

Rapport 550013001/2005

Metamodelleren bij het MNP-RIVM
Naar praktische toepassingen

P.H.M. Janssen, P.S.C. Heuberger, A. Tiktak

Contact:
Peter Janssen
MNP - IMP
Peter.Janssen@mnp.nl

Dit onderzoek werd verricht in opdracht van de Directeur-Generaal RIVM, in het kader van project S/550013: Methoden voor Modelleren.

MNP, Postbus 303, 3720 AH Bilthoven, telefoon: 030 - 274 274 5; fax: 030 - 274 44 79

Het rapport in het kort

Metamodelleren bij het MNP-RIVM: naar praktische toepassingen

Metamodellering is het benaderen van een ingewikkeld simulatiemodel door een eenvoudiger en sneller model. Voor onderzoekers vormt dit een belangrijk hulpmiddel om modellen te maken die, bijvoorbeeld via koppeling, integratie en opschaling, op brede schaal inzetbaar zijn voor integraal beleidsgericht onderzoek. Om hierbij verantwoorde resultaten te garanderen is het van belang dat dit proces vanuit methodologisch oogpunt met zorg gebeurt en ondersteund wordt door adequate technieken. Daartoe wordt een stappenplan voor metamodellering gepresenteerd en wordt een breed scala van methoden besproken die gebruikt kunnen worden bij de uitvoering hiervan. Praktijkervaringen met metamodellering binnen het MNP-RIVM en literatuuronderzoek laten zien dat aanpak en resultaten sterk toepassingsafhankelijk kunnen zijn. Belangrijke toekomstige uitdagingen zijn: het duidelijk aangeven en bewaken van kwaliteit, toepassingsbereik en transparantie van metamodellen; het breder gebruik van reeds beschikbare, meer geavanceerde technieken uit statistiek en systeemtheorie; metamodellering in nieuwe contexten (gekoppelde modellen; dynamiek in plaats en tijd; combinatie met expertkennis; optimalisatie; onzekerheidsanalyses).

Trefwoorden: simulatie; meta-model; modelreductie; onzekerheden; stappenplan

Abstract

Metamodelling at the MNP-RIVM: towards practical applications

Metamodelling concerns the approximation of a complex simulation model by a simpler and faster model. It is an important tool for researchers to obtain models which, e.g. by coupling, integration and upscaling, are widely applicable in policy-oriented research. To guarantee reliable results it is important to perform this process in a methodologically sound and careful way, supported by adequate techniques. For this reason a step-wise strategy for metamodelling is presented and a broad spectrum of methods for metamodel construction is discussed. Experiences with metamodelling within the MNP-RIVM and findings from literature illustrate that the approach and results can be very application-dependent. Several important future challenges have been identified for the MNP-RIVM: to explicitly document and assure the application area, quality and transparency of metamodelling; the use of already available, more advanced methods from statistics and system theory; metamodelling in new contexts (coupled models; spatial and temporal dynamics; links with expert knowledge, optimization, uncertainty analyses).

Keywords: simulation; meta-model; model reduction; uncertainties; step-wise strategy

Samenvatting

Simulatiemodellen spelen een belangrijke rol bij onderzoek van complexe beleidsproblemen rond milieu-, natuur- en duurzaamheidsvraagstukken. Deze modellen zijn echter niet zelden te gedetailleerd, reken- en data-intensief en vaak dienen ze vervangen te worden door een eenvoudiger model, een metamodel, om - gekoppeld, geïntegreerd of opgeschaald - op brede schaal inzetbaar te zijn. Om hierbij verantwoorde resultaten te garanderen is het van belang dat metamodellering vanuit methodologisch oogpunt met zorg gebeurt en ondersteund wordt door adequate technieken.

Daartoe is een stappenplan opgesteld, waarin expliciet doel en context van metamodellering aan bod komen, evenals de keuze van inputs/outputs, het genereren van data voor metamodelbouw, de keuze van metamodelleringstechniek(en), de evaluatie van het metamodel en het gebruik en beheer van het metamodel. Bij het bespreken van de technieken zijn een drietal hoofdklassen onderscheiden, namelijk (a) procesgebaseerde, (b) statistische en (c) systeemtheoretische modelreductie methoden en is een globale vergelijking van deze methoden gegeven, inclusief een bespreking van praktische toepassingsaspecten.

De huidige status van metamodellering bij het MNP-RIVM is in beeld gebracht door een aantal concrete praktijkstudies uit het afgelopen decennium te bespreken, waarin het gebruik van metamodelen een rol heeft gespeeld. Deze studies lagen op het gebied van nutriëntenmodellering, atmosferisch transport van verzurende stoffen, blootstelling aan ozon, klimaatmodellering, waterkwaliteitsmodellering en uitspoeling van bestrijdingsmiddelen. De daarbij gebruikte metamodelleringstechnieken variëren in mate van ingewikkeldheid, en het is opvallend dat er weinig aandacht was voor niet-lineariteiten en dat doorgaans dynamische aspecten ontbreken.

Op basis van literatuur en de besproken MNP-RIVM-ervaringen wordt geconcludeerd dat aanpak en resultaten van metamodellering sterk toepassingsafhankelijk kunnen zijn. Belangrijke toekomstige uitdagingen liggen er bij het duidelijk aangeven en bewaken van kwaliteit, toepassingsbereik en transparantie van metamodelen, bij het gebruik van reeds beschikbare, meer geavanceerde technieken uit statistiek en systeemtheorie, en bij metamodellering in nieuwe contexten (gekoppelde modellen; dynamiek in plaats en tijd; combinatie met expertkennis; optimalisatie; onzekerheidsanalyses). Met name zal hierbij centraal staan hoe de ontwikkeling van expertise op deze gebieden rechtstreeks te koppelen valt aan concrete en beleidsrelevante toepassingen.

INHOUD

1. INLEIDING.....	9
2. METAMODELLEREN: EEN STAPPENPLAN	11
3. METAMODELLEREN: EEN OVERZICHT VAN METHODEN.....	17
3.1. PROCES-GEbaseERDE METAMODELLERING.....	17
3.2. STATISTIEK-GEbaseERDE METAMODELLERING.....	17
3.3. SYSTEEMTHEORIE-GEbaseERDE METAMODELLERING.....	19
3.4. VERGELIJKING VAN METAMODELLERINGSMETHODEN.....	20
3.5. TOEPASSINGSASPECTEN	22
4. METAMODELLEREN BIJ HET MNP: ENKELE ERVARINGEN EN LEERPUNTEN..	29
4.1 METAMODELLERING VOOR STONE	29
4.2 METAMODELLERING VOOR ATMOSFERISCH MODEL OPS	34
4.3 METAMODEL VOOR ATMOSFERISCH MODEL EUROS	35
4.4 METAMODELLERING VOOR IMAGE.....	37
4.5 METAMODELLERING VOOR PCLAKE.....	40
4.6 METAMODELLEN VAN GEOPEARL.....	42
4.7 SLOTOPMERKINGEN	44
5. METAMODELLEREN: CONCLUSIES EN VOORUITZICHTEN	47
DANKWOORD	49
LITERATUUR	51
APPENDIX METHODEN VOOR MODELREDUCTIE	61
A.1. PROCES-GEbaseERDE MODELREDUCTIE.....	61
A.1.1. IRF: Impuls Responsie Functies	61
A.1.2. DBM – Data Based Mechanistic Modelling	62
A.1.3. Analytische approximaties.....	64
A.2. STATISTIEK-GEbaseERDE MODELREDUCTIE.....	65
A.2.1. Interpolatie in Look-Up Tables.....	66
A.2.2. Response Surfaces (polynoom regressie).....	66
A.2.3. Het Kriging model/ Gaussische Processen.....	68
A.2.4. GLMs en GAMs (Generalized Linear Models en Generalized Additive Models).....	69
A.2.5. CART (Classification And Regression Trees).....	71
A.2.6. MARS (Multivariate Adaptive Regression Splines)	72
A.2.7. Radial Basis Functions (RBF)	73
A.2.8. Neurale Netwerken	73
A.2.9. Support Vector Machines en Kernel-based learning.....	75
A.2.10. Fuzzy en Neuro-fuzzy methoden (inductive learning en reasoning)	76
A.2.11. Genetisch Evolutionaire Algoritmen.....	77
A.3. SYSTEEMTHEORETISCH GEbaseERDE MODELREDUCTIE	78
A.3.1. Lineaire Systemen.....	78
A.3.1.a. Gebalanceerde modelreductie	79
A.3.1.b. Hankel norm reductie	80
A.3.1.c. Singuliere Verstoringen (singular perturbations).....	81
A.3.1.d. Modal Truncation	82
A.3.1.e. Moment Matching/partiële realisatie/ Padé benadering/ rationale interpolatie	82
A.3.1.f. Systeemidentificatie.....	83
A.3.2. Niet-Lineaire Systemen.....	83
A.3.2.a. Niet-Lineair gebalanceerde reductie	83
A.3.2.b. Proper Orthogonal Decomposition (POD).....	84
A.3.2.c. Stuksgewijze Linearisatie.....	85
A.3.2.d. Niet-Lineaire Systeemidentificatie	85
A.3.2.e. Varia	86

1. Inleiding

Simulatiemodellen zijn tegenwoordig niet meer weg te denken bij onderzoek van complexe beleidsproblemen rond milieu-, natuur- en duurzaamheidsvraagstukken. Kennis over relevante mechanismen en processen en over belangrijke oorzaak-gevolg relaties is hierbij in computermodellen vastgelegd. Via berekeningen/simulaties met deze modellen kan inzicht in de problematiek en in mogelijke ‘oplossingen’ verschaft worden.

Bij het bouwen van deze simulatiemodellen is vaak veel specialistische proceskennis en informatie gebruikt. Niet zelden is daardoor het resulterend simulatiemodel te ingewikkeld, gedetailleerd en ondoorzichtig geworden voor directe (beleids)toepassing op de gewenste schaal. Ook kan de rekentijd en ‘datahonger’ (bijvoorbeeld wat betreft parameterwaarden en andere invoergegevens) van een model excessief groot geworden zijn, zodat het - ondanks ontwikkelingen op het vlak van sneller rekenen (grid-technologie, snellere processoren) - wenselijk is om het oorspronkelijke simulatiemodel te vervangen door een eenvoudiger model, een zogenaamd *metamodel*, ook wel model-proxy, surrogaat-model, gereduceerd model of model-emulator genoemd.

Zo’n metamodel kan bijvoorbeeld worden ingezet om op grotere tijd- en ruimteschaal (van plot-schaal, naar regionaal, landelijk, Europees etcetera) te rekenen, een groot aantal scenario’s te evalueren, onzekerheidsanalyses en grootschalige optimalisatie berekeningen mogelijk te maken. Hiermee wordt de praktische toepassingsschaal van het oorspronkelijke model sterk vergroot. Ook kan metamodellering worden gebruikt om bestaande kennis – die vaak vervat is in modellen en modelcomponenten op specifieke deeldomeinen – efficiënt in te zetten, te integreren en op te schalen ten behoeve van integrale (duurzaamheids)assessments en evaluaties. Zo kunnen metamodellen van specifieke procesmodellen nuttige bouwstenen aanleveren voor het bouwen van integrated assessment modellen (zie Rotmans en Van Asselt, 2001) of voor het bouwen van scanners, dat wil zeggen hoog geaggregeerde modelsystemen waarmee relevante indicatoren kunnen worden berekend ten behoeve van (semi-)kwantitatieve onderbouwing van (interactieve) beleidsadviezen.

Belangrijk uitgangspunt bij het opstellen van een metamodel voor een specifieke toepassing is dat het oorspronkelijke simulatiemodel (het referentiemodel) qua inhoud en kwaliteit in principe ook geschikt zou moeten zijn voor het specifieke beoogde doel, maar dat dit door praktische bezwaren (rekentijd, omvang, databehoeft) geen reële optie is. Voor een verantwoorde ontwikkeling en gebruik van een metamodel is het dus nodig om de gewenste gebruikscondities, toepassingsrange en kwaliteit van het metamodel (en het referentiemodel) helder en concreet te omschrijven. Ook het toetsen of deze wensen daadwerkelijk gerealiseerd worden bij de betreffende toepassing is een onlosmakelijk onderdeel van metamodellering. Speciaal aandachtspunt hierbij is het toepassingsbereik bij koppeling/integratie van diverse metamodellen.

Om dit proces methodologisch te ondersteunen is een stappenplan voor metamodellering opgesteld en is in kaart gebracht welke specifieke methoden nuttig gebruikt kunnen worden bij de uitvoering hiervan. In hoofdstuk 2 beschrijven we de hoofdlijnen van het stappenplan, waarna in hoofdstuk 3 (en appendix A) informatie wordt gegeven over methoden voor metamodellering en over praktische aspecten die bij de toepassing van deze technieken aan de orde zijn. In hoofdstuk 4 bespreken we kort een aantal metamodelleringsstudies die in het verleden binnen het MNP-RIVM hebben plaatsgevonden; we sluiten in hoofdstuk 5 af met algemene conclusies en suggesties voor toekomstig onderzoek en gebruik van metamodellering binnen het MNP-RIVM.

2. Metamodelleren: een stappenplan

In het hiernavolgende worden de hoofdlijnen gepresenteerd van een stappenplan voor het construeren van een metamodel. De precieze invulling is afhankelijk van een groot aantal factoren (onder andere doel/gebruik metamodel; beschikbaarheid van budget, tijd, gegevens, expertise, rekenkracht, software etcetera) die door het stappenplan expliciet aan de orde worden gesteld. Het stappenplan is geen strak keurslijf, maar vestigt de aandacht op een aantal punten die van belang zijn bij het proces van metamodellering. Ook wordt verwezen naar het volgende hoofdstuk waarin een aantal methoden voor het uitvoeren van metamodellering kort besproken wordt, en waar nader op een aantal praktische aspecten wordt ingegaan.

Stap 0: Begripsdefinitie: wat verstaan we onder metamodelleren?

We gebruiken het begrip metamodel in de volgende betekenis:

- een vereenvoudigde (minder complexe, snellere) versie van een zogenaamd moedermodel, of referentiemodel
- een model dat op een ruimere schaal (spatueel, temporeel) werkt dan een moedermodel, vaak op basis van geaggregeerde invoerdata en met geaggregeerde uitvoerdata als resultaat. Denk bijvoorbeeld aan een situatie waarbij het metamodel gebruikt zal worden voor geheel Nederland in plaats van een groot aantal modelruns met het moedermodel voor specifieke locaties.

Stap 1: Probleemdefinitie en –afbakening: waarom metamodelleren?

A. Doel

Wat is de beoogde toepassingssituatie en wat is het beoogde doel van de metamodellering, en waarom is dit specifiek van belang? Is het de bedoeling om een bestaand model geheel te vervangen, bijv door een model dat beter identificeerbaar is, om vervolgens het moedermodel ‘in de kast te leggen’¹ en het metamodel in te zetten bij eventuele additionele toekomstige modelverbetering/ aanpassing of is het metamodel alleen bedoeld om voor specifieke doeleinden als snelle versie van het moedermodel te functioneren? Denk bij dat laatste aan de mogelijkheid om beleidsvragen snel te beantwoorden, toepassing binnen integrated assessment modellen, maar ook aan mogelijke applicaties in optimalisatie-vraagstukken.

Is bij deze vraag voldoende onderzocht of metamodelleren van het oorspronkelijk moedermodel wel de juiste weg is om het gestelde doel te verwezenlijken? Zou ook volstaan kunnen worden met snellere machines, gridcomputing, verbetering van implementatie, opnieuw vereenvoudigd modelleren op basis van geaggregeerde variabelen, procesinzichten op hoofdlijnen (zie Bouwmans, Costanza et al., 2002), data-reductie etcetera? Denk bij het beantwoorden van deze vragen ook aan aspecten als:

- Wat is de vereiste tijd/geld/mankracht/software/expertise/uitbesteding die nodig is voor metamodelleren (is het uitvoerbaar)?
- Is er voldoende expertise voorhanden over het oorspronkelijke model om verantwoorde keuzes te maken bij het bouwen en evalueren van een metamodel?

¹ Hier is een waarschuwing op zijn plaats: De kwaliteit van het moedermodel zal bepalend zijn voor de kwaliteit van het metamodel. Het is daarom allereerst van groot belang om een goede inschatting te hebben van de kwaliteit van het moedermodel in relatie tot het beoogde toepassingsgebied van het metamodel. Daarnaast is het op zijn plaats om het moedermodel ‘actief’ te houden en de kwaliteit daarvan te verbeteren indien de toepassing dit vereist. Dus de strategie van het ‘in de kast leggen van het moedermodel’ kan op termijn contra-productief werken en leiden tot kennis-verlies en stagnatie.

- Hoe vaak zou het bouwen van een metamodel in de toekomst herhaald moeten worden (bijvoorbeeld na modelaanpassing)?
- Hoe vaak zal het metamodel toegepast worden?
- Moet er per uitvoervariabele mogelijk een apart metamodel gemaakt worden? Is er nagedacht over het beheer van de metamodellen?

B. Status

Wat is de status van het moedermodel? Is het uitontwikkeld, of is nog regelmatig aanpassing voorzien? Is het beheer en dergelijke goed geregeld (zonder stabiel en duurzaam beheerd referentiemodel geen zinnig metamodel)? Zijn er meetdata voorhanden? Wat is de kwaliteit (validatie, onzekerheid) en het toepassingsgebied (incl. beperkingen) van het moedermodel, en wat zijn de relevante kenmerken van het moedermodel (discontinuïteiten, niet-lineariteiten, schaalinvloeden, schaal koppelingen, feedbacks, relevante processen c.q. gevoelige parameters, discrete en continue variabelen, temporele en ruimtelijke schaal en dynamiek)? Wat is bekend over calibratie, onzekerheden en gevoeligheidsinformatie, databenodigheden, runtijd van het moedermodel? Zijn er (elders, internationaal) vergelijkbare (meta)modellen in gebruik; wat valt er te leren van die modellen?

C. Eisen

Kun je a priori vastleggen welke eisen er aan een metamodel (MM) dienen te worden opgelegd (vanuit het beoogde doel), zoals:

- Welke middelen (tijd, expertise, software, rekencapaciteit, budget) zijn beschikbaar om metamodellering uit te voeren?
- Wat is de gewenste toepassingsrange van het MM (bijvoorbeeld extrapoleren of enkel interpoleren; gewenste nauwkeurigheid)?
- Welke afwijkingen (ten aanzien van het moedermodel) zijn acceptabel (gewenste nauwkeurigheid)?
- Wat is de gewenste rekensnelheid en databehoefte voor het MM?
- Welke proceskennis/fysica (processen, parameters, schaalniveau's) wil je terug zien in het MM (transparantie, inzichtelijkheid, plausibiliteit)?
- Hoe ga je de kwaliteit van metamodel beoordelen (beoordelingsmaat)?
- Hoe ga je de onzekerheid van een MM in kaart brengen (in relatie tot de onzekerheid in het moedermodel)?
- Moet de constructie van het MM geautomatiseerd kunnen plaatsvinden, of geprotocolleerd? Zorg er altijd voor dat de activiteiten reproduceerbaar en goed gedocumenteerd zijn.
- Hoe flexibel moet metamodellering zijn (koppeling met expertkennis en gegevens)?
- Moet je het MM zodanig implementeren dat je een automatische waarschuwing krijgt als je je buiten het toepassingsdomein van het MM bevindt?

D. Waarschuwingen

- Qua '**validiteit**' geldt strikt genomen dat:

$$[\text{Toepassingsbereik (MM)} < \text{Toepassingsbereik (moedermodel)}]$$

MM is in feite enkel als interpolatie-instrument te gebruiken, en extrapolatie buiten het toepassingsbereik is pure speculatie. Daarom is het van *cruciaal belang* om je af te vragen hoe plausibel de toepassing van het 'moedermodel' is voor situaties waar het metamodel zal worden toegepast. Mogelijk kan op basis van informatie uit extra meetdata en extra proceskennis (evt. van andere modellen) het toepassingsbereik van het metamodel verder vergroot worden, zodat een mate van extrapolatie verantwoord is. Uiteraard vergt dit een goede argumentatie en evaluatie.

Bedenk dat het in de praktijk niet altijd mogelijk is om het toepassingsbereik van een model goed vast te stellen. Ook bij het moedermodel moet men er beducht op zijn dat dit een hybride kan zijn van harde en minder harde proceskennis, vermengd met (semi)empirische relaties waarvan het toepassingsbereik onduidelijk kan zijn.

- Qua *'pragmatisme'* geldt:
[Toepassingsbereik (MM) > Toepassingsbereik (moedermodel)]
 omdat het metamodel doorgaans met minder data gevoed hoeft te worden en aanzienlijk minder rekentijd vergt zodat het op bredere schaal toepasbaar is.
- Qua *'kennis'* over het moedermodel is het van belang om gedrag en eigenaardigheden, inclusief beperkingen, van dit model goed te kennen (d.w.z. een vorm van modelanalyse, bijvoorbeeld een gevoeligheidsanalyse, moet beschikbaar zijn). Denk daarbij ook aan kennis over 'non-excited modes', dat wil zeggen eigenschappen van het moedermodel (c.q. de werkelijkheid) die in de actueel beschikbare data niet tot uiting komen, maar die wel in het model 'verstopt' zitten ('slapende' dynamica). Zeker indien het metamodel gebruikt wordt voor het doorrekenen van 'extreme' scenario's, dan moet voorkomen worden dat door metamodellering de 'scherpe kanten' van het oorspronkelijke model zijn afgehaald.
- Qua *'consistentie'* moet goed worden vastgelegd welke veronderstellingen ten grondslag liggen aan de modelvorming (onder andere met betrekking tot schaalkeuze). Indien een metamodel in een (model)keten gebruikt gaat worden, dienen de veronderstellingen in de diverse keten-onderdelen op elkaar afgestemd te worden/zijn (bijvoorbeeld het gebruik van eenzelfde langjarig gemiddelde meteo in de diverse ketenmodellen).

Stap 2: Specificatie: wat metamodelleren?

- Wat wil je, gegeven het beoogde doel en de toepassingssituatie, benaderen van het moedermodel? Hierbij gaat het specifiek om de keuze van 'inputs'² (verklarende variabelen) en 'outputs' (verklaarde variabelen; targets, indicatoren) voor het metamodel. Deze keuzes maak je onder andere op basis van data-beschikbaarheid, meetbaarheid, relevantie, beïnvloedbaarheid (bijvoorbeeld management/stuur variabelen).
- Welke schaal/schalen-niveau's (temporeel, spatiaal, maar ook systemisch³) moeten in het MM zichtbaar zijn? Dit geldt zowel voor de gekozen outputs, maar ook voor de inputs en voor de gemetamodelleerde processen. Bedenk dat de schaalniveau's op elkaar afgestemd dienen te zijn en dat hierbij geen relevante dynamiek verloren mag gaan. Dit speelt met name bij aggregatie van grootheden (inputs, output-indicatoren): als er bijvoorbeeld sprake is van sterke niet-lineariteiten dan is het van belang om opschaling/aggregatie zo lang mogelijk uit te stellen ('calculate first/aggregate later') om te voorkomen dat er foutieve inschattingen gedaan worden (zie tekstbox 'Overbruggen schaalniveau's').
- Welke systeem/probleem-kenmerken moeten bij metamodellering zichtbaar zijn? Denk hierbij aan relevante processen/interacties, terugkoppelingen, schaalverbanden (bijvoorbeeld cross-scale interacties). In dit kader is het ook van belang om bij het werken met samengestelde indicatoren ('composite indicators') de afzonderlijke indicatoren in het metamodel op te nemen. Informatie hierover kan zicht geven op belangrijke deelaspecten en op de rol van trade-offs en synergiën tussen processen.
- In hoeverre moeten dynamische of (quasi)-steady state eigenschappen expliciet in het MM worden meegenomen?
- In hoeverre moet het oorspronkelijke model c.q. de modelketen bij metamodellering worden opgesplitst in deelcomponenten (submodellen), die vervolgens, na metamodellering van de afzonderlijke componenten, weer gekoppeld moeten worden (modulaire opbouw)? Hoe wil je deze partitionering en de daarop volgende metamodellering en koppeling tot stand brengen? Speciale aandacht voor schaalniveaus, terugkoppelingen en interacties (ook cross-scale interacties) is hierbij nodig.
- Wat moet er gerapporteerd en gedocumenteerd worden van het metamodelleringsproces?

² Met de generieke term 'inputs' beschrijven we een klasse van grootheden, die zowel modelparameters, begincondities, input functies of andere factoren kan omvatten.

³ D.w.z. de processchaal betreffend (wel of niet samenvoegen van processen).

Metamodellen zijn geschikt voor het overbruggen van verschillende schaalniveau's

Extrapolatie van modelresultaten naar hogere schaalniveaus vergt vaak een majeure investering. De reden hiervoor is dat het vanwege het niet-lineaire karakter van de meeste numerieke modellen niet mogelijk is één modelberekening te maken met gemiddelde eigenschappen als invoer. Leterme, Rounsevell et al., (2004) toonden dit aan in een studie naar de uitspoeling van bestrijdingsmiddelen in het stroomgebied van de Dyle in België. Zij berekenden de uitspoeling voor verschillende bodemkaartenheden met het model GeoPEARL (Tiktak, De Nie et al., 2002). Ze volgden hierbij twee routes:

- De uitspoeling werd berekend voor één gemiddeld profiel per bodemkaartenheid;
- De uitspoeling werd berekend voor alle beschikbare profielen uit het bodeminformatiesysteem. Met gebruikmaking van deze informatie werden frequentiediagrammen getekend.

De gemiddelde uitspoeling per bodemkaartenheid bleek in het eerste geval zwaar onderschat te worden! Het werken met geaggregeerde invoer(parameters) bij sterk niet-lineaire processen is dus niet adequaat; men dient verfijnd te rekenen, en daarna pas te aggregeren.

Om te voorkomen dat het oorspronkelijk numerieke model voor alle bodemprofielen gedraaid moet worden, kan gebruik gemaakt worden van metamodellen. (Vanclouster, Armstrong et al., 2003) pasten deze techniek toe in een Pan-Europese studie. Allereerst werd hiertoe een metamodel van GeoPEARL ontwikkeld (zie §4.6 voor details over de gebruikte methode voor Nederland; de toepassing op de Europese casus vereist een verschillende parametrisatie). Om ervoor te zorgen dat het metamodel op die schaal toepasbaar was, werden er bij de constructie van dit metamodel (uit het oorspronkelijk numeriek model dat initieel een punt-schaal model is) uitsluitend gegevens gebruikt die op die hogere resolutie beschikbaar zijn. Daar waar de toepassing van het oorspronkelijke model beperkt werd door zijn grote databehoeft, kan het metamodel dus adequaat worden ingezet. Op die manier kunnen metamodellen goed gebruikt worden om schaalovergangen naar hogere resolutieniveau's te overbruggen.

Stap 3: Methode: hoe metamodelleren?

In het volgende hoofdstuk wordt meer in detail ingegaan op de verschillende methoden die ter beschikking staan om metamodellen te construeren. Er wordt een onderscheid gemaakt in een drietal benaderingen/klassen, namelijk (a) proces-gebaseerde, (b) statistische en (c) systeemtheoretische metamodellerings technieken.

- (a) **Proces-gebaseerde metamodellering:** deze gebruikt expertkennis over het systeem, bijvoorbeeld vastgelegd in eenvoudige analytische modelrelaties, aangevuld met informatie over dynamische systeemkarakteristieken, zoals impulsresponsfuncties, transferfuncties, om tot een metamodel te komen waarin relevante proceseigenschappen expliciet herkenbaar zijn. Dit wordt vaak verwezenlijkt door de relevante processen die in het referentiemodel beschreven zijn, te vervangen door bovengenoemde vereenvoudigde beschrijvingswijzen. Deze laatste bevatten zogenaamde 'pseudo-fysische' 'effectieve parameters' die via calibratie op basis van simulatiedata met het referentiemodel vastgesteld moeten worden.

De proces-gebaseerde aanpak is met name bruikbaar voor 'extrapolatie' en om expertkennis explicieter in een metamodel vast te leggen. De toepassingsmogelijkheden worden beperkt doordat niet alle niet-lineaire dynamiek eenvoudig is weer te geven, en omdat vereenvoudigingen vaak slechts op geaggregeerde, opgeschaalde, steady state situaties toepasbaar zijn, en weinig informatie geven over heterogene individuele

kenmerken (vb. extremen) en niet-stationair transiënt gedrag. *Van groot belang is het om aan te geven onder welke voorwaarden de vereenvoudigingen adequaat zijn ('validatie' van de gehanteerde metamodel relaties).* In de praktijk worden vaak mengvormen gebruikt, waarin proces-georiënteerde relaties met (semi-)empirische metarelaties worden aangevuld.

- (b) **Statistisch-gebaseerde metamodellering:** Deze aanpak gebruikt statistische modelleringstechnieken om de relatie tussen verklarende (inputs) en verklaarde variabelen (outputs) uit het referentiemodel op een vereenvoudigde manier vast te leggen. Hierbij fungeert het referentiemodel als data genererend mechanisme, waaraan 'inputs' (zeg X) worden aangeboden en dat dan 'outputs' (zeg Y) uitrekent. Op deze manier kan een uitgebreide database van X en Y's gegenereerd worden, die representatief is voor de betreffende toepassingssituatie. Een breed scala van statistische technieken (interpolatie, response surface, regressiebomen, Kriging, GLIM, GAMs, CART, MARS, RBF etcetera zie appendix A) staat ter beschikking om de relatie tussen X en Y (bij benadering) vast te stellen. Keuzes die bij dit proces van statistische metamodellering aan de orde zijn betreffen onder andere:
- Keuze van de *toepassingssituatie* en van de bijbehorende relevante input- en output variabelen, zeg X en Y, evenals hun tijd-, ruimte en systemische schaal in het MM. Ook speelt bij het vaststellen van X en Y's de opsplitsing en analyse van het probleem/referentiemodel een grote rol (partitioneren; combineren; causatie). Zie ook voorgaande Stap 2.
 - Keuze van *inputruimte* (X) en de specifieke punten daarbinnen (keuze van 'experimenteel ontwerp') waarvoor het referentiemodel wordt doorgerekend. Hints: (a) pre-selectie van inputs op basis van gevoeligheidsanalyse; (b) keuze van inputs op basis van hun betekenis, meetbaarheid, beïnvloedbaarheid (stuur/management-variabelen).
 - Keuze van relatie tussen X en Y ('*vorm van het metamodel*'). Hints: onderzoek of niet-lineariteiten, discontinuïteiten, dynamische aspecten een belangrijke rol in het referentiemodel spelen, en in het MM zichtbaar moeten zijn. Keuze van transformaties en interacties die je in metamodel/regressiemodel opneemt. Volstaat het om één metamodel voor alle Y variabelen te maken of moet je meerdere metamodellen gebruiken?
 - Fitten van het metamodel: *fit-criterium* zodanig kiezen dat het een adequate maat geeft voor de overeenstemming met het referentiemodel, in relatie tot de beoogde MM-toepassing. Aandacht voor problemen bij het fitten (meerdere lokale optima; startwaarden; stopcriteria).
 - *Validatie* van het metamodel: vrijwel altijd kruisvalidatie, waarbij een deel van de simulatiedata wordt gebruikt om MM op te stellen, en een ander deel om MM te toetsen. Belang van geschikte performance-maten, hierbij. Voldoet metamodel in belangrijkste range? Hoe gaat het om met extremen? Zie ook §3.5.
 - *Analyse van het metamodel* (bijvoorbeeld gevoeligheidsanalyse): inzicht in de gevoeligheid van Y voor X levert nuttige informatie over de aard van de relatie tussen X en Y, en helpt om de plausibiliteit van het veelal black-box achtige MM te vergroten. Ook is het van belang om na te gaan of het MM gebruikt kan worden om een onzekerheidsanalyse van het referentiemodel uit te voeren. Zie ook §3.5.
- (c) **Systeem-theorie gebaseerde metamodellering:** Deze aanpak focust primair op de *dynamische* karakteristieken van het referentiemodel, en aspecten als selectie van relevante inputs, outputs en toestanden spelen ook hier een belangrijke rol. Ook is bij de keuze van MM-technieken de vorm waarin het referentiemodel beschreven en geïmplementeerd staat van belang, evenals het lineair of niet-lineair karakter ervan. Uiteraard spelen validatie en gevoeligheidsanalyse van het metamodel ook hier een belangrijke rol. Zie voor verdere toelichting §3.3, §3.5 en appendix A.

De keuze van metamodelleringsmethode zal o.a. afhangen van het metamodelleringsdoel c.q. de gewenste toepassingsrange, en van het referentiemodelgedrag. Ruwweg zou je de volgende heuristiek bij deze keuze kunnen hanteren:

- als het referentiemodel zich voornamelijk lineair gedraagt, dan is lineaire interpolatie of lineaire regressie vaak afdoende;
- bij sterke niet-lineariteiten is een andere methodiek vereist (bijvoorbeeld niet-lineaire statistische regressie- of data-mining technieken, met name als het aantal variabelen groot is);
- als dynamische aspecten een grote rol spelen in de toepassing, dan zijn metamodelleringsmethoden uit de systeemtheorie (bijvoorbeeld gebaseerd op niet-lineaire black-box systeemidentificatie) goede kandidaten;
- indien extrapolatie het voornaamste doel is dan is de expliciete inbreng van expertkennis bij metamodellering nodig.

Stap 4: Toepassen van metamodel

Aspecten die hierbij aan de orde komen zijn, (a) door wie zal het MM worden toegepast, (b) welke eisen zijn er aan de MM-kwaliteit gesteld (toepassingsgebied, nauwkeurigheid, transparantie, gevoeligheid, onzekerheid, zeggingskracht) en hoe wordt voorkomen dat het MM oneigenlijk gebruikt wordt; (c) hoe wordt er zodanig over de resultaten gecommuniceerd dat er geen ongeoorloofde conclusies getrokken worden (bijvoorbeeld keuze van schaal voor presentatie gegevens); (d) hoe is het beheer (technisch en inhoudelijk) geregeld?; (e) welke afspraken zijn er ten aanzien van updates?

3. Metamodelleren: een overzicht van methoden

In dit hoofdstuk geven we een kort overzicht van de veelheid van methoden die gebruikt kunnen worden bij metamodellering, waarbij we kort hun voornaamste eigenschappen en voor- en nadelen bespreken. We onderscheiden de methoden in een drietal hoofdklassen, namelijk (a) proces gebaseerde, (b) statistische en (c) systeem theoretische model reductie methoden, en eindigen met een globale vergelijking van deze methoden en een bespreking van aspecten die men bij het toepassen van deze methoden tegenkomt.

Omwille van de leesbaarheid wordt in dit hoofdstuk zo weinig mogelijk op details van de methoden ingegaan. In de appendix worden voor een aantal van de hier genoemde methoden nadere details gegeven.

3.1. Proces-gebaseerde metamodellering

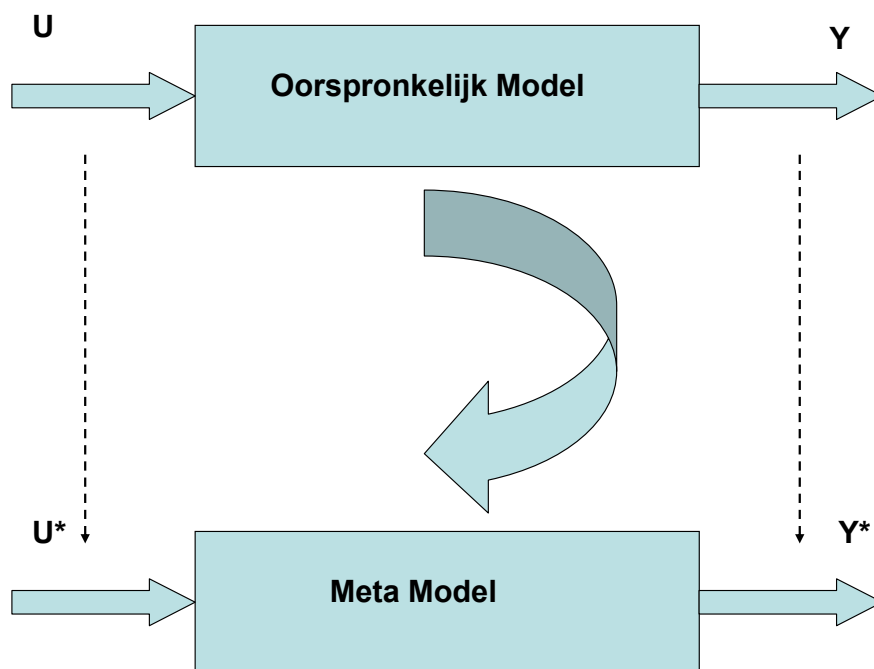
Bij deze vorm van modelreductie probeert men op basis van geabstraheerde procesinzichten - die bijvoorbeeld zijn afgeleid voor geaggregeerde of gesimplificeerde situaties (prototype-situaties) - eenvoudige modelrelaties af te leiden die gebruikt kunnen worden bij de constructie van een metamodel. Hierbij kan men bijvoorbeeld gebruik maken van *impuls-respons functie* of *transfer functie model* representaties (zie bijvoorbeeld Jury, 1982; Young en Lees, 1993; Joos, Bruno et al., 1996; Stewart en Loague, 2003; Stewart en Loague, 2004) of van eenvoudige *analytische modelrelaties* (Van der Zee en Boesten, 1991) die de meest relevante processen beschrijven en die via zogenaamde 'effectieve modelparameters' geparametriseerd zijn. De zo verkregen metamodelrelaties zijn vaak te zien als een hybride menging van empirische, data-gebaseerde informatie met proces-gerelateerde, mechanistische informatie en inzichten (zie ook Data Based Mechanistic Modelling, Young, 1998), en de parameters in deze modelrelaties worden doorgaans via calibratie bepaald op basis van simulatiedata van het referentiemodel, eventueel aangevuld met meetdata. Gebruik makend van deze vereenvoudigde modelrelaties kan de rekentijd en databehoeft van het oorspronkelijke gedetailleerde procesmodel verkleind worden.

Voorbeelden van toepassingen van proces gerelateerde modelreductie zijn bijvoorbeeld te vinden bij modellen voor klimaatonderzoek (Enting, Wigley et al., 1994; Joos, Bruno et al., 1996; Hooss, Voss et al., 2001), hydrologische processen (Young, 2001), bodemtransport (Jury, 1982; Forsman en Grimvall, 2003; Stewart en Loague, 2003; Stewart en Loague, 2004). De aanpak is bruikbaar bij dynamische modellen waarvan de relevante dynamiek tamelijk eenvoudig te beschrijven is, bijvoorbeeld lineair of van een niet-lineair karakter dat in een eenvoudige expliciete vorm uit te drukken is. Aan de parameters in het model kan vaak een 'pseudo-fysische' betekenis gegeven worden (karakteristieke tijdconstanten van sterk gelumpde processen), hoewel de processen als zodanig op het betreffend aggregatieniveau niet hoeven te bestaan, en de parameter waarden slechts na calibratie verkregen kunnen worden. De beperkingen van de techniek liggen onder andere bij de voorwaarden waaronder ze zijn afgeleid: niet-lineaire dynamiek kan slechts beperkt worden weergegeven en vaak is de techniek toegepast op geaggregeerde, opgeschaalde, steady-state situaties en geeft hij weinig zicht op heterogene individuele karakteristieken (bijvoorbeeld extremen) en op niet-stationair transient gedrag.

3.2. Statistiek-gebaseerde metamodellering

De crux van deze methoden bestaat er uit dat het oorspronkelijk referentiemodel als een soort 'virtueel laboratorium' gebruikt wordt, waarmee experimenten worden uitgevoerd die informatie opleveren over het gedrag van het referentiemodel. Het referentiemodel wordt hierbij als een 'data-genererend mechanisme' gebruikt, waaraan 'inputs' worden aangeboden en dat 'outputs' uitrekt. De relatie tussen inputs en outputs wordt vervolgens geanalyseerd

aan de hand van statistische methodes, en dit leidt tot de afleiding van een metamodel (zie Figuur 1).



Figuur 1: Relatie tussen het oorspronkelijk model en het metamodel; u en y zijn de inputs en outputs van het oorspronkelijk model; u^ en y^* zijn de (afgeleide) inputs van het metamodel, na selectie, voorbereiding, aggregatie.*

Opmerking

Omdat het referentiemodel vaak deterministisch zijn, is de relatie tussen inputs en outputs, zeg $Y=F(X)$, deterministisch, en richten de statistische methodes zich op het vinden van een adequate approximatie van deze relatie. Hierbij kunnen zowel de gekozen 'inputs' en 'outputs' tijdvariërende grootheden (tijdreeksen) representeren, en ook kan de relatie $F(\cdot)$ tijdafhankelijk zijn, maar in de praktijk wordt weinig expliciet rekening gehouden met het dynamisch karakter bij de constructie van het metamodel. Doorgaans wordt er naar tijd-gemiddelde- of steady-state situaties gekeken, of naar een beperkt aantal tijdstippen (bijvoorbeeld de toestand in 2010, 2030 etcetera). Een vollediger metamodelering van de dynamiek gebeurt daarentegen wel bij de systeemtheorie gebaseerde methoden die in de volgende paragraaf besproken worden.

Het verkregen metamodel hangt af van de keuze van de inputs en outputs en van de gebruikte statistische methode. Het brede scala van ter beschikking staande statistische technieken maakt het mogelijk om referentiemodellen van verschillende aard en toepassingsdomein te metamodeleren, en potentiële toepassingen kunnen dan ook op zeer breed vlak liggen zowel het ecologisch, economisch en sociaal domein omvattend.

Het statistisch metamodeleren van de input-output relaties heeft doorgaans een sterk empirisch, black-box karakter, en fysische kennis en betekenis komt hierin slechts beperkt tot uitdrukking. Deze situatie kan verbeterd worden door bij het opzetten van het metamodel (onder andere de keuze van inputs en outputs) gebruik te maken van kennis over relevante processen, systeemvariabelen en hun interacties. Ook helpt het om achteraf een

(gevoeligheids)analyse van het verkregen metamodel uit te voeren en zo meer (fysisch) inzicht over de relatie tussen inputs en outputs te krijgen.

In de appendix worden diverse statistische methoden besproken en komt ook het aspect van de keuze van inputs en outputs en van de evaluatie van het verkregen metamodel nader aan de orde. De belangrijkste methoden die daarbij besproken worden zijn (zie ook Hastie, Tibshirani et al., 2001):

- Interpolatie door middel van opzoek-tabellen (Look-Up Tables)
- Polynoom Regressie (Response Surface)
- Kriging en Gaussische Processen
- GLIM (Generalized Linear Interactive Modelling)
- GAMS (Generalized Additive Models)
- CART (Classification And Regression Trees)
- MARS (Multivariate Adaptive Regression Splines)
- RBF (Radial Basis Functions)
- NN (Neurale Netwerken)
- SVM (Support Vector Machines) en Kernel-based learning methoden
- Fuzzy en Neuro-Fuzzy methoden
- Genetische Evolutionaire Algoritmen

3.3. Systeemtheorie-gebaseerde metamodelering

De systeemtheorie richt zich voornamelijk op *dynamische* input/output systemen en modellen, i.e. processen die evolueren in de tijd, en ook hier kan het toepassingsgebied zich uitstrekken over een breed gebied, dat zowel het ecologisch, economisch en sociaal domein omvat. De betreffende dynamische processen worden in veel gevallen beschreven door middel van stelsels differentiaal- of differentievergelijkingen. Voor de deelklasse van lineaire tijdinvariante deterministische systemen (i.e. lineaire differentiaalvergelijkingen met constante coëfficiënten) bestaat een groot scala aan modelreductiemethoden, maar ook voor meer complexe systemen (niet-lineair, stochastisch, partiële differentiaalvergelijkingen etcetera) zijn in de literatuur veel methoden beschreven. Modelreductie komt hierbij meestal neer op het bepalen van een beperkter ('lagere orde') stelsel differentiaal- of differentievergelijkingen, waarmee het systeemgedrag benaderd kan worden.

Ruwweg gesproken kan men de systeemtheorie gebaseerde metamodeleringsmethoden indelen in:

- methoden die simulatiedata (input-output) van het dynamisch referentiemodel gebruiken om tot reductie te komen;
- methoden die zich niet op input-output simulatiedata richten, maar die rechtstreeks uit de modelvergelijkingen en karakteristieken van het referentiemodel een gereduceerde vorm afleiden.

De eerste klasse legt weinig extra eisen op aan het referentiemodel en bestaat uit methoden voor (niet-)lineaire black-box systeemidentificatie die sterk verwant zijn aan de statistische regressie, data-mining en data-reductie technieken. Bij de constructie van een metamodel wordt expliciet rekening gehouden met het dynamisch karakter van de achterliggende signalen, (zie bijvoorbeeld Sjöberg, Zhang et al., 1995; Nelles, 2000). De tweede klasse vereist dat het referentiemodel in een specifieke vorm beschreven en geïmplementeerd is, namelijk als differentie- of differentiaalvergelijkingen, die weergeeft hoe de toestand van het systeem zich door de tijd heen ontwikkelt onder de invloeden (inputs) waaraan het systeem blootstaat. Via state-of-the-art numeriek-algebraïsche algoritmen wordt uit deze zogenaamde toestandruimte beschrijving (of varianten daarvan) rechtstreeks een metamodel afgeleid, op

basis van projectie-operatoren. In toenemende mate vinden toepassingen op zeer grootschalige problemen plaats (bijvoorbeeld transportvergelijkingen beschreven door 2-D en 3-D partiële differentiaalvergelijkingen, waarvan het rekengrid bestaat uit vele tienduizenden componenten; (zie Antoulas en Sörensen, 2001). Nadere details rond methoden uit deze klassen staan in Appendix A gegeven, waarbij een onderscheid tussen lineaire en niet-lineaire systemen wordt gehanteerd:

- Lineaire systemen
 - Gebalanceerde modelreductie
 - Hankelnorm reductie
 - Singuliere verstoringstheorie ('singular perturbations')
 - Modal truncation
 - Moment matching/ Realisatie / Interpolatie
 - Systeemidentificatie
- Niet-lineaire systemen
 - Gebalanceerde modelreductie
 - POD (Proper Orthogonal Decomposition)
 - Stuksgewijze linearisatie
 - Systeemidentificatie

Potentiële toepassingen liggen met name bij modellen waarbij de interne dynamiek en koppeling van processen centraal staat en die als differentiaal- of differentievorm geformuleerd zijn, zoals bijvoorbeeld bij (grootschalige) transport- en stromingsmodellen, systeemdynamische procesmodellen uit de ecologische, economische, sociale en demografische hoek (bijvoorbeeld stock-and-flow modellen, (Serman, 2000)).

3.4. Vergelijking van metamodelleringsmethoden

Het overzicht in voorgaande paragrafen en in appendix A maakt duidelijk dat er zeer veel methoden ontwikkeld zijn en gebruikt worden voor metamodellering. Ook zijn er een aantal vergelijkende studies uitgevoerd waarin diverse methoden vergeleken zijn, met name voor de statistisch georiënteerde methoden. Vaak zijn deze studies echter voor een beperkt aantal voorbeelden uitgevoerd, en niet op een breed scala van complexiteiten, niet-lineariteiten en hoog-dimensionale (veel parameters) problemen gericht (Jin, Chen et al., 2001). De performance van de methoden is bijvoorbeeld bemeten aan de hand van een aantal criteria:

- *Nauwkeurigheid*: hoe goed wordt de systeemresponse (referentiemodel) benaderd over het toepassingsdomein?
- *Robuustheid*: hoe groot is de nauwkeurigheid bij verschillende probleemtipes (bijvoorbeeld niet-lineariteiten, andere omvang probleem, ander sample-aantal)?
- *Rekenkundige efficiëntie*: welke rekeninspanning is vereist voor constructie en gebruik van metamodel?
- *Transparantie*: welk inzicht geeft het metamodel in de belangrijke variabelen en hun interacties/relaties (interpreteerbaarheid en plausibiliteit; mate van black-box karakter)?
- *Conceptuele eenvoud*: hoe complex is de implementatie en het gebruik van het metamodel (vereiste expertise, data, software)?

Metamodelleringskeuze en geschiktheid zal daarnaast ook afhangen van de aard, vorm en eigenschappen van het referentiemodel (dynamische en ruimtelijke aspecten; discontinuïteiten, niet-lineariteiten; interacties en feedbacks; discrete en continue variabelen; rekentijd, databehoeftes). Verder speelt natuurlijk ook de beschikbaarheid van tijd, menskracht, expertise, data en software een belangrijke rol.

Naast bovengenoemde studie (Jin, Chen et al., 2001), worden ook in bijvoorbeeld Simpson, Peplinski et al., 2001; Daberkow en Mavris, 2002; Clarke, Griebisch et al., 2004; Simpson, Booker et al., 2004; Srivastava, Hacker et al., 2004; Tappenden, Chilcott et al., 2004 vergelijkingsresultaten gepresenteerd van statistische metamodelleringsmethoden. Nuttige vergelijkende informatie over de verschillende statistische methoden uit § 3.2 is ook te vinden in Verdonschot en Rebi Nijboer, 2002 en Moisen en Frescino, 2002, weliswaar in een wat andere context dan metamodellering, namelijk in het kader van data-analyse en -modellering van aquatische systemen en bosvegetatieonderzoek.

Tabel 1: Overzicht van de karakteristieken en gebruiksgeschiktheid van enkele statistische methoden voor metamodellering (vrij naar informatie uit Jin, Chen et al., 2001; Daberkow en Mavris, 2002; Clarke, Griebisch et al., 2004; Simpson, Booker et al., 2004; Srivastava, Hacker et al., 2004; Tappenden, Chilcott et al., 2004).

Model Keuze	Karakteristieken/ Gebruiksgeschiktheid
RSM	<ul style="list-style-type: none"> <input type="checkbox"/> Welbekend; Goed en eenvoudig te gebruiken <input type="checkbox"/> Het meest geschikt bij random errors <input type="checkbox"/> Geschikt bij toepassingen met < 10 factoren <input type="checkbox"/> Niet geschikt bij sterke niet-lineariteiten
MARS	<ul style="list-style-type: none"> <input type="checkbox"/> Flexibele software beschikbaar, lage kosten in modelbouw <input type="checkbox"/> Accurater dan lage orde polynoom regressie <input type="checkbox"/> Gevaar van overfitting door grote flexibiliteit. <input type="checkbox"/> Interpretatie kan moeilijk zijn; enige statistische expertise vereist
Neurale Netwerken/Radial Basis Functions	<ul style="list-style-type: none"> <input type="checkbox"/> Goed voor sterk niet-lineaire of zeer omvangrijke problemen (10.000 parameters) <input type="checkbox"/> Meest geschikt voor deterministische toepassingen <input type="checkbox"/> Hoge rekenlast (vaak meer dan 10000 training punten); het beste voor herhaald gebruik
Rule induction/inductive learning	<ul style="list-style-type: none"> <input type="checkbox"/> Beste als factoren en responsies discreet zijn <input type="checkbox"/> Vorm van modellen zijn regels of beslisboom; meer geschikt voor diagnose dan voor engineering design
Kriging, Gaussische Processen	<ul style="list-style-type: none"> <input type="checkbox"/> Zeer flexibel, maar complex <input type="checkbox"/> Geschikt voor deterministische toepassingen <input type="checkbox"/> Kan toepassingen met < 50 factoren aan <input type="checkbox"/> Vereist weinig modelruns, geschikt voor rekenintensieve modellen. <input type="checkbox"/> Er is beperkte ondersteuning bij implementatie
Support Vector Machines	<ul style="list-style-type: none"> <input type="checkbox"/> Flexibel; vereist keuze van Kernel functie (bijvoorbeeld Gaussisch) <input type="checkbox"/> Geen problemen met locale optima; theoretisch en algoritmisch elegant <input type="checkbox"/> Rekenkundig efficiënt en transparant. <input type="checkbox"/> Levert accurate en robuuste approximaties. <input type="checkbox"/> Software in toenemende mate beschikbaar; echter nog vaak onderzoeks-software.

In Tabel 1 is informatie over diverse statistische methodieken uit verschillende literatuurbronnen samengevoegd. De resultaten uit al deze studies laten zien dat alle methoden hun voor- en nadelen hebben, en dat er niet zoiets als een eenduidig beste is aan te wijzen. Bovendien kunnen de uitkomsten van vergelijkingsstudies sterk afhangen van de besproken cases en van de specifieke implementatie en toepassing van de geanalyseerde methoden. Wel kan gesteld worden dat polynoom-gebaseerde regressie methoden snel tekort zullen schieten als het referentiemodel belangrijke niet-lineariteiten vertoont of als de dimensie van de parameter ruimte groot is. Van Kriging of Gaussische Proces gebaseerde methodieken zijn goede resultaten gemeld bij het benaderen van rekenintensieve niet-lineaire modellen van beperkte dimensie (≤ 50 parameters), waarbij het aantal referentiemodelruns

beperkt diende te blijven in verband met de hoge rekentijd. Voor niet-lineaire referentiemodellen van hogere dimensie is niet-lineaire regressie via Radial Basis Functions en Neurale Netwerken bruikbaar. Vaak zijn hierbij echter veel runs met het referentiemodel nodig om een representatieve dataset voor het modelgedrag te krijgen voor adequate training van het neurale netwerk, en dient men beducht te zijn op problemen met het fitten van het meta-model (meerdere lokale optima; overfitting op de trainingsdataset ten gevolge van de vele instelparameters van het meta-model (overparametrisatie)). Recentelijk worden ook zeer goede resultaten gemeld met relatief nieuwe technieken uit de Statistical Learning Theory, zoals SVR (Support Vector Regression), zie Clarke, Griebisch et al., (2004). Door een ruimere formulering van het fitprobleem worden problemen met meerdere lokale optima vermeden, en kan efficiënt gebruik worden gemaakt van de meest informatieve datapunten. Deze theoretisch elegante aanpak is een beloftevolle ontwikkeling die naar verwachting in de toekomst meer aandacht zal krijgen bij meta-modellering.

Voor de systeemtheorie gebaseerde meta-modelleringmethoden zijn weinig breed georiënteerde vergelijkende studies uitgevoerd. In § 7.6 in Varga (2001) worden diverse systeemtheoretische softwaretools voor modelreductie kort vergeleken. Nelles (2000) geeft een kwalitatieve karakterisering en evaluatie van diverse methoden die gebaseerd zijn op niet-lineaire black-box systeemidentificatie. Hierbij gebruikt hij een gedifferentieerde set van evaluatiecriteria, waarbij zaken aan de orde komen als interpolatie- en extrapolatie gedrag, localiteit, gladheid, gevoeligheid voor ruis, parameter optimalisatie, on-line aanpassing, trainingsnelheid, rekentijd meta-model, omvang/dimensie, beperkingen, modelconstructie.

3.5. Toepassingsaspecten

We bespreken tot slot in het kort enkele belangrijke onderwerpen, waarmee men te maken krijgt bij praktische uitvoering van meta-modellering, en verwijzen naar de betreffende literatuur voor meer informatie:

Doel van meta-modellering

Centraal bij meta-modellering staat het doel en de context waarvoor het meta-model gebruikt zal worden. Deze bepalen voor een groot deel de eisen die aan het meta-model gesteld worden, bijvoorbeeld ten aanzien van gewenste nauwkeurigheid, tijd- en ruimte schaal, dynamisch karakter, plausibiliteit en inhoudelijke correctheid, gebruiksgemak (rekensnelheid, data-behoefte, maakbaarheid, aanpasbaarheid), interpreteerbaarheid en inzichtelijkheid, flexibiliteit (bijvoorbeeld koppeling met expertkennis en gegevens) etcetera. Vaak kan niet aan alle eisen in eenzelfde mate voldaan worden, en zal de keuze in veel gevallen *pragmatisch* zijn, waarbij de beschikbaarheid van resources (tijd, menskracht, expertise, software, data) een belangrijke rol spelen.

Het spreekt voor zich dat het referentiemodel *qua inhoud en kwaliteit* in principe ook bruikbaar moet zijn voor het gestelde doel; echter ten gevolge van beperkingen in gebruiksgemak van dit oorspronkelijke model is er bewust gekozen om gebruik te maken van een meta-model. In het algemeen kan gesteld worden dat het meta-model specifiek en meer gefocust zal zijn dan het referentiemodel. Ook kunnen voor diverse doelen/taken verschillende meta-modellen van eenzelfde referentiemodel nodig zijn: een meta-model dat adequaat is voor het ene doel kan voor een ander doel hopeloos tekort schieten. Dit maakt het belang duidelijk van het expliciet aangeven van de toepassingsrange van meta-model, en van zijn kwaliteit in relatie tot de gestelde doelen.

Keuze van inputs en outputs

De keuze van *welke variabelen* van het oorspronkelijke model te gebruiken bij de constructie van een meta-model hangt natuurlijk met het beoogde doel van het meta-model samen. Het is hierbij van belang om je af te vragen welke *processen en koppelingen* in het systeem wezenlijk zijn, en hoe herkenbaar en expliciet deze in een meta-model moeten terugkeren -

aldus het black-box karakter van metamodellen terugdringend (zie ook Davis en Bigelow, 2003). Ook de vraag naar het *schaalniveau* waarop deze grootheden moeten worden meegenomen in het metamodel speelt een grote rol. Een beschrijving die zich richt op een te ruwe schaal, en bijvoorbeeld relevante processen op een schaal erboven of eronder onvoldoende⁴ meeneemt loopt de kans dat essentiële aspecten in de spatiële en ruimtelijke dynamiek gemist worden.

In de praktijk zullen bij de selectie en aggregatie van de (invoer-, uitvoer-, toestands-) variabelen van het referentiemodel diverse factoren meespelen, onder andere:

- (a) in hoeverre is er (empirische) informatie beschikbaar over de (invoer)variabelen bij de beoogde toepassingen?
- (b) wat is hun betekenis en beleidsrelevantie (bijvoorbeeld stuur- en omgevingsvariabelen; indicatoren/targets)?
- (c) hoe sterk dragen de (invoer)variabelen bij aan de betreffende outputs/indicatoren;
- (d) welke verwaarlozingen (bijvoorbeeld ‘plat slaan’ van dynamiek en ruimtelijke variatie) zijn geoorloofd en welke informatie gaat daardoor verloren?

Gevoeligheidsanalyses van het referentiemodel kunnen nuttige informatie over deze aspecten opleveren en dragen bij tot een adequate screening en selectie van relevante invoervariabelen (zie Ratto, Tarantola et al., 2000; Ratto, Tarantola et al., 2000; Alam, McNaught et al., 2004).

Genereren van data voor metamodel bouw (keuze van experimenteel ontwerp)

In feite betreft dit het ‘experimenteren’ met het referentiemodel, waarbij een (groot) aantal situaties worden doorgerekend die de kale gegevens opleveren voor de constructie van het metamodel. In de praktijk komt dit vaak neer op het ‘trekken’ van waarden uit de gekozen input-ruimte⁵ en het uitvoeren van simulaties met het referentiemodel voor deze instellingen. Van belang is om een parameterdomein af te dekken dat representatief is voor de toepassing van het metamodel, dus niet te krap, maar ook niet te ruim. Ook dienen de trekkingen voldoende informatief en effectief zijn bij het opsporen van relevante factoren en hun interacties. Zo is het bijvoorbeeld belangrijk om informatie op te sporen over ‘slapende’ processen/parameters die bij veranderende omstandigheden (domain shifts; extreme scenario’s) plotseling een relevante rol kunnen spelen. Belangrijke facetten bij het uitvoeren van deze trekkingen (dat wil zeggen de keuze van ‘experimenteel ontwerp’) zijn onder andere het aantal runs dat nodig is, de wijze waarop de parameterinstellingen bepaald worden (systematisch, random etcetera), de eenvoud en uitvoerbaarheid van de ontwerpimplementatie, de flexibiliteit van het ontwerp (gebruik van expert-kennis, ‘adaptatie en leren’ bij de keuze van trekkingen). In de literatuur is uitgebreide informatie te vinden over designs voor computermodel-experimenten (zie bijvoorbeeld Simpson, Booker et al., 2004 en de referenties daarin). Ook het gebruik van adaptieve en sequentiële designs, waarbij al lerenderwijs gericht gezocht wordt in de relevante delen van de parameterruimte, is hierbij van belang (Jin, Chen et al., 2002; Turner, Campbell et al., 2003). Dit speelt onder andere bij toepassingen van metamodellering voor optimalisatie- en designdoeleinden, waar de interesse uitgaat naar die deelgebieden van de parameterruimte waar de doelfunctie optima heeft (Jin, Chen et al., 2003; Wang, 2003; Wang en Simpson, 2004; Qian, Seepersad et al., 2005); zie ook <http://endo.sandia.gov/DAKOTA/>. Zie verder Simpson, Mauery et al., 2001; Simpson, Dennis et al., 2002; Kleijnen, Sanchez et al., 2004; Kleijnen en van Beers, 2004; Kleijnen, 2005 waarbij met name design strategieën voor Kriging en Response Surface methoden aan bod komen.

Opmerking

Indien de invoerparameters X^* voor het metamodel geaggregeerde versies zijn van de variabelen X van het referentiemodel, dan vereist dit extra zorg bij de trekking. Immers, een

⁴ Een vaak gehanteerde vuistregel bij onderzoek naar schaal-aspecten, is om te kijken naar processen op de vigerende schaal, in relatie tot processen op een naast gelegen schaalniveau, erboven en eronder.

⁵ De invoerparameters van het referentiemodel die niet als invoer zijn geselecteerd voor het metamodel worden hierbij op hun nominale waarden gezet.

specifieke waarde X^* kan verwijzen naar verschillende achterliggende realisaties van X (vergelijk bijvoorbeeld een gemiddelde weektemperatuur die tot stand kan komen bij verschillende series dagtemperaturen).

- Indien de trekking initieel op het niveau van de X^* -en plaatsvindt, dan zal deze eerst moeten worden doorvertaald naar een (of meerdere) bijbehorende realisatie op het niveau van X en hierbij is in feite sprake van een *intrinsieke keuzevrijheid*. Doorgaans wordt deze keuzevrijheid in de praktijk ingeperkt door één specifieke keuze⁶ voor X te maken waarmee vervolgens het referentiemodel wordt doorgerekend, waarna de gesimuleerde modeloutputs Y uiteindelijk geaggregeerd/getransformeerd worden tot Y^* . Het spreekt voor zich dat deze specifieke keuze beargumenteerd moet worden, en dat de consequenties van deze invulling aangegeven moeten worden.
- Vindt de trekking daarentegen initieel op het niveau van de X -en plaats, (bijvoorbeeld trekken van een realisatie van een tijdreeks met dagtemperaturen), dan kan hiermee rechtstreeks de simulatie van het referentiemodel worden aangestuurd, die tot de output Y leidt. Na postprocessing levert dit de bijbehorende X^* en Y^* op het niveau van de metamodelparameters. In feite herbergt ook deze procedure een intrinsieke keuze in meerduidigheid, die bovendien niet vrij is van een zekere willekeur: immers, een andere realisatie op referentiemodel-niveau, zeg X' , die bijvoorbeeld tot een zelfde X^* leidt, heeft een bijbehorende referentiemodel output, zeg Y' , die niet noodzakelijk tot dezelfde Y^* hoeft te leiden. Deze meerduidigheid kan in feite discontinuïteiten in de relatie tussen X^* en Y^* veroorzaken, maar in de praktijk zullen deze vaak glad gestreken worden door de (statistische) metamodelleringsprocedure.

Keuze van metamodelleringstechniek

In voorgaande paragraaf is aangegeven dat er vele metamodelleringsmogelijkheden zijn, en dat er geen eenduidige 'beste keuze' is aan te geven. Veel zal afhangen van de eisen die een specifiek referentiemodel en toepassingssituatie stelt. Aspecten als de rekensnelheid van het referentiemodel, vorm van referentiemodel, aard van de inputs en outputs en van hun relatie⁷ spelen hierbij een rol. Bovendien zal de keuze ook voor een groot deel afhangen van de beschikbaarheid van expertise en data, de vertrouwdeheid met de metamodelleringstechniek, en wensen ten aanzien van transparantie en inzichtelijkheid van het metamodel. In situaties waarin er weinig acceptatie is voor black-box achtige metamodellen, zullen technieken moeten worden gebruikt die het mogelijk maken om proces- en expertkennis in handzame en flexibele vorm te incorporeren in de metamodellering. Zie Davis en Bigelow (2003) die het belang van 'motivated metamodels' onderstrepen.

Validatie/evaluatie aspecten

De cruciale vraag die hierbij gesteld wordt is hoe goed het metamodel voldoet in de belangrijkste toepassingsrange? Hoe gaat het bijvoorbeeld om met extremen? Hieronder ligt ook de vraag hoe goed het metamodel het referentiemodel benadert, dat wil zeggen hoe sterk lijken de metamodel resultaten op de resultaten van het referentiemodel in situaties die representatief zijn voor de beoogde toepassing. Voor dit soort evaluaties moeten testsituaties worden gedefinieerd en criteria op basis waarvan de resultaten vergeleken en beoordeeld worden. Hierbij is het van belang om vast te stellen wat de gewenste nauwkeurigheid van het metamodel is, rekening houdend met de mate waarin het referentiemodel 'de werkelijkheid' benadert. Zie bijvoorbeeld Meckesheimer, Booker et al. (2002), waarin het gebruik van deze zogenaamde 'leave-k-out' kruisvalidatie strategieën bij metamodel constructie en evaluatie wordt besproken. Zie verder ook Kleijnen en Sargent, 2000; Bayarri, Berger et al., 2002;

⁶ Bijvoorbeeld door het opleggen van een specifiek weekpatroon voor X dat tot het gemiddelde X^* leidt; het doorrekenen van een groot aantal realisaties X die tot dezelfde X^* leiden, om zo een soort gemiddelde Y^* van de bijbehorende outputs Y te bepalen, is doorgaans te rekenintensief.

⁷ Bijvoorbeeld zijn de inputs en outputs continue of discrete (bijvoorbeeld klassen) variabelen?; Is de relatie tussen inputs en outputs glad/continu of zijn er ook discontinuïteiten? Is deze relatie sterk niet-lineair? Welke temporele en spatiële aspecten hebben de inputs en outputs en hun relatie?

Easterling en Berger, 2002; Meckesheimer, Booker et al., 2002. In Jin, Chen et al. (2001) worden de resultaten tussen metamodel en referentiemodel vergeleken op basis van kwadratische fouten maat (RMSE), relatieve absolute foutmaat, en de relatieve maximale absolute fout. Ook kan een analyse van de schattingsresiduen mogelijke problemen aan het licht brengen.

Behalve de nauwkeurigheid van de fit spelen ook andere aspecten een rol bij de evaluatie van een metamodel, zoals plausibiliteit, inzichtelijkheid, gebruiksgemak (e.g. tijdbesparing, rekenefficiency, databehoeft, graad van complexiteit, vereiste expertise, software, restricties).

Metamodellering en onzekerheden

Doordat bij metamodellering het referentiemodel benaderd wordt door een eenvoudiger model voegt metamodellering onzekerheid toe. Het is niet altijd eenvoudig vast te stellen hoeveel dit is, en dit vereist vergelijking tussen het oorspronkelijk model en het metamodel voor diverse parameterinstellingen die representatief en relevant zijn voor de beoogde toepassing.

Ook dient vastgesteld te worden of deze toegevoegde onzekerheid substantieel is vergeleken met de onzekerheid die intrinsiek is aan het referentiemodel. Dit vereist allereerst dat er een onzekerheidsassessment van het oorspronkelijk model is uitgevoerd. Doorgaans zal dit overigens niet of nauwelijks het geval zijn, met name niet voor rekenintensieve referentiemodellen. In zo'n situatie kan wel - onder bepaalde voorwaarden - het metamodel gebruikt worden om een assessment van die onzekerheid te verkrijgen (zie Ratto, Tarantola et al., 2000; Chen, Jin et al., 2004; Jin, Chen et al., 2004). Hiervoor is wel vereist dat de onzekerheden die een rol spelen bij het oorspronkelijke model ook expliciet (bijvoorbeeld als input) herkenbaar zijn in het metamodel, en dat de onzekere parameters van het referentiemodel die niet meedoen bij het metamodel nauwelijks effect hebben op de onzekerheid in de outputs van het referentiemodel. Uiteraard dienen deze voorwaarden getoetst te worden.

Opmerking

De systeemtheoriegebaseerde metamodelleringstechnieken die werken met expliciete transformaties en projecties in de toestandsruimte voldoen **niet** aan deze voorwaarden: de parameters in het oorspronkelijke referentiemodel worden uit het zicht verloren bij deze transformaties, en zijn niet meer expliciet herkenbaar in het metamodel. Dit heeft tot gevolg dat een adequate onzekerheidsanalyse langs deze weg niet meer mogelijk is, omdat het niet helder is hoe onzeker de parameters van het metamodel zijn en hoe deze gerelateerd is aan de parameteronzekerheid in het referentiemodel.

Modelketens/composities

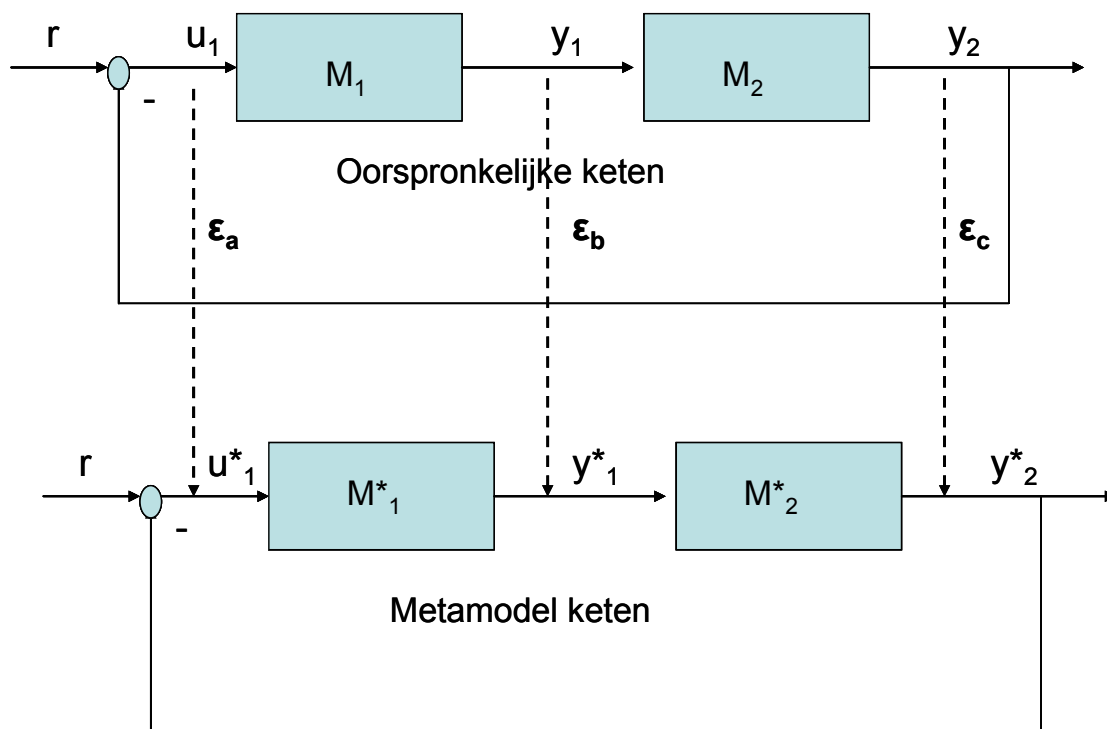
In situaties waarin sprake is van een modelketen of van submodellen die onderling (sterk) verbonden zijn tot een groot overall-model, is de vraag of we ons bij metamodelleren dienen te richten op de afzonderlijke (deel)modellen en deze allereerst moeten metamodelleren om ze daarna weer aan elkaar smeden, of dat we daarentegen moeten streven naar één groot metamodel voor de keten in zijn geheel.

Bij de laatstgenoemde strategie is het voorshands niet duidelijk of dit wel in één enkele slag te verwezenlijken is en of het mogelijk is om voldoende simulaties met de totale modelketen te doen om een representatieve database te construeren voor de metamodelbouw. Dit metamodel zou bovendien informatie moeten geven over relevante (sub)processen en hun interrelaties. Dit vereist zorgvuldige keuze van inputs en outputs bij metamodellering en vraagt bovendien om een aanpak die transparant en flexibel genoeg is om proceskennis expliciet toe te voegen (dat wil zeggen vermijden van een te hoog 'black-box' gehalte).

De eerste strategie lijkt uit het oogpunt van 'verdeel en heers'-strategie pragmatisch beter uitvoerbaar. Er kleven echter ook een aantal *serieuze problemen* aan, met name rond de vraag hoe de oorspronkelijke keten te partitioneren en de componenten vervolgens te metamodelleren en weer te koppelen. Compatibiliteit en consistentie in de keuze van

variabelen en schaal is hierbij belangrijk, waarbij achterliggende aannames met elkaar in overeenstemming zijn voor de verschillende ketenonderdelen (bijvoorbeeld keuze van langjarig gemiddelde meteo etcetera). Ook is zorg vereist voor (positieve) feedback-koppelingen, omdat de fouten en verwaarlozingen die ontstaan bij metamodellering van de submodellen versterkt kunnen worden in de feedback-keten (zie Figuur 2). De performance van de gekoppelde metamodel keten kan hierdoor sterk verslechteren, en relevante eigenschappen (bijvoorbeeld stabiliteit) van het oorspronkelijk modelsysteem kunnen op deze manier ‘verdwijnen’, met alle gevolgen van dien.

Het moge duidelijk zijn dat metamodellering van gekoppelde modellen extra uitdagingen stelt, niet in de laatste plaats aan afstemming, organisatie en samenwerking van betrokkenen. Welke metamodelleringsstrategie er uiteindelijk ook gekozen wordt, er zal altijd een gedegen evaluatie/validatie moeten plaatsvinden, zowel op het niveau van de relevante onderdelen, als op het niveau van het totaalsysteem.



Figuur 2: Fouten en verwaarlozingen bij metamodellering van modellen in een modelketen met feedback.

Gebruikerssuggesties/ operationele aspecten

Bij metamodellering is het verstandig om eerst na te gaan in hoeverre eenvoudige methoden (bijvoorbeeld lineaire regressie; lineaire interpolatie) al afdoende zijn als basis voor meta-model constructie, alvorens meer geavanceerdere technieken of proceskennis in te brengen. Eén en ander zal afhangen van niet-lineariteiten van het referentiemodel en van het extrapolatie-karakter van de beoogde toepassing. Ook is het gebruik van een adaptieve strategie om benaderende modelrelaties vast te stellen aan te bevelen, waarbij bijvoorbeeld naar wens gefocust wordt op die onderdelen van de parameter ruimte die er toe doen. Zie Wang en Simpson, 2004.

Daarnaast moeten we ons er rekenschap van geven dat metamodellering zorg en weerstand bij onderzoekers/modelleurs kan oproepen. Met name bestaat er bezorgdheid dat metamodellering leidt tot ongeoorloofde simplificaties, waarbij essentiële dynamiek ‘platgeslagen’ wordt en er te kort door de bocht wordt gemodelleerd. Ook het gebruik van

black-box technieken stuit op weerstand, omdat er geen proceskennis in te herkennen is. Bovendien is men vaak bang dat het metamodel oneigenlijk gebruikt gaat worden (oprekken van het toepassingsgebied) en een eigen leven gaat leiden en onvoldoende geactualiseerd wordt. Ook kan er bezorgdheid zijn dat anderen er met de oorspronkelijke onderzoeksresultaten vandoor gaan, en dat een metamodel de doodsteek is voor verdere modellering en kennisontwikkeling.

Naast deze ietwat negatief getinte aandachtspunten, biedt metamodellering ook veel extra kansen met name om modellen in te zetten in een bredere context en problematiek en bij integrale studies. Het is hierbij duidelijk dat metamodellering vergezeld dient te gaan van overwogen keuzes en een gedegen evaluatie. Met name voor het gebruik bij beleidstoepassingen is dit een belangrijke vereiste. Het dient helder te zijn wat de resultaten van een metamodel waard zijn, en welke conclusies wel of niet geoorloofd zijn. Ook is het van belang om de resultaten zodanig te communiceren dat er geen ongeoorloofde conclusies getrokken worden, en om bijvoorbeeld de schaalkeuze bij presentatie weloverwogen te doen, en de voorwaarden expliciet te maken waaronder de resultaten afgeleid zijn, met expliciete informatie over hun nauwkeurigheid. Het is duidelijk dat het gebruik van (meta)modellen een voortdurende toetsing aan nieuwe inzichten vereist, waarbij het raadzaam is om het achterliggend referentiemodel als benchmark te houden en dat, samen met het metamodel, zonodig te updaten naar nieuwe inzichten.

4. Metamodelleren bij het MNP: enkele ervaringen en leerpunten

In dit hoofdstuk geven we een overzicht van metamodelleringstoepassingen die in het verleden bij het MNP hebben plaatsgevonden. We geven hierbij kort de context aan waarbinnen metamodellering heeft plaats gevonden, en bespreken de gehanteerde methode en de verkregen resultaten per studie. De studies worden in volgorde van ingewikkeldheid van de gebruikte metamodelleringstechnieken gepresenteerd. We eindigen met enkele algemene bevindingen en leerpunten uit deze studies.

4.1 Metamodellering voor Stone

Het nutriënten-emissiemodel Stone (Versie 2.0) is een collectief model van RIVM, RIZA en Alterra voor de berekening van de emissie van stikstof (N) en fosfor (P) naar grond- en oppervlaktewater. Stone gebruikt als rekenbasis een ruimtelijke aggregatie, waarbij voor een groot aantal (6405) plots (dat wil zeggen unieke combinaties van hydrologische omstandigheden, bodemtype en landgebruik) N- en P-emissies worden doorgerekend waarin het *huidige* landgebruik als belangrijke factor is meegenomen. Aanpassing hiervan voor *andere ruimtelijke scenario's* zou een tijdrovende bezigheid zijn. Omdat dit soort scenario's essentieel onderdeel uitmaken van de Natuurverkenningen (NVK), is besloten ten behoeve van de NVK 2002, het model MetaStone (versie 1.0) te ontwikkelen, een metamodel gebaseerd op Stone, dat kan worden toegepast bij nationale ruimtegebruik varianten die afwijken van de huidige (Wortelboer, Rosenboom et al., 2005).

In feite bestaat het model MetaStone uit een opzoektabel (Look-Up Table), die zo uitgebreid is opgezet dat er een breed toepassingsdomein wordt afgedekt. De tabel is gemaakt door:

1. Een keuze te maken uit de invoerfactoren van Stone;
2. Stone berekeningen uit te voeren voor alle plots, gebruikmakend van een sample van deze invoerfactoren;
3. Deze uitvoerresultaten te aggregeren, gekoppeld aan de gebruikte invoerfactoren (zie hierna);
4. Invoer- en uitvoergegevens in een 'tabel' op te slaan.

Hierbij zijn de volgende invoerfactoren bij de modellering meegenomen (via 'expert judgement' werd de keuze beperkt tot de meest relevante factoren in relatie tot het beoogde doel van het kunnen evalueren van verander(en)de landgebruiksscenario's):

- a. Landgebruik (indeling in 3 klassen: gras, maïs/bouwland, natuur);
- b. Grondwatertrap (zoals in Stone en in de Nederlandse bodemkaart gebruikelijk);
- c. Bodemtype, deels gecombineerd met fysisch-geografische regio (klei, veen en 4 zandklassen);
- d. Kwel/infiltratie (4 klassen, zoals in Stone);
- e. (steady-state) Bemesting (discrete niveaus).

Met uitzondering van de laatste invoerfactor, die continue waarden kan aannemen, zijn alle invoerfactoren klassen.

Opmerking

In eerste instantie was geen rekening gehouden met meteo-gebieden en eventuele regionale verschillen in mestbewerking, met als gevolg dat voor al het zandgebied van Nederland op dezelfde manier werd gerekend en dat, mede door verwaarlozing van de invloed van meteo- en mestbewerkingsverschillen, de resultaten voor zand te grof werden. Vervolgens is echter het zandgebied opgedeeld naar de 4 fysisch-geografische zandregio's, waarbij de meteo- en mestbewerkingsinvloed dus impliciet in meegenomen is (van Gaalen, 2004).

Er zijn 4 Stone runs uitgevoerd waarbij de factoren landgebruik, grondwatertrap, bodemtype, en hydrologie gelijk zijn gehouden aan de huidige situatie; de verhoudingen tussen diverse N- en P- componenten zijn per plot (initiële 'oplading' per plot) vastgezet op de situatie zoals die voor 2000 in de Stone invoer wordt gebruikt; het bemestingsniveau verloopt voor de jaren 2000-2005 lineair van de 1995 situatie naar de opgegeven steady-state bemestingsituatie (vastgesteld binnen de NVK) en wordt vervolgens van 2005-2030 constant gehouden.

In feite zijn dus alleen de bemestingsniveaus verschillend per run. Deze niveaus zijn gerelateerd aan landgebruik en zijn ongeacht het bodemtype vastgesteld. De steady-state waarden gelden voor de periode 2005-2030, en zijn weergegeven in de volgende tabel:

Tabel 2: *Steady State bemestingswaarden in de diverse STONE-runs*

Gewas	N (kg/ha)				P (kg/ha)			
	<i>Run1</i>	<i>Run2</i>	<i>Run3</i>	<i>Run4</i>	<i>Run1</i>	<i>Run2</i>	<i>Run3</i>	<i>Run4</i>
<i>Gras</i>	105	130	315	360	15	15	25	30
<i>Mais</i>	95	150	190	125	15	15	20	20
<i>Bouwland</i>	95	150	190	125	15	15	20	20
<i>Natuur</i>	0	0	0	0	0	0	0	0

Op basis van deze 4 Stone runs konden in totaal 1235 unieke combinaties van waarden voor invoerfactoren onderscheiden worden. Daarnaast zijn ook nog enkele aanvullende 'generieke' combinaties doorgerekend om situaties te dekken van niet voorkomende combinaties. Omdat elke unieke combinatie kan refereren aan meerdere van de oorspronkelijke 6405 Stone plots, zijn de Stone uitkomsten gemiddeld over alle plots die bij zo'n unieke combinatie horen. Bij deze middeling is geen rekening gehouden met verschillen in areaalafmetingen tussen plots (ongewogen middeling; (persoonlijke communicatie Van Gaalen, 2004))

Van de aldus bepaalde gemiddelden worden de decade- (10 dagen, de kleinste tijdstap waarin binnen Stone gerekend wordt) en kwartaalwaarden in het jaar 2025 van de volgende variabelen bepaald, en als outputs voor de tabellen van MetaStone gebruikt:

- i. Het gemiddelde debiet van de oppervlaktewaterafvoer ('laterale drainage') in [mm];
- ii. De gemiddelde vrachten NO₃, NH₄ en organisch N in [kg/ha N] naar oppervlaktewater;
- iii. De gemiddelde vrachten PO₄ en organisch P in [kg/ha P] naar oppervlaktewater.

Opmerking

De keuze van 2025 als targetjaar lijkt gemaakt om - rekening houdend met de 15 weerjarencyclus van Stone - vergelijking met de situatie van 1995 mogelijk te maken. Er is van uit gegaan dat in 2025 een steady-state is bereikt wat betreft de uit- en afspoeling. Hierbij moet echter worden opgemerkt dat het geenszins duidelijk is of in 2025 al sprake is van een steady state, met name niet bij P in verband met het fosfaatbindend vermogen van de bodem. Ook is er natuurlijk in 1995 geen sprake van een steady-state situatie. Om een meer representatieve vergelijking te krijgen, zou het beter zijn indien zowel aan het begin als aan het eind een langjarig gemiddelde genomen zou zijn, dus bijvoorbeeld het (weerjarencyclus)

gemiddelde over 1986-2000 en het gemiddelde over 2031-2045. Ook is langer doorrekenen aan te bevelen omdat dan duidelijk kan worden of er echt wel sprake is van een steady-state situatie. Bij dit alles dient wel aangetekend te worden dat een dergelijke uitbreiding nog niet mogelijk was ten tijde van de beschreven studie (Persoonlijke communicatie Beusen, 2004).

Op bovenbeschreven manier is een opzoektabel met berekeningsresultaten samengesteld. Het achterliggende idee bij het gebruik van de opzoektabel is dat men nu de beschikking heeft over voldoende doorgerekende situaties om uitspraken te kunnen doen over de uitspoelingswaarden bij andere landgebruik-scenario's dan de huidige. Onder de aanname dat de invoer/uitvoerdata in de tabel betrekking hebben op 'steady state situaties' onder diverse bemestingsniveaus, *en* dat steady state situaties bij veranderd landgebruik daarmee overeenkomen⁸, kunnen de uitspoelingswaarden dus rechtstreeks uit de tabellen gelezen worden voor de betreffende combinaties van invoerfactoren. Om de methodiek verantwoord toe te passen, dient deze cruciale aanname op geldigheid getoetst te worden. Dat is tot dusverre niet gebeurd.

Wortelboer, Rosenboom et al. (2005) presenteren enkele validatie-resultaten waaruit onder andere blijkt dat de MetaStone berekeningen meer zijn uitgemiddeld dan de resultaten van Stone, doordat er een grovere aggregatie van invoer-factoren heeft plaatsgevonden. Een onderlinge vergelijking van Stone en MetaStone op basis van een ruwe classificatie (gelieerd aan doel- en normstelling) van outputkaartbeelden (zie Figuur 3 voor een voorbeeld), laat zien dat er voor 50-60% van het gridareaal in Nederland geen verschil is tussen de Stone en MetaStone outputklassen, terwijl voor een klein deel (< 10%) de resultaten van Stone en MetaStone onderling meer dan 1 klasse verschillen (zie Figuur 4). Merk hierbij wel op dat, met name voor P-totalen, de verschillen vooral optreden in problematische gebieden met hoge P-concentratiewaarden, zoals Noord- en Zuid-Holland en Zuid-Limburg. Daar lijkt metaSTONE *systematisch een onderschatting* van de problematiek te geven, hetgeen veroorzaakt lijkt door een te grove aggregatie/uitmiddeling bij MetaStone. Vergelijk ook de eerdere opmerking over GeoPEARL in de tekstbox in hoofdstuk 2.

Dit ondergraaft voor een deel de opmerking in Wortelboer, Rosenboom et al. (2005) dat de resultaten uit MetaStone 1.0 bruikbaar zijn voor de NVK, omdat men voornamelijk geïnteresseerd is in landsdekkende, indicatieve resultaten op een hoog geaggregeerd niveau, waarbij het vooral om *verschillen* tussen toekomstscenario's gaat en niet om absolute resultaten. Eén en ander maakt duidelijk dat verbetering van MetaStone gewenst is voor de hierboven beschreven toepassing; behalve aan een meer gefundeerde keuze van de invoerfactoren (bijvoorbeeld expliciet meenemen van drainage effecten bij plots) en door te rekenen instelwaarden, tijdframe en tijdmiddeling bij Stone simulaties, kan hierbij ook gedacht worden aan het gebruik van meer geavanceerde metamodelleringsmethoden dan opzoektabelen, en het doorrekenen van Stone over een langere tijdspanne om steady-state situaties te garanderen. In 2005 zal een studie hiernaar worden uitgevoerd.

Opmerking

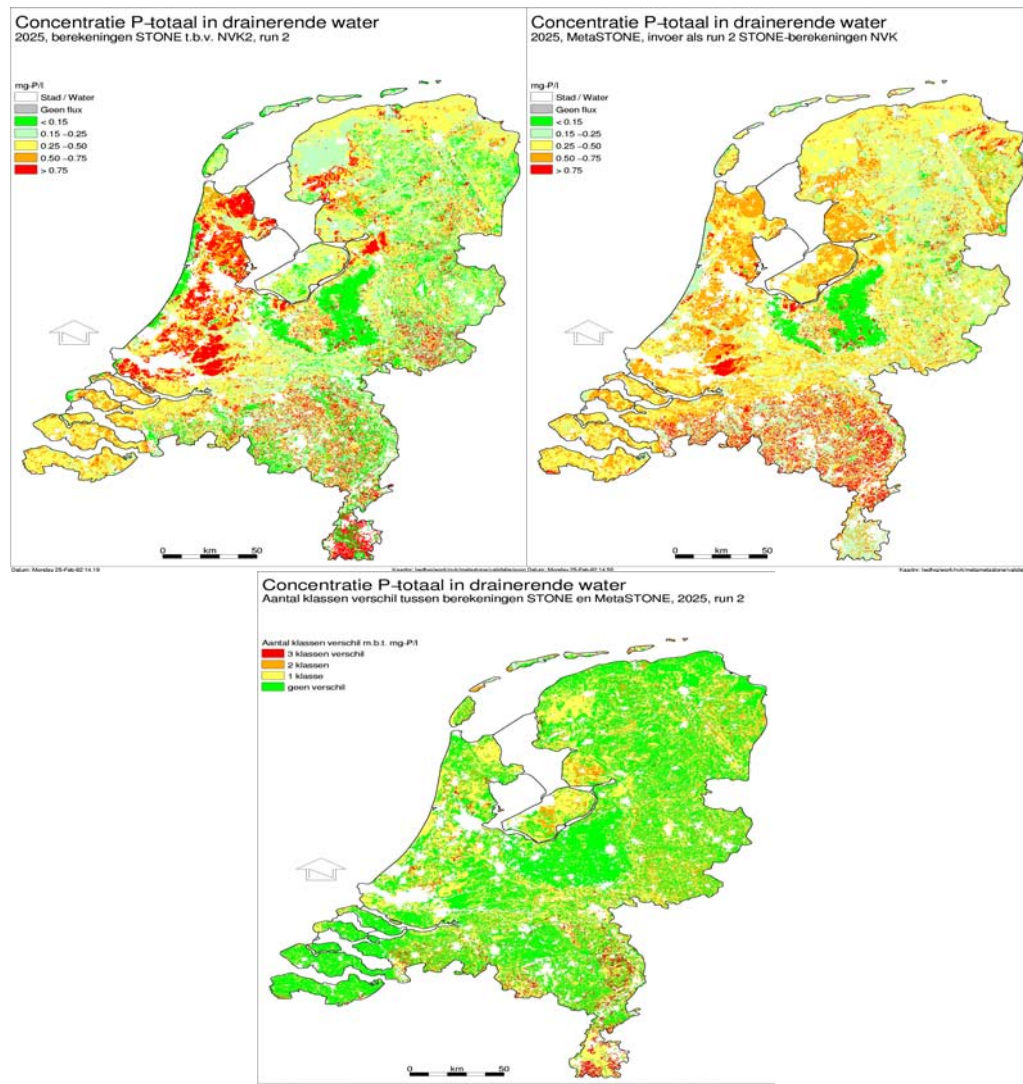
In dat laatste kader verwijzen we ook naar een alternatieve metamodelleringsstudie voor Stone die werd uitgevoerd in het kader van een Integraal Waterbeheerstudie van de STOWA (Stichting Toegepast Onderzoek Waterbeheer). Het betrof hier een toepassing op een meer 'locale' schaal (afwateringsgebieden) dan bij de NVK 2002. Binnen de STOWA studie (STOWA, 2003) stond het ontwikkelen van een vereenvoudigd rekensysteem ('module Waterkwaliteit') centraal waarmee eenvoudig een globaal beeld van de effecten van waterbeheer en landgebruik op de waterkwaliteit (o.a. uit- en afspoeling van nutriënten) bepaald kon worden. De schaal waarop de module zich richt is die van afwateringsgebieden

⁸ Dit betekent onder andere dat men uit gaat van de veronderstelling dat 'steady state' situaties niet anders zullen zijn, indien er sprake is van andere startcondities (bijv. initiële voorwaarden).

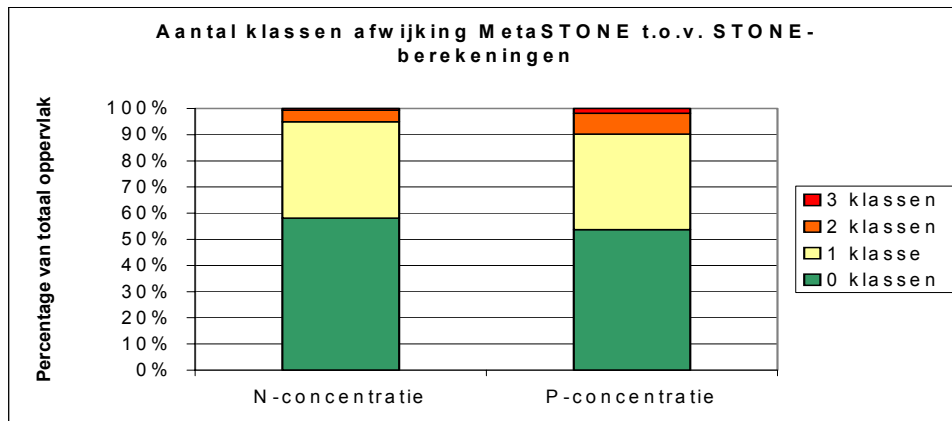
die minstens 2500 ha groot zijn, waarbij lokale gegevens over grondwaterstanden (GLG, GHG) en landgebruik in GIS kunnen worden ingelezen. Daarbij wordt uitgegaan van langjarig gemiddelde omstandigheden en evenwichtssituaties die op termijn ontstaan bij consequente voortzetting van het in 2003 geldende mestbeleid. De langjarige middeling heeft betrekking op periodes van 15 jaar, conform de in Stone gehanteerde cyclus in weerjaren. Als outputs worden de kwartaalwaarden beschouwd van de oppervlaktewaterafvoer en van de N- en P- belasting van oppervlaktewater vanaf de bodem. Deze worden bepaald aan de hand van simulaties van de 6405 verschillende Stone plots over de periode 1993-2038. De laatste 15 jaar (2024-2038) zijn gebruikt voor de bepaling van de langjarige middeling van de kwartaalwaarden. Als invoervariabelen bij de STOWA metamodelleringsstudie is de volgende set gebruikt, waarbij sommige variabelen op kwartaalbasis en andere op jaarbasis beschikbaar zijn:

1. GLG: gemiddeld laagste grondwaterstand op jaarbasis als absoluut getal (m);
2. GHG: gemiddeld hoogste grondwaterstand op jaarbasis als absoluut getal (m);
3. Gewas: Landgebruik op jaarbasis (1=mais, 2=akkerbouw, 3=natuur, 4=gras);
4. Bodem: 21 verschillende bodemfysische eenheden op jaarbasis;
5. Kwel13: Kwel op 13 meter diepte op kwartaalbasis (mm/kwartaal);
6. Wegz13: Wegzijging op 13 meter diepte op kwartaalbasis (mm/kwartaal);
7. Somkwel: Gedefinieerd als Kwel13 - Wegz12 op kwartaalbasis (mm/kwartaal);
8. Nckwel13: N-concentratie in de kwel op 13 meter diepte op kwartaalbasis (mg/l);
9. Pckwel13: P-concentratie in de kwel op 13 meter diepte op kwartaalbasis (mg/l);
10. SomP: Som P 0 - 1 meter op jaarbasis (kg/ha/jr).

De meeste invoervariabelen, behoudens gewas en bodem, nemen waarden aan op een continue schaal. Voor elk van de outputs, gedifferentieerd naar kwartaal, werd een meta-model bepaald via polynoomregressie naar de invoervariabelen, gebruikmakend van de input-output dataset van de STONE berekeningen van (maximaal) 6405 plots. Hierbij is onderzocht in hoeverre het gebruik van transformaties (log-schaal) op outputs en het meenemen van kwadratische en kruis effecten van inputs tot verbetering van de fit kan leiden. Een uitgebreide samenvatting van de verkregen resultaten staat in bijlage 1 van (STOWA, 2003) vermeld.



Figuur 3: Verschillen tussen STONE en MetaSTONE in concentratie P-totaal in het drainerende water bij de invoer van extra STONE-run 2



Figuur 4: Samenvatting resultaten validatie MetaSTONE

4.2 Metamodellering voor atmosferisch model OPS

Het Operationeel Prioritaire Stoffen model OPS is een statisch model voor de berekening van atmosferisch transport van onder andere verzurende stoffen, als NH₃, NO_y en SO₂. Het model wordt bijvoorbeeld vaak toegepast voor het bepalen van jaargemiddelde depositiecijfers, op basis van emissie- en meteogegevens. Het model berekent zowel gemiddelde concentraties in lucht als deposities vanuit de atmosfeer. Zie MODCAT (<http://modcat/>) en Van Jaarsveld en De Leeuw (1993); Van Jaarsveld (1995) voor een uitgebreidere beschrijving van het model.

OPS is lineair met betrekking tot de emissies, onder voorbehoud van gelijke uitstoothoogte en warmte inhoud van de emissiebron. Daarom kan eenmalig een *Source-Receptor-Matrix* (SRM) worden opgesteld, die per gridcel een relatie vastlegt tussen een (eenheids)emissie in punt/gridcel *i* en de resulterende concentratie/depositie in punt/gridcel *j*. De berekening van concentratie/depositie voor diverse emissies komt dan neer op een matrix/vectorvermenigvuldiging, hetgeen een aanzienlijke besparing in rekentijd oplevert (Vreeswijk, Aben et al., 2000). In formulevorm: Bij gegeven emissies worden de concentraties/deposities gegeven door: $O = E * M$ waarbij: *O* is de vector met de output (concentraties, deposities) in de punten *j*, *E* is de vector met de emissies in de punten *i*, *M* is de matrix met overdrachtscoëfficiënten (SRM). Voor het berekenen van SRM's is de zogenaamde SOREMA infrastructuur opgezet (Vreeswijk, Aben et al., 2000). Hierbij is het van belang dat een SRM afhangt van de warmte-inhoud en de uitstoothoogte van de emissies. Vaak worden er daarom voor een specifieke stof meerdere SRM's bepaald. Tevens wordt bij de bepaling van een SRM uitgegaan van specifieke meteocondities, in de praktijk is dat meestal een langjarig gemiddelde.

Een andere manier om van de lineariteit van OPS gebruik te maken is de zogenaamde SIGMA-methode. Hierbij worden van tevoren een groot aantal OPS basisberekeningen gedaan, waarmee kaartbeelden van concentratie en depositievelden worden geconstrueerd bij bepaalde basisemissieconfiguraties voor bijvoorbeeld specifieke doelgroepen en land(clusters). Omdat de lineariteit alleen geldt bij gelijke meteorologische condities en verontreinigingstoestand worden de berekeningen uitgevoerd op basis van lange termijn gemiddelde meteo. De concentratie/depositievelden voor nieuwe situaties met andere emissieconfiguraties per land/doelgroep combinatie worden berekend door deze basis OPS-resultaten op een bepaalde manier te schalen met de verandering in de emissies voor de betreffende combinaties. In formulevorm: $O_n = E_n/E_o * O_o$ waarbij: *O_n* = nieuwe output, *O_o* = oude output, *E_n* = nieuwe emissies, *E_o* = oude emissies. De geschaalde kaarten worden vervolgens gesommeerd (vandaar de naam SIGMA). Voorwaarde voor het toepassen van de SIGMA-methode is dat *ruimtelijke* verdeling van de emissies per land/doelgroep-combinatie niet is veranderd. Ook andere broneigenschappen, zoals uitworphoogte en warmteinhoud worden constant verondersteld. Dit beperkt natuurlijk het doorrekenen van toekomstige situaties waarin de emissiekarakteristieken wijzigen, bijvoorbeeld ongelijkmatige veranderingen in toekomstige verkeersintensiteiten, verschuivingen in industriële activiteiten.

In feite kan zowel SRM als SIGMA gezien worden als een soort impulsresponsie functie methode, maar dan zonder dynamische component (OPS is een statisch model). De SRM aanpak is zeer geschikt voor terugrekenen of optimalisatie. Dit is onder andere toegepast bij het bepalen van optimale NH₃ emissievelden (landbouw), met het doel om natuur zo goed mogelijk te beschermen tegen gevolgen van zure depositie (Dam, Heuberger et al., 2001).

4.3 Metamodel voor atmosferisch model EUROS⁹

Mensink, Delobbe et al. (2001) beschrijven een metamodel voor het Europese smogmodel EUROS, ten behoeve van een toepassing voor België waarbij de lange-termijn effecten van Europese emissie reducties rond troposferische ozon worden geëvalueerd¹⁰. De focus ligt hierbij met name op ozon-gerelateerde blootstellingsindicatoren, en in analogie met Heyes, Schöpp et al. (1995), wordt het volgende (kwadratische) 'regressie'-model gebruikt om een vijftal ozon-gerelateerde blootstellings-indicatoren te berekenen aan de hand van NO_x en VOC-emissies:

$$IND_{l,s,j} = k_{l,j} + \sum_i [a_{l,ij}(E_{VOC})_{s,i} + b_{l,ij}(E_{NOx})_{s,i} + c_{l,ij}(E_{NOx})_{s,i}^2 + d_{l,ij}(E_{NOx})_{s,i}(E_{VOC})_{s,i}],$$

met:

- l*: index voor indicator, $l = 1, \dots, 5$
- s*: index voor scenario, $s = 1, \dots, 5$, (base case, -30% NO_x, -30% VOC, -30% (NO_x & VOC), -50% (NO_x & VOC))
- i*: index voor land, $i = 1, \dots, 37$ (landen in Europa)
- j*: index voor receptor-gridcel, $j = 1, \dots, 15$ (15 cellen in België)
- IND*: indicator (AOT40¹¹ (*forest, crop*), AOT60, gemiddelde dagmaximum, aantal dagen [O₃] > 60 ppb)
- E_{NOx}*: NO_x-emissie
- E_{VOC}*: VOC-emissie
- k*: (onbekende) achtergrond-bijdrage aan indicator
- a, b, c, d*: onbekende coëfficiënten.

Hierbij is, in vergelijking met Heyes, Schöpp et al. (1995), de bijdrage van de natuurlijke (biogene) emissie tot de 15 gridcellen in België verwaarloosd. Daarnaast wordt deze formule hier direct voor de indicatoren gebruikt, terwijl Heyes, Schöpp et al. (1995) dit alleen voor jaargemiddelde ozonconcentraties doen. De resultaten die voor de metamodel constructie gebruikt worden hebben betrekking op de zomer-periode van 1997 (1 mei tot 30 september) en het resulterend metamodel wordt OZON97 genoemd.

Vanwege de relatief hoge rekenkosten met EUROS¹² is niet gekozen voor het uitvoeren van alle benodigde modelruns om de onbekende coëfficiënten *a, b, c, d, k* te bepalen, maar wordt er gebruik gemaakt van source-receptor matrices, berekend met behulp van het EMEP-model (EMEP, 1998). In deze SRM's wordt de verandering van ozonconcentraties in iedere gridcel gegeven t.g.v. een verandering in NO_x- of VOC emissies in een emitterend land. Er wordt verondersteld dat de onderlinge verhoudingen tussen landen hetzelfde blijven als bij de methode van Heyes, Schöpp et al. (1995), en bij de EMEP SRM's. Als deze veronderstelling gerechtvaardigd is, dan is het voldoende om voor één land de coëfficiënten *a, b, c, d, k* uit te rekenen; voor de andere landen kunnen de coëfficiënten dan uit de onderlinge verhoudingen van een EMEP-SRM berekend worden. Voor iedere indicator ($l=1, \dots, 5$) en iedere gridcel ($j=1, \dots, 15$), kunnen nu de 5 onbekende coëfficiënten *a, b, c, d, k* opgelost worden uit 5 vergelijkingen (dat wil zeggen voor scenario $s=1, \dots, 5$) met bekende emissies en berekende indicatoren (berekend uit slechts 5 EUROS-runs).

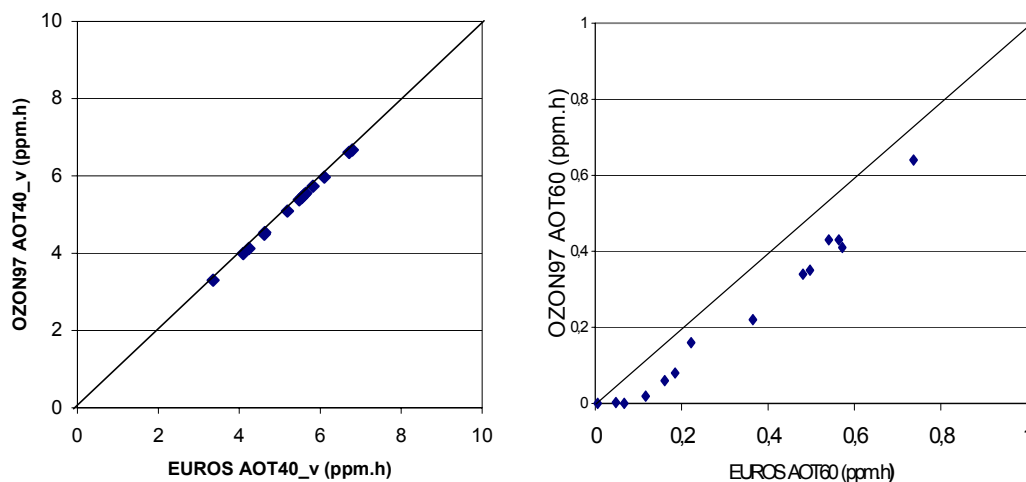
⁹ Deze informatie werd aangeleverd door Ferd Sauter.

¹⁰ Zie ook: <http://www.vito.be/milieu/milieustudies9e.htm>

¹¹ AOT40 is de geaccumuleerde blootstelling aan ozon boven een drempelwaarde van 40 ppb, gebaseerd op uurwaarden voor de periode onder beschouwing (1 mei-30 september, 1997).

¹² EUROS heeft een hogere resolutie in het modelgebied (België) dan het EMEP-model (stand van zaken in 2000). Bovendien wordt er een uitbreiding gehanteerd van jaargemiddelde ozonconcentraties naar verschillende ozon-indicatoren.

Mensink, Delobbe et al. (2001) gebruiken het NEC-scenario (National Emission Ceilings, Gothenburg protocol, (UN/ECE, 2000)) om het metamodel te valideren en zij concluderen onder andere dat voor AOT40 het metamodel goed overeenkomt met het originele EUROS-model, maar dat AOT60 erg onderschat wordt (zie Figuur 5). Dit heeft te maken met het feit dat de ozon-concentraties vaak onder de drempelwaarde van AOT60 liggen of er net boven. Dit impliceert een grote gevoeligheid van AOT60 voor kleine afwijkingen in ozon-concentraties.



Figuur 5: Vergelijking tussen de resultaten van Euros en het OZON97-metamodel voor de AOT40 en AOT60 waarden onder het NEC-scenario 2010. Bron: Mensink, Delobbe et al., 2001.

Bij de uitgevoerde analyse zijn een aantal *kanttekeningen* te plaatsen:

- Is de veronderstelling over het gebruik van EMEP-SRM's gerechtvaardigd? Wat is bijvoorbeeld de invloed van verschillende meteo-jaren op deze veronderstelling (EMEP gebruikt 5 jaar gemiddelde meteo)?
- Hoe sterk verschilt het NEC-scenario van de scenario's die gebruikt zijn voor de berekening van de coëfficiënten? De NEC 2010 gemiddelde emissiereductie over Europa bedraagt voor NO_x 41% en VOC 40% (ten opzichte van 1990). Voor België betreft dit NO_x 47%, VOC 56%, terwijl het gemiddelde van landen rond België (B,F,NL,D,UK) 54% (NO_x) en 61% (VOC) bedraagt. Voor een goede validatie zijn in feite meer scenario's nodig.
- Tot hoever gaat de geldigheid van het metamodel voor wat betreft de omvang van de emissiereducties?
- Wat is de gevoeligheid van het metamodel voor verschillende NO_x/VOC-ratio's?
- Heyes, Schöpp et al. (1995) stellen een indirecte methode voor ter berekening van AOT40-waarden: eerst jaargemiddelde ozon berekenen met een metamodel, dan lokale AOT40 parametriseren met behulp van hoogte, afstand tot de kust, lokale NO_x-emissie en andere parameters. Hoe verhoudt zich deze (nog niet uitgewerkte) methode tot de directe methode van Mensink, Delobbe et al. (2001)?

4.4 Metamodellering voor IMAGE

In de jaren negentig werden binnen het RIVM twee integrated assessment modellen voor internationaal milieuonderzoek ontwikkeld, te weten IMAGE 2 en TARGETS, die verschillende aspecten van mondiale milieuverandering simuleren. In 1997 ging een project van start waarbij harmonisatie van het TARGETS mondiale kringlopen model (CYCLES) met het IMAGE 2 Terrestrial Environment System (TES) en het Atmosphere-Ocean System (AOS) nagestreefd werd.

Uit ervaringen met beleidsmakers tijdens een reeks Delft Science Policy Dialogue sessies bleek echter al snel dat er met name behoefte bestond aan een modelinstrumentarium waarmee snel en op interactieve wijze lange termijn emissiescenario's en hun effecten op klimaatverandering doorgerekend konden worden. Dit leidde tot een nieuwe doelstelling van het harmonisatieproject, en besloten werd om op basis van het CYCLES model een soort *meta-IMAGE* te construeren, dat soortgelijke uitkomsten zou moeten geven als het IMAGE 2.1-model. Dit laatste model simuleert de complexe lange-termijn dynamiek van het biosfeer-klimaatstelsel op geografisch expliciete wijze ($0.5^{\circ} \times 0.5^{\circ}$ latitude-longitude grid), en had anno 1997 een runtijd van 12 uur op een Pentium PC, terwijl CYCLES zelf 40 seconden runtijd vergde (zie Den Elzen, 1998).

Meta-IMAGE zou een runtijd van hooguit enkele seconden dienen te hebben, en zou een aantal relevante klimaatindicatoren op *globale* schaal moeten kunnen berekenen, te weten de globale broeikasgasconcentraties en de globale gemiddelde temperatuur- en zeespiegelstijging. Afwijkingen met het referentiemodel IMAGE 2.1 op datzelfde globale aggregatieniveau dienden gering te zijn.

Om meta-IMAGE te bouwen werd gekozen voor aanpassing en herparametrisatie van het CYCLES model, waarbij (zie Figuur 6; Den Elzen, 1998):

- (1) de stikstof cyclus en enkele voor CYCLES specifieke terrestrische koolstofcyclus processen zoals bodem erosie werden *gedeactiveerd*¹³;
- (2) belangrijke modelparameters (onder andere refererend naar de terrestrische koolstof cyclus en naar klimaatgevoeligheid) *opnieuw (manueel) gecalibreerd* werden op basis van IMAGE 2.1 runs (zie § 3.2 in (Den Elzen 1998));
- (3) een verregaande *aggregatie* van de ruimtelijke schaal (w.b. terrestrische biosfeer) plaatsvond door slechts een viertal landgebruikstypes te onderscheiden: bossen, graslanden, landbouw en overigen, voor zowel de ontwikkelingslanden als ook voor de geïndustrialiseerde landen;
- (4) het oceaan submodel werd vervangen door een *vereenvoudigde impuls responsie representatie* (inclusief convolutie integraal) van het oceaan koolstofcyclus ('general circulation model' van Maier-Reimer en Hasselmann (1987), met aangepaste tijdsconstanten; zie ook Wigley, 1992).

Ter validatie werden de meta-IMAGE resultaten vergeleken met IMAGE 2.1 baseline scenario berekeningen. De resultaten lieten doorgaans slechts geringe afwijkingen tussen beide modellen zien (relatieve verschillen in de orde van enkele procenten; zie § 3.3 in Den Elzen (1998); zie ook Figuur 7).

De run-tijd van meta-IMAGE 2.1 lag in de orde-grootte van seconden. Meta-IMAGE 2.1 is als basis component gebruikt in het succesvolle interactieve beleidsondersteunende rekentool

¹³ Enkel de CO₂ fertilisatie feedback mode (CO2F-mode) en temperatuur feedbacks zijn meegenomen; de feedback effecten van nutriënten op CO₂ en N werden buitengesloten.

FAIR¹⁴, dat internationale beleidsmakers helpt bij het afwegen van mogelijkheden tot lastenverdeling rond de beoogde wereldwijde emissie reductie van broeikasgassen (zie o.a. Den Elzen, Berk et al., 1999; Den Elzen, 2000; Den Elzen, Schaeffer et al., 2002).

Verdere ontwikkelingen rond IMAGE en FAIR (Den Elzen, Schaeffer et al., 2002; Den Elzen en Lucas, 2003) hebben onder andere geleid tot vervanging van meta-IMAGE 2.1 door IMAGE-AOS (meta-IMAGE 2.2), terwijl de functionaliteit van FAIR verder uitgebreid werd met kosten-berekeningen die een kosten-evaluatie van beoogde beleidsmaatregelen mogelijk maakt (zie voor uitgebreide informatie rond FAIR de URL: <http://www.rivm.nl/fair/>).

Opmerking

Ook in het internationale klimaatonderzoek buiten het MNP is metamodellering op diverse wijzen en in diverse gedaanten toegepast:

- *Impuls-response modellering* is gebruikt in situaties waarin het gedrag van complexe modellen bij benadering lineair is. Doorgaans vereist dit dat er slechts sprake is van geringe afwijkingen rond een evenwichtoplossing of een nominale trajectorie, zodat men in het quasi-lineair werkgebied van het complexe model blijft, bijvoorbeeld.

$$C_a(t) = \int_{t_0}^t r_a(t-\tau) \cdot e(\tau) \cdot d\tau + C_a(t_0)$$

waarbij de convolutie integraal uitdrukt hoe de emissie-historie $e(\tau)$ over het traject van t_0 tot t , de huidige concentratie $C_a(t)$ bepaalt, vanuit de beginconditie $C_a(t_0)$; de impuls responsie $r_a(t)$ is hierbij bepaald door met het oorspronkelijke model na te gaan hoe een initiële koolstof emissie-input op tijdstip 0 de complexe koolstof cyclus bepaalt. Doorgaans is de impuls-responsie te schrijven als som van een aantal exponentiële termen:

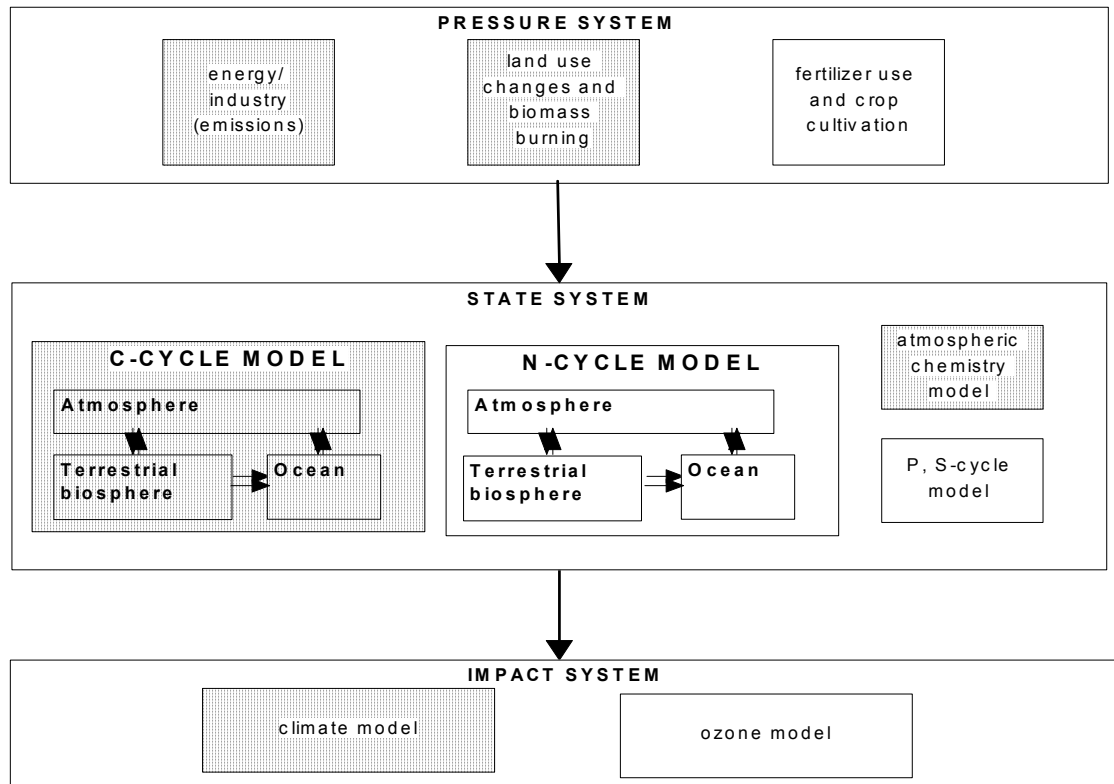
$$r_a(t) = \sum_{i=1}^K b_i \cdot e^{-t/\tau_i}$$

die de hoofd-dynamische componenten uitdrukken, met bijbehorende tijdconstanten (zie bijvoorbeeld Maier-Reimer en Hasselmann, 1987; Wigley, 1992).

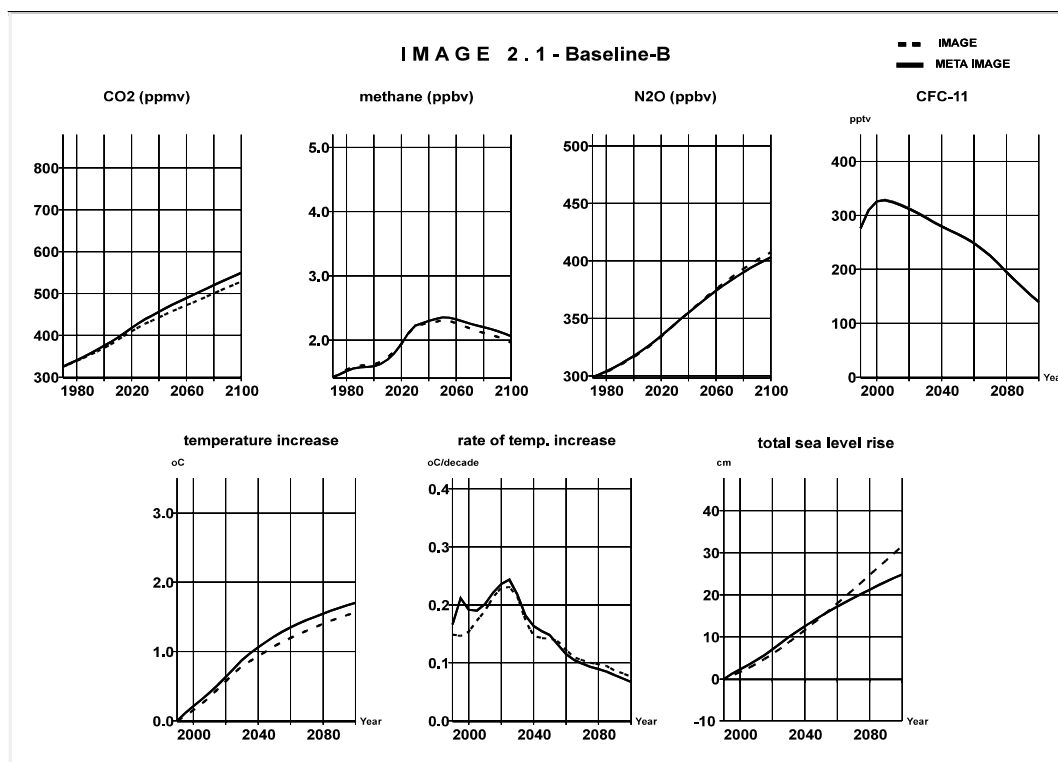
- Deze representatie kan ook worden gezien als de oplossing van een equivalent stelsel gewone differentiaalvergelijkingen, die uit een cascade van interacterende lagen is opgebouwd (*'box-model'*). Dit box-model wordt dan door een geschikte keuze van laagdikte en diffusiecoëfficiënt zo gecalibreerd dat - in de lineaire limietsituatie bij een kleine CO₂ emissie-impuls - de impuls-responsie functies (IRF's) van de 3D-koolstofcyclus gereproduceerd worden. De lagen zijn gekoppeld door koolstofluxen die op hun beurt worden aangedreven door de concentratieverschillen tussen de lagen (zie bijvoorbeeld de appendix in Hooss, Voss et al., 2001). Deze box-model representatie kan als basis gebruikt worden om *niet-lineariteiten* in bijvoorbeeld de dynamische atmosfeer-oceaan interactie expliciet mee te nemen in impuls responsie beschrijving via koppeling van deze laatste met separate vergelijkingen die de niet-lineaire aspecten beschrijven (zie Joos, Bruno et al., 1996).
- Hooss, Voss et al. (2001) beschouwen ook uitbreidingen waarbij expliciet ruimtelijke signaalcomponenten worden meegenomen, via een Empirical Orthogonal Functions (EOFs)-representatie. De coëfficiënten in deze representatie (de zogenaamde 'principale componenten') evolueren in de tijd, en hun evolutie wordt beschreven aan de hand van impuls responsie formulering, toegepast op getransformeerde variabelen (logaritmische relaties). Deze beschrijvingswijze is onder andere binnen het ICLIPS model gebruikt om

¹⁴ FAIR: Framework to Assess International Regimes for burden sharing, dat verder nog werd uitgebreid met andere vereenvoudigde klimaatmodellen (revised Brazilian model, koolstof cyclus model MAGICC en het Bern koolstof cyclus model).

- uitvoerbare ('viable') optimale paden door te rekenen ('tolerable windows' benadering; (zie Bruckner, Hooss et al., 2003)).
- Verwant hiermee zijn ook de optimale 'fingerprints' technieken, waarbij signaal/patroon-analyse plaatsvindt van klimaatveranderingsgegevens om het effect van antropogene invloed (c.q. de natuurlijke variatie) af te leiden en toe te wijzen (detectie en attributie van klimaatverandering). Dit gebeurt veelal via gebruik van multiële regressie op basis van decompositie van signaal in lineaire combinatie van componenten (bijvoorbeeld EOF's), met als doel het scheiden van ruis en signaal en het onderscheiden van een aantal onafhankelijke dimensies in het klimaatveranderings signaal. Dit kan worden toegepast op de uitkomsten van klimaatveranderingsmodellen (zie bijvoorbeeld Hasselmann, 1993; Von Storch en Navarra, 1995; Zwiers en IDAG, 2005).
 - Knutti, Stocker et al. (2003) gebruiken metamodellering op basis van neurale netwerken voor het verrichten van probabilistische analyses rond klimaatverandering. Deze methode vereist doorgaans een groot aantal runs. Knutti, Stocker et al. (2003) gebruiken bijvoorbeeld een 500 tal simulaties om een neuraal netwerk met 10 neuronen adequaat te schatten.
 - Webster en Sokolov (2000) propageren het gebruik van een zogenaamde 'orthogonale collocatie methode' (DEMM: Deterministic Equivalent Modeling Method; zie ook Webster, Forest et al., 2003) om de modeluitkomsten te decomponeren in signalen, die gebruikt kunnen worden voor het uitvoeren van onzekerheidsanalyses. Hiermee kan het aantal runs dat moet worden uitgevoerd om een adequaat beeld te krijgen van gemiddelde, variantie en percentielen van de verdeling aanzienlijk ingeperkt worden. De methodiek wordt gebruikt in combinatie met gevoeligheidsanalyse als een soort pre-screening stap om parameters te selecteren die er toe doen. Continuïteit in de responsies is vereist; discontinue toestands- overgangen worden slecht beschreven door DEMM.



Figuur 6: Het Pressure-State-Impact-Response systeemdiagram van het CYCLES model. De gearceerde onderdelen zijn meegenomen in het meta-IMAGE 2.1 model (bron: Den Elzen, 1998).



Figuur 7: De concentraties van CO_2 , CH_4 , N_2O and $CFC-11$, en de temperatuur- en zeespiegelstijging, voor het Baseline B scenario van IMAGE 2.1 en meta-IMAGE 2.1 (bron: Den Elzen, 1998)

4.5 Metamodellering voor PCLake

PCLake is een waterkwaliteitsmodel, dat onder andere gebruikt kan worden om het effect van voorgestelde beleids- en beheersmaatregelen op de waterkwaliteit van meren te berekenen. De grote gedetailleerdheid van het model leidt tot relatief lange rekestijden, hetgeen een bezwaar vormt als het model ingezet wordt in scenariostudies met een groot aantal simulaties. In Vleeshouwers, Janse et al. (2004) wordt een studie beschreven naar het opstellen van een meta-model PCLake, dat – bij benadering – dezelfde resultaten genereert, maar in een aanzienlijk kortere tijd. De operationele doelstelling van deze studie was om een geschikt meta-model te vinden voor de beschrijving van het effect van acht belangrijke omgevingsvariabelen (input) - waaronder belangrijke beheersvariabelen als fosfaatbelasting, doorstroomtijd en percentage moeraszone - op het chlorofylgehalte van het meer in de zomer in een steady-state situatie (output). Hierbij werden drie verschillende metamodelleringstechnieken toegepast: (1) regressieboom, (2) ‘radial basis function network’ en (3) interpolatie via meerdimensionale ‘look-up tables’. In een vergelijking tussen de drie methoden bleek dat interpolatie de meest nauwkeurige benadering gaf van de door PCLake gesimuleerde waarden. De overeenkomst tussen PCLake en het meest nauwkeurige meta-model, toegepast op een aselechte steekproef van 80000 punten uit de inputruimte, leverde een R^2_{fit} ¹⁵ van 0.965. De rekestijd van alle metamodellen varieerde van 1 – 2 milliseconden per

¹⁵ De R^2_{fit} is niet hetzelfde als de klassieke R^2 , in die zin dat het uitdrukt in hoeverre de uitkomsten van het originele model en de benaderende modellen voldoen aan een 1:1 verhouding.

berekening, tegen 9 seconden voor PCLake. Omdat het metamodel is opgesteld voor een breed bereik van invoervariabelen stellen de auteurs dat het bij integrale scenariostudies goed kan dienen om, ter vervanging van PCLake, voorspellingen te doen voor de waterkwaliteit in toekomstige situaties. Het metamodel gaat echter niet in op dynamische aspecten zoals de reactietijd. De methode kan echter wel worden aangepast voor andere uitvoervariabelen of indicatoren of voor vernieuwde modelversies. Dit vergt uiteraard een operationeel houden van het oorspronkelijke model.

In meer detail ingaand op de toegepaste metamodeleringsmethoden, merken we op dat bij de constructie van de meerdimensionale 'look-up tables' de invoer data zijn vergrid, waarbij de gridkeuze is gebaseerd op een gevoeligheidsanalyse van het PCLake model. Het aantal waarden per grid-as is omgekeerd evenredig gekozen aan de gevoeligheids-index van de corresponderende (invoer) parameter. Met andere woorden, hoe gevoeliger een model is voor veranderingen van een bepaalde parameter, des te fijner de gridverdeling voor die parameter. Voor gevoeligheidsanalyse van PCLake is gebruik gemaakt van de FAST methode uit het SIMLAB software pakket (Saltelli, Tarantola et al., 2004).

Voor het 'radial basis function network' is bij de keuze van de gehanteerde gaussische basisfuncties gebruik gemaakt van de resultaten van een regressieboom-analyse: het aantal, de radius (r) en het middelpunt (c) van de gaussische basisfuncties $\exp(-(x-c)^2/r^2)$ zijn bepaald op basis van de posities en afmetingen van de 'hyper-rechthoeken' in de partitie van de inputruimte die resulteerde uit de regressieboomanalyse, en die correspondeert met de eindbladeren in de regressieboom. Hiermee wordt een initiële verzameling van kandidaatnetwerken verkregen, waaruit het beste netwerkmodel wordt gekozen op basis van het BIC (Bayesian Information Criterion). Zie Orr, 1999; Orr, 1999 voor details.

Er zijn met elk van de drie methoden twee metamodellen bepaald, een op basis van 8 variabelen en een op basis van 5 variabelen (de belangrijkste 5 van de 8)¹⁶. De modellen zijn vervolgens vergeleken op basis van een onafhankelijke *testset* van 80000 simulatieruns met PCLake. Bij deze set zijn wel alle 8 variabelen gevarieerd. De kwaliteit (nauwkeurigheid) van de metamodellen wordt bepaald via de R^2_{fit} . Opvallend is daarbij dat de modellen die op 5 variabelen zijn gebaseerd het in *alle* gevallen beter doen dan de modellen die gebaseerd zijn op 8 variabelen (zie Tabel 3). Dit lijkt erop te wijzen dat er sprake is van *overparametrisatie* bij metamodellen gebaseerd op 8 variabelen: de extra vrijheidsgraden worden a.h.w. gebruikt om 'ruis' in de schattingsdataset mee te modelleren hetgeen op de onafhankelijke testdataset niet tot een betere fit leidt. Andere oorzaken kunnen zijn de beperkte kwaliteit van de data gebruikt voor metamodelering, en de beperkte kwaliteit van de testset. Verder dient te worden opgemerkt dat de berekende waarden van de R^2_{fit} sterk afhankelijk zijn van de range van outputwaarden. Het is de vraag in hoeverre het beeld zal veranderen indien de modellen vergeleken zouden worden op een beperktere output-range.

¹⁶ Hiermee wordt aangegeven dat de 8 (5) belangrijkste omgevingsvariabelen zijn gevarieerd, terwijl alle overige invoervariabelen constant gehouden zijn.

Tabel 3: Evaluatie van de verschillende metamodellen op basis van een onafhankelijke test-set van 80.000 runs. Weergegeven zijn de nauwkeurigheid, in termen van R^2_{fit} , en de modellerings-efficiency, in termen van de tijd die nodig is om een betreffende dataset met het referentiemodel (PCLAKE) te genereren, het meta-model te maken en te gebruiken. In de kolom techniek staat tussen haakjes het aantal invoervariabelen vermeld dat bij het meta-model betrokken is. (Bron: Vleeshouwers, Janse et al., 2004).

Techniek	R^2_{fit}	Tijd voor PCLAKE runs (in uren)	Tijd voor meta-modelconstructie (in uren)	Tijd voor 80.000 testruns (in uren)
Regressieboom (8)	0.884	186.0	0.05	0.02
(5)	0.911	218.2	0.05	0.02
RBF netwerk (8)	0.922	12.0	152.6	0.04
(5)	0.935	17.5	600	0.01
Interpolatie (8)	0.885	229.4	0.02	0.02
(5)	0.965	208.7	0.02	0.02

Uit de resultaten uit de tabel blijkt verder dat er zeer grote verschillen zijn in de tijd die nodig is om een schattingsdataset met PCLAKE te genereren, en de tijd die vervolgens nodig is om het meta-model te construeren op basis van deze dataset. Zo werden er voor de RBF netwerken ‘slechts’ 5000 tot 7000 PCLAKE runs gebruikt, terwijl voor de regressieboom analyse en de ‘look-up table’-constructie zo’n 70.000 tot 90.000 runs nodig waren. Het construeren van de metamodellen bleek daarentegen voor RBF netwerken een zeer tijdsinnende bezigheid te zijn, wat enerzijds veroorzaakt werd door de gebruikte software (Orr, 1999), maar anderzijds ook een gevolg is van de werkwijze om een groot aantal kandidaat-netwerken te evalueren. Vervolgonderzoek zou zich kunnen richten op de vraag in hoeverre de door Vleeshouwers, Janse et al. (2004) gepresenteerde resultaten maatgevend zijn voor het gebruik van de betreffende methoden, of dat ze voor een deel ook het gevolg zijn van de gehanteerde aanpak en implementatie.

4.6 Metamodellen van GeoPEARL

Het model GeoPEARL (Tiktak, De Nie et al., 2002) beschrijft het gedrag van bestrijdingsmiddelen in de bodem op de nationale schaal. GeoPEARL wordt gebruikt bij de evaluatie van beleidsplannen op het gebied van bestrijdingsmiddelen en bij de toelating van nieuwe bestrijdingsmiddelen. De kern van GeoPEARL is het PEARL model (Tiktak, Van den Berg et al., 2000). PEARL wordt gedraaid voor een groot aantal unieke combinaties van bodemtype, landgebruik en klimaatdistrict. Deze unieke combinaties worden ook wel plots genoemd. Het model geeft onder andere de water- en stofbalansen en percentielen van de concentratie van bestrijdingsmiddelen in uitspoelend water bij specifieke bestrijdingsmiddelen-toepassings-scenario’s. Het model kan gebruikt worden voor een groot aantal stoffen, waaronder vluchtige stoffen. Naast een nationale versie van GeoPEARL bestaat er ook een Europese versie (Tiktak, De Nie et al., 2004).

PEARL gebruikt een numerieke methode om de belangrijkste vergelijkingen op te lossen. Door de grote dynamische range van de processen in het bodemsysteem is het bovendien nodig om met kleine tijdstappen te rekenen. Hierdoor is de rekentijd van het model aanzienlijk; anno 2004 duurde een GeoPEARL run op nationale schaal ongeveer 24 uur. Hierbij dient wel te worden aangetekend dat nieuwe technologieën, zoals grid-computing, de rekentijd sindsdien aanzienlijk teruggebracht hebben. Desalniettemin bestaat er voor een aantal toepassingen behoefte aan een vereenvoudigd uitspoelingsmodel. Deze toepassingen betreffen Pan-Europese modelstudies en het (snel) doorrekenen van beleidsopties.

Bij het afleiden van een vereenvoudigd uitspoelingsmodel moet er rekening mee worden gehouden dat het gedrag van bestrijdingsmiddelen in de bodem een sterk niet lineair karakter heeft. Bovendien gaat het om verschillende middelen, die allemaal een ander gedrag vertonen. Het leeuwendeel van de klassieke bestrijdingsmiddelen wordt bijvoorbeeld vooral door organische stof gebonden, maar voor een aantal probleemstoffen geldt dit juist niet.

In het kader van het APECOP project (Vanclooster, Armstrong et al., 2003) zijn verschillende methoden voor modelvereenvoudiging getest. In eerste instantie is hierbij gekeken naar vereenvoudigde modellen die een fysische basis hebben. Deze modellen beschrijven de belangrijkste vergelijkingen meestal met een analytische in plaats van een numerieke oplossing. Dergelijke modellen veronderstellen een homogene bodem; onderscheid in bodemlagen is dus niet mogelijk. Ook wordt de neerslag constant in de tijd verondersteld. Uit modelanalyses (Tiktak, Boesten et al., 2002) is gebleken dat door deze aannames de uitspoeling met een factor van ruim 100 wordt onderschat. Bovendien bleek er een bijzonder slechte correlatie te zijn tussen de uitkomsten van PEARL en die van het analytische model. Dit laatste heeft te maken met het feit dat bodems in werkelijkheid zeer verschillend van opbouw kunnen zijn. Zo kan de organische stof bijvoorbeeld zeer ondiep of juist diep in het profiel aanwezig zijn. Door de slechte correlatie met de uitkomsten van PEARL en GeoPEARL was het niet mogelijk het analytisch model als screening model in te zetten voor het snel doorrekenen van bestrijdingsmiddellenscenario's.

De tweede groep modellen die zijn getoetst zijn *statistische* metamodellen (zie hoofdstuk 3). De database die nodig was voor het bouwen van het metamodel is gebaseerd op modelberekeningen met GeoPEARL. Hiertoe zijn op het MNP gridcomputingsysteem in totaal 100.000 (dat wil zeggen 20x5000) modelberekeningen uitgevoerd met 20 stoffen. Van deze modelberekeningen is de helft gebruikt als trainingsdataset en de andere helft om het metamodel te toetsen. De te voorspellen variabele is de gemiddelde concentratie in het ondiepe grondwater (op 1 meter diepte). Op basis van eerdere gevoeligheidsanalyses (Tiktak, Swartjes et al., 1994) zijn de belangrijkste onafhankelijke variabelen voor het metamodel geselecteerd. Dit zijn het organische stofgehalte van de bodem, het kleigehalte, de pH, de hoeveelheid beschikbaar water in het bodemprofiel, de gemiddelde grondwaterstand, de halfwaardetijd van de stof en de sorptieconstante. In eerste instantie zijn ook klimatologische gegevens beschouwd, maar de variatie binnen Nederland bleek te gering voor een significante bijdrage.

Verschiede soorten metamodellen (van simpel tot gecompliceerd) zijn op bovengenoemde dataset losgelaten. Eenvoudige statistische methoden, bijvoorbeeld multiple regressiemodellen, bleken het complexe modelgedrag niet goed te kunnen beschrijven. Bovendien bevatte de dataset veel modelberekeningen die een concentratie van 0 µg/L opleverde. Dit is met name het geval bij hoge organische stofgehalten. Het uiteindelijk afgeleide metamodel is daarom een hybride model, namelijk een combinatie van een logistische discriminant en een radial basis functions neuraal netwerk (Piñeros Garcet, Van der Linden et al., 2004). Het eerste model filtert onder andere op basis van het organische stofgehalte de concentraties van 0 µg/L uit de dataset. Het neuraal netwerk wordt op de rest van de dataset toegepast met alle bovengenoemde verklarende variabelen. Toepassing van het hybride metamodel op de validatieset leerde dat het hybride model een goede 'modelling efficiency' (EF) heeft (0.95), maar dat de Root Mean Square Error (RMSE) vrij hoog is (1.0602) (zoals gedefinieerd in Vanclooster, Boesten et al., 2000). Dit laatste wordt veroorzaakt doordat slechts 15% van de waarnemingen een echt hoge concentratie oplevert, waardoor met name bij heel lage concentraties (< 0.01 µg/L) grote afwijkingen ontstonden. De vraag rijst of met het metamodel éénduidig onderscheid gemaakt kan worden tussen gebieden waar de verwachte concentratie boven de normwaarde 0.1 µg/L ligt. Om deze vraag te beantwoorden zijn 90% betrouwbaarheidsintervallen rond de voorspellingen van het metamodel berekend (zie Piñeros Garcet, van der Linden et al., 2004 voor details).

De volgende conclusies konden getrokken worden:

- Als de door het metamodel voorspelde concentratie $< 0.08 \mu\text{g/L}$ is, dan is de door het numerieke model voorspelde concentratie vrijwel zeker kleiner dan $0.1 \mu\text{g/L}$;
- Als de door het metamodel voorspelde concentratie $> 0.17 \mu\text{g/L}$ is, dan is de door het numerieke model berekende concentratie vrijwel zeker groter dan $0.1 \mu\text{g/L}$;
- In het bereik $0.08\text{-}0.17 \mu\text{g/L}$ is niet met zekerheid te zeggen of de concentratie in het oorspronkelijk model boven of onder de norm ligt.

Indien rekening gehouden wordt met deze onzekerheidsmarges kan het metamodel gebruikt worden om onderscheid te maken tussen gebieden waar de concentratie onder, respectievelijk boven de norm ligt. De volgende vragen staan nog open voor vervolgonderzoek:

- Kan het metamodel worden gebruikt voor inverse modellering?
- Kan het metamodel worden ingezet in ruimtelijke optimalisatiestudies?
- Kan het metamodel gebruikt worden voor gevoeligheids- en onzekerheidsanalyse?

4.7 Slotopmerkingen

In het verleden is metamodellering binnen het MNP op een aantal fronten ingezet om tot snellere simulatiemodellen te komen die op een bredere schaal inzetbaar zijn bij beleidsgericht onderzoek. De daarbij gebruikte metamodelleringstechnieken variëren in mate van ingewikkeldheid, waarbij het opvallend is dat er doorgaans weinig aandacht was voor niet-lineariteiten en dat dynamische en ruimtelijke aspecten vaak ontbreken of slechts beperkt zijn ingevuld. Dit is deels te wijten aan het feit dat de oorspronkelijke modellen - waar metamodellen van gemaakt zijn - deze aspecten slechts in beperkte mate uitdrukken en dat deze bovendien van ondergeschikt belang geacht werden voor de beoogde toepassingen. Ook speelt een rol dat in de praktijk vaak pragmatische keuzes gemaakt moesten worden bij (meta)modellering, waarbij de beschikbare tijd en expertise het niet toeliet om meer geavanceerde metamodelleringsmethoden te beproeven. Bovendien werd hierbij vaak de voorkeur gegeven aan methoden die voor een gebruiker eenvoudig, plausibel en interpreteerbaar zijn, en dicht tegen het oorspronkelijk model aan liggen (bijvoorbeeld via gebruik van 'look-up tables'), boven meer geavanceerde methoden die specialistischer en black-box van karakter zijn. Gaandeweg blijkt er echter wel een groeiende aandacht te komen voor meer geavanceerdere metamodelleringsmethoden, die bijvoorbeeld het doorrekenen van diverse scenario's en overbruggen van schaalniveau's mogelijk maken (vergelijk de PCLake en GEOPEARL cases) met name in situaties waar de eenvoudige methoden niet meer goed voldoen bijvoorbeeld ten gevolge van niet-lineariteiten. Ook zijn in een beperkt aantal gevallen verschillende metamodelleringsmethoden met elkaar vergeleken (met name voor PCLake). Onduidelijk is echter in hoeverre de uitkomsten van deze vergelijking een goed beeld geven van de merites van de afzonderlijke metamodelleringsmethoden, of dat ze voor een groot deel bepaald zijn door de manier waarop de methoden geïmplementeerd en gebruikt zijn.

Uitdagingen voor de toekomst liggen met name op gebieden waarin niet-lineariteiten en/of temporele en ruimtelijke dynamiek een grote rol spelen en waarbij metamodellen gewenst zijn (bijvoorbeeld landgebruiksveranderingsmodellering, ecologische impactmodellering). Ook is de inzet van metamodellen bij inverse modellering, ruimtelijke optimalisatie, gevoeligheids- en onzekerheidsanalyse een potentieel relevant en grotendeels braakliggend terrein binnen het MNP. Dit is ook het geval voor situaties waarin er sprake is van gekoppelde modelsystemen, zoals bij integrale analyses: hierbij kan metamodellering van modelonderdelen vereist zijn om een adequate reikwijdte en toepassingsrange van de modelketen te garanderen.

Tenslotte zal ook het aangeven van kwaliteit en toepassingsgebied van het metamodel een blijvend aandachtspunt vormen. Om het draagvlak voor metamodelgebruik te vergroten is helder inzicht en transparantie in metamodelrelaties vereist. Dit speelt vooral indien de gebruikte metamodelleringsstechniek een sterk black-box karakter heeft. Een expliciete gevoeligheids-analyse van het uiteindelijke metamodel kan nuttige informatie leveren om de plausibiliteit van het (black-box) metamodel te vergroten.

5. Metamodelleren: conclusies en vooruitzichten

De resultaten uit de literatuur en de toepassingen binnen het MNP-RIVM maken duidelijk dat aanpak en resultaten van metamodellering sterk toepassingsafhankelijk zijn. Om te komen tot een verantwoorde ontwikkeling en gebruik van metamodellen, is een stappenplan opgesteld. Hierin komt expliciet doel en context voor metamodellering aan bod, evenals de keuze van inputs/outputs, het genereren van data voor metamodelbouw, de keuze van metamodellerings-techniek, de evaluatie van het metamodel en het vaststellen van toepassingsbereik, het gebruik en beheer (inclusief eventueel wenselijke updates) van het metamodel. Meerdere doelen kunnen leiden tot meerdere metamodellen van eenzelfde referentiemodel, en kwaliteit en geschiktheid van de metamodellen zal duidelijk afhangen van de kwaliteit van het referentiemodel (dat wil zeggen een slecht model leidt tot een slecht metamodel: GIGO (Garbage In Garbage Out)). Dit maakt het van belang om de kwaliteit van het referentiemodel, evenals de bijbehorende expertise (kennismanagement), op peil te houden en het metamodel aan te passen zodra dat nodig is. Hierdoor wordt voorkomen dat een metamodel teveel een eigen leven gaat leiden en onvoldoende geactualiseerd wordt. Ook is het helder aangeven van de toepassingsrange van het metamodel cruciaal, om onoordeelkundig gebruik te voorkomen. Daarbij hoort ook het verrichten van een a posteriori gevoeligheidsanalyse van het metamodel om inzicht in relevante metamodelrelaties te verschaffen, ter ondersteuning van transparantie en plausibiliteit.

Organisatorisch en operationeel stelt metamodellering flinke eisen en uitdagingen aan afstemming, organisatie en samenwerking, vooral bij het gezamenlijk opstellen van een metamodel, met name als de expertise verspreid is over diverse gebieden, mensen, onderzoeksgroepen en instituten, zoals vaak het geval is bij integrale studies. Zaken als het vaststellen van toepassingsrange, eindverantwoordelijkheid, beheer, gebruik, interpretatie etcetera zullen hierbij aan de orde moeten komen.

Ook werd duidelijk dat er vele methoden zijn die bij metamodellering kunnen worden ingezet, variërend van procesgebaseerde, tot meer black-box georiënteerde statistische methoden en op dynamisch systemen gerichte methoden. Het zal onder andere afhangen van specifieke probleemeigenschappen (bijvoorbeeld aard van systeem en zijn onderdelen¹⁷, complexiteit en niet-lineariteiten; data-beschikbaarheid; beschikbare tijd, menskracht, expertise, software) hoe haalbaar en succesvol het gebruik van specifieke metamodelleringsmethoden zal zijn. Met name bij niet-lineaire systemen verdient het aanbeveling om meer ervaring op te doen met moderne 'statistical learning' methoden als SVR (Support Vector Regression) en hun performance te vergelijken met RBF/NN (Radial Basis Functions; Neurale Netwerken) en Kriging/GP (Gaussische Processen). Ook is het een uitdaging om bij dynamische systemen explicieter de dynamische aspecten mee te nemen bij metamodellering, en hierbij relaties te leggen met systeemtheorie gerelateerde methoden en tot praktisch bruikbare methodieken te komen.

Hoewel toenemende mogelijkheden op het vlak van sneller rekenen (grid-technologie, snellere processoren) de behoefte aan metamodelering lijken te verkleinen, blijft bij de ontwikkeling van handzame (terug)rekenapplicaties en bij de uitvoering van integrale studies de inzet van metamodellen c.q. vereenvoudigde modellen gewenst om relevante kennis uit oorspronkelijke modellen compact en transparant te extraheren, te integreren en op te schalen. Ook komt het rekenen op grotere tijd- en ruimteschaal hierdoor meer binnen handbereik, terwijl ook sneller en adequater gereageerd kan worden op beleidsvragen, bijvoorbeeld om diverse scenario's te evalueren, terug te rekenen welke beleidsruimte er is om bepaalde doelen

¹⁷ Vergelijk bijvoorbeeld de gedragskarakteristieken van verschillende stoffen in de bodem: nutriënten zijn vaak wel goed via een lineair metamodel te beschrijven, terwijl bij pesticiden niet-lineariteiten een veel grotere rol spelen.

te bereiken etcetera Een eerste uitdaging ligt er voor het MNP om na te gaan op welke concrete en beleidsrelevante gebieden en voor welke modellen en modelsystemen metamodellering een rol kan spelen in de toekomst (bijvoorbeeld Natuurplanner, Stone, Eururalis, Duurzaamheidsscanner).

Te verwachten valt dat hierbij op een aantal thematische gebieden belangrijke vragen zullen liggen:

- ❑ Metamodellering en expertkennis: Hoe kan het 'black-box' karakter van veel metamodellerings-technieken verminderd worden door explicieter kennis in te brengen over relevante processen, systeemvariabelen/componenten en hun interacties? Dit zal de transparantie en plausibiliteit van metamodellen vergroten.
- ❑ Metamodellering voor gekoppelde modelsystemen (bijvoorbeeld 'modelketens'): Is het bijvoorbeeld beter om eerst metamodellen van de afzonderlijke (deel)modellen te maken, en deze aan elkaar te koppelen, of kan men beter streven naar 1 meta-model voor het geheel? Aan welke eisen moeten in het eerste geval de afzonderlijke metamodellen en hun koppelingen voldoen?
- ❑ Metamodelleren bij ruimtelijke modellering (bijvoorbeeld landgebruiksveranderingsmodellen): Hoe om te springen met schaalhierarchieën en interacties, en met het gebruik van zowel continue als discrete variabelen (klassen)?
- ❑ Metamodelleren voor hoogdimensionale problemen: Indien na een eerste screening nog een aanzienlijk aantal relevante factoren overblijven die in het meta-model moeten worden opgenomen, dan is partitionering of decompositie van het probleem gewenst, waarbij het wordt opgedeeld in kleinere deelproblemen die efficiënte experimentering en metamodellering toelaten (zie bijvoorbeeld Wang en Simpson, 2004).
- ❑ Metamodelleren en onzekerheid: Hoe bepaal je de 'modelfout' die je maakt door het referentiemodel te benaderen door het meta-model, en in hoeverre kan het meta-model gebruikt worden om onzekerheden van het referentiemodel vast te stellen?
- ❑ Metamodelleren en optimalisatie/terugrekenen: Aan welke eisen moet het meta-model voldoen om het referentiemodel te vervangen bij optimalisatie vraagstukken, en in hoeverre is het nuttig om bij de constructie van het meta-model gebruik te maken van informatie over het optimalisatiecriterium (bijvoorbeeld gebruik van adaptieve sampling van parameterruimte)?

Welke van deze thema's nader onderzoek vergen zal sterk afhangen van de concrete en beleidsrelevante toepassingen waarbij metamodellering een rol zal spelen. De hoofduitdaging bij een eventuele verdere uitbouw van de metamodelleringsexpertise binnen MNP ligt dan ook bij een adequate afstemming op de eisen die deze toepassingen stellen.

Dankwoord

We danken onder andere Wies Akkermans (Biometris, Wageningen) en Marnik Vanclooster (Université Catholique, Louvain) voor hun bijdrage aan onze gedachten- en ideeënvorming rond metamodellering. Ook danken we Ferd Sauter (LED) voor het aanleveren van de tekst rond Euros (&4.3) en Frank van Gaalen (LDL), Frits Kragt (LDL), Arthur Beusen (IMP) en Michel den Elzen (KMD) voor het leveren van informatie rond meta-STONE (&4.1) en meta-IMAGE (&4.4). Tot slot danken we Cecile van Dijk (DFB/DIV) en Ineke van den Brink (IMP) voor redactionele aanwijzingen en correcties. De verantwoordelijkheid van het hier gerapporteerde ligt uiteraard volledig bij de auteurs van dit rapport.

Literatuur

- Acharya, R. C. (2004). Upscaling of nonlinear reactive transport: From pore to core. Ph-D thesis. Wageningen University, Wageningen.
- Acharya, R. C., S. E. A. T. M. Van der Zee, et al. (2005). "Transport modeling of nonlinearly adsorbing solutes in physically heterogeneous porous media." Water Resour Research, **41**.
- Alam, F. M., K. R. McNaught, et al. (2004). Using Morris' randomized OAT design as a factor screening method for developing simulation metamodels. 2004 Winter Simulation Conference, Washington.
- Alis, Ö. F. en H. Rabitz (2001). "Efficient Implementation of High Dimensional Model Representations." J. Math. Chem, **29**(2): 127-142.
- Antoulas, A. C. en D. S. Sørensen (2001). "Approximation of large-scale dynamical systems: an overview." Int. J. Appl. Math. Comput. Sci., **11**(5): 1093-121.
- Babuška, R. en H. B. Verbruggen (1996). "An overview of fuzzy modelling for control." Control Engineering Practice, **4**(11): 1593-1606.
- Banerjee, I. en M. G. Ierapetritou (2002). "Design Optimization under Parameter Uncertainty for General Black-Box Models." Ind. Eng. Chem. Res., **41**: 6687-6697.
- Banerjee, I. en M. G. Ierapetritou (2003). "Parametric process synthesis for general nonlinear models." Computers and Chemical Engineering, **27**: 1499-1512.
- Barton, R. R. (1992). Metamodels for simulation of input-output relations. 1992 Winter Simulation Conference, Arlington, VA.
- Bayarri, M., J. O. Berger, et al. (2002). A framework for the validation of computer models. Workshop on Foundations for V&V in the 21st Century.
- Bennett, K. P., Campbell C. (2000). "Support Vector machines: Hype or Hallelujah?" SIGKDD Explorations, **2**(2): 1-13.
- Bernhardt, K. en K. W. Wirtz (2004). Reduction of complex models using data-mining and nonlinear projection techniques. Int. Environmental Modelling and Software Society Conference 2004, Osnabrück.
- Bontempi, G. en M. Birattari (2005). "From Linearization to Lazy Learning: A Survey of Divide-and-Conquer Techniques for Nonlinear Control." International Journal of Computational Cognition, **3**(1): 56-73.
- Bouwman, R., R. Costanza, et al. (2002). "Modeling the Dynamics of the Integrated Earth System and the Value of Global Ecosystem Services using the GUMBO model." Ecological Economics, **41**: 529-560.
- Breiman, L., R. Olshen, et al. (1984). Classification and Regression Trees, Wadsworth.
- Bruckner, T., G. Hooss, et al. (2003). "Climate system modelling in the framework of the tolerable windows approach: The ICLIPS climate model." Climatic Change, **56**: 119-137.
- Chen, P.-H., C.-J. Lin, et al. (2003). "A Tutorial on nu-Support Vector Machines." Applied Stochastic Models in Business and Industry (to appear in 2005).
- Chen, W., R. Jin, et al. (2004). Analytical Variance-Based Global Sensitivity Analysis in Simulation-Based Design under Uncertainty, DETC2004-57484. 2004 ASME Design Automation Conference, Salt Lake City, UT.
- Cheng, B. en D. M. Titterton (1994). "Neural Networks: a review from a statistical perspective." Statistical Science, **9**(1): 2-54.
- Clarke, S. M., J. H. Griebisch, et al. (2004). "Analysis of Support Vector Regression for Approximation of Complex Engineering Analyses." ASME Journal of Mechanical Design.
- Craig, P., M. Goldstein, et al. (2001). "Bayesian Forecasting for Complex Systems Using Computer Simulators." Journal of the American Statistical Association, **96**: 717-729.

- Cryer, S. A. en G. E. Applequist (2003). "Direct Treatment of Uncertainty: I: Applications in Aquatic Invertebrate Risk Assessment and Soil Metabolism for Chlorpyrifos." Environmental Engineering Science, **20**: 155-168.
- Cryer, S. A. en G. E. Applequist (2003). "Direct Treatment of Uncertainty: II: Applications in Pesticide Runoff, Leaching and Spray Drift Exposure Modeling." Environmental Engineering Science, **20**: 169-181.
- Daberkow, D. D. (2002). A formulation of metamodel implementation processes for complex system design. Ph-D. Georgia Institute of Technology, Georgia.
- Daberkow, D. D. en D. N. Mavris (2002). An investigation of metamodeling techniques for complex systems design, AIAA Paper 2002-5457. The 9th AIAA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Atlanta, GA.
- Dam, J. D. van (ed.), P. S. C. Heuberger, et al. (2001). Effecten van verplaatsing van agrarische ammoniakemissies: verkenning op provinciaal niveau. RIVM rapport 725501003, RIVM, Bilthoven.
- Daubechies, I. (1990). "The wavelet transform, time-frequency localization and signal analysis." IEEE Trans. Inf. Theory, **36**(5): 961-1005.
- Daubechies, I. (1992). Ten Lectures on Wavelets, SIAM.
- Davis, P. K. en J. H. Bigelow (2003). Motivated Metamodels: synthesis of cause-effect reasoning and statistical metamodeling. MR-1570, RAND corporation.
- De Brabanter, J. (2004). LS-SVM regression modelling and its applications. PhD. K.U.Leuven, Leuven, Belgium.
- Den Elzen, M. G. J. (1998). The meta-IMAGE 2.1 model: an interactive tool to assess global climate change. RIVM Report No. 461502020, RIVM, Bilthoven.
- Den Elzen, M. G. J. (2000). Exploring post-Kyoto climate regimes for differentiation of future commitments to stabilise greenhouse gas concentrations. RIVM Report no. 728001020, RIVM, Bilthoven.
- Den Elzen, M. G. J., M. Berk, et al. (1999). The Brazilian proposal and other options for International Burden Sharing: an evaluation of methodological and policy aspects using FAIR. RIVM Report No. 728001011, RIVM, Bilthoven.
- Den Elzen, M. G. J. en P. Lucas (2003). FAIR 2.0 - A decision-support tool to assess the environmental and economic consequences of future climate regimes. RIVM Report no. 550015001, RIVM, Bilthoven.
- Den Elzen, M. G. J., M. Schaeffer, et al. (2002). Responsibility for past and future global warming: time horizon and non-linearities in the climate system. RIVM Report no. 728001022, RIVM, Bilthoven.
- Dieckmann, U. en C. P. Williams (1992). Model approximation via dimension reduction. AAI Workshop on Approximation and Abstraction of Computational Theories, San Jose, California, USA.
- Dima, M. en G. Lohmann (2004). "Fundamental and derived modes of climate variability: concept and application to interannual time-scales." Tellus A, Dynamic Meteorology and Oceanography, **56**(3): 229-249.
- Dodd, T. J. en R. F. Harrison (2002). A new solution to Volterra series estimation. 2002 IFAC World Congress (CDROM).
- Dodd, T. J. en R. F. Harrison (2003). Estimating Volterra Filters in Hilbert Space. IFAC Conference on Intelligent Control Systems and Signal Processing (ICONS 2003), University of Algarve, Portugal.
- Easterling, R. G. en J. O. Berger (2002). Statistical Foundations for the Validation of Computer Models. Workshop on Foundations for V&V in the 21st Century.
- Eldred, M. S., A. A. Giunta, et al. (2004). DAKOTA: A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis. Version 3.2 Users Manual.
- EMEP (1998). <http://www.emep.int/ozone/webo3sr>, geraadpleegd op 8 december 2004.
- Enderlein, J. en R. Erdman (1997). "Fast fitting of multi-exponential decay curves." Optics Communications, **134**: 371-378.

- Enting, I. G., T. M. L. Wigley, et al. (1994). Future Emissions and Concentrations of Carbon Dioxide: Key Ocean/Atmosphere/Land Analyses. CSIRO Division of Atmospheric Research.
- Forsman, Å. en A. Grimvall (2003). "Reduced models for efficient simulation of spatially integrated outputs of one-dimensional substance transport models." Environmental Modelling & Software, **18**: 319-327.
- Friedman, J. H. (1991). "Multivariate Adaptive Regression Splines (with discussion)." Annals of Statistics, **19**(1): 1-141.
- Friswell, M. I., J. E. T. Penny, et al. (1995). "Model reduction using dynamic and iterated IRS techniques." Journal of Sound and Vibration, **186**(2): 311-323.
- Friswell, M. I., J. E. T. Penny, et al. (1995). "Using linear model reduction to investigate the dynamics of structures with local non-linearities." Mechanical Systems and Signal Processing, **9**(3): 317-328.
- Gevrey, M., I. Dimopoulos, et al. (2003). "Review and comparison of methods to study the contribution of variables in artificial neural network models." Ecological Modelling, **160**(3): 249-264.
- Gibbs, M. N. en D. J. C. MacKay (1997). Efficient implementation of gaussian processes. Cavendish Laboratory, Cambridge, UK.
- Glover, K. (1984). "All optimal Hankel-norm approximations of linear multivariable systems and their error bounds." International Journal of Control, **39**: 1115-1193.
- Goldstein, M. en J. M. Rougier (2003). Calibrated Bayesian Forecasting Using Large Computer Simulators. University of Durham, Durham.
- Goldstein, M. en J. C. Rougier (2005). "Probabilistic formulations for transferring inferences from mathematical models to physical systems." SIAM Journal on Scientific Computing, **26**: 467-487.
- Golub, G. H. en C. F. Van Loan (1996). Matrix Computations, 3rd ed. Baltimore, MD, Johns Hopkins University Press.
- Gunn, S. (1998). Support vector machines for classification and regression. Image, Speech & Intelligent Systems Group, University of Southampton.
- Hasselmann, K. (1993). "Optimal fingerprints for the detection of climate change." Journal of Climate, **6**: 1957-1971.
- Hastie, T. en R. Tibshirani (1990). Generalized Additive Models. London, Chapman and Hall.
- Hastie, T., R. Tibshirani, et al. (2001). The Elements of Statistical Learning: Data Mining, Inference and Prediction. New York, Springer Verlag.
- Hearst, M. A., Schölkopf B., et al. (1998). "Trends Controversies: Support vector machines." IEEE Intelligent Systems, **13**(4): 18-28.
- Heyes, C., W. Schöpp, et al. (1995). A simplified model to predict long-term ozone concentrations in Europe. IIASA, Laxenburg, Austria.
- Ho, T. S. en H. Rabitz (2003). "Reproducing kernel Hilbert space interpolation methods as a paradigm of high dimensional model representations: Application to multidimensional potential energy surface construction." J. Chem. Phys., **119**: 6433-6442.
- Hooimeijer, M. A. (2001). Reduction of complex computational models. PhD. Delft University of Technology., Delft, The Netherlands.
- Hooimeijer, M. A., A. W. Heemink, et al. (2000). Identifying Patterns in Complex Models for Model Reduction. 4th Int. Conf. on Hydro-Science and -Engineering, Seoul, Rep. of Korea.
- Hooss, G., R. Voss, et al. (2001). "A nonlinear impulse response model of the coupled carbon cycle-climate system (NICCS)." Climate Dynamics, **18**: 189-202.
- Hotelling, H. (1935). "The most predictable criterion." J. Educational Psychology, **26**: 139-142.
- Huber, K.-P. en M. R. Berthold (1998). Application of Fuzzy Graphs for Metamodeling. 7th IEEE International Conference on Fuzzy Systems, Anchorage, Alaska.
- Huber, K.-P., M. R. Berthold, et al. (1996). "Analysis of Simulation Models with Fuzzy Graph based Metamodeling." Performance Evaluation, **27&28**: 473-490.

- Isukapalli, S. S. en P. G. Georgopoulos (1999). Computational Methods for Sensitivity and Uncertainty Analysis for Environmental and Biological Models: Towards a Framework for an Exposure and Dose Modeling and Analysis System. Computational Chemodynamics Laboratory, Environmental and Occupational Health Sciences Institute, Piscataway, NJ.
- Isukapalli, S. S. en P. G. Georgopoulos (2001). Computational Methods for Sensitivity and Uncertainty Analysis for Environmental and Biological Models. Computational Chemodynamics Laboratory, Environmental and Occupational Health Sciences Institute, Piscataway, NJ.
- Isukapalli, S. S., A. Roy, et al. (1998). "Stochastic Response Surface Methods (SRSMs) for uncertainty propagation: Application to environmental and biological systems." Risk Analysis, **18**(3): 351-363.
- Isukapalli, S. S., A. Roy, et al. (2000). "Efficient sensitivity/uncertainty analysis using the combined Stochastic Response Surface Method and Automated Differentiation: Application to environmental and biological systems." Risk Analysis, **20**(5): 591-602.
- Jin, R., W. Chen, et al. (2001). "Comparative Studies of Metamodeling Techniques under Multiple Modeling Criteria." Structural and Multidisciplinary Optimization, **23**(1): 1-13.
- Jin, R., W. Chen, et al. (2002). On Sequential Sampling for Global Metamodeling in Engineering Design, DETC-DAC34092. 2002 ASME Design Automation Conference, Montreal, Canada.
- Jin, R., W. Chen, et al. (2003). "An Efficient Algorithm for Constructing Optimal Design of Computer Experiments." Journal of Statistical Planning and Inference.
- Jin, R., W. Chen, et al. (2004). Analytical Metamodel-Based Global Sensitivity Analysis and Uncertainty Propagation for Robust Design, , paper 2004-01-0429. SAE Congress, Detroit, MI.
- Johansen, T. A. en B. A. Foss (1995). "Identification of non-linear System Structure and Parameters using Regime Decomposition." Automatica, **31**: 321-326.
- Johansen, T. A. en B. A. Foss (1995). Local Modeling as a tool for semi-empirical and semi-mechanistic process modeling. Neural Networks for Chemical Engineers. A. B. Bulsari. Amsterdam, Elsevier: 297-334.
- Johansen, T. A. en B. A. Foss (1995). Semi-Empirical Modeling of Non-linear Dynamic Systems based on Identification of Operating Regimes and Local Models. Advances in Neural Networks for Control Systems. K. Hunt en G. Irwin. London, Springer-Verlag: 105-126.
- Johansen, T. A. en B. A. Foss (1997). "Operating regime based process modeling and identification." Computers and Chemical Engineering, **21**: 159-176.
- Joos, F., M. Bruno, et al. (1996). "An Efficient and Accurate Representation of Complex Oceanic and Biospheric Models of Anthropogenic Carbon Uptake." Tellus, **48B**(3): 397-417.
- Jury, W. A. (1982). "Simulation of solute transport using a transfer function model." Water Resour. Res., **18**: 363-368.
- Kennedy, M. C. en A. O'Hagan (2001). "Bayesian calibration of computer models." J. Roy. Statistic. Soc. Ser. B. Stat. Methodol., **63**(3): 425-464.
- Kennedy, M. C., A. O'Hagan, et al. (2001). Bayesian analysis of computer code outputs. Quantitative Methods for Current Environmental Issues. C. W. Anderson, C. V. Barnett, P. C. Chatwin en A. El-Shaarawi, Springer-Verlag.
- Kleijnen, J. P. C. (2005). "An overview of the design and analysis of simulation experiments for sensitivity analysis." European Journal of Operational Research, **164**(2): 287-300.
- Kleijnen, J. P. C., S. M. Sanchez, et al. (2004). A User's Guide to the Brave New World of Designing Simulation Experiments. Department of Information Systems and Management /Center for Economic Research (CentER), Tilburg University.
- Kleijnen, J. P. C. en R. G. Sargent (2000). "A methodology for the fitting and validation of metamodels in simulation." European Journal of Operational Research, **120**(1): 14-29.

- Kleijnen, J. P. C. en W. C. M. van Beers (2004). "Application-driven sequential designs for simulation experiments: Kriging metamodelling." Journal of the Operational Research Society, **55**(8): 876-883.
- Knutti, R., T. F. Stocker, et al. (2003). "Probabilistic climate change projections using neural networks." Climate Dynamics, **21**: 257-272.
- Kooperberg, C. en F. O'Sullivan (1994). The use of a statistical forecast criterion to evaluate alternative empirical spatial oscillation pattern decomposition methods in climatological fields. Dept. Statistics, Univ. Washington, Seattle, WA.
- Kosko, B. (1992). Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence, Prentice Hall.
- Krysl, P., S. Lall, et al. (2001). "Dimensional model reduction in non-linear finite element dynamics of solids and structures." International Journal for Numerical Methods in Engineering, **51**: 479-504.
- Lall, S., J. E. Marsden, et al. (1998). Empirical model reduction of controlled nonlinear systems. California Institute of Technology.
- Lek, S. en J. F. Guegan (1999). "Artificial Neural Networks as a tool in Ecological Modelling. An Introduction." Ecological Modelling, **120**: 65-73.
- Leterme, B., M. D. A. Rounsevell, et al. (2004). Development of a probabilistic data assimilation methodology for the assessment of groundwater contamination by pesticides at the catchment scale. IAS-conference on Groundwater Vulnerability Assessment and Mapping, Sosnowiec, Poland.
- Li, G., M. Artamonov, et al. (2003). "High Dimensional Model Representations Generated from Low Order Terms – 1p-RS-HDMR." Journal of Computational Chemistry, **24**: 647-656.
- Li, G., H. Rabitz, et al. (2003). "Correlation Method for Variance Reduction of Monte Carlo Integration in RS-HDMR." Journal of Computational Chemistry, **24**: 277-283.
- Li, G., H. Rabitz, et al. (2002). Efficient Sensitivity/Uncertainty Analysis Using the High Dimensional Model Representation (HDMR) Method: Application to the Princeton Groundwater Model.
- Li, G., C. Rosenthal, et al. (2001). "High Dimensional Model Representations." J. Phys. Chem. A, **105**: 7765-7777.
- Li, G., S. W. Wang, et al. (2002). "Global uncertainty assessments by high dimensional model representations (HDMR)." Chem. Eng. Sci., **57**: 4445-4460.
- Ljung, L. (1999). System Identification: Theory for the User, 2nd Ed. Upper Saddle River, NJ, Prentice Hall.
- Ljung, L. (2004). System Identification Toolbox 6.1, The Mathworks, Inc.
- Lorenz, E. N. (1956). Empirical orthogonal functions and statistical weather prediction. Dept. of Meteorology, MIT.
- Lucas, D. D. en R. G. Prinn (2004). "Parametric sensitivity and uncertainty analysis of dimethyl-sulfide oxidation in the remote marine boundary layer." Atmos. Chem. Phys. Discuss., **4**: 6379-6430.
- MacKay, D. J. C. (1997). Introduction to Gaussian Processes. Cavendish Laboratory, Cambridge, UK.
- Maier-Reimer, E. en K. Hasselmann (1987). "Transport and storage of CO₂ in the ocean - an inorganic ocean-circulation carbon cycle model." Climate Dynamics, **2**: 63-90.
- Martin, J. D. en T. W. Simpson (2003). A study on the Use of Kriging Models to Approximate Deterministic Computer Models, Paper No. DETC2003/DAC-48762. ASME 2003 Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Chicago, IL.
- Martin, J. D. en T. W. Simpson (2004). On the Use of Kriging Models to Approximate Deterministic Computer Models, Paper No. DETC2004/DAC-57300. ASME Design Engineering Technical Conferences - Design Automation Conference, Salt Lake City, UT.
- Mathworks (1994-2004). Matlab, The Language of Technical Computing, The Mathworks, Inc.

- McCullagh, P. en J. A. Nelder (1989). Generalized Linear Models, 2nd edition. London, Chapman and Hall.
- Meckesheimer, M., R. R. Barton, et al. (2001). "Metamodeling of Combined Discrete/Continuous Responses." AIAA Journal, **39**(10): 1955-1959.
- Meckesheimer, M., A. J. Booker, et al. (2002). "Computationally Inexpensive Metamodel Assessment Strategies." AIAA Journal, **40**(10): 2053-2060.
- Mensink, C., L. Delobbe, et al. (2001). A policy oriented model system for the assessment of longterm effects of emission reductions on ozone. 25th NATO/CCMS International Technical Meeting on Air Pollution Modelling and its Application, Louvain-la-Neuve.
- Mitaim, S. en B. Kosko (2001). "The Shape of Fuzzy Sets in Adaptive Function Approximation." IEEE Transactions on Fuzzy Systems, **9**(4): 637-656.
- Moisen, G. G. en T. S. Frescino (2002). "Comparing five modelling techniques for predicting forest characteristics." Ecological Modelling, **157**(2-3): 209-225.
- Moore, B. C. (1981). "Principal Component Analysis in Linear Systems: Controllability, Observability and Model Reduction." IEEE Trans. on Automat. Contr., **AC-26**, pp. 17-31, February 1981, **AC-26**: 17-31.
- Morris, C. W., A. Autret, et al. (2001). "Support vector machines for identifying organisms – a comparison with strongly partitioned radial basis function networks." Ecological Modelling, **146**: 57-67.
- Morris, M. D. en T. J. Mitchell (1992). Exploratory Designs for Computer Experiments. ORNL/TM-12045, Oak Ridge National Laboratory, Oak Ridge, TN.
- Murray-Smith, R. en T. A. Johansen, Eds. (1997). Multiple Model Approaches to Modeling and Control, Taylor and Francis.
- Nelles, O. (2000). Non-linear system identification. Berlijn-Heidelberg, Springer Verlag.
- Newman, A. J. en P. S. Krishnaprasad (2000). Computing Balanced Realizations for Nonlinear Systems. 14th International Symposium on Mathematical Theory of Networks and Systems, MTNS, Perpignan, France.
- O'Hagan, A., M. C. Kennedy, et al. (1998). Uncertainty analysis and other inference tools for complex computer codes (with discussion). Bayesian Statistics. M. Bernardo en e. al., Oxford University Press. **6**: 503-524.
- Oakley, J. (2002). "Eliciting Gaussian process priors for complex computer codes." Journal of the Royal Statistical Society: Series D (The Statistician), **51**(1): 81-97.
- Oakley, J. E. en A. O'Hagan (2002). "Bayesian inference for the uncertainty distribution of computer model outputs." Biometrika, **89**: 769-784.
- Oakley, J. E. en A. O'Hagan (2004). "Probabilistic sensitivity analysis of complex models: a Bayesian approach." Journal of the Royal Statistical Society: Series B, **66**(3): 751.
- Olden, J. D., M. K. Joy, et al. (2004). "An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data." Ecological Modelling, **178**: 389-397.
- Orr, M. J. L. (1999). Matlab functions for radial basis function networks. Edinburgh University, Edinburgh.
- Orr, M. J. L. (1999). Recent advances in radial basis function networks. Edinburgh University, Edinburgh.
- Orr, M. L. J. (1995). "Regularization in the selection of radial basis function centers." Neural Computation, **7**(3): 606-623.
- Pelckmans, K., J. A. K. Suykens, et al. (2002). LS-SVMlab Toolbox User's Guide. Internal Report 02-145. Leuven, Belgium, K.U.Leuven.
- Piñeros Garcet, J. D., A. M. A. van der Linden, et al. (2004). Development of a GeoPEARL metamodel. Université Catholique de Louvain, Unité de Genie Rural, Louvain-la-Neuve, Belgium.
- Powell, M. J. D. (1987). Radial basis functions for multivariable interpolation: a review. Algorithms for Approximation. J. C. Mason, Cox, M.G., Oxford University Press.

- Qian, Z., C. C. Seepersad, et al. (2005). "Building Surrogate Models Based on Detailed and Approximate Simulations." Transactions of the ASME, Journal of Mechanical Design.
- Rabitz, H. en Ö. F. Alis (1999). "General Foundations of High Dimensional Model Representations." J. Math. Chem., **25**: 197-233.
- Rabitz, H. en Ö. F. Alis (2000). Managing the Tyranny of Parameters in Mathematical Modelling of Physical Systems. Sensitivity Analysis. A. Saltelli, K. Chan en M. Scott. Chichester, John Wiley & Sons: 199-223.
- Rabitz, H., Ö. F. Alis, et al. (1999). "Efficient input-output model representations." Computer Phys. Comm., **117**: 11-20.
- Ratto, M., S. Tarantola, et al. (2000). Model reduction techniques for quantification of uncertainty, IMPACT/JRC/WP6/D16. Joint Research Centre of European Commission, Institute for Systems, Informatics and Safety, Ispra, It.
- Ratto, M., S. Tarantola, et al. (2000). Model reduction techniques for time series normalization, IMPACT/JRC/WP6/D17. Joint Research Centre of European Commission, Institute for Systems, Informatics and Safety, Ispra, It.
- Rewiński, M. J. en J. White (2003). "A trajectory piecewise-linear approach to model order reduction and fast simulation of nonlinear circuits and micromachined devices." IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, **22**(2): 155-70.
- Roth, K. en W. A. Jury (1993). "Modeling the transport of solutes to groundwater using transfer functions." J. Environ. Qual., **22**: 487-493.
- Rotmans, J. en M. B. A. van Asselt (2001). "Uncertainty and Integrated Assessment Modeling: a Labyrinthic Path." Integrated Assessment, **2**(2): 43-57.
- Rugh, W. J. (1981). Nonlinear Systems Theory, The Volterra-Wiener Approach, The Johns Hopkins University Press.
- Sacks, J., S. Schiller, et al. (1989). "Designs for computer experiments." Technometrics, **31**: 41-47.
- Sacks, J., W. Welch, et al. (1989). "Design and Analysis of Computer Experiments." Statistical Science, **4**: 409-435.
- Saltelli, A., K. Chan, et al. (2000). Sensitivity Analysis. Chichester, England, John Wiley & Sons, Ltd.
- Saltelli, A., S. Tarantola, et al. (2004). Sensitivity Analysis in Practice, A Guide to Assessing Scientific Models. Chichester, England, John Wiley & Sons Ltd.
- Scherpen, J. M. A. (1993). "Balancing for non-linear systems." Syst. & Contr. Let.: 143-153.
- Scherpen, J. M. A. (1994). Balancing for non-linear systems. PhD. University of Twente, The Netherlands.
- Schetzen, M. (1980). The Volterra and Wiener Theories of Nonlinear Systems, John Wiley & Sons.
- Schölkopf, B. en A. J. Smola (2002). Learning with Kernels. Cambridge, MA, MIT Press.
- Schoukens, J., J. G. Nemeth, et al. (2003). "Fast approximate identification of nonlinear systems." Automatica, **39**(7): 1267-1274.
- Simpson, T., L. Dennis, et al. (2002). "Sampling Strategies for Computer Experiments: Design and Analysis." International Journal of Reliability and Application, **2**(3): 209-240.
- Simpson, T. W., A. J. Booker, et al. (2004). "Approximation Methods in Multidisciplinary Analysis and Optimization: A Panel Discussion." Structural and Multidisciplinary Optimization, **27**(5): 302-313.
- Simpson, T. W., T. M. Mauery, et al. (2001). "Kriging Metamodels for Global Approximation in Simulation-Based Multidisciplinary Design Optimization." AIAA Journal, **39**(12): 2233-2241.
- Simpson, T. W., J. Peplinski, et al. (2001). "Metamodels for Computer-Based Engineering Design: Survey and Recommendations." Engineering with Computers, **17**(2): 129-150.

- Sjöberg, J., Q. Zhang, et al. (1995). "Nonlinear black-box modeling in system identification: A unified overview." *Automatica*, **31**(12): 1691-1724.
- Smola, A. J. en B. Schölkopf (1998). "On Kernel-based method for pattern recognition, regression, approximation and operator inversion." *Algorithmica*, **22**: 211-231.
- Smola, A. J. en B. Schölkopf (1998). *A Tutorial on Support Vector Regression*. Royal Holloway College, University of London, UK.
- Srivastava, A., K. Hacker, et al. (2004). "A method for using legacy data for metamodel-based design of large-scale systems." *Structural and Multidisciplinary Optimization*, **28**: 146-155.
- Sterman, J. D. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Boston, McGraw-Hill.
- Stewart, I. T. en K. Loague (2003). "Development of type transfer functions for regional-scale nonpoint source groundwater vulnerability assessments." *Water Resources Research*, **39**(12).
- Stewart, I. T. en K. Loague (2004). "Assessing Ground Water Vulnerability with the Type Transfer Function Model in the San Joaquin Valley, California." *J. Environ. Qual.*, **33**: 1487-1498.
- Stoer, J. en R. Bulirsch (1983). *Introduction to Numerical Analysis*. New York, Springer-Verlag.
- STOWA (2003). *Waterkwaliteit in Waterlood*. Rapportnummer 2003-02 (STOWA); Waterloodrapport 6, STOWA, Utrecht.
- Suykens, J. A. K., T. Van Gestel, et al. (2002). *Least Squares Support Vector Machines*. Singapore, World Scientific Publishing Co., Pte, Ltd.
- Tappenden, P., J. Chilcott, et al. (2004). "Methods for expected value of information analysis in complex health economic models: Developments on the health economics of beta interferon and glatiramer cetate for multiple sclerosis." *Health Technology Assessment*, **8**(27).
- Tatang, M. A., W. W. Pan, et al. (1997). "An efficient method for parametric uncertainty analysis of numerical geophysical model." *Journal of Geophysical Research-Atmospheres*, **102**(D18): 21925-21932.
- Tiktak, A., J. J. T. I. Boesten, et al. (2002). *Nationwide assessments of non-point source pollution with field-scale developed model: The pesticide case*. Proceedings of the 5th international symposium on spatial accuracy assessment in natural resources and environmental sciences (Accuracy 2002), Melbourne, 10-12 July 2002, pp. 17-31.
- Tiktak, A., D. S. de Nie, et al. (2004). "Assessment of the pesticide leaching risk at the Pan-European level: The EuroPEARL approach." *J. Hydrol.*, **289**: 222-238.
- Tiktak, A., D. S. de Nie, et al. (2002). "Modelling the leaching and drainage of pesticides in the Netherlands: The GeoPEARL model." *Agronomie*, **22**: 373-387.
- Tiktak, A., F. A. Swartjes, et al. (1994). *Sensitivity analysis of a model for pesticide leaching and accumulation*. Predictability and nonlinear modelling in natural sciences and economics, Kluwer Academic Publishers, Dordrecht, the Netherlands.
- Tiktak, A., F. van den Berg, et al. (2000). *Pesticide Emission Assessment at Regional and Local Scales: User Manual of FOCUS Pearl 1.1.1*. RIVM report 711401008, RIVM, Bilthoven, the Netherlands.
- Turner, C. J., M. I. Campbell, et al. (2003). *Generic sequential sampling for metamodel approximations*, DETC2003/CIE-48230. DETC'03: ASME 2003 Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Chicago, Illinois.
- UN/ECE (2000). <http://www.unece.org/env/lrtap>, geraadpleegd op 8 december 2004.
- Van Beers, W. C. M. en J. P. C. Kleijnen (2003). "Kriging for interpolation in random simulation." *Journal of the Operational Research Society*, **54**(3): 255-262.
- Van Beers, W. C. M. en J. P. C. Kleijnen (2004). *Kriging interpolation in simulation: a survey*. 2004 Winter Simulation Conference, Washington, D.C.
- Van der Zee, S. E. A. T. M. en J. J. T. I. Boesten (1991). "Effects of soil heterogeneity on pesticide leaching to groundwater." *Water Resources Research*, **27**: 3051-3063.

- Van Dooren, P. M. (1999). Gramian based model reduction of large-scale dynamical systems. Numerical Analysis 1999, Chapman and Hall, CRC Press.
- Van Jaarsveld, J. A. (1995). Modelling the atmospheric behaviour of pollutants on various spatial scales. UvU, Utrecht.
- Van Jaarsveld, J. A. en F. A. M. de Leeuw (1993). "OPS:an operational atmospheric transport model for priority substances." Environmental Software, **8**: 91-100.
- Vanclooster, M., A. Armstrong, et al. (2003). Effective approaches for predicting environmental concentrations of pesticides: The APECOP project. Proceedings of the XII International Symposium on Pesticide Chemistry: Pesticide in air, plant, soil and water systems, Piacenza, Italy.
- Vanclooster, M., J. J. T. I. Boesten, et al. (2000). "A European test of pesticide-leaching models: methodology and major recommendations." Agric. Water Manage., **44**: 1-19.
- Vapnik, V. (1998). Statistical Learning Theory. New York, Wiley.
- Vapnik, V., S. Golowich, et al. (1997). Support vector method for function approximation, regression estimation, and signal processing. Advances in Neural Information Processing Systems 9. M. C. Mozer, M. I. Jordan en T. Petsche. Cambridge, MA, MIT Press: 281-287.
- Vapnik, V. N. (1995). The Nature of Statistical Learning Theory. New York, Springer Verlag.
- Varga, A. (1991). Balancing-free square-root algorithms for computing singular perturbation approximations. 30th IEEE Conference on Decision and Control, Brighton, UK.
- Varga, A. (1995). "Enhanced modal approach for model reduction." Math. Modell. Syst., **1**: 91-105.
- Varga, A. (2001). Model reduction software in the SLICOT library. Applied and Computational Control, Signals, and Circuits. B. N. Datta. Boston,, Kluwer Academic Publishers. **2**: 239-282.
- Verdonschot, P. en R. Rebi Nijboer (2002). A review of model designs, deliverable 8. Alterra, Green World Research, Wageningen, The Netherlands.
- Vert, J.-P., K. Tsuda, et al. (2004). A primer on kernel methods. Kernel Methods in Computational Biology (Computational Molecular Biology). B. Schölkopf, K. Tsuda en J.-P. Vert: 35-70.
- Vleeshouwers, L. M., J. H. Janse, et al. (2004). A metamodel for PCLake. RIVM, report nr. 703715 007, RIVM, Bilthoven.
- Volkwein, S. (1999). Proper orthogonal decomposition and singular value decomposition. Institut für Mathematik, Universität Graz.
- Volkwein, S. (2004). Interpretation of proper orthogonal decomposition as singular value decomposition and HJB-based feedback design. 16th International Symposium on Mathematical Theory of Networks and Systems (MTNS2004), KU Leuven, Belgium.
- Von Storch, H., T. Bruns, et al. (1988). "Principal Oscillation Pattern analysis of the 30-60 day oscillation in general circulation model equatorial troposphere." J. Geo. Res., **93**(D9): 11022-11036.
- Von Storch, H. en A. Navarra, Eds. (1995). Analysis of Climate Variability: Applications of Statistical Techniques, Springer Verlag 3-10.
- Von Storch, H. en F. W. Zwiers (1999). Statistical Analysis in Climate Research, Cambridge University Press.
- Vreeswijk, D. G., J. M. M. Aben, et al. (2000). SOREMA 1.0.
- Wan, Y., T. J. Dodd, et al. (2003). Infinite Degree Volterra Series Estimation. Second International Conference on Computational Intelligence, Robotics and Autonomous Systems (CDROM), Singapore.
- Wang, G. G. (2003). "Adaptive Response Surface Method Using Inherited Latin Hypercube Design Points." Transactions of the ASME, Journal of Mechanical Design, **125**: 210-220.
- Wang, G. G. en T. W. Simpson (2004). "Fuzzy Clustering Based Hierarchical Metamodeling for Design Optimization." Engineering Optimization, **36**(3): 313-335.
- Wang, S. W., P. G. Georgopoulos, et al. (2003). "Random Sampling-High Dimensional Model Representation (RS-HDMR) with Nonuniformly Distributed Variables:

- Application to an Integrated Multimedia/Multipathway Exposure and Dose Model for Trichloroethylene." J. Phys. Chem. A, **107**: 4707-4716.
- Webster, M., C. Forest, et al. (2003). "Uncertainty Analysis of Climate Change and Policy Response." Climatic Change, **61**(3): 295-320.
- Webster, M., M. A. Tatang, et al. (2000). Application of the probabilistic collocation method for an uncertainty analysis of a simple ocean model. MIT.
- Webster, M. D. en A. P. Sokolov (2000). "A Methodology for Quantifying Uncertainty in Climate Projections." Climatic Change, **46**(4): 417-446.
- Welch, W. J., R. J. Buck, et al. (1992). "Screening, predicting, and computer experiments." Technometrics, **34**: 15-25.
- Wigley, T. M. L. (1992). "A simple inverse carbon cycle 'model'." Global Biochem. Cycles, **5**(4): 373-382.
- Willcox, K. en J. Peraire (2002). "Balanced Model Reduction via the Proper Orthogonal Decomposition." AIAA Journal, **40**,(11).
- Wortelboer, F. G., R. Rosenboom, et al. (2005). Ecologische Effectberekeningen voor de 2e Nationale Natuurverkenning; Aquatische systemen. RIVM Rapport (in prep), RIVM, Bilthoven.
- Young, P. C. (1998). "Data-based mechanistic modelling of environmental, ecological, economic and engineering systems." Environmental Modelling & Software, **13**: 105-122.
- Young, P. C. (1999). "Data-based mechanistic modelling, generalised sensitivity and dominant mode analysis." Computer Physics Communications, **117**: 113-129.
- Young, P. C. (2001). Data-based mechanistic modelling and validation of rainfall-flow processes. Perspectives in Hydrological Science. M. G. Anderson en P. D. Bates. Chichester, J. Wiley: 117-161.
- Young, P. C. (2001). Data-based mechanistic modelling of environmental systems. IFAC Workshop on Environmental Systems, Tokyo, Japan.
- Young, P. C. en A. Jarvis (2002). Data-based Mechanistic Modelling and State Dependent Parameter Models. Centre for Research on Environmental Systems and Statistics, Lancaster University, Lancaster, UK.
- Young, P. C. en M. J. Lees (1993). The active mixing volume: a new concept in modelling environmental systems Chapter 1. in V. Barnett and R. Turkman (eds.), Statistics and the Environment, J. Wiley & Sons: Chichester. Statistics and the Environment. V. Barnett en R. Turkman. Chichester, J. Wiley&Sons: 3-43.
- Young, P. C., P. G. McKenna, et al. (2001). "Identification of non-linear stochastic systems by state dependent parameter estimation." International Journal of Control, **74**(18): 1837-1857.
- Zwiers, F. W. en IDAG (2005). "Detecting and Attributing External Influences on the Climate System: A Review of Recent Advances." Journal of Climate (in press).

Appendix Methoden voor Modelreductie

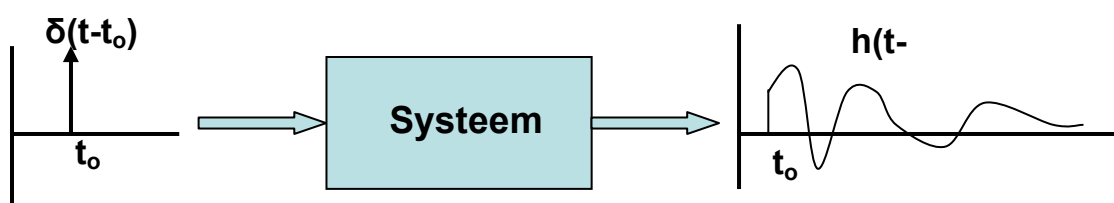
A.1. Proces-gebaseerde modelreductie

A.1.1. IRF: Impuls Responsie Functies

Het *input-output* gedrag van *lineaire dynamische systemen* kan compact gekarakteriseerd worden door de *impuls-responsie functie* (IRF of ook wel ‘Greense functie’), die weergeeft hoe de output ($y(t)$) van het systeem reageert op een specifieke impuls die op een bepaald tijdstip τ op de input (u) gegeven wordt. De output $y(t)$ van het betreffende lineair dynamische systeem bij een willekeurige input $u(t)$ kan worden uitgedrukt als convolutie-integraal:

$$y(t) = \int_{-\infty}^t h(t-\tau) \cdot u(\tau) \cdot d\tau \quad (1)$$

waarbij $h(t)$ de impuls responsie functie (zie Figuur 8) aanduidt.¹⁸



Figuur 8: De reactie van een systeem op een impuls in het inputsignaal.

Voor veel systemen blijkt de *impuls-responsie* goed benaderd te kunnen worden door bijvoorbeeld een som van enkele exponentiële termen:

$$h(t) = A_0 + \sum_{i=1}^k A_i \cdot e^{-t/\tau_i} \quad (2)$$

waarbij de karakteristieke coëfficiënten A_i en τ_i vast te stellen zijn aan de hand van u en y . Voorbeelden hiervan zijn te vinden bij bijvoorbeeld klimaat modellen, waarvan sommige geaggregeerde input-output relaties¹⁹ door eenvoudige impulsresponsies van bovenstaande vorm compact benaderd kunnen worden (zie bijvoorbeeld Maier-Reimer en Hasselmann, 1987; Enting, Wigley et al., 1994; Joos, Bruno et al., 1996). Andere vormen van impulsresponsies zijn ook mogelijk, zie bijvoorbeeld Jury, 1982; Roth en Jury, 1993; Stewart en Loague, 2003; Stewart en Loague, 2004 waarin impuls responsie modellen gepresenteerd worden ter beschrijving van het transport van verontreinigingen in de bodem onder steady-state stromings condities.

¹⁸ Gemakshalve beperken we ons tot een lineair, tijdsinvariant continue-tijd systeem, met 1 ingang en 1 uitgang, en veronderstellen we dat de input-output relatie causaal is (d.w.z. $h(t)=0$ voor $t<0$).

Generalisaties naar tijds(in)variante discrete systemen, en systemen met meerdere in- en uitgangen etcetera zijn mogelijk, maar deze bespreken we niet expliciet.

¹⁹ Bijv. ‘emissie \rightarrow concentratie’; ‘concentratie \rightarrow opwarming’; ‘opwarming \rightarrow zeespiegelstijging’ op globale schaal.

Voor het bepalen van de impulsresponsies op basis van $u(.)$ en $y(.)$ zijn diverse schattingstechnieken beschikbaar, variërend van bijvoorbeeld parametrische tijdreeksanalyse technieken tot niet-parametrische spectraal-analyse technieken²⁰ (zie bijvoorbeeld Ljung 1999; Ljung 2004).

Ook is een uitbreiding van de impuls responsie modellen mogelijk naar *niet-lineaire dynamische* modellen, bijvoorbeeld via het gebruik van Volterra reeks en Volterra kernel representaties, die een generieke impuls respons representatie leveren (Schetzen, 1980; Rugh, 1981), c.q. via Wiener-Hammerstein modellen waarbij de niet-lineariteit gescheiden wordt van de dynamische component (zie bijvoorbeeld hoofdstuk 18 in Nelles (2000)). De dimensionaliteit van deze representaties kan echter snel excessief groot worden, afhankelijk van de orde waarmee men het systeem wil benaderen, en het verdient aanbeveling om te zoeken naar zuinige en rekentechnisch goed hanteerbare procedures (zie ook Dodd en Harrison, 2002; Dodd en Harrison, 2003; Wan, Dodd et al., 2003 voor suggesties). Ook zijn andere manieren denkbaar om niet-lineair gedrag via impuls responsies te beschrijven. Zo presenteert Hooss, Voss et al. (2001) een mogelijkheid om door koppeling van lineaire impuls responsies met vereenvoudigde niet-lineaire modelrelaties te komen tot een eenvoudig metamodel voor klimaatmodellen.

A.1.2. DBM – Data Based Mechanistic Modelling

Deze methodiek, die al sinds de jaren '70 wordt ontwikkeld en gepropageerd door Peter Young en anderen (zie Young, 1998; Young en Jarvis, 2002 en de vele referenties daarin), richt zich op dynamische systemen, dat wil zeggen systemen waarbij de diverse grootheden functies zijn van de tijd. Daarbij wordt uitgegaan van de notie en ervaring dat vele processen, die gemodelleerd worden met behulp van zeer complexe (deterministische) stelsels wiskundige vergelijkingen, in de praktijk vaak zeer goed gerepresenteerd kunnen worden met behulp van efficiënte (lage orde) “*dominant mode*” modellen, waarin de dominerende gedragsdynamica wordt beschreven, terwijl de niet-dominante dynamica wordt verwaarloosd.

De kern van de methode is dat er door middel van statistische methoden (onder andere tijdreeksanalyse) een model wordt bepaald op basis van data, met zo weinig mogelijk a priori veronderstellingen, dat wil zeggen een inductieve benadering in tegenstelling tot hypothetisch-deductieve benadering. Vervolgens wordt geprobeerd om fysische (mechanistische) betekenis aan het resulterende model toe te kennen.

De benodigde data kunnen ofwel meetdata uit de praktijk zijn als ook simulatiedata afkomstig van een bestaand (complex) simulatiemodel. Zulke modellen zijn juist vaak expliciet ontwikkeld op basis van fysisch inzicht, zoals massa- en energiebalansen. In dit geval is de DBM-methode dus ook bruikbaar als modelreductiemethode.

Indien er een complex simulatiemodel voorhanden is, dient dit model eerst onderzocht te worden op onzekerheden en gevoeligheden. Young (1998) stelt voor om hiervoor gebruik te maken van Monte Carlo Simulatie (MCS), waartoe het gewoonlijk *deterministische* simulatiemodel in een ‘stochastische vorm’ wordt geschreven door onzekerheden toe te kennen aan inputs en modelparameters. Voor de dynamische inputsignalen gebeurt dit in de praktijk in de vorm van tijdreeksmodellen; aan parameters wordt vaak een verdelingsfunctie toegekend. Hierbij is expertkennis natuurlijk onontbeerlijk voor het bepalen/kiezen van de onzekerheidsspecificaties.

²⁰ Indien de impulsresponsie rechtstreeks beschikbaar is, dan kan voor het fitten van een multi-exponentieel model gebruik worden gemaakt van specifieke software (zie Nielsen, 2000, Enderlein en Erdman, 1997).

Op basis van Monte Carlo Simulatie²¹ worden vervolgens conclusies getrokken over de onzekerheidspropagatie, gevoeligheden van de diverse inputs en modelparameters en het belang van de diverse modelonderdelen (van het complexe simulatiemodel) in relatie tot het dominante procesgedrag. Bij het bepalen van de relevante modelonderdelen wordt gebruik gemaakt van zogenaamde Dominant Mode Analysis (DMA), waarbij de (ruisvrije) simulatiedata van het complexe simulatiemodel geanalyseerd worden door statistische data-reductietechnieken (Young, 1999). Zo worden lage orde (dominant mode) benaderingen verkregen van het complexe model.

Indien mogelijk wordt het verkregen model in een latere fase, als echte meetdata uit de praktijk voorhanden zijn, verder verbeterd c.q. gecalibreerd op deze data.

De auteurs wijzen op het grote belang van de kwaliteit van de data en dus ook van het experimentontwerp, i.e. de keuze van de inputsignalen die gebruikt worden om de simulatiedata te genereren. In Young (1998) wordt een voorbeeld gegeven van een niet-lineair simulatiemodel, waarvan de niet-lineaire componenten vrijwel irrelevant zijn voor de beschikbare meetdata, i.e. in de praktijk kunnen deze niet-lineaire componenten vaak verwaarloosd worden. Tevens wordt het belang van schaafeffecten benadrukt, i.e. op welke schaal dienen de diverse signalen gedefinieerd en geanalyseerd te worden. Dit hangt direct samen met het beoogde doel van het lage orde model.

Voor het bepalen van de lage orde modellen, vaak aangeduid als modelcalibratie, -optimalisatie of -identificatie, is een heel scala aan tools ontwikkeld en beschikbaar. In de meest eenvoudige vorm bestaat zo'n model uit een set van lineaire transferfuncties, met – indien vereist door niet stationair procesgedrag - tijdvariërende parameters. In de meest algemene vorm wordt gezocht naar een model van de vorm:

$$y_t = T_t + S_t + f(u_t) + N_t + e_t \quad (3)$$

waarbij het subscript t de tijdindex voorstelt. y_t is de waargenomen (of gesimuleerde) tijdreeks; T_t is een trend; S_t is een periodieke, seizoens- of cyclische component; $f(u_t)$ representeert de invloed van de exogene inputsignalen u_t ; N_t is een stochastische component (vaak gemodelleerd middels autoregressieve AR modellen of autoregressieve moving average ARMA modellen) en e_t is een irreguliere component die vaak wordt gemodelleerd als een normaal verdeeld witte ruis proces.

Zoals reeds gezegd kunnen de diverse componenten beschreven worden via tijdvariërende parameters, indien de waargenomen tijdreeksen niet-stationair zijn

Belangrijk aspect bij het construeren van de lage orde modellen is de validatie-stap, waarbij op basis van 'onafhankelijke' data de validiteit van het verkregen model beschouwd wordt (zie ook Young, 2001; Young, 2001; Young, McKenna et al., 2001).

Voor de meeste toepassingen zullen niet alle componenten van het algemene model nodig zijn. Het is zelfs zo dat in zijn meest algemene vorm dit model niet volledig identificeerbaar is.

Voor de invloed van de inputs wordt vaak gebruik gemaakt van lineaire en niet-lineaire transferfunctiemodellen. Er is software voor de methodiek beschikbaar in de vorm van een Matlab toolbox (Captain Toolbox). Deze omvat o.a. modelstructureselectie, schattingsmethoden, en validatie.

De gebruikte schattingsmethoden leveren ook weer een onzekerheidsband op voor de parameters van de diverse componenten van het lage orde model, wat in een later stadium kan worden gebruikt voor nadere analyse van het model.

²¹ D.w.z. herhaalde modelsimulaties voor getrokken waarden van inputs en modelparameters.

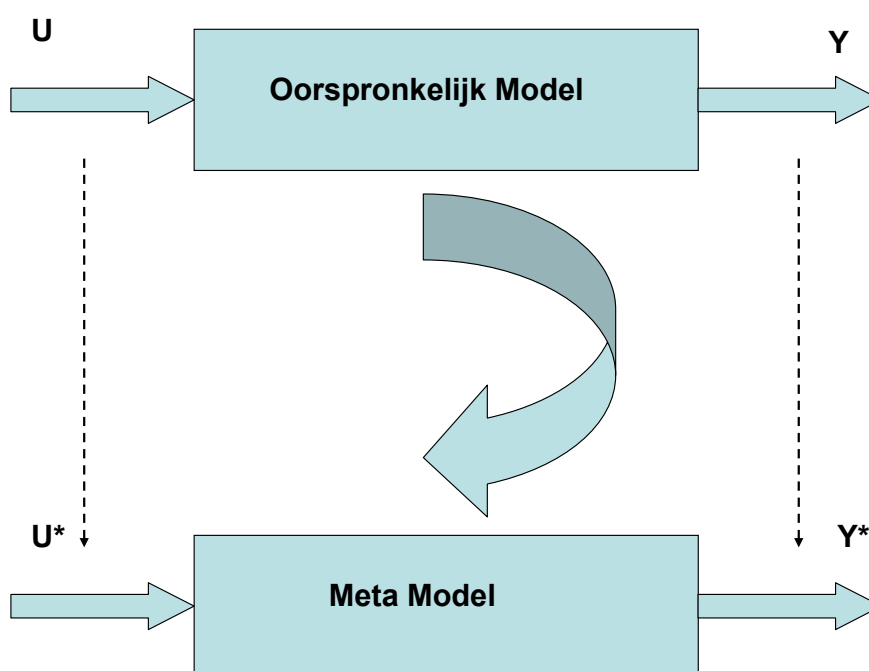
Zie Young (1998) voor een aantal voorbeelden van de methodiek, o.a. voor Global Carbon Cycle klimaatmodellen, neerslag-stroom modellen, analyse van historische klimaatgegevens etcetera

A.1.3. Analytische approximaties

Via analytische uitwerking van prototype-modelsituaties (zie bijvoorbeeld Jury, 1982; Van der Zee en Boesten, 1991) krijgt men inzicht in relevante modelrelaties die gebruikt kunnen worden om onderdelen van complexere modellen te benaderen. Vaak worden deze analytische prototypen afgeleid voor ideale condities of voor geaggregeerde situaties, bijvoorbeeld geen verticale heterogeniteit en steady state stroming, en een beperkt aantal processen. Toepassing in reële situaties vereist dan ook voorzichtigheid (zie Tiktak et al., 2002): enerzijds is er vaak calibratie nodig om de 'effectieve modelparameters' in de analytische meta-relaties adequate waarden te geven; anderzijds hebben de meta-relaties vaak te veel betrekking op geaggregeerde situaties om de diversiteit en interactie van processen die in werkelijkheid kunnen optreden adequaat weer te geven. Approximaties waarbij explicieter rekening gehouden wordt met deze heterogeniteit zijn mogelijk (vergelijk Acharya, 2004; Acharya, Van der Zee et al., 2005).

A.2. Statistiek-gebaseerde modelreductie

De crux van deze methoden bestaat er uit dat het oorspronkelijk referentiemodel als een soort ‘virtueel laboratorium’ gebruikt wordt, waarmee experimenten worden uitgevoerd die informatie opleveren over het gedrag van het referentiemodel. Het referentiemodel wordt hierbij als een ‘data-genererend mechanisme’ gebruikt, waaraan ‘inputs’²² (u) worden aangeboden en dat ‘outputs’ (y) uitrekt. Hierbij is doorgaans niet sprake van *alle* te onderscheiden inputs en outputs, maar van een selectie, zeg u^* , y^* , die voor de betreffende toepassing relevant geacht wordt. De relatie tussen deze inputs en outputs wordt geanalyseerd aan de hand van statistische methodes, en dit leidt tot de afleiding van een metamodel voor de betreffende input-output selectie (zie Figuur 9).



Figuur 9: Het oorspronkelijk model en het daarvan afgeleide metamodel met de bijbehorende input-output selectie.

Het verkregen metamodel hangt af van de keuze van de inputs en outputs en van de gebruikte statistische methode. In de onderstaande paragrafen bespreken we kort verschillende methoden die hiervoor gebruikt kunnen worden.

²² Met de generieke term ‘inputs’ beschrijven we een klasse van grootheden, die zowel modelparameters, begincondities, input functies of andere factoren kan omvatten. Inputs betreffen de ‘predictoren’; Outputs betreffen de ‘responsen’, of ‘predicties’.

A.2.1. Interpolatie in Look-Up Tables

Het idee is om voor een aantal parameterinstellingen modelsimulaties te berekenen, hiermee ‘opzoek-tabellen’ met berekeningsresultaten te creëren en deze resultaten vervolgens te gebruiken indien men modelberekeningen voor andere parameterwaarden nodig heeft. De uitkomsten hiervan worden dan bijvoorbeeld via opzoeken of interpolatie uit de opzoektabellen bepaald. In wezen is deze aanpak verwant met numerieke interpolatiemethoden (zie Stoer en Bulirsch, 1983).

In de praktijk heeft men te maken met een groot aantal keuzes, zoals: keuze van te variëren parameters en van parameterinstellingen (experimentdesign); keuze van de outputgegevens die men evalueert en opslaat; keuze van de interpolatie methode (bijvoorbeeld nearest neighborhood, (lineaire) interpolatie). De te maken keuzes zullen onder andere afhangen van de rekentijd en verdere gedragskarakteristieken (bijvoorbeeld niet-lineairiteiten) van het referentiemodel, terwijl ook aspecten als het gewenste toepassingsbereik en de gewenste nauwkeurigheid van het metamodel; beschikbaarheid van gegevens (input-parameters) voor het metamodel etcetera een rol spelen. Een gevoeligheidsanalyse van het referentiemodel kan behulpzaam zijn om te screenen welke parameters per sé betrokken moeten worden bij het maken van de ‘opzoek-tabellen’, met name omdat de ‘opzoek-tabellen’-methode last heeft van de ‘curse of dimensionality’ (dat wil zeggen het aantal door te rekenen combinaties groeit zeer snel met het aantal te onderscheiden input-parameters). In hoofdstuk 10 van Nelles (2000) wordt uitgebreider ingegaan op voor- en nadelen van deze methode.

A.2.2. Response Surfaces (polynoom regressie)

Polynoom-regressie modellen leveren de eenvoudigste manier van functie-approximatie, met hogere orde polynomen als voor de hand liggende uitbreiding van lineaire modelrelaties. De relatie tussen ‘inputs’ en de corresponderende model-‘output’ wordt bijvoorbeeld uitgedrukt als kwadratisch (dat wil zeggen tweede-orde) polynoom²³ :

$$\hat{y} = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_i \sum_{j<i} \beta_{ij} x_i x_j \quad (1)$$

Gebruiken we de resultaten van de verschillende modelsimulatie runs, dan kunnen de parameters β_i , β_{ij} in dit polynoom model bijvoorbeeld op basis van kleinste kwadraten regressie gefit worden.

Ook kan het belang van de verschillende factoren x_i in het regressiemodel worden ingeschat op basis van statistische toetsen²⁴ (dat wil zeggen verschilt β_i bijvoorbeeld significant van 0 in het genormaliseerde regressiemodel). Om een grote rekeninspanning te vermijden, met name bij grote dimensies (dat wil zeggen grote aantallen inputs, k), is het belangrijk om een voorselectie te maken van de meest belangrijke inputs. Hiervoor kunnen *screening* technieken gebruikt worden (zie bijvoorbeeld Saltelli, Chan et al., 2000). Daarmee kan ook overfitting vermeden worden; dit probleem is ook deels te voorkomen worden door expliciet penalty-termen aan het kwadratisch fit-criterium toe te voegen (zie onder andere Hastie, Tibshirani et al., 2001).

²³ We beperken ons bij de presentatie tot de situatie met 1 uitgang (uni-variaat). Generalisaties naar meerdere uitgangen zijn mogelijk, maar maken de presentatie en notatie snel ingewikkeld. Idem w.b. generalisaties naar hogere orde polynomen.

²⁴ Hierbij is echter een waarschuwing op zijn plaats: de ‘data’ X, Y zijn doorgaans deterministisch omdat ze vaak afkomstig van simulatie met deterministische modellen. Strikt genomen gelden de gangbare voorwaarden van statistische hypothesetoetsen dan niet, en is voorzichtigheid geboden bij het toepassen van statistische tests. Zie voor meer achtergrond de discussie rond ‘challenge 1’ in Simpson et al. (2004).

Daarnaast is bij bovenstaande Polynoom Regressie de keuze van de parametersettings x_i die bij constructie van het metamodel gebruikt worden van belang (zie Kleijnen, Sanchez et al., 2004; Simpson, Booker et al., 2004; Kleijnen, 2005 voor meer informatie en verdere referenties over dit ‘*experimental design*’ probleem). In Kleijnen, Sanchez et al. (2004) worden methoden genoemd om het verkregen metamodel te valideren.

Een groot nadeel van Polynoom Regressie is dat sterk niet-lineair gedrag vaak slecht gemodelleerd kan worden: óf men heeft snel last van oscillaties en instabiliteiten, met name bij hoge orde polynomen en als men het metamodel gaat gebruiken buiten het gebied waarop het gefit is, óf het aantal samples/runs dat men moet doorrekenen om het gedrag goed te benaderen is erg groot (Barton, 1992; Nelles, 2000 hoofdstuk 10).

Enkele alternatieven

Alternatieven zijn bijvoorbeeld gezocht in het gebruik van *stuksgewijze polynomen*, waarbij lage-orde polynomen op deelgebieden van de parameterange als benadering gebruikt worden, en zodanig aan elkaar worden ‘geplakt’ dat het resultaat een gladde functie is (*splines*, *MARS*, zie later). Dit levert doorgaans een robuustere approximatie op, en vermijdt voor een deel grote afwijkingen buiten de calibratierange.

Ook het gebruik van zogenaamde speciale basis functies ($h_i(x)$) als grondslag voor een eenvoudige benadering/expansie van de outputs ($Y = \sum \beta_i h_i(x)$) in termen van expliciete functies van de inputs kan de flexibiliteit in het modelleren verhogen. Basis functies als ‘*Gaussische radial basis functions*’ (zie later) en ‘*wavelets*’ zijn populaire keuzes (zie ook Hastie, Tibshirani et al. (2001), §5.8, §5.9, §6.7). Wavelets zijn met name geschikt in situaties waarbij verscheidene signalen met sterk variërende tijdschalen een rol spelen (‘*multi-resolution*’; zie Daubechies, 1990; Daubechies, 1992).

Een expansie die expliciet hogere orde en kruistermen meeneemt, en sterk verwant is aan de statistische variantie-analyse ANOVA, is de *HDMR High Dimensional Model Representation* (Rabitz en Alis, 1999; Rabitz, Alis et al., 1999; Li, Rosenthal et al., 2001):

$$Y = f(x) = f_o(x) + \sum_{i=1}^n f_i(x_i) + \sum_{1 \leq i < j \leq n} f_{ij}(x_i, x_j) + \dots + f_{12\dots n}(x_1, x_2, \dots, x_n) \quad (5)$$

waarbij f_o , $f_i(x_i)$, $f_{ij}(x_i, x_j)$ duiden op de gemiddelde response, de onafhankelijk bijdrage (main effect), de paarsgewijs gekoppelde bijdrage (two-way interacties) etcetera tot $f(x)$ van de diverse inputs. Deze component functies van de HDMR expansie worden optimaal geconstrueerd voor de $f(x)$, zodat de expansie snel convergeert. Vaak levert een tweede orde benadering al een adequate beschrijving. In de praktijk zijn verschillende manieren denkbaar om een dergelijke decompositie tot stand te brengen (bijvoorbeeld cut-HDMR, RS-HDMR, RKHS-HDMR; zie Alis en Rabitz, 2001; Li, Rosenthal et al., 2001; Ho en Rabitz, 2003; Li, Artamonov et al., 2003). De methodiek wordt behalve voor modelreductie (zie bijvoorbeeld Wang, Georgopoulos et al., 2003) met name ook gebruikt bij het uitvoeren van gevoeligheids- en onzekerheidsanalyses (Rabitz en Alis, 2000; Li, Rabitz et al., 2002; Li, Wang et al., 2002; Li, Rabitz et al., 2003) omdat de coëfficiënten in de verkregen expansie handzame informatie leveren over de bijdrage van de diverse onzekerheidsbronnen. Ook leveren de resulterende input-output metamodelrelaties een nuttig uitgangspunt voor optimaliserings- en designvraagstukken bij rekenintensieve modellen (Banerjee en Ierapetritou, 2002; Banerjee en Ierapetritou, 2003).

Een ander alternatief betreft een stochastische variant van de reponse surface methode, de zogenaamde *SRSM (Stochastic Response Surface Method)* van Isukapalli, Roy et al. (1998), waarbij de input-parameterverdeling allereerst naar een standaard multinormale verdeling wordt getransformeerd, en de model outputs vervolgens als reeksontwikkeling van Hermite polynomen in de getransformeerde inputs worden uitgedrukt. Hiermee kan een zuinig

geparametriseerde modelvorm verkregen worden, die het mogelijk maakt om, via koppeling met automatische differentiatie technieken, efficiënt gevoeligheids- en onzekerheidsstudies te doen (zie ook Isukapalli, Roy et al., 2000; Isukapalli en Georgopoulos, 1999; Isukapalli en Georgopoulos, 2001 en de web-interface voor de beschikbare software voor SRSM: <http://www.ccl.rutgers.edu/srsm.htm>).

Een verwante techniek, die ook gebruik maakt van orthogonale polynoom expansies om onzekerheidspropagatie op een zuinige wijze te bepalen is de *DEMM (Deterministic Equivalent Modeling Method)* of *Probabilistic Collocation Method*²⁵ (zie Tatang, Pan et al., 1997). Toepassingsstudies voor het uitvoeren van gevoeligheids- en onzekerheidsanalyses aan de hand van DEMM zijn gerapporteerd in Webster, Tatang et al., 2000; Webster en Sokolov, 2000; Cryer en Applequist, 2003; Cryer en Applequist, 2003; Lucas en Prinn, 2004. Isukapalli, Roy et al. (1998), hebben de DEMM met de SRSM vergeleken en beweren dat SRSM robuuster is en nauwkeuriger schattingen van de onzekerheden oplevert.

A.2.3. Het Kriging model/ Gaussische Processen

Oorspronkelijk ontwikkeld in de mijnbouw en de geostatistiek, is de Kriging methode een aanpak voor curve-fitten, interpolatie en respons-oppervlak-approximatie (zie onder andere Cressie, 1993). De relatie tussen inputs en outputs wordt beschreven in termen van:

$$\hat{y} = \sum \beta_j f_j(x) + Z(x) \quad (6)$$

waarbij $f_j(x)$ bekende regressie modellen zijn en β_j hun corresponderende parameters/coëfficiënten. Verondersteld wordt dat de functie $Z(x)$ een realisatie van een stochastisch proces is met gemiddelde nul en met spatiële covariantie gelijk aan:

$$\text{Cov}[Z(x_i), Z(x_j)] = \sigma^2 R(x_i, x_j) \quad (7)$$

waarbij σ^2 de variantie is en $R(.,.)$ de spatiële correlatie functie van het stochastisch proces Z is. Verschillende vormen van correlatiefuncties kunnen hierbij gekozen worden; vaak wordt de Gaussische correlatie functie gebruikt.

Op een bepaalde manier is $f_j(x)$ het 'globale' model over de hele design-ruimte, gefit uit de corresponderende modelrespons voor het referentie-model, terwijl $Z(x)$ de 'locale' afwijking van die globale vorm voorstelt, die er voor zorgt dat in de design-punten interpolatie plaatsvindt. De Kriging benadering levert de beste lineaire, zuivere predictor van Y in de design-punten x . Doorgaans worden de 'spatiële' functies $f_j(x)$ als constante termen gekozen, maar andere keuzes zijn ook denkbaar. De parameters in het model worden vaak bepaald op basis van maximum likelihood-schatting²⁶. Dit vereist niet-lineaire optimalisatie, en wordt normaliter op een iteratieve manier uitgevoerd. Deze optimalisatie kan veel rekentijd vergen, met name als de gesamplede dataset groot is (zie Jin, Chen et al., 2001; Simpson, Mauery et al., 2001).

Kriging methoden zijn flexibel, omdat er met een groot scala correlatie-functies gewerkt kan worden; bovendien kunnen ze nauwkeurige predicties/approximaties leveren van sterk niet-lineair of onregelmatig gedrag. Ook kunnen ze oscillatoire (multi-modale²⁷) responsies benaderen, iets dat polynoomregressie slechts in beperkte mate kan. Het grootste nadeel van de Kriging methode is dat het bepalen van het optimale Kriging model erg rekenintensief kan zijn, en dat er numerieke problemen (instabiliteiten) kunnen ontstaan als de gesamplede

²⁵ De coëfficiënten in de reeksexpansie worden bepaald door een stelsel lineaire vergelijkingen op te lossen. Dit stelsel is verkregen door het referentiemodel in zogenaamde collocatiepunten te evalueren, die in dit geval corresponderen met de wortels van de orthogonale Hermite polynomen

²⁶ Ook is er een Bayesiaanse schattingsopzet mogelijk.

²⁷ D.w.z. met meerdere maxima of minima.

punten te dicht bij elkaar liggen en de correlatie matrix bijna singulier wordt (Meckesheimer, Barton et al., 2001; Meckesheimer, Booker et al., 2002). Tegenwoordig is software voor het uitvoeren van Kriging in diverse statistische pakketten voorhanden (zie bijvoorbeeld <http://www.imm.dtu.dk/~hbn/dace/> voor een free-domain MATLAB toolbox). Voor informatie over experimenteel ontwerp voor Kriging, (zie Kleijnen, Sanchez et al., 2004; Kleijnen en van Beers, 2004; Simpson, Booker et al., 2004; Van Beers en Kleijnen, 2003; Van Beers en Kleijnen, 2004. In Kleijnen en Sargent (2000); Bayarri, Berger et al. (2002) en Easterling en Berger (2002) worden methoden voor validatie van metamodellen besproken, die ook voor Kriging modellen van belang zijn. Zie ook Martin en Simpson, 2003; Martin en Simpson, 2004.

De bovengenoemde manier van modelbenadering wordt ook wel *DACE* genoemd (*Design and Analysis of Computer Experiment*), een term die door Sacks, Schiller et al. (1989); Sacks, Welch et al. (1989) gelanceerd is. DACE heeft op het gebied van optimaal ontwerpstudies een hausse van toepassingen veroorzaakt met name in de werktuigbouwkundige en civiel-technische designstudies (zie Morris en Mitchell, 1992; Welch, Buck et al., 1992; Simpson, Pehlinski et al., 2001; Eldred, Giunta et al., 2004; Simpson, Booker et al., 2004).

Ook is door O'Hagan, Kennedy et al. (1998) een sterk verwante methodiek voor metamodellering geïntroduceerd, op basis van Gaussische processen/velden (GP) waarmee gladde functionele verbanden adequaat benaderd kunnen worden (zie ook Gibbs en MacKay, 1997; MacKay, 1997; Daberkow, 2002; Daberkow en Mavris, 2002). De GP-methodiek is uitgewerkt in een Bayesiaanse setting, en de 'model-emulator' die zo verkregen wordt kan bijvoorbeeld verder gebruikt worden bij calibratie en gevoeligheids/onzekerheidsanalyse van rekenintensieve modellen (zie Kennedy en O'Hagan, 2001; Kennedy, O'Hagan et al., 2001; Oakley, 2002; Oakley en O'Hagan, 2002; Oakley en O'Hagan, 2004). Recentelijk is software beschikbaar gesteld door Kennedy waarmee problemen tot en met 30 parameters geanalyseerd kunnen worden. Zie <http://www.sheffield.ac.uk/st1mck/code.html> en het verwante werk van Craig, Goldstein et al. (2001); Goldstein en M.Rougier (2003); Goldstein en Rougier (2005) en de URL <http://maths.dur.ac.uk/stats/bayeslin/>. Zie verder de website van Mark Gibbs (<http://wol.ra.phy.cam.ac.uk/mng10/GP/GP.html>) en David MacKay (<http://wol.ra.phy.cam.ac.uk/mackay/README.html#1997>) voor meer informatie en software, evenals de appendices in (Daberkow, 2002).

A.2.4. GLMs en GAMs (Generalized Linear Models en Generalized Additive Models)

Deze methoden betreffen uitbreidingen van het 'klassieke' lineaire regressiemodel, dat veel gebruikt wordt in de statistiek.

Bij lineaire modellen wordt doorgaans verondersteld dat de relatie tussen de random variabelen X en Y lineair is, en dat de componenten van Y onafhankelijke normaal verdeelde variabelen zijn met constante variantie σ^2 en verwachting:

$$E(Y | X_1, \dots, X_p) = \mu \quad \text{met} \quad \mu = X\beta = X_1\beta_1 + \dots + X_p\beta_p \quad (8)$$

Hierin is β de parametervector en in de praktijk zal men vaak proberen deze vector te schatten op basis van waarnemingen van X en Y.

Voor 'Generalized Linear Models' (GLMs) wordt de modelveronderstelling uitgebreid tot:

$$\begin{aligned} E(Y | X_1, \dots, X_p) &= \mu \\ \eta &= X\beta \\ \eta_i &= g(\mu_i) \end{aligned} \quad (9)$$

waarbij η de *lineaire predictor* is en de functie $g(\cdot)$ de monotoon differentieerbare *link-functie* is, die beschrijft hoe de verwachtingswaarde van Y gerelateerd is aan de lineaire predictor. Verder wordt verondersteld dat de response variabelen Y een verdeling hebben die tot de exponentiële familie behoort, waaronder o.a. de normale, de Poisson, de binomiale en de gamma verdeling vallen. De variantie van de response variabele hangt hierbij als volgt van het gemiddelde af:

$$\text{Var}(Y_i) = \frac{\varphi \cdot V(\mu_i)}{w_i} \quad (10)$$

waarbij $V(\cdot)$ de variantie-functie is, φ een dispersie parameter is en w_i het bekende gewicht voor elke observatie. φ is òf bekend (bijvoorbeeld voor de binomiale of de Poisson verdeling is $\varphi=1$), òf moet geschat worden. Merk op dat het oorspronkelijke lineaire model als bijzonder geval tevoorschijn komt als de eenheids-link functie $g(\mu)=\mu$ met normaal verdeelde response variabelen gekozen wordt.

Voor de link functie zijn er diverse mogelijkheden, waarvan er een aantal extra bekend zijn:

- | | |
|---------------------------|---|
| 1. Logit : | $\eta = \log[\mu/(1-\mu)]$ |
| 2. Probit: | $\eta = \Phi^{-1}(\mu)$ |
| | waarbij $\Phi(\cdot)$ de normale cumulatieve distributie functie is |
| 3. Power family: | $\eta = \mu^\lambda$ voor $\lambda \neq 0$ |
| | $\eta = \log(\mu)$ voor $\lambda = 0$ |
| 4. Log: | $\eta = \log(\mu)$ |
| 5. Complementary log-log: | $\eta = \log[-\log(1-\mu)]$ |

Voor het schatten van de parameters (β) wordt gebruik gemaakt van de maximum likelihood methode. Voor verdere details zie McCullagh en Nelder (1989). Zie de website <http://www.statsci.org/glm/software.html> voor informatie over statistische software voor GLMs.

In feite betreft GLM een deelklasse van niet-lineaire regressie met speciale eisen aan de niet-lineariteit, waardoor het mogelijk wordt om gebruik te kunnen maken van bepaalde nuttige aspecten uit de lineaire regressie analyse.

Bij de voorgaande vormen van niet-lineaire regressie wordt vaak verondersteld dat de vorm van de niet-lineariteit bekend is. Als we deze conditie loslaten, en de vorm van de niet-lineariteit *onbekend* veronderstellen, dan krijgen we een verdere generalisatie waarbij de niet-lineaire relatie niet-parametrisch uit de gegevens geschat moet worden. Zo'n generalisatie wordt bijvoorbeeld geleverd door de zogenaamde 'Generalized Additive Models (GAMs)'. GAMs is hierbij een uitbreiding op *additieve modellen* net zoals GLMs een uitbreiding van *lineaire modellen* vormt:

Het *additief model* is een verruiming van het lineair model waarbij verondersteld wordt dat de relatie tussen Y en de X_i gegeven is als (onbekend) additief effect van de afzonderlijke predictors X_i (ook interactie-effecten of synergieën kunnen door dit model beschreven worden door bijvoorbeeld producten $X_i X_j$ expliciet mee te nemen):

$$E(Y | X_1, \dots, X_p) = \mu \quad \text{met} \quad \mu = \alpha + f_1(X_1) + \dots + f_p(X_p) \quad (11)$$

waarbij de f_1, \dots, f_p onbekende gladde functies zijn, in niet-parametrische vorm.

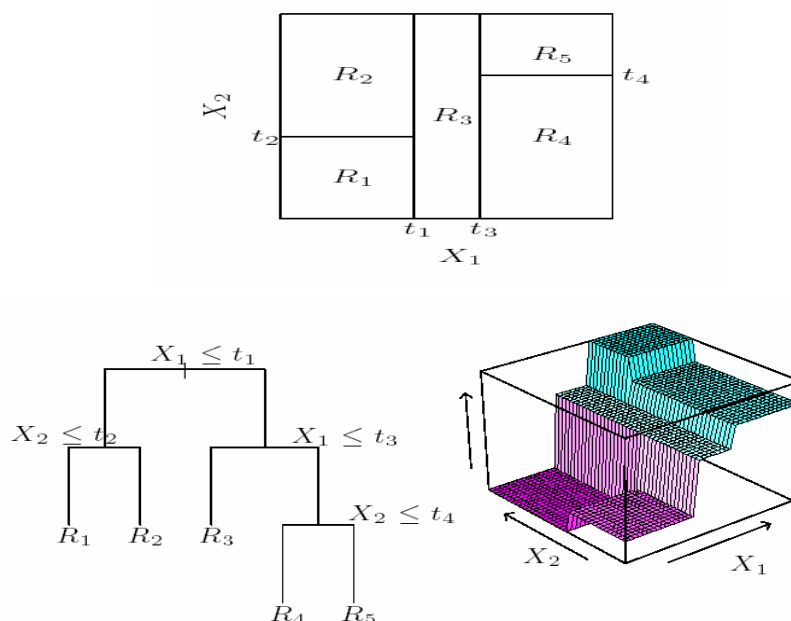
Het corresponderende *GAM* (*Generalized Additive Model*) wordt nu – geheel in analogie met de GLM - via een link-functie $g(\cdot)$ gegeven door:

$$\begin{aligned} E(Y | X_1, \dots, X_p) &= \mu \\ \eta &= \alpha + f_1(X_1) + \dots + f_p(X_p) \\ \eta_i &= g(\mu_i) \end{aligned} \quad (12)$$

Zie Hastie en Tibshirani (1990) voor details. GAMs zijn vooral nuttig voor exploratieve data-analyse, met name wanneer de onderzoeker weinig notie heeft van de specifieke functionele vorm van de relatie tussen de predictors en de respons. Via smoothing-achtige technieken schat GAMs niet-parametrische functionele relaties die doorgaans flexibeler zijn dan de parametrische variant die bijvoorbeeld door Response surface modellen geleverd wordt. GAMs kan een opstap leveren naar verdere analyse, bijvoorbeeld aan de hand van MARS. Zie met name ook Hastie, Tibshirani et al. (2001) voor meer informatie (vergelijk ook <http://jackman.stanford.edu/papers/gam.pdf>).

A.2.5. CART (Classification And Regression Trees)

CART²⁸ is een door (Breiman, Olshen et al., 1984) voorgestelde niet-parametrische regressietechniek waarbij de ruimte die door de invoervariabelen (predictor variabelen) wordt opgespannen, systematisch in regio's (veelal rechthoekige vlakken) wordt onderverdeeld terwijl de response Y door een simpel model (bijvoorbeeld een constante) in elk van deze vlakken gefit wordt. Op deze manier kan een conceptueel eenvoudige en vaak krachtige meta-relatie verkregen worden.



Figuur 10: Partities en CART: Het bovenste paneel laat de partitie van de twee-dimensionale predictor ruimte via recursief binair splitsen zien. De onderste panelen laten de corresponderende regressieboom zien en de plot met het resulterende predictieoppervlak. Vrij naar Figure 9.2 in Hastie, Tibshirani et al. (2001).

²⁸ In feite houden 'tree-based' methoden zich bezig met het verdelen van observaties/data in groepen die verschillen m.b.t. van belang zijnde variabele die van belang is.

Het proces van het creëren van zo'n partitie in de ruimte van invoervariabelen kan in de vorm van een beslisboom met eenvoudige beslisregels in termen van de invoervariabelen beschreven worden. De boom wordt een *classificatieboom* genoemd als de response variabele y kwalitatief is, en anders, als de responsie *kwantitatief* is, wordt het een *regressieboom* genoemd. De beginknoop van de boom wordt de wortel ('the root') genoemd, en het model wordt gefit via *binair recursief partitioneren*. Dit betekent dat de invoerdata successievelijk wordt gesplitst in linker en rechter 'vakken' waarbij de opsplitsregels ('*splitting rules*') gedefinieerd worden op basis van de predictor variabelen X . We splitsen de ruimte initieel in twee vakken, en modelleren de respons door het gemiddelde van Y te nemen in elk vak. Hierbij wordt de te splitsen variabele en het splitsingspunt zodanig gekozen dat de beste fit (van Y door zijn gemiddelde) bereikt wordt. Vervolgens wordt 1 of beide vakken weer in twee subvakken gesplitst, en het proces wordt voortgezet tot een bepaalde stopregel actief wordt. In Figuur 10 is een situatie geschetst waarbij we een continue respons Y hebben en twee inputs X_1 en X_2 die waarden aannemen op het eenheidsinterval: Een eerste opsplitsing gebeurt bij $x_1 < t_1$ waarbij t_1 nader geschat dient te worden. Dan is $\hat{f}(x) = a_1$ voor $x_1 \leq t_1$ en $\hat{f}(x) = a_2$ voor $x_1 > t_1$. Dan wordt het vlak $x_1 \leq t_1$ gesplitst bij $x_2 = t_2$ en het vlak $x_1 > t_1$ bij $x_1 = t_3$. Uiteindelijk wordt het vlak $x_1 > t_3$ gesplitst bij $x_2 = t_4$. Op deze manier resulteert een partitie in vijf regio's, die in Figuur 10 is aangeduid. Het corresponderende regressie model 'voorspelt' Y met een constante c_m in de betreffende regio R_m . Dit model kan ook worden weergegeven door de binaire boom in het linksonder-paneel van Figuur 10. Bij het maken van deze opsplitsing worden een groot aantal punten in het eenheidsvlak (inputparameter ruimte voor X_1 en X_2) gegenereerd, waarvoor de corresponderende responsies Y worden uitgerekend. Door de opsplitspunten zo te kiezen dat ze de 'waarde' van de split maximaal²⁹ maken, wordt de boom uitgebreid, totdat het punt bereikt is waarop de responsies allen dezelfde zijn binnen een knooppunt, of de data te beperkt is voor verdere opsplitsing. Bij het eindknooppunt is de respons waarde gelijk aan het gemiddelde of de meest voorkomende van de respons waarden voor continue resp. discrete grootheden. Het verder snoeien ('*pruning*') van de takken van de boom om overfitting te vermijden kan op verschillende manieren plaatsvinden. Snoeien op basis van kruisvalidatie is het meest populair (zie Hastie, Tibshirani et al., 2001).

In feite bepaalt CART een stuksgewijs constante benadering van het response oppervlak, en hoewel CART op een eenvoudige wijze vaak krachtige meta-relaties oplevert waarbij geen condities gesteld zijn aan de verdeling van de data, kleven er een aantal bezwaren aan: de resulterende opsplitsing van de regressieboom kan erg gevoelig zijn voor veranderingen in gegevens, hetgeen zich manifesteert in een grote variantie in de schattingen. Ook is het resulterend metamodel niet-glad/discontinu, hetgeen zijn performance beperkt. Dit gebrek aan gladheid kan opgeheven worden door de MARS procedure, die in zekere zin als modificatie van CART kan worden opgevat. Tenslotte blijkt ook het binaire partitie-principe van CART moeite te hebben met additieve structuren, iets waar MARS minder last van heeft (zie Hastie, Tibshirani et al., 2001).

A.2.6. MARS (Multivariate Adaptive Regression Splines)

Multivariate adaptieve regressie splines (MARS) (Friedman, 1991) is een flexibele niet-parametrische regressie methode die een stap verder gaat dan CART en de stuksgewijs constante functies van CART generaliseert tot gladde continue functies. Dit gebeurt door (multivariate) splines te fitten in de regio's R_m , en hun waarden met elkaar te verbinden op de grenzen van R_m . Hierbij wordt adaptief, uitgaande van stuksgewijs lineaire functies en hun producten, een verzameling basis functies geselecteerd om de response functie te benaderen.

²⁹ Voor classificatie problemen is de 'waarde' uitgedrukt als de reductie van de 'onzuiverheid' in het knooppunt, en voor regressieproblemen is het de 'reductie van de residuele kwadraten som' van de fit op Y .

Rekentechnisch verloopt dit via een voorwaartse en achterwaartse iteratieve procedure (voor details zie Friedman, 1991; Hastie, Tibshirani et al., 2001). Het achterliggende metamodel is:

$$\hat{y} = \sum_{m=1}^M a_m B_m(x) \quad (13)$$

waarbij a_m de expansie coëfficiënten zijn, en B_m de basis functies zijn van de vorm:

$$B_m(x) = \prod_{k=1}^{K_m} [s_{k,m}(x_{v(k,m)} - t_{k,m})]_+^q \quad (14)$$

met K_m het aantal factoren (interactie-orde), terwijl s en t de helling (± 1) en steunpunt coëfficiënten zijn, en $X_{v(k,m)}$, $1 \leq v(k,m) \leq n$ de betreffende inputparameter is, met q de orde van het stuksgewijze polynoom. De subscript '+' duidt op de getrunceerde vorm van de functie tussen de haken: $[F]_+$ is gelijk aan F als $F > 0$, en 0 anders.

De stuksgewijze polynomen in de basis-expansie leveren locale beschrijvingen, die het mogelijk maken om het response oppervlak zuiniger te beschrijven dan met globale polynomen. MARS is een middel om niet-lineariteiten van verschillende aard te modelleren, waarbij het aantal termen dat wordt meegenomen de nauwkeurigheid en de rekenkosten van het metamodel bepalen.

A.2.7. Radial Basis Functions (RBF)

Een algemene methodiek om niet-lineaire responsies te benaderen verloopt via het gebruik van *radiële basis functies*, waarmee *locaal* benaderingen kunnen worden uitgevoerd van de responsies. De output kan hierbij worden uitgedrukt als *lineaire* combinaties van radiaal symmetrische functies, de zgn. *basis functies* of '*kernels*', gebaseerd op een afstandsmetrick (bijvoorbeeld de Euclidische afstand):

$$\hat{y} = \sum_{j=1}^M D\left(\frac{\|x - x_{oj}\|}{\lambda_j}\right) \cdot \beta_j \quad (15)$$

waarbij de basis functies $D(\cdot)$ geïndexeerd zijn door een locatie- x_{oj} en een schaal parameter λ_j . Voor $D(\cdot)$ wordt vaak de standaard Gaussische dichtheidsfunctie gekozen, en er kunnen diverse aanpakken worden gehanteerd om de parameters te schatten c.q. te leren, bijvoorbeeld kleinste kwadraten. Zie § 6.7 in ook Hastie, Tibshirani et al. (2001) voor meer informatie. 'Radial basis function' approximaties geven een goede fit voor willekeurige contouren van zowel deterministische als stochastische response functies (Powell, 1987; Orr, 1995).

A.2.8. Neurale Netwerken

Neurale Netwerken (NN) hebben hun wortels in zowel de kunstmatige intelligentie als in de statistiek, en leveren een black-box statistische techniek voor algemene niet-lineaire regressie, waarmee willekeurige niet-lineaire modelresponsies benaderd kunnen worden (zie Cheng en Titterington, 1994); zie Lek en Guegan (1999) voor een overzicht van NN toepassingen binnen ecologisch modelleren en data-analyse). Het centrale idee is om lineaire combinaties van inputs als een soort afgeleide kenmerken ('features') uit de input/output data te extraheren, en de model response te zien als niet-lineaire functie van deze 'features'. Dit wordt vaak gerealiseerd als een twee-stappen regressie model, gepresenteerd in de vorm van

een netwerk diagram³⁰, waarbij naast de input en output laag een verborgen laag van tussenvariabelen, of afgeleide kenmerken wordt onderscheiden. Deze zijn een niet-lineaire functie van lineaire combinaties van de inputs:

$$\begin{aligned} Z_m &= \sigma(\alpha_{0m} + \alpha_m^T X), & m &= 1, \dots, M, \\ T_k &= \beta_{0k} + \beta_k^T Z, & k &= 1, \dots, K, \\ f_k(X) &= g_k(T), & k &= 1, \dots, K \end{aligned} \quad (16)$$

waarbij $Z=(Z_1, \dots, Z_M)$ en $T=(T_1, \dots, T_K)$ de tussenvariabelen of ‘features’ zijn, terwijl de niet-lineaire activeringsfunctie $\sigma(v)$ òf als sigmoïde $\sigma(v)=1/(1+e^{-v})$ gekozen wordt (Sigmoid Neural Networks), òf vaak ook als Gaussische basis functie $\frac{1}{\sqrt{2\pi}} \cdot e^{-v^2/2}$ (Radial Basis Function Networks).

In een regressie-context wordt de output functie $g_k(t)$, doorgaans gelijk gekozen aan de identiteit $g_k(t)=t_k$ en is de output dus een lineaire combinatie van de features Z_i . Deze features worden ook wel ‘hidden units’ genoemd, omdat hun waarden niet rechtstreeks worden gemeten; ze kunnen gezien worden als een soort basis ‘expansie’ (reeksontwikkeling) van de originele invoer ruimte, waarbij deze expansie bepaald wordt uit de data. De verzameling gewichten α_{0m} , α_m , β_{0k} , β_k wordt doorgaans zo gekozen dat de kwadraten som:

$V(\theta) = \sum_{k=1}^K \sum_{i=1}^N (y_{ik} - f_k(x_i))^2$ minimaal is. Bij minimalisatie van deze kwadraten som is

informatie over de gradiënt vereist die via een ‘back-propagation algoritme’ efficiënt bepaald kan worden. Doorgaans zijn gradiënt methoden voor optimalisatie traag, en is verbetering mogelijk door gebruik te maken van tweede orde optimalisatietechnieken. Met name zijn de geconjugeerde gradiënt en variabele metriek methoden geschikt, omdat hierbij de rekenintensieve expliciete bepaling van de tweede orde afgeleide (Hessiaan) vermeden wordt. Om overfitting ten gevolge van het groot aantal parameters te vermijden wordt er ook nog een penalty-term bij het kleinste kwadraten criterium gevoegd (‘regularisatie’). Deze zorgt ervoor dat scherpe pieken en randen vermeden worden en dat het resulterend metamodel een glad verloop krijgt. De keuze van het aantal verborgen eenheden (d.w.z. Z_1, \dots, Z_M) en lagen is ook een belangrijk issue. Ook kent het fit-criterium doorgaans meerdere locale optima, en is het dus nodig om de minimalisatie herhaald toe te passen, telkens met een andere (random gekozen) startwaarde. Voor een effectieve schaling van de gewichten (α , β) is het bovendien van belang om de inputs X_i te schalen/normaliseren. Zie onder andere hoofdstuk 11 in Hastie, Tibshirani et al. (2001) voor nadere informatie over training (calibratie) en ook testen (validatie) van Neurale Netwerk modellen. Neurale Netwerken leveren, verschillend van CART en MARS, gladde functies van reëelwaardige parameters. Voor toepassingen van Neurale Netwerken bij dynamische systemen, zie met name ook Nelles (2000) en Sjöberg, Zhang et al. (1995).

Er bestaan vele vormen van Neurale Netwerken, en er is ook veel software beschikbaar (vergelijk bijvoorbeeld de Neurale Netwerk toolbox van Matlab; zie ook <http://www.faqs.org/faqs/ai-faq/neural-nets/part6/> en <http://www.faqs.org/faqs/ai-faq/neural-nets/part5/>). Gezien deze veelheid, loert bij toepassing het gevaar van inadequate training en

³⁰ We beperken ons hier tot het meest gangbare ‘single hidden layer, feed-forward neural network’, in de vorm van MLPs (Multi Layer Perceptrons) of RBFs (Radial Basis Functions), als representant van ‘supervised learning’ methoden. Andere supervised learning varianten zijn denkbaar als Projection Pursuit Regression, evenals ‘unsupervised learning’ methoden als Self-Organizing Maps (SOM) die o.a. dienen voor efficiënte clustering, visualisatie en abstractie van datasets, zie bijv. Hastie et al. 2001, Haykin, 1999, Lek en Guégan, 1999, Nelles, 2000.

structuur keuze (variabelen, lagen) voor het netwerk. Ook is de methode sterk black-box georiënteerd, en is de structuur van het netwerk weinig transparant (in principe spelen alle inputs X_i mee in alle Z_m , en dragen deze vervolgens weer allen bij tot Y) hetgeen fysisch inzicht en interpretatie in de weg kan staan. NN worden doorgaans als minder effectief beoordeeld voor situaties waarin het primaire doel inzicht in de fysische processen en de rol van de individuele inputs is. Deze situatie kan verbeterd worden door methoden die het belang en de gevoeligheid van input-componenten in het NN vaststellen (zie Gevrey, Dimopoulos et al., 2003; Olden, Joy et al., 2004 voor een overzicht).

A.2.9. Support Vector Machines en Kernel-based learning

SVMs (Support Vector Machines) zijn een recent ontwikkelde patroon herkenningstools, waarbij 'supervised learning'³¹ gebruikt wordt om via hypervlakken (hyperplanes) een optimale splitsing tot stand te brengen tussen dataclusters in meerdimensionale ruimtes³² (zie Vapnik, 1995; Hearst, Schölkopf B. et al., 1998; Smola en Schölkopf, 1998; Vapnik, 1998; Bennett, 2000; Vert, Tsuda et al., 2004).

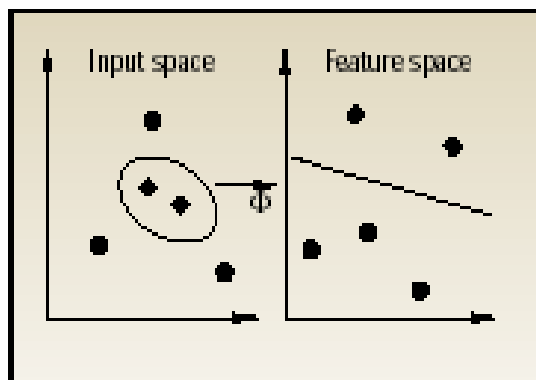
Via een niet-lineaire afbeelding ('mapping') (zie tekstbox) beeldt de SVM de training data af in een hoogdimensionale ruimte ('feature space', kenmerkenruimte), waarin het optimale hypervlak eenvoudig geconstrueerd kan worden, op basis van een klein aantal 'steunpunten' (de 'support vectors'). Hierbij worden de support vectors via een iteratief proces bepaald door data punten te zoeken op de grensvlakken van de te onderscheiden categorieën, en het iteratie proces te stoppen op een moment waarop het hypervlak binnen de foutdrempel geconstrueerd kan worden. Dit hypervlak in de 'feature space' is in feite een niet-lineair oppervlak in de oorspronkelijke 'input ruimte'. Bij ingewikkeldere situaties wordt de mapping naar de feature space bijvoorbeeld tot stand gebracht door gebruik van klassieke NN-kernels, zoals bijvoorbeeld de RBF of de sigmoïde kernel. Classificatie van nieuwe gegevens kan nu eenvoudig tot stand worden gebracht door deze aan te bieden aan het algoritme, en te bepalen aan welke kant van het hypervlak deze liggen.

Behalve voor classificatieproblemen, is SVM ook voor regressieproblemen (SVR: Support Vector Regression) bruikbaar, en kan worden opgevat als een zogenaamd 'kernel-induced nonlinear' methode (Vapnik, Golowich et al., 1997; Gunn, 1998; Smola en Schölkopf, 1998; Pelckmans, Suykens et al., 2002; Schölkopf en Smola, 2002; Suykens, Van Gestel et al., 2002; De Brabanter, 2004). Gebruik makend van een niet-lineaire kernel functie wordt het oorspronkelijk regressieprobleem na enige aanpassing als een convex programmeringsprobleem in een hoog-dimensionale ruimte geherformuleerd (zie bijvoorbeeld §12.3.5, §12.3.6 in Hastie, Tibshirani et al. (2001)). Dit probleem heeft een eenduidig (globaal) optimum. In de praktijk blijkt dat SVM hierbij vaak efficiënt met gegevens omgaan, en slechts een kleine fractie van de data gebruikt (de relevante, karakteristieke gegevens) om de oplossing tot stand te brengen. Hiermee bereiken SVM's een grote datareductie, en kan de validatie inspanning ook beperkt blijven tot een aantal belangrijke datapunten. SVM werkt zuiniger dan NN, door niet alle data te gebruiken voor de fit, en (ten gevolge van minder model parameters) hierbij overfitting te vermijden, terwijl tevens problemen met meerdere locale optima vermeden worden (ten gevolge van de convexe doelfunctie). De theoretisch elegante 'learning machines' (SVMs en Kernel machines) blijken neurale netwerken dan ook op een aantal gebieden te verdringen ('engineering, information retrieval and bioinformatics', zie bijvoorbeeld Schölkopf en Smola, 2002; Schölkopf, Tsuda

³¹ De taak bij 'statistical learning' ((Hastie, Tibshirani et al. 2001)) is om te beschrijven hoe gegevens georganiseerd of geclusterd zijn, en op basis van training data bijv. tot predicties te komen. Bij 'supervised learning' zijn expliciet output-variabelen beschikbaar die het leerproces kunnen sturen. In 'unsupervised learning' worden enkel 'features' waargenomen, en zijn geen metingen van de output beschikbaar.

³² SVMs brengen een optimale discriminatie tot stand tussen twee klassen, waarbij de 'beslissingsgrens' voor de classificatie zo ver mogelijk weg ligt van beide klassen.

en Vert, 2004). Clarke, Griebisch et al. (2004) vergelijken SVMs met andere metamodelleringsmethoden als RSM, MARS, RBF en Kriging, en wijzen op de goede performance van de SVMs (zie ook Morris, Autret et al., 2001). Voor software zie bijvoorbeeld Pelckmans, Suykens et al. (2002), en de algemene URL <http://www.support-vector.net/software.html>



Figuur 11:

Het idee van SV machines: beeld de training data niet-lineair af in een hoger-dimensionale 'feature space' via een afbeelding en construeer daar een scheidend hypervlak met een maximale scheidingsmarge. Dit komt overeen met een niet-lineair beslis-grensvlak in de input space'. Door gebruik van een 'kernel functie' is het mogelijk om het scheidend hypervlak te bepalen zonder expliciet de afbeelding te hoeven maken naar de 'feature space' (vrij naar Hearst, Schölkopf et al., 1998).

A.2.10. Fuzzy en Neuro-fuzzy methoden (inductive learning en reasoning)

Fuzzy modellen staan het gebruik van kwalitatieve expert-kennis in de vorm van regels toe, en kunnen daarmee transparantie en interpreteerbaarheid bij modelvorming bevorderen. Met name in situaties waarin onzekerheid over het bestudeerde systeem groot is en er van modellen daarom slechts beperkte nauwkeurigheid kan worden verwacht zijn, is het gebruik van fuzzy modellen op zijn plaats. Om nauwkeurigheid te verhogen kan ook nog gebruik worden gemaakt van meetdata. Deze kunnen gebruikt worden om de fuzzy model relaties te calibreren en zo nauwkeuriger af te stemmen op de werkelijkheid³³. Deze data-gedreven fuzzy modellen worden ook wel aangeduid als *neuro-fuzzy modellen* of *netwerken*. Er zijn vele combinaties van expert kennis met calibratie mogelijk, van een bijna volledig black-box model dat compleet getraind wordt door data tot een fuzzy model waarvan de fuzzy model relaties volledig door de expert gespecificeerd worden, met een aanvullende calibratie op meetgegevens. In Pfeiffer en Isermann (1994) staan criteria voor een succesvolle toepassing van fuzzy en neuro-fuzzy modellen gegeven. Zie verder ook Nelles (2000), hoofdstukken 12-14 en 19-20, en Sjöberg, Zhang et al. (1995), Babuška en Verbruggen (1996).

Uiteraard kunnen de in de literatuur vermelde methodieken voor het schatten van fuzzy-relaties uit empirische data en expert-kennis direct gebruikt worden voor synthetische data-sets die gegenereerd zijn door simulaties met het referentiemodel. Op deze manier kan metamodellering via fuzzy en neuro-fuzzy modellen tot stand komen. Ook kan in deze context gewerkt worden met fuzzy grafieken, enigzins in analogie met de look-up tables (§A.2.1), zie Huber, Berthold et al., 1996; Huber en Berthold, 1998.

Tot slot wijzen we op een link tussen support vector machines en fuzzy modellen, waarbij in Chen, Lin et al. (2003) wordt aangegeven hoe Support Vector machines en kernelmachines bruikbaar zijn bij het vinden van fuzzy relaties.

³³ In feite kunnen fuzzy relaties gebruikt worden om willekeurige functies te benaderen, en kunnen ze dus dienst doen als functieapproximator/metamodel (Kosko 1992; Mitaim en Kosko 2001).

A.2.11. Genetisch Evolutionaire Algoritmen

Bij optimalisatie- en optimale ontwerp-vraagstukken, worden tegenwoordig vaak genetisch evolutionaire algoritmen (genetische algoritmen, genetisch programmeren) ingezet om de zoekactie naar optimale oplossingen in grote parameterruimten gestalte te geven. Deze methoden zijn gestoeld op ideeën uit evolutionaire biologie en genetica (variatie, selectie, mutatie, fitness) om adaptief, proberender- en lerenderwijs (learning-by-doing) de parameterruimten te doorzoeken. Naast het belang van deze methoden voor het adaptief samplen/doorzoeken van de parameterruimte bij metamodel constructie (adaptive experiment design; zie Wang, 2003; Wang en Simpson, 2004; Qian, Seepersad et al., 2005) kunnen ze ook gebruikt worden bij de constructie van surrogaat-modellen en bij het uitvoeren van optimalisatie berekeningen (Alvarez, 2000; Jin, 2003; Affenzeller en Wagner, 2004a,b).

A.3. Systeemtheoretisch gebaseerde modelreductie

A.3.1. Lineaire Systemen

Binnen de systeem- en regeltheorie worden voornamelijk dynamische modellen/systemen beschouwd. Vooral voor lineaire systemen bestaat er een uitgebreide theorie. Lineaire systemen kunnen op diverse wijzen gerepresenteerd worden, zoals transferfunctiemodellen, impulse response modellen, ARMA (tijdreeks) modellen etcetera (Ljung, 1999; Ljung, 2004). Een populaire vorm is de zogenaamde toestandsvormrepresentatie (state space models), waarbij het model M beschreven wordt middels lineaire matrix differentiaalvergelijkingen (continue tijd) of differentievergelijkingen (discrete tijd) van de vorm:

$$\begin{array}{ll}
 \textit{Continu} & \textit{Discreet} \\
 \dot{x}(t) = Ax(t) + Bu(t) & x(t+1) = Ax(t) + Bu(t) \\
 y(t) = Cx(t) + Du(t) & y(t) = Cx(t) + Du(t)
 \end{array}
 \tag{17}$$

$$x \in R^n, y \in R^p, u \in R^m, A \in R^{n \times n}, B \in R^{n \times m}, C \in R^{p \times n}, D \in R^{p \times m}$$

waarbij x, u, y respectievelijk de toestand, input en output signalen van het model zijn.

Vaak wordt er ook nog een stochastische term aan de vergelijkingen toegevoegd om de invloed van ruisachtige signalen en verstoringen mee te modelleren. Indien de matrices $\{A, B, C, D\}$ constant zijn, spreekt men van tijdinvariante modellen, indien de matrices tijdafhankelijk zijn, noemt men het tijdvariante modellen.

Speciaal voor de tijdinvariante modellen³⁴ bestaat er een groot scala aan modelreductiemethoden, allen gericht op het bepalen van een model M_r met lagere toestandsdimensie $n_r \ll n$,

$$\begin{array}{l}
 \dot{z}(t) = \hat{A}z(t) + \hat{B}u(t) \\
 y_r(t) = \hat{C}z(t) + \hat{D}u(t), \quad z \in R^{n_r}
 \end{array}
 \tag{18}$$

dusdanig dat de approximatiefout $\|y - y_r\|$ in bepaalde zin zo klein mogelijk is. Vaak kijkt men voor evaluatie van een methode naar de zogenaamde oneindig-norm: $\|M - M_r\|_\infty$, waarbij $\|X\|_\infty$ de maximale versterking van een systeem X weergeeft (Glover 1984; Van Dooren 1999).

Hierbij is van belang dat vaak geen enkele *fysische betekenis* meer kan worden toegekend aan de elementen van de gereduceerde toestandsvector z . In de systeemtheorie worden toestanden vaak als 'hulpvariabelen' beschouwd, om het input-output gedrag te beschrijven, en staat het de gebruiker vrij om er wiskundige transformaties op toe te passen, zonder zich te bekommeren om fysische betekenis.

We veronderstellen hier voor de eenvoud dat we te maken hebben met stabiele systemen, wat erop neer komt dat de output altijd naar 0 zal convergeren, indien de input vanaf een bepaald moment op 0 gezet wordt. Stabiliteit wordt voor toestandsmodellen bepaald door de

³⁴ Zie Sandberg (2004) voor een recente verhandeling over modelreductie voor lineaire tijd-variante systemen.

eigenwaarden van de matrix A . Voor continue modellen geldt dat het reële deel van die eigenwaarden negatief moet zijn en voor discrete modellen moeten de eigenwaarden in absolute waarde strikt kleiner dan 1 zijn. Voor systemen die niet stabiel zijn, bestaan er wel varianten op de methoden die hier behandeld worden. Vaak komt dat neer op een splitsing van het systeem in een stabiel en instabiel deel, welke apart gereduceerd worden.

De meeste reductiemethoden gaan uit van deterministische modellen en zijn gericht op het zoeken naar een alternatief model van beperkte orde (lagere toestandsdimensie n_r). In het hiernavolgende bespreken we kort de meest gangbare methoden op dat gebied. Software voor het uitvoeren van deze modelreductie is o.a. verkrijgbaar in Matlab en daarop gebaseerde software (Mathworks 1994-2004).

A.3.1.a. Gebalanceerde modelreductie

Deze methode (Moore 1981; Van Dooren 1999) gaat uit van de concepten regelbaarheid en waarneembaarheid. Eenvoudig geformuleerd komt dit neer op de mate waarin de toestand x bestuurbaar is vanuit de input u , respectievelijk de mate waarin de waarde van een toestand x terug te vinden is in de output y .

Dit wordt geformaliseerd met behulp van de regelbaarheids-Gramian P en waarneembaarheids-Gramian Q , die voor continue tijd modellen gegeven zijn door:

$$P = \int_0^{\infty} (e^{At} B)(e^{At} B)^T dt \quad Q = \int_0^{\infty} (C e^{At})^T (C e^{At}) dt \quad (19)$$

De Gramian Q heeft de interpretatie dat voor een toestand $x(0)=x_0$ geldt dat –indien $u(t)=0$, $t>0$ – de bijbehorende output y voldoet aan (d.w.z. maat voor de ‘energie’ van de output):

$$\int_0^{\infty} y(t)^T y(t) dt = x_0^T Q x_0 \quad (20)$$

Voor de Gramian P geldt dat –indien we eisen dat $x(0)=x_0$ – de minimale benodigde input energie om x_0 te bereiken gegeven wordt door:

$$\min_{u \in L_2(-\infty, 0)} \int_{-\infty}^0 u(t)^T u(t) dt = x_0^T P^{-1} x_0 \quad (21)$$

De gebalanceerde modelreductie methode maakt gebruik van het feit dat modellen als (17) invariant (in *input/output* zin) zijn onder toestandstransformatie met behulp van een reguliere³⁵ (dat wil zeggen inverteerbare) ($n \times n$) matrix T ,

$$x \rightarrow Tx, \quad A \rightarrow TAT^{-1}, \quad B \rightarrow TB, \quad C \rightarrow CT^{-1}, \quad (22)$$

maar dat de Gramians P en Q niet invariant zijn ($P \rightarrow TPT^T$, $Q \rightarrow T^{-T}QT^{-1}$). Wel zijn de eigenwaarden van het product PQ onafhankelijk van transformatie.

De methode bepaalt een transformatiematrix T , dusdanig dat voor de nieuwe representatie (TAT^{-1}, TB, CT^{-1}) de bijbehorende Gramians gelijk en diagonaal zijn,

$$P = Q = \Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n) \quad , \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0 \quad (23)$$

De diagonaal-elementen $\{\sigma_k, k=1, \dots, n\}$ staan bekend onder de naam Hankel Singuliere Waarden en deze zijn zoals gezegd invariante systeemeigenschappen.

³⁵ D.w.z. inverteerbaar, i.e. de inverse matrix T^{-1} bestaat met $T^{-1} x T = Tx T^{-1} = I$.

Een dergelijk toestandsmodel (dus met $P=Q=\Sigma$) noemt men een *gebalanceerde realisatie*. Voor de bijbehorende toestandsvector $x=[x_1, \dots, x_n]^T$ geldt vervolgens dat de mate waarin x_i bestuurbaar is vanuit de input en waarneembaar is in de output wordt bepaald door de waarde van σ_i .

De reductie wordt nu gebaseerd op de notie dat elementen x_i met kleine singuliere waarden σ_i weinig invloed hebben op het input/output gedrag en verwaarloosd kunnen worden. In de praktijk zoekt men naar een punt in de reeks $\{\sigma_k, k=1, \dots, n\}$ waar een ‘knik’ optreedt, i.e. (\gg duidt op ‘veel groter dan’):

$$\sigma_r \gg \sigma_{r+1}, \quad (24)$$

en worden de singuliere waarden $\{\sigma_k, k=r+1, \dots, n\}$ gelijk gesteld aan 0. Indien we het gebalanceerde toestandsmodel van een model M conform partitioneren middels:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = [C_1 \quad C_2], \quad (25)$$

Dan wordt het gereduceerde model M_r gegeven door:

$$\begin{aligned} \dot{x}_r(t) &= A_{11}x(t) + B_1u(t) \\ y_r(t) &= C_1x(t) + Du(t) \end{aligned} \quad (26)$$

Deze reductiemethode is niet optimaal te noemen, in de zin dat er een norm zou bestaan die de approximatiefout minimaliseert, maar er kan wel de volgende afschatting van de fout gegeven worden:

$$\|M - M_r\|_\infty \leq 2(\sigma_{r+1} + \dots + \sigma_n) \quad (27)$$

Verder is de methode erg eenvoudig en numeriek robuust te implementeren en blijkt ze in de praktijk uitstekende resultaten op te leveren.

Er bestaan allerlei varianten van deze methode, o.a. balanceren met weging in het frequentie domein, combinatie van balanceren met singuliere verstoringen (zie later), stochastisch balanceren, closed loop balancing, etcetera

A.3.1.b. Hankel norm reductie

Deze methode (Glover, 1984) betreft de enige modelreductiemethode, die daadwerkelijk een norm voor de fout minimaliseert. De methode gaat uit van de Hankel operator H , de afbeelding die ‘verleden’ inputs afbeeldt op ‘toekomstige’ outputs. Voor discrete tijd systemen (zie (17)) kan deze operator worden weergegeven door een oneindige (Hankel) matrix:

$$H = \begin{bmatrix} CB & CAB & CA^2B & \dots \\ CAB & CA^2B & CA^3B & \dots \\ CA^2B & CA^3B & CA^4B & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix} \quad (28)$$

Indien we een singuliere waarden decompositie van deze matrix beschouwen: $H=U\Sigma V^T$, waarbij U en V unitaire matrices³⁶ zijn en Σ een diagonaal matrix met niet-negatieve elementen, dan kan aangetoond worden dat de positieve diagonaal elementen precies gelijk zijn aan de positieve waarden σ_i uit de vorige paragraaf (zie (23)). Vandaar ook de naam Hankel singuliere waarden.

De rang³⁷ van de matrix H is precies gelijk aan het aantal positieve singuliere waarden.

De methode zoekt een Hankel-matrix H_r van lagere rang, dusdanig dat het verschil $\|H - H_r\|_H$ in de *Hankel norm* ($\|\cdot\|_H$) minimaal is. Hierbij is de Hankel norm van een matrix de maximale singuliere waarde van die matrix.

Het blijkt dat er inderdaad zo'n Hankel-matrix bestaat en er is ook een goed algoritme voor deze methode, al is dit toch een stuk ingewikkelder dan het algoritme voor gebalanceerde modelreductie. Om de methodes enigszins te vergelijken kunnen we ook voor Hankelnorm reductie een foutafschatting geven in de oneindig norm, analoog aan (27):

$$\|M - M_r\|_\infty \leq (\sigma_{r+1} + \dots + \sigma_n) \quad (29)$$

De theoretische bovengrens voor de schattingsfout blijkt dus de helft kleiner dan bij de gebalanceerde modelreductie.

A.3.1.c. Singuliere Verstoringen (singular perturbations)

Het uitgangspunt bij singuliere verstoringstheorie is dat de dynamica van het systeem opgesplitst kan worden in een 'snel' en een 'langzaam' deel (meerdere tijdschalen³⁸). Dit wordt gerepresenteerd door de eigenwaarden van de toestandsmatrix A . De idee achter de bijbehorende modelreductiemethode (Varga 1991) is om het snelle deel van de dynamica te vervangen door een statische relatie, omdat de snelle transiënte effecten snel uitdempen. Op basis van de eigenwaarden van de matrix A wordt het model gepartitioneerd:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = [C_1 \quad C_2], \quad (30)$$

waarbij A_{22} de snelle eigenwaarden bevat. We kunnen het model dan schrijven als:

$$\begin{aligned} \dot{x}_1(t) &= A_{11}x_1(t) + A_{12}x_2(t) + B_1u(t) \\ \dot{x}_2(t) &= A_{21}x_1(t) + A_{22}x_2(t) + B_2u(t) \\ y(t) &= C_1x_1(t) + C_2x_2(t) + Du(t) \end{aligned} \quad (31)$$

De approximatie bestaat er dan in dat we $\dot{x}_2(t) = 0$ stellen (quasi-steady state), wat leidt tot:

$$x_2(t) = -A_{22}^{-1}(A_{21}x_1(t) + B_2u(t)) \quad (32)$$

³⁶ Een vierkante matrix U heet unitair als zijn geconjugueerd getransponeerde gelijk is aan zijn inverse, d.w.z. $U^*U=U^*U=I$, waarbij U^* de geconjugueerd getransponeerde van U is, d.w.z. $U^* = \bar{U}^T$.

³⁷ De rang van een matrix is gelijk aan het aantal onafhankelijke rijen of kolommen van de matrix. De rang van een Hankel matrix is gelijk aan de minimale dimensie van de toestandsruimte van het achterliggende betreffende systeem (C,A,B).

³⁸ Zie ook (Dieckmann en Williams 1992) waarin een modelreductietechniek voor een stelsel gekoppelde differentiaal vergelijkingen wordt voorgesteld die via attractoranalyses in het fase-vlak tot een separatie van snelle en langzame dynamiek komt en de langzame dynamiek als uitgangspunt voor het metamodel neemt.

en na substitutie hiervan tot het gereduceerde model:

$$\begin{aligned} \dot{z}(t) &= (A_{11} - A_{12}A_{22}^{-1}A_{21})z(t) + (B_1 - A_{12}A_{22}^{-1}B_2)u(t) \\ y_r(t) &= (C_1 - C_2A_{22}^{-1}A_{21})z(t) + (D - C_2A_{22}^{-1}B_2)u(t) \end{aligned} \quad (33)$$

Deze methode kan ook gecombineerd worden met balancing. In dat laatste geval kan de oneindig norm van de foutfunctie op dezelfde wijze als bij balanced reduction worden weergegeven.

Een belangrijke eigenschap van deze methode is dat het statische gedrag van het gereduceerde model gelijk is aan dat van het originele model.

A.3.1.d. Modal Truncation

Deze methode (Varga, 1995) lijkt op de singuliere verstoringsmethode, maar veronderstelt dat het model in modale vorm is gebracht, waarbij de A matrix in zogenaamde Jordan vorm staat (zie Golub en Van Loan, 1996). Hierbij wordt wederom een partitie gemaakt en wordt x_2 helemaal verwaarloosd. Het resulterende model is dan:

$$\begin{aligned} \dot{z}(t) &= A_{11}z(t) + B_1u(t) \\ y_r(t) &= C_1z(t) + Du(t) \end{aligned} \quad (34)$$

Ook hierbij is een afchatting van de fout te geven en wel:

$$\|M - M_r\|_\infty \leq - \sum_{i=r+1}^n \frac{\bar{\sigma}(C_i B_i)}{\text{Re}(\lambda_i)} \quad (35)$$

waarbij $\bar{\sigma}(X)$ de grootste singuliere waarde van X is, $\{\lambda_i\}$ de eigenwaarden van A en $C_i B_i$, de kolommen respectievelijk rijen van de matrices C en B zijn. Uit deze afchatting volgt dat het gevaarlijk kan zijn om alleen op basis van kennis van de eigenwaarden (modes) reductie toe te passen, omdat ook de waarde van $\bar{\sigma}(C_i B_i)$ van belang kan zijn.

A.3.1.e. Moment Matching/partiële realisatie/ Padé benadering/ rationale interpolatie

Deze methode (zie onder andere Antoulas en Sörensen, 2001) gaat uit van een transferfunctie beschrijving van een lineair systeem. Deze Laplace domein beschrijving is voor continue systemen te verkrijgen uit een toestandsrepresentatie via:

$$G(s) = D + C(sI - A)^{-1}B \quad (36)$$

Indien we deze functie $G(s)$ in een zogenaamde Laurent reeks ontwikkelen rond een punt s_0 in het complexe vlak:

$$G(s_0 + \sigma) = \eta_0 + \eta_1\sigma + \eta_2\sigma^2 + \eta_3\sigma^3 + \dots \quad (37)$$

Dan noemt men de η_k 's de *momenten* van het systeem in s_0 .

De methodiek beoogt nu een gereduceerd model met transferfunctie $\hat{G}(s)$ te vinden, waarvan de eerste N momenten samenvallen met de momenten van het oorspronkelijke systeem. Voor $s_0 = \infty$ staat dit probleem bekend als partiële realisatie of Padé benadering. Voor willekeurige s_0 spreekt men van rationale interpolatie. Voor de oplossing van deze problemen zijn efficiënte algoritmen (Lanczos en Arnoldi procedures) beschikbaar.

A.3.1.f. Systeemidentificatie

Systeemidentificatie (Ljung, 1999; Ljung, 2004) wordt vaak beschouwd als een methode om op basis van (gemeten) data een model voor het onderliggende proces te bepalen. Hierbij wordt uitgegaan van een dataverzameling $\{u(t), y(t), t=1, \dots, T\}$, een modelstructuur (bijvoorbeeld een toestandsvorm-model met vaste orde, of een ARMA model) en een criterium (vaak kwadratisch) om de modelfout, of voorspellingsfout te minimaliseren.

Zo kan men bijvoorbeeld postuleren dat de data gegenereerd worden door een vergelijking (i.e. modelstructuur):

$$\begin{aligned}x(t+1) &= Ax(t) + Bu(t) \\y(t) &= Cx(t) + Du(t) + e(t)\end{aligned}\tag{38}$$

waarbij $e(t)$ een foutsignaal is om de voorspellingsfout te modelleren. Laat θ een vector zijn met de (te schatten) elementen van de matrices $\{A, B, C, D\}$. Vervolgens wordt een criterium gekozen, bijvoorbeeld:

$$V(\theta, u, y, T) = \sum_{t=1}^T e^2(t)\tag{39}$$

en wordt een optimalisatieprobleem geformuleerd, waarbij die waarde van θ gezocht wordt die het criterium (kwadratische foutensom) minimaliseert. Voor de keuze van een goede modelstructuur, waaronder de dimensie van de toestandsvector, zijn ook weer een aantal methoden beschikbaar.

In deze vorm spreekt men vaak over ‘black box’ identificatie. Indien deze ‘data-based modelling’ methode gecombineerd wordt met fysisch/chemische modellering spreekt men van ‘grey box’ modellering (zie Ljung, Glad, 1994).

Ten behoeve van modelreductie kan bovenstaande methode natuurlijk ook toegepast worden op simulatiedata van een groot model, waarbij voor het gereduceerde model een modelstructuur gekozen wordt van veel lagere orde dan het oorspronkelijke referentiemodel.

A.3.2. Niet-Lineaire Systemen

Voor niet-lineaire systemen zijn er minder theoretische resultaten³⁹, maar toch zijn er vele applicaties te vinden.

A.3.2.a. Niet-Lineair gebalanceerde reductie

Dit is een uitbreiding van de methode voor lineaire systemen, waarbij 2 varianten onderscheiden worden:

1. op basis van model gebaseerde energie functies (generalisatie van Gramians) (Scherpen, 1993; Scherpen, 1994).

Hierbij wordt uitgegaan van een model:

$$\begin{aligned}\dot{x}(t) &= f(x(t)) + g(x(t))u(t) \\y(t) &= h(x(t))\end{aligned}\tag{40}$$

en worden de Gramians ‘theoretisch’ bepaald. Zie ook (Newman en Krishnaprasad, 2000).

³⁹ Voor specifieke toepassingsgebieden, bijv. voor mechanische constructies (structural dynamic modeling) zijn er specifieke modelreductie methoden ontwikkeld (zie bijv. (Friswell, Penny et al. 1995; Friswell, Penny et al. 1995) en zie ook Peter Avitabile’s website http://faculty.uml.edu/pavitabile/22.515/Model_Reduction_061904.pdf en http://faculty.uml.edu/pavitabile/22.515/Modal_Expansion_061904.pdf)

2. op basis van data, waarbij empirische Gramians bepaald worden op basis van een aantal model runs (Lall, Marsden et al., 1998).

Hierbij wordt uitgegaan van een algemenere vorm van een niet lineair model en worden een aantal ‘representatieve’ inputsignalen gekozen, waarmee het systeem gesimuleerd wordt. Vervolgens worden de Gramians van het systeem geschat middels het bepalen van de covariantie van de toestandsvector.

A.3.2.b. Proper Orthogonal Decomposition (POD)

Deze techniek heeft als doel om de dimensie van de toestandsruimte te verkleinen op basis van data. Laat een model gegeven zijn in de vorm:

$$\dot{x}(t) = f(x(t), u(t)); \quad y(t) = h(x(t), u(t)) \quad (41)$$

en veronderstel dat we bij gegeven u (uit metingen of simulaties) een verzameling datapunten:

$$X = [x(t_1) \quad x(t_2) \quad \dots \quad x(t_N)] \quad (42)$$

tot beschikking hebben. Deze matrix van ‘snapshots’ van de toestanden heeft dimensies $n \times N$, waarbij doorgaans $N \gg n$. Op basis van een singuliere waarden decompositie van deze matrix X en het verwaarlozen van kleine singuliere waarden, proberen we een lagere dimensionale toestandsruimte te creëren:

$$X = U \Sigma V^T \approx U_k \Sigma_k V_k^T \quad (43)$$

waarbij de U_k , Σ_k en V_k respectievelijk de $n \times k$, $k \times k$ en $n \times k$ deeltmatrix is van U , Σ , V die de componenten in de SVD (singuliere waarden decompositie; (Golub en Van Loan 1996)) van de snapshot-matrix X vormen. Hierbij is k de index waarbij er een forse knik optreedt in de aflopende reeks van singuliere waarden, d.w.z. $\sigma_1 \leq \sigma_2 \leq \dots \leq \sigma_k \gg \sigma_{k+1} \leq \dots \leq \sigma_n$ van de diagonaal matrix $\Sigma = \text{Diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$. De n -dimensionale toestand wordt nu geprojecteerd op de lager dimensionale ruimte (dimensie k) via:

$$x(t) \approx U_k \xi(t), \quad \text{met } \xi(t) \in R^k, \quad (44)$$

wat uiteindelijk leidt tot een gereduceerd model

$$\dot{\xi}(t) = U_k^T f(U_k \xi(t), u(t)); \quad y(t) = h(U_k \xi(t), u(t)) \quad (45)$$

De POD is een statistische patroon-analyse techniek om de dominante structuren te vinden in een ensemble van gegevens (vaak ruimtelijk verdeeld). Deze structuren kunnen gebruikt worden als orthogonale basis om het ensemble efficiënt te representeren. De POD staat ook bekend als de Karhunen-Loève decompositie, of de methode van empirische orthogonale (eigen)functies (EOF) of principale componenten analyse (PCA) (zie ook Lall, Marsden et al., 1998; Volkwein, 1999; Krysl, Lall et al., 2001; Willcox en Peraire, 2002; Volkwein, 2004).

Opmerking

Deze methoden worden o.a. ook toegepast voor zogenaamde autonome systemen, waarbij geen inputsignaal $u(t)$ wordt beschouwd. Voorbeelden betreffen het modelleren van geofysische en klimatologische processen, waarbij spatiële modellen/data gebruikt worden van de vorm $\dot{x}(t) = f(x(t))$, waarin de elementen van de vector $x(t)$ specifieke data voor een locatie, zoals een gridcel weergeven. PCA of EOF (Hotelling 1935; Lorenz 1956) beogen een patroon te vinden in de data waarmee zoveel mogelijk de variantie verklaard wordt, en wordt gebruikt om grote hoeveelheden data te reduceren. Veronderstel dat de bijbehorende snapshot matrix gecorrigeerd is voor het gemiddelde, dan kan bijvoorbeeld op basis van een eigenwaarde/singuliere waarde decompositie van de covariantie matrix Q_0 :

$$Q_0 = \sum_{k=1}^N x(t_k)x^T(t_k) = UDU^T = \sum_{j=1}^N d_j U_j U_j^T \quad d_1 \geq d_2 \geq \dots \geq d_N \geq 0 \quad (46)$$

een gereduceerde signaal worden vastgesteld. Data-projectie op de eerste K kolommen van de orthonormale matrix U levert een optimale benadering, in variantie termen, van het oorspronkelijk signaal $x(t)$:

$$\hat{x}(t) \equiv \sum_{j=1}^K U_j z_j(t), \quad \text{waarbij} \quad z_j(t) = U_j^T x(t) \quad (47)$$

een scalaire variabele van de tijd is. De variabele $\hat{x}(t)$ kan vervolgens als gereduceerde representant van $x(t)$ beschouwd worden, hetgeen een enorme reductie in data betekenen (zie Hooimeijer, Heemink et al., 2000; Hooimeijer, 2001).

Voor het bepalen van een gereduceerd model kan men ofwel theoretisch te werk gaan en deze transformatie $x(t) \rightarrow \hat{x}(t)$ direct verwerken in het model, analoog aan hetgeen hierboven beschreven werd, ofwel men kan een nieuw model ontwikkelen voor de data $\{\hat{x}(t)\}$, bijvoorbeeld via identificatie, neurale netwerken etcetera Hasselman (1988, 1993), Von Storch, Bruns et al. (1988) en Von Storch et al. (1988) gebruiken bijvoorbeeld tijdreeksmodellen (AR1 en ARMA) om de dynamiek van de gereduceerde signaaltermen (EOFs) te beschrijven, bijvoorbeeld PIP's (principal interaction patterns) en POP's (principal oscillation patterns). Zie ook Kooperberg en O'Sullivan (1994); Hooimeijer, Heemink et al. (2000). Zie Von Storch en Zwiers (1999) voor een overzicht van ruimtelijke patroon herkenningstechnieken binnen het klimaatonderzoek (zie ook Dima en Lohmann, 2004).

A.3.2.c. Stuksgewijze Linearisatie

Bij deze methode wordt het niet-lineaire model in diverse 'werkpunten' gelineariseerd (i.e. er wordt een lineaire model gemaakt, dat lokaal het niet-lineaire model goed benadert), en vervolgens worden die lineaire modellen gereduceerd middels 'standaard' lineaire methoden (Rewiński en White, 2003; Bontempi en Birattari, 2005). Het globale gereduceerde model is een gewogen som van de lokale gereduceerde modellen. Om te voorkomen dat er een zeer groot aantal linearisaties gemaakt moet worden, worden er vaak voor gekozen om een 'trainingsinput' signaal te bepalen en vervolgens langs het traject van de resulterende toestand een aantal werkpunten te kiezen. Een groot voordeel van deze aanpak is dat het mogelijk is om de hele reductieprocedure puur data-gebaseerd uit te voeren.

Zie ook Johansen en Foss, 1995; Johansen en Foss, 1995; Johansen en Foss, 1995; Johansen en Foss, 1997; Murray-Smith en Johansen, 1997 voor generalisaties van deze 'multi-model' methodiek waarbij ook gebruik wordt gemaakt van niet-lineaire deelcomponenten.

A.3.2.d. Niet-Lineaire Systeemidentificatie

Dit betreft een uitbreiding van systeemidentificatie voor niet-lineaire systemen, en omvat in feite ook de toepassing van fuzzy systemen, neurale netwerken, support vector machines en kernel gebaseerde methoden (zie onder andere Sjöberg, Zhang et al., 1995; Nelles, 2000; Suykens, Van Gestel et al., 2002; Schoukens, Nemeth et al., 2003). Tevens kan men onder deze noemer het schatten van parameters in een voorgeschreven (mogelijk fysisch gebaseerd) niet-lineaire structuur laten vallen. Het idee is om deze technieken toe te passen op data die gegenereerd zijn uit simulaties met het oorspronkelijke te reduceren niet-lineaire model.

A.3.2.e. Varia

Ook kunnen varianten van voorgaande methoden gehanteerd worden. Te denken valt bijvoorbeeld - in analogie met POD - aan het gebruik van data-mining en data-compressietechnieken, zoals niet-lineaire principale componenten analyse (PCA), om typische effectieve variabelen in modelsignalen (inputs, toestanden, outputs) op te sporen, waarvan het dynamisch systeemgedrag geanalyseerd kan worden aan de hand van systeemidentificatietechnieken. Dit kan een aanzet geven tot constructie van een gereduceerd model (zie bijvoorbeeld Bernhardt en Wirtz, 2004).