Frontiers in Science

TYPE Frontiers in Science Viewpoint
PUBLISHED 05 March 2024
DOI 10.3389/fsci.2024.1371049

EDITED AND REVIEWED BY
Justin Sheffield,
University of Southampton,
United Kingdom

*CORRESPONDENCE
Christina Tague
tague@ucsb.edu

RECEIVED 15 January 2024
ACCEPTED 26 February 2024
PUBLISHED 05 March 2024

CITATION
Tague C. Digital twins of the Earth:
can they keep up?
Front Sci (2024) 2:1371049.
doi: 10.3389/fsci.2024.1371049

Digital twins of the Earth: can they keep up?

Christina Tague*

Bren School of Environmental Science and Management, University of California, Santa Barbara, Santa Barbara, CA, United States

KEYWORDS

digital twins, hydrology, uncertainty, ecohydrology, observation

A Viewpoint on the Frontiers in Science Lead Article

A Digital Twin of the terrestrial water cycle: a glimpse into the future through high-resolution Earth observations

Key points

- First-generation Earth system digital twins, such as the Digital Twin Earth (DTE) for hydrology, create important opportunities for "learning by doing" that will ensure DTEs evolve to provide credible, reliable, and useful information.
- Recent DTEs for hydrology demonstrate the complexity of the cyberinfrastructure needed to support the integration of a diversity of high-resolution datasets—often through machine learning techniques—while also providing initial insights into how critical errors in these approaches might be identified.
- To remain useful, DTEs will need to be able to continuously evolve—this will require innovations in visualization, cross-disciplinary collaboration, and complementary tools that draw from advances in relevant research communities.

Recent advances in high-resolution spatial, temporal, and spectral datasets, in machine learning techniques and algorithms, and in large-scale data storage and analysis cyberinfrastructure have set the stage for the concept of a "digital twin" within Earth system science. Several recent papers have advocated for the construction of digital twins—data-model replicates of the biophysical environment—as these can be used for prediction tasks and the development of "what-if scenarios" (1, 2). Digital twin Earths (DTEs) utilize exciting cyberinfrastructure that strategically links multiple diverse observation and remote sensing datasets with physics-based models in new ways. This cyberinfrastructure is facilitated by machine learning and other software engineering and data science tools. The realization of digital twins, however, requires substantial human, cyberinfrastructure, and financial investment, and thus actual examples remain relatively rare. Brocca et al. (3) present a detailed example of a realized DTE, thus offering an opportunity to highlight emerging best practices and potential challenges.

A central feature of digital twins is the integration of high-resolution spatial and temporal data when creating the digital biophysical environment. This can create numerous data management complexities; one of the most challenging to address is dealing with potential errors and uncertainty in these datasets. All observational data is vulnerable to measurement errors, and the use of remote sensing data must also account for uncertainty in the translation of signals into meaningful variables. Brocca et al. (3) present several clear and well-conceived examples of evaluating dataset within the context of their DTE and, thus, offer an important precedent for DTE best practices. These are important first steps for building a comprehensive strategy to facilitate ongoing evaluation and communication of uncertainty across multiple datasets and understanding what this means for DTE predicted variables and generated "what-if" scenarios. Ultimately, the scientific community will need to ensure that DTEs can readily incorporate new datasets and readily reveal—and when needed, highlight—key data uncertainties. This will require continued thoughtful back-end software design to facilitate updates with new datasets and front-end interface design that can draw attention to not only a landscape of hydrologic variables but also to the assumptions and uncertainties associated with them.

A second important feature of digital twins is the coupling of multiple, diverse datasets with process-based or physical models. Here again Brocca et al. (3) offer a useful illustration of using hydrologic models in conjunction with observational datasets within DTEs, via both assimilation and calibration. As with observational datasets, one of the challenges with incorporating process-based models in DTEs is ensuring their veracity and facilitating the incorporation of advances in these models (or new, "better" models). Some physics-based models are well established, such as core models of snow melt. But other components of the hydrologic cycle, particularly models of how water is stored and how it moves through the subsurface, remain areas of active research. Research on hydrologic theory and physics-based models is focused on how to represent the structure and dynamics of subsurface flowpaths. These include strategies for accounting for preferential flowpaths or considering the subsurface as complex networks for storage that fill and spill (4, 5). While new remote sensing datasets, such as ground-penetrating radar and electrical resistivity tomography, provide insights into the subsurface, they are still insufficient for fully characterizing subsurface flow and storage (6). There are also advances in geomorphic theory that suggest what this structure might look like from a geoclimatic perspective (7). The use of datasets such as river discharge and observed evapotranspiration estimates can hide the need for actual theory and representation of the subsurface. However, as conditions change, understanding subsurface structure may become increasingly important (8, 9). DTEs that can incorporate new datasets and advances in theories to better represent subsurface storage and flowpaths will improve the realism of watershed representation. They will also ensure that the use of DTEs as a "sandbox" for exploring the hydrologic implications of land management choices and/or climate change account for how the subsurface can influence these sensitivities.

How DTEs account for ecologic processes presents some similar issues, even for DTEs like that of Brocca et al. (3) that are focused on hydrologic behavior such as flooding and soil moisture. Plant water use remains a fundamental component of the hydrologic cycle, which is complicated by the dynamic response of vegetation to water availability (10). While the DTE of Brocca et al. (3) updates vegetation using remote sensing products, these approaches do not necessarily account for systematic climate-driven changes to plant structure (such as biomass, leaf area, rooting depths, and density) and demography that might change future water use. However, ecology and ecophysiology science are advancing new theories on how vegetation adaptations to climate may shift structure, composition, and function beyond historic norms (11, 12). Recent work also includes evidence of within-species plasticity in water use characteristics (13) that might be particularly relevant for the current DTE of Brocca et al. (3). Consideration of ecohydrology as a coupled evolving system would have implications for physically based sub-models within a DTE and for which datasets may be most useful for characterizing vegetation (14).

As DTEs evolve and new ones are developed, consideration of best practices and how to confront key challenges will be critical to ensure that DTEs encode our best and evolving understanding of Earth systems. Which datasets to use, how to quantify their uncertainty, how to best link them with physical models, and even what physical processes and variables to include and at what scale are not questions that have conclusive answers. New observations, methods, and theories in Earth system science continue to evolve in parallel with DTE development. Thus, DTEs must be designed in a way that ensures updating is relatively easy and that it is easy for users to learn what the underlying data/models are. User interfaces and visualizations that expose underlying assumptions will be key (15). Perhaps the hardest challenge for ensuring that DTEs reflect the best available science is for DTE architects to keep pace with the rapid advancement of technique, data products, and theory in the multiple disciplines relevant to hydrology and, more broadly, Earth system science. To remain current, DTEs will require successful collaboration across multiple disciplines that go beyond the core DTE design team and engage a broader community (16). Best practices in team science and collaboration (17) can help; however, the rapid pace of scientific advances will remain a challenge. To address this information overload and complement DTEs, a knowledge-provisioning system that can identify new Earth system science relevant to the DTE (18) may be essential. DTEs, like all new technologies, have great potential, but now is the time to ensure that best practices in their design, evolution, and application maximize their benefits.

Acknowledgments

I thank Sloane Stevenson for her review of the manuscript.

Statements

Author contributions

CT: Conceptualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The author received funding from the National Science Foundation's Critical Zone Collaborative Research Cluster (Project #2012821) to support her review of critical zone science synthesis activities. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Conflict of interest

The author declares that the research was conducted in the absence of financial relationships that could be construed as a potential conflict of interest.

The author declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

1. Rigon R, Formetta G, Bancheri M, Tubini N, D'Amato C, David O, et al. HESS Opinions: participatory Digital eARth Twin Hydrology systems (DARTHs) for everyone–a blueprint for hydrologists. *Hydrol Earth Syst Sci* (2022) 26:4773–800. doi: 10.5194/hess-26-4773-2022

2. Li X, Feng M, Ran Y, Su Y, Liu F, Huang C, et al. Big data in earth system science and progress towards a digital twin. *Nat Rev Earth Environ* (2023) 4:319–32. doi: 10.1038/s43017-023-00409-w

3. Brocca L, Barbetta S, Camici S, Ciabatta L, Dari J, Filippucci P, et al. A Digital Twin of the terrestrial water cycle: a glimpse into the future through high-resolution Earth observations. *Front Sci* (2024) 1:1190191. doi: 10.3389/fsci.2023.1190191

4. Grant GE, Dietrich WE. The frontier beneath our feet. *Water Resour Res* (2017) 53:2605–9. doi: 10.1002/2017WR020835

5. Fan Y, Clark M, Lawrence DM, Swenson S, Band LE, Brantley SL, et al. Hillslope hydrology in global change research and earth system modeling. *Water Resour Res* (2019) 55:1737–72. doi: 10.1029/2018WR023903

6. Parsekian AD, Singha K, Minsley BJ, Holbrook WS, Slater L. Multiscale geophysical imaging of the critical zone. *Rev Geophys* (2015) 53:1–26. doi: 10.1002/2014RG000465

7. Pelletier JD, Barron-Gafford GA, Gutiérrez-Jurado H, Hinckley ES, Istanbulluoglu E, McGuire LA, et al. Which way do you lean? Using slope aspect variations to understand Critical Zone processes and feedbacks. *Earth Surf Processes Landf* (2018) 43:1133–54. doi: 10.1002/esp.4306

8. Kirchner JW. Getting the right answers for the right reasons: linking measurements, analyses, and models to advance the science of hydrology. *Water Resour Res* (2006) 42:2465. doi: 10.1029/2005WR004362

9. Fatichi S, Or D, Walko R, Vereecken H, Young MH, Ghezzehei TA, et al. Soil structure is an important omission in Earth System Models. *Nat Commun* (2020) 11:522. doi: 10.1038/s41467-020-14411-z

10. Newman BD, Wilcox BP, Archer SR, Breshears DD, Dahm CN, Duffy CJ, et al. Ecohydrology of water-limited environments: a scientific vision. *Water Resour Res* (2006) 42:W06302. doi: 10.1029/2005WR004141

11. Anderegg WRL, Kane JM, Anderegg LDL. Consequences of widespread tree mortality triggered by drought and temperature stress. *Nat Clim Change* (2013) 3:30–6. doi: 10.1038/nclimate1635

12. Xu X, Trugman AT. Trait-based modeling of terrestrial ecosystems: advances and challenges under global change. *Curr Clim Change Rep* (2021) 7:1–13. doi: 10.1007/s40641-020-00168-6

13. Ibsen PC, Santiago LS, Shiflett SA, Chandler M, Jenerette GD. Irrigated urban trees exhibit greater functional trait plasticity compared to natural stands. *Biol Lett* (2023) 19:20220448. doi: 10.1098/rsbl.2022.0448

14. Massoud EC, Hoffman F, Shi Z, Tang J, Alhajjar E, Barnes M, et al. Perspectives on artificial intelligence for predictions in ecohydrology. *Artif Intell Earth Syst* (2023) 2: e230005. doi: 10.1175/AIES-D-23-0005.1

15. Tague C, Frew J. Visualization and ecohydrologic models: opening the box. *Hydrol Processes* (2021) 35:e13991. doi: 10.1002/hyp.13991

16. Arora B, Kuppel S, Wellen C, Oswald C, Groh J, Payandi-Rolland D, et al. Building Cross-Site and Cross-Network collaborations in critical zone science. *J Hydrol* (2023) 618:129248. doi: 10.1016/j.jhydrol.2023.129248

17. Carey CC, Ward NK, Farrell KJ, Lofton ME, Krinos AI, McClure RP, et al. Enhancing collaboration between ecologists and computer scientists: lessons learned and recommendations forward. *Ecosphere* (2019) 10:e02753. doi: 10.1002/ecs2.2753

18. Tague C, Brandt WT. Critical zone science in the Western US—too much information? *Front Water* (2023) 5:1226612. doi: 10.3389/frwa.2023.1226612