DTO Side Event Report

The objective of the "Digital Twin of the Ocean" Phi-week side event was to set the floor and stimulate fruitful discussions for shaping the "Digital Twin of the Ocean" concept in the context of the ESA-EC joint Flagship initiative dedicated to Ocean Health. It was organised in two 1h30 sessions with 11 invited keynote talks from internationally recognized experts in observing (remote sensing, in-situ), modelling, forecasting, analysing the different ocean systems components (physics, biogeochemistry, biology, ecology) and their complex interactions. In each session, 10 minute invited talks were followed by 30 minutes of questions and open discussions, by the virtual audience. Session recordings are available on Brella Schedule for both the morning (https://youtu.be/7SfKtXHl9t4) and the afternoon (https://youtu.be/qj7bJJFFstg) sessions[1].

1 Digital Twin Ocean (DTO) concept elements

DTO shall be seen as a comprehensive digital representation of the ocean, based on interoperable and trusted real-time and historical data from a wide range of sources, state of the art modelling (Physics, Sea Ice, Biogeochemistry, Ecosystems), coupling and data assimilation and visualisation tools. It will allow people to dive into the ocean and explore it, understand its interactions, dynamics and evolution and provide knowledge to empower fact-based decision making for management purposes, but also to the public at large. The DTO is a unique opportunity to connect Science to Society.

To achieve this goal, the Ocean5D concept was introduced, where the first dimension is time (ranging from the past, to the near real-time to predictions days to months ahead, to future scenarios), the next three space dimensions would cover the Ocean basin (from pole-to-pole and its coasts and 'hinterland' to the deep abyss). The 5th dimension correspond to the societal dimension and to the different issues such a DTO would be able to tackle. To build this 5th dimension, the DTO should be envisioned as interactive fit-for-purpose virtual environments, i.e. a multitude of digital twins based on a common framework and infrastructure, allowing design and play scenarios for specific targets.

2. A DTO dedicated to Ocean Health

The side event was dedicated in particular to the use of the DTO to address issues related to Ocean Health, i.e. such a DTO shall provide a virtual reality showing the impact of human interactions on Ocean health and allow for the testing of "what if scenarios". A DTO should help for instance in the framework of Marine Spatial Planning, in particular to decide on aquaculture facilities or wind farms location, to regulate fishing activities (where, when, how, how much), or to set up Marine Protected Areas/networks for which knowledge of Biodiversity hotspots, Productivity, Connectivity, Feeding/breeding grounds, and Migration routes is needed. It is also crucially needed to help politics achieve a number of EU and UN goals related to Ocean Health, such as reducing the rate of biodiversity loss for instance, which is part of the EU 2030 Biodiversity strategy, or achieving the Good Environmental Status in Europe in the frame of the Marine Strategy Framework Directive. Finally, ecosystems health is threatened by multiple stressors (warming, increased CO_2 and sea-level rise, decrease in pH and ocean

[1] *Contact Point: Marie-Helen.Rio@esa.int*

O_2, changes in nutrient availability) characterized by non-linear impacts and non-additive responses. A DTO for Ocean Health therefore needs to provide advice on how to address these cumulative impacts.

To achieve this goal, biodiversity and ecosystem function should be included in any DTO for Ocean Health. In particular, it should allow streamlined sharing through open-access databases of biodiversity and ecosystem data, either *in situ* or remote sensing observations, or Essential Biodiversity Variables (Genetic composition, species populations, species traits, community composition, ecosystem structure, ecosystem function). Such information shall then be integrated with other ocean observations (physical, chemical), and further analysed through models in order to retrieve higher level indicators needed to develop scenarios for biodiversity and ecosystem services, thereby allowing the transfer of knowledge of data and products to governments and intergovernmental organizations to inform policies and practices.

3. DTO Building Blocks: Existing elements and missing pieces.

Observations are a necessary building block of the DTO. Many different programs exist at European and International level. Some challenges remain, such as the fact that few of these programmes are sustained or the access to marine data from industry that falls beyond the domain of public data. Also, in the specific case of biological data, there is a strong need to improve data standardization (through the promotion of best practices), and data access (through the promotion of open access databases). Finally, observations from smart sensors, Internet of Things, adopting Sensor Web Enablement (SWE) supporting FAIR, citizen science, crowd-sourced data should also be included.

The DTO concept can build on existing projects and activities. In particular, the Blue-Cloud project was introduced, whose objective is to pilot a cyber-platform, federating several data infrastructures and providing access to multidisciplinary data from observations and models, analytical tools and computing facilities. Copernicus Marine Service was presented as providing many building blocks of a DTO such as HPC and cloud infrastructure, state of the art modelling, coupling and data assimilation, user interaction and user engagement.

However, many challenges and new capabilities still need to be developed regarding HPC and cloud computing (exascale, massive data access), observations, improved digital representation of the ocean (in particular biology, coastal, high resolution, coupling) and the use of Artificial Intelligence (AI) and Machine Learning (ML) for getting more knowledge out of data and better data products. For instance, model data and observational data can be leveraged for designing new products with AI as for instance using physically constrained AI-based inversion/extrapolation techniques, or emulating prediction systems with AI to reduce the cost of ocean models and provide lightweight engine for users. AI also brings an opportunity for rethinking our approach to uncertainty quantification. Finally, hybrid ocean modelling with machine learning is now a very active field of research. Proof of concepts have been published showing that ML can help to design sub-grid-closures and represent unresolved scales in models. This opens the possibility to leverage high resolution frontier model simulations, coupled data assimilation and observations for designing ocean model components.

In order to tackle the Ocean Health related issues, an ensemble of background models needs to be included and coupled in the DTO, to predict the diversity of organisms and habitats present at different locations, and infer how any changes due to human use or climate change might affect each area. If DTO is to be used for management, all the building blocks need to be looked at separately. The current state of the art is to use physical models to provide input to biogeochemical models, which provide input (nutrients, primary production) to food web models, which provide estimates of plant and animal biomass, which provide input to socio-economic model, which provides estimates of market and social values of ecosystems. However, all these models need to work in a more integrated way. There is a need to better understand the two way coupling of models, including backwards links. As those models are characterised by different spatial and temporal scales this can be challenging. In addition, for each part of the system an ensemble of models is needed. This will help to address uncertainty in model design and input parameters as well as multiple spatial and temporal scales. One of the challenges will be how to link these different ensembles. Big data and AI can help with that.

Among the main priorities, the need to identify the functional links in ecosystems was mentioned, together with the provision of a spatial frame for biodiversity and ecosystem function (BEF) and the enhancement of BEF observing systems.

Among the different functional groups a specific focus was put on the role of phytoplankton, which represents 90% of the marine biomass, are at the base of the marine food web and hence of all oceanic biodiversity, and are themselves characterized by an immense biodiversity (already 5000 species identified). Phytoplankton provides many services for life on Earth as food for humans and marine animals, oxygen production, CO_2 removal from the atmosphere, and provision of biomedical products. However, phytoplankton production is threatened by climate change (reduction by up to 30% is expected at the end of the century in a worst-case scenario).

The impact of phytoplankton decrease on fish stock and on carbon uptake will strongly depend on the community structure. Phytoplankton traits (size, morphology, function) is therefore a key element to be included into a Digital Twin Ocean, both in term of observations and modelling.

In term of observations, there is a need to bridge the gap between satellite data (global coverage and high spatial and temporal resolution) and the crude description of phytoplankton type and genetic analysis (very fine description of the ecosystem but at the expense of spatial and temporal resolution). Machine learning could potentially be used to connect satellite resolution and gene resolution.

In term of modelling phytoplankton diversity, more complex ecosystem models are now being developed, including phytoplankton size, biogeochemical functional groups, temperature affinity.

Finally, it is essential to couple both observations and modelling of phytoplankton diversity with ocean general circulation models. In particular, sub-mesoscale dynamics is key in structuring ecosystems. There is overall a crucial, challenging need for observing and modelling at fine time/space resolution of both phytoplankton biodiversity and ocean dynamics, and how to link them together.

Compared to existing capabilities, the DTO should also be more end-to-end, allowing stronger engagement of users and citizen uptake by adding extra shells around the core services. It should expand the computing core, not only delivering fixed products, but also custom products and integrate What if / scenario capabilities (climate, ocean health and resources management). Since playing with models to build what-if scenarios request significant expertise, such an interactive environment needs to be developed collectively (i.e. co-designed). Innovative approaches should be implemented to leverage citizen awareness and uptake. This could include for instance impact gaming, or developing engagement and smart dialogues with many user communities, using visualisation and story board techniques for more understanding, knowledge transfer, and interaction. A key aspect will also be how to characterize and communicate uncertainty, which will be an essential component of any result or predictions, to societal end users in a useful manner.

Finally, the Ocean is one component of the Earth System, and in building the DTO, the other components should also be considered: DTO should fit in with the other Digital Twins of the Earth, and synergies between them need to be enabled.