Scientific needs to realise Digital Twin Earth

This paper has been elaborated by ESA on the basis of the finding of different sessions and side events organised during the Phi week 2020. It provides a summary of the main recommendations and discussion points addressing some of the major scientific needs for the realisation of Digital Twin Earth[1].

1 Introduction

In the next decades population growth and human activities are expected to amplify current pressures on critical resources such as fresh water and food, intensify the stress on land and marine ecosystems and increase environmental pollution and its impacts on health and biodiversity. These threats will be further exacerbated by global warming and the likely impacts of climate change in the Earth system. These, comprising rising sea level, increasing ocean acidification and more frequent and intense extreme events such as floods, heat waves and droughts, are expected not only to have a significant impact on critical resources, but mainly to endanger human lives and property, especially for most vulnerable populations.

Responding to these challenges, the EU Data Strategy proposes to launch “Destination Earth” as an ambitious initiative to “...bring together European scientific and industrial excellence to develop a very high precision digital model of the Earth. This ground-breaking initiative will offer a digital modelling platform to visualize, monitor and forecast natural and human activity on the planet in support of sustainable development thus supporting Europe’s efforts for a better environment as set out in the Green Deal.”

Europe has now a unique opportunity to realise such a vision. Destination Earth comes at the time where novel and emerging technologies, including High Performance Computing (HPC), cloud computing platform capabilities, Internet of Things (IoT), and Artificial Intelligence (AI) are transforming our global economy and societies. In addition, Europe is developing the most comprehensive and sophisticated space-based observation infrastructure in the world, including an extraordinary and complementary suite of sensors on board of the Copernicus Sentinel series, ESA’s Earth Explorers, the meteorological missions and variety of EO observation satellites launched and operated by national space agencies and private operators in Europe and worldwide.

Destination Earth will be a catalyst to integrate those technologies into a very high accuracy digital model of the Earth being able to monitor and simulate natural and human activity, and to develop and test scenarios that would enable more sustainable development and support European environmental policies and decision making.

[1] *Contact Point: Diego.Fernandez@esa.int*

Realising such an ambitious system, a Digital Twin of the Earth (DTE) or a set of Digital Twins (DTs) focused on specific policy needs, components of the Earth system or regions, would require a long term effort, institutional collaboration and dedicated investments to advance and integrate in an effective manner the different digital technologies, data infrastructure and processing capabilities that will represent the building blocks of DTE.

However, the success of such a system and hence its wide adoption and ultimate impact on society will strongly depend on its scientific credibility and its capacity to offer scientifically sound solutions that may match the needs of a wide variety of sectors and users expanding from decision makers and scientists to citizens and economic operators.

Therefore, it is critical to ensure that DTE design and implementation is based on scientific excellence and on a strong connection to the scientific community during all phases of its development. In fact, DTE may represent not only a catalyst for digital technologies and data infrastructures but also and accelerator and integrator of science. On the one hand, scientists, from a wide variety of different disciplines, will be a primary user community of such a system. On the other hand, scientist and academia should strongly contribute to drive the expansion of DTE across different sectors and Earth system domains. To achieve this:

- Investments in technology and digital infrastructures should be complemented by dedicated efforts in science devoted to enhance our capacity to observe, understand describe, and predict the complex and inter-connected natural and anthropogenic processes beyond current climate modelling and NWP capabilities planned at the core of DTE;

- DTE infrastructure evolution and expansion should be underpinned by an strong scientific component aimed at engaging the scientific community in a collective effort to establish a solid scientific basis for novel developments and specialised solutions across different sectors and Earth system components: carbon cycle, food systems and agriculture, ocean health, polar sciences, hydrology, biodiversity, etc..;

- Such a scientific component should capitalise on the scientific efforts, initiatives and programmes undertaken by different European and national institutions including ESA EO scientific activities and the recently launched EC-ESA joint Earth system science initiative addressing some of the most pressing scientific needs relevant for DTE evolution;

- DTE evolution and expansion should rely on an open science approach stablishing effective and inclusive mechanisms to ensure that the scientific community may participate in the DTE development process and that relevant advances in science may be shared, tested, plugged-in and eventually adopted within the DTE operational implementation well beyond its initial planned capabilities.

Realising the DTE vison will require significant advances in science across different but interconnected domains. In the following, this paper addresses the some generic scientific needs required to advance in the realisation of Digital Twin Earth with special focus on four specific areas where EO technology may provide a significant contribution[2]:

---

[2] The following points represent an elaboration of the main findings of the session "Earth system science in the Digital Twin Earth" complemented by the results of a consultation process organised on-line during the Phi week.

- *Advance Earth observation and the description of processes from observations*
- *Advance Earth system science and process understanding*
- *Advance sectorial modelling capabilities and the integration EO into specialised models*
- *Advance AI as a tool for Earth Observation and Earth system sciences*

2 *Advance Earth Observation and the description of processes from observations*

Advanced satellite-based Earth observation capabilities providing global and synoptic observations at unprecedented spatial and temporal scales across the full spectrum of the Earth system domains represent an essential element to realise the Digital Twin Earth vision. Information provided from satellites, complemented with in-situ data collected by consistently resourced global and regional networks, novel in-situ observation technologies (e.g., UAVs, smart sensors networks and internet of things) and citizen observations, shall represent the core of a dynamic reconstruction of our planet and its different processes and complex interactions required to realise the Digital Twin Earth ambition.

The first actionable layer of information of DTE, shall be a consistent, coherent and comprehensive "**4D data-driven reconstruction**" of the Earth system and its processes offering users with an advanced description of the past and the present state of our planet with accuracies and temporal and spatial scales compatible with decision making needs.

Realising such a "4D reconstruction" of Earth system at process level and capturing the complex interactions and feedbacks across the different components of the Earth system is an ambitious scientific endeavour that goes well beyond collecting different types of information into an heterogeneous data lake. It would require:

1. **Exploiting the novel capabilities and technological advances offered by the new Earth missions** and sensors to develop a new generation of EO products that may overcome some of the current observational gaps and limitations;
2. **Exploiting the huge synergistic capabilities** and the unique range of complementary information provided by the wide variety of different types of missions and sensors towards novel multi-mission products and datasets **that may not only capitalise of different observation types or frequencies but also different spatial scales and temporal coverage.**
3. Advancing towards the **development of process-based retrieval strategies** that may result in **novel multi-variate datasets providing a coherent and comprehensive description of complex Earth system processes** taken into account for potential redundancies and dependencies across variables**.**
4. Exploring **novel approaches to integrate multiple and heterogeneous data sources** including satellite-based products, in-situ data or citizen observations into a coherent information data structure ensuring consistency across variables whilst accounting for their individual uncertainties and spatial-temporal covariations.
5. **Ensure consistency of information across all levels of spatial and temporal extraction** offering a coherent framework to exploit the increasing number of EO based products and observations at different spatial and temporal scales.
6. **Ensure that the above developments are tied up with a sound characterisation of error and uncertainties** providing a solid basis for subsequent uses of the data and eventual assimilation in Earth system models.

In addition, the ambition of DTE to provide information at resolutions compatible with decision making implies the development of a novel capacity to observe, characterise and simulate local processes that may capture the human footprint in the Earth system and ecosystems. The increasing availability of high-resolution sensors (e.g., from Sentinel-1 and -2, commercial satellite constellations) and its combination with in-situ networks offers new opportunities to achieve this goal. However, the integration of such information into Earth system models pose a number of changes related to: 1) data limitations also in terms of error an uncertainty characterisation, 2) the complexity of the underlying processes and 3) limits in current modelling and assimilation approaches, among others.

Finally, it is important to note that observations alone are not sufficient to provide a complete and exhaustive description of the complex process that govern our planet. Several variables and processes cannot be observed or are only observed partially, with insufficient accuracy or at unsuitable spatial and temporal scales. Therefore, the above "4D data-driven reconstruction" shall be complemented with effective integration of data and different type of models. This may involve the estimation of higher level information products derived from model inversions, forward modelling, or data analysis and reanalysis based on data assimilation.

3 Advance Earth system science and process understanding

A basic element of Digital Twin Earth is the capacity to provide a scientifically sound description of complex processes at global, regional or local level and feedback mechanisms across the Earth system components including both natural phenomena and human activities. Therefore, Digital Twin Earth must be driven by the latest advances in Earth system science and underpinned by scientific excellence across a wide range of scientific domains and sectors.

The level of advancement in the fundamental scientific understanding of the Earth system is quite diverse across the wide spectrum of scientific disciplines and a comprehensive assessment of the major scientific gaps in Earth system science is beyond the scope of this paper. Nevertheless, a sound scientific implementation of Digital Twin Earth would require dedicated scientific efforts to address a number more generic needs where the latest advances in EO technology may provide an important contribution:

1. **Enhanced the scientific understanding of complex processes involving different components of the Earth system** across different scientific domains and traditionally distant disciplines including solid Earth, ocean, land processes, cryosphere and lower and upper atmosphere in an holistic manner.
2. **Enhanced the scientific understanding and characterisation of the role of human activities as an integral part of the Earth system**. This may require not only to incorporate new ways to inform models of the different impacts of human activities in the environment and understand its feedbacks but also to incorporate the human behaviour, political and socio-economic aspects in the way we describe the Earth system.
3. **Enhanced understanding and characterisation of the role of biology and ecosystems in the Earth and climate systems** accounting for the complex interactions and feedbacks across physics, biochemistry and ecosystem functioning.
4. **Enhanced understanding and characterisation of multi-scale connections** of complex processes from local scales to regional and global scale dynamics.

The rapidly expanding EO capabilities (e.g., Sentinel missions, Earth Explorers, meteorological missions, national and commercial satellites) offers an unique opportunity to address some of the above needs. In particular, the huge synergistic potential offered by the joint analysis of multi-variate EO datasets together with in situ observation and advanced models will provide a key source of new knowledge and an unprecedented potential to gain new insights on poorly known processes and complex interactions across the Earth system components. However, such a scientific advances and its subsequent adoption by the DTE operational implementation will not be possible without dedicated scientific efforts devoted to: e.g.,

- **Bringing together scientists from different sectors to jointly undertake far-reaching interdisciplinary research** connecting Earth processes, society and ecosystems and resulting in novel and enhanced capabilities and new science cases driving the DTE evolution and expansion;
- **Involving early adopters** (e.g., economic operators, public institutions, local communities) from the initial phases of the science process to accelerate the transition from science into actionable workflows that may be adopted by the DTE operational implementation;
- **Promoting Actionable-Science partnerships** fostering cooperation between academia and industrial partners to jointly develop novel DTE capabilities that may bring together advanced scientific solutions and innovative technologies.
- **Establishing an Open Science access to DTE infrastructure for the scientific community** allowing scientist to plug-in new elements, foster novel developments and testing new capabilities before being adopted by the DTE operational implementation.
- **Foster a community driven process** engaging the scientific community in all phases of the DTE definition, implementation, validation and evolution.

4 Advance sectorial modelling capabilities and the integration of EOs into specialised models

Digital Twin Earth shall be able to provide enhanced simulations, predictions and improved forecast capabilities at resolutions (in space and time) and accuracies compatible with the needs of policy makers to respond to the main challenges addressed by the Green Deal. Today, a new generation of models, such as high-resolution cloud-resolving models or convection permitting models, is emerging opening the door to more realistic representation of small scale dynamics. Digital Twin Earth shall capitalise on these advances and make a significant step forward by fostering an effective integration of advanced models, cutting edge AI methodologies and the latest EO observations capabilities to achieve high performance simulations at extreme scale.

Realising this vision represent a major scientific challenge that would require an effective combination of enhanced observations (satellite, in-situ, citizen observation) and estate-of-the-art models. Even if such an integration is well advanced in different domains, such as NWP, climate modelling or operational oceanography, significant efforts are still needed in many other sectors of the Earth sciences where the dialogue between the EO experts, Earth system scientist and modelling communities needs to be reinforced or are in its infancy. Therefore, a science pillar of DTE shall foster this dialogue, accelerate the integration of EOs and sectorial or specialised modelling and overcome some of the fundamental limitations in current modelling approaches to characterise different components of the Earth system.

Making an exhaustive analysis of the scientific needs required to advance modelling capabilities is beyond the scope of this document. However, a few key generic priorities can be summarised as follows:

- **Enhanced resolution of current modelling approaches** up to actionable scales in space and time compatible with decision making;
- **Advances in development of a fully coupled Earth system** model incorporating all domains (ocean, land, atmosphere and cryosphere) ensuring the effective and consistent integration across sectors and Earth system components.
- **Advances in consistency across spatial and temporal scales** ensuring coherence from global climate, ocean or land surface models down to regional or local scales simulations.
- **Advances in the data assimilation of novel EO data products** especially novel and multiple high-resolution data streams.
- **Advances in the incorporation of the human dimension** in the description of the Earth system, exploring novel approaches to connect human and social behaviour with Earth system modelling.

5 Advance AI as a tool for Earth system science

The field of Machine Learning has quickly evolved. AI is expected to radically improve current capabilities to simulate and predict complex phenomena and to overcome some of the main limitations of current data and modelling approaches in Earth and climate science.

AI can be instrumental at all levels of DTE implementation from the enhancement of the basic data layers to performing the most complex simulations. However, the use of AI for Earth system science is in its early stages. A major scientific effort is needed to overcome some fundamental limitations, reach a sufficient maturity level, gain confidence on AI capabilities and foster the adoption of AI as an additional tool for advancing our understanding of the Earth and its intertwined subsystems. In particular, major research needs include[3]:

- **Exploring the potential of AI for enhancing EO-based products** using AI as a tool for EO data analysis and enhanced or alternative retrieval methodologies;
- **Addressing data challenges** resulting from the wide variety and heterogeneity of data sources, the size of the data space, data gaps and scarcity of observations;
- **Learning and simulating complex models** accelerating our capacity to run simulations in real time.
- The development of **hybrid approaches** combining AI and physical modelling balancing between deriving parameters from first principles and learning directly from the data;
- **Advancing novel techniques such as reverse engineering or automated machine scientists** that may allow to learn the basic equations of a system from the data including variables and parameters of heterogeneous nature: e.g., environmental and socio-economic data;
- **Dealing with uncertainty** in modelling: e.g. understanding of model errors and error patterns through analysis of model-state and observation mismatch;
- **Equip Earth system science against non-rigorous uses of AI methods:** avoiding dubious results that undermine the general reputation of AI and Earth system research and advancing eXplainable AI (XAI).

---

[3] *From session "AI for understanding processes in intertwined Earth System Dynamics" organised by Miguel Mahecha (Leipzig University), Gustau Camps-Valls (Valencia Universitat) and Carsten Brockmann (Brockmann Consult GmbH).*