



ALTEERRA
WAGENINGEN UR



Validatie van grondwaterstandskaarten met de Landelijke Steekproef Kaarteenheden

Alterra-rapport 2440
ISSN 1566-7197

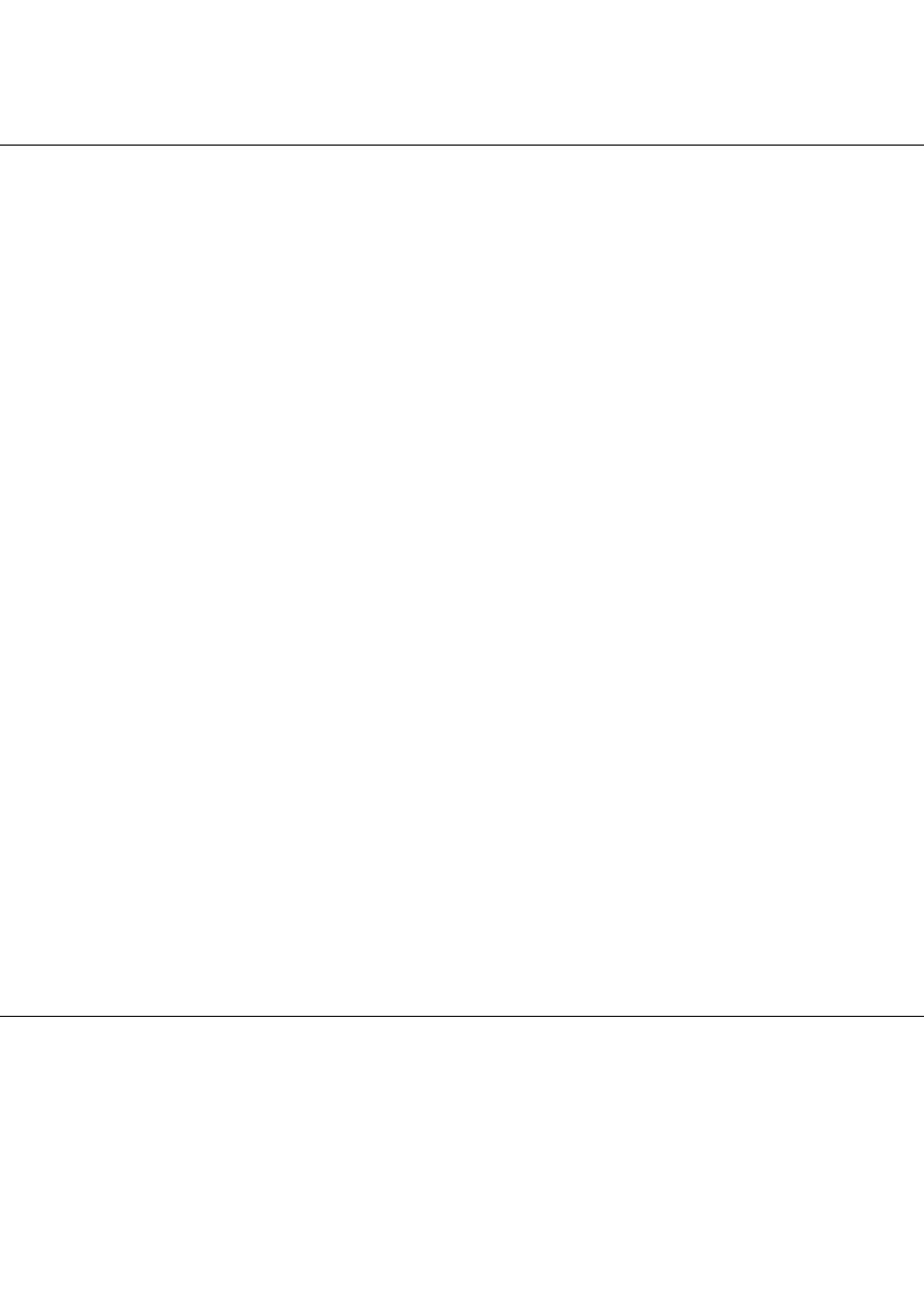
M. Knotters, T. Hoogland en D.J. Brus



ALTEERRA
WAGENINGEN UR

stowa

Validatie van grondwaterstandskarten met de
Landelijke Steekproef Kaarteenheden



Validatie van grondwaterstandskarten met de Landelijke Steekproef Kaarteenheden

M. Knotters, T. Hoogland en D.J. Brus

Alterra-rapport 2440

Alterra, onderdeel van Wageningen UR
Wageningen, 2013



Referaat

Knotters, M., T. Hoogland en D.J. Brus, 2013. *Validatie van grondwaterstandskarten met de Landelijke Steekproef Kaarteenheden*. Wageningen, Alterra, Alterra-rapport 2440. 65 blz.; 34 fig.; 21 tab.; 13 ref.

Kaarten met gemiddeld hoogste en laagste grondwaterstanden (GxG's) kunnen worden berekend met modellen zoals het Nationaal Hydrologisch Instrumentarium (NHI) en de hydrologische component van het modelinstrumentarium Stone V2.3 voor nutriënten-uitspoeling. De kaart van de grondwaterdynamiek (Gd-kaart) geeft naast GxG's ook standaardafwijkingen van de voorspelfouten. Met de GxG-schattingen uit de Landelijke Steekproef Kaarteenheden (LSK) zijn de GxG-voorspellingen van NHI 3.0, Hydrologie Stone V2.3 en Gd-kaart gevalideerd. GxG-voorspellingen van NHI 3.0 zijn gemiddeld relatief diep, vooral onder stuwwallen, de Hondsrug, de Brabantse Wal en de Maasterrassen. GxG-voorspellingen van Hydrologie Stone V2.3 zijn gemiddeld relatief ondiep, vooral onder stuwwallen, de Hondsrug en de Maasterrassen. GHG-voorspellingen van de Gd-kaart zijn gemiddeld relatief ondiep. NHI 3.0 onderschat de amplitude van seizoensfluctuatie van de freatische grondwaterstand in Oost-Nederland, en overschat deze amplitude in West-Nederland. Hydrologie Stone V2.3 en de Gd-kaart overschatten de amplitude van seizoensfluctuatie gemiddeld iets. Aanbevolen wordt om validatiecriteria te definiëren die verband houden met het gebruik van de modelvoorspellingen, en de validatie te baseren op een onafhankelijke set gemeten grondwaterstanden die zijn gemeten op locaties die volgens een kanssteekproef zijn geselecteerd.

Trefwoorden: freatische grondwaterstand, grondwaterdynamiek, kanssteekproef, NHI, Stone, stratificatie, validatie

ISSN 1566-7197

Dit rapport is gratis te downloaden van www.alterra.wur.nl (ga naar 'Alterra-rapporten'). Alterra Wageningen UR verstrekt geen gedrukte exemplaren van rapporten. Gedrukte exemplaren zijn verkrijgbaar via een externe leverancier. Kijk hiervoor op www.rapportbestellen.nl.

© 2013 Alterra (instituut binnen de rechtspersoon Stichting Dienst Landbouwkundig Onderzoek)
Postbus 47; 6700 AA Wageningen; info.alterra@wur.nl

- Overname, verveelvoudiging of openbaarmaking van deze uitgave is toegestaan mits met duidelijke bronvermelding.
- Overname, verveelvoudiging of openbaarmaking is niet toegestaan voor commerciële doeleinden en/of geldelijk gewin.
- Overname, verveelvoudiging of openbaarmaking is niet toegestaan voor die gedeelten van deze uitgave waarvan duidelijk is dat de auteursrechten liggen bij derden en/of zijn voorbehouden.

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

Alterra-rapport 2440
Wageningen, juni 2013

Inhoud

Woord vooraf	7
Samenvatting	9
1 Inleiding	13
1.1 Probleemstelling	13
1.2 Achtergrond	13
1.3 Doel en onderzoeksvragen	14
1.4 Leeswijzer	14
2 Materialen en methode	15
2.1 Landelijke Steekproef Kaarteenheden (LSK)	15
2.2 GxG-voorspellingen van NHI, Hydrologie Stone en Gd-kaart	16
2.3 Aanpak van validatie met LSK en vergelijking met Gd-kaart	18
2.4 Validatiematen	21
3 Resultaten	23
3.1 Vergelijking van NHI-voorspellingen met LSK	23
3.1.1 Resultaten landelijk	23
3.1.2 Resultaten per Gt-klasse	27
3.1.3 Resultaten per deelgebied van de LSK	28
3.1.4 Resultaten per toetsingsregio van NHI 3.0	28
3.2 Vergelijking van Hydrologie-Stone-voorspellingen met LSK	30
3.2.1 Resultaten landelijk	30
3.2.2 Resultaten per Gt-klasse	33
3.2.3 Resultaten per deelgebied van de LSK	34
3.2.4 Resultaten per toetsingsregio van NHI 3.0	34
3.3 Vergelijking van Gd-voorspellingen met LSK	35
3.4 Vergelijking van NHI-voorspellingen met Gd-voorspellingen	38
3.5 Vergelijking van Hydrologie-Stone-voorspellingen met Gd-voorspellingen	41
4 Conclusies en aanbevelingen	43
4.1 Conclusies	43
4.2 Aanbevelingen	46
Literatuur	47
Bijlagen	
1 Validatieresultaten NHI 3.0 per LSK-stratum	49
2 Validatieresultaten Hydrologie Stone V2.3 per LSK-stratum	53
3 Cumulatieve frequentieverdelingen van verschillen LSK - GxG-modellen	57
4 Histogrammen van GxG's uit LSK, NHI 3.0 en Hydrologie Stone V2.3	63

Woord vooraf

Dit rapport doet verslag van een validatiestudie, die voortkwam uit een initiatief van de Stichting Toegepast Onderzoek Waterbeheer (Stowa). Aanvankelijk beperkte de studie zich tot validatie van grondwaterstands-informatie die het Nationaal Hydrologisch Instrumentarium (NHI) genereert. Na presentatie van de eerste resultaten verzocht de projectgroep NHI de studie op een aantal onderdelen uit te breiden, en ook de grondwaterstands-informatie te valideren die gegenereerd wordt met Hydrologie Stone. De validatie is uitgevoerd op basis van de grondwaterstands-informatie uit de Landelijke Steekproef Kaarteenheden (LSK). Ruimtelijke patronen van NHI- en Hydrologie-Stone-voorspellingen zijn vergeleken met de kaart van de Grondwaterdynamiek (Gd-kaart), die daartoe ook gevalideerd is met behulp van de LSK.

Wij zijn een aantal mensen dank verschuldigd voor hun initiatief tot deze studie, voor hun kritische blik op de concepten van dit rapport, en voor hun opbouwende commentaar dat geleid heeft tot deze uiteindelijke publicatie: Joost Heijkers (Hoogheemraadschap de Stichtse Rijnlanden), Durk Klopstra (Stowa), Timo Kroon (Waterdienst), Jacco Hoogewoud (Deltares), Frank van der Bolt (Alterra Wageningen UR), Ab Veldhuizen (Alterra Wageningen UR) en Henk Vroon (Alterra Wageningen UR).

Wageningen, juni 2013

Martin Knotters, Tom Hoogland en Dick Brus

Samenvatting

Inleiding

Met het Nationaal Hydrologisch Instrumentarium en met de hydrologische component van het modelinstrumentarium voor nutriëntenuitspoeling Stone kunnen onder meer gemiddeld hoogste, laagste en voorjaars-grondwaterstanden (resp. GHG, GLG, GVG, samen GxG) ruimtelijk worden voorspeld. De kaart van de grondwaterstandsdynamiek (Gd) geeft voorspellingen van GxG's en standaardafwijkingen van de voorspelfouten. Er is zowel behoefte aan een algemene indicatie van de kwaliteit van deze GxG-voorspellingen, als aan een ruimtelijk beeld van de kwaliteit van voorspellingen.

Het doel van deze validatiestudie is om de kwaliteit van GxG-voorspellingen van NHI 3.0 en Hydrologie Stone V2.3 vast te stellen op basis van validatie met onafhankelijke waarnemingen uit de Landelijke Steekproef Kaarteenheden (LSK). Daarnaast zullen GxG-voorspellingen van NHI 3.0 en Hydrologie Stone V2.3 worden vergeleken met GxG-voorspellingen uit de Gd-karteringen, om inzicht te krijgen in de nauwkeurigheid op het niveau van specifieke cellen. Om te beoordelen of de Gd-kaart voor dit doel kan worden gebruikt zal ook de Gd-kaart worden gevalideerd aan de hand van de LSK.

Materialen en methode

Het bestand van de Landelijke Steekproef Kaarteenheden (LSK) bevat schattingen van de GHG en GLG voor 1193 locaties, die volgens een gestratificeerde enkelvoudig aselechte steekproef zijn geselecteerd. Bij de LSK zijn deelgebieden (strata) onderscheiden op basis van bodemtype en grondwatertrappen, zodat zowel een goede spreiding over verschillende bodemkundige en hydrologische situaties als een goede ruimtelijke spreiding is bewerkstelligd. De GxG's van deze locaties zijn geschat op basis van gerichte opnamen en stambuisregressie gedurende de periode 1988-2001. De standaardafwijking van de fout in GxG varieert van 10 tot 25 cm. Verder kunnen door *regression to the mean* ondiepe GxG's te diep worden geschat en diepe GxG's te ondiep, waardoor de ruimtelijke variatie in GxG's mogelijk wordt onderschat.

Met het NHI 3.0 en Hydrologie Stone V2.3 zijn GxG's berekend uit ruimtelijke voorspellingen van het grondwaterstandsverloop voor de periode 1997-2005, die samenvalt met de periode waarin de inventarisaties voor de kaart van de grondwaterstandsdynamiek (Gd) plaatsvonden. De NHI- en Hydrologie-Stone-voorspellingen van de GxG's zijn gemaakt voor een grid van 250x250 m.

De Gd-kaart heeft een resolutie van 25x25 m. De bestanden behorende bij de Gd-kaarten bevatten naast GxG-voorspellingen ook standaardafwijkingen van de voorspelfouten. De GxG's van de Gd-kaart zijn geschat door reeksen van acht jaar lang die eindigen in de periode 1997-2005 te verlengen tot reeksen van 30 jaar lang over de periode 1970-2000, waaruit de GxG's zijn geschat als gemiddelde van 30 xG3's (gemiddelden van de drie hoogste of laagste standen in een hydrologisch jaar, bij een halfmaandelijke meetfrequentie).

Uit een analyse van zes lange reeksen blijkt dat het effect van de verschillende perioden waarop GxG-schattingen zijn gebaseerd kan variëren van enkele centimeters tot circa 35 cm, waarbij de grootste verschillen mogelijk veroorzaakt worden door ingrepen in de hydrologie. Over het algemeen zijn de GxG's die zijn geschat volgens de procedure die bij de Gd-karteringen is gevolgd dieper dan de GxG's die zijn geschat op basis van waarnemingen uit de periode 1997-2005.

De aanpak van de validatie op basis van de LSK en de vergelijking van de NHI- en Hydrologie-Stone-voorspellingen begint met de definitie van gebieden (interessedomeinen) waarvoor een validatie en vergelijking

wordt uitgevoerd. Dit is het gebied waarbinnen tijdens de uitvoering van de LSK grondwaterstanden zijn gemeten (geen ontbrekende waarnemingen omdat het grondwater zich te diep bevond). Dit is vrijwel geheel Nederland met uitzondering van gebieden met diepe grondwaterstanden (delen van stuwwallen en duinen). Verder worden een aantal gebieden apart beschouwd: de grondwatertrappen (Gt's) volgens de LSK, de deelgebieden (strata) die bij de LSK zijn onderscheiden (combinaties van grondwatertrappen en bodemtypen), de drie toetsingsregio's die bij toetsing van NHI 3.0 zijn onderscheiden, en het gebied waarvoor een Gd-kaart is gemaakt (voornamelijk het pleistocene deel van Nederland). Vervolgens worden verschilwaarden berekend tussen de GxG's uit het LSK-bestand en de GxG's die zijn voorspeld met het NHI 3.0, Hydrologie Stone V2.3 en de Gd-methode. Met deze verschilwaarden worden de theoretisch berekende varianties van de voorspelfouten van de Gd-kaart gecorrigeerd. Vervolgens worden de NHI-voorspellingen en Hydrologie-Stone-voorspellingen vergeleken met de Gd-kaart, en wordt met de gecorrigeerde varianties van de voorspelfouten aangegeven waar verschillen significant zijn. Grote verschillen in modelvoorspellingen kunnen worden geïnterpreteerd als grote onzekerheid over de werkelijke GxG. Kleine verschillen hoeven echter niet te betekenen dat de onzekerheid over de werkelijke GxG klein is; dit is alleen met validatie vast te stellen.

De kwaliteit van modelvoorspellingen is met verschillende validatiematen uitgedrukt, die inzicht geven in systematische en toevallige verschillen en in de frequentieverdeling van de verschillen.

Resultaten

De NHI-voorspellingen van GxG's zijn gemiddeld dieper dan de LSK-schattingen. De verschillen zijn scheef verdeeld. Voor de GHG is het gemiddelde verschil tussen LSK-schatting en NHI-voorspelling -22 cm, en voor de GLG is dit -17 cm. Voor zowel GHG als GLG is de mediaan van de verschillen -7 cm. Een klein deel van de NHI-voorspellingen van zowel de GHG als de GLG is veel dieper dan de schattingen in de LSK: bij 2.5% van de GHG-voorspellingen is het verschil met LSK groter dan 179 cm en loopt op tot maximaal 1959 cm, en bij 2.5 % van de GLG-voorspellingen is het verschil met LSK groter dan 185 cm en loopt op tot maximaal 1963 cm. Vooral onder stuwwallen, de Hondsrug, de Brabantse Wal (bij Bergen op Zoom) en de Maasterrassen voorspelt NHI 3.0 significant diepere GxG's dan de Gd-kaart, bij een significantieniveau van 5%. De GLG-voorspellingen voor het Montferland vormen hierop een uitzondering: hier voorspelt NHI 3.0 significant ondiepere GLG's dan de Gd-kaart.

De Hydrologie-Stone-voorspellingen van GxG's zijn gemiddeld ondieper dan de LSK-schattingen. Voor de GHG is het gemiddelde verschil tussen LSK-schatting en Hydrologie-Stone-voorspelling 13 cm en is de mediaan van de verschillen 9 cm. Voor de GLG zijn zowel het gemiddelde verschil als de mediaan van de verschillen 7 cm. De GxG-voorspellingen van Hydrologie Stone V2.3 zijn vooral onder stuwwallen, de Hondsrug en de Maasterrassen significant ondieper dan de GxG-voorspellingen volgens de Gd-kaart.

De GHG-voorspellingen van de Gd-kaart zijn systematisch ondieper dan de GHG-schattingen van de LSK. Het gemiddelde van de verschillen is 14 cm, de mediaan is 7 cm. Voor de GLG kan geen systematisch verschil worden aangetoond.

In vergelijking met de LSK onderschat NHI 3.0 de amplitude van seizoensfluctuatie (GLG-GHG) voor het oosten van Nederland, en overschat deze amplitude voor het westen van Nederland. Hydrologie Stone V2.3 en de Gd-kaart geven gemiddeld genomen iets grotere amplitudes van seizoensfluctuatie dan de LSK. Uit de resultaten blijkt dat de variaties van de verschillen (uitgedrukt in SDE) in seizoensfluctuatie tussen enerzijds NHI 3.0, Hydrologie Stone V2.3 en Gd-kaart en anderzijds de LSK in dezelfde grootteorde liggen. Hier is niet direct een verklaring voor te geven.

Conclusies en aanbevelingen

Samenvattend kunnen uit deze studie de volgende conclusies worden getrokken over de nauwkeurigheid van GxG-voorspellingen volgens NHI 3.0, Hydrologie Stone V2.3 en de Gd-kaart:

1. GxG-voorspellingen van NHI 3.0 zijn gemiddeld genomen relatief diep. Diepe voorspellingen komen vooral voor onder stuwwallen, de Hondsrug, de Brabantse Wal en de Maasterrassen.
2. GxG-voorspellingen van Hydrologie Stone V2.3 zijn gemiddeld genomen relatief ondiep. Ondiepe voorspellingen komen vooral voor onder stuwwallen, de Hondsrug en de Maasterrassen.
3. GHG-voorspellingen van de Gd-kaart zijn gemiddeld genomen relatief ondiep.
4. NHI 3.0 onderschat de amplitude van seizoensfluctuatie van de freatische grondwaterstand in Oost-Nederland, en overschat deze amplitude in West-Nederland. Hydrologie Stone V2.3 en de Gd-kaart overschatten de amplitude van seizoensfluctuatie gemiddeld genomen iets.
5. Aan het criterium dat in tenminste 80% van een toetsingsregio de afwijking in GHG, GLG of amplitude van seizoensfluctuatie (dynamiek) niet groter mag zijn dan een bepaalde waarde voldoen de verschillen tussen voorspellingen van zowel NHI 3.0 als Hydrologie Stone V2.3 enerzijds en LSK-schattingen anderzijds in vrijwel geen van de gevallen. Alleen de GLG-voorspellingen van NHI 3.0 in sterk hellende gebieden voldoen aan het criterium van 80%. Aan het criterium dat is gesteld voor de mediaan van de afwijkingen wordt altijd voldaan, zowel bij NHI 3.0 als bij Hydrologie Stone V2.3.

Deze studie leidt tot de volgende aanbevelingen voor validatie van gebiedsdekkende informatie over de freatische grondwaterstand:

1. Voorafgaand aan validatie moeten onderbouwde validatiecriteria worden gedefinieerd die informatie geven over de bruikbaarheid van informatie over de freatische grondwaterstand bij de doelen waarvoor deze informatie wordt ingezet.
2. GxG's worden geschat en bevatten een schattingsfout. Het verdient daarom aanbeveling om niet te valideren op geschatte GxG's, maar op gemeten standen, of op parameters die direct uit gemeten standen zijn berekend, zoals xG3's (gemiddelden van de drie hoogste of laagste standen in een hydrologisch jaar, bij een halfmaandelijke meetfrequentie).
3. Om objectiviteit na te streven verdient het aanbeveling deze standen te meten op locaties die volgens een kanssteekproef zijn geselecteerd. Op deze manier kunnen zuivere schattingen van grootheden zoals gebiedsgemiddelde fouten worden gemaakt en kan de nauwkeurigheid van deze grootheden valide worden gekwantificeerd.
4. Om onafhankelijkheid na te streven verdient het aanbeveling dat de waarnemingen uit dit meetnet uitsluitend voor validatie van modellen van de grondwaterstand worden gebruikt en dus niet voor kalibratie beschikbaar worden gesteld.
5. Het verdient aanbeveling om de validatieset te gebruiken voor het beoordelen van de bruikbaarheid van alle modellen voor het genereren van informatie over de freatische grondwaterstand, inclusief regionale modellen en nieuwe modelversies.
6. Om de validatieset te kunnen inzetten voor beoordeling van regionale modellen verdient het aanbeveling om voor de betreffende gebieden het validatiemeetnet te verdichten.

1 Inleiding

1.1 Probleemstelling

Met het Nationaal Hydrologisch Instrumentarium (NHI; Snepvangers et al., 2008; Hoogewoud et al., 2013a, b) kunnen ondermeer gemiddeld hoogste, laagste en voorjaarsgrondwaterstanden (resp. GHG, GLG, GVG, samen GxG) ruimtelijk worden voorspeld voor elk 250x250m-pixel binnen Nederland, met uitzondering van de Waddeneilanden. Ook met de hydrologische berekeningen voor het modelinstrumentarium voor nutriënten-uitspoeling Stone (Van Bakel et al., 2008) kunnen GxG's gebiedsdekkend worden voorspeld. De kaart van de grondwaterstandsdynamiek (Gd) (Finke et al., 2005; Van Kekem et al., 2005) geeft voorspellingen van GxG's en standaardafwijkingen van de voorspelfouten.

Om de bruikbaarheid van modelvoorspellingen van de freatische grondwaterstand te kunnen beoordelen is het van belang de kwaliteit ervan objectief en onafhankelijk in een validatiestudie vast te stellen. Er is zowel behoefte aan een algemene indicatie van de kwaliteit van de GxG-voorspellingen, als aan een ruimtelijk beeld van de kwaliteit van voorspellingen.

1.2 Achtergrond

Bij validatie wordt gestreefd naar onafhankelijkheid en objectiviteit. Onafhankelijkheid bereiken we in de eerste plaats door modeluitkomsten te vergelijken met informatie die niet is gebruikt bij de modellering, dat wil zeggen bij de kalibratie van het model of één van zijn onderdelen. Verder streven we naar onafhankelijkheid door de validatie te laten uitvoeren door onderzoekers die niet bij de totstandkoming van NHI of Hydrologie Stone betrokken zijn geweest.

Objectiviteit streven we in de eerste plaats na door modelvoorspellingen te vergelijken met 'harde' informatie over de GxG. In principe is de GxG alleen foutloos te berekenen uit een meetreeks met een lengte van 30 jaar, die vrij is van effecten van ingrepen en geen meetfouten bevat. Deze reeksen zijn echter nauwelijks voorhanden, waardoor benaderingen zoals stambuisregressie (Te Riele en Brus, 1991) en tijdreeksmodellering (Knotters en Van Walsum, 1997) nodig zijn. De GxG's in de validatieset hebben hierdoor een bepaalde onnauwkeurigheid, die bekend moet zijn zodat er rekening mee kan worden gehouden bij de presentatie en interpretatie van de validatieresultaten. In de tweede plaats streven we objectiviteit na door de locaties met voorspelde GxG's te selecteren middels een kanssteekproef. Dit maakt een objectieve inschatting van een maat voor de kwaliteit van de GxG-voorspellingen voor heel Nederland mogelijk.

Validatie is vaak niet mogelijk door het ontbreken van een dataset die aan de bovenstaande eisen van onafhankelijkheid en objectiviteit voldoet. Voor validatie van GxG-voorspellingen kunnen echter de GxG-waarden uit de Landelijke Steekproef Kaarteenheden worden gebruikt (LSK, Finke et al., 2001; Visschers et al., 2007). De gegevens uit de LSK zijn niet gebruikt bij de totstandkoming van NHI en Hydrologie Stone, en daarom onafhankelijk. De LSK-gegevens zijn verzameld op locaties die volgens een kanssteekproef zijn geselecteerd. De objectiviteit van de gegevensverzameling is daardoor gegarandeerd, en omdat de selectiekansen bekend zijn kunnen gebiedsgemiddelden etc. van voorspelfouten zuiver worden geschat. Met de gegevens uit de LSK kunnen daarom NHI, Hydrologie Stone en andere modellen onafhankelijk en objectief worden gevalideerd.

1.3 Doel en onderzoeksvragen

Het doel van deze validatiestudie is om de kwaliteit van GxG-voorspellingen van NHI 3.0 en Hydrologie Stone V2.3 vast te stellen op basis van validatie met onafhankelijke waarnemingen uit de Landelijke Steekproef Kaarteenheden (LSK). Daarnaast zullen GxG-voorspellingen van NHI 3.0 en Hydrologie Stone V2.3 worden vergeleken met GxG-voorspellingen uit de Gd-karteringen, om inzicht te krijgen in de nauwkeurigheid op het niveau van specifieke cellen. Om te beoordelen of de Gd-kaart voor dit doel kan worden gebruikt wordt ook de Gd-kaart gevalideerd aan de hand van de LSK.

De volgende onderzoeksvragen worden beantwoord:

1. Hoe groot zijn de verschillen tussen GxG-voorspellingen volgens NHI 3.0, Hydrologie Stone V2.3 en Gd-kaart enerzijds en de GxG-schattingen uit de LSK anderzijds? Deze vraag wordt beantwoord voor geheel Nederland, per Gt-klasse en per bodemkundig-hydrologisch stratum dat is onderscheiden bij de LSK.
2. Waar verschillen GxG-voorspellingen volgens NHI 3.0 en Hydrologie Stone V2.3 van de GxG-voorspellingen volgens de Gd-kaart?
3. Hoe groot zijn de verschillen in amplitude van seizoensfluctuatie (GLG minus GHG) tussen enerzijds NHI-, Hydrologie-Stone- en Gd-voorspellingen en anderzijds LSK-schattingen?
4. In hoeverre beantwoorden de resultaten van de validatie van NHI 3.0 en Hydrologie Stone V2.3 op basis van de LSK aan de toetsingscriteria voor GHG, GLG en dynamiek die Hoogewoud et al. (2013b) hanteren voor toetsing van NHI 3.0?
5. Wat zeggen de validatieresultaten over de nauwkeurigheid waarmee NHI 3.0, Hydrologie Stone V2.3 en de Gd-kaart GxG's voorspellen, gezien de fouten in de GxG-schattingen uit de LSK?
6. Wat zeggen de validatieresultaten over de nauwkeurigheid waarmee NHI 3.0, Hydrologie Stone V2.3 en de Gd-kaart GxG's voorspellen, gezien de verschillen in rekenperioden die zijn gehanteerd bij deze GxG-voorspellingen en bij de GxG-schattingen volgens de LSK?
7. Wat kan op grond van de gevonden verschillen, en gegeven de nauwkeurigheid van de LSK-schattingen en het effect van verschillende rekenperioden worden geconcludeerd over de nauwkeurigheid van GxG-voorspellingen volgens NHI 3.0, Hydrologie Stone V2.3 en de Gd-kaart?

1.4 Leeswijzer

Hoofdstuk 2 beschrijft de gebruikte gegevens en de toegepaste methoden. Paragraaf 2.1 beschrijft de gegevens uit de Landelijke Steekproef Kaarteenheden die zijn gebruikt bij de validatie. Paragraaf 2.2 beschrijft de GxG-voorspellingen van NHI 3.0, Hydrologie Stone V2.3 en de Gd-kaart. De procedure die is gevolgd bij de validatie van GxG-voorspellingen van NHI 3.0, Hydrologie Stone V2.3 en Gd-kaart, en bij de vergelijking van NHI- en Hydrologie-Stone-voorspellingen met de Gd-kaart, staat beschreven in paragraaf 2.3.

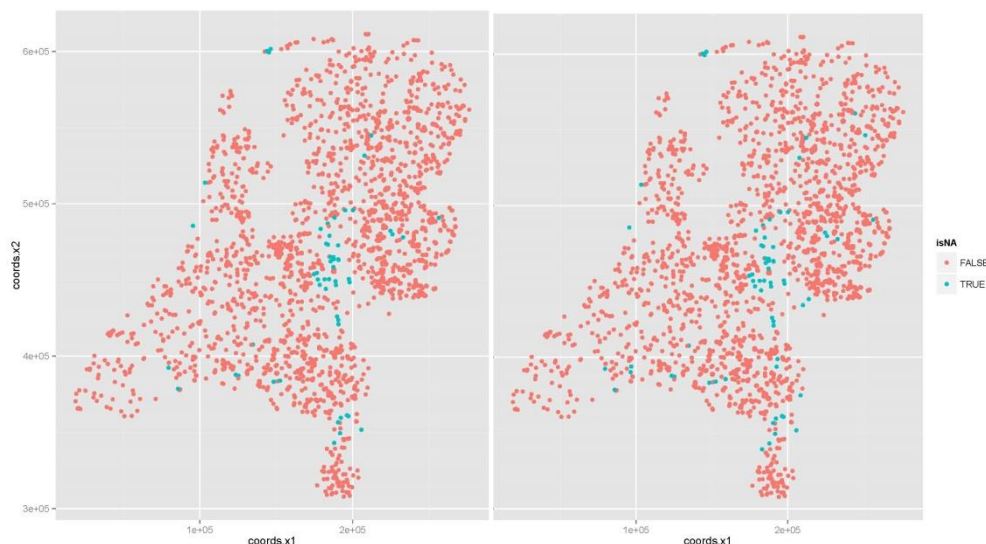
Hoofdstuk 3 geeft de resultaten van de validatie van GxG-voorspellingen met NHI 3.0 (paragraaf 3.1), Hydrologie Stone V2.3 (paragraaf 3.2) en Gd-kaart (paragraaf 3.3). Paragraaf 3.4 beschrijft de verschillen tussen GxG-voorspellingen met NHI 3.0 en de Gd-kaart, en paragraaf 3.5 beschrijft de verschillen tussen Hydrologie-Stone-voorspellingen van GxG's en de Gd-kaart.

Hoofdstuk 4 beschrijft de conclusies die uit deze validatie en vergelijking konden worden getrokken, en aanbevelingen voor validatie van modellen en kaarten van grondwaterstanden.

2 Materialen en methode

2.1 Landelijke Steekproef Kaarteenheden (LSK)

Het bestand van de Landelijke Steekproef Kaarteenheden (LSK) bevat schattingen van de GHG en GLG voor 1193 locaties, die volgens een gestratificeerde enkelvoudig aselechte steekproef zijn geselecteerd (LSK-A in Finke et al., 2001; Sample B in Visschers et al., 2007). De GxG's van deze locaties zijn geschat op basis van gerichte opnamen en stambuisregressie (Te Riele en Brus, 1991), gedurende de periode 1988-2001. Waarden van de GHG ontbreken voor 55 locaties, en van de GLG voor 66 locaties. Zeer waarschijnlijk hebben op deze locaties geen gerichte opnamen van de grondwaterstand plaatsgevonden, omdat deze te diep was. Dit blijkt uit de ligging van de ontbrekende waarden in stuwwallen, duinen etc., zie figuur 2.1 a en b. GHG's uit de LSK hebben een waarde van maximaal 357 cm, GLG's hebben een waarde van maximaal 455 cm. Bij de LSK zijn deelgebieden (strata) onderscheiden op basis van bodemtype en grondwatertrappen. Op deze manier is zowel een goede spreiding over verschillende bodemkundige en hydrologische situaties als een goede ruimtelijke spreiding nagestreefd.



Figuur 2.1

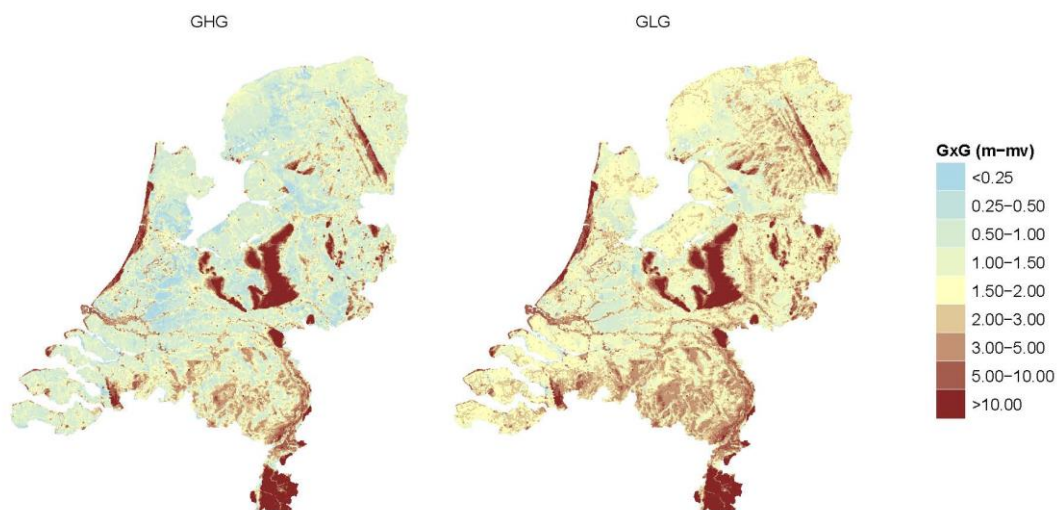
Ligging van de locaties van de Landelijke Steekproef Kaarteenheden, en ligging van de ontbrekende waarden (blauw) voor de GHG (links) en GLG (rechts)

Ritzema et al. (2012, pagina's 71-72) geven een overzicht van de fouten die op kunnen treden als met stambuisregressie en gerichte opnamen GxG's worden geschat. Op basis van verschillende onderzoeken concluderen zij dat bij stambuisregressie de standaardafwijking van de fout in GxG varieert van 10 tot 25 cm. GHG- en GLG-schattingen zijn ongeveer even nauwkeurig. Verder merken zij op dat als gevolg van *regression to the mean* ondiepe GxG's te diep worden geschat en diepe GxG's te ondiep. De ruimtelijke variatie in GxG's wordt hierdoor mogelijk onderschat.

2.2 GxG-voorspellingen van NHI, Hydrologie Stone en Gd-kaart

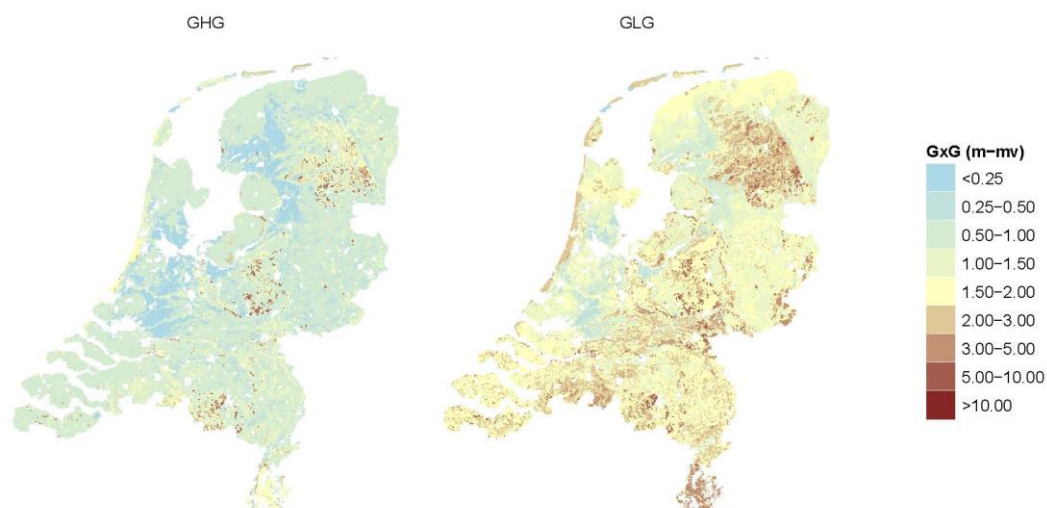
Met NHI 3.0 en Hydrologie Stone V2.3 zijn GxG's berekend uit ruimtelijke voorspellingen van het grondwaterstandsverloop voor de periode 1997-2005, zie figuur 2.2 en 2.3. Deze achtjarige periode valt samen met de periode waarin de inventarisaties voor de kaart van de grondwaterstandsdynamiek (Gd) plaatsvonden (Finke et al., 2005; Van Kekem et al., 2005). De NHI- en Hydrologie-Stone-voorspellingen van de GxG's zijn gemaakt voor een grid van 250x250 m.

De Gd-kaart heeft een resolutie van 25x25 m. De bestanden behorende bij de Gd-kaarten bevatten naast GxG-voorspellingen ook standaardafwijkingen van de voorspelfouten. Figuur 2.4 geeft de GxG-voorspellingen volgens de Gd-kaart.



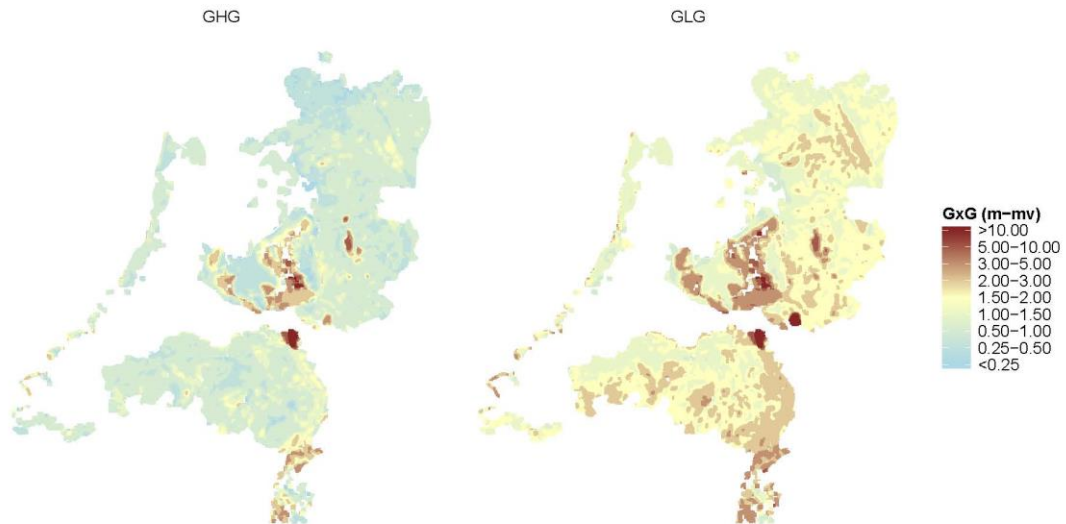
Figuur 2.2

GxG-voorspellingen met NHI 3.0 voor de periode 1997-2005



Figuur 2.3

GxG-voorspellingen met Hydrologie Stone V2.3 voor de periode 1997-2005



Figuur 2.4

GxG-voorspellingen volgens de Gd-kaart

De GxG's van NHI 3.0, Hydrologie Stone V2.3, Gd-kaart en LSK zijn gebaseerd op schattingen voor verschillende perioden. De GHG's en GLG's van NHI 3.0 en Hydrologie Stone V2.3 zijn berekend als het gemiddelde van de HG3's en LG3's (hoogste en laagste drie grondwaterstanden die jaarlijks optreden bij halfmaandelijke frequentie) over de periode 1997-2005. De GHG's en GLG's van de LSK zijn berekend uit reeksen van acht jaar lang, die eindigen in de periode dat de LSK is uitgevoerd, dus tussen 1988 en 2000. De GxG's van de Gd-kaart zijn geschat door reeksen van acht jaar lang die eindigen in de periode 1997-2005 te verlengen tot reeksen van 30 jaar lang over de periode 1970-2000, waaruit de GxG's zijn geschat als gemiddelde van 30 xG3's (gemiddelden van de drie hoogste of laagste standen in een hydrologisch jaar, bij een halfmaandelijke meetfrequentie). Om inzicht te krijgen in het effect van de verschillende perioden die zijn gehanteerd bij het schatten van GxG's, zijn voor zes verschillende locaties waar lange reeksen voor beschikbaar zijn, de GxG's geschat volgens de rekenperiode van NHI 3.0 en Hydrologie Stone V2.3, LSK en Gd-kaart. Tijdreeksmodellering is uitgevoerd met het programma Menyanthes (Von Asmuth et al., 2012). Tabel 2.1 geeft de resultaten, en figuur 2.5 geeft ter illustratie de tijdreeksgrafieken van de grondwaterstandsreeksen.

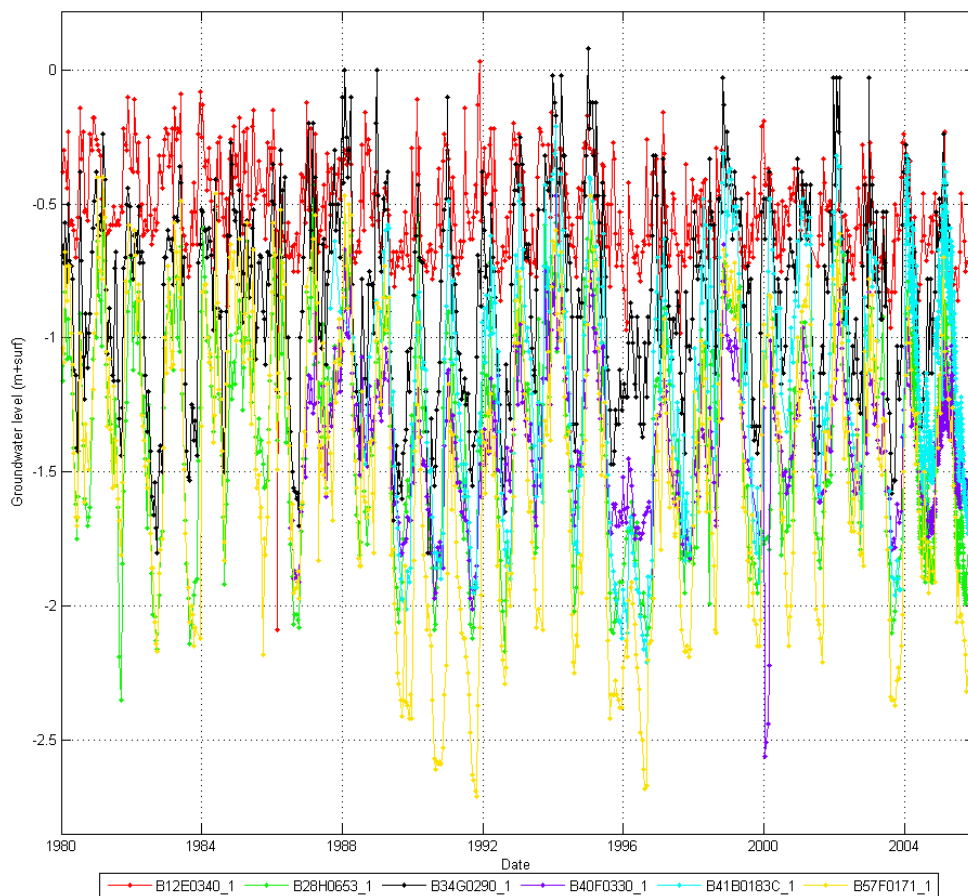
Tabel 2.1

Effect van verschillende rekenperioden op de GxG-schattingen. Waarden in cm.

Reeks	Rekenperiode					
	1997-2005 (NHI 3.0, Hydrologie Stone V2.3)		Kalibratie: acht jaar, eindigend tussen 1997 en 2005 Simulatie: 1970-2000 (Gd)			
	GHG	GLG	GHG	GLG	GHG	GLG
B12E0340_1	34	80	[24-39]	[80-81]	[24-31]	[65-79]
B28H0653_1	89	190	[93-103]	[198-208]	[78-92]	[177-199]
B34G0290_1	35	127	[38-42]	[133-142]	[30-46]	[129-151]
B40F0330_1	100	167	[112-122]	[176-184]	[100-108] ¹	[168-175] ¹
B41B0183_1	48	157	[74-85]	[188-193]	[57-68] ²	[178-188] ²
B57F0171_1	86	204	[102-107]	[224-230]	[64-107]	[189-239]

¹) Perioden van acht jaar, vanaf 1-4-1987; ²) Perioden van acht jaar, vanaf 1-4-1988

Uit tabel 2.1 blijkt dat het effect van de verschillende perioden waarop GxG-schattingen zijn gebaseerd kan variëren van enkele centimeters tot circa 35 cm. Over het algemeen zijn de GxG's die zijn geschat volgens de procedure die bij de Gd-karteringen is gevolgd dieper dan de GxG's die zijn geschat op basis van waarnemingen uit de periode 1997-2005. De grootste verschillen treden op bij B41B0183_1 en B57F0171_1. In beide gevallen lijkt een ingreep in de hydrologie hiervoor een verklaring. Uit figuur 2.5 blijkt namelijk dat tot ca. 1989 de standen van B28H0653 en B57F0171 vrij goed met elkaar overeenkomen, maar dat vanaf 1989 vooral in de zomerperiode in B57F0171 diepere standen worden gemeten. Mogelijk is bij B57F0171 daarom sprake van een effect van een ingreep, dat ook de variatie van GxG's verklaart voor de perioden van acht jaar die eindigen tussen 1988 en 2000 (laatste twee kolommen in tabel 2.1). Verder blijkt uit figuur 2.5 dat de standen in reeks B41B0183_1 vanaf circa 1999 relatief stijgen ten opzichte van de standen in de andere reeksen. Dit blijkt ook uit de relatief ondiepe GxG's die zijn berekend voor B41B0183_1 voor de periode 1997-2005 (tabel 2.1, kolom 2 en 3). Ook dit zou het effect van een ingreep in de hydrologie kunnen zijn.



Figuur 2.5

Tijdreeksgrafieken van de reeksen uit tabel 2.1 (uitgevoerd met Menyanthes, Von Asmuth et al., 2012)

2.3 Aanpak van validatie met LSK en vergelijking met Gd-kaart

De GxG-voorspellingen van NHI 3.0, Hydrologie Stone V2.3 en de Gd-kaart worden bij de validatie vergeleken met de GxG-schattingen uit de LSK. Omdat de LSK puntinformatie bevat geven de validatieresultaten informatie over de voorspelnaauwkeurigheid van NHI 3.0, Hydrologie Stone V2.3 en de Gd-kaart op puntschaal. Validatie

op de aggregatieniveaus van NHI 3.0, Hydrologie Stone V2.3 en de Gd-kaart is niet mogelijk omdat hiervoor herhaalde waarnemingen binnen de gridcellen nodig zijn waarmee de gridcelgemiddelden en hun standaardfout geschat kunnen worden (Stehman en Czaplowski, 1998; Brus et al., 2011).

De aanpak van de validatie op basis van de LSK is stapsgewijs als volgt:

1. Er worden gebieden (interessedomeinen) gedefinieerd waarvoor een validatie en vergelijking wordt uitgevoerd. De validatie en vergelijking richten zich op de volgende gebieden en deelgebieden:
 - a. het gebied waarbinnen tijdens de uitvoering van de LSK grondwaterstanden zijn gemeten (geen ontbrekende waarnemingen omdat het grondwater te diep stond). Dit is vrijwel geheel Nederland met uitzondering van gebieden met diepe grondwaterstanden (delen van stuwwallen en duinen, zie paragraaf 2.1). Op deze manier treedt geen vertekening op die zich voordoeft als ontbrekende waarnemingen worden genegeerd of vervangen door de maximale boordiepte. Vervangen van ontbrekende waarden door modelschattingen heeft geen voorkeur omdat dit afbreuk doet aan de objectiviteit van de validatie;
 - b. de grondwatertrappen (Gt's) volgens de stratificatie die bij de LSK is toegepast (zie tabel 2.2 voor de Gt-indeling);
 - c. de deelgebieden (strata) die bij de LSK zijn onderscheiden. Dit zijn combinaties van grondwatertrappen en bodemtypen;
 - d. de drie toetsingsregio's die bij toetsing van NHI 3.0 zijn onderscheiden (zie figuur 3.4; Hoogewoud et al., 2013b);
 - e. het gebied waarvoor een Gd-kaart is gemaakt. Dit is voornamelijk het pleistocene deel van Nederland. Voor dit gebied kunnen lokale GxG-voorspellingen van NHI 3.0 en Hydrologie Stone V2.3 worden vergeleken met voorspellingen van de Gd-kaart.Een klein deel van Nederland waar de bodemkaart associaties van Gt-klassen aangeeft is buiten beschouwing gebleven, omdat het aantal data hiervan beperkt was.
2. Er worden verschilwaarden berekend tussen de GxG's uit het LSK-bestand en de GxG's die zijn voorspeld met het NHI 3.0, Hydrologie Stone V2.3 en de Gd-methode.

De aanpak van de vergelijking tussen enerzijds GxG-voorspellingen met NHI 3.0 en Hydrologie Stone V2.3 en anderzijds de GxG-voorspellingen van de Gd-kaart is stapsgewijs als volgt:

1. De theoretisch berekende varianties van de voorspelfouten van de Gd-kaart worden gecorrigeerd met de resultaten van de validatie (zie kader).
2. NHI-voorspellingen en Hydrologie-Stone-voorspellingen worden vergeleken met de Gd-kaart, en er wordt aangegeven waar verschillen significant zijn. Grote verschillen in modelvoorspellingen kunnen worden geïnterpreteerd als grote onzekerheid over de werkelijke GxG.
3. Het effect van verschillen in aggregatieniveau tussen Gd enerzijds en NHI 3.0 en Hydrologie Stone V2.3 anderzijds wordt geanalyseerd door de Gd-voorspellingen ook te valideren op het aggregatieniveau van het NHI 3.0 en Hydrologie Stone V2.3. Hierbij wordt op twee verschillende manieren geaggregeerd:
 - a. GD-250, punt: validatie van de GxG-voorspellingen van de Gd-kaart voor 25x25m-cellen die grenzen aan de centra van het 250x250m-grid van het NHI 3.0 en Hydrologie Stone V2.3. Van de vier 25x25m-cellen die grenzen aan het middelpunt is telkens de cel linksonder genomen (standaard in ARC-GIS).
 - b. Validatie van het gemiddelde van de 100 25x25m-cellen van de Gd-kaart die binnen een 250x250m-cel van het NHI 3.0 en Hydrologie Stone V2.3 vallen.
4. Het effect van het gebruik van verschillende aggregatieniveaus in het AHN bij de omrekening van GxG t.o.v. NAP naar GxG t.o.v. maaiveld is geanalyseerd door GxG's t.o.v. NAP te valideren. Hierbij is voor LSK en GD het 25m-grid van het AHN gebruikt en voor NHI 3.0 het 250m-grid.

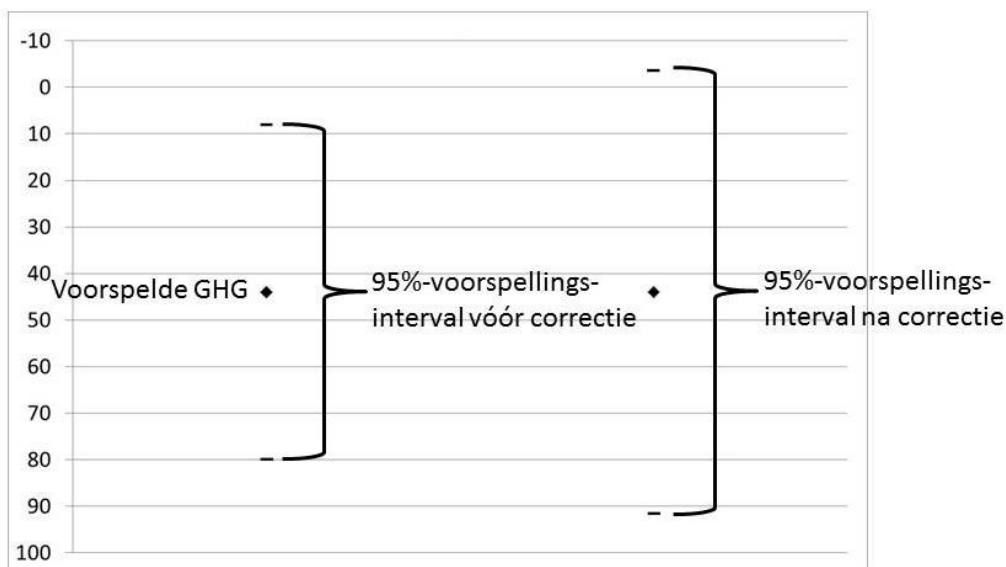
Correctie van standaardafwijkingen van voorspelfouten op de Gd-kaart

Kaarten van de grondwaterstandsdynamiek (Gd) zijn beschikbaar voor het pleistocene deel van Nederland. De kaarten geven voorspellingen van GxG's voor 25x25m-gridcellen, en geven een standaardafwijking van de voorspelfout (de wortel uit de zogeheten krigingvariantie). Deze krigingvariantie is gebaseerd op de veronderstelling dat het model van ruimtelijke variatie dat bij de interpolatie is gebruikt juist is, en geeft dus niet de nauwkeurigheid van dit model weer. Deze krigingvarianties kunnen echter worden vergeleken met de variantie van de voorspelfouten uit de validatie. Omdat de voorspelfouten die uit de validatie volgen ook de onzekerheid over het model van ruimtelijke variatie weergeven, kunnen deze worden gebruikt om de theoretisch berekende krigingvarianties te corrigeren. Vervolgens kunnen de NHI-voorspellingen en Hydrologie-Stonevoorspellingen worden vergeleken met de Gd-kaart, en kan op basis van de gecorrigeerde krigingvarianties worden aangegeven waar Gd-kaart enerzijds en NHI- en Hydrologie-Stonevoorspellingen anderzijds significant van elkaar verschillen. Grote verschillen kunnen worden geïnterpreteerd als grote onzekerheid over de werkelijke GxG. Kleine verschillen hoeven echter niet te betekenen dat de onzekerheid over de werkelijke GxG klein is; dit is alleen met validatie vast te stellen.

Als voorbeeld geven we een locatie waar de Gd-kaart een GHG-voorspelling geeft van 44 cm, met een standaardafwijking van de voorspelfout van 18 cm (krigingvariantie 324 cm²), gelegen in stratum 1501 van de LSK. Onder de veronderstelling dat de voorspelfouten normaal zijn verdeeld kan nu een 95%-voorspellingsinterval worden berekend van [44 cm +/- 2*18 cm], zie figuur 2.6, links.

De GHG-voorspellingen van de Gd-kaart hebben in stratum 1501 een krigingvariantie van gemiddeld 282 cm². De variantie van de verschillen tussen Gd-voorspellingen en LSK-schattingen van de GHG bedraagt in dit stratum 493 cm². De theoretisch berekende krigingvarianties worden in stratum 1501 nu gecorrigeerd met een factor 1.75 (=439/282). Voor de locatie in dit voorbeeld wordt de gecorrigeerde variantie van de voorspelfout 567 cm² (=1.75*324), en de gecorrigeerde standaardafwijking van de voorspelfout wordt 23,8 cm. Onder de veronderstelling dat de voorspelfouten normaal zijn verdeeld kan nu een 95%-voorspellingsinterval worden berekend van [44 cm +/- 2*23,8 cm], zie figuur 2.6, rechts.

Bij de bovenstaande correctie is geen rekening gehouden met de fout in de GxG-schattingen uit de LSK, waarvan de standaardafwijking door Ritzema et al. (2012) op 10 à 25 cm wordt geschat. Het is niet zonder meer duidelijk of deze fout tot over- of onderschatting van onzekerheid leidt. Nader onderzoek is nodig om het effect van deze fout te bepalen.



diepte (cm)

Figuur 2.6

95%-voorspellingsinterval rond een GHG-voorspelling van de Gd-kaart, vóór en na correctie op basis van de resultaten van validatie met de LSK

Tabel 2.2*Indeling in grondwatertrappen (Gt's). Waarden in cm.*

Gt	GHG	GLG
I	< 25	< 50
II	< 25	50-80
III	< 25	80-120
IIIb	25-40	80-120
IV	40-80	80-120
V	< 25	> 120
Vb	25-40	> 120
VI	40-80	> 120
VII	80-140	> 120
VIII	≥ 140	≥ 180

2.4 Validatiematen

De kwaliteit van modelvoorspellingen kan met verschillende validatiematen worden uitgedrukt. Deze paragraaf beschrijft welke validatiematen zijn gebruikt om de kwaliteit van GxG-voorspellingen weer te geven, en hoe deze validatiematen zijn gedefinieerd en berekend. Omdat de LSK is uitgevoerd volgens een steekproefopzet waarbij de selectiekansen van alle locaties in het studiegebied bekend zijn en groter dan 0, kunnen zuivere schattingen worden gemaakt van validatiematen, en kan de nauwkeurigheid van deze schattingen ook worden berekend. De volgende validatiematen zijn berekend:

- ME = gebiedsgemiddeld verschil tussen GxG volgens LSK en GxG volgens NHI 3.0, Hydrologie Stone V2.3 of Gd. Dit is een maat voor het systematische verschil of *bias*. Omdat de LSK is gebaseerd op een kanssteekproef kan ook de standaardfout van het geschatte gebiedsgemiddelde verschil worden berekend. Deze standaardfout kan worden gebruikt om te beoordelen of het gebiedsgemiddelde verschil significant van 0 afwijkt. Als het gebiedsgemiddelde verschil meer dan tweemaal de standaardfout is verwijderd van 0, dan kan het als significant worden beoordeeld (bij een significantieniveau van 5%).
- P50 = mediaan van de verschilwaarden. Als de verschilwaarden symmetrisch verdeeld zijn is deze gelijk aan de ME. Naarmate de verdeling schever is, verschillen ME en P50 meer van elkaar.
- RMSE = wortel uit het gemiddelde gekwadrateerde verschil, MSE, tussen GxG volgens LSK en GxG volgens NHI 3.0, Hydrologie Stone V2.3 of Gd. Dit is een maat voor het totale verschil of onnauwkeurigheid.
- SDE = standaardafwijking van de verschillen tussen GxG volgens LSK en GxG volgens NHI 3.0, Hydrologie Stone V2.3 of Gd. Dit is een maat voor het toevallige verschil of imprecisie.
- MAE = gemiddelde van de absolute verschillen tussen GxG volgens LSK en GxG volgens NHI 3.0, Hydrologie Stone V2.3 of Gd. Dit is evenals de RMSE een maat voor het totale verschil, maar minder gevoelig voor uitschieters.
- De cumulatieve frequentieverdeling van de verschilwaarden. Hieruit kunnen intervallen worden berekend waarbinnen een bepaald percentage van de fouten ligt (bijvoorbeeld 95%). Ook kunnen oppervlaktepercentages worden berekend waarbinnen de fout niet groter is dan een bepaalde waarde, bijvoorbeeld 50, 75 of 100 cm.

De procedure om uit de LSK-gegevens validatiematen te berekenen is als volgt:

1. Bereken verschilwaarden tussen LSK-waarden en NHI-, Hydrologie-Stone- of Gd-waarden van GxG voor locatie $i, i = 1 \dots n_h$ in stratum $h, h = 1 \dots H$:

$$\begin{aligned} z_{hi} &= \widehat{GxG}_{LSK,hi} - \widehat{GxG}_{NHI,hi}, \text{ of} \\ z_{hi} &= \widehat{GxG}_{LSK,hi} - \widehat{GxG}_{Stone,hi}, \text{ of} \\ z_{hi} &= \widehat{GxG}_{LSK,hi} - \widehat{GxG}_{Gd,hi}. \end{aligned} \quad (1)$$

2. Maak een schatting van de gemiddelde verschilwaarde voor het gedefinieerde interessedomein D . Gemiddelden worden geschat met de volgende formule (Finke et. al., 2001):

$$\hat{z}_D = \frac{\hat{z}'}{\hat{x}} = \frac{\sum_{h=1}^H a_h \cdot \sum_{i=1}^{n_h} z'_{hi}}{\sum_{h=1}^H a_h \cdot \sum_{i=1}^{n_h} x_{hi}} \quad (2)$$

waarin

- z'_{hi} de waarde van de doelvariabele z heeft op alle locaties binnen het domein en waarde 0 op alle locaties buiten het interessedomein,
- x_{hi} de waarde 1 heeft op alle locaties binnen het domein en 0 op alle locaties buiten het interessedomein,
- a_h de relatieve oppervlakten van de strata zijn, $h = 1 \dots H$.

De variantie van \hat{z}_D wordt geschat met:

$$v(\hat{z}_D) = \frac{1}{\hat{x}^2} \cdot \sum_{h=1}^H \frac{a_h^2}{n_h(n_h - 1)} \cdot \sum_{i=1}^{n_h} \left(\bar{d}_h - \frac{1}{n_h} \cdot \sum_{i=1}^{n_h} \bar{d}_h \right)^2 \quad (3)$$

waarin \bar{d}_h het gemiddelde is van $d = z'_{hi} - \hat{z}_D$ in stratum h , en

$$\hat{x} = \sum_{h=1}^H a_h \cdot \sum_{i=1}^{n_h} x_{hi} \quad (4)$$

de geschatte oppervlaktefractie van het interessedomein.

De standaardafwijking van de verschilwaarden, als maat voor het toevallige verschil, is berekend met

$$s_D(z) = \sqrt{\widehat{z_D^2} - (\hat{z}_D)^2 + v(\hat{z}_D)} \quad (5)$$

3. De validatiematen kunnen nu worden berekend. De ME wordt berekend met vergelijking (2). De MSE wordt analoog aan de ME berekend met vergelijking (2). De SDE wordt berekend met vergelijking (2), (3) en (5). De MAE wordt analoog aan de ME berekend met vergelijking (2). De cumulatieve frequentieverdeling van de verschilwaarden wordt als volgt berekend:
 - a. Stel uit de reeks verschilwaarden z_{hi} een gesorteerde reeks unieke waarden samen, zeg $z_j^*, j = 1 \dots n^*$.
 - b. Maak een 0/1-variabele $p_{hi,j}$, met waarde 1 als $z_{hi} \leq z_j^*$ en 0 als dat niet het geval is.
 - c. Bereken het gemiddelde van $p_{hi,j}$, zeg \hat{p}_j , analoog aan de berekening van ME met vergelijking (2).
 - d. Herhaal stap b. en c. voor alle $z_j^*, j = 1 \dots n^*$. De waarden van \hat{p}_j , uitgezet tegen z_j^* , geven de cumulatieve frequentieverdeling. De waarde van z_j^* die correspondeert met $\hat{p}_j = 0.5$ is de mediaan van de verschilwaarden.

Naast voorspelde GxG's valideerden we ook het voorspelde verschil tussen GLG en GHG, oftewel de voorspelde fluctuatieamplitude. We berekenen de volgende verschillen:

$$\begin{aligned} z_{hi} &= (GLG_{LSK,hi} - GHG_{LSK,hi}) - (GLG_{NHI,hi} - GHG_{NHI,hi}), \\ z_{hi} &= (GLG_{LSK,hi} - GHG_{LSK,hi}) - (GLG_{Stone,hi} - GHG_{Stone,hi}), \text{ en} \\ z_{hi} &= (GLG_{LSK,hi} - GHG_{LSK,hi}) - (GLG_{Gd,hi} - GHG_{Gd,hi}). \end{aligned}$$

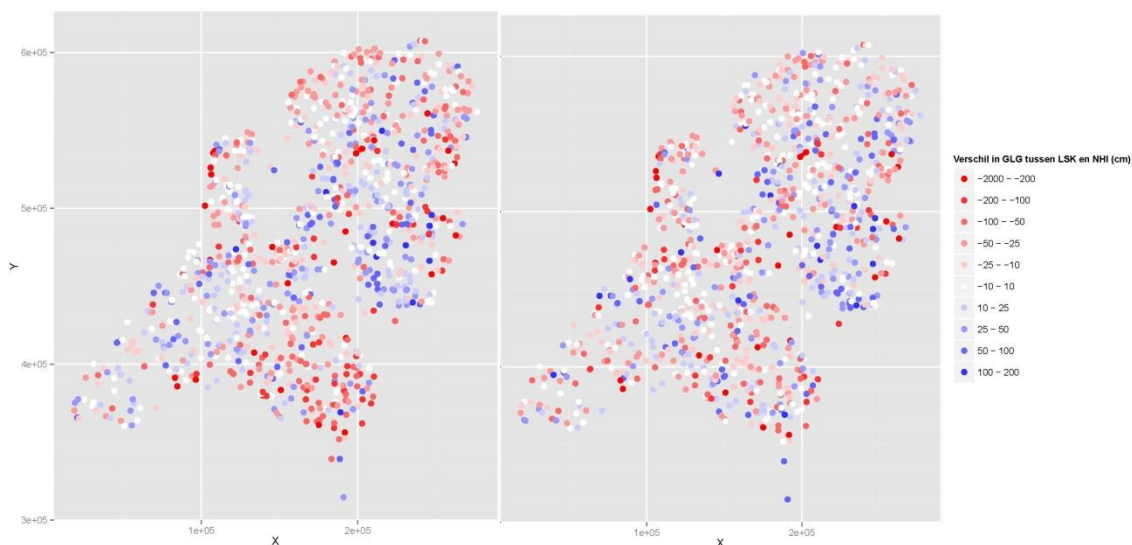
Uit deze verschilwaarden kunnen vervolgens validatiematen worden berekend volgens de stappen 1 t/m 3 hierboven.

3 Resultaten

3.1 Vergelijking van NHI-voorspellingen met LSK

3.1.1 Resultaten landelijk

Met de procedure die in paragraaf 2.4 is beschreven zijn validatiematen berekend die de kwaliteit van de GxG-voorspellingen van NHI 3.0 weergeven. Figuur 3.1 geeft kaartjes van de verschillen tussen GxG-voorspellingen met NHI 3.0 en de LSK-waarden. Rood geeft aan dat de NHI-voorspellingen dieper zijn dan de LSK-waarden, blauw geeft aan dat de NHI-voorspellingen ondieper zijn dan de LSK-waarden.



Figuur 3.1

Verskil tussen LSK-schattingen en NHI-voorspellingen van de GHG (links) en de GLG (rechts)

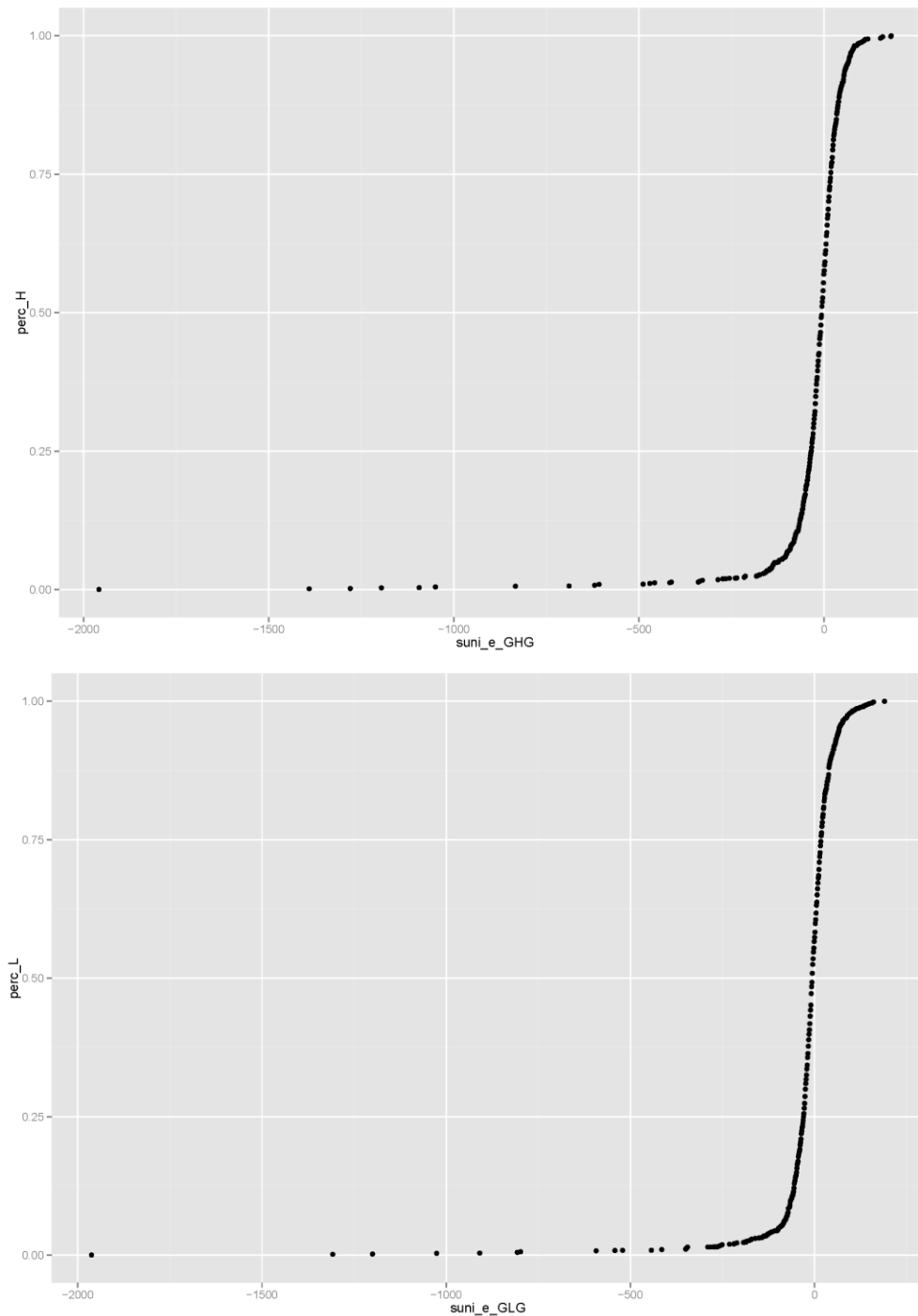
Tabel 3.1 geeft de berekende validatiematen. De negatieve ME's geven aan dat de NHI-voorspellingen systematisch dieper zijn dan de LSK-waarden. Deze ME's hebben kleine standaardfouten. Dit betekent dat ze significant van 0 verschillen bij een significantieniveau van 0.05: de ME is meer dan tweemaal de standaardfout verwijderd van 0. De mediaan van de verschillen is minder negatief dan de ME, waaruit blijkt dat de verschillen scheef verdeeld zijn. Dit blijkt ook uit de waarden voor het 2,5- en 97,5-percentiel, die de grenzen vormen van het interval waarbinnen 95% van de verschillen ligt. Een klein deel van de NHI-voorspellingen van zowel de GHG als de GLG is veel dieper dan de schattingen in de LSK: bij 2,5% van de GHG-voorspellingen is het verschil met LSK groter dan 179 cm en loopt op tot maximaal 1959 cm, en bij 2,5% van de GLG-voorspellingen is het verschil met LSK groter dan 185 cm en loopt op tot maximaal 1963 cm.

Tabel 3.1

Validatieresultaten in cm voor NHI-voorspellingen van GxG's, voor het gebied waarvoor LSK-schattingen van GxG's beschikbaar zijn. ME: gemiddelde fout. RMSE: root mean squared error. SDE: standaardafwijking van de fout. MAE: gemiddelde absolute fout. De Px-waarden geven de x-de percentielen van de cumulatieve frequentieverdeling aan. P50: mediaan. Standaardfouten tussen haakjes.

GHG					GLG				
ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5	ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5
-22(4)	119	117	49	-179; -35; -7; 18; 77	-17(3)	108	106	44	-185; -30; -7; 17; 90

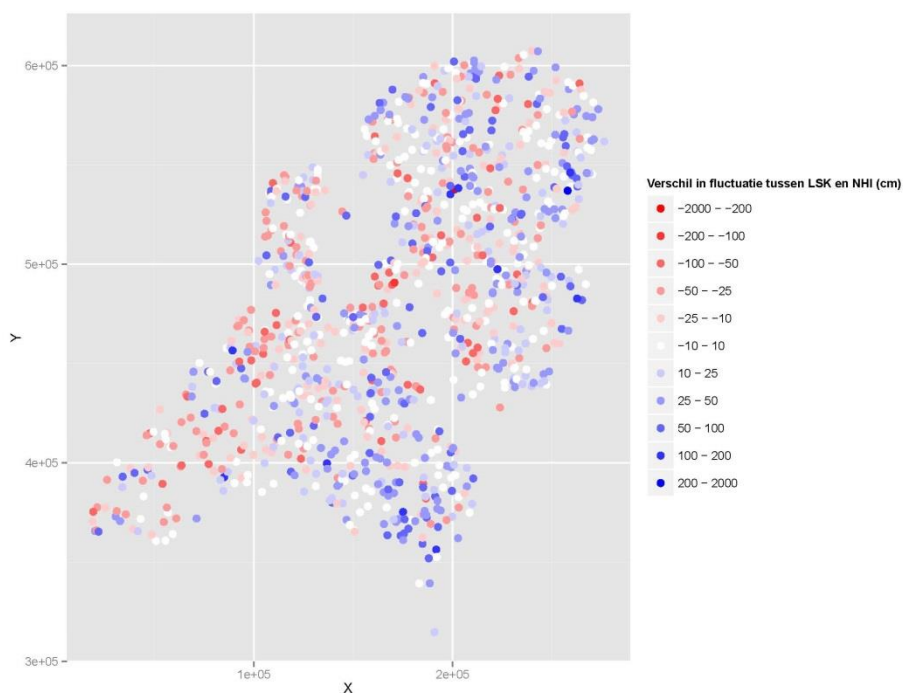
Ook de cumulatieve frequentieverdelingen van de verschillen in figuur 3.2 laten een scheve verdeling zien, in de richting van NHI-voorspellingen van GxG's die dieper zijn dan LSK-schattingen van GxG's. De figuren B3.1 t/m B3.4 in bijlage 3 laten de cumulatieve frequentieverdelingen zien voor de verschillen tussen LSK-schattingen en zowel NHI- als Stone-voorspellingen. In paragraaf 3.2.1 worden deze figuren verder toegelicht. Figuur B4.1 en B4.2 in bijlage 4 tonen de histogrammen van LSK-schattingen en NHI-voorspellingen van GHG en GLG.



Figuur 3.2

Cumulative frequentieverdeling van verschillen in NHI-voorspellingen en LSK-waarden voor GHG (boven) en GLG (onder). Waarden op x-as in cm.

Het verschil tussen NHI-voorspellingen van GLG en GHG, oftewel de voorspelde amplitude van de gemiddelde seizoensfluctuatie, is vergeleken met de amplitudes die uit de LSK-data volgen. Figuur 3.3 geeft een kaartje van de verschillen in amplitude volgens NHI 3.0 en LSK. In vergelijking met de LSK onderschat NHI 3.0 de amplitude van seizoensfluctuatie (GLG-GHG) voor het oosten van Nederland, en overschat deze amplitude voor het westen van Nederland. Tabel 3.2 geeft de berekende validatiematen.



Figuur 3.3

Verschil in fluctuatie (GLG-GHG) tussen NHI 3.0 en LSK. Positief: fluctuatie volgens NHI 3.0 kleiner dan volgens LSK. Negatief: fluctuatie volgens NHI 3.0 groter dan volgens LSK.

Tabel 3.2

Validatieresultaten in cm voor NHI-voorspellingen van de amplitude van seizoensfluctuatie (GLG-GHG), voor het gebied waarvoor LSK-schattingen van GxG's beschikbaar zijn. ME: gemiddelde fout. RMSE: root mean squared error. SDE: standaardafwijking van de fout. MAE: gemiddelde absolute fout. De Px-waarden geven de x-de percentielen van de cumulatieve frequentieverdeling aan. P50: mediaan. Standaardfouten tussen haakjes.

ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5
4(1)	39	39	30	-68; -22; 2; 27; 81

Uit tabel 3.2 blijkt dat NHI 3.0 de amplitude van de seizoensfluctuatie gemiddeld genomen licht onderschat. De vraag is of deze onderschatting van gemiddeld 4 cm relevant is voor de toepassingen van GxG-informatie uit NHI 3.0. Uit figuur 3.3 en de RMSE, SDE, MAE en de breedte van het interval waarin 95% van de verschilwaarden ligt (Tabel 3.2) blijkt dat er veel variatie in de mate van over- en onderschatting van de amplitude van seizoensfluctuatie is.

De GxG is per definitie ten opzichte van maaiveld. Om inzicht te krijgen in het effect van informatie over maaiveldhoogte op de voorspelnauwkeurigheid van de GxG, zijn ook GxG's ten opzichte van NAP berekend. De NHI-voorspellingen van GxG's zijn hierbij omgerekend naar NAP met behulp van het 250m-grid van het AHN, en de LSK-schattingen van GxG's zijn omgerekend naar NAP met behulp van het 25m-grid van het AHN. Tabel 3.3 geeft de validatiematen voor de GxG-voorspellingen ten opzichte van NAP.

Tabel 3.3

Validatieresultaten in cm voor NHI-voorspellingen van GxG's t.o.v. NAP, voor het gebied waarvoor LSK-schattingen van GxG's beschikbaar zijn. RMSE: root mean squared error. SDE: standaardafwijking van de fout. MAE: gemiddelde absolute fout. De Px-waarden geven de x-de percentielen van de cumulatieve frequentieverdeling aan. P50: mediaan. Standaardfouten tussen haakjes.

GHG					GLG				
ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5	ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5
24(4)	121	119	51	-76; -21; 7; 38; 192	18(4)	108	107	45	-91; -18; 7; 29; 186

Omdat de waarden in tabel 3.3 t.o.v. NAP zijn, is in vergelijking met tabel 3.1 het teken veranderd: positieve verschillen zijn nu negatief, *vice versa*. Verder blijkt dat voorspelling van GxG's t.o.v. NAP of maaiveld niet tot grote verschillen in (absolute) voorspelfouten leidt.

3.1.2 Resultaten per Gt-klasse

Tabel 3.4 geeft validatieresultaten voor de GHG, per Gt-klasse volgens de LSK. De grootste verschillen tussen NHI-voorspellingen en LSK-schattingen van de GHG komen voor in grondwatertrap Vb (zie tabel 2.1 voor de Gt-indeling). Grondwatertrap V en Vb hebben een grote amplitude van seizoensfluctuatie, die veroorzaakt kan zijn door bijvoorbeeld leemlagen. Hier kan mogelijk de oorzaak van de grote verschillen tussen NHI 3.0 en LSK worden gezocht. Opmerkelijk is echter dat de verschillen voor Gt V aanmerkelijk kleiner zijn dan voor Gt Vb. De kleinste verschillen tussen NHI-voorspellingen en LSK-schattingen van de GHG treden op bij Gt IIIb en IV. Deze Gt's hebben een geringe amplitude van seizoensfluctuatie, en komen vooral voor in gebieden met beheerste peilen.

Tabel 3.4

Validatieresultaten GHG-voorspellingen NHI 3.0, per Gt-klasse. ME: mean error (cm). RMSE: root mean squared error (cm). SDE: standard deviation of error (cm). MAE: mean absolute error (cm). n: aantal validatiepunten (-). Standaardfouten tussen haakjes.

Gt-klasse	ME(s.e.)	RMSE	SDE	MAE	n
I	-26(7)	78	74	41	115
II	-18(11)	150	149	38	200
IIIb	-11(4)	41	39	31	115
III	-8(7)	58	58	34	65
IV	-2(5)	55	55	32	130
Vb	-67(27)	231	223	90	67
V	-30(9)	108	105	47	139
VI	-16(4)	61	59	41	212
VII	-27(20)	162	161	67	93
VIII	-18(77)	126	146	22	65

Tabel 3.5 geeft validatieresultaten voor de GLG, per Gt-klasse. Ook bij de GLG komen de grootste verschillen tussen NHI 3.0 en LSK voor in Gt Vb, en de kleinste in Gt IIIb en IV.

Tabel 3.5

Validatieresultaten GLG-voorspellingen NHI 3.0, per Gt-klasse. ME: mean error (cm). RMSE: root mean squared error (cm). SDE: standard deviation of error (cm). MAE: mean absolute error (cm). n: aantal validatiepunten (-). Standaardfouten tussen haakjes.

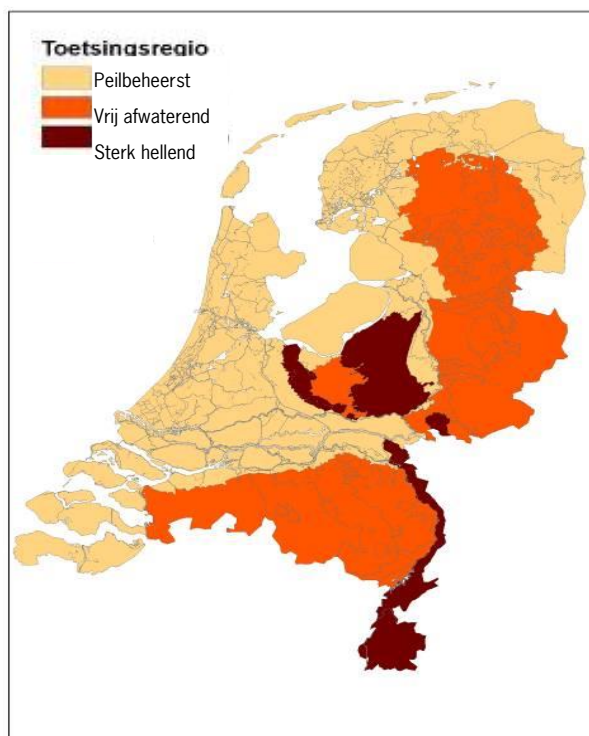
Gt-klasse	ME(s.e.)	RMSE	SDE	MAE	n
I	-32(7)	81	75	43	115
II	-25(11)	153	151	43	200
IIIb	-17(3)	41	37	30	115
III	-8(8)	61	61	38	65
IV	-3(5)	56	56	33	130
Vb	-54(24)	203	197	77	67
V	-13(8)	90	89	42	139
VI	-11(4)	59	58	40	212
VII	-10(20)	130	131	44	93
VIII	-11(76)	105	129	16	65

3.1.3 Resultaten per deelgebied van de LSK

Bijlage 1 geeft validatieresultaten per deelgebied (stratum) dat bij de LSK is onderscheiden. Tabel B1.1 geeft de resultaten voor de GHG, tabel B1.2 voor de GLG. In de tabellen zijn de deelgebieden aangegeven waarbinnen de kans op het voorkomen van stagnerende lagen binnen 120 cm groot is (mondelijke mededeling ing. H.R.J. Vroon, Alterra Wageningen UR). De diepte van 120 cm is gehanteerd, omdat de informatie van de bodemkaart die bij de stratificatie bij de LSK is gebruikt op deze diepte betrekking heeft. Uit de tabellen B1.1 en B1.2 blijken in zowel gebieden met grote kans op voorkomen van stagnerende lagen binnen 120 cm als erbuiten grote en kleine verschillen tussen LSK-schattingen en NHI-voorspellingen op te treden. Er kan geen verband worden aangetoond tussen de kans op voorkomen van stagnerende lagen binnen 120 cm en de grootte van de verschillen tussen LSK-schattingen en NHI-voorspellingen van de GxG.

3.1.4 Resultaten per toetsingsregio van NHI 3.0

Figuur 3.4 geeft de drie regio's waarvoor Hoogewoud et al. (2013b) kwaliteitsnormen geven waaraan NHI-voorspellingen van GxG's moeten voldoen. Tabel 3.6 geeft de resultaten voor toetsing van de NHI-voorspellingen aan de LSK-schattingen van GxG's op basis van de normen in Hoogewoud et al. (2013b).



Figuur 3.4

Ligging van de toetsingsregio's voor GxG-voorspellingen van NHI 3.0. Bron: Hoogewoud et al. (2013b)

Tabel 3.6

NHI 3.0 criteria voor GHG, GLG en dynamiek (ontleend aan tabel 2.2 uit Hoogewoud et al., 2013b), en resultaten van toetsing van NHI-voorspellingen aan GxG-schattingen van de LSK

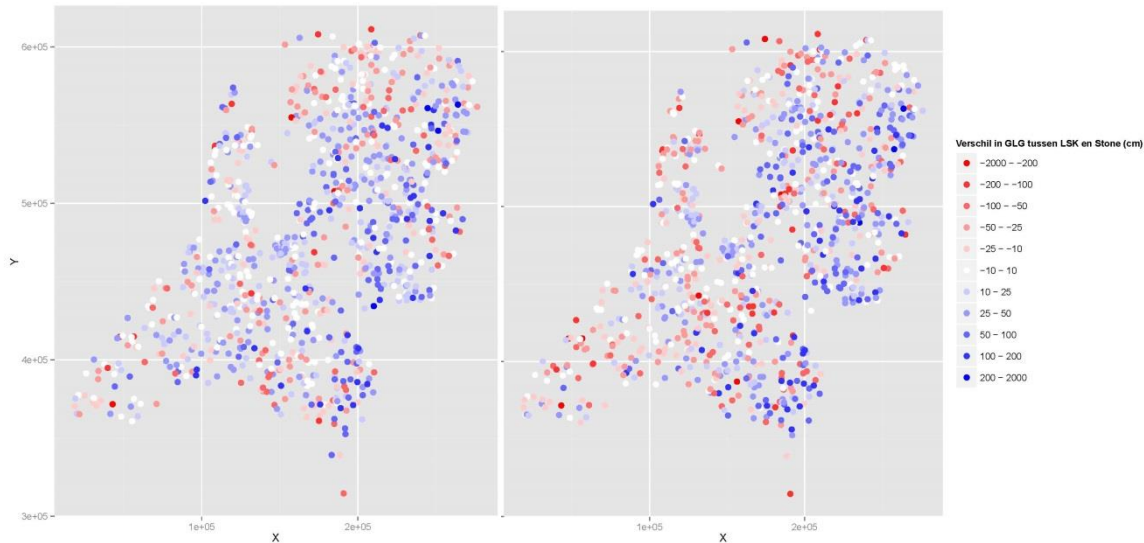
Indicator	Criterium	Toetsresultaten NHI 3.0-LSK
GHG (landelijk)	In 80% van de meetpunten in een gebied:	
	Peilbeheerst: max. 20 cm afwijking	43 %
	Vrij afwaterend: max. 35 cm afwijking	54 %
	Sterk hellend: max. 50 cm afwijking	49 %
	Afwijking van de mediaan < 15 cm	Mediaan = -7 cm
GLG (landelijk)	In 80% van de meetpunten in een gebied:	
	Peilbeheerst: max 30 cm afwijking	60 %
	Vrij afwaterend: max 50 cm afwijking	71 %
	Sterk hellend: max 100 cm afwijking	80 %
	Afwijking van de mediaan < 15 cm	Mediaan = -7 cm
Dynamiek (GHG-GLG)	In 80% van de meetpunten in een gebied:	
	Afwijking kleiner dan 25 % van de gemeten dynamiek	39 %
	Afwijking van de mediaan < 15 cm	Mediaan = 2 cm

Uit de resultaten in tabel 3.6 blijkt dat alleen bij de GLG in sterk hellende gebieden wordt voldaan aan het criterium van 80%. Aan het criterium dat is gesteld voor de mediaan van de afwijkingen wordt wel voldaan. Wel moet worden opgemerkt dat bij de toetsing in tabel 3.6 geen rekening is gehouden met de nauwkeurigheid van de LSK-schattingen van GxG's (zie paragraaf 2.1) en het effect van verschillen in rekenperiode tussen NHI 3.0 en LSK (zie tabel 2.2).

3.2 Vergelijking van Hydrologie-Stone-voorspellingen met LSK

3.2.1 Resultaten landelijk

De GxG-voorspellingen van het model Hydrologie Stone V2.3 zijn op dezelfde manier als de NHI-voorspellingen vergeleken met de LSK-schattingen. Volgens de procedure die in paragraaf 2.4 is beschreven zijn validatiematen berekend die de kwaliteit van de GxG-voorspellingen van Hydrologie Stone V2.3 weergeven. Figuur 3.5 geeft kaartjes van de verschillen tussen GxG-voorspellingen met Hydrologie Stone V2.3 en de LSK-waarden. Rood geeft aan dat de Hydrologie-Stone-voorspellingen dieper zijn dan de LSK-waarden, blauw geeft aan dat de Hydrologie-Stone-voorspellingen ondieper zijn dan de LSK-waarden.



Figuur 3.5

Verskil tussen LSK-schattingen en Hydrologie-Stone-voorspellingen van de GHG (links) en de GLG (rechts)

Tabel 3.7 geeft de berekende validatiematen. De positieve ME's geven aan dat de Hydrologie-Stone-voorspellingen systematisch ondieper zijn dan de LSK-waarden. Deze ME's hebben kleine standaardfouten, dit betekent dat ze significant van 0 verschillen bij een significantieniveau van 0.05. Uit de waarden voor het 2,5- en 97,5-percentiel, die de grenzen vormen van het interval waarbinnen 95 % van de verschillen ligt, blijkt dat de verschillen enigszins scheef zijn verdeeld in positieve richting.

Tabel 3.7

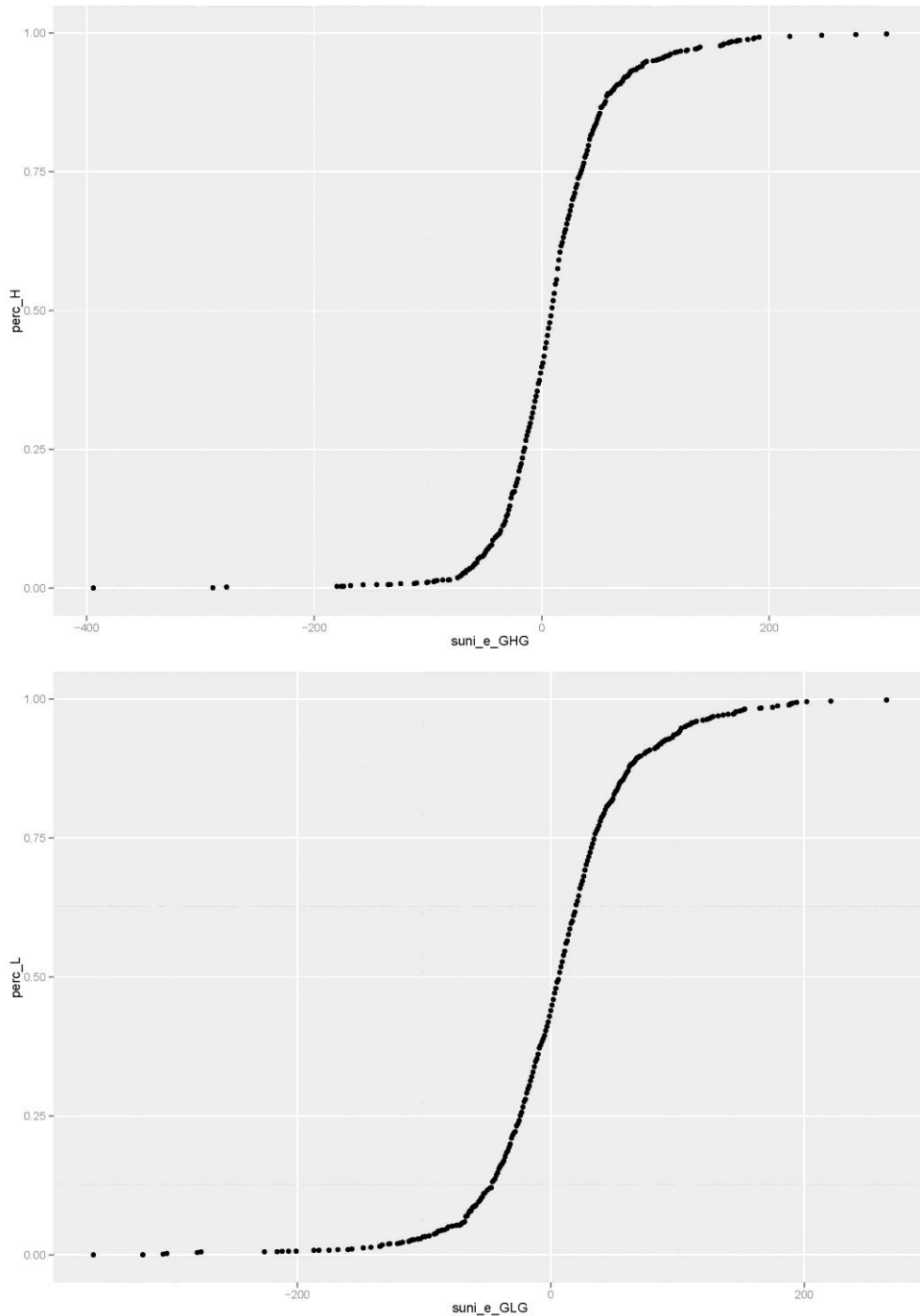
Validatieresultaten in cm voor Hydrologie-Stone-voorspellingen van GxG's, voor het gebied waarvoor LSK-schattingen van GxG's beschikbaar zijn. ME: gemiddelde fout. RMSE: root mean squared error. SDE: standaardafwijking van de fout. MAE: gemiddelde absolute fout. De Px-waarden geven de x-de percentielen van de cumulatieve frequentieverdeling aan. P50: mediaan. Standaardfouten tussen haakjes.

GHG					GLG				
ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5	ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5
13(2)	52	51	35	-70; -15; 9; 35; 143	7(2)	61	61	42	-111; -24; 7; 34; 146

De cumulatieve frequentieverdelingen van de verschillen in figuur 3.6 zijn enigszins scheef, in de richting van Hydrologie-Stone-voorspellingen van GxG's die ondieper zijn dan LSK-schattingen van GxG's. Hydrologie Stone V2.3 lijkt dus een iets 'natter' beeld te geven van de grondwaterstand dan volgt uit de LSK. Deze systematische verschillen zijn echter klein t.o.v. de verschillen als gevolg van verschillende rekenperioden (tabel

2.2). De figuren B3.1 tot en met B3.4 in bijlage 3 laten de cumulatieve frequentieverdelingen zien voor de verschillen tussen LSK-schattingen en zowel Stone- als NHI-voorspellingen. Uit deze figuren blijkt dat NHI 3.0 relatief vaak GxG's voorspelt die dieper zijn dan de LSK-schattingen, terwijl Hydrologie Stone V2.3 relatief vaak GxG's voorspelt die ondieper zijn dan de LSK-schattingen.

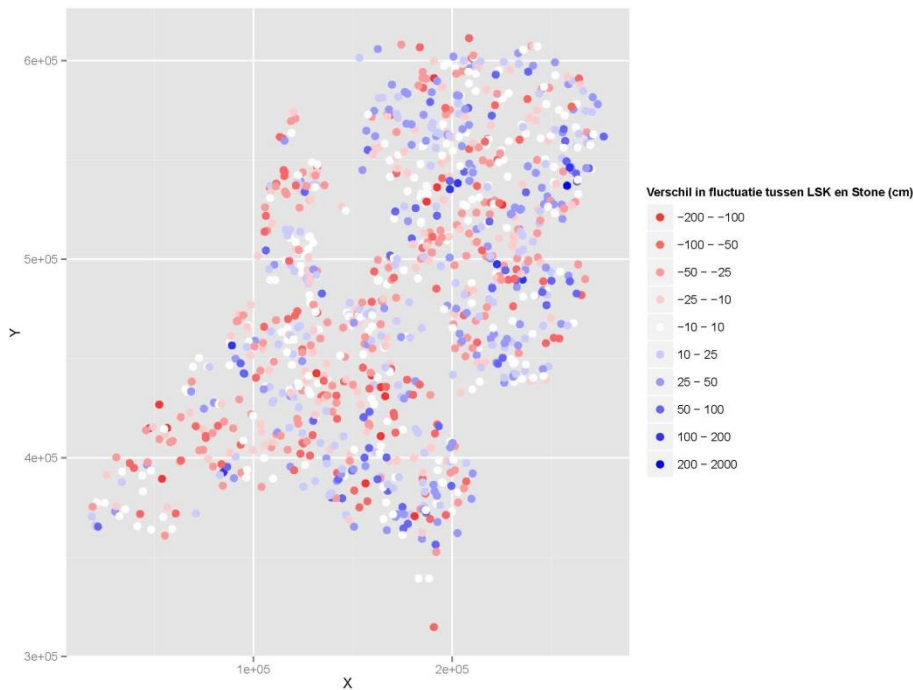
De figuren B4.3 en B4.4 in bijlage 4 tonen de histogrammen van LSK-schattingen en Hydrologie-Stone-voorspellingen van GHG en GLG.



Figuur 3.6

Cumulatieve frequentieverdeling van verschillen in Hydrologie-Stone-voorspellingen en LSK-waarden voor GHG (boven) en GLG (onder). Waarden op x-as in cm.

Het verschil tussen Hydrologie-Stone-voorspellingen van GLG en GHG, oftewel de voorspelde amplitude van de gemiddelde seizoensfluctuatie, is vergeleken met de amplitudes die uit de LSK-data volgen. Figuur 3.7 geeft een kaartje van de verschillen in amplitude volgens Hydrologie Stone V2.3 en LSK. Tabel 3.8 geeft de berekende validatiematen.



Figuur 3.7

Verschil in fluctuatie (GLG-GHG) tussen Hydrologie Stone V2.3 en LSK. Positief: fluctuatie volgens Hydrologie Stone V2.3 kleiner dan volgens LSK. Negatief: fluctuatie volgens Hydrologie Stone V2.3 groter dan volgens LSK.

Tabel 3.8

Validatieresultaten in cm voor Hydrologie-Stone-voorspellingen van de amplitude van seizoensfluctuatie (GLG-GHG), voor het gebied waarvoor LSK-schattingen van GxG's beschikbaar zijn. ME: gemiddelde fout. RMSE: root mean squared error. SDE: standaardafwijking van de fout. MAE: gemiddelde absolute fout. De Px-waarden geven de x-de percentielen van de cumulatieve frequentieverdeling aan. P50: mediaan. Standaardfouten tussen haakjes.

ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5
-4(1)	41	41	31	-93; -29; -2; 21; 74

Uit tabel 3.8 blijkt dat de Hydrologie Stone V2.3 de amplitude van seizoensfluctuatie van de grondwaterstand iets lijkt te overschatten: de ME is -4 cm, de mediaan slechts -2 cm. Het is echter de vraag of deze geringe gemiddelde overschatting relevant is. De over- of onderschatting varieert van locatie tot locatie, zoals uit figuur 3.7 en de overige informatie in tabel 3.8 blijkt.

Om inzicht te krijgen in het effect van informatie over maaiveldshoogte op de voorspelnauwkeurigheid van de GxG, zijn ook GxG's ten opzichte van NAP berekend. De Hydrologie-Stone-voorspellingen van GxG's zijn hierbij omgerekend naar NAP met het 250m-grid van het AHN, en de LSK-schattingen van GxG's zijn omgerekend naar NAP met het 25m-grid van het AHN. Tabel 3.9 geeft de validatiematen voor de GxG-voorspellingen met Hydrologie Stone V2.3, ten opzichte van NAP.

Tabel 3.9

Validatieresultaten in cm voor Hydrologie-Stone-voorspellingen van GxG's t.o.v. NAP, voor het gebied waarvoor LSK-schattingen van GxG's beschikbaar zijn. RMSE: root mean squared error. SDE: standaardafwijking van de fout. MAE: gemiddelde absolute fout. De Px-waarden geven de x-de percentielen van de cumulatieve frequentieverdeling aan. P50: mediaan. Standaardfouten tussen haakjes.

GHG					GLG				
ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5	ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5
-11(2)	65	64	40	-131; -36; -10; 18; 90	-5(2)	70	70	44	-130; -36; -5; 26; 124

Vergeleken met de validatieresultaten voor voorspellingen t.o.v. maaiveld in tabel 3.7, blijkt uit tabel 3.9 dat voorspelling van GxG's t.o.v. NAP niet tot kleinere voorspelfouten leidt. Alleen het teken verandert omdat het referentieniveau NAP is en niet maaiveld.

3.2.2 Resultaten per Gt-klasse

Tabel 3.10 geeft validatieresultaten voor de GHG, per Gt-klasse volgens de LSK. De grootste verschillen tussen Hydrologie-Stone-voorspellingen en LSK-schattingen van de GHG komen voor in grondwatertrap VII (zie tabel 2.1 voor de Gt-indeling). De kleinste verschillen tussen Hydrologie-Stone-voorspellingen en LSK-schattingen van de GHG treden op bij Gt III.

Tabel 3.10

Validatieresultaten GHG-voorspellingen met Hydrologie Stone V2.3, per Gt-klasse. ME: mean error (cm). RMSE: root mean squared error (cm). SDE: standard deviation of error (cm). MAE: mean absolute error (cm). n: aantal validatiepunten (-). Standaardfouten tussen haakjes.

Gt-klasse	ME(s.e.)	RMSE	SDE	MAE	n
I	1(5)	49	49	32	115
II	12(3)	40	38	24	200
IIIb	-4(5)	51	51	35	115
III	4(4)	29	29	22	65
IV	5(3)	37	37	28	130
Vb	14(6)	54	52	40	67
V	-7(3)	42	41	28	139
VI	8(3)	46	46	33	213
VII	46(7)	77	62	49	102
VIII	12(23)	42	46	13	65

Tabel 3.11 geeft de validatieresultaten per Gt-klasse voor de GLG. De grootste verschillen tussen Hydrologie Stone V2.3 en LSK blijken op te treden bij Gt V, Vb, IIIb en VII. Opmerkelijk is dat bij de 'natte' Gt's I en IIIb de Hydrologie-Stone-voorspellingen van de GLG systematisch dieper zijn dan de LSK-schattingen (negatieve ME's), terwijl bij de 'droge' Gt VII de Hydrologie-Stone-voorspellingen systematisch ondieper zijn (positieve ME).

Tabel 3.11

Validatieresultaten GLG-voorspellingen met Hydrologie Stone V2.3, per Gt-klasse. ME: mean error (cm). RMSE: root mean squared error (cm). SDE: standard deviation of error (cm). MAE: mean absolute error (cm). n: aantal validatiepunten (-). Standaardfouten tussen haakjes.

Gt-klasse	ME(s.e.)	RMSE	SDE	MAE	n
I	-13(6)	65	64	43	115
II	4(3)	44	44	29	200
IIIb	-19(6)	65	63	40	115
III	-3(5)	40	40	30	65
IV	-1(4)	45	45	32	130
Vb	6(8)	65	65	49	67
V	-3(6)	67	67	46	139
VI	6(4)	61	61	45	213
VII	30(8)	63	56	36	102
VIII	9(21)	34	39	9	65

3.2.3 Resultaten per deelgebied van de LSK

Bijlage 2 geeft validatieresultaten per deelgebied (stratum) dat bij de LSK is onderscheiden. Tabel B2.1 geeft de resultaten voor de GHG, tabel B2.2 voor de GLG. In de tabellen zijn de deelgebieden aangegeven waarbinnen de kans op het voorkomen van stagnerende lagen binnen 120 cm groot is (mondelijke mededeling ing. H.R.J. Vroon, Alterra Wageningen UR). De diepte van 120 cm is gehanteerd, omdat de informatie van de bodemkaart die bij de stratificatie bij de LSK is gebruikt op deze diepte betrekking heeft. Uit de tabellen B2.1 en B2.2 blijkt in zowel gebieden met grote kans op voorkomen van stagnerende lagen binnen 120 cm als erbuiten grote en kleine verschillen tussen LSK-schattingen en Hydrologie-Stone-voorspellingen op te treden. Er kan geen verband worden aangetoond tussen de kans op voorkomen van stagnerende lagen binnen 120 cm en de grootte van de verschillen tussen LSK-schattingen en Hydrologie-Stone-voorspellingen van de GxG.

3.2.4 Resultaten per toetsingsregio van NHI 3.0

De validatieresultaten voor GxG-voorspellingen van Hydrologie Stone V2.3 zijn getoetst aan de kwaliteitsnormen die Hoogewoud et al. (2013b) geven voor NHI 3.0. Tabel 3.12 geeft de resultaten, waaruit blijkt dat het criterium van 80% waarbinnen de afwijking niet groter is dan een bepaalde waarde in geen van de toetsingsregio's wordt gehaald. Aan het criterium voor de mediaan van de afwijkingen wordt wel voldaan.

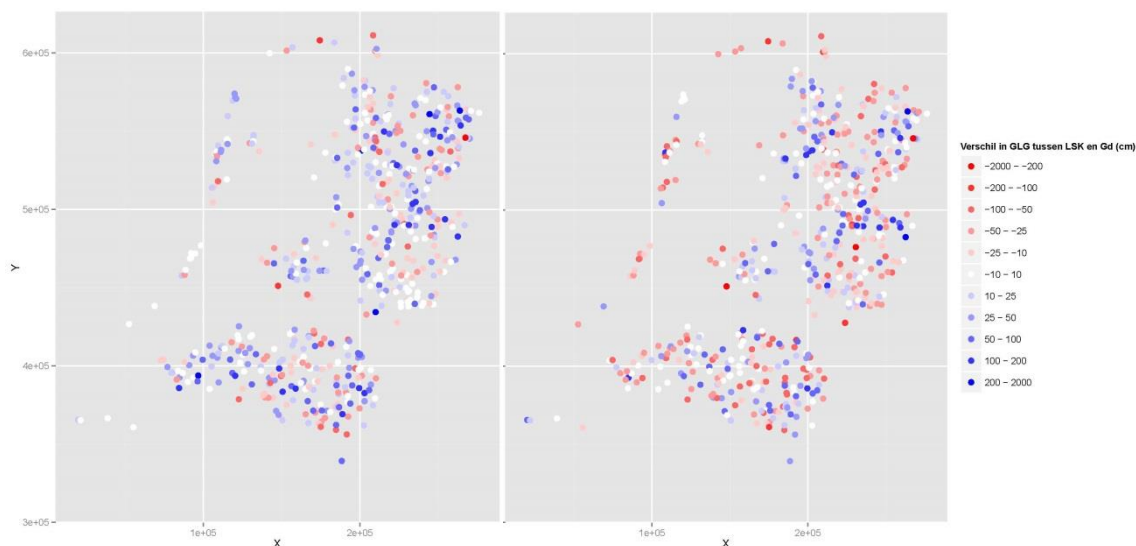
Tabel 3.12

NHI 3.0 criteria voor GHG, GLG en dynamiek (ontleend aan tabel 2.2 uit Hoogewoud et al., 2013b), en resultaten van toetsing van Hydrologie-Stone-voorspellingen aan GxG-schattingen van de LSK

Indicator	Criterium	Toetsresultaten Hydrologie-Stone - LSK
GHG (landelijk)	In 80% van de meetpunten in een gebied:	
	Peilbeheerst: max. 20 cm afwijking	36 %
	Vrij afwaterend: max. 35 cm afwijking	54 %
	Sterk hellend: max. 50 cm afwijking	39 %
	Afwijking van de mediaan < 15 cm	Mediaan = 9 cm
GLG (landelijk)	In 80% van de meetpunten in een gebied:	
	Peilbeheerst: max 30 cm afwijking	41 %
	Vrij afwaterend: max 50 cm afwijking	63 %
	Sterk hellend: max 100 cm afwijking	62 %
	Afwijking van de mediaan < 15 cm	Mediaan = 7 cm
Dynamiek (GHG-GLG)	In 80% van de meetpunten in een gebied:	
	Afwijking kleiner dan 25 % van de gemeten dynamiek	
	Afwijking van de mediaan < 15 cm	39 %
		Mediaan = -2 cm

3.3 Vergelijking van Gd-voorspellingen met LSK

Kaarten van de grondwaterstandsdynamiek (Gd) zijn beschikbaar voor het pleistocene deel van Nederland. De kaarten geven voorspellingen van GxG's voor 25x25m-gridcellen. Figuur 3.8 geeft een kaartje van de verschillen tussen GxG-voorspellingen volgens de Gd-kaart en GxG-schattingen volgens de LSK.



Figuur 3.8

Verskil tussen LSK-schattingen en Gd-voorspellingen van de GHG (links) en de GLG (rechts)

Tabel 3.13 geeft de validatieresultaten voor de GxG-voorspellingen van de Gd-kaart. Ter vergelijking zijn ook de validatieresultaten voor NHI- en Hydrologie-Stone-voorspellingen voor het gebied van de Gd-kartering opgenomen in tabel 3.12. NHI- en Hydrologie-Stone-voorspellingen hebben betrekking op 250x250m-cellen, terwijl Gd-voorspellingen betrekking hebben op 25x25m-cellen. Om het effect van dit verschil in

aggregatieniveau te analyseren zijn de Gd-voorspellingen geaggregeerd tot 250x250m-cellen. Deze aggregatie is op twee manieren uitgevoerd: door de waarde van een 25x25m-cel te nemen die grenst aan het midden van de 250x250m-cel (dit is in ARC-GIS standaard de cel links onder het midden), en door het gemiddelde te berekenen van de 100 25x25m-cellen die in een 250x250m-cel liggen. Uit de resultaten in tabel 3.12 blijkt dat beide manieren van aggregatie weinig effect hebben op de voorspelfouten: de waarden van de validatiematen veranderen weinig ten opzichte van de resultaten voor de niet-geaggregeerde Gd-voorspellingen. Dit duidt erop dat de Gd-voorspellingen weinig variatie vertonen binnen de 250x250m-cellen. Dit kan een effect zijn van het *smoothing*-effect van de interpolatiemethode kriging die bij de Gd-voorspellingen is gebruikt.

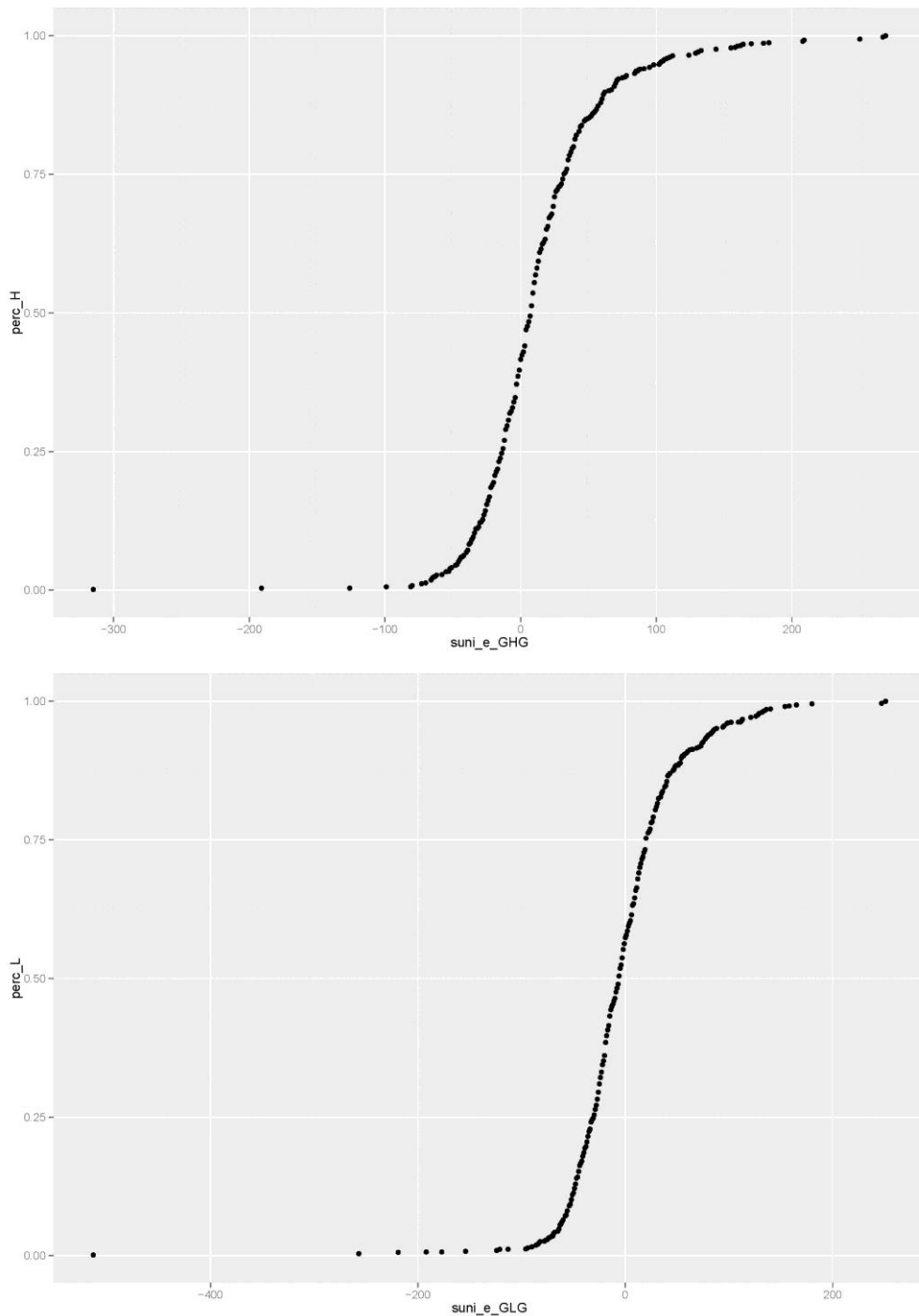
Tabel 3.13

Validatieresultaten in cm voor Gd-voorspellingen van GxG's, voor het gebied waarvoor 1) een Gd-kaart is, 2) LSK-schattingen van GxG's beschikbaar zijn. GD: validatieresultaten voor de GxG-voorspellingen van de Gd-kaart voor de 25x25m-cellen die met de LSK-punten samenvallen. ME: gemiddelde fout. RMSE: root mean squared error. MAE: gemiddelde absolute fout. De Px-waarden geven de x-de percentielen van de cumulatieve frequentieverdeling aan. P50: mediaan. Standaardfouten tussen haakjes.

Model	GHG					GLG				
	ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5	ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5
Gd	14(2)	52	50	34	-63; -14; 7; 32; 140	-2(2)	54	54	37	-85; -31; -6; 20; 128
Gd-250, punt*	16(2)	52	50	35	-62; -15; 8; 34; 145	-0(2)	53	53	38	-86; -32; -7; 22; 133
Gd-250, gemiddeld*	11(2)	56	55	37	-69; -20; 3; 30; 153	-5(2)	60	60	44	-111; -40; -11; 23; 152

**) GD-250, punt: validatieresultaten voor de GxG-voorspellingen van de Gd-kaart voor de 25x25m-cellen die met het 250x250m-grid van NHI 3.0 samenvallen. GD-250, gemiddeld: validatieresultaten voor het gemiddelde van de 100 25x25m-cellen van de Gd-kaart die binnen een 250x250m-cel van NHI 3.0 vallen.*

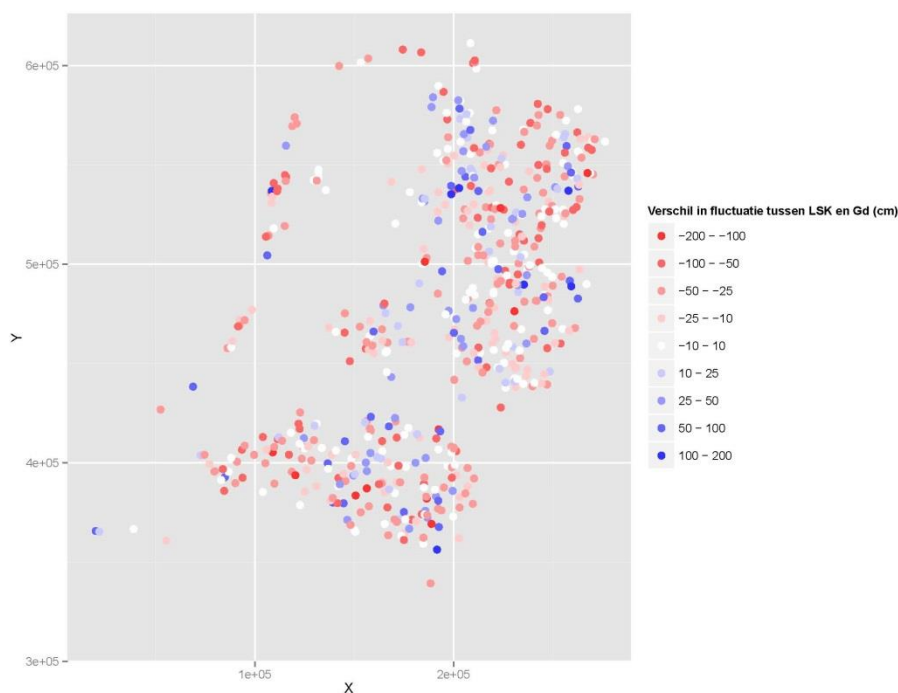
Figuur 3.9 geeft de cumulatieve frequentieverdelingen van de verschillen tussen GxG-voorspellingen volgens de Gd-kaart en GxG-schattingen volgens de LSK. De grafiek voor de GHG laat zien dat de voorspellingen volgens de Gd-kaart iets vaker ondieper dan dieper zijn dan de schattingen uit de LSK. Bij de GLG is dit andersom. De figuren B3.5 tot en met B3.8 in bijlage 3 laten de cumulatieve frequentieverdelingen zien voor de verschillen tussen LSK-schattingen en zowel Gd-, Stone- als NHI-voorspellingen. Uit deze figuur blijkt dat NHI 3.0 relatief vaak GxG's voorspelt die dieper zijn dan de LSK-schattingen, terwijl de Gd-kaart en Hydrologie Stone V2.3 relatief vaak GxG's voorspellen die ondieper zijn dan de LSK-schattingen. Bij de Gd-kaart geldt dit niet voor de GLG.



Figuur 3.9

Cumulative frequentieverdelingen van de verschillen tussen GxG-voorspellingen volgens de Gd-kaart en GxG-schattingen volgens de LSK. Boven: GHG. Onder: GLG. Waarden x-as in cm.

Het verschil tussen Gd-voorspellingen van GLG en GHG, oftewel de voorspelde amplitude van de gemiddelde seizoensfluctuatie, is vergeleken met de amplitudes die uit de LSK-data volgen. Figuur 3.10 geeft een kaartje van de verschillen in amplitude volgens de Gd-kaart en de LSK. Tabel 3.14 geeft de berekende validatiematen. De resultaten geven aan dat de Gd-kaart relatief vaak een grotere amplitude van fluctuatie weergeeft dan de LSK. Dit blijkt uit de negatieve waarde van de ME en de mediaan (P50) in tabel 3.14.



Figuur 3.10

Verschil in fluctuatie (GLG-GHG) tussen Gd en LSK. Positief: fluctuatie volgens Gd kleiner dan volgens LSK. Negatief: fluctuatie volgens Gd groter dan volgens LSK.

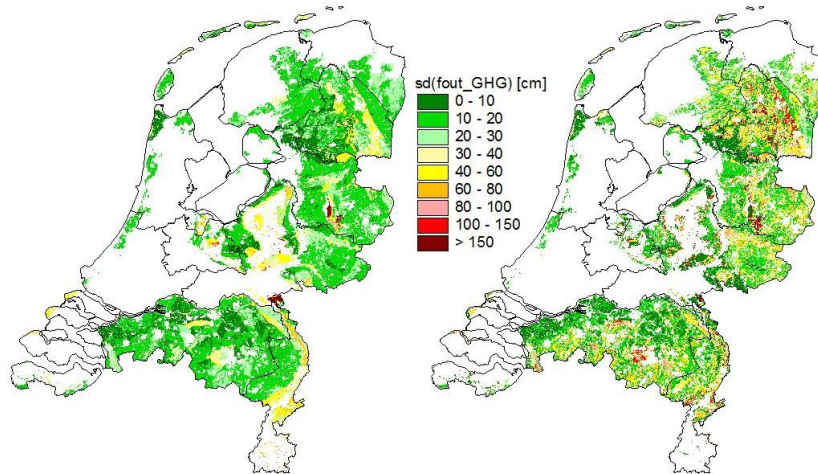
Tabel 3.14

Validatieresultaten in cm voor Gd-voorspellingen van de amplitude van seizoensfluctuatie (GLG-GHG), voor het gebied waarvoor zowel Gd-voorspellingen als LSK-schattingen van GxG's beschikbaar zijn. ME: gemiddelde fout. RMSE: root mean squared error. SDE: standaardafwijking van de fout. MAE: gemiddelde absolute fout. De Px-waarden geven de x-de percentielen van de cumulatieve frequentieverdeling aan. P50: mediaan. Standaardfouten tussen haakjes.

ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5
-1.3(2)	42	41	33	-94; -38; -15; 6; 74

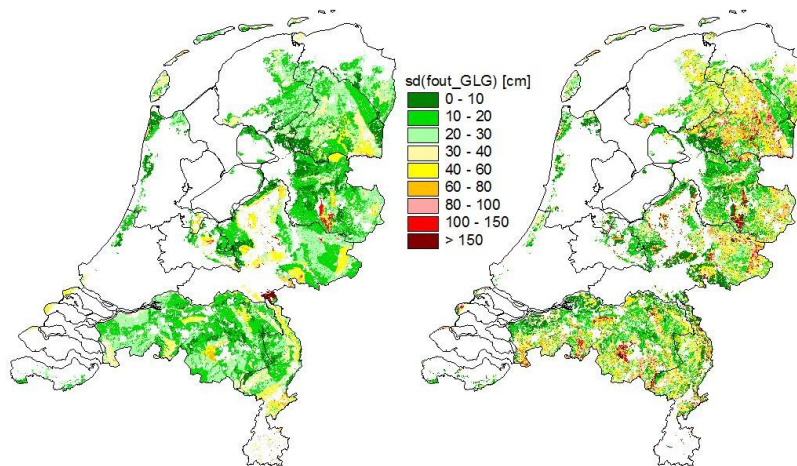
3.4 Vergelijking van NHI-voorspellingen met Gd-voorspellingen

Naast voorspellingen van GxG's voor 25x25m-gridcellen, geven de Gd-kaarten voor elke voorspelling een standaardafwijking van de voorspelfout (de wortel uit de zogeheten krigingvariantie). Zoals in paragraaf 2.3 is beschreven kunnen de krigingvarianties en -standaardafwijkingen worden aangepast aan de variantie en standaardafwijking van de voorspelfouten uit de validatie, zodat de onzekerheid over het model van ruimtelijke variatie wordt verdisconteerd. De figuren 3.11 en 3.12 geven de resultaten van deze aanpassing voor respectievelijk de standaardafwijking van de voorspelfout van GHG en GLG.



Figuur 3.11

Standaardafwijkingen van de voorspelfouten van GHG's volgens de Gd-kaart. Links: theoretisch berekende standaardafwijkingen. Rechts: standaardafwijkingen na correctie op basis van verschillen tussen Gd-kaart en LSK-schattingen van GHG's.

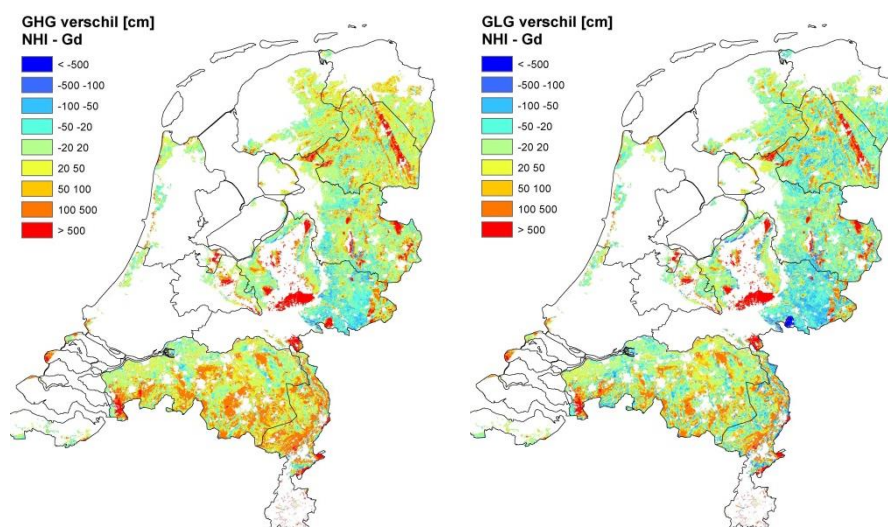


Figuur 3.12

Standaardafwijkingen van de voorspelfouten van GLG's volgens de Gd-kaart. Links: theoretisch berekende standaardafwijkingen. Rechts: standaardafwijkingen na correctie op basis van verschillen tussen Gd-kaart en LSK-schattingen van GLG's.

Vervolgens worden de NHI-voorspellingen vergeleken met de Gd-kaart, en kan op basis van de gecorrigeerde krigingvarianties worden aangegeven waar Gd-kaart enerzijds en NHI-voorspellingen anderzijds significant van elkaar verschillen. In deze gebieden lopen Gd-kaart en NHI 3.0 duidelijk uiteen en dit duidt op grote onzekerheid over de werkelijke GxG.

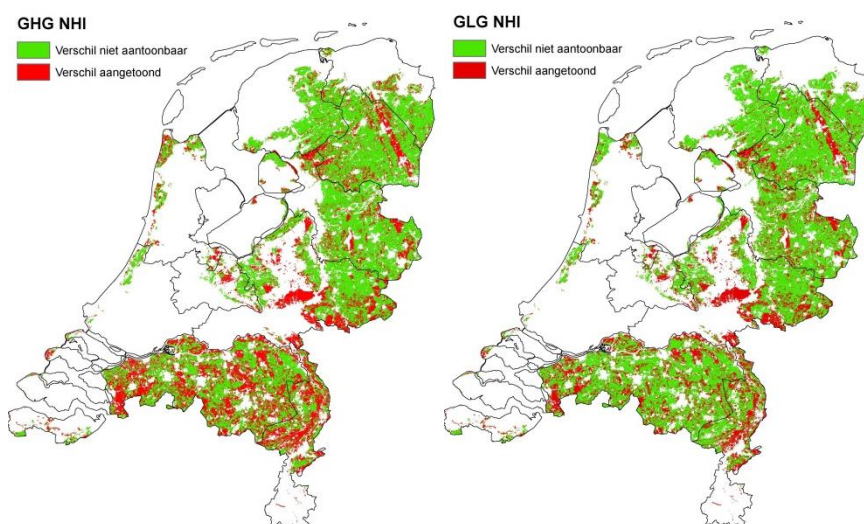
Figuur 3.13 geeft kaartjes van de verschillen tussen NHI- en Gd-voorspellingen van de GxG. De kaartjes laten zien dat onder stuwwallen, de Hondsrug, de Brabantse Wal (bij Bergen op Zoom) en de Maasterrassen NHI 3.0 diepere GxG's voorspelt dan de Gd-kaart. Opmerkelijk is echter dat de GLG-voorspellingen voor het Montferland hierop een uitzondering vormen: voor dit gebied voorspelt NHI 3.0 ondiepere GLG's dan de Gd-kaart. Omdat we hier twee modellen met elkaar vergelijken, kan niet op voorhand worden gezegd of hier sprake is van een fout in de voorspellingen van NHI 3.0 dan wel de Gd-kaart. In figuur 3.13 komt ook duidelijk de grens van het Oost-Nederlands plateau in de Achterhoek naar voren. Ten westen van deze grens zijn de NHI-voorspellingen vaak ondieper dan de Gd-voorspellingen, terwijl ten oosten van deze grens de NHI-voorspellingen vaak dieper zijn.



Figuur 3.13

Verschillen tussen NHI- en Gd-voorspellingen van de GHG (links) en de GLG (rechts). Positief verschil: NHI 3.0 dieper dan Gd. Negatief verschil: NHI 3.0 ondieper dan Gd.

Figuur 3.14 geeft aan waar de NHI-voorspellingen buiten het 95%-voorspellingsinterval van de Gd-kaart liggen en verschillen als significant zijn aan te merken. In 35% van de Gd-kaart vallen de NHI-voorspellingen van de GHG buiten het 95%-voorspellingsinterval dat de Gd-kaart aangeeft. Voor de GLG is dit percentage 29%.

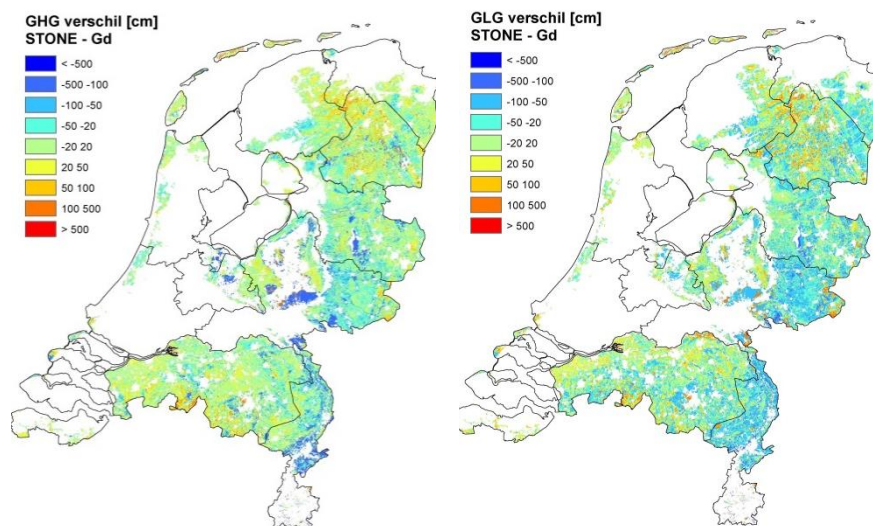


Figuur 3.14

Gebieden waar de NHI-voorspellingen significant van de Gd-voorspellingen verschillen, bij een 5%-significantieniveau. Links: GHG, rechts: GLG.

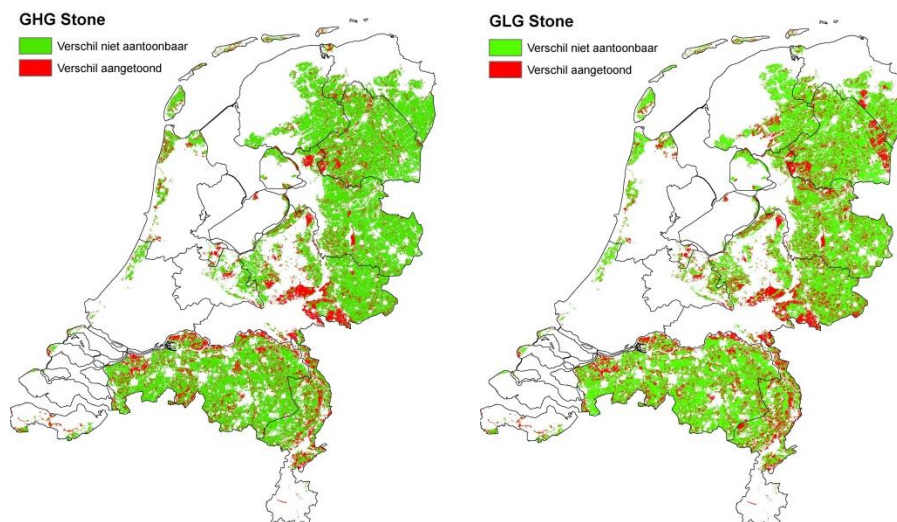
3.5 Vergelijking van Hydrologie-Stone-voorspellingen met Gd-voorspellingen

Figuur 3.15 geeft kaartjes van de verschillen tussen Hydrologie-Stone- en Gd-voorspellingen van de GxG. De kaartjes laten zien dat onder stuwwallen en in Oost-Nederland Hydrologie Stone V2.3 ondiepere GxG's voorspelt dan de Gd-kaart.



Figuur 3.15

Verschillen tussen Hydrologie-Stone- en Gd-voorspellingen van de GHG (links) en de GLG (rechts). Positief verschil: Hydrologie Stone V2.3 dieper dan Gd. Negatief verschil: Hydrologie Stone V2.3 ondieper dan Gd.



Figuur 3.16

Gebieden waar de Hydrologie-Stone-voorspellingen significant van de Gd-voorspellingen verschillen, bij een 5%-significantieniveau. Links: GHG, rechts: GLG.

Figuur 3.16 geeft aan waar de Hydrologie-Stone-voorspellingen buiten het 95%-voorspellingsinterval van de Gd-kaart liggen en verschillen als significant zijn aan te merken. In 20% van de Gd-kaart vallen de Hydrologie-Stone-voorspellingen van de GHG buiten het 95%-voorspellingsinterval dat de Gd-kaart aangeeft. Voor de GLG is dit percentage 25%.

4 Conclusies en aanbevelingen

4.1 Conclusies

Deze paragraaf beantwoordt de onderzoeksvragen in paragraaf 1.3 met een aantal conclusies.

Vraag 1: *Hoe groot zijn de verschillen tussen GxG-voorspellingen volgens NHI 3.0, Hydrologie Stone V2.3 en Gd-kaart enerzijds en de GxG-schattingen uit de LSK anderzijds? Deze vraag wordt beantwoord voor heel Nederland, per Gt-klasse en per bodemkundig-hydrologisch stratum dat is onderscheiden bij de LSK.*

Deze vraag is beantwoord met de resultaten die zijn gepresenteerd in hoofdstuk 3 en de bijlagen. Uit deze resultaten kunnen de volgende conclusies worden getrokken:

- 1.1. De GxG-voorspellingen van NHI 3.0 zijn systematisch dieper dan de GxG-schattingen van de LSK, en dit systematische verschil is significant bij een significantieniveau van 5% voor zowel GHG als GLG. Voor de GHG is het gemiddelde verschil tussen LSK-schatting en NHI-voorspelling -22 cm, en voor de GLG is dit -17 cm. Voor zowel GHG als GLG is de mediaan van de verschillen -7 cm.
- 1.2. Een klein deel van de NHI-voorspellingen van zowel de GHG als de GLG is veel dieper dan de schattingen in de LSK: bij 2.5% van de GHG-voorspellingen is het verschil met LSK groter dan 179 cm en loopt op tot maximaal 1959 cm, en bij 2.5% van de GLG-voorspellingen is het verschil met LSK groter dan 185 cm en loopt op tot maximaal 1963 cm.
- 1.3. De GxG-voorspellingen van Hydrologie Stone V2.3 zijn systematisch ondieper dan de GxG-schattingen van de LSK, en dit systematische verschil is significant bij een significantieniveau van 5% voor zowel GHG als GLG. Voor de GHG is het gemiddelde verschil tussen LSK-schatting en Hydrologie-Stone-voorspelling 13 cm en is de mediaan van de verschillen 9 cm. Voor de GLG zijn zowel het gemiddelde verschil als de mediaan van de verschillen 7 cm.
- 1.4. De GHG-voorspellingen van de Gd-kaart zijn systematisch ondieper dan de GHG-schattingen van de LSK, en dit systematische verschil is significant bij een significantieniveau van 5%. Het gemiddelde van de verschillen is 14 cm, de mediaan is 7 cm. Voor de GLG kan geen systematisch verschil worden aangetoond.

Vraag 2: *Waar verschillen GxG-voorspellingen volgens NHI 3.0 en Hydrologie Stone V2.3 van de GxG-voorspellingen volgens de Gd-kaart?*

Deze vraag is beantwoord in de figuren 3.13 en 3.14 voor NHI 3.0, en in de figuren 3.15 en 3.16 voor

Hydrologie Stone V2.3. Uit deze resultaten kan het volgende worden geconcludeerd:

- 2.1. Vooral onder stuwwallen, de Hondsrug, de Brabantse Wal (bij Bergen op Zoom) en de Maasterrassen voorspelt NHI 3.0 significant diepere GxG's voorspelt dan de Gd-kaart, bij een significantieniveau van 5%. De GLG-voorspellingen voor het Montferland vormen hierop een uitzondering: hier voorspelt NHI 3.0 significant ondiepere GLG's dan de Gd-kaart.
- 2.2. De GxG-voorspellingen van Hydrologie Stone V2.3 zijn vooral onder stuwwallen, de Hondsrug en de Maasterrassen significant ondieper dan de GxG-voorspellingen volgens de Gd-kaart.

Vraag 3: *Hoe groot zijn de verschillen in amplitude van seizoensfluctuatie (GLG minus GHG) tussen enerzijds NHI-, Hydrologie-Stone- en Gd-voorspellingen en anderzijds LSK-schattingen?*

Deze vraag is beantwoord voor NHI 3.0 in figuur 3.3 en tabel 3.2, voor Hydrologie Stone V2.3 in figuur 3.7 en tabel 3.8 en voor de Gd-kaart in figuur 3.10 en tabel 3.14. Uit deze resultaten kan het volgende worden geconcludeerd:

- 3.1. In vergelijking met de LSK onderschat NHI 3.0 de amplitude van seizoensfluctuatie (GLG-GHG) voor het oosten van Nederland, en overschat deze amplitude voor het westen van Nederland. Gemiddeld over Nederland wordt de amplitude iets onderschat (ME=4 cm). De verschillen variëren met een standaardafwijking van 39 cm.
 - 3.2. In vergelijking met de LSK overschat Hydrologie Stone V2.3 de amplitude van seizoensfluctuatie iets (ME=-4 cm). De verschillen variëren met een standaardafwijking van 41 cm.
 - 3.3. In vergelijking met de LSK overschat de Gd-kaart de amplitude van seizoensfluctuatie (ME=-13 cm). De verschillen variëren met een standaardafwijking van 41 cm.
- Uit de resultaten blijkt dat de verschillen in seizoensfluctuatie tussen enerzijds NHI 3.0, Hydrologie Stone V2.3 en Gd-kaart en anderzijds de LSK in dezelfde grootteorde liggen. Hier is geen directe verklaring voor te geven.

Vraag 4: *In hoeverre beantwoorden de resultaten van de validatie van NHI 3.0 en Hydrologie Stone V2.3 op basis van de LSK aan de toetsingscriteria voor GHG, GLG en dynamiek die Hoogewoud et al. (2013b) hanteren voor toetsing van NHI 3.0?*

Aan het criterium dat in tenminste 80% van een toetsingsregio de afwijking in GHG, GLG of amplitude van seizoensfluctuatie (dynamiek) niet groter mag zijn dan een bepaalde waarde, voldoen de verschillen tussen voorspellingen van zowel NHI 3.0 als Hydrologie Stone V2.3 enerzijds en LSK-schattingen anderzijds in vrijwel geen van de gevallen. Alleen de GLG-voorspellingen van NHI 3.0 in sterk hellende gebieden voldoen aan het criterium van 80%. Aan het criterium dat is gesteld voor de mediaan van de afwijkingen wordt altijd voldaan, zowel bij NHI 3.0 als bij Hydrologie Stone V2.3. Opgemerkt moet echter worden dat bij deze toetsing geen rekening is gehouden met de nauwkeurigheid van de LSK-schattingen van GxG's (zie paragraaf 2.1) en het effect van verschillen in rekenperiode tussen NHI 3.0 en Hydrologie Stone V2.3 enerzijds en LSK anderzijds (zie tabel 2.2).

Vraag 5: *Wat zeggen de validatieresultaten over de nauwkeurigheid waarmee NHI 3.0, Hydrologie Stone V2.3 en de Gd-kaart GxG's voorspellen, gezien de fouten in de GxG-schattingen uit de LSK?*

Vraag 1 en vraag 3 hadden betrekking op de verschillen tussen NHI 3.0, Hydrologie Stone V2.3 en de Gd-kaart enerzijds en de LSK anderzijds. De GxG's uit de LSK zijn echter schattingen die fouten kunnen bevatten. Wat zeggen deze verschillen dan over de nauwkeurigheid waarmee NHI 3.0, Hydrologie Stone V2.3 en Gd-kaart GxG's voorspellen?

Zoals beschreven in paragraaf 2.1 zijn de GxG's uit de LSK gebaseerd op stambuisregressie, en concluderen Ritzema et al. (2012, pp. 71-72) dat hierbij de standaardafwijking van de fout in GxG varieert van 10 tot 25 cm. Op de systematische fout in GxG-voorspelling over heel Nederland, uitgedrukt in ME of P50, hebben de fouten als gevolg van stambuisregressie waarschijnlijk weinig invloed. Wat het effect is op de toevallige fout in GxG-voorspellingen, uitgedrukt in SDE in de tabellen in hoofdstuk 3, is niet op voorhand te zeggen en moet nader worden onderzocht.

Ritzema et al. (2012) merken verder op dat als gevolg van *regression to the mean* ondiepe GxG's te diep worden geschat en diepe GxG's te ondiep. Dit heeft invloed op de resultaten per Gt-klasse (tabellen 3.4, 3.5, 3.10 en 3.11) en per LSK-stratum (bijlage 1 en 2). Uit tabel 3.4 blijkt bijvoorbeeld dat in Gt-klasse I NHI 3.0 de GHG gemiddeld 26 cm dieper voorspelt dan de LSK. Ten opzichte van de werkelijke GHG zijn de LSK-schattingen al te diep vanwege *regression to the mean*, dus hier kan worden aangenomen dat de systematische fout in werkelijkheid groter is dan 26 cm. Uit tabel 3.4 blijkt dat in Gt-klasse VIII NHI 3.0 de GHG gemiddeld 18 cm dieper voorspelt dan de LSK. Dit kan geheel of gedeeltelijk worden verklaard uit *regression to the mean* waardoor de LSK-schattingen te ondiep zijn.

Vraag 6: *Wat zeggen de validatieresultaten over de nauwkeurigheid waarmee NHI 3.0, Hydrologie Stone V2.3 en de Gd-kaart GxG's voorspellen, gezien de verschillen in rekenperiodes die zijn gehanteerd bij deze GxG-voorspellingen en bij de GxG-schattingen volgens de LSK?*

Het effect van verschillende rekenperiodes op schattingen en voorspellingen van de GxG is geanalyseerd in paragraaf 2.2. De GxG's voor de periode 1997-2005, waarmee gerekend is bij NHI 3.0 en Hydrologie Stone

V2.3, zijn meestal iets ondieper dan de GxG's die zijn berekend volgens de methode die bij de Gd-kaart is gevolgd of uit de reeksen van acht jaar die tijdens de opname van de LSK zijn gebruikt. Het volgende kan worden geconcludeerd:

- 6.1. De relatief diepe GxG-voorspellingen van NHI 3.0 t.o.v. LSK (conclusie 1.1) en Gd (conclusie 2.1) kunnen niet aan verschillen in rekenperiode worden toegeschreven.
- 6.2. De relatief ondiepe GxG-voorspellingen van Hydrologie Stone V2.3 t.o.v. LSK (conclusie 1.3) en Gd (conclusie 2.2) kunnen niet aan verschillen in rekenperiode worden toegeschreven.
- 6.3. De gemiddeld relatief ondiepe voorspelling van GHG's met de Gd-kaart (ME= 14 cm, tabel 3.13) kan niet worden verklaard uit verschillen in de rekenperiode tussen Gd-kaart en LSK.

Vraag 7: *Wat kan op grond van de gevonden verschillen, en gegeven de nauwkeurigheid van de LSK-schattingen en het effect van verschillende rekenperiodes, worden geconcludeerd over de nauwkeurigheid van GxG-voorspellingen volgens NHI 3.0, Hydrologie Stone V2.3 en de Gd-kaart?*

Samenvattend kunnen uit deze studie de volgende conclusies worden getrokken over de nauwkeurigheid van GxG-voorspellingen volgens NHI 3.0, Hydrologie Stone V2.3 en de Gd-kaart:

- 7.1. GxG-voorspellingen van NHI 3.0 zijn gemiddeld genomen relatief diep. Diepe voorspellingen komen vooral voor onder stuwwallen, de Hondsrug, de Brabantse Wal en de Maasterrassen.
- 7.2. GxG-voorspellingen van Hydrologie Stone V2.3 zijn gemiddeld genomen relatief ondiep. Ondiepe voorspellingen komen vooral voor onder stuwwallen, de Hondsrug en de Maasterrassen.
- 7.3. GHG-voorspellingen van de Gd-kaart zijn gemiddeld genomen relatief ondiep.
- 7.4. NHI 3.0 onderschat de amplitude van seizoensfluctuatie van de freatische grondwaterstand in Oost-Nederland, en overschat deze amplitude in West-Nederland. Hydrologie Stone V2.3 en de Gd-kaart overschatten de amplitude van seizoensfluctuatie gemiddeld genomen iets.

Ter verduidelijking van de onderlinge verschillen in voorspelnauwkeurigheid worden in tabel 4.1 de validatieresultaten samengevat voor het gebied waarvoor van zowel NHI 3.0, Hydrologie Stone V2.3 als de Gd-kaart voorspellingen zijn gemaakt.

Tabel 4.1

Validatieresultaten in cm voor Gd, NHI en Hydrologie-Stone-voorspellingen van GxG's, voor het gebied waarvoor 1) een Gd-kaart is, 2) LSK-schattingen van GxG's beschikbaar zijn. GD: validatieresultaten voor de GxG-voorspellingen van de Gd-kaart voor de 25x25m-cellen die met de LSK-punten samenvallen. ME: gemiddelde fout. RMSE: root mean squared error. MAE: gemiddelde absolute fout. De Px-waarden geven de x-de percentielen van de cumulatieve frequentieverdeling aan. P50: mediaan. Standaardfouten tussen haakjes.

Model	GHG					GLG				
	ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5	ME	RMSE	SDE	MAE	P2,5; P25; P50; P75; P97,5
Gd	14(2)	52	50	34	-63; -14; 7; 32; 140	-2(2)	54	54	37	-85; -31; -6; 20; 128
NHI 3.0	-32(6)	138	134	59	-326; -46; -11; 18; 91	-17(5)	120	118	51	-257; -30; -3; 25; 96
Hydrologie Stone V2.3	19(3)	60	57	40	-70; -14; 13; 41; 185	19(3)	67	64	48	-102; -18; 15; 52; 189

4.2 Aanbevelingen

Deze studie leidt tot de volgende aanbevelingen voor validatie van gebiedsdekkende informatie over de freatische grondwaterstand:

1. Het verdient aanbeveling om voorafgaand aan validatie onderbouwde validatiecriteria te definiëren die informatie geven over de bruikbaarheid van informatie over de freatische grondwaterstand bij de doelen waarvoor deze informatie wordt ingezet.
2. GxG's kunnen in principe alleen maar foutloos worden berekend als er een reeks is van foutloos gemeten grondwaterstanden, zonder missende waarnemingen, met een lengte van 30 jaar waarin het hydrologische regime niet door ingrepen is gewijzigd. Omdat reeksen die hieraan voldoen zelden voorhanden zijn worden GxG's geschat uit kortere reeksen, uit gerichte opnamen en stambuisregressie, en uit reeksen die met een tijdreeksmodel zijn verlengd tot 30 jaar. Als gevolg hiervan heeft een GxG altijd een bepaalde onnauwkeurigheid, die het validatieresultaat kan beïnvloeden. Het verdient daarom aanbeveling om niet te valideren op geschatte GxG's, maar op gemeten standen, of op parameters die direct uit gemeten standen zijn berekend, zoals xG3's (gemiddelden van de drie hoogste of laagste standen in een hydrologisch jaar, bij een halfmaandelijke meetfrequentie).
3. Om objectiviteit na te streven verdient het aanbeveling deze standen te meten op locaties die volgens een kanssteekproef zijn geselecteerd. Op deze manier kunnen zuivere schattingen van grootheden zoals gebiedsgemiddelde fouten worden gemaakt en kan de nauwkeurigheid van deze grootheden valide worden gekwantificeerd.
4. Om onafhankelijkheid na te streven verdient het aanbeveling dat de waarnemingen uit dit meetnet uitsluitend voor validatie van modellen van de grondwaterstand worden gebruikt en dus niet voor kalibratie beschikbaar worden gesteld.
5. Het verdient aanbeveling om de validatieset te gebruiken voor het beoordelen van de bruikbaarheid van alle modellen voor het genereren van informatie over de freatische grondwaterstand, inclusief regionale modellen en nieuwe modelversies.
6. Om de validatieset te kunnen inzetten voor beoordeling van regionale modellen verdient het aanbeveling om voor de betreffende gebieden het validatiemeetnet te verdichten.

Literatuur

Asmuth, J.R. von, K. Maas, M. Knotters, M.F.P. Bierkens, M. Bakker, Th.N. Olsthoorn, D.G. Cirkel, I. Leunk, F. Schaars en D.C. von Asmuth, 2012. Software for hydrogeologic time series analysis, interfacing data with physical insight. *Environmental Modelling and Software* 38: 178-190.

Bakel, P.J.T. van, H.Th.L. Massop, J.G. Kroes, J. Hoogewoud, M.J.H. Pastoors en T. Kroon, 2008. *Actualisatie hydrologie voor STONE 2.3; Aanpassing randvoorwaarden en parameters, koppeling tussen NAGROM en SWAP, en plausibiliteitstoets*. Wageningen, Wettelijke Onderzoekstaken Natuur & Milieu, WOt-rapport 57.

Brus, D.J., B. Kempen en G.B.M. Heuvelink, 2011. Sampling for validation of digital soil maps. *European Journal of Soil Science* 62: 394-407.

Finke, P.A., J.J. de Gruijter en R. Visschers, 2001. *Status 2001 Landelijke Steekproef Kaarteenheden en toepassingen*. Rapport 389, Alterra Wageningen UR.

Finke, P.A., D.J. Brus, M.F.P. Bierkens, T. Hoogland, M. Knotters en F. de Vries, 2005. Kartering van de grondwaterdynamiek met behulp van geo-informatie van hoge resolutie. *Stromingen* 11(1): 27-41.

Hoogewoud, J., J. Hunink, G. Prinsen, A. Veldhuizen en J. Verkaik, 2013a. *Veranderingsrapportage NHI 3.0. Beschrijving van de veranderingen in versie 3.0*. Rapport nr. 1206107-000, Deltares.

Hoogewoud, J.C., G.F. Prinsen, J.C. Hunink, A.A. Veldhuizen, F.J.E. van der Bolt en W.J. de Lange, 2013b. *Toetsingsrapportage NHI 3.0*. Rapport nr. 1206107-000, Deltares.

Kekem, A.J. van, T. Hoogland en J.B.F. van der Horst, 2005. *Uitspoelingsgevoelige gronden op de kaart; werkwijze en resultaten*. Rapport 1080, Alterra Wageningen UR.

Knotters, M. en P.E.V. van Walsum, 1997. Estimating fluctuation quantities from time series of water-table depths using models with a stochastic component, *Journal of Hydrology* 197: 25-46.

Riele, W.J.M. te en D.J. Brus, 1991. *Methoden van gerichte grondwaterstandmetingen voor het schatten van de GHG*. Rapport 158, DLO-Staring Centrum, Wageningen.

Snepvangers, J., A. Veldhuizen, G. Prinsen en J. Delsman, 2008. *Nationaal Hydrologisch Instrumentarium – NHI. Modelrapportage*. NHI, Hoofdrapport, 28 blz.

Stehman, S.V. en R.L. Czaplewski, 1998. Design and analysis for thematic map accuracy assessment: fundamental principles. *Remote sensing of the environment* 64: 331-344.

Visschers, R., P.A. Finke en J.J. de Gruijter, 2007. A soil sampling program for the Netherlands. *Geoderma* 139: 60-72.

Bijlage 1 Validatieresultaten NHI 3.0 per LSK-stratum

Tabel B1.1

Validatieresultaten GHG-voorspellingen NHI 3.0, per stratum. ME: mean error (cm). RMSE: root mean squared error (cm). SDE: standard deviation of error (cm). MAE: mean absolute error (cm). n: aantal validatiepunten (-). Standaardfouten tussen haakjes. Bodems waarin de kans op stagnerende lagen binnen 120 cm groot is, zijn donker aangegeven (ing. H.R.J. Vroon, Alterra Wageningen UR, mondelinge mededeling).

Stratum	Gt	Grondsoort	ME(s.e.)	RMSE	SDE	MAE	n
1501	I	Veen	-10(7)	55	54	28	56
1502	I	Veen	-40(24)	105	100	58	18
1503	I	Zand	-114(40)	160	120	114	9
1504	I	Zee- en rivierklei	-20(7)	31	25	22	12
1505	I	Alle	-18(15)	61	60	47	20
1601	II	Veen	2(4)	15	15	11	18
1602	II	Veen	-14(6)	40	38	29	34
1603	II	Veen	-6(6)	32	32	24	32
1604	II	Veen	-2(7)	33	34	26	24
1605	II	Zand, keileem	-34(53)	124	131	78	6
1606	II	Oude rivierklei, zand, leem, moerig	-14(14)	72	72	37	26
1607	II	Klei, zand, moerig	4(6)	34	35	22	33
1608	II	Zand	-2(5)	13	14	9	7
1609	II	Alle	-127(100)	452	446	143	20
1701	IIIb	Veen, moerig	-42(21)	98	91	59	18
1702	III	Veen, moerig	-14(8)	42	41	34	29
1703	III	Zand, keileem	-3(10)	45	46	27	23
1704	IIIb	Klei, leem	5(5)	32	32	25	47
1705	III	Klei, zand	-20(5)	37	31	29	36
1706	III	(Oude) rivierklei	-3(8)	40	41	33	27
1801	IV	Zand	12(8)	27	26	21	10
1802	IV	Zand, klei	13(9)	43	42	35	20
1803	IV	Veen, moerig	-5(7)	25	25	20	15
1804	IV	(Zee-)klei	6(5)	32	32	25	38
1805	IV	(Oude) rivierklei	-10(13)	38	39	26	9
1806	IV	Zeeklei, zand	10(8)	32	32	25	18
1807	IV	Alle	-46(25)	116	109	67	20
1901	Vb	Zand	-28(13)	60	55	45	19
1902	Vb	Zand	-2(25)	55	60	46	6
1903	Vb	Zand	-9(8)	18	18	13	5
1904	V	Zand	-46(41)	85	82	70	4
1905	Vb	Keileem, oude klei	-40(11)	71	59	51	27
1906	Vb	Klei, zand, beekdal	-44(47)	139	140	63	9
1907	Vb	Zand	-41(11)	46	24	41	5
1908	Vb	Veen, moerig	-86(61)	258	250	91	17
1909	Vb	Zeeklei	28(24)	44	42	36	3
1910	Vb	Zeeklei	-27(10)	52	45	43	22

Stratum	Gt	Grondsoort	ME(s.e.)	RMSE	SDE	MAE	n
1911	Vb	Zeeklei	4(9)	22	23	15	7
1912	Vb	Zeeklei	-11(11)	28	28	19	6
1913	V	Rivierklei	-23(3)	23	6	23	5
1914	Vb	Oude rivierklei	38(0)	38	0	38	4
1915	V	Zand	-1(0)	1	0	1	4
1916	Vb	Zeeklei met zware tussenlaag of ondergrond	3(12)	33	35	24	9
1917	V	Zeeklei met zware bovengrond	-13(15)	36	37	29	6
1918	V	Zeeklei	-6(16)	42	44	34	8
1919	V	Rivierklei met zware bovengrond	-4(12)	28	30	23	6
1920	V	Zand	-47(18)	72	58	56	10
1921	V	Alle	-149(73)	381	358	187	24
2001	VI	Zand	-14(11)	62	61	42	29
2002	VI	Zand	-45(20)	69	56	54	8
2003	VI	Keileem, oude klei	-20(24)	52	54	44	5
2004	VI	Zand, rivierklei	4(21)	48	52	45	6
2005	VI	Zand	-67(29)	117	100	80	12
2006	VI	Veen, moerig	1(11)	16	20	15	3
2007	VI	Veen, moerig met kleidek	7(4)	8	6	7	2
2008	VI	Zeeklei	19(11)	34	30	21	7
2009	VI	Zeeklei	-1(5)	32	33	27	45
2010	VI	Zeeklei	5(7)	27	27	23	14
2011	VI	Zeeklei met zware bovengrond, tussenlaag of ondergrond	6(12)	42	43	31	13
2012	VI	Rivierklei	-19(9)	42	38	34	17
2013	VI	Oude rivierklei	31(25)	39	35	31	2
2014	VI	Zeeklei	33(10)	36	17	33	3
2015	VI	Zand	-28(30)	66	67	36	5
2016	VI	Zand	-59(27)	103	88	72	11
2017	VI	Alle	-36(13)	79	71	55	30
2101	VII	Zand	11(24)	62	66	39	14
2102	VII	Zand	-44(0)	44	0	44	4
2103	VII	Zand	-54(51)	90	88	55	3
2104	VII	Zand, veen, keileem	17(24)	45	48	33	4
2105	VII	Zand	-165(198)	423	437	183	7
2106	VII	Zand	-125(78)	288	271	173	12
2107	VII	Zand	-11(56)	79	96	45	10
2108	VII	Zand	-90(25)	93	35	90	2
2109	VIII	Zand	-23(110)	133	171	41	12
2110	VIII	Zand	-7(38)	32	49	13	6
2111	VII	Zeeklei	30(10)	50	41	41	17
2112	VII	(Oude) rivierklei	8(18)	49	51	41	11
2113	VII	Leem, klei- en zandbrikgronden	-46(91)	102	129	91	2
2114	VII	Zand	-17(6)	18	8	17	2
2115	VII	Zand	-4(26)	14	29	4	5
2116	VIII	Zand	-24(608)	122	620	24	5
2117	VIII	Zand	-55(160)	235	279	62	19
2118	VIII	Zand	0(9)	2	9	0	23

Tabel B1.2

Validatieresultaten GLG-voorspellingen NHI, per stratum. ME: mean error (cm). RMSE: root mean squared error (cm). SDE: standard deviation of error (cm). MAE: mean absolute error (cm). n: aantal validatiepunten (-). Standaardfouten tussen haakjes. Bodems waarin de kans op stagnerende lagen binnen 120 cm groot is, zijn donker aangegeven (ing. H.R.J. Vroon, Alterra Wageningen UR, mondelinge mededeling).

Stratum	Gt	Grondsoort	ME(s.e.)	RMSE	SDE	MAE	n
1501	I	Veen	-19(7)	55	52	27	56
1502	I	Veen	-60(24)	117	103	66	18
1503	I	Zand	-103(41)	155	123	114	9
1504	I	Zee- en rivierklei	-8(7)	26	25	23	12
1505	I	Alle	-18(21)	84	84	77	20
1601	II	Veen	-4(5)	23	23	19	18
1602	II	Veen	-23(8)	49	44	30	34
1603	II	Veen	-25(7)	49	42	35	32
1604	II	Veen	-13(6)	33	30	24	24
1605	II	Zand, keileem	-33(47)	110	114	73	6
1606	II	Oude rivierklei, zand, leem, moerig	-20(15)	76	75	34	26
1607	II	Klei, zand, moerig	5(7)	40	40	27	33
1608	II	Zand	-34(9)	41	24	39	7
1609	II	Alle	-115(101)	456	453	155	20
1701	IIIb	Veen, moerig	-38(23)	102	97	63	18
1702	III	Veen, moerig	-13(6)	37	35	28	29
1703	III	Zand, keileem	-16(9)	43	41	29	23
1704	IIIb	Klei, leem	3(5)	35	35	28	47
1705	III	Klei, zand	-28(6)	44	35	36	36
1706	III	(Oude) rivierklei	-8(7)	38	38	26	27
1801	IV	Zand	6(8)	24	24	21	10
1802	IV	Zand, klei	9(8)	35	35	31	20
1803	IV	Veen, moerig	10(6)	25	23	18	15
1804	IV	(Zee-)klei	-4(5)	33	34	24	38
1805	IV	(Oude) rivierklei	-4(17)	47	50	39	9
1806	IV	Zeeklei, zand	20(11)	49	46	39	18
1807	IV	Alle	-43(25)	116	110	66	20
1901	Vb	Zand	-1(12)	53	54	40	19
1902	Vb	Zand	30(23)	60	57	50	6
1903	Vb	Zand	5(21)	42	47	30	5
1904	V	Zand	-27(43)	80	86	59	4
1905	Vb	Keileem, oude klei	-21(18)	94	93	49	27
1906	Vb	Klei, zand, beekdal	-13(28)	79	83	42	9
1907	Vb	Zand	-22(19)	44	43	22	5
1908	Vb	Veen, moerig	-37(48)	196	198	71	17
1909	Vb	Zeeklei	-34(11)	37	19	34	3
1910	Vb	Zeeklei	-17(8)	39	36	32	22
1911	Vb	Zeeklei	-24(10)	34	26	31	7
1912	Vb	Zeeklei	-18(7)	23	17	18	6
1913	V	Rivierklei	-7(7)	15	15	13	5
1914	Vb	Oude rivierklei	77(0)	77	NA	77	4
1915	V	Zand	-25(0)	25	0	25	4
1916	Vb	Zeeklei met zware tussenlaag of ondergrond	-14(8)	26	24	16	9
1917	V	Zeeklei met zware bovengrond	-8(5)	15	13	13	6

Stratum	Gt	Grondsoort	ME(s.e.)	RMSE	SDE	MAE	n
1918	V	Zeeklei	-12(15)	40	41	31	8
1919	V	Rivierklei met zware bovengrond	-99(17)	39	42	31	6
1920	V	Zand	-18(16)	51	51	41	10
1921	V	Alle	-124(65)	334	317	161	24
2001	VI	Zand	9(10)	56	56	42	29
2002	VI	Zand	-36(27)	81	77	51	8
2003	VI	Keileem, oude klei	-14(20)	43	45	37	5
2004	VI	Zand, rivierklei	2(18)	41	45	38	6
2005	VI	Zand	-45(29)	106	101	56	12
2006	VI	Veen, moerig	2(15)	22	26	20	3
2007	VI	Veen, moerig met kleidek	-4(28)	28	39	28	2
2008	VI	Zeeklei	12(7)	22	20	15	7
2009	VI	Zeeklei	-13(6)	39	37	30	45
2010	VI	Zeeklei	-8(9)	35	35	27	14
2011	VI	Zeeklei met zware bovengrond, tussenlaag of ondergrond	7(15)	51	53	32	13
2012	VI	Rivierklei	-19(8)	39	35	32	17
2013	VI	Oude rivierklei	37(53)	64	74	53	2
2014	VI	Zeeklei	64(29)	76	51	64	3
2015	VI	Zand	-5(44)	88	98	65	5
2016	VI	Zand	-29(26)	87	86	52	11
2017	VI	Alle	-28(12)	71	66	53	30
2101	VII	Zand	22(17)	50	62	28	14
2102	VII	Zand	-56(0)	56	0	56	4
2103	VII	Zand	9(30)	26	49	15	3
2104	VII	Zand, veen, keileem	1(35)	60	70	51	4
2105	VII	Zand	-100(207)	328	494	113	7
2106	VII	Zand	-51(93)	210	290	81	12
2107	VII	Zand	7(59)	64	141	26	10
2108	VII	Zand	-13(46)	47	64	46	2
2109	VIII	Zand	5(26)	19	63	5	12
2110	VIII	Zand	0(35)	27	65	12	6
2111	VII	Zeeklei	1(15)	60	62	43	17
2112	VII	(Oude) rivierklei	19(19)	50	61	29	11
2113	VII	Leem, klei- en zandbrikgronden	-	-	-	-	2
2114	VII	Zand	14(23)	26	32	23	2
2115	VII	Zand	-6(23)	15	44	6	5
2116	VIII	Zand	-24(593)	119	838	24	5
2117	VIII	Zand	-48(151)	220	487	56	19
2118	VIII	Zand	1(20)	5	41	1	23

Bijlage 2 Validatieresultaten Hydrologie Stone V2.3 per LSK-stratum

Tabel B2.1

Validatieresultaten GHG-voorspellingen Stone, per stratum. ME: mean error (cm). RMSE: root mean squared error (cm). SDE: standard deviation of error (cm). MAE: mean absolute error (cm). n: aantal validatiepunten (-). Standaardfouten tussen haakjes. De kleuren geven de grondsoorten aan. Bodems waarin de kans op stagnerende lagen binnen 120 cm groot is, zijn donker aangegeven (ing. H.R.J. Vroon, Alterra Wageningen UR, mondelinge mededeling).

Stratum	Gt	Grondsoort	ME(s.e.)	RMSE	SDE	MAE	n
1501	I	Veen	15(3)	27	23	21	56
1502	I	Veen	9(14)	58	59	42	18
1503	I	Zand	-51(21)	77	61	51	9
1504	I	Zee- en rivierklei	-8(17)	56	58	35	12
1505	I	Alle	-53(16)	85	68	62	20
1601	II	Veen	14(3)	20	15	16	18
1602	II	Veen	11(3)	20	18	16	34
1603	II	Veen	19(3)	24	14	20	32
1604	II	Veen	26(5)	34	23	31	24
1605	II	Zand, keileem	29(13)	40	30	38	6
1606	II	Oude rivierklei, zand, leem, moerig	12(5)	27	25	22	26
1607	II	Klei, zand, moerig	7(5)	26	26	19	33
1608	II	Zand	-38(59)	151	157	74	7
1609	II	Alle	5(8)	34	35	27	20
1701	IIIb	Veen, moerig	-9(8)	35	35	28	18
1702	III	Veen, moerig	1(7)	35	36	26	29
1703	III	Zand, keileem	14(6)	31	28	25	23
1704	IIIb	Klei, leem	9(4)	26	25	20	47
1705	III	Klei, zand	-32(11)	74	68	49	36
1706	III	(Oude) rivierklei	12(8)	41	40	34	27
1801	IV	Zand	18(8)	30	26	26	10
1802	IV	Zand, klei	16(5)	28	24	24	20
1803	IV	Veen, moerig	23(11)	47	42	36	15
1804	IV	(Zee-)klei	-7(5)	30	30	22	38
1805	IV	(Oude) rivierklei	-16(17)	52	52	39	9
1806	IV	Zeeklei, zand	9(10)	41	41	30	18
1807	IV	Alle	3(9)	41	42	33	20
1901	Vb	Zand	-6(8)	36	36	25	19
1902	Vb	Zand	3(16)	34	38	28	6
1903	Vb	Zand	8(15)	30	32	24	5
1904	V	Zand	-9(26)	46	52	45	4
1905	Vb	Keileem, oude klei	-21(8)	44	39	36	27
1906	Vb	Klei, zand, beekdal	7(11)	32	33	27	9
1907	Vb	Zand	-14(10)	25	23	19	5
1908	Vb	Veen, moerig	2(7)	29	29	24	17
1909	Vb	Zeeklei	-10(13)	21	23	19	3
1910	Vb	Zeeklei	-24(14)	70	67	35	22

Stratum	Gt	Grondsoort	ME(s.e.)	RMSE	SDE	MAE	n
1911	Vb	Zeeklei	-5(6)	16	16	15	7
1912	Vb	Zeeklei	-5(13)	30	32	28	6
1913	V	Rivierklei	3(4)	8	8	7	5
1914	Vb	Oude rivierklei	27(0)	27	0	27	4
1915	V	Zand	27(0)	27	0	27	4
1916	Vb	Zeeklei met zware tussenlaag of ondergrond	7(12)	34	36	30	9
1917	V	Zeeklei met zware bovengrond	-1(10)	22	24	19	6
1918	V	Zeeklei	-11(11)	31	31	27	8
1919	V	Rivierklei met zware bovengrond	4(7)	15	16	13	6
1920	V	Zand	-7(11)	34	35	29	10
1921	V	Alle	42(14)	80	70	67	24
2001	VI	Zand	8(7)	37	37	31	29
2002	VI	Zand	4(8)	21	23	18	8
2003	VI	Keileem, oude klei	48(54)	118	121	84	5
2004	VI	Zand, rivierklei	14(17)	42	43	33	6
2005	VI	Zand	26(21)	75	74	48	12
2006	VI	Veen, moerig	50(14)	60	35	50	6
2007	VI	Veen, moerig met kleidek	24(15)	44	40	38	7
2008	VI	Zeeklei	2(4)	26	26	20	45
2009	VI	Zeeklei	10(8)	31	30	22	14
2010	VI	Zeeklei	13(9)	33	31	28	13
2011	VI	Zeeklei met zware bovengrond, tussenlaag of ondergrond	-15(12)	50	49	36	17
2012	VI	Rivierklei	9(29)	30	41	29	2
2013	VI	Oude rivierklei	-20(53)	77	91	63	3
2014	VI	Zeeklei	21(11)	31	25	24	5
2015	VI	Zand	-10(16)	53	54	44	11
2016	VI	Zand	8(10)	54	54	43	30
2017	VI	Alle	43(22)	74	87	45	14
2101	VII	Zand	157(0)	157	0	157	4
2102	VII	Zand	84(48)	108	84	89	3
2103	VII	Zand	63(37)	89	74	65	4
2104	VII	Zand, veen, keileem	81(43)	121	123	81	7
2105	VII	Zand	85(20)	107	68	88	12
2106	VII	Zand	27(39)	64	111	29	10
2107	VII	Zand	92(26)	95	36	92	2
2108	VII	Zand	29(59)	80	168	35	12
2109	VIII	Zand	22(46)	48	104	22	6
2110	VIII	Zand	26(9)	46	39	33	17
2111	VII	Zeeklei	32(7)	43	35	32	22
2112	VII	(Oude) rivierklei	39(5)	6	7	5	2
2113	VII	Leem, klei- en zandbrikgronden	-3(33)	17	53	6	5
2114	VII	Zand	5(135)	27	191	5	5
2115	VII	Zand	20(25)	46	97	20	19
2116	VIII	Zand	2(74)	17	153	2	23
2117	VIII	Zand	15(3)	27	23	21	56
2118	VIII	Zand	9(14)	58	59	42	18

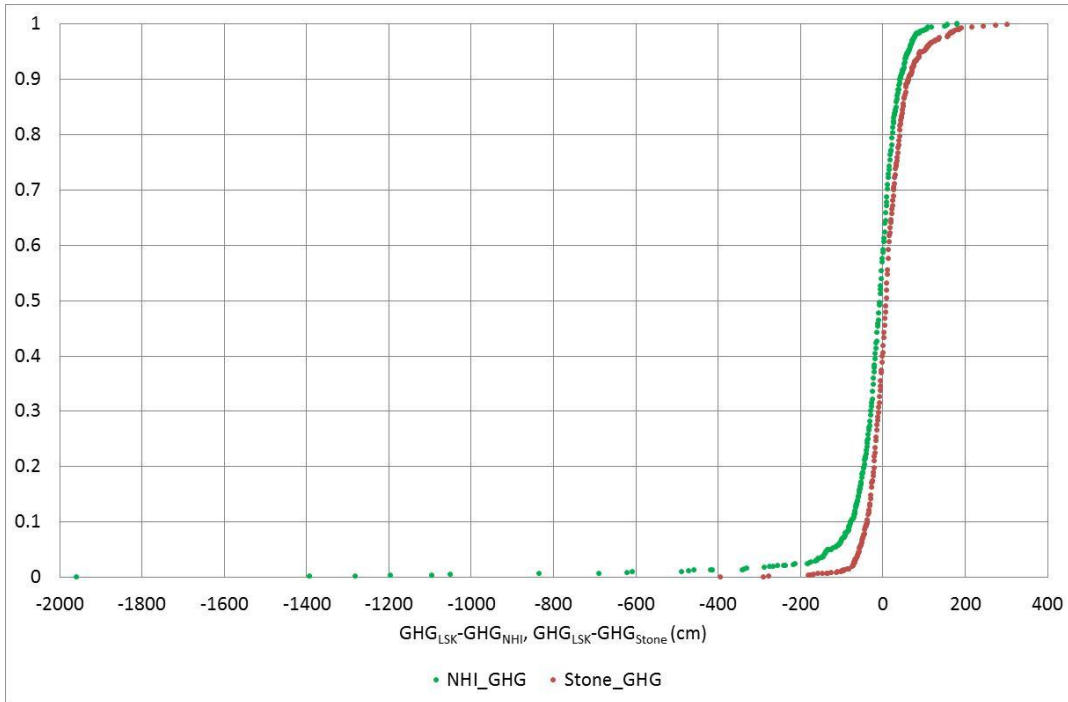
Tabel B2.2

Validatieresultaten GLG-voorspellingen Hydrologie Stone V2.3, per stratum. ME: mean error (cm). RMSE: root mean squared error (cm). SDE: standard deviation of error (cm). MAE: mean absolute error (cm). n: aantal validatiepunten (-). Standaardfouten tussen haakjes. De kleuren geven de grondsoorten aan. Bodems waarin de kans op stagnerende lagen binnen 120 cm groot is, zijn donker aangegeven (ing. H.R.J. Vroon, Alterra Wageningen UR, mondelinge mededeling).

Stratum	Gt	Grondsoort	ME(s.e.)	RMSE	SDE	MAE	n
1501	I	Veen	2(5)	38	38	28	56
1502	I	Veen	-16(20)	82	83	52	18
1503	I	Zand	-57(30)	102	89	74	9
1504	I	Zee- en rivierklei	-6(18)	60	63	42	12
1505	I	Alle	-76(21)	116	90	92	20
1601	II	Veen	19(5)	28	21	23	18
1602	II	Veen	8(4)	22	21	20	34
1603	II	Veen	19(4)	29	23	24	32
1604	II	Veen	12(8)	41	40	33	24
1605	II	Zand, keileem	11(23)	51	55	46	6
1606	II	Oude rivierklei, zand, leem, moerig	-4(8)	39	40	27	26
1607	II	Klei, zand, moerig	-9(7)	38	38	32	33
1608	II	Zand	-63(44)	124	115	63	7
1609	II	Alle	6(12)	54	56	37	20
1701	IIIb	Veen, moerig	-18(8)	39	35	30	18
1702	III	Veen, moerig	-9(9)	46	46	36	29
1703	III	Zand, keileem	-8(7)	34	34	26	23
1704	IIIb	Klei, leem	3(6)	40	40	30	47
1705	III	Klei, zand	-46(14)	95	85	53	36
1706	III	(Oude) rivierklei	-6(10)	54	54	41	27
1801	IV	Zand	-1(12)	35	37	30	10
1802	IV	Zand, klei	14(5)	25	22	19	20
1803	IV	Veen, moerig	36(13)	61	51	46	15
1804	IV	(Zee-)klei	-17(5)	33	29	23	38
1805	IV	(Oude) rivierklei	-16(31)	89	93	76	9
1806	IV	Zeeklei, zand	-2(9)	38	39	24	18
1807	IV	Alle	-1(11)	47	48	37	20
1901	Vb	Zand	12(12)	51	51	42	19
1902	Vb	Zand	50(15)	60	37	50	6
1903	Vb	Zand	18(36)	74	81	52	5
1904	V	Zand	5(32)	55	63	48	4
1905	Vb	Keileem, oude klei	-24(17)	90	88	70	27
1906	Vb	Klei, zand, beekdal	27(14)	47	41	42	9
1907	Vb	Zand	-4(25)	49	55	36	5
1908	Vb	Veen, moerig	36(13)	63	54	46	17
1909	Vb	Zeeklei	-40(15)	45	26	40	3
1910	Vb	Zeeklei	-33(18)	88	83	45	22
1911	Vb	Zeeklei	-17(12)	34	32	30	7
1912	Vb	Zeeklei	-23(7)	27	16	23	6
1913	V	Rivierklei	14(5)	17	11	14	5
1914	Vb	Oude rivierklei	-55(0)	55	NA	55	4
1915	V	Zand	-101(0)	101	0	101	4
1916	Vb	Zeeklei met zware tussenlaag of ondergrond	-3(6)	18	19	13	9
1917	V	Zeeklei met zware bovengrond	-7(10)	24	25	21	6

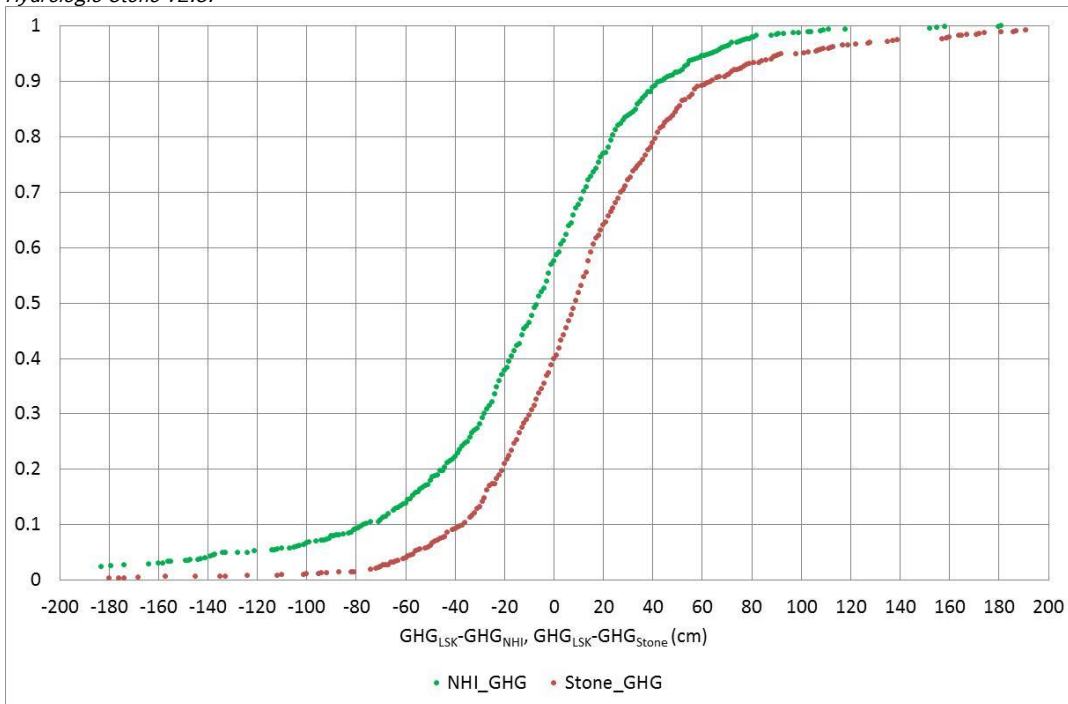
Stratum	Gt	Grondsoort	ME(s.e.)	RMSE	SDE	MAE	n
1918	V	Zeeklei	-7(15)	40	42	35	8
1919	V	Rivierklei met zware bovengrond	-18(13)	34	31	24	6
1920	V	Zand	-14(19)	59	60	41	10
1921	V	Alle	45(14)	82	71	67	24
2001	VI	Zand	32(7)	50	39	38	29
2002	VI	Zand	17(14)	40	38	35	8
2003	VI	Keileem, oude klei	74(51)	126	114	82	5
2004	VI	Zand, rivierklei	14(17)	41	43	37	6
2005	VI	Zand	37(15)	62	53	53	12
2006	VI	Veen, moerig	40(4)	41	10	40	6
2007	VI	Veen, moerig met kleidek	17(14)	39	38	33	7
2008	VI	Zeeklei	-10(6)	42	42	33	45
2009	VI	Zeeklei	-2(12)	43	45	34	14
2010	VI	Zeeklei	16(13)	48	47	33	13
2011	VI	Zeeklei met zware bovengrond, tussenlaag of ondergrond	-66(19)	100	77	76	17
2012	VI	Rivierklei	11(23)	25	33	23	2
2013	VI	Oude rivierklei	-21(84)	120	145	103	3
2014	VI	Zeeklei	50(11)	54	24	50	5
2015	VI	Zand	14(18)	58	59	49	11
2016	VI	Zand	-2(13)	70	72	53	30
2017	VI	Alle	49(19)	73	81	49	14
2101	VII	Zand	126(0)	126	NA	126	4
2102	VII	Zand	28(61)	58	102	28	3
2103	VII	Zand	21(31)	57	62	47	4
2104	VII	Zand, veen, keileem	42(37)	74	102	45	7
2105	VII	Zand	44(29)	80	101	47	12
2106	VII	Zand	28(53)	68	143	28	10
2107	VII	Zand	150(39)	155	55	150	2
2108	VII	Zand	7(41)	29	96	8	12
2109	VIII	Zand	26(55)	57	124	26	6
2110	VIII	Zand	1(7)	27	28	20	17
2111	VII	Zeeklei	10(6)	24	28	16	22
2112	VII	(Oude) rivierklei	31(23)	39	33	31	2
2113	VII	Leem, klei- en zandbrikgronden	-6(23)	15	43	6	5
2114	VII	Zand	4(110)	22	156	4	5
2115	VII	Zand	21(27)	48	104	21	19
2116	VIII	Zand	2(80)	19	165	2	23
2117	VIII	Zand	2(5)	38	38	28	56
2118	VIII	Zand	-16(20)	82	83	52	18

Bijlage 3 Cumulatieve frequentieverdelingen van verschillen LSK - GxG-modellen



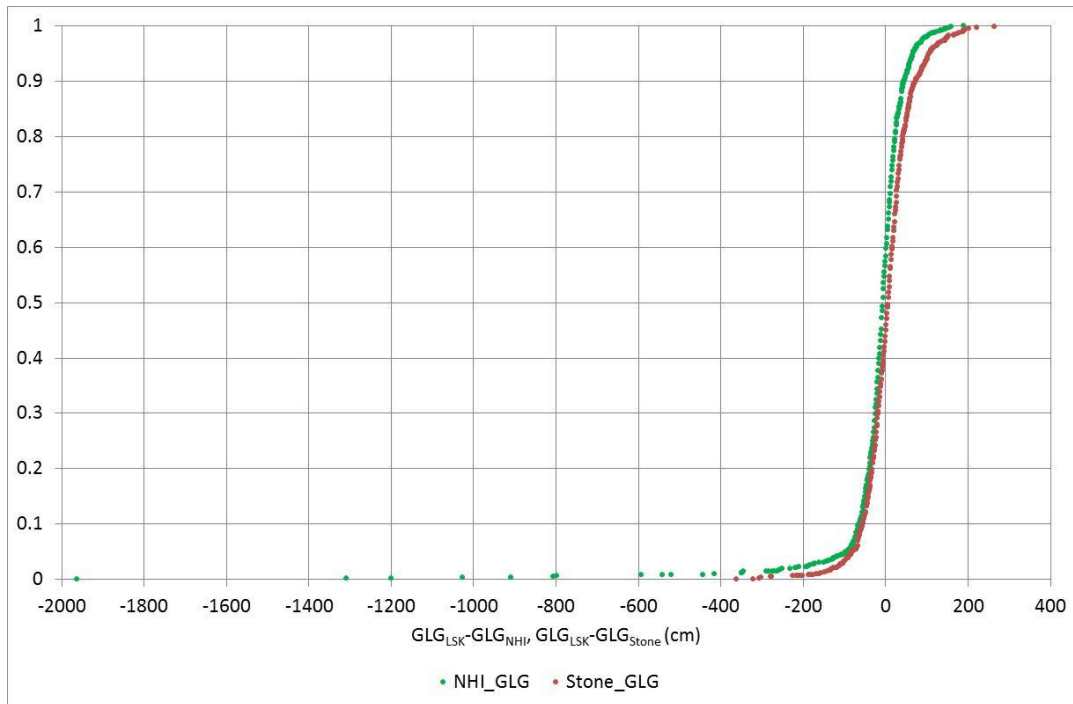
Figuur B3.1

Cumulative frequentieverdelingen van verschillen tussen GHG-schattingen met LSK en GHG-voorspellingen met NHI 3.0 en Hydrologie Stone V2.3.



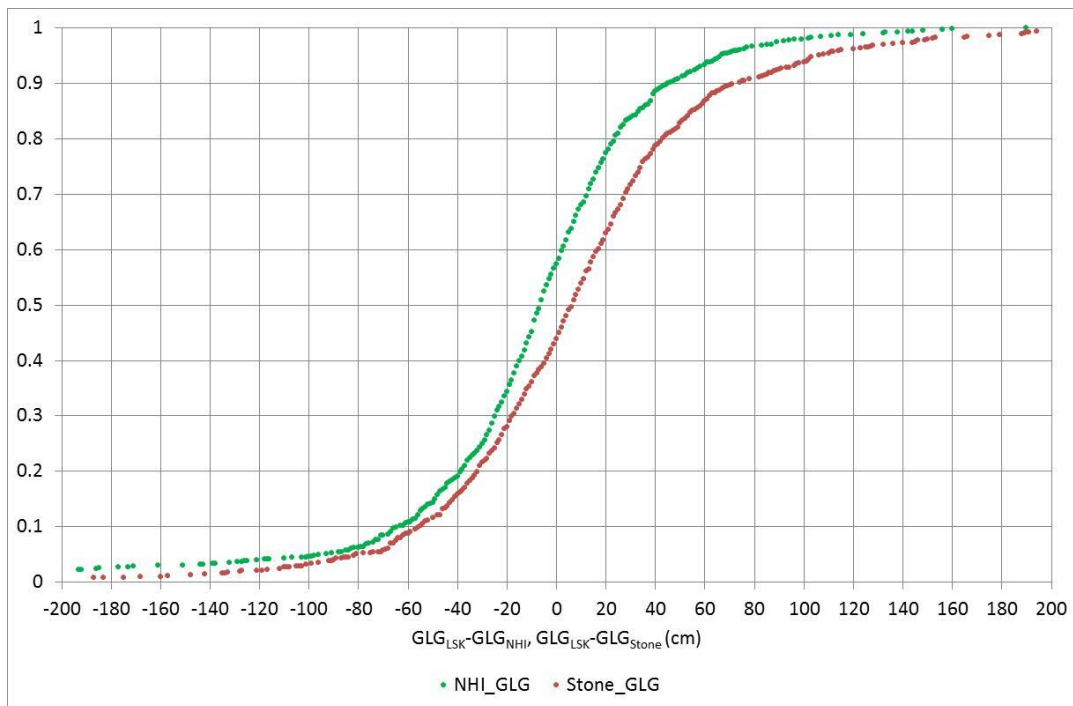
Figuur B3.2

Cumulative frequentieverdelingen van verschillen tussen GHG-schattingen met LSK en GHG-voorspellingen met NHI 3.0 en Hydrologie Stone V2.3, ingezoomd op verschillen tussen -200 en 200 cm. Figuur B3.1 geeft de volledige cumulatieve frequentieverdelingen.



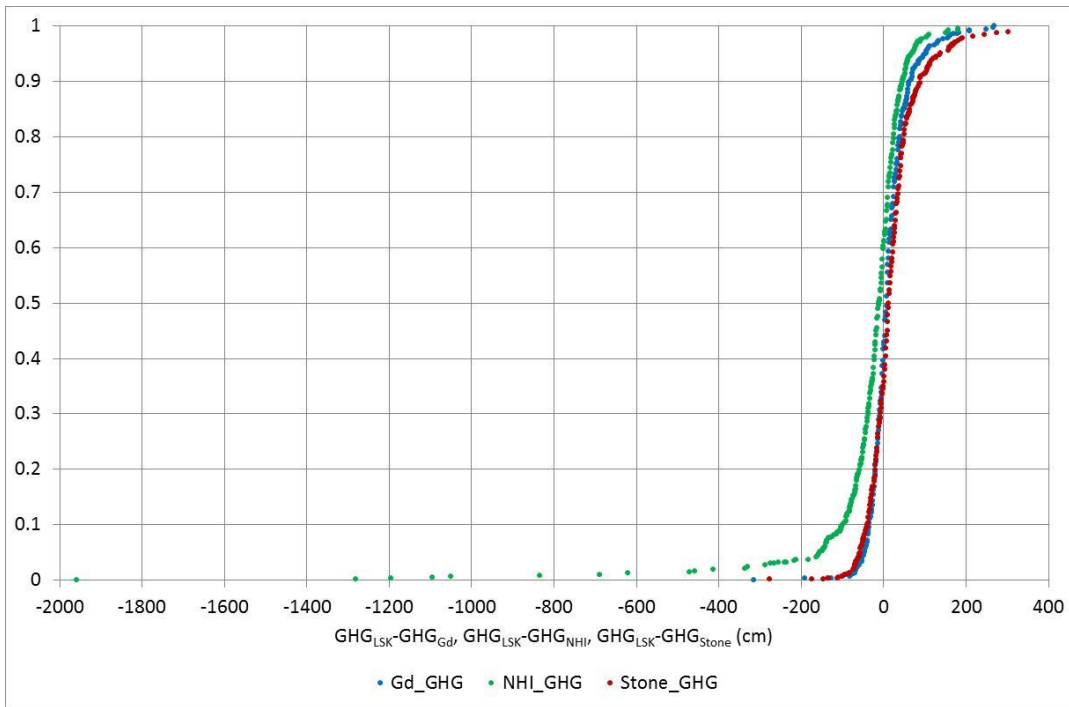
Figuur B3.3

Cumulative frequentieverdelingen van verschillen tussen GLG-schattingen met LSK en GLG-voorspellingen met NHI 3.0 en Hydrologie Stone V2.3.



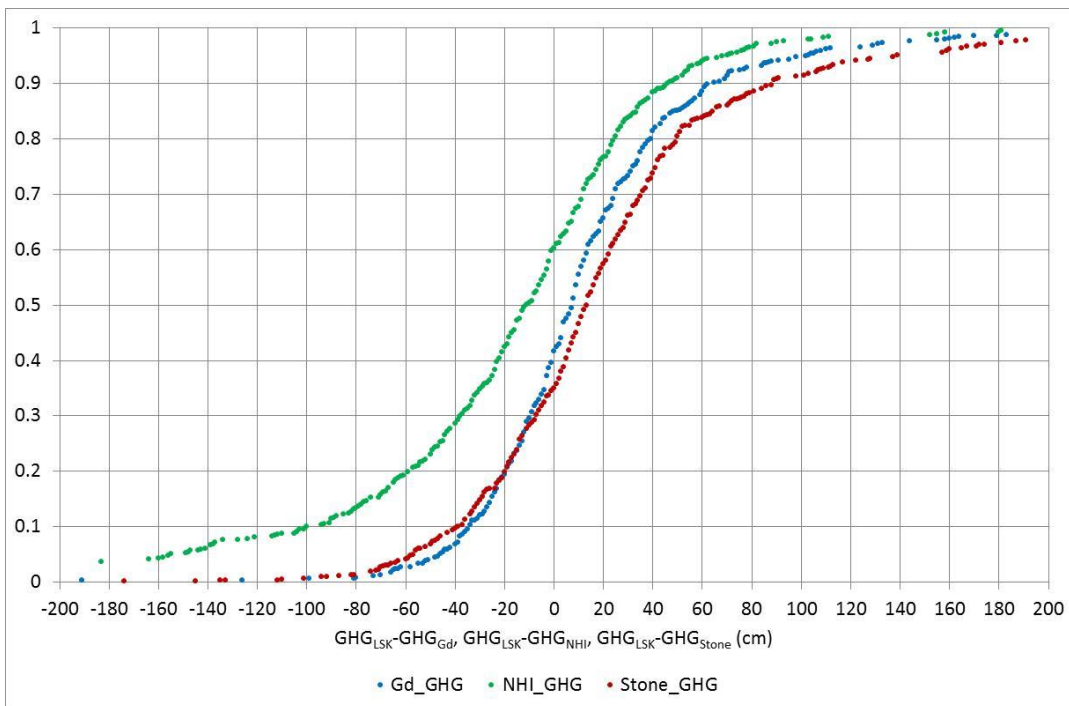
Figuur B3.4

Cumulative frequentieverdelingen van verschillen tussen GLG-schattingen met LSK en GHG-voorspellingen met NHI 3.0 en Hydrologie Stone V2.3, ingezoomd op verschillen tussen -200 en 200 cm. Figuur B3.3 geeft de volledige cumulatieve frequentieverdelingen.



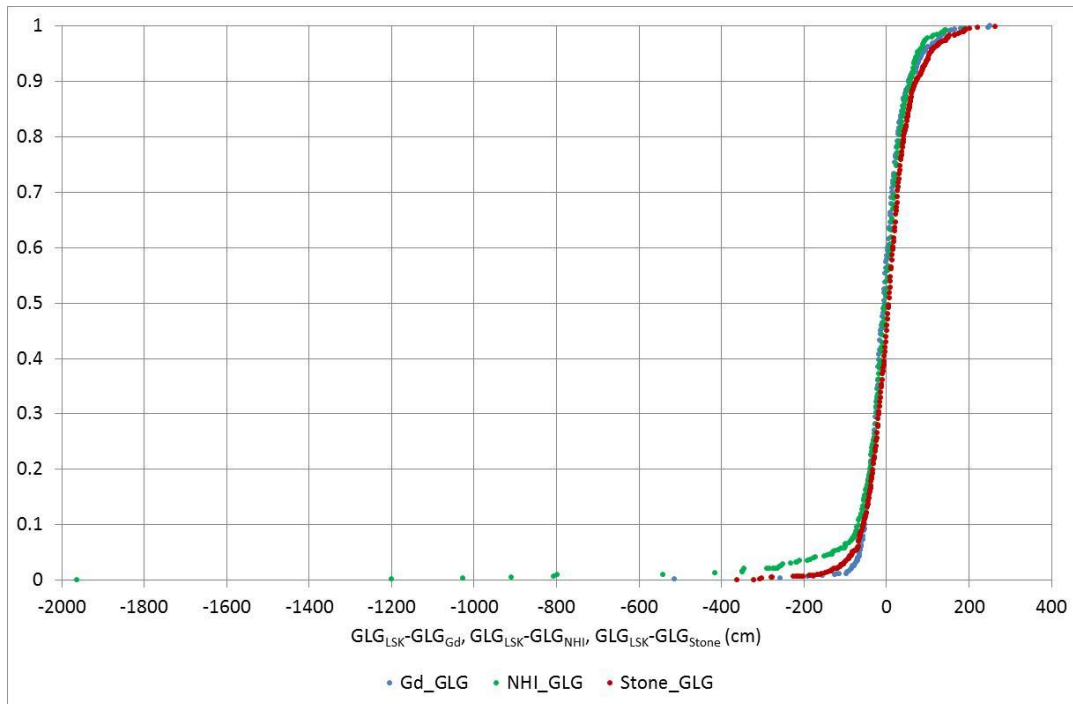
Figuur B3.5

Cumulatieve frequentieverdelingen van verschillen tussen GHG-schattingen met LSK en GHG-voorspellingen met de Gd-kaart, NHI 3.0 en Hydrologie Stone V2.3, voor het deel van Nederland waarvan een Gd-kaart is gemaakt.



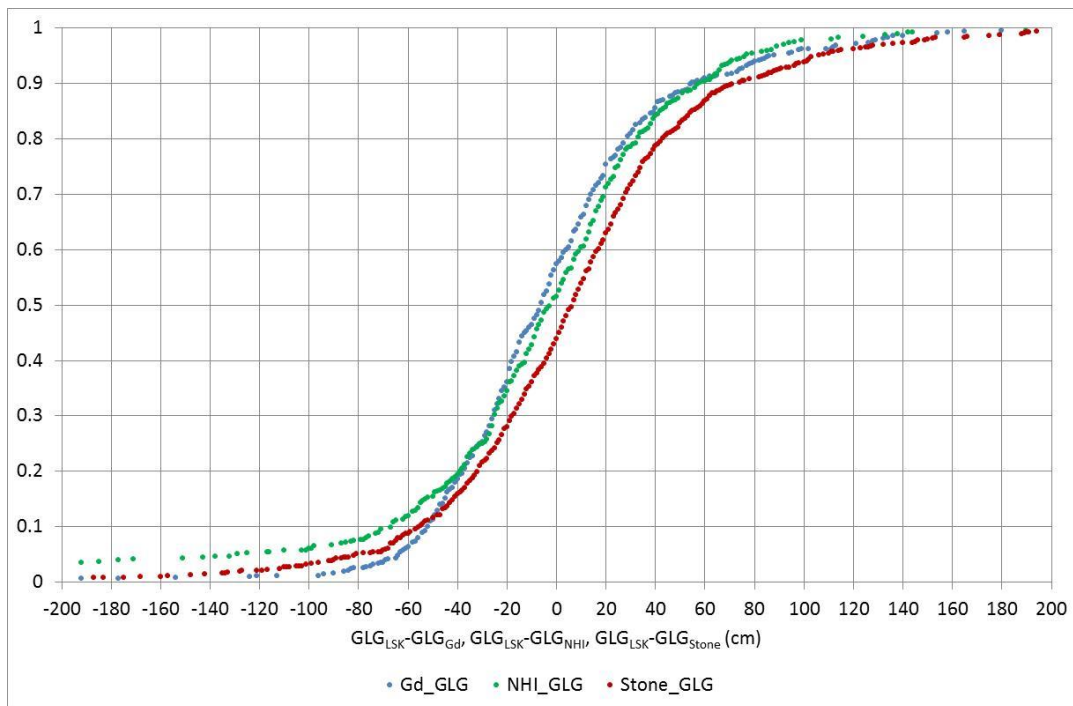
Figuur B3.6

Cumulatieve frequentieverdelingen van verschillen tussen GHG-schattingen met LSK en GHG-voorspellingen met de Gd-kaart, NHI 3.0 en Hydrologie Stone V2.3, voor het deel van Nederland waarvan een Gd-kaart is gemaakt, ingezoomd op verschillen tussen -200 en 200 cm. Figuur B3.5 geeft de volledige cumulatieve frequentieverdelingen.



Figuur B3.7

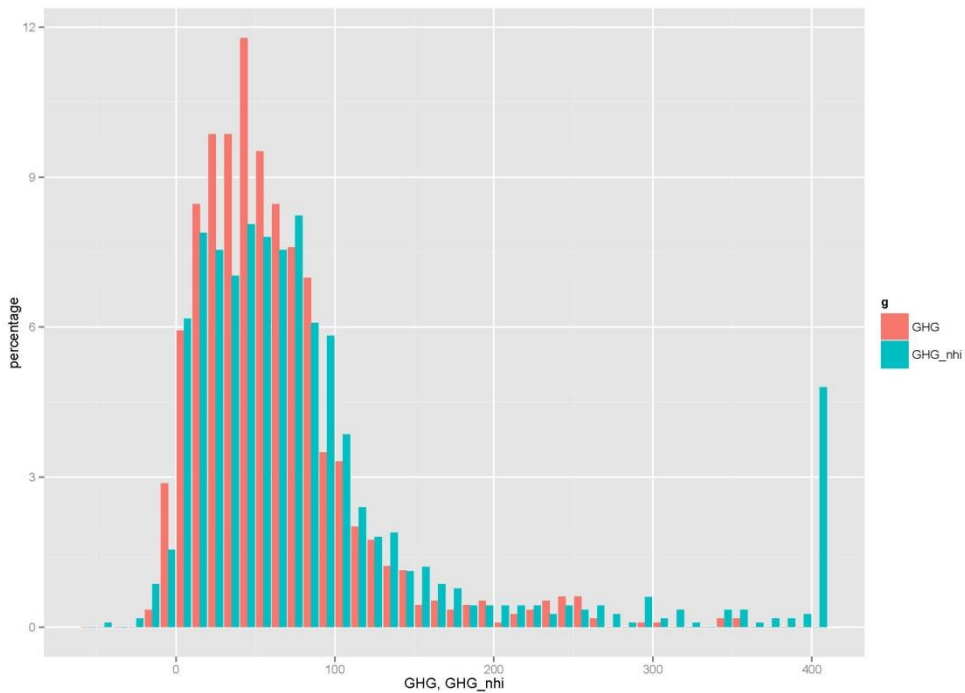
Cumulative frequentieverdelingen van verschillen tussen GLG-schattingen met LSK en GLG-voorspellingen met de Gd-kaart, NHI 3.0 en Hydrologie Stone V2.3, voor het deel van Nederland waarvan een Gd-kaart is gemaakt.



Figuur B3.8

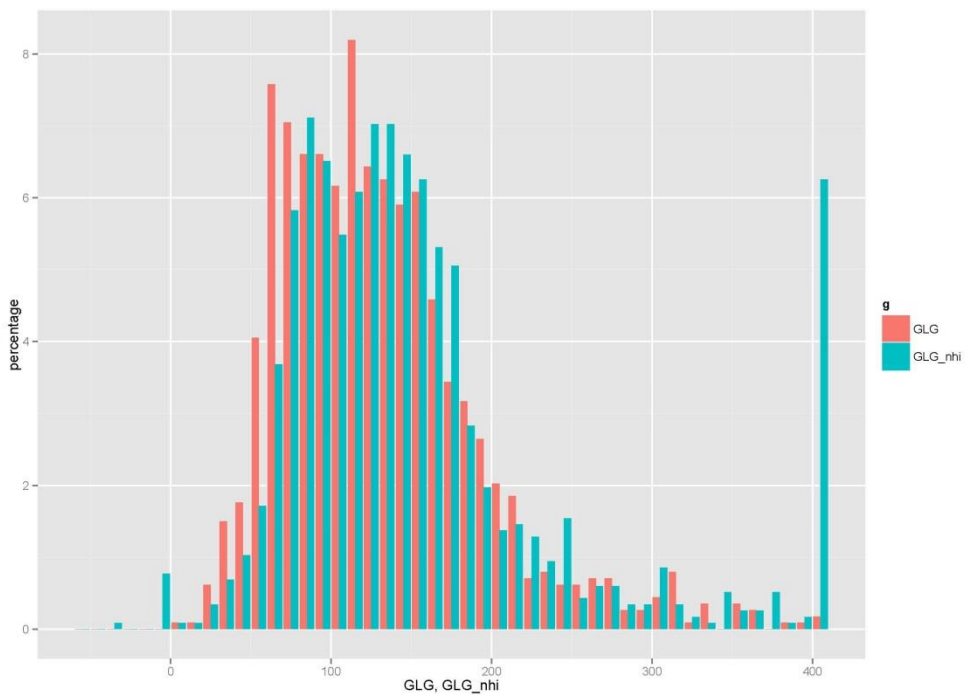
Cumulative frequentieverdelingen van verschillen tussen GLG-schattingen met LSK en GHG-voorspellingen met de Gd-kaart, NHI 3.0 en Hydrologie Stone V2.3, voor het deel van Nederland waarvan een Gd-kaart is gemaakt, ingezoomd op verschillen tussen -200 en 200 cm. Figuur B3.7 geeft de volledige cumulatieve frequentieverdelingen.

Bijlage 4 Histogrammen van GxG's uit LSK, NHI 3.0 en Hydrologie Stone V2.3



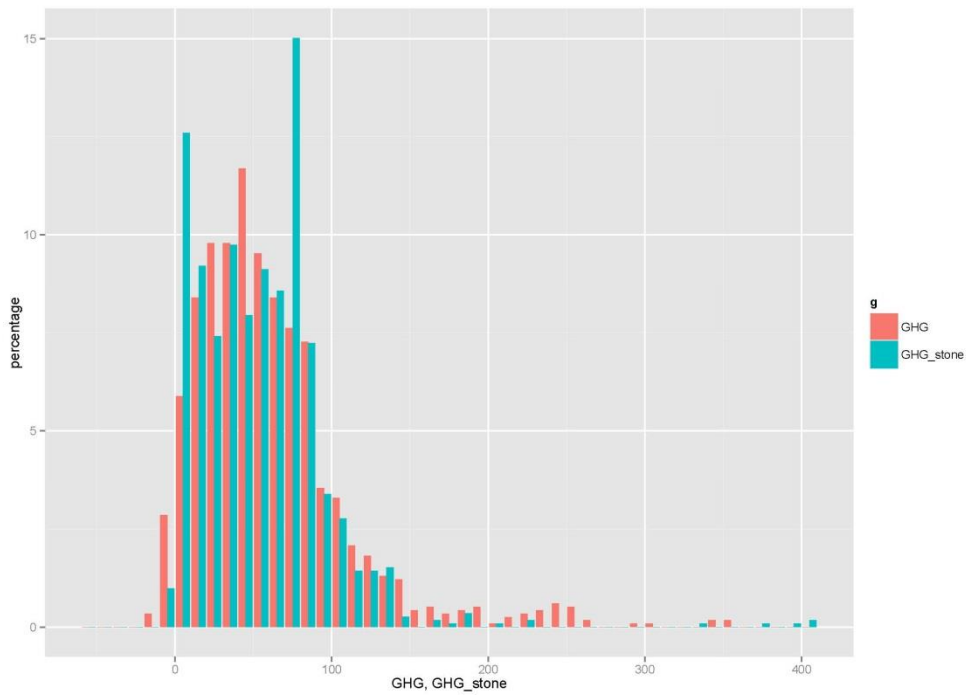
Figuur B4.1

Histogrammen van GHG's uit de LSK en NHI 3.0. De klassebreedte is 10 cm. Waarden groter dan 400 cm zijn samengevat in één klasse.



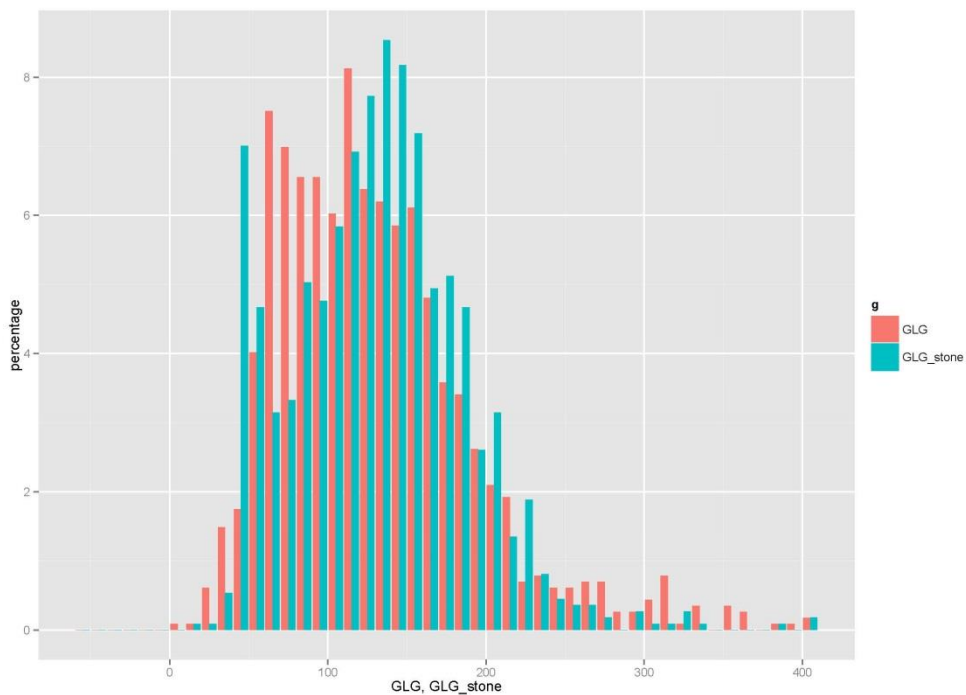
Figuur B4.2

Histogrammen van GLG's uit de LSK en NHI 3.0. De klassebreedte is 10 cm. Waarden groter dan 400 cm zijn samengevoegd in één klasse.



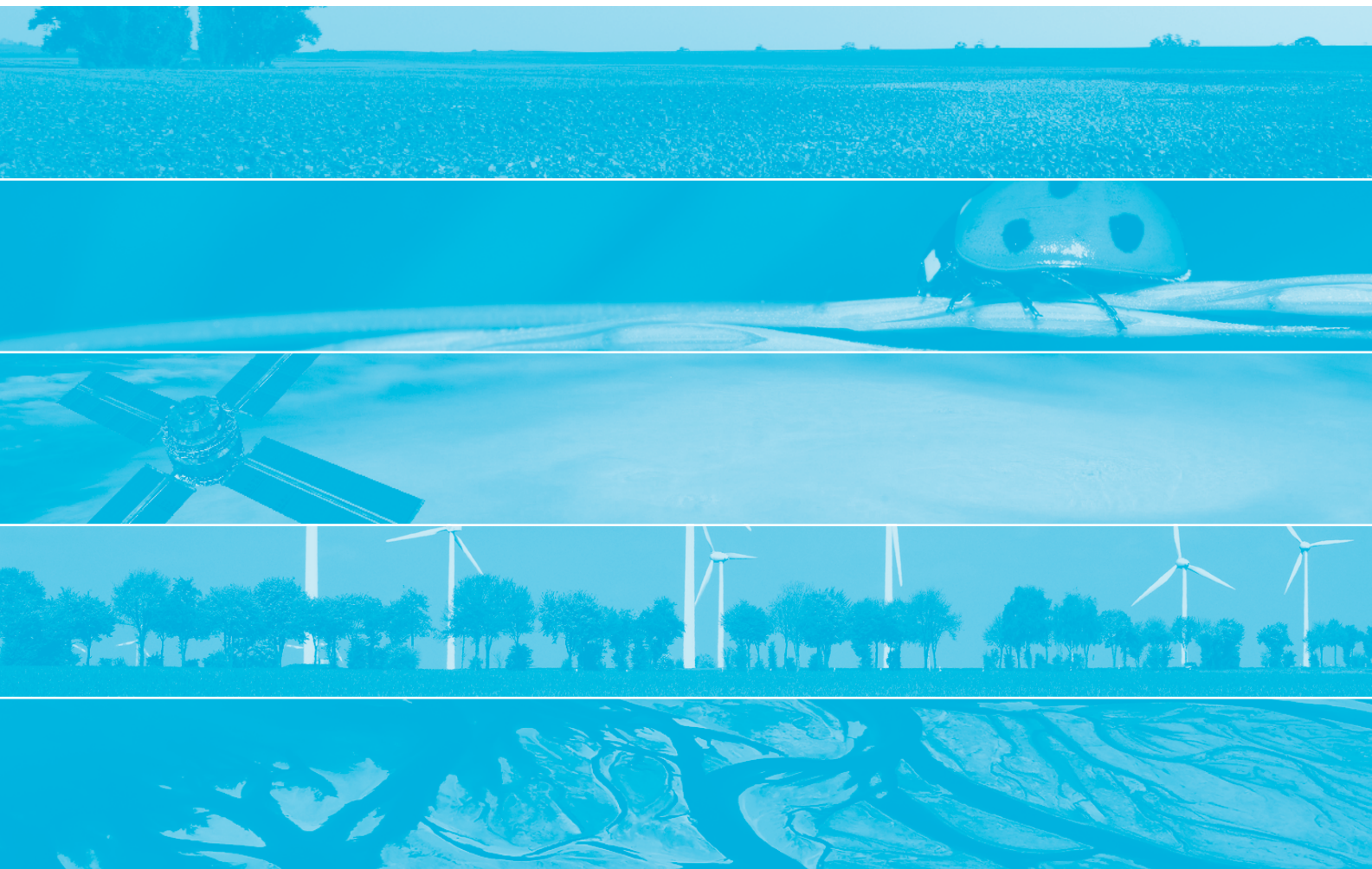
Figuur B4.3

Histogrammen van GHG's uit de LSK en Hydrologie Stone V2.3. De klassebreedte is 10 cm. Waarden groter dan 400 cm zijn samengevat in één klasse.



Figuur B4.4

Histogrammen van GLG's uit de LSK en Hydrologie Stone V2.3. De klassebreedte is 10 cm. Waarden groter dan 400 cm zijn samengevat in één klasse.



Alterra is onderdeel van de internationale kennisorganisatie Wageningen UR (University & Research centre). De missie is 'To explore the potential of nature to improve the quality of life'. Binnen Wageningen UR bundelen negen gespecialiseerde en meer toegepaste onderzoeksinstituten, Wageningen University en hogeschool Van Hall Larenstein hun krachten om bij te dragen aan de oplossing van belangrijke vragen in het domein van gezonde voeding en leefomgeving. Met ongeveer 40 vestigingen (in Nederland, Brazilië en China), 6.500 medewerkers en 10.000 studenten behoort Wageningen UR wereldwijd tot de vooraanstaande kennisinstellingen binnen haar domein. De integrale benadering van de vraagstukken en de samenwerking tussen natuurwetenschappelijke, technologische en maatschappijwetenschappelijke disciplines vormen het hart van de Wageningen Aanpak.

Alterra Wageningen UR is het kennisinstituut voor de groene leefomgeving en bundelt een grote hoeveelheid expertise op het gebied van de groene ruimte en het duurzaam maatschappelijk gebruik ervan: kennis van water, natuur, bos, milieu, bodem, landschap, klimaat, landgebruik, recreatie etc.

Meer informatie: www.wageningenUR.nl/alterra