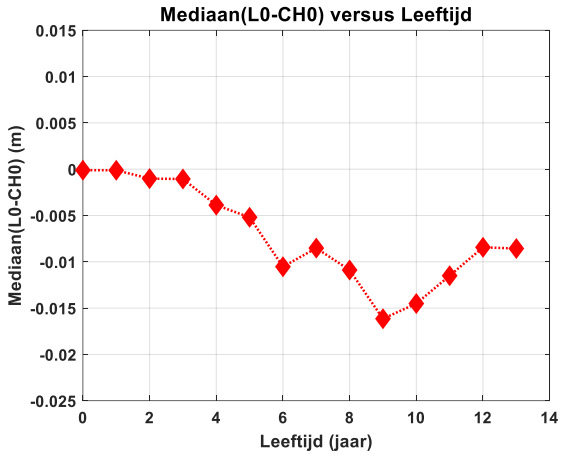
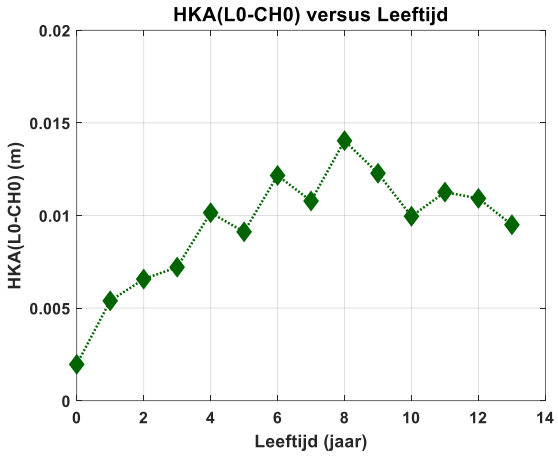
Kenmerken meetfout bij automatisch meten grondwaterstand door Provincie Overijssel





amo
Modellering en Optimalisatie

Icastat

Kenmerken meetfout bij automatisch meten grondwaterstand door Provincie Overijssel

Aanvrager:
Provincie Overijssel (ir. Thomas de Meij)

Auteurs:
drs. Paul K. Baggelaar (Icastat)
ir. Eit C.J. van der Meulen (AMO)

14 mei 2019

AMO (Adviesbureau Modellering en Optimalisatie)
Deldenerstraat 81
7551 AC Hengelo
Tel.: 074-7850124 / 06-24865321
Email : amo@amo-nl.com

INHOUD

SAMENVATTING **3**

1 INLEIDING **7**

1.1 ACHTERGRONDEN EN VRAAGSTELLING 7
1.2 ONTVANGEN BASISINFORMATIE 8
1.3 OVER DIT RAPPORT 8
1.4 STUDIEBEGELEIDING 8

2 THEORETISCHE AFLEIDING KENMERKEN MEETFOUT **10**

2.1 DEFINITIES RELEVANTE KENGETALLEN VAN DE MEETFOUT 10
2.2 ONDERDELEN EN FOUTENBRONNEN MEETPROCEDURE PROVINCIE OVERIJSSEL 12
2.2.1 Onderdelen meetprocedure Provincie Overijssel 12
2.3 MOGELIJKE FOUTENBRONNEN PROCEDURE-ONDERSTEUNENDE SOFTWARE 16
2.3.1 Schets van ideale combinatie van software en procedure 17
2.3.2 Omrekenen verschildruk naar waterhoogte 18
2.3.3 Omrekenen loggermetingen naar ander referentieniveau 19
2.3.4 Valideren van reeks ruwe loggermetingen (L0) 22
2.3.5 Conclusies over foutenbronnen procedure-ondersteunende software 22
2.4 MOGELIJKE INVLOED VAN CORRIGEREN RUWE LOGGERMETINGEN OP MEETFOUT 22
2.5 MEETFOUTEN KELLER-LOGGER DCX-22 AA 25
2.5.1 Principe van meten grondwaterstand met logger 25
2.5.2 Systematische fouten door instelling fysische constanten 26
2.5.3 Kenmerken meetfout druksensoren Keller DCX-22 AA volgens technische specificaties 29
2.5.4 Afleiden kenmerken gecombineerde meetfout druksensoren Keller DCX-22 AA 30
2.6 KERNPUNTEN VAN DIT HOOFDSTUK 35

3 EMPIRISCHE AFLEIDING KENMERKEN MEETFOUT **37**

3.1 BESCHOUWDE VERSCHILLEN TUSSEN LOGGERMETINGEN EN HANDMETINGEN 37
3.2 VERZAMELEN EN VOORBEWERKEN VAN LOGGERMETINGEN EN HANDMETINGEN 38
3.3 KENMERKEN VAN DE VERSCHILLEN TUSSEN LOGGERMETING EN HANDMETING 39
3.4 AFLEIDEN STANDAARDAFWIJKING VAN DE MEETFOUT 41
3.5 KERNPUNTEN VAN DIT HOOFDSTUK 44

4 ANALYSE INVLOEDSFACTOREN MEETFOUT LOGGERS **45**

4.1 LEVENSDUUR EN STORINGSFREQUENTIE LOGGERS 45
4.2 RELATIE LEEFTIJD LOGGER EN FUNCTIONEREN LOGGER 47
4.2.1 Relatie L0-CH0 en leeftijd logger 47
4.2.2 Trends in CL0-CH0 50
4.3 RELATIE L0-CH0 EN MAAND 52
4.4 RELATIE L0-CH0 EN WATERHOOGTE TEN OPZICHTE VAN DRUKSENSOR 53
4.5 RELATIE L0-CH0 EN HYDROLOGISCHE CONDITIES MEETLOCATIE 55
4.6 KERNPUNTEN VAN DIT HOOFDSTUK 60

5 OPZET STEEKPROEF NAAR KENMERKEN MEETFOUT **63**

5.1 VEREISTEN VAN STEEKPROEF NAAR KENMERKEN MEETFOUT LOGGER 63
5.2 RELATIE STEEKPROEFGROOTTE EN NAUWKEURIGHEID SCHATTER MEETONZEKERHEID 63
5.3 SELECTIE VAN LOGGERS VOOR DE STEEKPROEF 66
5.4 STATISTISCHE VERWERKING STEEKPROEFRESULTATEN 67
5.5 ENKELE ADVIEZEN OVER UITVOERING VAN DE STEEKPROEF 67
5.6 KERNPUNTEN VAN DIT HOOFDSTUK 68

6 AANBEVELINGEN 69

AANGEHAALDE LITERATUUR 70

BIJLAGE: TRENDPLOTS CL0-CH0 PERIODE 2004 T/M 2015 71

Samenvatting

Deze studie brengt op verzoek van de Provincie Overijssel de kenmerken in beeld van de meetfout bij meten van de grondwaterstand met een automatische drukopnemer volgens de procedure van de Provincie. Daarbij zijn antwoorden geformuleerd op een aantal vragen, zoals hieronder samengevat.

1. Wat zijn de langs theoretische weg afgeleide kenmerken van de meetfout en welke procescomponenten bepalen die kenmerken in belangrijke mate?

Aan de hand van een aantal technische specificaties van de Keller-logger DCX-22 AA is langs theoretische weg de standaardafwijking van zijn meetfout afgeleid, voor de situatie dat de logger nog geen kwaliteitsverslechtering vertoont. Deze blijkt dan tussen 0,8 en 1,3 cm waterhoogte te kunnen liggen (afhankelijk van de interpretatie van de technische specificaties van de logger).

In het algemeen kunnen de volgende factoren belangrijk zijn voor de kenmerken van de meetfout: meetcapaciteit (maximaal te meten druk), watertemperatuur, luchttemperatuur, waterhoogte ten opzichte van de druksensor, instellingen van de fysische constanten die worden gebruikt bij het omrekenen van verschildruk naar waterhoogte (soortelijk gewicht water en zwaartekrachtsversnelling), grootte en frequentie van druk- en temperatuurvariaties, trillingen, schokken en loggerleeftijd.

We hebben geen ernstige fouten geconstateerd van de procedure-ondersteunende software, te weten de Keller-software en de HydroData-module van Menyanthes. Er is wel sprake van een onvolkomenheid, die ontstaat doordat een controleloggermeting (CL0) met de Keller-software wordt omgezet van drukverschil naar waterhoogte en daarna naar meter ten opzichte van meetpunt of NAP, terwijl dat omzetten voor een ruwe loggermeting (L0) geschiedt met de HydroData-module van Menyanthes. Door niet-overeenkomende instellingen van deze softwarepakketten is CL0 echter weinig representatief voor de ruwe loggermetingen en daardoor speelt deze vanaf 2010 ook geen rol meer bij het valideren. Dit is inefficiënt, want het zo belangrijke validatieproces heeft baat bij zoveel mogelijk vergelijkingsmateriaal. Bovendien wordt CL0 vrijwel gelijktijdig met de controlehandmeting (CH0) gemeten, terwijl dat doorgaans niet geldt voor de dichtstbijzijnde L0.

Er zijn verder enkele onvolkomenheden geconstateerd in details van de uitvoering van de procedure, zoals:

- de starttijd van loggerdiepte en meetpuntshoogte blijkt in 747 gevallen (6,4% van alle controledatums) niet juist te zijn geregistreerd. Dit trad vooral op tot 2015 en kan in theorie hebben geleid tot fouten bij het omrekenen van de controleloggermeting en de controlehandmeting (beide oorspronkelijk uitgedrukt ten opzichte van meetpunt) naar een ander referentieniveau en fouten bij het omrekenen van loggermetingen (uitgedrukt in hoogte waterkolom ten opzichte van de sensordiepte) naar een ander referentieniveau, maar in de praktijk zal het nauwelijks zijn voorgekomen. Deze onvolkomenheid is overigens in 2015 opgelost;
- de criteria voor correctie van een ruwe loggerreeks op basis van de bij controles vastgestelde verschillen tussen ruwe loggermetingen (L0) en daaraan gepaarde handmetingen (CH0) zijn niet statistisch onderbouwd, wat kan leiden tot het beschadigen van meetinformatie. We adviseren nader te onderzoeken of gebruik kan worden gemaakt van technieken uit het vakgebied Statistische Procesbeheersing om het meetproces beheerst - dat wil zeggen, zonder kwaliteitsverslechtering - te houden. Daarvoor zal ondermeer kennis nodig zijn van de statistische

kenmerken van de verschilwaarde L0-CH0, representatief voor de fase dat het meetproces nog geen kwaliteitsverslechtering vertoont (zoals door drift). Het nader onderzoek dient ook na te gaan of statistische trendanalyse kan helpen het meetproces beheerst te houden.

2. Hoe groot zijn de verschillen tussen ruwe (ongevalideerde) loggermetingen en daaraan gepaarde controlehandmetingen?

De kenmerken van de ruwe meetfout zijn empirisch afgeleid uit de verschillen tussen ruwe loggermetingen (L0) en de daaraan gepaarde controlehandmetingen (CH0), die zijn uitgevoerd bij het circa halfjaarlijks uitlezen van een logger. Daarbij is verdisconteerd voor de in een andere studie bepaalde meetfout van de handmetingen (gemiddelde van 0,50 cm en standaardafwijking van 0,84 cm). Het gemiddelde verschil van L0 en CH0 blijkt -0,04 cm en de standaardafwijking van het verschil blijkt 3,63 cm. Na verdiscontering van de meetfout van CH0 wordt het gemiddelde van de ruwe meetfout van de logger geraamd op 0,46 cm en de standaardafwijking op 3,53 cm. Deze kengetallen hebben echter betrekking op een verzameling loggers die heterogeen is voor wat betreft de kwaliteit van de loggermeting.

De kengetallen kunnen bruikbaarder worden als we ze uitsplitsen naar homogenere deelverzamelingen loggers. De standaardafwijking van de ruwe meetfout van een logger van leeftijd 0 jaar wordt geraamd op minder dan 1,40 cm en het gemiddelde op 0,56 cm. Deze geraamde standaardafwijking van de meetfout is al duidelijk lager dan die van de heterogene verzameling loggers, wat zal komen doordat er bij leeftijd 0 jaar nog geen sprake zal zijn van sterke drift. De raming van een standaardafwijking van minder dan 1,40 cm sluit overigens goed aan bij de theoretisch afgeleide raming die uitkwam tussen 0,8 en 1,3 cm (zie het antwoord bij vraag 1). Deze kengetallen kunnen daarom bijvoorbeeld dienen als uitgangspunt voor correctiecriteria van ruwe loggermetingen.

3. Zijn de verschillen tussen ruwe (ongevalideerde) loggermetingen en daaraan gepaarde controlehandmetingen gerelateerd aan factoren zoals het seizoen, of de hydrologische condities van de meetlocatie?

We zijn voor een aantal factoren nagegaan of deze gerelateerd zijn aan het verschil tussen een ruwe loggermeting (L0) en een daaraan gepaarde controlehandmeting (CH0), aangezien dit verschil is op te vatten als een benadering van de meetfout (de meetfout is immers niet exact gelijk aan L0-CH0, doordat CH0 ook een meetfout heeft). Dit leverde de volgende bevindingen:

- Er blijkt sprake van een duidelijke relatie tussen L0-CH0 en de loggerleeftijd. Het centrum van L0-CH0 daalt tot een leeftijd van circa negen jaar, om daarna weer enigszins te stijgen. De spreiding van L0-CH0 neemt daarentegen duidelijk toe met de leeftijd. Dit laatste zal mede veroorzaakt zijn door drift.
- Er blijkt sprake van een seizoenseffect op L0-CH0, waarbij de mediaan van L0-CH0 in het voorjaar doorgaans het laagst is (tot -5 mm) en in de tweede jaarhelft het hoogst (tot -1 mm). De tegenstelling lente – tweede jaarhelft blijkt behalve voor de mediaan van L0-CH0 ook min of meer op te gaan voor het 25- en het 75-percentiel. De standaardafwijking van L0-CH0 lijkt nauwelijks onderhevig aan een seizoenseffect, maar dit kengetal is minder bruikbaar, doordat het sterk wordt beïnvloed door uitschieters.
- Zowel het centrum als de spreiding van L0-CH0 blijkt monotoon gerelateerd aan de gemeten waterhoogte van de bijbehorende controleloggermeting (CL0) ten opzichte van de druksensor. Als die waterhoogte toeneemt wordt het centrum van L0-CH0 negatiever, terwijl de spreiding van L0-CH0 toeneemt. Aangezien het denkbaar is dat dit causale relaties zijn, is het aan te bevelen de

loggerdiepte zo beperkt mogelijk te houden, uiteraard wel rekening houdende met de ter plekke geldende kansverdeling van de grondwaterstand.

- Ten slotte is met grafische analyses nagegaan of het verschil van L0 en CH0 wellicht gerelateerd is aan hydrologische karakteristieken van de meetlocatie. We vonden geen duidelijke relaties van L0-CH0 met de gemiddelde grondwaterstand ten opzichte van maaiveld, met de GHG, met de GHG-GLG en met de standaardafwijking van de grondwaterstand. Er bleek daarentegen wel sprake van een relatie met de GLG. Op locaties met een hoge GLG is het centrum van L0-CH0 doorgaans lager en de spreiding van L0-CH0 groter dan op locaties met een lage GLG.

Een andere indicatie van de invloed van de loggerleeftijd op de meetfout is het optreden van trend in de meetfout, wat we kunnen opvatten als een manifestatie van drift. We hebben trendanalyse uitgevoerd op alle reeksen CL0-CH0 (uniek betreffende loggernummer en combinatie peilbuis en filter) die voldoen aan de criteria dat de reeks minimaal vijf verschilwaarden CL0-CH0 omvat en minimaal vijf jaar lang is. Daaruit bleek dat 52 (18,4%) van de 283 geanalyseerde reeksen CL0-CH0 een statistisch significante trend vertoonden, met enigszins meer stijgingen (32) dan dalingen (20). De verdeling van de 283 geschatte trendhellingen is vrij symmetrisch, met - ongeacht statistische significantie - 139 positieve hellingen, 8 nulhellingen en 136 negatieve hellingen en met mediaan 0,00 cm/jaar en gemiddelde 0,02 cm/jaar. De maximale statistisch significante trend bedraagt 0,58 cm/jaar en de minimale bedraagt -1,10 cm/jaar. Als we alle statistisch niet-significante trendhellingen op nul zetten komt de standaardafwijking van de trendhelling uit op 0,14 cm/j. Dit is op te vatten als een statistische maat voor de drift die zich tot medio 2015 manifesteerde in het meetproces van CL0, ondanks de regelmatige aanpassingen van de installatiediepte, die juist bedoeld waren om drift tegen te gaan. Het feit dat ruim 18% van de geanalyseerde reeksen CL0-CH0 een statistisch significante trend vertoont, duidt op redelijk vaak optredende drift. Aangezien dit nog een onderschatting zal zijn van het werkelijke percentage drift (door de corrigerende aanpassingen van de installatiediepte), onderstreept dit de noodzaak om in ieder geval loggermetingen regelmatig te valideren en zonodig te corrigeren. Dat is het geval in het meetproces van de Provincie, al verdient het nog aanbeveling te onderzoeken of de correctiecriteria beter onderbouwd kunnen worden.

Er is ook een analyse uitgevoerd van de levensduur van de loggers, gedefinieerd als de loggerleeftijd bij vervanging vanwege een storing die niet is veroorzaakt door externe factoren (zoals werkzaamheden). Over de beschouwde periode (20 maart 2003 tot 30 januari 2018) is gebruik gemaakt van 1.165 loggers voor de 549 peilfilters waar we informatie over ontvingen. Dit betrof de 549 initieel geïnstalleerde loggers en 616 vervangingen daarvan. De levensduurverdeling blijkt niet symmetrisch, maar scheef naar rechts. De mediane levensduur is dientengevolge lager dan de gemiddelde levensduur (mediaan bedraagt 3,37 jaar en gemiddelde bedraagt 4,28 jaar). Verder blijkt dat in het eerste jaar 7,9% van de loggers is vervangen vanwege een storing, na drie jaar 43,9% en na vijf jaar 71,8%. De langste levensduur (bij vervanging) bedroeg 13,8 jaar.

Met de ons verstrekte gegevens over datum en reden van loggervervangingen hebben we ook de ontwikkeling van de storingsfrequentie als functie van de leeftijd gereconstrueerd. Het blijkt dat de storingsfrequentie van loggers tot de leeftijd vier jaar duidelijk toeneemt. Vervolgens neemt deze geleidelijk weer af tot de leeftijd tien jaar, waarna de storingsfrequentie weer toeneemt. Er is dus een relatief groot percentage uitvallers van jonge leeftijd, maar daarnaast zijn er ook meerdere overlevers, die aanvankelijk zonder veel problemen ouder worden. Deze laatste zijn blijkbaar

robuuster, mogelijk ook omdat ze in een gunstigere omgeving functioneren. Pas na tien jaar beginnen ook deze overlevers steeds meer mankementen te vertonen.

4. In welke mate veranderen de verschillen tussen loggermetingen en daaraan gepaarde controlehandmetingen door de correcties die bij validatie worden uitgevoerd?

De kenmerken van de meetfout van een logger ná validatie en zonodige correctie zijn afgeleid uit de verschillen tussen gevalideerde loggermetingen (L1) en de daaraan gepaarde controlehandmetingen (CH0). Het gemiddelde verschil van L1 en CH0 blijkt 0,01 cm en de standaardafwijking van het verschil blijkt 0,94 cm. Na verdiscontering van de meetfout van CH0 wordt het gemiddelde van de meetfout van de logger ná validatie en zonodige correctie geraamd op 0,51 cm, zodat er blijkbaar ook na validatie nog sprake is van enige overschatting. De standaardafwijking van de meetfout ná validatie en zonodige correctie wordt geraamd op minstens 0,42 cm. Het is niet mogelijk deze laatste nauwkeuriger te herleiden dan deze ondergrens, doordat de meetfout van L1 door de validatie (met zonodig correctie) gecorreleerd is geraakt aan de meetfout van CH0, zodat we ook de covariantie van de meetfouten van L1 en CH0 nodig hebben bij de berekening. Die covariantie is echter niet bekend.

Om de standaardafwijking van de meetfout van gevalideerde loggermetingen (L1) te kunnen vaststellen, zal eerst een steekproef moeten worden verzameld van handmetingen die niet worden gebruikt om loggermetingen te valideren. Zie daarvoor het antwoord bij vraag 5.

5. Wat is een geschikte steekproefopzet waarmee de kenmerken van de meetfout empirisch kunnen worden bepaald?

Om de standaardafwijking van de meetfout van gevalideerde loggermetingen zuiver te kunnen vaststellen dient eerst een steekproef te worden verzameld van handmetingen die niet worden gebruikt om loggermetingen te valideren. Deze handmetingen dienen dan te worden vergeleken met de daaraan gepaard te achten gevalideerde loggermetingen en de standaardafwijking van de resulterende set verschilwaarden dient vervolgens te worden gecorrigeerd voor de standaardafwijking van de meetfout van handmetingen.
De steekproefgrootte dient zodanig te zijn dat de standaardafwijking van de verschilwaarden voldoende nauwkeurig kan worden geraamd. Om een goede onderbouwing mogelijk te kunnen maken van die keuze (te doen in samenspraak met de Provincie), hebben we de relatie afgeleid tussen de steekproefgrootte (en daarmee ook de kosten) en de uit de steekproef volgende precisie van de standaardafwijking van het verschil.
Inmiddels is (bij de bespreking van het concept van dit rapport) gekozen voor een steekproef van 200 handmetingen van de grondwaterstand, uit te voeren gespreid over 2019.

1 Inleiding

1.1 Achtergronden en vraagstelling

Provincie Overijssel beheert een meetnet van peilbuizen om de toestand en ontwikkeling van de grondwaterstand en stijghoogte[1] in de provincie te kunnen bepalen. Oorspronkelijk werden de metingen handmatig uitgevoerd met een peilklokje aan een meetlint, op halfmaandelijkse basis (op of rond de 14e en de 28e). Maar vanaf eind 2003 is het meetnet geleidelijk voorzien van automatische drukopnemers (in het nu volgende meestal aangeduid als *loggers*), die de grondwaterstand eens per dag of eens per uur meten. Deze automatisering heeft als voordelen dat er mankracht wordt bespaard en dat er makkelijker hoogfrequent kan worden gemeten. Maar een nadeel is dat het meten complexer en daarmee ook foutgevoeliger is geworden. Er is immers geen sprake meer van een enkelvoudige handmeting van de grondwaterstand met een peilklokje aan een meetlint, maar van het simultaan met twee drukopnemers meten van de luchtdruk en het totaal van luchtdruk en waterdruk. Uit het verschil van deze drukmetingen wordt dan door de logger de grondwaterstand ten opzichte van een bepaald referentieniveau afgeleid.

Een probleem bij het automatisch meten van de grondwaterstand met een drukopnemer is dat de nauwkeurigheid van dat instrument kan veranderen. De praktijk heeft dan ook geleerd dat zijn metingen regelmatig gevalideerd[2] en daarbij zonodig gecorrigeerd moeten worden, om te voorkomen dat de meetfouten te groot worden (zie bijvoorbeeld [Von Asmuth, 2010, Post and Von Asmuth, 2013, Pleijter et al., 2015 en Von Asmuth en Van Geer, 2015]). De Provincie Overijssel hanteert al vrijwel vanaf het begin een procedure, waarbij de automatische metingen van een logger regelmatig (doorgaans circa halfjaarlijks) worden gevalideerd op basis van vergelijking met een controlehandmeting, indien nodig (bij relevante afwijking) gevolgd door correctie. Tot 2010 werd die procedure alleen ondersteund door de loggersoftware en daarna aanvullend ook door de module HydroData van het programma Menyanthes [Von Asmuth et al., 2010].

Vraagstelling

Op verzoek van de Provincie[3] richt onze studie zich op het in beeld brengen van de kenmerken van de meetfout bij automatisch meten van de grondwaterstand volgens de procedure van de Provincie Overijssel, waarbij antwoorden worden geformuleerd op de onderstaande vragen.

1. Wat zijn de langs theoretische weg afgeleide kenmerken van de meetfout en welke procescomponenten bepalen die kenmerken in belangrijke mate?
2. Hoe groot zijn de verschillen tussen ruwe (ongevalideerde) loggermetingen en daaraan gepaarde controlehandmetingen?

---

[1] Voor de leesbaarheid zullen we in het nu volgende verder geen onderscheid maken tussen *grondwaterstand* en *stijghoogte* en beide aanduiden als *grondwaterstand*.

[22] Onder *valideren* van metingen wordt in strikte zin verstaan het controleren of ze voldoen aan bepaalde eisen (zoals nauwkeurigheid). Maar omdat in het meetproces van de Provincie Overijssel het valideren van metingen bij gebleken relevante onnauwkeurigheid gepaard gaat met corrigeren, zullen we onder *gevalideerde metingen* impliciet ook verstaan *gevalideerde en zonodig gecorrigeerde metingen*. Het Protocol voor Datakwaliteitscontrole [Von Asmuth, 2018] stelt overigens *controleren en/of corrigeren* voor als alternatief voor *valideren*.

[3] Hier vertegenwoordigd door ir. Thomas de Meij, beleidsmedewerker geohydrologie van de Provincie Overijssel.

3. Zijn de verschillen tussen ruwe (ongevalideerde) loggermetingen en daaraan gepaarde controle-handmetingen gerelateerd aan factoren zoals het seizoen, of de hydrologische condities van de meetlocatie?
4. In welke mate veranderen de verschillen tussen loggermetingen en daaraan gepaarde controle-handmetingen door de correcties die bij validatie worden uitgevoerd?
5. Wat is een geschikte steekproefopzet waarmee de kenmerken van de meetfout empirisch kunnen worden bepaald?

1.2 Ontvangen basisinformatie

Voor deze studie hebben we de volgende basisinformatie van de Provincie Overijssel ontvangen:

1. Een aantal beschrijvingen van de procedure die de Provincie hanteert bij het automatisch meten van de grondwaterstand, inclusief de werkwijze bij het verzamelen van gegevens in het veld en bij validatie en correctie van metingen;
2. de grondwaterstandreeksen die volgens de meetprocedure van de Provincie Overijssel zijn verkregen met behulp van een logger, zowel vóór validatie (en eventuele aansluitende correctie), als daarna;
3. de metadata van de betreffende peilfilters en loggers, zoals coördinaten, filterstelling, maaiveldhoogte en meetpunthoogte van het peilfilter en druksensordiepte van de logger, evenals datum en tijdstip waarop vermelde metadata ingaan;
4. de bijbehorende controleloggermetingen en controlehandmetingen;
5. gegevens over datum en reden van loggervervangingen.

1.3 Over dit rapport

Na deze inleiding beschrijft hoofdstuk 2 een theoretische afleiding van kenmerken van de meetfout van de door de Provincie Overijssel gebruikte logger. En in hoofdstuk 3 leiden we die kenmerken empirisch af, aan de hand van een grote verzameling gepaarde logger- en handmetingen. Hoofdstuk 4 bevat analyses van de relaties van kenmerken van de meetfout van loggers met factoren, zoals leeftijd, maand, waterhoogte ten opzichte van druksensor en enkele hydrologische karakteristieken. Hoofdstuk 5 beschrijft de opzet van een steekproef naar de kenmerken van de meetfout van de door de Provincie Overijssel gebruikte logger. Hoofdstuk 6 geeft een aantal aanbevelingen voor de meetprocedure van de Provincie, gebaseerd op de bevindingen van onze studie. Dit rapport sluit af met de alfabetische lijst van de aangehaalde literatuur.

Het rapport heeft één bijlage. Deze toont de trendplots van de verschilwaardereeksen van controlemetingen van díe loggers waar deze reeks een statistisch significante trend vertoont (duidend op drift).

1.4 Studiebegeleiding

Onze speciale dank gaat uit naar de volgende personen:

- ir. Thomas de Meij, beleidsmedewerker geohydrologie van de Provincie Overijssel;
- ing. Joost Gooijer, onderzoeker waterhuishouding / medewerker grondwatermeetnet van de Provincie Overijssel;

- dr. ir. Jos von Asmuth, onderzoeker en contentmanager Menyanthes – Ecohydrologie van KWR Watercycle Research Institute.

Thomas en Joost waren onze eerste aanspreekpunten en vraagbakens bij de uitvoering van de studie en ze hebben samen met Jos het conceptrapport becommentarieerd. Dat laatste zal geen sinecure zijn geweest, gezien de omvang en vele technische details van het rapport.
We hebben ons best gedaan hun commentaren en suggesties zo goed mogelijk te verwerken. Niet alles kon echter worden meegenomen, ondermeer waar het teveel extra werk zou opleveren en waar het smaakkwesties, niet te objectiveren keuzen of verschillen van mening betrof. De verantwoordelijkheid voor eventuele resterende fouten en onvolkomenheden rust dan ook volledig bij ons en niet bij onze begeleiders.

Wij hopen dat dit rapport een bruikbare bijdrage kan zijn in het streven naar voldoende kwaliteitsborging bij het meten van de grondwaterstand met automatische drukopnemers.

2 Theoretische afleiding kenmerken meetfout

Op hoofdlijnen zijn er twee soorten aanpak om de kenmerken van de meetfout van een bepaalde meetprocedure af te leiden, namelijk [ISO, 1993]: i) de theoretische aanpak, gebaseerd op de foutenleer en ii) de empirische aanpak, gebaseerd op statistische analyse van een verzameling meetwaarden. In deze studie worden beide soorten aanpak toegepast om de kenmerken af te leiden van de meetfout bij het automatisch meten van de grondwaterstand door de Provincie Overijssel. Dit hoofdstuk richt zich op de theoretische aanpak, waarbij we starten met formele definities van relevante kengetallen van de meetfout. In hoofdstuk 3 bewandelen we de empirische weg.

2.1 Definities relevante kengetallen van de meetfout

Stel dat we de meting van een grondwaterstand aanduiden als *x* en de werkelijke grondwaterstand als τ (*true value*), dan zal gelden:

$$x = \tau + e \qquad [2.1]$$

waarin *e* de meetfout (*error*).

Als we beschikken over meerdere (onafhankelijke) metingen van één of meer grondwaterstanden, waarvan we ook de werkelijke waarden kennen, dan kunnen we relevante kengetallen van de meetfout vaststellen, zoals (zie ook figuur 2.1):

- de gemiddelde meetfout ($\bar{e}$), een schatting van μ, het gemiddelde van de meetfoutpopulatie, volgens:

  $$\bar{e} = \frac{\sum_{i=1}^{n} e_i}{n} \qquad [2.2]$$

  waarin e_i de meetfout van meting *i* en *n* het aantal metingen;

- de standaardafwijking van de meetfout (s_e) Dit is een schatting van σ, de standaardafwijking van de meetfoutpopulatie, volgens[4]:

  $$s_e = \sqrt{\frac{\sum_{i=1}^{n}(e_i - \bar{e})^2}{n-1}} \qquad [2.3]$$

  Dit kengetal is geschikt om de precisie (of ook wel herhaalbaarheid) van het meetproces weer te geven, zijnde de mate van spreiding van de meetfout. Hoe kleiner de standaardafwijking van de meetfout, hoe groter de precisie;

- de systematische meetfout (b_e)[5], dit is een schatting van β_e, de systematische afwijking van de meetfoutpopulatie, volgens:

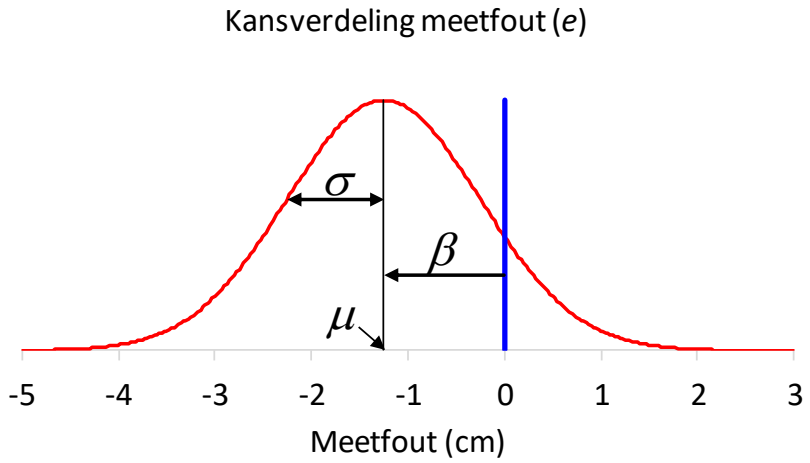
  $$b_e = \bar{e} = \bar{x} - \bar{\tau} \qquad [2.4]$$

  waarin $\bar{x}$ het gemiddelde van de metingen en $\bar{\tau}$ het gemiddelde van de werkelijke waarden van de grondwaterstanden. Dit kengetal kan de juistheid van het meetproces weergeven, zijnde de mate van overeenstemming tussen de (gemiddelde) meting en de werkelijke waarde.

[4] Dit is de gebruikelijke schatter van de standaardafwijking, aangezien deze zuiver is (dus zonder systematische fout) en het ook de meest precieze is van de zuivere schatters van de standaardafwijking.

[5] De *b* staat hier voor *bias*, de Engelse term voor systematische fout.

Figuur 2.1: Relevante kengetallen van de kansverdeling van de meetfout.

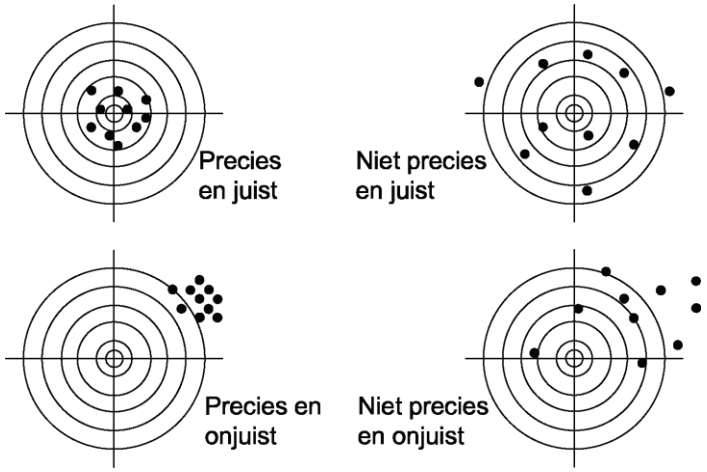


μ: gemiddelde, σ: standaardafwijking, β: systematische fout

> Meer over de term systematische fout
> De term systematische fout wordt - behalve om een belangrijk kengetal van de meetfout aan te duiden (zoals boven) – ook gebruikt om het sóórt meetfout aan te duiden. Met een systematische fout wordt dan bedoeld een afwijking ten opzichte van de werkelijke waarde die een systematische oorzaak heeft. Dit ter onderscheid van de toevallige fout. Als er alleen sprake is van een toevallige fout, zal bij herhaald meten zijn gemiddelde naar nul gaan naderen, terwijl dit bij een systematische fout niet het geval zal zijn. Bij meten met een logger kunnen systematische fouten optreden door effecten van ondermeer temperatuur, niet-lineariteit, hysterese en veroudering. Veroudering kan daarbij aanleiding geven tot een al dan niet geleidelijk opbouwende systematische fout, ook wel aangeduid als drift.

Ten slotte is er nog de nauwkeurigheid van het meetproces, die wordt bepaald door de combinatie van juistheid en precisie (zie figuur 2.2).[6]

Figuur 2.2: Samenhang tussen de begrippen nauwkeurigheid, precisie en juistheid. Linksboven is de nauwkeurigheid het grootst en rechtsonder het kleinst.



[6] In de statistische en metrologische literatuur zijn ook andere definities van *nauwkeurigheid* (Engels: *accuracy*) te vinden. Het is daarom zaak altijd eenduidig in een rapport te definiëren welke gehanteerd wordt.

Het begrip *meetonzekerheid*

In de metrologie (de leer van het meten) wordt voor het beschrijven van meetfouten ook een enigszins aangepaste terminologie gehanteerd, waarbij het begrip *meetonzekerheid* (*measurement uncertainty*) centraal staat. Volgens de formele metrologische definitie is de meetonzekerheid een kengetal voor de spreiding van meetwaarden die redelijkerwijs kunnen worden toegewezen aan dat wat gemeten is (zie bijvoorbeeld [ISO, 1993]).

De kwantificering van de meetonzekerheid kent twee (gerelateerde) uitdrukkingsvormen:

1. de standaardonzekerheid $u(x)$, bepaald als de standaardafwijking van meetwaarden die redelijkerwijs kunnen worden toegewezen aan dat wat gemeten wordt. Als er geen sprake is van een systematische fout stemt deze overeen met de boven gedefinieerde standaardafwijking van de meetfout (s_e), anders moet daar voor worden gecorrigeerd of op één of andere wijze rekening mee worden gehouden;
2. de uitgebreide onzekerheid $U(x)$, gedefinieerd als de helft van het interval dat met een grote betrouwbaarheid de werkelijke waarde zal omvatten van dat wat gemeten wordt. Deze volgt uit:

$$U(x) = k \cdot u(x) \qquad [2.5]$$

waarin k de dekkingsfactor. Doorgaans wordt k op 2 gesteld. Als de meetwaarden afkomstig zijn uit een normale kansverdeling mag worden aangenomen dat het interval $x \pm U(x)$ met circa 95% betrouwbaarheid de werkelijke waarde zal omvatten.

We zullen de standaardonzekerheid hanteren bij het theoretisch afleiden van kenmerken van de meetfout van de door de Provincie gehanteerde logger (de Keller DCX-22 AA), omdat dit de formules enigszins compacter maakt (zie § 2.5.4).

2.2 Onderdelen en foutenbronnen meetprocedure Provincie Overijssel

De theoretische afleiding van de kenmerken van de meetfout van een bepaalde meetprocedure omvat de volgende stappen:

1. het identificeren van alle onderdelen van de procedure;
2. het identificeren van de foutenbronnen van deze onderdelen;
3. het afleiden, herleiden (uit productspecificaties en/of calibratierapportages) of ramen van hun afzonderlijke bijdragen aan de meetfout en
4. het bepalen van de kenmerken van de meetfout, op basis van de theorie van de foutenvoortplanting.

De theoretische afleiding van de meetfout is doorgaans zeker niet triviaal en kan ook veel inspanning vergen, ondermeer door het raadplegen van technische documentaties, ook al zijn er ISO-normen die hiervoor enkele richtingen aangeven. Het probleem is echter dat we pas kunnen bepalen waar een invloed kan worden weggelaten, als we weten dat dit de theoretisch bepaalde kenmerken van de meetfout niet relevant vertekent.

2.2.1 Onderdelen meetprocedure Provincie Overijssel

Het door de Provincie Overijssel beheerde grondwatermeetnet omvat op dit moment van schrijven (12 juni 2018) 447 peilbuizen, met totaal 662 peilfilters waarin uurlijks de grondwaterstand wordt gemeten met een logger. De monitoring is gericht op het in beeld brengen van: i) de toestand en de

ruimtelijke en temporele veranderingen van het regionale grondwatersysteem en ii) de standplaatscondities in grondwaterafhankelijke natuurgebieden.

Een groot deel van het meetnet wordt al meerdere decennia bemeten. Aanvankelijk gebeurde dat halfmaandelijks (op of rond de 14e en de 28e), op basis van handmetingen. Van eind 2003 tot eind 2004 zijn de peilfilters door de Provincie voorzien van loggers en afgewerkt in een afsluitbare straatpot of schutkoker, waarna de meeste tot eind 2014 / begin 2015 dagelijks om 08:00 uur 's ochtends automatisch zijn bemeten. Op dat tijdstip zal het effect van verdamping namelijk minimaal zijn en is er doorgaans ook minder effect te verwachten van kleine grondwateronttrekkingen. Vanaf eind 2014 / begin 2015 worden alle peilfilters uurlijks bemeten.[7] Dit gold al vanaf 2012 voor de peilfilters in natuurgebieden (van het verdrogingsmeetnet), omdat de grondwatersystemen daar doorgaans meer dynamiek vertonen door de geringe dikte van de onverzadigde zone en het beperkte systeemgeheugen.

De Provincie volgt al vrijwel vanaf het begin een procedure om de kwaliteit van de drukopnemermetingen te borgen, gericht op het vermijden van fouten in beschikbaar gestelde grondwaterstandreeksen en op het minimaliseren van de noodzaak voor validatie en correctie achteraf. Daarbij worden loggermetingen, controlemetingen en correcties ook systematisch vastgelegd, zodat de gebruiker desgewenst uit de oorspronkelijke meetwaarden zelf een gevalideerde grondwaterstandreeks kan genereren. We ontvingen gedurende het uitvoeringstraject van deze studie drie beschrijvingen van deze procedure, inclusief de werkwijze bij het verzamelen van gegevens in het veld en bij validatie en correctie van metingen [Gooijer en De Meij, 2013, 2018a en 2018b]. Alle drie betreffen echter alleen de procedure die is gehanteerd vanaf 2010, toen - naast de reeds vanaf begin 2003 gehanteerde Keller-software - ook de HydroData-module van het programma Menyanthes aan de procedure is toegevoegd. De drie beschrijvingen verschillen op bepaalde details, wat aangeeft dat er gaandeweg ook veranderingen in de procedure zijn aangebracht.
De hoofdpunten van de procedure zoals gehanteerd vanaf 2010 zijn hieronder samengevat. Soms beschouwen we ook verschillen ten opzichte van de voorgaande procedure, toen er alleen nog Keller-software werd gebruikt.

Beschrijving procedure zoals gehanteerd vanaf 2010

1. Plaatsen van de logger, met de volgende vuistregels en handelingen:
   a. ga standaard uit van een logger met een meetbereik van 5 meter waterkolom, mits de amplitude van de grondwaterstand niet meer bedraagt dan vier meter. Ga anders uit van een logger met een meetbereik van 10 meter waterkolom;
   b. bepaal de gewenste kabellengte van de logger, op basis van de filterstelling en de in dat filter gemeten hoogste en laagste grondwaterstand. Ga standaard uit van een kabellengte van 4 meter, tenzij de laagste grondwaterstand meer dan ongeveer 3 meter onder de bovenkant van de peilbuis komt, in welk geval een logger met langere kabel moet worden gebruikt. Kies dan de kabellengte zodanig dat de druksensor op ongeveer 2,5 meter (de helft van het meetbereik van 5 meter) onder de gemiddeld optredende waterstand komt te hangen;
   c. voor ondiepere locaties worden bij de start kabellengtes op maat besteld, zodanig dat de loggerdiepte circa 10 cm boven de onderkant van het filter is;

[7] Gedurende de halfjaarlijkse uitleesronde die begon op 1 oktober 2014 zijn de loggers één voor één ingesteld op één keer per uur loggen.

d. bij vervanging wordt er gewerkt met standaard kabellengtes (1, 1,5, 2, 2,5, 3, 3,5 en 4 m). De kabel wordt dan zonodig aangepast aan de plaatselijke omstandigheden, door 'oplussen';
e. leg schriftelijk en fotografisch de topografische en technische kenmerken van het meetpunt vast. Relevante technische kenmerken zijn de coördinaten (in het Rijksdriehoekstelsel), de filterdiepten, meetpunthoogte en maaiveldhoogte;
f. bevestig een identificatieplaatje met peilbuiscode op de schutkoker van de peilbuis.

2. Ongeveer halfjaarlijks uitlezen van de logger. De beheerder controleert daarbij eerst of de batterij nog voldoende is geladen. En vervolgens neemt deze een handmeting om te kunnen controleren of de logger de grondwaterstand correct registreert. De handmeting en de daaraan gepaarde online-loggermeting worden ingevoerd in Menyanthes. Daarna wordt de uitgelezen reeks geïmporteerd in Menyanthes. Met behulp van de Viewer van Menyanthes worden eventuele driftcorrecties toegepast. Volgens de recentste beschrijving van de meetprocedure [Gooijer en De Meij, 2018b) zijn de details van deze handelingen:
   a. als het geen diep peilfilter betreft, verricht dan eerst een handmeting (in mm ten opzichte van het meetpunt), die dient ter controle van de nauwkeurigheid van de logger en sla deze controlemeting op in de HydroData-module van Menyanthes (bij 'Handmetingen'). Betreft het wél een diep peilfilter, verricht deze handmeting dan pas ná de online-loggermeting (na handeling d. hieronder);
   b. verbind de logger met een laptop die het hele jaar ingesteld is op wintertijd[8] en waarop zowel Keller-software als de module HydroData van Menyanthes is geïnstalleerd;
   c. stel met de Keller-software de logger in op de starttijd van de volgende meetreeks (vanaf eind 2014 / begin 2015 is dat doorgaans het volgende hele uur);
   d. lees met behulp van de 'Online knop' van de Keller-software de online-loggermeting van de grondwaterstand (hierna aangeduid als controleloggermeting) en sla deze op in de HydroData-module van Menyanthes (bij 'Handmetingen');
   e. lees het geheugen van de logger uit met de Keller-Reader ('lees laatste record') en importeer dit in de HydroData-module van Menyanthes. De Keller-software stelt deze data beschikbaar als een binair IDC-bestand, dat de metingen van verschillende kanalen kan bevatten, zoals drukverschil, totale druk, luchtdruk, watertemperatuur en luchttemperatuur[9];
   f. bekijk deze gegevens in de Viewer van de HydroData-module van Menyanthes, waarbij onder 'Drift' het verschil wordt getoond tussen de controlehandmeting en de laatste loggermeting, ook voor de voorgaande drie à vier uitleesmomenten (voorzover beschikbaar);
   g. betreft het de eerste controle- en inleesronde na het installeren van de logger, corrigeer dan de uitgelezen loggermetingen voor het verschil tussen loggermeting en handmeting door de instelling van de loggerdiepte in Menyanthes aan te passen, op basis van een visuele beoordeling middels de Viewer van Menyanthes (dit komt neer op het verticaal ophogen of verlagen van de gehele ingelezen reeks)[10];

[8] Als de laptop altijd is ingesteld op wintertijd zullen er geen discrepanties ontstaan door het verschil tussen zomertijd en wintertijd. Bij het met de Keller-Writer wegschrijven van de nieuwe starttijd naar het loggergeheugen wordt de loggertijd gesynchroniseerd met de laptoptijd.

[9] De beheerder kan bepalen welke gegevens deze kanalen weergeven. Vanaf circa 2010 worden de totale druk en de luchtdruk in de IDC-bestanden geregistreerd, via de kanalen P1 en P2.

[10] Tot de toevoeging van Menyanthes aan de procedure (2010) geschiedde dit door aanpassen van de loggerdiepte in de Keller-software (daar aangeduid als *installatiediepte*).

h. betreft het niet de eerste controle- en inleesronde na het installeren van de logger, onderneem dan zonodig actie op basis van verschillen tussen controlehandmetingen en loggermetingen. Daarbij worden volgens de recentste beschrijving van de meetprocedure [Gooijer en De Meij, 2018b] de volgende richtlijnen gehanteerd:
   - pas driftcorrecties toe om de loggerreeks zo goed mogelijk te laten aansluiten op de controlehandmetingen en ook zodanig dat er ongeveer evenveel positieve als negatieve afwijkingen zijn;
   - tracht de grootte van de driftcorrectie van de net ingelezen reeks beperkt te houden, omdat er nog geen grondwaterstand van de daarop aansluitende nieuwe reeks beschikbaar is. Deze restrictie geldt niet ten aanzien van de afwijkingen van de voorgaande drie à vier controlehandmetingen en de loggermetingen;
   - zorg bij voorkeur dat de reeks driftcorrecties een min of meer egale lijn vormen, bestaande uit stapjes van hooguit 5 mm. Zijn grotere stappen nodig, dan verdient het overweging lineaire correctie toe te passen;
   - vervang de logger als de controleafwijking groter dan 5 cm is ten opzichte van de vorige controleafwijking, óf als de logger tien jaar of ouder is én de afwijking groter dan 1,5 cm is ten opzichte van die van de vorige uitleesronde. Tracht hierbij visueel te boordelen vanaf wanneer de logger is gaan verlopen en verwijder de metingen vanaf dat tijdstip. Verwissel de logger en registreer het nieuwe loggernummer en ingangsdatum- en tijdstip in de HydroData-module van Menyanthes.

   Volgens de oudere beschrijvingen van de meetprocedure zijn er voorheen enigszins andere richtlijnen gehanteerd voor actie op basis van verschillen tussen controlehandmetingen en loggermetingen. Zo werd volgens [Gooijer en De Meij, 2013] bij een controleafwijking tussen 0,5 en 2 cm de helft daarvan gecorrigeerd en bij een controleafwijking tussen 2 en 4 cm een correctie ter grootte van de controleafwijking toegepast. De logger werd vervangen als de controleafwijking groter dan 4 cm was;
i. werk drie Excelbestanden bij:
   - geef in bestand *Route* aan dat de betreffende locatie is bezocht en eventueel wat van belang is voor het volgende bezoek;
   - geef in bestand *Dataverwerking* aan wát is uitgelezen en wat eventueel op kantoor nog nader moet worden gecontroleerd op betrouwbaarheid;
   - houd in bestand *Primair* bij welke locaties er zijn, wanneer deze zijn onderhouden, wanneer welke logger is geplaatst / vervangen en wanneer er is gewaterpast. Een deel van deze informatie staat ook in Menyanthes.

3. Om de zoveel jaar onderhouden van de meetpunten:
   a. voer elke tien jaar een waterpassing uit van de meetpuntshoogte van een peilbuis. Daartoe wordt jaarlijks 10% van de meetpunten door een externe partij gewaterpast;
   b. maak eens in de vijf jaar of tien jaar peilbuizen schoon:
      - eens in de vijf jaar worden landbouwbuizen (deze zijn niet dieper dan 5 meter) tijdens een reguliere controleronde schoongemaakt;
      - kleine reparaties worden direct uitgevoerd door de meetnetbeheerder;
      - verplaatsingen van landbouwbuizen (vaak op verzoek van landeigenaar) worden op korte termijn in eigen beheer uitgevoerd;
      - eens in de tien jaar worden diepere peilbuizen (dieper dan 5 meter) door een extern bedrijf doorgepompt en gecontroleerd;

- verplaatsingen van diepe peilbuizen worden door een boorbedrijf uitgevoerd.

4. Bijhouden logboek bij elk veldbezoek. Tot 2015 werden veel belangrijke gegevens analoog in het veld bijgehouden met een logboek. Maar vanaf 2015 worden deze (en ook andere) digitaal bijgehouden in het veld 'Opmerkingen' van de metadata in Menyanthes. Het betreft de volgende gegevens:
   - datum en tijd (automatisch);
   - loggernummer en kabellengte;
   - loggerwisseling;
   - handmeting en actuele loggermeting van de grondwaterstand;
   - uitgelezen reeksen;
   - driftcorrecties;
   - identificatie van de reeksen die als tuf-bestand zijn geëxporteerd naar TNO;
   - uitgevoerde waterpassingen;
   - uitgevoerde schoonmaakacties;
   - uitgevoerde reparaties.

Op basis van de boven beschreven meetprocedure komen in eerste instantie als mogelijke foutenbronnen naar voren: i) de procedure-ondersteunende software, ii) het corrigeren van ruwe loggermetingen en iii) de logger. Ook het waterpassen van de meetpuntshoogte zal een foutenbron van de meetprocedure zijn, ook al wordt dit weinig uitgevoerd (eens per 10 jaar). Maar deze zal hier verder niet worden beoordeeld. Belangrijke reden is dat we de studie moesten afbakenen, waarbij ook meespeelde dat we de waterpasfout niet empirisch kunnen evalueren met de ons ter beschikking staande meetgegevens.

Hieronder beoordelen we eerst of de procedure-ondersteunende softwarepakketten (Keller-software, vanaf 2010 aangevuld met de HydroData-module van Menyanthes) onvolkomenheden vertonen (§ 2.3). Daarna beoordelen we de mogelijke invloed van het corrigeren van ruwe loggermetingen op de meetfout (§ 2.4). Tenslotte leiden we in § 2.5 langs theoretische weg de kenmerken af van de meetfout van de door de Provincie Overijssel gehanteerde logger, de Keller DCX-22 AA. En in hoofdstuk 3 doen we dat laatste langs empirische weg.

2.3 Mogelijke foutenbronnen procedure-ondersteunende software

We zijn op zoek gegaan naar mogelijk door de software veroorzaakte fouten, niet alleen voor de huidige software-ondersteuning, maar ook voor de voorgaande. Sinds de start van de meetprocedure met loggers (2003) tot 2010 was er namelijk alleen nog ondersteuning met Keller-software, terwijl vanaf 2010 een deel van de ondersteuning is overgenomen door de Hydrodata-module van Menyanthes.

Bij de speurtocht is mede afgegaan op een verkenning van de volgende bestanden (van de Provincie Overijssel en KWR[11]):

[11] Verstrekt door Thomas de Meij en Joost Gooijer, beiden Provincie Overijssel en door Erwin Vonk, KWR.

- een binair uitvoerbestand van de Hydrodata-module van Menyanthes (aangeduid als Men-bestand). Dit omvat metagegevens van peilfilters en loggers, ruwe loggermetingen, handmetingen en gegevens om van de ruwe loggergegevens gevalideerde loggergegevens te maken (het bestand bevat zelf geen gevalideerde loggergegevens).[12] De ruwe loggermetingen zijn op dagelijkse of uurlijkse basis vanaf 2003 of later. Ze zijn afkomstig van 1.165 loggers, die over de periode 2003 tot februari 2018 functioneerden in 549 peilfilters van 379 peilbuizen. Het betrof vrijwel alleen Keller-loggers van het type DCX 22-AA[13];
- een csv-bestand met de gevalideerde uitvoeringen van de reeksen (een uitvoerbestand van Menyanthes). Al deze reeksen zijn na 2010 gevalideerd met de HydroData-module van Menyanthes (ook de delen van vóór 2010);
- een csv-bestand met ruwe loggermetingen (een uitvoerbestand van Menyanthes);
- een csv-bestand met controlemetingen (een uitvoerbestand van Menyanthes);
- IDC-bestanden van de ruwe loggerreeksen. Dit zijn binaire uitvoerbestanden van de Keller-logger die bij het uitlezen worden aangemaakt en naar keuze meetinformatie kunnen vermelden van verschillende meetkanalen van de logger (totale druk, luchtdruk, watertemperatuur en luchttemperatuur) over de voorgaande registratieperiode.

Onze speurtocht naar mogelijk door de software veroorzaakte fouten werd bemoeilijkt, doordat de beschrijvingen van de meetprocedure [Gooijer en De Meij, 2013, 2018a en 2018b] weinig details bevatten van de specifieke rollen van de twee ondersteunende softwarepakketten en van hun interacties en ook doordat er tijdens onze uitwerkingen geen detailbeschrijving beschikbaar was van de interne werking, gegevensstructuur en semantiek van de HydroData-module van Menyanthes.

2.3.1 Schets van ideale combinatie van software en procedure

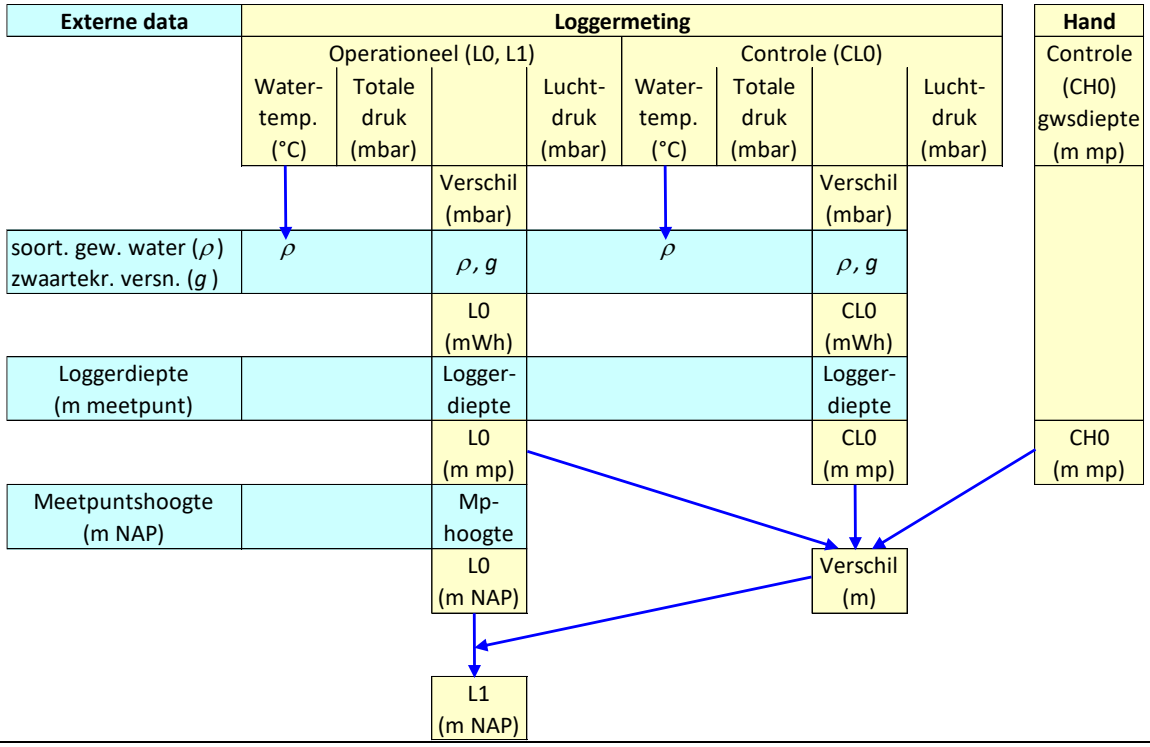
Om onvolkomenheden in de combinatie van procedure en software te kunnen identificeren schetsen we eerst het beeld van de ideale situatie, met behulp van onderstaande figuur 2.3.

Kernpunten van bovenstaande ideale situatie zijn de volgende:

1. Elke verschildruk van de twee druksensoren wordt correct omgerekend naar waterhoogte ten opzichte van sensordiepte, door middel van de van toepassing zijnde waarden van de daarvoor relevante fysische constanten (soortelijk gewicht van water ρ en zwaartekrachtversnelling g).
2. Elke loggermeting wordt correct omgerekend naar een ander referentieniveau, met de dan geldende loggerdiepte (m meetpunt) en meetpuntshoogte (m NAP).
3. Elke reeks met ruwe loggermetingen (L0) wordt bij validatie (naar de reeks L1) zonodig gecorrigeerd) voor een bij de controle geconstateerd relevant verschil met een controlehandmeting (CH0). Het corrigeren geschiedt door toevoeging van drift, een lineaire helling, of door handmatig verwijderen van waarden.

[12] Pas bij de bespreking van de verwerkingsresultaten kwam naar voren dat het verstrekte bestand abusievelijk ook gegevens bevatte van Vitens. Deze laatste hebben we toen verwijderd en de verwerking overgedaan.
[13] In de metadata staat het type logger meestal aangeduid als *unknown*. Volgens de Provincie zijn er slechts vijf Keller-loggers van het type DCX-22 A.

Figuur 2.3: Schets van ideale combinatie van software en procedure. De procesrichting is van boven naar beneden. De externe (blauwe) data is nodig voor omrekening naar een andere grootheid of een ander referentieniveau.



Toelichting
L0: Ruwe operationele loggermeting
mWh: meter waterhoogte (ten opzichte van loggerdiepte)
CL0: Controleloggermeting, gedaan alvorens L0-reeks uit te lezen (doorgaans na half jaar meten)
CH0: Controlehandmeting, gedaan vrijwel simultaan met CL0
L1: Operationele loggermeting na validatie. Deze volgt uit L0, na eventuele correctie op basis van verschil CL0 en/of laatste L0 met CH0. Maar als dat verschil te groot is, wordt de logger vervangen en wordt er geen of slechts een beperkte vorm van L1-reeks gemaakt.

We zijn voor elk van deze kernpunten nagegaan of er onvolkomenheden zijn van de combinatie van procedure en software, zonodig met onderscheid tussen de periode tot 2010 (met alleen Keller-software) en de periode vanaf 2010 (met toevoeging van de HydroData-module van Menyanthes). Onze bevindingen zijn hieronder beschreven.

2.3.2 Omrekenen verschildruk naar waterhoogte

Om de door de logger gemeten verschildruk van de twee druksensoren te kunnen omrekenen naar waterhoogte ten opzichte van de loggerdiepte zijn twee fysische constanten nodig (zie formule [2.8] in § 2.5), namelijk: i) het soortelijk gewicht van water (ρ) en ii) de zwaartekrachtsversnelling (g).

Het gehanteerde soortelijk gewicht moet daarbij in overeenstemming zijn met de watertemperatuur (en -samenstelling) en de gehanteerde zwaartekrachtsversnelling moet representatief zijn voor de regio (in dit geval Overijssel), anders ontstaat er een systematische fout. Maar voor zover wij hebben kunnen nagaan is in de periode tot 2010 vermoedelijk gebruik gemaakt van verschillende waarden voor het soortelijk gewicht van water en is gebruik gemaakt van de zwaartekrachtsversnelling representatief voor Zwitserland (zie § 2.5.2). Dit probleem speelt vanaf 2010 echter niet meer,

omdat toen alle meetreeksen (en met terugwerkende kracht ook de deelreeksen van vóór 2010) met de HydroData-module van Menyanthes met geschikte fysische constanten zijn omgerekend naar waterhoogten en vervolgens ook naar grondwaterstanden. De correctie van oudere deelreeksen was mogelijk doordat altijd basismetingen van de loggers zijn bewaard (in zogenaamde IDC-bestanden). Tot medio 2010 zijn ze bewaard als waterhoogten en daarná als drukken.[14]

Een probleem blijft nog wel bestaan bij de door de Keller-software uitgevoerde omrekening van het drukverschil van de controleloggermeting (CL0) naar waterhoogte, waarbij gebruik wordt gemaakt van de in die software aanwezige waarden van de fysische constanten. De zwaartekrachtsversnelling staat daarbij vast op de waarde die geldt voor Zwitserland, terwijl het soortelijk gewicht door de beheerder kan worden ingevuld. Dit zal de CL0-waarden minder vergelijkbaar maken met de ruwe loggermetingen (L0), die zijn omgerekend met de HydroData-module van Menyanthes. Die vergelijkbaarheid wordt overigens ook al verminderd door het regelmatig bijstellen van de installatiediepte in de Keller-software (zie ook verder in deze paragraaf).

2.3.3 Omrekenen loggermetingen naar ander referentieniveau

Bij het verwerken van ruwe controle- en loggermetingen met de HydroData-module van Menyanthes is de starttijd van loggerdiepte en meetpuntshoogte cruciaal om de basismetingen (in meter waterhoogte ten opzichte van loggerdiepte) correct om te kunnen rekenen naar meter ten opzichte van meetpunt en meter ten opzichte van NAP. Maar deze starttijd blijkt niet altijd juist te zijn geregistreerd. Tot 2010 zijn de startdatums van nieuwe metadata analoog door de beheerder bijgehouden, maar het tijdstip is daarbij vaak niet of niet juist ingevuld. En vanaf 2010 kon weliswaar de startdatum direct worden ingevoerd in Menyanthes, maar nog niet het tijdstip. Dit laatste kon vermoedelijk pas vanaf 2015 (zie verder), waarbij met terugwerkende kracht elke nog ontbrekende wijzigingstijd op 08:00 uur of 09:00 uur van de betreffende datum is gezet in de database van Menyanthes. Dit zijn ongelukkige keuzes geweest, want tenzij er sprake is van de initiële installatie van een logger (eerst geplaatste logger, dus geen loggerwisseling), zal het tijdstip van nieuwe metadata in principe altijd ná de uitgevoerde controlemetingen moeten zijn (uitzonderlijke situaties daargelaten), maar die controles zijn vrijwel nooit zo vroeg in de ochtend uitgevoerd. Onderstaande tabel 2.1 geeft een voorbeeld van het optreden van dit probleem.

Volgens onze analyse is er in 747 gevallen sprake van een starttijd van metadata die in de database van Menyanthes onterecht vóór de controletijd staat in plaats van daarná. Dit betreft 6,4% van alle controledatums. Het probleem trad vooral op tot 2015 (zie tabel 2.2).

[14] Persoonlijke mededeling van Joost Gooijer, per email van 9 november 2018.

Tabel 2.1: Historie van de metagegevens van peilfilter B16D0070-2, zoals beschikbaar in de database van Menyanthes. De eerste logger is geïnstalleerd op 20-3-2003 en daarna zijn er nog vier loggerwisselingen geweest. Voor een correcte omrekening van controle- en loggermetingen naar meter NAP zijn van belang de kolommen Startdatum, Maaiveld, Meetpunt en Loggerdiepte, maar in drie gevallen (oranje gekleurd) is het tijdstip van Startdatum te vroeg. De door ons gecorrigeerde tijdstippen zijn vermeld in de laatste kolom.

Controle-datum	Startdatum metadata	Maaiveld (m NAP)	Meetpunt (m NAP)	Logger		Gecorrigeerde startdatum metadata
				num-mer	diepte (m mp)	
	28-4-1986 00:00	-0,91	-0,99			
	7-10-1999 08:00	-0,91	-0,97			
20-11-2003 00:00	20-3-2003 08:00	-1,008	-1,069	1476	4,985	
18-8-2010 14:45	18-8-2010 09:00	-0,994	-1,052	10017	3,965	18-8-2010 14:46
24-7-2012 10:59	24-7-2012 08:00	-1,032	-1,11	10017	3,965	24-7-2012 11:00
29-8-2014 09:21	29-8-2014 08:00	-1,032	-1,11	15252	3,98	29-8-2014 09:22
18-11-2015 10:28	18-11-2015 10:31	-1,032	-1,11	16576	1,371	
27-9-2017 09:56	27-9-2017 10:30	-1,032	-1,11	539381	3,982	

De té vroege starttijd van metadata kan in theorie hebben geleid tot de volgende fouten:

- fouten bij het omrekenen van de controleloggermeting en de controlehandmeting (beide oorspronkelijk uitgedrukt ten opzichte van meetpunt) naar een ander referentieniveau. Deze fouten zullen optreden als de starttijd van de verandering van de meetpuntshoogte (m NAP) vóór de controletijd staat in plaats van daarná;
- fouten bij het omrekenen van loggermetingen (uitgedrukt in hoogte waterkolom ten opzichte van de druksensor) naar een ander referentieniveau. Deze fouten zullen optreden als de starttijd van de verandering van loggerdiepte (m ten opzichte van meetpunt) of meetpuntshoogte (m NAP) vóór de controletijd staat in plaats van daarná en zullen dan gelden voor de loggermetingen in de tussenliggende periode.

Het is echter niet waarschijnlijk dat deze fouten vaak zijn opgetreden. Voor wat betreft de eerste fout omdat er weinig sprake is van verandering van meetpuntshoogte en voor wat betreft de tweede fout omdat die vooral kan optreden bij meten op uurbasis, maar daar is pas algemeen sprake van vanaf 2015 en toen was het probleem met de starttijden al opgelost.

Tabel 2.2: Verdeling over de jaren en over de uren van de dag van de 747 gevallen dat de starttijd van metadata in de database onterecht vóór de controletijd staat in plaats van daarná. De kolom uur betreft de start van het uur waarbinnen de verkeerde starttijd valt. Voorbeeld: in 2010 vielen 48 van de verkeerde starttijden in het uur van 08:00:00 t/m 08:59:59.

Uur	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Totaal	%
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0,0%**
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0,0%**
2	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	**1**	**0,1%**
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0,0%**
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0,0%**
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0,0%**
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0,0%**
7	0	0	**4**	0	0	**5**	0	0	0	0	0	**1**	0	0	0	0	**10**	**1,3%**
8	0	**11**	**27**	**27**	**15**	**32**	**35**	**48**	**40**	**52**	**42**	**51**	**4**	0	0	0	**384**	**51,4%**
9	0	**10**	**64**	**17**	**21**	**35**	**28**	**26**	**23**	**27**	**18**	**27**	0	0	**1**	0	**297**	**39,8%**
10	0	0	**1**	**1**	0	0	**1**	0	**1**	**2**	**2**	**3**	**3**	**4**	**1**	0	**19**	**2,5%**
11	0	0	**2**	0	0	0	0	0	0	**1**	0	0	**1**	**1**	**1**	0	**6**	**0,8%**
12	0	**1**	0	0	0	0	**1**	0	**1**	0	0	**2**	0	**2**	**5**	0	**12**	**1,6%**
13	0	0	0	0	0	0	0	0	0	0	**1**	0	**2**	**1**	**2**	0	**6**	**0,8%**
14	0	0	0	0	0	0	0	0	0	0	**1**	**1**	**3**	**1**	0	0	**6**	**0,8%**
15	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	**1**	0	**2**	**0,3%**
16	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	**1**	**0,1%**
17	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	**1**	**0,1%**
18	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	**1**	**0,1%**
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0,0%**
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0,0%**
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0,0%**
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	**0,0%**
23	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	**1**	**0,1%**
Totaal	**0**	**22**	**98**	**45**	**36**	**72**	**66**	**75**	**65**	**82**	**66**	**86**	**13**	**10**	**11**	**0**	**747**	**100%**
%	**0,0%**	**2,9%**	**13,1%**	**6,0%**	**4,8%**	**9,6%**	**8,8%**	**10,0%**	**8,7%**	**11,0%**	**8,8%**	**11,5%**	**1,7%**	**1,3%**	**1,5%**	**0,0%**	**100%**	

Een bijzonder geval is verder nog de omrekening van de controleloggermeting in de Keller-software (CL0). Deze variabele wordt door de logger gemeten als drukverschil (mbar). De Keller-software rekent dit drukverschil eerst fysisch om naar de waterhoogte ten opzichte van de druksensor, middels soortelijk gewicht van water en zwaartekrachtsversnelling. En vervolgens naar meter ten opzichte van meetpunt, middels de loggerdiepte ten opzichte van meetpunt. Zo kan CL0 namelijk worden vergeleken met CH0, het resultaat van de controlehandmeting, die eveneens is uitgedrukt in meter ten opzichte van meetpunt. Het bijzondere van de omrekening van CL0 is dat deze is gebaseerd op de in de Keller-software beschikbare waarde van de loggerdiepte (daar aangeduid als *installatiediepte*) en dus niet op de in Menyanthes beschikbare waarde van de loggerdiepte, terwijl die laatste wordt gebruikt bij het omrekenen van de operationele loggermetingen (L0) van meter waterhoogte naar meter ten opzichte van meetpunt en/of NAP. Dit zal de vergelijkbaarheid van CL0 en L0 minder maken, tenzij de installatiediepte uit de Keller-software waarmee CL0 is omgerekend dezelfde is als de loggerdiepte uit de Menyanthes-software waarmee L0 is omgerekend. Maar uit onze verkenning bleek dit vaak niet het geval te zijn.

2.3.4 Valideren van reeks ruwe loggermetingen (L0)

Elke reeks met ruwe loggermetingen (L0) wordt bij validatie (naar de reeks L1) zonodig gecorrigeerd voor een bij de controle geconstateerd verschil met een controlehandmeting (CH0). Tot 2010 werd gecorrigeerd voor te grote verschillen door de softwarematige instelling van de loggerdiepte (in de Keller-software aangeduid als *installatiediepte*) aan te passen. Vanaf 2010 werd in geval van te grote verschillen daarnaast ook gecorrigeerd door in de HydroData-module van Menyanthes een drift en/of lineaire correctie toe te passen, dus zonder aanpassing van de loggerdiepte. Om consistente gevalideerde meetreeksen te krijgen zijn toen meteen ook met terugwerkende kracht de deelreeksen met ruwe loggermetingen (L0) van vóór 2010 gevalideerd (en daarbij zonodig gecorrigeerd) met de HydroData-module van Menyanthes.

De ons geleverde reeksen met gevalideerde loggermetingen (L1) bleken ook de ruwe controleloggermetingen (CL0) te omvatten. Dit geeft op bepaalde uitleesdagen een duidelijk afwijkende waarde in de reeks. Het is niet duidelijk of dit komt door een onvolkomenheid bij het samenstellen van de uitvoerreeks door de beheerder, of door een onvolkomenheid in de HydroData-module van Menyanthes.

2.3.5 Conclusies over foutenbronnen procedure-ondersteunende software

Uit deze inventarisatie komen geen ernstige fouten naar voren van de procedure-ondersteunende software, te weten de Keller-software en de HydroData-module van Menyanthes.
Er is wel sprake van een onvolkomenheid, die ontstaat doordat een controleloggermeting (CL0) met de Keller-software wordt omgezet van drukverschil naar waterhoogte en daarna naar meter ten opzichte van meetpunt of NAP, terwijl dat omzetten voor een ruwe loggermeting (L0) geschiedt met de HydroData-module van Menyanthes. Door niet-overeenkomende instellingen van deze softwarepakketten is CL0 echter weinig representatief voor de ruwe loggermetingen en daardoor speelt deze vanaf 2010 ook geen rol meer bij het valideren. Dit is inefficiënt, want het zo belangrijke validatieproces heeft baat bij zoveel mogelijk vergelijkingsmateriaal. Bovendien wordt CL0 vrijwel gelijktijdig met de controlehandmeting (CH0) gemeten, terwijl dat doorgaans niet geldt voor de dichtstbijzijnde L0.

2.4 Mogelijke Invloed van corrigeren ruwe loggermetingen op meetfout

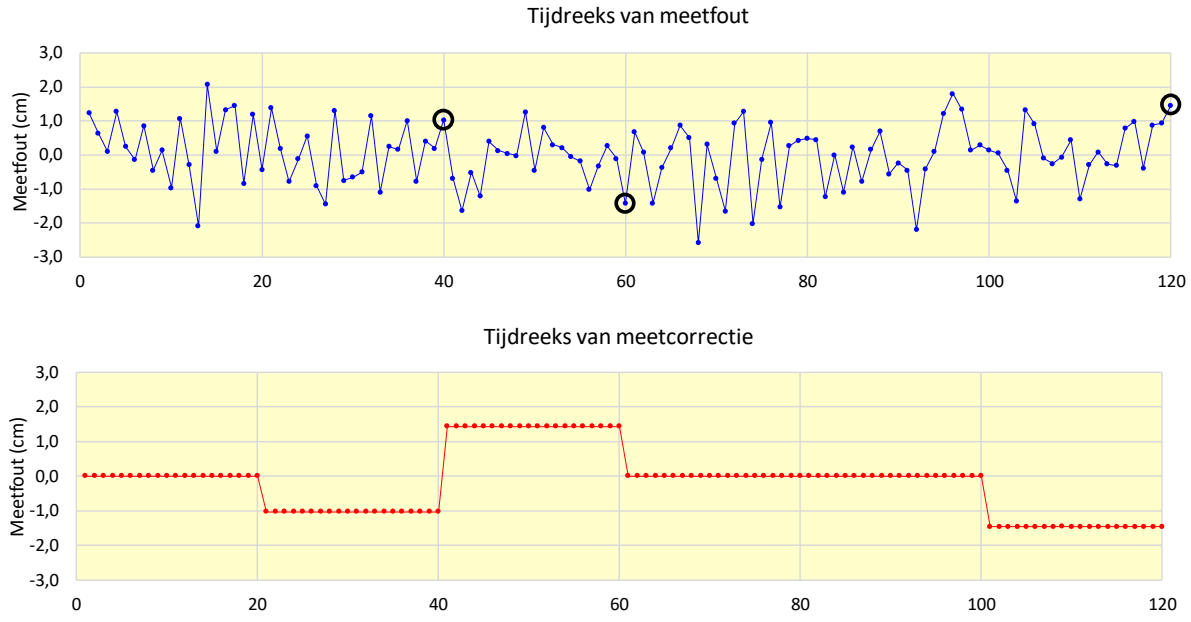
De in § 2.2.1 beschreven procedure voor corrigeren van ruwe loggermetingen gaat uit van de groottes en tekens (+ of -) van opeenvolgende verschillen tussen ruwe loggermetingen (L0) en daaraan gepaarde handmetingen (CH0). De correctiecriteria zijn echter niet statistisch onderbouwd en verder laten zijn vuistregels ruimte voor (subjectieve) invullingen.

Het ontwerp van een correctieprocedure dient bij voorkeur gebaseerd te zijn op kwaliteitseisen voor de uitvoer van het proces waar de procedure op wordt toegepast, die ook toetsbaar zijn. Voor wat betreft het proces van meten van de grondwaterstand met een logger, kunnen we bijvoorbeeld eisen dat de kwaliteit van het resultaat niet verslechtert, wat neerkomt op een beheerst meetproces (andersoortige of aanvullende eisen zijn uiteraard ook mogelijk). Het ligt dan voor de hand de correctiecriteria te baseren op de statistische kenmerken van de bij controle verkregen verschilwaarde van de ruwe loggermeting (L0) en de daaraan gepaarde controlehandmeting (CH0), die representatief zijn voor de fase dat de logger nog geen kwaliteitsverslechtering vertoont.

Als de correctiecriteria niet gebaseerd zijn op een degelijke statistische onderbouwing, ontstaat er een substantieel risico op informatieschade. Onderstaand tekstkader toont daar een voorbeeld van.

Voorbeeld van vertekening van resultaten door een onvoldoende onderbouwd correctiecriterium

Stel dat we met een logger dagelijks de grondwaterstand in een bepaald peilfilter meten, waarbij de logger tot dusverre nog geen kwaliteitsverslechtering vertoont (in de vorm van een groeiende systematische fout of iets dergelijks). De meetfout is dan stationair, met constante verwachtingswaarde (dit is het populatiegemiddelde van de meetfout) en constante standaardafwijking. De eerste plot hieronder toont een tijdreeks van 120 meetfouten van een dergelijk stationair meetproces, hier met verwachtingswaarde 0 cm en standaardafwijking 1 cm. Verder zijn de meetfouten afkomstig uit een normale kansverdeling.
Stel nu dat we uit kostenoverwegingen bij slechts 6 van de 120 metingen de meetfout kunnen vaststellen en beoordelen, door vergelijking met een simultane, foutloze handmeting (dus uit L0-CH0). Dit doen we voor elke laatste meting van een segment van 20 metingen (in de plot begrensd door verticale lijnen). En stel dat min of meer op gevoel een correctiecriterium voor de meetprocedure is opgesteld, dat luidt dat bij een absolute meetfout groter dan 1 cm elke meting van het betreffende segment moet worden gecorrigeerd door er die meetfout van af te trekken. Als we dat criterium toepassen op deze tijdreeks van de meetfout, blijkt correctie nodig voor 3 van de 6 segmenten. De betreffende te grote meetfouten op segmenteinden zijn omcirkeld in de eerste plot en de tijdreeks van de meetcorrectie is weergegeven in de tweede plot. Het betreft soms relatief forse correcties, die omdat ze segmentgewijs worden toegepast in feite systematische fouten van segmenten kunnen introduceren. Het moge duidelijk zijn dat in dit voorbeeld de correcties onnodig relevante schade toebrengen aan feitelijk valide meetinformatie, aangezien die is gegenereerd door een meetproces dat tot dusverre nog geen kwaliteitsverslechtering vertoont. Dit zou niet zijn gebeurd als de correctiecriteria gebaseerd waren op de statistische kenmerken van de meetfout voor de fase dat de logger nog geen kwaliteitsverslechtering vertoont (normale kansverdeling, met gemiddelde 0 cm en standaardafwijking 1 cm).

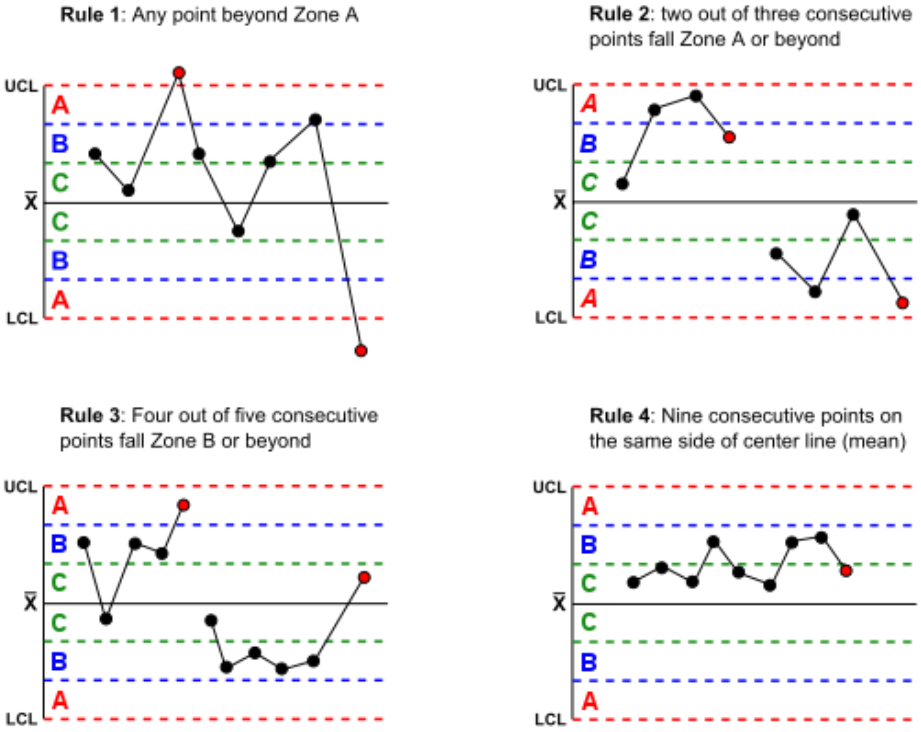


In de industriële wereld worden vaak methoden uit het vakgebied Statistische procesbeheersing[15] toegepast om te bewerkstelligen dat de procesuitvoer (het product) voldoet aan bepaalde kwaliteitseisen. Daarvoor zijn inmiddels meerdere statistische technieken beschikbaar gekomen, waarmee een proces zodanig kan worden gemonitord en beheerst dat voldaan wordt aan die eisen. Deze

[15] In het Engels aangeduid als Statistical Process Control (SPC).

technieken gebruiken meestal een zogenaamde beheerkaart die een aantal horizontale lijnen omvat, met in het midden de gemiddelde waarde (of doelwaarde) en verder waarschuwings- en actiegrenzen, die het gebied verdelen in zes even grote stroken (zie het voorbeeld in figuur 2.4).

Figuur 2.4: Voorbeeld van stelsel besliscriteria bij statistische procesbeheersing met een beheerkaart (in dit geval een Shewhart-kaart).



We adviseren nader te onderzoeken of ook bij het automatisch meten van de grondwaterstand met loggers gebruik kan worden gemaakt van dit soort technieken, om het meetproces beheerst te houden. Daarvoor zal ondermeer kennis nodig zijn van de statistische kenmerken van de verschilwaarde L0-CH0, representatief voor de fase dat het meetproces nog geen kwaliteitsverslechtering vertoont (zoals door drift). In de onderhavige studie is een eerste poging gedaan die kenmerken te ramen van het meetproces van de Provincie Overijssel, maar daar bleken nog diverse haken en ogen aan te zitten (zie § 4.2.1). Het nader onderzoek dient ook na te gaan of statistische trendanalyse kan helpen het meetproces beheerst te houden.

Knelpunt voor toepassing van een statistisch onderbouwde correctieprocedure kan zijn dat er momenteel alleen nog zeer laagfrequent (halfjaarlijks) informatie beschikbaar komt over de kwaliteit van het meetproces (in de vorm van verschilwaarden L0-CH0), terwijl de besliscriteria doorgaans meerdere verschilwaarden vergen (zie figuur 2.4). Dit impliceert dat optimale correcties pas na enkele jaren mogelijk zijn en dan met terugwerkende kracht toegepast moeten worden op een meerjarige deelreeks. Het kan dan enkele jaren duren voor een deelreeks optimaal gevalideerd (en zonodig aansluitend gecorrigeerd) is, wat uiteraard belemmeringen oplevert voor het gebruik van deze data.

2.5 Meetfouten Keller-logger DCX-22 AA

2.5.1 Principe van meten grondwaterstand met logger

Het principe van het meten van de grondwaterstand met een logger is dat de (hydrostatische) druk die de waterkolom uitoefent op de druksensor een functie is van de waterhoogte ten opzichte van de druksensor, volgens:

$$p_h = h \cdot \rho \cdot g \qquad [2.6]$$

waarin p_h de hydrostatische druk van de waterkolom (N/m^2), h de hoogte van de waterkolom (m), ρ het soortelijk gewicht van het betreffende water (kg/m^3) en g de zwaartekrachtsversnelling (m/s^2). De waterhoogte h (m) volgt dan uit:

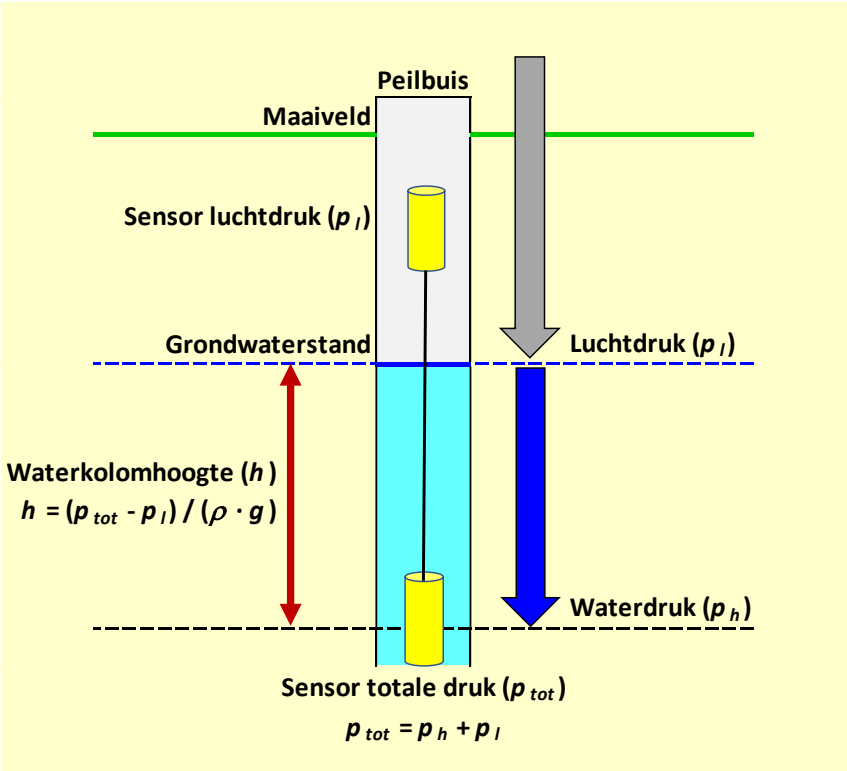
$$h = \frac{p_h}{\rho \cdot g} \qquad [2.7]$$

Deze waterhoogte kan vervolgens worden omgerekend naar de grondwaterstand ten opzichte van een bepaald referentieniveau. De precieze uitwerking van het bepalen van de waterhoogte hangt echter af van het soort logger en dan met name of er wel of geen fysische compensatie is voor de luchtdruk bij het meten van de waterdruk.

De Provincie Overijssel hanteert de Keller-logger van type DCX-22 AA, die geen fysische compensatie voor de luchtdruk heeft (zie figuur 2.5). Deze bevat daarom twee drukopnemers, één voor de luchtdruk (p_l) en één voor het totaal van waterdruk en luchtdruk ($p_{tot} = p_h + p_l$). De waterhoogte h (m) volgt dan uit hun verschildruk, volgens:

$$h = \frac{p_{tot} - p_l}{\rho \cdot g} \qquad [2.8]$$

Figuur 2.5: Meetprincipe van de Keller-logger DCX-22 AA, die twee drukopnemers hanteert.



2.5.2 Systematische fouten door instelling fysische constanten

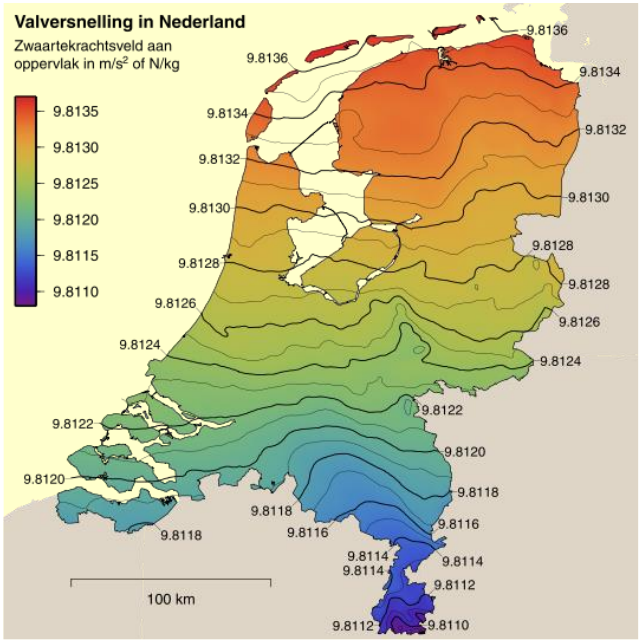
Net als bij veel andere meetinstrumenten is er ook bij de Keller DCX-22 AA sprake van foutenbronnen die een systematische fout kunnen geven (een fout met verwachtingswaarde ongelijk nul). Het betreft ondermeer de twee fysische constanten in formule [2.8], waarmee de verschildruk (p_{tot} - p_l) wordt omgerekend naar de waterhoogte (h), namelijk: 1) de zwaartekrachtsversnelling (g) en 2) het soortelijk gewicht van het betreffende water (ρ). Hun invloeden bespreken we hieronder. Daarnaast kan er ook sprake zijn van drift, een al dan niet geleidelijk opbouwende systematische fout door interne en externe factoren. We bespreken deze aan het eind van § 2.5.4. Andere bronnen van systematische fouten (hier verder niet besproken) zijn falen van de sensor en effecten van temperatuur, niet-lineariteit en hysterese.

Ad 1) Systematische fout door instelling zwaartekrachtsversnelling

Volgens de handleiding hanteert de software[16] van de Keller DCX-22 AA een vaste waarde voor de zwaartekrachtsversnelling (g). Deze staat daarin standaard op 9,80665 m/s^2. Zoals [Von Asmuth, 2010] in een eerdere theoretische foutenanalyse al vermoedde, geldt deze standaardwaarde voor een locatie ergens in Zwitserland, het thuisland van de firma Keller. Maar op wereldschaal varieert de waarde van g, voornamelijk als functie van de breedtegraad (geeft circa 0,5% verschil) en de hoogte (circa 0,003% verschil per 100 meter hoogteverschil). Verder kunnen de lokale topografie en getijdekrachten variaties geven. Aangezien daar door de Keller-logger geen rekening mee kan worden gehouden, zal deze bij gebruik buiten Zwitserland een systematische fout geven bij het omrekenen van waterdruk naar waterhoogte.

Figuur 2.6 toont de ruimtelijke verdeling van de zwaartekrachtsversnelling in Nederland, waaruit blijkt dat die in Overijssel gemiddeld circa 9,8128 m/s^2 bedraagt. Uit formule [2.8] volgt dan dat het automatisch toepassen van de Zwitserse zwaartekrachtsversnelling door Overijsselse Keller-loggers de waterhoogte met 0,06 cm per meter waterkolom overschat (0,3 cm per 5 meter waterkolom).

Figuur 2.6: Ruimtelijke verdeling van de zwaartekrachtsversnelling (g, in m/s^2) in Nederland.[17]



[16] DCX Data Logger. November 2013, Keller AG für Druckmesstechnik, 64 blz.

[17] Bron: Cavit, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=12805351.

Ad 2) Systematische fout door vaste waarde soortelijk gewicht van water

De recente handleiding (2013) vermeldt verder dat de software van de Keller DCX-22 AA standaard voor ρ, het soortelijk gewicht van het betreffende water, de waarde 998,207 kg/m^3 hanteert, het soortelijk gewicht van puur water bij 20 °C. Maar het soortelijk gewicht wijzigt zowel met de watersamenstelling (het zoutgehalte) als met de temperatuur. Het soortelijk gewicht is maximaal bij 3,98 °C en neemt af bij lagere en hogere temperaturen (zie tabel 2.3 voor dit laatste). De handleiding meldt hier alleen over dat de gebruiker de standaardwaarde in de software desgewenst kan vervangen door een andere waarde, dus blijkbaar is er geen optie voor automatisch verdisconteren van veranderingen van het soortelijk gewicht door temperatuurvariaties.

Tabel 2.3: Relatie tussen de temperatuur (T) en het soortelijk gewicht (ρ) van puur water. De laatste twee kolommen vermelden de systematische fout van een waterhoogte als die is bepaald met de standaardwaarde van het soortelijk gewicht die de Keller-logger hanteert, zijnde 998,207 kg/m^3 (die geldt voor een watertemperatuur van 20 °C).

		Syst. fout bij ρ 20 °C	
T (°C)	ρ (kg/m^3)	cm/m	cm/(5 m)
4	999,975	0,18	0,89
10	999,702	0,15	0,75
11	999,602	0,14	0,70
12	999,500	0,13	0,65
15	999,103	0,09	0,45
20	**998,207**	**0,00**	**0,00**
25	997,048	-0,12	-0,58
30	995,650	-0,26	-1,28

De laatste twee kolommen van tabel 2.3 vermelden voor acht watertemperaturen en twee waterkolomhoogten (1 en 5 meter) de systematische fout die wordt gemaakt bij het berekenen van de waterhoogte uit de waterdruk, als daarbij voor het soortelijk gewicht wordt uitgegaan van de standaardwaarde van de Keller DCX-22 AA (998,207 kg/m^3, die geldt voor zuiver water met een temperatuur van 20 °C). De gemiddelde temperatuur van het grondwater in Nederland zal afhangen van zowel de diepte als de locatie. Zo is bijvoorbeeld op circa 10 meter diepte de gemiddelde temperatuur van het grondwater ergens tussen 10 en 12 °C. Bij gebruik van de standaardwaarde kan de systematische fout dan tussen de 0,65 en 0,75 cm per 5 meter waterkolom bedragen. Maar bij een geschikte validatieprocedure zal deze doorgaans beperkt blijven tot de ruwe loggermetingen, omdat die bij validatie gedetecteerd en aansluitend gecorrigeerd zal worden.

Het basisbestand dat we voor deze studie van de Provincie ontvingen (uitvoer van Menyanthes) omvat gegevens van 15.619 uitleesmomenten van de Keller-loggers. Bij 2.065 van deze uitlezingen (13,2%) is in de kolom *density* een soortelijk gewicht van 998,200012207031 kg/m^3 vermeld en bij 365 van de uitlezingen (2,3%) is een soortelijk gewicht vermeld van exact 1.000 kg/m^3 (decimalen ontbreken). Beide waarden komen overigens verspreid voor over de hele meetperiode en beperken zich dus niet tot afgebakende tijdvensters. Bij de overige 13.189 uitlezingen (84,4%) is geen soortelijk gewicht vermeld. Het is echter onduidelijk of de voor het betreffende uitleesmoment vermelde waarde van het soortelijk gewicht destijds ook daadwerkelijk door de Keller-software is toegepast bij het omrekenen van gemeten drukverschillen naar waterhoogten. Deze twijfel betreft overigens alleen de controleloggermetingen (CL0). Alle ruwe operationele loggermetingen (L0), ook die van vóór 2010, zijn namelijk inmiddels met behulp van de HydroData-module van Menyanthes met

geschikt te achten fysische constanten omgezet naar waterhoogten. Daarbij rekent Menyanthes een door de logger gemeten drukverschil standaard om naar waterhoogte ten opzichte van de sensor-diepte met voor *g* de waarde 9,81225 m/s^2 (min of meer het gemiddelde voor Nederland) en voor ρ de waarde 999,9 kg/m^3 (het soortelijk gewicht van water bij 8 °Celsius).[18]

We hebben ten slotte nog voor verschillende grondwatertemperaturen bepaald tot welke systematische fout de instellingen van de fysische constanten in de HydroData-module van Menyanthes leiden ten opzichte van de ideale situatie dat de combinatie van het soortelijk gewicht van water bij de betreffende temperatuur en de zwaartekrachtsversnelling van Overijssel wordt gehanteerd. Uit deze in tabel 2.4 vermelde systematische fouten blijkt dat ze relatief beperkt zijn bij de doorgaans te verwachten grondwatertemperatuur tussen 10 en 12 °C, namelijk tussen -0,07 en -0,17 cm per 5 meter waterkolom.

Tabel 2.4: De systematische fout van een waterhoogte als die is bepaald uit het drukverschil van de logger op basis van de Menyanthes-instellingen van het soortelijk gewicht van water (ρ) van 999,9 kg/m^3 en een zwaartekrachtsversnelling (g) van 9,81225 m/s^2, ten opzichte van de ideale situatie dat dit gebeurt op basis van de ρ van water bij de betreffende temperatuur en de g van Overijssel (9,8128 m/s^2).

	Ideaal		Systematische fout	
T (°C)	ρ (kg/m^3)	g (m/s^2)	cm/m	cm/(5 m)
4	999,975	9,8128	0,013	0,07
10	999,702	9,8128	-0,014	-0,07
11	999,602	9,8128	-0,024	-0,12
12	999,500	9,8128	-0,034	-0,17
15	999,103	9,8128	-0,074	-0,37
20	998,207	9,8128	-0,164	-0,82
25	997,048	9,8128	-0,280	-1,40
30	995,650	9,8128	-0,419	-2,10

[18] Persoonlijke mededeling Jos von Asmuth, 23 oktober 2018.

2.5.3 Kenmerken meetfout druksensoren Keller DCX-22 AA volgens technische specificaties

De technische specificaties van de Keller DCX-22 AA zijn vermeld in tabel 2.5. Deze bevat ook enkele kenmerken van de meetfout. Die gelden echter voor de twee afzonderlijke druksensoren en dus nog niet voor hun gecombineerde werking. In § 2.5.4 werken we op basis van de theorie van de fouten-voortplanting uit wat de kenmerken van de meetfout zijn bij het combineren van de twee druksensoren en daarmee dus bij het combineren van hun meetfouten.

Tabel 2.5: Technische specificaties van de Keller DCX-22 AA. Bron: brochure Keller DCX-22 AA, versie 04/2018, www.keller-druck.com.

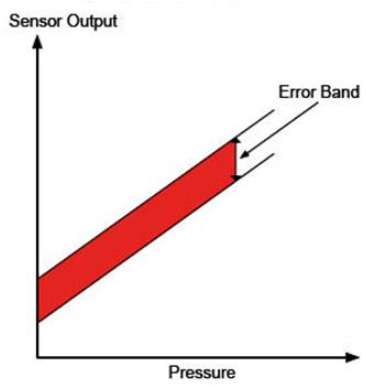
Measuring-/ Pressure Ranges	800…1800 mbar abs. for ranges up to 5 mWC (System Length: 5 m) 800…2300 mbar abs. for ranges up to 10 mWC (System Length: 10 m)		
Supply	Lithium-Battery 3,6 V (Type AA)	Temperature Measurement	Accuracy typ. ±0,5 °C
Battery Life *	10 years @ 1 measurement/hour	Shortest Measuring Range	1x per second
Output	RS 485 digital	Memory	57'000 measuring values @ storage interval ≤ 15 s, otherwise 28'000 measuring values (always with attributed time)
Electrical Connection	Fischer DEE 103A054		
Pressure Sensor Specifications	(for level sensor respectively air pressure sensor)	Material	Stainless steel 316L (DIN 1.4435) O-Ring: Viton®
Linearity	typ. 0,02 %FS		
Comp. Temperature Range	-10…40 °C	Weight: Housing / Probe	≈ 335 g / ≈ 200 g (without cable)
Error Band **	typ. 0,05 %FS *** max. 0,1 %FS	Tolerance System Length	± 2 cm
Resolution	max. 0,0025 %FS		
Long Term Stability	typ. 0,5 mbar		

* exterior influences could reduce battery life ** Linearity + Temperature Error *** optional max. 0,05 %FS

De Keller-brochure stelt verder dat de metingen van de druksensoren rekenkundig worden gecorrigeerd voor fouten van de druksensoren die komen door effecten van temperatuur en niet-lineariteit.

Een belangrijk kwantitatief kenmerk van de meetfout is de foutenband (*error band*). Volgens de formele definitie zou deze aan moeten geven wat de maximale meetfout is van het meetinstrument (*worst case error*). De foutenband strekt zich uit aan weerskanten van de ideale meetlijn (zie figuur 2.7) en als het meetinstrument voldoet aan de specificaties zou deze band alle mogelijke meetwaarden moeten omvatten.

Figuur 2.7: Principe van de foutenband. In het algemeen omvat deze de mogelijke meetwaarden (en daarmee ook mogelijke fouten) van het meetinstrument in zijn normale toepassingsgebied, mits het instrument niet defect is.



Tabel 2.5 specificeert de breedtes van twee uitvoeringen van de foutenband, namelijk:

- de karakteristieke foutenband (*typical error band*): volgens de definitie van deze term mag worden aangenomen dat de helft van de loggers van een productserie daaraan voldoen en dus geen fouten geven buiten die band;

- de maximale foutenband (*maximum error band*): volgens de definitie van deze term mag worden aangenomen dat alle loggers van een productserie daaraan voldoen en dus geen fouten geven buiten die band.

De in tabel 2.5 vermelde specificaties van de foutenband van de druksensoren van de Keller-logger DCX-22 AA houden ook rekening met fouten veroorzaakt door effecten van niet-lineariteit en temperatuur. Ze houden daarentegen geen rekening met effecten van hysterese en veroudering (zoals drift). De drift is apart gespecificeerd in tabel 2.5, als *Long Term Stability* (zie ook het laatste deel van § 2.5.4).

De foutenband is doorgaans uitgedrukt als %*FS* (ook voor de druksensoren van de Keller DCX-22 AA), waarbij *FS* staat voor *Full Scale*, de maximale uitslag van het meetinstrument. De grootte van de maximale fout is dan constant over het gehele meetbereik, ongeacht de meetwaarde.

Volgens de technische specificaties van de Keller DCX-22 AA bedraagt de karakteristieke foutenband van de druksensoren 0,05 %*FS* en de maximale foutenband 0,1 %*FS* (zie tabel 2.5). Als we uitgaan van de laatste foutenband (deze heeft een bredere geldigheid), geeft dit afhankelijk van de meetcapaciteit van de sensor (*FS* 1.800 of 2.300 mbar) de in tabel 2.6 vermelde maximale fouten die kunnen optreden.

Tabel 2.6: Maximale meetwaarden (FS, uitgedrukt in mbar en meter waterkolom) van de twee uitvoeringen van de druksensor van de Keller DCX-22 AA (geschikt voor een maximale druk van 1.800 respectievelijk 2.300 mbar) en hun maximale fouten (0,1 %FS).

	FS	
p_h (mbar)	1.800	2.300
h_{water} (m)	18,345	23,441
Max. fout	**0,1 %*FS***	
p_h (mbar)	1,8	2,3
h_{water} (m)	0,018	0,023

2.5.4 Afleiden kenmerken gecombineerde meetfout druksensoren Keller DCX-22 AA

Zoals toegelicht in § 2.5.1 volgt bij gebruik van de Keller DCX-22 AA de waterhoogte (h) uit de ratio van het verschil van de totale druk (p_{tot}) en de luchtdruk (p_l) en het product van het soortelijk gewicht van het water (ρ) en de zwaartekrachtsversnelling (g):

$$h = \frac{p_{tot} - p_l}{\rho \cdot g} \qquad [2.8]$$

Aangezien de aldus berekende waterhoogte volgt uit metingen van twee druksensoren, is er sprake van een combinatie van meetfouten. In deze paragraaf werken we de kenmerken uit van deze gecombineerde meetfout, met behulp van de theorie van de foutenvoortplanting. We richten ons daarbij alleen op de meetfout die nog niet beïnvloed is door kwaliteitsverslechtering van de logger, bijvoorbeeld in de vorm van drift (een al dan niet geleidelijk opbouwende systematische fout). Eerst werken we de standaardafwijking uit van de gecombineerde meetfout en daarna gaan we in op de drift die kan optreden.

Standaardafwijking van de gecombineerde meetfout

Stel dat een bepaalde variabele y een functie is van de variabelen x_1, x_2, …, x_n, volgens:

$$y = f(x_1, x_2 \ldots, x_n) \qquad [2.9]$$

Volgens de theorie van de foutenvoortplanting kan de standaardonzekerheid van y (aangeduid als $u(y)$)[19] als volgt worden bepaald uit de standaardonzekerheden en covarianties van x_1, x_2, …, x_n als:

$$u(y) = \sqrt{\sum_{i=1}^{n}\left(\frac{\partial y}{\partial x_i}\cdot u(x_i)\right)^2 + \sum_{i=1}^{n}\sum_{\substack{k=1\\k\neq i}}^{n}\left(\frac{\partial y}{\partial x_i}\cdot\frac{\partial y}{\partial x_k}\cdot u(x_i,x_k)\right)} \qquad [2.10]$$

waarin $u(x_i)$ de standaardonzekerheid van variabele x_i (i=1,2,...,n), $\partial y/\partial x_i$ de partiële afgeleide van y naar x_i (op te vatten als de gevoeligheid van y voor x_i) en $u(x_i,x_k)$ de covariantie van x_i en x_k, die volgt uit:

$$u(x_i, x_k) = r_{ik}\cdot u(x_i)\cdot u(x_k) \qquad [2.11]$$

waarin r_{ik} de (lineaire) correlatiecoëfficiënt van x_i en x_k.

Uit formule [2.10] volgt bijvoorbeeld ook de bekende formulering dat de variantie van een som van variabelen gelijk is aan de som van de varianties van de variabelen, mits de variabelen onderling onafhankelijk van elkaar zijn. Hun covarianties, vermeld in de tweede term van het rechterdeel van [2.10] zijn dan immers nul. Vaak wordt echter vergeten deze belangrijke voorwaarde van onderlinge onafhankelijkheid te vermelden.

We werken hieronder eerst de standaardonzekerheid uit van de teller van formule [2.8] en daar volgt dan eenvoudig de standaardonzekerheid van de waterhoogte uit ($u(h)$). De teller van [2.8] is het met de twee druksensoren vastgestelde drukverschil v_p, volgens:

$$v_p = p_{tot} - p_l \qquad [2.12]$$

Uit combineren van [2.10] en [2.12] volgt dan:

$$u(v_p) = \sqrt{\left(\frac{\partial v_p}{\partial p_{tot}}\cdot u(p_{tot})\right)^2 + \left(\frac{\partial v_p}{\partial p_l}\cdot u(p_l)\right)^2 + 2\cdot\frac{\partial v_p}{\partial p_{tot}}\cdot\frac{\partial v_p}{\partial p_l}\cdot u(p_{tot},p_l)} \qquad [2.13]$$

En als we de differentiaties uitwerken volgt hieruit:

$$u(v_p) = \sqrt{\left(u(p_{tot})\right)^2 + \left(u(p_l)\right)^2 - 2\cdot u(p_{tot},p_l)} \qquad [2.14]$$

De voor formule [2.14] benodigde standaardonzekerheden van respectievelijk p_{tot} en p_l kunnen we ramen uit de maximale fout van de druksensor, die volgt uit de technische specificaties. Volgens tabel 2.6 bedraagt de maximale fout van een druksensor van een Keller DCX-22 AA, geschikt voor een maximale druk van 1.800 mbar, 1,8 mbar. Om daaruit de standaardonzekerheid van een met die druksensor gemeten druk te kunnen afleiden, is het van belang of de 1,8 mbar letterlijk de maximale fout betreft (dus met een dekkingsgraad ter grootte van 100% van de meetfoutpopulatie), of een maximale fout bij een enigszins kleinere dekkingsgraad, zoals 99,9%, 99,5%, etc. (het met een experimentele proefopzet vaststellen van een maximale fout met een dekkingsgraad van 100% kan namelijk onpraktisch veel meetinspanning vergen). Dat is echter niet duidelijk, zodat we het hieronder voor verschillende dekkingsgraden van de foutenband zullen uitwerken.

[19] De standaardonzekerheid van een variabele y (zoals de waterhoogte in een bepaald peilfilter op een bepaald tijdstip) is de standaardafwijking van meetwaarden die redelijkerwijs kunnen worden toegewezen aan y.

Als we er van uitgaan dat de meetfout min of meer een normale kansverdeling volgt[20], dan zal gelden:

$$u(p_{tot}) = u(p_l) = \frac{1{,}8}{z_{dg}} \text{ mbar} \qquad [2.15]$$

waarin z_{dg} het percentiel van de standaardnormale kansverdeling[21] dat aansluit op de dekkingsgraad van de maximale foutenband. Een voorbeeld: voor een dekkingsgraad van 99% dient voor z_{dg} in bovenstaande formule $z_{99,5\%}$ te worden gehanteerd, het 99,5-percentiel van de standaardnormale kansverdeling (deze heeft de waarde 2,576). Uit combinatie van [2.14] en [2.15] volgt:

$$u(v_p) = \sqrt{2 \cdot \left(\frac{1{,}8}{z_{dg}}\right)^2 - 2 \cdot u(p_{tot}, p_l)} \qquad [2.16]$$

Het is niet eenvoudig om de tweede term in het rechterdeel van [2.16] te ramen (de covariantie van de meetfouten van p_{tot} en p_l). Daarvoor moeten we namelijk weten welke componenten van hun meetfouten gecorreleerd zijn, in welke mate ze gecorreleerd zijn en wat de bijdrage van die componenten aan de totale meetfout is, dit alles bovendien bij gebruik van de Keller DCX-22 AA in Overijsselse peilbuizen. Maar de technische specificaties van de Keller DCX-22 AA geven daar geen informatie over. Hieronder gaan we daarom na hoe ver we zelf met logisch redeneren kunnen komen.

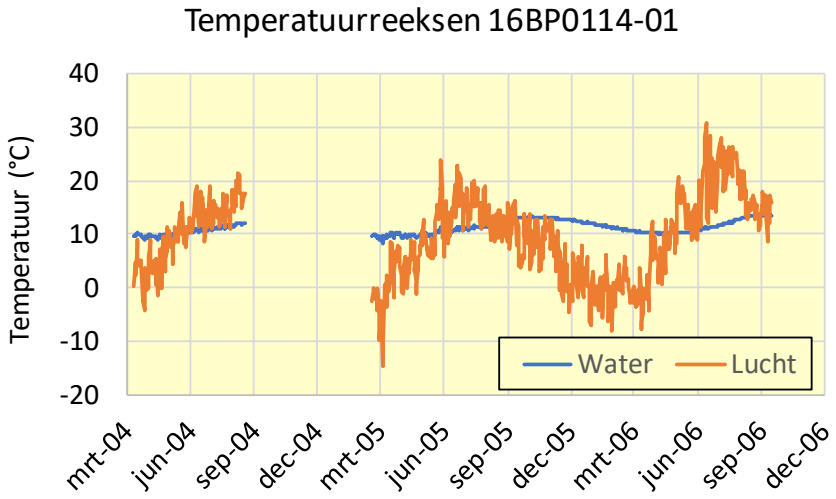
De relevantste foutenbronnen van een druksensor die nog geen kwaliteitsverslechtering heeft gekend zijn vermoedelijk temperatuurvariatie, niet-lineariteit en hysterese. En de belangrijkste daarvan is vermoedelijk de temperatuurvariatie, omdat die directe invloed uitoefent op allerlei componenten van het meetsysteem, zoals uitzetting en inkrimping van het sensormateriaal, het water en de behuizing. Verder heeft het thermo-elektrische effect (conversie van temperatuurverschil naar een elektrische spanning) ook invloed op de weerstanden, semigeleiders en elektrische verbindingen van de sensor [Gassmann, 2014].
Voor wat betreft de Keller DCX-22 AA, met een sensor voor totale druk (beïnvloed door de watertemperatuur) en een sensor voor luchtdruk (beïnvloed door de luchttemperatuur) is het niet onredelijk te veronderstellen dat de relevantste correlatie tussen foutenbronnen van de twee druksensoren die tussen de watertemperatuur en luchttemperatuur is. Bij een eerste globale verkenning constateerden we een niet-verwaarloosbare correlatiecoëfficiënt van de watertemperatuur en de luchttemperatuur in een peilbuis (ter grootte van circa 0,3, zie het tekstkader hieronder).

[20] Argumenten voor het benaderen van normaliteit door de meetfout zijn dat de gebruikte technische specificaties betrekking hebben op een logger die nog geen kwaliteitsverslechtering vertoont en er dus waarschijnlijk weinig sprake zal zijn van uitschieters, zoals door falen. Verder kan worden aangenomen dat er sprake is van meerdere foutenbronnen en dan zal volgens de Centrale Limietstelling de kansverdeling van de meetfout een normale verdeling gaan benaderen.

[21] De standaardnormale kansverdeling is een normale kansverdeling met gemiddelde nul en standaardafwijking 1. Elke normale kansverdeling kan worden gestandaardiseerd naar deze kansverdeling, door van elke waarde het gemiddelde af te trekken en dat verschil vervolgens te delen door de standaardafwijking.

Correlatie van watertemperatuur en luchttemperatuur in een peilbuis

Figuur 2.8: Verloop van watertemperatuur en luchttemperatuur in een peilbuis, zoals dagelijks om 08:00 uur geregistreerd door een Keller-logger DCX-22 AA, van 12 maart 2004 t/m 15 september 2006 (736 meetwaarden per reeks). Er ontbreekt circa een half jaar aan meetgegevens. Onder de figuur zijn enkele kengetallen van de twee reeksen vermeld.



Kengetal	Water °C	Lucht °C
Minimum	8,2	-14,8
Maximum	13,7	31,0
Gemiddelde	11,20	9,35
Standaardafwijking	1,19	7,56

Uit bovenstaande figuur blijkt dat beide temperatuurreeksen een seizoenpatroon vertonen in de vorm van een sinus, maar bij de luchttemperatuur is de sinus beduidend groter en bovendien komen daarop gesuperponeerd ook nog grote, kortdurende temperatuurschommelingen voor. Verder blijkt de watertemperatuur achter te lopen op de luchttemperatuur. De over deze periode bepaalde standaardafwijking van de luchttemperatuur is ruim zes maal groter dan die van de watertemperatuur, wat ook naar voren komt in de veel grotere extremen van de luchttemperatuur. Desondanks bedraagt de lineaire correlatiecoëfficiënt van deze twee temperatuurreeksen 0,33, wat niet verwaarloosbaar klein is.

Het is echter niet mogelijk om louter op basis van een niet-verwaarloosbare (positieve) correlatiecoefficiënt van de watertemperatuur en de luchttemperatuur in een peilbuis te concluderen dat dan ook de meetfouten van de twee druksensoren (positief) gecorreleerd zullen zijn. De meting van een Keller-sensor wordt namelijk intern gecorrigeerd voor de temperatuur, zodat de resterende meetfout een geringere relatie met de temperatuur zal vertonen en de meetfouten van de twee sensoren dan zelfs ongecorreleerd kunnen zijn. Maar omdat deze correcties doorgaans onvolkomen zullen zijn (te zwak of te sterk), is evenmin uit te sluiten dat ze kunnen leiden tot een (geringe) negatieve correlatie van de meetfouten van de twee druksensoren.[22]

We concluderen dat het niet mogelijk is de covariantie van de meetfouten van p_{tot} en p_l (de tweede term in het rechterdeel van [2.16]) te ramen. Maar het is wel denkbaar dat die covariantie hooguit beperkt kan zijn en wel om de volgende redenen: i) door de grote vertraging tussen luchttempera-

[22] Met dank aan Jos von Asmuth (KWR) voor het constructief meedenken over deze problematiek.

tuur en watertemperatuur kan hun correlatie niet groot zijn. De in figuur 2.8 weergegeven reeksen hebben bijvoorbeeld een lineaire correlatiecoëfficiënt van circa 0,3, wat gering is, zij het nog niet verwaarloosbaar; ii) door de interne correcties van sensormetingen voor de temperatuur, zal de meetfout van een sensor minder gerelateerd zijn aan de temperatuur en zullen de meetfouten van de twee sensoren nog minder aan elkaar gecorreleerd zijn dan de watertemperatuur en de lucht-temperatuur.

Onder de hierboven enigszins plausibel gemaakte vooronderstelling dat de covariantie van de meetfouten van de twee druksensoren hooguit beperkt is, laten we deze weg uit formule [2.16] en benaderen dan de standaardonzekerheid van het met de twee druksensoren vastgestelde drukverschil v_p volgens:

$$u(v_p) \cong \sqrt{2 \cdot \left(\frac{1{,}8}{z_{dg}}\right)^2} \qquad [2.17]$$

Deze raming kan vervolgens per dekkingsgraad met formule [2.8] worden omgerekend naar de standaardonzekerheid van de waterhoogte *h*, met de door Menyanthes gehanteerde instellingen van ρ en *g* (zie § 2.5.2). De standaardonzekerheid van de waterhoogte blijkt dan tussen 0,8 en 1,3 cm te kunnen liggen (zie tabel 2.7).

Tabel 2.7: Raming van de standaardonzekerheid van de waterhoogte (u(h)), als functie van de dekkingsgraad van de maximale foutenband. De ramingen gaan er van uit dat er geen relevante correlatie is van de meetfouten van de twee druksensoren.

Dekkings-graad	z_{dg} (-)	$u(p_{tot})$ $u(p_l)$ (mbar)	Raming $u(v_p)$ (mbar)	Raming $u(h)$ (cm Wh)
95%	1,960	0,918	1,299	1,32
97,5%	2,241	0,803	1,136	1,16
99%	2,576	0,699	0,988	1,01
99,5%	2,807	0,641	0,907	0,92
99,9%	3,291	0,547	0,774	0,79

z_{dg}: *z*-waarde bij de betreffende dekkingsgraad
Wh: Waterhoogte

Stel dat ondanks de interne correcties van de sensormetingen voor de temperatuur de meetfouten van de twee sensoren nog een lineaire correlatiecoëfficiënt van 0,2 hebben, dan volgt uit combineren van formules [2.16] en [2.11] dat de standaardonzekerheid van de waterhoogte circa 10% afwijkt van de in tabel 2.7 vermelde waarde. Bij een positieve correlatie zal de standaardonzekerheid circa 10% kleiner zijn dan vermeld, terwijl deze bij een negatieve correlatie circa 10% groter zal zijn dan vermeld.

Andere kanttekeningen bij de in tabel 2.7 vermelde ramingen van de standaardonzekerheid van de waterhoogte zijn:

- in formule [2.16] nemen we aan dat de standaardonzekerheden van totale druk en luchtdruk gelijk zijn, maar de sensor van de totale druk heeft te maken met een veel geringer temperatuurbereik dan de luchtdruksensor (zie figuur 2.8), zodat zijn standaardonzekerheid vermoedelijk ook kleiner zal zijn;

- de luchtdruksensor heeft soms te maken met zeer lage temperaturen, buiten zijn compensatiebereik (zie figuur 2.8), dat loopt van -10 tot 40 °C, waardoor deze soms een duidelijk grotere meetfout zal kunnen vertonen.

Deze afwijkingen kunnen elkaar enigszins opheffen, maar het is niet bekend in welke mate.

Drift, een al dan niet geleidelijk opbouwende systematische fout

Ten slotte kan een druksensor nog onderhevig zijn aan drift, een al dan niet geleidelijk opbouwende systematische fout. Dit verschijnsel kan samenhangen met interne factoren (veroudering) en externe factoren (zoals grootte en frequentie van druk- en temperatuurvariaties, trillingen en schokken). In de technische specificaties van de druksensoren van de Keller DCX-22 AA is deze aangeduid als *long term stability*[23] (zie tabel 2.5). Dit is op te vatten als een maat voor de geleidelijk opbouwende systematische fout onder normale omstandigheden. Bij veel meetinstrumenten betreft dit de vertekening na een periode van een jaar, al heeft Keller de periode hier niet gespecificeerd. Volgens de technische specificatie bedraagt de karakteristieke drift 0,5 mbar, wat neerkomt op circa 0,5 cm waterkolom. Maar deze geldt per druksensor, terwijl we hier te maken hebben met het verschil van twee druksensoren.
Een maat zoals *long term stability* is eigenlijk meer geschikt om algemene kenmerken van verschillende soorten meetinstrumenten te kunnen vergelijken en heeft weinig zeggingskracht voor individuele gevallen. Dit komt doordat de drift niet alleen wordt bepaald door interne factoren, maar ook door meerdere onvoorspelbare externe factoren, zoals grootte en frequentie van druk- en temperatuurvariaties, trillingen en schokken.

In hoofdstuk 4 zullen we op verschillende wijzen de invloed van drift op het functioneren van de logger kwantificeren. Zo wordt nagegaan welke relatie er bestaat tussen de leeftijd van de logger en de spreiding van de verschilwaarden L0-CH0 (die zal toenemen door drift). Verder wordt statistisch getoetst op trends van de verschillen tussen de controlemetingen CL0 en CH0, die zijn op te vatten als manifestatie van drift.

2.6 Kernpunten van dit hoofdstuk

- We hebben geen ernstige fouten geconstateerd van de procedure-ondersteunende software, te weten de Keller-software en de HydroData-module van Menyanthes.
- Er is wel sprake van een onvolkomenheid, die ontstaat doordat een controleloggermeting (CL0) met de Keller-software wordt omgezet van drukverschil naar waterhoogte en daarna naar meter ten opzichte van meetpunt of NAP, terwijl dat omzetten voor een ruwe loggermeting (L0) geschiedt met de HydroData-module van Menyanthes. Door niet-overeenkomende instellingen van deze softwarepakketten is CL0 echter weinig representatief voor de ruwe loggermetingen en daardoor speelt deze vanaf 2010 ook geen rol meer bij het valideren. Dit is inefficiënt, want het zo belangrijke validatieproces heeft baat bij zoveel mogelijk vergelijkingsmateriaal. Bovendien wordt CL0 vrijwel gelijktijdig met de controlehandmeting (CH0) gemeten, terwijl dat doorgaans niet geldt voor de dichtstbijzijnde L0.

[23] Deze aanduiding is enigszins verwarrend, want in feite is het eerder een maat voor de instabiliteit dan voor de stabiliteit.

- Er zijn verder enkele onvolkomenheden geconstateerd in details van de uitvoering van de procedure, zoals:
  - de starttijd van loggerdiepte en meetpuntshoogte blijkt in 747 gevallen (6,4% van alle controledatums) niet juist te zijn geregistreerd. Dit trad vooral op tot 2015 en kan in theorie hebben geleid tot fouten bij het omrekenen van de controleloggermeting en de controlehandmeting (beide oorspronkelijk uitgedrukt ten opzichte van meetpunt) naar een ander referentieniveau en fouten bij het omrekenen van loggermetingen (uitgedrukt in hoogte waterkolom ten opzichte van de sensordiepte) naar een ander referentieniveau, maar in de praktijk zal het nauwelijks zijn voorgekomen. Deze onvolkomenheid is overigens in 2015 opgelost;
  - de criteria voor correctie van een ruwe loggerreeks op basis van de bij controles vastgestelde verschillen tussen ruwe loggermetingen (L0) en daaraan gepaarde handmetingen (CH0) zijn niet statistisch onderbouwd, wat kan leiden tot het beschadigen van meetinformatie. We adviseren nader te onderzoeken of gebruik kan worden gemaakt van technieken uit het vakgebied Statistische Procesbeheersing om het meetproces beheerst - dat wil zeggen, zonder kwaliteitsverslechtering - te houden. Daarvoor zal ondermeer kennis nodig zijn van de statistische kenmerken van de verschilwaarde L0-CH0, representatief voor de fase dat het meetproces nog geen kwaliteitsverslechtering vertoont (zoals door drift). Het nader onderzoek dient ook na te gaan of statistische trendanalyse kan helpen het meetproces beheerst te houden.
- Aan de hand van een aantal technische specificaties van de Keller-logger DCX-22 AA is langs theoretische weg de standaardafwijking van zijn meetfout afgeleid, voor de situatie dat de logger nog geen kwaliteitsverslechtering vertoont. Deze blijkt dan tussen 0,8 en 1,3 cm waterhoogte te kunnen liggen, afhankelijk van wat de dekkingsgraad is van de maximale foutenband (dit is niet duidelijk uit de technische specificaties).

3 Empirische afleiding kenmerken meetfout

In dit hoofdstuk leiden we langs empirische weg de kenmerken af van de meetfout van de logger, op basis van de verschillen tussen loggermetingen en daaraan gepaarde controlehandmetingen. Daarbij wordt onderscheid gemaakt tussen de meetfout vóór en ná validatie (en zonodig daarbij toegepaste correctie).

3.1 Beschouwde verschillen tussen loggermetingen en handmetingen

Kenmerken van de meetfout kunnen empirisch worden afgeleid uit de verschillen tussen loggermetingen en daaraan gepaarde handmetingen van de grondwaterstand, waarbij dan aansluitend nog wel de meetfout van de handmetingen moet worden verdisconteerd. Dit zullen we doen aan de hand van drie datasets, elk met een groot aantal gepaarde loggermetingen en handmetingen. Omdat de drie datasets verschillen in het sóórt loggermeting (ruw of gevalideerd), zullen ze ook verschillende informatie opleveren over de mate waarin loggermetingen kunnen afwijken van de daaraan gepaarde handmetingen, zoals hieronder toegelicht.

Codering van het soort meting

Bij onze codering van het soort meting staat C staat voor controlemeting, L voor loggermeting, H voor handmeting, 0 voor ruwe meting en 1 voor gevalideerde (en zonodig gecorrigeerde) meting. De controleloggermeting (CL0) wordt ook wel aangeduid als de online-loggermeting.

Dataset 1: L0-CH0

Deze dataset omvat een combinatie van ruwe operationele loggermetingen (L0) en de daaraan gepaard te achten controlehandmetingen (CH0). De set omvat per uitleesmoment hooguit één ruwe operationele loggermeting - namelijk degene die het dichtst bij de controlehandmeting is genomen, zij het op voorwaarde dat er niet meer dan één uur tussen zit - en hooguit één controlehandmeting. Een andere voorwaarde is dat er voor elke operationele loggermeting (L0) ook een gevalideerde (en daarbij zonodig gecorrigeerde) versie (L1) beschikbaar is. Want als dat niet het geval is, was er blijkbaar voldoende aanleiding voor de beheerder om de prestaties van de logger ongeschikt te verklaren (wat doorgaans ook resulteert in vervanging van de logger). De betreffende L0 is in dat geval ongeschikt om informatie te verschaffen over de prestaties van de logger onder gebruikelijke operationele omstandigheden.
De verschilwaarden L0-CH0 van deze set geven informatie over de mate waarin ruwe operationele loggermetingen kunnen afwijken van de daaraan gepaarde handmetingen.

Dataset 2: L1-CH0

Deze set omvat een combinatie van gevalideerde (en daarbij zonodig gecorrigeerde) versies van de operationele loggermetingen (L1) en de daaraan gepaard te achten controlehandmetingen (CH0). De dataset is afgeleid uit dataset 1, door elke L0 te vervangen door zijn gevalideerde versie (L1).
De verschilwaarden L1-CH0 van deze set geven informatie over de mate waarin gevalideerde (en zonodig gecorrigeerde) loggermetingen kunnen afwijken van de daaraan gepaarde handmetingen. Dit zal echter een onderschatting opleveren van de werkelijke mate van afwijking tussen dit soort metingen, omdat het nog maar een selecte subset van gevalideerde loggermetingen betreft waarvan de gepaarde handmetingen bekend waren en ook gebruikt zijn om de correcties toe te passen.

Van alle andere gevalideerde loggermetingen (dit is veruit de meerderheid) waren echter geen gepaarde handmetingen beschikbaar, zodat hun ruwe uitvoeringen hooguit op basis van interpolatie tussen wél bekende afwijkingen konden worden gecorrigeerd.

Dataset 3: CL0-CH0

Deze dataset omvat een combinatie van controleloggermetingen (CL0) en de vrijwel gelijktijdig genomen controlehandmetingen (CH0), die zijn uitgevoerd bij het circa halfjaarlijks uitlezen van een logger. De set omvat per uitleesmoment hooguit één controleloggermeting en hooguit één controlehandmeting.

De verschilwaarden CL0-CH0 van deze set geven informatie over de mate waarin controleloggermetingen kunnen afwijken van de daaraan gepaarde handmetingen. We dienen ons echter wel te realiseren dat er gedurende de meetperiode zonodig tussentijdse correcties zijn toegepast op instellingen van de loggersoftware om de afwijkingen niet teveel uit de pas te laten lopen. De ruwe controleloggermetingen (totale druk en luchtdruk) worden namelijk door de Keller-software omgezet van drukverschil naar waterhoogte middels een fysische formule en daarna naar grondwaterstand ten opzichte van meetpunt middels de loggerdiepte (daar aangeduid als *installatiediepte*) ten opzichte van meetpunt. Maar als bij het uitleesmoment van de logger de controleloggermeting en de controlehandmeting teveel afweken, is de installatiediepte door de beheerder bijgesteld in de Keller-software, met als doel een kleinere afwijking te krijgen bij het volgende uitleesmoment. Deze bijstelling heeft geen effect meer op de controleloggermeting van dat uitleesmoment, maar zal wel effect hebben op die van het volgende uitleesmoment.[24] Het is daarom te verwachten dat controleloggermetingen (CL0) in het algemeen beter aansluiten op de handmetingen dan ruwe operationele loggermetingen (L0).

3.2 Verzamelen en voorbewerken van loggermetingen en handmetingen

De ruwe operationele loggermetingen (L0) zijn geëxtraheerd uit het eerder beschreven binaire uitvoerbestand van de HydroData-module van Menyanthes (zie § 2.3). De metingen zijn in dat bestand uitgedrukt in meter waterhoogte (ten opzichte van de sensordiepte), zoals destijds bepaald met de Keller-software uit de verschildruk van de twee sensoren. Om deze te kunnen afzetten tegen de controlehandmetingen (CH0) zijn ze door ons omgezet naar meter ten opzichte van NAP, na correctie van de verkeerde starttijden van metadata (zie het tekstkader hieronder). Daarbij is gebruik gemaakt van de loggerdiepte vermeld in de HydroData-module van Menyanthes.

Correctie voor verkeerde starttijden

Om bij de bovengenoemde omzettingen van het referentieniveau geen vertekeningen te krijgen door verkeerde starttijden van metadata (zie § 2.3.3), is elke starttijd van metadata die minder dan 24 uur eerder is dan het tijdstip van de controlemetingen op 1 minuut na dat laatste tijdstip gezet. Daarbij zijn echter uitgezonderd de gevallen dat er sprake is van een initiële loggerinstallatie (plaatsing van de eerste logger). Bij een initiële loggerinstallatie hebben de metagegevens immers betrekking op controlemetingen die pas ná de installatie kunnen zijn uitgevoerd en moeten de metagegevens dus gedateerd zijn vóór de controlemetingen.

[24] In deze zin kunnen controleloggermetingen (CL0) zijn beïnvloed door bij voorgaande uitleesmomenten uitgevoerde aanpassingen van de installatiediepte, bedoeld om ze niet teveel uit de pas te laten lopen met de controlehandmetingen. Het betrof dus geen volledige validatie (en zonodige correctie) zoals gehanteerd bij de ruwe operationele loggermeting (L0), zodat we het resultaat hier blijven aanduiden als CL0 en niet als CL1.

De gevalideerde operationele loggermetingen (L1) zijn afkomstig uit een csv-bestand (een exportbestand van Menyanthes), waarin ze zijn uitgedrukt in meter ten opzichte van NAP. Al deze reeksen zijn gevalideerd met de HydroData-module van Menyanthes, ook de delen van vóór 2010.

Ook de controleloggermetingen (CL0) en de vrijwel gelijktijdige controlehandmetingen (CH0) zijn afkomstig uit het eerder beschreven binaire uitvoerbestand van de HydroData-module van Menyanthes (zie § 2.3). Deze zijn daarin beide uitgedrukt in meter ten opzichte van meetpunt en zowel voor de eenduidigheid als voor grafische beoordelingen door ons omgezet naar meter ten opzichte van NAP, na correctie van de verkeerde starttijden van metadata.

Er zijn veel minder verschilwaarden L0-CH0 en L1-CH0 (elk 3.655) dan verschilwaarden CL0-CH0 (11.474), door de restrictie dat er niet meer dan één uur verschil mag zitten tussen het meettijdstip van L0/L1 en CH0 om ze als gepaard te mogen beschouwen. En omdat er tot eind 2014 / begin 2015 hoofdzakelijk eenmaal dagelijks is gemeten om 08:00 of 09:00 ’s ochtends, terwijl veel uitleesmomenten (en bijbehorende controlemetingen) doorgaans later waren, vallen de meeste metingen uit die periode af.

3.3 Kenmerken van de verschillen tussen loggermeting en handmeting

Van elk van de drie soorten loggermeting zijn kenmerken bepaald van de verschillen met de daaraan gepaarde controlehandmetingen. Deze zijn vermeld in tabel 3.1.

Tabel 3.1: Kenmerken van de verschillen tussen loggermetingen en de daaraan gepaarde controlehandmetingen, voor drie soorten loggermeting. De kolommen 2 en 3 bevatten de kenmerken van de 3.655 beschikbare verschilwaarden L0-CH0 en L1-CH0, waarbij het meettijdstip van L0 niet meer dan een uur verschilt van dat van CH0 en waarbij L1 de gevalideerde (en zonodig gecorrigeerde) versie van L0 is. Kolom 4 bevat de kenmerken van de verschilwaarden CL0-CH0, waarbij het dezelfde 3.655 CH0-waarden betreft die betrokken zijn in de voorgaande twee kolommen. Ten slotte bevat kolom 5 de kenmerken van alle 11.474 beschikbare verschilwaarden CL0-CH0.

	Dataset 1	Dataset 2	Dataset 3	
Kenmerk	**L0-CH0**	**L1-CH0**	**CL0-CH0**	**CL0-CH0**
Aantal waarden (-)	3.655	3.655	3.655	11.474
Minimum (m)	-0,7154	-0,3390	-0,2310	-0,2420
P5 (m)	-0,0352	-0,0090	-0,0280	-0,0310
Mediaan (m)	-0,0015	0,0000	0,0000	0,0000
Gemiddelde (m)	-0,0039	0,0001	0,0002	-0,0012
P95 (m)	0,0190	0,0100	0,0220	0,0250
Maximum (m)	0,5185	0,2520	0,4170	0,4170
Standaardafw (m)	**0,0363**	**0,0094**	**0,0225**	**0,0253**

P5 en P95 zijn het geschatte 5-percentiel en 95-percentiel van de achterliggende populatie.

Uit tabel 3.1 blijkt het volgende.

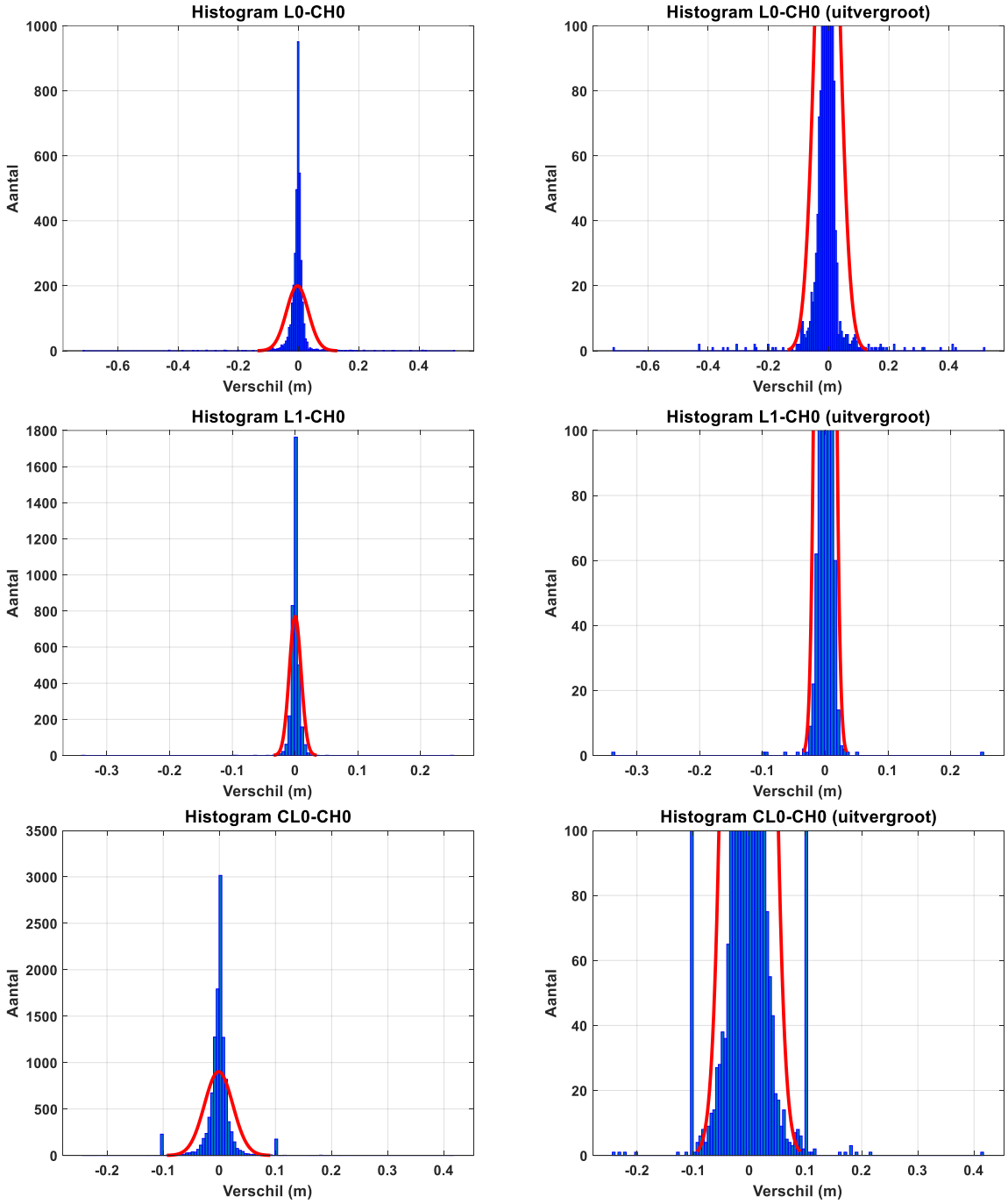
- Elk van de vier verzamelingen verschilwaarden bevat flinke uitschieters (van meerdere decimeters), zowel naar boven als naar beneden.
- De bulk van elke verdeling is relatief smal (90% of meer van de verschilwaarden bevindt zich in het interval van -3 en +3 cm).

- De standaardafwijking (blauw gekleurde cellen) is het grootst voor L0-CH0 (3,63 cm) en het kleinst voor L1-CH0 (0,94 cm). Die van CL0-CH0 zit daar tussenin (2,25 cm voor de deelset en 2,53 voor de hele set). Deze volgorde stemt overeen met de verwachting, aangezien L0 een ruwe loggermeting is en L1 zijn gevalideerde en zonodig gecorrigeerde versie, gebaseerd op het beoordelen van de verschilwaarden L0-CH0 van enkele opeenvolgende uitleesmomenten (waaronder dat van de betreffende L0). Verder zal CL0 doorgaans minder effectief zijn gecorrigeerd dan L0, zoals toegelicht in het laatste deel van § 3.1.
- Van elk van de vier verzamelingen is de geschatte populatiemediaan vrijwel 0. Alleen voor L0-CH0 wijkt deze licht af van 0 (-0,15 cm). Het geschatte populatiegemiddelde wijkt voor elk van de vier deelverzamelingen enigszins af van 0.
- De kenmerken van de subset van CL0-CH0 (met 3.655 verschilwaarden) zijn min of meer vergelijkbaar met die van de volledige set (met 11.474 verschilwaarden).

Figuur 3.1 toont de histogrammen van de drie datasets van verschilwaarden. De volgende punten vallen op aan deze histogrammen.

- Elk van de drie verdelingen wordt gekenmerkt door een sterke piek van het centrum en lange, maar zeer dunne staarten. Geen van de drie soorten verschilwaarde vertoont daarom een normale kansverdeling.
- Uit de gepiekte flanken van het histogram van CL0-CH0 blijkt dat er onverwacht vaak – gelet op de verder relatief gladde frequentieverdeling - verschillen tussen CL0 en CH0 optreden van nabij -10 en +10 cm. Bij gedetailleerdere beschouwing blijkt dat dit komt doordat er in 405 (3,53%) van de 11.474 gevallen sprake is van hetzij exact -10 cm verschil, of exact +10 cm verschil tussen CL0 en CH0. Uit navraag bij de Provincie blijkt dat het een door de beheerder toegepaste afknotting betreft. Bij een verschilwaarde kleiner dan -10 cm of groter dan +10 cm heeft deze de controleloggermeting (CL0) in de meeste gevallen op exact -10 cm respectievelijk +10 cm verschil van de bijbehorende controlehandmeting (CH0) gezet, omdat bij een dergelijk groot verschil de logger toch wordt vervangen en dan is de exacte verschilwaarde niet meer relevant.

Figuur 3.1: Histogram van achtereenvolgens dataset L0-CH0 (3.655 waarden), L1-CH0 (3.655 waarden) en CL0-CH0 (11.474 waarden), met voor elk rechts een uitvergroting van de onderkant. De horizontale as heeft klassen van 0,5 cm. De rode lijn toont de (theoretische) normale kansverdeling met het zelfde gemiddelde en dezelfde standaardafwijking als de betreffende dataset.



3.4 Afleiden standaardafwijking van de meetfout

Om uit de hiervoor afgeleide kenmerken van het verschil tussen loggermeting en controlehandmeting de kenmerken van de meetfout af te kunnen leiden, moet worden verdisconteerd dat ook de controlehandmeting onderhevig is aan een meetfout. Gelukkig zijn er empirische ramingen beschikbaar van die meetfout. Bij een factoriële proefopzet met zestien waarnemers en zestien peilbuizen in het Boelerveld (Overijssel) zijn namelijk de kengetallen van de meetfout van handmatig meten

van de grondwaterstand geraamd [Knotters et al, 2013]. Daarbij is onderscheid gemaakt tussen meten met: i) een akoestisch dompelklokje en ii) een elektronisch peilapparaat met geluidssignaal (de Provincie Overijssel gebruikt deze laatste voor de controlehandmetingen). Volgens genoemde studie kan de meetfout van die aanpak worden benaderd met een normale verdeling, met gemiddelde 0,5 cm en standaardafwijking 0,84 cm. In tabel 3.2 is uitgewerkt hoe uit deze kengetallen de kengetallen van de meetfout van een ruwe loggermeting (L0) zijn afgeleid.

Tabel 3.2: Afleiding van kengetallen van de meetfout van een ruwe loggermeting (L0, blauw gekleurd, laatste kolom).

Kengetal	verschil L0-CH0 [1)]	meetfout CH0 [2)]	meetfout L0 [3)]
Gemiddelde (cm)	**-0,04**	**0,50**	**0,46** [3)]
Standaardafwijking (cm)	**3,63**	**0,84**	**3,53** [4)]

L0 : ruwe loggermeting | **CH0** : controlehandmeting
1) empirische raming (deze studie)
2) empirische raming [Knotters et al., 2013]
3) volgt uit toepassing formule $\bar{x}_{L0-CH0} = \bar{e}_{L0} - \bar{e}_{CH0}$
4) volgt uit toepassing formule $s^2_{L0-CH0} = s^2_{e,L0} + s^2_{e,CH0}$

$\bar{x}_{L0-CH0}$	gemiddelde van L0-CH0
$\bar{e}_{L0}$ resp. $\bar{e}_{CH0}$	gemiddelde meetfout L0 resp. CH0
s^2_{L0-CH0}	variantie van L0-CH0
$s^2_{e,L0}$ resp. $s^2_{e,CH0}$	variantie meetfout L0 resp. CH0

Volgens bovenstaande afleiding wordt de standaardafwijking van de meetfout van de ruwe loggermeting geraamd op circa 3,5 cm en het gemiddelde op circa 0,5 cm. Dit laatste geeft aan dat er gemiddeld sprake is van een kleine systematische overschatting van de grondwaterstand door de logger. Deze kengetallen zijn echter afgeleid uit verschilwaarden van een groot aantal loggers, die verschillende leeftijden hebben en die ook onder verschillende omstandigheden functioneren, zodat ze ook zullen verschillen in kwaliteit van de ruwe loggermeting. De oudere loggers bijvoorbeeld zullen in veel gevallen drift vertonen, wat de standaardafwijking van de verschilwaarde L0-CH0 zal hebben vergroot. Dit alles beperkt de bruikbaarheid van de afgeleide kengetallen.

De kengetallen van de meetfout van een ruwe loggermeting kunnen bruikbaarder worden als we ze uitsplitsen naar homogenere deelverzamelingen loggers. In tabel 3.3 zijn de kengetallen van de meetfout van een ruwe meting van een logger van leeftijd 0 jaar afgeleid. We kunnen die kengetallen dan min of meer representatief achten voor de fase dat de logger nog geen kwaliteitsverslechtering vertoont (zoals door drift). Volgens de afleiding in tabel 3.3 wordt de standaardafwijking van de ruwe meetfout van een logger van leeftijd 0 jaar geraamd op minder dan 1,4 cm en het gemiddelde op circa 0,6 cm. Deze geraamde standaardafwijking van de meetfout is al duidelijk lager dan die van de heterogene verzameling loggers, wat zal komen doordat er bij leeftijd 0 jaar nog geen sprake zal zijn van sterke drift.

Tabel 3.3: Afleiding van kengetallen van de meetfout van een ruwe meting van een logger van leeftijd 0 jaar (L0, blauw gekleurd, laatste kolom). De kengetallen van L0-CH0 zijn afgeleid in § 4.2.1 en de standaardafwijking betreft vermoedelijk een overschatting, zodat de hier afgeleide standaard-afwijking van de meetfout dat ook zal zijn.

Kengetal	verschil L0-CH0 [1)]	meetfout CH0 [2)]	meetfout L0 [3)]
Gemiddelde (cm)	**0,06**	**0,50**	**0,56** [3)]
Standaardafwijking (cm)	**<1,63**	**0,84**	**<1,40** [4)]

L0 : ruwe loggermeting (loggerleeftijd 0 jaar) | **CH0** : controlehandmeting
1) empirische raming (deze studie)
2) empirische raming [Knotters et al., 2013]
3) volgt uit toepassing formule $\bar{x}_{L0-CH0} = \bar{e}_{L0} - \bar{e}_{CH0}$
4) volgt uit toepassing formule $s^2_{L0-CH0} = s^2_{e,L0} + s^2_{e,CH0}$

$\bar{x}_{L0-CH0}$	gemiddelde van L0-CH0
$\bar{e}_{L0}$ resp. $\bar{e}_{CH0}$	gemiddelde meetfout L0 resp. CH0
s^2_{L0-CH0}	variantie van L0-CH0
$s^2_{e,L0}$ resp. $s^2_{e,CH0}$	variantie meetfout L0 resp. CH0

Ten slotte zijn in tabel 3.4 de kengetallen afgeleid van de meetfout van een logger, ná validatie en zonodige correctie.

Tabel 3.4: Afleiding van kengetallen van de meetfout van een logger, ná validatie en zonodige correctie (L1, blauw gekleurd, laatste kolom).

Kengetal	verschil L1-CH0 [1)]	meetfout CH0 [2)]	meetfout L1 [3)]
Gemiddelde (cm)	**0,01**	**0,50**	**0,51** [3)]
Standaardafwijking (cm)	**0,94**	**0,84**	**>0,42** [4)]

L1 : gevalideerde loggermeting | **CH0** : controlehandmeting
1) empirische raming (deze studie)
2) empirische raming [Knotters et al., 2013]
3) volgt uit formule $\bar{x}_{L1-CH0} = \bar{e}_{L1} - \bar{e}_{CH0}$
4) volgt uit formule $s^2_{L1-CH0} = s^2_{e,L1} + s^2_{e,CH0} - 2 \cdot \mathrm{Cov}(e_{L1}, e_{CH0})$

$\bar{x}_{L1-CH0}$	gemiddelde van L1-CH0
$\bar{e}_{L1}$ resp. $\bar{e}_{CH0}$	gemiddelde meetfout L1 resp. CH0
s^2_{L1-CH0}	variantie van L1-CH0
$s^2_{e,L1}$ resp. $s^2_{e,CH0}$	variantie meetfout L1 resp. CH0
$\mathrm{Cov}(e_{L1}, e_{CHO})$	covariantie meetfout L1 en CH0

Volgens de afleiding in tabel 3.4 wordt de standaardafwijking van de meetfout van een logger ná validatie en zonodige correctie geraamd op minstens 0,4 cm en het gemiddelde op circa 0,5 cm. Het is niet mogelijk de standaardafwijking nauwkeuriger te herleiden. Dit komt doordat de meetfout van L1 door de validatie (met zonodig correctie) gecorreleerd is geraakt aan de meetfout van CH0, zodat we ook de covariantie van de meetfouten van L1 en CH0 nodig hebben bij de berekening. Die covariantie is echter niet bekend.
Om de standaardafwijking van de meetfout van gevalideerde loggermetingen (L1) zuiver te kunnen vaststellen, zal eerst een steekproef moeten worden verzameld van handmetingen die niet worden gebruikt om loggermetingen te valideren. Dit advies is verder uitgewerkt in hoofdstuk 5.

3.5 Kernpunten van dit hoofdstuk

- Langs empirische weg zijn de kenmerken afgeleid van de meetfout van de logger, op basis van de verschillen tussen loggermetingen en daaraan gepaarde controlehandmetingen, waarbij aansluitend de meetfout van de handmeting is verdisconteerd.
- Er is onderscheid gemaakt tussen drie soorten loggermeting, namelijk de ruwe meting (L0), de meting ná validatie en zonodige correctie (L1) en de controlemeting (CL0).
- Elk van de verzamelingen verschilwaarden bevat flinke uitschieters (van meerdere decimeters), zowel naar boven als naar beneden, maar de geschatte populatiemediaan is vrijwel 0 voor elk.
- De bulk van elke verdeling is relatief smal (90% of meer van de verschilwaarden bevindt zich in het interval van -3 en +3 cm).
- De standaardafwijking is het grootst voor L0-CH0 (3,63 cm) en het kleinst voor L1-CH0 (0,94 cm). Die van CL0-CH0 zit daar tussenin (2,25). Deze volgorde stemt overeen met de verwachting.
- Elk van de drie verdelingen wordt gekenmerkt door een sterke piek van het centrum en lange, maar zeer dunne staarten. Geen van de drie soorten verschilwaarde vertoont daarom een normale kansverdeling.
- Om uit de kenmerken van het verschil tussen loggermeting en controlehandmeting de kenmerken van de meetfout af te leiden, is verdisconteerd dat ook de controlehandmeting onderhevig is aan een meetfout. Volgens een studie op basis van een experimentele proefopzet kan de meetfout van die aanpak worden benaderd met een normale verdeling, met gemiddelde 0,5 cm en standaardafwijking 0,84 cm.
- De standaardafwijking van de meetfout van de ruwe loggermeting wordt dan geraamd op circa 3,5 cm en het gemiddelde op circa 0,5 cm. Deze kengetallen hebben echter betrekking op een verzameling loggers die heterogeen is voor wat betreft de kwaliteit van de loggermeting.
- De kengetallen kunnen bruikbaarder worden als we ze uitsplitsen naar homogenere deelverzamelingen loggers. De standaardafwijking van de ruwe meetfout van een logger van leeftijd 0 jaar wordt geraamd op minder dan 1,4 cm en het gemiddelde op circa 0,6 cm. Deze geraamde standaardafwijking van de meetfout is al duidelijk lager dan die van de heterogene verzameling loggers, wat zal komen doordat er bij leeftijd 0 jaar nog geen sprake zal zijn van sterke drift.
- De standaardafwijking van de meetfout van een logger ná validatie en zonodige correctie wordt geraamd op minstens 0,4 cm en het gemiddelde op circa 0,5 cm. Het is niet mogelijk de standaardafwijking nauwkeuriger te herleiden dan deze ondergrens. Dit komt doordat de meetfout van L1 door de validatie (met zonodig correctie) gecorreleerd is geraakt aan de meetfout van CH0, zodat we ook de covariantie van de meetfouten van L1 en CH0 nodig hebben bij de berekening. Die covariantie is echter niet bekend.

4 Analyse invloedsfactoren meetfout loggers

In dit hoofdstuk gaan we voor een aantal factoren na of deze gerelateerd zijn aan de meetfout van de loggers, of beter gezegd, aan een proxy[25] van de meetfout, namelijk het verschil tussen een ruwe (L0) loggermeting en een daaraan gepaarde handmeting (CH0, genomen binnen het uur van L0). De meetfout is immers niet alleen uit het verschil tussen een loggermeting en een daaraan gepaarde handmeting vast te stellen, doordat de handmeting ook een meetfout heeft (zie daarvoor § 3.4).
Er zijn relaties onderzocht van L0-CH0 met loggerleeftijd, maand, waterhoogte en vijf hydrologische karakteristieken, namelijk gemiddelde grondwaterdiepte, GHG, GLG, GHG-GLG en standaardafwijking van de grondwaterstand. Verder is geanalyseerd op trends van verschillen van controlelogger-metingen (CL0) en daaraan gepaarde handmetingen (CH0). De tot dusverre opgebouwde reeksen L0-CH0 zijn daar namelijk te kort voor (<5 jaar). We starten dit hoofdstuk echter met analyses van levensduur en storingsfrequentie van de loggers.

4.1 Levensduur en storingsfrequentie loggers

Uit de verstrekte data blijkt dat er over de beschouwde periode (20 maart 2003 tot 30 januari 2018) gebruik is gemaakt van 1.165 loggers voor de 549 peilfilters waar we informatie over ontvingen. Dit betrof de 549 initieel geïnstalleerde loggers en 616 vervangingen daarvan. Aan de hand van de verstrekte gegevens over datum en reden van loggervervangingen is de verdeling van de levensduur van de loggers vastgesteld. De levensduur is daarbij gedefinieerd als de loggerleeftijd bij vervanging vanwege een storing die niet is veroorzaakt door externe factoren (zoals werkzaamheden). In 380 gevallen is een logger vervangen zonder dat deze stoorde. Dit betrof provisorische plaatsingen, tot er een logger met kabellengte op maat beschikbaar was. Als een dergelijke logger later weer ergens is geplaatst, telt de periode dat deze niet geïnstalleerd was niet mee voor zijn leeftijd.

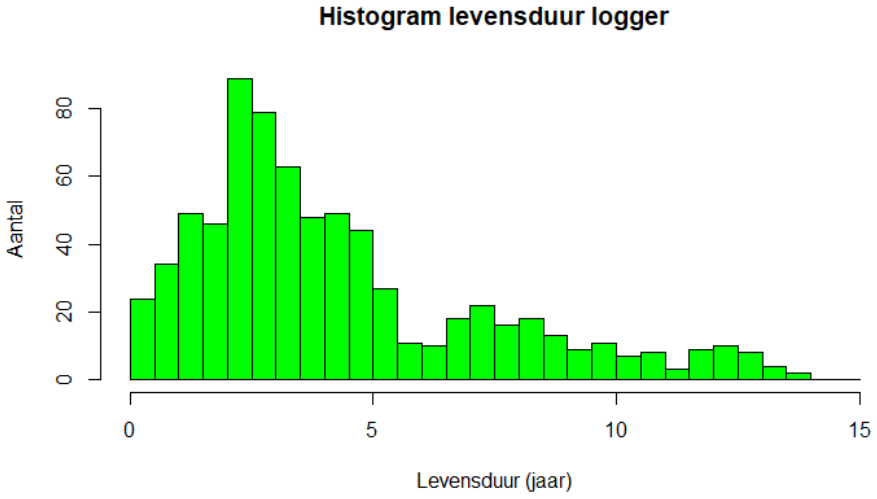
Figuur 4.1 toont de afgeleide verdeling van de levensduur, in de vorm van een histogram en tabel 4.1 geeft kenmerken van de cumulatieve verdeling van de levensduur.

Uit het histogram blijkt dat de levensduurverdeling niet symmetrisch, maar scheef naar rechts is. De mediane levensduur is dientengevolge lager dan de gemiddelde levensduur (mediaan bedraagt 3,37 jaar en gemiddelde bedraagt 4,28 jaar).

Uit tabel 4.1 blijkt dat in het eerste jaar 7,9% van de loggers is vervangen vanwege een storing, na drie jaar 43,9% en na vijf jaar 71,8%. De langste levensduur (bij vervanging) bedroeg 13,8 jaar.

[25] Een proxy is een variabele die gebruikt wordt om informatie te vergaren over een andere, aan de proxy gerelateerde variabele, die niet of niet eenvoudig meetbaar is.

Figuur 4.1: Histogram van de levensduur van een logger (leeftijd bij vervanging vanwege een storing die niet is veroorzaakt door externe factoren). De horizontale as heeft klassen van 0,5 jaar.



Tabel 4.1: Kenmerken van de cumulatieve verdeling van de levensduur van een logger (leeftijd bij vervanging vanwege een storing die niet is veroorzaakt door externe factoren).

Levensduur (jaar)	Aantal (-)	Relatief (%)
<0,5	24	3,3%
<1	58	7,9%
<2	153	20,9%
<3	321	43,9%
<4	432	59,1%
<5	525	71,8%
<6	563	77,0%
<7	591	80,8%

Levensduur (jaar)	Aantal (-)	Relatief (%)
<8	629	86,0%
<9	660	90,3%
<10	680	93,0%
<11	695	95,1%
<12	707	96,7%
<13	725	99,2%
<14	731	100,0%

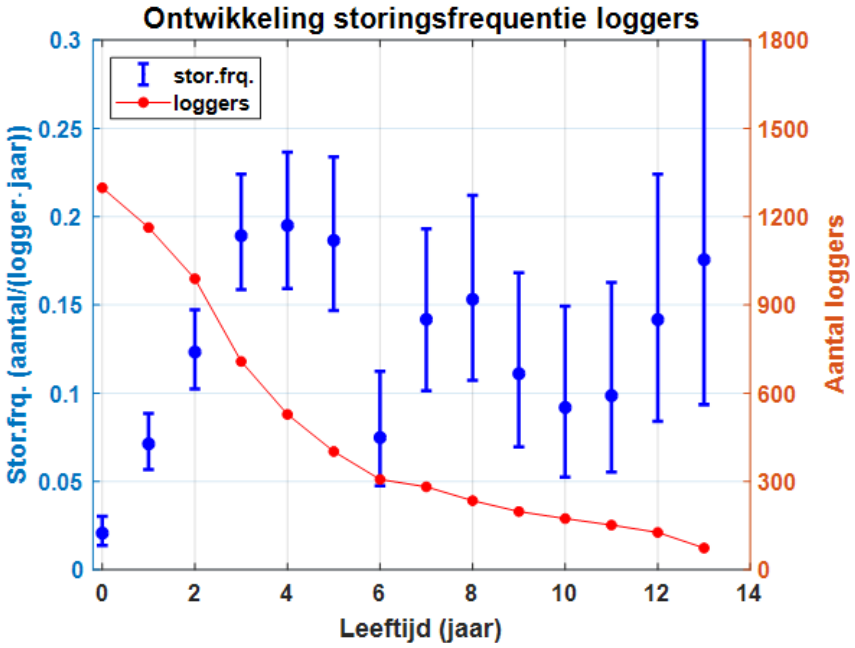
Met de ons verstrekte gegevens over datum en reden van loggervervangingen hebben we ook de ontwikkeling van de storingsfrequentie als functie van de leeftijd gereconstrueerd (zie figuur 4.2). De storingsfrequentie is daarbij per leeftijd afgeleid als:

$$sf_q = \frac{\sum_{j=2004}^{2017} s_{q,j}}{\sum_{j=2004}^{2017} n_{q,j}} \qquad [4.1]$$

waarin sf_q de storingsfrequentie (aantal/(logger·jaar)) van een logger van leeftijd q, $s_{q,j}$ het aantal storingen (-) van een logger van leeftijd q in jaar j van de gehanteerde registratieperiode (die hier loopt van 2004 t/m 2017)[26] en $n_{q,j}$ het totaal aantal loggers van leeftijd q in jaar j.
Vanuit theoretisch oogpunt is het aannemelijk dat het geregistreerd aantal storingen in een jaar afkomstig is uit de Poissonverdeling. Het 95%-betrouwbaarheidsinterval van een geschatte storings-frequentie kan dan worden berekend met de inverse van de chikwadraatverdeling. De betrouw-baarheid van een empirisch bepaalde storingsfrequentie zal toenemen met het aantal storingen en met het aantal loggers. We zien in figuur 4.2 dan ook dat de betrouwbaarste storingsfrequenties voorkomen bij de leeftijden die in de beschouwde periode het meest voorkwamen (0 t/m 2 jaar).

[26] Het jaar 2003 is niet als registratiejaar voor het aantal storingen meegenomen, omdat de (24) loggers in 2003 hoofdzakelijk zijn geplaatst in het najaar. Maar de gegevens van deze loggers tellen vanaf 2004 wel mee.

Figuur 4.2: Ontwikkeling van de storingsfrequentie van loggers, als functie van de leeftijd, zoals afgeleid met storingsgegevens van de periode 2004 t/m 2017. Van elke storingsfrequentie is tevens het 95%-betrouwbaarheidsinterval weergegeven (verticale blauwe lijn). De rode lijn geeft per leeftijd het aantal loggers aan waarover de storingsfrequentie is bepaald.



Uit figuur 4.2 blijkt dat de storingsfrequentie van loggers tot de leeftijd vier jaar duidelijk toeneemt. Vervolgens neemt deze geleidelijk weer af tot de leeftijd tien jaar, waarna de storingsfrequentie weer toeneemt. Er is dus een relatief groot percentage uitvallers van jonge leeftijd, maar daarnaast zijn er ook meerdere overlevers, die aanvankelijk zonder veel problemen ouder worden. Deze laatste zijn blijkbaar robuuster, mogelijk ook omdat ze in een gunstigere omgeving functioneren. Pas na tien jaar beginnen ook deze overlevers steeds meer mankementen te vertonen.

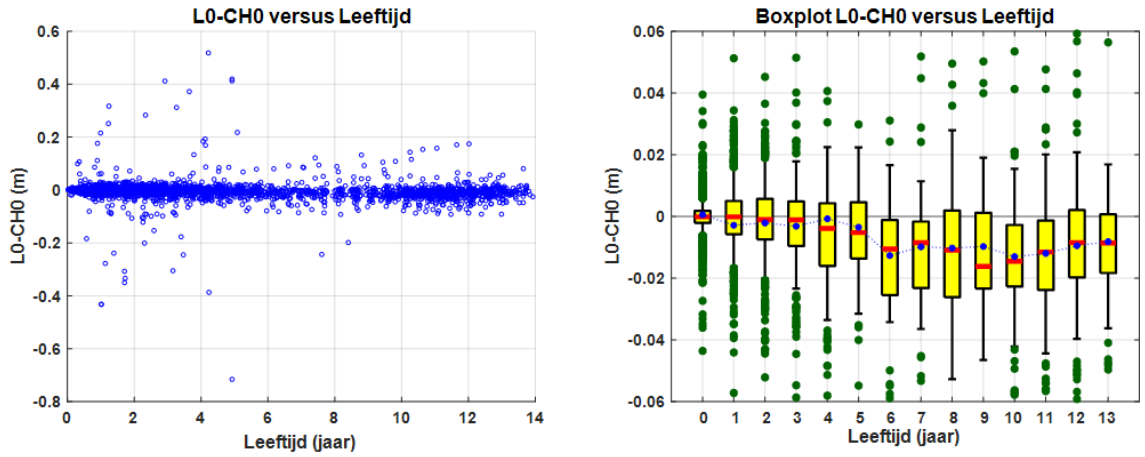
4.2 Relatie leeftijd logger en functioneren logger

In deze paragraaf gaan we op twee manieren na welke relatie er bestaat tussen de leeftijd van de logger en het functioneren van de logger. Eerst wordt visueel beoordeeld of er een relatie is tussen L0-CH0 en de leeftijd van de logger (§ 4.2.1). Daarna wordt statistisch getoetst op trends van de verschillen tussen de controlemetingen CL0 en CH0 (§ 4.2.2). Een dergelijke trendanalyse is nog niet mogelijk voor de verschillen tussen L0 en CH0, doordat die reeksen nog te kort zijn.

4.2.1 Relatie L0-CH0 en leeftijd logger

Ondermeer als er drift optreedt, zal er sprake zijn van een relatie van kenmerken van het verschil tussen L0 en CH0 met de leeftijd van de logger (tijd verstreken sinds de plaatsing). Om dit te kunnen beoordelen toont figuur 4.3 zowel het spreidingsdiagram als boxplots van L0-CH0 versus de leeftijd van de logger.

Figuur 4.3: Spreidingsdiagram en boxplots (uitvergroot) van L0-CH0 versus de leeftijd van de logger.



Toelichting op de boxplot

De boxplot (ook wel aangeduid als box-whisker-plot) is een handzame manier om de kenmerken van een onderzochte populatie zo compact mogelijk grafisch samen te vatten. Deze toont de posities van relevante populatiepercentielen van de beschouwde variabele, zoals geschat uit de meetwaarden. Het middendeel, de box (doos), loopt van het 25-percentiel naar het 75-percentiel, terwijl het 50-percentiel (ook wel de mediaan genoemd) is aangegeven als een horizontale rode streep in de box. De whiskers (snorharen) lopen van de box naar de uiteinden van de steekproef, maar niet verder dan de verste meetwaarde binnen maximaal eenmaal de boxlengte vanaf de box. Eventuele extremere waarden worden met een marker apart weergegeven boven de bovenste whisker of onder de onderste whisker. Verder is in onze boxplots ook het rekenkundig gemiddelde van de meetwaarden weergegeven, als een blauwe punt.

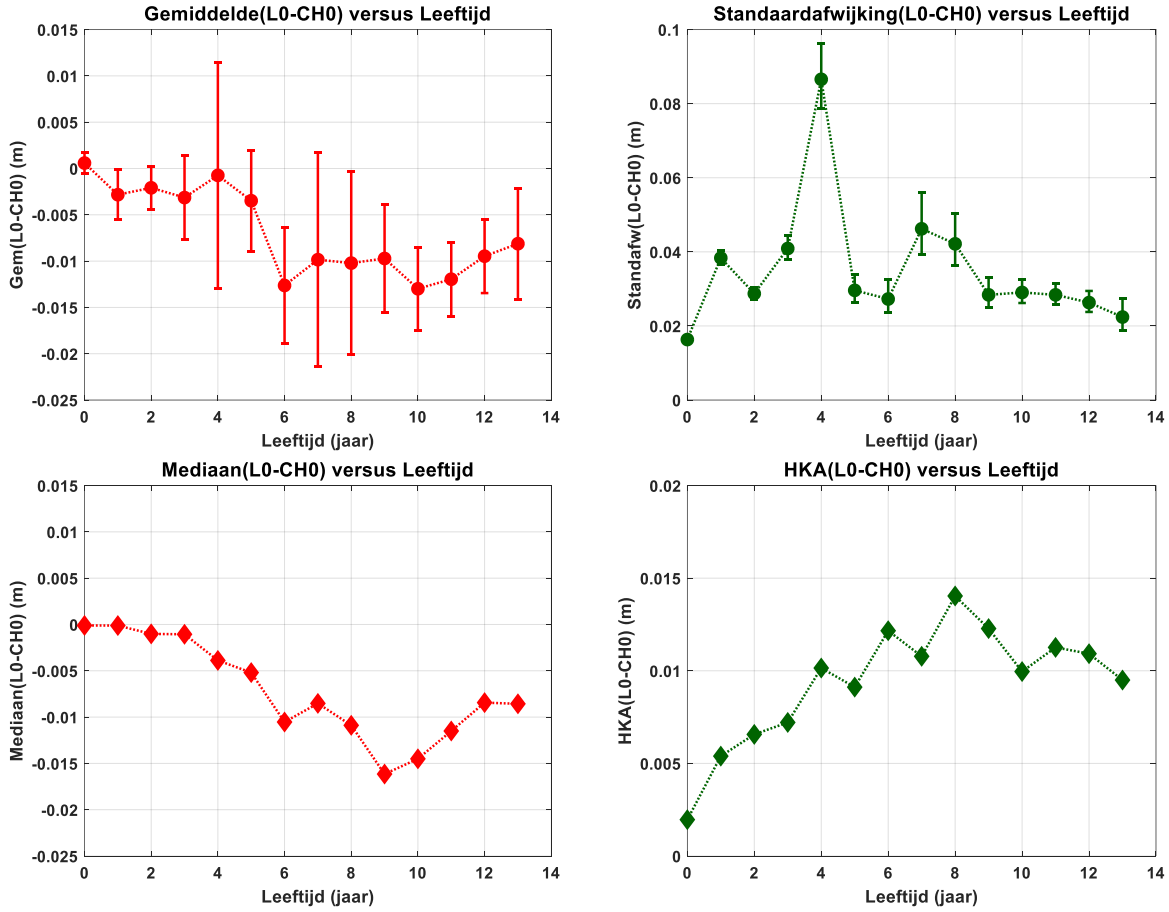
Uit de boxplots van figuur 4.3 blijkt dat zowel het gemiddelde als de mediaan van L0-CH0 licht negatiever wordt met de leeftijd.

We kunnen nog beter beoordelen op relaties van L0-CH0 met de leeftijd als zijn relevante kengetallen zijn uitgezet tegen de leeftijd (zie figuur 4.4). Daarvoor gebruiken we zowel kengetallen voor het centrum als voor de spreiding, namelijk:

- (rekenkundig) gemiddelde: het klassieke en bekendste kengetal voor het centrum van een variabele. In de figuur wordt tevens het 95%-betrouwbaarheidsinterval van zijn schatting weergegeven (bepaald met formule [5.9] van § 5.4), dat geldt in geval de variabele een normale kansverdeling heeft;
- standaardafwijking; het klassieke en bekendste kengetal voor de spreiding van een variabele. In de figuur is tevens het 95%-betrouwbaarheidsinterval van zijn schatting weergegeven (bepaald met formule [5.8] van § 5.4), dat geldt in geval de variabele een normale kansverdeling heeft;
- mediaan: een alternatief en robuuster kengetal voor het centrum van een variabele. Het is een schatting van de waarde die door 50% van de populatie wordt onderschreden en door 50% van de populatie wordt overschreden. De mediaan wordt niet of nauwelijks beïnvloed door uitschieters en is daarom te prefereren boven het gemiddelde als er sprake is van uitschieters;
- HKA (halvekwartielafstand): een alternatief en robuuster kengetal voor de spreiding van een variabele, bepaald als de helft van het verschil tussen het geschatte 75-percentiel en het geschatte 25-percentiel ((P75-P25)/2). De halvekwartielafstand wordt niet of nauwelijks beïnvloed door uitschieters en is daarom te prefereren boven de standaardafwijking als er sprake is van uitschieters.

Aangezien de set L0-CH0 veel en flinke uitschieters omvat, zal deze niet voor elke leeftijd voldoen aan een normale kansverdeling. De weergegeven 95%-betrouwbaarheidsintervallen van gemiddelde en standaardafwijking van deelsets kunnen daardoor vertekend zijn. Maar desondanks kunnen ze helpen om de betrouwbaarheden van kengetallen van deelsets onderling te kunnen vergelijken. Vanwege de uitschieters zullen we in het nu volgende onze oordelen over de relaties van L0-CH0 met bepaalde factoren vooral baseren op de verlopen van de robuuste kengetallen mediaan en HKA.

Figuur 4.4: Gemiddelde (links) en standaardafwijking (rechts) van L0-CH0 per loggerleeftijd, met tevens het 95%-betrouwbaarheidsinterval. Daaronder hun robuuste equivalenten, namelijk mediaan (links) en halvekwartielafstand (HKA, rechts). Ons oordeel baseert zich op de robuuste kengetallen.



In figuur 4.4 zien we dat het centrum van L0-CH0 vrijwel altijd onder nul ligt en tot een leeftijd van circa negen jaar daalt, om daarna weer enigszins te stijgen. De spreiding van L0-CH0 neemt daarentegen duidelijk toe met de leeftijd. Dit laatste zal mede veroorzaakt zijn door drift.

Deze uitsplitsing van kenmerken van de verschilwaarde L0-CH0 naar leeftijd stelt ook in staat de statistische kenmerken te verkennen van het meetproces, representatief voor de fase dat het meetproces nog geen kwaliteitsverslechtering vertoont (zoals door drift). Deze kenmerken kunnen dienen om statistisch onderbouwde correctiecriteria voor het meetproces af te leiden (zie § 2.4). In principe kunnen de kenmerken van L0-CH0 bij leeftijd nul jaar daar een beeld van geven. Er zijn 773 verschilwaarden L0-CH0 van leeftijd nul jaar. Het gemiddelde bedraagt 0,06 cm en de standaardafwijking 1,63 cm. Er is geen sprake van een normale kansverdeling, door een tiental uitschieters buiten het interval [-5 cm ; + 5cm]. Dit kan komen doordat bij de eerste controle- en inleesronde na het installeren van de logger de (nog provisorische) instelling van de loggerdiepte in Menyanthes is

aangepast voor het verschil tussen loggermeting en controlehandmeting (§ 2.2.1), wat extra variatie zal geven. In principe moet daar nog voor worden gecorrigeerd, maar dat voerde hier te ver.

4.2.2 Trends in CL0-CH0

Een andere indicatie van de invloed van de loggerleeftijd op het functioneren van de logger is het optreden van een trend in de meetfout, wat we kunnen opvatten als een manifestatie van drift. Onder trend verstaan we hier een permanente of semipermanente verandering van het niveau van de meetfoutreeks over een tijdshorizon van tenminste enkele jaren. Seizoensmatige veranderingen en kortstondige calamiteiten vallen daar dus niet onder.

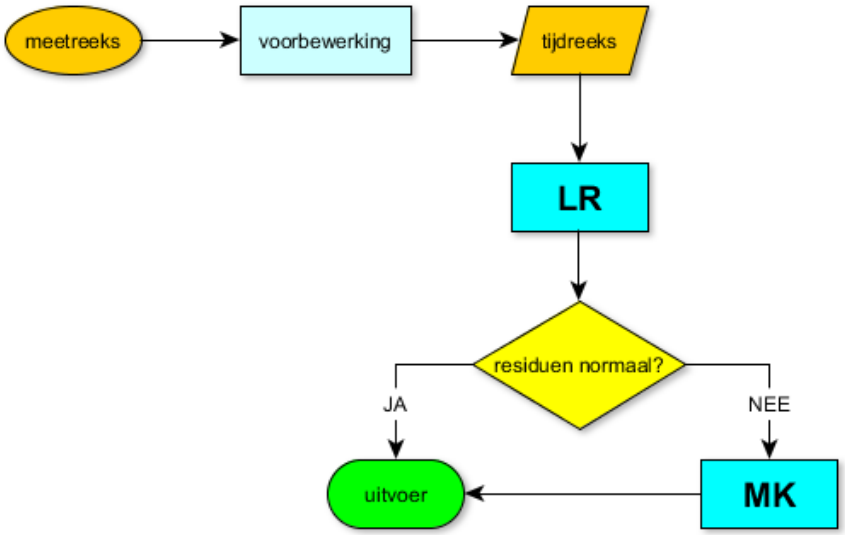
Het bleek niet mogelijk trendanalyse uit te voeren op de verschilreeksen van ruwe loggermetingen (L0) en daaraan gepaarde handmetingen (CH0). De tot dusverre opgebouwde reeksen L0-CH0 zijn daar namelijk te kort voor (<5 jaar). Dit komt doordat een verschilwaarde L0-CH0 alleen bruikbaar is als hun meetmomenten minder dan een uur verschillen, wat pas vanaf eind 2014 / begin 2015 grootschalig begon op te treden, omdat toen overal op uurbasis werd gemeten. Daarom hebben we trendanalyse toegepast op de verschilreeksen van de gepaarde controlemetingen CL0 en CH0. Die zijn namelijk beschikbaar van 2004 – 2018. Een nadeel is wel dat de CL0-waarden kunnen zijn beïnvloed door bij voorgaande uitleesmomenten uitgevoerde aanpassingen van de installatiediepte, bedoeld om ze niet teveel uit de pas te laten lopen met de controlehandmetingen (§ 3.1). Het is daarom te verwachten dat controleloggermetingen (CL0) in het algemeen beter aansluiten op de controlehandmetingen dan ruwe operationele loggermetingen (L0). Dit blijkt ook uit tabel 3.1. We moeten er daarom rekening mee houden dat onze trendanalyse een onderschatting zal opleveren van het optreden van trend (en daarmee drift).

We hebben de trendanalyses beperkt tot maximaal de periode 2004 t/m 2015, omdat vanaf medio 2015 is gestopt met het zonodig aanpassen van de installatiediepte van de logger in de Keller-software, waardoor het meetproces van CL0 vóór en ná medio 2015 niet meer vergelijkbaar is (dit zal de resultaten van de analyse moeilijker te interpreteren maken).
De trendanalyses zijn uitgevoerd op alle loggerreeksen (uniek betreffende loggernummer en combinatie peilbuis en filter) die voldoen aan de criteria dat de reeks minimaal vijf verschilwaarden CL0-CH0 omvat en minimaal vijf jaar lang is. Van de 1.526 loggerreeksen bleken er 283 te voldoen aan deze criteria. Bij de trendanalyse is getoetst op een monotone trend, dat wil zeggen op een overwegende daling of een overwegende stijging vanaf het startpunt van de reeks, ongeacht of dit in de vorm van een lineaire, convexe, of concave verandering is. Er is tweezijdig getoetst, dus op een verandering, ongeacht of het een daling of een stijging betreft. Daarbij is 95% betrouwbaarheid gehanteerd.

De reeksen bevatten doorgaans niet meer dan twee waarden per jaar, zodat er geen rekening hoefde te worden gehouden met autocorrelatie. En omdat er hooguit sprake is van geringe seizoen-effecten, kon worden volstaan met de lineaire regressietoets of de Mann-Kendalltoets. Om tot een passende keuze te komen is van elke meetreeks eerst een lineair regressiemodel geschat, waarbij de tijd de verklarende variabele is. Als vervolgens uit statistische toetsing (met de Lillieforstoets op normaliteit) bleek dat de modelresiduen afkomstig zijn uit een normale kansverdeling, is voor die meetreeks het resultaat van de lineaire regressietoets op trend genomen. Anders is de verdelings-vrije Mann-Kendalltoets toegepast. De aanpak is weergegeven in het stroomschema van figuur 4.5.

Als een meetreeks op trend is getoetst met de lineaire regressietoets, dan is de trend geschat als de lineaire regressiehelling en als is getoetst met de Mann-Kendalltoets, dan is de trend geschat als de Theilhelling [Theil, 1950; Sen, 1968]. Deze laatste is de mediaan van alle individuele hellingen tussen de afzonderlijke waarden in de tijdreeks en daarmee op te vatten als een verdelingsvrije schatter, die resistent is tegen de invloed van extremen.

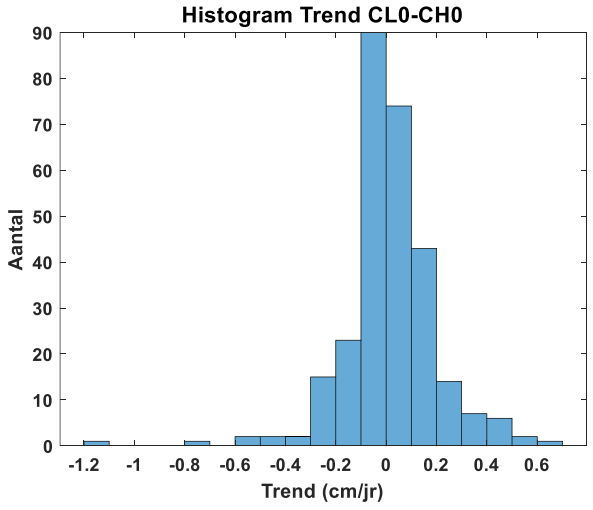
Figuur 4.5: Stroomschema van de hier toegepaste procedure voor trendanalyse.



LR = lineaire regressietoets - **MK** = Mann-Kendalltoets

De resultaten van de trendanalyse zijn samengevat in figuur 4.6, in de vorm van een histogram van de trendhellingen en de kenmerken van de trendhellingen en de statistisch significante trends.

Figuur 4.6: Histogram van de 283 bij trendanalyse geschatte trendhellingen van CL0-CH0 en de kenmerken van de trendhellingen. De laatste twee kolommen geven de kenmerken na verwijderen, respectievelijk op nul zetten van de statistisch niet-significante trendhellingen (Helling SNS).



Kenmerk	Helling	Helling SNS weg	Helling SNS => 0
Aantal	283	52	283
0	8	0	231
≠ 0	275	52	52 (18,4%)
>0	139	32	32
<0	136	20	20
Min (cm/j)	-1,10	-1,10	-1,10
Mediaan (cm/j)	0,00	0,12	0,00
Gemidd (cm/j)	0,02	0,04	0,01
Max (cm/j)	0,61	0,58	0,58
Standafw (cm/j)	0,18	0,33	0,14

SNS: Statistisch niet-significant

Uit de trendanalyse bleek dat 52 (18,4%) van de 283 geanalyseerde reeksen CL0-CH0 een statistisch significante trend vertoonden, met enigszins meer stijgingen (32) dan dalingen (20). De 52 betref-

fende trendplots zijn te zien in de bijlage van dit rapport. De verdeling van de 283 geschatte trendhellingen is vrij symmetrisch, met - ongeacht statistische significantie - 139 positieve hellingen, 8 nulhellingen en 136 negatieve hellingen en met mediaan 0,00 cm/jaar en gemiddelde 0,02 cm/jaar. De maximale statistisch significante trend bedraagt 0,58 cm/jaar en de minimale bedraagt -1,10 cm/jaar.

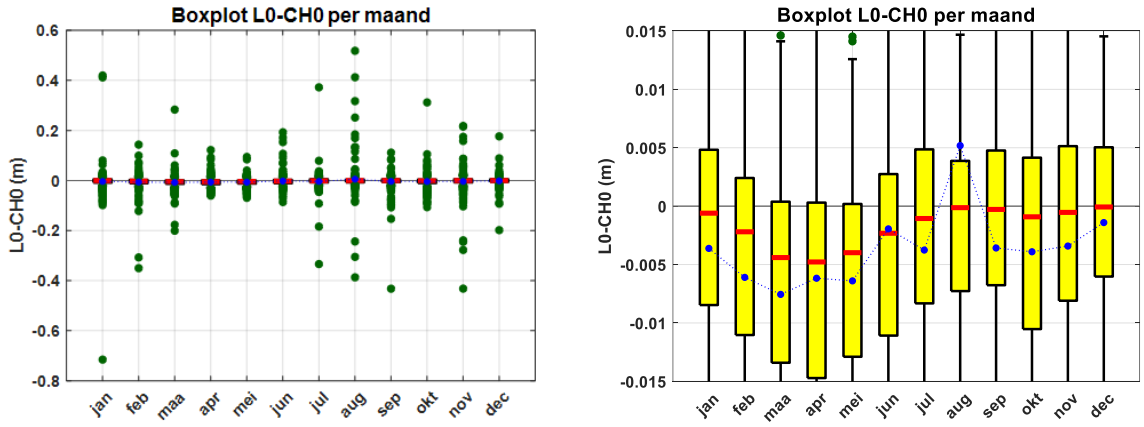
Als we alle statistisch niet-significante trendhellingen op nul zetten komt de standaardafwijking van de trendhelling uit op 0,14 cm/j. Dit is op te vatten als een statistische maat voor de drift die zich tot medio 2015 manifesteerde in het meetproces van CL0, ondanks de regelmatige aanpassingen van de installatiediepte, die juist bedoeld waren om drift tegen te gaan.

Het feit dat ruim 18% van de geanalyseerde reeksen CL0-CH0 een statistisch significante trend vertoont, duidt op redelijk vaak optredende drift. Aangezien dit nog een onderschatting zal zijn van het werkelijke percentage drift (zie eerder), onderstreept dit de noodzaak om loggermetingen regelmatig te valideren en zonodig te corrigeren. Zonder een strak validatieregiem zullen de metingen immers te onbetrouwbaar worden.

4.3 Relatie L0-CH0 en maand

Om na te kunnen gaan of het verschil tussen de ruwe loggermeting (L0) en de controlehandmeting (CH0) al dan niet afhangt van het seizoen toont figuur 4.7 voor elke maand de boxplot van het verschil.

Figuur 4.7: Boxplots op maandbasis van L0-CH0. De rechterfiguur is een uitvergroting.



Uit deze boxplots blijkt dat de mediaan van L0-CH0 in de lente doorgaans het laagst is (tot -5 mm) en in de tweede jaarhelft het hoogst (tot -1 mm).

We hebben ook statistisch getoetst op verschillen tussen de maandpopulaties van L0-CH0, met de Kruskal-Wallis-toets. Dit is het verdelingsvrije equivalent van eenweg-variantieanalyse (*oneway ANOVA*). De p-waarde van deze toets bleek <0,001, zodat de nulhypothese kon worden verworpen. Om vervolgens na te gaan wélke maanden verschillen in de populatie CL0-CH0 zijn alle maanden getoetst op onderlinge verschillen, met correctie voor het vergroten van het algemene risico op onterechte detectie van een verschil bij meervoudig toetsen. De toetsresultaten (*p*-waarden) zijn in

matrixvorm samengevat in tabel 4.2. Het blijkt dat er (bij een betrouwbaarheid van 95%) statistisch significante verschillen zijn tussen voorjaar (maart t/m mei) en tweede jaarhelft (juli t/m januari).

Tabel 4.2: De toetsresultaten (p-waarden) van het onderling toetsen op verschillen tussen de populaties L0-CH0 van maanden.

Maand	1	2	3	4	5	6	7	8	9	10	11
2	0,747										
3	0,000	0,392									
4	0,001	0,557	1,000								
5	0,001	0,657	1,000	1,000							
6	0,779	1,000	0,339	0,502	0,601						
7	1,000	0,949	0,015	0,033	0,046	0,959					
8	1,000	0,474	0,000	0,001	0,002	0,505	1,000				
9	0,994	0,151	0,000	0,000	0,000	0,168	0,998	1,000			
10	0,998	0,995	0,007	0,023	0,032	0,997	1,000	0,929	0,656		
11	1,000	0,305	0,000	0,000	0,000	0,335	1,000	1,000	1,000	0,896	
12	0,950	0,096	0,000	0,000	0,000	0,107	0,980	1,000	1,000	0,473	0,997

geen statistisch significant verschil (p>0,05)
statistisch significant verschil (p≤0,05)

Tabel 4.3 vermeldt per maand een aantal geschatte kengetallen van de populatie L0-CH0, waarbij per kengetal de cellen conditioneel zijn gekleurd op basis van de rangorde. Hieruit blijkt dat de tegenstelling lente – tweede jaarhelft behalve voor de mediaan ook min of meer opgaat voor het 25- en het 75-percentiel. De standaardafwijking van L0-CH0 lijkt nauwelijks onderhevig aan een seizoeneffect, maar dit kengetal is minder bruikbaar, doordat het sterk wordt beïnvloed door uitschieters.

Tabel 4.3: Geschatte kengetallen van de populatie L0-CH0 per maand, waarbij per kengetal de cellen conditioneel zijn gekleurd op basis van de rangorde.

	Kengetal van L0-CH0 (cm)				
Maand	P25	Mediaan	P75	Gemidd	Standafw
jan	-0,85	-0,06	0,48	-0,36	5,08
feb	-1,10	-0,22	0,24	-0,61	3,56
maa	-1,34	-0,44	0,04	-0,76	2,81
apr	-1,47	-0,48	0,03	-0,62	2,01
mei	-1,29	-0,40	0,02	-0,64	1,65
jun	-1,11	-0,23	0,27	-0,19	2,59
jul	-0,83	-0,11	0,49	-0,38	4,42
aug	-0,73	-0,01	0,39	0,52	7,78
sep	-0,68	-0,03	0,48	-0,36	3,21
okt	-1,05	-0,09	0,42	-0,39	2,55
nov	-0,81	-0,05	0,51	-0,34	3,74
dec	-0,60	-0,01	0,50	-0,14	2,48

hoogste waarde in kolom
laagste waarde in kolom

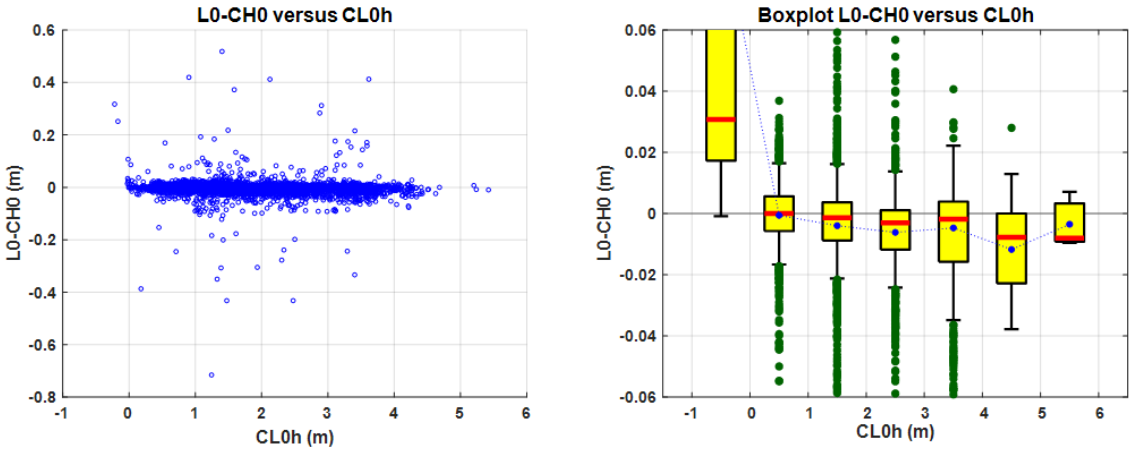
P25 is het 25-percentiel, Mediaan is het 50-percentiel en P75 is het 75-percentiel

4.4 Relatie L0-CH0 en waterhoogte ten opzichte van druksensor

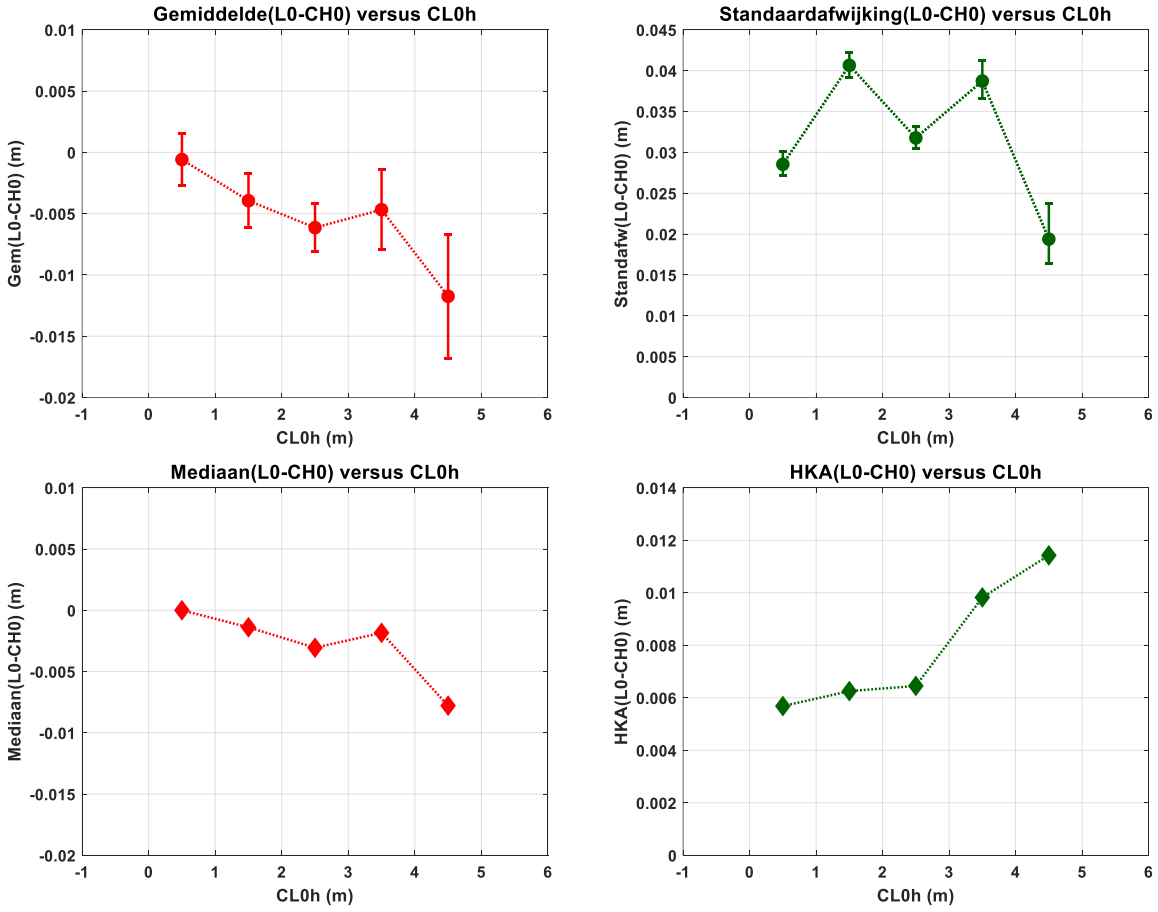
Uit fysische overwegingen is het denkbaar dat de meetfout van de logger gerelateerd is aan de gemeten waterhoogte ten opzichte van de druksensor (zie bijvoorbeeld de tabellen 2.3 en 2.4). Om dit te kunnen beoordelen is het verschil tussen L0 en CH0 in de figuren 4.8 en 4.9 op verschillende

wijzen grafisch uitgezet tegen de gemeten waterhoogte van de bijbehorende controleloggermeting (CL0) ten opzichte van de druksensor (CL0h). Deze laatste zal min of meer te vergelijken zijn met de waterhoogte van L0 ten opzichte van de druksensor (L0h) en was bovendien makkelijk beschikbaar, in tegenstelling tot L0h.

Figuur 4.8: Spreidingsdiagram en boxplots (uitvergroot) van L0-CH0 versus CL0h.



Figuur 4.9: Gemiddelde (links) en standaardafwijking (rechts) van L0-CH0 per meterklasse van CL0h, elk met tevens het 95%-betrouwbaarheidsinterval. Daaronder hun robuuste equivalenten, namelijk mediaan (links) en halvekwartielafstand (HKA, rechts). Ons oordeel baseert zich op de robuuste kengetallen.



Uit figuren 4.8 en 4.9 blijkt dat zowel het centrum als de spreiding van L0-CH0 monotoon gerelateerd is aan CL0h. Als de waterhoogte van CL0 ten opzichte van de druksensor toeneemt wordt het centrum van L0-CH0 negatiever, terwijl de spreiding van L0-CH0 toeneemt. Aangezien het denkbaar is dat dit causale relaties zijn, is het aan te bevelen de loggerdiepte zo beperkt mogelijk te houden, uiteraard wel rekening houdende met de ter plekke geldende kansverdeling van de grondwaterstand.

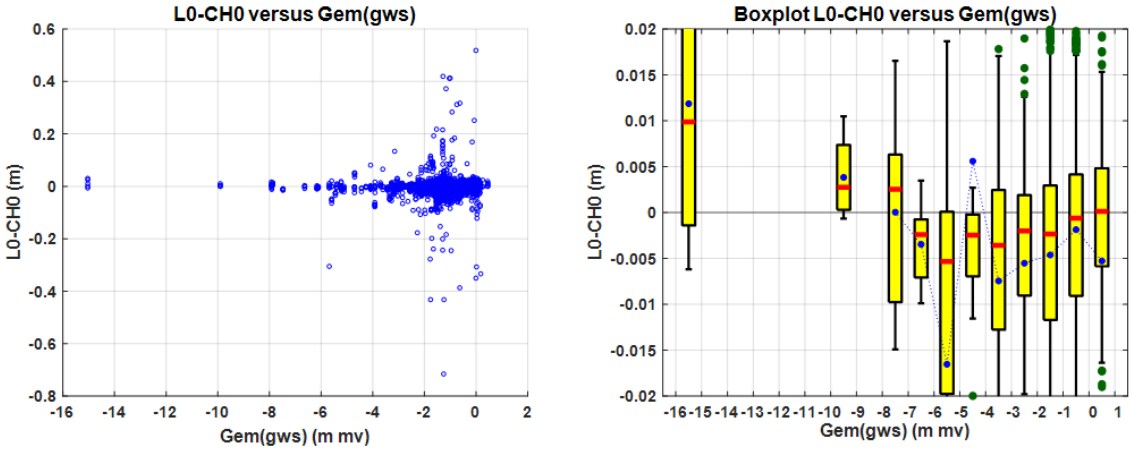
4.5 Relatie L0-CH0 en hydrologische condities meetlocatie

Ten slotte is nagegaan of het verschil tussen L0 en CH0 wellicht gerelateerd is aan hydrologische karakteristieken van de meetlocatie. We zijn dit met grafische analyses nagegaan voor gemiddelde grondwaterstand ten opzichte van maaiveld, GHG, GLG, GHG-GLG en standaardafwijking van de grondwaterstand.

N.B: De in deze paragraaf weergegeven betrouwbaarheidsintervallen van gemiddelde en standaardafwijking van L0-CH0 per klasse zijn nóg minder indicatief dan die in de voorgaande figuren, omdat er nu – behalve het reeds geconstateerde niet voldoen aan normaliteit – ook geen sprake zal zijn van onafhankelijke meetwaarden in een klasse (alle verschilwaarden L0-CH0 van een peilfilter vallen immers in dezelfde klasse).

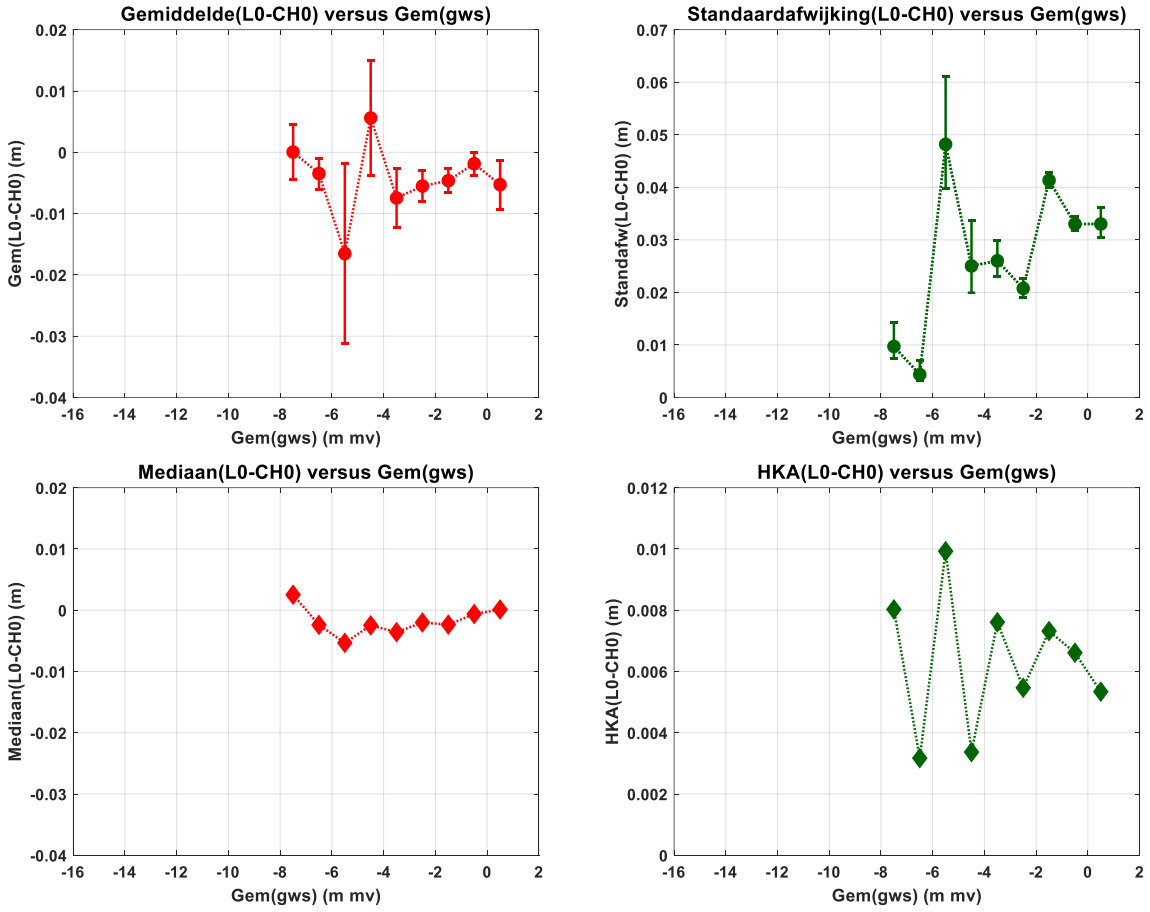
Relatie L0-CH0 en gemiddelde grondwaterstand ten opzichte van maaiveld

Figuur 4.10: Spreidingsdiagram en boxplots (uitvergroot) van L0-CH0 versus de gemiddelde grondwaterstand ten opzichte van maaiveld.



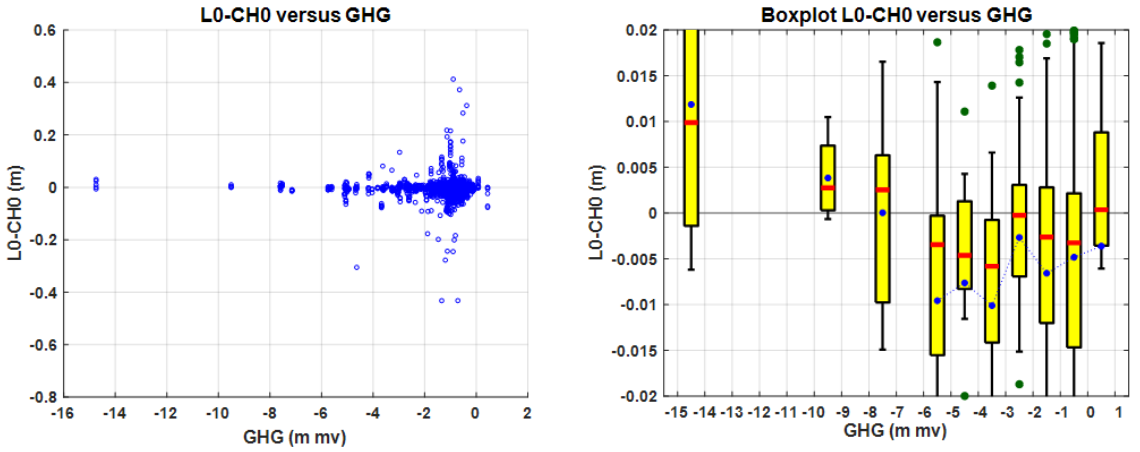
Uit de figuren 4.10 en 4.11 blijkt niet dat er sprake is van een duidelijke relatie tussen L0-CH0 en de gemiddelde grondwaterstand ten opzichte van maaiveld.

Figuur 4.11: Gemiddelde (links) en standaardafwijking (rechts) van L0-CH0 per meterklasse van de gemiddelde grondwaterstand ten opzichte van maaiveld, elk met tevens het 95%-betrouwbaarheids-interval. Daaronder hun robuuste equivalenten, namelijk mediaan (links) en halvekwartielafstand (HKA, rechts). Ons oordeel baseert zich op de robuuste kengetallen.

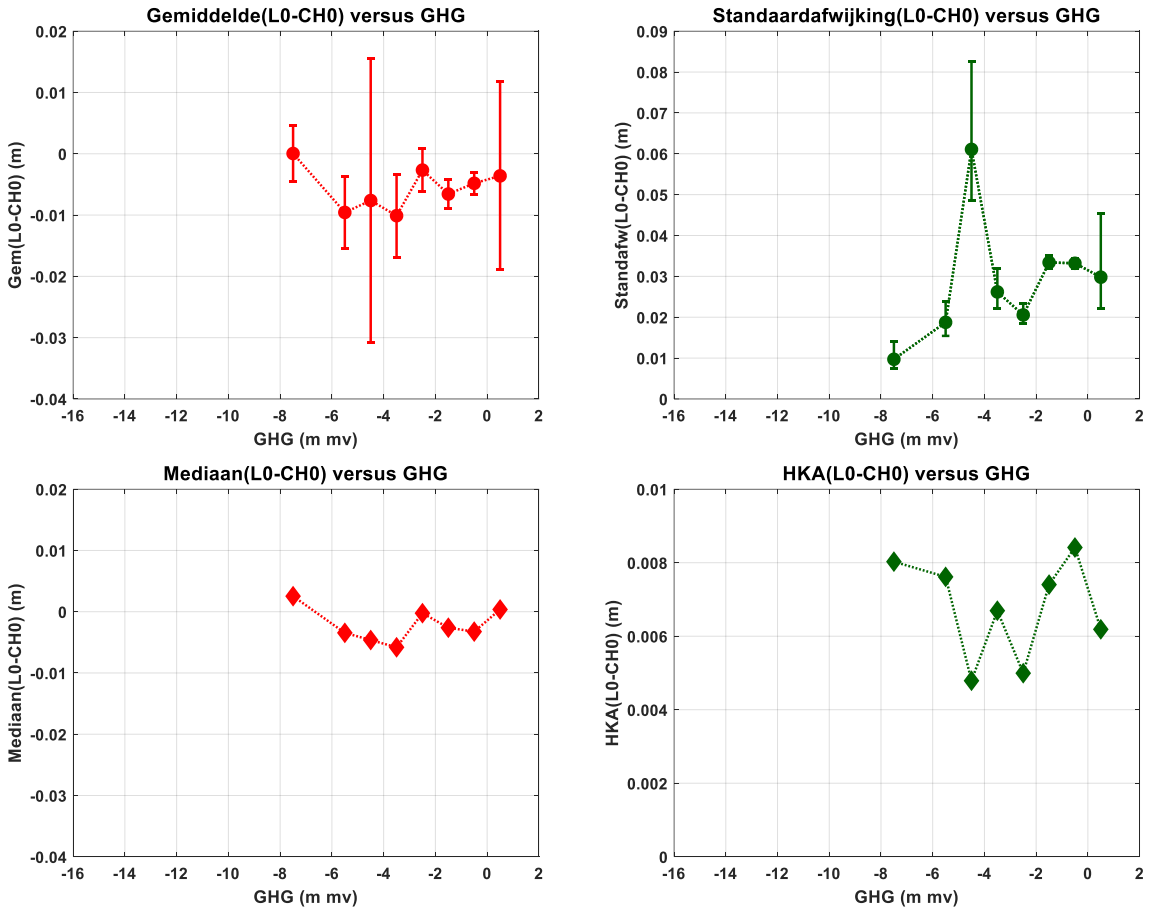


Relatie L0-CH0 en GHG

Figuur 4.12: Spreidingsdiagram en boxplots (uitvergroot) van L0-CH0 versus de gemiddelde hoogste grondwaterstand (GHG) ten opzichte van maaiveld.



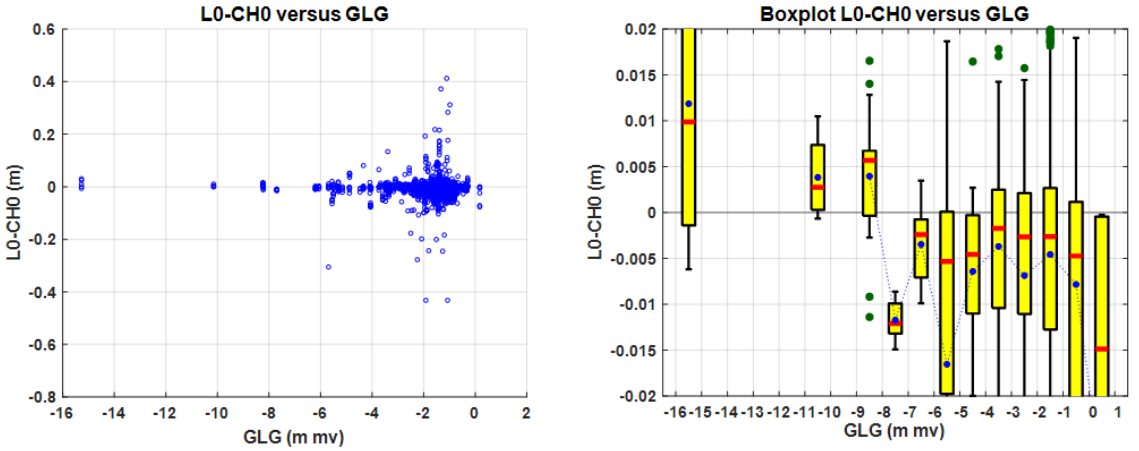
Figuur 4.13: Gemiddelde (links) en standaardafwijking (rechts) van L0-CH0 per meterklasse van de GHG, elk met tevens het 95%-betrouwbaarheidsinterval. Daaronder hun robuuste equivalenten, namelijk mediaan (links) en halvekwartielafstand (HKA, rechts). Ons oordeel baseert zich op de robuuste kengetallen.



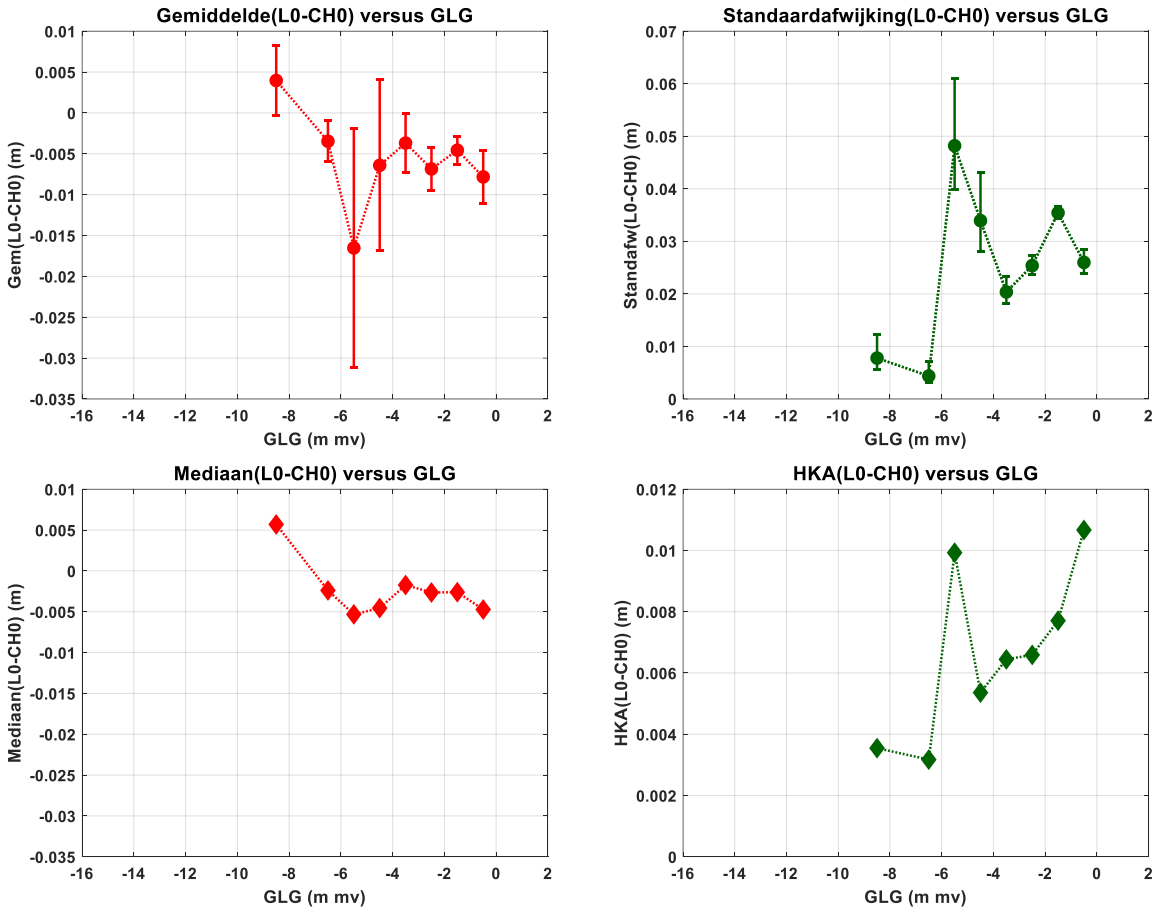
Uit bovenstaande figuren blijkt niet dat er sprake is van een duidelijke relatie tussen L0-CH0 en de GHG.

Relatie L0-CH0 en GLG

Figuur 4.14: Spreidingsdiagram en boxplots (uitvergroot) van L0-CH0 versus de gemiddelde laagste grondwaterstand (GLG) ten opzichte van maaiveld.



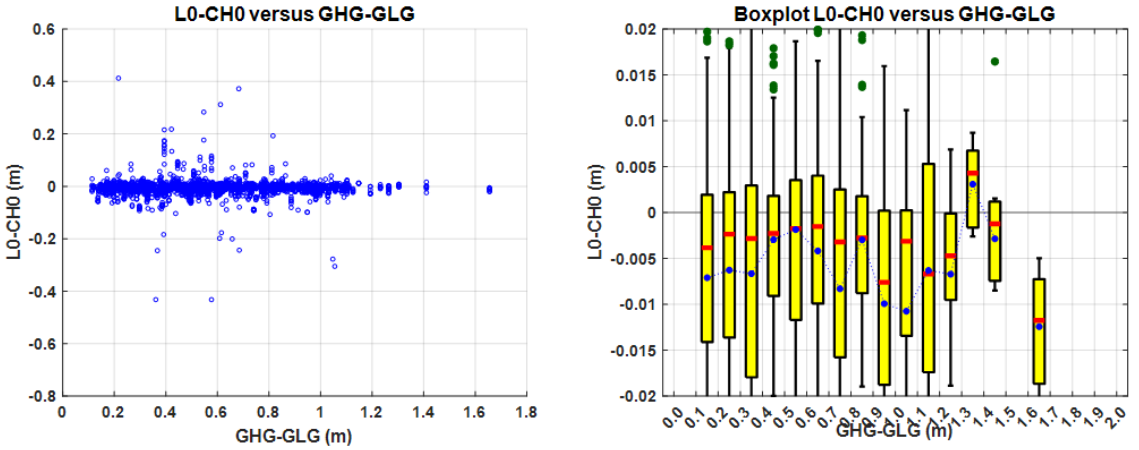
Figuur 4.15: Gemiddelde (links) en standaardafwijking (rechts) van L0-CH0 per meterklasse van de GLG, elk met tevens het 95%-betrouwbaarheidsinterval. Daaronder hun robuuste equivalenten, namelijk mediaan (links) en halvekwartielafstand (HKA, rechts). Ons oordeel baseert zich op de robuuste kengetallen.



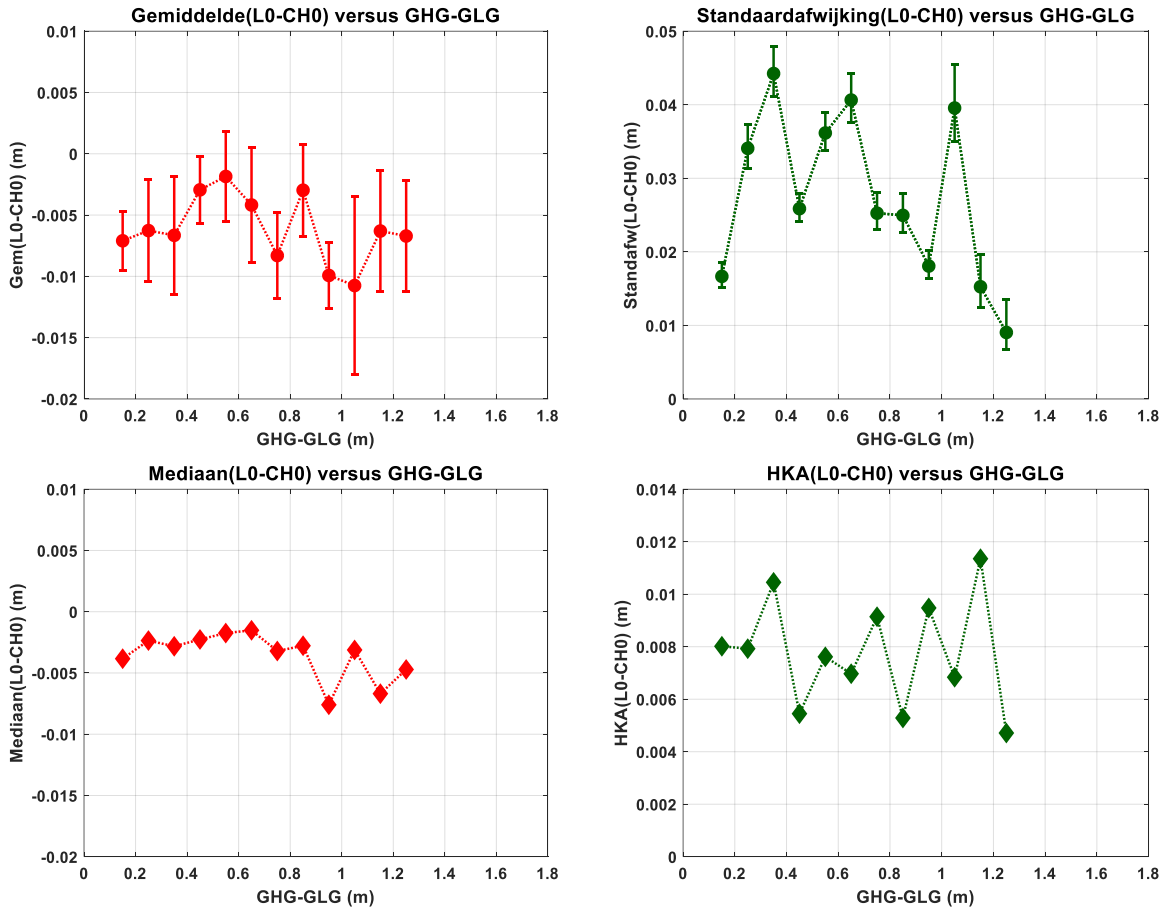
Uit bovenstaande figuren blijkt dat doorgaans op locaties met een hoge GLG het centrum van L0-CH0 enigszins lager en de spreiding van L0-CH0 groter is dan op locaties met een lage GLG.

Relatie L0-CH0 en GHG-GLG

Figuur 4.16: Spreidingsdiagram en boxplots (uitvergroot) van L0-CH0 versus het verschil van de gemiddelde hoogste en laagste grondwaterstand (GHG-GLG).



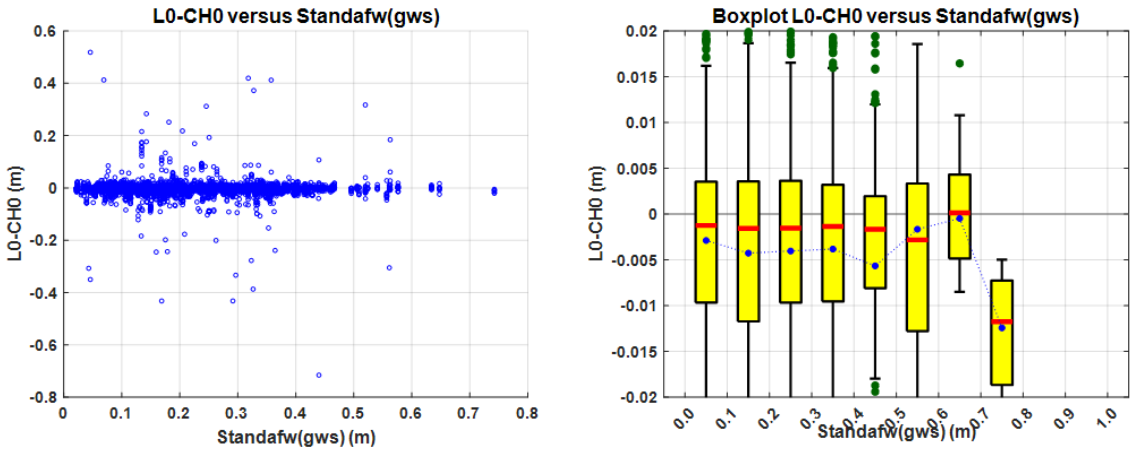
Figuur 4.17: Gemiddelde (links) en standaardafwijking (rechts) van L0-CH0 per decimeterklasse van het verschil tussen GHG en GLG, elk met tevens het 95%-betrouwbaarheidsinterval. Daaronder hun robuuste equivalenten, namelijk mediaan (links) en halvekwartielafstand (HKA, rechts). Ons oordeel baseert zich op de robuuste kengetallen.



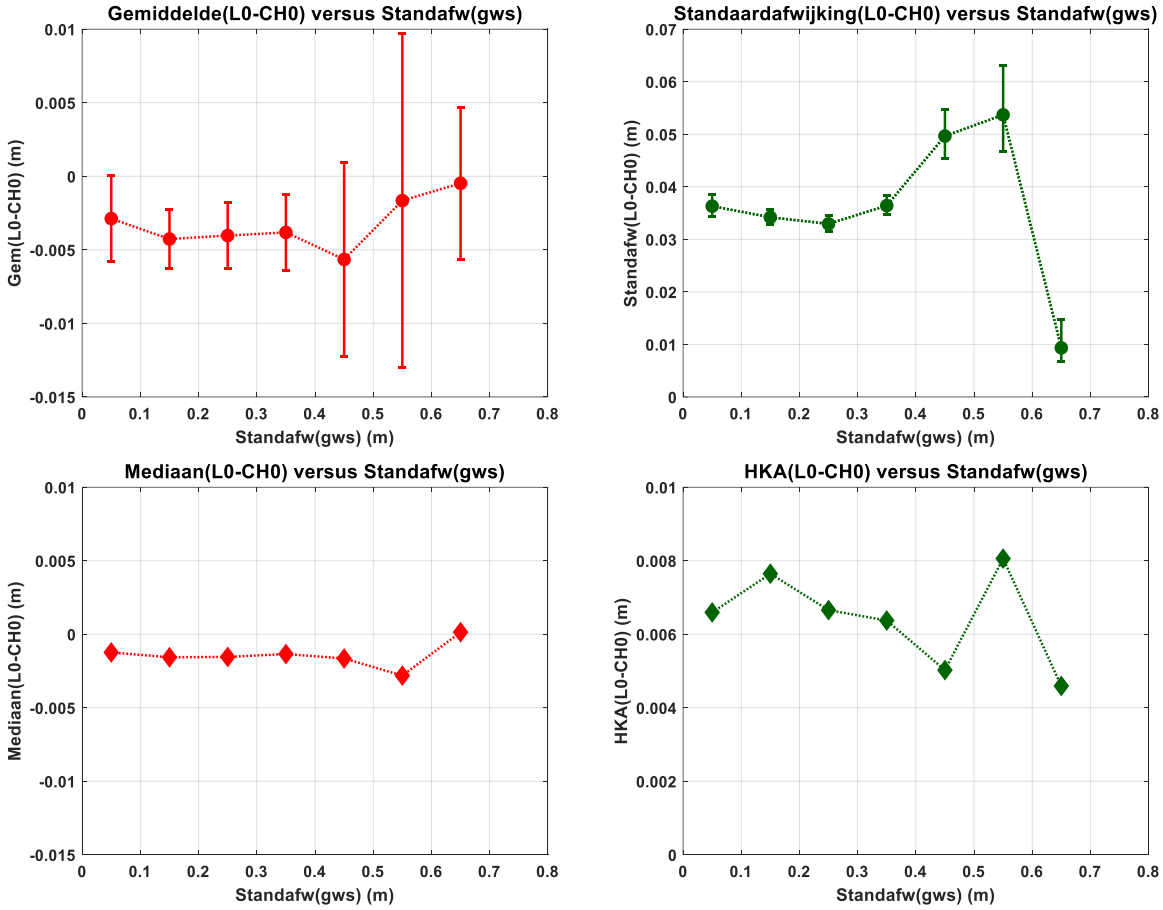
Uit bovenstaande figuren blijkt niet dat er sprake is van een duidelijke relatie tussen L0-CH0 en GHG-GLG.

Relatie L0-CH0 en standaardafwijking van de grondwaterstand

Figuur 4.18: Spreidingsdiagram en boxplots (uitvergroot) van L0-CH0 versus de standaardafwijking van de grondwaterstand.



Figuur 4.19: Gemiddelde (links) en standaardafwijking (rechts) van L0-CH0 per decimeterklasse van de standaardafwijking van de grondwaterstand, elk met tevens het 95%-betrouwbaarheidsinterval. Daaronder hun robuuste equivalenten, namelijk mediaan (links) en halvekwartielafstand (HKA, rechts). Ons oordeel baseert zich op de robuuste kengetallen.



Uit bovenstaande figuren blijkt niet dat er sprake is van een duidelijke relatie tussen L0-CH0 en de standaardafwijking van de grondwaterstand.

4.6 Kernpunten van dit hoofdstuk

- We zijn voor een aantal factoren nagegaan of deze gerelateerd zijn aan het verschil tussen een ruwe loggermeting (L0) en een daaraan gepaarde controlehandmeting (CH0), aangezien dit verschil is op te vatten als een proxy van de meetfout (de meetfout is immers niet exact gelijk aan L0-CH0, doordat CH0 ook een meetfout heeft). Dit leverde de volgende bevindingen:
  - Er blijkt sprake van een duidelijke relatie tussen L0-CH0 en de loggerleeftijd. Het centrum van L0-CH0 daalt tot een leeftijd van circa negen jaar, om daarna weer enigszins te stijgen. De spreiding van L0-CH0 neemt daarentegen duidelijk toe met de leeftijd. Dit laatste zal mede veroorzaakt zijn door drift.
  - Er blijkt sprake van een seizoenseffect op L0-CH0, waarbij de mediaan van L0-CH0 in het voorjaar doorgaans het laagst is (tot -5 mm) en in de tweede jaarhelft het hoogst (tot -1 mm). De tegenstelling lente – tweede jaarhelft blijkt behalve voor de mediaan van L0-CH0 ook min of meer op te gaan voor het 25- en het 75-percentiel. De standaardafwijking van L0-CH0 lijkt nauwelijks onderhevig aan een seizoenseffect, maar dit kengetal is minder bruikbaar, doordat het sterk wordt beïnvloed door uitschieters.

- Zowel het centrum als de spreiding van L0-CH0 blijkt monotoon gerelateerd aan de gemeten waterhoogte van de bijbehorende controleloggermeting (CL0) ten opzichte van de druksensor (CL0h). Als die waterhoogte toeneemt wordt het centrum van L0-CH0 negatiever, terwijl de spreiding van L0-CH0 toeneemt. Aangezien het denkbaar is dat dit causale relaties zijn, is het aan te bevelen de loggerdiepte zo beperkt mogelijk te houden, uiteraard wel rekening houdende met de ter plekke geldende kansverdeling van de grondwaterstand.
- Ten slotte is met grafische analyses nagegaan of het verschil van L0 en CH0 wellicht gerelateerd is aan hydrologische karakteristieken van de meetlocatie. We vonden geen duidelijke relaties van L0-CH0 met de gemiddelde grondwaterstand ten opzichte van maaiveld, met de GHG, met de GHG-GLG en met de standaardafwijking van de grondwaterstand. Er bleek daarentegen wel sprake van een relatie met de GLG. Op locaties met een hoge GLG is doorgaans het centrum van L0-CH0 enigszins lager en de spreiding van L0-CH0 groter dan op locaties met een lage GLG.

- Een andere indicatie van de invloed van de loggerleeftijd op de meetfout is het optreden van trend in de meetfout, wat we kunnen opvatten als een manifestatie van drift. We hebben trendanalyse uitgevoerd op alle reeksen CL0-CH0 (uniek betreffende loggernummer en combinatie peilbuis en filter) die voldoen aan de criteria dat de reeks minimaal vijf verschilwaarden CL0-CH0 omvat en minimaal vijf jaar lang is. Daaruit bleek dat 52 (18,4%) van de 283 geanalyseerde reeksen CL0-CH0 een statistisch significante trend vertoonden, met enigszins meer stijgingen (32) dan dalingen (20). De verdeling van de 283 geschatte trendhellingen is vrij symmetrisch, met - ongeacht statistische significantie - 139 positieve hellingen, 8 nulhellingen en 136 negatieve hellingen en met mediaan 0,00 cm/jaar en gemiddelde 0,02 cm/jaar. De maximale statistisch significante trend bedraagt 0,58 cm/jaar en de minimale bedraagt -1,10 cm/jaar. Als we alle statistisch niet-significante trendhellingen op nul zetten komt de standaardafwijking van de trendhelling uit op 0,14 cm/j. Dit is op te vatten als een statistische maat voor de drift die zich tot medio 2015 manifesteerde in het meetproces van CL0, ondanks de regelmatige aanpassingen van de installatiediepte, die juist bedoeld waren om drift tegen te gaan. Het feit dat ruim 18% van de geanalyseerde reeksen CL0-CH0 een statistisch significante trend vertoont, duidt op redelijk vaak optredende drift. Aangezien dit nog een onderschatting zal zijn van het werkelijke percentage drift (door de corrigerende aanpassingen van de installatiediepte), onderstreept dit de noodzaak om loggermetingen regelmatig te valideren en zonodig te corrigeren.
- Er is ook een analyse uitgevoerd van de levensduur van de loggers, gedefinieerd als de loggerleeftijd bij vervanging vanwege een storing die niet is veroorzaakt door externe factoren (zoals werkzaamheden). Over de beschouwde periode (20 maart 2003 tot 30 januari 2018) is gebruik gemaakt van 1.165 loggers voor de 549 peilfilters waar we informatie over ontvingen. Dit betrof de 549 initieel geïnstalleerde loggers en 616 vervangingen daarvan. De levensduurverdeling blijkt niet symmetrisch, maar scheef naar rechts. De mediane levensduur is dientengevolge lager dan de gemiddelde levensduur (mediaan bedraagt 3,37 jaar en gemiddelde bedraagt 4,28 jaar). Verder blijkt dat in het eerste jaar 7,9% van de loggers is vervangen vanwege een storing, na drie jaar 43,9% en na vijf jaar 71,8%. De langste levensduur (bij vervanging) bedroeg 13,8 jaar.
- Met de ons verstrekte gegevens over datum en reden van loggervervangingen hebben we ook de ontwikkeling van de storingsfrequentie als functie van de leeftijd gereconstrueerd. Het blijkt dat de storingsfrequentie van loggers tot de leeftijd vier jaar duidelijk toeneemt. Vervolgens neemt deze geleidelijk weer af tot de leeftijd tien jaar, waarna de storingsfrequentie weer toeneemt. Er is dus een relatief groot percentage uitvallers van jonge leeftijd, maar daarnaast zijn er ook meerdere overlevers, die aanvankelijk zonder veel problemen ouder worden. Deze laatste zijn

blijkbaar robuuster, mogelijk ook omdat ze in een gunstigere omgeving functioneren. Pas na tien jaar beginnen ook deze overlevers steeds meer mankementen te vertonen.

5 Opzet steekproef naar kenmerken meetfout

Deze studie beoogt de kenmerken in beeld te brengen van de meetfout bij het automatisch meten van de grondwaterstand door de Provincie Overijssel. In hoofdstuk 3 zijn daartoe langs empirische weg kenmerken van die meetfout afgeleid, middels statistische analyses van de verschillen tussen loggermetingen en daaraan gepaarde handmetingen van de grondwaterstand. Het betrof zowel kenmerken van de meetfout vóór correctie (de ruwe meetfout), als ná correctie (de meetfout na validatie). We concludeerden daarbij dat eerst een steekproef zal moeten worden verzameld van handmetingen die niet worden gebruikt om loggermetingen te valideren, om de standaardafwijking van de meetfout van gevalideerde loggermetingen zuiver te kunnen vaststellen. Dat advies is in dit hoofdstuk verder uitgewerkt.

5.1 Vereisten van steekproef naar kenmerken meetfout logger

Om de standaardafwijking van de meetfout van gevalideerde loggermetingen zuiver te kunnen vaststellen, dient: i) een set steekproefhandmetingen (SH) te worden vergeleken met de daaraan gepaard te achten gevalideerde loggermetingen (L1) en ii) de standaardafwijking van de resulterende set verschilwaarden (L1-SH) te worden gecorrigeerd voor de standaardafwijking van de meetfout van handmetingen. Deze steekproefhandmetingen dienen echter niet gebruikt te worden bij het valideren van de loggermetingen, zodat de meetfout van de loggermetingen ook na validatie onafhankelijk blijft van de meetfout van de steekproefhandmetingen.

De steekproefgrootte dient zodanig te zijn dat de standaardafwijking van het verschil L1-SH voldoende nauwkeurig kan worden geraamd. Dit is uitgewerkt in § 5.2.

Verder moet op verzoek van de Provincie Overijssel ook worden voldaan aan de volgende randvoorwaarden:

- de Provincie Overijssel mag geen kennis hebben van de meetlocaties en meettijdstippen van de steekproef, zodat het validatieproces daardoor niet beïnvloed kan worden;
- de Provincie mag vooraf aangeven dat bepaalde meetlocaties moeten worden uitgesloten van de steekproef, bijvoorbeeld omdat bekend is dat daar door een afwijkende meetopstelling minder betrouwbare metingen worden verricht;
- de steekproefopzet en -uitvoering dient voldoende gedetailleerd beschreven te zijn, zodat aanbesteding van het uitvoeren van de steekproef mogelijk is;
- de steekproefopzet dient ook na afronding van het onderzoek vertrouwelijk te blijven, om deze in de toekomst te kunnen blijven gebruiken. Er kunnen immers meetreeksen opnieuw worden gecorrigeerd op basis van het langjarige beeld van de verschillen met handmetingen, of door ontwikkeling van nieuwe technieken.

5.2 Relatie steekproefgrootte en nauwkeurigheid schatter meetonzekerheid

We spreken hier van een optimale steekproefopzet (gekenmerkt door het wát, wáár, wannéér en hóe meten) als met de steekproef de meetonzekerheid voldoende nauwkeurig kan worden geschat, tegen minimale inspanning (en daarmee kosten). Het is echter moeilijk om een objectief criterium te definiëren voor wat hier als '*voldoende nauwkeurig*' kan worden aangemerkt, zodat er een min of

meer subjectieve keuze moet wordt gedaan. Om een goede onderbouwing van die keuze mogelijk te kunnen maken, zullen we in deze paragraaf de relatie afleiden tussen de steekproefgrootte (en daarmee ook de kosten) en de uit de steekproef volgende precisie van de standaardafwijking van het verschil L1-SH.

Het ligt voor de hand dat de precisie van een geschatte standaardafwijking groter is naarmate die schatting gebaseerd is op meer meetwaarden. Dit lichten we hieronder toe, aan de hand van de stappen van de steekproef.

Stappen van de steekproef

1. Doe bij elk van *n* aselect gekozen loggers een willekeurige steekproefhandmeting (SH) van de grondwaterstand. Dit levert de set SH_1, SH_2, …, SH_n op. Registreer voor alle zekerheid bij elke steekproefhandmeting ook de bijbehorende ruwe loggermeting (L0). Dit levert de set ($L0_1$, $L0_2$, …, $L0_n$), die eventueel kan dienen om grote discrepanties te onderzoeken.

2. Verzamel circa een half jaar na de laatste steekproefhandmeting alle door de Provincie gevalideerde versies van de bij de steekproef betrokken loggerreeksen en registreer per reeks de gevalideerde meting die het dichtst bij de betreffende steekproefhandmeting ligt, mits het tijdsverschil niet meer dan één uur bedraagt. Dit levert de set $L1_1$, $L1_2$, …, $L1_n$.

3. Stel voor elke gevalideerde loggermeting (L1) en daaraan gepaarde steekproefhandmeting (SH) het verschil *v* vast:

$$v_i = \mathrm{L1}_i - \mathrm{SH}_i \quad \text{voor } i = 1, 2, \ldots, n \qquad [5.1]$$

4. Schat de standaardafwijking van de populatie van het verschil met de volgende schatter:

$$s_v = \sqrt{\frac{\sum_{i=1}^{n}(v_i - \bar{v})^2}{n-1}} \qquad [5.2]$$

   waarin s_v de schatting van de populatiestandaardafwijking van het verschil tussen L1 en SH, *i* de index van het verschilpaar (*i*=1, 2, …, *n*), *n* het aantal verschillen en $\bar{v}$ het gemiddelde verschil.

5. Bepaal de standaardafwijking van de meetfout van een gevalideerde loggermeting ($s_{e,\mathrm{L1}}$) door s_v te corrigeren voor de standaardafwijking van de meetfout van een handmeting ($s_{e,\mathrm{H}}$), volgens:

$$s_{e,\mathrm{L1}} = \sqrt{s_v^2 - s_{e,\mathrm{H}}^2} \qquad [5.3]$$

6. En bepaal het gemiddelde van de meetfout van een gevalideerde loggermeting ($\bar{e}_{\mathrm{L1}}$), volgens:

$$\bar{e}_{\mathrm{L1}} = \bar{v} + \bar{e}_{\mathrm{H}} \qquad [5.4]$$

Als de meetwaarden afkomstig zijn uit een normale kansverdeling en aselect zijn verkregen, dan geldt de volgende statistische relatie tussen de geschatte (s_v) en de werkelijke standaardafwijking van het verschil (σ_v):

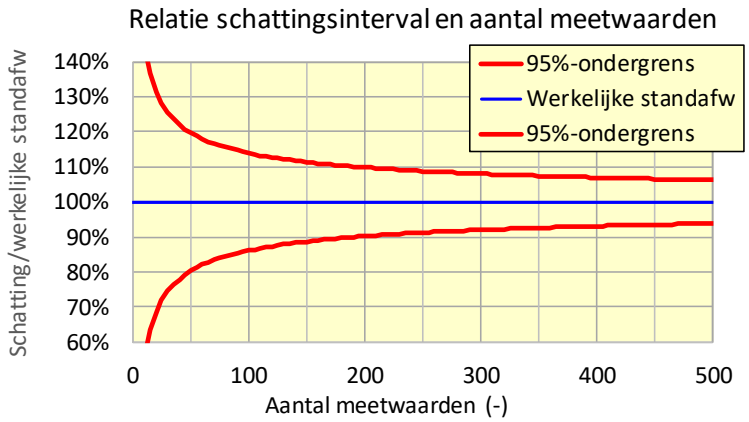
$$\frac{(n-1)\cdot s_v^2}{\sigma_v^2} \sim \chi^2_{n-1} \qquad [5.5]$$

waarin χ^2_{n-1} een willekeurige waarde uit de chikwadraat-kansverdeling voor *n*-1 vrijheidsgraden en *n* het aantal verschillen. Uit deze relatie volgen de grenzen van het relatieve schattingsinterval, waarbinnen de ratio s_v/σ_v zich in 95% van de gevallen zal bevinden:

$$\text{ondergrens 95\%-interval} = \sqrt{\frac{\chi^2_{0{,}025;n-1}}{n-1}} \text{ en bovengrens} = \sqrt{\frac{\chi^2_{0{,}975;n-1}}{n-1}} \qquad [5.6]$$

waarin $\chi^2_{0{,}025;n-1}$ en $\chi^2_{0{,}975;n-1}$ het 2,5- respectievelijk 97,5-percentiel van de chikwadraat-verdeling bij n-1 vrijheidsgraden. In figuur 5.1 is de relatie weergegeven tussen dit interval en het aantal meetwaarden.

Figuur 5.1: Relatie tussen het 95%-schattingsinterval van de standaardafwijking en het aantal meetwaarden waar de schatting op gebaseerd is.



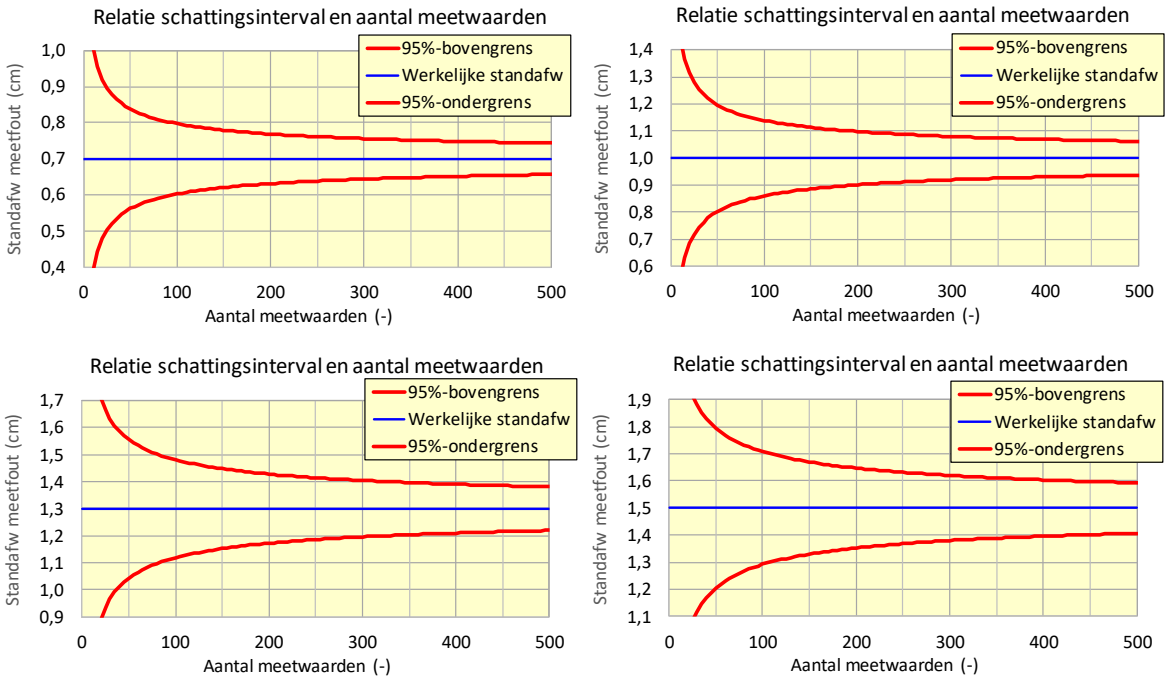
We zien in figuur 5.1 dat bij 100 steekproefhandmetingen er een kans van 95% is dat de schatting van s_v, de standaardafwijking van L1-SH, niet meer dan circa 14% afwijkt van σ_v, de werkelijke standaardafwijking van L1-SH. Bij 150 steekproefhandmetingen zal deze niet meer dan circa 11% afwijken en bij 200 steekproefhandmetingen niet meer dan circa 10%. De precisiewinst per extra steekproefeenheid is bij meer dan circa 150 steekproefhandmetingen dus nog maar gering.

Uit formule [5.6] volgt ook dat de schatting s_v in 95% van de gevallen in het volgende interval zal liggen:

$$\text{ondergrens 95\%-interval} = \sigma_v\sqrt{\frac{\chi^2_{0{,}025;n-1}}{n-1}} \text{ en bovengrens} = \sigma_v\sqrt{\frac{\chi^2_{0{,}975;n-1}}{n-1}} \qquad [5.7]$$

Ter illustratie toont figuur 5.2 uitwerkingen van dit interval voor enkele mogelijk realistische werkelijke standaardafwijkingen σ_v.

Figuur 5.2: Relatie tussen het 95%-schattingsinterval van de standaardafwijking en het aantal meetwaarden waar de schatting op gebaseerd is, voor een werkelijke standaardafwijking σ_v van respectievelijk 0,7, 1,0, 1,3 en 1,5 cm.



Dit laatste schattingsinterval moet overigens niet worden verward met het betrouwbaarheidsinterval, want dat is het interval rond een schatting (s_v), dat met een bepaalde betrouwbaarheid de werkelijke waarde (σ_v) zal bevatten. Ook de grenzen van dit interval volgen uit formule [5.5]. Zie daarvoor § 5.4.

5.3 Selectie van loggers voor de steekproef

- De selectie van loggers voor de steekproef bestaat uit de volgende stappen. Stap 1 zal worden genomen bij de bespreking van dit rapport. De overige stappen zullen worden gestart na accordering van dit rapport.

1. In samenspraak met de Provincie vaststellen van het aantal in de steekproef op te nemen loggers (en daarmee het aantal te nemen steekproefhandmetingen). Daarbij zullen zowel het beschikbare budget als de gewenste nauwkeurigheid in ogenschouw worden genomen. Geschikt hulpmiddel bij de afweging zijn de in de figuren 5.1 en 5.2 weergegeven relaties tussen de steekproefgrootte en de daarmee te verwachten precisie van de schatting van de standaardafwijking van het verschil tussen een gevalideerde loggermeting (L1) en een steekproefhandmeting (SH). Die precisie is in de figuren af te lezen uit de breedte van het schattingsinterval. Inmiddels is (bij de bespreking van het concept van dit rapport) gekozen voor een steekproef van 200 handmetingen van de grondwaterstand, uit te voeren gespreid over 2019.

2. Door de Provincie overleggen van een lijst met metadata van alle loggers van de te onderzoeken populatie (peilbuis, filter, coördinaten, filterstelling, maaiveldhoogte, meetpunthoogte en loggerdiepte). De Provincie kan daarbij bepaalde loggers bij voorbaat uitsluiten, bijvoorbeeld omdat

bekend is dat daar door een afwijkende meetopstelling minder betrouwbare metingen worden verricht.

3. Door ons selecteren van de loggers van de steekproef volgens het verlotingsprincipe, zodanig dat elke logger vooraf dezelfde kans heeft om te worden geselecteerd. We zullen geen stratificatie toepassen bij het selecteren. Stratificatie naar één of meer factoren kan de nauwkeurigheid van de steekproef bij dezelfde steekproefgrootte vergroten, mits elke factor een voldoende grote invloed heeft op de standaardafwijking van de steekproefvariabele (hier is dat L1-SH).

4. Beschikbaar stellen van de lijst met geselecteerde loggers aan de instantie die de steekproef gaat uitvoeren. De Provincie Overijssel mag deze lijst niet inzien, om te vermijden dat het validatie-proces daardoor beïnvloed kan worden.

5.4 Statistische verwerking steekproefresultaten

Als de steekproefresultaten beschikbaar zijn, dienen deze te worden verwerkt met de in § 5.2 beschreven stappen 3 t/m 6. Daarnaast kunnen ook de betrouwbaarheden van de onderliggende schattingen in beeld worden gebracht, zoals hieronder beschreven.

- Bepaal het 95%-betrouwbaarheidsinterval van de schatting van de standaardafwijking (s_v) van *v* = L1 - SH. Dit is het interval rond de schatting (s_v), dat met een bepaalde betrouwbaarheid de werkelijke waarde (σ_v) zal bevatten. De grenzen van dit interval volgen uit formule [5.5]:

  ondergrens 95%-interval $= s_v \sqrt{\frac{n-1}{\chi^2_{0{,}975;n-1}}}$ en bovengrens = $s_v \sqrt{\frac{n-1}{\chi^2_{0{,}025;n-1}}}$ [5.8]

- Bepaal ook het 95%-betrouwbaarheidsinterval van de schatting van het gemiddelde $\bar{v}$, volgens:

  ondergrens 95%-interval $= \bar{v} - t_{97{,}5\%;n-1} \cdot \frac{s_v}{\sqrt{n}}$ en bovengrens $= \bar{v} + t_{97{,}5\%;n-1} \cdot \frac{s_v}{\sqrt{n}}$ [5.9]

  waarin $t_{97{,}5\%;n\text{-}1}$ de waarde van de Student-t-verdeling met een eenzijdige onderschrijdingskans van 97,5% bij *n*-1 vrijheidsgraden. Als dit interval niet de waarde nul omvat, is er met 95% betrouwbaarheid sprake van een systematische afwijking tussen L1 en SH.

5.5 Enkele adviezen over uitvoering van de steekproef

- Er blijkt sprake van (beperkte) seizoensinvloeden op het verschil van loggermeting en handmetingen (zie § 4.3). Daarom is het beter de steekproef enigszins te spreiden over het jaar, om vertekening door seizoensinvloeden zoveel mogelijk te minimaliseren.
- Houd er bij de planning en kostenberekening rekening mee dat een ervaren waarnemer per dag 20 à 30 loggers kan bezoeken en daar de grondwaterstand handmatig kan meten.[27]
- Noteer bij het nemen van een handmeting zowel de systeemtijd als de loggertijd en registreer ook de loggermeting.

[27] Persoonlijke mededeling René Buijs, 6 juli 2018.

5.6 Kernpunten van dit hoofdstuk

- Om de standaardafwijking van de meetfout van gevalideerde loggermetingen zuiver te kunnen vaststellen dient eerst een steekproef te worden verzameld van handmetingen die niet worden gebruikt om loggermetingen te valideren. Deze handmetingen dienen dan te worden vergeleken met de daaraan gepaard te achten gevalideerde loggermetingen en de standaardafwijking van de resulterende set verschilwaarden dient vervolgens te worden gecorrigeerd voor de standaardafwijking van de meetfout van handmetingen.
- De steekproefgrootte dient zodanig te zijn dat de standaardafwijking van de verschilwaarden voldoende nauwkeurig kan worden geraamd. Om een goede onderbouwing mogelijk te kunnen maken van die keuze (te doen in samenspraak met de Provincie), hebben we de relatie afgeleid tussen de steekproefgrootte (en daarmee ook de kosten) en de uit de steekproef volgende precisie van de standaardafwijking van het verschil.
- Inmiddels is (bij de bespreking van het concept van dit rapport) gekozen voor een steekproef van 200 handmetingen van de grondwaterstand, uit te voeren gespreid over 2019.

6 Aanbevelingen

1. Zorg dat de installatiediepte in de Keller-software gelijk is aan de loggerdiepte in de HydroData-module van Menyanthes. Anders zijn de controleloggermetingen (CL0) minder representatief voor de ruwe loggermetingen (L0) en daardoor niet bruikbaar in het validatieproces van de Provincie. Dat is inefficiënt, want het validatieproces heeft baat bij zoveel mogelijk ijkmateriaal. Bovendien wordt CL0 vrijwel gelijktijdig met CH0 gemeten, terwijl dat doorgaans niet geldt voor de dichtstbijzijnde L0.

2. Zorg dat de bij de omrekening van drukverschil naar waterhoogte toegepaste waarden van soortelijk gewicht en zwaartekrachtsversnelling worden bewaard in de HydroData-module van Menyanthes. Dat stelt beter in staat de consistentie van een meetreeks te beoordelen. Nu ontbreekt deze informatie nog.

3. Onderzoek de mogelijkheid de kwaliteit van de loggermetingen te bewaken door middel van methoden uit de statistische procesbeheersing. Een optie is het gebruik van een controlekaart, waarin het verschil van loggermeting en daaraan gepaarde handmeting is uitgezet tegen de tijd en waarin waarschuwings- en/of actiegrenzen zijn weergegeven op basis van kennis van de kansverdeling van de ruwe meetfout, representatief voor de fase dat er nog geen sprake is van kwaliteitsverslechtering. In combinatie met voldoende en betrouwbare controlemetingen kan het oordeelkundig gebruik van deze kaart leiden tot een goede kwaliteitsbewaking van het meetproces. Dit is nodig, omdat de meetfout van deze loggers snel kan toenemen, vooral door drift. De momenteel gehanteerde criteria voor correctie van een meetreeks zijn niet statistisch onderbouwd.

4. Onderzoek ook de mogelijkheid of statistische trendanalyse kan helpen het meetproces beheerst te houden.

5. Betrek bij de kwaliteitscontrole ook de verschillende basismetingen van de logger, die uit te lezen zijn in de vorm van een binair IDC-bestand. Relevante basismetingen zijn de verschildruk, de totale druk, de luchtdruk, de watertemperatuur en de luchttemperatuur. Eventuele meetfouten kunnen soms beter gedetecteerd worden door kritische beoordeling van de fysische plausibiliteit van deze basismetingen.

Aangehaalde literatuur

- Gassmann, E. (2014): *Pressure-sensor fundamentals: Interpreting Accuracy and Error*. Chemical Engineering Progress (CEP), June 2014, pp. 37 – 45.

- Gooijer, J. en De Meij, T. (2013, 2018a en 2018b): *Werkwijze meten grondwaterstanden Provincie Overijssel*. Interne notitie Provincie Overijssel. Aan ons beschikbaar gestelde versies:
  - 2013 - 3 november 2013, 7 blz;
  - 2018a - 12 juni 2018, 8 blz.
  - 2018b - 31 oktober 2018, 6 blz.

- ISO (1993): *Guide to the Expression of Uncertainty in Measurement*. International Organization for Standardization, 1993, 101 blz.

- Knotters, M., De Meij, Th. en Pleijter, M. (2013). *Nauwkeurigheid van handmatig gemeten grondwaterstanden en stijghoogtes - Verslag van een veldexperiment*. Alterra Wageningen UR, rapport 2476, december 2013, 30 blz.

- Pleijter, M., Van Hamersveld, L. en Knotters, M. (2015). *Systematische fouten in metingen van grondwaterstanden met drukopnemers: verslag van een data-analyse*. Wageningen, Alterra Wageningen UR.

- Post, V.E.A. and Von Asmuth, J.R. (2013): *Review: Hydraulic head measurements—new technologies, classic pitfalls*. Hydrogeology Journal, DOI 10.1007/s10040-013-0969-0.

- Sen, P.K. (1968): *Estimates of the regression coefficient based on Kendall's tau*. Journ. Am. Statist. Assoc., 63, blz. 1379 - 1389, 1968.

- Theil, H. (1950): *A rank-invariant method of linear and polynomial regression analysis, 1,2 and 3*. Ned. Akad. Wetensch. Proc., 53, blz. 386 - 392, 521 - 525 en 1397 - 1412.

- Von Asmuth, J.R. (2010). *Over de kwaliteit, frequentie en validatie van druksensorreeksen (t.b.v. een optimale meetfrequentie en verwerkingsprotocol voor verdrogingsmeetnet Overijssel)*. Nieuwegein, november 2010, 44 blz.

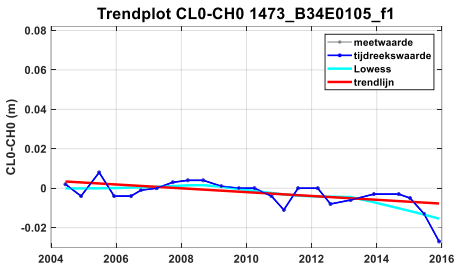
- Von Asmuth, J.R. (2018): *Kwaliteitsborging grondwaterstands- en stijghoogtegegevens: Protocol voor datakwaliteitscontrole (QC) (versie 2.0)*. Rapport PMB2018, Platform meetnetbeheerders grondwaterkwantiteit van de gezamenlijke provincies, Arnhem.

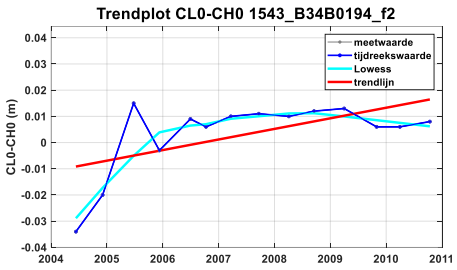
- Von Asmuth, J.R., Maas, K. Knotters, M. en Leunk, I. (2010): *Handleiding Menyanthes*. KWR, Nieuwegein, 52 blz.

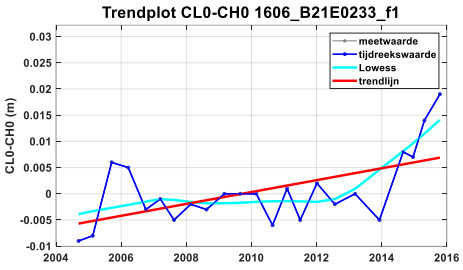
- Von Asmuth, J.R. en Van Geer, F.C. (2015): *Kwaliteitsborging grondwaterstands- en stijghoogtegegevens: systematiek en methodiek voor datakwaliteitscontrole (QC).* KWR-rapport 2015.004, april 2015, 56 blz.

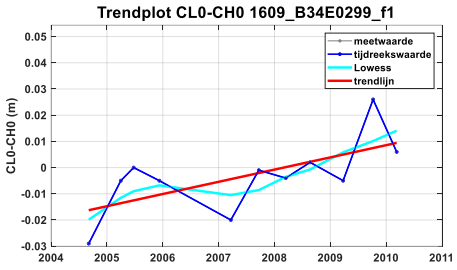
Bijlage: Trendplots CL0-CH0 periode 2004 t/m 2015

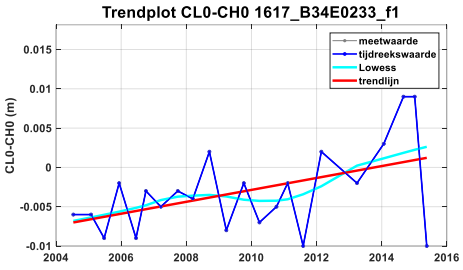
Trendplots van de verschilwaardereeksen van controlemetingen van díe loggers waar deze reeks een statistisch significante trend vertoont (duidend op drift).

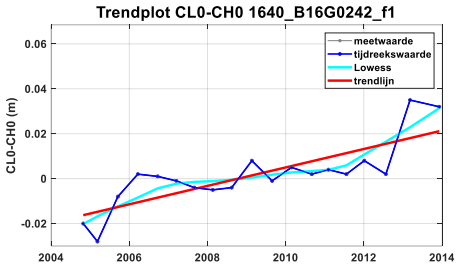


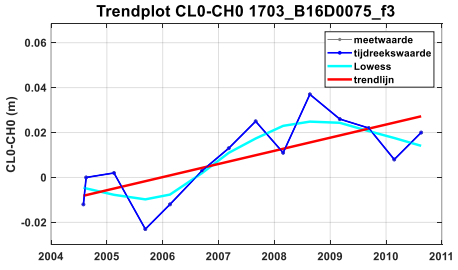


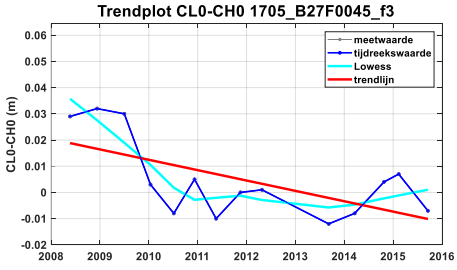


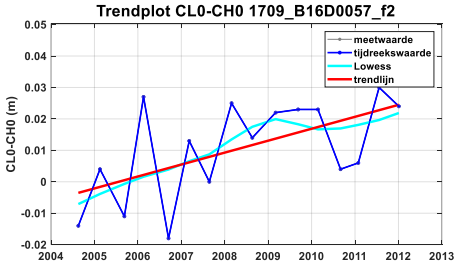


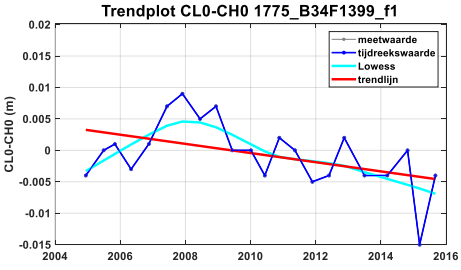


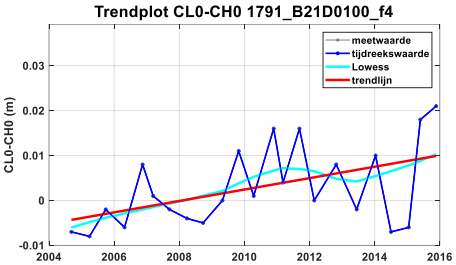




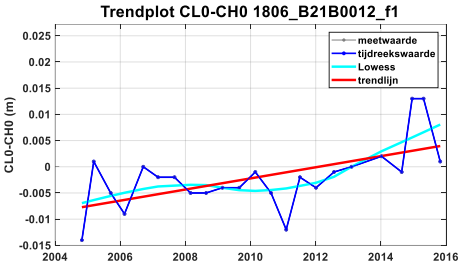




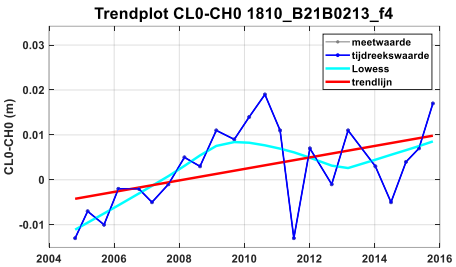




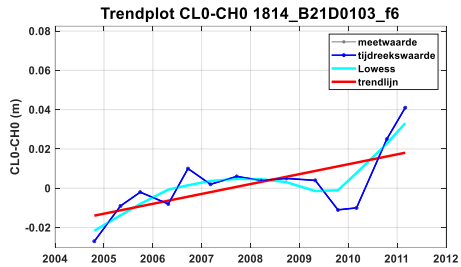
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0012711 eenheden per jaar.



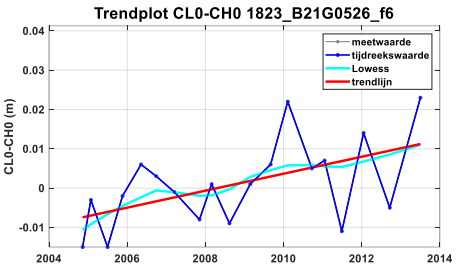
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0010612 eenheden per jaar.



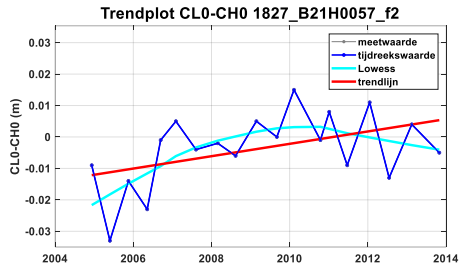
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0012795 eenheden per jaar.



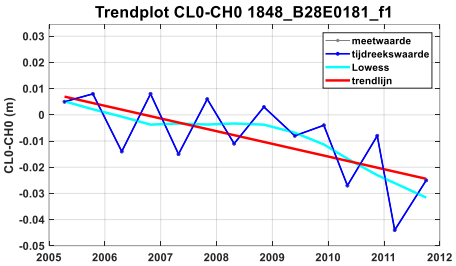
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0050315 eenheden per jaar.



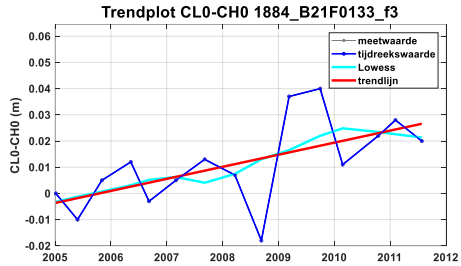
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.002159 eenheden per jaar.



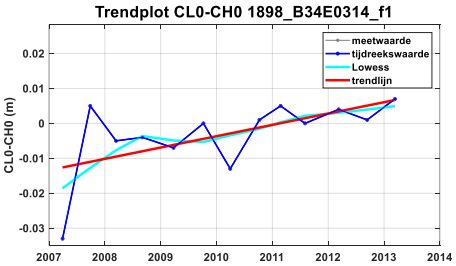
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.001967 eenheden per jaar.



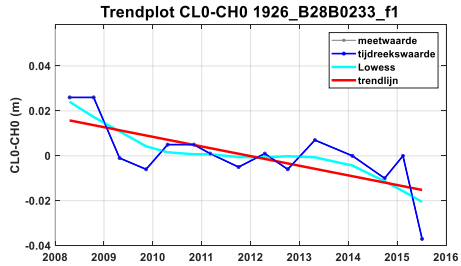
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0048419 eenheden per jaar.



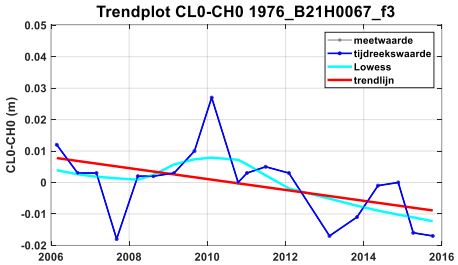
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0046039 eenheden per jaar.



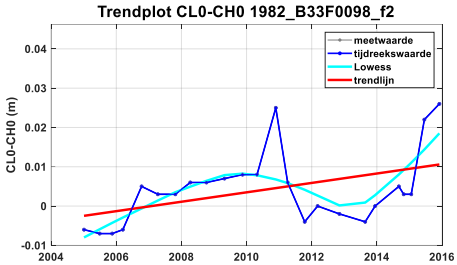
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0032436 eenheden per jaar.



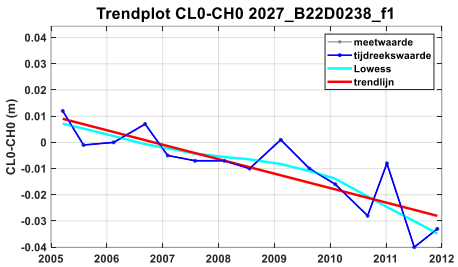
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0042915 eenheden per jaar.



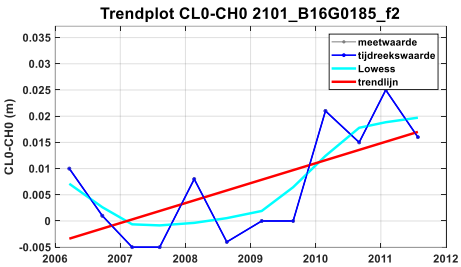
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0017321 eenheden per jaar.



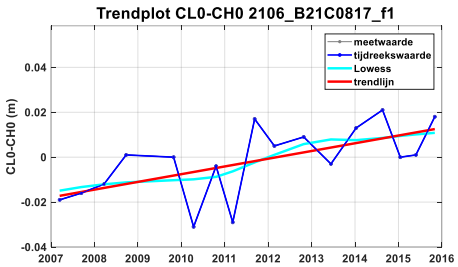
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0011973 eenheden per jaar.



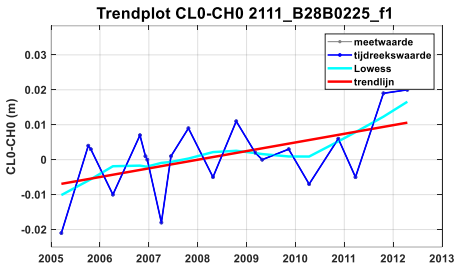
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0055191 eenheden per jaar.



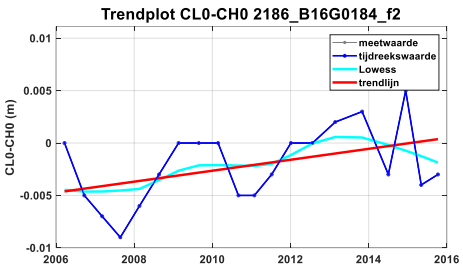
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0038075 eenheden per jaar.



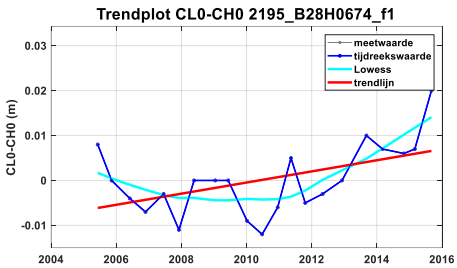
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0034416 eenheden per jaar.



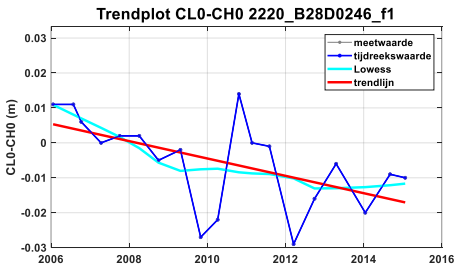
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0024737 eenheden per jaar.



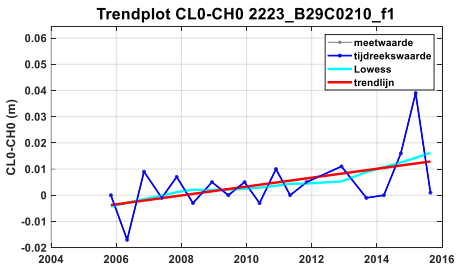
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.00052215 eenheden per jaar.



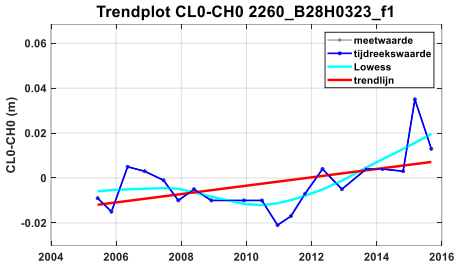
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0012394 eenheden per jaar.



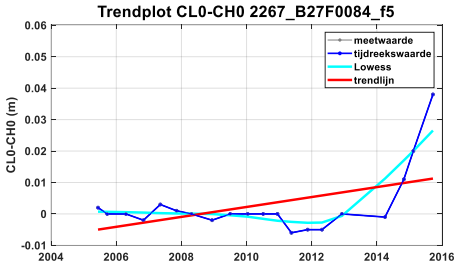
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0024838 eenheden per jaar.



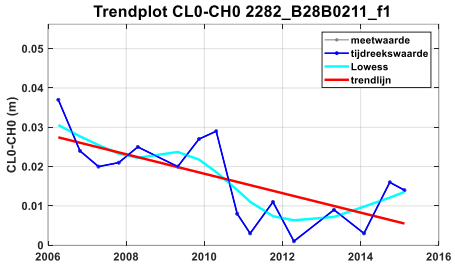
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0016956 eenheden per jaar.



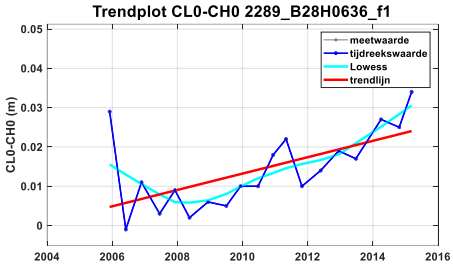
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0018624 eenheden per jaar.



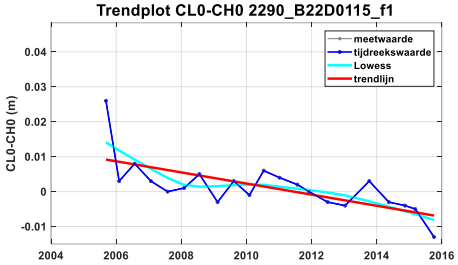
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.001579 eenheden per jaar.



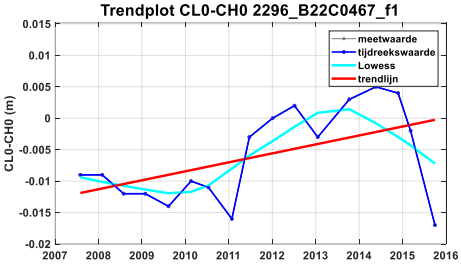
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0024768 eenheden per jaar.



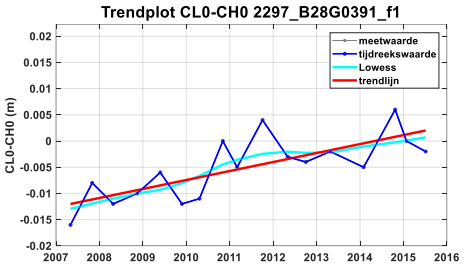
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0020821 eenheden per jaar.



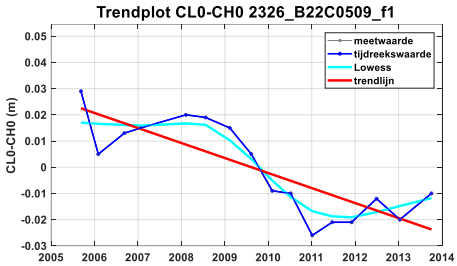
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0015922 eenheden per jaar.



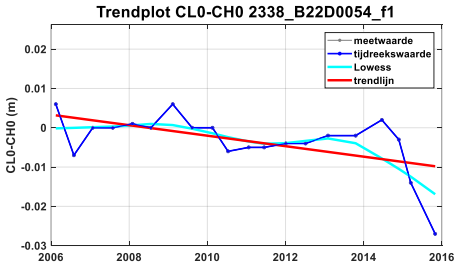
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.001425 eenheden per jaar.



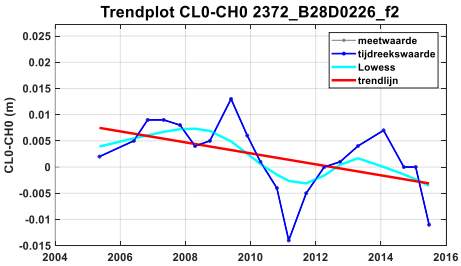
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0017144 eenheden per jaar.



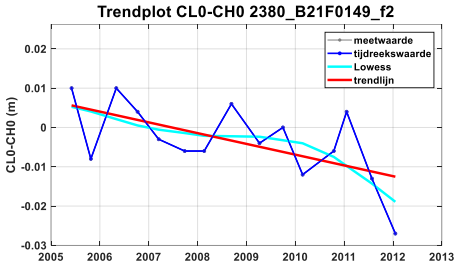
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0057302 eenheden per jaar.



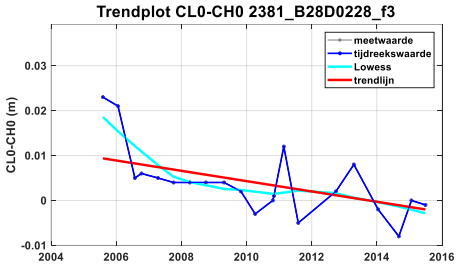
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0013349 eenheden per jaar.



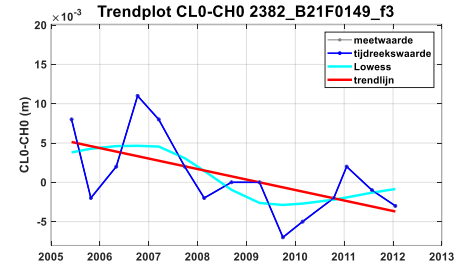
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0010475 eenheden per jaar.



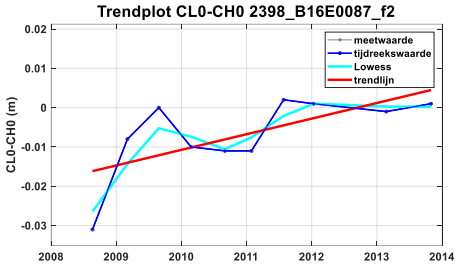
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0027379 eenheden per jaar.



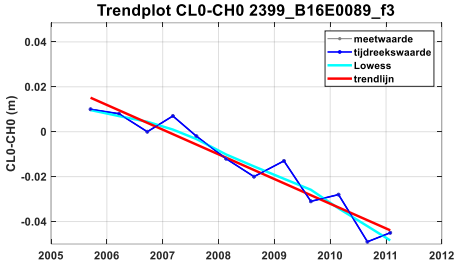
Resultaat MK-toets is
Oordeel = trend
De geschatte trend = -0.0011483 eenheden per jaar.



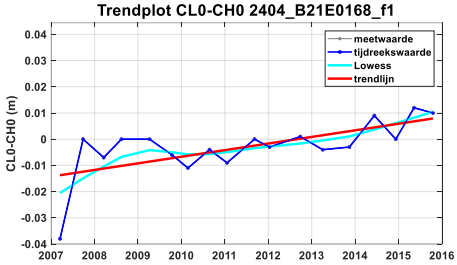
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0013338 eenheden per jaar.



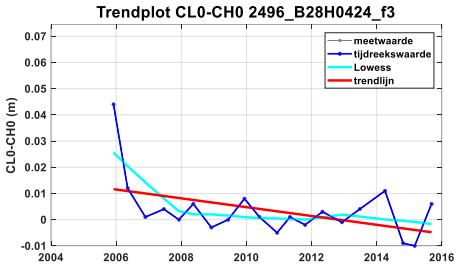
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0039721 eenheden per jaar.



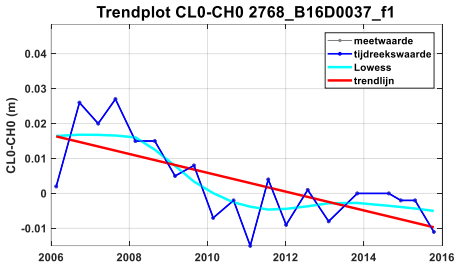
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.011 eenheden per jaar.



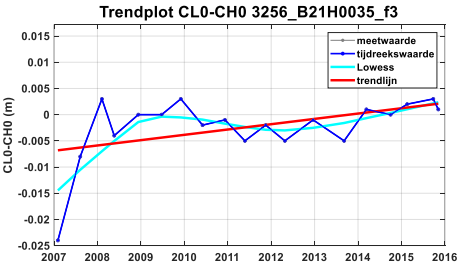
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0025212 eenheden per jaar.



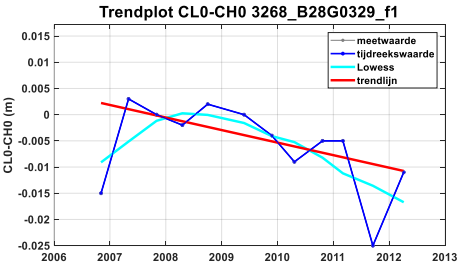
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0016835 eenheden per jaar.



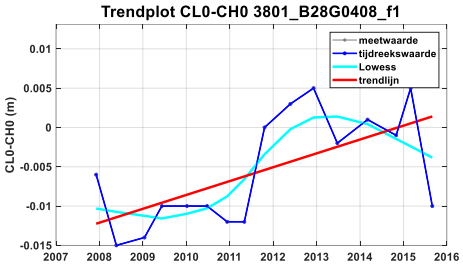
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = -0.0026928 eenheden per jaar.



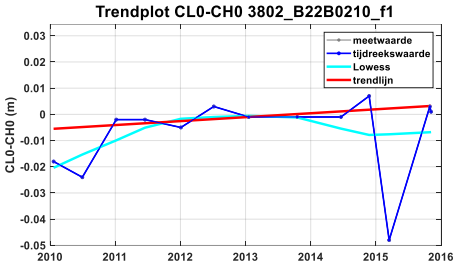
Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.001015 eenheden per jaar.



Resultaat MK-toets is
Oordeel = trend
De geschatte trend = -0.0023938 eenheden per jaar.



Resultaat LR-toets is
Oordeel = trend
De geschatte trend = 0.0017631 eenheden per jaar.



Resultaat MK-toets is
Oordeel = trend
De geschatte trend = 0.001504 eenheden per jaar.