



A solution (data architecture) for handling time-series data - sensor data (4D), its visualisation and the questions around uncertainty of this data

Martin Nayembil (1) and Andrew Barkwith (2)

(1) United Kingdom (mln@bgs.ac.uk), (2) United Kingdom (andr3@bgs.ac.uk)

Geo-environmental research is increasingly in the age of data-driven research. It has become necessary to collect, store, integrate and visualise more subsurface data for environmental research. The information required to facilitate data-driven research is often characterised by its variability, volume, complexity and frequency. This has necessitated the development of suitable data workflows, hybrid data architectures, and multiple visualisation solutions to provide the proper context to scientists and to enable their understanding of the different trends that the data displays for their many scientific interpolations.

However this data, predominantly time-series (4D) acquired through sensors and being mostly telemetered, poses significant challenges/questions in quantifying the uncertainty of the data.

To validate the research answers including the data methodologies, the following open questions around uncertainty will need addressing, i.e. uncertainty generated from:

- the instruments used for data capture;
- the transfer process of the data often from remote locations through telemetry;
- the data processing techniques used for harmonising and integration from multiple sensor outlets;
- the approximations applied to visualize such data from various conversion factors to include units standardisation

The main question remains: How do we deal with the issues around uncertainty when it comes to the large and variable amounts of time-series data we collect, harmonise and visualise for the data-driven geo-environmental research that we undertake today?