



Ruimtelijke modellering van plantengemeenschappen in Nederland

Ontwerp en kalibratie van een Random Forest-model voor de ruimtelijke
verdeling van plantengemeenschappen

W. de Winter, M. Knotters, S. Hennekens en C. Mùcher



ALTERRA
WAGENINGEN UR

Ruimtelijke modellering van plantengemeenschappen in Nederland

Ontwerp en kalibratie van een Random Forest-model voor de ruimtelijke verdeling van
plantengemeenschappen

W. de Winter¹, M. Knotters¹, S.M. Hennekens¹ en C.A. Mûcher¹

¹ WUR ESG Alterra

Dit onderzoek is uitgevoerd door Alterra Wageningen UR in opdracht van en gefinancierd door het Ministerie van Economische Zaken, in het kader van het programma Kennisbasis 14-Duurzame Ontwikkeling van de Groenblauwe Ruimte, thema 2-Groene ruimte en Biodiversiteit, project DynTerra (Dynamische Terrestrische Natuur).

Alterra Wageningen UR
Wageningen, november 2014

Alterra-rapport 2582
ISSN 1566-7197

De Winter, W., M. Knotters, S.M. Henneken en C.A. Mùcher, 2014. *Ruimtelijke modellering van plantengemeenschappen in Nederland; Ontwerp en kalibratie van een Random Forest-model voor de ruimtelijke verdeling van plantengemeenschappen*. Wageningen, Alterra Wageningen UR (University & Research centre), Alterra-rapport 2582. 80 blz.; 46 fig.; 3 tab.; 12 ref.

Trefwoorden: vegetatie, modellen, bodem, topografie, landgebruik, LVD, floristiek

Dit rapport is gratis te downloaden van www.wageningenUR.nl/alterra (ga naar 'Alterra-rapporten' in de grijze balk onderaan). Alterra Wageningen UR verstrekt geen gedrukte exemplaren van rapporten.

© 2014 Alterra (instituut binnen de rechtspersoon Stichting Dienst Landbouwkundig Onderzoek), Postbus 47, 6700 AA Wageningen, T 0317 48 07 00, E info.alterra@wur.nl, www.wageningenUR.nl/alterra. Alterra is onderdeel van Wageningen UR (University & Research centre).

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

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

Alterra-rapport 2582 | ISSN 1566-7197

Foto omslag: W. de Winter – iconoclastica.nl

Inhoud

1	Inleiding	5
2	Bronnen	6
2.1	Bronnen	6
2.1.1	Bodemkaart	6
2.1.2	Visueel Ruimtelijk InformatieSysteem (VIRIS)	7
2.1.3	Syntaxon-kaart op basis van FLORON-data	7
2.1.4	LGN-4	7
2.1.5	AHN en AHN-relatief	7
2.1.6	Landelijke Vegetatie Databank (LVD)	8
2.2	Analyse	8
3	Resultaten	10
3.1	Analyse op associatieniveau	10
3.1.1	Punten met Gt-code	10
3.1.2	Punten zonder Gt-code	11
3.2	Analyse op verbondsniveau	11
3.2.1	Punten met Gt-code	11
3.2.2	Punten zonder Gt-code	11
4	Pilot runs	12
4.1	Random-Forestmodel wel/geen verbond 16BB (<i>Arrhenatherion elatioris</i>)	12
4.1.1	punten met Gt-code	12
4.1.2	punten zonder Gt-code	12
4.2	Random-Forestmodel wel/geen verbond 16BA (<i>Alopecurion pratensis</i>)	12
4.2.1	punten met Gt-code	12
4.2.2	punten zonder Gt-code	12
4.3	Random-Forestmodel wel/geen verbond 43AA (<i>Alno-Padion</i>)	13
4.3.1	punten met Gt-code	13
4.3.2	punten zonder Gt-code	13
5	Validatie	14
5.1	Validatiemethode	14
5.2	Resultaten	14
5.3	voorspellingen voor heel Nederland	15
6	Conclusies en aanbevelingen	18
	Literatuur	19
	Bijlage 1 LGN-4	20
	Bijlage 2 Verbonden	22
	Bijlage 3 Figuren	24
	Bijlage 4 Validatie heel Nederland	48
	Bijlage 5 Voorspellingen heel Nederland	57

1 Inleiding

Elke plantensoort kent zijn eigen specifieke voorkeuren ten aanzien van zijn biotisch en abiotisch milieu. Wanneer men deze voorkeuren zou kennen en bovendien de ruimtelijke verspreiding van deze milieufactoren in kaart heeft gebracht, is een ruimtelijk expliciete voorspelling van het voorkomen van de soorten mogelijk. Botanici maken dan ook al eeuwenlang gebruik van deze methode om gericht naar nieuwe vindplaatsen van geselecteerde soorten te zoeken, ook al zijn hun datasets verre van compleet.

Een stap verder kun je gaan door niet de kans op het voorkomen van soorten uit abiotische en biotische factoren te voorspellen, maar uitspraken te doen over de biotische context zelf op basis van de beschikbare kennis. In dit rapport wordt dit proces methodisch uitgewerkt. Vegetatiekundige syntaxa zijn de biotische variabelen waarvan de ruimtelijke kansverdeling wordt geschat vanuit een aantal landsdekkende, abiotische, topografische en andere datasets. Om praktische redenen bepalen we ons hier tot vegetatietypes op het niveau van verbonden.

Met informatie uit ruimtelijke databases is een generieke methodiek ontwikkeld, waarmee de ruimtelijke verspreiding van plantengemeenschappen in Nederland door middel van statistische modellering kan worden voorspeld en vlakdekkend kunnen worden voorspeld uit digitale bestanden (topografie, bodem, landgebruik en hoogtemodellen), en waarmee de nauwkeurigheid van de ruimtelijke predicties kan worden gekwantificeerd.

2 Bronnen

2.1 Bronnen

Als verklarende datasets zijn gebruikt:

- de Bodemkaart van Nederland, schaal 1:50.000;
- VIRIS-rasterbestanden, afgeleid uit Top10-vector, met rastercellen van 25x25 meter;
- syntaxonkaarten op basis van de FLORON-database voor gridcellen van 1x1 kilometer;
- de nationale landgebruiksdatabase LGN met rastercellen van 25x25 meter;
- het Actueel Hoogtebestand Nederland (AHN-1) geaggregeerd naar gridcellen van 25x25 meter.

Deze bestanden worden in meer detail hieronder beschreven.

De doelvariabele, vegetatietype, is gelezen uit de Landelijke Vegetatie Database (LVD), een dataset met gegevens van ruim 135 000 locaties waarvan de vegetatie bekend is. Vegetatietype is een multinomiale klassevariabele, met van hoog naar laag de volgende aggregatieniveaus: klasse (42 categorieën), verbond (87 categorieën), associatie (215 categorieën) en subassociatie (367 categorieën).

Er zijn 54 potentiële predictorvariabelen:

- 7 kwalitatieve variabelen gerelateerd aan de bodemkaart;
- 1 kwalitatieve variabele gerelateerd aan de Gt-kaart;
- 44 kwantitatieve variabelen gerelateerd aan VIRIS;
- 1 kwalitatieve variabele gerelateerd aan LGN4;
- 1 kwantitatieve variabele (AHN-hoogte).

2.1.1 Bodemkaart

De Bodemkaart van Nederland, schaal 1 : 50 000, geeft informatie over belangrijke kenmerken van het bodemprofiel tot een diepte van 1,20 m. De grondwatertrappen (Gt's) geven informatie over de seizoensfluctuatie, uitgedrukt in gemiddeld hoogste en de gemiddeld laagste grondwaterstand. De gegevens van de Bodemkaart van Nederland, schaal 1 : 50 000, zijn verzameld tussen 1958 en 1999. Gedurende deze periode is de onderverdeling van de grondwatertrappen twee keer aangepast, waardoor de recente kaartbladen specifiekere informatie over het grondwaterstandsverloop bevatten dan de oudere kaartbladen. Bovendien kunnen de gegevens over de Gt verouderen als gevolg van veranderingen in het hydrologische systeem door ingrepen na de datum van opname. Het vectorbestand is opgebouwd met het GIS-pakket ARC/INFO. Doordat van elke kaartenheid de code in het bestand is opgesplitst, zijn er veel mogelijkheden voor specifieke selecties (de Vries et al., 2003). De bodemkundige informatie op de bodemkaart heeft betrekking op de aard en samenstelling van de bovengrond (grondsoort) met een verdere onderverdeling naar bodemvorming, veensoort, afwijkende lagen in het profiel, aanwezigheid van kalk en verstoringen door vergraving en egalisatie. De kaart geeft meestal geen bodemkundige informatie voor de stedelijke gebieden en voor bodems onder waterlichamen zoals meren, plassen en rivieren. Associaties op de bodemkaart zijn omgezet in bodemtypen volgens de Tabel in Bijlage 2. Een aantal omschrijvingen op de bodemkaart verdeeld over twee kolommen, beginnend met '|', zijn vervangen, telkens in kolom 'LETTER': afgrav, bebouw, bovland, dijk, egal, groeve, moeras, ophoog, terp, verwerk, water.

2.1.2 Visueel Ruimtelijk InformatieSysteem (VIRIS)

De VIRIS-bestanden vormen een serie gridbestanden met een ruimtelijke resolutie van 25 meter, afgeleid van Top10vector. Zo'n 50 objectsoorten ("TDN-codes") zijn omgezet naar een 25-metergrid (één rasterbestand voor elke TDN-code), waarbij voor elke gridcel de waarde aangeeft hoeveel vierkante meter van het betreffende topografisch element aanwezig is. De maximum waarde voor een gridcel van 25 bij 25 meter is dus 625 (m²).

Bij lijnelementen is de lengte per gridcel bepaald en weergegeven in strekkende meters. Van puntsymbolen is het aantal per gridcel vastgelegd.

Er zijn VIRIS-bestanden van 1996 en van 2000 t/m 2006 beschikbaar.

2.1.3 Syntaxon-kaart op basis van FLORON-data

Het FLORON-bestand bevat waarnemingen van plantensoorten over de periode 1975-2000 per kilometerhok. Op basis van sommatie van de `_trouwgraden_` van alle soorten die verbonden zijn aan een bepaalde associatie is de kans geschat dat de betreffende associatie in het kilometerhok voorkomt. De hoogst gesommeerde waarde is daarbij op 100 gesteld en alle andere waarden zijn daaraan gerelateerd. Deze standaardisatie is per associatie uitgevoerd en omdat de soortenrijkdom per gemeenschap aanzienlijk kan verschillen variëren dus ook de waarden voor gesommeerde trouwgraden sterk.

2.1.4 LGN-4

Het LGN4-bestand is het vierde nationale bestand (1999, 2000) in een serie landgebruiksbestanden waarvan de eerste in 1988 is ontwikkeld (zie Bijlage 1 op pag. 20 en www.lgn.nl).

De legenda van het LGN4-bestand is gelijk aan de legenda van het LGN3-bestand, met die uitzondering dat de klassen droge heide (code 13), overig open begroeid natuurgebied (code 14) en kale grond in natuurgebied (code 15) vervangen zijn door de LGN3plus natuurklassen.

Om de analyses met de random-forestmethode uit te kunnen voeren (zie 3 Analyse), mocht het aantal categorieën per klasse-variabele niet groter zijn 32. Voor dit onderzoek is zijn categorieën landgebruik samengevoegd om het aantal terug te brengen tot maximaal 32:

- 3, 4, 5, 6 en 10 (mais, aardappelen, bieten, graan, overige landbouwgewassen, bloembollen): samengevoegd tot klasse 300;
- 18, 19 en 26 (stedelijk bebouwd gebied, bebouwing in buitengebied, bebouwing in agrarisch gebied): samengevoegd tot klasse 301;
- 37 en 38: samengevoegd tot klasse 302;
- 32 en 33 (open en gesloten duinvegetatie): samengevoegd tot klasse 303.

2.1.5 AHN en AHN-relatief

Het Actueel Hoogtebestand Nederland (AHN) is een bestand met voor heel Nederland gedetailleerde en precieze hoogtegegevens. Deze hoogtes zijn bepaald met behulp van een laseraltimetrie vanuit een vliegtuig of helikopter. De laatste versie AHN-2 is nog niet geheel dekkend voor Nederland en daarom is AHN-1 gebruikt, dat is gemaakt in de periode 1996 tot 2003. Van de gemeten hoogtes is een aantal producten gemaakt; hier is het gridformaat van 25x25 meter gebruikt.

Voor het voorspellen van vegetatietypes is de bruikbaarheid van het hoogtebestand beperkt, aangezien de hoogte boven zeeniveau sterk geografisch gecorreleerd is. Voor planten zijn de ligging ten opzichte van het omringende terrein (vochtgradiënt) en de steilte (grondwaterstroming, expositie) van groter belang. Daarom is voor dit project een afgeleide van het AHN-1 gemaakt, waarin de relatieve hoogte van de cel is bepaald ten opzichte van het gemiddelde van het 500-meterkwadrant waarvan de cel het middelpunt vormt.

2.1.6 Landelijke Vegetatie Databank (LVD)

De Landelijke Vegetatie Databank (LVD) is een gegevensbestand over de plantengroei van Nederland. In de databank staan ongeveer 500.000 recente en historische vegetatiebeschrijvingen (vegetatieopnamen), in geautomatiseerde vorm bijeengebracht. De gegevens beslaan een periode van ruim 75 jaar en omvatten zowel aquatische als terrestrische begroeiingen, goed ontwikkelde plantengemeenschappen, maar ook verarmde gemeenschappen. De waarnemingen betreffen zowel het cultuurlandschap als de halfnatuurlijke en natuurlijke landschappen.

2.2 Analyse

Allereerst is multinomiale logistische regressie toegepast om vegetatietypen te voorspellen uit de predictorvariabelen (Hosmer en Lemeshow, 2000). Het voordeel van deze methode is dat de gekalibreerde parameterwaarden interpreteerbaar zijn met behulp van de `_odds ratio_`. De odds ratio is een maat van de grootte van een effect en beschrijft de sterkte van het verband tussen twee binaire waarden. De odds ratio geeft de relatieve waarschijnlijkheid aan. Op een zandgrond kan het bijvoorbeeld tweemaal zo waarschijnlijk zijn dat er een bepaalde plantengemeenschap voorkomt dan op een niet-zandgrond. De odds ratio is in dit voorbeeld twee. Toepassing van multinomiale logistische regressie op de beschikbare dataset leidt echter snel tot numerieke problemen, gezien het grote aantal variabelen, het grote aantal klassen per kwalitatieve variabele en het grote aantal missing values. Het aantal predictorvariabelen zou wellicht terug te brengen zijn met behulp van een multivariate analyse, maar het probleem daarbij is de aanwezigheid van kwalitatieve variabelen, de vele missing values en het verlies aan mogelijkheden om de gekalibreerde parameters te interpreteren.

Een alternatief is de random-forest-methode van Breiman en Cutler (zie Breiman, 2001). Met deze methode kan in korte tijd een grote dataset worden geanalyseerd. Missing values leiden niet tot problemen. De methode doet trekkingen uit de dataset en zet daarbij een aantal waarnemingen apart (out of bag, OOB), ontwikkelt een groot aantal classificatiebomen en evalueert de classificatiefout door predicties te vergelijken met waarnemingen in de 'out-of-bag' testset (OOB) die ca. een derde van de waarnemingen bestrijkt. Dit levert ook een maat voor de voorspelnaauwkeurigheid van het uiteindelijke model op (OOB error rate). De methode is 'black-box' in de zin dat er geen fysisch interpreteerbare parameters worden geschat. Door het effect van het weglaten van variabelen te evalueren wordt aangegeven welke variabelen belangrijk zijn (variable importance index) en dit helpt bij de interpretatie. Peters et al. (2007) vergeleken multinomiale logistische regressie en random forests bij de ruimtelijke predictie van vegetatietypen in België. Zij concludeerden dat de random-forest-methode nauwkeuriger predicties oplevert dan multinomiale logistische regressie.

Bij het construeren van een classificatieboom moet worden besloten waar de takken zich splitsen en tot hoever de boom 'doorgroeit'. Evenals bij een eenvoudig lineair regressiemodel wordt hierbij gezocht naar een beste fit, waarbij de responsvariabele zo veel mogelijk wordt verklaard uit de verklarende variabele. Een criterium dat wordt gebruikt bij het optimaliseren van classificatiebomen is de gini-index. Dit is een maat voor de onzuiverheid in de vertakkingen in een classificatieboom. We definiëren nu een gini-index voor een knooppunt m , dat een regio R_m met N_m waarnemingen vertegenwoordigt (Hastie et al., 2009, blz. 309). In dit knooppunt is

$$\hat{p}_{mk} = \frac{1}{N_m} \cdot \sum_{x_i \in R_m} I(y_i = k)$$

de fractie van waarnemingen behorend tot klasse k in knooppunt m . Alle waarnemingen in knooppunt m classificeren we in de meest voorkomende klasse in m ,

$$k(m) = \arg \max_k \hat{p}_{mk}$$

Hierdoor ontstaat een onzuiverheid, want niet alle waarnemingen zullen immers tot de meest voorkomende klasse behoren. Een maat voor de onzuiverheid is de gini-index:

Voor het geval er twee klassen zijn, en p de fractie waarnemingen in de tweede klasse is, dan is de gini-index gelijk aan

$$\text{gini-index} = \frac{1}{N_m} \cdot \sum_{k=1}^K \hat{p}_{mk}(1 - \hat{p}_{mk})$$

De variable importance index is gebaseerd op de gini-index. Hoe sterker de gini-index afneemt door het toevoegen van een variabele, hoe belangrijker deze variabele is.

De random-forest-methode kan, evenals multinomiale logistische regressie, worden uitgevoerd met het statistische programma R. Berekeningen met de volledige dataset van 135203 punten bleken numeriek niet mogelijk. Daarom is de dataset gesplitst in twee delen:

1. Punten waarvan de Gt-code bekend is (81740 locaties).
2. Punten waarvan de Gt-code niet bekend is (53463 locaties).

De motivatie voor deze splitsing is dat als er een Gt-code is, er in elk geval sprake is van terrestrische vegetatie en er ook een bodemkundige profielbeschrijving bekend is. Als er geen Gt-code is, dan is er sprake van water of moeras, dijken, bebouwing, bermen, hellingen, etc.

In eerste instantie is de aandacht gericht op een model voor de 81740 locaties waarvoor de Gt-klasse bekend is en de meeste bodemkundige informatie ook aanwezig is. Vervolgens is een aantal klassevariabelen geaggregeerd, omdat het aantal klassen per variabele niet meer dan 32 mag bedragen. De bodemkundige informatie is geaggregeerd tot de volgende klassen: veengronden, moerige gronden, moderpodzolgronden, humuspodzolgronden, leembrikgronden, oude kleibrikgronden, zandbrikgronden, enkeerdgronden, tuineerdgronden, zandgronden, kalkhoudende bijzonder lutumarme gronden, niet-gerijpte minerale zeekleigronden, niet-gerijpte minerale rivierkleigronden, zeekleigronden, rivierkleigronden, keileemgronden en een categorie overig. Verder is onderscheid gemaakt naar kalkloos en kalkhoudend.

Ook een aantal categorieën landgebruik moesten worden samengevoegd, om het aantal terug te brengen tot maximaal 32 (zie 2.2.4; Bijlage 1).

Vervolgens is gefocust op een model voor de 53463 locaties waarvoor geen Gt-code bekend is. Voor zover hiervoor een lettercode van de bodemkaart bekend is, is aan de bovenstaande klassen de volgende toegevoegd: oude mariene afzettingen, oude fluviaatiele afzettingen, afgrav, bebouw, bovland, dijk, groeve, kreekbedding, Linge-uiteerwaard, moeras, ophoog, petgat, strandwal, stuifzand, terp en water. De volgende klassen zijn verwijderd: moderpodzolgronden, humuspodzolgronden en enkeerdgronden. Op deze wijze bedraagt het aantal bodemklassen 32. De aggregatie van landgebruiksklassen is gelijk aan die bij de punten met Gt-code.

De aanpak is nu voor de twee subsets als volgt:

1. Maak met alle predictorvariabelen een classificatie m.b.v. de random-forest-methode.
2. Analyseer de variable importance. Verwijder de minst belangrijke variabelen en kalibreer een nieuw classificatiemodel.

De modellering is aanvankelijk uitgevoerd op het niveau van Verbond, en uiteindelijk op het niveau van Associatie. Het aantal (beslissings)bomen in het random-forest-model is gesteld op 500.

3 Resultaten

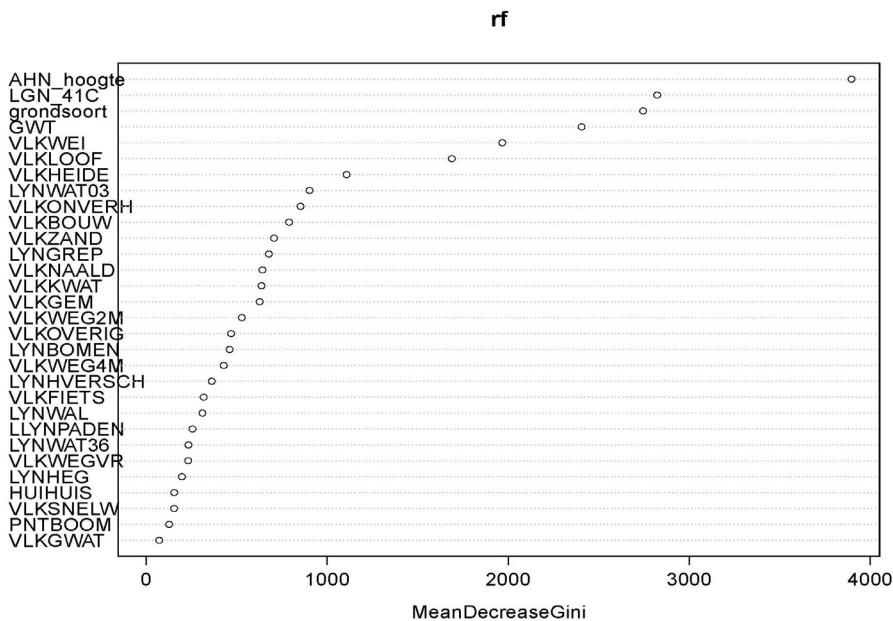
3.1 Analyse op associatieniveau

3.1.1 Punten met Gt-code

Een classificatiemodel met alle predictorvariabelen heeft een OOB error rate van 68%. Dit kan worden geïnterpreteerd als een verwachte kaartzuiverheid (percentage correct) van de voorspelde vegetatiekaart van 32%. Dit lijkt niet erg hoog, maar de mate van misclassificatie is niet verdisconteerd: één klasse verschil met de werkelijkheid is net zo erg als twee klassen verschil.

Figuur 1 geeft een overzicht van de variable importance. Hoe groter de waarde voor "MeanDecreaseGini", des te belangrijker de variabele.

Het blijkt dat AHN-hoogte, landgebruiksklasse volgens het LGN4, grondsoort volgens de bodemkaart, Gt-klasse volgens de bodemkaart en de categorieën VLKWEI, VLKLOOF en VLKHEIDE het meest bepalend zijn bij de predictie van associaties van vegetatie. Het aantal variabelen m dat random is geselecteerd om de optimale splitsing op een knooppunt te vinden is zes.



Figuur 1 Variable importance van het random-forestmodel voor de punten met Gt-code, met alle predictorvariabelen opgenomen in het model.

Figuur 3 (pag. 26) geeft de OOB-error rate versus het aantal bomen n. Hieruit blijkt dat een bos met 500 bomen groot genoeg is. In het vervolg is n daarom op 500 gesteld. Een classificatiemodel met alleen de belangrijkste predictorvariabelen heeft een OOB-error rate van 66%. Dat is dus iets beter dan het model met alle predictorvariabelen. Let wel dat de error rate met een soort kruisvalidatieprocedure wordt berekend en dus een schatting is. Het aantal variabelen m dat random is geselecteerd om de optimale splitsing op een knooppunt te vinden is vier. Figuur 4 geeft een overzicht van de variable importance in het uitgedunde model. Duidelijk blijkt dat AHN-hoogte veel verklaart.

3.1.2 Punten zonder Gt-code

Een random-forest-model met alle predictorvariabelen heeft een OOB error rate van 60%, wat we voorlopig interpreteren als een verwachte kaartzuiverheid (percentage correct) van 40%. Figuur 5 (pag. 27) geeft een overzicht van de variable importance. Ook hier verklaart de AHN-hoogte veel. Interessant is dat een predictorvariabele uit het VIRIS-bestand, VLKGWAT, naar boven komt als belangrijk. Er zijn zes predictorvariabelen m aselect gekozen om de optimale splitsing op een knooppunt te bepalen.

Een random-forestmodel met alleen de zeven belangrijkste predictorvariabelen heeft eveneens een OOB-error rate van 60%, voorlopig te interpreteren als een verwachte kaartzuiverheid (percentage correct) van 40%. Het aantal predictorvariabelen m dat aselect is gekozen om de optimale splitsing op knooppunten te bepalen bedraagt twee. Figuur 6 geeft een overzicht van de variable importance van het uitgedunde model.

3.2 Analyse op verbondsniveau

3.2.1 Punten met Gt-code

Een random-forestmodel met alle predictorvariabelen heeft een OOB error rate van 61% (percentage correct geclassificeerd 39%). Dat is dus een lichte verbetering t.o.v. classificatie op associatieniveau (OOB error rate 68%). Figuur 7 geeft een grafiek van de Variable Importance van het model met alle variabelen. Het model met alleen de belangrijkste variabelen heeft een OOB error rate van 62% (percentage correct geclassificeerd 38%). Figuur 8 geeft de Variable Importance van dit model.

Als relatieve maaiveldshoogte bij een zoekstraal van 250 m wordt toegevoegd als predictorvariabele, dan is de OOB error rate 53% (percentage correct geclassificeerd is 47%). Figuur 23 geeft de grafiek van de Variable Importance van het model met alle predictorvariabelen (m=7). Relatieve maaiveldshoogte is de op-één-na-belangrijkste predictorvariabele. Ook als alleen de belangrijkste predictorvariabelen in het model worden opgenomen is de OOB error rate 53% (m=2), zie Figuur 24 (pag. 36).

3.2.2 Punten zonder Gt-code

Een random-forest-model met alle predictorvariabelen heeft een OOB-error rate van 57% (percentage correct geclassificeerd 43%). Dat is dus een lichte verbetering t.o.v. classificatie op associatieniveau (OOB error rate 60%). Figuur 9 geeft een grafiek van de Variable Importance van het model met alle variabelen.

Als relatieve maaiveldshoogte bij een zoekstraal van 250 m wordt toegevoegd als predictorvariabele, dan is de OOB error rate 53% (percentage correct geclassificeerd is 47%). Figuur 25 geeft de grafiek van de Variable Importance van het model met alle predictorvariabelen (m=6). Relatieve maaiveldshoogte is de op-één-na-belangrijkste predictorvariabele. Als alleen de belangrijkste predictorvariabelen in het model worden opgenomen is de OOB error rate 52% (percentage correct is 48%, m=2), zie Figuur 26.

4 Pilot runs

4.1 Random-Forestmodel wel/geen verbond 16BB (Arrhenatherion elatioris)

4.1.1 punten met Gt-code

Om een kaart te maken van de verspreiding van verbond 16BB (*Arrhenatherion elatioris*) is een random-forest-model ontwikkeld voor de kwalitatieve responsvariabele wel/geen 16BB. Het aantal bomen n bedraagt 500. Het aantal aselect gekozen predictoren om per knooppunt de optimale splitsing te bepalen bedraagt $m=6$. Het volle model (met alle predictoren) heeft een OOB error rate van 2%, dus een verwachte kaartzuiverheid van 98%. Figuur 11 (pag. 30) geeft een grafiek van de variable importance. Het model met alleen de belangrijkste predictorvariabelen heeft eveneens een OOB error rate van 2%. Figuur 12 geeft een grafiek van de variable importance. In dit model is $m=3$. tabel 1 geeft de percentages correct geclassificeerd.

4.1.2 punten zonder Gt-code

Tabel Om een kaart te maken van de verspreiding van verbond 16BB is een random-forest-model ontwikkeld voor de kwalitatieve responsvariabele wel/geen 16BB. Het aantal bomen n bedraagt 500. Het aantal aselect gekozen predictoren om per knooppunt de optimale splitsing te bepalen bedraagt $m=6$. Het volle model (met alle predictoren) heeft een OOB error rate van 2%, dus een verwachte kaartzuiverheid van 98%. Figuur 13 geeft een grafiek van de variable importance. Het model met alleen de belangrijkste predictorvariabelen heeft eveneens een OOB error rate van 2%. Figuur 14 geeft een grafiek van de variable importance. In dit model is $m=2$. tabel 1 geeft de percentages correct geclassificeerd (Figuur 30 op pagina 41).

4.2 Random-Forestmodel wel/geen verbond 16BA (*Alopecurion pratensis*)

4.2.1 punten met Gt-code

Het aantal bomen n bedraagt 500. Het aantal aselect gekozen predictoren om per knooppunt de optimale splitsing te bepalen bedraagt $m=6$. Het volle model (met alle predictoren) heeft een OOB-error rate van 0.1%, dus een verwachte kaartzuiverheid van 99.9%. Figuur 15 (pag. 32) geeft een grafiek van de variable importance. Het model met alleen de belangrijkste predictorvariabelen heeft eveneens een OOB error rate van 0.1%. Figuur 16 geeft een grafiek van de variable importance. In dit model is $m=2$. tabel 1 geeft de percentages correct geclassificeerd.

4.2.2 punten zonder Gt-code

Het aantal bomen n bedraagt 500. Het aantal aselect gekozen predictoren om per knooppunt de optimale splitsing te bepalen bedraagt $m=6$. Het volle model (met alle predictoren) heeft een OOB error rate van 0.1%, dus een verwachte kaartzuiverheid van 99.9%. Figuur 17 geeft een grafiek van de variable importance. Het model met alleen de belangrijkste predictorvariabelen heeft eveneens een OOB error rate van 0.1%. Figuur 18 geeft een grafiek van de variable importance. In dit model is $m=2$. tabel 1 geeft de percentages correct geclassificeerd. (Figuur 29 pagina 40).

4.3 Random-Forestmodel wel/geen verbond 43AA (Alno-Padion)

4.3.1 punten met Gt-code

Het aantal bomen n bedraagt 500. Het aantal aselect gekozen predictoren om per knooppunt de optimale splitsing te bepalen bedraagt $m=6$. Het volle model (met alle predictoren) heeft een OOB error rate van 4%, dus een verwachte kaartzuiverheid van 96%. Figuur 19 (pag. 34) geeft een grafiek van de variable importance. Het model met alleen de belangrijkste predictorvariabelen heeft een OOB error rate van 3%. Figuur 20 geeft een grafiek van de variable importance. In dit model is $m=2$. tabel 1 geeft de percentages correct geclassificeerd (Figuur 31 pagina 42).

4.3.2 punten zonder Gt-code

Het aantal bomen n bedraagt 500. Het aantal aselect gekozen predictoren om per knooppunt de optimale splitsing te bepalen bedraagt $m=6$. Het volle model (met alle predictoren) heeft een OOB error rate van 1%, dus een verwachte kaartzuiverheid van 99%. Figuur 21 geeft een grafiek van de variable importance. Het model met alleen de belangrijkste predictorvariabelen heeft eveneens een OOB error rate van 1%. Figuur 22 geeft een grafiek van de variable importance. In dit model is $m=2$. Zie tabel 1 voor de percentages correct geclassificeerd.

Tabel 1

Percentages correct geclassificeerd, voor de aanwezigheid van vegetaties behorend tot verbond 16BA, 16BB en 43AA. De classificaties zijn gemaakt met random-forests met alleen de belangrijkste variabelen

verbond	deelgebied	percentage correct geclassificeerd		
		aanwezig	afwezig	totaal
16BA	Gt-code	65	100	100
16BA	Geen Gt-code	4	100	100
16BB	Gt-code	44	99	98
16BB	Geen Gt-code	28	100	98
43AA	Gt-code	37	99	97
43AA	Geen Gt-code	52	100	99

5 Validatie

De werking van het model is getoetst in het eerder gebruikte 2x2 km kwadrant bij Cortenoever (Gld). Van een deel van dit gebied is een gedetailleerde vegetatiekartering beschikbaar (SBB/Waardenburg) waarmee de predicties van het model gevalideerd kunnen worden (Figuur 27 op pagina 38).

5.1 Validatiemethode

Het model is twee maal gerund met de attributentabel van het testgebied als invoer, te weten eenmaal voor het deelgebied met grondwatertrap en eenmaal voor het resterende deel zonder grondwatertrap. Het gecombineerde resultaat van beide runs is een tabel met per 25m-cel de geschatte kans voor elk van de 88 verbonden. Per cel zijn de verbondskansen gerangnummerd, waarbij de hoogste kans het laagste rangnummer (1) krijgt. Wanneer nu de rangnummers worden opgezocht van de gevonden (=gekarteerde) verbonden, dan zou een gemiddelde van 1 indiceren voor een perfecte voorspelling, terwijl een gemiddelde van 44 (=88/2) zou duiden op random resultaat.

5.2 Resultaten

Voor 548 cellen in het testgebied kon de predictie vergeleken worden met een gekarteerd syntaxon (tabel 2). De validatieprocedure is tweemaal doorlopen: eenmaal direct met de predictie van het Random Forest-model en eenmaal na vermenigvuldiging van de kansen met de kansschattingen op basis van de soortverspreidingsgegevens per kilometerhok (Floron-data).

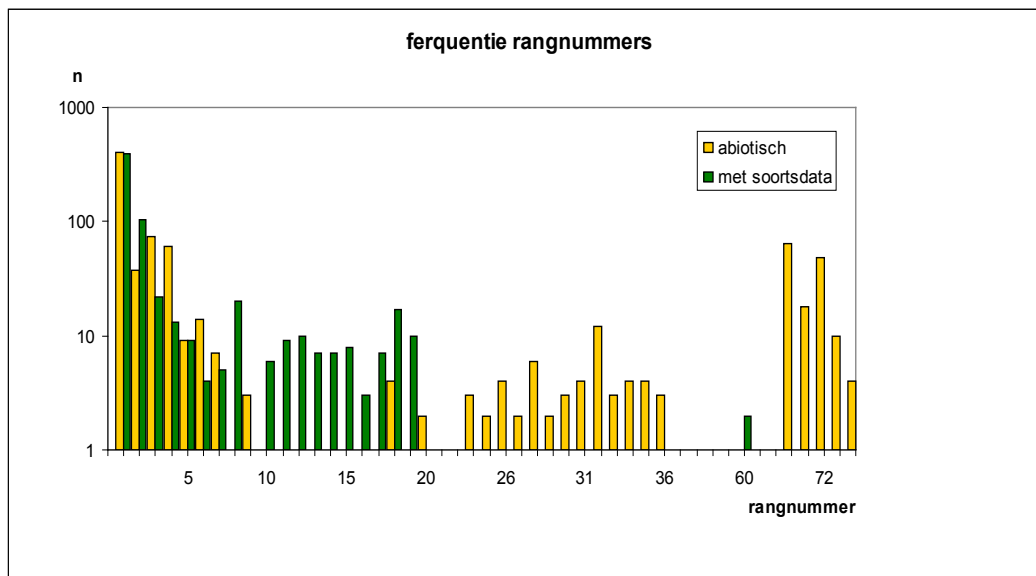
In beide gevallen ligt de mediaan op drie, ofwel in minstens de helft van het aantal cellen zit het gekarteerde syntaxon in de top-3 van de voorspelde syntaxa. Dit kan als een goed resultaat worden geïnterpreteerd.

Tabel 2

Validatie - rangnummer-ranges voor directe predictie en na combinatie met soort-verspreidingsdata

	alleen abiotisch	met soortdata
best	1	1
gemiddeld	16.34	4.00
slechtst	74	62
n	862	663
mediaan	3	2

Verder blijkt uit de validatie, dat de modelprestaties verschillen per syntaxon (tabel 3). Met name 30BA en 16BA worden minder goed voorspeld. De eerste toont een sterke verbetering als gevolg van het combineren met de soortdata; de tweede wordt er juist door geëlimineerd. Zie Figuur 28 voor de ruimtelijke spreiding van het resultaat.



Figuur 1 Rangnummercorrespondentie in het proefgebied, direct op basis van de predictie ("abiotisch", geel) en na combinatie met kansen op basis van soortverspreidingsdata ("met soortsddata", groen)

Tabel 3

Schattingsnauwkeurigheid in het proefgebied uitgesplitst per verbond, direct op basis van de predictie ("abiotisch") en na combinatie met kansen op basis van soortverspreidingsdata ("met soortsddata")

verbond	abiotisch				met soortsddata			
	n	best	gem.	slechtst	n	best	gem.	slechtst
12BA	366	1	2.37	34	370	1	1.52	19
14BC	7	4	11.42	28	62	4	13.24	61
16BA	144	70	71.09	74	-	-	-	-
16BB	205	1	1.24	4	141	1	1.26	3
30BA	55	24	32.55	72	55	8	17.36	62
38AA	17	1	9.29	20	19	1	5.84	13
43AA	32	1	3.50	9	16	1	1.50	5

5.3 voorspellingen voor heel Nederland

zie Bijlage 5, Bijlage 6

Voor het beoordelen van de Dyntaxa kanskaarten op nationaal niveau is een keuze gemaakt voor een 9-tal plantensociologische verbonden die floristisch, ecologisch en wat betreft verspreiding voor Nederland goed gedocumenteerd zijn.

- 06Ac - Hydrocotylo-Baldellion matig voedselarme, basische vennen
- 14Ca - Tortulo-Koelerion droge duingraslanden
- 16Aa - Junco-Molinion blauwgrasland
- 16Bb - Arrhenatherion elatioris glanshavergraslanden
- 19Aa - Nardo-Galion saxatilis heischraalgrasland
- 20Aa - Calluno-Genistion pilosae droge heide
- 26Ac - Armerion maritimae kwelder
- 39Aa - Alnion glutinosae broekbos
- 43Ab - Carpinion betuli eiken-haagbeukenbos

Voor de vergelijking is gebruik gemaakt van:

1. De actuele verspreiding van de gemeenschappen op basis van de vegetatieopnamen uit de Landelijke Vegetatie Databank (de feitelijke trainingset).
2. De geschatte verspreiding op basis van de FLORON-database (zie par. 2.2.3).

Omdat de verspreidingskaarten van SynBioSys een resolutie hebben van 1x1 km zijn de Dynterra-kaarten opgeschaald van 25x25m naar 1x1km, waarbij iedere km-cel de hoogste schattingswaarde van alle daartoe behorende 25x25 cellen heeft toegewezen gekregen.

De analyse bestond hieruit dat per verbond, van iedere opname of ieder kmhok met een FLORBASE-schatting > 60% bepaald is wat op die plek (kmhok) de Dynterra-score was. Van deze scores zijn cumulatieve frequenties bepaald die in aparte grafieken zijn opgenomen. Tevens wordt voor ieder verbond in twee kaarten de verspreiding getoond op basis van de opnamen, gecombineerd met de FLORBASE-schatting en de verspreiding op basis van Dynterra-scores.

En wordt in deze analyse aangenomen dat de opnamen/FLORBASE-schatting een betrouwbaar beeld geven van de verspreiding van de verbonden.

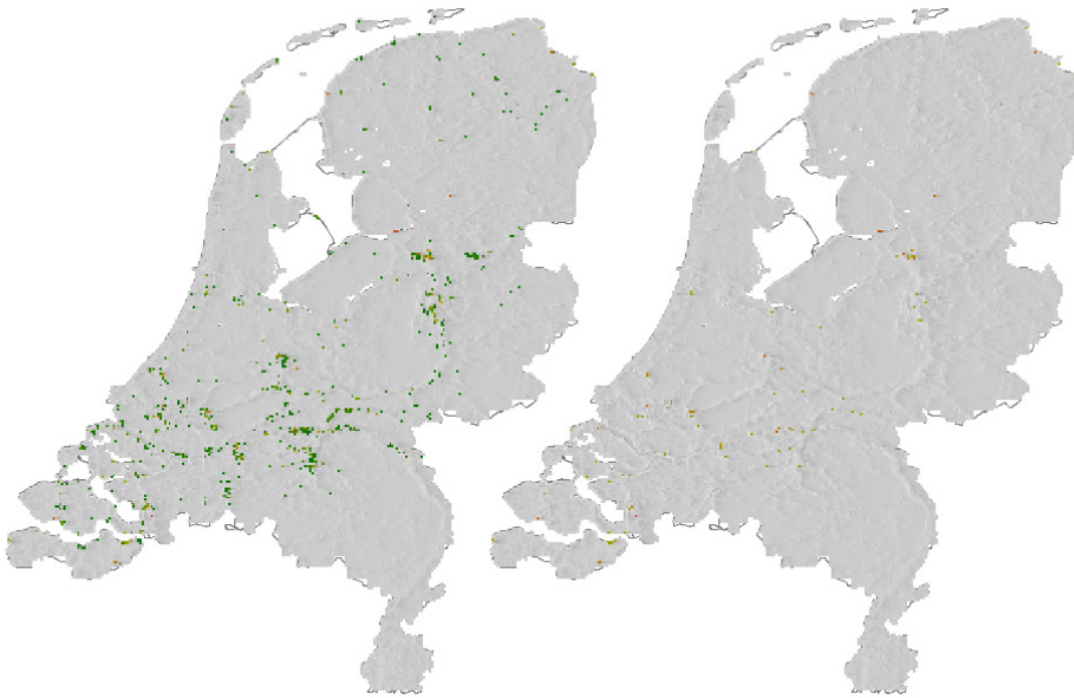
Wat in bijna alle gevallen opvalt is dat de trefkans van een verbond (op basis van opnamen of FLORBASE-schatting) nauwelijks samenvalt met hoge Dynterra-scores. Een uitzondering daarop zijn de verbonden 20Aa (droge heide) en 43Ab (eiken-haagbeukenbos). Beide vegetatietypen komen in Nederland met relatief grote aaneengesloten oppervlakten voor, in tegenstelling tot de andere verbonden die op veel plaatsen met enkele hectaren vertegenwoordigd zijn.

Vergelijken we de verspreidingspatronen gebaseerd op opnamen/FLORBASE-schatting en Dynterra-scores dan is er in een aantal gevallen een duidelijk discrepantie, soms een overlap, en in enkele gevallen een duidelijk overeenkomst.

Duidelijke verschillen in kaartbeelden zien we bij 06Ac (matig voedselarme vennen), 16Aa (blauwgraslanden) en 26Ac (kwelders). In het laatste geval wordt door het Dynterra-model ook een groot aantal plekken in het binnenland als kansrijk aangegeven, terwijl hier toch geen sprake is van saline bodems. Blijkbaar is de component zout niet goed verwerkt in het Dynterra-model.

Verbond 14Ca (droge duingraslanden) worden door het Dynterra-model goed voorspeld voor de kust, maar ook in het oosten van het land langs de grote rivieren. Dit is op zich niet vreemd aangezien er bodemkundig/ecologisch een hoge mate van overeenkomst bestaat tussen zangronden in de duinen en die langs de grote rivieren.

Kijken we naar de landelijke verspreiding van 16Bb (hooilanden met Frans raaigras - Figuur 3) dan valt het op dat er een hoge mate van overeenkomst bestaat tussen de patronen van beide datasets. Echter de cumulatieve frequenties (Figuur 41 pag. 51) laten zien dat kansrijke locaties volgens FLORBASE/opnamen samenvallen met relatief lage scores van Dyntaxa (tussen de 100 en 400; de hoogste Dyntaxa-score is 842). De verspreidingskaarten hieronder laten zien dat het probleem met de Dyntaxa-schaling is dat er maar weinig kmhokken zijn met een score > 400, en kmhokken met een waarde > 600 zijn al helemaal schaars. Daardoor ontstaat het beeld dat het Dyntaxa-model in het onderhavige geval geen goede predictie kan geven, in vergelijking met de verspreiding van opnamen/FLORBASE-schatting.



Figuur 2 16Bb *Arrhenatherion elatioris*: links: alle kmhokken met Dyntaxa score > 400; rechts: alleen kmhokken met Dyntaxa score > 400 (maximale waarde is 842)

Passen we een transformatie toe op de Dyntaxa-scores ($\log_{10}(x + 1) * 10$) dan verschuift de grafiek met cumulatieve frequenties naar rechts en ontstaat wél een beeld dat de vegetatieopnamen toegekend aan 16Bb redelijk goed samenvallen met hogere Dyntaxa-scores.

Conclusies:

- Op Dyntaxa-scores zou een log transformatie moeten worden toegepast om een meer evenwichtige schaalverdeling te krijgen;
- Het Dyntaxa-model werkt het beste voor typen die over grote oppervlakten voorkomen;
- De factor zout is niet goed verwerkt in het model;
- Typen met een complexe hydrologie zijn lastig te voorspellen met het Dyntaxa-model;
- De vereenvoudigde indeling van bodemtypen moet opnieuw worden beschouwd en een indeling op meer ecologisch grondslag worden doorgevoerd.

6 Conclusies en aanbevelingen

De bouw van een statistisch model voor vegetatietype op basis van een dataset met een groot aantal abiotische verklarende variabelen wordt vooral bemoeilijkt door de omvang van de dataset en in het bijzonder het grote aantal missing values en het grote aantal klassen van sommige kwalitatieve variabelen. Om deze redenen bleek het niet mogelijk om een regressiemodel te bouwen, wat in dit geval een multinomiaal logistisch model zou moeten zijn omdat de responsvariabele vegetatietype multinomiaal verdeeld is.

Multivariate analyses zijn niet of moeilijk toe te passen doordat er zowel kwalitatieve als kwantitatieve variabelen voorkomen en er veel missing values zijn.

Een Random-Forest model blijkt een goed alternatief te zijn. Maar ook hier loop je tegen numerieke grenzen aan in verband met de omvang van de dataset. Deze problemen zijn vermeden door de dataset in twee delen te splitsen: een deel waarvoor de Gt-code bekend is en een deel waarvoor de Gt-code niet bekend is.

De resultaten laten zien dat de verwachte kaartzuiverheid (percentage correct) rond de 30-35% zal liggen voor classificatie op associatieniveau. Dit is laag, maar bedacht moet worden dat het om predictie van associaties gaat (laag aggregatieniveau, 215 klassen) en dat de mate van misclassificatie niet in het percentage correct is verdisconteerd: een kleine fout telt net zo zwaar als een grote fout. Bij predicties op een hoger aggregatieniveau (verbond, 87 klassen) neemt het percentage correct iets toe.

Als de random-forestmethode wordt toegepast om predicties te doen of één bepaald vegetatietype wel of niet op een locatie voorkomt, dan worden hoge kaartzuiverheden bereikt. Hierbij moet echter een kanttekening worden geplaatst: het percentage correct is zo hoog omdat het model heel goed aangeeft dat het vegetatietype ergens niet voorkomt. Het percentage correct is daarom ook per klasse 'wel' en 'niet' aangegeven (i.e., consumer's accuracy, zie tabel 1).

De modellering is waarschijnlijk nog op een aantal punten te verbeteren:

1. De dataset is nu in twee delen gesplitst op basis van wel/geen Gt-code. De reden is dat het numeriek niet mogelijk is om de gehele dataset in één keer te analyseren. Onderzocht zou kunnen worden of een andere splitsing betere resultaten geeft.
2. AHN-hoogte blijkt een belangrijke variabele te zijn. Mogelijk kan het model verder worden verbeterd door relatieve maaiveldhoogte als predictorvariabele toe te voegen, dus de hoogteligging relatief tot een omgeving met een beperkte straal. Beekdalen etc. komen dan beter tot hun recht.
3. Met de huidige package in R, randomForest, treedt een bias op in de variabele-selectie wanneer variabelen verschillende schalen of verschillende aantallen categorieën hebben. Er is inmiddels echter een nieuw package verschenen, "party", waarmee dit waarschijnlijk verholpen is.

Literatuur

- Breiman, L., 2001. Random Forests. *Machine Learning* 45: 5-32.
- De Winter, W.; Hennekens, S., Mucher, S., 2011. Ruimtelijke modellering plantengemeenschappen. Casestudie voor de IJsselvallei. Deelrapportage KB-IV Veerkracht van Ecosystemen (WP3). Versie 1.0
- Groen, C.L.G.; Gorree, M.; Meijden, R. van der; Huele, R.; Zelfde, M. van 't, 1992. FLORBASE : een bestand van de Nederlandse flora, periode 1975-1990. Institute of Environmental Sciences, CML, Leiden.
- Hastie, T., Tibshirani en J. Friedman, 2009. The elements of statistical learning. Data mining, inference, and prediction. Second Edition. Springer, New York.
- Hazeu, G.W., 2005. Landelijk Grondgebruiksbestand Nederland (LGN5). Vervaardiging, nauwkeurigheid en gebruik. Wageningen, Alterra. Alterra-report 1213, 92p., 18 figs., 11 tables and 11 refs.
- Hazeu, G.W., 2005. Landelijk Grondgebruiksbestand Nederland (LGN5); monitoring landgebruik van 1995 -2004. *Geo-Info*, 10(2), 456-462.
- Hazeu, G.W., 2006. Land use mapping and monitoring in the Netherlands (LGN5). 2nd EARSeL Workshop on Land Use and Land Cover, 28-30 September 2006, Bonn, Germany. Conference proceedings.
- Hazeu, G.W., Schuiling, C., Dorland, G.J., Oldengarm, J. & H.A. Gijsbertse, 2010. Landelijk Grondgebruiksbestand Nederland versie 6 (LGN6). Vervaardiging, nauwkeurigheid en gebruik. Wageningen, Alterra. Alterra-rapport 2012, 132p., 20 figs., 9 tables and 15 refs.
- Hosmer, D.W. en S. Lemeshow, 2000. Applied logistic regression. Second edition. Wiley, New York etc.
- Peters, J., B. De Baets, N.E.C. Verhoest, R. Samson, S. Degroeve, P. De Becker en W. Huybrechts, 2007. Random forests as a tool for ecohydrological distribution modelling. *Ecological Modelling* 207: 304-318.
- F. de Vries, W.J.M. de Groot, T. Hoogland & J. Denneboom, 2003. De Bodemkaart van Nederland digitaal;. Toelichting bij inhoud, actualiteit en methodiek en korte beschrijving van additionele informatie. Wageningen, Alterra rapport 811.
- Wit, A.J.W. de, Heijden, Th.G.C. van der en H.A.M. Thunnissen, 1999. Vervaardiging en nauwkeurigheid van het LGN3-grondgebruiksbestand. Wageningen, DLO-Staring Centrum, Rapport 663.

Bijlage 1 LGN-4

code	hercodering	klasse	definitie
1	1	Agrarisch gras	Grasland binnen het stratum agrarisch gebied. Dit betreft voor het overgrote deel grasland dat gebruikt wordt voor agrarische productie, maar voor een deel betreft het ook erven van boerderijen en bedrijven, gras op dijken, wegbermen en andere met gras bedekte oppervlakken.
2	2	Mais	Agrarische percelen met het gewas maïs.
3	300	Aardappelen	Agrarische percelen met het gewas aardappelen. Hierbinnen wordt geen onderscheid gemaakt tussen pootaardappelen, consumptieaardappelen en fabrieksaardappelen.
4	300	Bieten	Agrarische percelen met het gewas suikerbieten. Deze klasse bevat zowel suikerbieten als voederbieten, maar geen 'rode bieten' die als tuinbouwgewas in de klasse 'overige landbouwgewassen' vallen.
5	300	Granen	Agrarische percelen met het gewas graan. Een verzamelklasse voor alle graangewassen: tarwe, gerst, have, rogge, enz. waarbij geen onderscheid is gemaakt naar zomergranen of wintergranen.
6	300	Overige gewassen	Bevat alle landbouwgewassen die niet binnen de voorgaande klassen vallen en niet tot de klasse 'bloembollen' behoren: Tuinbouwgewassen, boomkwekerijen, koolgewassen, hennep, koolzaad, enz.
8	8	Glastuinbouw	Glastuinbouw
9	9	Boomgaarden	Boomgaarden, zonder onderscheid naar hoogstam of laagstam en het type vrucht.
10	300	Bloembollen	Percelen met bloembollen. Hierbij is geen onderscheid gemaakt naar het type bloembol en ook geen onderscheid tussen voorjaars- of najaarsbollen.
11	11	Loofbos	Loofbos buiten het stratum stedelijk gebied. Deze klasse kan een natuurfunctie hebben, maar dit zal in veel gevallen niet het geval zijn.
12	12	Naaldbos	Naaldbos buiten het stratum stedelijk gebied. Deze klasse kan een natuurfunctie hebben, maar dat zal in veel gevallen niet het geval zijn.
16	16	Zoet water	
17	17	Zout water	
18	301	Stedelijk bebouwd gebied	Continue stedelijke bebouwing, inclusief straten, pleinen, kleine plantsoenen, smalle kanalen en tuinen. Een duidelijk voorbeeld is het centrum van Amsterdam waarbinnen geen kanalen of wegen te onderscheiden zijn. De grote parken (Vondelpark, Sarphatipark) zijn wel als park in LGN3 aanwezig, kleine plantsoenen (Marnixplantsoen) niet.
19	301	Bebouwing in buitengebied	Bebouwing buiten het stedelijk gebied die geen agrarische functie heeft. Hierbij moet worden gedacht aan: Gebouwen en landingsbanen van vliegvelden, Bungalowparken en campings, gebouwen in militaire terreinen, gebouwen van de elektriciteitsvoorziening, waterzuiveringsinstallaties, gebouwen in natuurgebieden, maar ook woningen in bossen.
20	20	Loofbos in bebouwd gebied	Loofbos binnen het stratum stedelijk gebied, dit betreft o.a. parken, recreatieterreinen en stroken bos langs wegen binnen stedelijk gebied.
21	21	Naaldbos in bebouwd gebied	Naaldbos binnen het stratum stedelijk gebied, dit betreft o.a. parken, recreatieterreinen en stroken bos langs wegen binnen stedelijk gebied.
22	22	Bos met dichte bebouwing	Dit is een stedelijke klasse waarbij de huizen in een dusdanig dicht bebost gebied geplaatst zijn, dat het gebied op een satellietbeeld niet is te onderscheiden van een gewoon bos. Een goed voorbeeld van deze klasse is Wageningen-Hoog (tussen Wageningen en Bennekom).
23	23	Gras in bebouwd gebied	Met gras bedekte gebieden met een stedelijke functie. Dit betreft parken, sportterreinen, recreatiegebieden, golfterreinen, maar bijvoorbeeld ook gras rond vliegvelden. Bij het vervaardigen van LGN3 is er bovendien voor gekozen om woonwijken in aanbouw (waar dus nog geen huizen staan, maar wel graafwerkzaamheden zijn begonnen) als 'gras in bebouwd gebied' te classificeren. Bij de volgende revisie van het LGN-bestand zullen deze woonwijken als 'stedelijk gebied' worden geclassificeerd.
24	24	Kale grond in bebouwd buitengebied	Dit is een kleine klasse die slechts in beperkte mate voorkomt, in de meeste gevallen betreft dit bouwputten en bijvoorbeeld de kalkgroeves in Zuid-Limburg.
25	25	Hoofdwegen en spoorwegen	De wegen in het LGN3-bestand zijn de snelwegen (A-wegen) en de overige 'doorgaande' wegen (N-wegen). Wegen en straten binnen woonwijken zijn niet als zodanig geklassificeerd en behoren bij de klasse 'Stedelijk bebouwd gebied'. Alle spoorlijnen in Nederland bevinden zich in het LGN3-bestand.

code	hercodering	klasse	definitie
26	301	Bebouwing in agrarisch gebied	Bebouwing binnen het stratum agrarisch gebied. Soms is de scheiding tussen 'bebouwing in agrarisch gebied' en 'stedelijk bebouwd gebied' nogal arbitrair. Dit is duidelijk te zien aan de lintbebouwing in het Groene Hart, waarbij sommige delen van het lint als 'bebouwing in agrarisch gebied' en andere delen als 'stedelijk bebouwd gebied' zijn geclassificeerd. Dit is ontstaan door de opdeling van strata in het oorspronkelijke BARS bestand (Basis Bestand Ruimtelijke Structuren) van de Rijksplanologische Dienst.
28	28	Gras in secundair bebouwd gebied	
30	30	Kwelders	Kwelders in beheer als natuurgebied.
31	31	Open zand in kustgebied	Gebieden langs de kust zonder vegetatie: stranden en open duinvalleien.
32	303	Open duinvegetatie	Duingebieden met een open vegetatiedek.
33	303	Gesloten duinvegetatie	Duingebieden met een gesloten vegetatiedek.
34	34	Duinheide	Duingebieden met een vegetatiedek van droge heide.
35	35	Open stuifzand	Open stuifzand
36	36	Heide	Heidegebieden met een vergrassing van minder dan 25%.
37	302	Matig vergraste heide	Gebieden met een vergrassing tussen de 25% en 75%.
38	302	Sterk vergraste heide	Gebieden met een vergrassing groter dan 75%.
39	39	Hoogveen	Levend, onvergraven hoogveen. Enkele kleine, sterk verdroogde veenrestanten zijn niet als zodanig geclassificeerd.
40	40	Bos in hoogveengebied	De bossen die zich binnen de hoogveengebieden bevinden zijn naar een code 40 gehercodeerd om te voorkomen dat er 'gaten' vallen in de hoogvenen. Door het samenvoegen van klassen 39 en 40 is het dus mogelijk om alle hoogveengebieden in Nederland te selecteren. Er is geen onderscheid gemaakt tussen naaldbos en loofbos.
41	41	Overige moerasvegetatie	Vegetatie in moerasgebieden die niet tot de klassen 42 en 43 behoren.
42	42	Rietvegetatie	Gebieden binnen moerassen die met riet begroeid zijn.
43	43	Bos in moerasgebied	De bossen die zich binnen de moerasgebieden bevinden zijn naar een code 43 gehercodeerd om te voorkomen dat er 'gaten' vallen in de moerasgebieden. Door het samenvoegen van klassen 41,42 en 43 is het dus mogelijk om alle moerasgebieden in Nederland te selecteren. Er is geen onderscheid gemaakt tussen naaldbos en loofbos, maar naaldbos komt vrijwel niet voor binnen moerasgebieden.
44	44	Veenweidegebieden	In Noord-Holland is een aantal veenweidegebieden die als natuurgebied beheerd worden naar code 44 gehercodeerd. Het is belangrijk om te realiseren dat dus niet alle veenweidegebieden als zodanig in LGN3plus aanwezig zijn. Veel gebieden in het Groene Hart komen in aanmerking voor de term 'veenweidegebied'. In dit geval is het dus de beheersinformatie die bepaald of een gebied als veenweidegebied in LGN3plus aanwezig is.
45	45	Overig open begroeid natuurgebied	Dit is een restklasse waarin natuurgebieden vallen die niet passen in de eerder genoemde klassen. In veel gevallen zijn die extensief beheerde graslanden of voormalige kweldergebieden (o.a. Lauwersmeer).
46	46	Kale grond in natuurgebied	Dit is een restklasse waarin delen van natuurgebieden vallen die niet of spaarzaam begroeid zijn, maar die niet direct tot de categorie stuifzanden behoren.

Bijlage 2 Verbonden

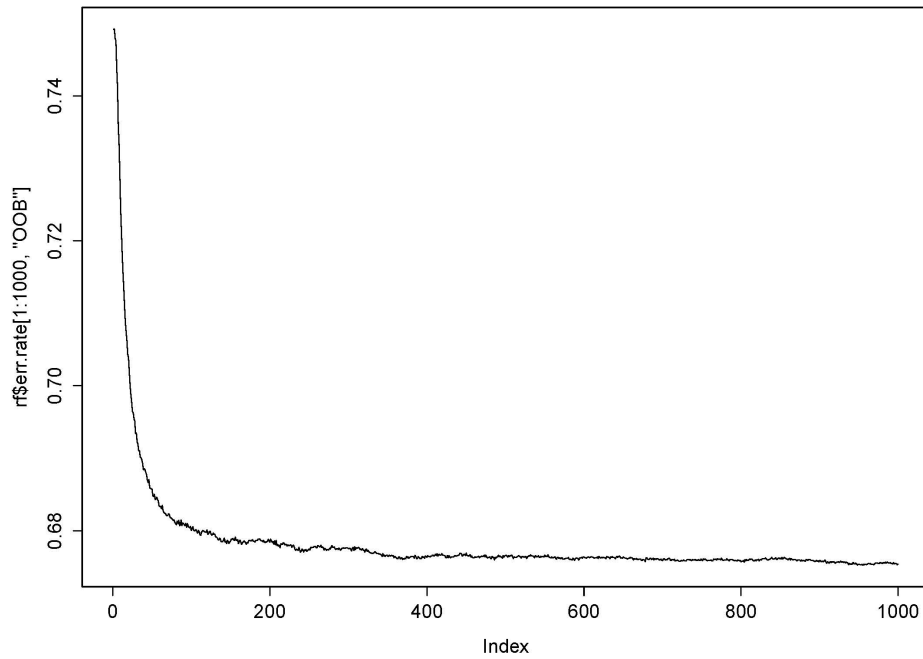
Code	Verbond	Verbond_nl
01AA	Lemnion minoris	Bultkroos-verbond
01AB	Lemnion trisulcae	Puntkroos-verbond
02AA	Ruppion maritimae	Ruppia-verbond
03AA	Zosterion	Zeegras-verbond
04AA	Nitellion flexilis	Glanswier-verbond
04BA	Charion fragilis	Verbond van Stekelharig kransblad
04BB	Charion vulgaris	Verbond van Gewoon kransblad
04CA	Charion canescentis	Verbond van Brakwater-kransblad
05AA	Zannichellion pedicellatae	Verbond van Gesteelde zannichellia
05BA	Nymphaeion	Waterlelie-verbond
05BB	Hydrocharition morsus-ranae	Kikkerbeet-verbond
05BC	Parvopotamion	Verbond der kleine fonteinkruiden
05CA	Ranunculion peltati	Verbond van Grote waterranonkel
06AA	Littorellion uniflorae	Oeverkruid-verbond
06AB	Potamion graminei	Verbond van Ongelijkbladig fonteinkruid
06AC	Hydrocotylo-Baldellion	Verbond van Waternavel en Stijve moerasweegbree
06AD	Eleocharition acicularis	Naaldwaterbies-verbond
07AA	Cardamino-Montion	Verbond van Bittere veldkers en Bronkruid
08AA	Sparganio-Glycerion	Vlotgras-verbond
08AB	Oenanthion aquaticae	Watertorkruid-verbond
08BA	Cicution virosae	Waterscheerling-verbond
08BB	Phragmition australis	Riet-verbond
08BC	Caricion gracilis	Verbond van Scherpe zegge
08BD	Caricion elatae	Verbond van Stijve zegge
09AA	Caricion nigrae	Verbond van Zwarte zegge
09BA	Caricion davallianae	Knobbies-verbond
10AA	Rhynchosporion albae	Verbond van Veenmos en Snavelbies
10AB	Caricion lasiocarpae	Draadzegge-verbond
11AA	Ericion tetralicis	Dophei-verbond
11BA	Oxycocco-Ericion	Hoogveenmos-verbond
12AA	Polygonion avicularis	Varkensgras-verbond
12BA	Lolio-Potentillion anserinae	Zilverschoon-verbond
13AA	Alyso-Sedion	Verbond van Vetkruiden en Kandelaartje
14AA	Corynephorion canescentis	Buntgras-verbond
14BA	Thero-Airion	Dwerghaver-verbond
14BB	Plantagini-Festucion	Verbond van Gewoon struisgras
14BC	Sedo-Cerastion	Verbond der droge stroomdalgraslanden
14CA	Tortulo-Koelerion	Duinsterretjes-verbond
14CB	Polygalo-Koelerion	Verbond der droge, kalkrijke duingraslanden
15AA	Mesobromion erecti	Verbond der matig droge kalkgraslanden
16AA	Junco-Molinion	Verbond van Biezeknoppen en Pijpestrootje
16AB	Calthion palustris	Dotterbloem-verbond
16BA	Alopecurion pratensis	Verbond van Grote vossestaart
16BB	Arrhenatherion elatioris	Glanshaver-verbond
16BC	Cynosurion cristati	Kamgras-verbond
17AA	Trifolion medii	Marjolein-verbond
18AA	Melampyrion pratensis	Verbond van Gladde witbol en Havikskruiden
19AA	Nardo-Galion saxatilis	Verbond der heischrale graslanden
20AA	Calluno-Genistion pilosae	Verbond van Struikhei en Kruipbrem
20AB	Empetrium nigri	Kraaihei-verbond
21AA	Parietario judaicae	Verbond van Klein glaskruid
21AB	Cymbalario-Asplenion	Muurleeuwebek-verbond
22AA	Atriplicion littoralis	Strandmelde-verbond
22AB	Salsolo-Honkenyion peploides	Loogkruid-verbond
23AA	Agropyro-Honkenyion peploides	Biestarwegras-verbond
23AB	Ammophilion arenariae	Helm-verbond
24AA	Spartinion	Slijkgras-verbond
25AA	Thero-Salicornion	Zeekraal-verbond
26AA	Puccinellion maritimae	Verbond van Gewoon kweldergras
26AB	Puccinellio-Spergularion salinae	Verbond van Stomp kweldergras
26AC	Armerion maritimae	Verbond van Engels gras
27AA	Saginion maritimae	Zeevetmuur-verbond
28AA	Nanocyperion flavescens	Dwergbiezen-verbond
29AA	Bidention tripartitae	Tandzaad-verbond

Code	Verbond	Verbond_nl
30AB	Fumario-Euphorbion	Verbond van Duivekervel en Kroontjeskruid
30BA	Aperion spicae-venti	Windhalm-verbond
30BB	Digitario-Setarion	Verbond van Vingergras en Naalbaar
31AA	Salsolion ruthenicae	Verbond van Kleverig kruiskruid
31AB	Arction	Kaasjeskruid-verbond
31BA	Onopordion acanthii	Verbond van Distels en Ruwbladigen
31CA	Dauco-Melilotion	Wormkruid-verbond
32AA	Filipendulion	Moerasspirea-verbond
32BA	Epilobion hirsuti	Verbond van Harig wilgeroosje
33AA	Galio-Alliarion	Verbond van Look-zonder-look
34AA	Carici piluliferae-Epilobion angustifolii	Wilgeroosjes-verbond
35AA	Lonicero-Rubion silvatici	Brummel-verbond
36AA	Salicion cinereae	Verbond der wilgenbroekstruwelen
37AA	Pruno-Rubion radulae	Verbond van Bramen en Sleedoorn
37AB	Carpino-Prunion	Verbond van Sleedoorn en Eenstijlige meidoorn
37AC	Berberidion vulgaris	Liguster-verbond
38AA	Salicion albae	Verbond der wilgenvloedbossen en -struwelen
39AA	Alnion glutinosae	Verbond der elzenbroekbossen
40AA	Betulion pubescentis	Verbond der berkenbroekbossen
41AA	Dicrano-Pinion	Verbond der naaldbossen
42AA	Quercion roboris	Zomereikverbond
42AB	Luzulo-Fagion	Verbond der Veldbies-Beukenbossen
43AA	Alno-Padion	Verbond van Els en Vogelkers
43AB	Carpinion betuli	Haagbeuken-verbond

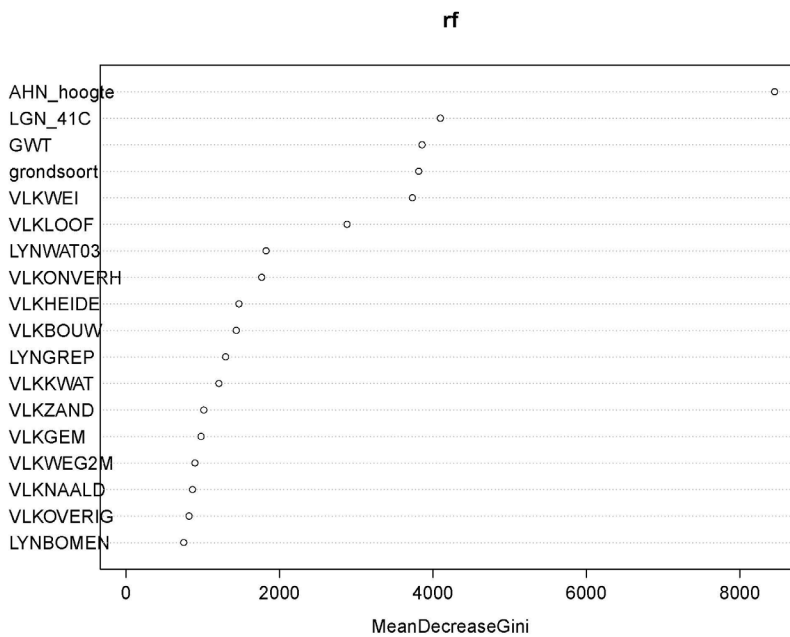
Bijlage 3 Figuren

Figuur 1	Variable importance van het random-forestmodel voor de punten met Gt-code, met alle predictorvariabelen opgenomen in het model	15
Figuur 2	Rangnummercorrespondentie in het proefgebied, direct op basis van de predictie ("abiotisch", geel) en na combinatie met kansen op basis van soortverspreidingsdata ("met soortsdata", groen)	15
Figuur 3	16Bb Arrhenatherion elatioris: links: alle kmhokken met Dynterra score > 400; rechts: alleen kmhokken met Dynterra score > 400 (maximale waarde is 842).....	17
Figuur 4	OOB-error rate versus aantal bomen.....	26
Figuur 5	Variable importance van het random-forestmodel voor de punten met Gt-code, met 18 meest belangrijke variabelen.	26
Figuur 6	Variable importance van het random-forestmodel voor de punten zonder Gt-code, met alle predictorvariabelen opgenomen in het model.....	27
Figuur 7	Variable importance van het random-forestmodel voor de punten zonder Gt-code, met de zeven belangrijkste predictorvariabelen opgenomen in het model.....	27
Figuur 8	Variable importance van het random-forestmodel op verbondsniveau voor de punten met Gt-code, met alle predictorvariabelen opgenomen in het model.....	28
Figuur 9	Variable importance van het random-forestmodel op verbondsniveau voor de punten met Gt-code, met de belangrijkste predictorvariabelen opgenomen in het model.	28
Figuur 10	Variable importance van het random-forestmodel op verbondsniveau voor de punten zonder Gt-code, met alle predictorvariabelen opgenomen in het model.	29
Figuur 11	Variable importance van het random-forestmodel op verbondsniveau voor de punten zonder Gt-code, met de belangrijkste predictorvariabelen opgenomen in het model.	29
Figuur 12	Variable importance van het random-forestmodel voor wel/geen 16BB, voor punten met Gt-code, met alle predictorvariabelen in het model.....	30
Figuur 13	Variable importance van het random-forestmodel voor wel/geen 16BB, voor punten met Gt-code, met de belangrijkste predictorvariabelen in het model.	30
Figuur 14	Variable Importance van het random-forestmodel wel/geen 16BB, voor punten zonder Gt-code, alle predictorvariabelen.	31
Figuur 15	Variable Importance van het random-forestmodel wel/geen 16BB, voor punten zonder Gt-code, belangrijkste predictorvariabelen.	31
Figuur 16	Variable Importance van het random-forestmodel wel/geen 16BA, voor punten met Gt-code, alle predictorvariabelen.	32
Figuur 17	Variable Importance van het random-forestmodel wel/geen 16BA, voor punten met Gt-code, belangrijkste predictorvariabelen.....	32
Figuur 18	Variable Importance van het random-forestmodel wel/geen 16BA, voor punten zonder Gt-code, alle predictorvariabelen.	33
Figuur 19	Variable Importance van het random-forestmodel wel/geen 16BA, voor punten zonder Gt-code, belangrijkste predictorvariabelen.	33
Figuur 20	Variable Importance van het random-forestmodel wel/geen 43AA, voor punten met Gt-code, alle predictorvariabelen.	34
Figuur 21	Variable Importance van het random-forestmodel wel/geen 43AA, voor punten met Gt-code, belangrijkste predictorvariabelen.....	34
Figuur 22	Variable Importance van het random-forestmodel wel/geen 43AA, voor punten zonder Gt-code, alle predictorvariabelen.	35

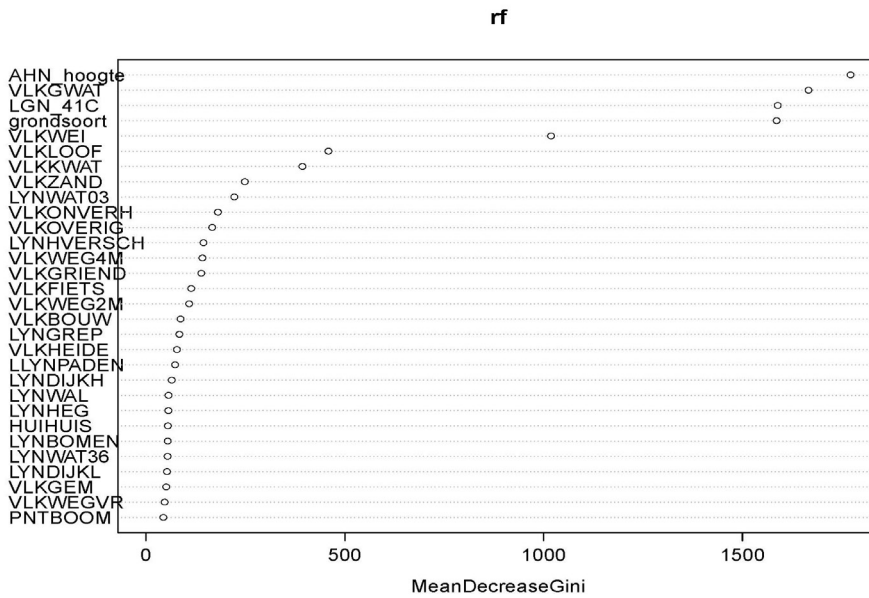
Figuur 23	Variable Importance van het random-forestmodel wel/geen 43AA, voor punten zonder Gt-code, belangrijkste predictorvariabelen.	35
Figuur 24	Variable Importance van het random-forestmodel op verbodsniveau, voor punten met Gt-code, alle predictorvariabelen.	36
Figuur 25	Variable Importance van het random-forestmodel op verbodsniveau, voor punten met Gt-code, belangrijkste predictorvariabelen.	36
Figuur 26	Variable Importance van het random-forestmodel op verbodsniveau, voor punten zonder Gt-code, alle predictorvariabelen.	37
Figuur 27	Variable Importance van het random-forestmodel op verbodsniveau, voor punten zonder Gt-code, belangrijkste predictorvariabelen.	37
Figuur 28	vegetatiekaart testgebied Cortenoever.	38
Figuur 29	hoogte van het rangnummer dat correspondeert met de gekarteerde vegetatie-eenheden.	39
Figuur 30	kanskaart 16BA (<i>Alopecurion pratensis</i>).	40
Figuur 31	kanskaart 16BB (<i>Arrhenatherion elatioris</i>).	41
Figuur 32	kanskaart 43AA (<i>Alno-Padion</i>).	42
Figuur 33	kanskaart 07AA (<i>Cardamino-Montion</i>).	43
Figuur 34	kanskaart 16BB (<i>Arrhenatherion elatioris</i>).	44
Figuur 35	kanskaart 39AA (<i>Alnion glutinosae</i>).	45
Figuur 36	kanskaart 41AA (<i>Dicrano-Pinion</i>).	46
Figuur 37	kanskaart 43AB (<i>Alno-Padion</i>).	47
Figuur 38	06Ac <i>Hydrocotylo-Baldellion</i> : a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (<i>Dynterra</i>); c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.	48
Figuur 39	14Ca <i>Tortulo-Koelerion</i> : a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (<i>Dynterra</i>); c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.	49
Figuur 40	16Aa <i>Junco-Molinion</i> : a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (<i>Dynterra</i>); c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.	50
Figuur 41	16Bb <i>Arrhenatherion elatioris</i> : a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (<i>Dynterra</i>); c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.	51
Figuur 42	19Aa <i>Nardo-Galion saxatilis</i> : a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (<i>Dynterra</i>); c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.	52
Figuur 43	20Aa <i>Calluno-Genistion pilosae</i> : a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (<i>Dynterra</i>); c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.	53
Figuur 44	26Ac <i>Armerion maritimae</i> : a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (<i>Dynterra</i>); c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.	54
Figuur 45	39Aa <i>Alnion glutinosae</i> : a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (<i>Dynterra</i>); c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.	55
Figuur 46	43Ab <i>Carpinion betuli</i> : a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (<i>Dynterra</i>); c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.	56



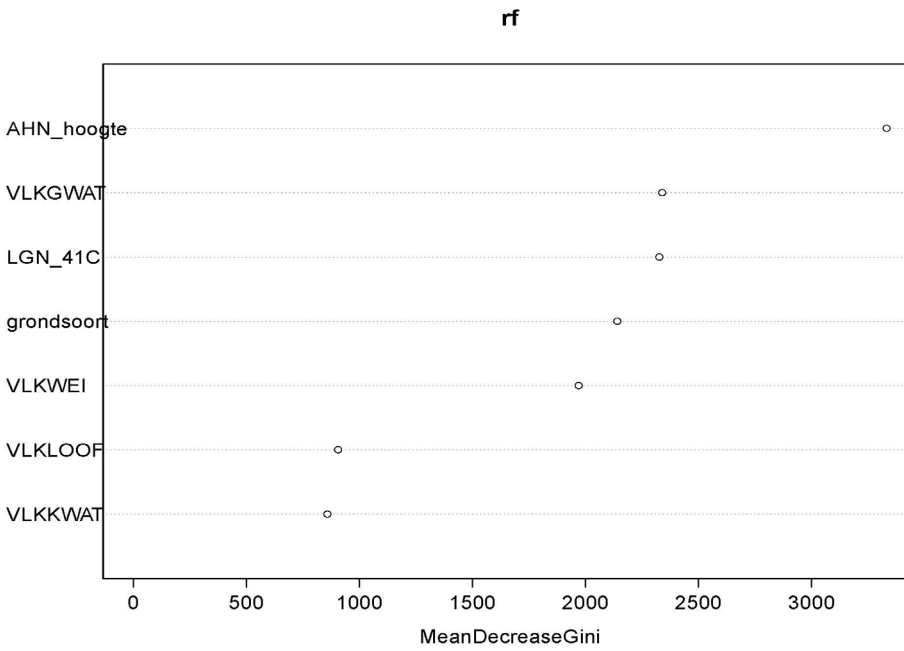
Figuur 3 OOB-error rate versus aantal bomen.



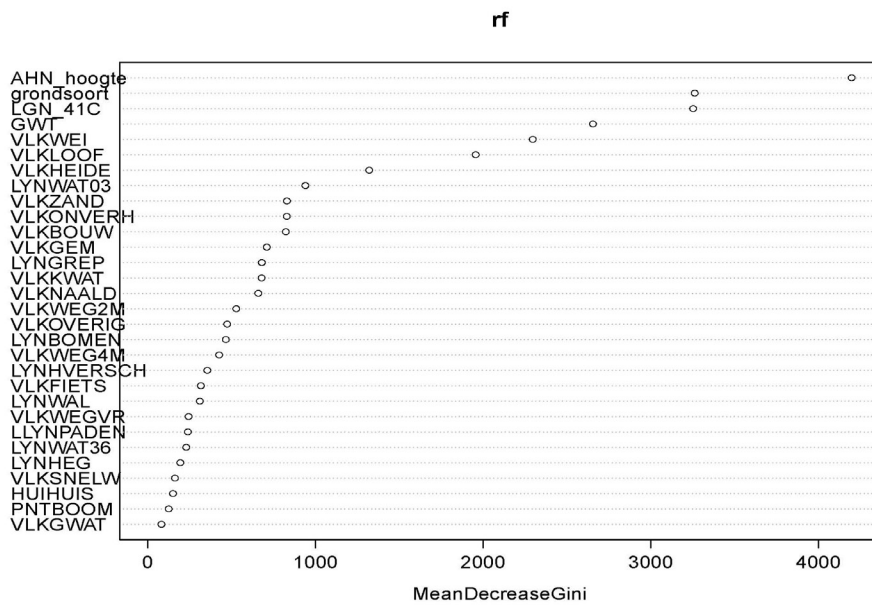
Figuur 4 Variable importance van het random-forestmodel voor de punten met Gt-code, met 18 meest belangrijke variabelen.



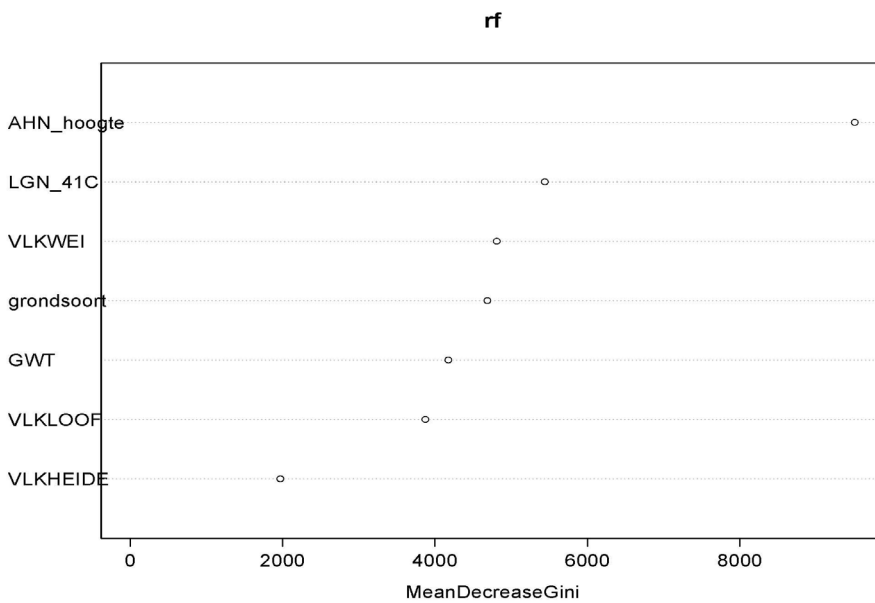
Figuur 5 Variable importance van het random-forestmodel voor de punten zonder Gt-code, met alle predictorvariabelen opgenomen in het model.



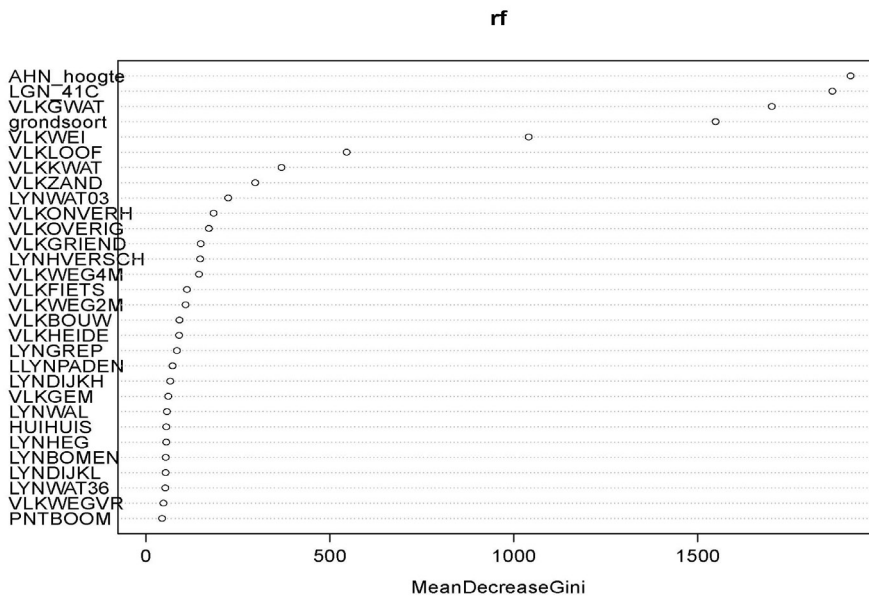
Figuur 6 Variable importance van het random-forestmodel voor de punten zonder Gt-code, met de zeven belangrijkste predictorvariabelen opgenomen in het model.



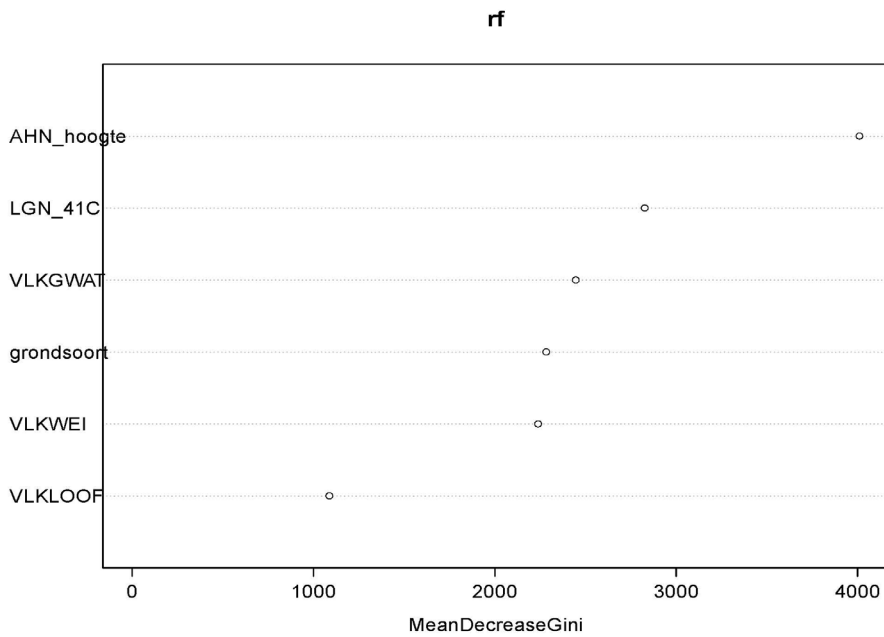
Figuur 7 Variable importance van het random-forestmodel op verbondsniveau voor de punten met Gt-code, met alle predictorvariabelen opgenomen in het model.



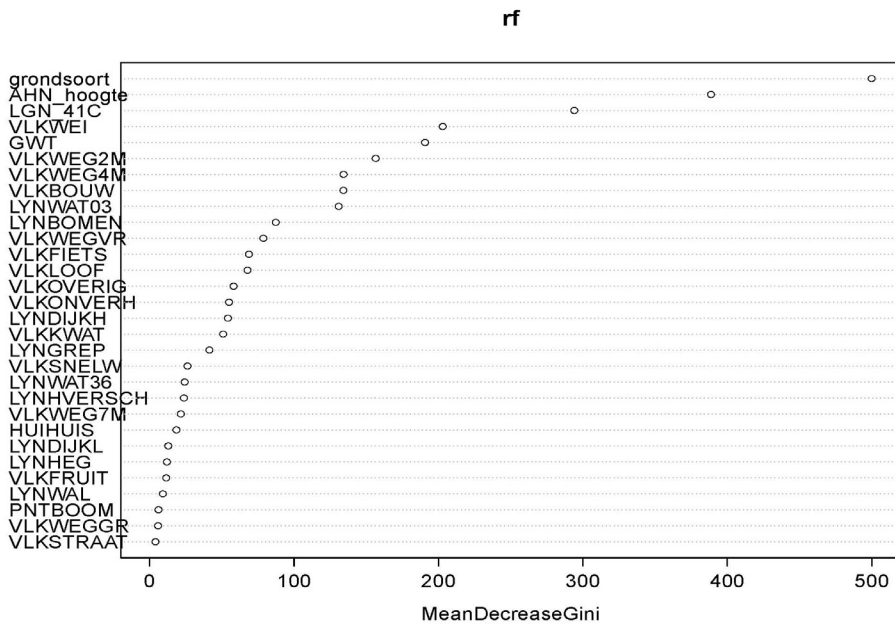
Figuur 8 Variable importance van het random-forestmodel op verbondsniveau voor de punten met Gt-code, met de belangrijkste predictorvariabelen opgenomen in het model.



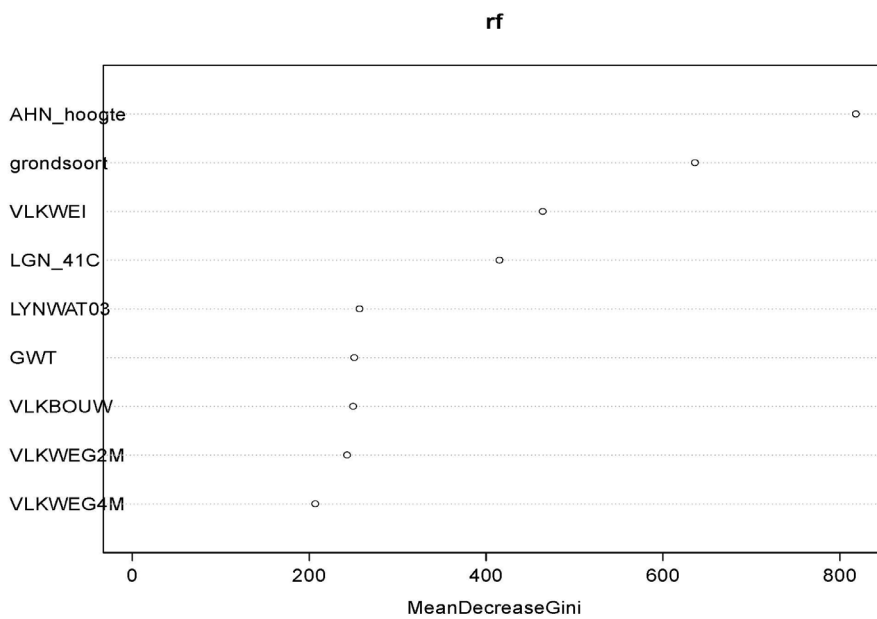
Figuur 9 Variable importance van het random-forestmodel op verbondsniveau voor de punten zonder Gt-code, met alle predictorvariabelen opgenomen in het model.



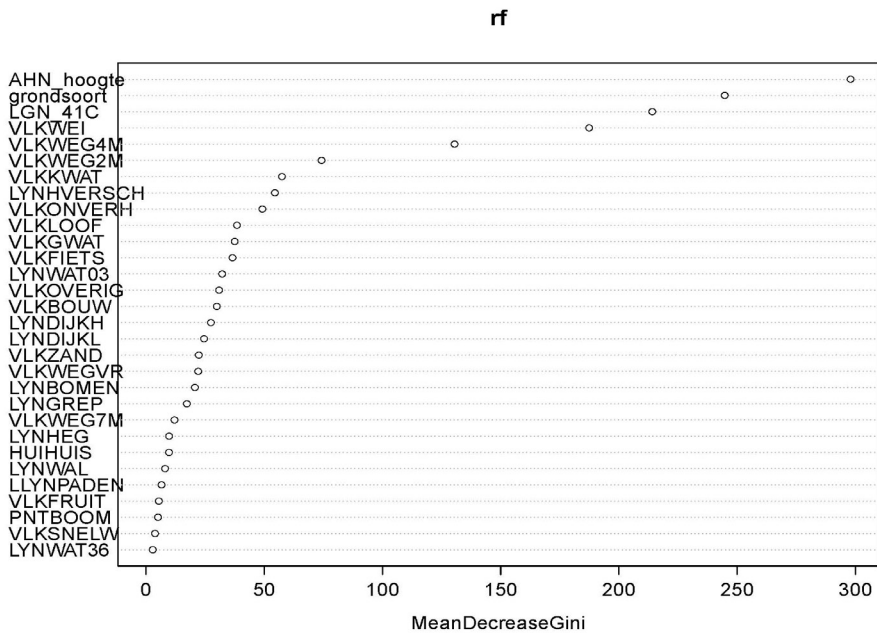
Figuur 10 Variable importance van het random-forestmodel op verbondsniveau voor de punten zonder Gt-code, met de belangrijkste predictorvariabelen opgenomen in het model.



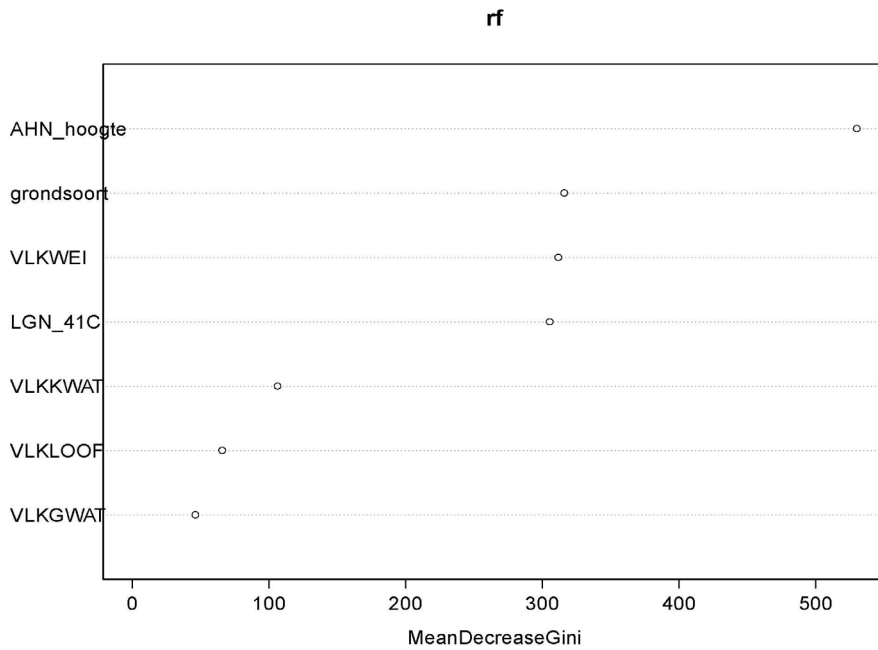
Figuur 11 Variable importance van het random-forestmodel voor wel/geen 16BB, voor punten met Gt-code, met alle predictorvariabelen in het model.



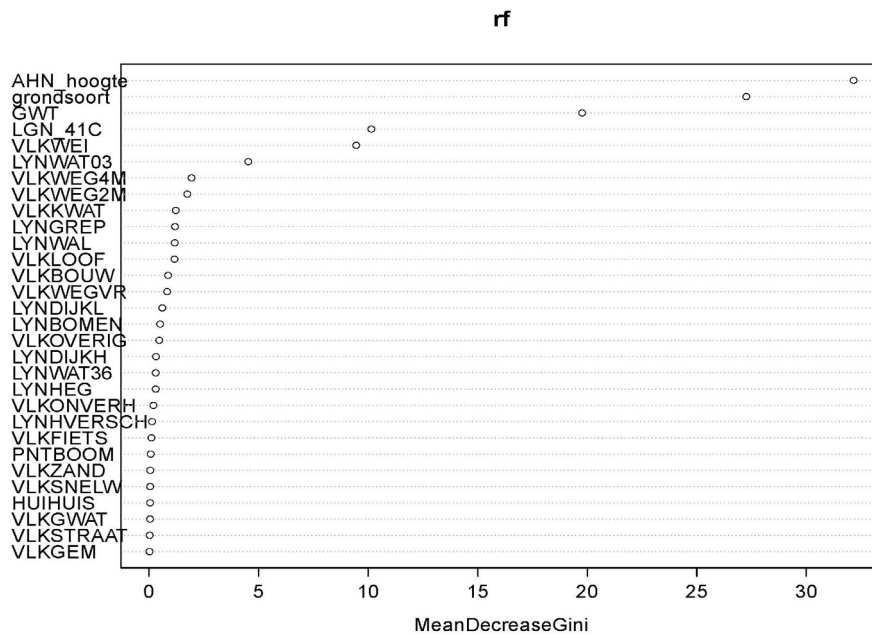
Figuur 12 Variable importance van het random-forestmodel voor wel/geen 16BB, voor punten met Gt-code, met de belangrijkste predictorvariabelen in het model.



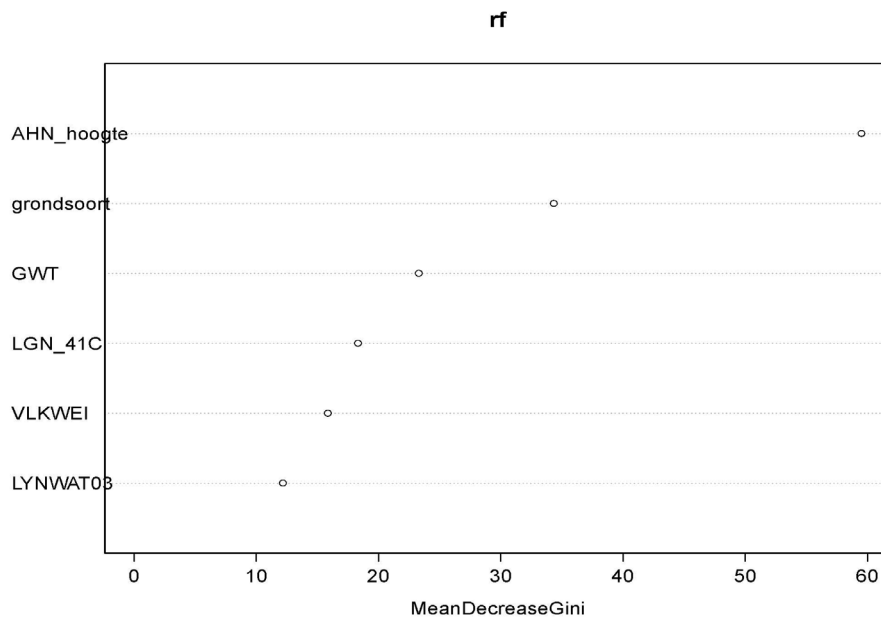
Figuur 13 Variable Importance van het random-forestmodel wel/geen 16BB, voor punten zonder Gt-code, alle predictorvariabelen.



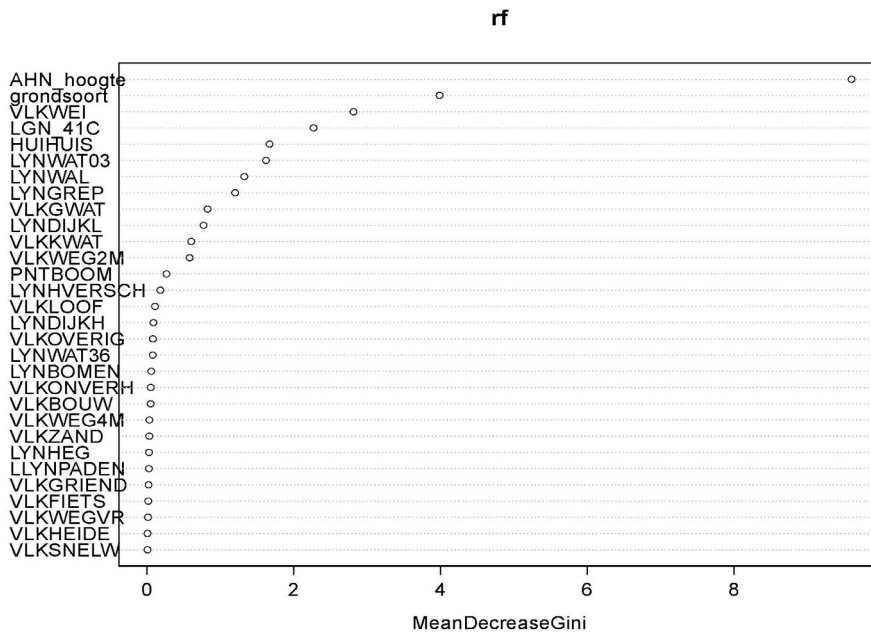
Figuur 14 Variable Importance van het random-forestmodel wel/geen 16BB, voor punten zonder Gt-code, belangrijkste predictorvariabelen.



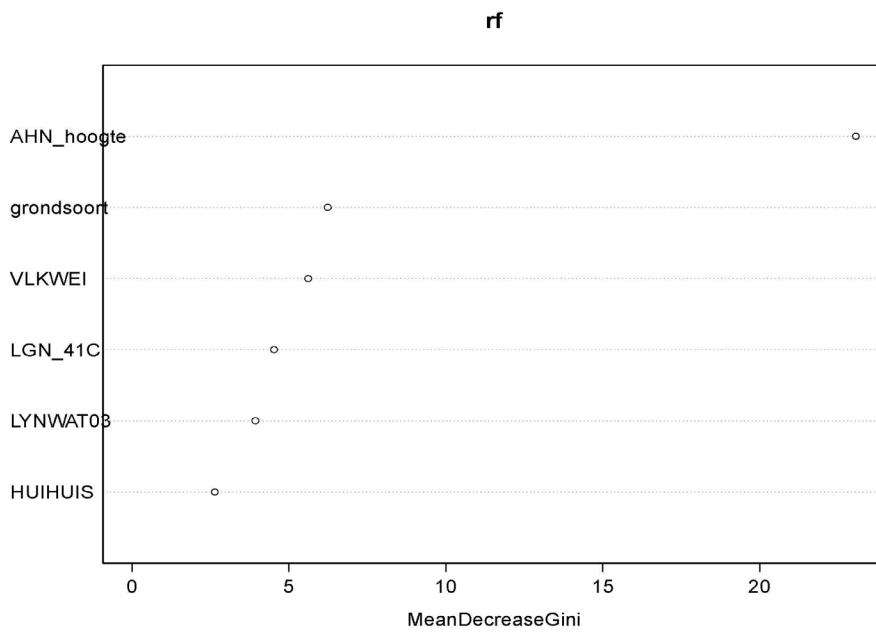
Figuur 15 Variable Importance van het random-forestmodel wel/geen 16BA, voor punten met Gt-code, alle predictorvariabelen.



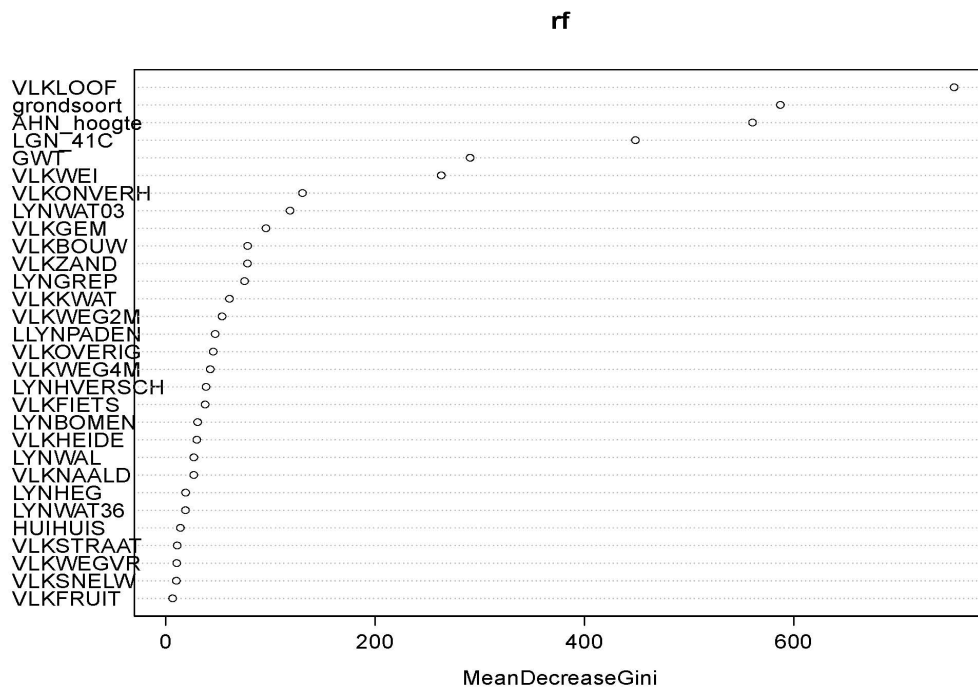
Figuur 16 Variable Importance van het random-forestmodel wel/geen 16BA, voor punten met Gt-code, belangrijkste predictorvariabelen.



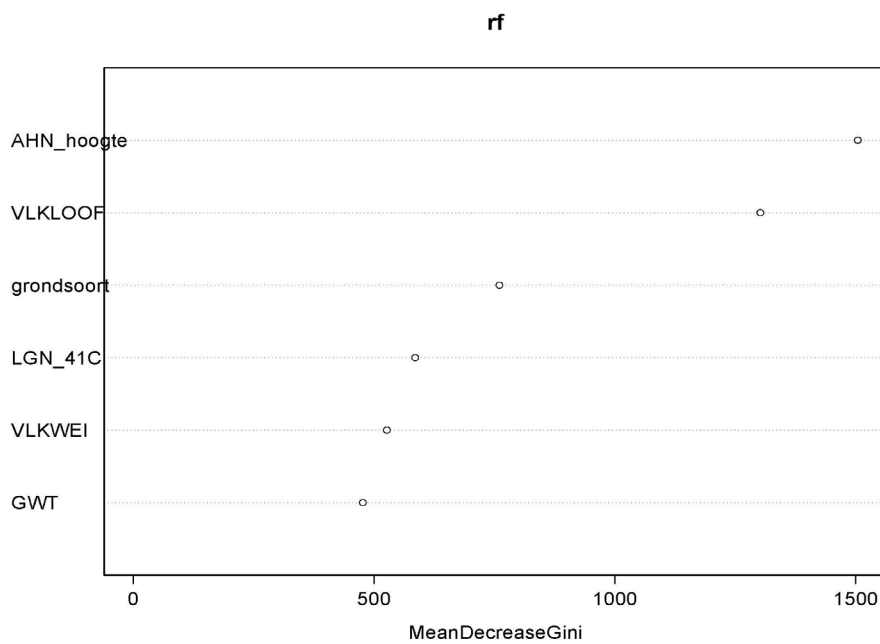
Figuur 17 Variable Importance van het random-forestmodel wel/geen 16BA, voor punten zonder Gt-code, alle predictorvariabelen.



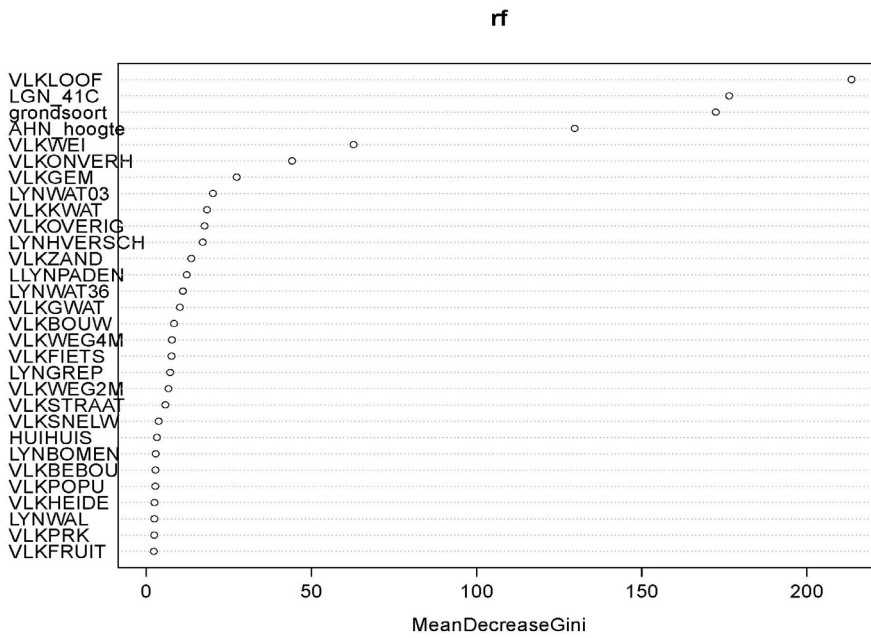
Figuur 18 Variable Importance van het random-forestmodel wel/geen 16BA, voor punten zonder Gt-code, belangrijkste predictorvariabelen.



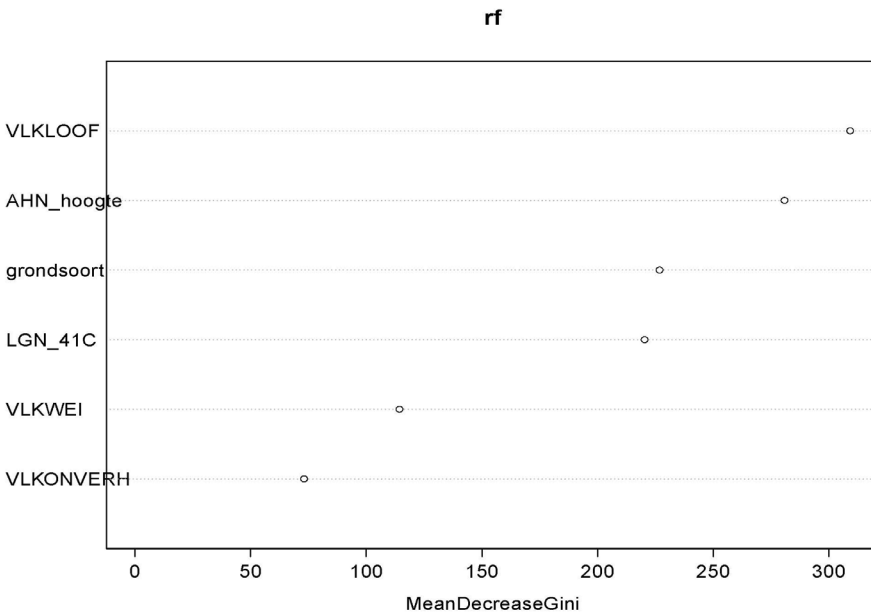
Figuur 19 Variable Importance van het random-forestmodel wel/geen 43AA, voor punten met Gt-code, alle predictorvariabelen.



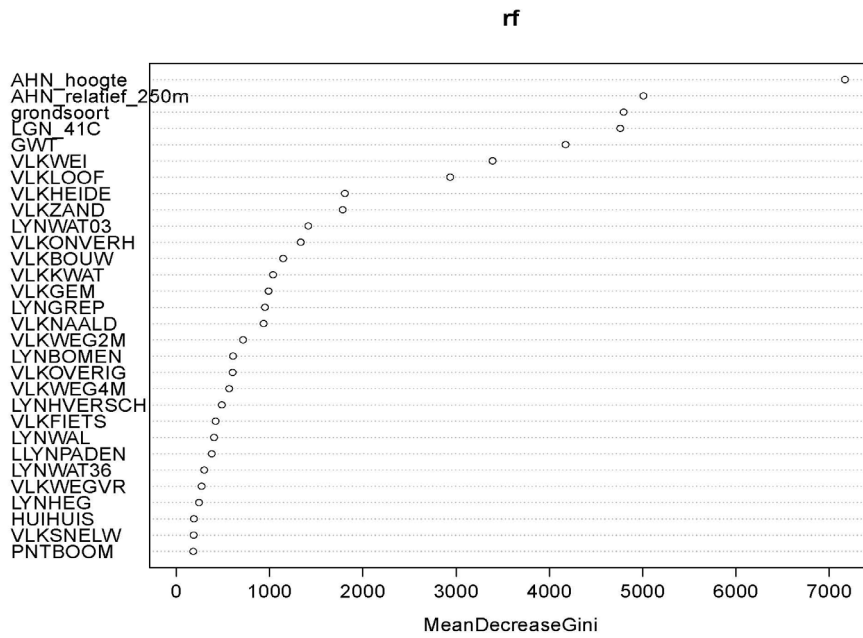
Figuur 20 Variable Importance van het random-forestmodel wel/geen 43AA, voor punten met Gt-code, belangrijkste predictorvariabelen.



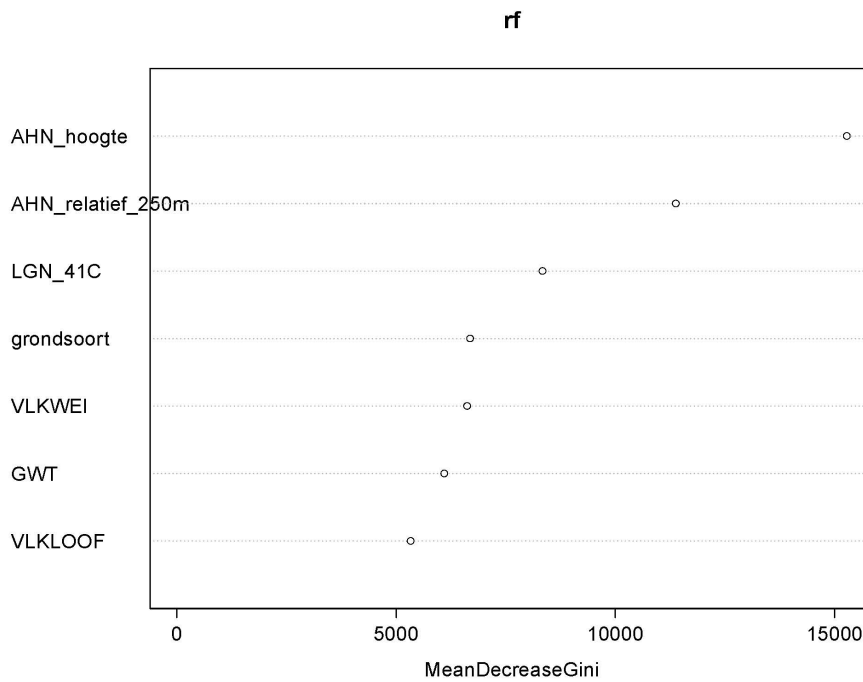
Figuur 21 Variable Importance van het random-forestmodel wel/geen 43AA, voor punten zonder Gt-code, alle predictorvariabelen.



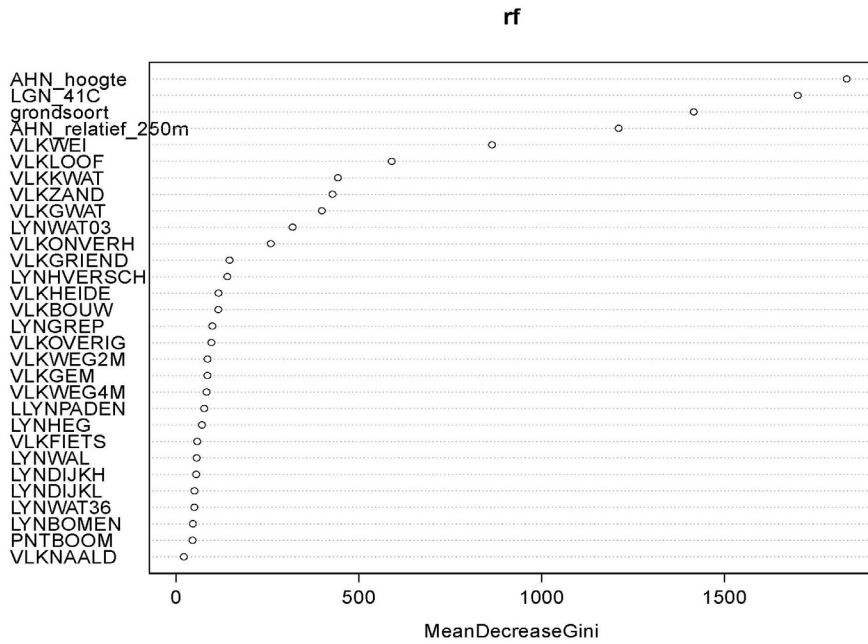
Figuur 22 Variable Importance van het random-forestmodel wel/geen 43AA, voor punten zonder Gt-code, belangrijkste predictorvariabelen.



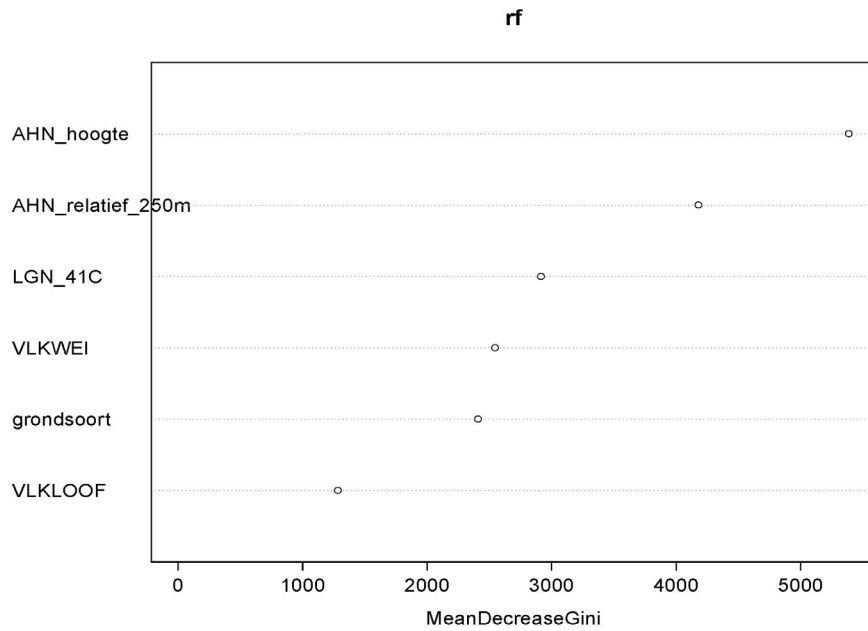
Figuur 23 Variable Importance van het random-forestmodel op verbondsniveau, voor punten met Gt-code, alle predictorvariabelen.



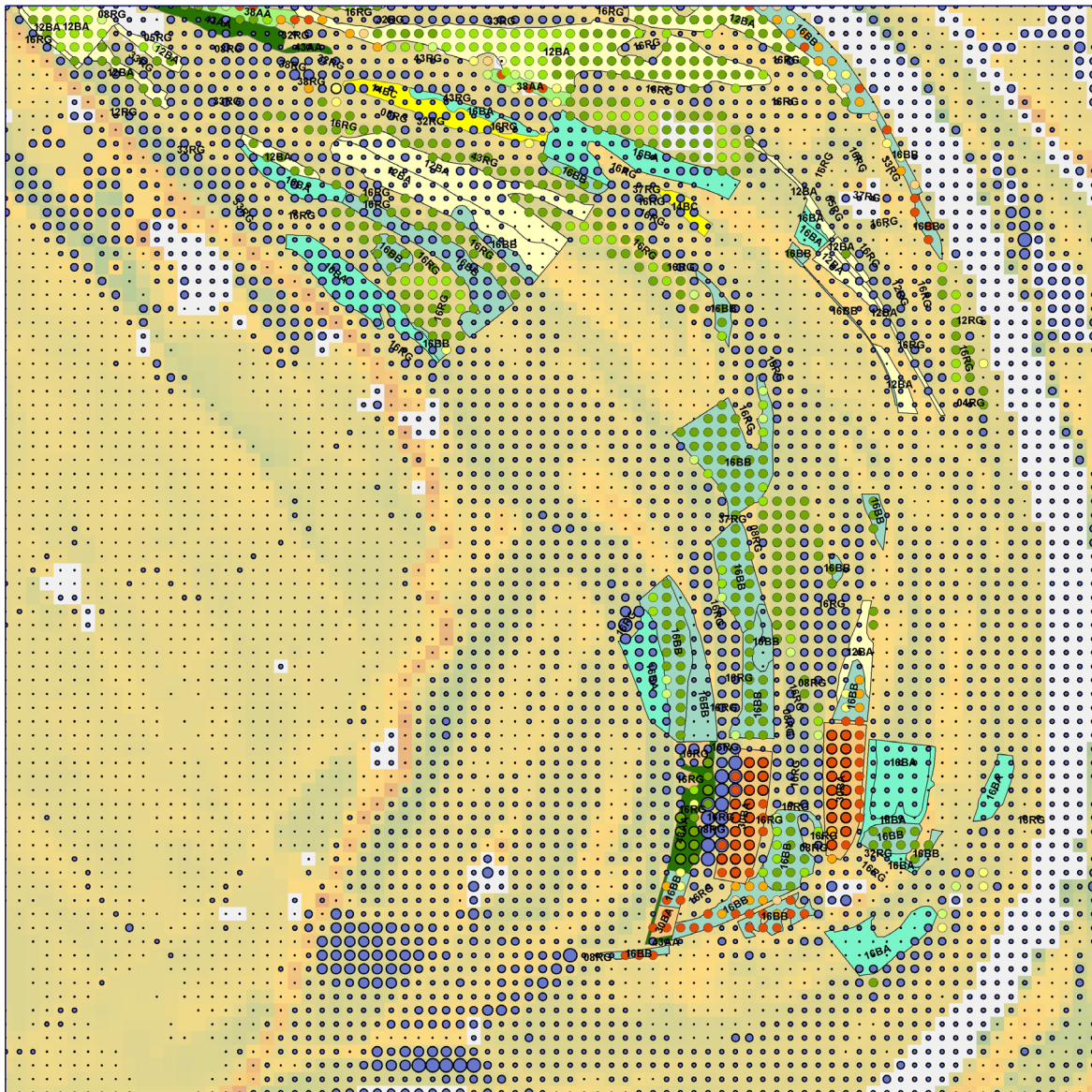
Figuur 24 Variable Importance van het random-forestmodel op verbondsniveau, voor punten met Gt-code, belangrijkste predictorvariabelen.



Figuur 25 Variable Importance van het random-forestmodel op verbondsniveau, voor punten zonder Gt-code, alle predictorvariabelen.

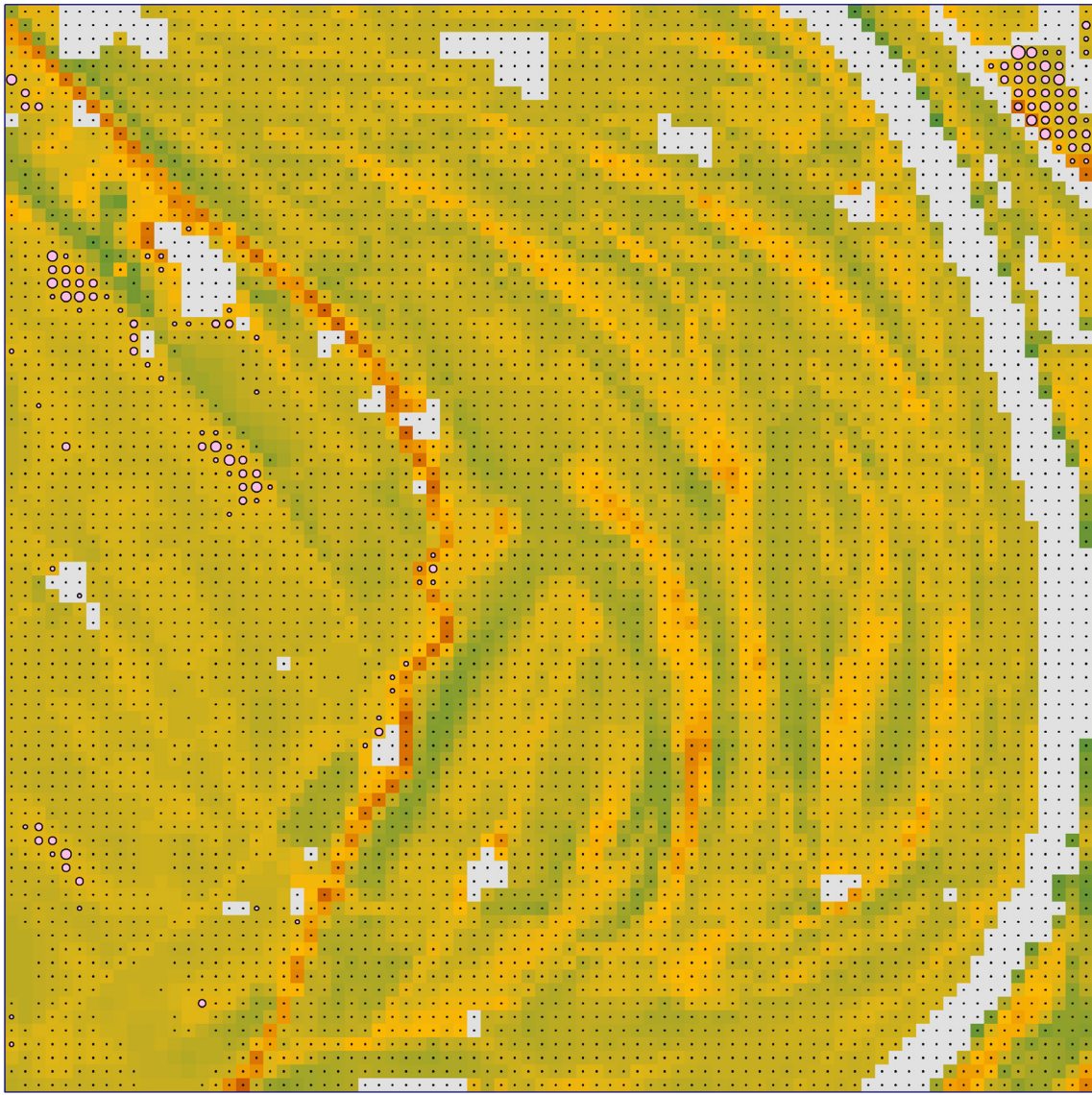


Figuur 26 Variable Importance van het random-forestmodel op verbondsniveau, voor punten zonder Gt-code, belangrijkste predictorvariabelen.



RANK	gecarteerd
● 1	12BA
● 2	14BC
● 3	16BA
● 4-5	16BB
● 6-7	30BA
● 8-10	38AA
● >10	43AA

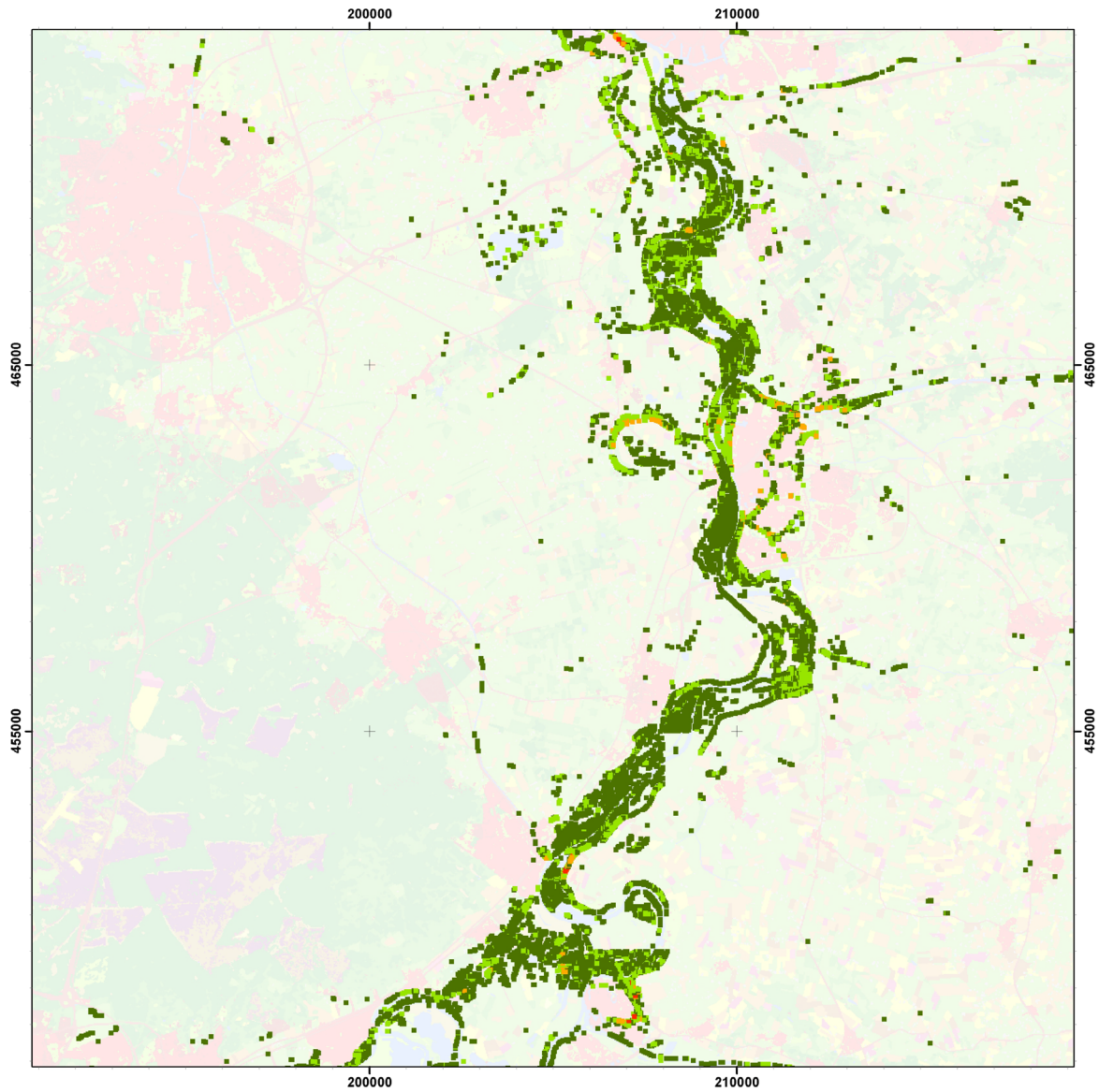
Figuur 28 hoogte van het rangnummer dat correspondeert met de gecarteerde vegetatie-eenheden.



16BA (kans)

- 0.00000
- 0.00001 - 0.00200
- 0.00201 - 0.00400
- ◐ 0.00401 - 0.00800
- ◑ 0.00801 - 0.01200

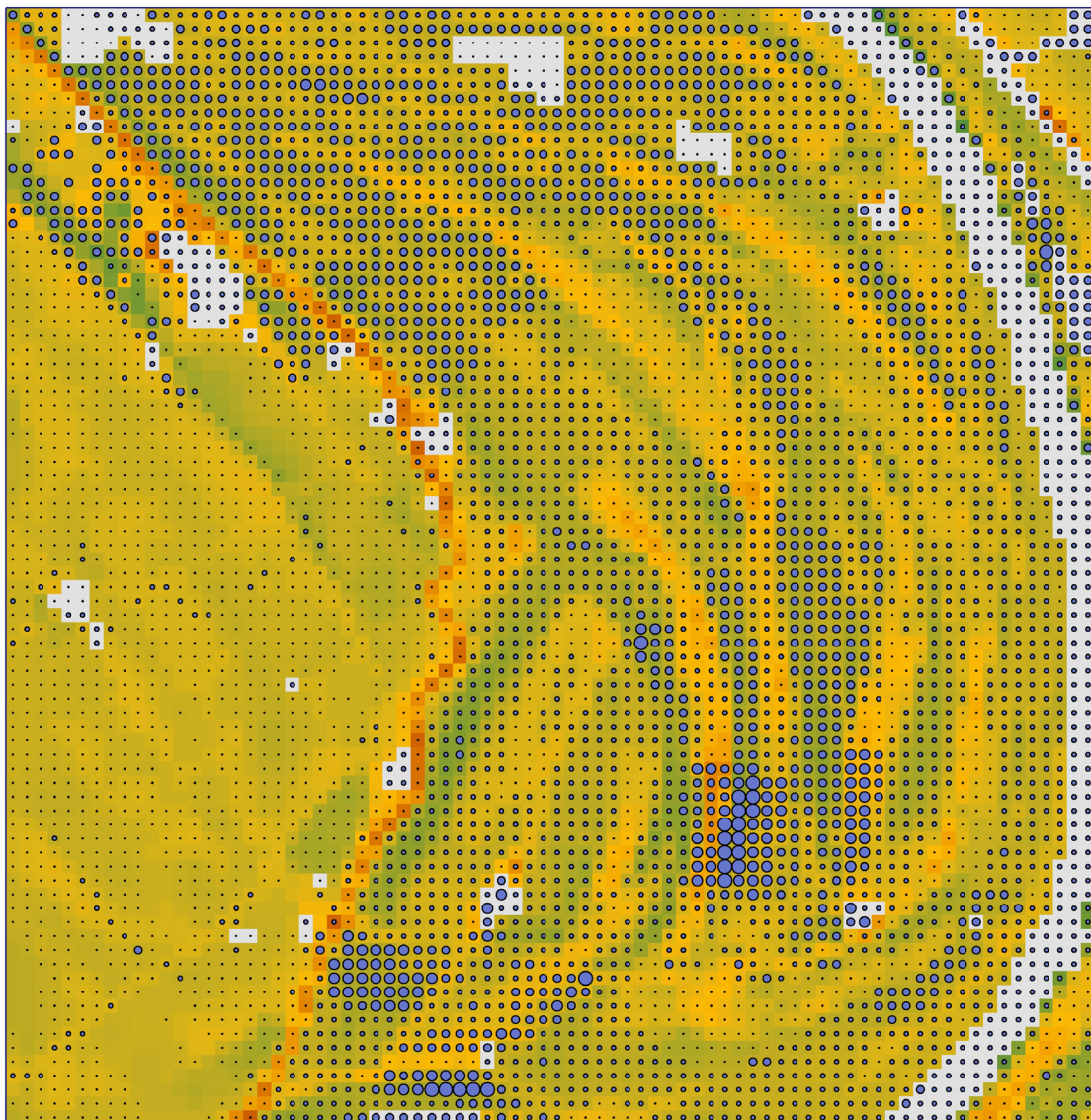
Figuur 29 kanskaart 16BA (*Alopecurion pratensis*).



p16BB



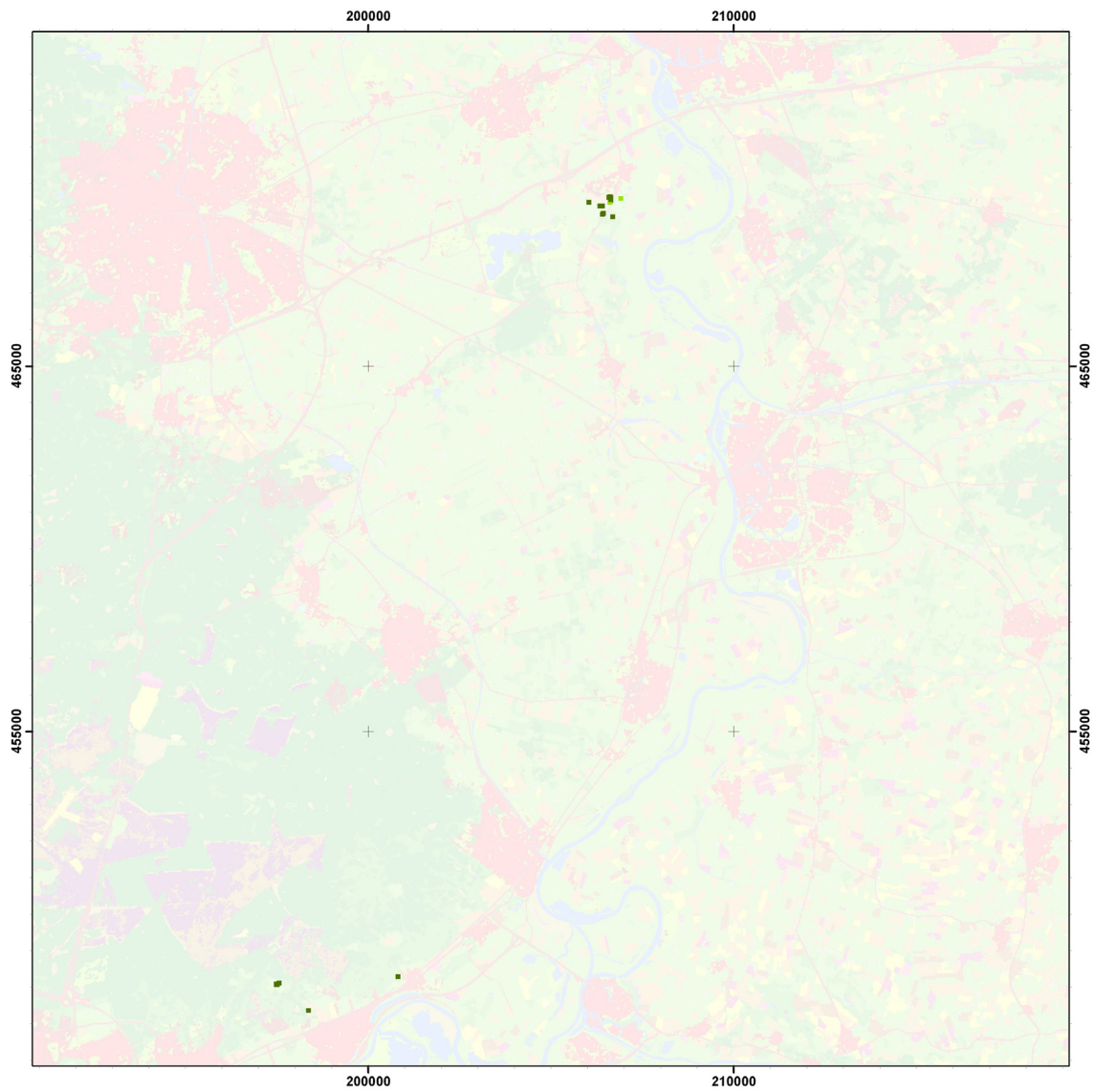
Figuur 30 kanskaart 16BB (*Arrhenatherion elatioris*).



43AA (kans)

- 0.00000 - 0.02200
- 0.02201 - 0.06200
- 0.06201 - 0.14800
- 0.14801 - 0.30800
- 0.30801 - 0.61400

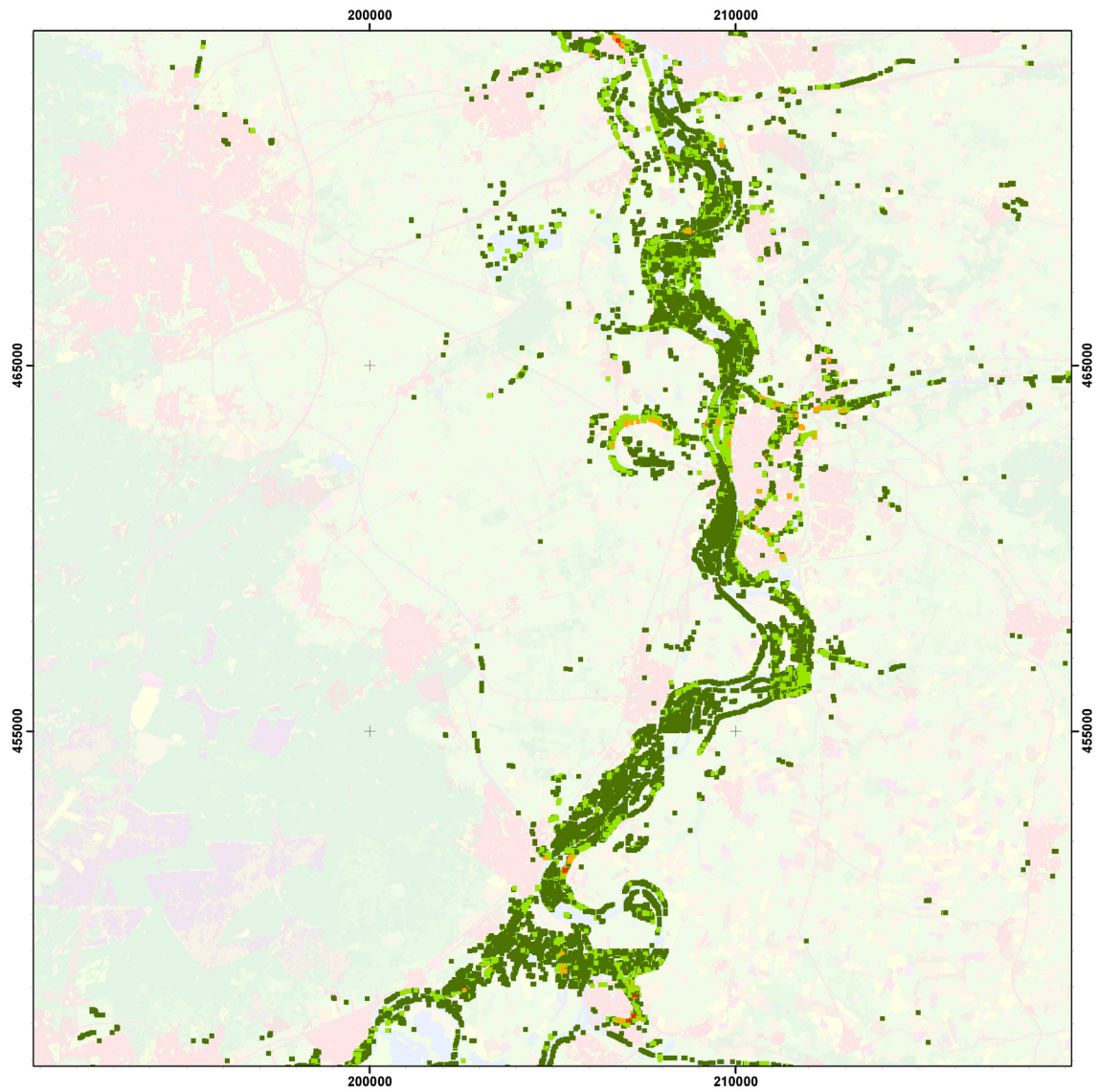
Figuur 31 kanskaart 43AA (Alno-Padion).



p07AA



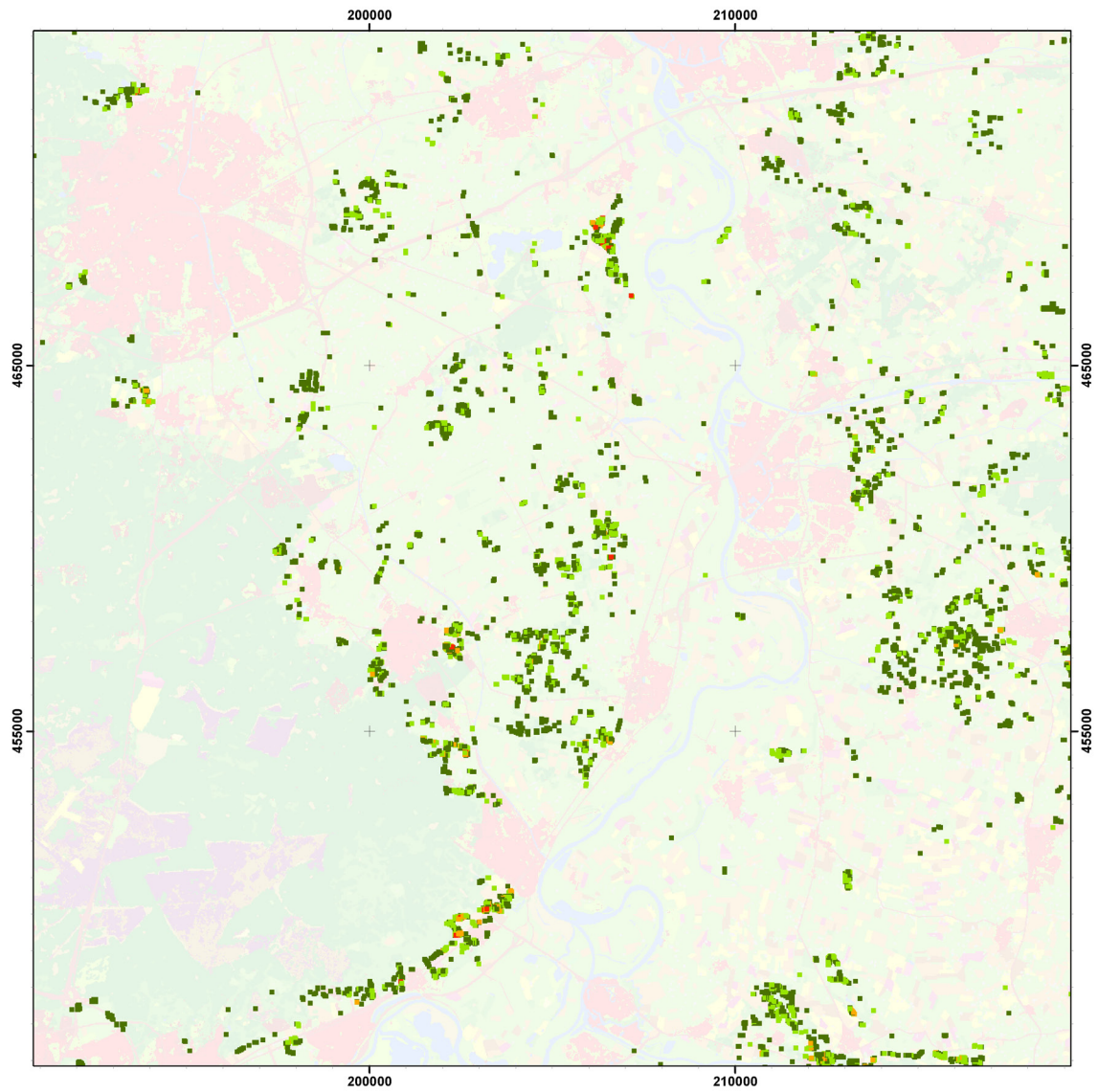
Figuur 32 kanskaart 07AA (Cardamino-Montion).



p16BB

- laag
- gemiddeld
- hoog
- zeer hoog

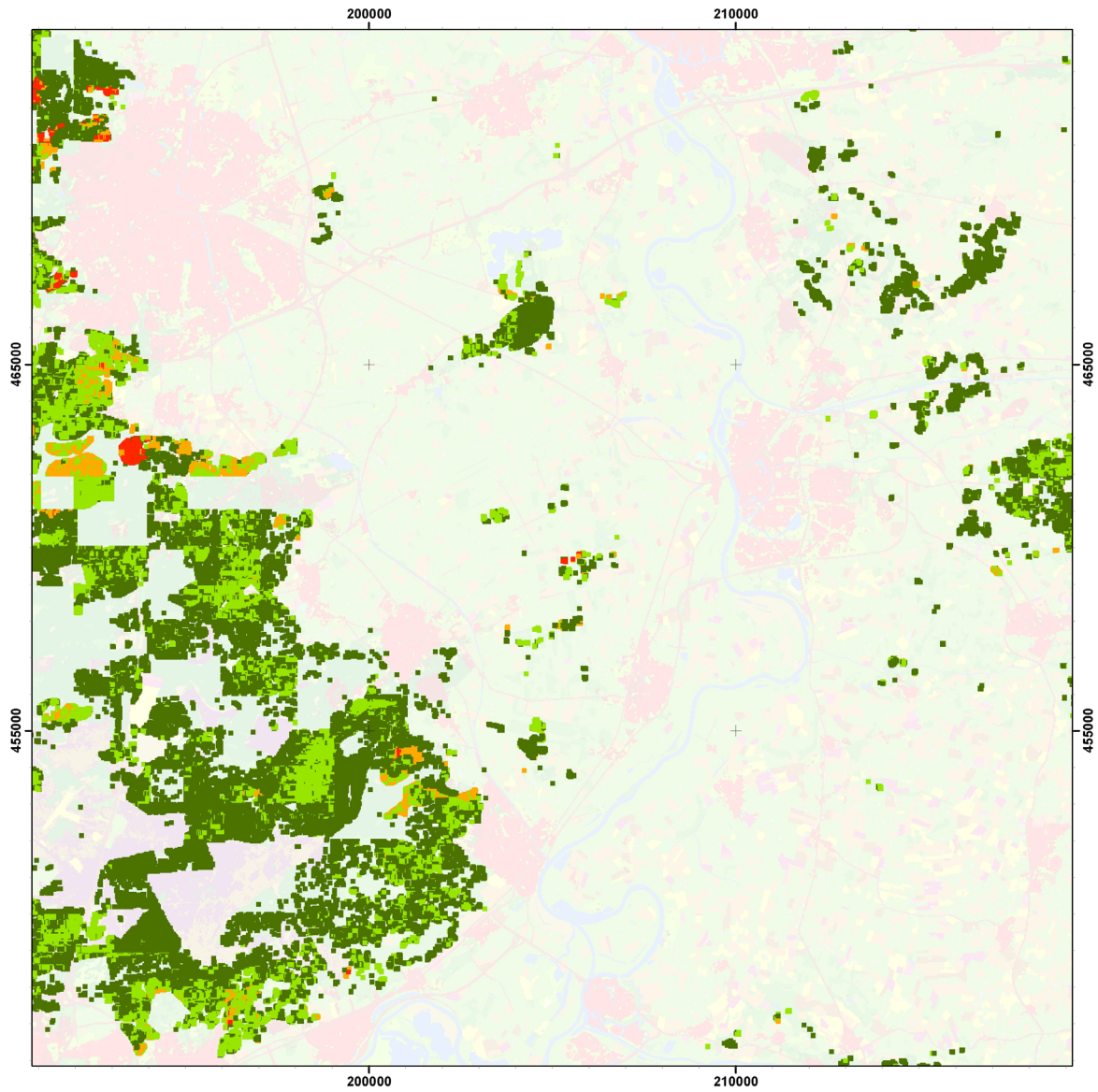
Figuur 33 kanskaart 16BB (*Arrhenatherion elatioris*).



p39AA



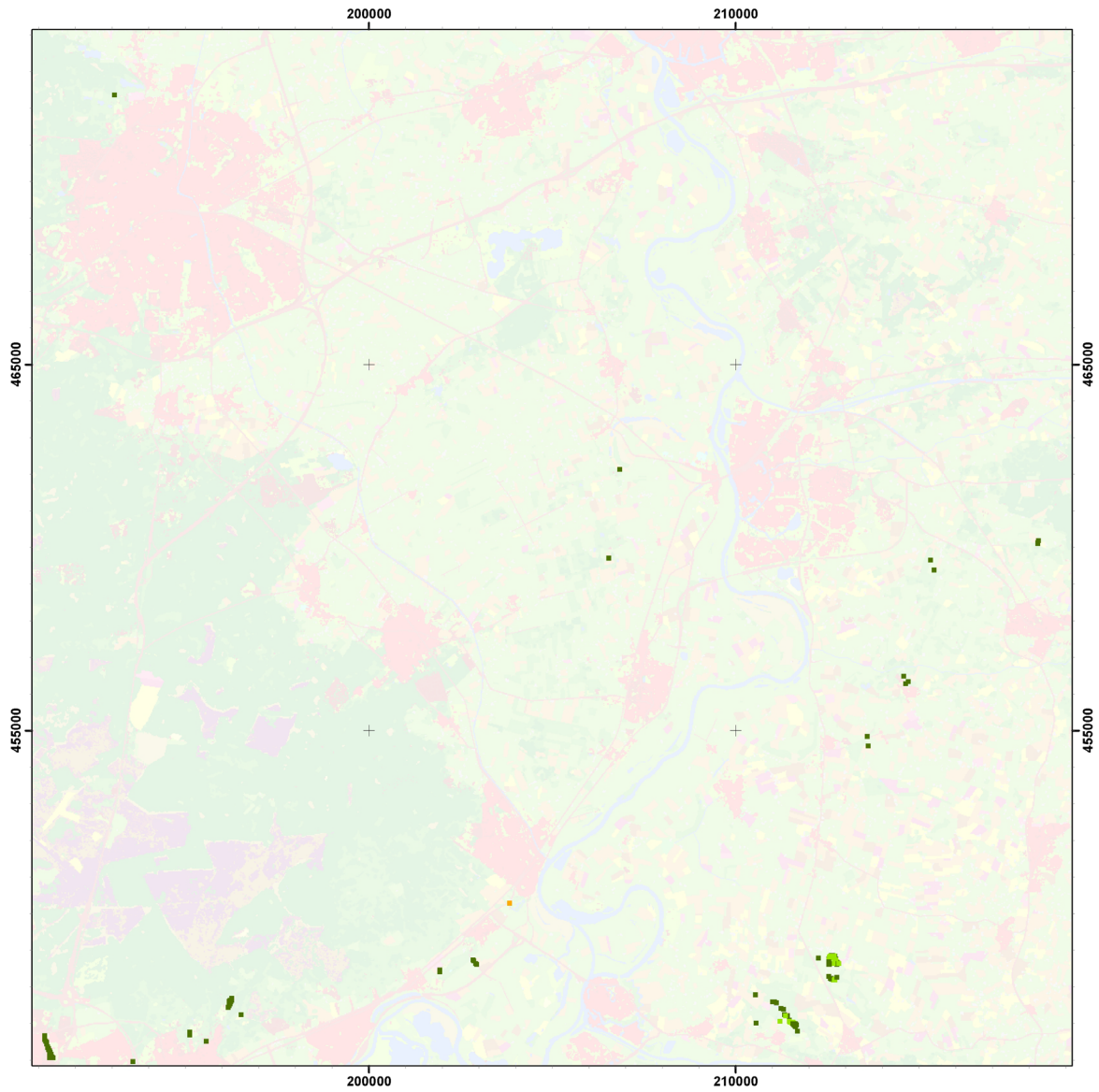
Figuur 34 kanskaart 39AA (*Alnion glutinosae*).



41AA



Figuur 35 kanskaart 41AA (Dicrano-Pinion).

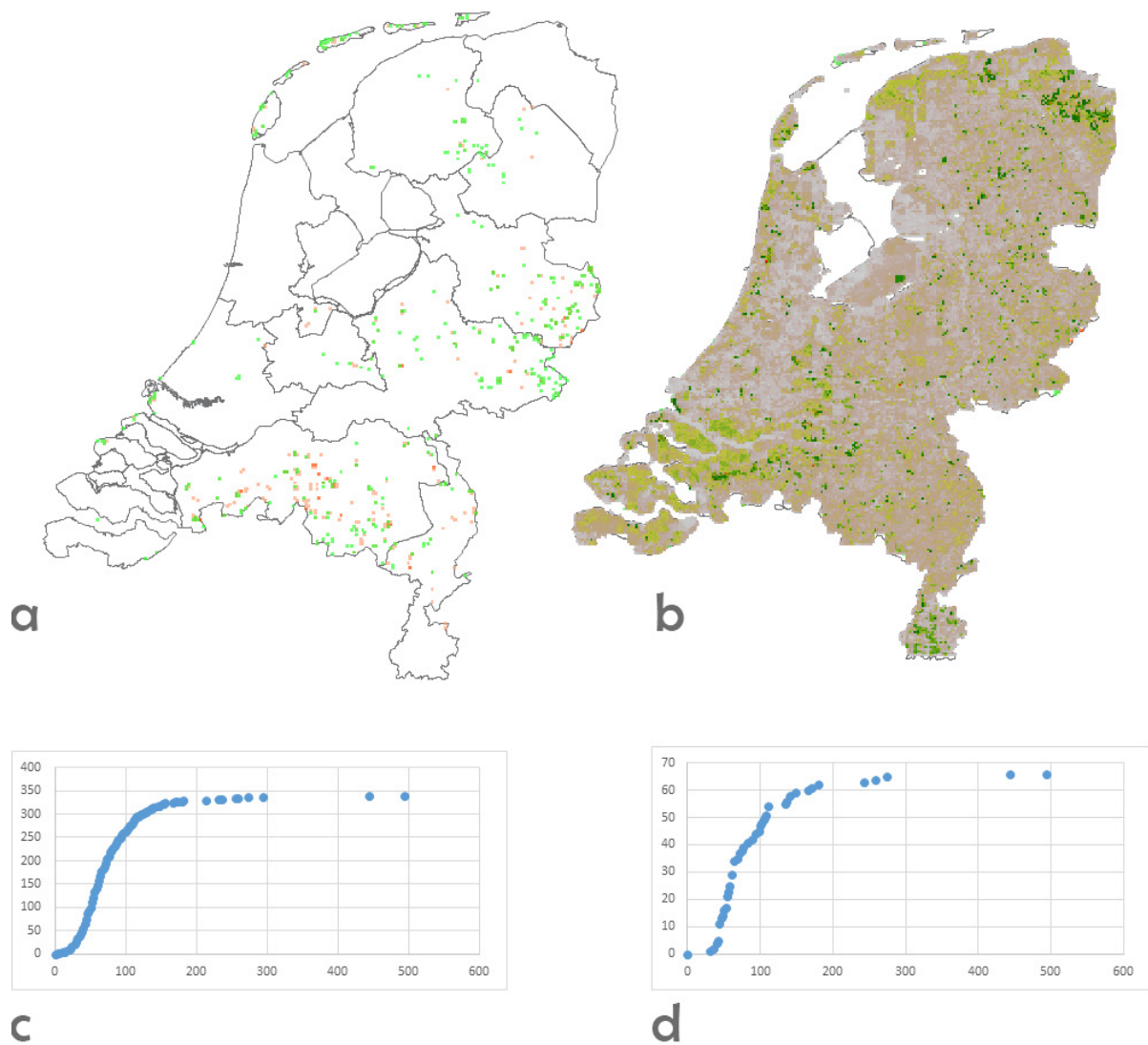


43AB

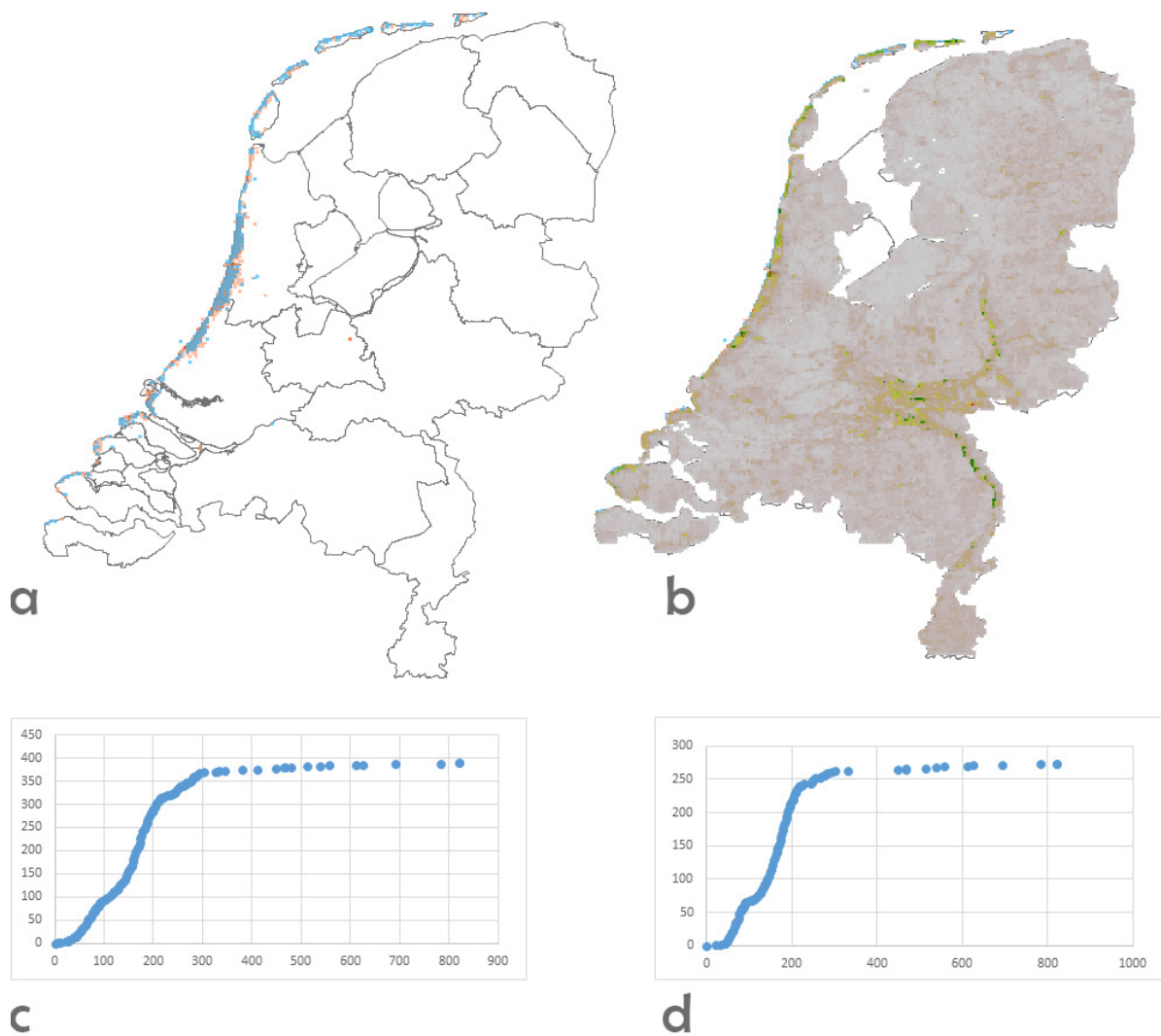


Figuur 36 kanskaart 43AB (Alno-Padion).

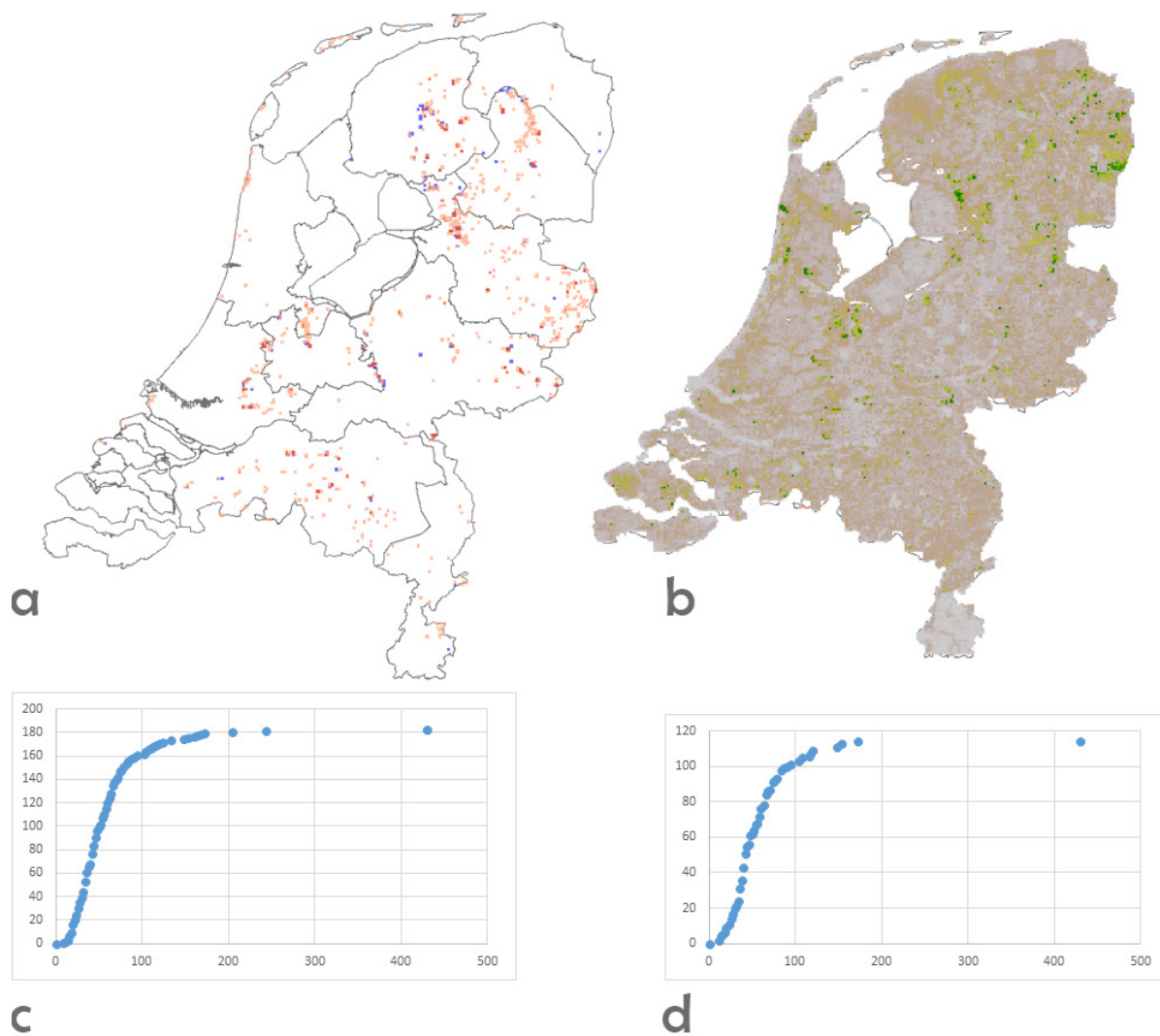
Bijlage 4 Validatie heel Nederland



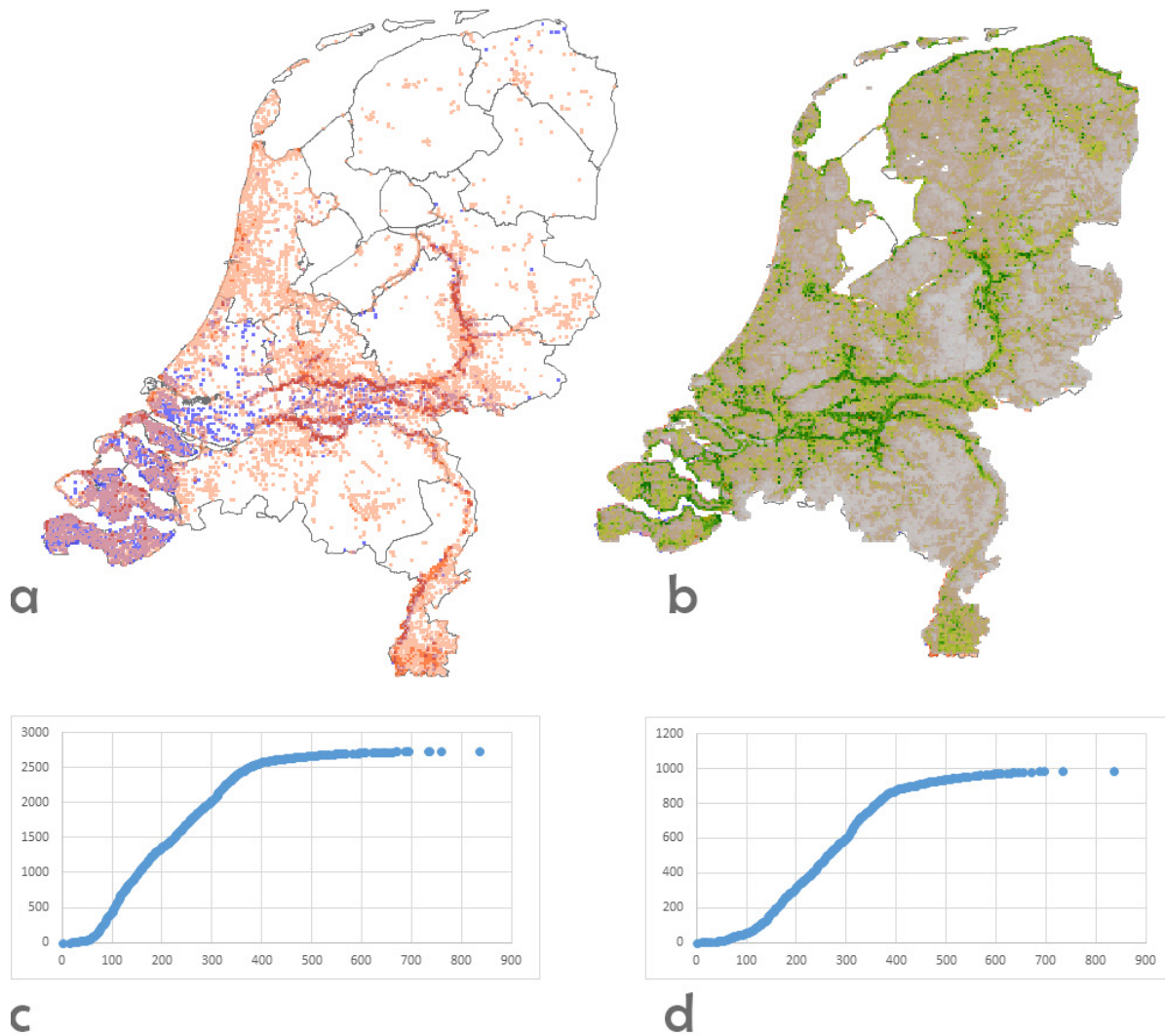
Figuur 37 06Ac Hydrocotylo-Baldellion: Opnames en voorspelling: a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (Dynterra). Cumulatieve frequenties versus de voorspelde score van het model: c Cumulatieve frequentie op basis van opnames; d Cumulatieve frequentie op basis van FLORBASE schatting.



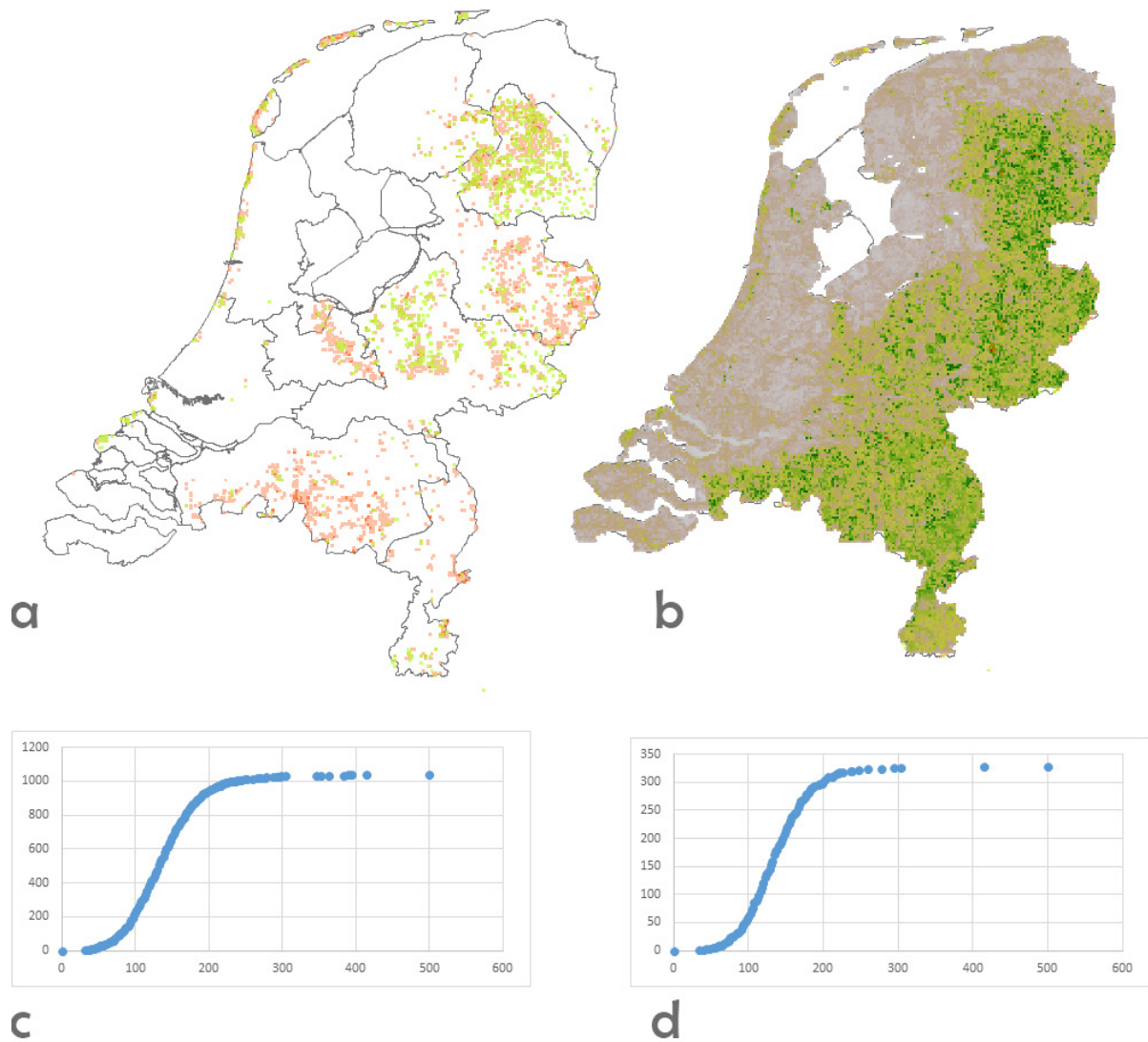
Figuur 38 14Ca Tortulo-Koelerion: Opnames en voorspelling: a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (Dynterra). Cumulatieve frequenties versus de voorspelde score van het model: c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.



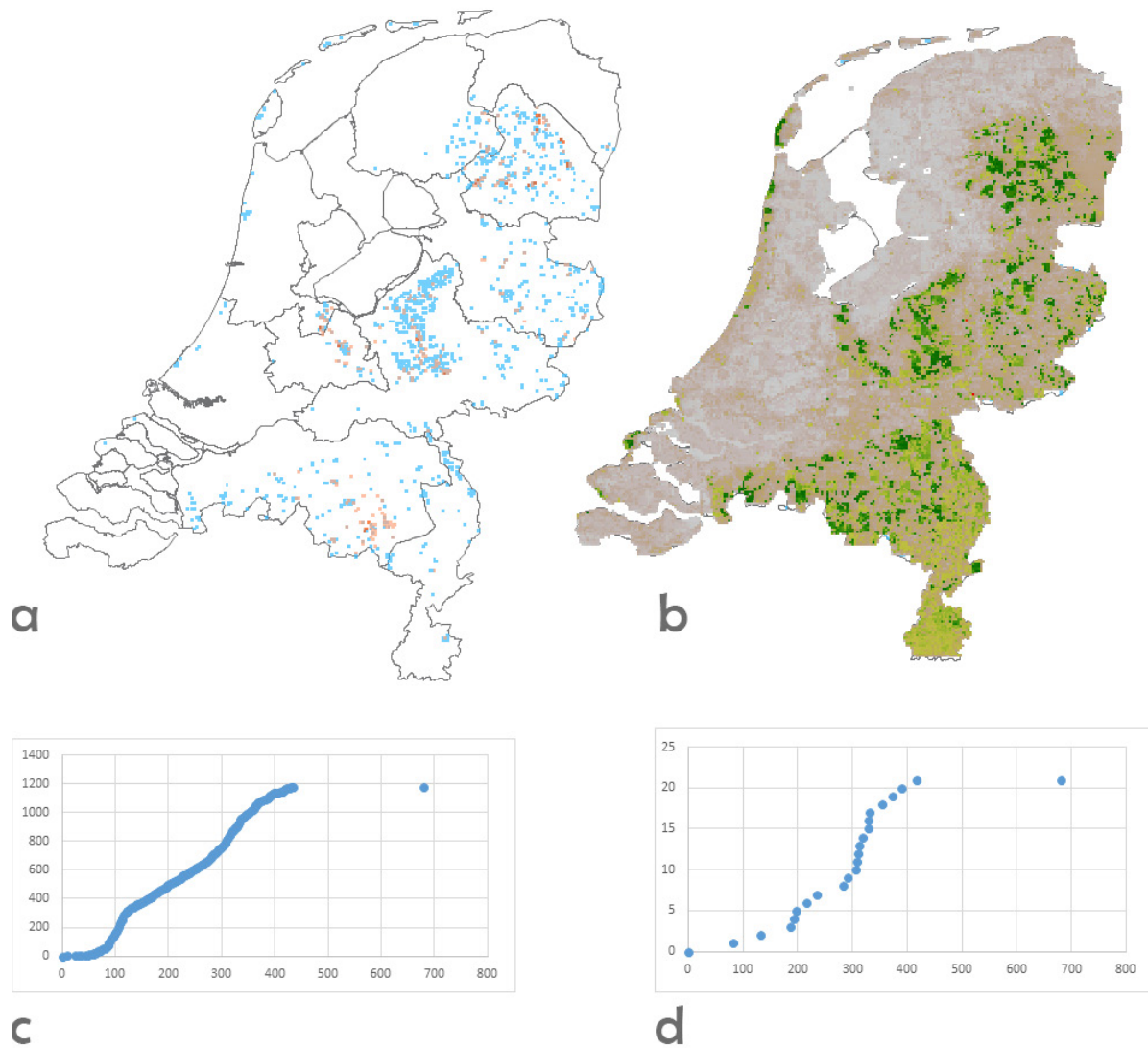
Figuur 39 16Aa *Junco-Molinion*: Opnames en voorspelling: a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (Dynterra). Cumulatieve frequenties versus de voorspelde score van het model: c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.



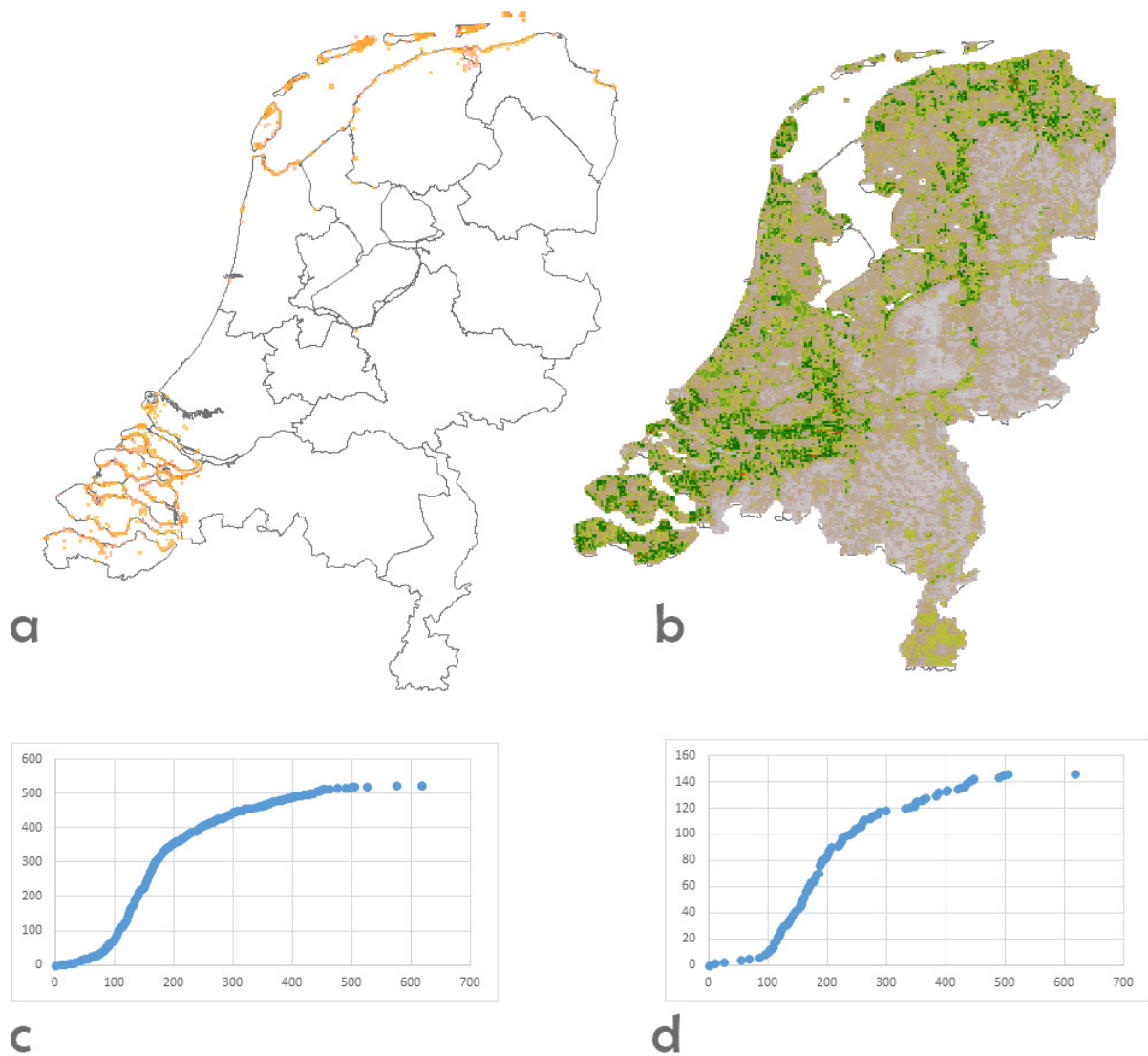
Figuur 40 16Bb *Arrhenatherion elatioris*: Opnames en voorspelling: a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (Dynterra). Cumulatieve frequenties versus de voorspelde score van het model: c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.



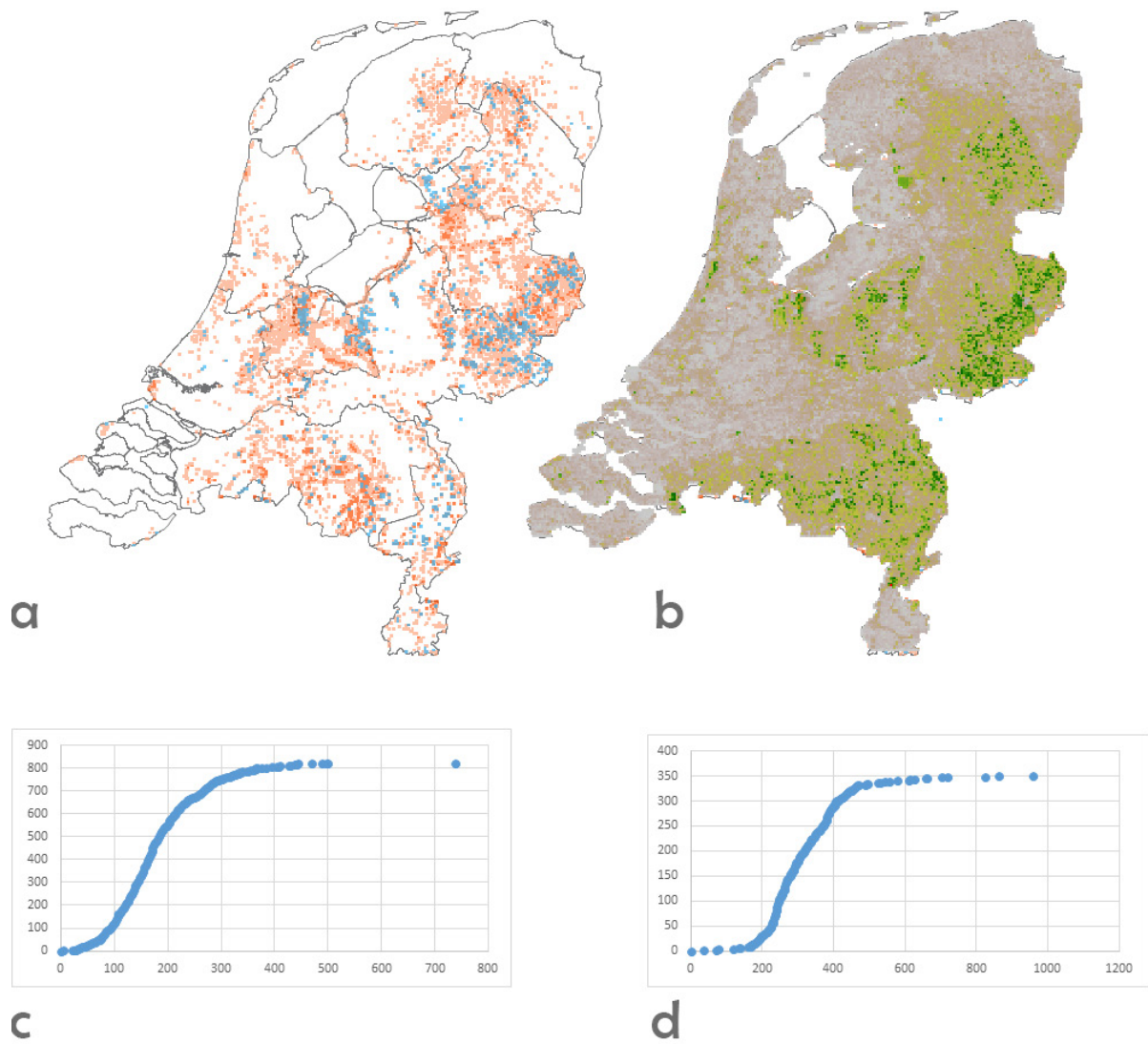
Figuur 41 19Aa *Nardo-Galium saxatilis*: Opnames en voorspelling: a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (Dynterra). Cumulatieve frequenties versus de voorspelde score van het model: c Cumulatieve frequentie op basis van opnames; d Cumulatieve frequentie op basis van FLORBASE schatting.



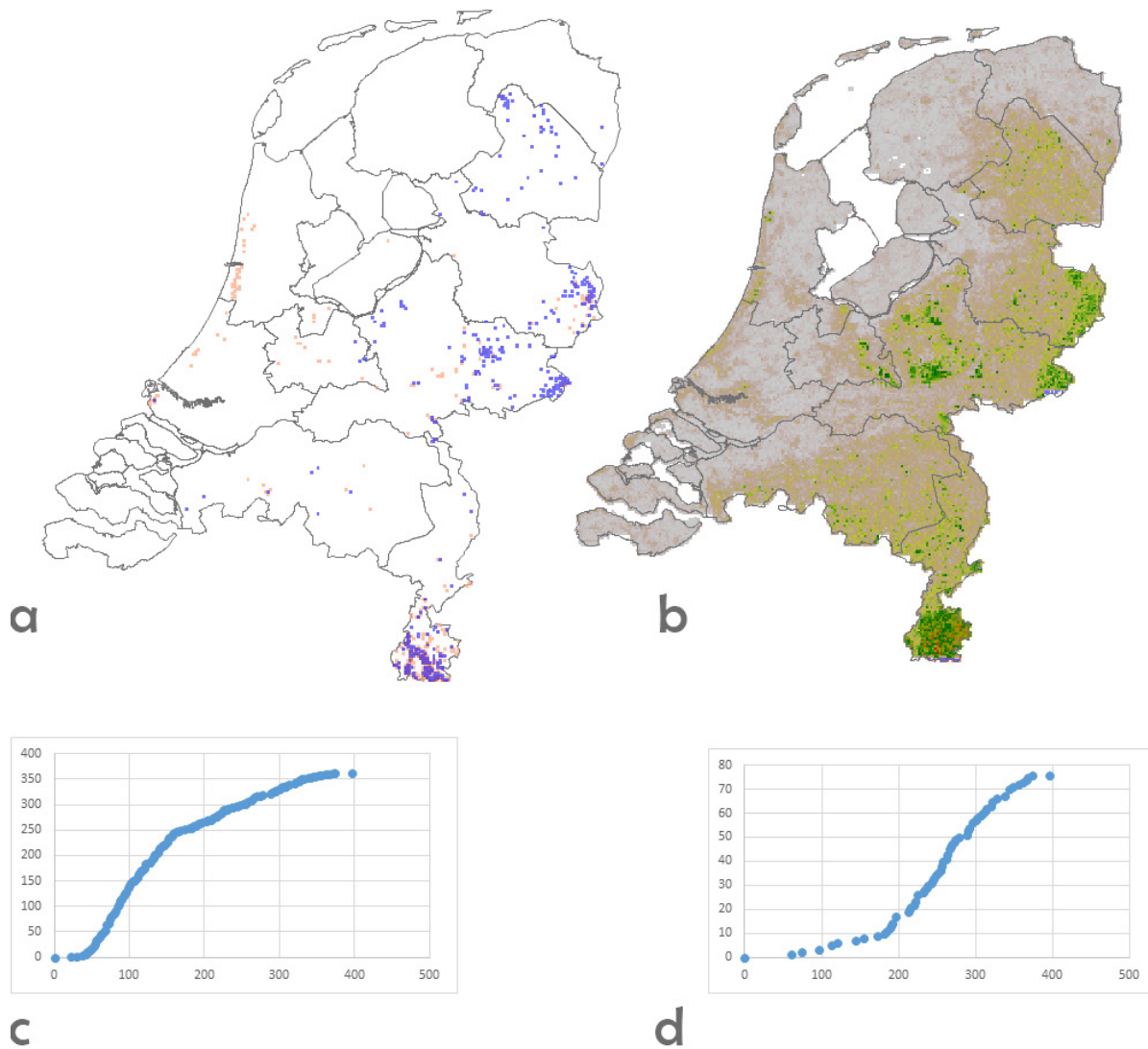
Figuur 42 20Aa *Calluno-Genistion pilosae*: Opnames en voorspelling: a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (Dynterra). Cumulatieve frequenties versus de voorspelde score van het model: c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.



Figuur 43 26Ac *Armerion maritima*: Opnames en voorspelling: a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (Dynterra). Cumulatieve frequenties versus de voorspelde score van het model: c Cumulatieve frequentie op basis van opnames; d Cumulatieve frequentie op basis van FLORBASE schatting.

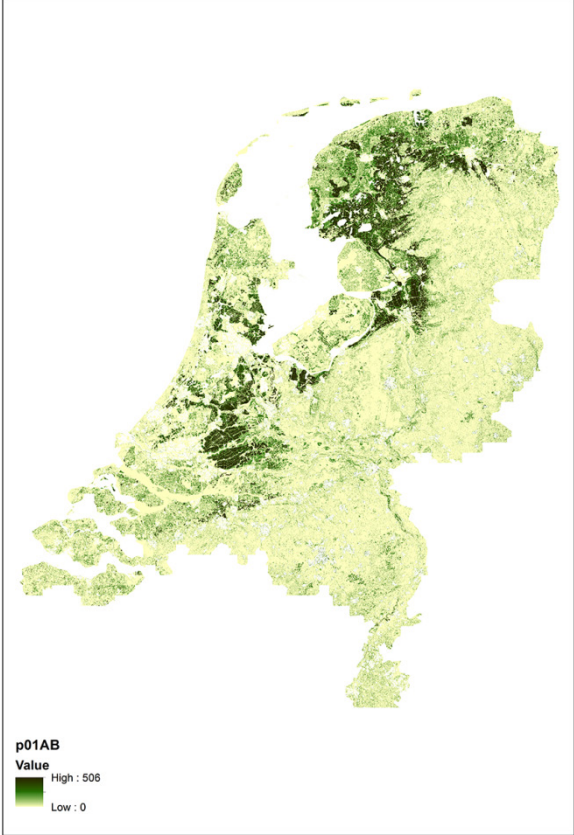
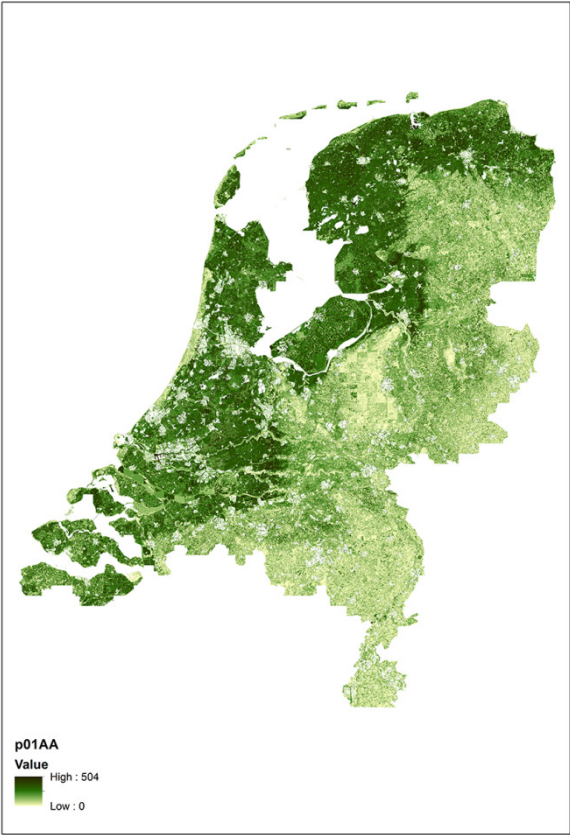


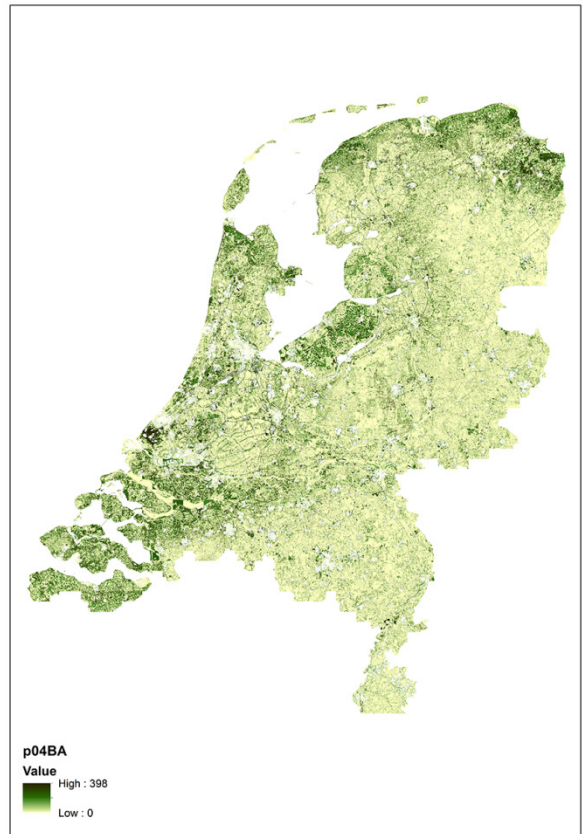
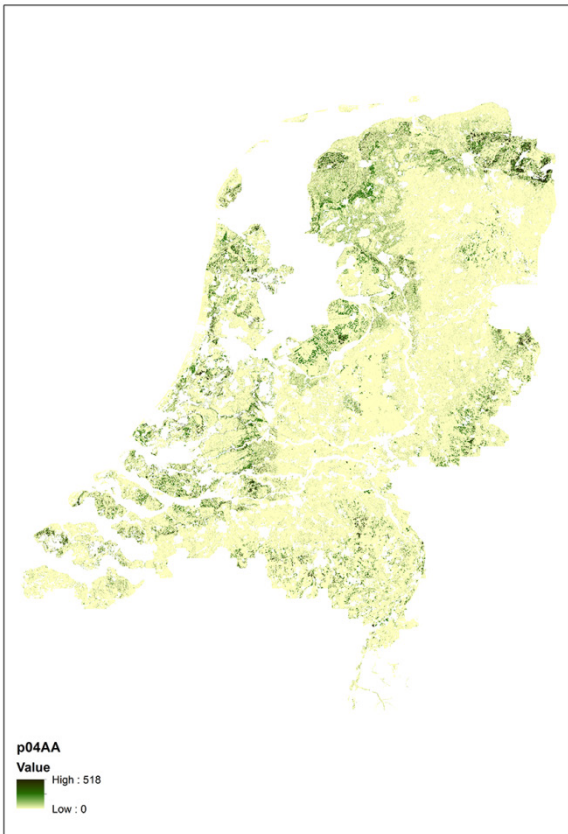
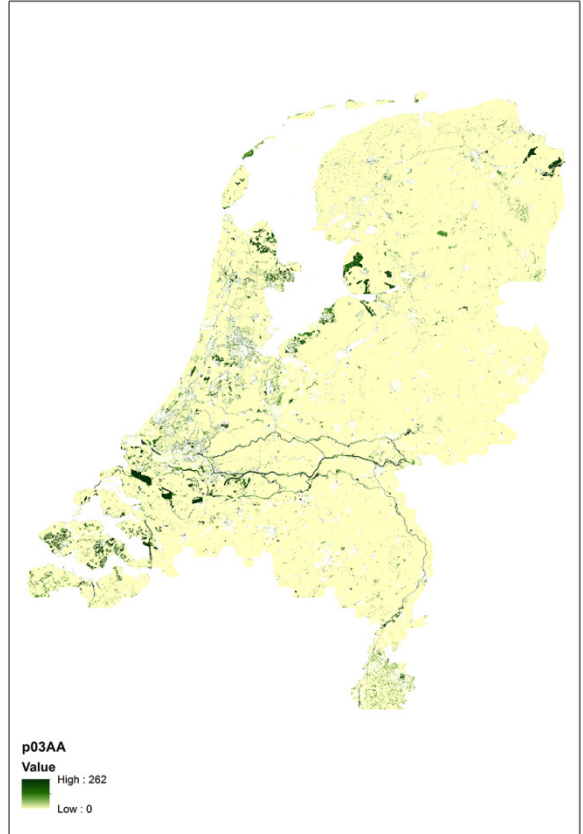
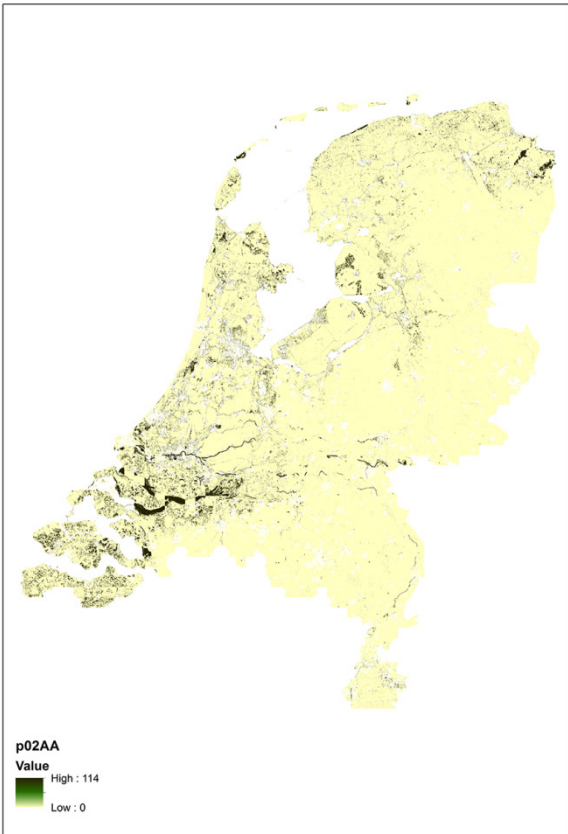
Figuur 44 39Aa *Alnion glutinosae*: Opnames en voorspelling: a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (Dynterra). Cumulatieve frequenties versus de voorspelde score van het model: c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.

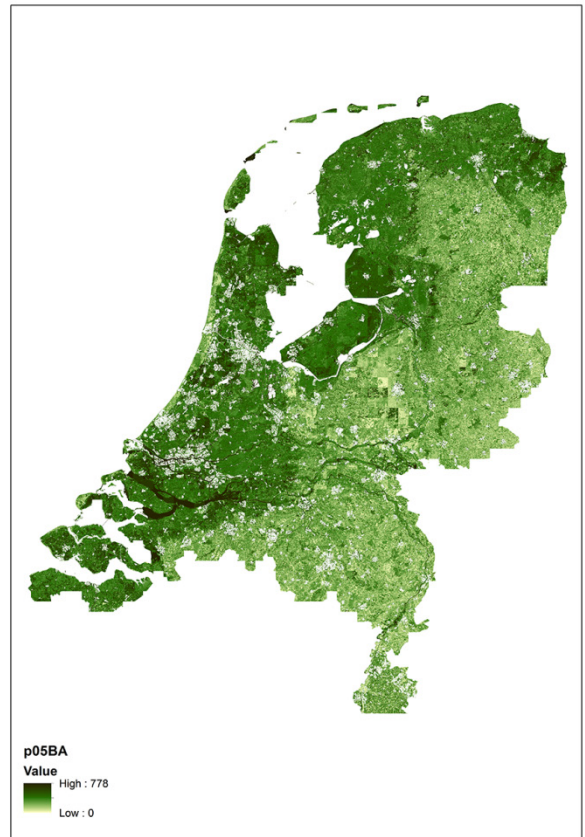
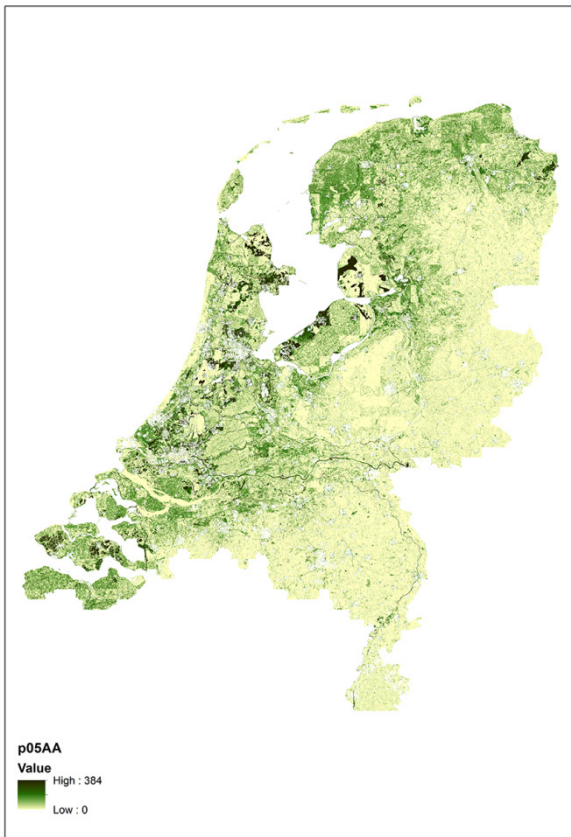
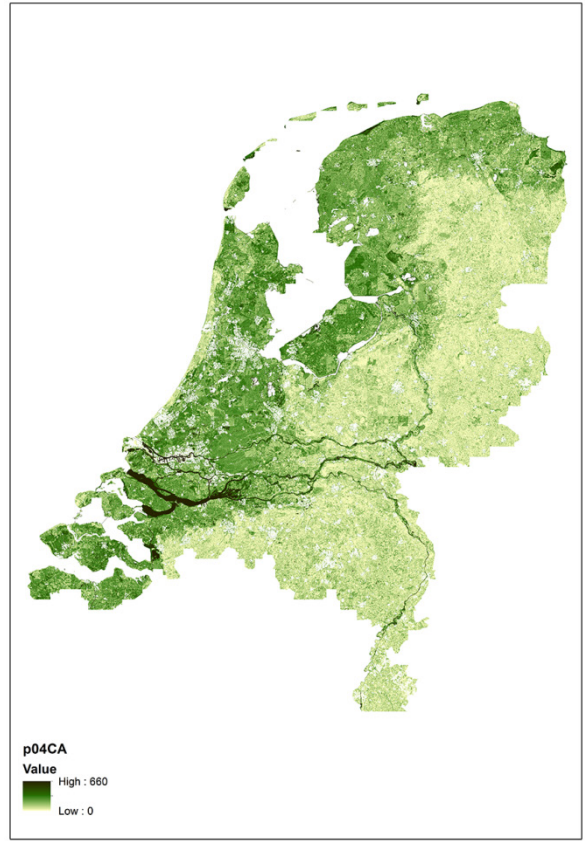
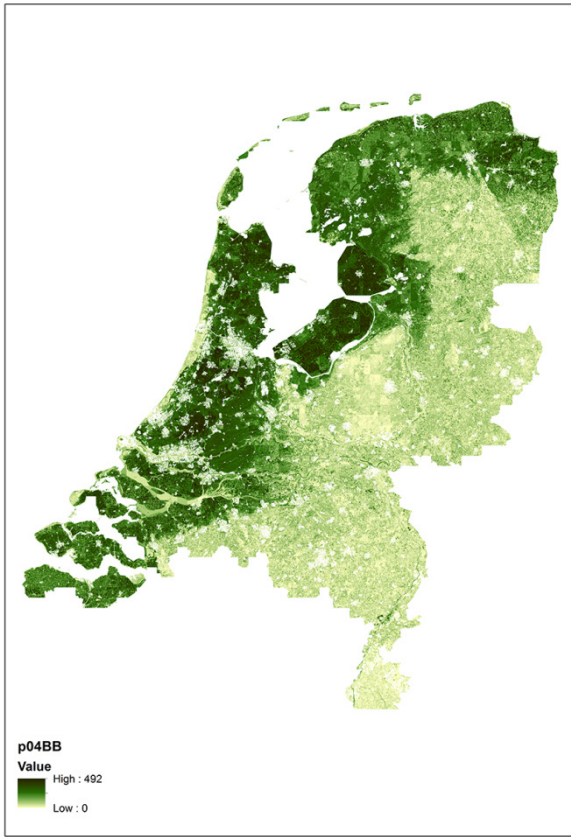


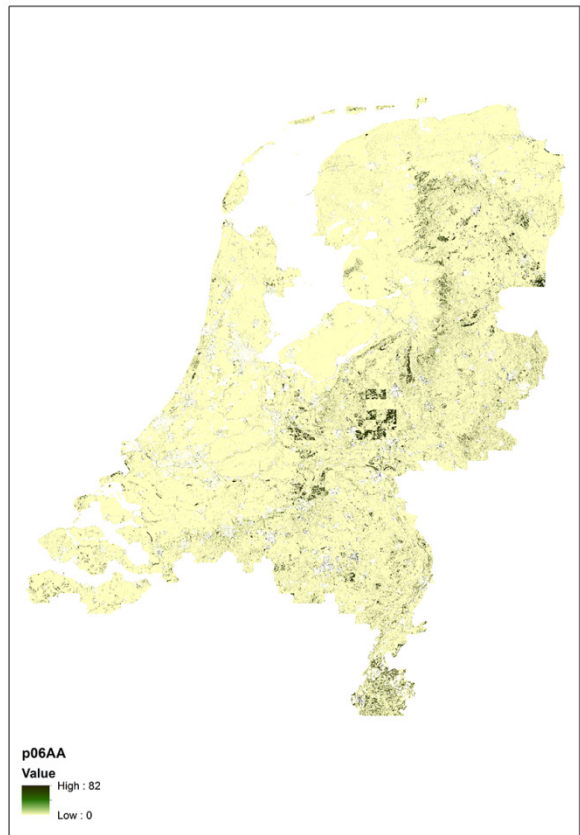
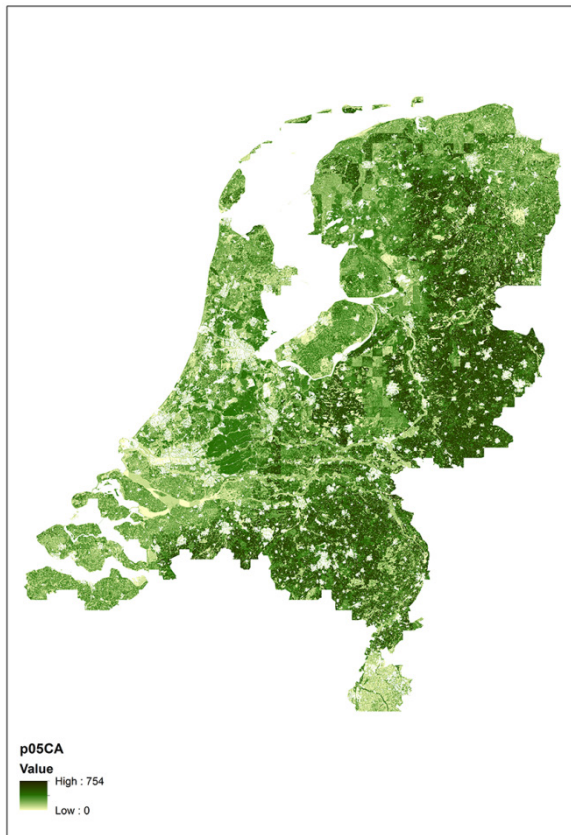
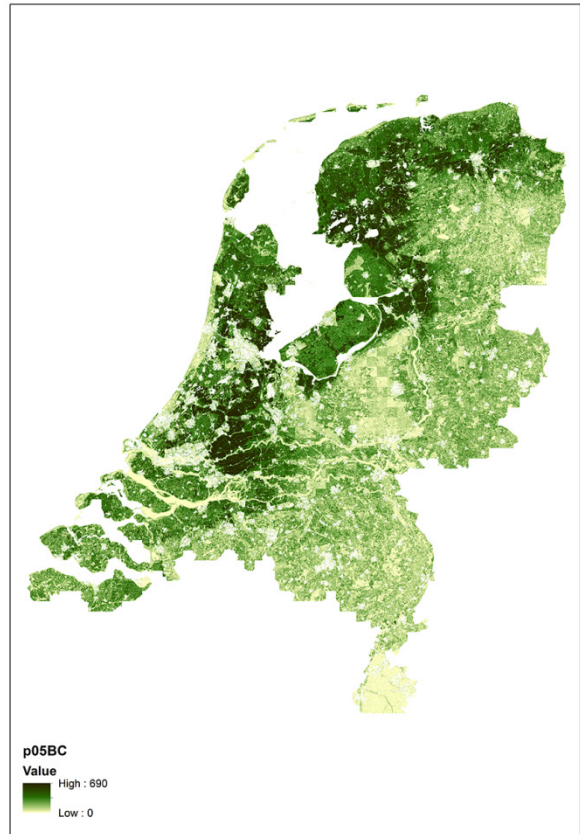
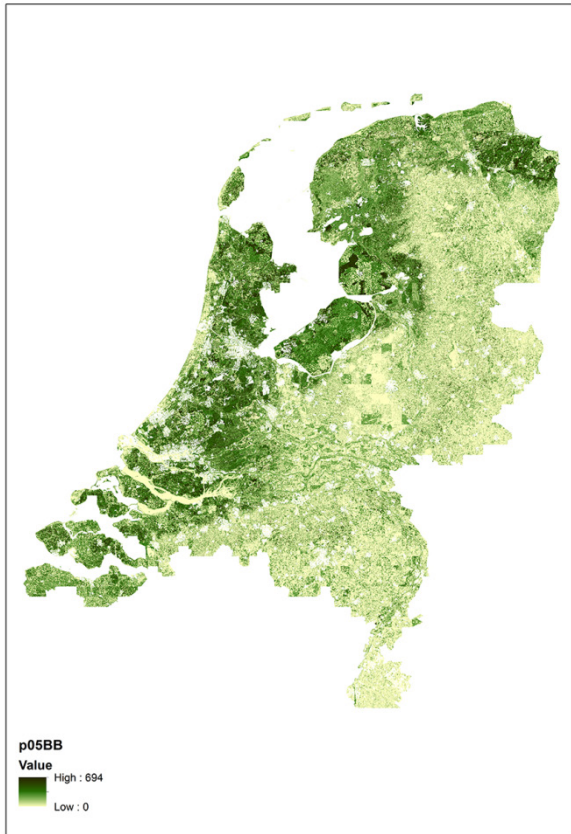
Figuur 45 43Ab *Carpinion betuli*: Opnames en voorspelling: a Verspreiding opnamepunten (groen) en geschat voorkomen (rood); b Kans op voorkomen op basis van abiotiek (Dynterra). Cumulatieve frequenties versus de voorspelde score van het model: c Cumulatieve frequentie op basis van opnamen; d Cumulatieve frequentie op basis van FLORBASE schatting.

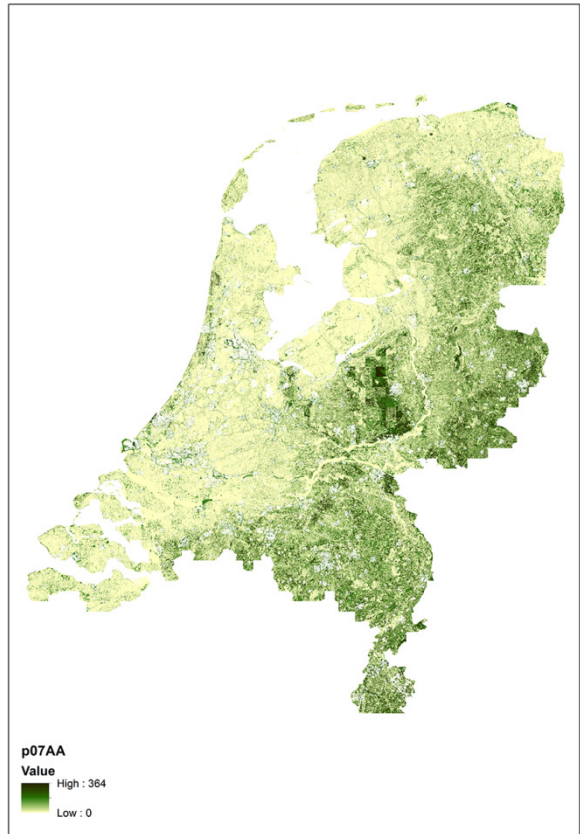
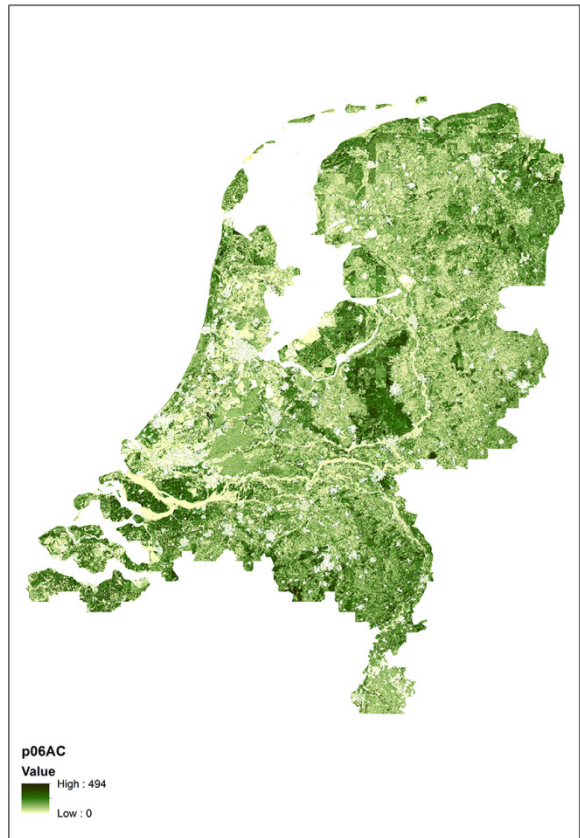
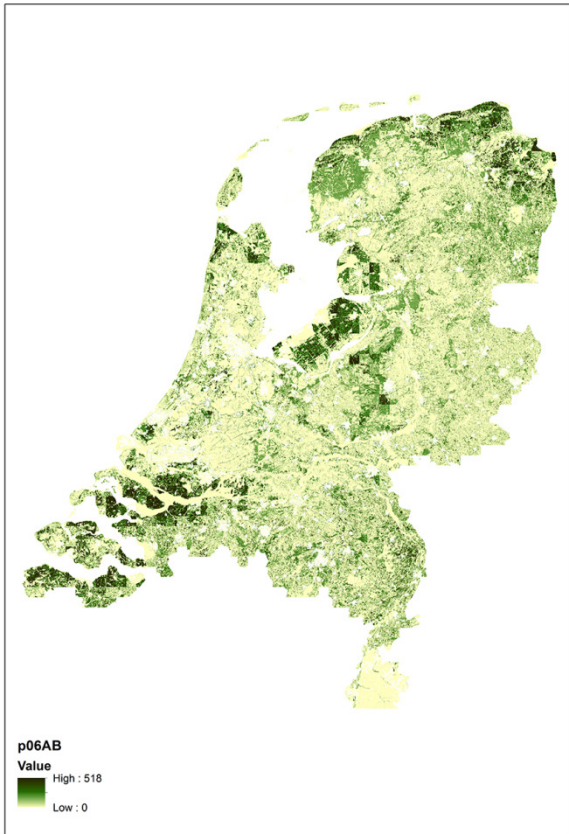
Bijlage 5 Voorspellingen heel Nederland

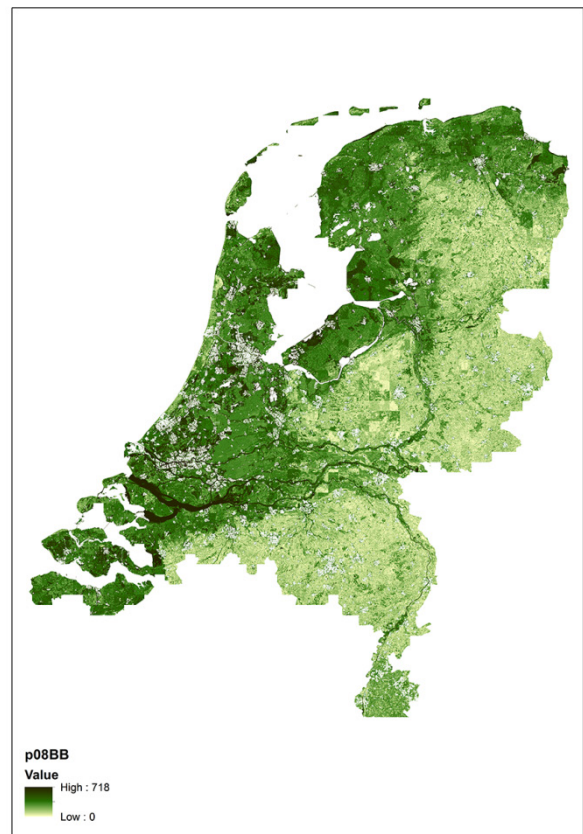
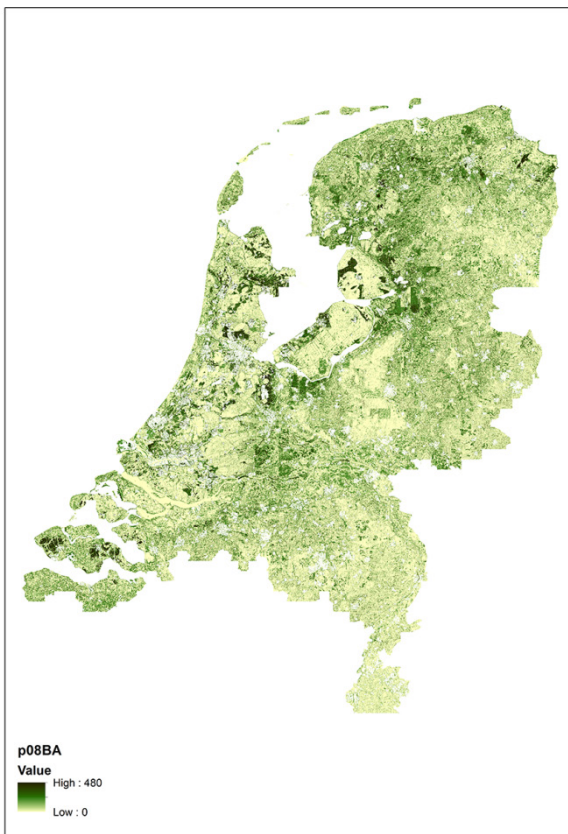
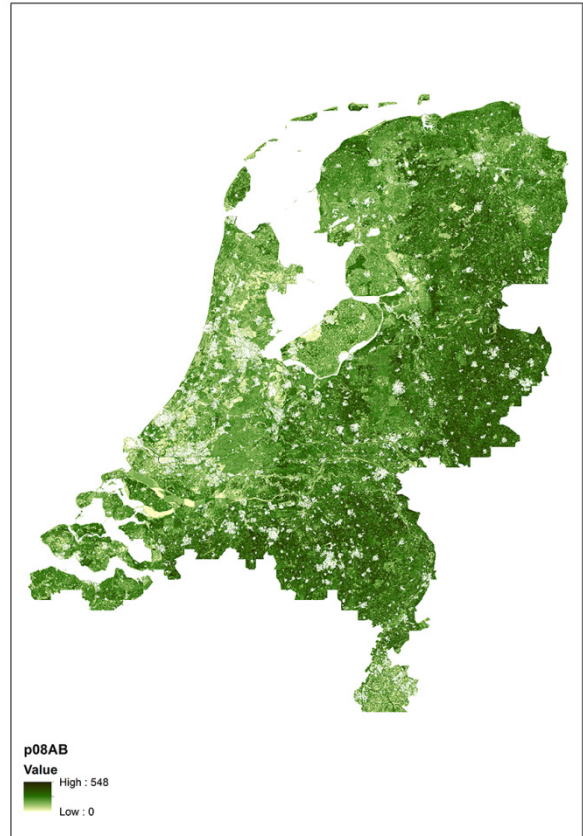


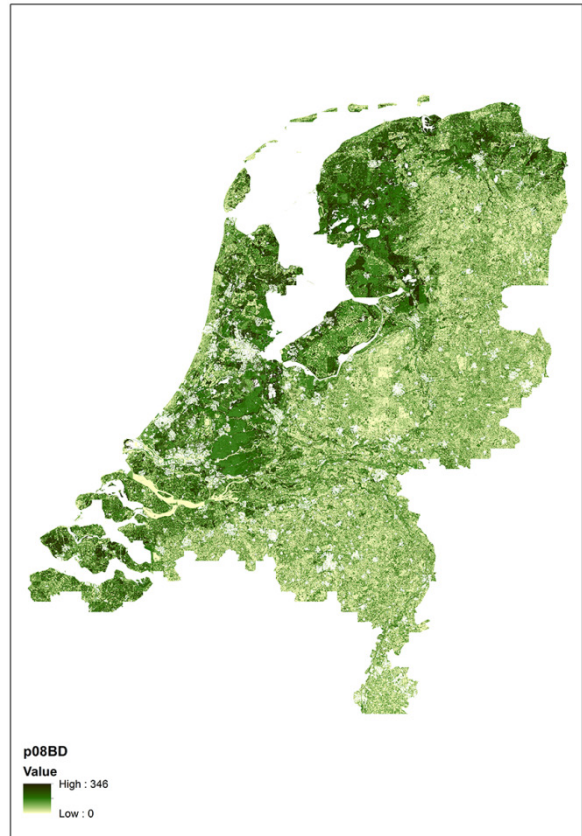


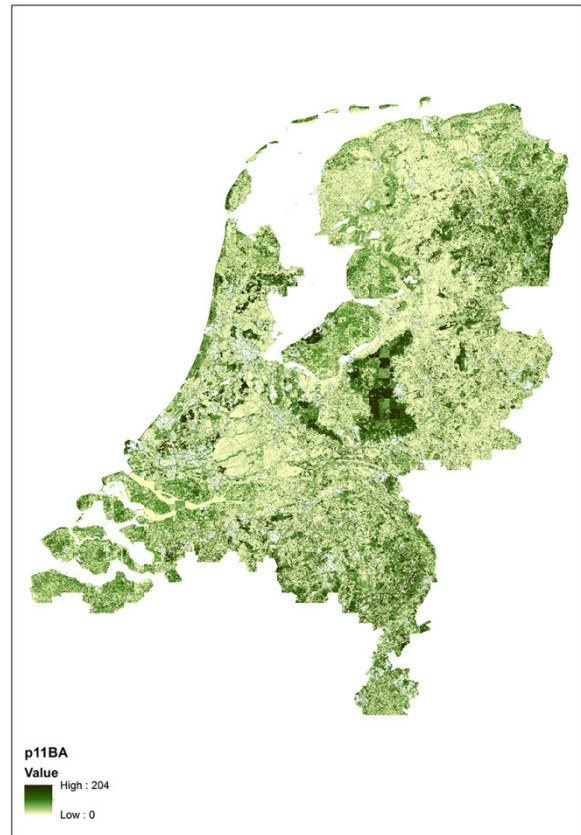
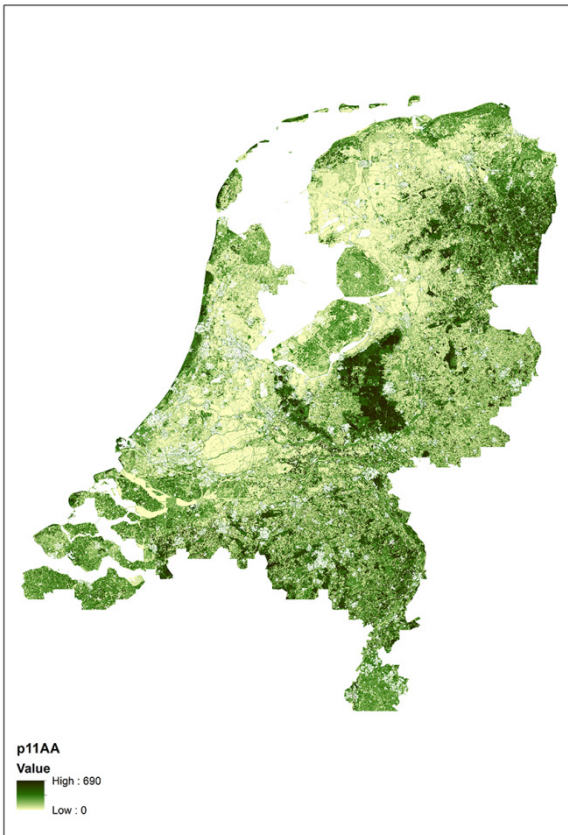
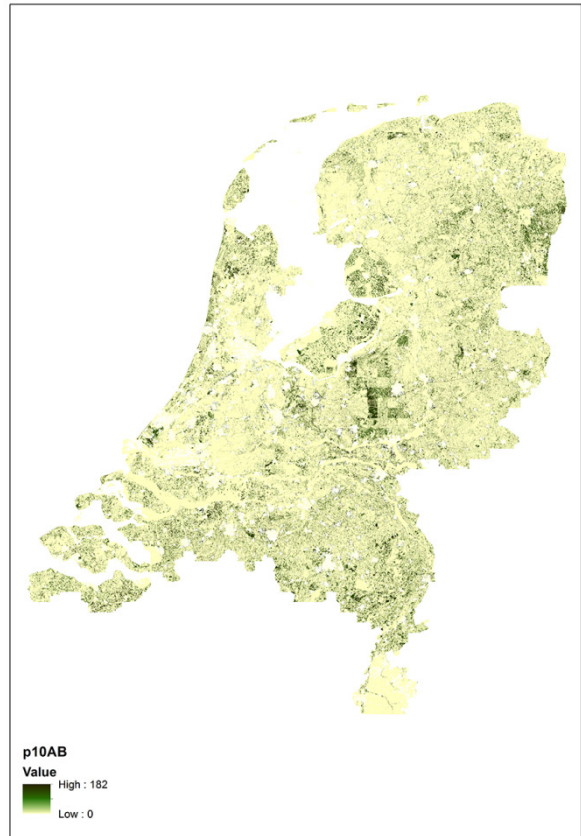
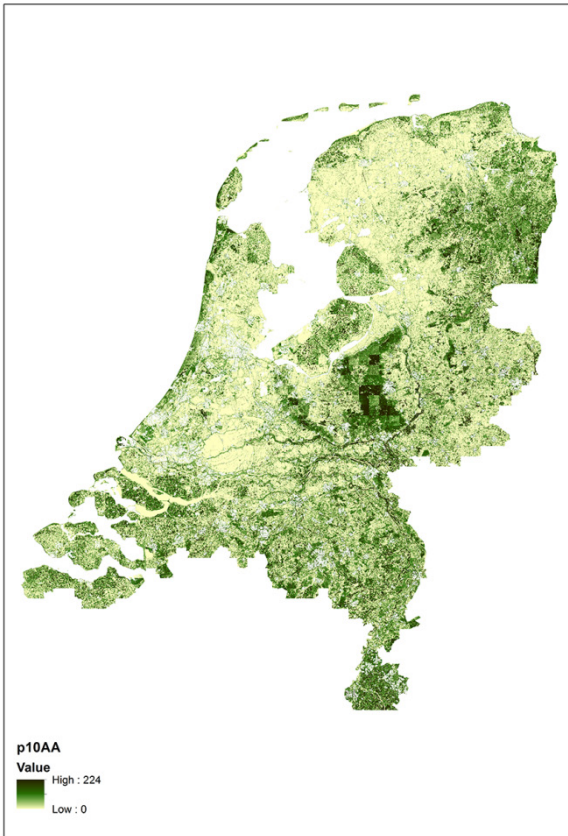


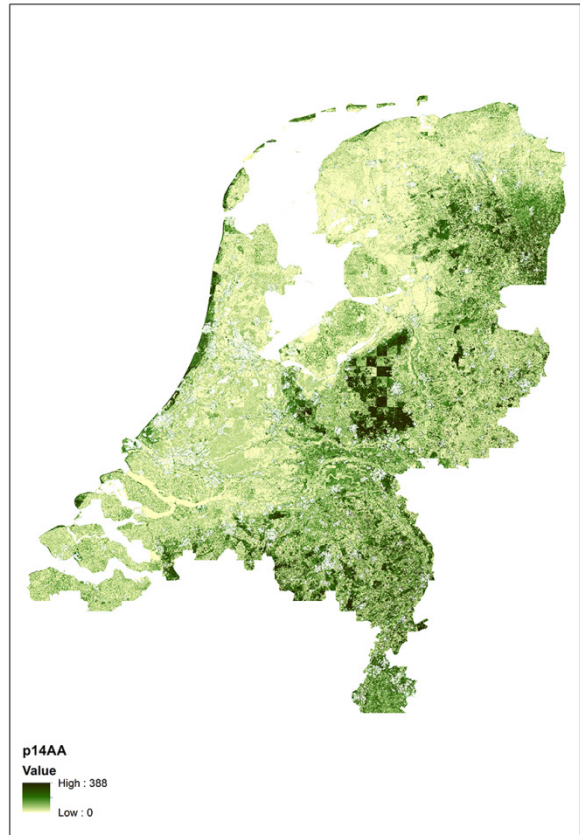
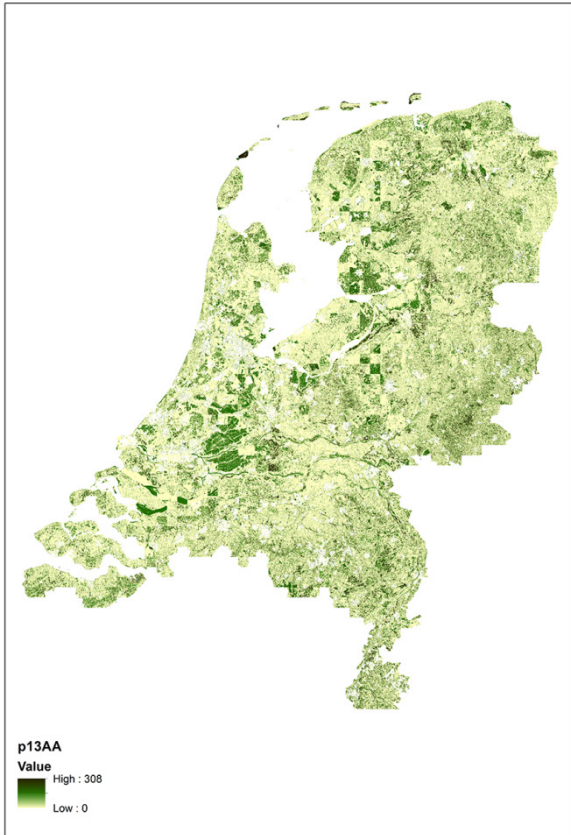
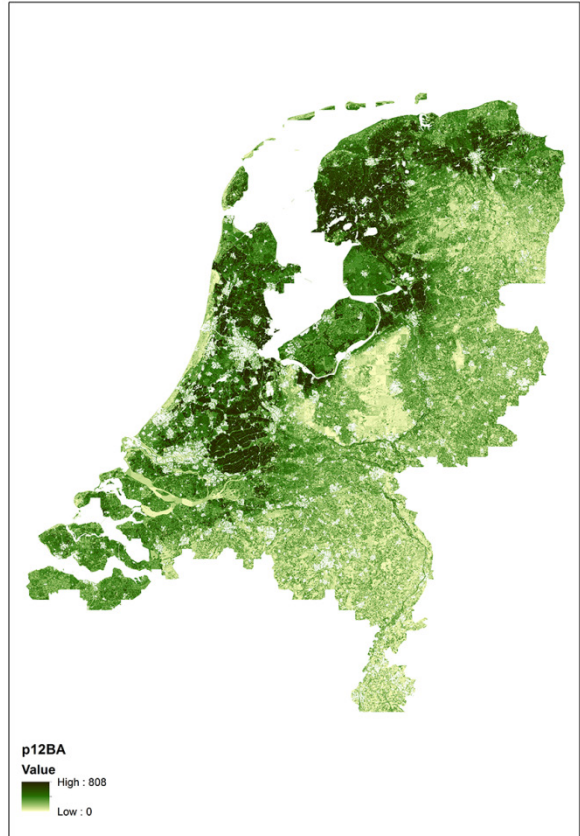


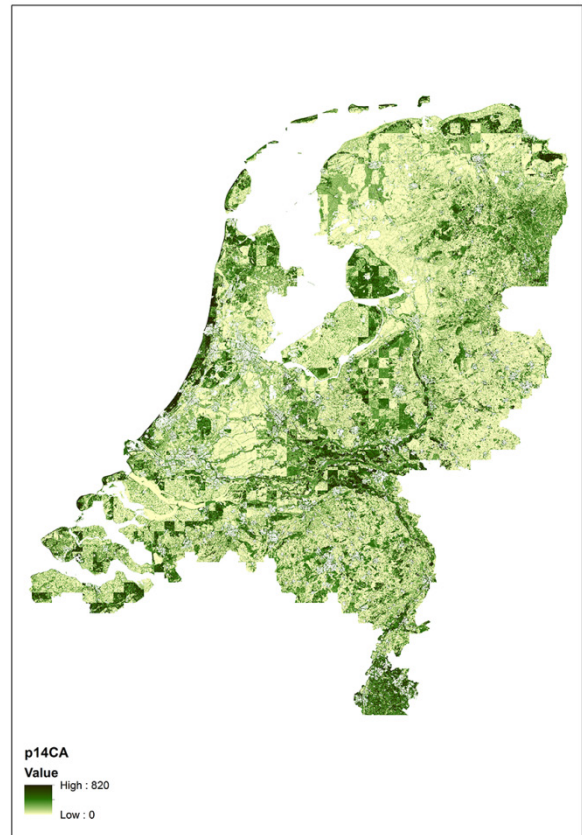
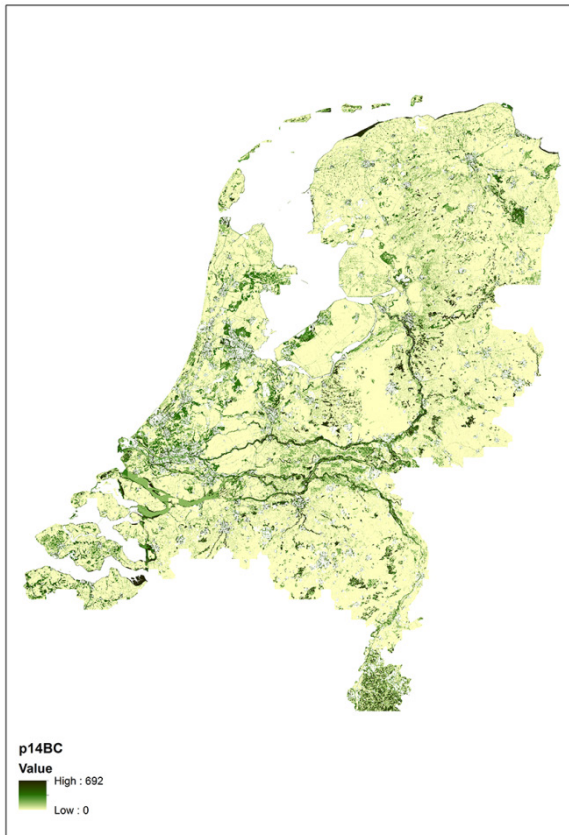
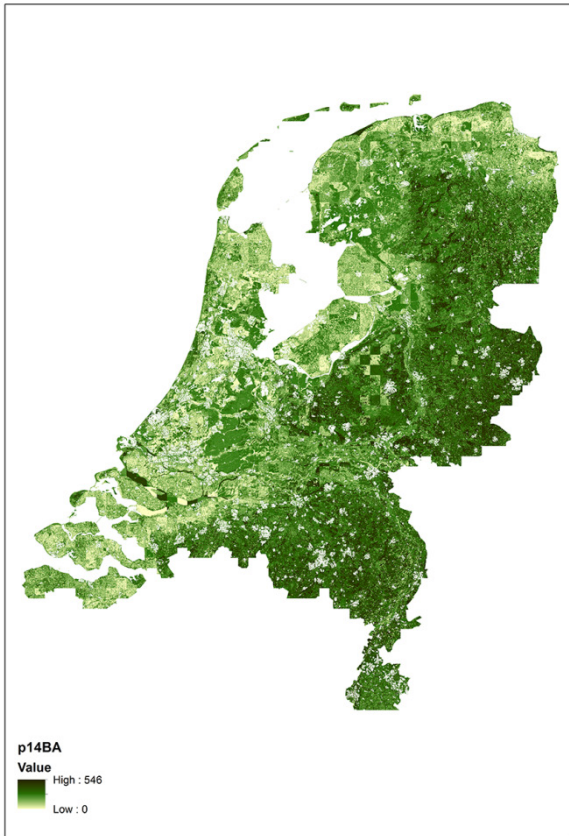


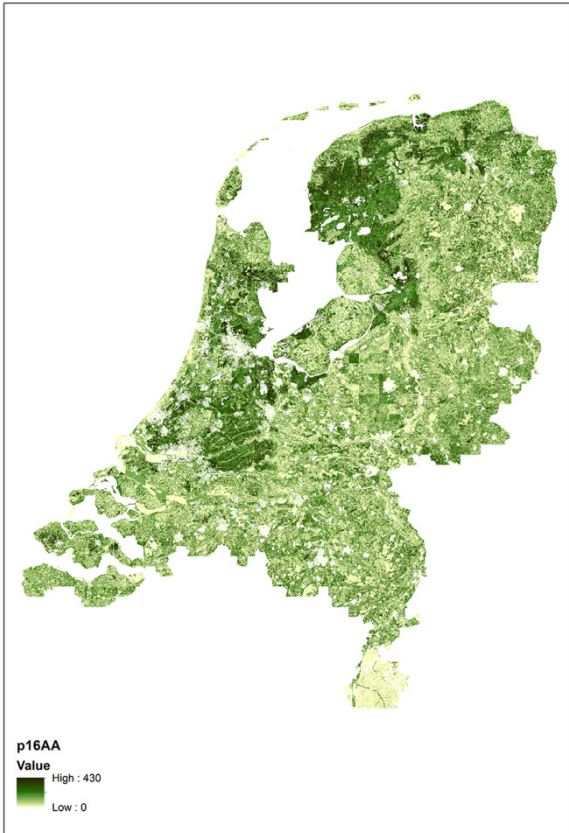
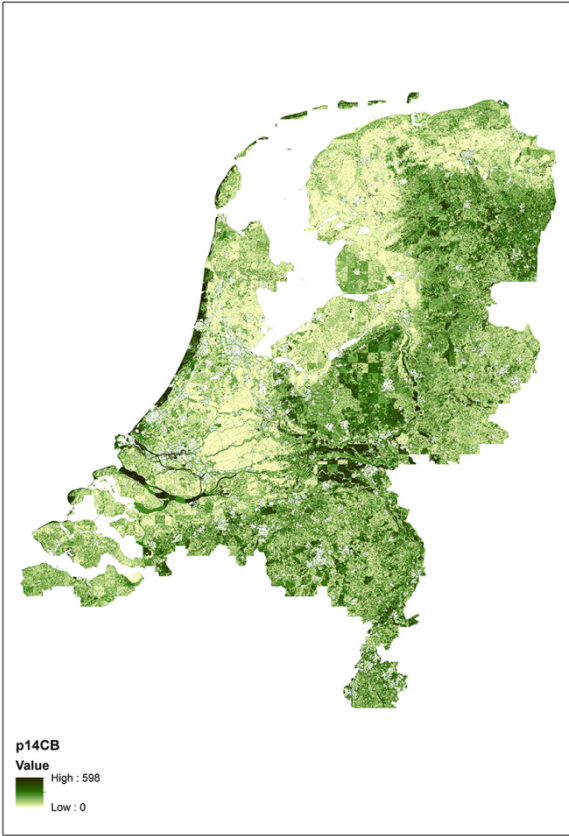


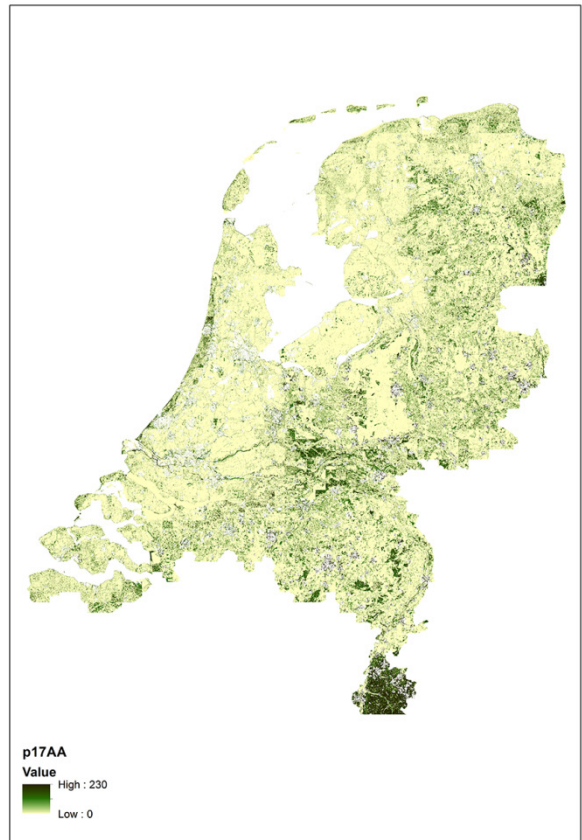
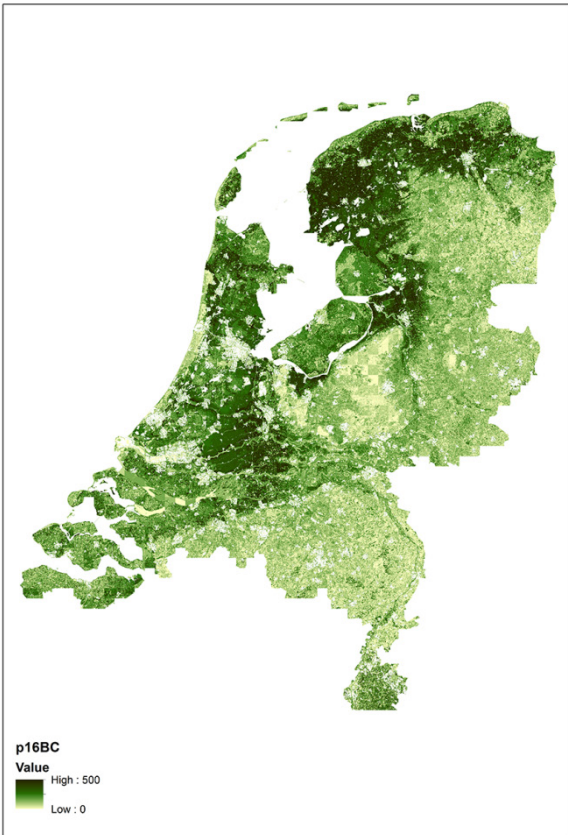
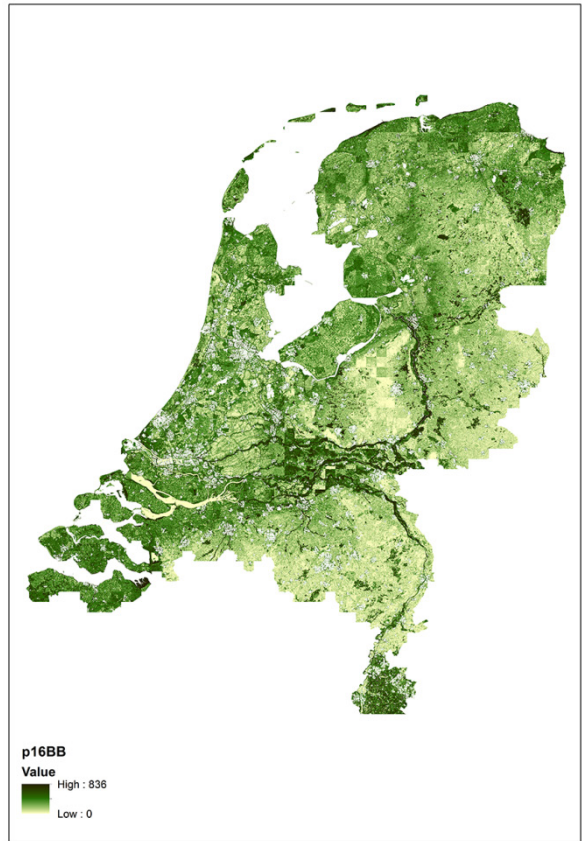
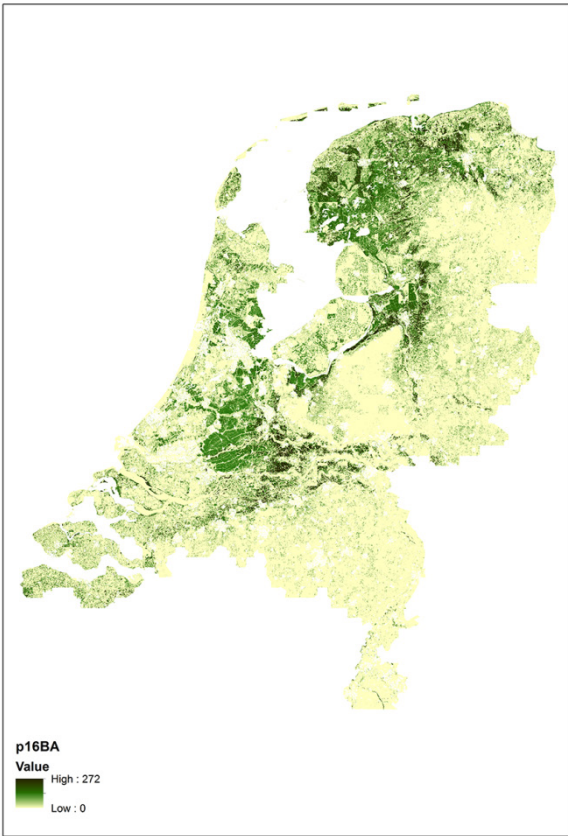


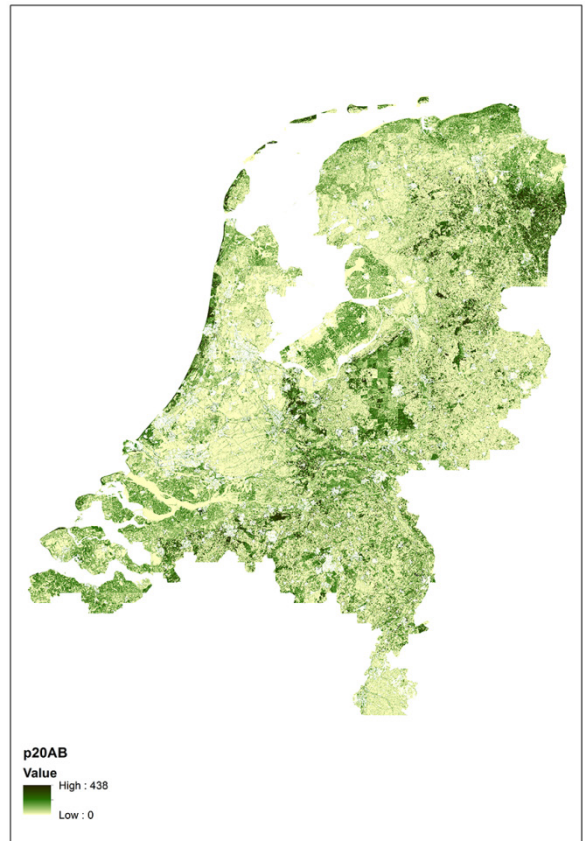
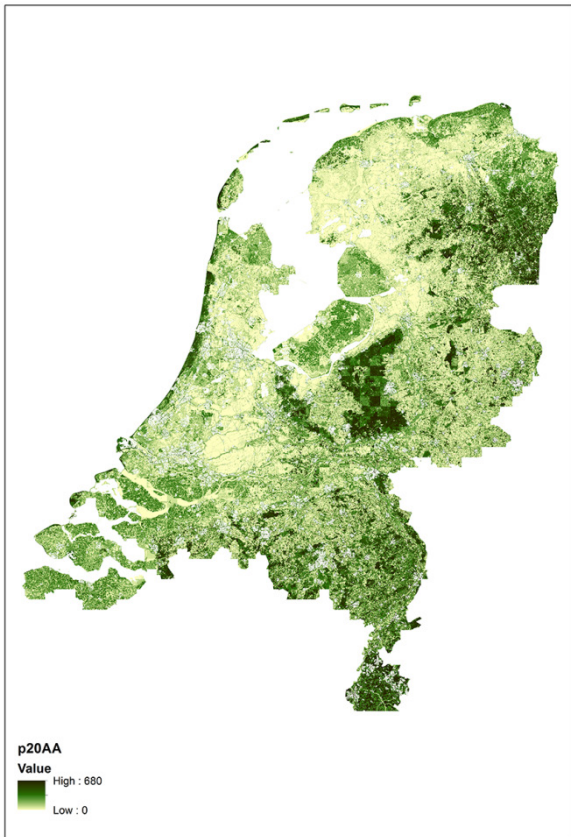
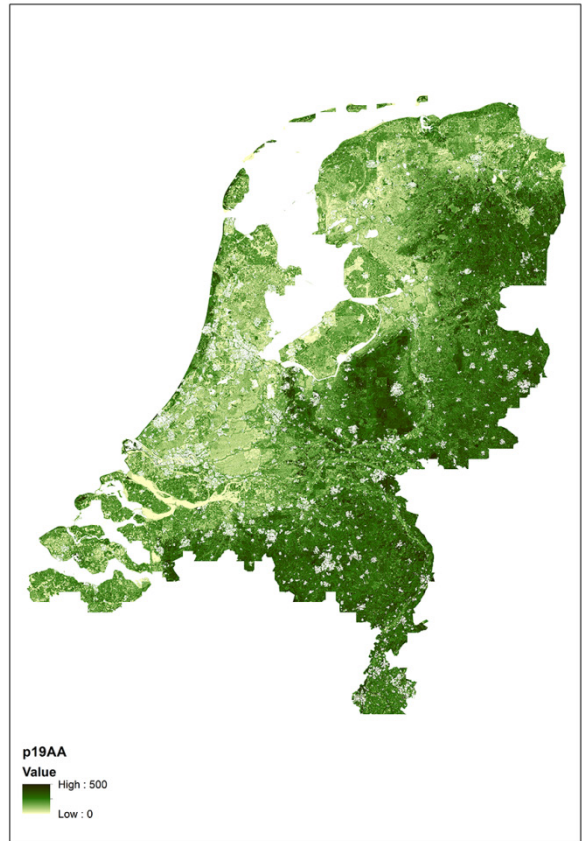
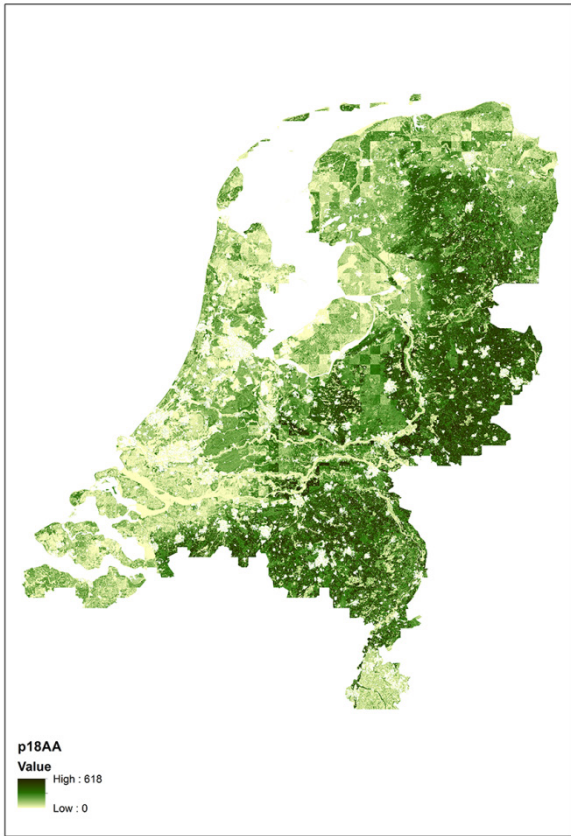


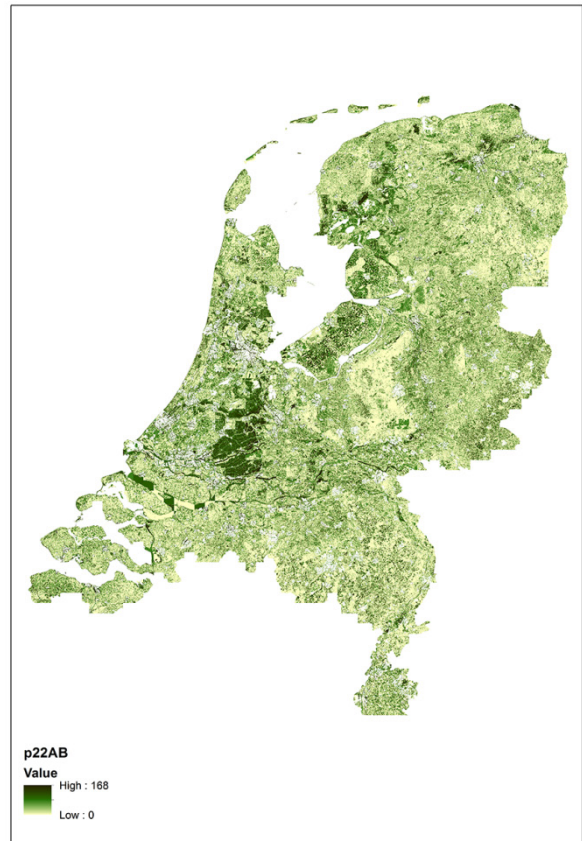
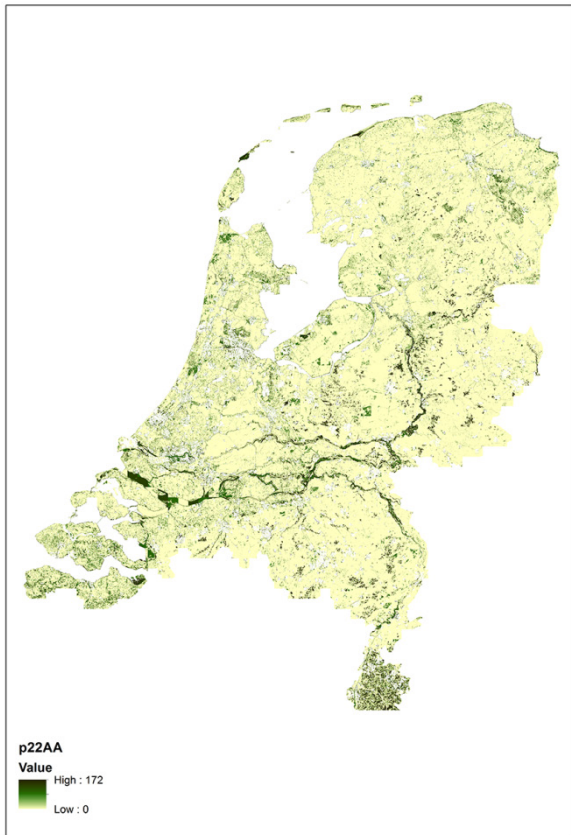


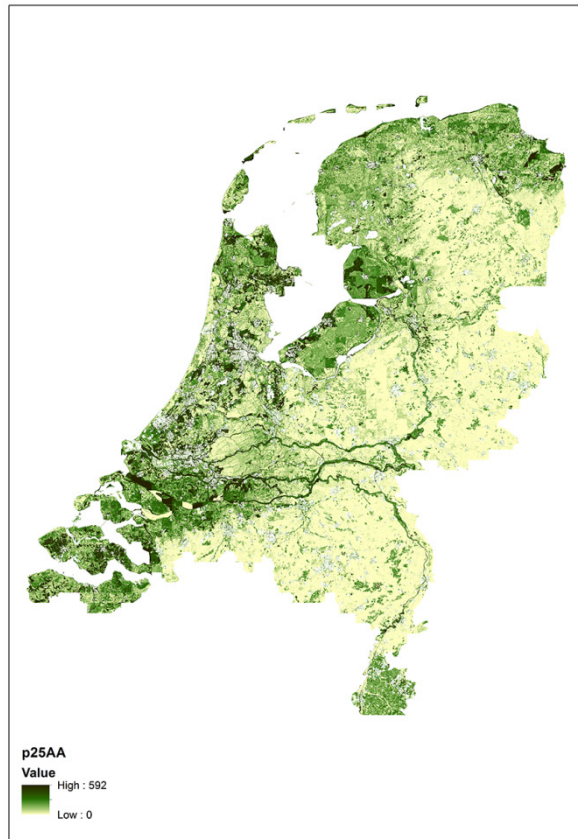
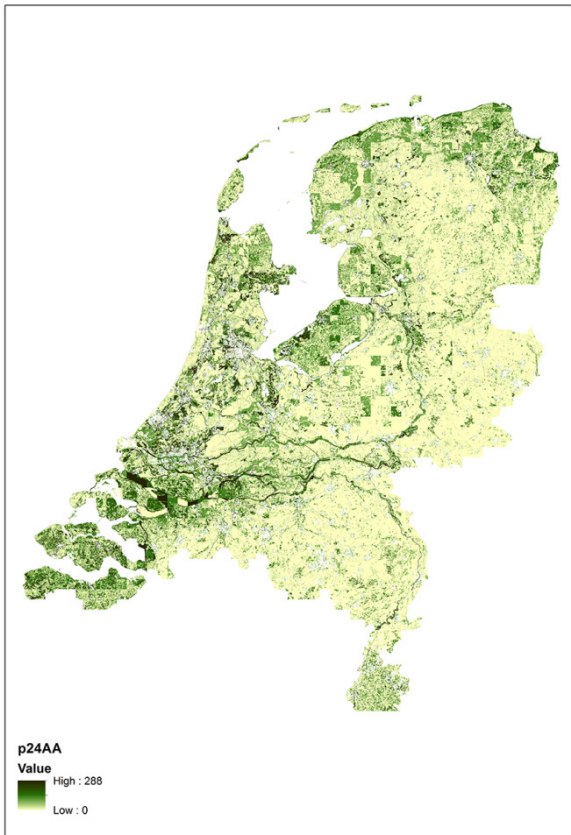
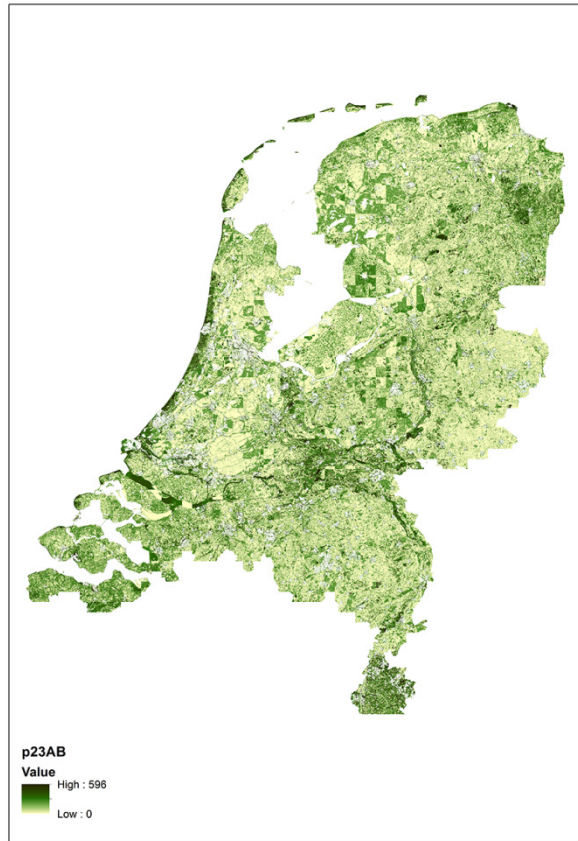
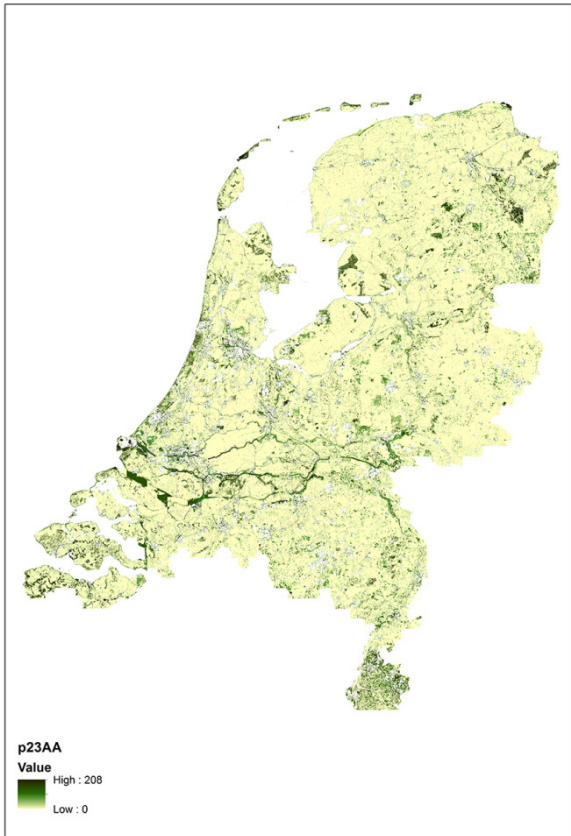


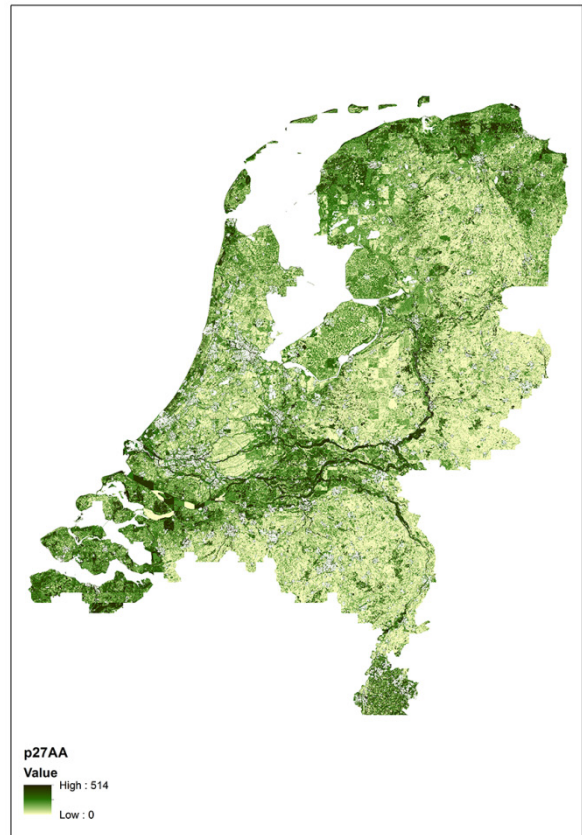
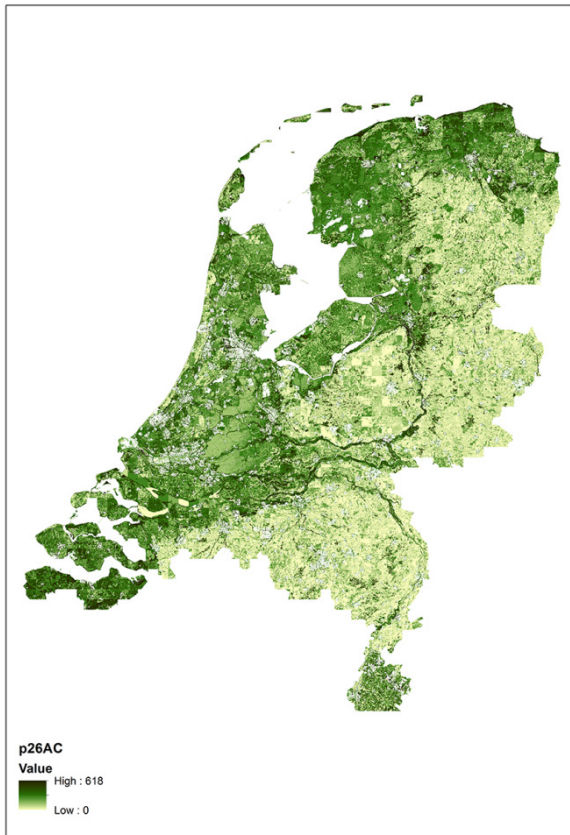
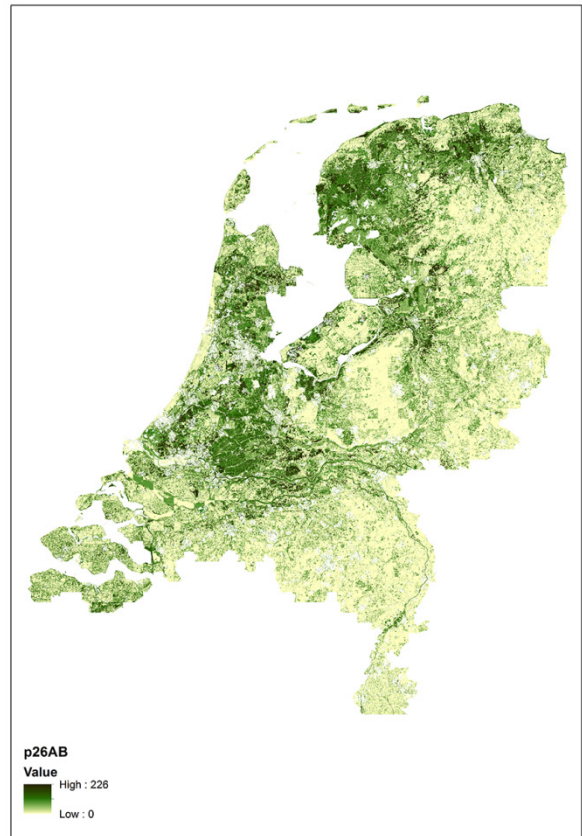
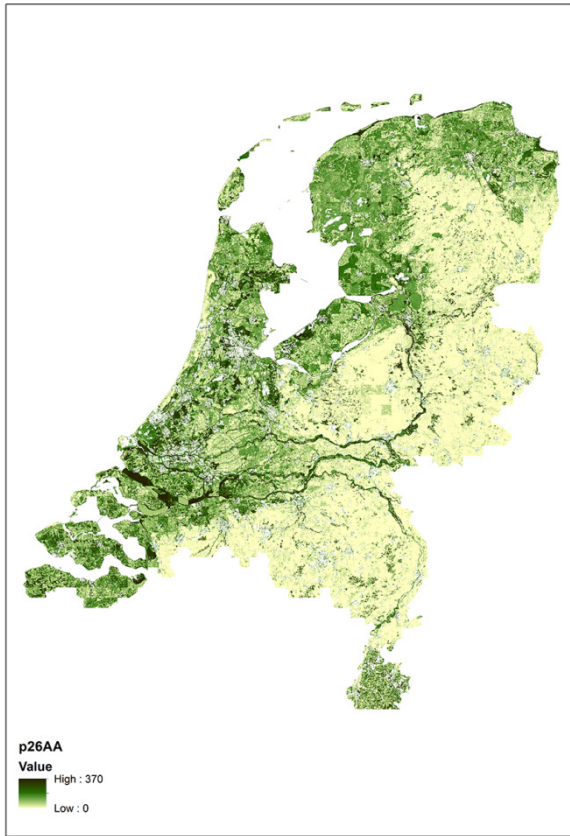


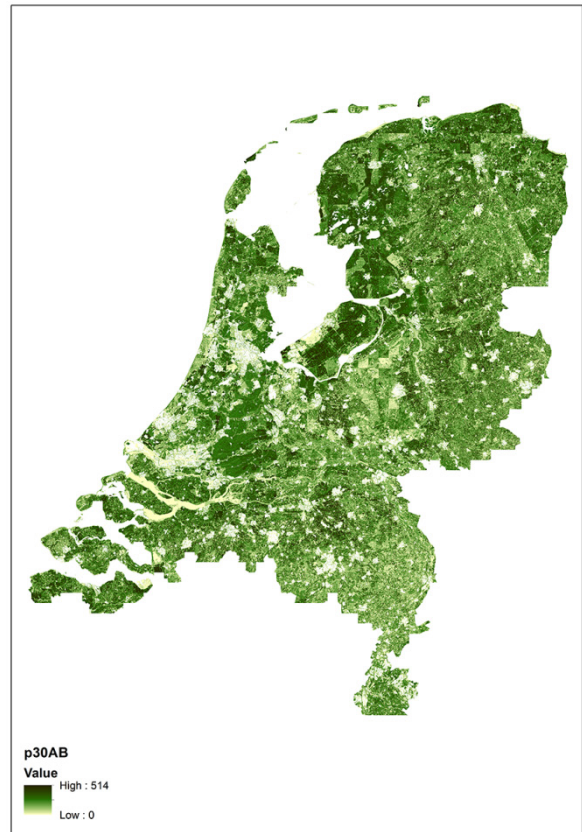
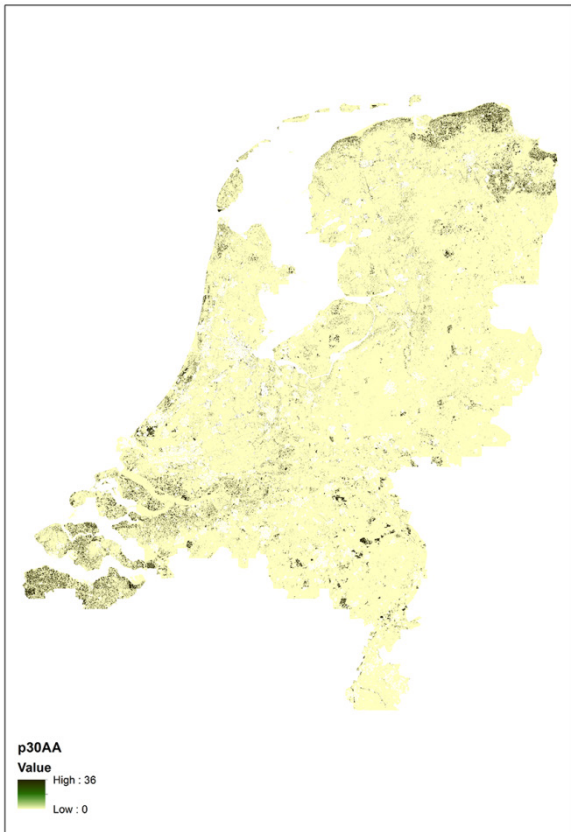
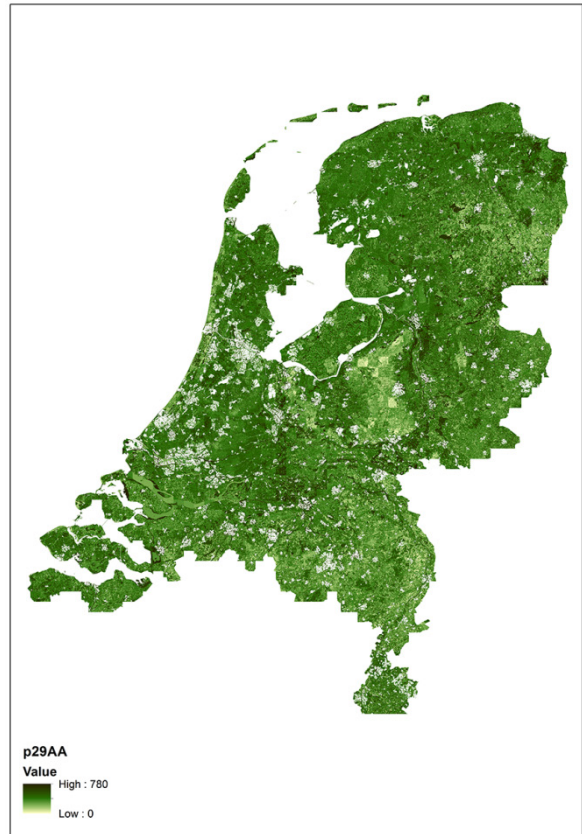
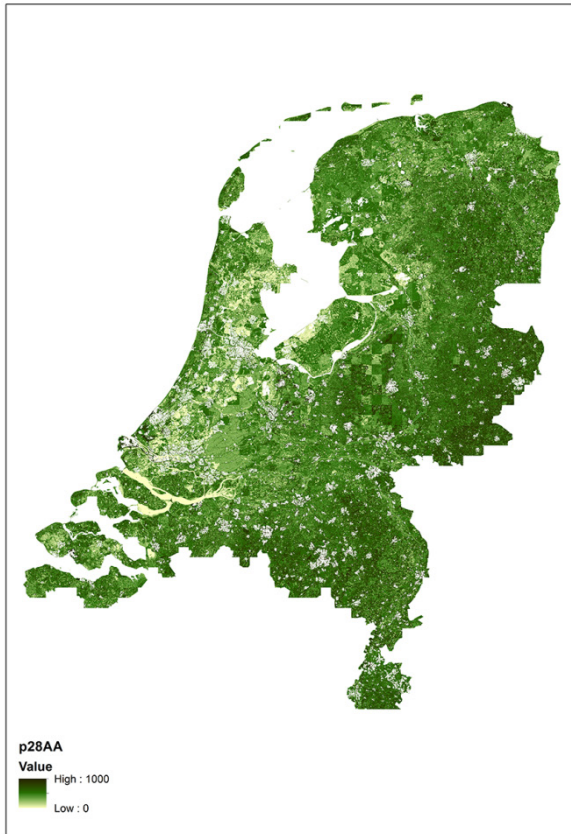


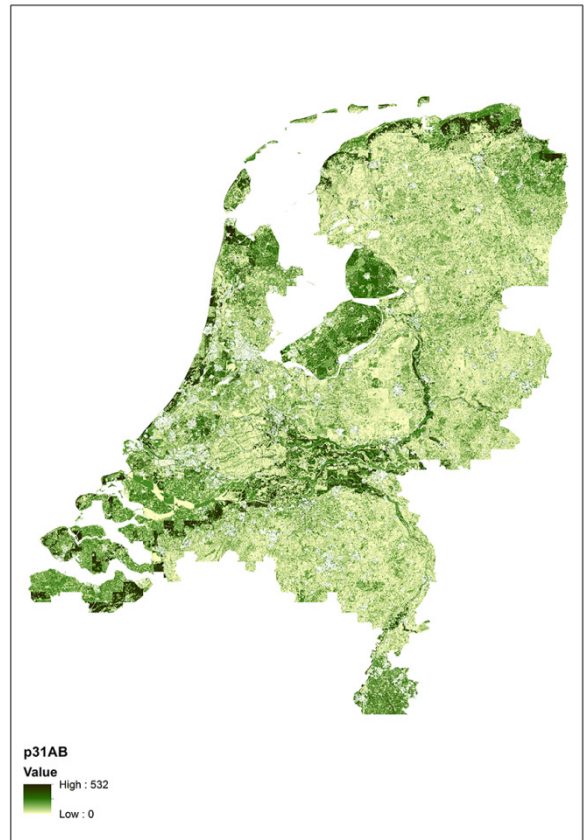
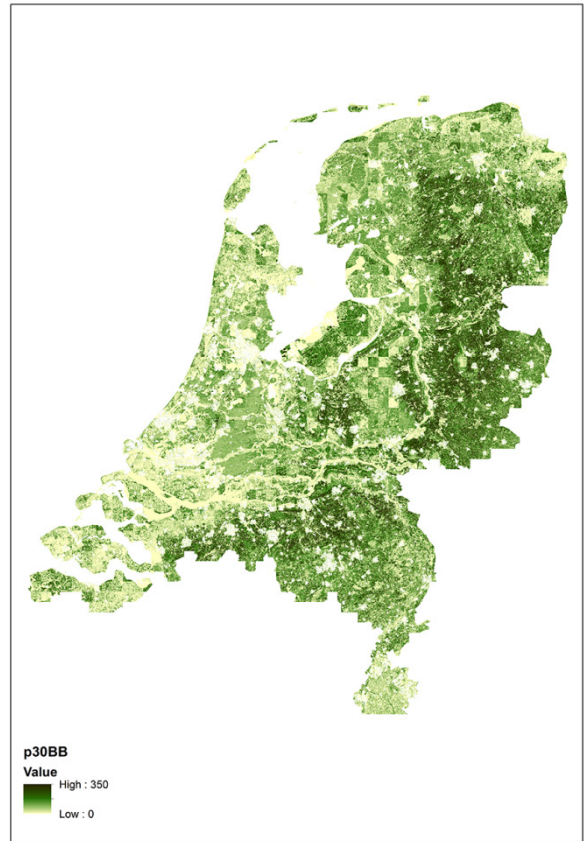
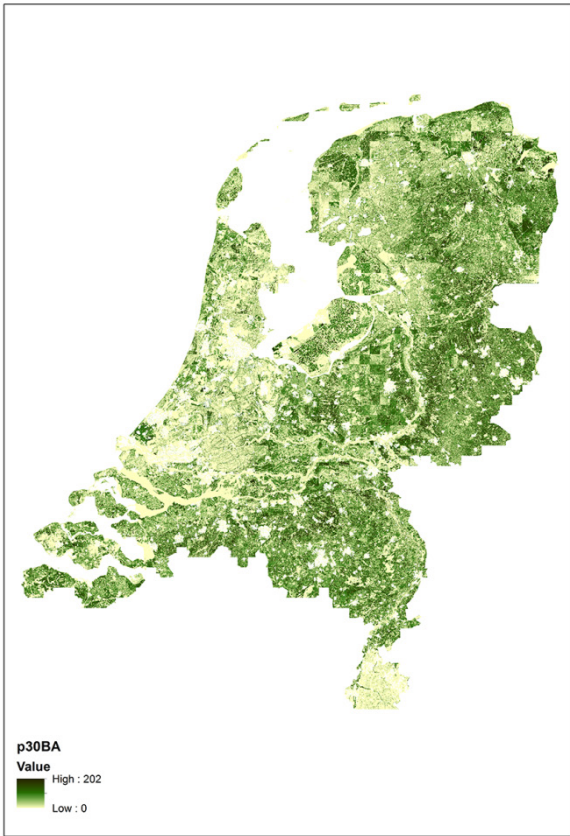


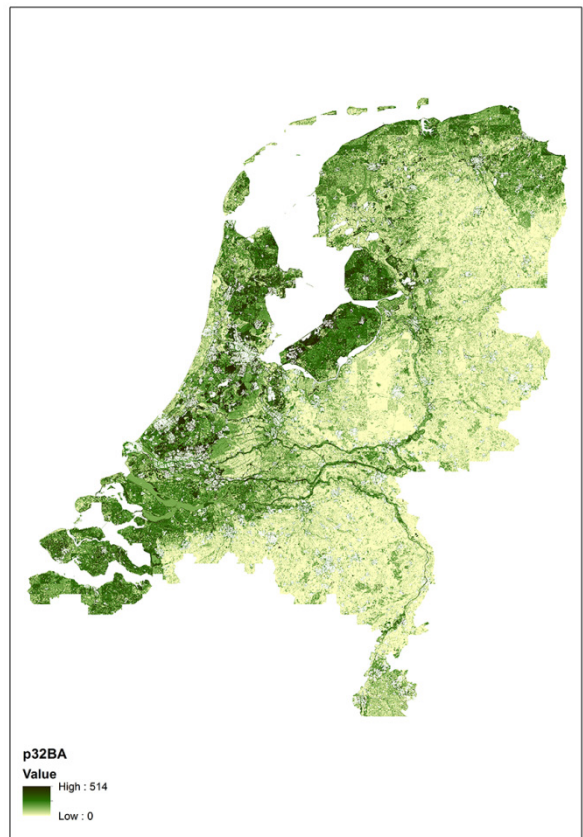
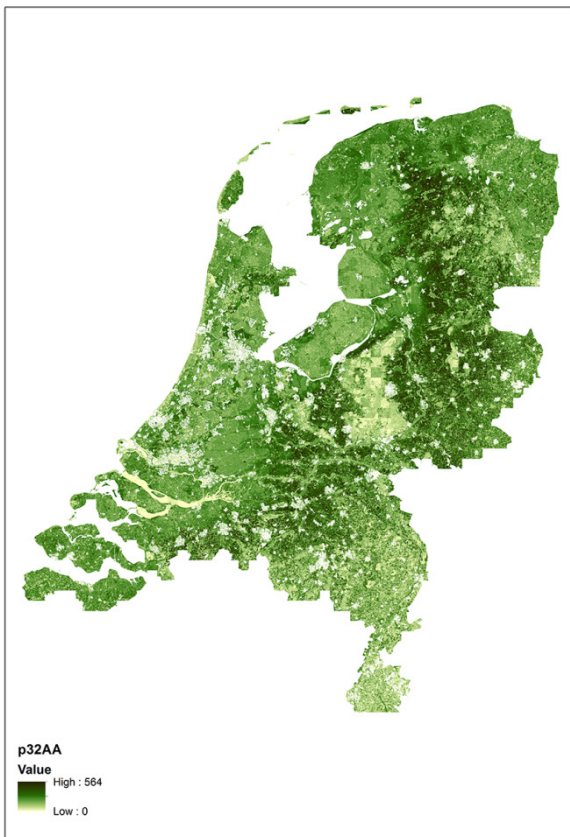
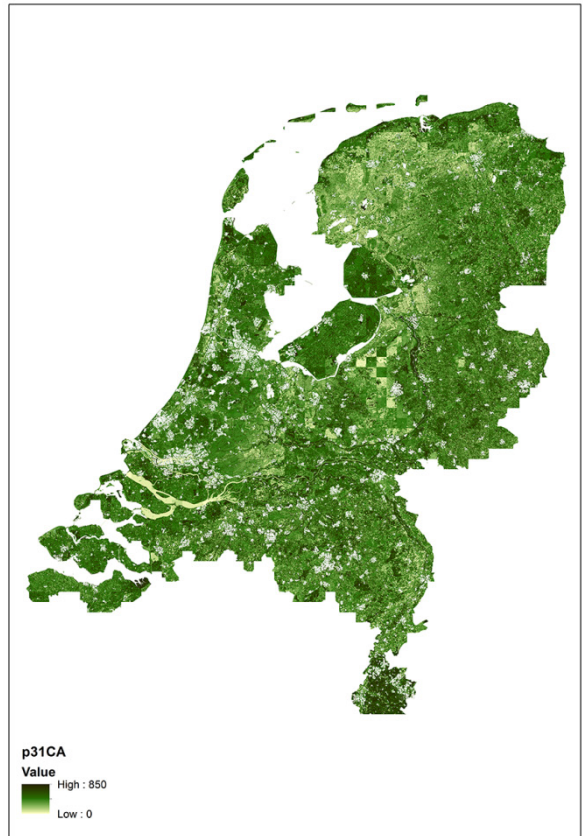
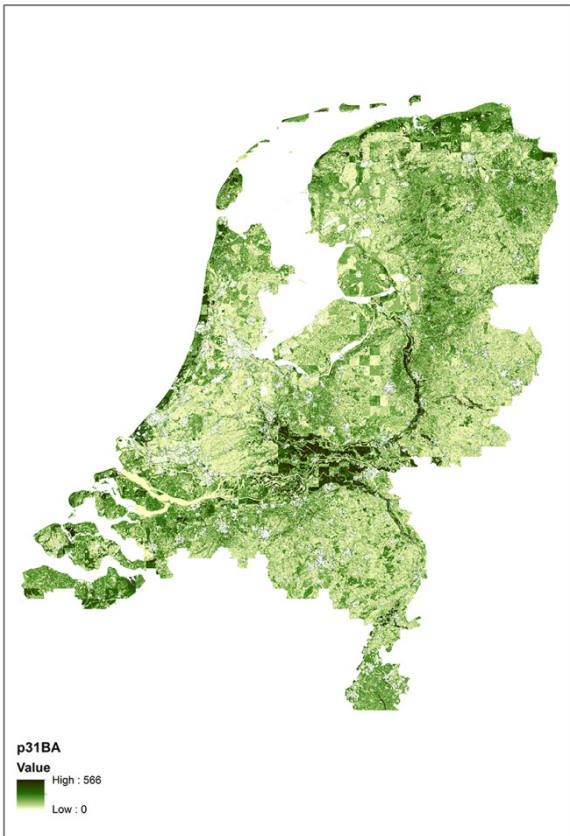




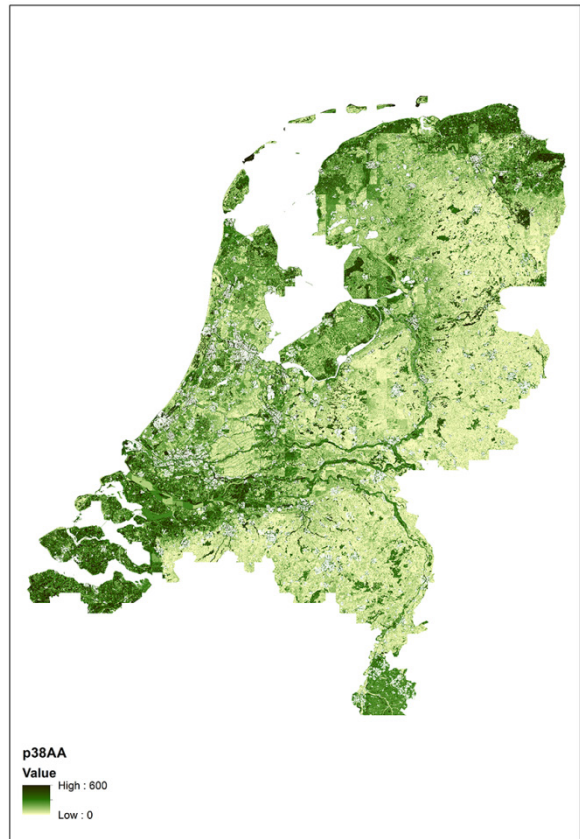
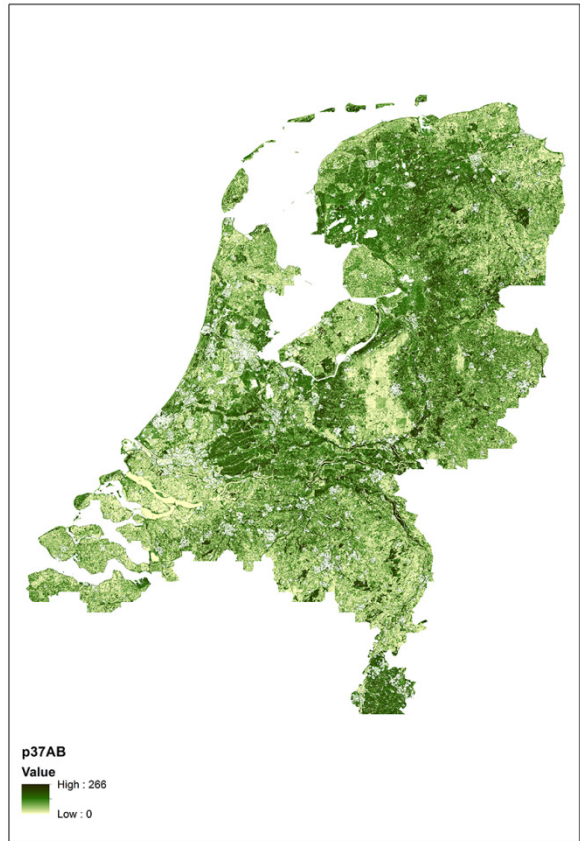
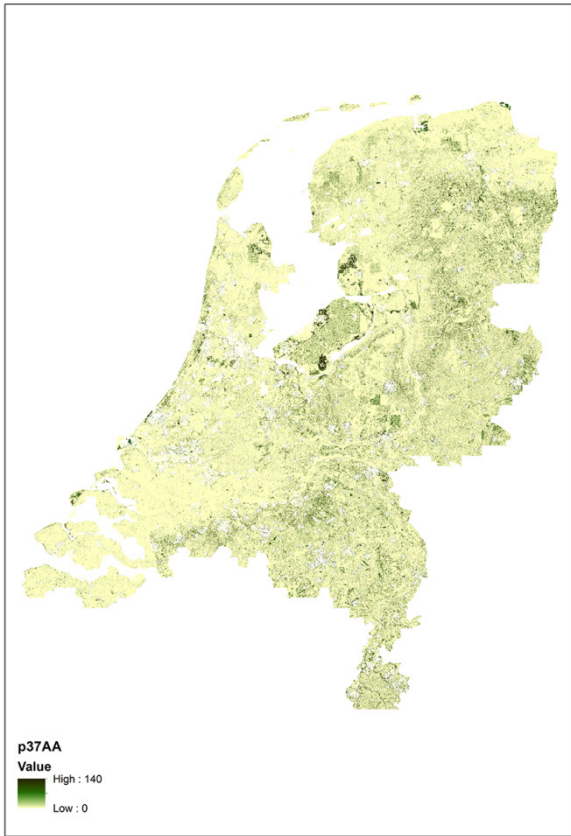


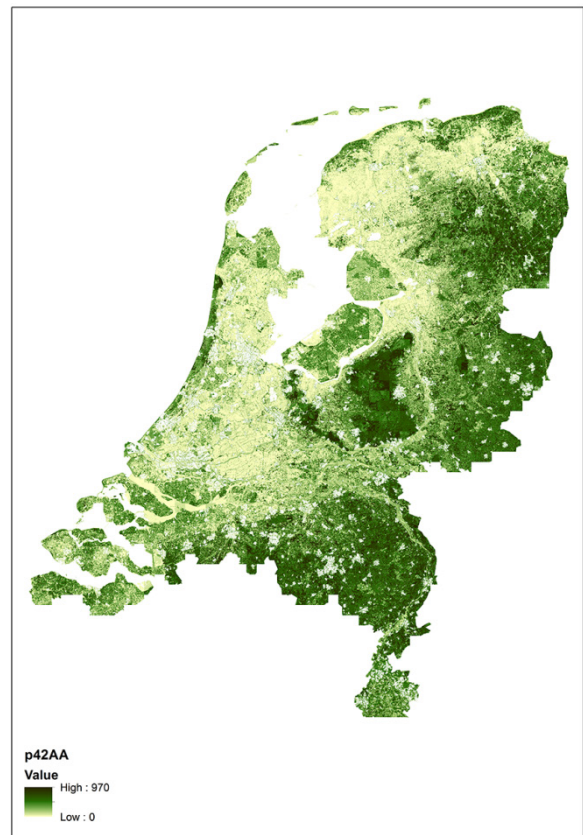
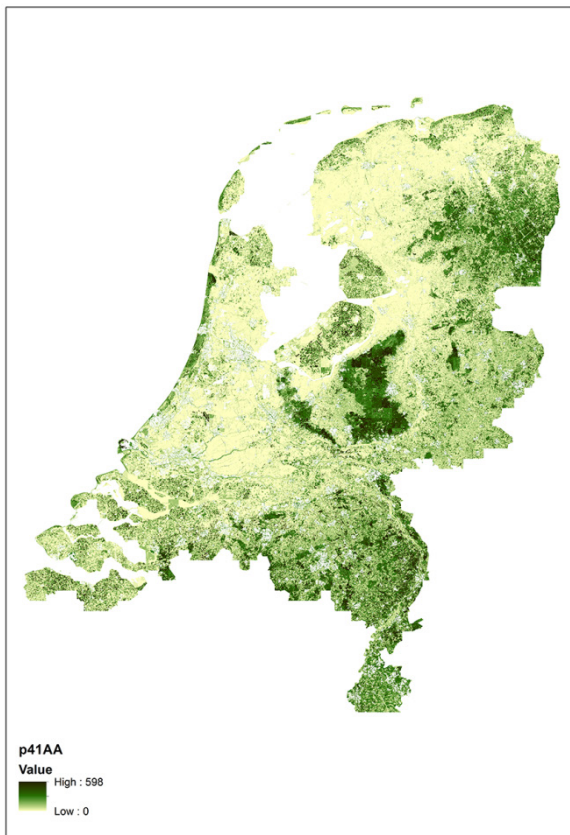
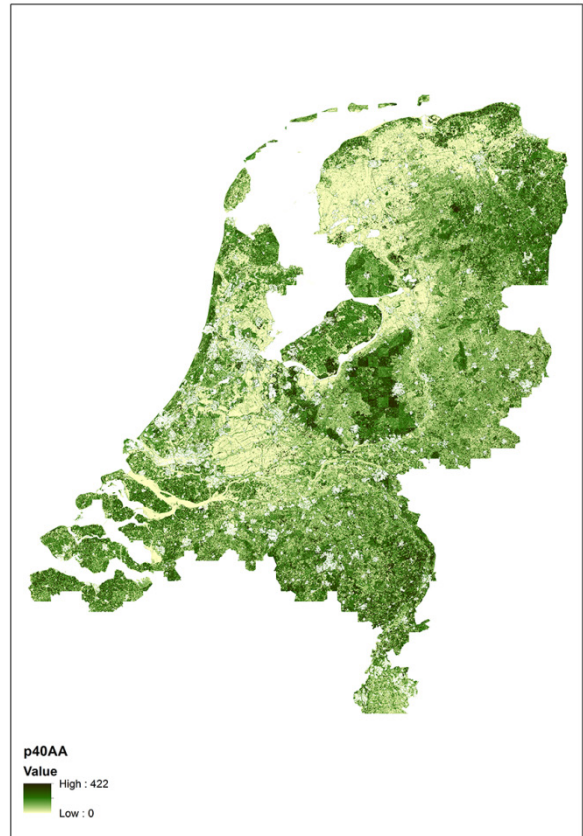
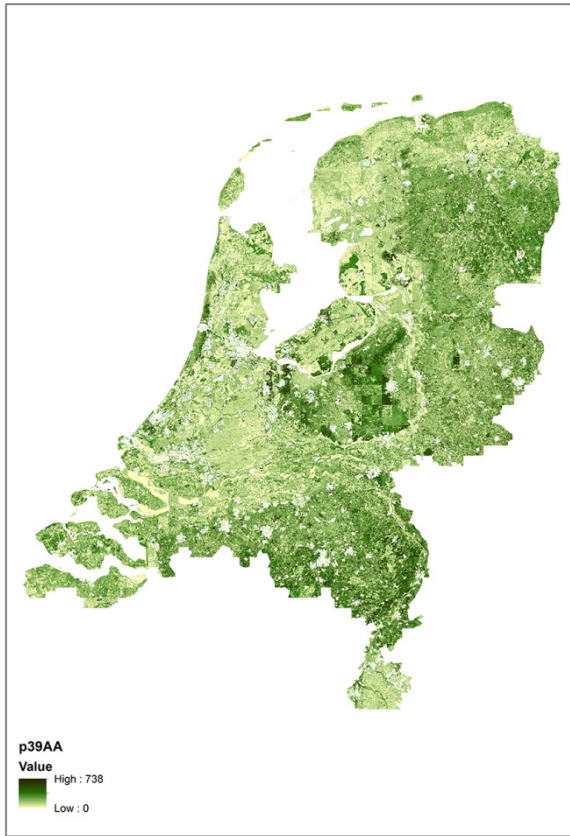


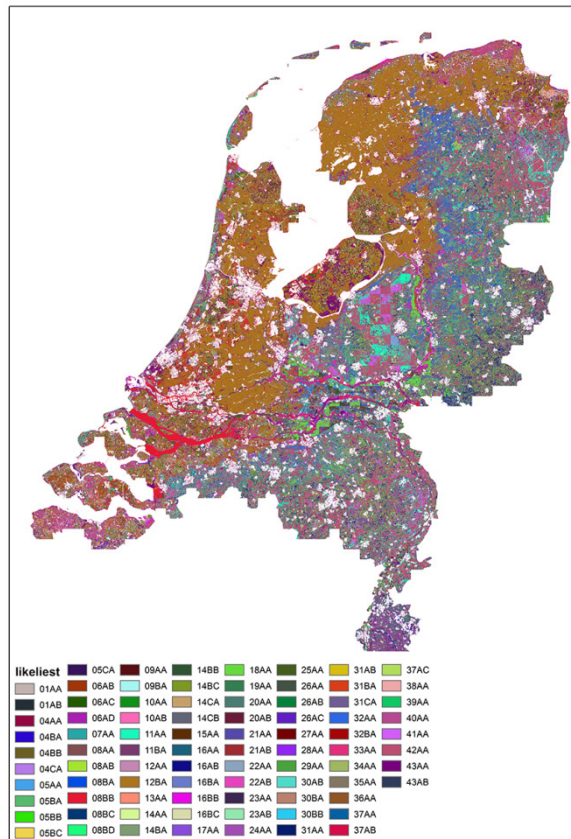
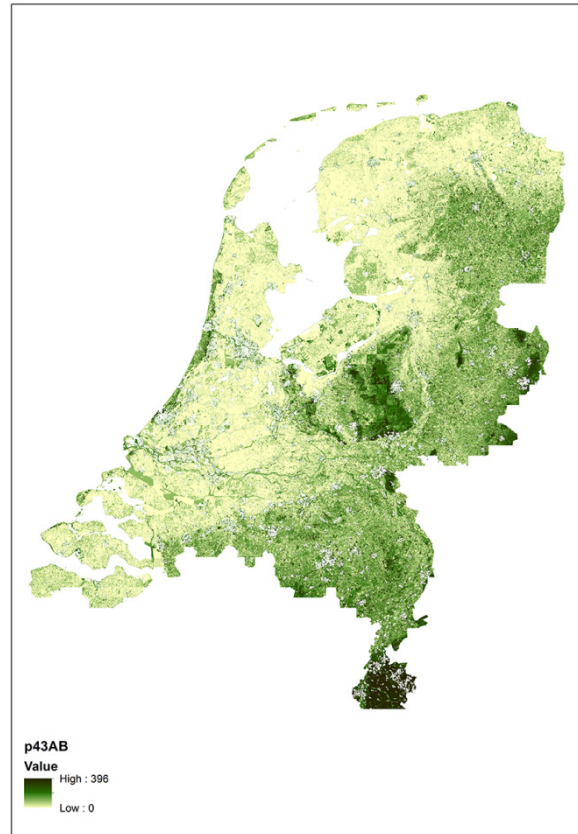
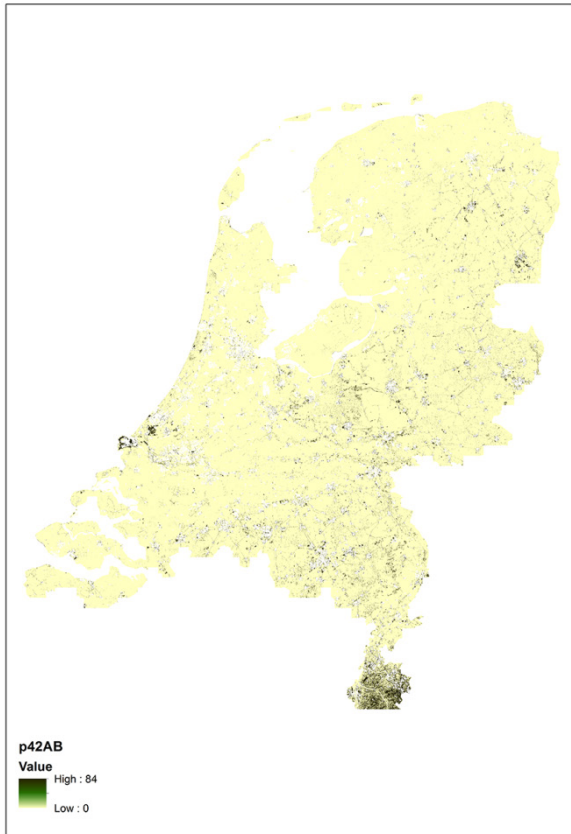












Alterra Wageningen UR
Postbus 47
6700 AA Wageningen
T 0317 48 07 00
www.wageningenUR.nl/alterra

Alterra-rapport 2582
ISSN 1566-7197



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

De missie van Wageningen UR (University & Research centre) is 'To explore the potential of nature to improve the quality of life'. Binnen Wageningen UR bundelen 9 gespecialiseerde onderzoeksinstituten van stichting DLO en Wageningen University hun krachten om bij te dragen aan de oplossing van belangrijke vragen in het domein van gezonde voeding en leefomgeving. Met ongeveer 30 vestigingen, 6.000 medewerkers en 9.000 studenten behoort Wageningen UR 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



Alterra Wageningen UR
Postbus 47
6700 AB Wageningen
T 317 48 07 00
www.wageningenUR.nl/alterra

Alterra-rapport 2582
ISSN 1566-7197

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

De missie van Wageningen UR (University & Research centre) is 'To explore the potential of nature to improve the quality of life'. Binnen Wageningen UR bundelen 9 gespecialiseerde onderzoeksinstituten van stichting DLO en Wageningen University hun krachten om bij te dragen aan de oplossing van belangrijke vragen in het domein van gezonde voeding en leefomgeving. Met ongeveer 30 vestigingen, 6.000 medewerkers en 9.000 studenten behoort Wageningen UR 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.

