



Wettelijke Onderzoekstaken Natuur & Milieu

Complexiteit van WUR-modellen en -bestanden

Toetsing van de EMC v1.0

| WOt-werkdocument 339

G.A.K. van Voorn, P.W. Bogaart, M. Knotters & D.J.J. Walvoort



WAGENINGEN UR
For quality of life

Complexiteit van WUR-modellen en -bestanden – Toetsing van de EMC v1.0

De reeks 'Werkdocumenten' bevat tussenresultaten van het onderzoek van de uitvoerende instellingen voor de unit Wettelijke Onderzoekstaken Natuur & Milieu (WOT Natuur & Milieu). De reeks is een intern communicatiemedium en wordt niet buiten de context van de WOT Natuur & Milieu verspreid. De inhoud van dit document is vooral bedoeld als referentiemateriaal voor collega-onderzoekers die onderzoek uitvoeren in opdracht van de WOT Natuur & Milieu. Zodra eindresultaten zijn bereikt, worden deze ook buiten deze reeks gepubliceerd.

Dit werkdocument is gemaakt conform het Kwaliteitshandboek van de WOT Natuur & Milieu.

Complexiteit van WUR- modellen en -bestanden

Toetsing van de EMC v1.0

G.A.K. van Voorn

P.W. Bogaart

M. Knotters

D.J.J. Walvoort

Werkdocument 339

Wettelijke Onderzoekstaken Natuur & Milieu

Wageningen, juli 2013

Referaat

Voorn, G.A.K. van, P.W. Bogaart, M. Knotters & D.J.J. Walvoort (2013). *Complexiteit van WUR-modellen en -bestanden. Toetsing van de EMC v1.0*. Wageningen, Wettelijke Onderzoekstaken Natuur & Milieu, WOt-werkdocument 339. 118 blz. 3 fig.; 10 tab.; 90 ref.; 2 bijl.

In dit werkdocument worden de resultaten behandeld van de toetsing van de evaluatielijst modelcomplexiteit (EMC v1.0) die in een eerdere fase is ontwikkeld. Deze lijst is bedoeld om de complexiteit te beoordelen van modellen en bestanden die de WOT Natuur & Milieu gebruikt om de effecten van natuur- en milieubeleid te evalueren. De toetsing heeft plaatsgevonden door de lijst toe te passen op een aantal casussen (bestaande modellen en bestanden), en daarbij te letten op zowel de inhoudelijke aspecten van de casus als de technische aspecten van de lijst. De resultaten hebben ertoe geleid dat een nieuwe (voorlopige) versie van de lijst (EMC v2.0) wordt voorgesteld in dit werkdocument.

Trefwoorden. modelcomplexiteit, evaluatielijst, natuurmodellen, modelleercyclus, modelkwaliteit

Auteurs

G.A.K. van Voorn: Biometris, PRI Wageningen UR

P.W. Bogaart, M. Knotters & D.J.J. Walvoort: Alterra Wageningen UR

©2013 **Biometris, PRI Wageningen UR**

Postbus 100, 6700 AC Wageningen

Tel: (0317) 48 07 98; fax: (0317) 48 35 54; e-mail: biometris@wur.nl

Alterra Wageningen UR

Postbus 47, 6700 AA Wageningen

Tel: (0317) 48 07 00; fax: (0317) 41 90 00; e-mail: info.alterra@wur.nl

De reeks WOt-werkdocumenten is een uitgave van de unit Wettelijke Onderzoekstaken Natuur & Milieu, onderdeel van Wageningen UR. Dit werkdocument is verkrijgbaar bij het secretariaat. **Het document is ook te downloaden via www.wageningenUR.nl/wotnatuurenmilieu**

Wettelijke Onderzoekstaken Natuur & Milieu, Postbus 47, 6700 AA Wageningen

Tel: (0317) 48 54 71; Fax: (0317) 41 90 00; e-mail: info.wnm@wur.nl; Internet: www.wageningenUR.nl/wotnatuurenmilieu

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever. De uitgever aanvaardt geen aansprakelijkheid voor eventuele schade voortvloeiend uit het gebruik van de resultaten van dit onderzoek of de toepassing van de adviezen.

Woord vooraf

Dit WOT-werkdocument verslaat de resultaten van de activiteiten binnen het project 'Evaluatie model-complexiteit' over de tweede helft van 2011 (WOT-04-011-016) en de eerste helft van 2012 (WOT-04-011-036.15). Deze activiteiten bestonden onder andere uit interviews met betrokkenen bij verschillende modellen en bestanden die bij Wageningen UR worden ontwikkeld en onder meer door het Planbureau voor de Leefomgeving (PBL) worden gebruikt.

Dank aan Paul Torfs, Claudia Brauer (beiden Wageningen University), Roel Velner (Royal Haskoning), Frank Weerts (Waterschap de Dommel), Ab Veldhuizen, Wieger Wamelink, Rogier Pouwels, Janet Mol (allen Alterra Wageningen UR), Arjen van Hinsberg, en Jaap Wiertz (beiden PBL) binnen die kaders. Dank aan Peter Janssen (PBL) en Harm Houweling (WOT Natuur & Milieu, Wageningen UR) voor kritische kanttekeningen bij de totstandkoming van het werkdocument.

Verdere publicaties die tot nog toe zijn verschenen over dit project zijn Bogaart *et al.* (2011), Van Voorn & Walvoort (2011), en Van Voorn *et al.* (2011), te vinden in de referenties binnen dit WOT-werkdocument. Casus bestanden: Martin Knotters; casus Wageningen-model: Patrick Bogaart; casus Meta-SWAP, NP, mNP: George van Voorn & Dennis Walvoort.

George van Voorn
Patrick Bogaart
Martin Knotters
Dennis Walvoort

Inhoud

Woord vooraf	5
Inhoud	7
Samenvatting	9
1 Inleiding	11
1.1 Probleemstelling	11
1.2 Projectdoelstelling	13
1.3 Kennisvraag en onderzoeksvragen	13
1.4 Beoogd projectresultaat	14
2 Bestandsdiscussies	15
2.1 Inleiding	15
2.2 Priapus	15
2.3 SIS+	17
2.4 Soil maps of the Netherlands 1:50,000 & 1:250,000	18
3 Wageningen-model	21
3.1 Inleiding	21
3.2 Basale gegevens Wageningen-model	21
3.3 Doel & toepassingen Wageningen-model	22
3.4 Systeemanalyse Wageningen-model	24
3.5 Conceptueel model Wageningen-model	27
3.6 Gegevens Wageningen-model	28
3.7 Formeel Wageningen-model	29
3.8 Numerieke implementatie Wageningen-model	32
3.9 Schematisering Wageningen-model	33
3.10 Gevoeligheden, onzekerheden, opbouw code	34
3.11 Kalibratie Wageningen-model	35
3.12 Validatie Wageningen-model	36
3.13 Samenvatting & evenwichtsbeoordeling Wageningen-model	37
4 MetaSWAP	39
4.1 Inleiding	39
4.2 Ingevulde vragenlijst MetaSWAP; basale gegevens	39
4.3 Doel & toepassingen MetaSWAP	40
4.4 Systeemanalyse MetaSWAP	42
4.5 Conceptueel model MetaSWAP	46
4.6 Gebruikte data en dataeisen bij MetaSWAP	48
4.7 Formeel model van MetaSWAP	49
4.8 Numerieke implementatie MetaSWAP	51
4.9 Schematisering MetaSWAP	53
4.10 Gevoeligheids- & onzekerheidsanalyse, opbouw code	54
4.11 Kalibratie MetaSWAP	56
4.12 Validatie MetaSWAP	57
4.13 Evenwichtsoordeel MetaSWAP, bevindingen	58
4.14 Verdere bevindingen	60

5	Natuurplanner	61
5.1	Inleiding	61
5.2	Ingevulde vragenlijst Natuurplanner; basale gegevens	61
5.3	Doel & toepassingen NP	63
5.4	Systeemanalyse NP	64
5.5	Conceptueel model NP	66
5.6	Data bij de NP	67
5.7	Formeel model NP	69
5.8	Numerieke implementatie NP	70
5.9	Schematisering NP	71
5.10	Gevoeligheden, onzekerheden, opbouw code NP	72
5.11	Kalibratie NP	73
5.12	Validatie NP	75
5.13	Evenwichtsoordeel Natuurplanner	76
6	Discussie	77
6.1	Inleiding	77
6.2	Punten ter verbetering van de lijst	77
6.3	Overige bevindingen	89
6.4	Aanzet tot aandachtspunten “evenwichtig modelleren”	89
6.5	Reflectie en discussiepunten	91
6.6	Vooruitzichten	93
	Literatuur	95
Bijlage 1	Evaluatielijst Modelcomplexiteit v1.0	101
Bijlage 2	Evaluatielijst Modelcomplexiteit v2.0	105

Samenvatting

Door diverse factoren, waaronder de steeds grotere beschikbaarheid van computerkracht, is de complexiteit van bestaande en nieuw ontwikkelde modellen en bestanden gemiddeld steeds verder toegenomen. De term 'complexiteit' duidt hierbij op een verzameling van aspecten die met elkaar verband houden, zoals de lengte van de computercode, bedienings- en gebruikersgemak, de tijd die benodigd is om een model te 'runnen', en het begrip dat men heeft van een model. Er bestaat een risico dat een model te complex is of wordt, waarbij zaken als testen en validatie steeds tijdrovender worden, en het algemene begrip van het model steeds meer gelimiteerd is. In dat geval daalt meestal ook het vertrouwen in de uitvoer van het model en de uitspraken die op die uitvoer gebaseerd zijn; het model is 'te complex' geworden.

Wanneer een model 'te complex' is hangt af van de context, en die wordt sterk bepaald door de toepassing (waar wordt het model voor gebruikt) en de beschikbaarheid van gegevens. Een model of bestand wordt in 'evenwicht' beschouwd wanneer de complexiteit van het model of bestand past bij de toepassing en de databeschikbaarheid, dat wil zeggen de uitvoer van het model of bestand geeft antwoorden met een gewenste betrouwbaarheid voor een bepaalde toepassing.

In dit WOT-werkdocument wordt gekeken naar de complexiteit van een aantal modellen en een modelketen die door het Planbureau voor de Leefomgeving (PBL) is of wordt gebruikt. Daarbij is gebruik gemaakt van een eerder ontwikkelde evaluatielijst voor modelcomplexiteit, de EMC v1.0. De resultaten van de uitgevoerde evaluaties voor modelcomplexiteit zijn vooral gebruikt om de functionaliteit van de evaluatielijst te toetsen, al leveren de resultaten ook nuttige informatie op wat betreft de complexiteit van de getoetste modellen. De resultaten zijn gebruikt om in bijlage 2 van deze rapportage de evaluatielijst op te waarderen van versie 1.0 naar versie 2.0. Daarnaast is een eerste versie met richtlijnen voor 'evenwichtig modelleren' bijgevoegd. Verder volgen er een aantal discussiepunten.

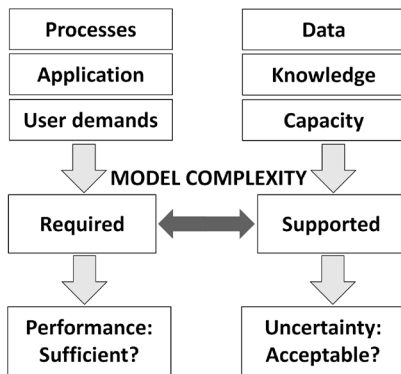
1 Inleiding

1.1 Probleemstelling

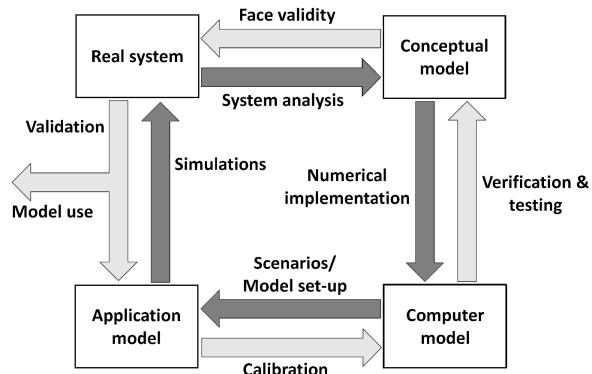
Bij verkenningen en evaluaties van het Nederlandse milieu- en natuurbeleid door het Planbureau voor de Leefomgeving (PBL) wordt voor een belangrijk deel gebruik gemaakt van modellen en bestanden die zijn ontwikkeld bij Wageningen UR. De verkenningen zijn in het algemeen gericht op de verschillen die kunnen optreden als gevolg van verschillende beleidsscenario's, terwijl de *ex ante* evaluaties gericht zijn op de vraag of beleidsdoelstellingen daadwerkelijk gerealiseerd zullen worden.

De meeste modellen en bestanden zijn vooral ontwikkeld vanuit het oogpunt van begrip van het systeem waarnaar gekeken wordt. Dit is een fundamenteel andere grondslag dan het gebruik van modellen en bestanden voor beleidsevaluaties en –verkenningen, die we hier als *toepassingsmodellen en –bestanden* aanduiden. De toepassing van deze modellen en bestanden stelt specifieke eisen aan deze modellen en bestanden, die niet *per se* gelden wanneer deze modellen en bestanden worden ontwikkeld voor het verkrijgen van een beter begrip van het systeem. Dit leidt vervolgens tot de vraag, of dat alle door PBL ingezette modellen en bestanden voldoen aan de door de toepassing gestelde eisen.

In de vorige fase van dit project “evaluatie modelcomplexiteit” is er een inventarisatie gemaakt van de eisen die voor toepassingsmodellen en –bestanden relevant kunnen zijn (Bogaart *et al*, 2011). Een aantal belangrijke aspecten zijn bv. het concept, de complexiteit van het model of bestand, het schaalniveau, de bruikbaarheid van het model of bestand, voldoende beschikbaarheid en voldoende kwaliteit van gegevens, etc. Daarbij is het concept ‘*evenwicht*’ geïntroduceerd. Dit concept is gebaseerd op de bekende trade-off tussen ‘bias’ en ‘variance’ – bij een beperkte beschikbaarheid van gegevens gaat het verminderen van systematische afwijkingen ten koste van de spreiding in modeluitvoer, en v.v. Echter, het concept ‘evenwicht’ is breder ingestoken.



Figuur 1.1



Figuur 1.2

Figuur 1.1 (links). Conceptueel schema van ‘evenwicht’, de trade-off tussen gewenste modelcomplexiteit en de beschikbare ondersteuning van de gewenste modelcomplexiteit; uitbreiding op eerder werk door Wagener *et al* (2001).

Figuur 1.2 (rechts). Conceptueel overzicht van de algemene modelleercyclus, vrij naar Rykiel (1996) en Refsgaard & Henriksen (2004).

Het concept 'evenwicht' is in principe een afweging tussen enerzijds de gewenste modelcomplexiteit, en anderzijds de beschikbare ondersteuning van deze gewenste complexiteit (zie Figuur 1.1; het principe is sterk ingegeven door eerder werk door Wagener *et al*, 2001). Bij het vaststellen van de gewenste complexiteit (de ene zijde van het 'evenwicht') wordt nadrukkelijk de toepassing van het model in ogenschouw genomen: deze bepaalt namelijk voor een belangrijk deel wat voor eisen er aan de modeluitvoer worden gesteld (o.a. keuze van variabelen, en betrouwbaarheidsmarges). Andere zaken die eisen stellen aan de gewenste modelcomplexiteit zijn o.m. gebruikerseisen (bv., het model moet voor trendanalyse worden gebruikt, of het moet flexibel genoeg zijn om met grote variaties in invoer om te gaan, etc.), en natuurlijk de correlaties die er tussen de verschillende processen in het gemodelleerde systeem bestaan. Bv., ecosystemen zijn over het algemeen inherent complex, en vereisen daarmee een afdoende complex model om het systeem reëel te beschrijven. Bij het vaststellen van de beschikbare steun voor de gewenste modelcomplexiteit (de andere kant van het 'evenwicht') wordt vooral gekeken naar beschikbare gegevens, maar ook naar expertkennis, en inzetbare capaciteit (denk aan personeel, computerkracht, etc.). We beschouwen vervolgens een model of bestand in 'evenwicht' wanneer het

- Complex genoeg is om tegemoet te komen aan uit de toepassing voorkomende eisen qua schaal, nauwkeurigheid, flexibiliteit en geldigheid binnen extern opgelegde randvoorwaarden;
- En tegelijkertijd eenvoudig genoeg om de onzekerheden terug te dringen tot aanvaardbare proporties door een adequate ondersteuning door data en expert-kennis.

Tevens is er geconstateerd dat er in de praktijk diverse factoren bestaan, die er gemakkelijk toe leiden dat er sterk van het evenwicht van een model of bestand afgeweken wordt (Van Voorn & Walvoort, 2011; Van Voorn *et al*, 2011).

Het concept 'evenwicht' ondervangt een aantal nadelen die kleven aan de traditionele, geautomatiseerde criteria voor modelselectie, zoals het Akaike criterium (AIC; Akaike, 1974; Van Voorn *et al*, 2011). Deze criteria zijn gericht op het verkrijgen van een optimaal model d.m.v. het vergelijken van de 'fit' van verschillende kandidaat-modellen op een specifieke dataset. Het 'optimale' model past daarbij enerzijds voldoende op de dataset, zonder dat het ook de flexibiliteit verliest om op andere, vergelijkbare datasets te passen, en anderzijds is het aantal vrijheidsgraden geminimaliseerd. Een technisch nadeel van deze criteria is dat er impliciet wordt aangenomen dat de vrijheidsgraden onafhankelijk van elkaar zijn, wat meestal niet zo is. Maar er zijn ook veel grotere nadelen:

- De criteria zijn niet toepasbaar voor databestanden of graadmeters;
- Er wordt niet expliciet gekeken naar bronnen van onzekerheid, en naar welke processen nadrukkelijk niet in het model zijn opgenomen;
- Deze criteria geven niet altijd een expliciet oordeel t.a.v. de toepassing(en) van het model of bestand, die bepaalt welke variabelen belangrijk zijn en op welke aspecten het model nadrukkelijk betrouwbaar moet zijn.

De bedoeling van het concept 'evenwicht' is er op gericht om bovenstaande nadelen te ondervangen voor de beoordeling of een model of bestand geschikt is voor een specifieke toepassing of niet.

Als hulp bij de beoordeling of een toepassingsmodel of -bestand in evenwicht is met bijbehorende toepassingen en beschikbare data, is in een volgende fase van het onderzoek een instrument ontwikkeld en getest. Dit instrument is de prototype lijst Evaluatielijst ModelComplexiteit (EMC v1.0; Van Voorn & Walvoort, 2011; Van Voorn *et al*, 2011). In deze lijst zijn criteria opgenomen, die erop zijn gericht om te beoordelen af aan de verschillende aspecten van 'evenwicht' is voldaan, m.a.w., of de complexiteit van het model of bestand in balans is in het licht van het gebruik ervan in beleidsevaluaties en -verkenningen. De opzet van deze lijst is gebaseerd op de algemene modelleercyclus (zie Figuur 1.2). De reden hiervoor is, dat in verschillende fasen van de modelleercyclus andere aspecten van de trade-off tussen gewenste en gesteunde modelcomplexiteit

een belangrijke rol spelen. Bijvoorbeeld in de systeemanalyse (omzetting van reëel systeem naar conceptueel model) spelen de eisen vanuit de toepassing en de beschikbare expertkennis een grote rol. Bij de implementatie van het conceptuele naar computermodel zijn zaken als rekenkracht en numerieke aspecten (zoals een afweging tussen snelheid en accuraatheid van subroutines) van belang. En bij kalibratie en validatie zijn bv. databeschikbaarheid en onzekerheden in invoer belangrijk.

Door de opzet van de lijst is het ook mogelijk om bij een analyse met de EMC v1.0 indirect aanwijzingen te verkrijgen op welke punten een model of bestand valt te verbeteren voor gebruik binnen de toepassing. De lijst levert daarmee ook een bijdrage aan de algemene verbetering of versterking van het PBL-modelleninstrumentarium.

1.2 Projectdoelstelling

In de laatste fase van 2011 is er een begin gemaakt met de toepassing van de ontwikkelde lijst EMC v1.0 op modellen en bestanden binnen het PBL-instrumentarium. Dit projectvervolg is onder meer bedoeld om de begonnen analyses af te maken, en hier lessen uit te trekken over nut en bruikbaarheid van de EMC-lijst, en verbeterpunten aan te geven voor deze lijst. Daarnaast is het de bedoeling om uit deze analyses leerpunten te halen voor verbetering of versterking van het PBL-instrumentarium om – naast casus-specifieke aanbevelingen – ook tot algemene richtlijnen te komen voor hoe een model of bestand ‘evenwichtig’ kan worden ontwikkeld voor een bepaalde toepassing binnen beleidsevaluaties en –verkenningen, eveneens met het oog op verbetering of versterking van het PBL-instrumentarium.

1.3 Kennisvraag en onderzoeksvragen

De centrale vraag gedurende dit project is:

- Zijn de modellen en bestanden die binnen het PBL-instrumenten vallen en gebruikt worden voor beleidsevaluaties en –verkenningen voldoende geschikt voor deze toepassingen?

Voor het beantwoorden van deze vraag is de EMC-lijst ontwikkeld, die in deze rapportage verder getest wordt door toepassing ervan op een aantal casussen.

Onderzoeksvragen die gelden bij de beantwoording van deze centrale vraag zijn:

- Welke eisen stelt de toepassing aan modellen en bestanden?
- Overlappen deze eisen met eisen aan modellen en bestanden gericht op begrip van het systeem, en zo nee, op welke additionele punten moet de aandacht gericht worden om een model of bestand geschikt te maken voor de toepassing?
- Hoe kunnen de modellen en bestanden geëvalueerd worden met het oog op geschiktheid voor de toepassing?
- Zijn de bestaande modellen en bestanden te eenvoudig of te complex voor gebruik binnen de toepassing?
- Zijn er algemene richtlijnen te geven voor de ontwikkeling van modellen en bestanden die voldoen aan alle eisen die gesteld worden door de toepassing, en daarmee voldoende geschikt zijn voor de toepassing?
- Is het mogelijk een klein en snel instrument te ontwikkelen dat gebruikt kan worden voor een snelle evaluatie van het gehele PBL-instrumentarium, waarin in elk geval de belangrijkste knelpunten voor een goede toepassing naar boven komen?

Een aantal van de bovenstaande onderzoeksvragen zijn al in meer of mindere mate beantwoord in vorige fasen van het project. De ontwikkelde lijst EMC v1.0 is echter op dit moment nog niet goed getest voor bestanden en modelketens (aaneenschakelingen van modellen en bestanden). Dit is wel relevant, omdat het PBL-instrumentarium ook veel bestanden bevat, en de bestanden en modellen bovendien veelal in modelketens zitten. De lijst zal worden toegepast op een aantal geselecteerde casussen voor deze testen. Als modelketen is de NatuurPlanner (Van der Hoek & Bakkenes, 2007) geselecteerd, omdat deze een grote rol binnen het beleidsonderzoek van PBL heeft.

1.4 Beoogd projectresultaat

Er worden analyses gemaakt van een aantal testcasussen (bestanden) en van relevante modellen en bestanden binnen het PBL-instrumentarium (o.a. MetaSWAP en de NatuurPlanner). Deze analyses zullen als basis dienen voor:

1. Verdere aanpassingen/verbeteringen van de evaluatielijst EMC v1.0 (Van Voorn *et al*, 2011), die leiden tot een nieuw versienummer van de EMC-lijst;
2. Casus-specifieke aanbevelingen voor eventuele verbetering van het model of bestand.

De resultaten van deze activiteiten zijn gepubliceerd in dit WOT-werkdocument, uitgezonderd de casus-specifieke aanbevelingen voor verbetering van de Natuurplanner en MetaNatuurplanner; die zullen in een toekomstige rapportage opgenomen worden.

Verder worden er in hoofdstuk 6 nog een aantal discussiepunten genoemd, en wordt aangegeven welke vooruitzichten er verder zijn voor dit project.

2 Bestandscasussen

2.1 Inleiding

Hoewel Van Voorn & Walvoort (2011) en Van Voorn *et al* (2011) vermelden dat het de bedoeling is dat de EMC-lijst ook geschikt is voor de evenwichtsevaluatie van bestanden, is dat in die rapportages nooit getoetst. In dit hoofdstuk worden daarom 3 casussen met bestanden doorgelopen, namelijk Priapus, SIS+, en de bodemkaarten van Nederland 1:50.000 en 1:250:000. Een deel van de informatie in dit hoofdstuk is overigens ook relevant voor de casus 'MetaSWAP' (zie hoofdstuk 4), omdat MetaSWAP gebruik maakt van bodemgegevens.

Noot: Dit verder Engelstalige hoofdstuk is opgesteld in de stijl van Van Voorn & Walvoort (2011).

2.2 Priapus

In this evaluation we consider Priapus, version 1.1, released 21-04-2008, containing data on soil hydraulic properties. For a description of Priapus we refer to Stolte *et al* (2007) and Verzandvoort *et al* (2012).

Goal/Application area

According to Stolte *et al* (2007) Priapus contains soil hydraulic characteristics, measurement data and meta data of soil samples collected and analysed by the former research institutes ICW, Stiboka and DLO-Staring Centrum and the current research institute Alterra (part of Wageningen University and Research centre). Parts of the data in Priapus form also part of one or more editions of the Staring series (Wösten *et al*, 1987; Wösten *et al*, 1994; Wösten *et al*, 2001). Priapus is not aimed to replace the Staring series, but to extend the Staring series:

1. Besides soil hydraulic characteristics, Priapus contains the original measurement data and meta-information. This enables the user of Priapus to derive soil hydraulic characteristics of a soil sample, with respect to certain restrictions, and to select soil samples on the basis of meta-information such as clay content and spatial co-ordinates.
2. The meta-information provides an indication of the quality of measurement data and soil hydraulic characteristics.
3. The design of Priapus allows to add soil hydraulic characteristics, measurement data and meta-information of new soil samples taken in the Netherlands and abroad. This enables use of the database abroad in future.

Priapus is aimed to be applied in studies on transport of water and substances in the unsaturated zone. The soil hydraulic characteristics in Priapus can be applied in all simulation models that describe water transport in the unsaturated zone numerically. Application areas include:

- Management and monitoring of water table depth and soil and water quality;
- Studies on the impacts of changes in climate and land use;
- Studies on the effects of water repellence of soils;
- Studies on the water balances of catchments.

Besides these applications related to water transport in the unsaturated zone, Priapus is applied in studies on soil erosion and soil compaction.

System analysis, assumptions, design of database

The main attributes in Priapus are the relationship between hydraulic conductivity and pressure head and the relationship between soil moisture content $K(h)$ and pressure head $\theta(h)$ (see also the

MetaSWAP-casus in Chapter 4). The system boundaries are defined only globally: the Netherlands, in future the world. It can be assumed that the system is restricted to unpaved area, and road shoulders, banks and public gardens are excluded.

Uncertainty about the $K(h)$ and $\theta(h)$ relationships of soil has the following sources:

- Errors in sample treatment;
- Errors in laboratory measurements. These errors are quantified by repeated measurements on the same sample, the results of which are stored in Priapus;
- Errors arising from the schematization by the model of Mualem-Van Genuchten (Mualem, 1976; Van Genuchten, 1980). This source of uncertainty can be quantified by measures of goodness-of-fit, which have not yet been stored in Priapus;
- Errors in the spatial co-ordinates of a sample location;
- Errors in the registration of the time of sampling;
- Errors in spatial prediction and spatial aggregation of $K(h)$ and $\theta(h)$ relationships.

Priapus does not provide quantitative information on the uncertainty about the $K(h)$ and $\theta(h)$ relationships of soil samples. Verzandvoort *et al* (2012, Chapter 6) give recommendations to include quantitative information on sources of uncertainty in Priapus, in order to make Priapus suitable for application in risk analysis.

The design of Priapus is described in detail in Stolte *et al* (2007, Section 5.2). Priapus contains 28 tables, 10 of which storing information on the soil samples, and 18 with descriptive information.

Data requirements

Currently, Priapus contains data of 96 samples with quality status A, taken at 39 locations. Besides laboratory measurements on the $K(h)$ and $\theta(h)$ relationship, the following information on soil samples is required:

- Spatial co-ordinates;
- Time of sampling;
- Sample depth, soil horizon;
- Geologic formation;
- Texture (clay content, silt content, sand content);
- Organic matter content.

Laboratory measurements of texture and organic matter content are available for 65 samples, taken at 29 locations.

For applications at a field scale Priapus currently contains insufficient data. For regional or national studies the support of data is insufficient and aggregations are needed. For worldwide applications the current extent of data is insufficient. It should be noted that Priapus also contains information on samples which do not meet the constraints of quality status A (Priapus contains information of 852 samples). These data might nevertheless meet the requirements of specific applications.

In the Staring series part of Priapus the soil hydraulic characteristics are aggregated to 36 so called 'bouwstenen' (building blocks), representing 18 different soil physical units for the topsoil and the subsoil, respectively. Although a clear definition of topsoil and subsoil is lacking, the most common interpretation is that topsoil characteristics represent the effective root zone and subsoil characteristics represent the parent material below the root zone (observe, that this is used in the MetaSWAP-casus discussed in Chapter 4).

Mathematical and numerical model

The data in Priapus are partly resulting from model calculations: the $K(h)$ and $\theta(h)$ -relationships are described by the model of Mualem-Van Genuchten (Mualem, 1976; Van Genuchten, 1980) and by cubical splines.

Calibration, analysis, and validation

Verzandvoort *et al* (2012, Sections 4.2 and 4.3) describe the calibration of the Mualem-Van Genuchten model and cubical splines to laboratory measurements on $K(h)$ and $\theta(h)$. A problem is extrapolation of the relationships outside the range of, say, 10 to 1000 cm absolute pressure head. Outside this range a limited number of laboratory measurements is available. Values of $K(h)$ and $\theta(h)$ fitted by the Mualem-Van Genuchten model or by cubical splines might differ largely outside this range. Validation results are available for estimated pore volumes, see Stolte *et al* (2007, Subsection 7.3.1).

Discussion of the case and advices for improvement

Data in Priapus are aimed to be applied in studies inside and outside the Netherlands, at field, regional, national and global scale. Basically, in its design Priapus is well balanced for most applications because data can be provided at the level of individual samples, and can be aggregated, for instance, to the level of building blocks of the Staring series or other levels depending on the application. Currently, the 96 samples with quality status A poorly cover all classes of texture, organic matter content and geological formations. Besides this, the samples are spatially clustered. Furthermore, samples were taken long ago (in various cases more than 30 years ago) and possibly not representing actual soil physical conditions. Recently efforts have been made to fill in the gaps in Priapus (Knotters *et al*, 2011).

Information on the accuracy of data in Priapus is not related to the applications. There is a need for validation studies on the effects of errors in applications, and for uncertainty analyses on the propagation of errors in model chains and decision processes.

2.3 SIS+

This evaluation concerns a map of pH-KCl, constructed by SIS+, version 01.0, released 29 October 2010, on the basis of data in the Dutch Soil Information System.

Goal/Application area

The goal is to construct a map of pH-KCl for the soil layer from 0 to 0.25 m. The period of interest is from 1990-01-01 to 2010-12-31 and the spatial extent is given in ESRI shape file provinces. The resulting map should have a spatial resolution of 1000 m x 1000 m. In this stage of development of SIS+ the aim of constructing a map is a proof-of-concept, i.e., to test the process of automatically generating maps from data stored in the Dutch Soil Information System. In future these maps will be generated for applications in nationwide studies on natural resources.

System analysis, assumptions, model concept

Walvoort (2010) and Brus *et al* (2010) give a system analysis of SIS+. The most important attributes are the target variable $y(s_0)$ at observation locations s_0 , and $y(s_i)$ at prediction locations s_i . In this case y is pH-KCl of the soil layer from 0 to 0.25 m. The system boundaries are defined by the soil layer from 0 to 0.25 m depth, the Netherlands, and the period from 1990-01-01 to 2010-12-31. Attributes and processes outside these boundaries are neglected in constructing the map.

The values of the target variable, pH-KCl, at unvisited locations are predicted by simple kriging with varying local means (Goovaerts, 1997). The varying local means are average values of pH-KCl within

soil physical PAWN-units (Wösten *et al*, 1988). The applied model of simple kriging with varying local means has two spatial dimensions.

Uncertainty exists on the true value of pH-KCl at unvisited locations. The uncertainty is quantified by the variance of the prediction error or kriging variance. This variance is minimized under the constraint that the mean error equals zero.

Analysis of the spatial structure of pH-KCl is relevant, because a model of spatial variation, semivariogram, is needed in computing interpolation weights and in quantifying the variance of the prediction error.

Data requirements

Measurements of pH-KCl in the soil layer from 0 to 0.25 m are needed, and a schematisation of soil into a restricted number of spatial units for which local means are calculated.

Data on pH-KCl are available in the Dutch Soil Information System for 1695 soil profiles. A schematisation of the soil into 23 spatial units is available from the PAWN-study (Wösten *et al*, 1988).

The pH-KCl values are weighted average values from all soil horizons which contribute to the soil layer from 0 to 0.25 m. The lengths with which soil horizons contribute to the soil layers are used as weights.

The activity of hydrogen ions (H⁺) is taken prior to the weighted averaging, and back-transformed to the pH scale afterward (Brus *et al*, 2010, p. 10, 15).

Calibration, analysis, and validation

A semivariogram model describing the spatial variation of pH-KCl was fitted to an experimental variogram that was calculated from the set of 1695 observations, after subtracting the local means of the PAWN-units. The experimental and fitted variogram were plotted and the goodness-of-fit was evaluated visually. The accuracy of the interpolations was evaluated by leave-one-out cross-validation (Walvoort, 2010, p. 23-29).

Discussion of the case and advices for improvement

Application in real research problems, for instance in the field of environmental policy, was not the starting point in constructing the pH-KCl map. In the developing stage of SIS+ the pH-KCl map was produced as a 'proof-of-concept' only. In this stage a generic procedure for constructing maps from data in the Dutch Soil Information System is aimed for. Maps generated by this procedure should be suitable to a variety of possible future applications.

Sensitivity analysis or uncertainty analysis for assumptions made in modelling the spatial structure has not taken place, but would be interesting to evaluate the effects of choices made in variogram modelling, stratification and kriging. This can help to evaluate the complexity of the spatial model.

2.4 Soil maps of the Netherlands 1:50,000 & 1:250,000

This case concerns the soil map of the Netherlands, digitally available via the website of the Dutch Soil Information System (<http://www.bodemdata.nl>). National soil maps at two different scales are available: 1 : 50,000 and 1 : 250,000. The evaluation is partly based on the information provided by this website.

Goal/Application area

Soil properties can differ largely between soil units, for instance with respect to crop growth. Therefore it is important to know the spatial distribution of soil units, i.e., to make soil maps. Since a changing landscape often means a changing soil profile the relationship between landscape and soil is important in soil mapping. A soil map represents different soil units by polygons with different colours. The scale of the soil map and the soil units being distinguished depend on the purposes for which the soil is mapped.

The soil maps provided by the Dutch Soil Information System enable to quickly make overviews of soil units and their areas for any area of interest within the Netherlands. Results of interpretations or model calculations can be visualised quickly, and can easily be combined with other digitally available spatial information such as land use maps (De Vries *et al*, 2003, p. 11). Soil maps are particularly important to the national government, provincial governments and district water boards in policy making, scenario studies and plan development in the areas of:

- Soil use and soil protection;
- Environmental management;
- Water management;
- Nature conservation.

(De Vries *et al*, 2003, p. 11). De Vries *et al* (2008) list in detail the users and applications of the Dutch Soil Information System, including the national soil maps 1:50,000 and 1:250,000. For applications in risk and uncertainty analysis, for instance in model chains, information on the accuracy of spatial predictions of soil types is needed, see, for instance, Brus and Heuvelink (2007) and Knotters *et al* (2010).

Information on soil type and water table class are directly relevant in estimating yield losses, for instance caused by lowering of the water table in an area of groundwater withdrawal. The estimated yield losses are based on tables giving percentages of crop growth reduction for combinations of soil types and water table classes (Werkgroep HELP-tabel, 1987; Brouwer and Huinink, 2002; De Vos *et al*, 2008; Bouwmans, 1990).

In soil protection and environmental management information on soil physical properties and water table depth are important. Soil physical properties are derived from data on texture and organic matter content using pedo-transfer functions (Vereecken *et al*, 1993), or using soil physical units from the Staring series (Wösten *et al*, 2001), for example the soil physical schematisation of the soil map 1 : 250,000 in the PAWN study (Wösten *et al*, 1988).

NATLES (Runhaar *et al*, 2003) is a program for land evaluation focused on nature conservation, that needs information on soil type and water table class from the soil map 1 : 50,000.

System analysis, assumptions, model concept

The soil maps, 1:50,000 and 1:250,000, are models of the Dutch soil. The most important attributes are

- Soil units, with information on soil type and texture of the topsoil, and additional information on reworked soils, composition of the soil profile, cover layers of sand or clay, and texture of the subsoil.
- Units of water table classes reflecting the seasonal fluctuation of water table depths.

In horizontal directions the system boundaries are the national border, and the borders between land and water and between land and built-up or paved areas. In the vertical direction the system boundaries are the ground surface and the surface at a depth of 1.20 m below the ground surface.

The model does not include a stochastic component describing the uncertainty about the true values of the attributes. Brus and Heuvelink (2007) give quantitative information on the uncertainty about

soil types. The Netherlands' Soil Sampling Program (NSSP, Finke *et al*, 2001; Visschers *et al*, 2007) gives quantitative information on the contents of a number of soil units, including the percentage correctly classified (also referred to as map purity).

At augering locations in the field the soil type is determined on the basis of a soil profile description. At least one soil profile description is made for each square centimetre of the soil map. Interpolation is based on expert judgement, using knowledge about the relationship between soil and landscape, and information on cultural history and geological and pedogenetical processes. Areas having relatively homogeneous soil properties are aggregated to map units.

The temporal, and largely seasonal, fluctuations of the water table relative to the ground surface are described by water table classes. These classes represent the average depths between which the water table fluctuates, given the prevailing hydrologic and climate conditions. Water table classes are estimated at point locations on the basis of hydromorphic profile characteristics, landscape characteristics, calibration to observed time series of water table depths and well-timed observations of water table depths in auger holes.

Data requirements

Soil profile descriptions and field observations are required in the mapping process and in applications of the soil data. For the soil map 1:50,000 at least one soil profile description per 25 hectares is made, whereas for the soil map 1:250,000 at least one soil profile description per 6.25 km² is made. Soil profile descriptions are amply available, from national soil surveys and from regional surveys at larger scales (1:10,000, at least one soil profile description per hectare; 1:25,000, at least one soil profile description per 6.25 hectares).

Discussion of the case and advices for improvement

In many applications the balance between required and provided information is good. It should be mentioned, however, that the information provided by the soil map was often a starting point. Therefore, possible limitations cannot easily be detected. Furthermore, it is often not known whether the soil map is sufficiently accurate to be used for a specific purpose. Little is known about the accuracy of soil maps and about the effects of map errors in applications. A first analysis on these effects is given by Knotters *et al* (2009) for applications of the soil map in estimating yield losses in areas with groundwater withdrawal.

Many attributes of the soil map are important for applications. For applications in modelling transport of water and substances in the unsaturated information on the environment in which soil layers were sedimented is more important than information on soil type.

For applications in rural development projects, in assigning areas that are sensitive to nitrate and phosphate leaching and in estimating compensations for crop yield reductions soil information on field or farm scale is needed. For applications in nature conservation information might even be needed on soil patterns within fields. It is assumed that the national soil maps 1 : 50,000 and 1 : 250,000 cannot be applied at such detailed scales. Only if fields and farms have large areal sizes these national soil maps might be suitable.

In rural development projects and in estimating compensations for crop yield reductions in areas with groundwater withdrawal the plan horizon is 30 years. Soil maps should thus inform about the average conditions of soil and water table over a period of 30 years length, representing the prevailing hydrologic and climatic conditions.

Currently, the national soil maps 1 : 50,000 and 1 : 250,000 cannot be applied in risk and uncertainty analysis because information on the map accuracy is lacking at a global as well as at a local level. The first efforts to quantify accuracy of the national soil maps have been made in the program to update the Dutch Soil Information System (Brus and Heuvelink, 2007).

3 Wageningen-model

3.1 Inleiding

In dit hoofdstuk wordt de casus 'Wageningen-model' beschreven, een eenvoudig hydrologisch neerslag-afvoermodel. De primaire functie van het Wageningen-model is om voor een (klein) stroomgebied de afvoer te voorspellen op basis van hydro-meteorologische forcering (neerslag en potentiële verdamping) en de dynamiek van de waterbalans binnen het stroomgebied. Het ligt voor de hand om dit model te vergelijken met MetaSWAP, het model dat in het volgende hoofdstuk wordt besproken. De momenteel voor onderzoeksdoeleinden alsmede operationeel ingezette versies van het Wageningen-model zijn onderworpen aan een evaluatie met de versie 1.0 van de checklist. Gezien het doel van deze evaluatie worden geen conclusies getrokken ten aanzien van het Wageningen-model zelf, enkel ten aanzien van de checklist.

Voor deze casus zijn zowel onderzoekers verbonden aan Wageningen Universiteit, waar het model in de jaren 70 en 80 is ontwikkeld, geïnterviewd, alsmede meer operationeel werkende onderzoekers, verbonden aan een adviesbureau en een waterschap. In januari 2012 zijn enkele ontwikkelaars en gebruikers van het Wageningen-model geïnterviewd. Dit waren Paul Torfs en Claudia Brauer (aangeduid als T&B), beide WU/HWM op 25 januari 2012 en Roel Velner (Royal Haskoning, nu Waterschap Rivierenland) en Frank Weerts (Waterschap de Dommel) (aangeduid als V&W) op 27 januari 2012.

3.2 Basale gegevens Wageningen-model

1 Geef aan: model/bestand, naam, versienummer, revisienummer, uitgiftedatum. Onder model verstaan we de broncode + invoergegevens.

Het Wageningen-model is een waterbalansmodel, geen tijdreeksmodel. Feitelijk bestaat het Wageningen-model niet. Er zijn verschillende conceptuele versies:

- De oorspronkelijke versie op basis van het j-model van Kraijenhoff van der Leur (1958);
- Een latere vereenvoudiging (Van Walsum, 1981) waarin de snelle en langzame impuls-responsefuncties zijn vervangen door twee lineaire reservoirs (en die tevens als verre voorouder kan worden beschouwd van MetaSWAP, zie het volgende hoofdstuk).

Ook zijn er meerdere implementaties:

- Oudere implementaties zijn doorgaans in FORTRAN;
- Ten behoeve van het onderwijs aan het IHE is een spreadsheetversie in Excel ontwikkeld;
- Voor het huidig onderzoek met het Wageningen-model maakt men gebruik van een implementatie in de statistisch/numerieke programmeertaal R (<http://www.r-project.org>).

Er vindt geen formeel versiebeheer plaats. *De facto* is de laatste versie van de FORTRAN implementatie gecompileerd en toegepast door Royal Haskoning in opdracht van Waterschappen De Dommel en Aa en Maas.

Dé numerieke versie van het Wageningen-model bestaat evenmin. In de loop van jaren zijn er verschillende versies geïmplementeerd, waarvan de volgende vermeldenswaardig zijn:

- Een versie in Fortran, waarin afvoer volgens de oorspronkelijk opzet wordt gemodelleerd, d.w.z. met een convectie-diffusie vergelijking voor de snelle afvoercomponent, en het Kraijenhoff van der Leur j-model voor de langzame (grondwater) afvoercomponent. Deze versie wordt niet actief

beheerd. De versie is in projecten door Royal Haskoning doorontwikkeld en gecompileerd (in opdracht van Waterschap De Dommel). Van alle bestaande versies komt deze het meest in aanmerking als de referentieversie.

- Een spreadsheet versie (in excel), geprogrammeerd ten behoeve van het onderwijs aan het UNESCO-IHE Institute for Water Education
- Verschillende versies in R, die voornamelijk binnen de leerstoelgroep Hydrologie en Kwantitatief Waterbeheer van Wageningen Universiteit worden ontwikkeld voor onderzoeksdoeleinden.

Enkele aspecten die wat aandacht verdienen zijn o.a.:

- De waterbalans voor de wortelzone wordt iteratief opgelost. In een eerste stap wordt het vochtgehalte S geactualiseerd op basis van neerslag N en potentiële verdamping E_p , waarbij de percolatie P en capillaire opstijging C op 0 worden gesteld. Op basis van deze eerste schatting van S worden P , C , en E_p berekend, en vervolgens wordt een tweede schatting van S gemaakt. Afhankelijk van de implementatie vindt een derde iteratieslag plaats.
- De impuls-responsefuncties voor $Q_{f,1}(t)$ en $Q_{s,1}(t)$ die in de Fortran-versie worden gebruikt (zie beschrijving conceptueel model) hebben oneindig lange staarten. In de praktijk worden de resulterende hydrografen echter na een bepaalde lengte (aantal tijdstappen) afgekapt. De resulterende waterbalansfout wordt niet altijd goed afgehandeld, wat eventueel tot een opbouw van afwijkingen kan leiden.

Evaluatie van de checklist:

Deze vraag, hoewel nuttig ten behoeve van dossiervorming (traceerbaarheid en transparantie), levert geen wezenlijke bijdrage aan de evenwichtsevaluatie.

3.3 Doel & toepassingen Wageningen-model

2A Wat was het doel van het bestand/model?

T&B: Het primaire doel van het Wageningen-model is het modelleren van het neerslag-afvoerproces, en de statistieken daarvan zoals maatgevende afvoer. Secundaire (achterliggende) doelen zijn in de loop van de tijd veranderd. Oorspronkelijk was dit de dimensionering van het watersysteem (slootdimensies en -afstanden, en dergelijke). Kwantificering met behulp van het Wageningen-model was een vooruitgang ten opzichte van de eerder gevolgde benadering op basis van vuistregels (bijvoorbeeld dat de maatgevende runoff-productie 14 l/s/ha bedraagt). Hedendaagse doelen zijn het berekenen van runoff binnen operationele openwatermodellen zoals SOBEK, en het bijdragen aan beleidsvragen, bijvoorbeeld met betrekking tot drainage, waar het Wageningen-model gezien de dynamische verdeling tussen snelle en langzame afvoer bij uitstek geschikt voor is.

V&W: Het Wageningen-model is vooral een hulpmiddel om waterbergingsgebieden in te richten, het gaat erom om betrouwbare uitspraken te verkrijgen hoe zo'n gebied functioneel ingezet kan worden. Een secundair doel is om (*ex ante*) hoogwatergolven te kunnen analyseren ten behoeve van het Model Regionaal Waterberging (MRW) in de context van het Nationaal Bestuursakkoord Water (NBW).

Kritiek:

Er is onduidelijkheid over de interpretatie van 'doel'. Het einddoel (bv. gebieds-inrichting) is iets anders dan het modeldoel (neerslag-afvoer modellering)

2B Wat was het beoogde toepassingsgebied, en waaruit blijkt dit?

T&B: Het toepassingsgebied van het Wageningen-model is het waterbeheer in vrij afwaterende laaglandstroomgebieden. Het model is niet geschikt voor sterk geaccidenteerd terrein (topografisch aangedreven stroming wordt niet beschreven door het model). Ook is het niet geschikt voor open

water (de hydraulica van het open water systeem wordt niet beschreven door het model)of voor poldergebieden waar water wordt ingelaten (inlaatwater kan niet in het model worden gediscoteerd).
V&W: Het model wordt binnen het studiegebied vooral ingezet op historische tijdreeksen (dus niet operationeel).

2C Wat zijn de toepassingen nu? Heeft het model/bestand een rol in een modelketen? Zo ja, welke?

T&B: De feitelijke toepassingen van het Wageningen-model vallen geografisch gezien binnen het beoogde toepassingsgebied: de vrij afwaterende kleine stroomgebieden in Pleistoceen Nederland en vergelijkbare gebieden net over de grens (met name in Duitsland). Procesmatig gezien wordt het Wageningen-model vooral ingezet voor vragen rondom wateroverlast (te veel water) en nauwelijks voor vragen over verdroging (watertekort). Het Wageningen-model wordt vooral ingezet voor het doorrekenen van lange tijdreeksen, en nauwelijks voor individuele gebeurtenissen.

Qua modelketens wordt het Wageningen-model in de praktijk vooral gekoppeld aan het SOBEK model, wat de standaard is in het regionale operationele waterbeheer. In de door Royal Haskoning ontwikkelde schil TRIWACO is het Wageningen-model één van de opties voor neerslag-afvoer. Als dusdanig fungeert het ook als second opinion ten opzicht van de andere neerslag-afvoer opties en draagt zodoende bij aan onzekerheidsanalyses binnen SOBEK.

V&W: Zij passen het Wageningen-model op regionale schaal met name toe vanuit het kader van MRW/NBW toetsing. Lokaal wordt het gebruikt ten behoeve van de inrichting van waterbeheersgebieden waar problemen zijn. Het Wageningen-model is hier een onderdeel van een modelketen. Het Wageningen-model verzorgt de neerslag-afvoer voor het landelijk gebied (Sobek RR; rainfall-runoff) wordt gebruikt voor het stedelijk gebied). Samen met een RWZI-module (rioolwaterzuiveringsinstallatie) vormen deze de invoer voor Sobek Channel Flow en een 2D maaiveld-overstromingsmodel.

Kritiek:

Feitelijk toepassingsgebied hoort meer bij beoogd toepassingsgebied. De rol in een modelketen is eigenlijk een vraag op zich (T&B).

2D Overlapt het beoogde toepassingsgebied alle daadwerkelijke toepassingen (ja/deels/nee)?

Grotendeels overlapt het daadwerkelijk toepassingsgebied het beoogde toepassingsgebied. Uitzonderingen zijn bijvoorbeeld een studie waarin het Wageningen-model in het kader van onderzoek naar de effecten van de Duitse bruinkoolwinning (en daaraan gepaarde grondwaterstandsverlaging) is toegepast in een stroomgebied in Limburg, waarin breuken en grondwaterstroming over de systeemgrens een rol speelt. Additionele informatie is hierover wel gewenst.

2E Indien relevant, in hoeverre overlappen de verschillende toepassingen met elkaar (goed/matig/slecht)?

T&B: Niet duidelijk was wat hiermee werd bedoeld.

V&W: Er is een ruimtelijke overlap tussen de lokale en regionale modellen.

Evaluatie van de checklist:

De vraag richt zich op het beoogd en feitelijk toepassingsgebied van het te evalueren model. Dit is echter niet direct relevant met betrekking tot de geschiktheid van het model voor het toepassingsgebied. Het zou beter zijn om eerst te kijken hoe het beoogde toepassingsgebied kan worden gekenmerkt, en vervolgens te zien in hoeverre het feitelijke toepassingsgebied aan deze kenmerken voldoet. Hiermee is ook direct de strekking van deelvraag E duidelijk. Verder zou deze

vraag met name gericht moeten zijn op het doel van het model: welke rol gaan de modelresultaten spelen in het verdere proces? Welke eisen ten aanzien van variabelen, acceptabele onzekerheden, tijd- en ruimteschalen komen hieruit voort? Onder welke voorwaarden kan het beoogde model hier aan voldoen?

3.4 Systeemanalyse Wageningen-model

3A Geef een systeemanalyse. Dit kan bv. een verwijzing naar een rapport zijn. Aspecten die daarin aan bod komen zouden moeten zijn: Wat zijn de belangrijkste attributen en processen van het systeem? Welke terugkoppelingen zijn aanwezig? Hoe zijn de grenzen van het systeem bepaald of gedefinieerd? Hoe wordt met relevante attributen en processen buiten deze grenzen omgegaan, bv. randvoorwaarden, constanten, of simpelweg genegeerd? Over welke processen, attributen, terugkoppelingen, etc. bestaat er onzekerheid? Hoe is die onzekerheid bepaald?

T&B: Hiervoor wordt doorverwezen naar de verschillende publicaties over het Wageningen-model. De belangrijkste attributen zijn toestandsvariabelen zoals waterberging (grondwater en bodemvocht) en systeemeigenschappen zoals porositeit en doorlaatbaarheidskarakteristieken. De belangrijkste processen zijn de infiltratie van neerslag in de bodem, verdamping door voornamelijk vegetatie, en de stroming van water langs verschillende stroombanen (over het oppervlak, ondiep door de bodem, als grondwater). De belangrijkste terugkoppelingen zijn die tussen berging enerzijds en verdamping en afvoer anderzijds. Afhankelijk van de hoeveelheid berging in het systeem en de intensiteit van de neerslag zijn verschillende van bovengenoemde stroombanen actief.

In het Wageningen-model zitten verschillende terugkoppelingen. Ten eerste de verdeling tussen snelle en trage afvoer, en ten tweede de capillaire opstijging, die beide een functie zijn van de berging in het 'trage' reservoir.

De grenzen van het systeem zijn aan de bovenkant het landoppervlak (inclusief vegetatie) en aan de onderkant de overgang van het aan het oppervlak liggend watervoerend pakket naar de eerste (niet of slecht doorlaatbare) scheidende laag. Lateraal zijn het 'bovenstrooms' de waterscheiding (waar per definitie de laterale flux nul is) en 'benedenstrooms' de overgang van het landsysteem naar het oppervlakte-watersysteem. Een probleem hierbij is dat sommige systeemgrenzen in principe een afgeleide zijn van de hydrologische fluxen, en dus per definitie dynamisch. Een voorbeeld hierbij is de waterscheiding die als laterale systeemgrens fungeert. Doorgaans wordt deze bepaald op basis van een topografische analyse. Een aanname hierbij is dat de waterstroming door topografie wordt gestuurd. Voor bijvoorbeeld oppervlakkige of ondiep-laterale stroming in reliëfrijke gebieden gaat dat vaak op. Voor grondwaterstroming is de topografie minder relevant, en wordt de positie van de waterscheiding in een grondwaterkoepel primair bepaald door de hydraulische structuur van de ondergrond en de afstanden tot de waterlopen die als randvoorwaarde voor het (lokale) grondwatersysteem fungeren. In een complexe setting vallen de waterscheiding voor het oppervlakkige/ondiepe water en die voor het diepere grondwater dus niet samen. In de praktijk wordt op basis van het dominante proces één statische waterscheiding gedefinieerd. Ook de onderrand is niet onduidelijk gedefinieerd. Water vanaf een gegeven diepte maakt geen deel meer uit van het lokale grondwatersysteem, maar van een grootschaliger regionaal systeem en zal pas ver buiten het onderzoeksgebied als kwel weer naar boven komen. In principe zou dus de grens tussen het 'lokaal uittredende grondwater' en het regionale grondwater als ondergrens gedefinieerd kunnen worden. Deze is in de praktijk echter slecht identificeerbaar, en bovendien grondwaterstands- en dus seizoen afhankelijk. In de praktijk wordt vaak gewerkt met modellen met een vaste en beperkte diepte, en de uitwisseling met het regionale grondwatersysteem wordt als randvoorwaarde beschreven, bv. gebruik makend van kalibratie.

In het Wageningen-model wordt geen rekening gehouden met grondwateruitwisseling over de systeemranden heen. Water komt als neerslag via de bovenrand binnen, en verdwijnt als afvoer of verdamping. Uitwisseling met een regionaal grondwatersysteem is dus niet mogelijk.

Attributen en processen buiten deze grenzen zijn, gezien bovenstaande met name stijghoogtes en/of grondwaterstanden in de omgeving van het modeldomein, zowel lateraal als in de diepere ondergrond.

Klassieke randvoorwaarden zijn de forcering van het systeem: neerslag, straling en overige relevante meteorologische variabelen.

Onzekerheden zijn wat betreft processen 1) de grondwaterstroming, waar binnen (de recente versie van) het Wageningen-model de onderliggende fysica vereenvoudigd is tot een lineair reservoir; 2) de verdampingsreductie onder droge omstandigheden. In het Wageningen-model is de wortelzone (waar te verdampen water aan wordt onttrokken) beschreven als een enkel (lumped) reservoir. 3) capillaire opstijging, die waarschijnlijk te simpel is beschreven. Terugkoppelingen zoals de dynamische verdeling tussen snelle en langzame afvoer lopen via een parameter zonder werkelijke fysische betekenis. Het Wageningen-model is gevoelig voor de begintoestand, terwijl het gezien het conceptuele karakter van het model niet eenvoudig is hier iets zinnigs over te zeggen.

V&W: De belangrijkste systeemvariabelen zijn gebiedseigenschappen zoals het landgebruik, doorlaatvermogen van de bodem, het voorkomen van eventuele hellingen, de lokatie van watergangen, de aanwezigheid van drainagemiddelen (en de hydraulische eigenschappen daarvan), en kenmerken van de regionale grondwaterstroming.

Grenzen van het systeemdomein zijn in beginsel de stroomgebieden van de rivier de Dommel. Oorspronkelijk was het domein opgedeeld in grote afwateringseenheden. Ten behoeve van een betere modellering zijn deze echter verkleind, op basis van informatie over maaiveldhoogte, het watergangenpatroon etc. De eenheden zijn dusdanig gekozen dat het stroomgebieden zijn.

Lacunes in de systeemanalyse zijn met name dat de regionale grondwaterstroming niet in rekening wordt gebracht. De grootste onzekerheid bestaat in de oppervlakte-waterafvoer, en het inschatten van de actuele evapotranspiratie in relatie tot de procesbeschrijving. Andere onzekere factoren zijn kwelfluxen, de terugkoppeling tussen grond- en oppervlaktewater, en neerslag. Met name de onzekerheid in afvoer bleek een ware eye-opener. In feite is niet goed bekend hoe groot deze onzekerheid is. Het is bijvoorbeeld maar de vraag of die kleiner dan 10% is.

Kritiek:

T&B: Vragen over randvoorwaarden, constanten etc. horen bij een modevaluatie, niet bij een systeemanalyse. Die dient enkel over het fysische systeem te gaan (natuurlijk wel conceptueel model)

Beschouw nu per toepassing

3B Welke systeem-analytische aspecten (attributen, processen, terugkoppelingen, etc.) zijn (direct) relevant voor de toepassing? En welke in mindere mate of niet? En hoe is die relevantie bepaald?

Toepassingen:

T&B: Het Wageningen-model is toegepast voor het waterschap Noorderzijlvest (NZV): Een eerdere modelstudie, uitgevoerd met SOBEK-RR (rainfall-runoff module) leidde tot ontevredenheid. In het model bleef te veel water in de polders achter, waardoor de berekende afvoeren lager waren dan in werkelijkheid. De verhouding afvoer-berging was dus fout. De vraag was of er hier sprake was van een verkeerd model, of een verkeerde toepassing daarvan. Het Wageningen-model werd ingezet als 2nd opinion.

V&W: Het Wageningen-model is toegepast voor zowel de inrichting van waterbeheersgebieden (lokaal) als wel als onderdeel van het MRW (regionaal)

Relevante systeem-analytische aspecten per toepassing:

T&B: In deze toepassing is de relatie tussen berging en afvoer het meest relevant. Capillaire opstijging is het minst relevante proces, aangezien de studie enkel op natte omstandigheden gericht was, waarbij er enkel neer- en zijwaartse waterbewegingen optreden. Ook bleek het oorspronkelijke Kraijenhoff van der Leur J-model te complex te zijn (wat betreft rekentijd, niet wat betreft parameters). Gezien de periodiek hoge peilen in het oppervlaktewater is de terugkoppeling tussen grond- en oppervlaktewater een relevant proces.

V&W: Alle aspecten zijn relevant. Inzichten hierin zijn verkregen door logisch te redeneren op basis van vak- en gebiedskennis.

3C Zijn alle voor de toepassing belangrijke aspecten als gegeven in B meegenomen in het model/bestand? En welke niet?

T&B: Nee. Het grondwater zit niet afdoende realistisch in het Wageningen-model. Ten eerste mist in het Wageningen-model een terugkoppeling tussen grondwater en oppervlaktewater. Ten tweede is in het onderzoeksgebied kwel significant, maar kan niet worden verdisconteerd in het model.

V&W: Nee: factoren als landgebruik, topografische hellingen en grondwaterstroming kunnen slechts in zeer beperkte zin worden meegenomen in het model. Hydraulische weerstanden worden wel in Sobek meegenomen, maar niet in het Wageningen-model.

Kritiek:

Voor de vraag kunnen verschillende antwoorden worden gegeven afhankelijk of men het model op zich evalueert, of het model in zijn rol binnen een modelketen, wat doorgaans het geval is.

3D Zijn er minder of niet belangrijke aspecten meegenomen in het model/bestand (ja/weinig/geen)?

T&B: Ja. Capillaire opstijging is een modelonderdeel dat niet relevant is voor deze toepassing. Verder worden alle modelcomponenten gebruikt.

V&W: Op de hoge heidegronden met diepe grondwaterstanden is capillaire opstijging niet relevant.

3E Beoordeel de mate van 'evenwicht' m.b.t. de systeemanalyse, gebaseerd op de antwoorden op 3C en 3D (goed/matig/slecht).

T&B: De mate van evenwicht wordt geschat als "redelijk". Het Wageningen-model is niet overcomplex, maar ook niet te simpel. Een ruimtelijk geaggregeerd model, wat het Wageningen-model is, lijkt toepasselijk, aangezien de ruimtelijke heterogeniteit beperkt is, en de eventuele meerwaarde van een ruimtelijk gedistribueerd model dus gering.

V&W: Het is moeilijk om heterogeniteit in gebiedskenmerken in het Wageningen-model te verwerken. Op regionale schaal is dit niet zo erg, maar op lokale schaal vormt dit wel een belemmering. In de praktijk wordt de grootte van de modeleenheden bepaald door de locaties van afvoermeetinstallaties, en niet door homogeniteit van het bijbehorende stoomgebied. Ondanks deze beperkingen kan men er mee uit de voeten. De resultaten en onzekerheden zijn voor zowel projectleider en andere stakeholders acceptabel. Op regionale schaal is de plausibiliteit en acceptatie wisselend, waarbij de vraag opkwam hoe goed de modelresultaten moeten zijn. En wat zou centraal moeten staan? De beschikbaarheid van data, of de meest relevante processen?

Evaluatie van de checklist.

De vragen in (A) die verwijzen naar randvoorwaarden, constanten, etc. horen niet bij een systeemanalyse, maar maken deel uit van een modevaluatie, en dienen dus te worden verschoven naar vraag 4.

3.5 Conceptueel model Wageningen-model

4A Geef het conceptuele model (voor een model) of ontwerpmodel (voor een bestand). Dit kan bv. een verwijzing naar een rapport zijn. Het conceptuele model legt op relatief informele wijze de relaties tussen componenten vast. Aspecten die aan bod kunnen komen zijn: Welke aannamen en vereenvoudigingen zijn er gemaakt? Wat zijn de ruimtelijke dimensies? Welke ruimtelijk-temporele aggregatieniveaus worden gebruikt? Welke zaken spelen een rol voor het bestand? Wat wordt er voor (des)aggregatie van data gebruikt? Voor modellen: Welk modeltype is gebruikt?

De kern van het model wordt gevormd door een tweetal afvoerprocessen. Een 'snel' proces, wat stroming over het oppervlak, door drains, greppels en beekjes etc. representeert, en een 'langzaam' proces, wat de (ondiepe) grondwaterstroming naar de hoofdbeek representeert. Deze twee processen worden gevoed door 'effectieve' neerslag, die op zijn beurt berekend wordt uit een vereenvoudigde waterbalans van een ondiep bodemvochtreservoir, wat vervolgens verdeeld wordt over deze twee afvoerprocessen. De resulterende afvoergolven worden opgeteld.

Kenmerkend voor het model zijn een tweetal terugkoppelingen. Ten eerste is de verdeling van effectieve neerslag over de snelle en langzame afvoerprocessen dynamisch, en wordt gemodelleerd als een functie van de hoeveelheid berging in een grondwaterreservoir (welke gekoppeld is aan de dynamiek van het langzame reservoir). De rationale hierachter is dat als na een langdurige regenbui de grondwaterspiegel stijgt, er relatief meer water via de snelle stroombanen wordt afgevoerd, dan gedurende een drogere periode met minder grondwaterberging.

Een tweede terugkoppeling wordt gevormd doordat onder droge omstandigheden water door middel van capillaire opstijging van het grondwaterreservoir naar het bodemvochtreservoir stroomt.

Beschouw nu per toepassing

4B Op welke (ruimtelijk-temporeel) aggregatieniveaus zijn antwoorden gewenst, gezien de toepassing? Welke ruimtelijke dimensies en welk modeltype/ontwerpmodel zijn gewenst?

T&B: Vanuit de toepassing Noorderzijvest (NZV) is een model op polderniveau gewenst. Concrete antwoorden dienen te worden gegeven in de vorm van afvoer naar het hoofdsysteem (de polderboezem).

V&W: Het Wageningen-model wordt gebruikt om inzicht te krijgen in het gedrag van relatief homogene stroomgebieden.

4C Komen de ruimtelijke dimensies, aggregatieniveaus, etc. overeen met diegene die vanuit de toepassing zijn gewenst (ja/matig/niet)?

T&B: Ja, het Wageningen-model is een ruimtelijk geaggregeerd model wat direct op de verlangde schaal kan opereren.

V&W: Nee, want de ruimte aggregatieschaal wordt primair bepaald door aan- en afwezigheid van meetgegevens.

Evaluatie van de checklist:

In (B) wordt niet gevraagd welke mate van dynamiek, terugkoppelingen etc. vanuit de toepassingen acceptabel zijn; in (C) wordt niet gevraagd of de in het model door-gevoerde vereenvoudigingen aan deze voorwaarden kunnen voldoen.

3.6 Gegevens Wageningen-model

5A Welke gegevens zijn benodigd, gezien de toepassing, en motiveer? Wees specifiek m.b.t. resolutie, nauwkeurigheid, schaal, etc., en vermijd algemene en ongemotiveerde opmerkingen als "file X moet gebruikt worden".

T&B: De gegevens die nodig zijn om het Wageningen-model te draaien (forceringsdata) zijn altijd: neerslag (P) en potentiële evapotranspiratie (PET). Afvoermetingen (Q) zijn nodig t.b.v. kalibratie. Belangrijk is dat P en PET op dezelfde tijdstippen beschikbaar zijn. P en PET zijn in feite ruimtelijke geaggregeerd over het modeldomein. De onderlinge relatie tussen P en PET , en het evenwicht daarin, is een aandachtspunt hierbij. De gewenste temporele resolutie van P , PET , en Q is liefst op uur-basis tot maximaal dagelijks (afhankelijk van de toepassing en het stroomgebied).

V&W: P , Q , en PET zijn op uur-niveau nodig. Voor P wordt het liefst gebruik gemaakt van ruimtelijk verdeelde data. Voor Q zijn zowel de afvoeren op stroomgebiedsschaal nodig, als per watergang, aangezien er doorgaans antwoorden worden verlangd op deelstroomgebiedsschaal. Van de aanwezige stuwen is het nodig te weten wanneer deze zijn ingezet/aangepast.

Er is niet gekeken naar de gewenste nauwkeurigheid.

5B Welke gegevens zijn beschikbaar, en in welke mate?

T&B: P en PET worden doorgaans verkregen door ruimtelijke selectie of interpolatie vanuit landelijke gegevens (meteostations), eventueel gecombineerd met radar-gebaseerde neerslagkaarten. PET is vaak niet beschikbaar op dezelfde (hogere) tijdsresolutie als neerslag; desaggregatie vindt plaats o.b.v. van een aangenomen sinus-vormige dagelijkse gang. Dat dit een relevante bron van onnauwkeurigheid is wordt onderkend.

Afvoermetingen zijn doorgaans afkomstig uit peilmetingen die worden omgerekend tot afvoermetingen, of uit meer directe stroomsnelheidsmetingen, bijvoorbeeld met een (horizontale) ADCP (Acoustic Doppler Current Profiler). Ook de betrouwbaarheid van afvoermetingen is een zorgkind. Aangezien veel waterschappen primair peilbeheerders zijn hebben afvoermetingen niet altijd de prioriteit gehad die ze verdienen, en zijn deze dientengevolge minder betrouwbaar dan gewenst.

Benodigde nabewerkingen op de ruwe data zijn dus interpolatie, synchronisatie en verschaling. Doorgaans wordt de tijdsresolutie van afvoermetingen als standaard genomen, en worden P en PET hiermee gesynchroniseerd. Uit ervaringen met het Wageningen-model blijkt dat de modelparameters tijdschaalafhankelijk zijn: parameters gekalibreerd met afvoer op uur-basis hebben een andere waarde dan wanneer de kalibratie gebruik maakt van afvoer op dag-basis (dit hier noemen of ergens anders? – zou ook moeten worden onderzocht, en heeft implicaties voor evenwicht).

V&W: Neerslagdata worden verkregen vanuit 10 tot 20 KNMI meetstations. Deze zijn doorgaans op dag-basis, en worden gedesaggregeerd naar uur-data op basis van de dynamiek van 1 meetstation dat op uur-basis meet. Voor PET wordt gebruik gemaakt van 1 KNMI station. Daggegevens worden hier gedesaggregeerd naar uur-data via deling. Met een onderliggend dag-nacht ritme wordt dus geen rekening gehouden, aangezien dat niet relevant wordt geacht. Afvoergegevens zijn op een aantal locaties op uur-basis beschikbaar, en elders enkel op dag-basis. Deze laatste worden gedesaggregeerd naar uur-gegevens.

5C Wat is uiteindelijk de invoer van het model of bestand?

T&B: De uiteindelijke invoer bestaat uit tijdreeksen van P , PET en eventueel Q , geformatteerd als ASCII tekstbestanden.

V&W: Idem, de modelinvoer wordt gevormd door bewerkte meetgegevens.

5D Komt de invoer overeen met de door de toepassing gewenste gegevens (ja/deels/nee)? Vergelijk de antwoorden op 5A, 5B en 5C. Let hierbij sterk op eenheden en de drie elementen van schaal (coverage, support, extent; zie Bierkens et al, 2000; Bogaart et al, 2011). Heeft de data de juiste dimensies en eenheden? Zo nee, is er een (des)aggregatiemethode voorhanden? Schaalproblemen doen zich overal voor door de 'vertaling' van attributen naar metingen, die gediscrètiseerd zijn in tijd en ruimte, en onderhevig zijn aan allerlei bronnen van meetruis.

T&B: Dat hangt af van de nabewerking. Doorgaans worden puntmetingen gebruikt als schatter van een gebiedsgemiddelde. Zeker voor hoogfrequente (per uur of korter) neerslaggegevens geldt dat de spatio-temporele variatie dusdanig hoog is dat verplaatsing en opschaling een significante bron van onzekerheid is. Het gebruik van neerslagradar zou hier verbetering in kunnen brengen, maar de markt in neerslagproducten is momenteel te omslachtig om deze data in kleine projecten te gebruiken.

V&W: Nee. tijdreeksen hebben vaak de verkeerde temporele resolutie, en moeten dus worden gedisaggregeerd (wat op zich een foutenbron is), en afvoermetingen zijn vaak op andere locatie dan waarvoor antwoorden verlangd worden.

Evaluatie van de checklist:

Er zou expliciet onderscheidt gemaakt moeten worden tussen de verschillende soorten data: bv. ten behoeve van forcering (bv neerslag), ten behoeve van *a priori* parameterisering (bv. ruimtelijk gedistribueerde bodemdata), en ten behoeve van kalibratie (bv. de parameters van het Wageningen-model). De centrale vraag hierbij is: stellen de data het model in staat om afdoende aan de eisen vanuit de doelstelling te voldoen? Bij forcering- en ruimtelijke data telt bv. de resolutie en precisie. Bij data ten behoeve van kalibratie telt de vraag of die data voldoende informatie in zich meedraagt zodat de *posterior* onzekerheden in modelparameters tot acceptabele proporties kunnen worden teruggedrongen. Er spreekt voor zich dat hier een overlap is met Vraag 10 (kalibratie).

3.7 Formeel Wageningen-model

Gebruik de antwoorden op vragen 5A tot 5D om vragen 6B tot 6D te antwoorden.

6A Geef het formeel model, en motiveer de keuze of verwijs naar paginanummers uit een gepubliceerd rapport of artikel. Het formeel model is de omzetting van concept naar bijvoorbeeld wiskundige vergelijkingen of een set rekenregels. De keuze van formeel model hangt af van de beoogde toepassing en de data, en wordt mede bepaald door de gestelde onzekerheidsmarges.

Zoals al opgemerkt, is er niet één uniek Wageningen-model.

Bodemvochtbalans

Het bodemvochtdeelmodel wordt gevormd door een eenvoudig boekhoudschema, waarin per tijdstap de waterbalans wordt bijgehouden, gekwantificeerd als de hoeveelheid bodemvocht $S[L]$

$$S_t = S_{t-1} + N_t - E_t - P_t + C_t$$

waarin $N[L]$ de neerslag, $E[L]$ de hoeveelheid verdamping, $P[L]$ de percolatie vanuit het bodemvochtreservoir naar grondwater en afvoer, en $C[L]$ de capillaire opstijging. De percolatie en capillaire opstijging zijn complementaire processen.

Percolatie

Percolatie vindt alleen plaats onder 'natte' omstandigheden, gedefinieerd door de verhouding tussen S en veldcapaciteit $F[L]$. Wanneer geldt dat $S > F$, dan P is gegeven door

$$P_t = \alpha S_t (S_t - F) / S_{max}$$

Waarin $\alpha[-]$ een dimensie-loze parameter, en $S_{max}[L]$ de bergingscapaciteit van het bodemvochtreservoir.

Capillaire opstijging

Onder droge omstandigheden, waarbij $S_t < F$, treedt capillaire opstijging op, gegeven door

$$C = \beta (F - S)G$$

Waarin $\beta[-]$ een dimensie-loze parameter en $G[L]$ de hoeveelheid berging in het grondwaterreservoir.

Verdamping

Verdamping E wordt beschreven als een lineaire reductie van de potentiële verdamping $E_{pot}[L]$

$$E = r E_{pot}$$

waarbij de reductiefactor r , afhankelijk van de versie van het Wageningen-model, op verschillende manieren wordt berekend. In de recente FORTRAN versie wordt gebruik gemaakt van een sinusvormige afhankelijkheid van de relatieve vochtgehalte S/F

$$r = 1 - [\cos\left(\frac{\pi S}{F}\right) + 1]/2$$

met $r = 1$ voor $S > F$, terwijl in bepaalde spreadsheetversies gebruik wordt gemaakt van een power-law model

$$r = (S/F)^y$$

Verdeling

Onder natte omstandigheden, dus wanneer $S > F$, wordt de percolatieflux P verdeeld over het snelle en het langzame afvoerproces, waarbij de snelle fractie f een lineaire functie is van de grondwaterberging G

$$f = \delta G$$

waarin $\delta[-]$ een parameter. Het snelle afvoerproces wordt dus gevoed door een flux fP , en het langzame proces door een flux $(1 - f)P$.

Afvoer

Snelle afvoer

In de oorspronkelijke versie van het Wageningen-model (Stricker en Warmerdam, 1982) wordt het snelle afvoer proces beschreven met een convectie-diffusievergelijking, waarbij de (eenheids)afvoer Qf op een afstand $x[L]$ en tijd $t[T]$ wordt berekend met de impulsie-response functie

$$Q_{f,1}(t) = x e^{-(x-At)^2/4Dt} / 2\sqrt{\pi Dt^3}$$

waarin D een diffusiecoëfficiënt en A een translatiecoëfficiënt.

In meer recentere versies is dit model vervangen door een lineair reservoir, welke rechtstreeks wordt opgelost voor de afvoerflux Q (d.w.z. dus zonder een expliciete toestandsvariabele voor de berging in het reservoir). Onder deze benadering is

$$Q_{f,i} = Q_{f,i-1}e^{-1/k_f} + f P (1 - e^{-1/k_f})$$

met k_f [?] een niet duidelijk gespecificeerde parameter.

Langzame afvoer

De dynamiek van het langzame afvoerproces is gebaseerd op eerder werk van eerder werk van Kraijenhoff van de Leur (1958), die op grond van de grondwaterhydraulica een vergelijking opstelde voor de hydrograaf als gevolg van niet-stationaire stroming van ondiep grondwater naar een beek of rivier

$$Q_{s,1}(t) = \frac{8}{\pi^2} \frac{1}{j} \sum_{n=1,3,5,\dots}^{\infty} e^{-n^2 t/j}$$

waarin j de zogenaamde reservoirconstante is, gegeven door

$$j = \frac{1}{\pi^2} \frac{\mu L^2}{kD}$$

met μ [-] de effectieve porositeit, L [L] de gemiddelde afstand tussen waterlopen en kD [$L^2 T^{-1}$] de transmissiviteit.

In latere versies van het Wageningen-model wordt ook het langzame afvoerproces beschreven als een lineair reservoir, zodat

$$Q_{f,i} = Q_{f,i-1}e^{-1/k_s} + f P (1 - e^{-1/k_s})$$

waarin parameter k_s een tijdconstante is.

Totale afvoer

Per tijdstap waarin percolatie optreedt, dus $P > 0$, worden de bijbehorende instantane hydrografen

$$Q_f(t) = f P Q_{f,1}(t)$$

en

$$Q_s(t) = (1 - f) P(t) Q_{s,1}(t)$$

berekend, welke voor alle tijdstappen worden opgeteld tot

$$Q(t) = Q_f(t) + Q_s(t)$$

Grondwaterberging

De dynamiek van de grondwaterberging G wordt bepaald door de aanvulling vanuit percolatie P en uitputting door het langzame afvoerproces Q_s

$$\frac{\Delta G}{\Delta t} = (1 - f)P - Q_s$$

6B Welke gegevens zijn benodigd voor dit formele model?

Voor het formele model zijn dezelfde data nodig als voor het conceptuele model: P , PET en Q , in de vorm van tijdreeksen, bv. $P(t) = P_1 \dots P_n$. Verder zijn waarden benodigd voor de modelparameters, zoals eerder genoemd.

Commentaar:

Wat is verschil met 5A?

6C Welke gegevens zijn beschikbaar, en in welke mate?

Voor wat betreft de data, zie de discussie in punt 5B. Voor wat betreft de modelparameters: doorgaans zijn deze niet beschikbaar, en deze moeten dus specifiek worden gekalibreerd voor de toepassing. Gewasfactoren ten behoeve van verdamping kunnen worden afgeleid van beschikbare landgebruiksgegevens (bv. kaarten).

Commentaar:

Wat is verschil met 5B?

6D Zijn de gegevens die benodigd én beschikbaar zijn ook gebruikt? Waarom wel/niet?

T&B: Ja, voor deze toepassing zijn alle relevante gegevens gebruikt.

V&W: Nee, momenteel worden verdampingsgegevens uit bv. satellietmetingen, radar-gebaseerde neerslagmetingen, en gedetailleerde gebiedseigenschappen niet gebruikt.

Evaluatie van de checklist:

De deelvraag (A) naar het formele model *an sich* is helder en noodzakelijk. Het zou echter goed zijn om ook hier te bepalen of het formele model in evenwicht is met het conceptuele model. Zijn de gebruikte vergelijkingen en/of balansen inderdaad de beste keuze om de essentie van de eerder vastgestelde relevante conceptualisaties te representeren? In geval van het Wageningen-model blijkt dat er verschillende keuzes gemaakt kunnen worden met betrekking tot de afvoerprocessen, die elk onder specifieke omstandigheden 'het beste' zijn. Er moet dus meer aandacht komen voor de keuzes die gemaakt worden bij het formele model. De deelvragen (B) t/m (D) hebben een sterke overlap met de corresponderende deelvragen uit Vraag 5. Je kunt je afvragen of ze bij Vraag 5 thuishoren, aangezien een conceptueel model misschien per definitie meer verhalend is, dan een kwantificering waarbinnen data een rol kan spelen. Voorgesteld wordt dus om de datavragen van Vragen 5 en 6 te combineren en aan Vraag 6 (of 7, zie aldaar) toe te wijzen.

3.8 Numerieke implementatie Wageningen-model

7A Geef het numeriek rekenmodel/gegevensbestand (de implementatie in software), en motiveer de keuze of verwijst naar paginanummers uit een gepubliceerd rapport of artikel.

Zoals al opgemerkt, de numerieke implementatie van het Wageningen-model bestaat niet. Er zijn 3 verschillende versies van het Wageningen-model (eerder besproken). De oorspronkelijke referentieversie in FORTRAN; de spreadsheetversie zoals gebruikt in het onderwijs, en recente implementaties in R.

Commentaar:

Hoe kan dit worden 'gegeven'? Opname van source code in evaluatie?

Antwoord:

"of verwijst naar paginanummers uit een gepubliceerd rapport of artikel"

7B Beschrijf de verificatie of verwijst naar paginanummers uit een gepubliceerd rapport of artikel. Verificatie moet niet verward worden met validatie. De eerste controleert de omzetting van formeel model naar code, de tweede toetst de resultaten van de code aan de werkelijkheid.

De FORTRAN versie van het Wageningen-model zou vroeger eens geverifieerd zijn. Verdere informatie hierover ontbreekt. Er is dus niet voldoende informatie op deze vraag te beantwoorden.

7C Beschrijf en motiveer de keuze voor numerieke rekenmethoden, en geef de gebruikte discretisatie of verwijst naar paginanummers uit een gepubliceerd rapport of artikel. Veelal bestaat er bij numerieke rekenmethoden een afweging tussen rekensnelheid en –nauwkeurigheid.

De keuze voor de rekenmethoden is voornamelijk historisch bepaald. Het oorspronkelijke Wageningen-model is een beetje een rare hybride tussen enerzijds een transferfunctie (het J-model) wat werd gebruikt “omdat iedereen — in navolging van James Dooge — dat destijds deed” en anderzijds een differentiaalvergelijking. Dit heeft nadelen: een transferfunctie wordt ergens in de tijd afgekapt, en heeft dus waterverlies tot gevolg. De latere vervanging van het J-model door een lineair reservoir is vooral ingegeven door een wens tot versimpeling, die verantwoord werd door het acceptabele verlies aan precisie bij afdoende grote tijdstappen.

Evaluatie van de checklist:

Er wordt niet afdoende doorgevraagd in hoeverre de numerieke implementatie (van het formele model) een bron van fouten en onzekerheden kan zijn. Impliciet komt dit wel aan de orde (“een afweging tussen rekensnelheid en -nauwkeurigheid”), maar er kan hier meer worden gevraagd. In feite is het numerieke model (de ‘broncode’) de enige versie van het model die er werkelijk tot doet, aangezien dit de versie is die feitelijk wordt gedraaid, en waarvan de uitvoer wordt geïnterpreteerd. In het numerieke model worden data ingelezen en verwerkt. Vaak vindt er nog een bepaalde vertaalslag van de data plaats (aggregatie bijvoorbeeld of interpolatie). Ook daar zouden vragen over gesteld moeten worden.

3.9 Schematisering Wageningen-model

De schematisering is bv. het indelen van de bodem in verschillende lagen (‘stratificatie’), of het clusteren van landgebruikstypen. De mate van detail die bij schematisering wordt gevolgd is afhankelijk van de toepassing, maar moet wel voldoende ondersteund worden door de data.

8A Indien van toepassing, beschrijf de schematisering, of geef relevante en specifieke referenties?

Verticale schematisering is niet aan de orde binnen het Wageningen-model. Er zijn geen mogelijkheden om verticale gelaagdheid in het model in te brengen. Hooguit kan aan de parameters zoals gerelateerd aan de wortelzone, het ‘snelle’ reservoir en het ‘langzame’ reservoir een interpretatie worden gegeven die aan gelaagdheid gerelateerd is.

Horizontale schematisering speelt slechts gedeeltelijk een rol. Op zich is het Wageningen-model een ‘gelumped’ model, en dus speelt ruimtelijke variatie binnen een enkele modelentiteit geen rol. Op grotere schaal kan het Wageningen-model toegepast worden op een verzameling stroomgebieden, waarbij voor elk stroomgebied andere parameters kunnen worden gedefinieerd. Schematisering moet dus gekoppeld worden aan (sub)stroomgebiedsindeling.

Een probleem hierbij is de rol van grondwater. In het Wageningen-model speelt grondwaterstand (impliciet) een grote rol, en dienen (sub)stroomgebiedsgrenzen dus daar gelegd te worden waar de (laterale) grondwaterfluxen afwezig of minimaal zijn.

**8B Wordt de schematisering voldoende ondersteund door de data (ja/matig/nee)?
Motiveer.**

Niet van toepassing (zie antwoord vraag 8A). Er is doorgaans afdoende data beschikbaar die zou kunnen dienen ter ondersteuning van modelschematisering, maar het Wageningen-model weet hier geen raad mee.

Evaluatie van de checklist:

Er wordt niet expliciet gevraagd of de gebruikte schematisering aansluit bij de eisen vanuit de toepassing. Aangezien vaak met beschikbare schematiseringen wordt gewerkt is dit zeker geen *a priori* uitgemaakte zaak. Effecten van schematisering zijn vaak niet goed gekend. Een illustratief voorbeeld is bv. de in veel hydrologische modellen gebruikte PAWN bodemschematisering waarbij slechts 2 bodemlagen worden onderscheiden. Er is dus geen ruimte voor dunne afwijkende bodemlagen, terwijl deze wel een grote invloed kunnen hebben. Een relatief dunne grofzandige laag kan bv. een remmende werking op de capillaire opstijging onder droge omstandigheden hebben.

3.10 Gevoeligheden, onzekerheden, opbouw code

Deze vraag is gericht op het bepalen welke parameters, forcing, invoer, delen van de code, etc. belangrijk zijn en welke overbodig zijn, en welke onderdelen een grote onzekerheid introduceren.

9A Zijn er gevoeligheids-/onzekerheidsanalyses uitgevoerd? Zo ja, beschrijf deze of verwijst naar paginanummers uit een gepubliceerd rapport of artikel en analyseer de resultaten. Sommige factoren kunnen weinig van belang blijken voor de toepassing, zodat er weinig moeite hoeft te worden gedaan om deze goed te kalibreren, of gebruikt kunnen worden om het model te vereenvoudigen. Andere factoren kunnen juist essentieel blijken.

T&B: Ja, in veel afstudeerprojecten zijn er gevoeligheids- en onzekerheidsanalyses uitgevoerd. Doorgaans is dit niet op formele wijze uitgevoerd, maar meer 'houtje touwtje'. Hieruit komt voornamelijk dat een goede begintoestand belangrijk is. Het model is 'schraal' ontworpen met betrekking tot overtollige parameters, alle zijn essentieel. Hooguit is de capillaire opstijging (en bijbehorende parameters) irrelevant voor natte toepassingen waar dit proces geen rol speelt. Het is niet goed bekend hoe belangrijk de dynamische verdeler f is. Het blijkt dat andere modellen onderpresteren door gebrek aan dit mechanisme, maar hoe belangrijk deze parameter voor het Wageningen-model is niet goed bekend.

V&W: Door ons zijn niet expliciet algemene gevoeligheids- en onzekerheidsstudies uitgevoerd. Deze zijn wel bekend uit andere studies. Op onderdelen zijn ze wel uitgevoerd.

9B (niet voor databestanden) Welke numerieke integratiemethode is gebruikt? Bespreek het prestatievermogen van de methode.

De belangrijkste waterbalansdifferentiaalvergelijking wordt met de eenvoudigst mogelijke techniek opgelost: de Euler-integratie. De differentiaalvergelijking behorende bij het langzame grondwaterreservoir wordt analytisch opgelost. Niet-lineaire drempel en feedbackmechanismen in de wortelzone worden door middel van een iteratieve verfijning opgelost. In de praktijk bleek deze benadering adequaat te zijn, en is deze niet verder verfijnd.

9C *Wat is de rol van elke module van het model/bestand, bv. een gewasgroei-deelmodel als onderdeel van een hydrologisch model, of stappen in het stroomschema of de scripts van een bestand? Bespreek dat in relatie met de toepassing. Zijn er onderdelen die overbodig zijn? Missen er nog onderdelen? Motiveer waarom.*

Deze vraag is niet van toepassing. Het Wageningen-model kent geen modules, tenzij je de afzonderlijke conceptuele bakjes (wortelzone, grondwater, oppervlaktewater) als dusdanig zou willen bestempelen.

9D *Beoordeel de complexiteit in termen van rekentijd en efficiëntie. Wat is de duur van een run? Hoeveel (tussentijdse) uitvoer is er? Hoe groot is de datastroom? Zou de code sneller/efficiënter kunnen, bv. door selectie, aggregatie, andere methode, andere modulaire opbouw, etc.?*

Het Wageningen-model draait heel snel op een typische PC uit 2012 kan 10 jaar aan data met een temporele resolutie van een uur in minder dan 15 minuten worden doorgerekend.

Evaluatie van de checklist:

Deze vraag is een allegaartje. Deelvraag (A) is onderdeel van de gehele onzekerheidsanalyse (samen met kalibratie en validatie). Deelvragen (B) en (D) horen thuis bij het numerieke model, Vraag 7.

3.11 Kalibratie Wageningen-model

De kalibratie is het proces waarbij de parameters, begincondities, etc. van het model of bestand van waarden worden voorzien. Kalibratie is vaak gebaseerd op gevoeligheids-analyses (zie vorige vraag).

10A *Indien beschikbaar, beschrijf de uitgevoerde kalibraties en motiveer de keuze, of verwijz naar paginanummers uit een gepubliceerd rapport of artikel. Geef ook aan welke doelfuncties (lokaal, globaal, deterministisch, stochastisch) zijn gekozen, of dat de betrouwbaarheid van de gekalibreerde parameters wordt gekwantificeerd, hoe, op welke data, of en hoe de betrouwbaarheidsintervallen zijn bepaald, etc. met eventuele verwijzingen naar de literatuur. Let op over-fitting, meervoudige lokale optima, en identificeerbaarheidsproblemen.*

T&B: Doorgaans wordt het Wageningen-model gekalibreerd met behulp van het Levenberg-Marquardt algoritme. Oorspronkelijk gebeurde dit op een mainframe met behulp van AMSL code. Aan het Levenberg-Marquardt werd de voorkeur gegeven, boven alternatieven zoals het Nelder-Mead Simplex algoritme omdat de eerste een betere performance geeft op de gladde responsoppervlak zoals dat zich bij het Wageningen-model voordoet.

Enkel de modelparameters worden gekalibreerd, niet de begintoestand van het model. Als doelfunctie wordt de *Sum of Squared Residuals* (SSR) van de gemeten en berekende afvoertijdreeks gebruikt. De in de hydrologie gebruikelijke Nash-Sutcliffe Model Efficiency (NSME) wordt nadrukkelijk niet gebruikt omdat de SSR meer nadruk legt op de hogere afvoeren, wat voor toepassingen van het Wageningen-model doorgaans de bedoeling is.

Op dit moment wordt nagedacht over een aanpassing van deze doelfunctie, bijvoorbeeld door met verschillende gewichten voor verschillende omstandigheden of afvoerclassen te gaan werken. Ook wordt overwogen om kalibratie op dag-basis uit te voeren, ook als uur-gegevens beschikbaar zijn. Dit om te voorkomen dat eventuele temporele verschuivingen een negatief effect op de kalibratie hebben.

De Hessiaan van het responsoppervlak wordt wel berekend, maar hier wordt verder niets mee gedaan. In principe is de Hessiaan te gebruiken voor informatie over onzekerheid, correlaties tussen

parameters, etc. Ook wordt het voorkomen van lokale optima niet in detail onderzocht. Wel wordt het eventueel voorkomen hiervan geïdentificeerd door vanuit verschillende beginwaarden te werken. Een analyse van de parameter-identificeerbaarheid is beperkt tot de invloed van initiële condities op gekalibreerde parameterwaarden. Hoewel het bestaan van correlaties tussen parameters niet uitputtend is onderzocht, blijkt dat er interactie bestaat tussen bijvoorbeeld de percolatiesnelheid en de coëfficiënt van het snelle reservoir.

V&W: het Wageningen-model wordt afzonderlijk gekalibreerd per bemeten deelstroomgebied. Binnen bemeten deelstroomgebieden worden meerdere (identieke) Wageningen-modeleenheden worden een SOBEK Channel Flow-module gekoppeld, en het geheel wordt vervolgens ook gekalibreerd. De kalibratie wordt doorgaans handmatig ('iteratief') uitgevoerd. Het kalibratieproces is beknopt gedocumenteerd in rapportages.

Voor de toetsing van het Wageningen-model, als onderdeel van de modelketen met Sobek CF, wordt ook gebruik gemaakt voor informatie over waterpeilen, en gegevens over inundatie, bijvoorbeeld voor het hoogwater uit 2002. Dit zijn vaak losse, individuele observaties of zelf herinneringen van mensen. Ook wordt er gebruik gemaakt van luchtfoto's. Het goed inschatten van de bijbehorende herhalingsperiode is hierbij wel een probleem. Desalniettemin wordt deze gegevensbron informatief beschouwd.

10B Welke eisen stelt de toepassing aan de nauwkeurigheid waarmee parameters, etc. wordt bepaald. Let ook op over-fitting.

Hier is nooit echt over nagedacht. Wel zijn (handmatig) de invloed van initiële condities bepaald. De resultaten hiervan zijn ter kennisgeving aangenomen; bij de toepassing Noorderzijlvest "verdween dit effect in de ruis".

10C Welke mogelijkheden leveren de data om waarden van parameters, etc. te bepalen. Denk hierbij aan resolutie, nauwkeurigheid, aantal gegevens, het aggregatieniveau van de gegevens, en let op over-fitting.

Afvoermetingen zijn allesbepalend voor de kalibratie. Grondwatermetingen zijn lastig mee te nemen, qua modelconcept: het modelconcept 'grondwaterberging' laat zich moeilijk kwantitatief vergelijken met metingen van grondwaterstand. Zelfs als deze relatie lineair zou zijn, levert dat een extra kalibratieparameter op, waarbij het nog maar de vraag is over er nog genoeg informatie overblijft. Zoals al gemeld in vraag 6, worden verdampingsgegevens uit bv. satellietmetingen, radar-gebaseerde neerslag-metingen, en gedetailleerde gebiedseigenschappen momenteel niet gebruikt. Ook grondwatermetingen worden nu niet betrokken bij het kalibratieproces, voornamelijk omdat geen goede procedure voorhanden is om dit te doen. Het betrekken van bv. de grondwaterdynamiek als soft data (*sensu* Seibert) zou naar verwachting wel de onzekerheid kunnen verkleinen.

Kritiek:

Overlap met 6D (?)

Evaluatie van de checklist:

Er wordt wel gevraagd welke doelfuncties zijn gebruikt, maar dit wordt niet in de context van de toepassing besproken, wat wel zou moeten. Welke doelfuncties zijn gewenst gezien de specifieke toepassing? Zijn deze gebruikt?

3.12 Validatie Wageningen-model

Wij definiëren hier een validatie als een toets of het model of bestand een redelijke representatie is van het werkelijke systeem dat wordt gemodelleerd. Dit geschiedt op basis van een set onafhankelijke gegevens (gegevens die nog niet in een eerder stadium bij het modelleren gebruikt zijn).

11A Indien beschikbaar, beschrijf de gedane validatiestudies of verwijst naar paginanummers uit een gepubliceerd rapport of artikel.

T&B: Voor de validatie van het Wageningen-model wordt verwezen naar de verschillende scripties. In het algemeen wordt gebruik gemaakt van een *split-sample* benadering waarbij gekalibreerd wordt een relatief droge deeltijdreeks, en gevalideerd op een relatief natte deeltijdreeks (en omgekeerd). Over het algemeen blijkt dat de resulterende doelfuncties (Nash-Sutcliffe Model Efficiency en *Sum of Squared Residuals*) voor kalibratie en validatie vergelijkbaar zijn. Er zijn geen aanwijzingen voor overfitting van het model.

V&W: Modelparameters die waren gekalibreerd op het hoogwater van 2002 zijn gebruikt voor een validatie op het hoogwater van 1995. De resultaten voor 1995 gaven aan dat er nog aandacht nodig was voor verbeteringen; de resultaten waren nog niet hiervan zijn bevredigend.

11B Onderbouw/beoordeel de waarde van de validatie(s). Hoe relevant zijn de verschillende validatiestudies voor de verschillende toepassing(en) van het model of bestand? Zijn alle toepassingen voldoende afgedekt door de studies? Is ook de betrouwbaarheid van de validatie gekwantificeerd? Welke toepassingen moeten opnieuw worden gevalideerd?

T&B: Dit werd een lastige vraag gevonden. Voor de meeste afstudeervakken (de context waarbinnen de meeste validatiestudies worden uitgevoerd) is modelvalidatie een verplicht onderdeel. Voor de toepassing op waterschap Noorderzijlvest bleek de validatie “problematisch” te zijn [het is onduidelijk wat hiermee precies bedoeld wordt].

V&W: De validatieberekeningen zijn nuttig om te beoordelen of het model al voldoende is afgeregeld voor een toepassing, en of dat er nog meer inspanning moet worden gepleegd voor modelontwikkeling.

Kritiek:

Het is niet duidelijk hoe de betrouwbaarheid van de validatie gekwantificeerd kan worden, zoals in de checklist wordt gesuggereerd.

Evaluatie van de checklist:

Het is mogelijk dat bij validatiestudies een andere doelfunctie wordt gebruikt dan bij de kalibratie. Hier zou expliciet naar gevraagd moeten worden, ook omdat het gebruik van een andere doelfunctie de essentie van de validatie zou kunnen zijn (andere doelfunctie i.p.v. ander ruimtelijk domein of tijdinterval).

3.13 Samenvatting & evenwichtsbeoordeling Wageningen-model

Dit laatste onderdeel is gericht op het geaggregeerde resultaat van de antwoorden op de verschillende vragen en de meningen van de verschillende betrokkenen, en kan indirect ook tot adviezen leiden ter verbetering van het model of bestand.

12A Beoordeel de mate van vertrouwen in het model/bestand in het algemeen, mede gebaseerd op wat er is gedaan aan testen, verificatie, kalibratie, validatie, gevoeligheids- en onzekerheidsanalyses, en motiveer deze.

T&B: Ondanks de “archaische” toestand van het Wageningen-model is er groot vertrouwen in het model. Het heeft nog veel potentie. Met name de dynamische verdeler f is een belangrijk concept wat in andere conceptuele waterbalansmodellen ontbreekt. Verder is de expliciet berekening van de Hessiaan een voordeel wat nog verder uitgebuit kan worden. Jaren ervaring met het Wageningen-model heeft laten zien dat het model acceptabele uitkomsten geeft, en geen “rare dingen doet”.

V&W: In de 'lokale' modellen heeft men meer vertrouwen dan in de 'regionale' modellen. De schematisering bij de eerste is duidelijk adequater dan bij de laatste. Onder andere de toenemende rekentijden bij de regionale modellen zijn daar debet aan. Ook is er doorgaans gebrek aan tijd om alle in het modeldomein voorkomende waterlopen in het veld te controleren, etc.

12B Zijn er veel missende zaken? Zijn er veel vragen waar de zaken als 'te eenvoudig' zijn beoordeeld? Motiveer.

T&B: Een aantal onderdelen worden gemist:

1. De koppeling met het regionaal grondwater;
2. De verticale verdeling van water is niet goed te koppelen met grond- en oppervlaktewatermetingen;
3. De koppeling van modeldomeinen met naastliggende domeinen.

V&W: Wat met name gemist wordt is een methode om (heterogene) gebiedskenmerken in het model mee te nemen, bijvoorbeeld door het modeldomein ruimtelijke te kunnen onderverdelen in sub-domeinen.

12C Zijn er veel overbodige componenten? Zijn er veel onderdelen waar het eenvoudiger zou kunnen? Motiveer.

T&B: Nee. Het model is dusdanig 'kaal' dat er geen overbodige onderdelen inzitten.

12D Is er voldoende datasteun, en waaruit blijkt dat? Motiveer

T&B: Over het algemeen is dit geen probleem. Aandachtpunten zijn de precisie van afvoermetingen en het onvermogen om grondwaterstandsmetingen te gebruiken.

V&W: Nee. Vooral afvoermetingen zijn de zwakke plek, juist tijdens hoogwater, worden afvoerpieken gemist, bijvoorbeeld omdat het water om een meet-stuw heen stroomt, of omdat de afvoeren buiten het meetbereik van de meetinstallatie valt. Ook is vanuit de kantonniers niet genoeg informatie wanneer stuwen omgegaan zijn, etc.

Het vermoeden bestaat dat onder extreme omstandigheden het model soms beter en/of betrouwbaarder is dan de meetgegevens.

12E Welke specifieke suggesties voor verbetering en data-aanlevering komen naar voren uit deze analyse?

T&B: Suggesties zijn de uitbreidingen die onder B, zijn genoemd, de precisie van afvoermetingen en de synchronisatie van afvoermetingen met metingen van evapo-transpiratie, alsmede de (ruimtelijke) representativiteit daarvan.

V&W: Zorg dat je je vragen (waar je wat voor data wilt hebben, wat je toepassingen zijn) heel helder hebt. Besteed ook genoeg tijd aan kwaliteitscontrole van de data. Koppel het ontwikkelen en kalibreren van modellen los van de toepassing. Eigenlijk zou je hiervoor een tool-kit van deelmodellen en concepten moeten ontwikkelen. Dan kun je per toepassing kiezen uit meerdere modellen, of nieuwe modellen (laten) maken indien gewenst. Verder is een goed databeheer cruciaal.

Evaluatie van de checklist:

Goede vragen, geen opmerkingen.

4 MetaSWAP

4.1 Inleiding

Het model MetaSWAP (Van Walsum & Groenendijk, 2008; Van Walsum & Veldhuizen, 2011) is een numeriek simulatiemodel dat de grondwaterstand in de onverzadigde zone berekent. Het is bedoeld als 'meta-model' van SWAP (Van Dam *et al*, 2008; Kroes *et al*, 2009), een zeer gedetailleerd model dat het transport van water in de onverzadigde zone beschrijft. Voor studies in Nederland is een dergelijk model van groot belang voor goed waterbeheer: Nederland kent een grote populatie met divers watergebruik, waaronder intensieve landbouw en chemische industrie, veel regenoverschot, en slechts kleine hoogte-variaties in de doorlaatbare sedimenten. SWAP is zeer complex in termen van rekenprocessen en gegevensstromen, waardoor de 'run time' van SWAP zodanig is, dat het absoluut onpraktisch is om het model veelvuldig te 'runnen' om het uitputtend te testen, kalibreren, en valideren, en daarnaast nog steeds nieuwe situaties door te rekenen. Er is daarom besloten om het model binnen toepassingen waarin veel gerekend moet worden te vervangen voor een 'meta-model', dat de eigenschappen van SWAP imiteert, maar tegelijkertijd een veel kleinere rekentijd heeft.

In deze casus is de vragenlijst EMC v1.0 afgelopen in een korte serie van interviews met Ab Veldhuizen, de front-office beheerder van MetaSWAP, aangevuld met informatie uit de Status A-beoordeling van MetaSWAP in 2011 (Van Walsum & Veldhuizen, 2011). Zowel de functionaliteit van de lijst als de complexiteit van MetaSWAP worden getoetst in deze casus. Waar nodig wordt buiten de lijst om extra informatie verschaft of wordt aangegeven dat informatie ontbreekt die nodig is voor een evenwichtsbeoordeling; er kan dan ook geconcludeerd worden dat de lijst op dat punt nog niet voldoende functioneel is.

4.2 Ingevulde vragenlijst MetaSWAP; basale gegevens

1 Geef aan: model/bestand, naam, versienummer, revisienummer, uitgiftdatum. Onder model verstaan we de broncode + invoergegevens.

MetaSWAP 7.1.4c/7.2.0, gebruikt in NHI v2.1, uitgebracht september 2011. Het verschil in de versies zit 'm in dat versie 7.1.4c specifiek is ontwikkeld om de verdamping gelijk aan die in SWAP te krijgen, terwijl door de ontwikkeling van MetaSWAP helder is geworden dat de verdamping in SWAP niet altijd correct kan worden berekend (Van Walsum & Veldhuizen, 2011, p. 16).

Opmerkingen bij vraag 1:

De naamgeving 'MetaSWAP' suggereert een meta-model. Het is een open vraag hier of dat de vragen heel anders moeten zijn voor een meta-model i.v.m. een 'gewoon' model.

Tevens is vooraf de vraag voorgelegd wat de a priori inschatting is m.b.t. 'evenwicht'. De meningen daarover lijken te zijn verdeeld: volgens sommigen is MetaSWAP de beste maar ook meest complexe oplossing is die nog hanteerbaar is ter vervanging van SWAP, terwijl een deel van de gebruikers aangeeft dat MetaSWAP toch te complex is om goed te gebruiken.

4.3 Doel & toepassingen MetaSWAP

2A. Wat was het doel van het bestand/model?

Deze vraag overlapt met vraag A2 van de Status A-evaluatielijst, die wordt gebruikt in de kwaliteitsslag van de WOT. Het doel is beschreven in Van Walsum & Veldhuizen (2011), p. 17. Functioneel gezien is het doel van MetaSWAP het vervangen van SWAP in toepassingen die veel rekentijd vergen, door de rekentijd te reduceren en de gegevensstromen beheersbaar te maken. Het eigenlijk doel van SWAP en dus van MetaSWAP is de beschrijving van de dynamiek van het grondwater in de onverzadigde zone (bovenste bodemlagen) als koppeling tussen atmosfeer (neerslag, verdamping) met het diepe grondwater. Daarnaast moet het op MODFLOW passen, het algemeen gebruikte model voor de beschrijving van de grondwaterdynamiek in de verzadigde zone.

2B. Wat was het beoogde toepassingsgebied, en waaruit blijkt dit?

Deze vraag is vergelijkbaar met vraag A3 van de Status A-evaluatielijst, waarin gevraagd wordt naar het toepassingsgebied. Het toepassingsgebied van MetaSWAP wordt beschreven in Van Walsum & Veldhuizen (2011), p. 18. MetaSWAP is bedoeld als geïntegreerde component in modelketens om tijdsreeksen van grondwaterstanden op bepaalde ruimtelijke en temporele schaal te berekenen in regionale waterstudies, bv. om de effecten van veranderingen in waterverdeling te kwantificeren. De minimale ruimtelijke schaal is 5x5 m², met een schematisering van enkele lagen diep; de toetsing van de temporele schaal van MetaSWAP is gericht op grondwaterstanden op het einde van de dag, en waterbalansen gesommeerd over een dag. In principe moet MetaSWAP daarmee op dag-basis worden gebruikt.

2C. Wat zijn de toepassingen nu? Heeft het model/bestand een rol in een modelketen? Zo ja, welke?

MetaSWAP zou kunnen worden gebruikt voor regionale studies in combinatie met SIMGRO en MODFLOW, maar dit is nooit gedaan. MetaSWAP wordt eigenlijk met name gebruikt als component in het NHI (Nationaal Hydrologisch Instrumentarium, zie <http://www.nhi.nu>). Daarin heeft het de rol om de waterdynamica in de onverzadigde zone te beschrijven, terwijl 3 andere modellen de verzadigde zone (MODFLOW) en het oppervlaktewater (MOZART, DM) beschrijven; MODFLOW levert daarbij invoer aan MetaSWAP, en MetaSWAP is een grensconditie van MODFLOW en koppelt dus ook weer invoer terug. De vraag wordt 'vertaald' naar wat de verschillende toepassingen van het NHI zijn. Het doel van het NHI is het op landelijke schaal in beeld brengen van de grond- en oppervlaktewaterstroming, waarbij het minimaal de kwaliteit en functionaliteit heeft van haar voorgangers. Het NHI (en dus MetaSWAP) wordt gebruikt voor: het beantwoorden van strategische en operationele vragen van de rijksoverheid m.b.t. de landelijke waterverdeling voor consumptie (bv. voor landbouw); effecten van klimaatverandering, zeespiegelstijging, en ingrepen in het watersysteem; en als hydrologische component in het berekenen van de waterkwaliteit. Daarnaast moet het NHI in de toekomst ook door andere waterbeheerders in Nederland gebruikt kunnen worden als raamwerk voor regionale of lokale studies. Het NHI rekent typisch met tijdstappen van een dag op 250x250 m². Een voorbeeld van een regionale toepassing is de casus "Baakse Beek" (Van der Sluijs *et al*, 2012; <http://www.baaksebeek.nl>), waarin het NHI inclusief MetaSWAP gebruikt wordt met een kleinere ruimtelijke resolutie (25x25 m², laagdikte 0.2 m en totale dikte van de onverzadigde zone 2 m), en op dag-basis, met een geclusterde tijdsstap van decades, over de periode 1980-2050.

2D. Overlapt het beoogde toepassingsgebied alle daadwerkelijke toepassingen (ja/deels/nee)?

De rol van MetaSWAP binnen het NHI lijkt in overeenstemming te zijn met het doel van SWAP en dus MetaSWAP (vraag 2A), namelijk het berekenen van bovenrand-flux-dichtheden, en de algemene toepassing gericht op watervraag lijkt wel te overlappen. De temporele en ruimtelijke schaal van de toepassingen van het NHI (vraag 2C) lijken in overeenstemming te zijn met de voorgestelde beperkingen van het toepassingsgebied voor MetaSWAP (vraag 2B) voor deze algemene toepassing,

maar de ruimtelijke schaal van natuurstudies is beduidend kleiner. Toepassingen op uur-basis of een ruimtelijke schaal kleiner dan 5x5 meter zouden niet binnen het opgegeven toepassingsgebied vallen. Het is zinnig om kritisch te zijn over het temporele bereik van de toepassingen, en over het bereik van de geldigheid van kalibraties van MetaSWAP. Zo wordt bv. in de casus Baakse Beek aangenomen dat onzekerheid m.b.t. de neerslag en meteo in de periode 1980-2050 voldoende is afgedekt door de IPCC-voorspellingen t.a.v. klimaatverandering (Van der Sluijs *et al*, 2012). Het is dan nog niet direct helder wat dat betekent voor de geldigheid van berekeningen door NHI/MetaSWAP. In principe zou MetaSWAP altijd “waar” moeten zijn, omdat de hydrologische processen niet veranderen, en klimaateffecten alleen doorwerken als veranderingen in zaken als neerslag en verdamping die niet direct effect op de modelstructuur hebben. Echter, MetaSWAP moet gekalibreerd worden op SWAP t.a.v. twee parameters, nl. de dikte van de capillair beïnvloede aggregatielaag en de Boesten-parameter (meer hierover verderop), en het is mogelijk dat bepaalde klimatologische veranderingen of menselijke ingrepen een her-kalibratie van MetaSWAP vereisen binnen het NHI t.a.v. bepaalde toepassingen.

Van Walsum & Veldhuizen (2011) wijzen erop, dat gezien de gevoeligheid van andere processen voor de waterbalansterm in MetaSWAP “de nodige terughoudendheid [moet] worden betracht bij het toepassen van MetaSWAP in situaties waar bovengrondse afvoer een cruciale rol speelt. Dat geldt met name voor situaties met extreme regenval op een uitgedroogde grond” (p. 10). De afgelopen jaren zijn er periodes van droogte in Nederland geweest (voorjaar 2007: 0,3 millimeter neerslag in De Bilt; voorjaar 2011: neerslagtekort in het voorjaar circa 100 mm, zie <http://nl.wikipedia.org/wiki/Droogte>, gevolgd door behoorlijke regenval. Zou dit patroon te maken hebben met bv. klimaatverandering, dan moet er dus ook kritisch gelet worden op de betrouwbaarheid van berekeningen binnen toepassingen als de casus Baakse Beek. En dus, hoewel het wel is beoogd dat het toepassingsgebied van het NHI effecten van klimaatverandering omvat, valt te concluderen dat die toepassing van MetaSWAP kritisch beschouwd moet worden. De percolatie (het watertransport in de onverzadigde zone) wordt weinig door temperatuur, maar sterk door stoftransport beïnvloed. In MetaSWAP wordt de temperatuur ook niet meegenomen. Stoftransport wordt ook niet meegenomen, wat voor toepassingen wel problematisch zou kunnen zijn, en dat wijst op nog een extra inperking van het toepassingsgebied.

2E. Indien relevant, in hoeverre overlappen de verschillende toepassingen met elkaar (goed/matig/slecht)?

MetaSWAP wordt bijna uitsluitend gebruikt als component binnen het NHI. De vraag moet daarom eigenlijk ‘vertaald’ worden naar de vraag, of dat de verschillende toepassingen van het NHI met elkaar overlappen. Dit is van belang, omdat verderop in meerdere vragen onderscheid wordt gemaakt tussen de verschillende toepassingen, indien er onderscheid valt te maken. Het beoogde toepassingsgebied van MetaSWAP lijkt alle toepassingen wel af te dekken (2D), wat nog niet betekent dat deze toepassingen elkaar sterk overlappen. Dit laatste lijkt niet erg het geval te zijn. Zaken t.a.v. zeespiegelstijging zijn anders dan wanneer het bv. de waterkwaliteit betreft: in het eerste geval wordt vooral gekeken naar de potentiële gevolgen voor de Nederlandse duinen en dijken en de daarmee samenhangende veiligheid, in het tweede geval wordt gekeken naar concentraties van nutriënten en gifstoffen in het water en de effecten daarvan op vooral vegetatie. Dit zijn zeer verschillende toepassingen, die andere eisen kunnen stellen aan de structuur van het model, de kalibratie, de benodigde gegevens, etc. De conclusie hier lijkt te zijn dat de overlap tussen verschillende toepassingen matig is, en dat bij het beantwoorden van vragen verderop onderscheid moet worden gemaakt tussen de verschillende toepassingen.

Opmerkingen bij vraag 2:

Er wordt in de lijst gevraagd of het beoogde toepassingsgebied alle toepassingen omvat (2D). Dat wil nog niet zeggen dat die toepassingen ook elkaar (grotendeels) overlappen, het punt waarnaar gevraagd wordt in vraag 2E. Dit laatste punt is wel relevant om vast te stellen: zou blijken dat de toepassingen elkaar niet overlappen, dan kan het wellicht verderop een nuttige suggestie blijken om

voor elke toepassing een aparte modellijn te creëren. Een reden hiervoor zou kunnen zijn dat het model voor elke toepassing opnieuw gekalibreerd moet worden, en deze kalibratie veel 'kosten' met zich mee-brengt. Bovendien wordt er in een aantal vragen verderop onderscheid gemaakt tussen de verschillende toepassingen, indien er meerdere toepassingen zijn. Overlappen de verschillende toepassingen elkaar sterk, dan is het mogelijk om dit onderscheid in de praktijk te negeren. Deze aspecten - het expliciet vaststellen of er verschillende toepassingen zijn, en of ze ook afzonderlijk behandeld moeten worden of niet - komen nog niet erg nadrukkelijk in de vragenlijst naar boven. Een vraag die toe te voegen valt is: "2F. Zijn er (dus) meer dan 1/verschillende toepassingen die in vragen verderop apart behandeld/beschouwd moeten worden?"

4.4 Systeemanalyse MetaSWAP

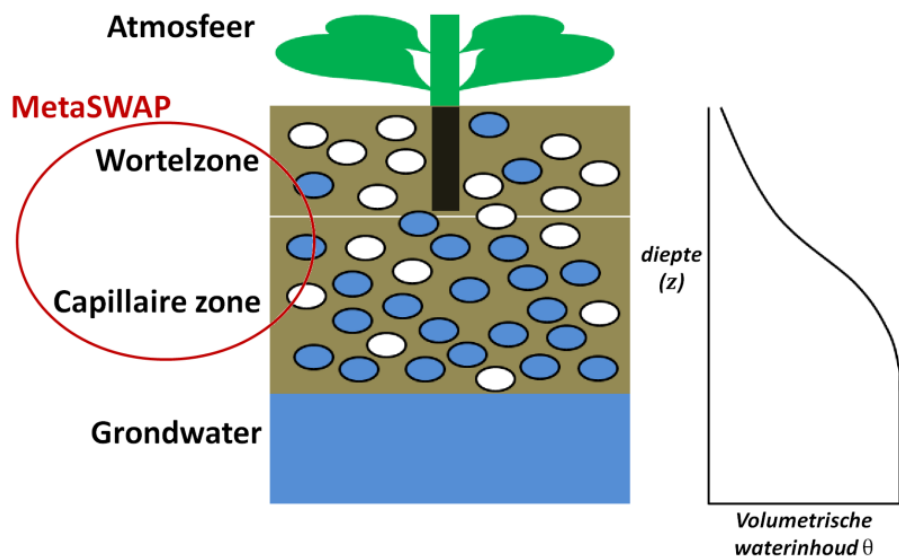
3A. Geef een systeemanalyse. Dit kan bv. een verwijzing naar een rapport zijn. Aspecten die daarin aan bod komen zouden moeten zijn: Wat zijn de belangrijkste attributen en processen van het systeem? Welke terugkoppelingen zijn aanwezig? Hoe zijn de grenzen van het systeem bepaald of gedefinieerd? Hoe wordt met relevante attributen en processen buiten deze grenzen omgegaan, bv. randvoorwaarden, constanten, of simpelweg genegeerd? Over welke processen, attributen, terugkoppelingen, etc. bestaat er onzekerheid? Hoe is die onzekerheid bepaald?

Het NHL kijkt naar de waterbalans van het oppervlaktewater per 10 dagen. De Rijn is de grootste aanvoerbron van water. Het oppervlaktewater wordt gebruikt als koel- en drinkwater, voor irrigatie, scheepvaart, om veengebieden op peil te houden, etc. Het peilbeheer in de zomer en winter is wel hetzelfde, het oppervlaktewaterpeil zelf wordt uitgerekend. Het watergebruik beïnvloed indirect ook de grondwaterstand: bv., in de zomer is er eigenlijk altijd een overschrijding van de potentiële verdamping t.o.v. neerslag, dat wordt aangevuld door grondwater. De bepaling van de berekening wordt ook bepaald uit MetaSWAP.

Grondwater in de diepe bodem staat indirect in contact met de atmosfeer via de onverzadigde zone. De onverzadigde zone is de bovenste bodemlaag, en bestaat uit grond met poriën, waarin zich lucht en water bevinden. Deze zone verdiept zich tot het punt waar het water in de poriën van de grond op atmosferische druk is, m.a.w. het grondwaterniveau. De drukhoogte (druk per specifiek gewicht, in lengte-eenheid) van het water in de onverzadigde zone is lager dan de atmosferische druk, en wordt zo gehouden door een combinatie van capillaire werking en adhesie. Neerslag voert water aan via de atmosfeer naar de onverzadigde bodem, en water verdampt naar de atmosfeer. De verdamping verloopt via de kale grond, maar ook via de vegetatie. Vegetatie heeft verder wortels in de onverzadigde zone die de capillaire werking versterken, en is daarmee belangrijk voor de opname van water uit de bodem en de verdamping ervan. De bovenkant van de onverzadigde bodem werkt feitelijk als een afvoer van water (als de neerslag genegeerd wordt), terwijl de onderkant als aanvoer werkt, omdat dit aan het grondwater gekoppeld is. In de onverzadigde bodem bestaat zo een 'profiel' waarin het vochtgehalte aan de onderkant praktisch gelijk is aan dat van de grondwaterzone, en aan de bovenkant een stuk lager door de afvoer (Figuur 4.1).

MetaSWAP beschrijft de onverzadigde zone, en fungeert als grensconditie van MODFLOW, dat de 3D grondwaterdynamiek beschrijft in de verzadigde zone. Neerslag die op de bodem valt wordt in zekere zin 'gebufferd'; de onverzadigde zone werkt als een vertraging- en afvlakkingscomponent (als ware het een koffiefilter). De onverzadigde zone wordt in MetaSWAP ingedeeld in verticale kolommen, die zijn gekoppeld aan het grondwaterniveau en aan het oppervlak, en die in de horizontale richtingen homogeen zijn. De kolommen hebben geen horizontale uitwisseling, omdat Nederland een laaglandregio is, waarin rond 85% van het oppervlak het grondwaterpeil zich binnen 2 meter van het grondoppervlak bevindt (Van Walsum & Groenendijk, 2008), waardoor kleine variaties (van, zeg, 0.2 m) al invloedrijk kunnen zijn, en er dus nauwelijks een dominante stroomrichting

aanwezig is. De kolommen communiceren daarentegen indirect met elkaar via het grondwater. Praktisch alle bodem is 'zachte' bodem met een behoorlijk hydraulisch geleidingsvermogen. Er zijn wel veel verschillende typen zachte bodems, waarvan de verschillende eigenschappen m.b.t. het adsorberen en geleiden van water sterk kunnen verschillen. De veelgemaakte vergelijking dat de Nederlandse bodem als een 'spons' fungeert is dus niet altijd correct, omdat bepaalde bodems zoals zandbodems erg star kunnen zijn; alleen veengrond kent een echte sponswerking. Er worden binnen MetaSWAP een aantal verschillende typen grond onderscheiden. De wortelzone kan in MetaSWAP worden gevarieerd, maar er wordt aangenomen dat de wortelzone kleiner is dan het grondwaterpeil. Verder is het mogelijk dat een kolom uit meerdere 'lagen' grond bestaat, bv., een kolom kan bestaan uit een laag zand bovenop een laag klei. Dit beïnvloedt uiteraard het vochtprofiel. Binnen een laag in een kolom wordt aangenomen dat de grond homogeen is.



Figuur 4.1. Een conceptueel diagram van de bodem. De grondwaterzone, waarvan hier wordt aangenomen dat het een constante verhouding water:grond heeft, staat in contact met de atmosfeer via de onverzadigde zone, die is verdeeld in een wortelzone, en een capillaire zone. De bodem bestaat uit grond en poriën, die met lucht (wit) of water (blauw) gevuld kunnen zijn. Hoe meer naar het oppervlak toe, hoe minder poriën water bevatten. Het vochtprofiel dat deze relatie geeft wordt door MetaSWAP beschreven.

Neerslag is een 'forcing' van het systeem, m.a.w., er is geen stochastisch model of procesmodel dat de variaties in neerslag genereert, maar de neerslag is een gegeven invoer. Het landgebruik beïnvloedt de waterinhoud in de onverzadigde zone, bv. door drainage of door de gebruikte gewassen. De verdamping van water uit de bodem via de kale grond en via de vegetatie wordt weliswaar expliciet meegenomen in MetaSWAP, maar de dynamica van vegetatiegroei is versimpeld; vegetatiegroei gaat volgens vastgelegde groeiprofielen en is dus niet dynamisch. Niet alle neerslag bereikt de bodem, omdat de vegetatie ook neerslag opvangt, die voor een deel weer verdampt zonder ooit op de grond terecht te komen. Diverse bodemflora en -fauna spelen verder een rol bij het verhogen of verlagen van bodemeigenschappen, bijvoorbeeld regenwormen en schimmels zijn zeer belangrijk voor de vorming van macroaggregaten in grond en humus die de waterretentie behoorlijk kunnen verhogen; humus kan tot 80-90% van het gewicht aan vocht bevatten (Olness & Archer, 2005). Ook dit wordt buiten beschouwing gelaten.

Beschouw nu per toepassing

3B. Welke systeem-analytische aspecten (attributen, processen, terugkoppelingen, etc.) zijn (direct) relevant voor de toepassing? En welke in mindere mate of niet? En hoe is die relevantie bepaald?

De enige toepassing waar naar gekeken wordt is de rol van MetaSWAP in het NHI. Een gevoeligheidsanalyse lijkt één van de beste manieren om de relatieve bijdrage en dus relevantie van onderdelen te bepalen, maar het gaat hier over de systeemanalyse, en niet over het model zelf, in welke vorm dan ook. De veruit belangrijkste toepassing van MetaSWAP is de rol die het speelt in het NHI. Het NHI rekent met tijdstappen van een dag en een ruimtelijke schaal van 250 bij 250 meter, terwijl de beschikbaarheid van oppervlaktewater wordt berekend met tijdstappen van 10 dagen. De bedoeling is om strategische en operationele vragen te beantwoorden op vooral landelijke schaal. Dat betekent dat de meegenomen processen en attributen vooral op een tijdschaal van 1-10 dagen en een ruimtelijke schaal van 250 bij 250 meter spelen.

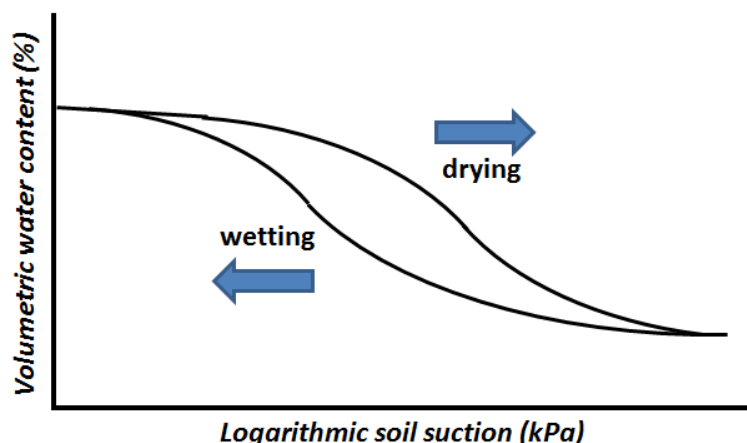
Het platte landschap en de ruimtelijke schaal van 250 bij 250 meter suggereren dat het zeer redelijk is om de horizontale uitwisseling van water tussen kolommen te negeren. De fluctuaties in het grondwater als gevolg van de dynamica in de onverzadigde zone zijn zeer relevant voor de toepassing, en MetaSWAP is ook precies bedoeld om de veranderingen in het grondwaterpeil te simuleren, en daarnaast de overdracht van water van bodem naar de oppervlakte. Vanwege de toepassing is het feitelijk niet van belang om een precieze en correcte beschrijving van de dynamica van het water in de onverzadigde zone te hebben, zolang de voorspellingen over de beschikbaarheid van grond- en oppervlaktewater op de tijd- en ruimtelijke schaal van de toepassing binnen geaccepteerde marges maar kloppen. Het is belangrijker om de grenscondities goed te hebben, d.w.z. de fluctuaties in het gemiddelde grondwaterpeil binnen eenheden van 250 bij 250 meter op een tijdschaal van een dag moeten redelijk kloppen, en daarmee de neerslag, de verdamping, en de hoeveelheden oppervlaktewater. De neerslag en verdamping kunnen berekend of gemeten worden, en in beide gevallen moet er wel gekeken worden naar zaken als afwatering en verschillen in landgebruik (het oppervlak in een cel van 205 bij 250 meter kan bv. worden gedomineerd door bebouwing, door een grasveld, etc.).

3C. Zijn alle voor de toepassing belangrijke aspecten als gegeven in B meegenomen in het model/bestand? En welke niet?

De in 3B genoemde relevante processen zijn meegenomen in de huidige versie van MetaSWAP. Van Walsum & Veldhuizen (2011), p. 19, geven een aantal beperkingen van MetaSWAP die relevant kunnen zijn.

Op de tijdschaal van 1-10 dagen van het NHI lijken macroporiën erg belangrijk te zijn, vooral voor veen- en kleigebieden, maar deze ontbreken nu nog. Hellingprocessen worden niet meegenomen, wat de toepasbaarheid van MetaSWAP/NHI voor met name Zuid-Limburg beperkt; dat gebied is dan ook niet meegenomen in de huidige versie van het NHI v2.2 (www.nhi.nu, d.d. 26 juni 2012). Processen die actief kunnen zijn bij infiltratie in een droge grond worden ook genegeerd; het model veronderstelt alleen het ontstaan van een infiltratiebeperking door verzadigde doorlaatbaarheid of verharding. Ook processen m.b.t. vriezen en dooien ontbreken, waardoor de waterbeschikbaarheid in de winter niet correct berekend kan worden.

Hysteresis ontbreekt ook nog, en zou op de tijdschaal van de toepassing misschien ook een rol van belang kunnen spelen. Verschillende bodems kennen een sterke 'hysteresis' (Figuur 4.2): het voorkomen van twee verschillende relaties tussen de zuigkracht van de bodem en de waterinhoud per volume (retentiecurves), afhankelijk van of dat de bodem aan het opdrogen is of juist natter wordt (o.a. Pham *et al*, 2005).



Figuur 4.2. Conceptueel diagram van 'hysteresis' in de bodem, naar Pham *et al* (2005). De relatie tussen de zuigkracht van de bodem en de volume-waterinhoud blijkt te verschillen, afhankelijk van of de bodem uitgedroogd wordt of er juist water wordt toegevoegd.

Hoewel hysteresis in de bodem als essentieel wordt beschouwd in de formatie en het terugkomen van preferente stromingen (Ritsema *et al*, 2000), wordt dit proces in MetaSWAP niet meegenomen (Van Walsum & Veldhuizen, 2011, p. 10). Het concept van MetaSWAP beperkt zich tot de Richards' vergelijking (zie verderop). Bovendien zijn gegevens om retentiecurves te kwantificeren schaars (Gebrenegus & Ghezzehei, 2011), waardoor er significante onzekerheid is over de kwantitatieve rol van hysteresis, en wordt aangenomen dat op de langere tijdschaal hysteresis geen rol van betekenis speelt.

3D. Zijn er minder of niet belangrijke aspecten meegenomen in het model/bestand (ja/weinig/geen)?

MetaSWAP is ontwikkeld om SWAP te vervangen in rekenintensieve toepassingen, maar hoeft voor aparte toepassingen, zoals bv. de rol in het NHI, niet *per se* de volledige functionaliteit van SWAP te dekken. Het is mogelijk dat aspecten van SWAP wel in MetaSWAP zijn opgenomen maar die niet voor het NHI van belang zijn. Ook uit een gevoeligheidsanalyse zou kunnen blijken dat bepaalde zaken geen belangrijke bijdrage leveren. Zoals al opgemerkt hoeft feitelijk het vochtprofiel niet goed bekend te zijn, zolang de fluctuaties in de grondwaterstand en het oppervlaktewater maar afdoende goed worden gesimuleerd op de gewenste aggregatieschalen. Hetzelfde geldt in zekere zin ook voor de typering van de grondtypes: voor de toepassing is het niet nodig om precies te weten wat voor type grond een eenheid bevat, zolang de invloed van de grond op de waterbeschikbaarheid maar afdoende klopt. Dit punt en andere punten worden uitvoeriger besproken verderop.

3E. Beoordeel de mate van 'evenwicht' m.b.t. de systeemanalyse, gebaseerd op de antwoorden op 3C en 3D (goed/matig/slecht).

Er ontbreken nog wat zaken (zoals macroporiën) in MetaSWAP, terwijl wellicht andere zaken al wat te complex zijn met het oog op de toepassing. Voor een goede evaluatie in het kader van evenwicht is het echter nodig MetaSWAP te vergelijken met modellen met een vergelijkbare rol. Naast een vergelijking met SWAP, valt er een zinnige vergelijking te maken tussen MetaSWAP en veel andere modellen voor de onverzadigde zone (Van Walsum & Groenendijk, 2008). Deze andere modellen geven meestal een indeling (schematisering) van de onverzadigde zone in één of twee lagen – iets waarin ze niet principieel verschillen van MetaSWAP. Echter, Van Walsum & Groenendijk (2008) noemen drie punten die als serieuze tekortkomingen worden ervaren:

1. Binnen die lagen vallen variaties in vochtgehalte weg door aggregatie ('lumping');
2. De capillaire stijging wordt teveel gesimplificeerd, bv. door het als een extinctie-functie te benaderen;

3. De bergingscapaciteit wordt constant verondersteld, hetgeen echter tot serieuze onbetrouwbaarheid in de simulaties leidt.

Deze punten wekken de suggestie dat het noodzakelijk is om een meer gedetailleerd vochtprofiel in de onverzadigde zone te simuleren om voldoende betrouwbare voorspellingen m.b.t. de fluctuaties in grondwaterspiegel gekoppeld aan de beschikbaarheid van oppervlaktewater te krijgen. M.a.w., een eenvoudige relatie tussen invoer en uitvoer lijkt die mate van betrouwbaarheid niet te kunnen reproduceren.

4.5 Conceptueel model MetaSWAP

4A. Geef het conceptuele model (voor een model) of ontwerpmodel (voor een bestand). Dit kan bv. een verwijzing naar een rapport zijn. Het conceptuele model legt op relatief informele wijze de relaties tussen componenten vast. Aspecten die aan bod kunnen komen zijn: Welke aannamen en vereenvoudigingen zijn er gemaakt? Wat zijn de ruimtelijke dimensies? Welke ruimtelijk-temporele aggregatieniveaus worden gebruikt? Welke zaken spelen een rol voor het bestand? Wat wordt er voor (des)aggregatie van data gebruikt? Voor modellen: Welk modeltype is gebruikt?

Het conceptuele model van MetaSWAP wordt beschreven door Van Walsum & Groenendijk (2008); Van Walsum *et al* (2010); Van Walsum & Veldhuizen (2011), §2.1, en de referenties daarin.

De beschrijving van grondwaterdynamica is gebaseerd op de fysische beschrijving van stromingen. Het formele model (Van Walsum & Groenendijk, 2008, en zie verderop) is gebaseerd op de bekende Richards' vergelijking (Richards, 1931), wat gebaseerd is op Darcy's Wet, wat weer gebaseerd is op de Navier-Stokes-vergelijkingen. Richards' vergelijking kan in verschillende vormen worden gegeven, maar één vorm is als een differentiaalvergelijking waarin het volumetrisch watergehalte (volume water gedeeld door volume totaal van een volume grond uit de onverzadigde zone) afhangt van de tijd. Het oplossen van de Richards' vergelijking is rekenkundig veeleisend, en wordt gedaan o.a. door het model SWAP (Van Dam *et al*, 2008). De oplossing van de differentiaalvergelijking hangt verder af van een aantal grenscondities, waaronder de hoogte van het grondwaterniveau. Onder gelijkblijvende omstandigheden is het mogelijk dat Richards' vergelijking een 'steady state' geeft, m.a.w., na voldoende tijd ontwikkelt zich een niet langer veranderende situatie; meestal wordt 10 dagen aangenomen als voldoende tijd hiervoor. Een 'steady state' kan berekend worden door de differentiaalvergelijking gelijk aan nul te stellen.

De zwaarte van de berekeningen is een serieuze beperking. De idee achter MetaSWAP is dat Richards' vergelijking niet direct wordt opgelost, maar wordt vervangen door twee gewone differentiaalvergelijkingen: één voor de variaties in de verticale kolom, die gebruik maakt van de *steady state*-vorm van de stroomvergelijking, en één voor de variaties in de tijd, die gebruik maakt van een waterbalans over de gehele kolom. Bovendien worden zoveel mogelijk berekeningen in een pre- en post-processing-stap gedaan. Voor elk bodemtype en elke wortelzonediepte wordt SWAP gedraaid voor een serie van potentiële fluxen voor de wortelzone en voor een serie van grondwaterstanden. De *steady state*-waarden worden, samen met de bijbehorende geleidingsparameters, grondwaterpeil, potentiële flux en potentiële opnamewaarde in de wortelzone, opgenomen in een database. Elk vochtprofiel (zoals in Figuur 2.1; bij een *steady state*) heeft dus een unieke combinatie van drukhoogte en grondwaterdiepte, terwijl de flux niet als onafhankelijke variabele wordt gebruikt.

Simulaties bestaan uit het continu doorlopen van twee stappen. Bij de *eerste stap* wordt het grondwaterniveau – de 'onderkant', die is gegeven door MODFLOW, het grondwatermodel dat de stroming in de verzadigde zone beschrijft – constant gehouden. De verandering in de opnamewaarde in de wortelzone leidt tot een verandering in de flux, maar de nieuwe flux weten we niet. Het vochtprofiel passende bij de grondwaterstand en opnamewaarde wordt uit de database afgelezen. Uit het vochtprofiel weten we ook de hoeveelheid water in de onverzadigde bodem. De infiltratie in de

bodem, de verdamping, en de actuele wortelopname als functie van de diepte worden ook bekend verondersteld (door metingen of anderszijds), en bepalen de *recharge*. Er wordt aangenomen dat het water instantaan over de wortelzone wordt verdeeld, en met een aanname over hoe de flux verandert gedurende de simulatietijdstap kan een waarde bepaald worden, waar het nieuwe, unieke vochtprofiel bij gezocht kan worden, en dus ook de nieuwe flux. Bij de *tweede stap* wordt dit nieuwe vochtprofiel gebruikt om de nieuwe waterinhoud aan het oppervlak te bepalen (de 'bovenkant'), die vervolgens wordt vastgehouden terwijl het grondwaterpeil als vrije variabele wordt gebruikt. M.b.v. een waterbalans wordt een tweede nieuw vochtprofiel bepaald, waarbij de nieuwe grond-waterstand wordt afgeleid. Afhankelijk van de wortelzonedikte wordt de onverzadigde zone als één groot compartiment behandeld, of worden de wortelzone en de capillaire zone daaronder apart behandeld. In het laatste geval worden ook twee profielen bepaald, die niet op elkaar hoeven aan te sluiten, omdat de capillaire zone "achterloopt" op de wortelzone wanneer de wortelzone dik is, m.a.w., de aanname is dat het water zich niet instantaan over de gehele onverzadigde zone verspreidt in dat geval.

T.o.v. SWAP zijn er een paar conceptuele wijzigingen in MetaSWAP (Van Walsum & Veldhuizen, 2011, p. 15-16 en p. 39):

- De Richards' vergelijking wordt niet direct opgelost (veel meer daarover bij de numerieke implementatie);
- De modellering van de plant-atmosfeerinteracties is zoals in SWAP, behalve wat betreft de interceptieverdamping. In SWAP wordt aangenomen dat interceptiewater binnen dezelfde dag verdampt, een aanname die niet erg realistisch is voor de winterse neerslag in Nederland. In MetaSWAP wordt een interceptiereservoir aangenomen, dat niet *per se* leeg hoeft te zijn binnen een dag;
- De verdamping bij 100% gewasbedekking is in MetaSWAP aangepast, omdat in SWAP de gewasverdamping verdwijnt indien gewassen onder water staan;
- Het vochtgehalte in het bovenste compartiment voor de kale-grondverdamping is bij MetaSWAP niet op hetzelfde schaalniveau beschikbaar als bij SWAP. Om tot dezelfde gemiddelde waarden te komen op de lange tijdschaal moet de zgn. 'Boesten-parameter' in MetaSWAP verlaagd worden t.o.v. de standaardwaarde van deze parameter. Dit geeft een parameter die gekalibreerd moet worden (zie verderop);
- MetaSWAP leunt op de aanname dat de drukhoogte in de zone beneden de wortelzone volledig wordt bepaald door de situatie in de wortelzone, in het geval er sprake is van capillaire opstijging. De dikte van deze capillaire zone hangt af van de hydraulische geleiding van de grond, en moet 'geijkt' worden m.b.v. SWAP, wat dus een tweede parameter geeft die gekalibreerd moet worden (zie verderop).

Beschouw nu per toepassing

4B. Op welke (ruimtelijk-temporeel) aggregatieniveaus zijn antwoorden gewenst, gezien de toepassing? Welke ruimtelijke dimensies en welk modeltype/ontwerpmodel zijn gewenst?

Voor het NHI zijn antwoorden gewenst op een tijdschaal van 10 dagen, en een tijdspanne van ca. 2050. De ruimtelijke eenheden hebben een grootte van 250 bij 250 meter, wat 550.000 eenheden levert voor het Nederland. Binnen die eenheden worden de horizontale ruimtelijke dimensies homogeen verondersteld. Een model moet dus voldoende betrouwbare temporele fluctuaties opleveren in de beschikbaarheid van oppervlaktewater en in het grondwaterpeil op die ruimtelijke en temporele schalen. Het verticale vochtprofiel zou in principe niet bekend hoeven te zijn, maar de bezwaren van Van Walsum & Groenendijk (2008) suggereren dat hier toch naar gekeken dient te worden. Het model moet dus een zeker vochtprofiel gebruiken om het grondwaterpeil en de beschikbare hoeveelheid water te bepalen. Echter, het expliciet oplossen van de Richards' vergelijking is in principe niet nodig, indien een andere methode wordt gebruikt om het vochtprofiel te bepalen (wat het geval is bij MetaSWAP).

4C. Komen de ruimtelijke dimensies, aggregatieniveaus, etc. overeen met die vanuit de toepassing zijn gewenst (ja/matig/niet)?

Het grondwaterpeil wordt dynamisch verondersteld in MetaSWAP, wat relevant voor de toepassing is. De dynamica inclusief de koppeling met MODFLOW speelt op een tijdstap van een dag, en resultaten worden geaggregeerd met een tijdstap van 10 dagen. De rekenschaal van de disaggregatie is in orde voor het NHL, maar nog te grof voor bv. natuurdoelen (25m²). MetaSWAP wordt echter nog niet gebruikt voor dergelijke regionale toepassingen, al is dat wel beoogd.

4.6 Gebruikte data en dataeisen bij MetaSWAP

5A. Welke gegevens zijn benodigd, gezien de toepassing, en motiveer? Wees specifiek m.b.t. resolutie, nauwkeurigheid, schaal, etc., en vermijd algemene en ongemotiveerde opmerkingen als "file X moet gebruikt worden".

Het NHL werkt met een ruimtelijke schaal van 250 bij 250 meter, en een tijdschaal van een dag. De gegevens zouden idealiter op deze ruimtelijke en temporele schaal moeten zijn. De gewenste gegevens voor invoer, kalibratie of validatie zijn grondwaterstanden, gegevens over afvoer van oppervlaktewater, neerslag en verdamping. Voor de verdamping en interceptie van neerslag is het van belang om voldoende gegevens te hebben over de dynamica in vegetatiegroei, evt. alleen als fluctuerende bovengrens-conditie. Verder is het nodig om de typering van de grond binnen eenheden te weten, het landgebruik binnen diezelfde eenheden, en de gemiddelde diepte van de wortel-zones. Idealiter zouden ook gegevens over fluxen beschikbaar zijn, of over vocht-profielen.

Opmerkingen bij vraag 5, specifiek 5A:

Er wordt in vraag 5A geen onderscheid gemaakt tussen waarvoor de gegevens gebruikt worden, bv. voor 'forcing', kalibratie, of validatie. Bij MetaSWAP is dit overigens een lastig punt, omdat MetaSWAP eigenlijk bijna alleen maar indirecte invoer gebruikt: MetaSWAP krijgt bijna alle invoer van SWAP, en wordt ook op basis van SWAP getest, gekalibreerd en gevalideerd.

5B. Welke gegevens zijn beschikbaar, en in welke mate?

Afvoermetingen zouden metingen kunnen geven op de gewenste aggregatieschaal. Ze worden nu op twee manieren gemeten: 1/ de hoogte van de waterstand bij stuwen, of 2/ het pompvolume per tijdseenheid bij gemalen. Dit zijn echter lastige metingen. Van Genuchten-parameters beschrijven de doorlatendheid van een kolom, maar niet van een oppervlak. De bodemeigenschappen worden typisch afgeleid uit laboratorium-proeven met bodemmonsters van 10 bij 10 cm. Data met een ruimtelijke schaal van 250 bij 250 m moeten dus afgeleid worden uit opschalingen. Dit leidt tot onzekerheden (zie verder vraag 9). Piezometers worden gebruikt voor metingen van de waterhoogte. Hoewel de kwaliteit van piezometers enorm is toegenomen, is een fundamenteel probleem dat het puntmetingen zijn en geen geaggregeerde metingen over grotere vlakken. De neerslag is beschikbaar via weerstations, maar de verdamping wordt nu voor een goed deel geschat o.b.v. vegetatiegroei, die weer opgelegd wordt als vaste groeipatronen. De meeste data wordt indirect gebruikt, omdat het eigenlijk in SWAP toegepast wordt, en SWAP vervolgens wordt gebruikt om de database met vochtprofielen aan te leggen (zie 5C en verderop).

5C. Wat is uiteindelijk de invoer van het model of bestand?

Omdat MetaSWAP een meta-model is, is dit een ietwat lastig punt. De uitvoer van SWAP is de enige directe invoer van MetaSWAP, omdat deze uitvoer is gebruikt bij het opstellen van de tabellen met de vochtprofielen en bijbehorende gegevens over de *steady states* (zie verderop). Verder worden er in MetaSWAP nog twee zaken gekalibreerd, maar dat gebeurt ook met SWAP-uitvoer. Men kan hier dus betogen dat de uiteindelijke invoer simpelweg de invoer uit SWAP is.

5D. Komt de invoer overeen met de door de toepassing gewenste gegevens (ja/deels/nee)? Vergelijk de antwoorden op 5A, 5B en 5C. Let hierbij sterk op eenheden

en de drie elementen van schaal (coverage, support, extent; zie Bierkens et al, 2000; Bogaart et al, 2011). Heeft de data de juiste dimensies en eenheden? Zo nee, is er een (des)aggregatiemethode voorhanden? Schaalproblemen doen zich overal voor door de 'vertaling' van attributen naar metingen, die gediscretiseerd zijn in tijd en ruimte, en onderhevig zijn aan allerlei bronnen van meetruis.

Omdat MetaSWAP een meta-model van SWAP is, zou de invoer overeenkomen met wat gewenst is, indien dit voor SWAP het geval is. Men kan in dit geval betogen dat de invoer dus volledig overeenkomt met wat gewenst is door de toepassing. De vraag wordt dus eigenlijk 'vertaald' naar of dat SWAP de juiste invoer gebruikt en kan gebruiken in het kader van het NHI.

Voor veel van de data is dit toch ietwat twijfelachtig. Gegevens over bodemverhoging zijn op 5 bij 5 m², en over landgebruik op 25 bij 25 m² beschikbaar (Van Walsum & Groenendijk, 2008). De neerslag is landelijk gezien wel goed gedekt, maar kan in regionale studies (nu nog niet gedaan) een probleem worden. De verdamping is een groter probleem (komt later nog aan de orde). De gewasgegevens van 19 gewassen zijn gebaseerd op veldproeven van voor 1987 (Feddes, 1987; gerefereerd in Van Walsum & Veldhuizen, 2011, hoofdstuk 6), en zijn wellicht dus verouderd. Gegevens over hydrologische eigenschappen zijn ook beperkt. De PAWN-schematisering (meer verderop) beperkt zich tot 21 soorten bodem, terwijl de bodemeigenschappen op laboratoriumschaal zijn bepaald (zoals boven besproken). Bij een *pilot* is bv. naar boven gekomen, dat twee aangrenzende cellen met een gelijke PAWN-indeling toch behoorlijke verschillen bleken te vertonen na metingen.

Een punt dat nu mist in de vragenlijst is welke alternatieven er mogelijk zijn. Satelliet-metingen zouden bv. gebruikt kunnen worden om meer nauwkeurige data te verkrijgen m.b.t. verdamping, en kunnen prima op de ruimtelijke schaal van 250 bij 250 m² verkregen worden (voor diverse bestanden worden inmiddels al fijnere resoluties aangehouden), en vermoedelijk ook op de juiste temporele schaal van 10 dagen. De satellietmetingen betreffen temperaturen aan het oppervlak, die een betere inschatting van de verdamping en van bodemvochtgehalten behelzen.

Opmerkingen bij vraag 5:

Er wordt nog niet expliciet gevraagd naar of de uitvoer past bij wat gewenst is, al wordt dat wellicht impliciet gedaan in vraag 4B/C, maar dan is het formele model noch de numerieke implementatie noch de data al behandeld. De volgorde van de lijst klopt op dit punt dan ook niet.

4.7 Formeel model van MetaSWAP

Gebruik de antwoorden op vragen 5A tot 5D om vragen 6B tot 6D te antwoorden.

6A. Geef het formeel model, en motiveer de keuze of verwijs naar paginanummers uit een gepubliceerd rapport of artikel. Het formeel model is de omzetting van concept naar bijvoorbeeld wiskundige vergelijkingen of een set rekenregels. De keuze van formeel model hangt af van de beoogde toepassing en de data, en wordt mede bepaald door de gestelde onzekerheidsmarges.

MetaSWAP omvat een aantal wiskundige vergelijkingen, die kunnen worden gevonden in Van Walsum & Groenendijk (2008); Van Walsum *et al* (2010); Van Walsum & Veldhuizen (2011). Het is weinig zinvol deze vergelijkingen hier te reproduceren. In plaats daarvan wordt hier verwezen naar de beschrijving van de numerieke implementatie in 7A. Wel is het van belang om te realiseren dat het formele model iets anders is dan de implementatie, zie Van Walsum & Groenendijk (2008) voor de details.

De basis van het formele model is de beschrijving van de waterstroomdynamica m.b.v. de Richards' vergelijking (Richards, 1931):

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[K(\theta) \left(\frac{\partial \psi}{\partial z} + 1 \right) \right]$$

Met K hydraulische geleiding [$L T^{-1}$], ψ drukhoogte [L], z [L] de diepte vanaf het oppervlak, en θ [$L^3 L^{-3}$] de hoeveelheid water per volume grond (bv. 0 is volledig droog). Achter deze vergelijkingen zitten meerdere aannames, waaronder

- De grondwaterstroomsnelheid is voldoende langzaam (Reynolds' getal < 1);
- De aquifer (watervoerende laag in de grond) is niet samendrukbaar, en de 'matrix' (fijnkorrelig materiaal) verandert niet gedurende het (numerieke) experiment;
- Water is niet samendrukbaar, zodat massa gelijk is aan dichtheid keer volume;
- De hydraulische geleiding is ruimtelijk uniform en isotroop (de materiaaleigenschappen hangen niet van de richting af);
- Externe belading op de aquifer (bv. atmosferische druk) is constant;
- De aquifer lekt niet.

Het is van belang om op te merken dat de Richards' vergelijking een driedimensionale partiële differentiaalvergelijking is, die in MetaSWAP benaderd wordt door het te vervangen door twee vergelijkingen (zie daarvoor Van Walsum & Groenendijk, 2008), en niet expliciet wordt opgelost door het creatief gebruik van een database (zie verder de implementatie). Daarnaast moet een waterbalans worden opgelost, waarin de hoeveelheid water in het vochtprofiel (de integraal van θ over z) wordt gesommeerd met wat er verder aan water is bijgekomen (de *recharge* R [$L T^{-1}$]) in een bepaalde tijdstap Δt [T], dat daarna weer herverdeeld wordt over wat er in de kolom achterblijft en wat er afgevoerd wordt.

Opmerkingen bij vraag 6A:

Er wordt bij de systeemanalyse (in vraag 3) gevraagd naar de aannames achter de analyse. Bij het formele model worden ook aannames en keuzes gemaakt, maar daar wordt nu niet expliciet naar gevraagd.

6B. Welke gegevens zijn benodigd voor dit formele model?

MetaSWAP gebruikt een database die gemaakt is m.b.v. SWAP (zie verderop), en heeft dus geen gegevens nodig voor vochtprofielen. MetaSWAP heeft wel het grondwaterpeil nodig, en totalen van de verzadigde stroming. Verder moeten de vocht karakteristieken worden bepaald per eenheid, en zijn uiteraard de gegevens m.b.t. verdamping, etc. benodigd.

6C. Welke gegevens zijn beschikbaar, en in welke mate?

De vocht karakteristieken zijn gegeven door de PAWN-schematisering (zie Tabel 2 in Van Walsum & Groenendijk, 2008; Van Walsum *et al*, 2010, en verdere referenties daarin). De vochtprofielen worden bepaald door het draaien van SWAP in de *pre-processing*, en die gegevens vervolgens in een database op te slaan (zie verderop). De meeste gegevens zijn overdraagbaar van SWAP naar MetaSWAP, en daarmee meestal voldoende beschikbaar. Het grondwaterpeil en de gegevens over de verzadigde water-stroming volgen uit MODFLOW. Echter, daadwerkelijke metingen van grondwater-standen zijn erg beperkt. Veel gegevens van waterbuizen zijn van onvoldoende kwaliteit (mededelingen BIS-symposium, 9 februari 2012); slechts de data van zo'n 100 waterbuizen verspreid over delen van Nederland zijn voldoende betrouwbaar gebleken. Daarmee is de dataondersteuning voor zowel waterkaarten als simulatiemodellen m.b.t. waterstroming erg zwak. In de interviews voor MetaSWAP is dit ook als zwak punt naar boven gekomen. Verder gelden natuurlijk de opmerkingen m.b.t. verdamping en de gewasgroei die zijn genoemd in 5D.

6D. Zijn de gegevens die benodigd én beschikbaar zijn ook gebruikt? Waarom wel/niet?

Alle gegevens die beschikbaar zijn lijken ook 'aangegrepen' te zijn. Er zijn namelijk geen financiële beperkingen of beperkingen in de sfeer van privacy. De ideeën m.b.t. het schatten van verdamping o.b.v. satellietdata is nog redelijk nieuw en niet uitgewerkt.

4.8 Numerieke implementatie MetaSWAP

7A. Geef het numeriek rekenmodel/gegevensbestand (de implementatie in software), en motiveer de keuze of verwijst naar paginanummers uit een gepubliceerd rapport of artikel.

Deze vraag moet breder geïnterpreteerd worden dan dat ie nu gesteld is. Hier hebben we het over de numerieke implementatie van het model, inclusief maar niet beperkt tot de vertaling van de wiskundige vergelijkingen naar code. In het geval van MetaSWAP omvat de numerieke implementatie ook een aantal zaken m.b.t. de *pre*- en *post-processing*. Een kort overzicht van de implementatie is te vinden in het stuk "*Model Implementation*" in Van Walsum & Groenendijk (2008).

Zoveel mogelijk rekenwerk wordt buiten het *online* rekenschema geplaatst. In de *pre-processing* stap wordt een database aangemaakt van alle *steady states* voor elk grond-type en mogelijke wortelzonediepte voor een reeks van potentiële fluxen voor de wortelzone en een reeks van grondwaterpeilen (zie Appendix 1 in Van Walsum *et al*, 2010, voor meer details). Elk *steady state* profiel heeft een unieke combinatie van drukhoogte en grondwaterpeil, en deze beide variabelen doen dienst als opzoekvariabelen voor de database. De flux wordt dus niet gebruikt als opzoekvariabele, maar volgt juist, samen met de waterinhoud in de wortelzone en de capillaire zone, uit de database na het invullen van de drukhoogte en het grondwaterpeil. In de *post-processing* worden gedetailleerde profielen voor drukhoogte en vochtinhoud geconstrueerd.

In de online berekeningen wordt de koppeling gemaakt met MODFLOW. De rekenkern van MetaSWAP is niet het oplossen van de Richards' vergelijking, maar het oplossen van een expliciete waterbalans. In de *eerste stap* wordt een *recharge* berekend die een functie is van infiltratie en verdamping aan het oppervlak, en een integraal van de wateropname door de wortels over de wortelzonediepte (zie verg. 3 in Van Walsum & Groenendijk, 2008). Verder wordt het vochtprofiel bepaald o.b.v. de grondwaterstand en de drukhoogte, en het vochtgehalte θ in de wortelzone wordt uitgerekend a.d.h.v. het geselecteerde vochtprofiel (verg. 2 in Van Walsum & Groenendijk, 2008). Het vochtgehalte wordt met de *recharge* over de tijdstap Δt gesommeerd en geeft daarmee het nieuwe vochtgehalte (verg. 4 of 7 in Van Walsum & Groenendijk, 2008, afhankelijk van het grondwaterpeil en de flux in de wortelzone, zie het schema in fig. 6 in dezelfde publicatie). Bij dit nieuwe vochtgehalte wordt het nieuwe vochtprofiel gezocht door het 'oplossen' van verg. 13 (Van Walsum & Groenendijk, 2008) d.m.v. inverse interpolatie in de database, waarin de nieuwe drukhoogte wordt bepaald, terwijl is aangenomen dat het grondwaterpeil gelijk blijft. Deze interpolatie levert ook de nieuwe flux, waarmee verg. 8 (Van Walsum & Groenendijk, 2008) kan worden opgelost, voor het bepalen van het watergehalte in de capillaire zone.

In de *tweede stap* wordt de informatie over de wijziging in grondwater *recharge* en de opslageigenschappen wordt doorgegeven aan MODFLOW, en MODFLOW retourneert wijzigingen in het grondwaterpeil en de totalen van de verzadigde stroming. De tweede stap behelst het oplossen van verg. 14 (Van Walsum & Groenendijk, 2008). Wederom is het nieuwe watergehalte onbekend, en wordt de database gebruikt om het nieuwe vochtprofiel en dus het nieuwe watergehalte te bepalen. Er blijven dan twee onbekenden over, namelijk de verzadigde stroom naar de kolom in de onverzadigde zone, en (dus) het nieuwe grondwaterpeil. Een detail van de numerieke implementatie is dat de meeste grondwatermodellen niet kunnen omgaan met de niet-lineaire relatie voor de coëfficiënt van de grondwateropslag, en dat hiervoor dus een iteratieve cyclus moet worden gecreëerd (Van Walsum & Groenendijk, 2008, p. 776). De implementatie in MetaSWAP daarvan levert ook de missende variabelen, waarna ook de uiteindelijke nieuwe drukhoogte kan worden bepaald. Een numeriek voorbeeld van de implementatie is gegeven in Van Walsum & Groenendijk (2008).

7B. Beschrijf de verificatie of verwijs naar paginanummers uit een gepubliceerd rapport of artikel. Verificatie moet niet verward worden met validatie. De eerste controleert de omzetting van formeel model naar code, de tweede toetst de resultaten van de code aan de werkelijkheid.

De vraag m.b.t. verificatie valt samen met vraag 17 van het Status A-formulier. MetaSWAP valt niet in zijn geheel te verifiëren, omdat het in de kern een “grote truc” omvat om zware berekeningen te omzeilen door gebruik te maken van een database. De implementatie van de gebruikte vergelijkingen valt uiteraard wel te verifiëren, en de gehele implementatie kan wel getest worden. Er ontbreekt hier echter een vraag naar testen, maar die bespreken we hier wel. De verificatie en testen van MetaSWAP zijn beschreven in Van Walsum & Veldhuizen (2011), p. 8 en hoofdstuk 5. Er is gebruik gemaakt van een *code checker*, en van waterbalansen per tijdstap per kolom en van het totale systeem. Verder zijn er stationaire berekeningen gemaakt, waarbij de bovenrand-voorwaarde constant werd gehouden, en er zijn berekeningstests gedaan. Er treedt waarschijnlijk enige numerieke ruis op door het interpoleren in de database, waardoor er kleine afwijkingen (ca. 5%) kunnen optreden tussen uitvoer van SWAP en MetaSWAP, maar verder zijn er geen resultaten gevonden die als problematisch zijn ervaren.

7C. Beschrijf en motiveer de keuze voor numerieke rekenmethoden, en geef de gebruikte discretisatie of verwijs naar paginanummers uit een gepubliceerd rapport of artikel. Veelal bestaat er bij numerieke rekenmethoden een afweging tussen rekensnelheid en –nauwkeurigheid.

De Richards' vergelijking is niet expliciet op te lossen en moet dus numeriek benaderd worden. De hele motivatie achter de ontwikkeling van MetaSWAP is dat dit numeriek benaderen van de Richards' vergelijking zwaar op computerkracht drukt. Er zijn diverse schema's opgesteld voor het numeriek benaderen van de Richards' vergelijking voor de beschrijving van de waterdynamiek in de onverzadigde zone (waaronder die in SWAP). De idee achter MetaSWAP is echter door het over een andere boeg te gooien en gebruik te maken van een database. Dit levert zeer veel tijdswinst op, met de prijs dat de meeste 'oplossingen' niet exact zijn, en dat er numerieke ruis optreedt door interpoleren in de database. Deze ruis (ca. 5 %) lijkt acceptabel binnen de context.

Verder is het *online* schema expliciet, en een bekend probleem met expliciete numerieke schema's voor het oplossen van differentiaalvergelijkingen is de gevoeligheid voor de grootte van de tijdstap (bv. de bekende Euler-methode). Daarom is in de implementatie ook een 'beveiliging' ingebouwd. De invloed van de tijdstap op de rekenresultaten is verder beschreven in Van Walsum & Veldhuizen (2011). Zij melden dat “uit de tests ... naar voren [komt] dat voor simulaties met als doel afvoerstatistieken af te leiden het nodig is om een tijdstap van 0.5 dag of kleiner te gebruiken. Indien daaraan wordt voldaan, dan blijkt de naar oppervlaktefractie (van het NHI) gewogen maximale afvoer voor 1966 door MetaSWAP slechts 5% af te wijken van de door SWAP berekende waarde” (Van Walsum & Veldhuizen, 2011, p. 10). Verder is het relevant, dat door de expliciete tijdstap de database ook afhankelijk is van de grootte van de tijdstap, m.a.w., voor elke tijdstapgrootte moet een aparte database gehanteerd worden.

Het meta-concept is nog steeds redelijk geheugenintensief (Van Walsum & Veldhuizen, 2011, p. 23-24). Echter, MetaSWAP is niet langzamer dan MODFLOW, en vergelijkingen tussen MetaSWAP en SWAP laten een factor 100 versnelling zien door de aanpak van MetaSWAP (het ophalen van oplossingen i.p.v. het iteratief berekenen) in combinatie met de koppelingsstrategie aan MODFLOW. In dit aspect lijkt MetaSWAP toch redelijk geoptimaliseerd te zijn, met de twee kanttekeningen dat 1/ er wellicht methoden zijn om de Richards' vergelijking veel sneller op te lossen dan de *finite difference*-methode die in SWAP is gebruikt, 2/ het voor de toepassing nodig is om het vochtprofiel met dit detail te behandelen (de opmerkingen door Van Walsum & Groenendijk, 2008, suggereren dus van wel) en de data ook beschikbaar is voor de ondersteuning van dat detail. Verder merken we nog op, dat in het kader van het NMDC (Nationale Modellen en Data Centrum) gekeken wordt naar versnelling van MetaSWAP en de andere modellen binnen het NHI d.m.v. parallel programmeren in OpenMP.

Opmerkingen bij vraag 7:

Zoals al opgemerkt moet vraag 7 breder ingestoken worden dan alleen de numerieke vergelijkingen; de hele implementatie moet uitgelegd én beoordeeld worden. Bij MetaSWAP zijn de *pre- en post-processing* belangrijke stappen, die niet alleen goed moeten verlopen, maar ook onderdeel zijn van de evaluatie in termen van evenwicht, omdat ze deel uitmaken van het oplossen van het formele model.

Er moet ook gevraagd worden naar uitgevoerde testen, omdat daaruit beperkingen kunnen blijken, die weer relevant voor de toepassing kunnen zijn.

Vragen 6B, C en D kunnen gekopieerd worden naar vraag 7. Het lijkt geen zin te hebben om te vragen naar de gegevens die nodig zijn voor het formele model, als het formele model vervolgens deels omzeild wordt door de numerieke implementatie. Echter, de keuze om te omzeilen kan natuurlijk (deels) zijn ingegeven door het gebrek aan bruikbare data, al is die keuze hier vooral ingegeven door de numerieke kant van de zaak (snelheid).

4.9 Schematisering MetaSWAP

De schematisering is bv. het indelen van de bodem in verschillende lagen ('stratificatie'), of het clusteren van landgebruikstypen. De mate van detail die bij schematisering wordt gevolgd is afhankelijk van de toepassing, maar moet wel voldoende ondersteund worden door de data.

8A. Indien van toepassing, beschrijf de schematisering, of geef relevante en specifieke referenties?

De schematisering van SIMGRO, waar MetaSWAP onderdeel van is, wordt expliciet beschreven in Van Walsum *et al* (2010), met name figuur 1 en 3, dat weer gebaseerd is op de schematisering van Querner & Van Bakel (1989). Verticaal wordt de bodem in profielen ingedeeld, bestaande uit meestal 2 of 3 lagen. Elke laag bevat één van de 21 bodemtypen uit de PAWN-indeling, weergegeven in Van Walsum & Groenendijk (2008), tabel 2, en gebaseerd op het werk van Wösten *et al* (1988). De PAWN-indeling kent 4 hoofdtypen bodem (löss, zand, klei, veen) die verder zijn onderverdeeld in de 21 bodemtypen. Deze bevatten de geclusterde bodem-fysische eigenschappen als vocht-gehalte en doorlaatbaarheid. De wortelzone valt meestal samen met de bovenste laag. Voor het NHI wordt de bodem horizontaal verdeeld in cellen van 250 bij 250 meter, dat 550.000 eenheden oplevert voor heel Nederland (uitgezonderd de Wadden en Zuid-Limburg), terwijl de bodem inclusief MODFLOW uit 31 lagen is opgebouwd.

**8B. Wordt de schematisering voldoende ondersteund door de data (ja/matig/nee)?
Motiveer.**

Voor de verzadigde zone worden pompproeven gebruikt in het veld op een flink grotere schaal. De bodem-fysische relaties voor de onverzadigde zone zijn echter afgeleid van laboratoriummetingen met monsters op kleine schaal (ca. 10 bij 10 cm). MetaSWAP is gebaseerd op de PAWN-schematisering uit 1988 van 1:250.000 (Wösten *et al*, 1988), die feitelijk niet zou mogen worden gebruikt op de ruimtelijke schaal die voor het NHI wordt gebruikt. De 21 bodemtypen zijn ingedeeld naar diverse karakteristieken als vochtgehalte en doorlaatbaarheid. De bruikbaarheid van de bodem-fysische karakteristieken voor simulatiemodellen als MetaSWAP is vergroot door het opstellen van de 'Staringreeks', die de karakteristieken beschrijft als analytische vergelijkingen, die echter statistisch gefit zijn en zeker niet "absoluut" zijn. Wösten *et al* (2001) wijzen er dan ook op, dat er richtlijnen bij de Staring-reeks worden vermeld voor het gebruik ervan (de verwijzing is naar een verouderde pagina). Er bestaat de neiging om de schematisering steeds verder te verfijnen, en inmiddels sluit de schematisering van regionale studies niet aan bij de oude PAWN-indeling, omdat het rekenwerk gedetailleerder is dan de PAWN-schematisering. MetaSWAP wordt echter vooralsnog niet gebruikt voor regionale studies. De conclusie lijkt hier te zijn, dat de ondersteuning van de schematisering zwak is.

4.10 Gevoeligheids- & onzekerheidsanalyse, opbouw code

Deze vraag is gericht op het bepalen welke parameters, forcing, invoer, delen van de code, etc. belangrijk zijn en welke overbodig zijn, en welke onderdelen een grote onzekerheid introduceren.

9A. Zijn er gevoeligheids-/onzekerheidsanalyses uitgevoerd? Zo ja, beschrijf deze of verwijs naar paginanummers uit een gepubliceerd rapport of artikel en analyseer de resultaten. Sommige factoren kunnen weinig van belang blijken voor de toepassing, zodat er weinig moeite hoeft te worden gedaan om deze goed te kalibreren, of gebruikt kunnen worden om het model te vereenvoudigen. Andere factoren kunnen juist essentieel blijken.

Een gevoeligheidsanalyse van MetaSWAP is beschreven door Van Walsum & Veldhuizen (2011), hoofdstuk 8. Daarin is onderscheid gemaakt tussen de parameters voor tijd- en ruimtelijke discretisatie, en die voor eigenschappen van de kolommen. De breedte van de kolommen bepaalt o.a. de mate van beïnvloeding door aanliggende kolommen. De verschillen tussen de uitkomsten van SWAP en MetaSWAP worden groter naarmate de wortelzone dikker wordt, maar dat geldt vooral voor een paar combinaties van eenheden die eco- of agro-hydrologisch niet of nauwelijks voorkomen. Met het oog op de toepassing is dit dus geen serieus probleem. De gevoeligheid van de dikte van de capillaire zone blijkt nogal verschillend te zijn voor de verschillende van de 21 eenheden.

Er zijn geen onzekerheidsanalyses uitgevoerd met MetaSWAP. In de interviews is een poging gedaan om enkele bronnen van onzekerheid samen te vatten.

1. De PAWN-schematisering en diens discrete aard wordt als een grote zwakke plek ervaren. De data m.b.t. de bodemeigenschappen is typisch afgeleid van bodemmonsters van 10 bij 10 cm², terwijl deze voor het NHI naar 250 bij 250 m² worden geaggregeerd. De profielen zijn feitelijk geclusterde bouwstenen, die bovendien conceptuele secundaire processen missen. De discrete aard van de eenheden laat af en toe grote 'sprongen' in de eigenschappen zien, bv. lichte klei kent weinig verdampingsreductie, terwijl zware klei 30% verdampingsreductie kent. De keuze om een eenheid als lichte of zware klei te classificeren heeft daarmee grote consequenties; in een *pilot* is gedemonstreerd dat aaneenliggende eenheden volgens de PAWN-classificatie gelijk zijn, terwijl de data laat zien dat er behoorlijke verschillen zijn. Wellicht kan de aanpak van het indelen van de bodem in vaste laagjes o.b.v. 21 eenheden vervangen worden door een indeling o.b.v. bepaalde eigenschappen, die veel 'natuurlijker' bij MetaSWAP/NHI passen en een ruimere keuze aan opties toelaat.
2. Er zit ook onzekerheid in de neerslag-invoer. De invoer is nu 250 bij 250 m², maar de metingen zijn 'puntmetingen' met buisjes met een doorsnede van een aantal cm. Deze data is nog te verbeteren met neerslagradar, waar men de daggemiddelde neerslag meet, en het patroon in neerslag beschrijft (bv. constante motregen, of een flinke regenbui). De onzekerheid in verdamping is echter veel groter dan neerslag. Hier worden twee redenen voor gegeven:
 - Verdamping is moeilijk direct te meten, omdat verdamping wordt afgeleid uit gewichtsverlies of een Penman-vergelijking (Van Walsum *et al*, 2010, p. 21). Alle factoren die invloed uitoefenen op verdamping hebben dan invloed (temperatuur, wind);
 - Verdamping wordt in Nederland alleen gemeten bij hoofdstations; per saldo een stuk of 12 puntmetingen voor het Nederland.

Daarnaast zijn de huidige gewassen niet altijd meer vergelijkbaar met de gewassen die zijn gebruikt om de parameters te bepalen voor verdamping door gewassen; zoals al genoemd zijn de gewasgegevens van 19 gewassen gebaseerd op veldproeven van voor 1987 (Feddes, 1987).

3. Grote onzekerheid zit 'm verder nog in de te kalibreren parameters. De Boesten-parameter en de dikte van de capillaire zone zijn twee parameters die nog moeten worden gekalibreerd (is eerder

genoemd, en wordt nog behandeld bij de vraag over kalibratie). Deze parameters missen echter een procesmatige basis, en daarom is er onzekerheid over hun “geldigheid”.

4. Voor een (groot) deel van de toepassingen van het NHI speelt de factor ‘mens’ een belangrijke rol in vraag en aanbod van water (dit werd bv. ook genoemd in de onzekerheidsstudie van de Baakse Beek, Van der Sluijs *et al*, 2012): de keuze van gewassen en ander landgebruik, klimaatverandering als gevolg van menselijk handelen die de verdamping beïnvloed, het geforceerd veranderen van waterlopen, etc. Dit is ook een bron van onzekerheden, die zeker belangrijk is met het oog op de toepassing.

Opmerkingen bij vraag 9A:

Wellicht is het zinnig om een expliciete vraag toe te voegen om onzekerheden in te schatten, of in elk geval om de bronnen van onzekerheid expliciet te maken. Dit is nu hierboven gedaan als voorbeeld.

9B. (niet voor databestanden) Welke numerieke integratiemethode is gebruikt? Bespreek het prestatievermogen van de methode.

Opmerkingen bij vraag 9B:

Deze vraag lijkt erg veel op vraag 7C (“Beschrijf en motiveer de keuze voor numerieke rekenmethoden”), en wordt daarom overgeslagen. Vraag 9B kan eventueel vervangen worden door iets als: “Zijn er bijzonderheden of zwakke plekken door de keuzes in de implementatie?”

9C. Wat is de rol van elke module van het model/bestand, bv. een gewasgroei-deelmodel als onderdeel van een hydrologisch model, of stappen in het stroomschema of de scripts van een bestand? Bespreek dat in relatie met de toepassing. Zijn er onderdelen die overbodig zijn? Missen er nog onderdelen? Motiveer waarom.

MetaSWAP is een grensconditie van MODFLOW, en bestaat uit het ‘oplossen’ van het vochtprofiel in combinatie met de ondergrens (grondwaterpeil; MODFLOW) en de bovengrens, die afhangt van neerslag en verdamping. De gewasgroei-module voor de bovengrens is een integraal onderdeel van MetaSWAP (en tevens de inspiratie voor het eerste voorbeeld in deze vraag). Het is een eenvoudige module, die eerder te simpel dan te complex is. Voor de toepassing van het NHI hoeft dat geen probleem te zijn, mits het geen scenario’s betreft die grote variaties in gewasgroei behelzen, direct of indirect. Dit lijkt echter wel het geval te zijn: het NHI wordt ingezet voor klimaatstudies, waarin de gewasgroei sterk beïnvloed wordt door de temperatuur en CO₂. Bovendien laat de klimaat-tijdreeks van Nederland een behoorlijke spreiding zien in (de ruimtelijke en temporele verdeling van) neerslag en verdamping in natte en droge jaren. Daarnaast is er nog de genoemde onzekerheid in de factor ‘mens’ t.a.v. welke gewassen waar en hoe groeien. Overigens gelden deze opmerkingen niet alleen voor de gewasgroei-module, maar ook in mindere mate voor MetaSWAP in het geheel. De tegenvraag is dan, of dat een complexere gewasgroei-module dit zou kunnen opvangen. Dit is ook twijfelachtig, en vermoedelijk zal de complexiteit van het model dan alleen maar worden verhoogd zonder een toename in dataondersteuning of een beter voorspellings-vermogen.

9D. Beoordeel de complexiteit in termen van rekentijd en efficiëntie. Wat is de duur van een run? Hoeveel (tussentijdse) uitvoer is er? Hoe groot is de datastroom? Zou de code sneller/efficiënter kunnen, bv. door selectie, aggregatie, andere methode, andere modulaire opbouw, etc.?

Kwantitatieve informatie over de methode die wordt gebruikt in MetaSWAP is te vinden in Van Walsum & Veldhuizen (2011), p. 23-24. De bevindingen van tests suggereren dat de code van MetaSWAP voldoende efficiënt is; in elk geval is het tot dusver niet gelukt om MetaSWAP veel te versnellen d.m.v. parallel programmeren. MetaSWAP is bedoeld om SWAP te vervangen als versnelling in bepaalde toepassingen, zoals het NHI. MetaSWAP is in elk geval niet langzamer dan MODFLOW, i.t.t. SWAP; over de vergelijking met de andere modellen in het NHI is nu geen inschatting te maken. De “prijs” voor de aanpak met MetaSWAP is een paar procent afwijking in de uitvoer van MetaSWAP t.o.v. SWAP in bepaalde gevallen bij bepaalde tijdstapgroottes. De vraag is, of dat die paar procent

acceptabel is binnen de toepassing, en of het überhaupt merkbaar is in vergelijking met de dataondersteuning, de uitvoer van de andere modellen, of de onzekerheden in het NHI. Dit laatste lijkt nauwelijks het geval te zijn: bv. de PAWN-schematisering lijkt voor grotere afwijkingen te zorgen, en de genoemde bronnen van onzekerheid in vraag 9A lijken ook voor grotere spelingen dan 5% in oppervlakte- en grondwaterstanden te kunnen zorgen.

Opmerkingen bij vraag 9:

Vraag 9 mist de vergelijking met data en vooral de toepassing, wat essentieel is voor de beoordeling in termen van evenwicht.

4.11 Kalibratie MetaSWAP

De kalibratie is het proces waarbij de parameters, begincondities, etc. van het model of bestand van waarden worden voorzien. Kalibratie is vaak gebaseerd op gevoeligheids-analyses (zie vorige vraag).

10A. Indien beschikbaar, beschrijf de uitgevoerde kalibraties en motiveer de keuze, of verwijst naar paginanummers uit een gepubliceerd rapport of artikel. Geef ook aan welke doelfuncties (lokaal, globaal, deterministisch, stochastisch) zijn gekozen, of dat de betrouwbaarheid van de gekalibreerde parameters wordt gekwantificeerd, hoe, op welke data, of en hoe de betrouwbaarheidsintervallen zijn bepaald, etc. met eventuele verwijzingen naar de literatuur. Let op over-fitting, meervoudige lokale optima, en identificeerbaarheidsproblemen.

MetaSWAP is een meta-model van SWAP, en heeft (zoals al eerder genoemd) 2 parameters die nog apart gekalibreerd moeten worden: de Boesten-parameter, en de dikte van de capillaire zone. Dit wordt per eenheid (van de 21 eenheden) gedaan, tegen de uitvoer van SWAP. Voor de kalibratie wordt gebruik gemaakt van een doelfunctie die de grondwaterstand tegen de verdamping afzet. Er is geen geautomatiseerde software als PEST gebruikt; de doelfunctie is 'afgetast' in de parameter ruimte. De details van de kalibratie zijn weergegeven in hoofdstuk 6 van Van Walsum & Veldhuizen (2011). Omdat MetaSWAP tegen de uitvoer van SWAP wordt gekalibreerd zijn er eigenlijk geen problemen met de datavoorziening. Er zouden problemen kunnen zijn t.a.v. identificeerbaarheid, maar die gelden dan waarschijnlijk ook voor SWAP. Er zijn geen betrouwbaarheidsintervallen bepaald.

10B. Welke eisen stelt de toepassing aan de nauwkeurigheid waarmee parameters, etc. wordt bepaald. Let ook op over-fitting.

Met het oog op de toepassing binnen het NHI zijn de eisen voor MetaSWAP dat de uitvoer van het meta-model zoveel mogelijk overeenkomt met de uitvoer van SWAP. Dat staat los van de eisen die de toepassing van het NHI stelt aan de nauwkeurigheid waarmee de parameters, etc. binnen SWAP dan wel MetaSWAP moeten worden bepaald. Om die marges te bepalen moet eerst helder zijn wat de toegestane speling is in de voorspellingen m.b.t. de waterpeilen, en hoe die speling zich vertaalt naar de verschillende modelparameters, begincondities, etc. Daarvoor is het ook nodig om de gevoeligheden bepaald te hebben. De kans bestaat, dat de mogelijkheden om o.b.v. de data parameters, etc. te identificeren niet zodanig zijn, dat de gewenste marges haalbaar zijn. De marges zullen afhangen van de toepassing (bv. effecten van klimaatverandering, zeespiegelstijging, waterkwaliteit, etc., zie vraag 2C). Meer hierover bij het onderdeel 'validatie'.

10C. Welke mogelijkheden leveren de data om waarden van parameters, etc. te bepalen. Denk hierbij aan resolutie, nauwkeurigheid, aantal gegevens, het aggregatieniveau van de gegevens, en let op over-fitting.

De uitvoer van SWAP geldt als data in dit geval, en daarmee is in principe de data voor kalibratie ongelimiteerd. De kalibratie op SWAP-uitvoer is echter wel arbeidsintensief, en bovendien moeten er zeer veel simulaties gedraaid worden om de database aan te maken. Verder geldt dit alleen voor het

ondergrondse deel; het bovengrondse deel van MetaSWAP is gelijk aan SWAP (op de interceptieverdamping na, zie eerdere opmerkingen), en is dus wel degelijk beperkt in de mogelijkheden o.b.v. de data. Bovendien is de betrouwbaarheid van MetaSWAP door het meta-concept afhankelijk van de betrouwbaarheid van SWAP en de data die ervoor zijn gebruikt. M.a.w., SWAP moet voldoende gevalideerd zijn (zie ook volgende vraag) om vertrouwen in de geldigheid van de uitvoer van MetaSWAP te krijgen. De antwoorden bij vragen 5, 6 en 9 laten verder al zien dat er op het gebied van dataondersteuning nog wel een aantal verbeteringen mogelijk zijn. Echter, omdat er geen onzekerheidsstudies zijn gedaan, is het moeilijk om deze vraag kwantitatief te beantwoorden.

4.12 Validatie MetaSWAP

Wij definiëren hier een validatie als een toets of het model of bestand een redelijke representatie is van het werkelijke systeem dat wordt gemodelleerd. Dit geschiedt op basis van een set onafhankelijke gegevens (gegevens die nog niet in een eerder stadium bij het modelleren gebruikt zijn).

11A. Indien beschikbaar, beschrijf de gedane validatiestudies of verwijs naar paginanummers uit een gepubliceerd rapport of artikel.

De validatie van MetaSWAP gebeurt in principe tegen SWAP, omdat het de uitvoer van SWAP moet reproduceren, en SWAP daarmee de “werkelijkheid” is geworden. SWAP is op zijn beurt tegen metingen gevalideerd van grondwaterstand en verdamping. Voor temperaturen onder 0 graden geeft MetaSWAP geen betrouwbare voorspellingen meer. De overeenkomst tussen de uitvoer van SWAP en MetaSWAP is goed tot zeer goed, behalve voor kleigrond. Een aantal combinaties doet 't niet goed, maar die hebben geen eco- of agro-hydrologische betekenis (eerder genoemd). De validatie van MetaSWAP wordt besproken in hoofdstuk 7 van Van Walsum & Veldhuizen (2011). MetaSWAP is verder ook indirect gevalideerd in het kader van het NHI. De uitkomsten van het NHI worden afgezet tegen 'harde' en 'zachte' data, die worden gegroepeerd om betrouwbaarheidsmarges te genereren. De belangrijkste toepassingsvalidatie is de zoetwaterverkenning, en de uitkomsten vallen tot nog toe binnen door gebruikers bepaalde marges. Dit is een kwalitatieve maat, gebaseerd op door beheerders gestelde verdelingsregels, waarbij gekeken wordt naar het aantal keren dat er aan de watervraag van één of meerdere consumenten als industrie of landbouw niet voldaan kan worden (mondelijke correspondentie, Ab Veldhuizen).

11B. Onderbouw/beoordeel de waarde van de validatie(s). Hoe relevant zijn de verschillende validatiestudies voor de verschillende toepassing(en) van het model of bestand? Zijn alle toepassingen voldoende afgedekt door de studies? Is ook de betrouwbaarheid van de validatie gekwantificeerd? Welke toepassingen moeten opnieuw worden gevalideerd?

MetaSWAP/NHI is nog niet voor alle toepassingen gevalideerd, maar de gedane studies zijn wel zeer relevant. Er is een redelijk vertrouwen in het vermogen van MetaSWAP om betrouwbare uitvoer (waterstanden) te genereren. Van Walsum & Veldhuizen (2011, p. 55) melden verder, dat er in enkele gevallen in de validatiestudie een grote afwijking is tussen de uitvoer van SWAP en MetaSWAP. Die wordt echter veroorzaakt door een fout in het concept van SWAP, waarin een homogene grond zonder preferente stroming wordt verondersteld, waarin korte zomerbuien niet diep genoeg doordringen. In MetaSWAP worden buien toegevoegd bij de waterbalans over de gehele wortelzone, waardoor de capillaire opstijging vermindert en er een effect ontstaat dat op preferente stroming lijkt, wat meer in overeenstemming met de werkelijkheid is.

4.13 Evenwichtsoordeel MetaSWAP, bevindingen

Dit laatste onderdeel is gericht op het geaggregeerde resultaat van de antwoorden op de verschillende vragen en de meningen van de verschillende betrokkenen, en kan indirect ook tot adviezen leiden ter verbetering van het model of bestand.

Deze vraag moet specifiek met het oog op de toepassing gesteld worden, en eigenlijk moet deze vraag dus ook per toepassing opnieuw bekeken worden. In dit geval ligt de nadruk op het NHI, dus is dit niet echt nodig.

12A. Beoordeel de mate van vertrouwen in het model/bestand in het algemeen, mede gebaseerd op wat er is gedaan aan testen, verificatie, kalibratie, validatie, gevoeligheids- en onzekerheidsanalyses, en motiveer deze.

MetaSWAP is behoorlijk getoetst; er is dus wel vertrouwen bij ontwikkelaar en gebruiker. Wel is er verschil van inzicht in de beoordeling van de complexiteit: terwijl bv. Deltares van mening is dat MetaSWAP simpeler moet, is men bij Waterkwaliteit juist van mening dat MetaSWAP meer op SWAP zou moeten lijken. Ironisch genoeg lijkt MetaSWAP daarmee wel aardig het midden te houden in termen van complexiteit. Naar de mening van de ontwikkelaar is de code wel in orde, maar ontbreekt er nog veel aan de ondersteuning door data. Versnellen van de code is verder weinig zinvol: MetaSWAP is een grensconditie op MODFLOW, en is op dit moment zeker niet trager dan MODFLOW. Het is dan ook niet de bottleneck binnen het NHI, en investeren in versnellen is daarmee niet erg zinvol. Overigens zijn er (zoals al gemeld) pogingen binnen het NMDC om het NHI te versnellen via parallel programmeren.

12B. Zijn er veel missende zaken? Zijn er veel vragen waar de zaken als 'te eenvoudig' zijn beoordeeld? Motiveer.

Een aantal missende zaken zijn genoemd in 3C en 3D, maar die betreffen vooral fysische aspecten, als hysteresis en vries- en dooiprocessen. Hysteresis wordt toch gezien als een belangrijk proces. Wösten *et al* (2001) bv. vermelden dat: “[g]ezien de rol van hysteresis als veroorzaker van preferente waterstroming in gronden, is te verwachten dat in de toekomst zowel de uitdrogings- als de bevochtigingscurve bepaald gaat worden [in de Staringreeks]” (p. 13). Om dit aspect te dekken zou de database van MetaSWAP uitgebreid moeten worden met twee vochtprofielen per set selectievariabelen, en een ‘geheugen’ om de keuze tussen de uitdrogings- en de bevochtigings-curve te kunnen maken (anders is het niet mogelijk om te bepalen of de kolom in een uitdrogings- of vernattingsfase zit). Dit maakt het model wel weer complexer en trager, maar het is nog niet helder hoeveel beter de voorspellingen daarmee worden voor de toepassing. Voor klei- en veengronden is preferente stroming een belangrijk aspect. De preferente stroming, nu nog niet in MetaSWAP, wordt beschreven door een morfologische functie die poriedichtheid als functie van diepte beschrijft (mondelinge correspondentie, Ab Veldhuizen), maar dit houdt weer veel meer parameters en dus complexiteit in. Een ‘oplossing’ is wellicht om dit op de een of andere manier in de database met SWAP-uitvoer te verwerken, m.a.w. het zit wel in SWAP maar niet direct in MetaSWAP verwerkt.

Voor de toepassing van waterbeheer binnen toekomstscenario's, bv. met klimaatverandering, moet ook breder gekeken worden. Hierbij valt te denken aan fysisch georiënteerde zaken als de invloed van klimaatverandering op neerslag en verdamping, maar ook aan andere zaken, zoals socio-economische veranderingen. Deze zijn echter meer relevant voor de toepassing van het NHI als geheel, dan voor MetaSWAP als onderdeel van het NHI. Een aantal aspecten die nog bij MetaSWAP bekeken zouden kunnen worden betreffen vooral data-aspecten, waaronder de opschaling van puntmetingen naar vlakken. Verder kan ook gedacht worden aan de verdamping door vegetatie, die wellicht te weinig dynamisch is. In elk geval zijn de parameterwaarden inmiddels ruim verouderd, maar de gewasgroei-module is wellicht te simpel t.o.v. de hydrologie achter MetaSWAP. Verder zijn de opmerkingen relevant die genoemd zijn bij de “verkennende” onzekerheidsanalyse (vraag 9A), bv.

klimaat effecten zouden marginaal kunnen zijn t.o.v. de onzekerheid m.b.t. de reacties van boeren op de veranderde waterbeschikbaarheid.

12C. Zijn er veel overbodige componenten? Zijn er veel onderdelen waar het eenvoudiger zou kunnen? Motiveer.

Er lijken niet echt overbodige componenten te zijn, al zou de gewasmodule buiten MetaSWAP geplaatst kunnen worden. De gebruikte aanpak in MetaSWAP is wel als "sympathiek" aan te duiden, en lijkt een redelijk evenwichtige oplossing voor het probleem m.b.t. het numeriek benaderen van de Richards' vergelijking. De vraag is wellicht beter te beantwoorden als ze in een historische context wordt geïnterpreteerd. De "voorganger" van SWAP en MetaSWAP is CAPSIM/CAPSEV, dat conceptueel een grondwaterfilter beschrijft, waarbij de diepere bodem een bergingscoëfficiënt is (Wesseling, 1991). CAPSEV gebruikt al het idee van een *steady state* in de capillaire opstijging, maar werd toch te simplistisch bevonden voor het voldoende beschrijven van de stromingen in de onverzadigde zone. MetaSWAP zit qua complexiteit en reken-snelheid tussen CAPSEV en SWAP in. Een kritische noot is nog wel te maken m.b.t. het hele bestaansrecht van MetaSWAP; in het hypothetisch geval dat SWAP met andere (nieuwere) methoden voor het numeriek benaderen veel te versnellen valt, dan is MetaSWAP wellicht in zijn geheel niet nodig. Maar dat neemt dan niet weg dat de complexiteit van MetaSWAP veel minder is dan die van SWAP.

Opmerkingen bij vragen 12B en C:

De vragen zijn wellicht beter te beantwoorden als ze in een historische context worden gesteld en/of door vergelijking met alternatieve modellen. Een vergelijking met in de peer-reviewed literatuur gepubliceerde modellen lijkt hierbij een voor de hand liggende optie.

12D. Is er voldoende datasteun, en waaruit blijkt dat? Motiveer

Bij een aantal vragen (5, 6, 9) is al naar boven gekomen dat er op een aantal punten een zwakke onderbouwing is door de data. De PAWN-schematisering wordt als zeer problematisch ervaren, er zijn weinig betrouwbare metingen van het grondwaterpeil (hoewel de metingen tegenwoordig semi-continu zijn, kunnen buizen fout staan, filters kunnen verkeerd staan t.o.v. de verschillende bodemlagen, buizen kunnen droogvallen en kapotgaan, er zijn bodemkundige veranderingen als huizen- en wegenbouw, etc.), er is een typisch punt-vlakprobleem (van punt-data van buizen naar vlakken van 250 bij 250 m²), de verdamping is weinig onderbouwd, en de gewasgroei-gegevens zijn verouderd. Bovendien is de gewasgroei-module eenvoudig, en wellicht te eenvoudig voor een toepassing als de verkenning van effecten van klimaatverandering; anderzijds zijn voor een deel data voor die laatste toepassing ook al lastig, omdat het een toekomstprojectie betreft, waarover weinig zekerheid bestaat (bv. wanneer het data over het toekomstig atmosferisch gehalte CO₂ betreft).

12E. Welke specifieke suggesties voor verbetering en data-aanlevering komen naar voren uit deze analyse?

De complexiteit van MetaSWAP lijkt redelijk overeen te stemmen met wat voor de toepassing binnen het NHI nodig is, maar de datasteun voor MetaSWAP (eigenlijk vooral SWAP) is nog problematisch. Een aantal suggesties om dit punt te verbeteren zijn genoemd in de antwoorden bij de verschillende vragen. De verdamping kan wellicht een stuk beter geschat worden o.b.v. satellietmetingen. De PAWN-schematisering zou waar-schijnlijk vervangen moeten worden voor een indeling naar 'natuurlijke' eigenschappen die veel flexibeler is. Nieuwe proeven voor het verkrijgen van gegevens m.b.t. de groei van de huidige gewassen zijn nodig. Bovenal is het nodig om een voldoende dekking in ruimte en tijd te krijgen van de waterstanden. Op het vlak van het model zelf zijn er suggesties om het concept van MetaSWAP te 'verbeteren' (bv. hysteresis), maar het loont de moeite om ook te zien of SWAP zelf te versnellen valt door bv. een numeriek wiskundige naar de integratiemethode te laten kijken.

Opmerkingen bij vraag 12E:

Deze vraag zou uitgebreid kunnen worden. Men zou willen weten wat voor andere opties overwogen zijn, en waarom er niet voor die opties gekozen is uiteindelijk.

Opmerkingen bij vraag 12:

Het hangt nogal af van de toepassing of het model evenwichtig gevonden wordt of niet. Echter, in vraag 12 wordt geen specifiek onderscheid gemaakt t.a.v. de toepassing. Het is nodig om expliciet bij te voegen dat vraag 12 *per toepassing* opnieuw moet worden bekeken.

4.14 Verdere bevindingen

Enig pragmatisme is wel vereist voor het gebruik van de lijst; het is toch vooral een hulpmiddel, en niet een doel op zich. Desalniettemin zijn er wel enkele opmerkingen te maken t.a.v. de functionaliteit van de lijst n.a.v. deze casus.

De inbedding in de historische context mist nu nog volledig, maar kan zeer nuttige informatie bevatten. Wat is er met de vorige 6 versies van MetaSWAP gebeurd? Betreffen de verschillen tussen de versies vooral verschillen in implementatie (bv. het oplossen van bugs), of zijn er ook conceptuele wijzigingen gemaakt? En zo ja, waarom zijn die gemaakt? Er is een mogelijkheid dat oudere versies evenwichtiger zijn dan de nieuwste, om de een of andere reden. Voor een goede evaluatie in het kader van evenwicht is het ook nodig MetaSWAP te vergelijken met modellen met een vergelijkbare rol.

Het botst nog een beetje tussen data, schematisering, en de volgorde waarin vragen over data, schematisering en model worden gesteld. Waarschijnlijk moet men eerst weten hoe het model eruit ziet en geïmplementeerd is, voordat duidelijk te bepalen is wat voor invoer er in moet. De schematisering is ook deel van de implementatie, maar is ook apart van belang, omdat aanbod van data en toepassing met elkaar kunnen botsen waar het data in het algemeen, en schematisering in het bijzonder betreft. Er wordt verder ook niet gevraagd naar alternatieve opties m.b.t. data, bv. is het mogelijk om uitvoer van een ander model of bestand te gebruiken, of zijn er andere datasets?

Er zou nog een expliciete vraag 12F toegevoegd kunnen worden: "Beoordeel nu (gebruikmakend van alle voorgaande antwoorden) of het model evenwichtig is of niet (complexiteit i.r.t. data en toepassing). Zo ja, motiveer. Zo nee, geef een aantal (gemotiveerde) punten waarop het model te verbeteren valt, eventueel met suggesties hoe."?

Er is nogal een verschil tussen bv. conceptueel model en numerieke implementatie. Bv., een model kan erg simpel conceptueel zijn maar behoorlijk complex om goed te implementeren (Richards' vergelijking is een voorbeeld daarvan), maar andersom kan ook. De lijst is wel opgezet met die filosofie, maar dit komt nog niet erg expliciet in de huidige opzet van de EMC naar boven. Eigenlijk zou de lijst bij elke stap in de modelleercyclus expliciet moeten vragen naar een beoordeling van evenwicht.

5 Natuurplanner

5.1 Inleiding

De Natuurplanner is een belangrijke modelketen binnen het instrumentarium van het PBL, dat gebruikt wordt om een beeld te krijgen van de biodiversiteit binnen Nederland en de condities voor die biodiversiteit. Omdat het een modellenketen betreft die zeer veel daadwerkelijke simulatietijd kost, is er ook een 'kleiner' model verschenen, de MetaNatuurplanner (Pouwels *et al*, 2012), die bedoeld is om de Natuurplanner in elk geval voor een aantal toepassingen te vervangen.

De EMC v1.0 (Van Voorn *et al*, 2011) is tot dusverre nog niet getoetst tegen een modelketen. Zowel de Natuurplanner als de MetaNatuurplanner zijn geëvalueerd in het kader van dit project, met als doel inzicht te krijgen in het functioneren van de lijst EMC v1.0 wanneer het wordt toegepast op een modelketen. In dit hoofdstuk wordt de functionaliteit van de lijst geëvalueerd aan de hand van de Natuurplanner; de inhoudelijke bevindingen t.a.v. de Natuurplanner en Meta-Natuurplanner worden in een andere, nog te verschijnen rapportage behandeld. Bij deze toets is de lijst ingevuld a.d.h.v. interviews met ontwikkelaars en gebruikers. Wederom wordt in principe de indeling van de lijst gevolgd, maar wordt er afgeweken indien dit nodig wordt geacht. De resultaten zijn bij het uitgeven van dit WOt-werkdocument nog niet teruggekoppeld aan de geïnterviewde personen, wat de belangrijkste reden is dat in dit hoofdstuk alleen de bevindingen opgenomen voor de functionaliteit van de lijst.

5.2 Ingevulde vragenlijst Natuurplanner; basale gegevens

1 Geef aan: model/bestand, naam, versienummer, revisienummer, uitgiftedatum. Onder model verstaan we de broncode + invoergegevens.

De Natuurplanner (afgekort NP) is geen model maar een modelketen, die bestaat uit meerdere modellen en databestanden, die sequentieel gekoppeld zijn – d.w.z., de uitvoer van het ene bestand of model is invoer voor de volgende in de rij. Er zijn verschillende versies van de NP. De eerste versie en de ontstaansgeschiedenis worden beschreven in Latour *et al* (1997), p. 8 'samenvatting'. De tweede versie is beschreven en uitgebreid getest door Bakkenes *et al* (2003). Deze versie is gebruikt voor de tweede Natuurverkenning en bevat ook zware metalen, die in de volgende versie, versie 3, zijn verdwenen. De derde versie, NP 3.0, is beschreven door Van der Hoek & Bakkenes (2007). De resultaten van NP v3.0 zijn wel gebruikt maar nooit gepubliceerd in een Natuurverkenning (communicatie J. Wiertz); in plaats daarvan is de Meta-Natuurplanner gebruikt (zie het volgende hoofdstuk). De complexiteitsanalyse richt zich desalniettemin in principe op deze laatste versie (NP v3.0), al is het relevant om de gehele geschiedenis te overzien.

Opmerkingen bij vraag 1:

Om de functionaliteit van de EMC v1.0 ook op modelketens te richten, is het nodig om hier ook een overzicht van de verschillende modules te geven. Dit overzicht kan dan gebruikt worden als startpunt om doel, toepassing, concepten, etc. van de verschillende modules te inventariseren.

Uitbreiding vraag 1: Indien de casus een modelketen is, beschrijf ook kort uit welke modules de modelketen bestaat.

Versie 3.0 van de NP bevat diverse modules, verdeeld over twee 'takken': een aquatische en een terrestrische. Alleen de terrestrische is afdoende geïntegreerd, beschreven (Van der Hoek &

Bakkenes, 2007), en gebruikt, dus alleen dit deel wordt hier beschouwd. Een grafisch overzicht van de volledige NP is gegeven door Van der Hoek & Bakkenes (2007), Figuur 1, p. 13. Elke module bevat een model en een 'schil' die het model aanstuurt en in- en uitvoer behandelt. De *terrestrische* keten van NP v3.0 bestaat uit:

1. SMART/SUMO, een combinatie van SMART2 (Kros, 2002) en SUMO (Wamelink, 2008b). De rol van deze module is de bepaling van a-biotische factoren zoals de zuurgraad en de beschikbaarheid van stikstof, en de terugkoppeling met de vegetatie;
2. De P2E-module, tegenwoordig geïntegreerd met MOVE (zie hieronder), waarin de a-biotische factoren worden geconverteerd naar zgn. 'Ellenberg'-getallen. Deze getallen zijn afgeleid uit correlaties tussen a-biotische condities en soorten flora en fauna;
3. Drie 'biotische' modules die de kans op voorkomen van verschillende soorten als uitvoer leveren: MOVE (Van Adrichem *et al*, 2010), voor planten;
4. VLINDERMOVE (Oostermeijer & Van Swaay, 1998), voor vlinders;
5. LARCH (Pouwels *et al*, 2002), voor fauna;
6. BIODIV (Van der Hoek *et al*, 2000), die de uitvoer van de 3 'biotische' modules integreert tot een algemene 'natuurwaarde' voor de kwantificering van de biodiversiteit.

Bovenstaande opsomming is kort weergegeven in Tabel 5.1. Overigens is later nog het model DIMO toegevoegd als 'correctie' op P2E/MOVE, maar deze ontbreekt hier in de evaluatie.

Tabel 5.1. Tabel met een opsomming van de modules uit de NP v3.0, met sleutelreferenties, een korte beschrijving van de rol, en de kwaliteitsstatus anno 2012.

Naam	Referentie	Rol	Kwaliteit
SMART2	Kros (2002)	Stikstof en pH in de bodem	Status A Normenkader
SUMO	Wamelink (2008), Wamelink <i>et al</i> (2008)	Vegetatie en beheer	Status A Normenkader
P2E/MOVE	Van Adrichem <i>et al</i> (2010)	Vertaling a-biotische factoren naar kans op voorkomen vegetatie	Status A Normenkader
VLINDERMOVE	Oostermeijer & Van Swaay (1998)	Vertaling a-biotische factoren naar kans op voorkomen vlinders	-
LARCH	Pouwels <i>et al</i> (2002) Pouwels <i>et al</i> (2008)	Vertaling a-biotische factoren naar kans op voorkomen fauna	Status A
BIODIV	Van der Hoek <i>et al</i> (2000)	Integratie resultaten tot graadmeter	Normenkader

Omdat de NP geen afzonderlijk model is maar een verzameling van gekoppelde modellen en bestanden, is de NP als zodanig nooit beoordeeld voor Status A (voor Status A, zie Jansen *et al*, 2004). De modules SMART2, SUMO, en LARCH hebben wel Status A (maar worden in 2012 opnieuw beoordeeld in dat kader); P2E/MOVE is in 2011-2012 beoordeeld voor Status A; BIODIV is in 2010 in het Normenkader van PBL beoordeeld; VLINDERMOVE is niet beoordeeld naar een bestaand kwaliteitscriterium. Een beoordeling met Status A betekent, dat er in elk geval heldere documentatie beschikbaar is waarin de module beschreven wordt, die bruikbaar is voor het invullen van de EMC-lijst, en bovendien dat er iets gedaan is aan zaken als verificatie, kalibratie, validatie, en gevoeligheidsanalyse, die allemaal aan bod komen in de evenwichts-analyse.

Opmerkingen bij uitbreiding vraag 1:

In deze casus is het zeer behulpzaam dat er een overzicht was van de opbouw van de modelketen, en tevens dat er van de meeste modules een overzicht bestond van de documentatie in de vorm van een ingevulde checklist voor Status A of Normenkader. Zonder deze informatie zou het veel meer moeite kosten om deze casus te toetsen.

5.3 Doel & toepassingen NP

2A. Wat was het doel van het bestand/model?

De NP is bedoeld als het ecologisch instrumentarium op de nationale schaal, waarmee effecten kunnen worden berekend van veranderingen in milieu-, water-, ruimedruk, en natuurbeheer op de biodiversiteit (Van der Hoek & Bakkenes, 2007, p. 9). Dit moet gezien worden in het kader van de Ecologische Hoofdstructuur en Natura2000, de Nederlandse en Europese doelstelling om behoud van natuur en biodiversiteit binnen de EU te waarborgen. Het instrumentarium moet dus in staat zijn om diverse drijvende processen die de biodiversiteit (kunnen) beïnvloeden te modelleren, en kwantitatieve uitspraken te maken m.b.t. die invloeden op de biodiversiteit.

2B. Wat was het beoogde toepassingsgebied, en waaruit blijkt dit?

De toepassing van de NP is, in overeenstemming met het doel, gericht op het kwantificeren van de veranderingen in biodiversiteit door veranderingen in drijvers van processen die de biodiversiteit beïnvloeden. De ruimtelijke schaalgrootte is 250 bij 250 m² (identiek aan MetaSWAP/NHI), en de tijdstapgrootte is een jaar.

2C. Wat zijn de toepassingen nu? Heeft het model/bestand een rol in een modelketen? Zo ja, welke?

Opmerkingen bij vraag 2A, B, en C:

Vraag 2A en B moeten eigenlijk ook *per module* beantwoord worden. Het doel en de toepassing van de modelketen zelf hoeft niet overeen te stemmen met die van de afzonderlijke modules, vooral wanneer de modules ontwikkeld zijn voordat de modelketen samengesteld werd. Het is zinnig om een tabel samen te stellen en per module doel en toepassing in te vullen. De referenties zijn meestal direct uit de verschillende checklists gehaald. Vraag 2C kan overgeslagen worden als de casus een modelketen betreft: de toepassing van de modules is dan de rol die ze vervullen binnen de modelketen.

Uitbreiding vraag 2A: Wat is het doel van elk van de aparte modules?

Uitbreiding vraag 2B: Wat zijn de toepassingen van elk van de aparte modules?

Een kort overzicht met referenties naar doel en toepassing per module is gegeven in Tabel 5.2.

Tabel 5.2. Tabel met referenties van de doelen en toepassingen van de verschillende modules uit de NP v3.0.

Module	Doel/Referentie doel	Toepassing/Referentie toepassingen
SMART2	De Vries <i>et al</i> (1989), p. 351 Kros <i>et al</i> (1995), H. 1	Mol <i>et al</i> (2006), <i>preface</i> Mol (2005), H. 2
SUMO	Wamelink (2007), H. 4 Wamelink (2008b), H. 1	Wamelink (2008b), §1.2
P2E/MOVE	Van Adrichem <i>et al</i> (2010), H. 1	Wamelink <i>et al</i> (2009), H. 1
VLINDERMOVE	Oostermeijer & Van Swaay (1998), p. 271	Oostermeijer & Van Swaay (1998), p. 278
LARCH	Pouwels <i>et al</i> (2008), §2.2.1	Pouwels <i>et al</i> (2008), H. 4
BIODIV	Van der Hoek <i>et al</i> (2000) Reijnen <i>et al</i> (2010), §1.1	Van der Hoek <i>et al</i> (2000) Reijnen <i>et al</i> (2010), §3.1 + 5.1

2D. Overlapt het beoogde toepassingsgebied alle daadwerkelijke toepassingen (ja/deels/nee)?

2E. Indien relevant, in hoeverre overlappen de verschillende toepassingen met elkaar (goed/matig/slecht)?

Opmerkingen bij vraag 2D en E:

Bij vraag 2D en E is het nuttig om een matrix samen te stellen, waarin de verschillende zaken van de verschillende modules met elkaar vergeleken kunnen worden. Bij grote verschillen (en dus een magere overlap) tussen de doelen en toepassingen van de verschillende modules onderling en met die van de NP kan al voorzichtig de conclusie getrokken worden dat het evenwicht van de modelketen wellicht niet in orde is, of in elk geval moeilijk te bepalen zal zijn (Tabel 5.3).

Tabel 5.3. Opzet matrixtabel voor het onderling vergelijken van doel en toepassing van de NP v3.0 en de modules daarin. De overlap zou aangegeven kunnen worden met kleur (groen/geel/rood), in woorden (goed/matig/slecht), of anderszids.

Doel\Toepassing	SMART2	SUMO	P2E/MOVE	VLINDERMOVE	LARCH	BIODIV	NP
SMART2							
SUMO							
P2E/MOVE							
VLINDERMOVE							
LARCH							
BIODIV							
NP							

5.4 Systemanalyse NP

3A. Geef een systemanalyse. Dit kan bv. een verwijzing naar een rapport zijn. Aspecten die daarin aan bod komen zouden moeten zijn: Wat zijn de belangrijkste attributen en processen van het systeem? Welke terugkoppelingen zijn aanwezig? Hoe zijn de grenzen van het systeem bepaald of gedefinieerd? Hoe wordt met relevante attributen en processen buiten deze grenzen omgegaan, bv. randvoorwaarden, constanten, of simpelweg genegeerd? Over welke processen, attributen, terugkoppelingen, etc. bestaat er onzekerheid? Hoe is die onzekerheid bepaald?

Bij de systemanalyse is het de bedoeling dat met een ‘frisse blik’ gekeken wordt naar wat het doel is van het model, wat dus als uitvoer geproduceerd zou moeten worden, en welke processen daarvoor meegenomen zouden moeten worden. De bedoeling van de NP is om de kwaliteit van de natuur op de Nederlandse nationale schaal in te schatten en te evalueren. De kwaliteit van de natuur is daarbij opgevat als ‘biodiversiteit’, dus het aantal soorten planten en dieren. Er wordt gekeken naar de effecten van veranderingen in milieu-, water-, ruimedruk, en natuurbeheer op de biodiversiteit (Van der Hoek & Bakkenes, 2007, p. 9). Dit moet gezien worden in het kader van de Ecologische Hoofdstructuur en Natura2000, de Nederlandse en Europese doelstelling om behoud van natuur en biodiversiteit binnen de EU te waarborgen. De NP moet dus in staat zijn om diverse drijvende processen die de biodiversiteit (kunnen) beïnvloeden te modelleren, en kwantitatieve uitspraken te maken m.b.t. die invloeden op de biodiversiteit.

De eerste aanname is dat biodiversiteit inderdaad een goede indicator is voor de stand van de natuur in het algemeen. BIODIV is bedoeld om de “kwaliteit” van biodiversiteit in te schatten o.b.v. de NCI (“natural capital index”; Ten Brink, 2000, H2; Van der Hoek *et al*, 2000), een mathematisch product van natuurlijke gebieden (ook wel ecosysteem-*kwantiteit*, het % natuurlijk areaal in een land of regio) en een maat van ecosysteem-*kwaliteit* (% van een geselecteerd nul-niveau, bv. de stand in 1950) van die gebieden. BIODIV selecteert soorten uit de drie ‘biotische’ modules (MOVE voor planten, VLINDERMOVE voor vlinders, en LARCH voor fauna), en interpreteert de mate van voorkomen van de geselecteerde soorten als de natuurkwaliteit van een gebied. Deze keuze zou gemotiveerd moeten worden.

Aannemende dat bovenstaande aanname terecht is, dan is verder de vraag welke processen in belangrijke mate de biodiversiteit beïnvloeden. Beleid zal dan gericht worden op die processen, om te proberen of negatieve invloeden kunnen worden verminderd. Daarvoor is dan wel voldoende inzicht nodig in die processen.

Beschouw nu per toepassing

3B. Welke systeem-analytische aspecten (attributen, processen, terugkoppelingen, etc.) zijn (direct) relevant voor de toepassing? En welke in mindere mate of niet? En hoe is die relevantie bepaald?

3C. Zijn alle voor de toepassing belangrijke aspecten als gegeven in B meegenomen in het model/bestand? En welke niet?

3D. Zijn er minder of niet belangrijke aspecten meegenomen in het model/bestand (ja/weinig/geen)?

Opmerkingen bij vraag 3B, C en D:

Deze vragen zijn onverminderd geldig, maar kunnen op twee niveaus bekeken worden. Men kan deze vragen beantwoorden met het oog op de NP. Maar wederom is het ook zinnig om *per module* te kijken welke aspecten van belang zijn, en welke minder of niet. Vragen 3B, C, en D kunnen dus ook gesteld worden per module, en voor een goede evenwichtsevaluatie is dit vermoedelijk zelfs noodzakelijk. In deze casus veronderstellen we verder dat er sprake is van één bepaalde toepassing (al moet dit nog blijken uit de antwoorden op vraag 2); in het geval dat er sprake is van meerdere toepassingen, dan moet er *per toepassing* een evenwichtsevaluatie gedaan worden. De verschillende vragen zouden verder nog kunnen vragen naar de motivering achter de keuzes om processen wel of niet mee te nemen. Dit kan behulpzame informatie verschaffen over mogelijke alternatieven voor of aanpassingen aan het model.

De bedoeling is effecten te bekijken op biodiversiteit van veranderingen in milieu-, water- en ruimtedruk en natuurbeheer, om daarbij effecten in te kunnen schatten van beleid en/of toekomstige socio-economische ontwikkelingen. De ruimtelijke schaal van de NP is 250 bij 250 m², de tijdstap is een jaar, de temporele horizon is ca. half 21^{ste} eeuw. Vegetatie is een belangrijk element m.b.t. biodiversiteit. Voor een deel vormt het zelf de biodiversiteit, voor een ander deel wordt de biodiversiteit sterk gevormd door de vegetatie: waardplanten, begrazing door fauna, schaduwwerking, de capillaire werking van de wortelzone, beheer via maaien, vermesting, hitte en koude, etc. zijn allemaal zaken die direct of indirect via de vegetatie werken. De vegetatie wordt sterk beïnvloed door – en heeft op haar beurt weer een sterke invloed op – grondwater en de aanwezigheid van diverse elementen in de bodem. De aanwezigheid van de geïntegreerde module SMART/SUMO in de NP is daarmee ook logisch. De vraag is dan nog of dat die module de belangrijke processen beschrijft op de gewenste tijd- en ruimteschalen.

Voor de biodiversiteit wordt verder gekeken naar de kans op voorkomen van soorten planten (MOVE), vlinders (VLINDERMOVE), en fauna (LARCH). Vanuit een functioneel oogpunt dekt dit de biodiversiteit in een behoorlijke mate, omdat het ofwel kernsoorten bevat (planten = vegetatie), gevoelige soorten (vlinders zijn veelal erg gevoelig voor veranderingen), of soorten die bovenin de voedselpiramide zitten (zoogdieren, vogels, reptielen). De laatste categorie is een goede indicator, omdat ze het eerste verdwijnt bij een achteruitgang in het leefgebied. Deze groepen dekken verder ook de biodiversiteit vanuit een pragmatisch oogpunt, omdat ze meestal beter te detecteren en/of te classificeren zijn dan andere soorten.

3E. Beoordeel de mate van 'evenwicht' m.b.t. de systeemanalyse, gebaseerd op de antwoorden op 3C en 3D (goed/matig/slecht).

Opmerkingen bij vraag 3E:

Om vraag 3E te beantwoorden is het nodig om de uitkomsten o.b.v. de vergelijkings-matrix uit vraag 2D/E mee te nemen. Er moet dus eigenlijk ook naar die vragen verwezen worden voor het geven van een initiële evenwichtsbeoordeling.

Laten we aannemen dat uit vraag 2D/E naar voren is gekomen, dat de doelen en toepassingen van de NP en de aparte modules goed overeenstemmen. Dit geeft ons een referentie voor het bepalen van het evenwicht. De vraag is dan of de relevante processen zijn meegenomen in de verschillende modules, en of er geen modules zijn die onevenredig veel 'zwaarder' zijn dan andere. Een onbalans in een modelketen als de NP zou bv. kunnen ontstaan, wanneer de ene module 25 dynamische processen meeneemt, terwijl de andere module slechts enkele correlaties afleidt, waarvan niet helder is hoe die dan in de tijd nog verder veranderen. Het aspect van datasteun mist dan nog in de analyse; dat komt aan bod in vraag 5 en 6.

5.5 Conceptueel model NP

4A. Geef het conceptuele model (voor een model) of ontwerpmodel (voor een bestand). Dit kan bv. een verwijzing naar een rapport zijn. Het conceptuele model legt op relatief informele wijze de relaties tussen componenten vast. Aspecten die aan bod kunnen komen zijn: Welke aannamen en vereenvoudigingen zijn er gemaakt? Wat zijn de ruimtelijke dimensies? Welke ruimtelijk-temporele aggregatieniveaus worden gebruikt? Welke zaken spelen een rol voor het bestand? Wat wordt er voor (des)aggregatie van data gebruikt? Voor modellen: Welk modeltype is gebruikt?

Het conceptuele model kan goed samengevat worden als een combinatie van Figuur 1 in Van der Hoek & Bakkenes (2007) met een korte beschrijving van de modules en hun rol in de NP. Zie in dit geval de uitbreiding op vraag 1 en Tabel 5.1. Er vinden uiteindelijk een aantal aggregatiestappen plaats; deze staan uitgelegd in Ten Brink *et al.*(2002), p. 20.

Beschouw nu per toepassing

4B. Op welke (ruimtelijk-temporeel) aggregatieniveaus zijn antwoorden gewenst, gezien de toepassing? Welke ruimtelijke dimensies en welk modeltype/ontwerpmodel zijn gewenst?

Er zijn antwoorden gewenst op landelijke schaal, in tijdstappen van een jaar, met projecties tot ca. halverwege de 21^{ste} eeuw. Dat wil niet zeggen dat deze uitvoer ook zo gevraagd en gegeven zou moeten worden. Gegevens worden in de NP eerst per fysisch-geografische regio en per soort geaggregeerd, daarna wordt een ongewogen middeling bepaald van de soortindexen van drie grote soortgroepen en uiteindelijk van de soortgroep-indexen. Het ontwerpmodel zou in theorie niets meer kunnen zijn dan de biodiversiteit, gemeten in bv. het gemiddeld aantal soorten per natuurtype, per jaar. Voor de inschatting van effecten van beleid en socio-economische ontwikkelingen is het echter van belang dat er processen worden meegenomen die de biodiversiteit beïnvloeden. Het model moet dus meer kunnen dan de gegevens over voorkomen aggregeren. Verder is het relevant om enige ruimtelijke differentiatie in het oog te houden. Een landelijke beleidsingreep kan bv. op ruimtelijk verschillende posities verschillende effecten hebben, die elkaar op landelijk geaggregeerd gezichtspunt opheffen. Bovendien moet dat dan ook temporeel bekeken worden. Men kan zich bv. voorstellen dat een effect van een ingreep op de korte termijn negatief lijkt te zijn, maar op de lange termijn positief. In dit licht bezien zou een dynamisch procesmodel in principe de voorkeur hebben.

4C. Komen de ruimtelijke dimensies, aggregatieniveaus, etc. overeen met die vanuit de toepassing zijn gewenst (ja/matig/niet)?

De NP aggregereert de gegevens tot een landelijke graadmeter, wat in overeenstemming is met wat de wens lijkt te zijn. Dat wil nog niet automatisch zeggen dat de stappen die vóór de aggregatie plaatsvinden ook 'goed' zijn; er kunnen bv. fouten gemaakt worden in deze stappen. Verder is het nadeel – zoals al boven genoemd – dat voor een goede beoordeling van effecten van landelijk beleid en (internationale) socio-economische ontwikkelingen het belangrijk is om ruimtelijk en temporeel te differentiëren.

Opmerkingen bij vraag 4:

Deze vragen lijken onverminderd van toepassing voor een modelketen.

5.6 Data bij de NP

5A. Welke gegevens zijn benodigd, gezien de toepassing, en motiveer? Wees specifiek m.b.t. resolutie, nauwkeurigheid, schaal, etc., en vermijd algemene en ongemotiveerde opmerkingen als "file X moet gebruikt worden".

5B. Welke gegevens zijn beschikbaar, en in welke mate?

5C. Wat is uiteindelijk de invoer van het model of bestand?

5D. Komt de invoer overeen met de door de toepassing gewenste gegevens (ja/deels/nee)? Vergelijk de antwoorden op 5A, 5B en 5C. Let hierbij sterk op eenheden en de drie elementen van schaal (coverage, support, extent; zie Bierkens et al, 2000; Bogaart et al, 2011). Heeft de data de juiste dimensies en eenheden? Zo nee, is er een (des)aggregatiemethode voorhanden? Schaalproblemen doen zich overal voor door de 'vertaling' van attributen naar metingen, die gediscretiseerd zijn in tijd en ruimte, en onderhevig zijn aan allerlei bronnen van meetruis.

Opmerkingen bij vraag 5:

Zoals al opgemerkt in de casus 'MetaSWAP' is het relevant om ook de data te beschouwen die nodig is in latere stappen in het modelleerproces, dus voor het formele model (voor suggesties voor dynamische formuleringen, voor het afleiden van correlaties, etc.), de numerieke implementatie (voor invoer, zoals scenario's en *forcing* data), de kalibratie (voor het bepalen van parameterwaarden), en de validatie (voor het onafhankelijk beoordelen of het model de werkelijkheid binnen de toepassing afdoende goed beschrijft). Er moet daarom onderscheid gemaakt worden tussen verschillende typen data (bv. invoer, kalibratie, procesbeschrijving, correlaties afleiden,...). Bovendien is het voor een modelketen van belang om te weten, welke data door welke modules gebruikt wordt. Het is in theorie goed mogelijk dat er voor een deel van de modules relatief veel gegevens beschikbaar zijn ter ondersteuning van hun complexiteit, terwijl andere modules relatief 'data-arm' zijn. Dit kan ook weer consequenties hebben voor de mogelijkheden tot kalibratie en validatie, maar ook voor de relatieve onzekerheden m.b.t. de uitvoer van de modelketen. Daarnaast is voor de 2^{de} Natuurverkenning gebruik gemaakt van verschillende scenario's als invoer van de NP; met data kan dus ook scenario's bedoeld worden.

De invoer van de verschillende modules is beschreven door Van der Hoek & Bakkenes (2007). Verder kunnen de Status A-lijsten gebruikt worden wat betreft de verwijzingen naar de in- en uitvoer (vragen over in- en uitvoer), en de beschrijvingen van de kalibratie en validatie (zie onderdelen 'kalibratie' en 'validatie'). Die verwijzingen geven nog niet *per se* een overzicht van alle data, noch geven ze informatie over eventuele schaalproblemen, informatie-inhoud van de data, of suggesties over of er een goede overlap is tussen de beschikbare gegevens en de gewenste gegevens. De vragen zijn dus erg relevant, zowel voor de NP in het geheel als de afzonderlijke modules, maar alleen vraag 5C is redelijk eenvoudig te beantwoorden.

In Tabel 5.4 wordt een overzicht gegeven van de in de NP gebruikte data o.b.v. de gegeven referenties. Daarnaast wordt aangegeven welke data, gezien de rol van de module, gewenst zou zijn. Omdat het een modelketen betreft waarin de uitvoer van de ene module ook (een deel van) de invoer van een andere module vormt, is het relevant om ook de uitvoer te noemen.

Tabel 5.4. Tabel voor het maken van een overzicht van de gegevens die in een model gebruikt worden of gewenst zijn. De laatste kolom geeft de mate van overeenstemming tussen gewenste data en gebruikte data. Een slechte overlap tussen de twee is een indicatie dat de dataondersteuning van het model te wensen overlaat, en dat het model wellicht niet evenwichtig is. De opmerkingen zoals die gemaakt zijn bij de casus 'MetaSWAP' gelden hier nog meer: voor het volledig invullen van deze tabel is het eigenlijk nodig om ook de behoefte aan gegevens te weten voor het formele model, de numerieke implementatie, kalibratie en validatie.

Model	Gebruikte data	Gewenste data	Match?
NP	4 invoerscenario's	Invoerscenario's (bv. projecties m.b.t. klimaat & socio-economische zaken); Natuurtypen; Data voor onafhankelijke validatie (is lastig voor graadmeter; toetsen tegen expert-kennis?), en valideren modules	
SMART2	Mol <i>et al</i> (2006): §2.3 invoerparameterwaardes §4.2 invoer: waterbalans, fluxes depositie, fysische en chemische eigenschappen, vegetatiedata §4.3 uitvoer: tijdreeksen van nutriëntbeschikbaarheid, balansen, concentraties, pH, productie van biomassa Kros (2002): §2.3 & §3.3 kalibratie & validatie	Gegevens voor het kwantificeren van processen en concentraties relevante moleculen en ionen; Grondwaterstanden; Gegevens voor kalibratie & validatie	
SUMO	Wamelink (2007): H4, bijlage 1 parameters Kalibratie & validatie De Vries <i>et al</i> (2007): kalibratie App. 5 & 6 parameters Wamelink (2008b): H2 parameters 4.2 invoer; 4.3 uitvoer	Betrouwbare schattingen aantallen en soorten vegetatie; Neerslag, begrazing, beheer; Gegevens voor kalibratie & validatie	
P2E/MOVE	Wamelink <i>et al</i> (2009c): §2.3 invoer; H4 uitvoer Van Adrichem <i>et al</i> (2010): §2.3-2.5 parameters H2-3 kalibratie H4 kruisvalidatie Wamelink <i>et al</i> (2001): Validatie (oudere versie MOVE)	Gegevens kwantificering processen; Frequenties planten plaats & tijd; Ecologische gegevens per natuurtype, (voedsel, verspreiding, voortplanting); Gegevens voor kalibratie & validatie	
VLINDERMOVE	Oostermeijer & Van Swaay (1998): Observaties soorten vlinders per vaste hokgrootte, in een bepaalde tijd	Frequenties vlinders plaats & tijd; Ecologische gegevens per natuurtype, (voedsel, verspreiding, voortplanting); Gegevens voor kalibratie & validatie	
LARCH	Pouwels <i>et al</i> (2008): §3.3 parameters; §3.4 invoer; §3.7 uitvoer H4, §5.1 kalibratie §5.3 validatie	Frequenties dieren plaats & tijd; Ecologische gegevens per natuurtype, (voedsel, verspreiding, voortplanting), vooral ook kritische dichtheden; Gegevens voor kalibratie & validatie	
BIODIV	Van der Hoek <i>et al</i> (2000): Polygonenkaart FGR; §5.2 Data FLORON, CBS	Natuurtypen; Uitvoer uit onderliggende modules; Gegevens voor kalibratie & validatie	

De meest rechtse kolom is om een kwalitatief oordeel te geven of er een goede mate van overeenstemming tussen de gebruikte en de gewenste data is. Als bij verschillende modules grote afwijkingen zijn, dan moet nog bekeken worden wat voor afwijkingen dit zijn. In het geval dat de gebruikte data de gewenste data volledig overlapt, dan kan dit eventueel betekenen dat het model veel complexer is dan nodig is, al hoeft dat niet het geval te zijn. In het geval dat de gebruikte data veel minder is dan wat gewenst is, dan kan dit op verschillende soorten gebreken duiden. Voorbeelden zijn:

- Er missen data van een bepaald type, bv. er zijn geen waterstanden, terwijl die wel nodig zijn;
- De data zijn wel van het goede type maar er is te weinig dekking, waardoor bv. een correlatieverband onbetrouwbaar is.

Niet alle mogelijk problemen worden afgedekt met deze tabel. Zo kan het zijn dat de data van het goede type zijn en in voldoende mate, maar dat de betrouwbaarheid te laag is. Bv., er zijn wel waterstanden, maar die zijn dusdanig slordig gemeten dat ze toch niet bruikbaar zijn voor invoer, kalibratie, of validatie, met het oog op de toepassing.

5.7 Formeel model NP

Gebruik de antwoorden op vragen 5A tot 5D om vragen 6B tot 6D te antwoorden.

6A. Geef het formeel model, en motiveer de keuze of verwijst naar paginanummers uit een gepubliceerd rapport of artikel. Het formeel model is de omzetting van concept naar bijvoorbeeld wiskundige vergelijkingen of een set rekenregels. De keuze van formeel model hangt af van de beoogde toepassing en de data, en wordt mede bepaald door de gestelde onzekerheidsmarges.

Het formele model is een vertaling van het conceptuele model, zoals hierboven kort beschreven. De NP bestaat uit verschillende modules, die verschillende formaliseringen van de concepten kennen. Een kort overzicht van de modules en (de referenties naar) hun beschrijvingen is gegeven in Tabel 5.5, samengesteld m.b.v. ingevulde Status A-lijsten (vragen over 'theorie'). Motivaties voor keuzes achter het opstellen van het formele model ontbreken meestal in de referenties, en moeten dus achterhaald worden d.m.v. interviews met de programmeurs.

Tabel 5.5. Tabel met een kort overzicht van de formele modules en referenties naar de volledige beschrijvingen.

Module	Formeel model & sleutelreferenties
SMART2	Enkel-Haags bodemverzuringmodel met nutriëntencycli. Partiële differentiaal- en algebraïsche vergelijkingen van relaties in- en uitvoer, en beschrijvingen van reactie-gelimiteerde en evenwichtsprocessen in de bodem (H, Al, divalente basiskationen, K, Na, NH ₄ , NO ₃ , SO ₄ , HCO ₃ , Cl). Mol <i>et al</i> (2006), §2.1 en 2.2
SUMO	Dynamisch model dat competitie voor licht en stikstof als de drijvende processen achter successie beschouwd, en groei biomassa en hoogte van 6 functionele groepen onder beheer beschrijft. Feitelijk een forward Euler met een tijdstap van een jaar, gecombineerd met een aantal sommaties (algebraïsche vergelijkingen). Wamelink <i>et al</i> (2009a), p. 2
P2E/MOVE	Statistisch model, waarin de kans op voorkomen van <i>planten</i> bepaald wordt o.b.v. afgeleide correlaties tussen soorten en 6 factoren (waaronder Ellenberggetallen). Wamelink <i>et al</i> (2009c), H1; Van Adrichem <i>et al</i> (2010), H1-2 (2.5 correlatieformule)
VLINDERMOVE	Statistisch model, waarin de kans op voorkomen van <i>vlinders</i> bepaald wordt o.b.v. afgeleide correlaties tussen soorten en 6 factoren (waaronder Ellenberggetallen). Oostermeijer & Van Swaay (1998), §2.3 & Figuur 3
LARCH	Genereert potentieel (niet actueel!) habitatnetwerk voor soorten onder de aanname van een constant landschap. Opdam <i>et al</i> (2003)
BIODIV	De NCI-methode: Ten Brink (2000) Graadmeter: Van der Hoek <i>et al</i> (2000), §1.3-3.2.2

- 6B. Welke gegevens zijn benodigd voor dit formele model?**
6C. Welke gegevens zijn beschikbaar, en in welke mate?
6D. Zijn de gegevens die benodigd én beschikbaar zijn ook gebruikt? Waarom wel/niet?

Opmerkingen bij vraag 6B, C en D:

Een kritische opmerking hier is in hoeverre een formeel model anders is dan een conceptueel model m.b.t. de evaluatie van het evenwicht, en dus m.b.t. veranderingen in de behoefte aan gegevens. Een conceptueel model kan bv. een diagram met boxen en pijlen zijn, terwijl het formele model dan de vertaling daarvan in bv. wiskundige vergelijkingen is. Bijvoorbeeld, we hebben een diagram met box 1 en box 2, en een pijl van box 1 naar 2. Box 1 stelt een populatie prooidieren voor, box 2 een populatie roofdieren. De pijl geeft aan dat populatie 2 van populatie 1 eet; de zgn. 'functionele respons'. De keuze van de uiteindelijke functionele respons kan sterk afhangen van de beschikbare gegevens voor de beschrijving van de functionele respons.

Het grootste punt m.b.t. de benodigde data in de NP is vermoedelijk t.a.v. de ruimtelijke dekking. De relaties voor in- en uitvoer in SMART, het afleiden van de correlaties in MOVE en VLINDERMOVE, en het in beeld brengen van het habitatnetwerk vereisen veel data, vooral wanneer onderscheid moet worden gemaakt tussen de verschillende kandidaat-relaties. Bv. om de correlaties in (VLINDER)MOVE af te leiden zijn per kilometer- of uur-hok observaties van alle soorten nodig. Hierbij moet gekeken worden hoe om te gaan met de vele 'nullen' (soort niet geobserveerd in het betreffende hok). Een tweede punt is dan hoe de afgeleide correlaties gevalideerd kunnen worden (vraag 11 EMC-lijst).

5.8 Numerieke implementatie NP

7A. Geef het numeriek rekenmodel/gegevensbestand (de implementatie in software), en motiveer de keuze of verwijst naar paginanummers uit een gepubliceerd rapport of artikel.

7B. Beschrijf de verificatie of verwijst naar paginanummers uit een gepubliceerd rapport of artikel. Verificatie moet niet verward worden met validatie. De eerste controleert de omzetting van formeel model naar code, de tweede toetst de resultaten van de code aan de werkelijkheid.

7C. Beschrijf en motiveer de keuze voor numerieke rekenmethoden, en geef de gebruikte discretisatie of verwijst naar paginanummers uit een gepubliceerd rapport of artikel. Veelal bestaat er bij numerieke rekenmethoden een afweging tussen rekensnelheid en –nauwkeurigheid.

Opmerkingen bij vraag 7:

N.a.v. de opmerkingen bij eerdere casussen is de kop van de vraag al vervangen voor "numerieke implementatie". Zoals is aangegeven in de casus 'MetaSWAP' ontbreekt in deze vraag nog het aspect m.b.t. testen. Testen zijn echter wel van belang om na te gaan of dat de implementatie voldoende is; een implementatie die niet voldoende is kan eigenlijk bij voorbaat ook niet goed (of beter: succesvol) gevalideerd worden. Vragen 7B en C moeten beide breder geïnterpreteerd worden, wat betekent dat ze ook opnieuw geformuleerd moeten worden. Verificatie verwijst in de praktijk meestal naar het toetsen van de uitkomsten van een numerieke methode tegen een 'utopia', bv. een formule of balans die eenvoudig genoeg is om ook expliciet (algebraïsch) op te lossen. Voor SMART en SUMO gaat dit zeker op, maar voor de correlatiemodellen (MOVE, VLINDERMOVE) niet. Voor die laatste gevallen zijn de vragen hier daarom anders geïnterpreteerd. Een 'verificatie' voor correlaties kan bv. bestaan uit het testen of gelijke correlaties worden afgeleid op basis van dezelfde data in een heel andere implementatie, bv. door 'handmatig' stappen na te rekenen in zoiets als Excel; bij het optreden van verschillen is in elk geval duidelijk dat er ergens fouten worden gemaakt. Verder mist de nadruk op

het aspect ‘toepassing’ ook in deze vragen. Bij de numerieke implementatie van een formeel model worden diverse keuzes gemaakt die bepalen wat voor soort software er uiteindelijk ontwikkeld wordt. Die keuzes (zouden sterk moeten) worden beïnvloed door de toepassing van het model. Het enige voorbeeld van een keuze dat nu wordt gegeven is in vraag 7C: “*tussen rekensnelheid en – nauwkeurig-heid*”. In de praktijk kan echter gemakkelijk blijken dat keuzes zijn gemaakt op basis van hele andere motieven, bv. omdat de programmeur de taal goed beheerst, of om financiële redenen. Vraag 7C kan ook eigenlijk samengevoegd worden met vraag 9D.

Hieronder is een tabel (Tabel 5.6) te vinden met details m.b.t. de numerieke implementaties van de modules, en de testen en verificaties daarvan, samengesteld m.b.v. de ingevulde lijsten voor Status A (vragen over ‘technische documentatie’, m.n. vraag A6: “Is er een globale beschrijving van de werking van het computerprogramma?”, en ‘verificatie en testen software’). De NP als geheel is geïmplementeerd in ArisFlow, met de bedoeling de opties voor gebruikers te vergroten (Van der Hoek & Bakkenes, 2007). Motivaties voor keuzes voor de technische implementatie zijn in de referenties meestal niet te vinden, en de tabel moet dus aangevuld worden met interviews met de programmeurs.

Tabel 5.6. Tabel met een kort overzicht van de numerieke implementaties en referenties naar de volledige beschrijvingen.

Module	Numerieke implementatie & sleutelreferenties
NP	In ArisFlow, zie Van der Hoek & Bakkenes (2007) Testen worden beschreven in Bakkenes <i>et al</i> (2003)
SMART2	Numerieke implementatie en gebruik zijn beschreven in Mol <i>et al</i> (2006). Verificatie en tests verlopen via test-sets, die centraal in versiebeheer zijn opgeslagen. Bakkenes <i>et al</i> (2003), §2.1 & 3.1: Testen
SUMO	Wamelink (2008b), H2: beschrijving; H4: testen Wamelink <i>et al</i> (2008), H3 Bakkenes <i>et al</i> (2003), §2.1 & 3.1: Testen
P2E/MOVE	Stappen MOVE4 beschreven in Van Adrichem <i>et al</i> (2010), App. 1-4. Wamelink <i>et al</i> (2009c), H3: gebruikersbeschrijving MOVE4 Wamelink (2009) beschrijft testen MOVE4: dimensieanalyse, afvangen fouten, code controle, testberekening Bakkenes <i>et al</i> (2003), §2.2-2.3 & 3.2-3.3: Testen P2E-MOVE 3.2
VLINDERMOVE	Bakkenes <i>et al</i> (2003), §2.4 & 3.4: Testen
LARCH	Pouwels <i>et al</i> (2008), H3, inclusief testen. LARCH bestaat overigens op zijn beurt ook weer uit modelcomponenten, waarvoor de lijst doorlopen zou kunnen worden.
BIODIV	Implementatie in Delphi (Normenkader). Bakkenes <i>et al</i> (2003), §2.5 & 3.5: Testen

5.9 Schematisering NP

De schematisering is bv. het indelen van de bodem in verschillende lagen ('stratificatie'), of het clusteren van landgebruikstypen. De mate van detail die bij schematisering wordt gevolgd is afhankelijk van de toepassing, maar moet wel voldoende ondersteund worden door de data.

8A. Indien van toepassing, beschrijf de schematisering, of geef relevante en specifieke referenties?

De NP is een 'ruimtelijke' modelketen, die werkt met een horizontale schematisering van cellen van 250 bij 250 m², en tijdstappen van een jaar. Binnen een cel functioneren alle modellen als puntmodellen, dat wil zeggen de in- en uitvoer zijn 'homogeen' geldig binnen een cel. De NP is niet landsdekkend; het beschrijft alleen terrestrische natuurgebieden, die zo'n 7% van het landoppervlak van Nederland uitmaken. De module SMART binnen de NP beschrijft nog een derde ruimtelijke dimensie; de partiële differentiaalvergelijkingen beschrijven een afhankelijkheid voor de bodemdpte.

**8B. Wordt de schematisering voldoende ondersteund door de data (ja/matig/nee)?
Motiveer.**

Zoals al genoemd bij vraag 6B-D is er een grote databehoeftte om de ruimtelijke schematisering te kunnen beschrijven. SMART bv. wordt gekalibreerd per natuurtype, wat een grote databehoeftte met zich meeneemt. Oostermeijer & Van Swaay (1998) merken op, dat VLINDERMOVE alleen maar geldig is op landelijk geaggregeerd niveau, i.v.m. het clusteren van alle observatiedata van vlinders voor het afleiden van de correlaties. Dit duidt niet op een optimale ondersteuning van een ruimtelijk expliciet model met data. Op deze manier zouden alle modules afgelopen moeten worden.

5.10 Gevoeligheden, onzekerheden, opbouw code NP

Deze vraag is gericht op het bepalen welke parameters, forcing, invoer, delen van de code, etc. belangrijk zijn en welke overbodig zijn, en welke onderdelen een grote onzekerheid introduceren.

9A. Zijn er gevoeligheids-/onzekerheidsanalyses uitgevoerd? Zo ja, beschrijf deze of verwijst naar paginanummers uit een gepubliceerd rapport of artikel en analyseer de resultaten. Sommige factoren kunnen weinig van belang blijken voor de toepassing, zodat er weinig moeite hoeft te worden gedaan om deze goed te kalibreren, of gebruikt kunnen worden om het model te vereenvoudigen. Andere factoren kunnen juist essentieel blijken.

Een overzicht van gevoeligheids- en onzekerheidsanalyses wordt gegeven in Tabel 5.7.

Tabel 5.7. Tabel met een kort overzicht van de gevoeligheids- en onzekerheidsanalyses t.a.v. de NP met referenties naar de volledige beschrijvingen.

Module	GA/OA & sleutelreferenties
NP (geheel)	Van der Hoek & Heuberger (2006): SA SMART/SUMO-P2E-MOVE Wamelink <i>et al</i> (2011): OA SMART2-SUMO-P2E-MOVE4
SMART2	Kros (2002; onderliggende referentie Kros <i>et al</i> , 1993) OA van RESAM, voorloper van SMART. Mol <i>et al</i> (2006), §2.5: SA loofbos op zand. Verder als onderdeel van de NP.
SUMO	Wamelink (2008a): GA en SA
P2E/MOVE	Wamelink <i>et al</i> (2009c), §6.1
VLINDERMOVE	Geen (zover duidelijk is)
LARCH	Pouwels <i>et al</i> (2008), §5.2
BIODIV	Geen (zover duidelijk is)

Opmerkingen bij vraag 9A:

Er wordt gevraagd naar de referenties van de gevoeligheids- en onzekerheidsanalyses (GA/OA), en ook "analyseer de resultaten". Deze vraag zou duidelijker gesteld kunnen worden, met iets meer nadruk op wat de resultaten van de GA/OA zijn, en hoe die bruikbaar kunnen zijn voor de evenwichtsanalyse. M.b.t. dat laatste is het relevant om ook de data en toepassing goed te beschouwen. Bv., de vraag zou geherformuleerd kunnen worden als: "Geef de referenties naar eventuele gevoeligheids- en/of onzekerheidsanalyses. Wat waren de resultaten van deze studies? Welke parameters, invoer, factoren, etc. zijn gevoelig of onzeker, en welke niet? Welke zouden eventueel weggelaten kunnen worden? Motiveer dit, met name in het licht van de toepassing en de beschikbare gegevens".

9B. (niet voor databestanden) Welke numerieke integratiemethode is gebruikt? Bespreek het prestatievermogen van de methode.

Opmerkingen bij vraag 9B:

Zoals al opgemerkt bij de casus 'MetaSWAP' vervalt deze vraag.

9C. Wat is de rol van elke module van het model/bestand, bv. een gewasgroei-deelmodel als onderdeel van een hydrologisch model, of stappen in het stroomschema of de scripts van een bestand? Bespreek dat in relatie met de toepassing. Zijn er onderdelen die overbodig zijn? Missen er nog onderdelen? Motiveer waarom.

Opmerkingen bij vraag 9C:

Deze vraag is zeer relevant voor een modelketen, maar de positie lijkt in dit geval 'te laat'. De bedoeling van de vraag in de lijst t.a.v. een 'gewoon' model is om de complexiteit van de code onder de loep te nemen, met het oog op mogelijke reductie of uitbreiding/toevoeging van onderdelen. In dit geval hebben we deze vraag deels al beantwoord met het overzicht dat is toegevoegd bij vraag 1, inclusief Tabel 5.1. De vraag is echter niet 'nutteloos', omdat er aandacht is voor het aspect van de toepassing. De vergelijking tussen de NP en de mNP is nu precies waar het om gaat bij deze vraag. Immers, de mNP is bedoeld als een versimpeling van de NP. Inhoudelijke opmerkingen t.a.v. dit kunnen echter ook bij 9D geplaatst worden.

9D. Beoordeel de complexiteit in termen van rekentijd en efficiëntie. Wat is de duur van een run? Hoeveel (tussentijdse) uitvoer is er? Hoe groot is de datastroom? Zou de code sneller/efficiënter kunnen, bv. door selectie, aggregatie, andere methode, andere modulaire opbouw, etc.?

Opmerkingen bij vraag 9D:

Ook hier is het relevant om in de vraag expliciet de nadruk op de toepassing te leggen, terwijl de vraag nu nogal technisch ingestoken is, door de nadruk te leggen op de rekentijd en efficiëntie. De toepassing speelt een cruciale rol t.a.v. de keuzes die gemaakt worden. Bv., een versnelling van de code is altijd zinnig, maar het heeft alleen prioriteit als dit voor de toepassing van belang is. De opmerkingen die bij deze vraag gemaakt worden zijn zeer essentieel voor de evenwichtsanalyse t.a.v. de numerieke implementatie. Bovendien is de numerieke implementatie het product waar de gebruikers mee werken (i.t.t. bv. het formele model, dat in de praktijk vooral alleen door de ontwikkelaars wordt beschouwd), wat het belang van deze vraag extra onderstreept.

Zoals bij vraag 9C is opgemerkt, is de vergelijking tussen de mNP en de NP feitelijk ook gericht op de vraag hoeveel complexiteit nu nodig is voor de toepassing, in dit geval m.n. het gebruik van het model als graadmeter voor biodiversiteit in de Natuur-verkenning. Een aspect van 'evenwicht' is de complexiteit van de code, en het uiteindelijke doel van deze vraag is om tot inzichten te komen hoe deze vorm van complexiteit verminderd kan worden (met het oog op winst in snelheid of inzicht in de werking van het programma), terwijl er nog altijd voldaan wordt aan de eisen van de toepassingen. Een aardig detail bv. is de reductie van de regels code van de module SUMO met 20% in een herprogrammeringsproject binnen de Kwaliteitsslag (Wamelink et al, 2008, p. 11), dat naast snelheidswinst ook heeft geleid tot een grotere betrouwbaarheid en inzichtelijkheid.

5.11 Kalibratie NP

De kalibratie is het proces waarbij de parameters, begincondities, etc. van het model of bestand van waarden worden voorzien. Kalibratie is vaak gebaseerd op gevoeligheids-analyses (zie vorige vraag).

10A. Indien beschikbaar, beschrijf de uitgevoerde kalibraties en motiveer de keuze, of verwijz naar paginanummers uit een gepubliceerd rapport of artikel. Geef ook aan welke doelfuncties (lokaal, globaal, deterministisch, stochastisch) zijn gekozen, of dat de betrouwbaarheid van de gekalibreerde parameters wordt gekwantificeerd, hoe, op welke data, of en hoe de betrouwbaarheidsintervallen zijn bepaald, etc. met eventuele verwijzingen naar de literatuur. Let op over-fitting, meervoudige lokale optima, en identificeerbaarheidsproblemen.

In Tabel 5.8 wordt een overzicht gegeven van de referenties m.b.t. kalibratie, samengesteld m.b.v. ingevulde lijsten voor Status A (vragen over 'kalibratie'). Een algemeen geldende opmerking voor de NP is, dat kalibratie lastig is omdat langjarige tijdsreeksen ontbreken. Dit geldt m.n. voor de procesmodellen, die de eerste stappen in de modelketen (SMART, SUMO) vormen.

Tabel 5.8. Tabel met een kort overzicht van de formele modules en referenties naar de volledige beschrijvingen.

Module	Kalibratie & sleutelreferenties
SMART2	Kros (2002), p. 106-111; p. 122-127. Gebruik PEST (Doherty, 2005); doelfuncties beschreven.
SUMO	SUMO onderscheid 6 vegetatietypes. Het model wordt gekalibreerd voor elke combinatie van vegetatietype en functioneel planttype. Kalibratie beschreven in Wamelink <i>et al</i> (2009a), §3.
P2E/MOVE	Van Adrichem <i>et al</i> (2010), H2: aanpak, H3: resultaten.
VLINDERMOVE	Oostermeijer & Van Swaay (1998)
LARCH	Pouwels <i>et al</i> (2008), §5.1
BIODIV	Niet van toepassing: alle kalibratie vindt plaats in de onderliggende modellen.

10B. Welke eisen stelt de toepassing aan de nauwkeurigheid waarmee parameters, etc. wordt bepaald. Let ook op over-fitting.

De NP is een graadmeter van de biodiversiteit op landelijk niveau. Zolang het ook als zodanig wordt gebruikt zullen veel kleine variaties of afwijkingen tegen elkaar wegvallen bij de aggregaties. Bovendien zijn er ook marges t.a.v. de biodiversiteit, bv. veel soorten hebben een behoorlijk bereik van a-biotische waarden waarbinnen ze kunnen gedijen. Daarnaast is de vraag wat de afwijkingen betekenen t.o.v. de processen die niet (goed) meegenomen zijn in de modellering. Dit zal betekenen dat er voor veel specifieke parameters, etc. marges zijn t.a.v. de benodigde nauwkeurigheid.

10C. Welke mogelijkheden leveren de data om waarden van parameters, etc. te bepalen. Denk hierbij aan resolutie, nauwkeurigheid, aantal gegevens, het aggregatieniveau van de gegevens, en let op over-fitting.

Een zinnige manier om hier tegenaan te kijken, is een vergelijking tussen de gewenste marges – bepaald door de toepassing, zoals afgeleid in de vorige vraag – en de haalbare marges, die wordt bepaald door de beschikbare gegevens en hun betrouwbaarheid. Indien de gewenste marges door de haalbare marges overlapt worden, dan is er geen probleem m.b.t. dit aspect van evenwicht (let wel, dit zegt nog niets over eventuele over-fitting of over-complexiteit van het model). Andersom is er wel een mogelijk groot probleem, wanneer de gewenste marges niet door de data afgedekt worden. Een mogelijke aanpak is dan om de ondersteuning met data te verbeteren, het bereik van de toepassing aan te passen, en/of de complexiteit van het model te reduceren.

In een aantal gevallen is de ondersteuning met data bij de NP een lastig punt. We herhalen hier bv. de opmerking van Oostermeijer & Van Swaay (1998) m.b.t. de geldigheid van VLINDERMOVE.

5.12 Validatie NP

Wij definiëren hier een validatie als een toets of het model of bestand een redelijke representatie is van het werkelijke systeem dat wordt gemodelleerd. Dit geschiedt op basis van een set onafhankelijke gegevens (gegevens die nog niet in een eerder stadium bij het modelleren gebruikt zijn).

11A. Indien beschikbaar, beschrijf de gedane validatiestudies of verwijst naar paginanummers uit een gepubliceerd rapport of artikel.

11B. Onderbouw/beoordeel de waarde van de validatie(s). Hoe relevant zijn de verschillende validatiestudies voor de verschillende toepassing(en) van het model of bestand? Zijn alle toepassingen voldoende afgedekt door de studies? Is ook de betrouwbaarheid van de validatie gekwantificeerd? Welke toepassingen moeten opnieuw worden gevalideerd?

De NP is eigenlijk een 'graadmeter' van de biodiversiteit. Procesmodellen bv. zijn redelijk te valideren (in theorie althans), wanneer ze een proces beschrijven op een bepaalde plek, dat ook geldig is op een andere plek of op een ander tijdstip. Gegevens van de verschillende plekken of tijdstippen kunnen voor verschillende zaken gebruikt worden, bv. de ene dataset is gebruikt voor de kalibratie, en een andere voor de validatie. Voor graadmeters ligt dit lastiger, omdat er meestal geen direct vergelijkingsmateriaal is. Vergelijk dit met de AEX-index, die een graadmeter van de stand van de economie is, maar die niet met directe metingen van de 'AEX' is te valideren. Validatie vereist dan een creatievere aanpak, bv. vergelijking van de AEX-waarden in het verleden tegen de koers van de olieprijs. De "validatie" behelst dan een meer kwalitatieve dan kwantitatieve toetsing van het model. De NP als geheel is niet gevalideerd (A. van Hinsberg, mondelinge mededeling).

Indien de modelketen niet in zijn geheel is gevalideerd, dan is de beste optie dat in elk geval de verschillende modules apart zijn gevalideerd, bij voorkeur in de context van dezelfde toepassing. Hieronder is een tabel (Tabel 5.9) te vinden met referenties naar de validatiestudies van de aparte modules, samengesteld m.b.v. Status A-formulieren (vragen over 'kalibratie'). De beschouwing van de validatiestudies in het licht van de toepassing behoeft in een aantal gevallen nog verdere uitwerking.

Tabel 5.9. Tabel met een kort overzicht van de validatiestudies en referenties daarnaar.

Module	Validatiestudies & sleutelreferenties, bevindingen
SMART2	Validatie tegen gemanipuleerde punten in stroomgebied Risdalsheia met naaldwouden (Kros, 2000, p. 111-116), tegen een intensief gemonitorde site met sparren (Kros, 2000, §2.4.3-.4), en in andere studies, inclusief tegen andere modellen van verschillende complexiteit. SMART lijkt redelijk te functioneren voor lange-termijn trends in bodemchemie.
SUMO	Wamelink <i>et al</i> (2009a), §4.
P2E/MOVE	Wamelink <i>et al</i> (2009c), §6.2: Percentage verklaarde variantie, Pearson correlatie, en regressieanalyse. Van Adrichem <i>et al</i> (2010), H4: Kruisvalidatie
VLINDERMOVE	Geen specifieke validatiestudie bekend.
LARCH	Pouwels <i>et al</i> (2008), §5.3. Voor verschillende soorten zijn aanpassingen voorgesteld in parameterwaarden of invoer.
BIODIV	Van der Hoek <i>et al</i> (2000) §5.3 vergelijking van uitkomsten met referentieanalyses, en H6 binnen de toepassing VIJNO (Vijfde Nota Ruimtelijke Ordening).

Opmerkingen bij vraag 11:

Ook hier is het relevant om zowel naar de validatiestatus van de aparte modules als van de modelketen als geheel te kijken. Additionele aspecten die hier aan bod zouden moeten komen zijn:

- Zijn alle modules gevalideerd? Welke wel/niet?
- Zijn de modules gevalideerd in de context van de toepassing van de modelketen?
- Is de modelketen zelf gevalideerd in de context van de toepassing? En hoe?

5.13 Evenwichtsoordeel Natuurplanner

Dit laatste onderdeel is gericht op het geaggregeerde resultaat van de antwoorden op de verschillende vragen en de meningen van de verschillende betrokkenen, en kan indirect ook tot adviezen leiden ter verbetering van het model of bestand.

12A. Beoordeel de mate van vertrouwen in het model/bestand in het algemeen, mede gebaseerd op wat er is gedaan aan testen, verificatie, kalibratie, validatie, gevoeligheids- en onzekerheidsanalyses, en motiveer deze.

12B. Zijn er veel missende zaken? Zijn er veel vragen waar de zaken als 'te eenvoudig' zijn beoordeeld? Motiveer.

12C. Zijn er veel overbodige componenten? Zijn er veel onderdelen waar het eenvoudiger zou kunnen? Motiveer.

12D. Is er voldoende datasteun, en waaruit blijkt dat? Motiveer

E. Welke specifieke suggesties voor verbetering en data-aanlevering komen naar voren uit deze analyse?

Het inhoudelijk evenwichtsoordeel van de NP, en de mNP, wordt besproken in een later te verschijnen WOT-werkdocument (Van Voorn & Walvoort, 2013).

6 Discussie

6.1 Inleiding

In dit laatste hoofdstuk kijken we naar de bevindingen die per casus zijn gedaan t.a.v. de verschillende vragen op de evaluatielijst EMC v1.0 (compleet weergegeven in Bijlage 1) en de verdere bevindingen, bv. de ervaringen bij het invullen van de lijst. Daarna volgen er nog conclusies, vooruitzichten en suggesties voor verder onderzoek. Veel punten ter verbetering van de lijst zijn opgenomen in de voorgestelde versie 2.0 van de EMC-lijst, die in bijlage 2 is weergegeven. Daarbij willen we benadrukken dat dit een voorstel betreft, en dat er geen sprake is van de 'definitieve' EMC.

6.2 Punten ter verbetering van de lijst

Uit de verschillende casussen zijn meerdere punten naar voren gekomen waarop de EMC v1.0 te verbeteren valt. Bij deze opmerkingen beschouwen we tevens nog even de punten die opgemerkt zijn bij de evaluatie door experts op de eerste versie van de lijst EMC v0.1 (Van Voorn & Walvoort, 2011), om te kijken of ook aan deze punten inmiddels tegemoet is gekomen. Aangezien voor het merendeel van de vragen de volgorde en opbouw wel in orde is bevonden, behandelen we al deze punten per vraag, met suggesties hoe de vraag verbeterd kan worden.

Algemene suggesties voor de EMC v2.0:

- Leg uit dat de term 'model' breed geïnterpreteerd dient te worden; het kan immers model, bestand, meta-model, modelketen, of een graadmeter inhouden.
- Maak onderscheid tussen de verschillende niveaus van 'vragen' om de communicatie te vergemakkelijken. De nummers 1 t/m 12 kunnen aangeduid worden met 'onderdeel', de letters A t/m X met 'deelvragen', en de vragen per deelvraag met bv. 'detailvragen'.
- Leg duidelijk uit wat met 'evenwicht' bedoeld wordt, en dat de lijst een middel is om te bevorderen dat een evenwichtig model wordt bereikt.

1. BASALE GEGEVENS

Geef aan: model/bestand, naam, versienummer, revisienummer, uitgiftedatum. Onder model verstaan we de broncode + invoergegevens.

In de casus 'Wageningen-model' werd opgemerkt dat deze vraag, hoewel nuttig ten behoeve van dossiervorming, geen wezenlijke bijdrage aan de evenwichtsevaluatie levert. Deze opmerking is terecht, maar duidt er ook op dat deze vraag uitgebreid zal moeten worden. De volgende punten zouden overwogen moeten worden voor opname binnen deze vraag, dan wel in een toelichting bij deze vraag:

- Het is de vraag, of de lijst nu slaagt in haar bedoeling om de benodigde informatie voor de beoordeling van 'evenwicht' te leveren. Het is vermoedelijk zinnig om expliciet aan de lezer uit te leggen dat het gaat om de oorspronkelijke keuzes achter alle modelleerstappen te evalueren in de context van toepassing en dataondersteuning, om modelleers als het ware "los te weken" van eventuele automatismen t.a.v. het modelleerwerk.
- Om de vraag te beantwoorden wat 'evenwicht' behelst, is het ook van belang om uit te leggen wat 'complexiteit' inhoudt, en wat onder 'data' en 'doel/toepassing' wordt verstaan. Een aantal punten uit het review-commentaar uit Van Voorn & Walvoort (2011) komt daarmee ook aan bod. Er kan bv. enige aandacht geschonken worden aan de verschillende definities (>30) en aspecten van complexiteit.

- Zoals al opgemerkt in de review-commentaren, is de context van de geschiedenis nog niet opgenomen in de lijst. Informeel is dit wel meegenomen in verschillende casussen, en daaruit blijkt dat de historische context van belang is om een oordeel over 'evenwicht' te maken. In de review-commentaren is voorgesteld dat een graduele toename in de complexiteit van het model waarschijnlijk de beste resultaten m.b.t. 'evenwicht' geeft. Deze voorgestelde aanpak is overeenstemmend met de veelgebruikte aanpak in modelselectie, waarin de complexiteit van een model steeds vergroot wordt terwijl de fit op de data wordt geëvalueerd. In de praktijk begint een casus echter meestal met een numerieke code van een versie die later dan de eerste versie is, i.p.v. een systeemanalyse en doelstelling. Op de achtergrond speelt dan ook nog het zogenaamde "Concorde-effect" (doorgaan op de ingeslagen weg, ook al is dat niet langer rationeel verdedigbaar), dat conservatisme t.a.v. het model in de hand werkt. Het is dus waarschijnlijk beter om te vragen naar het model, alle bestaande versies, een kort overzicht van de verschillen daartussen, en de motivaties achter de verschillen.
- Er werd in review-commentaar aangegeven dat er niet *per se* een duidelijk 'evenwicht' hoeft te bestaan. Deze verwachting wordt gemotiveerd uit het feit, dat de beoordeling t.a.v. modelcomplexiteit op vele criteria is (of kan zijn) gebaseerd. Onze verwachting is wel, dat door het minimaal eenmaal doorlopen van de modelleercyclus a.d.h.v. de lijst, de modelcomplexiteit evenwichtiger zal zijn dan bij het niet doorlopen van de lijst. Bovendien kan het gebruik van de lijst altijd nog leiden tot het blootleggen van pijnpunten, waarop het model verbeterd kan worden.
- Voor verschillende soorten modellen en databestanden zijn verschillende vragen meer of minder van toepassing (bv. in de casus 'MetaSWAP', of een modelketen als de NP), en evenzo is gebleken dat de volgorde niet altijd optimaal is. Dit zou ondervangen kunnen worden door in vraag 1 expliciet te vragen naar wat de te analyseren casus is, en de analist de keuze te laten motiveren. Een alternatief is om in elk geval aan te geven, dat per casus niet alle vragen even relevant hoeven te zijn.
- In de interviews bij een aantal casussen is ook voorgelegd aan de geïnterviewde personen wat de *a priori* inschatting t.a.v. 'evenwicht' was. Gegeven dat de term voldoende uitgelegd is, kan deze vraag in de lijst opgenomen worden. Het biedt dan informatie of, n.a.v. de beantwoording van de lijst, de originele inschatting gewijzigd is. Een alternatief is om te vragen naar bekende "pijnpunten" of zwakke plekken; dit is weliswaar niet helemaal hetzelfde als het beantwoorden van de vraag of het model in evenwicht is, maar een dergelijke vraag is vermoedelijk wel herkenbaarder voor modelleers.

Oplossing/aanbeveling: Kort samengevat, lijkt het zinnig om bij deze vraag een korte *introductie* toe te voegen over het onderwerp en doel van de lijst, en de vraag in elk geval *uit te breiden* zodat het de (noodzakelijke) informatie oplevert t.a.v. de historische context en de keuzes die in het verleden gemaakt zijn t.a.v. het model en modelgebruik.

Een specifiek knelpunt is de bevinding in de NP-casus dat het nodig is om een kort overzicht op te stellen van de verschillende modules in de modelketen. Dit knelpunt ontstaat omdat, zoals al gezegd, de analyse plaatsvindt op het product 'Natuurplanner', dat al een geschiedenis heeft en een numerieke implementatie betreft.

Oplossing/aanbeveling: Dit kan wellicht afgedekt worden door uitbreiding van de vraag zoals in de NP-casus voorgesteld: "*Indien de casus een modelketen is, beschrijf ook kort uit welke modules de modelketen bestaat*".

Verder is er nog een opmerking geplaatst: "*Wie is verantwoordelijk voor de code?*". Hoewel dit erop wijst wie de veranderingen gaat doorvoeren, heeft deze vraag geen betrekking op de inhoudelijke analyse m.b.t. 'evenwicht'. Dit is meer een vraag voor de borging van modellen, en wordt niet meegenomen in de actualisatie van de checklist.

2. DOEL & TOEPASSINGEN

2A. Wat was het doel van het bestand/model?

2B. Wat was het beoogde toepassingsgebied, en waaruit blijkt dit?

2C. Wat zijn de toepassingen nu? Heeft het model/bestand een rol in een modelketen? Zo ja, welke?

2D. Overlapt het beoogde toepassingsgebied alle daadwerkelijke toepassingen (ja/deels/nee)?

2E. Indien relevant, in hoeverre overlappen de verschillende toepassingen met elkaar (goed/matig/slecht)?

Er is onduidelijkheid over de interpretatie van 'doel' in vraag 2A. Het 'einddoel' is wat anders dan het 'modeldoel' (bv. gebiedsinrichting en neerslag-afvoermodellering). Dit is een begrijpelijke opmerking, die erop wijst dat er onderscheid gemaakt moet worden tussen het model zelf en zijn 'reden van bestaan'. Dit onderscheid is terecht, maar dit wijst op een semantisch punt, aangezien de term 'beoogde toepassing' deze laatste categorie dekt. In bovenstaand voorbeeld is de functie van het model de neerslag-afvoermodellering, en de gebiedsinrichting de studie waarin het model een rol moet gaan spelen, dus de beoogde toepassing. Zodra het model dan ook is toegepast in de studie, dan valt die studie onder het feitelijke toepassingsgebied van het model (vraag 2C).

Oplossing/aanbeveling: Voeg een kleine toelichting hierover bij.

Er is opgemerkt dat beoogd en feitelijk toepassingsgebied bij elkaar horen, maar dit is nadrukkelijk niet het geval. De reden waarom deze scheiding niet goed opgaat, is omdat het model in de historische context bekeken moet worden, dus: Wordt het model op dit moment toegepast waar het origineel voor ontwikkeld was? Als dit niet het geval is, dan moet gevraagd worden naar een betoog waarom dit gebruik gerechtvaardigd is. Vraag 2C is niet relevant, indien de checklist wordt gevolgd bij de 'geboorte' van een model.

Oplossing/aanbeveling: Een kleine toelichting zou dit kunnen verhelderen. Het is waarschijnlijk beter om de term 'beoogd' te vervangen voor de term 'oorspronkelijk'.

Er wordt in de lijst gevraagd of het beoogde toepassingsgebied alle toepassingen omvat (2D). Dat wil nog niet zeggen dat die toepassingen ook elkaar (grotendeels) overlappen, het punt waarnaar gevraagd wordt in vraag 2E. Dit laatste punt is wel relevant om vast te stellen: zou blijken dat de toepassingen elkaar niet overlappen, dan kan het wellicht verderop een nuttige suggestie blijken om voor elke toepassing een aparte modellijn te creëren. Een reden hiervoor zou kunnen zijn dat het model voor elke toepassing opnieuw gekalibreerd moet worden, en deze kalibratie veel kosten met zich meebrengt. Bovendien wordt er in een aantal vragen verderop onderscheid gemaakt tussen de verschillende toepassingen, indien er meerdere toepassingen zijn. Overlappen de verschillende toepassingen elkaar sterk, dan is het mogelijk om dit onderscheid in de praktijk te negeren. Deze aspecten – het expliciet vaststellen of er verschillende toepassingen zijn en of ze ook afzonderlijk behandeld moeten worden of niet – komen nog niet erg nadrukkelijk in de vragenlijst naar boven.

Oplossing/aanbeveling: Voeg een vraag toe: "*Zijn er (dus) meer dan 1/verschillende toepassingen die in vragen verderop apart behandeld/beschouwd moeten worden?*". Verder kan expliciet worden opgemerkt dat het de bedoeling is om per toepassing de gehele evaluatielijst opnieuw te doorlopen. Het is namelijk een reële mogelijkheid dat er per toepassing een duidelijk ander evenwicht bestaat.

De beschrijving van de toepassing zou eigenlijk ook in kwantitatieve termen moeten plaatsvinden. Vraag 2E richt zich op het beoogd en feitelijk toepassingsgebied van het te evalueren model, maar dit is nog redelijk ambigu. Voor de bepaling van 'evenwicht' is het van belang om te weten, of het model kan voldoen aan de eisen vanuit de toepassing, en of dat de eisen van de toepassing niet veranderen. Een model kan zonder grote problemen worden ingezet in een andere toepassing (bv. een tweede gebiedsstudie), mits de eisen aan het model daarvoor niet veranderen. Indien dit wel het geval is, dan moet uit validatie blijken dat het model aan de nieuwe eisen kan voldoen. Een praktisch

probleem is wel, dat verschillen kunnen zitten in subtiele zaken als acceptabele normen en onzekerheden.

Oplossing/aanbeveling: Een toelichting met enige uitleg m.b.t. de kwantificering van de toepassing, zoals: "Welke rol gaan de modelresultaten spelen in het verdere proces? Welke eisen ten aanzien van variabelen, acceptabele onzekerheden, tijd- en ruimte-schalen komen hieruit voort? Onder welke voorwaarden kan het beoogde model hier aan voldoen?". Dit kan bijgevoegd worden aan de extra vraag die voorgesteld is in de vorige alinea.

In vraag 2C horen de twee vragen eigenlijk niet bij elkaar. De vraag, of een model een rol in een modelketen speelt, is wel terecht, maar zou apart gesteld moeten worden.

Oplossing/aanbeveling: Er moet een aparte vraag komen: "*Heeft het model/bestand een rol in een modelketen? Zo ja, beschrijf deze*", die eenvoudig als NVT beantwoord kan worden.

Nergens in de lijst wordt expliciet gevraagd naar de uitvoer van het model of de modelketen, terwijl dit feitelijk meestal precies is waar het om gaat. Er zou dus een expliciete vraag moeten worden geformuleerd t.a.v. de uitvoer van het model, wat deze is, wat gewenst is vanuit de toepassing, en of de uitvoer van het huidige model past bij de toepassing. Immers, deze vraag is de kern van de evenwichtsbeoordeling!

Oplossing/aanbeveling: Er moet een expliciete vraag naar de uitvoer van het model komen. De meest relevante plaats waar de wensen t.a.v. de uitvoer van het model genoemd zou kunnen worden, is al bij het aangeven van doel en toepassing van het model in vraag 2; vraag X: "*Wat is de gewenste uitvoer van het model, gezien doel en toepassing? Motiveer*". Het gebruik van het model bepaalt (of zou moeten bepalen) wat de uitvoer van het model is. Verderop (bij vraag 7 of 9) moet dan nog een additionele vraag komen of de uitvoer van het model ook inderdaad is wat het zou moeten zijn met het oog op 'evenwicht': "*Wat is de uitvoer van het model?*", en: "*Beschouw het antwoord op vraag X. Klopt de uitvoer van het model met de wensen vanuit doel en toepassing? Waarom wel/niet?*".

Een aantal punten zijn specifiek opgekomen in de NP-casus. Voor een modelketen is het zinnig om *per module* de verschillende vragen te beantwoorden. Doel, beoogd toepassingsgebied, en feitelijk toepassingsgebied *naast* de rol van het model in de modelketen kunnen per module verschillen, en hoeven niet te overlappen met de toepassing van de modelketen. In de casus bleek het zinnig om een matrixtabel samen te stellen en per module doel en toepassing in te vullen, en daarna de overlap kritisch te beschouwen.

Oplossing/aanbeveling: Voeg de vraag toe: "*Wat is de overlap in doel, beoogde en daadwerkelijke toepassing tussen de modules van de modelketen en de modelketen zelf?*" met de suggestie om een matrixtabel samen te stellen, of in elk geval steeds te kijken naar de overlap tussen de uitvoer van een model met de invoer van het daaropvolgende model in de modelketen. Wanneer er duidelijke discrepanties zijn, dan is de aanbeveling de toepassing van de modelketen terug te brengen tot het punt dat de discrepanties verdwijnen (in het ergste geval betekent dat, dat de modelketen niet langer geschikt is voor een toepassing).

3. SYSTEEMANALYSE

3A. Geef een systeemanalyse. Dit kan bv. een verwijzing naar een rapport zijn. Aspecten die daarin aan bod komen zouden moeten zijn: Wat zijn de belangrijkste attributen en processen van het systeem? Welke terugkoppelingen zijn aanwezig? Hoe zijn de grenzen van het systeem bepaald of gedefinieerd? Hoe wordt met relevante attributen en processen buiten deze grenzen omgegaan, bv. randvoorwaarden, constanten, of simpelweg genegeerd? Over welke processen, attributen, terugkoppelingen, etc. bestaat er onzekerheid? Hoe is die onzekerheid bepaald?

Beschouw nu per toepassing

3B. Welke systeem-analytische aspecten (attributen, processen, terugkoppelingen, etc.) zijn (direct) relevant voor de toepassing? En welke in mindere mate of niet? En hoe is die relevantie bepaald?

3C. Zijn alle voor de toepassing belangrijke aspecten als gegeven in B meegenomen in het model/bestand? En welke niet?

3D. Zijn er minder of niet belangrijke aspecten meegenomen in het model/bestand (ja/weinig/geen)?

3E. Beoordeel de mate van 'evenwicht' m.b.t. de systeemanalyse, gebaseerd op de antwoorden op 3C en 3D (goed/matig/slecht).

Kleine adaptatie: Breidt de tekst "beschouw nu per toepassing" iets uit met: "Indien er duidelijk meerdere toepassingen zijn (zie vraag 2C), ...".

De systeemanalyse geeft de benodigde informatie waaruit blijkt welke keuzes zijn gemaakt bij het opstellen van het conceptuele model i.r.t. de beoogde toepassing, en of die keuzes logisch zijn t.a.v. die beoogde toepassing. Verder willen we weten of dat die keuzes dan acceptabel zijn met het oog op die toepassing. Wat zijn de belangrijkste variabelen? Een groot deel van de vragen in vraag 3A zou eigenlijk naar vraag 4 moeten verhuizen. Daarentegen zou er expliciet gevraagd kunnen worden naar de keuzes in de selectie van variabelen, en de motivatie achter die keuzes. De vraag is verder ook 'te lang', en het is beter om in een korte toelichting uit te leggen wat een systeemanalyse inhoudt.

Oplossing/aanbeveling: Pas vraag 3A aan, en voeg een korte introductie toe: "Geef een systeemanalyse. Dit kan bv. een verwijzing naar een rapport zijn. Welke keuzes t.a.v. de selectie zijn gemaakt? Geef motivaties bij die keuzes". Vraag 3B moet dan ook meteen aangepast worden in dit licht.

Een belangrijk punt is ook hoe een systeemanalyse gemaakt moet worden. Het is lang niet voor elke modelleur (of stakeholder) duidelijk wat criteria zijn om de relevantie van een proces te bepalen. Er zijn waarschijnlijk ook geen heldere criteria te geven, omdat het belang om een proces op te nemen in een model of niet sterk afhangt van de eisen die door de toepassing gesteld worden. Feitelijk moet met een 'face validation' (expert-review) beoordeeld worden, of dat een systeemanalyse goed doordacht is of niet.

Oplossing/aanbeveling: Geef bij vraag 3B de richtlijn dat een evaluatie van de systeemanalyse door experts gewenst is.

Er volgen ook nog een aantal punten uit de NP-casus. De vragen 3B, C, D en E zijn onverminderd geldig, maar kunnen zowel vanuit het oogpunt van de gehele modelketen of vanuit de verschillende modules beantwoord worden. Daarbij is het relevant om dit te spiegelen tegen de verschillende toepassingen. Stel het aantal modules op N , en het aantal duidelijk verschillende toepassingen op M . Er moeten dan eigenlijk maximaal $(N + 1) M$ situaties beschouwd worden (+1 vanwege de gehele modelketen). Het is zinnig om ook meteen te verwijzen naar de vergelijkingsmatrix uit vraag 2D/E, aangenomen dat mensen van dit idee gebruik maken. Er moet dus eigenlijk ook naar die vragen verwezen worden voor het geven van een initiële evenwichtsbeoordeling.

Oplossing/aanbeveling: Een oplossing is om een splitsing toe na vraag 3A: "Indien de casus een modelketen is: alvorens vragen 3B t/m 3E te beantwoorden, adviseren wij om eerst een matrix van $(N + 1) \times M$ op te stellen, waarin N het aantal modules en M het aantal toepassingen, of in elk geval steeds te kijken naar de overlap tussen de uitvoer van een model met de invoer van het daaropvolgende model. Beschouw vragen 3B t/m E voor de gehele modelketen en per module.". Een waarschijnlijk betere oplossing is om te herhalen dat per toepassing de gehele evaluatielijst opnieuw doorlopen moet worden.

Bij vraag 3E wordt gevraagd naar een initiële inschatting van het evenwicht. Het is op zich relevant om de vraag al in dit stadium te stellen, maar het probleem is dat het aspect van datasteun nog mist in de analyse, want dat komt pas aan bod in vraag 5 en 6 in de huidige evaluatielijst. Het is echter niet zinvol om de vragen m.b.t. data een eerdere plek op de lijst te geven, aangezien er eerst een beschrijving van het conceptuele model moet zijn (dat het resultaat van de systeemanalyse is), voordat er helderheid is welke eisen er eigenlijk t.a.v. de gegevens zijn.

Oplossing/aanbeveling: Pas de vraag aan, zodat niet expliciet gevraagd wordt naar 'evenwicht', maar naar de keuzes die gemaakt zijn bij de systeemanalyse, en de logica en motivatie daarachter, waaronder in het licht van de toepassing. *Maar zie de opmerkingen bij 5/DATA.*

Adaptatie: Bij vraag 5 behandelen we nog zaken t.a.v. data, die naar 3 verhuisd zouden kunnen worden (zie aldaar).

4. CONCEPTUEEL MODEL of ONTWERPMODEL

4A. Geef het conceptuele model (voor een model) of ontwerpmodel (voor een bestand). Dit kan bv. een verwijzing naar een rapport zijn. Het conceptuele model legt op relatief informele wijze de relaties tussen componenten vast. Aspecten die aan bod kunnen komen zijn: Welke aannamen en vereenvoudigingen zijn er gemaakt? Wat zijn de ruimtelijke dimensies? Welke ruimtelijk-temporeel aggregatieniveaus worden gebruikt? Welke zaken spelen een rol voor het bestand? Wat wordt er voor (des)aggregatie van data gebruikt? Voor modellen: Welk modeltype is gebruikt?

Beschouw nu per toepassing

4B. Op welke (ruimtelijk-temporeel) aggregatieniveaus zijn antwoorden gewenst, gezien de toepassing? Welke ruimtelijke dimensies en welk modeltype/ontwerpmodel zijn gewenst?

4C. Komen de ruimtelijke dimensies, aggregatieniveaus, etc. overeen met die vanuit de toepassing zijn gewenst (ja/matig/niet)?

Kleine adaptatie: Evenals bij vraag 3, breidt de tekst "beschouw nu per toepassing" iets uit met: "Indien er duidelijk meerdere toepassingen zijn (zie vraag 2C), ...".

Veel vragen uit vraag 3A zouden verhuisd moeten worden naar vraag 4A. Bij 4A zou vervolgens eigenlijk veel hiervan in een toelichting gezet kunnen worden. De toelichting zou moeten beschrijven wat de variabelen zijn, dus zijn het

- Constanten (vast);
- Parameters (constant in de tijd maar in principe te variëren, bv. met kalibratie);
- Forcering (externe aandrijving, bv. grondwaterstanden die volgen uit een ander model);
- Cascade-type/response (toestandsvariabelen die alleen reageren op invoer);
- Terugkoppelingstype (toestandsvariabelen die ook feedback geven naar invoervariabelen);

mét een duidelijke nadruk op de schaal. De opmerkingen over 'schaal', die nu bij de data staan, kunnen dus ook naar deze vraag/toelichting verhuizen. In vraag 4B wordt dan nu evenmin gevraagd welke mate van dynamiek, terugkoppelingen etc. vanuit de toepassingen acceptabel zijn.

Oplossing/aanbeveling: Pas vraag 3A aan, en voeg een introductie toe. Pas evenzo vraag 3B binnen die context aan. Verhuis daarnaast de bovenbedoelde onderdelen naar vraag 4A, en voeg daar een korte introductie toe.

Er wordt niet gevraagd of de in het model doorgevoerde vereenvoudigingen aan de voorwaarden kunnen voldoen, die bij 4B worden gevraagd. Dit is wel nodig om antwoord te kunnen geven op vraag 4C.

Oplossing/aanbeveling: Voeg tussen vraag 4B en 4C een extra vraag toe: “*Stellen de in het model doorgevoerde vereenvoudigingen het model in staat om te voldoen aan de gewenste voorwaarden?*”

5. DATA

5A. *Welke gegevens zijn benodigd, gezien de toepassing, en motiveer? Wees specifiek m.b.t. resolutie, nauwkeurigheid, schaal, etc., en vermijd algemene en ongemotiveerde opmerkingen als “file X moet gebruikt worden”.*

5B. *Welke gegevens zijn beschikbaar, en in welke mate?*

5C. *Wat is uiteindelijk de invoer van het model of bestand?*

5D. *Komt de invoer overeen met de door de toepassing gewenste gegevens (ja/deels/nee)? Vergelijk de antwoorden op 5A, 5B en 5C. Let hierbij sterk op eenheden en de drie elementen van schaal (coverage, support, extent; zie Bierkens et al, 2000; Bogaart et al, 2011). Heeft de data de juiste dimensies en eenheden? Zo nee, is er een (des)aggregatiemethode voorhanden? Schaalproblemen doen zich overal voor door de ‘vertaling’ van attributen naar metingen, die gediscrificeerd zijn in tijd en ruimte, en onderhevig zijn aan allerlei bronnen van meetruis.*

Een algemeen probleem met vraag 5 is, dat ie eigenlijk te laat gesteld wordt. In de praktijk zullen al keuzes, die gemaakt worden bij de selectie in de systeemanalyse om tot een conceptueel model te komen, deels beïnvloed worden door de beschikbaarheid van data. Bovendien moet bij bijna elk onderdeel in de modelleercyclus het model afgezet worden tegen de data, net als bij de toepassing. De vraag kan gesplitst worden, maar eigenlijk moet bij veel vragen een additionele vraag worden toegevoegd t.a.v. de data. Bv., bij de systeemanalyse in vraag 3 kan apart worden gevraagd naar de motivatie achter keuzes, en de rol van de data en databeschikbaarheid daarin.

Uit de verschillende casussen kwam verder duidelijk naar voren dat de rol van de verschillende data veel duidelijker uitgelegd zou moeten worden. Immers, data kunnen worden gebruikt voor verschillende variabelen (constant, parameters, forcering, cascade-type, terugkoppelingstype), en voor verschillende zaken (afleiding processen, forcering, kalibratie, validatie). De centrale vraag hierbij is: stellen de data het model in staat om afdoende aan de eisen vanuit de doelstelling te voldoen? Bij forcering- en ruimtelijke data telt bv. de resolutie en precisie. Bij data ten behoeve van kalibratie telt de vraag of die data voldoende informatie in zich meedraagt zodat de *posterior* onzekerheden in modelparameters tot acceptabele proporties kunnen worden teruggedrongen. Er moet dus expliciet naar de rol van de verschillende data gevraagd worden. Dit zou een onderdeel van de systeemanalyse kunnen worden gemaakt, omdat keuzes t.a.v. conceptuele modellen (maar ook voor formele modellen) in de praktijk vaak worden ingegeven door de beschikbare data. De casus ‘MetaSWAP’ laat overigens nog een opvallend punt zien, namelijk dat er eigenlijk alleen maar synthetische data is gebruikt bij het opstellen van het model (het is begonnen als een meta-model van SWAP), terwijl synthetische data in principe onbeperkt beschikbaar is. Wellicht is het zinvol om expliciet een toevoeging te maken m.b.t. dit aspect. Met het oog op de beoordeling van evenwicht en het geven van mogelijke oplossingen, is het daarnaast relevant om te vragen naar alternatieven voor data.

Oplossing/aanbeveling: Verwijder vraag 5. In plaats daarvan wordt bij alle relevante onderdelen vanaf vraag 3 een verwijzing gezet naar de volgende set vragen:

“1/ *Welke gegevens zijn nodig voor dit onderdeel, gezien doel en toepassing? Motiveer. Let op resolutie, schaal, etc., en leg uit wat de rol van de data zou zijn.*”

“2/ *Welke gegevens zijn beschikbaar, en in welke mate?*”

“3/ *Wat is uiteindelijk de data/invoer die gebruikt wordt? Motiveer.*”

Vraag 5D wordt onder vraag 12 geplaatst, omdat dat gaat over de beoordeling van evenwicht, en de mogelijke opties ter verbetering van de casus.

6. FORMEEL MODEL (niet voor databestanden)

Gebruik de antwoorden op vragen 5A tot 5D om vragen 6B tot 6D te antwoorden.

6A. Geef het formeel model, en motiveer de keuze of verwijst naar paginanummers uit een gepubliceerd rapport of artikel. Het formeel model is de omzetting van concept naar bijvoorbeeld wiskundige vergelijkingen of een set rekenregels. De keuze van formeel model hangt af van de beoogde toepassing en de data, en wordt mede bepaald door de gestelde onzekerheidsmarges.

6B. Welke gegevens zijn benodigd voor dit formele model?

6C. Welke gegevens zijn beschikbaar, en in welke mate?

6D. Zijn de gegevens die benodigd én beschikbaar zijn ook gebruikt? Waarom wel/niet?

Er was veel commentaar in de casussen t.a.v. de (vermeende) duplicatie van vragen m.b.t. de data. In dit kader verwijzen we hier naar de opmerkingen hierboven. Vragen 6B t/m D worden dus "herschreven" in die trend.

Oplossing/aanbeveling: Aanpassing vragen 6B t/m D als gesuggereerd bij vraag 5.

Een gewenst aspect is om te bepalen, of het formele model een correcte omzetting van het conceptuele model is. Zijn de vergelijkingen en/of balansen die gebruikt zijn inderdaad de beste keuze, om de essentie van de eerder vastgestelde relevante conceptualisaties te representeren? Wat zijn de uiteindelijke aannames en vereenvoudigingen? Er moet meer aandacht komen voor de keuzes die gemaakt worden bij het formele model, bv. of dat variabelen deterministisch of stochastisch zijn, en van welk type (constant, forcering, etc.).

Oplossing/aanbeveling: Breid de vraag uit van: "motiveer de keuze" naar: "bespreek nadrukkelijk de opties, keuzes, aannames en vereenvoudigingen bij het opstellen van het formele model". Er moet een extra vraag 6E komen: "Beoordeel de mate van 'evenwicht' m.b.t. het formele model. Zijn de gemaakte keuzes, aannames en vereenvoudigingen achter het formele model in overeenstemming met doel, toepassing, en data?"

In het licht van reductie van de vragenlijst is het te overwegen of onderdelen 5 en 6 niet simpelweg samengevoegd moeten worden. Een additionele motivatie volgt uit de al eerder besproken modelleercyclus (Figuur 1.2): veel auteurs en modelleers zien geen fundamenteel verschil tussen het conceptuele en het formele model. We komen hier nog op terug bij de verdere discussie.

7. NUMERIEK REKENMODEL of BESTAND

7A. Geef het numeriek rekenmodel/gegevensbestand (de implementatie in software), en motiveer de keuze of verwijst naar paginanummers uit een gepubliceerd rapport of artikel.

7B. Beschrijf de verificatie of verwijst naar paginanummers uit een gepubliceerd rapport of artikel. Verificatie moet niet verward worden met validatie. De eerste controleert de omzetting van formeel model naar code, de tweede toetst de resultaten van de code aan de werkelijkheid.

7C. Beschrijf en motiveer de keuze voor numerieke rekenmethoden, en geef de gebruikte discretisatie of verwijst naar paginanummers uit een gepubliceerd rapport of artikel. Veelal bestaat er bij numerieke rekenmethoden een afweging tussen rekensnelheid en –nauwkeurigheid.

Wijziging: Vraag 7 en 8 worden omgedraaid, zie opmerkingen bij vraag 8.

Het is relevant hier te vragen naar de gehele numerieke implementatie, en niet alleen het model (rekenhart). Bv. in de casus 'MetaSWAP' bleek dat de pre- en post-processing essentiële onderdelen van het gehele concept zijn. Het is niet heel inzichtelijk om naar de gehele code te vragen. Praktischer is het om naar een flow-diagram of meta-code te vragen.

Oplossing/aanbeveling: Wijzig vraag A tot: "Geef een overzicht met pseudo-code van de numerieke implementatie."

Aansluitend moet ook gevraagd worden naar testen t.a.v. delen van of de gehele code, naast verificatie van het rekenhart.

Oplossing/aanbeveling: Wijzig vraag B tot: *“Geef een overzicht van de verificatie en testen waaruit blijkt dat de numerieke implementatie een goede omzetting van formeel model naar code is. Verwijs evt. naar paginanummers uit een gepubliceerd rapport of artikel”*.

De evaluatie t.a.v. ‘evenwicht’ is juist van toepassing op de numerieke implementatie. Immers, dat is het ‘model’ waarmee in de praktijk (lees: voor toepassingen) altijd wordt gewerkt. In dit kader worden vragen 9C en 9D ook hier naartoe verhuisd. Er kan ook meer informatie worden gevraagd over in hoeverre de numerieke implementatie een bron van fouten en onzekerheden kan zijn. De vraag m.b.t. de afweging tussen rekensnelheid en –nauwkeurigheid is een zeer terechte, maar ook nog erg beperkt. Er kunnen allerlei zaken een rol spelen buiten de numerieke rekenmethodes om, bv. in de vertaalslag van data naar invoer, of van rekenuitkomsten naar uitvoer. De vragen t.a.v. data zijn relevant bij deze vraag in het kader van de invoer, en tevens moet er sterk de nadruk liggen op de toepassing. Bv., een versnelling van de code is altijd zinnig, maar het heeft alleen prioriteit als dit voor de toepassing van belang is.

Oplossing/aanbeveling: Vervang de laatste vraag door een aantal vragen:

1/ Vraag 9C wordt bijgevoegd bij vraag A: *“Geef een overzicht met pseudo-code van de numerieke implementatie. Beschrijf de rol van elke module van het model/bestand, bv. een gewasgroei-deelmodel als onderdeel van een hydrologisch model, of stappen in het stroomschema of de scripts van een bestand. Verwijs eventueel naar een gepubliceerd rapport of artikel met de broncode.”*

2/ Vraag C wordt gewijzigd en uitgebreid: *“Beschrijf en motiveer de keuzes die gemaakt zijn bij de numerieke implementatie (zoals numerieke integratiemethode, opbouw modules, etc.). Bv. bij numerieke rekenmethoden bestaat er vaak een afweging tussen rekensnelheid en –nauwkeurigheid. Verwijs eventueel naar paginanummers uit een gepubliceerd rapport of artikel.”*

3/ Vraag 9D is de basis voor een additionele vraag D: *“Beoordeel de complexiteit van de numerieke implementatie in relatie tot de toepassing. Belangrijke aspecten zijn onder meer: Zijn de goede keuzes gemaakt met het oog op de toepassing? Zijn alle onderdelen nodig? Welke missen/zijn overbodig? Wat is de duur van een run? Hoeveel (tussentijdse) uitvoer is er? Hoe groot is de datastroom? Zou de code sneller/efficiënter kunnen, bv. door selectie, aggregatie, andere methode, andere modulaire opbouw, etc.?”*

8. SCHEMATISERING

De schematisering is bv. het indelen van de bodem in verschillende lagen (‘stratificatie’), of het clusteren van landgebruikstypen. De mate van detail die bij schematisering wordt gevolgd is afhankelijk van de toepassing, maar moet wel voldoende ondersteund worden door de data.

8A. Indien van toepassing, beschrijf de schematisering, of geef relevante en specifieke referenties?

8B. Wordt de schematisering voldoende ondersteund door de data (ja/matig/nee)? Motiveer.

Het lijkt zinniger om deze vraag te stellen, voordat de numerieke implementatie aan de beurt komt. Immers, de schematisering zou toegepast moeten zijn in de implementatie, anders is er geen reden om deze te bespreken. De schematisering kan bovendien als onderdeel van het conceptuele/formele model gezien worden.

Oplossing/aanbeveling: Verwissel vraag 8 en 7 van plaats. Een alternatieve oplossing kan zijn, om deze vraag samen te voegen met vraag 4 (conceptueel model).

Een uitbreiding van deze vraag is ook zinnig. Er wordt niet expliciet gevraagd of de gebruikte schematisering aansluit bij de eisen vanuit de toepassing. Aangezien vaak met beschikbare schematiseringen wordt gewerkt (bv. PAWN) i.p.v. wat gewenst zou zijn, is dit zeker geen triviaal

punt. Het is vooral belangrijk om te weten, wat de onderliggende heterogeniteit is, en hoe deze vereenvoudigd is weergegeven.

Oplossing/aanbeveling: Breid vraag A uit met: *“Let specifiek op de onderliggende heterogeniteit, en hoe deze is vereenvoudigd”*. Voeg een vraag tussen A en B toe: *“Welke schematisering zou gezien de toepassing gewenst zijn?”* Voeg daar meteen een laatste vraag bij: *“Stemt de daadwerkelijke schematisering van het model overeen met wat gewenst is vanuit de toepassing? Waarom wel/niet?”*

9. GEVOELIGHEDEN, ONZEKERHEDEN & OPBOUW CODE

Deze vraag is gericht op het bepalen welke parameters, forcing, invoer, delen van de code, etc. belangrijk zijn en welke overbodig zijn, en welke onderdelen een grote onzekerheid introduceren.

9A. Zijn er gevoeligheids-/onzekerheidsanalyses uitgevoerd? Zo ja, beschrijf deze of verwijz naar paginanummers uit een gepubliceerd rapport of artikel en analyseer de resultaten. Sommige factoren kunnen weinig van belang blijken voor de toepassing, zodat er weinig moeite hoeft te worden gedaan om deze goed te kalibreren, of gebruikt kunnen worden om het model te vereenvoudigen. Andere factoren kunnen juist essentieel blijken.

9B. (niet voor databestanden) Welke numerieke integratiemethode is gebruikt? Bespreek het prestatievermogen van de methode.

9C. Wat is de rol van elke module van het model/bestand, bv. een gewasgroei-deelmodel als onderdeel van een hydrologisch model, of stappen in het stroomschema of de scripts van een bestand? Bespreek dat in relatie met de toepassing. Zijn er onderdelen die overbodig zijn? Missen er nog onderdelen? Motiveer waarom.

9D. Beoordeel de complexiteit in termen van rekentijd en efficiëntie. Wat is de duur van een run? Hoeveel (tussentijdse) uitvoer is er? Hoe groot is de datastroom? Zou de code sneller/efficiënter kunnen, bv. door selectie, aggregatie, andere methode, andere modulaire opbouw, etc.?

Deze vraag moet worden uitgekled en wellicht geheel worden verwijderd. Allereerst, een evaluatie van de code in inhoudelijke zin is iets anders is dan een evaluatie van het evenwicht. Dit houdt wel in, dat de vragen t.a.v. de complexiteit van de numerieke implementatie niet over de volle breedte kunnen worden beschouwd, omdat de informatie t.a.v. de analyse van het model nog mist. T.a.v. de onzekerheden: Het beoordelen van onzekerheden kan beter in het kader van validatie plaatsvinden, en dus in een later stadium dan bij deze vraag. T.a.v. de gevoeligheidsanalyse: Dit is meer een techniek dan een onderdeel van de modelleercyclus. Een gevoeligheidsanalyse kan verschillende rollen hebben, bv. in combinatie met kalibratie en/of het kwantificeren van onzekerheden. Het is daarom wellicht niet zinnig om dit onderdeel zo expliciet te beschouwen t.a.v. ‘evenwicht’.

Oplossing/aanbeveling: Alleen vraag A blijft over. Er wordt een nieuwe vraag B geformuleerd: *“Indien er geen onzekerheidsanalyses zijn gedaan, geef dan een overzicht hier van de mogelijke bronnen van onzekerheden. Schat, als mogelijk, ook in wat hun invloed kan zijn.”* Een alternatief is om de vraag geheel te schrappen, waarbij onderdelen terugkomen in andere vragen.

Vraag A kan helderder geformuleerd worden. Bovendien is het relevant om ook de data en toepassing goed te beschouwen.

Oplossing/aanbeveling: herformuleer de vraag als: *“Geef de referenties naar eventuele gevoeligheids- en/of onzekerheidsanalyses. Wat waren de resultaten van deze studies? Welke parameters, invoer, factoren, etc. zijn gevoelig of onzeker, en welke niet? Welke zouden eventueel weggelaten kunnen worden? Motiveer dit, met name in het licht van de toepassing en de beschikbare gegevens”*.

10. KALIBRATIE (niet voor databestanden)

De kalibratie is het proces waarbij de parameters, begincondities, etc. van het model of bestand van waarden worden voorzien. Kalibratie is vaak gebaseerd op gevoeligheids-analyses (zie vorige vraag).

10A. Indien beschikbaar, beschrijf de uitgevoerde kalibraties en motiveer de keuze, of verwijfs naar paginanummers uit een gepubliceerd rapport of artikel. Geef ook aan welke doelfuncties (lokaal, globaal, deterministisch, stochastisch) zijn gekozen, of dat de betrouwbaarheid van de gekalibreerde parameters wordt gekwantificeerd, hoe, op welke data, of en hoe de betrouwbaarheidsintervallen zijn bepaald, etc. met eventuele verwijzingen naar de literatuur. Let op over-fitting, meervoudige lokale optima, en identificeerbaarheidsproblemen.

10B. Welke eisen stelt de toepassing aan de nauwkeurigheid waarmee parameters, etc. wordt bepaald. Let ook op over-fitting.

10C. Welke mogelijkheden leveren de data om waarden van parameters, etc. te bepalen. Denk hierbij aan resolutie, nauwkeurigheid, aantal gegevens, het aggregatieniveau van de gegevens, en let op over-fitting.

Deze vraag is erg op de data gericht, maar ontbeert de focus op de toepassing, terwijl kalibratie van een model altijd plaatsvindt in het kader van een toepassing. Er wordt nu nog niet gevraagd naar de vrije parameters, en welke parameters eigenlijk de meeste aandacht zouden moeten krijgen met de kalibratie. Er is dan ook niet helder waarom die parameters de aandacht hebben, bv. omdat dat volgt uit de gevoeligheidsanalyse, in verband met de toepassing of databeschikbaarheid, etc. Er wordt wel gevraagd welke doelfunctie(s) is/zijn gebruikt (meestal is er maar één doelfunctie), maar dit wordt niet in de context van de toepassing besproken, wat wel zou moeten. Welke doelfuncties zijn gewenst gezien de specifieke toepassing? Zijn deze ook gebruikt? In het geval van synthetische data (bv. bij meta-modellen) zijn deze vragen beperkt relevant. Echter, het is twijfelachtig of dat nu specifiek voor deze klasse van modellen de vragen erg aangepast moeten worden. Dit zou al ondervangen kunnen worden met de zinsnede: “indien van toepassing”.

Oplossing/aanbeveling: Breid vraag 10B uit met: “Welke eisen stelt de toepassing aan de kalibratie, bv. welke doelfuncties zijn gewenst gezien de specifieke toepassing?” Voeg een vraag 10D toe: “Stemt de uitgevoerde kalibratie overeen met wat gewenst is vanuit de toepassing, en ondersteund wordt met data?”

11. VALIDATIE

Wij definiëren hier een validatie als een toets of het model of bestand een redelijke representatie is van het werkelijke systeem dat wordt gemodelleerd. Dit geschiedt op basis van een set onafhankelijke gegevens (gegevens die nog niet in een eerder stadium bij het modelleren gebruikt zijn).

11A. Indien beschikbaar, beschrijf de gedane validatiestudies of verwijfs naar paginanummers uit een gepubliceerd rapport of artikel.

11B. Onderbouw/beoordeel de waarde van de validatie(s). Hoe relevant zijn de verschillende validatiestudies voor de verschillende toepassing(en) van het model of bestand? Zijn alle toepassingen voldoende afgedekt door de studies? Is ook de betrouwbaarheid van de validatie gekwantificeerd? Welke toepassingen moeten opnieuw worden gevalideerd?

Het is niet helder hoe de betrouwbaarheid van een validatie gekwantificeerd zou kunnen worden in de meeste gevallen. Verder is het mogelijk dat bij validatiestudies een andere doelfunctie wordt gebruikt dan bij de kalibratie. Hier zou expliciet naar gevraagd moeten worden, ook omdat het gebruik van een andere doelfunctie de essentie van de validatie zou kunnen zijn. De doelfunctie is echter maar één aspect van de validatie. In de praktijk zijn er diverse modellen (bv. graadmeters, (delen van) natuur- en milieuprocesmodellen, en Complex Adaptieve Systemen) waar validatie zeer lastig is, principieel dan wel in de praktijk.

Oplossing/aanbeveling: Vraag 11B zou specifiek uitgebreid kunnen worden met de opmerkingen t.a.v. doelfuncties: *“Welke doelfuncties zijn gebruikt in de validatie? Motiveer”*.

Een specifiek punt t.a.v. modelketens is dat het relevant is om zowel naar de validatiestatus van de aparte modules als van de modelketen als geheel te kijken. Een aantal additionele aspecten die hier aan bod zouden moeten komen zijn genoemd in de NP-casus.

Oplossing/aanbeveling: Geef een extra vraag 11C: *“Indien het een modelketen betreft: Zijn alle modules gevalideerd? Welke wel/niet? Zijn de modules gevalideerd in de context van de toepassing van de modelketen?”*

12. SAMENVATTING en ALGEMEEN OORDEEL

Dit laatste onderdeel is gericht op het geaggregeerde resultaat van de antwoorden op de verschillende vragen en de meningen van de verschillende betrokkenen, en kan indirect ook tot adviezen leiden ter verbetering van het model of bestand.

12A. Beoordeel de mate van vertrouwen in het model/bestand in het algemeen, mede gebaseerd op wat er is gedaan aan testen, verificatie, kalibratie, validatie, gevoeligheids- en onzekerheidsanalyses, en motiveer deze.

12B. Zijn er veel missende zaken? Zijn er veel vragen waar de zaken als ‘te eenvoudig’ zijn beoordeeld? Motiveer.

12C. Zijn er veel overbodige componenten? Zijn er veel onderdelen waar het eenvoudiger zou kunnen? Motiveer.

12D. Is er voldoende datasteun, en waaruit blijkt dat? Motiveer

12E. Welke specifieke suggesties voor verbetering en data-aanlevering komen naar voren uit deze analyse?

Een aantal van deze vragen, en met name 12B en C, is wellicht beter te beantwoorden wanneer de historische context van het model in de evenwichtsanalyse wordt beschouwd. Een andere optie is door het model te vergelijken met andere modellen die betrekking hebben op dezelfde toepassing. Het hangt verder nogal af van de toepassing of het model evenwichtig gevonden wordt of niet. In vraag 12 wordt nu geen specifiek onderscheid gemaakt t.a.v. de toepassing.

Oplossing/aanbeveling: Breid de introductie van vraag 12 uit: *“Beschouw bij elk van de vragen in elk geval doel, (de verschillende) toepassingen, databeschikbaarheid, en de geschiedenis van de casus.”*

Een expliciete vraag die mist is simpelweg wat nu het oordeel m.b.t. ‘evenwicht’ is. Die vraag moet ook gesteld worden, voordat vraag 12E gesteld wordt.

Oplossing/aanbeveling: Verschuif vraag 12E naar F. Vraag E wordt: *“Beoordeel nu (gebruikmakend van alle voorgaande antwoorden) of het model evenwichtig is of niet (complexiteit i.r.t. data en toepassing). Motiveer”*.

Vraag 12E kan goed uitgebreid worden. Men zou willen weten wat voor andere opties er overwogen zijn, en waarom er dan niet voor die opties gekozen is. Dat kan ook mede afhangen van de geschiedenis van het model, of die van de toepassing (voor doorlopende toepassingen worden meer dan eens ook steeds dezelfde – inmiddels verouderde – modellen gebruikt). Het moet niet alleen in het kader van de analyse bekeken worden, maar ook suggesties van de ondervraagden en auditeurs kunnen meegenomen worden. Verder is het van belang om ook kritisch te zijn bij het geven van de suggesties, d.w.z. de suggesties moeten zelf ook onderbouwd worden.

Oplossing/aanbeveling: Breid de vraag uit: “*Motiveer de suggesties. Welke opties zijn overwogen? Welke zijn redelijk, mogelijk, en zinnig? Beschouw de keuzes en aannames die gemaakt zijn bij het model, en die gemaakt worden bij de suggesties*”.

6.3 Overige bevindingen

De aanpak in dit WOt-document was een dubbelzijdige toets, d.w.z., zowel de casus als de evaluatielijst werden getoetst. Het nadeel daarvan is, dat niet altijd alle gewenste informatie boven water kwam in een casus. Dat werd deels gecompenseerd doordat het invullen van de lijst zelf vaak in meerdere stappen gebeurd is, die ook gescheiden zijn geweest in de tijd. Het maken van notities voor de verschillende vragen was soms toch nog erg tijdrovend, en in een aantal gevallen is het geven van een antwoord zeker niet triviaal gebleken, en was verdere uitleg m.b.t. de bedoeling van de vraag nodig. Soms is op suggestie van de interviewer dan wel de geïnterviewde afgeweken van de vragenlijst. De vragen op de lijst zijn uiteindelijk nog apart volledig ingevuld door gebruik te maken van de notities in combinatie met verdere referenties. Verder leert de ervaring dat het zinvol is om in elk geval één iteratieve slag te maken bij het invullen van de lijst – iets dat ook telkens blijkt uit het invullen van de Status A-checklist in de kwaliteitsborging bij de WOT, zeker daar waar het mensen betreft die geen eerdere ervaring met een dergelijke lijst hebben.

Voor het invullen van de lijst en de bijbehorende evaluatie van een modelketen lijkt het erop, dat het zinniger is om eerst te bepalen uit welke modules de modelketen bestaat, en dan voor elk van deze modules apart een overzicht in te vullen met zaken als doel, toepassing, modelbeschrijving, validatie, etc., alvorens tot de analyse van de modelketen in zijn geheel over te gaan. Dit is van belang, omdat zelfs als er korte beschrijvingen van de modules in de beschrijving van de modelketen zelf zijn, deze beschrijvingen altijd in de context van die modelketen zullen zijn. Echter, we zoeken naar zaken waardoor het evenwicht van het model verschuift, en zaken als het doel en de toepassing van de modelketen kunnen anders zijn dan dat ze van origine waren voor de aparte modules. Dit soort zaken moet dus gecontroleerd worden, bv. is het doel van module X in overeenstemming met dat van de modelketen, of wellicht is module Y wel gevalideerd, maar niet voor de toepassing van de modelketen. In de casus ‘Natuurplanner’ was het zeer behulpzaam dat er van bijna alle modules al gegevens beschikbaar waren via ingevulde formulieren voor Status A en Normenkader (de lijsten die gebruikt worden bij de bestaande kwaliteitsborging bij resp. de WOT en het PBL); gegevens konden in veel gevallen redelijk eenvoudig overgezet worden bij het beantwoorden van de vragen van de EMC-lijst. Het is twijfelachtig of de casus doenbaar was geweest bij het ontbreken van die vooraf ingevulde lijsten. Zoals al gezegd, is het voor de analyse van de complexiteit van de modelketen zeer relevant om ook de status van de modules onderling goed te kunnen vergelijken (bv. of alle modules geverifieerd en gevalideerd zijn, en of er gevoeligheids- en onzekerheidsanalyses zijn gedaan, etc.), en het is onzeker of dat dit gelukt was zonder de bestaande overzichten in de vorm van ingevulde Status A- en Normenkader-lijsten.

6.4 Aanzet tot aandachtspunten “evenwichtig modelleren”

Nu de EMC v2.0 ontwikkeld is, is het de vraag wat er met de resultaten gedaan zou kunnen en moeten worden. Op basis van de in dit project gebruikte literatuur en de bevindingen binnen de casussen, is er een eerste versie van richtlijnen opgesteld met als beoogd doel het stimuleren van de vergroting van modelevenwicht.

Algemene aandachtspunten

- De toepassing is uiteindelijk allesbepalend voor de keuzes die gemaakt worden in het modelleerproces. Maak deze expliciet. Waarom krijgen bepaalde delen van het gemodelleerde systeem meer aandacht dan andere? Waarom wordt er gekozen voor een bepaalde schaal?

Waarom wordt er een bepaalde doelfunctie gebruikt tijdens de modelkalibratie? Waarom wordt er voor een bepaalde validatiestrategie gekozen?

- Documenteer zoveel mogelijk de gemaakte keuzes bij alle modelleerstappen, waaronder alle aannames en vereenvoudigingen, en de motivaties achter de gemaakte keuzes, aannames, en vereenvoudigingen, liefst ook aangevuld een overzicht van belangrijke overwogen alternatieven.

Aandachtspunten bij toepassingen

- Maak een goede systeemanalyse. Beredeneer vanuit de toepassing welke systeemcomponenten (toestandsvariabelen en relaties) onmisbaar, relevant of irrelevant zijn. Controleer of verschillende stakeholders hier hetzelfde over denken. Bepaal welke schaal (zowel in ruimte als in tijd het meest relevant is). Bepaal welke precisie en accuraatheid gewenst of vereist is, en dus welke onzekerheid acceptabel is.
- Probeer op basis van deze redenering, en data-aspecten (zie onder) een adequate modeluitvoer, doelfuncties voor validatie, en kalibratieprocedure op te stellen. Maak indien mogelijk gebruik van multi-objective doelfuncties en gebruik kalibratietechnieken die de maximale hoeveelheid informatie uit de beschikbare data kunnen halen (Vrugt *et al*, 2002; Wagener *et al*, 2003; 2009a).
- Ook voor “academische” toepassingen geldt dat er een helder modelleerdoel gedefinieerd kan worden, bv. de identificatie van de eerste orde-invoer, de stuurbaarheid van het systeem, robuustheid tegen verstoringen, attributie (d.w.z. de identificatie van causale factoren), enz. Er is altijd ook een doel achter het doel, bv. er is identificatie van de eerste orde-invoer gewenst, omdat men daar op wil gaan sturen.

Aandachtspunten bij data

- Inventariseer welke data beschikbaar zijn, inclusief “soft data” zoals kwalitatieve of semi-kwantitatieve observaties (Seibert *et al*, 2002).
- Analyseer op welke wijze de beschikbare data gerelateerd is aan systeem-variabelen en bijbehorende toepassingsafhankelijke relevantie. Zijn data primair gekoppeld aan niet-relevante systeemeigenschappen? Overweeg dan om deze niet te gebruiken.
- Bedenk of de huidige modelleeropzet maximaal gebruik maakt van de beschikbare data. Denk hierbij aan het verbeteren van de modeldiscretisatie of het gebruik binnen een kalibratieprocedure: aanvullende data kunnen worden gebruikt om multi-objective kalibraties uit te voeren. Kwalitatieve data kunnen ook in een kalibratie procedure worden gebruikt, bv. om “behavioural” parametersets te scheiden van “non-behavioural”, om zodanig onzekerheden te kunnen verkleinen. Door op een slimme wijze data en modelvariabelen te transformeren kan onderscheid worden gemaakt tussen alternatieve mechanismen en modellen (Kirchner *et al*, 1996).
- Indien blijkt dat de beschikbare data niet afdoende is om modelparameters adequaat te identificeren, of om tussen concurrerende modelstructuren onderscheid te kunnen maken, zodat posterior onzekerheden niet tot door de toepassing gesuggereerde acceptabele betrouwbaarheidsintervallen kunnen worden teruggebracht, overweeg dan additionele data te vergaren, eventueel in het kader van lange-termijn planning en follow-up studies.
- Zorg voor afdoende balans tussen de verschillende rollen van data: data die sturend zijn voor de ruimtelijke discretisatie van modellen; data die gebruikt worden ten behoeve van kalibratie, en “onafhankelijke” data die gebruikt kunnen worden voor de validatie van het model.

Aandachtspunten bij modelstructuur

- Probeer binnen de modelstudie de beschikbare ruimte te gebruiken om enerzijds zo goed mogelijk aan te sluiten bij de wensen vanuit de toepassing (schaal, uitvoer, doelfuncties, relevante variabelen en processen, etc.) en anderzijds bij de mogelijkheden van de data (maximalisering van de informatie-extractie). Gebruik hiervoor bovengenoemde technieken (Vrugt *et al*, 2002; Wagener *et al*, 2003; 2009a).

- Voer een iteratieve reductie van het model uit. Probeer onderdelen uit het model te halen die geen bijdrage leveren aan de toepassing, maar wel een bron van onzekerheid vormen. Elke aanpassing kan worden geëvalueerd door middel van een validatie.
- Analyseer de koppeling tussen de modelcomponenten (modules, sub-modellen) met een matrix. Zijn er verschillende interpretaties van gelijke namen van variabelen, schaalverschillen, etc. die verbeterd kunnen worden? Probeer dat dan. Zijn de koppelingen tussen modules één- of multi-directioneel? Is er, gezien systeemanalyse een grond, en gezien de toepassing een wens voor terugkoppeling tussen de modules? Zo ja, probeer deze te implementeren.
- Geef adequate aandacht aan de ruimtelijke discretisatie van het model en de bijbehorende bepaling van parameterwaarden. Probeer een balans te vinden tussen wat gewenst is vanuit de toepassing en wat mogelijk is vanuit de data. Indien er wordt aangenomen dat ruimtelijke eenheden intern homogeen zijn, controleer dan of dat zo is. Normaliter is er heterogeniteit op alle schalen. Controleer of de feitelijke heterogeniteit binnen model-eenheden geen te grote bron van onzekerheden en onnauwkeurigheden vormt. Dit kan worden gedaan door de voorgestelde discretisatie te vergelijken met een fijnere discretisatie gevolgd door aggregatie tot de oorspronkelijke, voorgestelde discretisatie.
- Harmoniseer numerieke paradigma's, m.a.w., combineer bij voorkeur bv. geen evenwichtoplossingen met dynamische benaderingen, of partiële differentiaal-vergelijkingen met cellulaire automaten. Ditzelfde geldt idealiter t.a.v. het gebruik van verschillende computertalen: hoewel er diverse redenen te verzinnen zijn om gebruik te maken van verschillende 'talen' voor hetzelfde model (bv. gebruik van verschillende functionaliteiten), is het te prefereren dat alles in dezelfde taal wordt geschreven, voor het voorkomen van "vertaalbugs" door incompatibiliteiten.

Aandachtspunten ter vergroting kwaliteit

- Voer een validatiestudie, onzekerheidsanalyses en verdere onderdelen van de evenwichtsanalyse uit. Identificeer op deze wijze de zwakke plekken in het model, vanuit de context van de beoogde toepassing. Gebruik hiervoor moderne inzichten in modelvalidatie (Rykiel, 1996; Refsgaard en Henriksen, 2004; Wagener *et al*, 2009b).
- Gebruik deze analyse om te inventariseren welke verbeteringen (bv. in modelstructuur, data-acquisitie, validatie-inspanningen) het meest effectief zijn. In de praktijk is het gegeven beschikbaar budget hierin een dominante randvoorwaarde.

6.5 Reflectie en discussiepunten

In hoeverre zijn de opgeworpen onderzoeksdoelen gehaald in deze studie? Welke vragen kunnen we nu beantwoorden? En wat moet er nog gebeuren?

De ontwikkelde evaluatielijst EMCv2.0 is een tool, die bedoeld is om te evalueren of de modellen en bestanden van het PBL geschikt zijn met het oog op de toepassing. De lijst kan in elk geval de discussie stroomlijnen om de antwoorden op een aantal onderzoeksvragen vast te stellen, waaronder:

- Welke eisen stelt de toepassing aan modellen en bestanden?
- Overlappen deze eisen met eisen aan modellen en bestanden gericht op begrip van het systeem, en zo nee, op welke additionele punten moet de aandacht gericht worden om een model of bestand geschikt te maken voor de toepassing?

De in Van Voorn & Walvoort (2011) gestarte casussen en de analyses daarvan zijn in dit werkdocument opgenomen. Er zijn veel nuttige punten naar boven gekomen voor verbetering van de lijst; tevens zijn er bevindingen die inzicht geven in het 'evenwicht' van de casus ("Zijn de bestaande modellen en bestanden te eenvoudig of te complex voor gebruik binnen de toepassing?"). Aandachtspunten t.a.v. verbetering van het PBL-instrumentarium staan vermeld in de voorgaande sectie (6.4).

De centrale vraag binnen dit project (“Zijn de modellen en bestanden die binnen het PBL-instrumenten vallen en gebruikt worden voor beleidsevaluaties en –verkenningen voldoende geschikt voor deze toepassingen?”) is niet integraal beantwoord. Bij een aantal casussen (o.a. Natuurplanner en MetaNatuurplanner) is de complexiteit van de casus aanzienlijk, is er een aanzienlijk aantal belanghebbenden, en lijken er meerdere toepassingen te zijn, die niet altijd even helder aangegeven kunnen worden. Deze factoren maken dat het beoordelen van het ‘evenwicht’ van deze casussen niet triviaal is. Er zijn aandachtspunten voor verbetering, maar deze moeten nog verder uitgewerkt worden om daadwerkelijk te kunnen spreken over algemene richtlijnen ter verbetering van het PBL-instrumentarium. Verder is er nog niet gewerkt aan een klein en snel instrument voor een snelle evaluatie van het gehele PBL-instrumentarium, met de bedoeling om de belangrijkste knelpunten t.a.v. modelcomplexiteit naar boven te halen.

Er zijn nog een aantal punten uit de expert-review uit Van Voorn & Walvoort (2011) waar verder nog niet veel mee gedaan is:

- Vragen moeten niet alleen relevant zijn, maar ook goed gesteld worden. Bij deze dubbelzijdige toets is er al wel op gelet, maar in sommige gevallen zouden vragen toch nog ambigu kunnen blijken;
- De suggestie om met spelontwikkelaars te spreken. Hier is verder niets mee gedaan;
- Het gebruik van een wiki-systeem om modellen te verbeteren. Dit is een interessante suggestie, en bv. Deltares maakt ook gebruik van een wiki-systeem. Dit punt zou bij de toekomstige ‘richtlijnen en suggesties’ opgenomen kunnen worden;
- Modellen kunnen ook gebruikt worden om aanwijzingen te geven hoe zo efficiënt mogelijk data verzameld kan worden. Dit is een ‘systems & control’-oogpunt, dat echter pas relevant wordt als er voldoende vertrouwen in een model is.

De ervaring leert dat lijsten nooit in staat zijn om alle gewenste informatie boven water te krijgen. Men zou zich zelfs kunnen afvragen of het niet zinniger is om de oorzaken achter de constante toename in modelcomplexiteit (zoals opgesomd in Van Voorn *et al*, 2011) aan te pakken. De evaluatielijst zou daarvoor ook bruikbaar kunnen zijn. Sowieso is enig pragmatisme vereist bij het gebruik van de lijst, en men kan de lijst dan ook misschien beter opvatten als een hulpmiddel bij het bewustwordingsproces t.a.v. modelcomplexiteit, en niet als een doel op zich. Het belangrijkste aspect van modelcomplexiteit is, om in te schatten of de best mogelijke keuzes zijn gemaakt in het licht van de toepassing en databeschikbaarheid. Subjectiviteit speelt altijd een rol, ook in het formuleren en gebruiken van modellen. Zolang echter de inschatting met de evaluatielijst redelijk lukt, dan lijkt de opzet achter de lijst geslaagd.

Het concept ‘evenwicht’ zelf moet ook kritisch beschouwd worden. Er is bv. nogal een verschil tussen een conceptueel of formeel model en de numerieke implementatie. Een voorbeeld is de Richards’ vergelijking uit de casus ‘MetaSWAP’: het formele model bestaat uit een enkele partiële differentiaalvergelijking en twee grenscondities, maar aangezien er geen algebraïsche oplossing voor lijkt te zijn, bestaat er een ongelooflijke variatie aan (behoorlijk complexe) numerieke implementaties om dit formele model te benaderen. Dit heeft consequenties voor het gebruik van het concept ‘evenwicht’. Sterker nog, de motivatie om de checklist op te stellen a.d.h.v. de modelleercyclus is ingegeven doordat er bij verschillende stappen in de modelleercyclus verschillende aspecten een dominante rol spelen t.a.v. modelcomplexiteit en de balans daarin. Wat nog wel aandacht behoeft in toekomstig onderzoek is een goede inventaris van de belangrijke aspecten in elke fase. Echter, de peer-reviewed literatuur lijkt voornamelijk gericht te zijn op “enkelvoudige” trade-offs, waarvan de AIC (Akaike, 1974) – de trade-off tussen het aantal parameters als maat voor modelcomplexiteit en de relatieve winst in fit op de data bij toenemende complexiteit – een zeer prominente is.

Daarnaast is het relevant om in het oog te houden, dat het goed mogelijk is dat hetzelfde model (in de praktijk vaak dezelfde code) voor meerdere toepassingen wordt gebruikt. Het is niet moeilijk een situatie te verzinnen waarin het model een evenwicht heeft in het licht van de ene toepassing, terwijl

het gezien een andere toepassing helemaal geen goed evenwicht heeft. Dit geeft aan dat er bij een model dat wel of niet in 'evenwicht' is, niet noodzakelijkerwijs ook gesproken kan worden over resp. een 'goed' of 'slecht' model. Een consequentie hiervan is, dat men modellen als dusdanig niet kan kwalificeren als 'evenwichtig' (of niet), maar enkel modellen i.r.t. hun toepassingen. Er moet daarom eigenlijk per toepassing een nieuwe evenwichtsevaluatie gemaakt worden.

6.6 Vooruitzichten

Voor het verdere verloop van het project zijn in elk geval de volgende zaken voorzien:

1. Publicatie van het onderzoek in de internationale wetenschappelijke peer-reviewed literatuur;
2. Het opstellen van een minimumchecklist;
3. Een mogelijke opname van de EMC-lijst in de WOT/WUR-kwaliteitsborging;
4. Het inpassen van het onderzoek binnen de nieuwe onderzoeksthema's binnen WOT/PBL.

1. *Publicatie wetenschappelijke literatuur*

Het is beoogd om in 2013 over het onderzoek te publiceren in de internationale wetenschappelijke peer-reviewed literatuur. Dit is in de eerste plaats bedoeld om het onderzoek te toetsen tegen wat er bekend is over modelcomplexiteit en aspecten daarvan, om er zeker van te zijn dat het onderzoek aansluit bij de meest recente kennis, maar ook om het onderwerp te laten landen bij een breder publiek. Daarvoor is nog een bredere inventaris nodig van de verschillende aspecten die spelen t.a.v. modelcomplexiteit bij de verschillende stappen in de modelleercyclus.

2. *Opstellen minimumchecklist*

Na twee rondes van toetsing – de eerste ronde is beschreven in Van Voorn & Walvoort (2011), de tweede in deze rapportage – is de verwachting, dat de functionaliteit van de EMC v2.0 wel redelijk getest is. Echter, de 'complexiteit' van de lijst zelf is inmiddels ook aanzienlijk: er zijn redelijk wat onderdelen, deelvragen, en detailvragen. Dit verhoogt de drempel aanzienlijk om de lijst in te vullen, en verhoogt het risico dat mensen op een punt "struikelen", terwijl dat misschien niet de kern van de zaak is. Met in het achterhoofd het principe, dat men grofweg de eerste 80% van de informatie verkrijgt met de eerste 20% van de moeite, is het beoogd om een minimumchecklist op te stellen. Deze 'mini-EMC' moet de kern van de zaak bevatten in maximaal 5 à 10 vragen, waarbij veel van de onderliggende vragen en begeleidende informatie naar een bijbehorende appendix wordt verschoven. De lijst zal dan niet in staat zijn om direct alle pijnpunten boven te halen, maar moet wel bruikbaar zijn om een snel onderscheid te maken tussen casussen die 'in orde', 'redelijk tot matig' of 'onder de maat' zijn (met daarbij acht te slaan op het punt, dat dit van de toepassing afhangt). Na een dergelijke eerste schifting kan dan altijd nog de 'grote broer' EMC v2.0 gebruikt worden voor een diepgaandere analyse van de casus.

Tabel 6.1. Mogelijke opzet voor een minimumchecklist voor de beoordeling van 'evenwicht' van een model, mede gebaseerd op figuren 1.1 en 1.2. Per fase worden in kolom 2 de eisen vanuit de toepassing vertaald naar de gewenste modelcomplexiteit. In kolom 3 wordt de aanwezige steun voor deze modelcomplexiteit beschreven. In kolom 4 wordt de beoordeling in termen van 'evenwicht' gegeven. In kolom 5 wordt deze beoordeling onderbouwd.

Toepassing?:				
Fase	MC gewenst	MC gesteund	Saldo	Motivatie
Systeemanalyse/ Conceptueel model				
Numeriek model				
Kalibratie/ Toepassingsmodel				
Validatie				
Eindoordeel 'evenwicht':				

Het opstellen van deze minimumchecklist gaat hand in hand met het verder inbedden van dit onderwerp in de peer-reviewed literatuur. Het lijkt namelijk een voor de hand liggende keuze om in de minimumchecklist de belangrijkste aspecten op te nemen t.a.v. modelcomplexiteit zoals die in de peer-reviewed literatuur worden genoemd, eventueel aangevuld met enkele punten die als grote leermeester worden ervaren. Een suggestie is om deze minimumchecklist als een tabel vorm te geven, bijvoorbeeld in bovenstaande Tabel 6.1.

3. Opname in kwaliteitsborging

Het huidige hulpmiddel in de kwaliteitsborging van modellen en bestanden is het al eerder genoemde 'Status A'. Deze evaluatielijst is vooral gericht op documentatie en beheer van de casus, terwijl inhoudelijke controle vooral 'onofficieel' gebeurt. De EMC v2.0 zou een bruikbaar hulpmiddel kunnen zijn voor een meer inhoudelijk gerichte borging van modellen en bestanden. Oppervlakkig delen de Status A-lijst en EMC v2.0 veel vragen met elkaar, maar bij EMC is het de bedoeling dat de vragen veel dieper op de inhoud ingaan. Het is nog wel zinnig om de twee lijsten verder inhoudelijk te vergelijken, om te kijken welke vragen eventueel zouden kunnen worden overgeslagen, als voor een casus eerst de Status A-lijst is ingevuld, alvorens de EMC-lijst wordt toegepast. In de casus NP is overigens veelvuldig gebruik gemaakt van de ingevulde Status A-lijsten voor verschillende modules, maar in die gevallen bleek de informatie zelden inhoudelijk erg bruikbaar. Andersom zal een casus, die is onderworpen aan een evaluatie met EMC, vermoedelijk al op veel punten (ruim) voldoen aan de eisen voor Status A.

4. Onderzoek binnen nieuwe thema's inpassen

Er is bij PBL en het ministerie meer aandacht gekomen voor de bredere inkadering van modellen en bestanden in onderzoek. Naast de wetenschappelijke borging van modellen en bestanden in meer technische zin, is het ook relevant dat modellen en bestanden aan andere eisen voldoen. Drie aspecten die genoemd worden in de peer-reviewed literatuur zijn '*legitimacy*', '*credibility*' (geloofwaardigheid), en '*salience*' (mate van belangrijkheid voor de beschouwer). *Credibility* is een aspect dat al voor een goed deel afgedekt wordt in de huidige kwaliteitsborging (Status A) en in EMC v2.0. In principe is een model geloofwaardiger naarmate de onderbouwing beter is. Dat wil nog niet zeggen dat het model of zijn gebruik geldig is – dat hangt immers van de toepassing af! – en of dat verschillende stakeholders rondom de toepassing van het model zich herkennen in het model en modelgebruik.

Het aspect *legitimacy* komt nu al voor een deel terug in de expliciete opname van doel en toepassing in de evenwichtsanalyse, terwijl *salience* slechts impliciet enigszins aan bod komt in EMC v2.0. Beide aspecten zouden veel explicieter kunnen worden opgenomen in een EMC-variant, bv. door het opnemen van vragen voor het in kaart brengen van de verschillende stakeholders en hun visies. Die punten zouden dan verwerkt kunnen worden rondom 'doel', 'toepassing', en 'systeemanalyse'. Daarbij is dan wel een nieuwe definitie nodig van 'evenwicht', waarin naast de *credibility* ook de andere twee aspecten afdoende aan bod komen. Een 'evenwichtig' model is in dit kader wellicht een model, dat voldoende complex is met het oog op het geven van antwoorden voor een toepassing, voldoende gesteund door gegevens, niet over-complex, maar bovendien correct gebruikt, en in staat om de verschillende visies en belangen van de belanghouders voldoende te beschouwen.

Literatuur

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Contr.* 19, 716-723.
- Bakkenes, M., D. C. J. van der Hoek, J. R. M. Alkemade (2003). Documentatie testrapport modelketen NATUURPLANNER. RIVM-rapport 500002001/2003.
- Bierkens, M. F .P., P.A. Finke & P. de Willigen (2000). Upscaling and Downscaling Methods for Environmental Research. Kluwer Academic Publishers, Dordrecht.
- Bierkens, M. F .P., P.A. Finke & P. de Willigen (2000). Upscaling and Downscaling Methods for Environmental Research. Kluwer Academic Publishers, Dordrecht.
- Bogaart, P. W., G. A. K. van Voorn, L. M. W. Akkermans (2011). *Evenwichtsanalyse modelcomplexiteit – een verkennende studie*. WOt-werkdocument 226. WOT Natuur & Milieu, Wageningen UR, Wageningen.
- Bouwmans, J.M.M. (1990). Achtergrond en toepassing van de TCGB-tabel: een methode voor het bepalen van de opbrengstdepressie van grasland op zandgrond als gevolg van een grondwaterstandsverlaging. Utrecht, Technische Commissie Grondwater Beheer.
- Brouwer, F., J.T.M. Huinink (2002). Opbrengstdervingpercentages voor combinaties van bodemtypen en grondwatertrappen. Wageningen, Alterra-rapport 429.
- Brus, D.J., G.B.M. Heuvelink (2007). Towards a Soil Information System with quantified accuracy. Three approaches for stochastic simulation of soil maps. Wageningen, Wettelijke onderzoekstaken Natuur & Milieu, Rapport 58.
- Brus, D.J., R. Vašát, G.B.M. Heuvelink, M. Kotters, F. de Vries, D.J.J. Walvoort (2010). Towards a Soil Information System with quantified accuracy. Wageningen, Wettelijke Onderzoekstaken, WOt-werkdocument 197.
- De Vos, J.A., P.J.T. van Bakel, I.E. Hoving (2008). Waterpas nat- en droogteschade-berekeningen ten behoeve van landbouwkundige doelrealisatie; plan van aanpak. Wageningen, Alterra-rapport 1653.
- De Vries, W., M. Posch, J. Kämäri (1989). Simulation of the long-term soil response to acid deposition in various buffer ranges. *Water, Air, and Soil Pollution* 48, 349-390.
- De Vries, F., W.J.M. de Groot, T. Hoogland, J. Denneboom (2003). De Bodemkaart van Nederland digitaal. Toelichting bij inhoud, actualiteit en methodiek en korte beschrijving van additionele informatie. Wageningen, Alterra-rapport 811.
- De Vries, W., G. W. W. Wamelink, G. J. Reinds, H. J. J. Wieggers, J. P. Mol-Dijkstra, J. Kros, G. J. Nabuurs, A. Pussinen, S. Solberg, M. Dobbertin, D. Laubhann, H. Sterba, M. van Oijen (2007). Assessment of the relative importance of nitrogen deposition, climate change and forest management on the sequestration of carbon by forests in Europe. Alterra-rapport 1538, Wageningen.
- De Vries, F., G. Mol, M.J.D. Hack-ten Broeke, G.B.M. Heuvelink and F. Brouwer (2008). Het Bodemkundig Informatie Systeem van Alterra. Overzicht van het gebruik en wensen voor verbetering van de informatie. Alterra-rapport 1709, Wageningen.
- Doherty, J. (2005). PEST – Model-independent parameter estimation. User manual: 5th Edition. Watermark Numerical Computing. <http://www.pesthomepage.org/Downloads.php>

- Feddes, R. A. (1987). *Crop factors in relation to Makkink reference-crop evapotranspiration*. In: Hooghart, J. C. (ed.), *Proceedings and information / Netherlands Organization for Applied Scientific Research, Committee on Hydrological Research* 39, 33-45.
- Finke, P.A., J.J. de Gruijter, R. Visschers (2001). Status 2001 Landelijke Steekproef Kaarteenheden en toepassingen. Wageningen, Alterra-rapport 389.
- Goovaerts, P., 1997. *Geostatistics for natural resources evaluation*. Oxford University Press.
- Gebrengus, T., T. A. Ghezzehei (2011). An index for degree of hysteresis in water retention. *Soil Sci. Soc. Am. J.* 75, 2122-2127.
- Jansen, J. M. L., J. M. Halbertsma, J. A. P. Heesterbeek, H. Houweling, M. J. W. Jansen (2004). *Kwaliteitsborging databestanden en modellen – Balanceren tussen chaotische dynamiek en geordende stilstand*. Alterra-rapport 956, Wageningen UR, Wageningen.
- Kirchner, J. W., Hooper, R. P., Kendall, C., Neale, C., Leavesley, G., Hooper, R. P., & Neal, C. (1996). Testing and validating environmental models. *Science of The Total Environment*, 183(1-2), 33-47. doi:10.1016/0048-9697(95)04971-1
- Knotters, M., H. Vroon, A. van Kekem, T. Hoogland (2009). Deciding on the detail of soil survey in estimating crop yield reduction. In: Devillers, R. and H. Goodchild (Eds.). *Spatial data quality. From process to decisions*. CRC Press, Boca Raton, p. 117-125.
- Knotters, M., D. Brus, G. Heuvelink, B. Kempen, F. de Vries, D. Walvoort (2010). Vaste grond onder de voeten? *Bodem* 5: 22-25.
- Knotters, M., D.J. Brus, S.J.E. Verzandvoort, M. Heinen, 2011. Aanvullende bodem-fysische gegevens voor BIS-Nederland. Wageningen, Alterra-rapport 2245.
- Kraijenhoff van de Leur, D. A. (1958). A study of non steady groundwater flow with special reference to a reservoir-coefficient. *De ingenieur* 70, 87-94.
- Kroes, J. G., J. C. van Dam, P. Groenendijk, R. F. A. Hendriks, C. M. J. Jacobs (2009). *SWAP version 3.2. Theory description and user manual*. Alterra-rapport 1649, update 2, augustus 2009. Alterra Wageningen UR, Wageningen.
- Kros, J. (2002). Evaluation of biogeochemical models at local and regional scale. PhD-thesis, Wageningen University. ISBN 9058085767 - p. 284.
- Kros, J., W. de Vries, P. H. M. Janssen, C. I. Bak (1993). The uncertainty in forecasting trends of forest soil acidification. *Water, Air, and Pollution* 66, 29-58.
- Kros, J., G. J. Reinds, W. de Vries, J. B. Latour, M. J. S. Bollen (1995). Modeling of soil acidity and nitrogen availability in natural ecosystems in response to changes in acid deposition and hydrology. SC-DLO report 95, Wageningen UR, Wageningen.
- Latour, J. B., I. G. Staritsky, J. R. M. Alkemade, J. Wiertz (1997). DE NATUURPLANNER – Decision Support Systeem natuur en milieu versie 1.1. RIVM-rapport 711901019, RIVM, Bilthoven.
- Mol-Dijkstra, J. P. (2005). Ontwikkeling en beheer van SMART2-SUMO – Ontwikkelings- en beheersplan en versiebeheerprotocol. WOT-werkdocument 6, Wageningen UR, Wageningen.
- Mol-Dijkstra, J. P., J. Kros, G. J. Reinds, M. Posch, H. J. J. Wieggers (2006). Model description and Users guide SMART2 version 3.4. Alterra rapport 1425, Wageningen UR, Wageningen.
- Mualem, Y. (1976). A new model predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research* 12: 513-522.
- Ollness, A., D. Archer (2005). Effect of organic carbon on available water in soil. *Soil Science* 170, 90-101.

- Oostermeijer, J. G. B., C. A. M. van Swaay (1998). The relationship between butterflies and environmental indicator values: a tool for conservation in a changing landscape. *Biological Conservation* 86, 271-280.
- Opdam, P., J. Verboom, R. Pouwels (2003). Landscape cohesion: an index for the conservation potential of landscapes for biodiversity. *Landscape Ecology* 18, 113-126.
- Pham, H. Q., D. G. Fredlund, S. L. Barbour (2005). A study of hysteresis models for soil-water characteristic curves. *Can. Geotech. J.* 42, 1548-1568.
- Pouwels, R., R. Jochem, M. J. S. M. Reijnen, S. R. Hensen, J. G. M. van der Greft (2002). LARCH voor ruimtelijk ecologische beoordelingen van landschappen. Alterra-rapport 492, Wageningen University & Research, Wageningen.
- Pouwels, R., J. G. M. van der Greft, M. H. C. van Adrichem, H. Kuipers, R. Jochem, M. J. S. M. Reijnen (2008). LARCH Status A. WOT-werkdocument 107, Wettelijke Onderzoekstaken Natuur & Milieu, Wageningen.
- Pouwels, R., M. van Eupen & H. Kuipers (2012). *MetaNatuurplanner 2.0 – achtergronddocument t.b.v. NVK2011*. Te verschijnen als WOT-werkdocument. WOT Natuur & Milieu, Wageningen UR, Wageningen.
- Querner, E. P., P. J. T. van Bakel (1989). *Description of the regional groundwater flow model SIMGRO*. Wageningen, DLO-Staring Centrum, report 7.
- Refsgaard, J. C., & Henriksen, H. J. (2004). Modelling guidelines—terminology and guiding principles. *Advances in Water Resources*, 27(1), 71–82. doi:10.1016/j.adwatres.2003.08.006
- Reijnen, M. J. S. M., A. van Hinsberg, M. L. P. van Esbroek, B. de Knecht, R. Pouwels, S. van Tol, J. Wiertz (2010). Natuurwaarde 2.0 land. graadmeter natuurkwaliteit landeco-systemen voor nationale doeleinden. WOT-rapport 110, WOT Natuur & Milieu, Wageningen UR, Wageningen.
- Rykiel, E. J. (1996). Testing ecological models: the meaning of validation. *Ecological Modelling*, 90(3), 229–244. doi:10.1016/0304-3800(95)00152-2
- Richards, L. A. (1931). Capillary conduction of liquids through porous mediums. *Physics* 1, 318-333.
- Ritsema, C. J., J. C. van Dam, J. L. Nieber, L. W. Dekker, K. Oostindie, T. S. Steenhuis (2000). *Preferential flow in water repellent sandy soils: Principles and modeling approaches*. Proceedings from the 2nd international symposium on preferential flow, Honolulu, Hawaii, January 3-5, 2001, p. 129-132.
- Runhaar, J.H., H. Kuijpers, H.L. Boogaard, E.P.A.G. Schouwenberg, P.C. Jansen (2003). Natuurgericht Landevaluatiesysteem (NATLES) versie 2.1. Wageningen, Alterra-rapport 550.
- Seibert, J., & McDonnell, J. J. (2002). On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. *Water Resources Research*, 38(11), 1241. doi:10.1029/2001WR000978
- Stolte, J., Wesseling J.G., S. Verzandvoort (2007). Kwaliteitsdocumentatie voor de verkrijging van Status A voor de gegevens van de Staringreeks zoals opgenomen in het gegevensbestand Priapus. Versie 1. Wageningen, Alterra-rapport 1522.
- Ten Brink (2000). Biodiversity indicators for the OECD Environmental Outlook and Strategy – A feasibility study. RIVM-report 402001014, Bilthoven.
- Ten Brink, B. J. E., A. van Hinsberg, M. de Heer, D. J. C. van der Hoek, B. de Knecht, O. M. Knol, W. Ligtoet, M. J. S. M. Reijnen, R. Rosenboom (2002). Technisch ontwerp Natuurwaarde 1.0 en toepassing in Natuurverkenning 2. RIVM-rapport 408657007, Bilthoven.

- Van Adrichem, M. H. C., F. G. Wortelboer, G. W. W. Wamelink (2010). MOVE. MOdel for terrestrial Vegetation version 4.0. WOT-werkdocument 153, Wettelijke Onderzoekstaken Natuur & Milieu, Wageningen.
- Van Dam, J. C., P. Groenendijk, R. F. A. Hendriks, J. G. Kroes (2008). Advances of modeling water flow in variably saturated soils with SWAP. *Vadose Zone J.* 7, 640-653.
- Van der Hoek, D. C. J., M. Bakkenes, J. R. M. Alkemade (2000). Natuurwaardering in de Natuurplanner. Toepassing voor de VIJNO. RIVM-rapport 408657004. RIVM, Bilthoven.
- Van der Hoek, D. C. J. & P. S. C. Heuberger (2006). Gevoeligheidsanalyse Natuur-planner – Van complex tot simpel. MNP rapport 500067001/2006.
- Van der Hoek, D. C. J. & M. E. Bakkenes (2007). *Natuurplanner 3.0 – Beschrijving en handleiding*. MNP rapport 500067002/2007.
- Van der Sluijs, J. P., J. A. Wardekker, P. Janssen (2012). *Integraal waterbeheer – kritische zone & onzekerheden*. Rapport binnen NMDC-innovatieproject, deelresultaat kwalitatieve onzekerheidsanalyse Baakse Beek.
- Van Genuchten, M. Th. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* 44, 892-898.
- Van Voorn, G. A. K., D. J. J. Walvoort (2011). *Evaluation of an evaluation list for model complexity*. WOT-werkdocument 272. WOT Natuur & Milieu, Wageningen UR, Wageningen.
- Van Voorn, G. A. K., D. J. J. Walvoort (2013). Complexiteitsanalyse Natuurplanner en MetaNatuurplanner. Te verschijnen als WOT-werkdocument.
- Van Voorn, G. A. K., D. J. J. Walvoort, M. Knotters, P. W. Bogaart, H. Houweling, P. H. M. Janssen (2011). *Een beoordelingslijst voor de complexiteit van modellen en bestanden*. WOT-paper 11. WOT Natuur & Milieu, Wageningen UR, Wageningen.
- Van Walsum, P. E. V. (1981). A simple model for the Hupselse Beek drainage-basin. Afstudeerscriptie, Dept. Hydraulica en Stroomgebiedhydrologie, Wageningen University & Research.
- Van Walsum, P. E. V. & P. Groenendijk (2008). Quasi steady-state simulation of the unsaturated zone in groundwater modeling of lowland regions. *Vadose Zone Journal* 7, 769-781.
- Van Walsum, P. E. V. & A. A. Veldhuizen (2011). *MetaSWAP_V7_2_0 – rapportage van activiteiten ten behoeve van certificering met Status A*. WOT-werkdocument 276. WOT Natuur & Milieu, Wageningen UR, Wageningen.
- Van Walsum, P. E. V., A. A. Veldhuizen, P. Groenendijk (2010). *SIMGRO 7.1.0. – Theory and model implementation*. Alterra-rapport 913.1, Alterra, Wageningen UR, Wageningen, actualisatie 2010.
- Vereecken, H., J. Diels, J. van Orshoven, J. Feyen, J. Bouma (1993). Functional evaluation of pedotransfer functions for the estimation of soil hydraulic properties. *Soil Science Society of America Journal* 56: 1371-1378.
- Verzandvoort, S., H. Vroon, J. Wesseling, G. Bakker, K. Oostindie, G. Stoffelsen, N. Heidema, G. Heuvelink (2012). Naar een database van bodemhydraulische karakteristieken voor Nederland. Wageningen, Alterra-rapport 2238.
- Visschers, R., P.A. Finke, J.J. de Gruijter (2007). A soil sampling program for the Netherlands. *Geoderma* 139: 60-72.
- Vrugt, J. A., Bouten, W., Gupta, H. V., & Sorooshian, S. (2002). Toward improved identifiability of hydrologic model parameters: The information content of experimental data. *Water Resources Research*, 38(12), 1312. doi:10.1029/2001WR001118

- Wagener, T., D. P. Boyle, M. J. Lees, H. S. Wheatler, H. V. Gupta, S. Sorooshian (2001). A framework for the development and application of hydrological models. *Hydrology and Earth System Sciences* 5(1), 13-26.
- Wagener, T., McIntyre, N., Lees, M. J., Wheatler, H. S., & Gupta, H. V. (2003). Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis. *Hydrological Processes*, 17(2), 455–476. doi:10.1002/hyp.1135
- Wagener, T., van Werkhoven, K., Reed, P., & Tang, Y. (2009a). Multiobjective sensitivity analysis to understand the information content in streamflow observations for distributed watershed modeling. *Water Resources Research*, 45(2), W02501. doi:10.1029/2008WR007347
- Wagener, T., Reed, P., Van Werkhoven, K., Tang, Y., & Zhang, Z. (2009b). Advances in the identification and evaluation of complex environmental systems models. *Journal of Hydroinformatics*, 11(3–4), 266. doi:10.2166/hydro.2009.040
- Walvoort, D. (2010). Prediction of topsoil pH-KCl for the Netherlands. Wageningen, Alterra-WUR, intern werkdocument.
- Wamelink, G. W. W. (2007). Simulation of vegetation dynamics as affected by nitrogen deposition. PhD-thesis, Wageningen UR, Wageningen.
- Wamelink, G. W. W. (2008a). Gevoeligheids- en onzekerheidsanalyse van SUMO. WOT-werkdocument 98, Wettelijke Onderzoekstaken Natuur & Milieu, Wageningen.
- Wamelink, G. W. W. (2008b). Technical documentation for SUMO2 v. 3.2.1. WOT-werkdocument 108, Wettelijke Onderzoekstaken Natuur & Milieu, Wageningen.
- Wamelink, G. W. W. (2009). Testdocument voor MOVE4 v1.0. Intern Alterra-document, Wageningen.
- Wamelink, G. W. W., H. van Oene, J. P. Mol-Dijkstra, J. Kros, H. F. van Dobben, F. Berendse (2001). Validatie van de modellen SMART2, SUMO 1, NUCOM en MOVE op site-, regionaal en nationaal niveau. Alterra-rapport 065, Wageningen.
- Wamelink, G. W. W., Mol-Dijkstra, J. P., Reinds, G. J. (2008). Herprogrammeren van SUMO2 - verbeteringen in het kader van de modelkwaliteitsslag. WOT-werkdocument 109, Wettelijke Onderzoekstaken Natuur & Milieu, Wageningen.
- Wamelink, G. W. W., H. F. van Dobben, F. Berendse (2009a). Vegetation succession as affected by decreasing nitrogen deposition, soil characteristics and site management: A modelling approach. *Forest Ecology and Management* 258, 1762-1773.
- Wamelink, G. W. W., H. J. J. Wieggers, G. J. Reinds, J. Kros, J. P. Mol-Dijkstra, M. van Oijen, W. de Vries (2009b). Modelling impacts of changes in carbon dioxide concentration, climate and nitrogen deposition on carbon sequestration by European forests and forest soils. *Forest Ecology and Management* 258, 1794-1805.
- Wamelink, G. W. W., R. M. Winkler, F. G. Wortelboer (2009c). User documentation MOVE4 v1.0. WOT-werkdocument 154, Wettelijke Onderzoekstaken Natuur & Milieu, Wageningen.
- Wamelink, G. W. W., L. M. W. Akkermans, D. J. Brus, G. B. M. Heuvelink, J. P. Mol, E. P. A. G. Schouwenberg (2011). Uncertainty analysis of SMART2-SUMO-P2E-MOVE4: the nature planner soil and vegetation model chain. WOT-werkdocument 108, Wettelijke Onderzoekstaken Natuur & Milieu, Wageningen.
- Werkgroep HELP-tabel (1987). De invloed van de waterhuishouding op de landbouwkundige productie. Mededeling 176. Landinrichtingsdienst, Utrecht.
- Wesseling, J. G. (1991). *CAPSEV – Steady state moisture flow theory. Program description. User Manual*. Wageningen, SC-rapport 37.

- Wösten, J. H. M., F. de Vries, J. Denneboom, A. F. van Holst (1988). *Generalisatie en bodemfysische vertaling van de bodemkaart van Nederland, 1 : 250 000, ten behoeve van de PAWN-studie*. STIBOKA, rapport 2055.
- Wösten, J. H. M., M. H. Bannink, J. Beuving (1987). Waterretentie- en doorlatendheidskarakteristieken van boven- en ondergronden in Nederland: de Staringreeks. Wageningen, Stiboka-rapport 1932, ICW-rapport 18.
- Wösten, J. H. M., G. J. Veerman en J. Stolte (1994). Waterretentie- en doorlatendheidskarakteristieken van boven- en ondergronden in Nederland: de Staringreeks. Vernieuwde uitgave 1994. Wageningen, Staring Centrum-DLO, Technisch Document 18.
- Wösten, J. H. M., G. J. Veerman, W. J. M. de Groot, J. Stolte (2001). *Waterretentie- en doorlatendheidskarakteristieken van boven- en ondergronden in Nederland: de Staringreeks*. Nieuwe uitgave 2001, Alterra-rapport 153, Wageningen UR, Wageningen.

Bijlage 1 Evaluatielijst Modelcomplexiteit v1.0

BASALE GEGEVENS

Geef aan: model/bestand, naam, versienummer, revisienummer, uitgiftedatum. Onder model verstaan we de broncode + invoergegevens.

DOEL & TOEPASSINGEN

- A. Wat was het doel van het bestand/model?
- B. Wat was het beoogde toepassingsgebied, en waaruit blijkt dit?
- C. Wat zijn de toepassingen nu? Heeft het model/bestand een rol in een modelketen? Zo ja, welke?
- D. Overlapt het beoogde toepassingsgebied alle daadwerkelijke toepassingen (ja/deels/ nee)?
- E. Indien relevant, in hoeverre overlappen de verschillende toepassingen met elkaar (goed/matig/slecht)?

SYSTEEMANALYSE

- A. Geef een systeemanalyse. Dit kan bv. een verwijzing naar een rapport zijn. Aspecten die daarin aan bod komen zouden moeten zijn: Wat zijn de belangrijkste attributen en processen van het systeem? Welke terugkoppelingen zijn aanwezig? Hoe zijn de grenzen van het systeem bepaald of gedefinieerd? Hoe wordt met relevante attributen en processen buiten deze grenzen omgegaan, bv. randvoorwaarden, constanten, of simpelweg genegeerd? Over welke processen, attributen, terugkoppelingen, etc. bestaat er onzekerheid? Hoe is die onzekerheid bepaald?

Beschouw nu per toepassing

- B. Welke systeem-analytische aspecten (attributen, processen, terugkoppelingen, etc.) zijn (direct) relevant voor de toepassing? En welke in mindere mate of niet? En hoe is die relevantie bepaald?
- C. Zijn alle voor de toepassing belangrijke aspecten als gegeven in B meegenomen in het model/bestand? En welke niet?
- D. Zijn er minder of niet belangrijke aspecten meegenomen in het model/bestand (ja/weinig/geen)?
- E. Beoordeel de mate van 'evenwicht' m.b.t. de systeemanalyse, gebaseerd op de antwoorden op 3C en 3D (goed/matig/slecht).

CONCEPTUEEL MODEL of ONTWERPMODEL

- A. Geef het conceptuele model (voor een model) of ontwerpmodel (voor een bestand). Dit kan bv. een verwijzing naar een rapport zijn. Het conceptuele model legt op relatief informele wijze de relaties tussen componenten vast. Aspecten die aan bod kunnen komen zijn: Welke aannamen en vereenvoudigingen zijn er gemaakt? Wat zijn de ruimtelijke dimensies? Welke ruimtelijk-temporele aggregatieniveaus worden gebruikt? Welke zaken spelen een rol voor het bestand? Wat wordt er voor (des)aggregatie van data gebruikt? Voor modellen: Welk modeltype is gebruikt?

Beschouw nu per toepassing

- B. Op welke (ruimtelijk-temporeel) aggregatieniveaus zijn antwoorden gewenst, gezien de toepassing? Welke ruimtelijke dimensies en welk modeltype/ontwerpmodel zijn gewenst?
- C. Komen de ruimtelijke dimensies, aggregatieniveaus, etc. overeen met die vanuit de toepassing zijn gewenst (ja/matig/niet)?

DATA

- A. Welke gegevens zijn benodigd, gezien de toepassing, en motiveer? Wees specifiek m.b.t. resolutie, nauwkeurigheid, schaal, etc., en vermijd algemene en ongemotiveerde opmerkingen als "file X moet gebruikt worden".

- B. Welke gegevens zijn beschikbaar, en in welke mate?
- C. Wat is uiteindelijk de invoer van het model of bestand?
- D. Komt de invoer overeen met de door de toepassing gewenste gegevens (ja/deels/nee)? Vergelijk de antwoorden op 5A, 5B en 5C. Let hierbij sterk op eenheden en de drie elementen van schaal (coverage, support, extent; zie Bierkens *et al*, 2000; Bogaart *et al*, 2011). Heeft de data de juiste dimensies en eenheden? Zo nee, is er een (des)aggregatiemethode voorhanden? Schaalproblemen doen zich overal voor door de 'vertaling' van attributen naar metingen, die gediscretiseerd zijn in tijd en ruimte, en onderhevig zijn aan allerlei bronnen van meetruis.

FORMEEL MODEL (niet voor databestanden)

Gebruik de antwoorden op vragen 5A tot 5D om vragen 6B tot 6D te antwoorden.

- A. Geef het formeel model, en motiveer de keuze of verwijst naar paginanummers uit een gepubliceerd rapport of artikel. Het formeel model is de omzetting van concept naar bijvoorbeeld wiskundige vergelijkingen of een set rekenregels. De keuze van formeel model hangt af van de beoogde toepassing en de data, en wordt mede bepaald door de gestelde onzekerheidsmarges.
- B. Welke gegevens zijn benodigd voor dit formele model?
- C. Welke gegevens zijn beschikbaar, en in welke mate?
- D. Zijn de gegevens die benodigd én beschikbaar zijn ook gebruikt? Waarom wel/niet?

NUMERIEK REKENMODEL of BESTAND

- A. Geef het numeriek rekenmodel/gegevensbestand (de implementatie in software), en motiveer de keuze of verwijst naar paginanummers uit een gepubliceerd rapport of artikel.
- B. Beschrijf de verificatie of verwijst naar paginanummers uit een gepubliceerd rapport of artikel. Verificatie moet niet verward worden met validatie. De eerste controleert de omzetting van formeel model naar code, de tweede toetst de resultaten van de code aan de werkelijkheid.
- C. Beschrijf en motiveer de keuze voor numerieke rekenmethoden, en geef de gebruikte discretisatie of verwijst naar paginanummers uit een gepubliceerd rapport of artikel. Veelal bestaat er bij numerieke rekenmethoden een afweging tussen rekensnelheid en –nauwkeurigheid.

SCHEMATISERING

De schematisering is bv. het indelen van de bodem in verschillende lagen ('stratificatie'), of het clusteren van landgebruikstypen. De mate van detail die bij schematisering wordt gevolgd is afhankelijk van de toepassing, maar moet wel voldoende ondersteund worden door de data.

- A. Indien van toepassing, beschrijf de schematisering, of geef relevante en specifieke referenties?
- B. Wordt de schematisering voldoende ondersteund door de data (ja/matig/nee)? Motiveer.

GEVOELIGHEDEN, ONZEKERHEDEN & OPBOUW CODE

Deze vraag is gericht op het bepalen welke parameters, forcing, invoer, delen van de code, etc. belangrijk zijn en welke overbodig zijn, en welke onderdelen een grote onzekerheid introduceren.

- A. Zijn er gevoeligheids-/onzekerheidsanalyses uitgevoerd? Zo ja, beschrijf deze of verwijst naar paginanummers uit een gepubliceerd rapport of artikel en analyseer de resultaten. Sommige factoren kunnen weinig van belang blijken voor de toepassing, zodat er weinig moeite hoeft te worden gedaan om deze goed te kalibreren, of gebruikt kunnen worden om het model te vereenvoudigen. Andere factoren kunnen juist essentieel blijken.
- B. (niet voor databestanden) Welke numerieke integratiemethode is gebruikt? Bespreek het prestatievermogen van de methode.
- C. Wat is de rol van elke module van het model/bestand, bv. een gewasgroei-deelmodel als onderdeel van een hydrologisch model, of stappen in het stroomschema of de scripts van een

- bestand? Bespreek dat in relatie met de toepassing. Zijn er onderdelen die overbodig zijn? Missen er nog onderdelen? Motiveer waarom.
- D. Beoordeel de complexiteit in termen van rekentijd en efficiëntie. Wat is de duur van een run? Hoeveel (tussentijdse) uitvoer is er? Hoe groot is de datastroom? Zou de code sneller/efficiënter kunnen, bv. door selectie, aggregatie, andere methode, andere modulaire opbouw, etc.?

KALIBRATIE (niet voor databestanden)

De kalibratie is het proces waarbij de parameters, begincondities, etc. van het model of bestand van waarden worden voorzien. Kalibratie is vaak gebaseerd op gevoeligheids-analyses (zie vorige vraag).

- A. Indien beschikbaar, beschrijf de uitgevoerde kalibraties en motiveer de keuze, of verwijfs naar paginanummers uit een gepubliceerd rapport of artikel. Geef ook aan welke doelfuncties (lokaal, globaal, deterministisch, stochastisch) zijn gekozen, of dat de betrouwbaarheid van de gekalibreerde parameters wordt gekwantificeerd, hoe, op welke data, of en hoe de betrouwbaarheidsintervallen zijn bepaald, etc. met eventuele verwijfsingen naar de literatuur. Let op over-fitting, meervoudige lokale optima, en identificeerbaarheidsproblemen.
- B. Welke eisen stelt de toepassing aan de nauwkeurigheid waarmee parameters, etc. wordt bepaald. Let ook op over-fitting.
- C. Welke mogelijkheden leveren de data om waarden van parameters, etc. te bepalen. Denk hierbij aan resolutie, nauwkeurigheid, aantal gegevens, het aggregatieniveau van de gegevens, en let op over-fitting.

VALIDATIE

Wij definiëren hier een validatie als een toets of het model of bestand een redelijke representatie is van het werkelijke systeem dat wordt gemodelleerd. Dit geschiedt op basis van een set onafhankelijke gegevens (gegevens die nog niet in een eerder stadium bij het modelleren gebruikt zijn).

- A. Indien beschikbaar, beschrijf de gedane validatiestudies of verwijfs naar paginanummers uit een gepubliceerd rapport of artikel.
- B. Onderbouw/beoordeel de waarde van de validatie(s). Hoe relevant zijn de verschillende validatiestudies voor de verschillende toepassing(en) van het model of bestand? Zijn alle toepassingen voldoende afgedekt door de studies? Is ook de betrouwbaarheid van de validatie gekwantificeerd? Welke toepassingen moeten opnieuw worden gevalideerd?

SAMENVATTING en ALGEMEEN OORDEEL

Dit laatste onderdeel is gericht op het geaggregeerde resultaat van de antwoorden op de verschillende vragen en de meningen van de verschillende betrokkenen, en kan indirect ook tot adviezen leiden ter verbetering van het model of bestand.

- A. Beoordeel de mate van vertrouwen in het model/bestand in het algemeen, mede gebaseerd op wat er is gedaan aan testen, verificatie, kalibratie, validatie, gevoeligheids- en onzekerheidsanalyses, en motiveer deze.
- B. Zijn er veel missende zaken? Zijn er veel vragen waar de zaken als 'te eenvoudig' zijn beoordeeld? Motiveer.
- C. Zijn er veel overbodige componenten? Zijn er veel onderdelen waar het eenvoudiger zou kunnen? Motiveer.
- D. Is er voldoende datasteun, en waaruit blijkt dat? Motiveer
- E. Welke specifieke suggesties voor verbetering en data-aanlevering komen naar voren uit deze analyse?

Bijlage 2 Evaluatielijst Modelcomplexiteit v2.0

De aanpassingen voor de lijst voor Evaluatie Modelcomplexiteit, besproken in hoofdstuk 6, zijn opgenomen in de voorgestelde nieuwe versie van de lijst, die in deze bijlage wordt weergegeven. Nogmaals willen we benadrukken, dat dit de huidige stand van zaken weergeeft, en niet de definitieve versie van de EMC betreft.

INTRODUCTIE

Deze evaluatielijst is gericht op het verkrijgen van een inhoudelijk oordeel t.a.v. de balans (het zgn. 'evenwicht') tussen de complexiteit van een model i.r.t. de toepassing van het model, de beschikbare kennis, en de beschikbare data (let op: het woord 'model' wordt breed opgevat, en kan een model, bestand, meta-model, graadmeter, of zelfs modelketen inhouden). We beschouwen een model in 'evenwicht' wanneer het

- Complex genoeg is om tegemoet te komen aan uit de toepassing voorkomende eisen qua schaal, nauwkeurigheid, flexibiliteit en geldigheid binnen externe factoren;
- En tegelijkertijd eenvoudig genoeg om de onzekerheden terug te dringen tot aanvaardbare proporties door een adequate ondersteuning door data en expert-kennis.

Parsimonie – het principe dat een model niet ingewikkelder zou moeten zijn dan nodig is om het gemodelleerde systeem of probleem goed te beschrijven of begrijpen – is al langer onderkend. Echter, met name door de opkomst van computers is de complexiteit van modellen een belangrijk punt geworden. Belangrijke zaken die snel limiterend worden zijn o.a.

- Data-ondersteuning;
- Ironisch genoeg, computerrekenkracht (en daarmee de mogelijkheden voor testen, modelanalyse, en validatie);
- Modelbegrip (en daarmee samenhangend het vertrouwen van stakeholders in de uitkomsten van het model).

Er zitten vele aspecten aan de complexiteit van een model. Er bestaan nochtans geen criteria om 'evenwicht' kwantitatief te benaderen die de vele aspecten tegelijk in ogenschouw nemen; in plaats daarvan beoogt deze evaluatielijst een manier te bieden om tot een kwalitatieve beoordeling van het evenwicht te komen. Verder wordt hier benadrukt, dat een evenwichtsbeoordeling *an sich* niet zegt of dat een model 'goed' of 'slecht' is; het vertelt alleen of de complexiteit van het model redelijk is in vergelijking met de toepassing van het model, en de gegevens die beschikbaar zijn. In de praktijk kunnen bij de analyse wel zwakke plekken t.a.v. het model zelf of het modelgebruik naar boven komen, en/of suggesties ter verbetering van het model of modelgebruik.

De vragenlijst is vooral gericht op het verkrijgen van inzicht in de gemaakte keuzes bij de ontwikkeling van het model; de geschiedenis van het model en die van de toepassing van het model zijn dus belangrijk om mee te nemen in de evenwichtsevaluatie. De lijst kan het beste worden ingevuld door een klein team, dat is samengesteld uit in elk geval een onafhankelijke auditeur, een persoon aan de kant van de modelontwikkeling, en een persoon aan de kant van het modelgebruik. De lijst bestaat uit 11 onderdelen, die elk uit een aantal deelvragen bestaan, veelal met detailvragen en/of een kleine toelichting. Verdere informatie over de achtergrond van deze lijst is te vinden in WOT-paper 11 (Van Voorn *et al*, 2011) en aanhangende bronnen.

1. BASISGEGEVENS

Geef aan: *Type model, naam, versienummer, revisienummer, uitgiftedatum. Onder model verstaan we de broncode + invoergegevens (die nodig zijn om het model te draaien). Bespreek ook de*

geschiedenis van het model en van de toepassing van het model. Indien het geëvalueerde model een modelketen is, beschrijf dan ook kort uit welke modules de modelketen bestaat, en wat hun rol in de keten is.

2. DOEL & TOEPASSINGEN

2A Wat is het originele doel van het model?

Deze deelvraag richt zich op de originele reden waarom het model ontwikkeld is, bv. "dit model was destijds bedoeld om de neerslag-afvoerdynamica te beschrijven in een gebiedsstudie."

2B Wat was het oorspronkelijk beoogde toepassingsgebied? Waaruit blijkt dit? Specificeer de toepassing zoveel mogelijk.

Deze deelvraag richt zich op de originele inkadering van het model, bv. de gebiedsstudie in bovenstaand voorbeeld. Bij kwantificering van de toepassing wordt bedoeld: Welke rol gaan de resultaten spelen in het verdere proces? Welke eisen ten aanzien van variabelen, acceptabele onzekerheden, tijd- en ruimte-schalen komen hieruit voort? Onder welke voorwaarden kan het beoogde model hier aan voldoen?

2C Wat zijn de toepassingen nu? Specificeer dit zoveel mogelijk.

Deze deelvraag richt zich op de rol van het model in huidige toepassingen, die niet *per se* overeen hoeven te stemmen met de originele toepassing en het originele doel. Toepassingen kunnen heel verschillende eisen stellen aan het model, bv. de neerslag-afvoerdynamica in gebied 1 is dermate anders dan in gebied 2, dat verschillende processen de boventoon voeren. Als de eisen van de verschillende toepassingen significant verschillen, dan is er sprake van meerdere toepassingen, die ook apart beschouwd zouden moeten worden. Er moeten dan ook meerdere evenwichtsevaluaties uitgevoerd worden.

2D Heeft het model een rol in een modelketen? Zo ja, beschrijf deze. Let wederom op de specificering.

Een rol in een modelketen/modelketen kan een belangrijke toepassing van een model zijn, die mogelijk weer aparte eisen aan het model stelt.

2E Is het model zelf een modelketen? Zo ja, geef dan ook een duidelijk overzicht van de verschillende doelen en (originele/beoogde/huidige) toepassingen van de verschillende modules. Specificeer de toepassingen zoveel mogelijk.

Een modelketen bestaat uit meerdere modellen en bestanden. Die kunnen om verschillende redenen ontwikkeld zijn, of een andere rol in de modelketen hebben dan waarvoor ze eigenlijk ontwikkeld waren.

2F Beschouw de antwoorden op de voorgaande deelvragen. Wat is de mate van overlap tussen oorspronkelijk doel en (beoogde/huidige) toepassingen. Zijn er (dus) meer dan 1/verschillende toepassingen die in vragen verderop apart behandeld/beschouwd zouden moeten worden? Indien ja, dan moet daar bij de beantwoording van de vragen vanaf nu rekening mee gehouden worden.

Als de evaluatie op een modelketen wordt uitgevoerd, dan moet dat expliciet aan de orde komen in deze vraag.

2G Wat is de gewenste uitvoer van het model? Motiveer.

Uitvoer kan breed opgevat worden hier, bv. waarden van variabelen op een bepaald tijdstip of bepaalde locatie, getallen in een databestand, of 'stoplichtwaarden'. Wat is gewenst vanuit de toepassing(en)?

3. SYSTEEMANALYSE

3A Geef een systeemanalyse. Dit kan bv. een verwijzing naar een rapport zijn. Welke keuzes t.a.v. de selectie zijn gemaakt (van hoe het systeem en zijn omgeving zijn opgebouwd, welke aspecten en componenten er onderdeel van uit maken, hoe ze in relatie tot elkaar staan, etc.)? Geef motivaties bij die keuzes.

Bij de systeemanalyse komen zaken aan bod als welke variabelen van belang zijn om mee te nemen, en welke fluxen, processen en terugkoppelingen moeten worden beschouwd. Hoe relevant deze zaken zijn in het licht van de toepassing? Dit geeft informatie die relevant is om een oordeel te geven t.a.v. de keuzes die zijn gemaakt om tot een conceptueel model te komen.

Indien er meerdere duidelijk verschillende toepassingen zijn, behandel deze dan ook apart bij het beantwoorden van onderstaande vervolgvragen.

Indien het hier een modelketen betreft: alvorens onderstaande deelvragen te beantwoorden, is het overzichtelijk om eerst een matrix van $(N + 1) \times M$ op te stellen, waarin N het aantal modules en M het aantal toepassingen is. Beschouw onderstaande deelvragen voor de gehele modelketen en per module.

3B Welke systeem-analytische aspecten zijn (direct) relevant voor de toepassing? En welke in mindere mate of niet? En hoe is die relevantie bepaald?

3C Welke gegevens zijn waarschijnlijk nodig voor het model, gezien doel en toepassing? Motiveer. Let op resolutie, eenheden, schaal, etc., en leg uit wat de rol van de data zou zijn.

Data kunnen op verschillende wijzen in het model worden gebruikt, nl. voor constanten (onveranderlijke grootheden), parameters (onveranderlijke grootheden die nog ingesteld moeten worden per toepassing), forcering (veranderlijke grootheden die niet door het model zelf beschreven worden), cascade-type variabelen (veranderlijke variabelen die beïnvloed worden door invoer, bv. forcering) of terugkoppelingstype-variabelen (veranderlijke grootheden die beïnvloed worden door invoer en zelf ook weer andere grootheden beïnvloeden), en voor verschillende zaken (afleiding processen, forcering, kalibratie, validatie). De centrale vraag is: stellen de data het model in staat om afdoende te voldoen aan de eisen vanuit de doelstelling? Deze vraag is weliswaar hier nog niet te beantwoorden, maar het is zinvol om relevante informatie t.a.v. beschikbaarheid en gebruik van data hier alvast op te sommen. Een belangrijk aspect is 'schaal', waarin onderscheid gemaakt kan worden tussen 'coverage' (dekkingsgraad, bv. metingen op uur-basis), 'support' (ondersteuning van de dekkingsgraad, b.v., de coverage op uur-basis is op basis van metingen die 1 seconde duren), en 'extent' (bereik, bv. van 0 tot 10 uur; zie Bierkens *et al*, 2000; Bogaart *et al*, 2011).

3D Welke gegevens zijn beschikbaar, en in welke mate?

Welke data is er, en welke niet? Heeft de beschikbare data de juiste schaal, dimensies, eenheden? Zo nee, is er een (des)aggregatiemethode voorhanden? Problemen met 'schaal' doen zich veelal voor door de 'vertaling' van attributen naar metingen, door discretisatie in tijd en ruimte, en het onderhevig zijn aan allerlei bronnen van meet-ruis (bv. het meten van de grondwaterstand, waarin de meting 'grondwaterstand' om allerlei redenen niet de daadwerkelijke 'grondwaterstand' correct hoeft weer te geven).

3E Wat is uiteindelijk de data/invoer die gebruikt wordt? Motiveer waarom, en beschrijf de rol (forcering, randvoorwaarde, kalibratie, etc.) van de verschillende data.

3F Geef een initiële beoordeling van het 'evenwicht' op basis van de systeemanalyse, gebaseerd op de antwoorden op de voorgaande deelvragen en de deelvragen uit onderdeel 2/.

Geef een eerste indruk van het 'evenwicht' van het model i.r.t. de toepassing. Zijn er zinnige en goed gemotiveerde keuzes gemaakt in het licht van wat gewenst is (doel, toepassing) en wat beschikbaar is (kennis, data)?

3G Afhankelijk van de voorgaande antwoorden, zijn er suggesties ter verbetering van de systeemanalyse of het model te geven? Motiveer.

4. CONCEPTUEEL MODEL/ONTWERPMODEL

4A Geef het conceptuele model. Dit kan bv. een verwijzing naar een rapport zijn.

Het conceptuele model legt op relatief informele (m.a.w., kwalitatieve) wijze de relaties tussen componenten vast. Relevante aspecten zijn: Wat zijn de belangrijkste attributen en processen van het systeem, en welke zijn minder of niet belangrijk? Wat zijn de variabelen, en van welk type zijn ze (constant, parameters, forcering, cascade-type/response, terugkoppelingstype)? Welke terugkoppelingen zijn aanwezig? Hoe zijn de grenzen van het systeem bepaald of gedefinieerd? Hoe wordt met relevante attributen en processen buiten deze grenzen omgegaan, bv. worden ze meegenomen als randvoorwaarden, constanten, of worden ze simpelweg genegeerd? Waarom? Over welke processen, attributen, terugkoppelingen, etc. bestaat er onzekerheid? Hoe is die onzekerheid ingeschat? Welke aannamen en vereenvoudigingen zijn er gemaakt? Wat zijn de ruimtelijke dimensies? Welke schaalproblemen spelen er? Welke ruimtelijk-temporele aggregatieniveaus worden gebruikt? Voor bestanden: welke zaken spelen een rol voor het bestand? Wat wordt er voor (des)aggregatie van data gebruikt?

4B Zijn alle voor de toepassing belangrijke aspecten als gegeven in 3A en 3B (de systeemanalyse) meegenomen in het conceptuele model?

Welke zaken zijn meegenomen na de systeemanalyse, en waarom? Is die keuze terecht geweest? Het gaat hier met name over keuzes op een fundamenteel niveau: zou er gebruik moeten worden gemaakt van een "gelumped" of balans-model, of moet er gebruik gemaakt worden van een ruimtelijk expliciet, dynamisch model? Vragen over de exacte spatio-temporele resolutie komen later aan de orde.

4C Zijn er minder of niet belangrijke aspecten als gegeven in 3A en 3B (de systeemanalyse) meegenomen in het conceptuele model (ja/weinig/geen)?

Welke zaken zijn weggelaten na de systeemanalyse, en waarom? Is die keuze terecht geweest?

Indien er duidelijk meerdere toepassingen zijn, beschouw per toepassing:

4D Op welke (ruimtelijk-temporeel) aggregatieniveaus zijn antwoorden gewenst, gezien de toepassing? Welke ruimtelijke dimensies en welk modeltype/ontwerpmodel zijn gewenst?

4E Welke gegevens zijn nodig voor dit onderdeel, gezien doel en toepassing? Motiveer. Let op resolutie, schaal, etc., en leg uit wat de rol van de data zou zijn.

4F Welke gegevens zijn beschikbaar, en in welke mate?

Veelal wordt expertkennis ingezet bij het vaststellen van het conceptuele model.

4G Wat is uiteindelijk de data/invoer die gebruikt wordt? Motiveer.

4H Stellen de in het model doorgevoerde vereenvoudigingen het model in staat om aan de gewenste voorwaarden te voldoen? Komen de gebruikte ruimtelijke dimensies, aggregatieniveaus, etc. overeen met diegene die vanuit de toepassing zijn gewenst (ja/matig/niet)?

4I Afhankelijk van de voorgaande antwoorden, zijn er suggesties ter verbetering van het conceptuele model/ontwerpmodel te geven? Motiveer.

5. FORMEEL MODEL (niet relevant voor eenvoudige bestanden)

5A Geef het formele model, of verwijst naar paginanummers uit een gepubliceerd rapport of artikel. Bespreek nadrukkelijk de opties, keuzes, aannames en vereenvoudigingen bij het opstellen van het formele model.

Het formele model is de omzetting van conceptueel model naar bijvoorbeeld wiskundige vergelijkingen of een set rekenregels. Leg uit wat de variabelen zijn (type, deterministisch of stochastisch), wat voor vergelijkingen, welke vereenvoudigingen en aannames er zijn gemaakt, grenscondities, welke keuzes zijn gemaakt, welke motivaties daarachter zitten, wat verdere opties waren, etc.

Indien er duidelijk meerdere toepassingen zijn, beschouw per toepassing:

5B Welke gegevens zijn nodig voor dit onderdeel, gezien doel en toepassing? Motiveer. Let op resolutie, schaal, etc., en leg uit wat de rol van de data zou zijn.

Deze analyse is sterk gekoppeld aan de beschrijving van het formele model. Immers, hieruit volgt welke variabelen niet door het model zelf worden berekend, maar extern moeten worden gespecificeerd.

5C Welke gegevens zijn beschikbaar, en in welke mate?

Geef hierbij aan per variabele/proces welke data beschikbaar zijn, en welke informatie deze bevatten. Men kan denken hierbij aan bv. een extern gegevensbestand, het resultaat van een kalibratieprocedure (met bijbehorende informatie-extractie, zie vraag 9 t.a.v. kalibratie), etc.

5D Zijn er beschikbare gegevens die niet gebruikt zijn? Motiveer.

Men kan hierbij bv. denken aan restricties gerelateerd aan privacy, bronbescherming, of financiën, maar ook aan het weglaten van data die weinig informatie bevatten (bv. gebaseerd op de Fisher-informatie die volgt uit een gevoeligheidsanalyse).

5E Beoordeel de mate van 'evenwicht' a.d.h.v. het formele model. Zijn de gemaakte keuzes, aannames en vereenvoudigingen achter het formele model in overeenstemming met doel, toepassing, en data?

De keuze van formeel model hangt af van de beoogde toepassing en de data, en wordt mede bepaald door de gewenste onzekerheidsmarges en andere eisen die de toepassing stelt aan het model.

5F Afhankelijk van de voorgaande antwoorden, zijn er suggesties ter verbetering van het formele model te geven? Motiveer.

6. SCHEMATISERING

6A Indien het model ruimtelijk gedistribueerd is, beschrijf de gebruikte schematisering (discretisatie en bijbehorende bepaling van parameterwaarden), of geef relevante en specifieke referenties? Let specifiek op de onderliggende heterogeniteit, en hoe deze is vereenvoudigd.

In de praktijk bestaat er altijd heterogeniteit in het systeem dat met het model beschreven wordt. De schematisering is bv. het indelen van de bodem in verschillende lagen ('stratificatie'), of het clusteren van landgebruikstypen (let op: dit is dus zeker ook relevant voor bestanden). De mate van detail die bij schematisering wordt gevolgd is afhankelijk van de toepassing, maar moet wel voldoende ondersteund worden door de data. Deze vraag is eigenlijk altijd van toepassing, maar een keuze zou kunnen zijn, dat het model een set van gewone differentiaalvergelijkingen betreft: de schematisering is dan dat alles als een puntsysteem wordt beschouwd.

Indien er duidelijk meerdere toepassingen zijn, beschouw per toepassing:

6B Welke schematisering zou gezien de toepassing gewenst zijn? Motiveer.

6C Stemt de daadwerkelijke schematisering van het model overeen met wat gewenst is vanuit de toepassing? Waarom wel/niet?

6D Wordt de schematisering voldoende ondersteund door de data? Motiveer.

6E Afhankelijk van de voorgaande antwoorden, zijn er suggesties ter verbetering van de schematisering te geven? Motiveer.

7. NUMERIEKE IMPLEMENTATIE

7A Geef een overzicht van de code van de numerieke implementatie. Beschrijf de rol van elke module van het model/bestand, bv. een gewasgroei-deelmodel als onderdeel van een hydrologisch model, of stappen in het stroomschema of de scripts van een bestand. Verwijs eventueel naar een gepubliceerd rapport of artikel met de broncode.

De numerieke implementatie omvat niet alleen het rekenhart, maar ook bv. de stappen voor pre- en post-processing. Dit is dus ook relevant voor bestanden, omdat daarin altijd pre- en post-processing plaatsvindt, bv. het toekennen van attributen aan ruimtecoördinaten.

Indien er duidelijk meerdere toepassingen zijn, beschouw per toepassing de volgende vragen:

7B Beschrijf en motiveer de keuzes die gemaakt zijn bij de numerieke implementatie (zoals numerieke integratiemethode, opbouw modules, etc.).

Bv. bij numerieke rekenmethoden bestaat er vaak een afweging tussen rekensnelheid en nauwkeurigheid. Verwijs eventueel naar paginanummers uit een gepubliceerd rapport of artikel.

7C Geef een overzicht van de verificatie en testen waaruit blijkt dat de numerieke implementatie een goede omzetting van formeel model naar code is. Verwijs evt. naar paginanummers uit een gepubliceerd rapport of artikel.

7D Welke gegevens zijn nodig voor de numerieke implementatie, en dan met name de invoer, gezien doel en toepassing? Motiveer. Let op resolutie, schaal, etc., en leg uit wat de rol van de data zou zijn. Veel modellen hebben bv. gegevens nodig om daadwerkelijk te draaien, zoals bv. begincondities (waarden van veranderlijke grootheden op het begintijdstip) of gegevens t.a.v. massa-balansen. Dit geldt ook voor modellen, bv. voor bestanden die natuurtypen aangeven op een ruimtelijke schaal van 25-bij-25 meter zijn ook invoergegevens over natuurtypen of gegevens voor de afleiding van die natuurtypen op die schaal nodig.

7E Worden de gewenste data zoals aangegeven bij 7D ook daadwerkelijk gebruikt? Waarom (niet)?

7F Wat is de uitvoer van het model?

'Uitvoer' kan breed opgevat worden. Bv. variabelen kunnen uitvoer zijn, of kaarten uit een databank, etc.

7G Vergelijk met het antwoord op deelvraag 2G. Stemt de uitvoer van het model overeen met wat gewenst was?

7H Beoordeel de complexiteit van de numerieke implementatie in relatie tot de toepassing. Motiveer.

De evaluatie t.a.v. 'evenwicht' in de complexiteit in relatie tot de toepassing is juist relevant voor de numerieke implementatie. Immers, het computermodel is het model waarmee in de praktijk (lees:

voor toepassingen) altijd wordt gewerkt. Dit staat in principe los van de beoordeling, of dat het conceptuele model een evenwichtige complexiteit heeft; in theorie kan een conceptueel model erg eenvoudig zijn, maar toch een zeer ingewikkelde numerieke implementatie kennen. Belangrijke aspecten zijn onder meer: Klopt de uitvoer van het model met de wensen vanuit doel en toepassing? Waarom wel/niet? Zijn de goede keuzes gemaakt met het oog op de toepassing? Welke onderdelen missen/zijn overbodig? Wat is de duur van een run? Hoeveel (tussentijdse) uitvoer is er? Hoe groot is de datastroom? Zou de code sneller/efficiënter kunnen, bv. door selectie, aggregatie, andere methode, andere modulaire opbouw, etc.? Bv., een versnelling van de code is altijd zinnig, maar het heeft alleen prioriteit als dit voor de toepassing van belang is.

7I Afhankelijk van de voorgaande antwoorden, zijn er suggesties ter verbetering van de numerieke implementatie te geven? Motiveer.

8. GEVOELIGHEIDSANALYSE & ONZEKERHEIDSANALYSE

Indien er duidelijk meerdere toepassingen zijn, beschouw per toepassing:

8A Geef de referenties naar eventuele gevoeligheids- en/of onzekerheidsanalyses. Welke methoden zijn gebruikt voor de uitvoering? Wat waren de resultaten van deze studies (bv. welke epistemische en aleatorische onzekerheden zijn er, of wat waren de gevoelige parameters)? Motiveer dit, met name in het licht van de toepassing en de beschikbare gegevens.

Deze vraag is gericht op het bepalen welke parameters, forcering, invoer, delen van de code, etc. belangrijk zijn en welke overbodig zijn, en welke onderdelen een grote onzekerheid introduceren. Sommige factoren kunnen weinig van belang blijken voor de toepassing, zodat er weinig moeite hoeft te worden gedaan om deze goed te kalibreren, of gebruikt kunnen worden om het model te vereenvoudigen. Andere factoren kunnen juist essentieel blijken. Welke parameters, invoer, factoren, etc. zijn gevoelig en onzeker, en welke niet? Welke zouden eventueel weggelaten kunnen worden? Geef aan welke methoden zijn gebruikt, bv. "one-at-a-time", variantie-gebaseerde methoden, decompositie (zoals FAST of Sobol'), bifurcatie-analyse. Geef ook aan welke nominale sets en waarschijnlijkheidsverdelingen zijn gebruikt voor de verschillende parameters, en motiveer alle keuzes.

8B Indien er geen onzekerheidsanalyses zijn gedaan, geef dan een overzicht hier van de mogelijke bronnen van onzekerheden. Schat, indien mogelijk, ook in wat hun invloed kan zijn t.a.v. de toepassing.

Een bruikbaar format voor een onzekerheidsanalyse is de Leidraad Onzekerheid van het PBL. Probeer, waar mogelijk, aan te geven wat de zwakke plekken in het geheel zijn, en hoe relevant die zijn voor de toepassing van het model. Bv. de schematisering is zwak onderbouwd, er is grote onzekerheid over de waarden van de forcering, een proces is constant verondersteld bij gebrek aan kennis en data, etc.

8C Afhankelijk van de voorgaande antwoorden, zijn er suggesties ter verbetering van de gevoeligheids- of onzekerheidsanalyses, of het model zelf te geven? Motiveer.

9. KALIBRATIE

9A Indien beschikbaar, beschrijf de uitgevoerde kalibraties, of verwijz naar pagina's uit een gepubliceerd rapport of artikel. Wat is uiteindelijk de data/invoer die gebruikt wordt?

Kalibratie is het proces waarbij de parameters, begincondities, etc. van het model of bestand van waarden worden voorzien, door middel van een al dan niet geautomatiseerd, iteratief proces. De term 'kalibratie' is lastig waar het regressie-modellen betreft, omdat deze modellen eigenlijk het resultaat van de regressie op data zijn. Kalibratie wordt vaak gecombineerd met gevoeligheidsanalyses. Wat zijn de vrije parameters? Welke doelfunctie(s) is/zijn gekozen (lokaal,

globaal, deterministisch, stochastisch)? Let op over-fitting, meervoudige lokale optima, en identificeerbaarheidsproblemen. Geef ook aan wanneer er synthetische data is gebruikt, zoals in het geval van meta-modellen.

Indien er duidelijk meerdere toepassingen zijn, beschouw per toepassing:

9B Welke eisen stelt de toepassing aan de kalibratie? Welke doelfunctie(s) past/passen het beste bij de toepassing? Welke gegevens zijn nodig voor dit onderdeel, gezien doelfunctie) en toepassing? Motiveer.

De kalibratie vindt altijd plaats in het kader van een toepassing. Welke nauwkeurigheid wordt vereist? Let op resolutie, schaal, etc., en leg uit wat de rol van de kalibratie-data zou zijn. De keuze voor doelfunctie is meestal een minimalisatie van de MSE/RMSE ('Mean Squared Error' en de wortel daarvan, respectievelijk), maar er zijn vele andere mogelijkheden, die tot significant andere resultaten kunnen leiden.

9C Welke mogelijkheden leveren de data t.a.v. de kalibratie? Motiveer.

Voor kalibratie zijn gegevens nodig om waarden van parameters, etc. te bepalen. Denk hierbij aan resolutie, nauwkeurigheid, aantal gegevens, het aggregatieniveau van de gegevens, en let op over-fitting. Passen de beschikbare gegevens bij de gewenste doelfunctie(s)?

9D Stemt de uitgevoerde kalibratie overeen met wat gewenst is vanuit de toepassing, en ondersteund wordt met data? Motiveer.

Welke doelfunctie is gebruikt? En welke was gewenst? Bv., in de praktijk wordt veel gebruik gemaakt van de MSE of RMSE, maar voor de toepassing zou een andere maat meer geschikt kunnen blijken. Waren de gewenste data ook beschikbaar? Etc.

9E Afhankelijk van de voorgaande antwoorden, zijn er suggesties ter verbetering van de kalibratie te geven? Motiveer.

10. VALIDATIE

Indien er duidelijk meerdere toepassingen zijn, beschouw per toepassing de volgende vragen:

10A Indien beschikbaar, beschrijf de gedane validatiestudies of verwijst naar paginanummers uit een gepubliceerd rapport of artikel.

Wij definiëren hier een validatie als een toets of het model of bestand een redelijke representatie is van het werkelijke systeem dat wordt gemodelleerd. Dit geschiedt op basis van een set onafhankelijke gegevens (gegevens die nog niet in een eerder stadium bij het modelleren gebruikt zijn). Een andere veelvoorkomende optie is door het model en de uitvoer ervan te vergelijken met (de uitvoer van) andere modellen/bestanden die betrekking hebben op dezelfde toepassing (hierbij is het dan van belang, dat men vertrouwen heeft in de kwaliteit van de modellen/bestanden, waarmee vergeleken wordt).

10B Indien het een modelketen betreft: Zijn alle modules gevalideerd? Welke wel/niet? Zijn de modules gevalideerd in de context van de toepassing van de modelketen?

10C Welke gegevens zouden idealiter nodig zijn voor een adequate validatie, gezien doel en toepassing? Motiveer. Let op resolutie, schaal, etc.

Bedenk welke gegevens gewenst zouden zijn, gegeven onbeperkte financiële middelen en ongelimiteerde mogelijkheden, en leg uit hoe ze gebruikt zouden moeten worden om tot een maximaal haalbare validatie van het model te komen.

10D Welke gegevens zijn daadwerkelijk beschikbaar voor validatie, en in welke mate?

In de praktijk zijn gegevens voor validatie vaak gelimiteerd, omdat de data beperkt zijn en alle data nodig zijn voor de eerdere stappen in het modelleringsproces. Data worden ook wel 'hergebruikt', bv. in kruisvalidatie.

10E Wat is uiteindelijk de data/invoer die gebruikt wordt? Motiveer.

Door de praktijk zal het antwoord op deze deelvraag veelal samenvallen met het antwoord op de vorige deelvraag. Er kunnen echter redenen zijn waarom beschikbare data niet is gebruikt, bv. omdat de betrouwbaarheid van de data te wensen overlaat.

10F Onderbouw/beoordeel de waarde van de validatie(s). Motiveer.

Hoe relevant zijn de verschillende validatiestudies voor de verschillende toepassing(en) van het model of bestand? Zijn alle toepassingen voldoende afgedekt door de studies? Indien mogelijk, is ook de betrouwbaarheid van de validatie gekwantificeerd, en hoe? Welke toepassingen moeten opnieuw worden gevalideerd? Bij validatiestudies kunnen andere doelfuncties zijn gebruikt dan bij de kalibratie. In de praktijk zijn er diverse modellen (bv. graadmeters, (delen van) natuur- en milieuprocesmodellen, en Complex Adaptieve Systemen) waar validatie zeer lastig is, principieel dan wel in de praktijk.

10G Afhankelijk van de voorgaande antwoorden, zijn er suggesties ter verbetering van de validatie of het model zelf te geven? Motiveer.

11. SAMENVATTING en ALGEMEEN OORDEEL

Dit laatste onderdeel is gericht op het geaggregeerde resultaat van de antwoorden op de verschillende deelvragen bij alle voorgaande onderdelen, en de meningen van de verschillende betrokkenen, en kan idealiter ook al leiden tot adviezen ter verbetering van het model zelf of analyses van het model. Beschouw bij elk van de vragen in elk geval doel, (de verschillende) toepassingen, databeschikbaarheid, en de historische context van het model.

11A Beoordeel de mate van vertrouwen die verschillende stakeholders in het model/bestand hebben. Motiveer.

Dit oordeel wordt mede gebaseerd op wat er is gedaan aan testen, verificatie, kalibratie, validatie, gevoeligheids- en onzekerheidsanalyses, etc. Bruikbare informatie kan worden geleverd door aanwezige leden van het team, bv. wat zij weten over hoe anderen over het model denken.

Indien er duidelijk meerdere toepassingen zijn, beschouw per toepassing:

11B Zijn er veel missende zaken? Zijn er veel vragen waar de zaken als 'te eenvoudig' zijn beoordeeld? Motiveer.

11C Zijn er veel overbodige componenten? Zijn er veel onderdelen waar het eenvoudiger zou kunnen? Motiveer.

11D Is er voldoende datasteun, en waaruit blijkt dat? Motiveer.

11E Beoordeel nu (gebruikmakend van alle voorgaande antwoorden) of het model evenwichtig is of niet (complexiteit i.r.t. data en toepassing). Motiveer

11F Welke specifieke suggesties voor verbetering van het model of modelgebruik, data-aanlevering, en verder te ontwikkelen expert-kennis komen naar voren uit deze analyse?

Maak ook gebruik van de antwoorden op eerdere deelvragen t.a.v. (aspecten van) modelcomplexiteit (deelvragen 3F, 3G, 4I, 5E, 5F, 6E, 7H, 7I, 8C, 9E, en 10G). Motiveer de suggesties. Beschouw de keuzes en aannames die gemaakt zijn bij de verschillende onderdelen. Welke opties zijn overwogen? Welke zijn redelijk, mogelijk, en zinnig? Wat valt te verbeteren?

11G Doorgaand op de vorige deelvraag: Is er een schatting te maken van de kosten om de suggesties uit te voeren? Motiveer.

Het antwoord op deze vraag is vrijblijvend, maar kan in elk geval keuzes bieden met het oog op eventuele doorontwikkeling van het model en het gebruik van het model in de toekomst.

Verschenen documenten in de reeks Werkdocumenten van de Wettelijke Onderzoekstaken Natuur & Milieu vanaf 2010

Werkdocumenten zijn verkrijgbaar bij het secretariaat van Unit Wettelijke Onderzoekstaken Natuur & Milieu, te Wageningen. T 0317 – 48 54 71; E info.wnm@wur.nl

De werkdocumenten zijn ook te downloaden via de WOT-website www.wageningenUR.nl/wotnatuurenmilieu

2010

- 174** *Boer de, S., M.J. Bogaardt, P.H. Kersten, F.H. Kistenkas, M.G.G. Neven & M. van der Zouwen.* Zoektocht naar nationale beleidsruimte in de EU-richtlijnen voor het milieu- en natuurbeleid. Een vergelijking van de implementatie van de Vogel- en Habitatrichtlijn, de Kaderrichtlijn Water en de Nitraatrichtlijn in Nederland, Engeland en Noordrijn-Westfalen
- 175** *Jaarrapportage 2009.* WOT-04-001 – Koepel
- 176** *Jaarrapportage 2009.* WOT-04-002 – Onderbouwend Onderzoek
- 177** *Jaarrapportage 2009.* WOT-04-003 – Advisering Natuur & Milieu
- 178** *Jaarrapportage 2009.* WOT-04-005 – M-AVP
- 179** *Jaarrapportage 2009.* WOT-04-006 – Natuurplanbureauafunctie
- 180** *Jaarrapportage 2009.* WOT-04-007 – Milieuplanbureauafunctie
- 181** *Annual reports for 2009;* Programme WOT-04
- 182** *Oenema, O., P. Bikker, J. van Ham, E.A.A. Smolders, L.B. Sebek, M. van den Berg, E. Stehfest & H. Westhoek.* Quicksan opbrengsten en efficiëntie in de gangbare en biologische akkerbouw, melkveehouderij, varkenshouderij en pluimveehouderij. Deelstudie van project 'Duurzame Eiwitvoorziening'
- 183** *Smits, M.J.W., N.B.P. Polman & J. Westerink.* Uitbreidingsmogelijkheden voor groene en blauwe diensten in Nederland; Ervaringen uit het buitenland
- 184** *Dirkx, G.H.P. (red.).* Quick responsefunctie 2009. Verslag van de werkzaamheden
- 185** *Kuhlman, J.W., J. Luijt, J. van Dijk, A.D. Schouten & M.J. Voskuilen.* Grondprijkskaarten 1998-2008
- 186** *Slangen, L.H.G., R.A. Jongeneel, N.B.P. Polman, E. Lianouridis, H. Leneman & M.P.W. Sonneveld.* Rol en betekenis van commissies voor gebiedsgericht beleid
- 187** *Temme, A.J.A.M. & P.H. Verburg.* Modelling of intensive and extensive farming in CLUE
- 188** *Vreke, J.* Financieringsconstructies voor landschap
- 189** *Slangen, L.H.G.* Economische concepten voor beleidsanalyse van milieu, natuur en landschap
- 190** *Knotters, M., G.B.M. Heuvelink, T. Hoogland & D.J.J. Walvoort.* A disposition of interpolation techniques
- 191** *Hoogeveen, M.W., P.W. Blokland, H. van Kernebeek, H.H. Luesink & J.H. Wisman.* Ammoniakemissie uit de landbouw in 1990 en 2005-2008
- 192** *Beekman, V., A. Pronk & A. de Smet.* De consumptie van dierlijke producten. Ontwikkeling, determinanten, actoren en interventies.
- 193** *Polman, N.B.P., L.H.G. Slangen, A.T. de Blaeij, J. Vader & J. van Dijk.* Baten van de EHS; De locatie van recreatiebedrijven
- 194** *Veeneklaas, F.R. & J. Vader.* Demografie in de Natuurverkenning 2011; Bijlage bij WOT-paper 3
- 195** *Wascher, D.M., M. van Eupen, C.A. Mûcher & I.R. Geijzendorffer.* Biodiversity of European Agricultural landscapes. Enhancing a High Nature Value Farmland Indicator
- 196** *Apeldoorn van, R.C., I.M. Bouwma, A.M. van Doorn, H.S.D. Naeff, R.M.A. Hoefs, B.S. Elbersen & B.J.R. van Rooij.* Natuurgebieden in Europa: bescherming en financiering
- 197** *Brus, D.J., R. Vasat, G. B. M. Heuvelink, M. Knotters, F. de Vries & D. J. J. Walvoort.* Towards a Soil Information System with quantified accuracy; A prototype for mapping continuous soil properties
- 198** *Groot, A.M.E. & A.L. Gerritsen, m.m.v. M.H. Borgstein, E.J. Bos & P. van der Wielen.* Verantwoording van de methodiek Achtergronddocument bij 'Kwalitatieve monitor Systeeminnovaties verduurzaming landbouw'
- 199** *Bos, E.J. & M.H. Borgstein.* Monitoring Gesloten voer-mest kringlopen. Achtergronddocument bij 'Kwalitatieve monitor Systeeminnovaties verduurzaming landbouw'
- 200** *Kennismarkt 27 april 2010;* Van onderbouwend onderzoek Wageningen UR naar producten Planbureau voor de Leefomgeving
- 201** *Wielen van der, P.* Monitoring Integrale duurzame stallen. Achtergronddocument bij 'Kwalitatieve monitor Systeeminnovaties verduurzaming landbouw'
- 202** *Groot, A.M.E. & A.L. Gerritsen.* Monitoring Functionele agrobiodiversiteit. Achtergrond-document bij 'Kwalitatieve monitor Systeeminnovaties verduurzaming landbouw'
- 203** *Jongeneel, R.A. & L. Ge.* Farmers' behavior and the provision of public goods: Towards an analytical framework
- 204** *Vries, S. de, M.H.G. Custers & J. Boers.* Storende elementen in beeld; de impact van menselijke artefacten op de landschapsbeleving nader onderzocht
- 205** *Vader, J. J.L.M. Donders & H.W.B. Bredenoord.* Zicht op natuur- en landschapsorganisaties; Achtergronddocument bij Natuurverkenning 2011
- 206** *Jongeneel, R.A., L.H.G. Slangen & N.B.P. Polman.* Groene en blauwe diensten; Een raamwerk voor de analyse van doelen, maatregelen en instrumenten
- 207** *Letourneau, A.P., P.H. Verburg & E. Stehfest.* Global change of land use systems; IMAGE: a new land allocation module
- 208** *Heer, M. de.* Het Park van de Toekomst. Achtergronddocument bij Natuurverkenning 2011
- 209** *Knotters, M., J. Lahr, A.M. van Oosten-Siedlecka & P.F.M. Verdonschot.* Aggregation of ecological indicators for mapping aquatic nature quality. Overview of existing methods and case studies
- 210** *Verdonschot, P.F.M. & A.M. van Oosten-Siedlecka.* Graadmeters Aquatische natuur. Analyse gegevenskwaliteit Limnodata
- 211** *Linderhof, V.G.M. & H. Leneman.* Quicksan kosteneffectiviteitsanalyse aquatische natuur
- 212** *Leneman, H., V.G.M. Linderhof & R. Michels.* Mogelijkheden voor het inbrengen van informatie uit de 'KRW database' in de 'KE database'
- 213** *Schrijver, R.A.M., A. Corporaal, W.A. Ozinga & D. Rudrum.* Kosteneffectieve natuur in landbouwgebieden; Methode om effecten van maatregelen voor de verhoging van biodiversiteit in landbouwgebieden te bepalen, een test in twee gebieden in Noordoost-Twente en West-Zeeuws-Vlaanderen
- 214** *Hoogland, T., R.H. Kemmers, D.G. Cirkel & J. Hunink.* Standplaatsfactoren afgeleid van hydrologische model uitkomsten; Methode-ontwikkeling en toetsing in het Drentse Aa-gebied
- 215** *Agricola, H.J., R.M.A. Hoefs, A.M. van Doorn, R.A. Smidt & J. van Os.* Landschappelijke effecten van ontwikkelingen in de landbouw
- 216** *Kramer, H., J. Oldengarm & L.F.S. Roupioz.* Nederland is groener dan kaarten laten zien; Mogelijkheden om 'groen' beter te inventariseren en monitoren met de automatische classificatie van digitale luchtfoto's
- 217** *Raffe, J.K. van, J.J. de Jong & G.W.W. Wamelink (2011).* Kostenmodule Natuurplanner; functioneel ontwerp en software-validatie
- 218** *Hazeu, G.W., Kramer, H., J. Clement & W.P. Daamen (2011).* Basiskaart Natuur 1990rev
- 219** *Boer, T.A. de.* Waardering en recreatief gebruik van Nationale Landschappen door haar bewoners
- 220** *Leneman, H., A.D. Schouten & R.W. Verburg.* Varianten van natuurbeleid: voorbereidende kostenberekeningen; Achtergronddocument bij Natuurverkenning 2011
- 221** *Knegt, B. de, J. Clement, P.W. Goedhart, H. Sierdsema, Chr. van Swaay & P. Wiersma.* Natuurkwaliteit van het agrarisch gebied

- 222** *Kamphorst, D.A. & M.M.P. van Oorschot.* Kansen en barrières voor verduurzaming van houtketens
- 223** *Salm, C. van der & O.F. Schoumans.* Langetermijneffecten van verminderde fosfaatgiften
- 224** *Bikker, P., M.M. van Krimpen & G.J. Rummelink.* Stikstofverteerbaarheid in voeders voor landbouwhuisdieren; Berekeningen voor de TAN-excretie
- 225** *M.E. Sanders & A.L. Gerritsen (red.).* Het biodiversiteitsbeleid in Nederland werkt. Achtergronddocument bij Balans van de Leefomgeving 2010
- 226** *Bogaart, P.W., G.A.K. van Voorn & L.M.W. Akkermans.* Evenwichtsanalyse modelcomplexiteit; een verkennende studie
- 227** *Kleunen A. van, K. Koffijberg, P. de Boer, J. Nienhuis, C.J. Camphuysen, H. Schekkerman, K.H. Oosterbeek, M.L. de Jong, B. Ens & C.J. Smit (2010).* Broedsucces van kustbroedvogels in de Waddenzee in 2007 en 2008
- 228** *Salm, C. van der, L.J.M. Boumans, D.J. Brus, B. Kempen & T.C. van Leeuwen.* Validatie van het nutriëntenemissiemodel STONE met meetgegevens uit het Landelijk Meetnet effecten Mestbeleid (LMM) en de Landelijke Steekproef Kaartenheden (LSK).
- 229** *Dijkema, K.S., W.E. van Duin, E.M. Dijkman, A. Nicolai, H. Jongerius, H. Keegstra, L. van Egmond, H.J. Venema & J.J. Jongsma.* Vijftig jaar monitoring en beheer van de Friese en Groninger kwelderwerken: 1960-2009
- 230** *Jaarrapportage 2010.* WOT-04-001 – Koepel
- 231** *Jaarrapportage 2010.* WOT-04-002 – Onderbouwend Onderzoek
- 232** *Jaarrapportage 2010.* WOT-04-003 – Advisering Natuur & Milieu
- 233** *Jaarrapportage 2010.* WOT-04-005 – M-AVP
- 234** *Jaarrapportage 2010.* WOT-04-006 – Natuurplanbureaufunctie
- 235** *Jaarrapportage 2010.* WOT-04-007 – Milieuplanbureaufunctie
- 236** *Arnouts, R.C.M. & F.H. Kistenkas.* Nederland op slot door Natura 2000: de discussie ontrafeld; Bijlage bij WOT-paper 7 – De deur klemt
- 237** *Harms, B. & M.M.M. Overbeek.* Bedrijven aan de slag met natuur en landschap; relaties tussen bedrijven en natuurorganisaties. Achtergronddocument bij Natuurverkenning 2011
- 238** *Agricola, H.J. & L.A.E. Vullings.* De stand van het platteland 2010. Monitor Agenda Vitaal Platteland; Rapportage Midterm meting Effectindicatoren
- 239** *Klijn, J.A.* Wisselend getij. Omgang met en beleid voor natuur en landschap in verleden en heden; een essayistische beschouwing. Achtergronddocument bij Natuurverkenning 2011
- 240** *Corporaal, A., T. Denters, H.F. van Dobben, S.M. Hennekens, A. Klimkowska, W.A. Ozinga, J.H.J. Schaminée & R.A.M. Schrijver.* Stenoeciteit van de Nederlandse flora. Een nieuwe parameter op grond van ecologische amplitudo's van de Nederlandse plantensoorten en toepassingsmogelijkheden
- 241** *Wamelink, G.W.W., R. Jochem, J. van der Graft-van Rossum, C. Grashof-Bokdam, R.M.A. Wegman, G.J. Franke & A.H. Prins.* Het plantendispersiemodel DIMO. Verbetering van de modellering in de Natuurplanner
- 242** *Klimkowska, A., M.H.C. van Adrichem, J.A.M. Jansen & G.W.W. Wamelink.* Bruikbaarheid van WNK-monitoringgegevens voor EC-rapportage voor Natura 2000-gebieden. Eerste fase
- 243** *Goossen, C.M., R.J. Fontein, J.L.M. Donders & R.C.M. Arnouts.* Mass Movement naar recreatieve gebieden; Overzicht van methoden om bezoekersaantallen te meten
- 244** *Spruijt, J., P.M. Spoorenberg, J.A.J.M. Rovers, J.J. Slabbekoorn, S.A.M. de Kool, M.E.T. Vlaswinkel, B. Heijne, J.A. Hiemstra, F. Nouwens & B.J. van der Sluis.* Milieueffecten van maatregelen gewasbescherming
- 245** *Walker, A.N. & G.B. Woltjer.* Forestry in the Magnet model.
- 246** *Hoefnagel, E.W.J., F.C. Buisman, J.A.E. van Oostenbrugge & B.I. de Vos.* Een duurzame toekomst voor de Nederlandse visserij. Toekomstscenario's 2040
- 247** *Buurma, J.S. & S.R.M. Janssens.* Het koor van adviseurs verdient een dirigent. Over kennisverspreiding rond phytophthora in aardappelen
- 248** *Verburg, R.W., A.L. Gerritsen & W. Nieuwenhuizen.* Natuur meekoppelen in ruimtelijke ontwikkeling: een analyse van sturingsstrategieën voor de Natuurverkenning.
- Achtergronddocument bij Natuurverkenning 2011
- 249** *Kooten, T. van & C. Klok.* The Mackinson-Daskalov North Sea EcoSpace model as a simulation tool for spatial planning scenarios
- 250** *Bruggen van, C., C.M. Groenestein, B.J. de Haan, M.W. Hoogeveen, J.F.M. Huijsmans, S.M. van der Sluis & G.L. Velthof.* Ammoniakemissie uit dierlijke mest en kunstmest 1990-2008. Berekeningen met het Nationaal Emissiemodel voor Ammoniak (NEMA)
- 251** *Bruggen van, C., C.M. Groenestein, B.J. de Haan, M.W. Hoogeveen, J.F.M. Huijsmans, S.M. van der Sluis & G.L. Velthof.* Ammoniakemissie uit dierlijke mest en kunstmest in 2009. Berekeningen met het Nationaal Emissiemodel voor Ammoniak (NEMA)
- 252** *Randen van, Y., H.L.E. de Groot & L.A.E. Vullings.* Monitor Agenda Vitaal Platteland vastgelegd. Ontwerp en implementatie van een generieke beleidsmonitor
- 253** *Agricola, H.J., R. Reijnen, J.A. Boone, M.A. Dolman, C.M. Goossen, S. de Vries, J. Roos-Klein Lankhorst, L.M.G. Groenemeijer & S.L. Deijl.* Achtergronddocument Midterm meting Effectindicatoren Monitor Agenda Vitaal Platteland
- 254** *Buiteveld, J. S.J. Hiemstra & B. ten Brink.* Modelling global agrobiodiversity. A fuzzy cognitive mapping approach
- 255** *Hal van R., O.G. Bos & R.G. Jak.* Noordzee: systeemdynamiek, klimaatverandering, natuurtypen en benthos. Achtergronddocument bij Natuurverkenning 2011
- 256** *Teal, L.R.* The North Sea fish community: past, present and future. Background document for the 2011 National Nature Outlook
- 257** *Leopold, M.F., R.S.A. van Bemmelen & S.C.V. Geelhoed.* Zeevogels op de Noordzee. Achtergronddocument bij Natuurverkenning 2011
- 258** *Geelhoed, S.C.V. & T. van Polanen Petel.* Zeezoogdieren op de Noordzee. Achtergronddocument bij Natuurverkenning 2011
- 259** *Kuijs, E.K.M. & J. Steenbergen.* Zoet-zoutovergangen in Nederland; stand van zaken en kansen voor de toekomst. Achtergronddocument bij Natuurverkenning 2011
- 260** *Baptist, M.J.* Zachte kustverdediging in Nederland; scenario's voor 2040. Achtergronddocument bij Natuurverkenning 2011
- 261** *Wiersinga, W.A., R. van Hal, R.G. Jak & F.J. Quirjns.* Duurzame kottervisserij op de Noordzee. Achtergronddocument bij Natuurverkenning 2011
- 262** *Wal J.T. van der & W.A. Wiersinga.* Ruimtegebruik op de Noordzee en de trends tot 2040. Achtergronddocument bij Natuurverkenning 2011
- 263** *Wiersinga, W.A. J.T. van der Wal, R.G. Jak & M.J. Baptist.* Vier kijkrichtingen voor de mariene natuur in 2040. Achtergronddocument bij Natuurverkenning 2011
- 264** *Bolman, B.C. & D.G. Goldsborough.* Marine Governance. Achtergronddocument bij Natuurverkenning 2011
- 265** *Bannink, A.* Methane emissions from enteric fermentation in dairy cows, 1990-2008; Background document on the calculation method and uncertainty analysis for the Dutch National Inventory Report on Greenhouse Gas Emissions
- 266** *Wyngaert, I.J.J. van den, P.J. Kuikman, J.P. Lesschen, C.C. Verwer & H.H.J. Vreuls.* LULUCF values under the Kyoto Protocol; Background document in preparation of the National Inventory Report 2011 (reporting year 2009)
- 267** *Helming, J.F.M. & I.J. Terluin.* Scenarios for a cap beyond 2013; implications for EU27 agriculture and the cap budget.
- 268** *Woltjer, G.B.* Meat consumption, production and land use. Model implementation and scenarios.
- 269** *Knegt, B. de, M. van Eupen, A. van Hinsberg, R. Pouwels, M.S.J.M. Reijnen, S. de Vries, W.G.M. van der Bilt & S. van Tol.* Ecologische en recreatieve beoordeling van toekomstscenario's van natuur op het land. Achtergronddocument bij Natuurverkenning 2011.
- 270** *Bos, J.F.F.P., M.J.W. Smits, R.A.M. Schrijver & R.W. van der Meer.* Gebiedsstudies naar effecten van vergroening van het Gemeenschappelijk Landbouwbeleid op bedrijfseconomie en inpassing van agrarisch natuurbeheer.
- 271** *Donders, J., J. Luttkik, M. Goossen, F. Veeneklaas, J. Vreke & T. Weijsschede.* Waar gaat dat heen? Recreatiemotieven, landschapskwaliteit en de oudere wandelaar. Achtergronddocument bij Natuurverkenning 2011.
- 272** *Voorn G.A.K. van & D.J.J. Walvoort.* Evaluation of an evaluation list for model complexity.

- 273** *Heide, C.M. van der & F.J. Sijtsma.* Maatschappelijke waardering van ecosysteemdiensten; een handreiking voor publieke besluitvorming. Achtergronddocument bij Natuurverkenning 2011
- 274** *Overbeek, M.M.M., B. Harms & S.W.K. van den Burg (2012).* Internationale bedrijven duurzaam aan de slag met natuur en biodiversiteit.; voorstudie bij de Balans van de Leefomgeving 2012.
- 275** *Os, J. van; T.J.A. Gies; H.S.D. Naeff; L.J.J. Jeurissen.* Emissieregistratie van landbouwbedrijven; verbeteringen met behulp van het Geografisch Informatiesysteem Agrarische Bedrijven.
- 276** *Walsum, P.E.V. van & A.A. Veldhuizen.* MetaSWAP_V7_2_0; Rapportage van activiteiten ten behoeve van certificering met Status A.
- 277** *Kooten T. van & S.T. Glorius.* Modeling the future of het North Sea. An evaluation of quantitative tools available to explore policy, space use and planning options.
- 279** *Bilt, W.G.M. van der, B. de Knegt, A. van Hinsberg & J. Clement (2012).* Van visie tot kaartbeeld; de kijkrichtingen ruimtelijk uitgewerkt. Achtergronddocument bij Natuurverkenning 2011
- 280** *Kistenkas, F.H. & W. Nieuwenhuizen.* Rechtsontwikkelingen landschapsbeleid: landschapsrecht in wording. Bijlage bij WOT-papier 12 – 'Recht versus beleid'
- 281** *Meeuwssen, H.A.M. & R. Jochem.* Openheid van het landschap; Berekeningen met het model ViewScape.
- 282** *Dobben, H.F. van.* Naar eenvoudige dosis-effectrelaties tussen natuur en milieucondities; een toetsing van de mogelijkheden van de Natuurplanner.
- 283** *Gaaff, A.* Raming van de budgetten voor natuur op langere termijn; Achtergronddocument bij Natuurverkenning 2011.
- 285** *Vries, P. de, J.E. Tamis, J.T. van der Wal, R.G. Jak, D.M.E. Slijkerman and J.H.M. Schobben.* Scaling human-induced pressures to population level impacts in the marine environment; implementation of the prototype CUMULEO-RAM model.
- 2012**
- 286** *Keizer-Vlek, H.E. & P.F.M. Verdonschot.* Bruikbaarheid van SNL-monitoringgegevens voor EC-rapportage voor Natura 2000-gebieden; Tweede fase: aquatische habitattypen.
- 287** *Oenema, J., H.F.M. Aarts, D.W. Bussink, R.H.E.M. Geerts, J.C. van Middelkoop, J. van Middelaar, J.W. Reijs & O. Oenema.* Variatie in fosfaatopbrengst van grasland op praktijkbedrijven en mogelijke implicaties voor fosfaatgebruiksnormen.
- 288** *Troost, K., D. van de Ende, M. Tangelder & T.J.W. Ysebaert.* Biodiversity in a changing Oosterschelde: from past to present
- 289** *Jaarrapportage 2011.* WOT-04-001 – Koepel
- 290** *Jaarrapportage 2011.* WOT-04-008 – Agromilieue
- 291** *Jaarrapportage 2011.* WOT-04-009 – Natuur, Landschap en Platteland
- 292** *Jaarrapportage 2011.* WOT-04-010 – Balans van de Leefomgeving
- 293** *Jaarrapportage 2011.* WOT-04-011 – Natuurverkenning
- 294** *Bruggen, C. van, C.M. Groenestein, B.J. de Haan, M.W. Hoogeveen, J.F.M. Huijsmans, S.M. van der Sluis & G.L. Velthof.* Ammoniakemissie uit dierlijke mest en kunstmest in 2010; berekeningen met het Nationaal Emissiemodel voor Ammoniak (NEMA).
- 295** *Spijker, J.H., H. Kramer, J.J. de Jong & B.G. Heusinkveld.* Verkenning van de rol van (openbaar) groen op wijk- en buurtniveau op het hitte-eilandeffect
- 296** *Haas, W. de, C.B.E.M. Aalbers, J. Kruit, R.C.M. Arnouts & J. Kempenaar.* Parknatuur; over de kijkrichtingen beleefbare natuur en inpasbare natuur
- 297** *Doorn, A.M. van & R.A. Smidt.* Staltypen nabij Natura 2000-gebieden.
- 298** *Luesink, H.H., A. Schouten, P.W. Blokland & M.W. Hoogeveen.* Ruimtelijke verdeling ammoniakemissies van beweiden en van aanwenden van mest uit de landbouw.
- 299** *Meulenkamp, W.J.H. & T.J.A. Gies.* Effect maatregelen reconstructie zandgebieden; pilotgemeente Gemert-Bakel.
- 300** *Beukers, R. & B. Harms.* Meerwaarde van certificeringsschema's in visserij en aquacultuur om bij te dragen aan het behoud van biodiversiteit
- 301** *Broekmeyer, M.E.A., H.P.J. Huiskens, S.M. Hennekens, A. de Jong, M.H. Storm & B. Vanmeulebrouk.* Gebruikers-handleiding Auditrail Natura 2000.
- 302** *Bruggen van, C., C.M. Groenestein, B.J. de Haan, M.W. Hoogeveen, J.F.M. Huijsmans, S.M. van der Sluis & G.L. Velthof.* Ammonia emissions from animal manure and inorganic fertilisers in 2009. Calculated with the Dutch National Emissions Model for Ammonia (NEMA)
- 303** *Donders, J.L.M. & C.M. Goossen.* *Recreatie in groen blauwe gebieden.* Analyse data Continu Vrijtijdsonderzoek: bezoek, leeftijd, stedelijkheidsgraad en activiteiten van recreanten
- 304** *Boesten, J.J.T.I. & M.M.S. ter Horst.* Manual of PEARLNEQ v5
- 305** *Reijnen, M.J.S.M., R. Pouwels, J. Clement, M. van Esbroek, A. van Hinsberg, H. Kuipers & M. van Eupen.* EHS Doelrealisatiegraadmeter voor de Ecologische Hoofdstructuur. Natuurkwaliteit van landecosysteemtypen op lokale schaal.
- 306** *Arnouts, R.C.M., D.A. Kamphorst, B.J.M. Arts & J.P.M. van Tatenhove.* Innovatieve governance voor het groene domein. Governance-arrangementen voor vermaatschappelijking van het natuurbeleid en verduurzaming van de koffieketen.
- 307** *Kruseman, G., H. Luesink, P.W. Blokland, M. Hoogeveen & T. de Koeijer.* MAMBO 2.x. Design principles, model, structure and data use
- 308** *Koeijer de, T., G. Kruseman, P.W. Blokland, M. Hoogeveen & H. Luesink.* MAMBO: visie en strategisch plan, 2012-2015
- 309** *Verburg, R.W.* Methoden om kennis voor integrale beleidsanalyses te combineren.
- 310** *Bouwma, I.M., W.A. Ozinga, T. v.d. Sluis, A. Griffioen, M.P. v.d. Veen & B. de Knegt.* Dutch nature conservation objectives from a European perspective.
- 311** *Wamelink, G.W.W., M.H.C. van Adrichem & P.W. Goedhart.* Validatie van MOVE4.
- 312** *Broekmeyer, M.E.A., M.E. Sanders & H.P.J. Huiskens.* Programmatische Aanpak Stikstof. Doelstelling, maatregelen en mogelijke effectiviteit.
- 313**
- 314** *Pouwels, P. C. van Swaay, R. Foppen & H. Kuipers.* Prioritaire gebieden binnen de Ecologische Hoofdstructuur voor behoud doelsoorten vlinders en vogels.
- 315** *Rudrum, D., J. Verboom, G. Kruseman, H. Leneman, R. Pouwels, A. van Teeffelen & J. Clement.* Kosteneffectiviteit van natuurgebieden op het land. Eerste verkenning met ruimtelijke optimalisatie biodiversiteit.
- 316** *Boone, J.A., M.A. Dolman, G.D. Jukema, H.R.J. van Kernebeek & A. van der Knijff.* Duurzame landbouw verantwoord. Methodologie om de duurzaamheid van de Nederlandse landbouw kwantitatief te meten.
- 317** *Troost, K., M. Tangelder, D. van den Ende & T.J.W. Ysebaert.* From past to present: biodiversity in a changing delta
- 318** *Schouten, A.D., H. Leneman, R. Michels & R.W. Verburg.* Instrumentarium kosten natuurbeleid. Status A.
- 319** *Verburg, R.W., E.J.G.M. Westerhof, M.J. Bogaardt & T. Selnes.* Verkennen en toepassen van besluitvormingsmodellen in de uitvoering van natuurbeleid.
- 2013**
- 320** *Woltjer, G.B.* Forestry in MAGNET; a new approach for land use and forestry modelling.
- 321** *Langers, F., A.E. Buijs, S. de Vries, J.M.J. Farjon, A. van Hinsberg, P. van Kampen, R. van Marwijk, F.J. Sijtsma, S. van Tol.* Potenties van de Hotspotmonitor om de graadmeter Landschap te verfijnen
- 322** *Verburg, R.W., M.J. Bogaardt, B. Harms, T. Selnes, W.J. Oliemans.* Beleid voor ecosysteemdiensten. Een vergelijking tussen verschillende EU-staten
- 323** *Schouten, M.A.H., N.B.P. Polman & E.J.G.M. Westerhof.* Exploring green agricultural policy scenarios with a spatially explicit agent-based model.
- 324** *Gerritsen, A.L., A.M.E. Groot, H.J. Agricola, W. Nieuwenhuizen.* Hoogproductieve landbouw. Een verkenning van motivaties, knelpunten, condities, nieuwe organisatievormen en de te verwachten bijdragen aan natuur en landschap
- 325** *Jaarrapportage 2012.* WOT-04-008 – Agromilieue
- 326** *Jaarrapportage 2012.* WOT-04-009 – Informatievoorziening Natuur (IN)
- 327** *Jaarrapportage 2012.* WOT-04-010 – Balans van de Leefomgeving (BvdL)
- 328** *Jaarrapportage 2012.* WOT-04-011 – Natuurverkenning (NVK)
- 329** *Goossen, C.M., F. Langers, T.A. de Boer.* Relaties tussen recreanten, ondernemers en landschap
- 330** *Bruggen, C. van, P. Bikker, C.M. Groenestein, B.J. de Haan,*

- M.W. Hoogeveen, J.F.M. Huijsmans, S.M. van der Sluis & G.L. Velthof.* Ammoniakemissie uit dierlijke mest en kunstmest in 2011. Berekeningen met het Nationaal Emissiemodel voor Ammoniak (NEMA).
- 331** *Dirkx, G.H.P. & W. Nieuwenhuizen.* Histland. Historisch-landschappelijk informatiesysteem
- 332** *Ehlert, P.A.I., T.A. van Dijk & O. Oenema.* Opname van struviet als categorie in het Uitvoeringsbesluit Meststoffenwet. Advies.
- 333** *Ehlert, P.A.I., H.J. van Wijnen, J. Struijs, T.A. van Dijk, L. van Schöll, L.R.M. de Poorter.* Risicobeoordeling van contaminanten in afval- en reststoffen bestemd voor gebruik als covergistingmateriaal
- 334** *Verdonschot R.C.M., J.H. Vos J.H. & P.F.M. Verdonschot.* Exotische macrofauna en macrofyten in de Nederlandse zoete wateren; voorkomen en beleid in 2012.
- 335** *Commissie Deskundigen Meststoffenwet.* Protocol beoordeling stoffen Meststoffenwet. Versie 3.1
- 336** *Ehlert, P.A.I., L. Posthuma, P.F.A.M. Römken, R.P.J.J. Rietra, A.M. Wintersen, H. van Wijnen, T.A. van Dijk, L. van Schöll, J.E. Groenenberg.* Appraising fertilisers: Origins of current regulations and standards for contaminants in fertilisers. Background of quality standards in the Netherlands, Denmark, Germany, United Kingdom and Flanders
- 337** *Greft-van Rossum, J.G.M. van der, M.J.S.M. Reijnen, W.A. Ozinga, R. Pouwels, M. van Eupen, A.M.G. de Bruijn, H. Kuipers, S.M. Hennekens & A.H. Malinowska.* Water-, milieu- en ruimtecondities vaatplanten; Implementatie in Model for Nature Policy MNP 2.0.
- 338** *Vos, C.C., R. Pouwels, M. van Eupen, H.T. Lemaris Meeuwsen, W.A. Ozinga, M. Sterk, M. Wallis de Vries.* Operationalisering van het begrip 'veerkracht van ecosystemen'. Een empirische verkenning voor planten en dagvlinders.
- 339** *Voorn van, G.A.K., P.W. Bogaart, M. Knotters, D.J.J. Walvoort.* Complexiteit van WUR-modellen en -bestanden; Toetsing van de EMC v1.0
- 341** *Knegt de, B., J.G.M. van der Greft-van Rossum, S.M. Hennekens, G.B.M. Heuvelink.* Trends van zeldzame plantensoorten voorspelt.



Thema Natuurverkenning

Wettelijke Onderzoekstaken

Natuur & Milieu

Postbus 47

6700 AA Wageningen

T (0317) 48 54 71

E info.wnm@wur.nl

www.wageningenUR.nl/

wotnatuurenmilieu

De WOT Natuur & Milieu voert wettelijke onderzoekstaken uit op het beleidsterrein natuur en milieu. Deze taken worden uitgevoerd om een wettelijke verantwoordelijkheid van de minister van Economische Zaken te ondersteunen. De WOT Natuur & Milieu werkt aan producten van het Planbureau voor de Leefomgeving, zoals de Balans van de Leefomgeving en de Natuurverkenning. Verder brengen we voor het ministerie van Economische Zaken adviezen uit over (toelating van) meststoffen en bestrijdingsmiddelen, en zorgen we voor informatie voor Europese rapportageverplichtingen over biodiversiteit.

De WOT Natuur & Milieu is onderdeel van de internationale kennisorganisatie Wageningen UR (University & Research centre). De missie is 'To explore the potential of nature to improve the quality of life'. Binnen Wageningen UR bundelen 9 gespecialiseerde onderzoeksinstituten van stichting DLO en Wageningen University hun krachten om bij te dragen aan de oplossing van belangrijke vragen in het domein van gezonde voeding en leefomgeving. Met ongeveer 30 vestigingen, 6.000 medewerkers en 9.000 studenten behoort Wageningen UR wereldwijd tot de aansprekende kennisinstellingen binnen haar domein. De integrale benadering van de vraagstukken en de samenwerking tussen verschillende disciplines vormen het hart van de unieke Wageningen aanpak.

