A Matter of Scale: Harnessing Imperfect Data to Understand Groundwater Across Space and Time

27/08/2026

DONALD JOHN MACALLISTER + MANY OTHERS
DONMAC@BGS.AC.UK

British Geological Survey

Overview

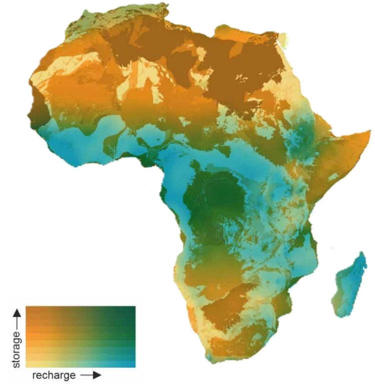
Groundwater in space

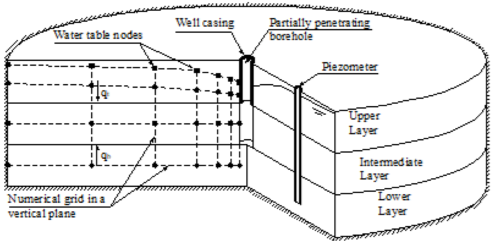
- Groundwater at the continental scale
- Groundwater at the regional, national and sub-national scale
- Groundwater at the borehole scale

Groundwater in space and time

- Groundwater over the long-term
- Groundwater forecasting
- Groundwater during climate extremes

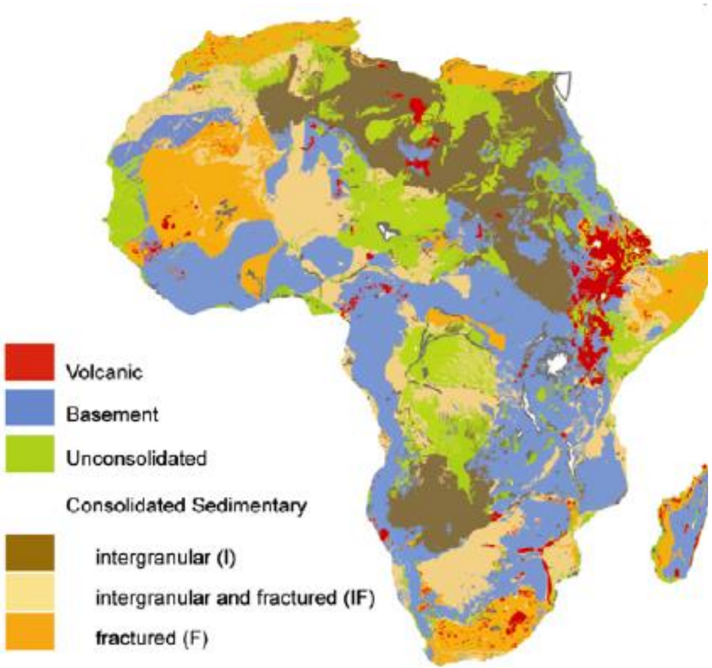
What role does Artificial Intelligence (AI) have?

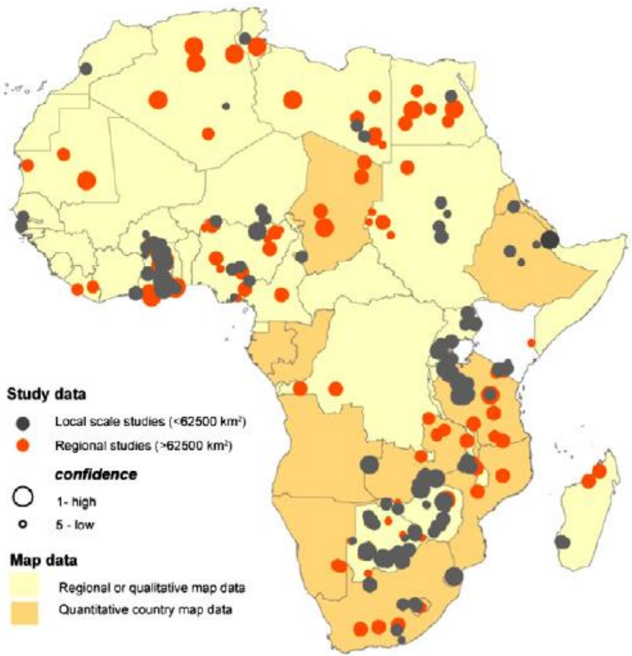




Groundwater at a continental scale

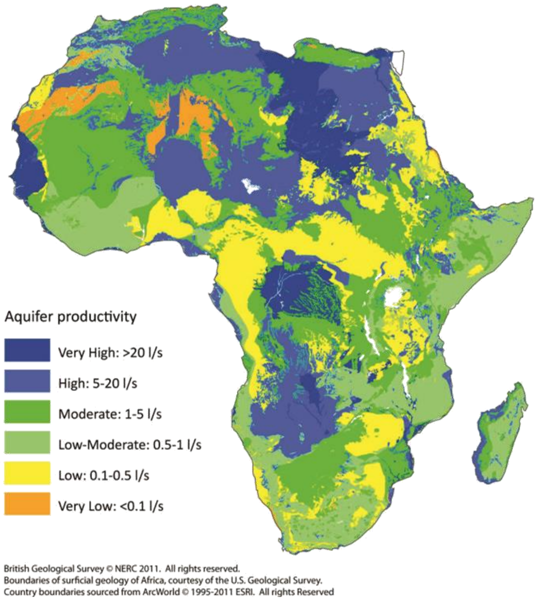
Quantitative maps of groundwater resources



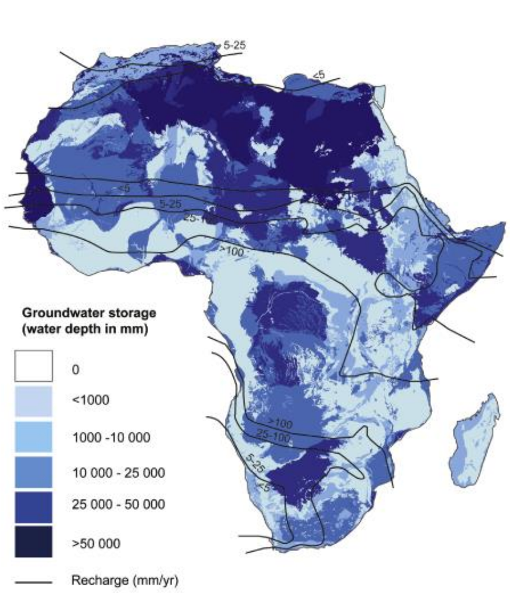


MacDonald, Alan M., et al. "Quantitative maps of groundwater resources in Africa." *Environmental Research Letters* 7.2 (2012): 024009. https://doi.org/10.1088/1748-9326/7/2/024009

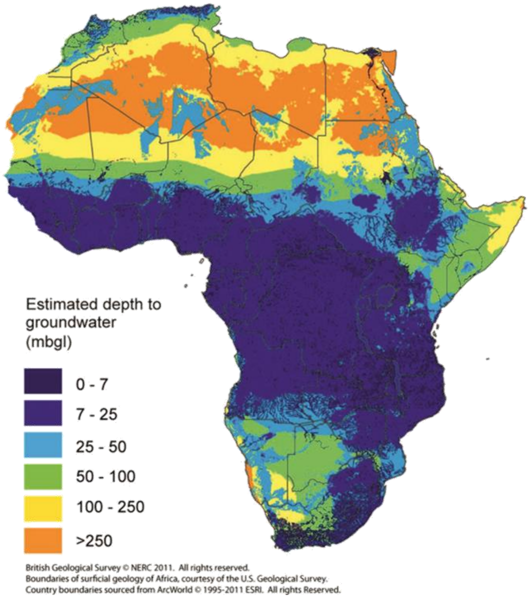
Quantitative maps of groundwater resources



Aquifer productivity



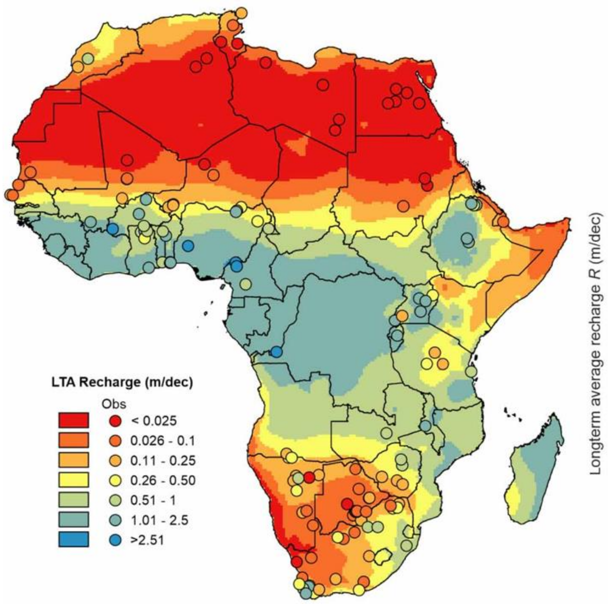
Aquifer storage

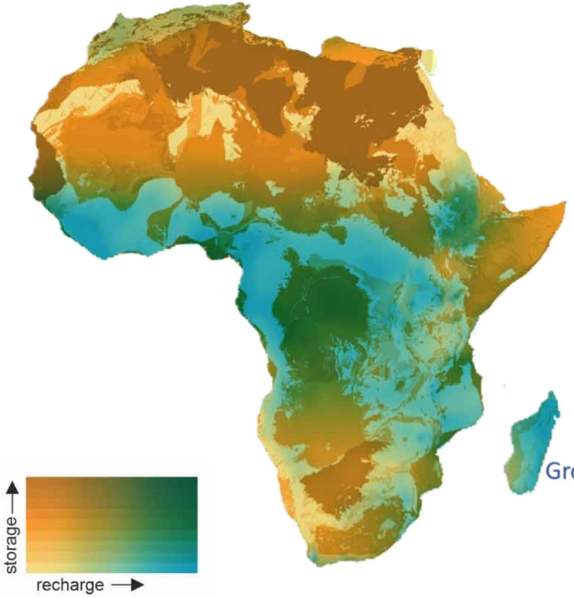


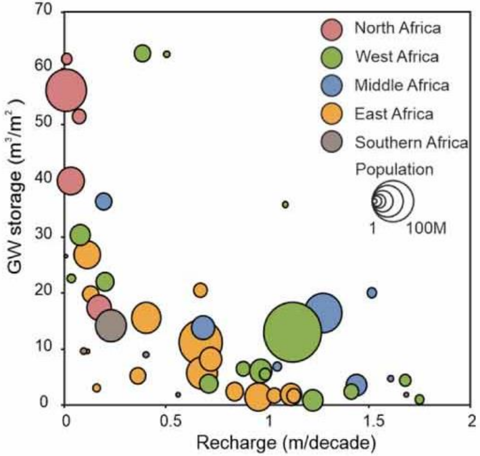
Depth to GW

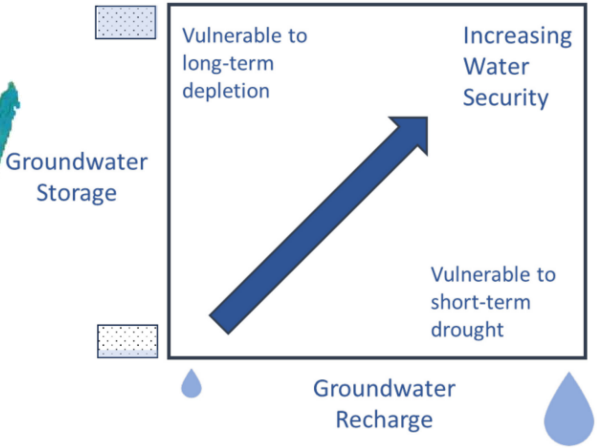
MacDonald, Alan M., et al. "Quantitative maps of groundwater resources in Africa." *Environmental Research Letters* 7.2 (2012): 024009. https://doi.org/10.1088/1748-9326/7/2/024009

Groundwater recharge and groundwater security





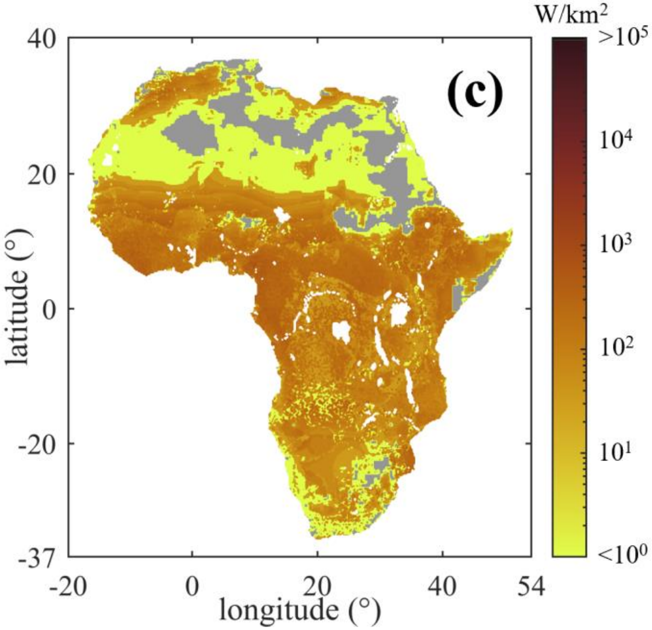




MacDonald, Alan M., et al. "Mapping groundwater recharge in Africa from ground observations and implications for water security." Environmental Research Letters 16.3 (2021): 034012. https://doi.org/10.1088/1748-9326/abd661

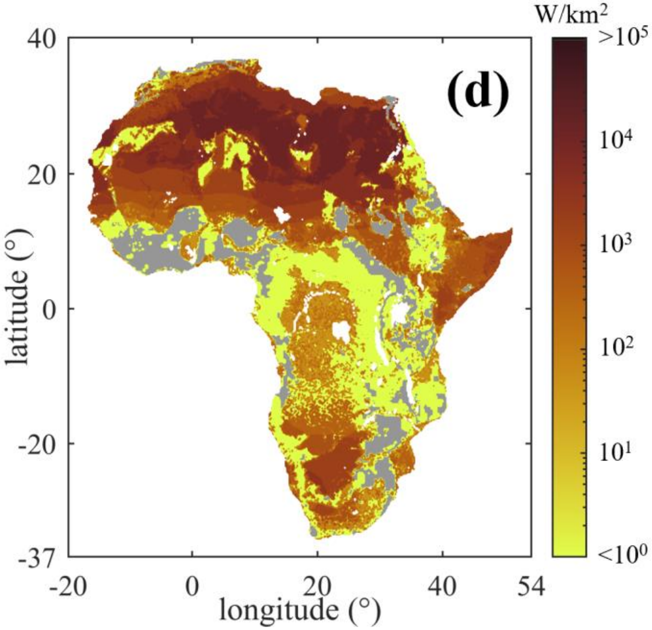
Maximum installable power for groundwater development

25% of annual recharge / year



96 W/km² remaining

0.1% of storage / year

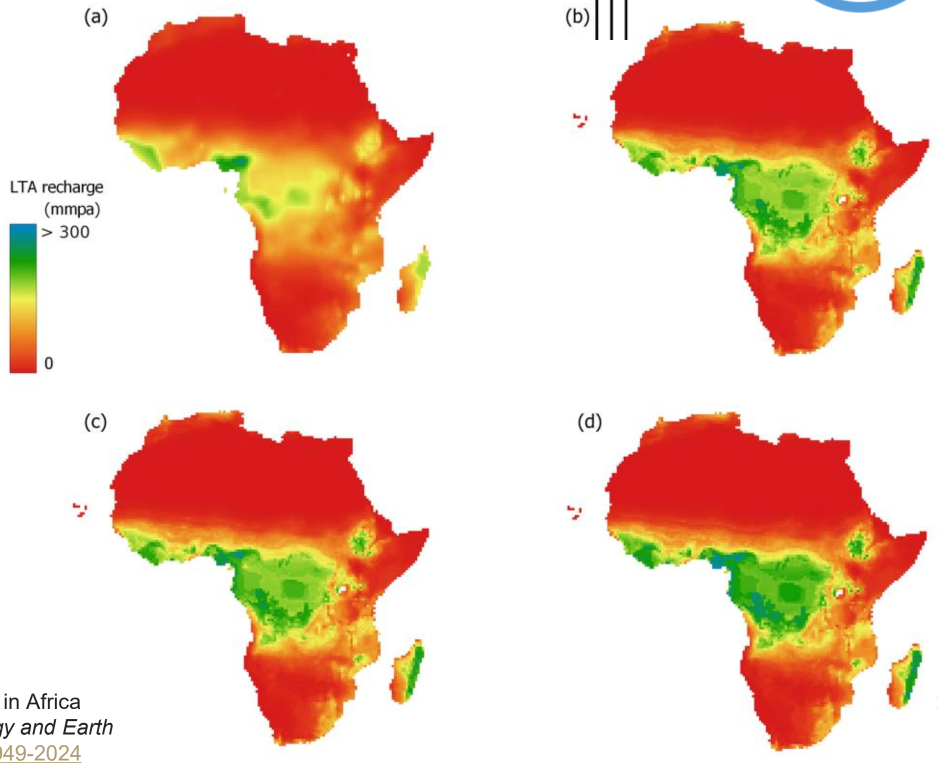


1980 W/km² remaining

Zuffinetti, G., et al. 2024. A method for estimating maximum safe installable power for groundwater extraction with application to Africa. *Science of the Total Environment.* https://doi.org/10.1016/j.scitotenv.2024.177062

Estimating continental recharge using random forests

- Linear Mixed Model vs **Random Forest** Model at 0.5°:
  - RF reproduces LMM
  - RF better captures high spatial variability => useful in humid regions where small scale variability is expected.

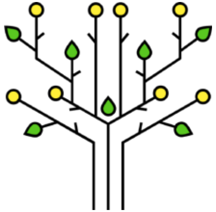


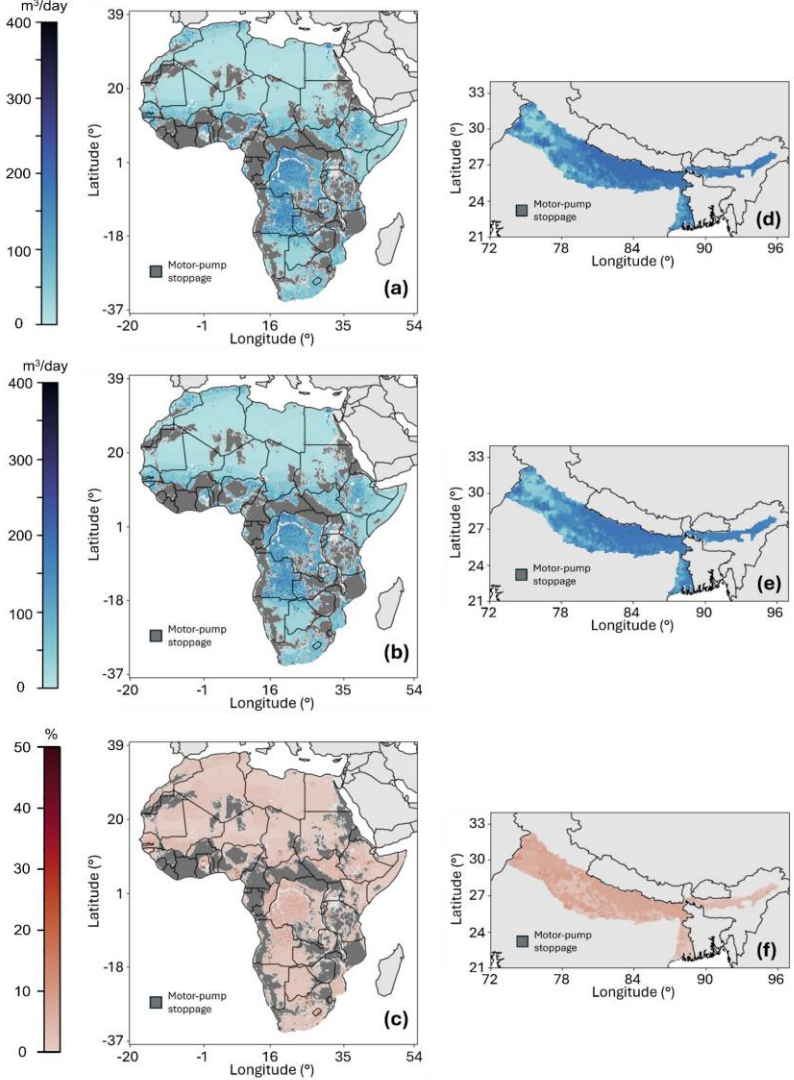
- Higher resolution maps produced at 0.1°, but limited input data means these models have high uncertainty => finer scale not automatically better

Pazola, Anna, et al. "High-resolution long-term average groundwater recharge in Africa estimated using random forest regression and residual interpolation." *Hydrology and Earth System Sciences* 28.13 (2024): 2949-2967. https://doi.org/10.5194/hess-28-2949-2024

Machine learning for speed at scale

- Large-scale physics-based model for solar water pumping
- Emulators decreased computation time by > 1500 (**Gradianet Boosting Classifier** and **Random Forest Regressor**)
- Emulators achieved high accuracy: R2 ≥ 0.99, NRMSE ≤ 5%.
- Pre-size and position solar water pumps at regional, national or even continental scale:





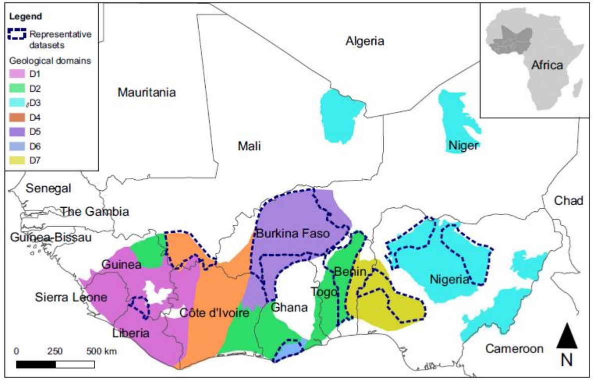
Zuffinetti, Guillaume, et al. "Large-scale modeling of solar water pumps using machine learning." *Applied Energy* 406 (2026): 127268.
https://doi.org/10.1016/j.apenergy.2025.127268

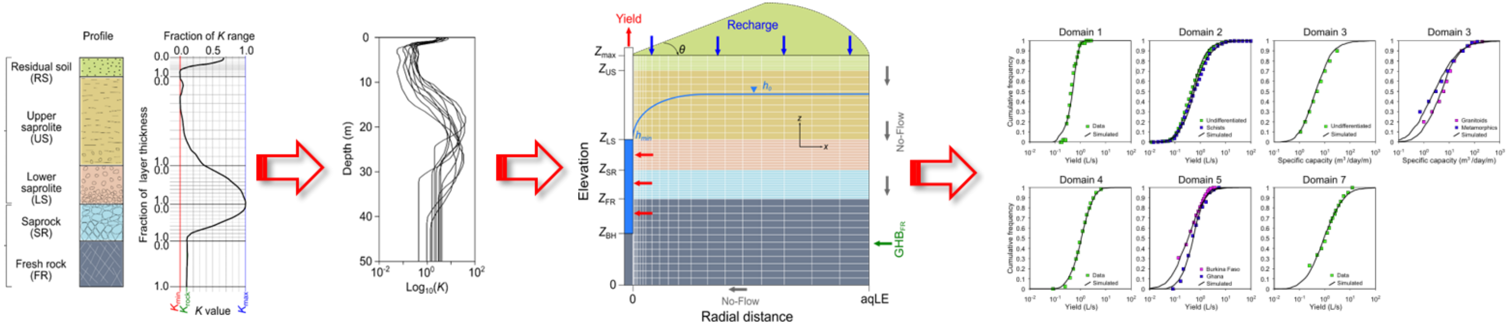
Groundwater at the regional scale

Mapping the basement in West Africa

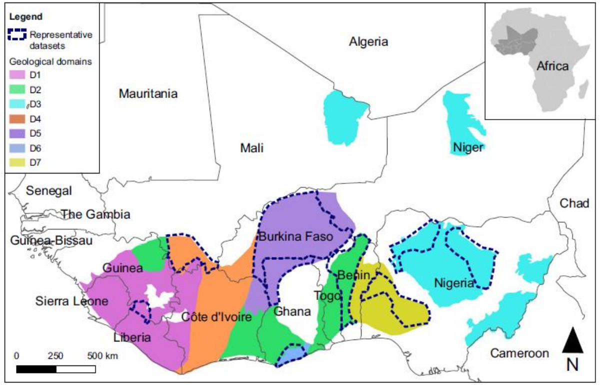
Bianchi, Marco, et al. "Assessing regional variation in yield from weathered basement aquifers in West Africa and modelling their future groundwater development and sustainability." *Hydrogeology Journal* 31.2 (2023): 257-274. https://doi.or/10.1007/s10040-023-02594-w

- Defining basement domains based on the nature of weathering and fracturing
- Geological, climatic and landscape controls
- Stochastic parameterisation of typologies (thickness, Kmax, Kmin) in MODFLOW
- Define yield probabilities
- Input data: High quality geological data but localised
- Validation data: low quality yield data but widespread

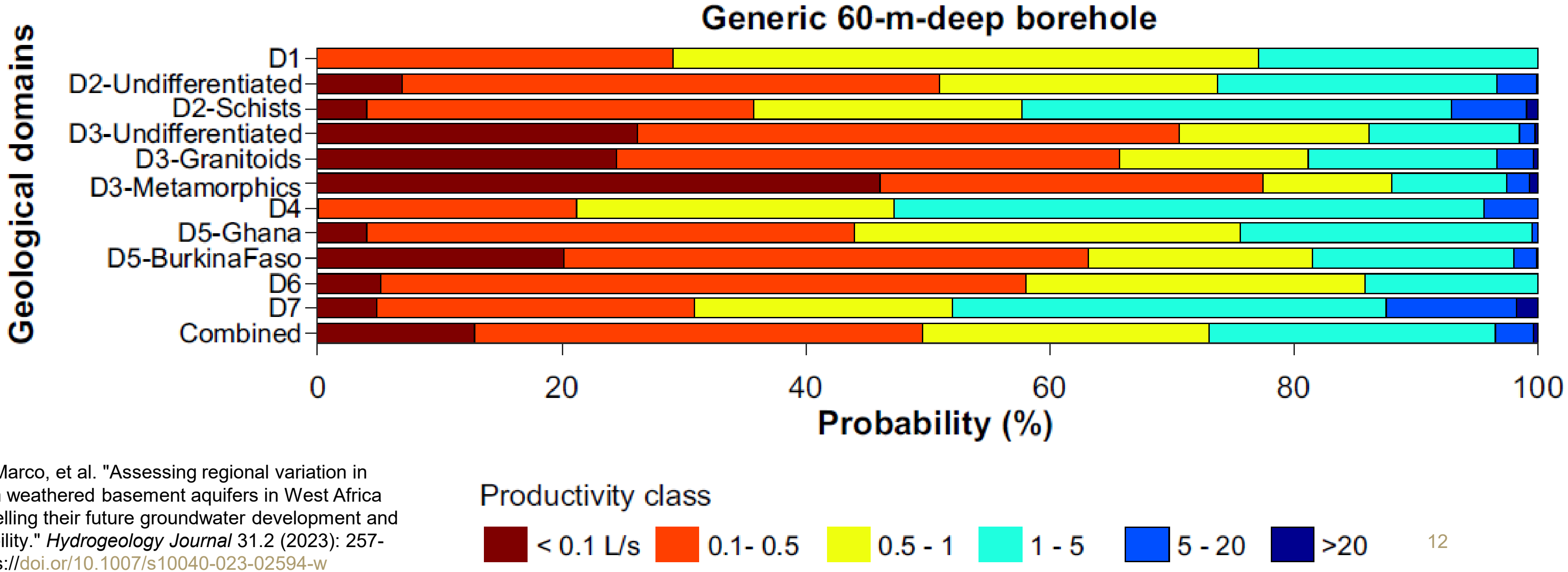




Mapping the basement in West Africa



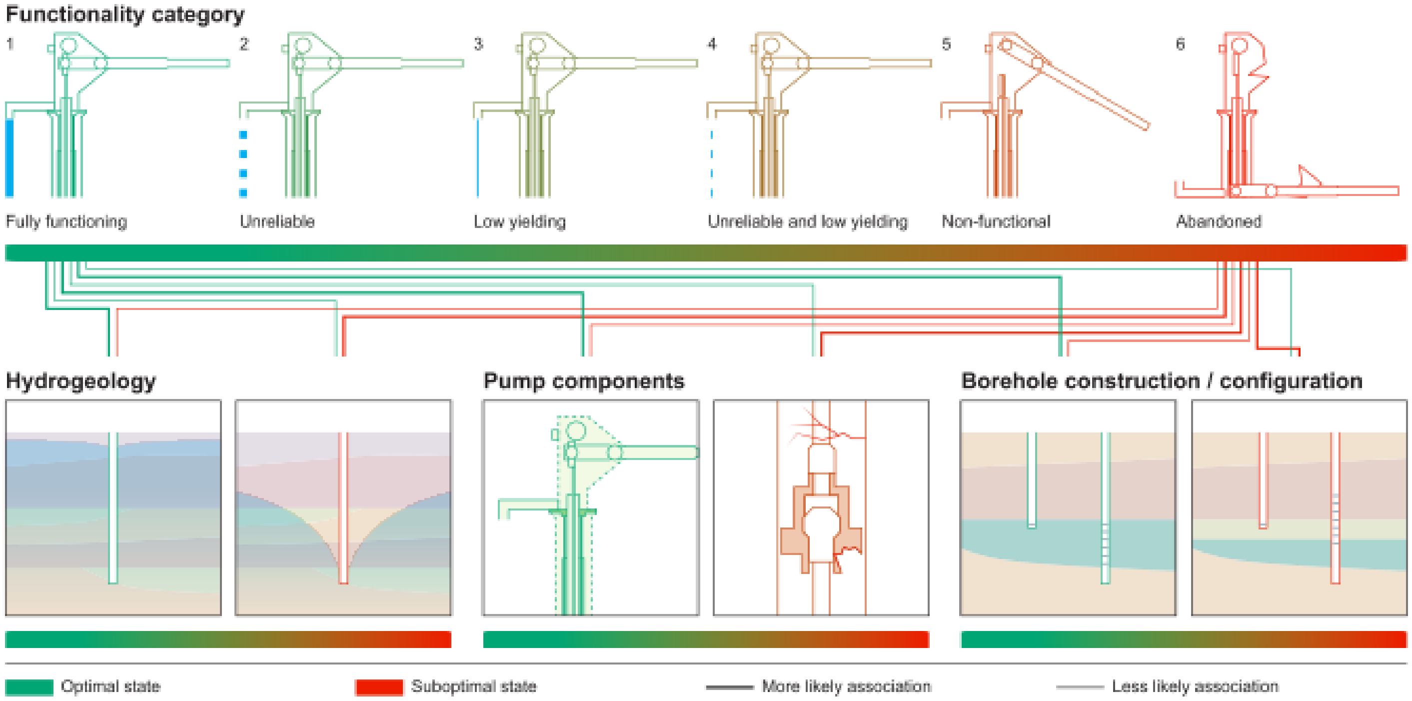
- Significant regional differences
- Over all 25% can sustain > 1 L/s
- Several have >5% of > 5 L/s



Bianchi, Marco, et al. "Assessing regional variation in yield from weathered basement aquifers in West Africa and modelling their future groundwater development and sustainability." *Hydrogeology Journal* 31.2 (2023): 257-274. https://doi.or/10.1007/s10040-023-02594-w

Groundwater at the borehole scale

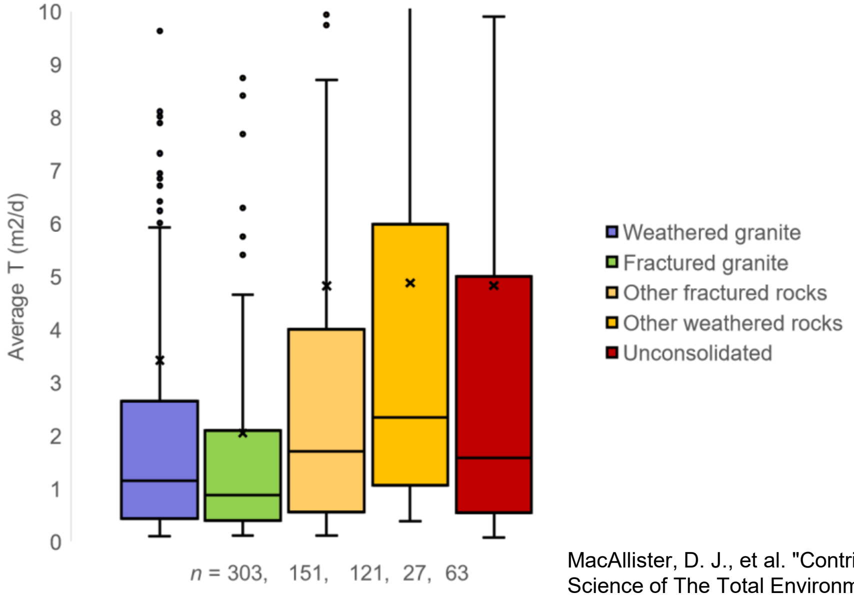
What determines handpump borehole functionality?

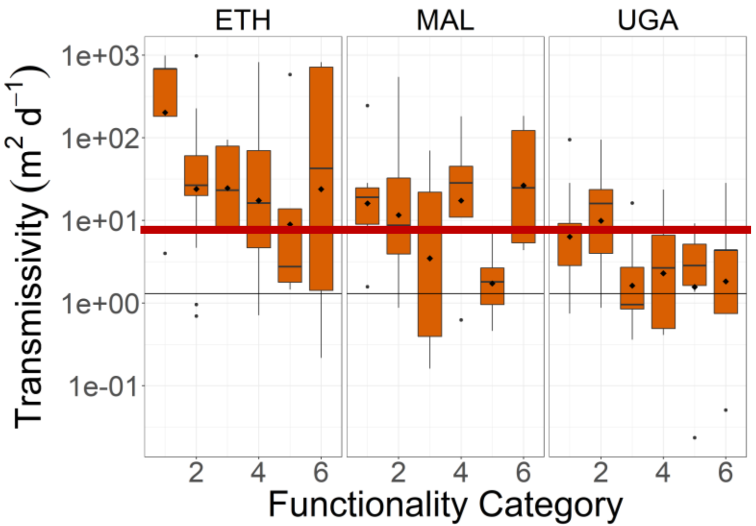


MacAllister, D. J., et al. "Contribution of physical factors to handpump borehole functionality in Africa." Science of The Total Environment 851 (2022): 158343. https://doi.org/10.1016/j.scitotenv.2022.158343

Transmissivity

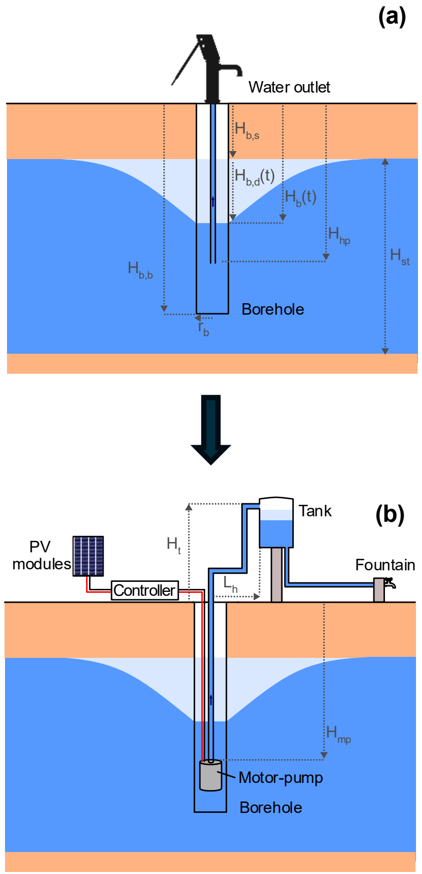
- Transmissivity derived from non-standard pumping tests using BGSPT (generalised well function numerically approximated using Bessel functions).
- BGSPT good for low-quality pumping test data.
- Transmissivity results have implications for larger reticulated systems (e.g. solar pumps).

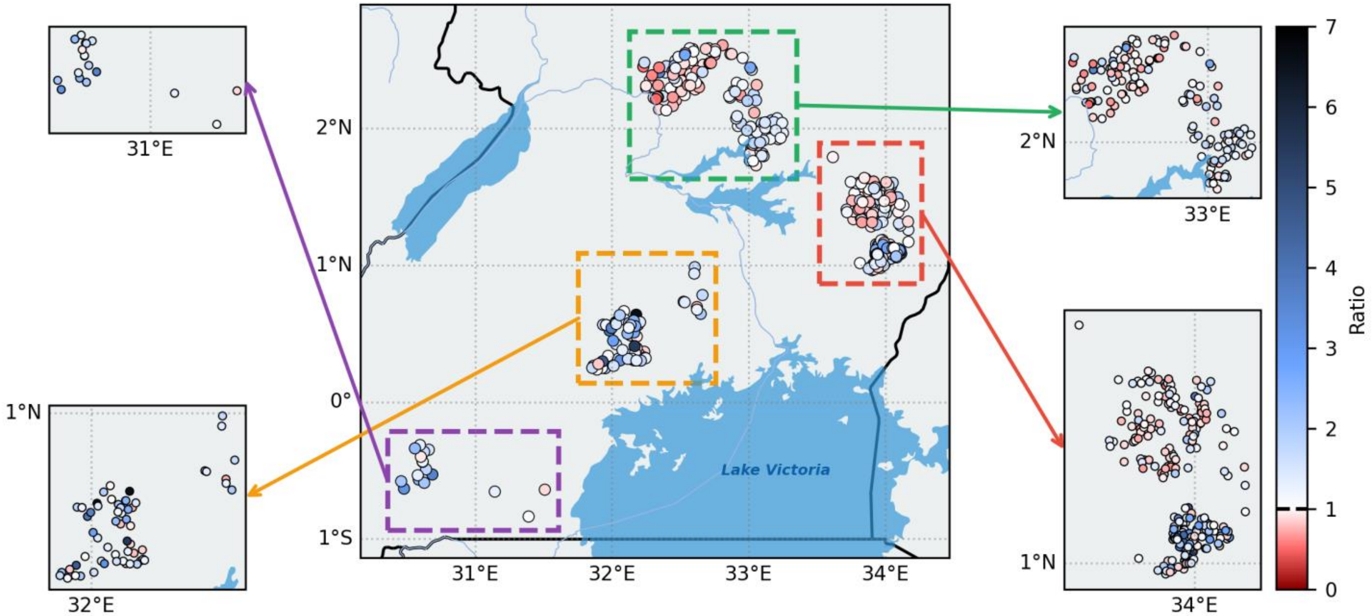




MacAllister, D. J., et al. "Contribution of physical factors to handpump borehole functionality in Africa." Science of The Total Environment 851 (2022): 158343. https://doi.org/10.1016/j.scitotenv.2022.158343

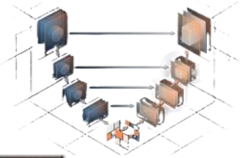
Hand-to-solar retrofit



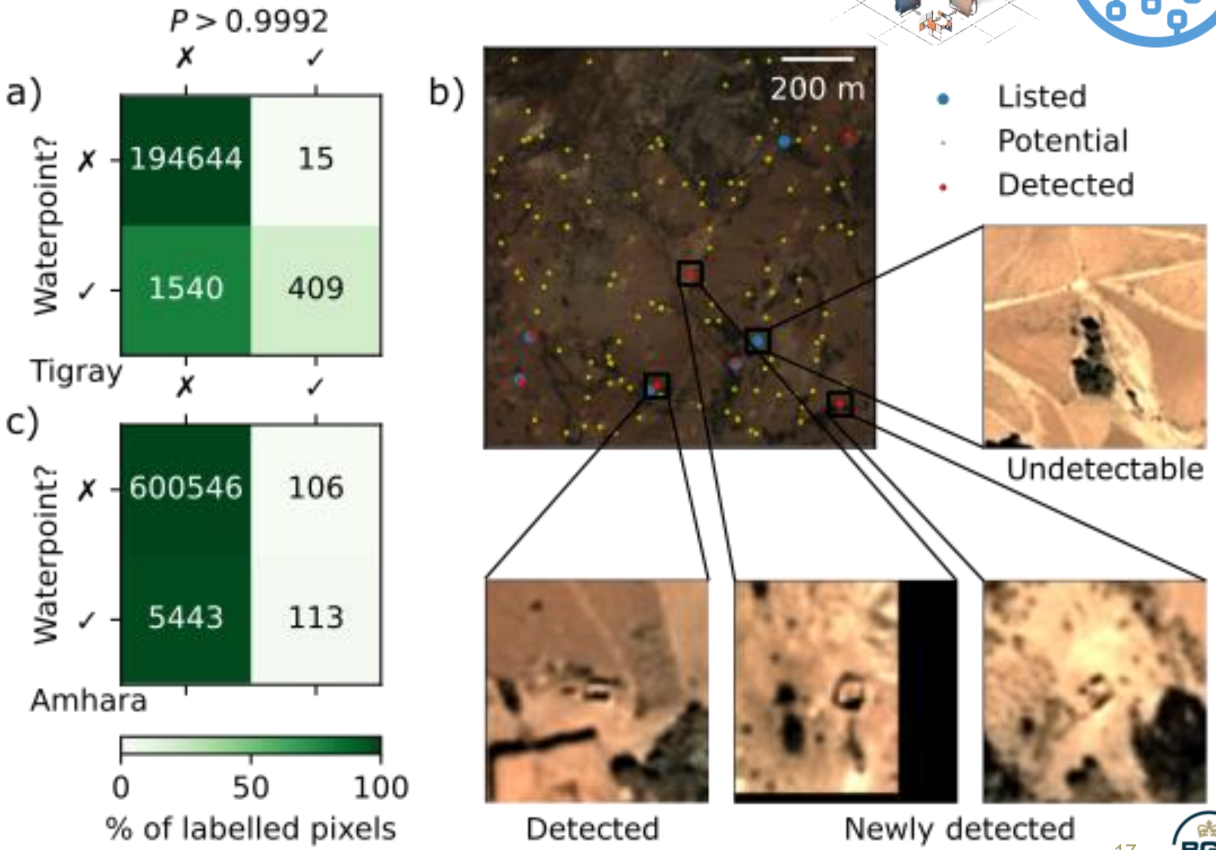


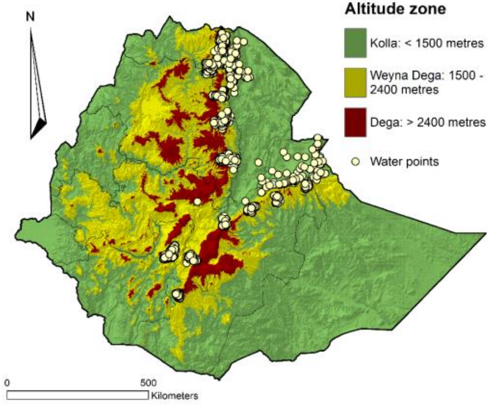
- Ratio > 1 solar produces more water and vice versa
- Retrofitting handpumps with solar pumps can increase daily supply at 64% of sites
- Retrofitting allows productive uses (e.g. small-scale irrigation)

Remote mapping of rural water sources



- **U-net** architecture
- Trained on data from Ethiopia and WPDX
- 83% prediction accuracy in Tigray



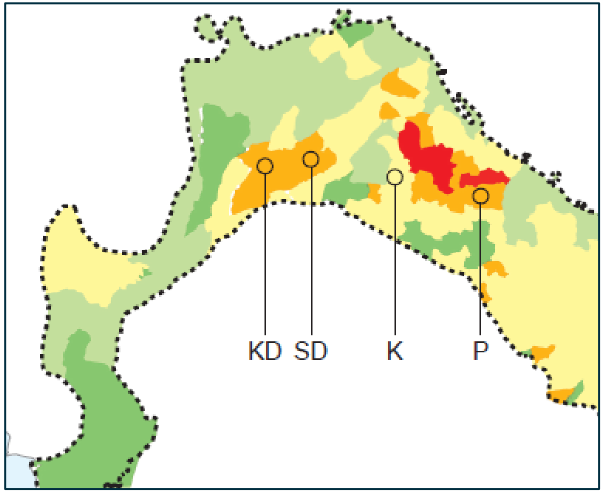


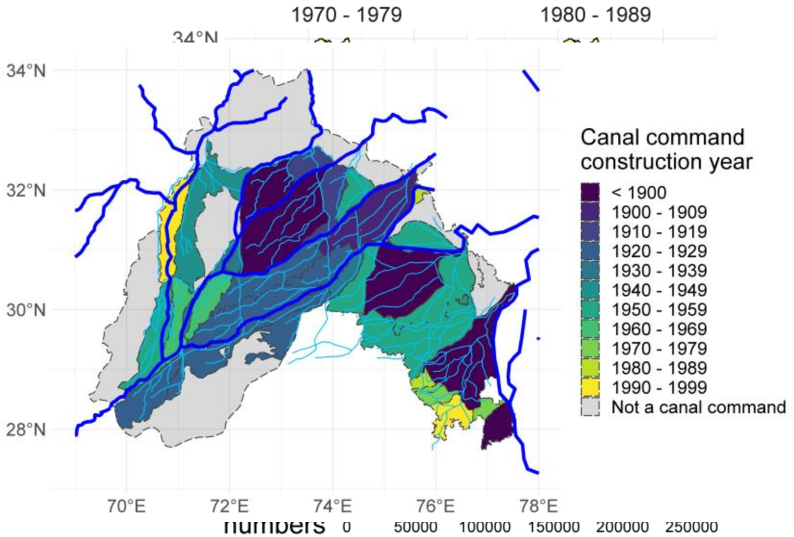
Groundwater in space and time

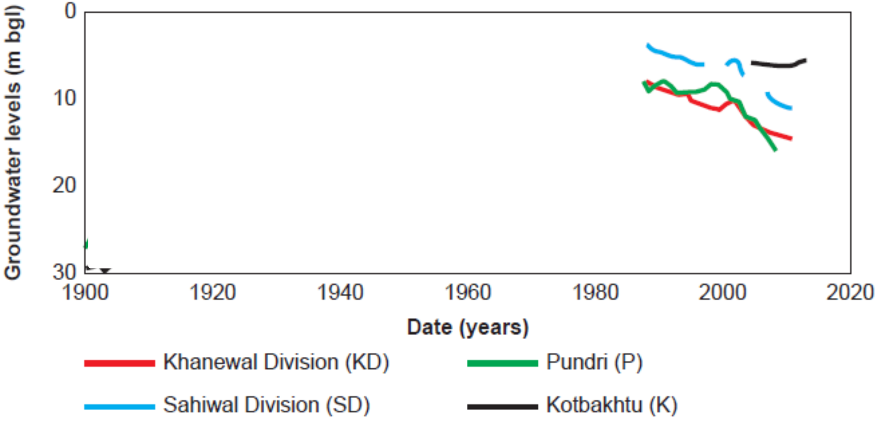
Groundwater over the long-term

A century of groundwater level change

- Groundwater depletion is a major concern in the 21st century.
- Long-history of water resource development in the Indus basin.



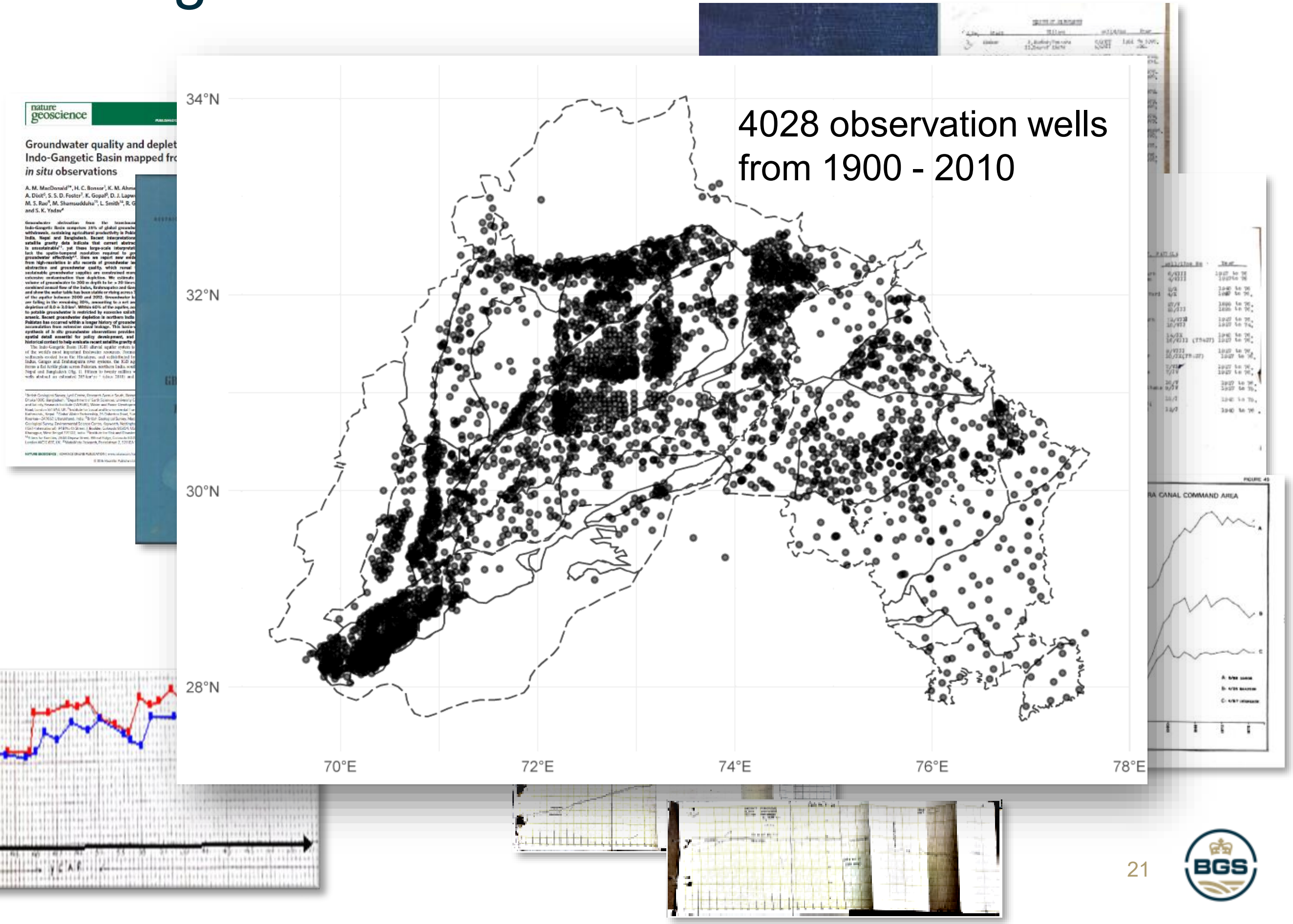


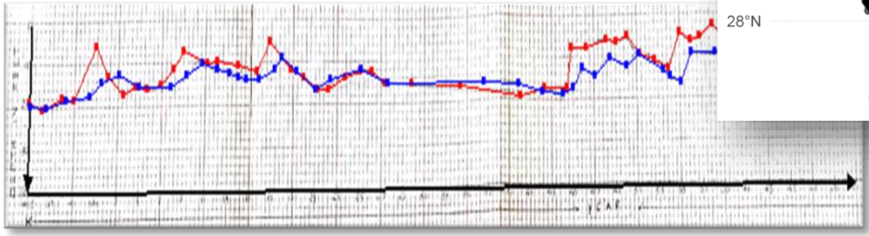


MacAllister, D. J., et al. "A century of groundwater accumulation in Pakistan and northwest India." Nature Geoscience 15.5 (2022): 390-396. https://doi.org/10.1038/s41561-022-00926-1

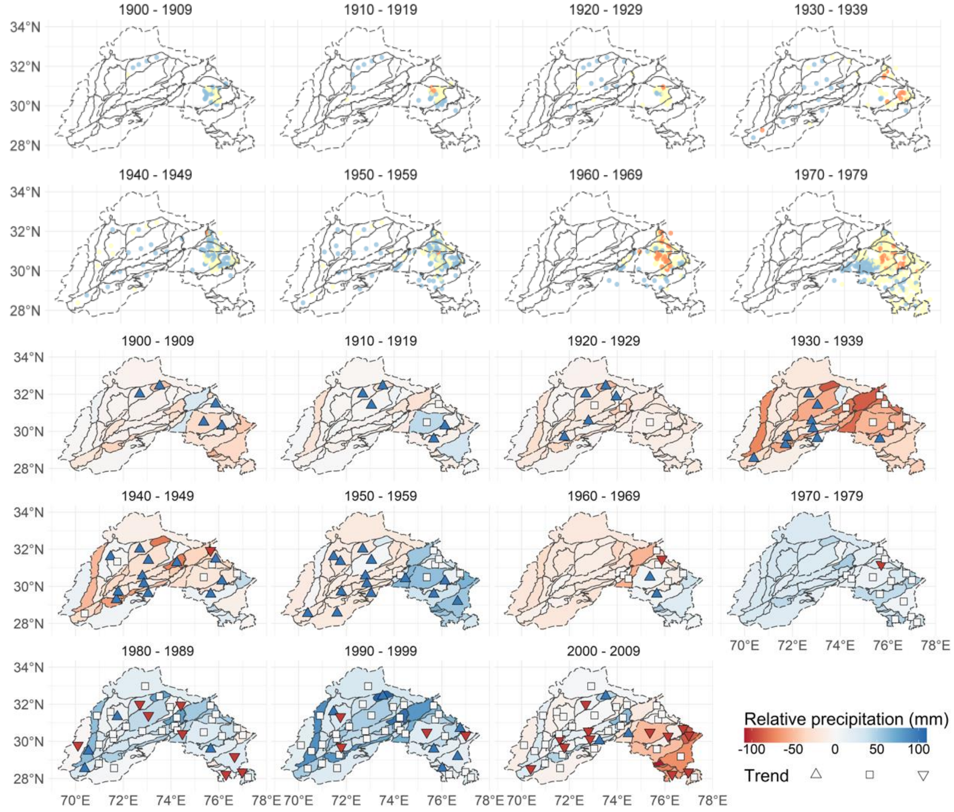
Compilation of historic groundwater data

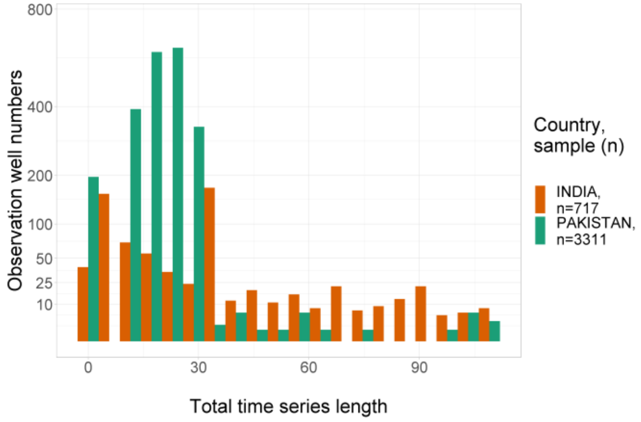
- > 100 scans:
  - old reports
  - books
  - government archives
  - digital and hard copy
  - Ex-director agricultural Punjab

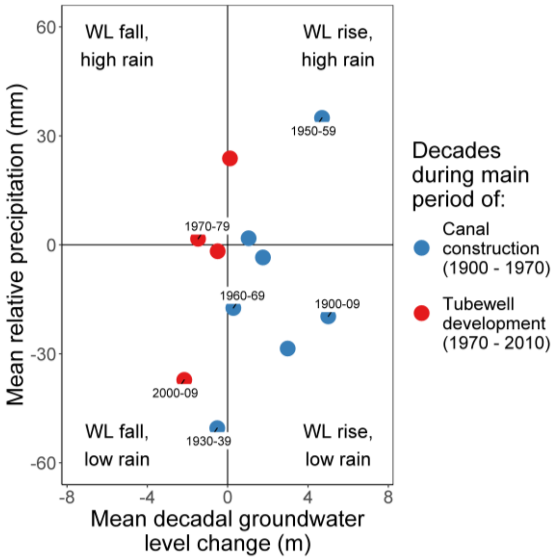




Groundwater level decadal trends



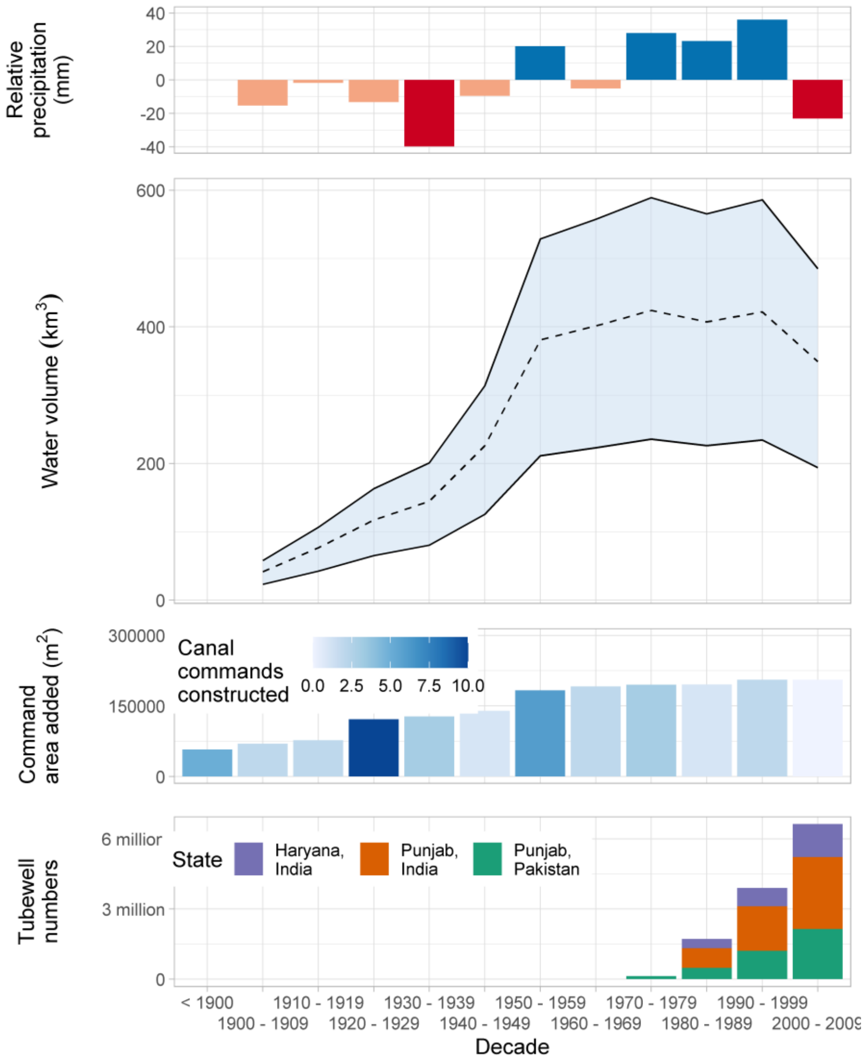




MacAllister, D. J., et al. "A century of groundwater accumulation in Pakistan and northwest India." Nature Geoscience 15.5 (2022): 390-396. https://doi.org/10.1038/s41561-022-00926-1

A century of accumulation

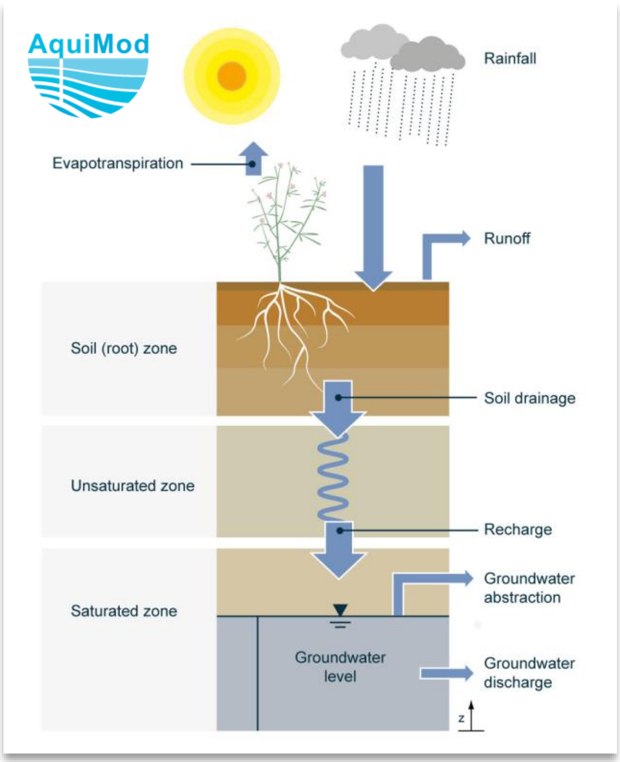
- 1900 – 1960:
  - Rainfall below period mean.
  - **125,000 km²** of canal area added.
  - Net groundwater accumulation: **350 km³**
- 1960 – 1990:
  - Rainfall above period mean.
  - Little additional canal area added.
  - Tubewell development began.
  - Groundwater stabilised.
- 1990 – 2010:
  - Rainfall below period mean.
  - Groundwater depletion: **75 km³**



MacAllister, D. J., et al. "A century of groundwater accumulation in Pakistan and northwest India." Nature Geoscience 15.5 (2022): 390-396. https://doi.org/10.1038/s41561-022-00926-1

Groundwater forecasting

Time series modelling with simple models



FLOOD FORECASTING CENTRE

A working partnership between

Environment Agency | Met Office

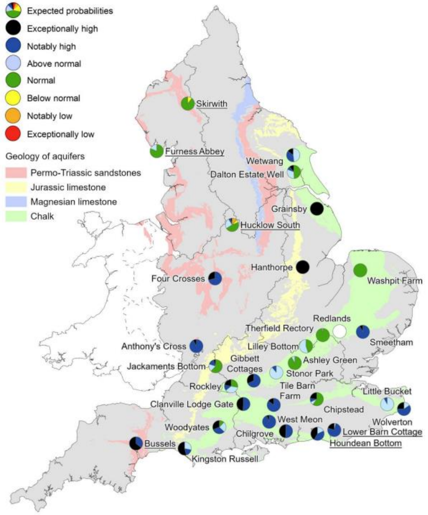
- 50 AquiMod models run daily
- 15 day forecasts
- Groundwater flood alert triggers

Research and analysis

Water situation: February 2026 summary

Updated 11 March 2026

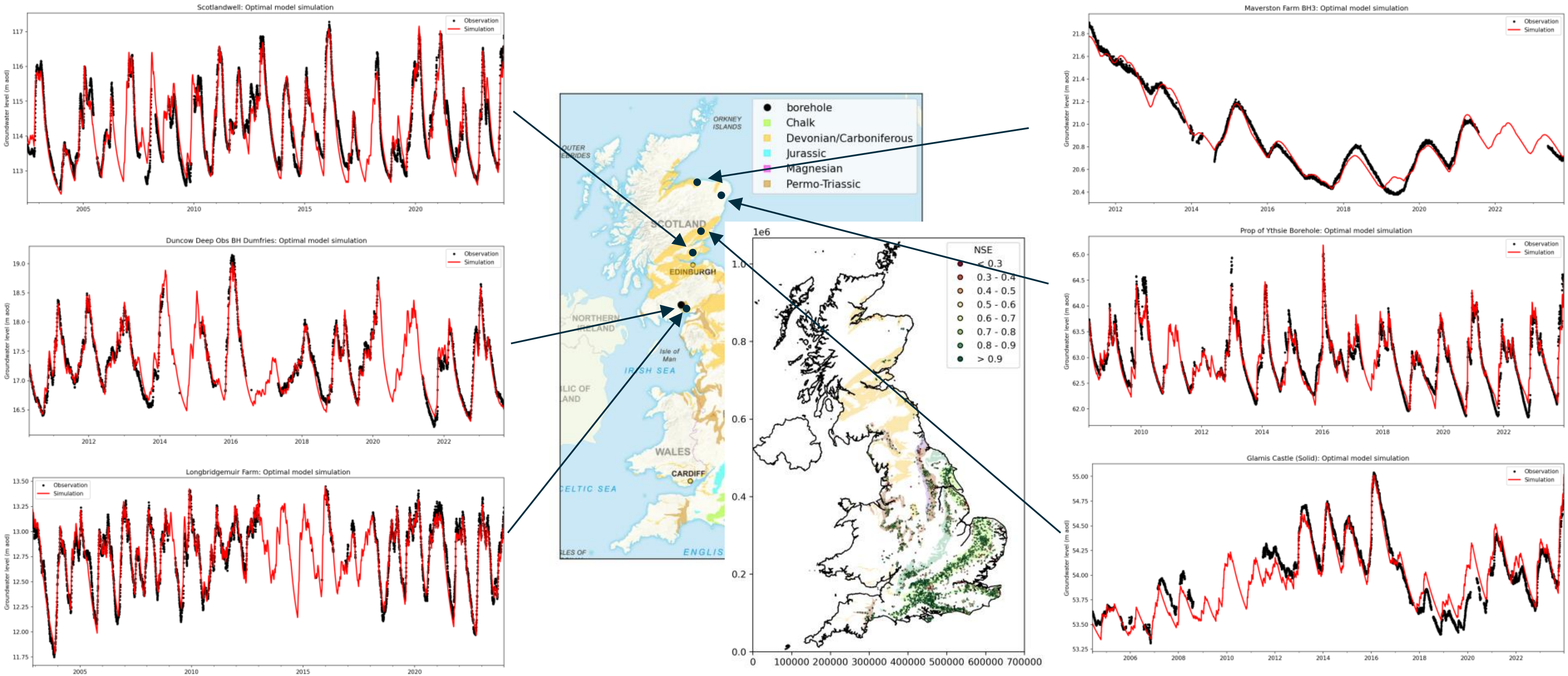
- AquiMod run monthly for 30 indicator boreholes
- 24-month probabilistic forecasts of GWLs



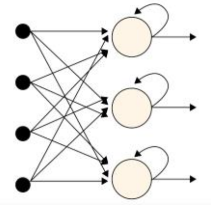
UK Hydrological Outlook

Time series modelling with simple models

- Models used to reconstruct GWLs to 1890 will now be used to run future climate projections to assess groundwater level response to future floods and droughts

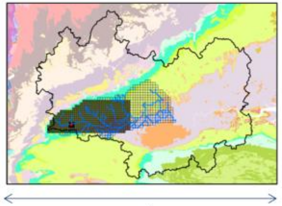


Forecasting with Machine Learning

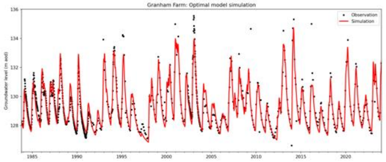


- Estimating the sustainable yield of groundwater supply sources in the future.
- Forecasting the performance of groundwater supply sources during a drought.

Regional Aquifer



200 km

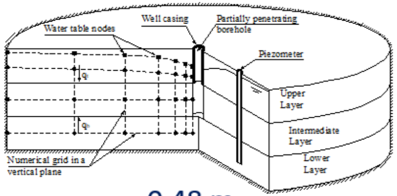


AquiMod

Boundary conditions

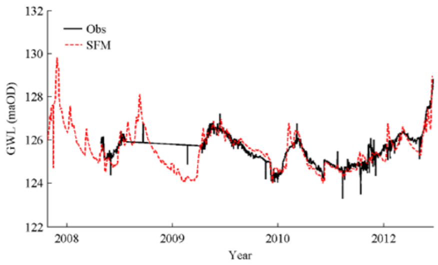
GWL + recharge

SPIDERR



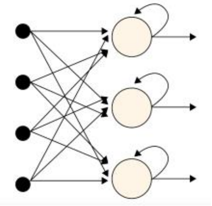
0.48 m

750 m

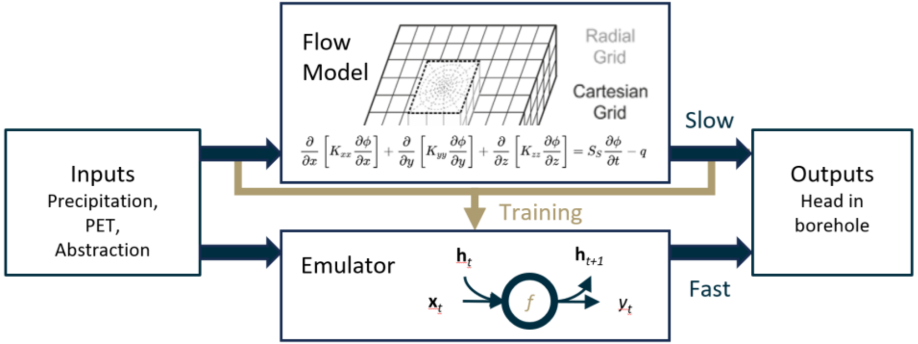


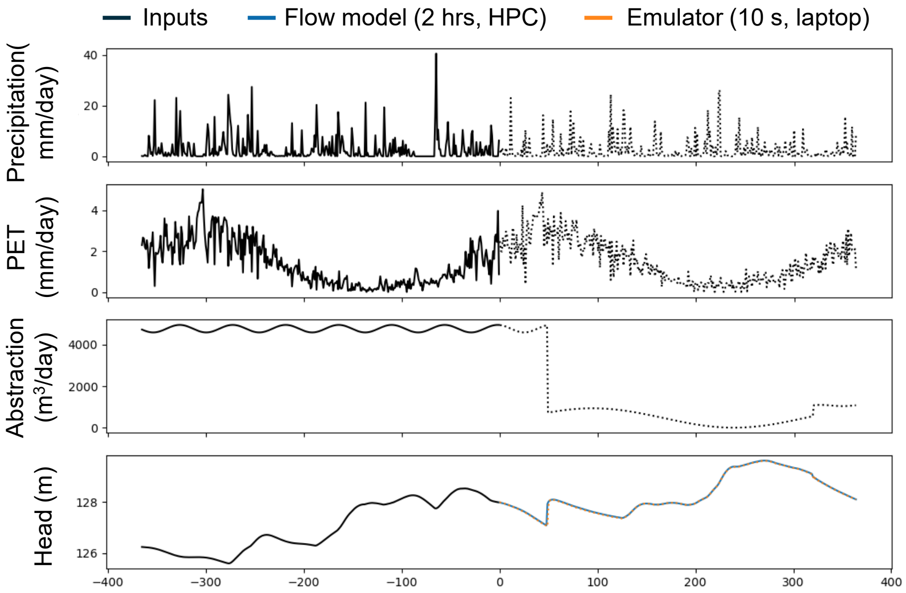
Upton, K. A., et al. "An integrated modelling approach for assessing the effect of multiscale complexity on groundwater source yields." *Journal of Hydrology* 588 (2020): 125113. https://doi.org/10.1016/j.jhydrol.2020.125113

Forecasting with Machine Learning



- Recurrent neural network
- Emulator trained on input-output relationship of SPIDERR-AquiMod

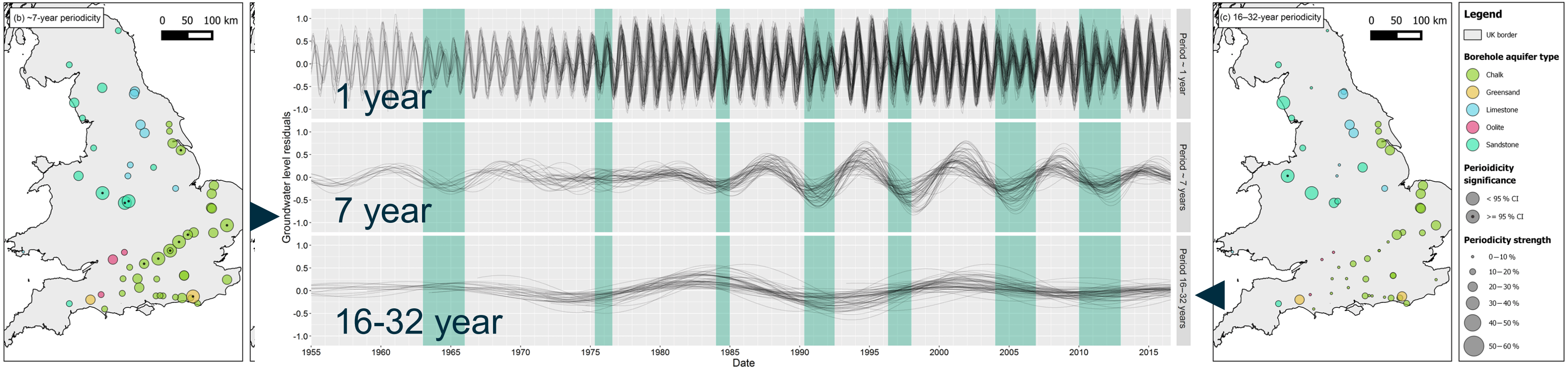




Forecasting groundwater drought?

Groundwater drought prediction?

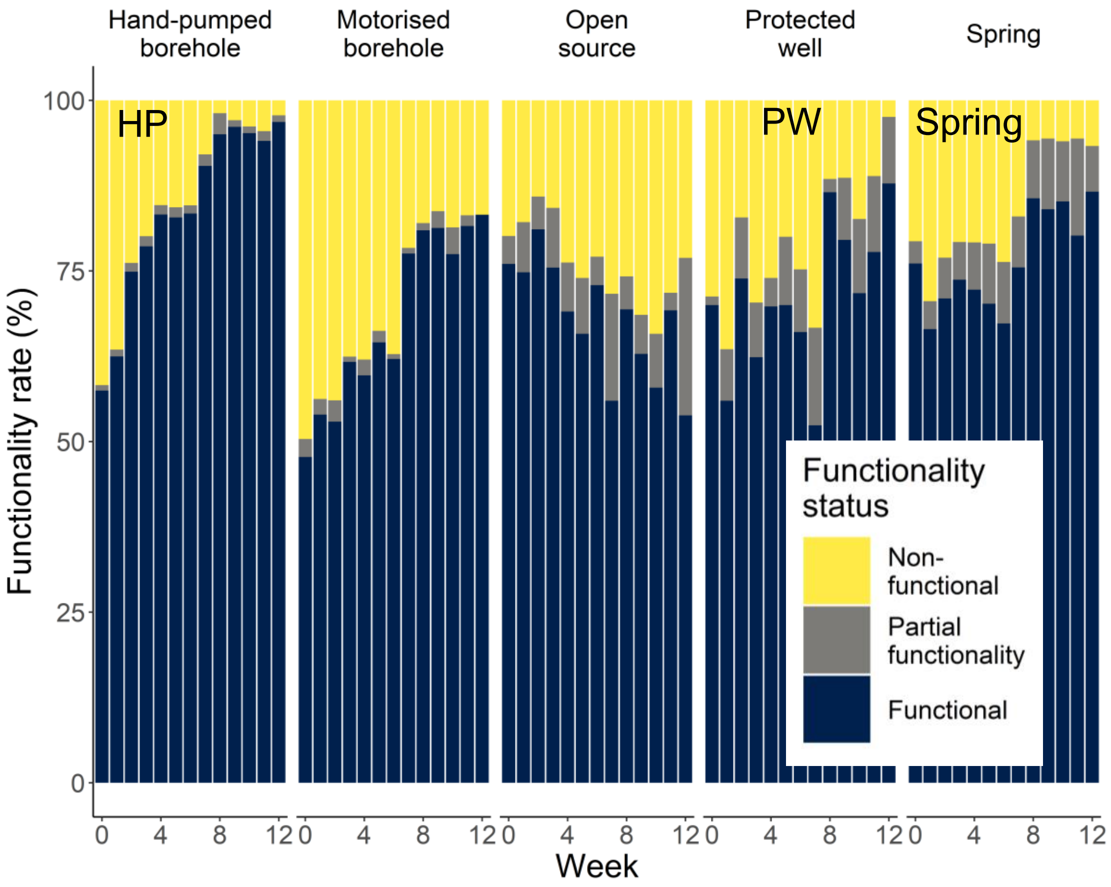
- Multi-annual variability in groundwater levels assessed using monthly data and time series from 21 to 181 years (53 years average) => link GW to climate surrogates.
- 40% of storage variability was attributed to the North Atlantic Oscillation.
- UK droughts coincide with minimum phase in 7 year NAO groundwater cycles.
- NAO is modulated by East Atlantic pattern, explaining the 16-32 year periodicity.

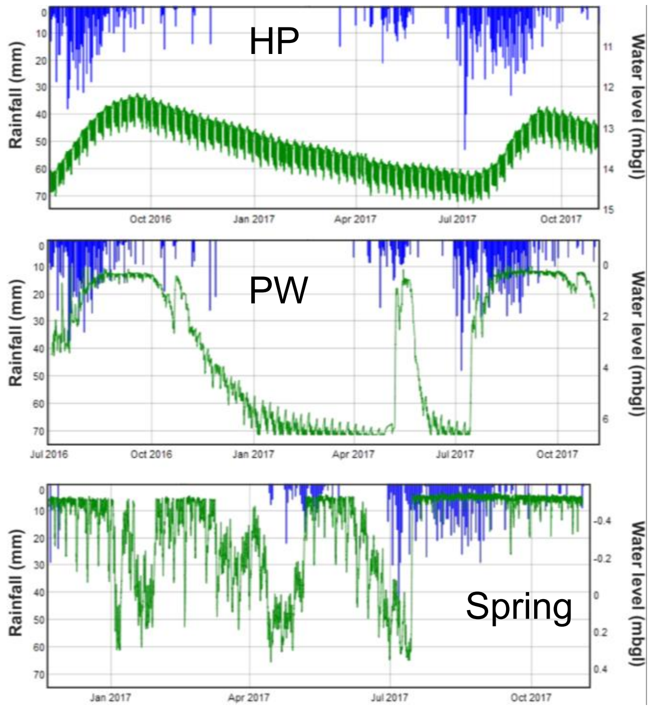


Rust, William, et al. "Understanding the potential of climate teleconnections to project future groundwater drought." *Hydrology and Earth System Sciences* 23.8 (2019): 3233-3245. https://doi.org/10.5194/hess-23-3233-2019

Groundwater access during drought

Functionality of water source types during drought

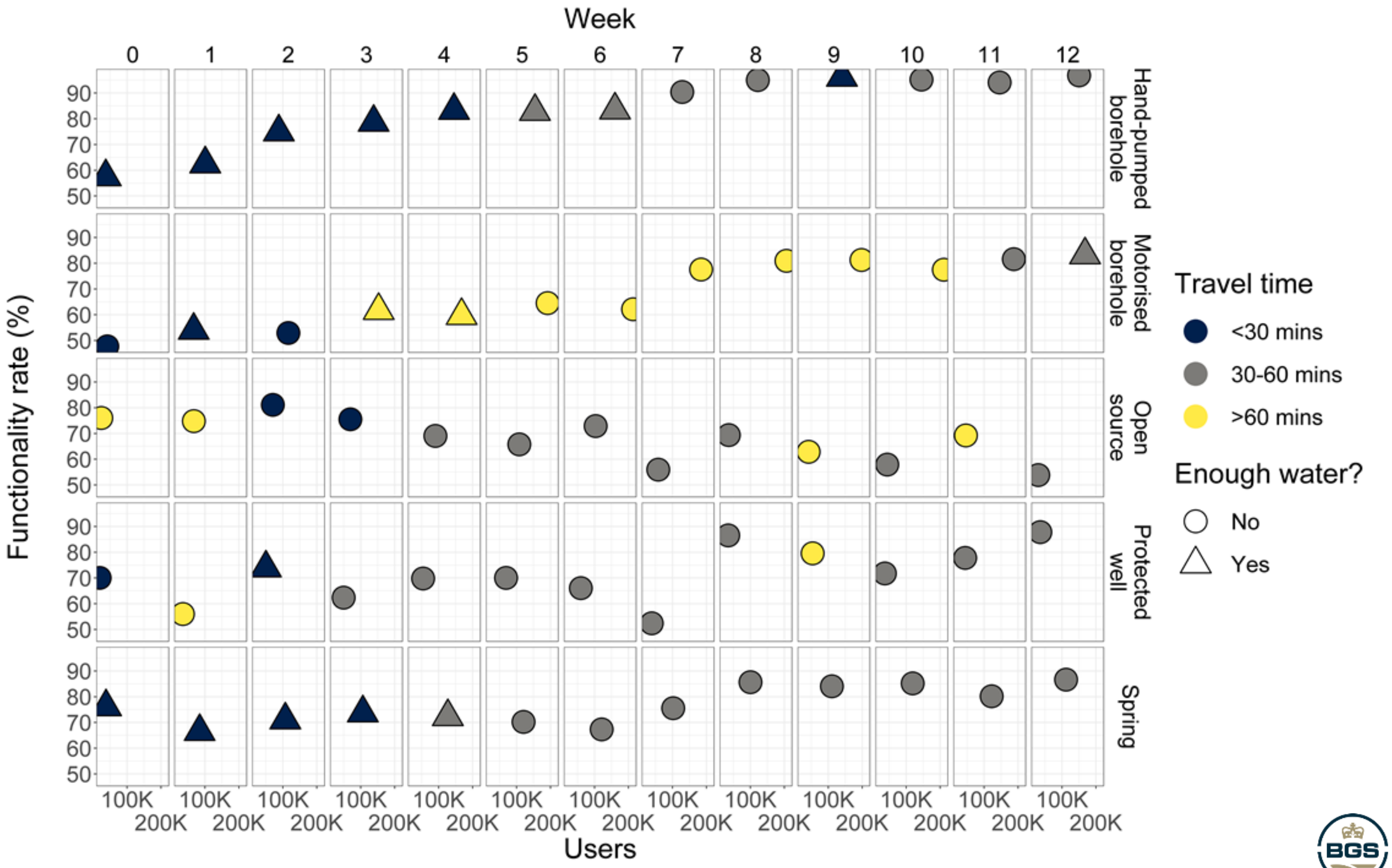




MacAllister, D.J., et al. Comparative performance of rural water supplies during drought. Nat Commun 11, 1099 (2020). https://doi.org/10.1038/s41467-020-14839-3

Summary of temporal impact of drought

- **Handpumps:**
  - Most accessible.
  - Quickest functionality increase (O&M).
  - Highest functionality.
  - Adequate water.
  - Serve 100,000s of people.

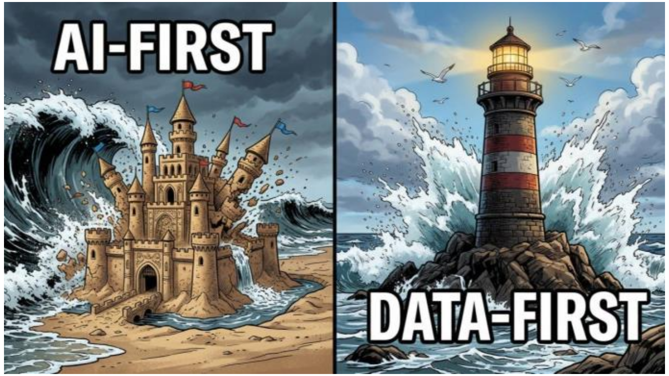


MacAllister, D.J., et al. Comparative performance of rural water supplies during drought. Nat Commun 11, 1099 (2020). https://doi.org/10.1038/s41467-020-14839-3

Summary and conclusions

Summary

- **Scale matters** and **scale depends** on the **questions** being asked and the **data** available.
- Putting in the effort to **collate spatial** (from national to continental) and **temporal** (from a few years to more than a century) data can produce novel and **impactful datasets**.
- The tools are important, but what really matters is our **hypothesis, frameworks for analysis and conceptualisation of groundwater systems**.
- With **imperfect data** and **well-designed analysis**, we can reveal **fundamental insights** into groundwater processes in **space and time**.
- AI has it's place, often increasing the speed and scale of the data analysis, but….
- …doing the **simple things can** produce as much insight, plus…
- …the outputs are **only as good as the data that goes in** (as always!).



British Geological Survey

THANK YOU

Any questions?
donmac@bgs.ac.uk