National Oceanography Centre
NATURAL ENVIRONMENT RESEARCH COUNCIL

The Big Data Dilemma Inquiry

Response coordinators:

Dr Jennifer Riley (Jennifer.sian.riley@noc.ac.uk)

Input:

Mr Kevin Forshaw (kevin.forshaw@noc.ac.uk)
Mr Steve Hall (sph@noc.ac.uk)

About us:

The National Oceanography Centre1 is part of the Natural Environment Research Council (NERC) and is the UK national focus for Oceanography. It undertakes large-scale, long-term oceanographic research from coast to deep ocean. It provides national capability in oceanographic sciences (including sustained ocean observing, integrated seafloor/risk/habitat mapping, large-scale ocean/self/ecosystem model development and technology innovation) and manages on behalf of NERC marine scientific research facilities, large research infrastructure and data assets – including the global-class research ships RRS James Cook and RRS Discovery, the National Marine Equipment Pool (NMEP), the Marine Autonomous and Robotics Systems Facility (MARS), the British Oceanographic Data Centre (BODC), the British Ocean Sediment Core Research Facility (BOSCORF) and the Permanent Service for Mean Sea Level (an international facility).

Declaration of interests:

NOC welcomes the opportunity to respond to the BIS Consultation on the long-term capital investment. The National Oceanography Centre (NOC) is a focal point for UK Marine Science and many of its programmes are funded through public money, accessed via NERC and ESPRC as well as other government departments such as DEFRA. This response presents the views of NOC and may not necessarily reflect those of our parent body NERC.

Submission date:

3 September 2015

Consultation Response:

A significant amount of data is currently collected from different sensors in the oceans. This includes data from floating buoys such as under the Argo programme[1], ships such as those under the Voluntary Observing Ship (VOS) scheme organised through the World Meteorological Organisation (WMO)[2], autonomous surface and underwater vehicles and from offshore oil and gas platforms. However, the current ability of the UK researchers to integrate large-scale modelling with complimentary large-scale observation datasets is not maximised for advancements in world-leading research and improvements in observational exploitation and predictive capability, of scientific research for the public good. Currently, the use of this data is sporadic and erratic, in terms of frequency, intensity and impact. With the imminent

[1] www.argo.net
[2] www.vos.noaa.gov

availability of greater bandwidths in satellite transmission and the forecast of cheaper transmission charges, an opportunity is about to open to facilitate better use of the diverse data sets.

The key to crucial advances in observational capability include:

- Improving data transmission from ships and remote platforms for real time, high quality data analysis
- Improving the usability of information already collected

The key to crucial advances in predictive capability, include:

- Improving down-scaling methodologies, to enable models to be used for predictive capacity at societally relevant spatial scales.
- Increasing modelling and observation resolution, so that truly effective predictive regional assessments can be made at scales of 5-20 km.

Improved predictive and observational capability will enable greater understanding of:

- High-productivity aquaculture practice,
- Fisheries productivity (associated with ocean eddies),
- Storm-surge and flooding risk,
- Environmental change,
- Hazard risk assessment
- Ship efficiency and safety

Such an increase in the knowledge base will ensure that appropriate mitigation and adaptation strategies to climate change are in place in the UK. To realize such a goal requires significantly increased investment in big data capital to support the necessary architecture and capacity of integrating large scale, high-resolution observational datasets with nested high-resolution physical and biogeochemical models. Such an investment will have notable benefits for:

- Government (e.g., transport infrastructure investment),
- Industry (e.g., environmental hazard re-insurance, changing food production practice),
- Wider society benefits for environmental resilience

In order to exploit data from ocean sensors and ships a cloud based computing infrastructure/architecture capable of super scalability to record data from globally distributed, diverse data sources (from different ship-board systems), whilst ensuring end-to-end data security and privacy is needed.

Such data architecture would consist of:

- Data collectors - separating the time critical low bandwidth data for immediate transmission from the larger finer grained data, which can be transmitted later, ensuring data transmission bandwidths are not exceeded. This will ensure that time sensitive data is ready for analysis, for example related to ocean weather and sea state, whilst finer grained data is available later for retrospective analysis and higher accuracy ocean, ship and climate modelling.

- Online brokers - act as a cache to cope with data bursts as well as a router, directing data to the appropriate transformation destinations. The data from the brokers is transformed into formats suitable for storage and processing and added to long-term cloud based storage. By using a cloud based storage and compute model we can easily make the underlying data sets accessible to all interested parties as well as reduce the overall infrastructure costs.