

**Quickscan: onzekerheid en  
het D-Hydro  
modelinstrumentarium**

Inventarisatie en aanbevelingen





# **Quickscan: onzekerheid en het D-Hydro modelinstrumentarium**

**Inventarisatie en aanbevelingen**

K.D. Berends

11200569-012



**Titel**

Quickscan: onzekerheid en het D-Hydro modelinstrumentarium

|                      |                |                       |                 |
|----------------------|----------------|-----------------------|-----------------|
| <b>Opdrachtgever</b> | <b>Project</b> | <b>Kenmerk</b>        | <b>Pagina's</b> |
| RWS-WVL              | 11200569-012   | 11200569-012-ZWS-0014 | 24              |

**Trefwoorden**

D-HYDRO, zesde generatie, modelinstrumentarium, onzekerheid, nauwkeurigheid, kalibratie, validatie, verificatie


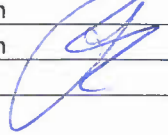
**Samenvatting**

In het kader van het ontwikkelen van het zesdegeneratiemodelinstrumentarium voor Rijkswaterstaat met de nieuwe D-HYDRO software suite worden verschillende kennisvragen onderzocht. Het onderwerp dat in dit rapport centraal staat is onzekerheid in D-Hydromodellen. Onzekerheid is een onduidelijk, diffuus begrip dat door verschillende gebruikers verschillend wordt geïnterpreteerd. Daarom heeft dit rapport als eerste doel om de rol en definitie van onzekerheid binnen de verschillende D-Hydro pilots te inventariseren en in context van modelgebruik te plaatsen. Daartoe zijn vraaggesprekken gevoerd met zes verschillende experts uit alle D-HYDRO pilottoepassingen. Het tweede doel van het rapport is het doen van aanbevelingen om te komen tot een gedragen richtlijn voor het adresseren van onzekerheden in het zesdegeneratiemodelinstrumentarium. Op basis van de inventarisatie worden drie concrete aanbevelingen gedaan:

1. Het opzetten van een standaardprocedure voor modelgebruik, verificatie, validatie en onzekerheid.
2. Het kwantificeren van modelonzekerheid voor een pilottoepassing, bij voorkeur een WBI-toepassing.
3. Het opstellen van een heldere 'bijsluiter' voor het communiceren van de toepasbaarheid, nauwkeurigheid en onzekerheid van modellen.

**Referenties**

-

| Versie      | Datum        | Auteur       | Paraaf  | Review           | Paraaf | Goedkeuring | Paraaf  |
|-------------|--------------|--------------|---|------------------|--------|-------------|---|
| 1 (concept) | juli 2017    | K.D. Berends |   | F.L.M. Diermanse |        | Johan Boon  |   |
| 2           | oktober 2017 | K.D. Berends |  | E.L.M. Diermanse | FD     | Johan Boon  |  |

**Status**

definitief



## Inhoud

|  |            |
|--|------------|
| <b>1 Inleiding</b>   | <b>1</b>   |
| 1.1 Aanleiding   | 1          |
| 1.2 Doelstelling en beperkingen  | 1          |
| 1.3 Werkwijze  | 1          |
| 1.4 Leeswijzer   | 2          |
| <b>2 Context</b>   | <b>3</b>   |
| 2.1 Overzicht van het D-Hydro modelinstrumentarium                                       | 3          |
| 2.2 Doel van de D-Hydro modellen   | 3          |
| 2.3 Gebruikers en belanghebbenden  | 4          |
| 2.4 Inbedding van modelresultaten in beleidsprocessen                                    | 4          |
| <b>3 Nauwkeurigheid, kalibratie en onzekerheid</b>                                       | <b>7</b>   |
| 3.1 Definitie van onzekerheid  | 7          |
| 3.2 Kalibratie, verificatie, validatie en onzekerheid                                    | 8          |
| 3.3 Bronnen van onzekerheid  | 9          |
| 3.4 Omgaan met onzekerheid   | 10         |
| <b>4 Conclusies en aanbevelingen</b>   | <b>11</b>  |
| 4.1 Een gestandaardiseerde bewijsprocedure   | 11         |
| 4.2 Pilot-toepassingen modelonzekerheid  | 12         |
| 4.2.1 Kwantificatie van modelonzekerheid binnen WBI                                      | 12         |
| 4.2.2 Onderzoeken van communicatie en inbedding van modelonzekerheid in beleidsprocessen | 12         |
| 4.3 Een bijsluiters voor modeltoepasbaarheid in gewone taal                              | 12         |
| <b>5 Referenties</b>   | <b>13</b>  |
| <br><b>Bijlage(n)</b>  |            |
| <b>A Verslag overleg met WL Antwerpen</b>  | <b>A-1</b> |
| <b>B Correspondentie</b>   | <b>B-1</b> |
| B.1 Commentaar Rijkswaterstaat   | B-1        |
| B.2 Antwoord Deltares  | B-3        |





# 1 Inleiding

## 1.1 Aanleiding

Deltares werkt in opdracht van Rijkswaterstaat (RWS) een nieuw zesdegeneratiemodelinstrumentarium gebaseerd op de nieuwe D-HYDRO suite software. Binnen dit overkoepelende D-HYDRO project worden nieuwe modellen ontwikkeld voor de Nederlandse watersystemen. D-HYDRO is het “modelleringsplatform van de toekomst” voor hydrodynamica, morfodynamica en golven. Deze software-suite wordt de vervanger van het door RWS gebruikte SIMONA pakket en het SOBEK-RE softwarepakket. RWS heeft Deltares gevraagd om de mogelijkheden van de nieuwe software te onderzoeken. Een van de kennisvragen die binnen dit project zijn gesteld is hoe om te gaan met onzekerheden.

In Spruyt et al. (2016) is reeds aangeraden om onzekerheidsanalyse op te nemen als integraal onderdeel in de bouw van zesdegeneratiemodellen. In dat rapport is een verkenning uitgevoerd voor de rivieren-pilot. In gesprekken met klankbordgroepen en expert-modellereurs bleken echter verschillende opvattingen te bestaan over de definitie van onzekerheid. Hieruit blijkt een behoefte om het onderwerp ‘onzekerheid’ eerst D-HYDRO-breed te definiëren, voordat een stap kan worden gemaakt naar een gedragen handelingsstrategie.

## 1.2 Doelstelling en beperkingen

Het doel van dit rapport is het inventariseren van de rol van onzekerheid binnen de verschillende D-Hydro pilots en het doen van aanbevelingen om te komen tot een gedragen richtlijn voor het adresseren van onzekerheden in het zesdegeneratiemodelinstrumentarium.

In deze ‘quickscan’ wordt geprobeerd de overeenkomsten en verschillen tussen de verschillende pilots te belichten en waar dat kan te duiden in de context van onzekerheidsliteratuur. Het benoemen en kwantificeren van alle relevante onzekerheidsbronnen valt buiten de beperkingen van dit rapport. Ook wordt niet geprobeerd om sterk aan onzekerheid verwante termen als ‘gevoeligheid’ en ‘robuustheid’ binnen een omvattende terminologie te duiden. In bijlage B is inhoudelijke correspondentie over ‘gevoeligheid’, ‘robuustheid’, en ‘modelfout’ opgenomen.

## 1.3 Werkwijze

Om een D-Hydro dekkend begrip te krijgen van het ‘probleem’ van onzekerheid zijn interviews afgenomen met de verschillende trekkers van de D-HYDRO pilots. In deze vraaggesprekken is de lijn gevolgd van de ‘*Guidance for uncertainty assessment and communication*’ (Peterson et al. 2013) van het Planbureau voor de Leefomgeving, waarin de volgende zaken centraal staan:

- Wat is de context? Waarvoor worden de modellen gebruikt? Wie zijn de belanghebbenden?
- Wordt er beleid gebaseerd basis van modelresultaten? Is dit beleid gevoelig voor kleine veranderingen in modelresultaten?
- Wat betekent onzekerheid in de context van het model? Zijn bronnen van onzekerheid bekend, en wat zijn ze?

In tabel 2.1. is een overzicht opgenomen van de pilottrekkers. In dit rapport worden deze gesprekken samengevat en in de context geplaatst van relevante literatuur. De belangrijkste observaties worden benoemd en op basis hiervan worden aanbevelingen gedaan voor vervolgstappen.

#### **1.4 Leeswijzer**

In hoofdstuk 2 bespreken we de context van de D-HYDRO modellen: welke modellen zijn er, waarvoor worden ze gebruikt en wie zijn de voornaamste gebruikers? In hoofdstuk 3 gaan we in op de begrippen kalibratie, validatie en onzekerheid. In hoofdstuk 4 worden conclusies getrokken en aanbevelingen gedaan voor vervolgstappen.

## 2 Context

### 2.1 Overzicht van het D-Hydro modelinstrumentarium

Binnen D-Hydro worden modellen ontwikkeld voor (Nederlandse) estuaria, rivieren, zeeën en meren. In Tabel 1 is een overzicht gegeven van de verschillende domeinen, cases in de pilots en de Deltares pilot-trekkers waarmee voor deze Quicksan is gesproken.

Tabel 1 Overzicht D-Hydro modellen, gebieden en Deltares pilottrekkers.

| Domein     | Pilot-Case                        | Pilottrekker                         |
|------------|-----------------------------------|--------------------------------------|
| Estuaria   | Oosterschelde<br>Rijn-Maasmonding | Theo van der Kaaij<br>Ymkje Huismans |
| Rivieren   | Maas                              | Mohammed Yossef<br>Jurjen de Jong    |
| Meren      | Markermeer & Veluwerandmeren      | Menno Genseberger                    |
| Zee & Kust | Noordzee                          | Firmijn Zijl                         |

Daarnaast is overleg gepleegd met Joris Vanlede en Jiri Nossent van het Waterbouwkundig Laboratorium Antwerpen. Een verslag van dit overleg is opgenomen in bijlage A.

### 2.2 Doel van de D-Hydro modellen

Numerieke modellen kunnen gebruikt worden voor vele toepassingen. Het doel van het model bepaalt doorgaans de eisen aan bijvoorbeeld nauwkeurigheid of resolutie, waaraan het model moet voldoen. Uit gesprekken met de pilottrekkers is naar voren gekomen dat de D-Hydrmodellen vele doelen hebben.

Voor **estuaria** zijn 2D en 3D Delft3D-modelschematisaties<sup>1</sup> opgezet. Deze modellen zijn gekoppeld aan SWAN (windgolfmodel). Afhankelijk van de vraag wordt het 2D of het 3D model gebruikt. Voor het 2D model is de wettelijke beoordeling van waterkeringen (WBI) een belangrijke toepassing, waarbij extreme waterstanden aan de teen van de dijk worden berekend. Andere toepassingen zijn scheepvaartstroming, getijpoorten (tijdvensters waarbinnen schepen binnen kunnen varen). Op termijn zijn ook operationele voorspelling, waterkwaliteit, morfologie en koelwaterlozing gewenste toepassingen.

Voor **rivieren** worden 1D (SOBEK) en 2D (Delft3D) modellen gebruikt. Afhankelijk van de vraag wordt het 1D of het 2D model gebruikt. De primaire toepassing van het 1D model is operationele voorspelling en waterverdeling. Het 2D model wordt gebruikt voor waterveiligheid, vergunningverlening, waterkwaliteit (bv. chemische verontreiniging, scheepsongevallen). Toekomstig gebruik van hetzelfde model is voorzien voor morfologische voorspellingen, bijvoorbeeld voor voorspelling van navigatiediepte en studies naar baggerbezwaren.

<sup>1</sup> In dit rapport wordt 'modelschematie' en 'model' beide gebruikt voor hetzelfde: een rekenkundig model gebouwd met specifieke software, voor een specifiek systeem en een specifieke periode (bijvoorbeeld: sobek-rijn-j15\_5-v1). Dit wordt onderscheiden van modelsoftware (bv SOBEK of Delft3D) en modelsuites (SIMONA, D-HYDRO).

Voor **meren** zijn 1D, 2D en 3D modellen opgezet. Er wordt onderscheid gemaakt in vier verschillende doelen. (1) waterstanden en golfoploop voor WBI en operationele voorspelling, (2) waterverdeling binnen het Nationale Watermodel, (3) stofverspreiding voor calamiteiten, drinkwatervoorziening; en (4) ingreep-effectstudies voor het meerjarenprogramma infrastructuur, ruimte en transport (MIRT) en milieueffectrapportages (MER). Afhankelijk van het doel wordt een 1D, 2D of 3D model gebruikt.

Voor de **Noordzee** is een 2D Delft3D model opgezet dat gebruikt wordt voor verschillende toepassingen. Het voorspellen van waterstanden tijdens normale condities voor de Hydro-Metocentra (HMC's) en voor extreme condities door de stormvloedwaarschuwingsdienst (voorheen SVSD, nu het watermanagementcentrum WMCN-kust). Andere toepassingen zijn scenariostudies (zeespiegelstijging, kennisontwikkeling), en stromingsvoorspellingen voor *search & rescue* en scheepvaart.

Een gemeenschappelijk thema is dat de D-HYDRO modellen voor veel toepassingen worden gebruikt. Er bleek een convergentie plaats te vinden naar *all-purpose* modellen, waar zo veel mogelijk hetzelfde model gebruikt voor alle soorten toepassingen. Uit de gesprekken bleek tegelijkertijd (momenteel) nog een discrepantie te bestaan tussen de *all-purpose* ambitie en de geteste nauwkeurigheid van de modellen. Bijna alle modellen zijn primair getest op waterstanden, maar worden ook gebruikt voor stroomsnelheden waarop de modellen niet, of slechts beperkt, op gevalideerd zijn. De reden hiervoor verschilt per gebied. Bij de **Noordzee** en **estuaria** wordt niet of beperkt beschikbare meetdata (vooral van afvoeren en snelheden) en beperkte tijd om modellen uitvoeriger te testen genoemd. Bij **rivieren** wordt beperkte data genoemd, maar ook de meer fundamentele kwestie dat het systeem sinds de laatste hoge afvoer, door invloed van Ruimte voor de Rivier, erg veranderd is. Bij de **meren** wordt genoemd dat de locatie van de modeluitvoer voor bijvoorbeeld WBI niet de locaties zijn waar de modelnauwkeurigheid getest kan worden door een beperkt aantal meetlocaties, die bovendien niet aan de teen van de dijk liggen.

### 2.3 Gebruikers en belanghebbenden

De primaire gebruiker en belanghebbende van de D-HYDRO-modellen is Rijkswaterstaat. Binnen Rijkswaterstaat worden de modellen gebruikt door het Wettelijk Beoordelingsinstrumentarium (WBI), regionale diensten (beheer en onderhoud, vergunningverlening) en de watermanagementcentra (operationele voorspelling). Ook wordt Deltares zelf genoemd als belanghebbende, vanuit haar verantwoordelijkheid als kennisinstituut om verantwoord met modellen om te gaan en onzekerheid eerlijk en duidelijk te communiceren.

### 2.4 Inbedding van modelresultaten in beleidsprocessen

Onzekerheid van modelresultaten kan, afhankelijk van hoe de modelresultaten worden gebruikt, invloed hebben op beslissingen. De meeste pilottrekkers geven aan dat onzekerheid in modelresultaten kan leiden tot andere beslissingen. Dit speelt vooral voor rivieren, maar ook meren en estuaria. Het is daarom belangrijk om te weten, in welke beleidsprocessen- en projecten de modellen worden gebruikt. De belangrijkste processen die worden genoemd is het Wettelijk Beoordelingsinstrumentarium (**WBI**), beheer en onderhoud (**beno**), het rivierkundig beoordelingskader (**RBK**), operationele voorspelling (**RWsOS**, **HMC**, **WMCN**) en beslis- en ondersteund systeem Maeslant/Europoortkering (**BOS**).

Alle pilottrekkers noemen **WBI** als een van de belangrijkste gebruikers. Binnen WBI wordt specifiek gevraagd om modelonzekerheid: 'wat is de onzekerheid van het model als de randvoorwaarden bekend zijn?'. De pilottrekkers gaven allen aan moeite te hebben gehad met het beantwoorden van deze vraag. Dit komt enerzijds door de specifieke vraagstelling, doordat gevraagd wordt naar onzekerheid buiten gekalibreerd bereik, en omdat zulke analyses een grote rekenkundige investering vragen.

Een voorbeeld van een **beno** toepassing in estuaria is de analyse van aantal dagen dat de maximaal toelaatbare saliniteit wordt overschreden. Voor de rivieren, en in mindere mate estuaria, wordt aan de hand van modelresultaten besloten over vergunningverlening in het rivierkundig beoordelingskader (**RBK**). In zowel beno als vergunningverlening speelt onzekerheid vaak geen expliciete rol. De interpretatie van modelresultaten kan wel sterk verschillen per toepassing. In het kader van RBK worden modelresultaten geïnterpreteerd als (soms tot op de millimeter) nauwkeurig, terwijl in het genoemde beno voorbeeld men geïnteresseerd is in een orde-grootte antwoord. Dat modelresultaten onnauwkeurig zijn, wordt dan (soms impliciet) bekend verondersteld.

In operationele voorspelling (**RWsOS**, **HMC** en **WCMN**) en **BOS**-Maeslantkering worden ook D-HYDRO modellen gebruikt. Er is vanuit deze processen (in tegenstelling tot WBI) geen expliciete vraag naar de onzekerheid van het model. Binnen deze processen wordt onzekerheid wel meegenomen in de afgegeven voorspellingen, maar hier is geen terugkoppeling naar de model-ontwikkelaars.



### 3 Nauwkeurigheid, kalibratie en onzekerheid

Uit de interviews bleek dat over het begrip ‘onzekerheid’ bij de pilottrekkers verschillende opvattingen bestaan. Zo definieert iemand het als gerelateerd aan nauwkeurigheid en het verschil tussen model en meting. Een ander zegt dat er een keus moet worden gemaakt tussen kalibratie of onzekerheid en dat de twee niet naast elkaar kunnen bestaan. Weer een ander ziet onzekerheid als een afgebakend deel van de totale onzekerheid dat enkel te maken heeft met het model.

In dit hoofdstuk wordt de relatie tussen modelnauwkeurigheid en onzekerheid uitgewerkt in het kader van de D-Hydro pilots en relevante literatuur.

#### 3.1 Definitie van onzekerheid

In de breedste zin van het woord is onzekerheid ‘*elke afwijking van het onbereikbare doel van volledig deterministische kennis van het bestudeerde systeem*’ (Walker, 2003, eigen vertaling). In de wetenschappelijke literatuur wordt er onderscheid gemaakt tussen verschillende ‘onzekerheden’. Er zijn verschillende typologieën (onderverdelingen) (zie hiervoor bijvoorbeeld Skinner, 2014), maar voor deze Quicksan beperken wij ons, conform Draper (1995)<sup>2</sup>, tot een onderscheid tussen voorspelonzekerheid (*predictive uncertainty*) en modelonzekerheid (*model uncertainty*).

**Voorspelonzekerheid** betreft het verschil tussen meting en model. Dit is van belang voor systemen waarbij vaak voorspellingen worden afgegeven (zoals operationele voorspelling binnen RWsOS) en waarbij het verschil tussen meting en model kan worden afgeschat op basis van lange tijdreeksen (validatie) of operationeel met het Kalman-filter en, meer generiek, data-assimilatie.

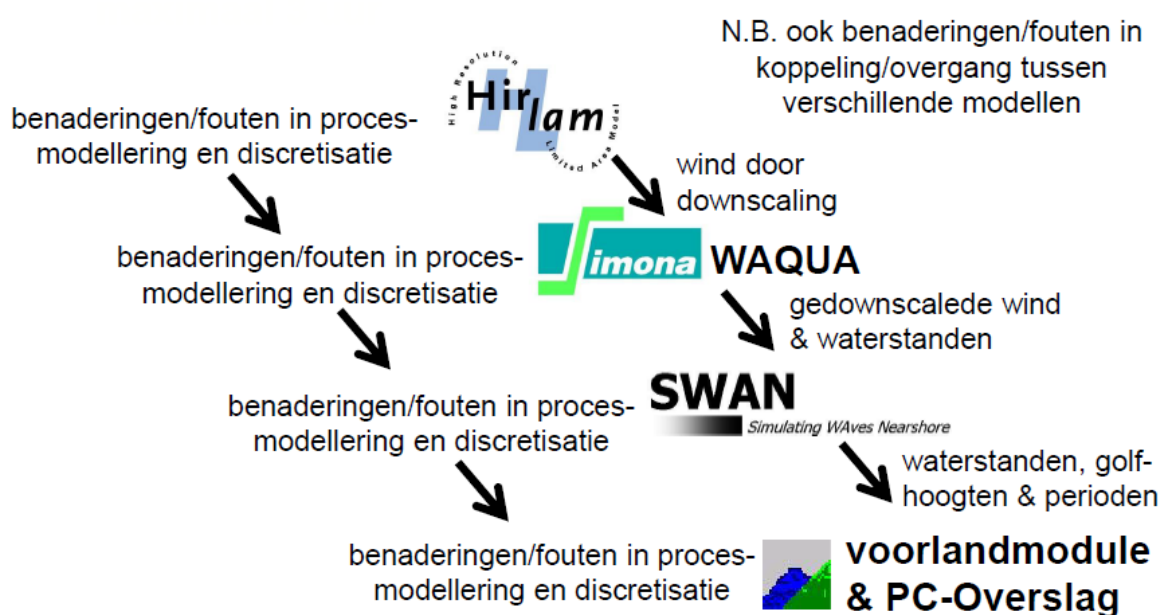
**Modelonzekerheid** betreft onzekerheid in de modelresultaten door onzekerheid in de modelopzet of modelkeuze. Dit is van belang bij toepassingen waarbij metingen niet beschikbaar zijn waardoor voorspelonzekerheid niet kan worden ingeschat of geen betekenis heeft. Het gebruik van meerdere modellen (modelmiddelen of *model averaging*) in klimaatstudies is hier een voorbeeld, maar ook extrapolatie naar extreem hoge of lage afvoeren, extreme stormen of ontwerpstudies.

Uit de gesprekken met de trekkers bleek dat binnen het D-Hydro modelinstrumentarium zowel voorspelonzekerheid als modelonzekerheid belangrijk zijn. Voorspelonzekerheid wordt binnen alle pilots genoemd binnen het kader van modelvalidatie (zie ook paragraaf 3.2). Voorspelonzekerheid is belangrijk voor operationele voorspelling (RWsOS, WCMN) en beheer- en onderhoud toepassingen die vallen binnen gemeten condities. Een bemoeilijkende factor in het schatten van modelnauwkeurigheid is het gebruik van ‘model-cascades’: het koppelen van verschillende modellen waarin elke stap onzekerheden met zich mee brengt (zie het voorbeeld van figuur 1).

---

<sup>2</sup> Draper (1995) definieerde nog een derde onderscheid: *scenario-onzekerheid*. Een voorbeeld hiervan is bijvoorbeeld welke afvoer hoort bij een 1/1250 terugkeerkans.

Modelonzekerheid is voor alle pilots vooral belangrijk bij toepassingen binnen het WBI, omdat daar expliciet wordt gevraagd naar extreme (onbemeten) condities waarvoor geen metingen beschikbaar zijn. Daarnaast wordt binnen de rivieren en estuaria modelonzekerheid genoemd als voornaam probleem in het kader van extrapolatie naar niet-bemeten condities (bij rivieren betreft dit zowel lage als hoge afvoeren) voor beno en RBK toepassingen. In rivieren wordt bovendien de veranderlijkheid van het natuurlijk systeem genoemd als een vorm van extrapolatie. Een voorbeeld hiervan is dat de Rijn sinds de hoogst bemeten afvoer (in de winter van 1995) flink is veranderd, wat de nauwkeurigheid van gebruik van de op de afvoergolf van 1995 gekalibreerde parameters kan beïnvloeden.



Figuur 1 Modellenketen binnen RWsOS-Meren

### 3.2 Kalibratie, verificatie, validatie en onzekerheid

Alle pilottrekkers noemen kalibratie en validatie als een vast en belangrijk onderdeel van het modelleerproces. Uit de gesprekken bleek dat er niet altijd een helder onderscheid bestaat tussen de begrippen kalibratie, verificatie, validatie en onzekerheidsanalyse.

Een goed modelproces (of 'Good modelling practice') wordt vaak gezien als een serie stappen (van Waveren, 1999; Jakeman, 2006) of een cyclus (Rykiel, 1996) dat doorlopen moet worden om tot een model te komen. **Kalibratie** is hierin het 'afregelen' van het model met als doel het verbeteren van de modelnauwkeurigheid. Vaak wordt kalibratie uitgevoerd op parameters waarvan de precieze waarde niet met zekerheid kan worden gemeten of vastgesteld. Dit is veelal bodemruwheid (rivieren, Noordzee, estuaria), diffusievelden (estuaria), en in mindere mate bodemdiepte (Noordzee). Door middel van kalibratie wordt voor deze parameters een waarde vastgesteld.

Omdat kalibratie de modelnauwkeurigheid ---het verschil tussen model en meting--- verhoogt, kan kalibratie kan hierdoor ook gezien worden als het reduceren van de *voorspelonzekerheid*. Tegelijkertijd is er een spanningsveld tussen *modelonzekerheid* en kalibratie. Omdat kalibratie in de praktijk leidt één waarde voor een parameter, wordt de modelonzekerheid 'platgeslagen'. Methodes die een vorm van kalibratie en modelonzekerheid proberen te verenigen (Beven 2001; Montanari, 2004), worden binnen D-HYDRO niet gebruikt.



Over de rol van ‘validatie’ of ‘verificatie’ van modellen is veel gedebatteerd, onder meer of dit wel mogelijk is (Oreskes, 1994). Een pragmatisch onderscheid wordt gemaakt door Rykiel (1996). In zijn interpretatie is **verificatie** bewijs dat de numerieke implementatie juist is: dit kan getest worden door *unit-tests* en vergelijking met meetdata of analytische oplossingen. Voor de D-HYDRO-software wordt in verificatiedocumenten de juistheid van de software getoetst aan (geïdealiseerde) testdata. Een voorbeeld is de test of de getijvoortplanting in een halfgesloten bekken overeenkomt tussen Delft3D Flexible Mesh en het (semi-) analytische Taylor-model. **Validatie** (preciezer: operationele validatie / *operational validation*) is de test of het model geschikt is voor de toepassing waar het model voor bedoeld is. In deze opvatting kan een model dus niet algemeen ‘gevalideerd’ worden<sup>3</sup>. Er kan hoogstens een nauwkeurighedsverwachting worden afgegeven voor de omstandigheden waaronder het model is getest.

Binnen D-Hydro wordt validatie vaak uitgevoerd als een onafhankelijke test (in de pilot **meren**, waarin vaak niet wordt gekalibreerd) of na kalibratie (overige pilots) om aan te tonen dat het model een bepaalde nauwkeurigheid heeft bereikt. In deze context kan validatie worden gezien als een schatting van de voorspelonzekerheid – binnen de grenzen van het geteste bereik en gelimiteerd tot de geteste modeluitvoerlocaties.

### 3.3 Bronnen van onzekerheid

Er zijn verschillende oorzaken die de voorspel- en modelonzekerheid kunnen beïnvloeden. De meest genoemde bronnen zijn randvoorwaarden, modelschematisatie en modelparameters. Over het algemeen is er bij de pilottrekkers een groot vertrouwen in de wijze waarop modellen de werkelijke waterbeweging benaderen. Voor andere processen dan waterbeweging (hydrodynamica) bestaat er nog (grote) onzekerheid over de werking van het natuurlijk systeem. Dit betreft vooral zoutindringing, (sediment-) transport en morfologie. Voor hydrodynamica is de algemene opvatting dat de fysica bekend (genoeg) is. Onzekerheid in hydrodynamica (stroomsnelheden, waterstanden) komen voornamelijk uit de *schematisatie*<sup>4</sup>.

Modelspecifieke bronnen die vaak zijn genoemd zijn waarden of grootheden binnen een model die andere processen parametriseren en dus niet direct gemeten kunnen worden. De voornaamste voorbeelden hiervan zijn ruwheids- en diffusieparameters. Andere bronnen die genoemd worden zijn randvoorwaarden (lateralen, afvoeren, windvelden), submodellen (turbulentie-, ruwheids- en dispersiemodellen), discretisatie, bodemligging en vegetatiekaarten. Tot slot wordt door alle trekkers het gebruik van modellen buiten het validatiebereik genoemd als een bron van onzekerheid.

Bovenstaand overzicht biedt geen uitputtend overzicht van onzekerheden. Raamwerken zoals de ‘onzekerheidsmatrix’ van Walker (2003) zouden gebruikt kunnen worden om per toepassing een gestructureerd overzicht te maken van onzekerheidsbronnen. Een voorbeeld van een dergelijke toepassing voor rivieren is gegeven door Warmink (2011), die bovenstaande matrix toepaste in een expert-onderzoek voor rivierkundige modellen (specifiek: WAQUA-Waal). Er bestaat onder de pilottrekkers over het algemeen een groot vertrouwen dat de belangrijkste bronnen van onzekerheid bekend zijn, of eenvoudig geïdentificeerd kunnen worden.

<sup>3</sup> Het begrip “modelvalidatie” heeft in deze context dus eigenlijk geen betekenis.

<sup>4</sup> Het begrip ‘schematisatie’ wordt hier opgevat als het geheel van geometrie en parameters. Een schematisatie kan worden opgevat als een ‘tekening van het natuurlijk systeem’ in een modelstelsel. Zie ook een eerdere voetnoot hierover.

### 3.4 Omgaan met onzekerheid

Er wordt momenteel op verschillende manieren omgegaan met onzekerheid. Bij alle modellen wordt modelnauwkeurigheid (als een maat voor voorspelonzekerheid) uitgedrukt via indicatoren zoals de *root mean square error* (afgekort als RMS, RMSE of RMSD) of andere statistische testen. In de meeste gesprekken werd het risico genoemd dat gebruikers gerapporteerde modelnauwkeurigheid te breed interpreteren: (ver) buiten geteste bereiken, voor ongeteste grootheden of voor ongeteste locaties.

Voor toepassingen waarin modelnauwkeurigheid niet kan worden vastgesteld wordt de (model)onzekerheid zelden gekwantificeerd, maar vaak toegelicht in de tekst. Soms wordt een bereik gegeven waarin antwoorden kunnen liggen. Dit bereik wordt gebaseerd op *expert judgement*, soms ondersteund door gevoeligheidsanalyses of analytische methoden.

De manier van communicatie hangt ook af van de opdrachtgever: niet alle belanghebbenden zijn geïnteresseerd in communicatie van onzekerheid. Een vaak terugkerend commentaar is dat WBI een schatting van de modelonzekerheid wil, terwijl er vanuit beno- en RBK-toepassingen geen of weinig interesse is in onzekerheid. Het communiceren van modelresultaten als 'absolute waarheid', met een (te) hoog aantal significante cijfers, wordt als ongewenst gezien.

Er zijn mogelijk ook culturele verschillen die verschillende verwachtingspatronen kunnen verklaren. Vanuit hydrodynamica (voornamelijk waterstanden) wordt vaak centimeters, of zelfs millimeters nauwkeurigheid verwacht. Het communiceren van onzekerheid (dat de eerder in de orde van enkele decimeters wordt geschat) kan hier een breuk vormen met huidige verwachtingen en beleidsprocessen. In andere disciplines, zoals zoutindringing, waterkwaliteit en morfologie, is de verwachting van modelnauwkeurigheid niet zo groot. Ook hier worden modelresultaten vaak vrij precies gecommuniceerd (bijvoorbeeld het 'aantal dagen overschrijding van een grenswaarde'), maar wordt aangenomen dat er een wederzijds begrip bestaat dat deze waarden niet precies moeten worden opgevat.

Er is grote overeenstemming onder de pilottrekkers dat onzekerheden relevant zijn in hun pilot en dat de invloed van onzekerheid op de robuustheid van conclusies steeds duidelijk gecommuniceerd moet worden: zelfs als de gebruikers hier niet expliciet om vragen. Voor verdere analyse van onzekerheden blijkt een behoefte aan kwantificatie in het kader van WBI, een gestandaardiseerde communicatievorm ('disclaimer'), duidelijke en goed gedocumenteerde softwareverificatie, en meer metingen (m.n. in estuaria en meren) om de nauwkeurigheid voor belangrijke locaties en processen te kunnen afschatten.

## 4 Conclusies en aanbevelingen

Het doel van dit rapport is het inventariseren van de rol van onzekerheid binnen de verschillende D-Hydro pilots en het doen van aanbevelingen om te komen tot een gedragen, consistente richtlijn voor het adresseren van onzekerheden in het zesdegeneratiemodelinstrumentarium. Hiertoe zijn er gesprekken gehouden met de trekkers van de vier D-HYDRO pilots (tabel 2.1) van **rivieren, meren, Noordzee** en **estuaria**. Op basis van deze gesprekken worden de volgende observaties gemaakt:

- De modellen die binnen het D-HYDRO modelinstrumentarium ontwikkeld worden zijn in toenemende mate bedoeld voor meerdere toepassingen en binnen meerdere beleidsprocessen.
- Niet voor alle toepassingen kan de nauwkeurigheid worden gekwantificeerd. In combinatie met de vele toepassingen van D-HYDRO modellen kan dit leiden tot een te hoge verwachting van de nauwkeurigheid van de rekenmodellen door gebruikers.
- Er bestaan verschillende inzichten over wat onzekerheid precies is; de verschillende opvattingen kunnen geduid worden als een verschil tussen voorspelonzekerheid (*predictive uncertainty*) en modelonzekerheid (*model uncertainty*).
- Er is een behoefte aan een gestructureerde aanpak op het gebied van communicatie en kwantificatie van onzekerheid binnen het modelproces.
- In veel praktische toepassingen wordt onderkend dat er sprake is van grote onzekerheid, maar tegelijkertijd dat er geen behoefte is om deze te kwantificeren.

Op basis van de in dit rapport gepresenteerde quickscan worden drie concrete aanbevelingen gedaan voor vervolgstappen. Deze worden afzonderlijk besproken paragraaf 4.1 t/m 4.3

### 4.1 Een gestandaardiseerde bewijsprocedure

Modellen in het D-HYDRO modelinstrumentarium zullen voor veel toepassingen gebruikt gaan worden, binnen verschillende beleidsprocessen. Modelresultaten zullen daardoor door een brede, heterogene gebruikersgroep gebruikt worden. Het is daarom belangrijk om de nauwkeurigheid en onzekerheid van deze modellen zo expliciet en gestandaardiseerd mogelijk vast te leggen. Daarom wordt aanbevolen om een D-HYDRO validatiestandaard te adopteren die voor elke D-HYDRO modelschematisatie wordt doorlopen. Het doel van deze standaard is om claims van nauwkeurigheid transparant, herkenbaar en herleidbaar te documenteren. Belangrijke onderdelen van een standaard kunnen zijn:

- Volledige en goed gedocumenteerde *verificatiedocumenten* van de gebruikte D-Hydro softwareversie waarnaar verwezen wordt.
- Voor elke modelversie<sup>5</sup> een *validatiedocument* met:
  - o Een beschrijving van de validatie van data die gebruikt is om het model op te zetten
  - o Een lijst van alle doelen waarvoor het model gebruikt wordt.
  - o Voor elke doel een claim van hetzij nauwkeurigheid (voorspelonzekerheid) voor bemeten bereiken en een (kwantitatieve of kwalitatieve) beschrijving van de modelonzekerheid voor onbemeten bereiken.
  - o Voor elke claim bewijs, of een verwijzing naar bewijs.

<sup>5</sup> Hierbij wordt de versie van de schematisatie bedoeld, niet de softwareversie. Zie ook eerdere voetnoten over deze terminologie

## 4.2 Pilot-toepassingen modelonzekerheid

In tegenstelling tot voorspelonzekerheid wordt modelonzekerheid zelden (onderbouwd) gekwantificeerd. Dit wordt echter wel als relevant bevonden door de pilottrekkers (bijvoorbeeld in het kader van beno of vergunningverlening) of gevraagd door de opdrachtgever (binnen WBI). Kwantificatie, waar mogelijk, van de modelonzekerheid kan subjectiviteit wegnemen. Het kwantificeren van modelonzekerheid is echter sterk afhankelijk van de vraag of toepassing, waarvoor de kwantificatie gebruikt zal worden. Er zijn verschillende toepassingen geïdentificeerd waarvoor kwantificatie van modelonzekerheid van belang kan zijn. Niet voor alle toepassingen is de vraagarticulatie, of inbedding van onzekerheid in beleidsprocessen, al duidelijk.

### 4.2.1 Kwantificatie van modelonzekerheid binnen WBI

Voor WBI is de vraag en inbedding al duidelijk. Daardoor is WBI de eerste kandidaat voor kwantificatie van de modelonzekerheid voor een pilot-toepassing, ter vervanging van of aanvulling op huidige schattingen van de onzekerheid. Kwantificatie voor één toepassing kan bovendien helpen de communicatie, inbedding en vraagarticulatie van modelonzekerheid voor andere toepassingen te verduidelijken.

### 4.2.2 Onderzoeken van communicatie en inbedding van modelonzekerheid in beleidsprocessen

Voor andere toepassingen, zoals effectstudies en beno-toepassingen waarbij geen validatie kan worden uitgevoerd is zowel de vraag (waarop geeft een onzekerheidsanalyse antwoord?) en de inbedding in beleid, onduidelijk. Voor deze toepassingen wordt aangeraden om eerst deze context te verduidelijken, zodat [a] de meerwaarde van het kwantificeren van onzekerheid expliciet kan worden gemaakt en [b] duidelijk wordt gemaakt dat de analyses niet per definitie veel complexer wordt door het meenemen van onzekerheden.

## 4.3 Een bijsluiter voor modeltoepasbaarheid in gewone taal

Zonder duidelijke communicatie van modelnauwkeurigheid en modelonzekerheid wordt geen helderheid geschept. Validatiedocumenten, onzekerheidsbanden en kansintervallen zijn vaak bijzonder informatiedicht en niet eenvoudig te begrijpen voor (alle) gebruikers. Gelet op de brede insteek van de D-Hydro-modellen wordt aangeraden om bij documenten een 'bijsluiter' te leveren waarin in heldere, niet-modeltechnische taal de toepasbaarheid van modellen wordt beschreven. Het taal- en termgebruik en algemene opzet bij deze bijsluiter is bij voorkeur uniform over alle modellen in het D-HYDRO-modelinstrumentarium.

Een noodzakelijk onderdeel van deze bijsluiter is een vastgelegde richtlijn voor het gebruik van potentieel ambigue of subjectieve kwaliteitstermen ('goede overeenkomst', 'onzekere uitkomsten'). Het koppelen van deze termen aan meetbare, objectieve variabelen helpt om precies en begrijpelijk te communiceren over modelkwaliteit en onzekerheid. Deze richtlijn zou in dezelfde geest het document '*IPCC guidance notes on consistent treatment of uncertainty*', opgesteld kunnen worden.

## 5 Referenties

- Beven, K. & Freer, J., (2001), Equifinality, data assimilation and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, *Journal of Hydrology*, 249, 11-29
- Draper, D. (1995), Assessment and Propagation of Model Uncertainty, *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 45-97
- Jakeman, A. J.; Letcher, R. A. & Norton, J. P. (2006), Ten iterative steps in development and evaluation of environmental models, *Environmental Modelling & Software*, 21, 602-614
- Montanari, A. & Brath, A. (2004), A stochastic approach for assessing the uncertainty of rainfall-runoff simulation, *Water Resources Research*, 2004
- Oreskes, N.; Shrader-Frechette, K. & Belitz, K. (1994), Verification, Validation and Confirmation of Numerical Models in the Earth Sciences, *Science*, 263, 641-646
- Petersen, A.C., Janssen, P.H.M., van der Sluijs, J.P., Risbey, J.S., Ravetz, J.R., Wardekker, J.A., Martinson Hughes, P.H.M. (2013), *Guidance for uncertainty assessment and communication*, 2nd Edition, PBL
- Rykiel, E. J. (1996), Testing ecological models: the meaning of validation, *Ecological Modelling*, 90, 229-244
- Skinner, D.J.C., Rocks, S.A., Pollard, S.J.T. (2014), A review of uncertainty in environmental risk: characterising potential natures, locations and levels. *Journal of Risk Research* 17(2), 195-219
- Spruyt, A., Minns, T., Yossef, M., Kerkhoven, D., Zijl, F., Genseberger, M. (2016), *Advies voor algemeen functioneel ontwerp voor de 6<sup>e</sup>-generatie modellen van RWS*, Deltares rapport 1230071-011-ZWS-0009
- Walker, W.E., Harremoes, P., Rotmans, J., van der Sluijs, J.P., van Asselt, M.B.A., Janssen, P., Kreyer von Krauss, M.P. (2003), *Defining uncertainty. A Conceptual Basis for Uncertainty Management in Model-Based Decision Support*, *Integrated Assessment* 4(1), 5-17
- Warmink, J. J., van der Klis, H., Booij, M. J. & Hulscher, S. J. M. H. (2011), Identification and quantification of uncertainties in a hydrodynamic river model using expert opinion elicitation, *Water Resources Management*, 25, 601-622
- van Waveren, R. H.; Groot, S.; Scholten, H.; van Geer, F. C.; Wösten, J. H. M.; Koeze, R. D. & Noort, J. J. (1999), *Good Modelling Practice Handbook*, STOWA report 99-05, RIZA-report 99.036



## A Verslag overleg met WL Antwerpen

Hieronder is het verslag met kenmerk 1120056-012-ZWS-0020 integraal opgenomen.



|  |   |                                |                                       |
|--|---|--------------------------------|---------------------------------------|
| <b>Datum verslag</b><br>2 oktober 2017 | <b>Ons kenmerk</b><br>11200569-012-ZWS-0020 | <b>Project</b><br>11200569-012 | <b>Opgemaakt door</b><br>Koen Berends |
|--|---|--------------------------------|---------------------------------------|

|   |                              |
|---|------------------------------|
| <b>Datum bespreking</b><br>17 juli 2017 | <b>Aantal pagina's</b><br>30 |
|---|------------------------------|

**Vergadering**  
Overleg onzekerheid in D-Hydro modellen

**Aanwezig**  
Koen Berends, Joris Vanlede, Jiri Nossent

### *Vooraf*

Dit is een verslag van het overleg n.a.v. het conceptrapport "11200569-012-ZWS-0014-r-v1-Quickscan onzekerheid" tussen Koen Berends (Deltares) en Joris Vanlede en Jiri Nossent (beiden Waterbouwkundig Laboratorium Antwerpen). Joris Vanlede (JV) is de coördinator voor numerieke modellering binnen het Waterbouwkundig Laboratorium Antwerpen (WLA). Jiri Nossent (JN) houdt zich voornamelijk bezig met hydrologische modellen. Koen Berends (KB) is auteur van het genoemde rapport. De kopjes in dit verslag zijn achteraf toegevoegd om het lezen te vergemakkelijken.

### *Doel van het overleg*

De reden voor dit overleg komt voort uit een wens tot samenwerking. Hydraulische modellen zijn volgens JV in veel opzichten zijn voor een deel 'uitontwikkeld'. De volgende stap is 'onzekerheid', waarin door het WLA graag wordt samengewerkt met ontwikkelingen binnen D-Hydro.

### *Toelichting op het conceptrapport*

KB begint het overleg met het kort toelichten van het conceptrapport, met daarin aandacht voor het verschil tussen modelonzekerheid en voorspelonzekerheid. JN merkt binnen de hydrologie dit verschil ook. Soms gaat het om voorspelonzekerheid, maar als het om kwantificatie vooral de modelonzekerheid-aanpak. Volgens JV zit de uitgating vooral erin als de klant met een vraag komt als "Wat als ik deze bocht aanpak?". Het effect is misschien 1 cm. Maar is dat wel relevant op bijvoorbeeld 10 cm onzekerheid op de waterstand? Maar met zware modellen kunnen we niet zomaar 100+ berekeningen doen om dat uit te zoeken.

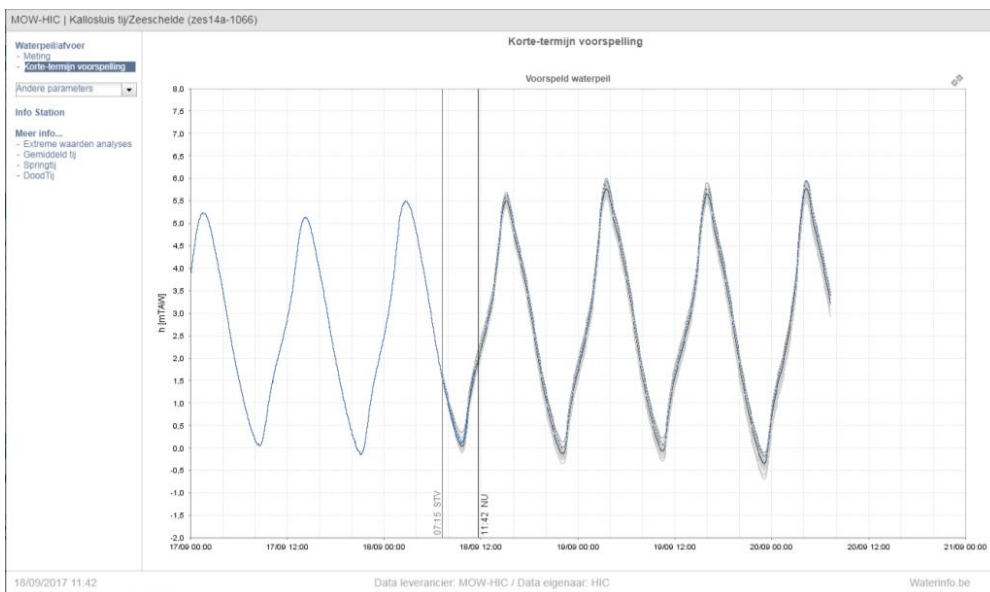
De samenwerking tussen Nederland en Vlaanderen komt ter sprake. JV zegt dat er in het Scheldegebied al jarenlang grote en goede samenwerking is met NL. De samenwerking is ook gevat binnen een Memorandum of Understanding (MOU) tussen het Vlaams Gewest en

Rijkswaterstaat rond samenwerking bij het gebruik, het beheer, onderhoud en de ontwikkeling van het modelinstrumentarium voor waterstandsverwachtingen, getij, golfklimaat, sedimenttransport en morfologie voor het getijgebied Schelde en Kust, Het MOU is in werking getreden op 01/01/2014. Momenteel loop er een project waarbij WLA NEVLA gaan onderhouden via Baseline. Dit project is echt een samenwerking tussen Deltares, RWS en WLA. Binnen het framework van het MOU zijn er ook mogelijkheden om deze modellen als testcase te gebruiken.

KB noemt dat een belangrijk doel van het rapport was om tot een gezamenlijk idioom te komen: zodat we elkaar begrijpen als we over 'onzekerheid' praten. JV vindt dit een belangrijk doel, en herkent dezelfde problematiek binnen het WLA. Hij denkt dat dit rapport een goede aanzet is. JV is samen met Universiteit van Gent (prof. Nico Verhoest & Katrien van Eerdenbrugh) bezig om een workshop op te zetten over onzekerheid. Onderwerpen zijn bijvoorbeeld onzekerheid in QH relaties, afleiden van sedimentconcentraties uit regressies. Hoe onzeker zijn deze resultaten? JV hoopt met behulp van deze opleiding een kritische massa te bereiken en zo ook binnen het WLA een gezamenlijk idioom te kweken. Het doel is dus hetzelfde: dezelfde taal gaan gebruiken.

JN geeft aan dat het onderwerp 'onzekerheid' een moeilijk verhaal is om in de praktijk te brengen. JV benadrukt het probleem met rapporteren van effecten van 1 cm. De tegenstand die hij ondervindt met het benoemen van onzekerheid is geen onwil, maar vooral inertie. Het meenemen van onzekerheden rond resultaten in beleidscontext is nog geen gangbare praktijk, waardoor standaard wordt teruggevallen op de 'deterministische manier van denken'.

We bespreken wat er al *wel* wordt gedaan met onzekerheid. JN noemt voorspellingen die worden afgegeven aan de beheerder van de Zeeschelde. Er wordt dan een range meegegeven (bv +/- 10 cm). Die range wordt ingeschat op ervaring. Daarnaast wordt onzekerheid afgegeven binnen het platform *waterinfo.be* (figuur 1). Deze onzekerheid is gebaseerd op oude metingen en een tool die onzekerheidsbanden genereert.



Figuur 1 - Voorbeeld van communicatie van onzekerheden op *waterinfo.be*: Voorspelde waterstand en onzekerheid in grijze band errond.



### *Modellen van het WLA*

Om een beter begrip te krijgen van de context, vraagt KB naar de verschillende modellen die het WLA gebruikt.

- Er zijn verschillende modellen voor het Scheldesysteem. Een MIKE11 1D takkenmodel gedreven door waterstand in Vlissingen. Die geeft elke uur voorspellingen af. Er bestaat ook een 2D model, NEVLA, waarin wind wordt meegenomen. Die draait ook elke 6 uur.
- Er wordt samengewerkt met het HMC (NL), die hun voorspellingen doorgeven.
- Er is software die de verschillende datastromen en modelvoorspellingen samenvoegt. De expert-van-dienst beslist op basis van expert-analyse welke datastroom gebruikt wordt voor voorspellingen. De voorspellingen van het MIKE11 model worden door data-assimilatie verbeterd. De voorspellingen zijn in ieder geval 2 dagen vooruit.
- Daarnaast is er een hele tak studiemodellen, zoals 2D/3D water- en golfmodellen (SWAN en MIKE21Boussinesq), sedimenttransportmodellen (TELEMAC, DELFT3D), kust-morfologische modellen (XBeach). Deze modellen worden gebruikt voor ingrepeffectstudies (“wat als ik een vaarweg verleg?”).
- WLA doet momenteel niets of niet veel met waterkwaliteitsmodellen (buiten sediment en zout), dat zit bij de Vlaamse MilieuMaatschappij (VMM, onderdeel van Vlaamse Overheid, verantwoordelijk voor de onbevaarbare waterlopen). Dit heeft wel een kennislacune veroorzaakt op het gebied van waterkwaliteit op de bevaarbare waterlopen (bijvoorbeeld stofverspreiding van contaminatie in bevaarbare wateren).
- Zoutindringing is wel relevant. Dat wordt gekalibreerd/gevalideerd. Dit zijn primair 2D/3D modellen zoals NEVLA (SIMONA/TRIWAQ) en SCALDIS (3D TELEMAC). We gebruiken beide modellen voor voorspelling van zout, zowel pulsen als menggedrag. Desalniettemin is de vraag naar zoutberekeningen klein.
- Naast de Schelde is er de Grensmaas, en de rivier de IJzer. Ook de Noordzee, maar het Belgisch continentaal plat is federaal gebied. Daarmee is dit het expertisegebied van MUM/BMM OD Natuur – en niet de Vlaamse overheid. We werken wel met hun samen, want kustveiligheid is Vlaamse bevoegdheid. Voor kustveiligheid spelen vragen als: als ik mijn specie ergens anders stort, hoe moet ik mijn stranden dan efficiënt suppleren. Waar haal ik het zand vandaan? Daarnaast extreme stormen en overslag.

### *Wat betekent onzekerheid voor jou?*

JV haalt ‘drempelwaardes’ aan: waterstanden waarbij alarm moet worden afgegeven. Dicht bij Antwerpen Centrum is bijvoorbeeld een parkeerterrein dat af en toe kan overstromen. Bij alarm moet deze dicht en de auto’s geëvacueerd. Dat is dus een harde drempel, en heel zichtbaar. Voorspellers moeten alle bronnen samenvoegen en dan een voorspelling maken voordat de sms eruit gaat over wel of niet evacueren. JN noemt dat ze afgelopen stormseizoen tweemaal een onder-verwachting hadden: de beheerder wil betere voorspellingen. Als mogelijke oplossing noemt JV het gebruik van meerdere weervoorspellingen. Momenteel wordt gerekend met 1 voorspelling, maar hebben er meerdere. We zouden met meerdere kunnen doorrekenen en dan op basis daarvan gaan voorspellen. De enige manier waar we nu meerdere weersvoorspellingen gebruiken is heuristische terugkijken naar historische reeksen.

JN noemt dat op afvoeren (van rivier- en beeksystem in Vlaanderen) onzekerheid wordt berekend op basis van een *data-based model*. Deze onzekerheid wordt gecommuniceerd via waterinfo.be.

Op de vraag of bij ingrepeffectstudies ook onzekerheid wordt gecommuniceerd, zegt JV dat er intern nog draagvlak voor moet komen. Hij vraagt zich ook af of de klant hier wel op zit te wachten. Een voorbeeld hiervan is het MER traject voor de bevaarbaarheid van de Schelde. Als onzekerheden zomaar worden meegecommuniceerd, worden rapporten mogelijk vatbaar voor aanvallen door juristen en MER-deskundigen. Als klanten in MER-context denken, willen ze maar 1 antwoord. Als modelleur is het eerlijk om onzekerheid mee te communiceren, maar voorzichtigheid is geboden.

Een heel ander probleem is kalibratie/validatie in combinatie met ingrepeffect onzekerheid. Tot slot is er nog zuivere meetonzekerheid. Wij meten bijvoorbeeld turbiditeit en drukken het uit in gesuspendeerd sediment. Deze relatie is niet 1 op 1, maar gekalibreerd. Zelfde met Q/H relaties. Daar zit onzekerheid in.

*Wat zou je als eerst doen, als middelen geen obstakel waren?*

JV zou focussen op het communiceren van modelresultaten. Met verschillende personen rond de tafel gaan en bespreken: "Wat noemen wij een goed model"? Welke Bias/RMSE is goed? Wat is minder? Dat zit in de lijn van het 'Good modelling practice handbook' (van Waveren, 1999), maar die dekken dat niet af. Ik wil naar de klant kunnen zeggen: op basis van deze objectieve maatstaf is dit model *goed*.

JN zou zich toespitsten op modelonzekerheden, en het vertalen van technieken van hydrologische modellering naar meer complexe modellen. Daar is hij al deels mee bezig op het gebied van grondwatermodellering, waar hij een PhD student begeleidt.

*Hebben jullie nog aanvullingen op het conceptrapport?*

JV zou graag komen tot een tabel waarin kwalitatieve beschrijvingen (zoals 'erg onzeker', 'onzeker', 'zeker', etc) worden gekoppeld aan kwantitatieve maatstaven, in dezelfde lijn als IPCC rapporten. Het is belangrijk om dit vuur gaande te houden.

## B Correspondentie

Onderstaande correspondentie heeft plaats gevonden n.a.v. de conceptversie van dit rapport. Hierin komen aanvullende termen als gevoeligheid, robuustheid en fout aan bod. Het commentaar en repliek van de auteur ten aanzien van deze termen is hier opgenomen.

### B.1 Commentaar Rijkswaterstaat

Ik heb de rapportage gelezen. Ik vind het een goed geschreven en helder verhaal. Ik ben blij dat de definitie van de verschillende onzekerheden ook expliciet worden weergegeven. Het woord onzekerheid kan inderdaad op verschillende manieren worden geïnterpreteerd en zo weet de lezer meteen waar het om gaat. Het rapport discussieert goed welke typen onzekerheden onderscheiden kunnen worden.

Ik herken ook de behoefte om onzekerheden beter te kwantificeren. Ik ken inderdaad weinig kaders waarin expliciet wordt gevraagd naar onzekerheid, maar uiteindelijk worden er wel vaker vragen over gesteld (denk aan de MER Zandwinning, aan de discussies over het effect van de verdieping NWW, etc.). Ik ben dus ook eens met de aanbeveling om expliciet aandacht hieraan te besteden.

Ik denk wel dat kijken naar onzekerheden niet voldoende is. Onzekerheid gaat m.i. hand in hand met fout, gevoeligheid, en robuustheid. Ik definieer dan (model)fout als de fout die je bewust maakt omdat je bepaalde modelkeuzes maakt omdat je model anders niet bruikbaar is. Het perfecte voorbeeld is 1D of 2D rekenen in plaats van 3D, of seizoen gemiddeld in plaats van volledig dynamisch, of zelfs een eindigde resolutie. Een deel van de modelfout is in bepaalde omstandigheden op te vangen door kalibratie, maar een deel ook niet (ten minste in specifieke condities). Idealiter zou je elke keer de modelfout zo goed mogelijk moeten kwantificeren.

De gevoeligheid is de mate waarin bepaalde modeluitkomsten worden beïnvloed op het moment dat sommige parameters variëren. Zeker op het moment dat de parameters niet met zekerheid zijn bekend (en dus goede kandidaten zijn als kalibratieparameters) is het van belang om de gevoeligheid voor die parameters te kwantificeren. Zo kan je dan ook je kalibratie beginnen met de parameters waarvoor het model het meest gevoelig voor is. Toen ik nog bij Deltares werkte hebben we in een project geprobeerd combinaties van parameterinstellingen voor slibverspreiding te variëren in realistische ranges (in een soort kleinschalige Monte-Carlo analyse) om de gevoeligheid van het slibmodel te kwantificeren. Dit was uiteindelijk niet erg succesvol, waarschijnlijk omdat de parameters waarmee toen gespeeld is niet de meest relevant waren. Toen hebben we wel automatische procedures voor gebruikt, dat was wel handig.

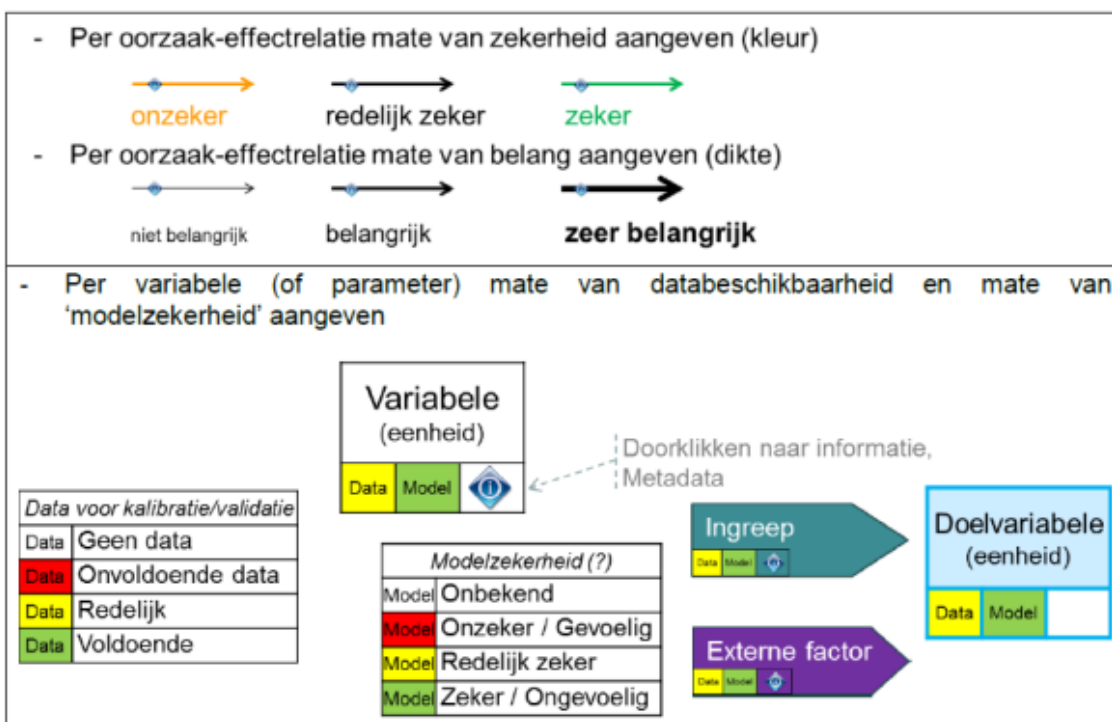
Gevoeligheid heeft ook direct te maken met onzekerheden in de forcering. We gaan vaak modellen kalibreren tot een paar centimeters maar ik ben ooit verteld dat de afvoermetingen bij Lobith een onzekerheidsmarge tot 10% hebben (zonder een bewijs te hebben gezien). Wat is de gevoeligheid van onze modellen op bijv. waterstanden voor deze onzekerheid? En hoe vertaalt het zich door naar de onzekerheid in de modelresultaten?

Hiermee raken we direct ook de onzekerheid in metingen. Deze is bijzonder lastig boven tafel te krijgen (kunnen wij bijvoorbeeld meteen de onzekerheid in waterstands- en zoutmetingen? Ik in ieder geval niet). De enige manier om dat zo kwantitatief mogelijk te kunnen doen is om

in de metadata van de gegevens de onzekerheid zo kwantitatief mogelijk aan te geven. En die moet ook bij Deltares bekend zijn. Ik heb het gevoel dat we bijna een Baseline voor metingen zouden moeten samenvoegen die bij Deltares rechtstreeks beschikbaar is, anders krijgen we waarschijnlijk alleen maar kwalitatieve beschouwingen. Trouwens over Baseline gesproken, weten we ook de onzekerheid van de gegevens die daarin staan?

Naast gevoeligheid heb je nog het concept van robuustheid. Een mooi voorbeeld ervan hebben we ooit aangepakt bij een scenariostudie voor slib en primaire productie in het Eems-Dollard. Op basis van meetgegevens alleen was er niet mogelijk om met zekerheid vast te stellen welke set parameterinstellingen tot de beste resultaten leidde bij de validatie. Daarom zijn scenario's gedraaid met verschillende (in dit geval 3) sets parameterinstellingen. De mate waarin de scenario uitkomsten onderling varieerden gaf toen een indicatie voor de robuustheid van de resultaten. Dit hebben we toen ook objectivering van scenarioberekeningen genoemd (er is een rapport ervan als je dat wil).

Ooit is er ook gekeken naar de robuustheid van het modelleren van effectketens (vooral op ecologie gericht maar op zich kan het ook prima worden toegepast op fysische processen), waarin werd aangegeven met welke mate van zekerheid de relatie tussen proces 1 en proces 2 was bekend, en of de meetgegevens voldoende waren om het model ermee te valideren/kalibreren (zie figuur als voorbeeld, daar heb ik ook het rapport van). Dit zou een onderdeel dat een Baseline voor validatiegegevens kunnen zijn.



*Figuur 3.4 Principes van de verrijkte effectketen: Toevoegen van informatie aan de oorzaak-effectrelatie (boven) en aan de variabelen (onder). De doelvariabele wordt met blauwe kleurcodering onderscheiden. Met Ingreep wordt datgene bedoeld waarvan primair het effect gevraagd wordt. Met Externe factor worden de overige factoren bedoeld die van invloed zijn op de doelvariabele.*

Ik denk dat er structureel explicietere aandacht nodig is voor die verschillende concepten in de ontwikkeling en evaluatie van modellen. Het nadeel is dat er hiervoor in het algemeen grote aantallen simulaties nodig zijn (met verschillende sets parameterinstellingen, met kleine

variaties in de forceringen, etc.). In praktijk kunnen we niet het gebruik van modellen laten wachten op uitkomsten van grote ensembles. Daarom heb ik het gevoel dat we het kwantificeren van onzekerheden e.d. moeten inbouwen in de ontwikkelingsfase van modellen, idealiter in een geautomatiseerde procedure met zo min mogelijk handmatige acties.

## B.2 Antwoord Deltares

Het conceptrapport heeft, noodzakelijkerwijs, een beperkte scope: het onderwerp leent zich (nog) niet voor een alomvattend rapport dat alle discussie uit de lucht neemt. Sterker, de discussie zelf is in mijn ogen het belangrijkste product. Hierom heb ik gekozen quickscan door interviews, om de huidige zienswijzen te inventariseren en integreren. De keerzijde daarvan is dat het geen uitgebreid literatuuronderzoek is waarin elke term uitvoerig wordt gedefinieerd. Dit is m.i. ook nog niet aan de orde – ook in de wetenschappelijke literatuur zijn termdefinities soms diffuus.

Desalniettemin wil ik reageren op de drie termen die je uitlicht: fout, gevoeligheid en robuustheid. Ik ben het met je eens dat deze termen nauw samenhangen met “onzekerheid”. In het rapport belicht ik de termen inderdaad niet an sich – als losstaand begrip – maar komen enkele onderdelen wel terug.

(i) Fout definieer je als bewuste ‘modelfout’. Elk model is volgens deze definitie fout, omdat elk model per definitie een versimpeling/schematisatie van de werkelijkheid is. Je kan daar op twee manieren mee omgaan. Dit fout zou goed mogelijk verkleinen (door goed kiezen van het model en kalibratie) en die fout vervolgens afschatten. In het rapport wordt die afgeschatte fout gedefinieerd als voorspelonzekerheid. Een andere manier om hiermee om te gaan is om het model zelf onzeker te veronderstellen – bijvoorbeeld door belangrijke parameters stochastisch te beschrijven. In het rapport zou dit vallen onder modelonzekerheid. In het laatste geval is het natuurlijk de uitdaging om vast te stellen dat we geen fout over het hoofd zien, en de invoerparameters zo beschrijven, dat de werkelijkheid in de onzekere envelop valt. Idealiter zou een 2D model dan een smallere envelop geven dan een 1D model, omdat minder processen geparametriseerd hoeven te worden. Zoals je zegt: voor beide gevallen is het belangrijk om de fout/onzekerheid te kwantificeren. Ik zou dit dan ook graag willen oppakken in het kader van aanbeveling 4.1: een gestandaardiseerde bewijsprocedure.

(iia) Gevoeligheid is nauw verwant aan modelonzekerheid. De twee zijn enkel te onderscheiden doordat hun vraagstelling anders is. Beide werken met variërende parameters en beide hebben direct te maken met onzekerheid in forcering. Beide zijn ook bijzonder afhankelijk van de keuze van de distributie waarbinnen de parameters variëren: een enkele parameter tussen twee extreme waarden variëren zegt immers weinig over de gevoeligheid van die parameter in de situatie waarvoor het model wordt gebruikt – hoewel zo’n analyse vaak wel model-diagnostische waarde heeft. Ik ben hier niet op ingegaan in het rapport, maar gevoeligheid is een belangrijk onderwerp dat hand-in-hand gaat met onzekerheid.

(iib) De onzekerheid in metingen is een belangrijk punt, waar in m.i. zeker een uitdaging ligt. Zoutmetingen zijn bijvoorbeeld niet direct: de elektrische geleidbaarheid wordt gemeten, en dit wordt via relaties omgezet naar zoutgehalte. Hetzelfde argument geldt voor sedimenttransport. Dit punt werd aangedragen door WL-Antwerpen. Over onzekerheid in Baseline is verspreid al wat bekend – bijvoorbeeld de nauwkeurigheid van ecotoopkaarten, maar een alomvattend beeld ontbreekt. Ik zal dit als extra aanbeveling opnemen onder 4.2.

(iii) Robuustheid definieer je, volgens mij, als gevoeligheid van conclusies voor onzekerheden in de methodiek. Het voorbeeld dat je noemt, valt m.i. binnen wat in het rapport modelonzekerheid wordt genoemd: onzekere parameters die leiden tot onzekere uitkomsten. Robuustheid is daarmee niet zozeer een kwalificatie van het model maar van de conclusies die (mede) worden getrokken op basis van modelresultaten. Dit was ook een belangrijk onderwerp in de interviews. Met name de aanwezigheid van 'drempels' of 'beslisgrenzen' is hierbij belangrijk. Voorbeelden die werden genoemd zijn harde doelen voor centimeters waterstandsverlaging, evacuatiewaterstanden voor parkeerplaatsen in Antwerpen, en aantal dagen overschrijding van het maximale zoutgehalte voor zoetwaterinname. Robuustheid kan dus op meerdere manieren geïnterpreteerd worden. Enerzijds als 'good modelling practice' – een goede analyse beschrijft immers gevoeligheden en onzekerheden die de robuustheid van de conclusies onderschrijven. In dat kader zijn aanbevelingen 4.1 en 4.3 erg belangrijk. Anderzijds kan je praten over de robuustheid van besluitvorming onder onzekerheid. Dit is een belangrijk onderwerp, maar viel buiten de scope van dit rapport (zie aanbeveling 4.2.2).

(iv) Tot slot merk je op dat in de praktijk we niet kunnen wachten op uitkomsten van grote ensembles, en dat kwantificatie daarom ingebouwd moet worden in de ontwikkelfase. Dit is niet bij voorbaat mogelijk: je kan niet echt spreken over 'de gevoeligheid of onzekerheid van een model', omdat elke gevoeligheids- en onzekerheidsanalyse afhankelijk is van conditie waaronder de analyse plaatsvindt. Ik heb hier ook over geschreven in de onzekerheidsbijlage van het rapport D-Hydro Pilot Rivieren. Voor andere toepassingen ligt kwantificatie tijdens modelbouw meer voor de hand, m.n. toepassingen waar voorspelonzekerheid belangrijk is. Dit wordt in zekere zin al gedaan – nauwkeurigheidsmaten als RMSE zijn immers een maat voor variantie van de fout.