

Building large-scale 3D coastal groundwater models with iMOD-WQ and global datasets

Daniel Zamrsky¹, Gualbert H.P. Oude Essink^{2,1},
 Jude King², Wahdan Achmad Syaehuddin³,
 Mitchel Kwint⁴, Marc F.P. Bierkens^{1,2}



Utrecht University



Deltires



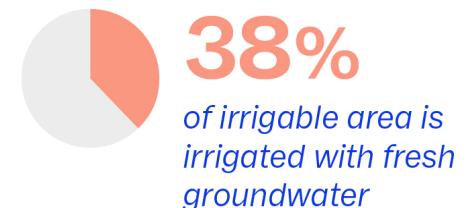
WAGENINGEN
UNIVERSITY & RESEARCH



THE HAGUE
UNIVERSITY OF
APPLIED SCIENCES

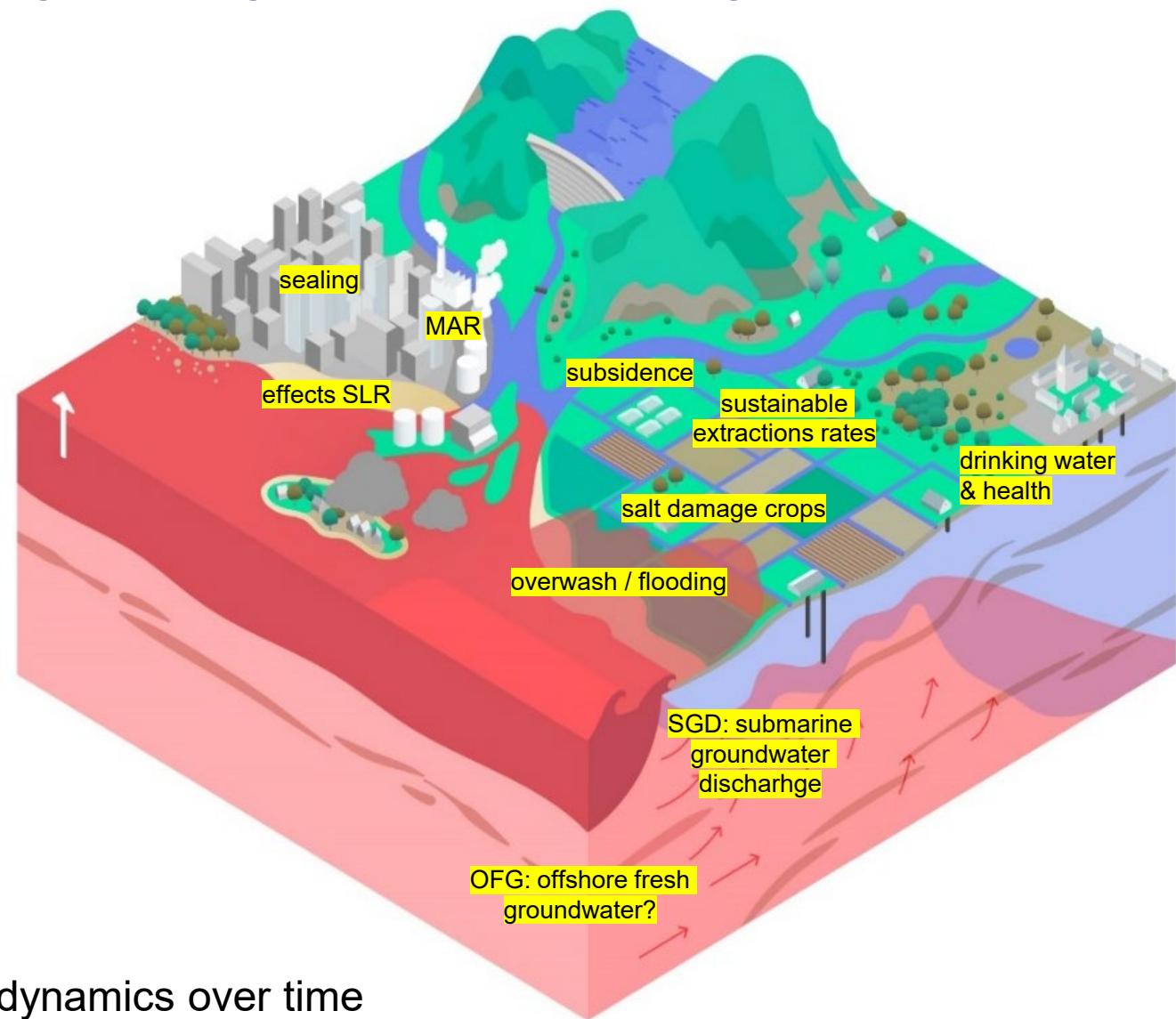
Why this coastal groundwater model initiative?

- Coastal fresh groundwater is **main water resource** for ~50% of the world population in the coastal zone
- Groundwater is important for **drinking water, agriculture, industry**, as well as **ecosystems & river baseflows**
- **Right now**, fresh groundwater resources are threatened by **excessive pumping**, and **reduced replenishments**
- Projected climate change impacts, inducing **sea-level rise**, will worsen this situation
- We need quantified **storylines** on fresh groundwater availability under stress in **data-poor** coastal zones
- These storylines should be **linked to droughts, land subsidence, flooding, (human) health and biodiversity**



Applications, components and insights large-scale coastal groundwater model

- Components:
 - groundwater quantity
 - groundwater salinity
 - subsidence (into 2024)
 - heat transport (later, >2024)
 - groundwater quality (later, >>2024)



- Insights:
 - into (supra) - regional coastal groundwater dynamics over time
 - into understanding current and future state of transboundary fresh groundwater resources
 - into identifying potential hotspots for fresh groundwater shortages

Characteristics large-scale coastal groundwater models LCGMs(1km² scale)

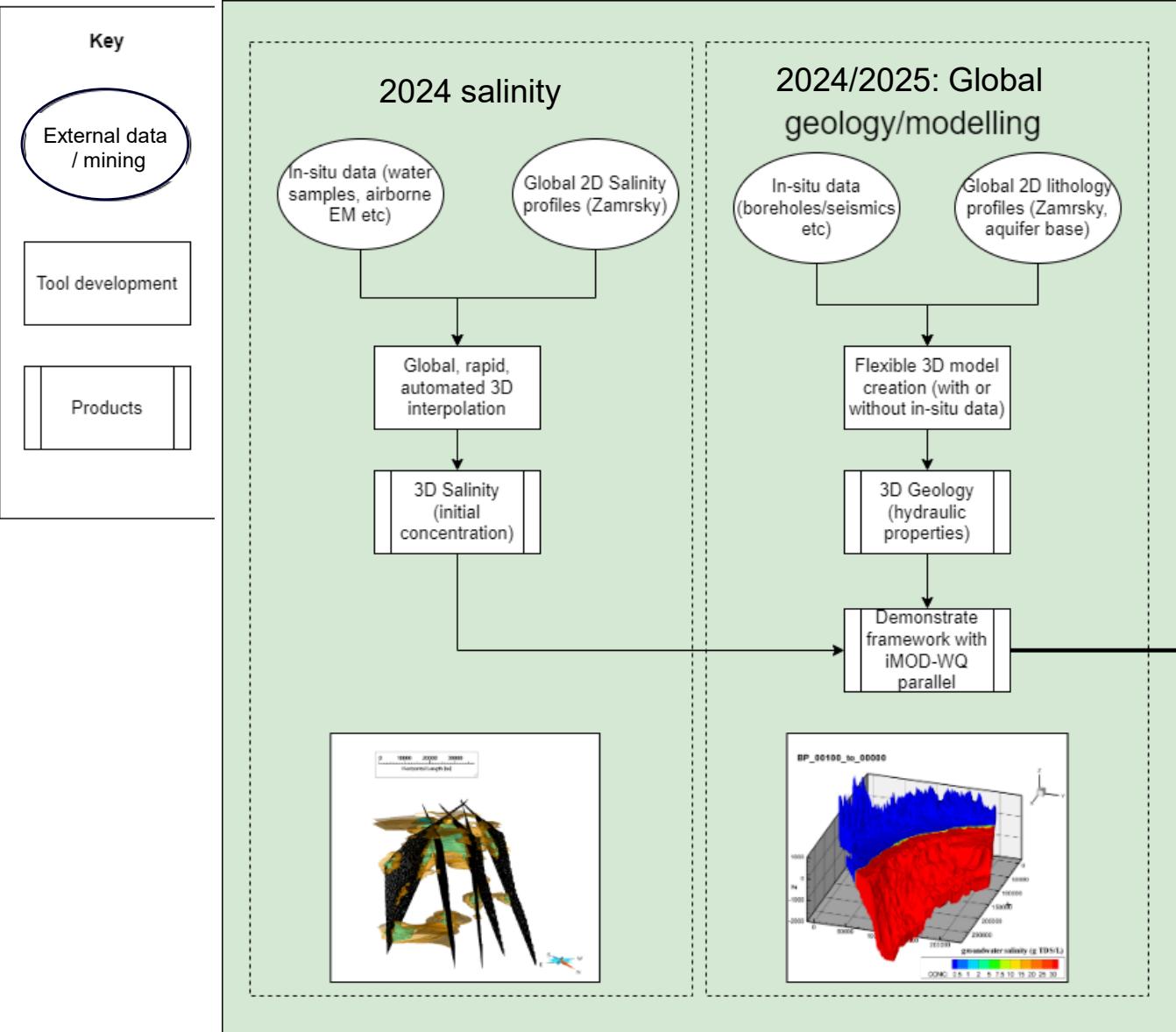
1. Models can cover areas >> **10.000 km²**
2. Typically **cell sizes** of **1*1km²**
3. **Multiple model layers** (to properly represent groundwater salinity and subsidence)
4. Our LCGM building tool use **open-source tools like Python**
5. **Global datasets** are used, providing 1st-order approximations of groundwater conditions
in data scarce regions
6. **High-performance computing** opens up new possibilities
7. **Parallelization of SEAWAT** (iMOD-WQ): important breakthrough in speeding up variable-density groundwater flow and salt transport modeling
8. **Simulation** groundwater salinity dynamics over **full glacial-interglacial cycle** (e.g. 125 ka).

Characteristics large-scale coastal groundwater models LCGMs(1km² scale)

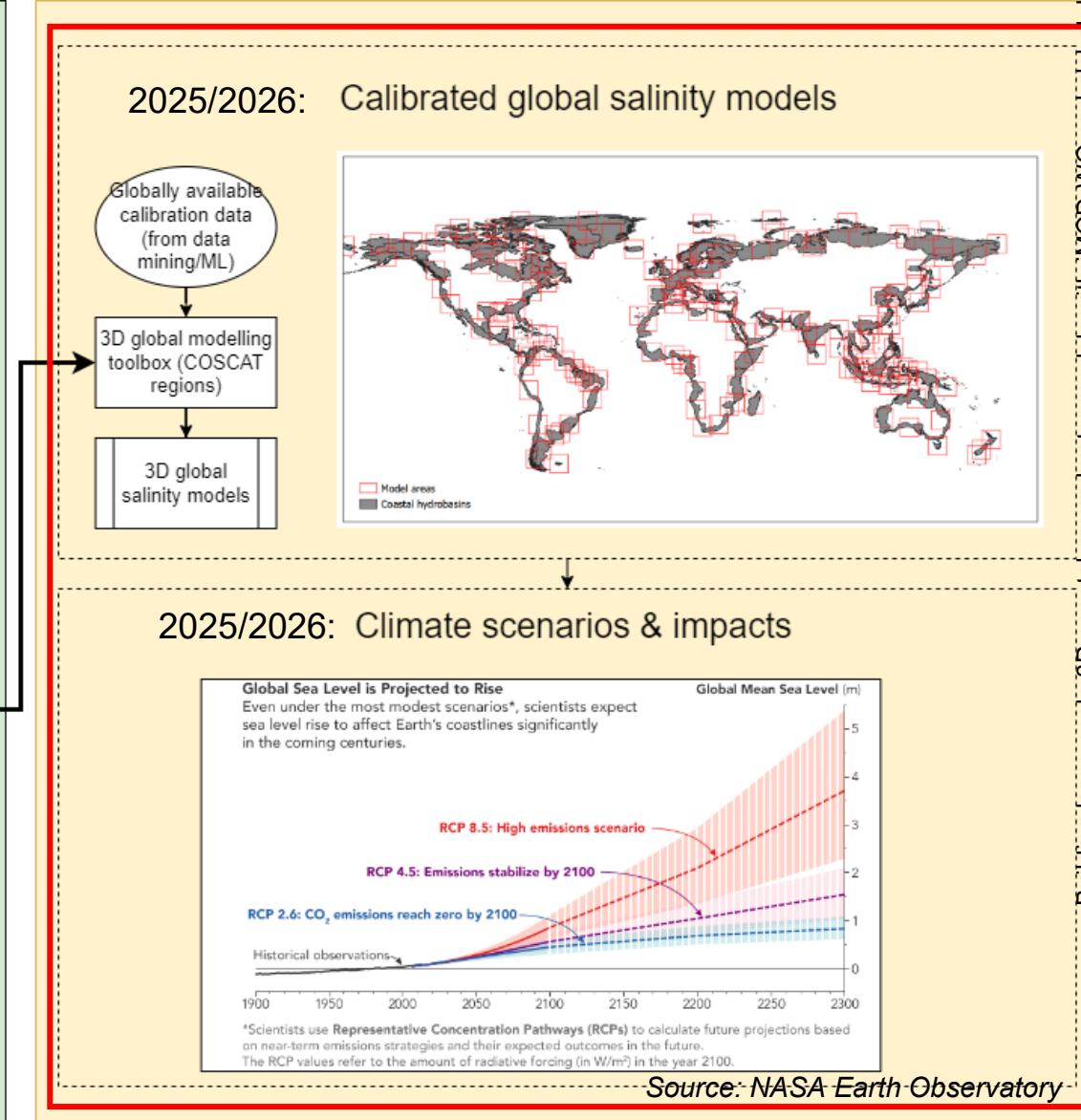
9. **HydroBASINS global-watershed-boundaries** dataset are used to delineate the boundaries
10. **Offshore continental shelves** are also covered (manually outlined and added to the selected HydroBASINS).
11. **Top elevation** derived from a **global DEM dataset (GEBCO)**
12. **Bottom elevation** estimated by:
 - a. the bottom of the unconsolidated sediment formations
 - b. sedimentary rock formations (limited to siliciclastic lithology)
13. When **local hydrogeological input data** is available (e.g. borelogs, groundwater salinity, extractions), tools like **GEMPY** are used to **improve the LCGMs**

Planning large-scale coastal groundwater model building tool

Phase 1: Tool development, demonstration

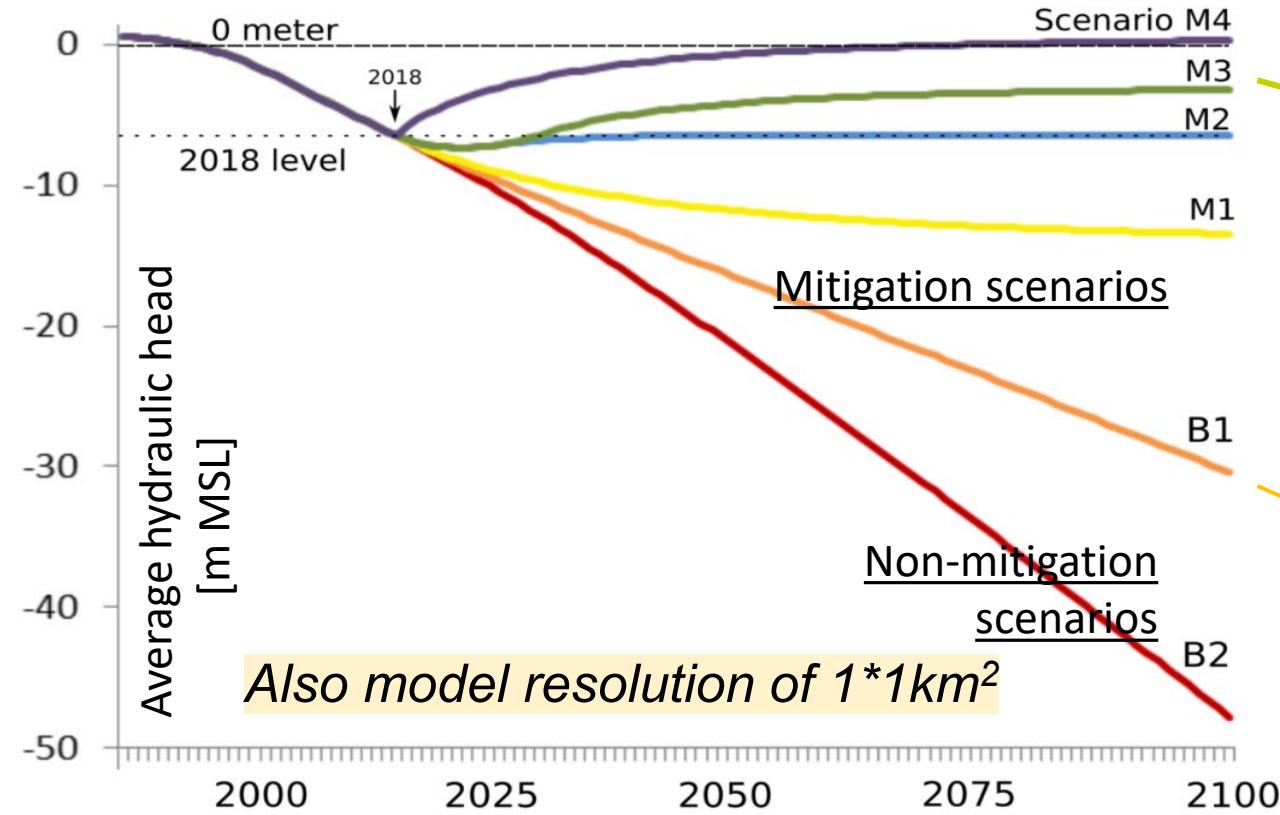


Phase 2: Global modelling and scenarios

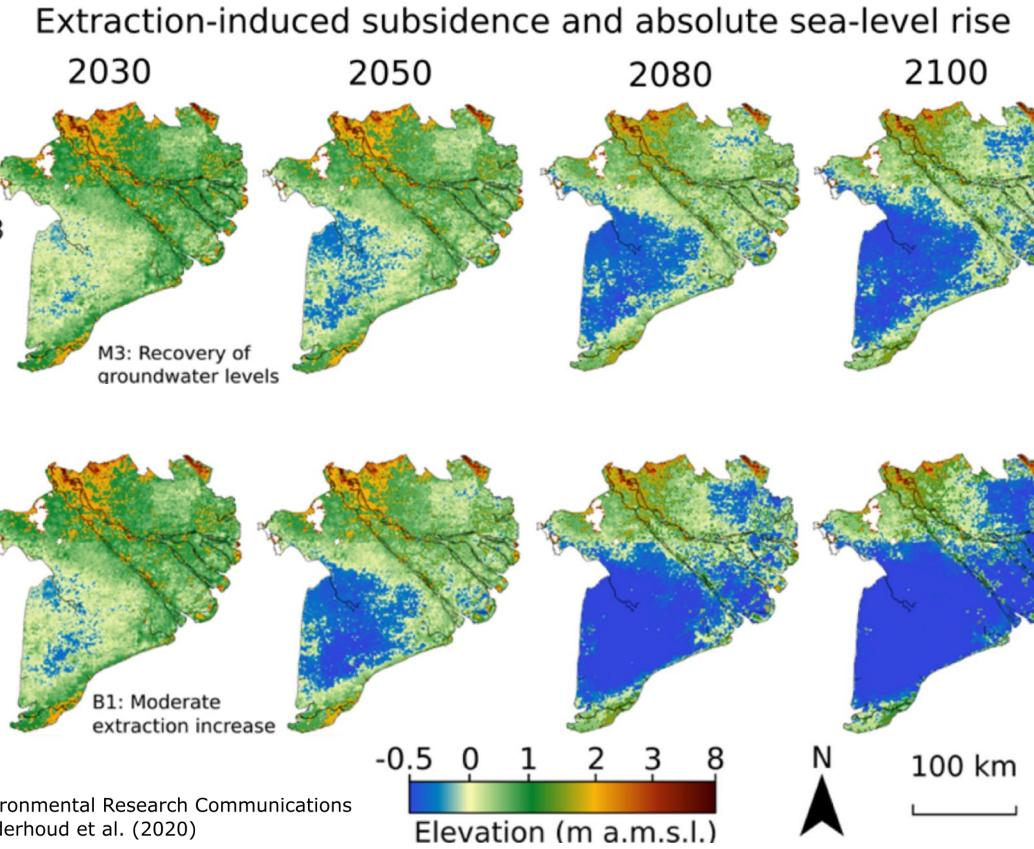


Example storyline: Pathways to demonstrate the future Mekong delta: linking groundwater extraction → subsidence → increased flood risk

Scenarios of future groundwater extraction pathways



Future elevation of the delta



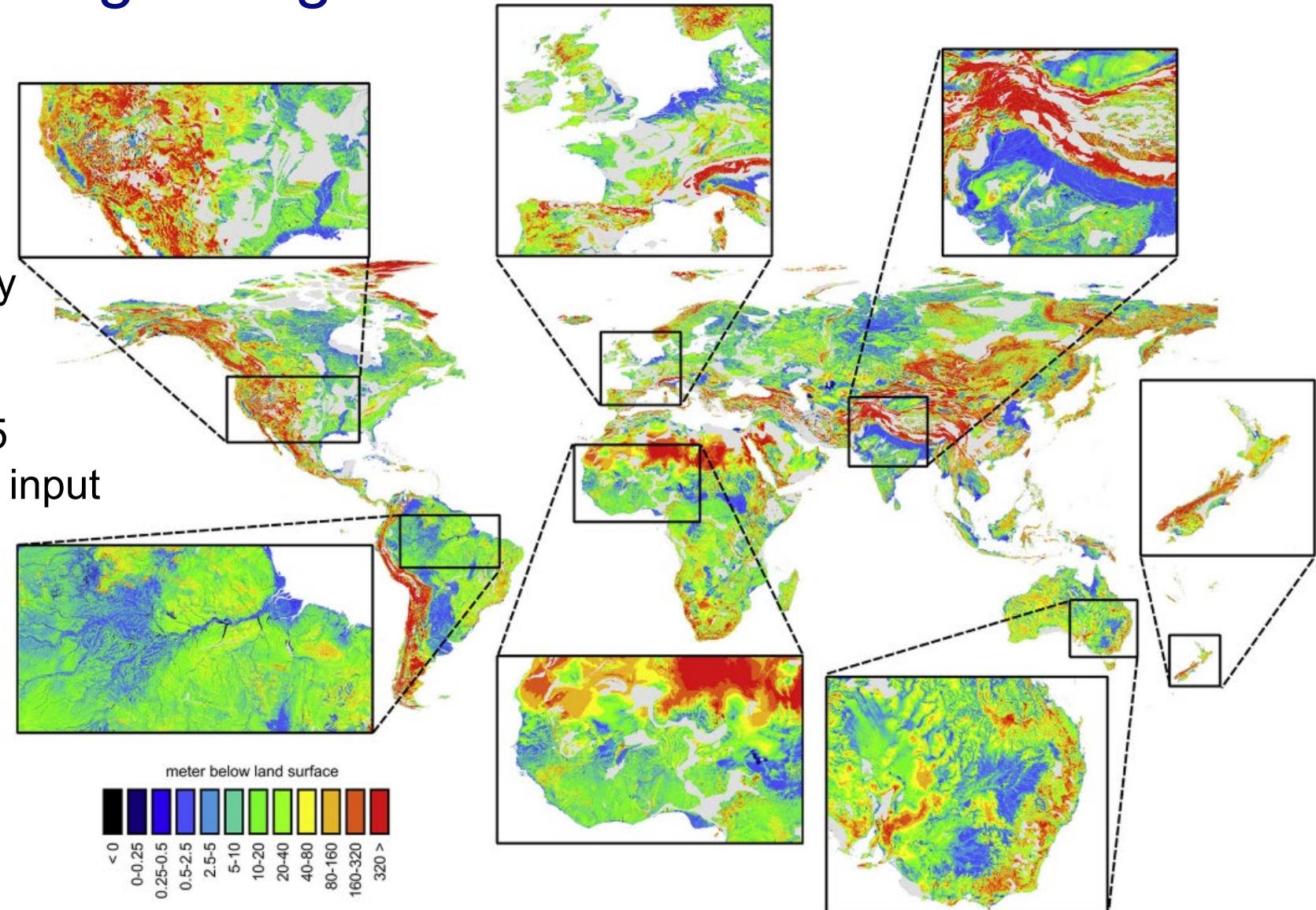
Environmental Research Communications
Minderhoud et al. (2020)

Among others based on this research, 'Decree 167' has been implemented in Vietnam in the **Law of Water Resources**: develop and implement zoning plans to **restrict groundwater overexploitation**

Example 1: 1*1km² global groundwater model

Components:

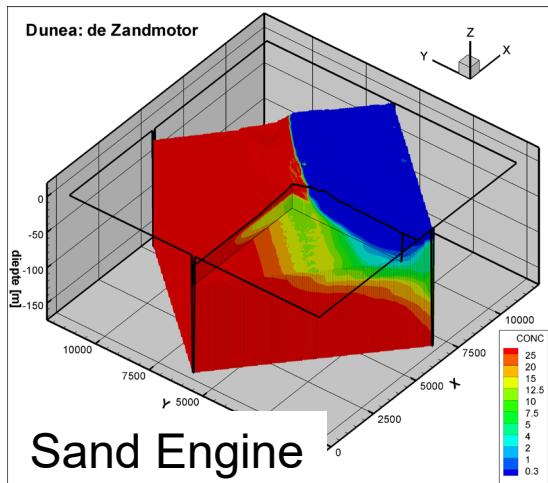
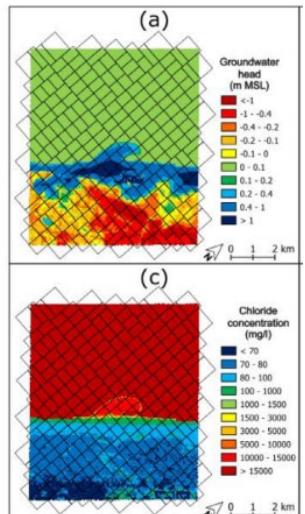
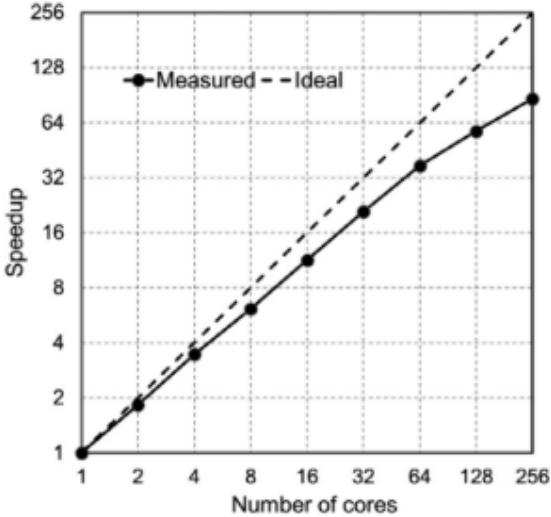
- quantitative groundwater only
- 278 million active cells
- two model layers
- simulating period 1958–2015
- daily time steps and monthly input
- 12 nodes, 384 cores
- Snellius supercomputer
- maximum 16 hr simu time!



Building large-scale 3D coastal groundwater models with iMOD-WQ and global datasets

Example 2: parallel computing plus smart model parameters

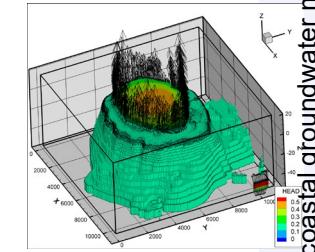
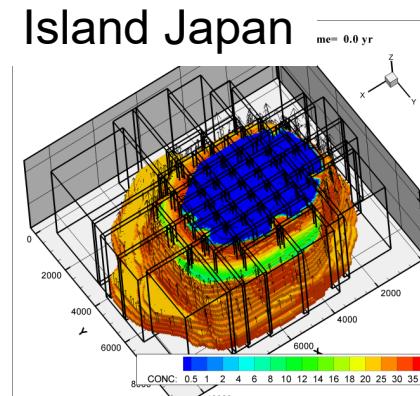
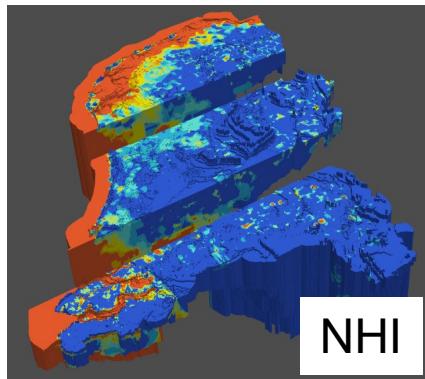
- Split into (tens of) partitions, leading to a significant reduction in computation time
- Speed-ups of at least 10 up to 100 times, depending on cores, solver iterations and data exchange efficiencies



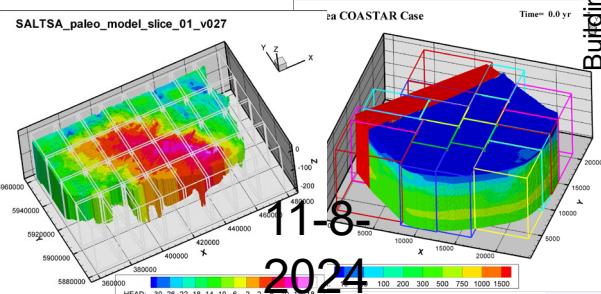
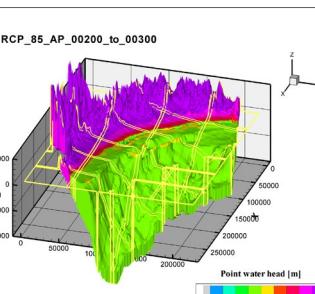
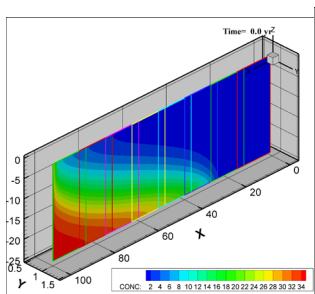
Sand Engine

Three examples:

1. Sand Engine: from 1hr 47min 55sec \rightarrow 2min 40sec: **40***
2. NHI fresh-salt: from \sim 30 days to \sim 2days: **15***
3. Island Japan: from 5d0h36m to 5m59s: **1209***



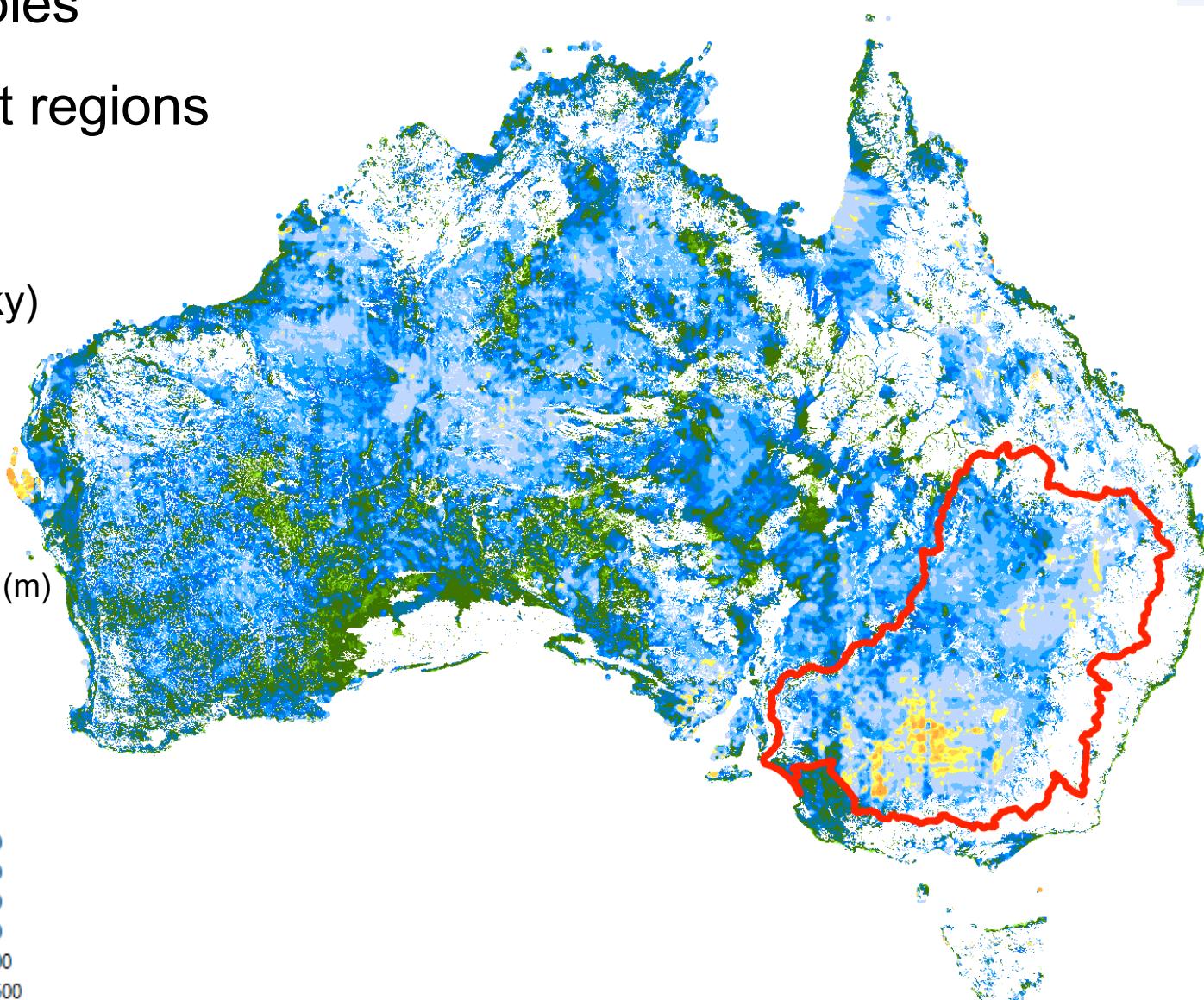
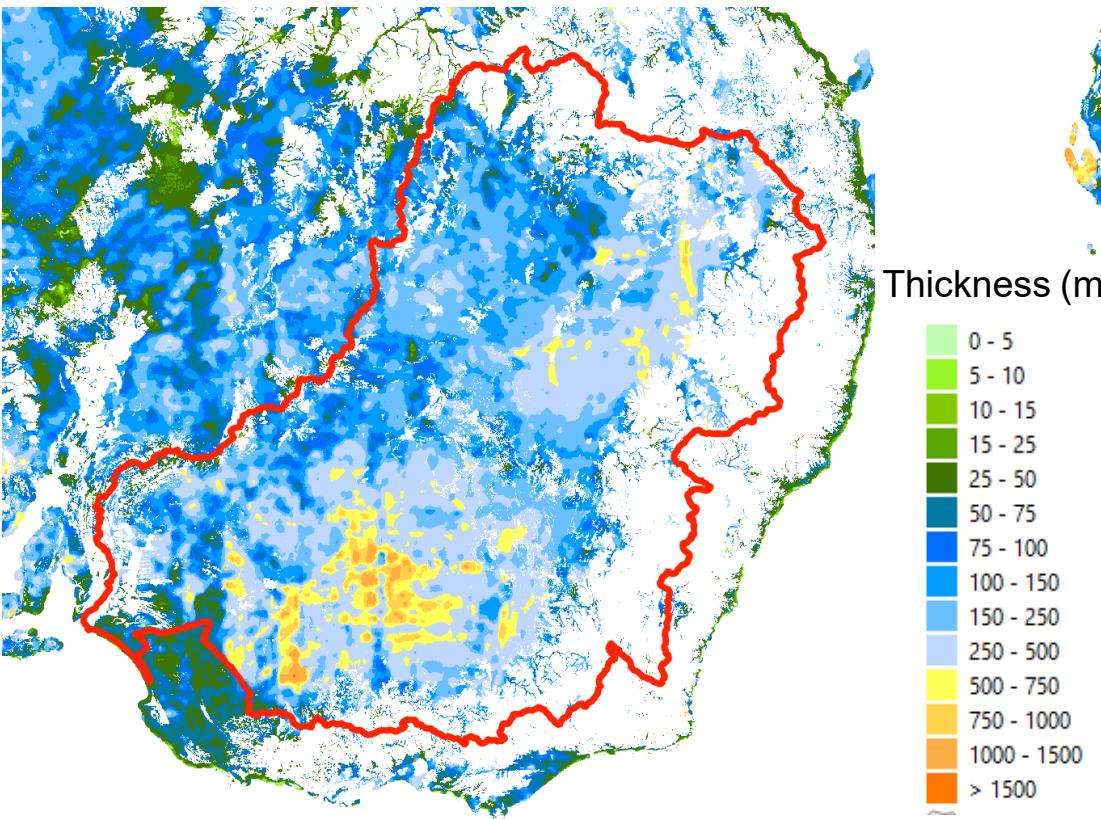
Verkaik, J. et al., 2021. Adv. Water Resources



Example 3: Improving geology, focus sediment thickness

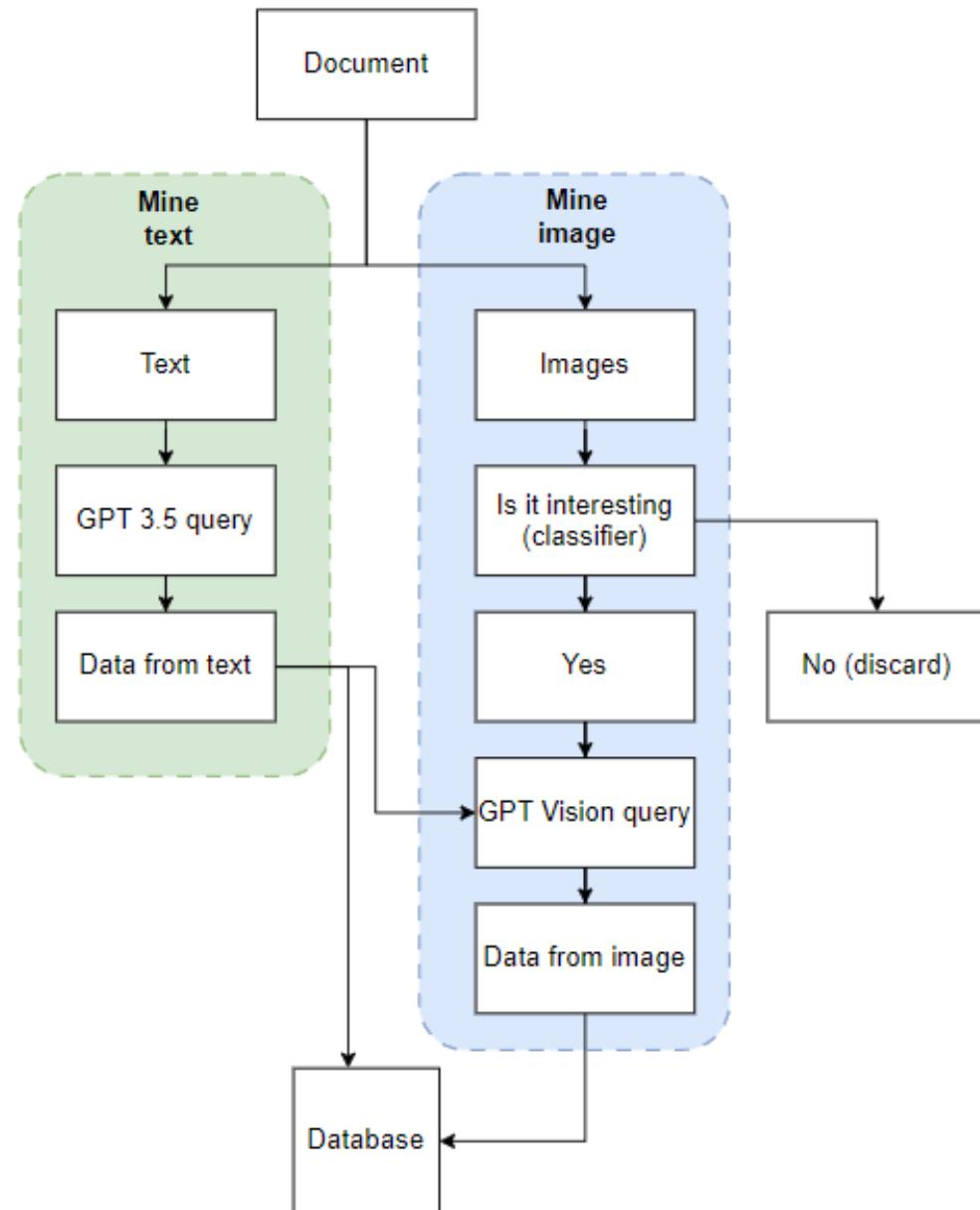
- we validated with ~40.000 boreholes
- decent results in deeper sediment regions
- example Australia

(for more information: check with Daniel Zamrsky)



Example 4: Data mining hydrogeology

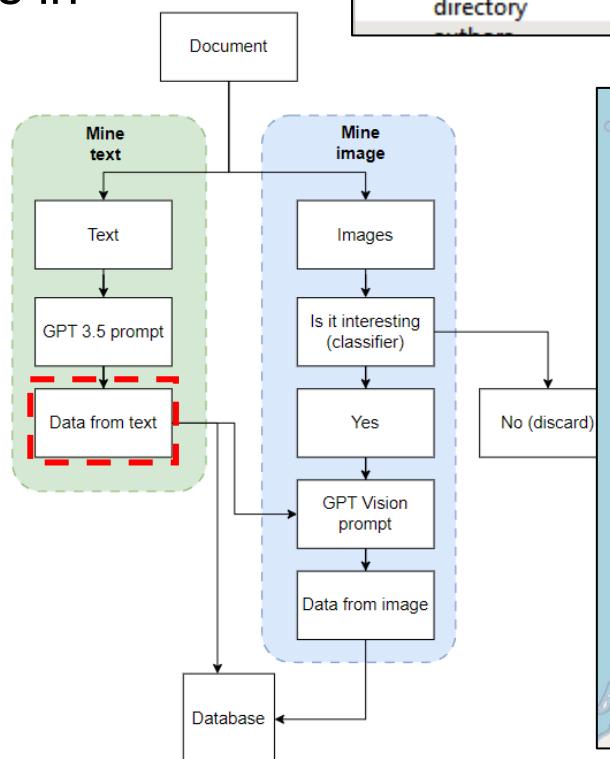
- Global models need data for **hydrogeological properties, validation, calibration**, etc.
- **Without enough data**, models and their **projections are highly uncertain** / people do not trust them.
- Recently available Python APIs for Large Language Models (LLMs), e.g., Open AI GPT3.5/GPT4 Vision).
- Architecture designed with FAIR practices in mind
- Short scripts, adaptable to different LLMs
- Exciting opportunities to apply these methods



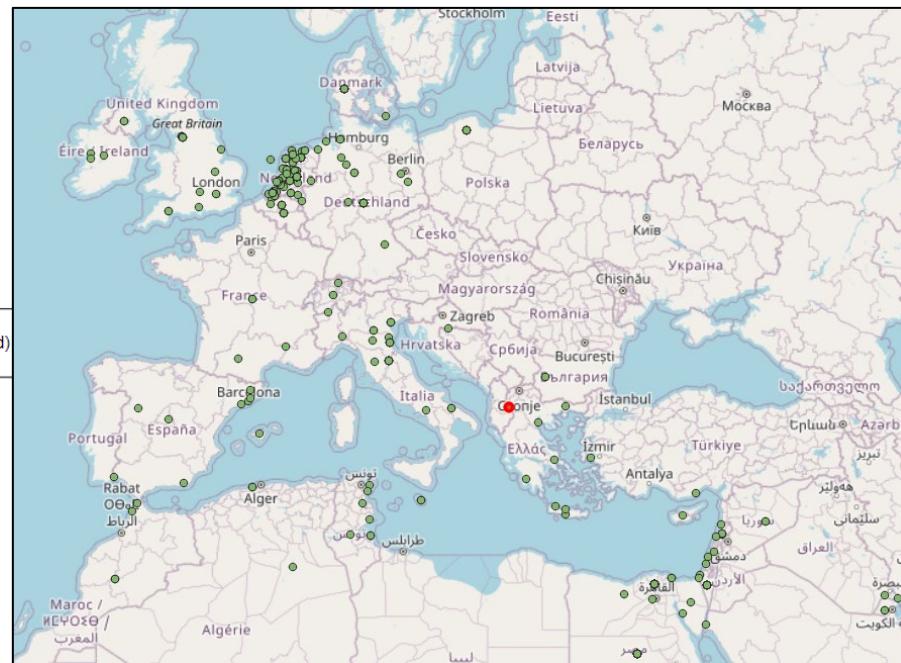
Example 4: Data mining hydrogeology

Structured data from text

- LLM response is processed using a parser and added to a structured database
- Bing maps API turns place names in coordinates
- Here: result from mining ~2000 documents
- Five orders faster & cheaper



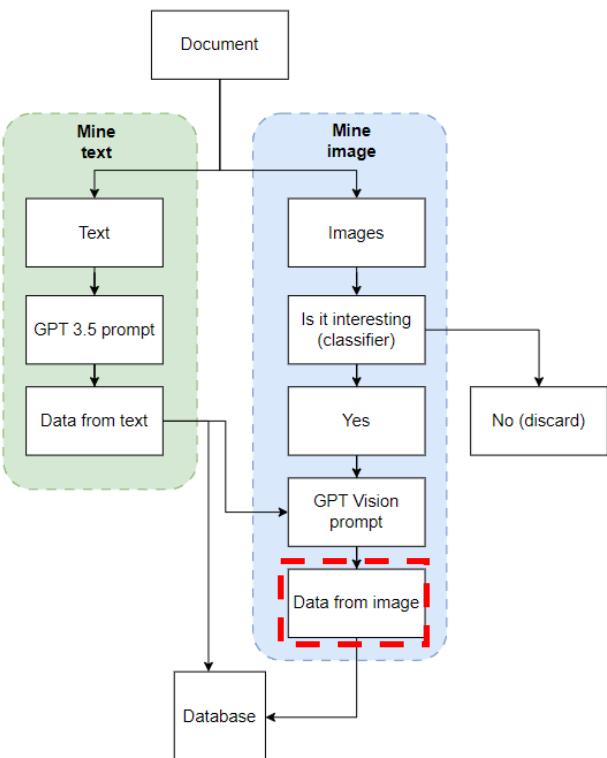
Data_Mine_V1	
title	Borehole logging and seismic data from Lake Ohrid (North...
field_1	1
(Derived)	Borehole logging and seismic data from Lake Ohrid (North...
(Actions)	
title	Borehole logging and seismic data from Lake Ohrid (North...
year	2022
country	North Macedonia/Albania
region	Lake Ohrid
url	https://doi.org/10.1016/j.quascirev.2021.107295
keywords	'age-depth modelling', 'downhole methods', 'seismic inte...
data_links	[]
Y_country	41,13994598000000
X_country	20,06507682999999
Y_region	41,06011580999999
X_region	20,73127556000000
directory	C:\Users\king_je\Projects\Text_Mining\pdfs\1-s2.0-S02773...
...	NULL



Example 4: Data mining hydrogeology

Structured data from images

- Uses multimodal capabilities (**GPT4 Vision**)
- Quantitative data needs careful structuring
- It can format and also reject incomplete or poorly structured data



Example parsed JSON output from LLM for one borehole interval

```
{  
    "top_m": 0.0,  
    "bot_m": 3.0,  
    "litho_main": "TOPSOIL",  
    "litho_description": "TOPSOIL GRAVEL, rare boulders",  
    "x_coordinate": 1661007,  
    "y_coordinate": 5315067,  
    "ground_level": 27.0,  
    "grid_reference": "NZTM"  
},
```

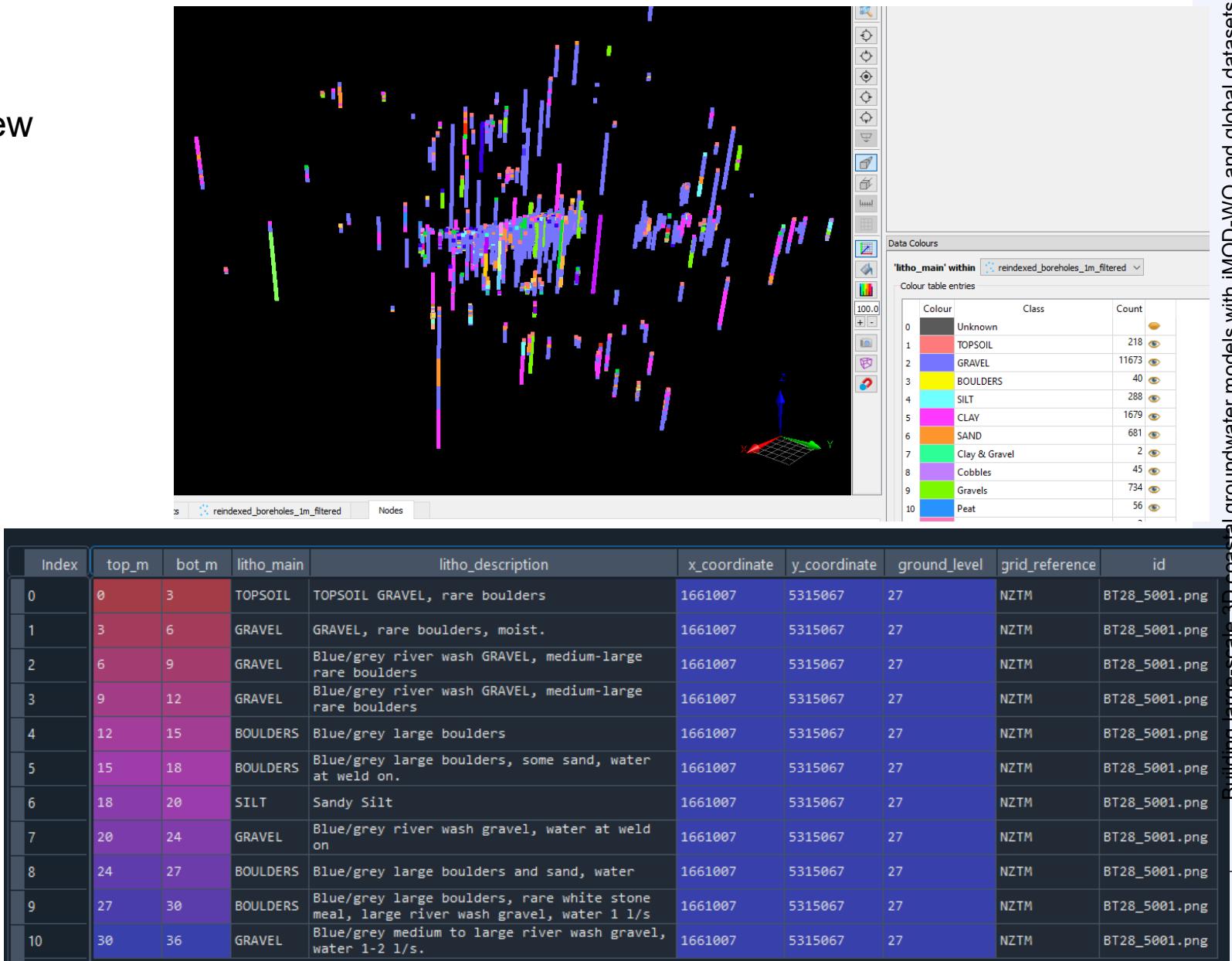
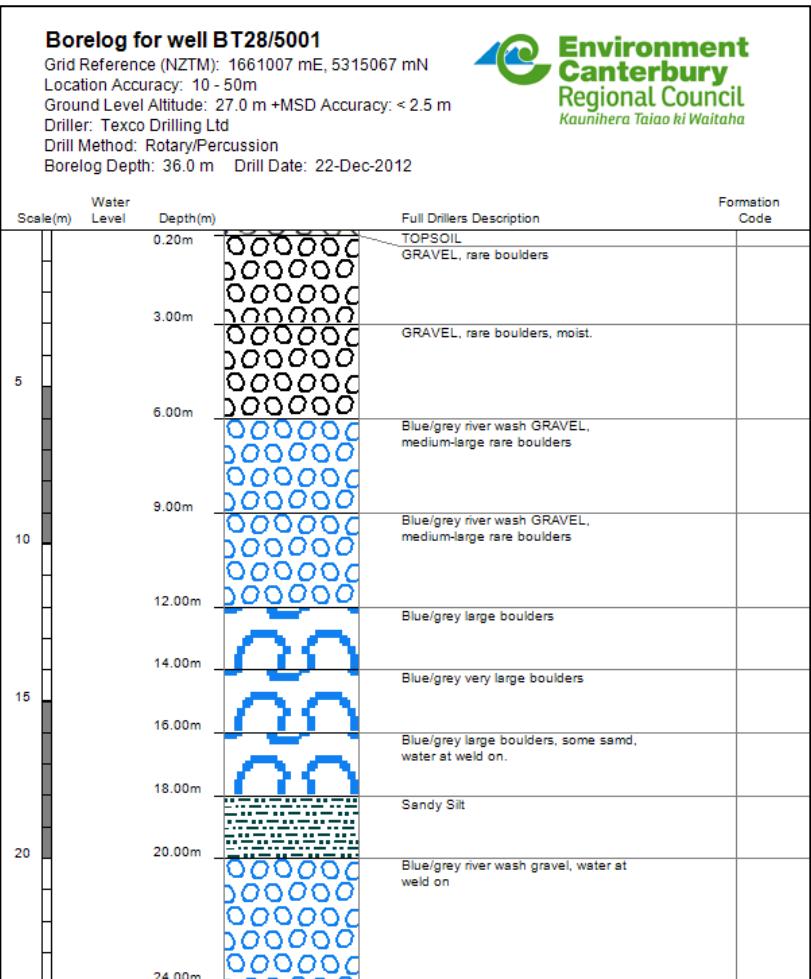
Parsed output added to database

df	top_m	bot_m	litho_main	...	ground_level	grid_reference	id
0	0.0	3.0	TOPSOIL	...	27.0	NZTM	BT28_5001.png
1	3.0	6.0	GRAVEL	...	27.0	NZTM	BT28_5001.png
2	6.0	9.0	GRAVEL	...	27.0	NZTM	BT28_5001.png
3	9.0	12.0	GRAVEL	...	27.0	NZTM	BT28_5001.png
4	12.0	15.0	BOULDERS	...	27.0	NZTM	BT28_5001.png
5	15.0	18.0	BOULDERS	...	27.0	NZTM	BT28_5001.png
6	18.0	20.0	SILT	...	27.0	NZTM	BT28_5001.png
7	20.0	24.0	GRAVEL	...	27.0	NZTM	BT28_5001.png
8	24.0	27.0	BOULDERS	...	27.0	NZTM	BT28_5001.png
9	27.0	30.0	BOULDERS	...	27.0	NZTM	BT28_5001.png
10	30.0	36.0	GRAVEL	...	27.0	NZTM	BT28_5001.png

Example 4: Data mining hydrogeology

Structured data from images

Tested on ~500 borehole images from New Zealand, provided by Utrecht University



Example 4: Data mining hydrogeology

Structured data from images

Handles handwritten logs (mostly), also in other languages

FORMATION LOG SHEET			
Date drilled:	26-5-195 -	Hole number:	T100 C = W.W.3503.
Date logged:	26-5-195 - 21-5-195	Project:	CAPM
Geologist:	EDE	Location:	3.45 m off 20° 44' 2" S / 20° 27' E
Latitude:	70° 26.665' S	Longitude:	20° 44.517' E
From	To	Rock Type	Description
0	2 m	Soil / overburden	
2	11 m	Grey, dark white sand	coarse - fine (silty) sands; poor sorting subangular - rounded fragments poorly consolidated (possibly some volcanic cement).
11	22	yellow/brown Sand	to whitish-greyish brown fine-coarse sized; subangular- rounded poorly consolidated; Quartz is coated Iron oxide?
22	29	Silicified (lahar)	in white sandy matrix; consolidated bioclastic fragments up to pebble size (Ø up to 10-15 cm); some laminae
29	102	Oxidised, weathered basalt	x fine subangular - fine subangular; general grey fragments white bands: Calcite (HCl + foam) x red brown bands: Limonite? Iron oxide?
102	125	less weathered chlorite schist	Quartz band smaller & occurs elsewhere more bluish grey in colour
125	126	"fresh" chlorite schist (EOT)	Schistosity, Silky gloss; quartz up to 3% (t) pyrite, calcite (from with HCl) Quartz etc. w/ graphite in some places



top_m	bot_m	litho_main	litho_description
0	2	Soil/Organic	Coarse-fine (silty) sands, poor sorting
2	11	Grey silt with some sand	Subangular to rounded fragments, poorly consolidated (possibly some volcanic glass)
11	22	Yellow/brown Sand	Coarse to fine grain brown, ferroan coated,...
22	29	Silicified (Lahar)	White sandy matrix (poorly sorted, bimodal)...
29	102	Oxidised weathered basalt	Some scoriacity, general clay fragments, wh...
102	125	Basalt	Fresh basalt, small to medium olivine phenocrysts, black glassy groundmass
125	126	Frank chlorite schist	Epidote, calcite (HCl), quartz, schistosity, silty, (possibly up to 30 (Gt))

Example 4: Data mining hydrogeology

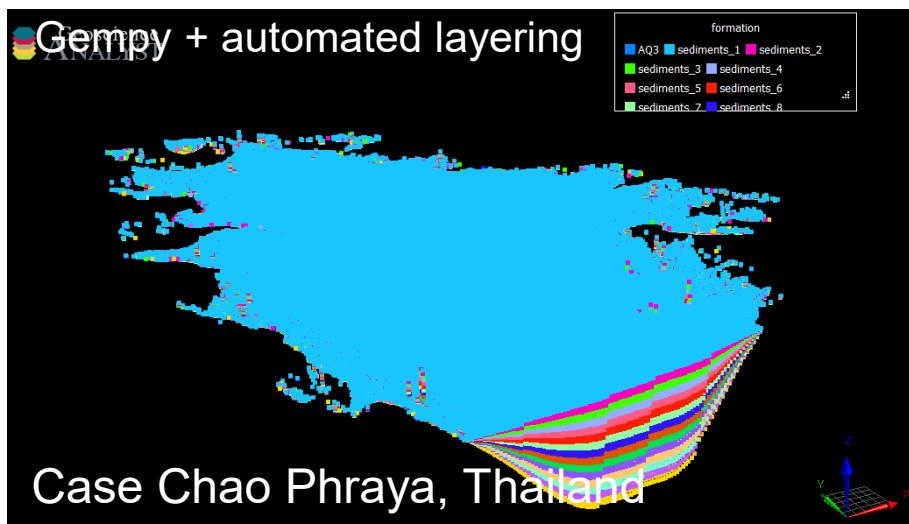
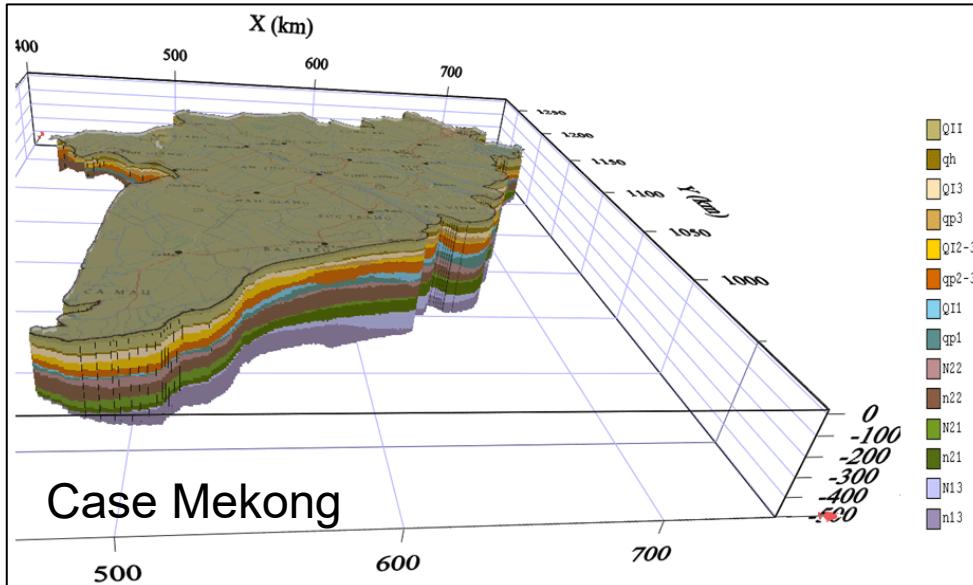
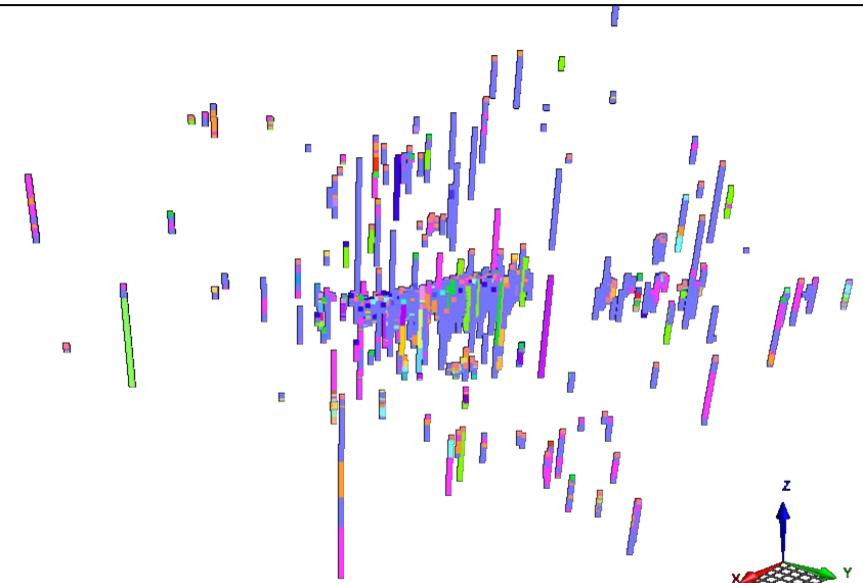


Using the data to construct large-scale hydrogeological models

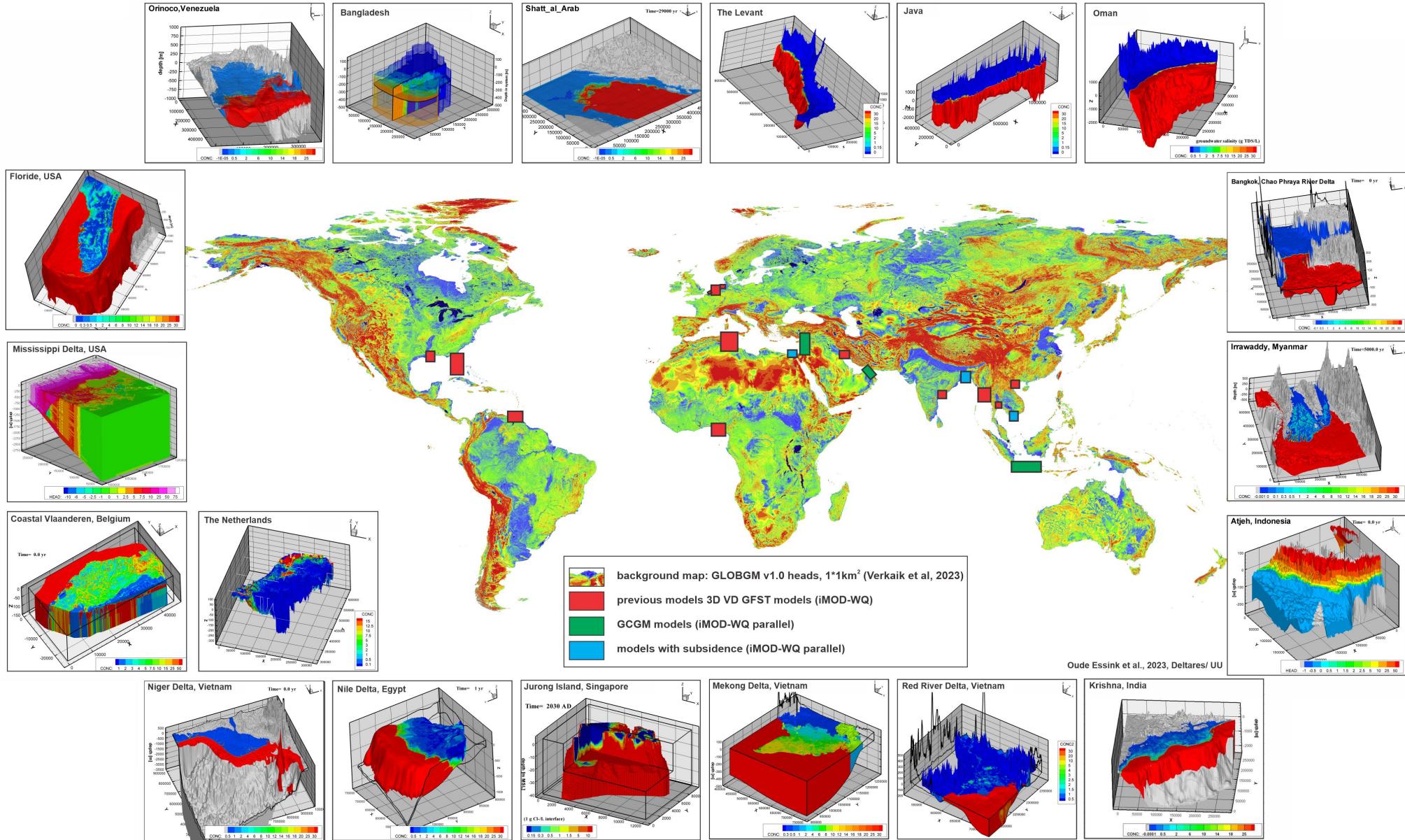
Models can (mainly) parameterised in two ways with data-mining:

1. Hydrogeological properties
2. Initial salinity distributions

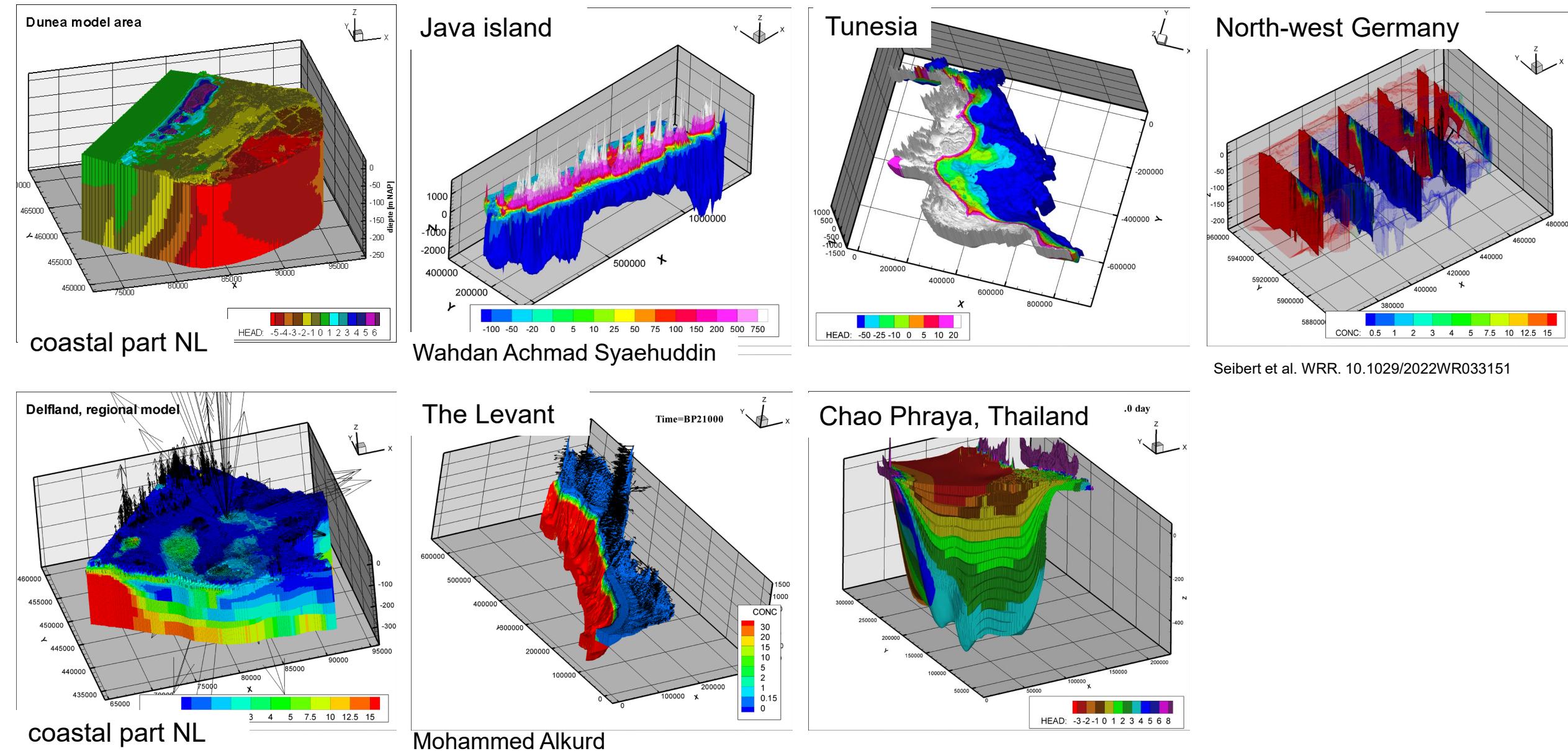
Demonstration of 3D interpolation for hydrogeological properties

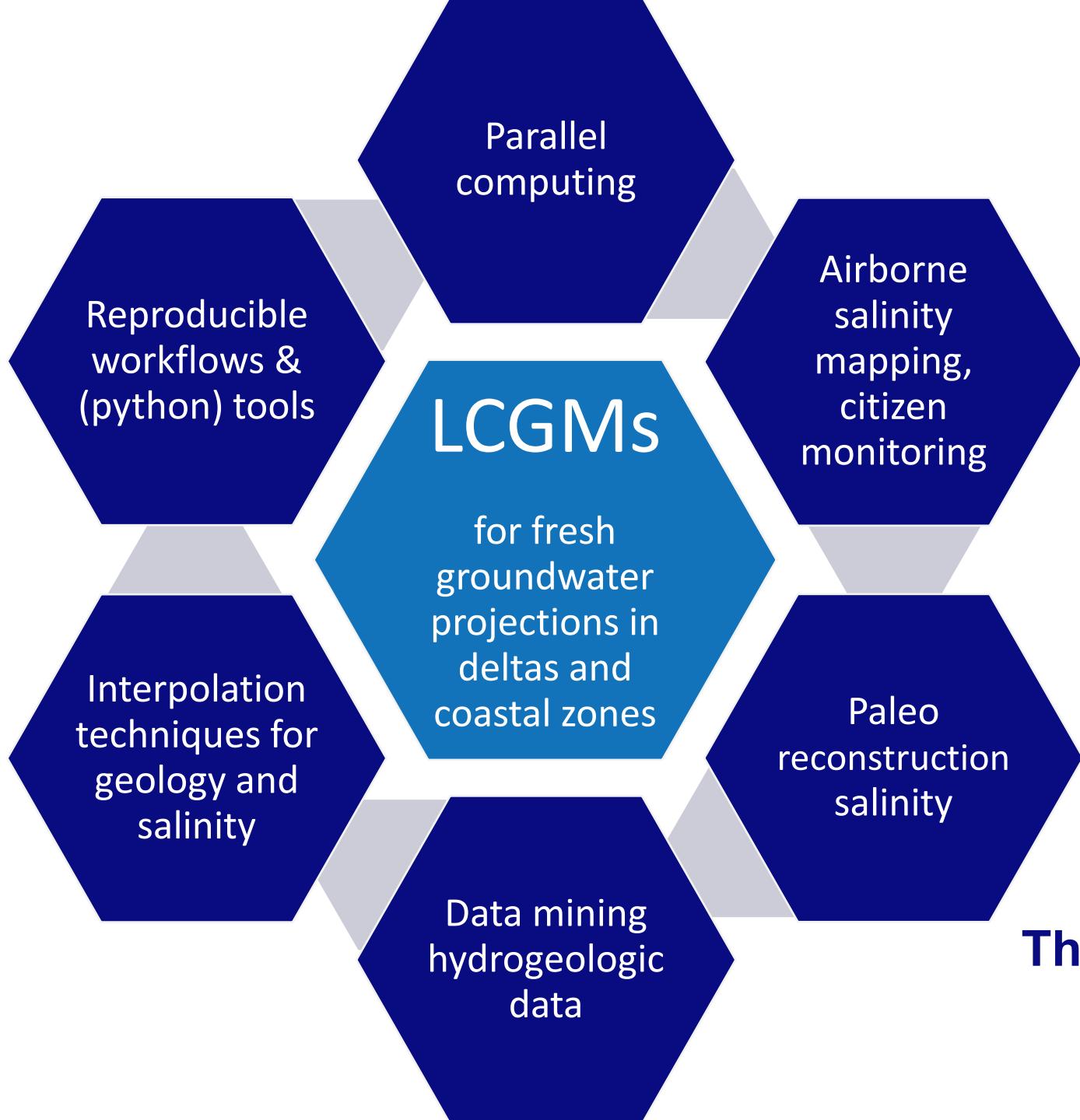


Example 5: some 1*1km² large-scale groundwater models



Example 5: some 1*1km² large-scale groundwater models



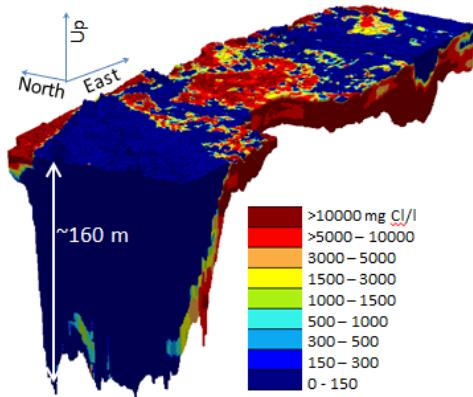
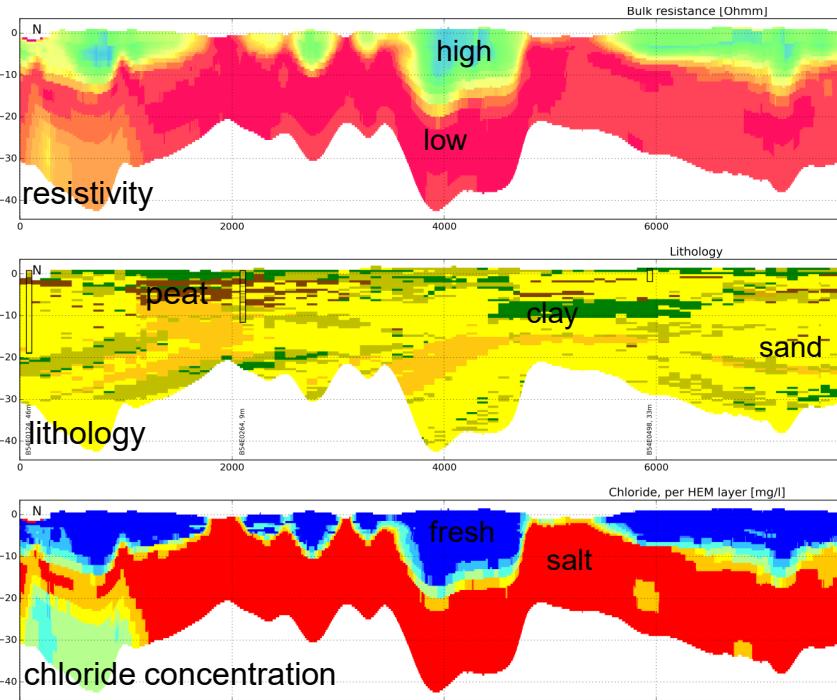


**Thank for your attention
Questions?**

Airborne groundwater salinity mapping



FRESHEM Zeeland



Method:

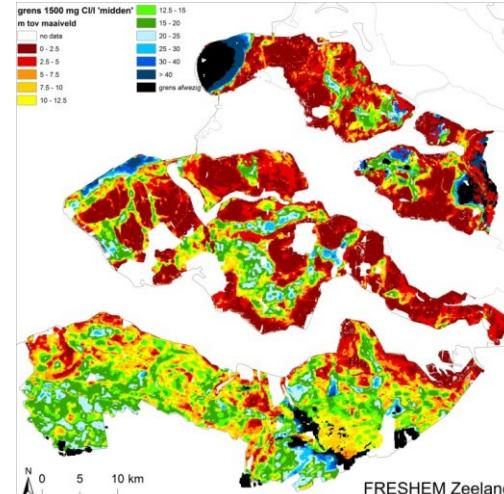
Combination helicopter measurements with knowledge about subsurface and processes in fresh-saline groundwater, and geostatistical mapping via (multiple) indicator kriging.

Results:

- Mapping of 3D groundwater salinity and clay layers

Applications:

- strategic fresh groundwater users & policy makers
- support ASR (COASTAR) in coastal zone
- identify brackish water potential
- improve groundwater models & monitoring



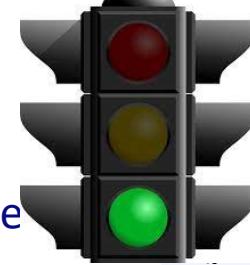
International:

- Project in Flanders, Belgium
- Pilot Mekong, Vietnam

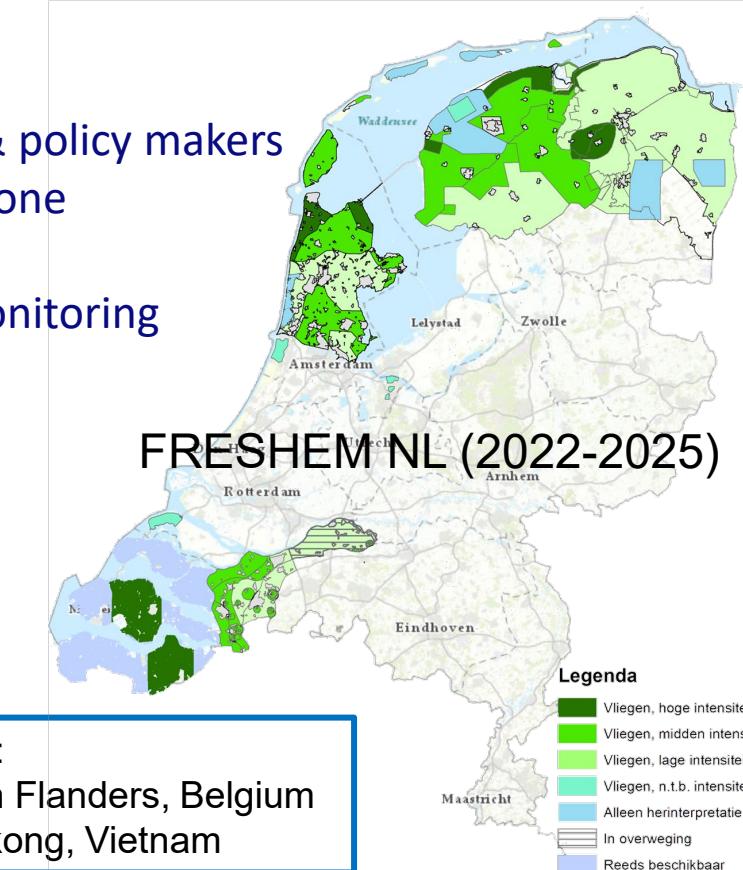
Deltares

TNO innovation
for life

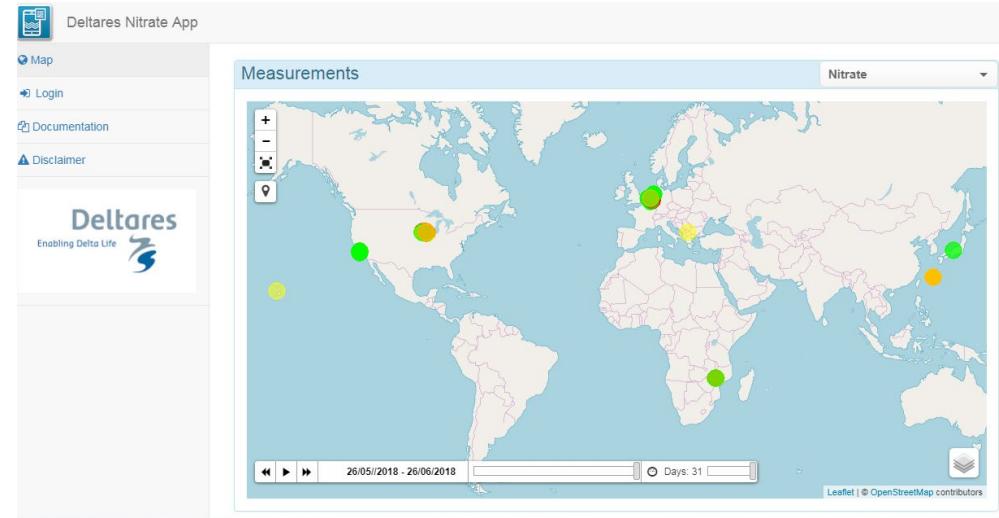
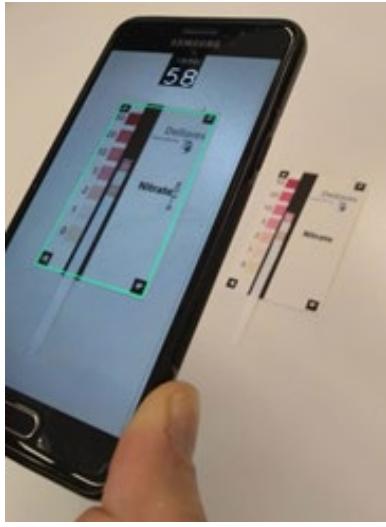
BGR



Building large-scale 3D coastal groundwater models with iMOD-WQ and glob...



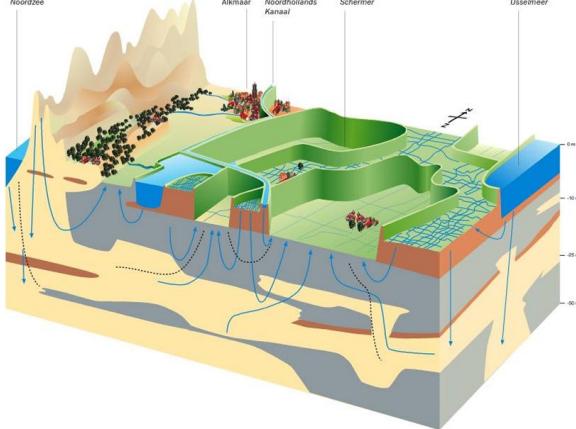
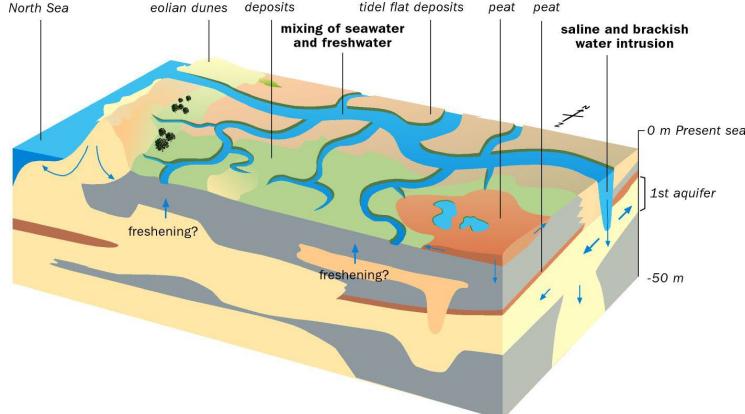
Citizen science, using simple devices and webportals



Building large-scale 3D coastal groundwater models with iMOD-WQ and global

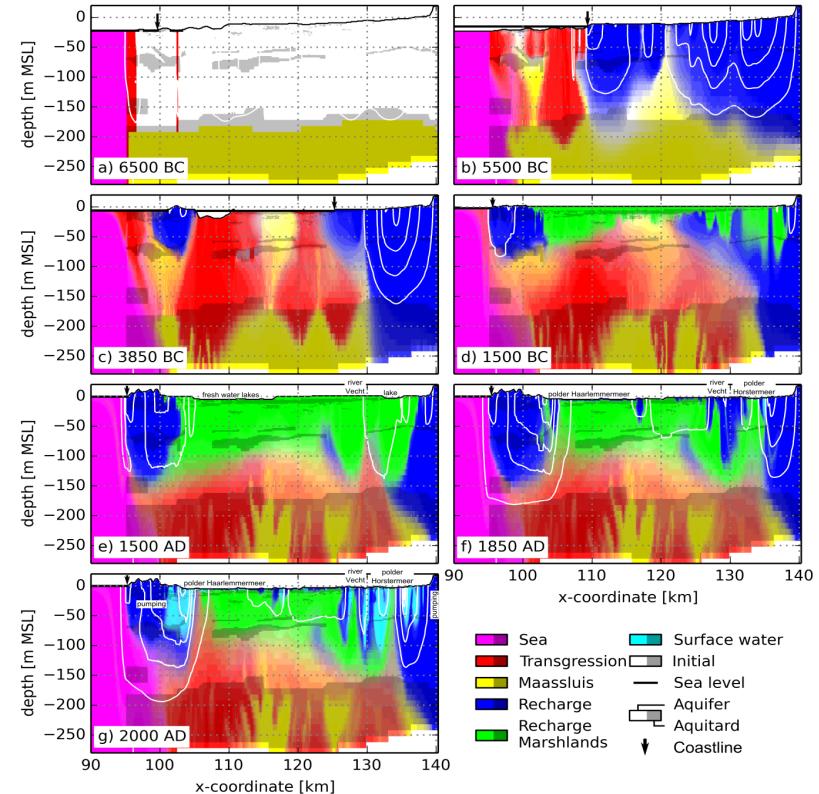
Paleo-reconstructions groundwater salinity

To simulate reconstructions of past hydrological conditions in (data-poor) areas, improving understanding of present groundwater salinity.

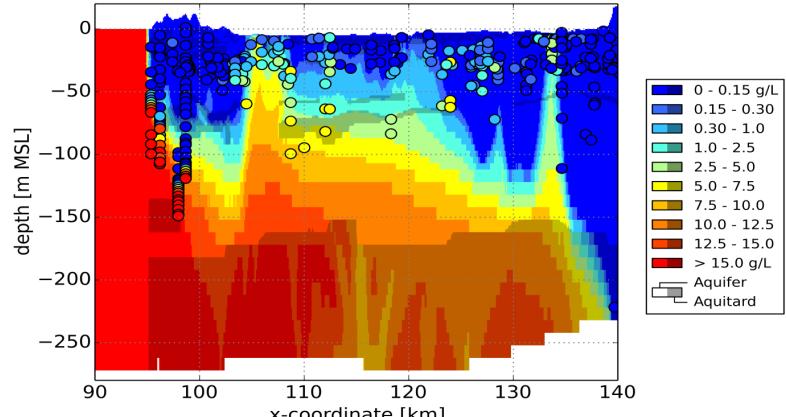


2D → 3D

Origin of groundwater resources



Compare with data



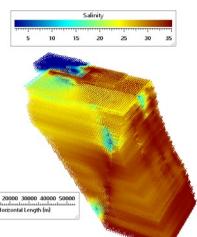
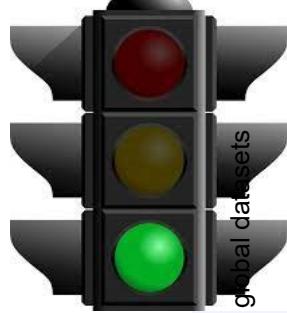
Delsman et al., 2014, HESS

Building large-scale 3D coastal groundwater models with iMOD-WQ and

Why now?

High-resolution global coastal groundwater salinity models are now possible:

1. **Parallel groundwater salinity modelling** (iMOD-WQ / SEAWAT).
2. **Fast Airborne EM groundwater salinity mapping in 3D**, (e.g., FRESHM), citizen science data collection at high TRL.
3. **Paleo reconstructions of past hydrological conditions in data-poor areas**, (possible due to parallel computer), resulting in improved understanding of present groundwater salinity.
4. **More open hydrogeologic data available** (advanced **text mining**, open-source **webportals**).
5. **Advanced techniques for rapid 3D interpolation** of coastal geology and groundwater salinity, and hydrogeological model parameters.
6. **Fully scripted reproducible modelling workflows, clipping & refining** (e.g., iMOD-Python), aiding regular updating and stakeholder trust in model results.
7. **And: groundwater community initiatives**, like Groundwater Model Portal (GroMoPo) (e.g. poster EGU23-12340)



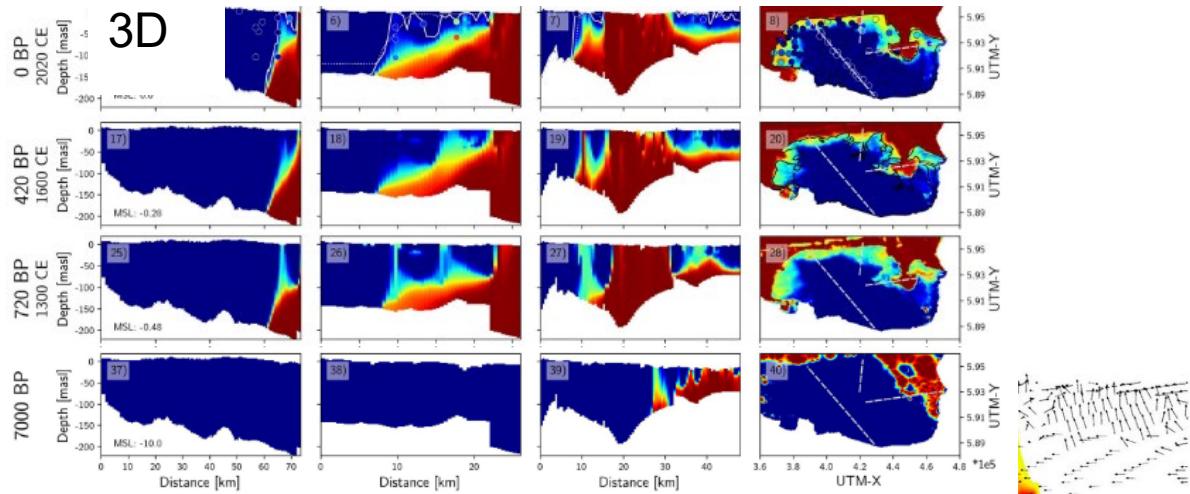
Building large-scale 3D coastal groundwater models with iMOD-WQ and global datasets

Paleo-reconstructions groundwater salinity

Parallel computer power is utilized to simulate 3D reconstructions of past hydrological conditions in (data-poor areas), improving understanding of present groundwater salinity.



Northwest Germany

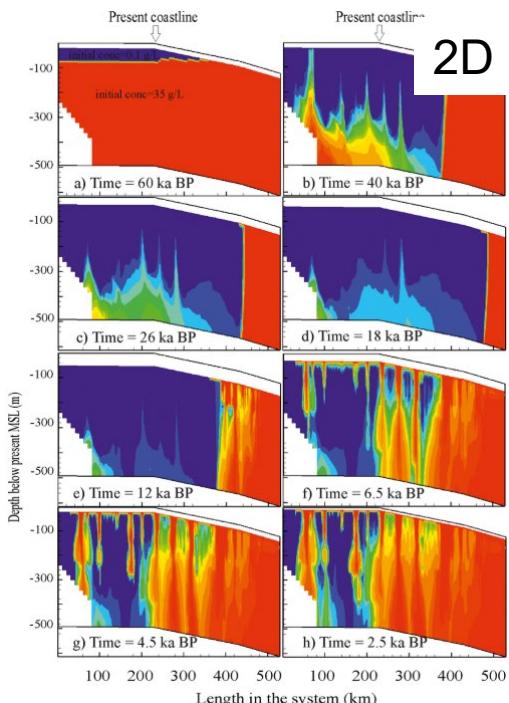


Seibert et al., 2023 WRR

Origin of sources and ~age dating

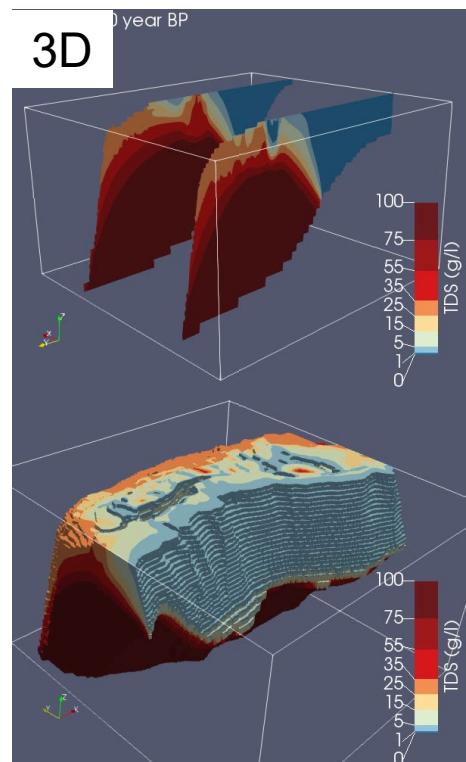


Mekong delta

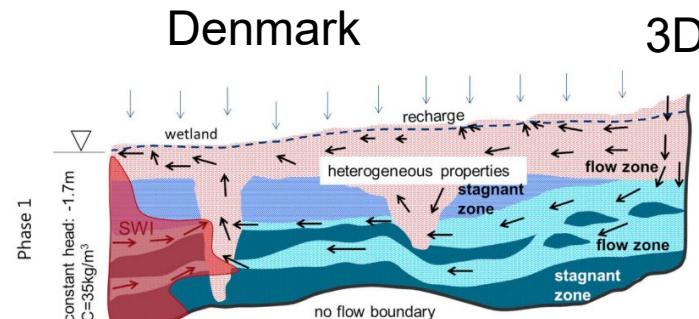


Hung et al., 2019 JoH, RS

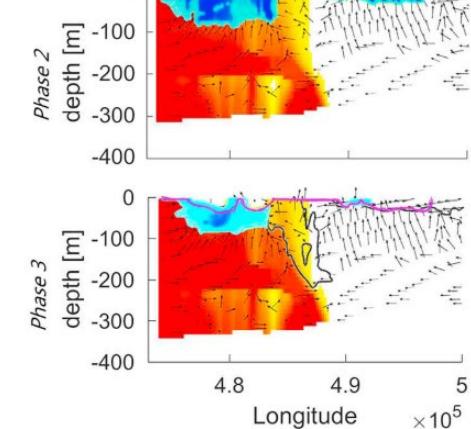
Nile delta



Van Engelen et al., 2019. HESS



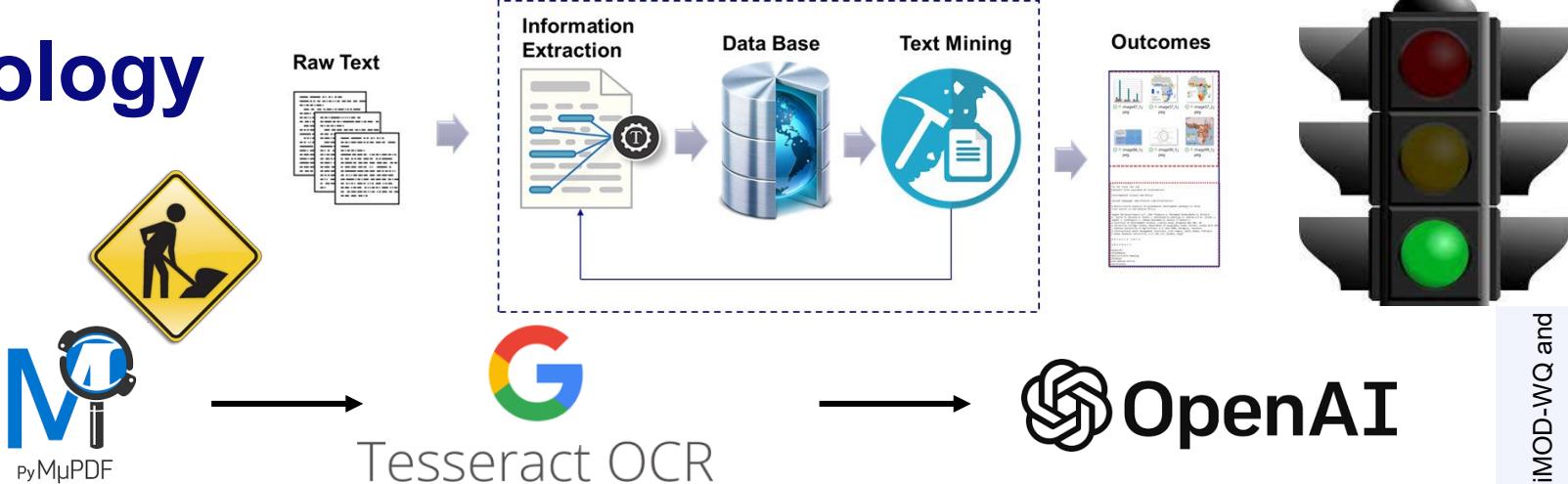
Meyer, et al., 2019



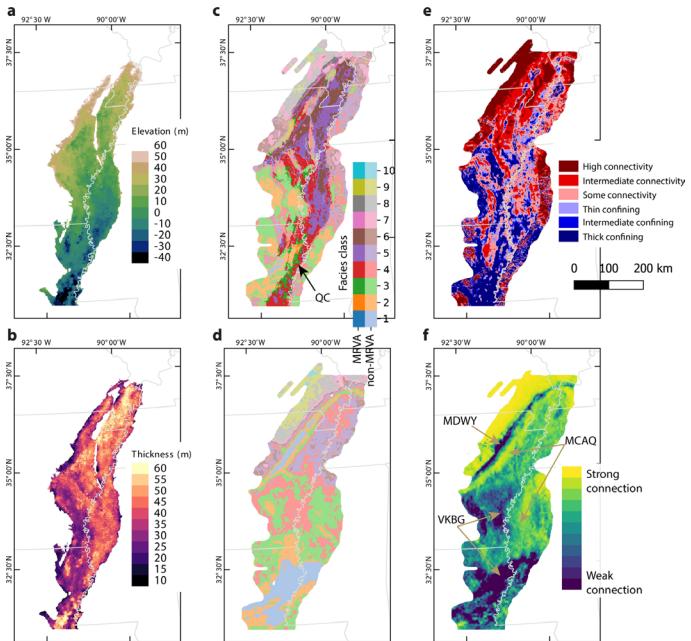
Data mining hydrogeology

Extracting information from images.

Automated in Python using Tesseract
and OpenAI API tools



Input image



Output text (Tesseract)

```
Text editor - 21
35°00' N
32°30' N
92°30' W 90°00' W 92°30' W 90° 00' W
37°30' N
37°30' N
37°30' N
35°00' N
35°00' N
35°00' N
HB High connectivity
|| Intermediate connectivity
||| Some connectivity
I) thin confining
|| Intermediate confining
||| Thick confining
32°30' N
32°30' N
32°30' N
00 200 km
```

Extracted data (OpenAI)

User: "If there are coordinates in the text, extract them as minimum and maximum coordinate pairs:

System: "Example output: X: 100, 100.
Y: 100, 100"

```
Text editor - 21
Minimum coordinate pair: 32°30' N, 90°00' W
Maximum coordinate pair: 37°30' N, 92°30' W
Image: 2021_Airborne geophysical surveys of the lower Mississi_6.png
```

Data mining hydrogeology

Extracting information text.

Automated in Python using OpenAI API

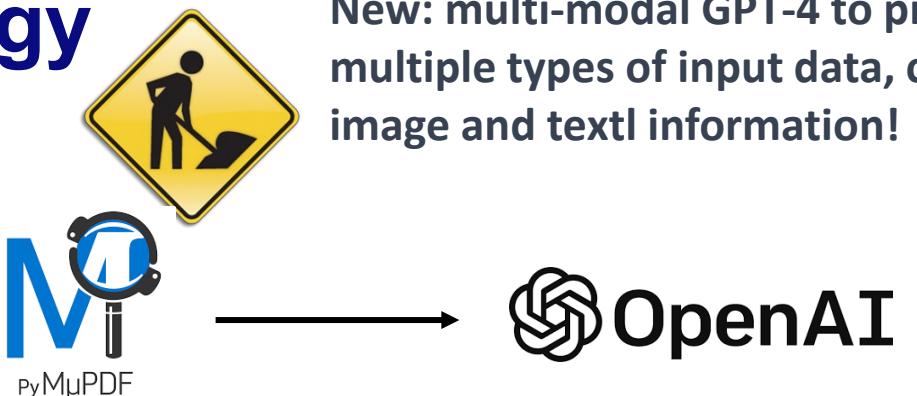
Input text

The screenshot shows a journal article page from 'Advances in Water Resources' volume 160 (2022) page 104118. The page includes the Elsevier logo, a tree illustration, and the journal title. The abstract discusses joint estimation of groundwater salinity and hydrogeological parameters using variable-density groundwater flow, salt transport modelling, and airborne electromagnetic surveys. It lists authors Jude King, Tobias Mulder, Gualbert Oude Essink, and Marc.F.P. Bierkens. A note at the bottom indicates the article is part of a special issue on 'Large-scale 3D coastal groundwater models with MOD-WQ and'.

Joint estimation of groundwater salinity and hydrogeological parameters using variable-density groundwater flow, salt transport modelling and airborne electromagnetic surveys

Jude King^{a,b,*}, Tobias Mulder^a, Gualbert Oude Essink^{a,b}, Marc.F.P. Bierkens^{a,b}

^a Utrecht University, Department of Physical Geography, Utrecht, the Netherlands



New: multi-modal GPT-4 to process multiple types of input data, combining image and textl information!



Extracted data (OpenAI)

User: "Extract model parameters from the text in tabular format:"

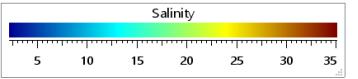
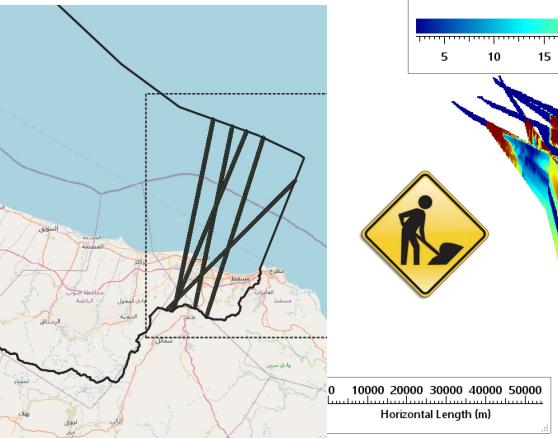
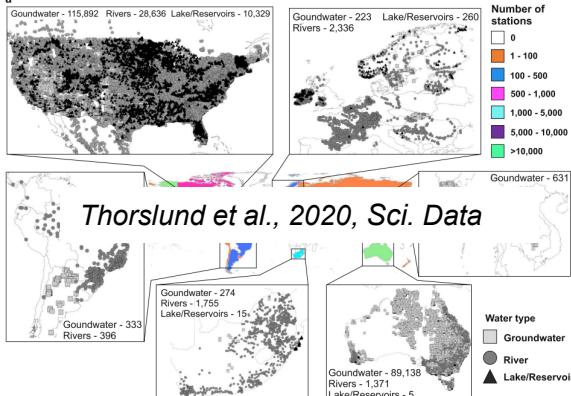
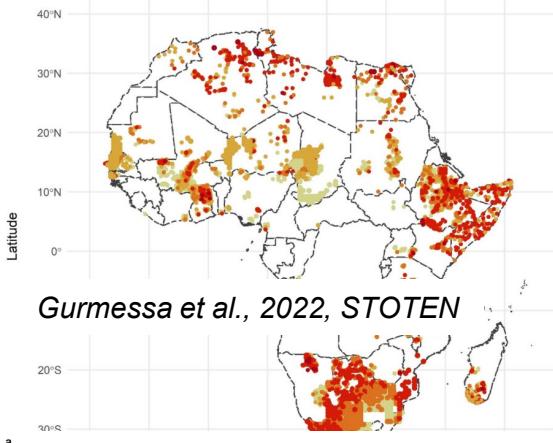
System: "Example: X,Y,Z,1,2,3"

Initial Estimate	Actual Value	Predicted Value (difference actual)
---	---	---
Kh Aquifer (m/day)	1	10 10.63291 (0.63291)
Kv Aquitard (m/day)	0.3	3 3.312433 (0.312433)
Kh Aquitard (m/day)	0.001	0.01 0.011838 (0.001838)
Kv Aquitard (m/day)	0.0005	0.005 0.005831 (0.000831)
Porosity	0.1	0.35 0.386181 (0.036181)

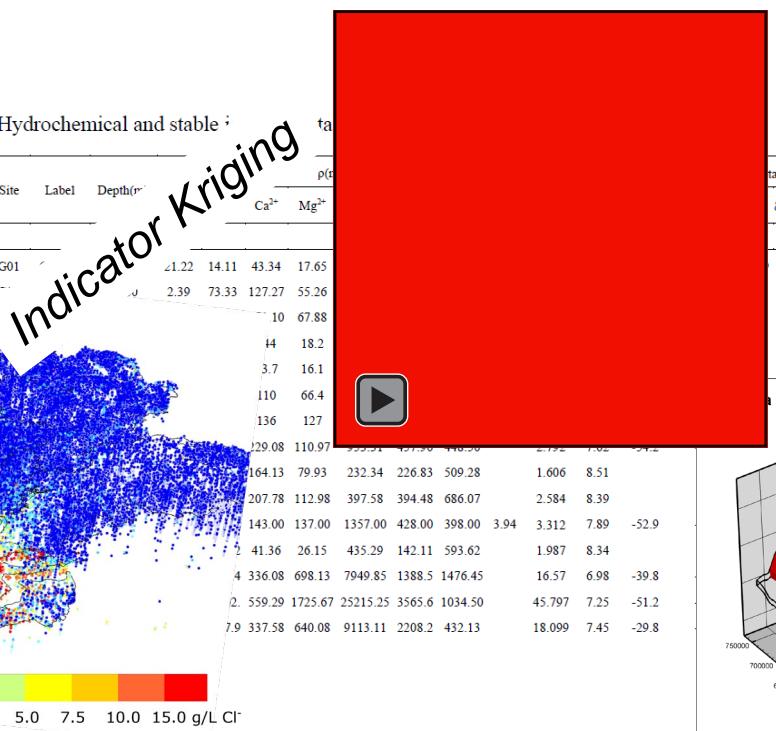
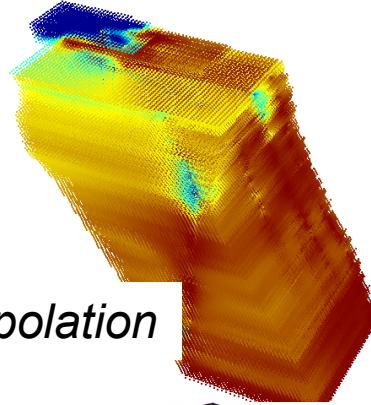
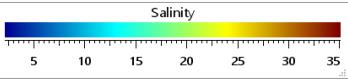
Kh Aquifer | Horizontal hydraulic conductivity of the aquifer | 10 (m/day)
Kh Aquitard | Horizontal hydraulic conductivity of the aquitard | 0.01 (m/day)
Kh/Kv | Anisotropy | 3.3 (aquifer) 2 (aquitard)
Porosity | Porosity | 35 (%)
Recharge winter | Recharge in winter, higher values denote ASR areas | 0.003 m/day (ASR areas), 0.0015 m/day (other areas)
Recharge summer | Recharge in summer, negative value denotes evaporation | -0.0005 m/day
Well Extraction winter | Groundwater extraction in winter | 0 m3/day per model cell
Well Extraction summer | Groundwater extraction in summer | -0.625 m3/day per model

Combining techniques for 3D groundwater salinity

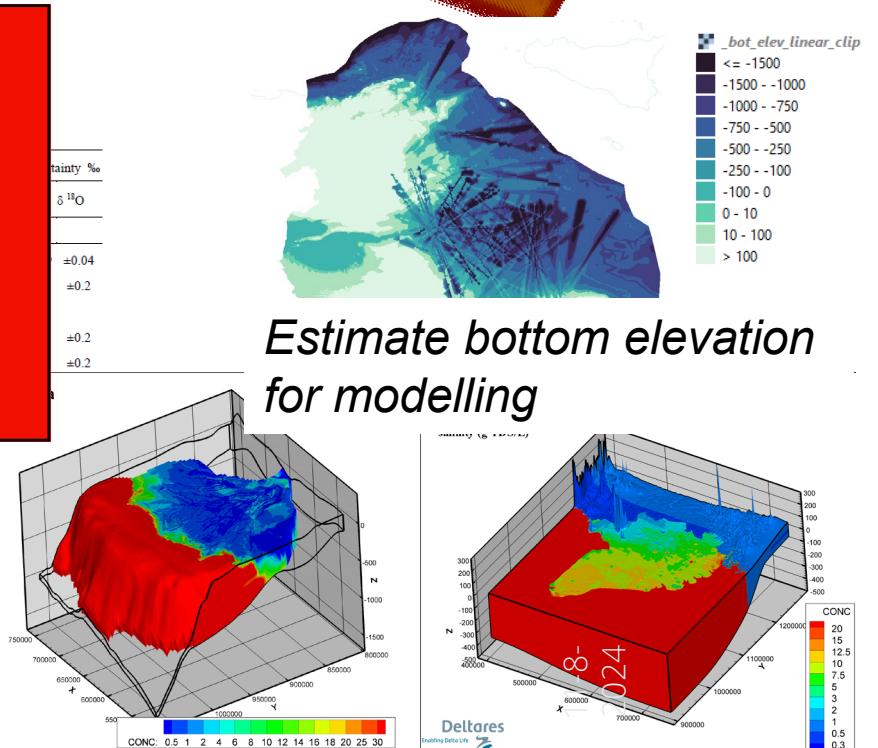
- Airborne EM surveys
- Text mining for pdfs and webportals
- Rapid, automated interpolations
- Paleo reconstructions modelling
- (Citizen science salinity monitoring)



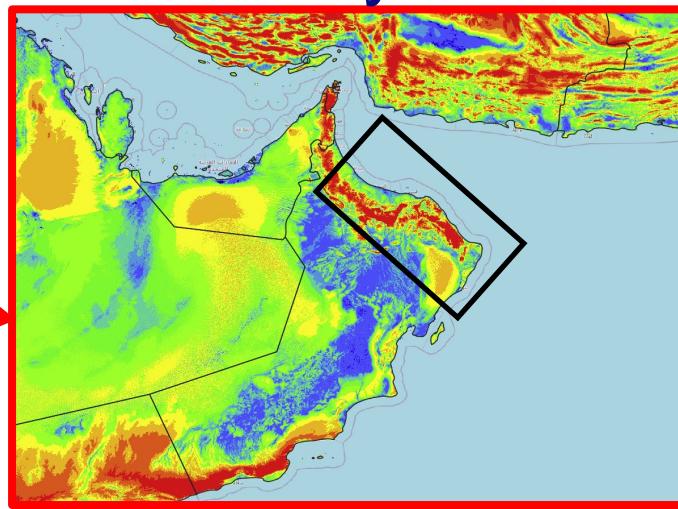
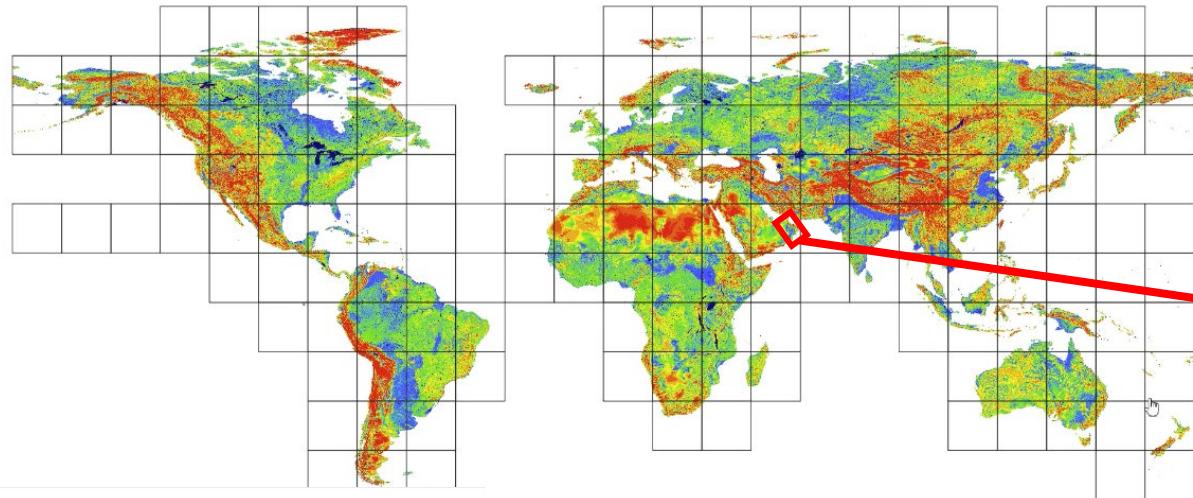
Fast interpolation



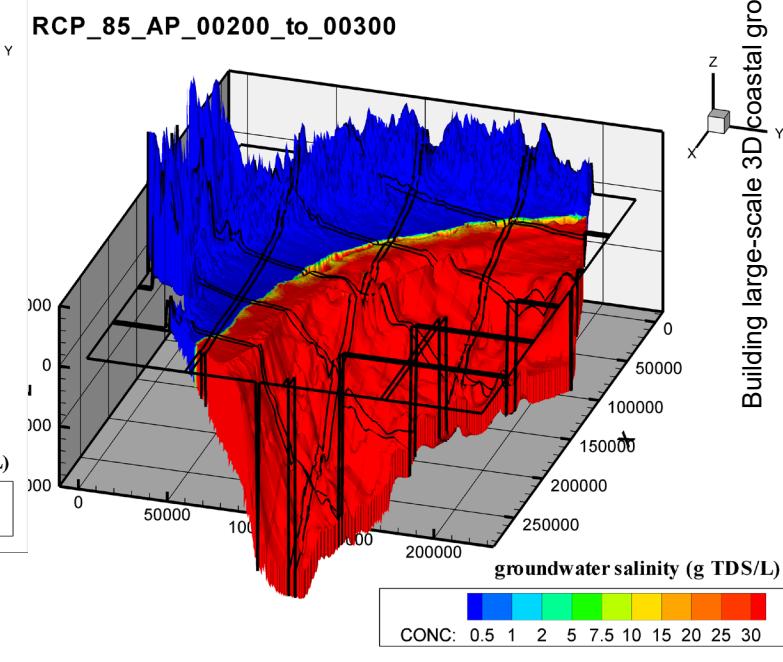
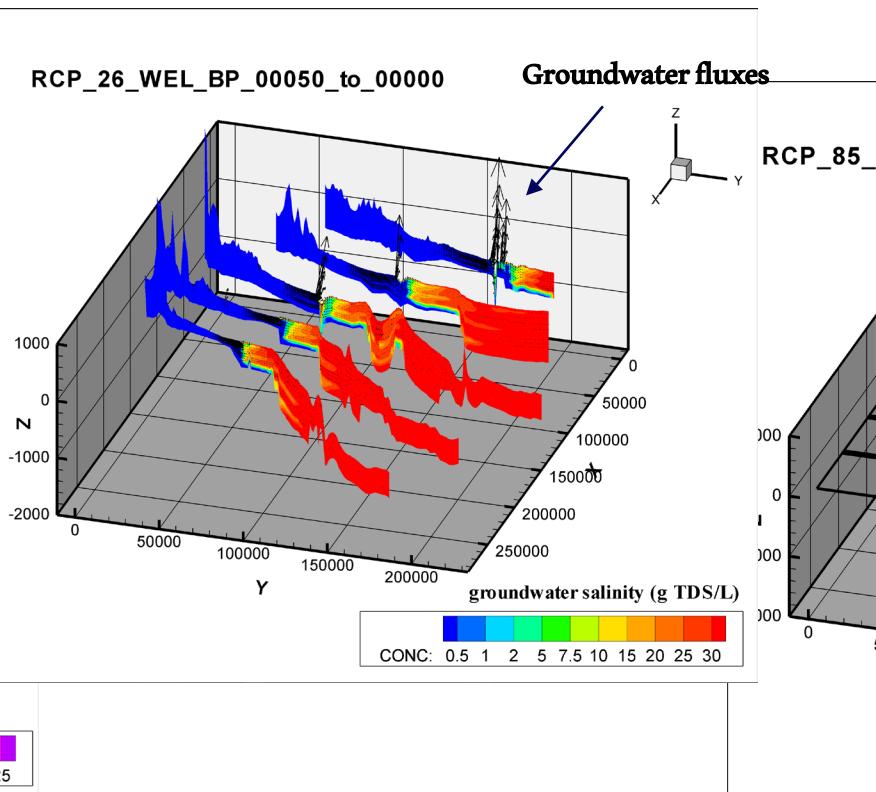
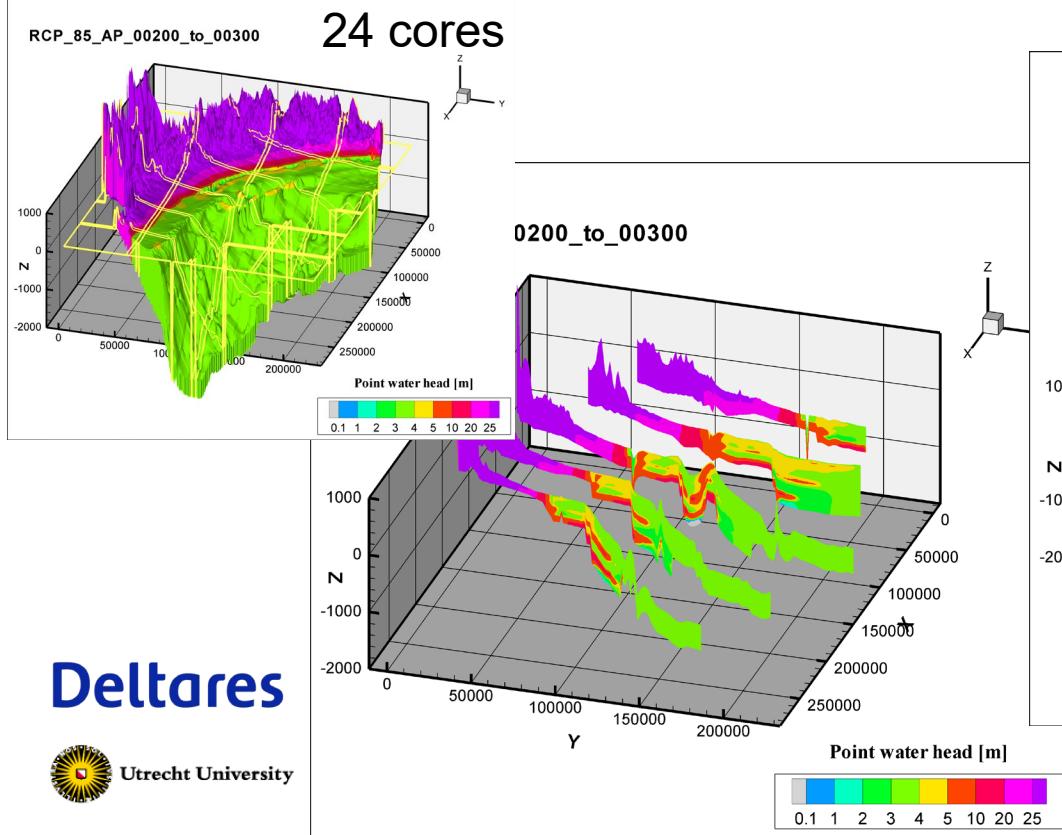
Indicator Kriging



Test 1 Making a regional 3D groundwater salinity model

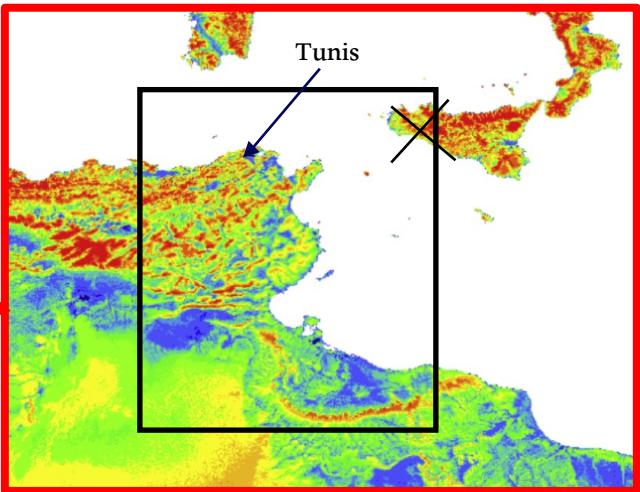
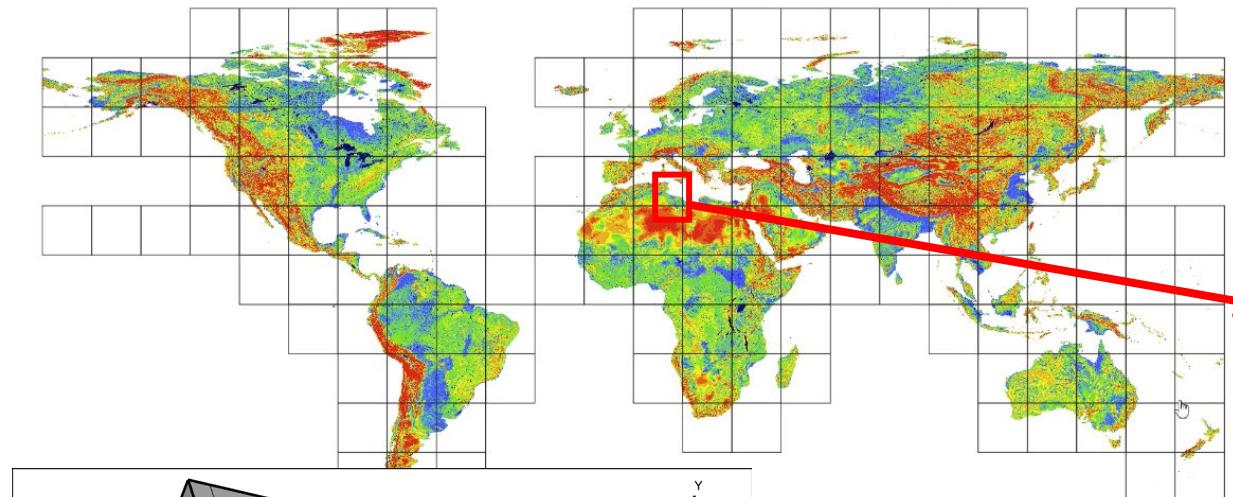


- Oman case
- 223*274*12 cells of 1*1km²;
- Simulating groundwater salinity paleo-reconstruction (120kyrs) and 300 yrs into the future including extractions (using PCR-GLOBWB).
- Computation time: < 1day parallel on only 24 cores; using supercomputer Snellius, but even on a laptop it is doable

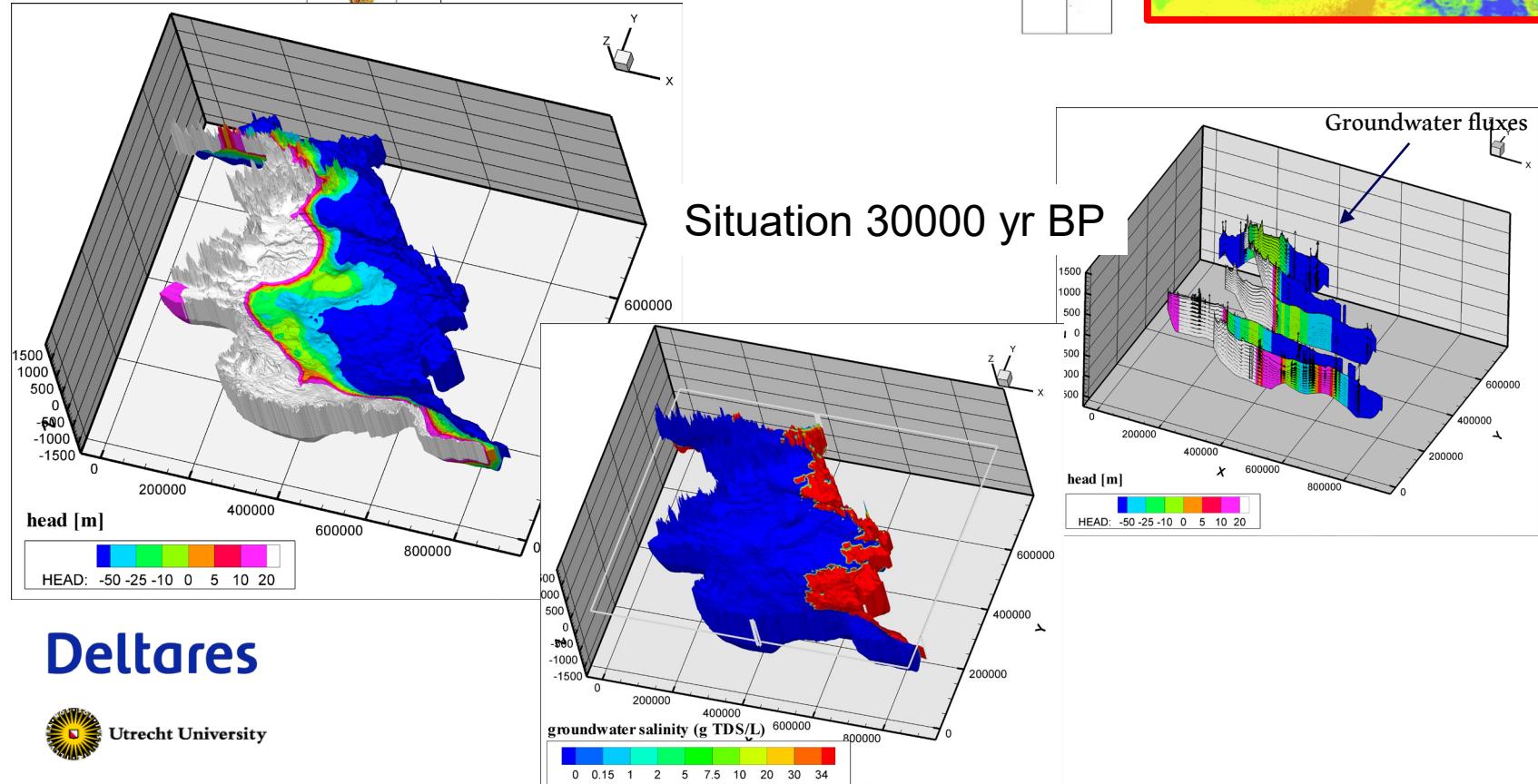


Building large-scale 3D/coastal groundwater models with IMOD-WQ and global datasets

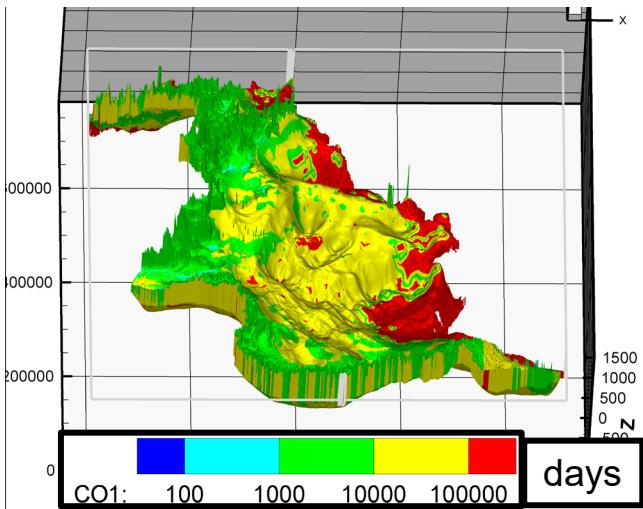
Test 2 Making a regional 3D groundwater salinity model



- Tunisia-Lybia case
- 907*747*12 cells of 1*1km²
- Simulating groundwater salinity paleo-reconstruction
- For now 30-20kyrs BP.
- Testing parallel on 62-128 cores; using supercomputer Snellius



Automated STEPSIZE analysis!



- Advection,

$$\Delta t \leq \frac{R}{|v_x|/\Delta x + |v_y|/\Delta y + |v_z|/\Delta z}$$

Deltares

Thank for you attention

Questions?

Gualbert Oude Essink, Daniel Zamrsky, Marc Bierkens

gualbert.oudeessink@deltares.nl

pdf on wiki freshsalt.deltares.nl

More information:

Parallel SEAWAT, imod-python and 3D viewer:

- <https://oss.deltares.nl/web/imod/about-imod5>
 - Verkaik, J. et al., 2021. Distributed memory parallel computing of three-dimensional variable-density groundwater flow and salt transport. *Adv. Water Resour.* 154, 103976. <https://doi.org/10.1016/j.advwatres.2021.103976>
 - https://deltares.github.io/iMOD-Documentation/python_index.html
 - https://deltares.github.io/iMOD-Documentation/viewer_index.html



iMOD Python



iMOD Viewer

Reproducibility and transparency, Gitlab

- <https://gitlab.com/deltares/imod/nhi-fresh-salt>
- Delsman, J.R. et al 2023. Reproducible construction of a high-resolution national variable-density groundwater salinity model for the Netherlands. *Environ. Model. Softw.* 105683. <https://doi.org/10.1016/j.envsoft.2023.105683>
- 3D Paleo-reconstruction groundwater salinity and iMOD-WQ
 - Seibert, S.L. et al., 2023. Paleo-hydrogeological modeling to understand present-day groundwater salinities in a low-lying coastal groundwater system (Northwestern Germany). *Water Resour. Res.* <https://doi.org/https://doi.org/10.1029/2022WR033151>
 - Van Engelen, J., Verkaik, J., King, J., Nofal, E.R., Bierkens, M.F.P., Oude Essink, G.H.P., 2019. A three-dimensional palaeohydrogeological reconstruction of the groundwater salinity distribution in the Nile Delta Aquifer. *Hydrol. Earth Syst. Sci.* 23, 5175–5198. <https://doi.org/10.5194/hess-2019-151>

**EGU23-7607
HS8.2.6**

A study on the suitability and quantitative potential of aquifer storage and recovery and brackish water extraction in Dutch coastal areas.

Iija America -
van den
Heuvel et al

**EGU23-15557
HS8.2.6**

Monitoring & simulation groundwater salinity due to extractions in a coastal aquifer

Thijs
Hendrikx et al

**EGU23-17249
HS8.2.6**

Effects surface water boundary condition scaling on modelled groundwater salinity and salt fluxes

Ignacio Farias
et al

**EGU23-2859
HS8.2.6**

Assessing impact of climate change and anthropogenic factors on future salinization; a case in Northwestern Germany)

Stephan L.
Seibert et al

**EGU23-1844
Henry Darcy
Medal Lecture**

Global Water Resources and the Limits to Groundwater Use

Marc Bierkens

And PICO Zamrskly et al.: **EGU23-11444 HS8.2.6**



Deltares

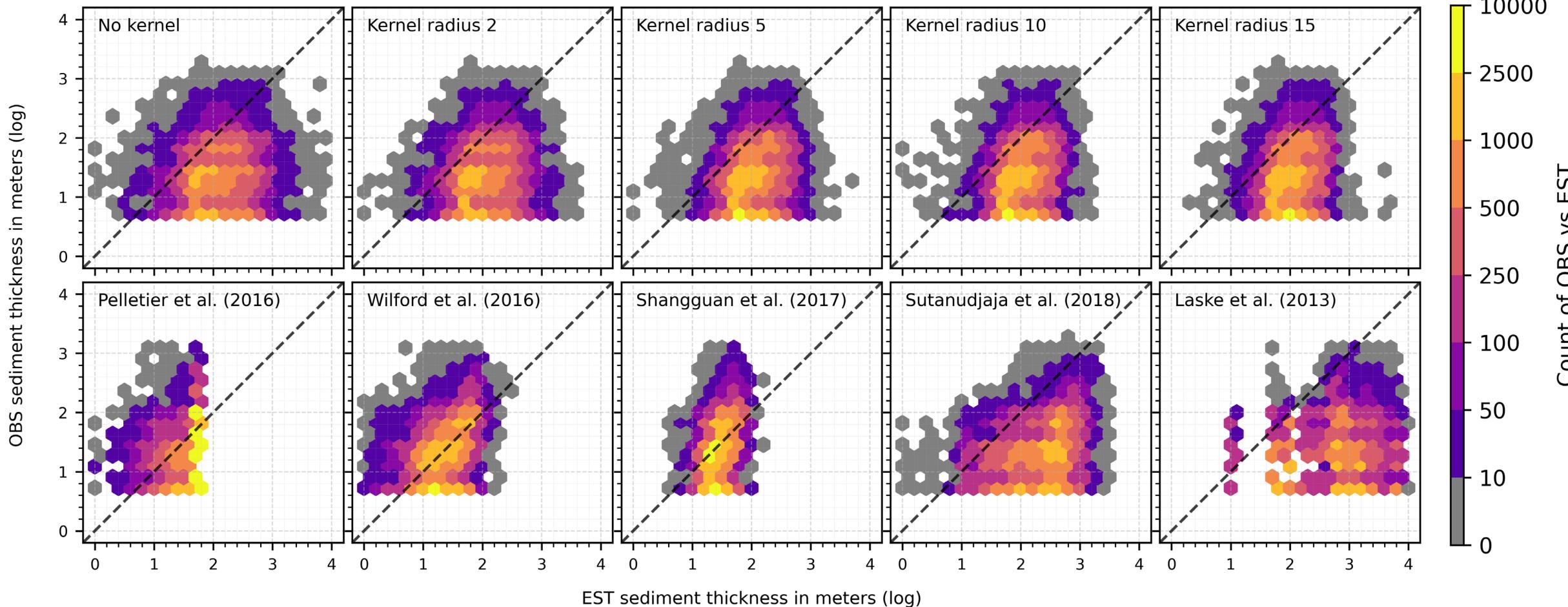
Orange issues

- Calibration, validation, verification.
- Tekst mining: IPR of articles.
- Interferences with local hydrogeological communities, some same regional scale.

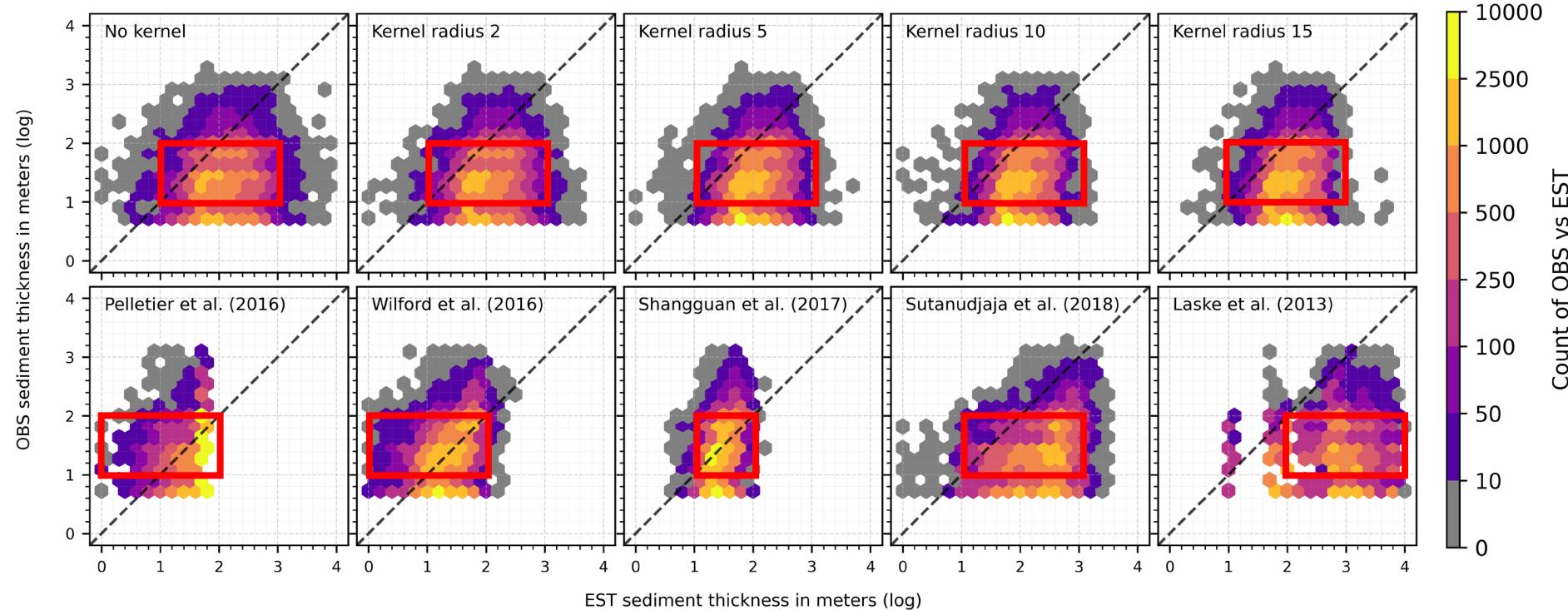


Example 3: Validation with ~40000 boreholes

- Bedrock must be reached, borehole located in unconsolidated sediment lithology (GLIM)
- Bias towards shallow boreholes – deeper boreholes might not reach bedrock..



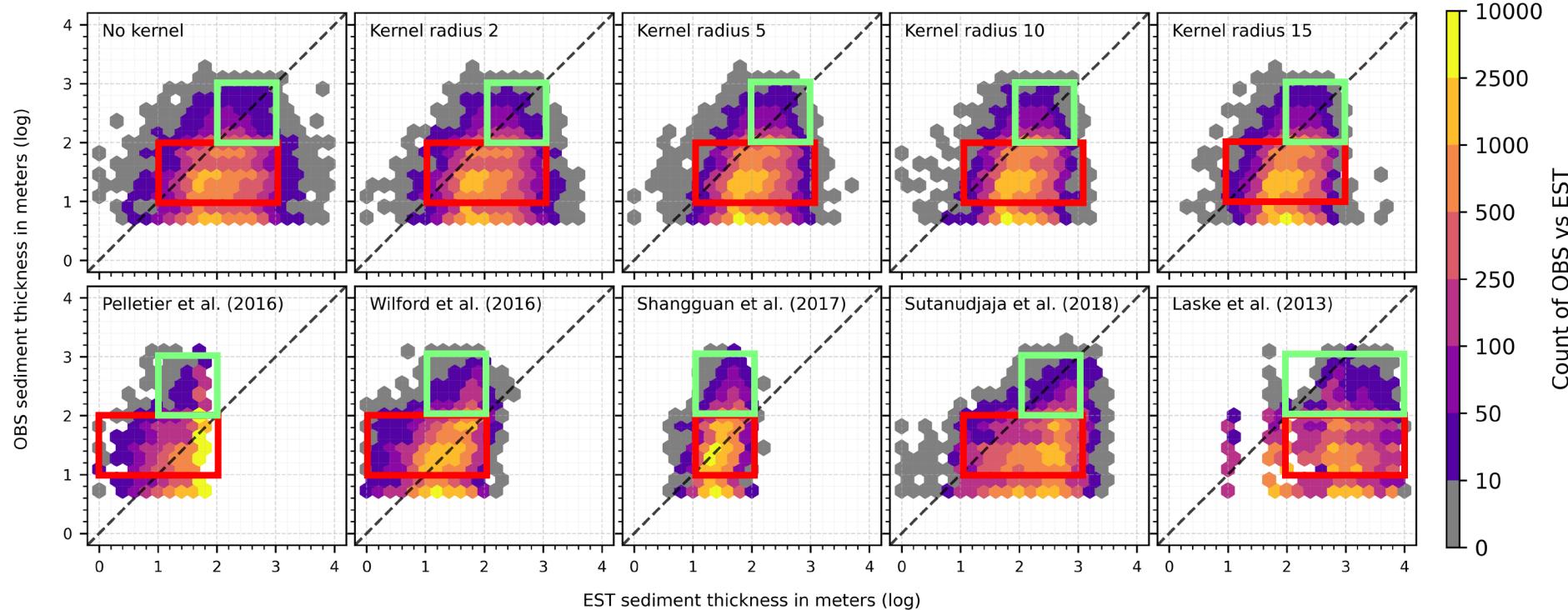
Example 3: Validation with ~40000 boreholes



Range 10m – 100m OBS thickness

- Lots of overestimations by 1 order of magnitude
- Better fit in regolith datasets, most are machine learning using the borehole data as training data.. (bias)
- PCR-GLOB and CRUST 1.0 overestimate even more, understandable..

Example 3: Validation with ~40000 boreholes



Range 100m – 1000m OBS thickness

- Pretty good fit overall in this range – way lower number of borelogs with this sediment depth..
- Regolith datasets underestimate quite a lot (hundreds of meters), due to bias in training data?
- PCR-GLOB and CRUST 1.0 better than the regolith data but still can overestimate quite a lot

Developments into LCGM version 1.0

Early 2023 -----> Early 2024 -----> Late 2024

