



HAL
open science

Geoscientific benchmarks for AI methods.

Nicolas Gilardi, Lisa Baulon, Hugo Breuillard, Musaab Khalid, Alexis Plunder, Alex Vella

► **To cite this version:**

Nicolas Gilardi, Lisa Baulon, Hugo Breuillard, Musaab Khalid, Alexis Plunder, et al.. Geoscientific benchmarks for AI methods.. NuTS First Workshop, May 2023, Villeurbanne, France. . hal-04104786

HAL Id: hal-04104786

<https://brgm.hal.science/hal-04104786v1>

Submitted on 24 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GEOSCIENTIFIC BENCHMARKS FOR AI METHODS

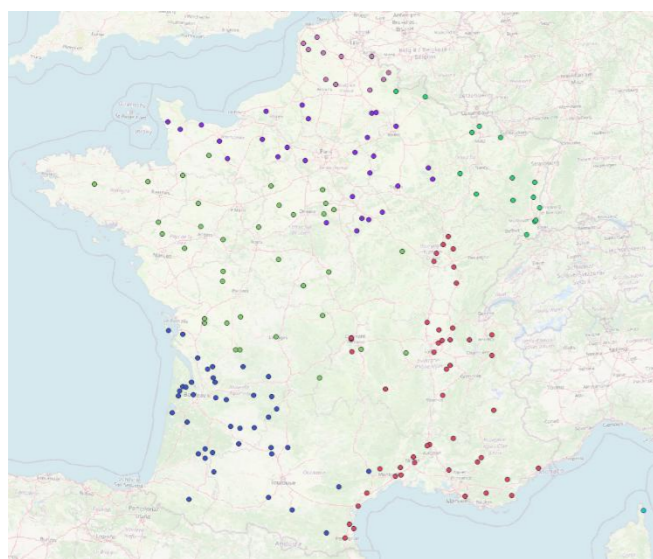
30/05/23
Référence
du poster

Authors

Nicolas Gilardi, BRGM - DNG; Lisa Baulon, BRGM - DEPA; Hugo Breuillard, BRGM - DNG;
Musaab Khalid, BRGM - DNG; Alexis Plünder, BRGM - DGR; Alex Vella, BRGM - DGR;

The BRGM (French National Geological Survey) has a wealth of geoscientific data. Through the AMIGO research project, we wish to make part of these data available to the scientific community in order to advance the development of AI methods adapted to geoscientific problems.

Ground Water Levels

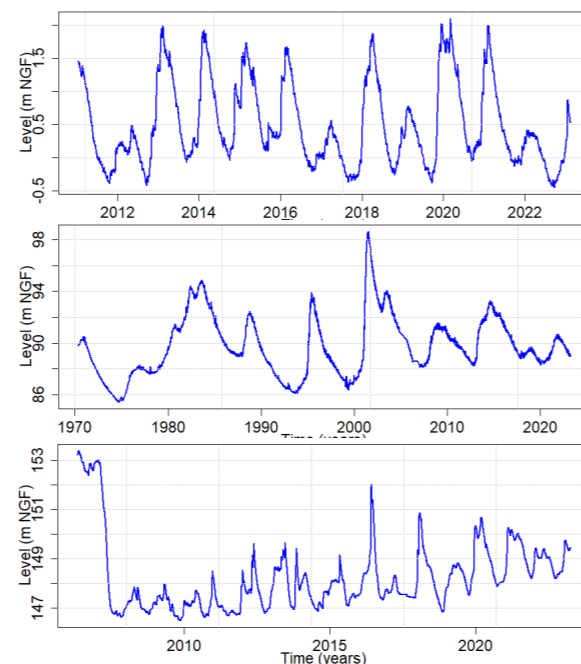


Dataset containing time series of water levels in France :

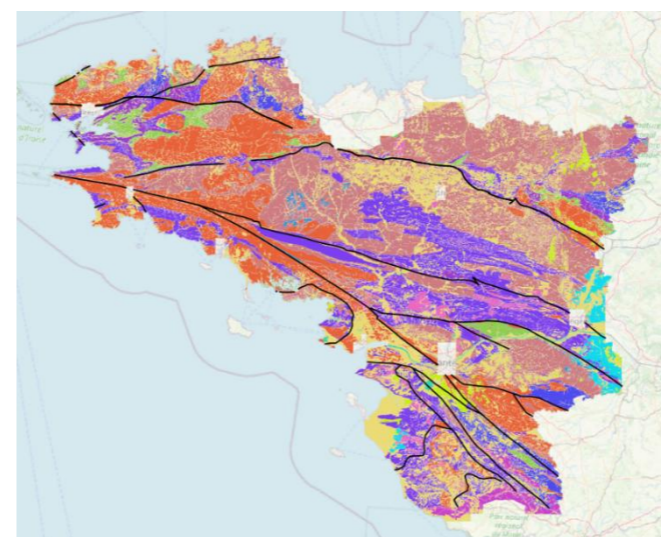
- 7 watersheds;
- 78 hydrologic areas;
- 180 boreholes, of which 37 are used for France's water table monitoring;
- 30 years long time series (on average).

Water level time series are grouped by periodicity, water table properties, human impact, etc.

The objective is to forecast the water level up to 1 year.



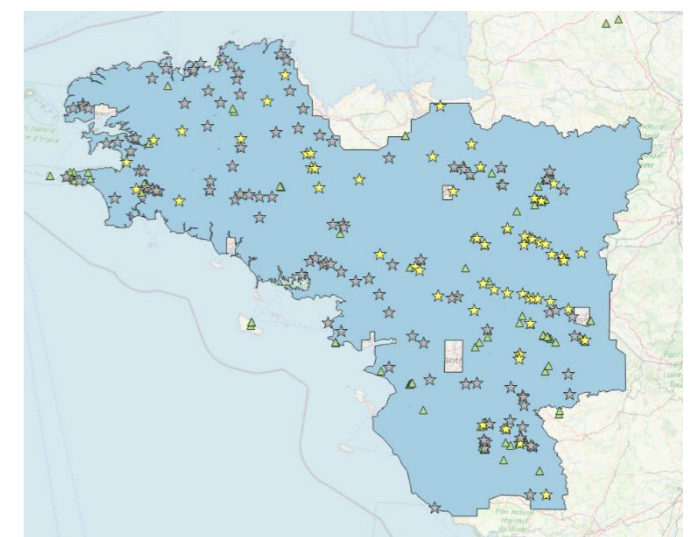
Mineral Potential Mapping



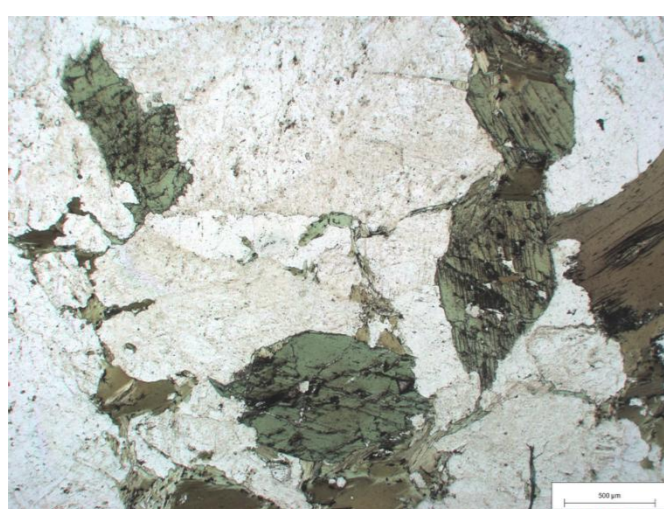
Dataset containing geology and mineral resources of the Armorican Massif:

- Vector data from 1/50000 geology map;
- surface lithology, main faults, rivers, etc.;
- gold and antimony occurrences and non-occurrences.

The objective is to map the probability of occurrence of gold in the Armorican Massif.



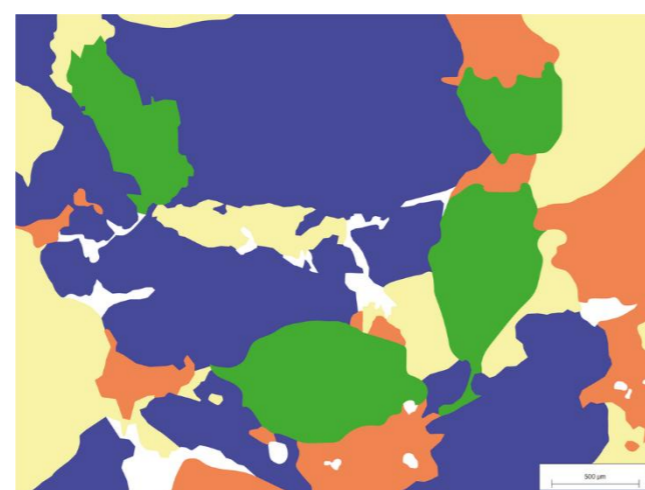
Thin Sections



Dataset containing thin sections of various rocks, with mineral type identified by experts.

The dataset is under construction. The plan is to identify a few tenth of minerals allowing automatic identification of hundreds of rocks.

The objective is to provide a proportion of each mineral phase in the image.



Summary

The BRGM is creating reference datasets for various geoscientific problems: ground water level prediction, mineral potential mapping, thin section analysis.

Each dataset will serve as a benchmark to compare solutions to solve these tasks, for instance using AI methods.

When ready, these benchmarks will be open to the geoscience and data science communities.

In parallel, BRGM is doing a constant survey of new AI methods, and is adapting these methods to geoscience tasks, and evaluate them using these benchmarks.