Towards intelligent revolution in Earth surface system modeling with deep learning

Min Chen[1,22,24,†], Zhen Qian[1,23,24], Niklas Boers[2,25,26], Anthony J. Jakeman[3], Albert J. Kettner[4], Martin Brandt[5], Mei-Po Kwan[6], Michael Batty[7], Wenwen Li[8], Rui Zhu[9], Wei Luo[10], Daniel P. Ames[11], Susan M. Cuddy[12], Sujan Koirala[13], Fan Zhang[14], Carlo Ratti[15], Michael C. Barton[16], Jian Liu[1,23,24], Teng Zhong[1,23,24], Junzhi Liu[17], Yongning Wen[1,23,24], Songshan Yue[1,23,24], Zhuo Sun[1,23,24], Zhixin Zhang[18], Zhiyi Zhu[1,23,24], Zaiyang Ma[1,23,24], Yuanqing He[1,23,24], Kai Xu[1,23,24], Chunxiao Zhang[19], Jian Lin[20], Hui Lin[27], and Guonian Lü[1,23,24,†]

[1] Key Laboratory of Virtual Geographic Environment (Ministry of Education of PRC), Nanjing Normal University, Nanjing, Jiangsu, China

[2] School of Engineering & Design, Technical University of Munich, Munich, Germany

[3] Fenner School of Environment and Society, Australian National University, Canberra, Australia

[4] Institute of Arctic and Alpine Research (INSTAAR), University of Colorado, Boulder, CO, USA

[5] Department of Geosciences and Natural Resource Management, University of Copenhagen, Copenhagen, Denmark

[6] Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, China

[7] The Bartlett Centre for Advanced Spatial Analysis (CASA), University College London (UCL), London, UK

[8] School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, USA

[9] Systems Science Department, Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), Singapore

[10] Department of Geography, National University of Singapore, Singapore

[11] Department of Civil and Construction Engineering, Brigham Young University, Provo, UT, USA

[12] CSIRO Environment, Canberra, Australia

[13] Max Planck Institute for Biogeochemistry, Jena, Germany

[14] Department of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong SAR, China

[15] Department of Urban Studies and Planning, Massachusetts Institute of Technology, Cambridge, USA

[16] School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, USA

[17] Center for the Pan-Third Pole Environment, Lanzhou University, Lanzhou, China

[18] College of Geography & Marine, Nanjing University, Nanjing, Jiangsu, China

[19] School of Information Engineering, China University of Geosciences in Beijing, Beijing, China

[20] Department of Geography and Resource Management, The Chinese University of Hong Kong

[21] School of Geography and Environment, Jiangxi Normal University, Nanchang, China

[22] International Research Center of Big Data for Sustainable Development Goals, Beijing, China

[23] State Key Laboratory Cultivation Base of Geographical Environment Evolution, Nanjing, Jiangsu, China

[24] Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing, Jiangsu, China
[25] Potsdam Institute for Climate Impact Research, Potsdam, Germany
[26] Global Systems Institute and Department of Mathematics, University of Exeter, Exeter, UK
[27] Key Laboratory of Poyang Lake Wetland and Watershed Research, Ministry of Education, Jiangxi Normal University, Nanchang, China

†e-mail of Min Chen: chenmin0902@njnu.edu.cn

†e-mail of Guonian Lü: gnlu@njnu.edu.cn

Abstract | The Earth surface system and its dynamics are changing through nature-human interactions. Earth surface system modeling (ESSM) is essential for understanding Earth surface processes pertaining to the past, present and future, and for assisting in decision-making. Deep learning (DL), with its outstanding strength for data-driven inference, shows promise in assisting ESSM by exploiting information from big observational data. In this Perspective, we discuss current ESSM demands and DL potentials before examining hybrid ESSM, a new research paradigm that integrates DL strengths into ESSM. By overcoming subjective bias and deployment problems in current integration processes, we envision an intelligent revolution in ESSM. We illustrate a conceptual framework to automatically generate customized, scalable and accurate solutions for given ESSM tasks, based on modeling-related knowledge and DL strengths. We conclude by discussing potential prospects for ESSM when integrated with DL for identifying pathways toward a sustainable future.

Introduction

The Earth surface system encompasses dynamics on spatial scales ranging from sub-millimeter to global and in temporal scale from milliseconds to billions of years[1,2]. The Earth surface system consists of various components, such as hydrological, geological, (near-surface) atmospheric, biological and social subsystems (Fig. 1), which preserve interconnected and inter-constrained interactions through energy fluxes, material fluxes and information fluxes[3,4]. To understand the underlying mechanisms and anticipate chain reactions, ancient philosophers to current scientists have studied the interactions between nature and the human realm[5,6].

Based on computational techniques and mathematical models (typically, physically, (semi-)empirically or statistically based), Earth surface system modeling (ESSM, Fig. 1) is a primary tool for representing and quantifying the spatiotemporal variations and internal interactions of the Earth surface across the past, present and future[7–9]. The scientific lifecycle of ESSM can be generally described as having five methodological stages, namely (i) problem definition and contextualization, (ii) data preparation and processing, (iii) model development and integration, (iv) model evaluation and optimization, and (v) model simulation and application[10,11]. These stages may need to be attended to iteratively, and all are important for ensuring that the key processes are addressed and the modeling is suitable for the purpose[12,13]. However, ESSM is confronted with technical challenges due to the vast volume of data available, creating analytical barriers and necessitating the adoption of sophisticated technologies to overcome computational bottlenecks[14,15].

Fgiure1. **An illustration of integrating Earth surface system modeling and deep learning to analyze current scientific challenges.** The figure shows the various subsystems of the Earth surface system and how they connect and interact. The Earth surface system dynamics and interactions can be interpreted and predicted by Earth surface system modeling and deep learning methods to better understand frontier challenges such as climate change, natural resource exploitation, health and environment, and the sustainable city.

Deep learning (DL, Fig. 1), using the power of deep neural networks for prediction accuracy, computational efficiency, and the ability to process multimodal data, has revolutionized several research fields, including computer vision, natural language processing, and protein structure prediction[16]. This data-driven approach has also found applications in geosciences[17–19], demonstrating its potential to address the analytical and computational challenges faced by ESSM research[20,21]. However, the data-intensive nature of DL has inherent "black box" drawbacks due to its underlying abstract formalisms, whereas ESSM relies more on process-based and interpretable representations. Moreover, Earth sciences face unique challenges arising from the heterogeneous and noisy observed data, which often yields an incomplete view of Earth surface processes[21]. Despite the abundance of raw data, labeled and preprocessed data are scarce, mainly due to technical barriers and labor-intensive processes[22]. In light of the "bitter lesson's" emphasis on data quality[23], it is challenging for DL models to recognize patterns and generate trustworthy trends from noisy data with few labels without adding prior domain expertise and physical principles[20,24].

Hybrid ESSM, which combines the strengths of ESSM and DL, is a current research trend that has resulted in groundbreaking discoveries (for example, emulating Earth surface processes in high resolution[25,26]) and an improved understanding of frontier challenges (Fig. 1)[12–14]. While enhancing the efficiency of analyzing from observational data and accelerating discovery in ESSM[27,28], hybrid ESSM has also broadened the application range of DL, such as information extraction from remote sensing imagery and climate variable prediction[21]. However, existing research has currently focused more on combining approaches at the model-integration level, rather than adopting a holistic approach that encompasses the modeling lifecycle; the latter helps to better understand and solve given tasks. The potential for subjective bias towards one paradigm over the other can lead to an inadequate balance between the two paradigms, potentially impeding their successful integration. In addition, the incompatibility of model deployment can result in computational bottlenecks, posing another substantial obstacle.

In this Perspective, we discuss the challenges of existing ESSM research from a geographical perspective, as well as the opportunities presented by DL. Further, integration modes and shortcomings of hybrid ESSM are examined. Based on the modeling-related knowledge and the DL strengths, we propose a conceptual framework for intelligently managing the ESSM lifecycle and investigate a potential application case, with the aim of reducing current technical barriers. Finally, we look at future directions toward advancing ESSM research through its integration with DL.

Challenges of current ESSM

Numerous process-based models have been developed and applied in ESSM throughout the evolution of Geosciences. In order to analyze more comprehensive issues involving numerous processes, communities have developed a series of integrated models based on ESSM that can depict interactions among multiple subsystems[29]. Table 1 lists prominent modeling applications in distinct domains. As indicated below, we have identified four significant challenges that ESSM is currently facing.

Table 1 | Example conventional ESSM approaches and new DL options to scientific problems in various domains.

Domain	Scientific challenge	Example conventional ESSM approaches	DL-integrated options
Hydrological system	Rainfall-runoff simulation	SAC-SMA[30]	LSSVM-HHO[31] (multilayer Perceptron (MLP) based)
	Groundwater modeling	MODFLOW[32]	CNN-BiLSTM[33] (convolutional neural network (CNN) and long short term memory network (LSTM) based)
Geological system	Soil erosion modeling	WEPP[34]	ANFIS[35] (MLP based)
	Sediment estimation	SEDD[36]	SediNet[37] (MLP based)
Atmospheric system	Air quality assessment	Gaussian Plume Model[38]	AQC-Net[39] (CNN based)
	Weather prediction	WRF[40]	Graph neural network[41]
Biological system	Forest carbon estimation	SEIB-DGVM[42]	FLUXCOM[43] (MLP based)
	Wetland monitoring	WSM[44]	Bootstrap-BP neural network[45] (MLP based)
Social system	Epidemic spread modeling	Susceptible-infected-susceptible Model[46]	LSTM[47]

	Human migration simulation	Gravity Model[48]	Deep Gravity model[49] (MLP based)

Completeness of understanding problems. Understanding the dynamics of the Earth surface system, which exhibit self-organization, emergent, and hierarchical properties, should consider the intrinsic interactions and feedbacks between different subsystems[50,51]. In ESSM, macroscopic problems are often decomposed hierarchically into less complex and more manageable ones to facilitate analysis and problem-solving, while underlining the importance of interactions and emergent properties across different scales[52,53]. Yet, some current methodologies in ESSM, particularly those designed for large-scale simulations, may not fully capture the intrinsic connections among related subsystems, potentially resulting in reductionist approach[54,55]. Furthermore, these methods could lead to incomplete understanding and computational challenges. Specifically, decomposed subproblems with too few geographic objects (for example, landforms, vegetation, river) in subsystems might not provide a comprehensive view of the relevant Earth surface states[7,56]; but in contrast, those with a large number of geographic objects may not necessarily address the nonlinearity problem effectively and could introduce additional computational complexities[4457,58].

Capability of handling big data. A plethora of sensors continue to proliferate unstructured observational data that capture states, fluxes and interactions of the Earth's surface[59]. They include Earth-observation satellites, the global positioning system, in situ observations, and social media; all of these generate quintillions of bytes every day[60,61]. Although this data availability has created numerous opportunities for ESSM studies, it has also led to unprecedented technological obstacles because of Big data's "five Vs" characteristics, namely, volume, variety, veracity, velocity and value[62,63]. It is generally difficult to fully process the various data sources and further extract deep-level patterns, let alone discover knowledge from them, utilizing conventional ESSM approaches[64].

Precision of modeling dynamics. The construction of process-based models, particularly when involving multiple subsystems and (semi-) empirical representations, is largely dependent on an expert's perspective[65]. So, model architecture and configuration are potentially affected by subjectivity and are prone to bias, errors and unexpected simulation results[66]. This is also accentuated when the derived models consist of physical, (semi-)empirical, or statistical models that may struggle with effectively addressing complex nonlinear dynamics[67,68]. Although data assimilation strategies can enhance the performance of these models, the pace of creating data frequently far exceeds the ability of models to assimilate it sensibly[20].

Efficiency of computational technology. The computational efficiency of process-based models is crucial, particularly for high resolution or (near-) real-time modeling (for example, natural disaster assessment), where delays in results caused by large time overheads could potentially impact decision-making processes[69,70]. Hardware-wise, current ESSM research often relies on multiple central processing units (CPU)-based computers or supercomputers, which have been outperformed by expanding computational demands[71]. A three-year study of fine-grained climate simulations on supercomputers shows that GPUs outperform CPUs by at least an order of magnitude during high-resolution simulations[72]. Regarding software, ESSM lifecycle processes typically require manual operations or intermediate data transfers, which can impede the computing pipeline. In addition, some models with computationally expensive modules, such as the solution of optimization problems and partial differential equations, necessitate time-intensive iterative simulations.

DL strengths

As a specific subfield of artificial intelligence, DL comprises a large class of approaches based upon the different variations of deep neural network architectures. For example, convolutional neural networks, architectures that focus on local connections through multi-dimensional convolutions, are often used to extract patterns from various data modalities (for instance, 1D convolutions for sequences, 2D convolutions for images, and 3D convolutions for videos)[16]. Recurrent neural networks, particularly those equipped with memory cells known as Long Short-Term Memory (LSTM) networks[73], are commonly adept at learning features and long-term dependencies from sequential inputs[74]. More sophisticated networks, like graph neural networks, generative adversarial networks, and transformers, expand the applicability of neural networks beyond relatively specific uses and demonstrate greater flexibility and adaptability for various tasks[41,75,76]; in particular, transformers have been shown to be applicable across diverse purposes with outstanding performance in geoscientific applications, such as modeling spatio-temporal patterns of climate variables[77] and tectonic plate movement[78].

Compared to conventional process-based models, deep neural networks generally exhibit superior prediction performance in terms of fitting observational data[16]. Although it is important to acknowledge that these networks typically have limited interpretability for understanding decision processes[79,80], with the research community actively working to address these shortcomings, the characteristics of deep learning still pave the way for data-driven discovery of patterns in Earth surface system dynamics. Table 1 contains some existing examples of DL-integrated ESSM options for the different domains. Since the introduction of DL in 2006, most research areas have witnessed its development, with the number of published papers related to these methods increasing annually (see Fig.2). In some fields, the volume of papers based on DL published in the last three years is nearly half of the total published using these methods over the past decade or more. On a broader note, the opportunities that DL brings to mitigate the challenges of ESSM can be seen from four perspectives, as described in the following sections.

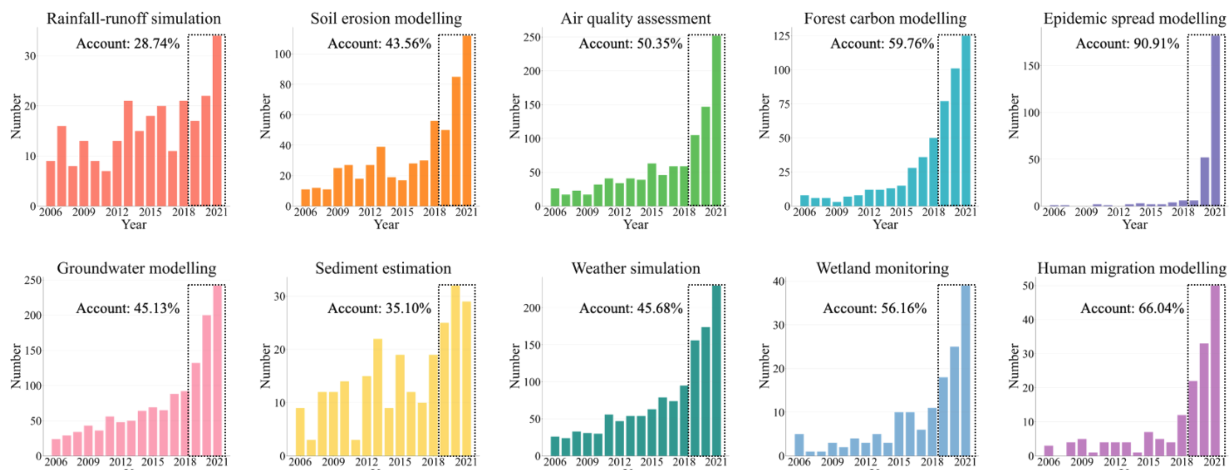


Figure 2. **Quantity of published research papers utilizing deep learning in various subsystems from 2006 to 2021.** Each subfigure displays the proportion of articles published in the last three years relative to the total number of articles in that category. The statistical data are collected from Web of Science.

Maximum use of multimodal data. Data derived across space and time are often characterized by multimodalities; that is, they are multi-source, heterogeneous, unstructured, or multi-temporal[81]. Integrating information from various modalities into a homogeneous space helps uncover distinctive characteristics and explain the observed processes[82]. Techniques for multimodal data fusion are numerous. Those techniques that rely heavily on manual encoding with domain-specific expertise inevitably impair the fusion results[83]. In contrast, deep neural networks can adapt to unstructured multimodal data and uncover complicated correlations among them[84]. The ability to tackle the challenges of ESSM using this aspect of DL is a major advantage. For instance, DL-based approaches can fuse the various multimodal data derived from decomposed problems, thereby affording an efficient way to understand Earth's surface processes more comprehensively.

Self-adaptive feature representation. The data generated by natural laws exhibit considerable uncertainty and high dimensionality[20,85]. To extract information from and understand such data, scientific communities have a strong interest in representing their features. Traditional methods like Scale-invariant Feature

Transform (SIFT), Term Frequency - Inverse Document Frequency (TF-IDF), and Principal Component Analysis (PCA) commonly extract low- or mid-level features and are only suitable for certain workloads[86]. In contrast, DL-based approaches have received attention in geoscientific applications due to the self-adaptive feature learning mechanism (commonly based on supervised learning and labeled data). Specifically, deep neural networks can uncover patterns and relationships from data, such as interpreting various objects within complex backgrounds in observed images, that may be challenging to formulate using traditional methods based on our a priori knowledge[87,88]. This facilitates the extraction of deep-level features without tedious feature engineering. Further, unsupervised or self-supervised approaches can automatically adapt to latent domains in heterogeneous data at a fraction of manual and computational cost[89,90]. Modelers can use pre-trained models on public datasets like ImageNet[91] to transition to geoscientific applications, reducing time-consuming labeling efforts.

***Superior fitting precision*.** DL-based approaches perform well in complex Earth surface system dynamics as universal functional approximators[92]. For example, DL-based forecasting or nowcasting of climate variables (e.g., precipitation, temperature and humidity) can achieve better results, spatially and temporally, including the exact timing, location and intensity[90,93]. On the other hand, traditional models such as optical flow frequently struggle to effectively capture nonlinear climate dynamics (for example, moist convection and cloud formation)[93,94], which can be attributed to the separation of internal processes and the presence of nonoptimizable parameters[95]. Some studies have also attempted to shift the paradigm for specific tasks to enhance their performance, such as visual question-answering for geographic scenes[96], synthetic spatiotemporal data generation[97], and extreme weather prediction[98], that seem impossible for traditional process-based models through customized networks. All of the preceding examples rely on the ability of deep neural networks to fit with superior precision. There is however one large caveat to recognize here in that, as with all modeling, the parameterization of deep neural networks depends on the training dataset(s), which greatly affects fitting performance[99]. Biases embedded in training data could get encoded into the model, making it essential to consider data quality and the conditions that affect their parameterizations and extracted patterns[100,101].

High inferencing speed. It is undeniable that training deep neural networks require a significant amount of time[102]. However, the inferencing speed of trained networks can be orders of magnitude faster than conventional process-based models[103], such as numerical methods, which frequently require lengthy simulation durations to yield reliable outcomes[72,104]. The computational efficiency of these conventional models can be significantly enhanced with trained networks as a substitute[17]. End-to-end network architecture and parallel computing explain inferencing's computational advantage. First, end-to-end setups enable networks to learn complex representations of data, from inputs to targets, by feeding given data directly without manual manipulations, thereby being highly beneficial for large-scale simulation[105]. Second, the data in deep neural networks are usually structured as a couple of tensors or matrices, which is suitable for parallel computation[106]. The resulting inferencing speed can be increased by several orders of magnitude with GPUs and TPUs[107].

Integrating ESSM and DL

The integration of ESSM and DL offers a promising avenue for advancing our understanding of Earth surface system dynamics. While these two approaches have distinct research paradigms—theory-simulation-driven and data-driven—they complement each other in principle[28]. ESSM offers a strong theoretical foundation for interpreting and representing Earth surface processes but may struggle to handle complex dynamics in the context of big observational data. Conversely, DL excels at uncovering information and fitting trends in large datasets, though it lacks interpretive equations and physical constraints. Hybrid ESSM leverages the strengths of both approaches, demonstrating enhanced prediction and interpretability capacities, potentially expediting the discovery of underlying Earth surface system dynamics and interactions[108,109].

Existing hybrid ESSM research primarily focuses on integrating process-based models with deep neural networks during the stage of model development and integration in the modeling lifecycle. The main integration modes can be categorized into three fundamental modes: the cascading mode, the parallel mode, and the embedding mode (Fig. 3). It is worth noting that complex tasks often require a combination of these fundamental modes.

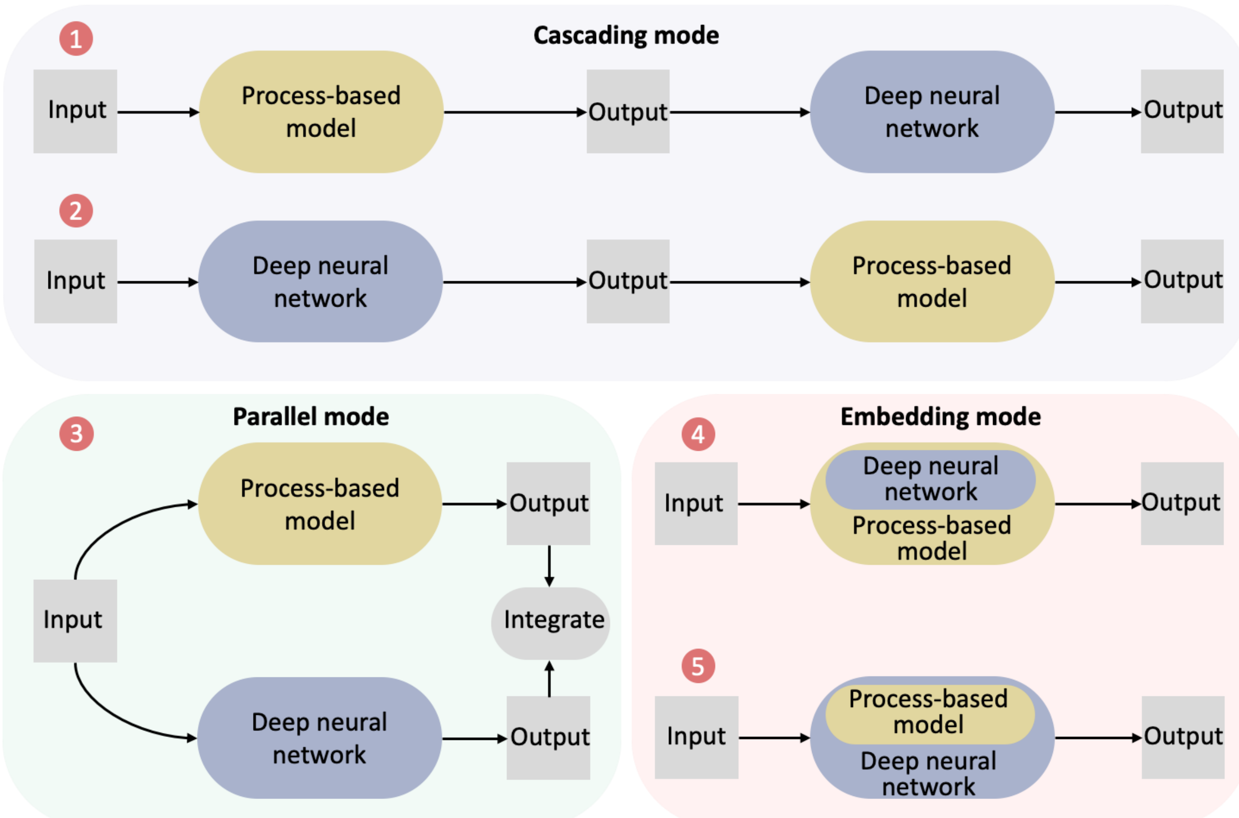


Figure 3. **Computational logics of hybrid models.** The cascading mode is a computational pipeline consisting of process-based models and deep neural networks that runs sequentially and transmits intermediate results. There are two cases according to the sequential order of the models. The parallel mode is when both types of models are run simultaneously. The embedding mode is when the two types of models are embedded into each other's models as plug-in modules. According to the embedding relationship, they can also be divided into two cases.

Cascading mode. The cascading mode is a computational pipeline consisting of process-based models and deep neural networks that run sequentially and transmit intermediate results. This mode has two cases.

In the first case, the process-based model is executed before the deep neural network (diagram 1 in Fig. 3). Using a process-based model to produce training data or perform feature engineering for a deep neural network and the latter's ability to downscale the output variables of the former are two typical functions. For instance, process-based models can filter high-quality samples based on physics-based criteria or construct simulated datasets for training deep neural networks to achieve high prediction accuracy with less ground truth data[110,111]. Moreover, deep neural networks can statistically downscale the coarse outputs of process-based models, which is crucial for predicting climate variables[112,113] and reconstructing real-world landscapes[114,115].

In the second case, the deep neural network is utilized first, followed by the process-based model (diagram 2 in Fig. 3). As an example, process-based models can constrain or refine deep neural network outputs to adhere to physical mechanisms[116,117]. In addition, deep neural networks can be used to calibrate process-based models to reduce parameterization complexity when solving partial differential equations[118,119].

Parallel mode. In the parallel mode, process-based models and deep neural networks are executed concurrently (diagram 3 in Fig. 3). This mode has three practical uses due to its parallel nature: (i) solving complex issues by dividing and conquering, (ii) processing multimodal data and (iii) parallel computing. Specifically, the divide-and-conquer strategy, generally built for decomposed sub-problems, simultaneously employs the process-based model and deep neural network to address the challenges at which they excel[120,121]. Also, process-based models generally process data in specific file formats (e.g., Shapefile and NetCDF) more efficiently than deep neural networks in terms of preprocessing and encoding these raw datasets. Therefore, from the standpoint of computational efficiency, it is promising to use process-based models or deep neural networks to process the data they can handle most efficiently while performing tasks involving heterogeneous data sources[122,123]. Parallel computing cannot only employ supercomputer technology to boost computational performance[124] but also partition the modeling environment, thereby avoiding incompatibilities caused by heterogeneous computing resources between process-based models and deep neural networks[125,126].

Embedding mode. The embedding mode allows process-based models and deep neural networks to be plug-and-play components[127–129]. Specifically, the two approaches will be seen as plug-ins, complementing one other. The embedding mode can be further subdivided into two cases.

The first case involves incorporating deep neural networks as surrogate modules into process-based models (diagram 4 in Fig. 3). The trained deep neural networks can be neural surrogates or solvers for difficult-to-compute submodules, such as based on partial differential equations[130], optimization procedures[131], and high-dimensional tasks[118]. Thus, the local modules of process-based models can be automatically parameterized and modified[132], thereby improving computational efficiency and accurately solving complex systems[133,134].

The second case refers to the integration of process-based models into deep neural networks (diagram 5 in Fig. 3) to incorporate physical mechanisms and principles and construct physics-informed architectures[135], such as Physics-Informed Neural Networks (PINNs)[92]. For example, designing specific loss functions to optimize networks is a straightforward and effective way to constrain inferred results to confirm to domain-specific understanding[136]. Some studies have investigated methods for determining the network's structure (e.g., hidden layers) based on domain laws or physical techniques. Although this is not an easy task, groundbreaking results have been achieved, such as neural ordinary differential equations[137] and geographically weighted artificial neural network[138]. In addition, another promising application of research is the incorporation of physical restrictions into deep neural networks to determine new equations that characterize Earth surface dynamics[139].

Shortcomings

Despite many years of sustained research, hybrid ESSM is still in its infancy. Highly heterogeneous data, insufficient ground truth data, and low interpretability of outcomes have been previously described as the main challenges[20]. This section examines further theoretical and practical shortcomings in existing hybrid ESSM studies, to identify opportunities for significant improvements in hybrid ESSM capabilities.

Restricted integration scenarios. Existing hybrid ESSM studies concentrate mainly on model-level integration. Nevertheless, the ESSM lifecycle is more comprehensive and generally includes the five indispensable stages in the scientific methodology summarized in the introduction. These stages are all essential for determining the quality and relevance of the solution so that the modeling is suitable for the purpose, such as is captured by the notions of usability, feasibility and reliability[12]. Future studies can focus on systematically organizing knowledge about the lifecycle processes in ESSM, encompassing the physical mechanisms behind Earth surface processes, data sources, model structures, and computational methodologies. By integrating this prior knowledge with deep learning techniques, an intelligent question-answering and recommendation system can be developed to assist users in generating accurate and customized solutions for their specific tasks. In this envisioned process, modeling-related knowledge would guide and constrain the inference of deep learning models, while deep learning techniques could potentially uncover new discoveries and insights by leveraging the existing knowledge base.

Subjectivity in the modeling lifecycle. Subjective factors can be the primary obstacles to achieving highly accurate outcomes in hybrid ESSM. As noted previously, researchers are prone to use their expertise or criteria, likely making the modeling logic less precise and potentially biased. For example, modelers can favor physical or numerical models, whereas others with a strong background in DL prefer a more data-driven approach. Both may lead to suboptimal hybrid models for a specific task[140]. Another underlying challenge is that numerous innovative ideas and techniques about DL continue to inundate scientific communities, necessitating researchers to comprehend the most current technical advancements[141]. When it comes to choosing configurations (e.g., architectures or hyperparameters) for deep neural networks, many experienced ESSM researchers might be at a loss.

Incompatible computational environment. Incompatibilities between ESSM and DL in terms of hardware, software stack, and operating environment could impair computational efficiency. Specifically, process-based models are often executed on multi-CPU computers or high-performance computing facilities[142], whereas the training and inferencing phases of deep neural networks are typically deployed in GPU-based and container-based (e.g., Docker) environments[143]. Further, process-based models, particularly mechanistic ones, were until recently often constructed using Fortran and C++, whereas deep neural networks in specific environments employ Python and packages like Tensorflow and PyTorch. This latter distinction has become less problematic as many scientists are starting to embrace Python and the emerging technique of scientific machine learning (SciML) developed by Julia[144]. But these discrepancies in development and deployment methodologies generally result in separating DL and ESSM workloads. As a result, it significantly impacts data and message transmission and limits the computing capacity of hybrid ESSM.

Towards intelligent ESSM

Constructing appropriate and effective solutions to complex ESSM tasks is generally challenging. An initial undertaking is to fully understand the problem contexts and associated geographic objects. Handling big and multimodal data, especially extracting useful information or knowledge from it, is also a laborious task. Further, it is essential to focus on the trade-offs between model complexity and computational efficiency, as well as to calibrate the derived models and quantify or at least indicate model performance including uncertainty aspects. Finally, when applying constructed models, computational environments and software stacks are not easy to comprehend for those domain experts who are often not also experts in computation. Given the possible challenges and shortcomings analyzed earlier, we aim to start an intelligent revolution in ESSM that automatically delivers customized, scalable and accurate solutions to given ESSM tasks so as to lower technical barriers. We propose a conceptual framework intending to direct the entire modeling lifecycle automatically and intelligently for specific tasks (Box 1).

Box 1 | **The construction of the conceptual framework.**

In a homogeneous environment, adaptive guidance plans are intelligently generated to guide modeling tasks and allocate modular resources throughout the modeling lifecycle. These plans are the outputs of question-answering mechanisms and recommendation functions, powered by a modeling-related knowledge repository and a DL computing system. The repository organizes knowledge like domain theory, computational resources and model configurations, extracted from peer-reviewed literature, web corpora, and expert input. The DL computing system employs advanced strategies to predict user preferences (user-item interactions) in modeling, enabling acquisition of the most desired resources. Moreover, the knowledge repository provides a priori knowledge to constrain the prediction results of the DL computing system, while the latter can provide inferencing aid to uncover and infer unknown knowledge from given material to improve the former's completeness. The adaptive guidance plans function as follows:

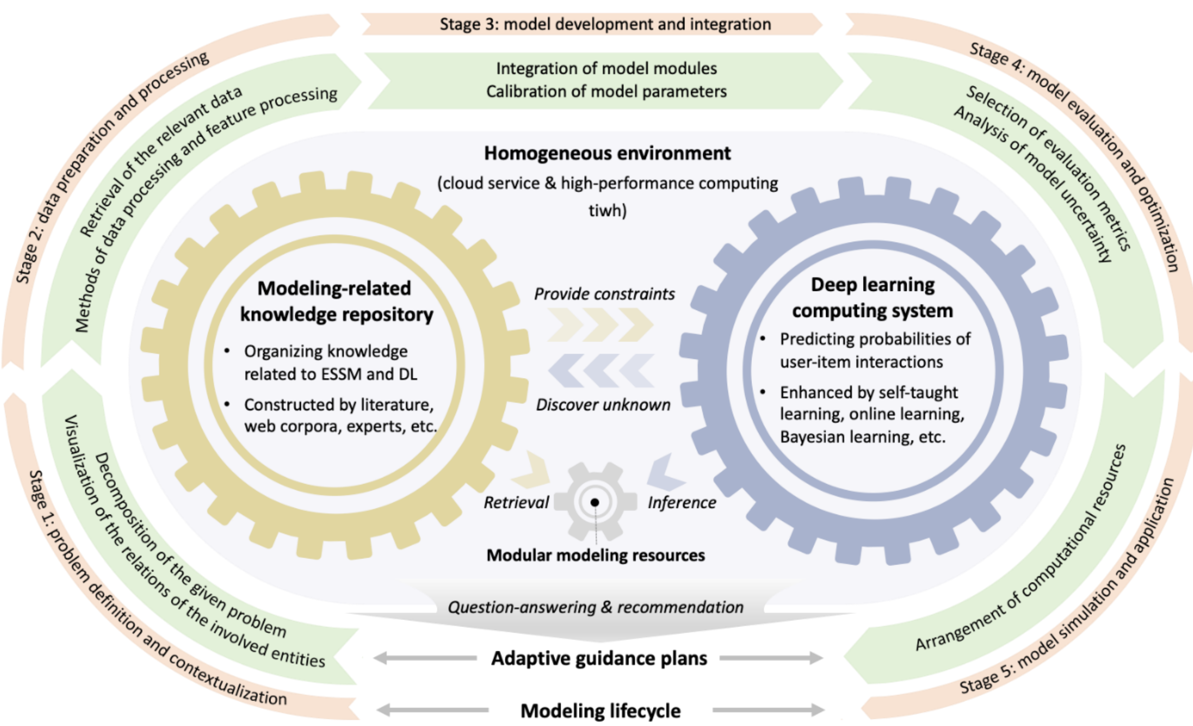
Stage 1. Problem definition and contextualization. A given ESSM topic to investigate can be hierarchically decomposed into multiple sub-analyses or interactions of the subsystems involved. Geographic objects and their relations in space and time can be recognized and visualized automatically, giving modelers a clearer understanding of what to analyze before investigation commences.

Stage 2. Data preparation and processing. Relevant data is retrieved, with adaptive recommendations for processing techniques. Therefore, beneficial patterns and interior knowledge can be acquired from various types of data rather than through manual manipulation.

Stage 3. Model development and integration. A modular strategy organizes model modules, allowing knowledge-based reasoning methods to build customized models for specific needs. Automatic methods, including calibration and uncertainty estimation, improve prediction results and computational efficiency.

Stage 4. Model evaluation and optimization. Suitable metrics for evaluating performance will also be selected at this stage. The probabilistic inferencing and other methods will be deployed to estimate the statistical confidence in the various models and other ways to represent uncertainties.

Stage 5. Model simulation and application. Based on the characteristics of the data and models, it is proposed that the CPU, GPU, memory, storage and network resources be dynamically scheduled to enhance computational performance and efficiency.



Modeling-related knowledge repository. The purpose of the modeling-related knowledge repository is to organize the diverse knowledge required to model Earth surface processes, not only about the paradigm of conventional methods, but also about the DL techniques. For example, there should be knowledge about geophysical mechanisms in the subsystems, as well as modeling information such as data, methods, models and computational logic. This repository could have the advantage of not only organizing a vast amount of knowledge but also serving as an a priori medium and constraint to improve performance of the DL computing, thereby enabling the discovery of previously undiscovered information.

In constructing this knowledge repository strategies combine both "bottom-up" and "top-down" designations. The former uses natural language processing and computer vision methods to automatically extract dynamic knowledge about concepts, entities and relationships of the Earth surface system from publicly available authoritative data (e.g., peer-reviewed research literature and web corpora). The latter depends on the domain expertise of specialists. That is, it relies on the crowdsourced participation of experts or communities to build the repository's structure and properties, encode existing knowledge, and to provide this in collaboration. Moreover, the "bottom-up" strategy can generate more comprehensive and up-to-date knowledge, and the construction process can be automated, but it relies on existing information and knowledge extraction technologies. On the other hand, the "top-down" strategy relies on expert knowledge, which could be more scientific but biased and would typically involve tedious manual manipulation. These two constructions have distinct characteristics and can be combined to create a comprehensive knowledge repository.

Deep learning computing system. The DL computing system and the modeling-related knowledge repository ideally are built in a homogeneous environment. There are two potential practical advantages from a computational standpoint. First, a homogeneous environment is likely to be efficient for interaction and communication between these two components since it avoids the issue of incompatible local computing facilities. Second, such a homogeneous environment can continuously acquire data from the Internet and crowdsourcing, update deep neural networks online and dynamically complete the modeling-related knowledge repository.

In order to anticipate the probability value of modeling resources being used through deep neural networks, it must first comprehend the demands of modelers based on context analysis and historical user-item interaction records. The DL computing system can also be used during the "bottom-up" development, and to infer unknown or missing information among the existing knowledge based for instance on Bayesian learning strategies, all of which can be used to complete the knowledge repository. Data deficiency, that is, the lack of high-quality ground truth data, will be a foreseeable challenge for the DL computing system. In this case, (semi- or un-) supervised learning or self-taught learning offer promising solutions. The deep neural networks would update their parameters based on unlabeled data and maintain lifelong learning on the backend, facilitating self-renewal of the conceptual framework.

Adaptive guidance plans. Crucially, the framework entails question-answering mechanisms and recommendation functions to generate adaptive modeling guidance plans. The question-answering mechanisms should grasp modeling descriptions and requirements and then acquire answers, whereas the recommendation function adaptively relates to the next modeling steps. The adaptive guidance plans are expected to serve as clear guidelines for issues that arise throughout the entire modeling lifecycle.

Potential application case. There are many uncertainties in modeling human activities in the Earth surface system because human behavior is heterogeneous and variously influenced not only by themselves but also by their social and natural environment[145]. And unlike purely biophysical systems human behavior is not governed by scientific laws. Moreover, most of the data for such modeling are derived from questionnaires, survey observations, interviews, expert opinion and global positioning systems, thereby incorporating a large degree of subjective and systematic bias[146]. Here, human behavior, especially in the context of the post-COVID-19 pandemic era, will be used to demonstrate the proposed framework.

The rapid transmission of COVID-19 and its unprecedented worldwide scope have radically altered contemporary societies. We have entered the post-COVID-19 era during which all countries live with COVID. Consequently, we are interested in questions like what are the various impacts of the changed travel behaviors adapted to the post-COVID-19 era? Understanding this question will assist in analyzing how the pandemic affected the social and natural environment to promote sustainable development. We employ our conceptual framework to guide the modeling of this problem, with potential guidance plans as illustrated in Fig. 5.

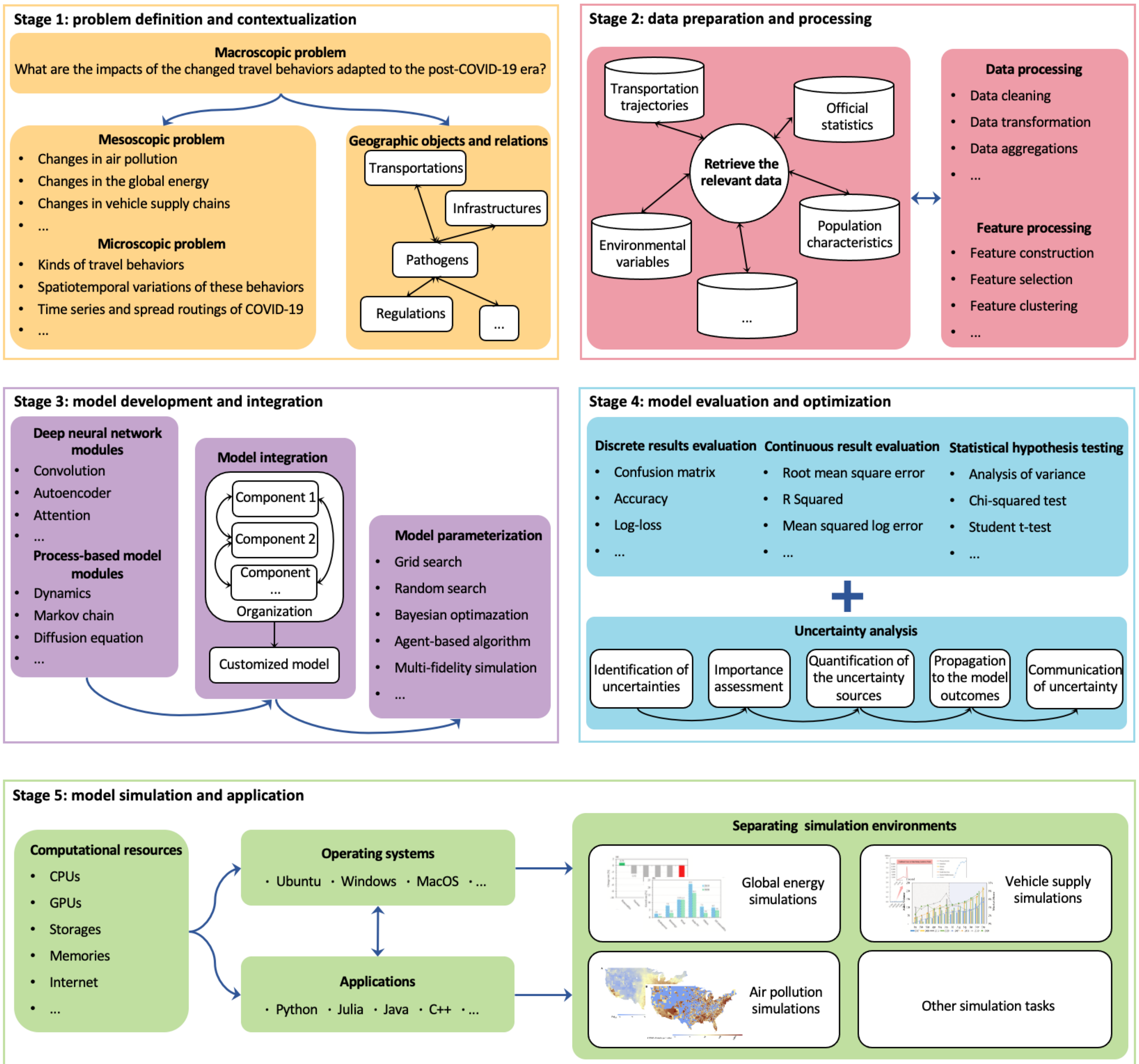


Figure 5. **Potential plans for guiding the modeling for analyzing impacts of travel behaviors adapted to the post-COVID-19 era.** The guidance plans are generated by our conceptual framework and can direct modelers to build customized solutions to solve given problems. The guidance can be divided into five stages according to the ESSM lifecycle.

First, this macroscopic problem can be automatically subdivided into mesoscopic issues (such as changes in air pollution, global energy supply and consumption, or vehicle supply chains) and additional microscopic issues (e.g., different varieties of travel behaviors, spatiotemporal variations in these behaviors, time series trajectories and spread routings of COVID-19, and interrelationships between COVID-19 variants and travel behaviors). By visualizing the geographic objects and relationships involved, the decomposed sub-problems help modelers comprehensively understand the raw problem.

Second, multimodal data collection and processing will be a significant technical challenge. This conceptual framework should capture and extract real-time information from open platforms such as administrative websites, social media platforms and news websites, to recommend various sorts of up-to-date data and their associated processing methods. By assembling modular deep neural networks and process-based models, customized model architectures for distinct challenges can be created. Through automated parameterization of models, the potential of these models for addressing associated issues can be enhanced. The guidance plans should also include metrics and methodologies for outcome evaluation and uncertainty analysis, thereby enhancing trustworthiness. Finally, the guidance plans should stress application effectiveness and computational efficiency. Computing resources can be automatically aggregated in cyberspace and loaded with suitable operating systems and software applications to increase the compatibility of computational pipelines and efficacy of the analysis.

Summary and future perspectives

Integration of ESSM and DL approaches is an emerging paradigm for understanding Earth surface system dynamics. Most research focuses on integrating process-based models and deep neural networks into hybrid models, rather than exploring the advantages of a comprehensive approach that covers all modeling lifecycle stages. Moreover, integration success could be affected by subjective biases in the modeling processes and incompatible computational environments between the approaches. In this paper, we have examined the state and characteristics of studies of ESSM, DL, and their current hybrids before presenting a conceptual framework that we envision to be an intelligent revolution in ESSM. This aims to intelligently create customized, scalable, and accurate solutions for modeling Earth surface processes by integrating the ESSM knowledge and DL capabilities.

Our framework shares similarities with ChatGPT[147–149] in the ability to automatically generate customized responses based on user inputs by leveraging deep learning techniques, but it is specifically designed for the ESSM field. Notable differences between our framework and ChatGPT include the output form (multimodal outputs and modeling resource assignment vs. pure-text outputs), technical foundation (knowledge-constrained inference vs. inference by large-scale deep neural networks) and learning strategy (online self-learning vs. periodic background updates)[147–149].

In conclusion, the integration of ESSM and DL is across multiple disciplinaries and still an evolving field of science, thus aspiring to advance capability and capacity through the collaboration of an open scientific community and to increase the trustworthiness of the results through advanced tools and good practices. Here, we present recommendations for the future development of the ESSM and DL integration to make it more accessible, transparent and trustworthy.

Open community. ESSM's interdisciplinarity necessitates an open community for open knowledge, open resources (for example, datasets, codes and models) and open research cooperation. Research organizations, such as the OMF (Open Modeling Foundation)[150], the OpenGMS (Open Geographic Modeling and Simulation)[151] and CSDMS (Community Surface Dynamics Modeling System)[152] already encourage collaboration and sharing. Hopefully among others, these environments will facilitate the collaboration of scientists from various disciplines to address complex problems. Building a virtual online platform for researchers to experiment and discuss will also enhance the transparency and reproducibility of modeling.

Trustworthiness of outcomes. The black box nature of DL networks presents a unique challenge for geoscientific applications, as they are not easily interpretable despite producing superior results. Explainable or interpretable artificial intelligence using explanatory approaches (for example, layer-wise relevance propagation, integrated gradients, and occlusion analysis) do however allow users to understand internal mechanics of deep neural networks[153]. Merging process-based models with domain-specific knowledge as surrogates in deep neural networks can further increase the transparency of what might otherwise be black boxes[154]. Related research projects are still evolving, but there remains a significant trade-off between model performance in terms of explainability and simulation accuracy of model outputs.

Moving forward, our framework anticipates the intelligent development of customized models; however, the pathway may not align entirely with geoscientists' logic, as generated solutions predominantly depend on the inference results of deep neural networks. Therefore, it advocates not only enhancing the accuracy of the DL computing system based on specific objective functions, but also implementing contextually-appropriate logic constraints that are compliant with the mindset of major geoscientists. These considerations should be taken into account throughout the entire modeling lifecycle, ultimately enhancing the trustworthiness of results and outcomes.

Moreover, two typical characteristics reduce confidence in the predictive accuracy of ESSM. The first is the difficulty in accurately simulating certain extreme events due to the highly dynamic character of the Earth's surface system[155]. Second, climate change and technological progress, such as the capacity of humans to move sediments, could disrupt observed data, posing additional challenges to the efficacy of created models[156,157]. To mitigate these issues, maintaining regular updates of models and software is crucial, as is utilizing data assimilation, lifelong learning techniques, and explainable or interpretable artificial intelligence. In addition, acquiring the up-to-date and widespread data and processing the vulnerable observations using hybrid models can also effectively improve modeling performance.

Ultimately, recognizing that uncertainty will always be present, enhancing the trustworthiness and credibility of results requires adherence to good modeling practices[12,158]. These include deliberating on fitness for purpose, applying systematic procedures, characterizing and discussing uncertainties, justifying choices, and clearly stating assumptions and limitations[10]. Ensuring transparency through thorough documentation further strengthens the reliability of the outcomes[12,159].

References

1. Phillips, J. D. *Earth surface systems*. (Blackwell Oxford, 1999).
2. Knight, J. & Harrison, S. The impacts of climate change on terrestrial Earth surface systems. *Nature Climate Change* **3**, 24–29 (2013).
3. Murray, A. B. *et al.* Geomorphology, complexity, and the emerging science of the Earth’s surface. *Geomorphology* **103**, 496–505 (2009).
4. Phillips, J. D. Amplifiers, filters and geomorphic responses to climate change in Kentucky rivers. *Climatic change* **103**, 571–595 (2010).
5. Suzuki, D. *The sacred balance: Rediscovering our place in nature*. (Greystone Books Ltd, 2022).
6. Allen, P. A. *Earth surface processes*. (John Wiley & Sons, 2009).
7. Chen, M. *et al.* Geographic modeling and simulation systems for geographic research in the new era: Some thoughts on their development and construction. *Sci. China Earth Sci.* **64**, 1207–1223 (2021).
8. Luttge, A., Arvidson, R. S., Fischer, C. & Kurganskaya, I. Kinetic concepts for quantitative prediction of fluid-solid interactions. *Chemical Geology* **504**, 216–235 (2019).
9. Pelletier, J. D. *Quantitative modeling of earth surface processes*. (Cambridge University Press, 2008).
10. Jakeman, A. J., Letcher, R. A. & Norton, J. P. Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling & Software* **21**, 602–614 (2006).
11. Ma, Z. *et al.* Activity-based process construction for participatory geo-analysis. *GIScience & Remote Sensing* **58**, 180–198 (2021).
12. Hamilton, S. H., Pollino, C. A., Stratford, D. S., Fu, B. & Jakeman, A. J. Fit-for-purpose environmental modeling: Targeting the intersection of usability, reliability and feasibility. *Environmental Modelling & Software* **148**, 105278 (2022).
13. Chen, M. *et al.* Position paper: Open web-distributed integrated geographic modelling and simulation to enable broader participation and applications. *Earth-Science Reviews* **207**, 103223 (2020).
14. Lee, C. A., Gasster, S. D., Plaza, A., Chang, C.-I. & Huang, B. Recent developments in high performance computing for remote sensing: A review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **4**, 508–527 (2011).

15. Li, S. *et al.* Geospatial big data handling theory and methods: A review and research challenges. *ISPRS Journal of Photogrammetry and Remote Sensing* **115**, 119–133 (2016).

16. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).

17. Goldstein, E. B., Coco, G. & Plant, N. G. A review of machine learning applications to coastal sediment transport and morphodynamics. *Earth-Science Reviews* **194**, 97–108 (2019).

18. Qian, Z. *et al.* Deep Roof Refiner: A detail-oriented deep learning network for refined delineation of roof structure lines using satellite imagery. *International Journal of Applied Earth Observation and Geoinformation* **107**, 102680 (2022).

19. Li, W. & Hsu, C.-Y. GeoAI for Large-Scale Image Analysis and Machine Vision: Recent Progress of Artificial Intelligence in Geography. *ISPRS International Journal of Geo-Information* **11**, 385 (2022).

20. Reichstein, M. *et al.* Deep learning and process understanding for data-driven Earth system science. *Nature* **566**, 195–204 (2019).

21. Camps-Valls, G., Tuia, D., Zhu, X. X. & Reichstein, M. *Deep learning for the Earth Sciences: A comprehensive approach to remote sensing, climate science and geosciences*. (John Wiley & Sons, 2021).

22. Bergen, K. J., Johnson, P. A., de Hoop, M. V. & Beroza, G. C. Machine learning for data-driven discovery in solid Earth geoscience. *Science* **363**, eaau0323 (2019).

23. Sutton, R. The bitter lesson. *Incomplete Ideas (blog)* **13**, (2019).

24. Razavi, S. *et al.* Coevolution of machine learning and process-based modelling to revolutionize Earth and environmental sciences: A perspective. *Hydrological Processes* **36**, e14596 (2022).

25. Bolton, T. & Zanna, L. Applications of Deep Learning to Ocean Data Inference and Subgrid Parameterization. *J. Adv. Model. Earth Syst.* **11**, 376–399 (2019).

26. Beucler, T., Rasp, S., Pritchard, M. & Gentine, P. Achieving Conservation of Energy in Neural Network Emulators for Climate Modeling. Preprint at http://arxiv.org/abs/1906.06622 (2019).

27. Kadow, C., Hall, D. M. & Ulbrich, U. Artificial intelligence reconstructs missing climate information. *Nature Geoscience* **13**, 408–413 (2020).

28. Razavi, S. Deep learning, explained: Fundamentals, explainability, and bridgeability to process-based modelling. *Environmental Modelling & Software* **144**, 105159 (2021).

29. Fisher, R. A. *et al.* Vegetation demographics in Earth System Models: A review of progress and priorities. *Global change biology* **24**, 35–54 (2018).

30. Burnash, R. & others. The NWS River Forecast System-catchment modeling. *Computer models of watershed hydrology.* 311–366 (1995).

31. Tikhamarine, Y. *et al.* Rainfall-runoff modelling using improved machine learning methods: Harris hawks optimizer vs. particle swarm optimization. *Journal of Hydrology* **589**, 125133 (2020).

32. Harbaugh, A. W. *MODFLOW-2005, the US Geological Survey modular ground-water model: the ground-water flow process*. vol. 6 (US Department of the Interior, US Geological Survey Reston, VA, USA, 2005).

33. Ali, A. S. A., Ebrahimi, S., Ashiq, M. M., Alasta, M. S. & Azari, B. CNN-Bi LSTM neural network for simulating groundwater level. *Environ. Eng* **8**, 1–7 (2022).

34. Laflen, J. M., Lane, L. J. & Foster, G. R. WEPP: A new generation of erosion prediction technology. *Journal of soil and water conservation* **46**, 34–38 (1991).

35. Senanayake, S. & Pradhan, B. Predicting soil erosion susceptibility associated with climate change scenarios in the Central Highlands of Sri Lanka. *Journal of Environmental Management* **308**, 114589 (2022).

36. Ferro, V. & Porto, P. Sediment delivery distributed (SEDD) model. *Journal of hydrologic engineering* **5**, 411–422 (2000).

37. Buscombe, D. SediNet: a configurable deep learning model for mixed qualitative and quantitative optical granulometry. *Earth Surf. Process. Landforms* **45**, 638–651 (2020).

38. Turner, D. B. *Workbook of atmospheric dispersion estimates: an introduction to dispersion modeling*. (CRC press, 2020).

39. Zhang, Q., Fu, F. & Tian, R. A deep learning and image-based model for air quality estimation. *Science of the Total Environment* **724**, 138178 (2020).

40. Skamarock, W. C. *et al. A description of the advanced research WRF version 2*. (2005).

41. Keisler, R. Forecasting global weather with graph neural networks. *arXiv preprint arXiv:2202.07575* (2022).

42. Sato, H., Itoh, A. & Kohyama, T. SEIB–DGVM: A new Dynamic Global Vegetation Model using a spatially explicit individual-based approach. *Ecological Modelling* **200**, 279–307 (2007).

43. Jung, M. *et al.* The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific data* **6**, 74 (2019).

44. Mondal, B. *et al.* Urban expansion and wetland shrinkage estimation using a GIS-based model in the East Kolkata Wetland, India. *Ecological indicators* **83**, 62–73 (2017).

45. Wang, X. *et al.* Estimation of soil salt content (SSC) in the Ebinur Lake Wetland National Nature Reserve (ELWNNR), Northwest China, based on a Bootstrap-BP neural network model and optimal spectral indices. *Science of the Total Environment* **615**, 918–930 (2018).

46. Bailey, N. T. & others. *The mathematical theory of infectious diseases and its applications*. (Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE., 1975).

47. Yang, Z. *et al.* Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *Journal of thoracic disease* **12**, 165 (2020).

48. Karemera, D., Oguledo, V. I. & Davis, B. A gravity model analysis of international migration to North America. *Applied economics* **32**, 1745–1755 (2000).

49. Simini, F., Barlacchi, G., Luca, M. & Pappalardo, L. A deep gravity model for mobility flows generation. *Nature communications* **12**, 6576 (2021).

50. Werner, B. T. Complexity in Natural Landform Patterns. *Science* **284**, 102–104 (1999).

51. Anderson, P. W. More is different: broken symmetry and the nature of the hierarchical structure of science. *Science* **177**, 393–396 (1972).

52. Werner, B. T. & McNamara, D. E. Dynamics of coupled human-landscape systems. *Geomorphology* **91**, 393–407 (2007).

53. Klir, G. J. & Simon, H. A. *The architecture of complexity*. (Springer, 1991).

54. Heymann, M. & Dahan Dalmedico, A. Epistemology and politics in Earth system modeling: Historical perspectives. *Journal of Advances in Modeling Earth Systems* **11**, 1139–1152 (2019).

55. Shen, C. & Phanikumar, M. S. A process-based, distributed hydrologic model based on a large-scale method for surface–subsurface coupling. *Advances in Water Resources* **33**, 1524–1541 (2010).

56. Lü, G. *et al.* Geographic scenario: a possible foundation for further development of virtual geographic environments. *International Journal of Digital Earth* **11**, 356–368 (2018).

57. Haarsma, R. J. *et al.* High resolution model intercomparison project (HighResMIP v1. 0) for CMIP6. *Geoscientific Model Development* **9**, 4185–4208 (2016).

58. Worley, P. H. *et al.* Performance of the community earth system model. in *Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis* 1–11 (2011).

59. Zhang, Z. *et al.* Vectorized rooftop area data for 90 cities in China. *Sci Data* **9**, 66 (2022).

60. Lee, J.-G. & Kang, M. Geospatial Big Data: Challenges and Opportunities. *Big Data Research* **2**, 74–81 (2015).

61. Ansari, S. *et al.* Unlocking the Potential of NEXRAD Data through NOAA's Big Data Partnership. *Bulletin of the American Meteorological Society* **99**, 189–204 (2018).

62. Ge, Y. *et al.* Progress of big geodata. *Science Bulletin* **67**, 1739–1742 (2022).

63. Yin, S. & Kaynak, O. Big data for modern industry: challenges and trends [point of view]. *Proceedings of the IEEE* **103**, 143–146 (2015).

64. Qian, Z. *et al.* Vectorized dataset of roadside noise barriers in China using street view imagery. *Earth Syst. Sci. Data* **14**, 4057–4076 (2022).

65. Clark, M. P., Kavetski, D. & Fenicia, F. Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resources Research* **47**, (2011).

66. Beven, K. & Freer, J. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of hydrology* **249**, 11–29 (2001).

67. Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G. & Yacalis, G. Could machine learning break the convection parameterization deadlock? *Geophysical Research Letters* **45**, 5742–5751 (2018).

68. Tang, Y., Reed, P., Wagener, T. & Van Werkhoven, K. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. *Hydrology and Earth System Sciences* **11**, 793–817 (2007).

69. Di Baldassarre, G., Schumann, G. & Bates, P. Near real time satellite imagery to support and verify timely flood modelling. *Hydrological Processes: An International Journal* **23**, 799–803 (2009).

70. Kucera, P. A. *et al.* Precipitation from Space: Advancing Earth System Science. *Bull. Amer. Meteor. Soc.* **94**, 365–375 (2013).

71. Zhang, K. *et al.* Quantifying the photovoltaic potential of highways in China. *Applied Energy* **324**, 119600 (2022).

72. Fuhrer, O. *et al.* Near-global climate simulation at 1 km resolution: establishing a performance baseline on 4888 GPUs with COSMO 5.0. *Geoscientific Model Development* **11**, 1665–1681 (2018).

73. Kratzert, F., Klotz, D., Brenner, C., Schulz, K. & Herrnegger, M. Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrol. Earth Syst. Sci.* **22**, 6005–6022 (2018).

74. Wei, X., Zhang, L., Yang, H.-Q., Zhang, L. & Yao, Y.-P. Machine learning for pore-water pressure time-series prediction: Application of recurrent neural networks. *Geoscience Frontiers* **12**, 453–467 (2021).

75. Khodayar, M. & Wang, J. Spatio-temporal graph deep neural network for short-term wind speed forecasting. *IEEE Transactions on Sustainable Energy* **10**, 670–681 (2018).

76. Bihlo, A. A generative adversarial network approach to (ensemble) weather prediction. *Neural Networks* **139**, 1–16 (2021).

77. Zhou, L. & Zhang, R.-H. A self-attention–based neural network for three-dimensional multivariate modeling and its skillful ENSO predictions. *Science Advances* **9**, eadf2827 (2023).

78. Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y. & Beroza, G. C. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature communications* **11**, 1–12 (2020).

79. Gilpin, L. H. *et al.* Explaining Explanations: An Overview of Interpretability of Machine Learning. Preprint at http://arxiv.org/abs/1806.00069 (2019).

80. Montavon, G., Samek, W. & Müller, K.-R. Methods for interpreting and understanding deep neural networks. *Digital Signal Processing* **73**, 1–15 (2018).

81. Gómez-Chova, L., Tuia, D., Moser, G. & Camps-Valls, G. Multimodal classification of remote sensing images: A review and future directions. *Proceedings of the IEEE* **103**, 1560–1584 (2015).

82. Zhu, R. *et al.* Deep solar PV refiner: A detail-oriented deep learning network for refined segmentation of photovoltaic areas from satellite imagery. *International Journal of Applied Earth Observation and Geoinformation* **116**, 103134 (2023).

83. Hong, D. *et al.* More diverse means better: Multimodal deep learning meets remote-sensing imagery classification. *IEEE Transactions on Geoscience and Remote Sensing* **59**, 4340–4354 (2020).

84. Fan, R. *et al.* Fine-scale urban informal settlements mapping by fusing remote sensing images and building data via a transformer-based multimodal fusion network. *IEEE Transactions on Geoscience and Remote Sensing* **60**, 1–16 (2022).

85. Ives, A. R. *et al.* Statistical inference for trends in spatiotemporal data. *Remote Sensing of Environment* **266**, 112678 (2021).

86. Li, X., Zhang, C. & Li, W. Building block level urban land-use information retrieval based on Google Street View images. *GIScience & Remote Sensing* **54**, 819–835 (2017).

87. Zhang, K. *et al.* Using street view images to identify road noise barriers with ensemble classification model and geospatial analysis. *Sustainable Cities and Society* **78**, 103598 (2022).

88. Zhong, T. *et al.* Assessment of solar photovoltaic potentials on urban noise barriers using street-view imagery. *Renewable Energy* **168**, 181–194 (2021).

89. Bousmalis, K., Silberman, N., Dohan, D., Erhan, D. & Krishnan, D. Unsupervised pixel-level domain adaptation with generative adversarial networks. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 3722–3731 (2017).

90. Hess, P., Drüke, M., Petri, S., Strnad, F. M. & Boers, N. Physically constrained generative adversarial networks for improving precipitation fields from Earth system models. *Nature Machine Intelligence* **4**, 828–839 (2022).

91. He, X., Chen, Y. & Ghamisi, P. Heterogeneous transfer learning for hyperspectral image classification based on convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing* **58**, 3246–3263 (2019).

92. Raissi, M., Perdikaris, P. & Karniadakis, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics* **378**, 686–707 (2019).

93. Shi, X. *et al.* Deep learning for precipitation nowcasting: A benchmark and a new model. *Advances in neural information processing systems* **30**, (2017).

94. Stevens, B. & Bony, S. What Are Climate Models Missing? *Science* **340**, 1053–1054 (2013).

95. Gao, Z. *et al.* Deep learning and the weather forecasting problem: Precipitation nowcasting. *Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences* 218–239 (2021).

96. Lobry, S., Marcos, D., Murray, J. & Tuia, D. RSVQA: Visual question answering for remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing* **58**, 8555–8566 (2020).

97. Chai, S., Xu, Z., Jia, Y. & Wong, W. K. A robust spatiotemporal forecasting framework for photovoltaic generation. *IEEE Transactions on Smart Grid* **11**, 5370–5382 (2020).

98. Boers, N. *et al.* Complex networks reveal global pattern of extreme-rainfall teleconnections. *Nature* **566**, 373–377 (2019).

99. Sambasivan, N. *et al.* "Everyone wants to do the model work, not the data work": Data Cascades in High-Stakes AI. in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* 1–15 (ACM, 2021). doi:10.1145/3411764.3445518.

100. Goldstein, E. B. *et al.* Labeling Poststorm Coastal Imagery for Machine Learning: Measurement of Interrater Agreement. *Earth and Space Science* **8**, (2021).

101. Geiger, R. S. *et al.* 'Garbage In, Garbage Out' Revisited: What Do Machine Learning Application Papers Report About Human-Labeled Training Data? *Quantitative Science Studies* **2**, 795–827 (2021).

102. Samsi, S., Mattioli, C. J. & Veillette, M. S. Distributed deep learning for precipitation nowcasting. in *2019 IEEE High Performance Extreme Computing Conference (HPEC)* 1–7 (IEEE, 2019).

103. Rasp, S., Pritchard, M. S. & Gentine, P. Deep learning to represent subgrid processes in climate models. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 9684–9689 (2018).

104. Flato, G. M. Earth system models: an overview. *WIREs Clim Change* **2**, 783–800 (2011).

105. Brandt, M. *et al.* An unexpectedly large count of trees in the West African Sahara and Sahel. *Nature* **587**, 78–82 (2020).

106. Tung, F. & Mori, G. Deep neural network compression by in-parallel pruning-quantization. *IEEE transactions on pattern analysis and machine intelligence* **42**, 568–579 (2018).

107. Jouppi, N. P., Young, C., Patil, N. & Patterson, D. A domain-specific architecture for deep neural networks. *Communications of the ACM* **61**, 50–59 (2018).

108. Shen, H. & Zhang, L. Mechanism-learning coupling paradigms for parameter inversion and simulation in earth surface systems. *Sci. China Earth Sci.* (2023) doi:10.1007/s11430-022-9999-9.

109. Hunter, J. M. *et al.* Framework for developing hybrid process-driven, artificial neural network and regression models for salinity prediction in river systems. *Hydrol. Earth Syst. Sci.* **22**, 2987–3006 (2018).

110. Lv, X. *et al.* BTS: a binary tree sampling strategy for object identification based on deep learning. *International journal of geographical information science* **36**, 822–848 (2022).

111. Sun, Z. *et al.* Improving the Performance of Automated Rooftop Extraction through Geospatial Stratified and Optimized Sampling. *Remote Sensing* **14**, 4961 (2022).

112. Vandal, T. *et al.* Deepsd: Generating high resolution climate change projections through single image super-resolution. in *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining* 1663–1672 (2017).

113. Kurth, T. *et al.* Exascale deep learning for climate analytics. in *SC18: International Conference for High Performance Computing, Networking, Storage and Analysis* 649–660 (IEEE, 2018).

114. Lanaras, C., Bioucas-Dias, J., Galliani, S., Baltsavias, E. & Schindler, K. Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network. *ISPRS Journal of Photogrammetry and Remote Sensing* **146**, 305–319 (2018).

115. Zhang, Y., Yu, W. & Zhu, D. Terrain feature-aware deep learning network for digital elevation model superresolution. *ISPRS Journal of Photogrammetry and Remote Sensing* **189**, 143–162 (2022).

116. Eslami, E., Choi, Y., Lops, Y., Sayeed, A. & Salman, A. K. Using wavelet transform and dynamic time warping to identify the limitations of the CNN model as an air quality forecasting system. *Geoscientific Model Development* **13**, 6237–6251 (2020).

117. Duan, P. *et al.* Fusion of dual spatial information for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing* **59**, 7726–7738 (2020).

118. Han, J., Jentzen, A. & E, W. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences* **115**, 8505–8510 (2018).

119. Gagne, D. J., Christensen, H. M., Subramanian, A. C. & Monahan, A. H. Machine learning for stochastic parameterization: Generative adversarial networks in the Lorenz'96 model. *Journal of Advances in Modeling Earth Systems* **12**, e2019MS001896 (2020).

120. Zhu, X. *et al.* Comparison of two optimized machine learning models for predicting displacement of rainfall-induced landslide: A case study in Sichuan Province, China. *Engineering geology* **218**, 213–222 (2017).

121. Rasp, S. Coupled online learning as a way to tackle instabilities and biases in neural network parameterizations: general algorithms and Lorenz 96 case study (v1. 0). *Geoscientific Model Development* **13**, 2185–2196 (2020).

122. Qian, Z., Liu, X., Tao, F. & Zhou, T. Identification of Urban Functional Areas by Coupling Satellite Images and Taxi GPS Trajectories. *Remote Sensing* **12**, 2449 (2020).

123. Amini, A., Dolatshahi, M. & Kerachian, R. Adaptive precipitation nowcasting using deep learning and ensemble modeling. *Journal of Hydrology* **612**, 128197 (2022).

124. Li, G. & Choi, Y. HPC cluster-based user-defined data integration platform for deep learning in geoscience applications. *Computers & Geosciences* **155**, 104868 (2021).

125. Sun, A. Y. & Scanlon, B. R. How can Big Data and machine learning benefit environment and water management: a survey of methods, applications, and future directions. *Environmental Research Letters* **14**, 073001 (2019).

126. See, S. & Adie, J. Challenges and opportunities for a hybrid modelling approach to earth system science. *CCF Trans. HPC* **3**, 320–329 (2021).

127. Venkatakrishnan, S. V., Bouman, C. A. & Wohlberg, B. Plug-and-play priors for model based reconstruction. in *2013 IEEE Global Conference on Signal and Information Processing* 945–948 (IEEE, 2013).

128. Shen, H. *et al.* Coupling Model- and Data-Driven Methods for Remote Sensing Image Restoration and Fusion: Improving physical interpretability. *IEEE Geosci. Remote Sens. Mag.* **10**, 231–249 (2022).

129. Goldstein, E. B. & Coco, G. Machine learning components in deterministic models: hybrid synergy in the age of data. *Front. Environ. Sci.* **3**, (2015).

130. Gelbrecht, M., Boers, N. & Kurths, J. Neural partial differential equations for chaotic systems. *New J. Phys.* **23**, 043005 (2021).

131. Mahjoubi, S., Barhemat, R., Guo, P., Meng, W. & Bao, Y. Prediction and multi-objective optimization of mechanical, economical, and environmental properties for strain-hardening cementitious

composites (SHCC) based on automated machine learning and metaheuristic algorithms. *Journal of Cleaner Production* **329**, 129665 (2021).

132. Goldstein, E. B., Coco, G., Murray, A. B. & Green, M. O. Data-driven components in a model of inner-shelf sorted bedforms: a new hybrid model. *Earth Surf. Dynam.* **2**, 67–82 (2014).

133. Gelbrecht, M., White, A., Bathiany, S. & Boers, N. Differentiable Programming for Earth System Modeling. Preprint at http://arxiv.org/abs/2208.13825 (2022).

134. Bar-Sinai, Y., Hoyer, S., Hickey, J. & Brenner, M. P. Learning data-driven discretizations for partial differential equations. *Proceedings of the National Academy of Sciences* **116**, 15344–15349 (2019).

135. Kraft, B., Jung, M., Körner, M., Koirala, S. & Reichstein, M. Towards hybrid modeling of the global hydrological cycle. *Hydrol. Earth Syst. Sci.* **26**, 1579–1614 (2022).

136. Zubov, K. *et al.* NeuralPDE: Automating Physics-Informed Neural Networks (PINNs) with Error Approximations.

137. Chen, T. Q., Rubanova, Y., Bettencourt, J. & Duvenaud, D. K. Neural Ordinary Differential Equations.

138. Hagenauer, J. & Helbich, M. A geographically weighted artificial neural network. *International Journal of Geographical Information Science* **36**, 215–235 (2022).

139. Li, H. & Weng, Y. Physical equation discovery using physics-consistent neural network (pcnn) under incomplete observability. in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* 925–933 (2021).

140. Ma, Z. *et al.* Customizable process design for collaborative geographic analysis. *GIScience & Remote Sensing* **59**, 914–935 (2022).

141. Zhang, W. *et al.* Application of machine learning, deep learning and optimization algorithms in geoengineering and geoscience: comprehensive review and future challenge. *Gondwana Research* (2022).

142. Massonnet, F. *et al.* Replicability of the EC-Earth3 Earth system model under a change in computing environment. *Geoscientific Model Development* **13**, 1165–1178 (2020).

143. Warnat-Herresthal, S. *et al.* Swarm learning for decentralized and confidential clinical machine learning. *Nature* **594**, 265–270 (2021).

144. Rackauckas, C. & Nie, Q. Differentialequations. jl–a performant and feature-rich ecosystem for solving differential equations in julia. *Journal of open research software* **5**, (2017).

145. Luo, W. & MacEachren, A. M. Geo-social visual analytics. *Journal of spatial information science* 27–66 (2014).

146. Jusup, M. *et al.* Social physics. *Physics Reports* **948**, 1–148 (2022).

147. Stokel-Walker, C. AI bot ChatGPT writes smart essays-should academics worry? *Nature* (2022).

148. Thorp, H. H. ChatGPT is fun, but not an author. *Science* **379**, 313–313 (2023).

149. van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R. & Bockting, C. L. ChatGPT: five priorities for research. *Nature* **614**, 224–226 (2023).

150. Barton, C. M. *et al.* How to make models more useful. *Proceedings of the National Academy of Sciences* **119**, e2202112119 (2022).

151. Chen, M. *et al.* Teamwork-oriented integrated modeling method for geo-problem solving. *Environmental modelling & software* **119**, 111–123 (2019).

152. Tucker, G. E. *et al.* CSDMS: a community platform for numerical modeling of Earth surface processes. *Geoscientific Model Development* **15**, 1413–1439 (2022).

153. Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J. & Müller, K.-R. Explaining deep neural networks and beyond: A review of methods and applications. *Proceedings of the IEEE* **109**, 247–278 (2021).

154. Irrgang, C. *et al.* Towards neural Earth system modelling by integrating artificial intelligence in Earth system science. *Nat Mach Intell* **3**, 667–674 (2021).

155. Trapp, R. J. *et al.* Changes in severe thunderstorm environment frequency during the 21st century caused by anthropogenically enhanced global radiative forcing. *Proceedings of the National Academy of Sciences* **104**, 19719–19723 (2007).

156. Diffenbaugh, N. S. & Field, C. B. Changes in Ecologically Critical Terrestrial Climate Conditions. *Science* **341**, 486–492 (2013).

157. Syvitski, J. P., Vörösmarty, C. J., Kettner, A. J. & Green, P. Impact of humans on the flux of terrestrial sediment to the global coastal ocean. *science* **308**, 376–380 (2005).

158. Goswami, B. *et al.* Abrupt transitions in time series with uncertainties. *Nat Commun* **9**, 48 (2018).

159. Zhu, Z. *et al.* Documentation strategy for facilitating the reproducibility of geo-simulation experiments. *Environmental Modelling & Software* **163**, 105687 (2023).

Acknowledgements

We are grateful to Markus Reichstein for his valuable comments on the manuscript. We thank the editors for handling our manuscripts. G.L. acknowledges funding by the National Natural Science Foundation of China under Grant No. 41930648. M.C. acknowledges funding from the National Key R&D Program of China under No. 2022YFF0711604. N.B. acknowledges funding by the Volkswagen foundation, by the European Union's Horizon 2020 research and innovation program under grant agreement No. 820970 and under the Marie Sklodowska-Curie grant agreement No. 956170, as well as by the German Federal Ministry of Education and Research under grant No. 01LS2001A.

Competing interests

The authors declare no competing interests.

Author contributions

M.C. and Z.Q. initiated the writing and co-led the design and writing of the article. G.L. conceptualized and supervised this article. All co-authors provided input on the manuscript text, figures and discussion of scientific content.