Aanpassing en uitbreiding controlemethodiek BRT

Aanbevelingen om de methodiek van de externe kwaliteitscontrole voor de
BasisRegistratie Topografie aan te passen en uit te breiden

M.H. Storm, J. Davidse, D.J. Brus

WAGENINGEN
UNIVERSITY & RESEARCH

Aanpassing en uitbreiding controlemethodiek BRT

Aanbevelingen om de methodiek van de externe kwaliteitscontrole voor de Basisregistratie Topografie aan te passen en uit te breiden

M.H. Storm, J. Davidse, D.J. Brus

Dit onderzoek is uitgevoerd door Wageningen Environmental Research (Alterra) in opdracht van en gefinancierd door het Kadaster.

Wageningen Environmental Research
Wageningen, juni 2017

Rapport 2796
ISSN 1566-7197

Storm, M.H., J. Davidse, D.J. Brus, 2017. *Aanpassing en uitbreiding controlemethodiek BRT; Aanbevelingen om de methodiek van de externe kwaliteitscontrole voor de Basisregistratie Topografie aan te passen en uit te breiden*. Wageningen, Wageningen Environmental Research, Rapport 2796. 38 blz.; 7 fig.; 18 tab.; 12 ref.

Er is in de Kadasterwet vastgelegd dat de kwaliteit van de Basisregistratie Topografie (BRT) eenmaal in de drie jaar door een externe deskundige moet worden gecontroleerd. De tot nu toe gehanteerde methodiek is om meerdere redenen niet meer toereikend. Als eerste is onderzocht hoe de huidige methode is aan te passen aan het nieuwe gegevensmodel van TOP10NL. Vervolgens is onderzocht hoe ervoor gezorgd kan worden dat van elke objectklasse voldoende objecten in de steekproef zitten, zodat de resultaten van controle betrouwbaarder zijn. Ten slotte is ook onderzocht hoe de methode uit te breiden is met controlemethodieken voor de kleinschalige producten.

In the Land Registry Act it is stated that the quality of the Key Register Topography (BRT) has to be audited by an external expert once every three years. The method which is used up till now, is not sufficient anymore for several reasons. First, it is studied how to adjust the current method to the new data model of TOP10NL. Next, it is investigated in what way the sample design can be adjusted to ensure enough objects of each object class for a more reliable result of the quality audit. Finally, a study has been done to be able to expand the current method in such a way that is applicable for the small-scale products.

Trefwoorden: steekproef, kwaliteit, basisregistratie, gegevensmodel, controlemethodiek

Dit rapport is gratis te downloaden van <http://dx.doi.org/10.18174/417956> of op www.wur.nl/environmental-research (ga naar 'Wageningen Environmental Research' in de grijze balk onderaan). Wageningen Environmental Research verstrekt *geen* gedrukte exemplaren van rapporten.

© 2017 Wageningen Environmental Research (instituut binnen de rechtspersoon Stichting Wageningen Research), Postbus 47, 6700 AA Wageningen, T 0317 48 07 00, E info.alterra@wur.nl, www.wur.nl/environmental-research. Wageningen Environmental Research is onderdeel van Wageningen University & Research.

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

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

Wageningen Environmental Research Rapport 2796 | ISSN 1566-7197

Foto omslag: Shutterstock

Inhoud

	Samenvatting	5
1	Inleiding	7
2	Deelonderzoek 1: Aanpassing van de huidige methode aan het nieuwe gegevensmodel van TOP10NL	8
3	Deelonderzoek 2: Aanpassing van de huidige methode om te zorgen voor voldoende objecten in de steekproef voor elke objectklasse	19
	3.1 Inleiding	19
	3.2 Onderzoek	20
	3.3 Alternatieve steekproefopzet	22
	3.3.1 Alternatief 1: gecombineerde steekproef	22
	3.3.2 Alternatief 2: vergroten van de steekprofeenheden	24
	3.4 Advies	25
4	Deelonderzoek 3: Uitbreiding van de huidige methode met controle-methodieken voor de kleinere schalen	26
	4.1 Actualiteit	27
	4.1.1 Beschrijving	27
	4.1.2 Methode	27
	4.1.3 Toepasbaarheid	27
	4.2 Logische consistentie	27
	4.2.1 Beschrijving	27
	4.2.2 Methode	27
	4.2.3 Toepasbaarheid	28
	4.3 Positionele nauwkeurigheid	28
	4.3.1 Beschrijving	28
	4.3.2 Methode	28
	4.3.3 Toepasbaarheid	28
	4.4 Volledigheid	29
	4.4.1 Beschrijving	29
	4.4.2 Methode	29
	4.4.3 Toepasbaarheid	29
	4.5 Thematische nauwkeurigheid	31
	4.5.1 Beschrijving	31
	4.5.2 Methode	32
	4.5.3 Toepasbaarheid	32
	4.6 Bruikbaarheid	32
	4.6.1 Beschrijving	32
	4.6.2 Methode	33
	4.6.3 Toepasbaarheid	34
	4.7 Overzicht van mogelijke controles	35
	Literatuur	37

Samenvatting

Er is in de Kadasterwet vastgelegd dat de kwaliteit van de Basisregistratie Topografie (BRT) eenmaal in de drie jaar door een externe deskundige moet worden gecontroleerd. De tot nu toe gehanteerde methodiek is om meerdere redenen niet meer toereikend. Dit rapport beschrijft de resultaten van drie deelonderzoeken:

- Aanpassing van de huidige methode aan het nieuwe gegevensmodel van TOP10NL.
- Aanpassing van de huidige methode om te zorgen voor voldoende objecten in de steekproef voor elke objectklasse.
- Uitbreiding van de huidige methode met controlemethodieken voor de kleinere schalen.

De wijziging van het gegevensmodel van TOP10NL heeft als gevolg dat er in een externe controle meer tijd besteed zal moeten worden aan de bepaling van het kwaliteitsaspect 'thematische nauwkeurigheid'. Naar verwachting zal de externe audit in totaal ongeveer 10 dagen extra kosten t.o.v. de vorige audits bij een gelijkblijvende steekproef.

Het aantal objecten in de steekproef is van invloed op de betrouwbaarheid van de resultaten van de externe audit. In eerdere externe audits was het aantal objecten in bepaalde objectklassen nog erg laag, wat leidde tot een geringe betrouwbaarheid van de resultaten. In dit rapport wordt inzicht gegeven in het effect van een grotere steekproefomvang op de betrouwbaarheid. Er is door het Kadaster van tevoren niet aangegeven wat het significantieniveau of onderscheidend vermogen moet zijn en daarom kan er geen advies worden gegeven voor een specifieke steekproefomvang. De keuze voor de steekproefomvang is aan het Kadaster zelf. In deze keuze zal ook het kostenaspect meewegen. Een grotere steekproefomvang leidt tot meer werk en dus hogere kosten. In dit rapport wordt een inschatting gegeven van de extra tijd die voor een controle nodig is bij uitbreiding van de steekproefomvang.

Een andere manier om de betrouwbaarheid van de resultaten te vergroten, is door een andere steekproefopzet toe te passen. Er wordt een gecombineerde steekproefopzet voorgesteld, wat leidt tot een hoger onderscheidend vermogen voor de objectklasse 'spoorbaandelen'. Deze gecombineerde steekproefopzet zal iets meer werk betekenen voor een volgende externe audit. Een te kleine steekproefomvang kan het positieve effect van de gecombineerde steekproef op de betrouwbaarheid echter tenietdoen. Nader onderzoek kan uitwijzen bij welke steekproefaantallen dit negatieve effect groter is dan het positieve effect van het toepassen van de gecombineerde steekproefopzet.

Literatuuronderzoek heeft geleid tot een of meerdere methodes om de kwaliteitsaspecten 'actualiteit', 'logische consistentie', 'positionele nauwkeurigheid', 'volledigheid', 'thematische nauwkeurigheid' en 'bruikbaarheid' te toetsen. Er wordt per kwaliteitsaspect beschreven of en hoe deze methoden toe te passen zijn op de raster- en/of objectgerichte kleinschalige producten. Er wordt beschreven welk nader onderzoek er eventueel nog nodig is om de methoden praktisch te kunnen uitvoeren.

1 Inleiding

De Basisregistratie Topografie (BRT) maakt onderdeel uit van het Stelsel van Basisregistraties. Kenmerkend voor de basisregistraties is dat er een wettelijke basis voor bestaat en dat zij alle moeten voldoen aan een bepaalde kwaliteit. Om deze kwaliteit te waarborgen, dient er regelmatig een kwaliteitscontrole uitgevoerd te worden. In de wet is vastgelegd dat voor de BRT de kwaliteitscontrole eenmaal in de drie jaar uitgevoerd dient te worden door een externe expert. De jaren dat er geen externe controle plaatsvindt, dient er intern een controle plaats te vinden (Kadasterwet, 2016).

De externe controle omvatte tot nu toe enkel TOP10NL. Het is gebleken dat met de huidige methodiek (Storm et al. 2012) er voor bepaalde objectklassen een te laag aantal objecten in de steekproef zit om de kwaliteit hiervan met hoge betrouwbaarheid te kunnen geven. Daarnaast is enige tijd geleden het gegevensmodel van TOP10NL veranderd en dat heeft invloed op de methodiek. Ten slotte zullen bij de volgende externe audit ook de kleinschalige objectgerichte producten en rasterproducten van de BRT gecontroleerd moeten worden. Daar bestond nog geen methodiek voor. Om een goede en volledige externe audit op de kwaliteit van de BRT uit te voeren, zal de controlemethodiek dus aangepast en uitgebreid moeten worden.

Dit rapport beschrijft de resultaten van drie deelonderzoeken:

- Aanpassing van de huidige methode aan het nieuwe gegevensmodel van TOP10NL.
- Aanpassing van de huidige methode om te zorgen voor voldoende objecten in de steekproef voor elke objectklasse.
- Uitbreiding van de huidige methode met controlemethodieken voor de kleinere schalen.

De volgende hoofdstukken gaan in op deze drie deelonderzoeken. In hoofdstuk 2 worden de resultaten beschreven van het eerste deelonderzoek. Hierin staat welke aanpassingen er nodig zijn in de methodiek m.b.t. het kwaliteitsaspect 'thematische nauwkeurigheid' en wat dit betekent voor de hoeveelheid werk.

Deelonderzoek 2, beschreven in hoofdstuk 3, toont aan op welke wijze de steekproef die gebruikt wordt in de controlemethodiek kan worden aangepast, zodat de betrouwbaarheid van de resultaten van een externe audit verhoogd wordt.

Ten slotte worden in hoofdstuk 4 de methoden beschreven die geschikt zijn om de kwaliteit van de raster- en objectgerichte producten op kleinere schalen te toetsen. Dit is het resultaat van deelonderzoek 3.

2 Deelonderzoek 1: Aanpassing van de huidige methode aan het nieuwe gegevensmodel van TOP10NL

Sinds de TOP10NL levering van november 2015 is een nieuwe versie (1.2) van het gegevensmodel van TOP10NL in gebruik. De wijzigingen t.o.v. voorgaande versies zijn dusdanig dat de huidige methode voor de externe controle (Storm et al. 2011) niet meer goed aansluit bij het gegevensmodel.

De huidige methodiek is nu zo opgezet dat voor het kwaliteitsaspect 'thematische nauwkeurigheid' elk attribuut van elke objectklasse minimaal eens per drie externe controles wordt meegenomen. Omdat er wijzigingen zijn in de objectklassen en attributen, is er uitgezocht wat de gevolgen van de wijziging van het gegevensmodel op de methodiek is. Uitgangspunt daarbij is dat ook in de nieuwe methodiek alle attributen eens per drie externe controles worden meegenomen.

Er komen verschillende soorten wijzigingen in het gegevensmodel. Sommige modelwijzigingen betreffen enkel een naamswijziging. Het attribuut 'versieBeginTijd' heet bijvoorbeeld in gegevensmodel 1.2 'tjdstipRegistratie'. In dit soort gevallen kan de huidige methode voortgezet worden, waarbij er rekening wordt gehouden met de naamswijziging. In andere gevallen is er sprake van geheel nieuwe attributen (zoals 'isBAGNaam') of nieuwe attributen ontstaan uit eerdere attributen (zoals 'naam', ontstaan uit 'naamNL' en 'naamFries'). Ook zijn er attributen opgeheven (zoals 'dimensie').

In samenwerking met het Kadaster is voor elk attribuut onderzocht en vastgelegd of en hoe goed het attribuut is te controleren. Bepaalde attributen zijn niet of nauwelijks te controleren m.b.v. luchtfoto's of cyclorama's. Daarom is ook onderzocht of attributen op een andere wijze gecontroleerd kunnen worden. Ook is een schatting gemaakt in welke mate de hoeveelheid werk van een externe controle wordt beïnvloed.

Dit rapport beschrijft een aangepast methodiek voor de controle van de thematische nauwkeurigheid.

Hieronder wordt eerst per attribuut van elke TOP10NL-klasse van het nieuwe gegevensmodel (versie 1.2) aangegeven wat het bronmateriaal is voor een controle. Dit kunnen orthofoto's, cyclorama's of andere nader beschreven bronnen zijn. Daarnaast wordt aangegeven of het attribuut nooit gecontroleerd hoeft te worden, bij elke externe controle gecontroleerd moet worden of dat d.m.v. een steekproef bepaald moet worden of een attribuut gecontroleerd moet worden. In dit laatste geval wordt ervoor gezorgd dat het attribuut minimaal eens per drie externe controles wordt meegenomen.

Over het algemeen geldt dat de attributen die vaak belangrijk worden geacht voor analyses en selecties altijd gecontroleerd moeten worden en minder belangrijke attributen eens per drie externe controles.

Indien een attribuut niet gecontroleerd kan worden, dan wordt de reden hiervan aangegeven.

In de laatste drie kolommen wordt aangegeven of een attribuut (eventueel onder een andere naam) al in de eerdere externe audits van de data van september 2011 en november 2014 is gecontroleerd. Uit de voorgaande kolommen wordt vervolgens afgeleid of het attribuut in de komende externe audit van de data van november 2017 gecontroleerd moet worden. Attributen die altijd gecontroleerd worden, zullen in de volgende audit uiteraard gecontroleerd moeten worden en ook de attributen die m.b.v. een steekproef bepaald worden, maar die in de voorgaande twee externe audits nog niet aan bod zijn gekomen.

Tabel 1 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de klasse 'Algemeen' ('GeoObject'). Deze attributen komen voor in elk van de volgende objectklassen.

Attribuutnaam	Bronmateriaal			Controle			Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd	Steekproef	Sep 2011	Nov 2014	Nov 2017
bronactualiteit			lijst met unieke brontypen, bronbeschrijving, etc.			X	nee	nee	ja
bronbeschrijving			lijst met unieke brontypen, bronbeschrijving, etc.			X	nee	nee	ja
bronnauwkeurigheid			lijst met unieke brontypen, bronbeschrijving, etc.			X	nee	nee	ja
brontype			lijst met unieke brontypen, bronbeschrijving, etc.			X	nee	nee	ja
eindRegistratie				X			n.v.t.	n.v.t.	n.v.t.
lokaalID				X			n.v.t.	n.v.t.	n.v.t.
mutatieType				X			n.v.t.	n.v.t.	n.v.t.
namespace				X			n.v.t.	n.v.t.	n.v.t.
objectBeginTijd				X			n.v.t.	n.v.t.	n.v.t.
objectEindTijd				X			n.v.t.	n.v.t.	n.v.t.
tdnCode			Definitiebestand / sql-script			X	nee	nee	ja
tijdstipRegistratie				X			n.v.t.	n.v.t.	n.v.t.
visualisatieCode			definitiebestand			X	nee	nee	ja

De attributen 'eindRegistratie' en 'objectEindTijd' kunnen niet worden gecontroleerd, want deze worden niet uitgeleverd.

Voor de attributen 'lokaalID', 'mutatieType', 'objectBeginTijd' en 'tijdstipRegistratie' bestaat geen bronmateriaal om het attribuut te kunnen controleren.

De attributen 'tdnCode' en 'visualisatieCode' worden minder belangrijk geacht, omdat deze worden afgeleid van de andere attributen.

Het attribuut 'namespace' is voor elk object hetzelfde. De controle hiervan is onderdeel van het kwaliteitsaspect 'logische consistentie'. Ook de juiste notatie van 'objectBeginTijd' en 'tijdstipRegistratie' is onderdeel van het kwaliteitsaspect 'logische consistentie'.

De controle van de attributen kan, voor zover deze uitvoerbaar zijn, min of meer automatisch. Het eenmalig maken van de scripts kost wat tijd, met name voor het attribuut 'tdnCode', maar kunnen daarna eenvoudig uitgevoerd worden voor elke objectklasse. Dit kost dan verder weinig tijd.

De geschatte extra tijd t.o.v. de vorige externe audits is drie dagen.

Tabel 2 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de objectklasse 'Wegdeel'.

Attribuutnaam	Bronmateriaal			Controle		Steekproef	Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd		Sep 2011	Nov 2014	Nov 2017
aantalRijstroken	x	x				X	nee	ja	nee
afritnaam		x	RWS?			X	nee	nee	ja
afritnummer		x	RWS?			X	nee	nee	ja
A-wegnummer		x	RWS?			X	nee	nee	ja
brugnaam		x				X	nee	nee	ja
E-wegnummer		x	RWS?			X	nee	nee	ja
fysiekVoorkomen	x	x				X	ja	nee	nee
gescheidenRijbaan	x	x				X	nee	nee	ja
hoofdverkeersgebruik	x	x			X		ja	ja	ja
hoogteniveau	x	x				X	nee	ja	nee
isBAGnaam			BAG			X	n.v.t.	n.v.t.	ja
knooppuntnaam		x	RWS			X	nee	ja	nee
naam		x	BAG			X	n.v.t.	n.v.t.	nee
N-wegnummer		x	RWS?			X	nee	nee	ja
status	x	x				X	ja	nee	nee
S-wegnummer		x	RWS?			X	ja	nee	nee
tunnelnaam		x				X	ja	nee	nee
typeInfrastructuur			TOP10NL		X		ja	ja	ja
typeWeg	x	x	TOP10NL		X		ja	ja	ja
verhardingsbreedteklasse	x		objectcatalogus			X	ja	ja	nee
verhardingstype	x	x				X	nee	nee	ja

De waarde van het attribuut 'typeInfrastructuur' bepaalt mede de toegestane combinaties van geometrieën (en vice versa) en wordt daarom belangrijk geacht. Voor het attribuut 'verhardingsbreedteklasse' geldt dat niet altijd de werkelijke verhardingsbreedte de klasse bepaalt. Dit staat beschreven in de objectcatalogus van TOP10NL.

Het attribuut 'naam' hoeft in de audit van de data van november 2017 niet gecontroleerd te worden, omdat dit attribuut ontstaan is uit de eerdere attributen 'straatnaamFries' en 'straatnaamNL', welke beide al eerder zijn gecontroleerd.

De controle van de objectklasse 'wegdeel' kost naar verwachting één dag meer dan de vorige audits. Er moeten meer attributen gecontroleerd moeten worden (12 tegenover 9 en 10 in de vorige audits). De extra attributen (met name namen en nummers) zijn echter lang niet altijd gevuld en zijn optioneel, dus dit beperkt de hoeveelheid extra werk.

Tabel 3 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de objectklasse 'Spoorbaanddeel'.

Attribuutnaam	Bronmateriaal			Controle		Steekproef	Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd		Sep 2011	Nov 2014	Nov 2017
aantalSporen	x	x				x	ja	nee	nee
baanvaknaam			Prorail			x	nee	nee	ja
brugnaam		x	Prorail/BGT			x	nee	nee	ja
elektrificatie	x	x				x	ja	nee	nee
fysiekVoorkomen	x	x				x	nee	ja	nee
hoogteniveau	x	x				x	ja	ja	nee
spoorbreedte	x	x				x	nee	ja	nee
status	x	x				x	nee	ja	nee
tunnelnaam		x	Prorail/BGT			x	nee	nee	ja
typeInfrastructuur	x	x	TOP10NL			x	ja	ja	nee
typeSpoorbaan	x	x	Prorail/BGT/wettelijke status		x		ja	ja	ja
vervoerfunctie	x	x	?			x	nee	nee	ja

De waarde van het attribuut 'typeInfrastructuur' is direct af te leiden uit de geometrie (punt of lijn) van het object in TOP10NL en wordt daarom als minder belangrijk gezien.

De verwachting is dat de tijdsbesteding die nodig is voor de externe audit van de data van november 2017 niet heel anders zal zijn dan de vorige twee externe audits.

Tabel 4 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de objectklasse 'Waterdeel'.

Attribuutnaam	Bronmateriaal			Controle		Steekproef	Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd		Sep 2011	Nov 2014	Nov 2017
breedteklasse	x					X	nee	nee	ja
brugnaam		x	BGT			X	nee	ja	nee
functie	x	x				X	ja	nee	nee
fysiekVoorkomen	x	x				X	ja	nee	nee
getijdeInvloed			objectcatalogus			X	n.v.t.	n.v.t.	nee
hoofdafwatering			TOP10NL			X	nee	nee	ja
hoogteniveau	x	x				X	nee	nee	ja
isBAGnaam			BAG			X	n.v.t.	n.v.t.	ja
naamFries		x	BAG, lijst met Friese waternamen van Provincie Fryslân, Hydrografische Dienst kaarten			X	ja	nee	nee
naamNL		x	BAG, Hydrografische Dienst kaarten			X	nee	ja	nee
naamOfficieel		x	BAG, lijst met Friese waternamen van Provincie Fryslân, Hydrografische Dienst kaarten			X	n.v.t.	n.v.t.	ja
sluisnaam		x	BGT			X	nee	ja	nee
typeWater	x	x			X		ja	ja	ja
vaarwegklasse			RWS			X	n.v.t.	n.v.t.	ja
voorkomen	x	x				X	nee	nee	ja

In de objectcatalogus van TOP10NL staat beschreven welke objecten welke waarde krijgen voor het attribuut 'getijdeInvloed'. Dit attribuut heette eerder 'stroomrichting' en dat attribuut is in een van de eerdere audits al gecontroleerd, dus hoeft niet in de audit van de data van november 2017 meegenomen te worden.

De verwachting is dat de komende externe controle van de objectklasse 'waterdeel' iets meer tijd zal kosten t.o.v. de vorige externe controles, omdat het aantal te controleren attributen iets hoger is en omdat de controle van het attribuut 'breedteklasse' vrij veel tijd kost doordat de breedte van de objecten gemeten moet worden. De geschatte extra tijd t.o.v. de vorige externe audits is één dag.

Tabel 5 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de objectklasse 'Gebouw'.

Attribuutnaam	Bronmateriaal			Controle		Steekproef	Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd		Sep 2011	Nov 2014	Nov 2017
fysiekVoorkomen	x	x				X	n.v.t.	n.v.t.	ja
gebruiksdoel		x	BAG			X	n.v.t.	n.v.t.	ja
hoogte		x				X	nee	nee	ja
hoogteklasse		x				X	nee	nee	ja
hoogteniveau	x	x				X	ja	nee	nee
naam		x				X	n.v.t.	n.v.t.	ja
soortnaam	x	x				X	n.v.t.	n.v.t.	ja
status		x				X	nee	nee	ja
typeGebouw	x	x			X		ja	ja	ja

Het aantal te controleren attributen is in de audit van de data van november 2017 een stuk hoger dan bij de vorige externe audits. Dit komt door een aantal nieuwe attributen en doordat hoogte en hoogteklasse eerder als niet controleerbaar werden beschouwd, maar nu wel goed te controleren zijn. Hoogte meten kost relatief veel tijd, maar is zelden gevuld (in de TOP10NL-levering van juni 2016 staat slechts twee maal een echte waarde). De controle hiervan kost dus geen extra tijd, mits hier geen grote wijzigingen in komen. De extra attributen kosten naar schatting wel één dag extra controleren.

Tabel 6 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de objectklasse 'Terrein'.

Attribuutnaam	Bronmateriaal			Controle		Steekproef	Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd		Sep 2011	Nov 2014	Nov 2017
fysiekVoorkomen	x	x				X	ja	nee	nee
hoogteniveau	x	x				X	nee	ja	nee
naam		x				X	n.v.t.	n.v.t.	ja
typeLandgebruik	x	x			X		ja	ja	ja
voorkomen	x	x				X	nee	ja	nee

Doordat er in de audit van de data van november 2017 minder attributen gecontroleerd hoeven te worden dan in de vorige externe audits, zal de controle waarschijnlijk één dag minder kosten.

Tabel 7 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de objectklasse 'Inrichtingselement'.

Attribuutnaam	Bronmateriaal			Controle			Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd	Steekproef	Sep 2011	Nov 2014	Nov 2017
breedte	x					X	n.v.t.	n.v.t.	ja
hoogte			RD kernnetpunten			X	nee	nee	ja
hoogteniveau	x	x				X	ja	nee	nee
naam		x				X	n.v.t.	n.v.t.	ja
nummer		x	RD kernnetpunten			X	nee	ja	nee
soortnaam		x				X	n.v.t.	n.v.t.	ja
typeInrichtingselement	x	x			X		ja	ja	ja

Het meten van de hoogte kost relatief veel tijd, maar het aantal objecten in de steekproef waarvoor dit nodig is, zal waarschijnlijk beperkt zijn (naar verwachting hooguit enkele tientallen). De totaal te besteden tijd aan de controle van de objectklasse 'Inrichtingselement' wordt naar verwachting een halve dag meer dan bij de vorige externe audits.

Tabel 8 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de objectklasse 'Reliëf'.

Attribuutnaam	Bronmateriaal			Controle			Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd	Steekproef	Sep 2011	Nov 2014	Nov 2017
functie	x	x				X	nee	ja	nee
hoogteklasse		x				X	nee	nee	ja
hoogteniveau	x	x				X	nee	nee	ja
typeRelief	x	x			X		ja	ja	ja

Het meten van de hoogte t.b.v. het attribuut 'hoogteklasse' kost relatief veel tijd. Daardoor zal de totale tijdsbesteding voor deze controle iets hoger zijn dan bij de vorige externe audits. Geschat wordt dat dit één dag meer kost.

Tabel 9 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de objectklasse 'Hoogte'.

Attribuutnaam	Bronmateriaal			Controle			Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd	Steekproef	Sep 2011	Nov 2014	Nov 2017
hoogte		x	AHN		X		nee	nee	ja
referentievlak				X			n.v.t.	n.v.t.	n.v.t.
typeHoogte	x				X		ja	ja	ja

Voor het attribuut 'referentievlak' bestaat geen bronmateriaal om het te kunnen controleren.

T.o.v. de vorige externe audits is er een attribuut minder om te controleren, maar wel moet de hoogte gemeten worden. Dat kost relatief veel tijd. De totaal benodigde tijd voor de controle van de objectklasse 'Hoogte' blijft naar verwachting ongeveer gelijk aan de tijd die nodig was in de vorige externe audits.

Tabel 10 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de objectklasse 'RegistratiefGebied'.

Attribuutnaam	Bronmateriaal			Controle			Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd	Steekproef	Sep 2011	Nov 2014	Nov 2017
naamFries			BRK			X	nee	nee	ja
naamNL			BRK			X	nee	nee	ja
naamOfficieel			BRK		X		n.v.t.	n.v.t.	ja
nummer			BRK/CBS			X	nee	nee	ja
typeRegistratiefGebied			BRK		X		ja	ja	ja

Het vorige controleprotocol ging enkel uit van orthofoto's en cyclorama's als bronmateriaal. Het controleren van de namen was daarmee niet mogelijk. M.b.v. externe bronnen kan de controle van namen wel. Daardoor zijn er in de audit van de data van november 2017 meer attributen te controleren en zal dit meer tijd kosten t.o.v. de vorige externe audits. Dit tijd gaat met name zitten in het verkrijgen en verwerken van de gegevens uit de BRK en het vergelijken hiermee. Het betreft echter een relatief laag aantal objecten, waardoor de extra benodigde tijd geschat wordt op anderhalve dag.

Tabel 11 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de objectklasse 'GeografischGebied'.

Attribuutnaam	Bronmateriaal			Controle			Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd	Steekproef	Sep 2011	Nov 2014	Nov 2017
naamFries		x				X	ja	nee	nee
naamNL		x				X	nee	nee	ja
typeGeografischGebied	x	x			X		ja	ja	ja

Met betrekking tot de attributen 'naamFries' en 'naamNL' dient gezegd te worden dat deze namen slechts in enkele gevallen op cyclorama's te zien zullen zijn. Gezien het beperkte belang van de attributen gaat het te ver om alle cyclorama's in het gebied te onderzoeken op het voorkomen van de naam. Enkel de cyclorama's op voor de hand liggende locaties (zoals voor bezoekers interessante plaatsen als parkeerplaatsen en picknickplekken) zullen bekeken worden.

De verwachte tijdsduur voor de controle van 'GeografischGebied' blijft gelijk aan de tijd die nodig was in de vorige externe audits.

Tabel 12 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de objectklasse 'Plaats'.

Attribuutnaam	Bronmateriaal			Controle			Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd	Steekproef	Sep 2011	Nov 2014	Nov 2017
aantalInwoners			Document met door CBS berekende aantallen			X	nee	nee	ja
bebouwdeKom	x	x	BAG			X	n.v.t.	n.v.t.	ja
isBAGwoonplaats			BAG			X	n.v.t.	n.v.t.	ja
naamFries		x	BAG			X	ja	nee	nee
naamNL		x	BAG			X	nee	nee	ja
naamOfficieel		x	BAG		X		n.v.t.	n.v.t.	ja
typeGebied	x	x			X		ja	ja	ja

De objectklasse 'Plaats' is nieuw in TOP10NL, maar plaatsen kwamen eerder al voor als onderdeel van 'GeografischGebied'. Het attribuut 'typeGebied' is bij vorige externe audits al gecontroleerd als 'typeGeografischGebied'.

Door de nieuwe attributen waarbij externe bronnen geraadpleegd moeten worden, kost de controle naar verwachting behoorlijk wat extra tijd t.o.v. de vorige externe audits. Bovendien kent 'typeGebied' meer waarden, wat ook extra tijd kost. Het aantal objecten zal echter beperkt zijn. Geschat wordt dat de controle anderhalve dag extra kost.

Tabel 13 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de objectklasse 'FunctioneelGebied'.

Attribuutnaam	Bronmateriaal			Controle			Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd	Steekproef	Sep 2011	Nov 2014	Nov 2017
naamFries		x				X	nee	ja	nee
naamNL		x				X	ja	nee	nee
soortnaam		x	diverse externe bronnen (campings, mosselbanken, etc.)			X	n.v.t.	n.v.t.	ja
typeFunctioneelGebied	x	x			X		ja	ja	ja

De tijdsbesteding voor de controle van 'FunctioneelGebied' zal naar verwachting niet anders zijn dan in eerdere externe audits.

Tabel 14 Selectie van attributen voor controle op de thematische nauwkeurigheid voor de objectklasse 'Plantopografie'.

Attribuutnaam	Bronmateriaal			Controle			Audit		
	Orthofoto	Cyclorama	Anders	Nooit	Altijd	Steekproef	Sep 2011	Nov 2014	Nov 2017
Naam			Externe bronnen (RWS, ruimtelijke plannen, ...)			X	n.v.t.	n.v.t.	ja
typeObject			Externe bronnen (RWS, ruimtelijke plannen, ...)		X		n.v.t.	n.v.t.	ja

Dit betreft een nieuwe objectklasse met nieuwe attributen. De controle hiervan kost dus extra tijd t.o.v. de vorige externe audits. Vanwege het beperkte aantal objecten en attributen valt de extra tijd erg mee. Geschat wordt dat er een halve dag extra voor nodig is.

De wijziging van het gegevensmodel van TOP10NL heeft als gevolg dat er meer tijd besteed zal moeten worden aan de bepaling van de thematische nauwkeurigheid. In de tabel hieronder wordt per objectklasse de hoeveelheid meer of minder werk aangegeven t.o.v. de vorige audits bij een gelijkblijvende steekproef. In totaal zal de externe audit naar verwachting ongeveer tien dagen extra kosten.

Tabel 15 De hoeveelheid meer of minder werk t.o.v. vorige audits.

Objectklasse	Geschat verschil t.o.v. vorige audits (in dagen)
Algemeen	3
Wegdeel	1
Spoorbaanddeel	0
Waterdeel	1
Gebouw	1
Terrein	-1
Inrichtingselement	0.5
Relief	1
Hoogte	0
Registratief gebied	1.5
Geografisch gebied	0
Plaats	1.5
Functioneel gebied	0
Plantopografie	0.5
Totaal	10

Een wijziging van de steekproef, bijvoorbeeld n.a.v. de uitkomsten van deelonderzoek 2 (zie het volgende hoofdstuk), kan effect hebben op de hoeveelheid werk. Dit effect is in bovenstaand overzicht niet meegenomen.

3 Deelonderzoek 2: Aanpassing van de huidige methode om te zorgen voor voldoende objecten in de steekproef voor elke objectklasse

Een van de aanbevelingen uit het rapport van de laatste externe audit (Storm et al. 2015) is om bij toekomstige controles de steekproef zodanig op te zetten dat het minimum aantal objecten per objectklasse groter is. Dit i.v.m. een mogelijk vertekend beeld van de kwaliteit indien het aantal objecten laag is. Dit speelde bij de laatste audit bij de objectklassen 'functionele gebieden', 'geografische gebieden' en 'spoorbaandelen'.

Omdat veel functionele gebieden nu vlakvormig worden (i.p.v. puntvormig), is de verwachting dat het aantal functionele gebieden bij een volgende audit voldoende zal zijn. Geografische gebieden en spoorbaandelen hebben bij de huidige methode in een volgende audit waarschijnlijk weer een laag aantal objecten.

De vraag vanuit het Kadaster is om uit te zoeken wat een relevant minimum aantal objecten in een steekproef is om betrouwbare gegevens te krijgen over de kwaliteit van de BRT en hoe de steekproef dan eventueel aangepast moet worden.

In dit onderzoek wordt ervan uitgegaan dat het gebied met steekproefeenheden begrensd wordt door het gebied waarvan luchtfoto's (en ook cyclorama's) bestaan. In een eerdere release van TOP10NL kwamen objecten voor ver op de Noordzee, buiten het gebied waarvan luchtfoto's of cyclorama's beschikbaar zijn. Het ging om de Natura 2000-gebieden 'Doggersbank', 'Klaverbank' en 'Friese Front' die als Functioneel Gebied zijn opgeslagen. Inmiddels zijn deze geen onderdeel meer van TOP10NL. Mochten er in de toekomst wel weer objecten ver op de Noordzee worden opgenomen in TOP10NL, dan kan het nodig zijn om het gebied met steekproefeenheden uit te breiden. Uitbreiding van het gebied met steekproefeenheden zou leiden tot een hoger aantal 'lege' steekproefeenheden, waarmee het onderscheidend vermogen, zoals beschreven aan het einde van dit hoofdstuk, lager zou uitvallen.

3.1 Inleiding

De resultaten van een onderzoek naar de kwaliteit zijn betrouwbaar als:

- in werkelijkheid aan de norm wordt voldaan en de resultaten ook aangeven dat aan de norm wordt voldaan, en
- als in werkelijkheid niet aan de norm wordt voldaan, ook de resultaten aangeven dat niet aan de norm wordt voldaan.

De betrouwbaarheid kan worden aangegeven met de kans op een foute conclusie. Dit is het geval als:

- er onterecht geconcludeerd wordt dat niet aan de norm wordt voldaan (fout van de eerste soort), en als
- er onterecht geconcludeerd wordt dat wel aan de norm wordt voldaan (fout van de tweede soort).

Er is door het Kadaster niet kwantitatief vastgesteld welke betrouwbaarheid de resultaten van de kwaliteitscontrole moeten hebben. Het resultaat van dit onderzoek geeft dus niet een exact aantal minimum objecten dat nodig is in een steekproef. Wel geeft dit onderzoek inzicht in het effect van de steekproefomvang en het aantal objecten in een steekproef op de betrouwbaarheid. Dit inzicht kan helpen bij het bepalen van een steekproefomvang waarbij de betrouwbaarheid van de resultaten acceptabel is.

3.2 Onderzoek

Het aantal objecten in de steekproef wordt o.a. bepaald door het aantal geselecteerde steekproefgebieden van 500m x 500m.

Dit onderzoek omvat een simulatiestudie op basis van de gegevens in TOP10NL en op basis van de resultaten van de twee externe audits. De simulatiestudie is enkel uitgevoerd op de objectklassen Spoorbaandelen en Geografische gebieden, omdat dit objectklassen zijn met tot nu toe weinig objecten in de steekproef van de externe audits en daarmee een lagere betrouwbaarheid van de resultaten.

Voor alle voorkomende objecten van deze objectklassen is een indicator gesimuleerd die aangeeft of het object in het bestand wel of niet juist is. Deze indicator is gesimuleerd door te trekken uit een Bernoulli-verdeling met een 'kans op succes' van 0,05. De kans op succes moet hier worden geïnterpreteerd als de 'kans dat een object als onjuist wordt aangegeven'. Dit is gebaseerd op de norm van 95% juist die het Kadaster heeft vastgesteld in het Kwaliteitshandvest (Kadaster, 2016).

Vervolgens zijn steekproeven geloot met een omvang van 50, 100, 200, 300 400 en 500 eenheden. De steekprofeenheden zijn de gebieden van 500m x 500m. De steekproeven zijn getrokken volgens de methode zoals beschreven in de huidige controlemethodiek (Storm et al. 2015), dus met een gestratificeerde steekproefopzet, met provincies als strata. Het aantal gelote steekprofeenheden in een stratum is evenredig met het totaal aantal steekprofeenheden van dat stratum (evenredige allocatie). Binnen strata zijn de gebieden geloot met kansen evenredig aan het totaal aantal objecten van alle objectklassen binnen de gebieden, zonder teruglegging.

In onderstaande figuren is te zien hoeveel objecten van de objectklassen Spoorbaandeel en Geografisch Gebied er zitten in de gelote steekprofeenheden bij verschillende steekproefomvang (50, 100, 200, 300 400 en 500 eenheden). De boxplots geven het resultaat van de 500 simulaties voor elke genoemde steekproefomvang, waarbij in elke simulatie de indicator van de objecten die aangeeft of een object wel of niet juist is, opnieuw is bepaald.

In de boxplots is duidelijk te zien dat het aantal objecten toeneemt bij een grotere steekproefomvang. Ook is te zien dat bij eenzelfde steekproefomvang het aantal spoorbaandelen lager is dan het aantal geografische gebieden. Daarnaast is ook een grote spreiding te zien in het aantal objecten. Bij een steekproefomvang van 500 gebieden kan het aantal spoorbaandelen variëren tussen ongeveer 20 en 55 objecten en het aantal geografische gebieden tussen ongeveer 50 en 350.

Figuur 1 Het aantal spoorbaandeel-objecten (links) en geografisch gebied-objecten (rechts) bij een steekproefomvang van 50, 100, 200, 300, 400 en 500 gebieden voor 500 simulaties per steekproefomvang.

Gegeven een steekproef is het totaal aantal onjuist weergegeven objecten geschat en het totaal aantal objecten (ongeacht of deze goed of fout zijn weergegeven). De ratio van deze twee schattingen is een schatting van de populatiefractie onjuist weergegeven objecten. Voor iedere steekproefomvang is de steekproefname en schatting 500 keer herhaald. Dit levert per steekproefomvang 500 schattingen van de fractie onjuist weergegeven objecten op. Zie figuur 2.

In deze figuur is duidelijk te zien dat bij een kleine steekproefomvang (van 50 gebieden) de geschatte fractie onjuist weergegeven objecten kan variëren van 0 tot 1. M.a.w.: De conclusie van een kwaliteitscontrole kan bij deze steekproefomvang luiden dat alles juist is, maar ook dat alles onjuist is, terwijl in werkelijkheid 95% juist is. Het groter de steekproefomvang, hoe kleiner de spreiding in de waarde van de geschatte fractie onjuiste objecten. Uit deze simulatiestudie blijkt dat met name het aantal uitschieters naar boven (die dus een grote fractie onjuiste objecten suggereren) minder wordt bij een grotere steekproefomvang.

Figuur 2 De geschatte fractie onjuiste spoorbaandeel-objecten (links) en geografisch gebied-objecten (rechts) bij een steekproefomvang van 50, 100, 200, 300, 400 en 500 gebieden voor 500 simulaties per steekproefomvang.

De kans op een onjuiste conclusie in een kwaliteitsonderzoek lijkt dus kleiner te worden naarmate de steekproefomvang toeneemt. Zoals in het begin al aangegeven, bestaan er twee foute conclusies:

- onterecht concluderen dat niet aan de norm wordt voldaan (fout van de eerste soort);
- onterecht concluderen dat wel aan de norm wordt voldaan (fout van de tweede soort).

De kansen op deze twee fouten worden aangeduid met de Griekse letters α en β . α wordt het *significantieniveau* van een toets genoemd, $1-\beta$ wordt aangeduid met het *onderscheidend vermogen* van de toets. α wordt vooraf gekozen; in dit geval 0,05, 0,10 en 0,20. Het onderscheidend vermogen is allereerst afhankelijk van de werkelijke fractie onjuist weergegeven objecten. Des te groter deze fractie, des te groter het onderscheidend vermogen. Het onderscheidend vermogen hangt ook af van de gekozen α . Des te kleiner α (kans op fout van eerste soort), des te groter β (kans op fout van tweede soort), m.a.w. des te kleiner het onderscheidend vermogen. Verder hangt het onderscheidend vermogen af van de steekproefomvang. Des te groter de omvang, des te groter het onderscheidend vermogen. De vraag naar de benodigde steekproefomvang kan nu nader gespecificeerd worden als: Hoeveel eenheden zijn nodig om, gegeven een α , een gewenst onderscheidend vermogen te realiseren? Omdat het lastig kan zijn een keuze te maken voor een bepaalde α en β , hebben we voor drie waarden van α grafieken gemaakt van het onderscheidend vermogen als functie van de steekproefomvang. In deze berekening is voor de werkelijke fractie onjuist weergegeven objecten gekozen voor 10%: het dubbele van wat de norm van het Kwaliteitshandvest toelaat.

Het onderscheidend vermogen zoals in onderstaande grafieken is weergegeven, is het gemiddelde van 500 schattingen bij een steekproefomvang van 50, 100, 200, 300, 400 en 500 gebieden.

Figuur 3 Het onderscheidend vermogen als functie van de steekproefomvang en het significantieniveau alfa, voor spoorbaandelen (links) en geografische gebieden (rechts).

Het onderscheidend vermogen neemt toe met de steekproefomvang. Ook is duidelijk het effect van alfa te zien: des te groter alfa (het significantieniveau), des te groter het onderscheidend vermogen. Voor een betrouwbaar resultaat wil je een zo klein mogelijke alfa en een zo groot mogelijk onderscheidend vermogen ($1-\beta$).

In het algemeen kan gezegd worden dat het onderscheidend vermogen laag is. Bij een alfa van 0,10 is voor spoorbaandelen een steekproefomvang van 500 nodig om een onderscheidend vermogen van 0,80 te realiseren. Voor geografische gebieden is het onderscheidend vermogen nog veel lager. Het lage onderscheidend vermogen wordt veroorzaakt door het kleine aantal objecten dat wordt geselecteerd. In veel steekproefeenheden komen geen spoorbaandelen voor, wat als gevolg heeft dat slechts een beperkt aantal spoorbaandelen voorkomt binnen de steekproef. Voor geografische gebieden is het aandeel lege steekproefeenheden nog groter, waardoor nog minder objecten van dit type worden geselecteerd.

3.3 Alternatieve steekproefopzet

Een hoger onderscheidend vermogen wordt mogelijk indien de steekproef een groter deel steekproefeenheden mét spoorbaandeel-objecten of geografisch gebied-objecten bevat. Dit is te bereiken door een andere steekproefopzet. Hierin zijn twee alternatieven denkbaar, namelijk het combineren van twee steekproeven of het vergroten van de steekproefeenheden.

3.3.1 Alternatief 1: gecombineerde steekproef

Het eerste alternatief bestaat uit twee steekproeven die gecombineerd worden. Een van de steekproeven richt zich daarbij op de objectklasse die weinig voorkomt, dus op spoorbaandelen of geografische gebieden. Omdat de steekproeven 'area-based' zijn, is de spreiding van de steekproefgebieden met spoorbaandelen en geografische gebieden over Nederland van belang. In onderstaande figuur is duidelijk het verschil in spreiding te zien tussen steekproefgebieden met spoorbaandelen (links) en met geografische gebieden (rechts).

Figuur 4 De spreiding van steekproefgebieden met een of meer spoorbaandelen (links) en van steekproefgebieden met een of meer geografische gebieden (rechts).

De hieronder beschreven varianten van alternatief 1 zijn redelijk goed uitvoerbaar voor spoorbaandelen, omdat van spoorbaandelen goed af te bakenen is waar deze liggen en omdat eventuele nieuwe spoorbaandelen over het algemeen zullen aansluiten op reeds bestaande spoorbaandelen. Voor geografische gebieden zijn de varianten echter niet goed toe te passen, omdat deze overal in het land kunnen voorkomen. Geografische gebieden zullen daarom net als elke andere veel voorkomende objectklasse worden behandeld.

Alternatief 1 kent twee varianten (*a* en *b*). De verschillen zitten in het gebied dat in de twee steekproeven wordt gebruikt en in de te controleren objectklassen in die steekproeven.

Bij variant *a* wordt Nederland opgedeeld in twee delen, namelijk het deel van Nederland waarin bekend of aannemelijk is dat daar spoorbaandelen voorkomen (zie figuur 4) en het deel van Nederland waarin bekend of aannemelijk is dat daar geen spoorbaandelen voorkomen (dus de rest van Nederland).

De steekproef op het deel van Nederland waarin bekend of aannemelijk is dat daar spoorbaandelen voorkomen kent een relatief hoog onderscheidend vermogen. Die steekproef is echter niet geschikt om een kwantitatief oordeel te geven over omissies, want daarvoor zal heel Nederland in de steekproef meegenomen moeten worden. Er is dus nog een tweede aanvullende steekproef nodig waarin gebieden uit het andere deel van Nederland worden geselecteerd.

Er is nog wel een aandachtspunt bij deze variant. Een klein aantal objecten per stratum kan namelijk een negatief effect hebben op de betrouwbaarheid van de resultaten. Het aantal strata neemt toe van 12 naar 24. In de tot nu gebruikte methodiek vormden de provincies de strata. In variant *a* kent elke provincie twee strata, namelijk het deel met spoorbaandelen en het deel zonder spoorbaandelen. De kans op een klein aantal objecten per stratum wordt daarmee vergroot.

Een nadeel van variant *a* is dat de uitkomst van een audit voor een belangrijk deel door het Kadaster beïnvloedt zou kunnen worden door te zorgen dat de kwaliteit van de gebieden waarin spoorbaandelen(/geografische gebieden) voorkomen op orde is. Uiteraard is dit afhankelijk van de verhouding van het aantal gebieden in de ene steekproef en in de andere steekproef.

Variante *b* kent dit nadeel niet. Deze verschilt van variant *a* in die zin dat de steekproef op het deel van Nederland zonder spoorbaandelen vervangen wordt door een steekproef op heel Nederland. Daarnaast

wordt de steekproef die de gebieden bevat waarin bekend of aannemelijk is dat daar spoorbaandelen voorkomen alleen gebruikt voor spoorbaandelen. De resultaten van de audit op die gebieden heeft dus geen invloed op de controleresultaten van andere objectklassen.

Bij variant b is het wel van belang om te zorgen dat de tweede steekproef (met enkel de gebieden met spoorbaandelen) niet dezelfde steekproefgebieden oplevert als de eerste steekproef (over heel Nederland), om te voorkomen dat een zelfde gebied twee maal gecontroleerd zou moeten worden.

In onderstaande tabel staan de verschillen tussen varianten a en b.

Tabel 16 De verschillen tussen de varianten a en b van alternatief 1.

	Eerste steekproef Gebied	Te controleren objectklassen	Tweede steekproef Gebied	Te controleren objectklassen
Variant a	Deel van Nederland zonder spoorbaandelen	Alle	Deel van Nederland met spoorbaandelen	Alle
Variant b	Heel Nederland	Alle	Deel van Nederland met spoorbaandelen	Spoorbaandelen

Het wordt aanbevolen om voor de volgende externe audits te kiezen voor variant b, welke bestaat uit:

- Een steekproef van een nader te bepalen aantal gebieden in heel Nederland.
- Een steekproef van een nader te bepalen aantal gebieden waarin (naar alle waarschijnlijkheid) spoorbaandelen voorkomen.

De steekproeven worden op eenzelfde wijze worden uitgevoerd als bij de vorige audits, dus gestratificeerd naar provincie, waarbij de kans dat een gebied geloot wordt, bepaald wordt door het aantal objecten in het gebied (hoe groter het aantal objecten in het gebied, hoe groter de kans om geloot te worden).

Doordat alternatief 1 bestaat uit twee steekproeven, zal dit iets meer werk vereisen t.o.v. de tot nu gebruikte steekproefopzet.

De hier beschreven steekproef opzet verandert niets m.b.t. het aantal geselecteerde geografische gebieden t.o.v. de tot nu toe gebruikte steekproefopzet, maar het aantal spoorbaandelen in de steekproef zal wel toenemen. De betrouwbaarheid van de resultaten van een audit op de kwaliteit van TOP10NL zal op deze manier toenemen voor de objectklasse Spoorbaandelen en min of meer gelijk blijven voor de overige objectklassen.

Uiteraard is het ook mogelijk om voor de aanpak te kiezen die gebruikt is in de vorige externe audit, waarbij voor de objectklassen met weinig objecten in de steekproef (spoorbaandelen en geografische gebieden) de steekproef wordt uitgebreid met een van tevoren vastgesteld aantal gebieden. Hierbij bestaat echter de kans dat in die extra gebieden geen of nog steeds een heel beperkt aantal spoorbaandelen of geografische gebieden voorkomen.

3.3.2 Alternatief 2: vergroten van de steekprofeenheden

Om te zorgen voor minder 'lege' steekprofeenheden zonder objecten en daarmee een hoger onderscheidend vermogen, zou er ook voor gekozen kunnen worden om de steekproefgebieden te vergroten naar bijvoorbeeld 1000m x 1000m. Om de controle in een zelfde tijd te voltooien, moet de steekproefomvang omlaag, wat het onderscheidend vermogen weer laat zakken. Bijkomend nadeel is dat de spreiding van de geselecteerde gebieden minder is en daarmee wellicht ook minder representatief voor de variatie die Nederland kent. Daarom wordt dit alternatief niet aangeraden.

3.4 Advies

Om te bepalen hoe groot de steekproefomvang moet zijn voor acceptabele betrouwbaarheid van auditresultaten, kan gekeken worden naar de grafieken in figuur 3. Voor een betrouwbaar resultaat wil je een zo klein mogelijk significantieniveau (alfa) en een zo groot mogelijk onderscheidend vermogen (1-bèta). Het lastige hierbij is dat het onderscheidend vermogen daalt bij een lager significantieniveau.

Andere grafieken die kunnen helpen bij de bepaling van de gewenste steekproefomvang staan in figuur 2. Een grotere steekproefomvang leidt tot een lagere spreiding van het geschatte aantal onjuiste objecten. De spreiding is een maat voor betrouwbaarheid. Hoe kleiner de spreiding, hoe groter de betrouwbaarheid.

Zowel in figuur 2 als in figuur 3 is te zien dat een grotere steekproefomvang de betrouwbaarheid toeneemt. Er is door het Kadaster van tevoren niet aangegeven wat het significantieniveau of onderscheidend vermogen moet zijn en daarom kan in dit advies geen specifieke steekproefomvang voor de audit genoemd worden. De keuze voor de steekproefomvang is aan het Kadaster zelf. In deze keuze zal ook het kostenaspect meewegen. Een grotere steekproefomvang leidt tot meer werk en dus hogere kosten.

In de laatste audit bestond de steekproef voor spoorbaandelen en geografische gebieden uit 77 gebieden (met 51 spoorbaandelen en 10 geografische gebieden) en in de eerste externe audit uit 60 (met 8 spoorbaandelen en 19 functionele gebieden).

Op basis van het aantal objecten per steekproefgebieden en de benodigde controletijd per object is een heel globale schatting te maken van de benodigde extra tijd per steekproefgebied.

Voor spoorbaandelen is de schatting dat (elke) 20 extra controlegebieden in de steekproef leiden tot ongeveer een uur extra werk.

Voor geografische gebieden is de schatting dat (elke) 125 extra controlegebieden in de steekproef leiden tot ongeveer een uur extra werk.

Van de genoemde alternatieve steekproefopzetten is enkel het eerste alternatief het overwegen waard. Het tweede alternatief, waarbij de steekproefgebieden worden vergroot, levert eigenlijk geen enkel voordeel op t.o.v. de tot nu toe gebruikte steekproefopzet.

Het eerste alternatief t.o.v. de huidige steekproefopzet heeft als voordeel dat het, voor de objectklasse Spoorbaandelen, tot een hoger onderscheidend vermogen leidt en daarmee tot een hogere betrouwbaarheid van de auditresultaten voor deze objectklasse. Variant *a* van alternatief 1 kent echter een aantal (mogelijke) nadelen:

- De uitkomst van een audit zou voor een belangrijk deel door het Kadaster beïnvloedt kunnen worden door te zorgen dat de kwaliteit van de gebieden waarin spoorbaandelen voorkomen op orde is, zeker indien het aantal gebieden in de steekproef over het deel van Nederland met spoorbaandelen relatief hoog is ten opzichte van het aantal gebieden in de steekproef over de rest van Nederland.
- Als gevolg van de toename van de strata (namelijk van 12 naar 24), zal er bij een gelijkblijvende omvang van de steekproef t.o.v. de vorige externe audits een lager aantal objecten per stratum zijn. Dit kan een negatief effect hebben op de betrouwbaarheid van de resultaten. Nader onderzoek kan uitwijzen bij welke steekproefaantallen dit negatieve effect groter is dan het positieve effect van het toepassen van de gecombineerde steekproefopzet.

Variant *b* kent deze mogelijke nadelen niet, dus is het aan te bevelen om daarvoor te kiezen.

Voor zowel variant *a* als voor variant *b* geldt dat er twee steekproeven dienen te worden genomen, terwijl in de huidige opzet een enkele steekproef volstaat. Dit betekent iets meer werk voor de statisticus die de steekproeven neemt.

4 Deelonderzoek 3: Uitbreiding van de huidige methode met controlemethodieken voor de kleinere schalen

In de eerstvolgende externe controle dienen ook de kleinschalige producten meegenomen worden. De huidige methode is daar nog niet op voorbereid en zal dus aangepast moeten worden. Het Kadaster ziet graag een uitwerking van een methode voor de producten op schaal 1:50.000 welke vervolgens ook toe te passen is voor de andere kleinschalige producten. Omdat er voor de GML (TOP50NL) en het rasterbestand (TOP50raster) andere gebruikstoepassingen zijn, zullen er andere kwaliteitscriteria gelden voor deze producten.

Het onderzoek moet op basis van literatuur duidelijkheid geven over welke criteria van toepassing kunnen zijn en welke onafhankelijk gecontroleerd of gemeten kunnen worden. Uiteraard wordt er ook gekeken naar de controles die het Kadaster op dit moment al uitvoert, zoals de controle op de plaatsing van teksten in TOPraster-bestanden.

Het doel van dit onderzoek is niet om te komen tot criteria waarmee te bepalen is of iets goed of fout is. Dit kan op een later moment worden bepaald. De implementatie van de generalisatieregels zal ook niet worden gecontroleerd. Het gaat erom te onderzoeken of de gebruikte generalisatiemethode tot het gewenste resultaat leidt.

Dit deelonderzoek omvatte de volgende onderdelen:

- Literatuuronderzoek naar welke criteria er van toepassing kunnen zijn op kleinschalige digitale kaartproducten;
- Literatuuronderzoek naar welke methoden er bestaan om op onafhankelijke wijze de bovengenoemde criteria te toetsen;
- Onderzoek naar de haalbaarheid en bruikbaarheid van de verschillende methoden om toegepast te worden op TOP25raster, TOP50raster en TOP50NL.

Om de kwaliteit van een product te kunnen beoordelen, zul je van tevoren moeten bepalen wat de criteria zijn waartegen je het product beoordeelt. Dit geldt ook voor de kleinschalige kaartproducten van het Kadaster. Omdat het Kadaster hiervoor geen kwaliteitscriteria heeft vastgelegd, is in de eerste stap van dit deelonderzoek gezocht naar mogelijk toe te passen criteria.

De huidige controlemethodiek kijkt naar de volgende aspecten, die ook genoemd worden in de richtlijn ISO19157:2013 Geografische Informatie Data kwaliteit.

Voor TOP10NL als geheel:

- Actualiteit

Per objectklasse in TOP10NL:

- Logische consistentie
- Positionele nauwkeurigheid
- Volledigheid (omissie en commissie)
- Thematische nauwkeurigheid

In dit onderzoek is getracht om hierbij aansluiting te vinden, hoewel er vele andere indelingen voor kwaliteitsaspecten bedacht, zoals bijvoorbeeld die indeling gebruikt door Lonskiy en Hopfstock (2015). Zij bieden een hiërarchisch model waarin de criteria zijn in te delen. Zij onderscheiden op het hoogste niveau zeven componenten van kaartkwaliteit, namelijk: mathematische basis, volledigheid, betrouwbaarheid, geometrische nauwkeurigheid, cartografisch ontwerp en wetenschappelijke en sociale waarde. Met name het onderwerp 'cartografisch ontwerp' is voor kleinschalige kaartproducten interessant. Dit aspect is voor een belangrijk deel onder te brengen in het aspect 'bruikbaarheid' van ISO19157 en zal in dit onderzoek meegenomen worden.

In de volgende paragrafen wordt per kwaliteitsaspect uit ISO19157 beschreven wat ermee bedoeld wordt, welke methoden er in de literatuur zijn gevonden om dit kwaliteitsaspect te meten en of dit is toe te passen op de kleinschalige producten van de BRT.

4.1 Actualiteit

4.1.1 Beschrijving

De Basisregistratie Topografie dient een actualiteit van twee jaar te hebben. Dit betekent dat de producten van de BRT gebaseerd moeten zijn op informatie van maximaal twee jaar oud. Alle kleinschalige producten van de BRT worden gebaseerd op TOP10NL. TOP10NL wordt bijgehouden m.b.v. luchtfoto's. Uiteindelijk zijn de opnamedata van de luchtfoto's bepalend voor de actualiteit van de kleinschalige producten.

4.1.2 Methode

Het Kadaster publiceert op haar website (<https://www.kadaster.nl/-/top10nl>) vijf keer per jaar (na elke levering) een BRT Actualiteitskaart. Op deze kaart wordt voor elk product van de BRT aangegeven van wanneer de luchtfoto is.

Het kwaliteitsaspect 'actualiteit' van TOP10NL wordt gerapporteerd als het percentage kaartbladen dat voldoet aan de eis van 2-jaren-actualiteit. Met kaartblad wordt daarbij een in noord- en zuidversie opgedeeld 1:25.000-kaartblad bedoeld.

Voor de kleinere schalen is een vergelijkbare methode te hanteren, want op de BRT Actualiteitskaart is de kaartbladindeling voor de verschillende schalen zichtbaar. De actualiteit van elk product kan dan gerapporteerd worden als het percentage kaartbladen dat voldoet aan de 2-jaren-actualiteit.

Voorgesteld wordt om voor de berekening van de actualiteit de kaartbladindeling van het betreffende product te gebruiken. Voor de berekening van de actualiteit van TOP25raster worden dan de (hele) 1:25.000-kaartbladen gebruikt en voor TOP50NL en TOP50raster de 1:50.000-kaartbladen.

4.1.3 Toepasbaarheid

De beschreven methode is specifiek voor de BRT-producten. Daarmee is het te gebruiken voor alle BRT-producten, mits deze vermeld worden op de BRT Actualiteitskaart en deze actualiteitskaart beschikbaar en vindbaar is.

4.2 Logische consistentie

4.2.1 Beschrijving

Logische consistentie is de mate van overeenstemming met regels voor de gegevensstructuur, voor de toekenning van attributen en voor relaties (naar Jakobsson en Giversen, z.j.). Logische consistentie kent een aantal subelementen, zoals conceptuele consistentie, domeinconsistentie en logische consistentie.

4.2.2 Methode

Een controle op de logische consistentie wordt gebaseerd op de regels die gelden voor het desbetreffende product. Een topologische regel kan bijvoorbeeld zijn: alle objecten met hoogteniveau '0' van de objectklassen 'terrein', 'waterdeel' en 'wegdeel' vormen samen een landsdekkend geheel.

Een regel die betrekking heeft op de conceptuele consistentie is bijvoorbeeld 'Een object van de objectklasse 'Gebouw' kan enkel als puntobject voorkomen'. De domeinconsistentie kan bepaald worden m.b.v. de gegevenscatalogus of een xsd-bestand.

4.2.3 Toepasbaarheid

Logische consistentie is een kwaliteitsaspect dat enkel van belang is voor de objectgerichte producten en niet op de rasterproducten.

Indien de betreffende regels voor een bepaald product beschikbaar zijn, kunnen deze gebruikt worden om een controle op uit te voeren. Dit geldt voor elk objectgerichte product. Een aantal regels kan m.b.v. scripts of modellen geautomatiseerd uitgevoerd worden, maar er zullen ook regels zijn die op handmatige wijze gecontroleerd moeten worden.

4.3 Positionele nauwkeurigheid

4.3.1 Beschrijving

Hierbij gaat het om de nauwkeurigheid van de positie van de weergegeven gegevens op de kaart. Dit is voor veel toepassingen en gebruikers van belang. Informatie hierover dient dan ook beschikbaar te zijn, of op zijn mist opvraagbaar bij de makers. Deze informatie is vaak gerelateerd aan de manier waarop brondata zijn ingewonnen en vervolgens verwerkt zijn. Dit kan opgeslagen worden in attribuutinformatie, in metadata of in een productbeschrijving. En waar mogelijk kan dit gekwantificeerd worden met een maat van nauwkeurigheid.

De geometrische nauwkeurigheid kan bijvoorbeeld als volgt worden vastgelegd: "De maximale geometrische afwijking is 50 meter t.o.v. de referentie." Als referentie kan bijvoorbeeld de werkelijkheid, de luchtfoto of TOP10NL worden bedoeld. Dit moet uiteraard wel duidelijk worden vastgelegd.

4.3.2 Methode

De metadata of productbeschrijving dienen informatie te bevatten over de geometrische nauwkeurigheid.

De geometrische nauwkeurigheid van TOP50NL zou vervolgens getoetst kunnen worden m.b.v. de in de norm genoemde referentie, dus m.b.v. luchtfoto's (net als voor TOP10NL) of het TOP10NL-bestand zelf.

Indien de referentie 'de werkelijkheid' is, zouden er veldmetingen moeten worden gedaan, wat zeer arbeidsintensief is. Bovendien dient er dan rekening mee gehouden te worden dat de toestand in het veld veranderd kan zijn t.o.v. de laatste herziening. Om deze redenen is het niet aan te raden om 'de werkelijkheid' als referentie te gebruiken.

Indien de luchtfoto's of TOP10NL een goede en bekende nauwkeurigheid hebben, kunnen deze goed als referentie worden gebruikt.

Bij het toetsen van de positionele nauwkeurigheid van rasterproducten dient er rekening mee te worden gehouden dat een rasterproduct de cartografische representatie bevat van objecten. Lijnen en punten worden met een bepaalde dikte weergegeven en vervolgens ook nog verrasterd. Positionele afwijkingen van punt- en lijnobjecten in een raster zijn daardoor minder makkelijk te beoordelen.

Als referentie voor een raster zou het objectgerichte product op dezelfde schaal gebruikt kunnen worden (dus bijvoorbeeld voor TOP50raster zou TOP50NL als referentie gebruikt kunnen worden), ware het niet dat het productieproces voor rasterproducten en objectgerichte producten geïntegreerd is, waardoor de positie van de cartografische representatie van objecten in het rasterproduct altijd overeen zal komen met de positie van de objecten in het objectgerichte product. Het op deze wijze toetsen van de positionele nauwkeurigheid heeft dus weinig nut en is niet aan te raden.

4.3.3 Toepasbaarheid

Het valt eenvoudig te controleren of er informatie over de geometrische nauwkeurigheid aanwezig is in bijvoorbeeld de metadata van het product, als waardes in een attribuutveld of in de productbeschrijving.

Omdat de gegevensmodellen voor de verschillende objectgerichte producten verschillen, is het bij gebruik van TOP10NL als referentie wel noodzakelijk zijn om te weten welke objecten in TOP10NL

worden omgezet naar objecten in kleinschalige producten: de 'mapping rules' moeten gedocumenteerd zijn. Anders zal het mogelijk niet altijd duidelijk zijn met welke geometrie in TOP10NL de geometrie van een object in een kleinschalig product (bijv. TOP50NL) vergeleken moet worden.

Bij gebruik van luchtfoto's als referentie speelt bovenstaand probleem uiteraard niet.

4.4 Volledigheid

4.4.1 Beschrijving

Met betrekking tot de inhoud is een logisch criterium of de informatie compleet is. De vraag is dan of alle informatie, labels en relaties aanwezig zijn die aanwezig moeten zijn. Dit strekt verder dan de pure inhoud van de gegevens op de kaart zelf. Het gaat ook om de beschikbaarheid van metadata en achtergrondinformatie bij het product, zoals informatie over hoe bepaalde klassen zijn ontstaan.

4.4.2 Methode

Om te controleren of alle gewenste informatie, labels en relaties aanwezig zijn, kan er gekeken worden naar een selectie van de data die representatief is voor het geheel, waarbij er een vergelijking gemaakt kan worden met de brondata van waaruit gegeneraliseerd is, bijvoorbeeld op de wijze zoals beschreven door Bard (2004). Door het karakteriseren van brondata en gegeneraliseerde data en deze karakteristieken met elkaar te vergelijken op het niveau van objecten, groepen van objecten of op de het niveau van de hele dataset, kan er een kwantitatief en kwalitatief oordeel worden gegeven over de generalisatie.

Een andere wijze waarop de gegeneraliseerde data vergeleken kunnen worden met de brondata, is door op basis van 'mapping rules' te controleren of de gegeneraliseerde data juist zijn. In de mapping rules dient dan te staan waaruit een object (met een bepaalde set attributen) in het gegeneraliseerde product ontstaan kan zijn met daarbij ook informatie over hoe ver dit object verplaatst mag zijn. Er kunnen dan controles worden uitgevoerd door de objecten in het gegeneraliseerde product te vergelijken met de objecten in het bronbestand en te toetsen of dit juist is.

Een derde optie om een oordeel te kunnen geven over inhoudelijke volledigheid, is het toepassen van een van de clutteranalyses, zoals beschreven door Touya, Decherf et al. (2015). De clutteranalyses ('edge density', 'subband entropy', 'quad tree complexity' en 'segmentation clutter') geven een numerieke waarde die elk een maat voor 'complexiteit' zijn.

4.4.3 Toepasbaarheid

Om de methode van Bard toe te kunnen passen, zal er bepaald moeten worden welke karakteristieken van belang zijn en welke drempelwaarden (thresholds) toegepast moeten worden om iets als juist of onjuist te beoordelen. Bepaling van de drempelwaarden kan in veel gevallen wellicht op basis van productspecificaties, indien deze beschikbaar zijn. In figuur 5 staan voorbeelden van karakteristieken die Bard noemt. Het microniveau heeft betrekking op objecten, het mesoniveau heeft betrekking op een groep objecten en het macroniveau heeft betrekking op een hele dataset. Deze methode voor controle zal voor een groot deel te automatiseren zijn, maar kan enkel toegepast worden op de objectgerichte TOPNL-producten.

Level of analysis	Characterisation		Evaluation	
	<u>Property</u>	<u>Measure</u>	<u>Function</u>	<u>Threshold and tolerance</u>
Micro	Size	Area		Minimum size: 400 m ² Tolerance: ±10%
	Shape – Elongation – Concavity ratio	$\frac{width}{length}$ $\frac{area}{convex\ area}$		Tolerance elongation: 10% Minimum Concavity ratio: 0.6 Tolerance Concavity ratio: ±10%
	Position	Centroid		Tolerance: 10 m
	Orientation	Wall Orientation		Tolerance: 5°
	Granularity	Shortest edge		Minimum granularity Tolerance: ±10%
Meso	Density	Building density		Density final > Density final
	Proximity – Between buildings – Between building and roads		No function	Minimum distance: 10 m Tolerance: ±10%
	Semantic re-partition	Buildings type (number, area) Number, area	Preservation	No threshold
Macro	Structure – Alignment – Free space		Preservation	No threshold
	Size distribution	Building size distribution	Preservation	No threshold

Figuur 5 Voorbeelden van karakteristieken die gebruikt kunnen worden om de kwaliteit van een gegeneraliseerd product te beoordelen (Bard 2004).

Een controle gebaseerd op 'mapping rules' is enkel mogelijk indien deze zijn vastgelegd. Op dit moment is dat nog niet het geval.

Aan de hand van een 'mapping' document kan er gekeken worden welke klassen of objecten er na generalisatie niet meer voor zouden komen in het kleinschalige product. Ook kan er dan (al dan niet met steekproeven) bepaald worden of de juiste objecten of klassen op de juiste locaties zijn ontstaan. Ook deze methode zal voor een groot deel te automatiseren zijn. De methode is enkel toepasbaar op de objectgerichte TOPNL-producten. Eenzelfde methodiek kan toegepast worden voor de controle op de thematische nauwkeurigheid.

De clutteranalyses (Touya, Decherf et al. 2015) zijn enkel toe te passen op de TOPraster-producten. Elke methode geeft voor (een deel van) een kaart een waarde die de 'complexiteit' aangeeft. De genoemde methodes zijn 'quad tree clutter', 'segmentation based clutter', 'edge density' en 'sub band entropy'. Drie van de methoden kunnen een visueel resultaat tonen voor het gebied, zoals te zien in figuur 6.

Figuur 6 Visuele resultaten van drie verschillende clutteranalyses (Touya, Decherf et al. 2015).

Om tot een oordeel te komen, zal er voor de gebruikte clutteranalyse(s) een drempelwaarde moeten worden bepaald om te bepalen wanneer (een deel van) een generaliseerd product juist of onjuist is. Het bepalen van een drempelwaarde blijft subjectief en daarmee zijn de clutteranalyses eigenlijk niet geschikt om gebruikt te worden in een externe audit op de kwaliteit van de producten. Eventueel kunnen de methoden wel door het Kadaster toegepast worden om mogelijke ongewenste clutter in de output van het generalisatieproces te herkennen. De methodes lijken echter niet eenvoudig te implementeren en Touya et al. geven zelf aan dat er nog veel gedaan moet worden voordat de genoemde methoden te gebruiken zijn t.b.v. generalisatie-evaluaties.

4.5 Thematische nauwkeurigheid

4.5.1 Beschrijving

Onder thematische nauwkeurigheid verstaan we de juiste vulling van attributen. Indien een attribuut verplicht is gesteld, zal het een waarde moeten hebben. Indien een attribuut een waarde heeft, moet deze waarde overeenkomen met wat er in het referentiemateriaal staat.

Thematische nauwkeurigheid heeft enkel betrekking op de objectgerichte producten (de TOPNL-reeks). Het is niet van toepassing op de rasterproducten, omdat hier de kleur (RGB-waarde) het enige attribuut is. Deze kleur wordt direct afgeleid van de gegevens in het objectgerichte product en is afhankelijk van het cartografisch ontwerp. De (controle op de) kwaliteit hiervan wordt behandeld in de paragraaf over 'bruikbaarheid'.

4.5.2 Methode

In de basis kan dezelfde methodiek gebruikt worden als voor TOP10NL. Dit houdt in dat er een steekproef wordt genomen van gebieden (zie hoofdstuk 2 voor een voorstel tot aanpassing hiervoor) en binnen deze gebieden worden alle objecten gecontroleerd. De steekproef die voor TOP10NL gebruikt wordt, kan ook voor de kleinere schalen gebruikt worden.

Voor een controle van de kleinschalige producten is het belangrijk een keuze te maken in wat er als referentiemateriaal wordt gebruikt. Er zijn drie mogelijkheden:

- Veldwaarnemingen
- Luchtfoto's en evt. cyclorama's
- TOP10NL

Bij gebruik van de eerste optie (veldwaarnemingen) is het grote nadeel dat datgene wat er in het veld te zien is, veranderd kan zijn sinds de opnamedatum van de luchtfoto die voor TOP10NL (en daarmee ook voor de kleinschalige producten) is gebruikt. Daarnaast is een controle m.b.v. veldwaarnemingen zeer arbeidsintensief. Om deze redenen worden veldwaarnemingen ook niet gebruikt in de externe controle van TOP10NL.

Voor TOP10NL worden de luchtfoto's en cyclorama's die bij de herziening zijn gebruikt als referentiemateriaal gebruikt voor de externe controle. Ditzelfde bronmateriaal zou gebruikt kunnen worden voor de kleinschalige producten.

De kleinschalige producten van de BRT worden alle afgeleid van TOP10NL. Dit product kan als referentiemateriaal gebruikt worden, waarbij er wel op gelet moet worden dat de release datum van het kleinschalige product hetzelfde moet zijn als van TOP10NL. Het grote voordeel bij gebruik van TOP10NL als referentiemateriaal is dat een controle op de thematische nauwkeurigheid grotendeels of mogelijk zelfs volledig automatisch kan worden uitgevoerd nadat dit eenmalig is ingericht.

4.5.3 Toepasbaarheid

De beschreven methode is toepasbaar op alle schalen van de BRT. Er zal per schaal wel bepaald moeten worden of alle attributen van alle objectklassen gecontroleerd worden of dat er, net als bij TOP10NL, een beperkt aantal attributen altijd wordt gecontroleerd en de rest in roulatie.

Een controle op de thematische nauwkeurigheid kan enkel indien definities van objectklassen, attributen en attribuutwaarden gedocumenteerd zijn, bijvoorbeeld in een of meerdere objectcatalogi.

Indien TOP10NL als referentiemateriaal gebruikt wordt, dienen ook de 'mapping rules' vastgelegd te zijn die bepalen wat de relatie is tussen objecten in TOP10NL en de kleinschalige objectgerichte producten.

4.6 Bruikbaarheid

4.6.1 Beschrijving

De bruikbaarheid ('usability') is, zoals de term al aangeeft, afhankelijk van het beoogde gebruik. Elk gebruik kent specifieke wensen en eisen m.b.t. de producten van de BRT. Het valt buiten de scope van dit onderzoek wat de specifieke wensen en eisen zijn voor de verschillende toepassingen waarin de kleinschalige BRT-producten gebruikt worden. Zaken die in het algemeen gelden voor kleinschalige producten worden wel meegenomen in dit onderzoek. In dit kader worden de volgende aspecten behandeld:

- Kleinst herkenbare object
- Aantal objecten
- Cartografisch ontwerp

4.6.2 Methode

De juiste schaal wordt in het algemeen bepaald in relatie tot het kleinste herkenbare object op de kaart. Een generieke en veelgebruikte regel voor rasters is: deel de deler van de kaart schaal door 1000 om de nog herkenbare afmeting te bepalen. De helft van deze waarde is dan de gewenste resolutie voor de kaart (Tobler, 1987) in meters. Dit impliceert dan ook de minimumafmeting van te herkennen objecten in de kaart.

Met 'Töpfer's Radical Law' (Töpfer and Pillewizer, 1966) kan worden berekend hoeveel objecten er op een gegeneraliseerde kaart zouden moeten staan, gebaseerd op het aantal objecten in de bronkaart. Het resultaat moet uiteraard als indicatie worden gezien, maar grote afwijkingen kunnen duiden op fouten. De formule van 'Töpfer's Radical Law' is als volgt:

$$n_f = n_a \sqrt{M_a / M_f}$$

waarbij:

n_f = het aantal objecten op de gegeneraliseerde kaart

n_a = het aantal objecten op de schaal van de bronkaart

M_a = de noemer van de schaal van de bronkaart

M_f = de noemer van de schaal van de gegeneraliseerde kaart

Het product moet goed te interpreteren zijn. Omdat er veel soorten data worden gecombineerd in één afbeelding, kan dit lastig zijn. Goed bruikbare criteria m.b.t. het cartografisch ontwerp zijn opgesteld door onder anderen Bertin (1967). Er wordt hierbij gekeken of grafische variabelen correct zijn gebruikt. Daarnaast zijn de visuele variabelen leesbaarheid, kleur, lettertypes, grootte ook van belang. Uiteindelijk gaat het erom hoe goed het product kan worden geïnterpreteerd door gebruikers. In figuur 7 is een overzicht gegeven van de toepasbaarheid van grafische variabelen in een kaartproduct (Nijholt, 2008). Per grafische variabele wordt de toepasbaarheid aangegeven.

Om het juiste gebruik van de variabelen van Bertin te controleren, kunnen moeilijk te lezen kaartlocaties geïdentificeerd worden. In *Identifying areas of a map that are difficult to read* van Olsson, Pippig et al. (2011) worden een drempelwaarde- en een cluster methode voor dit doeleinde beschreven.

Nijholt (2008) gaat met zijn *Handboek Geo-visualisatie* uitvoerig in op het gebruik van grafische en visuele aspecten voor kaarten. Of aan criteria voldaan wordt, hangt af van het doel en de doelgroep van de kaart. Bertin formuleert grafische variabelen die als basis voor evaluatie zouden kunnen dienen. Het aantal visuele klassen dient bij het weergeven van kwantitatieve verschillen gelijk of kleiner te zijn dan het ingewonnen aantal klassen. Aan de kant van de lezer zit er ook een beperking. Hieronder wordt het maximum aantal door de oog-brein-combinatie te onderscheiden klassen weergegeven, afhankelijk van of het verschijnsel door punten (bijvoorbeeld soorten steden), lijnen (bijvoorbeeld soorten transportstromen) of door vlakken (bijvoorbeeld soorten grondgebruik) wordt beschreven.

Tabel 17 Maximum aantal door de oog-brein-combinatie te onderscheiden klassen.

Variabele	Punten	Lijnen	Vlakken
Grootte	4	4	5
Grijswaarde	3	4	5
Kleur	7	7	8
Richting	4	2	-
Vorm	-	-	-

Als het gaat om symbolen zijn de belangrijkste variabelen vorm, kleur, grijswaarde en grootte. Voor vlakken zijn ook de variabelen textuur en volume belangrijk. De term grafische variabele moet niet verward worden met visuele variabele. Een (aantal) grafische variabele(n) bepaalt (bepalen) hoe het symbool eruitziet/hoe het symbool visueel overkomt. Grafische variabelen staan de kaartenmaker tot dienst bij het in kaart brengen van de objecten uit werkelijkheid. Niet alle grafische variabelen zijn even goed bruikbaar in elke situatie.

4.6.3 Toepasbaarheid

De formule van Tobler is toe te passen op alle rasterproducten. Als voorbeeld: de resolutie van TOP50raster zou volgens de formule van Tobler (1987) minimaal $50.000/2000 = 25$ m moeten zijn. De formule is ook de andere kant op te gebruiken. De afmeting van het minimaal te herkennen object vermenigvuldigd met 2000 moet in dit geval 50.000 (of lager) zijn. Hier wordt overigens met minimaal herkenbaar object niet hetzelfde bedoeld als een object in TOP50NL. Het gaat om de representatie ervan in het rasterproduct. Een wegdeel-object wordt bijvoorbeeld weergegeven met wegcontouren. De wegcontour kan dan gebruikt worden als minimaal herkenbaar object. Om de formule toe te kunnen passen, zal bekend moeten zijn wat het kleinste 'object' in de representatie is.

De formule van Töpfer's Radical Law is enkel toe te passen op de objectgerichte producten, want het aantal objecten moet bekend zijn. Indien het wordt toegepast op TOP50NL, dan leidt dat tot het volgende:

$$n_f = n_a \sqrt{10000/50000}$$
$$n_f = 0.447n_a$$

Op een schaal van 1:50.000 zouden dus volgens deze formule iets minder dan de helft van de objecten moeten voorkomen als op 1:10.000. De formule is op vergelijkbare manier toe te passen op de andere kleinschalige producten.

De formule is toe te passen op de hele dataset, maar ook op delen (die bijv. m.b.v. een steekproef zijn geselecteerd) van een dataset. Uiteraard zal een berekening op basis van echte data zelden exact uitkomen op dezelfde waarde als met Töpfer's Radical Law. Om dit toe te kunnen passen in een controle (waarbij moet kunnen worden beoordeeld of het product goed of fout is), zal er een bandbreedte moeten worden bepaald waarbinnen een product nog goed bevonden wordt.

Of de kleinschalige kaarten duidelijk en goed te interpreteren zijn, kan getest worden tegen de omschrijvingen van tabel 17 en het overzicht van de variabelen in onderstaande figuur.

grafische variabelen	punten	lijnen	vlakken	opmerkingen
1) variatie in vorm (lijnstijl / abstract / figuratief)			nvt / zie 9)	voor kwalitatieve verschillen (soorten)
2) variatie in kleur				voor kwalitatieve verschillen
3) variatie in grijswaarde (donkerte / lightness)				voor kwantitatieve verschillen (minder/meer)
4) variatie in verzadiging (color intensity)				gebruik dit alleen in combinatie met 2) en 3)
5) variatie in grootte			een anamorfose 	voor kwantitatieve verschillen
6) variatie in textuur / structuur		nvt		voor kwalitatieve verschillen. Optie 2) is beter, maar dit is voor vlakken een zwart-wit alternatief
7) variatie in richting		nvt		voor kwalitatieve verschillen. Verschillen zijn lastig te zien
8) variatie in grein	nvt	nvt		voor kwantitatieve verschillen; uit de tijd van de plak-rasters. Niet gebruiken
9) variatie in volume				voor kwantitatieve verschillen. In combinatie met 3) nog sterker effect
10) overige variabelen:	Doorzichtigheid: denk aan risicogebied of plannen over luchtfoto's heen Focus: of: onscherpte; denk een een symbool voor mist Schaduw: om 3D-effect te suggereren / versterken; een regenwolk is voorspeld boven een regio			kunnen vorige grafische variabelen versterken. Gebruik niet de mate van deze variabelen, maar hanteer deze variabelen wel of niet
toepasbaarheid:				
	ideaal	goed, soms beperkingen	lastig toepasbaar	niet

Figuur 7 Grafische variabelen. Handboek Geo Visualisatie (2016).

4.7 Overzicht van mogelijke controles

In deze paragraaf wordt in een enkel overzicht weergegeven welke controles er mogelijk zijn voor de rasterproducten en objectgerichte producten van de Basisregistratie Topografie. In onderstaande tabel staat per controle methode of het toe te passen is op de rasterproducten en/of objectgerichte producten en wat er nodig is om de methode toe te kunnen passen. De laatste kolom geeft aan of en welk nader onderzoek er eventueel nog nodig is.

Tabel 18 Overzicht van de toepasbaarheid van controlemethodes.

<i>Kwaliteitsaspect</i>				
controlemethodiek	TOPraster	TOPNL	Voorwaarde om methode toe te kunnen passen	Nader onderzoek
<i>Actualiteit</i>				
Actualiteitskalender	✓	✓	Product dient opgenomen te zijn in actualiteitskalender	Nee
<i>Logische consistentie</i>				
Op basis van regels	-	✓	Regels moeten gedocumenteerd zijn	Ja (documentatie)
<i>Positionele nauwkeurigheid</i>				
Luchtfoto als referentie	✓	✓	Beschikbaarheid van luchtfoto's Normen moeten bekend zijn	Ja (normen bepalen)
TOP10NL als referentie	✓	✓	Weten welke objecten in TOP10NL worden omgezet naar objecten in kleinschalige producten ('mapping rules') Normen moeten bekend zijn	Ja (normen bepalen, documenteren 'mapping rules')
<i>Volledigheid</i>				
Methode van Bard	-	✓	Karakteristieken en drempelwaarden vastgelegd	Ja (bepalen karakteristieken en drempelwaarden)
Op basis van 'mapping rules'	-	✓	Gedocumenteerde 'mapping rules'	Ja (documenteren 'mapping rules')
Clutter analyses	✓	-	Software geïmplementeerd voor analyse(s)	Ja (haalbaarheidsanalyse, implementatie)
<i>Thematische nauwkeurigheid</i>				
Methodiek TOP10NL	-	✓	Catalogus beschikbaar	Ja (catalogus opstellen)
<i>Bruikbaarheid</i>				
Tobler	✓	-	Omvang van kleinste 'object' is bekend	Ja (uitzoeken wat kleinste 'object' is)
Töpfer's Radical Law	-	✓	Bandbreedte van correcte waarden vastgelegd	Ja (bandbreedte bepalen)
Gebruik van grafische variabelen	✓	-		Nee

✓: direct toepasbaar

√: toepasbaar na nader onderzoek wanneer aan voorwaarden wordt voldaan

-: niet toepasbaar

Het toetsen van de actualiteit m.b.v. de actualiteitskalender en het toetsen van de bruikbaarheid aan de hand van grafische variabelen is op dit moment al toepasbaar. Er wordt aan de eventuele voorwaarden voldaan en er is geen nader onderzoek voor nodig. Alle overige toetsingsmogelijkheden zijn pas toepasbaar na nader onderzoek wanneer voldaan wordt aan de voorwaarden.

Literatuur

- Bard, S., 2004. "Quality assessment of cartographic generalisation." Transactions in GIS no. 8 (1): 63-81.
- Bertin, J., 1967. *Sémiologie Graphique: Les diagrammes, les réseaux, les Cartes*. Mouton: La Haye.
- Nijeholt, T., 2016. Handboek Geo-visualisatie. Kaarten maken met een GIS. Geraadpleegd op 14 oktober 2016 via <https://nl.wikibooks.org/wiki/Geo-visualisatie/Symbologie#Symbologie>.
- Jakobsson, A., J. Giversen, z.j. Guidelines for implementing the ISO 19100 geographic information quality standards in national mapping and cadastral agencies. Eurogeographics.
- Kadaster, 2016. Kwaliteitshandvest. Geraadpleegd op 27 oktober 2016 via <http://www.kadaster.nl/web/Kwaliteitshandvest.htm>.
- Kadasterwet, 2016. Geraadpleegd op 30 juni 2016 via <http://wetten.overheid.nl/jci1.3:c:BWBR0004541&hoofdstuk=1a&titeldeel=5&artikel=7u&z=2016-01-18&g=2016-01-18>.
- Lonskiy, I.I., A. Hopfstock, 2015. "Complex estimation of quality of cartographical products." Geraadpleegd op 17 oktober 2016 via <http://www.eurogeographics.org/sites/default/files/SDMQ2015-Ivan.pdf>
- Olsson, P., K. Pippig, L. Harrie, H. Stigmar, 2011. Identifying areas of a map that are difficult to read. Geraadpleegd op 17 oktober 2016 via <https://www.researchgate.net/publication/259196963>.
- Storm, M.H., M. Knotters, D. Brus, 2012. *Controlemethodiek Basisregistratie Topografie*. Wageningen: Alterra, Wageningen UR.
- Tobler, W., 1987. "Measuring spatial resolution." Proceedings, Land Resources Information Systems Conference, Beijing: 12-16.
- Töpfer, F., W. Pillewizer, 1966. "The principles of selection." The Cartographic Journal 3(1): 10-16.
- Touya, G., B. Decherf, M. Lalanne, M. Dumont, 2015. "Comparing Image-Based Methods for Assessing Visual Clutter in Generalized Maps." 9th International Symposium on Spatial Data Quality (ISSDQ 2015), La Grande Motte, France.

Wageningen Environmental Research
Postbus 47
6700 AA Wageningen
T 0317 48 07 00
www.wur.nl/environmental-research

Wageningen Environmental Research
Rapport 2796
ISSN 1566-7197

De missie van Wageningen University & Research is 'To explore the potential of nature to improve the quality of life'. Binnen Wageningen University & Research bundelen Wageningen University en gespecialiseerde onderzoeksinstituten van Stichting Wageningen Research hun krachten om bij te dragen aan de oplossing van belangrijke vragen in het domein van gezonde voeding en leefomgeving. Met ongeveer 30 vestigingen, 5.000 medewerkers en 10.000 studenten behoort Wageningen University & Research wereldwijd tot de aansprekende kennisinstellingen binnen haar domein. De integrale benadering van de vraagstukken en de samenwerking tussen verschillende disciplines vormen het hart van de unieke Wageningen aanpak.

To explore
the potential
of nature to
improve the
quality of life

Wageningen Environmental Research
Postbus 47
6700 AB Wageningen
T 317 48 07 00
www.wur.nl/environmental-research

Rapport 2796
ISSN 1566-7197

De missie van Wageningen University & Research is 'To explore the potential of nature to improve the quality of life'. Binnen Wageningen University & Research bundelen Wageningen University en gespecialiseerde onderzoeksinstituten van Stichting Wageningen Research hun krachten om bij te dragen aan de oplossing van belangrijke vragen in het domein van gezonde voeding en leefomgeving. Met ongeveer 30 vestigingen, 5.000 medewerkers en 10.000 studenten behoort Wageningen University & Research wereldwijd tot de aansprekende kennisinstellingen binnen haar domein. De integrale benadering van de vraagstukken en de samenwerking tussen verschillende disciplines vormen het hart van de unieke Wageningen aanpak.