SCIENCE CHINA
Earth Sciences

•RESEARCH PAPER•

https://doi.org/10.1007/s11430-022-9999-9

Mechanism-learning coupling paradigms for parameter inversion and simulation in earth surface systems

Huanfeng SHEN[1,2†] & Liangpei ZHANG[3*]

[1] *School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China;*
[2] *Key Laboratory of Geographic Information System (Ministry of Education), Wuhan University, Wuhan 430079, China;*
[3] *State Key Laboratory of Information Engineering, Survey Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China*

Received March 27, 2022; revised August 9, 2022; accepted September 8, 2022; published online January 19, 2023

Abstract Building the physics-driven mechanism model has always been the core scientific paradigm for parameter estimation in Earth surface systems, and developing the data-driven machine learning model is a crucial way for paradigm transformation in geoscience research. The coupling of mechanism and learning models can realize the combination of "rationalism" and "empiricism", which is one of the most concerned research hotspots. In this paper, for remote sensing inversion and dynamic simulation, we deeply analyze the internal bottleneck and complementarity of mechanism and learning models and build a coupling paradigm framework with mechanism-learning cascading model, learning-embedded mechanism model, and mechanism-infused learning model. We systematically summarize ten specific coupling methods, including preprocessing and initialization, intermediate variable transfer, post-refinement processing, model substitution, model adjustment, model solution, input variable constraints, objective function constraints, model structure constraints, hybrid, etc., and analyze the main existing problems and future challenges. The research aims to provide a new perspective for in-depth understanding and application of the mechanism-learning coupling model and provide theoretical and technical support for improving the inversion and simulation capabilities of parameters in Earth surface systems and serving the development of Earth system science.

Keywords Mechanism model, Machine learning, Model coupling, Remote sensing inversion, Numerical simulation

Citation: Shen H, Zhang L. 2023. Mechanism-learning coupling paradigms for parameter inversion and simulation in earth surface systems. Science China Earth Sciences, 66, https://doi.org/10.1007/s11430-022-9999-9

1. Introduction

Problems such as climate change and environmental pollution in the Earth surface process profoundly affect the production, life, and health of human beings. To deeply understand the complex natural and humanistic phenomena on Earth surface systems and to promote sustainable social and economic development, comprehensive, complete, and continuous sensory data is required (Research Group of Geoscience Development Strategy, 2009). Satellite remote sensing inversion and dynamic numerical simulation are two important means to obtain macroscopic and continuous parameter data of Earth surface systems (Chen et al., 2019). How to continuously improve the accuracy and capability of remote sensinzg inversion and numerical simulation is a key fundamental issue in scientific research on Earth surface systems.

No matter through remote sensing inversion or dynamic numerical simulation, building physically interpretable mechanism models has always been a core scientific paradigm (De Bézenac et al., 2019). In remote sensing inversion, quantitative inversion based on the physical process of radiative transfer is the main way to obtain the parameters of multiple spheres such as hydrosphere, pedosphere, atmo-

† Corresponding author (email: shenhf@whu.edu.cn)
* Corresponding author (email: zlp62@whu.edu.cn)

sphere, and biosphere. Researchers have developed a large number of remote sensing inversion methods with strict physical mechanisms (Li, 2005; Liang et al., 2016; Li et al., 2016) and released a series of quantitative remote sensing parameter products at global and regional scales (Zhang et al., 2016). In the numerical simulation, scientists from various countries have built a variety of atmospheric numerical models (Skamarock et al., 2005), land surface process models (Meng and Dai, 2013), hydrological models (Arnold et al., 1998), etc., and developed Earth System Simulator based on supercomputing platforms (Chen et al., 2005; Qiu, 2021). In a word, mechanism models based on physical driving are the "main framework" for the inversion and simulation of parameters of Earth surface systems (De Bézenac et al., 2019) and an indispensable tool for geoscience knowledge discovery (Karpatne et al., 2017b).

In recent years, geosciences have witnessed a major revolution from being a data-poor field to a data-rich field (Karpatne et al., 2019), and people's ability to acquire and produce spatiotemporal data is far greater than the ability to process, analyze, and understand it (Reichstein et al., 2019). In this context, the fourth scientific paradigm based on big data has quietly emerged and has become crucial support for geoscience research (Guo et al., 2014; Song, 2016; Cheng et al., 2018; Deng et al., 2020; Zhou et al., 2021). Artificial intelligence technology represented by machine learning is developing rapidly, is considered the "Golden Key" to tapping the potential of big data (Guo R et al., 2020; Chen J et al., 2021; Li et al., 2022), and has received extensive attention and rapid development in the fields of satellite remote sensing and numerical simulation (Hsieh and Tang, 1998; Li and Ye, 2005; Gong, 2009; Härter and de Campos Velho, 2010; Zhang, 2018). In the data fusion contest organized by the IEEE Geoscience and Remote Sensing Society, the deep learning model has won championships in most tracks in recent years (Huang et al., 2021). And in quantitative applications, the machine learning model has been widely used in the remote sensing inversion of dozens of parameters (Guo Q et al., 2020; Yuan et al., 2020; Letu et al., 2020; Ran et al., 2021). At the same time, machine learning has also been successfully applied to the simulation and prediction of surface processes such as atmosphere (Navares and Aznarte, 2020), hydrology (Petty and Dhingra, 2018), ocean (De Bézenac et al., 2019), etc., and has shown great application potential. Given this, machine learning is expected to become a crucial framework for unleashing data-driven potential and accelerating scientific discovery (Karpatne et al., 2017b), and some scholars believe it has pushed geoscience research to the threshold of dramatic progress (Bergen et al., 2019).

Obviously, the machine learning model supported by big data has already made an impact on the orthodox mechanism model (Pei et al., 2019), and some scholars even believe that it may lead to "the end of theory" (Anderson, 2008). However, some scholars insist that the "big data hubris" problem (Lazer et al., 2014) is serious, and the effectiveness of machine learning is overestimated. For example, Google released the neural network precipitation forecast model MetNet (Sønderby et al., 2020), claiming that the neural network model already outperforms the mechanism model in the 8-hour forecast. But it has been questioned a lot in academia, and scholars believe that at least for long-term forecasting, large-scale forecasting, etc., the neural network model still cannot replace the mechanism model (Witt et al., 2020; Chantry et al., 2021). For the application of machine learning in geoscience, well-known journals such as Nature and Science have recently successively published papers (Bergen et al., 2019; Reichstein et al., 2019; Bauer et al., 2021), arguing that many factors, including the complexity, interactivity, multi-scale characteristics of the geoscience process, the uncertainty of data, the scarcity of realistic samples, etc., make machine learning model still unable to replace mechanism model, but the two models have natural complementary advantages, coupling the two is a promising development direction!

However, there are great challenges in coupling the explicit mechanism model with the implicit learning model. Although some research progress has been made, there is still a lack of a standard and unified paradigm framework. In this article, based on fully summarizing existing works, we build a systematical mechanism-learning coupling paradigm framework, analyze the characteristics and potentials of different coupling methods, and look forward to future challenges. We aim to provide theoretical and application references for related research, promote the development of remote sensing inversion and numerical simulation technology of parameters in Earth surface systems and provide support for improving the estimation ability of parameters in Earth surface systems and serving the development of Earth system science.

2. Advantages and bottlenecks of the current models

2.1 Mechanism model

"Mechanism" can be broadly understood as any knowledge that expresses the attributes or elements relationships of geographic objects (von Rueden et al., 2023), which includes not only physical knowledge but also geometric constraints and geological laws. Mechanism models follow objective laws to establish explicit associations between inputs and outputs, helping people to recognize and understand the physical world (Karpatne et al., 2017b). The classical quantitative remote sensing inversion methods are based on models such as atmospheric radiative transfer, which estab-

lishes the correlation between the electromagnetic wave signals of the earth observation and parameters, and realizes the area perception of Earth surface systems. The numerical simulation system obtains the continuous evolution of geographic objects in time and space through their intrinsic physical process and dynamic mechanism (Li et al., 2007). It can be seen that the mechanism model can describe the internal characteristics of the system clearly, and its outstanding advantages are rigorous theory, (relatively) stable models, and interpretable results, but it also has its insurmountable shortcomings:

(1) Limitations of mechanism understanding. Earth surface system is a complex giant system with multi-element mixing, multi-scale coupling, and multi-process interweaving (Chen M et al., 2021), and the existing mechanism models are still difficult to accurately describe all geoscience processes as some physical processes are still unknown. For example, there is still a lack of remote sensing mechanism inversion models for many parameters (such as air temperature, $PM_{2.5}$, etc.); and some sub-processes in the numerical model cannot be accurately modeled physically, where the reduction or approximation often leads to uncertainty.

(2) Underdetermined system problem. Even if the mechanism of some geoscience processes is relatively clear, the parameter inversion is often an underdetermined system (that is, the number of observation equations is less than the number of unknowns), which makes it very difficult to solve the model and usually requires some assumptions, but when the assumptions do not match the reality, it will bring a large solution error. For example, the remote sensing inversion of surface temperature is to use N observations (number of bands) to solve the ill-conditioned problem of N+1 unknowns (N surface emissivity and surface temperature).

(3) Computational burden problem. The computational complexity of some mechanistic processes is enormous. For example, in the atmospheric model of the US National Center for Atmospheric Research, the calculation of physical processes takes about 70% of the total model computations (Krasnopolsky et al., 2005). If the requirements in terms of resolution and consistency are further improved, the amount of calculation will increase exponentially, which will bring great application troubles.

2.2 Learning model

The machine learning model simulates the human "induction" and "inference" process through "training" and "prediction" and realizes the modeling and solution of typical problems. Different from the explicit expression of the mechanism model, the learning model establishes the implicit association between variables through training data, namely the "black box" model. One of the key advantages of the learning model is that when the mechanism is unknown, data-driven modeling can be performed directly skipping the understanding of the physical process. It can usually obtain high modeling accuracy when sufficient training data is available. In addition, although machine learning is time-consuming in the training phase, it generally has high computational efficiency in the testing application phase, which has also become one of its outstanding advantages. Nevertheless, machine learning models still have many limitations, especially in complex geoscience applications, which often have the following problems:

(1) Insufficient generalization. Lack of sufficient training samples is the most common problem in geoscience applications of machine learning. Overfitting is prone to occur when machine learning models learn complex geoscientific processes on limited data; even if the training samples show high modeling accuracy, the accuracy of the test application will be greatly reduced. Moreover, when the actual numerical range, variable relationship, etc., are not covered by the training samples, the prediction results are more likely to have great deviations, which is a typical problem of insufficient generalization ability.

(2) Insufficient transferability. Regionality is an essential feature of geography, and different regions are not only represented by differences in geographical elements but also by differences in the relationship between various elements. Thus, machine learning models trained in one area are often difficult to transfer to other areas for application. In addition, elements of Earth surface system and their interrelationships are in the process of continuous change, and the influence of human activities has made the changes more severe, so the models of different time spans in the same area are often difficult to generalize. Furthermore, insufficient scale transferability is another dilemma in geoscience applications.

(3) Insufficient interpretability. The goal of scientific research is not only to develop a usable model, but also to discover the internal causal relationships and driving patterns between different variables, and use them to explain theories and hypotheses, thereby promoting the advancement of scientific knowledge (Karpatne et al., 2017a). A particular problem of machine learning is the lack of interpretability. Although it can obtain relatively high accuracy under certain conditions, it cannot explain the internal mechanism process.

It can be seen from the above analysis that although the mechanism model and the learning model have their respective advantages, they have insurmountable shortcomings, and there is an obvious natural complementarity between the two (Ganguly et al., 2014; Wu et al., 2015). The coupling of mechanism and learning models can realize the combination of "rationalism" and "empiricism", can effectively adjust the "bias" of the mechanism model, and avoid the "hubris" of the learning model (Chantry et al., 2021), so it is an inevitable choice.

3. Coupling paradigms of mechanism and learning models

The coupling of mechanism and learning models has recently become a research hotspot in various fields. In fact, since the end of the last century, whether in the field of numerical simulation (Chevallier et al., 1999) or remote sensing inversion (Aires et al., 2001), there have been ideas and successful cases of the coupling of mechanism and learning models; however, limited by the level of cognition and technical conditions, the research in this direction has not received enough attention and development. Until recently, with the re-emergence of neural networks, especially deep learning technology, the mechanism-learning coupling has become a research hotspot in many fields, including geoscience.

In recent years, many terms have appeared in literature to express the coupling of mechanism and learning models, which can be represented by any combination of optional words in each of the three columns shown in Figure 1, such as "*Physics Informed Machine Learning*". However, the various combination terms above place too much emphasis on "learning" and put "mechanism" on the back burner. Actually, there are various coupling modes between the two, where the proportions of "mechanism" and "learning" are different, so it would be best to maintain a balance between the two, for which Shen et al. (2022) proposed the expression of "*Coupling of Mechanism and Learning*".

In this article, we propose that in the inversion and simulation of geoscience parameters, the coupling of mechanism and learning models can be classified into three basic paradigms: mechanism-learning cascading model, learning-embedded mechanism model, and mechanism-infused learning model (hereinafter referred to as cascading, embedded, infused), as shown in Figure 2. Mechanism learning cascading is to connect these two models in series, and the output of one model is used as the input of the other model. Learning-embedded mechanism model uses the mechanism model as the main and the learning model as a supplement, which embeds the learning model into the mechanism model to replace or optimize the original uncertain process. Mechanism-infused learning model takes the learning model as the main framework and integrates physical knowledge into it to realize the constraint guidance of the learning process. In addition, the three paradigms can be combined into a hybrid mode to take advantage of them.

3.1 Mechanism-learning cascading model

The simplest coupling between the mechanism model and the learning model is cascading, which is a straightforward combination through sequential modeling. According to the functional stage and importance of the two models in the whole system, it can be divided into three specific approaches: preprocessing and initialization, intermediate variable transfer, and post-refinement.

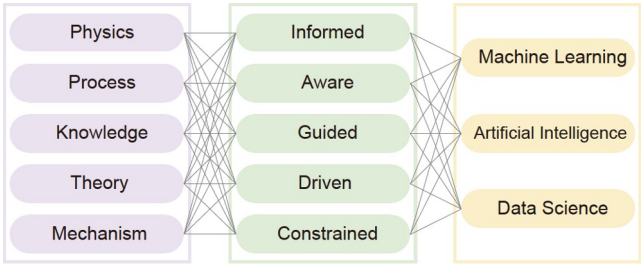

Figure 1 English terms of mechanism-learning coupling.

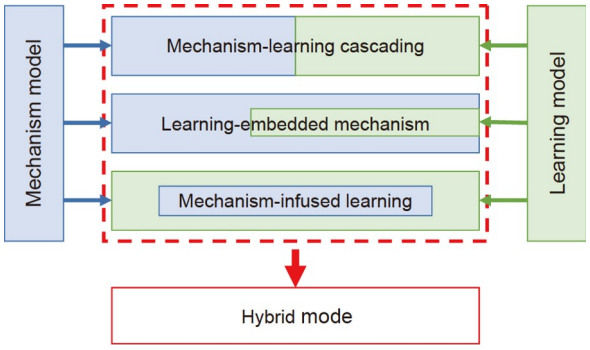

Figure 2 The basic paradigm of mechanism-learning coupling.

3.1.1 *Preprocessing and initialization*

(1) Quality control. Quality control of the input data in the mechanism model using the learning model can effectively improve the accuracy of subsequent parameter estimation. For example, remotely sensed data often suffer from noise, missing pixels, etc. Before the parameter inversion based on the mechanism model, machine learning is firstly used for image denoising, gap-filling, and other processing, which can effectively improve the accuracy and reliability of the mechanism model.

(2) Parameter optimization. Mechanism models often require multiple input parameters, and their effectiveness is significantly limited by the accuracy of the input parameters (Zhang et al., 2012). Machine learning can be used to obtain more accurate parameters to provide better initial conditions for subsequent calculations in the mechanism model. For example, Beck et al. (2016) proposed a scheme for regionalization of hydrologic model parameters, which was successfully applied on a global scale; Sawada (2020) used a Gaussian process regression model to optimize the parameters of the land surface model, effectively improving the performance of the model simulation.

(3) Sample generation. In many geoscience applications, the real database required for training machine learning models is often difficult to obtain. At this point, the me-

chanism model can be used to generate training samples. For example, Aires et al. (2001) generated a training database with the radiative transfer model in microwave remote sensing. The parameters such as atmospheric water vapor, land surface temperature, and emissivity were then retrieved based on the machine learning method. Besides, radiative transfer models have been widely used to generate sample databases for machine learning in applications such as land surface temperature retrieval based on thermal infrared remote sensing (Mao et al., 2007), leaf area index retrieval based on optical remote sensing (Campos-Taberner et al., 2016), gross primary productivity retrieval (Wolanin et al., 2019), and vegetation water content retrieval (Trombetti et al., 2008).

(4) Transfer learning. To avoid overfitting due to insufficient real samples, the mechanism model can be used to generate coarser training datasets for pre-training. When the model is relatively stable, it is then finely trained based on a small number of high-precision real samples (Figure 3). To predict lake water temperature, Jia et al. (2021) simulated datasets with a physics-based General Lake Model, which were used to pre-train a long short-term memory network, effectively reducing the reliance on real samples (Read et al., 2019).

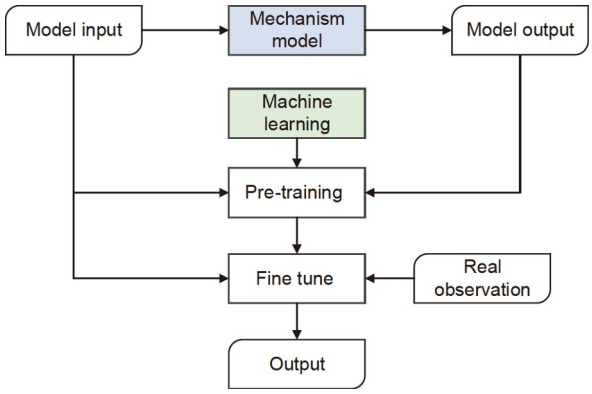

Figure 3 The coupling approach of transfer training.

3.1.2 *Intermediate variable transfer*

Some parameters affected by factors such as mechanism cognition and technical limitations are difficult to obtain by full physical processes. At this point, the joint application of the mechanism model and the learning model can be realized through the transfer of intermediate variables. That is, the intermediate variables are calculated using the mechanism model, and then the target parameters are estimated using the machine learning model. For example, effective full-physics inversion models in remote sensing are still lacking for parameters such as near-surface air temperature and atmospheric $PM_{2.5}$. However, land surface temperature and aerosol optical depth have strong correlations with air temperature and $PM_{2.5}$, respectively, which have relatively mature physical inversion methods. Therefore, the land surface temperature and aerosol optical depth estimated by mechanism models can be used as inputs to machine learning models for the retrieval of air temperature and $PM_{2.5}$ (Shen et al., 2018, 2020), as shown in Figure 4. In addition, intermediate variables can also be estimated by numerical simulations of dynamic models (Xiao et al., 2017). For example, Liang et al. (2020) forecasted the concentration of Chlorophyll using a long short-term memory network, based on six water quality variables simulated by the water quality model.

3.1.3 *Post-refinement*

In order to enhance the outputs of mechanism models such as remote sensing inversion and dynamic simulation in terms of accuracy, resolution, etc., machine learning models can be used for post-refinement, which is a relatively traditional way of coupling the mechanism model with the learning model. It includes various specific categories such as error correction, downscaling, and ensemble optimization.

(1) Error correction. Machine learning-based error correction methods have been widely used for the calibration of parameters from remote sensing inversions and model simulations. The accuracy or consistency of the original outputs is improved with post-correction by establishing a mapping relationship between the model outputs and ground observations or other reference data. Rasp and Lerch (2018) used neural networks for the systematic error correction of ensemble weather forecasts, which improved the original model considerably in terms of accuracy and efficiency; Ivatt and Evans (2020) corrected the outputs of the atmospheric chemistry transport model through gradient-boosted regression trees, which effectively improved the simulation accuracy of ozone; Noori et al. (2020) calibrated the outputs of the SWAT hydrological model using a machine learning approach with station observations as references, effectively

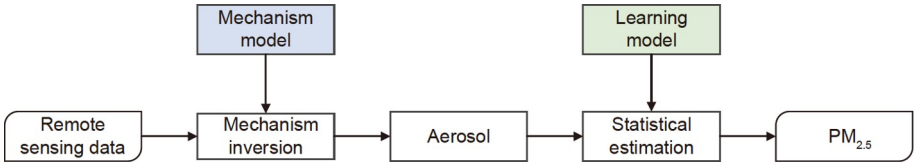

Figure 4 Schematic diagram of $PM_{2.5}$ retrieval with mechanism-learning cascading.

improving the simulation accuracy of three key water quality parameters.

(2) Downscaling. The spatial resolution of data from large-scale remote sensing inversions and model simulations is generally coarse, which cannot easily meet the requirements of fine monitoring and analysis. Based on the inversion or simulation of mechanism models, machine learning can be further used for downscaling to improve the spatial resolution of the outputs. Currently, machine learning has become a general downscaling method for remote sensing parameters such as precipitation (Wang et al., 2021), soil moisture (Alemohammad et al., 2018), and surface temperature (Li et al., 2019). Meanwhile, machine learning models such as the neural network (Wilby et al., 1998; Cannon, 2011) and support vector machine (Ghosh, 2010) have been widely used to downscale numerical simulation data. In addition to conventional downscaling methods, machine learning super-resolution techniques in image processing have also been introduced to improve the resolution of Earth System Model outputs (Vandal et al., 2017).

(3) Ensemble optimization. Due to the limitations of mechanism cognition and differences in parameterization schemes, the output results of different mechanism models often have large inconsistencies. Combining the outputs from different models is an effective way to obtain more reliable results. In the field of machine learning, ensemble learning can achieve "drawing upon the strengths of others" by combining multiple machine learners, which is widely used for land cover classification and mapping in remote sensing (Du and Samat, 2013). Similarly, machine learning can also realize ensemble optimization of multiple mechanism models, as shown in Figure 5. Monteleoni et al. (2011) integrated the predictions of multiple climate models based on the hidden Markov model with improved accuracy; On this basis, McQuade and Monteleoni (2012) further developed a multi-model ensemble framework with higher spatial resolution. Krasnopolsky and Lin (2012) used a neural network to integrate multiple models, which effectively improved the accuracy of precipitation forecasting.

3.2 Learning-embedded mechanism model

Leveraging the advantages of the mechanism model, such as physical interpretability, the learning model is embedded into the mechanism model to replace, adjust or optimize the solution of the original uncertain sub-processes. It is a typical coupling paradigm with a dominant mechanism model and a supplementary learning model, which is a hot topic in current mechanism-learning coupling research.

3.2.1 Model replacement

Model replacement is a coupling approach in which machine learning is used to replace sub-processes of the mechanism model, as shown in Figure 6. In the modeling process of the mechanism model, especially the dynamic model, the spatial scale of some sub-processes is likely to be smaller than that of the original model, so it is difficult to directly model with a strict mechanism model. Thus, it is necessary to design an applicable parameterization scheme for expression. Parameterization is a processing scheme for the indirect expression of physical processes that cannot be directly modeled and is an approximate or idealized representation of complex physical processes (Stensrud, 2007). Therefore, the parametrization of the model is fundamentally different from the

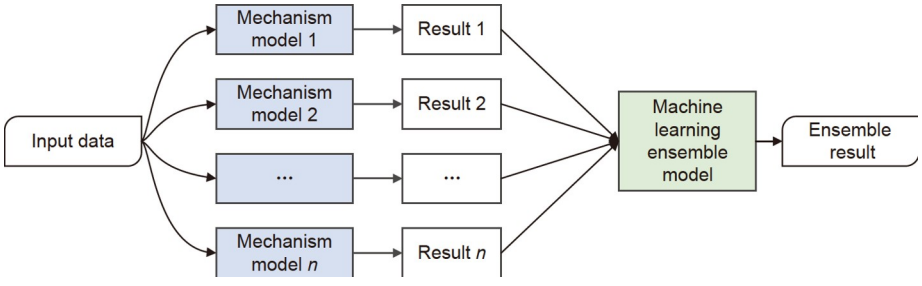

Figure 5 The coupling approach of ensemble optimization.

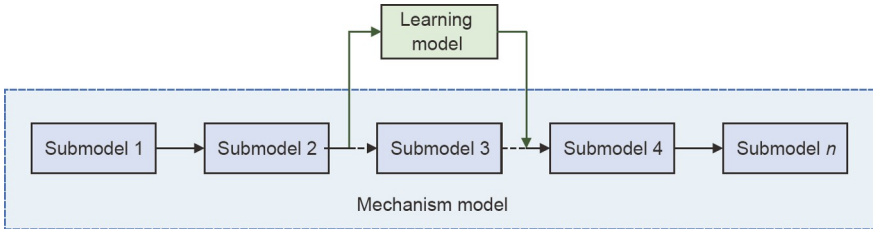

Figure 6 The coupling approach of model replacement.

aforementioned parameter optimization. The most widely used coupling approach for model replacement is that the machine learning model is used to replace the parameterization scheme in the mechanism model.

(1) Model "emulator". Since the computation of some parameterization schemes is very time-consuming, a common alternative is to build a machine learning "emulator" of the mechanism model by training the input-output data pairs of the sub-model to improve the computational efficiency, so that it has close accuracy and higher efficiency compared to the original model. Chevallier et al. (1999) applied machine learning to the construction of a new generation of radiative transfer models. Embedding the multi-layer perceptron into the physical modeling process to replace long-wave radiation from the top of the atmosphere to the land surface. The computational efficiency is 22 times higher than that of the traditional band model and 10^6 times higher than that of the line-by-line model. Subsequently, the method and its improvements were operationally applied to the 4D Variational Assimilation System of the European Centre for Medium-Range Weather Forecasts. For the Community Atmosphere Model (CAM) of the National Center for Atmospheric Research, Krasnopolsky et al. (2005) introduced a neural network to simulate and replace the original long-wave radiation parameterization method and further applied it to the parameterization of convection and other processes (Krasnopolsky et al., 2013), which can improve the computational efficiency by 10–10^5 times compared to the original model (Krasnopolsky, 2020). Based on the GEOS-Chem atmospheric chemical transport model, Keller and Evans (2019) attempted to replace its chemical integrator with a random forest approach, resulting in a viable alternative that provided an important foundation for efficiency optimization.

Since the machine learning model can replace some subprocesses of the mechanism model and achieve similar accuracy as the original model, it is natural to wonder whether they can replace more sub-processes or even the whole complex mechanism model. Sargsyan et al. (2014) used a sparse learning method to simulate the land surface model, which showed certain application potential. Based on the global forecast system of the National Centers for Environmental Prediction, Krasnopolsky et al. (2009) attempted to replace all subprocesses except radiative transfer with machine learning models and found that not all outputs could reach the level of the original model. Dueben and Bauer (2018) constructed an atmospheric model emulator using deep learning, which performed well for regional short-term predictions, but struggled to achieve the expected results for long-term predictions. Scher and Messori (2019) showed that using machine learning to replace the whole mechanism process of atmospheric models including a seasonal cycle remains challenging.

(2) Model "enhancer". Machine learning alternatives can further improve the estimation accuracy if sufficient real samples exist. In the simulation of ocean parameters, Bolton and Zanna (2019) further optimized the model by introducing real observations and machine learning. Even with only local observations, the prediction accuracy of the model can be improved on a large scale. Hunter et al. (2018) effectively improved the prediction of salinity by embedding the neural network and simple regression model in the simulation of river parameters. Kraft et al. (2022) embedded a neural network into global hydrological models for the simulation of parameters such as soil moisture, groundwater, and snow, and obtained better local adaptivity than the mechanism model. It can be seen that replacing uncertain mechanism processes with machine learning is an effective way to enhance models by fully utilizing high-precision ground-based observations, satellite remote sensing images, and other data.

However, the training samples required for machine learning are often difficult to obtain. To this end, simulated data can be generated using a higher-resolution mechanism model, which can be used as "pseudo-observation" data to train the learning model, and then the trained model can be applied to the lower-resolution mechanism model, as shown in Figure 7. This approach has been widely used in parameterization schemes of atmospheric models (Krasnopolsky et al., 2013; Schneider et al., 2017; Brenowitz and Bretherton, 2018). It has been proved to be effective in capturing the spatiotemporal information on the sub-grid scale, obtaining higher accuracy than the original parameterization schemes, and even having better prediction ability for extreme events (Krasnopolsky et al., 2009).

3.2.2 Model adjustment

As mentioned above, existing global and regional dynamic models often contain complex parameterization schemes, leading to uncertainties in model outputs (Li et al., 2007). Data assimilation technology can integrate direct or indirect observations from different sources and at different resolutions within the dynamic framework to effectively adjust the trajectory of the mechanism model, thus enhancing the accuracy and predictability of the model (Li X et al., 2020, 2021). Variational approaches and Bayesian filtering are two types of data assimilation methods commonly used at present. Some scholars have mathematically analyzed the theoretical equivalence of data assimilation and machine learning (Bonavita et al., 2021). In recent years, the application of machine learning methods to data assimilation has become a hot research direction. Based on data assimilation, embedding machine learning methods into the dynamic framework of model simulation is an effective way to couple mechanism and learning. The difference between this approach and the aforementioned model replacement is that it optimally adjusts the model instead of directly replacing the

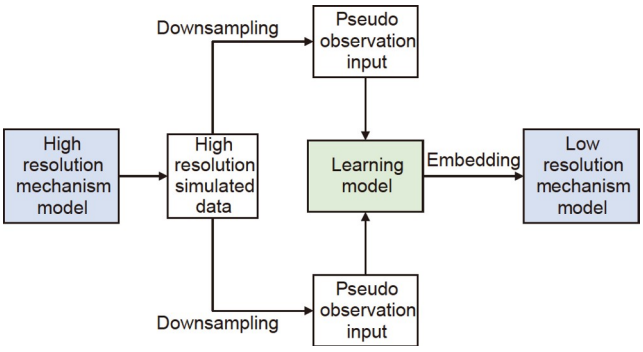

Figure 7 The model replacement with pseudo observations.

original mechanism process.

Hsieh and Tang (1998) earlier used machine learning for data assimilation in meteorological and oceanographic models. Researchers have theoretically explored data assimilation using machine learning models such as the neural network (Härter and de Campos Velho, 2008) and support vector machine (Gilbert et al., 2010), and have gradually applied them to real applications. There are three main ways for data assimilation with machine learning: Firstly, using machine learning to simulate the existing assimilation algorithm, the purpose of which is to improve the efficiency of assimilation processing. For example, the neural network approach is 274 times more efficient than the Local Ensemble Transform Kalman filter for the same accuracy in a global surface temperature assimilation study (Cintra et al., 2016). Secondly, developing new machine learning assimilation methods. For example, Lu et al. (2018) effectively improved the prediction accuracy of precipitation using a neural network-based assimilation algorithm. Finally, combining machine learning with the existing data assimilation method to improve the applicability of the model through error correction (Bonavita and Laloyaux, 2020; Farchi et al., 2021).

3.2.3 Model solution

In some parameter estimation processes, the optimal model is often established based on the forward process and the related physical mechanism, and solved by gradient descent iterative process, etc. However, problems often arise in the specific solving process, such as the gradient cannot be calculated or the computation is too large even if it can be solved. In this case, the model can be optimally solved by machine learning. In terms of theoretical research, machine learning is applied to solve partial differential equations, which has received extensive attention in the field of applied mathematics (Han et al., 2018). In terms of applications, Davis et al. (1993) trained the scattering model using a neural network to obtain a conversion model from parameters to bright temperature in the retrieval of passive microwave snow parameters, which was used in an iterative inversion algorithm. Based on a similar solution idea, various parameters were further retrieved, such as soil moisture, near-surface air temperature, vegetation water content, etc. (Davis et al., 1995). Venkatakrishnan et al. (2013) developed a "plug-and-play" mechanism-learning coupling method, which can embed the machine learning model into the iterative solution of variational optimization for remote sensing applications such as SAR data reconstruction (Alver et al., 2019) and multi-source data fusion (Dian et al., 2021).

3.3 Mechanism-infused learning model

The third type of coupling paradigm is to integrate the mechanism knowledge into the machine learning model, that is, machine learning is the main framework, and the mechanism knowledge is used to constrain the learning process. The whole model adopts an "end-to-end" computing form. According to the application mode and role of mechanism constraints in the machine learning model, it can be divided into input variable constraints, objective function constraints, model structure constraints, etc., as shown in Figure 8 (taking neural network as an example).

3.3.1 Input variable constraints

Input variable constraints refer to the introduction of new input variables into the machine learning model through the calculation of the mechanism model or the guidance of mechanism knowledge, so that the learning process is more in line with specific mechanism constraints. For example, in the study of Karpatne et al. (2017b) (see Figure 9), the driving data is used as input to carry out the mechanism simulation of the dynamics, and then the output data of the mechanism simulation and the original driving data are used as the input variables of the machine learning model. At this point, there is a corresponding physical mapping relationship between the two sets of input variables in machine learning. Experiments show that this coupling method has higher prediction accuracy than pure data-driven models. Li et al. (2017) introduced a spatiotemporal correlation factor into the input variables in the remote sensing parameters inversion, which effectively took into account the first law of geography and imposed effective spatiotemporal geography constraints on the machine learning model.

3.3.2 Objective function constraints

In general, machine learning solves the model by minimizing the objective function. Therefore, introducing mechanism constraints to the objective function is a straightforward and widely used integration method (Kashinath et al., 2021). Overall, the objective function of the mechanism-constrained neural network can be summarized as the following basic form (Karpatne et al., 2017b; Willard et al., 2020):

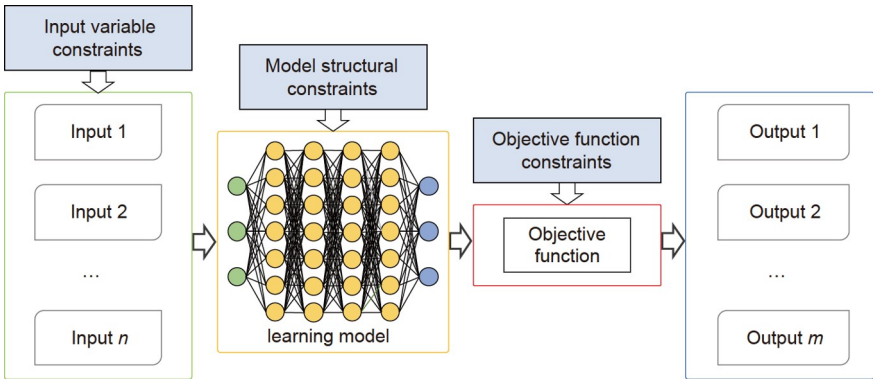

Figure 8 Mechanism constraints of the neural network model.

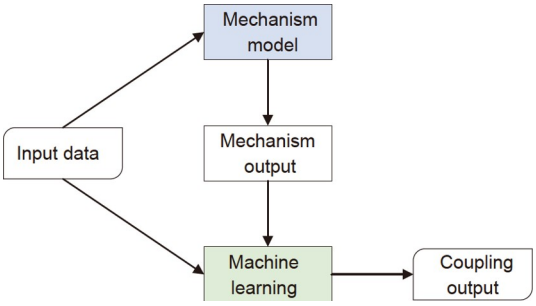

Figure 9 A form of input variable constraint.

$$L = L_{\mathrm{d}}(x_{\mathrm{true}} - x_{\mathrm{pred}}) + \alpha R(w) + \beta L_{\mathrm{phy}}(x_{\mathrm{pred}}), \quad (1)$$

where, the first item L_{d} represents the supervision error between the real sample data x_{true} and the model prediction data x_{pred}, which can be defined as the sum of squared errors, absolute error, cross entropy, etc.; the second term $R(w)$ is a general regularization term, used to compress the solution subset, w is the solution parameter of the model; the third term $L_{\mathrm{phy}}(x_{\mathrm{pred}})$ is the constraint imposed on the basis of general regularization based on specific mechanism knowledge to further narrow the search space for parameter solving and overcome the overfitting problem (Reichstein et al., 2019); α and β are hyperparameters used to adjust the weights.

The $L_{\mathrm{phy}}(x_{\mathrm{pred}})$ can directly impose corresponding constraints according to the distribution characteristics of the predictor variable x_{pred}. Erichson et al. (2019) added Lyapunov stability constraints to the objective function, which effectively reduced the uncertainty of sea surface temperature prediction. To strengthen the constraint ability, the variable z is widely introduce that has a mechanism relationship with x_{pred}. It can be either a model input variable or other related variables. Then, the constrain can be expressed in the form of $L_{\mathrm{phy}}=\mathcal{L}(z, A, x_{\mathrm{pred}})$, where A is the mechanism correlation model, and $\mathcal{L}$ is the penalty function. For example, Karpatne et al. (2017b) made full use of the physical relationship equation of temperature and density into the lake temperature simulation, and applied the relationship constraint of density to depth in the construction of L_{phy}. This method is further improved to construct L_{phy} based on constraints on the input-output heat flux, so that the predicted temperature and lake water environmental changes conform to the law of conservation of energy (Read et al., 2019; Jia et al., 2021). In the long-wave radiation simulation process, Beucler et al. (2019) simultaneously considered the conservation laws of heat, mass, solar radiation, and surface radiation, and imposed corresponding physical constraints in the objective function. Besides, in the fusion and downscaling of remote sensing data, the forward model between the input data y and the output data x_{pred} can be used for the construction of L_{phy} (Lin et al., 2022), such as $L_{\mathrm{phy}} = \left\| y - Ax_{\mathrm{pred}} \right\|^2$, that is, the model is constrained by the known relationship matrix **A** to improve the fidelity of the model solution.

In specific applications, the introduction of domain knowledge can be achieved by directly improving L_{d}. For example, when the target variable of machine learning does not have corresponding real sample data, the objective function cannot be directly constructed. However, if there is an associated variable z that has a definite mechanism relationship with the target variable, the objective function can be constructed indirectly based on the mechanism relationship between the two variables, such as $L_{\mathrm{d}} = \left\| z - Bx_{\mathrm{pred}} \right\|^2$, where B is transformation models between variables. De Bézenac et al. (2019) took the motion field parameters as the target variable of the neural network in the estimation of sea temperature, and used the physical relationship between the sea temperature and motion field parameters to establish an energy function to realize the joint solution of the two. Without introducing associated variables, Li T et al. (2021) established a spatiotemporal geographic weighting constraint function $L_{\mathrm{d}} = \left\| w(x_{\mathrm{true}} - x_{\mathrm{pred}}) \right\|^2$ (w is the spatiotemporal weight) in remote sensing quantitative inversion,

that is, by taking into account the autocorrelation characteristics of the variables, the inversion accuracy of the model is effectively improved.

3.3.3 Model structural constraints

Generally, the solution process of machine learning is a "black box". Imposing constraints on the "black box" is a way of introducing mechanism knowledge. This kind of coupling is a challenging way, which needs to have a clear understanding of the internal structure of machine learning and mechanism process, and it is also necessary to find the sweet spot between them. Li T et al. (2020) developed a spatiotemporal geographic weighted learning method, which improved the pattern layer and summation layer of the neural network structure, and multiplied the weighted summation nodes and arithmetic summation nodes by the corresponding spatiotemporal weights respectively to fully consider the geoscience laws associated with space and time. This method is similar to the aforementioned objective function spatiotemporal constraint method (Li T et al., 2021), but this method is difficult to apply to other neural network structures. In the simulation of lake water temperature, Daw et al. (2020) added an activation function directly behind the original long short-term memory neural network, and used the output of the activation function to illustrate the constraint relationship between lake water depth and density. Beucler et al. (2019) added a physical relationship expressing energy conservation to the back of the neural network (Figure 10) as a model structure constraint method and compared it with an objective function constraint method. It is shown that both methods can effectively improve the simulation of longwave radiation.

Besides neural network models, some machine learning models have been proposed for model structure constraints. For example, Gaussian process regression is a nonparametric model that regresses data using a Gaussian process prior. It works well for small sample data and can analyze forecast uncertainty (Willard et al., 2020). Camps-Valls et al. (2018) put physical constraints on the relationship between multiple variables by introducing differential equations into the Gaussian process machine learning model for the multi-output regression problem, and took the leaf area index and photosynthetically active radiation as example to verify the effectiveness of the model.

3.4 Comparison and hybrid application of coupling paradigms

As mentioned above, the coupling between the mechanism model and the learning model includes three basic paradigms: cascade, embedding, and integration, each of which has its own advantages and limitations. The advantage of the cascading paradigm is that it is simple to apply, does not require any changes to the internal processes of the two models. It is suitable for most application scenarios, and can significantly improve accuracy. However, the theoretical breakthrough of the cascade paradigm is limited, it lacks a fundamental solution to the model problem, and it is difficult to give full play to the complementary advantages of the models. Relatively speaking, the embedding paradigm can make targeted improvements according to the shortcomings of the mechanism model. Since it maintains the basic structure of the mechanism model and has stronger physical interpretability, it is more suitable for application scenarios where the mechanism model is relatively mature. However, replacing the sub-process of mechanistic models with machine learning requires a large number of intermediate variables for training, and the availability of these data becomes a limiting condition in some applications. The integration paradigm maintains the "end-to-end" computing framework of the machine learning model, and realizes efficient processing and application by integrating the mechanism knowledge into the learning model. It is more suitable for scenarios where the mechanism model is immature and there are a large number of real training samples. However, it is difficult to modify the "black box" model structure of machine learning, and the introduction of mechanism knowledge will be restricted accordingly.

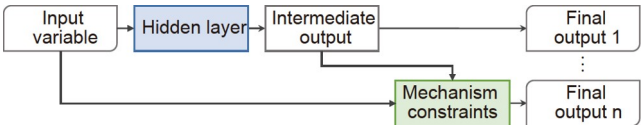

Figure 10 Structural constraints of neural network models.

Therefore, the three coupling paradigms have no absolute advantages or disadvantages, which have different application scenarios for different conditions, and they can also be mixed in some applications. For example, Schneider et al. (2017) used neural networks to replace parameterization schemes in Earth System Models (ESM), which belongs to the embedding coupled paradigm. However, it incorporates physical constraints into the objective function of the neural network, which further uses the integration coupled paradigm. Read et al. (2019) integrated three coupling paradigms in the study of water temperature estimation (Jia et al., 2021). First, the cascade coupling paradigm is applied to generate simulated data based on the mechanism model, which is further used for the pre-training of the machine learning model; secondly, in the fine training stage, the driving data and the output of the mechanism model are used as the input of the machine learning model, that is, the integration coupling paradigm; Furthermore, a physical constraint is added to the energy function of the machine learning model, that is, the embedding coupling paradigm. Through the joint application of different coupling paradigms, the complementary

advantages of the mechanism model and the learning model can be fully utilized.

4. Main issues and challenges

Although the mechanism-learning coupling has been explored in the field of remote sensing inversion and numerical simulation, gratifying progress has been made in some typical applications. However, this research direction is still in a relatively primary development stage, and needs to be further developed in both breadth and depth. In terms of breadth, there are relatively many studies on atmospheric numerical simulation, but there are still few studies on land surface process and hydrological simulation, remote sensing parameter inversion, and comprehensive breakthroughs in all directions are urgently needed. In terms of depth, there is still a huge research space on how to embed a robust learning process in the mechanism model and how to integrate more complex mechanism knowledge into the machine model architecture. Under the background of the rapid development of geoscience big data and artificial intelligence, the coupling research of mechanism model and learning model faces unprecedented opportunities and challenges, including but not limited to:

(1) Integrated learning and fusion of multi-source heterogeneous geoscience big data. Machine learning is a data-driven computing paradigm, so the coupling of mechanism and learning is largely dependent on the available reference data. Although various types of data such as ground-based observations, remote sensing observations, numerical simulations, and social perception have been developed rapidly, the existing model coupling research is mostly aimed at single-type or few-type data. How to take into account the differences and complementarity of multi-source heterogeneous data in terms of accuracy, scale, spatiotemporal continuity, and integrate machine learning modeling and fusion applications is an important development trend (Zhang and Shen, 2016). Besides, due to the lack of real reference data, making full use of multi-source and multi-scale observation and simulation data to obtain more sufficient training samples through transfer learning and active learning is an effective way to improve the performance of existing models.

(2) Selective surrogate modelling of mechanism process. Using machine learning to surrogate the uncertain sub-process in the mechanism model is a coupling paradigm with great potential, which maintains the original physical process mechanism and has strong physical interpretability. However, in the specific implementation process, which sub-process mechanism model is replacement? When is the replacement required? Which machine model to use instead? These are all affected by a variety of factors, and currently it is mainly selected by domain experts based on experience, which brings certain uncertainties. Therefore, the future development trend is to construct adaptive learning alternative mechanisms for mechanistic processes. That is, according to the operation of the mechanism model and the current data conditions, the computer automatically determines which machine learning model to use, when and where to replace the mechanism sub-process (von Rueden et al., 2020).

(3) New architecture of deep learning network for geosciences. As the most representative machine learning method at present, deep learning is becoming an important technical support for the estimation of earth surface parameters and the coupling of mechanism and learning. However, existing methods are mainly based on general neural network architectures, such as Google neural network (GoogLeNet) (Szegedy et al., 2015), densely connected convolutional network (DenseNet) (Huang et al., 2017), residual network (ResNet) (He et al., 2016), deep belief network (DBN) (Hinton et al., 2006), etc. On the original basis, the introduction of mechanism or prior knowledge through structural modification will affect the improvement effect of the model due to the limitation of the inherent network structure. Therefore, how to design a new deep neural network architecture according to the uniqueness of geoscience applications is a challenging direction to break through the existing limitations.

(4) The coupling of geoscience knowledge graph and machine learning. Knowledge-guided machine learning methods are an important development direction. The expression of knowledge has various forms, and knowledge graph is one of the most concerned directions at present (von Rueden et al., 2023). Geoscience Knowledge Graph is a geoscience knowledge base and "inference engine" that can be understood and calculated by machines. It is a knowledge system formed by effectively organizing relevant geoscience knowledge in a structured graph mode, which is used to express various geographical entities, concepts and the semantic relationships between them (Zhou et al., 2021). As we all know, symbolism, connectionism and behaviorism are the three major schools of artificial intelligence, and knowledge graph and deep learning are the representatives of symbolism and connectionism, respectively. The combination of the two in the field of geoscience is bound to have great application potential.

5. Conclusion

For the inversion and simulation of parameters in Earth surface systems, mechanism models often have problems such as cognitive limitations, underdetermined systems, and computational burdens, while learning models often have shortcomings in generalization, transferability, and inter-

pretability. The coupling of mechanism and learning models can effectively adjust the "bias" of mechanism models and avoid the "hubris" of learning models (Chantry et al., 2021), which is an important concern in many disciplines, such as geoscience. In this article, for the remote sensing inversion and model simulation of parameters, we establish a coupling paradigm framework of mechanism-learning cascading model, learning-embedded mechanism model, mechanism-infused learning model, and their hybrid applications. We systematically summarize ten coupling modes based on specific application examples, and look forward to prospects such as integrated learning and fusion, selective surrogate modelling, new deep learning network architecture, and the coupling of knowledge graph and deep learning. Mechanism-learning coupling is a combination of "rationalism" and "empiricism", which will become a "booster" for the development of Earth science research (Bergen et al., 2019). It is noteworthy that the mechanism model involves rigorous geoscientific processes and physical derivation, while the learning model needs to establish a complex information transmission mechanism. Thus, there is an urgent need for multi-disciplinary cross-integration to break through the key scientific problems of mechanism-learning coupling, improve the accuracy and efficiency of inversion and simulation of parameters in Earth surface systems, and show stronger support in Earth system scientific research and response to resource and environmental problems.

Acknowledgements *The research and writing of this article have benefited from the inspiration, advice, and help of many experts in related fields, as well as the discussion with our research team members and their assistance in pictures and texts. I would like to express my sincere thanks to them here. This work was supported by the National Natural Science Foundation of China (Grant No. 42130108).*

References

Aires F, Prigent C, Rossow W B, Rothstein M. 2001. A new neural network approach including first guess for retrieval of atmospheric water vapor, cloud liquid water path, surface temperature, and emissivities over land from satellite microwave observations. J Geophys Res, 106: 14887–14907

Alemohammad S H, Kolassa J, Prigent C, Aires F, Gentine P. 2018. Global downscaling of remotely sensed soil moisture using neural networks. Hydrol Earth Syst Sci, 22: 5341–5356

Alver M B, Saleem A, Cetin M. 2019. A novel plug-and-play SAR reconstruction framework using deep priors. Boston: Proceedings of the 2019 IEEE Radar Conference (RadarConf)

Anderson C. 2008. The end of theory: The data deluge makes the scientific method obsolete. Wired Magazine, 16: 16–07

Arnold J G, Srinivasan R, Muttiah R S, Williams J R. 1998. Large area hydrologic modeling and assessment part I: Model development. J Am Water Resour Assoc, 34: 73–89

Bauer P, Dueben P D, Hoefler T, Quintino T, Schulthess T C, Wedi N P. 2021. The digital revolution of Earth-system science. Nat Comput Sci, 1: 104–113

Beck H E, van Dijk A I J M, de Roo A, Miralles D G, McVicar T R, Schellekens J, Bruijnzeel L A. 2016. Global-scale regionalization of hydrologic model parameters. Water Resour Res, 52: 3599–3622

Bergen K J, Johnson P A, de Hoop M V, Beroza G C. 2019. Machine learning for data-driven discovery in solid Earth geoscience. Science, 363: aau0323

Beucler T, Rasp S, Pritchard M, Gentine P. 2019. Achieving conservation of energy in neural network emulators for climate modeling. arXiv preprint, arXiv:190606622. https://doi.org/10.48550/arXiv.1906.06622

Bolton T, Zanna L. 2019. Applications of deep learning to ocean data inference and subgrid parameterization. J Adv Model Earth Syst, 11: 376–399

Bonavita M, Geer A, Laloyaux P, Massart S, Chrust M. 2021. Data assimilation or machine learning? ECMWF Newsletter, No. 167

Bonavita M, Laloyaux P. 2020. Machine learning for model error inference and correction. J Adv Model Earth Syst, 12: e2020MS002232

Brenowitz N D, Bretherton C S. 2018. Prognostic validation of a neural network unified physics parameterization. Geophys Res Lett, 45: 6289–6298

Campos-Taberner M, García-Haro F J, Camps-Valls G, Grau-Muedra G, Nutini F, Crema A, Boschetti M. 2016. Multitemporal and multi-resolution leaf area index retrieval for operational local rice crop monitoring. Remote Sens Environ, 187: 102–118

Camps-Valls G, Martino L, Svendsen D H, Campos-Taberner M, Muñoz-Marí J, Laparra V, Luengo D, García-Haro F J. 2018. Physics-aware Gaussian processes in remote sensing. Appl Soft Comput, 68: 69–82

Cannon A J. 2011. Quantile regression neural networks: Implementation in R and application to precipitation downscaling. Comput Geoscis, 37: 1277–1284

Chantry M, Christensen H, Dueben P, Palmer T. 2021. Opportunities and challenges for machine learning in weather and climate modelling: Hard, medium and soft AI. Phil Trans R Soc A, 379: 20200083

Chen C, Zhang Z, Lin H. 2005. Earth simulator and simulation research progress (in Chinese). Adv Earth Sci, 20: 1135–1142

Chen F, Fu B, Xia J, Wu D, Wu S, Zhang Y, Sun H, Liu Y, Fang X, Qin B, Li X, Zhang T, Liu B, Dong Z, Hou S, Tian L, Xu B, Dong G, Zheng J, Yang W, Wang X, Li Z, Wang F, Hu Z, Wang J, Liu J, Chen J, Huang W, Hou J, Cai Q, Long H, Jiang M, Hu Y, Feng X, Mo X, Yang X, Zhang D, Wang X, Yin Y, Liu X. 2019. Major advances in studies of the physical geography and living environment of China during the past 70 years and future prospects. Sci China Earth Sci, 62: 1665–1701

Chen J, Liu W, Wu H, Li S, Yan L. 2021. Smart surveying and mapping: Fundamental issues and research agenda (in Chinese). Acta Geodaet Cartogr Sin, 50: 995–1005

Chen M, Lv G, Zhou C, Lin H, Ma Z, Yue S, Wen Y, Zhang F, Wang J, Zhu Z, Xu K, He Y. 2021. 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

Cheng C, Shi P, Song C, Gao J. 2018. Geographic big-data: A new opportunity for geography complexity study (in Chinese). Acta Geogr Sin, 73: 1397–1406

Chevallier F, Chéruy F, Scott N A, Chédin A. 1999. A neural network approach for a fast and accurate computation of a longwave radiative budget. J Appl Meteorol, 37: 1385–1397

Cintra R, de Campos Velho H, Cocke S. 2016. Tracking the model: Data assimilation by artificial neural network. Vancouver: Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN). 403–410

Davis D T, Zhengxiao Chen D T, Jenq-Neng Hwang D T, Tsang L, Njoku E. 1995. Solving inverse problems by Bayesian iterative inversion of a forward model with applications to parameter mapping using SMMR remote sensing data. IEEE Trans Geosci Remote Sens, 33: 1182–1193

Davis D T, Chen Z, Tsang L, Hwang J N, Chang A T C. 1993. Retrieval of snow parameters by iterative inversion of a neural network. IEEE Trans Geosci Remote Sens, 31: 842–852

Daw A, Thomas R Q, Carey C C, Read J S, Appling A P, Karpatne A. 2020. Physics-guided architecture (PGA) of neural networks for quantifying uncertainty in lake temperature modeling. In: Proceedings of the 2020 SIAM International Conference on Data Mining. 532–540

De Bézenac E, Pajot A, Gallinari P. 2019. Deep learning for physical processes: Incorporating prior scientific knowledge. J Stat Mech-Theory Exp, 2019: 124009
Deng M, Cai J, Yang W, Tang J, Yang X, Liu Q, Shi Y. 2020. Spatio-temporal analysis methods for multi-modal geographic big data (in Chinese). J Geo-inform Sci, 22: 41–56
Dian R, Li S, Kang X. 2021. Regularizing hyperspectral and multispectral image fusion by CNN denoiser. IEEE Trans Neural Netw Learn Syst, 32: 1124–1135
Du P, Samat A. 2013. Multiple instance ensemble learning method for high-resolution remote sensing image classification (in Chinese). J Remote Sens, 17: 77–97
Dueben P D, Bauer P. 2018. Challenges and design choices for global weather and climate models based on machine learning. Geosci Model Dev, 11: 3999–4009
Erichson N B, Muehlebach M, Mahoney M W. 2019. Physics-informed autoencoders for Lyapunov-stable fluid flow prediction. arXiv preprint, arXiv:190510866. https://doi.org/10.48550/arXiv.1905.10866
Farchi A, Laloyaux P, Bonavita M, Bocquet M. 2021. Using machine learning to correct model error in data assimilation and forecast applications. Q J R Meteorol Soc, 147: 3067–3084
Ganguly A R, Kodra E A, Agrawal A, Banerjee A, Boriah S, Chatterjee S, Chatterjee S, Choudhary A, Das D, Faghmous J, Ganguli P, Ghosh S, Hayhoe K, Hays C, Hendrix W, Fu Q, Kawale J, Kumar D, Kumar V, Liao W, Liess S, Mawalagedara R, Mithal V, Oglesby R, Salvi K, Snyder P K, Steinhaeuser K, Wang D, Wuebbles D. 2014. Toward enhanced understanding and projections of climate extremes using physics-guided data mining techniques. Nonlin Processes Geophys, 21: 777–795
Ghosh S. 2010. SVM-PGSL coupled approach for statistical downscaling to predict rainfall from GCM output. J Geophys Res, 115: D22102
Gilbert R C, Richman M B, Trafalis T B, Leslie L M. 2010. Machine learning methods for data assimilation. Comput Intell Architect Complex Eng Syst. New York: ASME Press. 105–112
Gong P. 2009. Some Frontier Issues in Remote Sensing Science and Technology (in Chinese). J Remote Sens, 13: 13–23
Guo H, Wang L, Chen F, Liang D. 2014. Scientific big data and digital Earth. Chin Sci Bull, 59: 1047–1054
Guo Q, Jin S, Li M, Yang Q, Xu K, Ju Y, Zhang J, Xuan J, Liu J, Su Y, Xu Q, Liu Y. 2020. Application of deep learning in ecological resource research: Theories, methods, and challenges. Sci China Earth Sci, 63: 1457–1474
Guo R, Lin H, He B, Zhao Z. 2020. GIS framework for smart cities. Geomat Inform Sci Wuhan Univ, 45: 1829–1835
Han J, Jentzen A, E W. 2018. Solving high-dimensional partial differential equations using deep learning. Proc Natl Acad Sci USA, 115: 8505–8510
Härter F P, de Campos Velho H F. 2008. New approach to applying neural network in nonlinear dynamic model. Appl Math Model, 32: 2621–2633
Härter F P, de Campos Velho H F. 2010. Multilayer perceptron neural network in a data assimilation scenario. Eng Appl Comput Fluid Mech, 4: 237–245
He K M, Zhang X Y, Ren S Q, Sun J. 2016. Deep residual learning for image recognition. Seattle, WA: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 770–778
Hinton G E, Osindero S, Teh Y W. 2006. A fast learning algorithm for deep belief nets. Neural Computation, 18: 1527–1554
Hsieh W W, Tang B. 1998. Applying neural network models to prediction and data analysis in meteorology and oceanography. Bull Amer Meteorol Soc, 79: 1855–1870
Huang G, Liu Z, Van Der Maaten L, Weinberger K Q. 2017. Densely connected convolutional networks. Honolulu: 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2261–2269
Huang X, Li J, Yang J, Zhang Z, Li D, Liu X. 2021. 30 m global impervious surface area dynamics and urban expansion pattern observed by Landsat satellites: From 1972 to 2019. Sci China Earth Sci, 64: 1922–1933
Hunter J M, Maier H R, Gibbs M S, Foale E R, Grosvenor N A, Harders N P, Kikuchi-Miller T C. 2018. 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
Ivatt P D, Evans M J. 2020. Improving the prediction of an atmospheric chemistry transport model using gradient-boosted regression trees. Atmos Chem Phys, 20: 8063–8082
Jia X, Willard J, Karpatne A, Read J S, Zwart J A, Steinbach M, Kumar V. 2021. Physics-guided machine learning for scientific discovery: An application in simulating lake temperature profiles. ACM IMS Trans Data Sci, 2: 1–26
Karpatne A, Atluri G, Faghmous J H, Steinbach M, Banerjee A, Ganguly A, Shekhar S, Samatova N, Kumar V. 2017a. Theory-guided data science: A new paradigm for scientific discovery from data. IEEE Trans Knowl Data Eng, 29: 2318–2331
Karpatne A, Ebert-Uphoff I, Ravela S, Babaie H A, Kumar V. 2019. Machine learning for the geosciences: Challenges and opportunities. IEEE Trans Knowl Data Eng, 31: 1544–1554
Karpatne A, Watkins W, Read J, Kumar V. 2017b. Physics-guided neural networks (PGNN): An application in lake temperature modeling. arXiv preprint, arXiv:171011431. https://doi.org/10.48550/arXiv.1710.11431
Kashinath K, Mustafa M, Albert A, Wu J L, Jiang C, Esmaeilzadeh S, Azizzadenesheli K, Wang R, Chattopadhyay A, Singh A, Manepalli A, Chirila D, Yu R, Walters R, White B, Xiao H, Tchelepi H A, Marcus P, Anandkumar A, Hassanzadeh P, Prabhat P. 2021. Physics-informed machine learning: Case studies for weather and climate modelling. Phil Trans R Soc A, 379: 20200093
Keller C A, Evans M J. 2019. Application of random forest regression to the calculation of gas-phase chemistry within the GEOS-Chem chemistry model v10. Geosci Model Dev, 12: 1209–1225
Kraft B, Jung M, Körner M, Koirala S, Reichstein M. 2022. Towards hybrid modeling of the global hydrological cycle. Hydrol Earth Syst Sci, 26: 1579–1614
Krasnopolsky V. 2020. Using machine learning for model physics: An overview. arXiv preprint, arXiv.2002.00416. https://doi.org/10.48550/arXiv.2002.00416
Krasnopolsky V M, Fox-Rabinovitz M S, Belochitski A A. 2013. Using ensemble of neural networks to learn stochastic convection parameterizations for climate and numerical weather prediction models from data simulated by a cloud resolving model. Adv Artif Neural Syst, 2013: 1–13
Krasnopolsky V M, Fox-Rabinovitz M S, Chalikov D V. 2005. New approach to calculation of atmospheric model physics: Accurate and fast neural network emulation of longwave radiation in a climate model. Mon Weather Rev, 133: 1370–1383
Krasnopolsky V M, Lin Y. 2012. A neural network nonlinear multimodel ensemble to improve precipitation forecasts over continental US. Adv Meteorol, 2012: 1–11
Krasnopolsky V M, Lord S J, Moorthi S, Spindler T. 2009. How to deal with inhomogeneous outputs and high dimensionality of neural network emulations of model physics in numerical climate and weather prediction models. Atlanta: Proceedings of the International Joint Conference on Neural Networks. 1668–1673
Lazer D, Kennedy R, King G, Vespignani A. 2014. The parable of google flu: Traps in big data analysis. Science, 343: 1203–1205
Letu H, Shi J, Li M, Wang T, Shang H, Lei Y, Ji D, Wen J, Yang K, Chen L. 2020. A review of the estimation of downward surface shortwave radiation based on satellite data: Methods, progress and problems. Sci China Earth Sci, 63: 774–789
Li T, Shen H, Yuan Q, Zhang L. 2020. Geographically and temporally weighted neural networks for satellite-based mapping of ground-level $PM_{2.5}$. ISPRS J Photogrammetry Remote Sens, 167: 178–188
Li T, Shen H, Yuan Q, Zhang L. 2021. A locally weighted neural network constrained by global training for remote sensing estimation of $PM_{2.5}$. IEEE Trans Geosci Remote Sens, doi: 10.1109/TGRS.2021.3074569

Li T, Shen H, Yuan Q, Zhang X, Zhang L. 2017. Estimating ground-level $PM_{2.5}$ by fusing satellite and station observations: A geo-intelligent deep learning approach. Geophys Res Lett, 44: 11,985–11,993
Li W, Ni L, Li Z L, Duan S B, Wu H. 2019. Evaluation of machine learning algorithms in spatial downscaling of modis land surface temperature. IEEE J Sel Top Appl Earth Observ Remote Sens, 12: 2299–2307
Li X. 2005. Retrospect, prospect and innovation in quantitative remote sensing (in Chinese). J Henan Univ-Nat Sci: 49–56
Li X, Huang C, Che T, Jin R, Wang S, Wang J, Gao F, Zhang S, Qiu C, Wang C. 2007. Progress and prospects of land surface data assimilation system research in China (in Chinese). Prog Nat Sci, 17: 163–173
Li X, Liu F, Fang M. 2020. Harmonizing models and observations: Data assimilation in Earth system science. Sci China Earth Sci, 63: 1059–1068
Li X, Ma H, Ran Y, Wang X, Zhu G, Liu F, He H, Zhang Z, Huang C. 2021. Terrestrial carbon cycle model-data fusion: Progress and challenges. Sci China Earth Sci, 64: 1645–1657
Li X, Ye J. 2005. Cellular automata for simulating complex land use systems using neural networks (in Chinese). Geogr Res, 24: 19–27
Li X, Zheng D, Feng M, Chen F. 2022. Information geography: The information revolution reshapes geography. Sci China Earth Sci, 65: 379–382
Li Z, Duan S, Tang B, Wu H, Ren H, Yan G, Tang R, Leng P. 2016. Review of methods for land surface temperature derived from thermal infrared remotely sensed data (in Chinese). J Remote Sens, 20: 899–920
Liang S, Cheng J, Jia K, Jiang B, Liu Q, Liu S, Xiao Z, Xie X, Yao Y, Yuan W, Zhang X, Zhao X. 2016. Recent progress in land surface quantitative remote sensing (in Chinese). J Remote Sens, 20: 875–898
Liang Z, Zou R, Chen X, Ren T, Su H, Liu Y. 2020. Simulate the forecast capacity of a complicated water quality model using the long short-term memory approach. J Hydrol, 581: 124432
Lin L P, Li J, Shen H F, Zhao L L, Yuan Q Q, Li X H. 2022. Low-resolution fully polarimetric SAR and high-resolution single-polarization sar image fusion network. IEEE Trans Geosci Remote Sens, 60: 1–17
Lu J, Hu W, Zhang X. 2018. Precipitation data assimilation system based on a neural network and case-based reasoning system. Information, 9: 106
Mao K, Shi J, Li Z L, Tang H. 2007. An RM-NN algorithm for retrieving land surface temperature and emissivity from EOS/MODIS data. J Geophys Res, 112: D21102
McQuade S, Monteleoni C. 2012. Global climate model tracking using geospatial neighborhoods. AAAI, 26: 335–341
Meng C, Dai Y. 2013. Development and verification of a bulk urbanized land surface model (in Chinese). Chin J Atmos Sci, 37: 1297–1308
Monteleoni C, Schmidt G A, Saroha S, Asplund E. 2011. Tracking climate models. Statistical Analy Data Min, 4: 372–392
Navares R, Aznarte J L. 2020. Predicting air quality with deep learning LSTM: Towards comprehensive models. Ecol Inf, 55: 101019
Noori N, Kalin L, Isik S. 2020. Water quality prediction using SWAT-ANN coupled approach. J Hydrol, 590: 125220
Pei T, Liu Y, Guo S, Shu H, Du Y, Ma T, Zhou C. 2019. Principle of big geodata mining (in Chinese). Acta Geogr Sin, 74: 586–598
Petty T R, Dhingra P. 2018. Streamflow hydrology estimate using machine learning (SHEM). J Am Water Resour Assoc, 54: 55–68
Qiu C. 2021. China's first earth system simulation large-scale scientific device opened (in Chinese). China Youth Daily
Ran Y, Li X, Cheng G, Nan Z, Che J, Sheng Y, Wu Q, Jin H, Luo D, Tang Z, Wu X. 2021. Mapping the permafrost stability on the Tibetan Plateau for 2005–2015. Sci China Earth Sci, 64: 62–79
Rasp S, Lerch S. 2018. Neural networks for postprocessing ensemble weather forecasts. Mon Weather Rev, 146: 3885–3900
Read J S, Jia X, Willard J, Appling A P, Zwart J A, Oliver S K, Karpatne A, Hansen G J A, Hanson P C, Watkins W, Steinbach M, Kumar V. 2019. Process-guided deep learning predictions of lake water temperature. Water Resour Res, 55: 9173–9190
Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, Prabhat N. 2019. Deep learning and process understanding for data-driven Earth system science. Nature, 566: 195–204
Research Group of Geoscience Development Strategy, Department of Geosciences, Chinese Academy of Sciences. 2009. Report on China's Geoscience Development Strategy in the 21st Century. Beijing: Science Press
von Rueden L, Mayer S, Beckh K, Georgiev B, Giesselbach S, Heese R, Kirsch B, Walczak M, Pfrommer J, Pick A, Ramamurthy R, Garcke J, Bauckhage C, Schuecker J. 2023. Informed machine learning—A taxonomy and survey of integrating prior knowledge into learning systems. IEEE Trans Knowledge Data Eng, 35: 614–633
von Rueden L, Mayer S, Sifa R, Bauckhage C, Garcke J. 2020. Combining machine learning and simulation to a hybrid modelling approach: Current and future directions. In: Berthold M, Feelders A, Krempl G, eds. Advances in Intelligent Data Analysis XVIII. Cham: Springer International Publishing. 548–560
Sargsyan K, Safta C, Najm H N, Debusschere B J, Ricciuto D, Thornton P. 2014. Dimensionality reduction for complex models via Bayesian compressive sensing. Int J Uncertain Quant, 4: 63–93
Sawada Y. 2020. Machine learning accelerates parameter optimization and uncertainty assessment of a land surface model. J Geophys Res-Atmos, 125: e2020JD032688
Scher S, Messori G. 2019. Weather and climate forecasting with neural networks: Using general circulation models (GCMs) with different complexity as a study ground. Geosci Model Dev, 12: 2797–2809
Schneider T, Lan S, Stuart A, Teixeira J. 2017. Earth system modeling 2.0: A blueprint for models that learn from observations and targeted high-resolution simulations. Geophys Res Lett, 44: 12,396–12,417
Shen H, Jiang M, Li J, Zhou C, Yuan Q, Zhang L. 2022. Coupling model- and data-driven methods for remote sensing image restoration and fusion: Improving physical interpretability. IEEE Geosci Remote Sens Mag, 10: 231–249
Shen H, Jiang Y, Li T, Cheng Q, Zeng C, Zhang L. 2020. Deep learning-based air temperature mapping by fusing remote sensing, station, simulation and socioeconomic data. Remote Sens Environ, 240: 111692
Shen H, Li T, Yuan Q, Zhang L. 2018. Estimating regional ground-level $PM_{2.5}$ directly from satellite top-of-atmosphere reflectance using deep belief networks. J Geophys Res-Atmos, 123: 13,875–13,886
Skamarock W, Klemp J, Dudhia J, Gill D, Barker D, Wang W, Powers J. 2005. A Description of the Advanced Research WRF Version 2. Technical Report. Report No. NCAR/TN 468+STR
Sønderby C K, Espeholt L, Heek J, Dehghani M, Oliver A, Salimans T, Agrawal S, Hickey J, Kalchbrenner N. 2020. Metnet: A neural weather model for precipitation forecasting. arXiv preprint, arXiv:200312140. https://doi.org/10.48550/arXiv.2003.12140
Song C. 2016. On paradigms of geographical research (in Chinese). Prog Geogr, 35: 1–3
Stensrud D J. 2007. Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models. Cambridge: Cambridge University Press. 449
Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. 2015. Going Deeper with Convolutions. Boston: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 1–9
Trombetti M, Riaño D, Rubio M A, Cheng Y B, Ustin S L. 2008. Multi-temporal vegetation canopy water content retrieval and interpretation using artificial neural networks for the continental USA. Remote Sens Environ, 112: 203–215
Vandal T, Kodra E, Ganguly S, Michaelis A, Nemani R, Ganguly A R. 2017. DeepSD: Generating high resolution climate change projections through single image super-resolution. Halifax: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Part F129685: 1663–1672
Venkatakrishnan S V, Bouman C A, Wohlberg B. 2013. Plug-and-play priors for model based reconstruction. Austin: 2013 IEEE Global Conference on Signal and Information Processing. 945–948
Wang F, Tian D, Lowe L, Kalin L, Lehrter J. 2021. Deep learning for daily precipitation and temperature downscaling. Water Res, 57:

e2020WR029308
Wilby R L, Wigley T M L, Conway D, Jones P D, Hewitson B C, Main J, Wilks D S. 1998. Statistical downscaling of general circulation model output: A comparison of methods. Water Resour Res, 34: 2995–3008
Willard J, Jia X, Xu S, Steinbach M, Kumar V. 2020. Integrating physics-based modeling with machine learning: A survey. arXiv preprint, arXiv:200304919. https://doi.org/10.48550/arXiv.2003.04919
Witt C, Tong C, Zantedeschi V, Martini D, Kalaitzis F, Chantry M, Watson-Parris D, Bilinski P. 2020. RainBench: Towards global precipitation forecasting from satellite imagery. 35th AAAI Conference on Artificial Intelligence, AAAI 2021. 17A: 14902–14910
Wolanin A, Camps-Valls G, Gómez-Chova L, Mateo-García G, van der Tol C, Zhang Y, Guanter L. 2019. Estimating crop primary productivity with Sentinel-2 and Landsat 8 using machine learning methods trained with radiative transfer simulations. Remote Sens Environ, 225: 441–457
Wu Z, Chai Y, Dang A, Gong J, Gao S, Yue Y, Li D, Liu L, Liu X, Liu Y, Long Y, Lu F, Qin C, Wang H, Wang P, Wang W, Zhen F. 2015. Geography interact with big data: Dialogue and reflection (in Chinese). Geogr Res, 34: 2207–2221
Xiao Q, Wang Y, Chang H H, Meng X, Geng G, Lyapustin A, Liu Y. 2017. Full-coverage high-resolution daily $PM_{2.5}$ estimation using MAIAC AOD in the Yangtze River Delta of China. Remote Sens Environ, 199: 437–446
Yuan Q, Shen H, Li T, Li Z, Li S, Jiang Y, Xu H, Tan W, Yang Q, Wang J, Gao J, Zhang L. 2020. Deep learning in environmental remote sensing: Achievements and challenges. Remote Sens Environ, 241: 111716
Zhang B. 2018. Remotely sensed big data era and intelligent information extraction (in Chinese). Geomat Inform Sci Wuhan Univ, 43: 1861–1871
Zhang L, Shen H. 2016. Progress and future of remote sensing data fusion (in Chinese). J Remote Sens, 20: 1050–1061
Zhang T, Huang C, Shen H. 2012. Sensitivity and parameters optimization method of soil parameters to soil moisture in common land model (in Chinese). Adv Earth Sci, 27: 678–685
Zhang Z, Tang P, Li H Y, Feng Z. 2016. Refined domain model for multisource data synergized quantitative remote sensing production system (in Chinese). J Remote Sens, 20: 184–196
Zhou C, Wang H, Wang C, Hou Z, Zheng Z, Shen S, Cheng Q, Feng Z, Wang X, Lv H, Fan J, Hu X, Hou M, Zhu Y. 2021. Geoscience knowledge graph in the big data era. Sci China Earth Sci, 64: 1105–1114

(Responsible editor: Xin LI)