

# Interviews over gebruik van AI in waterstand voorspellingen

Bert den Oudsten

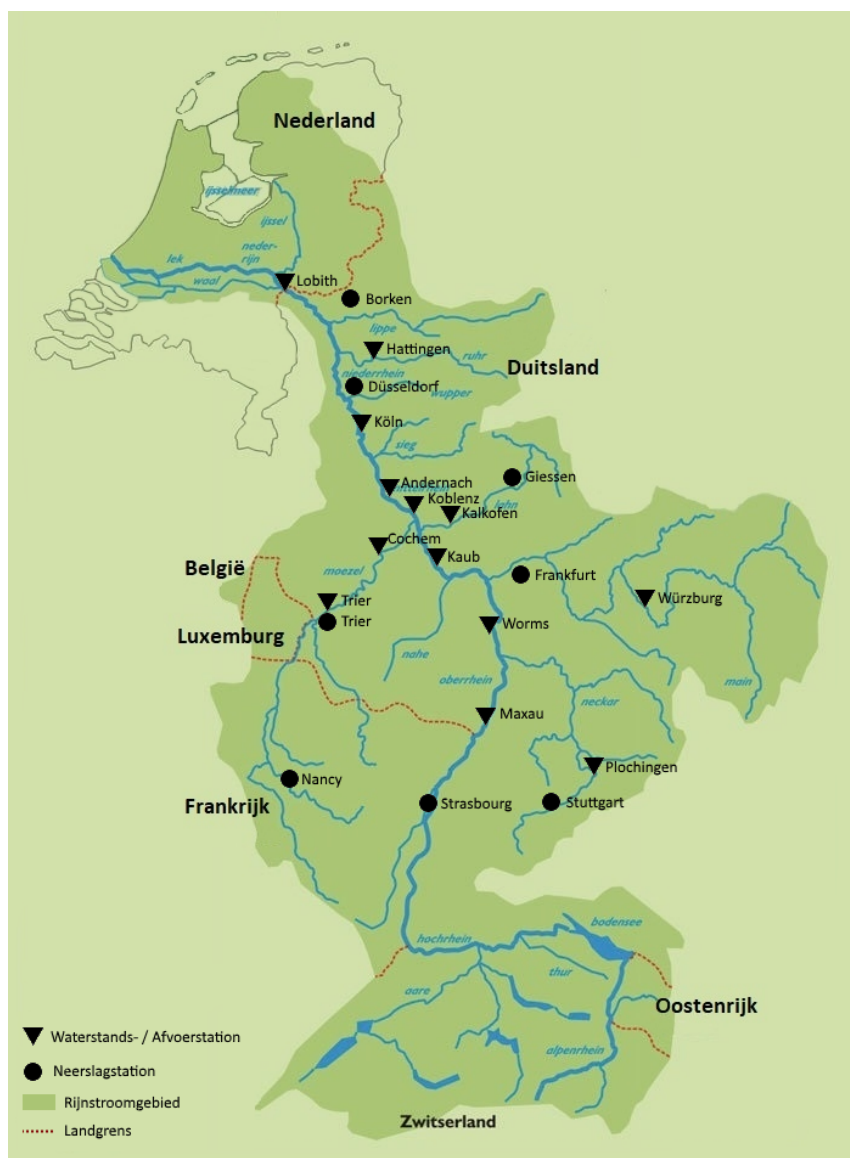
April 2019

# Contents

|   |                                            |    |
|---|--------------------------------------------|----|
| 1 | Introduction                               | 2  |
| 2 | LobithW                                    | 3  |
| 3 | LobithAI                                   | 6  |
| 4 | Vergelijking modellen                      | 8  |
| 5 | Toekomst                                   | 9  |
| 6 | Conclusie                                  | 9  |
| A | Interview met de ontwikkelaar van LobithAI | 10 |
| B | Interview met experts bij Rijkswaterstaat  | 13 |

# 1 Introduction

In dit verslag worden meerdere interviews beschreven. Het eerste interview is met de ontwikkelaar van LobithAI omdat hij er veel technische kennis van heeft. Vanaf nu zal er naar hem worden gerefereerd als geïnterviewde 1. Ook zijn twee werknemers van Rijkswaterstaat geïnterviewd, zij zijn experts in het voorspellen van waterstanden omdat ze daar dagelijks mee werken. In de rest van dit verslag zal er naar hun worden gerefereerd als geïnterviewde 2 en 3. Op basis van de antwoorden die de geïnterviewden hebben gegeven en literatuurbronnen worden de verschillende aspecten van LobithW en LobithAI beschreven. Daarna worden beide modellen met elkaar vergeleken en wordt er beschreven hoe de geïnterviewden denken over de toekomst van het voorspellen van waterhoogten met behulp van AI.



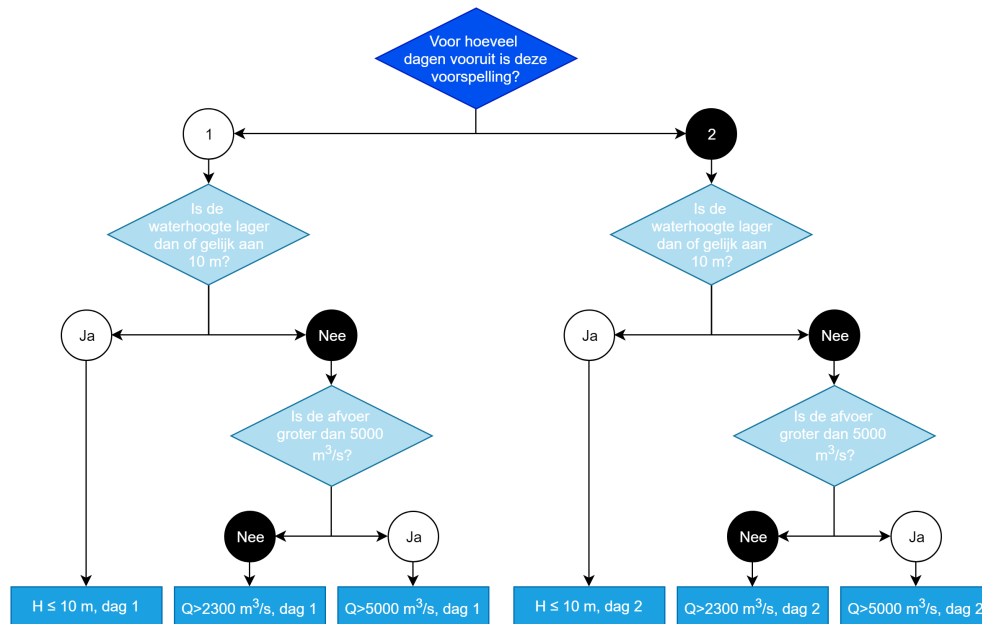
Figuur 1: De Stations waarvan de data wordt gebruikt als input voor LobithW en LobithAI in het stroomgebied van de Rijn.

## 2 LobithW

LobithW is een model dat ontwikkeld is om de waterstand bij Lobith tot vier dagen vooruit te voorspellen met tijdstappen van een dag. Tegenwoordig worden alleen de voorspellingen voor de eerste twee dagen nog maar gebruikt omdat fysisch gebaseerde modellen na twee dagen beter presteren. Het model is geïmplementeerd

in RWSOS dus wordt het dagelijks gebruikt door Rijkswaterstaat bij het voorspellen van waterstanden in Nederland.

LobithW is een multi lineair regressie model. Dat betekent dat lineaire statistische relaties, die bij de kalibratie van het model zijn bepaald, worden gebruikt om de waterstanden te bepalen. De inputs van het model bestaan uit waterstandmetingen en afvoeren in de Rijn bij en bovenstrooms van Lobith en neerslagvoorspellingen in verschillende deelstroomgebieden van de Rijn. Voor de afvoeren en waterstanden worden metingen van de dag zelf tot 4 dagen daarvoor gebruikt. Voor de neerslaggegevens worden de voorspellingen voor de dag zelf tot 2 dagen in de toekomst. De gebruikte stations zijn te zien in Figure 1. Uiteindelijk is na onderzoek gebleken dat niet al deze stations meerwaarde hebben om toe te voegen, dus zijn sommige uiteindelijk niet meegenomen.



Figuur 2: Een diagram van de relaties tussen de verschillende deelmodellen van LobithW.

Het model bestaat uit verschillende deelmodellen die in verschillende situaties worden gebruikt. Voor elke dag vooruit zijn er verschillende deelmodellen, waarbij de keuze voor het specifieke deelmodel wordt bepaald door de waterstand en afvoer op de dag van de voorspellingen. Er zijn voor elke dag verschillende deelmodellen voor waterstanden lager of gelijk aan 10 m, een afvoer tussen de 2300 en 5000 m<sup>3</sup>/s en een afvoer van meer dan 5000 m<sup>3</sup>/s. Een waterhoogte van 10 meter is gelijk aan een afvoer van 2300 m<sup>3</sup>/s. Een diagram van de verschillende deelmodellen en het gebruik ervan is te zien in Figure 2. De verschillende deelmodellen gebruiken ook verschillende inputs omdat er andere statistische relaties zijn bepaald voor de verschillende situaties. Deze verschillende deelmodellen met bijbehorende inputs zijn te zien in Table 1 en 2. In deze tabel zijn behalve de stations ook Noord(P) en Zuid(P) te vinden. Dit is de neerslagverwachtingen voor het noordelijke en zuidelijke stroomgebied van de Rijn in Duitsland. Een H betekent dat de waterhoogte wordt gebruikt, een Q betekent de afvoer en P staat voor de voorspelde neerslag. -24 betekent dat het over de

waarde van 24 uur voor het voorspelmoment gaat en logischerwijs betekent +24 dan de voorspelde waarde van 24 uur na het voorspelmoment.

| <b>H <math>\leq</math> 10 m, dag 1</b> | <b>Q &gt; 2300 m<sup>3</sup>/s, dag 1</b> | <b>Q &gt; 5000 m<sup>3</sup>/s, dag 1</b> |
|----------------------------------------|-------------------------------------------|-------------------------------------------|
| Lobith(H)                              | Lobith(Q)                                 | Lobith(Q)                                 |
| Lobith(H - 24)                         | Lobith(Q - 24)                            | Lobith(Q - 24)                            |
| Andernach(H - 24)                      | Lobith(Q - 48)                            | Hattingen(Q)                              |
| Andernach(H - 48)                      | Hattingen(Q)                              | Köln(Q)                                   |
| Hattingen(H)                           | Hattingen(Q - 24)                         | Köln(Q - 24)                              |
| Hattingen(H - 24)                      | Köln(Q)                                   | Borken(P + 24)                            |
| Hattingen(H - 48)                      | Köln(Q - 24)                              | -                                         |
| Köln(H)                                | Plochingen(Q - 96)                        | -                                         |
| Köln(H - 24)                           | Dusseldorf(P + 24)                        | -                                         |
| Köln(H - 48)                           | Dusseldorf(P + 48)                        | -                                         |
| Dusseldorf(P + 24)                     | -                                         | -                                         |
| Dusseldorf(P + 48)                     | -                                         | -                                         |

Tabel 1: De stations die worden gebruikt voor de verschillende deelmodellen van LobithW op dag 1

| <b>H <math>\leq</math> 10 m, dag 2</b> | <b>Q &gt; 2300 m<sup>3</sup>/s, dag 2</b> | <b>Q &gt; 5000 m<sup>3</sup>/s, dag 2</b> |
|----------------------------------------|-------------------------------------------|-------------------------------------------|
| Lobith(H)                              | Lobith(Q)                                 | Lobith(Q)                                 |
| Lobith(H - 24)                         | Lobith(Q - 24)                            | Lobith(Q - 24)                            |
| Andernach(H - 24)                      | Andernach(Q)                              | Andernach(Q)                              |
| Andernach(H - 48)                      | Hattingen(Q)                              | Hattingen(Q)                              |
| Hattingen(H)                           | Hattingen(Q - 24)                         | Köln(Q)                                   |
| Hattingen(H - 24)                      | Köln(Q)                                   | Köln(Q - 24)                              |
| Kalkofen(H)                            | Köln(Q - 24)                              | Trier(Q)                                  |
| Kalkofen(H - 24)                       | Trier(Q)                                  | Trier(Q - 24)                             |
| Köln(H)                                | Trier(Q - 24)                             | Dusseldorf(P + 24)                        |
| Köln(H - 24)                           | Dusseldorf(P + 24)                        | Frankfurt(P + 24)                         |
| Trier(H)                               | Noord(P)                                  | -                                         |
| Trier(H - 24)                          | -                                         | -                                         |
| Noord(P)                               | -                                         | -                                         |

Tabel 2: De stations die worden gebruikt voor de verschillende deelmodellen van LobithW op dag 2

De data die gebruikt is voor de laatste kalibratie van LobithW zijn afkomstig uit de periode van 1980 tot en met 2011. In deze tijdreeks zitten vrij extreme waterhoogten en afvoeren zoals de bijna-overstromingen van '93 en '95 en droge jaren zoals '76 en '83. Met deze kalibratie voldoet het model 90% van de voorspellingen aan de criteria van een afwijking van 5 cm voor een dag vooruit en een afwijking van 10 cm voor 2 dagen vooruit.

Zoals al gezien kan worden aan het feit dat een gedeelte van het originele model niet meer gebruikt wordt, heeft dit model net als elk ander model wel nadelen. In dit geval was het dat andere modellen beter presteer-

den. Een ander nadeel van dit model is dat het opnieuw kalibreren erg veel tijd kost. Het model lijkt soms ook rare dingen te doen omdat het puur op statistische relaties berust en niet op natuurkundige processen. Hierdoor zitten er soms onwaarschijnlijke relaties in het model. Zo wordt in het deelmodel  $H \leq 10$  m voor dag 1 Hattingen( $H - 48$ ) meegenomen. Dat betekent de waarde op 48 uur voor het voorspelmoment. Omdat er voor een dag vooruit wordt voorspelt, zit er dus 72 uur tussen het meetmoment en de voorspelde waterstand. Zoals te zien is in Figure 3, duurt het echter minder dan een dag voor het water om van Hattingen naar Lobith te gaan. Dit betekent dat het water al lang voorbij Lobith is maar toch wordt het meegenomen in de berekening. Een groot voordeel van dit model is wel dat het berekeningen erg snel kan uitvoeren omdat het alleen maar lineaire vermenigvuldigingen zijn.

### 3 LobithAI

Als MSc-afstudeeropdracht heeft geïnterviewde 1 bij Rijkswaterstaat het model LobithAI gemaakt. Dit model is gemaakt om te onderzoeken wat de mogelijkheden zijn van het gebruik van Artificial Intelligence voor het voorspellen van waterstanden. Het is vooral gemaakt als vergelijking met LobithW, dus daarom gebruikt het bijvoorbeeld ook dezelfde meetstations als inputs.

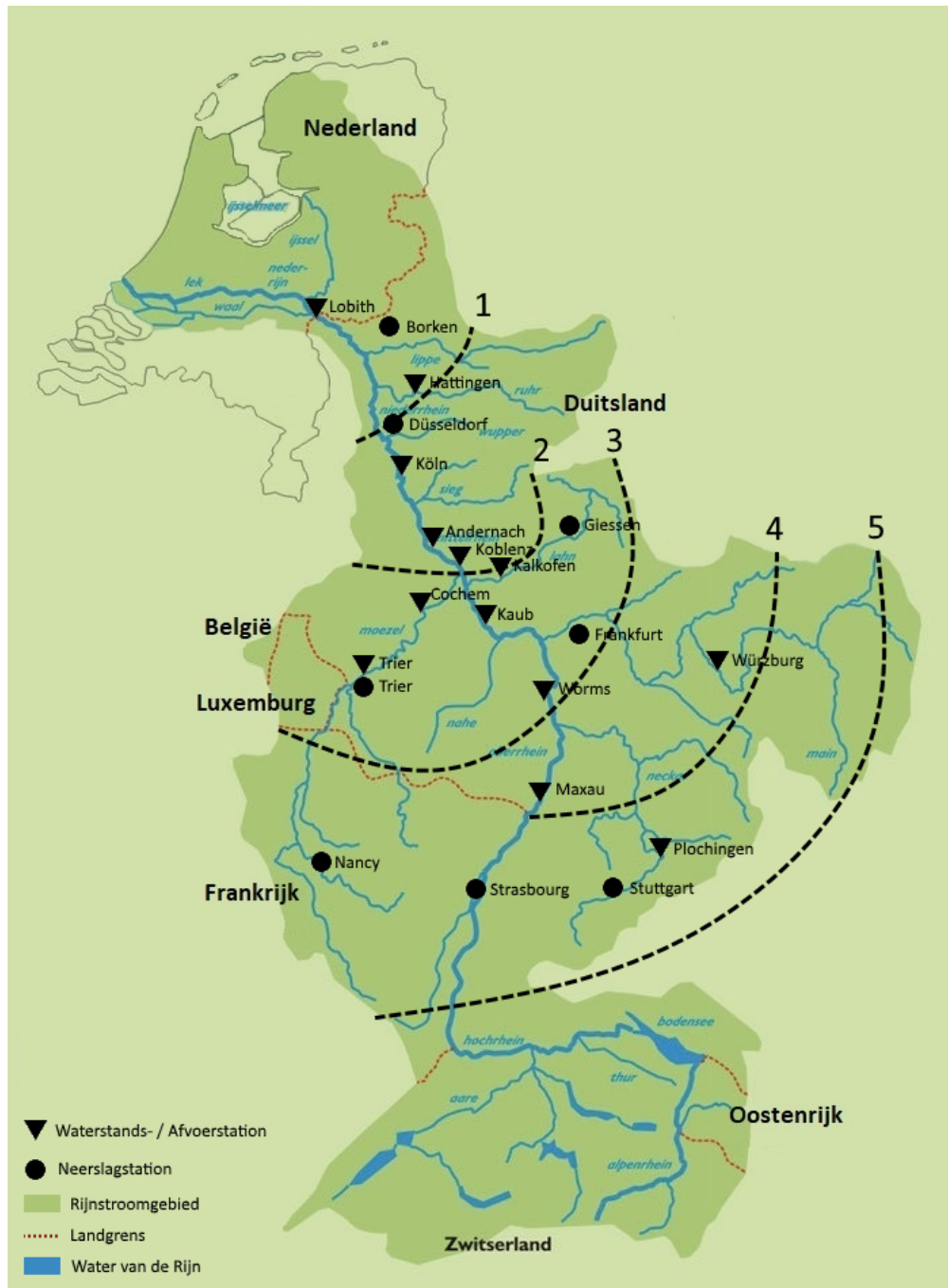
LobithAI is een neuraal netwerk. Een neuraal netwerk is opgebouwd uit verschillende neuronen. Deze neuronen krijgen een input(s) in de vorm van getal(len) binnen. Deze input(s) worden dan verwerkt via een activation function en resulteren zo in een output. Deze output wordt dan weer naar verschillende neuronen in een volgende laag doorgegeven. LobithAI is een Deep Neural Network, wat betekent dat er meerdere lagen elkaar opvolgen.

Net als LobithW maakt het model gebruik van de waterhoogte en neerslaggegevens van verschillende meetstations. Deze stations zijn de zelfde stations als die bij LobithW, echter zijn bij LobithW sommige van deze stations uiteindelijk niet meegenomen. Deze stations zijn hier nog wel meegenomen, dus alle stations uit Figure 1 worden gebruikt als input voor LobithAI. Ook worden bij LobithW alleen metingen van specifieke dagen voor de stations, dat is bij LobithAI ook niet het geval, daar worden alle dagen van alle stations meegenomen. Van alle stations die de waterhoogte meten wordt de waterhoogte op de dag van de voorspelling meegenomen en de vier dagen daarvoor. Voor de neerslagen gegevens worden de weervoorspellingen voor de dag van de voorspelling en neerslagvoorspellingen tot twee dagen in de toekomst meegenomen. Het netwerk heeft zelf, door getraind te zijn met gemeten data, bepaald wanneer welke inputs belangrijk zijn. In het model zitten 8 lagen waarbij de input alleen maar rechtstreeks naar de volgende laag wordt doorgegeven.

De data die is gebruikt om LobithAI te trainen is zoveel mogelijk gelijk gehouden aan de kalibratiedata van LobithW om de vergelijking tussen de twee modellen zo eerlijk mogelijk te houden. Exact dezelfde data was helaas niet meer beschikbaar waardoor dat niet precies kon gebeuren.

Uit een snelle vergelijking tussen de twee modellen suggereerde dat LobithAI over het algemeen beter presteerde, maar hier is nog wel meer onderzoek voor nodig om dit echt te kunnen vaststellen. Dit zou namelijk bijvoorbeeld ook kunnen komen door het verschil in kalibratiedata.

LobithAI wordt op dit moment niet gebruikt omdat de prestaties ervan eerst nog bewezen moeten worden. Als daaruit blijkt dat het op bepaalde domeinen echt nauwkeuriger voorspeld dan LobithW, zal het



Figuur 3: De grenzen van looptijden tot Lobith.



misschien naast LobithW in het RWSOS kunnen gaan draaien of het zelfs volledig vervangen. Uit gesprek met Rijkswaterstaat blijkt dat zij daar ook zeker voor open staan, als blijkt dat de resultaten aanzienlijk beter zijn.

In tegenstelling tot LobithW is dit model een black box, wat als een nadeel gezien kan worden. Het heeft echter als voordeel dat het heel makkelijk opnieuw gekalibreerd kan worden door met nieuwe data te trainen. Met de juiste kennis kan het model heel makkelijk worden uitgebreid om extra soorten input zoals de temperatuur of sneeuwhoogte mee te nemen in de berekening. Op dit moment is het model ook nog erg snel in het berekenen van de resultaten, maar dat wordt waarschijnlijk wel minder als er extra tijdstappen of nieuwe soorten input worden toegevoegd.

## 4 Vergelijking modellen

Er zijn meerdere verschillen tussen LobithW en LobithAI en beide hebben hun eigen voor- en nadelen. Ook al wordt LobithW nu gebruikt in het RWSOS, betekent dat niet per se dat het beter is dan LobithAI. Dit komt meer vanwege het feit dat LobithAI nog maar een paar maanden ontwikkeld is en dus nog niet uitgebreid getest is waardoor de exacte prestaties van het model nog niet bekend zijn. Hieronder worden de verschillen tussen LobithW en LobithAI nogmaals samengevat.

In LobithW zitten af en toe rare relaties maar omdat het allemaal lineaire verbanden zijn is het in ieder geval duidelijk wat de invloed van meetstations is op de voorspelling. Dit betekent dat experts die hier bij Rijkswaterstaat mee werken er rekening mee kunnen houden omdat het duidelijk is wat er gebeurt. Bij LobithAI worden er nog meer dagen voor alle stations meegenomen, dus ook dagen waar helemaal geen relaties tussen zou moeten zitten. Het is hier echter heel lastig om deze relaties goed vast te stellen omdat de invloed van de verschillende stations verschilt voor verschillende initiële waarden. De relaties met de meetstations zijn immers niet constant maar hangen af van alle waarden op de dag van de voorspellingen. Dat betekent dat dit model een echte black box is. Dit kan lastiger zijn om mee te werken, maar de experts bij Rijkswaterstaat hebben aangegeven dat geen groot probleem te vinden omdat veel modellen voor hun nu ook al een black box zijn. De HBV en SOBEK modellijnen, die ook worden gebruikt voor voorspellingen, zouden geen black box hoeven te zijn, maar omdat ze bij Rijkswaterstaat niet voldoende kennis hebben over die modellen is het voor hun wel een black box. Dit hoeft niet per se een probleem te zijn omdat er bij Deltares genoeg mensen zijn die er wel verstand van hebben, maar het maakt het gebruik van het model niet makkelijker.

Uit de eerste analyse van geïnterviewde 1 blijkt dat LobithAI in het normale bereik beter werkt dan LobithW. Er is echter nog niet echt gekeken naar het gedrag bij zeer extreme waarden. Het onderzoek daarnaar zal in het andere deel van dit verslag beschreven worden. Dit is dus hoopgevend voor LobithAI maar nog geen definitieve uitkomst.

Ook is LobithAI veel makkelijker opnieuw te kalibreren in het geval dat dat nodig is. Dit verschil is zeer aanzienlijk aangezien LobithAI in een uur opnieuw gekalibreerd kan worden als de data in de juiste vorm is en de laatste kalibratie van LobithW een onderzoek van een BSc student van 10 weken was. Het uitbreiden van het model zal bij LobithAI ook veel makkelijker zijn omdat het dan ook opnieuw gekalibreerd moet worden.

## 5 Toekomst

Ook de toekomst van dit soort modellen is besproken in de interviews. Uiteraard gaat LobithAI nog niet gebruikt worden in RWsOS zolang er niet bewezen is of het beter werkt dan LobithW. Waarschijnlijk zal LobithAI in de huidige vorm LobithW nog niet volledig gaan vervangen. De eerste stap in het gebruik van Artificial Intelligence in het voorspellen van waterhoogtes zal er waarschijnlijk op neerkomen dat een model als LobithAI naast LobithW zal gaan draaien. Zo kan het vertrouwen in LobithAI worden vergroot en kunnen de resultaten altijd nog makkelijk worden vergeleken. Zeker omdat beide modellen in de huidige vorm weinig rekenkracht kosten is het geen probleem om ze naast elkaar te draaien.

De volgende stap is het uitbreiden van LobithAI. In een neurale netwerk kunnen heel veel verschillende soorten input. Daardoor wordt het mogelijk om ook hele andere dingen mee te nemen die wel invloed hebben op de waterhoogte. Denk daarbij aan dingen als; de sneeuwhoogte op bepaalde plekken, temperatuur, stand van de meren in Zwitserland of de dagelijkse verdamping. Ook zou het model kunnen worden uitgebreid door meer tijdstappen te berekenen, verder in de toekomst voorspellen, data van meer dan 4 dagen geleden meenemen of meer meetstations te gebruiken. Als het model op de juiste manier met zulke uitbreidingen en toevoegingen wordt verbeterd zal het waarschijnlijk beter gaan functioneren. Als dat het geval is kan die verbeterde versie van LobithAI LobithW zeker vervangen.

Een nadeel van al deze uitbreidingen is wel dat het berekenen van voorspellingen veel complexer wordt. Dit zou betekenen dat nieuwe betere server nodig zijn en dat kost tijd en geld. Rijkswaterstaat moet daar wel toe bereid zijn voordat een verbeterde LobithAI LobithW kan vervangen. Toch kan dit volgens de geïnterviewde personen binnen afzienbare tijd gebeuren.

Ook zouden andere delen van RWsOS kunnen worden vervangen door modellen of delen van modellen met AI. De experts bij Rijkswaterstaat zien vooral meerwaarde op plekken waar menselijke invloeden in rol spelen. Een voorbeeld daarvan is het gebruik van stuwen in Wallonië. De informatie daarover is heel lastig te verkrijgen, maar een model met AI zou daar waarschijnlijk goede voorspellingen kunnen maken.

## 6 Conclusie

Uit de vergelijking hierboven lijkt te komen dat LobithAI beter is dan LobithW. Dat is echter nog niet zeker omdat de belangrijkste factor daarvoor nog niet bekend is. Dat is hoe goed het model functioneert bij extreme waarden. Op zulke momenten is een goed functionerend model het belangrijkste, omdat er dan risico's zijn voor onder andere de scheepvaart en mensen die langs de rivieren wonen. Want het model kan nog zo snel kalibreren, als de voorspellingen niet goed zijn is dat irrelevant.

Toch is de kans wel groot dat een model als LobithAI binnen afzienbare tijd gebruikt gaat worden voor waterstandvoorspellingen omdat het er zeer veelbelovend uitziet.

## A Interview met de ontwikkelaar van LobithAI

**Welkom, fijn dat ik jou mag interviewen. Het interview bestaat eigenlijk uit twee delen. De eerste vragen gaan over LobithAI en daarna over de vergelijking tussen LobithW en LobithAI. LobithAI is een Neural Network, maar ook een Deep Learning Network. Het verschil tussen die termen is mij niet helemaal duidelijk, kan je dat een beetje uitleggen?**

Een neurale netwerk is kortgezegd opgebouwd uit verschillende neuronen. De input van zo'n neuron is gewoon een getalwaarde zoals de waterhoogte bij bijvoorbeeld Lobith de laatste vijf dagen. Die vijf getallen worden dan met een bepaalde berekening omgezet in een waarde. Dat is de functie van een neuron. De uitkomsten van deze neuronen wordt weer gebruikt als input van andere neuronen, die waarden gaan dan door naar de volgende laag van het model. Door dingen aan te passen aan de neuronen krijg je uiteindelijk een som die, hoe verder je door gaat steeds dichterbij de verwachte waarde komt.

Een Deep Learning Network is in principe hetzelfde alleen in plaats van een grote box neuronen zijn ze verdeeld in lagen. Zodat de output van de ene laag telkens de input van de volgende laag is. Er is dus bijvoorbeeld geen directe verbinding tussen laag een en vijf. Als een model veel lagen achter elkaar heeft, dan is er sprake van een Deep Learning Network.

**Wat zijn de voordelen van zo'n Deep Learning Network ten opzichte van een standaard Neural Network?**

Hoe meer lagen een model heeft, hoe meer berekeningen je kan doen, hoe preciezer je een berekening kunt hebben. Als je kijkt op het eerste niveau, dan kan je wel waarden optellen met elkaar, maar dan kan je nog niet zeggen; stel je hebt een input van nul en een input van een dan kan je wel zeggen dat het totaal een is, maar dan kan je niet zeggen dat altijd als de ene een is en de andere nul moet er een uitkomen. In een laag krijg je dat niet voor elkaar.

**Dus door meer lagen kan het model beter de verwachte uitkomst bepalen, maar wordt de kans op overfitting dan ook niet groter?**

Ja klopt, dus daar moet je een goede balans tussen zien te vinden.

**Uit hoeveel lagen bestaat LobithAI?**

Volgens mij zitten er zo'n zeven of acht lagen in. Dat is niet heel groot. De input is nu per dag, maar als dat per uur zou zijn, komen er een stuk of 6 lagen bij.

**Waarom heb je die grootte gekozen?**

Dat komt door het aantal inputs. De techniek die ik heb gebruikt, de Temporal Convolution zorgt er ongeveer voor dat elke laag verder de grootte halveerde. Er zijn 13 locaties dus daarom zijn er ongeveer zes lagen nodig.

**Hoe goed het model werkt hangt natuurlijk heel erg af van de trainingsdata, welke data heb je daarvoor gebruikt?**

Waterstanden van Lobith en een aantal locaties daarvoor en de data van 2011 tot eind 2018 is daarvoor gebruikt.

**Daar zit, behalve de droogte in 2018, niet echt hele extreme data bij. Hoe weet je dat hoe het model op dat bereik functioneert?**

Er zaten wel een paar pieken met hoogwater bij maar niet heel extreem, want zulke data had ik simpelweg

niet beschikbaar.

Er waren dus wel een of twee pieken met ongeveer 15 meter hoog water en daar presteerde het beter dan LobithW.

**Zou het model beter werken als je meer data als trainingsdata zou hebben gebruikt?**

Ja, tot op zekere hoogte wel, maar ik had dat ik niet beschikbaar dus kon dat ook gewoon niet gebruiken. Het is echter wel zo dat de rivier de Rijn veranderd over de jaren heen, dus als je de waterstanden van bijvoorbeeld 1993 hebt, komt er wel net iets anders uit.

**Hoelang duurt het ongeveer om het model te trainen?**

Ik had de data opgesplitst in 25 stukken, met trainingsdata en validatiedata. En om met een zo'n validatiedeel te trainen duurde ongeveer een uurtje volgens mij. Dus dat gaat dan om die zeven jaar behalve 3 maanden die gebruikt worden voor validatie en dat gaat er dan meerdere keren doorheen totdat de resultaten met de validatiedata niet meer beter worden.

**LobithAI is eigenlijk bedoeld als mogelijke opvolger voor LobithW toch?**

Dat was mijn opdracht niet. Mijn opdracht was om te kijken; wat kan er? Wat is de toegevoegde waarde? Zijn er punten waarop het beter is en zijn er punten waarop het slechter is? Dan mogen ze daarna zelf weer beslissen wat ze er mee willen.

**Welk van de twee modellen is makkelijker om te gebruiken?**

Zodra je een scriptje hebt geschreven is het voor allebei input erin en klaar. Dus dat maakt niet zoveel uit eigenlijk. Op het moment is LobithW makkelijker omdat dat al in RWsOS zit, maar zodra LobithAI er ook in zit is het net zo makkelijk.

**Bij welk model is het opnieuw kalibreren makkelijker?**

Dat is waarschijnlijk LobithAI, want dat kan je in een uur opnieuw trainen als je de data beschikbaar hebt. De laatste keer dat LobithW gekalibreerd was, was dat de scriptie van een bachelor student, dus dat kost wel iets meer tijd dan een uurtje. Dat komt vooral omdat voor LobithW een paar locaties voor een paar dagen worden uitgekozen en LobithAI neemt alles mee en gebruikt de trainingsdata om de invloed van die locaties te bepalen.

**Zou het model daardoor ook langer mee kunnen gaan?**

Zolang je het blijft kalibreren lijkt me dat wel.

**Welke van de twee modellen is sneller?**

Met LobithW zal het niet zoveel verschillen, want dat duurt voor beide modellen maar ongeveer een halve seconde. Maar in vergelijking met bijvoorbeeld het SOBEK model is het wel een stuk sneller.

**Als ik jouw verslag goed herinner is LobithAI over het algemeen beter dan LobithW, klopt dit?**

Klopt, maar de vergelijking is niet helemaal goed gelukt omdat de data gewoon verschillend was. De data waarmee LobithW is gekalibreerd is van 1980 tot 2010 dacht ik en die data had ik niet. Dus daarom is de vergelijking niet helemaal eerlijk.

**Als de kalibratie wel met de zelfde data gebeurd zou zijn, welk model zou dan beter werken denk je?**

Ik denk dat LobithAI dan net iets beter is omdat LobithW, als lineair regressie model, maar een paar berekeningen heeft dat het uit kan voeren, waardoor LobithAI iets flexibeler kan zijn in het omgooien van berekeningen en daardoor speciale gebeurtenissen misschien net op een andere manier kan berekenen.

**Kan LobithAI zoals het nu is geïmplementeerd worden in RWsOS of zou het daarvoor nog aangepast moeten worden?**

Ik weet niet precies RWsOS in elkaar zit, dus dat weet ik eigenlijk gewoon niet.

**Denk jij dat LobithAI of iets vergelijkbaars door Rijkswaterstaat gebruikt gaat worden?**

Ja ik denk het uiteindelijk wel. Ik denk wel daar er nog meer van het soort onderzoeken dat wij aan het doen zijn gaan komen. Rijkswaterstaat is ook wel op de veiligheid en dubbelcheck gericht dus willen ze graag een ander model dat dingen op een andere manier berekend om mee te kunnen vergelijken. Op de langere termijn gaat het waarschijnlijk LobithW vervangen omdat het kalibreren zoveel makkelijker is.

**Hoelang denk je dat dat nog gaat duren?**

Een jaar of vijf denk ik.

**Denk je dat er ook andere delen van het RWsOS vervangen kunnen worden door modellen met Artificial Intelligence?**

Ik denk het wel ja.

**Welke van de twee modellen zou makkelijker uit te breiden zijn?**

LobithAI is met de juiste kennis makkelijk om uit te breiden omdat het dan ook opnieuw gekalibreerd zou moeten worden en dat veel sneller is bij LobithAI. Ook andere dingen zoals sneeuwhoogte en temperatuur kunnen prima toegevoegd worden.

## B Interview met experts bij Rijkswaterstaat

**Welkom, fijn dat ik jullie mag interviewen. Ik heb het interview in twee delen opgesplitst. Het eerste deel gaat over LobithW, want ik weet ook nog niet precies hoe dat werkt. Daarna heb ik wat vragen over de vergelijking tussen LobithW en LobithAI en Artificial Intelligence in het algemeen.**

**Als ik het goed heb begrepen is LobithW een meervoudig lineair regressie model, wat houdt dat in?**

Een meervoudig lineair regressie model houdt in dat je heel veel data van het Rijngebied, zowel neerslagen, afvoeren en waterstanden, in een grote bak gooit en je laat statistische relaties bepalen door een stukje software. In ons geval is dat met SPSS gedaan. En dan komen daar lineaire relaties uit van bijvoorbeeld; nul komma zoveel keer de afvoer bij Keulen plus nul komma zoveel keer de afvoer bij Ruhrort is gelijk aan de afvoer morgen bij Lobith. En de data die er in gaat is op dag basis.

**Dus dat lijkt eigenlijk best wel veel op Artificial Intelligence, alleen dan zijn de verbanden altijd lineair?**

Ja klopt.

**LobithW wordt nu gebruikt in RWsOS toch?**

Ja, het is een deel van het RWsOS systeem en wordt specifiek gebruikt voor de verwachting van de eerste twee dagen bij Lobith. LobithW is ooit ontwikkeld als een model dat voor vier dagen gebruikt werd en dat is het ook heel lang. Maar op een gegeven moment zijn we over gegaan op wat meer fysisch gebaseerde modellen; HBV, SOBEK en dat soort dingen. En de conclusie was dat na een dag of twee die fysische modellen beter werden. Maar voor die eerste twee dagen is een statistisch model eigenlijk niet te verslaan in een stroomgebied als de Rijn.

**Voor de Maas bijvoorbeeld zou dat misschien wel anders zijn?**

Voor de Maas hebben we dat nooit geprobeerd.

Maar het vermoeden is dat dat niet zo werkt. Want bij de Rijn werkt dat voor die eerste twee dagen omdat eigenlijk al het water er al in zit. De meteo data is vooral bij de eerste twee dagen van min of meer marginaal belang. Ook niet helemaal zo maar het meeste water zit er al in.

Onze aanname is eigenlijk dat tot anderhalve dag vooruit de afvoer wordt bepaald door wat er al in de rivier zit. Daarna begint de gemeten neerslag een rol te spelen en nog verder vooruit begint ook de verwachte neerslag een rol te spelen. En dan worden die statistische relaties steeds minder belangrijk. Fysische processen worden dan maatgevend.

**Wat zijn de inputs voor LobithW precies?**

Dat zijn waterstanden, afvoeren, gemeten en voorspelde neerslagen.

**En welke stations worden daar voor gebruikt?**

Niet heel veel eigenlijk er zijn volgens mij zes meteo meetpunten en voor de waterstanden zijn er een stuk of 10. We kunnen zo ook wel even op de computer kijken wat het precies is, want voor de eerste en de tweede dag gaan er ook verschillende dingen in. En voor de verwachte neerslag wordt het stroomgebied in twee stukken opgeknipt, Noordelijk en zuidelijk. Dan wordt er op basis van een aantal gridpunten een

neerslagverwachting voor het noordelijke deel, dat is ten noorden van Frankfurt, en voor het zuidelijke deel, ten zuiden van Frankfurt. Die neerslag verwachting gaat tot vier dagen vooruit.

### **Wat is de data voor de kalibratie van LobithW?**

Een jaartje of vijf geleden is er dertig jaar aan data gebruikt om te kalibreren.

### **In hoeverre zaten daar ook extreme waarden bij?**

Ja best wel, '93 en '95 zaten daar bij en zaten ook hele droge jaren bij zoals '76 en '83.

### **Zou het model beter werken als er meer data was voor de kalibratie?**

Je ziet dat het elke keer een klein beetje beter wordt. Als je de kalibratie doet, en in het verleden hebben we dat elke vijf jaar gedaan maar daarmee zijn we een beetje slordig geworden eigenlijk de laatste tijd, dan bepaal je ook kwaliteitscriteria en je ziet dan dat het iedere keer een heel klein beetje beter wordt. Hoe meer data je erin stopt, hoe meer dat model leert.

### **En hoe nauwkeurig is het model op dit moment ongeveer?**

De criteria zijn afhankelijk van de zichttijd. Dus een dag, twee, drie of vier dagen vooruit gelden verschillende criteria. Maar we gebruiken het nu alleen nog maar voor de eerste twee dagen. En dan is het criterium dat het voor de een binnen de 5 centimeter moet zijn en voor 2 dagen binnen de 10 centimeter en dan 30 en 40 centimeter. En in 90 procent van de gevallen halen we dat.

### **Dus die 90 procent gaat met elke nieuwe kalibratie een klein stukje omhoog?**

Dat wordt iedere keer een heel klein beetje beter ja.

In hoeverre het echt beter wordt is een beetje moeilijk te bepalen. Dat is dan mijn persoonlijke mening, zeg ik maar direct even bij. Omdat je namelijk wel ziet dat er nog behoorlijk kwaliteitsverschil zit in welk afvoerbereik je zit. Want vaak zie je dat als je er meer data in stopt dat het tussen 8 en 10 meter verder verbeterd terwijl dat eigenlijk al vrij goed is. Maar als je dan iets hebt zoals afgelopen zomer met extreem lage afvoeren, dan doet het eigenlijk niks want zo'n tijdsserie zit er niet in.

Je ziet dat het model het goed doet bij een geleidelijk afvoerterloop en waar het slechter is zijn bochten, van dalend naar stijgend en van stijgend naar dalend. Daar zie je het model altijd doorschieten.

### **En denken jullie dat dat beter zou kunnen met Artificial Intelligence?**

Ik heb geen idee.

Ik denk het wel, en waarom denk ik dat. Ik denk a) dat er meer ruimte en mogelijkheden in zitten. Nu met SPSS heb je meervoudig lineair regressie, maar dat komt er eigenlijk nooit uit. Maar eigenlijk zou je dat veel dynamischer kunnen doen. Je zou bijvoorbeeld kunnen zeggen; Ik kan me voorstellen dat die relaties bij 11 meter Lobith andere zijn dan bij 8 meter. En dat is er bij een het huidige model niet in en ik denk dat dat een hele chique manier is om tot 10.000 kub een meer evenwichtige kwaliteit aan te kunnen koppelen. Dus dat je daardoor ook in meer extreme veranderingen dat het doorschieten van het model dat dat er meer af gaat.

Wat je eigenlijk zou willen met AI is dat; Je hebt de initiële conditie vandaag en dat het model op basis daarvan gaat kijken, wat zijn dan de goede relaties en die dan toepast in de verwachtingen.

En b) wat je ziet, bij die meervoudige lineaire regressie, nemen we een beperkt aantal stations mee. Dan zijn er niet zo heel veel, een stuk of 15 en uiteindelijk blijken er 8 of 10 relevant te zijn. Maar je zou eigenlijk nog een hele hoop andere informatie er ook in kunnen stoppen. Ik noem iets over sneeuw, ik noem iets over bodemvocht, iets over de stand van de meren in Zwitserland, verdamping of temperatuur. Je kunt er gewoon veel meer dingen instoppen die ook van invloed zijn.

**Maar op een gegeven moment moet je natuurlijk wel keuzes maken. Je kan niet altijd alles meenemen omdat daar simpelweg de beschikbaarheid van data en rekenkracht nog niet goed genoeg voor zijn, hoe zou je daar dan toch goede keuzes in kunnen maken?**

AI kan je wel zelf laten kiezen.

Je kan natuurlijk wel logisch naden; Wat zijn nu echt de invloedsfactoren. En dat is meer dan alleen de afvoer bij Keulen of de afvoer bij Bazel.

Als je begint met de kijken naar welke stromen er nou eigenlijk in een fysisch model zitten. Al dan niet ingevoerd, al dan niet gefixt in bakjes of in een factor terwijl we daar wel al data voor hebben.

We weten tegenwoordig ook al veel meer dan bijvoorbeeld tien jaar geleden. Bijvoorbeeld sneeuw, dat zit nu dus in een bakje ingebakken maar we hebben sneeuw ook via modeldata. Hartstikke mooi, dus als je dan nu die data gebruikt in plaats van een vastgestelde waarde dan denk ik dat je een sprint gaat maken in kwaliteit. In die situatie natuurlijk. Je kan bijvoorbeeld misschien ook een correlatie vinden tussen de verkoop van het aantal sigaretten in de zomer maar dat is natuurlijk niet iets dat relevant is. Er zijn mensen die daar ook heel exotisch over denken maar ik denk dat als je nu kijkt van wat voor aspecten zitten er nu in een fysisch model. Eigenlijk zou je die dan in die database van data moeten instoppen en dan daaruit relaties hebben door AI te laten bepalen, dan hebben we al een hele mooie LobithAI 1.0 denk ik. En in de verdere toekomst, ook met toenemende rekenkracht, dan kan het model ook leren wat is nu echt relevant en dat dan ook in het voorspelproces mee te nemen. Dat je zegt van we hebben nu de initiële conditie en de AI bepaald dan op dat moment wat hij dan op die dag de relevante regressie vindt. En dat is wel iets voor later, want dat is echt nog wel erg complex, maar dat is wel waar ik denk dat die LobithAI echt vol tot zijn recht komt.

**LobithAI, zoals het model gemaakt is door geïnterviewde 1, lijkt qua inputs heel erg op LobithW, maar dat is dus eigenlijk nog niet genoeg om AI maximaal te kunnen benutten?**

Ja dat vermoeden heb ik wel, vooral dat je dus ook andere informatie mee zou moeten nemen.

Dan wordt het pas echt interessant. Want 1 cm beter bij 12 meter op de derde dag maakt niet zoveel uit, want er komt ook nog best wel veel expert-adjustment bij kijken. Als het zo is dat wij het vertrouwen hebben dat het model over het gehele afvoerbereik goed functioneert en dat het maximale eruit haalt wat er is aan kennis, dan maken we echt een stap. En vooral als je model hebt dat dat over de volle breedte kan dan heb je ook iets wat nog niet heel veel gedaan is. Dan heb je ook echt wel een mooi model waar iedereen zijn vingers bij af likt denk ik.

**Wordt er ergens anders wel al zoiets gebruikt voor waterhoogtevoorspellingen of loopt Nederland wel voorop op dat gebied?**

Ik denk niet dat wij echt voorop lopen, het type model wat wij gebruiken, en combinatie van een hydrologisch en hydraulisch model is vrij gangbaar. De combinatie van een statistisch en een fysisch model is wel



bijzonderder, die heb ik nog niet zo vaak gezien. Maar dat komt omdat de rivier zich daar ook wel voor leent want de Rijn is eigenlijk een luie rivier. De Donau bijvoorbeeld dat klinkt ook wel heel leuk, maar daar stroomt alles ook veel sneller, dus die doorstroomtijden zijn daar ook allemaal veel korter. Dus daarom is het voorspellen voor de Rijn ook veel makkelijker dan bijvoorbeeld de Thames. We hebben daarmee ook wel gewoon randvoorwaardelijk geluk, wat ons zeker mogelijkheden geeft.

### **Zou het moeilijk zijn om dit te implementeren in FEWS/RWsOS?**

Dat kan je eigenlijk het beste aan je collega's bij Deltares vragen, maar ik neem aan van wel omdat ze ondertussen de wereld aan model al in FEWS hebben geïmplementeerd.

Implementatie heeft natuurlijk ook met de hardware kant te maken. Het ligt er dus aan hoe uitgebreid het model is. Als je elke ochtend bij de initiële conditie de relaties laat bepalen dan zit daar wel een heel stuk 'bare metal' achter te knallen natuurlijk. Dus wat het uiteindelijk gaat worden en wat je dan nodig hebt aan hardware dat moet welk op een bepaald moment bekend zijn, want de doorlooptijd om het dan ook uiteindelijk te krijgen kan wel lang zijn bij Rijkswaterstaat.

We moeten natuurlijk rekening houden met de rekentijd, we moeten gewoon binnen twee uur het proces van data inwinning tot het maken van ons bericht gedaan hebben. Dus als wij straks een model hebben dat vier uur aan het stampen is, daar hebben we er niks aan. Het is zegmaar mooi als we het weten maar als het de situatie beschrijft van twee dagen geleden, dan hebben we er niks aan. Dan wisten we heel goed wat het had moeten zijn.

Ik zie echt wel meerdere stappen omdat er dan natuurlijk ook nog onderzoek gedaan moet worden, van wat is nou de set die relevant is? Maar de techniek en de investeringen die daar bij horen, moeten daar ook simultaan in meelopen.

### **Een model op basis van AI is natuurlijk een black box, levert dat bij jullie geen problemen op met het gebruikt ervan?**

Nee, want dat hebben we bij statistische modellen ook. Daar zitten ook relaties in die eigenlijk helemaal niet kunnen kloppen. De afvoer bij Lobith morgen heeft bijvoorbeeld een relatie met de afvoer bij Bazel van gisteren, maar dat kan helemaal niet vanwege de stroomsnelheid van de rivier. Dus in wezen is LobithW ook een black box, dus kan daar net zo goed mee gewerkt worden als met AI.

En voor mij ook hoor, jij weet er nog iets meer van. Bij de oprichting van Deltares zijn de modellen naar Deltares gegaan. Voor die tijd werden de modellen gewoon binnen Rijkswaterstaat ontwikkeld, dus wisten wij hier nog beter wat de modellen deden. Maar sinds die tijd zijn de modellen zoals SOBEK en HBV eigenlijk ook gewoon een black box.

Ik vind jouw vraag heel goed en interessant ook, want ik ben daar niet gelukkig mee dat AI een black box is, maar er veranderd niks want dat is nu ook al zo voor ons. Maar nu zouden we nog naar Deltares kunnen bellen met de vraag hoe iets zit, maar met AI weet niemand het meer. Maar daarom is het denk ik ook goed om de SOBEK HBV lijn om die er bij te houden zodat de resultaten daar altijd mee vergeleken kunnen worden.

### **Maar meerdere modellen naast elkaar draaien kost wel weer extra rekenkracht en dus extra geld. Is dat geen probleem?**

Dat is zeker waar, dus daar kijken we ook altijd kritisch naar. Maar op dit moment is er voor gekozen om en het statistische model en de fysische modellen allebei te gebruiken. De modellen zoals we ze nu gebruiken werken met tijdstappen van een dag, maar eigenlijk willen we naar tijdstappen van een uur. Dat zou natuurlijk ook weer meer rekenkracht gaan kosten. Maar daar moeten we natuurlijk een balans in zien te gaan vinden van wat er nog zinvol is om te doen.

**Om weer even terug te gaan naar de black box. Stel je voor dat er gevaarlijk hoog water komt en jullie doen dan een voorspelling. Waar ligt dan de verantwoordelijkheid?**

De verantwoordelijkheid voor het uitgeven van de waarschuwing ligt bij ons. Wij voegen daar altijd onze expert judgement aan toe. Dus we zullen nooit ineens een waterstand uitgeven die zes meter te hoog ligt bijvoorbeeld. Dat hebben we wel door dat dat niet klopt.

Maar wij zijn inderdaad verantwoordelijk voor de kwaliteit van de uitgifte. Mocht het misgaan, dan worden wij daar dus op aangesproken en het is ook nooit zo dat we klakkeloos een modeluitkomst overnemen. Daar wordt altijd nog kritisch naar gekeken voor de uitgifte. Dus als een model als LobithAI gebruikt zou worden blijft de verantwoordelijkheid hetzelfde.

**Zouden jullie liever de modellen zoals ze nu zijn blijven gebruiken of willen jullie wel van AI gebruikmaken?**

Ik zie het meer als een toevoeging, niet als vervanging. Vooral omdat we wel willen begrijpen wat er gebeurt. Dus de fysische modellen zou ik wel willen blijven gebruiken.

AI is voor mij meer een middel, het is voor mij geen doel. Mijn werk wordt voor mij niet beter als er AI in mijn systeem zit. Mijn werk wordt beter als ik betere voorspellingen en kortere tijdstappen krijg. Of dat met AI of een multi multi lineair regressie model of met een nieuw heel nauwkeurig beschreven fysisch model vanuit de TU Delft maakt me niet uit. Als het maar beter wordt.

**Zouden er ook nog andere delen of modellen in RWsOS vervangen kunnen worden door AI?**

Ik had er net al op gehint, maar stuwmeergedrag dat is informatie die heel moeilijk te verkrijgen is vanuit Wallonië, daar zie ik ook echt wel grote meerwaarde. Vooral over de menselijke invloeden op het systeem; hoe worden stuwen ingezet. Je ziet bijvoorbeeld dat ze bij laag water op een hele andere manier met hun kunstwerken omgaan. Waterbesparende schutten en dat soort dingen. Daar wordt nu in de modellen geen rekening mee gehouden, dus moeten we achteraf gaan uitvogelen hoe dat zit. Terwijl over die menselijke invloeden die je dus niet in het model kan vatten, AI wel een zinvolle uitspraak kan doen. Dus ik denk dat het meerwaarde kan hebben om delen van fysische modellen op AI te baseren.

**Als ik het zo hoor zijn jullie wel enthousiast over AI in modellen, denken jullie ook dat het binnenkort al gebruikt gaat worden in RWsOS?**

Dat klopt maar we willen natuurlijk wel eerst zien hoe goed het model werkt. En als het blijkt dat dat meerwaarde heeft dan gaan we dat zeker gebruiken. We denken zeker dat het kansrijk is.

We hebben op dit moment een behoefte niet meer helemaal aansluit bij wat we nu hebben. Dus de drang om iets anders te hebben dat beter aansluit op de vraag is groot, vooral bij extreem lage afvoeren net als de wens voor kleinere tijdstappen.