

Van: Inwoner (*geanonimiseerd door de griffie*)

Verzonden: dinsdag 28 oktober 2025 13:47

Aan: Griffie-Hillegom <Griffie-Hillegom@HLTsamen.nl>

Onderwerp: De illusie van een betrouwbare stikstof-modelwerkelijkheid

Geachte griffie medewerkster,

Bijgaand stuur ik u een belangrijk wetenschappelijk rapport over stikstof, geschreven door dhr. Meester, hoogleraar Waarschijnlijkheidsrekening aan de Vrije Universiteit Amsterdam.

Omdat deze informatie belangrijk is voor diverse thema's die ook in Hillegom invloed uitoefenen op de bewoners van Hillegom hierbij mijn vriendelijk verzoek om dit belangrijke rapport aan het college en de leden van de gemeenteraad te sturen. Tevens met het verzoek aan een ieder om dit rapport aandachtig te lezen en te bedenken welke veranderingen in beleid en visie dit rapport te weeg kan brengen.

Met vriendelijke groet,

Inwoner (*geanonimiseerd door de griffie*)

De illusie van een betrouwbare stikstof-modelwerkelijkheid

Een onderzoek naar de onzekerheden in het wetenschappelijke stikstofdiscours;
Aerius/OPS, Kritische Depositiewaardes, Natuurdoelanalyses, en de adviezen
van de Ecologische Autoriteit

Ronald Meester¹

17 oktober 2025

¹Ronald Meester is hoogleraar Waarschijnlijkheidsrekening aan de Vrije Universiteit Amsterdam. Hij publiceert ruim en breed over kansrekening, statistiek, het gebruik van wiskundige modellen en over de relatie tussen wetenschap, techniek, levensbeschouwing en samenleving. Hij treedt regelmatig als forensisch statistisch expert op bij rechtszaken om te adviseren over de gebruikte logica, berekeningen en modellen, en adviseert verschillende gremia. Recente publicaties zijn bijvoorbeeld *Van aardbeving tot zoönose – over de inzet van modellen voor beleid*, *De onttovering van AI – een pleidooi voor het gebruik van gezond verstand* (beide met Marc Jacobs, Mazirel Pers), *Kan dat geen toeval zijn – een kritische blik op statistische bewijsvoering* (met Klaas Slooten, AUP), en *Wetenschap als nieuwe religie* (Ten Have). Hij doceert naast de VU bij SSR (het opleidingsinstituut van de rechterlijke macht) en De Filosofische School Nederland, en is een veelgevraagd spreker. Hij neemt actief deel aan het publieke debat over de inzet van modellen.

Samenvatting

Ik heb allereerst vastgesteld dat de natuurdoelanalyses (NDA's) en de adviezen van de Ecologische Autoriteit (EA) blijken geven van een 'maakbaar' wereldbeeld waarin metaforische knoppen zijn bedacht waar we aan kunnen draaien teneinde een bepaald doel te bereiken. Kritische depositiewaardes (KDW's) zijn in dit wereldbeeld unieke en exacte getallen. Uitkomsten van Aerius/OPS berekeningen zijn reëel en worden zonder marges of onzekerheden gebruikt. De natuur zal herstellen als de depositie onder de KDW komt, en boven de KDW zal of kan stikstofdepositie 'significante' schade aan de natuur toebrengen.

Ik heb bijzondere aandacht geschonken aan de bepaling van deze KDW's en de daarmee samenhangende drempelwaardes. De statistiek kan niet zeggen wat 'kritisch' is, om verschillende redenen: er zijn grote methodologische problemen, en een beroep op het streven naar statistische significante veranderingen en/of een definitie in die termen heb ik gemotiveerd afgewezen. Statistische significantie heeft weinig te maken met dingen die voor ons of voor de natuur belangrijk zouden kunnen zijn.

De gedachte dat er een unieke en wel gedefinieerde KDW zou bestaan voor individuele habitats is onjuist. Het is evident dat er te veel stikstofdepositie kan zijn, voor elke redelijk criterium, om de simpele reden dat van alles te veel kan zijn. Echter, ik heb geen enkele wetenschapper gesproken die vanuit die wetenschap een unieke (exacte) waarde verdedigt. Er zijn wel wetenschappers die unieke waardes propageren omdat de politiek en de rechterlijke macht deze accepteert of zelfs eist. Dat heeft echter niets met wetenschap te maken; wetenschappers zouden een dergelijke verleiding moeten weerstaan. Het huidige gebruik van KDW's, in de NDA's, in de adviezen van de EA en in de vergunningverlening, is daarmee vanuit statistisch perspectief onwetenschappelijk, ondeugdelijk en dus niet verdedigbaar.

Vervolgens heb ik de onzekerheden en onbepaaldheden van het Aerius/OPS rekenmodel zelf bestudeerd. Zoals elk model geeft Aerius/OPS slechts terug wat de makers erin hebben gestopt. De prestaties van het model op individueel niveau bij vergunningverlening worden omgeven met onzekerheden die niet kwantificeerbaar zijn, omdat gemodelleerde (toename van) deposities te klein zijn. Dat de onzekerheden hier veel

groter zullen zijn dan bij landelijke (globale) gemiddeldes is een wiskundig en geen empirisch gegeven. Deze globale modeluitkomsten worden zelf al omgeven met tientallen procenten onzekerheid, voortkomend uit veelsoortige onzekerheden. Controle van modelprecisie en -onzekerheid is niet mogelijk omdat er geen meetdata voor stikstofdepositie voorhanden zijn – deze worden primair afgeleid uit atmosferische concentraties waarvan de onzekerheid zelf ook moeilijk te kwantificeren is. Voor alle duidelijkheid, stikstofdepositie wordt niet zélf gemeten.

Samengenomen concludeer ik dat we met de combinatie van de onzekerheden in de KDW's en de door het model berekende deposities van Aerius/OPS feitelijk niet weten wat we aan het doen zijn. We zijn blind modeluitkomsten aan het volgen, zonder de mogelijkheid te hebben om te controleren waar we mee bezig zijn. Er wordt veel beweerd over stikstof, maar er is vrijwel geen enkele kwantitatieve claim die daadwerkelijk controleerbaar is. Het werken met de combinatie KDW/Aerius acht ik vanwege dit alles niet wetenschappelijk. Uitspraken over proefondervindelijke stand van zaken met verstrekende juridische consequenties (stikstofemissie en –depositie) behoren namelijk verifieerbaar te zijn, en dat zijn ze niet.

Een antwoord dat dit “het beste is wat we hebben” is onbevredigend en onduidelijk omdat het uitgaat van de gedachte dat de wetenschap móet leveren. In plaats van net te doen alsof we precieze informatie hebben en daar beleid met zeer grote maatschappelijke, juridische en particuliere gevolgen op baseren, kunnen we beter accepteren dat we bepaalde dingen niet kunnen weten, en dat ‘de’ wetenschap hierop ook geen antwoord gaat geven. Wetenschap wordt in dit dossier overvraagd.

Ook langs de weg van de beginselen van behoorlijk bestuur laat ik zien dat modelgebruik in dit dossier problematisch is. De aard van wiskundige modellen maakt dat niet ingezien kan worden dat aan deze beginselen wordt voldaan in het huidige beleid.

Wetenschappers horen te interveniëren als hun resultaten incorrect geïnterpreteerd of zelfs misbruikt worden. Dat interveniëren gebeurt in dit discours te weinig of zelfs helemaal niet.

Samenvattend is de model-wetenschappelijke benadering van het stikstofdiscours zowel praktisch, theoretisch als filosofisch problematisch. We hebben een benadering nodig die vanuit de empirie werkt in plaats van langs de weg van wiskundige modellering; de Europese Habitatrichtlijn geeft ons die ruimte. Het dossier moet van het bord van de

modelleurs af, en terug op het bord van beleidsmakers. Deze dienen verantwoordelijkheid te nemen voor beslissingen op basis van een maatschappelijk en politiek debat in plaats van op basis van een onwetenschappelijke omgang met modellen. Wetenschap en bestuur dienen gescheiden te blijven; wetenschappers moeten hun bevindingen publiek maken zonder zich af te vragen wat de rechter ervan zal vinden. Ik doe een voorstel voor een aantal normerende regels waaraan modelgebruik gehouden zou moeten zijn, en sluit af met een stappenplan om dit alles in Nederland te realiseren binnen de mogelijkheden die de Europese regelgeving biedt.

1 - Inleiding en algemene beschouwing

De staatssecretaris van Landbouw, Visserij, Voedselzekerheid en Natuur heeft mij gevraagd om een notitie op te stellen waarin:

- gereflecteerd wordt op de wijze waarop in de huidige handreiking NDA's (en dus toegepast in de huidige NDA's) met de statistische relevantie van de berekende stikstofdepositie op Natura 2000-gebieden wordt omgegaan en aanbevelingen te doen om hier verbetering in aan te brengen ten behoeve van de nieuwe handreiking;
- het gebruik van stikstofmodellen in een historisch perspectief wordt geplaatst;
- de algemene principes rondom het gebruik van statistische modellen in hoofdlijnen worden geschetst.

Voor u ligt de notitie die ik hierover heb opgesteld. Uiteraard ben ik niet de eerste die vanuit de wetenschap een kritiek geeft op het dossier. De commissie Hordijk besloot al in 2020 tot de ongeschiktheid van het Aerius/OPS rekenmodel voor het berekenen van stikstofdepositie zoals dat gebruikt wordt voor vergunningverlening.² De onzekerheden in de modeluitkomsten zijn volgens die commissie veel te groot, zeker gezien de huidige drempelwaarde die bij het wel of niet verlenen van toestemming wordt gebruikt.

De Nieuwe Denktank schreef hierover enkele jaren later in 2023 het volgende:³

“De onzekerheden hebben tot gevolg dat het stikstofmodel, hoe gesofisticeerd het ook mag zijn, een toepassingsbereik heeft: uitkomsten buiten dit toepassingsbereik zijn zo onzeker en zeggen zo weinig over de werkelijkheid dat zij niet als basis kunnen dienen voor de besluitvorming. [...] Dat leidt tot schijnzekerheid, willekeur en rechtsongelijkheid.”

Over de Nederlandse praktijk wat betreft stikstofberekeningen zegt hetzelfde rapport dat deze⁴

² [Meer meten, robuuster rekenen | Rapport | Rijksoverheid.nl](#)

³ *Uit de stikstofcrisis – verantwoord omgaan met onzekerheid*, www.denieuwedentank.nl, blz. x.

⁴ *Ibid*, blz. xi.

“[...] niet getuigt van het zich beter of strikter houden aan Europese natuurwetgeving, het getuigt evenmin van het belang van de natuur. Het getuigt enkel van een onwetenschappelijke omgang met modeluitkomsten.”

Het rapport spreekt verder van een toestemmingsverlening in Nederland die “zich afspeelt in een parallel, modelmatig universum dat los staat van de werkelijkheid” en vindt Nederland “het onhandigste jongetje van de klas”.⁵

Dat zijn voorwaar allemaal geen milde oordelen. De Nieuwe Denktank doet ook suggesties over hoe we hier mee om zouden moeten gaan. Deze figureren allemaal binnen het huidige juridische kader, waarin het gebruik van het Aerius/OPS rekenmodel in de Omgevingsregeling is vastgelegd⁶ - daarover later meer. Kort gezegd komt het advies van de Nieuwe Denktank erop neer dat we een drempelwaarde zouden moeten kiezen tussen de 1 en 35 mol/ha/jaar in plaats van de huidige 0,005 mol/ha/jaar.

Wat betekent deze drempelwaarde? Welnu, Nederland hanteert een zogenaamde kritische depositie waarde (KDW) die ruwweg aan zou moeten geven hoeveel stikstofdepositie een bepaalde habitat kan verdragen zonder schade.⁷ Als deze KDW (bijna) is overschreden geldt een drempelwaarde voor stikstofdepositie voor activiteiten. In Nederland is deze drempelwaarde dus 0,005 mol/ha/jaar.

We hebben in principe de vrijheid om e.e.a. zo in te richten, maar ook om deze drempelwaarde te verhogen zoals De Nieuwe Denktank voorstaat. Immers, de Europese Habitatrichtlijn verplicht lidstaten om te onderzoeken of nieuwe plannen “significante” invloed kunnen hebben op Natura 2000-gebieden. Hoe die beoordeling plaatsvindt wordt aan de lidstaten overgelaten. Frankrijk en Italië bijvoorbeeld, kijken zelfs helemaal niet naar stikstofdepositie. Denemarken en Duitsland rekenen wel aan depositie, maar hun drempelwaardes zijn met respectievelijk 35 en 21 mol/ha/jaar onvergelijkbaar met die in Nederland. Volgens De Nieuwe Denktank en haar adviseur Arthur Petersen zou deze

⁵ Ibid, blz. xi.

⁶ Zie bijvoorbeeld Artikel 4.15 in [wetten.nl - Regeling - Omgevingsregeling - BWBR0045528](https://wetten.nl/-Regeling-Omgevingsregeling-BWBR0045528).

⁷ Ik zal straks preciezer worden.

benadering de vergunningverlening in Nederland weer op gang kunnen brengen, en is deze benadering volgens hen ook wetenschappelijk verantwoord.⁸

Het feit dat in sommige landen stikstofdepositie helemaal geen rol speelt in hun natuurbeleid is interessant. We spreken in Nederland vaak over het ‘stikstofbeleid’ maar strikt gesproken hebben we dat niet. We hebben ‘natuurbeleid’, en hebben ons verplicht om een aantal natuurgebieden te beschermen. Hoe we dat doen en wat we onder “beschermen” precies verstaan, is aan Nederland zelf. Er is geen Europese regel die ons dwingt om alleen (of überhaupt) naar stikstofdepositie te kijken. We hebben met stikstof een specifieke *maatstaf* gekozen die het beleid bepaalt. Deze keuze is natuurlijk niet neutraal, maar tegelijkertijd kunnen er goede redenen voor zijn.

Het feit dat een specifieke maatstaf vaak het beleid bepaalt staat bekend als de ‘Wet van Goodhart’.⁹ Volgens deze wet kunnen maatstaven heel makkelijk zomaar een eigen leven gaan leiden, en wordt het beleid feitelijk op de maatstaf gericht in plaats van op het oorspronkelijke beleidsdoel. In het natuurdebat lijkt dat in Nederland inderdaad gebeurd te zijn: de maatstaf ‘stikstof’ is de rol gaan vervullen van de oorspronkelijke doelstelling.

Is stikstof een goede maatstaf? Ik ben geen ecooloog dus in directe zin kan en zal ik daar weinig over zeggen. Maar in het debat over het natuur- en stikstofbeleid spelen onzekerheden en statistiek een grote rol, en het is dáár waar mijn bijdrage betrekking op heeft. Ook zonder specifieke ecologische kennis kan een statisticus de situatie beoordelen. Ik geef twee voorbeelden ter illustratie hiervan.

Allereerst is het in de praktijk samenvatten van een extreem gecompliceerde situatie in slechts één parameter (hier stikstofdepositie) in het algemeen zelden zinvol – er wordt te veel informatie en context weggelaten, en een statisticus heeft hier veel ervaring mee. Ten tweede is het zo dat een begrip als ‘statistische significantie’ (dat we vaak tegenkomen wanneer het over veranderingen in het veld gaat) in zichzelf niet veel betekenis heeft. Het is namelijk zo dat statistisch significante resultaten in het veld irrelevant kunnen zijn. Dat punt

⁸ Zie ook Arthur Petersen, *Expertoordeel onderbouwing maximale rekenafstand bij project-specifieke berekeningen van stikstofdeposities*, geschreven in opdracht van het Ministerie van Infrastructuur en Waterstaat. 9 november 2022; https://www.ucl.ac.uk/steapp/sites/steapp/files/dutch_-_expert_judgement_arthur_petersen_11-11-202235.pdf.

⁹ Zie bijvoorbeeld D. Manheim en S. Garrabrant, *Categorizing variants of Goodhart’s law*, arXiv:1803.04585, 2018 en S. Kerr, *On the folly of rewarding A, while hoping for B*, The Academy of Management Executive 9, 1995.

adresseer ik straks uitgebreid; langs deze weg kan begrepen worden dat de huidige statistische omgang met stikstof geen goede maatstaf kán zijn.

In deze notitie onderzoek ik een aantal zaken in relatie tot statistische onzekerheid van stikstofdepositie en de bijbehorende vergunningverlening in de context van de natuurdoelanalyses (NDA's) en de corresponderende adviezen van de Ecologische Autoriteit (EA). Je zou dit rapport dan ook kunnen zien als een aanvulling en een vervolg op de rapportage van de commissie Hordijk, met de nadruk op de vraag hoe we de modeluitkomsten en berekeningen statistisch moeten duiden; vanuit het perspectief van onzekerheid dus.

Het rapport is als volgt opgebouwd. Allereerst heb ik onderzocht hoe in de huidige NDA's en de bijbehorende adviezen van de Ecologische Autoriteit omgegaan wordt met onzekerheid van de berekeningen van het Aerius/OPS rekenmodel en de 'kritische depositie waardes' (de KDW's) en drempelwaardes. Dat beschrijf ik in Hoofdstuk 2.

In Hoofdstuk 3 vindt u het verslag van mijn onderzoek naar KDW's. Waar komen die vandaan, wat betekenen ze, hoe worden ze gebruikt en welke onzekerheden worden daar wel of niet in meegenomen? Daarna, in Hoofdstuk 4, richt ik mijn aandacht op het Aerius/OPS rekenmodel zelf, en onderzoek ik de onzekerheden in dat model en de manier waarop het model wordt gebruikt. De combinatie van de KDW's, drempelwaardes en het Aerius/OPS model bepaalt de gang van zaken van de vergunningverlening in Nederland, en mijn onderzoek leidt tot een totaalbeeld over de onzekerheden die met het rekenen met Aerius/OPS in combinatie met vergunningverlening gepaard gaan.

Vervolgens ga ik in Hoofdstuk 5 in op de aard van modellen in het algemeen, en ga ik na hoe de beginselen van behoorlijk bestuur richtinggevend zijn voor het werken met wiskundige modellen, en met Aerius/OPS in het bijzonder. Mijn onderzoek laat hierbij heldere grenzen zien voor wat betreft de vraag wat we wel en wat we niet aan een model kunnen uitbesteden.

Het werken met modellen in een beleidsomgeving is in het algemeen geen sinecure. In een eerder publicatie¹⁰ hebben Marc Jacobs en ikzelf een lijst van aandachtspunten samengesteld die bij elk modelgebruik belangrijk zijn. Ik zal deze lijst hier in Hoofdstuk 6

¹⁰ Marc Jacobs en Ronald Meester, *Van aardbeving tot zoönose – over de inzet van modellen voor beleid*, Mazirel Pers 2023.

gebruiken om uit te leggen wat in het algemeen de gevaren zijn van het gebruik van modellen voor beleid en deze toepassen op dit dossier.

Ik eindig in Hoofdstuk 7 met een samenvatting en een advies over het gebruik van Aerius/OPS, drempelwaardes en KDW's in relatie tot de NDA's, de adviezen van de EA en de vergunningverlening.

2 - De NDA's en de adviezen van de EA

Ten behoeve van dit rapport heb ik 133 NDA's met corresponderende adviezen van de EA bestudeerd. De studie is naar de EA-adviezen is gedeeltelijk geautomatiseerd uitgevoerd. Ik was vooral geïnteresseerd in het 'wereldbeeld' dat spreekt uit de NDA's en de adviezen; hieronder wordt duidelijk wat ik daarmee bedoel. Ik bespreek eerst de NDA's en daarna de adviezen van de EA.

De NDA's

Bij het opstellen van NDA's worden Aerius berekeningen veelvuldig en routinematig gebruikt. De uitkomsten van deze berekeningen en ook de waarden van de KDW's worden hierbij letterlijk genomen; er zijn vrijwel geen opmerkingen die blijken geven van een bepaalde mate van bewustzijn over onzekerheid hierover. Ik zal deze beweringen staven met enkele representatieve citaten.

"De kamsalamander is gevoelig voor stikstof wanneer door eutrofiëring zuurstofgebrek optreedt in zijn leefgebied, habitatype H3130. Dat gebeurt pas bij een hogere depositie dan de KDW van het habitatype H3130 (571 mol N/ha/jaar)."¹¹

"De kritische depositiewaarde (KDW) voor het habitatype Stuifzandheiden, Droge heiden en voor Zandverstuivingen is 714 mol N/ha/jaar."¹²

"De kritische depositiewaarde van Zwakgebufferde vennen voor stikstof is 500 mol/ha/jr."¹³

"Wanneer er teveel stikstof valt, zal de kwaliteit en uiteindelijk ook het oppervlakte van de habitattypen afnemen. Voor ieder habitatype en leefgebied is landelijk een zogeheten Kritische Depositie Waarde (KDW) vastgesteld. Valt er meer stikstof dan

¹¹ NDA Korenburgerveen.

¹² NDA Maasduinen.

¹³ Ibid.

de KDW, dan gaat het habitatype er op termijn in kwaliteit en oppervlakte op achteruit.”¹⁴

“De hierna getoonde prognoses zijn verkregen door het gebruik van AERIUS Monitor 2022. De getoonde percentages geven aan welk deel van het totale gekarteerde oppervlakte overbelast is, dat wil zeggen waar de stikstofdepositie boven de KDW (Kritische Depositiewaarde) ligt.”¹⁵

“De KDW van het habitatype H2330 van 714 mol N/ha/jr wordt zowel in het referentiejaar 2020 alsook in de prognose van 2030 overschreden.”¹⁶

“Als gevolg van overschrijding van de KDW bestaat het risico dat de kwaliteit van het habitatype wordt aangetast door verzuring en/of vermesting door stikstofdepositie, met mogelijke gevolgen voor bodem en waterkwaliteit.”¹⁷

“De KDW is de grens waarboven schade aan het habitatype optreedt.”¹⁸

“Bij een toename van de stikstofdepositie boven de kritische depositiewaarde kan de veenmosvegetatie niet alle stikstof meer vastleggen; het ‘veenmosfilter’ is dan verzadigd geraakt met stikstof.”¹⁹

“De kritische depositie waarde (KDW) wordt daarbij als volgt gedefinieerd: ‘De grens waarboven het risico bestaat dat de kwaliteit van het habitat significant wordt aangetast door de verzurende en/of vermestende invloed van atmosferische stikstofdepositie.’”²⁰

¹⁴ NDA Wijnjeterperschar.

¹⁵ NDA Eilandspolder.

¹⁶ NDA Brabantse Wal.

¹⁷ Ibid.

¹⁸ NDA Bargerveen; deze definitie is overigens onjuist.

¹⁹ Ibid.

²⁰ NDA Brabantse Wal.

Het beeld is hiermee wel duidelijk. De KDW wordt zonder omhaal gebruikt als een unieke waarde; merk op dat bijvoorbeeld in het laatste citaat gesproken wordt over “de” grens waarboven iets gebeurt. Onder de KDW is er binnen deze visie geen probleem, daarboven wel. Die grenzen zijn heel expliciet en numeriek. In de NDA Buurserzand en Haaksbergerveen vinden we bijvoorbeeld KDW’s ter grootte van respectievelijk 1071, 500, 1429, 1143, 714, 1214, 571, 1768 en 1857 mol/ha/jaar, allemaal voor verschillende soorten habitats. Dat deze KDW’s echt letterlijk genomen worden, wordt in vele NDA’s expliciet bevestigd. In dezelfde NDA als zojuist lezen we bijvoorbeeld dat

“De langdurige overschrijding van de KDW heeft geleid tot een ernstige aantasting van de structuur en het functioneren van Natura 2000-habitats, maar ook van buiten Natura 2000-gebieden gelegen stikstofgevoelige natuur.”

Voor de verantwoording van deze en dergelijke claims wordt veelal verwezen naar artikelen uit de wetenschappelijke literatuur.²¹ Desalniettemin bevestigde een van de auteurs van dit artikel, ecooloog Roland Bobbink, dat KDW’s “ranges” zijn, en “geen unieke waardes”.²² Dat is uiteraard een veel plausibelere gedachte dan unieke waardes.

Bij de beoordeling van de vraag of de KDW’s overschreden worden, wordt geheel in lijn met de eerder genoemde Omgevingsregeling het rekenmodel Aeries/OPS gebruikt. Dat mag geen verbazing wekken. Sterker nog, het wordt voorgeschreven door diezelfde Omgevingsregeling. Het is daarom niet nodig om dit feit met een lange lijst van citaten te illustreren. Het volgende citaat is exemplarisch:

“Bij de beoordeling van de mate van (over)belasting door stikstofdepositie op de instandhoudingsdoelen is gebruik gemaakt van AERIUS Monitor 2022.”²³

Aeries/OPS is een deterministisch model: als je er twee keer hetzelfde ‘instopt’ komt er ook twee keer hetzelfde uit. Toch lijkt het dat in sommige NDA’s bepaalde

²¹ Bijvoorbeeld naar Bobbink, R., G. van Dijk, E. Remke & H. Tomassen (2022). *Herstelbaarheid van door stikstofdepositie aangetaste Natura 2000-habitattypen: een overzicht*. Onderzoekcentrum B-WARE, Nijmegen. Rapportnummer RP-21.117.21.95.

²² Persoonlijke communicatie.

²³ NDA Brabantse Wal.

onzekerheidsmarges worden gerapporteerd. In de NDA Landgoederen Brummen vinden we bijvoorbeeld deze figuur:



Er wordt bij de figuur opgemerkt dat dit de gemiddelden en de bijbehorende 10% en 90% percielen zijn. Dit lijkt in eerste instantie een uitspraak over onzekerheid te zijn, maar dat is toch niet zo. Wat hier gebeurt is dat gekeken wordt naar ‘hexagonen’ (dat zijn zeshoekige stukjes land waarin een gebied is ingedeeld). Bijvoorbeeld, in 2018 geldt dat Aerius zegt dat in 10% van de hexagonen de depositie lager is dan 1271 mol/ha/jaar en in 10% hoger dan 1956 mol/ha/jaar. Het is dus geen onzekerheid in de Aeriusberekening zelf, maar een observatie over het geheel van (deterministische) deposities op alle relevante hexagonen.

De adviezen van de EA

Alvorens na te gaan hoe de EA adviseert en welke taal men daarvoor gebruikt eerst een paar woorden over de EA zelf. Volgens het instellingsbesluit²⁴ is een van haar taken

“te adviseren over relevante wetenschappelijke inzichten voor besluitvorming ten aanzien van brede ecologische vraagstukken, met inbegrip van vraagstukken over klimaat en abiotische randvoorwaarden voor natuur zoals bodem, water en atmosferische depositie.”

²⁴ [Instellingsbesluit Ecologische Autoriteit.pdf](#), Artikel 3b.

Hoewel de EA naar eigen zeggen zelf geen onderzoek doet is het natuurlijk niet mogelijk om te adviseren over wetenschappelijke vraagstukken zonder die wetenschap helemaal te doorgronden, in alle facetten die nodig zijn voor haar taak. In die zin is de EA een wetenschappelijk instituut, of zou het moeten zijn.

Elke wetenschapper dient zich te verhouden tot onzekerheid, zeker wanneer er kwantitatieve beweringen en uitspraken worden gedaan aan de hand van modellen. Om te onderzoeken in welke mate de EA-adviezen aandacht besteden aan het thema onzekerheid in de context van Aeries, heb ik 133 documenten met EA-adviezen geanalyseerd, met behulp van de AI software van Elicit. Deze documenten zijn als PDF gedownload van de publieke website van de Ecologische Autoriteit.²⁵ Binnen Elicit gebruikten we drie zoekprompts:

1. “Wordt er expliciet gereflecteerd op onzekerheid of betrouwbaarheid van metingen en berekeningen specifiek door het rekenmodel Aeries, inclusief termen zoals onzekerheid, bandbreedte, significantie, skill, prestatie, betrouwbaarheid of bias? Het gaat hierbij uitsluitend om reflecties op het Aeries-model zelf.”
2. “Expliciete reflecties op Aeries in de context van onzekerheid, bijvoorbeeld aan de hand van termen zoals onzekerheid, bandbreedte, significantie, skill, prestatie, betrouwbaarheid of bias”
3. “Statistische onzekerheid Aeries”

Deze zoekopdrachten leverden respectievelijk 0, 16 en 8 documenten met matches op. Na validatie bleek echter dat geen van deze resultaten daadwerkelijk aansloot op de onderzoeksvraag. Het beperkte aantal bruikbare resultaten vormde aanleiding om een diepgaandere analyse uit te voeren met behulp van een geavanceerder Large Language Model (LLM).

Doel van deze uitgebreide analyse is om zoveel mogelijk kandidaat-citaten te identificeren die potentieel relevant zijn voor de onderzoeksvraag, om deze citaten daarna zelf te beoordelen. Hiervoor hebben we een analysemethode ontwikkeld waarin gebruik wordt gemaakt van de meest geavanceerde OpenAI GPT modellen. Deze methode

²⁵ ecologischeautoriteit.nl.

combineert geautomatiseerde AI-tekstanalyse met menselijke validatie en omvat de volgende drie stappen:

1. Extractie van de tekst uit de EA-advies PDF documenten.
2. Tekstanalyse met een AI-zoekopdracht (prompt)
3. Menselijke beoordeling van de LLM output op relevantie voor de onderzoeksvraag.

Voor de technische details van deze procedure verwijs ik naar Appendix A.

Ik verwijs naar Appendix B voor de volledige collectie van geselecteerde citaten. Hieronder vat ik de citaten samen - aan de hand van voorbeelden. De 'taalomgeving' is hierbij van belang, omdat ze iets vertelt over de filosofie van de EA. Taal serieus nemen is voor mij een belangrijk hermeneutisch principe.

De EA adviseert kennelijk vanuit een deterministisch wereldbeeld, zoals Aerius inderdaad een deterministisch model is. Er is geen spoor te vinden van onzekerheid in Aerius-uitkomsten, de KDW's en de consequenties die deze uitkomsten hebben op de onderzochte ecologie. Er wordt dus, metaforisch gesproken, een ecologie getoond waaraan 'knoppen' zijn verbonden waar naar geloven aan kan worden gedraaid, met bijbehorende schijnbaar voorspelbare ecologische uitkomsten. Dat is in zichzelf een boeiende constatering die ik aan de hand van een groot aantal voorbeelden zal staven. Na elk voorbeeld volgt mijn commentaar.

Voorbeeld 1: "Daarbij wordt het voorbehoud gemaakt dat er daadwerkelijke stikstofreductie plaatsvindt zoals in het AERIUS-model wordt voorspeld, en dat bij het terugdringen van de nitraatuitspoeling in landbouwgronden (met groot effect op habitats) er prioriteit wordt gegeven aan een stevige en voortvarende aanpak in deelgebied Noord." (5042 Bunder- en Elslooërbos, blz. 6)

Het bijvoeglijk naamwoord "daadwerkelijk" is empirisch van aard, maar verwijst hier naar het Aerius-model dat blijkbaar (wiskundig en met de input van NEMA²⁶ getallen en KDW) voorspellend te werk kan gaan. Dat leidt tot een verwarrende zinssnede die een empirische 'hardheid' suggereert die niet kan bestaan gezien de modelmatige verwijzing. De

²⁶ Het National Emission Model for Agriculture, zie Hoofdstuk 4.

“voortvarende aanpak” onderstreept deze tweeslachtigheid.

Voorbeeld 2: “De NDA analyseert de effecten van de huidige en toekomstige stikstofdepositie. De NDA beschrijft dat in 2020 voor vrijwel alle habitattypen sprake was van een overschrijding van de kritische depositiewaarden (KDW’s). Alleen op een deel van het habitatype alluviale bossen is geen sprake van overschrijding. Voor 2030 verwacht de NDA, op basis van Aeries Monitor, dat de stikstofdepositie op 100% van het oppervlak van zwakgebufferde vennen (tevens leefgebied van kamsalamander), droge heiden, heischrale graslanden en blauwgraslanden nog steeds boven de KDW ligt. Voor de andere habitattypen is er sprake van een gedeeltelijke overschrijding van de KDW’s, of is geen sprake meer van overschrijding.” (5126 Stelkampsveld, blz. 11)

Hier wordt een vergelijkbare deterministische ecologie geventileerd. De zin “Alleen op een deel van het habitatype alluviale bossen is geen sprake van overschrijding” straalt een grote stelligheid uit. Die stelligheid wordt tweeledig als er wordt opgemerkt dat “op basis van AeriesMonitor, dat de stikstofdepositie op 100% van het oppervlak van zwakgebufferde vennen (tevens leefgebied van kamsalamander), droge heiden, heischrale graslanden en blauwgraslanden nog steeds boven de KDW ligt.” Zowel Aeries als KDW zijn empirisch en modelmatig eenduidig, aldus de EA, waarmee beleid kan worden vastgesteld en uitgevoerd.

Voorbeeld 3: “Voor een analyse van de veranderingen in atmosferische stikstofdepositie is het nuttig om ook de metingen van de ammoniakconcentratie in de lucht van het RIVM in een grafiek te tonen en te bespreken (RIVM-MAN16). Dit is geen essentiële ecologische informatie voor deze versie van de NDA, maar wordt geadviseerd in de volgende versies mee te nemen. Beschrijf daarin wat de relatie is tussen de uitkomsten van stikstofdepositie uit AERIUUS en de ammoniakmetingen van het RIVM.” (5087 Drentsche Aa-gebied, blz. 11)

Op welke wijze MAN-data met Aeries moet wordt vergeleken is een vraag op zich, waarbij onzekerheid een cruciale rol speelt. Hoe de ruwe MAN-gegevens (die niet openbaar zijn) worden omgezet in atmosferische concentraties is een belangrijke vraag. Ook de koppeling van de bewerkte MAN-data met Aeries uitkomsten moet worden besproken. Dit alles is voor mij op dit moment niet na te gaan, maar dat hierbij ook (grote) onzekerheid een rol speelt

leidt geen twijfel.

Voorbeeld 4: “Verdere verslechtering moet ook worden voorkomen door beperking van de stikstofdepositie. De huidige depositie ligt in het gebied nu tussen 1497 en 1936 mol N/ha/jaar. De meest kritische depositiewaarde (van H3130 Zwakgebufferde vennen) is veel lager, namelijk 500 mol N/ha/jaar (dit betreft de aangepaste KDW, zie tekstkader hieronder). De Ecologische Autoriteit rekent uit dat de stikstofdepositie dus met 75% omlaag moet om de natuurdoelen te behalen. De NDA presenteert alleen de KDW en huidige (en verwachte) deposities, maar verbindt hier verder geen (gekwantificeerde) conclusies aan.” (5083 Lonnekermeer, blz. 10)

De EA interpoleert lineair tussen KDW en Aeries calculaties zonder enige reflectie op onzekerheden in beide. Daarnaast wordt KDW hier specifiek opgevat als een harde grens. Eén van de definities van de KDW luidt echter als volgt: “De kritische depositie is een kwantitatieve schatting van de blootstelling aan één of meer verontreinigende stoffen, waar beneden geen significante schadelijke effecten optreden aan gespecificeerde gevoelige elementen in het milieu, volgens de huidige stand van kennis.”²⁷ Hoewel deze definitie tal van zaken ontbeert - zoals een tijdsschaal waarop schade van een bepaald gedefinieerde soort zichtbaar moet worden - kan het absoluut niet als een harde grens worden opgevat zoals de EA hier poneert.²⁸

Voorbeeld 5: “Voetnoot 7: Voor het bepalen in welke mate stikstofdepositie een drukfactor is kan gebruik gemaakt worden van de AERIUS-calculator. Kijk daarbij ook naar verschillen in stikstofdepositie op hexagoonniveau. Die verschillen kunnen namelijk relevant zijn bij de keuze voor herstelmaatregelen (hoe groter de overschrijding van de kritische depositiewaarde, hoe ingrijpender de maatregelen die nodig zijn). Beoordeel op basis van gebiedskennis ook de mate waarin historische belasting van stikstof een beperking vormt voor de ontwikkeling van de habitat en het leefgebied. Bedenk dat de effecten van stikstofdepositie door droogte nog versterkt kunnen worden. In het advies van de TEO wordt ingegaan op relevantie van verschillende overschrijdingsklassen van de kritische

²⁷ <https://edepot.wur.nl/245248>

²⁸ Zie het volgende Hoofdstuk.

depositiewaarde (KDW) die bij de beoordeling van deze drukfactor relevant zijn.” (5007 Broekvelden, blz. 6)

Hexagoon stikstofdepositie wordt zonder reflectie als maat genomen ten opzichte van andere hexagoon stikstofdeposities en de keuze voor herstelmaatregelen in relatie tot KDW: “hoe groter de overschrijding van de kritische depositiewaarde, hoe ingrijpender de maatregelen die nodig zijn”. Deze formulering verradt een rotsvast vertrouwen in de gedachte dat een overschrijding van de KDW een maat is voor de actie die ondernomen zou moeten worden. Maar zoals ik later zal laten zien is dat onjuist, niet alleen omdat onzekerheden verbloemd worden maar ook omdat de statistiek en de modelering in zichzelf geen enkele uitspraak kunnen doen over wat zou moeten gebeuren en op basis waarvan, nog los van de onzekerheden in de resultaten van genomen maatregelen.

Voorbeeld 6: “Hoogveen heeft een zeer lage kritische depositiewaarde van 500 mol N/ha/jaar. Deze wordt momenteel sterk overschreden: de huidige stikstofdeposities zijn gemiddeld 1283 en 1428 mol N/ha/jaar voor actieve en herstellende hoogvenen. Dit is in de vegetatie te zien in de vorm van vergrassing, toename van adelaarsvaren en opslag van bomen. De atmosferische stikstof komt nog bovenop de stikstof die als gevolg van interne eutrofiering door verdroging (zie hierboven). De NDA presenteert alleen deze cijfers, maar maakt vervolgens niet expliciet duidelijk wat de totale benodigde stikstofreductie in het gebied is. De Ecologische Autoriteit rekent zelf uit dat de KDW dus met een factor 2,5 tot 3 wordt overschreden. Dat betekent dat er een reductie nodig is van zo’n 65%.” (5077 Aamsveen, blz. 12)

Dit voorbeeld is vergelijkbaar met Voorbeeld 4. Er wordt door de EA eenzelfde lineaire vergelijking gemaakt tussen KDW, stikstofdepositieberekeningen door Aerius en de gedachte stikstofreductie. Ook hier laat de EA geen bewustwording zien van de experimentele realiteit van KDW-bandbreedtes, onzekerheden daarin en alle onzekerheden in Aerius modelleringen.

Voorbeeld 7: “Ondanks het ontbreken van gegevens, kunnen sommige conclusies wel degelijk al getrokken worden, bijvoorbeeld omdat de mate van onzekerheid kleiner is dan de

omvang van het effect.” (Meerdere EA documenten, zoals bijvoorbeeld 5008 Oldenzaal)

De vraag is hoe dat kan, in de context van (i) ontbrekende gegevens, (ii) de bijbehorende maatregel-onzekerheden, en (iii) de niet in kaart gebrachte onzekerheden van stikstofdepositiemodellering en KDW onzekerheden. De hoop is waarschijnlijk dat de mate van onzekerheid kleiner is dan de omvang van het effect, maar dan moeten beide tot op zekere hoogte fundamenteel kenbaar kunnen zijn en binnen grenzen in beeld zijn gebracht.

Voorbeeld 8: “Uit de natuurdoelanalyse blijkt dat op vrijwel alle habitattypen en leefgebieden in het Natura 2000-gebied sprake is van overbelasting met stikstof. De Ecologische Autoriteit constateert dat de stikstofbelasting per habitatype in de NDA niet volledig is beschreven. Zo is niet duidelijk voor welk jaar de stikstofbelasting is beschreven en of er sprake is van lichte, matige of sterke overbelasting door stikstof. Bovendien ontbreekt voor sommige habitattypen een prognose van de situatie in 2030. Dit is van belang omdat voor het vaststellen van de omvang van aanvullende bronmaatregelen de achtergronddepositie in 2030 de referentie is.” (5111 Dwingelderveld, blz. 9)

De EA poneert hier zekerheden over stikstofdeposities per habitatype. De Aeries berekeningen die betrekking hebben op geprojecteerde stikstofemissies voor 2030 worden zonder enige twijfel geaccepteerd.

Voorbeeld 9: “Door B-Ware is in 2022 voorgesteld om de KDW van habitattypen van de lage tot hoge kwelder met een derde te verlagen. Dat geldt in ieder geval voor het habitatype Schorren en zilte graslanden (H1330). Stikstof is voor dit habitatype dus mogelijk een hogere drukfactor dan op basis van de gebruikte versie van AERIUS geconcludeerd wordt.” (5093 Greveling, blz. 12)

Aeries wordt ook hier contrafeitelijk ingezet: - *als* de KDW lager moet zijn *dan* is de stikstof drukfactor potentieel hoger. De EA lijkt niet geïnteresseerd te zijn in de vraag hoe nauwkeurig Aeries en de KDW's feitelijk kunnen zijn.

Voorbeeld 10: “Onlangs zijn de KDW's voor verschillende habitattypen en leefgebieden van

soorten op basis van nieuw wetenschappelijk onderzoek nader ingevuld. Dit is verwerkt in de nieuwe versie van AERIUS Monitor (5 oktober 2023), samen met onder andere de nieuwe cijfers over de totale stikstofdepositie op Natura 2000-gebieden. De nieuwe KDW's en nieuwe cijfers over de totale depositie waren op het moment van opstellen van de NDA nog niet beschikbaar. Toch zijn deze relevant voor de actuele opgave voor dit gebied.” (5043 Geleenbeekdal, blz. 12)

Ondanks deze voorbeelden is er in de adviezen van de EA sporadisch sprake van de herkenning van onzekerheden. Hier is een voorbeeld:

Voorbeeld 11: “De grootste opgave tot het komen tot systeemherstel is het reduceren van stikstofdepositie (via lucht en via water) dat leidt tot vermessing en verzuring. Er zijn daarom bronmaatregelen nodig die de stikstofuitstoot flink verlagen. De Ecologische Autoriteit onderschrijft daarbij dat de kleine daling die wordt voorspeld in het AERIUS-model grote onzekerheden heeft. Ook de NDA geeft terecht aan dat de afgelopen jaren de feitelijke daling van stikstofdepositie stagneert. Er zijn daarom (aanvullende) maatregelen nodig om de stikstofbelasting sterk te reduceren. Ook als stikstofreductie is bereikt, zal er nog vele jaren sprake van na-ijleffecten zijn. Geef aan in welke mate de stikstofbelasting moet worden teruggedrongen.” (5025 Bemeler- en Schiepersberg, blz. 8)

Het is onduidelijk over welke onzekerheden het hier gaat, en de argumentatie is moeilijk navolgbaar. Er wordt wel over onzekerheden gesproken maar er lijkt geen consequentie aan verbonden te worden.

Tenslotte merk ik op dat er bij de adviezen van de EA sprake lijkt van een zekere tunnelvisie. Ook als Aeries géén overschrijding van de KDW vaststelt, worden de gevolgen in het veld tóch aan stikstofbelasting toegeschreven. Hier is een voorbeeld.

Voorbeeld 12: “De NDA stelt dat stikstofdepositie, ten opzichte van andere drukfactoren, voor het gebied maar een beperkt knelpunt is. Ook hexagonen die in Aeries geen overschrijding van de KDW laten zien blijken in het veld wel degelijk de negatieve effecten van stikstofoverbelasting, zoals verzuuring, te ondervinden.” (5094 Krammer-Volkerak, blz. 12)

De modelmatige uitkomsten van Aerius per hexagoon geeft een ‘onderschrijding’ aan van de KDW die desalniettemin in twijfel wordt gebracht door de gesignaleerde “verruiging” die wordt gezien als een stikstofoverbelasting. Dat is verwarrend omdat in dit geval Aerius geen uitsluitsel geeft over stikstofbelasting; de gesignaleerde “verruiging” wordt één-op-één gekoppeld aan stikstofoverbelasting zonder duidelijke argumentatie.

Verificatie en conclusie

Ter illustratie en bevestiging van onze bevindingen hebben we een ChatGPT prompt ontworpen om het beeld dat we hebben te verifiëren. De “citaten” waarover gesproken wordt zijn de citaten die onze eerdere zoektocht hebben opgeleverd.

“Dit is een overzicht van citaten uit ecologische autoriteit adviezen die inzicht geven in hoe omgegaan wordt met statistische onzekerheid en het AERIUS model. Doe een grondig onderzoek en generaliseer de top 10 unieke meest voorkomende thema's over onzekerheid en kwantificeer aantal citaten in thematieken, en ook in hoeveel unieke EA adviezen + het percentage van de adviezen het voorkomt. Vul de sheet aan met deze gegevens en output het bestand in een downloadable csv bestand. Geef ook een oordeel over omgang met statistische onzekerheid van de EA. Geef ook aanvullende informatie of opvallendheden terug. In elke bewering moet je ook aantallen kwantificeren. Ga door totdat je de hele csv gevuld hebt en een betrouwbaar resultaat hebt.”

Uit het antwoord bleek dat het belangrijkste thema “KDW and Aerius” was. Maar één van de gerapporteerde 10 thema's betrof “Lack of reflection on statistical uncertainty in model outputs”, en daar trof ik deze passage aan:

“Strikingly, almost none of the NDAs or EA discussions explicitly address the statistical uncertainty of the AERIUS model results (e.g. no confidence intervals, no error margins). This absence is so consistent that it becomes a theme in itself – essentially a systematic omission. In only a couple of advisories did the EA point out that there was no reflection on the reliability or uncertainty margins of the calculated deposition. In those cases the EA clearly states that the NDA failed to discuss

statistical uncertainties (implicitly suggesting this is something that should be included). The low frequency of this theme shows that statistical uncertainty is rarely made explicit in these documents.”

Deze samenvatting is consistent met onze eigen bevindingen. Ik concludeer dat de EA adviseert vanuit het perspectief van vrijwel totale maakbaarheid, specifiek langs de lijn van stikstofdepositie: niet alleen zijn de KDW's duidelijke en unieke getallen, maar ook maakt Aerius/OPS precies duidelijk aan welke metaforische knoppen gedraaid moet of kan worden om een bepaald doel te bereiken. De vraag die voorligt is of dit wereldbeeld gerechtvaardigd is.

Ik merk op dat deze kennelijke maakbaarheid blijkt uit wat de EA zélf opschrijft. In mijn gesprek met de EA bleek echter dat de EA de KDW's veel minder letterlijk zou willen nemen dan uit hun schrijven blijkt, en dat ze de KDW's eerder als ordegroottes zou willen zien. Dat moge zo zijn, maar dat blijkt zoals gezegd niet uit hun publikaties. Ik schreef al eerder dat ik als hermeneutisch principe taal serieus neem: er staat wat er staat en ik lees niet tussen de regels door.

De komende twee hoofdstukken gaan nu dieper in op respectievelijk de KDW's en Aerius/OPS. Ik zal laten zien dat het deterministische en maakbare wereldbeeld zoals dat uit de adviezen van de EA opdoemt niet onderbouwd kan worden.

3 - KDW's, statistiek en significantie

KDW's zijn ranges, geen unieke waarden

Het is duidelijk dat de KDW's een cruciale rol spelen in de hele problematiek. Deze waarden bepalen immers of een bepaalde activiteit met een gegeven berekende depositie wel of niet kan plaatsvinden. In Nederland gebruiken we voor de KDW's unieke waarden, dat wil zeggen dat berekende deposities onder of boven die waarde liggen. Echter, elke wetenschapper begrijpt dat een dergelijk scherp omslagpunt niet kan bestaan, en dat KDW's in feite 'ranges' zijn – ik schreef dat al eerder. Toch produceren wetenschappers wel degelijk unieke waarden. In Van Dobben et al. lezen we bijvoorbeeld:²⁹

“Critical loads for nitrogen are increasingly used as standards in legislation. Standards have to be unique values, whereas critical loads are usually ranges. We produce unique values by combining simulated and empirical values.”

In Bobbink et al. lezen we:³⁰

“The procedure yielded N critical loads for most Annex I habitat types in the Netherlands. In most cases, empirical ranges and/or reliable model estimates were available, and critical loads could be set based on published information. In about 70 per cent of all the habitats, the models yielded critical load values within the given empirical range. In those cases where simulated critical loads were outside the empirical range, the critical load was set to the nearest extreme of the CL_{emp} N values. The difference between the middle of the CL_{emp} N range and the modelled values per habitat type was on average less than 1 kg/ha/year 60 out of the 75 habitat types found in The Netherlands were sensitive to N deposition (CLN < 34 kg N/ha/year). These critical loads for N deposition are the given standard in The Netherlands, and are also accepted by the government and by the judges of the highest administrative court. And because these critical loads have unique values

²⁹ H.F. van Dobben et al., *Revision of nitrogen critical loads for Nature 2000 Habitat types in The Netherlands*, Science of Total Environment 974,(2025) 179203 (mijn cursivering).

³⁰ Roland Bobbink et al., *Review and revision of empirical critical loads of nitrogen for Europe*, blz. 326 -327, 2022; [Publikationen | Umweltbundesamt](#), (mijn cursivering).

instead of ranges, it is clear whether or not the critical load of a specific habitat is exceeded. The habitat-specific critical loads are used together with detailed maps of habitat occurrences and detailed deposition maps to reveal exceedances across Natura 2000 nature reserves.”

De citaten zijn veelzeggend. Alle auteurs zijn zich wel degelijk bewust van onzekerheid in de KDW's of beter gezegd: van het feit dat 'de' KDW niet kan bestaan. Echter, uit de (door mij) gecursiveerde gedeeltes blijkt dat deze wetenschap bij de toepassing ervan de facto onmiddellijk overboord gezet wordt, en dat de KDW's in de praktijk unieke waarden krijgen, zonder ranges en zonder onzekerheidsmarges. Bovendien wordt kennelijk geaccepteerd dat rechters en overheid in Nederland deze unieke waarden gebruiken, en probeert men het bevoegd gezag van dienst te zijn door dergelijke unieke waarden te produceren.

De geschiedenis van dit dossier verraadt dat de gekozen wetenschappelijke aanpak in ieder geval mede het gevolg is geweest van de gekozen beleidsrichting. Al in de Hinderwet uit 1981 komen afstandstabellen voor in het kader van depositiewaarden. Om beleid te maken vanuit de filosofie van deposities is een bepaalde kwantificering nodig waar de wetenschap voor moest zorgen. De vraag is dan: kan de wetenschap dat?

Een van de definities van de kritische depositiewaarde luidt als volgt: “de hoeveelheid depositie die een ecosysteem over langere tijd kan verdragen zonder dat significante schade optreedt aan de structuur of het functioneren van dat systeem.”³¹ Hoewel er zoals gezegd zeker bewustzijn is dat de KDW geen uniek getal kan zijn (voor een gegeven ecosysteem en voor een gegeven soort depositie), is er nog een ander, fundamenteeler, probleem met deze en gelijksoortige definities. Wat hier bedoeld wordt met “significante” schade is namelijk niet erg duidelijk, want er zijn twee heel verschillende manieren om dat woord “significant” te verstaan. Ik zal dit nu uitgebreid bespreken omdat het (verkeerd) gebruik van het woord “significant” exemplarisch is voor de technocratische manier waarop Nederland met de stikstofproblematiek omgaat. Het volgende stuk is niet bedoeld om het belang van goed natuurbeleid te ontkennen of te frustreren. Integendeel: het technisch gebruik van het begrip “significant” kan juist schadelijke effecten hebben, ook of juist op natuurbehoud, die we allemaal liever zouden vermijden.

³¹ [Kritische depositiewaarde - Wikipedia](#)

Significantie

“Significant” kan als gewoon Nederlands woord staan voor zoiets als “veelbetekenend”³², “belangrijk”, “groot”, “rijk aan implicaties”, “samenhangend”, “betekenisvol”, en vergelijkbare woorden. Wanneer iemand in gewone spreek- of schrijftaal zegt: “Dit is significant”, dan kondigt hij of zij aan dat er iets heeft plaatsgevonden dat de moeite van onze aandacht waard is. Ik noem deze invulling van het begrip “belangrijk-significantie”. Zo was de aanval op de Twin Towers zeker belangrijk-significant, want deze had een zeer grote impact op de geopolitieke dynamiek van de decennia erna.

Maar er is een tweede invulling van het woord “significant”, namelijk zoals gebruikt in de wetenschap en in het bijzonder in de statistiek. In de Dikke van Dale wordt deze tweede betekenis ook expliciet genoemd, en omschreven als “verantwoorde conclusies toelatend”. Binnen de statistiek bestaat er een methode waarmee, zoals algemeen gedacht, bepaald zou kunnen worden welke wetenschappelijk verantwoorde conclusies uit data getrokken kunnen worden: als de data verantwoorde conclusies toelaten, dan heten deze conclusies “significant”. Ik noem dit, naar analogie van belangrijk-significant, dan ook “statistisch-significant”.

Deze twee verschillende contexten van het woord “significant” lijken wellicht op het eerste gezicht in elkaars verlengde te liggen. Immers, als iets kennelijk statistisch-significant wordt bevonden, en dus verantwoorde conclusies impliceert, waarom zou dat iets dan ook niet automatisch belangrijk-significant zijn? Andersom, als iets belangrijk-significant is, waarom zou de statistiek dat eveneens niet laten zien?

Ik zal hieronder proberen uit te leggen dat statistisch-significant, zoals deze doorgaans wordt gebruikt, op geen enkele manier iets te maken heeft met “verantwoorde conclusies toelaten” zoals Van Dale definieert, en al helemaal niet met belangrijk-significant. Dit wetende zijn de problemen duidelijk: een statisticus concludeert dat een bepaald effect significant is. Hij bedoelt dat als statistisch-significant, maar de leek die het leest kan niet anders dan denken dat dit over belangrijk-significant gaat en het misverstand is geboren. De meeste mensen zullen het woord “significant” lezen als belangrijk-significant, want dat is hoe het woord gebruikt wordt in normaal taalgebruik. Echter, zoals we zullen zien

³² Dit is de beschrijving in de Dikke van Dale.

bedoelt men vaak wel degelijk statistisch-significant. Ik zal nu eerst uitleggen wat statistisch-significant is, en in iets meer detail de verwarring bespreken.

Stel je bent als onderzoeker geïnteresseerd in de vraag of een bepaalde stikstofdepositie veroorzaakt wordt door een bepaalde bron, zeg een stal met vee. Hoe gaat een statisticus dan te werk? Meestal gaat dat als volgt. De onderzoeker begint met de zogenaamde “nulhypothese”: de bedoelde mogelijke oorzaak is *niét* de oorzaak van de depositie. Als nu deze aanname zeer moeilijk verenigbaar lijkt met de waargenomen data, dan trekt men veelal de conclusie dat de oorspronkelijke aanname – de nulhypothese – onjuist is én dat de bedoelde oorzaak dus juist *wél* de bron is.

Deze procedure gaat formeel ongeveer als volgt. De mate waarin de data opmerkelijk is *als de nulhypothese waar zou zijn*, wordt numeriek uitgedrukt met een zogenaamde ‘p-waarde’. Een kleine p-waarde betekent dat de uitkomst opmerkelijk is *als de nulhypothese waar zou zijn*. Anders gezegd: Een kleine p-waarde betekent dat de feitelijke uitkomst van ons experiment ‘onverwacht’ is als de nulhypothese waar is. Omdat de uitkomst toch is opgetreden, ‘geloven’ we daarom niet meer in die nulhypothese.³³

Echter, deze logica van de p-waarde faalt. Dat is niet moeilijk te bevatten, omdat een kleine p-waarde, of een p-waarde van welke grootte dan ook, niet logisch impliceert dat de nulhypothese onwaar is of zelfs onwaarschijnlijk. Immers, de p-waarde zegt iets *onder de aanname* dat de nulhypothese waar is.

Deze kritiek is uiteraard algemeen bekend in de statistiek,³⁴ maar om de een of andere reden blijft de legende van het belang van statistisch-significantie voortleven.

De nadelige gevolgen van het hanteren van de p-waarde en statistisch-significantie zijn nauwelijks te overschatten. In een belangrijke publicatie rekende epidemioloog John Ioannidis voor dat het niet anders kan dan dat de *meeste* wetenschappelijke hypotheses die statistisch-significant zijn bevonden, helemaal niet waar zijn.³⁵

Het statistiekwoord heeft geen enkele relatie met het Nederlandse succeswoord, hooguit bij toeval kunnen beide samenvallen. Omdat er geen verband is tussen statistische theorie en de Nederlandse taal, ontstaat, zoals nu duidelijk moet zijn, de drogreden van

³³ De technische details doen er hier niet toe, en laten we achterwege.

³⁴ Zie bijvoorbeeld Richard Royal, *Statistical Evidence*, CRC Press 1996; Ronald Meester en Klaas Slooten, *Kan dat geen toeval zijn? – een kritische blik op statistisch bewijs*, AUP 2021; William M. Briggs, *Uncertainty: the soul of probability, modeling & statistics*, Springer 2016.

³⁵ J. P. A. Ioannidis, *Why most published research findings are false*, Plos Medicine 2005.

dubbelzinnigheid. Dit gebeurt wanneer het statistiek-woord wordt verward met het succeswoord, wanneer “statistisch-significant” wordt verward met “belangrijk-significant”.

De situatie wordt verergerd door wat we de toevals- of correlatiedrogreden zouden kunnen noemen. Alle wetenschappers weten, en het is inderdaad een elementair principe van de logica, dat correlatie niet logischerwijs oorzaak en gevolg impliceert. Toch is de verleiding nauwelijks te weerstaan om wanneer een kleine p-waarde is gevonden, de correlatie waarop de kleine p-waarde is berekend, een oorzaak te noemen. De correlatiedrogreden wordt verergerd wanneer deze wordt gekoppeld aan de drogreden van dubbelzinnigheid. Want niet alleen wordt een correlatie ten onrechte beschouwd als een oorzaak, het wordt beschouwd als een belangrijk-significante oorzaak.

Er zijn echter nog meer redenen om niet naar statistiek-significantie te kijken. Naast vragen over de waarheid en onzekerheid in elke stelling of voorgestelde oplossing, moeten er namelijk ook altijd kosten en verwachte baten van eventueel handelen worden overwogen. Daarbij moeten we minimaal met de volgende dingen rekening houden:

1. Wat is de schade en hoeveel onzekerheid hebben we daar eigenlijk over? Deze onzekerheid is nodig om te kunnen beoordelen of acties eigenlijk wel gewenst en doelgeschikt zijn.
2. Wat is de mogelijke schade of gevolg van de oplossing? De oplossing kan variëren van “niets doen” tot een hele reeks aan activiteiten. De exacte voorgestelde oplossing of oplossingen moeten duidelijk worden gespecificeerd. Vage oplossingen zijn wetenschappelijk nutteloos, hoe politiek vruchtbaar ze ook mogen zijn (in de politiek loont het vaak om vaag en onduidelijk te zijn). Bijvoorbeeld, laten we veronderstellen dat een oplossing is om een grote boerderij, in de buurt van een locatie waar vermoed wordt dat groei van ‘onkruid’ is toegenomen vanwege hoge stikstofemissies, te sluiten. Deze oplossing kan al dan niet zeker zijn in haar beweerde effecten. De onzekerheid van dit beoogde effect moet worden vermeld. In de meeste gevallen is het zelden zo dat de oplossing zekere effecten sorteert.
3. Wat zijn de kosten? Het sluiten van een boerderij zal grote kosten met zich meebrengen. Deze moeten nauwkeurig worden gespecificeerd, niet alleen in de directe kosten, zoals het verlies van levensonderhoud voor velen en het

daaropvolgende verdwijnen van voedselproducten die de boerderij produceerde, maar ook in de indirecte kosten, zoals de kansen op verhoogde kosten van voedsel, de kosten van het geven van meer macht aan de staat, en dergelijke. Al deze kosten zullen onzekerheden met zich meebrengen en die moeten allemaal worden vermeld.

4. Wat zijn de voordelen? Het voordeel kan bijvoorbeeld zijn dat de kans op ongewenste groei van 'onkruid' wordt verminderd. Er zal echter, net als bij de kosten, slechts een kans zijn dat de voordelen worden gerealiseerd. De onzekerheden moeten daarbij ook worden vermeld.

Dit alles betekent dat we sowieso met een complexe wetenschappelijke situatie te maken hebben, die niet zinvol samengevat kan worden in een p-waarde of in welk ander getal dan ook. Alle genoemde vragen dienen bij een beoordeling meegenomen te worden.

KDW's en significantie

Als het inderdaad zo is dat een p-waarde (de waarde die bepaalt of een bevinding statistisch-significant is) niet veel zegt over de vraag hoe belangrijk het effect is, dan ligt het voor de hand om in combinatie met KDW's de belangrijk-significantie interpretatie te kiezen. Helaas is het echter zo dat dit in feite onmogelijk is, omdat de KDW werd geboren in een puur statistische omgeving. Hier is de oorspronkelijke definitie:³⁶

"A quantitative estimate of an exposure to one or more pollutants below which significant harmful effects on specified sensitive elements of the environment do not occur according to present knowledge".

Dit is vanwege de gevraagde kwantitatieve natuur niet als belangrijk-significant te begrijpen. Je kunt zelfs stellen dat wanneer een wetenschapper het woord "significant" gebruikt, het eigenlijk altijd statistisch-significant moet zijn, want dat is de enige manier om *als wetenschapper* het woord te gebruiken. Immers, wetenschap kan ons niet vertellen wat belangrijk-significant is en wat niet, want daar gaat de wetenschap niet over. Dit wordt

³⁶ Nilsson, S.I. en P. Grenfeldt (eds.), 1988. Critical Loads for Sulphur and Nitrogen. NORD 1988: 97. Kopenhagen: Nordic counsel of ministers.

bevestigd door bijvoorbeeld het volgende citaat uit een rapport van Wageningen University:³⁷

“Anders dan in de studie naar de door struikheide gedomineerde heiden, waren er in deze studie wél voldoende waarnemingen in de laagste depositieklasse om statistisch significante uitspraken te doen bij lage depositieniveaus. [...] De totale soortenrijkdom nam significant af met toenemende stikstofdepositie. De sterkste afname lijkt op te treden tussen de klassen 15-20 kg N/ha/j. en 20-25 kg N/ha/j. Deze conclusie hangt wel af van de precieze lijn die wordt getrokken door de punten: een lineaire regressielijn is significant, maar een exponentiële afname is significanter.”

Het is evident dat significantie hier als statistisch concept gebruikt wordt.

De Europese Commissie stelt echter, en mijns inziens terecht, dat significantie primair moet worden verstaan als “belangrijk-significantie”. Dit blijkt zonneklaar uit twee e-mails die Daniela Stoecheva, Press Offices for Environment, Maritime Affairs and Fisheries stuurde toen haar werd gevraagd naar de manier waarop we volgens haar het begrip significantie moeten duiden:

“In the terms you are using below, I would tend to think the interpretation would be as ‘important significant’.”³⁸

“The guidance documents on the interpretation of Article 6 and methodological guidance on Article 6(3) (4) provide some clarification how “significance” should be understood. The analysis of significance of effect has to be done on a case-by-case basis and in view of the conservation objectives set for the site. What may be significant for one site may not be significant for another. Significance of the effects will vary depending on factors such as the magnitude of impact, the type, extent, duration, intensity, timing, probability, cumulative effects and the vulnerability of the habitats and species concerned. For example, a loss of a hundred square meters of

³⁷ G.W.W. Wamelink et al., *Relaties tussen de hoeveelheid stikstofdepositie en de kwaliteit van habitattypen*, Rapport 3089, 2021, blz. 21 respectievelijk blz. 25.

³⁸ Persoonlijke communicatie, 5 oktober 2023.

habitat may be significant for a small rare orchid site, while a similar loss in a large steppe site may be insignificant if it does not affect the site's conservation objectives.”³⁹

Dit antwoord laat aan duidelijkheid niet veel te wensen over. We zoeken niet naar statistisch-significante effecten, maar naar effecten die wij mensen belangrijk of veelbetekenend achten. Dat zijn geen zaken waar een wetenschappelijk antwoord op te geven valt. Immers, het zijn mensen die iets ‘belangrijk’ vinden; daar gaat wetenschap helemaal niet over en de statistiek al helemaal niet.

Terugkerend naar de KDW's moet nu duidelijk zijn dat statistisch-significantie *in zichzelf* niet gebruikt kan worden bij het beantwoorden van de vraag wat nu precies kritiek is en wat niet.⁴⁰ Het begrip ‘kritiek’ kan niet aan de hand van een statistisch concept worden gedefinieerd, maar moet gebaseerd zijn op het ‘veld’: biologie, chemie, ecologie, ethiek en ongetwijfeld nog andere zaken. Zodra een goed idee van ‘kritiek’ algemeen wordt geaccepteerd, is het daarna natuurlijk wel mogelijk om met behulp van statistiek iets te zeggen over geobserveerde of verwachte waarnemingen.

Met andere woorden: statistiek *in zichzelf* kan niet bepalen wat wij belangrijk vinden en wat niet; die keuzes zijn niet aan de wetenschap maar aan de maatschappij. Statistiek mag ook niet voor dat doel misbruikt worden.

Hoe worden KDW's bepaald?

Het is natuurlijk zo dat ‘te veel’ stikstof schadelijk kan zijn, om de simpele reden dat dit voor alles geldt. Te veel stikstof kan een negatief effect hebben op de biodiversiteit. Hoeveel nodig is om schadelijk te zijn, en zelfs hoe je “schadelijk” definieert, is niet makkelijk vast te stellen en de literatuur is daar absoluut niet eensluidend over. Je bent ofwel aangewezen op kleinschalige experimenten waarvan de relevantie op grotere ruimte- en tijdschaal onduidelijk is, ofwel je probeert een model te maken waarbij je kijkt naar de kans dat er iets schadelijks gebeurt. Zo'n kans betekent dan echter alleen maar iets binnen het model zelf en de betekenis voor het veld is dan vaak helemaal niet duidelijk. In mijn studie naar het

³⁹ Persoonlijke communicatie, 17 oktober 2023.

⁴⁰ Een dergelijke conclusie werd ook getrokken in William M. Briggs en Jaap C. Hanekamp, *Nitrogen critical loads: critical reflections on past experiments, ecological endpoints, and uncertainties*, Dose-response DOI: 10.01177/15593258221075513 (2022).

gebruik van KDW's kwam ik veel verschillende benaderingen tegen.⁴¹ Er is voor zover ik heb kunnen nagaan geen gezamenlijke en algemeen geaccepteerde methodologie om de KDW's vast te stellen. Ik noem een aantal zaken die mij op zijn gevallen.

In sommige benaderingen werd een KDW met een model vastgesteld.⁴² Dat vind ik in zichzelf al een probleem, want een model geeft alleen maar terug wat je er zelf in hebt gestopt en is daardoor noodzakelijk enigszins circulair, maar bovendien wordt 'schadelijk' hier doorgaans in termen van p-waardes gedefinieerd. Die methodologie heb ik hierboven al afgewezen.

Ook is er vaak sprake van 'opschaling' van empirische bevindingen.⁴³ Experimenten op een zeer klein stukje grond en gedurende korte tijd (denk aan 1 m² en een periode van enkele weken) worden zonder verder commentaar opgeschaald naar minimaal hectares en periodes van jaren. Dat kan statistisch niet goed gaan, want meetfouten en verdere onzekerheden worden bij de opschaling natuurlijk enorm uitvergroot.

Verder lijkt het erop dat de onzekerheid van allerlei externe factoren vaak niet gehonoreerd wordt. Zo kwam ik een artikel tegen waarin de zogenaamde 'achtergronddepositie' op een vast getal gezet werd, terwijl die grootte enorm zal variëren in de loop van de tijd.⁴⁴

Tenslotte lijkt het erop dat onderzoekers simpelweg conclusies trekken waarvan ik onmogelijk in kan zien dat ze uit de experimenten zouden volgen. In een bepaald experiment⁴⁵ werden verschillende potten van 24 × 24 × 32 cm gebruikt waarin extra stikstof werd geïnjecteerd, van 0 tot 4 g/m²/jaar, met een "verwaarloosbaar" achtergrondniveau van stikstofdepositie. Er werden verschillen gevonden bij de hoogste stikstofniveaus in de plantbiomassa van twee soorten Sphagnum. Maar gezien de volledig kunstmatige en kleinschalige aard van het experiment, lijkt het moeilijk om, zoals de auteurs doen, te concluderen dat "hoge niveaus van stikstofdepositie inderdaad verantwoordelijk lijken te

⁴¹ Zie bijvoorbeeld R. Bobbink en J.P. Hettelingh, *Review and revision of empirical critical loads and dose-response relationships*, Proceedings of an expert workshop, 23-25 June, 2010.

⁴² Bijvoorbeeld in Banin L, Bealey B, Smith R, Sutton M, Campbell C, Dise N, *Quantifying Uncertainty in Critical Loads*. Technical report, CEH Report to SEPA. Lancaster, UK: Centre for Ecology & Hydrology; 2014.

⁴³ Bijvoorbeeld in Aerts R, Wallen B, Malmer N, *Growth-limiting nutrients in sphagnum-dominated bogs subject to low and high atmospheric nitrogen supply*. J Ecol. 1992;80: 11-39.

⁴⁴ Bijvoorbeeld in Bragazza L, Tahvanainen T, Kutnar L, et al. *Nutritional constraints in ombrotrophic sphagnum plants under increasing atmospheric nitrogen deposition in Europe*. New Phytol. 2004;163: 609-616.

⁴⁵ Bijvoorbeeld in Tomassen HBM, Smolders AJP, Lamers LPM, Roelofs JGM, *Stimulated growth of betula pubescens and molinia caerulea on ombrotrophic bogs: role of high levels of atmospheric nitrogen deposition*. J Ecol. 2003;91:357-370.

zijn voor de waargenomen snelle vegetatieveranderingen in (echte) ombrotrofe veengebieden.”

Al deze problemen betekenen onder meer dat de KDW's eenvoudigweg niet als unieke welgedefinieerde getallen kunnen worden gerepresenteerd, en doen alsof dat wel zo is acht ik onwetenschappelijk. Dat omstandigheden veranderen als gevolg van de depositie van stikstof is in *zichzelf* niet van belang. Het belang daarvan moet op een andere manier blijken.

Ik merk tenslotte op dat het bepalen van KDW's zoals boven beschreven niet via de fauna verloopt, maar via de flora. Dat heeft natuurlijk een begrijpelijke pragmatische reden.

Historisch perspectief

Hoe de KDW's toch zo belangrijk hebben kunnen worden is een verhaal apart dat ik hier niet helemaal zal vertellen, maar waarschijnlijk heeft de VN conferentie in Sofia in 1988 hierbij een grote rol gespeeld. In 2008 (!) concludeerde de commissie Trojan het volgende:⁴⁶

“De kritische depositowaarde is in Nederland als een te absolute waarde gehanteerd. Deze waarde moet in de toekomst niet meer gezien worden als een doelstelling die behaald moet worden, zoals dat ook in de omliggende landen niet gebeurt.”

Mijn enige bezwaar tegen deze uitspraak is dat er gesproken wordt over “de” kritische depositiewaarde. Ik heb hiervoor uitgelegd dat “de” waarde niet bestaat. Een jaar later, in 2009 doet de commissie-Huys het nog eens dunnetjes over:⁴⁷

“Naar onze mening heeft de kritische depositiewaarde in het Nederlandse beleid een te grote aandacht gekregen en is die waarde te stringent geformuleerd en toegepast.”

Het mocht allemaal niet baten. In 2020 neemt de Tweede Kamer de ‘Stikstofwet’ aan. In Artikel 1.12a lid 2 van die wet lezen we:

⁴⁶ Kamerstuk 30 654, nr. 51, <https://zoek.officielebekendmakingen.nl/kst-30654-51-b2.pdf>, blz. 26.

⁴⁷ Adviesgroep Huys, *Meer dynamiek bij de uitvoering van nationale en Europese natuurwetgeving: perspectief van een programmatische aanpak*, blz. 9.

“De in het eerste lid bedoeld omgevingswaarden⁴⁸ zijn resultaatsverplichtingen.”

Vanuit een wetenschappelijk perspectief is dit artikel niet te verdedigen, zoals ik heb betoogd. Maar eigenlijk is de situatie nog ondoorzichtiger en gecompliceerder. Want hoe kom je er achter of je aan de “resultaatsverplichting” hebt voldaan? Hoe kun je aan de wet voldoen? Het antwoord is dat je met Aerius/OPS moet berekenen of het resultaat gehaald is.

Enigszins tot mijn verbazing heb ik over het toepassen van Aerius alleen in de Omgevingsregeling artikelen kunnen vinden.⁴⁹ Zo luidt Artikel 4.15 uit die Omgevingsregeling als volgt:

“Artikel 4.15 (methode berekenen stikstofdepositie Natura 2000-activiteit)

1. Dit artikel is van toepassing op het bepalen van de stikstofdepositie bij het vaststellen of een activiteit door het veroorzaken van stikstofdepositie als een Natura 2000-activiteit moet worden aangemerkt.
2. Op het berekenen van de stikstofdepositie is AERIUS Calculator van toepassing.”

In hoeverre “van toepassing” zijn tot een verplichting leidt om de uitkomsten van een model letterlijk te nemen of zelfs maar te gebruiken laat ik aan juristen over. Mijn conclusie is dat mijn wetenschappelijke en statistische bezwaren tegen het gebruik van Aerius/OPS tot een eenvoudige aanpassing van de Omgevingsregeling zouden moeten leiden. De stikstofwet is moeilijker aan te passen, dus het probleem van de KDW's is hardnekkiger.

Daarmee wordt het dus belangrijk om te onderzoeken hoe het zit met de betrouwbaarheid van uitkomsten van het Aerius/OPS rekenmodel. Kan dat model wel leveren wat van haar gevraagd wordt in dit dossier? Dat onderzoek ik in het volgende hoofdstuk.

⁴⁸ Hier worden de KDW's mee bedoeld.

⁴⁹ [wetten.nl - Regeling - Omgevingsregeling - BWBR0045528](https://wetten.nl/Regeling-Omgevingsregeling-BWBR0045528).

4 - Onzekerheid en voorspellend vermogen van Aerius/OPS

Aerius/OPS is een deterministisch model. Het voorspelt atmosfeerconcentraties aan de hand van emissiegetallen, in combinatie met tal van parameters. Als je twee keer dezelfde input geeft komt er twee keer hetzelfde uit. Dat er toch zoiets als ‘onzekerheid’ rond Aerius/OPS bestaat heeft een aantal redenen. Allereerst maakt Aerius gebruik van de output van andere modellen, zoals bijvoorbeeld de met NEMA geschatte geaggregeerde emissies. Verder is het zo dat stikstofdepositie niet gemeten kan worden, maar afgeleid wordt uit andere grootheden zoals atmosferische concentraties. Die andere grootheden en de afleiding zélf leveren ook weer onzekerheden op.

Een wiskundig model kan de werkelijkheid nooit volledig of vaak zelfs maar bij benadering beschrijven, en het feit dat we daarvoor een model gebruiken brengt inherent onzekerheden met zich mee over de adequaatheid van het model. Deze verschillende vormen van onzekerheid, die lang niet allemaal gekwantificeerd kunnen worden, bespreek ik hieronder in iets meer detail.

Meetonzekerheid van de depositie zelf⁵⁰

Wellicht denken veel mensen dat je depositie van stikstof kunt meten zoals je bijvoorbeeld de temperatuur meet, maar in werkelijkheid is dat niet zo. Natuurlijk is er zelfs bij het meten van temperatuur sprake van onzekerheid bij het aflezen, maar deze vallen in het niet bij wat er moet gebeuren om stikstofdepositie te ‘meten’.⁵¹ Ik citeer hier enkele opmerkingen hierover van de website van het RIVM.⁵²

“De bodem en planten kunnen ammoniak en stikstofoxiden direct uit de lucht opnemen. Luchtwervelingen zorgen ervoor dat stoffen naar de grond bewegen (turbulentie). Dit noemen we droge depositie. Met de COTAG-mast meten we droge depositie van ammoniak.

⁵⁰ In dit rapport heb ik het over droge depositie, die het leeuwendeel van de totale depositie behelst. Natte depositie is veel beter te meten maar van ondergeschikt belang.

⁵¹ Ik schrijf ‘meten’ tussen aanhalingstekens omdat, zoals zo duidelijk wordt, er in feite geen depositie gemeten wordt. Het wordt afgeleid uit andere zaken.

⁵² [Metingen van droge depositie van ammoniak | RIVM](#)

Bovenop de COTAG-mast staat een windmeter. Deze meet de turbulentie in de lucht. Twee kasten die aan de mast hangen meten op twee hoogten hoeveel ammoniak er in de lucht zit. Met deze meetgegevens berekent het RIVM hoeveel ammoniak er direct uit de lucht in de natuur terechtkomt.

Het RIVM gebruikt de meteorologische gradiëntmethode om de droge depositie van NH_3 te bepalen met een COTAG-mast. In deze methode is de hoeveelheid droge depositie een combinatie van het concentratieverschil van NH_3 op twee hoogten en een mate van turbulentie.

Als er depositie is, is de concentratie direct boven het oppervlak lager dan op grotere hoogte. Dat komt doordat de vegetatie en bodem ammoniak opnemen. Er kan in plaats van depositie ook emissie optreden. In dat geval is de concentratie direct boven de bodem en vegetatie hoger dan op grotere hoogte.

De mate van turbulentie bepaalt de snelheid waarmee ammoniak van of naar de bodem en vegetatie gaat. Dit hangt af van de stabiliteit van de atmosfeer. De stabiliteit van de atmosfeer verandert gedurende de dag maar ook door het jaar heen.

Overdag en in de zomer wanneer de zon de aarde opwarmt is er veel turbulentie. Opstijgende bellen warme lucht zorgen voor een gemakkelijke uitwisseling met het aardoppervlak. De atmosfeer is onstabiel. In de nacht of in de winter is er bijna geen turbulentie. Dit komt omdat de lucht dicht aan het aardoppervlak kouder is dan op hogere hoogte. Er is dus dan ook weinig uitwisseling en menging, de atmosfeer is stabiel.

Als de wind langs bomen, gebouwen of andere obstakels waait, kunnen er ook turbulente wervelingen ontstaan. Hoe hoger de windsnelheid hoe meer turbulentie. Als dit de belangrijkste oorzaak van de turbulentie is, dan is de atmosfeer neutraal. De COTAG meet onder twee turbulente situaties: onstabiel en neutraal. Onder situaties met zeer weinig turbulentie (stabiel) en zeer veel turbulentie (zeer onstabiel) staat de COTAG in de 'uitstand'. Tijdens stabiele situaties is er zeer weinig menging van de lucht. Daarom is het moeilijk de menging van de lucht goed te meten. De droge depositie van ammoniak is dan ook klein. Maar ook onder zeer onstabiele situaties staat het COTAG-systeem in de 'uitstand'. In dat geval is er wel veel menging, maar is het NH_3 -verschil zo klein dat het niet goed meetbaar is.

Doordat de COTAG een deel van de tijd in de ‘uitstand’ staat, meet de COTAG niet alle droge depositie. In 2023 hebben we de methode om tot de totale droge depositie van ammoniak te komen verbeterd. Sindsdien gebruiken we voor de tijden dat de COTAG niet meet, de waarde van de meest passende turbulentieklasse. Staat de COTAG uit vanwege stabiele omstandigheden, dan gebruiken we de meetwaarde van neutraal. Bij zeer onstabiele omstandigheden gebruiken we de meetwaarde van onstabiel.”

Het zal nu voor iedereen duidelijk zijn nu dat de depositie niet direct gemeten kan worden, maar een afgeleide is van tal van andere metingen (elk met hun eigen onzekerheid) die bovendien ook nog eens modelmatig (via de genoemde gradiëntmethode) tot een berekende depositie leiden. Dit kan niet anders betekenen dan dat de ‘metingen’ zelf met grote onzekerheden gepaard gaan. Hoe groot deze zijn weet ik niet, en dat zal ook moeilijk vast te stellen zijn.⁵³ Ik merk daarbij bovendien op dat deze ‘metingen’ op slechts vijf Natura-2000 gebieden plaatsvinden.

Voor een toegankelijk overzicht van de hele Aeries/OPS omgeving verwijs ik verder naar Figuur 1 van de eerdergenoemde rapportage van de commissie Hordijk.

Ik merk voor de volledigheid op dat deze ‘metingen’ algemene en globale ‘metingen’ van jaargemiddeldes zijn – het zijn géén ‘metingen’ over de depositie als gevolg van een bepaalde, specifieke, activiteit. ‘Metingen’ daarvan bestaan simpelweg niet, alleen berekeningen via Aeries/OPS.

NEMA-onzekerheid

Het National Emission Model for Agriculture (NEMA) berekent landbouwemissie (naast andere emissiebronnen) naar de lucht, en is een input voor Aeries/OPS. Zo lezen we:⁵⁴

“Het NEMA berekent de uitstoot van stoffen voor bijvoorbeeld stallen, mestopslag, en het gebruik van mest. Het NEMA wordt ook gebruikt om emissies zoals methaan

⁵³ Maar zie Jan Willem Erisman, *Acid deposition to nature areas in The Netherlands: part I: methods and results*, water, Air and Soil pollution 71 (1993) voor schatting op basis van een andere techniek. De onzekerheden over de metingen bedragen volgens hem tientallen procenten.

⁵⁴ [Methodology for estimating emissions from agriculture in the Netherlands. Calculations for CH₄, NH₃, N₂O, NO_x, NMVOC, PM₁₀, PM_{2.5} and CO₂ using the National Emission Model for Agriculture \(NEMA\) – Update 2021 | RIVM](#)

uit verschillende dieren en mest te berekenen. Dit model wordt elk jaar aangepast aan de nieuwste wetenschappelijke inzichten.”

Natuurlijk gaan deze berekeningen met de nodige onzekerheden gepaard:

“Met behulp van onzekerheidsanalyses wordt een bandbreedte aangegeven waarbinnen de berekende emissies met een 95%-betrouwbaarheid (waarschijnlijkheid) zullen liggen. De onzekerheidsanalyses hebben betrekking op activiteitendata en emissiefactoren, maar niet op gehanteerde proces- en modelformuleringen.”⁵⁵

Wat betekent dit precies? Welnu, de emissies worden ‘berekend’ als

$$\text{emissie} = \text{activiteit} \times \text{emissiefactor}$$

Voor activiteiten en emissiefactoren is per bron een onzekerheid ingeschat, en er wordt dus aangenomen dat het proces volgens deze formule verloopt. Dit is een grote vereenvoudiging van de werkelijkheid. Er zijn allerlei factoren die effect hebben op het proces emissie (wind, temperatuur enzovoort) waardoor deze formule van beschrijving van het proces emissie de werkelijkheid onmogelijk goed kan beschrijven. Dus er is niet alleen sprake van onzekerheid binnen de getoonde formule zélf, maar ook over de reikwijdte van deze formule. De formule ‘past’ niet goed op de werkelijkheid. De onzekerheid in de eenvoudige formulering van het proces emissie is echter niet meegenomen in de onzekerheidsanalyse, zoals de auteurs zelf dus ook erkennen.

In deze onzekerheidsanalyses is verder de onzekerheid in naleving en handhaving niet of nauwelijks meegenomen. Ook is nog geen rekening gehouden met de gewijzigde onzekerheid in de emissie uit stallen met emissiearme vloeren en mestkelders.

Voor de methode van onzekerheidsberekening en de schattingen van de onzekerheden in activiteitendata en emissiefactoren wordt verwezen naar Van der Zee et al.

⁵⁵ C. Van Bruggen et al., *Emissies naar lucht uit de landbouw berekend met NEMA voor 1990-2022*, WOT-Technical report 264, blz. 62, 2024.

(2024).⁵⁶ We vinden daar een onzekerheidsanalyse voor 2020 met de methode van ‘propagation of error’.⁵⁷ Wat betekent dat precies? Deze methode berekent de onzekerheden van een aantal gezamenlijke bronnen met de individuele onzekerheden als input. Als je een aantal verschillende schattingen bij elkaar optelt of juist met elkaar vermenigvuldigt, en je doet aannames over de (on)afhankelijkheid van de verschillende bronnen, dan kun je met elementaire kansrekening uitrekenen hoe groot de totale onzekerheid wordt. Maar ik benadruk dat deze berekening volledig binnen het model plaatsvindt, en dus theoretisch en niet empirisch is.

De NEMA-schattingen zijn dus zeer onzeker, en deze onzekerheid zal een weerslag hebben op de onzekerheid over de realiteitswaarde van Aerius/OPS-berekeningen.

Voorspelt Aerius goed?

In 2023 publiceerde het RIVM een rapport over de onzekerheid in de berekeningen van Aerius/OPS.⁵⁸ Het is een nogal technisch rapport. Het rapport gaat niet over de onzekerheid van specifieke projecten, maar over onzekerheid in *gemiddelden* op een bepaalde locatie over langere tijd – het gaat over jaren:

“De eerste stap is het bepalen van de onzekerheid voor de zes stikstofcomponenten. We doen dit door jaarcijfers van de OPS berekeningen te vergelijken met de beschikbare metingen.”⁵⁹

Op specifieke locaties wordt de totale depositie een jaar lang ‘gemeten’. Vervolgens wordt uitgerekend hoe groot het modelgemiddelde over een jaar is. Dat laatste gebeurt door de depositie onder verschillende omstandigheden te berekenen door het model,⁶⁰ en dan het gewogen gemiddelde hiervan te berekenen op basis van hoe vaak elk van deze omstandigheden gemiddeld in een jaar voorkomen. Op basis hiervan wordt op een

⁵⁶ Van der Zee, T.C., A. Bleeker, C. van Bruggen, W. Bussink, H.J.C. van Dooren, C.M. Groenestein, J.F.M. Huijsmans, H. Kros, Lagerwerf, L.A., K. Oltmer, M. Ros, M.W. van Schijndel, L. Schulte-Uebbing & G.L. Velthof (2024). *Methodology for the calculation of emissions from agriculture. Calculations for methane, ammonia, nitrous oxide, nitrogen oxides, non-methane volatile organic compounds, fine particles and carbon dioxide emissions using the National Emission Model for Agriculture (NEMA)*. RIVM report 2024-0015. RIVM, Bilthoven.

⁵⁷ Zie bijvoorbeeld [Microsoft Word - V1_Ch3_Uncertainty_final_V2.doc](#)

⁵⁸ [Uncertainty in the determined nitrogen deposition in the Netherlands. Status report 2023 | RIVM](#)

⁵⁹ Ibid, blz. 13.

⁶⁰ De omstandigheden kunnen tot op zekere hoogte in het Aerius/OPS model ‘ingevoerd’ worden. Natuurlijk is dat ook in zichzelf al een benadering.

specifieke manier een uitspraak gedaan over de onzekerheid van modelberekeningen. Op blz. 17 van het rapport lezen we dit:

“Dit resulteert in een gemiddelde onzekerheid in de totale stikstofdepositie van ongeveer 60 - 70% (2σ) voor een willekeurige locatie in Nederlands natuurgebied. Dit betekent 900 - 1100 mol/ha/j. Deze 60% à 70% komt goed overeen met de eerdere schatting. Het voordeel ten opzichte van de eerdere onzekerheidsschattingen is dat de definitie van de onzekerheidsmaat nu veel scherper is. Merk op dat dit de onzekerheid is van de waarde zelf, dus in absolute zin. Er is 5% kans dat de gerapporteerde modelwaarde 60-70% lager of hoger is. Maar in veel gevallen worden modelwaarden relatief met elkaar vergeleken. Voorbeelden hiervan zijn vergelijking van verschillende jaren in trendanalyses, of het vergelijken van het wel of niet meenemen van bepaalde bronnen. In deze situaties vallen veel van de bovenstaande onzekerheden tegen elkaar weg. Hierdoor zal de onzekerheid kleiner worden.”

Het RIVM erkent dus dat het model er wat betreft het gemiddelde er 60% à 70% naast kan zitten. Ik vermoed overigens dat de totale onzekerheid nog een stuk groter is dan het RIVM hier rapporteert, omdat de eerdergenoemde onzekerheid van de ‘metingen’ niet wordt meegenomen.

Het RIVM-rapport rept dus helemaal niet over de onzekerheid van de Aerius berekeningen van specifieke lokale activiteiten, zoals dat voor de vergunningverlening wel nodig is, maar over geaggregeerde jaarcijfers. Dat maakt dat de gerapporteerde onzekerheid geen directe functie heeft in het kader van deze vergunningverlening. Echter, we leren er toch wel wat van. De onzekerheid van berekeningen voor specifieke activiteiten zal immers zeker groter zijn - het nemen van een gemiddelde werkt immers ‘absorberend’ in de zin dat extremen minder groot worden, en de verwachte schommelingen dus kleiner. Het punt dat ik wil maken, is dat door het nemen van gemiddeldes, een model er beter uit kan komen te zien dan op het niveau van individuele voorspellingen. Dat is een cosmetisch (wiskundig) fenomeen en niet feitelijk toegenomen modelprecisie.

In een eerder rapport van het RIVM⁶¹ toont de auteur een figuur (hun Figuur 8.3) waaruit blijkt dat het door Aerius/OPS berekende patroon van gemiddelden in zekere mate wel overeenkomen met de praktijk, in de zin dat een hogere ‘meting’ doorgaans ook een hogere berekende depositie betekent. Maar dat kan een eenvoudiger model ook, en bovendien staat dat nog mijlenver af van een zinvolle toepassing bij vergunningverlening.

Deze kritiek op Aerius is noch nieuw, noch origineel. In 2022 werd al dergelijk kritiek op Aerius geformuleerd, en werd zelfs gesteld dat het eenvoudige ‘mean model’ beter werkt dan Aerius zelf, op basis van empirische data.⁶² In dat ‘mean model’ neem je als voorspelling gewoon de gemiddelde waarde tot nu toe. De kritiek op Aerius/OPS in dit artikel gaat overigens veel verder dan dat, daarover straks meer.

Dit erkent het RIVM overigens allemaal volmondig. Ik heb tijdens mijn gesprek op het RIVM begrepen dat er ook gewerkt wordt aan een onzekerheidsanalyse op individueel activiteitsniveau. Die analyse kan echter geen empirische analyse zijn omdat de hoeveelheden waarom het gaat niet te meten zijn. Het is een model-analyse; in de wiskunde noemen we dat een sensitiviteitsanalyse, waarbij wordt gekeken hoe de uitkomsten variëren bij het variëren van de inputparameters. Hoewel dat in zichzelf interessant kan zijn (maar ook circulair omdat we die onzekerheden er bij de constructie van het model zelf ingestopt hebben) geeft die methode geen inzicht in de empirische onzekerheden. Die zijn simpelweg niet te bepalen. Dat reduceert uiteraard het belang van zo’n analyse.

Het RIVM rapporteert dus over de onzekerheid van gemiddelde uitkomsten en komt daarbij tot onzekerheden die groot zijn. Kennelijk is het niet zomaar waar dat een gecompliceerd model goed kan voorspellen. Dat is overigens heel gebruikelijk in de wetenschap. De meeste mensen denken dat hoe gedetailleerder een model wordt, hoe beter het ook wordt, maar dat is doorgaans helemaal niet het geval.

Feit is dus dat we geen idee hebben hoe goed Aerius werkt in individuele gevallen, om de simpele reden dat de hoeveelheden waar het daar om gaat niet te meten zijn. De lezer moet goed beseffen dat de onzekerheden die we daar mogen verwachten erg groot zijn, groter dan de onzekerheden die het RIVM rapporteert over de gemiddelde waardes. Dit acht ik voor een vergunningsinstrument niet acceptabel.

⁶¹ J. Jaarsveld, *Description and validation of ops-pro 41*, Technisch rapport 500045001/2004.

⁶² Zie William M. Briggs, Jaap C. Hanekamp en Geesje Rotgers, *Criticizing AERIUS/OPS Model Performance*, 2022.

Met een klein experiment met de Aerius Online Calculator⁶³ kun je overigens zelf uitzoeken hoeveel depositie Aerius voorspelt bij een agarische activiteit. Ik heb e.e.a. zelf geprobeerd, en kwam tot de conclusie dat bijvoorbeeld 300 koeien nooit meer dan zo'n 3 mol stikstof/ha/jaar opleverden, en dat is gezien de gestelde achtergronddepositie van enkele duizenden mol/ha/jaar niet alleen praktisch verwaarloosbaar, maar ook niet te meten. Het gebruik van Aerius/OPS is dus uitsluitend een modeloefening waarbij we geen idee hebben over de realiteitswaarde. De onzekerheid die het RIVM zelf rapporteert in de jaarcijfers laten al zien dat de onzekerheid op individueel niveau groter zullen zijn dan de 60% - 70% die het RIVM rapporteert. Hoe groot kunnen we echter niet weten. Voor een instrument dat bij vergunningsverlening voor individuele projecten, met unieke KDW's, gebruikt wordt, is een dergelijke onzekerheid onaanvaardbaar.

Samenvattend: de prestaties van het model op individueel niveau bij vergunningverlening worden omgeven met onzekerheden die niet kwantificeerbaar zijn, omdat gemodelleerde deposities onmeetbaar klein zijn. Dat de onzekerheden hier veel groter zullen zijn dan bij landelijke (globale) gemiddeldes is nu eenmaal een wiskundig gegeven. Deze globale modeluitkomsten worden zelf al omgeven met vele tientallen procenten onzekerheid, voortkomend uit verschillende onzekerheden. Controle van modelprecisie en -onzekerheid is niet mogelijk omdat er geen meetdata voor stikstofdepositie voorhanden zijn. Daar komt de onzekerheid van meetresultaten van atmosferische concentraties nog eens bovenop, die in dit specifieke dossier moeilijk kwantificeerbaar zijn. En de 'gemeten' depositie is dan een afgeleide van deze atmosferische concentraties.

In het eerste citaat uit het RIVM rapport wordt, tenslotte, nog gesproken over het "wegvallen" van onzekerheden tegen elkaar. Dat is alleen, en alleen dan, eventueel mogelijk wanneer de resultaten inderdaad gebruikt worden voor vergelijking. Immers, in die gevallen mag je hopen dat de vergeleken waarden op dezelfde manier zullen afwijken van de berekeningen als gevolg van correlaties van onzekerheden. Bij het berekenen van de depositie van individuele projecten is "wegvallen" echter simpelweg niet aan de orde, en dat zal ik hier dan ook niet verder bespreken.

⁶³ [Alles over AERIUS Calculator | Producten | AERIUS](#)

Validatie van Aeries

Na de bespreking van de grote onzekerheden in berekende deposities, rijst vanzelf de vraag of Aeries/OPS wel goed gevalideerd is. Valideren van een model is geen eenvoudige opgave, en wat je daarvoor precies moet doen staat niet in een algemeen geaccepteerde handleiding. Wel is het zo dat validatie altijd neerkomt op het vergelijken van voorspellingen met daadwerkelijke observaties, maar er zijn veel manieren om zo'n vergelijking te maken.⁶⁴ Het blijkt daarbij dat het simpele mean model dat ik eerder besprak, beter uit die validatietests komt dan Aeries/OPS zelf.

Ik heb geprobeerd om na te gaan welke validatiestudies er voor Aeries/OPS zijn uitgevoerd, maar de resultaten daarvan waren niet erg bevredigend. In het kader van dit rapport is het niet nodig om hier gedetailleerd op in te gaan, de studie van het RIVM zelf is hierover voor ons informatief genoeg.⁶⁵

De recente voorlichting aan de regering van de Raad van State

Dit is nu het goede moment om kort in te gaan op de voorlichting die de Raad van State publiceerde op 26 mei 2025. Ik citeer:⁶⁶

“Het eventueel niet kunnen meten van stikstofdepositie onder de grens van 1 mol/ha/jaar, betekent echter niet dat de omvang daarvan niet modelmatig kan worden berekend en daarmee redelijk kan worden ingeschat. Het betekent ook niet dat die depositie er niet is. Het AERIUS-model en de daarop gebaseerde Calculator geven de best mogelijke inschatting van de stikstofdepositie op een Natura 2000-gebied, ook voor individuele projecten.”

Laten we deze opvatting bespreken aan de hand van de kennis die we nu hebben verzameld. Ik zal de drie zinnen in het citaat één voor één bespreken.

⁶⁴ In het eerder geciteerde rapport van Briggs, Hanekamp en Rotgers bijvoorbeeld, wordt een lijst met mogelijke validatievergelijkingen gepresenteerd.

⁶⁵ De geïnteresseerde lezer kan William M. Briggs, Jaap C. Hanekamp en Geesje Rotgers, *Criticizing AERIUS/OPS Model Performance*, 2022 raadplegen voor een overzicht.

⁶⁶ [Voorlichting over mogelijkheden introductie rekenkundige ondergrens stikstofdepositie. - Raad van State](#)

“Het eventueel niet kunnen meten van stikstofdepositie onder de grens van 1 mol/ha/jaar, betekent echter niet dat de omvang daarvan niet modelmatig kan worden berekend en daarmee redelijk kan worden ingeschat.”

Het woordje “eventueel” is overbodig: we kunnen onder die grens niet ‘meten’. Het is vervolgens natuurlijk juist dat je ondanks die meetproblemen kunt uitrekenen wat je maar wilt. Maar de Raad van State maakt niet aannemelijk dat deze berekening “redelijk” is. In dit rapport heb ik uitgelegd dat het tegenovergestelde het geval is.

“Het betekent ook niet dat die depositie er niet is.”

Dit is, uiteraard, volkomen juist maar ik ken niemand die ooit het tegengestelde heeft beweerd. De Raad van State wekt hier de indruk een stroman te bestrijden.

“Het AERIUS-model en de daarop gebaseerde Calculator geven de best mogelijke inschatting van de stikstofdepositie op een Natura 2000-gebied, ook voor individuele projecten.”

Ik heb uitgelegd dat een beroep op “de best mogelijke inschatting” niet acceptabel is, als die inschatting ons gewoon geen informatie geeft. Bij individuele projecten zegt de Aeries-berekening niets; daarvoor zijn de onzekerheden van het geheel veel te groot. De Raad van State ziet onvoldoende in dat we hier te maken hebben met een modelexercitie die noch met enige functionele precisie kan worden uitgevoerd noch empirisch verifieerbaar is.

De problemen staan niet op zichzelf

In mijn analyse heb ik de stikstofpraktijk en de empirische en theoretische (statistische) kaders die daarvoor de afgelopen jaren in stelling zijn gebracht onderzocht. Ik heb geconstateerd dat beide ernstig tekortschieten. Daarmee lijkt de kous af. Echter, mijn en andere kritieken op de huidige gang van zaken, zouden aanleiding kunnen geven om het stikstofdossier op een andere manier tóch modelmatig te blijven benaderen, bijvoorbeeld via emissie-modellen in plaats van depositie. Dat is geen goed idee, omdat alle problemen die ik heb besproken dan opnieuw zullen opdoemen, hoewel wellicht in een andere verschijningsvorm.

Je zou kunnen stellen dat de problemen die ik heb signaleerd symptoomverschijnselen zijn, en dat het onderliggende probleem dieper zit. In het stikstofdiscours wordt de wetenschap gevraagd om antwoorden te geven op vragen die de

wetenschap ten principale niet kán beantwoorden. De wetenschap hoort daar verantwoordelijkheid voor te nemen, maar doet dat in mijn optiek niet of te weinig – ik schreef dat al eerder. Elke puur modelmatige benadering van het stikstofdiscours is uiteindelijk tot mislukken gedoemd is. Een volledige bespreking van dit punt vraagt veel meer ruimte dan ik hier tot mijn beschikking heb. Ik noem het hier kort omdat dit punt in mijn ervaring vaak over het hoofd wordt gezien.

De kwestie is dat stikstof als enkelvoudige beleidsparameter voor natuurbeleid een myriade aan consequenties heeft voor tal van sociaal-economische activiteiten in Nederland die nooit modelmatig inzichtelijk gemaakt kunnen worden. Dat is deels terug te voeren, zoals ik heb laten zien, op het circulaire proces van niet-begrepen en inaccuraat toegepaste statistische kaders die hebben geleid tot een fictieve precisering van de impact van stikstofverbindingen op de Nederlandse natuur.

De wereld is verder te complex en kan niet worden gereduceerd tot of gestuurd vanuit stikstof. Noch gereduceerde stikstofemissies van alle economische actoren, noch alle bijbehorende technologische en sociaal-economische consequenties, noch natuurkwaliteit kunnen worden begrepen vanuit het gevoerde stikstofbeleid. Bedenk dat dit beleid, met niet meer dan KDW's (met niet-bestaande ecologische precisie), en Aerius/OPS (dat geen stikstofemissies, transport en deposities met enige functionele nauwkeurigheid kan modelleren) wordt geïmplementeerd. Met het stikstofbeleid worden politieke en beleidsmatige pogingen gedaan om aan de hand van één chemisch element - stikstof - talloze economische activiteiten van hogerhand te 'stroomlijnen' en te structureren ten behoeve van de natuur (de Wet van Goodhart indachtig; zie boven).

Dit is geen kwestie van de noodzaak van voortschrijdend inzicht of grotere behoefte aan empirische en mathematische kennis. Het huidige stikstofbeleid kan nooit datgene opleveren wat wordt beoogd of beloofd. Dat is dan ook de belangrijkste reden dat het niet zinvol is om nog gecompliceerdere modellen te gaan beschouwen dan Aerius/OPS. Wetenschap geeft ons nuttige informatie over bepaalde vragen, maar is naar haar aard ook beperkt. Dat is geen diskwalificatie van wetenschap, maar juist een van haar sterke punten: om nuttig te kunnen zijn moet de wetenschap zich beperken tot haar natuurlijk (metaforische) habitat. In het stikstofdiscours wordt deze voorwaarde met voeten getreden.

5 - Modelgebruik voor beleid in het algemeen

De aard van modellen

Waarom maken en gebruiken we eigenlijk wiskundige modellen? We willen begrijpen wat het gevolg zal zijn van bepaalde interventies, of wat er gebeurt bij juist het uitblijven daarvan. We willen snappen in hoeverre stikstofdepositie wordt beperkt als gevolg van bepaalde maatregelen, we willen weten hoe het klimaat zich zal ontwikkelen als we niets zouden doen, of we willen weten of windmolens niet te veel hinderlijk geluid gaan genereren wanneer ze op een specifieke locatie worden geplaatst. Omdat de werkelijkheid extreem complex is maken we wiskundige modellen die hopelijk op de cruciale punten iets over de werkelijkheid kunnen zeggen. Echter, zelfs als dat lukt is de huidige werkelijkheid beschrijven nog heel iets anders is dan de toekomst voorspellen. Dat dit niet makkelijk kan zijn is een understatement.

Modellen worden door mensen gemaakt, en hierover is meestal helemaal geen consensus. Als een overheidsorgaan kiest voor de inzet van een model bij beleid, dan wordt kritiek daarop vaak gepareerd met dat 'dit het beste is wat we hebben'. Behoorlijk bestuur, rechtmatig en verstandig gebruik van modellen vragen echter om meer: niet alleen om goed doordachte modellen, maar ook om kennis en inzicht bij ambtenaren, juristen en bestuurders over de kwaliteit en (uit)werking van die modellen en de gevolgen die deze kunnen hebben.

Het is hierbij belangrijk dat bestuurders en juristen het besef ontwikkelen dat geen enkel model te interpreteren is zonder context. Wat helpt is te starten met het stellen van een aantal basale vragen over de gebruikte modellen. Wie waren de makers? Welke opdracht hadden ze? In hoeverre heeft de politiek zich wel of niet bemoeid met de opdracht? Welke aspecten van de realiteit vonden de makers dermate belangrijk dat ze deze in hun model hebben verwerkt en welke niet? Welke motivatie hadden ze daarvoor? Is het model mede ingegeven door bekende fysische, biologische, fysiologische, sociale, psychologische, economische of andere processen die we goed begrijpen of beheersen, en hoe controversieel is ons begrip daar over? Zijn er data die aan de keuzes binnen het model ten grondslag lagen of niet? Welke grootheden hebben de makers gekwantificeerd, en op basis waarvan gebeurde dat? Welke aspecten van het model zijn echt betrouwbaar, en

welke zijn meer gebaseerd op een inschatting? Zijn de ingrediënten van het model zelf goed begrepen? Hoeveel vertrouwen hebben de makers zelf in het voorspellend vermogen van hun model, en waarop is dat vertrouwen precies gebaseerd? En *last but not least*, welke onzekerheden zijn er in de uitkomsten en hoe zijn die bepaald?

Een wiskundig model is dus zeker geen objectief bolwerk van kennis, maar een meestal organisch gegroeid geheel van methodes, ideeën, gedachtes, keuzes, belangen, patronen en conclusies. Dat betekent dat een simpele verwijzing naar een modeluitkomst niet zomaar als argument voor bepaalde beleidsbeslissingen gebruikt kan worden; daarvoor ontbreekt eenvoudigweg de nodige objectiviteit en zitten er te veel keuzes in het model.

Verstandig en wijs gebruiken van wiskundige modellen is daarom geen sinecure, zeker niet voor bestuurders, juristen en beleidsmakers die daarmee weinig ervaring hebben en daarbij ook weinig ondersteuning krijgen. In een eerdere publicatie⁶⁷ schreef ik al dat een model om al deze redenen nooit het eindpunt van een gesprek kan zijn, maar alleen het begin. Met een model maak je duidelijk welke aspecten jij belangrijk vindt, en dat is nuttig om een gesprek op gang te krijgen en een startpunt te hebben voor het ontwikkelen van beleid.

Echter, een model zien als begin van een gesprek is moeilijk verenigbaar met het opnemen van een model in wet- en regelgeving. De eerste conclusie is dan ook dat een model naar haar aard nooit zomaar mag worden ingezet, ook niet als 'de' wetenschap beweert dat er niets beters is. Overigens bestaat 'de' wetenschap niet. Er bestaan wetenschappers die iets beweren, maar niet zelden spreken ze elkaar behoorlijk tegen. Dat hoort bij wetenschap maar heeft natuurlijk gevolgen. Want wie volg je dan als je zegt dat je de wetenschap volgt?

Reële doelstellingen

Doelstellingen kunnen per definitie niet in modeltermen worden geformuleerd – doelstellingen gaan immers over de werkelijkheid en niet over modellen van die werkelijkheid. Zo kun je bijvoorbeeld als doel formuleren dat een bepaalde uitstoot (een fysisch gegeven, al dan niet bekend) met 50% zou moeten dalen. In veel gevallen is meten aan die uitstoot echter niet zo makkelijk of zelfs onmogelijk, en zou je modellen willen

⁶⁷ Marc Jacobs en Ronald Meester, *Van aardbeving tot zoönose – over de inzet van modellen voor beleid*, Mazirel Pers 2022.

gebruiken om de reductie van de uitstoot enigszins te monitoren. Zo'n modelbenadering is zinvol mits het model interpreteerbare uitkomsten genereert die op de een of andere manier verdedigbaar te verbinden zijn met de realiteit.

Deze observatie heeft consequenties. In het stikstofdossier wordt, zoals ik uitgebreid heb besproken, de KDW als richtlijn gebruikt, een grens aan de stikstofdepositie waar je onder zou moeten blijven om natuurverslechtering te voorkomen. Echter, of aan deze grenswaarde is voldaan wordt vaak alleen maar door het model uitgerekend. De drempelwaarde voor overschrijding is soms zo laag dat we geen mogelijkheid hebben om te meten of deze waarde gehaald wordt, en dan is het uitsluitend een modelexercitie.⁶⁸ De gezondheid van de natuur wordt op deze manier gereduceerd tot modelberekeningen, zonder observationele bevestiging.

Modellen zijn niet objectief, en de relatie met de realiteit is lang niet altijd even duidelijk. Soms is er overduidelijk geen relatie met de realiteit maar wordt het model toch gebruikt.⁶⁹ Een modelparameter als richtlijn is dus altijd omgeven met een bepaalde subjectiviteit, terwijl de verwijzing ernaar een bepaalde objectiviteit suggereert. Het model zegt het immers. Maar als een model eenmaal een juridische status heeft, kan een rechter niet anders dan dit model gebruiken.

Onzekerheid

Als we een bepaald proces wiskundig modelleren, dan is het uiteraard de bedoeling dat we de realiteit zo nauwkeurig mogelijk beschrijven. Hierbij worden noodzakelijkerwijs veel keuzes gemaakt: je kunt niet alle aspecten die mogelijk een rol spelen meenemen in je modellering; het doel van een model is immers juist om de beschrijving zodanig te vereenvoudigen dat deze te analyseren of te simuleren is. Dit alles leidt tot verschillende types onzekerheid:

1. De keuze van het model zelf. Deze keuze hangt af van de kennis die experts hebben van het te modelleren proces, maar ook van persoonlijke voorkeuren. Verschillende keuzes leiden tot potentieel verschillende uitkomsten. De onzekerheid die hiervan

⁶⁸ Zie bijvoorbeeld [Aanbiedingsbrief expertoordeel Arthur Petersen en peer reviews rekenkundige ondergrens stikstof | Kamerstuk | Rijksoverheid.nl](#). Dit is natuurlijk nog afgezien van de kritiek die ik op KDW's als richtinggevend heb gegeven.

⁶⁹ Tijdens de coronapandemie gebeurde dit bijvoorbeeld bij modellen die berekenden dat een QR-toegangsbewijs bepaalde gevolgen zou hebben, zie [Uitgeverij Ten Have | Wetenschap als nieuwe religie](#).

het gevolg is laat zich niet of nauwelijks zinvol wiskundig beschrijven. Het is iets waar elke modelmaker of –gebruiker zich bewust van dient te zijn en waarvoor verantwoordelijkheid genomen moet worden.

2. Als het model vastligt, is er doorgaans onzekerheid over de parameters ervan. Het kan niet genoeg benadrukt worden dat dit alleen onzekerheid binnen het model betreft. Immers, de parameters bestaan alleen maar in het model. Soms worden parameters geschat op basis van data, soms op basis van bekende wetten uit de fysica. Een schatting alleen, zonder enig inzicht in de onzekerheid die we daarover hebben is eigenlijk betekenisloos. Deze onzekerheid wordt in verschillende statistische paradigma's anders uitgedrukt. Je kunt bijvoorbeeld een betrouwbaarheidsinterval geven, maar vaak is het beter een zogenaamd geloofwaardigheidsinterval te maken. De verschillen tussen die soorten intervallen zijn subtiel maar belangrijk. De onzekerheid die deze intervallen uitdrukken dient ook weer geïnterpreteerd te worden. Elke gekwantificeerde onzekerheid moet beargumenteerd worden.
3. Er kan grote onzekerheid ontstaan bij meetactiviteiten en andere vormen van datavergaring.

De combinatie van de punten hierboven maakt dat modeluitkomsten op een zeer ingewikkelde manier onzeker zijn, en maken dat deze onzekerheden zelfs niet zomaar geformuleerd kunnen worden. Het gaat hier trouwens altijd over *onze* onzekerheid als mensen; we hebben het hier niet over de vraag of de werkelijkheid zélf onzeker is of niet maar om onze kennis ervan. Niet alleen is de interpretatie lastig, maar ook zal elke poging om de uiteindelijke onzekerheden te kwantificeren op allerlei aannames gebaseerd moeten zijn.

Deze beschouwing was in grote lijnen onderdeel van een artikel dat ik op uitnodiging van een juridisch tijdschrift had geschreven. Ondanks het feit dat ik uitgenodigd was om hierover een bijdrage te schrijven, weigerde de redactie mijn manuscript. Om de hardnekkigheid van bepaalde vooroordelen te illustreren, citeer uit haar argumentatie hiervoor:

“De redactie vindt ook dat de waarde van modellen nogal wordt weg gerelativeerd. Zo vindt zij de suggestie dat modellen [...] tot uitdrukking brengen wat de maker van het model erin heeft gestopt niet geloofwaardig.⁷⁰ Volgens de redactie is dit verregaande relativisme van modellen eenzijdig, waarmee alles ter discussie kan worden gesteld. Daarbij komt dat ook de rest van de wetgeving niet uit een objectieve bron voortvloeit, maar de uitkomst is van een continu politiek en maatschappelijk debat. De redactie vraagt zich dan ook af: is het dan zo verkeerd als aan een model een bepaalde visie ten grondslag ligt?”

Vooraf de laatste vraag is interessant en belangrijk. Mijn antwoord is dat aan elk model een visie ten grondslag ligt, en dat dit zeker niet ‘verkeerd’ is. Maar het is wel verkeerd om die visie als het ware te objectiveren door de uitkomsten van zo’n model letterlijk te nemen als ware het de realiteit. Een model is altijd een bepaalde afspiegeling van de realiteit, en als je anders naar die realiteit kijkt zie je ook andere dingen. Dat is uiteindelijk de reden dat ik er voor pleit om wiskundige modellen slechts te laten informeren, en ze nooit op te nemen in wet- en regelgeving.

De belangrijkste passage uit bovenstaand citaat is toch wel dat het kennelijk lastig is te begrijpen dat modellen de weerslag zijn van dat wat de makers erin hebben gestopt. Maar het kan gewoon niet anders zijn, want modellen rekenen slechts uit wat de makers hebben bedacht – meer niet. Dat heeft werkelijk niets met “relativisme” te maken, en het is cruciaal dat beleidsmakers dat doorzien.

Modellen en de algemene beginselen van behoorlijk bestuur⁷¹

De overheid is in Nederland gehouden te handelen in overeenstemming met het recht. Het betreft (i) specifieke wettelijke regels die voor bepaalde overheidshandelen gelden, (ii) hoger recht (zoals het EVRM, het EU-recht en de Grondwet) en (iii) de algemene beginselen van behoorlijk bestuur.⁷² Als in besluitvorming door de overheid gebruik wordt gemaakt van modellen, dan is van belang dat het gebruik van modellen in overeenstemming is met de

⁷⁰ Het is bijzonder dat de redactie niet kon accepteren wat echt algemeen en voor iedereen duidelijk moet zijn. Een model is de weerslag van wat de maker erin stopt, en kan ook niets anders zijn.

⁷¹ Deze sectie bevat passages uit het rapport *Een richtlijn voor het gebruik van modellen voor beleid voor ambtenaren, juristen, bestuurders en modellers*, door Marc Jacobs en Ronald Meester, opgesteld in opdracht van de Provincie Zuid-Holland.

⁷² Van Wijk/Konijnenbelt & Van Male 2014, p. 264-266.

voor de overheid geldende algemene beginselen van behoorlijk bestuur. Ik noem een aantal van deze beginselen en geef daarbij aan wat het gevolg daarvan is voor het gebruik van modellen in het algemeen maar vooral voor het gebruik van Aerius/OPS.

Zorgvuldige voorbereiding (art. 3:2 Awb)

“Bij de voorbereiding van een besluit vergaart het bestuursorgaan de nodige kennis omtrent de relevante feiten en de af te wegen belangen.”

Een model kan nooit alle relevante factoren en omstandigheden bevatten en ‘meenemen’. Die beperking van modellen kan afbreuk doen aan een zorgvuldige voorbereiding van het besluit. Een model, en Aerius/OPS in het bijzonder, kan dan alleen gebruikt worden als het bestuursorgaan weet heeft van die relevante feiten. In het onderhavige dossier valt hier bijvoorbeeld ook kennis van KDW’s onder: wat betekenen die KDW’s echt? Als de kennis ontbreekt dat KDW’s geen unieke getallen kunnen zijn, dan is dat mijns inziens op te vatten als een belangrijk gebrek aan kennis dat zorgvuldig gebruik van het model in de weg staat. Ook gebrekkige statistische kennis staat een dergelijk zorgvuldige proces in de weg.

Fair play

“Bestuursorganen moeten werken zonder vooringenomenheid, personen die een persoonlijk belang bij een besluit hebben, mogen de besluitvorming niet beïnvloeden en bestuursorganen mogen burgers niet de mogelijkheid ontnemen om voor hun belang op te komen door te beslissen op een in dit opzicht ongunstig gekozen moment.”⁷³

De overheid moet zich onbevangen opstellen waar het gaat om de inhoud van modellen. Als een model niet ter discussie kan staan en er ook geen vragen over kunnen worden gesteld, is het niet mogelijk om na te gaan in hoeverre (het gebruik van) een model vooringenomen is. Dat staat op gespannen voet met het fair play-beginsel.

⁷³ Van Wijk/Konijnenbelt & Van Male 2014, p. 282-283 en p. 207-208.

In het geval van het gebruik van Aerius/OPS is moeilijk in te zien dat dit fair play beginsel gehonoreerd wordt. De berekeningen en aannames zijn voor niemand te verifiëren of te controleren. Als een modeluitkomst niet strookt met de waargenomen realiteit dan is er eigenlijk geen manier waarop burgers dat kunnen inbrengen op een manier die enige invloed heeft.

Ik erken dat het moeilijk zal zijn om burgers te informeren over wat Aerius/OPS als model nu werkelijk wel en niet zegt. Ik zie hier ook een taak voor het RIVM zélf. Het RIVM is zich bewust van de onzekerheden van hun rapportages en zou mijns inziens moeten interveniëren wanneer de overheid deze onzekerheden niet adresseert of benoemt. De NDA's en de adviezen van de EA bevatten veel uitspraken die als vaststaand worden gepresenteerd, maar dat onmogelijk kunnen zijn. Het is in mijn ogen de verantwoordelijkheid van de wetenschap om in te grijpen als beleidsmakers de resultaten van deze wetenschap onjuist interpreteren of gebruiken.

Verbod van détournement de pouvoir (art. 3:3 Awb)

“Het bestuursorgaan gebruikt de bevoegdheid tot het nemen van een besluit niet voor een ander doel dan waarvoor die bevoegdheid is verleend.”

Vrij vertaald naar modellen betekent dit dat een model alleen gebruikt kan worden voor het doel waarvoor het is ontworpen. Is dit niet het geval, dan is sprake van strijd met het verbod van détournement de pouvoir.

Dit is een bijzonder saillant beginsel in het geval van Aerius/OPS, vooral in relatie tot de eerder genoemde wet van Goodhart. Deze wet beschrijft het fenomeen dat een maatstaf belangrijker kan worden dan het oorspronkelijke doel. Stikstof is in zichzelf een doel geworden, terwijl het toch echt over breder natuurbeleid gaat. We verkeren nu in de situatie dat het gebruik van Aerius/OPS verdedigd wordt als beleid voor natuurbescherming. Dat is echter onjuist.

Aerius/OPS was het wetenschappelijke antwoord op een beleidsvraag uit de oorspronkelijke Hinderwet, waarin uitgangspunt was dat er een relatie was tussen activiteit en depositie elders, en dat hierover vanuit de wetenschap iets over gezegd moest worden. Daar gaat het model over, en nergens anders over. Het model gaat niet over natuurbeleid.

Draagkrachtige en kenbare motivering (art. 3:46 Awb en art. 3:47 Awb)

“Een besluit moet een motivering bevatten; de motivering moet het besluit dragen, wat wil zeggen dat de wijze waarin in het besluit de feiten op de regel zijn toegepast, correct en overtuigend moet zijn.”

Van een besluit of basis van een model kan niet gezegd worden dat het draagkrachtig is gemotiveerd omdat toepassing is gegeven aan het model. In zekere zin geldt dat de overtuigingskracht van het besluit altijd een afgeleide is van de kennis van het model. We hebben gezien dat Aerius/OPS niet in staat is om op een betrouwbare manier uitspraken te doen over stikstofdeposities wanneer het over vergunningverlening gaat. Dat betekent vervolgens dat er geen deugdelijke of overtuigende motivering gegeven kan worden voor het wel of juist niet afgeven van vergunningen.

Rechtszekerheid

“Regels moeten duidelijk zijn geformuleerd en daadwerkelijk worden toegepast, de toepassing moet consistent zijn, zodat de uitvoering ervan voorspelbaar is en de betekenis ervan niet verschilt al naar gelang degene die de regel toepast.”⁷⁴

Modellen kunnen updates krijgen of op andere manieren worden aangepast waardoor de toepassing ervan tot een andere uitkomst leidt. In gevallen waarbij AI een rol speelt, kan door gewijzigde input een model zich anders gaan gedragen. In het licht van het rechtszekerheidsbeginsel kan dat problematisch zijn.

Op dit punt is er geen probleem met het toepassen van Aerius, want het model is in principe overal hetzelfde.

⁷⁴ Van Wijk/Konijnenbelt 2014, p. 305.

Gelijkheid

“Gelijke gevallen worden gelijk behandeld en ongelijke gevallen ongelijk naar de mate waarin zij verschillen.”

Verschillende bestuursorganen kunnen uiteenlopende modellen gebruiken voor dezelfde beleidsvragen. Dat kan een strijd opleveren met het gelijkheidsbeginsel.

Dit lijkt mij bij het gebruik van Aerius/OPS niet problematisch omdat het model altijd wordt toegepast.

Materiële zorgvuldigheid

“Bij beslissingen die de overheid neemt, moeten alle relevante aspecten bij een besluit worden meegenomen; besluiten moeten zo weinig mogelijk schade veroorzaken: de lasten die voor iemand uit een besluit voortvloeien mogen niet zwaarder zijn dan strikt noodzakelijk en besluiten mogen evenmin op sommige (categorieën van) personen, in vergelijking met anderen, onevenredige lasten opleggen.”⁷⁵

Besturen op basis en met behulp van modellen kan op gespannen voet staan met dit beginsel, dat ook wel ‘het beginsel van de minste pijn’ wordt genoemd, bij voorbeeld omdat onvoldoende rekening is gehouden met de implicaties van de toepassing van het model voor bepaalde personen.

In het Aerius/OPS dossier is deze zorgvuldigheid nadrukkelijk in het geding, vanwege de statistische onzekerheden die ik besproken heb. De onevenredigheid van lasten voor agrariërs bij een zeer geringe berekende (niet gemeten!) overschrijding van de KDW versus de enorme financiële en juridische gevolgen is vanwege dit beginsel niet acceptabel. Bovendien is het niet mogelijk dat “alle relevante aspecten” in het model zitten. Wat wel of niet relevant is kan niet aan modelmakers worden overgelaten, omdat dit het maken van maatschappelijke keuzes veronderstelt.

⁷⁵ Van Wijk/Konijnenbelt 2014, p. 327-328.

Evenredigheid (art 3:4 lid 2 Awb)

“De voor een of meer belanghebbenden nadelige gevolgen van een besluit mogen niet onevenredig zijn in verhouding tot de met het besluit te dienen doelen.”

De verhouding tussen het effect en het middel moeten in balans zijn. Dat is iets waar je in veel modellen tevergeefs naar zult zoeken. Niet alleen zal het model slechts bij benadering iets zeggen over een bepaalde grootte, maar belangrijker is dat ‘balans’ een menselijke en niet objectiveerbaar begrip is dat zich niet door een model laat bepalen. De keuze wát een goede balans is, is altijd mensenwerk. Dat staat haaks op de huidige Aeries/OPS-praktijk waarin een model, een berekening dus, het laatste woord heeft, en dat daardoor een extreem technocratisch karakter heeft gekregen.

6 - Vragen bij elk modelgebruik

In een eerdere publicatie ontwierpen Marc Jacobs en ikzelf een lijst van 11 vragen die bij elk gebruik van modellen voor beleid opgeld doen.⁷⁶ In die publicatie gingen we ook al in op de 11 vragen in relatie tot het stikstof discours, maar met de kennis die mijn onderzoek nu heeft opgeleverd zal ik de vragen opnieuw kort adresseren. Hieronder volgen de vragen met mijn reflectie. De vragen figureren rondom de centrale these in ons boek, dat modellen het begin vormen van een gesprek, en niet het einde.

1. Adresseert het model eigenlijk wel de juiste vraag? Wat denken we te kunnen winnen met het gebruiken van een model in vergelijking met beleidskeuzes zonder een model te maken?

Een uitleg waarom stikstof wordt berekend en gemodelleerd in plaats van overal gemeten, vinden we op de website van het RIVM:⁷⁷

“Het is niet mogelijk om de concentratie en depositie van stikstof in heel Nederland met metingen in kaart te brengen. Dat kost te veel geld en is praktisch onuitvoerbaar. Om toch een landelijk dekkend beeld te geven, en om trends en toekomstverwachtingen te kunnen maken, werkt het RIVM met rekenmodellen. Computers voeren berekeningen uit op basis van modellen.”

Wat we dus kunnen winnen, aldus het RIVM, is dat we informatie kunnen krijgen over die (vele) locaties waar we niet kunnen meten. We hebben echter gezien dat dit onjuist is. De berekeningen kunnen deze belofte niet waarmaken – daarvoor is er veel te veel onzekerheid in het spel. Bovendien wordt bij het gebruik van modellen verondersteld dat deze modellen

⁷⁶ Marc Jacobs en Ronald Meester, *Van aardbeving tot zoönose – over het gebruik van modellen voor beleid*, Mazirel Pers 2023.

⁷⁷ <https://www.rivm.nl/stikstof/berekenen-en-modelleren>

antwoorden gaan geven die we vervolgens kunnen inzetten voor het oplossen van allerlei maatschappelijke problemen. Modellen zijn hiervoor niet geschikt.

2. Wat beoogt het model precies te modelleren? Is het een voorspellend model? Is het een beschrijvend model?

Het model probeert om drie ogenschijnlijk samenhangende vragen tegelijkertijd te beantwoorden: “Waar zien we hoeveel stikstofdepositie, en waar komt deze vandaan?” Het model probeert dus te beschrijven wat er gebeurd is, maar doet dat op een voorspellende manier door te modelleren hoeveel stikstof er wordt uitgestoten én waar de stikstof dan terecht komt. Het gebruik van het model veronderstelt impliciet dat we deze vragen kunnen beantwoorden. Gezien mijn analyse is deze veronderstelling niet gerechtvaardigd.

3. Als het model moet bijdragen aan besluitvorming, is er dan nagedacht over de risico's en consequenties van eventuele slecht uitpakkende beslissingen?

“Het kabinet heeft besloten dat iedereen de pijn zal moeten voelen. Dat daarmee wellicht méér gedaan wordt dan nodig is, willen de ambtenaren op de koop toe nemen, vanwege de onzekerheid van de modellen en omdat het omgekeerde nog minder wenselijk is: dat uiteindelijk de opgave naar boven moet worden bijgesteld en er dus meer gedaan moet worden dan eerder gedacht. Dat wil zeggen dat ook veehouders die niet zoveel bijdragen aan het stikstofprobleem, bijvoorbeeld omdat ze ver van natuurgebieden boeren, een (kleine) bijdrage zullen moeten leveren. In het goedkoopste, efficiëntste scenario zouden 5300 boeren moeten worden uitgekocht. Andere bedrijven hoeven dan niet te krimpen. Dat vergt wel ‘zeer gericht’ beleid, schrijven de ambtenaren, waarbij boeren in bijvoorbeeld de Gelderse Vallei zo’n beetje allemaal zouden moeten stoppen en andere doelen, bijvoorbeeld op het gebied van klimaat en waterkwaliteit, niet zijn meegewogen”.⁷⁸

⁷⁸ <https://www.rtlnieuws.nl/nieuws/politiek/artikel/5322446/stikstofberekeningen-stikstof-boeren-landbouw-veeteelt-protesten>

Dit citaat redeneert sterk vanuit de behoefte van de overheid, maar niet vanuit de mensen die de gevolgen van het overheidsbeleid zullen ondervinden. De besluiten van de overheid steunen sterk op technocratische processen waarin menselijkheid al snel achter de horizon verdwijnt. Een eventueel besluit om 'niets te doen' betekent overigens niet dat er niet nagedacht wordt. Het betekent dat een keuze wordt gemaakt om niet te handelen omdat de consequenties van een besluit of niet opwegen tegen de voordelen of onduidelijk zijn. Het is niet gezegd dat we per se iets moeten doen.

Dit is allemaal niet nieuw; ik wil de lezer in dit verband de volgende waarschuwing van James R. Otteson niet onthouden:⁷⁹

"The [...] reason that we should prefer individuals' judgments about how to arrange their own time, talent, and treasure over the judgments of third parties, however expert, is because of the incentives involved. If I follow your proposed policy and it does not benefit me, or perhaps even harms me, on whom do the negative consequences fall? It is not on you; it is on me. But because you did not get that feedback, it may not incline you to correct, or even respond to, this consequence of your policy. I, on the other hand, who did suffer the negative consequence, would presumably respond to it, but since I am not making the decisions about the policy, my response has muted effectiveness."

4. Is er consensus in de wetenschappelijke literatuur of onder wetenschappers over het gebruik van dergelijke modellen?

Dit is een gevaarlijke vraag. Wat we er níet mee bedoelen is dat consensus zou betekenen dat een model om die reden aanvaardbaar of acceptabel zou zijn. De vraag is ook waarschuwend bedoeld: als er consensus geclaimd wordt, wat zou dat dan betekenen? Wetenschappelijke consensus – wat dat ook precies betekent – functioneert alleen maar binnen de wetenschap. Het is best mogelijk dat 'de' wetenschap het er over eens is dat er niets beters zou zijn dan, zeg, Aerius/OPS. Maar dat betekent helemaal niet dat dat model

⁷⁹ James R. Otteson, *The end of socialism*, 2015.

daarmee dus geschikt zou zijn voor het doel waar het voor nodig is. De opmerking dat ‘dit het beste is wat we hebben’, door politici veel gebezigd, is om die reden niet zinvol. Er is namelijk helemaal geen a priori noodzaak om je tot de wetenschap te wenden – sterker nog – het is vaak evident duidelijk dat de wetenschap de gestelde vragen over wat wel of niet goed is om te doen, simpelweg niet kán beantwoorden; zie ook het vorige hoofdstuk.

5. Wat weten we precies over het realiteitsgehalte van het model waar we over praten?

Hier heb ik het in dit rapport uitgebreid over gehad – het realiteitsgehalte is onvoldoende voor de toepassingen waar het voor bedoeld is.

6. Wie heeft het model geconstrueerd, en welke belangen hebben de makers hier eventueel zelf bij?

Het Aeries/OPS model komt van het RIVM. Het lijkt mij belangrijk om over belangen niet naïef te zijn. Het RIVM heeft er als instituut vanzelfsprekend belang bij dat het model gebruikt wordt. Het feit dat Aeries/OPS niet zomaar toegankelijk is, is daar wellicht een gevolg van. Ik stel mij persoonlijk op het standpunt dat wetenschap die naam niet mag dragen als niet alles transparant en toegankelijk is. Dat is in dit dossier zeker niet het geval.

7. Wie worden geraakt door de beleidsvoornemens van de overheid? Zitten deze mensen of groepen aan tafel? Hebben zij de gelegenheid gehad het model diepgaand te bestuderen, en zijn ze in de gelegenheid gebracht om kritiek te leveren of alternatieven aan te dragen?

Een diepgaande studie van de gebruikte modellen is voor de meeste mensen die er belang bij hebben onmogelijk. Dit pleit er in zichzelf al voor om voorzichtig te zijn met het gebruik van modellen bij beleid.

8. Waarom heeft de overheid voor dit specifieke model gekozen? Zijn er alternatieven onderzocht?

Het is onduidelijk of er meerdere modellen zijn onderzocht. Een kort artikel in de Telegraaf van 2019 kopt met een citaat van minister Carola Schouten: “We hebben geen beter stikstofmodel”.⁸⁰ Argumenten worden niet aangedragen. Maar zelfs als dat waar zou zijn, dan betekent dat, zoals ik heb uitgelegd, helemaal niet dat we dan dit model dan maar zouden moeten gebruiken.

9. Kunnen de gesprekspartners zich vinden in de aannames van het model, en waarom dan precies?

Vanwege de complexiteit van het model is het antwoord op de vraag een duidelijk nee.

10. Is het überhaupt te verwachten dat we van modellen veel leren in dit dossier, en waarom precies?

Deze vraag heb ik impliciet en expliciet besproken in dit verslag. Ik heb geen aanwijzingen gezien dat we van de modellen iets leren. Daar waren ze ook niet voor bedoeld. De modellen hadden als doel om als een rekenmachine ingezet te worden. Wat we leren zijn numerieke uitkomsten, maar we hebben gezien dat deze uiterst onbetrouwbaar zijn, en is het dus zo dat we er uiteindelijk vrijwel niets van leren.

11. Is er een ethische component in het spel, en hoe verhoudt het model zich hiertoe?

Bij nader inzien is de formulering van de vraag enigszins bevreedend. Het model zelf verhoudt zich natuurlijk nergens toe. De vraag is eigenlijk of het model een kwestie betreft die een ethische component heeft. Dat is hier zeker het geval, en vragen daarover zijn uiteraard niet het domein waarover modellen iets te zeggen zouden hebben. Dit is reden te

⁸⁰ <https://www.telegraaf.nl/nieuws/393895100/schouten-we-hebben-geen-beter-stikstofmodel>

meer om af te stappen van een rigide en technocratisch proces waarin modellen het voor het zeggen hebben. Dit wil natuurlijk niet zeggen dat er bij beleid voor sommige mensen of groepen geen bijkomende schade kan ontstaan, dat is meestal niet helemaal te vermijden. Maar voor die schade dienen beleidsmakers verantwoordelijkheid te nemen, en deze kan niet worden afgewend op 'de' wetenschap. Ethiek is mensenwerk, geen modellenwerk.

7 - Conclusies en advies

Mijn onderzoek naar de statistische aspecten van het gebruik van Aerius/OPS heeft heel wat kennis opgeleverd. Ik vat de belangrijkste conclusies nog eens samen.

Ik heb allereerst vastgesteld dat de NDA's en de adviezen van de EA blijken te geven van een 'maakbaar' wereldbeeld, waarin metaforische knoppen zijn bedacht waar we aan kunnen draaien teneinde een bepaald doel te bereiken. KDW's zijn in dit wereldbeeld unieke en exacte getallen. Uitkomsten van Aerius/OPS berekeningen zijn reëel en worden allemaal zonder marges of onzekerheden gebruikt. De natuur zal herstellen als de depositie onder de KDW komt, en boven de KDW zal depositie 'significante' schade aan de natuur kunnen toebrengen.

Het is een simplistisch wereldbeeld met weinig of geen realiteitswaarde. Echter, de overheid gaat er gretig in mee. Zo staat in de toelichting van de stikstofkaart het volgende:⁸¹

"De aanpak van de stikstofcrisis vergt veel hervormingen, waaronder in de landbouw. Het huidige regeerakkoord bevat het doel om in 2030 de stikstofdepositie op 74% van de Natura2000-gebieden naar een onschadelijk niveau te hebben gebracht, oftewel, onder de Kritische Depositiewaarde (KDW) van elk gebied. In 2050 moet dit 100% van de Natura2000-gebieden zijn."

Op 21 juli 2022 berichtte RTL nieuws hierover als volgt:⁸²

"Er zijn veel verschillende manieren om de stikstofdoelen te halen. Uiteindelijk is voor het ministerie van Landbouw maar één doel heilig: dat 74 procent van de natuur in 2030 zodanig verbeterd is dat de kritische depositiewaarde niet langer wordt overschreden."

In het licht van wat ik in dit rapport heb besproken klinkt dit "heilig" bijna hallucinant.

⁸¹ Document niet meer online te vinden maar wel in mijn bezit.

⁸² [Goedkoper de stikstofdoelen halen: het kan, maar is het wenselijk?](#)

Ik heb bijzondere aandacht geschonken aan de bepaling van deze KDW's. De statistiek kan niet bepalen wat een kritische waarde zou moeten zijn, om de simpele reden dat dit een waardeoordeel is dat alleen mensen kunnen geven. Een beroep op het streven naar statistische significante veranderingen en/of een definitie in die termen heb ik terzakekundig afgewezen. Statistische significantie heeft weinig of niets te maken met dingen die voor ons of voor de natuur belangrijk zouden kunnen zijn en de wijzen waarop we dat wetenschappelijk zouden kunnen vaststellen.

De gedachte dat er een unieke en welgedefinieerde KDW zou bestaan is onjuist. Het is zeker juist dat er 'te veel' stikstofdepositie kan zijn, voor elk redelijk criterium, om de simpele reden dat van alles te veel kan zijn. Echter, geen enkele wetenschapper die ik heb gelezen of gesproken, verdedigt vanuit die wetenschap een unieke waarde. Ik heb wel - en ik heb daarover gerapporteerd - wetenschappers gezien die unieke waardes propageren omdat de politiek en de rechterlijke macht deze accepteert of zelfs eist. Dat heeft echter natuurlijk niets met wetenschap te maken; wetenschappers zouden een dergelijke verleiding moeten weerstaan. Het huidige gebruik van KDW's (en drempelwaardes) is daarmee vanuit statistisch perspectief onwetenschappelijk, ondeugdelijk en daarmee dus niet verdedigbaar.

We hebben dus een situatie waarin geen enkele wetenschapper publiekelijk wil erkennen dat de KDW's niet deugen, terwijl ze in persoonlijke gesprekken wel degelijk grote twijfels uiten. Dat is een bijzondere en ook ongewenste situatie. In het al eerder aangehaalde instellingsbesluit van de EA lezen we bijvoorbeeld dit in de toelichting:

"De advisering door de Ecologische Autoriteit, gelet op de rol die is benoemd in het coalitieakkoord, heeft betrekking op gevoelige thematiek, waarvoor een gezaghebbend advies nodig is. De wetenschappelijke kennis en expertise die hiervoor nodig zijn, mag niet beïnvloed kunnen worden door politiek bestuurlijke visies en keuzes."

Als de EA dus twijfelt aan de waarde van de KDW's, en dat doet ze, dan dient ze dit publiek te verkondigen. Dat geldt in elke omstandigheid, voor elke wetenschapper. Geuite twijfels in individuele gesprekken blijven zonder gevolgen voor de samenleving, terwijl de KDW's zoals publiekelijk verstaan dat juist wel (ten onrechte) hebben.

Vervolgens heb ik de onzekerheden en onbepaaldheden van het Aerius/OPS rekenmodel zélf bestudeerd. Zoals elk model geeft Aerius/OPS slechts terug wat de makers er in hebben gestopt. Zelfs als we uitkomsten middelen over ruimte en/of tijd, wijken de gevonden waardes vele tientallen procenten af van wat 'gemeten' wordt. Daar komen de onzekerheden van die 'meet'resultaten nog eens bovenop, die in dit specifieke dossier groot maar niet of nauwelijks kwantificeerbaar zijn. Hoe het model presteert op het individuele niveau bij vergunningverlening weten we dus helemaal niet, want de berekende toenames van deposities zijn zo klein dat we die niet kunnen verifiëren met 'metingen'. Wat we wel weten is dat de onzekerheden hier nog groter zullen zijn dan bij de gemiddelden.

Samengenomen concludeer ik dat we met de combinatie van de onzekerheden in de KDW's en de berekende deposities van Aerius/OPS feitelijk niet weten wat we aan het doen zijn. We zijn blind modeluitkomsten aan het volgen zonder de mogelijkheid te hebben om te controleren waar we mee bezig zijn. Er wordt veel beweerd over stikstof, maar er is vrijwel geen enkele kwantitatieve claim die daadwerkelijk controleerbaar is. Het werken met de combinatie van KDW's en Aerius/OPS acht ik vanwege dit alles niet wetenschappelijk.

Dat conflicteert met de gangbare praktijk in de NDA's, de adviezen van de EA en de vergunningverlening. We hebben immers gezien dat uit de NDA's en de adviezen van de EA een maakbaar wereldbeeld opdoemt; een wereld die goed te beschrijven en te controleren zou zijn door wiskundige modellen.

Mijn analyse van de onzekerheden in zowel de KDW's als de modelberekeningen van Aerius/OPS laten echter zien dat dit wereldbeeld afgewezen moet worden. Het is eenvoudigweg niet het geval dat berekeningen van Aerius/OPS in combinatie met de KDW's tot een werkbare of zelfs maar zinvolle procedure kunnen komen. Het is evident het geval dat de wetenschap niet kan leveren wat de EA voorstaat in haar adviezen.

Een antwoord dat dit "het beste is wat we hebben" is onduidelijk omdat het uitgaat van de gedachte dat de wetenschap móet en kán leveren. In plaats van net te doen of we precieze informatie hebben en daar beleid met zeer grote maatschappelijke en particuliere gevolgen op baseren, kunnen we ook accepteren dat we bepaalde dingen niet kunnen weten, en dat 'de' wetenschap hierop geen antwoord zal geven.

Ook langs de weg van de beginselen van behoorlijk bestuur heb ik laten zien dat modelgebruik hierbij problematisch is. De aard van wiskundige modellen maakt dat aan deze beginselen niet voldaan kan worden.

Tenslotte heb ik gesuggereerd dat de wetenschap ten principale overvraagd wordt in het stikstofdiscours. De kennis die ze geacht wordt te leveren, kán ze niet leveren. Wetenschappers horen te interveniëren als hun resultaten incorrect geïnterpreteerd of zelfs misbruikt worden. Dat gebeurt in dit discours te weinig en dat is de wetenschap aan te rekenen.

Samenvattend is de model-wetenschappelijke benadering van het stikstofdiscours zowel praktisch, theoretisch als filosofisch problematisch, en is dus een andere benadering nodig.

Advies

Gezien de statistische beperkingen die mijn onderzoek zichtbaar hebben gemaakt ligt de eerste conclusie voor de hand: Aerius/OPS levert uitkomsten die zo onzeker zijn dat deze niet wetenschappelijk en statistisch verdedigbaar toegepast kunnen worden in de NDA's, de adviezen van de EA en voor vergunningverlening. Werken met deze software komt neer op het blind varen in een onbekend territorium.

De oplossing van deze situatie is zeker niet om in plaats van Aerius/OPS een ander, eventueel nog geavanceerder model te gaan gebruiken. Er is namelijk voor welk geavanceerder of ander model dan ook, eveneens geen manier om uitkomsten te verifiëren. Het probleem is niet dat Aerius/OPS een 'verkeerd' of 'fout' model is, maar dat de situatie vanuit de wetenschap niet op een bevredigende manier bestudeerd kan worden op het aggregatieniveau dat voor de NDA's, de EA adviezen en de vergunningverlening nodig is. Geavanceerdere modellen zijn veelal helemaal niet beter dan eenvoudige.

De oplossing van deze situatie is óók niet om in plaats van stikstofdepositie naar emissies te gaan kijken binnen een emissie-rekenmodel zoals het NEMA. De problemen die we hebben besproken doen daar immers ook allemaal opgeld en gaan de onzekerheden niet wegnemen of verminderen. Als we dit zouden doen (en er zijn aanwijzingen dat deze richting overwogen wordt) dan blijven we in de modellenwerkelijkheid werken en lossen we het probleem van de onwetenschappelijkheid niet op. Er verandert dan niet zo veel, en het probleem wordt dan alleen maar verplaatst.

De oplossing van De Nieuwe Denktank om veel hogere drempelwaardes te gaan gebruiken is weliswaar een stap in de goede richting, maar als we dat doen blijven Aerius/OPS en de KDW's voor vergunningverlening gehandhaafd, en blijven we dus denken

en werken binnen het discours van een modellenwerkelijkheid. Het feit dat De Nieuwe Denktank alle drempelwaardes van 1 – 35 mol/ha/jaar allemaal wetenschappelijk verantwoord vindt, geeft in zichzelf trouwens ook al aan dat de grenzen van een wetenschappelijke benadering allang bereikt zijn.

Waarom zouden we blijven rekenen met Aerius/OPS, emissie-modellen of KDW's als de uitkomsten ervan zo overduidelijk onbetrouwbaar en onzeker zijn? We zijn in dit dossier de maat der dingen kwijtgeraakt, en hebben als land de weg van de modellenwerkelijkheid gekozen in de hoop dat de wetenschap en haar modellen ons hierbij zouden kunnen leiden. Maar mijn onderzoek laat (in navolging van andere studies) zien dat dit ijdele hoop is – de wetenschap wordt hier overvraagd. De wetenschap gaat, nogmaals, geen antwoorden opleveren.

Mijn advies is dan ook om helemaal af te zien van de modellenwerkelijkheid, en de KDW's, Aerius/OPS en andere modellen niet langer te gebruiken. Het dossier moet van het bord van de modelleurs af, en terug op het bord van mensen, maatschappij, beleidsmakers en politici. Het beleid moeten we niet langer laten bepalen door een fictieve modellenwerkelijkheid, maar door de empirie – de waarneming dus, in de breedste zin van het woord. De staat van de natuur kan niet worden afgelezen door een model van welke aard dan ook. Ik legde ook al uit dat de werkelijkheid zo ongelooflijk complex is dat we het beleid onmogelijk vorm kunnen geven vanuit één specifiek chemisch element – die filosofische boodschap moeten we niet vergeten.

Het feit dat ik adviseer om af te zien van het werken in een modelwerkelijkheid impliceert dat ik adviseer om een wetenschappelijk en maatschappelijk goed onderbouwde nieuwe methodiek op te zetten voor het doel en de vorm van de NDA's, de adviezen van de EA en de vergunningverlening.

De methodiek moet wetenschappelijk verdedigbaar zijn *voor zover* de wetenschappelijke bemoeienis met het dossier kan reiken. Dat impliceert allereerst dat er naast ecologen ook vertegenwoordigers van andere wetenschappelijke disciplines aan boord moeten zijn, zoals datawetenschappers, statistici, sociologen, organisatiedeskundigen, risico-analisten, et cetera. Het dossier omvat veel verschillende maatschappelijke aspecten die veel verder reiken dan slechts ecologie. Ik onderstreep echter het feit dat veel vragen geen wetenschappelijk antwoord zullen hebben omdat die vragen naar hun aard niet wetenschappelijk zijn. Dat betekent alles bij elkaar niet dat wetenschappers geen rol dienen

te spelen, maar wel dat die rol in proportie moet worden gezien. Wetenschap is in deze benadering dienend, niet leidend.

Politiek en bestuur enerzijds, en wetenschap anderzijds, dienen verder strikt gescheiden te zijn, en houden elk hun eigen verantwoordelijkheid. Dat EA moet, wil ze gezaghebbend zijn in het wetenschappelijk discours, ook de kennis in huis halen om ook echt volwaardig wetenschappelijk te kunnen adviseren. Het feit dat de EA zelf geen onderzoek doet betekent namelijk niet dat ze het wetenschappelijk discours alleen maar als een gegeven moet zien. Ze moet dit discours doorgronden in al haar facetten en er een filosofie over hebben. Een wetenschappelijk advies moet geen rekening houden met de vraag wat een rechter ervan zal vinden.

We hebben verder concrete en normerende regels nodig voor het werken met modellen. Hieronder doe ik een voorzet voor een aantal regels die daarbij op de een of andere manier gegarandeerd zouden moeten worden.

1. Overheidsbeleid mag nooit de indruk wekken dat wetenschappelijke inzichten vaststaande feiten zijn.
2. Wetenschap is een onderzoekend en voortgaand proces. In de omgang met wetenschappelijke bevindingen geldt als uitgangspunt dat die bevindingen geen vaststaande feiten opleveren maar inzichten en opvattingen waarbij altijd tegenspraak en weerwoord mogelijk en zelfs nodig zijn.
3. Wetenschap kan naar haar aard niet op alle vragen een antwoord geven; die mogelijkheid wordt nadrukkelijk opengehouden.
4. De overheid mag niet een bepaalde wetenschappelijke overtuiging tot waarheid verheffen en tot leidend verklaren.
5. De overheid stelt zich open voor kritiek op, en publiceert de inzichten waarop haar beleid berust.
6. Modeluitkomsten hebben nooit het laatste woord.
7. Gebruikte modellen voor overheidsbeleid staan aantoonbaar in relatie tot de werkelijkheid.
8. De berekeningen en aannames in modellen zijn openbaar te verifiëren en/of te controleren.

9. De modeltoelichting benoemt omvang en aard van onzekerheden, en vermeldt de factoren en omstandigheden die niet in het model zijn meegenomen.
10. Het materiaal waarmee modellen werken, wordt steeds getoetst op bruikbaarheid en juistheid. Modellen worden waar nodig geactualiseerd en aangepast, met behoud van rechtszekerheid.

Tenslotte schets ik een concreet stappenplan voor een nieuwe manier om het natuurbeleid in Nederland vorm te geven. Eerst benoem ik de hoofdpunten die ik daarna kort uitwerk.

1. Breng in kaart welke richtlijnen, handreikingen en adviezen de Europese Commissie geeft/heeft, als alternatief voor Aerius/OPS, de KDW's, de drempelwaardes en de vergunningverlening.
2. Onderzoek hoe e.e.a. in het buitenland is vormgegeven en onderzoek of buitenlandse benaderingen in Nederland ook haalbaar en/of wenselijk zijn.
3. De samenhang tussen Natuurdoelanalyse, Advies van de Ecologische Autoriteit en de vergunningverlening moet bezien worden. Een natuurrapportage is van meerwaarde als deze gekoppeld kan worden aan de vergunningverlening.
4. Stel een multidisciplinaire werkgroep samen die adviseert over de juridische haalbaarheid, de uitwerking en de uitvoering van de hiervoor benoemde punten.

Ad 1.

De Europese Commissie baseert zich in haar documenten op de staat van de natuur, en daarbij in hoge mate op veldgegevens. Voor 2015 werd ook in Nederland op die manier naar het dossier gekeken; er zijn verschillende documenten die daarvan getuigen. De 'Leeswijzer Natura 2000 profielen' van september 2014 is daar een voorbeeld van.⁸³ Ik geef ter illustratie enkele citaten uit die leeswijzer.

⁸³https://www.natura2000.nl/sites/default/files/profielen/Habitattypen_profielen_algemene_documenten/Leeswijzer%20Natura2000%20profielendoc%202014.pdf

“Natura 2000 aanwijzingsbesluiten stellen instandhoudingdoelstellingen vast voor onder meer de kwaliteit van habitattypen in een Natura 2000 gebied. De kwaliteit van habitattypen wordt bepaald door vier aspecten (zie Natura 2000 profielendocument), te weten:

- Definiërende vegetatietypen
- Typische soorten
- Abiotische randvoorwaarden
- Overige Kenmerken van goede structuur en functie

Deze notitie gaat in op één van de vier kwaliteitsaspecten, namelijk de typische soorten. Voor juridisch houdbare beheerplannen en vergunningen is het van belang om voldoende aandacht te geven aan de kwaliteitsaspecten van de instandhoudingsdoelstelling.”

[...]

“Naast effecten via de abiotiek en verstoring kunnen echter ook versnippering en habitatdestructie het vóórkomen van typische soorten beïnvloeden.

- Bij versnippering gaat het erom dat het leefgebied van typische soorten zodanig wordt opgedeeld dat een habitatype zijn functionele omvang verliest en de typische soorten eruit verdwijnen. Er is dan ook een effect op de structuur en functie (functionele samenhang) van een habitatype. De toetst aan structuur en functiekenmerken van een habitatype volstaat dus.
- Bij habitatdestructie verdwijnen typische soorten bijvoorbeeld als er vegetaties verdwijnen, of de structuren en functie van het habitatype (waar de typische soorten van afhankelijk zijn) negatief worden beïnvloed. Ook voor deze potentiële invloeden zal in de praktijk dus een toets aan de andere kwaliteitsaspecten (structuur en functie of vegetaties) veelal volstaan.”

Deze citaten verraden een veel completere manier om naar het beheer van Natura 2000 gebieden te kijken dan nu het geval is. In lijn met die gedachte moeten we ervoor

waken om niet opnieuw aan de wet van Goodhart ten prooi te vallen.⁸⁴ Omdat ons natuurbeleid uiteindelijk gericht is op biodiversiteit kan dat zomaar leiden tot een nieuwe tunnelvisie waarbij het beleid gereduceerd wordt tot een getal dat gehaald moet worden. Het document gaat hier ook op in:

“Verschuivingen in typische soorten met lokaal verschijnen en verdwijnen binnen een gebied, maar ook tussen gebieden, zijn een natuurlijk fenomeen.”

Met andere woorden, we moeten geen doelstellingen hanteren die tegen de aard van ‘de natuur’ ingaan. Natuur is nu eenmaal dynamisch, en binnen bepaalde marges zou ons dat niet moeten of hoeven te verontrusten. In geen geval kan natuurbeleid gereduceerd worden tot een aantal getalsmatige doelen; het vereist het vakmanschap, wijsheid, overzicht en kennis.

Wat betreft de vergunningverlening geeft De Europese Commissie een richtsnoer voor een “passende beoordeling” voor vergunningaanvragen.⁸⁵ Het volgende citaat is representatief:

“Bij de passende beoordeling van plannen of projecten die gevolgen kunnen hebben voor Natura 2000-gebieden, moeten gegarandeerd alle elementen die bijdragen aan de natuurlijke kenmerken van het gebied volledig in aanmerking worden genomen, zowel bij het bepalen van de referentietoestand als in de fasen die leiden tot de vaststelling van mogelijke gevolgen, verzachtende maatregelen en resteffecten.

In dit verband moet ervoor worden gezorgd dat in de passende beoordeling alle elementen aan de orde komen die bijdragen aan de natuurlijke kenmerken van het gebied zoals vermeld in de instandhoudingsdoelstellingen en het standaardgegevensformulier van het gebied, en dat de beoordeling is gebaseerd op de beste beschikbare wetenschappelijke kennis.

⁸⁴ Zie Hoofdstuk 1.

⁸⁵ Deze “passende beoordeling” wijkt sterk af van de “passende beoordeling” zoals de Nederlandse provincies die hanteren. In die laatste hebben berekeningen over stikstof een prominente plaats gekregen, maar dat is op Europees niveau niet het geval. Zie bijvoorbeeld <https://www.bij12.nl/onderwerp/stikstof/passende-beoordeling/>; <https://www.noord-holland.nl/bestanden/pdf/Stappenplan%20vergunning%20stikstof.pdf>; <https://www.brabant.nl/@10772/handreiking-passende-beoordeling/>.

De vereiste gegevens moeten actueel zijn en kunnen waar van toepassing de volgende kwesties omvatten:

- structuur, functie en rol van de natuurwaarden van het gebied;
- de oppervlakte, representativiteit en mate van instandhouding van de habitattypen in het gebied;
- populatieomvang, mate van isolatie, ecotype, genenpool, leeftijdsstructuur en mate van instandhouding van de in het gebied aanwezige soorten die worden vermeld in bijlage II bij de habitatrichtlijn of van de vogelsoort waarvoor een bepaalde SBZ voor vogels is aangewezen;
- andere in het gebied vastgestelde natuurwaarden en ecologische functies; en
- bedreigingen die van invloed zijn op of een mogelijk risico vormen voor in het gebied aanwezige habitats en soorten.

In de passende beoordeling moet melding worden gemaakt van alle mogelijke gevolgen van het plan of project die significant kunnen zijn voor het gebied, met inbegrip van cumulatieve en andere gevolgen die kunnen ontstaan door de combinatie van het plan of project dat wordt beoordeeld, met andere plannen of projecten.”

Deze richtlijn geeft dus zeker wel enige houvast voor een ander beleid in Nederland. In feite keren we tot op zekere hoogte terug naar het natuurbeleid waar een aanzet voor werd gemaakt voor de introductie van de KDW's en Aeries. We erkennen dan dat het beleid aan de hand van onwetenschappelijke KDW-waardes en niet-realistische modellen een vergissing was.

Ad 2.

In mijn rapport heb ik al kort verwezen naar het buitenland. Frankrijk en (Noord-)Italië maken gebruik van een ecologische, kwalitatieve analyse als passende beoordeling, in plaats van stikstofmodellen. Hun aanpak is hier wellicht haalbaar, aangepast naar Nederlandse omstandigheden. Er is geen reden om het wiel opnieuw uit te vinden, en een uitgebreidere studie lijkt mij zinvol en zelfs nodig. Politici dienen het 'hele plaatje' te zien en te begrijpen. Het feit dat Nederland (en Vlaanderen, dat de Nederlandse werkwijze in belangrijke mate

kopieerde) zich als enige land heeft overgeleverd aan modeluitkomsten verraadt een beperkt zicht op de realiteit. Het lijkt mij dus cruciaal om te onderzoeken hoe andere landen het probleem aanpakken; kennelijk zijn er veel meer opties om binnen de Europese richtlijnen te opereren – zie ook Ad 1.

Ad 3.

Het dossier heeft veel facetten. In dit rapport heb ik de wetenschappelijke kant bestudeerd, maar alles hangt met alles samen. Een goede samenwerking tussen de verschillende onderdelen is dus van groot belang, en uiteraard dienen natuuranalyse en vergunningverlening op elkaar afgestemd te zijn. Wil de EA met gezag kunnen blijven adviseren, dan dienen haar rapportages de kennis en filosofie van de EA zelf te reflecteren.

Ad 4.

Deze werkgroep dient te bestaan uit wetenschappers en andere deskundigen die zich a priori committeren aan een advies zonder KDW's en rekenmodellen. Leden van deze werkgroep kunnen ook nadrukkelijk internationaal aangezocht worden, maar het voorzitterschap zou in Nederlandse handen moeten liggen. Het verdient de voorkeur om de leden niet te werven uit het groepje wetenschappers dat de afgelopen decennia nauw bij het stikstofbeleid betrokken was; er is nieuw elan en nieuwe creativiteit nodig, zonder mogelijke ballast uit het verleden. Aangezien de problematiek multidisciplinair is, dient de werkgroep dat ook te zijn. Leden van de werkgroep handelen op persoonlijke titel. De adviesvraag aan deze werkgroep zou zowel inhoudelijk als procesmatig moeten zijn: waar gaan we ons op concentreren, hoe dan we dat inhoudelijk, en hoe richten we de processen in?

Tot slot: ik realiseer mij terdege dat dit alles een grote omslag in ons denken over het stikstofdossier zou betekenen. Ik ben er echter van overtuigd dat een dergelijke aanpak meer recht doet aan de feitelijke problematiek dan de fictieve modellenwerkelijkheid waar we ons nu in hebben begeven.

Appendix A

Technische details van het zoekproces in de adviezen van de EA

Voor elk van de 133 EA-advies PDF-documenten werd de volledige tekstinhoud geëxtraheerd met behulp van de Python-bibliotheek PyPDF2. Een Python-script (`extract_pdf_text.py`) automatiseerde dit proces. Het script opende elk PDF-bestand, itereerde door alle pagina's, extraheerde de tekst per pagina via de `extract_text()` methode, en voegde de teksten samen tot één doorlopende tekst per document. Om de oorspronkelijke structuur enigszins te behouden en de leesbaarheid voor de latere LLM-analyse te verbeteren, werden dubbele newline-karakters ingevoegd tussen de tekst van opeenvolgende pagina's. De geëxtraheerde platte tekst van elk document werd opgeslagen in een afzonderlijk UTF-8 gecodeerd tekstbestand.

De kern van de analyse werd uitgevoerd met behulp van gpt-4.5 en gpt-4o via de OpenAI API om relevante citaten over AERIUS-gerelateerde onzekerheid te identificeren.

Prompt Engineering

Een cruciale stap was het ontwerpen van een effectieve prompt om het LLM te instrueren met een zoekopdracht. De prompt is opgebouwd uit de instructie en de tekst van de EA adviezen. Hieronder staat de instructie beschreven:

“Analyseer deze PDF en geef aan of, waar en hoe statistische onzekerheid wordt besproken in relatie tot het AERIUS-model of de uitkomsten van dit model. Geef specifieke passages met paginanummers pdf en ruime context. Geef per citaat ook aan hoe dit zich verhoudt tot AERIUS model en statistische onzekerheid. Geef ook aan of dit impliciet of expliciet beschreven staat. Wees kritisch of er echt een match is, ik ben met name geïnteresseerd of gereflecteerd wordt op statistische onzekerheid. Als er geen directe relatie met AERIUS bestaat moet je deze weglaten. Geef concrete citaten en vat ze kort samen in gewone taal. Geef per gevonden passage in deze structuur:

- Paginanummer/paragraaf;
- Citaat uit de tekst;
- Samenvatting in gewone taal;
- Is het expliciete of impliciete onzekerheid?

Sluit af met een tabeloverzicht met paginanummer, onderwerp, statistische onzekerheid en de explicietheid duidelijk maakt.”

Deze basis prompt werd aangevuld met de volledige geëxtraheerde tekst van één PDF document van een EA advies.

Uitvoering

Een Python-script orkestreerde het analyseproces. Het script itereerde door de geëxtraheerde .txt-bestanden, las de inhoud, formatteerde de gebruikersprompt, en stuurde een request naar de OpenAI API⁸⁶ (chat.completions.create endpoint). De model temperatuurparameter van de API werd ingesteld op een lage waarde (0.2) om de consistentie en reproduceerbaarheid van de output te bevorderen. De analyse maakte gebruik van gpt-4.5-preview (voor 78 documenten) en gpt-4o (voor 55 documenten). Een vergelijkende analyse toonde aan dat gpt-4.5-preview meer citaten identificeerde (gemiddeld 3,62 per document; 99% van documenten met minsten 1 citaat) dan gpt-4o (gemiddeld 2,09 per document; 84% van documenten met minstens 1 citaat). In een directe vergelijking van 24 documenten vond gpt-4.5-preview in 16 gevallen meer citaten. De door het LLM gegenereerde Markdown-tekst werd voor elk geanalyseerd document opgeslagen als een .md-bestand. De bestanden zijn ook geëxporteerd naar een PDF voor beoordeling.

Om een totaaloverzicht te creëren van gevonden citaten werd er gebruik gemaakt van reguliere expressies (regex), om de beoogde data te isoleren. In totaal werden 490 citaten geëxtraheerd uit de bestanden; zie Appendix B.

Elk EA analyse werd individueel beoordeeld en de relevantie van de citaten afgewogen. Waar nodig werd de volledige EA analyse geraadpleegd, en ik enkele gevallen ook de bijbehorende NDA's.

⁸⁶ chat.completions.create endpoint

Software, Bibliotheken en Omgeving

- AI Tools: Elicit, OpenAI GPT
- Programmeertaal: Python (versie 3.11)
- Core Python Bibliotheken: os, re, glob, time
- PDF Extractie: PyPDF2
- API Interactie: openai
- Data Manipulatie & Excel Export: pandas, openpyxl
- Markdown naar PDF Conversie: markdown, weasyprint
- Archivering: zip (standaard Linux command-line utility)
- Overig: tabulate (voor output formatting in vergelijkingsscript)

Verantwoording en dankwoord

Mijn dank gaat uit naar Roland Bobbink, William M. Briggs, Rob Brussee, Jan Willem Erisman, Thomas van Goethem, Jaap Hanekamp, Yannic Hanekamp, Ronald Hoogerbrugge, Marc Jacobs, Hans Mommaas, Quinten Pluymaeckers, Geesje Rotgers, Willemijn Smal en Gerard Velthof voor hun bereidheid om hun kennis, ideeën en visie met mij te delen.

Vanzelfsprekend ben ik, en ik alleen, verantwoordelijk voor de inhoud van dit rapport.