de modelvoorspellingen wordt namelijk gekwantificeerd middels de spreiding in de gesimuleerde GxG. In een gegeven situatie zou dus voor elke buis zowel het lineaire als niet lineaire model gefit kunnen worden en op basis van de spreiding in de gesimuleerde GxG's kan dan bepaald worden welk model in die situatie beter voldoet. Mogelijkerwijs kan het zelfs zo zijn dat binnen eenzelfde gebied sommige buizen met het lineaire en andere met het niet-lineaire model worden beschreven. Kanttekening hierbij is wel dat dit alles veronderstelt dat de stochastische parameters van de modellen correct zijn gekalibreerd.

Enige voorzichtigheid bij het interpreteren van de resultaten van dit hoofdstuk is geboden. De resultaten hebben strikt genomen immers alleen betrekking op de gekozen toepassing. Deze toepassing wijkt echter niet af van wat in het algemeen wordt gebruikt en het is dan ook niet onaannemelijk dat vergelijkbare uitkomsten voor andere gebieden zullen worden verkregen.

4.9 Conclusies

De belangrijkste conclusies van dit hoofdstuk staan hieronder puntsgewijs gegeven. Kanttekening hierbij is dat ze zijn gebaseerd op één, overigens als representatief te beschouwen, onderzoek. Dit maakt de conclusies niet ondeugdelijk, maar enige voorzichtigheid bij de interpretatie ervan is geboden.

- de invloed van het gebruik van een lineair dan wel niet-lineair tijdreeksmodel op de uitkomsten van de Gd-methode zijn niet verwaarloosbaar;
- het lineaire tijdreeksmodel is gemakkelijker en minder tijdrovend in het gebruik dan het niet-lineaire model;
- voor het in dit hoofdstuk uitgevoerde onderzoek is het lineaire model aanmerkelijk beter in staat de GxG te voorspellen dan het niet-lineaire model;
- de meest logische verklaring voor de matige prestatie van het niet-lineaire model is het gebrek aan fysische kennis over het onderhavige systeem om de parameters van het niet-lineaire model naar behoren te kunnen schatten;
- het niet-lineaire model zou alleen dan moeten worden toegepast wanneer voldoende fysische kennis over het systeem beschikbaar is. Enige scepsis over de praktische haalbaarheid hiervan is op zijn plaats gezien de kosten gemoeid met de verkrijging van fysische kennis en de verhoudingsgewijs veel hogere nauwkeurigheid van het lineaire model bij het huidige gebruik;
- er is geen reden het lineaire tijdreeksmodel in de Gd-methode te vervangen door een niet-lineair tijdreeksmodel. Bij toepassingen met geringe kennis over fysische aspecten van het hydrologische systeem moet dit zelfs worden afgeraden.

5 Effect van het meettijdstip en de locatie van stambuizen op de kwaliteit van met stambuisregressie voorspelde GxG

5.1 Inleiding

Nadat, op basis van een gekalibreerd tijdreeksmodel, de GxG voor een set van stambuizen is bepaald, wordt een regressiemodel gefit met deze GxG als responsvariabele en de gemeten grondwaterstand als predictorvariabele. Dit stambuisregressiemodel wordt vervolgens gebruikt om op een groot aantal locaties, waar op dezelfde dag de grondwaterstand is gemeten, de GxG te voorspellen. In dit experiment is onderzocht wat het effect is van het meettijdstip (meetdag) en het effect van de locaties van de stambuizen op de kwaliteit van regressievoorspellingen van de GxG.

5.2 Gebruikte data

Voor dit onderzoek is gebruik gemaakt van 102 stambuizen gelegen in het waterschap Regge en Dinkel. Alle stambuizen zijn in het veld op ligging gecontroleerd. Vervolgens is met de methode die is beschreven in § 2.3.3 een lineair tijdreeksmodel gekalibreerd, waarmee op basis van neerslaggegevens een tijdreeks van 30 jaar is gesimuleerd. Deze gesimuleerde tijdreeks van grondwaterstanden is gebruikt om de klimaatsrepresentatieve GxG van deze stambuizen te bepalen.

5.3 Onderzoeksopzet

5.3.1 Meettijdstippen

Het effect van de meetdag op de nauwkeurigheid van de regressievoorspellingen van de GxG is onderzocht voor zes dagen in het voorjaar van 1997, te weten 16 januari, 30 januari, 16 februari, 28 februari, 16 maart en 30 maart, en voor zes dagen in de herfst van 1997, te weten 16 juli, 30 juli, 16 augustus, 30 augustus, 16 september en 30 september. Omdat niet in alle buizen precies op deze dagen een grondwaterstand is gemeten, is een schatting van de grondwaterstand gemaakt op basis van metingen voorafgaand aan de beoogde dag en een predictie met het gekalibreerde tijdreeksmodel.

5.3.2 Steekproefopzetten

In de praktijk wordt het regressiemodel met de GxG als responsvariabele en de waargenomen grondwaterstand als predictorvariabele gefit op de gegevens van ongeveer 20 stambuizen. In het studiegebied zijn 102 stambuizen beschikbaar, en dit biedt de mogelijkheid onderzoek te doen naar het effect van de locatie van de

kalibratiestambuizen in de geografische ruimte en eigenschappenruimte op de nauwkeurigheid van de regressievoorspellingen. Dit is onderzocht door een groot aantal keren volgens een bepaalde steekproefopzet 20 stambuizen uit de set van 102 te selecteren, waarmee vervolgens het model is gefit (gekalibreerd). Op het hoogste niveau kan onderscheid gemaakt worden tussen selectie door middel van loting (kanssteekproefname, *probability sampling*) en gerichte selectie gebaseerd op een model. Kanssteekproeven zijn in dit onderzoek volgens twee steekproefstrategieën uitgevoerd, te weten enkelvoudig aselekt zonder teruglegging (SI), en gestratificeerd, enkelvoudig aselekt zonder teruglegging, d.w.z. volledig aselekt binnen strata (STSI). In geval van STSI zijn de volgende stratificaties onderzocht:

1. stratificatie op basis van de Gt volgens de Bodemkaart van Nederland 1 : 50 000 (STSI(Gt));
2. stratificatie op basis van de gemiddelde grondwaterstand in stambuizen (STSI(h));
3. stratificatie op basis van geografische ligging (STSI(x)).

Het aantal strata is in alle gevallen vier. De vier Gt-strata zijn gevormd door de oorspronkelijke Gt's van de 102 stambuizen te groeperen (zie Tabel 9). De stratificatie naar gemiddelde grondwaterstand en geografische ligging is uitgevoerd door clusteranalyse van de 102 buizen met behulp van *k-means* (Brus et al., 2003). Het totaal aantal buizen per stratum is verschillend (Tabel 10). Het aantal gelote steekproefbuizen per stratum kan evenredig zijn met het totaal aantal buizen in het stratum, ook wel proportionele allocatie genoemd. Het effect van proportionele en niet-proportionele allocatie in combinatie met *Ordinary Least Squares* (zie § 5.2.3) is onderzocht door naast evenredige aantallen ook gelijke aantallen steekproefbuizen per stratum (alle vier strata vijf buizen) te selecteren bij de steekproefopzettingen STSI(Gt) en STSI(h). In geval van Gt-stratificatie bestaat stratum één echter uit slechts vier buizen, één minder dan het gewenste aantal van vijf. Dit heeft tot gevolg dat de aantallen buizen per stratum in dit geval niet exact gelijk zijn. Door selectie van één extra buis in stratum 3 is het totaal aantal buizen wel op 20 gebracht.

Uit de theorie van *experimental design* volgt dat gerichte selectie van tien buizen met een minimale waarde voor de predictorvariabele (de grondwaterstand op een bepaalde dag) en tien buizen met een maximale waarde van de predictor optimaal is voor het fitten van een enkelvoudig lineair model. Omdat in de praktijk bij de selectie van de stambuizen nog niet bekend is op welke dag de gerichte opnames van de grondwaterstand zullen worden verricht, is niet de grondwaterstand op een bepaalde dag gebruikt als selectiecriteria, maar de gemiddelde grondwaterstand die is bepaald door alle beschikbare standen in de meetperiode te middelen. In deze steekproefopzet worden dus de tien buizen met de kleinste gemiddelde grondwaterstand en de tien buizen met de grootste gemiddelde grondwaterstand geselecteerd. In deze steekproefopzet worden de buizen dus niet geloot. Tabel 11 geeft een overzicht van de onderzochte steekproefopzettingen.

Tabel 9 Groepering van grondwatertrappen tot vier grondwatertrap-strata

Gt	II	II*	III	III*	IV	V	V*	VI	VII	VII*
Groep	1	1	2	2	2	3	3	3	4	4

Tabel 10 Aantal beschikbare en geselecteerde stambuizen per stratum voor evenredige, gelijke en bijna gelijke aantallen met stratificatie op Gt, gemiddelde grondwaterstand (h) en geografische ligging (X).

Groep	Gt			h			X	
	Beschikbaar	Evenredig	Bijna gelijk	Beschikbaar	Evenredig	Gelijk	Beschikbaar	Evenredig
1	5	1	4	14	3	5	17	4
2	28	5	5	37	7	5	37	7
3	45	9	6	32	6	5	31	6
4	24	5	5	19	4	5	17	3

Tabel 11 Selectie wijze voor de onderzochte steekproefopzetten.

Steekproefopzet	Loten	Stratificatie	Allocatie	Selectie-kenmerk
SI	Ja	-	-	
STSI(Gt,p)	Ja	Gt-bodemkaart	Proportioneel	
STSI(Gt,c)	Ja	Gt-bodemkaart	Bijna gelijke aantallen	
STSI(h,p)	Ja	Gemiddelde grondwaterstand in buis	Proportioneel	
STSI(h,c)	Ja	Gemiddelde grondwaterstand in buis	Gelijke aantallen	
STSI(x,p)	Ja	Geografische ligging	Proportioneel	
Gericht(h)	Nee	-	-	Gemiddelde grondwaterstand in buis

5.3.3 Fitten van het regressiemodel

De stambuisregressiemodellen zijn gefit met *Ordinary Least Squares* (OLS). Hierbij wordt verondersteld dat de variantie van de residuen constant is (homoscedasticiteit), en dat de residuen onderling onafhankelijk zijn. Bij de gestratificeerde aselechte steekproeven zijn de *design*-gewichten (inverse van de kans dat stambuis wordt geselecteerd) dus niet gebruikt als gewicht bij het fitten van het model omdat dit bij de reguliere Gd-projecten ook niet wordt gedaan. In geval van proportionele allocatie zijn de verschillen tussen de *design*-gewichten relatief klein. Hierdoor zullen de met OLS en *Weighted Least Squares* (WLS) geschatte regressiecoëfficiënten nauwelijks verschillen, en zal de vertekening van de OLS-schatter van de populatieregressiecoëfficiënten klein zijn. In geval van niet-proportionele allocatie kan de vertekening van de OLS-schatter van de populatieregressiecoëfficiënten aanzienlijk zijn, met als gevolg ook vertekende voorspellingen van de GxG op de gerichte opname locaties (Särndal, 1980).

5.4 Validatie

5.4.1 Leave-one-out

De geschiktheid van het meettijdstip en van de steekproefopzetten voor de selectie van de kalibratiebuizen is onderzocht door met de gefitte regressiemodellen de GxG te voorspellen op een punt met bekende GxG, en de voorspelde waardes te vergelijken met de werkelijke waarde. Dit is gedaan door om de beurt één buis apart te houden, en de 20 kalibratiebuizen te selecteren uit de resterende 101 buizen (*leave-one-out* validatie). In geval van selectie door middel van loting volgens één van bovengenoemde steekproefopzetten zijn, gegeven een validatiebuis, 100 steekproeven van 20 stambuizen getrokken (zonder teruglegging) uit de resterende 101 buizen. Voor elke steekproef zijn zes enkelvoudig lineaire modellen gefit voor de GHG (voor elke voorjaarsgrondwaterstand een model) en zes modellen voor de GLG (voor elke najaarsgrondwaterstand een model). In totaal is dus per steekproefopzet $102 \cdot 100 \cdot 12 = 122400$ keer een model gefit, 61200 keer voor de GHG en 61200 keer voor de GLG. In geval van de gerichte (modelgebaseerde) selectie van 20 buizen (Gericht(h)) is $102 \cdot 12 = 1224$ keer een model gefit.

5.4.2 Kwaliteitsmaten

Uit de verschillen tussen de voorspelde GxG's en werkelijke GxG voor de validatiebuizen zijn de volgende kwaliteitsmaten berekend:

1. de modelvariantie (variantie over realisaties van het regressiemodel) van de voorspelfout gemiddeld over de 102 stambuizen en over de 100 steekproeven (gemiddelde modelvariantie):

$$AVE = \frac{1}{NS} \sum_{i=1}^N \sum_{s=1}^S \sigma_{is}^2, \quad (5.1)$$

waarin N het aantal stambuizen is ($N = 102$), en S het aantal steekproeven is ($S = 100$);

2. de gekwadraterde fout gemiddeld over de 102 stambuizen en over de 100 steekproeven (gemiddelde gekwadraterde fout):

$$MSE = \frac{1}{NS} \sum_{i=1}^N \sum_{s=1}^S E_{is}^2. \quad (5.2)$$

3. de fout gemiddeld over de 102 stambuizen en over de 100 steekproeven (systematische fout):

$$ME = \frac{1}{NS} \sum_{i=1}^N \sum_{s=1}^S E_{is}. \quad (5.3)$$

4. de steekproefvariantie (variantie over de 100 steekproeven) van de geschatte regressiecoëfficiënten (steekproefvariantie regressiecoëfficiënten), gemiddeld over de 102 buizen:

$$\begin{aligned}\text{var}(b_0) &= \frac{1}{N} \sum_{i=1}^N \frac{1}{S-1} \sum_{s=1}^S (b_{0s} - m(b_0))^2 \\ \text{var}(b_1) &= \frac{1}{N} \sum_{i=1}^N \frac{1}{S-1} \sum_{s=1}^S (b_{1s} - m(b_1))^2,\end{aligned}\tag{5.4}$$

waarin b_{0s} en b_{1s} de geschatte regressiecoëfficiënten zijn van steekproef s , en $m(b_0)$ en $m(b_1)$ het gemiddelde is van de geschatte regressiecoëfficiënten over de 100 steekproeven. In geval van gerichte selectie van kalibratie-stambuizen (Gericht(h)) is er, gegeven een validatiepunt, slechts één steekproef geselecteerd. Dit betekent dat de eerste drie kwaliteitsmaten niet gemiddeld zijn over 100 steekproeven maar alleen over de 102 stambuizen, en dat de vierde kwaliteitsmaat niet bestaat.

In de praktijk wordt de nauwkeurigheid van de regressievoorspellingen zelden door middel van validatie bepaald, maar wordt de nauwkeurigheid gekwantificeerd met de berekende modelvariantie. Vergelijking van de gemiddelde modelvariantie en de gemiddelde gekwadraterde fout geeft inzicht in de kwaliteit van de berekende modelvariantie als maat voor de gemiddelde gekwadraterde fout. De gemiddelde modelvariantie is op twee manieren berekend, n.l. met de restvariantie (variantie van de residuen) geschat op basis van de 20 kalibratie-stambuizen (geschatte restvariantie van model gefit met 20 buizen), en met de restvariantie geschat op basis van alle 102 stambuizen (geschatte restvariantie van model gefit met alle 102 buizen). De eerste modelvariantie is interessant omdat in de praktijk de modelvariantie op deze manier wordt berekend, de tweede is interessant omdat deze inzicht geeft in de waarde van de modelvariantie als maat voor de gemiddelde gekwadraterde fout in gevallen waarin we de restvariantie zouden kennen (potentiële waarde).

In de praktijk wordt de steekproef niet herhaald, maar wordt met één steekproef van stambuizen het model gefit. Daarom is ook gekeken naar de steekproefvariantie van de modelvariantie, d.w.z. de variantie van de berekende modelvariantie bij herhaalde steekproefname van kalibratiestambuizen volgens steekproefopzet p , berekend volgens:

$$\begin{aligned}MVVE &= \frac{1}{N} \sum_{i=1}^N VVE_i; \\ VVE_i &= \frac{1}{S-1} \sum_{s=1}^S (\sigma_{is}^2 - AVE_i)^2; \\ AVE_i &= \frac{1}{S} \sum_{s=1}^S \sigma_{is}^2.\end{aligned}\tag{5.5}$$

Voor de steekproefopzet Gericht(h) bestaat deze steekproefvariantie niet. Immers, gegeven een validatiebuis, is er geen herhaling van steekproefname. Hoe kleiner de steekproefvariantie, hoe kleiner de kans dat bij een individuele steekproef de berekende modelvariantie sterk afwijkt van de werkelijke modelvariantie.

5.5 Resultaten

Tabel 12 geeft weer de gemiddelde gekwadrateerde fout voor de GHG voor de zes meettijdstippen en de zes steekproefopzetten. Kijken we naar de gemiddelden per meetdatum in de meeste rechtse kolom en die in de onderste rij met gemiddelden per steekproefopzet, dan valt op dat de verschillen tussen de meettijdstippen veel groter zijn dan die tussen de steekproefopzetten. Hetzelfde is te zien in Tabel 13 voor de GLG.

Het beste tijdstip voor de GHG is in dit geval 30 januari. Dit geldt voor alle steekproefopzetten. Voor de GHG is de gemiddelde gekwadrateerde fout voor 30 januari (157.3cm^2) aanzienlijk kleiner dan voor de overige tijdstippen. Uit de neerslagcijfers van het KNMI blijkt dat 30 januari vooraf werd gegaan door een lange periode met nauwelijks neerslag. De voorspelfouten voor de GHG bleken groter te worden, naarmate er kort voor de gerichte opname meer neerslag was gevallen. Voor de GLG is de gemiddelde gekwadrateerde fout het kleinst voor 16 augustus (165.7cm^2), maar voor 30 augustus en 16 september is deze niet veel groter. Opvallend is dat voor de GLG het beste meettijdstip direct voorafgegaan wordt door het slechtste meettijdstip (278.8cm^2). Voor de GHG komt de meetdag met de kleinste gemiddelde gekwadrateerde fout niet overeen met de dag waarop het gemiddelde verschil, het gemiddelde absolute verschil, of het gemiddelde gekwadrateerde verschil tussen de grondwaterstand en de GHG minimaal is (Tabel 14). Op de beste dag staat de grondwaterstand gemiddeld 21 cm onder GHG-niveau. Voor selectie van een geschikt meettijdstip is het dus nodig het model te fitten en de restvariantie te schatten. Voor de GLG valt de dag met de kleinste restvariantie wel samen met de dag met het kleinste gemiddelde absolute of gekwadrateerde verschil tussen GLG en de gemeten grondwaterstand. Voor de GLG is het tijdstip met de kleinste restvariantie wel het tijdstip waarop deze verschillen ook het kleinst zijn. Een verband tussen de voorspelfout van de GLG en de hoeveelheid neerslag voorafgaand aan de datum van de gerichte opname is minder duidelijk dan bij de GHG. De redenen hiervoor zijn dat 1) de dikkere onverzadigde zone bufferend werkt en dus het effect van neerslag kort voor de gerichte opname op de grondwaterstand dempt, en 2) neerslagpatronen in de zomer convectief zijn, waardoor het verband tussen KNMI-data en voorspelfouten in de GLG minder duidelijk is dan in de winter bij de GHG en adjectieve neerslag.

Tabel 12 Gemiddelde gekwadrateerde fout (MSE) van voorspelde GHG (cm²) voor zes meettijdstippen en zeven steekproefopzettingen

Datum	SI	STSI(Gt,p)	STSI(Gt,c)	STSI(h,p)	STSI(h,c)	STSI(x,p)	Gericht(h)	Gem.
16-jan	181.7	180.3	176.8	176.5	173.6	181.2	190.4	180.1
30-jan	157.7	156.3	153.1	152.9	150.2	157.7	173.3	157.3
16-feb	196.9	197.0	195.7	194.1	193.9	195.3	191.3	194.9
28-feb	243.1	244.0	242.3	240.7	241.1	241.4	238.7	241.6
16-mrt	296.8	296.6	296.0	293.7	296.9	296.4	294.4	295.8
30-mrt	339.1	336.5	334.5	329.9	329.9	339.9	389.5	342.8
Gem.	235.9	235.1	233.1	231.3	230.9	235.3	246.3	235.4

Tabel 13 Gemiddelde gekwadrateerde fout (MSE) van voorspelde GLG (cm²) voor zes meettijdstippen en zeven steekproefopzettingen

Datum	SI	STSI(Gt,p)	STSI(Gt,c)	STSI(h,p)	STSI(h,c)	STSI(x,p)	Gericht(h)	Gem.
16-jul	214.9	211.7	213	212.5	211.0	213.2	227.7	214.9
30-jul	278.5	274.8	275	271.6	268.4	275.8	307.5	278.8
16-aug	167.7	165.9	165.2	163.9	160.8	167.7	169.0	165.7
30-aug	171.3	169.6	168.9	167.4	164.6	170.6	185.4	171.1
16-sep	178.1	176.1	175.9	175.0	171.7	176.5	165.4	174.1
30-sep	271.6	268.5	271.1	266.2	264.5	270.9	258.7	267.4
Gem.	213.7	211.1	211.5	209.4	206.8	212.4	219.0	212.0

Het tijdstip waarop de grondwaterstand wordt gemeten is veel meer bepalend voor de kwaliteit van de regressievoorspellingen dan de steekproefopzet waarmee de stambuizen worden geselecteerd in de geografische en eigenschappenruimte. Stratificatie op basis van de gemiddelde grondwaterstand is voor zowel de GHG als de GLG iets beter dan stratificatie op basis van de grondwatertrap (Gt) volgens de Bodemkaart van Nederland 1 : 50 000, maar de verschillen zijn nauwelijks vermeldenswaardig. Verder valt op dat de gemiddelde gekwadrateerde fout voor zowel de GHG als de GLG het grootst is bij de gerichte steekproef. Dit wordt geheel verklaard door de systematische fout in de voorspelde GxG die bij deze steekproefopzet optreedt (Tabellen 15 en 16). Voor alle kanssteekproefopzettingen is de systematische fout verwaarloosbaar klein, ook in geval van (bijna) gelijke aantallen per stratum, maar voor de gerichte steekproefopzet vormt de systematische fout een niet te verwaarlozen deel van de gemiddelde gekwadrateerde fout. Omdat grondwaterstanden en GxG beneden maaiveld zijn weergegeven als negatieve waarden, betekent een negatieve waarde voor de systematische fout een te diepe GxG ten opzichte van maaiveld. Deze systematische fout hangt samen met een systematische fout in de geschatte regressiecoëfficiënten (Tabel 17). De werkelijke gemiddelden (waarden voor de regressiecoëfficiënten wanneer alle 102 buizen worden gebruikt bij het fitten van het model) zijn voor de GHG 8.7 (b_0) en 0.888 (b_1), en voor de GLG -18.8 (b_0) en 1.041 (b_1). De verschillen tussen de gemiddelden van de geschatte regressiecoëfficiënten over de 100 steekproeven en deze werkelijke gemiddelden zijn klein en niet significant ($\alpha = 0.05$), maar voor de gerichte steekproefopzet relatief het grootst. Zowel het intercept als de helling wordt overschat, wat de te diepe voorspelde GxG verklaart.

Tabel 14 Restvariantie van het stambuisregressie model voor GHG (eerste zes meettijdstippen) en GLG (laatste zes meettijdstippen), gefit op alle 102 buizen, en het verschil, het absolute verschil en het gekwadrateerde verschil tussen de grondwaterstand en de GxG, gemiddeld over de 102 buizen

Datum	MSres (cm ²)	Gem. verschil (cm)	Gem. abs. verschil (cm)	Gem. gekwad. verschil (cm ²)
16-jan	162.1	-27.5	27.6	924
30-jan	140.8	-21.0	21.4	601
16-feb	165.9	-12.0	13.3	337
28-feb	203.4	-7.5	11.3	267
16-mrt	259.4	1.3	10.8	262
30-mrt	300.4	-16.1	19.8	554
16-jul	187.1	36.8	37.3	1538
30-jul	245.1	39.1	39.6	1767
16-aug	147.5	25.0	25.8	772
30-aug	150.1	28.5	29.3	959
16-sep	157.5	38.2	38.9	1622
30-sep	242.7	52.7	53.1	3049

Tabel 15 Systematische fout (ME) van voorspelde GHG (cm) voor zes meettijdstippen en zeven steekproefopzetten

Datum	SI	STSI(G _{t,p})	STSI(G _{t,c})	STSI(h _p)	STSI(h _c)	STSI(x _p)	Gericht(h)	Gem.
16-jan	0.29	0.32	-1.25	-0.03	-0.70	0.24	-4.77	-0.84
30-jan	0.26	0.31	0.01	0.01	-0.51	0.24	-4.90	-0.66
16-feb	-0.04	-0.04	0.76	-0.14	-0.27	-0.08	-1.03	-0.12
28-feb	-0.08	-0.16	0.47	-0.10	0.07	-0.25	-0.06	-0.01
16-mrt	-0.02	-0.15	0.31	-0.06	0.08	-0.08	-0.81	-0.11
30-mrt	0.25	0.14	-1.09	-0.03	-0.80	0.27	-7.12	-1.20
Gem.	0.11	0.07	-0.13	-0.06	-0.35	0.05	-3.12	-0.49

Tabel 16 Systematische fout (ME) van voorspelde GLG (cm) voor zes meettijdstippen en zeven steekproefopzetten

Datum	SI	STSI(G _{t,p})	STSI(G _{t,c})	STSI(h _p)	STSI(h _c)	STSI(x _p)	Gericht(h)	Gem.
16-jul	0.19	0.04	1.44	-0.08	-0.78	0.25	-5.55	-0.64
30-jul	0.19	0.04	1.68	-0.21	-1.34	0.48	-7.20	-0.91
16-aug	0.17	0.04	0.99	-0.10	-0.71	0.25	-4.18	-0.50
30-aug	0.09	0.02	0.84	-0.20	-1.03	0.29	-5.57	-0.80
16-sep	-0.09	-0.17	1.46	-0.18	-0.45	0.12	-1.65	-0.14
30-sep	-0.18	-0.32	2.13	-0.10	0.38	-0.07	1.65	0.50
Gem.	0.06	-0.06	1.42	-0.14	-0.66	0.22	-3.75	-0.41

Tabel 17 Gemiddelde ($m(\cdot)$) en variantie ($var(\cdot)$) over de 100 steekproeven van de geschatte regressiecoëfficiënten (b_0 : intercept, b_1 : helling) van het enkelvoudig lineaire model voor de GHG (predictor grondwaterstand op 30 januari) en GLG (predictor grondwaterstand 16 augustus), gefit op 20 buizen, voor zeven steekproefopzetten

		SI	STSI(Gt,p)	STSI(Gt,c)	STSI(h,p)	STSI(h,c)	STSI(x,p)	Gericht(h)
GHG	$m(b_0)$	8.3	8.6	8.4	9.0	10.0	7.8	10.8
	$m(b_1)$	0.884	0.886	0.886	0.892	0.905	0.879	0.954
	$var(b_0)$	30.9	26.4	26.3	24.3	12.2	36.0	-
	$var(b_1)$	0.0029	0.0021	0.0023	0.0020	0.0013	0.0030	-
GLG	$m(b_0)$	-19.1	-18.9	-14.4	-18.3	-17.3	-19.2	-16.5
	$m(b_1)$	1.038	1.040	1.063	1.044	1.055	1.037	1.083
	$var(b_0)$	106.9	112.4	91.2	81.1	51.5	120.5	-
	$var(b_1)$	0.0050	0.0051	0.0044	0.0041	0.0026	0.0057	-

In Tabel 18 zijn de modelvarianties van de voorspelde GHG (voorspeld met de grondwaterstand van 30 januari) en GLG (voorspeld met de grondwaterstand van 16 augustus) weergegeven, zoals berekend met de geschatte restvariantie en met de werkelijke restvariantie van Tabel 14. Deze modelvariantie is gemiddeld over de 100 steekproeven, en over de 102 buizen.

Tabel 18 Gemiddelde gekwadrateerde fout (MSE), en gemiddelde modelvariantie (AVE) van GHG (voorspeld met grondwaterstand van 30 januari) en GLG (voorspeld met grondwaterstand van 16 augustus) berekend met geschatte restvariantie en met werkelijke restvariantie. Ook is de steekproefvariantie van de modelvariantie (berekend met geschatte restvariantie) (MVVE) en de variatiecoëfficiënt (ρ) weergegeven

		SI	STSI (Gt,p)	STSI (Gt,c)	STSI (h,p)	STSI (h,c)	STSI (x,p)	Gericht (h)
GHG	MSE	157.7	156.3	153.1	152.9	150.2	157.7	173.3
	AVE (berekend met geschatte MSres)	155.9	158.1	136.5	153.2	131.6	155.9	39.8
	AVE (berekend met werkelijke MSres)	156.6	156.3	156.9	155.1	153.5	156.7	150.3
	MVVE	6963 (0.54)	7025 (0.53)	5207 (0.53)	6832 (0.54)	6078 (0.59)	6850 (0.53)	-
GLG	MSE	167.7	165.9	165.2	163.9	160.8	167.7	169.0
	AVE (berekend met geschatte MSres)	161.7	163.8	135.1	159.2	145.0	159.6	87.7
	AVE (berekend met werkelijke MSres)	164.2	163.9	164.3	162.5	160.7	164.4	157.4
	MVVE	8477 (0.57)	8726 (0.57)	6194 (0.58)	8366 (0.57)	7567 (0.60)	8110 (0.56)	-

Uit Tabel 18 blijkt dat behalve voor steekproefopzet Gericht(h) het verschil tussen de gemiddelde modelvariantie, berekend met de werkelijke restvariantie, en de gemiddelde gekwadraterde fout klein is. Het verschil tussen deze twee kwaliteitsmaten voor Gericht(h) kan verklaard worden door de systematische fout, die niet wordt meegenomen in de modelvariantie. De gemiddelde modelvariantie berekend met de geschatte restvariantie is een goede maat voor de gemiddelde gekwadraterde fout voor de steekproefopzetten enkelvoudig aselekt (SI) en gestratificeerd enkelvoudig aselekt met proportionele allocatie (STSI(*,p)). Voor een gestratificeerde enkelvoudig aselekte steekproef met (bijna) gelijke aantallen buizen per stratum (STSI(*,c)), en met name voor gerichte selectie op basis van gemiddelde grondwaterstand (Gericht(h)), wordt de gemiddelde gekwadraterde fout onderschat wanneer de modelvariantie wordt berekend met de geschatte restvariantie. Dit kan mogelijk verklaard worden door heteroscedasticiteit, d.w.z. de variantie van de residuen is niet constant, maar afhankelijk van het niveau van de predictor. In geval van gestratificeerde steekproef met niet-evenredige allocatie en gerichte selectie zijn de insluitkansen van buizen ongelijk, waardoor in geval van heteroscedasticiteit de schatter van de restvariantie onzuiver wordt. Door te zorgen voor gelijke insluitkansen wordt de methode minder gevoelig voor de aanname van homoscedasticiteit.

De bijdrage van de onzekerheid over de regressiecoëfficiënten in de voorspelfout kan worden bepaald door vergelijking van de gemiddelde model-variantie (berekend met de werkelijke restvariantie) met de restvariantie (Tabel 16). Voor de GHG varieert dit verschil van 9.5 cm² voor Gericht(h) tot 16.1 cm² (STSI(Gt,c)), voor de GLG van 9.9 cm² voor Gericht(h) tot 16.9 cm² (STSI(x,p)). Dit zijn verhoudingsgewijs kleine bijdrages (6 -10% van gemiddelde model-variantie).

De variatie in model-variantie (berekend met geschatte restvariantie) tussen de 100 steekproeven, uitgedrukt in de variatiecoëfficiënt, is voor alle steekproefopzetten erg groot (0.50-0.60) en ongeveer gelijk (voor STSI(h,c) is ze het grootst). Wanneer 30 kalibratiebuizen worden geselecteerd met steekproefopzet STSI(Gt,p) neemt deze variatiecoëfficiënt af naar 0.40 (GHG) en 0.43 (GLG) (Tabel 19) . Met 40 buizen neemt ze verder af naar 0.32 (GHG) en 0.34 (GLG).

Tabel 19 Steekproefvariantie van de geschatte regressiecoëfficiënten (b_0 : intercept, b_1 : helling) van het stambuisregressie model voor de GHG (predictor grondwaterstand op 30 januari) en GLG (predictor grondwaterstand 16 augustus), en steekproefvariantie van de geschatte modelvariantie van de voorspelfout, MVVE, gefit op 20, 30 en 40 buizen geselecteerd met steekproefopzet STSI(G_t, p)

Aantal	GHG			GLG		
	$\text{var}(b_0)$	$\text{var}(b_1)$	MVVE	$\text{var}(b_0)$	$\text{var}(b_1)$	MVVE
20	26.4	0.0021	7025 (0.53)	112.4	0.0051	8726 (0.57)
30	14.4	0.0013	3703 (0.40)	52.1	0.0025	4608 (0.43)
40	10.9	0.00093	2399 (0.32)	31.5	0.0016	2954 (0.34)

5.6 Discussie

De representativiteit van beschikbare OLGA-buizen voor het te karteren gebied werd als belangrijk ingeschat. Met het experiment zoals hier uitgevoerd over de selectie van stambuizensets, kan hierop maar gedeeltelijk antwoord worden gegeven omdat alleen onderzocht is hoe uit beschikbare OLGA-buizen het best een set stambuizen kan worden samengesteld. Het gegeven dat nauwelijks OLGA-buizen in akkerbouwpercelen en gedraineerde percelen beschikbaar zijn terwijl voorspellingen met stambuisregressie wel voor deze gebieden worden uitgevoerd, maakt beantwoording van de complete onderzoeksvraag in dit project niet mogelijk.

In deze studie zijn de 102 stambuizen beschouwd als de te onderzoeken populatie. In de praktijk zijn we niet alleen geïnteresseerd in de 102 stambuizen, maar in het hele gebied waaruit deze stambuizen afkomstig zijn, in dit geval het waterschap Regge en Dinkel. Het doel is het schatten van het regressiemodel voor dit hele onderzoeksgebied, d.w.z. het regressiemodel dat zou worden verkregen wanneer de GxG en de grondwaterstand op alle punten in het gebied bekend zou zijn, en het model met al deze punten zou worden gefit.

Uit dit deelonderzoek blijkt dat een gestratificeerde steekproef met (groepen van) Gt's als strata zo gek nog niet is. Ook blijkt dat ongelijke insluitkansen van de stambuizen in combinatie met OLS tot sterk vertekende schattingen van de modelvariantie kan leiden. Daarom is het van belang om bij Gt-gestratificeerde steekproeven van stambuizen de aantallen buizen evenredig met de oppervlakte van de Gt-strata te kiezen.

5.7 Conclusies

De volgende conclusies kunnen uit het experiment dat in dit hoofdstuk is beschreven worden getrokken:

- De keuze van de dag waarop de grondwaterstand in de stambuizen (en op de plekken waar de GxG moet worden voorspeld) wordt gemeten is veel meer bepalend voor de kwaliteit van de regressievoorspellingen dan de methode waarmee de subset van stambuizen wordt geselecteerd.
- Voor de GHG is het gemiddelde absolute of gekwadrateerde verschil tussen de grondwaterstand en de GHG in de stambuizen geen goed criterium voor het

selecteren van een geschikt meettijdstip. Hiervoor is het nodig te kijken naar de restvariantie van het regressiemodel. Voor de GLG kan in dit geval het beste tijdstip ook bepaald worden met het gemiddelde absolute of gekwadrateerde verschil tussen grondwaterstand en GLG in de stambuizen.

- Het is van belang dat in de dagen voorafgaand aan de gerichte opname er geen neerslag is gevallen.
- Gerichte selectie op basis van de gemiddelde grondwaterstand is ongeschikt als methode voor selectie van een substeekproef van kalibratiebuizen uit het DINO-bestand, omdat het leidt tot vertekende schattingen van het regressiemodel, en daarmee tot systematische fouten in de regressievoorspellingen. Bovendien leidt deze methode tot sterk vertekende schattingen (onderschatting) van de model-variantie van de voorspelfout.
- Gestratificeerde selectie met (bijna) gelijke aantallen buizen per stratum (d.w.z. niet-evenredige allocatie) leidt niet tot substantiële systematische fouten in de regressievoorspellingen, maar wel tot vertekende schattingen (onderschattingen) van de modelvariantie van de voorspelfout. Evenredige allocatie leidt wel tot zuivere schattingen van de modelvariantie van de voorspelfout.
- In geval van proportionele allocatie, levert stratificatie op basis van de gemiddelde grondwaterstand de nauwkeurigste voorspellingen op (gemiddelde gekwadrateerde voorspelfout kleinst), maar de verschillen met Gt-stratificatie en geografische stratificatie zijn klein.
- Kwantificering van de modelvariantie van de voorspelfout bij 20 buizen is erg onnauwkeurig (variatiecoëfficiënt 0.55). Door meer buizen te selecteren voor kalibratie zal de nauwkeurigheid van de voorspellingen zelf niet veel verbeteren, maar die van de schattingen van de nauwkeurigheid (modelvariantie) wel. Wanneer kwantificering van nauwkeurigheid belangrijk is wordt aangeraden om te streven naar 40 stambuizen voor het fitten van het regressiemodel. Naar verwachting zal de nauwkeurigheid van de geschatte model-variantie ook verbeteren wanneer het lukt om de restvariantie te reduceren door twee grondwaterstanden (meting rond GHG en rond GLG) als predictor te gebruiken in het regressiemodel. Dit is echter hier niet onderzocht.

6 Het effect van gerichte opnamelocaties op de nauwkeurigheid van gebiedsdekkende GxG-voorspellingen

6.1 Inleiding

Bij opdrachtgevers voor een Gd-kartering bestaat vaak de wens om met een bepaalde, voor een specifieke toepassing noodzakelijke, nauwkeurigheid uitspraken te doen over de grondwaterstandsdynamiek. Bij de aanwijzing van uitspoelingsgevoelige percelen voor de mestwetgeving wordt, bijvoorbeeld, op basis van gekwantificeerde onzekerheid van gebiedsdekkende GxG voorspellingen met een bepaalde zekerheid een perceel als al dan niet uitspoelingsgevoelig aangewezen (De Gruijter et al., 2004). Om aan te kunnen sluiten bij dit soort toepassingen is het van belang te weten welke factoren de nauwkeurigheid van de voorspelde GxG beïnvloeden en om een correcte inschatting te kunnen maken van deze nauwkeurigheid. In dit hoofdstuk staan daarom de volgende drie onderzoeksvragen centraal:

1. Heeft een toename van de waarnemingsdichtheid bij de gerichte opname een positief effect op de nauwkeurigheid van gebiedsdekkende voorspellingen? Deze vraag wordt in dit hoofdstuk onderzocht door zowel toevallige als systematische fouten in de voorspelde GxG te bepalen bij verschillende waarnemingsdichtheden.
2. In de huidige Gd-karteringen worden preferent locaties in natte Gt's geallokeerd. Opdrachtgevers hechten meer belang aan betrouwbare uitspraken in natte dan in droge Gt's en de gedachte is dat dit door preferente allocatie in natte Gt's bereikt kan worden. Wat is het effect op de nauwkeurigheid van de GxG-voorspellingen van preferente allocatie in natte terreindelen t.o.v. proportionele allocatie over natte en droge terreindelen?
3. Niet alleen de nauwkeurigheid van de voorspellingen zelf is van belang, maar ook een correcte inschatting van deze nauwkeurigheid. Stemt de nauwkeurigheid die bij de gehanteerde methode voor Gd-kartering wordt gekwantificeerd overeen met de nauwkeurigheid die uit validatie is bepaald?

In dit experiment wordt in acht steekproefopzetten stapsgewijs de waarnemingsdichtheid en allocatie over Gt-classes gevarieerd en wordt door middel van validatie het effect hiervan op de predictie van de GLG vastgesteld. Ook wordt op basis van validatie beoordeeld of de gekwantificeerde nauwkeurigheid overeenstemt met de waargenomen nauwkeurigheid.

6.2 Studiegebied en dataverzameling

Het studiegebied voor dit experiment bestaat uit een deelgebied van ongeveer 2 000 ha gelegen nabij Markelo in de provincie Overijssel en is aangegeven in Figuur 5 en 6. Het deelgebied is begrensd op basis van de bodemkaart 1 : 50 000 en bestaat uit klei op fijn zand, leemarm zand, sterk lemig fijn zand op keileem en enkeerdgronden. Het grondwaterniveau varieert volgens de 1 : 50 000 bodemkaart

tussen Gt-klassen III en VII* en het gebied heeft een gemiddelde maaiveldshoogte van 10.50 m + NAP en de standaarddeviatie van maaiveldshoogten bedraagt 1.3 m. Op 530 locaties is tussen juni en september 2002 in een boorgat eenmalig de grondwaterstand ten opzichte van maaiveld gemeten, de zogenaamde gerichte opname. Daarnaast is op 24 meetreeksen van grondwaterstanden in stambuizen in en nabij het studiegebied, tijdreeksanalyse uitgevoerd waarmee voor deze buizen een klimaatsrepresentatieve GLG is berekend. Met behulp van stambuisregressie (zie § 2.4.1.) is voor elke meetdatum het verband tussen de gemeten stand en de klimaatsrepresentatieve GLG bepaald. De gefitte regressierelaties zijn gebruikt voor het voorspellen van klimaatsrepresentatieve GLG's op de 530 meetlocaties waar eenmalig een grondwaterstand is gemeten.

6.3 Opzet van het experiment

6.3.1 Steekproefopzet

De 530 meetlocaties zijn gestratificeerd aselekt geloot binnen de aanwezige Gt-klassen. Uit de 530 beschikbare meetlocaties zijn, proportioneel naar areaal van de Gt-klassen, gestratificeerd aselekt 79 meetlocaties geloot als validatieset.

Uit de resterende 451 locaties zijn gestratificeerd aselekt tien maal, bij benadering, 20, 40, 80 en 160 locaties geloot. Dit is op twee manieren gedaan:

1. proportioneel naar het areaal van de Gt-klassen binnen het studiegebied, en
2. met een oververtegenwoordiging van locaties in natte Gt-klassen, verder aangeduid met preferent.

De aanduiding van de steekproefopzet en het aantal locaties worden gegeven in Tabel 20.

Door afronding naar een geheel aantal waarnemingen per Gt-klasse wijkt het totaal aantal waarnemingen soms af van het gewenste aantal. Met elke steekproefopzet met een waarnemingsdichtheid van 20, 40, 80 of 160 en verdeling over de Gt-klassen (proportioneel of preferent geallokeerd) zijn tien steekproeven uitgevoerd. De tien steekproefresultaten zijn telkens gemiddeld, zodat toevallige afwijkingen tussen de steekproeven teniet worden gedaan.

Per Gt-klasse resulteert dat, afhankelijk van de waarnemingsdichtheid en het type allocatie over Gt-klassen, in het aantal locaties zoals vermeld in Tabel 21. Het cijfer na de "D" in de codering duidt op de waarnemingsdichtheid in aantal waarnemingen per km². De "_A" of "_P" erna staat voor respectievelijk allocatie proportioneel naar het areaal van de Gt-klassen binnen het studiegebied of preferente allocatie in natte Gt-klassen.

Tabel 20 Aanduiding van de steekproefopzet en aantal locaties () voor verschillende meetdichtheden en allocatie

Steekproefopzet	1/km ²	2/km ²	4/km ²	8/km ²
Allocatie proportioneel	D1_A	D2_A	D4_A	D8_A
Naar Gt-areaal	(21)	(39)	(79)	(161)
Allocatie preferent in	D1_P	D2_P	D4_P	D8_P
Natte Gt's	(20)	(40)	(80)	(160)

Tabel 21 Aantal meetlocaties per Gt-klasse voor de verschillende steekproefopzetten

Gt	D1_A	D2_A	D4_A	D8_A	D1_P	D2_P	D4_P	D8_P
III	4	8	16	33	7	15	29	59
III*	8	15	30	61	8	16	31	63
IV	3	5	11	22	2	3	7	14
VII	2	4	8	17	1	2	4	8
VII*	4	7	14	28	2	4	8	16

6.3.2 Toepassing van de Gd-methode bij verschillende steekproefopzetten

De 80 (8 · 10) steekproeven die volgens bovengenoemde werkwijze worden gealloceerd, worden gebruikt om de GLG met behulp van universal kriging, gebruikmakend van vlakdekkende hulpinformatie, te voorspellen (zie § 2.6). Hiervoor wordt allereerst voor elk van de 80 steekproeven gezocht naar het 'beste' lineaire verband tussen de GLG op meetlocaties en de beschikbare gebiedsdekkende hulpinformatie. De set kandidaatmodellen bestaat uit modellen met één tot vijf predictoren, die worden geselecteerd uit vijf groepen, met een maximum van één per groep. Het totaal aantal predictoren bedraagt 15 (De Gruijter et al. 2004). Afhankelijk van waarnemingsdichtheid en preferente dan wel proportionele allocatie over de Gt-klassen leidt dit tot een selectie van predictoren zoals weergegeven in Tabel 22. In totaal zijn 61 verschillende modellen –combinaties van predictoren- geselecteerd. Hieruit blijkt dat selectie van predictoren sterk kan verschillen bij eenzelfde steekproefopzet. Naarmate de waarnemingsdichtheid toeneemt treedt een zekere stabilisatie op, omdat het dan vaker voorkomt dat een predictor geen enkele van de tien keer gekozen wordt of juist alle tien keer. De predictor 'dtdrain' - drooglegging – wordt bijvoorbeeld tien van de tien keer en 'Grepd12' wordt nul van de tien keer geselecteerd bij de hoogste waarnemingsdichtheid.

Tabel 22 Aantal keren van de tien steekproeven dat een predictor wordt geselecteerd voor de verschillende steekproefopzetten

Groep	Predictor	D1_A	D2_A	D4_A	D8_A	D1_P	D2_P	D4_P	D8_P
1	hgt	4	2	2	1	3	4	0	1
2	Rht4	0	1	1	2	2	2	4	1
2	Rht12	1	1	0	1	1	1	2	0
2	Rht16	1	1	1	0	3	2	0	0
2	Rht20	2	3	3	3	1	3	0	7
3	Dtdrain	4	6	9	10	5	4	9	10
3	Ctdrain	3	0	1	0	1	2	0	0
4	Slootd12	2	2	0	1	1	0	1	0
4	Slootd16	1	0	1	1	1	1	1	0
4	Grepd12	2	1	1	0	2	0	1	0
4	Grepd16	0	1	0	3	1	0	0	1
5	Bghg	3	2	0	2	0	2	0	2
5	Bglg	2	3	3	4	4	1	4	3
5	Bergh	1	0	1	1	2	1	2	0
5	Bergl	0	0	1	2	1	1	0	3

Als alle 451 locaties worden gebruikt voor modelselectie, dan worden Rht20, Dtdrain en Bglg als predictoren geselecteerd. Het percentage verklaarde variantie bedraagt voor dit model 66% en de voorspelfout van de regressie is 26.3 cm. Op basis van dit model worden op de meetlocaties residuen bepaald. Van deze residuen wordt vervolgens het gestandaardiseerd variogram berekend.

Het gevonden variogram behorend bij dit model wordt gestandaardiseerd door de *nugget*- en de *sill*-parameter te delen door de restvariantie van het regressiemodel (zie § 2.6.2). Het gestandaardiseerde variogram heeft een *nugget*- en een *sill*-parameter van respectievelijk 0.35 en 0.51 cm². Per steekproef wordt dit gestandaardiseerde variogram gestandaardiseerd door vermenigvuldiging van de *nugget*- en de *sill*-parameter met de restvariantie van het regressiemodel dat geselecteerd is voor de betreffende steekproef.

Omdat de GLG-voorspellingen op meetpunten zijn behept met een pseudomeetfout (afkomstig uit de stambuisregressie, zie § 2.3.1) kan het variogram ook hiervoor worden gecorrigeerd. In deze studie wordt bij universal kriging zowel gebruik gemaakt van een variogram dat voor de gemiddelde variantie van de pseudomeetfout is gecorrigeerd als van een ongecorrigeerd variogram. Deze twee varianten zullen we verder aanduiden met de toevoegingen ‘met meetfoutcorrectie’ en ‘zonder meetfoutcorrectie’. Meetfoutcorrectie vindt plaats door de *nugget*-variantie van het variogram te verminderen met de gemiddelde variantie van de pseudomeetfout. Als daardoor de *nugget*-variantie negatief zou worden dan wordt deze op nul gezet.

Voor alle 80 steekproeven wordt de GLG geïnterpoleerd door middel van universal kriging, waarbij de geselecteerde regressiemodellen en het gefitte variogram worden gebruikt. De GLG's die voor de 79 validatielocaties zijn voorspeld met behulp van universal kriging, zijn vergeleken met de GLG's die voor deze locaties zijn geschat met behulp van stambuisregressie. Van de GLG-voorspellingen die met behulp van

stambuisregressie zijn verkregen op deze validatielocaties, is ook een ‘pseudo-meetfout’ gekwantificeerd in de vorm van een geschatte variantie van de voorspelfout. Deze is vergeleken met de modelvariantie die is berekend voor de GLG-voorspellingen die met universal kriging zijn uitgevoerd.

6.3.3 Validatie

Per steekproef worden de systematische fout en de variantie van de systematische fout geschat. Eerst wordt de systematische fout (ME_h) per stratum geschat,

$$ME_h = \frac{1}{n_h} \sum_{i=1}^{n_h} E_i \quad (6.1)$$

waarin n_h het aantal validatielocaties in stratum h is en E_i het verschil tussen de "waargenomen" en voorspelde GLG. Vervolgens wordt de gemiddelde, of systematische fout (ME) voor de steekproef geschat, waarbij rekening wordt gehouden met de verschillende insluitkansen per stratum door middel van wegingsfactoren (W_h), die worden berekend uit het quotiënt van het areaal in een stratum (A_h) en het areaal van het totale gebied (A), te rekenen (Cochran, 1977). Deze weging is nodig omdat de steekproef slechts bij benadering zelfwegend is. Er vindt immers een afronding naar een geheel aantal validatielocaties per stratum plaats, waardoor de aantallen niet exact proportioneel zijn aan de arealen. In formulevorm is de berekening van de systematische fout als volgt samengevat:

$$ME = \sum_{h=1}^H W_h \cdot ME_h, \quad (6.2)$$

waarin H het aantal strata is, en

$$W_h = \frac{A_h}{A}. \quad (6.3)$$

De geschatte variantie van de systematische fout per stratum (VME_h , $h=1\dots H$) en de geschatte variantie van de systematische fout voor iedere steekproef (VME) worden berekend met (Cochran, 1977):

$$VME_h = \frac{1}{n_h(n_h - 1)} \sum_{i=1}^{n_h} (E_i - ME_h)^2 \quad (6.4)$$

$$VME = \sum_{h=1}^H W_h^2 \cdot VME_h \quad (6.5)$$

Daarna wordt voor elk van de acht steekproefopzetten de geschatte systematische fout (AME) en de geschatte variantie van de systematische fout ($AVME$) bepaald voor de tien steekproeven per steekproefopzet. Dit gebeurt door middeling over het aantal steekproeven (S) dat met een bepaalde steekproefopzet is uitgevoerd:

$$AME = \frac{1}{S} \sum_{s=1}^S ME \quad (6.6)$$

$$AVME = \frac{1}{S} \sum_{s=1}^S VME, \quad (6.7)$$

$$ASME = \sqrt{AVME}.$$

Voor telkens tien steekproeven per steekproefopzet is bepaald hoe groot de variantie van de fout op validatielocaties (VE) is, als maat voor de toevallige fout. Hiervoor is allereerst de gekwadeerde fout van de GLG voorspellingen per stratum (MSE_h) berekend. Ook bij de berekening van de gekwadeerde fout (MSE) per steekproef moet rekening worden gehouden met de insluitkansen per Gt-stratum door het gebruik van wegingsfactoren (W_h). De variantie van de fout per steekproef (VE) wordt berekend uit de gemiddelde gekwadeerde fout (MSE), het kwadraat van de systematische fout (ME) en de variantie van de systematische fout (VME) (Marsman en de Gruijter, 1986):

$$MSE_h = \frac{1}{n_h} \sum_{i=1}^{n_h} E_i^2 \quad (6.9)$$

$$MSE = \sum_{h=1}^H W_h \cdot MSE_h \quad (6.10)$$

$$VE = MSE - (ME)^2 + VME \quad (6.11)$$

Belangrijk is om na te gaan in hoeverre het model in staat is de nauwkeurigheid van de voorspellingen te kwantificeren. Dit is geanalyseerd met behulp van zogenaamde Z-scores. Voor elke validatielocatie (i) is een Z-score berekend door de fout (E_i) te delen door de geschatte standaardafwijking van de voorspelfout (σ_i), berekend als de wortel uit de krigingvariantie op locatie i :

$$Z_i = \frac{E_i}{\sigma_i} \quad (6.12)$$

Als de variantie van deze Z-score, berekend over alle validatielocaties, rond de één ligt vindt geen onder- of overschatting van de nauwkeurigheid plaats. Als de variantie van de Z-score groter is dan één, dan is de modelschatting van de nauwkeurigheid een overschatting. Als de modelschatting van de nauwkeurigheid correct is bedraagt de variatie één. De nul-hypothese is daarom:

$$H_0 : VAR(Z) = 1. \quad (6.13)$$

Uit de Z-score's is als volgt een toetsingsgrootte per steekproef afgeleid (Snedecor en Cochran, 1989):

$$\chi^2 = \sum_{i=1}^n (Z_i - \bar{Z})^2. \quad (6.14)$$

6.4 Resultaten

6.4.1 Systematische fout

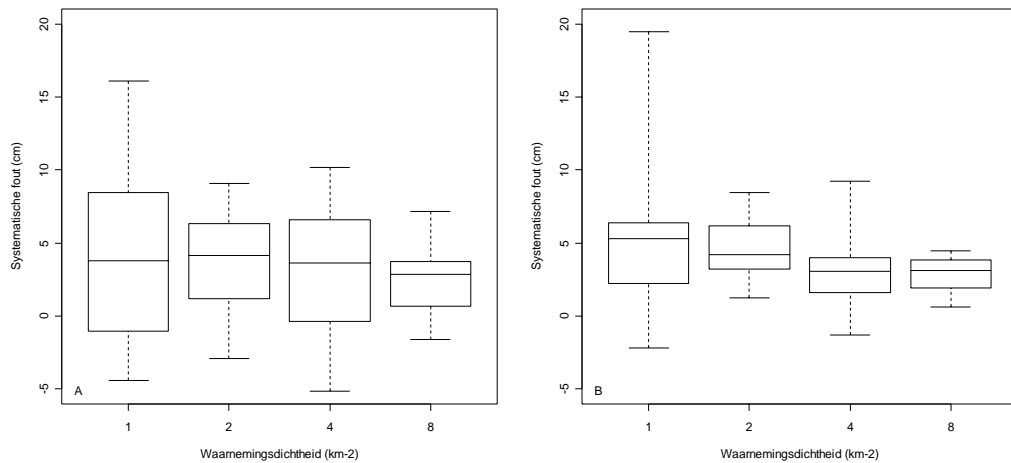
Tabel 23 geeft de systematische fouten (AME) en standaardafwijkingen van de systematische fouten per steekproefopzet ($ASME$). De systematische fouten blijken voor alle steekproefopzetten positief te zijn. Dit betekent dat de GLG-voorspellingen gemiddeld te droog zijn. De systematische fout per steekproefopzet vertoont een afname bij toenemende waarnemingsdichtheid en ook de spreiding in systematische fouten neemt af bij een toenemende waarnemingsdichtheid (zie figuur 17). Bij lagere waarnemingsdichtheden blijkt voor individuele steekproeven met een bepaalde

steekproefopzet vaker een significante systematische fout op te treden. Het aantal van de tien steekproeven met een bepaalde steekproefopzet dat een significante systematische fout heeft is ook in Tabel 23 weergegeven. Het risico op een systematische fout bij een individuele steekproef is bij een waarnemingsdichtheid van één locatie per km² het grootst; vier tot vijf van de tien keer. Dit is overigens de waarnemingsdichtheid die wordt gehanteerd bij een gangbare regionale Gd-kartering

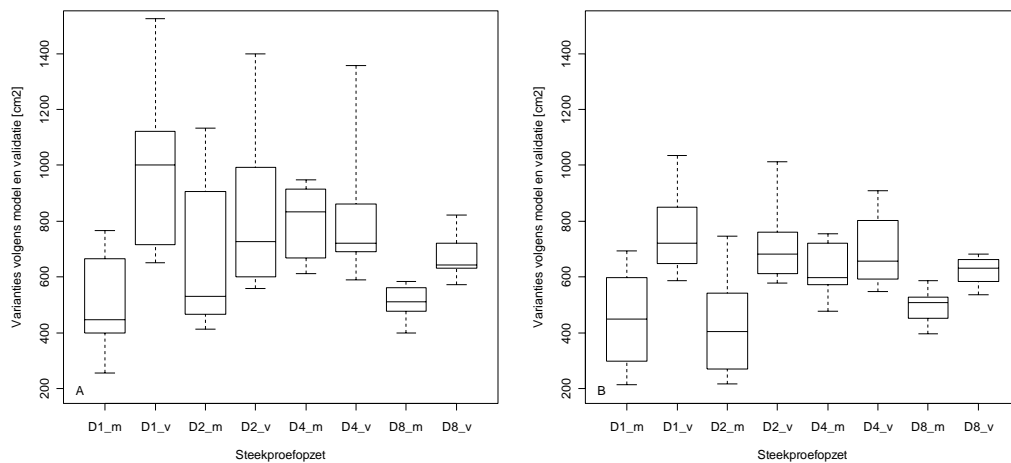
Tabel 23 Systematische fout en standaardafwijking van de systematische fout per steekproefopzet en het aantal maal van de tien dat een significante systematische fout optreedt. Het cijfer staat voor de waarnemingsdichtheid [aantal per km²] en _A of _P staat voor respectievelijk allocatie proportioneel naar het areaal van de Gt-klassen of preferente allocatie in natte Gt-klassen

Steekproefopzet	D1_A	D1_P	D2_A	D2_P	D4_A	D4_P	D8_A	D8_P
Gemiddelde systematische fout (AME)	4.01	5.22	3.43	4.62	3.24	3.26	2.57	2.87
Standaardafwijking v/d systematische fout (ASME)	3.44	3.09	3.14	2.98	3.13	2.94	2.85	2.73
Aantal steekproeven met significante ($\alpha=0.05$) systematische fouten	4	5	3	4	4	2	2	0

Figuur 17 geeft de validatieresultaten voor de systematische fouten grafisch weer. De resultaten van de validatie bij proportionele allocatie staan in Figuur 17a en voor preferente allocatie in natte Gt's in Figuur 17b. De spreiding in de systematische fout, zoals die optreedt over de tien steekproeven van een steekproefopzet, vermindert bij een toenemende waarnemingsdichtheid. Bij de laagste waarnemingsdichtheid, die overeenkomt met de gangbare dichtheid bij recente Gd-karteringen, kunnen de systematische fouten oplopen tot 19 cm.



Figuur 17 Systematische fouten in voorspellingen op validatielocaties voor verschillende waarnemingsdichtheden bij proportionele allocatie over Gt (a) en bij preferente allocatie in natte Gt's (b). De spreiding tussen de tien steekproeven is met box plots weergegeven met de mediaan in het midden, en het 1e en 3e kwartiel als begrenzing van de box. De uiteinden geven het minimum en maximum aan.



Figuur 18 Precisie volgens het model (*m*) en variantie van de fout uit validatie (*v*) weergegeven per steekproefopzet voor proportionele allocatie (18a) en voor preferente allocatie in natte Gt-classes (18b)

6.4.2 Toevallige fout

Figuur 18 geeft de validatieresultaten voor de toevallige fouten weer, zowel voor preferente allocatie in natte Gt's als voor proportionele allocatie. De variantie van de fout varieert van 500 tot 1500 cm² tussen de individuele steekproeven in een steekproefopzet. Dit geeft aan dat de precisie van de voorspellingen sterk afhankelijk is van de locaties van de aselekt gelote steekproefpunten. Bij steekproefopzetten met een hogere waarnemingsdichtheid neemt de variantie van de fout af, zodat de nauwkeurigheid van voorspellingen toeneemt. Ook de spreiding in de variantie van de fout neemt af bij toenemende waarnemingsdichtheid (Figuur 18).

6.4.3 Modelschattingen van de voorspelnaauwkeurigheid

In Tabel 24 staan de resultaten van de validatie van de modelschattingen van de voorspelnaauwkeurigheid. Uit de χ^2 -toets voor varianties van de Z-scores blijkt dat bij steekproefopzetten met waarnemingsdichtheden van één en twee waarnemingen per km² (D1_ en D2_) de nauwkeurigheid significant ($\alpha=0.05$) door het model wordt overschat. Dit blijkt eveneens voor de steekproefopzetten met een waarnemingsdichtheid van acht waarnemingen per km² (D8_). Als gekeken wordt naar het aantal van de tien steekproeven per steekproefopzet waarvoor een significante overschatting van de nauwkeurigheid optreedt, dan valt op dat dit aantal afneemt bij toenemende waarnemingsdichtheid. Het risico op een significante overschatting van de nauwkeurigheid is voor een individuele steekproef bij een waarnemingsdichtheid van één locatie per km² het grootst; namelijk zes tot acht van de tien keer. Bij deze waarnemingsdichtheid, die overigens wordt gehanteerd bij een gangbare regionale Gd-kartering, wordt de nauwkeurigheid met een factor twee door het model overschat.

Tabel 24 Variantie van Z-score, χ^2 -toetsingsgrootte en het aantal steekproeven met een significante overschatting van de nauwkeurigheid per steekproefopzet

Steekproefopzet	D1_A	D1_P	D2_A	D2_P	D4_A	D4_P	D8_A	D8_P
VAR(Z-score)	2.144	2.097	1.262	1.914	0.947	1.052	1.221	1.215
χ^2	1621.1	1537.6	975.4	1503.4	725.2	821.1	955.0	956.2
Aantal steekproeven met verschil ($\alpha=0.05$)	8	6	5	7	2	3	2	3
Aantal steekproeven met verschil ($\alpha=0.01$)	8	5	2	5	0	1	1	1

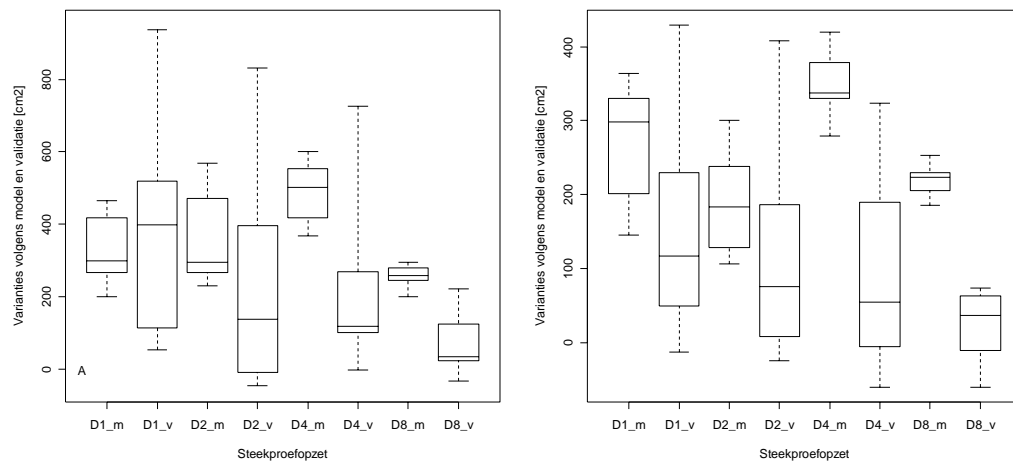
De overschatting van de nauwkeurigheid wordt waarschijnlijk veroorzaakt doordat bij het standaardiseren van het variogram op basis van de restvariantie van het geselecteerde regressiemodel niet met het juiste aantal vrijheidsgraden rekening wordt gehouden. Vooralsnog is echter niet duidelijk hoe het juiste aantal vrijheidsgraden kan worden bepaald op basis van groepsgewijze modelselectie van maximaal vijf predictoren uit de 15 beschikbare predictoren. Uit de variantie van de Z-score blijkt dat vooral bij lagere waarnemingsdichtheden de overschatting van het aantal vrijheidsgraden van invloed is op een juiste schatting van de nauwkeurigheid van voorspellingen.

6.4.4 Correctie voor de meetfout in GxG

In recente Gd-karteringen wordt bij de gebiedsdekkende voorspelling van de GLG gecorrigeerd voor zogenaamde pseudomeetfouten, door middel van de gekwadrateerde voorspelfout van de GLG-voorspellingen uit stambuisregressie op meetlocaties, $VAR(mf)$. Deze correctie wordt uitgevoerd door de nuggetparameter van het variogram model te verminderen met de gemiddelde variantie van de pseudomeetfout op alle gebruikte meetlocaties. Als op basis van dit gecorrigeerde variogram krigingvoorspellingen en -varianties worden berekend spreken we over ‘voor meetfout gecorrigeerde’ voorspellingen en krigingvarianties. Bij validatie moet dan op de validatielocaties ook een correctie voor deze pseudomeetfout plaatsvinden. Dit gebeurt als volgt:

$$VE_c = VE - VAR(mf) \quad (6.15)$$

Uit de validatie blijkt dat bij de lagere waarnemingsdichtheden de nauwkeurigheid redelijk wordt ingeschat, terwijl bij hogere waarnemingsdichtheden een overschatting optreedt (Figuur 19).



Figuur 19 Variantie volgens het model ($_m$) en variantie van de fout uit validatie ($_v$) na correctie voor meetfouten per steekproefopzet voor proportionele allocatie in Gt-klassen (19a) en voor preferente allocatie in natte Gt-klassen (19b)

Opvallend is dat na correctie voor meetfouten de geschatte modelvariantie groter is dan de variantie van de fout die uit de validatie volgt (zie Figuur 19), terwijl zonder correctie voor meetfouten het omgekeerde het geval is (zie Figuur 18). De reden hiervan is dat het variogram niet in alle gevallen voor de gehele pseudomeetfout wordt gecorrigeerd. Als de nuggetparameter van het variogram door vermindering met de gemiddelde variantie van de pseudomeetfout negatief zou worden, dan is deze namelijk op nul gezet. In deze studie bleek de gemiddelde variantie van de pseudomeetfout in alle steekproeven groter dan de oorspronkelijke *nugget* van het variogrammodel, waardoor altijd onvolledige correctie plaats vond. Bij correctie voor meetfouten op de validatielocaties wordt de variantie van de fout echter wel altijd verminderd met de gemiddelde variantie van de pseudomeetfout (zie formule 6.15). De reden dat de variantie van de fout uit validatie en de modelvariantie bij de laagste

waarnemingsdichtheid toch redelijk overeenkomen heeft te maken met de verhouding tussen de nuggetparameter van het variogram en de gemiddelde variantie van de pseudomeetfout. In andere studiegebieden hoeft geen redelijke overeenkomst op te treden. Het lijkt erop dat door correctie voor meetfouten onbedoeld de overschatting van de nauwkeurigheid ten gevolge van de overschatting van het aantal vrijheidsgraden bij de selectie van regressiemodel (zie § 6.3.3.) wordt gecompenseerd. Correctie is dus wel nodig, maar het lijkt erop dat hier sprake is van overcompensatie, veroorzaakt door de aanname dat de pseudomeetfouten ruimtelijk ongecorrleerd zijn.

6.4.5 Proportionele of preferente allocatie per Gt-klasse

In Figuur 17 valt op dat de systematische fouten bij preferente allocatie in natte Gt's (Figuur 17b) kleiner zijn dan bij proportionele allocatie (Figuur 17a). Ook de spreiding in de systematische fouten is bij een vergelijkbare waarnemingsdichtheid kleiner bij preferente allocatie. Het lijkt erop dat, bij gelijke waarnemingsdichtheid, de kans op een systematische fout bij een steekproef met proportionele allocatie over de Gt's groter is dan bij een steekproef met preferente allocatie in natte Gt's. Bij waarnemingsdichtheden van één, twee en vier waarnemingen per km² is ook de variantie van de fout bij preferente allocatie in natte Gt's (Figuur 18b) lager dan bij areaal proportionele allocatie (Figuur 18a). Een verklaring voor deze bevindingen is niet gevonden.

6.5 Discussie

Bij de interpretatie van de resultaten die in dit hoofdstuk zijn gepresenteerd moet worden opgemerkt dat het gehele experiment is gebaseerd op de gebiedsdekkende voorspelling van alleen GLG binnen één deelgebied. Bij de keuze van het deelgebied is wel gekeken naar voldoende variatie aan bodemtypen en variatie in hoogteligging. In een gangbare regionale Gd-kartering worden ongeveer 40 - 60 strata tegelijk gekarteerd voor zowel GHG, GVG als GLG, zodat meer variatie in gebiedskenmerken verwacht mag worden. Het aantal extra metingen die nodig zijn voor dit experiment liet het echter niet toe een groter studiegebied bestaande uit meerdere deelgebieden te gebruiken.

De resultaten van dit experiment voor wat betreft geselecteerde predictoren en modelvarianties komen in een Gd-kartering van een groter gebied frequent voor. Dit geeft aan dat het om een 'normaal' deelgebied gaat. Het lijkt daarom acceptabel de uitkomsten, zij het met enige terughoudendheid, te gebruiken om meer algemene conclusies te trekken.

6.6 Conclusies

- Systematische fouten zijn voor alle steekproefopzetten positief (ongeveer vier cm), wat aangeeft dat de GLG-voorspellingen gemiddeld te droog zijn. De systematische fout kan oplopen tot 19 cm. Het risico op een significante systematische fout is bij een waarnemingsdichtheid van één locatie per km² het grootst; vier tot vijf van de tien keer. Het risico op significante systematische fouten vertoont een afname bij toenemende waarnemingsdichtheid tot tussen nul en twee van de tien keer bij een waarnemingsdichtheid van acht per km². De spreiding in systematische fouten neemt ook af bij een toenemende waarnemingsdichtheid.
- De variantie van de fouten varieert van 500 tot 1500 cm² tussen de 20 steekproeven met een waarnemingsdichtheid van één per km². Dit geeft aan dat de nauwkeurigheid van voorspellingen bij deze waarnemingsdichtheid sterk afhankelijk is van de ligging van de steekproefpunten. Bij steekproefopzetten met hogere waarnemingsdichtheden neemt de variantie van de fout af tot tussen 500 en 800 cm² bij een waarnemingsdichtheid van acht per km².
- Bij steekproefopzetten met een waarnemingsdichtheid van één, twee en acht waarnemingen per km² blijkt een significante ($\alpha=0.05$) overschatting van de nauwkeurigheid op te treden.
- Het risico op een significante overschatting van de nauwkeurigheid is voor een individuele steekproef met een waarnemingsdichtheid van één locatie per km² het grootst; zes tot acht van de tien keer. Bij deze waarnemingsdichtheid, die wordt gehanteerd bij een standaard-Gd-kartering, wordt de nauwkeurigheid in dit onderzoek met een factor twee overschat.
- Uit het verschil tussen de modelschatting van de nauwkeurigheid en de nauwkeurigheid die uit validatie wordt berekend, blijkt dat vooral voor lagere waarnemingsdichtheden de overschatting van het aantal vrijheidsgraden bij selectie van het regressiemodel een nadelige invloed heeft op de schatting van de nauwkeurigheid van de voorspellingen. Vooralsnog is echter niet duidelijk hoe het juiste aantal vrijheidsgraden kan worden bepaald op basis van groepsgewijze modelselectie van maximaal vijf predictoren uit de 15 beschikbare predictoren.
- Na correctie voor 'pseudomeetfouten' in de geschatte GLG op gerichte opnamelocaties (feitelijk zijn dit de voorspelffouten ten gevolge van de stambuisregressie) is de modelvariantie hoger dan de variantie die uit de validatie volgt. Indien correctie achterwege blijft dan treedt het omgekeerde op.
- Bij de correctie voor meetfout is sprake van overcompensatie, veroorzaakt door de aanname dat de pseudo-meetfouten ruimtelijk ongecorrleerd zijn. Bij lage waarnemingsdichtheden wordt het effect van deze overcompensatie deels opgeheven door het effect van het te grote aantal gehanteerde vrijheidsgraden, zodat de variantie uit validatie en de modelvariantie bij de laagste waarnemingsdichtheden toch redelijk met elkaar overeenkomen.
- Het lijkt erop dat, bij eenzelfde waarnemingsdichtheid, de kans op een systematische fout bij een steekproef met proportionele allocatie over de Gt's groter is dan bij een steekproef met preferente allocatie in natte Gt's. Ook een overschatting van de nauwkeurigheid lijkt bij preferente allocatie minder vaak op te treden. Een verklaring hiervoor is niet gevonden.

7 Conclusies en aanbevelingen

In dit project is onderzocht hoe de Gd-methodiek kan worden verbeterd. Uit een uitgebreide inventarisatie van mogelijkheden tot verbetering zijn de volgende drie als kansrijk beoordeeld en nader onderzocht:

1. tijdreeksmodellering van grondwaterstandsreeksen voor de voorspelling van klimaatsrepresentatieve GxG;
2. stambuisregressie voor de omzetting van gemeten grondwaterstanden naar klimaatsrepresentatieve GxG;
3. vlakdekkende voorspelling van klimaatsrepresentatieve GxG.

Op basis van de resultaten van de drie deelonderzoeken kan het volgende worden geconcludeerd en aanbevolen:

- het toepassen van niet-lineaire in plaats van lineaire tijdreeksmodellen leidt niet tot betere voorspellingen van de GxG;
- de keuze van het tijdstip voor de gerichte opname van de grondwaterstand heeft veel meer invloed op de kwaliteit van de GxG-voorspellingen dan de selectie van stambuizen. Een goed criterium voor het kiezen van een geschikte dag voor gerichte opname van de grondwaterstand blijkt de restvariantie van het gefitte stambuisregressiemodel te zijn. Daarnaast is het van belang dat in de dagen voorafgaand aan de gerichte opname geen neerslag is gevallen;
- kanssteekproeven voor de selectie van stambuizen zijn aan te bevelen boven gerichte selectie, omdat de laatste leidt tot systematische fouten in de geschatte regressiecoëfficiënten en de GxG-voorspellingen. Bij kanssteekproeven treden deze fouten niet op;
- zowel bij gestratificeerde kanssteekproeven met niet-evenredige allocatie als bij gerichte selectie wordt de onzekerheid over de werkelijke GxG onderschat (de modelvariantie is kleiner dan de variantie van de voorspelfout die door validatie is vastgesteld). Gestratificeerde kanssteekproeven met proportionele allocatie leveren daarentegen zuivere schattingen van de variantie van de voorspelfout op en zijn daarom aan te bevelen boven gerichte selectie en steekproeven met niet-evenredige allocatie;
- de nauwkeurigheid waarmee de onzekerheid over de GxG wordt gekwantificeerd neemt toe met het aantal geselecteerde stambuizen. Bij 20 stambuizen is de nauwkeurigheid laag (variatiëcoëfficiënt 0.55). Een aantal van tenminste 40 stambuizen wordt aanbevolen als nauwkeurige kwantificering van de onzekerheid belangrijk is (variatiëcoëfficiënt ca. 0.3);
- het aantal gerichte-opnamelocaties per km² is van invloed op de grootte van zowel de systematische en de toevallige fout in de GxG-voorspellingen. Bij een dichtheid van één per km² (die tot nog toe gebruikt werd) kunnen systematische fouten zelfs oplopen tot ca. 19 cm en kan de variatie van de fouten 500 à 1500 cm² bedragen. Aanbevolen wordt de dichtheid van de gerichte-opnamelocaties te verhogen tot de gewenste nauwkeurigheid wordt bereikt;

- de nauwkeurigheid van de GxG-voorspellingen blijkt met een factor twee te worden overschat bij een dichtheid van één gerichte-opnamelocatie per km². Naarmate het aantal gerichte-opnamelocaties toeneemt, wordt de nauwkeurigheid van de GxG-voorspellingen beter geschat (betere kwantificering van de onzekerheid);
- de kwantificering van de onzekerheid kan verder worden verbeterd door 'pseudo-meetfouten' (fouten in de GxG-voorspellingen op gerichte-opnamelocaties) te verdisconteren, en door het juiste aantal vrijheidsgraden bij de selectie van het regressiemodel voor de relatie tussen GxG en hulpinformatie in rekening te brengen. Naar dit laatste is nader onderzoek noodzakelijk.

Ten slotte dient te worden opgemerkt dat bij deze studie geen fundamentele fouten in de Gd-methode aan het licht zijn gekomen. Een grondige herziening van de methode is dan ook niet nodig. Met de relatief kleine aanpassingen die hier worden aanbevolen kan de GxG nauwkeuriger worden voorspeld en de onzekerheid over de werkelijke GxG beter worden gekwantificeerd.

Literatuur

- Bierkens, M., M. Knotters en F. van Geer (1999). Tijdreeksanalyse nu ook toepasbaar bij onregelmatige meetfrequenties. *Stromingen* 5(2):43–54.
- Bierkens, M.F.P., W.A. Bron en M. Knotters (2002). *VIDENTE 1.1: a graphical user interface and decision support system for stochastic modelling of water table fluctuations at a single location*. Alterra rapport 613, Wageningen.
- Bracht, M. van (1988). *OLGA: On Line Groundwater Archief*. Rapport PN88-11, DGV-TNO.
- Brus, D.J., J.J. de Gruijter, en J.W. van Groenigen (2003). Designing spatial coverage samples by the k-means clustering algorithm. In: *Proceedings of the 8th International FZK/TNO conference on contaminated soil (ConSoil 2003)*, Gent, Belgium, May 2003, 504–509.
- Cochran, W.G. (1977). *Sampling techniques*. Wiley, New York.
- Cohen, A. (1991). *Truncated and censored samples: theory and applications*. Dekker Inc., New York.
- Davis, J. (2002). *Statistics and data analysis in geology*. Wiley, New York, third edition.
- De Bruin, H. (1987). Van Penman naar Makkink. In Hooghart, J., redactie, *Neerslag en Verdamping*, CHO-TNO Mededeling 39. CHO-TNO, Den Haag.
- De Gruijter, J.J., J.B.F. van der Horst, G.B.M. Heuvelink, M. Knotters en T. Hoogland (2004). *Grondwater opnieuw op de kaart. Methodiek voor de actualisering van grondwaterstands-informatie en perceelsclassificatie naar uitspoelingsgevoeligheid voor nitraat*. Rapport 915, Alterra.
- Finke, P., M. Bierkens, D. Brus, J. van der Gaast, T. Hoogland, M. Knotters en F. de Vries (2002). *Klimaatrepresentatieve grondwaterdynamiek in Waterschap Mark en Weerijds*. Rapport 387, Alterra.
- Hoogland, T., Finke, P.A., De Vries, F. (2003). *Actualisatie grondwatertrappenkaart waterschap Rijn en IJssel*. Rapport 126, Alterra.
- Knotters, M. (2001). *Regionalised time series models for water table depths*. PhD thesis, Wageningen Universiteit.
- Knotters, M. en M. Bierkens (1999). Hoe lang moet je de grondwaterstand meten om iets over de dynamiek te weten? *Stromingen* 5(4):5–12.

- Knotters, M. en P. van Walsum (1994). *Uitschakeling van weersinvloeden bij de karakterisering van het grondwaterstandsverloop*. Rapport 350, Staring Centrum.
- Mallows, C.L. (1966). *Choosing a Subset Regression*. Joint Statistical Meetings, Los Angeles, CA 1966.
- Marsman, B.A.. en J.J. de Gruijter (1986). *Quality of soil maps; a comparison of soil survey methods in a sandy area*. Soil survey papers No. 15, Soil survey institute, Wageningen.
- Oude Voshaar, J. en J. Stolp, (1997). *Schatting van GHG en GLG van tijdelijke peilbuizen met korte meetreeksen*. Technisch Document 30, Staring Centrum.
- Pebesma E. en C. Wesseling (1998). Gstat, a program for geostatistical modelling, prediction and simulation. *Computers and Geosciences* **24**(1): 17-31; <http://www.gstat.org>
- Riele, W. te en D. Brus (1992). *Het gebruik van fysisch-geografische voorinformatie bij de ruimtelijke voorspelling van grondwaterstanden en grondwaterkarakteristieken (GHG en GLG)*. Rapport 209, Staring Centrum.
- Riele, W. te, E. Querner, M. Knotters en A. Pomper (1995). *Geostatistische interpolatie van grondwaterstandsdiepten met behulp van fysisch-geografische informatie en de resultaten van een regionaal stromingsmodel*. Rapport 414, Staring Centrum.
- Särndal, C-E. (1980). On pi-inverse weighting versus best linear unbiased weighting in probability sampling. *Biometrika* **67**: 639-650.
- Sluijs, P. van der (1982). De grondwatertrap als karakteristiek van het grondwaterstandsverloop. *H₂O* **15**: 42-46.
- Sluijs, P. van der (1990). Hoofdstuk 11: Grondwatertrappen. In: Locher, W.P. en H. de Bakker (red.) *Bodemkunde van Nederland deel 1: Algemene bodemkunde*. Malmberg Den Bosch.
- Snedecor, G.W. en Cochran, W.G. (1989). *Statistical Methods*. Iowa State University Press, Ames Iowa, Eight edition.
- Wesseling, J.G. (1991). *CAPSEV: steady state moisture flow theory; program description and user manual*. Winand Staring centre, Report 37, Wageningen.
- Wösten, J.H.M., F. de Vries, J. Denneboom en A.F. van Holst (1988). *Generalisatie en bodemkundige vertaling van de bodemkaart 1 : 250.000, ten behoeve van de PAWN-studie*. Rapport 2055, Stiboka. Wageningen.