Future Generation Computer Systems 112 (2020) 449–466

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Participatory intercomparison strategy for terrestrial carbon cycle models based on a service-oriented architecture

Songshan Yue [a,b,c], Min Chen [a,b,c,*], Jie Song [a,b,c], Wenping Yuan [d], Tiexi Chen [e], Guonian Lü [a,b,c], Chaoran Shen [a,b,c], Zaiyang Ma [a,b,c], Kai Xu [a,b,c], Yongning Wen [a,b,c], Hongquan Song [f]

[a] *Key Laboratory of Virtual Geographic Environment (Nanjing Normal University), Ministry of Education, Nanjing, 210023, China*
[b] *State Key Laboratory Cultivation Base of Geographical Environment Evolution (Jiangsu Province), Nanjing Normal University, Nanjing, 210023, China*
[c] *Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing Normal University, Nanjing, 210023, China*
[d] *Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, School of Atmospheric Sciences, Sun Yat-sen University, Guangdong, Zhuhai 519082, China*
[e] *Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disaster, School of Geographical Sciences, Nanjing University of Information Science & Technology, Nanjing 210044, China*
[f] *Key Laboratory of Geospatial Technology for the Middle and Lower Yellow River Regions (Henan University), Ministry of Education, Kaifeng, 475004, China*

ARTICLE INFO

Article history:
Received 2 February 2020
Received in revised form 14 April 2020
Accepted 26 May 2020
Available online 28 May 2020

Keywords:
Participatory modeling
Model-data intercomparison project
Reusability and reproducibility
Model service
Terrestrial carbon cycle

ABSTRACT

Terrestrial carbon cycle models are important tools for simulating carbon exchange; however, there are still significant uncertainties in the simulation results of different models. Model-data intercomparison has therefore been widely recognized as an effective approach for evaluating model performance and acquiring a more reliable understanding of the terrestrial carbon cycle. Although considerable efforts have been made in establishing model intercomparison projects (MIPs), existing MIPs still experience limitations in supporting teams of researchers working collaboratively online and ensuring the reproducibility of model experiments. This article proposes a participatory intercomparison strategy based on a service-oriented architecture (SOA), which aims to offer a web-based platform for researchers to construct participatory intercomparison (PIC) projects. The three fundamental components of a PIC project are the PIC topic, PIC instance, and PIC task. The PIC topic is used to help participants co-design backgrounds, goals, and comparison protocols. The PIC instance is used to help participants provide models, observations and benchmark data as reusable services. The PIC task is used to help participants formulate comparison workflows and acquire customized comparison results. Using the proposed strategy, a PIC project can be easily created and maintained by a group of geographically distributed participants. Reusability of models and data can be achieved through the proposed service-based wrapping method, and the reproducibility of model-data comparison experiments can be achieved through the workflow-based comparison method. A web platform named "P-MIP" was implemented and tested that helps researchers collaborate online, and a demonstrative PIC project was constructed to verify the feasibility and capability of the proposed strategy.

1. Introduction

The carbon emissions generated by human activities exceed the adjustment capacity of Earth itself; therefore, atmospheric CO_2, which is a primary greenhouse gas, concentration has increased and caused global warming [1–5]. To capture and understand the factors that affect changes in the terrestrial biosphere, an increasing number of studies have focused on the terrestrial carbon cycle [6–12]. The terrestrial carbon cycle indicates the migration path of carbon, which is generally related to the atmosphere, plant photosynthetic organs, plant sustaining organs, litterfall, soil and so on [6,13]. This cycle is extremely complex and involves many physical, chemical and biological processes and is also affected by a range of different environmental parameters, such as temperature, precipitation, radiation, atmospheric CO_2 concentration, land use and land cover change (LUCC) [14–18].

Due to the complexity of the terrestrial carbon cycle, understanding and simulating such a cycle requires knowledge of

* Corresponding author.
E-mail address: chenmin0902@163.com (M. Chen).

https://doi.org/10.1016/j.future.2020.05.044
0167-739X/

Earth as a system [19]. Constructing quantitative models has been widely recognized as an effective approach. Since Craig [20] used two carbon pools to simulate the carbon balance of terrestrial ecosystems in the global carbon cycle model in the 1950s, several terrestrial carbon models that improved our understanding of the territorial carbon cycle have been developed and studied. Terrestrial biogeochemical models (TBMs) and dynamic global vegetation models (DGVMs) are useful tools for simulating carbon fluxes at local, regional and global scales [21–31]. Although these models included several detailed carbon exchange processes, such as photosynthesis, autotrophic respiration, heterotrophic respiration, soil respiration, and vegetation distribution, the simulated carbon fluxes remain uncertain [32–34]. Therefore, model-data intercomparison has been recognized as an effective approach to acquire a more reliable understanding of the terrestrial carbon cycle [35–37].

Focusing on different topics, regions and scales, a range of model-data intercomparison projects (MIPs) have been established [38–40], through which the performances of different models can be evaluated. The comparison results can also foster improvements in terrestrial carbon cycle models. MIPs generally compare gross primary productivity (GPP), net primary productivity (NPP) and net ecosystem productivity (NEP); however, current MIPs still have limitations, especially regarding their support for researchers working collaboratively online and their ability to ensure the reproducibility of model experiments. Online collaboration is important for collecting resources and knowledge from researchers distributed worldwide, while reproducibility is essential to ensure the fairness and openness of comparison results [41,42]. To obtain more reliable simulation results, models are continually improved, and more accurate *in situ* measurements are increasingly explored. Therefore, model-data intercomparison needs to be considered as an ongoing and progressive activity, which requires that the legacy model experiments be reproducible and that individual researchers' comparison activities are able to be integrated easily for synthetic analyses. In this context, this paper proposes a participatory intercomparison strategy for terrestrial carbon cycle models. Using this strategy, a variety of participatory intercomparison (PIC) projects can be constructed not only by research organizations but also by participants in an open web environment. Researchers can customize PIC projects, and intercomparison activities can be conducted in an open, configurable and reproducible manner.

The remainder of this article is structured as follows. In Section 2, the motivation and basic framework of the proposed strategy are explained. The detailed methods to construct a PIC project are presented in Section 3. Section 4 introduces the implementation of a prototype system and related experiments. Finally, the discussion and conclusions of the study are presented in Section 5.

2. Basic design of the PIC strategy

Existing MIPs (see Supplementary A) have contributed to an improved understanding of the terrestrial carbon cycle and have supported refinements to individual models. Despite having a variety of goals, all these MIPs are formulated and conducted by certain organizations or groups. This form of project maintenance is designed to satisfy the demands for comparing model data within a standard framework (i.e., being fair to each involved model). However, this practice also leads to the following limitations:

(1) **Research barriers**. We have observed that past MIP hosts usually hold meetings or workshops to discuss the intercomparison protocol or publish intercomparison reports; this requires gathering various researchers in one location; consequently, financial support or government funding are required. The costs of project maintenance leads to practical barriers for individual researchers who want to build a new MIP. In addition, after an MIP is established, other modelers have fewer opportunities to contribute to new models and conduct further comparison tasks (which is why some MIPs are designed in several phases—so that the design can include additional models and researchers). These barriers emerge not only among researchers but also among different MIPs. Because the existing MIPs are managed in a centralized manner, it is difficult to integrate different MIPs to acquire more synthetic intercomparison results.

(2) **Reusability and reproducibility**. It is essential to set up an intercomparison topic for a MIP around which the observations can be collected and model simulation experiments can be executed. Focusing on a predefined topic, existing MIPs typically contribute their model-data comparison results as reports or articles (supplemented by figures or tables). In this way, model results and comparison results are difficult to reuse: reports and articles provide less original information, making it more difficult to apply historical comparison results to other model-data comparison work. Although some MIPs have made the driver data (i.e., the datasets input to a model) and the model results downloadable, the parameter configuration settings for the involved models were often not available. This lack of sufficient model configuration information leads to low reproducibility, meaning that a model simulation/experiment with some minor parameter modifications must be reconstructed from the very beginning.

Overall, the current MIP construction approach can be considered as a centralized strategy. To overcome the high research barriers and low reusability and reproducibility obstacles, the proposed participatory intercomparison strategy is designed based on a service-oriented architecture (SOA), through which PIC projects can be collaboratively constructed in an open web environment. Within a PIC project, various participants can contribute models and data as reusable web services, design model experiments as reproducible web applications and conduct model-data comparisons as customizable workflows. Members of the general public can also join a PIC project and provide resources or simply observe how the comparison process is conducted. This inspired the name "participatory intercomparison".

Fig. 1 shows the basic framework of the PIC strategy. Based on the SOA, a PIC project is constructed via collaboration. The dependent resources (i.e., models, data, data processing methods, and comparison methods) for obtaining intercomparison results are contributed by different participants in the web environment. Original models and data are wrapped and organized as model services and data services, respectively. Model services are presented as executable web applications that allow users to modify parameters to obtain customized results. Data services are data accessible through web request commands. In addition, data processing methods (e.g., converting data format, extracting data within a specific time range, and merging data in different regions) and comparison methods (e.g., statistical comparison, time series comparison, and spatial distribution comparison) are reusable algorithms that can be executed online.

Supported by these resources, a PIC project is constructed with one PIC topic, multiple PIC instances, and multiple PIC tasks. Within the PIC topic, backgrounds, goals and comparison protocols can be coedited by all participants. After the PIC topic is recognized by all participants, different PIC instances can be created. A PIC instance is a comparable element that can exist in three types: observation, benchmark and model experiment outputs. As all PIC instances are contributed by participants, they can also be reused when constructing a new PIC project. Different participants can create customized PIC tasks for a PIC project by selecting PIC instances. Within a PIC task, the detailed metrics (e.g., GPP, NPP, and NEP), data (referred to as PIC instances), data processing methods (e.g., clipping the data in PIC

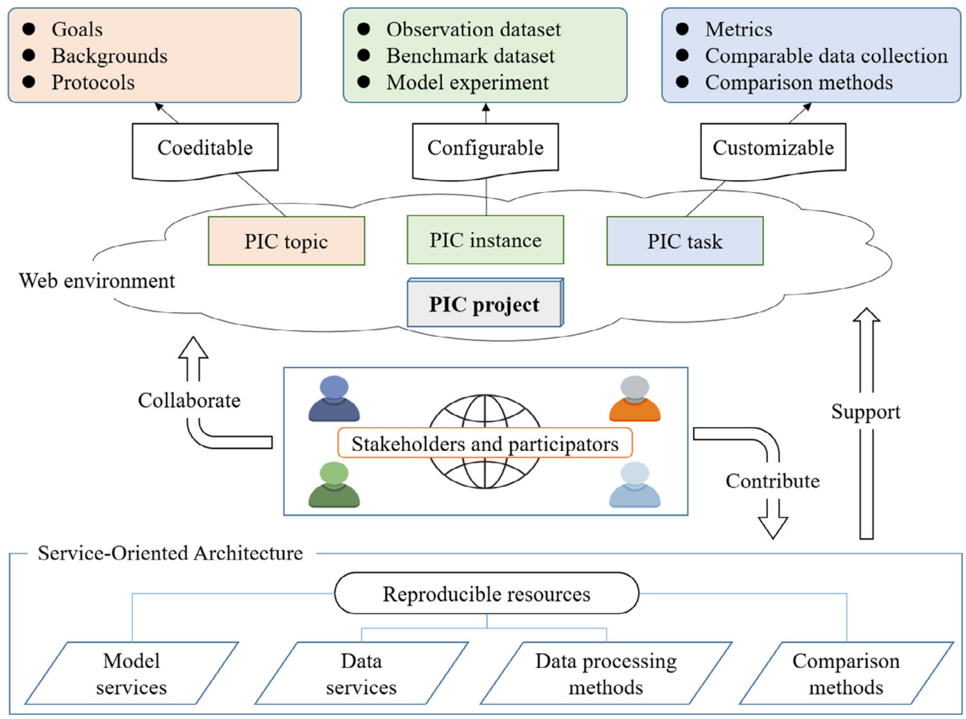


Fig. 1. Basic framework of the PIC strategy.

instances with a certain spatial extent), and comparison methods (e.g., comparing the spatial distribution of different models' output NPPs through maps) are combined into a workflow to obtain the intercomparison results. Such a workflow can be applied to create new PIC tasks with customized modifications. With the help of web communication, the proposed PIC topic, PIC instance and PIC task can reduce the barriers in constructing new MIPs and improve the reusability and reproducibility of model-data comparison experiments.

3. PIC project

3.1. PIC topic

To initialize a PIC project, it is necessary to identify the motivation and objectives since they present the information that attracts different researchers to join and become participants. Two corresponding parts, backgrounds and goals, are therefore designed in a PIC topic. When establishing a PIC project, the backgrounds and goals are first defined by the creator, and then all the participants can propose ideas to update them. After confirming the backgrounds and goals, the intercomparison protocols for a PIC project need to be formulated.

To ensure the consistency and fairness of model-data comparison, many conditions should be considered when drafting protocols, which include but are not limited to the standard environmental driver data, consistent biome classifications, reasonable procedures of model simulation and expected output variables of potential participating models. Traditionally, the protocols of an MIP are organized in documents, and detailed agreements are given as descriptive texts. Protocol documents in existing MIPs are diverse in both organization and expression aspects. Therefore, it is not easy to integrate one MIP protocol with another. To help participants formulate protocols for different PIC projects in a consistent manner, a structural protocol framework is proposed in this article.

As shown in Fig. 2, the protocol framework is designed with two layers: an information layer and a convention layer. The

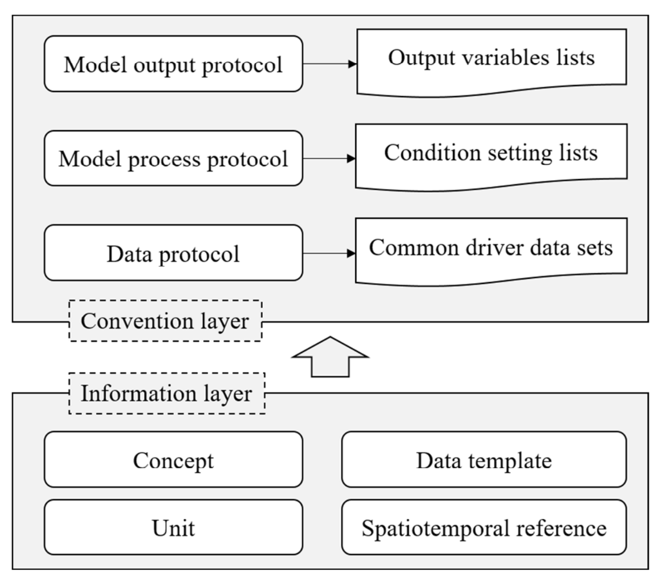


Fig. 2. Protocol framework for a PIC topic.

information layer mainly focuses on providing unambiguous descriptions for protocols via participants' collaboration, while the convention layer mainly focuses on helping participants discuss and formulate protocols. There are four repositories in the information layer:

(1) Concept repository. This repository is a vocabulary repository that explains the words and phrases presented in the detailed protocol items. For example, a low-latitude area might be defined as "generally, a low-latitude area refers to the spatial extent of the earth's surface from the equator to 30 degrees north–south latitude. However, in this project, a low-latitude area refers to the extent from the equator to 20 degrees north–south latitude".

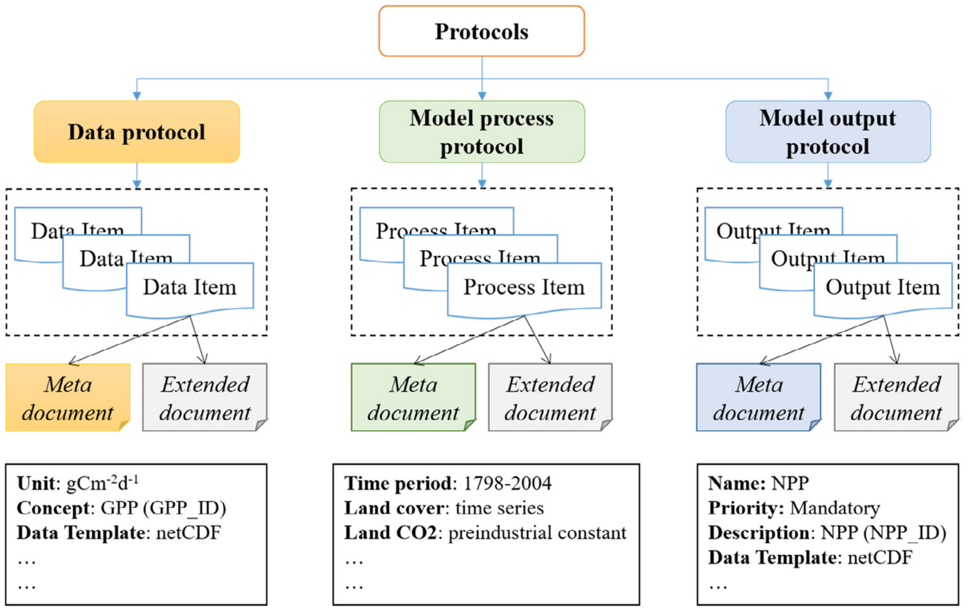


Fig. 3. Structure of the data, process and output protocols.

(2) Unit repository. This repository stores units that can be used to describe variables. For example, the unit Kelvin is described as "The reference point that defines the Kelvin scale is the triple point of water at 273.16 K (0.01 °C; 32.02 °F). Kelvin is defined as 1/273.16 of the difference between these two reference points". The symbol for Kelvin is also given as K.

(3) Data template repository. This repository is designed to hold data format explanations that are used for driver data or to explain model outputs with detailed protocol information. For example, the netCDF format is described as "a data format that is short for the network common data form, which contains dimensions, variables, attributes and data contents". The file suffix ".nc" is also given.

(4) Spatiotemporal reference repository. This repository is designed to hold descriptions such as the spatial scale, spatial resolution, spatial coordinate system, temporal range, temporal resolution, and temporal system. For example, the spatial coordinate system WGS84 is defined as "short for World Geodetic System 1984, where coordinates are represented in degrees".

All the descriptive information in the above four repositories is organized into lists of entries. Based on the information layer, the convention layer is designed with three types of protocols, as shown in Fig. 3.

(1) Data protocol. This protocol provides datasets for participants and supports execution of different models with consistent environmental driver data. In the data protocols, each dataset is organized as a data item that contains data file(s) and two auxiliary documents: a meta-document and an extended document. The meta-document consists of a key–value list, such as "unit: g cm^{-2} d^{-1}", "concept: GPP", and "data template: netCDF". The extended document contains more descriptive information about the dataset.

(2) Model process protocol. This protocol mainly focuses on the condition settings for the model simulation process. A model process protocol includes one or more process items, each of which is also constructed with the meta-document and extended document. The meta-document contains the key–value (condition name: option) list. For example, a process item can be the combination of "time period: 1798–2004", "land cover: time series", and "land CO2: preindustrial constant".

(3) Model output protocol. This protocol is designed to indicate which variables models need to generate. A model output protocol includes several output items. Similar to a data item and process item, an output item is also constructed with a meta-document and extended document. The key–value structure still applies to the meta-document of an output item; however, the variable's name must be included in the key–value list. For example, an output item can be the combination of "name: NPP", "unit: g cm^{-2} d^{-1}", "priority: mandatory", "description: NPP_ID in the concept repository" and "data template: netCDF".

The meta-documents for the above three types of protocol items are all constructed based on the same key–value structure, allowing the protocol information to be automatically displayed as tables in web pages to help various participants understand them and contribute new protocols. The extended documents are used to help participants express more detailed ideas and explanations about the proposed protocols. In addition, when establishing a new PIC project, these data items, process items and output items can also be reused during assemblies to form new protocols. Based on this consistent structure, the protocols of different PIC projects can be easily and clearly compared to help researchers understand them.

3.2. PIC instance

PIC instances are the fundamental comparable elements for model performance evaluation. According to the existing MIPs, model performances can be evaluated in the following ways: intermodel comparison, model-observation comparison and model-benchmark comparison. For intermodel comparison, the output variables of different models are collected and compared in a many-to-many manner; for model-observation comparison, the model output variables are compared with the observed data in a many-to-one manner; and for model-benchmark comparison, the model output variables are compared with benchmark data in a many-to-one manner (the benchmark data are generated based on observation datasets or on a certain model's outputs). Consequently, there are three types of PIC instances:

(1) PIC instances of the observation type, which are created by uploading associated observation datasets. Observation datasets are generated through various information acquisition devices; for example, a daily temperature dataset is an observation dataset collected from a range of weather stations.

(2) PIC instances of the benchmark type, which are created by uploading associated benchmark datasets. Benchmark datasets contain widely recognized results that can act as references for model performance evaluations and are normally acquired by applying certain data processing methods in observation datasets or model results. For example, original temperature data (at different points) can be processed into grid data using interpolation methods; then, the grid data constitute a benchmark dataset.

(3) PIC instances of the model type, are created by uploading the associated model and configuring the parameters needed to execute the model. Such instances are executable models, and the associated model results can be compared with observation data, benchmark data and the results of other models.

As shown in Fig. 4, when participants upload observation or benchmark datasets to create a PIC instance, the meta-documents and extended documents (referring to the protocols in the PIC topic) should also be provided so that other participants can understand the offered data. The data files and any related meta-documents and extended documents are combined into a data package and stored in a data repository. Each data package is assigned a unique identifier (ID), and the corresponding data services are automatically generated and published in the PIC instance. Subsequently, the detailed data contents and descriptive information can be accessed through the representational state transfer (REST) application programming interface (API) in a web environment. Specifically, a meta-document must contain the name, unit, and data template information so that other participants can understand the data.

A PIC instance of the model type consists primarily of the uploaded model and the configuration document used to execute the model. As shown in Fig. 4, the associated drivers, parameters and outputs of the model are organized in the configuration document. With the help of the configuration documents, a model can be wrapped as a reusable model service, and the output data are stored and published as accessible data services in the web environment. Through the REST API, participants can customize the drivers and parameters in the configuration document, and the corresponding output data can be compared with other PIC instances. The configuration document is key to conducting model experiments in a web environment. Considering that different models have diverse execution processes and input/output (I/O) actions, a state-event-based method is employed for wrapping models as standard model services and organizing the configuration documents [43,44].

Fig. 5(a) presents the state-event-based method. Using this method, the model execution process is organized as a flow of switching states, where each state indicates an execution step. The execution process starts with the *Initialize* state and terminates at the *Finalize* state. Between *Initialize* and *Finalize*, a sequence of states composes the entire execution process. *EnterState* and *LeaveState* indicate the beginning and end of a step, respectively. Between *EnterState* and *LeaveState*, a range of events may be triggered that correspond to I/O actions. There are two types of events: *RequestData* (for model input) and *ResponseData* (for model output). Based on this wrapping method, different models can be wrapped as model services that have consistent APIs. According to the state-event structure, the APIs include *OnInitialize*, *OnEnterState*, *OnLeaveState*, *OnRequestData*, *OnResponseData*, and *OnFinalize*. To provide a model service's API information for model users, extensible markup language (XML) is employed to construct the configuration document (as shown in Fig. 5(b)).

In the configuration document, the state structure should be consistent with the model process protocols, events of the input type should be consistent with the data protocols, and events of the output type should be consistent with the model output protocols. To conduct a model experiment (i.e., execute the model), the configuration document is used as the main user interface: (1) data and parameters are assigned to corresponding *event* nodes of the "request" type; (2) after a model service starts, the required data and parameters can be requested through the *event* nodes (where the "name" attribute serves as the unique identifier); and (3) when intermediate results or final results are calculated, the corresponding *event* nodes of the "response" type are filled with the generated file paths or values. These results can be reused by different PIC tasks, enabling comparisons with other PIC instance outputs.

A web application with a graphical user interface (GUI) is supplied to help participants edit the configuration document. Customized model results can be obtained by modifying the data and parameter configurations. Based on the web application, a PIC instance can be easily duplicated to create a new instance, and the new PIC instance can also be reconfigured. Therefore, model experiments can be conducted in an open and reproducible manner.

3.3. PIC task

In a PIC project, based on the designed protocols, various PIC tasks can be created and executed. Because the model outputs include multiple variables, the metrics to be compared (e.g., GPP, NEP, NPP, and leaf area index (LAI)) must be defined first when creating a PIC task. Then, different PIC instances can be selected and imported to the PIC task. After determining the metrics and related PIC instances, detailed comparison logic should be designed. The model-data comparison work can be conducted by assessing various aspects, for example, comparing the time series of different model results at a specific site, comparing the spatial distribution of different model results, and comparing the Nash–Sutcliffe coefficient [45] of different model results. To satisfy these various demands, we propose a workflow-based method to help participants design the comparison logic of a PIC task.

As shown in Fig. 6, a comparison workflow that defines which datasets are compared and how the comparison results are generated is designed and edited via an interactive panel. When designing a comparison workflow, there are two mandatory components (the datasets to be compared and the quantitative comparison methods) and one optional component (the data processing methods). The dataset component comes from the selected PIC instances (which could be observations, benchmark data and model outputs). The comparison method component is a set of algorithms that require a list of comparable datasets and provide statistical or visual comparison results. The data processing method component is a set of algorithms that require input data and provide processed output data. The data processing method component is needed when the original datasets are not compatible with a comparison method.

Traditionally, comparison methods and data processing methods are implemented as offline tools, and they are strongly connected with specific model-data comparison tasks. To help participants conduct PIC tasks online, it is essential to make these methods function through reusable web services with standard APIs. Therefore, we designed two normalized interfaces for creating and using comparison methods and data processing methods (as shown in Fig. 7).

(1) Based on the Python language, the function interface defines the main entrance point for invoking a method. The inputs, parameters and outputs of a method must be transferred sequentially via the system command arguments. An "Execute" function is then invoked to carry out the calculation. By implementing a customized "Execute" function with the help of the Python programming template, participants can create method tailored to their needs.

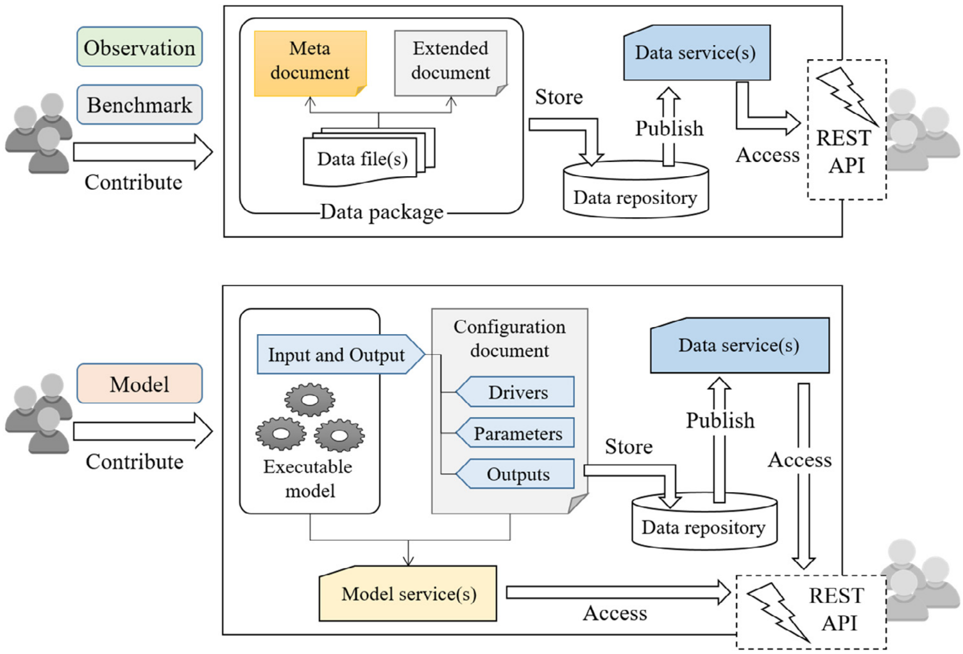


Fig. 4. Workflow for creating PIC instances of different types.

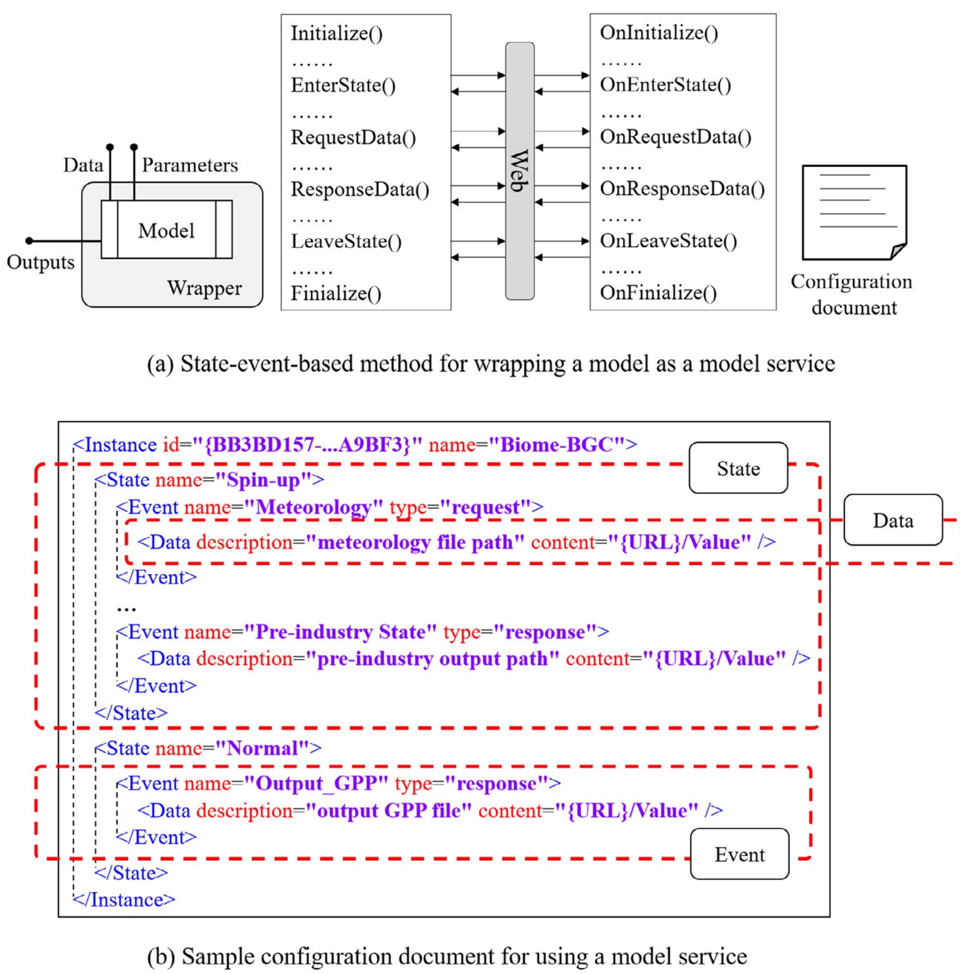


Fig. 5. Basic framework of reusable model services in a PIC instance.

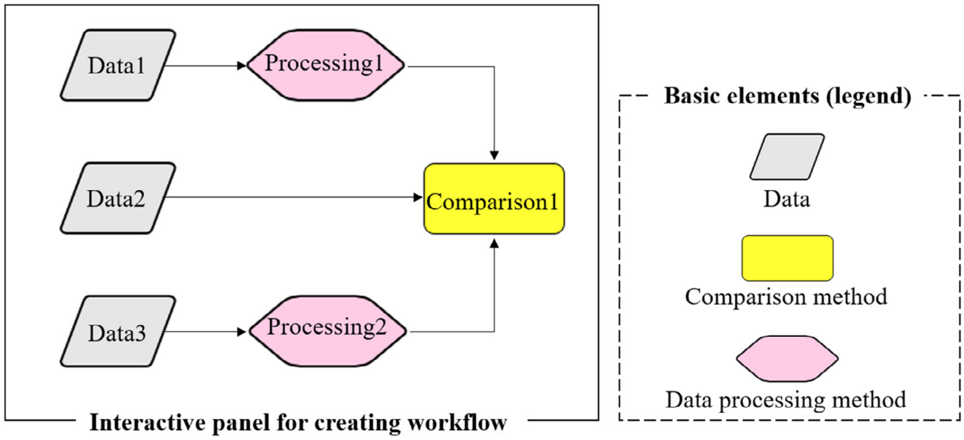


Fig. 6. Workflow for model-data comparison.

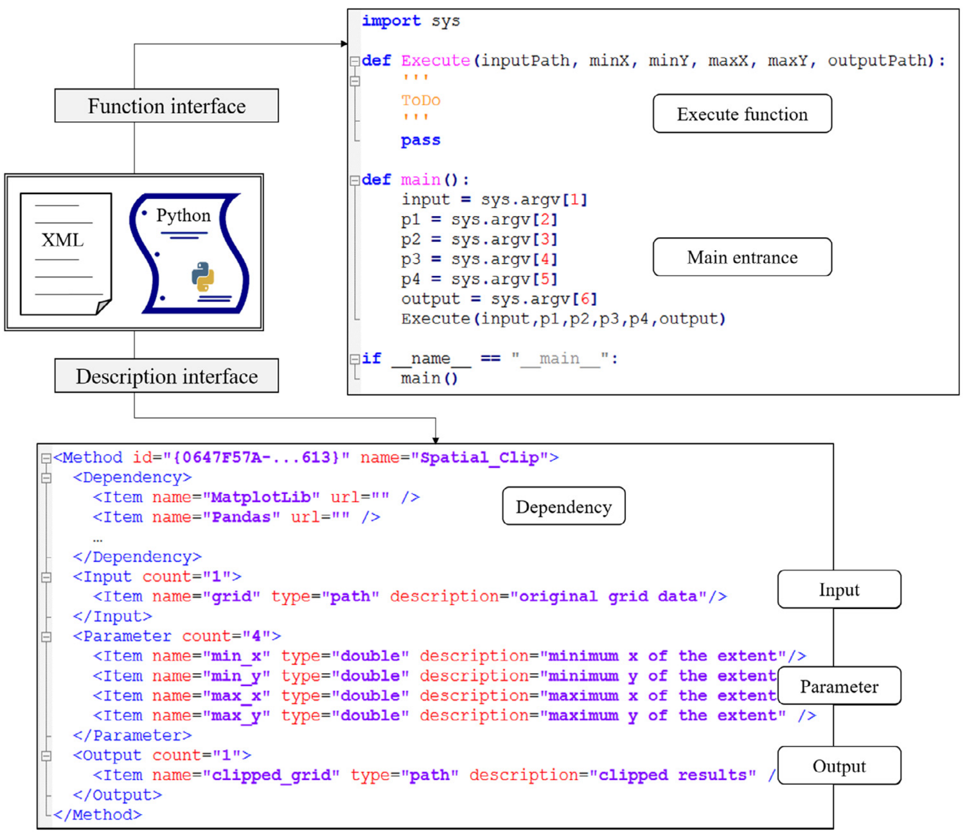


Fig. 7. Normalized interfaces for comparison methods and data processing methods.

(2) Another interface is the description interface, which focuses primarily on providing information regarding method use (in both human-readable and machine-interpretable forms). XML markup is employed to structurally describe the associated dependencies, inputs, parameters and outputs. The item nodes within the dependency section are used to document the Python packages a method requires. Universal resource links (URLs) are provided for downloading the packages and installing them on backend servers in the web environment. The item nodes within the input, parameter and output sections are used to help users understand the method and support the formation of GUI tools in the web application.

Because the Python script and XML document implement the functional and descriptive interfaces, respectively, they need to be packaged into a .zip file. The .zip file is contributed to the PIC task by the participator; then, a corresponding service can be generated automatically. In the interactive panel for creating workflows (shown in Fig. 6), each shape links to a detailed dataset or method, and the dependency relationships among these components are indicated by arrows. Because model-data comparison can be conducted at both site and grid levels, several different comparison workflows can be constructed to help researchers evaluate model performances and analyze model simulation results. The comparison workflow for a PIC instance should also be stored in the web environment so that all participants can check it out and repeat the comparison experiments. Storing such comparison workflows is also an important aspect of facilitating open and reproducible model-data comparison research.

4. Experiments

4.1. Web application for PIC projects

In this study, a prototype web platform named the Participatory Model-data Intercomparison Platform (P-MIP) was established to help various researchers create customized PIC projects (http://geomodeling.njnu.edu.cn/PMIP/cmp-projectlist). This platform is endorsed by the Open Geographic Modeling and Simulation (OpenGMS) project (http://geomodeling.njnu.edu.cn). Fig. 8 presents the basic architecture of the web application. Overall, the web application includes three fundamental modules: a service module, a management module and a PIC project module. Detailed technical implementations are presented in Supplementary B.

In the service module, the model service containers and data service containers are used to publish model and data services, all which are connected to the portal server. Through the portal server, users can access all the resources and services in the web platform (under user permission controls). Based on the service module, the management functionalities are provided to the management module, and the execution functionalities are provided to the PIC project module.

In the management module, the project management submodule is mainly used to help users create, delete, discover, join and maintain PIC projects; the resource management submodule is mainly used to help users contribute and utilize resources (e.g., models, data, and documents); and the user management submodule is mainly designed to maintain user information, such as new user registrations, participant invitations and participant role assignments (e.g., project leader, steering council member, or scientific committee member). Overall, the management module provides project maintenance functionalities for the PIC project module.

In the PIC project module, the PIC topic, PIC instance and PIC task are implemented as individual submodules, which allows participants to play different roles in contributing resources, conducting model experiments and analyzing comparison results. (1) One and only one PIC topic submodule can be created when establishing a PIC project. With the help of the communication toolset, descriptive information and protocols can be discussed online. (2) In the PIC instance submodule, a model service engine was developed to execute models online. Via the web application, settings of the drivers, parameters and output variables can be edited, and the XML-based configuration document can be generated and transferred to the backend server. Thus, the model service engine executes the model according to the specified configuration. (3) In the PIC task submodule, a comparison workflow engine is developed to execute a model-data comparison online. The comparison workflow can be constructed visually through an interactive panel, and the corresponding XML-based document is generated that records the related datasets, comparison methods and data processing methods and then transferred to the backend server. The comparison workflow engine parses the configuration file and then executes the workflow.

Regarding the relationships among the three submodules, a PIC instance can be linked to a PIC task, and a PIC task can link several PIC instances; all PIC instances and PIC tasks are linked to a PIC topic via an associated PIC project. Fig. 9 shows some snapshots of the web application.

4.2. PIC case of Biome-BGC, IBIS and LPJ

An experimental model-data intercomparison case was studied based on the proposed strategy. As shown in Fig. 10, six participators located in five different places are working on the web application. The PIC topic was co-designed by four participators; then, another researcher (participant E) joined the project and contributed PIC instances. Based on the PIC topic and PIC instances, different PIC tasks were customized by the existing and new participants to compare the model data from various aspects.

Table 1
Basic characteristics of the models involved in the case study.

	IBIS [25]	Biome-BGC [23]	LPJ [27]
Type	DGVM	TBM	DGVM
Time step of the simulation	Hourly	Daily	Daily
Time step of driver data	Daily	Daily	Daily
Spatial resolution	0.5°, 1°, 2°, 4°	No restriction	0.5°
Programming language	Fortran 77	C	Fortran 90

Three different terrestrial carbon cycle models are involved in this case; IBIS, Biome-BGC and LPJ were employed to create a PIC project, as shown in Table 1. Regarding the time step of the simulation, IBIS is hourly, Biome-BGC is daily, and the original LPJ is monthly. In this work, a modified LPJ-Daily [46] was used. The spatial resolution of IBIS can be 0.5°, 1°, 2°, or 4°; Biome-BGC has no specific restriction regarding spatial resolution; however, LPJ is limited to 0.5°. IBIS and LPJ were developed using the Fortran programming language, while Biome-BGC was developed with the C programming language. Technically, all these models can be executed on a system running either Windows or Linux, but they are usually executed on a Linux system. In addition to these basic similarities and heterogeneities, the models also vary in terms of their driver data, biome classification, output variables, and other detailed conditions.

4.2.1. PIC topic of the demonstrative PIC project

The name of this project is "*A model-data comparison project for historical simulation*", and the background is summarized as follows. "*Various terrestrial carbon cycle models have been developed to simulate and predict carbon exchange. Some are static vegetation models, while some are dynamic vegetation models. In addition, the biogeochemistry, biophysical and biography processes in these models are also diverse. Thus, the differences presented in different models' simulation results need to be analyzed.*" The goals of this project are to "*evaluate the IBIS, Biome-BGC and LPJ models' historical simulations at the site level and grid level*". This information can be accessed and edited through the PIC topic web page.

To avoid ambiguities in people's understanding of the related expressions, the four repositories (concept, units, data templates and spatiotemporal references) are constructed within the project. As shown in Fig. 11, the GPP, NPP, NEP, LAI, AT-Neu and other terms or phrases are organized in the concept repository. The gram (g), kilogram (kg), centimeter (cm), meter (m), Kelvin (K), null (NA), gram per centimeter square per day (g cm^{-1} d^{-1}) and other units are defined in the unit repository. The netCDF, shapefile and GeoJSON formats are described in the data template repository. The WGS84 spatial coordinate system and UTC temporal coordinate system are given in the spatiotemporal repository. All these resource items are described in a structural and compatible multilanguage manner.

In terms of the comparison protocols, the common driver datasets were contributed to the project. For example, the meteorological, soil, elevation, CO_2 concentration and biome classification data are shared as data services; the corresponding meta-documents and extended documents can also be accessed to understand these datasets. In addition, the spin-up and simulation processes are described in the model process protocols, and the output variables are detailed in the model output protocols. All these protocols are collaboratively formulated by the participants, as shown in Fig. 12, and they can be reused throughout the PIC platform.

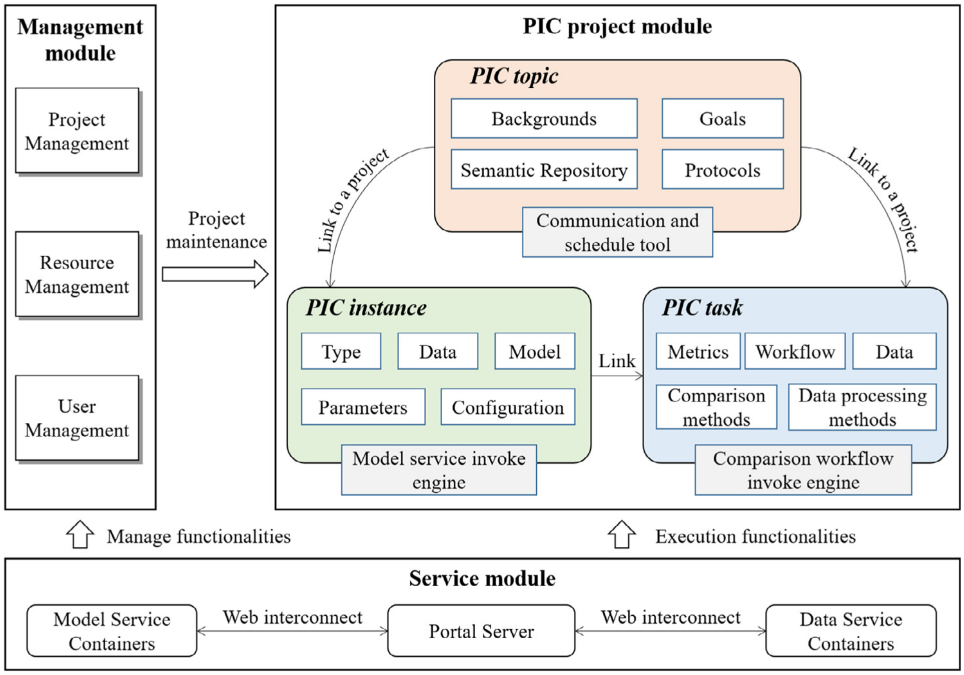


Fig. 8. Architecture of the web application for PIC projects.

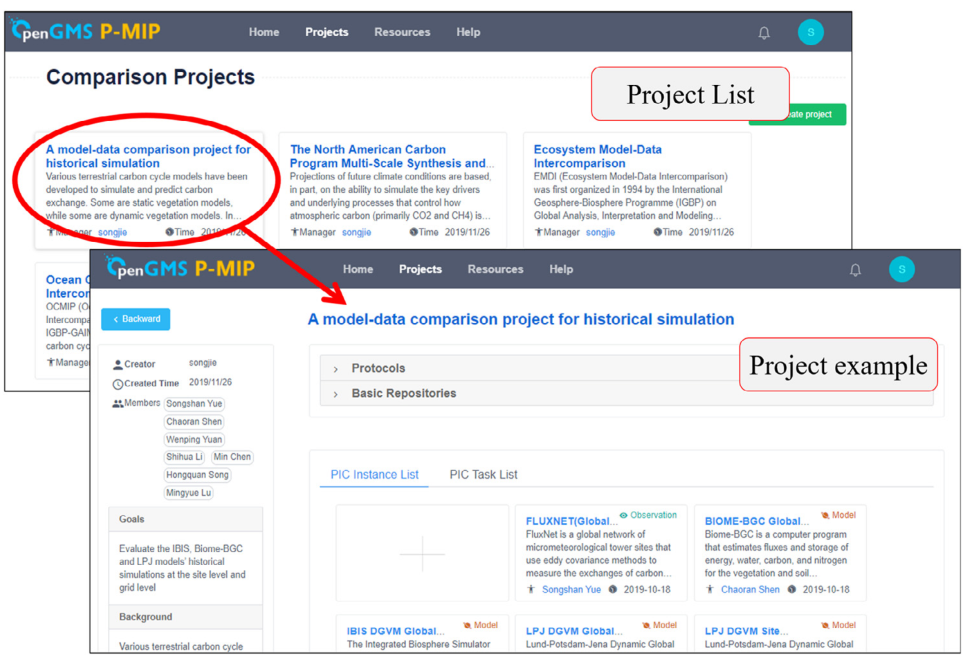


Fig. 9. Snapshots of the P-MIP web application.

4.2.2. PIC instances of the demonstrative PIC project

According to the PIC topic, three PIC instances of model type were created by three different participants, which corresponded to the IBIS, Biome-BGC and LPJ models. In each of these PIC instances, the settings of the drivers, parameters and output variables can be configured via a web page GUI. Thus, the reproducibility of a model experiment can be checked by all participants.

In addition, a PIC instance of observation type was created based on the Fluxnet datasets (https://fluxnet.fluxdata.org/data/fluxnet2015-dataset/). The Fluxnet datasets offer observations from 231 sites, among which the observations of 129 sites were used after removing errors. The time range of each site in the Fluxnet network varies, while the temporal resolution for all sites is daily.

In addition, the Moderate Resolution Imaging Spectroradiometer (MODIS) MOD17A2 products [47] were employed to create a PIC instance of benchmark type. The GPP and NPP are offered through such PIC instances. The time range of the MOD17A2 products is 2000–2015, and the time resolution is 8 days.

As shown in Fig. 13, all the PIC instances were created by different participants. For a PIC instance of model type, it can be duplicated into a new PIC instance and then re-invoke the related model with modified parameters. By selecting different PIC instances, comparison tasks can be conducted from multiple

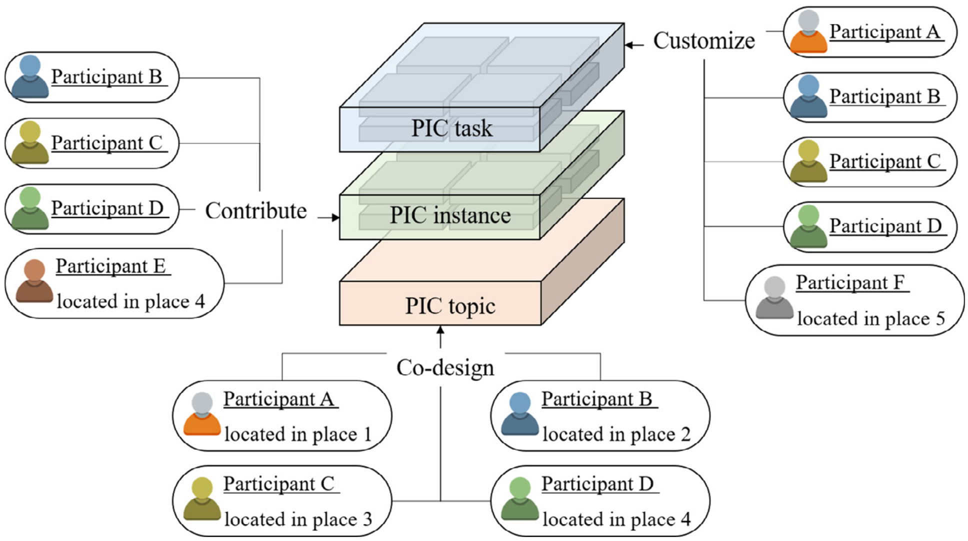


Fig. 10. Experimental case conducted by different participants.

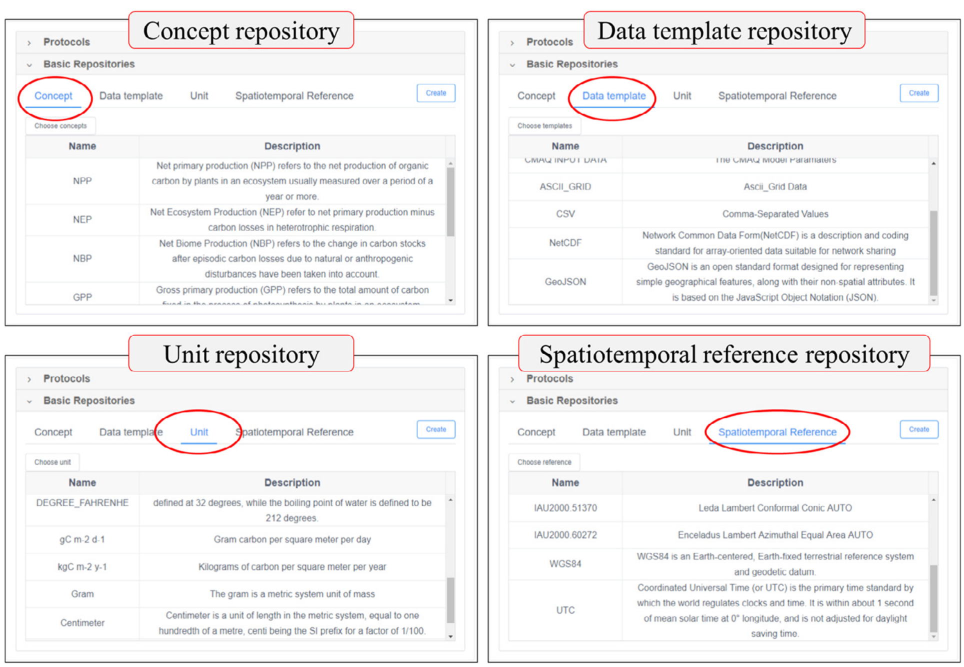


Fig. 11. A PIC topic and the four basic repositories.

aspects, such as comparing outputs of one model using different drivers or parameters, comparing outputs of different models, and comparing models' outputs with observation/benchmark datasets.

4.2.3. PIC tasks in the demonstration PIC project

Based on the co-designed protocols and the collaboratively created PIC instances, two typical comparison tasks were conducted that focused on comparisons at the site and grid levels.

4.2.3.1. Site-level comparison task. In the site-level comparison task, the three PIC instances of model type, one PIC instance of observation type (site AT-Neu), and one PIC instance of benchmark type were selected for intercomparison. The site AT-Neu, located at 11.3175° longitude and 47.1167° latitude in Austria (http://www.google.com/maps/place/47%C2%B007'00.1%22N+11%C2%B019'03.0%22E/@47.1818197,10.4704477,7.92z/data=!4m5!3m4!1s0x0:0x0!8m2!3d47.1167!4d11.3175), belongs to the grassland biome system, and it includes 11 years of observation data (from 2002 to 2012). A comparison workflow was formulated, as shown in Fig. 14. Several data processing methods were implemented to calculate statistical factors from the GPPs of different PIC instances. In the workflow, the mean (MEAN), standard deviation (SD), root mean square error (RMSE), correlation coefficient (COEF), and Nash–Sutcliffe coefficient (NSE) were used. A comparison method ("Table Comparison" in the workflow) was used to generate a table that presents all the PIC instance factors structurally. Another comparison method, named the "Taylor diagram" [48], was used to generate a diagram comparing the SD, RMSE and COEF via a pie chart. In addition, a third comparison method, named "Time Series", was used to generate a line chart that presented the GPP trends over time. These comparison results are all accessible online, as shown in Fig. 15.

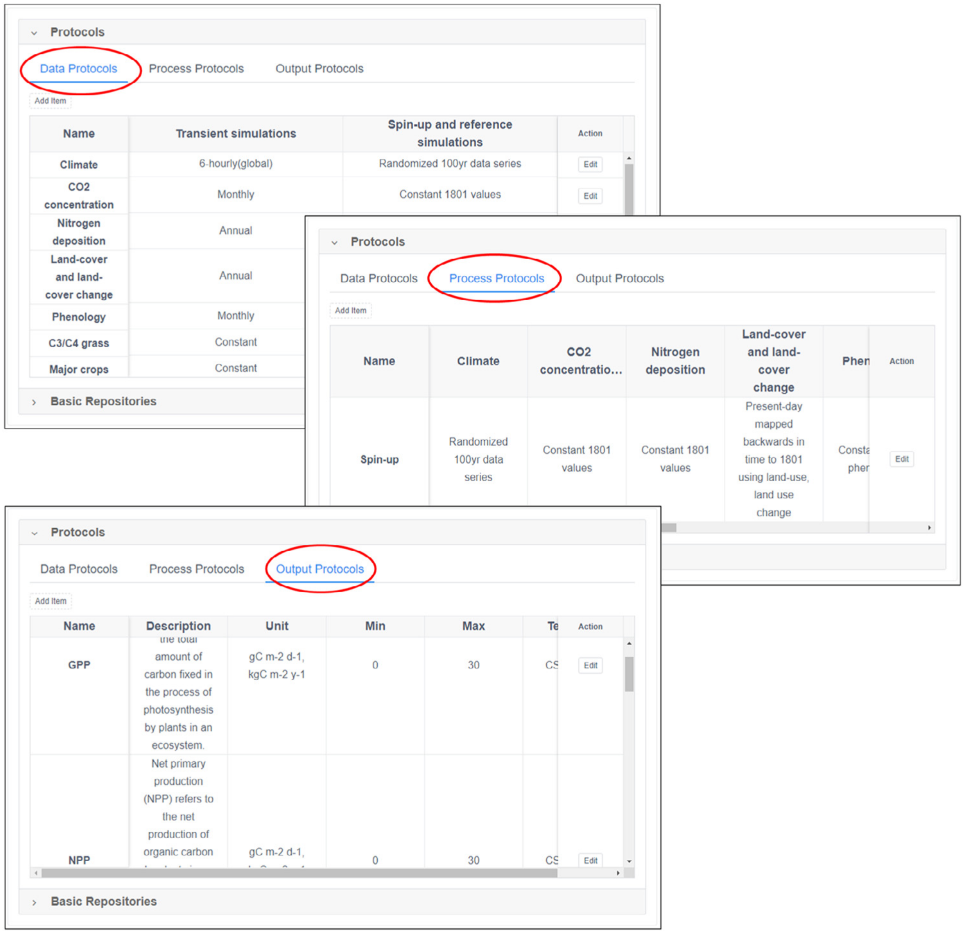


Fig. 12. PIC topic and model-data comparison protocols.

Based on Fig. 15(a), the comparison results can be analyzed from different aspects. For the MEAN, Biome-BGC is the closest to the observation data, while those for IBIS and MOD17A2 are high and those for LPJ are low. For the SD, the simulation results of IBIS and MOD17A2 are relatively similar, as are the results for Biome-BGC and LPJ. Regarding the RMSE, Biome-BGC has the smallest value. The COEF between the simulation results of IBIS, Biome-BGC and LPJ and the observed values are above 0.69, which means that their simulation performances are good. For the NSE, Biome-BGC has a value of 0.57, while the values of LPJ, IBIS and MOD17A2 are relatively similar. In Fig. 15(b), the RMSE, SD, and COEF are combined to form a comprehensive evaluation. According to Fig. 15(c), the GPPs of IBIS, Biome-BGC, LPJ, MOD17A2 and Fluxnet are consistent regarding their time distribution trends (reaching peaks in summer and lows in winter). However, the GPP ranges vary widely. The peaks of IBIS and Fluxnet are relatively close (approximately 16), while the others have smaller peaks (approximately 8). Overall, it is difficult to determine which model(s) is better for the AT-Neu site. Participants can analyze the results and discuss the model performances from different aspects.

4.2.3.2. Grid-level comparison task. In the grid-level comparison task, the three PIC instances of model type and the MOD17A2 PIC instance were selected for intercomparison. Comparison workflows were constructed, as shown in Fig. 16. There are three comparison branches in this case: temporal evolution, spatial distribution, and spatiotemporal trends.

(1) Regarding the temporal evolution comparison, the data processing method named "Grid Extraction" was implemented to calculate a list of grid data within a certain range, and the data processing method named "Period Statistics" was implemented to calculate the minimum–maximum-average values of the grid over a time range. The "Grid Extraction" method requires the latitudinal extent, which, in this case, was set for three regions: mid-high latitudes in the Northern Hemisphere (20°N to 85°N), mid-high latitudes in the Southern Hemisphere (20°S to 60°S), and low latitudes (20°S to 20°N). The "Period Statistics" method requires a time step and time range parameters, which were set to one month and from 1982 to 2014, respectively. A comparison method named "Seasonal Variation" was implemented as a reusable web service. This method requires a list of value arrays and generates a line chart to present the comparison results. This comparison method was employed for the three regions, and the corresponding comparison results are shown in Fig. 17.

(2) Regarding the spatial distribution comparison, the data processing method ("Grid Average") was used, which requires a list of grid results and calculates the average value of each grid cell. A comparison method named "Map Distribution" was implemented as a reusable web service and linked to the "Grid Average" (using GPP as the grid input). This comparison method requires a list of grid data (in netCDF format) and generates the maps for each grid. The map-based comparison results are shown in Fig. 18.

(3) Regarding the spatiotemporal trend comparison, the data processing method "Grid Extraction" was used, and the latitudinal extent parameter was set to 60°S to 85°N. The data processing method "Period Statistics" was used with the time step set to one month and the time range set from 2014 to 2014 (that is, 12 months in one year). The average value list generated

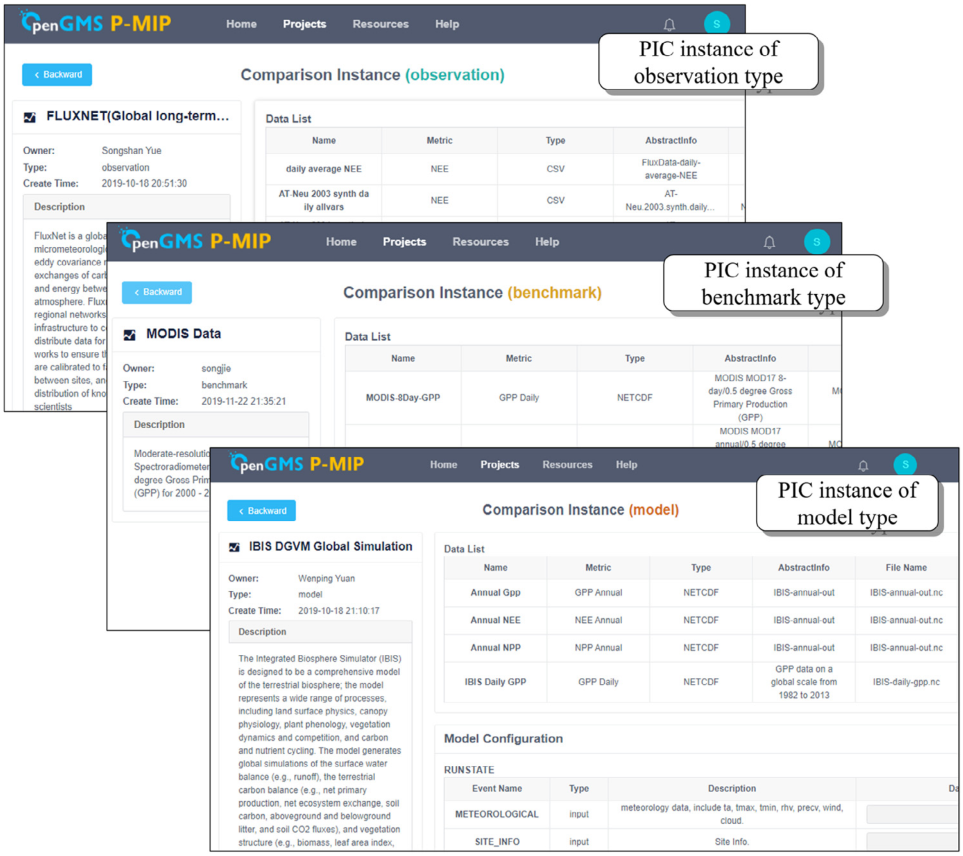


Fig. 13. PIC instances in the demonstration project.

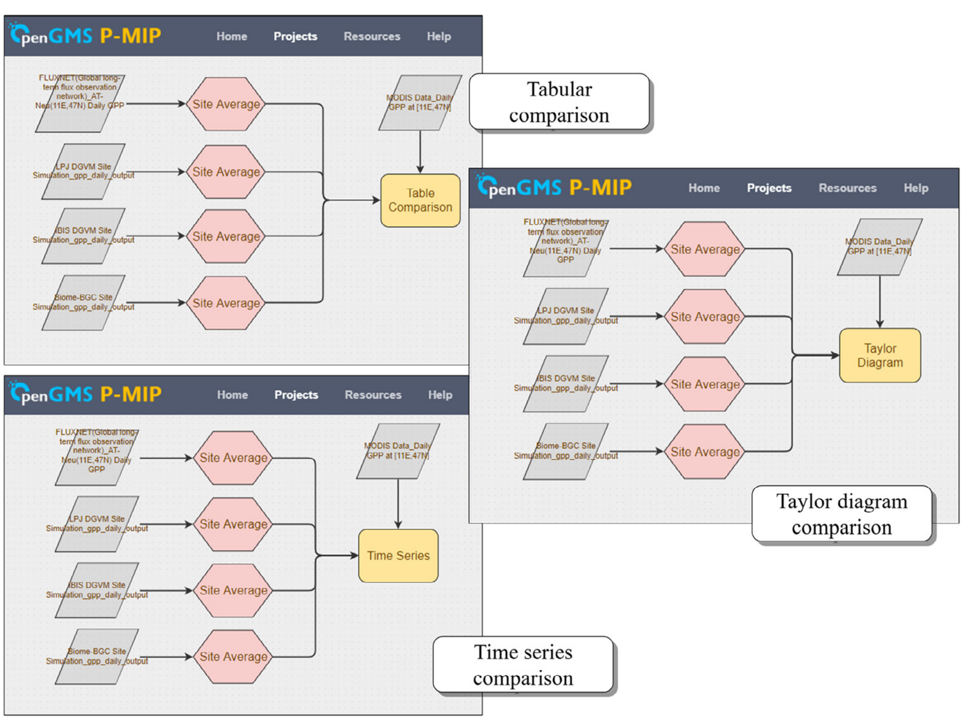
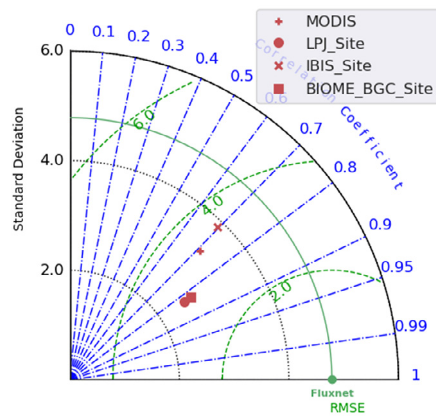


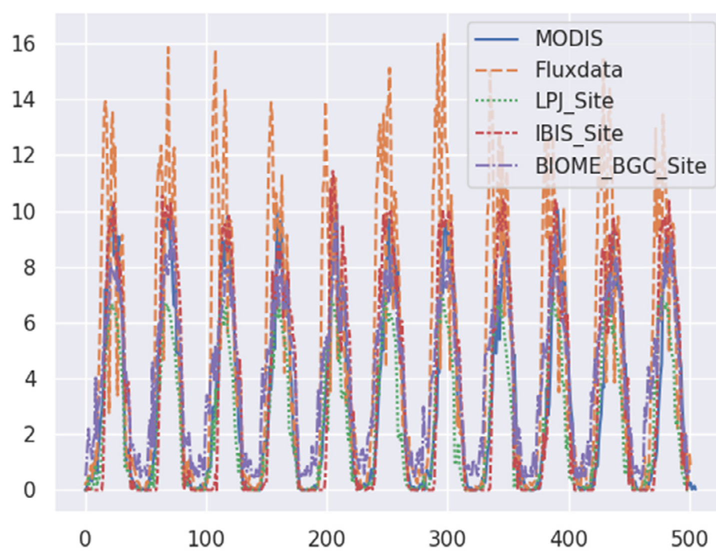
Fig. 14. Site-level comparison workflow.

Facors	LPJ_Site	IBIS_Site	BIOME_BGC_Site	MODIS	Fluxnet
MEAN	2.39	3.58	4.05	3.28	4.99
STD	2.52	3.88	2.67	3.34	4.78
RMSE	4.01	3.76	3.13	3.78	0.0
COEF	0.83	0.69	0.82	0.71	1.0
NSE	0.3	0.38	0.57	0.38	0.0

(a) Tabular comparison results



(b) Taylor diagram comparison results



(c) Time series comparison results

Fig. 15. Site-level comparison results.

by the "Period Statistics" form the input for the comparison method "Seasonal Latitude Distribution"; the latter generates a chart presenting the spatial distribution in the vertical direction and the temporal evolution in the horizontal direction. The spatiotemporal trend comparison results are shown in Fig. 19.

As shown in Fig. 17, the basic seasonal trend in GPP is consistent among the IBIS, Biome-BGC, LPJ and MOD17A2 models. However, some differences exist in the value ranges of the four models. In the mid-high latitudes in the Northern Hemisphere, Biome-BGC, LPJ and MOD17A2 are similar, but IBIS has a much higher value range. In the mid-high latitudes in the Southern Hemisphere, the four methods have fewer differences in the spring and autumn but more differences in the winter and summer. In the low-latitude region, there are significant differences in the value ranges of the four methods (IBIS has the highest range, while LPJ and Biome-BGC have ranges much lower than the other two methods). Overall, the four methods are more consistent in the mid-high latitude region than in the low latitude region.

The spatial distributions of the four similar patterns can be seen in Fig. 18. Regions with higher GPP are distributed in wet

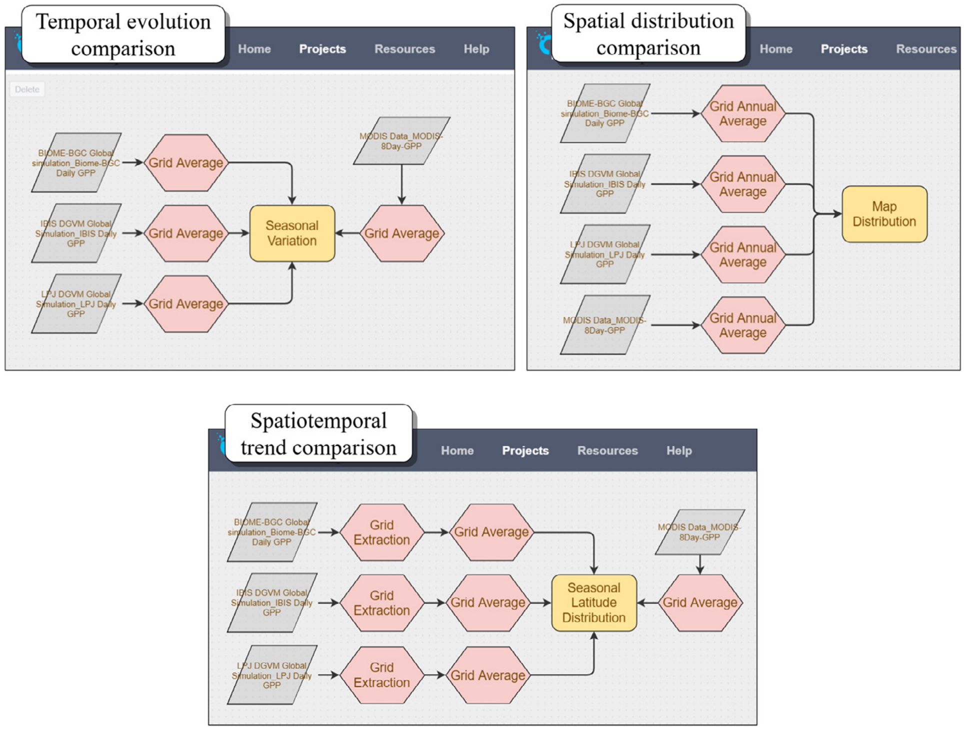


Fig. 16. Grid-level comparison workflow.

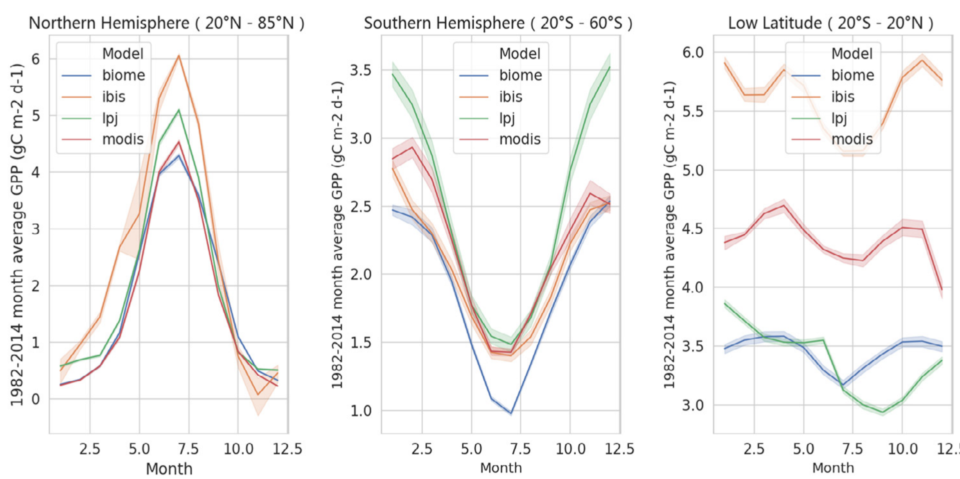


Fig. 17. Temporal evolution comparison results.

tropical areas (e.g., the Amazon, Central Africa and Southeast Asia), where the temperature and humidity meets the needs of photosynthesis. Regions with lower GPP are distributed in cold and arid areas, reflecting that temperature and water are important for plant growth. However, the map comparison results also show that the GPP simulation results in mid-high latitude areas are much more consistent than those in low latitude areas. Similar conclusions were reported by Anav et al. [49] and Malavelle et al. [50].

Fig. 19 indicates that GPP has a seasonal difference in latitude, and all four instances indicate that the growing season is longer in the equatorial region compared to that of other regions. Summer is the only growing season in the mid-high latitude area. The IBIS-simulated GPP has fewer seasonal fluctuations in the equatorial region than in other regions (i.e., the growing season lasts throughout the year), resulting in a much higher value than those of the other three models. The LPJ-simulated GPP is distributed more uniformly in time and space, and the GPP values in the equatorial region and in the mid-high latitude region are roughly the same in the growing season. In addition, the GPP values of Biome-BGC and MOD17A2 are relatively similar.

Through site- and grid-level comparison PIC tasks, participants can analyze the model results and evaluate model performances from different aspects. All the data processing methods and comparison methods are contributed by participants and can be reused in all PIC tasks in the project. New PIC tasks can also be created by duplicating an existing task and modifying the comparison workflow to acquire different comparison results.

5. Conclusions and future work

This article presented a PIC strategy for terrestrial carbon cycle models. To help researchers collaboratively conduct model-data comparison tasks in a web environment, we proposed a set of

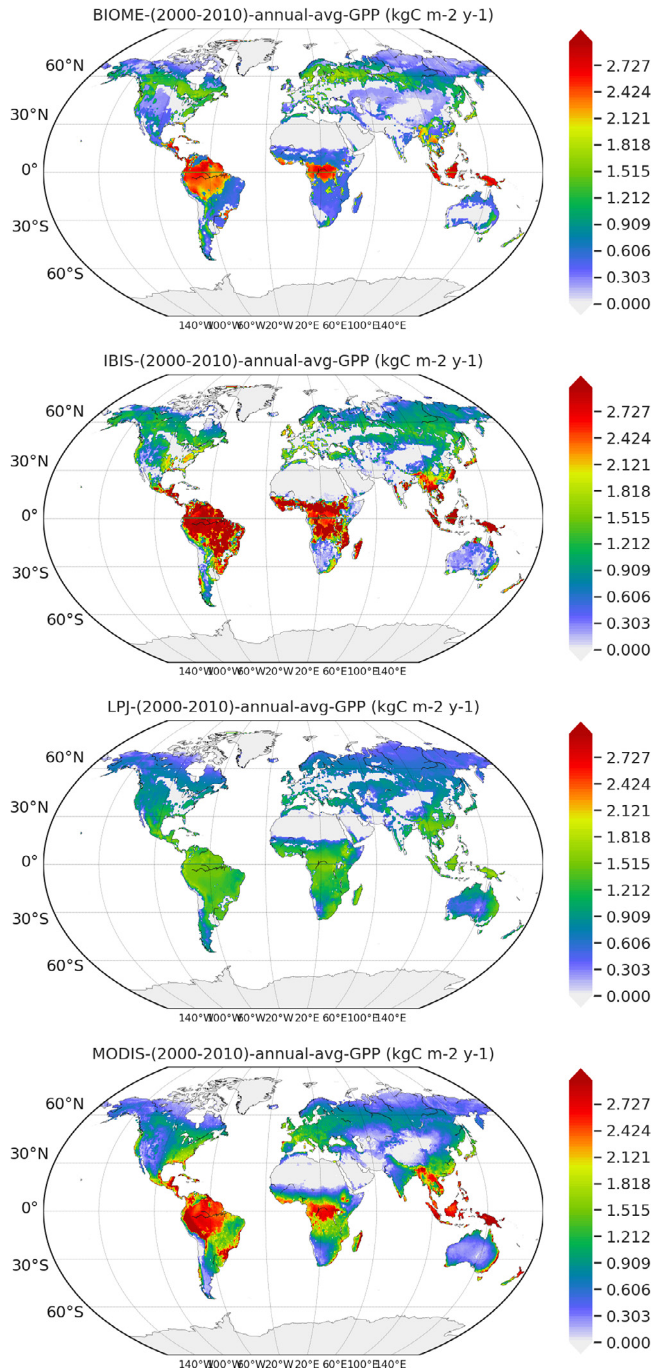


Fig. 18. Spatial distributions of the comparison results.

methods to create and maintain a PIC project, which includes the PIC topic, PIC instance and PIC task. The backgrounds, goals, and comparison protocols of a PIC project are organized and co-designed in the associated PIC topic; the models and observations to be compared are organized and executed in different PIC instances; and the detailed intercomparison processes are organized and conducted in different PIC tasks. All the participating models, observation data, benchmark data, data processing methods and comparison methods are wrapped into web services, enabling the model and comparison experiments to be reusable and reproducible. Based on the proposed strategy, communication barriers among researchers and resource sharing barriers among different MIPs can be reduced; thus, the cost of constructing a new MIP can be decreased.

Following the work presented in this paper, the PIC strategy could be employed to perform model-data comparisons in other study areas. For example, a PIC project could be established to compare different clustering methods (e.g., DBSCAN, K-Means, OPTICS, DenStream). In this case, PIC instances of model type could be created for each involved clustering method, and PIC instances of benchmark type could also be used if the methods were applied to a synthetic dataset with predefined clusters. However, PIC instances of observation type are usually not used because clusters are not typically separated in real-world datasets. Moreover, we envision that PIC projects can also be established for education purposes, providing help to beginning modelers trying to understand how to use models [51,52].

Due to the complexity of terrestrial carbon cycle modeling work, future participatory model-data intercomparison work is in high demand, especially from the following aspects:

(1) Integration of existing MIPs. A web-based method for constructing and conducting MIPs was proposed to help various researchers collaborate online. Through this approach, the dependent resources of a PIC project can be contributed by all participants (in a collaborative manner). On the one hand, participants should be able to offer their own models, datasets, and algorithms; on the other hand, many resources have already been developed by recent MIPs, and these resources can also be shared as reusable web services to help researchers conduct PIC projects more conveniently. Issues of intellectual property and license control should also be considered.

(2) Model, data and parameter recommendation. When conducting model-data comparison work, it is essential to collect suitable models and datasets for comparison. By manually creating PIC instances, participants can offer their own resources to collect comparable models and datasets for a project. Because different models and data can be shared as reusable services in the web environment, another alternative that needs to be considered is recommending existing model and data services based on the protocol formulated for a PIC project. Moreover, parameter recommendations based on similarities to existing model experiments can also improve the convenience of conducting a PIC project [53].

(3) Exception handling in model services. Execution errors occur frequently when executing terrestrial carbon cycle models; such errors can be caused by numerous conditions, such as inputting unsuitable driver data, applying unsuitable spatial/temporal scales, or failing to consider parameter configurations. In a web-based participatory intercomparison environment, it is important to inform the participants about the model limits. Based on such information, participants can modify and refine their related model experiments. Therefore, studies about the fault tolerant and situation handling are necessary to facilitate participatory model-data intercomparison.

CRediT authorship contribution statement

Songshan Yue: Methodology, Writing - original draft, Writing - review & editing. **Min Chen:** Conceptualization, Writing - review & editing, Funding acquisition. **Jie Song:** Methodology, Software. **Wenping Yuan:** Validation, Data curation. **Tiexi Chen:** Writing - review & editing. **Guonian Lü:** Supervision. **Chaoran Shen:** Investigation, Resources. **Zaiyang Ma:** Methodology, Software. **Kai Xu:** Visualization. **Yongning Wen:** Project administration. **Hongquan Song:** Data curation, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

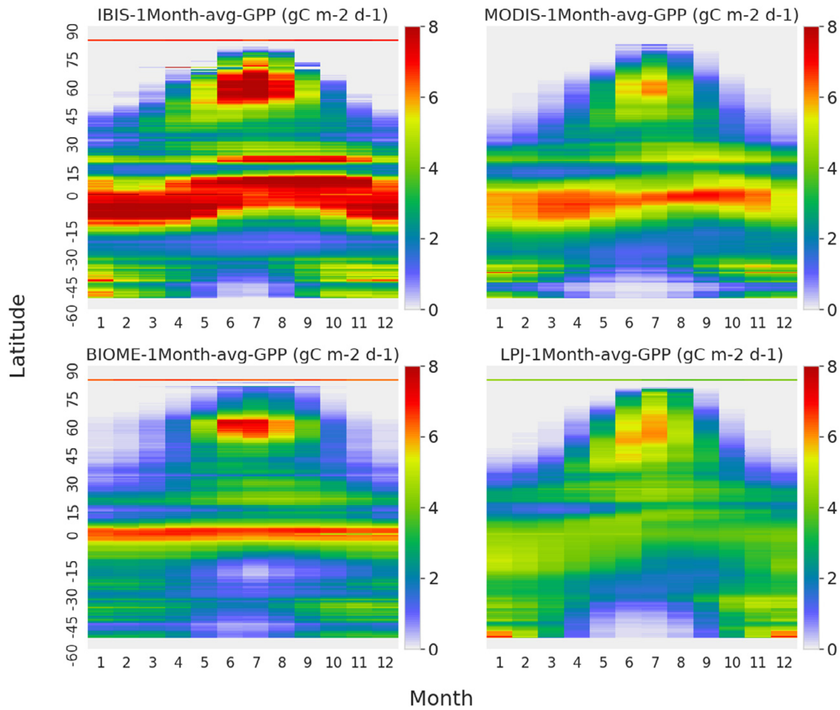


Fig. 19. Spatiotemporal trends of the comparison results.

Acknowledgments

We appreciate the detailed suggestions and comments from the editor and the anonymous reviewers. The work described in this article was supported by the following research programs: the National Natural Science Foundation (NSF) of China [grant numbers 41701441, 41930648, 31570464 and U1811464], the NSF for Excellent Young Scholars of China [grant number 41622108], the Priority Academic Program Development of Jiangsu Higher Education Institutions [grant number 164320H1 16], the National Key R&D Program of China [grant number 2017YFB0504000] and the Open Fund of Key Laboratory of Geosp atial Technology for the Middle and Lower Yellow River Regions (Henan University), Ministry of Education (grant number GTYR201807).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.future.2020.05.044.

References

[1] P.M. Cox, R.A. Betts, C.D. Jones, S.A. Spall, I.J. Totterdell, Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model, Nature 408 (6809) (2000) 184.
[2] A.D. McGuire, S. Sitch, J.S. Clein, et al., Carbon balance of the terrestrial biosphere in the twentieth century: Analyses of CO2, climate and land use effects with four process-based ecosystem models, Glob. Biogeochem. Cycles 15 (1) (2001) 183–206.
[3] J.G. Canadell, C. Le Quéré, M.R. Raupach, et al., Contributions to accelerating atmospheric CO2 growth from economic activity, carbon intensity, and efficiency of natural sinks, Proc. Natl. Acad. Sci. 104 (47) (2007) 18866–18870.
[4] S. Sitch, P. Friedlingstein, N. Gruber, et al., Recent trends and drivers of regional sources and sinks of carbon dioxide, Biogeosciences 12 (3) (2015) 653–679.
[5] K. Elofsson, M. Gren, Cost-efficient climate policies for interdependent carbon pools, Environ. Model. Softw. 101 (2018) 86–101.
[6] D.S. Schimel, Terrestrial ecosystems and the carbon cycle, Glob. Change Biol. 1 (1) (1995) 77–91.
[7] M. Cao, F.I. Woodward, Dynamic responses of terrestrial ecosystem carbon cycling to global climate change, Nature 393 (6682) (1998) 249.
[8] S. Sitch, C. Huntingford, N. Gedney, et al., Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five dynamic global vegetation models (DGVMs), Global Change Biol. 14 (9) (2008) 2015–2039.
[9] T. Le Toan, S. Quegan, M.W.J. Davidson, et al., The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle, Remote Sens. Environ. 115 (11) (2011) 2850–2860.
[10] B. Quesada, A. Arneth, E. Robertson, N. de Noblet-Ducoudré, Potential strong contribution of future anthropogenic land-use and land-cover change to the terrestrial carbon cycle, Environ. Res. Lett. 13 (6) (2018) 064023.
[11] J. Song, S. Wan, S. Piao, A.K. Knapp, A.T. Classen, S. Vicca, …, P. Kardol, A meta-analysis of 1, 119 manipulative experiments on terrestrial carbon-cycling responses to global change, Nat. Ecol. Evol. 3 (9) (2019) 1309–1320.
[12] J. Qi, X. Du, X. Zhang, S. Lee, Y. Wu, J. Deng, et al., Modeling riverine dissolved and particulate organic carbon fluxes from two small watersheds in the northeastern United States, Environ. Model. Softw. 124 (2020) 104601.
[13] T.M. Smith, W.P. Cramer, R.K. Dixon, R. Leemans, R.P. Neilson, A.M. Solomon, The global terrestrial carbon cycle, Water Air Soil Pollut. 70 (1–4) (1993) 19–37.
[14] W.R. Emanuel, G.G. Killough, Modeling terrestrial ecosystems in the global carbon cycle with shifts in carbon storage capacity by land-use change, Ecology 65 (3) (1984) 970–983.
[15] R.M. Gifford, Plant respiration in productivity models: conceptualisation, representation and issues for global terrestrial carbon-cycle research, Funct. Plant Biol. 30 (2) (2003) 171–186.
[16] B. Poulter, D.C. Frank, E.L. Hodson, N.E. Zimmermann, Impacts of land cover and climate data selection on understanding terrestrial carbon dynamics and the CO 2 airborne fraction, Biogeosciences 8 (8) (2011) 2027–2036.
[17] Y. Wen, X. Liu, et al., Determining the impacts of climate change and urban expansion on terrestrial net primary production in China, J. Environ. Manag. 240 (2019) 75–83.
[18] W. Ye, F. Zhang, Y. Bai, Z. Du, R. Liu, A tile service-driven architecture for online climate analysis with an application to estimation of ocean carbon flux, Environ. Model. Softw. 118 (2019) 120–133.
[19] P. Falkowski, R.J. Scholes, E.E.A. Boyle, et al., The global carbon cycle: a test of our knowledge of earth as a system, Science 290 (5490) (2000) 291–296.
[20] H. Craig, The natural distribution of radiocarbon and the exchange time of carbon dioxide between atmosphere and sea, Tellus 9 (1) (1957) 1–17.
[21] W.J. Parton, J.W. Stewart, C.V. Cole, Dynamics of c, n, p and s in grassland soils: a model, Biogeochemistry 5 (1) (1988) 109–131.

[22] W.J. Parton, J.M.O. Scurlock, D.S. Ojima, T.G. Gilmanov, R.J. Scholes, D.S. Schimel, ., A. Kamnalrut, Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide, Glob. Biogeochem. Cycles 7 (4) (1993) 785–809.

[23] S.W. Running, E.R. Hunt Jr., Generalization of a forest ecosystem process model for other biomes, BIOME-bcg, and an application for global-scale models, in: Scaling Physiological Processes: Leaf to Globe, Academic Press, San Diego, CA, 1993, pp. 141–157.

[24] M. Chiesi, F. Maselli, M. Moriondo, L. Fibbi, M. Bindi, S.W. Running, Application of BIOME-BGC to simulate mediterranean forest processes, Ecol. Model. 206 (1–2) (2007) 179–190.

[25] J.A. Foley, I.C. Prentice, N. Ramankutty, S. Levis, D. Pollard, S. Sitch, A. Haxeltine, An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics, Glob. Biogeochem. Cycles 10 (4) (1996) 603–628.

[26] R.P. Neilson, A model for predicting continental-scale vegetation distribution and water balance, Ecol. Appl. 5 (2) (1995) 362–385.

[27] S. Sitch, B. Smith, I.C. Prentice, et al., Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, Global Change Biol. 9 (2) (2003) 161–185.

[28] G. Pranuthi, S.K. Dubey, S.K. Tripathi, Comparison of different models for estimation of net primary productivity in India, J. Agrometeorol. 2 (14) (2012) 105–110.

[29] H. Lin, Y. Zhang, Evaluation of six methods to predict grassland net primary productivity along an altitudinal gradient in the Alxa Rangeland, Western Inner Mongolia, China, Grassl. Sci. 59 (2) (2013) 100–110.

[30] P. Morales, M.T. Sykes, I.C. Prentice, et al., Comparing and evaluating process-based ecosystem model predictions of carbon and water fluxes in major European forest biomes, Global Change Biol. 11 (12) (2005) 2211–2233.

[31] B. Byrne, Monitoring the Carbon Cycle: Evaluation of Terrestrial Biosphere Models and Anthropogenic Greenhouse Gas Emissions with Atmospheric Observations (Doctoral dissertation), 2018.

[32] C. Beer, M. Reichstein, E. Tomelleri, et al., Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate, Science 329 (5993) (2010) 834–838.

[33] P. Friedlingstein, M. Meinshausen, V.K. Arora, C.D. Jones, A. Anav, S.K. Liddicoat, R. Knutti, Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks, J. Clim. 27 (2) (2014) 511–526.

[34] D.N. Huntzinger, A.M. Michalak, C. Schwalm, et al., Uncertainty in the response of terrestrial carbon sink to environmental drivers undermines carbon-climate feedback predictions, Sci. Rep. 7 (1) (2017) 4765.

[35] P. Friedlingstein, P. Cox, R. Betts, et al., Climate–carbon cycle feedback analysis: results from the C4MIP model intercomparison, J. Clim. 19 (14) (2006) 3337–3353.

[36] S. Piao, S. Sitch, P. Ciais, et al., Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO2 trends, Global Change Biol. 19 (7) (2013) 2117–2132.

[37] D. Kim, M.I. Lee, S.J. Jeong, J. Im, D.H. Cha, S. Lee, Intercomparison of terrestrial Carbon fluxes and Carbon use efficiency simulated by CMIP5 Earth System Models., Asia-Pac. J. Atmos. Sci. 54 (2) (2018) 145–163.

[38] L.G.G. De Gonçalves, J.S. Borak, M.H. Costa, S.R. Saleska, I. Baker, N. Restrepo-Coupe, ., M.A. Arain, Overview of the large-scale biosphere–atmosphere experiment in Amazonia Data Model Intercomparison Project (LBA-DMIP), Agric. For. Meteorol. 182 (2013) 111–127.

[39] F.M. Hoffman, J.T. Randerson, I.Y. Fung, P.E. Thornton, Y.H. Lee, C.C. Covey, ., S.W. Running, The Carbon-land model intercomparison project (C-LAMP): A protocol and evaluation metrics for global terrestrial biogeochemistry models, 2008.

[40] C.D. Jones, V. Arora, P. Friedlingstein, L. Bopp, V. Brovkin, J. Dunne, ., M. Jung, C4mip–the coupled climate–Carbon cycle model intercomparison project: experimental protocol for cmip6, Geosci. Model Dev. 9 (8) (2016) 2853–2880.

[41] H. Lin, M. Chen, G. Lu, Q. Zhu, J. Gong, X. You, ., M. Hu, Virtual geographic environments (VGEs): a new generation of geographic analysis tool, Earth-Sci. Rev. 126 (2013) 74–84.

[42] M. Chen, S. Yue, G. Lü, H. Lin, C. Yang, Y. Wen, et al., Teamwork-oriented integrated modeling method for geo-problem solving, Environ. Model. Softw. 119 (2019) 111–123.

[43] S. Yue, M. Chen, Y. Wen, G. Lu, Service-oriented model-encapsulation strategy for sharing and integrating heterogeneous geo-analysis models in an open web environment, ISPRS J. Photogramm. Remote Sens. 114 (2016) 258–273.

[44] F. Zhang, M. Chen, D.P. Ames, C. Shen, S. Yue, Y. Wen, G. Lü, Design and development of a service-oriented wrapper system for sharing and reusing distributed geoanalysis models on the web, Environ. Model. Softw. 111 (2019) 498–509.

[45] J.E. Nash, J.V. Sutcliffe, River flow forecasting through conceptual models part I - a discussion of principles, J. Hydrol. 10 (3) (1970) 282–290, http://dx.doi.org/10.1016/0022-1694(70)90255-6.

[46] S. Venevsky, S. Maksyutov, SEVER: A modification of the LPJ global dynamic vegetation model for daily time step and parallel computation, Environ. Model. Softw. 1 (22) (2007) 104–109.

[47] F.A. Heinsch, M. Reeves, P. Votava, S. Kang, C. Milesi, M. Zhao, ., J.S. Kimball, Gpp and npp (mod17a2/a3) products nasa modis land algorithm, MOD17 User's Guide (2003) 1–57.

[48] K.E. Taylor, Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res.: Atmos. 106 (D7) (2001) 7183–7192.

[49] A. Anav, P. Friedlingstein, C. Beer, P. Ciais, A. Harper, C. Jones, ., S. Piao, Spatiotemporal patterns of terrestrial gross primary production: A review, Rev. Geophys. 53 (3) (2015) 785–818.

[50] F.F. Malavelle, J.M. Haywood, L.M. Mercado, G.A. Folberth, N. Bellouin, S. Sitch, P. Artaxo, Studying the impact of biomass burning aerosol radiative and climate effects on the amazon rainforest productivity with an earth system model, Atmos. Chem. Phys. 19 (2) (2019) 1301–1326.

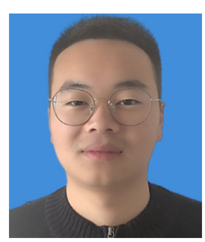
[51] A. Zhu, F. Zhao, P. Liang, C. Qin, Next generation of GIS: must be easy, Ann. GIS (2020) http://dx.doi.org/10.1080/19475683.2019.1670735.

[52] G. Lü, M. Batty, J. Strobl, H. Lin, A. Zhu, M. Chen, Reflections and speculations on the progress in Geographic Information Systems (GIS): a geographic perspective, Int. J. Geogr. Inf. Sci. 33 (2) (2019) 346–367.

[53] Y. Zhu, J. Yang, Automatic data matching for geospatial models: a new paradigm for geospatial data and models sharing, Ann. GIS 25 (4) (2019) 283–298.

Songshan Yue received his Ph.D. degree in Cartography and Geographic Information System in 2016. He is currently an associate professor in the School of Geography at Nanjing Normal University (Nanjing, China) and a member of OpenGMS (Open Geographic Modeling and Simulation) research group. His research interests are in the areas of geographic modeling, geographic information system and virtual geographic environment.

Min Chen received his Ph.D. degree in Cartography and Geographic Information System in 2009. He is currently a professor in the School of Geography at Nanjing Normal University (Nanjing, China) and the team leader of OpenGMS research group. His research interests are in virtual geographic environment, geographic modeling and simulation, and geographic Information System.



Jie Song is currently a M.S. candidate at the School of Geography at Nanjing Normal University (Nanjing, China). His research interests are in the areas of Terrestrial Carbon Cycle Modeling and Distributed Computing.

Wenping Yuan is a profess or in the School of Atmospheric Sciences at Sun Yat-sen University (Guangzhou, China) and a member of the Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies. His research interests are in terrestrial ecological modeling, integrated modeling of terrestrial ecological and hydrological systems.

Tiexi Chen received his Ph.D. degree in Vrije University Amsterdam 2014 and he is currently a professor in the School of Geographical Sciences at Nanjing University of Information Science & Technology. His research interests are in the carbon cycle of large-scale terrestrial ecosystems and the constraints of climate change on ecosystems.

Guonian Lü is Deputy Director of Chinese Academic Committee of Geographical Society, a professor in the School of Geography at Nanjing Normal University (Nanjing, China) and he is also one of the scientific steering member in the OpenGMS research group. His research interests are in the areas of geographic science, cartography, geographic information system and virtual geographic environment.

Chaoran Shen received his M.S. degree in Geographic Information System in 2019. His main research field interests are in the areas of Terrestrial Carbon Cycle Modeling, Web-based model application, Parallel Computing in Geo-Science.

Zaiyang Ma is currently a Ph.D. candidate at the School of Geography at Nanjing Normal University (Nanjing, China). His research interests are in the areas of Participatory Modeling and Distributed Computing.

Kai Xu is currently a Ph.D. candidate at the School of Geography at Nanjing Normal University (Nanjing, China). His research interests are in the areas of Participatory Modeling and Distributed Computing.

Yongning Wen is a profess or in the School of Geography at Nanjing Normal University (Nanjing, China) and member of OpenGMS (Open Geographic Modeling and Simulation) research group. His research interests are in the areas of geographic information system, cartography, and web-based model sharing and integration.

Hongquan Song is an associate professor in the College of Environment and Planning of Henan University and the key Laboratory of Geospatial Technology for the Middle and Lower Yellow River Regions (Ministry of Education). His research interests are in the areas of environmental modeling, video surveillance for geospatial modeling and GIS.