British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

Gateway to the Earth

EGU2016:Information in Earth Sciences

A solution for handling time-series data

Visualisation and the questions around uncertainty of the data

Martin Nayembil & Andrew Barkwith

British Geological Survey

mln@bgs.ac.uk & andr3@bgs.ac.uk

Acknowledgements:

Phil Meldrum, Tony Myers, Graham Smith, Carl Watson, Simon Burden, James Sorensen

Introduction

- The BGS focuses on public-good science for government, and research to understand earth and environmental processes.

- Our future use of the subsurface: for groundwater, energy and waste disposal - depends on much greater understanding of subsurface processes.

- To better manage these activities safely and sustainably. Manage environmental change and be resilient to environmental hazards.

- Sub-surface monitoring: Instrumenting the Earth so that we understand geological processes in real time.

- Use these new data to model and forecast the geological processes that matter to lives and livelihoods, in cities and rural areas, in Britain and internationally.

- A step change in the way that we ingest, process and visualise-serve data to support the data driven research.

- Require more processing power for the ‘big data’ that will be generated.

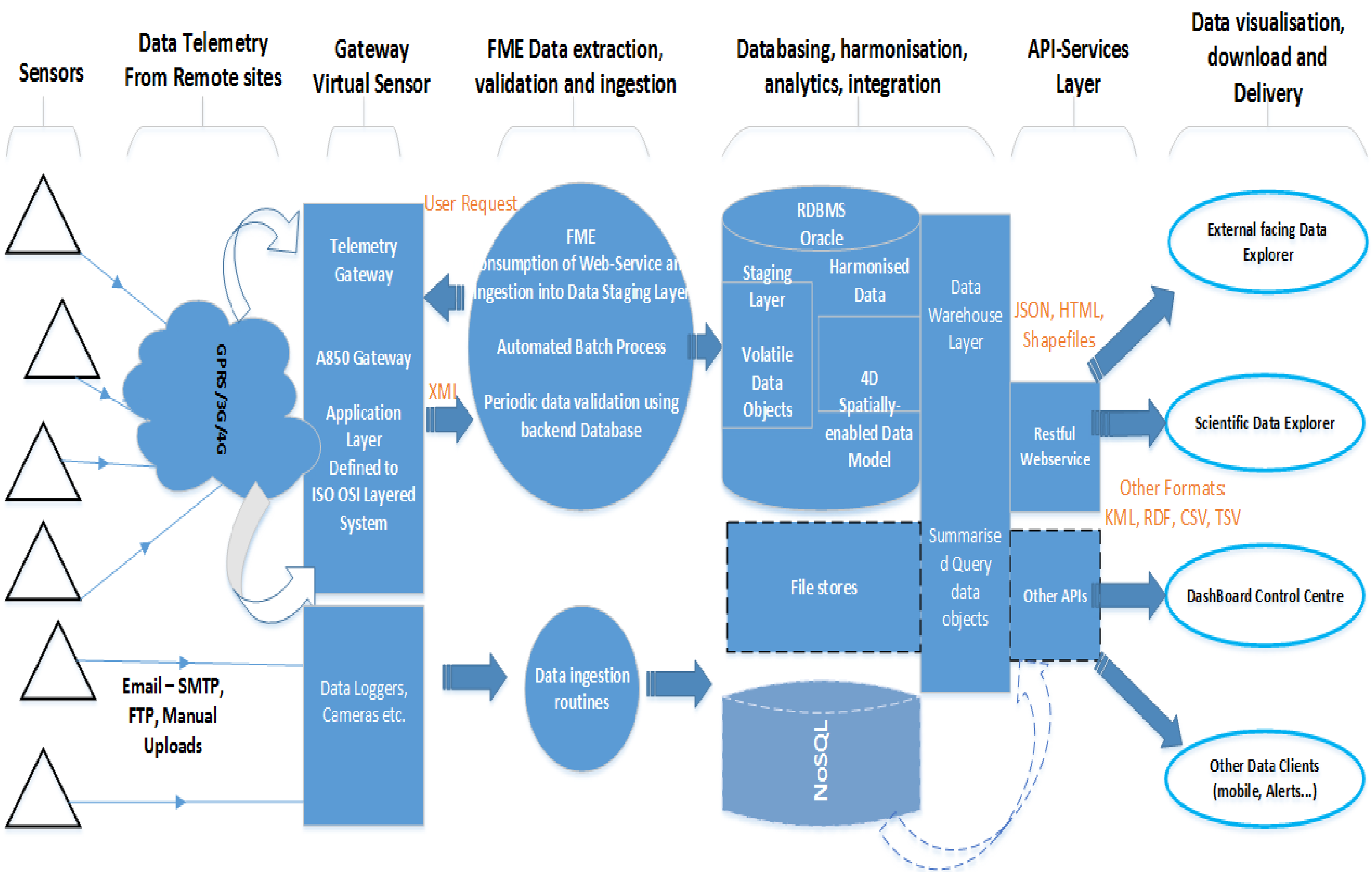
The Data Challenge

- To harvest and ingest data from any accessible "**Virtual SENSOR**" into a central data HUB for geo-environmental data
  - Near Real-time
  - Manual sensor data downloads (e.g. data loggers)
  - Historical Sensor data in various formats (e.g. csv, hsql)
  - Regenerated- Interpreted data outputs from telemetered sensor data delivered as file formats (e.g. via email)

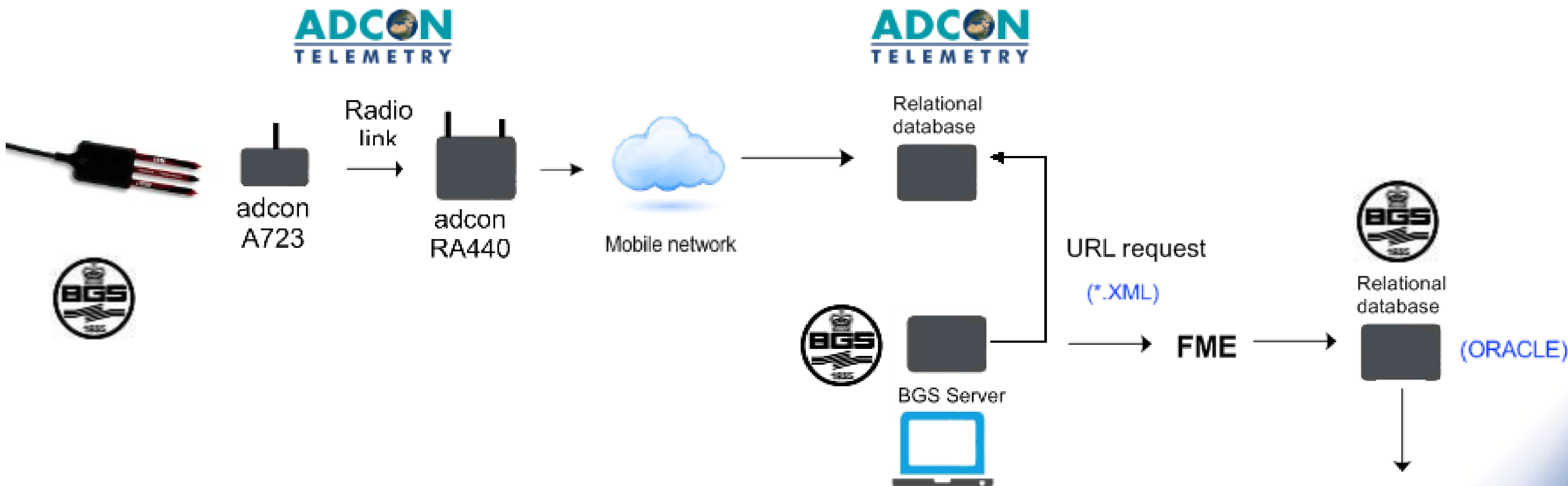
Component parts of the Architecture

- Open data architecture
- Data is harvested (e.g. near real-time – telemetry)
- Data staging – *"Virtual Sensor"*
- Data extraction, validation and ingestion
- Databasing, integration, harmonisation, validation
- Data Warehouse Layer
- API – Services layer
- Data visualisation, download and delivery layer

Sensor Data Architecture



Example: BGS deployed *Sensor – ADCON data logger and telemetry approach*



Web interface

Setup & Configure

Access, Model, Visualization

BGS Clients

(Cached JSON files)

(NOSQL)

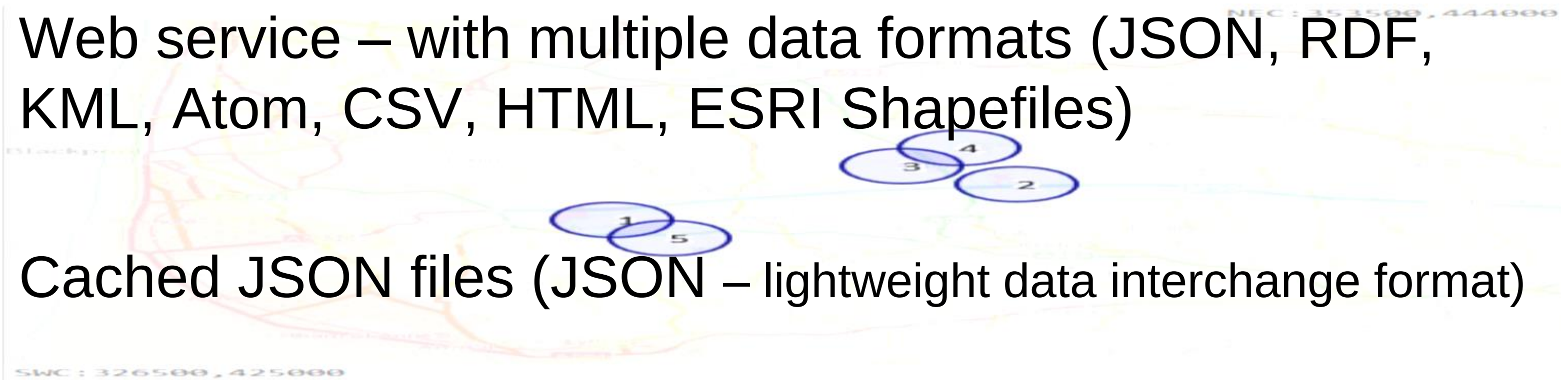
BGS: Sub-surface sensor Commercial data logger & telemetry

BGS: data access; *near-real time* data requests via webservices, email, FTP

BGS: data management FME transfer to relational database

Visualisation of time-series data

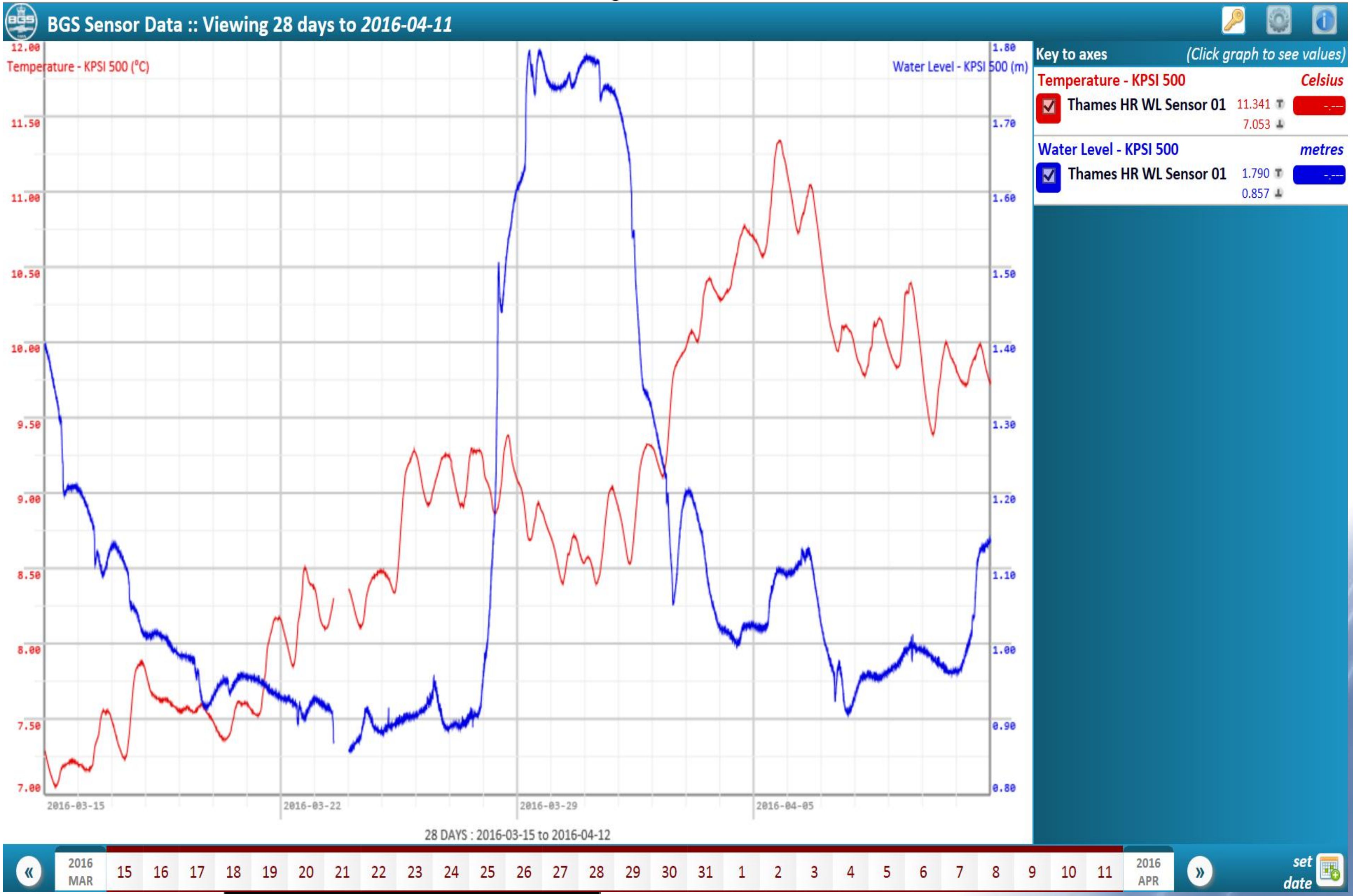
- Data warehouse layer – Query objects
- Web service – with multiple data formats (JSON, RDF, KML, Atom, CSV, HTML, ESRI Shapefiles)



- Cached JSON files (JSON – lightweight data interchange format)
- Time-series graphing clients for multiple parameters (e.g. water temperature, water level, soil moisture,

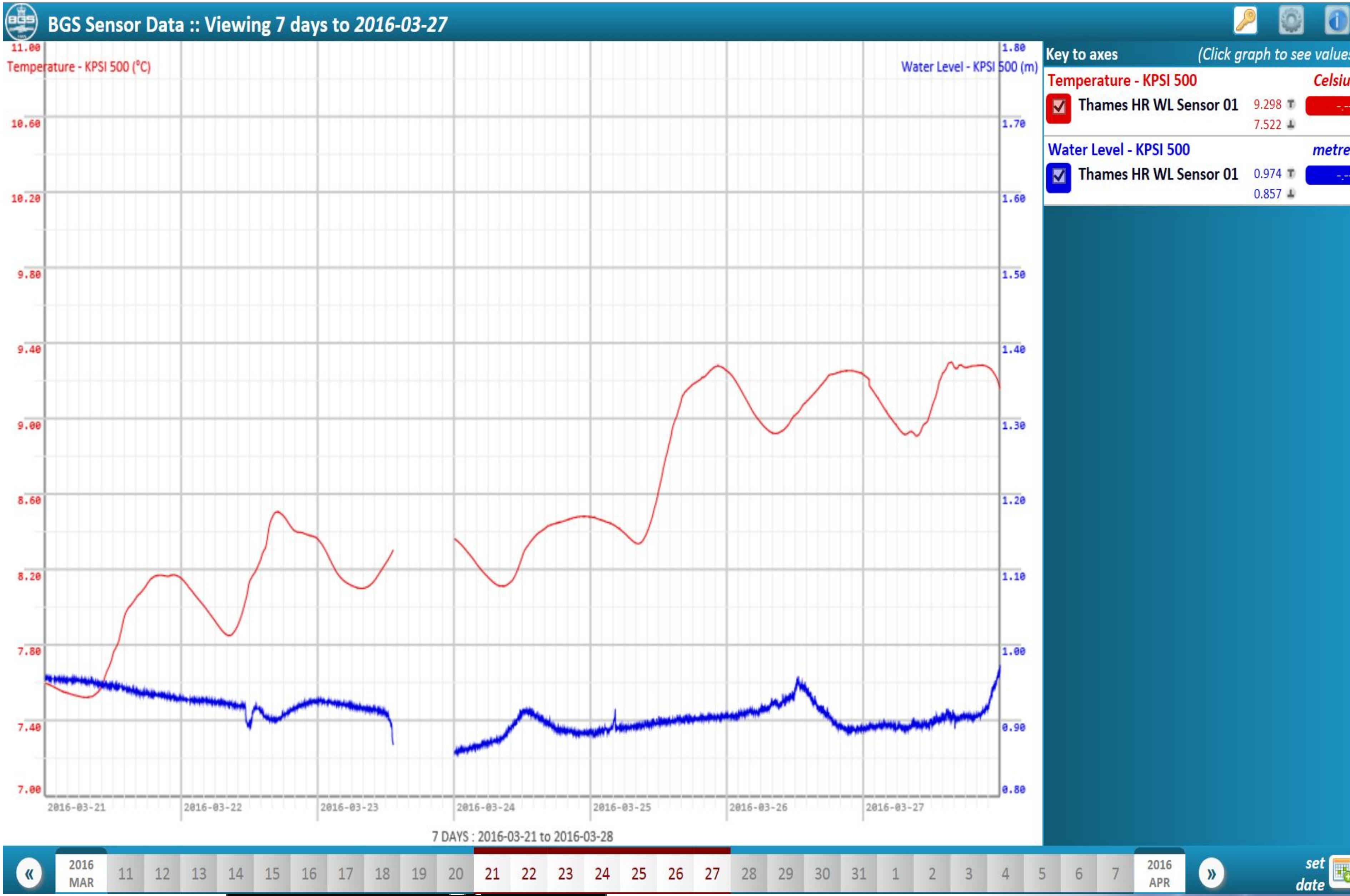
Visualisation: Temperature vs Water Level – Weekly

single sensor

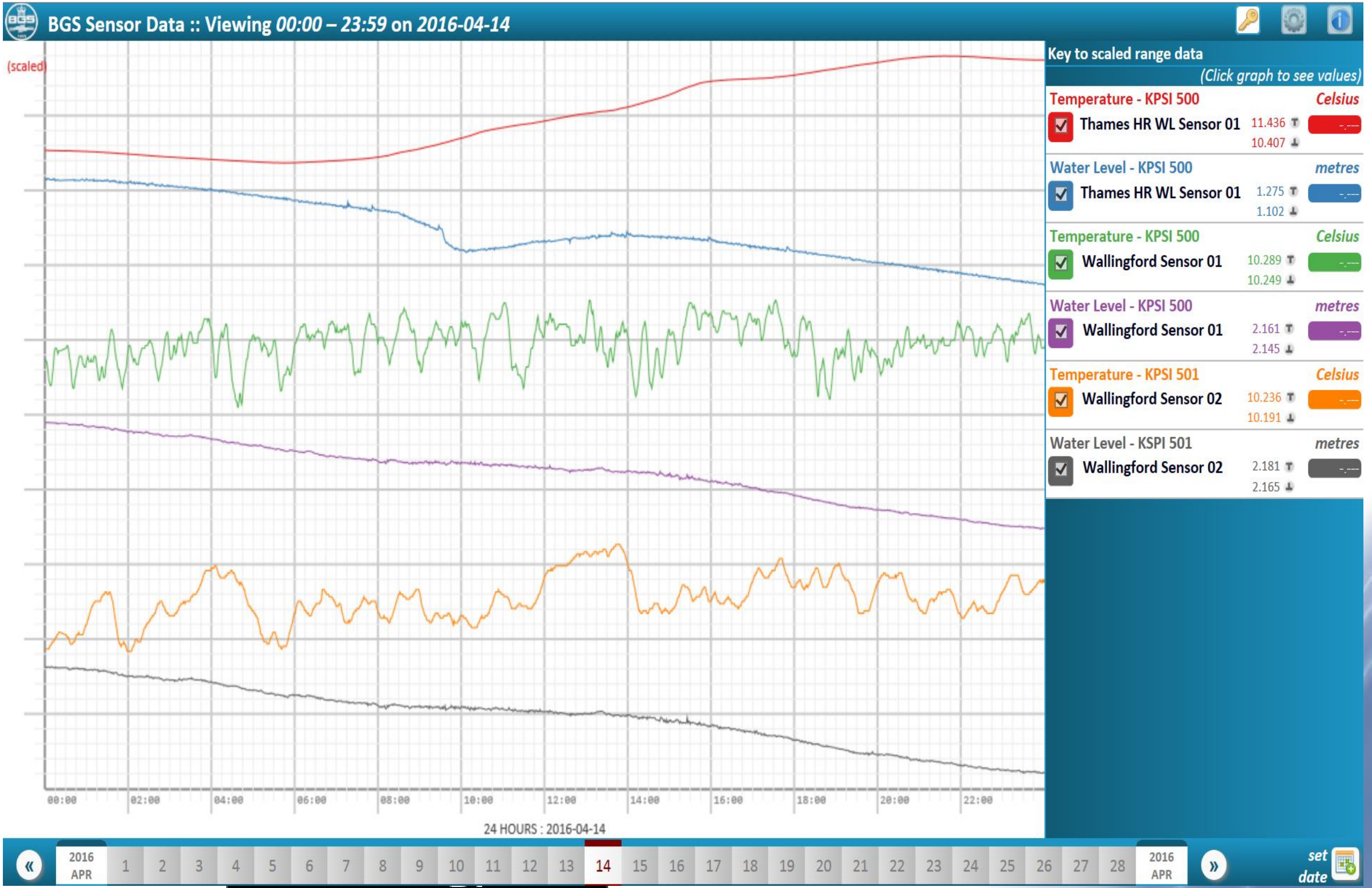


Visualisation: Temperature vs Water Level – daily

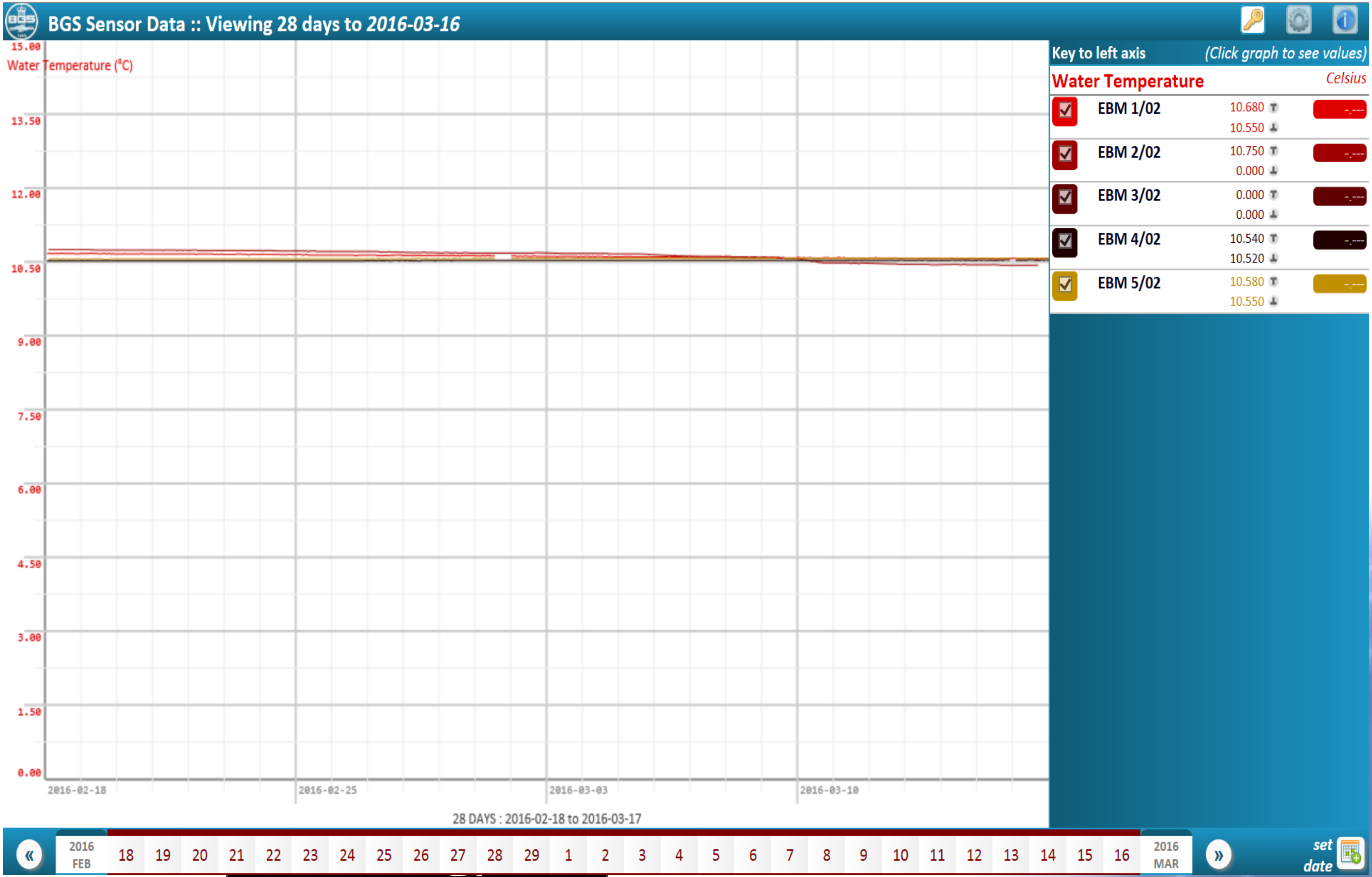
single sensor



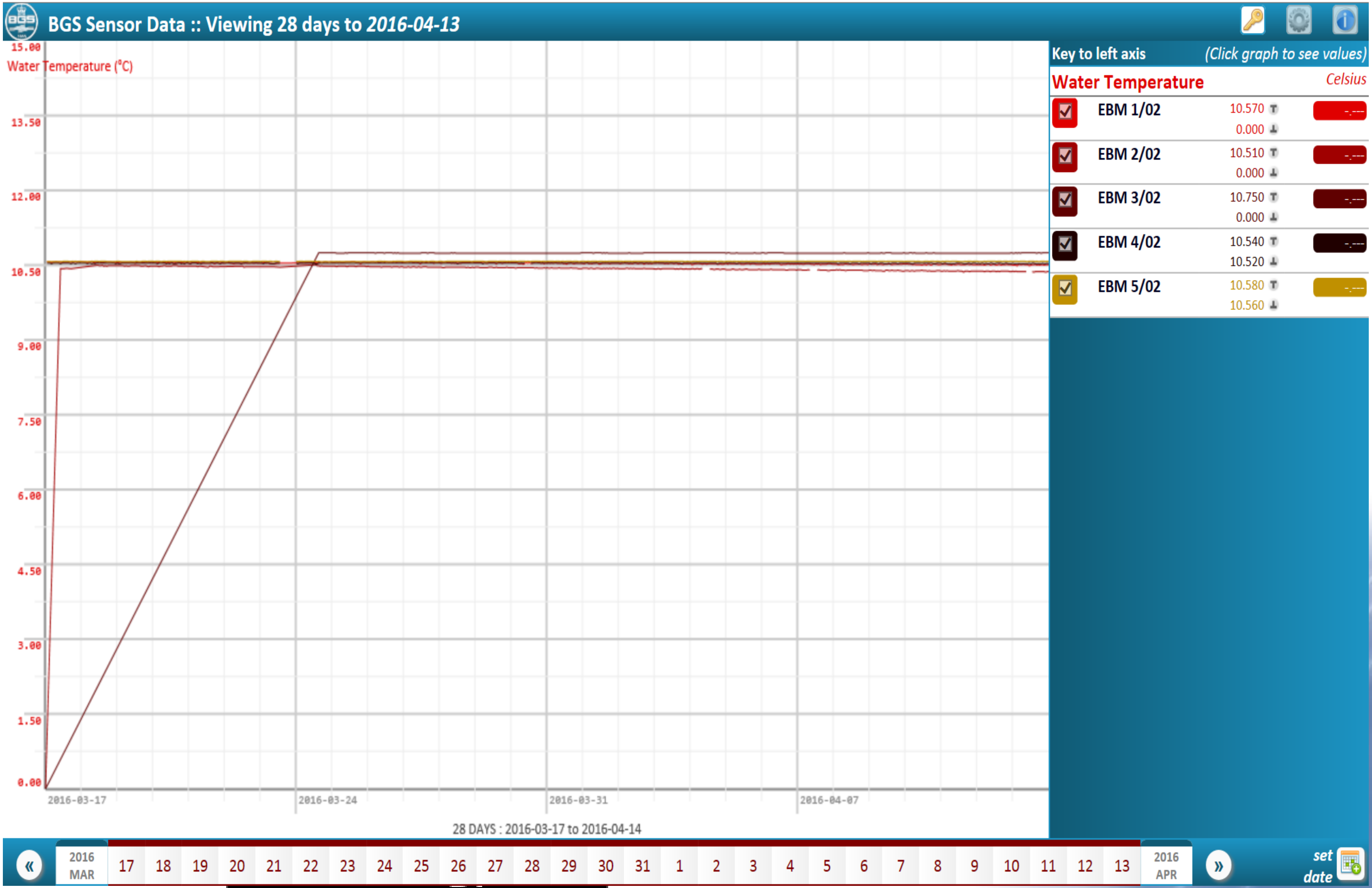
Visualisation: Temperature vs Water Level
multiple sensors



Visualisation: Water Temperature
multiple sensors



Visualisation: Water Temperature
multiple sensors



Uncertainty: How do we quantify and communicate uncertainty?

- Possible Sources of data errors - uncertainty:
  - Surrounding environment of the sensors
    - Signal variability/noise, interferences, calibration drift, deviation from lab simulation or other references
  - Instrumentation errors
  - Notification errors
  - Data transfer errors
  - Data conversion errors
  - Many other possible sources of error

An intuitive visualisation of uncertainty

Uncertainty – A description of the degree of accuracy of the final corrected data?

How can this be quantified?

How can we standardise on a methodology for doing this with time-series data and possibly apply to other datasets?

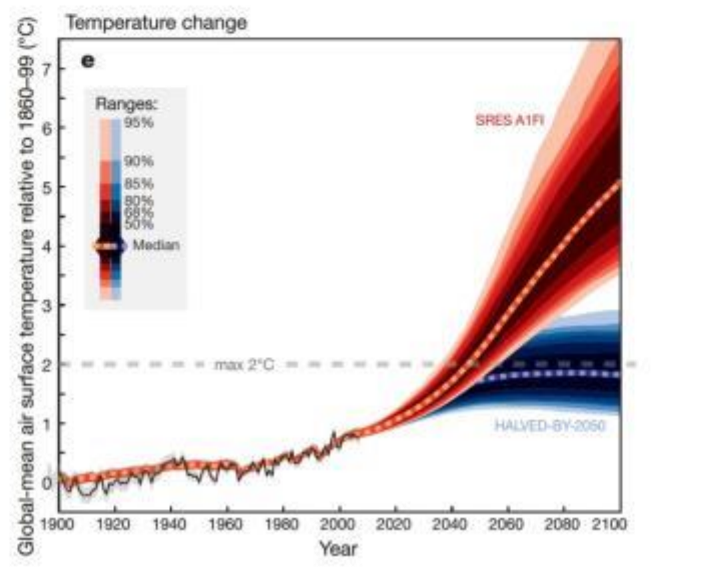
How can this be expressed as part of the visualisation?

Point – easy (1D line)

1D line – easy (2D surface)

2D surface – ok (3D volume or 2D surfaces)

3D volume – hard (second 3D volume)??

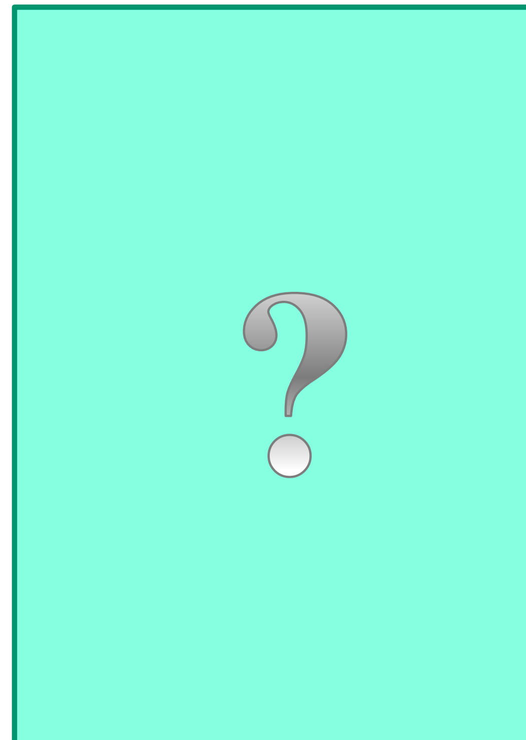


The problem - we can't scale traditional methods of visualisation. We need a novel way to represent the uncertainty.

An intuitive visualisation of uncertainty

2D – masking

3D – ??



Thank You!

- Martin Nayembil, BGS, mln@bgs.ac.uk
- Andy Barkwith, BGS, andr3@bgs.ac.uk