



Vlaanderen
is wetenschap



Modelinstrumentarium voor het voorspellen van overwinterende aantallen wintertaling (Anas crecca) in de Boven- Zeeschelde

Deelrapport voor het Integraal plan Boven-Zeeschelde

Vanoverbeke Joost, Van Reyckegem Gunther, Van Braeckel Alexander & Van den Bergh Erika

**INSTITUUT
NATUUR- EN BOSONDERZOEK**

Auteurs:

Vanoverbeke Joost, Van Reyckegem Gunther, Van Braeckel Alexander & Van den Bergh Erika

Instituut voor Natuur- en Bosonderzoek

Het Instituut voor Natuur- en Bosonderzoek (INBO) is het Vlaams onderzoeks- en kenniscentrum voor natuur en het duurzame beheer en gebruik ervan. Het INBO verricht onderzoek en levert kennis aan al wie het beleid voorbereidt, uitvoert of erin geïnteresseerd is.

Vestiging:

Herman Teirlinckgebouw

INBO Brussel

Havenlaan 88 bus 73, 1000 Brussel

www.inbo.be

e-mail:

joost.vanoverbeke@inbo.be

Wijze van citeren:

Vanoverbeke J., Van Reyckegem G., Van Braeckel A. & Van den Bergh E. (2019). Modelinstrumentarium voor het voorspellen van overwinterende aantallen wintertaling (Anas crecca) in de Boven-Zeeschelde - Deelrapport voor het Integraal plan Boven-Zeeschelde. Rapporten van het Instituut voor Natuur- en Bosonderzoek 2019 (15). Instituut voor Natuur- en Bosonderzoek, Brussel.
DOI: doi.org/10.21436/inbor.14517871

D/2019/3241/093

Rapporten van het Instituut voor Natuur- en Bosonderzoek 2019 (15)

ISSN: 1782-9054

Verantwoordelijke uitgever:

Maurice Hoffmann

Druk:

Managementondersteunende Diensten van de Vlaamse overheid

Foto cover:

Yves Adams / Vilda

Dit onderzoek werd uitgevoerd in opdracht van:

Agentschap Natuur en Bos



Modelinstrumentarium voor het voorspellen van overwinterende aantallen wintertaling (*Anas crecca*) in de Boven- Zeeschelde

Deelrapport voor het Integraal plan Boven-Zeeschelde

**Joost Vanoverbeke, Gunther Van Ryckegem, Alexander Van Braeckel & Erika
Van den Bergh**

Rapporten van het Instituut voor Natuur- en Bosonderzoek 2019 (15)
doi.org/10.21436/inbor.14517871

Inhoudstafel

1	INLEIDING	8
2	MATERIAAL EN METHODE	11
2.1	ECOLOGISCH ACHTERGROND	11
2.2	GEGEVENS	12
2.2.1	Vogeltellingen	12
2.2.2	Europese trend	15
2.2.3	Strengheid van winter	15
2.2.4	Habitatkenmerken	15
2.2.4.1	Oppervlakte en breedte van de ecotopen	15
2.2.4.2	Helling	16
2.2.4.3	Spreiding droogvalduur	16
2.2.4.4	Maximale vloedsnelheid	17
2.2.4.5	Luwte Index	17
2.3	METHODIEK	18
2.3.1	Generalized linear mixed models	18
2.3.2	Andere statistische technieken	19
2.4	VERKENNENDE ANALYSE	19
2.4.1	Vogeltellingen	19
2.4.2	Lineair verband tussen afhankelijke en onafhankelijke variabelen	20
2.4.3	Ruimtelijke patronen in de habitatkenmerken	20
2.4.4	Multicollineariteit	20
2.5	ANALYSE VAN AANTALLEN	21
2.5.1	Kalibratie	21
2.5.1.1	Globaal overzicht	21
2.5.1.2	Opbouw van de GLMMs	22
2.5.1.3	Model score	23
2.5.1.4	Cross-validatie	23
2.5.1.5	Stapsgewijze modelopbouw	24
2.5.1.6	Bepalen van het optimale model	25
2.5.2	Modelassumpties	25
2.5.3	Berekenen van voorspelde waarden en confidentie interval	26
2.5.4	Validatie	27
3	RESULTATEN	28
3.1	VERKENNENDE ANALYSE	28
3.1.1	Vogeltellingen	28
3.1.2	Habitatkenmerken	31
3.1.2.1	Gebieden verwijderd uit de dataset	31
3.1.2.2	Lineair verband tussen aantallen en habitatvariabelen	31
3.1.2.3	Ruimtelijke variatie	31
3.1.2.4	Multicollineariteit tussen habitatvariabelen	31
3.1.3	Vergelijking van statistische technieken	33
3.2	ANALYSE VAN AANTALLEN	33
3.2.1	Kalibratie	33
3.2.2	Modelassumpties	37
3.2.3	Validatie	37
4	DISCUSSIE	40

5	CONCLUSIES.....	43
6	REFERENTIES.....	44
	BIJLAGE 1: WEERGAVE VAN DE RUIMTELIJKE VARIATIE IN DE HABITATKARAKTERISTIEKEN	46
	BIJLAGE 2: LIJST VAN ALLE BEREKENDE MODELLEN VOOR WINTERTALING IN FUNCTIE VAN HABITATKENMERKEN	50

Lijst van figuren

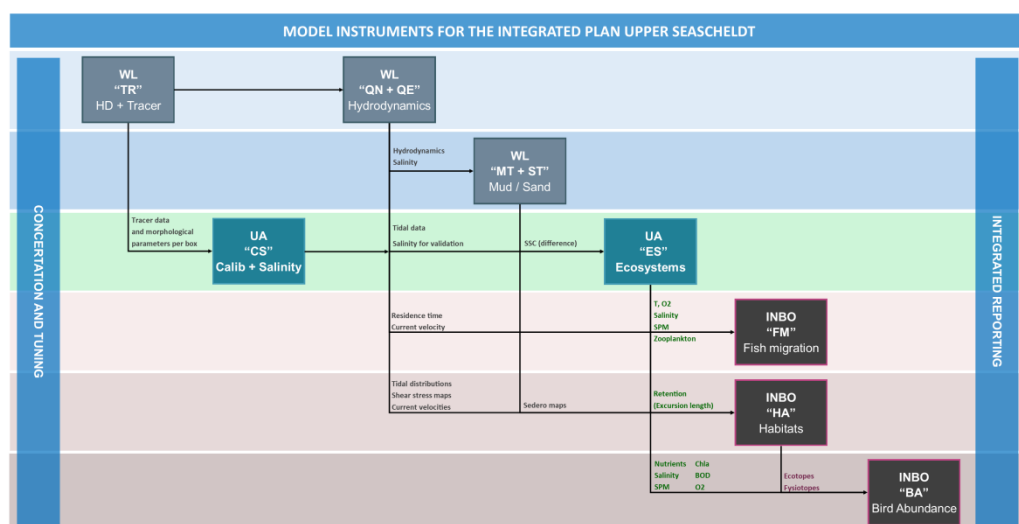
Figuur 1-1: Sequentie van doorrekeningen en datastroom met modelinstrumenten binnen het Integraal plan Boven-Zeeschelde.	8
Figuur 1-2: Historische trends van de belangrijkste soorten eenden in de Zeeschelde (uitgedrukt als % van de Europese populatie).	9
Figuur 2-1: Overzichtskaart per saliniteitszone van de vogeltelgebieden en OMES-segmenten langsheen de Zeeschelde. De rood omlijnde gebieden zijn de vogeltelgebieden, de grotere genummerde zones de OMES-segmenten.	14
Figuur 2-2: Globaal overzicht van de gevolgde methode voor het opstellen van een voorspellend model voor vogelaantallen in functie van habitat- en ecosysteemkenmerken.	21
Figuur 2-3: Schematische voorstelling van de methode van K-fold cross-validatie. Score = LG score.	24
Figuur 3-1: Aantal wintertaling per OMES segment (boven) en per saliniteitszone (onder). Gemiddelde over de wintermaanden (oktober-maart) \pm 95% confidentie interval. ...	29
Figuur 3-2: Maandeffect in preliminaire analyse (GLMM) met enkel ruimtelijke en temporele afhankelijkheidsstructuur. Het maandeffect is een schatter van de afwijking van de log aantallen binnen elke maand ten opzichte van het algemeen gemiddelde. De boxplots geven de variatie weer voor dezelfde maand in de verschillende jaren (2007-2012).	30
Figuur 3-3: Overzicht van de paarsgewijze correlaties tussen de habitatkarakteristieken in de Boven-Zeeschelde.	32
Figuur 3-4: Relatie tussen aantal wintertaling (\pm 95%) en habitatkenmerken op basis van het finale model.	36
Figuur 3-5: Relatie tussen residuen en verklarende habitatkenmerken voor wintertaling.	37
Figuur 3-6: Relatie tussen voorspelde aantallen en waargenomen aantallen wintertaling in de telgebieden. Vergelijking tussen validatie en kalibratie set. Punten geven het gemiddelde per telgebied en winter (gemiddeld over wintermaanden) weer voor de waargenomen aantallen. De grijze band geeft het 95% betrouwbaarheidsinterval weer van de voorspelde waarden. Links: voorspellingen zonder Europese trend in de regressievergelijking. Rechts: voorspellingen zonder Europese trend in de regressievergelijking.	38
Figuur 3-7: Relatie tussen voorspelde aantallen en waargenomen aantallen wintertaling per OMES zone. Vergelijking tussen validatie en kalibratie set. Punten geven het gemiddelde per OMES zone en winter (gemiddeld over wintermaanden) weer voor de waargenomen aantallen. De grijze band geeft het 95% betrouwbaarheidsinterval weer van de voorspelde waarden. Links: voorspellingen zonder Europese trend in de regressievergelijking. Rechts: voorspellingen zonder Europese trend in de regressievergelijking.	39

Lijst van tabellen

Tabel 1: habitatkenmerken voor de vogeltelgebieden langsheen de Zeeschelde, waaruit tijdens de verkennende analyse een verdere selectie wordt gemaakt.	18
Tabel 2: verkennende inschatting van de ruimtelijk en temporele variatie in aantallen vogels. De waarden geven de grootteorde (i.e. 10^x geeft de verhouding weer tussen hoogste en laagste aantallen binnen het 95% betrouwbaarheidsinterval) van de verschillen in waargenomen aantallen op ruimtelijke schaal (telgebieden) en temporele schaal (winterjaren).	30
Tabel 3: Variance Inflation Factors (VIF) voor de habitatkarakteristieken in de Boven-Zeeschelde. Links: VIF waarden indien alle initieel geselecteerde habitatkarakteristieken worden in rekening gebracht. Rechts: VIF waarden indien enkel rekening wordt gehouden met de selectie van karakteristieken die zal worden gebruikt in verdere analyses.	32
Tabel 4: gemiddelde en standaard deviatie gebruikt voor herschaling van de variabelen in de modelselectie.....	34
Tabel 5: geschatte regressie coëfficiënten β voor de relatie tussen aantallen wintertaling en verklarende variabelen in het finale model. Merk op dat deze parameters van toepassing zijn op de herschaalde variabelen (zie boven). Waarden tussen haakjes geven parameterschattingen wanneer Europese trend wordt meegenomen in het model.	34
Tabel 6: Geschatte onverklaarde variatie geassocieerd met telgebieden (ruimtelijke variatie). Waarden γ_1 geven voor elk telgebied de afwijking ten opzichte van de geschatte waarde ($\log Y$) op basis van enkel de verklarende variabelen ('fixed' component van het model). Telgebieden zijn gerangschikt van stroomopwaarts naar stroomafwaarts.	35
Tabel 7: Inschatting van de ruimtelijk en temporele variatie in aantallen vogels. De waarden geven de grootteorde (i.e. 10^x geeft de verhouding weer tussen hoogste en laagste aantallen binnen het 95% betrouwbaarheidsinterval) van de verschillen in waargenomen aantallen op ruimtelijke schaal (telgebieden), temporele schaal (winters). Zonder: verkennende inschatting van de variatie zonder habitatvariabelen in het model. Finaal: residuele (kan niet worden toegeschreven aan habitatvariabelen) ruimtelijke en temporele variatie in het finale model.	41
Tabel 8: voorspelde aantallen in functie van variatie in kenmerkwaarden. Laag = 2.5% percentiel van de kenmerkwaarden; Hoog = 97.5% percentiel van de kenmerkwaarden. Voor de overige kenmerken in de lineaire vergelijking wordt de waarde gefixeerd op de 50% percentiel. Gebiedsspecifieke correctiefactoren (ruimtelijke random variatie) worden hier niet in rekening gebracht.	42

1 Inleiding

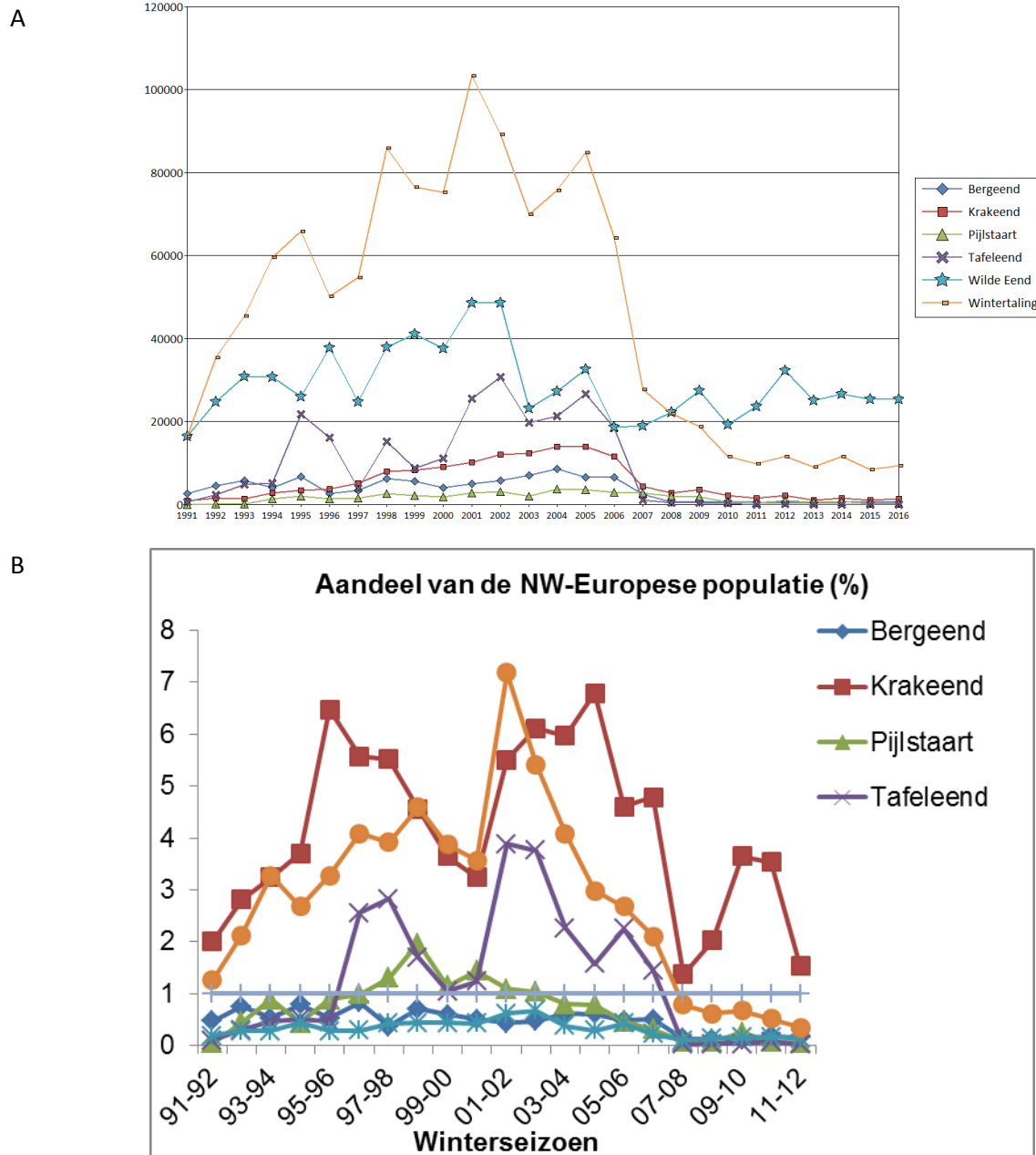
Het Integraal Plan van de Boven-Zeeschelde beschrijft een set van morfologische aanpassingen met als doel het systeem tegen 2050 te verbeteren in relatie tot scheepvaart, veiligheid, ecologie, onderhoud en bijkomende functies. Hiervoor zijn een aantal alternatieven uitgewerkt die ingrepen behelzen met meer of minder grote impact op de huidige morfologie. Binnen dit project worden een aantal numerieke en statistische modellen ontwikkeld of verbeterd, die gebruikt zullen worden om de effecten van voorgestelde morfologische ingrepen op hydrodynamiek, sediment transport, ecosysteem functioneren, waterkwaliteit, habitatkwaliteit en fauna en flora in te schatten ('Model instruments for the Integrated Plan Upper Seascheldt' (IMDC et al., 2015)) (Figuur 1-1) en zodoende het meest gewenste alternatief te bepalen. Het huidige document behandelt de opbouw van een statistisch model (INBO "BA – Bird Abundance" in Figuur 1-1) om de effecten van morfologische ingrepen op kenvogelsoorten te voorspellen.



Figuur 1-1. Sequentie van doorrekeningen en datastroom met modelinstrumenten binnen het Integraal plan Boven-Zeeschelde.

De Schelde is een belangrijke pleisterplaats voor (water)vogels en biedt een grote variëteit en oppervlakte aan rust- en voedselgebieden. Naast een aantal permanente bewoners van de slikken en schorren, worden deze ook intensief gebruikt door doortrekkers en overwinteraars, waaronder eendensoorten zoals wintertaling en bergeend (Figuur 1-2). Voor een aantal van deze eendensoorten maakten de populaties in de Zeeschelde tot een aantal jaar geleden één van de grootste, zo niet de grootste, fractie uit van de Europese populatie tijdens de winter (zie Figuur 1-2B). De slikken maken dan ook deel uit van de habitat- en vogelrichtlijnen. Veel vogelsoorten zijn afhankelijk van de rust (weinig verstoring) en het rijke voedselaanbod in de slikken. Dat voedselaanbod bestaat voor een groot deel uit macrozoöbenthos. Veranderingen

in beschikbaarheid van de slikken en naburige schorren (veranderingen in oppervlakte, getijkaracteristieken, ...) of de kwaliteit van slikken en water (zuurstof, saliniteit, ...) kunnen gevoelige verschuivingen teweeg brengen in de aantallen vogels, inclusief gevoelige soorten die als indicatoren fungeren voor de kwaliteit van het Schelde ecosysteem.



Figuur 1-2: Historische trends van de belangrijkste soorten eenden in de Zeeschelde (uitgedrukt als % van de Europese populatie).

De focus in de huidige studie ligt op wintertaling (*Anas crecca*) als indicatorsoort. Deze eendensoort komt in relatief grote aantallen voor en de Zeeschelde fungeert als een belangrijk overwinteringsgebied. Andere eendensoorten die voorkomen langs de Zeeschelde, zoals bijvoorbeeld tafeleend, pijlstaart of wilde eend, zijn minder geschikt omdat zij sinds 2007 in te lage aantallen voorkomen (tafeleend & pijlstaart) of de slikken minder gebruiken als foerageergebied en er zodoende minder afhankelijk van zijn (wilde eend). De krakeend zou een andere interessante soort kunnen zijn. Op basis van een verkennende analyse (Onkelinx et al. 2008) neemt deze soort echter een zeer gelijkaardige niche in als wintertaling. Omdat wintertaling de grootste absolute winteraantallen en de grootste daling relatief tot de Europese populatie vertoont (Figuur 1-2), is er beslist om enkel de wintertaling te weerhouden binnen deze studie. In een preliminaire analyse werd nagegaan of ook de bergeend kan fungeren als modelsoort. Zoals voor een aantal andere soorten komt deze soort recent echter in te lage aantallen voor in de Boven Zeeschelde om verder mee te nemen.

Om een onderbouwde inschatting te maken van de mogelijke effecten van geplande aanpassingen aan de morfologie van de Schelde op de vogels in het systeem, wordt een modelinstrument ontwikkeld dat op basis van correlatieve modellering (lineaire regressie) de relatie legt tussen relevante habitatkenmerken en de aantallen overwinterende eenden in de Boven-Zeeschelde. De nadruk ligt hierbij op het optimaliseren van voorspellingen, wat zich reflecteert in de gebruikte technieken; gegevens gebruikt voor kalibratie van de parameters en voor selectie/validatie van de modellen worden zoveel mogelijk gescheiden (zie paragraaf 2.5). Voor het opstellen van de modellen wordt gebruik gemaakt van gegevens over vogelaantallen en potentieel relevante habitatkenmerken voor de periode 2007-2012. Het zou interessant zijn om ook kennis omtrent het ecosysteem en het voedselweb van de Zeeschelde te betrekken in de modellering. Deze kennis en vooral de link met de abundantie van overwinterende eenden is echter nog schaars. Ondanks sterke aanwijzingen dat er in het verleden belangrijke veranderingen in de aantallen overwinterende eenden zijn opgetreden in respons op veranderingen in de waterkwaliteit, wordt dit dus niet verder meegenomen.

De keuze voor het gebruik van correlatieve modellering wordt eveneens opgelegd door het feit dat er nog onvoldoende kennis is van het systeem en de causale verbanden om een gefundeerd causaal model op te stellen. Het achterhalen van deze causale relaties, evenals het belang van de bentisch-pelagische koppeling en de link met de rest van het voedselweb in de Zeeschelde is het onderwerp van lopend onderzoek.

2 Materiaal en methode

2.1 Ecologisch achtergrond

De meeste eendensoorten zijn in de Zeeschelde afhankelijk van de zachte substraten op de slikken om te foerageren (Van Ryckegem et al., 2006). Fytobenthos en vooral benthische macro-invertebraten zijn in deze gebieden een belangrijke voedselbron voor foeragerende eenden in de Zeeschelde (Van Ryckegem et al., 2006). Een grote oppervlakte aan slikken met zachte bodems is in dit opzicht dus zeer belangrijk. Bovendien is het aangewezen dat deze slikken niet onderhevig zijn aan al te zware verstoring door onder meer menselijk toedoen (bijvoorbeeld recreatie langsheen de oever of scheepvaart op de Schelde). Ook de kwaliteit van de slikken is belangrijk. Niet alleen dient er voldoende voedselaanbod te zijn, maar het voedsel moet ook in voldoende mate toegankelijk zijn. De toegankelijkheid van voedsel op een slik wordt onder meer bepaald door het gelijkmatig droogvallen van slik gedurende een getijcyclus. Grondeleenden zoals wintertaling, kraakeend en bergeend zeven het slib door lamellen in hun bek en het gelijkmatig vrijkomen van slik zorgt ervoor dat er steeds slib aanwezig is met een goede waterbalans om te zeven (meestal op of net boven de waterlijn, Van Ryckegem et al., 2006). Het voedselaanbod is mede afhankelijk van de dynamiek van het systeem. Gebieden met hoge stroomsnelheden (hoogdynamisch) bij vloed vertonen hoger zandpercentage en zijn vermoedelijk qua foerageerlocatie minder interessant omdat macrobenthos (een belangrijke voedselbron) er minder aanwezig is dan in laagdynamische gebieden (Van Braeckel et al., 2018).

Ook de waterkwaliteit en het ecosysteem-functioneren beïnvloeden het voedselaanbod op de slikken. Het waterzuiveringsprogramma van de bovenloop van de Schelde en zijrivieren loopt inderdaad samen met opmerkelijke veranderingen in het ecologisch functioneren binnen de Zeeschelde. De aanvang van waterbehandeling in de jaren negentig zorgde voor een afbouw van hypereutrofe en anoxische condities (Cox et al., 2009), en resulteerde in de opbouw van een gemeenschap met een hoge densiteit aan benthische invertebraten. Verdere verbetering van de waterkwaliteit met de inwerkingtreding van de waterzuiveringsstations van Brussel-Zuid en Brussel-Noord op de Zenne, respectievelijk in 2000 en 2007, zorgde voor verdere verbetering van de waterkwaliteit met hogere zuurstofwaarden en het verdwijnen van de anoxische zone rond de Rupelmonding. Deze veranderingen in waterkwaliteit lopen parallel met een drastische daling in de densiteit van benthische macro-invertebraten en watervogels en het opnieuw verschijnen van vis en hyperbenthische invertebraten in de Zeeschelde. Zoals eerder vermeld worden deze historische tendensen en de mogelijke link met het ecosysteem-functioneren niet verder meegenomen in de huidige analyse. Er wordt eerder gefocust op de huidige toestand (laatste decennium) waarin de aantallen overwinterende eenden vrij stabiel zijn. Preliminair analyses hebben aangetoond dat er voor deze periode er geen sterk verband is tussen de aantallen overwinterende eenden en ecosysteemvariabelen.

De habitatkenmerken die tijdens de modelopbouw worden beschouwd als mogelijke verklarende variabelen voor de aantallen vogels (zie paragraaf 2.2) zijn zo gekozen dat ze representatief zijn voor de verschillende hierboven besproken aspecten die de kwaliteit van een slik als foerageerhabitat voor watervogels bepalen. Dit betekent niet noodzakelijk dat al deze variabelen ook in het finale model een rol zullen spelen. Een deel van de aangehaalde relaties zijn hypothetisch en nog niet sterk genoeg onderbouwd via waarnemingen of

experimenten. Vaak zijn zij het onderwerp van lopend onderzoek. Aan de hand van modelselectie, welke een onderdeel is van de kalibratie, wordt de verzameling van mogelijke verklarende variabelen gereduceerd tot een beperkte set van variabelen die aanwijsbaar bijdragen tot onderbouwde voorspellingen van het aantal vogels (zie paragraaf 2.5).

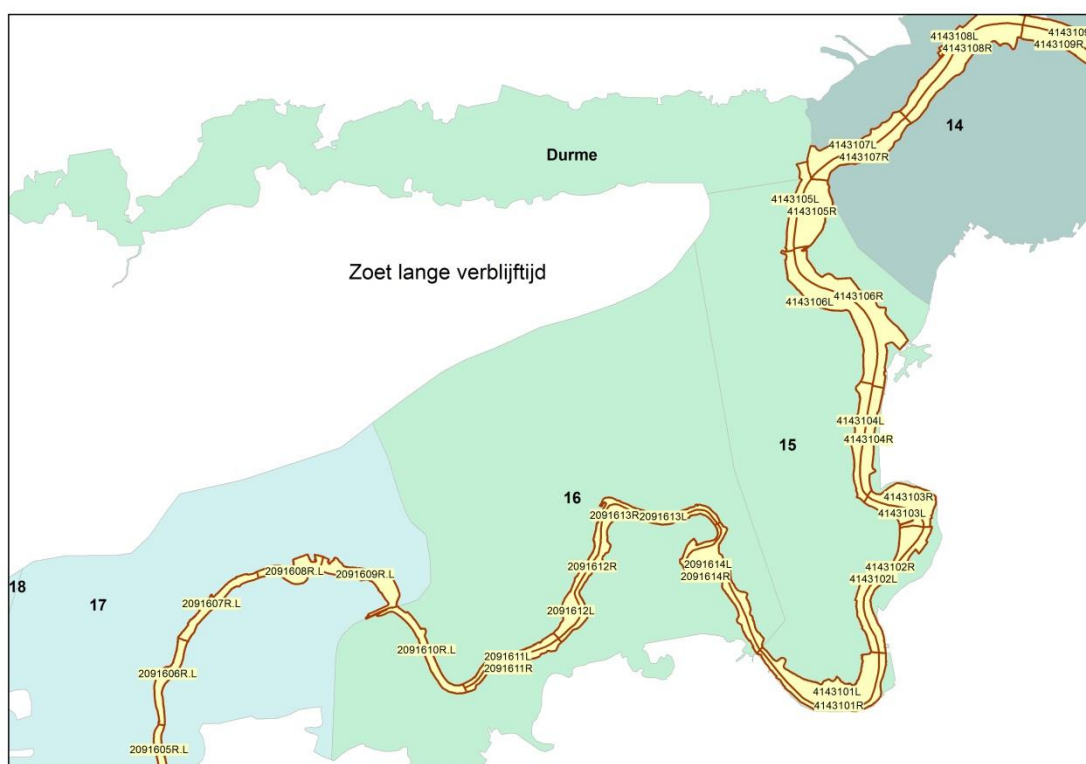
2.2 Gegevens

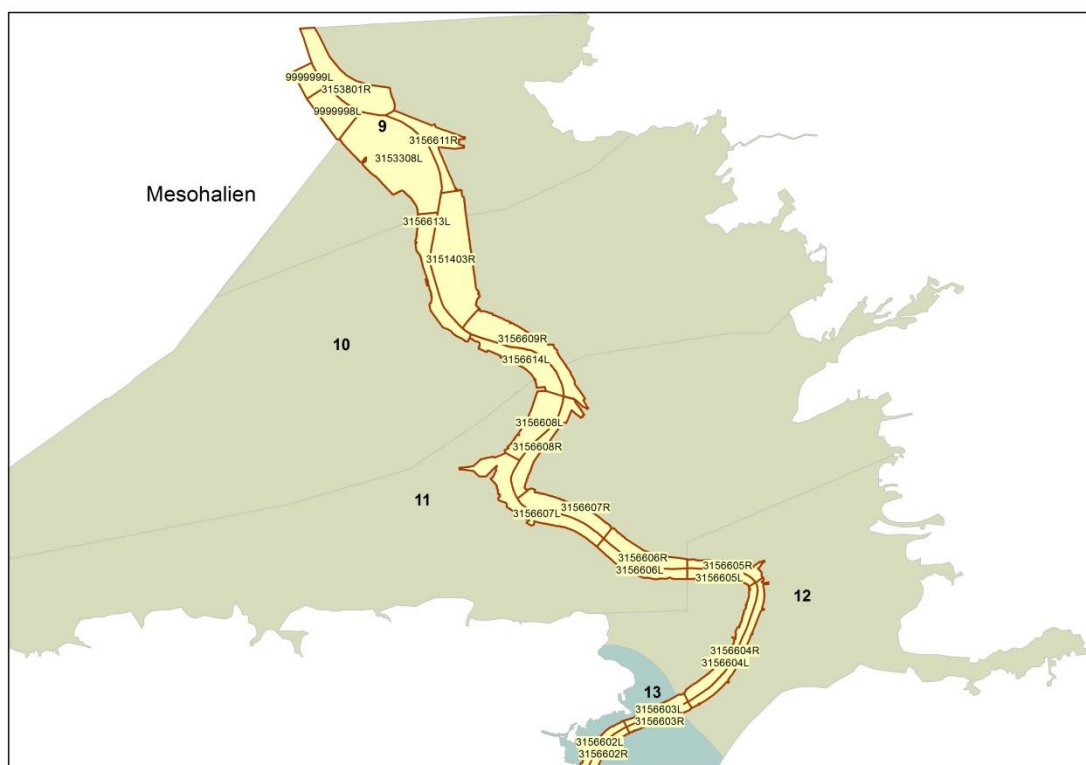
Naast de gegevens voor de vogeltellingen worden in eerste instantie een aantal variabelen geselecteerd die *a priori* een mogelijk (causaal) verband vertonen met de aantallen eenden en waarvoor gegevens beschikbaar zijn in voldoende ruimtelijke en temporele resolutie. Deze variabelen worden hieronder beschreven. Gedurende de verkennende analyses en modelselectie wordt een verdere selectie gemaakt van variabelen waarvoor blijkt dat ze een onafhankelijke bijdrage leveren tot het voorspellen van de vogelaantallen. Er wordt voor de modelselectie enkel gewerkt met gegevens voor de Boven-Zeeschelde.

2.2.1 Vogeltellingen

De telgegevens van wintertaling zijn gehaald uit de INBO-watervogeldatabank <http://watervogels.inbo.be/>. Sinds oktober 1991 tellen medewerkers van het INBO maandelijks het aantal watervogels langs de Zeeschelde vanaf de Belgisch-Nederlandse grens tot Gent. De tellingen gebeuren vanaf een boot en rond laag tij, wanneer de slikken maximaal vrij liggen. Omdat het niet haalbaar is om het volledige onderzoeksgebied grondig te tellen tijdens de periode van laag tij, worden de telling gespreid over drie dagen. De teldagen worden steeds gegroepeerd in het midden van de maand.

De Zeeschelde is opgedeeld in een aantal vogeltelgebieden met een aslengte van 1 tot 3.6 km (ongeveer 2km gemiddelde lengte) (Figuur 2-1), waarbinnen de vogels per soort geteld worden. Stroomafwaarts van Dendermonde zijn er afzonderlijke telgebieden voor linker- en rechteroever. De grens tussen de telgebieden voor linker- en rechteroever ligt hierbij op de thalweg. Stroomopwaarts van Dendermonde zijn linker- en rechteroever telkens samengevoegd binnen hetzelfde telgebied. De resultaten van deze tellingen worden jaarlijks gerapporteerd (bv. Van Ryckegem, 2015). In het huidige project ligt de focus op de aantallen van overwinterende eenden die de slikken langsheen de Boven-Zeeschelde gebruiken om te foerageren. Meer bepaald worden de tellingen voor wintertaling, de meest abundante soort, gebruikt uit de wintermaanden (oktober tot en met maart) tijdens de periode 2007/2008 tot 2012/2013. Een winterjaar_x loopt van oktober jaar_x tot maart jaar_{x+1}.





Figuur 2-1: Overzichtskaart per saliniteitszone van de vogeltelgebieden en OMES-segmenten langs de Zeeschelde. De rood omlijnde gebieden zijn de vogeltelgebieden, de grotere genummerde zones de OMES-segmenten.

2.2.2 Europese trend

De aantallen wintertaling zijn niet enkel onderworpen aan ruimtelijke en temporele variatie binnen de Zeeschelde maar zijn ook afhankelijk van tendensen op grotere ruimtelijke schaal. Om na te gaan in hoeverre de waargenomen patronen in de Zeeschelde een weerspiegeling zijn van de dynamieken op Europees niveau worden Europese trends in rekening gebracht als potentiële verklarende variabele. Gegevens voor Europese trends werden bekomen uit de 6^{de} editie van de 'Conservation Status Review'¹. Zij weerspiegelen het relatieve verschil in aantallen tussen jaren met als ijkpunt de aantallen in 2003.

2.2.3 Strengheid van winter

Ook de strengheid van winters kan potentieel een effect hebben op het aantal foeragerende vogels in de Zeeschelde. Het vorstgetal V van IJnsen (1991) wordt hier gebruikt als maat voor de strengheid van de winter:

$$V = \frac{v^2}{363} + \frac{2y}{3} + \frac{10z}{9}$$

waarbij V : het aantal vorstdagen (dagen met minimumtemperatuur $< 0^{\circ}\text{C}$); y : het aantal ijsdagen (dagen met maximumtemperatuur $< 0^{\circ}\text{C}$) en z : het aantal zeer koude dagen (dagen met minimumtemperatuur $< -10^{\circ}\text{C}$). De temperatuurgegevens zijn afkomstig van het KMI².

Voor alle winters wordt het vorstgetal meegenomen als mogelijke verklarende variabele tijdens modelselectie.

2.2.4 Habitatkenmerken

Voor alle vogeltelgebieden (Figuur 2-1) werden de habitatkenmerken geëxtraheerd uit gebiedsdekkende databronnen verzameld binnen het INBO-MONEOS-programma dat jaarlijks gerapporteerd wordt. Uit deze data wordt een initiële selectie gemaakt (Tabel 1) die tijdens de verkennende analyse nog verder wordt verfijnd om multicollineariteit tussen variabelen te vermijden (zie paragraaf 20).

2.2.4.1 Oppervlakte en breedte van de ecotopen

Wintertaling foerageert op slikken gekenmerkt door zacht substraat (verder slik zacht genoemd). Naarmate de oppervlakte van dit ecotoop toeneemt is er dus meer ruimte en meer voedsel om op te foerageren. Aangezien er in de analyses wordt gestandaardiseerd naar

¹ www.unep-aewa.org

² <http://www.meteo.be/meteo/view/nl/1317239-Voorbije+jaren.html>

eenheid afstand langsheen de lengteas van de rivier (aslengte; zie paragraaf 2.5.1.2), wordt er echter gewerkt met breedte van de slikken. Voor eenzelfde lengte langsheen de rivier bepaalt de breedte immers de totale oppervlakte. Ook de breedte van het ondiep subtidaal en van de schorren kunnen eventueel van belang zijn. Deze dienen als maat voor de kans op verstoring van foeragerende vogels op de slikken. Indien het ondiep subtidaal smal is, is er een grotere kans op verstoring door bijvoorbeeld scheepvaart op de rivier. Indien het schor smal is, is er meer kans op verstoring van de rustende en foeragerende vogels op de slikken door recreanten aanwezig op de dijk. Gemiddelde breedte per ecotoop en vogelgebied wordt berekend aan de hand van de oppervlakte en aslengte als:

$$breedte_{ecotoop} = \frac{oppervlakte_{ecotoop}}{aslengte_{telgebied}}$$

Voor de gebieden stroomopwaarts van Dendermonde, waar linker- en rechteroever samen worden geteld als één gebied, wordt de breedte van het schor en ondiep subtidaal nog eens gedeeld door twee. Dit reflecteert beter de bufferende werking langs beide oevers. Langs elke oever is het ondiep subtidaal of schor immers slechts half zo breed als de totale breedte voor dat telgebied. De oppervlakte van de ecotopen wordt berekend als het aantal m² dat aan elk ecotoop wordt toegekend in elk telgebied. De oppervlakte van de habitats wordt geëxtraheerd uit de ecotopenkaart van 2010 (Van Braeckel, 2013). Bij de slikken wordt hierbij geen onderscheid gemaakt tussen laag, middelhoog en hoog slik. Deze categorieën worden samengebracht tot één oppervlakte (en één breedte). Indien er in de periode 2007-2012 door bijkomende ontpolderingen of in werking treden van gereduceerde gecontroleerde getijzones (GGG) gebieden zijn bijgekomen, worden deze zones ook in rekening gebracht vanaf het jaar dat de uitbreiding tot stand kwam.

2.2.4.2 Helling

De helling van de slikken is een proxy voor de geschiktheid van het slik als foerageergebied. Hoe steiler de helling is, hoe minder stabiel en duurzaam de oever is als foerageer- en rustgebied. Bovendien bepaalt de helling ook hoe breed op elk moment de zone rond de waterlijn is waar optimaal kan gefoerageerd worden. De eenden foerageren immers bij voorkeur aan of in de buurt van de waterlijn (Van Ryckegem et al., 2006). De zone waarin optimaal kan gefoerageerd worden door eenden is smaller naarmate de helling steiler is. De helling van het zacht slik wordt voor elk telgebied berekend aan de hand van het gemiddelde per transect voor laag en middelhoog slik op basis van gecombineerde lidar en bathymetrische hoogtemetingen uit de periode 2009 - 2011.

2.2.4.3 Spreiding droogvalduur

De spreiding in droogvalduur (*SpD*) geeft een maat voor de beschikbaarheid van de slikken om te foerageren. Indien er een brede spreiding is in de droogvalduren binnen een slik, betekent dit dat er een groot deel van de tijd gedurende een getijdencyclus oppervlakte beschikbaar is die net is drooggevallen en optimaal is om te foerageren in de waterlijn. Indien er slechts een kleine spreiding is in droogvalduur, dan komt al het slik op ongeveer hetzelfde moment droog

te liggen en is er maar een beperkte periode waarin er optimaal kan gevoerageerd worden. Op basis van de droogvalduurkaart 2010 wordt de spreiding in droogvalduur over de volledige breedte van het slik voor elk telgebied als volgt berekend:

$$SpD = \frac{1}{\sum p_i^2}$$

waarbij p_i de proportie is van de i-de droogvalduur klasse. Er wordt gewerkt met 20 droogvalduurklassen (van 0 tot 100% in stappen van 5%). SpD houdt zowel rekening met de range van droogvalduurklassen binnen een slik als met de gelijkmatige verdeling van de droogvalduurklassen.

2.2.4.4 Maximale vloodsnelheid

De maximale vloodsnelheid van het subtidaal gebied wordt gebruikt als proxy voor de hydrodynamiek van de naburige slikken. Hoogdynamische zones (hoge stroomsnelheid) worden verondersteld minder interessant te zijn voor watervogels om te foerageren omdat er een minder abundante benthosgemeenschap aanwezig is dan in laagdynamische zones (lage stroomsnelheid, Van Braeckel et al., 2018). Aangezien de meeste zones vloedgedomineerd zijn, is de maximale vloodsnelheid overwegend de hoogste stroomsnelheid op het slik in de Schelde. De maximale vloodsnelheid (MV) wordt afgeleid uit het Nevla model (Maximova et al., 2013) met een gestructureerd modelrooster en wordt berekend als het gemiddelde in het subtidaal per telgebied. In toekomstige scenarioberekeningen binnen het integraal plan Boven-Zeeschelde zullen wel rechtstreeks stroomsnelheden op het slik gebruikt worden, afkomstig uit het SCALDIS waar slikken en ondiepe waterzones beter in vervat zitten (Smolders et al., 2015). Tijdens modelopbouw en kalibratie werden geen gegevens afkomstig van het SCALDIS model gebruikt, omdat die op dat ogenblik nog niet beschikbaar waren.

2.2.4.5 Luwte Index

Gebieden waar veel bezinking mogelijk is (luwe gebieden) zijn potentieel interessant als foerageergebieden, omdat organisch materiaal er cumuleert en kan dienen als voedselbron voor de benthische gemeenschap. Op basis van de maximale vloeddominantie VD (maximale vloodsnelheid / maximale ebsnelheid) en de gemiddelde maximale stroomsnelheid MS (gemiddelde van de maximale vloodsnelheid en maximale ebsnelheid) wordt er een maat voor de dynamiek berekend die voor elk gebied de globale luwte weergeeft:

$$VD \geq 1 \rightarrow LI = \frac{1}{VD \cdot MS}$$

$$VD < 1 \rightarrow LI = \frac{VD}{MS}$$

Deze luwte index LI neemt af naarmate de ongelijkheid in stroomsnelheid tussen eb en vloed toeneemt en naarmate de gemiddelde maximale stroomsnelheid toeneemt. Deze index is dus lager (minder luw) naarmate de verschillen tussen eb- en vloodsnelheden groter zijn en de gemiddelde maximale stroomsnelheid wordt gedomineerd door extreme stroomsnelheden

tijdens eb of vloed. De luwte index is afgeleid uit modelresultaten van het NEVLA model (Maximova et al. 2013) voor het subtidaal per telgebied.

	Schor	Slik Zacht substr	Ondiep subtidaal	Algemeen
Breedte (m)	X	X	X	
Helling (%)		X		
Spreiding droogvalduur		X		
Maximale vloodsnelheid (m/s)				X
Luwte Index				X

Tabel 1: habitatkenmerken voor de vogeltelgebieden langs de Zeeschelde, waaruit tijdens de verkennende analyse een verdere selectie wordt gemaakt.

2.3 Methodiek

2.3.1 Generalized linear mixed models

De kern van de hier uitgevoerde analyses betreft het toepassen van Generalized Linear Mixed Models (GLMM). De relatie tussen verklarende variabelen (in dit geval habitat en ecosysteemvariabelen) en afhankelijke variabele (in dit geval vogelaantallen) wordt hierbij beschreven aan de hand van een lineaire vergelijking, met intercept en richtingscoëfficiënt(en). Deze vergelijking geeft eenvoudig uitgedrukt een vaste verhouding weer tussen verklarende variabelen en de respons.

Er wordt bij een GLMM rekening gehouden met het feit dat waarden van de afhankelijke variabele niet noodzakelijk voldoen aan een normaalverdeling, maar dat de spreiding rond de verwachte (voorspelde) waarden een andere verdeling kan hebben. Aangezien hier met tellingen van aantal vogels wordt gewerkt, wordt de verdeling *a priori* verondersteld Poisson te zijn. In de praktijk wijkt de verdeling bij biologische gegevens hier vaak enigszins van af en wordt er uitgegaan van een negatief binomiale verdeling, die een grotere spreiding van de waarden rond het gemiddelde toelaat (i.e. een grotere spreiding van individuele waarnemingen (tellingen) rond het verwachte/voorspelde gemiddelde). Hierbij spreekt men van overdispersie ten opzichte van de *a priori* verwachte Poisson verdeling.

Vaak is er ook een 'vertaling' nodig tussen de regressie vergelijking en de afhankelijke variabele; de zogenaamde link functie. Deze link functie garandeert dat de afhankelijke variabele kan uitgedrukt worden binnen een theoretisch verloop van $-\infty$ tot $+\infty$, zoals vereist voor een lineaire regressie. Aantallen, bijvoorbeeld, hebben een theoretisch verloop van 0 tot $+\infty$. Door een logaritmische transformatie (log link functie) toe te passen kan dit omgezet worden in een theoretisch verloop van $-\infty$ tot $+\infty$.

In een GLMM kunnen er ook variabelen toegevoegd worden die de ruimtelijke en temporele afhankelijkheidsstructuur in de gegevens beschrijven. Ruimtelijke afhankelijkheid wordt bijvoorbeeld veroorzaakt door het feit dat er meerdere tellingen binnen hetzelfde gebied gebeuren; temporele afhankelijkheid heeft te maken met het feit dat meerdere tellingen (voor verschillende gebieden) gebeuren binnen dezelfde maand en/of hetzelfde jaar. Deze afhankelijkheid moet in rekening worden gebracht om jaar (of maand) en gebiedseffecten die niet gerelateerd zijn tot de variabelen waarin we geïnteresseerd zijn (habitat en ecosysteemvariabelen) weg te filteren. Voor deze ruimtelijke en temporele variabelen worden eveneens parameters geschat (de zogenaamde random variabelen in het model) die de spreiding in de gegevens (vogelaantallen) capteren die wordt veroorzaakt door deze factoren.

2.3.2 Andere statistische technieken

In een initiële fase van de methodologie ontwikkeling werd er nagegaan of een aantal andere technieken een alternatief kunnen vormen voor GLMM. In eerste instantie werd er aandacht besteed aan N-mixture models (Kéry & Royle, 2015). Dit is in feite een speciaal geval van Generalized Linear (Mixed) Models gerelateerd aan modellen met zero-inflation, waarbij de regressie bestaat uit twee gekoppelde vergelijkingen. Eén vergelijking beschrijft de kans om al dan niet vogels waar te nemen in een gebied (binomiale respons) en een tweede vergelijking beschrijft het aantal waargenomen vogels. Dit kan lijden tot preciezere voorspellingen indien de waargenomen aantallen laag zijn. Er werd ook onderzocht of er kan gewerkt worden met Boosted Regression Trees (Elith et al. 2008). Boosted Regression Trees (BRTs) combineren aspecten van regressie technieken en van machine learning. Dit is een veelbelovende techniek voor het maken van goede voorspellingen die minder last hebben van overparameterisatie. Bovendien kunnen ook gemakkelijk niet-lineaire responsen in rekening worden gebracht zonder nood aan transformaties of polynomen.

2.4 Verkennende analyse

2.4.1 Vogeltellingen

De maandelijkse tellingen voor wintertaling worden uitgezet per saliniteitszone en per OMES segment om een eerste idee te krijgen van de temporele en ruimtelijke patronen in de aantallen. Vervolgens wordt een eerste statistische analyse gedaan waarin enkel rekening gehouden wordt met factoren gelinkt aan de ruimtelijke en temporele afhankelijkheidsstructuur binnen de gegevens. Meer bepaald wordt er een Generalized Linear Mixed Model (GLMM) uitgevoerd op de aantallen, waarbij telgebied (ruimtelijke afhankelijkheid) en winter (jaartal) en wintermaand (temporele afhankelijkheid) als random variabelen worden ingevoerd (voor meer details, zie 2.5). Op basis van deze verkennende analyse kan een eerste inschatting gemaakt worden van het belang van deze factoren, zonder rekening te houden met potentiële verklarende variabelen gerelateerd tot habitat kenmerken.

Aan de hand van deze initiële analyse zal een verdere afbakening gemaakt worden van de wintermaanden die in de uiteindelijke analyse gebruikt zullen worden. Hiervoor wordt een selectie gemaakt van de maanden met de hoogste aantallen die gemiddeld genomen het minst

van elkaar verschillen (i.e. de geschatte random effecten voor de beoogde maanden zijn gelijkaardig). Als link functie wordt in de analyse de log link gebruikt (de voorspelde waarden uit het model zijn log getransformeerde verwachte aantallen). In eerste instantie wordt een Poisson respons verdeling verondersteld. Indien deze verdeling de variatie niet correct inschat (overdispersie van de respons variatie) wordt ook een analyse uitgevoerd met een negatief binomiale respons verdeling (zie paragraaf 2.3.1).

2.4.2 Lineair verband tussen afhankelijke en onafhankelijke variabelen

De voorgestelde analyse aan de hand van GLMM veronderstelt dat de verbanden tussen de afhankelijke variabele (aantal vogels) en de verklarende variabelen (habitatkenmerken en ecosysteemvariabelen), al dan niet getransformeerd, lineair zijn. Zowel voor habitatkenmerken als ecosysteemvariabelen wordt via een aantal verkennende grafieken nagegaan of de variabelen dienen getransformeerd te worden om de lineariteit van de relatie te maximaliseren. Hierbij wordt telkens de relatie onderzocht ten opzichte van logaritmisch getransformeerde aantallen van vogels, aangezien deze transformatie ook wordt toegepast in de uiteindelijke modelselectie (aan de hand van GLMM; zie paragraaf 2.3.1).

2.4.3 Ruimtelijke patronen in de habitatkenmerken

Aan de hand van een aantal verkennende figuren wordt de ruimtelijke variatie in de habitatkenmerken weergegeven. Voor ruimtelijke variatie wordt er gekeken op twee hiërarchische niveaus: telgebieden (1) binnen OMES segmenten (2).

2.4.4 Multicollineariteit

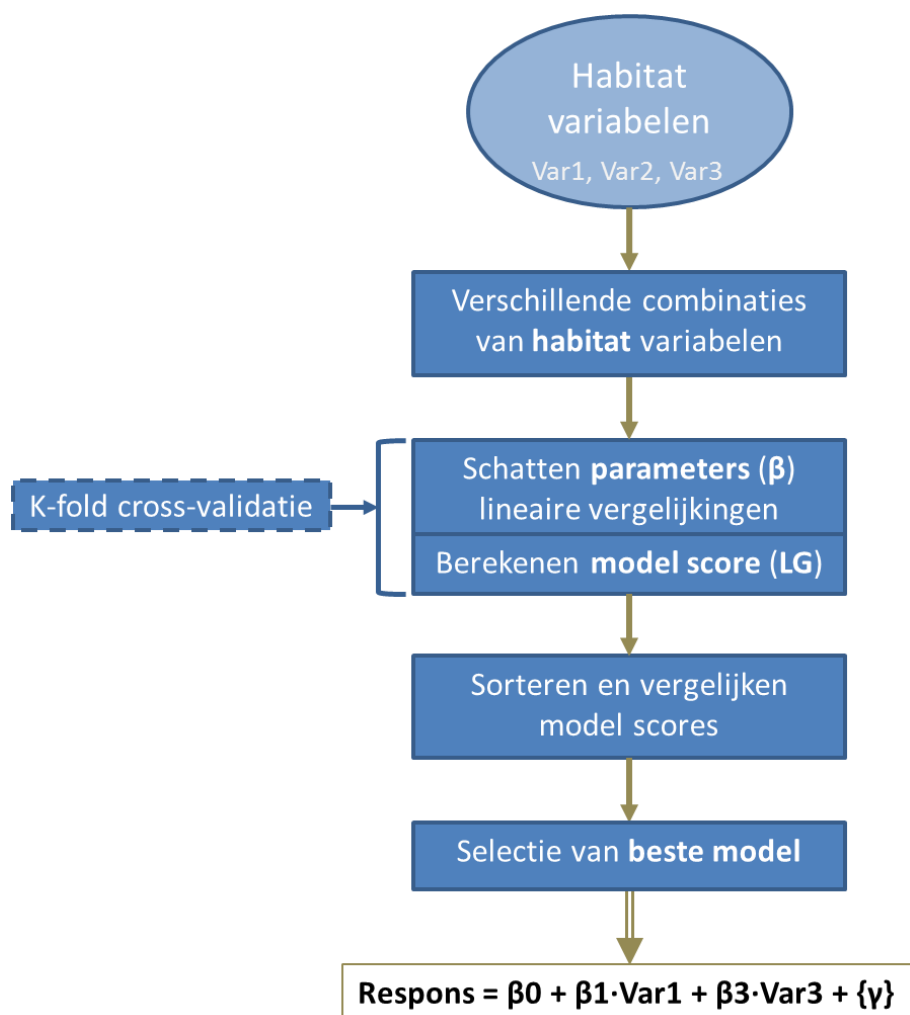
Indien er sterke correlaties bestaan tussen de verklarende variabelen onderling (multicollineariteit), kan dit problemen geven voor de analyse, aangezien er geen onderscheid kan gemaakt worden tussen het belang van de variabelen die sterk zijn gecorreleerd (noch door de analyse methode noch voor de interpretatie). Voorafgaand aan de analyse wordt daarom eerst de correlatiestructuur tussen de verklarende variabelen bekeken, en wordt er een verdere selectie gemaakt van variabelen die minimaal zijn gecorreleerd. Dit gebeurt aan de hand van paarsgewijze correlaties en aan de hand van Variance Inflation Factors (VIF). Het voordeel van VIF ten opzichte van paarsgewijze correlaties is dat VIF voor elke variabele een unieke maat geven voor de globale correlatie met alle andere variabelen. Op basis van VIF kan het aantal variabelen stapsgewijs worden verkleind tot alle VIF onder een aanvaardbare waarde liggen. Uit de praktijk blijkt dat voor ecologische studies de bovengrens voor VIF waarden vrij streng moet gehouden worden om betrouwbare analyses te kunnen uitvoeren. Er wordt hier een bovengrens voor de VIF van 2-2.5 gehanteerd.

2.5 Analyse van aantallen

2.5.1 Kalibratie

2.5.1.1 Globaal overzicht

In Figuur 2-2 wordt een overzicht getoond van de gevolgde methodiek. In de volgende paragrafen wordt dan verder ingegaan op de verschillende componenten van de methodiek.



Figuur 2-2: Globaal overzicht van de gevolgde methode voor het opstellen van een voorspellend model voor vogelaantallen in functie van habitat- en ecosysteemkenmerken.

De statistische analyse is erop gericht om een zo goed mogelijk voorspellend model te vinden van de aantallen vogels op de slikken op basis van habitatkarakteristieken, om zo toekomstige beheersscenario's tegen elkaar af te wegen. Voor het opstellen van het voorspellend model wordt een procedure van modelselectie gehanteerd (paragraaf 2.5.1.5). Hierbij worden modellen met verschillende combinaties van potentiële verklarende variabelen met elkaar vergeleken. Op basis van die vergelijking kan dan een model worden opgesteld dat enkel de relevante verklarende variabelen bevat om voorspellingen te doen in verband met de aantallen watervogels langsheen de Zeeschelde. De modellen die met elkaar worden

vergeleken, bevatten een verschillende selectie uit de habitatkarakteristieken. Voor elke combinatie van verklarende variabelen worden de parameters geschat van een Generalized Linear Mixed Model (GLMM; paragraaf 2.3.1 en 2.5.1.2). Deze parameters worden gebruikt om voorspellingen te doen over het aantal vogels in functie van de betrokken verklarende variabelen. De vergelijking tussen potentiële modellen gebeurt op basis van een logaritmische score (LG score; paragraaf 2.5.1.3) die een maat is voor de afwijking tussen waargenomen aantallen en voorspelde aantallen. Op basis van deze score worden potentiële modellen gerangschikt, waarna de beste modellen worden geselecteerd (paragraaf 2.5.1.6). Aan de hand van de x beste modellen wordt dan een finaal gemiddeld model samengesteld om zo goed mogelijke voorspellingen te doen voor toekomstige situaties (paragraaf 2.5.1.6).

Procedures voor modelselectie zoals hier gehanteerd, kunnen gevoelig zijn voor overparameterisatie, waarbij het model dat als beste naar voor komt complexer is dan nodig en een aantal verklarende variabelen bevat die niet veel bijdragen tot het voorspellend vermogen. Daarom worden een aantal stappen ondernomen om de kans op overparameterisatie te minimaliseren. Overparameterisatie kan vooral optreden als dezelfde gegevens gebruikt worden zowel voor modelschatting (parameterschatting) als voor het scoren van het model. Daarom wordt de score voor elk model berekend via cross-validatie, waarbij gegevens voor het schatten van de parameters en voor het berekenen van de score worden gescheiden (paragraaf 2.5.1.4). Bovendien wordt er gewerkt naar een finaal model dat een weerspiegeling is van de meest voorkomende variabelen in de modellen met de beste score (paragraaf 2.5.1.6). Variabelen die zelden bijdragen tot een goede score worden dus niet geselecteerd.

Voor de kalibratie worden niet alle gegevens gebruikt; een deel wordt opzij gehouden voor externe validatie van het finale model dat uit de modelselectie naar voor treedt. De dataset voor kalibratie is samengesteld uit de winterjaren 2008-2011. De externe validatieset bestaat uit de winterjaren 2007 en 2012. Op die manier zijn zowel de kalibratieset als de validatieset representatief voor de gehele meetperiode.

De volledige analyse wordt uitgevoerd in de statistische omgeving R (2016).

2.5.1.2 Opbouw van de GLMMs

Het effect van habitatkarakteristieken op het aantal waargenomen vogels wordt geanalyseerd aan de hand van Generalized Linear Mixed Models (GLMM). Er wordt dus een lineair verband verondersteld tussen de verklarende variabelen en de aantallen vogels. De geschatte parameters (regressiecoëfficiënten) uit het model kunnen vervolgens gebruikt worden om voorspellingen te doen van aantallen in functie van de betrokken verklarende variabelen. *A priori* worden in de GLMM analyses telgebied, winterjaar en wintermaand gebruikt als structurele random variabelen om rekening te houden met de ruimtelijke en temporele afhankelijkheidsstructuur van de gegevens. Hierbij wordt maand niet hiërarchisch genest in jaar, maar wordt elke combinatie winterjaar – wintermaand als uniek beschouwd. Om verschillen in de lengte van telgebieden op te vangen wordt in elke analyse de aslengte (log getransformeerd) van de telgebieden langsheen de rivier in rekening gebracht als offset. Dit betekent dat er *a priori* wordt verondersteld dat bij een verdubbeling van de aslengte ook het

aantal getelde vogels verdubbelt. Dit is een correctie voor verschillen in lengte tussen de vogeltelgebieden en komt erop neer dat een richtingscoëfficiënt (β) van de lineaire relatie wordt opgelegd gelijk aan 1 tussen de lengte van het telgebied (aslengte) en het aantal waargenomen vogels. Analyse en resultaten worden dus geïnterpreteerd per eenheid van afstand langsheen de lengteas van de rivier. De verdere keuze van verklarende variabelen gerelateerd tot habitatkarakteristieken hangt af van de verkennende analyses en het potentieel optreden van multicollineariteit. Als link functie voor aantallen in de GLMM wordt de log-link functie gebruikt (de voorspelde waarden uit het model zijn log getransformeerde aantallen). Uit verkennende analyses bleek dat een model met een Poisson responsdistributie de respons variatie niet correct inschat (overdispersie van de responsvariatie). Daarom wordt voor de eigenlijke analyses een negatief binomiale responsvariatie verondersteld. De GLMMs worden berekend aan de hand van de module lmer (Bates et al., 2015)) in de statistische omgeving R (2016).

2.5.1.3 Model score

Het voorspellend vermogen van verschillende modellen wordt met elkaar vergeleken aan de hand van een model score die aangeeft hoe goed het model aansluit bij de waargenomen aantallen. Als model score wordt de logaritmische score (LG) gebruikt. De logaritmische score is een ‘*strictly proper scoring rule*’ (zie Gneiting & Raftery, 2007) en wordt voor elk model m berekend als:

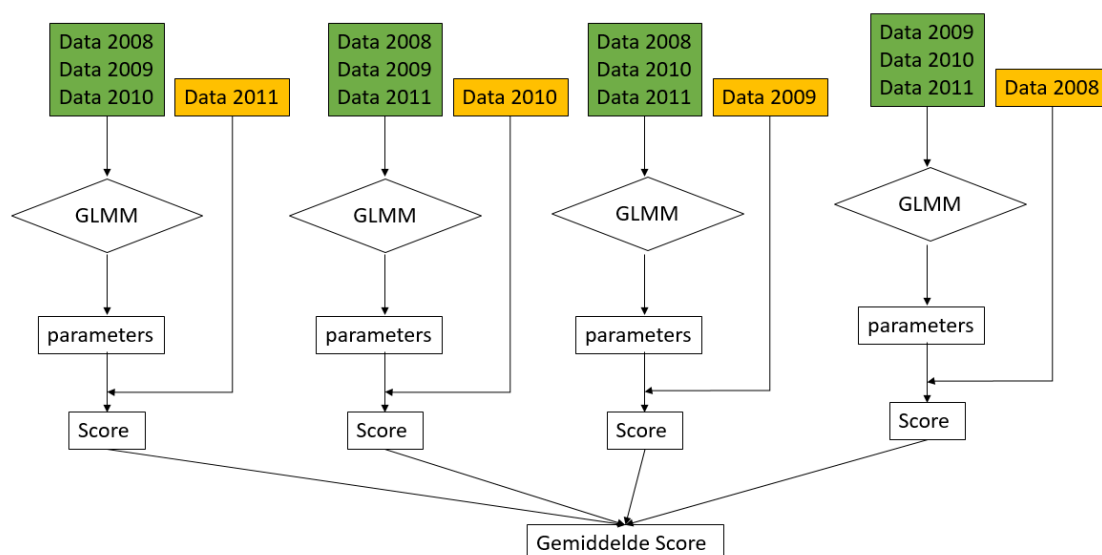
$$LG_m = -gem[\ln(p_m(x))]$$

waarbij $p(x)$ voor elke observatie x de kans weergeeft dat deze voorkomt, gegeven de verwachtingen gebaseerd op model m . De score is een maat voor de afwijking tussen waargenomen aantallen en voorspelde aantallen. Hoe lager deze score hoe beter de voorspellingen aansluiten bij de waargenomen aantallen. In het geval van een negatief binomiale respons distributie wordt $p(x)$ berekend op basis van de *probability density* functie met gemiddelde e (voorspelde waarde) en *size* parameter s (spreidingsmaat). Parameter s wordt voor elk model m verkregen uit de output van de GLMM analyse. Om rekening te houden met de onzekerheid geassocieerd met de voorspelde waarde e , wordt deze waarde 1000 maal bemonsterd uit de normaal verdeling van verwachte waarden met gemiddelde μ_e en standaard deviatie σ_e . Beide parameters worden voor elk model m bekomen uit de output van de GLMM analyse. De finale score wordt dan berekend als het gemiddelde over de 1000 bekomen waarden. Preliminaire analyses hebben aangetoond dat dit stabielere en meer betrouwbare resultaten oplevert dan rechtstreeks LG te berekenen op basis van μ_e .

2.5.1.4 Cross-validatie

Indien dezelfde dataset wordt gebruikt om modelparameters te schatten en de afwijking tussen waargenomen en voorspelde waarden te bepalen, is er een risico tot overparameterisatie waarbij het optimale model verklarende variabelen bevat die niet substantieel bijdragen tot het voorspellen van waarden extern aan de kalibratie set. Om dit te vermijden wordt gebruik gemaakt van k-fold cross-validatie (Kohavi, 1995) waarbij de schatting

van de parameters en de berekening van de model score gebeuren op afzonderlijke subsets uit de kalibratieset (Figuur 2-3). De dataset voor kalibratie wordt hiervoor onderverdeeld in vier groepen ($k = 4$) die overeenkomen met de vier winters in de kalibratieset (2008-2011). Tijdens cross-validatie worden voor een gegeven combinatie van verklarende variabelen vier GLMMs berekend, waarbij telkens één groep wordt weggelaten. De groep die niet wordt gebruikt voor de parameterschattingen, wordt dan telkens gebruikt om de model score te berekenen voor de afwijking tussen waargenomen en voorspelde waarden; i.e. de gegevens gebruikt voor het berekenen van de score zijn niet gebruikt voor de parameterschatting. Meer concreet worden er dus voor elke combinatie van verklarende variabelen vier parameterschattingen gedaan en vier scores berekend. Tenslotte wordt het gemiddelde over deze vier scores genomen. Deze cross-validatie dient enkel om de scores te berekenen. De uiteindelijke parameterschattingen geassocieerd met de gegeven set van verklarende variabelen worden berekend aan de hand van de volledige kalibratie set.



Figuur 2-3: Schematische voorstelling van de methode van K-fold cross-validatie. Score = LG score.

2.5.1.5 Stapsgewijze modelopbouw

Op basis van de hierboven beschreven procedures wordt een stapsgewijze modelopbouw uitgevoerd, waarbij stelselmatig verklarende variabelen worden toegevoegd of verwijderd uit de modelformulering. Dit resulteert dus in lineaire vergelijkingen met verschillende combinaties van parameters, aangezien elke variabele is geassocieerd met een parameter voor de bijhorende richtingscoëfficiënt (β).

De iteratieve procedure voor modelopbouw loopt als volgt (pseudocode):

1. Bereken een initieel model (GLMM) met enkel de structurele random variabelen en zonder verklarende variabelen.

2. Bereken de LG score voor dit model aan de hand van cross-validatie.
3. Voeg dit model toe aan de lijst van reeds berekende modellen (voorlopig dus slechts 1 model; tijdens de iteraties van stappen 4 - 6 worden hier modellen aan toegevoegd).
4. Voor alle nieuwe modellen x uit de top 10 van de lijst met reeds berekende modellen (laagste LG scores):
 - 4.1. Voor alle termen (variabelen v) die nog niet in model x zitten:
 - 4.1.1. Voeg de term toe aan model x en creëer zo een nieuw model y.
 - 4.1.2. Ga na of het nieuwe model y reeds is berekend.
 - 4.1.3. Indien niet, voeg toe aan een lijst met nieuwe modellen y.
 - 4.2. Voor alle termen die reeds in model x zitten:
 - 4.2.1. verwijder de term uit model x en creëer zo een nieuw model y.
 - 4.2.2. Ga na of het nieuwe model y reeds is berekend.
 - 4.2.3. Indien niet, voeg toe aan de lijst met nieuwe modellen y.
 - 4.3. Voor alle nieuwe modellen y:
 - 4.3.1. bereken het nieuwe model y (GLMM).
 - 4.3.2. Bereken de LG score voor dit model y aan de hand van cross-validatie.
 - 4.3.3. Voeg model y toe aan de lijst van reeds berekende modellen.
5. Sorteer de lijst met berekende modellen van laagste naar hoogste score.
6. Indien in de huidige iteratie nieuwe modellen werden toegevoegd in de top 10 van de lijst met reeds berekende modellen, ga terug naar 4.

2.5.1.6 Bepalen van het optimale model

Op basis van de procedure voor modelopbouw wordt een reeks modellen bekomen, gerangschikt op voorspellend vermogen aan de hand van hun model score (LG score). Voor alle modellen in deze lijst wordt nagegaan welke modellen niet significant verschillen in voorspellend vermogen van het best scorende model (eerste model in de lijst, laagste LG score). Om te bepalen of een model significant verschilt van het best scorende model, worden voor beide modellen de individuele LG scores berekend voor alle waarnemingen in de kalibratie dataset. Op deze gepaarde set van LG waarden wordt dan een eenzijdige paarsgewijze Fisher-Pitman permutatie test toegepast. Via deze test wordt nagegaan of de LG scores van het te testen model systematisch slechter zijn dan de LG scores voor het best scorende model. Op deze manier wordt een lijst van modellen geselecteerd die equivalent zijn in voorspellend vermogen aan het best scorende model (i.e. de LG scores zijn niet significant lager dan voor het best scorende model). Voor deze lijst van modellen wordt dan voor elke potentiële verklarende variabele het percentage berekend van modellen waarin deze variabele voorkomt. Het optimale model wordt tenslotte bekomen door die variabelen te behouden die in meer dan 50% van de equivalente modellen voorkomen.

2.5.2 Modelassumpties

Waarden voor de dispersiefactor van het finale model worden bekeken. De dispersiefactor geeft de verhouding weer tussen de residuele χ^2 en de residuele vrijheidsgraden van het

model. Indien de waarde van deze verhouding beduidend groter is dan 1 (> 1.2) dan is er sprake van overdispersie (de spreiding van de respons variatie wordt niet adequaat gemodelleerd) en is het model niet betrouwbaar.

Er wordt ook nagegaan of er duidelijke niet lineaire verbanden zijn tussen de residuen van het model en de verklarende variabelen. Indien dit het geval is, wijst dit er op dat de relatie tussen deze variabele en de respons niet lineair is, en wordt niet voldaan aan de *a priori* assumptie van lineariteit van de verbanden. Deze inspectie gebeurt aan de hand van grafieken.

2.5.3 Berekenen van voorspelde waarden en confidentie interval

Berekening van voorspelde aantallen gebeurt aan de hand van de regressievergelijking bekomen uit de modelschatting van het beste model. Aangezien voorspelde waarden worden berekend in de log schaal ($\log \hat{Y}$) dient de inverse transformatie (exponent) toegepast te worden om de eigenlijke aantallen \hat{Y} te bekomen. Om voorspellingen te doen in functie van aangeleverde gegevens worden de bekomen parameters voor random variabelen met betrekking tot temporele afhankelijkheid (winterjaar en wintermaand) niet gebruikt. Tijdsafhankelijke effecten kunnen immers niet voorspeld worden voor 2050. De bekomen schattingen zijn een gemiddelde over effecten van wintermaand en winterjaar. Ruimtelijke effecten gebonden aan de telgebieden, die niet kunnen verklaard worden door de habitatkenmerken, worden wel meegenomen in de voorspellingen. De voorspellingen gebeuren immers op dezelfde gebieden als gebruikt in de kalibratie. Dit verkleint de fout op de geschatte aantallen. Hierbij wordt ervan uitgegaan dat deze ruimtelijke effecten niet variabel zijn in de tijd.

Naast de schattingen van voorspelde aantallen worden ook betrouwbaarheidsintervallen rond de schattingen berekend. Voor elke voorspelde waarde gebeurt de berekening van het betrouwbaarheidsinterval als volgt:

- 1 Op basis van de variantie-covariantie matrix voor de parameters (richtingscoëfficiënten) uit het optimale model (enkel de parameters uit het 'fixed' deel van het model en niet de parameters geassocieerd met random variabelen) wordt een waarde getrokken uit de geschatte waarschijnlijkheidsdistributie (normaal verdeling) voor iedere parameter.
- 2 Op basis van deze parameterschatters wordt dan de voorspelde waarde berekend voor elk vogeltelgebied (inclusief correcties met betrekking tot gebied (ruimtelijke random variatie)).
- 3 De bovenstaande procedure wordt 1000 keer herhaald en zodoende wordt de waarschijnlijkheidsdistributie van de voorspelde waarde bekomen waaruit betrouwbaarheidsintervallen kunnen worden berekend.
- 4 Tenslotte worden de betrouwbaarheidsintervallen teruggetransformeerd (inverse functie: exponent) om de intervallen in vogelaantallen te bekomen.

De bovenvermelde methode laat toe om betrouwbaarheidsintervallen te berekenen voor individuele telgebieden. Deze methode kan ook geëxtrapoleerd worden om voorspellingen te

doen op het niveau van OMES zones (of saliniteitszones). De methode om voorspellingen te doen op een grotere geografische schaal is als volgt:

- 1 Op basis van de bovenstaande procedure (stap 1-2) wordt voor elk telgebied een voorspelde waarde berekend.
- 2 De teruggetransformeerde voorspelde aantallen worden opgeteld per geografische zone (OMES zone of saliniteitszone).
- 3 Deze procedure wordt 1000 keer herhaald en op basis van deze berekeningen kunnen de betrouwbaarheidsintervallen voor voorspelde aantallen worden berekend op de gewenste geografische schaal.

2.5.4 Validatie

Zoals hoger aangegeven worden de gegevens voor de winters 2007 en 2012 niet gebruikt voor kalibratie maar apart gehouden voor validatie. Zowel voor de validatieset als voor de kalibratieset worden de waargenomen gemiddelde aantallen (gemiddeld over de getelde maanden per winter) uitgezet ten opzichte van de voorspelde aantallen en betrouwbaarheidsintervallen. Dit gebeurt zowel voor de waarnemingen per telgebied als voor de waarnemingen per OMES zone.

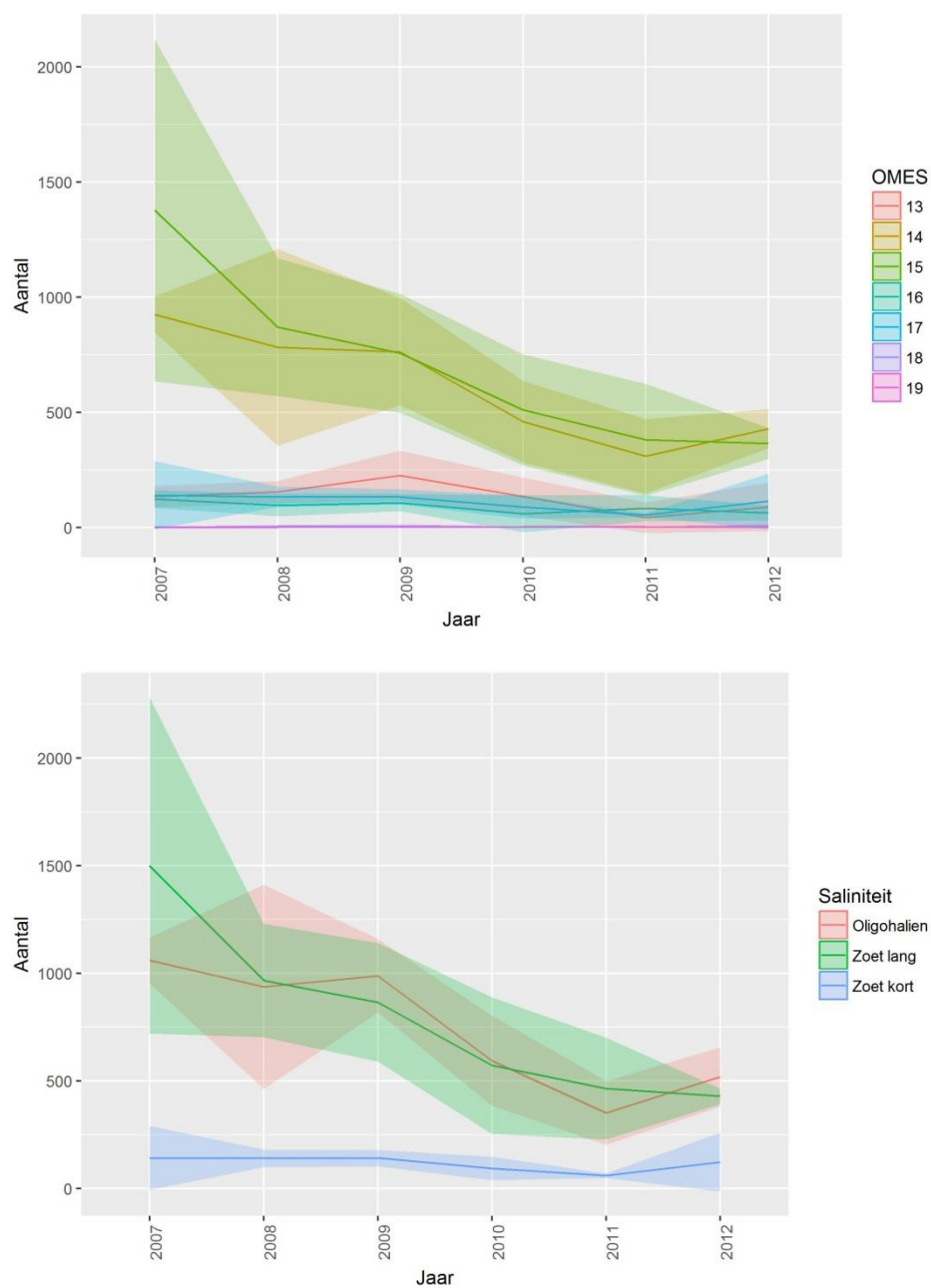
3 Resultaten

3.1 Verkennende analyse

3.1.1 Vogeltellingen

In Figuur 3-1 worden de aantallen wintertaling voor de wintermaanden (oktober-maart) uitgezet per OMES of saliniteitszone (zie Figuur 2-1 voor de situering van de zones) en per winterjaar (jaar van aanvang van de winter). De hoogste aantallen wintertaling worden waargenomen in OMES 14 en 15 (Oligohalien en Zoet lange verblijftijd respectievelijk). Er is een daling waar te nemen in de aantallen met de tijd. Deze daling in aantallen is echter klein in vergelijking met de historisch waargenomen daling wanneer winters vóór 2007 worden vergeleken met winters na 2007.

Preliminare GLMM analyses op de aantallen, waarin enkel wordt rekening gehouden met factoren gelinkt aan de ruimtelijke en temporele afhankelijkheidsstructuur (random variabelen voor telgebied, winterjaar en wintermaand genest in winterjaar) en met de offset voor aslengte, geven aan dat er een negatief binomiale respons verdeling moet gebruikt worden. Analyses met een Poisson verdeling geven een sterke overdispersie van de responsvariatie zodat deze resultaten niet betrouwbaar zijn. In de rest van de analyses zal dus steeds gewerkt worden met een negatief binomiale respons verdeling. Aan de hand van de preliminaire analyses (met negatief binomiale respons verdeling) kan er een inschatting gemaakt worden van het belang van ruimtelijke en temporele patronen in het bepalen van de aantallen, zonder in te gaan op de oorzaken van deze patronen (veranderingen in habitatkenmerken of andere onverklaarde veranderingen). De preliminaire analyses tonen aan dat voor wintertaling vooral ruimtelijke factoren (verschillen tussen telgebieden) zorgen voor grote verschillen in de aantallen (bijna drie grootteordes verschil, zie Tabel 2). Temporele factoren (verschillen tussen winterjaren) zijn slechts verantwoordelijk voor verschillen in aantallen van minder dan één grootteorde (Tabel 2).

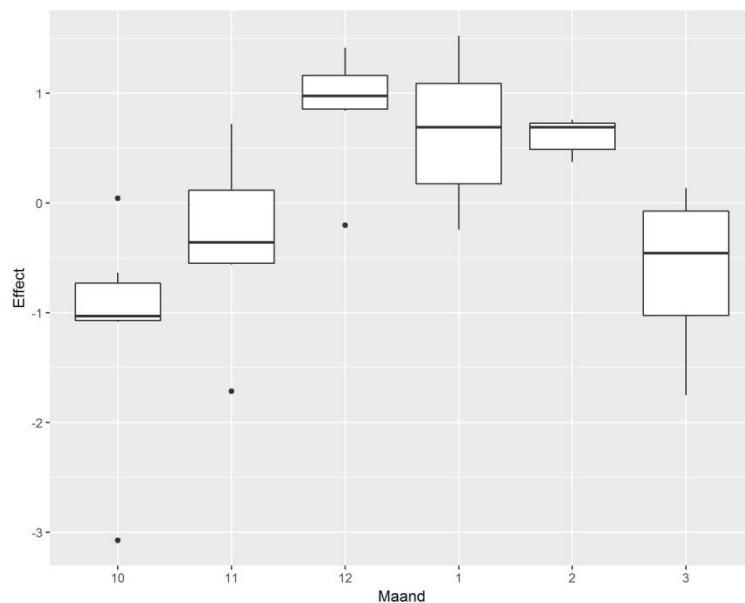


Figuur 3-1: Aantal wintertaling per OMES segment (boven) en per saliniteitszone (onder). Gemiddelde over de wintermaanden (oktober-maart) \pm 95% confidentie interval.

grootteorde	
Gebied	2.8
Winter	0.7

Tabel 2: verkennende inschatting van de ruimtelijk en temporele variatie in aantallen vogels. De waarden geven de grootteorde (i.e. 10^x geeft de verhouding weer tussen hoogste en laagste aantallen binnen het 95% betrouwbaarheidsinterval) van de verschillen in waargenomen aantallen op ruimtelijke schaal (telgebieden) en temporele schaal (winterjaren).

Op basis van preliminaire analyses kan ook de variatie tussen de wintermaanden (oktober - maart) over de jaren heen in kaart worden gebracht (Figuur 3-2). Hieruit blijkt dat de hoogste aantallen wintertaling worden waargenomen in de maanden december-januari-februari. Er wordt dan ook geopteerd om enkel verder te werken met de maanden december-januari-februari in de analyses om de variatie tussen maanden te minimaliseren.



Figuur 3-2: Maandeffect in preliminaire analyse (GLMM) met enkel ruimtelijke en temporele afhankelijkheidsstructuur. Het maandeffect is een schatter van de afwijking van de log aantallen binnen elke maand ten opzichte van het algemeen gemiddelde. De boxplots geven de variatie weer voor dezelfde maand in de verschillende jaren (2007-2012).

3.1.2 Habitatkenmerken

3.1.2.1 Gebieden verwijderd uit de dataset

In gebieden **2071101_R.L**, **2071102_R.L** (OMES 19) en gebied **2071105_R.L** (OMES 18) komt zo goed als geen intertidaal zacht substraat voor ($< 50 \text{ m}^2$). Hierdoor zijn er ook geen goede gegevens voor de helling en de spreiding in droogvalduren. Ook deze gebieden worden uit de analyse verwijderd. In deze gebieden zijn in de maanden december-februari zeer weinig wintertaling aangetroffen (minder dan 20% van de tellingen en nooit meer dan 10 individuen). Tenslotte valt ook gebied **2055801_R.L** (OMES 19) weg omdat hier geen gegevens voorhanden zijn omtrent spreiding in droogvalduur. In dit gebied werd in geen enkele telling in de maanden december-februari wintertaling aangetroffen.

3.1.2.2 Lineair verband tussen aantallen en habitatvariabelen

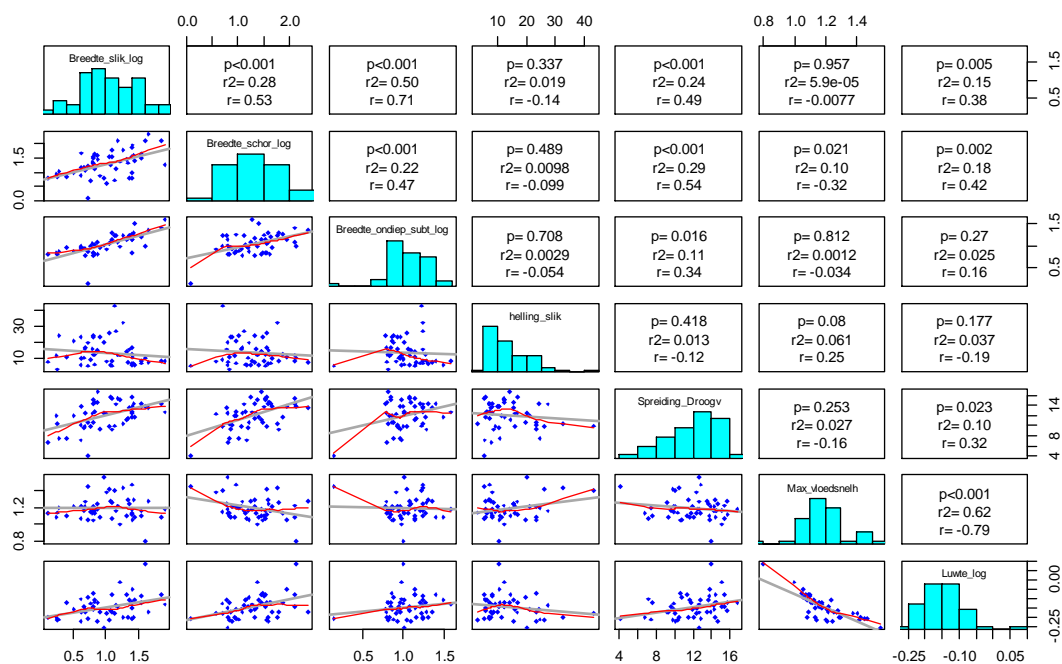
Op basis van verkennende analyses van de relaties tussen aantallen en al dan niet getransformeerde habitatvariabelen worden voor verdere analyse de breedte zacht slik, breedte schor en luwte-index logaritmisch getransformeerd. Op die manier worden de relaties tussen de (log getransformeerde) aantallen en de habitatvariabelen maximaal gelineariseerd.

3.1.2.3 Ruimtelijke variatie

Ruimtelijke variatie in (log getransformeerde) habitatkenmerken uit zich vooral op het niveau van de telgebieden (Figuren Bijlage 1). Voor de meeste habitatvariabelen zijn er ook tendensen waar te nemen op grotere ruimtelijke schaal (vb. minder slikken en schorren in het stroomopwaartse deel van de Zeeschelde), maar de variatie op kleinere schaal primeert.

3.1.2.4 Multicollineariteit tussen habitatvariabelen

Indien alle initieel geselecteerde habitatvariabelen samen worden vergeleken zijn er nog een aantal sterke correlaties, wat resulteert in hoge VIF waarden voor bepaalde variabelen (Figuur 3-3 en Tabel 3). De breedte van het ondiep subtidaal is sterk gecorreleerd met de breedte van het slik ($R^2 = 0.50$), en wordt niet verder meegenomen. De luwte index is sterk (negatief) gecorreleerd met de maximale vloodsnelheid ($R^2 = 0.62$). Bij de selectie wordt enkel de luwte index als maat voor dynamiek behouden. Deze bevat zowel informatie over de maximale stroomsnelheid als over de vloeddominantie. Wanneer de gecorreleerde variabelen worden verwijderd, blijven de VIF waarden onder het vooropgestelde criterium van 2-2.5 (Tabel 3).



Figuur 3-3: Overzicht van de paarsgewijze correlaties tussen de habitatkarakteristieken in de Boven-Zeeschelde.

VIF		VIF	
Breedte_slik_zacht_log	4.34	Breedte_slik_zacht_log	1.57
Breedte_schor_log	1.89	Breedte_schor_log	1.71
Breedte_ondiep_subt_log	2.50		
Helling_slik_zacht	1.13	helling_slik_zacht	1.05
Spreiding_Droogv	1.55	Spreiding_Droogv	1.54
Max_vloedsnelh	4.61		
Luwte_log	5.08	Luwte_log	1.31

Tabel 3: Variance Inflation Factors (VIF) voor de habitatkarakteristieken in de Boven-Zeeschelde. Links: VIF waarden indien alle initieel geselecteerde habitatkarakteristieken worden in rekening gebracht. Rechts: VIF waarden indien enkel rekening wordt gehouden met de selectie van karakteristieken die zal worden gebruikt in verdere analyses.

3.1.3 Vergelijking van statistische technieken

In een initiële fase van de conceptualisatie van de methodologie werden een aantal technieken met elkaar vergeleken. Naast GLMM werd er ook geëxperimenteerd met N-mixture models en boosted regression trees. Deze methoden werden echter niet verder uitgewerkt. Voor N-mixture models bleek de voorhanden zijnde software niet in staat de grote hoeveelheid aan gegevens efficiënt te verwerken. Boosted regression trees hebben het nadeel dat ze als een 'black-box' fungeren waarbij de exacte relaties tussen verklarende variabelen en afhankelijke variabele niet expliciet zijn wat de interpretatie bemoeilijkt. Bovendien kunnen met de voorhanden zijnde software geen random variabelen ('mixed model') worden toegevoegd om rekening te houden met de ruimtelijke en temporele afhankelijkheidsstructuur van de gegevens.

3.2 Analyse van aantallen

3.2.1 Kalibratie

Om numerieke problemen te vermijden tijdens de parameterschattingen aan de hand van GLMM worden de verklarende variabelen herschaald naar gemiddelde 0 en standaard deviatie 1:

$$V_{sc} = \frac{V - gem}{sd}$$

Tabel 4 geeft voor elke variabele de waarden gebruikt voor deze herschaling. Deze herschaling heeft ook het voordeel dat het relatieve belang van de verklarende variabelen kan vergeleken worden aan de hand van de geschatte regressieparameters.

Het finale model is een combinatie van 8 modellen die niet significant slechter scoren dan het model met de laagste LG score (zie Bijlage 2 voor een lijst van alle berekende modellen). De bepalende habitatkenmerken zijn breedte zacht slik, helling van het slik en spreiding in droogvalduur (SpD). De bijhorende parameters zijn weergegeven in Tabel 5. Op basis van de gevolgde methode worden breedte van het schor en luwte index worden niet weerhouden als bepalende variabelen. Ook de Europese trend in aantallen wintertaling wordt geselecteerd in het finale model na kalibratie. Het model inclusief Europese trend geeft echter een minder goede validatie dan een model zonder Europese trend (onderschatting van de aantallen; zie verder paragraaf 3.2.3). Daarom wordt er toch besloten om de Europese trend niet verder mee te nemen. Aangezien alle variabelen herschaald zijn naar gemiddelde 0 en standaard deviatie 1, geeft de parameterschatter β (helling van een rechte; Tabel 5) een indicatie van het belang van elke variabele voor het bepalen van het aantal vogels. Hieruit blijkt dat de breedte van het zacht slik het meest bepalende habitatkenmerk is voor de aantallen wintertaling (Figuur 3-4A). Hoe breder het slik, hoe meer wintertaling, voor een zelfde afstand langsheen de rivieras. Ook de spreiding in droogvalduur heeft een positief effect op de aantallen (Figuur 3-4B). De helling van het slik tenslotte heeft een negatief effect op de aantallen (Figuur 3-4C).

Het finale model (zonder Europese trend) met alle relevante variabelen inclusief alle random variabelen is weergegeven in de volgende vergelijking:

$$\log \hat{Y} = \log_{10} \text{Aslengte} + \beta_0 + \beta_1 \cdot \log_{10} \text{Breedte Slik} + \beta_2 \cdot \text{Helling Slik} + \beta_3 \cdot \text{SpD} + \gamma_1(\text{Gebied}) + \gamma_2(\text{Winter}) + \gamma_3(\text{Maand})$$

Voor de uiteindelijke voorspellingen op basis van de aangeleverde gegevens wordt geen rekening gehouden met random variabelen die betrekking hebben op tijdsafhankelijke variatie. Tijdsafhankelijke effecten kunnen immers niet voorspeld worden voor 2050. De vergelijking zonder deze tijdsafhankelijke random variabelen is dan als volgt:

$$\log \hat{Y} = \log_{10} \text{Aslengte} + \beta_0 + \beta_1 \cdot \log_{10} \text{Breedte Slik} + \beta_2 \cdot \text{Helling Slik} + \beta_3 \cdot \text{SpD} + \gamma_1(\text{Gebied})$$

Waarden voor β parameters worden weergegeven in Tabel 5. Waarden γ_1 geassocieerd met elk telgebied zijn weergegeven in Tabel 6.

	gem	sd
vorstgetal	11.94	5.97
Log₁₀(eurotrend)	0.27	0.03
Log₁₀(breedte slik)	1.15	0.53
Log₁₀(breedte schor)	1.31	0.48
helling slik	13.86	8.22
SpD	12.29	2.78
log₁₀(luwte)	-0.13	0.07

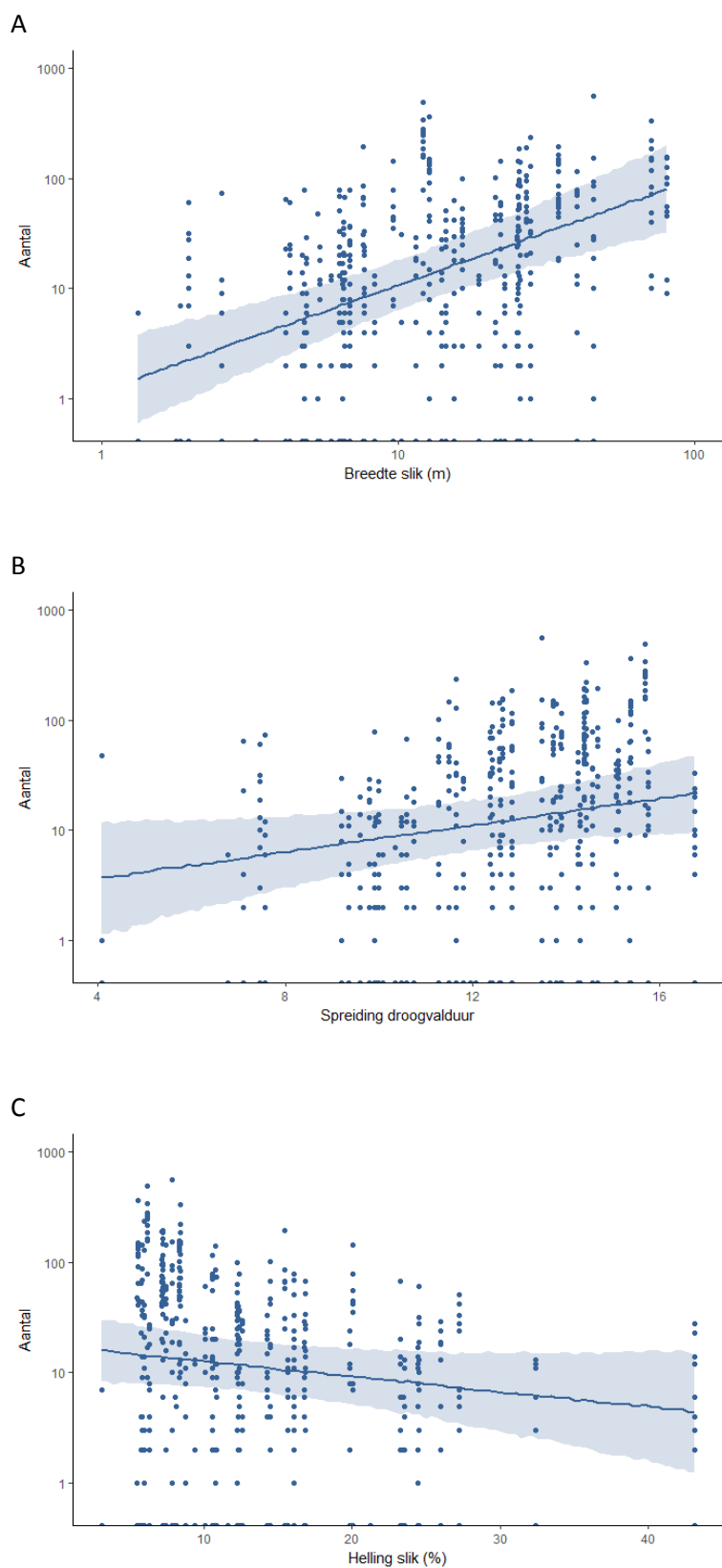
Tabel 4: gemiddelde en standaard deviatie gebruikt voor herschaling van de variabelen in de modelselectie.

		Schatter	Std. Error
Intercept	β_0	-0.55 (-0.85)	0.27 (0.27)
log₁₀Breedte Slik	β_1	1.16	0.23
Helling Slik	β_2	-0.26	0.16
SpD	β_3	0.39	0.19
(log₁₀EuroTrend)	(β_4)	(-0.59)	(0.31)

Tabel 5: geschatte regressie coëfficiënten β voor de relatie tussen aantallen wintertaling en verklarende variabelen in het finale model. Merk op dat deze parameters van toepassing zijn op de herschaalde variabelen (zie boven). Waarden tussen haakjes geven parameterschattingen wanneer Europese trend wordt meegenomen in het model.

OMES	Telgebied	γ_1	OMES	Telgebied	γ_1	OMES	Telgebied	γ_1
19	2071103_R.L	-2.11	16	2091613_L	0.06	14	4143107_R	0.04
19	2071104_R.L	-0.99	16	2091613_R	-0.01	14	4143108_L	-0.31
18	2071106_R.L	-0.61	16	2091614_L	0.38	14	4143108_R	0.24
18	2091601_R.L	-0.26	16	2091614_R	-0.22	14	4143109_L	0.13
18	2091602_R.L	-1.19	15	4143101_L	-0.50	14	4143109_R	-0.15
17	2091603_R.L	2.59	15	4143101_R	-0.78	14	4140106_R	0.08
17	2091604_R.L	1.07	15	4143102_L	1.22	14	4140501_L	0.07
17	2091605_R.L	1.44	15	4143102_R	-0.28	14	4143110_L	0.08
17	2091606_R.L	-0.13	15	4143103_L	1.56	14	4143110_R	0.27
17	2091607_R.L	0.49	15	4143103_R	-0.19	13	4143111_L	-0.88
17	2091608_R.L	-0.23	15	4143104_L	0.17	13	4143111_R	1.10
17	2091609_R.L	0.07	15	4143104_R	-0.08	13	3156601_L	0.18
16	2091610_R.L	-2.25	15	4143106_L	-0.56	13	3156601_R	-1.66
16	2091611_L	1.07	15	4143106_R	-0.88	13	3156602_L	0.32
16	2091611_R	0.15	15	4143105_L	2.02	13	3156602_R	0.12
16	2091612_L	0.72	15	4143105_R	0.19	13	3156603_L	-1.17
16	2091612_R	0.38	14	4143107_L	-0.62	13	3156603_R	0.04

Tabel 6: Geschatte onverklaarde variatie geassocieerd met telgebieden (ruimtelijke variatie). Waarden γ_1 geven voor elk telgebied de afwijking ten opzichte van de geschatte waarde ($\log \hat{Y}$) op basis van enkel de verklarende variabelen ('fixed' component van het model). Telgebieden zijn gerangschikt van stroomopwaarts naar stroomafwaarts.



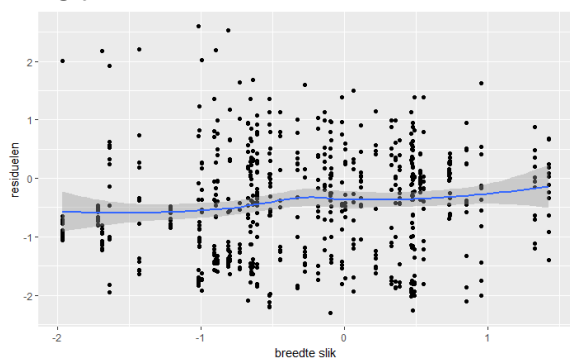
Figuur 3-4: Relatie tussen aantal wintertaling ($\pm 95\%$) en habitatkenmerken op basis van het finale model.

3.2.2 Modelassumpties

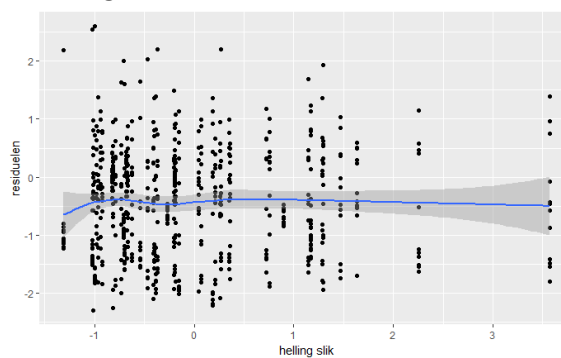
De dispersiefactor van het model voor wintertaling is 0.85. Er zijn dus geen aanwijzingen voor overdispersie.

Aan de hand van visuele inspectie van de grafieken (Figuur 3-5) werden geen noemenswaardige (niet-lineaire) tendensen in de relatie tussen verklarende variabelen en residuen waargenomen.

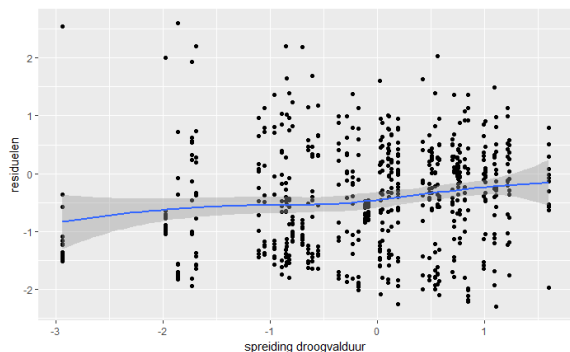
A: \log_{10} Breedte Slik



B: Helling Slik



C: SpD



Figuur 3-5: Relatie tussen residuen en verklarende habitatkenmerken voor wintertaling

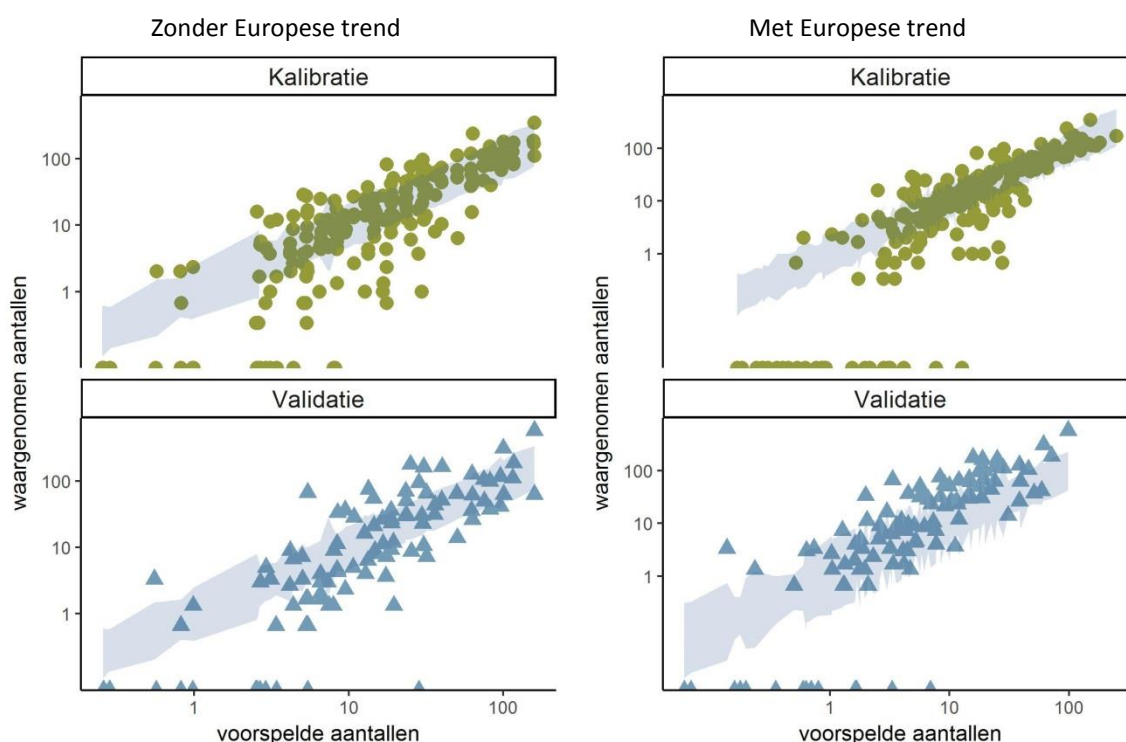
3.2.3 Validatie

In elk van de onderstaande figuren geven de punten het gemiddelde weer van de waargenomen aantallen (gemiddeld over drie wintermaanden per winterjaar), voor elk winterjaar afzonderlijk en voor elk telgebied of OMES zone. De grijze band geeft telkens het 95% betrouwbaarheidsinterval rond het voorspelde gemiddelde.

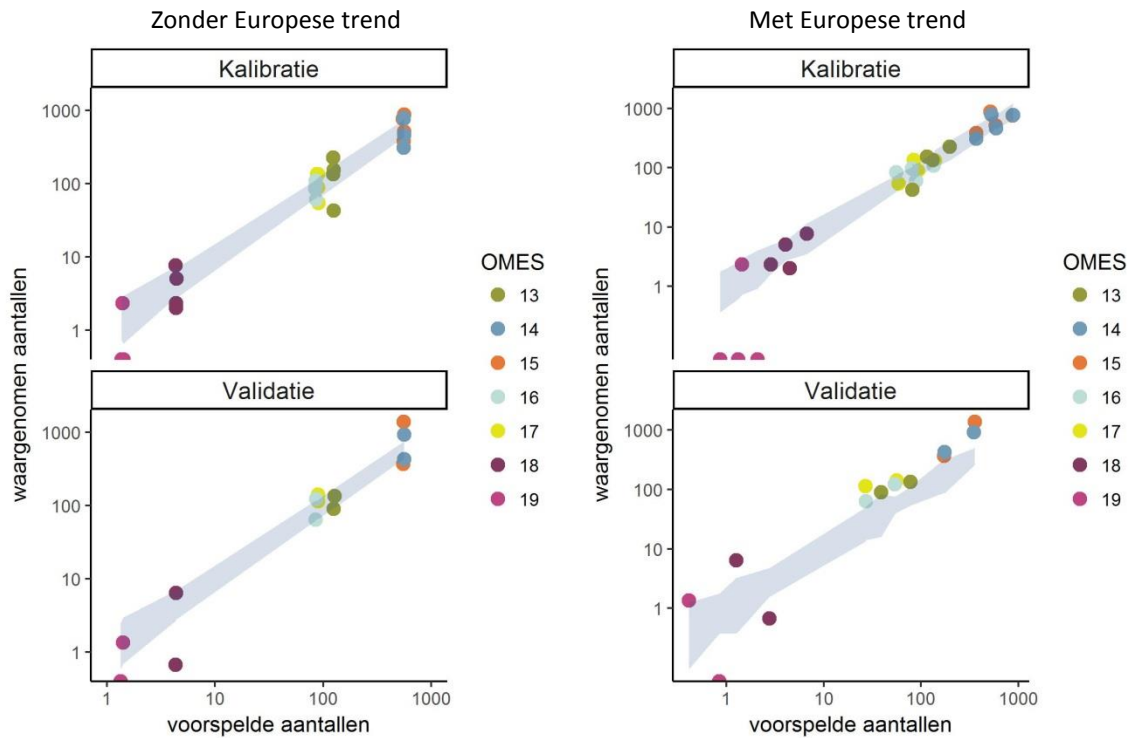
Er wordt een vergelijking gemaakt tussen het finale model zonder Europese trend en het finale model met Europese trend. Zonder Europese trend in de regressievergelijking liggen zowel voor de kalibratieset als voor de validatieset de punten per winter en vogeltelgebied grotendeels binnen de voorspelde 95% betrouwbaarheidsintervallen rond het gemiddelde

(Figuur 3-6). Vanwege de aard van de gegevens (aantallen) kan er een grote variatie optreden tussen de individuele maandelijkse tellingen en kunnen deze sterk afwijken van het eigenlijke gemiddelde. Aangezien de waargenomen gemiddeldes hier slechts zijn berekend op drie datapunten (drie maanden per winter), kunnen deze schatters van het waargenomen gemiddelde ook nog sterk afwijken van het eigenlijke gemiddelde. Het feit dat een aantal punten buiten de betrouwbaarheidsintervallen liggen is dus niet problematisch. Ook de nulwaarden voor het waargenomen gemiddelde zijn niet in tegenspraak met de verwachtingen. Deze nulwaarden komen grotendeels voor bij voorspelde gemiddeldes lager dan 5. Indien de Europese trend wel in rekening wordt gebracht in de regressievergelijking zijn de voorspellingen voor de kalibratieset nog steeds goed. Voor de validatieset is er echter een systematische onderschatting van de aantallen. Zoals eerder vermeld, wordt er daarom voor gekozen om de Europese trend niet verder mee te nemen.

Zonder Europese trend komen ook de waargenomen aantallen voor wintertaling per OMES zone zeer sterk overeen met de verwachtingen zowel voor kalibratie als voor validatie (Figuur 3-7). In Figuur 3-7 is bovendien duidelijk te zien dat voorspellingen rekening houdend met Europese trend een onderschatting geven van de waargenomen aantallen voor de validatieset.



Figuur 3-6: Relatie tussen voorspelde aantallen en waargenomen aantallen wintertaling in de telgebieden. Vergelijking tussen validatie en kalibratie set. Punten geven het gemiddelde per telgebied en winter (gemiddeld over wintermaanden) weer voor de waargenomen aantallen. De grijze band geeft het 95% betrouwbaarheidsinterval weer van de voorspelde waarden. Links: voorspellingen zonder Europese trend in de regressievergelijking. Rechts: voorspellingen met Europese trend in de regressievergelijking.



Figuur 3-7: Relatie tussen voorspelde aantallen en waargenomen aantallen wintertaling per OMES zone. Vergelijking tussen validatie en kalibratie set. Punten geven het gemiddelde per OMES zone en winter (gemiddeld over wintermaanden) weer voor de waargenomen aantallen. De grijze band geeft het 95% betrouwbaarheidsinterval weer van de voorspelde waarden. Links: voorspellingen zonder Europese trend in de regressievergelijking. Rechts: voorspellingen met Europese trend in de regressievergelijking.

4 Discussie

Op basis van de beschreven methodiek werd voor wintertaling een statistisch model opgesteld met als streefdoel een instrument te genereren voor het voorspellen van aantallen op basis van onafhankelijke, externe gegevens. Het hoofddoel van de huidige oefening is immers om een beleidsondersteunend instrumentarium te voorzien dat kan worden ingezet om de effecten in te schatten van geplande morfologische ingrepen in de Boven-Zeeschelde. Die ingrepen hebben tot doel het systeem tegen 2050 te verbeteren in relatie tot ecologie, veiligheid, scheepvaart, onderhoud en bijkomende functies. De beoogde modeltrein (MTR) vormt een aaneenschakeling van modellen die de processen in de Schelde weerspiegelen, gaande van hydrodynamiek en sedimenttransport (berekend door het Waterbouwkundig Laboratorium; WL) over waterkwaliteit en pelagische ecosysteem functioneren (berekend door Universiteit Antwerpen; UA) tot habitatkwaliteit van slikken en schorren en het functioneren van de hogere trofische niveaus (INBO). Hierbij worden de modellen verder in de trein gevoed met resultaten uit voorafgaande modules.

Voor de vogelaantallen wordt input gebruikt uit de habitatmodellen (INBO, Van Braeckel et al. 2018), die op hun beurt gevoed worden door input uit het hydrodynamisch model en het model voor sedimenttransport (WL). Op die manier worden een aantal alternatieven voor morfologische ingrepen met elkaar vergeleken, waarbij ook rekening wordt gehouden met verschillende scenario's van klimaatverandering ('Model instruments for the Integrated Plan Upper Seascheldt' (IMDC et al., 2015)).

Globaal genomen wordt een groot deel van de waargenomen ruimtelijke variatie in aantallen wintertaling verklaard door de geselecteerde habitatvariabelen in het finale model. Er is een sterke reductie in de variatie toegekend aan ruimtelijke random variabelen (telgebied) in vergelijking met een model zonder habitatkenmerken. Dit betekent dat een groot deel van de waargenomen variatie in aantallen tussen telgebieden kan verklaard worden door de gekozen habitatvariabelen in het finale model. Tabel 7 geeft een vergelijking van de niet verklaarde ruimtelijke en temporele variatie (= toegekend aan telgebieden of winters) tussen het model zonder verklarende variabelen (zie ook Tabel 2) en het finale model. Voor ruimtelijke variatie is de onverklaarde component met ongeveer één grootteorde gedaald (i.e. 10 keer lager) in het finale model. Voor onverklaarde temporele variatie is er geen verschil tussen de modellen, wat logisch is aangezien er geen temporele variatie in de habitatkenmerken aanwezig is.

Het finale model voor wintertaling bevat drie variabelen gerelateerd tot habitatkwaliteit: breedte van de slikken met zacht substraat, helling van de slikken en spreiding in droogvalduur. Door de herschaling van de variabelen in de analyse is het mogelijk om de richtingscoëfficiënten (β) uit de lineaire vergelijking (zie paragraaf 3.2.1) met elkaar te vergelijken naar grootte en richting en zo het relatieve belang van de verschillende variabelen te bepalen (zie Tabel 5). De factor met de grootste invloed op de aantallen wintertaling is de oppervlakte aan beschikbaar slik ($\beta = 1.16$) (Breedte aan slik per eenheid afstand langs de lengteas van de rivier). In Tabel 8 is te zien dat het berekende aantal wintertalingen in een telgebied varieert tussen 2 vogels in een klein gebied tot 79 vogels in een groot gebied (gegeven een waarde gelijk aan de 50% percentiel voor de overige variabelen in de vergelijking). Deze variabele resulteert dus in de grootste range aan berekende vogelaantallen.

Naast het feit dat grotere (bredere) slikken grotere aantallen wintertaling aantrekken geven de resultaten ook aan dat de kwaliteit van het slik belangrijk is. Zo resulteert een steile helling van het slik in minder aantallen wintertaling en een grote spreiding in droogvalduur zorgt voor een hoger aantal vogels. Een steilere helling is geassocieerd met een minder stabiele en duurzame oever en is dus minder geschikt als foerageer- en rustgebied. Een grote spreiding in droogvalduur betekent dat het slik een vrij egaal profiel heeft (niet hol of bol) en zorgt ervoor dat er gedurende een groot deel van de getijcyclus slik aanwezig is dat een goede verhouding tussen water en substraat bevat (i.e. rond de waterlijn) en zodoende optimaal is voor grondeleenden zoals wintertaling om in te foerageren. De observaties in verband met droogvalduur sluiten aan bij onderzoek naar foerageergedrag bij steltlopers in de Westerschelde (Vanoverbeke & Van Ryckegem, 2015; de Jong, et al., 2016). Ook hier werd aangetoond dat vooral kleinere soorten gebieden verkiezen met een grote spreiding in droogvalduur.

In een eerste model werd naast de habitatkenmerken ook de Europese trend in aantal wintertalingen meegenomen als bepalende variabele. Dit resulteerde echter in een systematische onderschatting van de voorspelde aantallen voor de validatieset. Daarom werd toch besloten om de Europese trend niet verder mee te nemen. Dit vereenvoudigt ook het maken van voorspellingen voor de toekomstscenario's, aangezien er geen voorspellingen voorhanden zijn voor deze Europese trend in vogelaantallen.

	Zonder habitatkenmerken	Finaal model met habitatkenmerken
Gebied	2.8	1.7
Winter	0.7	0.7

Tabel 7: Inschatting van de ruimtelijk en temporele variatie in aantallen vogels. De waarden geven de grootteorde (i.e. 10^x geeft de verhouding weer tussen hoogste en laagste aantallen binnen het 95% betrouwbaarheidsinterval) van de verschillen in waargenomen aantallen op ruimtelijke schaal (telgebieden), temporele schaal (winters). Zonder: verkennende inschatting van de variatie zonder habitatvariabelen in het model. Finaal: residuele (kan niet worden toegeschreven aan habitatvariabelen) ruimtelijke en temporele variatie in het finale model.

Habitatkenmerken	Kenmerkwaarde	Aantal
Breedte slik (zacht substraat)	Laag	2
	Hoog	79
Helling slik	Laag	15
	Hoog	6
Spreiding in droogvalduur	Laag	5
	Hoog	20

Tabel 8: voorspelde aantallen in functie van variatie in kenmerkwaarden. Laag = 2.5% percentiel van de kenmerkwaarden; Hoog = 97.5% percentiel van de kenmerkwaarden. Voor de overige kenmerken in de lineaire vergelijking wordt de waarde gefixeerd op de 50% percentiel. Gebiedsspecifieke correctiefactoren (ruimtelijke random variatie) worden hier niet in rekening gebracht.

5 Conclusies

Er werd gekozen om te werken met lineaire regressie om habitatkenmerken te linken aan het aantal wintertaling in de telgebieden. Om het voorspellend karakter van de modellen te optimaliseren werd een procedure uitgewerkt aan de hand van cross-validatie waarbij het voorspellend vermogen van mogelijke modellen wordt vergeleken via de logaritmische score. Op basis van de verkennende analyses en de voorgestelde procedure voor modelselectie zijn de volgende verklarende variabelen gekomen voor het voorspellen van het aantal wintertaling: Breedte van het slik (zacht substraat), helling van het slik en spreiding in droogvalduur. Aan de hand van lineaire modellen die deze variabelen linken aan het aantal vogels kan een groot deel van de ruimtelijke variatie in aantallen in de kalibratie en validatie set verklaard worden.

6 Referenties

- Bates D., Maechler M., Bolker B. & Walker S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- Cox T. J. S., Maris T., Soetaert K. E. R., Conley D. J., Damme S. V., Meire P., Middelburg J. J., Vos M. & Struyf E. (2009). A macro-tidal freshwater ecosystem recovering from hypereutrophication: the Schelde case study. *Biogeosciences*, 6.
- de Jong D., Van Ryckegem G. & Van den Bergh E. (2016). Vogels en hun habitat: waarom kiezen onze doelsoorten voor de Schelde? VNSC Scheldesymposium 23 november 2016 – sessie Natuur.
- Elith J., Leathwick J. R. & Hastie T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4), 802-813.
- IJnsen, F. (1991). Karaktergetallen van de winters vanaf 1707. *Zenit* 18: 65-73.
- IMDC, Technum, WL, UA & INBO (2015). Modelling instruments for the Integrated Plan Upper Seascheldt. I/NO/11448/14.165/DDP.
- Kéry M. & Royle J. A. (2015). Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1: Prelude and Static Models. Academic Press.
- Kohavi R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Ijcai*, 14(2), 1137-1145.
- Maximova T., Vanlede J., Plancke Y., Verwaest T. & Mostaert F. (2013). Habitatmapping ondiep water Zeeschelde: Deelrapport 2 - Numeriek 2D model. Version 1.2. WL Rapporten, 00_028. Flanders Hydraulics Research. Antwerpen.
- Onkelinx T., Van Ryckegem G., Bauwens D., Quataert P. & Van den Bergh E. (2008). Potentie van ruimtelijke modellen als beleidsondersteunend instrument m.b.t. het voorkomen van watervogels in de Zeeschelde. Rapport INBO.R.2008.34. Instituut voor Natuur- en Bosonderzoek, Brussel.
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Smolders S., Maximova T., Vanlede J., Verwaest T., & Mostaert F. (2015). Integraal Plan Bovenzeeschelde: Subreport 1 – 3D Hydrodynamisch model Zeeschelde en Westerschelde. Version 1.0. WL Rapporten, 13_131. Flanders Hydraulics Research: Antwerp, Belgium.
- Van Braeckel, A. (2013). Geomorfologie – Fysiotopen – Ecotopen. P. 89-102. In Van Ryckegem G. (red.). MONEOS – Geïntegreerd datarapport INBO: toestand Zeeschelde 2012. Monitoringsoverzicht en 1^{ste} lijnsrapportage Geomorfologie, diversiteit Habitats en diversiteit Soorten. Rapport INBO.R.2013.26. Instituut voor Natuur- en Bosonderzoek, Brussel.

Van Braeckel A., Speybroeck J., Vanoverbeke J., Van Ryckegem G. & Van den Bergh E. (2018). Habitatmapping Zeeschelde subtidaal. Ecologie en ecotopen in het subtidaal. Relatie tussen bodemdieren en hydro- en morfodynamiek. Rapport INBO.R.x Instituut voor Natuur- en Bosonderzoek, Brussel.

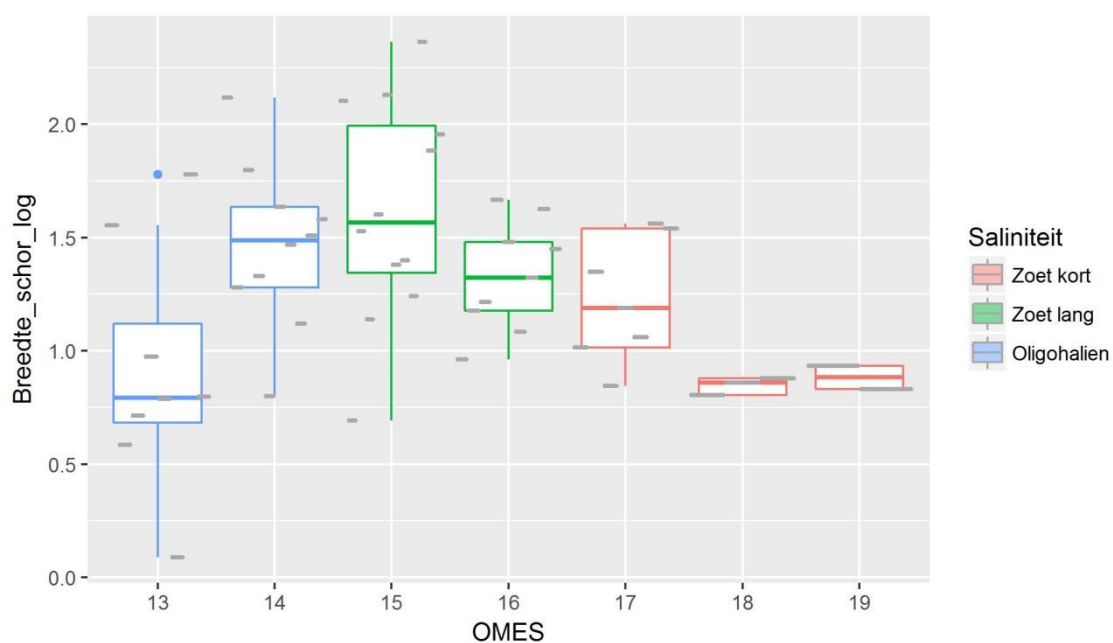
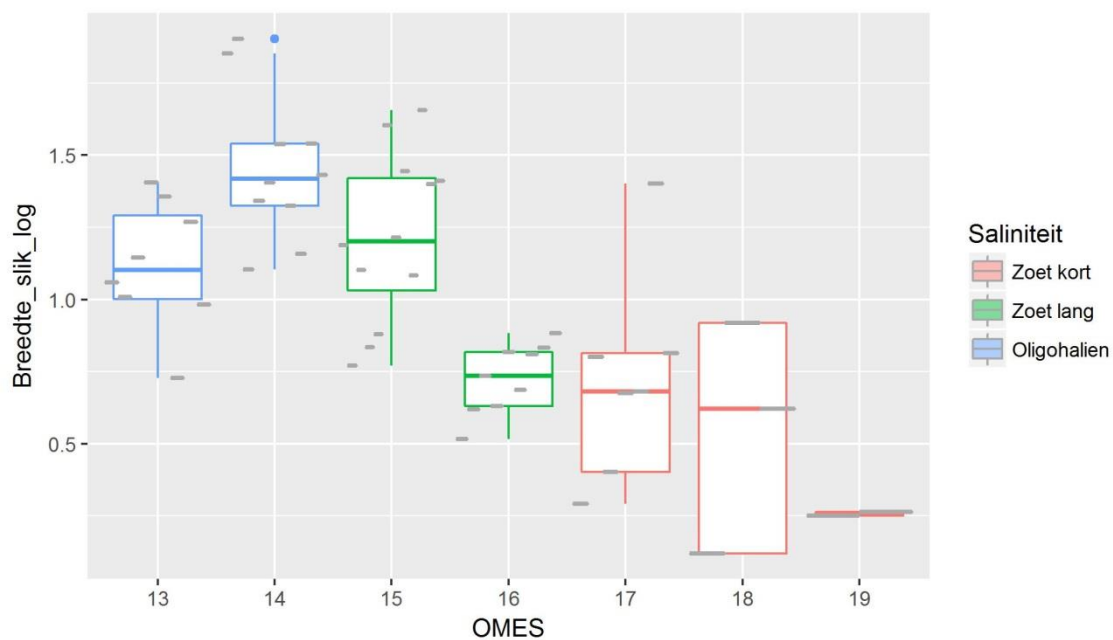
Van Braeckel A., Vanoverbeke J., Elsen R. & Van Ryckegem G. (2018). Integraal plan Boven-Zeeschelde. Deelrapport Habitatmodel: methodiek en calibratie ACT2013. Rapport INBO.R.x Instituut voor Natuur- en Bosonderzoek, Brussel.

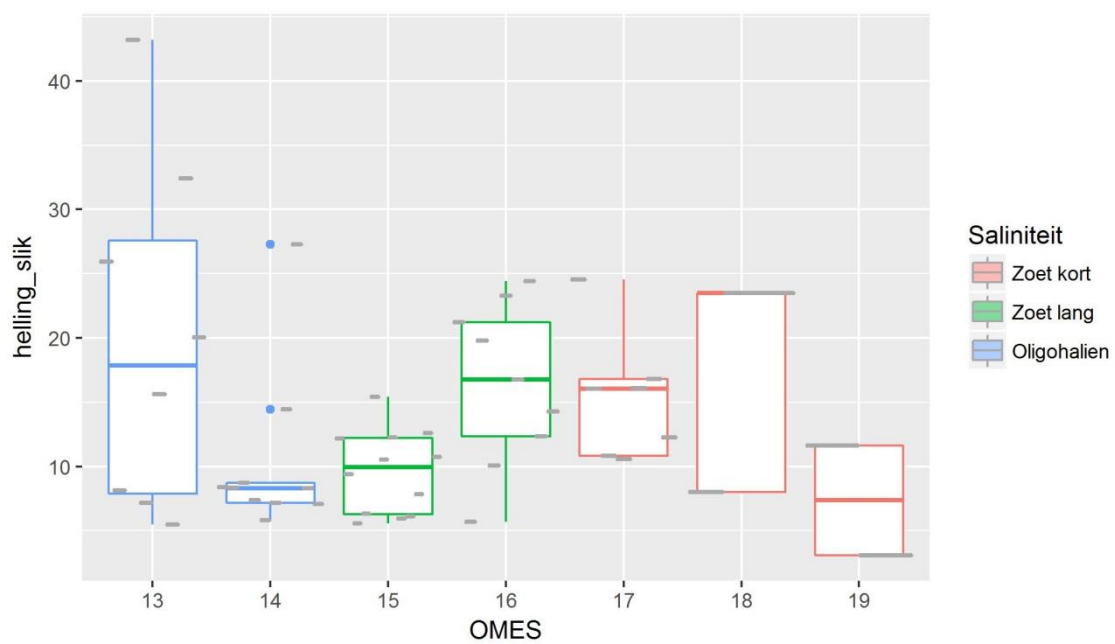
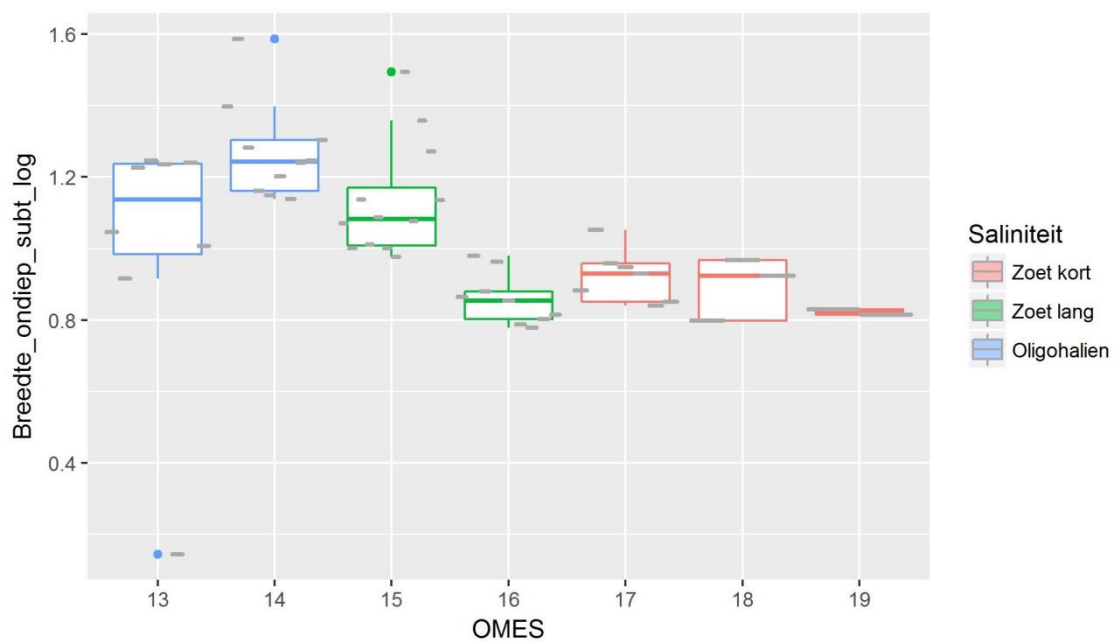
Vanoverbeke J. & Van Ryckegem G. (2015). Statistische analyse van het gebruik van het litoraal door steltlopers in de Westerschelde. Rapport INBO.R.2015.11358580. Instituut voor Natuur- en Bosonderzoek, Brussel.

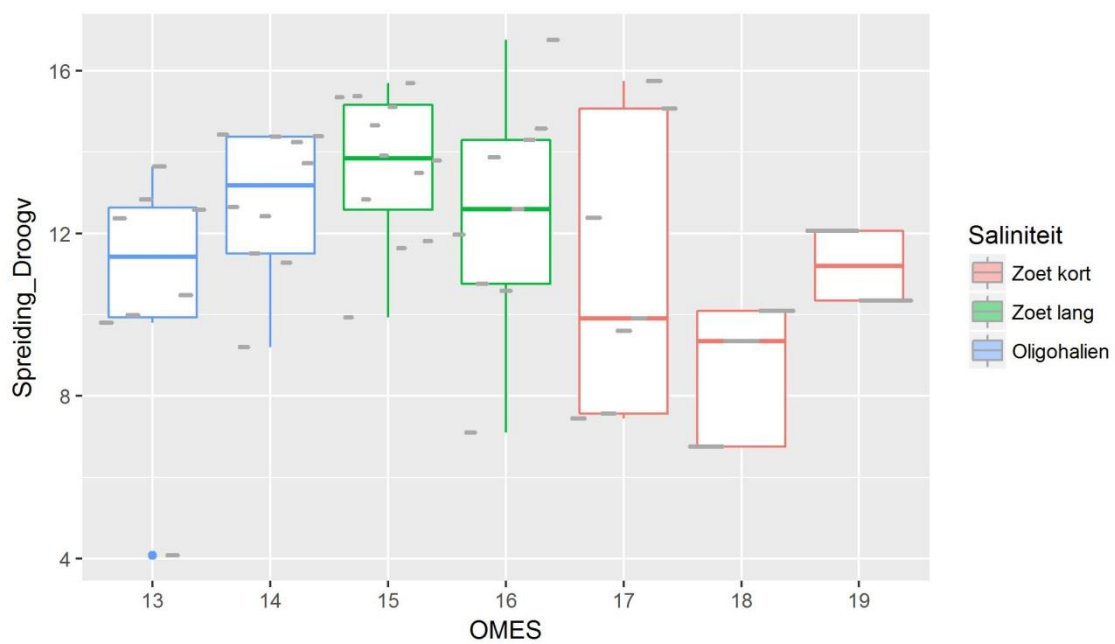
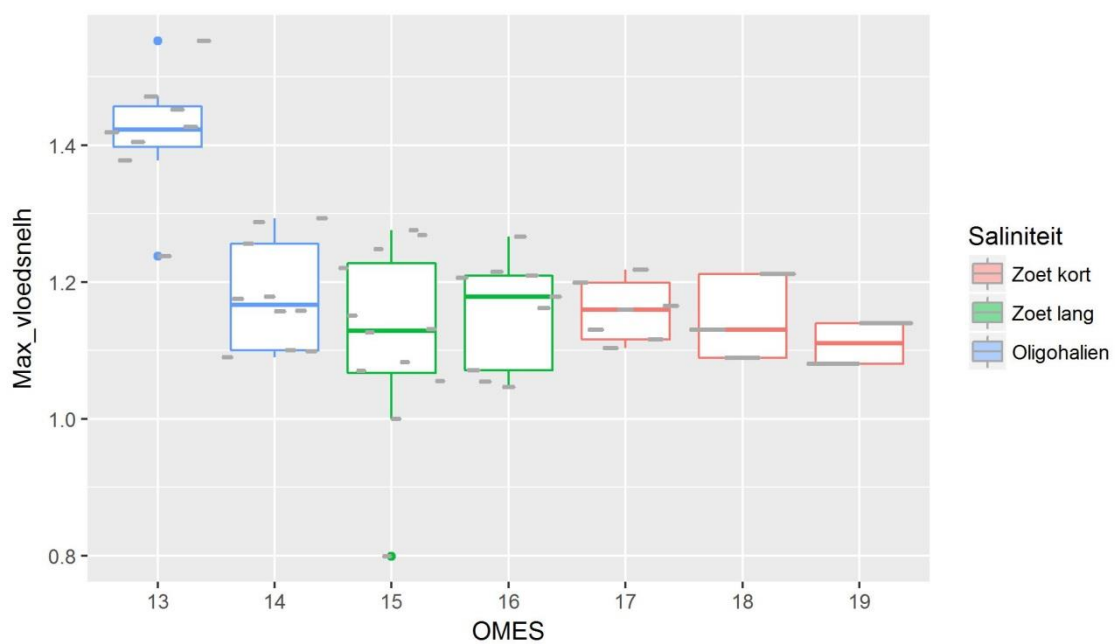
Van Ryckegem G. (2015). Watervogels. p. 80-88 *In* Van Ryckegem G. (red.). MONEOS – Geïntegreerd datarapport INBO: toestand Zeeschelde 2014. Monitoringsoverzicht en 1^{ste} lijnsrapportage Geomorfologie, diversiteit Habitats en diversiteit Soorten. Rapport 8990774. Instituut voor Natuur- en Bosonderzoek, Brussel.

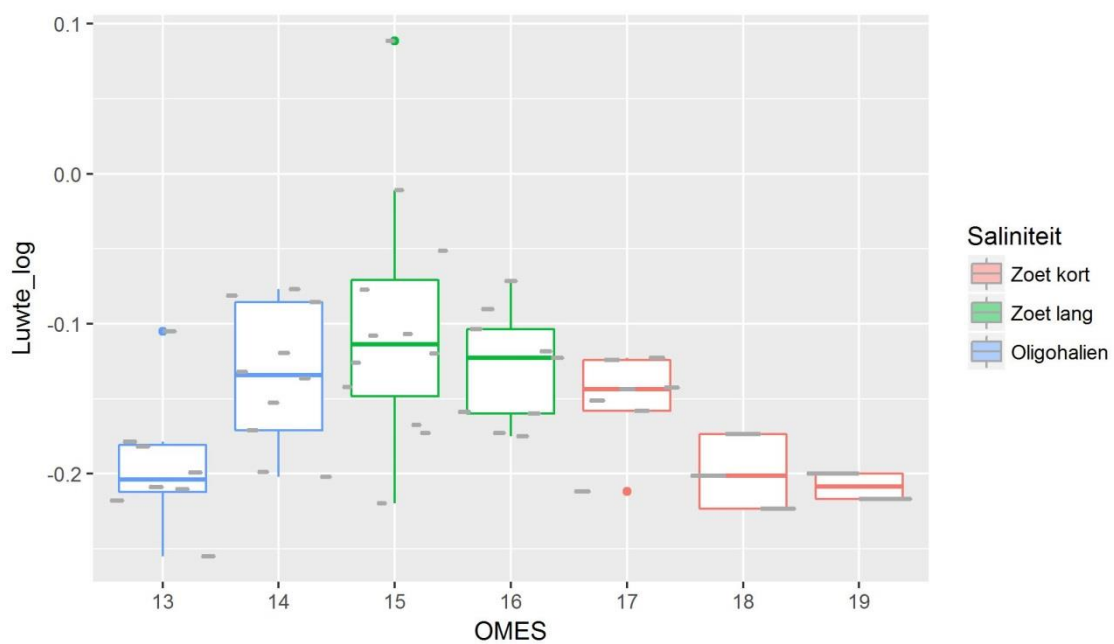
Van Ryckegem G. , De Regge N. & Van den Bergh E. (2006). Voedseleecologie en gedrag van overwinterende watervogels langs de Zeeschelde. Rapport INBO.R.2006.28. Instituut voor Natuur- en Bosonderzoek, Brussel.

Bijlage 1: Weergave van de Ruimtelijke variatie in de habitatkarakteristieken









Bijlage 2: lijst van alle berekende modellen voor wintertaling in functie van habitatkenmerken

De modellen zijn geordend volgens oplopende LG score. Een lagere LG score betekent een beter voorspellend vermogen van het model.

I: intercept; Br SI: breedte slik; Br Sch: breedte schor; H: helling slik; SpD: spreiding droogvalduur; L: luwte index

Pars: aantal parameters in het model (exclusief intercept); LG: gemiddelde LG score

Rank: rangschikking van de modellen op basis van de LG score. Modellen die niet significant verschillen van het beste model (laagste LG score) zijn in het vet en met een * aangeduid.

I	Br SI	Br Sch	H	SpD	L	Pars	LG	Rank
-0.54573	1.162032		-0.26105	0.390897		3	4.030439	1*
-0.54763	1.167415		-0.26327	0.393589	-0.0181	4	4.042139	2
-0.57076	1.071227	0.219987	-0.25887	0.31171		4	4.053849	3*
-0.53282	1.207927			0.406409		2	4.054102	4*
-0.5299	1.198562			0.401823	0.029321	3	4.067769	5
-0.55857	1.114741	0.22521		0.3251		3	4.076215	6*
-0.56151	1.119564	0.2311		0.326527	-0.02274	4	4.089621	7
-0.51919	1.403703		-0.28083			2	4.095753	8*
-0.56634	1.187396	0.341298	-0.27115			3	4.096169	9*
-0.57363	1.199059	0.356485	-0.27776		-0.0545	4	4.103037	10
-0.51483	1.38584		-0.2749		0.045654	3	4.111016	11
-0.82108			-0.34731	0.856617		2	4.113294	12*
-0.55323	1.23728	0.352263				2	4.118413	13*
-0.50397	1.462091					1	4.137191	14
-0.83064		0.554291	-0.32605	0.565236		3	4.146519	15
-0.55351	1.237787	0.352868			-0.00213	3	4.151614	16
-0.7788			-0.31219	0.793297	0.243377	3	4.153497	17
-0.49544	1.42172				0.096564	2	4.160009	18
-0.81621				0.900654		1	4.176691	19
-0.82689		0.576511		0.595019		2	4.178144	20
-0.7624				0.811728	0.307608	2	4.215371	21
-0.80035		0.528862		0.577869	0.146773	3	4.236818	22
-0.87499		0.863713	-0.36329			2	4.302072	23
-0.9001			-0.45023			1	4.341437	24
-0.84672		0.808384	-0.34289		0.14813	3	4.351396	25
-0.8725		0.905998				1	4.422666	26
-0.83063		0.820026			0.220832	2	4.471717	27
-0.79573			-0.35953		0.518179	2	4.489186	28
-0.89667						0	4.591374	29
-0.77729					0.599273	1	4.670001	30