The Water AI partnership: Recycling energy industry data to find new water resources

Marina Flores[1], Claudia Bertoni[1], Andrew Walker[1], Chia-Hsin (Wendy) Tsai[1], Orla Marnell[2]
[1]Department of Earth Sciences, University of Oxford.
[2]Wood Mackenzie.

The University of Oxford Department of Earth Sciences EARTHSCIENCES

UNIVERSITY OF OXFORD

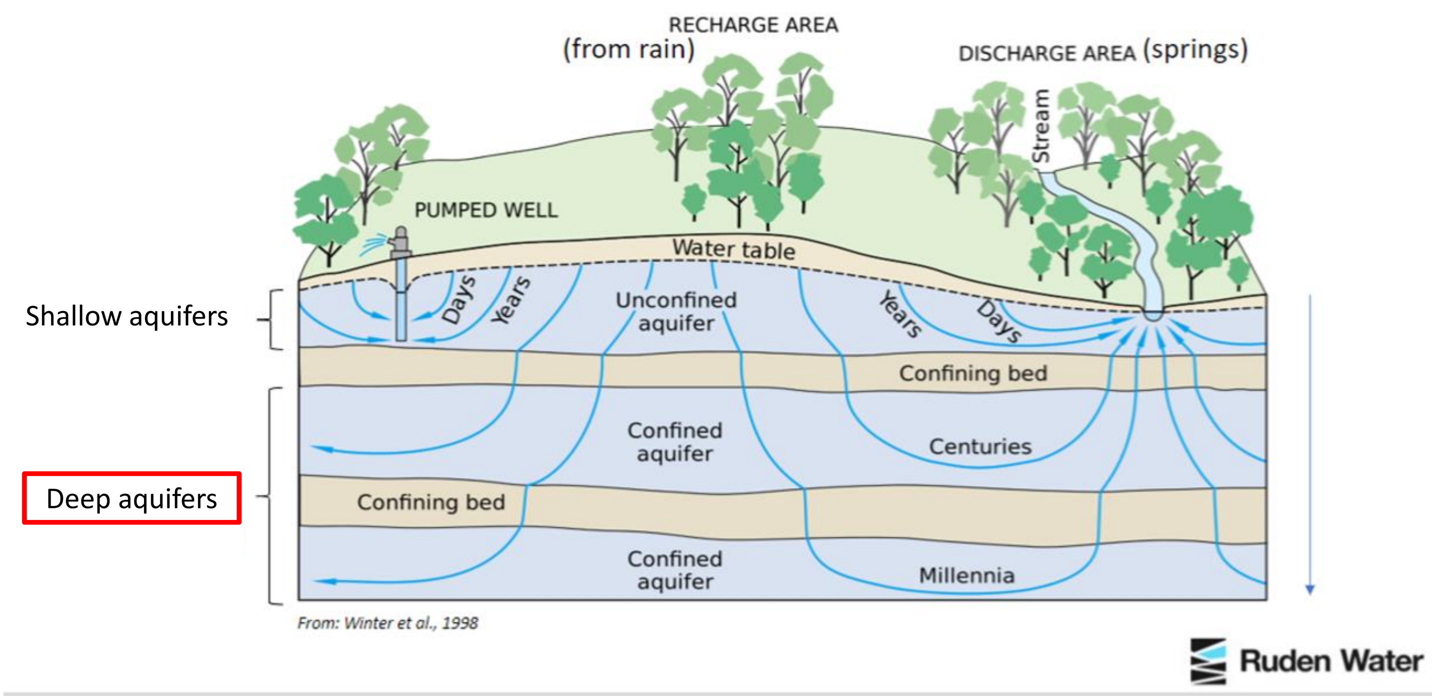
Introduction

As the global population continues to grow the demand for freshwater is increasing, while the supply is decreasing due to factors such as climate change and overuse of traditional freshwater sources. Can unconventional, ultra-deep aquifers and offshore fresh groundwater resources help addressing the water crisis?

'Water AI' is an Oxford EPSRC Impact Accelerator Account-funded project, in partnership with Wood Mackenzie. The main project aim is to focus on identification of unconventional water resources by:

1) Repurposing oil & gas data for water resources
2) Using Machine Learning (ML) to predict the distribution of unconventional fresh groundwater
3) Enhancing visualisation of data with GIS and Virtual Reality (VR) .

Motivation: Reliable deep groundwater sources. Unconventional. We need deeper data.

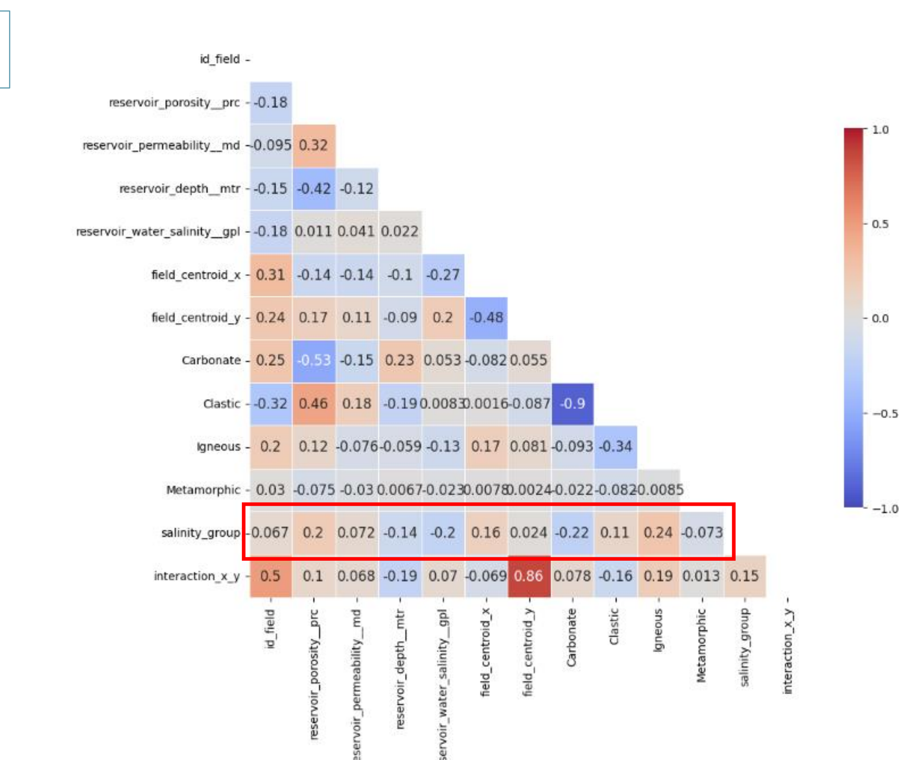


Machine Learning (ML)

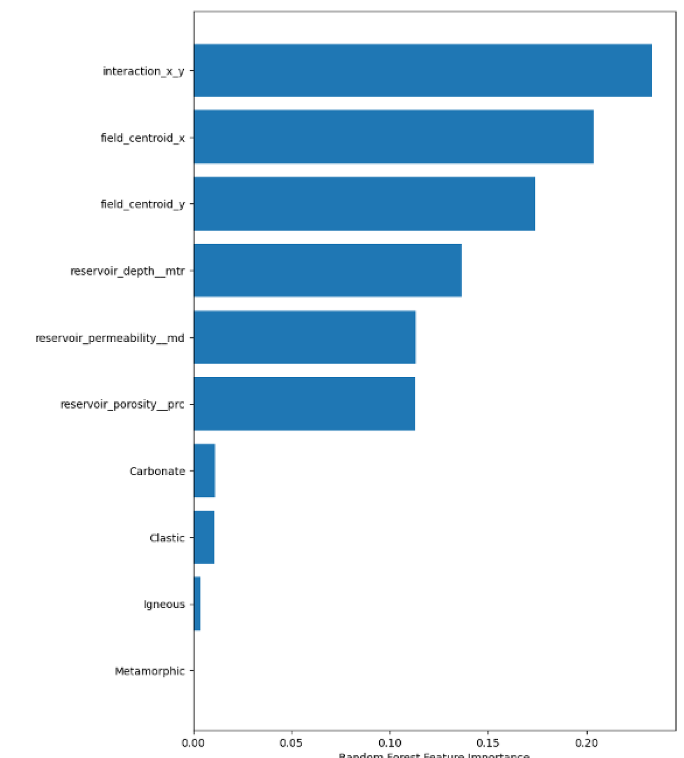
ML enables computers to learn from data and make predictions or decisions without being explicitly programmed. For the project we tested different types of ML models including supervised, unsupervised methods and a neural network. The ML process is iterative, initially with data exploration, and then refining the feature engineering step and pre-processing techniques. The insights gained from ML can help uncover relationships not initially considered. The method that gave us the best results was a supervised random forest classifier, where the salinity values were factorised by their potential end use, e.g. potable 0-10 gpl.

- In fig. A selected parameters were analysed to find relationships. Lithology (categorical data) was encoded to be compared to numerical data, and included in the ML models. The correlation matrix shows the correlation coefficients between pairs of variables in a dataset. The red colour indicates a strong positive, and the blue colour a strong negative correlation. Note there are no strong correlations between the salinity groups and the other variables.
- Fig. B shows the feature importance for the random forest classifier, i.e. the features that contribute the most to the predictions of salinity group.
- Fig. C shows the classification results where 65% of the test datapoints were predicted correctly, however an increased number of datapoints would help to improve this model.

A



B



C

Predicted Values \ Actual Values	1. Potable water/low salinity (0-10 gpl)	2. Agricultural & desalination uses (10-20gpl)	3. Cheaper desalination potential (20-35gpl)	4. More expensive desalination potential (35-100gpl)	5. Limited industrial use brine (100+ gpl)	All
1. Potable water (0-10 gpl)	15	1	0	3	2	**21**
2. Agricultural & desalination uses (10-20gpl)	4	6	1	1	0	**12**
3. Cheaper desalination potential (20-35gpl)	2	1	6	1	1	**11**
4. More expensive desalination potential (35-100gpl)	3	4	2	19	8	**36**
5. Limited industrial use brine (100+ gpl)	0	0	0	0	19	**19**
All	**24**	**12**	**9**	**24**	**30**	**99**

Geographic Information Systems (GIS)

GIS is a powerful technology that leverages the spatial dimension to unlock patterns, relationships, and opportunities that traditional data visualisation approaches might overlook. In this project, we've combined salinity values from Wood Mackenzie and academic databases, to perform a spatial analysis relative to depth and other parameters.

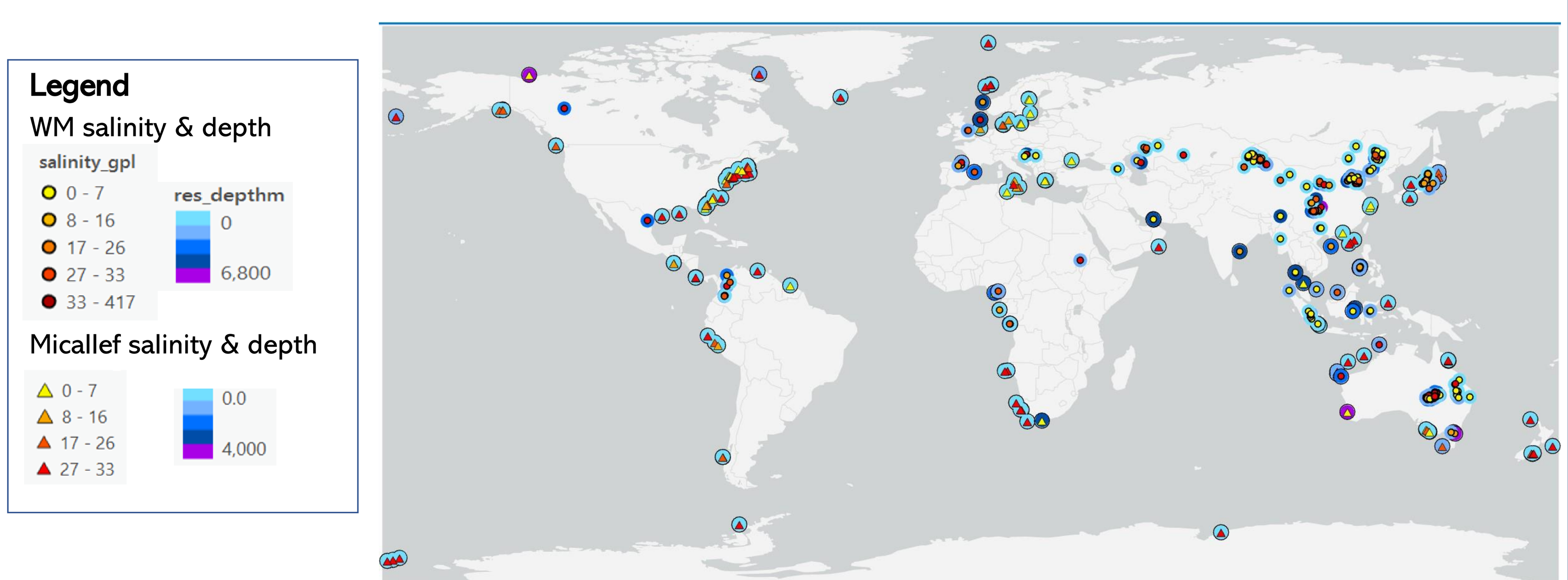


Figure 1. Global salinity map showing industry-partner data and the database by Micallef et al, 2020.

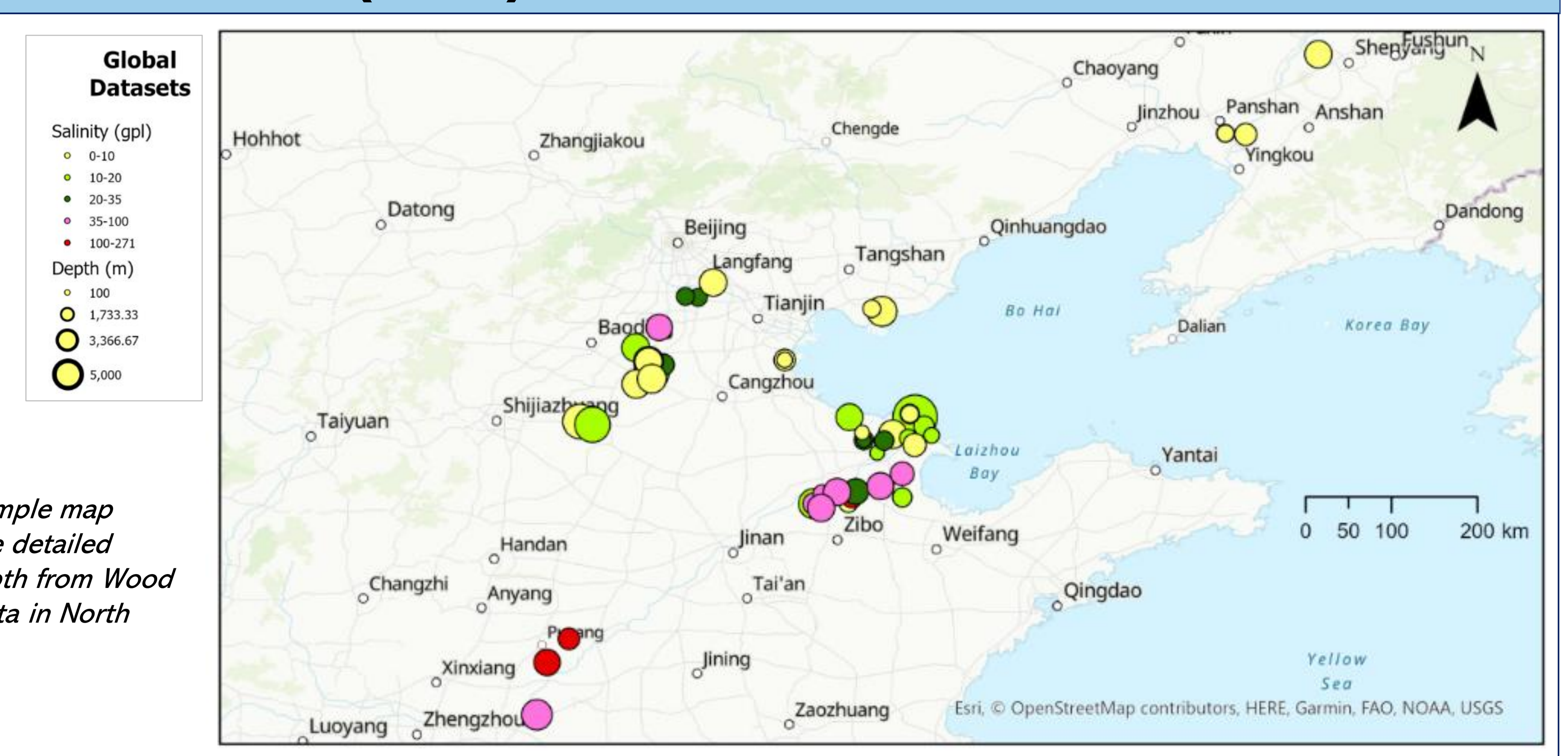


Figure 2. Example map showing more detailed salinity at depth from Wood Mackenzie data in North China.

The analysis indicates basins of interest with low salinity at 1000-5000 m depth in Asia, south Australia, and south and west coasts of Africa.

Virtual Reality (VR)

Our VR model is currently under development.

VR is an immersive technology that offers a simulated, interactive experience in a 3D digital environment. In Geosciences, VR presents a transformative tool for visualising the subsurface. We are working on converting complex geological data into immersive 3D models, to gain a deeper understanding of subsurface structures, and ultra-deep aquifers.

VR allows users to navigate through these models in real-time, enhancing spatial comprehension and facilitating data-driven decision-making, while enhancing collaboration among experts.

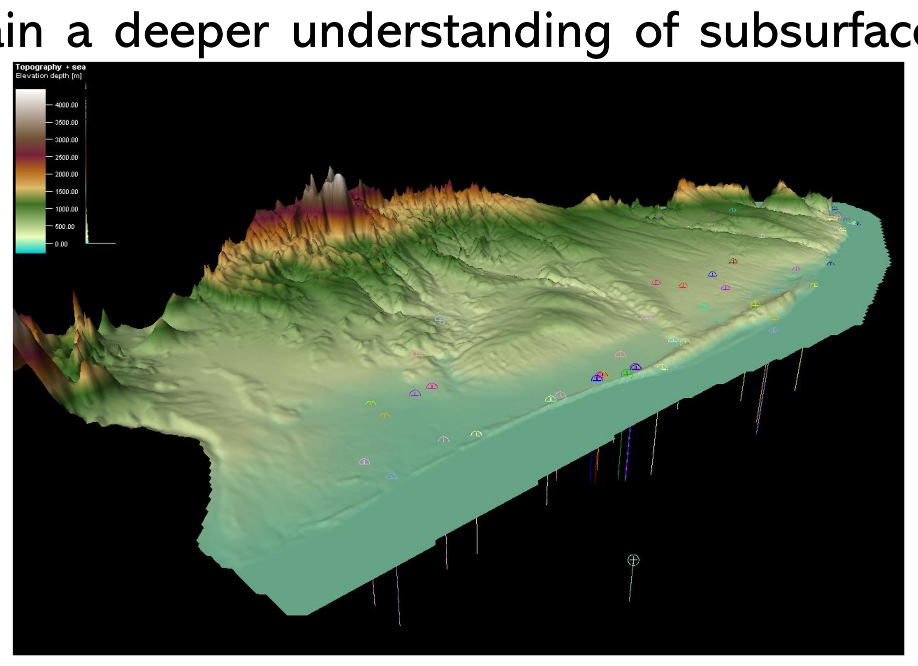


Figure 3. Example 3D map showing the topography in Somalia with the drilling locations.

Conclusions

Our project uses a combination of technologies that allow a novel perspective on the investigation of unconventional groundwater resources. GIS analysis allows us to identify the distribution at global scale of this resource. Data analysis and ML have shown potential to predict salinity groups relevant for human consumption and industrial uses. However, to increase the accuracy of the model, we would need greater volumes of data across all salinity groups.

Throughout the project, Earth Sciences Oxford have collaborated with Wood Mackenzie and have exchanged knowledge on workflows and methods. This type of collaboration is increasingly important in the context of the energy transition, where the vast datasets held by the O&G industry could be repurposed for other applications.

References

Bertoni et al., 2020. *Seismic reflection methods in offshore groundwater research. Geosciences, 10(8)*

Micallef et al., 2021. *Offshore freshened groundwater in continental margins. Rev. of Geophysics, 59(1).*

Winter et al., 1998. *Ground water and surface water: A single resource. USGS Circular 1139.*