FEATURE ARTICLE: BY INVITATION OF THE EDITOR-IN-CHIEF

Destination Earth: High-Performance Computing for Weather and Climate

Nils Wedi, Peter Bauer, Irina Sandu, Jörn Hoffmann, Sophia Sheridan, and Rafael Cereceda, *European Centre for Medium-Range Weather Forecasts, 53175, Bonn, Germany*

Tiago Quintino and Daniel Thiemert, *European Centre for Medium-Range Weather Forecasts, Reading, RG2 9AX, U.K.*

Thomas Geenen, *European Centre for Medium-Range Weather Forecasts, 40128, Bologna, Italy*

Destination Earth is the first grand effort to define and deploy digital twins of the Earth system. The European Commission is making this important, multiyear investment to develop this new type of information system, blending the physical and digital worlds. The scale of computational resources and data flows is unprecedented, and so are the challenges and the opportunities. Digital twins of Earth will support decision making faced with weather extremes and climate change adaptation as well as provide to users the means to interact, modify, and create their own tailored information. Building on the latest science and technology advances, this article describes the steps to realize the dream of preparing a more resilient society faced with unprecedented climate change in the decades to come.

DIGITAL TWINS OF THE EARTH SYSTEM

In the field of Earth system science, interest has soared in both "digital twins" and turning scientific knowledge into societal benefits, particularly in view of large (multi-)national investments toward the so-called green transition of our society.[a–c] This has led to sometimes simplistic views of what digital twin technology means and where its true potential lies.

Digital twins originate from industry, where they can help to optimize the design of a production process (e.g., in a factory) or a technical component (e.g., an engine) and/or help to optimize the operation of such a production process or technical component.[1] Digital twins can rely on simulations only or on simulations combined with observations.

Advances in numerical weather prediction are feeding the ongoing effort to create digital replicas of the Earth system. Two notable examples are the European Union's Destination Earth (DestinE) initiative and the National Oceanic and Atmospheric Administration's digital twin for Earth observations. Encapsulating the latest science and technology advances, the DestinE twins aim to provide near-real-time information on extremes and climate change adaptation in a wider digital environment where many users can interact with, modify, and ultimately create their own tailored information.

Recent work has demonstrated that global coupled storm-resolving (or kilometer-scale) simulations are no longer a dream and can contribute to building these information systems. This has been made possible thanks to recent advances in Earth system modeling, supercomputing, and the ongoing adaptation of weather and climate codes for hardware accelerators.

These advanced simulations start to represent essential climate processes, e.g., detailed inland water and land use representation, deep convection, and mesoscale ocean eddies that, today, need to be fully parameterized even at the highest resolution used in global weather and climate information production.

[a] https://reform-support.ec.europa.eu/what-we-do/green-transition_en
[b] https://www.state.gov/driving-the-clean-energy-transition-a-progress-report-on-implementing-u-s-efforts-to-advance-clean-energy/
[c] https://www.weforum.org/agenda/2023/01/davos23-how-can-we-accelerate-the-energy-transition-in-the-asia-pacific-region-3-leaders-weigh-in/

Digital Object Identifier 10.1109/MCSE.2023.3260519
Date of current version 1 June 2023.

These simulation outputs combined with novel, data-driven deep learning advances open a window into the future by significantly increasing the realism and timeliness in delivering Earth system information. Despite the significant computing and data challenges, including both memory-intensive and extreme-scale workflows, there is a real prospect of better supporting a broad range of applications and users in global to local warning systems, complementing existing climate change mitigation and adaptation efforts.

In essence, a digital twin allows users to test and trial solutions. This makes it possible to predict impacts given changing scenarios. The reliability of these predictions depends critically on the realism of the simulations and the accuracy and representativeness of the observations. The digital twin and the real world are connected in both directions: the twin is shaped by the information it receives on the evolution of the real world, and it influences how the real world evolves through providing new information.

THE COMPLEXITY AND SCALE OF EARTH'S TWINS ARE MANAGED THROUGH DISTRIBUTED RESOURCES.

Digital twin technology can be applied to Earth stewardship, supporting society in becoming more resilient to the evolution of the climate in the years and decades ahead. This is crucial, given that climate change impacts are likely to dramatically affect the availability of clean water and food as well as increase risks to health, lives, and property. For this application, the twin will be mostly based on simulations, but long records of past observations can help tune imperfect models and bridge the science with the impacts to society. The purpose is to develop adaptation options and advise policy makers on the changes required to transform society, making it resilient to climate change. For this to succeed, models and observations need to include both physical and human-driven impacts. On shorter timescales, such as dealing with environmental extremes days to months ahead, the twin relies heavily on observations to create possible scenarios for emergency response agencies to plan for, potentially saving lives and improving safety.

To achieve technical success, a digital twin must generate high-quality data through simulations and observations while also offering the speed, interactivity, and tools necessary to extract information and trial different scenarios. To become a digital twin, in fact, an information system must achieve a sufficient level of data quality and interactivity.[2] Since neither simulation models nor observations will be perfect, it is essential to quantify uncertainty when assessing whether the impacts of different action scenarios can be meaningfully distinguished and remain physically realizable. Rather than designing a single digital twin, it is more realistic and efficient to consider a system of twins that are connected, complementary, and continually updated.[3]

For the Earth system, these conditions result in unprecedented footprints for computing and data, most likely at the exascale (i.e., exa-floating-point operations per second) with multiple petabytes of daily data production). Digital twins will necessarily be highly resource intensive due to the complexity of the models and the vast number of diverse observations but also the need for fast throughput to minimize the time to decision, the need for trialing different scenarios, and the diversity of information required by a growing number of applications. The complexity and scale of Earth's twins are managed through distributed resources. This requires interoperable data and software infrastructures so that the science–technology–societal impact nexus can be fully exploited. Upscaling the interconnectivity and interactions with dedicated but distributed resources is new and a paradigm change compared to the present practice of an individual-center approach supported by its dedicated computing resources.

Substantial investments in digital technologies are a prerequisite for the successful implementation of Earth system digital twins. The existential threat of climate change demands that we scale up our dedicated resources to address this challenge. Machine learning (ML) has become an essential tool in many areas, enabling us to handle diverse observations, accelerate computing, and extract as well as translate specific information from vast amounts of data.[4] Increasingly, these methods are being used to replicate entire forecasting systems, which are largely based on physics-based models and observations for training, such as those parts of the digital twin systems discussed here.

A digital twin of Earth is an information system that serves a range of users. It offers scientists a highly effective tool to improve simulation models and fuses the available information from simulations and observations. It connects physical weather and climate models with impact models to interact with the system and define bounding solutions. Paired with a constantly improving level of scientific confidence and data quality, the digital twin offers an information system architecture based on generic interfaces and modular software infrastructures. This approach improves the dissemination and technology transfer of methodologies and facilitates access to (super-)computing and data handling resources that typically require specialist knowledge to be efficiently exploited.

Nonexpert users may ask questions that trigger interactions with Earth system models and observations. However, these interactions are not necessarily direct but, rather, through exploiting existing or on-demand data, delivered as fast and effectively as possible, while also allowing users to add their own data analysis and value-adding tools. In policy making, it is crucial to have standardized workflows and methods to monitor the quality of information used and to ensure that the information is fully traceable. To achieve this, Earth system digital twins will leverage the processes and standardization schemes that weather and climate services have been developing for decades, building and sharing best practices.

The ultimate challenge will be to align these expectations and challenges with a relatively short delivery time. Operating such an information system in evolving complex national and international governance structures is equally demanding while continuously embracing both technological and scientific progress.

UNCERTAINTY ABOUT THE FUTURE

Local extreme events are inherently uncertain, and their potential occurrence in future climate scenarios is even more so. Unfortunately, the tools available today to derive global climate projections do not enable us to easily focus on local events.

To quantify uncertainty in numerical weather prediction, we rely on ensembles with perturbed initial conditions and apply additional perturbations over time during the evolving forecast trajectories of the model simulations. The resulting products derived from the set of scenarios quantify the statistical likelihood of weather and extreme events and allow us to extend the predictable forecast range of mean changes for weather parameters, such as temperature and pressure systems.

In the context of both weather and climate scenarios, events are classified by probabilities obtained from statistics of simulations. They may be classified by the likely return period of extreme events, such as flooding or heat waves, or the likely increase or decrease of mean changes in temperature and precipitation. Recently, other approaches, such as storylines, are being encouraged[5,6] to address projections of low-likelihood events that, nevertheless, bear a high financial or even existential risk, such as flooding or energy grid failures. Storylines are utilized in a way to easily comprehend best- and worst-case scenarios and facilitate planning and decision-making processes. As such, storylines are perfectly suited for digital twins. Computing resources may be focused on events that can be derived with a much faster time to information and, therefore, aid the continuous enhancement of the statistical underpinning of longer-term trends. The computational resource requirements (and, equally, the data processing requirements) for an ensemble of kilometer-scale digital twins or multiple multidecade kilometer-scale scenarios are large, with data production rates well exceeding 1 PB/day. Intelligent data processing and reduction methods that explicitly consider a range of equally likely outcomes are, therefore, necessary.

Providing on-demand capabilities to constrain regional or local uncertainty and downscale to the needs of the impact sector applications highlights the need for programmable data pipelines.

These pipelines will enable the reduction and translation of complex weather and climate uncertainty information to the essential elements, which may differ for each application. To scale this process, some responsibility for collecting digital twin information will need to shift from production to application owners, allowing for an increasing number of users to participate.

DESTINE—A GATEWAY TO DIGITAL TECHNOLOGY

DestinE[d] is the first concerted effort to implement and operate digital twins of the Earth system at the required scale. It will contribute to achieving the objectives of the twin transition, green and digital, in alignment with the European Commission's Green Deal and Digital Strategy, by exploiting present and future multinational investments in digital technology infrastructures and providing end-to-end capabilities that combine Earth system, computer technology, and application science. Although presently limited to Europe, DestinE should be seen as the nucleus for the larger international collaboration and data network that are required to achieve the necessary intellectual and financial global support.[7]

In its first implementation phase, DestinE focuses on climate change adaptation and weather-induced extremes.[e] The activity is carried out through a partnership between the European Centre for Medium-Range Weather Forecasts (ECMWF), the European Space

[d]https://digital-strategy.ec.europa.eu/en/policies/destination-earth
[e]https://stories.ecmwf.int/destination-earth/index.html

Agency (ESA), and the European Agency for Meteorological Satellites (EUMETSAT). It is funded from the Digital Europe Programme managed by the European Commission's Directorate-General for Communications Networks, Content, and Technology and planned for a seven- to 10-year period between implementation and operation. The 2.5-year implementation phase is funded with €150 million, while the necessary computational resources are expected to be provided by the European High-Performance Computing Joint Undertaking (EuroHPC JU) Joint Undertaking,[f] which manages the implementation of a €7 billion investment program toward European exascale infrastructures.

One critical aspect of DestinE is its approach to investing in the science, computational, and application-oriented aspects of digital twins simultaneously. The ECMWF has formed partnerships with large European consortia, bringing together research and high-performance computing (HPC) centers, academia, and service providers to achieve this goal. In addition, contracts have been awarded for the provision of technology components and use cases. Weather and climate dwarfs, which are mini applications resembling relevant algorithms, provide a means for partners to collaborate on focused adaptation and science topics. These dwarfs allow rapid progress on testing and verification for HPC use and adaptation without the complexity of full applications. This innovative approach fosters collaboration and accelerates progress in developing digital twins of the Earth system, bringing us closer to achieving our goals of a resilient future.

The digital twins will use so-called *data bridges* placed near some of the largest EuroHPC computing installations [IT Science Centre Finland (CSC) in Finland, Italian Supcercomputing Consortium (Cineca) in Italy, and Barcelona Supcercomputing Centre (BSC) in Spain]. They contribute to an interactive data space where the digital twin output streams described feed ML methods and where users can interact with workflows and data. This exemplifies the philosophy of placing the compute near the data. The data bridges will handle petabyte/day data rates and tens-of-petabytes data volumes, with the bulk of computing on HPC. Substantial investments are also made in smarter methods for data handling, compression, and information retention to keep the bridges affordable and to maintain fast access rates.

In support of DestinE, EUMETSAT will break up information silos by integrating data bridges into a wider concept of distributed data and resources. Meanwhile, the ESA will invest in a service platform that complements DestinE and enables users to register and interact with DestinE output and tools. The ECMWF will develop and implement digital twins and foundational software infrastructures that enable exascale computing and data handling across DestinE. The first end-to-end demonstration, including links to selected impact applications across all components, is expected in 2023.

[f]https://eurohpc-ju.europa.eu/index_en

THE SCIENCE OF EARTH SYSTEM DIGITAL TWINS

There are high expectations for digital twins to improve physical modeling across local to global scales, blurring the lines between the physical and digital worlds. They are like interactive, user-driven digital factories that offer a continuously changing perspective of the past, present, and future of the Earth system.

DestinE will provide global, continuous, and on-demand configurable components of a digital twin for weather extremes (Extremes DT). These components are expected to deliver capabilities for a global European monitoring and prediction framework that supports decision making from the global scale to continental, country, coastline, catchment, and city scales in response to meteorological, hydrological, and air quality extremes. Extremes DT will combine weather, hydrology, and air quality observation and simulation capabilities at a qualitatively new level, also providing uncertainty quantification. Additionally, the Climate Change Adaptation DT is expected to deliver multi-decade Earth system simulations based on a global storm/eddy-resolving numerical Earth system simulation capability with more than one modeling system. This will be demonstrated with throughput rates that enable the timely delivery of climate information for policy adaptation.

Challenges and opportunities arise from the enhanced resolution of certain Earth system components (up to kilometer scales, e.g., 1–4-km global grid spacing). The change requires a corresponding improved representation of Earth system processes in the model. Moreover, the closely associated assimilation of observations at fine scales challenges traditional approximation procedures of assimilation, such as 4-D variational assimilation.

Another challenge is the focus and integration of selected impact sector elements (e.g., atmospheric composition, hydrology, and renewable energy planning) in the Earth system modeling framework at the kilometer scale and the design of interfaces across disciplines. This combines with the need for evaluating, quantifying, and communicating the uncertainty associated with digital twin data across disciplines. Relevant research topics include uncertainty quantification supported by ML and communicating uncertainty through artificial intelligence.[4] Equally, the distribution and fusion of data generating subject-relevant new information can be supported through ML (e.g., health sector distribution risks, the management of renewable energy resources, and transport-derived dynamic motion data) and (re-)estimating poorly quantified emissions data from simulated responses for environmental protection.

Finally, digital twins aim to be production systems, providing multiple interacting components with both continuous and on-demand capabilities. This requires a rigorous research-to-operations translation process, with continuous development and integration pipelines, to manage both scientific advances and the continuous technological adaptation to emerging computing architectures.

NOVEL WAYS OF MANAGING THE DELUGE OF DATA

Today's operational global predictions are completed within an hour for models with about 10 million grid points, 100 vertical layers, and 10 prognostic variables, initialized using 100 million observations per day. These calculations run on hundreds of nodes of general-purpose CPUs offered by vendors in the United States, Asia, and Europe. The ensembles needed for uncertainty quantification multiply both the compute and data burdens. Figure 1 of Bauer et al.[8] illustrates the elements of an operational weather prediction workflow. With the sustained progression of computing power, science, algorithm, and software development, the ECMWF has doubled its horizontal resolution of the global prediction system approximately every eight years. This progress is to be substantially accelerated with digital twins.[9,10]

The challenges in terms of technology adaptation are continuously increasing. With the growing demand to improve the time to information, digital twins are required to substantially improve their time to solution while not exceeding reasonable limits on energy consumption to derive the information. Complex digital twins require sustainable solutions that can be ported to emerging (super-)computing infrastructures, together with supporting resilient prediction and projection services. Solutions are actively pursued on alternative numerical methods, algorithms, programmatic handling of data structures, and domain-specific languages that separate science and the HPC specialization. The latter will be achieved via programming models that support the ease of adaptation to the growing variety of heterogeneous processing, memory, and interconnect technologies. The need to access distributed compute and data resources (and cloud-native infrastructures) increases the demand on those highly specialized weather and climate prediction systems that were originally designed to only run within well-defined network boundaries.

ML may address some but not all of these challenges. ML modules can support the execution speed of Earth system digital twin components through emulation, data compression in communication, and supporting the process of observational quality control and assimilation.[11]

Just as news feeds and social media changed the way we interact with news items, there is a desire to add, consolidate, stream, and reshape regular and timely updates on the constantly evolving climate and weather extreme scenarios. This substantially reduces the time to information, which is anticipated to be 1–24 h for extremes predictions and, perhaps, three-monthly updates for decadal climate prediction and even multidecade projections. Production rates in excess of 1 PB/day challenge classical paradigms and user expectations with respect to both data access and data preservation for long periods of time. This requires intelligently reducing the complexity of the generated data, adding more semantic annotation to metadata, and creating a feed of information rather than an archive.

To tackle this deluge of data, systems and services supporting the new digital twins must use novel approaches that depart from the traditional producer–consumer paradigms of models run on dedicated HPC platforms. It may no longer be affordable to a priori produce large datasets, store them in high-performance storage systems, and delegate the analysis phase to postprocessing tools that simply consume data from that storage. While this approach will still be used, the trend is toward a more datacentric approach, where data consumers are able to intercept the data earlier in the pipeline and, likely, before the data arrive to storage.

To achieve this, data processing pipelines need to become more dynamic and flexible. The earliest interception possible would be for data consumers to inject processing algorithms directly inside of models, thus accessing the high-resolution, high-frequency data right in the memory space of the (parallel) model while it executes. This can be achieved by defining application programming interfaces that implement a separation of concerns without loss of efficiency and with dynamic loading of extensions, in the spirit of how computer-aided engineering tools allow the integration of user-defined functions. This approach is also not dissimilar to using couplers to bring together multiple components of the Earth system to exchange information.

Alternatively, data can be intercepted as they come out of a model, typically in the input–output path, by executing algorithms integrated in pipelines just before writing to storage. This is changing the view toward data as a stream, which can be tapped into and captured as it becomes available. It can be done both by software (e.g., Maestro[g]) or by hardware, as the advent of smart network interfaces has opened the door to in-network processing. Here, again, it is possible to supplement the approach with processes that are user provided and that will consume the data stream before the data hit storage. By doing this, it may even be possible to avoid having to store the whole dataset, focusing storage on different criteria, such as maximizing information content, what is most expensive to recreate, or what is required for policy traceability (provenance), with a view to store the minimal information that allows reconstruction of the larger original dataset.

This focuses on accessing data "just in time," which provides the challenge that all parts of the pipeline need to be ready to consume data. What happens for those components that are not ready? An alternative is to access reduced data once they are already in storage. To achieve this, modern systems focus on data sharing and chunking that the minimum datum necessary to read to obtain a particular piece. If a data consumer only requires information on Europe, there is no need to read the information for the whole globe. The downside of such an approach is that often they require a priori decisions on the sharing or chunking strategy, which may suit some applications well but may be less optimal for other particular data users. This is where novel data cube access technologies are employed, which construct fine-grained indexes that allow almost (if not fully) dynamic chunking and, thus, access information at a higher granularity, potentially combined with the duplication of subselected datasets.

A final avenue may exploit anticipated "gaps" in the dataset that can be filled with ML algorithms specifically trained to bridge them using the available data and pretrained ahead of dataset production. It may be that, in the future, datasets are made up not only of data but also of the tools that (re-)generate the data, raising additional technical challenges, such as reproducibility, provenance, performance, and data governance extensions.

Moreover, a refocus and scaling with the increasing number of applications can be achieved by requiring them to extract and collect the essential information they need. This does not entirely remove the responsibility of the data provider to archive and contribute to the long-term preservation of unique datasets [e.g., ECMWF's fifth generation atmospheric reanalysis of the global climate (ERA-5) reanalysis and observational data records]. Scientific rigor underlying key decisions about Earth will also still require a full provenance trace.

Digital twins armed with the additional power of continuous ML allow essential information to be recreated, potentially at a fraction of the original production cost. This opens the door to fundamentally redesigning the data management of climate information, as is required for adaptation to a continuously changing environment.

CONNECTING MACHINES WITH PEOPLE

The user expectations for digital twins are vast: findable, accessible, interoperable, and reusable data principles and related standards; free and open data, delivered at lightning speed; digital services underpinned with provenance information and trust; geographical diversity and accessibility, serving both public and private entities; and defined limitations, ethical responsibilities, and liabilities associated with its data use for policy and multistructured decision making.

Of course, weather and climate is not the only field producing unprecedented amounts of data. Policy making requires a diverse range of different complex information streams to be merged before a decision can be derived. Applications such as health, energy, transport, environmental protection, and extreme-events preparedness all require the fusion of diverse data sources. Initial steps have been taken to more closely connect the production HPC sites through (data) bridges, which are cloud infrastructure and services feeding data into a federated domain-specific data holding, a data lake. These are still siloed, and, thus, the need exists to extend and federate different information sources further afield into a topical *data space*. The goal is to define common data access patterns, tools and services that connect seamlessly and securely across different data spaces. Among the technological challenges and opportunities are the provision of cloud compute facilities that optimize data access and the delivery of information.

Developing digital twins is a challenging task, and so is connecting the complex data production with applications that collect, compare, fuse, analyze, and translate information from various sources before delivering it through well-defined interfaces to users or software applications. These users may be application owners who have their own user community and institutional priorities, or they may be individual users accessing specific digital twin data. To facilitate co-design, digital twin data streams are explicitly exposed to impact sector applications, and negotiating with application owners

[g]https://www.maestro-data.eu/project/

BY INVITATION OF THE EDITOR-IN-CHIEF

Digital twins of Earth, such as Destination Earth, aim to be a powerful tool for decision and policy making. They can visualize, monitor, and forecast natural and human activity as well as optimize disaster prevention or adaptation strategies. They also give nonexperts full access to data and analytics toolkits using a scalable cloud-based platform. This article shows how high-performance computing (HPC) enables digital twins of Earth, integrating observation, modeling, and simulation, to deliver accurate predictions of future developments in different application areas. HPC is essential in capturing the complex processes of the Earth system. As we face challenges from climate change and environmental degradation, we need to act fast and smart. Digital twins of Earth can help us enhance our resiliency and adaptability in a changing world.

Lorena Barba
Editor-in-Chief

regarding the required level of adaptation and design is conducted on both sides. Software-defined infrastructures enable users to define their workflows with the necessary technical and scientific resources provided by DestinE platforms.

The feedback and control of digital twin scenarios combined with a predefined time to information further exacerbates the challenges. Equally, one must assume a range of users since not every user can launch their own Earth system-defining digital twin, especially considering the large but limited supercomputing resources. However, the success of digital twins for local policy decisions will also be measured in terms of providing access to, e.g., low-income countries, especially as there will likely only be a limited number of Earth system digital twins, running on dedicated supercomputing platforms, predominantly in high-income countries. Given the complexity of the simulations and the vast amounts of data produced, the challenges are funding and how to facilitate access to a share of these dedicated (super-)computing resources to a particular user group with urgent needs [e.g., this could be facilitated through World Meteorological Organisation (WMO) and widening its information system strategy WIS2.0[h]]. To increase the number of users, expensive ML (pre-)training may need to be a shared resource, combined with highly cost-effective inference models that can be executed and manipulated in much larger numbers.

DestinE will offer a variety of user journeys, from traditional data analysis of digital twin data using standard analysis and visualization tools to immersive experiences that allow users to interact with the data. While users can access data through downloads into data servers that cache and redistribute or fuse topical information, direct access or subsetting of data is also attractive due to the large size of digital twin data and the semantic description of data access. Rendering and processing data close to the digital twin data is possible but resource limited, and accessing data through semantic description messages may be preferred. By collaborating with users early in the development phase through specific use cases in the domains of renewable energy management, flood risk management, or air quality management, data interfaces to digital twins are exposed and improved.

In addition, DestinE aims to build interfaces and bridges with selected domains and define standards and options for interoperability to support users who may not be experts in relevant domains but want to create new storylines quickly. Future phases will need to support interfaces for orchestration across different areas of expertise and across different data spaces to meet these demands.

[h]https://library.wmo.int/doc_num.php?explnum_id=4620

ACKNOWLEDGMENTS

The authors would like to thank all of the DestinE contractors and the partners at EUMETSAT and the ESA for the continuing collaboration. The authors would also like to acknowledge the ECMWF staff involved in the diverse DestinE activities. The work presented in this article was carried out with funding by the European Union under the "Destination Earth" initiative. This article is an extended version of an SC22 plenary talk.

REFERENCES

1. F. Tao and Q. Qi, "Make more digital twins," *Nature*, vol. 573, no. 7775, pp. 490–491, Sep. 2019, doi: 10.1038/d41586-019-02849-1.

2. P. Bauer, B. Stevens, and W. Hazeleger, "A digital twin of Earth for the green transition," *Nature Climate Change*, vol. 11, pp. 80–83, Feb. 2021, doi: 10.1038/s41558-021-00986-y.
3. S. A. Niederer et al., "Scaling digital twins from the artisanal to the industrial," *Nature Comput. Sci.*, vol. 1, pp. 313–320, May 2021, doi: 10.1038/s43588-021-00072-5.
4. P. D. Dueben and P. Bauer, "Deep learning to improve weather predictions," in *Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences*, G. Camps-Valls, D. Tuia, X. X. Zhu, and M. Reichstein, Eds. Hoboken, NJ, USA: Wiley, 2021, pp. 204–217, doi: 10.1002/9781119646181.ch14.
5. J. Sillmann et al., "Event-based storylines to address climate risk," *Earth Future*, vol. 9, no. 2, Feb. 2021, Art. no. e2020EF001783, doi: 10.1029/2020EF001783.
6. A. Sanchez-Benitez, H. Goessling, F. Pithan, T. Semmler, and T. Jung, "The July 2019 European heat wave in a warmer climate: Storyline scenarios with a coupled model using spectral nudging," *J. Climate*, vol. 35, no. 8, pp. 2373–2390, Mar. 2022, doi: 10.1175/JCLI-D-21-0573.1.
7. J. Slingo et al., "Ambitious partnership needed for reliable climate prediction," *Nature Climate Change*, vol. 12, pp. 499–503, Jun. 2022, doi: 10.1038/s41558-022-01384-8.
8. P. Bauer, P. Dueben, T. Hoefler, T. Quintino, T. Schulthess, and N. Wedi, "The digital revolution of Earth-system science," *Nature Comput. Sci.*, vol. 1, pp. 104–113, Feb. 2021, doi: 10.1038/s43588-021-00023-0.
9. T. Schulthess, P. Bauer, N. Wedi, O. Fuhrer, T. Hoefler, and C. Schär, "Reflecting on the goal and baseline for exascale computing: A roadmap based on weather and climate simulations," *Comput. Sci. Eng.*, vol. 21, no. 1, pp. 30–41, Jan./Feb. 2019, doi: 10.1109/MCSE.2018.2888788.
10. N. Wedi et al., "A baseline for global weather and climate simulations at 1 km resolution," *J. Adv. Model. Earth Syst.*, vol. 12, no. 11, pp. 1–17, Nov. 2020, doi: 10.1029/2020MS002192.
11. A. Geer, "Learning earth system models from observations: Machine learning or data assimilation?" *Philos. Trans. Roy. Soc. A*, vol. 379, no. 2194, 2021, Art. no. 20200089, doi: 10.1098/rsta.2020.0089.

NILS WEDI is the digital technology lead for Destination Earth (DestinE) at the European Centre for Medium-Range Weather Forecasts (ECMWF), 53175, Bonn, Germany. His research interests include numerical weather prediction, high-performance computing, and Earth system modeling. Wedi received his Ph.D. degree in meteorology from Ludwig Maximilian University of Munich. He is the cochair of the WMO Working Group on Numerical Experimentation. Contact him at nils.wedi@ecmwf.int.

PETER BAUER is the director of DestinE at the ECMWF, 53175, Bonn, Germany, responsible for implementing the ECMWF's contribution to this European Commission-funded flagship action. His research interests include satellite observations and data assimilation, Earth system modeling, and high-performance computing. Bauer received his Ph.D. degree in meteorology from the University of Hamburg. Contact him at peter.bauer@ecmwf.int.

IRINA SANDU is the science lead and appointed director for DestinE at the ECMWF, 53175, Bonn, Germany. Her research interests include boundary layer processes, surface drag, and global storm-resolving simulations. Sandu received her Ph.D. degree in meteorology from Meteo-France and Université Paul Sabatier, Toulouse. Contact her at irina.sandu@ecmwf.int.

JÖRN HOFFMANN is the DestinE application partnership lead at the ECMWF, 53175, Bonn, Germany. His research interests include satellite-derived information and impact sector applications of weather and climate knowledge. Hoffmann received his Ph.D. degree in geophysics from Stanford University. Contact him at joern.hoffmann@ecmwf.int.

SOPHIA SHERIDAN is a communications manager for DestinE and Copernicus Services at the ECMWF, 53175, Bonn, Germany. Her research focuses on demonstrating the impact and showcasing the value of scientific advancement, connecting people with science. Sheridan received a bachelor's degree in archaeology from the University of Bangor. Contact her at sophia.sheridan@ecmwf.int.

RAFAEL CERECEDA is a science and content writer for Copernicus and DestinE at the ECMWF, 53175, Bonn, Germany. His research focus is communicating the fast-evolving digital environment. Cereceda received his master's degree in journalism from Universidad Complutense de Madrid. Contact him at rafael.cereceda@ecmwf.int.

TIAGO QUINTINO is the head of the Development Section at the ECMWF, Reading, RG2 9AX, U.K. His research interests include high-performance computing, cloud technology, and data analytics for exascale weather forecasting. Quintino received his Ph.D. degree in applied sciences from Katholieke Universiteit Leuven. Contact him at tiago.quintino@ecmwf.int.

DANIEL THIEMERT is the program integration lead for DestinE at the ECMWF in Reading, RG2 9AX, U.K. He is responsible for the DestinE program management and budget/resource planning. Thiemert received his Ph.D. degree in computer science from the University of Reading. Contact him at daniel.thiemert@ecmwf.int.

THOMAS GEENEN is the technology partnership lead for DestinE at the ECMWF, 40128, Bologna, Italy. He has a background in solid earth system modeling from the University of Utrecht and brings 20 years of experience in HPC and HPDA. Contact him at thomas.geenen@ecmwf.int.