



Voorspellend model voor besmettingsrisico van HPAI-virus bij commerciële pluimveebedrijven in Nederland

J.L. Gonzales, E. de Freitas Costa, W.H.G.J. Hennen, R. Petie en A.R.W. Elbers



WAGENINGEN
UNIVERSITY & RESEARCH

Rapport 2301353

31 Maart 2023

Voorspellend model voor besmettingsrisico van HPAI-virus bij commerciële pluimveebedrijven in Nederland

J.L. Gonzales¹, E. de Freitas Costa¹, W.H.G.J. Hennen², R. Petie¹ en A.R.W. Elbers¹

¹ Wageningen Bioveterinary Research, Lelystad; ² Wageningen Economic Research, Den Haag

Arjan Stegeman (Faculteit Diergeneeskunde, Universiteit Utrecht) was in dit project als adviseur betrokken. Gerdien van Schaik (Royal GD, Deventer) is bij de review van de eindresultaten betrokken geweest.

Dit onderzoek is uitgevoerd door Wageningen Bioveterinary Research in opdracht van - en gefinancierd door - het Ministerie van Landbouw, Natuur en Voedselkwaliteit, in het kader van het Beleidsondersteunend onderzoekthema 'Duurzame voedselvoorziening & -productieketens & Natuur' (Projectnr. BO-43-111-089/ KD-2022-073).

Wageningen Bioveterinary Research
Lelystad, 31 Maart 2023

Rapport 2301353

Dit rapport is gratis te downloaden op www.wur.nl/bioveterinary-research (onder Wageningen Bioveterinary Research publicaties).

© 2022 Wageningen Bioveterinary Research

Postbus 65, 8200 AB Lelystad, T 0320 23 82 38, E info.bvr@wur.nl, www.wur.nl/bioveterinary-research.
Wageningen Bioveterinary Research.

Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt worden door middel van druk, fotokopie, microfilm of op welke wijze dan ook zonder voorafgaande toestemming van de uitgever of auteur.

Wageningen Bioveterinary Research Rapport 2301353

Inhoud

Samenvatting	7
1 Achtergrond	9
2 Werkwijze	11
2.1 Data	11
2.1.1 Landschapsvariabelen en vogelsoortendichtheid	11
2.1.2 HPAIv in pluimvee	11
2.2 Identificatie van HPAI risicogebieden	13
2.2.1 Voorspelling (kunstmatige intelligentie) modellen	13
2.2.2 Selectie van Cases en Controls	15
2.2.3 Selectie en ontwikkeling van voorspelmodel	15
2.2.3.1 Random Forest model	15
2.2.3.2 Risicobeoordelingsmodel	16
3 Resultaten	17
3.1 Random Forest model	17
3.1.1 Verkenning van de variabelen	17
3.1.2 Model validatie en selectie	18
3.2 Risicobeoordelingsmodel	22
3.2.1 Identificatie van risicofactoren en Odds Ratio-Schatting	22
4 Discussiepunten	25
5 Conclusies	27
6 Aanbevelingen	29
Literatuur	31
Bijlage A1	33
Bijlage A2	35
Bijlage A3	39
Bijlage A4	43



Samenvatting

In dit onderzoek zijn voorspellende modellen gemaakt die een voorspelling geven van waar in Nederland het grootste risico – en daarmee ook het minste risico – aanwezig is voor het optreden van besmetting van hoog pathogene aviaire influenza virus (HPAIV) bij commerciële pluimveebedrijven. Daarbij wordt gekeken welke combinatie van landschapskarakteristieken, zoals de aanwezigheid van water en landbouw grasland, en dichtheid van wilde vogelsoorten de beste voorspelling geeft voor het optreden van HPAI-besmettingen op commerciële pluimveebedrijven. De vogelsoorten zijn geselecteerd gebaseerd op hun gevoeligheid voor HPAIV. Er zijn twee verschillende voorspelmodel methodes gebruikt:

1. kunstmatige intelligentie (KI, "machine learning") modellen en,
2. risicobeoordeling (pathway) model.

In dit onderzoek zijn we van het beoordelen van het besmettingsrisico op bedrijfsniveau naar het beoordelen van het besmettingsrisico op gebiedsniveau voor heel Nederland gegaan, door het besmettingsrisico te kwantificeren in een klein gebied (100 x 100 m, 1 ha) dat groot genoeg is om het gebied te representeren rond een pluimveebedrijf en klein genoeg om over het algemeen maar 1 pluimveebedrijf te bevatten. Heel Nederland is daarom opgedeeld in een rooster met cellen van één hectare groot, iedere cel die een pluimveebedrijf bevat is gebruikt om correlaties te vinden tussen HPAIV besmettingen en landschapskarakteristieken en vogeldichtheden. Met deze aanpak kan het besmettingsrisico ook worden voorspeld voor gebieden in Nederland waar geen pluimveebedrijven aanwezig zijn.

Voor elke rastercel van 1 hectare werd informatie van 25 landschapskarakteristieken en de dichtheid (aantallen per oppervlakte-eenheid) van 82 wilde vogelsoorten verzameld. Allereerst werden verschillende kunstmatige intelligentie modellen ontwikkeld en onderzocht op hun voorspellende kwaliteiten. Het best presterende model was het Random Forest model. Met deze methode zijn vervolgens zeven voorspelmodellen geconstrueerd met verschillende variabelen als input. Daarna werd nogmaals het best presterende model geselecteerd, wat een gevoeligheid had van 83% en een specificiteit van 74%. In deze context duidt de gevoeligheid op het vermogen van het model om een HPAI-besmetting in een pluimveebedrijf correct te voorspellen. De specificiteit verwijst daarentegen naar het correct voorspellen welke gebied/bedrijf niet-besmet is. Met het best presterende model is een risicokaart van Nederland gemaakt.

Ten tweede is er een risicobeoordelingsmodel ontwikkeld. Voor dit model zijn 18 landschapsvariabelen en vogel-dichtheidsvariabelen gebruikt die een significant associatie met het HPAI-besmettingsrisico in pluimveebedrijven hadden. Opnieuw is er een risicokaart gemaakt op basis van de resultaten van deze analyse.

Zowel het Random Forest model als het risicobeoordelingsmodel komen overeen in de identificatie van gebieden met hoge of lage risico's van HPAI-besmetting van commerciële pluimveebedrijven:

- Een hoog risico op HPAI-besmetting bij pluimveebedrijven is te zien in de kustregio's en laaggelegen gebieden in Groningen, Friesland, Flevoland, Noord- en Zuid-Holland en Zeeland. Verder is er een verhoogd risico op HPAI-besmetting bij pluimveebedrijven te zien in de stroomgebieden van de grote rivieren in het midden van het land: de IJssel, Neder-Rijn, Waal en Maas. Verder is er een hoog risico op HPAI-besmetting bij pluimveebedrijven te zien in Limburg ter hoogte van de Maasplassen bij Roermond.
- Er is een laag risico op HPAI-besmetting bij pluimveebedrijven te zien voor de hooggelegen delen van Groningen, Drenthe, Overijssel, Gelderland, Utrecht, Noord-Brabant en Limburg.
- De aanwezigheid van water en agrarische graslanden in de directe omgeving van pluimveebedrijven vergroten het risico op HPAI-besmetting bij pluimveebedrijven. Dit risico neemt tevens toe als de oppervlakte aan water en grasland in de directe omgeving van het pluimveebedrijf toeneemt.
- De aanwezigheid van Wilde Eenden, Bergeenden, Slobeenden, Brandganzen en Knobbelswanen in de directe omgeving van pluimveebedrijven lijkt geassocieerd met het risico op HPAI-besmettingen bij pluimveebedrijven.

De risicokaarten uit dit project kunnen gebruikt worden om surveillance te verbeteren, door actievere vormen van surveillance op de aanwezigheid van HPAIv onder wilde vogels en op pluimveebedrijven te vergroten in de gebieden met een hoog HPAI-besmettingsrisico op pluimveebedrijven.

Verder kunnen de risicokaarten van nut zijn bij het plannen van de ruimtelijke inrichting van het landschap (nieuwe locaties van te starten pluimveebedrijven).

De voorspellingen van HPAI-besmettingsrisico's zouden moeten worden bijgewerkt wanneer veranderingen in landgebruik (landschapskarakteristieken) of veranderingen in dichtheid van wilde vogelsoorten in gebieden in Nederland in de toekomst worden waargenomen.

1 Achtergrond

Recent is er een rapport uitgekomen m.b.t. risicofactoren voor introductie van hoog-pathogene aviaire influenzavirus (HPAIV) bij commerciële pluimveebedrijven in Nederland in de periode 2014-2022. (WUR rapport 2211632 door Gonzales et al., 2022). In het rapport zijn kaarten opgenomen met de kans op introductie van HPAIV in Nederland, waarin risicofactoren de variatie verklaren in HPAIV-uitbraken op commerciële pluimveebedrijven. Het doel van de analyse was het bepalen van het risico op introductie op bedrijfsniveau, gebruikmakend van specifieke data per bedrijf. De afstand tot water was bijvoorbeeld bepaald voor ieder pluimveebedrijf in Nederland. De hoge mate van detail op bedrijfsniveau beperkte de mogelijkheid tot extrapolatie of voorspellingen te doen m.b.t. het HPAI-introductierisico voor regio's waar geen pluimveebedrijven voorkomen omdat er aannames gemaakt moesten worden over de locatie van het imaginaire pluimveebedrijf en het landgebruik rond het imaginaire pluimveebedrijf. Een vervolgvraag vanuit het Ministerie van Landbouw, Natuur en Voedselkwaliteit (LNV) luidde of er risicokaarten gemaakt kunnen worden die een voorspelling geven waar in Nederland het grootste HPAI-besmettingsrisico is voor commerciële pluimveebedrijven. Een dergelijke risicokaart zou van nut kunnen zijn bij het plannen van surveillance en ruimtelijke inrichting.

In vergelijking met het risicofactorenonderzoek werd bij het voorspelmodel minder nadruk gelegd op hoe significant een variabeel geassocieerd is met het besmettingsrisico. De uiteindelijke landschapsvariabelen in het risicofactorenonderzoek hadden een significantie van minimaal $p < 0.05$. In het voorspelmodel onderzoek werd er gekeken welke combinatie van "relevante" variabelen de beste voorspelling gaf voor het optreden van HPAIV-besmetting op commerciële pluimveebedrijven. Daarbij konden ook risicofactoren, die in het risicofactorenonderzoek niet significant waren ($p > 0.05$) toch een bijdrage leveren aan het voorspelmodel waardoor het uiteindelijke voorspellend vermogen werd gemaximaliseerd. Een mogelijke verbetering die werd gebruikt bij het voorspelmodel in vergelijking met het risicofactorenonderzoek was dat voor wat betreft de wilde vogeldichtheidsgegevens gebruik gemaakt werd van soortgegevens, en niet van soortgroepen (dus b.v. Wilde Eendendichtheid i.p.v. grondeleendendichtheid). De verwachting was dat daarmee de voorspelling beter zou kunnen zijn. Daarnaast werden er nieuwe potentiële risicofactoren toegevoegd.

Daarmee is de doelstelling van het onderzoeksproject als volgt:

Het ontwikkelen van een voorspellend model voor de identificatie van gebieden met een hoog HPAI-besmettingsrisico op commerciële pluimveebedrijven in Nederland, om daarmee risicokaarten kunnen genereren



2 Werkwijze

2.1 Data

2.1.1 Landschapsvariabelen en vogelsoortendichtheid

Er zijn verschillende landschaps- en wilde vogelvariabelen meegenomen in de analyse (Tabel 1). De keuze van de verschillende landschapsvariabelen is gedaan op basis van de eerdere risicoanalyse rapportage (Gonzales et al., 2022), aangevuld met enkele nieuwe variabelen. De bronnen van de gebruikte landschapsvariabelen waren:

- Basisregistratie Percelen (BRP) <https://data.overheid.nl/dataset/10674-basisregistratie-gewaspercelen-brp->
- Basisregistratie Grootchalige Topografie (BGT): <https://www.geobasisregistraties.nl/basisregistraties/grootchalige-topografie>
- Bestand Bodemgebruik CBS: <https://data.overheid.nl/dataset/25362-beschikbaarheid-zoet-grondwater-verzilting>
- Natuurnetwerk Nederland (NNN): <https://www.rijksoverheid.nl/onderwerpen/natuur-en-biodiversiteit/natuurnetwerk-nederland>
- Data verzilting in Nederland: <https://data.overheid.nl/dataset/25362-beschikbaarheid-zoet-grondwater-verzilting>
- Kaart Hoogveen: Wösten, H., Brouwer, F., & Veraart, J. (2020). Kanskaart voor bescherming van koolstofvoorraad en CO₂-emissiereductie in natte natuur en multifunctionele klimaatbuffers: technische achtergrondmemo.

Gegevens over wildevogeldichtheid zijn afkomstig van Sovon. De wilde vogel data (Sovon 2018) is gebaseerd op tellingen van wilde vogels uitgevoerd in de winterperiodes van 2012 t/m 2015. De originele data had een resolutie van 5 x 5 km en is voor dit project geïnterpoleerd naar een 100 m x 100 m resolutie. De wilde vogelsoorten zijn geselecteerd op hun gevoeligheid voor HPAIv (APHA et al. 2017; Schreuder et al., 2022). De lijst met gebruikte soorten is beschikbaar in Bijlage 4, Tabel A2.

De landschapsvariabelen en vogeldata hebben een resolutie van 100 x 100 m (1 ha). Dit betekent dat voor de analyse Nederland werd verdeeld in rasters met een celgrootte van 100 X 100 m voor elk van de bovengenoemde databronnen. Voor elke raster werden gegevens voor de landschapsvariabelen en wildevogeldichtheid verzameld. Voor een gedetailleerde uitleg van de gegevensvoorbereiding, zie bijlage 1. Voor de analyse werd in totaal informatie voor 96 variabelen verzameld.

2.1.2 HPAI in pluimvee

Locatiegegevens van pluimveebedrijven en hun geschiedenis van besmetting met HPAI vanaf januari 2014 tot oktober 2022 zijn gebruikt. Alleen primaire besmettingen werden meegenomen, alle HPAI-besmettingen die mogelijk het gevolg waren van tussenbedrijfstransmissie, bevestigd door genetisch onderzoek van WBVR of epidemiologisch onderzoek van de NVWA en/of WBVR, werden uitgesloten. Tussen januari 2014 en oktober 2022 zijn in totaal 96 primaire besmettingen (HPAI-besmette pluimveebedrijven) geweest op ongeveer 2115 actieve pluimveebedrijven in Nederland.

De geografische ligging van elk bedrijf werd gebruikt om het bedrijf te koppelen aan een cel in de bovengenoemde hectarerrasters (zie 2.1.1). Dit raster was klein genoeg om de lokale omgeving rond een pluimveebedrijf te beschrijven. Verder werd een cel in het raster geclassificeerd als positief als een HPAI-besmet pluimveebedrijf aanwezig was gedurende de studieperiode (2014-2022). Omgekeerd werd een cel als negatief aangemerkt bij afwezigheid van besmettingen.

Tabel 1 Overzicht van landschapsvariabelen vogeldichtheden gebruikt in de analyse. Alle variabelen hebben een ruimtelijke resolutie van een hectare (ha, 100 x 100 m)¹.

Variabele	Bron	Beschrijving
Water (in %)	BGT	Aandeel water in % binnen 1 ha pixel
Wateroppervlak in radius 500m	BGT	Som van 10 m ² water pixels binnen 500m van randen van 1ha pixel
Water Reilly Index ²	BGT	Som van 10 m ² water pixels gedeeld door afstand ² in radius 500m van randen 1 ha pixel
Grasland (in %)	BGT	Aandeel van landbouw grasland in % binnen 1 ha (bron BGT)
Grasland (in %)	BRP	Aandeel van landbouw grasland in % binnen 1 ha (bron BRP)
Grasland in radius 1000m (ha)	BRP&BGT	Som 10 m ² grasland pixels binnen 1000 m van randen van 1 ha pixel
Bos en bomen (in %)	BBG	Aandeel van bos en bomen in % binnen 1 ha
Bos en bomen (ha) binnen 500m	BBG	Som 10 m ² bos en bomen pixels binnen 500 m van randen van 1 ha pixel
Bos en bomen afstand	BBG	Afstand centrum van 1 ha pixel tot dichtstbijzijnde 10 m ² bos en bomen
Bos en bomen Reilly index ²	BBG	Som bos/boom opp. gedeeld door afstand ² in radius 500 m van randen 1 ha pixel
NNN (in %)	RVO	Aandeel 'Natuur Netwerk Nederland' (NNN) in % binnen 1 ha
NNN oppervlak binnen 500m (ha)	RVO	Som van 10 m ² NNN pixels binnen 500 m van randen 1 ha pixel
NNN Reilly index	RVO	Som van 10 m ² NNN pixels gedeeld door afstand ² in radius 500 meter van randen 1 ha pixel
NNN afstand	RVO	Afstand centrum dichtstbijzijnde 10 m ² NNN pixel
Fruitbomen in radius 1000m (ha)	BRP	Som opp. fruitbomen in ha binnen 1000 m van randen 1 ha pixel
Bomenteelt in radius 1000m (ha)	BRP	Som opp. bomenteelt in ha binnen 1000 m van randen 1 ha pixel
Suikerbieten binnen 1000m (ha)	BRP	Som opp. suikerbieten in ha binnen 1000 m van randen 1 ha pixel
Granen binnen 1000m (ha)	BRP	Som opp. granen in ha binnen 1000 m van randen 1 ha pixel
Mais binnen 1000m (ha)	BRP	Som opp. Mais in ha binnen 1000 m van randen 1 ha pixel
Gewassen rest binnen 1000m (ha)	BRP	Som opp. gewassen rest in ha binnen 1000 m van randen 1 ha pixel
Zoutige grond (in %)	BBG&RIVM	Aandeel zoutige grond (in %) binnen 1 ha uit BBG en RIVM database
Zoutige kwelder (in %)	BGT	Aandeel zoutige kwelder (in %) binnen 1 ha pixel
Duinen	BGT	Aandeel duinen in % binnen 1 ha pixel
Moeras (in %)	BGT	Aandeel moeras in % binnen 1 ha pixel
Riet (in %)	BGT	Aandeel riet in % binnen 1 ha pixel
Hoogveen (in %)	BGT	Aandeel hoogveen in % binnen 1 ha pixel
Natte natuur (in %)	BBG	Aandeel natte natuur in % binnen 1 ha pixel
Bevolkingsdichtheid	CBS	Aantal mensen per ha
Bevolking binnen 1000m	CBS	Aantal mensen binnen straal van 1000 m van randen 1 ha pixel
Windmolens binnen 500m	BGT	Aantal windmolens in een straal van 500 m van randen 1 ha pixel
Duikende	Sovon	Aantal duikende gedisaggregeerd van 5 x 5 km naar 1 ha
ganzen	Sovon	Aantal ganzen gedisaggregeerd van 5 x 5 km naar 1 ha
Grondelende	Sovon	Aantal grondelende gedisaggregeerd van 5 x 5 km naar 1 ha
Zwanen	Sovon	Aantal zwanen gedisaggregeerd van 5 x 5 km naar 1 ha
Steltlopers	Sovon	Aantal steltlopers gedisaggregeerd van 5 x 5 km naar 1 ha
Meeuwen	Sovon	Aantal meeuwen gedisaggregeerd van 5 x 5 km naar 1 ha
Roofvogels en uilen	Sovon	Aantal roofvogels en uilen gedisaggregeerd van 5 x 5 km naar 1 ha
Overige watervogels	Sovon	Aantal overige watervogels gedisaggregeerd van 5 x 5 km naar 1 ha
Kraaien	Sovon	Aantal kraaien gedisaggregeerd van 5 x 5 km naar 1 ha

¹ Een gedetailleerde beschrijving van de landschapsvariabelen kan worden gevonden in de bron (zie sectie 2.1.1). Een beschrijving van hoe de gegevens van iedere variabele werd gebruikt is te vinden in annex A1.

² Reilly index: binnen een gegeven afstandsradius, combineert de Reilly index informatie van een specifieke variabele in het gebied (b.v. oppervlakte aan water of oppervlakte aan bomen) en de afstand vanaf een referentiepunt tot het centroid van de rastercel van 1 ha.

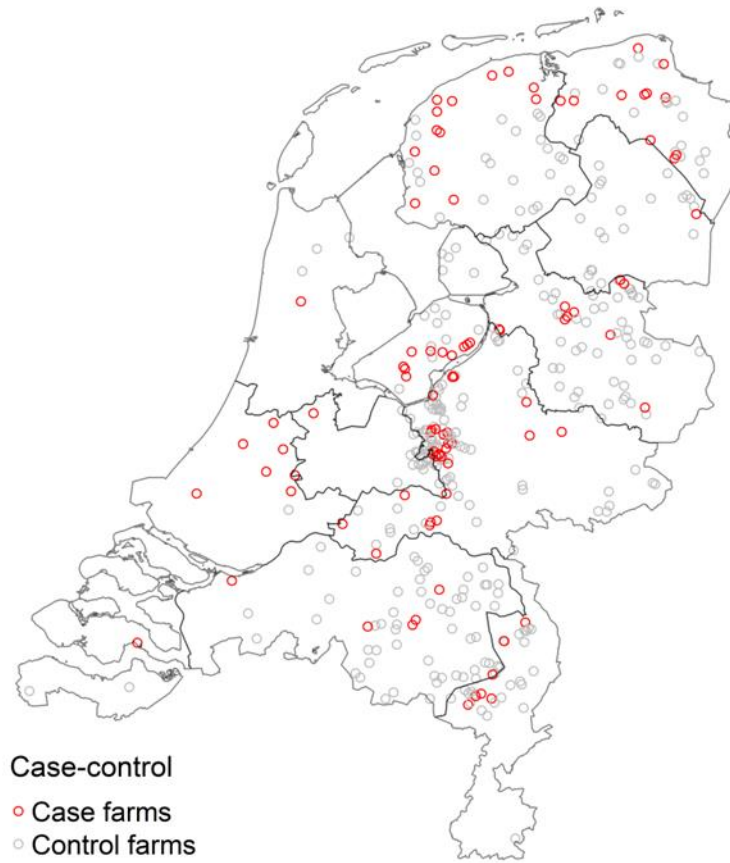
2.2 Identificatie van HPAI- risicogebieden

2.2.1 Voorspelling (kunstmatige intelligentie) modellen

Voor deze analyse is het optreden van HPAI-besmettingen bij pluimvee in een gebied gebruikt als proxy voor de aanwezigheid van HPAI-besmettingsdruk in dat specifieke gebied. En dit kan op zijn beurt geassocieerd zijn met de mate van HPAI-besmetting van wilde vogels. Het gebruik van alle Nederlandse pluimveebedrijven (besmet en niet besmet) om de voorspellingsmodellen te ontwikkelen zal resulteren in onjuiste modellen vanwege de lage mate van voorkomen van HPAI-besmettingen bij pluimveebedrijven: slechts vier percent (96/2115) van de data is van besmette bedrijven. Het gebruik van een laag aantal besmette pluimveebedrijven (en daarmee een hoog aantal niet-besmette pluimveebedrijven) voor een dergelijke analyse staat bekend "Class imbalance". Om op een goede manier om te gaan met deze onbalans is er gekozen voor een "case-control" benadering zoals eerder gedaan door Schreuder et al. (2022).

2.2.2 Selectie van Cases en Controls

Alle primair HPAI-besmette pluimveebedrijven in de periode 2014 – 2022 werden geselecteerd als cases (n = 96). Niet-HPAI-besmette pluimveebedrijven werden gekozen op een afstand van minimaal 5 km en maximaal 100 km afstand van een HPAI-besmet pluimveebedrijf en gepaard aan een besmet pluimveebedrijf op basis van het pluimveebedrijfstype en jaar. Het optimale aantal controlebedrijven, ten opzichte van de cases, is bepaald zodat verschillen in blootstelling tussen de controlebedrijven en uitbraakbedrijven een minimaal onderscheidingsvermogen (power) hadden van 80% en een betrouwbaarheid van 95%. Er is aangenomen dat de kans op blootstelling aan een modelvariabele 0.5 was (Appendix A2). De verhouding van case tot controlebedrijven werd gevarieerd van 1 : 2 tot 1 : 6, waardoor het onderscheidende vermogen 70,3% (twee controles per case), 76,% (drie controles per case), 80,3% (vier controles per case), 82,4% (vijf controles per case) of 83,8% (zes controles per case) was. Het gebruik van 4 controles per case garandeert een onderscheidend vermogen van 80%. Het selecteren van meer dan 4 controles per case zou, ondanks een hoger onderscheidend vermogen, de onbalans vergroten van de data. Daarom werden er vier controlebedrijven geselecteerd per besmet pluimveebedrijf (Figuur 1). Elk case- en controlebedrijf is gekoppeld aan de hectarerasters met gegevens van de landschapskarakteristieken en wilde vogeldichtheid (zie 2.1.1).



Figuur 1 Geografische spreiding van case en controle pluimveebedrijven

2.2.3 Selectie en ontwikkeling van voorspellingsmodel

Het doel van de modellen was het risico op HPAI-besmettingen op Nederlandse commerciële pluimveebedrijven te voorspellen op basis van landschapsvariabelen en de dichtheid van verschillende soorten wilde vogels. Alle variabelen werden beoordeeld aan de hand van de verschillen tussen cases en controles. Verder werd de Spearman-correlatie tussen de variabelen onderzocht en de associatie tussen elke variabele en besmettingsrisico geëvalueerd door univariabele, conditionele logistische regressies. Voor het ontwikkelen van de modellen werd de dataset in tweeën gesplitst:

- **Training** (75%) dataset. – 305 pluimveebedrijven, 72 cases en 233 controles.
- **Test** (25%) dataset. – 107 pluimveebedrijven, 24 cases en 83 controles.

De trainingsdataset is gebruikt om de modellen te ontwikkelen, terwijl de testdataset is gebruikt om de voorspellende waarde van de modellen te toetsen. De onderstaande voorspellende modellen werden getest:

- Logistische Regressie,
- Partial Least Squares Discriminant Analysis,
- Naive Bayes,
- Neural Networks en
- Random Forest.

Deze modellen werden allen ontwikkeld gebruikmakend van de landschapskarakteristieken en de gegroepede wilde vogelsoorten data (Tabel 1). Het vermogen van deze modellen om een juiste classificatie te geven werd beoordeeld door het bepalen van de "Area Under the Curve" (AUC). Op basis van deze parameter werd het beste model geselecteerd. Het model met de hoogste AUC was het "Random Forest" model, wat vervolgens werd geselecteerd voor verdere ontwikkeling. We zullen dus verder in dit rapport alleen de methoden en resultaten van het Random Forest-model beschrijven.

2.2.3.1 Random Forest model

Verschillende Random Forest modellen werden ontwikkeld met behulp van de training dataset. Zeven verschillende modellen werden getraind om het effect van verschillende variabelen in het model te onderzoeken (Tabel 2). De getrainde modellen werden geselecteerd op basis van de AUC. Daarna werden de modellen gebruikt om de besmette en niet-besmette rastercellen te voorspellen (classificeren). Deze voorspellingen werden vervolgens vergeleken met de werkelijke status van de rastercel om voorspellende prestatieparameters zoals de AUC, de gevoeligheid (Se) en specificiteit (Sp) te bepalen.

Tabel 2 Beschrijving van de Random Forest machine learning-modellen die zijn getraind om HPAI-besmetting van Nederlandse commerciële pluimveebedrijven te voorspellen.

Model	Variabelen in het model opgenomen
1	Alleen landschapsvariabelen
2	Alle landschapsvariabelen en de gegroepede wilde vogelsoorten ¹
3	Alle landschapsvariabelen en de gegroepede wilde vogelsoorten, alleen individuele vogelsoorten binnen de "Gronddeleenden" groep
4	Alle landschapsvariabelen en de gegroepede wilde vogelsoorten, alleen individuele vogelsoorten binnen de "ganzen" groep
5	Alle landschapsvariabelen en de gegroepede wilde vogelsoorten, alleen individuele vogelsoorten binnen de "Zwanen" groep
6	Alle landschapsvariabelen en alle individuele wilde vogelsoorten
7	Alleen wilde vogelsoorten met een statistisch significante ($p < 0.05$) associatie met HPAI-besmetting in pluimvee

¹ Zie Bijlage A4 voor een complete lijst van vogelsoorten behorend bij elke wilde vogelsoortengroep.

De modellen zijn getraind door gebruik te maken van een vijfvoudige kruisvalidatie aanpak, waarin vijf maal een verschillende test- en trainingsdataset gemaakt is om zo de invloed van de dataselectie te kunnen bepalen. De hyperparameters van het model zijn geoptimaliseerd door gebruik te maken van een zoekrooster. Vervolgens is het model getraind met de optimale hyperparameters. Het belang van de modelvariabelen is bepaald door gebruik te maken van de "permutation variable importance" methode op alle "trees" in de random forest, genormaliseerd voor de standaardfout.

De voorspellende waarde van de modellen is bepaald door de "area under the curve" (AUC) te bepalen van de Receiver Operating Characteristic (ROC) curve, waarin de ROC een drempelwaarde-onafhankelijke maat is voor het compromis tussen gevoeligheid en specificiteit.

Alle analyses werden uitgevoerd in R (Core team, 2022) gebruikmakend van de softwaremodules "caret", "mlbench", en "ranger". De AUC werd beoordeeld m.b.v. de softwaremodule "ROCR" in R. Het best voorspellende model wordt gebruikt voor het besmettingsrisico voor het hele land op hectare niveau.

2.2.3.2 Risicobeoordelingsmodel

Voor dit model is het HPAI-besmettingsrisico (Rha_i) voor elke rastercel i (onafhankelijk van de aan- of afwezigheid van een pluimveebedrijf) berekend als:

$$Rha_i = \frac{\sum R_i W_i}{\sum W_i}$$

Waarbij R_i de waarde is van de risicofactor i , geschaald tussen 0 en 1, en W_i de wegingsfactor is van risicofactor i .

Voor de selectie van de belangrijkste risicofactoren R_i en het schatten van de wegingsfactoren W_i werden univariabele logistische regressiemodellen gefit op data van de hectarerastercellen die pluimveebedrijven bevatten, zowel besmet als niet-besmet ($n = 2115$). De geschatte ruwe odds ratios (OR) voor de significante risicofactoren werden gebruikt als wegingsfactoren, met

$$W_i = OR_i \text{ if } OR \geq 1$$

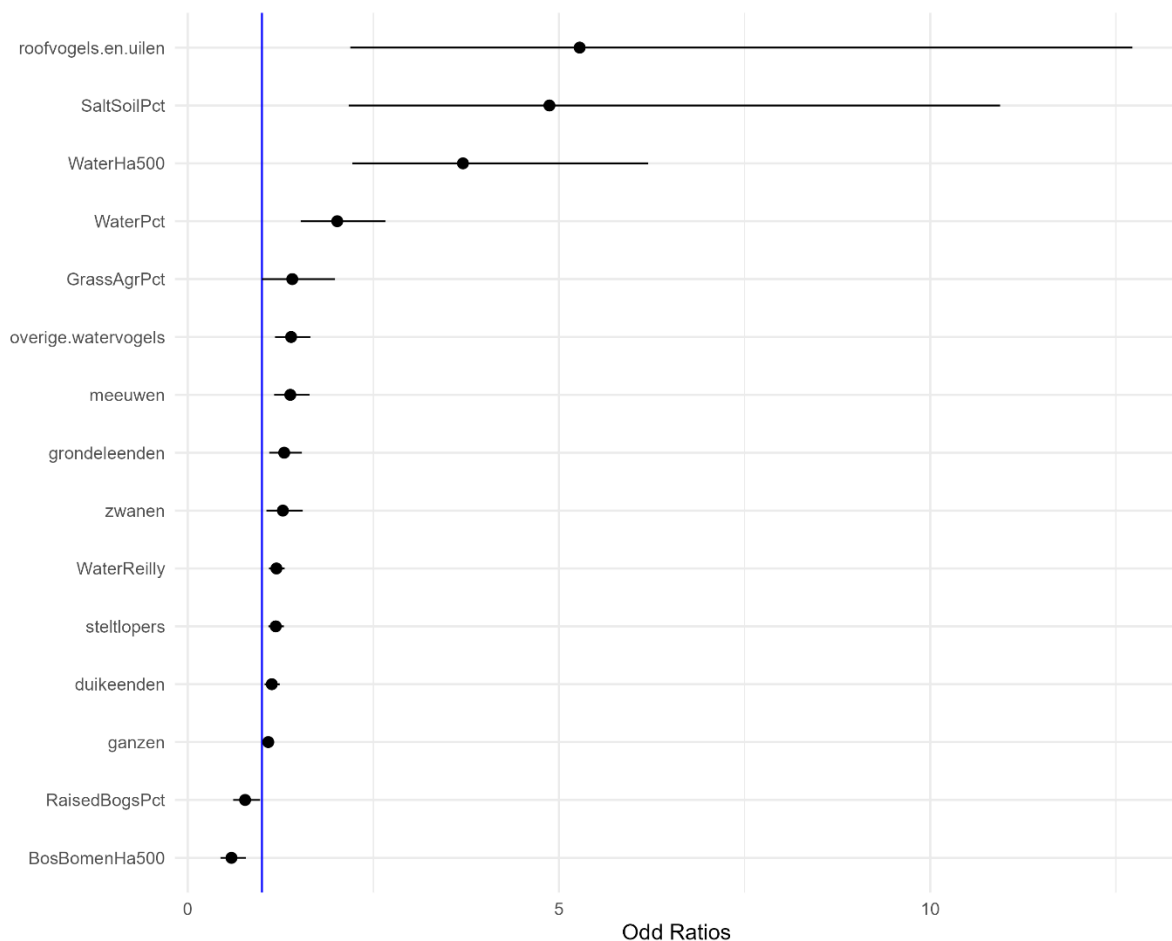
$$W_i = 1/OR_i \text{ if } OR < 1$$

3 Resultaten

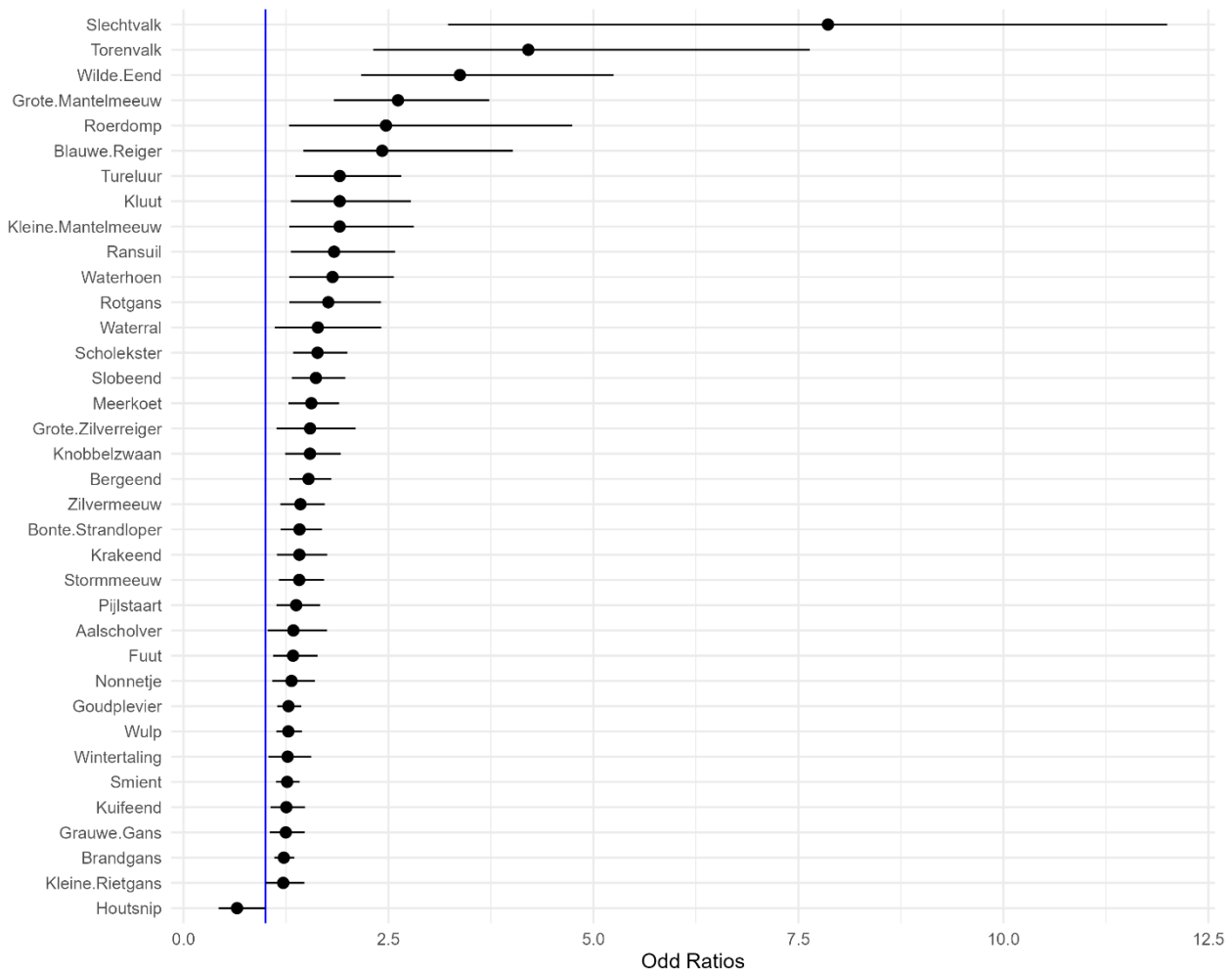
3.1 Random Forest model

3.1.1 Verkenning van de variabelen

De verdeling van de waarden van landschapsvariabelen en populatiedichtheid van de wilde vogels voor de cases (HPAI-besmette pluimveebedrijven) en controles (niet-HPAI-besmette pluimveebedrijven) is weergegeven in Bijlage 2 Tabel A1. De verdeling van de waarden van de verschillende wilde vogelsoorten is te zien in Figuur A1. De resultaten van de univariate analyses (logistische regressie) zijn weergegeven in Figuren 2 en 3. Alleen de statistisch significante ($p < 0,05$) variabelen zijn gepresenteerd voor landschapsvariabelen, gegroepeerde wilde vogelsoorten dichtheden (Figuur 2) en voor individuele wilde vogelsoort dichtheden (Figuur 3). Variabelen die positief geassocieerd zijn met het risico op besmetting hebben een odds ratio (OR) > 1 en voor negatief geassocieerde variabelen is de OR < 1 .



Figuur 2 Gemiddelde Odds ratios en 95% betrouwbaarheidsintervallen uit de univariate analyses van landschapskarakteristieken en dichtheden van gegroepeerde vogelsoorten. Alleen variabelen met een significante ($p < 0,05$) risico op HPAI-besmetting van pluimveebedrijven zijn gevisualiseerd. De verticale blauwe lijn representeert de OR waarde = 1.



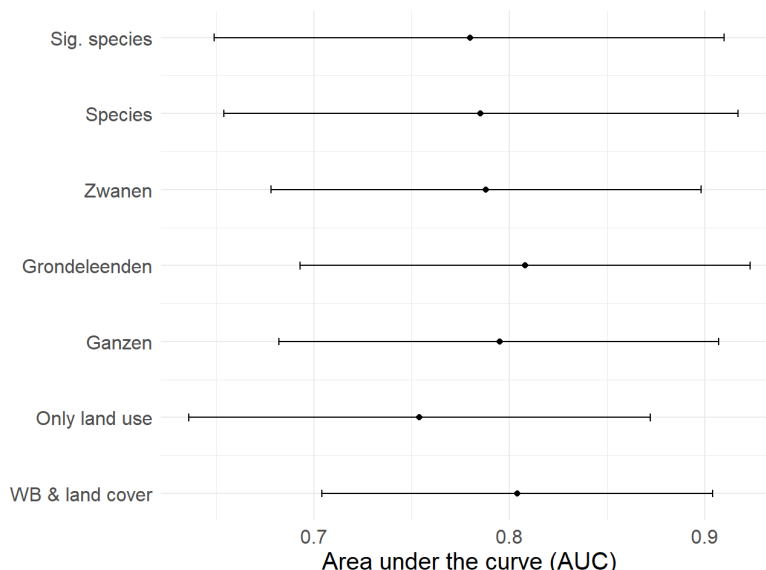
Figuur 3 Gemiddelde Odds ratios en 95% betrouwbaarheidsintervallen uit de univariate analyses van individuele vogelsoorten. Alleen variabelen met een significante ($p < 0,05$) risico op HPAI-besmetting van pluimveebedrijven zijn gevisualiseerd. De verticale blauwe lijn representeert de OR waarde = 1.

3.1.2 Model validatie en selectie

Alle modellen (zie Tabel 1 voor beschrijving van de modellen) hadden vergelijkbare voorspelende prestaties (Figuur 4). De gemiddelde AUC van de zes verschillende modellen varieerde van 0,754 (Model1: Alle landschapsvariabelen) tot 0,808 (Model 3: Alle landschapsvariabelen en de gegroepeerde wilde vogelsoorten, alleen individuele vogelsoorten binnen de "Grondelenden" groep). In Tabel 3 is de gevoeligheid (Se) en specificiteit (Sp) van alle modellen samengevat.

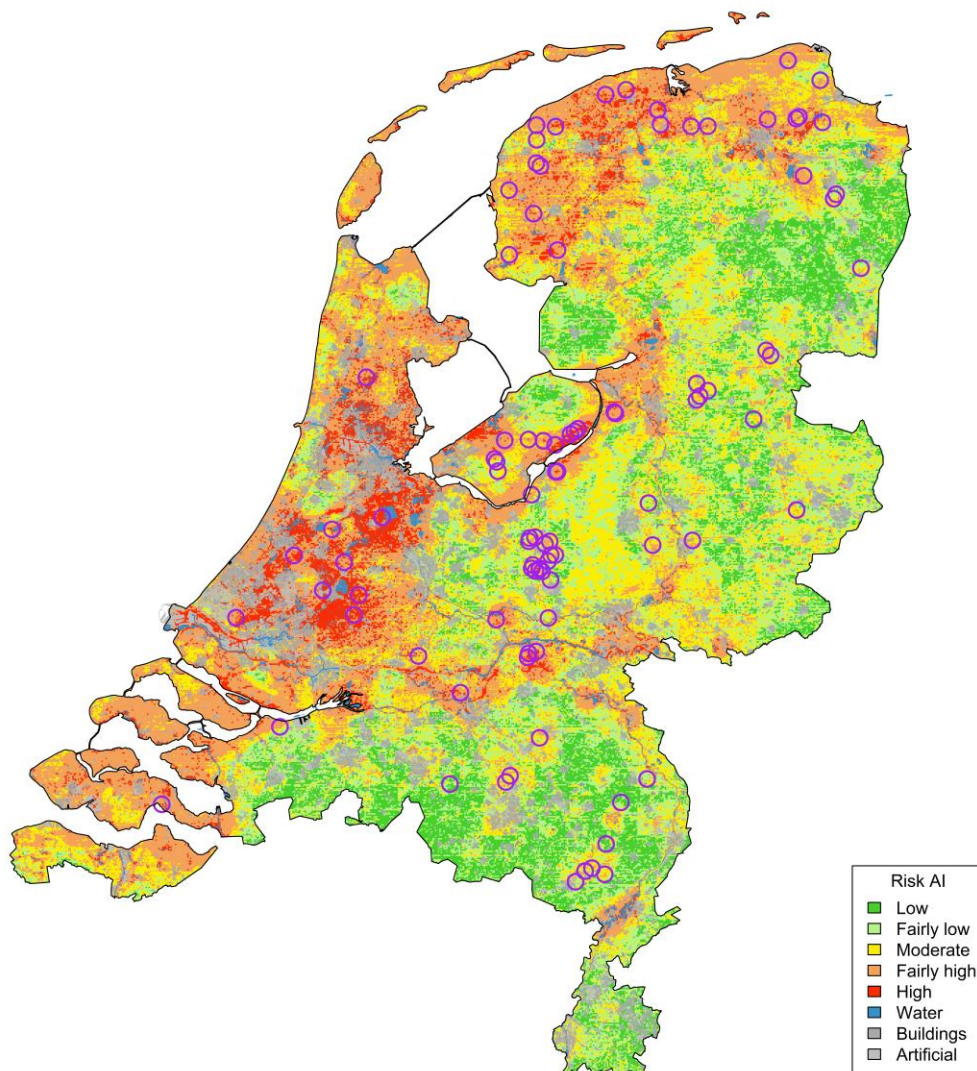
Tabel 3 Voorspellende prestaties van zeven geselecteerde Random Forest modellen

Parameter	Landschapsvariabelen & wilde vogelsoorten gegroepeerd	Alleen landschapsvariabelen	Landschapsvariabelen & wilde vogelsoortgroepen & ganzen individueel	Landschapsvariabelen & wilde vogelsoortgroepen & grondeleenden individueel	Landschapsvariabelen & wilde vogelsoortgroepen & zwanen individueel	Landschapsvariabelen & alle vogelsoorten individueel	Alleen wilde vogelsoorten met significant associatie met HPAI besmetting in pluimvee
AUC	0.804	0.754	0.795	0.808	0.788	0.785	0.780
Gevoeligheid	0.826	0.783	0.783	0.783	0.783	0.783	0.783
Specificiteit	0.738	0.613	0.713	0.775	0.713	0.738	0.750
cut off	0.22	0.218	0.221	0.247	0.207	0.264	0.262
Aantal variabelen	33	24	39	40	35	82	36



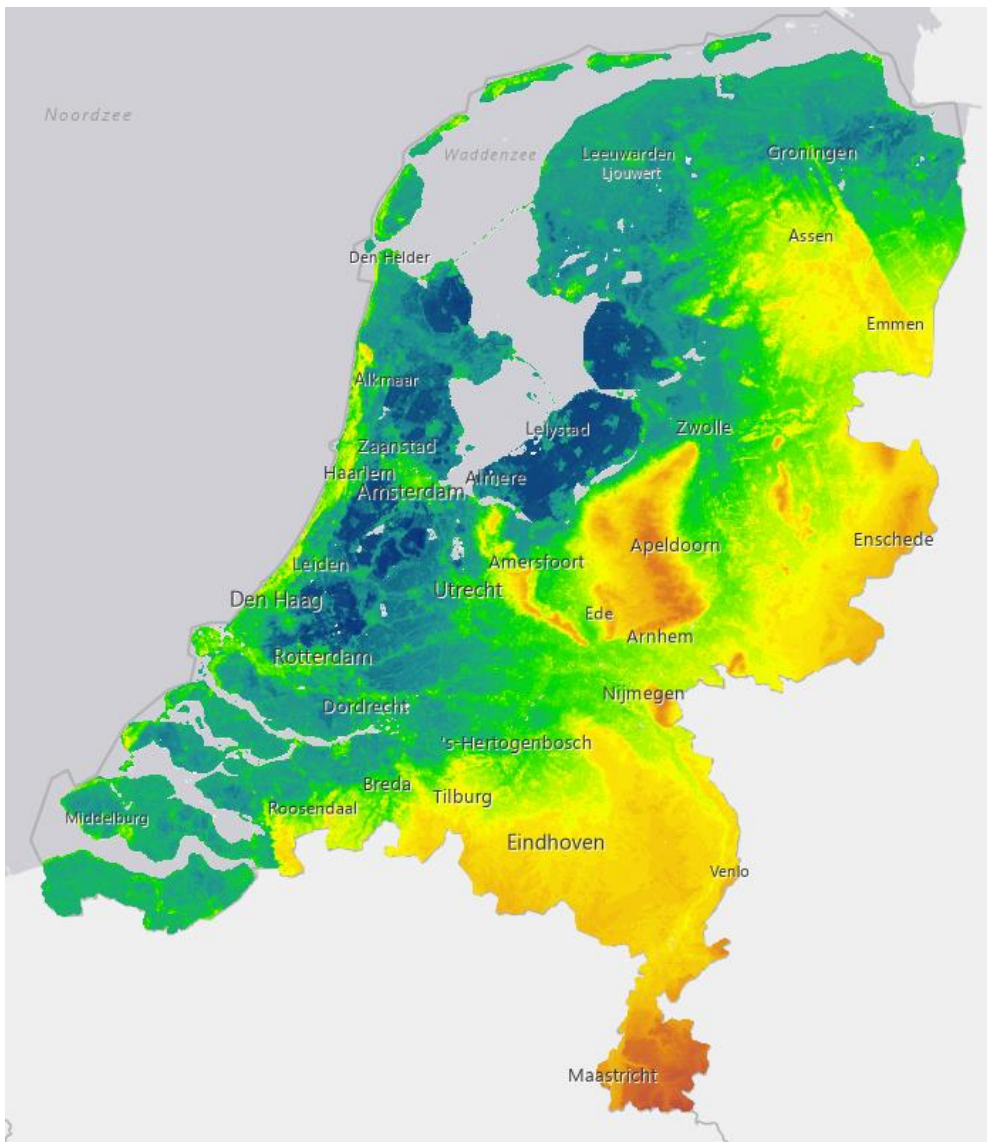
Figuur 4 Verdeling van AUC voor de verschillende voorspelmodellen validatie. "Sig. species" refereert naar het model met alleen vogelsoorten met een statistisch significante associatie met HPAI-besmetting van pluimveebedrijven (Figuur 3). "Species" refereert naar het model met de individuele vogelsoorten en "WB & land cover" verwijst naar het model met landschapsdata en gegroepeerde vogelsoortdata. De overige regels zijn de modellen waarin een enkele vogelsoortgroep is opgesplitst in de individuele vogelsoorten.

Van alle modellen was het model gebaseerd op landschapsvariabelen & gegroepeerde wilde vogelsoorten geselecteerd om de risicogebieden voor besmetting met HPAI in Nederland te identificeren. Dit model was verder gevalideerd met een subset van de test dataset met alleen gegevens van de jaren 2020 – 2022. De gevoeligheid en specificiteit van het model was 0,95. Daarbij kunnen wij verwachten dat de voorspellingsprestaties voor de aankomende jaren goed (betrouwbaar) zullen zijn. Naast de verschillende wilde vogel variabelen, waren de belangrijk variabelen voor alle modellen de variabelen die de aanwezigheid van water kwantificeren (Reilly index, Wateroppervlak in radius 500m, % Water). In bijlage A3 is het belang van de 15 belangrijkste variabelen gegeven, voor alle zeven voorspelmodellen (Figuur A2). De geografische verdeling van het risico in Nederland is gepresenteerd in Figuur 5 en Figuur A3 van bijlage A3.



Figuur 5 HPAI-besmettingsrisico op pluimveebedrijven in Nederland volgens het best presterende Random Forest model (paarse cirkels representeren HPAI-besmette pluimveebedrijven). Om de interpretatie te vereenvoudigen, werden de gekwantificeerde risicoscores gecategoriseerd in categorieën variërend van laag tot hoog. HPAI-besmettingsrisico is uitgedrukt in de volgende risico-categorieën: low (laag) = $[0.0, 0.12]$, Fairly Low (redelijk laag) = $(0.12, 0.22]$, Moderate (gemiddeld) = $(0.22, 0.35]$, Fairly high (redelijk hoog) = $(0.35, 0.6)$ en High (hoog): > 0.6 . Een ROC analyse identificeerde een risico-waarde van 0.22 als drempelwaarde voor scheiding tussen besmette en niet-besmette grids. Daarom zijn grids met een waarde ≤ 0.22 geclassificeerd als laag risico.

Om de oriëntatieduiding van de gebieden met hoog en laag HPAI-besmettingsrisico op pluimveebedrijven voor de lezer makkelijker te maken voegen wij hierbij een kaart van Nederland toe op basis van hoogte t.o.v. de zeespiegel.



Figuur 6 Hoogtekaart van Nederland (bron: Actueel Hoogtebestand Nederland, <https://www.ahn.nl>). Het kleurverloop gaat van donkerblauw (< -6 m t.o.v. zeespiegel) naar bruin (> 250 m t.o.v. de zeespiegel).

De resultaten van Figuur 5 laten zien dat er een hoog risico op HPAI-besmetting bij pluimveebedrijven is te zien in de kustregio's en laaggelegen gebieden in Groningen, Friesland, Flevoland, Noord- en Zuid-Holland en Zeeland. Verder is er een verhoogd risico op HPAI-besmetting bij pluimveebedrijven te zien in de stroomgebieden van de grote rivieren in het midden van het land: de IJssel, Neder-Rijn, Waal en Maas. Verder is er een hoog risico op HPAI-besmetting bij pluimveebedrijven te zien in Limburg ter hoogte van de Maasplassen bij Roermond. Er is een laag risico op HPAI-besmetting bij pluimveebedrijven te zien voor de hooggelegen delen van Groningen, Drenthe, Overijssel, Gelderland, Utrecht, Noord-Brabant en Limburg.

3.2 Risicobeoordelingsmodel

3.2.1 Identificatie van risicofactoren en Odds Ratio-schatting

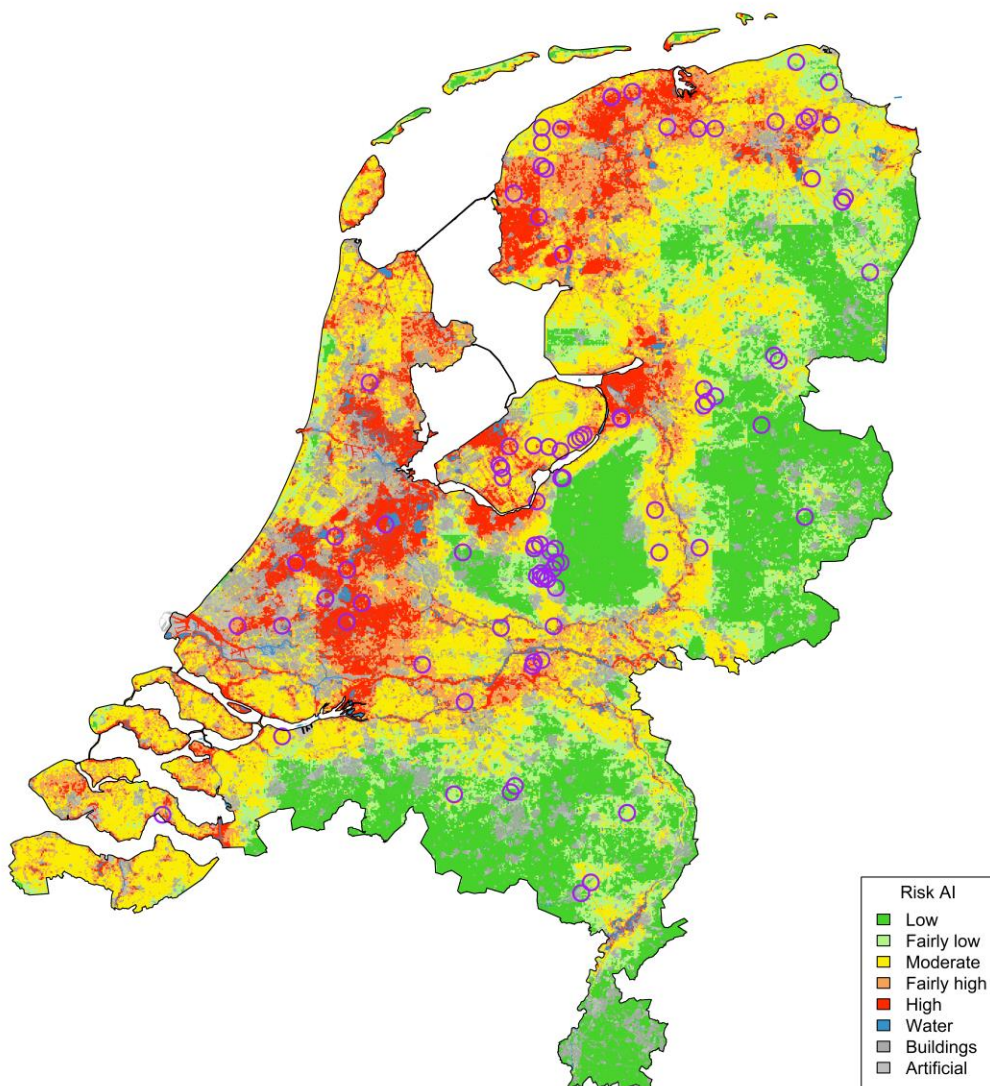
Als eerste werd een univariate logistische regressieanalyse uitgevoerd met de gegevens van alle 2115 pluimveebedrijven om significante risicofactoren voor een besmetting met HPAI-infectie te identificeren en de bijbehorende OR te kwantificeren. De significante risicofactoren zijn samengevat in Tabel 4.

Tabel 4 Significante risicofactoren (landschapskarakteristieken en wilde vogel dichtheden) voor besmetting met HPAI virus op commerciële pluimveebedrijven in Nederland (2014-2022), met OR=Odds Ratio, en 95% betrouwbaarheidsinterval (LCL – UCL).

Risicofactor	OR	LCL	UCL	p waarde	Transformatie ^a
Roofvogels en uilen	3.30	1.71	6.32	0.00	Ln
Zoutige grond %	2.91	1.84	4.57	0.00	Ln
Water-Ha500	2.72	1.95	3.77	0.00	Ln
Water %	1.83	1.45	2.31	0.00	Ln
Zwanen	1.54	1.30	1.83	0.00	Ln
Overige watervogels	1.49	1.28	1.78	0.00	Ln
Riet %	1.46	0.98	1.99	0.03	
Meeuwen	1.40	1.22	1.66	0.00	Ln
Grondeleenden	1.40	1.20	1.64	0.00	Ln
Duikeenden	1.22	1.12	1.32	0.00	Ln
Steltlopers	1.18	1.10	1.26	0.00	Ln
Ganzen	1.06	1.01	1.12	0.03	Ln
Gras landbouw %	1.02	1.01	1.03	0.00	
Water-Reilly	1.00	1.00	1.00	0.00	
Bevolkingsdichtheid per km ²	0.78	0.67	0.93	0.00	Ln
Hoogveen %	0.75	0.57	0.93	0.02	Ln
BosBomen-Ha500	0.65	0.50	0.83	0.00	Ln
Kraaien	0.42	0.23	0.75	0.00	Ln

a De kolom Transformatie laat zien welke variabelen werden Log-getransformeerd (Ln) om te passen in het univariate logistische regressiemodel en hoe deze gegevens worden gebruikt in het risicobeoordelingsmodel.

Vervolgens, om het gebiedsrisico voor HPAI in elk raster van 100 X 100 m in Nederland te kwantificeren (risicoscores), hebben we de gegevens gebruikt voor de significante variabelen genoemd in Tabel 4 en de OR-waarde gebruikt als hun gewichten (zie 2.3). In Figuur 6 is de ruimtelijke verdeling van dit risico in Nederland weergegeven. Om de interpretatie te vereenvoudigen, werden de gekwantificeerde risicoscores (Figuur A4, bijlage 3) gecategoriseerd in categorieën variërend van laag tot hoog. Via een ROC (receiver operating characteristic)-analyse werd de risicoscore 0,49 geïdentificeerd als de waarde die besmette rastercellen scheidt van niet-besmette rastercellen. Rastercellen met een risicoscore lager dan 0,49 werden daarom gecategoriseerd als laagrisico-cellen. De AUC waarde van het gehele model was 0,73 (95% betrouwbaarheidsinterval (btbh): 0,69 – 0,78).



Figuur 7 HPAI-besmettingsrisico op pluimveebedrijven in Nederland volgens het risicobeoordelingsmodel (paarse cirkels representeren HPAI-besmette pluimveebedrijven). HPAI-besmettingsrisico is uitgedrukt in de volgende risico-categorieën: low (laag) = $[0.069, 0.358]$, Fairly low (redelijk laag) = $(0.358, 0.49]$, Moderate (gemiddeld) = $(0.49, 0.574]$, Fairly high (redelijk hoog) = $(0.574, 0.646]$ en High (hoog): > 0.646 . ROC analyse identificeerde een risico-waarde van 0.49 als drempelwaarde voor scheiding tussen besmette en niet-besmette rastercellen. Daarom zijn rastercellen met een waarde ≤ 0.49 geclassificeerd als laag risico.

De resultaten van Figuur 7 laten zien dat er een hoog risico op HPAI-besmetting bij pluimveebedrijven is te zien in de kustregio's en laaggelegen gebieden in Groningen, Friesland, Flevoland, Noord- en Zuid-Holland en Zeeland. Verder is er een verhoogd risico op HPAI-besmetting bij pluimveebedrijven te zien in de stroomgebieden van de grote rivieren in het midden van het land: de IJssel, Neder-Rijn, Waal en Maas. Verder is er een hoog risico op HPAI-besmetting bij pluimveebedrijven te zien in Limburg ter hoogte van de Maasplassen bij Roermond. Er is een laag risico op HPAI-besmetting bij pluimveebedrijven te zien voor de hooggelegen delen van Groningen, Drenthe, Overijssel, Gelderland, Utrecht, Noord-Brabant en Limburg.



4 Discussiepunten

In dit onderzoek zijn modellen gemaakt die een voorspelling geven waar in Nederland het grootste risico – en daarmee ook het minste risico – is voor potentiële HPAI-besmetting op commerciële pluimveebedrijven. Wij hebben twee verschillende soorten modellen ontwikkeld en ondanks de verschillen in methodiek laten ze toch een overeenkomstig beeld zien in de gebiedsverdeling van de risico's. Een hoog risico op HPAIv besmetting op pluimveebedrijven is te zien in de kustregio's en laaggelegen gebieden in Groningen, Friesland, Flevoland, Noord- en Zuid-Holland en Zeeland. Verder is er een verhoogd risico te zien in de stroomgebieden van de grote rivieren in het midden van het land: de IJssel, Neder-Rijn, Waal en Maas. Wat verder opvalt is een verhoogd in Limburg ter hoogte van de Maasplassen bij Roermond. Verder laten beide modellen een laag risico zien voor de hooggelegen delen van Groningen, Drenthe, Overijssel, Gelderland, Utrecht, Noord-Brabant en Limburg.

Verschillen in resultaten tussen de voorspelmodellen worden gevonden op de Waddeneilanden, de Veluwe, de Achterhoek, Twente en Zeeland. Het Random Forest model kent hogere risico's toe aan deze gebieden vergeleken met het risicobeoordelingsmodel.

Zeeland vormde in het voorgaande rapport (Gonzales et al. 2022) een aandachtsgebied omdat daar de aanwezigheid van water en watervogels hoog is, maar het aantal HPAI-besmettingen op pluimveebedrijven tegen de verwachting in laag is. In beide modellen is te zien dat het HPAI-besmettingsrisico in gebieden in Zeeland niet homogeen is. Beide modellen geven aan dat binnen Zeeland de risico's het laagst zijn in Zeeuws-Vlaanderen terwijl het oostelijk deel van Zuid-Beveland en Tholen hoge risico's hebben. In het Random Forest model worden er hogere risico's toegekend vergeleken met het risicobeoordelingsmodel.

Met het ontwikkelen van de modellen lag de nadruk op het optimaliseren van de voorspellende waarde, en niet op het verkrijgen van begrip over welke variabelen gekoppeld zijn aan toe- of afnames van risico's of de achterliggende mechanismen. Dit gezegd hebbende is er toch inzicht opgedaan over de correlatie van de variabelen met de risico's door de univariate analyses die gedaan zijn (Tabel 1) en het "belang" van de variabelen in de Random Forest modellen (Figuur A2).

Uit de wildevogeldata bleken hogere dichtheden van Wilde Eenden, Bergeenden, Slobeenden, Brandganzen en Knobbelzwanen in de omgeving van pluimveebedrijven statistisch significant gecorreleerd te zijn met het optreden van HPAIv besmettingen op pluimveebedrijven in de univariate analyses. Verder kwam naar voren dat hun voorspellende waarde in de Random Forest modellen groot was. Deze soorten zijn eerder ook geïdentificeerd als belangrijke soorten in een onderzoek gedaan op data van 2014 tot 2020 (Schreuder et al., 2022). Van de bovengenoemde vogelsoorten was de dichtheid van Wilde Eenden in de omgeving van pluimveebedrijven positief gecorreleerd met het risico op besmettingen met laag-pathogene aviaire influenza virus op pluimveebedrijven (Gonzales et al., ongepubliceerd). Daarnaast bezochten grote hoeveelheden Wilde Eenden de uitloop van een regelmatig LPAI-besmet pluimveebedrijf in de herfst-winterperiode (Elbers & Gonzales 2020; 2021). Aannemende dat Wilde Eenden een belangrijke rol spelen bij het HPAI-besmettingsrisico van pluimveebedrijven, zou actieve surveillance gericht op de aanwezigheid van Wilde Eenden bij de pluimvee-wilde vogel interface in HPAI-hoog risicogebieden als een vroegwaarschuwingssysteem kunnen optreden in de herfst-winterperiode wanneer grote hoeveelheden trekkende wilde watervogels arriveren in Nederland.

We zien ook een significante associatie tussen de dichtheid van roofvogels/uilen soorten en het HPAI-besmettingsrisico. Daarnaast was de dichtheid van roofvogels/uilen ook een belangrijke variabele in de voorspelmodellen. Roofvogels en uilen worden eigenlijk altijd in betrekkelijk lage dichtheden aangetroffen in het Nederlandse landschap. Het zijn geen groepsvogels zoals bijv. eenden en ganzen. De samenstelling van de groep roofvogels en uilen is gevarieerd. Er zitten soorten bij die gespecialiseerd zijn in het vangen van kleine

knaagdieren (zoals b.v. muizen, ratten) en er zijn ook soorten bij die andere prooidieren vangen (waaronder (kleine) vogels). De (iets) hogere dichtheden van roofvogels/uilen in de omgeving van HPAI-besmette pluimveebedrijven duiden waarschijnlijk op de aanwezigheid van meer prooidieren in de omgeving van de HPAI-besmette pluimveebedrijven in vergelijking met niet-besmette pluimveebedrijven.

Van de landgebruiksvariabelen waren de variabelen die te maken hadden met het oppervlakte aan water belangrijk in alle besmettingsrisicovoorspelmodellen, dat gold ook voor de oppervlakte aan agrarisch gras en variabelen die de teelt van verschillende gewassen beschrijven. De aanwezigheid van water en gras is een geschikte habitat voor wilde vogels zoals eenden, ganzen en zwanen en daarmee een belangrijke voorwaarde voor hun aanwezigheid in de directe omgeving van pluimveebedrijven. Deze observaties zijn in lijn met de bevindingen uit ons vorige onderzoek (Gonzales et al., 2022), waar dezelfde variabelen werden geïdentificeerd als significante risicofactoren voor de besmetting van pluimvee met HPAIV.

In de toekomst jaarlijks een update kunnen blijven doen van nauwkeurige besmettingsrisicovoorspellingen hangt ook af van de kwaliteit en beschikbaarheid van data. De data over landgebruik kan in principe jaarlijks bijgewerkt worden met de nieuwste landbouw- en landinrichtingsdata. Daarmee kan er ook een een-op-een koppeling plaatsvinden van de vogelgriepuitbraakdata met de fingerende landschapsdata. Er kan helaas geen een-op-een koppeling plaatsvinden van de vogelgriepuitbraakdata met vogeltellingen omdat deze vogeltellingen grotendeels ontbreken, we hebben slechts van enkele jaren (de winterperiodes van 2012 t/m 2015) beschikbare vogeldata kunnen gebruiken. Ook is er geen rekening gehouden met mogelijk grote veranderingen in aantallen wilde vogels gedurende de winter, met name bij vogelsoorten die grotendeels buiten Nederland broeden. Dit zal zeker een bijdrage leveren aan de onzekerheid van het model, maar betere data zijn er op dit moment niet. In het recente risicofactorenonderzoek (Gonzales et al., 2022) was het type pluimveehouderij een significante risicofactor voor HPAI-besmetting bij pluimveebedrijven. Kalkoen- en eendenbedrijven hadden een significant hoger besmettingsrisico vergeleken met legkippen- en fokbedrijven, terwijl het risico voor vleeskuikenbedrijven lager was. Omdat we in de voorspelmodellen HPAI-besmettingen in alle pluimveebedrijfstypen als een proxy gebruiken voor het besmettingsniveau in een gebied, zou het specifieke risico voor één bepaald type pluimveehouderij de toekenning van gebiedsrisico's kunnen beïnvloeden. Om voor dit effect te corrigeren, hebben we voor de cases en de controles overeenkomstige bedrijfstypen gebruikt. Om het risico van pluimveebedrijfstype expliciet als variabele in het model mee te nemen zouden we de voorspellingen moeten beperken tot bedrijfsniveau, en daarmee de risicovoorspellingen beperken tot alleen die gebieden waar pluimveebedrijven aanwezig zijn. Echter het risicobeoordelingsmodel, hoewel met een wat lagere accuraatheid (AUC = 0,73; 95% btbh: 0,69 – 0,78) dan het geselecteerde Random Forest voorspelmodel (AUC = 0,80; 95% btbh: 0,71 – 0,91), is eenvoudig en flexibel waardoor er gemakkelijk risicofactoren toegevoegd of verwijderd kunnen worden met het voortschrijden van het inzicht in de toekomst. Het besmettingsrisico dat kan worden toegeschreven aan pluimveebedrijfstype kan daarom makkelijk worden toegevoegd en worden berekend voor die rastercellen waarin pluimveebedrijven aanwezig zijn.

De modellen in hun huidige vorm vormen hun voorspellingen aan de hand van gebeurtenissen uit het verleden en zijn bedoeld om risico's in te schatten in de (nabije) toekomst. Veranderingen in één of meer van de variabelen, zoals landgebruik en de ruimtelijke verdeling van wilde vogels kunnen de voorspellingen van besmettingsrisico's doen veranderen. Daarnaast hebben we veranderingen gezien in de circulatie van het virus wat betreft de infectieusiteit en pathogeniciteit bij pluimvee en wilde vogels (Beerens et al 2021) en de epidemiologie van het verspreiding (EFSA, 2023). Er was ruime viruscirculatie tussen oktober 2020 en 2022 (HPAIV H5N8 en het huidige HPAIV H5N1) met als gevolg een grote mortaliteit onder wilde vogels waarbij een grotere diversiteit aan vogelsoorten slachtoffer werd dan in de voorgaande jaren (EFSA, 2023). Voorheen waren de besmettingen van HPAIV in de tijd gelinkt aan de jaarlijks terugkerende vogeltrek in de herfst en de winter van de verschillende trekkende watervogelsoorten. Echter in 2022 zijn er ook HPAI-besmettingen waargenomen in wilde vogels en pluimveebedrijven tijdens de zomermaanden, wat een teken kan zijn van blijvend geïnfecteerde populaties wilde vogels. In de toekomst kan het dan ook belangrijk worden om ook de zomerdata van de vogeltellingen te gebruiken. De veranderende epidemiologie verzwakt de voorspellende waarde van de modellen die hier zijn gepresenteerd, het is daarom belangrijk om de modellen aan te passen zodra nieuwe observaties en databronnen beschikbaar zijn, om de voorspellingen te blijven verbeteren in de toekomst.

5 Conclusies

- Wij hebben twee verschillende modellen (Random Forest en risicobeoordelingsmodel) ontwikkeld om gebieden met een hoog HPAI-besmettingsrisico op commerciële pluimveebedrijven in Nederland te identificeren. Beide modellen komen over het algemeen overeen in de identificatie van regio's met hoge of lage HPAI-besmettingsrisico's op commerciële pluimveebedrijven in de herfst en winter periode. Een hoog risico op HPAI-besmetting bij pluimveebedrijven is te zien in de kustregio's en laaggelegen gebieden in Groningen, Friesland, Flevoland, Noord- en Zuid-Holland en Zeeland. Verder is er een verhoogd risico op HPAI-besmetting bij pluimveebedrijven te zien in de stroomgebieden van de grote rivieren in het midden van het land: de IJssel, Neder-Rijn, Waal en Maas. Verder is er een hoog risico op HPAI-besmetting bij pluimveebedrijven te zien in Limburg ter hoogte van de Maasplassen bij Roermond.
- Er is een laag risico op HPAI-besmetting bij pluimveebedrijven te zien voor de hooggelegen delen van Groningen, Drenthe, Overijssel, Gelderland, Utrecht, Noord-Brabant en Limburg.
- De aanwezigheid van water en agrarische graslanden in de directe omgeving van pluimveebedrijven vergroten het risico op HPAI-besmetting bij pluimveebedrijven. Dit risico neemt tevens toe als de oppervlakte aan water en grasland in de directe omgeving van het pluimveebedrijf toeneemt.
- De aanwezigheid van Wilde Eenden, Bergeenden, Slobeenden, Brandganzen en Knobbelzwanen lijkt geassocieerd met het risico op HPAI-besmettingen bij pluimveebedrijven.



6 Aanbevelingen

- Voorspelmodellen zouden in ieder geval jaarlijks opnieuw moeten worden aanpast met aanvulling van de nieuwe uitbraakgegevens en data van risicofactoren.
- De voorspellende waarde van de modellen kan verbeterd worden door jaarlijkse data over dichtheid van wilde vogels in Nederland beschikbaar te hebben voor deze modellering.
- Een goede toevoeging zou zijn om de variatie in dichtheid van de verschillende vogelsoorten gedurende het jaar te gebruiken (seizoenveranderingen).
- De risicokaarten uit dit project kunnen gebruikt worden om surveillance te verbeteren, door actievere vormen van surveillance op de aanwezigheid van HPAIv onder wilde vogels en op pluimveebedrijven te vergroten in de gebieden met een hoog HPAI-besmettingsrisico op pluimveebedrijven.
- De risicokaarten kunnen van nut zijn bij het plannen van de ruimtelijke inrichting van het landschap (nieuwe locaties van te starten pluimveebedrijven).



Literatuur

- Animal and Plant Health Agency (UK); Erasmus Medical Centre (NL); Friedrich Loeffler Institute (DE); Istituto Zooprofilattico Sperimentale delle Venezie (IT); Linnaeus University (SE); Wageningen University (NL). Report about HPAI introduction into Europe, HPAI detection in wild birds and HPAI spread between European holdings in the period 2005–2015. (2017). EFSA Supporting Publ., 14, 1284E.
- Beerens N, Germeraad EA, Venema S, Verheij E, Pritz-Verschuren SBE, Gonzales JL (2021). Comparative pathogenicity and environmental transmission of recent highly pathogenic avian influenza H5 viruses, *Emerging Microbes & Infections*, 10:1, 97-108, DOI: 10.1080/22221751.2020.1868274.
- EFSA , ECDC, EURL, Adlhoch C, Fusaro A, Gonzales JL, Kuiken T, Marangon S, Niqueux É, Staubach C, Terregino C, Aznar I, Muñoz Guajardo I and Baldinelli F (2023). Scientific report: Avian influenza overview September–December 2022. (2023). *EFSA Journal*, 21(1):7786, 63 pp. <https://doi.org/10.2903/j.efsa.2023.7786>
- Elbers ARW & Gonzales JL (2020) Quantification of visits of wild fauna to a commercial free-range layer farm located in an avian influenza hot-spot area assessed by video-camera monitoring. *Transboundary Infectious Diseases* 67, 661- 677. DOI: 10.1111/tbed.13382.
- Elbers ARW & Gonzales JL (2021) Efficacy of an automated laser for reducing wild bird visits to the free range area of a poultry farm. *Sci Rep* 11, 12779. <https://doi.org/10.1038/s41598-021-92267-z>
- Gonzales JL, Hennen WHGJ, Petie R, de Freitas Costa E, Beerens N, Slaterus R, Kuiken T, Stahl J, Elbers ARW (2022). Risicofactoren voor introductie van HPAI-virus op Nederlandse commerciële pluimveebedrijven, 2014-2022. Wageningen Bioveterinary Research Report 2211632.
- R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Schreuder J, de Knecht HJ, Velkers FC, Elbers ARW, Stahl J, Slaterus R, Stegeman JA, de Boer WF (2022). Wild bird densities and landscape variables predict spatial patterns in HPAI outbreak risk across the Netherlands. *Pathogens*, 11, 549. <https://doi.org/10.3390/pathogens11050549>.
- Sovon Dutch Center for Field Ornithology. *Bird Atlas of the Netherlands*, 1st ed.; Kosmos Publishers: Utrecht, The Netherlands; Antwerp, Belgium, 2018; pp. 1–640.

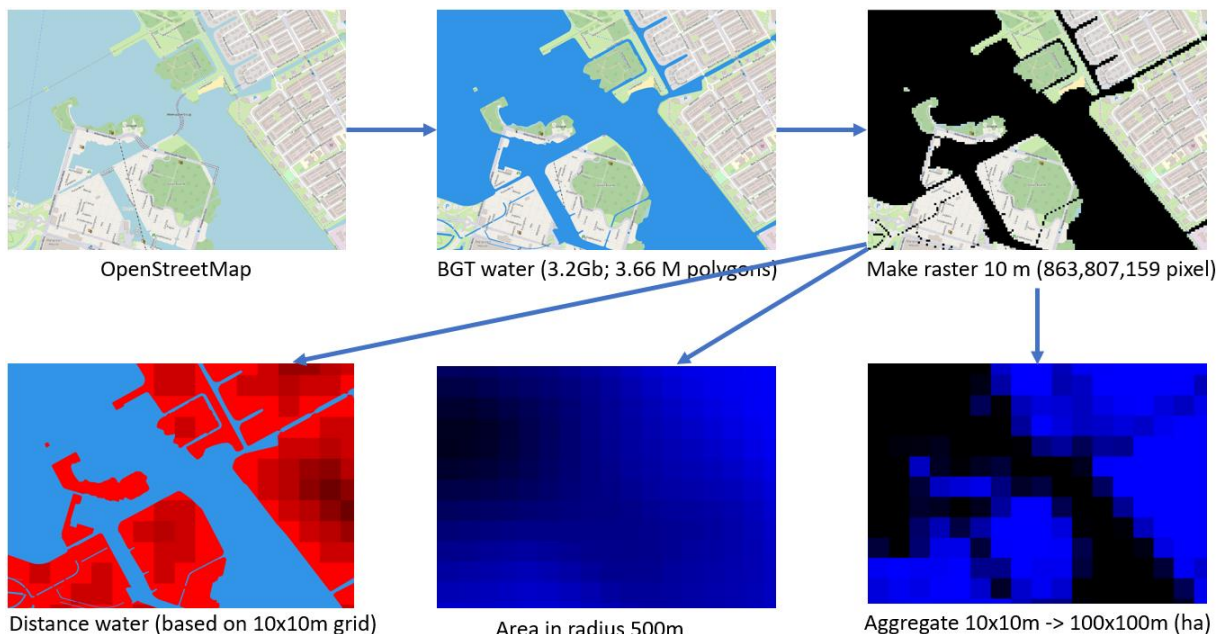


Bijlage A1

Toelichting GIS bewerkingen t.b.v. maken van de kaarten

De GIS bewerkingen die nodig zijn om kaarten te maken wordt geïllustreerd met onderstaande figuur. Ter oriëntatie wordt als achtergrond de kaart van OpenStreetMap gebruikt (kaart links boven). Het aspect water uit het geografisch data bestand Basisregistratie Grootchalige Topografie (BGT) wordt gebruikt als voorbeeld. De hieruit verkregen waterkaart bevat 3,66 miljoen vlakken (slootjes, etc.) met een bestandsomvang van 3.2 Gb. Van deze waterkaart wordt een raster met een resolutie van 10x10 meter gemaakt, zie rechtsboven. Voor elke cel van 10x10 meter wordt bepaald of deze meer of minder voor de helft uit water bestaat. Hierdoor is een rasterkaart voor water gemaakt die bestaat uit cellen met of zonder water, waarde 1 of 0. Met deze zeer gedetailleerd rasterkaart voor water kunnen als voorbeeld de volgende kaarten gemaakt worden:

- Percentage water dat elke oppervlakte van 1 hectare bevat (zie kaart rechts beneden). Het doel van het project is een risicokaart op hectare niveau te maken, voor elke hectare worden het aantal 10x10 meter cellen met water bij elkaar geteld. Als er bijvoorbeeld 73 cellen water hebben (en dus $100-73=27$ geen water), dan bestaat die hectare voor 73% uit water.
- Oppervlakte water in hectare in een straal van 500 meter (zie kaart beneden midden). Vanuit het centrum van 1 hectare (i.e. midpoint van 100x100 meter cel) worden alle water cellen van 10x10 meter (i.e. 0,01 ha) in een straal van 500 meter bij elkaar geteld, het resultaat is het aantal ha water binnen deze straal van 500 meter.
- Afstand hectare tot dichtstbijzijnde water (zie kaart links beneden). Vanuit het centrum van 1 hectare wordt het dichtstbijzijnde waterelement (i.e. 10x10 meter cel met water) gezocht, de afstand tot dit element in meters is de waarde op de kaart.





Bijlage A2

Tabel A1 Overzicht van de verdeling van waarden landschapsvariabelen bij cases en controles

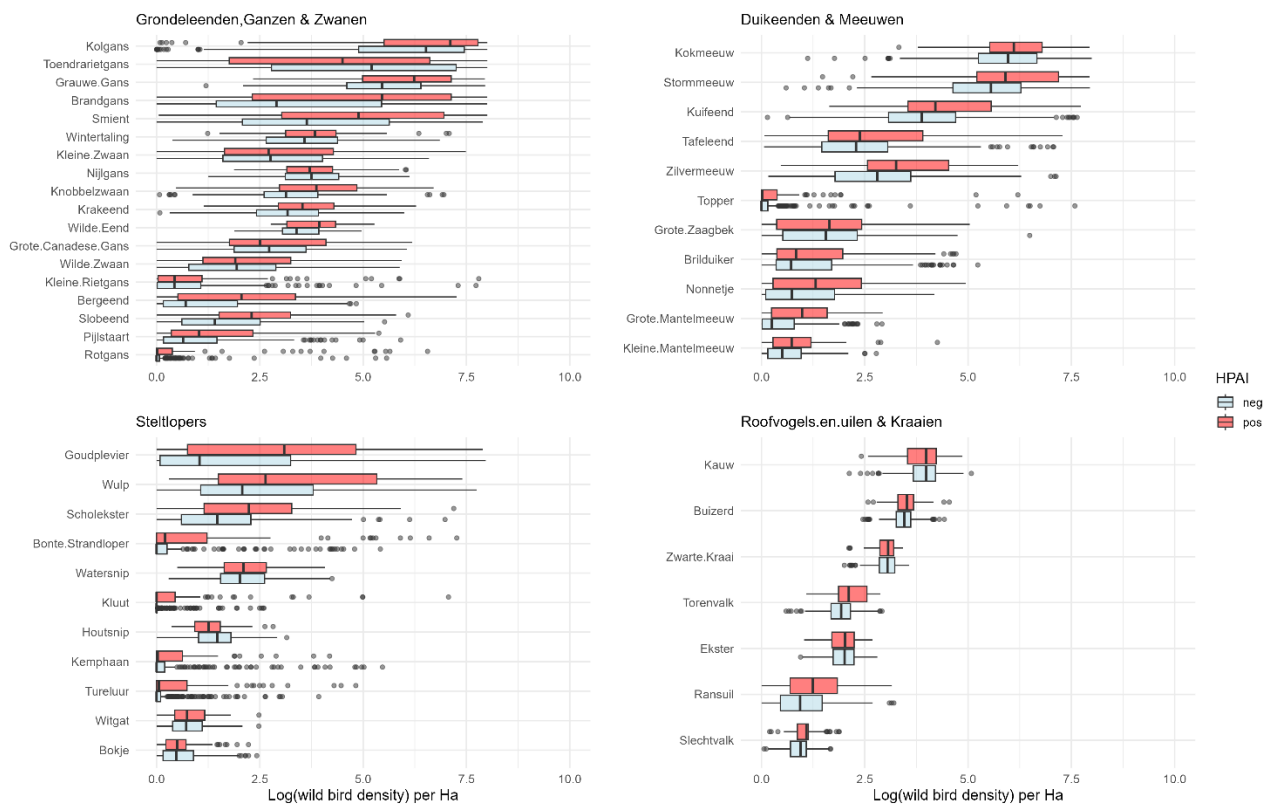
HPAI-status	Variabelen	n	min	max	gemid.	mediaan	Q1	Q3
case	Water %	96	0	47	6.4	5	1	9.3
control	Water %	316	0	32	2.9	2	0	5
case	Water - Ha500	96	0	39.9	5.2	3.6	2.1	6.5
control	Water - Ha500	316	0.03	15.0	3.0	2.4	1.6	3.6
case	Water Reilly	96	0	4702.4	632.0	500.3	0.5	926.0
control	Water Reilly	316	0.001	3201.8	277.9	200.2	0.3	500.1
case	Gras %	96	0	92	47.0	47	25	69
control	Gras %	316	0	99	41.9	42.4	23.5	60.0
case	BosBomen %	96	0	0	0	0	0	0
control	BosBomen %	316	0	85	1.8	0	0	0
case	BosBomen - Ha500	96	0	26.2	1.9	0	0	0.9
control	BosBomen - Ha500	316	0	47.5	4.1	0.5	0	4.8
case	NNN %	96	0	29	0.8	0	0	0
control	NNN %	316	0	76	1.6	0	0	0
case	NNN-Ha500	96	0	80.7	4.2	0	0	2.3
control	NNN-Ha500	316	0	67.8	4.6	0	0	3.3
case	Zoutige grond %	96	0	24.0	0.6	0	0	0.1
control	Zoutige grond %	316	0	1.8	0.1	0	0	0
case	Zoutige kwelder %	96	0	0.1	0.001	0	0	0
control	Zoutige kwelder %	316	0	0.2	0	0	0	0
case	Duinen %	96	0	0	0	0	0	0
control	Duinen %	316	0	0	0	0	0	0
case	Moeras %	96	0	1	0.02	0	0	0
control	Moeras %	316	0	1.8	0.01	0	0	0
case	Riet %	96	0	5.1	0.13	0	0	0
control	Riet %	316	0	2	0.038	0	0	0
case	Hoogveen %	96	0	81	3.4	0	0	0
control	Hoogveen %	316	0	84.2	6.3	0	0	0.3
case	Natte natuur %	96	0	20	0.8	0	0	0
control	Natte natuur %	316	0	23	0.6	0	0	0
case	Windturbines	96	0	9	0.3	0	0	0
control	Windturbines	316	0	5	0.1	0	0	0
case	Gras landbouw %	96	0	100	34.9	23.5	5.1	60.5
control	Gras landbouw %	316	0	100	38.4	28.5	4.8	75.3
case	Bos	96	0	100	10.9	0	0	0
control	Bos	316	0	100	13.8	0	0	0

Vervolg Tabel A1 Overzicht van de verdeling van waarden landschapsvariabelen bij cases en controles

HPAI-status	Variabelen	n	min	max	gemid.	mediaan	Q1	Q3
case	Gras- Ha1000	96	0	319.0	120.2	115.1	49.8	178.1
control	Gras- Ha1000	316	0	320.8	128.4	114.7	46.1	205.4
case	BosBomen Reilly	96	0	10012.2	1093.0	0	0	0.6
control	BosBomen Reilly	316	0	10013.4	1384.3	0.004	0	1.6
case	NNN Reilly	96	0	10013.4	1855.7	0.108	0	2.6
control	NNN Reilly	316	0	10013.4	2020.3	0	0	80.2
case	Fruitbomen - Ha1000	96	0	32.8	1.648	0	0	0.01
control	Fruitbomen - Ha1000	316	0	80.6	2.692	0	0	0.16
case	Bomenteelt - Ha1000	96	0	22.7	0.51	0	0	0
control	Bomenteelt - Ha1000	316	0	21.3	0.586	0	0	0
case	Suikerbiet - Ha1000	96	0	57.6	9	1.1	0	14.0
control	Suikerbiet - Ha1000	316	0	69.5	8.2	0.8	0	11.2
case	Granen - Ha1000	96	0	143.4	16.5	3.6	0	22.3
control	Granen - Ha1000	316	0	157.6	15.5	2.9	0	16.1
case	Mais - Ha1000	96	0	96.3	18.7	10.7	2.7	28.4
control	Mais - Ha1000	316	0	108.6	17.3	11.2	1.3	26.2
case	Overige gewassen -Ha1000	96	0	146.1	27.2	8.7	0.04	44.2
control	Overige gewassen - Ha1000	316	0	138.9	21.2	5.2	0	29.1
case	Bevolkingsdichtheid	96	0	3397.6	573.7	242.9	162.4	738.7
control	Bevolkingsdichtheid	316	0	4092	696.0	381.2	137.7	838.6
case	Bevolking binnen 1000 m	96	0	10870.5	1992.1	1040.5	542.2	2361.4
control	Bevolking binnen 1000 m	316	0	13650.1	2416.8	1397.1	623.2	3089.6

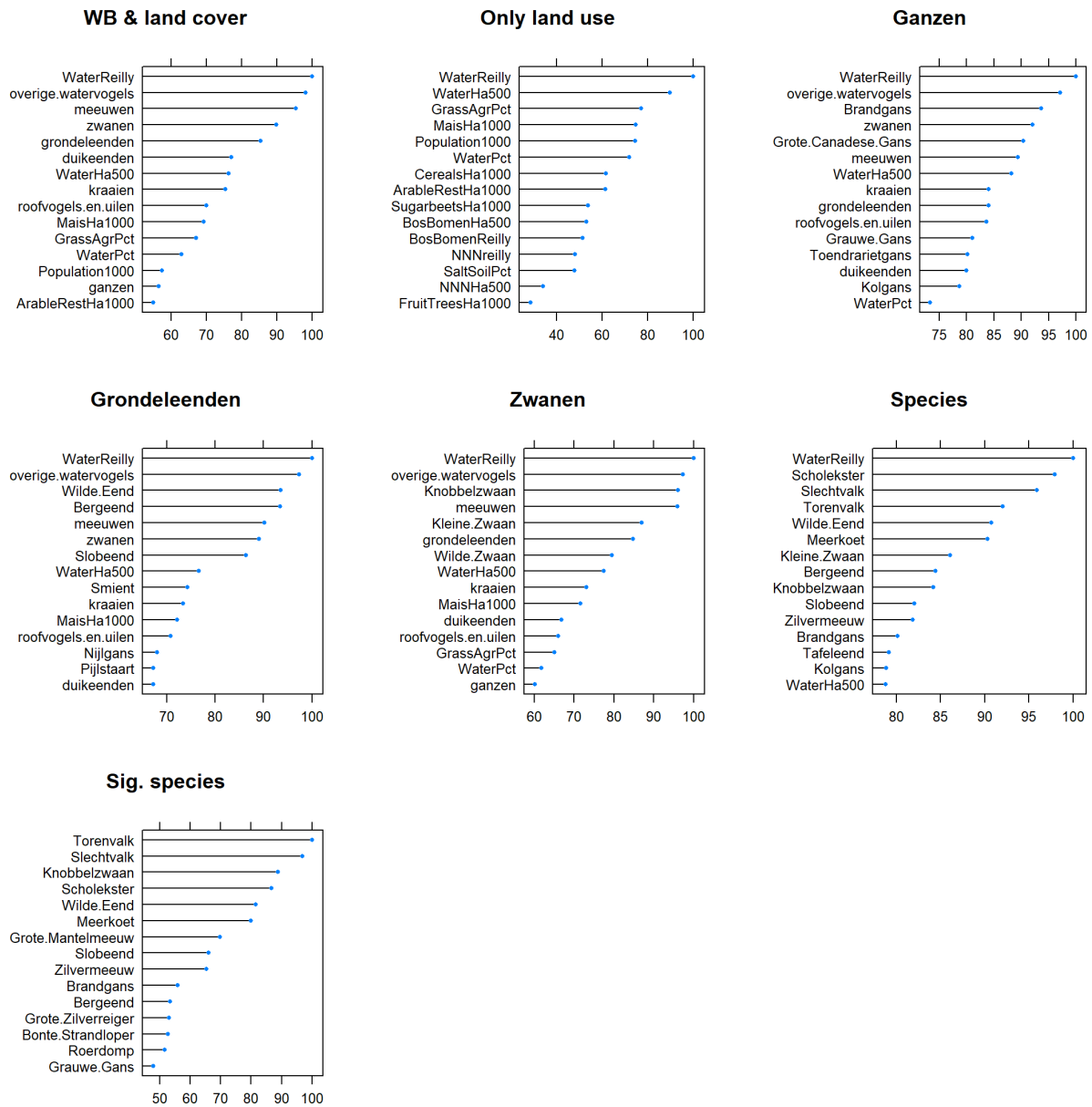
Vervolg Tabel A1 *Overzicht van de verdeling van waarden vogelsoortgroep dichtheden bij cases en controles*

hpai-status	Variabelen	n	min	max	gemid.	mediaan	Q1	Q3
case	duikeenden	96	0	3168.3	377.1	58.5	0	386.1
control	duikeenden	316	0	5044.4	212.1	32.0	0	134.6
case	ganzen	96	0	8973.6	2635.9	1321.2	0	5264.3
control	ganzen	316	0	8689.9	1688.8	0	0	3111.4
case	grondeleenden	96	0	5087.4	994.4	443.6	185.705	1584.2
control	grondeleenden	316	0	4003.7	478.1	226.2	121.052	565.4
case	zwanen	96	3.9	2157.9	249.2	72.4	30.926	194.4
control	zwanen	316	0	1522.6	121.9	61.8	24.541	130.4
case	steltlopers	96	0	4533.0	394.4	0	0	46.3
control	steltlopers	316	0	3997.0	107.9	0	0	0
case	meeuwen	96	0	4656.1	1447.2	1032.8	634.487	2283.4
control	meeuwen	316	0	5024.5	1020.2	657.0	260.941	1408.5
case	roofvogels.en.uilen	96	26.0	128.0	49.1	45.4	37.812	56.2
control	roofvogels.en.uilen	316	0	110.6	42.6	40.2	34.674	47.6
case	overige.watervogels	96	0	2836.9	528.0	237.0	152.196	673.6
control	overige.watervogels	316	0	1799.9	246.7	161.9	79.095	292.6
case	kraaien	96	23.5	155.8	79.1	80.7	61.018	98.7
control	kraaien	316	18.1	195.5	80.7	79.4	60.808	97.9

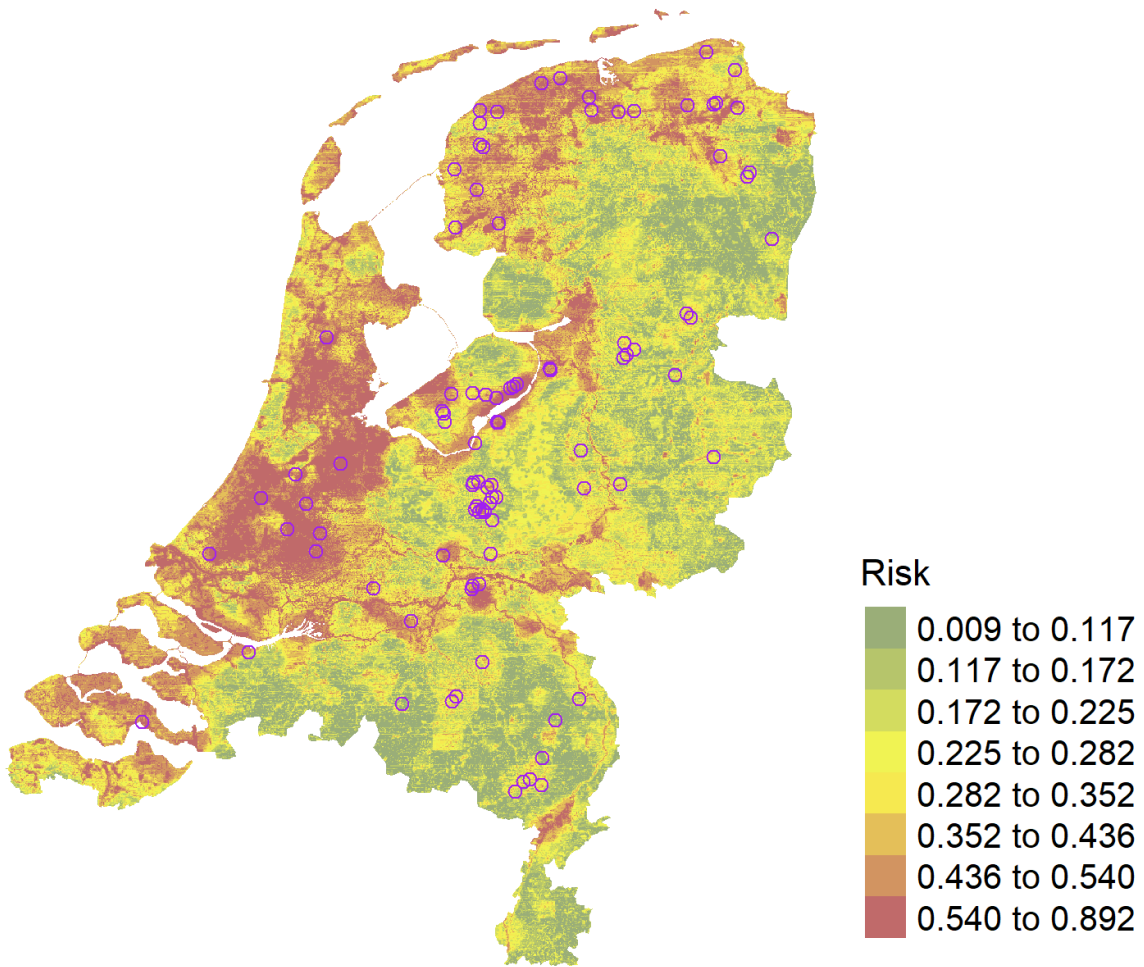


Figuur A1 Log-getransformeerde boxplot-verdelingen van wilde vogelsoorten (groepen) bij HPAI-onbesmette (controls; blauwe kleur), en bij HPAI-besmette (cases; rode kleur) gebieden/pluimveehouderijen.

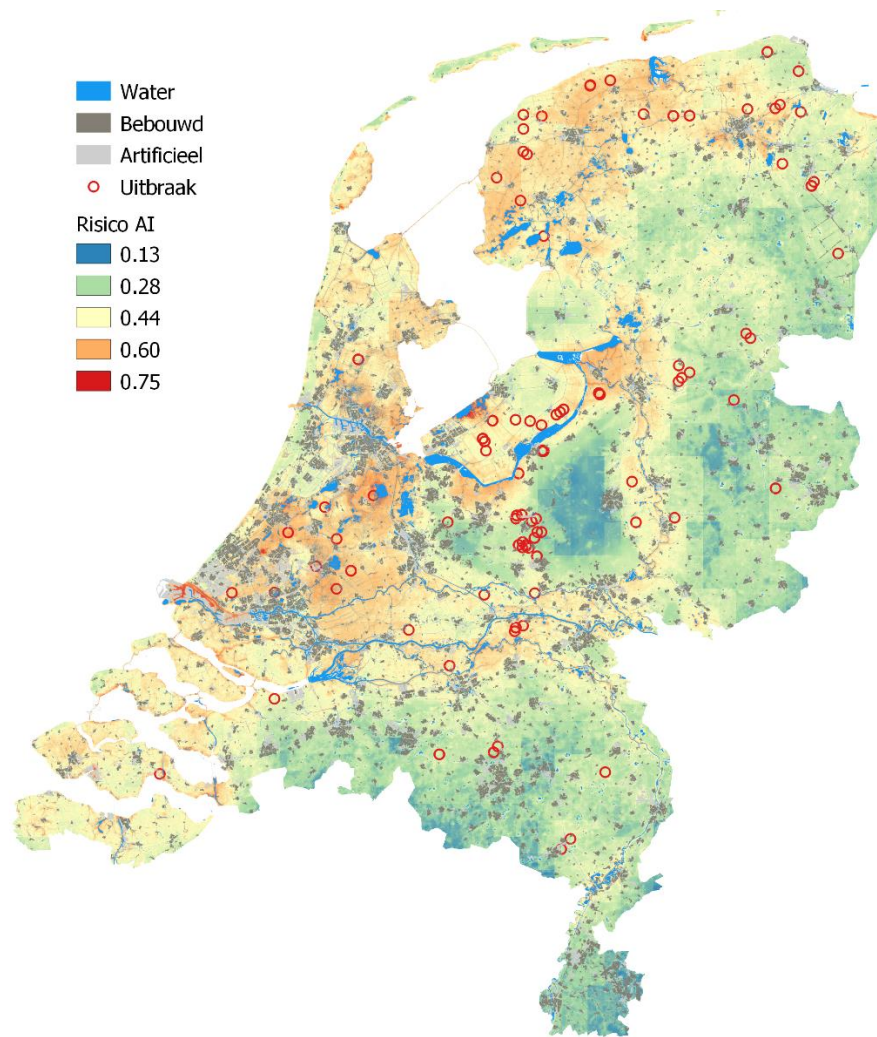
Bijlage A3



Figuur A2 Belang van variabelen (15 meest belangrijk variabelen) voor alle zeven willekeurige voorspelmodellen.



Figuur A3 Relatieve HPAI-besmettingsrisico op pluimveebedrijven volgens het best presterende Random Forest model (op een doorlopende schaal). Gebieden in groen zijn gebieden met laag risico in vergelijking met bruine gebieden. Paarse cirkels representeren HPAI-besmette pluimveebedrijven.



Figuur A4 Relatieve HPAI-besmettingsrisico op pluimveebedrijven volgens het risicobeoordelingsmodel (op een doorlopende schaal). Gebieden in groen zijn gebieden met laag risico in vergelijking met bruine gebieden. Rode cirkels representeren HPAI-besmette pluimveebedrijven.



Bijlage A4

Table A2 Lijst van wilde vogelsoorten binnen samengestelde soortgroepen die gebruikt zijn bij de analyse.

Soortgroep	Orde	Familie	Naam
Duikeenden	Anseriformes	<i>Anatidae</i>	Tafeleend
			Kuifeend
			Toppereend
			Grote Zaagbek
			Middelste Zaagbek
			Brilduiker
			Nonnetje
			Topper
			Toendrarietgans
			Kleine Rietgans
Ganzen	Anseriformes	<i>Anatidae</i>	Grauwe Gans
			Kolgans
			Canadese Gans
			Rotgans
			Brandgans
			Grote Canadese Gans
			Nijlgans
			Bergeend
			Krakeend
			Smient
Grondeleenden	Anseriformes	<i>Anatidae</i>	Wilde Eend
			Slobeend
			Pijlstaart
			Wintertaling
			Knobbelzwaan
			Kleine Zwaan
			Wilde Zwaan
			Eider
			Zwarte Zeeend
			IJseend
Zwanen	Anseriformes	<i>Anatidae</i>	Grote Zee eend
			Drieteenmeeuw
			Kokmeeuw
			Stormmeeuw
			Grote Mantelmeeuw
			Grote Burgemeester
			Zilvermeeuw
			Kleine Mantelmeeuw
			Pontische Meeuw
			Geelpootmeeuw
Zee-eenden	Anseriformes	<i>Anatidae</i>	
Meeuwen	Charadriiformes	<i>Laridae</i>	

Soortgroep	Orde	Familie	Naam		
Steltlopers	Charadriiformes	<i>Charadriidae</i> <i>Recurvirostridae</i>	Goudplevier		
			Kluut		
		<i>Scolopacidae</i>	Scholekster		
			Kievit		
			Zilverplevier		
			Houtsnip		
			Bokje		
			Watersnip		
			Regenwulp		
			Wulp		
			Tureluur		
			Bosruiter		
			Drieteenstrandloper		
			Kleine Strandloper		
			Bonte Strandloper		
Kemphaan					
Witgat					
Roofvogels en uilen	Anders	<i>Accipitridae</i>	Sperwer		
			Havik		
			Blauwe Kiekendief		
			Zeearend		
			Buizerd		
			Torenavalk		
		<i>Falconidae</i>	Smelleken		
			Slechtvalk		
			Ransuil		
		Andere vogels	Anders	<i>Ardeidae</i>	Roerdomp
					Blauwe Reiger
					Grote Zilverreiger
<i>Ciconiidae</i>	Ooievaar				
	Aalscholver				
<i>Phalacrocoracidae</i>	Kuifaalscholver				
	Dodaars				
<i>Podicipedidae</i>	Fuut				
	Geoorde Fuut				
	<i>Rallidae</i>			Waterral	
		Waterhoen			
		Meerkoet			
Kraaien	Passeriformes	<i>Corvidae</i>	Gaai		
			Ekster		
			Kauw		
			Zwarte Kraai		
			Roek		
			Raaf		
Duiven	anders	<i>Columbidae</i>	Turkse Tortel		
Fazanten	anders	<i>Phasianidae</i>	Fazant		

Wageningen Bioveterinary Research
Postbus 65
8200 AB Lelystad
T 0320 23 82 38
info.bvr@wur.nl
wur.nl/bioveterinary-research

Wageningen Bioveterinary Research
Rapport 2301353

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, 6.800 medewerkers (6.000 fte) en 12.900 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.