Kwaliteit van Geo-informatie in theorie en praktijk

Een onderzoek naar de aspecten van kwaliteit van geo-informatie en een analyse van de attribuut nauwkeurigheid van een dataset

J.A. van Popering

Afstudeerscriptie ten behoeve van het verkrijgen van de
Master of Science graad in Geografische Informatie Systemen (UNIGIS)
aan de Faculteit van Aard- en Levens wetenschappen
Vrije Universiteit Amsterdam

Mentoren:
Prof. Dr. H.J. Scholten
Drs. A. J. Wagtendonk

Maart 2004

Samenvatting

Deze scriptie verwoordt een onderzoek dat is verricht naar kwaliteit van geo-informatie in theorie en praktijk.
De onderzoeksvraag waar in dit onderzoek antwoord op gegeven wordt is of de kwaliteit van geo-informatie op een dusdanige wijze gewaarborgd wordt zodat deze aansluit op de nominale grondslag of op de vereiste kwaliteit ten behoeve van het bedrijfsproces.
Om hierop te kunnen antwoorden zijn een aantal deelonderzoeken uitgevoerd. Het eerste deelonderzoek is ingegaan op de vraag wat kwaliteit van geo-informatie nu eigenlijk is, hoe dit getoetst kan worden en welke nationale en internationale standaarden er bestaan. Hierna is bij een viertal Nederlandse organisaties die veel geo-informatie inwinnen en beheren via een kwalitatief onderzoek bekeken hoe zij de kwaliteit van hun geo-informatie borgen. Tot slot is bij waterschap Zeeuwse Eilanden onderzocht in hoeverre de attribuutnauwkeurigheid van een dataset overeenkomt met de veronderstelling en eisen en wensen ten aanzien van de kwaliteit.
De uitkomsten van het onderzoek zijn divers. Geconcludeerd mag worden dat inzicht in de werkelijke kwaliteit van geo-informatie zeer belangrijk is. Enerzijds vanwege het kunnen bepalen of een dataset geschikt is voor een bepaalde toepassing, anderzijds vanwege het kunnen controleren of een bepaalde inventarisatie conform de eisen is uitgevoerd. Door een juist inzicht in de kwaliteit van een dataset kunnen aanzienlijke kosten bespaard worden en kan een reële inschatting gemaakt worden van eventuele gevolgen van het gebruiken van een bepaalde kwaliteit geo-informatie.

Daarnaast is gebleken dat in veel gevallen de werkelijke kwaliteit van geo-informatie niet getoetst wordt en dat er geen normen gesteld worden. Aangezien er gebleken is dat er een aanzienlijk verschil bestaat tussen de werkelijke en gewenste/veronderstelde kwaliteit van de attribuutnauwkeurigheid van de geo-informatie, mag er niet aangenomen worden dat dit bij niet-getoetste kwaliteitsparameters wel overeen zal komen. Hierdoor wordt het belang van het stellen en toetsen van normen van kwaliteit van geo-informatie zeer hoog geacht. Door inzicht in de kwaliteit van geo-informatie is het ook mogelijk om het inwin- of beheersproces te verbeteren en hiermee de kwaliteit te verhogen.

De hoofdconclusie van het onderzoek is dat de kwaliteit van geo-informatie nog niet op een dusdanige wijze gewaarborgd wordt zodat deze aansluit op de nominale grondslag of op de vereiste kwaliteit ten behoeve van het bedrijfsproces. Wel worden er aanbevelingen gedaan om hierin verbeteringen aan te brengen.

Hoofdstukindeling

DISCLAIMER 5

VOORWOORD 6

1 INLEIDING EN AANLEIDING 8

2 KWALITEIT EN GEO-INFORMATIE 17

3 KWALITEITSASPECTEN VAN GEO-INFORMATIE BIJ NEDERLANDSE ORGANISATIES 34

4 KWANTITATIEF ONDERZOEK NAAR ATTRIBUUT NAUWKEURIGHEID 56

5 CONCLUSIES EN AANBEVELINGEN 98

LITERATUURLIJST 103

LIJST MET GERAADPLEEGDE INTERNETSITES 105

BIJLAGE 1 INTERVIEW BIJ TOPOGRAFISCHE DIENST KADASTER 107

BIJLAGE 2 INTERVIEW BIJ HET KADASTER 114

BIJLAGE 3 INTERVIEW BIJ DE ADVIESDIENST GEO-INFORMATIE EN ICT 119

BIJLAGE 4 INTERVIEW BIJ WATERSCHAP ZEEUWSE EILANDEN 124

BIJLAGE 5 TOTSTANDKOMING BESTAND DIJKBEKLEDINGEN 130

BIJLAGE 6 OVERZICHT VAN GEMETEN OBJECTCODERINGEN 134

BIJLAGE 7 PROCESSCHEMA TOEVOEGEN STATUS 135

BIJLAGE 8 PROCESSCHEMA TOEVOEGEN ZICHTBAARHEID 136

BIJLAGE 9 TOEVOEGEN GEBIED AAN WATERKERINGENBESTAND 137

BIJLAGE 10 PROCESSCHEMA KIEZEN TE TOETSEN VLAKKEN 138

BIJLAGE 11 LIJST MET CODERINGEN VAN STEENBEKLEDINGEN 140

BIJLAGE 12 HET GENEREREN VAN HET CENTROIDENBESTAND 142

BIJLAGE 13 HET GEREEDMAKEN VAN DE GEO-EXPLORER 143

BIJLAGE 14 CONVERSIE VAN GEO-EXPLORER NAAR WERKSTATION 144

BIJLAGE 15 VERWERKEN INVENTARISATIE TOT RESULTAAT 145

BIJLAGE 16 VERGELIJKEN BESTANDEN T.B.V. CONVERSIEFOUTEN 146

BIJLAGE 17 INVULLEN VAN DE SOORTEN FOUTEN 147

BIJLAGE 18 CONTROLE NIET TE VERKLAREN FOUTEN 148

BIJLAGE 19 CONFUSIONMATRIX VAN ALLE WAARNEMINGEN 149

BIJLAGE 20 CONFUSIONMATRIX VAN FOUTEN 51, 52 EN 53 150

BIJLAGE 21 BEPALEN NAUWKEURIGHEID GEO-EXPLORER 151

BIJLAGE 22 VRAGENLIJST BEPALEN BENODIGDE KWALITEIT..... 152

BIJLAGE 23 INVULLEN KENMERKEN VEREISTE KWALITEIT..... 156

BIJLAGE 24 VERGELIJKINGSPERCENTAGE VEREISTE KWALITEIT..... 157

Inhoudsopgave

DISCLAIMER **5**

VOORWOORD **6**

1 INLEIDING EN AANLEIDING **8**

1.1 PROBLEEMSTELLING 8
1.1.1 Doelstelling *8*
1.1.2 Vraagstelling *9*
1.2 ONDERZOEKSONTWERP 10
1.2.1 Methode *10*
1.2.2 Object *14*
1.3 OPBOUW VAN HET RAPPORT 15

2 KWALITEIT EN GEO-INFORMATIE **17**

2.1 INLEIDING 17
2.2 KWALITEIT EN GEO-INFORMATIE 17
2.2.1 Geo-informatie *17*
2.2.2 Kwaliteit *18*
2.2.3 Modellering *19*
2.3 KWALITEITSASPECTEN 20
2.3.1 Voornorm - geografische informatie- gegevensbeschrijving – kwaliteit (VNK) *20*
2.3.2 NCDCDS *23*
2.3.3 Spatial Data Transfer Standard (SDTS) *24*
2.3.4 DIGEST *25*
2.3.5 Referentiemodel Kwaliteit van geo-informatie (RAVI) *27*
2.4 VERSCHILLEN TUSSEN DE STANDAARDEN 29
2.5 ASPECTEN VAN KWALITEIT VAN GEO-INFORMATIE 31
2.5.1 Algemeen *31*
2.5.2 Kwaliteitsparameters *31*
2.5.3 Overige kwaliteitsaspecten *32*
2.6 TESTEN VAN DE VERSCHILLENDE KWALITEITSPARAMETERS 32
2.7 CONCLUSIE 33

3 KWALITEITSASPECTEN VAN GEO-INFORMATIE BIJ NEDERLANDSE ORGANISATIES **34**

3.1 INLEIDING 34
3.2 KWALITEIT VAN GEO-INFORMATIE BIJ TOPOGRAFISCHE DIENST KADASTER 34
3.2.1 Inleiding *34*
3.2.2 Kwaliteit van geo-informatie *36*
3.3 KWALITEIT VAN GEO-INFORMATIE BIJ KADASTER 39
3.3.1 Inleiding *39*
3.3.2 Kwaliteit van geo-informatie *42*
3.3.2.1 Algemeen 42
3.3.2.2 Kadastrale kaart 43
3.3.2.3 Grootschalige Basiskaart van Nederland 44
3.4 KWALITEIT VAN GEO-INFORMATIE BIJ DE ADVIESDIENST GEO-INFORMATIE EN ICT 46
3.4.1 Inleiding *46*
3.4.2 Kwaliteit van geo-informatie *47*
3.5 KWALITEIT VAN GEO-INFORMATIE BIJ WATERSCHAP ZEEUWSE EILANDEN 50
3.5.1 Inleiding *50*

3.5.2 Kwaliteit van geo-informatie 51
3.6 OVEREENKOMSTEN/VERSCHILLEN TUSSEN DE ORGANISATIES 53
3.7 CONCLUSIE 54

4 KWANTITATIEF ONDERZOEK NAAR ATTRIBUUT NAUWKEURIGHEID 56

4.1 INLEIDING 56
4.2 BESCHRIJVING BESTAND STEENBEKLEDINGEN 56
4.2.1 Inleiding 56
4.2.2 Procesbeschrijving inwincampagne 59
4.2.3 Het product "steenbekleding" 60
4.3 MODELLERING 63
4.3.1 Inleiding 63
4.3.2 Onderdelen modellering 63
4.4 KEUZE TE ONDERZOEKEN ATTRIBUTEN EN METHODEN 63
4.4.1 Te onderzoeken attributen 63
4.4.2 Methode van analyse 64
4.5 GENEREREN STEEKPROEF 67
4.5.1 Inleiding 67
4.5.2 Status 68
4.5.3 (On)Zichtbaar 69
4.5.4 Gebieden 69
4.5.5 Grootte van de steekproef 70
4.5.6 Kiezen van vlakken 71
4.6 METHODE VAN INVENTARISATIE EN VERWERKING 75
4.7 VOORBEREIDING INVENTARISATIE 77
4.7.1 Inleiding 77
4.7.2 Gereedmaken steekproefbestand 78
4.7.3 Gereedmaken topografie 78
4.7.4 Oriënteren gebruik Geo-Explorer 78
4.8 INVENTARISATIE IN HET VELD 79
4.9 VERWERKING INVENTARISATIE 80
4.10 VERGELIJKING VAN WAARDEN EN ANALYSE SOORT FOUTEN 81
4.11 ANALYSE RESULTAAT KWALITATIEF ONDERZOEK 84
4.11.1 Resultaat 84
4.11.2 Bereikte nauwkeurigheid 87
4.11.3 Resultaat methode van inwinning 88
4.12 VASTGESTELDE KWALITEIT VERSUS BENODIGDE KWALITEIT 90
4.12.1 Inleiding 90
4.12.2 Werkprocessen 91
4.12.3 Kenmerken die de impact beïnvloeden 91
4.12.4 Mogelijke gevolgen bij misclassificatie 92
4.12.5 Bepaling kwaliteitseisen 94
4.13 CONCLUSIE 95

5 CONCLUSIES EN AANBEVELINGEN 98

5.1 INLEIDING 98
5.2 LITERATUURONDERZOEK 98
5.3 KWALITATIEF ONDERZOEK 99
5.4 KWANTITATIEF ONDERZOEK 99
5.5 "OVERALL" CONCLUSIE 100
5.6 AANBEVELINGEN 101

LITERATUURLIJST 103

LIJST MET GERAADPLEEGDE INTERNETSITES 105

BIJLAGE 1 INTERVIEW BIJ TOPOGRAFISCHE DIENST KADASTER 107

BIJLAGE 2 INTERVIEW BIJ HET KADASTER 114

BIJLAGE 3 INTERVIEW BIJ DE ADVIESDIENST GEO-INFORMATIE EN ICT.....119

BIJLAGE 4 INTERVIEW BIJ WATERSCHAP ZEEUWSE EILANDEN.....124

BIJLAGE 5 TOTSTANDKOMING BESTAND DIJKBEKLEDINGEN 130

BIJLAGE 6 OVERZICHT VAN GEMETEN OBJECTCODERINGEN 134

BIJLAGE 7 PROCESSCHEMA TOEVOEGEN STATUS 135

BIJLAGE 8 PROCESSCHEMA TOEVOEGEN ZICHTBAARHEID 136

BIJLAGE 9 TOEVOEGEN GEBIED AAN WATERKERINGENBESTAND 137

BIJLAGE 10 PROCESSCHEMA KIEZEN TE TOETSEN VLAKKEN 138

BIJLAGE 11 LIJST MET CODERINGEN VAN STEENBEKLEDINGEN.....140

BIJLAGE 12 HET GENEREREN VAN HET CENTROIDENBESTAND 142

BIJLAGE 13 HET GEREEDMAKEN VAN DE GEO-EXPLORER 143

BIJLAGE 14 CONVERSIE VAN GEO-EXPLORER NAAR WERKSTATION 144

BIJLAGE 15 VERWERKEN INVENTARISATIE TOT RESULTAAT 145

BIJLAGE 16 VERGELIJKEN BESTANDEN T.B.V. CONVERSIEFOUTEN 146

BIJLAGE 17 INVULLEN VAN DE SOORTEN FOUTEN 147

BIJLAGE 18 CONTROLE NIET TE VERKLAREN FOUTEN 148

BIJLAGE 19 CONFUSIONMATRIX VAN ALLE WAARNEMINGEN 149

BIJLAGE 20 CONFUSIONMATRIX VAN FOUTEN 51, 52 EN 53 150

BIJLAGE 21 BEPALEN NAUWKEURIGHEID GEO-EXPLORER 151

BIJLAGE 22 VRAGENLIJST BEPALEN BENODIGDE KWALITEIT.....152

BIJLAGE 23 INVULLEN KENMERKEN VEREISTE KWALITEIT.....156

BIJLAGE 24 VERGELIJKINGSPERCENTAGE VEREISTE KWALITEIT.....157

Disclaimer

De in deze scriptie gepresenteerde resultaten zijn gebaseerd op mijn eigen onderzoek aan de Faculteit der Aard- en Levenswetenschappen van de Vrije Universiteit Amsterdam. Alle bijstand die is verkregen zowel van individuele personen als organisaties is vermeld. Daarnaast zijn alle gebruikte gepubliceerde en ongepubliceerde bronnen opgenomen in de literatuurlijst.

Deze scriptie is niet eerder gebruikt voor het verkrijgen van een graad aan een instituut.

Getekend:

Sjaak van Popering,
Goes, 8 maart 2004

Voorwoord

Voor u ligt een onderzoek naar de toepassing van kwaliteit van geo-informatie in theorie en praktijk. Dit onderzoek vormt de afsluiting van een periode van bijna 3,5 jaar waarin ik met veel plezier aan de Vrije Universiteit Amsterdam gestudeerd heb.

Tijdens deze jaren heb ik veel mensen leren kennen en ben ik ook veel mensen dank verschuldigd. Ik wil beginnen met het UNIGIS team te bedanken, met name Mathilde, Margaret, Bart en Eleni. Bijzondere dank ben ik verschuldigd aan Henk en Alfred die mijn onderzoek op een heel plezierige wijze begeleid hebben waardoor dit jaar voorbij is gevlogen.

Naast bovenstaande wil ik ook een aantal collega's bij waterschap Zeeuwse Eilanden bedanken, te beginnen met de co-sponsor van mijn studie, Joan de Visser. Daarnaast wil ik ook Hans Brouwer, Hans Filius, Frans van der Maas en Wim Veldhuis bedanken. Ook zij zijn als leidinggevende/verantwoordelijke betrokken geweest.
Voor de input van het kwalitatieve gedeelte wil ik mijn collega's Johan van Cranenburgh en Willem Rijn bedanken. Voor de input en het meedenken ten behoeve van het kwantitatieve onderzoek gaat mijn dank uit naar Ad Beaufort en Hans van der Sande. Vooral van hun rijkelijk voorradige kennis heb ik veel gebruik mogen maken.

Tijdens het onderzoek heb ik ook input gehad van anderen. Voor het kwalitatieve onderzoek ben ik dank verschuldigd aan Gerard Vrijkotte van Topografische Dienst/Kadaster, Auke Hoekstra van het Kadaster en Adri den Boer van de Adviesdienst Geo-informatie en ICT. Voor het meedenken en voor de belangstelling wil ik ook Martin Schlüter en Hans Hainje van ESRI-Nederland en Co Drost van Arcadis Vastgoedinformatie bedanken. Door laatstgenoemde is ook de Geo-explorer ter beschikking gesteld.

Tot slot wil ik graag mijn vrouw Jeanette en zoon Finn bedanken. Zij zijn beiden een grote energie- en inspiratiebron voor mij. Ondanks dat ik getracht heb mijn studie zo weinig mogelijk mijn gezinsleven te laten beïnvloeden, geloof ik niet dat dat altijd goed gelukt is. Veel dank voor het geduld en het geloof. Hopelijk kan ik mijn zoon Finn later net zo inspireren als dat hij mij de afgelopen 3,5 jaar gedaan heeft. Om hiermee een eerste stap te zetten wil ik dit afstudeeronderzoek dan ook aan hem opdragen.

Sjaak

1 Inleiding en aanleiding

In het kader van het afronden van de UNIGIS opleiding dient voor het verkrijgen van een MSc-graad in Geografische Informatie Systemen, wetenschappelijk onderzoek verricht te worden op dit vakgebied. Om een goed afstudeeronderwerp te vinden is in eerste instantie een lange lijst opgesteld met potentiële onderwerpen. Na lang wikken en wegen is besloten om onderzoek te gaan doen naar kwaliteit van geo-informatie in theorie en praktijk. In dit eerste hoofdstuk wordt beschreven waarom nu juist dit een goed onderwerp is voor uitgebreid onderzoek en hoe aan dit onderzoek verder vorm is gegeven.

1.1 Probleemstelling

1.1.1 Doelstelling

Geo-informatie wordt binnen veel organisaties gebruikt ten behoeve van de ondersteuning van belangrijke bedrijfsprocessen. Met betrekking tot de kwaliteit van deze geo-informatie valt tijdens het inwin- of aankoopproces een aantal keuzes te maken. De mate van kwaliteit van geo-informatie bepaalt de hoogte van de kosten van de inwinning en het beheer van deze geo-informatie. Des te hoger de gewenste kwaliteit, des te hoger de kosten.
Hierdoor is voor het ondersteunen van een bedrijfsproces inzicht in de kwaliteit van een dataset noodzakelijk. Zonder dit inzicht is het niet mogelijk bewust af te wegen of een dataset voor het beoogde doel ingezet kan worden of kunnen benodigde kwaliteitseisen niet vooraf omschreven of achteraf getoetst worden.

Omdat de uitgaven van waterschappen volledig door burgers vergoed worden, dient er zeker hier zo efficiënt mogelijk met middelen te worden omgegaan. In dat kader is het belangrijk inzicht in de vereiste en aanwezige kwaliteit van geo-informatie te hebben of te verkrijgen en dient er aandacht besteed te worden aan een inwinmethodiek die erop gericht is zo efficiënt mogelijk de gewenste geo-informatie te verkrijgen.

Het doel van dit onderzoek is om een indruk te krijgen welke kwaliteitsaspecten van geo-informatie van belang zijn met betrekking tot een bepaalde dataset, in hoeverre de borging hiervan reeds geïmplementeerd is binnen diverse organisaties en hoe de vereiste kwaliteit van een aspect van geo-informatie overeen komt met de werkelijke kwaliteit.

De resultaten uit dit onderzoek kunnen dienen ter verbetering van de afstemming van kwaliteit van geo-informatie op de benodigde kwaliteit ten behoeve van het bedrijfsproces. Deze verbetering kan enerzijds leiden tot een besparing van kosten voor de aanschaf of inwinning van geo-informatie en anderzijds tot een verbetering van het bedrijfsproces doordat de juiste kwaliteit informatie gebruikt wordt. Daarnaast wordt met dit onderzoek ook een bewustwording binnen zowel de waterschapsorganisatie als binnen andere organisaties nagestreefd om het aspect “kwaliteit van geo-informatie” de vereiste aandacht te laten toekomen.

Wegens de beperkt beschikbare tijd voor dit onderzoek zal de vergelijking van de vereiste en werkelijke kwaliteit alleen worden onderzocht voor een beperkte dataset en voor één aspect van kwaliteit van geo-informatie. De dataset die onderzocht gaat worden heeft betrekking op de waterkeringen van Waterschap Zeeuwse Eilanden. Deze zijn met name interessant omdat de dataset onder andere gebruikt wordt voor het toetsen van de waterkering op veiligheid. Inzicht in de kwaliteit van de data is met name dan belangrijk en relevant. Omdat andere waterschappen ook druk bezig zijn met het toetsen van de waterkeringen is dit onderzoek voor meerdere partijen relevant.
Het kwaliteitsaspect waarop de dataset onderzocht gaat worden betreft de nauwkeurigheid van attribuutwaarden. Dit is interessant omdat dit aspect van kwaliteit van geo-informatie in tegenstelling tot bijvoorbeeld volledigheid of positionele nauwkeurigheid nog niet vaak onderzocht is. Tevens wordt met behulp van dit onderzoek een methode ontwikkeld die in de toekomst gebruikt kan worden voor het bepalen van de attribuutnauwkeurigheid van andere datasets.

1.1.2 Vraagstelling

Uit het oogpunt van efficiency en vanwege de vereiste kwaliteit ten behoeve van de toepassing dient de benodigde kwaliteit of nominale grondslag van een bepaalde geo-dataset overeen te komen met de werkelijke kwaliteit van deze dataset. Daarom luidt de hypothese van dit onderzoek als volgt:

De kwaliteit van geo-informatie wordt op een dusdanige wijze gewaarborgd dat deze aansluit op de nominale grondslag of op de vereiste kwaliteit ten behoeve van het bedrijfsproces.

Deelvragen:

- Waaruit bestaat het begrip kwaliteit van geo-informatie
- Hoe wordt kwaliteit van geo-informatie bij het Kadaster, Topografische Dienst Kadaster (TDK), Adviesdienst Geo-informatie en ICT (AGI) en Waterschap Zeeuwse Eilanden gewaarborgd
- Welke kwaliteit van geo-informatie van een bepaalde dataset is vereist voor de ondersteuning van het werkproces waarvoor deze dataset is aangeschaft of ingewonnen
- Welke werkelijke kwaliteit van geo-informatie bezit deze dataset
- In hoeverre komt de vereiste kwaliteit van geo-informatie overeen met de werkelijke kwaliteit van geo-informatie
- Is het mogelijk om, door het gebruiken van andere inwinmethodieken, het maken van fouten te voorkomen of te beperken en hierdoor een hogere kwaliteit te verkrijgen

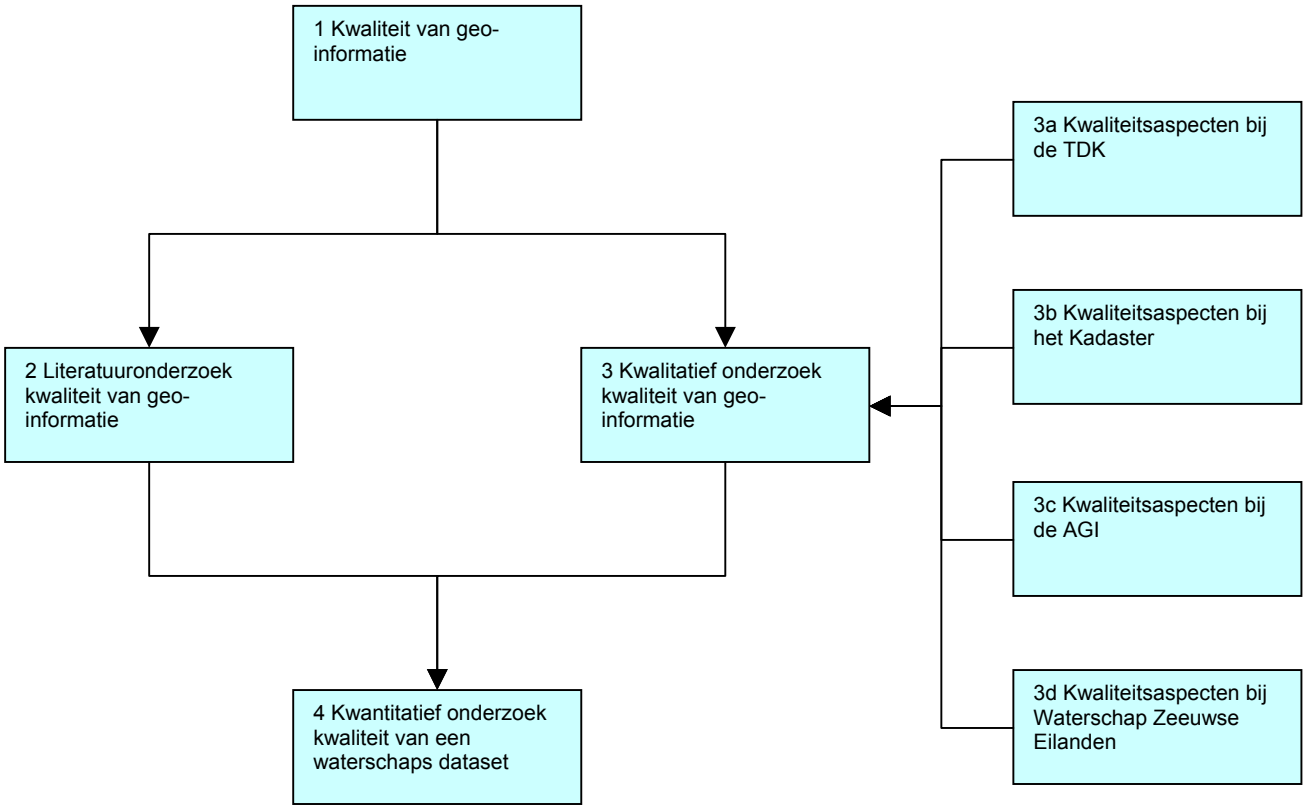
1.2 Onderzoeksontwerp

1.2.1 Methode

Het onderzoek is te kenschetsen als gedeeltelijk inventariserend en gedeeltelijk vergelijkend. Inventariserend is het gedeelte waarbij onderzocht wordt welke aspecten er bestaan rondom het begrip geo-informatie en hoe deze kwaliteit bij een aantal organisaties gewaarborgd wordt. Het vergelijkende deel bestaat uit het analyseren van de gegevens zoals die bij een organisatie in het geografische informatie systeem voorkomen en zoals die in de werkelijkheid voorkomen.

De methode van onderzoek zal bestaan uit een literatuurstudie, een kwalitatief onderzoek d.m.v. interviews en een gedetailleerd kwantitatief praktijkonderzoek. Dit laatste onderzoek zal bestaan uit een vergelijking van de geo-informatie in GIS met de werkelijkheid. Op basis hiervan zal de werkelijke kwaliteit vergeleken worden met de kwaliteit zoals die benodigd is ten behoeve van de bedrijfsprocessen.

In de volgende figuur wordt een en ander verduidelijkt:



Figuur 1 De geschematiseerde weergave van het onderzoek

De literatuurstudie (2) zal zich richten op wat er op dit gebied al onderzocht is en wat er met kwaliteit van geo-informatie bedoeld wordt. Tevens gaat het erom te onderzoeken hoe kwaliteit van geo-informatie getoetst kan worden, uit welke onderdelen de kwaliteit van geo-informatie is opgebouwd en welke standaarden er bestaan. Deze studie zal uitgevoerd worden met behulp van internet en op universiteitsbibliotheken.

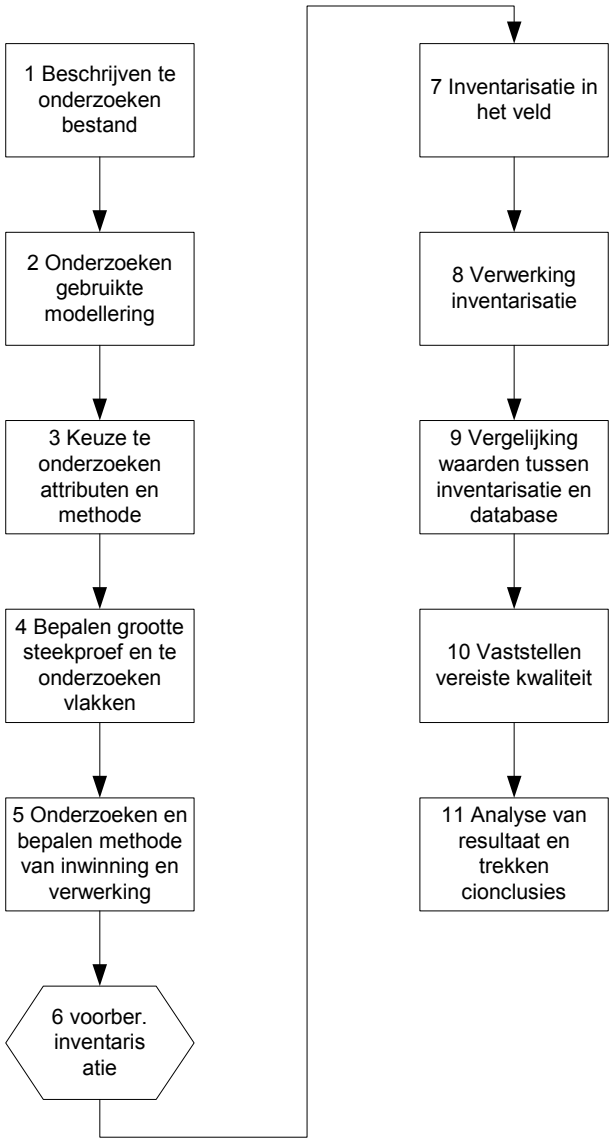
Het kwalitatief onderzoek (3) zal zich richten op de waarborging van kwaliteit in de praktijk (3a-3d). In dit onderdeel wordt in de praktijk onderzocht hoe er momenteel omgegaan wordt met de kwaliteit van geo-informatie. De Topografische dienst kadaster (TDK), het Kadaster en de Adviesdienst Geo-informatie en ICT zijn grote leveranciers en beheerders van geo-informatie in Nederland. Met name daarom wordt onderzocht hoe deze organisaties omgaan met het begrip kwaliteit van geo-informatie. Bij waterschap Zeeuwse Eilanden zal dit kwalitatieve onderzoek ook uitgevoerd worden. Bij deze organisatie wordt veel geo-informatie gebruikt en wordt ook behoorlijk wat ingewonnen. Door middel van interviews zal dit kwalitatieve onderzoek uitgevoerd worden.

Het kwantitatief onderzoek (4) zal uitgevoerd worden bij Waterschap Zeeuwse Eilanden en zal bestaan uit het analyseren van de attribuut nauwkeurigheid van een dataset. Deze zal vergeleken worden met de veronderstelde attribuut nauwkeurigheid. De dataset die hiervoor gebruikt zal gaan worden betreft de steenbekledingen van de waterkeringen. Hieronder is een foto weergegeven met hierop die steenbekledingen.

Foto 1 De steenbekleding van een waterkering

Met de attribuut nauwkeurigheid wordt de nauwkeurigheid van de waarden van de attributen van een object bedoeld. Concreet betekent dit het onderzoeken van de juistheid van de attribuutwaarden met betrekking tot de steenbekledingen van de waterkeringen.

In de volgende figuur staan de verschillende stappen in dit kwantitatieve onderzoek.



Figuur 2 De verschillende stappen in het kwantitatieve onderzoek

Hieronder wordt ieder stap in het kwantitatieve onderzoek toegelicht:

1. *Beschrijven te onderzoeken bestand*

   In deze stap wordt het te onderzoeken bestand beschreven. Belangrijke zaken zijn hier:

   - Wanneer en hoe is het bestand ingewonnen
   - Welke attributen zijn aanwezig in dit bestand en in welke categorieën vallen deze (ordinaal, nominaal, etc)
   - Aantallen vlakken en de verdeling in soorten binnen deze vlakken

2. *Onderzoeken gebruikte modellering*
   Voor het vertalen van de werkelijkheid naar een databestand wordt een model van deze werkelijkheid ontworpen. Om de attribuut nauwkeurigheid te kunnen analyseren dient het gebruikte model van de werkelijkheid bekend te zijn.

3. *Van welke attributen gaat de nauwkeurigheid geanalyseerd worden en welke methode dient hiervoor gebruikt te worden.*
   In het de onderzoeken bestand zijn diverse attributen aanwezig. De meest belangrijke gaan getest worden. Voor verschillende categorieën attributen zijn verschillende manieren om de kwaliteit hiervan vast te stellen.

4. *Bepalen steekproef*
   Voor het bepalen van de mate van de attribuutnauwkeurigheid dient van een aantal vlakken de waarden in de database vergeleken te worden met de werkelijkheid, met inachtneming van het modelleringproces. Hiertoe dienen een aantal zaken onderzocht te worden:
   - Hoeveel vlakken dienen er per soort onderzocht te worden om een representatieve steekproef te hebben. Eventueel kan dit ook nog per gebied verschillen.
   - Welke vlakken gaan onderzocht worden.

5. *Bepalen methode van inventarisatie en verwerking*
   Voor de inventarisatie van de waarden in de werkelijkheid zijn een aantal mogelijkheden. Iedere mogelijkheid heeft voor en nadelen. De volgende eisen worden aan de methode gesteld:
   - De inventarisatie moet een hoge juistheidfactor bezitten
   - De inventarisatie moet in relatief korte tijd uitgevoerd kunnen worden
   - De verwerking van de inventarisatie en de vergelijking met de database moet op korte termijn uitgevoerd kunnen worden en moet een hoge juistheidfactor bezitten.

6. *Voorbereiding op inventarisatie*
   Installeren software, converteren data, maken domeinwaarden, testen inwinning en verwerking

7. *Inventarisatie in het veld*
   Samen met een waterkeringendeskundige waarden inventariseren in het veld.

8. *Verwerking van inventarisatie*

9. *Vergelijking van geïnventariseerde waarden met de waarden in de database*
   Hiervoor is het belangrijk de goede methoden te gebruiken zoals bijvoorbeeld een confusion matrix, Cohen's Kappa of een RMSE.

10. *Vaststellen van de vereiste kwaliteit*
    Door middel van interviews en vragenlijsten zal deze vastgesteld worden.

11. *Analyse van resultaat en trekken conclusies*

1.2.2 Object

De kwaliteit van geo-informatie is van belang voor alle geo-informatie die gebruikt wordt binnen een organisatie. Binnen vele organisaties wordt voor vele doeleinden en op vele plaatsen geo-informatie gebruikt. Omdat in het kader van dit onderzoek de tijd ontbreekt om heel breed een onderzoek te wijden aan de kwaliteit van geo-informatie, wordt er beperkt tot een kwalitatief onderzoek en een gedetailleerd kwantitatief onderzoek. Het kwalitatieve onderzoek zal uitgevoerd worden bij de Topografische Dienst Kadaster, het Kadaster, de Adviesdienst Geo-informatie en ICT en waterschap Zeeuwse Eilanden. Het kwantitatieve onderzoek zal alleen uitgevoerd worden bij waterschap Zeeuwse Eilanden.

1.3 Opbouw van het rapport

In dit eerste hoofdstuk is beschreven wat er in dit onderzoek onderzocht gaat worden en waarom dit allemaal gedaan wordt.
In hoofdstuk 2 wordt beschreven wat kwaliteit van geo-informatie nu eigenlijk betreft, welke standaarden er gebruikt worden en op welke wijze kwaliteit van geo-informatie vastgesteld en getoetst kan worden.
Hoofdstuk 3 gaat in op het kwalitatieve onderzoek. Hierin worden de bevindingen van de interviews met de diverse organisaties beschreven en worden de verschillen tussen de organisaties inzichtelijk gemaakt.
Het kwantitatieve gedeelte van het onderzoek wordt uitgewerkt in hoofdstuk 4. Hierin wordt zowel beschreven welk bestand nu eigenlijk getoetst is, hoe deze toetst uitgevoerd is en wat de bevindingen zijn. Dit wordt mede gedaan op basis van een vergelijking met de benodigde kwaliteit. Het vaststellen van deze benodigde kwaliteit wordt ook in dit hoofdstuk uitgewerkt.

Tot slot worden in hoofdstuk 5 alle conclusies en bevindingen op een rij gezet en worden aanbevelingen gegeven.

2 Kwaliteit en geo-informatie

2.1 Inleiding

In dit hoofdstuk wordt uitgebreid ingegaan op kwaliteit van geo-informatie. Diverse aspecten komen hierbij aan de orde zoals onder andere wat is kwaliteit precies, wat wordt er onder geo-informatie verstaan, welke aspecten van kwaliteit van geo-informatie kunnen onderscheiden worden.
Uit de literatuurstudie naar bovenstaande blijkt dat er diverse aanpakken en definities zijn met betrekking tot de aspecten rondom kwaliteit van geo-informatie. Ook op deze verschillende aspecten wordt dieper ingegaan.

Belangrijk bij het begrip kwaliteit is dat het gemeten kan worden. Pas na meting kan er iets over de kwaliteit gezegd worden. Meting van kwaliteit kan echter op diverse manieren plaatsvinden. Belangrijk is echter dat de uitkomst van de kwaliteitsmeting gebruikt kan worden als toets om te bepalen of bepaalde informatie geschikt is voor gebruik binnen een bepaalde toepassing. In dit hoofdstuk zullen ook de diverse methoden die gebruikt kunnen worden voor het meten van kwaliteit behandeld worden.

2.2 Kwaliteit en Geo-informatie

2.2.1 Geo-informatie

Het begrip geo-informatie wordt als verzamelterm gebruikt voor bestaande termen als vastgoedinformatie, geografische informatie en ruimtelijke informatie. Gesteld kan worden dat geo-informatie, informatie is over geo-objecten. Dit begrip kan conform Heres (1998) in een tweetal zaken onderverdeeld worden:

- Informatie over de positie, vorm en oriëntatie van objecten;
- Informatie over de classificatie, identificatie en kenmerken van deze objecten

In het “Referentiemodel Kwaliteit van geo-informatie” van de RAVI wordt geo-informatie aangeduid als de overkoepelende term voor vastgoed, geografische, geometrische en ruimtelijk informatie.

2.2.2 Kwaliteit

Van het begrip kwaliteit bestaan een heleboel verschillende definities. Hierna volgt een aantal voorbeelden:

- De mate waarin iets aan kenbaar gemaakte behoeften voldoet (Heres, 1998)
- De mate waarin aan de behoeften en verwachtingen wordt voldaan.
- Het geheel van karakteristieken van een product die bijdragen aan het in staat zijn om aan vastgestelde en impliciete eisen te voldoen (ISO 8402)
- “Fitness for use” (Chrisman 1983)
- “How closely data values represent reality given the chosen model for representing that reality” (Salge 1995)

Belangrijk bij het begrip kwaliteit is dat het gemeten kan worden. Pas na meting kan er iets over de kwaliteit gezegd worden. Meting van kwaliteit kan echter op diverse manieren plaatsvinden. Belangrijk is echter dat de uitkomst van de kwaliteitsmeting gebruikt kan worden als toets om te bepalen of bepaalde informatie geschikt is voor een bepaalde toepassing.

Het verkrijgen van inzicht in de kwaliteitsaspecten is onder andere om de volgende redenen belangrijk:

- *Voor het bepalen van geschiktheid ten behoeve van een bepaalde bewerking.*
  Productaansprakelijkheid heeft hier ook mee te maken. Door het duidelijk maken van de kwaliteit van een bepaalde dataset worden ook de mogelijkheden en wellicht nog belangrijker de beperkingen, van het bestand in beeld gebracht.

- *Voor het toetsen van een bestand aan de gestelde eisen.*
  Bij de inwinning van datasets dient door middel van controles onderzocht te worden of de ingewonnen data ook voldoet aan de vooraf gestelde eisen. Inzicht in de betrouwbaarheid van en de methodes die gebruikt zijn om de kwaliteit vast te stellen is dan noodzakelijk. Deze kwaliteit wordt ook wel meta-kwaliteit genoemd.

- *Door de complexiteit van bestanden is het steeds moeilijker om zonder kwaliteitsaspecten inzicht in de kwaliteit te krijgen.*
  Ten tijde van analoge producten kon door middel van bijvoorbeeld de ondergrond en bijschriften op kaart of tekening bepaald worden wat de kwaliteit van een dataset betrof.

Met digitale producten is dat allemaal een stuk lastiger. Bijschriften zijn vaak niet aanwezig en tevens is het product op vele schalen bruikbaar.

- *Voor het stellen van interne doelstellingen*
  Ten behoeve van certificering of andere interne kwaliteitsdoelstellingen is inzicht in de kwaliteit noodzakelijk om te kunnen meten of aan de certificering of doelstelling voldaan wordt.

- *Voor levering aan derden*
  Ten behoeve van bijvoorbeeld regelgeving kan het zijn dat geo-informatie aan derden geleverd wordt. Inzicht in de kwaliteit is dan ook essentieel.

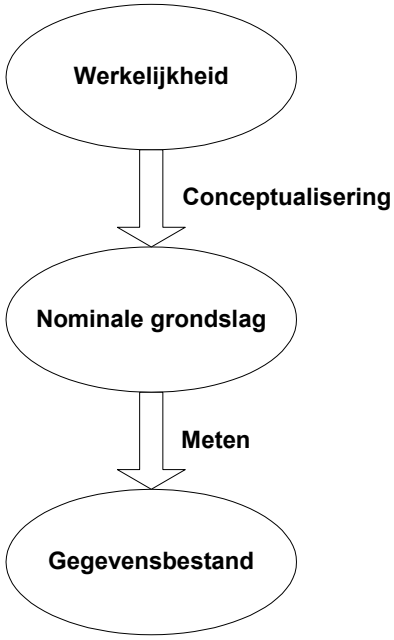
Er bestaan veel parameters die de kwaliteit van geo-informatie beschrijven. In veel standaarden worden de aspecten van kwaliteit van geo-informatie benoemd. Hiertussen zitten verschillen. In paragraaf 2.3 worden deze verschillende standaarden verder uitgewerkt en met elkaar vergeleken.

2.2.3 Modellering

Omdat de werkelijkheid te complex is om in een computer te representeren worden er vereenvoudigde modellen van de werkelijkheid gemaakt. Afhankelijk van de toepassing waarvoor de dataset bedoeld is, kunnen verschillende modellen gemaakt worden. Deze modellen kunnen op een aantal punten verschillen:

- De nauwkeurigheid waarmee een object gemodelleerd wordt, ingestoken op de voor de toepassing noodzakelijke nauwkeurigheid.
- De keuze voor opname van soort objecten in het model die van belang zijn voor de toepassing

Voor het analyseren van de kwaliteit van een bestand kan niet zomaar de werkelijkheid vergeleken worden met het database bestand. Tijdens de realisatie van het bestand zijn er allerlei keuzes gemaakt om de werkelijkheid in een database te representeren. Dit betreft voornamelijk het vereenvoudigen van de werkelijkheid zodat dit in een database past. De vertaling van de werkelijkheid naar een modelmatige omgeving heet modelleren (zie onderstaande figuur).



Figuur 3 De stappen van modellering van de werkelijkheid in een database

2.3 Kwaliteitsaspecten

In deze paragraaf worden diverse internationale en nationale standaarden en uitwerkingen van kwaliteit van geo-informatie behandeld. Op basis hiervan wordt in paragraaf 2.5 een algemene definitie van kwaliteitsaspecten gedefinieerd.

2.3.1 Voornorm - geografische informatie- gegevensbeschrijving – kwaliteit (VNK)

Eind 1998 is de Voornorm – geografische informatie – gegevensbeschrijving – kwaliteit ontwikkeld en is door de technische commissie CEN/TC287 voorbereid.
Deze voornorm is opgesteld om in een drietal behoeftes te voorzien. Dit betreft de volgende:

- Het kunnen beschrijven van de kwaliteit van een dataset voor leveranciers van geo-informatie.
- Voor het kunnen bijsluiten van kwaliteitsinformatie bij levering aan een gebruiker.
- Voor het kunnen beschrijven van de kwaliteitsbehoeftes van gebruikers.

In deze voornorm worden verschillende kwaliteitsaspecten onderscheiden. De volgende aspecten worden onderscheiden:

- *"Lineage"*
  Dit is een tekstuele beschrijving van de afkomst van een dataset. De onderdelen die hier onderscheiden worden zijn:
  → Productie
  Hierbij gaat het zowel om de organisatie van de producent alsmede de datum van de totstandkoming van het product.

  → Bron en bewerkingsproces
  Hierbij gaat het om informatie over hoe, door wie, wanneer en waarom de productie is uitgevoerd.

- *Het gebruik*
  Een beschrijving voor welke toepassingen de geografische dataset reeds gebruikt is. De volgende onderdelen worden hier onderscheiden:
  → De naam van de organisatie die de dataset gebruikt heeft
  → Het soort gebruik waarvoor de dataset ingezet is
  → Voorwaarden of beperkingen die tijdens het gebruik ondervonden zijn
  → De datum waarop het gebruik heeft plaatsgevonden

- *Homogeniteit*
  Een tekstuele en kwalitatieve omschrijving van verwachte of uit onderzoeken blijkende uniformiteit van kwaliteitsparameters.

- *Kwaliteitsparameters die de dataset beschrijven in relatie tot haar nominale grondslag*
  Hierbinnen worden de volgende onderscheiden:

  → Positionele nauwkeurigheid
  - Relatieve horizontale nauwkeurigheid
  - Relatieve verticale nauwkeurigheid
  - Gebruikers gedefinieerd
    Hierbij zijn diverse mogelijkheden. Voorbeelden zijn onder andere absolute nauwkeurigheid, nauwkeurigheid van de pixel positie en de relatieve positie nauwkeurigheid van een entiteit ten opzichte van aanliggende entiteiten.

→ Semantische nauwkeurigheid

Dit aspect kan opgebouwd worden uit de volgende onderdelen:

- De nauwkeurigheid van classificatie conform de entiteitklasse
- Overeenstemming van attribuut van een entiteit
- Gebruikersgedefinieerd

→ Temporele nauwkeurigheid

Deze kan op diverse wijzen gedefinieerd worden. De verschillende wijzen zijn:

- De nauwkeurigheid van het meten van de tijd in bijvoorbeeld uren, weken of jaren.
- De datum van de laatste update
- De hoeveelheid wijzigingen per tijdseenheid
- De gemiddelde tijd tussen de wijziging in de nominale grondslag en verwezenlijking in de dataset
- De status van de data in relatie tot de tijd. Is de data actueel of gedateerd.

→ Volledigheid

De volledigheid kan ook op diverse wijzen gerepresenteerd worden. Conform de voornorm dient in een van de volgende inzicht gegeven te worden:

- Omissie
- Overcompleetheid
- Gebruikersgedefinieerd

→ Logische consistentie

Deze kwaliteitsparameter beschrijft de mate waarin een geografische dataset voldoet aan zijn specificaties voor wat betreft de interne structuur.

De verschillende onderdelen die hieronder vallen zijn:

- Geometrische consistentie

  Een percentage van in hoeverre de geometrie voldoet aan de opgestelde specificaties. Voorbeelden hiervan zijn:

  - Er mogen geen dubbele punten in het bestand voorkomen
  - Binnen een straal van 2 meter mogen geen meerdere punten voorkomen
- Topologische consistentie

  Een percentage van in hoeverre de geometrie voldoet aan de opgestelde specificaties. Voorbeelden hiervan zijn:

  - Alle vlakken dienen gesloten te zijn.
  - Alle vlakken dienen een unieke identificatiecode te bezitten

- Semantische consistentie
  Dit betreft een percentage van ruimtelijke relaties tussen verschillende soorten entiteiten die voldoet aan de gestelde specificaties. Voorbeelden hiervan zijn:
  - Een waterkering ligt aan het water
  - Een benzinestation ligt aan een weg
- Formaat consistentie
  Voldoet het formaat van de geografische dataset aan de standaarden.
- De validiteit van waarden van een of meer attributen ten opzichte van haar gedefinieerd domein. Hierin kunnen zich verschillende situaties voordoen:

Binnen de voornorm wordt ook het begrip metakwaliteit onderkend. Metakwaliteit betreft hier informatie over de definiëring van de kwaliteit. Binnen de voornorm worden de volgende aspecten hieronder verstaan:

- *De mate van vertrouwelijkheid van de kwaliteitsinformatie*
  Dit kan zowel in tekst als in statische indicatoren weergegeven worden.

- *De mate van betrouwbaarheid van de kwaliteitsinformatie*

- *Een beschrijving van de methodieken die gebruikt zijn om de kwaliteitsinformatie te bepalen.*

- *Het in beeld brengen van het gevolg van het abstraheren van de werkelijkheid naar de nominale grondslag*

2.3.2 NCDCDS

Het “National (VS) Committee for Digital Cartographic Data Standards” (NCDCDS) is in 1982 gestart met het ontwikkelen van een standaard betreffende kwaliteitsaspecten van digitale cartografische data. In 1988 is een eerste voorstel hiervoor opgeleverd. Later is dit eerste voorstel opgenomen in de in 1992 ontwikkelde “Spatial Data Transfer Standard”. Zie hiervoor de volgende paragraaf.

2.3.3 Spatial Data Transfer Standard (SDTS)

The SDTS is bedoeld voor en biedt de mogelijkheid om ruimtelijke data te kunnen uitwisselen tussen verschillende computerplatforms. SDTS is bekrachtigd door de National Institute of standards en Technology (NIST). Dit heeft geleid tot een Federal Information Processing Standard (FIPS 173) die gebruikt wordt door onder andere U.S. Geological Survey, U.S. Census Bureau en het Army Corps of Engineers.

In de SDTS wordt de noodzakelijkheid van kwaliteit van data onderkend vanuit het perspectief van of een bepaalde dataset geschikt is voor een bepaalde toepassing. Van hieruit wordt ook het begrip kwaliteit van geo-informatie beschreven. Binnen de SDTS wordt hieronder de volgende parameters onderkend:

- *Lineage*
  Hierbinnen kan informatie zoals bron, de verschillende procedurestappen en de bijhoudingsactiviteiten die uitgevoerd zijn, vermeld worden.

- *Positionele nauwkeurigheid*
  Hierbinnen wordt informatie verstrekt over in hoeverre de coördinaten in de dataset overeenkomen met de werkelijke positie. Er wordt onderscheid gemaakt tussen horizontale en verticale coördinaten.

- *Attribuut nauwkeurigheid*
  Informatie over de nauwkeurigheid van attribuutdata.

- *Logische consistentie*
  Dit betreft een indicatie voor de grafische kwaliteit en de topologische integriteit van de dataset.

- *Volledigheid*
  Deze parameter betreft hier informatie over selectiecriteria voor het wel opnemen van entiteiten in de dataset en het beschrijven van andere gebruikte definities of gebruikte regels. Binnen volledigheid wordt ook het verschil tussen relatie-volledigheid en object-volledigheid onderscheiden.

2.3.4 DIGEST

The digital Geographic Information Exchange Standard (DIGEST) is een internationale standaard die ontwikkeld is met het doel om digitale geografische data tussen diverse geografische informatie systemen uit te kunnen wisselen.

De ontwikkeling van DIGEST is in 1983 gestart en is door de jaren heen doorontwikkeld. Sinds het begin van de 90-er jaren van de vorige eeuw wordt DIGEST gebruikt bij samenwerkingsprojecten tussen landen. Inmiddels is DIGEST de NATO standaard geworden.

Binnen DIGEST wordt ook het onderwerp kwaliteit behandeld. Als reden hiervan wordt aangegeven dat het noodzakelijk is om inzicht in de kwaliteit te hebben van uit te wisselen data omdat dat noodzakelijk is voor het integreren in de ontvangende database zonder dubbelzinnigheid.

Onderstaande verschillende onderdelen van kwaliteit worden door DIGEST onderscheiden en voorgeschreven.

- *Specificatie*
  Hierbinnen wordt indien nodig aangegeven welke editie gebruikt is voor het genereren van het uitwisselingsbestand en welke versie van DIGEST hiervoor gebruikt is.

- *Bron*
  De bron beschrijft welke databron gebruikt is voor het genereren van het uitwisselingsbestand. Verder worden hier ook naam, brondatum, producent en controle informatie toegevoegd.

- Positionele nauwkeurigheid
  Hierbinnen dient een indicatie van zowel de horizontale als verticale nauwkeurigheid weergegeven te worden. Dit gebeurt door zowel voor de XY-richting als de Z-richting een standaardafwijking in te vullen.

- *Attribuut nauwkeurigheid*
  Hierbinnen wordt de nauwkeurigheid en betrouwbaarheid van attributen weergegeven. Deze worden voor een kwalitatief attribuut weergegeven door de procentuele aantallen

correcte attribuutwaarden. Voor een kwantitatief attribuut wordt de nauwkeurigheid weergegeven door de standaardafwijking.

- *Actualiteit*
  Dit betreft de datum waarbinnen data in de database is gemodificeerd of ingebracht. Deze wordt in de notatie Jaarmaanddag (jjjjmmdd) weergegeven. Eventueel is nog een toevoeging van uur mogelijk. De notatie wordt dan jjjjmmdduu

- *Logische consistency*
  Dit betekent dat de logische regels van structuur en attributen zijn gevolgd en het beschrijft de overeenkomst comptabiliteit tussen betreffende data en andere data van de dataset.

- *Object compleetheid*
  Dit betreft de mate waarin objecten zijn opgenomen in de dataset.

- *Attribuut compleetheid*
  De mate waarin attributen zijn opgenomen in de dataset.

Bij de datakwaliteit onderdelen worden vervolgens ook nog de “clipping indicator”, “security classification” en “releasability” aangegeven. Niet dat dit onderdelen van kwaliteit betreffen maar vanwege de gelijkenis die zij vertonen met dit soort onderwerpen.
Met de clipping indicator wordt aangegeven hoe vaak betreffend object “geclipped” is. De waarde 0 betekent dat een object niet “geclipped” is.
Binnen de “security classification” kan de mate van beveiliging worden aangegeven voor het ongeautoriseerd gebruik van betreffend object. De waarden kunnen varieren van T (“top secret”) tot U (“unclassified”).
Met “releasability” kan informatie opgeslagen worden die betrekking heeft op de uitgevoerde distributie of op de versie.

De onderdelen van kwaliteit worden op diverse niveaus opgeslagen. Binnen DIGEST worden de volgende niveaus onderkend:

- Volume niveau
- Dataset niveau
- Object niveau
- Topologisch / ruimtelijk data record niveau

- Attribuut niveau

In onderstaande tabel is aangegeven op welk niveau welke kwaliteitsparameter beschreven worden.

Onderdeel kwaliteit	Volume niveau	Dataset niveau	Object niveau	Topologisch niveau	Attribuut niveau
Specificatie	X	X			
Bron		X	X	X	X
Positionele nauwkeurigheid		X	X	X	
Attribuut nauwkeurigheid					X
Actualiteit		X	X	X	X
Logische consistency		X			
Attribuut compleetheid		X			
Object compleetheid		X			

Tabel 1 De beschrijvingsniveaus van de verschillende kwaliteitsparameters binnen DIGEST

2.3.5 Referentiemodel Kwaliteit van geo-informatie (RAVI)

Het referentiemodel kwaliteit van geo-informatie is in 1994 (?) een eerste aanzet geweest van de RAVI om het Europese kwaliteitsmodel de CEN/TC287 te vertalen naar een Nederlandse variant. De aspecten die in dit referentiemodel onderscheiden worden zijn de volgende:

- *Idealisatie*
  Hierbij gaat het hoe goed een object uit de werkelijkheid geabstraheerd kan worden in de nominale grondslag. De onderdelen hierbij zijn:
  → De mate van generalisatie
  → De meetkundige precisie waarmee een object gedefinieerd is
  → De herkenbaarheid van een object.

- *Bronvermelding*
  Het betreft hier de geschiedenis in de volle breedte. De volgende onderdelen kunnen hierbinnen onderscheiden worden:

→ Beschrijving bronmateriaal (tijdstippen en methodes van inwinning)
→ Aangebrachte transformaties en conversies (soort, functie en parameters)
→ Verantwoordelijke organisatie (naam en adres van producent of bewerker)

- *Nauwkeurigheid*
  In dit geval heeft het aspect nauwkeurigheid betrekking op waarnemingsfouten die gemaakt worden bij de inwinning van gegevens over objecten. Het beschrijft de kans op een juiste benoeming van een waarde. De volgende onderdelen worden onderscheiden:

  → Beschrijvende informatie
  Deze kan vervolgens weer onderverdeeld worden in:
  - Discrete waarden
  - Continue waarden

  → Geometrische informatie
  Deze kan vervolgens weer onderverdeeld worden in:
  - Continue waarden (lijnen, punten en vlakken)
  - Discrete waarden (pixels en grid-punten)

  → Grafische informatie
  In het Europese model is aan dit onderdeel geen aandacht besteed.

- *Volledigheid*
  Het aspect volledigheid geeft de kans weer dat een onderdeel van de nominale grondslag ook daadwerkelijk in het gegevensbestand voorkomt en dat een onderdeel van het gegevensbestand ook daadwerkelijk in de nominale grondslag voorkomt. Er worden diverse onderdelen van volledigheid onderscheiden:

  → Overcompleetheid
  → Omissie
  → Foutief

- *Consistentie*

  → Statische consistentie
  In hoeverre voldoet het gegevensbestand aan de gestelde voorwaarden in de gegevensspecificatie.

→ Dynamische consistentie
In hoeverre levert het proces op wat ervan verwacht wordt.

- *Betrouwbaarheid*
  Dit betreft twee onderdelen:
  → Een maat voor hoe goed de waarden van de kwaliteitsaspecten overeenkomen met de actuele situatie in het gegevensbestand.
  → Een maat voor hoe groot de kans is dat er een fout ontdekt wordt.

- *Mutatiefrequentie*
  Dit aspect betreft de maat voor de veranderingen in een bepaald tijdsverloop. Het is een maat voor de juistheid van de kwaliteitsparameters op een bepaald moment.

2.4 Verschillen tussen de standaarden

In de volgende tabel worden van de verschillende standaarden de onderkende kwaliteitsaspecten naast elkaar gepresenteerd en kunnen daarom goed vergeleken worden.

	Standaard			
	Spatial Data Transfer Standard	DIGEST	Referentiemodel kwaliteit van geo-informatie	Voornorm-geografische informatie-gegevensbeschrijving-kwaliteit
Kwaliteitsparameters				
Positionele nauwkeurigheid	X	X	X	X
Attribuut nauwkeurigheid	X	X	X	X
Temporele nauwkeurigheid		X	X	X
Volledigheid	X	X	X	X
Logische consistentie	X	X	X	X
Overige kwaliteitsindicatoren				
Overzicht van gebruikers				X
Metakwaliteit			X	X
Lineage	X	X	X	X
Homogeniteit				X

Tabel 2 De onderkende kwaliteitsaspecten van de verschillende standaarden

Met betrekking tot de verschillende standaarden vallen een aantal zaken op:

- Op het onderkennen van de temporele nauwkeurigheid bij SDTS na worden binnen de verschillende standaarden dezelfde kwaliteitsparameters onderkend.

- Binnen de verschillende kwaliteitsparameters worden binnen de verschillende standaarden op verschillende detailniveaus invulling gegeven. Een voorbeeld om dit te illustreren is de invulling van de positionele nauwkeurigheid van de VNK. Deze is veel gedetailleerder dan de invulling van de positionele nauwkeurigheid van SDTS. De invulling van VNK is zeer gedetailleerd en maakt bijvoorbeeld onderscheid tussen absolute en relatieve nauwkeurigheid.

- Niet bij alle standaarden wordt er onderscheid gemaakt in kwaliteitsparameters die iets zeggen over de kwaliteit van het bestand of die informatie geven over de totstandkoming

van het bestand. Op basis van de laatstgenoemde categorie kan vanzelfsprekend ook een oordeel worden geveld over de toepasbaarheid van een dataset.

- De verschillende kwaliteitsparameters kunnen op diverse niveaus, bijvoorbeeld object of datasetniveau, ingevuld worden. De meeste standaarden geven hier geen invulling aan.

- Dezelfde kwaliteitsparameters of beschrijvingen worden verschillende namen gegeven. Voorbeelden hiervan zijn de volgende:
    - In de VNK wordt semantische nauwkeurigheid onderkend. De omschrijving hiervan is echter grotendeels dezelfde als het bij de andere standaarden onderkende aspect “attribuut nauwkeurigheid”.
    - Binnen de diverse standaarden wordt er zowel over temporele nauwkeurigheid als over actualiteit gesproken.

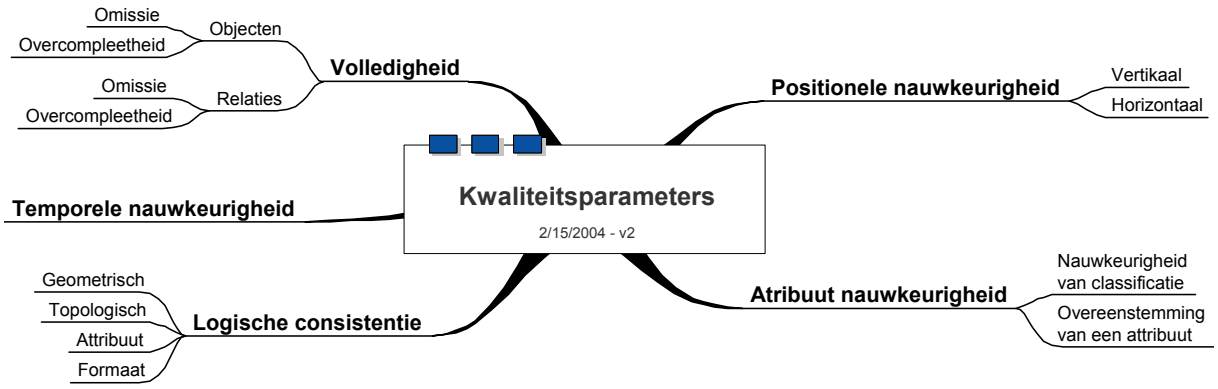
2.5 Aspecten van kwaliteit van geo-informatie

2.5.1 Algemeen

Vanwege de verschillen tussen de diverse standaarden is hieronder een overzicht gegeven van de verschillende aspecten van kwaliteit. Dit overzicht is gemaakt op basis van de overeenkomsten tussen de verschillende standaarden.
Zoals reeds eerder in dit hoofdstuk is beschreven is de belangrijkste reden voor het inzicht verlenen in kwaliteit van geo-informatie, het bepalen of een dataset geschikt is voor een bepaalde toepassing. Het zogenaamde bepalen van de “fitness for use”. Mede hierdoor zijn het niet alleen kwaliteitsparameters die van belang zijn maar zijn er ook andere aspecten van kwaliteit die bij kunnen dragen aan het bepalen van inzicht in de “fitness for use”. Deze categorie aspecten zijn samengevoegd onder de noemer “overige kwaliteitsaspecten”.

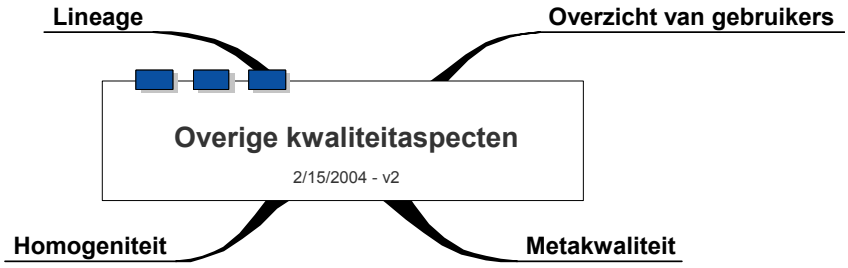
2.5.2 Kwaliteitsparameters

In onderstaande figuur is weergegeven welke kwaliteitsparameters onderkend worden. Voor wat betreft naamgeving is er aangesloten op de meest voorkomende binnen de standaarden.



Figuur 4 Schematische weergave van kwaliteitsparameters

2.5.3 Overige kwaliteitsaspecten



Figuur 5 Schematische weergave van overige aspecten van kwaliteit

2.6 Testen van de verschillende kwaliteitsparameters

In deze paragraaf wordt aangegeven op welke wijze een kwaliteitsparameter getest kan worden en op welke wijze een parameter weergegeven kan worden.

<u>Positionele nauwkeurigheid</u>

De positionele nauwkeurigheid van een object of objecten binnen een dataset kunnen bepaald worden door de objecten te vergelijken met een dataset die een hogere nauwkeurigheid bezit of door de objecten te vergelijken met de werkelijkheid. De mate van nauwkeurigheid wordt weergegeven door de standaardafwijking of de “root mean square error” (RMSE). Zoals weergegeven is in figuur 4 kan deze nauwkeurigheid zowel betrekking hebben op de verticale of horizontale positionering.

Attribuut nauwkeurigheid

De attribuutnauwkeurigheid kan bestaan uit de mate waarin een object correct geclassificeerd is of de mate waarin een attribuut overeenkomt met de werkelijkheid. Afhankelijk van of het attribuut betrekking heeft op een kwalitatief of kwantitatief element, kan de mate van attribuutnauwkeurigheid verschillend weergegeven worden. Voor een kwantitatief attribuut kan de nauwkeurigheid weergegeven worden door een standaardafwijking of een root mean square error. Voor een kwalitatief attribuut kan de nauwkeurigheid weergegeven worden in procentueel correcte waarden of door middel van Cohen's kappa. In paragraaf 4.4.2 wordt het weergeven van kwalitatieve attributen verder uitgewerkt.

Volledigheid

De volledigheid van een dataset, voor zowel objecten als relaties kan weergegeven worden in percentages aanwezig ten opzichte van het totaal. Toetsing hiervan kan plaatsvinden door het onderzoeken van de volledigheid van een representatieve steekproef van het gehele bestand.

Overige parameters

De parameters logische consistentie en temporele nauwkeurigheid kunnen zeer divers van aard zijn en hierdoor op diverse wijzen weergegeven worden. Voor bijvoorbeeld de temporele nauwkeurigheid is het mogelijk om deze weer te geven door middel van een aantal verschillende items. Reeds bij de VNK worden er diverse weergegeven.
Ook de parameter logische consistentie bestaat uit diverse aspecten die ieder op een verschillende wijze, in de vorm van tekst, beschreven worden. Toetsing hiervan kan plaatsvinden door te controleren of deze tekst inderdaad overeenkomt met de werkelijkheid.

2.7 Conclusie

In voorgaand hoofdstuk is onderzocht wat kwaliteit en geo-informatie nu eigenlijk zijn en waarom het noodzakelijk is om inzicht in de kwaliteit van een dataset te bezitten. De voornaamste redenen zijn om te kunnen bepalen of een dataset geschikt is voor een bepaalde toepassingen en om na de inwinning van een dataset te kunnen controleren of deze conform de specificaties is uitgevoerd.

Verder zijn er in dit hoofdstuk ook een viertal standaarden behandeld. Opmerkelijk hierbij is dat er in eerste instantie veel overeenkomsten tussen deze standaarden lijken te bestaan. Na een grondige bestudering blijken er toch grote verschillen in aanwezig te zijn.
Vooral de Voornorm – geografische informatie – gegevensbeschrijving – kwaliteit is zeer uitgebreid en biedt voldoende houvast om op een juiste wijze kwaliteit van geo-informatie te kunnen opslaan.

Tot slot is er weergegeven hoe de diverse parameters die een indicatie geven voor de kwaliteit weergegeven en getoetst kunnen worden. Naast deze parameters is ook inzicht in de metakwaliteit, huidige gebruikers, homogeniteit en de “lineage” van een bestand zeer waardevol om inzicht in de kwaliteit te bewerkstelligen.

3 Kwaliteitsaspecten van geo-informatie bij Nederlandse organisaties

3.1 Inleiding

In het voorgaande hoofdstuk zijn een aantal standaarden behandeld die ingaan op het begrip kwaliteit van geo-informatie. In dit hoofdstuk wordt voor een viertal organisaties beschreven hoe zij invulling geven aan kwaliteit van geo-informatie.
Van deze vier organisaties zijn de Topografische dienst kadaster, het Kadaster en de Adviesdienst Geo-informatie en ICT uitgekozen omdat zij op grote schaal geo-informatie produceren en dit aan veel gebruikers, tegen betaling, beschikbaar stellen. Waterschap Zeeuwse Eilanden is uitgekozen omdat hier ook het kwantitatieve onderzoek uitgevoerd wordt. Het is dan met name interessant om te zien hoe deze organisatie haar kwaliteitsborging heeft georganiseerd en dit te relateren aan de Topografische dienst kadaster, het Kadaster en de Adviesdienst Geo-informatie en ICT.

Van iedere hierboven genoemde organisatie is een interview afgenomen van een medewerker van het bedrijf die op de hoogte is van de interne ontwikkelingen en borging van kwaliteitsaspecten van geo-informatie. De uitwerkingen van deze interviews zijn opgenomen in de bijlagen 1 tot en met 4 voor respectievelijk de Topografische dienst kadaster, het Kadaster, de Adviesdienst Geo-informatie en ICT en waterschap Zeeuwse Eilanden.
In de volgende paragrafen zal op basis van de uitgewerkte interviews de zaken met betrekking tot kwaliteit van geo-informatie beschreven worden. In paragraaf 3.6 zullen alle verschillen en overeenkomsten op een rij gezet worden terwijl in paragraaf 3.7 de conclusies getrokken zullen worden.

3.2 Kwaliteit van Geo-informatie bij Topografische Dienst Kadaster

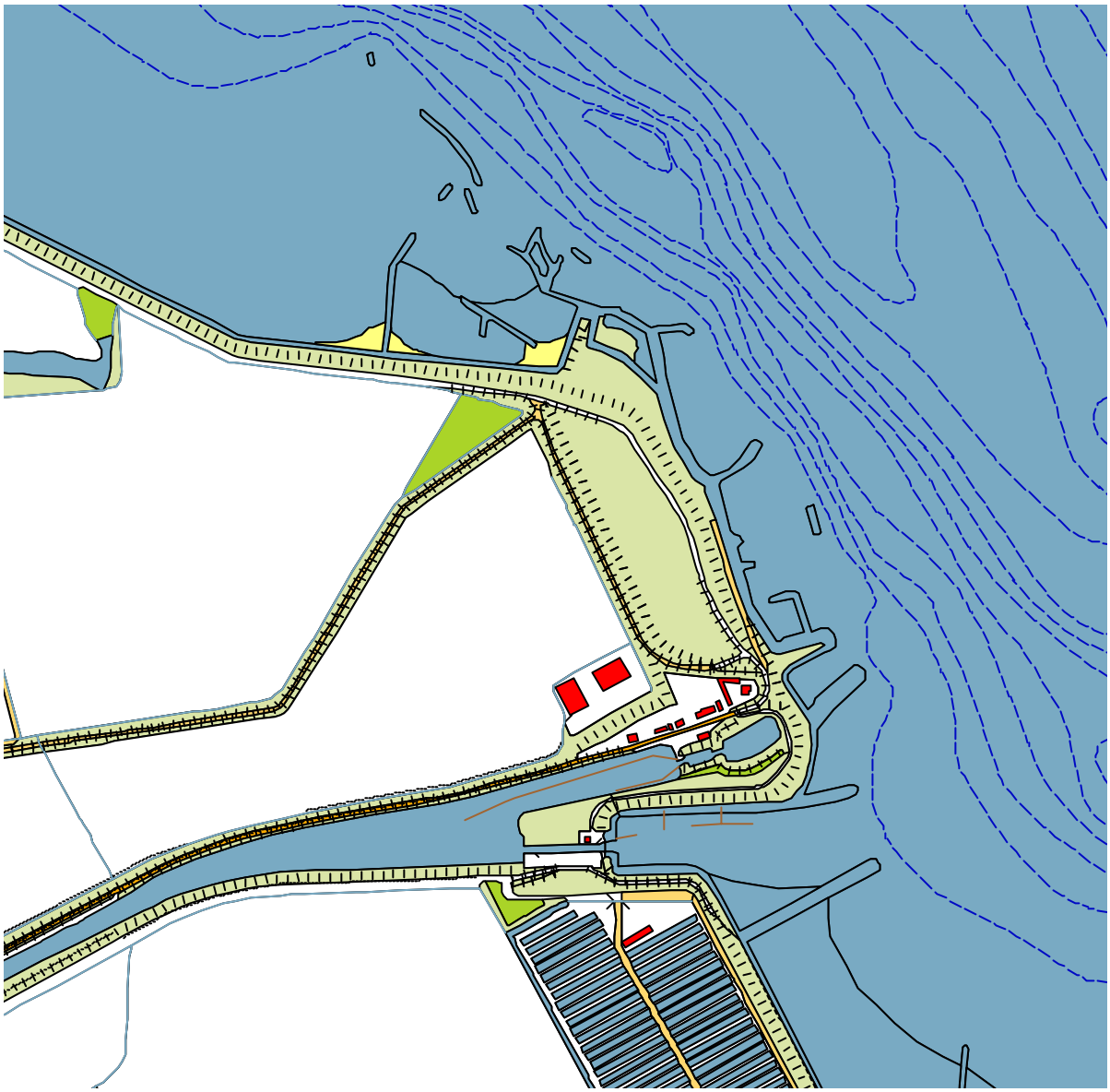
3.2.1 Inleiding

De Topografische Dienst Kadaster (TDK), voorheen de Topgrafische Dienst Nederland, verzamelt, verwerkt en verstrekt topografische informatie van het gehele Nederlandse grondgebied. Zij doet dit door diverse topografische producten aan te bieden die variëren van topografische bestanden op standaard schalen tot speciale bestanden voor militaire en

civiele toepassingen. Alle topografische producten zijn geattributeerd met soorten objecten en overige kenmerken.
De Top10Vector is de meest gebruikte en mede daardoor de “core business” van de TDK. Alle andere producten worden uiteindelijk hiervan afgeleid. Eens in de 4 jaar wordt dit bestand bijgewerkt. Het proces van vervaardiging ziet er als volgt uit:

- Het maken van luchtfoto’s en de verwerking ervan met behulp van fotogrammetrische technieken.
- Het verzamelen van topografische informatie door middel van terreinverkenning.
- Het opbouwen van een digitaal topografisch Basisbestand (TOP10vector).
- De cartografische verwerking van TOP10vector tot verschillende topografische bestanden en kaartseries op verschillende schalen.

Een fragment van de Top10Vector is hierna weergegeven:



Figuur 6 Fragment van de Top10Vector (© TDK)

3.2.2 Kwaliteit van geo-informatie

Binnen TDK is er niet één afdeling die verantwoordelijk is voor het borgen van de kwaliteit. Wel is er één afdeling die het gehele proces van de kaartproductie voor haar rekening neemt. Hierdoor valt de verantwoordelijkheid van de kwaliteit van de geo-informatie binnen deze afdeling, waardoor het enigszins gemakkelijk te reguleren is.

TDK onderscheidt de volgende aspecten van kwaliteit van geo-informatie:

- Logische consistentie

Bij TDK begint het maken van een consistent bestand met het opleiden, begeleiden en instrueren van de topografen. Regels die hieraan ten grondslag liggen zijn opgenomen in het handboek verkenningsvoorschriften. In dit handboek staan regels die bij de verkenning in acht dienen te worden genomen. Deze regels dienen echter eenduidig toegepast te worden omdat het uiteindelijke eindproduct ook tussen topografen onderling consistent dient te zijn. Te denken valt bijvoorbeeld aan de classificatie, wanneer betreft een verzameling bomen een bos is en wanneer betreft het losse bomen.

Van ieder verkend gedeelte worden er 3 afdrukken gemaakt om een bepaald object of objectsoort zoals bijvoorbeeld een categorie weg beter te onderzoeken op consistentie van codering of volledigheid van digitaliseren. Dit laatste onderdeel hoort overigens bij het aspect volledigheid, wat later behandeld wordt. Is de gehele weg wel gedigitaliseerd en heeft deze weg over haar gehele loop dezelfde codering. Op deze wijze wordt het gehele bestand consequent gecontroleerd.
Naast bovenstaande maatregelen om de consistentie zo goed mogelijk te borgen heeft TDK er ook nog voor gekozen om de ene medewerker te laten digitaliseren en dit door een andere medewerker te laten kijken. Vooral na het invoeren van deze maatregel leidde dit tot de nodige problemen. Na een aanloopperiode werkt het zogenaamde collegiaal nakijken naar behoren.
Binnen TDK worden geen normen gesteld met betrekking tot de te accepteren mate van logische consistentie.

- <u>Volledigheid</u>
  Zoals reeds vermeld onder “logische consistentie” wordt er door middel van het maken van 3 afdrukken visueel gecontroleerd op volledigheid. Daarnaast worden er ook nog andere visuele controles uitgevoerd. Er wordt onder andere consequent op basis van de luchtfoto’s gecontroleerd of alles gedigitaliseerd is.
  Het controleren van volledigheid is een erg moeilijk onderdeel omdat er ook bepaald dient te worden wanneer iets wel en wanneer iets niet meegenomen dient te worden. Als bijvoorbeeld bij de bouw van een huis alleen de fundering aanwezig is, dient dit dan meegenomen als gebouw of nog niet. Hiervoor zijn wel richtlijnen opgesteld maar er komen altijd grensgevallen voor.
  Binnen TDK worden geen normen gesteld met betrekking tot de te accepteren mate van volledigheid.

- <u>Temporele nauwkeurigheid</u>

Het hoofdproduct van TDK, de TOP10Vector, heeft een updatecyclus van 4 jaar. Iedere jaar wordt er van 25 % van Nederland luchtfoto's gemaakt die de uiteindelijke bron vormen van het eindproduct. Na het maken van de luchtfoto's gaat er door het interne productieproces nog een behoorlijke tijd overheen voordat deze 25 % van de complete kaart bijgewerkt is en uitgeleverd kan worden aan klanten.
Om deze interne doorlooptijd te verkorten worden sinds enige tijd de luchtfoto's ook gescand afgeleverd aan TDK. Hierdoor kunnen de topografen met een laptop veel sneller het veld in om nieuwe topografie in te tekenen. De doorlooptijd van het gehele proces van kaartvervaardiging wordt hierdoor sterk verkort. Vroeger duurde het herzien van de kaarten jaren. Dit is nu (in het digitale tijdperk) teruggebracht tot maanden. Op het moment dat de producten nu op de markt worden gebracht zijn ze dus ook minder verouderd en zijn dus kwalitatief beter.
Binnen TDK worden naast de updatecyclus van eens in de 4 jaar geen normen gesteld met betrekking tot de te accepteren mate van temporele nauwkeurigheid. Er wordt getracht de kaartproducten na het maken van de luchtfoto's, zo snel mogelijk ter beschikking te stellen.

- <u>Positionele nauwkeurigheid</u>
  Normen met betrekking tot de te accepteren mate van positionele nauwkeurigheid worden nog niet gesteld. Wel wordt een bepaalde kwaliteit van de luchtfoto's voorgeschreven. Dit betreft onder andere de mate van scherpte, contrast en eventuele sneeuwbedekking.
  Daarnaast wordt er visueel gecontroleerd op afwijkingen in de positionele nauwkeurigheid. In eerste instantie gebeurt dit door de topograaf zelf en later door het collegiaal nakijken.
  Naast bovenstaande wordt er ook door middel van softwarematige controles gecontroleerd op slivers en gaps. Dit betreffen naast topologische fouten ook positionele onnauwkeurigheden.
  Bovenstaande controles worden structureel uitgevoerd, dus niet steekproefsgewijs.
  Momenteel loopt er een onderzoek om de positionele nauwkeurigheid van de TOP10Vector te gaan bepalen op basis van de Grootschalige BasisKaart van Nederland (GBKN) die een hogere mate van positionele nauwkeurigheid bezit. Op basis van de uitkomsten van dit onderzoek zal getracht worden om normen voor de positionele nauwkeurigheid van de TOP10Vector te gaan bepalen.

- <u>Attribuut nauwkeurigheid</u>

Bij het hiervoor reeds beschreven visueel collegiaal nakijken wordt op een grove wijze ook gecontroleerd op juistheid van classificatie. Dit nakijken op een luchtfoto is een zeer beperkte controle van classificatie omdat het merendeel van de classificaties niet van de luchtfoto onderscheiden kan worden.
Normen met betrekking tot de te accepteren mate van attribuut nauwkeurigheid worden niet gesteld.

3.3 Kwaliteit van Geo-informatie bij Kadaster

3.3.1 Inleiding

Het Kadaster, dat haar hoofdvestiging heeft in Apeldoorn, heeft als wettelijke taak onder meer het bijhouden en beheren van de eigendomssituatie. Dit houdt in het beheren en bijhouden van de Openbare Registers en de Kadastrale Registratie met inbegrip van de kadastrale kaart.
Hiernaast beheert het Kadaster ook voor diverse regionale Stichtingen GBKN de Grootschalige BasisKaart van Nederland (GBKN). Dit is een zeer gedetailleerde digitale grootschalige topografische kaart. Deze kaart bevat zowel geometrische informatie als beschrijvende informatie over deze geometrie. Afnemers van de GBKN zijn voornamelijk lokale overheden en nutsbedrijven.
Het bijwerken van de GNKN wordt als regel periodiek uitgevoerd. Vooraf worden mutaties opgespoord. De betrokken afnemers kunnen ook bij de opsporing en meting van mutaties worden ingeschakeld. Het Kadaster zorgt ervoor dat deze mutaties of nog ingemeten worden en daarna, of direct verwerkt worden in de GBKN. Het is voor klanten mogelijk zowel regelmatig als op ad-hoc basis deze mutaties of de complete GBKN bestanden af te nemen.

De kadastrale kaart is een digitale kaart met alle kadastrale perceelsgrenzen, perceelsaanduidingen, hoofdgebouwen en diverse plaatselijke benamingen en huisnummers. De hoofdgebouwen en plaatselijke benamingen worden als regel overgenomen van de GBKN. Deze kaart wordt naast lokale overheden en nutsbedrijven ook veelvuldig door particulieren en notariskantoren afgenomen. Deze kadastrale kaart bevat ook beschrijvende informatie. Te denken valt bijvoorbeeld aan de opnamedatum van een object en de verwijzing naar brongegevens (zoals het kadastrale veldwerk). Op basis van de kadastrale aanduiding (de combinatie gemeente-sectie-perceelnummer) is het mogelijk om informatie uit de Kadastrale Registratie te raadplegen, zoals bijvoorbeeld de zakelijk gerechtigde, de kadastrale oppervlakte en de betrokken rechten.

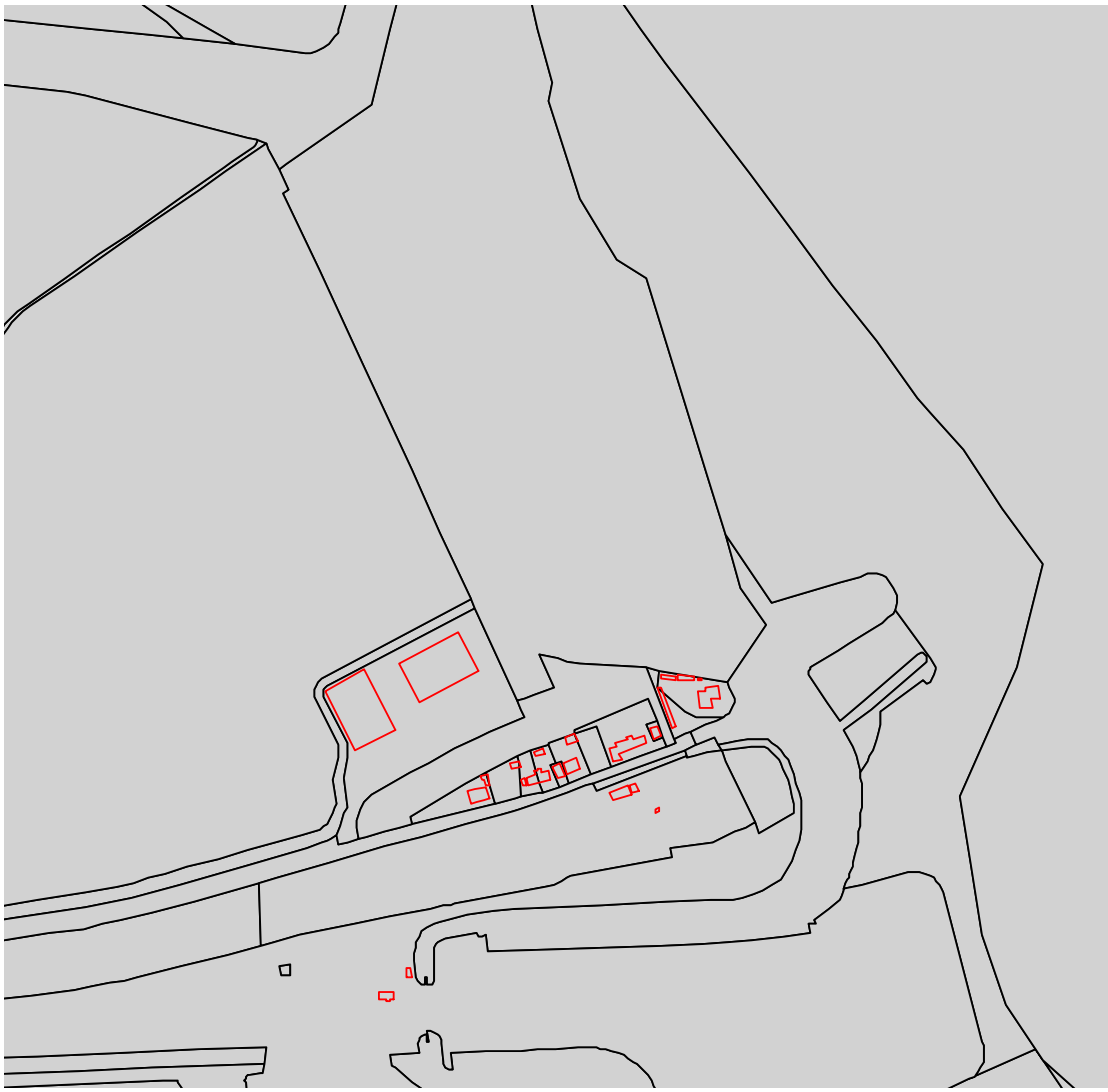
Het bijwerken van de Kadastrale kaart gebeurt op basis van de door notariskantoren aangeboden akten. Als voorbeeld kan gedacht worden aan het verkopen van een gedeelte van een kadastraal perceel. Vervolgens worden deze mutaties door het Kadaster ingepland en gezamenlijk met andere mutaties in het betreffende gebied, ingemeten en verwerkt in de kadastrale kaart.
Ook voor de kadastrale kaart geldt dat het voor klanten mogelijk is om zowel regelmatig als op ad-hoc basis bijgewerkte bestanden te verkrijgen.

Voor de kadastrale kaart streeft het Kadaster er naar om binnen 1 jaar aangemelde mutaties in de verschillende producten verwerkt te hebben. Voor de GBKN verschilt dat waarop de mutaties betrekking hebben. De volgende situaties kunnen zicht voordoen:

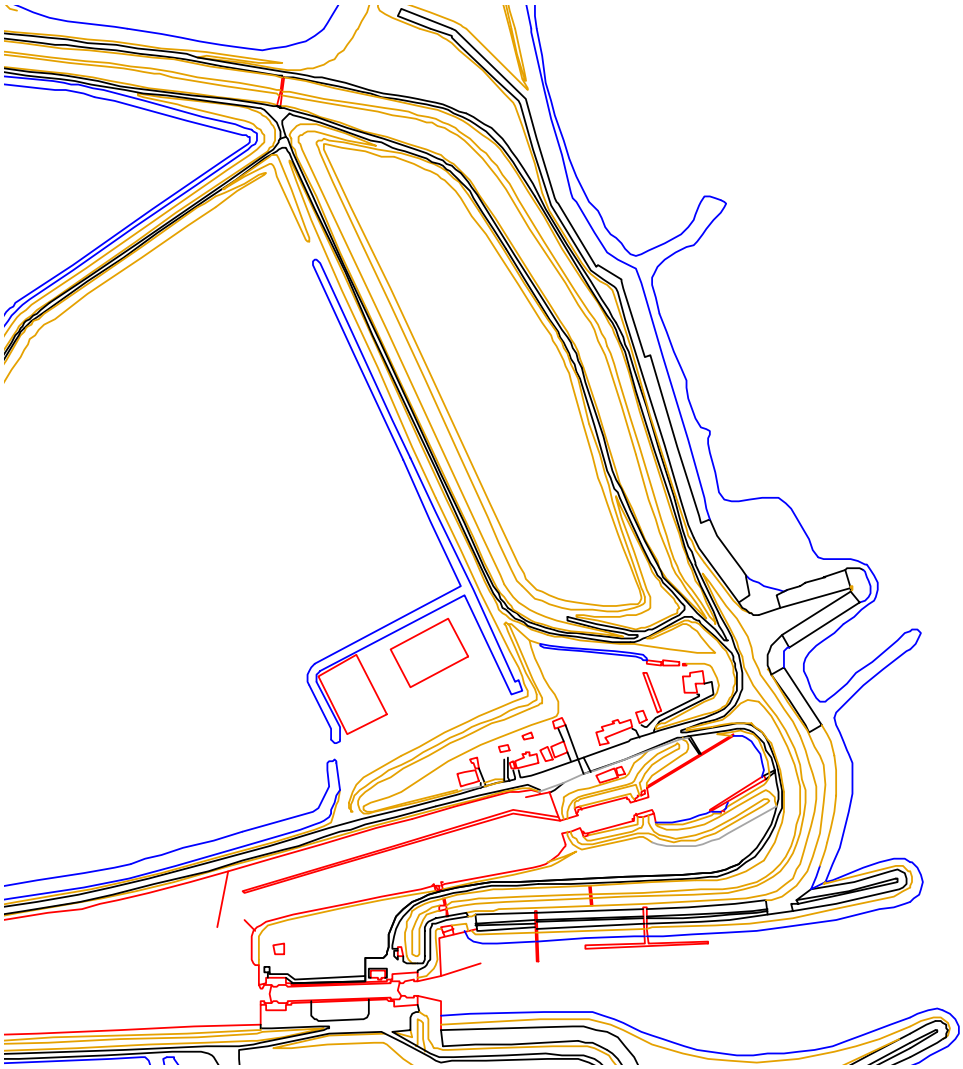
- Mutaties binnen de bebouwde kom → binnen één jaar
- Mutaties buiten de bebouwde kom → binnen twee jaar
- Mutatie van hoofdgebouwen → binnen 6 maanden

Bovenstaande betreft echter een landelijke aanbeveling. In principe wordt de frequentie van bijhouding bepaald door de regionale gebruikers van het kaartbestand.

Hierna is zowel een fragment van de kadastrale kaart als de GBKN weergegeven.



Figuur 7 Een fragment van de kadastrale kaart (© Kadaster)



Figuur 8 Een fragment van de GBKN (© Regionale Stichting GBKN Zeeland)

3.3.2 Kwaliteit van geo-informatie

3.3.2.1 Algemeen

Bij het kadaster is er niet één centrale afdeling die verantwoordelijk is voor het borgen van de kwaliteit of die hierop toeziet. Wel is het zo dat de benodigde kwaliteit van het Kadaster door middel van wetten, besluiten en centraal vastgestelde uitvoeringsregelingen vastgelegd is. Door middel van controles en steekproeven wordt regelmatig getoetst of de gewenste kwaliteit ook gewaarborgd en gerealiseerd wordt. Jaarlijks wordt door de kwaliteitscoördinatoren (1 in elke regionale vestiging en de Concernstaven) een plan opgesteld met daarin de voorgenomen acties en onderzoeken
De kwaliteitseisen worden continue toegepast in ieder inwintraject en elk bijhoudingsproces. Door het veranderen van methodes en inzichten kan het wel zo zijn dat die eisen aangepast

worden aan de op dat moment geldende inzichten. Het Handboek Technische Werkzaamheden van het Kadaster (HTW) is een van de handleidingen waarin de kwaliteitsspecificaties voor inwintrajecten staan beschreven. Verder zijn er diverse uitvoeringsregelingen en kwaliteitshandboeken waarin ook specificaties en richtlijnen staan aangegeven voor de verschillende werkprocessen.

3.3.2.2 Kadastrale kaart

Voor de kadastrale kaart worden een aantal kwaliteitsaspecten onderkend. Voor al deze kwaliteitsaspecten heeft het Kadaster normen gesteld waaraan ieder kwaliteitsaspect zou moeten voldoen. De volgende kwaliteitsaspecten worden voor de kadastrale kaart onderkend:

- <u>Positionele nauwkeurigheid</u>
  Een kadastrale grens is (een deel van) de begrenzing van een kadastraal perceel en moet geometrisch voldoen aan de eisen zoals die zijn beschreven in de HTW (hoofdstuk 2.6.1):
  - de afbeelding in de kadastrale kaart geschiedt met een grafische precisie,
  - grenzen worden gemeten met een puntprecisie die beter is dan de idealisatieprecisie van die punten
  - zodanig dat de grens in het terrein reconstrueerbaar is.

  De controles die uitgevoerd worden om te onderzoeken in hoeverre de normen gehaald worden beschreven in de HTW. Het komt er op neer er tijdens de inwinning meer gegevens opgenomen worden dan strikt noodzakelijk. Deze extra informatie wordt bij de verwerking statistisch getoetst. De nieuwe situatie wordt pas geaccepteerd als geconstateerde discrepanties binnen de gestelde normen vallen. Incidenteel vindt er bovendien een controle plaats (de zgn. kwaliteitsaudit) of het vereiste werkproces met inbegrip van de vereiste toetsing correct is uitgevoerd.

  Voor de positionele nauwkeurigheid van de overige elementen, met name de hoofdgebouwen gelden dezelfde normen als bij de GBKN.

- <u>Consistentie</u>
  Normen voor wat betreft de consistentie van de kadastrale kaart worden op diverse onderdelen gesteld. Voorbeelden zijn onder andere:

- Zijn alle percelen gesloten
- Tussen percelen bestaat geen open, niet geregistreerde ruimte.
- Percelen overlappen elkaar niet
- Alle percelen dienen een landelijke unieke aanduiding te hebben
- Een perceelnummer wordt maar eenmalig gebruikt

Al deze aspecten zijn automatisch geborgd in de werkprocessen.

- Volledigheid
  Voor de kadastrale kaart betreft dit kwaliteitsaspect zowel ontbrekende als overcomplete objecten. Voor verschillende objectsoorten binnen de kadastrale kaart zijn verschillende normen opgenomen. De normen voor een perceelsgrens (100%) verschilt bijvoorbeeld van een bijgebouw of een huisnummer (<100%). De volledigheid van kadastrale grenzen inclusief bijbehorende beschrijvende informatie is automatisch geborgd in de werkprocessen. Voor wat betreft de andere objecten in de kadastrale kaart is de volledigheid afhankelijk van de kwaliteit van de GBKN. Deze overige objecten betreffen namelijk bouwwerken die identiek zijn aan de inhoud van de GBKN. Controles op de volledigheid hiervan vindt alleen plaats als er voor de inmeting mutaties in het veld geïnventariseerd worden. In sommige regio's worden mutaties niet in het terrein verzameld maar verzameld in een mutatieregistratiecentrum. Dan is de volledigheid afhankelijk van de kwaliteit van de registratie. Incidenteel ontstaan verbeteringen in de volledigheid door het melden door gebruikers (het 'piepsysteem').

- Actualiteit
  De actualiteit van de kadastrale kaart wordt continue gevolgd door gebiedsgewijs bij te houden in hoeverre de aangemelde wijzigingen in de begrenzing van eigendommen zijn verwerkt. Hierdoor is het inzicht in de actualiteit van de grenzen altijd aanwezig.
  Voor de objecten afkomstig uit de GBKN ligt dit anders. De actualiteit wordt bepaald door de regionale gebruikers van het bestand en kan dus verschillen. De actualiteit is mede afhankelijk van de registratie van meldingen van mutaties.

3.3.2.3 Grootschalige Basiskaart van Nederland

De kwaliteitsaspecten die voor de GBKN onderscheiden worden staan vermeld in de objectencatalogus GBKN. Voor al deze kwaliteitsaspecten zijn normen opgesteld waaraan voldaan zou moeten worden.

Bij de eerste vervaardiging van de GBKN is de kaart in fasen gecontroleerd of het gerealiseerde binnen de gestelde kwaliteitsnormen viel. In de verdere bijhouding van de GBKN wordt de positionele precisie visueel gecontroleerd en zeer incidenteel met het nemen van steekproeven in het veld. Dit gebeurt wel eens als het project bijvoorbeeld een complete nieuwbouwwijk betreft. Daarnaast worden soms ook de resultaten beoordeeld van de rekenkundige toetsing van de aansluiting/inpassing van nieuwe objecten in de bestaande situatie.
Van de andere aspecten van kwaliteit wordt regelmatig steekproefsgewijs gecontroleerd of de gestelde normen ook gehaald worden. Dit betreft voornamelijk visuele controles en incidenteel controles in het veld.

De volgende kwaliteitsaspecten worden voor de GBKN door het Landelijk Samenwerkingsverband GBKN geadviseerd:

- Actualiteit
  De wijze waarop de actualiteit van de GBKN wordt gerealiseerd en de frequentie van de bijhouding wordt bepaald door de regionale gebruikers van het kaartbestand. In paragraaf 3.3.1 is op de gestelde normen reeds dieper ingegaan. De actualiteit wordt gebiedsgewijs geregistreerd door de GBKN-beheerder.

- Identificeerbaarheid
  Hiermee wordt de herkenbaarheid, generalisatie, idealisatie en leesbaarheid bedoeld. Dit betreft een algemene kwaliteitseis die op bestandsniveau bepaald wordt. Hier wordt geen kwaliteit van individuele objecten mee beschreven.

- Wijze van inwinning
  Hierbij zijn een aantal mogelijkheden, onder andere fotogrammetrische inwinning, scanningstechnieken en digitalisering van bestaand kaartmateriaal. Dit aspect van kwaliteit is niet van belang voor het gebruik van het bestand.

- Toegankelijkheid
  De toegankelijkheid van de digitale geometrischge en beschrijvende informatie is vastgelegd in de voorschriften met betrekking tot het gebruikte uitwisselingsformaat. Dit geldt voor het gehele bestand.

- Volledigheid
  Voor het definiëren van normen van kwaliteit wordt de GBKN opgesplitst in 3 delen:

- Opdelingsobjecten
- Inrichtingselementen
- Hoofdgebouwen

Voor alle drie zijn normen opgesteld waaraan de GBKN zou moeten voldoen. De normen worden opgenomen in de bestekken voor de bijhouding van de GBKN. Controles op de volledigheid vindt alleen plaats als er voor de inmeting mutaties in het veld geïnventariseerd worden. In sommige regio's worden mutaties niet in het terrein verzameld maar verzameld in een mutatieregistratiecentrum. Dan is de volledigheid afhankelijk van de kwaliteit van de registratie. Incidenteel ontstaan verbeteringen in de volledigheid door het melden door gebruikers (het 'piepsysteem').

- Consistentie
  Hiermee wordt binnen de GBKN de eenduidigheid van de digitale standaardpresentatie aangeduid. De normen worden opgenomen in de bestekken voor de bijhouding van de GBKN en vervolgens kan na oplevering van de bijhouding steekproefsgewijs worden gecontroleerd door de GBKN-beheerder.

- Precisie
  Hiermee wordt de standaardafwijking bedoeld van de nauwkeurigheid van de geometrie van objecten. De normen hiervoor zijn opgesteld en afgestemd op de oorspronkelijke kaartschaal van de analoge GBKN. De normen worden opgenomen in de bestekken voor de bijhouding van de GBKN en kunnen vervolgens na oplevering van de bijhouding steekproefsgewijs worden gecontroleerd door de GBKN-beheerder.

3.4 Kwaliteit van Geo-informatie bij de Adviesdienst Geo-informatie en ICT

3.4.1 Inleiding

De Adviesdienst Geo-informatie en ICT (AGI) is één van de specialistische diensten van Rijkswaterstaat, de uitvoeringsorganisatie van het Ministerie van Verkeer en Waterstaat. De AGI ondersteunt de kerntaken van het ministerie en rijkswaterstaat door ervoor te zorgen dat

zij kunnen beschikken over goede, gecertificeerde en gestandaardiseerde geo-informatie en ICT infrastructuur.

Bij de AGI worden diverse soorten geo-informatie ingewonnen en beheerd. Het betreffen de volgende:

- Digitale Topografische Bestanden (DTB)
- Het Actueel Hoogtebestand Nederland (AHN)
- TOPhoogteMD
- Het waterstaatkundig Informatiesysteem
- Grondslagpunten
- Normaal Amsterdams Peil (NAP)
- Informatie Systeem Waterschappen
- Dijkringenbestand
- Vegetatiekaarten

Al deze bestanden bestaan uit zowel geometrie als (globale) beschrijvende informatie van deze geometrie. Omdat voor de AGI de nadruk van de werkzaamheden ligt bij het inwinnen en leveren ten behoeve van de DTB's beperkt dit onderzoek zich alleen tot deze DTB's.

De inwinning van de DTB's wordt door de hoofdafdeling topografische geo-informatie uitbesteed of voert het in eigen beheer uit. In principe wordt het bijwerken van de DTB's iedere 5 jaar uitgevoerd. Eventueel is het mogelijk om door de regionale directies de mutatiecyclus te laten verkorten. Zij dienen dit dan echter wel zelf te bekostigen. In de natte infrastructuur zijn momenteel dermate grote financiële problemen dat de kans aanwezig is dat de mutatiecyclus van eens in de 5 jaar verlengd zal worden naar bijvoorbeeld eens in de 7 jaar. Het veranderen van de muteercyclus gebeurt op basis van budgetgrootte en is derhalve niet afgestemd op het werkproces.

3.4.2 Kwaliteit van geo-informatie

De hoofdafdeling topografische geo-informatie ziet toe op de kwaliteitsborging en besteed inwinwerk uit of voert het zelf uit. In het algemeen worden er bij interne inventarisaties weinig controles uitgevoerd. Bij uitbestedingen van de inventarisaties of verwerkingen vinden deze wel plaats. Er is geen centrale afdeling die op deze controles toeziet. Ten behoeve van de uitbesteding zijn er productvoorwaarden opgesteld. Deze voorwaarden zijn verwoord in de Productspecificatie uitbesteding fotogrammetrie. Zij gelden

niet alleen voor uitbesteding van werken maar worden ook als interne voorwaarde meegenomen.
Ten behoeve van het borgen van de kwaliteit is er in deze productspecificatie een grote hoeveelheid aan voorwaarden gesteld met betrekking tot de inwinning van de geo-informatie. Dit betreft onder andere voorwaarden betreffende de volgende onderdelen:

- Het vliegplan
- Het schijvenplan
- De paspunten
- De diapositieven
- De gescande beelden

Voor alle aldus beschreven voorwaarden worden ook toetsen beschreven waarmee gecontroleerd kan worden in hoeverre aan de gestelde voorwaarden voldaan wordt. Deze beschrijvingen zijn zeer gedetailleerd. Er wordt bijvoorbeeld ook aandacht besteed aan de grootte van steekproeven en aan een trapsgewijze nauwkeurigheidseis. Een voorbeeld hiervan is dat 60% van het verschil tussen werkelijkheid binnen de standaardafwijking moet liggen, 90 % binnen 2 maal de standaardafwijking en 95 % binnen 3 maal de standaardafwijking.

Om de interne verwerkingskwaliteit te borgen is er ook een instructie opgesteld waarin precies staat beschreven hoe de werkzaamheden uitgevoerd dienen te worden. Deze instructie heet de Inwin- en verwerkingsinstructie.

Bij de AGI worden de volgende aspecten van kwaliteit van geo-informatie onderkend:

- <u>Positionele nauwkeurigheid</u>
  Doordat er voornamelijk gewerkt wordt aan de DTB's liggen de eisen voor wat betreft de positionele nauwkeurigheid vast in de productspecificatie van de DTB's. Door middel van vereffeningen wordt hierop gecontroleerd en ook worden deze vereffeningen steekproefsgewijs optisch gecontroleerd. Dit staat beschreven in de productspecificaties uitbesteding fotogrammetrie. De normen waaraan het product dient te voldoen zijn ook in dit document vastgelegd. Hierbij wordt op diverse manieren differentiaties in de vereiste mate van nauwkeurigheid onderscheiden. Het betreft de volgende:
  - Onderscheid tussen het horizontale en verticale vlak
  - Onderscheid tussen absolute en relatieve nauwkeurigheid

- Afhankelijk van de mate van idealisatie kan een object in drie verschillende klassen geplaatst worden. Wegbelijning is goed idealiseerbaar en valt in klasse 1. Taludlijnen zijn slecht idealiserbaar en vallen daarom in klasse 3

- Consistentie

  Bij de AGI betreft dit consistentie met betrekking tot topologie. De eisen voor de opbouw van de topologie liggen vast in het datamodel DTB. De topologie wordt bij de AGI zelf altijd opgebouwd in het Infocam-systeem van Leica, wat de basis checkt v.w.b. 'spaghetti' (check tijdens procesgang). Tevens zijn er controles op het vlakkennet op maaiveldniveau: geen overlappingen of ontbreken van vlakken. Dit kwaliteitsaspect wordt structureel gecontroleerd.

  Daarnaast wordt er voor de bestandsopbouw ook gecontroleerd op een logische opbouw conform de Inwin- en verwerkingsinstructie. Hierbij wordt onder andere gelet op de volgende situaties:
  - Punten zonder z-hoogte
  - Niet gesloten vlakken
  - Dubbele coördinaten
  - Dubbele symbolen.

  Het verwerkingsproces is zo ingericht dat alle data gecontroleerd wordt op bestandstechnische aspecten voor het de database in gaat.

- Attribuut nauwkeurigheid

  Voor wat betreft de attribuutnauwkeurigheid dient 100 % van de objecten goed geclassificeerd te zijn. Onregelmatig vinden er steekproefsgewijs optische controles plaats. Parameters van de kwaliteit van de werkelijke attribuutnauwkeurigheid zijn niet bekend. Er is dus bijvoorbeeld niet bekend hoeveel procent van een bepaalde categorie foutief geclassificeerd is.

- Volledigheid

  Normen voor de vereiste volledigheid zijn aanwezig. Hierbij wordt onderscheid gemaakt tussen objecten die goed of slecht idealiseerbaar zijn. De volledigheid van objecten, attributen of relaties hiertussen binnen een bestand wordt visueel getoetst met de foto als ondergrond. Dit levert alleen een controle op basis van volledigheid van objecten, niet van attributen of relaties tussen deze. Parameters van de werkelijke kwaliteit van volledigheid zijn niet bekend. Er is dus niet bekend hoeveel procent van een bepaald object in een bepaald gebied naar verwachting ontbreekt.

3.5 Kwaliteit van Geo-informatie bij Waterschap Zeeuwse Eilanden

3.5.1 Inleiding

Waterschap Zeeuwse Eilanden zorgt dagelijks voor de zeeweringen, het oppervlaktewater en de plattelandswegen op Sint-Philipsland, Tholen, Schouwen-Duiveland, Walcheren, Noord-Beveland en Zuid-Beveland. Dit werkgebied is verdeeld in zes werkgebieden. Het hoofdkantoor is te vinden in Goes.
Ten behoeve van deze werkprocessen is een grote hoeveelheid geo-informatie benodigd. Deze kan opgedeeld worden in de volgende twee hoofdsoorten:

Basis- of referentiebestanden.
Dit zijn bestanden die van derden worden betrokken en die binnen het waterschap voor diverse toepassingen gebruikt worden. Deze bestanden worden dus alleen beheerd door het waterschap en niet ingewonnen. De volgende kunnen onderscheiden worden:

- Grootschalige Basiskaart van Nederland (geleverd door Stichting GBKN)
- Kadastrale kaart (geleverd door het Kadaster)
- Top10Vector, Top50Vector, Top25 Raster (geleverd door de Topgrafische Dienst Kadaster)
- Bodemkaart (geleverd door Alterra)
- Nationaal Wegenbestand (geleverd door Rijkswaterstaat)
- Actueel Hoogtebestand van Nederland (geleverd door de Adviesdienst Geo-informatie en ICT)

Bestanden die de infrastructuur van het waterschap beschrijven
Deze bestanden worden zowel ingewonnen als beheerd en bijgehouden door het waterschap zelf en kunnen in de volgende subcategorieën opgedeeld worden:

- Bestanden die betrekking hebben op de waterkeringen. Het gaat te ver om alle objecten die hierbinnen onderscheiden kunnen worden te beschrijven. Wel kan de volgende opdeling gehanteerd worden:
    - Objecten die betrekking hebben op de constructieopbouw
    - Objecten die betrekking hebben op de vooroever
    - Objecten die betrekking hebben op referentiestelsels
    - Objecten die betrekking hebben op bekledingen

Bijzondere kenmerken van bovenstaande bestanden is dat een aantal driedimensionaal beschikbaar zijn. Dit betreffen de categorieën vooroever en bekledingen. Van alle categorieën is zowel beschrijvende als geometrische informatie beschikbaar. De geometrische weergave is zowel in vlakken, lijnen als punten.

- Bestanden die betrekking hebben op het beheer van de wegen. Dit betreft een kleinere groep geo-informatie. Het betreft hier alleen de hartlijnen van en het groen langs de wegen. Deze bestaan beide uit zowel geometrische informatie als beschrijvende informatie en is in twee dimensies weergegeven. De geometrische weergave van de weg is in lijnen. Groen is geometrisch weergegeven als lijnen en punten.

- Bestanden die betrekking hebben op het beheer van het oppervlaktewater. Ook hier gaat het te ver om alle objecten die hierbinnen onderscheiden kunnen worden, te beschrijven. Wel kan de volgende groepsindeling gemaakt worden:
    - Objecten die betrekking hebben op het oppervlaktewatersysteem. Deze kunnen vervolgens weer onderverdeeld worden in meetpunten, het waterlopenstelsel, peil- en afvoergebieden en kunstwerken.
    - Objecten die betrekking hebben op de watersysteemketen. Dit betreffen de riool(pers)leidingen, de zuiveringen en de kunstwerken.

Bijzondere kenmerken van bovenstaande categorieën is dat deze allemaal in twee dimensies gerepresenteerd zijn en dat het zowel geometrische informatie als beschrijvende informatie hierover betreft. De geometrische weergave is zowel vlakken, lijnen als punten.

Momenteel wordt er een plan opgesteld om voor al deze geo-informatie de beheer en bijhouding te organiseren.

3.5.2 Kwaliteit van geo-informatie

Er is niet één centrale afdeling die verantwoordelijk is voor het borgen van kwaliteit. Iedere opdrachtgever is zelf verantwoordelijk voor het borgen van de kwaliteit. Vaak wordt deze rol binnen de diverse aandachtvelden ingevuld door applicatiebeheerders. Geo-informatie en ook niet–geo-informatie die binnen de applicatie horen, zijn de verantwoording van de applicatiebeheerders.

Verder is het zo dat als er bij een inwintraject aan de cluster Geo-informatie een bepaalde kwaliteit geeist wordt, het de verantwoording van deze cluster is om hieraan te voldoen. Zij dient zelf hiervoor eventueel maatregelen te treffen. Wel wordt vaak maar een beperkt aantal kwaliteitseisen gevraagd. Een goede onderbouwing van deze gevraagde kwaliteitseisen ontbreekt vaak.
De volgende aspecten worden onderkend:

- Positionele nauwkeurigheid
  Van de onderkende kwaliteitsaspecten wordt bij nieuwe inwintrajecten alleen de positionele nauwkeurigheid gedefinieerd. De aldus gesteld norm wordt getoetst door regelmatig tijdens een grootschalige inwinning, infrastructuur dubbel te meten en te onderzoeken of dit overeenkomt. Daarnaast worden er vereffeningen uitgevoerd. Deze moeten de toets van gestelde eisen en voorwaarden doorstaan.
  Naast bovenstaande controlemetingen worden er ook log-bestanden van vereffeningen van landmetingen onderzocht. Deze geven ook een indicatie van de positionele nauwkeurigheid van de meting. Niet alle fouten worden hiermee gevonden.
  Het definiëren van de vereiste positionele nauwkeurigheid wordt bij grootschalige inwinprojecten beschreven in het bestek. De controletoetsing wordt niet apart beschreven. De opdrachtnemer moet door middel van berekeningen aantonen dat aan de gestelde kwaliteitseisen met betrekking tot positionele nauwkeurigheid is voldaan.

- Attribuut nauwkeurigheid
  Tijdens het toetsen in het veld van de positionele nauwkeurigheid worden ook attributen meegenomen zoals bijvoorbeeld soort kunstwerk, diameter van duikers en soort steenbekleding van een waterkering.
  De normen waaraan dit aspect van kwaliteit zou moeten voldoen worden niet expliciet gedefinieerd. Impliciet wordt er door opdrachtgevers vanuit gegaan dat de attribuut nauwkeurigheid 100 % is.
  Bij grootschalige inwinprojecten wordt er daarnaast ook gecontroleerd op het voorkomen van onlogische waarden. Te denken valt bijvoorbeeld aan duikerdiameters boven een bepaalde waarde. Daarnaast vinden er ook visuele controles plaats waarbij gelet wordt op codering en bijvoorbeeld het voorkomen van afwijkende waarden.

- Volledigheid
  Bij grote inwinprojecten wordt er ook gecontroleerd op de volledigheid. Dit gebeurt door het uitvoeren van een voor- en/of naverkenning die volledig of steekproefsgewijs uitgevoerd kan worden. De normen waaraan dit kwaliteitsaspect zou moeten voldoen

worden niet expliciet gedefinieerd. Impliciet wordt er door opdrachtgevers vanuit gegaan dat deze foutloos zijn.

- Actualiteit
  Omdat het inwinprojecten zijn en de inwinning gedurende het project loopt, bestaat er inzicht in de actualiteit van gegevens. Voor het actueel houden van ingewonnen gegevens wordt momenteel een uitwerking gemaakt die onder andere gericht is op de toepassing waarvoor deze gegevens gebruikt worden. Om dit goed te kunnen te kunnen doen dienen ook de vereiste kwaliteitsaspecten in beeld te worden gebracht, geredeneerd vanuit het bedrijfsproces.
  Normen met betrekking tot de gestelde actualiteit zijn nog niet expliciet gedefinieerd. Na het opzetten van de beheer en bijhouding van de geo-informatie zal dit wellicht wel het geval kunnen zijn. Ook controles voor wat betreft het toetsen aan deze normen dient dan opgezet te worden.

3.6 Overeenkomsten/verschillen tussen de organisaties

Op basis van het kwalitatieve onderzoek is in onderstaande tabel aangegeven welke aspecten van kwaliteit door betreffende organisatie onderkend wordt. Tevens is aangegeven of voor deze aspecten normen worden gesteld en of deze ook getoetst worden.

	Organisatie				
Kwaliteitsparameter	TDK	Kadaster Kad kaart	GBKN	AGI	WZE
Positionele nauwkeurigheid					
Attribuut nauwkeurigheid					
Logische consistentie					
Temporele nauwkeurigheid					
Volledigheid					

geen invulling
wel invulling, geen normen gesteld
normen gesteld
geen normen gesteld, wel gecontroleerd
normen getoetst

Tabel 3 De vergelijking van parameters van de diverse organisaties

Het bepalen of een parameter nu wel of niet getoetst wordt, is niet altijd even duidelijk. Er bestaat een grijs gebied. Als voorbeeld kan het kwaliteitsaspect “volledigheid” van de kadastrale kaart van het Kadaster dienen. Van een objectsoort in de kadastrale kaart wordt de volledigheid geborgd in het productieproces. De volledigheid is hierdoor automatisch 100%. Van een andere objectsoort binnen de kadastrale kaart wordt de volledigheid alleen bij het inwinnen van mutaties gecontroleerd. Op basis hiervan is bepaald dat de normen niet getoetst worden, voor een gedeelte is dat dus niet het geval.

In tabel 3 zijn er een aantal zaken die opvallen. Het betreffen de volgende:

- Alle organisaties onderkennen ten minste vier van de vijf kwaliteitsparameters. Alleen de TDK onderkent alle parameters maar toetst daarentegen als enige organisatie geen enkele parameter.
- Er bestaat tussen de organisaties een grote verscheidenheid aan de wijze van invulling van de verschillende parameters.
- Normen met betrekking tot de positionele nauwkeurigheid worden in het algemeen gesteld en worden ook getoetst.
- In 13 van de 25 gevallen worden er normen gesteld. In 8 gevallen worden deze normen ook getoetst.
- Bij de AGI worden van vier van de vijf kwaliteitsparameters normen gesteld en ook daadwerkelijk getoetst.
- In vier gevallen wordt aan een parameter van kwaliteit geen invulling gegeven.
- In vijf gevallen wordt er wel gecontroleerd maar zijn er geen normen gesteld. Wanneer de uitkomsten van de controles acceptabel of niet acceptabel zijn is dus niet bekend.

3.7 Conclusie

In voorgaand hoofdstuk is het kwalitatief onderzoek beschreven naar hoe een viertal organisaties in Nederland een invulling geven aan het aspect kwaliteit van geo-informatie. Deze vier organisaties zijn zeer druk met het inwinnen van geo-informatie.

Al deze organisaties onderkennen het belang van het in beeld hebben van de kwaliteit van geo-informatie. Dit wordt mede onderstreept doordat alle organisaties ten minste vier van de vijf kwaliteitsparameters onderkennen. Uit het onderzoek blijkt tevens dat voor de helft van alle kwaliteitsparameters normen gesteld worden en dat in iets meer dan de helft van deze

gevallen ook wordt getoetst of deze normen ook gehaald worden. Dit betekent echter wel dat in meer dan 2/3 van de gevallen, inzicht in de werkelijke kwaliteitsparameters niet voor handen is.

Bij drie van de vier organisaties is er niet één centrale afdeling die toeziet op een goede borging van de kwaliteit van geo-informatie. Opvallend is dat de enige organisatie die wel een centrale afdeling bezit voor het toezien op de borging van de kwaliteitsaspecten, in vier van de vijf gevallen zowel de normen van kwaliteitsparameters stelt en deze ook toets. Deze organisatie betreft de AGI.

Wat verder geconcludeerd kan worden is dat er door de meeste organisaties veel aandacht besteed wordt aan het opstellen van normen en het toetsen van de positionele nauwkeurigheid. In totaal komt dit in 4 van de 5 gevallen voor.

4 Kwantitatief onderzoek naar attribuut nauwkeurigheid

4.1 Inleiding

Zoals reeds in de doelstelling is verwoord, wordt binnen veel organisaties geo-informatie gebruikt ten behoeve van de ondersteuning van belangrijke werkprocessen. Voor een juiste ondersteuning van deze werkprocessen wordt een bepaalde kwaliteit van de geo-informatie gevraagd.

In dit hoofdstuk wordt voor een dataset met betrekking tot de steenbekledingen van waterkeringen van Waterschap Zeeuwse Eilanden onderzocht voor welke werkprocessen deze dataset gebruikt wordt. Op basis hiervan wordt de vereiste kwaliteit, benodigd voor een goede ondersteuning van de werkprocessen, bepaald. Van deze dataset van de steenbekledingen wordt zowel schematisch als in geschrift de totstandkoming beschreven en wordt voor het kwaliteitsaspect “attribuutnauwkeurigheid” een toets gedaan in hoeverre deze aansluit bij de gewenste situatie. De volgende vragen worden er aan de dataset gesteld:

- Wat is de foutenfrequentie
- Wat is de aard van de fouten
- Wat is de impact van deze fouten
- Wat is de oorzaak van deze fouten

Naast bovenstaande is dit onderzoek ook interessant omdat er een methodiek van werken wordt uitgedacht hoe de attribuutnauwkeurigheid van deze dataset getoetst kan worden. Daarnaast worden er diverse categorieën met fouten bepaald en wordt de gebruikte wijze van inwinning vergeleken met de inwinning met een mobiel GIS.

4.2 Beschrijving bestand steenbekledingen

4.2.1 Inleiding

Waterschap Zeeuwse Eilanden heeft circa 380 kilometer primaire waterkering. Deze kunnen er zoals op de volgende foto’s uitzien.

Foto 2 Een haventje met op de achtergrond de primaire waterkering

Foto 3 De primaire waterkering

In 1998 is gestart met een campagne om informatie betreffende de primaire waterkeringen van waterschap Zeeuwse Eilanden in te winnen. Deze campagne is geïnitieerd vanuit een drieledig doel:

- Voor het ondersteunen van de bedrijfsprocessen met betrekking tot waterkeringen is er meer actuele en accurate informatie benodigd over deze waterkeringen. De werkprocessen waarvoor de steenbekledingen gebruikt worden zijn de volgende:
  - Het verlenen van vergunningen of ontheffingen
  - Onderhouden en beheren van infrastructuur zoals glooiingen, duinen, paalhoofden, grondlichamen en kunstwerken
  - Uitvoeren van peilingen
  - Calamiteitenbestrijding
  - Aanleg van waterkeringen
  - De beschikking hebben over en het bijhouden van een beheerregister

- Om de kans op overstromingen in beeld te brengen, dient iedere vijf jaar een toets op de sterkte van de waterkering uitgevoerd te worden. Ten behoeve van het uitvoeren van deze toets was meer informatie over de waterkering benodigd.

- In het kader van verbeteringswerken aan de waterkeringen was er behoefte aan veel informatie over waterkeringen. De verbeteringswerken hadden onder andere betrekking op de vervanging van een soort steenbekleding. Bij het in beeld hebben van de soorten bekledingen zijn veel vragen reeds in een vroeg stadium te beantwoorden. Voorbeelden van deze vragen zijn:
  - hoeveel ligt er van een bepaalde soort steenbekleding
  - waar ligt een bepaalde soort steenbekleding

Ten behoeve van het uitvoeren van bovenstaande werkzaamheden dient de in te winnen geo-informatie aan diverse kwaliteitseisen te voldoen. De eisen die gesteld zijn, betreffen de volgende:

- Een standaardafwijking van 2 centimeter voor de verticale positionele nauwkeurigheid;
- Een standaardafwijking van 5 cm voor de horizontale positionele nauwkeurigheid
- Een attribuutnauwkeurigheid van 100%

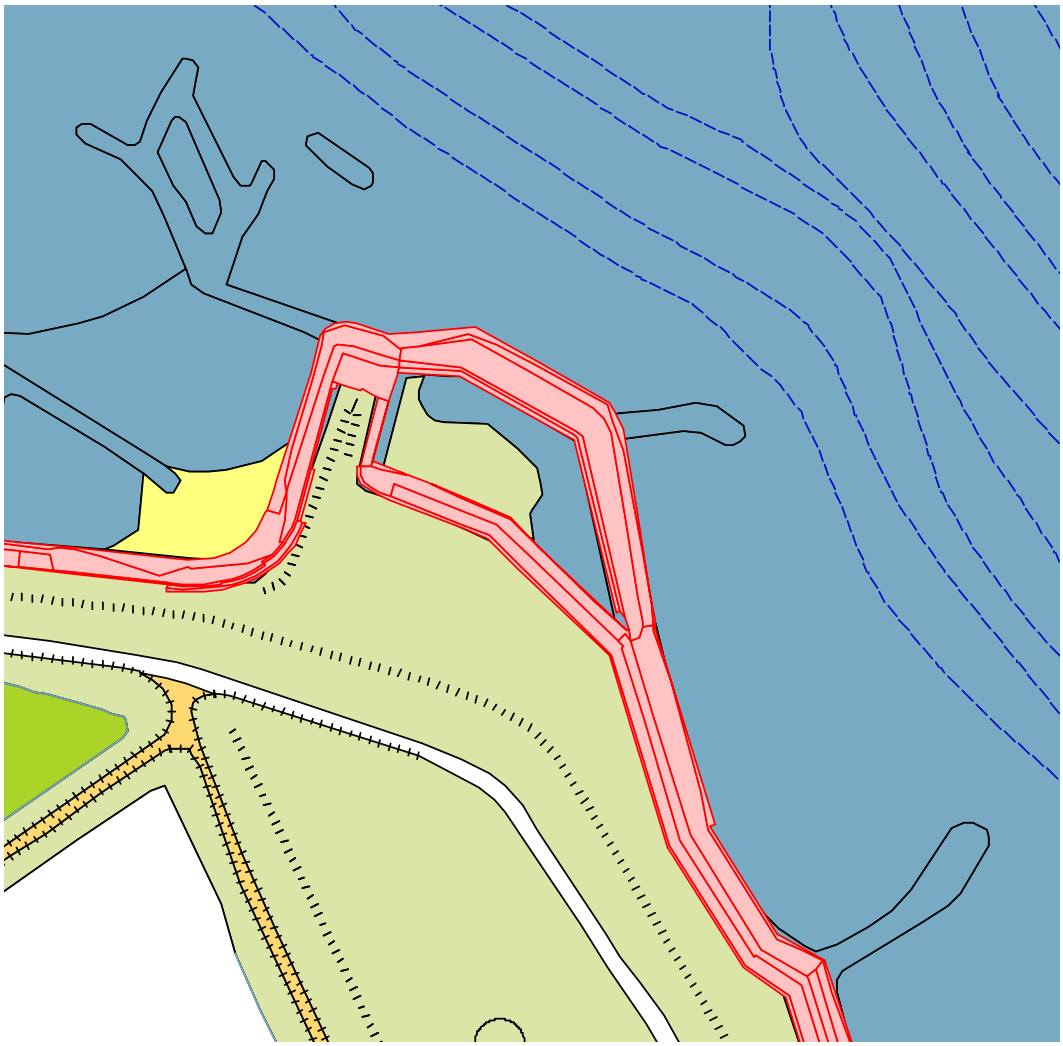
4.2.2 Procesbeschrijving inwincampagne

De inwincampagne bestond uit diverse acties. In bijlage 5 is het processchema inclusief beschrijving opgenomen van alle 23 uitgevoerd acties.

Al deze acties hebben uiteindelijk geresulteerd in een 3-dimensionaal bestand waarin zeer gedetailleerd de complete waterkering is opgenomen. Naast informatie over de steenbekledingen zijn ook andere objecten ingewonnen. Voorbeelden hiervan zijn onder andere:

- Het referentiestelsel, wat soms fysiek door middel van dijkpalen op de waterkering is uitgezet.
- Grasbekleding
- Kunstwerken op de waterkering

Hieronder is een uitsnede van het bestand met de steenbekledingen inclusief topografische ondergrond bijgevoegd.



Figuur 9 Een uitsnede van het bestand steenbekledingen met Top10Vector ondergrond

4.2.3 Het product ”steenbekleding”

Na het gereedkomen van diverse gedeelten van het bestand in 1999 is het direct ingezet voor het uitvoeren van de 5-jaarlijkse veiligheidstoetsing. Dit is begonnen met de Westerschelde en later is ook de Noordzeekust getoetst. De voorbereidingen van de toetsing van de Oosterschelde zijn momenteel in gang gezet. Deze reeds uitgevoerde toetsing is gedaan binnen een speciaal hiervoor ontwikkelde Arc-View applicatie.

Na het gereedkomen van het gehele bestand in 2000 is begonnen met het converteren van het bestand naar INTWIS. INTWIS is het INTegraal Waterschaps Informatiesysteem voor waterorganisaties gebouwd voor en door deze organisaties. Het is opgebouwd uit de standaardsoftware Arc-View, Oracle, SDE en Oracle Forms. In INTWIS kunnen de diverse organisaties hun geo-informatie opslaan op een dusdanige wijze dat deze informatie aangewend kan worden voor het ondersteunen van diverse bedrijfsprocessen en het organisatiebreed ontsluiten van geo-informatie. Tegelijkertijd met de inwincampagne heeft ook de ontwikkeling van de INTWIS-module “Waterkeringen” plaatsgevonden die gericht is op bedrijfsprocessen betrekking hebbend op de waterkeringen.

Inmiddels hebben op het bestand met de steenbekledingen diverse mutaties plaatsgevonden. De reden hiervan is dat er verbeteringswerken uitgevoerd zijn. Hierdoor zijn sommige gedeelten van het bestand pas in 2003 of zelfs nog niet geconverteerd.

Om het bestand te kunnen toetsen is een conversie vanuit INTWIS naar Arc-View uitgevoerd. Dit bestand heeft de naam steen.shp gekregen en bezit de volgende attributen:

Kolomnaam	**Type**	**Omschrijving**
Shape		Soort entiteiten in het bestand, in dit geval PolygonZ, een polygonenbestand met hoogte-informatie.
Szk_id	Numeriek	De identificatiesleutel uit de database. Via deze sleutel worden relaties met andere tabellen gelegd. Szk staat voor SteenZettingenKeringen.
Szkident	String	De zichtbare codering van het betreffende object.
Pvomsch	String	Het omschrijvingen veld. Dit is ook wel gebruikt voor maken van opmerkingen.
Bkecode	Numeriek	De code met de soort bekleding.
Bkenaam	String	De naam van de soort bekleding.

Tabel 4 De attributen van het bestand steenbekledingen

In dit bestand zijn 5639 records aanwezig die ieder verwijzen naar een steenbekledingsvlak die in grootte kunnen variëren tussen de 0,4 en 58779,6 m².

Ook zijn er grote verschillen betreffende de aantallen voorkomende soorten en de oppervlaktes van de verschillende soorten. Hierna zijn in een tabel de aantallen opgenomen per voorkomende soort. Naast deze aantallen is ook een sommatie gemaakt van het totale oppervlakte van een voorkomende soort steenbekleding.

Toplaagtype	Omschrijving	Aantal	Oppervlakte in m2
100	Asfaltbeton	1	1205
700	Breuksteen, gepenetreerd met asfalt (vol en zat)	57	63628
800	Baksteen/betonsteen, gepenetreerd met asfalt (vol en zat)	1	112
1000	Betonblokken met afgeschuinde hoeken of gaten erin	1	209
1100	Betonblokken zonder openingen	556	296421
1101	Betonblokken zonder openingen, gepentreerd met asfalt	2	1272
1110	Haringmanblokken	515	659143
1120	Diaboolblokken	33	22711
1130	gebakken steen	4	752
1132	Gebakken steen gepenetreerd met beton	4	1795
1140	betonblokken system Pitt	4	1907
1142	Betonblokken system Pitt ingegoten met beton	1	47
1170	Spijkerglooiing	1	180
1200	Open blokkenmatten, afgestrooid met granulair materiaal	4	281
1300	Blokkenmatten zonder openingen	4	2149
1700	Doorgroeisteen, beton	164	72488
1800	Breuksteen, gepenetreerd met cementbeton of colloidaal beton, (vol en zat)	13	1220
2200	Bestorting van grof grind en andere granulaire materialen	4	2781
2500	Breuksteen, (stortsteen)	19	12333
2600	Basalt, gezet	787	535277
2601	Basalt, gezet, ingegoten met gietasfalt	154	111039
2602	Basalt, gezet, ingegoten met colloidaal beton of cementbeton	204	82266
2700	Betonzuilen en andere niet rechthoekige blokken	3	549
2701	Betonzuilen of niet rechthoekige blokken, ingegoten met gietasfalt	1	392
2710	Basalton	81	99713
2711	Basalton, ingegoten met gietasfalt	15	5136
2720	PIT Polygoon zuilen	4	360
2721	PIT Polygoon zuilen, ingegoten met gietasfalt	2	304
2730	Hydroblock	19	70092
2740	Basalton met ecolaag	27	33338
2750	Hydroblock met ecolaag	8	76725
2800	Natuursteen, gezet	4	124
2801	Natuursteen, gezet, en ingegoten met gietasfalt	2	761
2802	Natuursteen, gezet, en ingegoten met beton	2	305
2810	Vilvoordse	522	185519
2811	Vilvoordse, ingegoten met gietasfalt	83	25440
2812	Vilvoordse, ingegoten met beton	360	179665
2813	Vilvoordse, overlaagd met asfalt gepenetreerde stortsteen	51	81335
2820	Lessinische	129	34848
2821	Lessinische, ingegoten met gietasfalt	12	12509
2822	Lessinische, ingegoten met beton	101	31010
2830	Doornikse	159	51371
2831	Doornikse, ingegoten met gietasfalt	10	3027
2832	Doornikse, ingegoten met beton	12	6164
2840	Petit graniet	208	60020
2841	Petit graniet, ingegoten met gietasfalt	17	4161
2842	Petit graniet, ingegoten met beton	13	1942
2843	Petit graniet	1	2353
2850	Graniet	80	75550
2851	Graniet, ingegoten met gietasfalt	11	3932
2852	Graniet, ingegoten met beton	20	4012
2861	Grauwacke ingegoten met gietasfalt	1	74
2870	Doorniks met gekantelde patronen	24	29604
2871	Doorniks met gekantelde patronen, ingegoten met gietasfalt	3	2478
2872	Doorniks met gekantelde patronen, ingegoten met beton	1	430
2900	Koperslakblokken	34	14028
3200	Klinkers, beton of gebakken.	110	64941
3210	tegels	45	6061
3220	dakpannen	8	3112
3230	Maassteen	2	189
3900	Zetwerk, ratjetoe	42	9540
5900	diverse constructies	3	1490
9900	onbekend	876	1178992
	Totaal	5639	4230812

Tabel 5 De verschillende soorten steenbekledingen met voorkomende aantallen en oppervlakte

4.3 Modellering

4.3.1 Inleiding

Zoals reeds in paragraaf 2.2.3 is beschreven, is het nodig om een model van de werkelijkheid te maken als de werkelijkheid in een database gerepresenteerd wordt. Om de kwaliteit van een dataset te toetsen is het belangrijk deze modellering in ogenschouw te nemen. De werkelijkheid kan niet zomaar vergeleken worden met de databasewaarden.

4.3.2 Onderdelen modellering

Ten tijde van het modelleren van de werkelijkheid in digitale vorm zijn er een aantal keuzes gemaakt om dit mogelijk te maken. Het betreft de volgende:

- De waterkering is opgebouwd uit diverse over elkaar heen liggende vlakken. Een voorbeeld is eerst het zandlichaam, dan een kleilaag en vervolgens de steenbekleding. Voor het modelleren van de waterkering is er gekozen voor een visueel model vanuit het topografische vlak in plaats van bijvoorbeeld op basis van een onderlaag. De toplaag is bepalend geweest voor het definiëren van individuele objecten.

- Vlakken kleiner dan 5 m² zijn tijdens de inventarisatie niet meegenomen. Voor het doel van de inwinning zijn deze niet van belang. In principe kunnen deze vlakken dus ook niet terugkomen in de database.

4.4 Keuze te onderzoeken attributen en methoden

4.4.1 Te onderzoeken attributen

Het vaststellen van de attribuutnauwkeurigheid van een dataset kan gebeuren door de waarden in de database te vergelijken met waarden met een hogere nauwkeurigheid. Dit kan zowel een kaart zijn of ten opzichte van een meer nauwkeurige inwinning in het veld. In dit geval is er echter geen kaart beschikbaar waarop de attributen met een hogere nauwkeurigheid staan afgebeeld. Hierdoor is de enige methode om de kwaliteit vast te stellen door in het veld een nauwkeurige inventarisatie uit te voeren en deze te vergelijken

met de waarden in de database. Vanzelfsprekend moet de nauwkeurigheid van de inventarisatie hoger zijn dan de nauwkeurigheid van de dataset.

Om met inachtneming van het modelleringproces later een vergelijking tussen de werkelijke kwaliteit en de aanwezige kwaliteit te kunnen uitvoeren, dienen de volgende attribuutwaarden in het veld ingewonnen te worden:

- *Bkenaam*
  Dit betreft de soort steenbekleding van de waterkering.

- *Oppervlakte < 5 m²*
  Dit is van belang voor de modelleringslag die is uitgevoerd. Vlakken kleiner dan 5 m² zijn namelijk niet meegenomen in de inventarisatie.

- *Leeftijd*
  Het is mogelijk dat na de inwinning van de vlakken door bijvoorbeeld stormschade stukken van een nieuwe toplaag zijn voorzien. Om al bij voorbaat rekening te houden met bijvoorbeeld een kwaliteitsafwijking vanwege dit fenomeen is het praktisch reeds tijdens de inwinning mee te nemen als een vlak jonger dan 5 jaar is. Deze 5 jaar is gekozen omdat de inventarisatie 5 jaar geleden is uitgevoerd. Veranderingen in niet verbeterde trajecten zijn nog niet in het databasebestand doorgevoerd.
  Bij een verschil tussen een geïnventariseerde waarde en waarde in de database kan onderzocht worden of de oorzaak van deze afwijking zit in het feit van verandering door de tijd.

- *Opmerking*
  Mocht er een opmerking geplaatst moeten worden, dan is ook deze mogelijkheid aanwezig.

4.4.2 Methode van analyse

Voor het analyseren van de attribuutnauwkeurigheid van een kwalitatief attribuut van een dataset is het noodzakelijk om waarden in de database te vergelijken met waarden van een hogere nauwkeurigheid, met inachtneming van de modellering. Deze waarden met een hogere nauwkeurigheid kunnen zowel in het veld zijn ingewonnen of afkomstig zijn uit een

dataset met een hogere nauwkeurigheid dan de te toetsen dataset. Voor het gemak worden deze waarden met een hogere nauwkeurigheid aangeduid als waarden uit het veld.
De vergelijking tussen de werkelijke waarden en de waarden in de database kan door middel van een tweetal methodes verwerkt worden tot een getal dat een indicatie geeft voor de attribuutnauwkeurigheid. Het betreft de confusionmatrix en Cohen's kappa.

De confusionmatrix

In een confusionmatrix worden horizontaal de classificaties zoals die in het veld zijn waargenomen gepresenteerd. Verticaal komen de classificaties conform de database (zie tabel 6) .

	classificatie in het veld				totaal
		klasse 1	klasse 2	klasse 3	
classificatie in	klasse 1	A	B	C	J
de database	klasse 2	D	E	F	K
	klasse 3	G	H	I	L
Totaal		M	N	O	P

Tabel 6 Het concept van een confusionmatrix

In de matrix worden de aantallen dat een bepaalde klasse in het veld als een bepaalde klasse in de database is gecodeerd, vermeld in de confusionmatrix. Bijvoorbeeld aantal b is het aantal malen dat een object in het veld met klasse 2, in de database is gecodeerd als klasse 1. De cellen die op de diagonaal liggen (A, E en I) zijn de goed gecodeerde objecten, de waarden die niet goed gecodeerd zijn, zijn de overige cellen. De waarde P is het totaal aantal waarnemingen dus of M + N + O of J + K + L.

Uit bovenstaande kan de attribuutnauwkeurigheid van een dataset als volgt gedefinieerd worden: (A + E + I) / P * 100

De nauwkeurigheid van classificatie van een bepaalde klasse kan als volgt worden berekend. De nauwkeurigheid van klasse 1 = A (de goed gecodeerde objecten) / M (het aantal in de database als klasse 1 gecodeerde objecten) * 100. Dit kan op dezelfde wijze voor de andere klassen gedaan worden.

De omissie van een klasse geeft in procenten weer hoeveel werkelijke waarden van een bepaalde klasse er niet als zodanig geclassificeerd zijn. Dit is het tegenovergestelde van de nauwkeurigheid. De berekening van de omissie is dan ook 100 – de nauwkeurigheid. Deze kan ook berekend worden door (D + G) / M * 100.

De overcompleetheid van een klasse geeft in procenten weer hoe vaak in de database een bepaald object een bepaalde klasse is gegeven, terwijl deze in werkelijkheid een andere klasse bezit. De overcompleetheid kan berekend worden door van een bepaalde klasse de in de database verkeerd geclassificeerde aantallen te delen door de totaal in de database geclassificeerde aantallen voor die bepaalde klasse. Bovenstaande ingevuld voor de klasse 1 in tabel 6 maakt (B + C) / J * 100.

De betrouwbaarheid van een klasse geeft in procenten weer hoe vaak voor een bepaalde klasse, een bepaalde waarde in de database ook goed geclassificeerd is. Voor klasse 1 wordt dat in het voorbeeld in tabel 6: A / J * 100.

Cohen's Kappa

Cohen's kappa is een index die door J. Cohen ontwikkeld is en die een indicatie geeft van de attribuut nauwkeurigheid van een dataset. Het idee achter deze methode is dat een gedeelte van het classificatieproces een gevolg kan zijn van toeval. Cohen's kappa index geeft aan in hoeverre een bepaalde dataset door willekeur tot stand is gekomen. De index kan variëren tussen 0 en 1, waarbij 0 volledig willekeurig betreft en 1 volledig betrouwbaar.
De waarde van de kappa index neemt af bij een toename van het aantal klassen.

De berekening van de kappa-index kan voor de inventarisatie in tabel 6 als volgt uitgevoerd worden:

- Vermenigvuldig voor iedere klasse de in de database geclassificeerde aantallen met de in het veld geclassificeerde aantallen en deel deze door het aantal in alle klassen voorkomende classificaties. Voor de waarden in tabel 6 wordt dat:
    - klasse 1 -> J * G / P
    - klasse 2 -> K * N / P
    - klasse 3 -> L * O / P

- Tel de waarden van de verschillende klassen bij elkaar op. Dit wordt de waarde Q.

- Bereken de kappa-index door de volgende formule: ((A + E + I)-Q) / (P – Q).

4.5 Genereren steekproef

4.5.1 Inleiding

Voor het vergelijken van de werkelijke waarden met de waarden in de database is het nodig een inventarisatie in het veld uit te gaan voeren. Nu kan natuurlijk het gehele bestand opnieuw ingewonnen worden en dan vergeleken worden met hetgeen in de database zit. Dit is echter veel werk en is ook niet noodzakelijk om een goed beeld te krijgen van deze vergelijking. Het is namelijk ook mogelijk om met behulp van een aantal aselect gekozen vlakken een uitspraak te doen over een complete dataset. Deze verzameling van aselect gekozen vlakken wordt de steekproef genoemd. De complete dataset, dus alle vlakken, wordt de populatie genoemd.

Het analyseren van de attribuutnauwkeurigheid met behulp van een confusionmatrix of Cohen's kappa kan op basis van een steekproef uitgevoerd worden. Voordat met het uitvoeren van de steekproef begonnen kan worden dienen er echter eerst nog een aantal vragen beantwoord te worden en dienen er diverse processen doorlopen te worden.

Er zijn een aantal keuzes te maken met betrekking tot de te nemen steekproef. Dit betreft de volgende:

- *Worden alle vlakken geïnventariseerd of alleen een deel van de vlakken*
  Door de grote diversiteit aan soorten vlakken wordt het qua tijdsplanning moeilijk om alle vlakken te gaan onderzoeken. Er wordt dus gekozen voor het nemen van een steekproef. De berekening van de grootte van deze steekproef en het hiermee praktisch omgaan wordt later behandeld in paragraaf 4.5.5.

- *Wordt er per soort onderzocht wat de attribuut nauwkeurigheid is*
  Gezien de diversiteit aan soorten en het grote verschil in voorkomende aantallen per soort is het binnen de gestelde tijd niet mogelijk om alle soorten te gaan onderzoeken. De steekproef wordt dus gemaakt zonder rekening te houden met de verschillende soorten die hierbinnen voorkomen en het aantal keren dat een bepaalde soort voorkomt. Na het onderzoek is het dus moeilijk om een conclusie over een betreffende soort te

kunnen trekken. Wel kan er een conclusie over het bestand als geheel getrokken worden, voor wat betreft de attribuut nauwkeurigheid.

- *Wordt er een steekproef genomen op basis van een ruimtelijke indeling.*
  Omdat er in de diverse gebieden door verschillende personen geïnventariseerd is, en er verschillen in status zijn tussen de diverse gebieden, is het wellicht interessant om te kijken of er verschillen in de diverse gebieden aanwezig zijn.

- *Op welke wijze dienen vlakken onderzocht te worden.*
  Wordt er trajectgewijs vlakken onderzocht of wordt er gericht naar een aantal vlakken gekeken die verspreid liggen over het gehele gebied. Het onderzoeken van het gehele gebied zal meer tijd in beslag nemen waardoor minder vlakken onderzocht kunnen worden. Afhankelijk van het aantal te onderzoeken vlakken zal bekeken worden voor welke aanpak gekozen wordt.
  Naast de keuze van welke vlakken onderzocht dienen te worden dient ook onderzocht te worden hoe en waar een vlak getoetst wordt. Met dit “hoe” wordt de methode bedoeld waarmee de toets-inventarisatie uitgevoerd gaat worden. Met het “waar” wordt bedoeld waar in het vlak de classificatie geïnventariseerd wordt. Gebeurt dit in het middelpunt van het vlak, willekeurig of op nog een andere plaats. De keuze hiervan kan eventuele gevolgen met zich meebrengen ten aanzien van het introduceren van een andere aspect van kwaliteit zoals bijvoorbeeld positionele nauwkeurigheid.

4.5.2 Status

Het bestand steen.shp is zoals te lezen valt onder paragraaf 4.2.1 onder andere benodigd voor verbeteringswerken. Inmiddels zijn een aantal van deze werken reeds uitgevoerd. Dit leidt tot de volgende verschillende situaties en daarmee samenhangend de verschillende statussen:

- Vlakken waar geen verbeteringswerken zijn uitgevoerd

- Vlakken die door werken verbeterd zijn en waarvan de wijzigingen geïnventariseerd zijn en reeds verwerkt zijn in het bestand met de steenbekledingen

- Vlakken die door werken verbeterd zijn maar waarvan de inventarisatie en bijwerking in het bestand nog niet afgerond is

Bij de verbeterde vlakken is in het kader van de verbeteringswerken van waterkeringen de waterkering opnieuw bekleed. Hierdoor zijn naar verwachting betere en accuratere informatie beschikbaar. Door de gebruikers wordt hier in ieder geval van uitgegaan en dat maakt het interessant om te controleren of dit nu ook het geval is.

Omdat bovenstaande verschillen belangrijk zijn met betrekking tot eventuele uitkomsten van het onderzoek worden de verschillende categorieën door middel van het aangeven van de statussen van elkaar onderscheiden. In bijlage 7 staat een schema en beschrijving van welke stappen gemaakt zijn om dit te realiseren.

4.5.3 (On)Zichtbaar

Op diverse plaatsen wordt na verloop van tijd zand afgezet op de waterkering of wordt de teen van de dijk bestort met bestortingsmaterialen. Hierdoor is niet te zien welk soort vlak er onder het zand aanwezig is. Deze informatie, maar ook de geometrie van het vlak is op basis van archiefstukken gegenereerd en in het steenbekledingenbestand terechtgekomen. Dit is in het processchema van de totstandkoming van het steenbekledingenbestand in bijlage 5 terug te vinden.
Omdat het geen zin heeft om onzichtbare vlakken te gaan toetsen, deze zijn immers in het veld niet terug te vinden, dient er in het bestand aangegeven te worden of een betreffend vlak zichtbaar danwel onzichtbaar is. In bijlage 8 is te zien hoe dit ten uitvoer is gebracht.

4.5.4 Gebieden

Zowel ruimtelijk als qua inventarisatie valt er onderscheid te maken in een aantal gebieden. Het betreft de volgende:

- Oosterschelde – Noord
- Oosterschelde – Zuid
- Noordzee
- Westerschelde - West
- Westerschelde - Oost

Voor de Westerschelde is de opdeling in gebieden ook al in de status vertaald. Alle vlakken in het gebied Westerschelde-West zijn “verbeterd” terwijl alle vlakken in het gebied

Westerschelde-Oost ”inbewerking” zijn. Bij het toekennen van de gebiedsinformatie is deze statusinformatie gebruikt. In bijlage 9 staat in het processchema beschreven hoe de toedeling in bovenstaande gebieden is uitgevoerd.

4.5.5 Grootte van de steekproef

Voor het berekenen van de grootte van de steekproef zijn een aantal elementen van belang. Het betreft de volgende:

- *Hoe groot is de populatie?*
  De populatie is 5639.

- *Wat is de gewenste betrouwbaarheid?*
  De gewenste betrouwbaarheid is 95%.

- *Wat is de gewenste nauwkeurigheid?*
  De gewenste nauwkeurigheidsmarge = 5%.

- *Wat is de proefschatting of een vlak goed gecodeerd is?*
  Ingeschat wordt dat deze circa 80% is.

Uit bovenstaande volgt dat:

a = 0.05 (is 5% nauwkeurigheid)
z = 1,96 (behorend bij 95% betrouwbaarheid)
p = 0,8 (is 80% proefschatting)

De formule voor het berekenen van de steekproefgrootte is:

$$\text{steekproefgrootte } n = \frac{z^2 \cdot p(1-p)}{a^2}$$

$$= \frac{(1{,}96)^2 \cdot 0{,}8\ (1-0{,}8)}{0{,}0025}$$

$$= \frac{0{,}61}{0{,}0025}$$

$$= 245{,}9$$

Als dit aantal groter is dan 10% van de populatie mag er nog een correctiefactor toegepast worden. Aangezien 245,9 kleiner is dan 10% van 5639 mag deze correctiefactor in dit onderzoek niet toegepast worden.

Er dienen dus minimaal 246 vlakken getoetst te worden om de gestelde nauwkeurigheid en betrouwbaarheid te behalen en de steekproef representatief te laten zijn voor het gehele bestand.

4.5.6 Kiezen van vlakken

Nu de steekproefgrootte bepaald is, is de volgende stap te onderzoeken hoe hier praktisch invulling aan kan worden gegeven.
In paragraaf 4.5.1 is reeds aangegeven dat er een keuze zal moeten worden gemaakt tussen of een beperkt aantal vlakken onderzoeken of het onderzoeken van een of meerdere trajecten. Gezien de grootte van de steekproef en de geplande tijdsinspanning is het niet uitvoerbaar om de toets op willekeurig in het gebied verspreide vlakken te gaan uitvoeren. Dit betekent dat er in het kader van uitvoerbaarheid van het onderzoek er helaas gezondigd wordt op de aselectheid van de gekozen vlakken. Wel is getracht deze zondiging zo gering mogelijk te houden.

Een praktische oplossing is om te kiezen voor trajecten. De geplande tijd kan dan nuttig besteed worden in plaats van veel reisuren te maken en weinig inventarisatiewerkzaamheden uit te voeren. Omdat de inventarisatie ruimtelijk van aard is, ligt het voor de hand om ook de keuze voor de te onderzoeken vlakken ruimtelijk te maken.

Hierdoor is er voor de aanpak gekozen om diverse vlakken te kiezen en rond deze vlakken andere te toetsen vlakken te kiezen. Op deze manier blijft de keuze van de te toetsen vlakken grotendeels aselect maar is de uitvoering toch praktisch.

Voor het bepalen van welke vlakken getoetst dienen te gaan worden zijn de volgende uitgangspunten gebruikt:
Het onderzoeksgebied is opgedeeld in 5 deelgebieden. Van de deelgebieden 3 en 5 is geen informatie over de steenbekledingen beschikbaar. Het heeft geen nut deze te gaan toetsen. Alleen dus de deelgebieden 1,2 en 4 kunnen worden dus getoetst. Dat betreft de deelgebieden Oosterschelde-Noord, Oosterschelde-Zuid en Westerschelde-West. De redenen om deze geografische deelgebieden ook als te toetsen deelgebied te beschouwen zijn de volgende:

- Het gedeelte van de Westerschelde betreft een verbeterd gedeelte. Verwacht wordt dat de informatie over dit gedeelte beter is dan bij de niet verbeterde gedeeltes.
- In de drie gekozen deelgebieden zijn verschillen in personen die geïnventariseerd hebben. Zoals de deelgebieden nu gekozen zijn, geeft dat eenduidigheid voor wat betreft de persoon van inventarisatie.
- Door het meten van de soort steenbekleding binnen de steekproef die gebaseerd is op de 3 deelgebieden wordt er geen structurele fout geïntroduceerd, maar maakt het de praktische invulling van inventarisatie wel een stuk eenvoudiger.

Het verschil in aantallen vlakken tussen de deelgebieden is groot. Zie hiervoor de tabel hieronder.

Deel	Omschrijving	Aantal
1	Oosterschelde-Noord	1849
2	Oosterschelde-Zuid	1402
3	Noordzee	342
4	Westerschelde-West	581
5	Westerschelde-Oost	379

Tabel 7 De aantallen vlakken per gebied

Op basis van bovenstaande tabel is er gekozen voor het kiezen van vlakken in een bepaalde verhouding. Deze verhouding is voor de deelgebieden Oosterschelde-Noord Oosterschelde-Zuid, en Westerschelde-West respectievelijk 2 : 2 : 1. In de deelgebieden Oosterschelde-Noord en Oosterschelde-Zuid worden in ieder gebied 2 toetszoneringen gemaakt. In Westerschelde-West wordt er één toetszonering gemaakt.

Het kiezen van de te toetsen vlakken binnen de toetszoneringen is uitgevoerd door binnen ieder deelgebied één of twee vlakken te selecteren. Dit aantal is afhankelijk van welk deelgebied het betreft. Dit selecteren is gedaan door binnen Microsoft Excel met de aselect functie. Met behulp van deze functie kan binnen een range een willekeurig getal gekozen worden. Voor deelgebied 1 is er bijvoorbeeld een vlak gekozen uit id's 1 tot en met 1849. Aangezien de aselectfunctie werkt met behulp van floating getallen dient er gekozen worden voor een getal tussen de 0,5 en de 1849,5. De functie ziet er dan als volgt uit:

randomgetal = aselect() * (b-a)+a.

In deze formule is *a* de onderste begrenzing. Met het getal *b* wordt de bovenste begrenzing vastgelegd.

De ingevulde formule ziet er dan als volgt uit:

Randomgetal = aselect () * (1849,5 – 0,5) +0,5
In de tabel hieronder staan de willekeurig gekozen waarden.

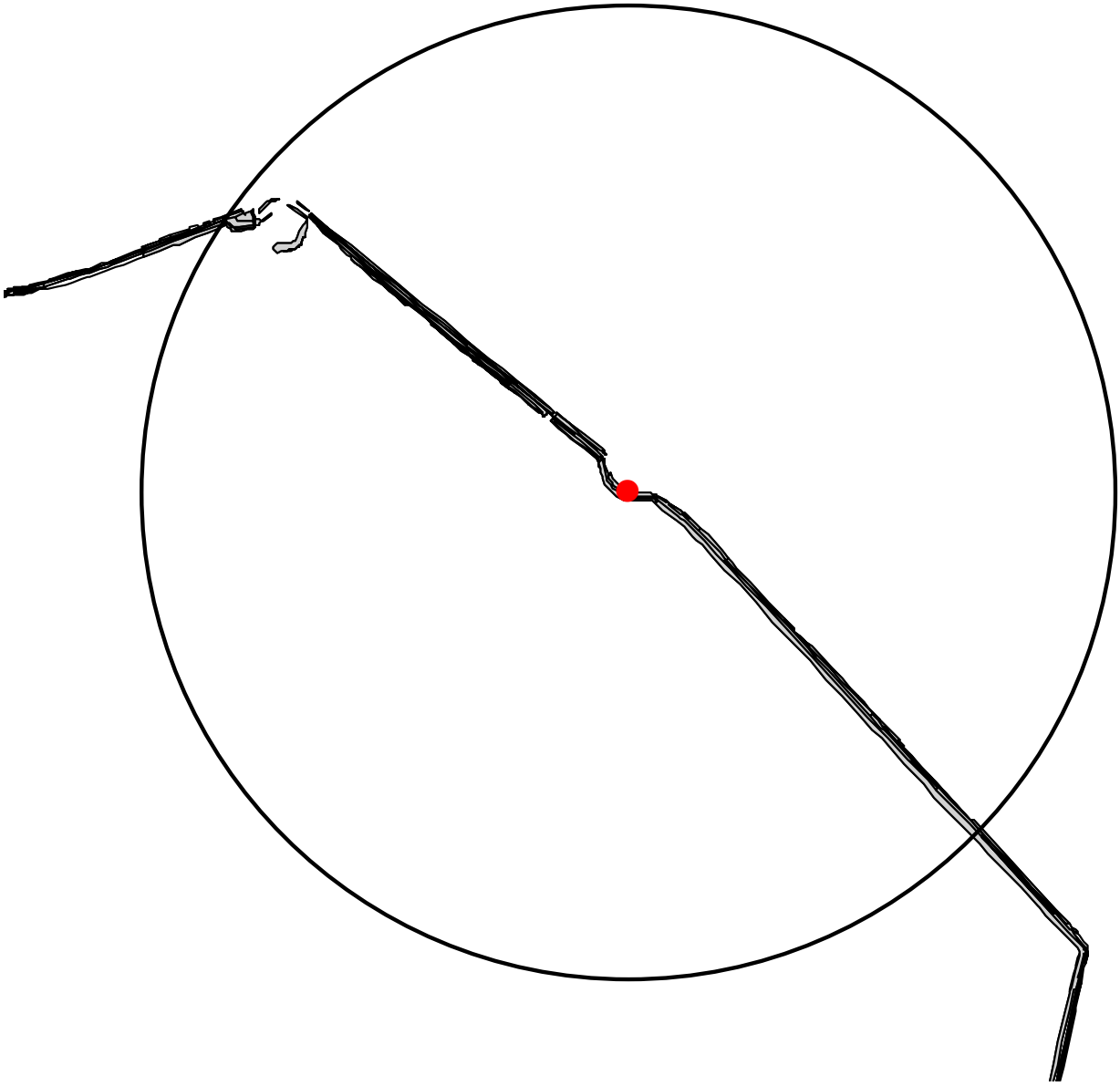
Deel	Omschrijving	Id
1	Oosterschelde-Noord	1465
1	Oosterschelde-Noord	69
2	Oosterschelde-Zuid	623
2	Oosterschelde-Zuid	780
4	Westerschelde-West	532

Tabel 8 De aselect gekozen vlakken

De te toetsen vlakken binnen de toetszoneringen zijn bepaald door met het gekozen vlak als middelpunt een cirkel te trekken met een dusdanige afmeting dat er binnen de cirkel 60 te toetsen vlakken vallen. Dit aantal van 60 is gebaseerd op de benodigde steekproefgrootte. In totaal wordt er nu 60 * 5 = 300 vlakken getoetst. Dit is ruim voldoende om een representatieve toets uit te voeren. In figuur 10 is bovenstaande methodiek voor een van de toetszoneringen van het deelgebied Oosterschelde-Noord weergegeven.

In theorie is het mogelijk dat er binnen de 2 * 60 te kiezen vlakken binnen hetzelfde deelgebied, dezelfde vlakken voorkomen, dus overlappen in toetszoneringen. Het komt er

dan op neer dat er geen 120 vlakken gekozen worden maar minder. Mocht dat voorkomen dan wordt de cirkel van het eerstgekozen willekeurige vlak dusdanig vergroot dat er toch 120 te toetsen vlakken bepaald worden.
Op deze manier blijft de keuze van de vlakken willekeurig maar wordt wel aangesloten op de steekproefgrootte.



Figuur 10 De methodiek voor het selecteren van de te toetsen vlakken

In bijlage 10 zijn in een processchema alle stappen weergegeven om rond de aselect gekozen vlakken de andere 295 te toetsen vlakken te kiezen.

4.6 Methode van inventarisatie en verwerking

Om de inventarisatie binnen gestelde tijd en op een juiste wijze uit te voeren is er voor de methode van inventarisatie en verwerking een aantal eisen gedefinieerd. Het betreft de volgende:

- De inventarisatie moet een hoge juistheidfactor bezitten. Dit betekent dat er gekozen moet worden voor een methode die betrouwbaar is en die een geringe kans op fouten met zich meebrengt.
- De inventarisatie moet in betrekkelijk weinig tijd uitgevoerd kunnen worden.
- De verwerking achteraf en het vergelijken van de inventarisatie met de waarden in de database moet redelijk snel uitgevoerd kunnen worden.

Naast bovenstaande eisen bestaat er wel de uitdaging om met moderne technieken en mogelijkheden te werken. Bij gelijke geschiktheid van mogelijkheden gaat de voorkeur uit naar het kiezen van een moderne oplossing.

Grofweg bestaan er voor het inventariseren van attributen van ruimtelijke informatie drie methoden. Dat betreft analoge, digitale en een combinatie van analoge en digitale methodes. Hieronder vindt een beschrijving van de ruimtelijke inventarisatie methoden plaats:

Analoge inventarisatiemethode
De analoge methode van het inventariseren van attribuutdata van ruimtelijke objecten betreft het op basis van een analoge kaart inventariseren van op de kaart aanwezige objecten. Op de kaart is naast de objecten ook topografische data geplaatst om in het veld te kunnen oriënteren. De inventarisatie van de attribuutdata vindt plaats door middel van het toekennen van een unieke identificatiesleutel per object en het analoog noteren van de attribuutwaarden.

Gecombineerde analoog/digitale inventarisatiemethode
Deze inventarisatiemethoden kunnen in diverse vormen voorkomen. In vergelijking tot de analoge inventarisatiemethode is het bijvoorbeeld mogelijk om of het noteren van de attribuutwaarden digitaal met een portable computer uit te voeren, of een digitale kaart op een portabele computer mee het veld in te nemen.

Digitale inventarisatiemethode

Bij deze inventarisatiemethode is zowel de benodigde kaart om te oriënteren als het noteren van de attribuutwaarden in het veld volledig digitaal. Toch zijn hier ook nog diverse mogelijkheden:

- Kaart en notatie beschrijvende informatie gescheiden, oriëntatie op basis van topografie in de kaart
- Kaart en notatie beschrijvende informatie gescheiden, oriëntatie op basis van GPS en topografie
- Kaart en notatie beschrijvende informatie gekoppeld, oriëntatie op basis van topografie in de kaart
- Kaart en notatie beschrijvende informatie gekoppeld, oriëntatie op basis van GPS en topografie

Onafhankelijk van soort inventarisatiemethode gelden de volgende uitgangspunten:

- Het verwerken van een analoge inventarisatie kost meer tijd dan wanneer dit reeds digitaal in het veld wordt uitgevoerd.
- Het werken met gescheiden kaart en beschrijvende informatie, onafhankelijk of dit nu digitaal of analoog is, werkt fouten in de hand.
- Oriëntatie op basis van GPS biedt meer voordelen dan oriëntatie op basis van alleen topografie. Ondanks de soms nog beperkte nauwkeurigheid van de GPS in relatie met de gedetailleerdheid van de topografie kan de GPS in ieder geval gebruikt worden voor een globale orientatie. Omdat er vlakken getoetst dienen te worden die bijvoorbeeld maar 1 meter breed zijn en er beschikt kan worden over een GPS waarbij de afwijking ook ongeveer die grootte bedraagt, kan er niet blind op de GPS gevaren worden. Het is echter mogelijk de GPS te gebruiken als aanvulling is op de topografie. Dit kan door onder de GPS positiebepaling ook de topografie te leggen.

Op basis van bovenstaande is er gekozen voor de digitale inventarisatiemethode. Daarnaast is er vanwege de kleinere kans op fouten ook gekozen voor een geïntegreerde kaart en beschrijvende informatie. Vanwege een optimale positiebepaling zal er ook tijdens de inwinning gebruik gemaakt worden van een GPS-ontvanger. Daarnaast zal in het veld ook (digitale) topografie gebruikt worden ter bevordering van de nauwkeurigheid van de oriëntatie.

Bovenstaande digitale inwinmethode gecombineerd met GPS-ontvanger is verenigd in een Geo-explorer XT van Trimble. Op de volgende pagina is een foto hiervan afgebeeld. Op

deze geo-explorer is Arc-Pad beschikbaar. Dit programma is geschikt als mobiele GIS-omgeving waarin de door de GPS gemeten positie zichtbaar is.

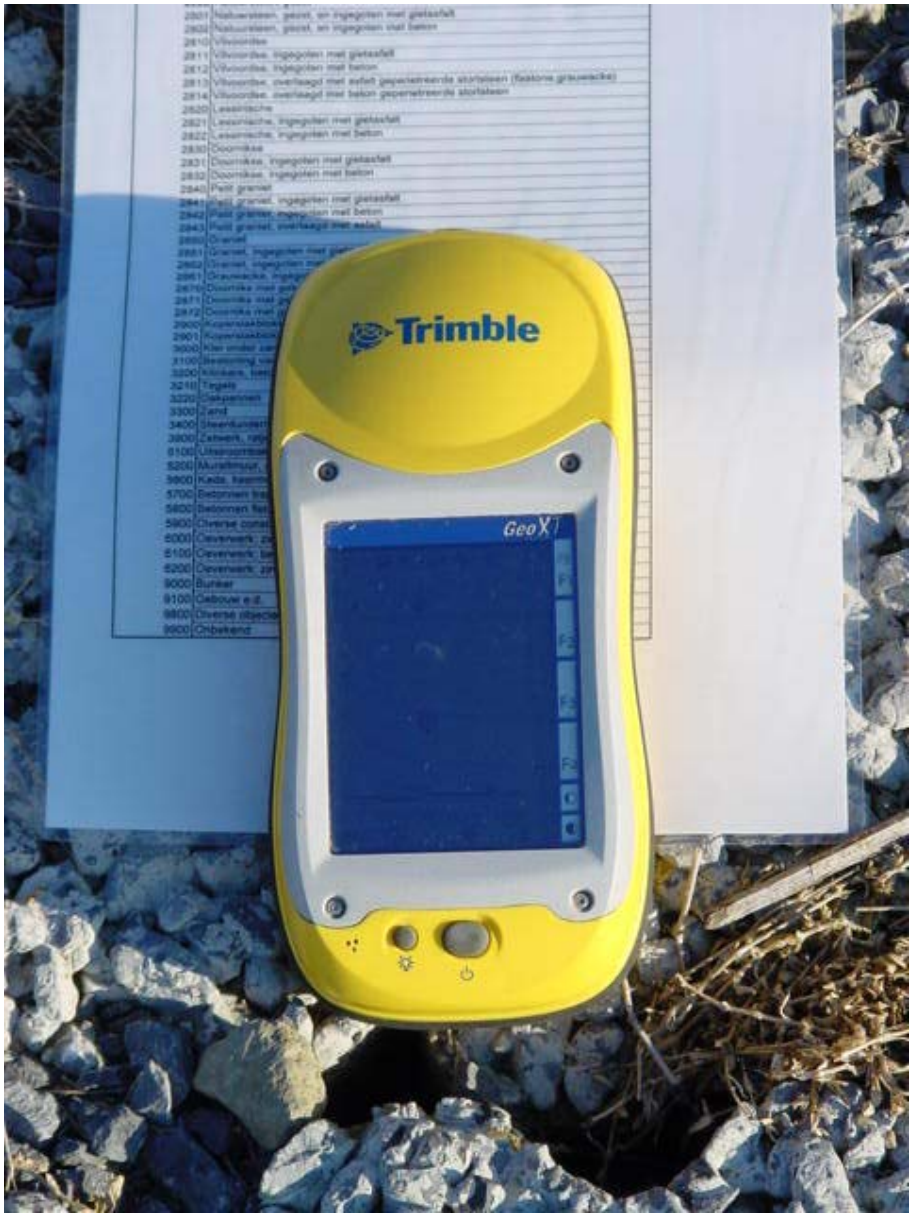


Foto 4 De geo-explorer met inventarisatielijst op een ondergrond van basalton met ecolaag

4.7 Voorbereiding inventarisatie

4.7.1 Inleiding

Om in het veld de inventarisatie op een juiste en adequate wijze uit te voeren, zijn diverse voorbereidingen nodig. Tevens zijn deze nodig om in een later stadium de kans zo gering mogelijk te laten zijn dat er niet voldoende of op een verkeerde wijze ingewonnen is.

In de volgende paragrafen worden alle voorbereidende acties inclusief nut of noodzaak beschreven.

4.7.2 Gereedmaken steekproefbestand

Uitgangspunt voor de toetsing in het veld is de shapefile Steekproef.shp. Dit bestand bestaat uit diverse vlakken die getoetst dienen te worden. Met het toetsen van deze vlakken komt er echter direct een ander probleem om de hoek kijken. Waar moet de soort steenbekleding in het vak getoetst worden? Als dit op diverse plaatsen in het vlak gebeurt, wordt het onderzoek naar attribuutnauwkeurigheid vertroebeld doordat er zich fouten in de begrenzingen kunnen voordoen. Om dit probleem te ondervangen is er met behulp van Arc-View en ARC/INFO voor ieder vlak een centroid gecreëerd. Om tevens een objectieve toets te kunnen uitvoeren wordt er geen steenbekleding-informatie aanwezig in de database mee het veld in genomen. Wel worden er een aantal kolommen toegevoegd om de leeftijd, oppervlakte kleiner dan 5 m², het steenbekledingstype en een opmerking in te kunnen vullen. Om geen fouten te maken met decimalen worden alle coderingen met een factor 100 vermenigvuldigd. Hierdoor dienen er voor de soort steenbekleding alleen gehele getallen ingevoerd te worden. Deze lijst is opgenomen in bijlage 11.
In bijlage 12 is het complete processchema om tot dit centroidenbestand te komen, weergegeven.

4.7.3 Gereedmaken topografie

Voor de oriëntatie in het veld wordt gebruik gemaakt van de geo-explorer met behulp van de GPS die hierin aanwezig is. Daarnaast is er echter ook de topografie van de waterkering meegenomen. Vanwege het makkelijk kunnen toekennen van “symbology” aan lijnenbestanden is er gekozen om de verschillende steenzettingsvlakken als lijnenweergave mee te nemen in het veld en niet als vlakweergave.

4.7.4 Oriënteren gebruik Geo-Explorer

Voordat aan een inventarisatie begonnen kan worden, is het noodzakelijk om een oriëntatie uit te voeren met betrekking tot het gebruik van de Geo-Explorer. Deze oriëntatie betreft de volgende onderdelen:

- Hoe kan te onderzoeken data naar de Geo-explorer gebracht worden en hoe kan geïnventariseerde data naar een werkstation overgebracht worden. Ten behoeve van het onderzoek kwam dit proces neer op het overbrengen van de steekproefcentroiden en de topografische data naar de Geo-Explorer. Na het inventariseren dienden de geïnventariseerde informatie die gekoppeld is aan de centroiden van de vlakken naar het werkstation geconverteerd te worden.
- In een Geo-explorer is het mogelijk om een aantal instellingen te doen die de nauwkeurigheid van de meting bepalen. In het onderzoek is gebruik gemaakt van de op een na nauwkeurigste instelling. Dit betreft een 3 dimensionale meting.
- Er zijn meerder mogelijkheden om de Geo-explorer in te richten en de inwinning uit te voeren en te registreren. Hierin dient de meest wenselijke methode en lay-out uitgedacht te worden. Onder dit punt kunnen tal van items vallen. Voorbeelden zijn onder andere:
    - welke functies dienen er beschikbaar te zijn
    - welke toetsenbord werkt het gemakkelijkst
    - op welke schaal kan er het gemakkelijkst geïnventariseerd worden
- Hoe kan er voor gezorgd worden dat de positiebepaling meeloopt met de inventarisatie
- Hoe lang kan met een accu geïnventariseerd worden.

Al bovenstaande punten dienen vooraf getest te worden. In het veld dient er zoveel mogelijk geïnventariseerd te worden en zo weinig mogelijk tijd aan het bijstellen van zaken.

In bijlage 13 is aangegeven hoe de betreffende data naar de Geo-explorer geconverteerd is. Hierbij is gebruik gemaakt van Arc-Pad. Na deze conversie is alles gereed voor inventarisatie.

4.8 Inventarisatie in het veld

Gedurende twee dagen zijn de steenbekledingssoorten van de vlakken behorend bij de steekproef geïnventariseerd. Dit inventariseren is door twee personen gedaan. Persoon 1 heeft de bediening van de Geo-explorer en de oriëntatie in het veld voor zijn rekening genomen. Persoon 2 heeft het analyseren van de soort steenbekleding uitgevoerd. Hieronder zijn de verschillende handelingen beschreven.

- Het opstarten van de Geo-explorer. Voor het zoeken van de te gebruiken satellieten is het nodig om circa 10 minuten voor aanvang van de inventarisatie de geo-explorer aan te zetten (actie persoon 1).

- Bepalen welk vlak geïnventariseerd gaat worden (actie persoon 1).
- Met behulp van de geo-explorer de centroide van betreffend vlak opzoeken en naar deze plaats toenavigeren. Op de Geo-explorer is het mogelijk om de positie op het scherm mee te laten lopen met de actuele positie in het veld. Op deze wijze is het vrij gemakkelijk om de plaats van een centroide in het veld te bepalen (actie persoon 1).
- Het beoordelen van de soort steenbekleding. Dit laatste bleek in een aantal gevallen erg lastig te zijn. Af en toe waren er hulpmiddelen benodigd zoals een plamuurmes om er achter te komen welke steensoort van toepassing was (actie persoon 2).
- Opzoeken van de van toepassing zijnde steensoort in een tabel (actie persoon 2). In deze tabel is voor iedere steensoort een codering weergegeven. Voor het gemak zijn alle waarden met de factor 100 vermenigvuldigd. Hierdoor hoeven er in het veld alleen gehele getallen ingevoerd te worden waardoor dit eenvoudiger wordt en wellicht dus ook de kans op tikfouten verminderd.
- Opslaan van de codering van de soort steenbekleding in de Geo-explorer. Indien van toepassing dan is ook een opmerking ingevuld en is aangegeven indien een vlak kleiner is dan 5 m² en of de leeftijd jonger is dan 5 jaar (actie persoon 1).

Hierna kan doorgegaan worden met het inventariseren van het volgende vlak.

Vanwege de redelijk onbekendheid met de geo-explorer is er voor gekozen om geïnventariseerde informatie regelmatig uit te lezen. De reden om dit te doen was om geïnventariseerde informatie te waarborgen en dus niet het risico te lopen deze in een later stadium kwijt te raken door diverse omstandigheden zoals bijvoorbeeld het defect of zoek raken van de Geo-explorer.Tijdens de inventarisatie zijn er aan een bestand waarden toegevoegd. Van dit bestand is echter drie maal een back-up gemaakt:

- Aan het einde van de eerste dag → dag1
- Aan het eind van de morgen van de tweede dag → dag 2a
- Aan het eind van de middag van de tweede dag → dag 2b

4.9 Verwerking inventarisatie

Na de inventarisatie in het veld dient deze informatie dusdanig verwerkt te worden dat zij vergeleken kan worden met de informatie die aanwezig is in de database. Om dit te kunnen doen is zij geconverteerd van de Geo-explorer naar het werkstation.
Hierna zijn alle drie de bestanden geconverteerd naar een shapefile en vervolgens samengevoegd tot het bestand “Merge1.shp” zie voor het processchema bijlage 14.

Merge1 is het bestand waarin alle geïnventariseerde informatie van de twee dagen in aanwezig is.

4.10 Vergelijking van waarden en analyse soort fouten

In bijlage 15 zijn de diverse acties weergegeven die het vergelijken mogelijk maken van geïnventariseerde waarden en waarden zoals die aanwezig zijn in de database. Als eerste is er begonnen met het koppelen van de inventarisatiewaarden aan het oorspronkelijke steekproefbestand. Dit is gedaan door op basis van een gemeenschappelijk veld, de codering van het vlak, een join te leggen.
Hierna zijn er twee kolommen aangemaakt. De eerste kolom “bkecode*100” is nodig om ook de waarden van de steenbekleding in de oorspronkelijke database een factor 100 groter kunnen maken. Pas dan kunnen de oorspronkelijke waarden vergeleken worden met de geïnventariseerde waarden. In de tweede kolom “verschil” is het verschil berekend tussen deze oorspronkelijke waarden en de geïnventariseerde waarden.
Voor het gemakkelijk kunnen bekijken van de soort fout en het invullen van de fouten in de confusionmatrix is er ook een kolom gemaakt waarin zowel de oorspronkelijke als de geïnventariseerde steenbekledingssoort is weergegeven. De naam van deze kolom is “inv+org”.

De volgende stap in het proces is het beoordelen tot welke categorie verschillen tussen inventarisatie en database behoren. Binnen deze fouten valt er onderscheid te maken tussen fouten die onafhankelijk van de vlaksoort zijn gemaakt of vlakken die afhankelijk van de vlaksoort zijn gemaakt. Alleen van de fouten die gemaakt zijn tijdens de inventarisatie heeft de fout een relatie met de steenbekledingssoort. De andere fouten zijn onafhankelijk van steensoort tot stand gekomen. De volgende categorieën worden onderscheiden

1. *Onzichtbare vlakken*
   Tijdens het onderzoek zijn er vlakken geselecteerd voor de steekproef waarvan de zichtbare status in het veld niet klopte. Een aantal vlakken bleek onder breuksteen of zand te liggen en was dus onzichtbaar. Deze fout in het zichtbaar/onzichtbaar definiëren, heeft geen invloed op dit onderzoek. Van deze vlakken kon immers de classificatie niet achterhaald worden en viel dus ook niet te vergelijken. Dit aantal vlakken is van de steekproef afgetrokken. In totaal betroffen dit 3 vlakken zodat de grootte van de steekproef hiermee op 297 is uitgekomen.

2. *Fouten als gevolg van nalaten van actualisatie*
   Het complete inwintraject heeft plaatsgevonden in de jaren 1999 en 2000. Op een aantal plaatsen hebben inmiddels veranderingen plaatsgevonden. Om deze te kunnen registreren is tijdens de inventarisatie van het onderzoek de leeftijd jonger of ouder dan 5 jaar meegenomen. Hierdoor kunnen fouten in het bestand, voortkomend uit een gebrek aan actualisatie, vastgesteld worden.

3. *Fouten in interne verwerkingsproces*
   Na de inventarisatie in 1999/2000 zijn er diverse acties op kantoor uitgevoerd. Deze zijn terug te vinden in het processchema opgenomen in bijlage 5. In dit interne proces zijn ook fouten gemaakt. Omdat de inventarisatieformulieren nog steeds beschikbaar zijn, is het mogelijk deze te achterhalen. Is er wel goed geïnventariseerd (te controleren op de inventarisatieformulieren), en betreft het geen onzichtbaar vlak of is het een fout uit een van de andere categorien, dan is er sprake van een fout in het interne verwerkingsproces. Fouten die hierbinnen vallen, kunnen de volgende zijn:
   - Verwisseling van formulieren
   - Fouten tijdens digitaliseren van formulieren, bijvoorbeeld transcriptiefouten

4. *Fouten in conversieproces*
   Na het gereedkomen van het bestand in 2000 is het bestand in een afgeleide vorm gebruikt ten behoeve van het uitvoeren van de toetsing. Het bestand ten behoeve van INTWIS is pas 2 jaar later na diverse bewerkingen gereedgekomen. De fouten die binnen deze categorie vallen zijn voornamelijk ontstaan door het verwisselen van identificatiecoderingen en/of het verschuiven van labels. In bijlage 16 is aangegeven hoe de twee bestanden met elkaar vergeleken zijn.

   Het bestand dat gebruikt is voor het uitvoeren van de toetsing van de waterkering op veiligheid, bestaat uit vier verschillende bestanden. Deze bestanden zijn samengevoegd tot een bestand. Omdat de originele toetsbestanden het gehele gebied van waterschap Zeeuwse Eilanden betreft, is met behulp van een clipactie, het samengevoegde toetsbestand uitgesneden op het steekproefbestand (resultaat). Dit heeft geresulteerd in het bestand “controleclip.shp”. Hierna is zowel de informatie van de inventarisatie + database (uit resultaat.shp) alsmede de toetsinformatie (uit controleclip.shp) door middel van een spatial join gekoppeld aan de centroiden.
   Alle vlakken waar de identificatiecoderingen uit “resultaat.shp” niet overeenkomen met “controleclip.shp” zijn fouten veroorzaakt tijdens het conversieproces.

5. *Fouten in inventarisatie in het veld*
   Tijdens de inventarisatie in het veld zijn er fouten gemaakt. Dit kunnen zowel fouten zijn die betrekking hebben op het verkeerd beoordelen van de soort steenbekleding, het verkeerde vlak beoordelen alsmede het verkeerd noteren van de soort steenbekleding.

   Omdat deze fouten diverse soorten van importantie bezitten zijn de fouten met de waarden 5 opgesplitst in de volgende categorieën:
   Fout 51 → fouten waarbij alleen de penetratie van de steenbekleding verkeerd is geïnventariseerd
   Fout 52 → fouten waarbij de soort steenbekleding verkeerd is gekozen maar die wel binnen dezelfde categorie steenbekleding vallen. De eerste twee nummers van de steenbekleding geeft de categorie aan. Bijvoorbeeld alle 2800 nummers betreffen natuursteen.
   Fout 53 → fouten waarbij een verschillende soort steenbekleding uit een verschillende categorie geïnventariseerd is.

   Bovenstaande fouten zijn verschillend van importantie. Fouten 51 zijn de minst zware fouten, fouten 53 zijn de zwaarste fouten.

6. *Fouten door een gebrek aan duidelijke definitie*
   In een aantal gevallen blijkt er sprake te zijn van verschillende interpretaties. De definities zijn niet altijd duidelijk. Fouten of onduidelijkheden hierdoor vallen onder deze categorie.

7. *Fout door het onbekend zijn van de steenbekledingssoort*
   Van een vlak van de 297 getoetste vlakken was in de database de steenbekledingssoort niet bekend. Waarom de waarde niet is ingevuld kon niet worden achterhaald.

Om de soort fout in het bestand te kunnen verwerken is de kolom “soort_fout aangemaakt. In bijlage 17 is schematisch weergegeven hoe het invullen van de diverse soorten fouten in is uitgevoerd.

Na de inventarisatie zijn van niet te verklaren fouten door bijvoorbeeld gebrek aan actualisatie of verkeerd overnemen van het formulier, de bijbehorende vlakken nogmaals in het veld gecontroleerd. Na deze extra controle blijven er een aantal verschillen bestaan tussen de waarden in de database en de inventarisatie die alleen te wijten kunnen zijn aan een verkeerde inventarisatie in het terrein. In het processchema in bijlage 18 zijn de verschillende stappen op een rij gezet.

4.11 Analyse resultaat kwalitatief onderzoek

4.11.1 Resultaat

Op basis van de vergelijking tussen de inventarisatie en de waarden zoals die in de database voorkomen is een “confusion matrix” gemaakt zoals deze in paragraaf 4.4.2 behandeld is. Deze confusionmatrix is in bijlage 19 opgenomen. Op basis van de confusionmatrix is ook Cohen’s Kappa uitgerekend. Deze heeft een waarde van 0,7. Dit geeft aan dat een gedeelte van het classificatieproces het gevolg kan zijn van willekeur en dat de classificatie dus niet geheel betrouwbaar is. In het algemeen wordt een waarde van 0,7 aangemerkt als een “voldoende”. Hierbij dient echter aangemerkt te worden dat de waardering “voldoende” alleen gegeven kan worden in relatie tot de toepassing waarvoor het inwin- of classificatieproces uitgevoerd is.

In de confusionmatrix zijn de in het veld geclassificeerde vlakken uitgezet tegen de in de database geclassificeerde vlakken. Uit de confusionmatrix blijkt dat ruim 72 % van de 297 getoetste vlakken in de database correct geclassificeerd zijn. Dit betekent dat bijna 28 % verkeerd geclassificeerd is. Wat de consequenties hiervan zijn voor het gebruiken van het bestand wordt in paragraaf 4.12 behandeld. Daar wordt de benodigde kwaliteit vastgesteld en vergeleken met de vastgestelde kwaliteit.

Ook is in de confusionmatrix per klasse procentueel aangegeven hoe vaak er sprake is van omissie en overcompleetheid en hoe het met de betrouwbaarheid en nauwkeurigheid is gesteld. Hier geldt echter dat deze waarden vertroebeld worden door de grote verschillen in de aantallen getoetste soorten. Wel kan er voor een aantal soorten een behoorlijke indicatie gegeven worden voor wat betreft de nauwkeurigheid en betrouwbaarheid van classificatie en hiermee samenhangend respectievelijk de omissie en overcompleetheid.

Verder kan in de confusionmatrix ook globaal inzicht verkregen worden welke klassen vaak met welke klassen verwisseld worden. Dit inzicht wordt echter door de volgende twee redenen vertroebeld:

- Er zitten grote verschillen in de aantallen malen dat een bepaalde soort steenbekleding getoetst is. Dit varieert van een steenbekledingssoort die in de database 0 keer voorkomt

(maar wel in het veld is tegengekomen) tot een soort die 42 malen voorkomt. Van de soorten die in het veld zijn tegengekomen variëren die aantallen tussen de 0 (maar die wel in de database voorkomt) en de 44 aantallen.

- Behalve de fouten die zijn ontstaan uit een verkeerde inventarisatie missen de andere fouten een argument om bepaalde klassen met elkaar te verwisselen en dus ook een argument om hier dus een logica achter te zoeken. Waarom zou het verwisselen van een identificatie bij een bepaalde bekledingssoort vaker voorkomen dan bij een andere bekledingssoort. In de confusionmatrix zijn echter alle soorten fouten door elkaar heen opgenomen.

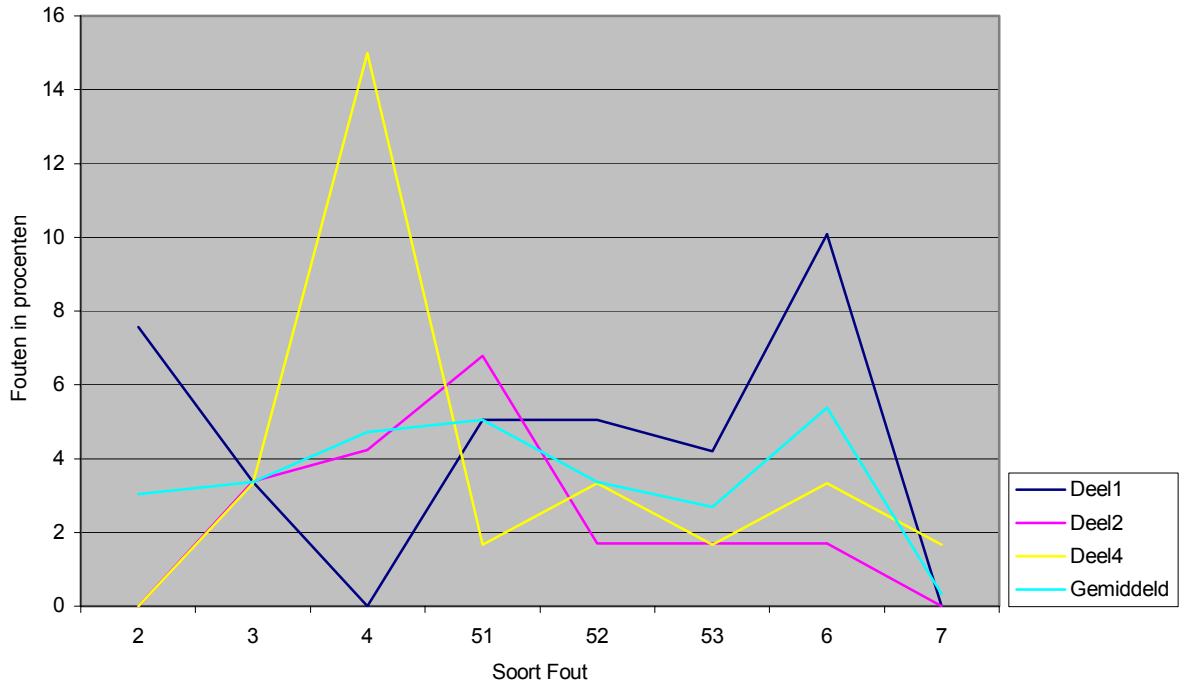
Ondanks de eerste reden van vertroebeling is er in bijlage 20 een confusionmatrix opgenomen met hierin alleen de fouten uit de klassen 51, 52 en 53. Bepaalde verwisselingfouten lijken gemakkelijker gemaakt te worden dan andere verwisselingfouten. Ondanks de grote verschillen in de aantallen die binnen de bekledingssoorten getoetst zijn, vallen een paar dingen op.

- Fout 51, de verwisseling van de soort penetratie, komt bijna alleen bij basalt (2600) en vilvoordse (2810) voor. In totaal zijn er 15 van deze fouten gevonden. Hiervan kwamen er 7 binnen de steenbekledingssoort basalt voor, 6 binnen de steenbekledingssoort vilvoordse en 2 binnen de steenbekledingssoort doornikse.
- Fout 52, de verwisseling van steenbekledingen binnen dezelfde categorie van steenbekledingssoorten, komt voornamelijk voor bij natuursteen. Van de 9 gevonden fouten komen er 8 voor binnen de categorie natuursteen.
- Fout 53, de verwisseling van vlakken buiten dezelfde categorie steenbekleding, komt alleen bij de eerste vlakken uit de confusionmatrix voor. Dit betekent bij de soorten Breuksteen (700), Betonblokken en platen (1100 + 1400), gras (2000) en basalt (2600 + 2602). Bij de andere vlakken komt deze fout niet voor.

In de hieronder geplaatste tabel 9 is zowel in absolute aantallen als procentueel aangegeven hoe vaak een bepaalde soort fout in welk deel is waargenomen. Onder de tabel is dit grafisch in een grafiek weergegeven.

onderdeel	deel1		deel2		deel4		totaal	
	absoluut (aantal)	% (aantal)	absoluut (aantal)	% (aantal)	absoluut (aantal)	% (aantal)	absoluut (aantal)	% (aantal)
goed geclassificeerd	77	64.7	95	80.5	42	70.0	214	72.1
fout te wijten aan nalaten actualisatie	9	7.6	0	0.0	0	0.0	9	3.0
fout in het interne verwerkingsproces	4	3.4	4	3.4	2	3.3	10	3.4
fout in het conversieproces	0	0.0	5	4.2	9	15.0	14	4.7
inventarisatiefout mbt penetratie	6	5.0	8	6.8	1	1.7	15	5.1
inventarisatiefout mbt steensoort binnen dezelfde categorie	6	5.0	2	1.7	2	3.3	10	3.4
inventarisatiefout mbt steensoort buiten dezelfde categorie	5	4.2	2	1.7	1	1.7	8	2.7
fout door gebrek aan definitie	12	10.1	2	1.7	2	3.3	16	5.4
fout door het onbekend zijn van de bekledingssoort in de database	0	0.0	0	0.0	1	1.7	1	0.3
Totaal	119	100.0	118	100.0	60	100.0	297	100.0

Tabel 9 De soort fout in procenten gemiddeld en per deel



Figuur 11 De soort fout in procenten gemiddeld en per deel

Uit zowel de tabel als de grafiek vallen een aantal zaken af te leiden. Het betreffen de volgende:

- Er zitten grote verschillen tussen de aantallen goed geclassificeerde vlakken per deel. Hierbij komt deel 2, Oosterschelde-Zuid als beste naar voren terwijl deel 1, Oosterschelde-Noord, als slechtste naar voren komt. deel 4, de Westerschelde valt hier ongeveer tussenin.

- Om te verduidelijken hoe per onderdeel de score is geweest is dit in tabel 9 in kleur per deel aangegeven. Hierbij is in groen de beste score weergegeven, in oranje de tweede score en in rood de minst beste score. Hier valt op dat deel 1 in 5 van de 8 soorten fouten als slechtste scoort waarvan een keer dan wel gedeeld slechtste. De delen 2 en 4 zijn hierin ongeveer vergelijkbaar.

- Vooral uit de grafiek is goed af te lezen dat bij deel 4 veel fouten zijn waargenomen die te wijten zijn geweest aan fout 4, het conversieproces. Deze fout is enigszins verklaarbaar door het feit dat dit een verbeterd traject heeft betroffen. Hierdoor heeft het conversieproces onder andere bestaan uit het combineren van “oud” en “nieuw”. Blijkbaar is dit een erg grote foutenbron omdat dit in 15 % van de getoetste vlakken niet goed is gegaan. Dit is precies de helft van alle fouten in deel 4. Andersom valt op dat bij deel 1 geen fouten zijn gemaakt in het conversieproces. De oorzaak hiervan is dat er voor dit gedeelte nog weinig geconverteerd is.

- Ook valt in de grafiek het gebrek aan actualisatie bij deel 1 op. In 7.6 % van de getoetste vlakken zijn hierdoor fouten opgetreden. De reden waarom in dit gedeelte deze foutencategorie zo groot is, is vanwege het aanleggen van een buitendijks fietspad in de periode tussen oorspronkelijke inventarisatie en dit onderzoek.

- Binnen deel 1 is ruim 10 % van de getoetste vlakken verkeerd geclassificeerd door gebrek aan een duidelijke definitie. Dit had voornamelijk te maken met het veelvuldig toepassen van de steenbekledingssoort “zetwerk, ratjetoe”. De omschrijving hiervan is niet geheel vastgelegd maar er wordt zetwerk waarin diverse steensoorten zijn gebruikt mee bedoeld. Uit dit onderzoek blijkt echter dat deze omschrijving teveel ruimte laat om een correcte classificatie te kunnen uitvoeren.

4.11.2 Bereikte nauwkeurigheid

In paragraaf 4.5.5 is op basis van de gewenste betrouwbaarheid, gewenste nauwkeurigheid, populatie en proefschatting een steekproefgrootte van 246 vlakken berekend. In deze paragraaf wordt de behaalde nauwkeurigheid van dit onderzoek berekend en worden de consequenties hiervan bepaald.

De uitgangspunten zijn de volgende:

- De gewenste betrouwbaarheid is 95% → z = 1.96
- De behaalde nauwkeurigheid voor invulling van de proefschatting is 0,721. Dit komt overeen met de berekende 72,1 % nauwkeurigheid.
- De steekproefgrootte is 297. Van 3 van de 300 getoetste vlakken bleek het niet mogelijk te zijn deze in het veld te classificeren. Deze 3 zijn dus buiten dit onderzoek gehouden.

De formule voor het berekenen van de behaalde nauwkeurigheid is een afgeleide van de formule die eerder gebruikt is voor het berekenen van de steekproefgrootte. De formule luidt als volgt:

$$\text{behaalde nauwkeurigheid} \quad a^2 = \frac{z^2 \cdot p(1-p)}{n}$$

$$a^2 = \frac{(1{,}96)^2 \cdot 0{,}72\,(1-0{,}72)}{297}$$

$$a = 5{,}1\ \%$$

Deze 5,1 % is lager dan de gewenste 5 % nauwkeurigheid. Omdat dit verschil echter dermate klein is, heeft dit weinig invloed op de uitkomst van dit onderzoek. In plaats van dat de attribuut nauwkeurigheid van het getoetste bestand 72,1 % ± 5 % is, is het nu 72,1 % ± 5,1 %. De nauwkeurigheid van het getoetste bestand ligt dus tussen de 67,0 % en de 77,2 % bij een betrouwbaarheid van 95 %.

4.11.3 Resultaat methode van inwinning

Het onderzoeken van de attribuutnauwkeurigheid van de soort steenbekledingen is met een andere inwinmethode uitgevoerd dan de oorspronkelijke inwinning. De onderbouwing van deze keuze is uitgewerkt in paragraaf 4.6.
Doordat in het analysetraject niet verklaarbare verschillen tussen database en onderzoek nogmaals in het veld zijn ingewonnen, kan er ook een vergelijking gemaakt worden tussen de methodes van inwinning.
In het processchema in bijlage 21 zijn de stappen aangegeven die geleidt hebben tot het bepalen van de nauwkeurigheid van de inventarisatie met de geo-explorer.

Aangezien de controle-inventarisatie alleen uitgevoerd is op de fouten die niet verklaard konden worden (fout nummer 5) kan ook alleen hiervoor de vergelijking gemaakt worden. De andere fouten zijn immers verklaarbaar geweest en hiervan wordt dan ook aangenomen dat deze met de geo-explorer correct ingewonnen zijn. In werkelijkheid hoeft dit echter niet zo te zijn maar is het mogelijk maar niet waarschijnlijk dat in zowel de oorspronkelijke als in de inventarisatie met de geo-explorer dezelfde fouten zijn gemaakt. Wat dus zuiver vergeleken wordt is de nauwkeurigheid van de oorspronkelijke inventarisatie met de inventarisatie met geo-explorer.
Uit de vergelijking van de gecontroleerde vlakken met de inventarisatie met de geo-explorer blijkt in 9 vlakken een fout te zijn gemaakt. Deze fout kan door diverse oorzaken ontstaan zijn. De volgende mogelijkheden kunnen zich hebben voorgedaan:

- Een fout bij het vaststellen van de soort steenbekleding
- Een afleesfout bij het opzoeken van de code behorende bij de soort steenbekleding
- Een communicatiestoornis tussen voorlezen van code en intikken van code
- Een intikfout bij het intikken van de code
- Het beoordelen van het verkeerde glooiingsvlak door het verkeerd inschatten van de positie

Achteraf is het niet mogelijk om voor alle 9 de fouten te bepalen welke fout er gemaakt is. Wel is van een vlak bekend dat er tijdens de onderzoeksinventarisatie van een ander vlak de soort steenbekleding bepaald is. Wat verder nog opvalt is dat in 5 van de 9 vlakken er in werkelijkheid sprake was van een vlak met basalt als bekledingssoort. In 3 van de 9 vlakken werd tijdens de controle als steensoort “ratjetoe” toegekend. Hierdoor wordt aangegeven dat deze vlakken op meerdere manieren geïnterpreteerd kunnen worden.

Belangrijk is nu op welke wijze de methodes van inventarisatie met elkaar vergeleken kunnen worden. Fouten die ontstaan zijn door gebrek aan actualisatie dienen niet meegenomen te worden omdat deze niet zozeer iets met de methode van inwinning te maken hebben maar meer met de methodiek van beheer en bijhouding. Ook de fouten die te maken hebben met gebrek aan definitie dienen niet meegenomen te worden. Ook dit staat immers los van de methode van inwinning.
De andere fouten die voorkomen hebben echter wel te maken met de methode van inwinning. Het betreffen de volgende:

- Fouten in het interne verwerkingsproces
- Fouten in het conversieproces
- Fouten tijdens de inventarisatie

De vergelijking tussen de beide methodes ziet er dan als volgt uit:

methode	oorspronkelijke inventarisatie		inventarisatie mbv geo-explorer	
	absoluut	procentueel	absoluut	procentueel
aantallen fouten	57	19.2	6	2.0

Tabel 10 De vergelijking tussen beide inventarisatiemethodes in procenten en absoluut

Uit dit onderzoek blijkt dus dat de kans op fouten volgens de oorspronkelijke methode bijna 10 maal zo groot is als met de inventarisatie met de geo-explorer. Wat hierbij wel aangetekend dient te worden is dat er door de zeer goede aanwezige materiedeskundigheid in dit onderzoek wellicht minder fouten zijn gemaakt bij het determineren van de soort steenbekleding. Wel is door het gebruik van de geo-explorer de kans op fouten bij vastlegging en verwerking van de soort steenbekleding vele malen kleiner. Daarnaast wordt ook de kans op fouten door het determineren van een verkeerd vlak verminderd door de GPS-positiebepaling van de geo-explorer.

4.12 Vastgestelde kwaliteit versus benodigde kwaliteit

4.12.1 Inleiding

Naast het bepalen van de kwaliteit van een dataset is het ook van belang om deze te vergelijken met de benodigde kwaliteit van betreffende dataset. De benodigde kwaliteit dient

immers “fitness for use” te zijn (Chrisman 1983). Hiermee wordt bedoeld dat onder andere voor wat betreft de kwaliteit, de data aan moet sluiten op gestelde eisen, geredeneerd vanuit de toepassing of toepassingen.

Tijdens de definiëring van de eisen om het steenbekledingenbestand in te gaan winnen is er nagedacht over enkele aspecten van kwaliteit. Over de eisen omtrent de attribuutnauwkeurigheid van de dataset zijn geen expliciete eisen gedefinieerd. Wel is er impliciet over nagedacht, wat vervolgens geleid heeft tot de gebruikte methode van inwinning. Om deze vereiste kwaliteit van de soort steenbekleding alsnog te kunnen bepalen is een workshop georganiseerd met materiedeskundigen en is na deze workshop een vragenlijst opgesteld die door vertegenwoordigers van twee afdelingen ingevuld zijn. Dit betreffen de afdelingen Planvorming en de afdeling Beheer en Onderhoud van de sector Wegen en Waterkeringen. Medewerkers van deze afdelingen zijn de belangrijkste gebruikers van dit bestand.
Een voorbeeld van deze vragenlijst is opgenomen in bijlage 22. In de volgende paragrafen zijn de resultaten van dit kwalitatieve onderzoek welke geleid hebben tot de bepaling van de benodigde kwaliteit op een rij gezet.

4.12.2 Werkprocessen

Om de benodigde kwaliteit te kunnen bepalen is het nodig om in beeld te krijgen voor welke werkprocessen het bestand met steenbekledingen exact gebruikt wordt. Dit betreft de volgende werkprocessen:

- Het uitvoeren van een rationeel glooiingbeheer dat er op gericht is tegen zo gering mogelijke kosten een goede glooiing in stand te houden.
- Het toetsen van de steenbekleding op veiligheid. Dit gebeurt onder andere als voorbereiding op werken in het kader van Project Zeeweringen. Dit project houdt het verbeteren van dijkvakken met te zwakke steenbekledingen in. Daarnaast dient sowieso inzicht te bestaan in de sterkte van de steenglooiingen.
- Het beoordelen van werken van derden. Hiervoor wordt in het systeem op betreffende plaats bekeken hoe de situatie eruit ziet en welke soort steenbekleding er toegepast is.
- Het maken van ontwerpen ten behoeve van het vernieuwen van glooiing. Hiervoor wordt informatie uit het steenbekledingenbestand gebruikt.

4.12.3 Kenmerken die de impact beïnvloeden

Tussen de grootte van de impact bij een misclassificatie en een aantal kenmerken van het betreffende glooiingsvlak bestaat een verband. Hieronder zijn deze verschillende kenmerken op een rij gezet die de impact van een misclassificatie beïnvloeden. Tevens zijn binnen deze kenmerken de grenzen aangegeven waarbij de impact wijzigd. De volgende kenmerken worden onderkend:

- *Oppervlakte van een glooiingsvlak*
  Een onjuist geclassificeerd groot glooiingsvlak heeft meer impact dan een verkeerd geclassificeerd klein vlak. Door de afdelingen is er meer dan één grens aangegeven. De volgende grenzen zijn aangegeven:
  - Oppervlakte van het vlak kleiner dan 300 m²
  - Oppervlakte van het vlak groter dan 300 m² maar kleiner dan 1000 m²
  - Oppervlakte van het vlak groter dan 1000 m²

- *Functie*
  Op de plaatsen waar glooiingsvlakken met steenbekledingen voorkomen kunnen twee functies onderscheiden worden. Afhankelijk van de functie kan de impact van een misclassificatie groter of kleiner zijn. De volgende functies worden onderscheiden:
  - Waterkerende functie
  - Functie als strekdam

- *Reeds verbeterd*
  Of een glooiingsvlak in het kader van de verbeteringswerken wel of niet verbeterd is, beïnvloedt de impact van een misclassificatie. Als een misclassificatie in een nog niet verbeterd glooiingsvlak voorkomt, bestaat er een kans dat deze misclassificatie alsnog achterhaald wordt. Hierdoor kunnen toekomstige extra kosten door fouten tijdens het ontwerpproces en bij het rationeel glooiingbeheer bespaard worden.

- *Geografische ligging*
  Sommige misclassificaties kunnen, vanwege de geografische ligging van de glooiingsvlakken waarop zij betrekking hebben, een grotere impact veroorzaken dan andere misclassificaties. De volgende verschillende geografische liggingen worden onderscheiden:
  - Glooiingsvlakken liggend aan de Westerschelde of Noordzee

- Glooiingsvlakken liggend achter de Oosterscheldekering (dus in de Oosterschelde)

4.12.4 Mogelijke gevolgen bij misclassificatie

Nu de werkprocessen van het steenbekledingenbestand en de kenmerken die van invloed zijn op de impact van een verkeerde classificatie in beeld zijn, kan bepaald worden wat de eventuele gevolgen zouden kunnen zijn bij het verkeerd classificeren van de soort steenbekleding van een glooiingsvlak. De volgende gevolgen zouden zich kunnen voordoen:

- *Het trekken van verkeerde conclusies omtrent de sterkte*
  Alle primaire waterkeringen worden op sterkte en dus op veiligheid getoetst. Bij een misclassificatie wordt met verkeerde informatie gerekend en zal dus ook de uitkomst niet juist zijn. De gevolgen hiervan kunnen zijn dat zwakke steenglooiingen niet in beeld komen en dus ook niet aangepakt worden. De gevolgen hiervan kunnen extra stormschade zijn en als een misclassificatie aan een aantal voorwaarden voldoet dan zou het eventueel mogelijk kunnen zijn dat de misclassificatie uiteindelijk zelfs zou leiden tot een dijkdoorbraak met alle gevolgen vandien. De voorwaarden die dan aan een vlak gesteld worden zijn:
  - De oppervlakte moet groter zijn dan 1000 m².
  - Het vlak moet gelegen zijn aan de Westerschelde of Noordzee.
  - Het vlak moet een waterkerende functie bezitten.
  - De werkelijk aanwezige steenbekleding zou van een steensoort dienen te zijn die onvoldoende sterk is, terwijl deze in de database als een voldoende sterke soort geclassificeerd is.

  Van de getoetste vlakken is onderzocht of hierbij een dergelijk vlak aanwezig is. Dit is niet het geval. Wel is een vlak gevonden die aan de eerste drie voorwaarden voldoet. Aan de vierde voorwaarde wordt niet voldaan omdat het vlak in werkelijkheid een betere kwaliteit bezit dan volgens de database. Aangezien de misclassificatie echter ontstaan is door een fout in het conversieproces, waarbij twee identificatienummers zijn verwisseld, kan niet worden uitgesloten dat in het overige bestand, vlakken die aan alle vier de voorwaarden voldoen bestaan. In een zeer extreem geval zouden deze de aanleiding kunnen vormen tot een dijkdoorbraak.

- *Ten onrechte verbeteren van goede glooiingen*

In het kader van verbeteringswerken is het mogelijk dat, goede glooiingen die als slechte glooiingen geclassificeerd zijn, ten onrechte verbeterd worden. Hierdoor worden er onnodige kosten gemaakt. Momenteel wordt getracht deze vergissingen zoveel mogelijk te voorkomen door gebruik te maken van lokale kennis bij medewerkers en verkenning in het veld voordat de werken aanvangen. De kans dat er dus daadwerkelijk ten onrechte glooiingen verbeterd worden lijkt zeer klein.
Om echter een indicatie te geven van de kosten die hiermee gemoeid zouden kunnen zijn, is van de 297 getoetste vlakken onderzocht hoeveel van deze vlakken er geclassificeerd zijn als vlak met een steensoort van onvoldoende sterkte en die in werkelijkheid wel uit een steensoort van voldoende sterkte bestaat. In totaal bleek dit bij 14 vlakken het geval te zijn. Dit komt overeen met een percentage van 4,7 %. Aangezien de Oosterschelde, het gebied waar nog verbeteringswerken uitgevoerd dienen te worden, een oppervlakte aan zichtbare steenbekledingen bezit van 2.078.500 m², zou dit betekenen dat hiervan 4,7 % ten onrechte verbeterd zou kunnen worden. Deze 4,7 % leidt tot een oppervlakte van 97690 m². Er vanuit gaande dat een m² steenbekleding ca. € 100,-- kost, betekent dit een extra kostenpost van bijna € 10.000.000,-- als er vanuit het systeem zonder overige menselijke tussenkomst en procedures, ontworpen en verbeterd zou worden.

- *Het creëren van wantrouwen in het systeem*
  Op het moment dat er veel fouten in een geografische informatiesysteem aanwezig zijn daalt het vertrouwen hierin en neemt het wantrouwen toe. De gevolgen hiervan kunnen divers zijn:

  - Afname van het gebruik
  - Het niet doorgeven van mutaties
  - Het niet verder ontwikkelen van functionaliteit.

  De gevolgen hiervan kunnen bestaan uit het slechter renderen van de gedane investering tot eventueel mislukken van het project waarbij de volledige investering in zowel tijd, middelen als kennis tenietgedaan wordt. In de enquêtes is aangegeven dat als er in meer dan 20 % van de gevallen onjuiste informatie wordt verstrekt, de ontwikkeling van het GIS in een negatieve spiraal terechtkomt.

- *Verkeerd adviseren inzake het ontwerpproces*
  Een verkeerde inzage in steenbekledingssoorten kan ook het proces van ontwerpen van nieuwe glooiingen frustreren. De gevolgen hiervan kunnen tot extra kosten leiden

vanwege niet geplande werkzaamheden en tot eventueel over moeten doen van ontwerpwerkzaamheden. Daarnaast werkt het ook frustratie bij medewerkers in de hand.

- *Lagere efficiency van rationeel glooiingenbeheer*
  Door misclassificaties kloppen aannames die gebruikt worden voor het nastreven van een optimaal beheer niet. Hierdoor wordt er dus ook geen optimaal beheer uitgevoerd. Het gevolg hiervan zijn extra kosten.

4.12.5 Bepaling kwaliteitseisen

Op basis van het combineren van de informatie die in de bovenstaande paragrafen beschreven is en de informatie uit de enquêtes, is het mogelijk om de kwaliteitseisen betreffende de attribuutnauwkeurigheid van de soort steenbekleding van een glooiingsvlak te bepalen. In de volgende tabel is de vereiste attribuutnauwkeurigheid weergegeven

Oppervlakte in m2	Functie	Verbeterd	Vereiste attribuutnauwkeurigheid in % Oosterschelde	Vereiste attribuutnauwkeurigheid in % Westerschelde
< 300	waterkering	j	82	87
< 300	waterkering	n	80	85
< 300 en > 1000	waterkering	j	95	98
< 300 en > 1000	waterkering	n	90	95
> 1000	waterkering	j	98	99
> 1000	waterkering	n	95	98
< 300	strekdam	j	80	82
< 300	strekdam	n	80	80
< 300 en > 1000	strekdam	j	82	82
< 300 en > 1000	strekdam	n	80	80
> 1000	strekdam	j	85	85
> 1000	strekdam	n	80	80

Tabel 11 De vereiste kwaliteit van de attribuutnauwkeurigheid van de steenbekledingen

Hieruit blijkt dat het niet mogelijk is om één kwaliteitskenmerk te geven voor het bestand met de steenglooiingen. Dit kwaliteitskenmerk is afhankelijk van de hierboven reeds beschreven kenmerken. Als er een vergelijking gemaakt zou worden tussen de vereiste kwaliteit en de aanwezige kwaliteit, dan zou er op bovenstaande kenmerken getoetst moeten worden. Dit betekent dat er een steekproef gekozen zou moeten worden die met bovenstaande kenmerken rekening houdt. Omdat dit inzicht pas uit dit onderzoek naar voren is gekomen, is het bij aanvang niet mogelijk geweest hier rekening mee te houden. Tevens zou de beschikbare tijd niet voldoende zijn geweest omdat er een veel grotere steekproef genomen zou moeten worden om over ieder kenmerk iets te kunnen zeggen.

Om nu toch een vergelijking te kunnen maken tussen de vereiste kwaliteit en de vastgestelde kwaliteit is onderzocht in welke categorie in de tabel met kenmerken een getoetst vlak valt. Op basis van de voorkomende aantallen in een bepaalde categorie is ook het percentage vereiste kwaliteit gehanteerd. Er is dus een gewogen vereiste kwaliteit bepaald op basis van alle kenmerken voor de in de steekpoef getoetste vlakken. In bijlage 23 is het processchema weergegeven van de verschillende stappen in dit proces. In bijlage 24 is de tabel weergegeven waarin het vergelijkingspercentage weergegeven wordt. Dit vergelijkingspercentage bedraagt 84,4 % en kan dus vergeleken worden met de geconstateerde 72,1 %.

4.13 Conclusie

Uit dit kwantitatief onderzoek kan geconcludeerd worden dat er een aanzienlijk verschil bestaat tussen de vereiste attribuutnauwkeurigheid van de steenbekledingen van de waterkeringen en de kwaliteit zoals die momenteel binnen GIS beschikbaar is. De huidige kwaliteit is onvoldoende. De gevolgen hiervan kunnen aanzienlijk zijn. Varierend van afname van het vertrouwen in het systeem tot toename van kosten door minder efficient rationeel glooiingenbeheer en in een zeer extreem geval zou een misclassificatie zelfs kunnen leiden tot een dijkdoorbraak.

Uit dit onderzoek is echter ook inzicht ontstaan in welke fouten gemaakt worden die uiteindelijk de misclassificatie veroorzaken. In totaal zijn er 6 verschillende foutensoorten aangemerkt die ieder op hun eigen wijze wellicht aangepakt kunnen worden om hiermee de nauwkeurigheid te verhogen. Verder is er ook inzicht gekregen in de verschillen per deel. Deze zijn in een aantal opzichten aanzienlijk en wellicht is het hierdoor eenvoudiger verbeteringen aan te brengen. Het aanreiken van verbeteringen voor al deze fouten is buiten dit onderzoek gehouden maar zou aanvullend alsnog verder uitgewerkt kunnen worden. Waar wel een verbetering gesignaleerd is, betreft het inzetten van een andere methodiek van inventarisatie. In plaats van het inventariseren met analoge kaarten, kan er veel verbetering behaald worden uit het inventariseren met behulp van een mobiel GIS met GPS-orientatie

Daarnaast zijn er ook een aantal andere verbeteringen mogelijk. Het betreft de volgende:

- Voor het inventariseren van de soort steenbekleding dienen goed opgeleide en geinstrueerde materiedeskundigen ingezet te worden. Uit het onderzoek is gebleken dat

een aanzienlijk gedeelte van de fouten voortkomt uit het verschillend interpreteren van steenglooiingen.

- Het opstellen van een eenduidige definitie voor het determineren van de verschillende steenbekledingen. Wanneer is een vlak bijvoorbeeld ratjetoe en wanneer geef je het de naam van de meest voorkomende steensoort. Dit soort grenzen kan duidelijker aangegeven worden.
- Bij het naverkennen van de inventarisatie geeft het een beter resultaat als er zonder waarden op de kaart ingewonnen wordt. Juist in de meer onduidelijke situaties bestaat er bij het laten checken van waarden de neiging om het oorspronkelijk ingewonnenen te handhaven. Bij het “blind” inventariseren ontstaat een meer realistisch beeld.

Tot slot is er in dit hoofdstuk een methodiek ontwikkeld die het mogelijk maakt om één aspect van kwaliteit van een dataset vast te stellen en deze vervolgens te vergelijken met de vereiste kwaliteit voor dit aspect, gereredeneerd vanuit de werkprocessen. Dit biedt veel houvast bij een evaluatie van de huidige manier van werken rondom de inwinning en het beheer van geo-informatie.

5 Conclusies en aanbevelingen

5.1 Inleiding

In voorgaande hoofdstukken is onderzoek verricht naar kwaliteit van geo-informatie. Wat betreft dit nu eigenlijk, welke nationale en internationale standaarden bestaan er en hoe kan kwaliteit van geo-informatie gemeten worden. Tevens is onderzocht hoe bij een aantal organisaties in Nederland vorm wordt gegeven aan het borgen van kwaliteit van geo-informatie. Bij een van deze organisaties is ook onderzocht in hoeverre deze borging ook leidt tot de gewenste situatie. Met andere woorden: Komt de kwaliteit overeen met de gestelde wens.

In dit hoofdstuk wordt kort ingegaan op de resultaten uit de diverse deelonderzoeken. Deze zijn uitgebreid behandeld in de hoofdstukken 2 tot en met 4.
Hierna wordt op basis hiervan een “overall” conclusie getrokken en worden een aantal aanbevelingen gedaan.

5.2 Literatuuronderzoek

In het literatuuronderzoek is onderzocht wat kwaliteit en geo-informatie nu eigenlijk zijn en waarom het noodzakelijk is om inzicht in de kwaliteit van een dataset te bezitten. De voornaamste redenen hiervoor zijn om te kunnen bepalen of een dataset geschikt is voor een bepaalde toepassing en om na de inwinning van een dataset te kunnen controleren of deze conform de specificaties is uitgevoerd.

In dit hoofdstuk zijn er een viertal standaarden behandeld die ingaan op kwaliteit van geo-informatie. Opmerkelijk hierbij is dat er in eerste instantie veel overeenkomsten tussen deze standaarden lijken te bestaan. Na een grondige bestudering blijken er toch ook grote verschillen in aanwezig te zijn.
Vooral de Voornorm – geografische informatie – gegevensbeschrijving – kwaliteit is zeer uitgebreid en biedt voldoende houvast om op een juiste wijze kwaliteit van geo-informatie te kunnen borgen. Op basis van de parameters uit deze norm is onderzocht hoe deze weergegeven en getoetst kunnen worden.

5.3 Kwalitatief onderzoek

Mede op basis van de resultaten uit het literatuuronderzoek is in het kwalitatief onderzoek onderzocht hoe een viertal organisaties in Nederland invulling geven aan het aspect kwaliteit van geo-informatie.
De conclusie uit dit hoofdstuk is dat alle vier de organisaties het belang van het in beeld hebben van de kwaliteit van geo-informatie onderkennen. Dit wordt mede onderstreept doordat alle organisaties ten minste vier van de vijf kwaliteitsparameters onderkennen. Uit het onderzoek blijkt tevens dat voor de helft van alle kwaliteitsparameters normen gesteld worden en dat in iets meer dan de helft van deze gevallen ook wordt getoetst of deze normen ook gehaald worden. Dit betekent echter wel dat in meer dan 2/3 van de gevallen, inzicht in de werkelijke kwaliteitsparameters niet voor handen is.
Bij drie van de vier organisaties is er niet één centrale afdeling die toeziet op een goede borging van de kwaliteit van geo-informatie. Opvallend is dat de enige organisatie die wel een centrale afdeling bezit voor het toezien op de borging van de kwaliteitsaspecten, in vier van de vijf gevallen zowel de normen van kwaliteitsparameters stelt en deze ook toetst. Deze organisatie betreft de AGI.
Wat verder geconcludeerd kan worden is dat er door de meeste organisaties veel aandacht besteed wordt aan het opstellen van normen en het toetsen van de positionele nauwkeurigheid. In totaal komt dit in 4 van de 5 gevallen voor.

5.4 Kwantitatief onderzoek

In dit onderzoek is vastgesteld dat er een aanzienlijk verschil bestaat tussen de vereiste en gewenste attribuutnauwkeurigheid van de steenbekledingen van de waterkeringen en de kwaliteit zoals die momenteel binnen GIS beschikbaar is. De aanwezige kwaliteit is beneden de eisen en wensen. De gevolgen hiervan kunnen aanzienlijk zijn. Variërend van afname van het vertrouwen in het systeem tot toename van kosten door minder efficiënt rationeel glooiingenbeheer en in een zeer extreem geval zou een misclassificatie zelfs kunnen leiden tot een dijkdoorbraak.

Uit dit onderzoek is echter ook inzicht ontstaan in welke fouten gemaakt worden die uiteindelijk de misclassificatie veroorzaken. In totaal zijn er 6 verschillende foutensoorten aangemerkt die ieder op hun eigen wijze wellicht aangepakt kunnen worden om hiermee de nauwkeurigheid te verhogen. Van het inzetten van een mobiel GIS met GPS koppeling is

aangetoond dat hiermee voor bepaalde categorieën reeds een reductie van 90 % van de fouten gerealiseerd kan worden.

Tot slot is er in dit hoofdstuk een methodiek ontwikkeld die het mogelijk maakt om één aspect van kwaliteit van een dataset vast te stellen en deze vervolgens te vergelijken met de vereiste kwaliteit voor dit aspect, geredeneerd vanuit de werkprocessen. Dit biedt veel houvast bij een evaluatie van de huidige manier van werken rondom de inwinning en het beheer van geo-informatie.

5.5 "Overall" conclusie

Inzicht in de werkelijke kwaliteit van geo-informatie is zeer belangrijk. Enerzijds vanwege het kunnen bepalen of een dataset geschikt is voor een bepaalde toepassing, anderzijds vanwege het kunnen controleren of een bepaalde inventarisatie conform de eisen is uitgevoerd. Door een juist inzicht in de kwaliteit van een dataset kunnen tevens aanzienlijke kosten bespaard worden en kan een reële inschatting gemaakt worden van eventuele gevolgen van het gebruiken van een bepaalde kwaliteit geo-informatie.

Uit het onderzoek blijkt dat voor wat betreft de maatregelen die genomen worden ten behoeve van de kwaliteitsborging van de geo-informatie, waterschap Zeeuwse Eilanden niet onderdoet voor ten minste 2 van de 3 andere onderzochte organisaties.
Uit het kwantitatief onderzoek bij waterschap Zeeuwse Eilanden blijkt echter dat de attribuutnauwkeurigheid behoorlijk afwijkt van de veronderstelling en de gestelde eisen.
Op basis van deze twee feiten mag er niet vanuit gegaan worden dat deze afwijking tussen gewenste/vereiste en werkelijke kwaliteit niet bij andere ongetoetste parameters van kwaliteit en bij andere organisaties voorkomt. Het veronderstellen van een bepaalde kwaliteit door te veronderstellen het inventarisatieproces te beheersen, biedt geen garantie voor het werkelijk borgen van de kwaliteit. Voor de parameters waarbij de werkelijke kwaliteit niet getoetst wordt, mag er niet aangenomen worden dat deze conform de verwachting zijn. Dit onderzoek staaft het tegenovergestelde.

Op basis van inzicht in de werkelijke kwaliteit van geo-informatie is het door het reduceren van bepaalde fouten mogelijk om de kwaliteit te verhogen. Aangezien 3 van de 4 onderzochte organisaties op geringe schaal de aspecten van kwaliteit van geo-informatie toetsen en er voor diverse aspecten geen normen aanwezig zijn, mag er gevreesd worden

voor de werkelijkheid. Dat dit naast bovenstaande gebrekkig inzicht ook tot grote kosten en zelfs gevaren met zicht mee kan brengen blijkt tevens uit dit onderzoek.

Uit het onderzoek is tevens gebleken dat er een goede standaard aanwezig is om te gebruiken als kapstok voor het ophangen van kwaliteit van geo-informatie. Binnen deze standaard, de Voornorm – geografische informatie – gegevensbeschrijving – kwaliteit, worden alle aspecten van kwaliteit beschreven. Tevens zijn er in het onderzoek methodes aangereikt om de werkelijke kwaliteit te kunnen toetsen.

In de inleiding is de hypothese gegeven waar in dit onderzoek antwoord op gegeven dient te worden. Deze luidt als volgt:

De kwaliteit van geo-informatie wordt op een dusdanige wijze gewaarborgd dat deze aansluit op de nominale grondslag of op de vereiste kwaliteit ten behoeve van het bedrijfsproces.

Gezien de resultaten uit dit onderzoek is de conclusie dat de verwachting dat de kwaliteit van geo-informatie dusdanig gewaarborgd wordt dat deze aansluit op de nominale grondslag of op de vereiste kwaliteit ten behoeve van het bedrijfsproces, bij niet getoetste kwaliteit gering is. Er is geen enkele aanleiding om te veronderstellen dat dit wel het geval is, terwijl er wel diverse aanwijzingen en bevindingen zijn die het tegendeel bewijzen.
In gevallen waarin de kwaliteit getoetst wordt is dit niet het geval. Hier is wel inzicht in de werkelijke kwaliteit.

5.6 Aanbevelingen

Uit dit onderzoek is het mogelijk een aantal aanbevelingen te destilleren. Sommige zijn reeds eerder genoemd. Hieronder zijn ze allemaal weergegeven:

- Ten behoeve van het borgen van de kwaliteit kunnen er een aantal zaken verbeterd worden. Als eerste kan aangesloten worden op de standaard “Voornorm – geografische informatie – gegevensbeschrijving – kwaliteit”. Deze biedt goede mogelijkheden als kapstok voor het ophangen van de kwaliteitsborging van geo-informatie. Daarnaast verdient het de aanbeveling om bij inwintrajecten breder kwaliteitsaspecten en normen hiervan te definiëren en deze ook naderhand te toetsen. Gebleken is dat een bepaald uitgangspunt met betrekking tot kwaliteit en processen die hiertoe zouden moeten leiden in werkelijkheid wel eens anders kunnen zijn.

- Bij het kwantitatief onderzoek bij waterschap Zeeuwse Eilanden is inzicht gekregen in de verschillen in kwaliteit per gebiedsdeel en in de voorkomende soorten fouten. Hoewel het aanreiken van verbeteringen voor al deze fouten buiten dit onderzoek is gehouden, lijkt het erop dat er met relatief weinig inspanning veel verbeteringen gemaakt zouden kunnen worden. In een vervolgonderzoek zou dit en andere mogelijke verbeteringen verder uitgewerkt dienen te worden.

- Door het inzetten van een andere methodiek van inventarisatie kunnen er veel fouten tijdens het inwin- en conversieproces voorkomen worden. In plaats van het inventariseren met analoge kaarten, kan er veel verbetering behaald worden uit het inventariseren met behulp van een mobiel GIS met GPS-orientatie.

- Voor het inventariseren van de soort steenbekleding dienen goed opgeleide en geïnstrueerde materiedeskundigen ingezet te worden. Uit het onderzoek blijkt dat het determineren een lastig proces is dat bij niet voldoende instructie subjectief kan zijn.

- Het opstellen van een eenduidige definitie voor het determineren van de verschillende steenbekledingen. Wanneer is een vlak bijvoorbeeld ratjetoe en wanneer geef je het de naam van de meest voorkomende steensoort. Dit soort grenzen kan duidelijker aangegeven worden.

- Bij het naverkennen van een inventarisatie geeft het een beter resultaat als er zonder waarden op de kaart ingewonnen wordt. Juist in de meer onduidelijke situaties bestaat er bij het laten checken van waarden de neiging om oorspronkelijk ingewonnen informatie te handhaven. Bij het “blind” inventariseren ontstaat een meer realistisch beeld.

Literatuurlijst

- Baarda, D.B., De Goede, M.P.M. en Van der Meer-Middelburg, A.G.E., 1996, basisboek open interviewen, Wolters-Noordhoff B.V., Groningen
- Buijs, A., 1989, Statistiek om mee te werken, Stenfert Kroese, Leiden/Amsterdam
- Burrough, P.A., 1986, *Principles of Geographical Information Systems for Land Resources Assessment*, Clarendon Press, Oxford
- Campbell, H. en Masser, I., 1995, GIS and Organizations, Taylor & Francis ltd, London
- Chrisman, N.R., 1983, *The role of quality Information in the long-term functioning of a GIS*, Proceedings of Autocart06. volume2, Falls Church
- Cohen, J., 1960, *A coefficient of agreement for nominal scales*, Educational and Psychological Measurement 20, 37-46
- Cotteleer, C., Gardebroek, K., Vrolijk, H.C.J. en Dol, W., 2003, Opfriscursus Statistiek, LEI, Den Haag
- Fotheringham, A.S., Bunsdon, C. en Charlton, M., 2002, *Quantitative Geography*, Sage Publications, London
- Genderen, J.L. van en Lock, B.F., 1977, *Testing land-use map accuracy in Photogrammetric Engineering and Remote Sensing 43*
- Goodchild, M.F. en Gopal, S., 1994, *Accuracy of Spatial Databases*, Taylor & Francis ltd, London
- Griffith, D.A., Layne, L.J., 1999, *A casebook for statistical data analysis,* Oxford university Press, New York
- Grimshaw, D.J., 2000, *Bringing Geographical Information Systems into Business,* John Wiley & Sons inc., New York
- Haining, R., 2003, *Spatial Data Analysis*, Cambridge University Press, Cambridge
- Heres, L., 1998, Kwaliteit van geo-informatie, Nederlandse commisie voor geodesie, Delft
- Heywood, I., Cornelius, S. en Carver, S., 1998, *An Introduction to Geographical Information Systems*, Pearson Education Limited, Harlow
- Juran, J.M., Godfrey A. Blanton en Godfrey, A. Blanford, 1998, *Juran's Quality Handbook,* McGraw-Hill Professional, New York
- Landelijk samenwerkingsverband GBKN, 1998, Objectencatalogus GBKN, Landelijk samenwerkingsverband GBKN, Apeldoorn
- Leeuw, A.C.J. de, 2001, Bedrijfskundige methodologie management van onderzoek, Koninklijke Van Gorcum B.V., Assen

- Longley, P.A., Goodchild, M.F., Maguire, J.D. en Rhind, D.W., 1999, *Geographical Information Systems, Principles and technical issues*, John Wiley & Sons inc., New York
- Longley, P.A., Goodchild, M.F., Maguire, J.D. en Rhind, D.W., 2001, *Geographic Information Systems and Science*, John Wiley & Sons ltd, Chichester
- Meetkundige Dienst, 2002, Productspecificaties uitbesteding fotogrammetrie, intern document, Meetkundige Dienst, Delft
- Nederlands Normalisatie Instituut, 1998, Voornorm Geografische Informatie Kwaliteit, Nederlands Normalisatie Instituut, Delft
- Polman, J., en Salzmann, M.A., 1996, Handleiding voor de technische werkzaamheden van het Kadaster (HTW), Apeldoorn.
- Ravi, referentiemodel vastgoed, Stichting Ravi overlegorgaan voor vastgoedinformatie, Amersfoort
- Ravi, 1994, Kwaliteit in perspectief, overzicht van de behoefte aan regisatratie van kwaliteitskenmerken over geo-informatie in Nederland, Stichting Ravi overlegorgaan voor vastgoedinformatie, Amersfoort
- Ravi, 1995, Europese standaardisatie met betrekking tot geo-informatie, Ravi participatie in de Europese normalisatie commissie CEN/TC287, Stichting Ravi overlegorgaan voor vastgoedinformatie, Amersfoort
- Reeve, D. en Petch, J., 1999, *GIS Organisations and People*, Taylor & Francis ltd, London
- Salge, F., Guptill, S.C. en Morrisson, J.L., 1995, *Elements of spatial data quality* , Elsevier Science, Oxford
- Verschuren, P.J.M., 2002, de probleemstelling voor een onderzoek, Uitgeverij Het Spectrum, Utrecht
- Zhang, J. en Goodchild, M., 2002, *Uncertainty in Geographical Information*, Taylor & Francis, New York

Lijst met geraadpleegde internetsites

- Adviesdienst Geo-Informatie en ICT, http://www.minvenw.nl/rws/agi/home, laatst geraadpleegd op 29 januari 2004
- Burger, R., bibliography about Data Quality and Quality Management in GIS,http://www.clearinghouse.com.uy/~carlos/bibliography.htm, laatst geraadpleegd op 28 oktober 2003
- Cassini International Workshop, Data Quality in Geographic Information,http://www-cassini.univ-mrs.fr/DataQuality97.html, laatst geraadpleegd op 23 oktober 2003
- DIGEST, http://www.digest.org, laatst geraadpleegd op 9 februari 2004
- European Committee for Standardization, http://www.cenorm.be/cenorm/index.htm, laatst geraadpleegd op 27 oktober 2003
- Federal Geographic Data Committe, http://www.fgdc.gov, laatst geraadpleegd op 18 december 2003
- Kadaster, http://www.kadaster.nl, laatst geraadpleegd op 25 januari 2004
- Landelijk Samenwerkingsverband Grootschalige Basiskaart Nederland, http://www.gbkn.nl, laatst geraadpleegd op 29 januari 2004
- National Center for Geographical Information & Analysis, http://www.ncgia.ucsb.edu, laatst geraadpleegd op 4 januari 2004
- Nederlands Normalisatie Instituut, http://www2.nen.nl, laatst geraadpleegd op 26 december 2003
- Norheim, R.A., *Distinguishing Positional and Attribute Error in Two old Growth Forest Mapping Projects*, http://staff.washington.edu/norheim/oldgrowth/distinguishing.html, laatst geraadpleegd op 28 december 2003
- RAVI, http://www.ravi.nl, laatst geraadpleegd op 9 februari 2004
- The American Cartographer, Proposed Standard for Digital Cartographic Data Quality, http://faculty.washington.edu/chrisman/G460/DataQual.html, laatst geraadpleegd op 8 februari 2004
- USGS, What is SDTS, http://mcmcweb.er.usgs.gov/sdts/whatsdts.html, laatst geraadpleegd op 3 februari 2004
- Waterschap Zeeuwse Eilanden, http://www.wze.nl, laatst geraadpleegd op 30 januari 2004

Bijlage 1 Interview bij Topografische Dienst Kadaster

Inleiding

Dit betreft een interview dat gehouden is met Gerard Vrijkotte, senior beleidsadviseur bij de Topografische Dienst Kadaster op 22 juli 2003 te Emmen.

- *Voorstellen*

  Gerard Vrijkotte is werkzaam als senior beleidsadviseur bij de afdeling Onderzoek en Ontwikkeling (O&O) bij de Topografische dienst kadaster (TDK) te Emmen. Deze afdeling is verantwoordelijk voor de lange termijn projecten gericht op het verbeteren van producten en productieprocessen. Daarnaast ondersteunt O&O de beleidsbepaling betreffende alle relevante bedrijfsprocessen.

- *Uitleg over doel en probleemstelling en schetsen kader*

  Uitleg over de doel- en probleemstelling conform het onderzoeksplan. Dit is ter voorbereiding aan het interview verstrekt.

- *Wijze van interview en tijdsplanning*

  Er is een structuur opgesteld om de gewenste informatie in beeld te krijgen. Deze structuur is verwoord in een aantal vragen. In een tijdsbestek van circa 2 uur wordt getracht deze vragen te beantwoorden. Er is ruimte om eventueel dieper op bepaalde zaken door te gaan.

Interview

- *Welke soorten geo-informatie worden er ingewonnen en beheerd. Betreft dit zowel geografische informatie als beschrijvende informatie over deze geografische gegevens;*

  Er worden diverse kaartproducten bij TDK gemaakt. De belangrijkste series zijn de topografische kaarten op de schalen 1:10.000, 1 : 25.000, 1 : 50.000 en 1 : 250.000. Daarnaast worden verschillende speciale kaarten gemaakt voor militaire en civiele gebruikers.

De core business is de vervaardiging van TOP10vector dit is de topografische kaart op een schaal 1 : 10.000. Deze wordt eens in de 4 jaar herzien. De andere topografische bestanden zijn afgeleid van TOP10vector. De top25Raster is een verkleining, de Top50Vector is gegeneraliseerd van TOP10vector.
Alle topografische bestanden zijn geattributeerd met soorten objecten en overige kenmerken.

- *Hoe ziet deze inwinning er organisatorisch globaal uit, wie is er opdrachtgever, wie leidt de projecten;*

Eens in de 4 jaar wordt TOP10vector herzien. Dit betekent dat ieder jaar een kwart van Nederland gefotografeerd wordt. De opdrachtgever hiervoor is de Bevelhebber Landstrijdkrachten hiervoor vertegenwoordigt door de Dienst Geografie Koninklijke Landmacht (DGKL). Met het DGKL wordt ieder jaar een convenant afgesloten betreffende wat er dat jaar geproduceerd gaat worden.

De cyclus van 4 jaar is gebaseerd op wat TDK organisatorisch aankan en wat vanuit de historie zo gegroeid is. Bij sommige klanten bestaat de behoefte dat vaker de producten van TDK herzien worden. Het verkorten van de periode van herziening levert wel een forse verhoging van kosten met zich mee. Door het efficienter maken van het proces wordt er op termijn wel naar toegewerkt om de herzieningscyclus te verhogen naar 2 jaar.

Binnen TDK is de afdeling productie verantwoordelijk voor het productieproces en leidt ook de inwinprojecten.

- *Welke afdeling is verantwoordelijk voor het borgen van de kwaliteit en wie ziet daarop toe;*

Er is niet een afdeling die verantwoordelijk is voor het borgen van de kwaliteit. Wel is er een afdeling die de kaartproductie van TDK realiseert. Hierdoor is het makkelijker om kwaliteitsproblemen organisatorisch op te lossen. De afdeling productie is namelijk verantwoordelijk voor de kwaliteit van het product en ziet dus ook toe dat dit op een juiste manier gebeurt.

Het borgen van de kwaliteit bij TDK omvat vele zaken. Volgens TDK begint dit bij het instrueren en opleiden van medewerkers. Voor de 35 topografen die werkzaam bij TDK

betreft dit het leren toepassen van de verkenningsregels. Daarnaast dienen deze regels ook eenduidig toegepast te worden. Het te maken bestand dient ook tussen topografen onderling consistent te zijn. Om dit zo goed mogelijk te realiseren, heeft TDK er voor gekozen om de ene medewerker te laten digitaliseren en een andere medewerker dit na te laten kijken. Vooral na het invoeren van deze maatregel leidde dit tot de nodige problemen. Inmiddels werkt het zogenaamde collegiaal nakijken naar behoren.

- *Welke diverse aspecten van kwaliteit van geo-informatie worden er onderkend;*

  Bij TDK worden de kwaliteitsaspecten logische consistentie, volledigheid, temporele nauwkeurigheid, positionele nauwkeurigheid en attribuut nauwkeurigheid onderscheiden.

- *Worden al deze soorten kwaliteit bij nieuwe inwintrajecten ook gedefinieerd en wordt hierop ook getoetst. Zoja, op welke wijze vindt deze toetsing plaats van kwaliteit van geo-informatie.Staat dat ergens beschreven;*

  Nieuwe inwintrajecten wijken niet af van eerdere inwintrajecten. De eisen die gedefinieerd zijn, zijn dus voor alle trajecten gelijk. In onderstaande opsomming staat voor alle door TDK onderkende kwaliteitsaspecten welke normen hieraan gesteld worden, of hier eventueel een toetsing van plaatsvind en of deze twee zaken ook ergens beschreven staan.

  <u>Logische consistentie</u>
  Normen met betrekking tot dit aspect worden niet gesteld. Door middel van opleiding en begeleiding wordt wel getracht een bestand zo consistent mogelijk te maken. Verder staan in het handboek verkenningsvoorschriften de regels die voor de verkenning in acht dienen te worden genomen. Te denken valt bijvoorbeeld aan de classificatie, wanneer iets een bos is en wanneer het losse bomen betreft of wanneer het een weg is met een bepaalde klasse en wanneer het een andere klasse betreft.
  Verder worden er 3 afdrukken per bestand gemaakt om een bepaald object of objectsoort zoals bijvoorbeeld een categorie weg beter te onderzoeken op consistentie van codering of volledigheid van digitalisatie. Is de gehele weg wel gedigitaliseerd en heeft deze weg over haar loop dezelfde codering. Op deze wijze wordt het gehele bestand consequent gecontroleerd.

  <u>Volledigheid</u>

Normen met betrekking tot dit aspect worden niet gesteld. Wel wordt de volledigheid visueel en beperkt softwarematig gecontroleerd. Deze visuele controles gebeuren op verschillende manieren. Er wordt onder andere gecontroleerd op basis van de luchtfoto's of alles gedigitaliseerd is. Verder worden er 3 afdrukken per bestand gemaakt om een bepaald object of objectsoorten zoals bijvoorbeeld een categorie weg beter te onderzoeken op volledigheid van digitalisatie. Is de gehele weg wel gedigitaliseerd en heeft deze weg over haar loop dezelfde codering. Dit is een erg moeilijk onderdeel omdat er ook bepaald dient te worden wanneer iets wel en wanneer iets niet meegenomen dient te worden. Als bijvoorbeeld bij de bouw van een huis alleen de fundering er nog ligt, wordt dit dan meegenomen als gebouw of nog niet. Hiervoor zijn wel richtlijnen opgesteld maar er komen altijd grensgevallen voor.

Temporele nauwkeurigheid

Normen met betrekking tot dit aspect worden niet gesteld. Wel wordt er getracht om een bestand zo snel mogelijk na de inwinning op de markt te brengen.
Om de doorlooptijd te verkorten worden sinds enige tijd de luchtfoto's ook gescand afgeleverd aan TDK. Hierdoor kunnen de topografen met een laptop veel sneller het veld in om nieuwe topografie in te tekenen. De doorlooptijd van het gehele proces van kaartvervaardiging wordt hierdoor sterk verkort. Vroeger duurde het herzien van de kaarten jaren. Dit is nu teruggebracht tot maanden. Op het moment dat de producten nu op de markt worden gebracht zijn ze dus ook minder verouderd en zijn dus kwalitatief beter.

Vroeger was het proces als volgt:

Anologe luchtfoto -> aanpassen bestaande en intekenen nieuwe topografie -> scannen -> in veld onzekere en niet van luchtfoto zichtbare zaken controleren en eventueel aanpassen -> digitaliseren

Nu is het proces:
Digitale luchtfoto ->voor zover mogelijk binnen digitaliseren van nieuwe topografie en aanpassen van bestaande topografie -> in het veld onzekere en niet van luchtfoto zichtbare zaken controleren en eventueel digitaal aanpassen

In 2003 is het proces voor de eerste keer veranderd. In de winter 2003/2004 zullen kentallen gegenereerd worden om te zien hoeveel efficienter het proces nu is geworden.

Positionele nauwkeurigheid

Normen met betrekking tot dit aspect in het eindproduct worden nog niet gesteld. Wel wordt een bepaalde kwaliteit van de luchtfoto's voorgeschreven. Deze kwaliteit wordt achteraf ook gecontroleerd.

Het werk van de topograaf wordt in eerste instantie door de topograaf zelf en later door het collegiaal nakijken visueel gecontroleerd.

Samen met het Kadaster is TDK momenteel bezig om op basis van de GBKN als referentiebestand de positionele nauwkeurigheid van TOP10vector te onderzoeken. Uit dit onderzoek zal blijken hoe goed of hoe slecht de positionele nauwkeurigheid van TOP10vector is ten opzichte van de GBKN. In eerste instantie was het de bedoeling om de vergelijking te maken met een bestand wat op basis van fotogrammetrie gemaakt zou worden. Hetzelfde object zou door meerder operatoren gekarteerd worden waardoor een zogenaamd nulbestand zou ontstaan. De TOP10vector zou dan met dit bestand worden vergeleken. Vanwege capaciteitsproblemen is gekozen om de vergelijking met de GBKN te maken. Op basis van de resultaten uit dit onderzoek zullen normen betreffende de positionele nauwkeurigheid opgesteld worden.

Naast bovenstaande wordt er ook door middel van softwarematige controles gecontroleerd op slivers en gaps. Dit betreffen naast topologische fouten ook positionele onnauwkeurigheden.

Attribuut nauwkeurigheid

Zoals hiervoor reeds is beschreven wordt er voor wat betreft positionele nauwkeurigheid visueel collegiaal nagekeken. Dit betreft voornamelijk visuele controles op structurele basis, dus niet steeksproefgewijs. Hierbij wordt ook op een grove manier gecontroleerd op juistheid van classificatie. Dit nakijken op een luchtfoto is een zeer beperkte controle van classificatie omdat het merendeel van de classificaties niet van de luchtfoto onderscheiden kan worden. Normen met betrekking tot de attribuut nauwkeurigheid worden niet gesteld.

- *Wat is de verwachting hoe er in de toekomst met kwaliteitsaspecten van geo-informatie omgegaan zal worden;*

De inwintechnieken worden steeds beter waardoor de positionele nauwkeurigheid verbeterd kan worden. Dit geeft kwaliteitsconflicten tussen bestaande en nieuwe topografie. Mede hierdoor is het aspect meta-data erg van belang. Na verloop van tijd kan het erg belangrijk worden wanneer een object is gedigitaliseerd en op basis van welke luchtfoto.

- *Overige zaken….*

Onder andere in opdracht van TDK is Pepijn van Oort aan de Universiteit van Wageningen bezig met een onderzoek naar het vertalen van aspecten van kwaliteit naar een model. Met andere woorden het bepalen van een instrumentarium waarmee de kwaliteit van een dataset bepaald kan worden.

Op het nemen van de luchtfoto's na wordt het gehele proces van het produceren van de diverse bestanden in eigen beheer uitgevoerd..

TDK is ook druk bezig met het realiseren van meta-data bij hun bestanden. Hiervoor hebben ze een korte-termijn-oplossing die bestaat uit een meta-data omgeving die bij ESRI gebouwd wordt. Voor de lange-termijn-oplossing wordt gedacht aan een oplossing conform een CEN of ISO normering. Het belang van het adequaat bijhouden van meta-data wordt als zeer belangrijk onderkend.

Door het gebruiken van gescande luchtfoto's met hoge resolutie is de mogelijkheid ontstaan om meer detail te onderscheiden dan vroeger. Hierdoor kan topografie nauwkeuriger ingetekend worden maar kunnen ook afwijkingen in bestaande topografie gesignaleerd worden. Als dit nu allemaal aangepast zou worden zou het realiseren van een nieuwe herziening te veel tijd in beslag nemen en zouden klanten die infrastructuur vastgelegd hebben aan TDK-objecten dit allemaal nog een keer dienen te doen. Daarom zijn er regels opgesteld wanneer iets wat afwijkt wel aangepast dient te worden en wanneer niet.

Voorlopig worden er nog geen normeringen opgesteld voor de overige elementen van kwaliteit. Uit het onderzoek met het Kadaster wordt dus alleen een positionele nauwkeurigheid vastgesteld. Op basis hiervan zal een norm voor de positionele nauwkeurigheid voor toekomstige inwinningen opgesteld worden.

Afronding

- *Verdere stappen in het onderzoek*

Toelichting op andere stappen conform het onderzoeksplan

- *Tijdsplanning onderzoek*

Toelichting op tijdsplanning conform onderzoeksplan

- *Resultaten*

Na afronden onderzoek zal een exemplaar van de scriptie toegestuurd worden.

Bijlage 2 Interview bij het Kadaster

Inleiding

Dit betreft een interview dat gehouden is met Auke Hoekstra, senior beleidsmedewerker bij het Kadaster op 2 juli 2003 te Apeldoorn.

- *Voorstellen*

Auke Hoekstra is senior beleidsmedewerker bij de afdeling Beleid en Marketing van de Concernstaf Vastgoedinformatie en geodesie van het Kadaster te Apeldoorn. De werkzaamheden betreffen de zorg voor beleidsontwikkeling en waarvan met name geografische aspecten, strategie voor ontwikkeling nieuw producten en vernieuwingsstrategie voor redesign van kern registraties (kadastrale kaart, GBKN in relatie met kadastrale registratie)

- *Uitleg over doel en probleemstelling en schetsen kader*

  Uitleg over de doel- en probleemstelling conform het onderzoeksplan. Dit is ter voorbereiding aan het interview verstrekt.

- *Wijze van interview en tijdsplanning*

  Er is een structuur opgesteld om de gewenste informatie in beeld te krijgen. Deze structuur is verwoord in een aantal vragen. In een tijdsbestek van circa 2 uur wordt getracht deze vragen te beantwoorden. Er is ruimte om eventueel dieper op bepaalde zaken door te gaan.

Interview

- *Welke soorten geo-informatie worden er ingewonnen en beheerd. Betreft dit zowel geografische informatie als beschrijvende informatie over deze geografische gegevens;*

  Bij het Kadaster worden eigenlijk twee soorten geo-informatie ingewonnen en beheerd. De eerste soort is de Kadastrale kaart met beschrijvende informatie. Dit is een eigen product.

De tweede soort is de Grootschalige Basiskaart van Nederland (GBKN). Dit is een grootschalige topografische kaart waarvoor het Kadaster in opdracht van regionale stichtingen de beheer en bijhouding voor haar rekening neemt. Ook in de GBKN is beschrijvende informatie aanwezig.

- *Hoe ziet deze inwinning er organisatorisch globaal uit, wie is er opdrachtgever, wie leidt de projecten;*

  Binnen het Kadaster zijn er regionaal projectteams die de inwinningen en verwerking van mutaties voor de Kadastrale Kaart voor haar rekening nemen. De inwinning vindt plaats nadat bekend is geworden dat de kadastale perceelssituatie is veranderd. Het Kadaster heeft de wettelijke taak om de eigendomssituatie van percelen bij te houden en te beheren.
  De inwinning en verwerking van de GBKN-mutaties wordt grotendeels door middel van uitbesteding uitgevoerd. De begeleiding hiervan wordt door projectcoördinatoren van het Kadaster uitgevoerd. De regionale stichtingen, de eigenaren van de GBKN, zijn opdrachtgever voor de beheer en bijhouding. De verwerkte mutaties worden weer geleverd aan de leden van de betreffende stichting.

- *Welke afdeling is verantwoordelijk voor het borgen van de kwaliteit en wie ziet daar op toe;*

  Er is niet één centrale afdeling die verantwoordelijk is voor het borgen van de kwaliteit of die hierop toeziet. Wel is het zo dat de benodigde kwaliteit van het Kadaster door middel van voorschriften en bestekken vastgelegd is. Door middel van controles en steekproeven wordt regelmatig getoetst of de gewenste kwaliteit ook gewaarborgd en gerealiseerd wordt. Jaarlijks wordt door de kwaliteitscoördinatoren (1 in elke regionale vestiging en de Concernstaven) een plan opgesteld met daarin de voorgenomen acties en onderzoeken. Input voor deze plannen kan ook en vooral worden geleverd door uitvoerders en het management. De centraal vastgestelde jaarplannen kunnen vervolgens worden aangepakt waarbij intern opgeleide auditoren veelal de onderzoeken uitvoeren in een beperkt aantal dagen. Deze auditoren hebben daarnaast gewoon hun reguliere werk.

- *Welke diverse aspecten van kwaliteit van geo-informatie worden er onderkend;*

De kwaliteitsaspecten die het Kadaster onderscheid voor de kadastrale producten zijn juistheid, consistentie, volledigheid en actualiteit. Van ieder product van het Kadaster zijn deze vereiste kwaliteits aspecten in principe bepaald.
Naast deze Kadastrale producten is het Kadaster ook de beheerder van de GBKN. De kwaliteitsaspecten die voor de GBKN onderscheiden worden staan vermeld in de objectencatalogus GBKN en zijn actualiteit, identificeerbaarheid, wijze van inwinning, toegankelijkheid, volledigheid, consistentie en precisie.

- *Worden al deze soorten kwaliteit bij nieuwe inwintrajecten ook gedefinieerd en wordt hierop ook getoetst. Zoja, op welke wijze vindt deze toetsing plaats van kwaliteit van geo-informatie.Staat dat ergens beschreven;*

Door het continue bijhoudingsproces liggen de kwaliteitseisen vast in voorschriften en bestekken en gelden voor ieder inwintraject. Door het veranderen van methodes en inzichten kan het wel zo zijn dat die eisen aangepast worden aan de op dat moment geldende inzichten.
Het Handboek Technische Werkzaamheden van het Kadaster (HTW) is een van de handleidingen waarin de kwaliteitsspecificaties voor inwintrajecten staan beschreven. Verder zijn er diverse kwaliteitshandboeken waarin ook specificaties en richtlijnen staan aangegeven. Het belang van de kwaliteit van de gegevens wordt door het Kadaster onderkend. Deze bepalen immers de kwaliteit van de producten en diensten van het Kadaster.
Er worden tijdens het inwinproces controles uitgevoerd. Hierdoor wordt er een toets op de kwaliteit van inwinningen uitgevoerd. Er wordt gewerkt aan een plan om de uitkomsten van deze toetsen te registreren in een landelijk overzicht. De kwaliteit van de geometrie is volgens het Kadaster goed gewaarborgd. Door middel van steekproeven is de positionele nauwkeurigheid recent nog uitgebreid getoetst. De andere kwaliteitseisen worden doorgaans visueel getoetst en incidenteel ook in het veld.

Naast bovenstaande is voor het verkrijgen van inzicht in de kwaliteit van kadastrale gegevens een project opgestart voor metadata. Daarin worden de volgende fasen onderscheiden:

- Beoordelen en vaststellen indeling kadastrale gegevens en kwaliteitscriteria
- Vaststellen huidig en ontbrekend inzicht in gegevenskwaliteit
- Specificeren van vervolgacties

- *Wat is de verwachting hoe er in de toekomst met kwaliteitsaspecten van geo-informatie omgegaan zal worden;*

  Er wordt gewerkt aan de uitbreiding van de data met meta-informatie (zie hiervoor). Het voornaamste doel hiervan is om richting de gebruikers aan te geven wat de data voorstelt, hoe actueel deze is, hoe nauwkeurig deze is en voor welke toepassingen de data geschikt is. Het niveau waarop het Kadaster de meta-data wil realiseren is op objectsoortniveau. Per object heeft geen zin omdat het gelijksoortige objecten betreft met dezelfde meta-data.

- *Overige zaken….*

  In de Grootschalige Basis Kaart van Nederlandf (GBKN) en Digitale Kadastrale Kaart (DKK) wordt nu ook al informatie over de verschillende objecten meegeleverd. Onderdelen van deze informatie betreft onder andere precisie, idealisatie en nauwkeurigheid. Deze informatie is echter voornamelijk gericht op het eigen beheer en bijhoudingsproces. De meta-data zoals die onder de toekomstverwachting behandeld wordt is voornamelijk gericht op gebruiker van de data en dus niet op de beheerder.

  Het belang van kwaliteitsaspecten wordt onderkend bij het Kadaster en de verwachting is dat deze aspecten alleen maar belangrijker zullen gaan worden doordat er steeds meer mensen gaan werken met GIS die minder inzicht hebben in de processen en data in het systeem.

  Bij het definieren van de interne kwaliteitseisen valt het op dat het Kadaster vaak het criterium hanteert van 100 % gegevenskwaliteit. Ondanks dat dit niet gegarandeerd haalbaar is heeft het Kadaster dit toch gedefinieerd. De volgende argumenten liggen hieraan ten grondslag:
  - Het schept duidelijkheid richting de medewerkers wat er van hun wordt verwacht;
  - Vanuit haar wettelijke opgedragen taak wordt ervanuit gegaan dat deze juist uitgevoerd wordt en dus een 100 % gegevenskwaliteit bezit.
    Deze kwaliteitseis wordt ieder jaar door een realistische doelstelling ondersteund. Dit zijn percentages die liggen in de buurt van de huidige kwaliteit.

In het algemeen is het wel zo dat het toetsen van data erg veel geld kost. In de mate van toetsen wordt een afweging gemaakt tussen nut, noodzaak en wensen van de klant.

Afronding

- *Verdere stappen in het onderzoek*

Toelichting op andere stappen conform het onderzoeksplan

- *Tijdsplanning onderzoek*

Toelichting op tijdsplanning conform onderzoeksplan

- *Resultaten*

Na afronden onderzoek zal een exemplaar van de scriptie toegestuurd worden.

Bijlage 3 Interview bij de Adviesdienst Geo-informatie en ICT

Inleiding

Dit betreft een interview dat gehouden is met Adri den Boer, senior adviseur geo-informatie bij de Adviesdienst Geo-informatie en ICT op 2 juli 2003 te Delft.

- *Voorstellen*

  Adri den Boer is senior adviseur geo-informatie bij de afdeling Advies en Onderzoek GIS bij de Adviesdienst Geo-informatie en ICT (AGI) te Delft. De missie van deze afdeling is 'zorgen dat Verkeer & Waterstaat optimaal profiteert van geografische informatie zodat zij haar missie zo goed mogelijk kan vervullen'. Vanuit deze functie participeert Adri den Boer inzake standaardisatie voor RWS bijvoorbeeld in de NNI-normcommissie GI (351.240), de stichting Ravi en het NCGI. Ook is de AGI medefinancier en hij van daaruit begeleider van een AIO bij de UW (Kwaliteit van geo-informatie).

- *Uitleg over doel en probleemstelling en schetsen kader*

  Uitleg over de doel- en probleemstelling conform het onderzoeksplan. Dit is ter voorbereiding aan het interview verstrekt.

- *Wijze van interview en tijdsplanning*

  Er is een structuur opgesteld om de gewenste informatie in beeld te krijgen. Deze structuur is verwoord in een aantal vragen. In een tijdsbestek van circa 2 uur wordt getracht deze vragen te beantwoorden. Er is ruimte om eventueel dieper op bepaalde zaken door te gaan.

Interview

- *Welke soorten geo-informatie worden er ingewonnen en beheerd. Betreft dit zowel geografische informatie als beschrijvende informatie over deze geografische gegevens;*

  De AGI kent de volgende geo-informatie die zij inwinnen en beheren:

- Digitale Topografische Bestanden (DTB)
- Het Actueel Hoogtebestand Nederland (AHN)
- TOPhoogteMD
- Het waterstaatkundig Informatiesysteem
- Grondslagpunten
- Normaal Amsterdams Peil (NAP)
- Informatie Systeem Waterschappen
- Dijkringenbestand
- Vegetatiekaarten

Van alle bestanden betreft het zowel geometrie als (globale) beschrijvende informatie van deze geometrie. Omdat voor de AGI de nadruk van de werkzaamheden ligt bij het inwinnen en leveren ten behoeve van de DTB's wordt er in de rest van het onderzoek beperkt tot deze DTB's.

- *Hoe ziet deze inwinning er organisatorisch globaal uit, wie is er opdrachtgever, wie leidt de projecten;*

De hoofdafdeling topografische geo-informatie besteed inwinwerkzaamheden uit of voert het in eigen beheer uit. Het hoofdkantoor van Rijkswaterstaat is opdrachtgever voor de inwinning en beheerswerkzaamheden.
In principe wordt het bijwerken van de DTB's iedere 5 jaar uitgevoerd. Als de regionale directies de mutatiecyclus willen verkorten dienen zij dit zelf te bekostigen. In de natte infrastructuur zijn momenteel dermate grote financiele problemen dat de kans aanwezig is dat de mutatiecyclus van 5 jaar verlengd zal worden naar bijvoorbeeld 7 jaar. Het veranderen van de muteercyclus gebeurt op basis van budgetgrootte en is derhalve niet afgestemd op het werkproces. Inzicht in de muteercyslus van objecten is niet aanwezig

- *Welke afdeling is verantwoordelijk voor het borgen van de kwaliteit en wie ziet daar op toe;*

De hoofdafdeling topografische geo-informatie ziet toe op de kwaliteitsborging en besteed inwinwerk uit of voert het zelf uit.
In het algemeen worden er bij interne inwinningen weinig controles uitgevoerd. Bij uitbestedingen van de inwinningen of verwerkingen vinden deze wel plaats. Er is geen centrale afdeling die op deze controles toeziet.

- *Welke diverse aspecten van kwaliteit van geo-informatie worden er onderkend;*

  Bij de AGI worden alle aspecten van kwaliteit van geo-informatie onderkend. Door bezuinigen wordt er in de inwintrajecten voornamelijk gekeken naar positionele nauwkeurigheid. Aan topologische consistentie, atttribuut nauwkeurigheid en compleetheid wordt ad-hoc invulling gegeven terwijl aan de andere aspecten van kwaliteit geen tijd wordt besteed. Wel is het zo dat er signalen vanuit de regionale directies richting de AGI komen dat er bepaalde kwaliteitsaspecten ontbreken en dat hierin graag inzicht verkregen zou worden.

- *Worden al deze soorten kwaliteit bij nieuwe inwintrajecten ook gedefinieerd en wordt hierop ook getoetst. Zoja, op welke wijze vindt deze toetsing plaats van kwaliteit van geo-informatie.Staat dat ergens beschreven;*

  Doordat er voornamelijk gewerkt wordt aan de DTB's liggen de eisen voor wat betreft de positionele nauwkeurigheid vast in de specificatie van de DTB's. Door middel van vereffeningen wordt hierop gecontroleerd en ook worden deze vereffeningen steeksproefgewijs optisch gecontroleerd. Dit staat beschreven in de productspecificaties uitbesteding fotogrammetrie.

  Inwinningen die uitbesteed worden, worden beter gecontroleerd dan bij interne inwinning. Er wordt echter niet gesuggereerd dat het resultaat dan ook beter is. Argumenten om dit te staven zijn niet voorhanden. Logisch lijkt het echter wel. De AGI heeft routine in deze inwintrajecten terwijl dat voor veel bedrijven vaak incidenteler is. Aan de andere kant kunnen fouten eenvoudig gemaakt worden juist omdat het routinematige werkzaamheden betreft.

  De eisen voor de opbouw van de topologie liggen vast in het datamodel DTB. De topologie wordt bij de AGI zelf altijd opgebouwd in het Infocam-systeem van Leica, wat de basis checkt v.w.b. 'spaghetti' (check tijdens procesgang). Tevens zijn er controles op het vlakkennet op maaiveldniveau: geen overlappingen of ontbreken van vlakken.

  Voor wat betreft de eisen en controles op attribuutnauwkeurigheid ligt er niets vast. Onregelmatig vinden er steeksproefgewijs optische controles plaats.
  De compleetheid van een bestand wordt visueel getoetst met de foto als ondergrond. Dit levert wel alleen een controle op op basis van compleetheid van objecten, niet van

attributen. Parameters van de kwaliteit van compleetheid zijn niet bekend. Er is dus niet bekend hoeveel procent van een bepaald object in een bepaald gebied naar verwachting ontbreekt.

- *Wat is de verwachting hoe er in de toekomst met kwaliteitsaspecten van geo-informatie omgegaan zal worden;*

  In verband met bezuinigingen is hier niet echt veel over te zeggen. Logischerwijs zou er ook naar andere aspecten van kwaliteit onderzoek verricht moeten worden. Gezien de financiele perikelen lijkt dit op korte termijn niet tot de realiteit te gaan behoren.

- *Overige zaken….*

  In het kader van beheer en bijhouding van geo-informatie dient er inzicht te worden verkregen in de muteerbaarheid van verschillende objecten. Hoe vaak dient er ingewonnen en gemuteerd te worden. Eventueel kan dit worden afgestemd op de wijze van inwinnen. Er kan bijvoorbeeld gekozen worden voor een objectgerichte of een gebiedsgerichte wijze van inwinning. Momenteel is bij de AGI dit inzicht in muteerbaarheid niet voor handen.

  De automatisering en digitalisering heeft niet tot gevolg gehad dat de taakverdeling tussen de AGI + regionale directies is gewijzigd. De AGI levert het DTB als halfproduct aan de regionale directies. In KernGIS, een door Rijkswaterstaat ontwikkelde GIS – applicatie, wordt de DTB door de regionale directies verder gedetailleerd , opgebouwd en aangevuld met beschrijvende informatie. De grens tussen wat de AGI doet en wat de regionale directies doen, ligt op het gebied van kennis van de materie. Wat zonder specifieke materiedeskundigheid onderscheiden kan worden, wordt door de AGI ingewonnen. Voor objecten waarvoor wel materiedeskundigheid nodig is wordt de detaillering uitgevoerd door de regionale directies. Het verwerken van nieuwe DTB's in deze omgevingen is nog niet geregeld. Dit is een behoorlijk complexe klus.

  De AGI bestaat sinds 1931. De reden hiervan is de opkomst van luchtfoto's waardoor er een hierop gespecialiseerde dienst nodig was. De inwinning van de AGI is voor 95 % gebaseerd op fotogrammetrie en voor 5 % op basis van andere inwinmethodes zoals terrestrische metingen, sonarlodingen etc..

Afronding

- *Verdere stappen in het onderzoek*
  Toelichting op andere stappen conform het onderzoeksplan

- *Tijdsplanning onderzoek*
  Toelichting op tijdsplanning conform onderzoeksplan

- *Resultaten*
  Na afronden onderzoek zal een exemplaar van de scriptie toegestuurd worden.

Bijlage 4 Interview bij Waterschap Zeeuwse Eilanden

Inleiding

Dit betreft een uitwerking van een interview dat gehouden is met Johan van cranenburgh en Willem Rijn op 15 december 2003 te Goes.

- *Voorstellen*

  Johan van Cranenburgh is coordinator van de cluster Geo-informatie binnen de afdeling Automatisering en Geo-informatie. Deze afdeling maakt onderdeel uit van de sector Bedrijfsvoering. Willem Rijn is projectleider Geo-informatie en werkzaam binnen dezelfde cluster. Deze cluster verzorgd het inwinnen van diverse geo-informatie en thematische kaartvervaardiging. Daarnaast wordt het Waterschaps GIS beheerd en worden toepassingen van Geo-informatie georganiseerd en geimplementeerd.

- *Uitleg over doel en probleemstelling en schetsen kader*

  Uitleg over de doel- en probleemstelling conform het onderzoeksplan. Dit is ter voorbereiding aan het interview verstrekt.

- *Wijze van interview en tijdsplanning*

  Er is een structuur opgesteld om de gewenste informatie in beeld te krijgen. Deze structuur is verwoord in een aantal vragen. In een tijdsbestek van circa 2 uur wordt getracht deze vragen te beantwoorden. Er is ruimte om eventueel dieper op bepaalde zaken door te gaan.

Interview

- *Welke soorten geo-informatie worden er ingewonnen en beheerd. Betreft dit zowel geografische informatie als beschrijvende informatie over deze geografische gegevens;*

  Bij waterschap Zeeuwse Eilanden kan de de geo-informatie opgedeeld worden in twee hoofdsoorten. Het betreffen de volgende:

  <u>Basis- of referentiebestanden</u>.

Dit zijn bestanden die van derden worden betrokken en die binnen het waterschap voor diverse toepassingen gebruikt worden. Deze bestanden worden dus alleen beheerd door het waterschap en niet ingewonnen. De volgende kunnen onderscheiden worden:

- Grootschalige Basiskaart van Nederland (geleverd door Stichting GBKN)
- Kadastrale kaart (geleverd door het Kadaster)
- Top10Vector, Top50Vector, Top25 Raster (geleverd door de Topgrafische Dienst Nederland)
- Bodemkaart (geleverd door Alterra)
- Nationaal Wegenbestand (geleverd door Rijkswaterstaat)
- Actueel Hoogtebestand van Nederland (geleverd door de Adviesdienst Geo-informatie en ICT)

<u>Bestanden die de infrastructuur van het waterschap beschrijven</u>

Deze bestanden worden zowel ingewonnen als beheerd en bijgehouden door het waterschap zelf en kunnen in de volgende subcategorien opgedeeld worden:

- Bestanden die betrekking hebben op de waterkeringen. Het gaat te ver om alle objecten die hierbinnen onderscheiden kunnen worden te beschrijven. Wel kan de volgende opdeling gehanteerd worden:
    - Objecten die betrekking hebben op de constructieopbouw
    - Objecten die betrekking hebben op de vooroever
    - Objecten die betrekking hebben op referentiestelsels
    - Objecten die betrekking hebben op bekledingen

  Bijzondere kenmerken van bovenstaande bestanden is dat een aantal drie-dimensonaal beschikbaar zijn. Dit betreffen de categorien vooroever en bekledingen. Van alle categorien is zowel beschrijvende als geometrische informatie beschikbaar. De geometrische weergave is zowel in vlakken, lijnen als punten.

- Bestanden die betrekking hebben op het beheer van de wegen. Dit betreft een kleinere groep geo-informatie. Het betreft hier alleen de hartlijnen van en het groen langs de wegen. Deze bestaan beide uit zowel geometrische informatie als beschrijvende informatie en is in twee dimensies weergegeven. De geometrische weergave van de weg is in lijnen. Groen is geometrisch weergegeven als lijnen en punten.

- Bestanden die betrekking hebben op het beheer van het oppervlaktewater. Ook hier gaat het te ver om alle objecten die hierbinnen onderscheiden kunnen worden, te beschrijven. Wel kan de volgende groepsindeling gemaakt worden:
    - Objecten die betrekking hebben op het oppervlaktewatersysteem. Deze kunnen vervolgens weer onderverdeeld worden in meetpunten, het waterlopenstelsel, peil- en afvoergebieden en kunstwerken.
    - Objecten die betrekking hebben op de watersysteemketen. Dit betreffen de riool(pers)leidingen, de zuiveringen en de kunstwerken.

  Bijzondere kenmerken van bovenstaande categorien is dat deze allemaal in twee dimensies gerepresenteerd zijn en dat het zowel geometrische informatie als beschrijvende informatie hierover betreft. De geometrische weergave is zowel vlakken, lijnen als punten.

- *Hoe ziet deze inwinning er organisatorisch globaal uit, wie is er opdrachtgever, wie leidt de projecten;*

  De cluster Geo-informatie is een centrale cluster binnen het waterschap die dienstverlenend optreedt met betrekking tot de inwinning van geo-informatie en het verzorgen van oplossingen op het gebied van geo-informatievoorziening.
  Voor de inwinning van geo-informatie kan iedereen binnen en eventueel buiten het waterschap opdrachtgever zijn. In principe leiden medewerkers van de de cluster “Geo-informatie” de inwinning van nieuwe geo-informatie en het ontwikkelen van systemen en functionaliteit. In de praktijk komt het er echter op neer dat waar de nadruk ligt van een project, ook de leiding van dat project komt te liggen. Regelmatig komt het voor dat door andere redenen zoals bijvoorbeeld beschikbare capaciteit de leiding over projecten wordt bepaald.
  Bij grootschalige geo-data projecten zijn naast de cluster geo-informatie, voor de beschrijvende informatie, de functionele afdelingen betrokken en verantwoordelijk.

- *Welke afdeling is verantwoordelijk voor het borgen van de kwaliteit en wie ziet daar op toe;*

  Er is niet één centrale afdeling die verantwoordelijk is voor het borgen van kwaliteit. Iedere opdrachtgever is zelf verantwoordelijk voor het borgen van de kwaliteit. Vaak wordt deze rol binnen de diverse aandachtvelden ingevuld door applicatiebeheerders.

Geo-informatie en ook niet–geo-informatie die binnen de applicatie horen, zijn de verantwoording van de applicatiebeheerders.
Verder is het zo dat als er bij een inwintraject aan de cluster Geo-informatie een bepaalde kwaliteit geeist wordt, het de verantwoording van deze cluster is om hieraan te voldoen. Zij dient zelf hiervoor eventueel maatregelen te treffen. Wel wordt vaak maar een beperkt aantal kwaliteitseisen gevraagd. Een goede onderbouwing van deze gevraagde kwaliteitseisen ontbreekt vaak.

- *Welke diverse aspecten van kwaliteit van geo-informatie worden er onderkend;*

De volgende aspecten worden onderkend:
- Positionele nauwkeurigheid
- Attribuut nauwkeurigheid
- Volledigheid
- Actualiteit
- Semantische nauwkeurigheid

- *Worden al deze soorten kwaliteit bij nieuwe inwintrajecten ook gedefinieerd en wordt hierop ook getoetst. Zoja, op welke wijze vindt deze toetsing plaats van kwaliteit van geo-informatie. Staat dat ergens beschreven;*

Van de onderkende kwaliteitsaspecten wordt bij nieuwe inwintrajecten alleen de positionele nauwkeurigheid gedefinieerd. Het proces wordt getoetst door regelmatig tijdens een grootschalige inwinning, infrastructuur dubbel te meten en te onderzoeken of dit overeenkomt. Het definieren van de grootte van de toets wordt op basis van gevoel gedaan. Statistische onderbouwing om later gefundeerd een kwaliteitskenmerk aan een dataset te kunnen hangen ontbreekt. Wel is het zo dat met betrekking tot de positionele nauwkeurigheid er vereffeningen worden uitgevoerd. Deze moeten de toets van gestelde eisen en voorwaarden doorstaan.
Naast bovenstaande controlemetingen worden er ook log-bestanden van vereffeningen van landmetingen onderzocht. Deze geven ook een indicatie van de positionele nauwkeurigheid van de meting. Niet alle fouten worden hiermee gevonden. Een voorbeeld van een fout die niet gevonden zal worden is als tijdens de meting de spiegel op een verkeerd punt geplaats wordt. De meting zelf is goed, alleen is het verkeerde gemeten.

Tijdens het toetsen in het veld worden ook attributen meegenomen zoals bijvoorbeeld soort kunstwerk, diameter van duikers en soort steenbekleding van een waterkering. De kwaliteitsparameters attribuut nauwkeurigheid, volledigheid en actualiteit worden niet expliciet gedefinieerd. Impliciet wordt er door opdrachtgevers vanuit gegaan dat deze foutloos zijn. Naast attribuutnauwkeurigheid wordt er ook bij grote inwinprojecten gecontroleerd op de volledigheid. Dit gebeurt door voor- en/of naverkenning die vervolgens volledig of steeksproefgewijs uitgevoerd kan worden.

Bij de verwerking binnen wordt er bij grootschalige inwinprojecten ook gecontroleerd op het voorkomen van onlogische waarden. Te denken valt bijvoorbeeld aan duikerdiameters boven een bepaalde waarde of gemeten waterstanden in een dwarsprofiel die niet gelijk zijn en waarbij het water dus niet horizontaal zou liggen. Daarnaast vinden er ook visuele controles plaats waarbij gelet wordt op codering, consistentie en bijvoorbeeld het voorkomen van afwijkende waarden.

Omdat het inwinprojecten zijn en de inwinning gedurende het project loopt, bestaat er inzicht in de actualiteit van gegevens. Voor het actueel houden van ingewonnen gegevens wordt momenteel een uitwerking gemaakt die onder andere gericht is op de toepassing waarvoor deze gegevens gebruikt worden. Om dit goed te kunnen te kunnen doen dienen ook de vereiste kwaliteitsaspecten in beeld te worden gebracht, geredeneerd vanuit het bedrijfsproces.

Het definieren van de vereiste positionele nauwkeurigheid wordt bij grootschalige inwinprojecten beschreven in het bestek. De controletoetsing wordt niet apart beschreven. De opdrachtnemer noet door middel van berekeningen aantonen dat aan de gestelde kwaliteitseisen met betrekking tot positionele nauwkeurigheid is voldaan.

- *Wat is de verwachting hoe er in de toekomst met kwaliteitsaspecten van geo-informatie omgegaan zal worden;*

Het belang van het implementeren van kwaliteitsborging wordt onderkend. In eerste instantie zal beschreven gaan worden welke aspecten van kwaliteit geborgd dienen te worden. Dit onderzoek zal hiertoe de nodige input leveren. Belangrijk hierbij is dat de vereiste kwaliteit geredeneerd vanuit de werkprocessen gedefinieerd dient te worden. In een dialoog met de opdrachtgevers zal bewustwording voor wat betreft kwaliteitsaspecten nagestreefd worden en zal deze voor toekomstige inwinprojecten gedefinieerd dienen te worden.

Naast bovenstaande zal ook aan momenteel reeds beschikbare geo-informatie een kwaliteitsstempel toegekend dienen te worden. Hoe dit precies zal gebeuren is nu nog niet bekend.

- *Overige zaken….*

  Naast applicatiespecifieke informatie, wordt er ook informatie onderscheiden die meervoudig door verschillende applicaties gebruikt wordt. Binnen het waterschap is deze informatie aangeduid als “basisregistratie” informatie. De verantwoording voor het beheren en bijhouden maar ook het borgen van de kwaliteit van deze informatie is neergelegd bij de afdeling Automatisering & Geo-informatie. Op termijn zal op basis van onderzoek hieraan invulling worden gegeven mede om te kunnen bepalen voor welke toepassing deze informatie geschikt is. Moeilijkheid hierbij is dat er onderzocht dient te worden wat de benodigde kwaliteit voor de basisregistratie dient te worden terwijl deze voor meerdere appliacties en dus verschillende processen gebruikt gaan worden. Vanwege efficiencymotieven dient er wellicht niet uitgegaan te worden van uitzonderingssituaties. Te denken valt bijvoorbeeld aan een proces dat een zeer hoge nauwkeurigheid vereist maar waarvan de informatie maar eens in de 4 jaar gebruikt wordt. Voor een ander proces hoeft bijvoorbeeld de nauwkeurigheid niet hoog te zijn, maar dient het wel redelijk actueel te zijn. Als deze beide gegevens in de basisregistratie ondergebracht zouden worden, zou dat kunnen leiden tot het beheren en bijhouden van een zeer nauwkeurig gegeven met een hoge mate van actualiteit. De kosten hiervan zouden veel hoger kunnen zijn dan het separaat beheren en bijhouden. Belangrijk is om te bepalen wanneer een proces een uitzondering is en wanneer niet.

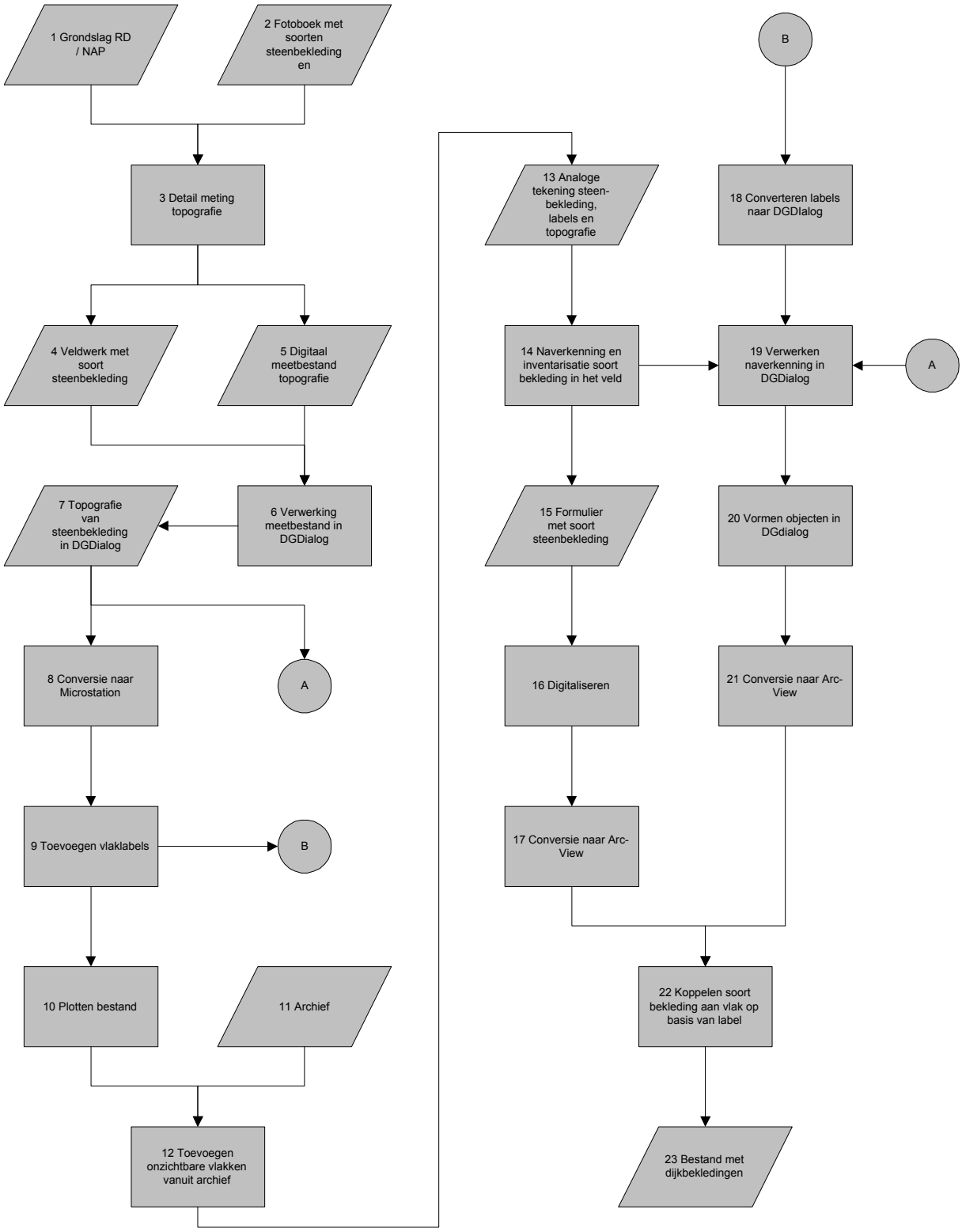
Afronding

- *Verdere stappen in het onderzoek*
  Toelichting op andere stappen conform het onderzoeksplan

- *Tijdsplanning onderzoek*
  Toelichting op tijdsplanning conform onderzoeksplan

- *Resultaten*
  Na afronden onderzoek zal een exemplaar van de scriptie toegestuurd worden.

Bijlage 5 Totstandkoming bestand dijkbekledingen



1 Grondslag RD / NAP

Om veldmetingen in een later stadium te kunnen inpassen of vergelijken met andere bestanden dienen metingen te worden uitgevoerd ten opzichte van referentiestelsels. Door het doen van metingen op basis van referentiestelsels kunnen metingen ingepast worden in reeds bestaande topografie of bijvoorbeeld binnen GIS vergeleken te worden met andere objecten of datasets.
De referentiestelsels die voor de inwincampagne zijn gebruikt is het Rijksdriehoekstelsel (RD) voor de horizontale positionering en het Normaal Amsterdams Peil (NAP) voor de verticale positionering. Om op deze referentiestelsels te kunnen aansluiten zijn langs de waterkeringen diverse afgeleide meetpunten gemaakt; het meetnet is verdicht.

1 Fotoboek met soorten steenbekledingen

De veldmetingen zijn uitgevoerd door landmeetkundigen zonder specifieke deskundigheid van waterkeringen. Om het toch mogelijk te maken om de diverse bekledingssoorten van elkaar te onderscheiden is een fotoboek gemaakt. In dit boek zijn foto's van alle voorkomende bekledingssoorten opgenomen.

2 Detailmeting topografie

Van de waterkering is de topografie door middel van detailmetingen ingemeten. Dit betekent dat alle zichtbare objecten, met inachtneming van modellering, op de waterkering zijn ingemeten. In bijlage 6 is een overzicht opgenomen met hierin alle objectsoorten die onderscheiden worden in het landmeetproces.

3 Veldwerk met soort steenbekleding

Bij het landmeetproces hoort ook het maken van een veldwerk waarop analoog de situatie geschetst wordt. Op dit veldwerk zijn van de steenbekledingsvlakken de soorten aangegeven. Dit betekent dat alle vlakken in het veld individueel benoemd zijn.

4 Digitaal meetbestand topografie

Naast het veldwerk is een resultaat van de veldmeting een digitaal bestand met hierin coderingen en waarnemingen (richting en afstand) van de gemeten punten.

5 Verwerking meetbestand in DGDialog

Voordat er iets met het digitaal meetbestand (5) gedaan kan worden, dient het eerst verwerkt te worden in een landmeetprogramma. In dit geval is dat gebeurd binnen DGDialog. Binnen dit programma zijn controles op de landmetingen uitgevoerd, xyz-coordinaten van de gemeten coördinaten berekend en afwijkingen binnen gestelde normen vereffend.
Wanneer nodig is het veldwerk gebruikt om onduidelijkheden op te lossen.

6 Topografie van steenbekleding in DGDialog

Het resultaat van actie 6 is een lijnen en punten bestand binnen DGDialog betreffende de topografie van de waterkeringen.

7 Conversie naar Microstation

Omdat Microstation breder ingezet wordt binnen Waterschap Zeeuwse Eilanden, is de topografie van de waterkeringen geconverteerd van DGDialog naar Microstation.

8 Toevoegen vlaklabels

Om in een latere actie (22) informatie aan de verschillende vlakken met steenbekledingen te kunnen toevoegen is een uniek label aan het vlak toegekend.

9 Plotten bestand

Ten behoeve van het naverkennen van de topografie en het inventariseren van informatie over de soort steenbekleding in het veld is de topografie van de dijk inclusief de toegekende labels geplot.

10 Archief

In het archief is veel materiaal aanwezig waaruit de opbouw van de waterkeringen is te achterhalen. Dit betreft voornamelijk analoge ontwerp of revisietekeningen die net voor of na de bouw van de waterkering zijn gemaakt.

11 Toevoegen onzichtbare vlakken vanuit archief

Bij de inmetingen zijn alleen zichtbare objecten ingemeten. In veel gevallen loopt de waterkering door onder de zogenaamde bestortingslijn. Dit is de grenslijn tussen de waterkering en de vooroever. Om toch in beeld te krijgen tot hoe ver de waterkering onder de zandlijn doorloopt, is getracht deze informatie uit het archief op te diepen. Vervolgens is deze informatie analoog toegevoegd aan de plots van de dijk.

12 Analoge tekening steenbekleding, labels en topografie

Het resultaat uit de vorige actie is een analoge tekening met de steenbekleding. Ter oriëntatie is ook overige topografie op de plot geplaatst.

13 Naverkenning en inventarisatie soort bekleding in het veld

Met de analoge tekening als leidraad is in het veld door materiedeskundigen verkend of deze ook inderdaad klopt. Er is verkend op volledigheid en juistheid. Tegelijkertijd hiermee is bekeken welke soort steenbekleding het betreft.

14 Formulier met soort steenbekleding

De inventarisatie van de soort steenbekleding is op een apart formulier verwerkt.

15 Digitaliseren

Na ontvangst op kantoor is het inwinformulier handmatig in een spreadsheet verwerkt.

16 Conversie naar Arc-View

Vervolgens is de spreadsheet als tabel geconverteerd naar Arc-View.

17 Conversie naar DGDialog

De labels zijn geconverteerd van Microstation naar DGDialog.

18 Verwerken naverkenning in DGDialog

De uit de naverkenning komende opmerkingen worden in de topografie van de steenbekledingen in DGDialog verwerkt. Dit resulteert in het toevoegen aan of aanpassen van de begrenzingen van de steenbekledingen. Tevens zijn de op de plots aangegeven onzichtbare vlakken geconstrueerd in het bestand.

19 Vormen objecten in DGDialog

Op basis van de begrenzingen van de steenbekledingen worden er vlakken gecreëerd. Vooral omdat dit 3 dimensies betreffen, is dit een lastige klus.
De labels worden als teksten bij de vlakken opgenomen.

20 Conversie naar Arc-View

Het in actie 20 gemaakte bestand wordt geconverteerd van DGDialog naar Arc-View. Tevens worden de geplaatste labelteksten als attribuut aan de objecten gekoppeld.

21 Koppelen soort bekleding aan vlak op basis van label

Binnen Arc-View wordt de informatie betreffende soort steenbekleding gekoppeld aan het steenbekledingsvlak. Dit gebeurt op basis van het label dat zowel aan de beschrijvende als aan de geometrische informatie is toegevoegd.

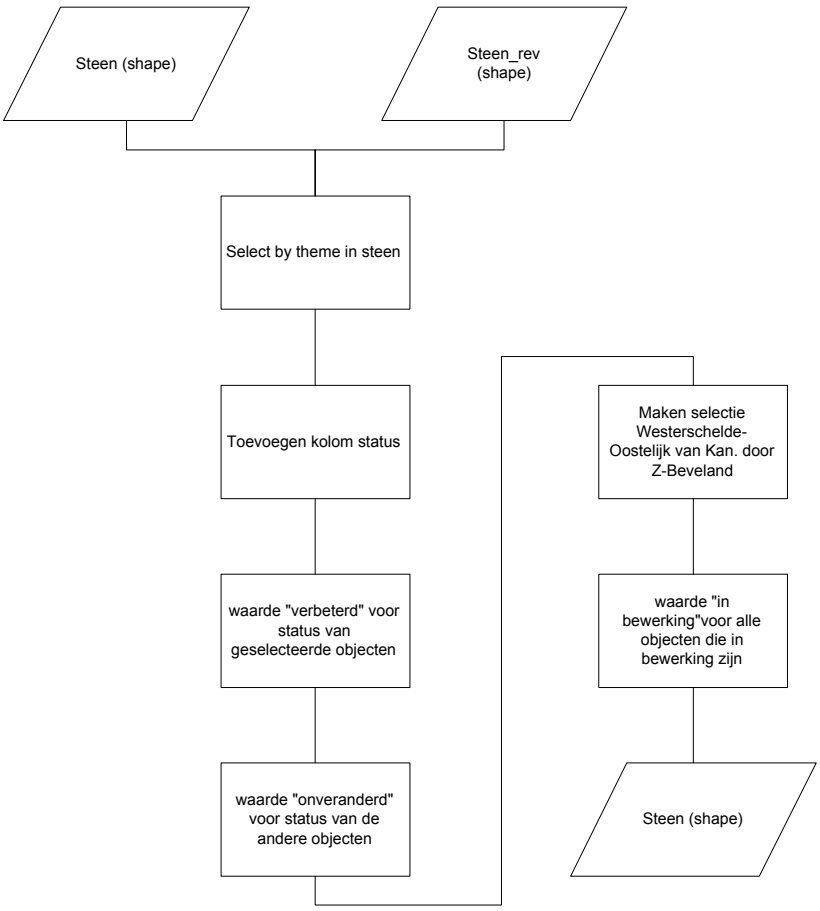
22 Bestand met steenbekledingen

Bijlage 6 Overzicht van gemeten objectcoderingen

Nr topo	Omschrijving	Nr profiel
100	bebouwing	
	bebouwing links	481
	bebouwing rechts	482
112	bijgebouw	
	bijgebouw links	483
	bijgebouw rechts	484
115	fundatie	
116	bunker	
120	kunstwerk	
300	grens	*474*
351	hoogtepunt	
355	hoogtelijn	
370	leiding	
380	elec.kabel	
381	telecom. kabel	
382	overige kabels	
383	waterleiding	
384	gasleiding	
385	riolering	
386	persleiding	
400	grondslagpunt	
403	meetlijn	*470*
409	overige grondslag	
410	hoofdraaipaal	*471*
415	dijkpaal/raaipaal	*476*
420	meetpunt	
450	ligging dwarsprofiel	
455	sliblaag	
456	ligging dwarsprofiel uitmeting	
458	sliblaag uitmeting	
460	ligging lengteprofiel	
601	spoorrail	
602	as spoor	431
603	as weg	432
605	bomenrij/groep	464
619	overig lijnv. object	
620	straatmeubilair	
621	lantaarnpaal/lichtmast	
622	paal/steen	
623	kolk	
624	putdeksel	
626	boom	*466*
628	rioolput (insp.put)	
629	verkeerslicht	
630	verkeersbord	

Nr topo	Omschrijving	Nr profiel
631	verklikker transportleiding	
639	overig punt object	
650	brug	
655	coupure	
660	duiker	440
662	peilschaal	
675	stuw	443
685	damwand	463
690	beschoeiing	498
691	rietbegrenzing	*469*
700	terreinafscheiding	
702	raster	*465*
703	hek	462
704	haag	*473*
705	schutting	441
706	houtwal	
707	midden sloot/greppel	433
709	insteek sloot	
710	slootbodem	
750	kant verharding	*467*
	kant gesloten verharding links	485
	kant gesloten verharding rechts	486
	kant open verharding links	487
	kant open verharding rechts	488
	kant sierbestrating links	495
	kant sierbestrating rechts	496
753	kant tegel	
	kant tegel links	489
	kant tegel rechts	490
755	kant half verhard	
	kant halfverhard links	491
	kant halfverhard rechts	492
756	kant onverhard	
	kant onverhard links	493
	kant onverhard rechts	494
770	kant water	497
800	kruinlijn	
801	teenlijn	442
802	tussenberm	
807	stortsteenlijn	
808	zandlijn	
810	scheidingslijn taludbekleding	*468*
811	perkoenrij	*472*
812	trap	
813	betonband	*477*
814	teenschot	*480*
815	paalhoofd	
816	perkoenrij met betonband	

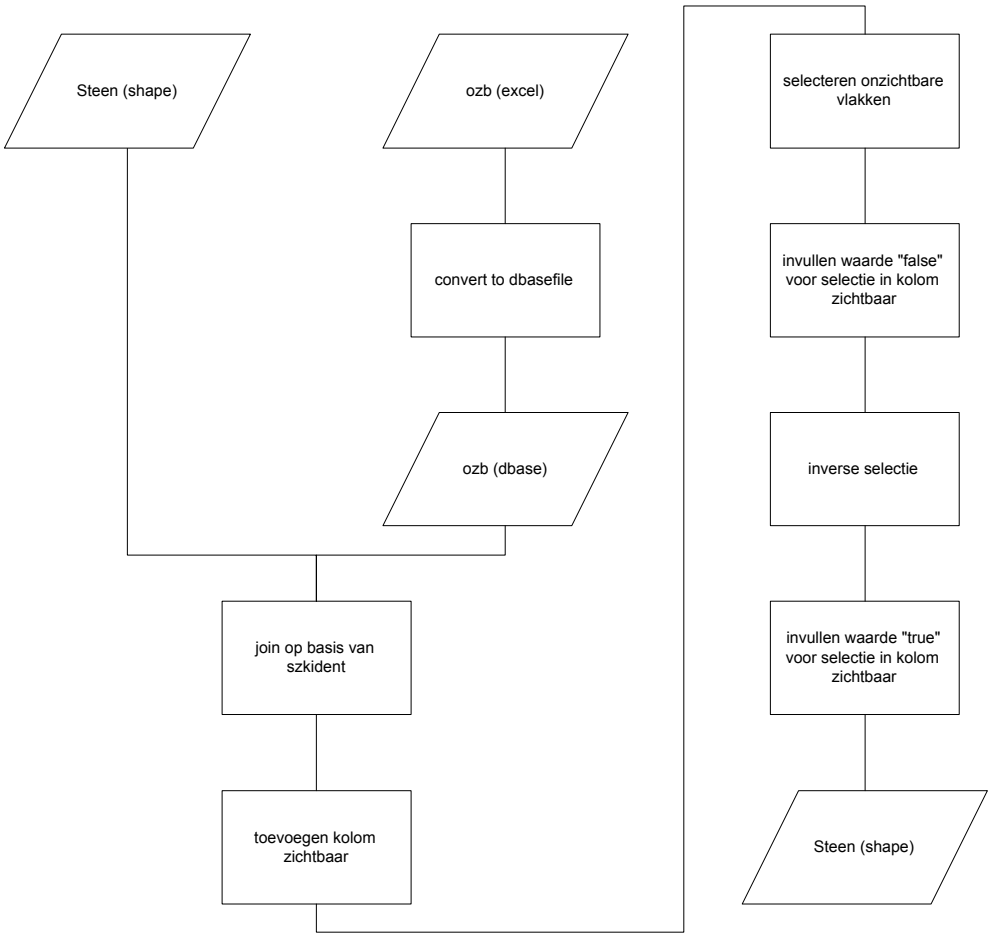
Bijlage 7 Processchema toevoegen status



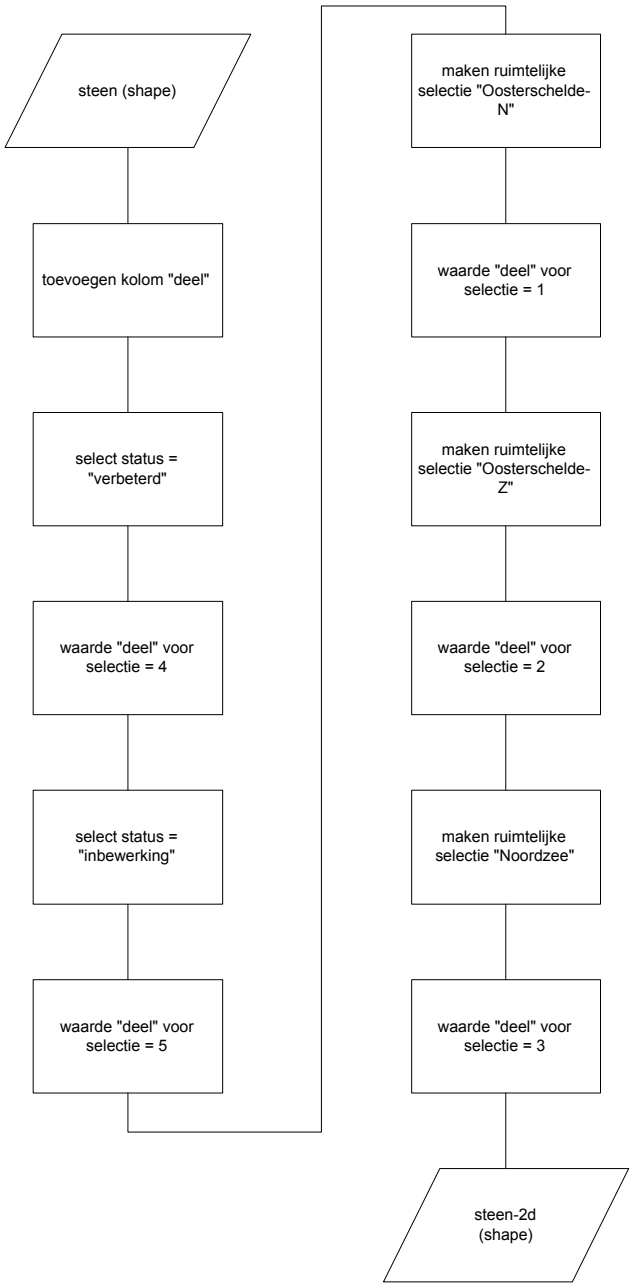
Op basis van de shapefile Steen.shp en de shapefile Steen_rev.shp is aan de vlakken die in de eerste categorie vallen, de status “onveranderd” toegekend. Aan de tweede categorie is de status “verbeterd” toegekend. Aan de derde categorie is de status “in bewerking” toegevoegd.

Het selecteren van de categorie “verbeterd” is gedaan met behulp van een bestand waarin alleen verbeterde vlakken aanwezig waren (Steen_rev.shp). Dit is een automatische actie geweest met behulp van standaard Arc-View functionaliteit.
Het selecteren van de categorie “in bewerking” is gedaan op basis van een ruimtelijke selectie. Dit betreft alle vlakken langs de Westerschelde, oostelijk van het kanaal door Zuid-Beveland. Deze ruimtelijke selectie is handmatig gemaakt.

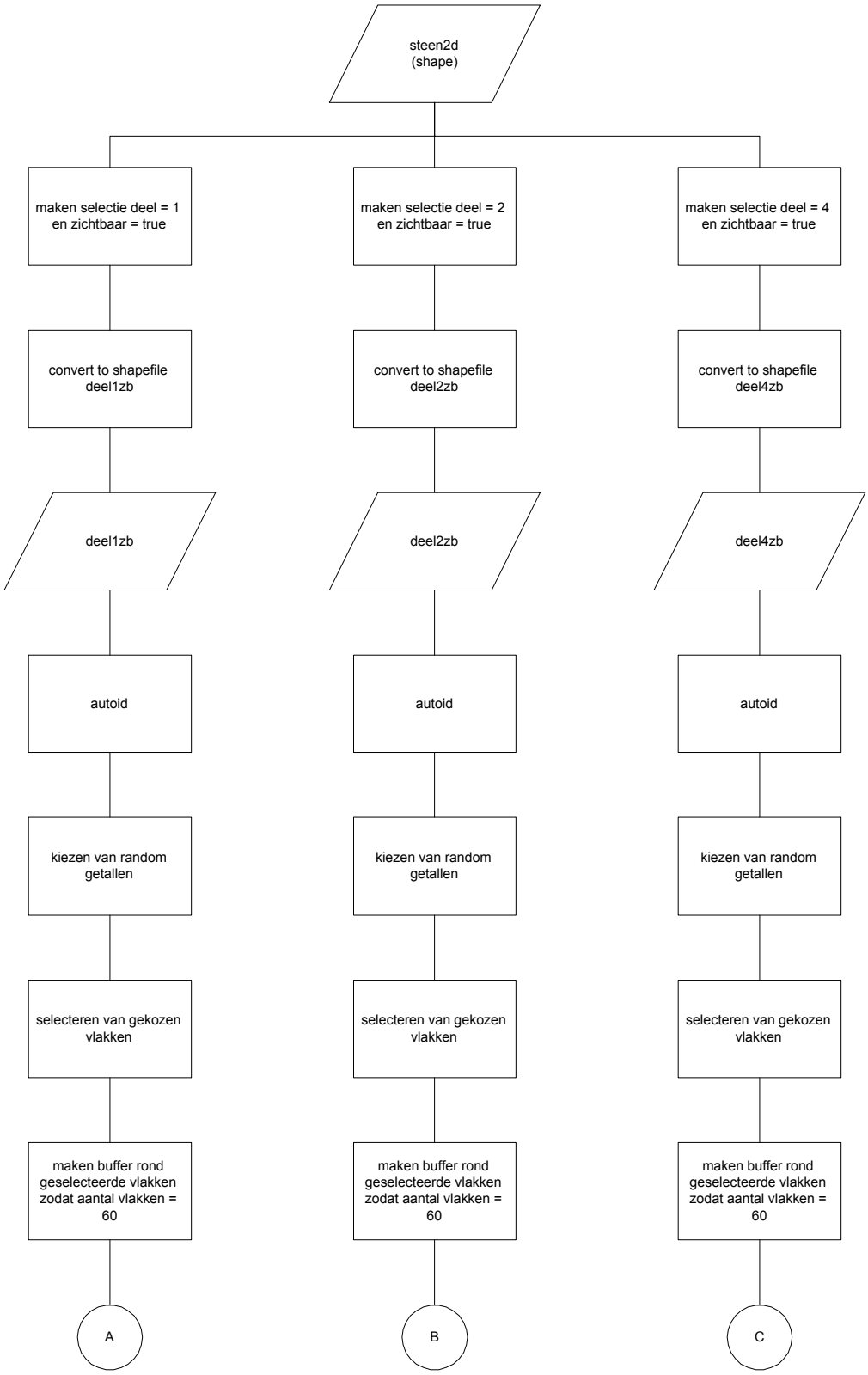
Bijlage 8 Processchema toevoegen zichtbaarheid

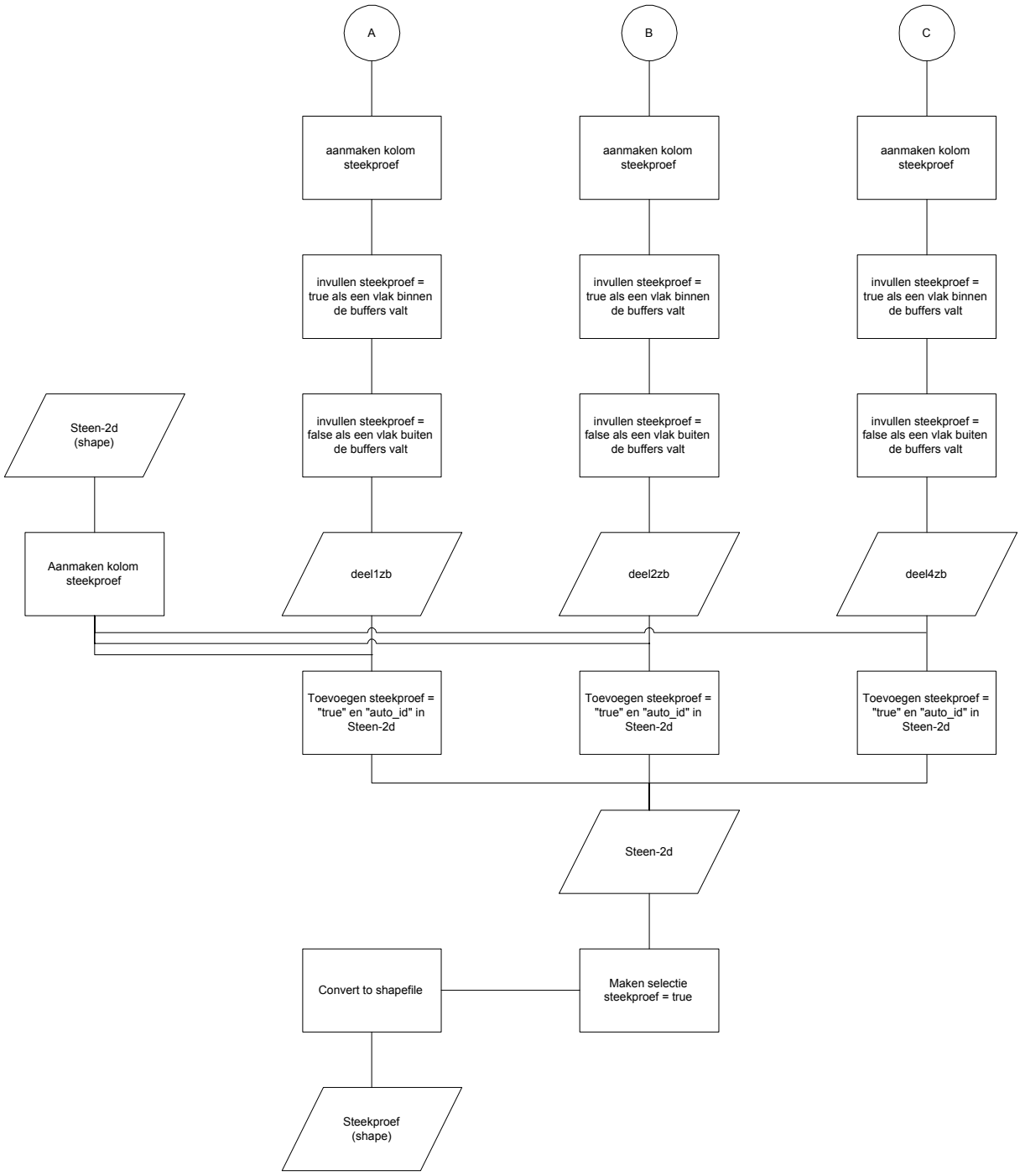


Bijlage 9 Toevoegen gebied aan waterkeringenbestand



Bijlage 10 Processchema kiezen te toetsen vlakken



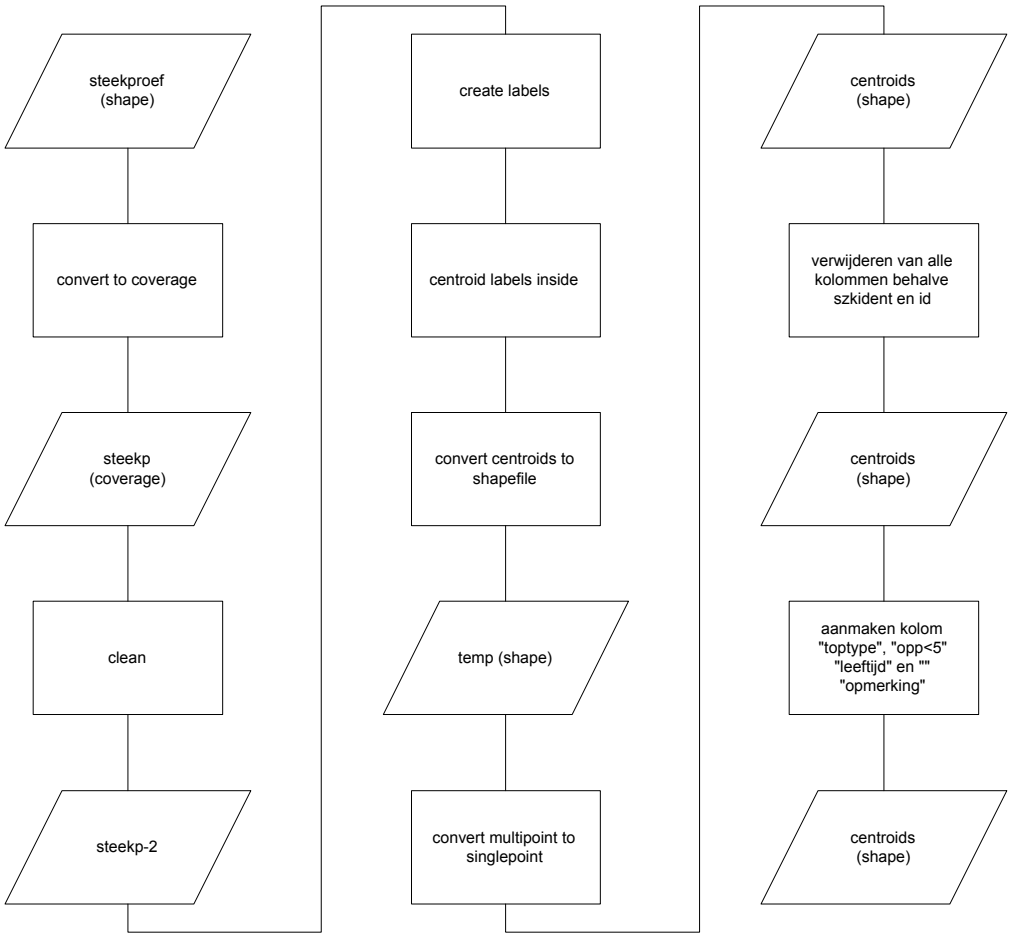

A
B
C
aanmaken kolom steekproef
aanmaken kolom steekproef
aanmaken kolom steekproef
invullen steekproef = true als een vlak binnen de buffers valt
invullen steekproef = true als een vlak binnen de buffers valt
invullen steekproef = true als een vlak binnen de buffers valt
Steen-2d (shape)
invullen steekproef = false als een vlak buiten de buffers valt
invullen steekproef = false als een vlak buiten de buffers valt
invullen steekproef = false als een vlak buiten de buffers valt
Aanmaken kolom steekproef
deel1zb
deel2zb
deel4zb
Toevoegen steekproef = "true" en "auto_id" in Steen-2d
Toevoegen steekproef = "true" en "auto_id" in Steen-2d
Toevoegen steekproef = "true" en "auto_id" in Steen-2d
Steen-2d
Convert to shapefile
Maken selectie steekproef = true
Steekproef (shape)

Bijlage 11 Lijst met coderingen van steenbekledingen

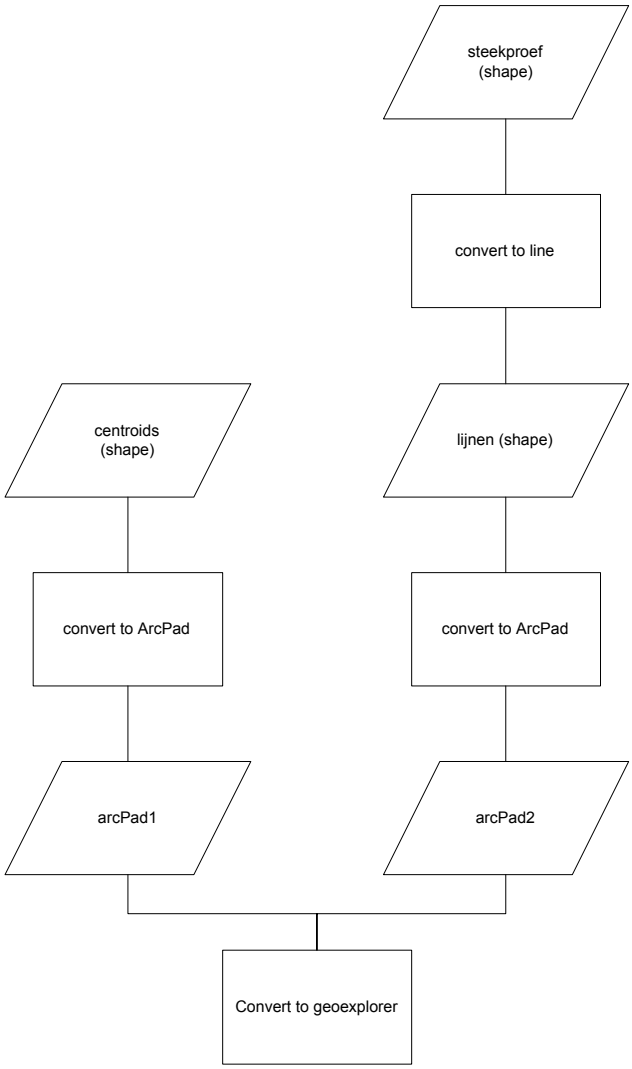
Toplaagtype	Omschrijving
100	Asfaltbeton
200	Mastiek
300	Dicht steenasfalt
400	Open geprefabriceerde steenasfaltmatten
500	Open steenasfalt
510	Fixstone (open steenasfalt)
600	Zandasfalt (tijdelijk of in onderlaag)
700	Breuksteen, gepenetreerd met asfalt (vol en zat)
710	Grauwakke (Breuksteen), gepenetreerd met asfalt (vol en zat)
800	Baksteen/betonsteen, gepenetreerd met asfalt (vol en zat)
900	Breuksteen, gepenetreerd met asfalt (patroonpenetratie)
1000	Betonblokken met afgeschuinde hoeken of gaten erin
1010	Betonblokken met grote afgeschuinde hoeken (5 cm)
1100	Betonblokken zonder openingen
1101	Betonblokken zonder openingen, gepentreerd met asfalt
1110	Haringmanblokken
1120	Diaboolblokken
1130	Gebakken steen
1131	Gebakken steen, gepentreerd met asfalt
1132	Gebakken steen, gepentreerd met beton
1140	Betonblokken system Pitt
1141	Betonblokken system Pitt, gepentreerd met asfalt
1142	Betonblokken system Pitt, gepentreerd met beton
1150	Betonblokken zonder openingen gekanteld
1160	Haringmanblokken gekanteld
1200	Open blokkenmatten, afgestrooid met granulair materiaal
1300	Blokkenmatten zonder openingen
1400	Betonplaten van cementbeton of gesloten colloidaal beton, (in situ gestort)
1410	Muraltglooiing
1500	Colloidaal beton, (open structuur)
1600	Betonplaten, (prefab)
1700	Doorgroeisteen, beton
1800	Breuksteen, gepenetreerd met cementbeton of colloidaal beton, (vol en zat)
1900	Breuksteen, met patroonpenetratie van cementbeton of colloidaal beton
2000	Gras, gezaaid
2100	Gras, zoden of gezaaid, in kunstofmatten
2200	Bestorting van grof grind en andere granulaire materialen
2300	Grove granulaire materialen c.q. breuksteen verpakt in metaalgaas
2400	Fijne granulaire materialen c.q. zand/grind verpakt in geotextiel
2500	Breuksteen, (stortsteen)
2600	Basalt, gezet
2601	Basalt, gezet, ingegoten met gietasfalt
2602	Basalt, gezet, ingegoten met colloidaal beton of cementbeton
2700	Betonzuilen en andere niet rechthoekige blokken
2701	Betonzuilen of niet rechthoekige blokken, ingegoten met gietasfalt
2702	Betonzuilen of niet rechthoekige blokken, ingegoten met beton
2710	Basalton
2711	Basalton, ingegoten met gietasfalt
2712	Basalton, ingegoten met beton
2720	PIT Polygoon zuilen
2721	PIT Polygoon zuilen, ingegoten met gietasfalt
2730	Hydroblock

Toplaagtype	Omschrijving
2731	Hydroblock, ingegoten met gietasfalt
2740	Basalton met ecolaag
2750	Hydroblock met ecolaag
2800	Natuursteen, gezet
2801	Natuursteen, gezet, en ingegoten met gietasfalt
2802	Natuursteen, gezet, en ingegoten met beton
2810	Vilvoordse
2811	Vilvoordse, ingegoten met gietasfalt
2812	Vilvoordse, ingegoten met beton
2813	Vilvoordse, overlaagd met asfalt gepenetreerde stortsteen (fixstone,grauwacke)
2814	Vilvoordse, overlaagd met beton gepenetreerde stortsteen
2820	Lessinische
2821	Lessinische, ingegoten met gietasfalt
2822	Lessinische, ingegoten met beton
2830	Doornikse
2831	Doornikse, ingegoten met gietasfalt
2832	Doornikse, ingegoten met beton
2840	Petit graniet
2841	Petit graniet, ingegoten met gietasfalt
2842	Petit graniet, ingegoten met beton
2843	Petit graniet, overlaagd met asfalt
2850	Graniet
2851	Graniet, ingegoten met gietasfalt
2852	Graniet, ingegoten met beton
2861	Grauwacke, ingegoten met gietasfalt
2870	Doorniks met gekantelde patronen
2871	Doorniks met gekantelde patronen, ingegoten met gietasfalt
2872	Doorniks met gekantelde patronen, ingegoten met beton
2900	Koperslakblokken
2901	Koperslakblokken gepenetreerd met asfalt
3000	Klei onder zand
3100	Bestorting van natuursteenmassa
3200	Klinkers, beton of gebakken.
3210	Tegels
3220	Dakpannen
3300	Zand
3400	Steenfundering, gebonden
3900	Zetwerk, ratjetoe
5100	Uitstroombak
5200	Muraltmuur, dijkmuur
5600	Kade, keermuur, kistdam
5700	Betonnen trap
5800	Betonnen fietspad
5900	Diverse constructies
6000	Oeverwerk: zinkstuk
6100	Oeverwerk: bestorting
6200	Oeverwerk: zinkstuk + bestorting
9000	Bunker
9100	Gebouw e.d.
9800	Diverse objecten
9900	Onbekend

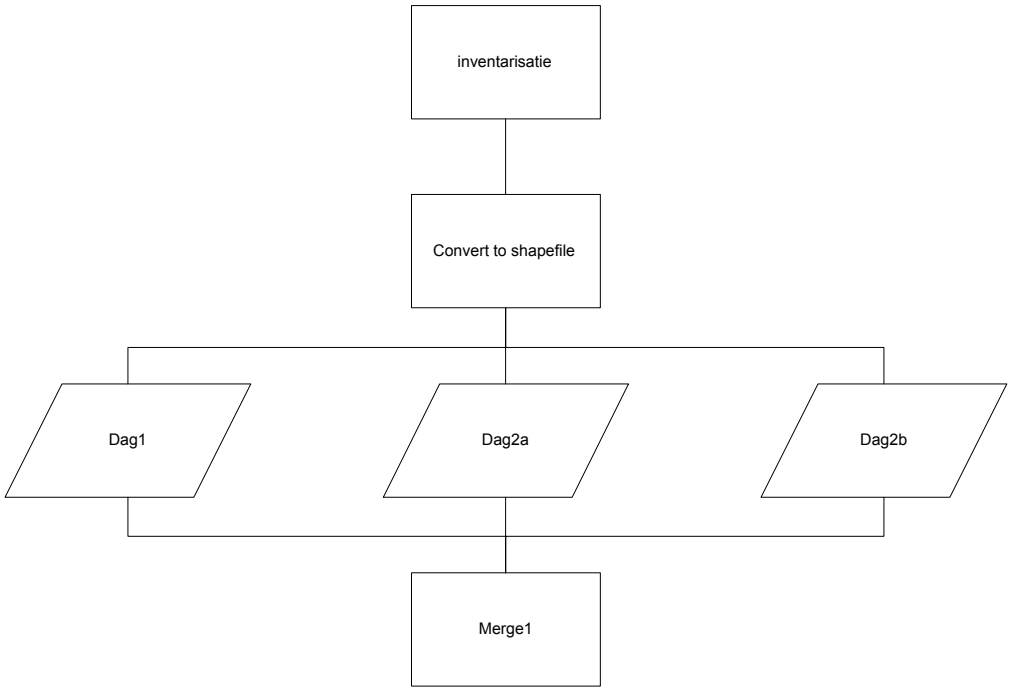
Bijlage 12 Het genereren van het centroidenbestand



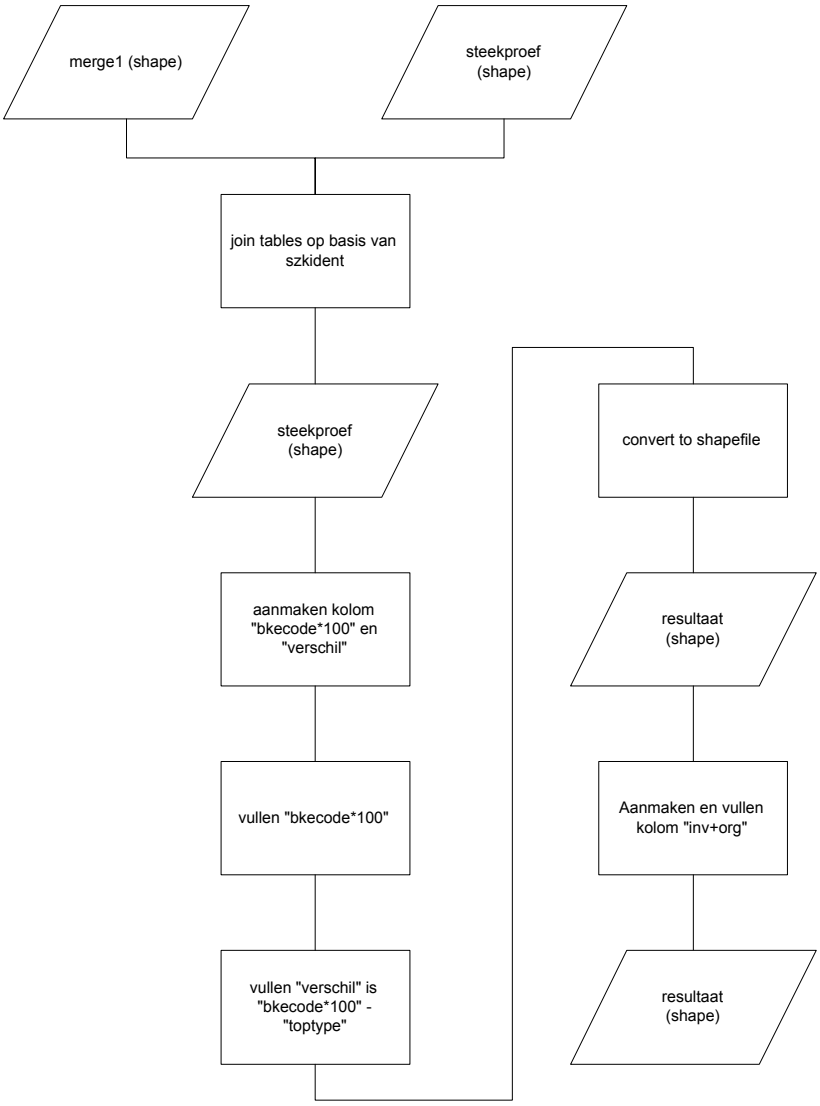
Bijlage 13 Het gereedmaken van de Geo-Explorer



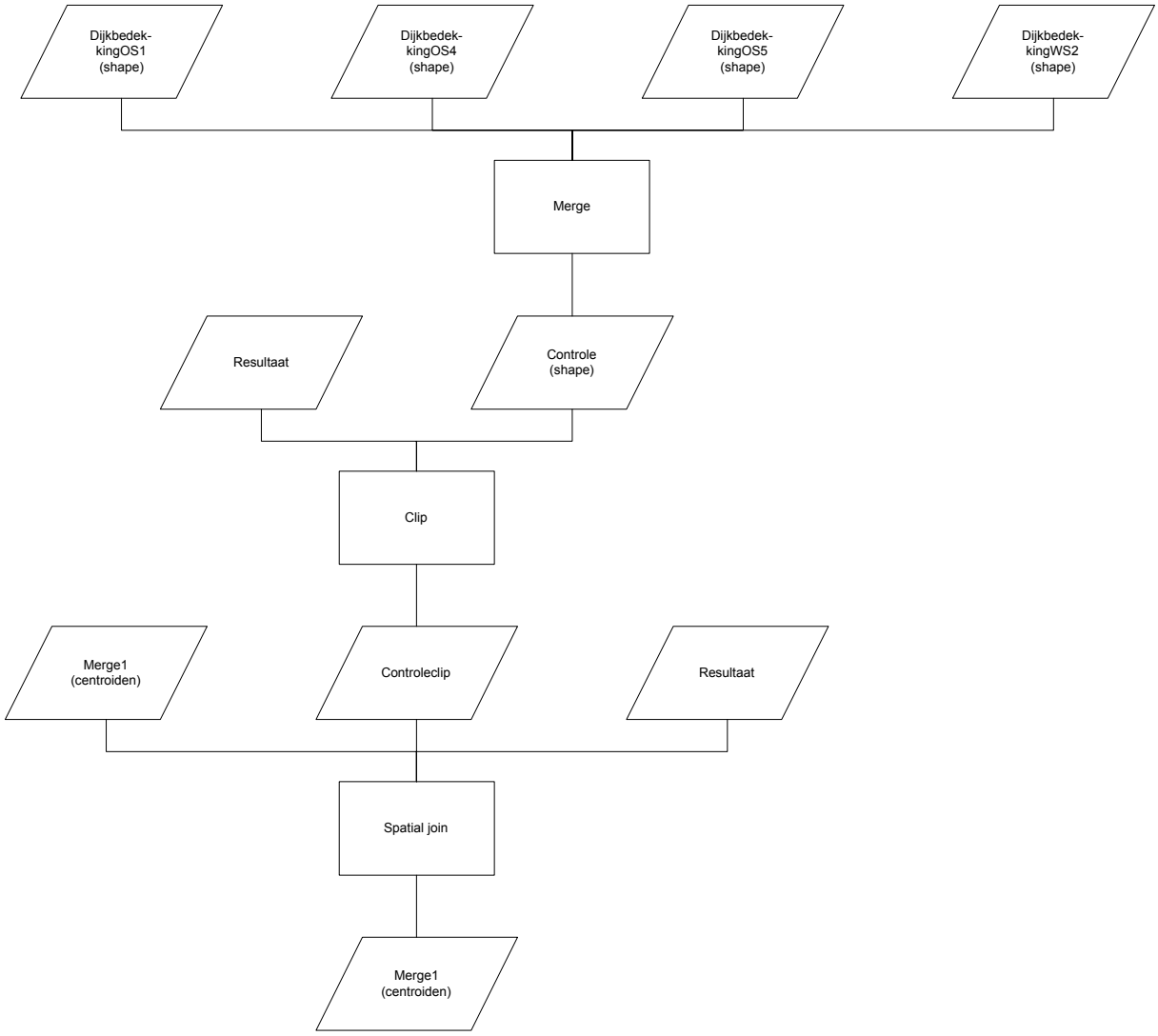
Bijlage 14 Conversie van Geo-explorer naar werkstation



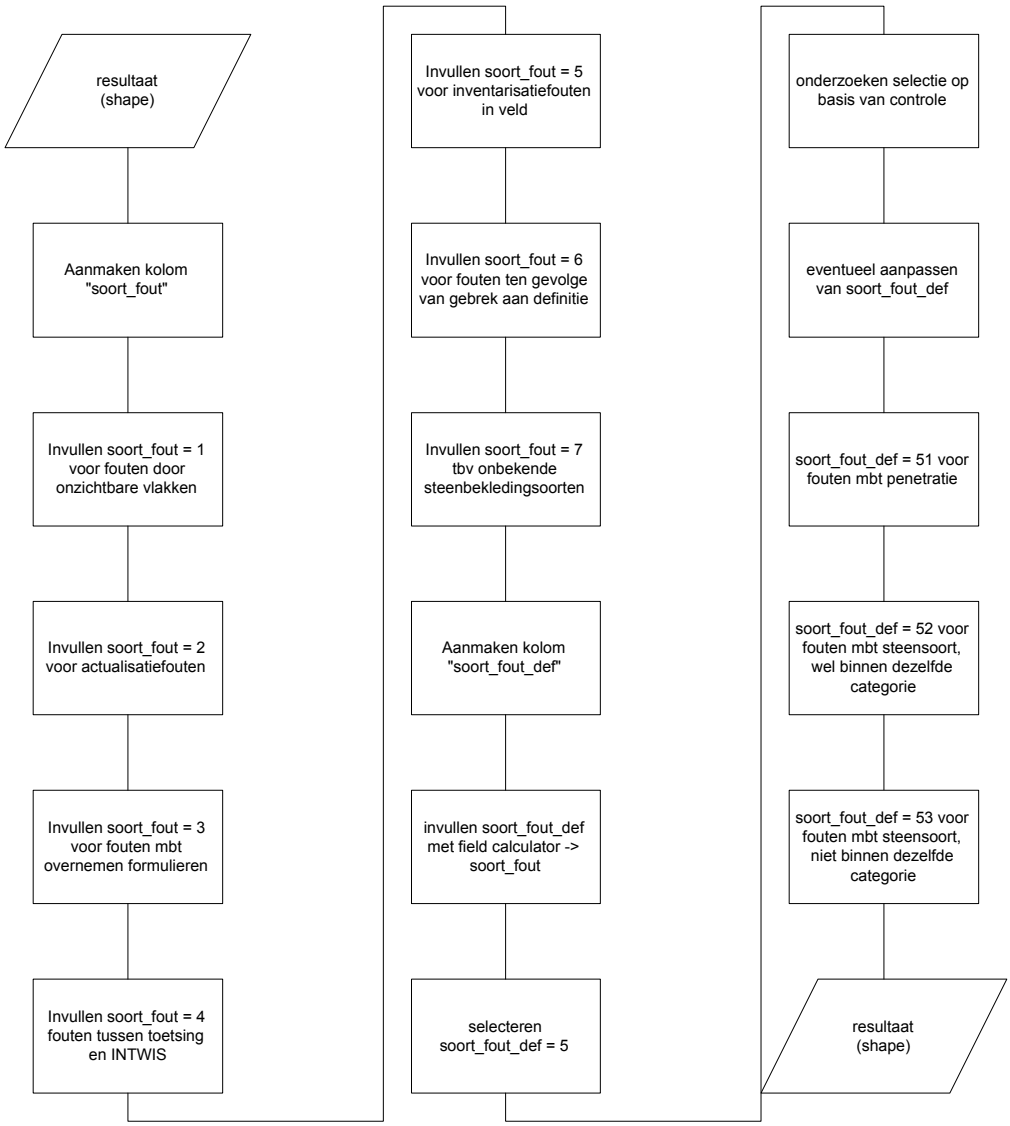
Bijlage 15 Verwerken inventarisatie tot resultaat



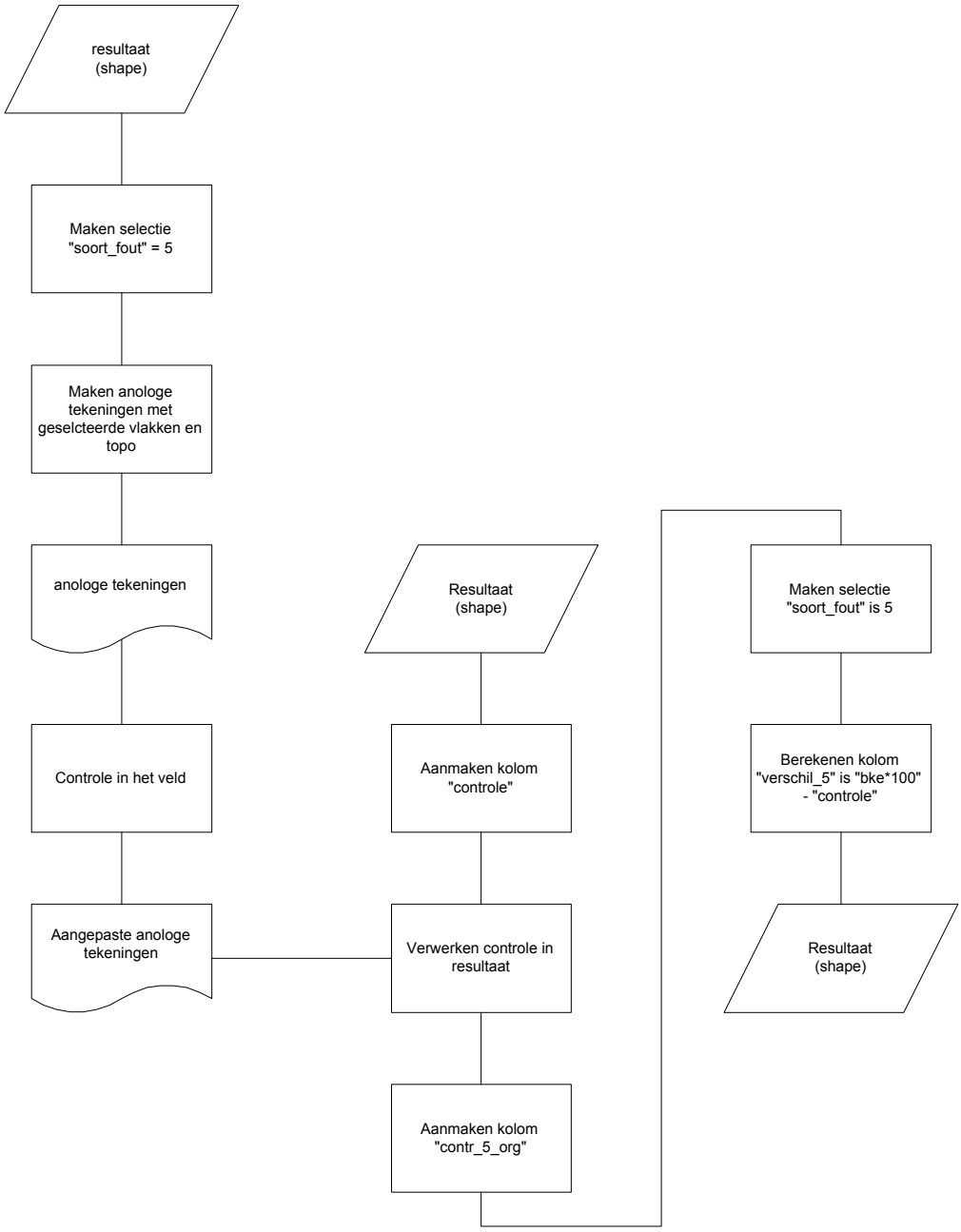
Bijlage 16 Vergelijken bestanden t.b.v. conversiefouten



Bijlage 17 Invullen van de soorten fouten



Bijlage 18 Controle niet te verklaren fouten



Bijlage 19 Confusionmatrix van alle waarnemingen

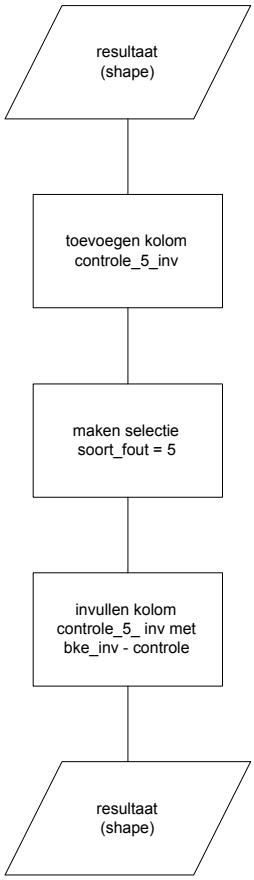
			classificatie in het veld																																									
			700	1100	1110	1130	1132	1400	1700	2000	2600	2601	2602	2701	2710	2730	2740	2750	2800	2802	2810	2811	2812	2813	2820	2821	2822	2830	2831	2832	2840	2841	2842	2851	2861	3200	3900	9800	9900	totaal
classificatie in de	700		0																																1						1
database	1100			29	1		1	1			1																											2		35
	1110			1	20																										1							3		25
	1130					0				1																														1
	1132						0																																	0
	1400							0																																0
	1700			1					10																															11
	2000									0																														0
	2600										36	1	4																										1	42
	2601		1					1			2	10																								1			1	16
	2602										2		15																									3		20
	2701													0	1																									1
	2710														4		1																							5
	2730															4																							2	6
	2740																2																							2
	2750																	4																						4
	2800																		0																					0
	2802																			0			1																	1
	2810										2										15	3	3				1	1									3		2	30
	2811		1																			4																			5
	2812									1			1						1		1	2	26								1								1		34
	2813																							2																	2
	2820																								1													1			2
	2821																									1															1
	2822																			1			1		1		1	1										2			7
	2830		1																									4		2										1	8
	2831																												1												1
	2832																													3											3
	2840																														5										5
	2841																															5									5
	2842																																0					1			1
	2851																											1			1			0							2
	2861																																		0						0
	3200					1																														10		3		14
	3900				1						1												2														2			6
	9800																																					0		
	9900			1																																			0	1
	totaal		3	32	22	1	1	2	10	2	44	11	20	0	5	4	3	4	1	1	16	9	33	2	2	1	2	7	1	6	7	5	0	0	1	11	9	12	7	297
nauwkeurigheid (%)			0	91	91	0	0	0	100	0	82	91	75		80	100	67	100	0	0	94	44	79	100	50	100	50	57	100	50	71	100			0	91	22	0	0	72.1
omissie (%)			100	9	9	100	100	100	0	100	18	9	25		20	0	33	0	100	100	6	56	21	0	50	0	50	43	0	50	29	0			100	9	78	100	100	27.9
overcompleetheid (%)			100	17	20	100			9		14	38	25	100	20	33	0	0		100	50	20	24	0	50	0	86	50	0	0	0	0	100	100		29	67	100	100	
betrouwbaarheid (%)			0	83	80	0			91		86	63	75	0	80	67	100	100		0	50	80	76	100	50	100	14	50	100	100	100	100	0	0		71	33	0	0	
			0	4	2	0	0	0	0	0	6	1	1	0	0	0	0	0	0	0	2	0	4	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	21
kappa		1																																						

Bijlage 20 Confusionmatrix van fouten 51, 52 en 53

		classificatie in het veld																																				
		700	1100	1110	1130	1132	1400	1700	2000	2600	2601	2602	2701	2710	2730	2740	2750	2800	2802	2810	2811	2812	2813	2820	2821	2822	2830	2831	2832	2840	2841	2842	2851	2861	3200	3900	9800	9900	totaal
classificatie in de	700	0																																					0
database	1100		29				1			1																													31
	1110		1	20																																			21
	1130				0				1																														1
	1132					0																																	0
	1400						0																																0
	1700		1					10																															11
	2000								0																														0
	2600									36	1	3																											40
	2601									2	10																												12
	2602									1		15																											16
	2701												0																										0
	2710													4		1																							5
	2730														4																								4
	2740															2																							2
	2750																4																						4
	2800																	0																					0
	2802																		0			1																	1
	2810									2										15	1	3					1												22
	2811																				4																		4
	2812											1						1		1	1	26							1										31
	2813																						2																2
	2820																							1															1
	2821																								1														1
	2822																		1			1				1	1												4
	2830	1																									4		2										7
	2831																											1											1
	2832																												3										3
	2840																													5									5
	2841																														5								5
	2842																															0							0
	2851																													1			0						1
	2861																																	0					0
	3200																																		10				10
	3900																																			2			2
	9800																																				0		
	9900																																					0	0
	totaal	1	31	20	0	0	1	10	1	42	11	19	0	4	4	3	4	1	1	16	6	31	2	1	1	1	6	1	6	6	5	0	0	0	10	2	0	0	247
nauwkeurigheid (%)		0	94	100			0	100	0	86	91	79		100	100	67	100	0	0	94	67	84	100	100	100	100	67	100	50	83	100				100	100			11.1
omissie (%)		100	6	0			100	0	100	14	9	21		0	0	33	0	100	100	6	33	16	0	0	0	0	33	0	50	17	0				0	0			88.9
overcompleetheid (%)			6	5	100			9		10	17	6		20	0	0	0		100	32	0	16	0	0	0	75	43	0	0	0	0		100		0	0	100		
betrouwbaarheid (%)			94	95	0			91		90	83	94		80	100	100	100		0	68	100	84	100	100	100	25	57	100	100	100	100		0		100	100	0		

goed
fout 51
fout 52
fout 53

Bijlage 21 Bepalen nauwkeurigheid Geo-explorer



Bijlage 22 Vragenlijst bepalen benodigde kwaliteit

Vraag 1

Voor welke werkprocessen gebruik jij het bestand met de steenbekledingen en dan met name voor wat betreft het toplaagtype?

Werkprocessen:

Vraag 2

Voor welke andere werkprocessen wordt bovenstaand bestand ook gebruikt (door anderen binnen jouw gezichtsveld)?

Werkprocessen:

Vraag 3

Wat kunnen de gevolgen zijn van een verkeerde invulling van het toplaagtype? Geef ook aan of dit wel eens voorgekomen is?

Gevolgen Wel eens voorgekomen j/n

Vraag 4

Is de grootte van de gevolgen afhankelijk van kenmerken van het steenbekledingsvlak? Zoja, welke kenmerken hebben hier dan invloed op. Als hulpmiddel hierbij heb ik een aantal voorbeelden opgesomd.

Bestaat er een verband tussen grootte van het gevolg en de afmetingen van een vlak. Bijvoorbeeld een verkeerd geclassificeerd groot vlak kan grotere gevolgen hebben dan een verkeerd geclassificeerd klein vlak. Het kenmerk in dit geval is “afmeting”!

Maakt het bijvoorbeeld uit of een steenbekledingsvlak hoog of laag op de dijk ligt. Een voorbeeld zou kunnen zijn dat als een verkeerd geclassificeerd vlak boven 2 meter + N.A.P. ligt, de mogelijke gevolgen kleiner zijn dan voor een verkeerd geclassificeerd vlak dat lager dan 2 meter + N.A.P. ligt. Het kenmerk in dit geval is “hoogte op waterkering”!

Maakt het uit of het of een verkeerd geclassificeerd vlak op een waterkering ligt of op een niet waterkering, bijvoorbeeld strekdam. Het kenmerk in dit geval is “waterkerende functie”.

Maakt het uit of een verkeerd geclassificeerd vlak op een verbeterd gedeelte ligt of op een nog niet verbeterd gedeelte. Het kenmerk in dit geval is “status van verbetering”.

Wellicht kloppen bovenstaande kenmerken niet of is het mogelijk andere kenmerken te definieren. Schrijf deze op. Uiteindelijk gaan ze de benodigde kwaliteit bepalen. Het is natuurlijk ook mogelijk dat er geen kenmerken onderscheiden kunnen worden maar dat alle vlakken als eenzelfde beschouwd worden. Ga dan direct door naar vraag 6.

Kenmerken:

Vraag5

Probeer van alle kenmerken, waarvan jij van mening bent dat die invloed hebben op de grootte van het gevolg van een verkeerd geclassificeerd vlak, aan te geven waar de grenzen liggen van meer of minder gevolg. Voor waterkering of niet waterkering is dat eenvoudig. Bij een verband tussen de afmetingen van de vlakken ligt dat iets anders. Probeer hiervoor ook een grens te definieren. Afgelopen dinsdag hadden we een grens van 300 m² gedefinieerd, maar misschien zijn er meerdere grenzen te definieren.

Kenmerk	Grens

Vraag6

Geef voor ieder kenmerk en mogelijkheid binnen het kenmerk aan in hoeveel % van de gevallen de classificatie correct moet zijn. Dit kan er dan bijvoorbeeld als volgt uitzien:

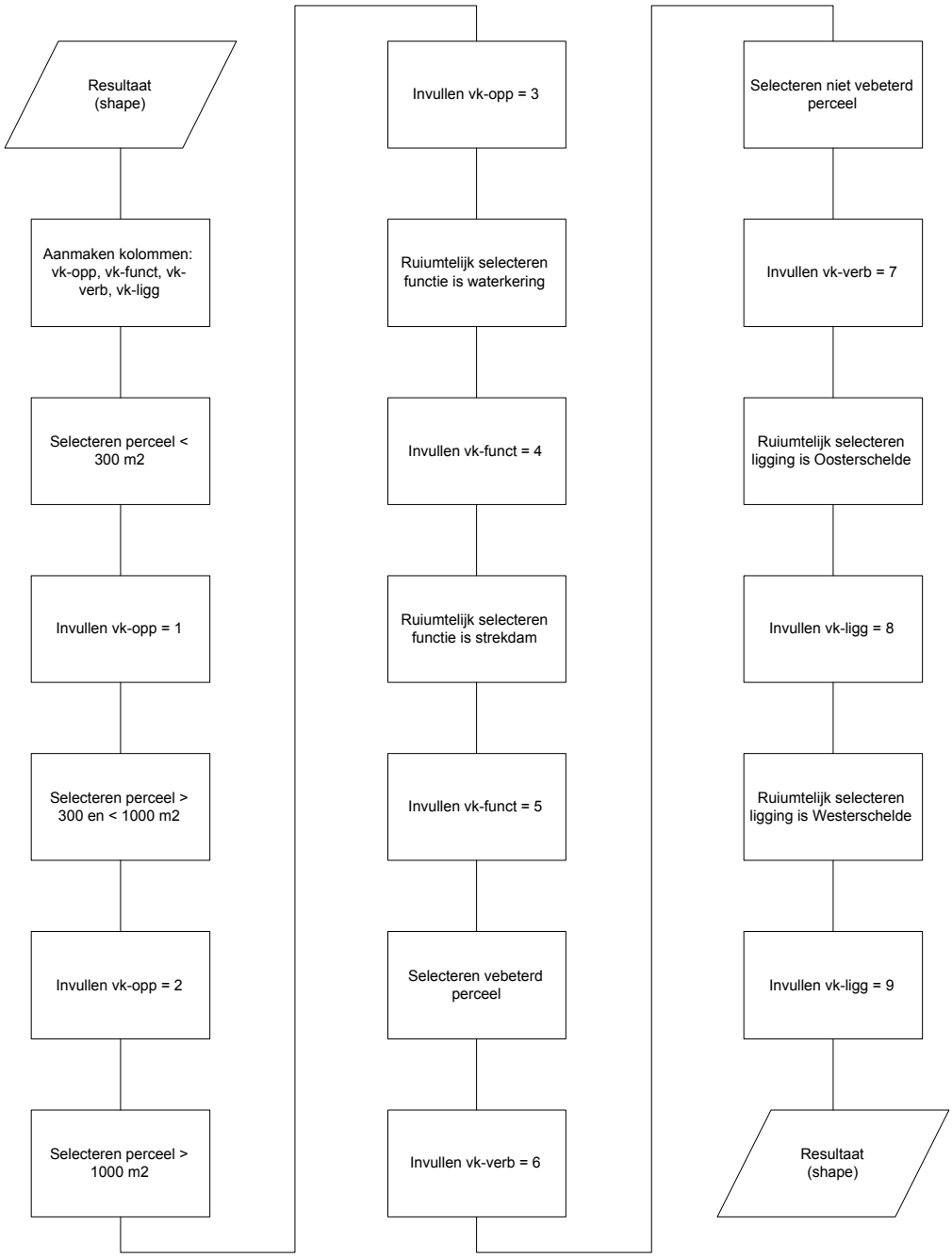
Kenmerk	Mogelijkheid	Correct
Afmeting	>300 m²	95 %
	<300 m²	80 %
Soort	Waterkering	98 %
	Strekdam	60%

Als er geen kenmerken zijn gedefinieerd dient hier voor het gehele bestand aangegeven te worden in hoeveel % van de gevallen de classificatie correct moet zijn.

Kenmerk	Mogelijkheid	Correct

Op basis van de laatste antwoorden kan nu de benodigde kwaliteit gedefinieerd worden, die afhankelijk van het aantal kenmerken meervoudig kan zijn. Bijvoorbeeld: vlakken met een waterkerende functie groter dan 300 m² dienen in 95 % van de gevallen goed gecodeerd te zijn. Vlakken met een niet-waterkerende functie dienen in 60 % van de gevallen goed gecodeerd te zijn. Vlakken met een waterkerende functie kleiner dan 300 m² dienen in 80 % van de gevallen goed gecodeerd te zijn.

Bijlage 23 Invullen kenmerken vereiste kwaliteit



Bijlage 24 Vergelijkingspercentage vereiste kwaliteit

Oppervlakte in m2	Functie	Verbeterd	Geografische ligging	Vereiste attribuutnauwkeurig-heid in %	Aantallen	Vereiste nauwkeurigheid * aantallen in %
1	4	6	8	82	0	0.0
1	4	7	8	80	101	27.2
1	4	6	9	87	7	2.0
1	4	7	9	85	0	0.0
2	4	6	8	95	0	0.0
2	4	7	8	90	44	13.3
2	4	6	9	98	3	1.0
2	4	7	9	95	0	0.0
3	4	6	8	98	0	0.0
3	4	7	8	95	21	6.7
3	4	6	9	99	20	6.7
3	4	7	9	98	0	0.0
1	5	6	8	80	0	0.0
1	5	7	8	80	52	14.0
1	5	6	9	82	23	6.4
1	5	7	9	80	0	0.0
2	5	6	8	82	0	0.0
2	5	7	8	80	14	3.8
2	5	6	9	82	4	1.1
2	5	7	9	80	0	0.0
3	5	6	8	85	0	0.0
3	5	7	8	80	5	1.3
3	5	6	9	85	3	0.9
3	5	7	9	80	0	0.0
				Totaal	297	84.4