

Wat gaat AI opleveren voor het NHI?

*NHI-dag 18 januari 2024 -
Blik op de Toekomst met het oog op Innovatie, Verbinding & Community Building*

Bennie Minnema

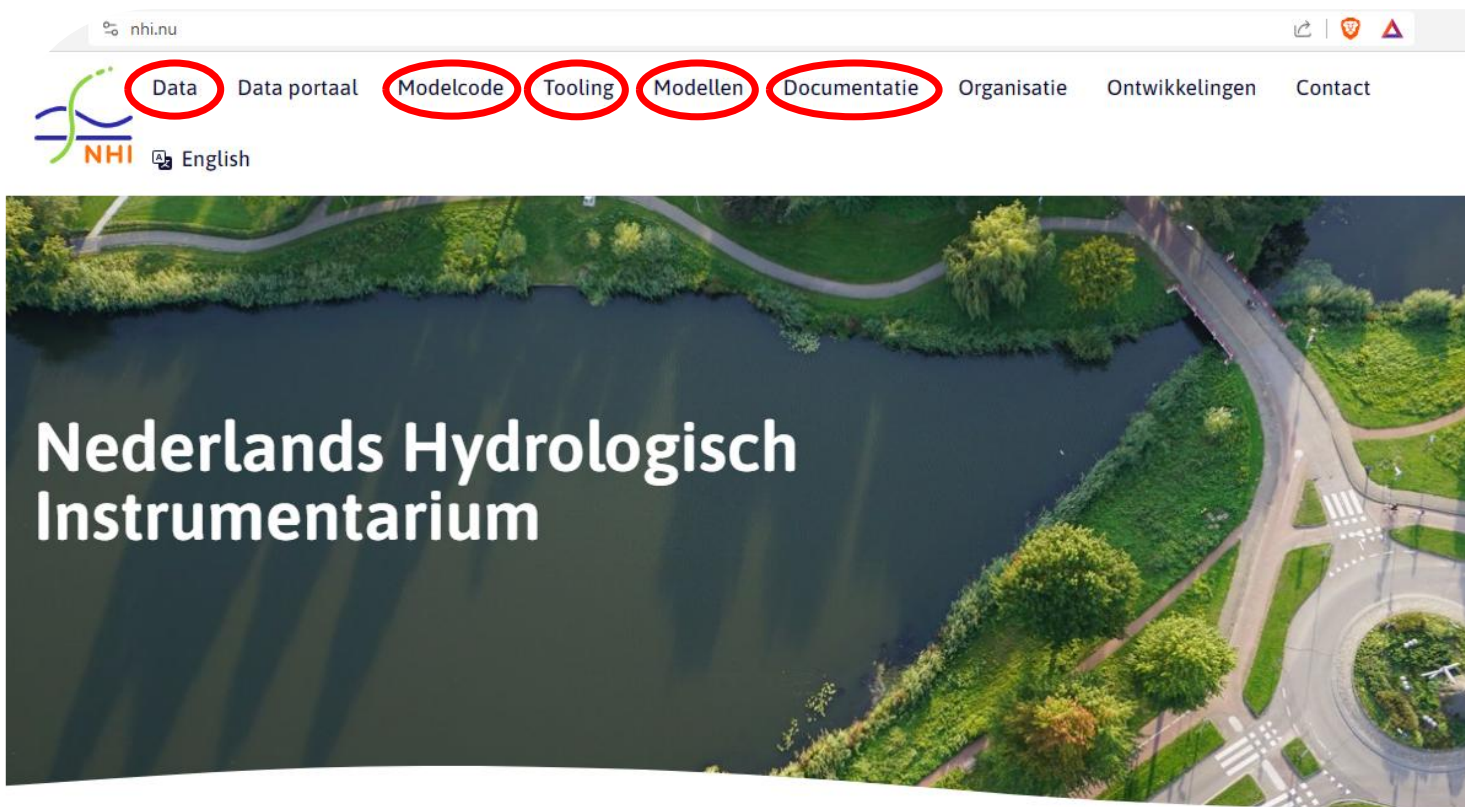
*Met dank aan
Antonio Moreno Rodenas – Enabling Technologies Data Science*

AI: Artificiële of kunstmatige Intelligentie

Definities (bron: Chat-GPT):

- Artificiële Intelligentie (AI) is de theorie en ontwikkeling van computersystemen die taken kunnen uitvoeren waarvoor normaal *menselijke intelligentie* nodig is, zoals visuele waarneming, spraakherkenning, besluitvorming en vertaling tussen talen. ... en zichzelf herhaaldelijk kan verbeteren op basis van vergaarde informatie.
- Machine learning: een subset van AI waarbij algoritmen in de loop van de tijd hun prestaties kunnen verbeteren als ze aan meer gegevens worden blootgesteld.
- Deep learning: een subset van machine learning waarbij meerlaagse neurale netwerken leren van enorme hoeveelheden gegevens
- Neuraal netwerk: een reeks algoritmen die worden gebruikt als proces bij machinaal leren en die patronen en relaties in grote hoeveelheden gegevens kunnen herkennen. Neurale netwerken gebruiken een logische structuur geïnspireerd door het menselijk brein en vormen de basis voor deep learning-algoritmen.

Scope van deze presentatie: AI n NHI



Artificiële Intelligentie als onderdeel van 'data science'

AI-'werkgebieden':

1. Generatieve AI
 - Large Language Models (LLM's)
 - b.v. ChatGPT van OpenAI: geavanceerde chatbot
2. Het AI-gebaseerd verwerken en analyseren van grote (meet)datasets
 - Foto's, video's, remote sensing data
3. 'Data-driven' modelbouw
 - Classificatie
 - Regressie
 - Clustering
4. Emulators of 'Surrogate Models'
 - Nabootsen van bestaande numerieke modellen

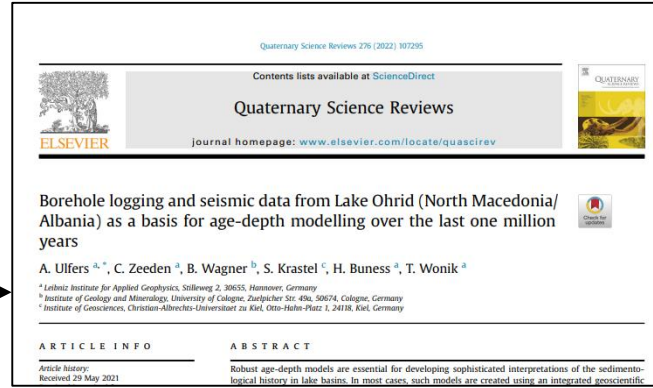
Artificiële Intelligentie als onderdeel van 'data science'

1. Generatieve AI
 - Large Language Models (LLM's)
 - b.v. ChatGPT van OpenAI: geavanceerde chatbot
2. AI-gebaseerd verwerken en analyseren van grote (meet)datasets
 - Foto's, video's, remote sensing data
3. 'Data-gedreven' modelbouw
 - Classificatie
 - Regressie
 - Clustering
4. Emulators of 'Surrogate Models'
 - Nabootsen van bestaande numerieke modellen

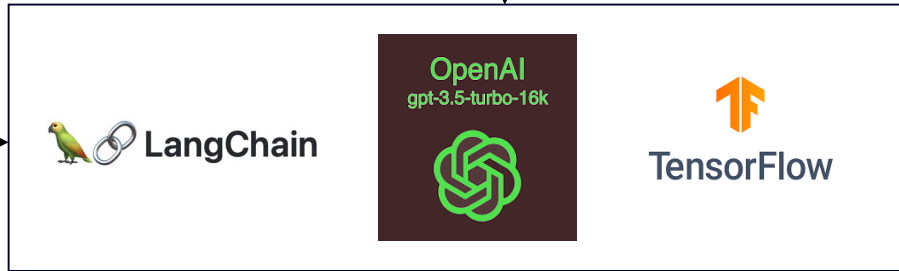
Extractie van grondwater data uit 'grey literature'

Data mining based on OpenAI API (ChatGPT)

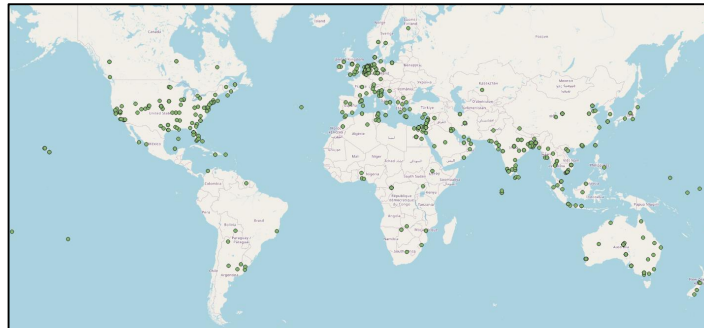
Document with text/images (e.g. reports, engineering files, scientific articles)



Python script with APIs



Contextual query: (e.g. give me all borehole and seismic essays in Albania)



Catalogued and georeferenced data

Deltares

Jude King, Deltares
Ali Meshgi, Deltares
Gualbert Oude Essink, Deltares

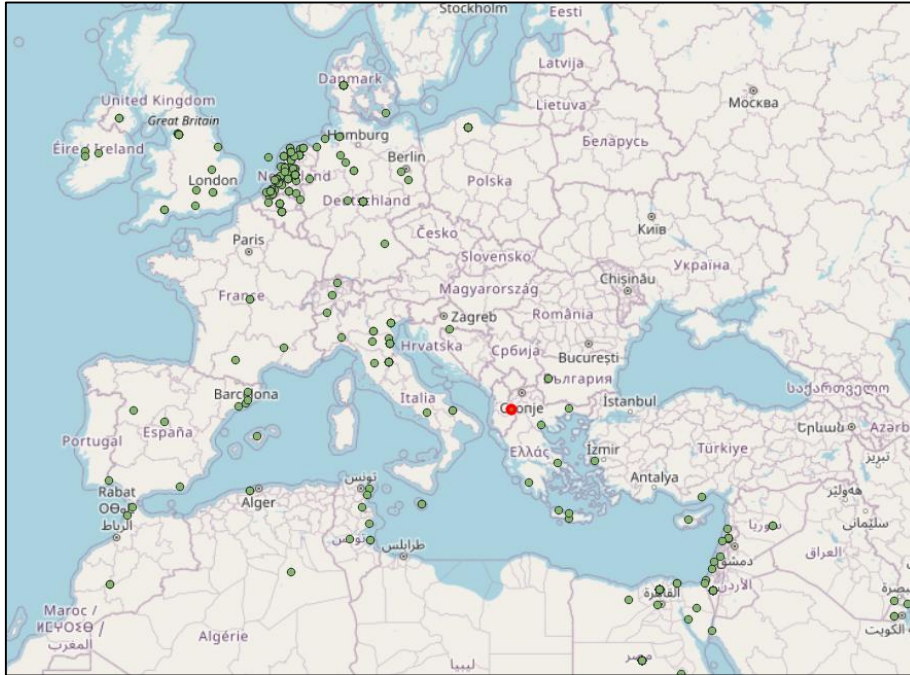
- Custom data mining routine developed for extraction of groundwater model data
- Based on **large language models** and trained **neural network image classifier**
- Automated extraction of parameters of interest, including georeferencing
- Automated classification of image type (work in progress!)
- Computation time ~6hrs per 1000 documents
- API calls and prompts handled by LangChain
- Tested on 2000+ documents

Data_Mine_V1	
title	Borehole logging and seismic data from Lake Ohrid (North...
field_1	1
title	Borehole logging and seismic data from Lake Ohrid (North...
year	2022
country	North Macedonia/Albania
region	Lake Ohrid
url	https://doi.org/10.1016/j.quascirev.2021.107295
keywords	['age-depth modelling', 'downhole methods', 'seismic inte...
data_links	[]
Y_country	41,139945980000000
X_country	20,065076829999999
Y_region	41,060115809999999
X_region	20,731275560000000
directory	C:\Users\king_je\Projects\Text_Mining\pdfs\1-s2.0-S02773...

Extractie van grondwater data uit 'grey literature'

Output in GIS

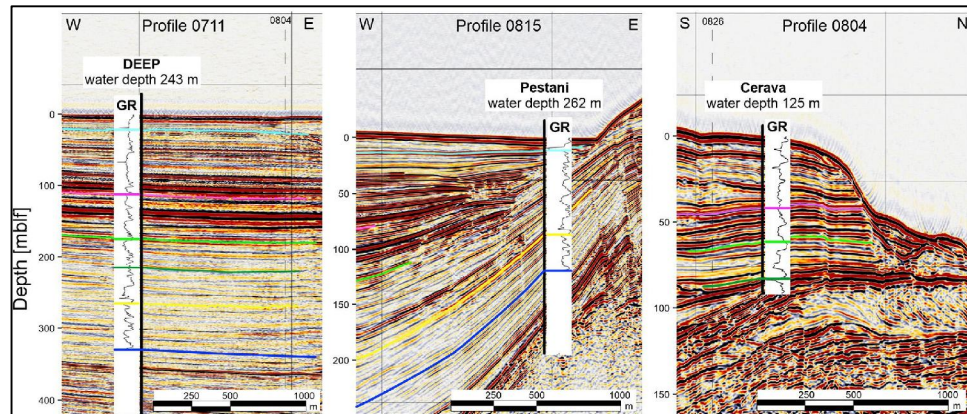
1. Select georeferenced data



2. Check information

Data_Mine_V1	
title	Borehole logging and seismic data from Lake Ohrid (North...
(Derived)	
(Actions)	
field_1	1
title	Borehole logging and seismic data from Lake Ohrid (North...
year	2022
country	North Macedonia/Albania
region	Lake Ohrid
url	https://doi.org/10.1016/j.quascirev.2021.107295
keywords	['age-depth modelling', 'downhole methods', 'seismic inte...
data_links	[]
Y_country	41,139945980000000
X_country	20,065076829999999
Y_region	41,060115809999999
X_region	20,731275560000000
directory	C:\Users\king_je\Projects\Text_Mining\pdfs\1-s2.0-S02773...

3. Extract automatically classified images



Sneller betere softwarecode schrijven

Gemakkelijker schrijven van (tooling)software:

- [GitHub Copilot](#): ontwikkeld door GitHub in samenwerking met OpenAI, werkt als een virtuele mede-programmeur die ontwikkelaars helpt om sneller betere code te schrijven.

Recent (januari 2024):

- OpenAI:
"We're launching the GPT Store to help you find useful and popular custom versions of ChatGPT"

Artificiële Intelligentie als onderdeel van 'data science'

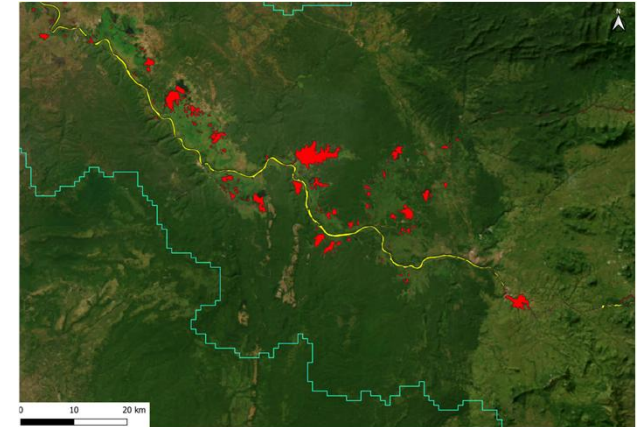
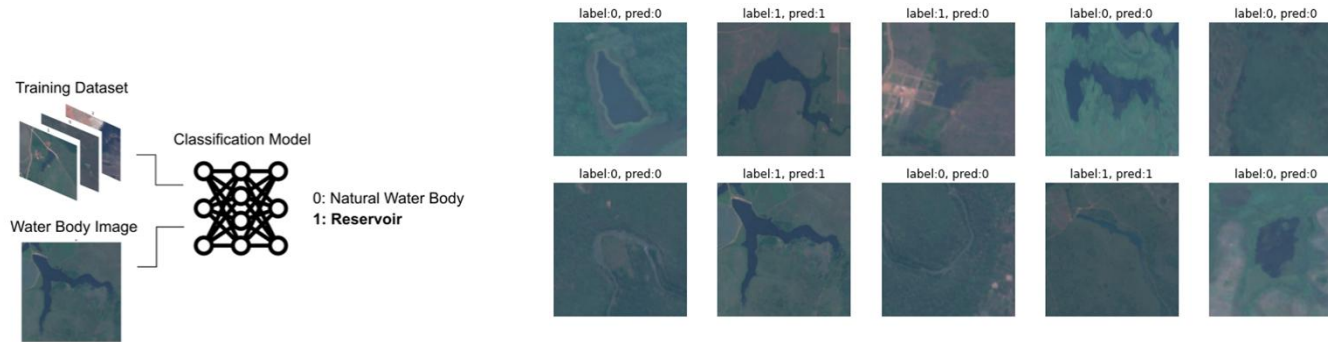
1. Generatieve AI
 - Large Language Models (LLM's)
 - b.v. ChatGPT van OpenAI: geavanceerde chatbot
2. Verwerken en analyseren van grote (meet)datasets
 - Foto's, video's, remote sensing data; voorbeeld www.globalwaterwatch.earth
3. 'Data-gedreven' modelbouw
 - Classificatie
 - Regressie
 - Clustering
4. Emulators of 'Surrogate Models'
 - Nabootsen van bestaande numerieke modellen

Computer Vision CNN: Automated detection of infrastructure from remote sensing (Dams detection)

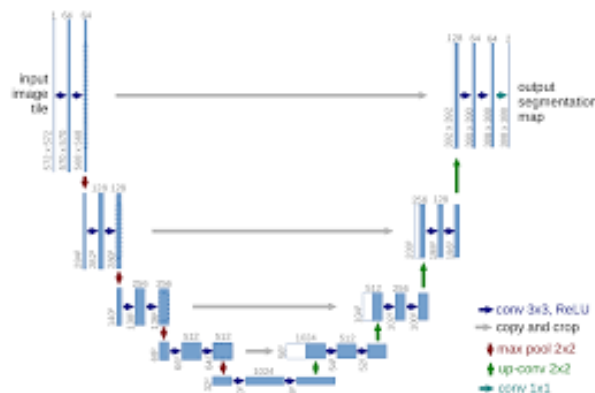
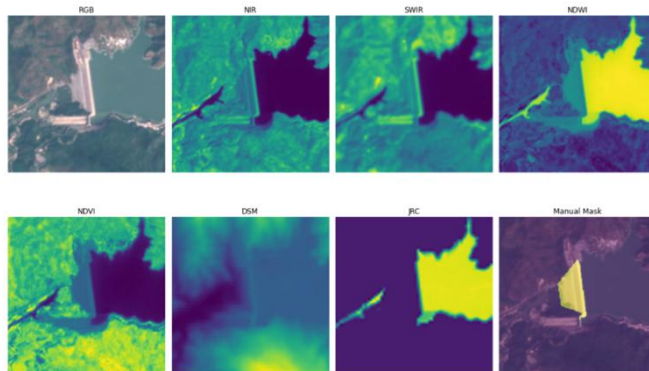
Isabel Gerritsma
 Antonio Moreno Rodenas
 Arjen Haag
 Mike Woning
 Deltares

- Missing infrastructure (dams) in public/government records. Knowing the location, type and characteristics of dams is essential for quantifying and managing risk.

1- Automated detection/classification of reservoirs



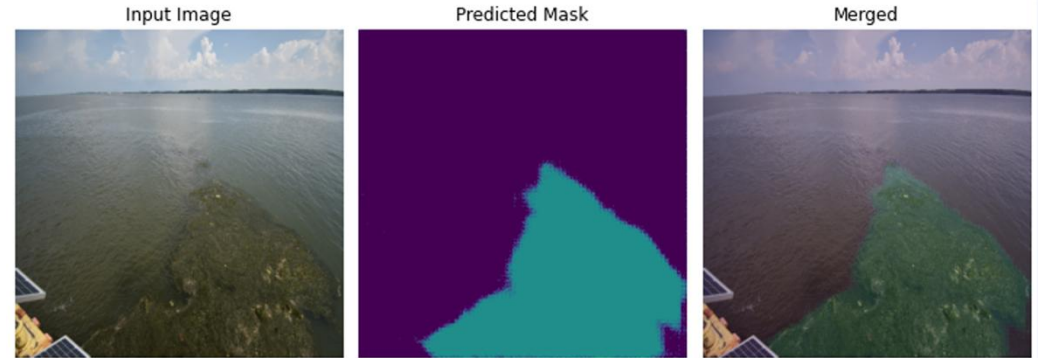
2- Detection and segmentation of dam bodies and dam type (relevant for failure probability)



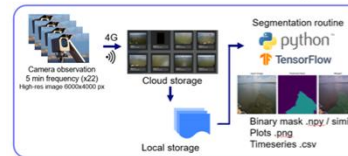
Computer Vision CNN: Applications in the field

Bas van Vossen
Antonio Moreno Rodenas
Deltares

Early warning systems voor drijvende vegetatie-events



Working scheme automated detection eelgrass



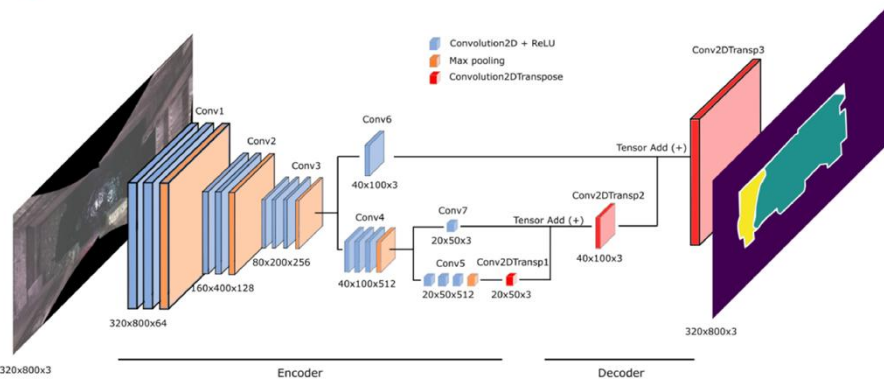
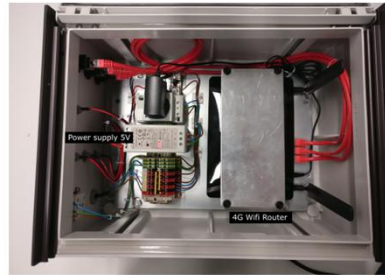
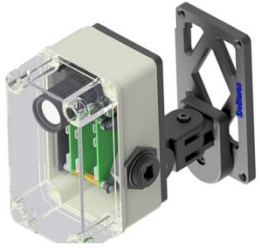
Visualization and integration

FATracker

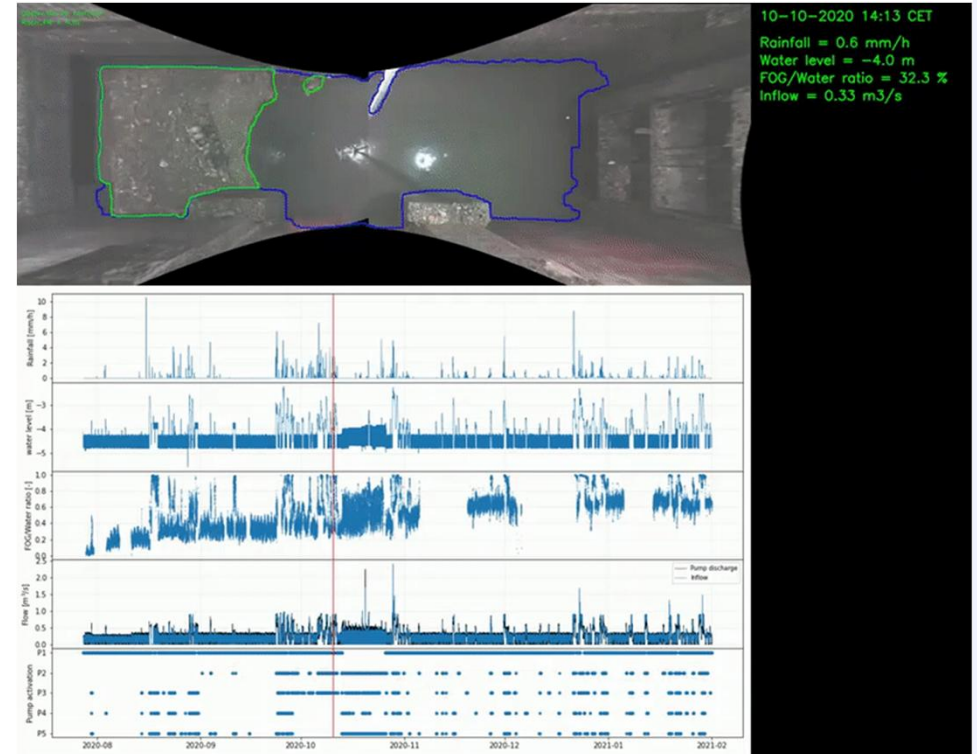
Antonio Moreno Rodenas, Deltares
Alex Duinmeijer, Ingenieursbureau Gemeente Rotterdam
Francois Clemens-Meijer, Deltares

Pumping station fat-oil-grease detection

Erik Laurentzen, Gemeente Arnhem
Raymond de Wit, Waterschapsbedrijf Limburg



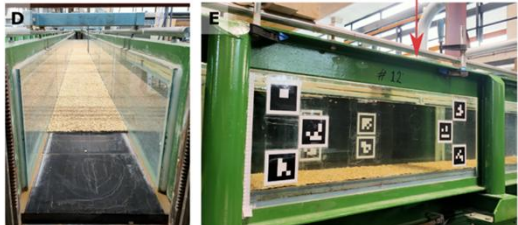
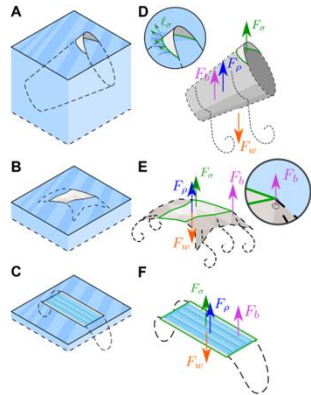
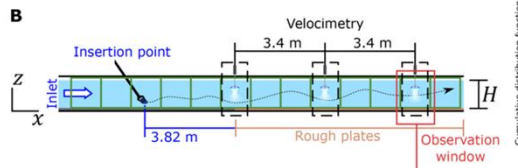
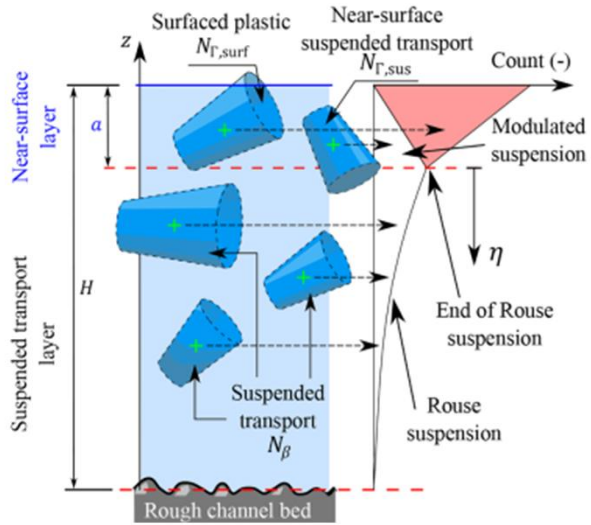
- Convolutional Neural Network (Encoder-Decoder architecture)
- VGG - Imagenet weights
- Retrained for FOG detection with 124 manually labelled images



FATracker wint [RIONED innovatieprijs 2022](#)

Computer Vision CNN: Laboratory setups

Hydraulic Laboratory – Automated plastics detection and 3D path reconstruction.
Image processing for targeted object detection and time-space tracking.



Valero, D., Belay, B. S., Moreno-Rodenas, A., Kramer, M., & Franca, M. J. (2022). The key role of surface tension in the transport and quantification of plastic pollution in rivers. *Water Research*, 226, 119078.

Artificiële Intelligentie als onderdeel van 'data science'

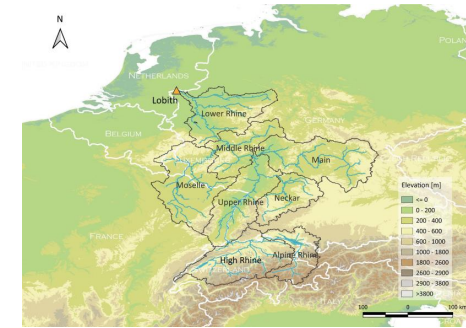
1. Generatieve AI
 - Large Language Models (LLM's)
 - b.v. ChatGPT van OpenAI: geavanceerde chatbot
2. AI-gebaseerd verwerken en analyseren van grote (meet)datasets
 - Foto's, video's, remote sensing data
3. 'Data-gedreven' modelbouw
 - Classificatie
 - Regressie
 - Clustering
4. Emulators of 'Surrogate Models'
 - Nabootsen van bestaande numerieke modellen

Operational low-flow forecasting model bij Lobith

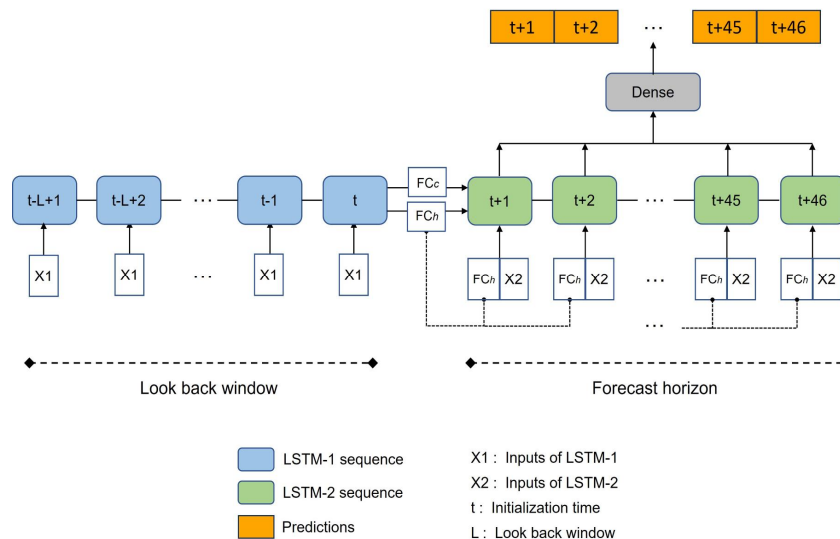
zeer recent voorbeeld, zojuist (17-1-24) gepubliceerd

Deng J., Couasnon A., Dahm R., Hrachowitz M., van Heeringen K.J., Korving H., Weerts A. and Taormina R. (2024). Operational low-flow forecasting using LSTMs. *Front. Water* 5:1332678. <https://doi.org/10.3389/frwa.2023.1332678>

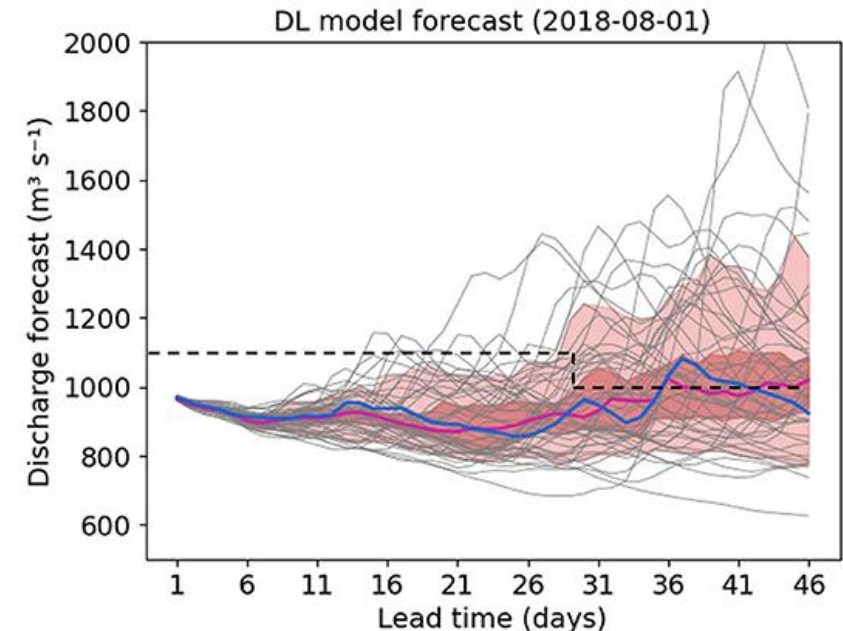
Een nieuwe, op LSTM gebaseerde modelarchitectuur is ontworpen om zowel historische waarnemingen als voorspelde meteorologische gegevens te gebruiken om de laagwatervoorspelling voor de Rijn bij Lobith uit te voeren op dagbasis met 'leadtimes' tot 46 dagen vooruit.



Eight subbasins of the Rhine River basin upstream of Lobith.



Model architecture



Forecast results of the DL model with SEAS5 initialized on 2018-08-01

Afvoerhydrologie

- Long-Short-Term-Memory (LSTM)-neurale netwerken doen het al enkele jaren kwalitatief beter dan de traditionele Rainfall-Runoff modellen:
 - Een RR-model is een 'black box model', er zitten geen deterministische concepten in, neerslag en benedenstroomse afvoer worden 'gefit'
 - Een neurale netwerk is een vorm van 'multi-dimensionale regressie' en 'ontdekt' daarbij patronen die niet in het gefitte 'black-box' model zitten.
 - Recente HESS-artikelen, door Kratzert et al. :
 - [Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets](#)
 - [Toward Improved Predictions in Ungauged Basins: Exploiting the Power of Machine Learning](#)
- Google + Kratzert:
 - organiseert 22-23 januari 2024 het 5^e(!) jaarlijkse online workshop '*Flood Forecasting Meets Machine Learning Workshop*'
 - neuralhydrology.github.io

Artificiële Intelligentie als onderdeel van 'data science'

1. Generatieve AI
 - Large Language Models (LLM's)
 - b.v. ChatGPT van OpenAI: geavanceerde chatbot
2. AI-gebaseerd verwerken en analyseren van grote (meet)datasets
 - Foto's, video's, remote sensing data
3. 'Data-gedreven' modelbouw
 - Classificatie
 - Regressie
 - Clustering
4. Emulators of 'Surrogate Models'
 - Nabootsen van bestaande numerieke modellen

Hydrodynamica; emulatie

2D flood modeling:

- D-Flow FM, 3Di, HEC-RAS, TUFLOW, etc. bevatten fysieke deterministische concepten, voor b.v. 'overstroming over maaiveld'; echter, lange rekestijden.

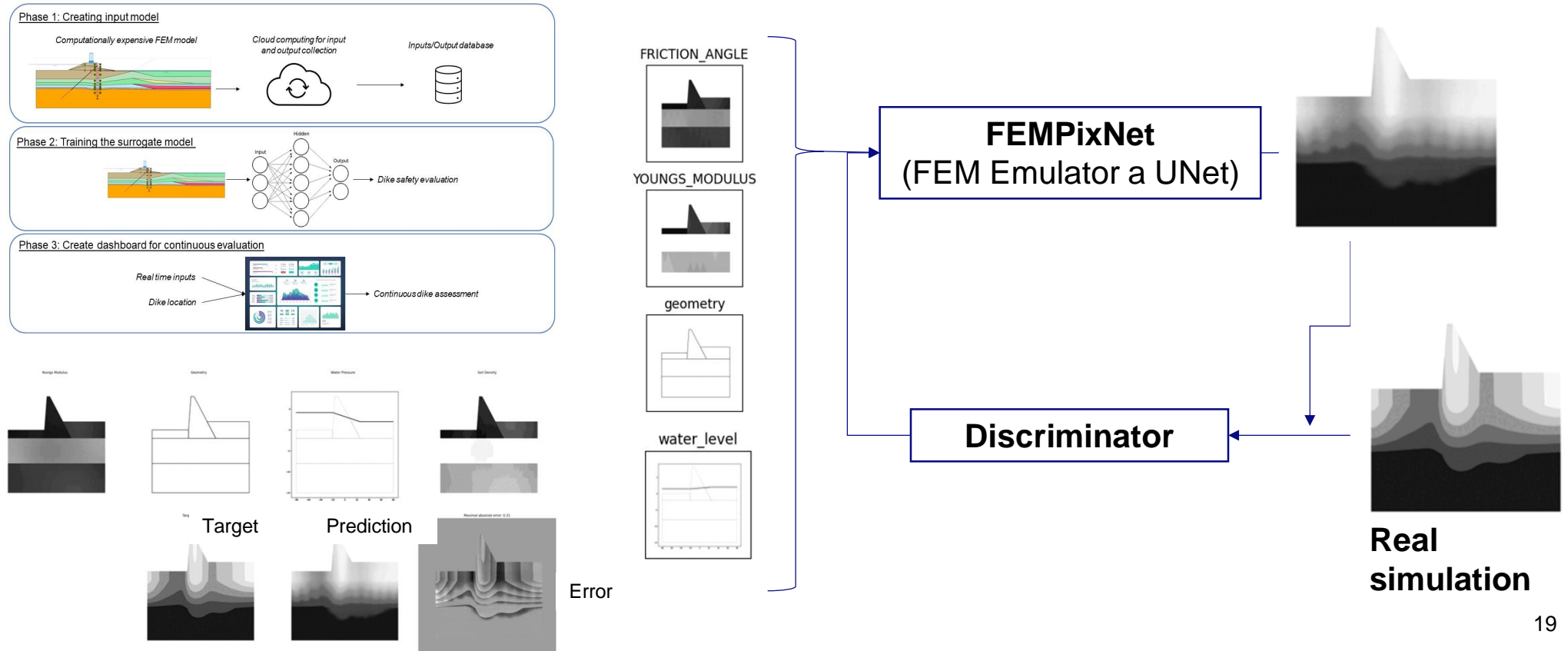
Zoeken naar snellere benaderingen:

- Recent symposium (27 september 2023) 'Emulating 2D flood modeling': <https://softwaredays.deltares.nl/-/symposium-floods-ai-2023>;
 - er worden al orde van grootte rekestijdwinsten geboekt o.a. gebruik makend van LSTM's-emulatoren
 - onderzoek lijkt nog maar net op gang te zijn gekomen, er is nog veel te doen.

Model Emulation: GAN based emulator of a sheet pile installation (FEMPixNET)

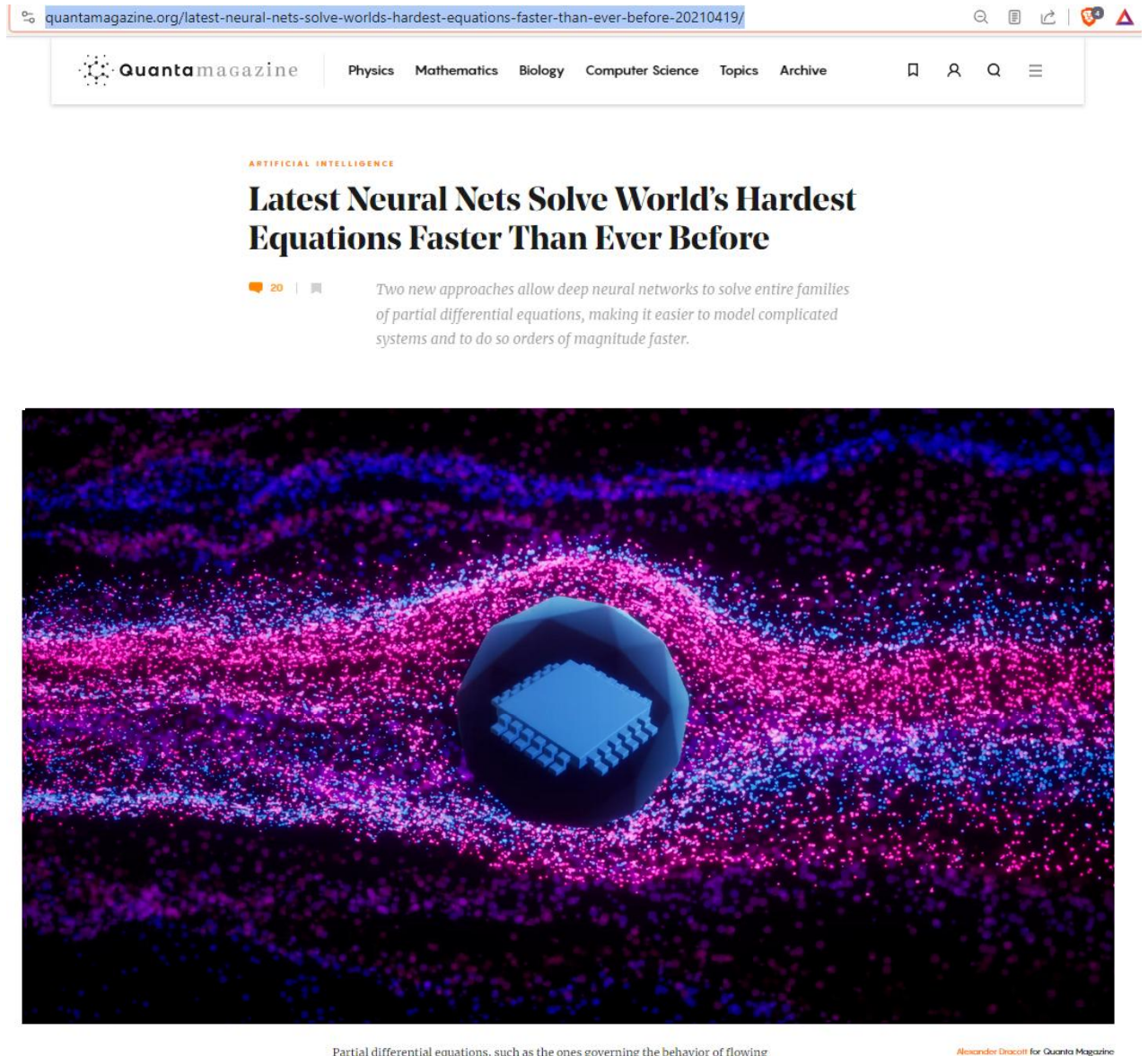
Eleni Smyrniou
Bruno Zuado Coelho
Aron Noordam
Elham Mahmoudi
Deltares

Computationally expensive FEM model of a sheet pile (Kratos + STEM)
5000 simulations (cloud-based as training set)



Internationaal ...

- Algemeen wetenschappelijk onderzoek om partiele differentiaal vergelijkingen veel sneller op te lossen
- O.a. [Brown University, Faculty Applied Mathematics, USA](#)
 - Fourier Neural Operator
 - DeepONet

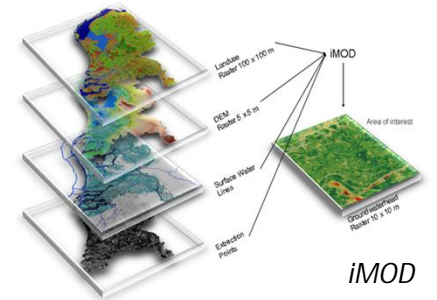


The image is a screenshot of a web browser displaying a Quanta Magazine article. The browser's address bar shows the URL: quantamagazine.org/latest-neural-nets-solve-worlds-hardest-equations-faster-than-ever-before-20210419/. The Quanta Magazine logo is visible in the top left of the page, along with navigation links for Physics, Mathematics, Biology, Computer Science, Topics, and Archive. The article title is "Latest Neural Nets Solve World's Hardest Equations Faster Than Ever Before" under the category "ARTIFICIAL INTELLIGENCE". A sub-headline reads: "Two new approaches allow deep neural networks to solve entire families of partial differential equations, making it easier to model complicated systems and to do so orders of magnitude faster." Below the text is a large, vibrant image of a neural network visualization, featuring a central blue chip-like structure surrounded by a complex, multi-colored (pink, purple, blue) particle-like field. At the bottom of the image, there is a caption: "Partial differential equations, such as the ones governing the behavior of flowing" and a credit: "Alexander Dracott for Quanta Magazine".

Hoe ben ikzelf eigenlijk in de AI-wereld verzeild geraakt?

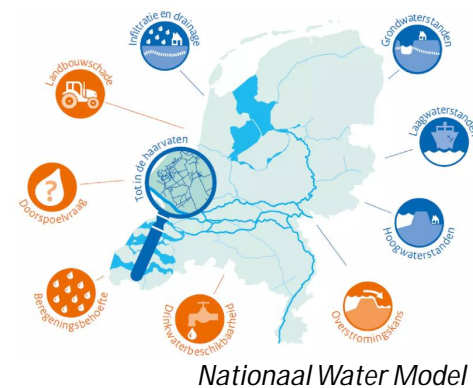
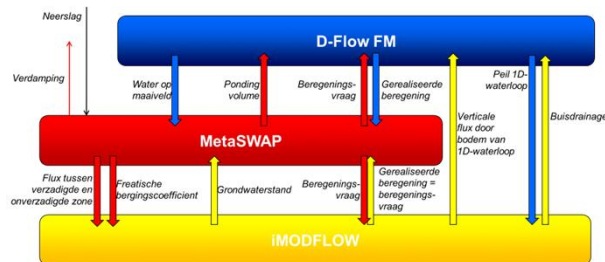
Sinds ~1990 geohydroloog:

- Bouw van grondwatermodellen
 - NAGROM; NAtionaal GRONDwater Model, o.b.v. Analytische Elementen Methode (AEM, MLAEM, MVAEM)
 - Strijd tussen AEM, Eindige Elementen Methode, Eindige Differentie Methode => MODFLOW (USGS)
 - Landelijk Hydrologisch Model LHM
 - Parameter optimalisatie (o.a. de 'Representer'-methode van Johan Valstar)
 - Regionale modelbouw o.b.v. iMOD: MIPWA, MORIA, IBRAHYM, HYDROMEDAH, AMIGO, AZURE
 - Nationaal Water Model NWM



Rekentijd is vaak nog steeds een bottleneck!

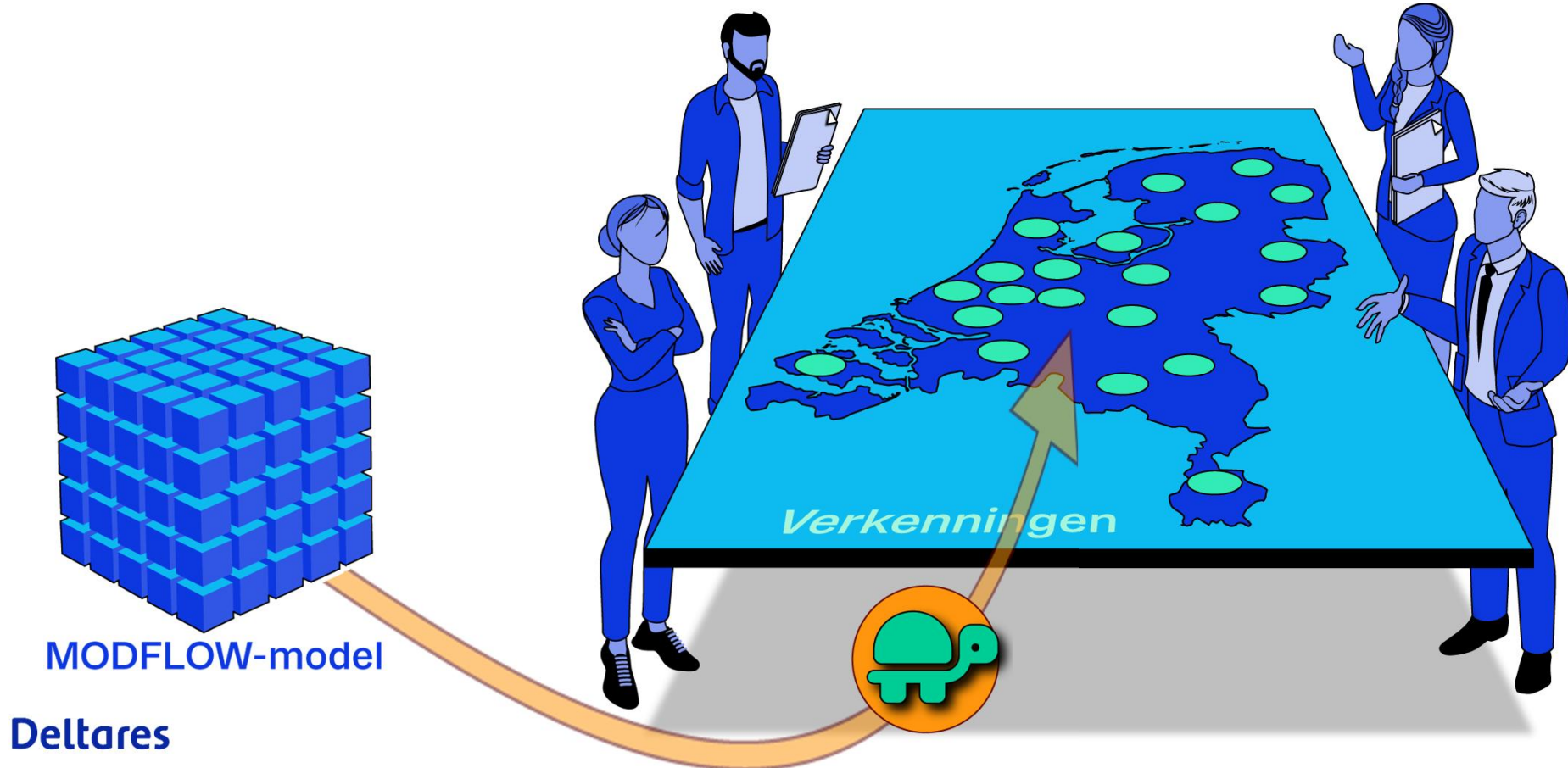
- Koppeling grond- en oppervlaktewatermodellen
D-HYDRO FM – MetaSWAP - MODFLOW



Transitieopgaven

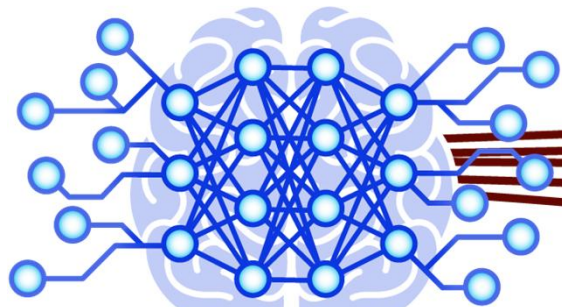


Transitieopgaven



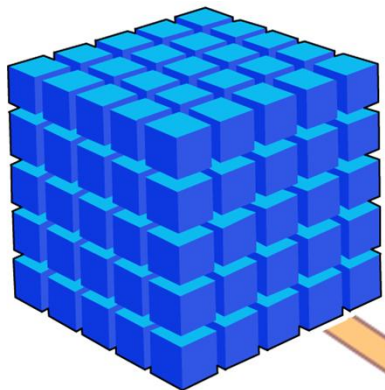
MODFLOW-model

Deltares



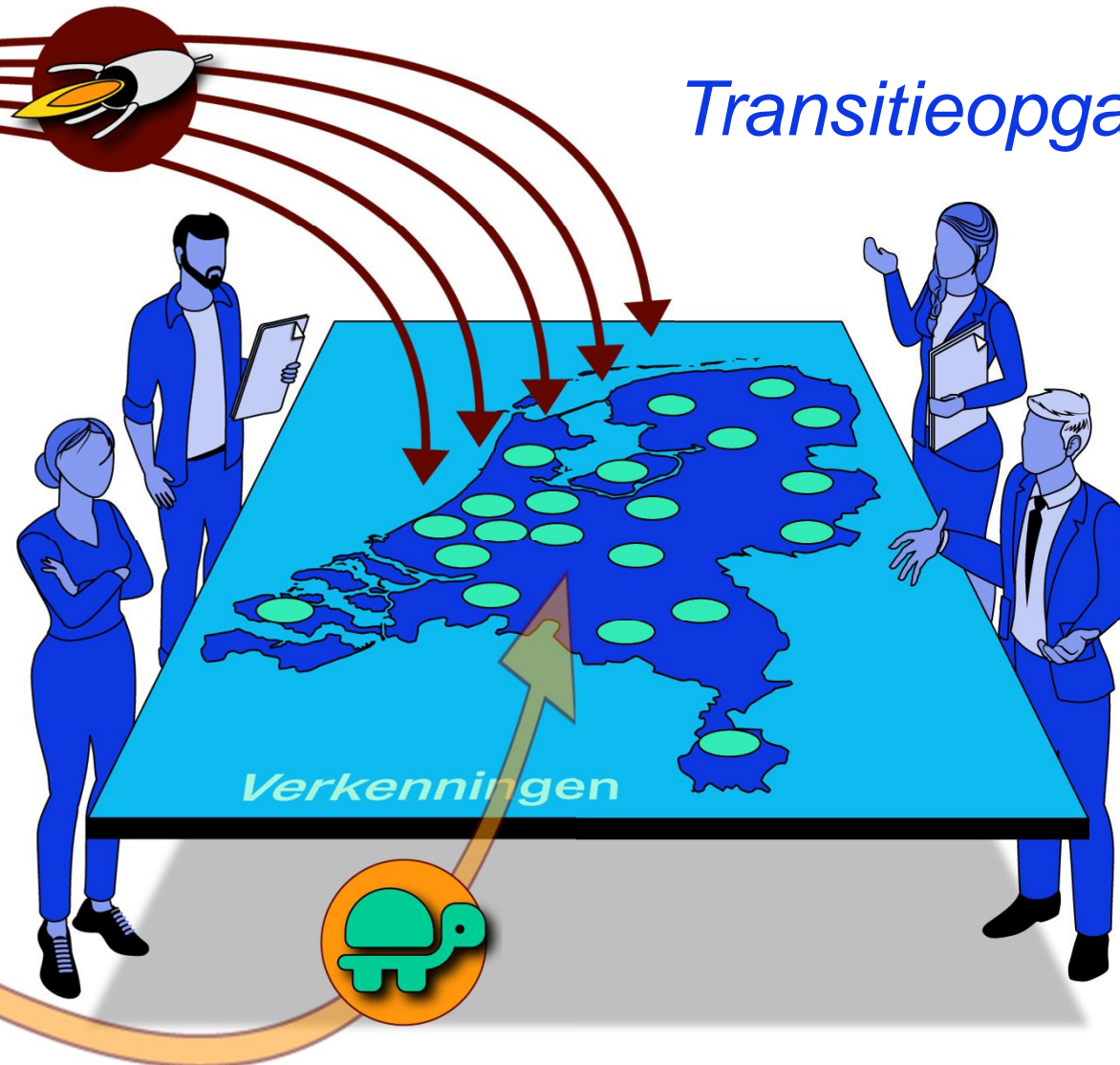
Neuraal Netwerk-versie
van MODFLOW-model

Deep learning



MODFLOW-model

Deltares



Transitieopgaven

Verkenningen

De techniekontwikkeling; hoe we gestart zijn

Bennie Minnema
Jonathan Nuttall

- Eind 2017:
Deltares Software Innovation Challenge award
(seed money: ~10 kEuro)

HOME EN / IDEA PLATFORM / CHALLENGE: SOFTWARE INNOVATION / PHYSICAL NEURAL NETWORKING IN GROUNDWATER

Physical Neural Networking in groundwater

Share

Edit

Follow this idea



Description

To develop a new type of - neural network + fuzzy logic based - groundwater modelling technique: using existing transient three-dimensional (MODFLOW-) groundwater models as input to train a neural network version of the same model including only essential physical building blocks. This new approach is already developed in the field of reservoir petroleum engineering. Real-world cases show that a single model run of these 'neuro-fuzzy' models are usually measured in seconds or in some cases minutes. The 'neuro-fuzzy' models mimic the underlying deterministic models with an accuracy of around 85%-90%, which is more than sufficient for groundwater modelling.

Benefit

The technical advantages are: - The neuro-fuzzy version of your model runs in seconds instead of hours/days; - Detailed 3D non-linear groundwater modelling in a (near) real-time operational setting, running massive amounts of scenarios within limited amount of time, do on-site real-time risk analysis during infrastructural projects, include huge numbers of variations in boundary conditions you never dared to dream about before... - Training the neural-fuzzy version of your groundwater model in a machine learning environment basically combines brute-force parameter optimization with cross-validation methods. It results in models that really predict optimally and the accuracy of the models can be quantified - The approach potentially supersedes 'traditional' Proper Orthogonal Decomposition (POD), Impulse Response- and 'proxy'-methods because it includes non-linear behaviour by including physics-based building blocks (the 'fuzzy' part of the approach). - It allows implementing a chosen scenario in the underlying deterministic groundwater model to fine-tune planned infrastructural measures in groundwater management.

Impact

The general impacts are: - It generates the perspective of real-time 3D groundwater modelling on 'domestic' portable devices like smart phones and tablets => easy access to full-fledged 3D groundwater models for everybody. - Enormous possibilities for local on-site risk-analysis in an infrastructural environment, decreasing total costs - It sustains the ongoing investments in building detailed deterministic numerical 3D groundwater models, since these models are the necessary input the neural-fuzzy versions. - In general, it strengthens the position of groundwater modelling as part of water management. The potential impacts for Deltares: - Developing this new potentially game changing modelling technique can illustrate the innovative capabilities of Deltares in the field of applied research. - A new extreme fast modelling technique allows us to initiate projects aiming to make

About this idea

Published by:



Bennie Minnema

Published On:

18 Sep 2017

Challenge:

Challenge: Software Innovation

Theme:

Software Innovation

Progress



Interesting Link(s)

Following this idea (3)

De techniekontwikkeling; hoe we gestart zijn

Bennie Minnema
Jonathan Nuttall

- Eind 2017:
Deltares Software Innovation Challenge award
(seed money: ~10 kEuro)
- 2018-2019:
 - Connectie met DSC en Rijkswaterstaat
 - Zoektocht naar funding: RWS ~50 + 40 + SO
 - Samenwerking met Olie & Gas-industrie:
😊 😐 😞: nog een erg gesloten wereld.
 - Gestart met eenvoudige grondwater modellen
 - Brug slaan tussen geohydrologen en wiskundigen



De techniekontwikkeling; hoe we gestart zijn

Bennie Minnema
Jonathan Nuttall

- Eind 2017:
Deltares Software Innovation Challenge award
(seed money: ~10 kEuro)
- 2018-2019:
 - Connectie met DSC en Rijkswaterstaat
 - Zoektocht naar funding: RWS ~50 + 40 + SO
 - Samenwerking met Olie & Gas-industrie:
😊 ☹️ 😞: nog een erg gesloten wereld.
 - Gestart met eenvoudige grondwater modellen
 - Brug slaan tussen geohydrologen en wiskundigen
- Medio 2019: eerste resultaten gepresenteerd
op MODFLOW and more conference



From: Lucky Luke, De Spookstad, René Goscinny, Belgium

De techniekontwikkeling; hoe we gestart zijn

Bennie Minnema
Jonathan Nuttall

- Eind 2017:
Deltares Software Innovation Challenge award
(seed money: ~10 kEuro)
- 2018-2019:
 - Connectie met DSC en Rijkswaterstaat
 - Zoektocht naar funding: RWS ~50 + 40 + SO
 - Samenwerking met Olie & Gas-industrie:
😊 😐 😞: nog een erg gesloten wereld.
 - Gestart met eenvoudige grondwater modellen
 - Brug slaan tussen geohydrologen en wiskundigen
- Medio 2019: eerste resultaten gepresenteerd
op MODFLOW and more conference
- 2020-2024: promotieonderzoek van Maria Luisa Taccari





Journal of Hydrology
Available online 12 December 2023, 130551
In Press, Journal Pre-proof [What's this?](#)



Research papers

Developing a cost-effective emulator for groundwater flow modeling using deep neural operators

[Maria Luisa Taccari](#)^a  , [He Wang](#)^b, [Somdatta Goswami](#)^c, [Mario De Florio](#)^c,
[Jonathan Nuttall](#)^d, [Xiaohui Chen](#)^a, [Peter K. Jimack](#)^b

TKI-project: *AI-kennis voor grondwatermodellen*



Doel:

- Een neurale netwerk-versie maken van bestaande MODFLOW-modellen

zodanig dat:

- Het consistent is met het onderliggende MODFLOW-model waarbij:
- getrainde neurale netwerk 100-1000 keer zo snel is als het MODFLOW-model

=>

- Ideaal voor quick-scan doeleinden
- Ideaal voor parameter optimalisaties en Monte Carlo simulaties
- 'iPAD'-/internet-versies van grondwatermodellen

Info: bennie.minnema@deltares.nl en [TKI-site](#)

TKI-project *AI-kennis voor grondwaterverkenningen*

1^e jaar:

- Testen verschillende typen neurale netwerken (UNet, GAN, CNN, FNO, DeepONet, LSTM) voor kleine(re) (hypothetische) testmodellen.

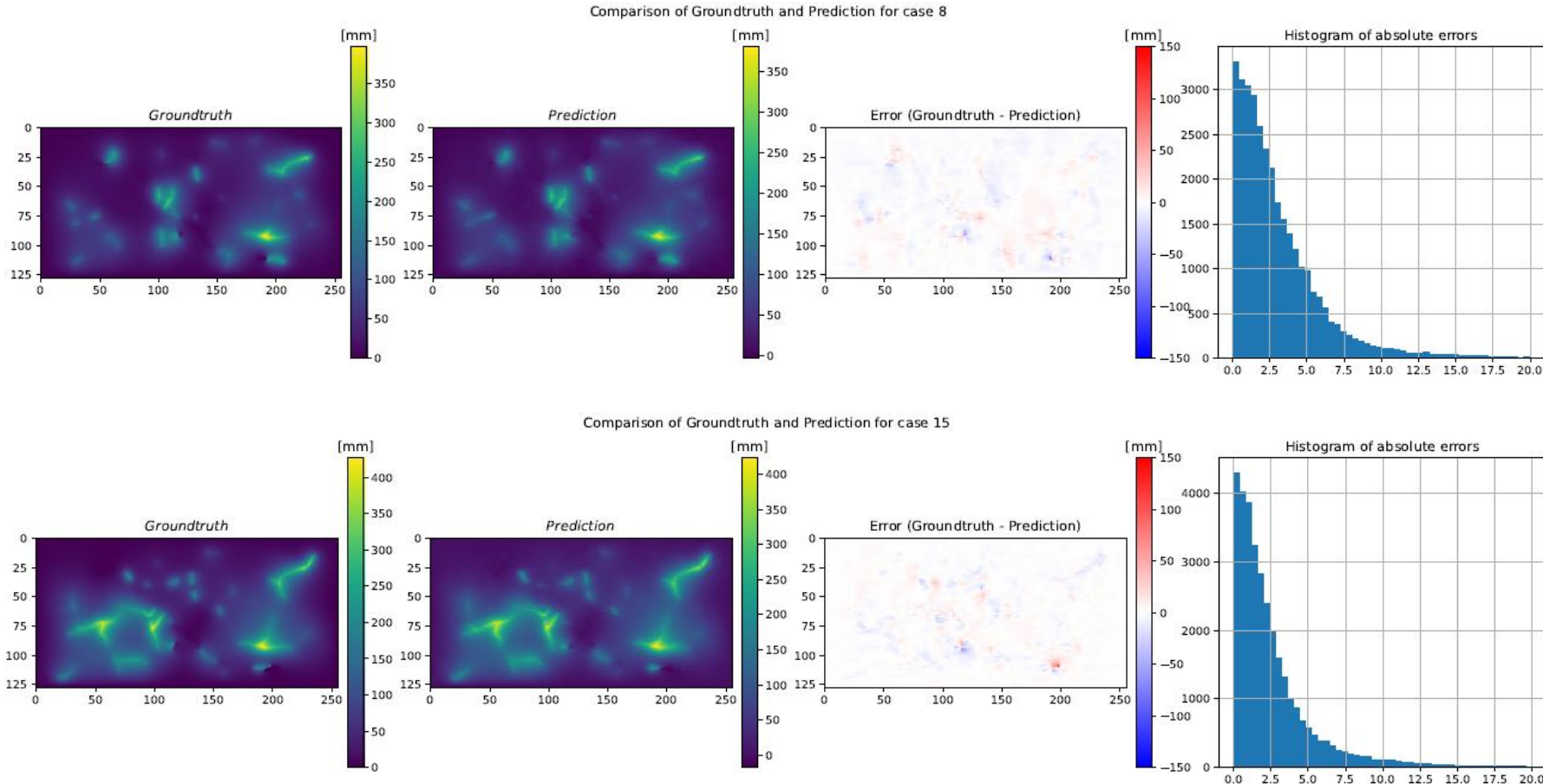
2^e jaar:

- Opschalen naar 'real-world' casussen

Meer info:

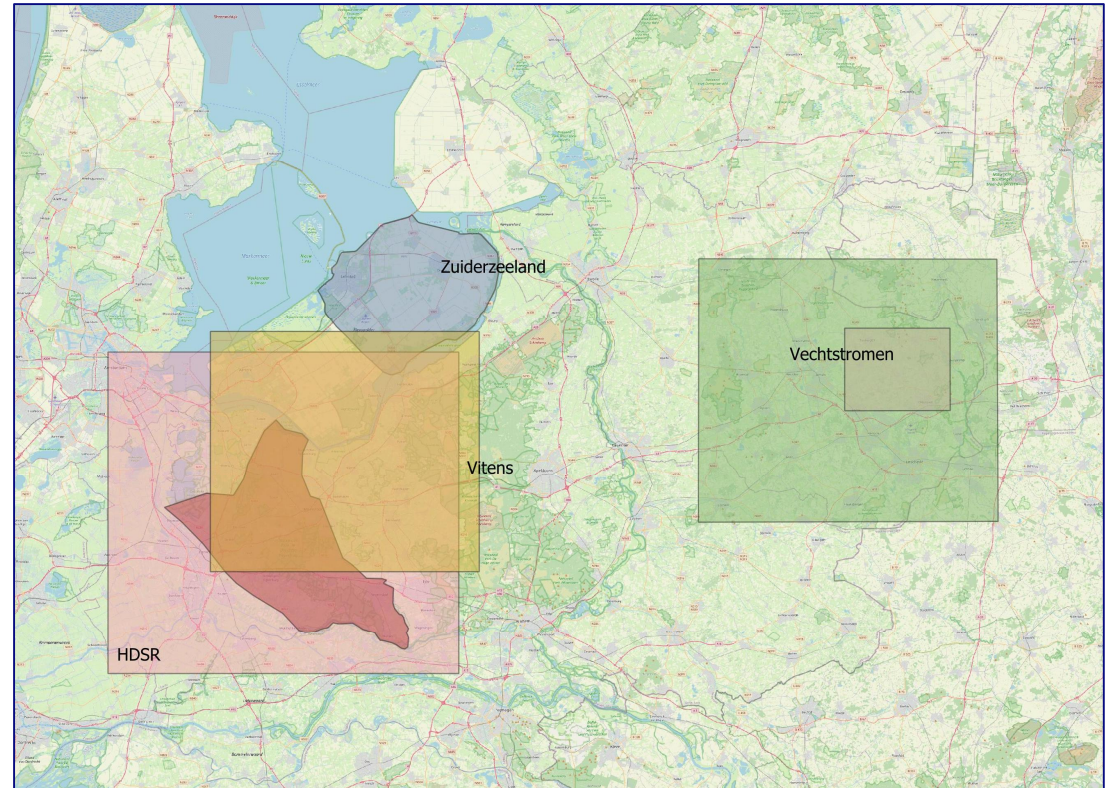
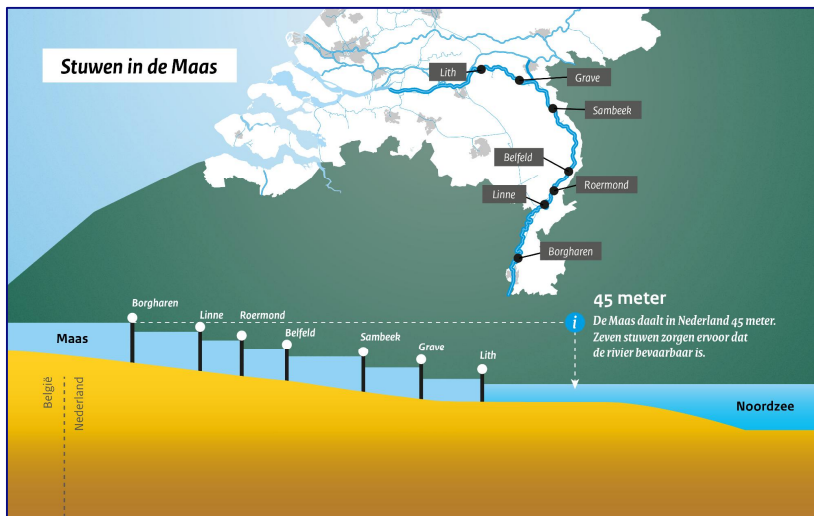
<https://tkideltatechnologie.invoermodule.nl/project/ai-kennis-voor-grondwaterverkenningen/>

Voorbeeld: Fourier Neural Operator (FNO); effecten op GLG voor testmodel met omvang van 256 kolommen bij 128 rijen



TKI-project *AI-kennis voor grondwaterverkenningen*

'Real-world' casussen:



TKI-project *AI-kennis voor grondwaterverkenningen*

- Einddoel:
dat we in staat zijn razendsnelle emulators te maken van bestaande MODFLOW-modellen.
- Maar ook:
generieke aspecten kunnen hergebruiken voor belendende domeinen, bijvoorbeeld bij het simuleren van riviermorfologie.

(nabije) toekomstplannen

1. Enabling Technologies team, Deltares

- Maakt op dit moment een internationale inventarisatie 'AI for water management' (in opdracht van UNESCO-IHP), incl. update van het overzicht wie binnen Deltares met welke AI-technieken bezig zijn => kennisuitwisseling.
- voorstel: deze inventarisatie en update uitbreiden binnen NHI-kader.

2. AI-grondwater symposium dag, in week 47 (18-22 november 2024, onderdeel van Deltares Software Days).

Oproep:

- Alvast na te denken over een AI en grondwater gerelateerde presentaties.
- Alvast nadenken over het opzetten van AI en (grond)water-gerelateerde vervolgprojecten.

Dank u voor uw aandacht!

Voor tips & feedback, email naar Bennie.Minnema@deltares.nl