The application of machine learning methods to aggregate geochemistry predicts quarry source location: a case study from the Irish aggregate industry.

1. Background

Pyrite is often found in the sub-floor rock aggregate layer of Irish houses. Oxidation of pyrite forms H₂SO₄, which reacts with limestone aggregate and results in gypsum formation. Gypsum crystallization causes delamination and expansion of aggregate material.

12,500 tons thought to be affected: €500,000 to remediate each home. Overall cost to Irish Government: €562,500,000.

2. Analytical techniques

- **SEM-EDS**: In-situ major element analysis of the pyrite crystal was undertaken using a Tescan MIRA3 Field Emission Gun Scanning Electron Microscope coupled to an energy dispersive spectrometer (EDS) at the Materials Characterisation Laboratory in Trinity College Dublin.
- **IRM$: Isotope analyses were carried out at the University of Alberta, Canada, using a Micromass Quatro II TOF mass spectrometer operated in negative ion mode. Samples were analysed using a 25% diluted sample with 5% internal water at 120 °C. Acidity was 50% of acid at 25°C, and 100% of acid at 100°C.
- **LA-Q-ICPMS**: On-site major element analysis of the pyrite was carried out using a New Wave UP213 femtosecond laser coupled to a ThermoFisher Scientific iCAP-Qc mass spectrometer. Peaks were pumped using 200 Hz, 3 ns pulses with a 100% duty cycle.

3. Principle Component Analysis

PCA is an unsupervised statistical method that rotates and shuffles a data matrix into dimensions where axes (components) of greatest variability can be identified.

For PCA, if the p-pm values of the various trace minor and major elements are examined in isolation, and not in the context of the entire ablated volume, misleading determinations may result.

As a result, the data used in PCA has been transformed using centred log ratio (CLR) transformation. This includes a residual value (or "dilution") of non-essential elements to sum to unity (i.e. 1 million ppm). CLR is calculated as the log of the original value divided by the geometric mean of that element across the entire dataset. CLR can be calculated using ioGAS, CoDaPack or R.

4. Machine learning

- **The AQCM is a bespoke piece of machine learning software, tailored to the dataset generated through the geochemical analysis of both rich aggregate and pyrite. This software has been made open source and can be used for the prediction of any geochemical database. Further information regarding how to operate the AQCM and its potential uses can be found by referencing the QR code attached to this poster.**

- **Published Work**

Published in the Quarterly Journal of Engineering Geology and Hydrogeology, this paper illustrated that a clear geochemical correlation exists between pyrite from 4 different aggregate quarry sources located in Eastern Ireland. The results from this paper indicated that further research into the geochemistry of pyrite was needed in order to discover which technique provided the clearest compositional variation between the quarry sources.

Study area

- Pyrite to be analysed from 8 different quarry sources (A-H), located in east of Ireland in counties Meath (MH), Dublin (D), Kildare (K) and Wicklow (WW).