

Research challenges and opportunities for using big data in global change biology

Jiayang Xia¹  | Jing Wang^{1,2}  | Shuli Niu³ 

¹Zhejiang Tiantong Forest Ecosystem National Observation and Research Station, Research Center for Global Change and Ecological Forecasting, School of Ecological and Environmental Sciences, East China Normal University, Shanghai, China

²Shanghai Institute of Pollution Control and Ecological Security, Shanghai, China

³Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China

Correspondence

Jiayang Xia, School of Ecological and Environmental Sciences, East China Normal University, Shanghai 200241, China.
Email: jyxia@des.ecnu.edu.cn

Funding information

National Natural Science Foundation of China, Grant/Award Number: 31722009; National Key R&D Program of China, Grant/Award Number: 2017YFA0604600

Abstract

Global change biology has been entering a big data era due to the vast increase in availability of both environmental and biological data. Big data refers to large data volume, complex data sets, and multiple data sources. The recent use of such big data is improving our understanding of interactions between biological systems and global environmental changes. In this review, we first explore how big data has been analyzed to identify the general patterns of biological responses to global changes at scales from gene to ecosystem. After that, we investigate how observational networks and space-based big data have facilitated the discovery of emergent mechanisms and phenomena on the regional and global scales. Then, we evaluate the predictions of terrestrial biosphere under global changes by big modeling data. Finally, we introduce some methods to extract knowledge from big data, such as meta-analysis, machine learning, traceability analysis, and data assimilation. The big data has opened new research opportunities, especially for developing new data-driven theories for improving biological predictions in Earth system models, tracing global change impacts across different organismic levels, and constructing cyberinfrastructure tools to accelerate the pace of model-data integrations. These efforts will uncork the bottleneck of using big data to understand biological responses and adaptations to future global changes.

KEYWORDS

big data, Earth system model, global change biology, machine learning, model uncertainty

1 | INTRODUCTION

One fundamental question in global change biology is how the biosphere responds and adapts to global environmental changes (Sage, 2020; Shugart & Woodward, 2011). However, understanding interactions between biological systems and global environmental changes remains a major scientific challenge. To tackle this challenge, both biological and environmental data are growing at an exponential rate, promoting global change biology into a big data era. Such big data includes not only observations (Baldocchi, 2020; Schimel, Schneider, & JPL Carbon and Ecosystem Participants,

2019) but also simulation outputs from modern modeling systems (Hoffman et al., 2017; Reichstein et al., 2019). The explosion of big data has boosted new ways of research which are not applicable by traditional studies at the site level. In this paper, we introduce big databases at different organismic levels from gene to biosphere and discuss how the big data has benefited the researches in global change biology.

The concept of big data is commonly characterized by multiple Vs (Nativi et al., 2015), mainly including a large amount of data (Volume), speed of data accumulation (Velocity), heterogeneity of data source and structure (Variety), data quality

and uncertainty (Veracity), effective presentations of complex data (Visualization), and valuable information of data (Value). Transforming data with the first five Vs into Value is the ultimate goal of using big data in all research areas. Massive data sets have been publicly available in both fields of biology and environmental change over recent decades. These data sets range from genomic to global scale. For example, the GenBank database has surpassed one billion sequence records (Sayers et al., 2019), the transmission rate of remote sensing data has exceeded hundreds of terabytes per day (Agapiou, 2017), and the modeling data in the 6th phase of the Climate Model Intercomparison Project (CMIP) are estimated to reach 30 petabytes (Stockhouse & Lautenschlager, 2017). The exponential growth of data has benefited from the technological advances in the past few decades, such as large-scale computing (Liu, 2013), enhanced computational capacity and data storage (Chen, Mao, & Liu, 2014), gene sequencing (Franzosa et al., 2015), and satellite remote sensing (Schimel et al., 2019). Extracting knowledge from these big databases needs not only new analytical techniques but also a new research paradigm in global change biology (Niu, Wang, Wang, Xia, & Yu, 2020). For example, most published and ongoing observational studies in global change biology are driven by scientific hypotheses. In the CMIPs, only a small proportion of the modeling data have ever been analyzed. However, in the big data era, sophisticated algorithms and statistical tools are developed to discover new patterns directly from the data. In the previous decades, analyses on big data have led to fruitful achievements, such as identifications of general response patterns of biological processes to global changes, discoveries of emergent biological patterns and mechanisms, and predictions of biological feedbacks to global climate change.

Over the past decade, we have seen many reviews on the applications of big data and related technology in different subjects. They broadly cover topics on Bayesian statistical learning for biology (Yau & Campbell, 2019), deep learning for geoscience (Reichstein et al., 2019) and bioinformatics (Li, Huang, et al., 2019), data integration in aquatic ecology (Durden, Luo, Alexander, Flanagan, & Grossmann, 2017), data gaps in macroecology (Wüest et al., 2019), and big data analytics for healthcare (Wang, Kung, & Byrd, 2018) and animal movement ecology (Thums et al., 2018). Yet important research challenges and opportunities remain undiscussed for global change biology. In fact, global change biology addresses scientific questions at scales from the gene up to the biosphere, and there is an increasing trend of using big data in recent studies. Thus, the purpose of this paper is to review the research challenges and opportunities for using big data in global change biology. We mainly focus on reporting how big data has been used to (a) identify general patterns of biological responses to global changes based on data from genomic to ecosystem level, (b) discover new emergent biological phenomena and mechanisms on the regional and global scales, and (c) predict biosphere dynamics and their feedback to climate change.

2 | GENERAL PATTERNS OF BIOLOGICAL RESPONSES TO GLOBAL CHANGES FROM DATA ON SCALES FROM GENE TO ECOSYSTEM

2.1 | Genetic adaptation of populations to global environmental changes

Many big databases have been built for archiving genomic data of different organisms on the Earth. The GenBank sequence database, which is maintained by the National Center for Biotechnology Information, now provides nucleotide sequences for >420,000 formally distinct species (Sayers et al., 2019). Biological populations respond to environmental changes in three major ways, including dispersal to unaffected areas, phenotypic plasticity, and genetic adaptation (Gienapp, 2020). Thus, big genomic data offer the opportunity to detect how fast populations can genetically adapt to new environmental conditions, which is critical for maintaining fitness and demographic stability and therefore for avoiding extinction. For example, *Arabidopsis thaliana*, the first plant species whose genome was sequenced in 2000, has shown adaptational lag to a warming climate in common garden experiments at four European sites (Wilczek, Cooper, Korves, & Schmitt, 2014). On the contrary, other experimental evidence has shown fast evolution or genetic adaptation of plants (Franks, 2011; Franks, Sim, & Weis, 2007; Mousavi-Derazmahalleh et al., 2019). In the published studies in *Global Change Biology*, the genomic approach has been mainly used on terrestrial biota (Figure 1). The most investigated global change factors are climate warming and elevated atmospheric CO₂ concentration (Figure 1a). For example, some recent studies have used genomic data to analyze the warming impacts on soil microbial activity in a mountain glacier (Wang et al., 2020), local adaptation in a terrestrial-breeding frog (Cummins, Kennington, Rudin-Bitterli, & Mitchell, 2019), and terrestrial N₂O emissions (Li et al., 2020). It is also clear that most studies have used genomic data to study global change impacts on the ecosystem rather than tissues or organ level (Figure 1).

Global changes can alter selection in both direction and strength, leading to either maladaptation or fast genetic adaptation (Razgour et al., 2019; Siepielski et al., 2017). However, it is still difficult to uncover the molecular mechanisms causing maladaptation or adaptation to environmental changes even when a large volume of genomic data are available. We still lack knowledge about either the particular genes related to each key adaptive trait in most wild species or which type of genetic variation in a gene can elevate the fitness of individuals (McCall & Irwin, 2006; You et al., 2013). Such limitation has been highlighted by a recent review on mapping genetic adaptation of wild populations to global change (Gienapp, 2020). In addition, the change of key adaptive traits can be caused by changes in both genetic composition and epigenetic variation (i.e., phenotypic plasticity; Bock et al., 2015), but genomic data are only effective in detecting genetic variation. These aspects are likely to restrict the applications of genomic techniques in bridging global changes and the corresponding biological responses on the species and population levels.

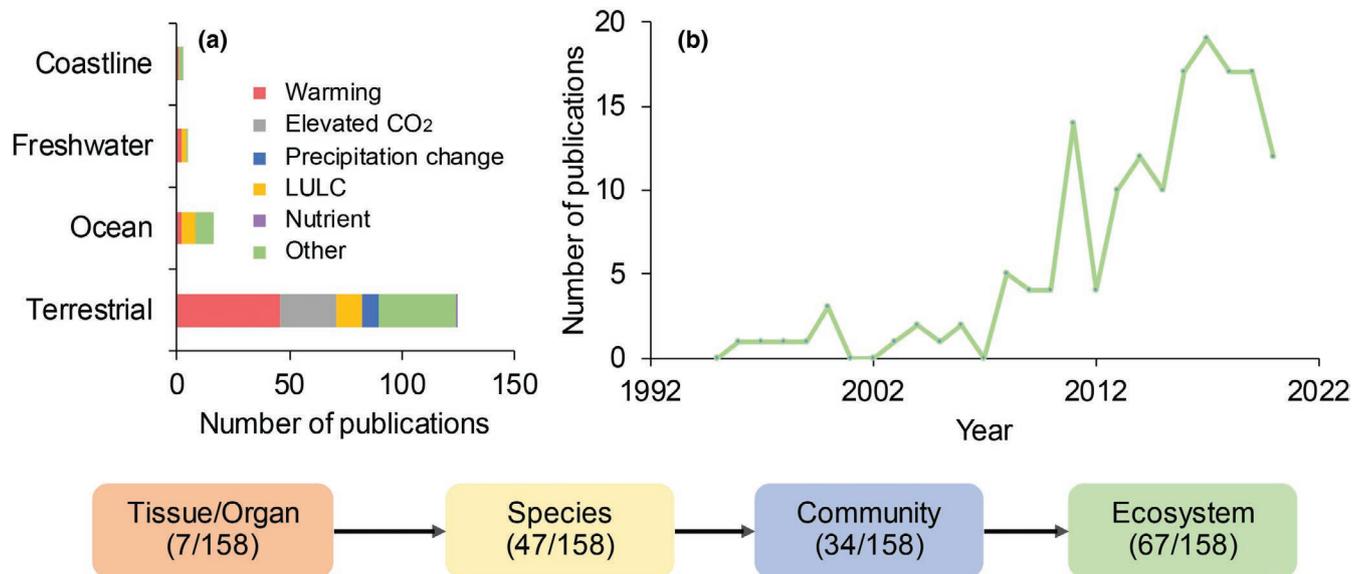


FIGURE 1 A literature survey on the studies published in *Global Change Biology* by using genomic data during 1995–2020. The numbers in the boxes represent the number of studies on different organismic levels. In panel a, the LULC stands for land use and land cover changes, and “other” represents studies on other global change factors. In panel b, the green line indicates the number of published papers based on genomic data. The results were searched with the keyword of “gene”, “genetic” or “genomic”

Nevertheless, global changes have facilitated drastic range expansion of many species, which might display a large range of trait variation over a gradient of environmental factors. It thus provides chances for linking genes with adaptive traits and determining the functional genes under the selection of changing environments (Bay et al., 2018; Nallu et al., 2018). Also, the increasing occurrence of extreme weather and climate events offers experiments by nature to examine the evolutionary responses of species. For example, the European heatwaves have driven adaptive evolutionary shifts in insects (Rodríguez-Trelles, Tarrío, & Santos, 2013) and fishes (Wegner, Kalbe, Milinski, & Reusch, 2008). In addition, whole-genome bisulfite sequencing can be used to evaluate the level of epigenetic variation. When it combines genomic sequencing techniques, we are able to assess the relative roles of genetic and epigenetic variation in fast adaptation to global changes (Hawes et al., 2019). Moreover, new generation sequencing technology has been successfully applied to demonstrate the different impacts of global changes on different taxa, such as soil microbes, trees, and animals (Zhou et al., 2016). These new advances offer important insights into how and why populations respond and adapt to future global environmental changes.

2.2 | Response patterns of species-level traits and processes

The estimate of species richness on the Earth has a large range from 0.5 to 100 million (Caley, Fisher, & Mengersen, 2014; Jiang, 2016). Due to such a high species diversity, the volume of data on species-level traits and processes has been increasing fast in the last decades. For example, trait measurements provide multiple structural and functional characteristics for plant species (Kattge et al., 2020;

Violle et al., 2007). The global plant trait database (TRY) was released in 2007 (Kattge et al., 2011) and has reached 11.8 million trait records in its fifth version (Kattge et al., 2020). The Botanical Information and Ecology Network (version 4.1.1) integrates >20.6 million records on plant distribution, abundance, and traits for 485,902 species on the globe (Enquist, Condit, Peet, Schildhauer, & Thiers, 2016). Phenology is another example of the rapid increase in data volume on the species level. The Pan European Phenology project has monitored >12 million phenological records over 265 plant species in Europe since 1868 (Templ et al., 2018). The USA National Phenology Network also maintains >10 million records of plant and animal phenology observations across the United States (Crimmins, Crimmins, Gerst, Rosemartin, & Weltzin, 2017). The large volume of species-level data has great value in improving our understanding of biological responses to global changes. The TRY database, for instance, has been widely used for exploring the impacts of global change factors on plant diversity (Harrison, Spasojevic, & Li, 2020), physiology (De Kauwe et al., 2016), phenology (König et al., 2018), and biomass (Shao et al., 2019).

Associated with the rapid increases in volume and value, most species-level data sets have encountered variety and veracity issues. The variety of species-level data is usually induced by large heterogeneity of growth conditions or measuring approaches. For example, plant and animal phenology have been monitored in both natural and urban regions (Harrison, Gibbs, & Winfree, 2019; Li, Stucky, et al., 2019; Wohlfahrt, Tomelleri, & Hammerle, 2019). Also, there are large differences in species-level biomass (Poorter et al., 2016) and its response to environmental changes (Xu, Yan, et al., 2019a) between controlled and field experiments. Plant photosynthesis is one of the most important processes in driving biological responses to global changes but has been measured by different portable

photosynthesis systems, such as LI 6400 and LI 6800 (Li-Cor), GFS-3000 (Walz), CID-CI340 (Bio-Science), CIRAS-3 (PP Systems) and LCI-SD (ADC BioScientific). Although there are some standard manuals or protocols to measure species traits or processes (e.g., Hendry & Grime, 1993), the uneven veracity or quality of specie-level data is another challenge for analyzing data of different species across space. In the TRY database, for instance, a large proportion of leaf traits data could be measured in the shade rather than in full sun as reported by the data contributors (Keenan & Niinemets, 2016).

The species-level big data provides the opportunity to study the linkages between plant traits and ecosystem functions (Fraser, 2020). In fact, plant functional traits are important contributors to the simulation uncertainty of ecosystem productivity in current terrestrial biosphere models (Cui et al., 2019; Xia et al., 2017). A promising approach is to collect environmental data and measure multiple ecosystem functions associated with the species traits. Some recent studies have constructed data sets of trait-function combinations for multiple species based on the measurements in global change experiments

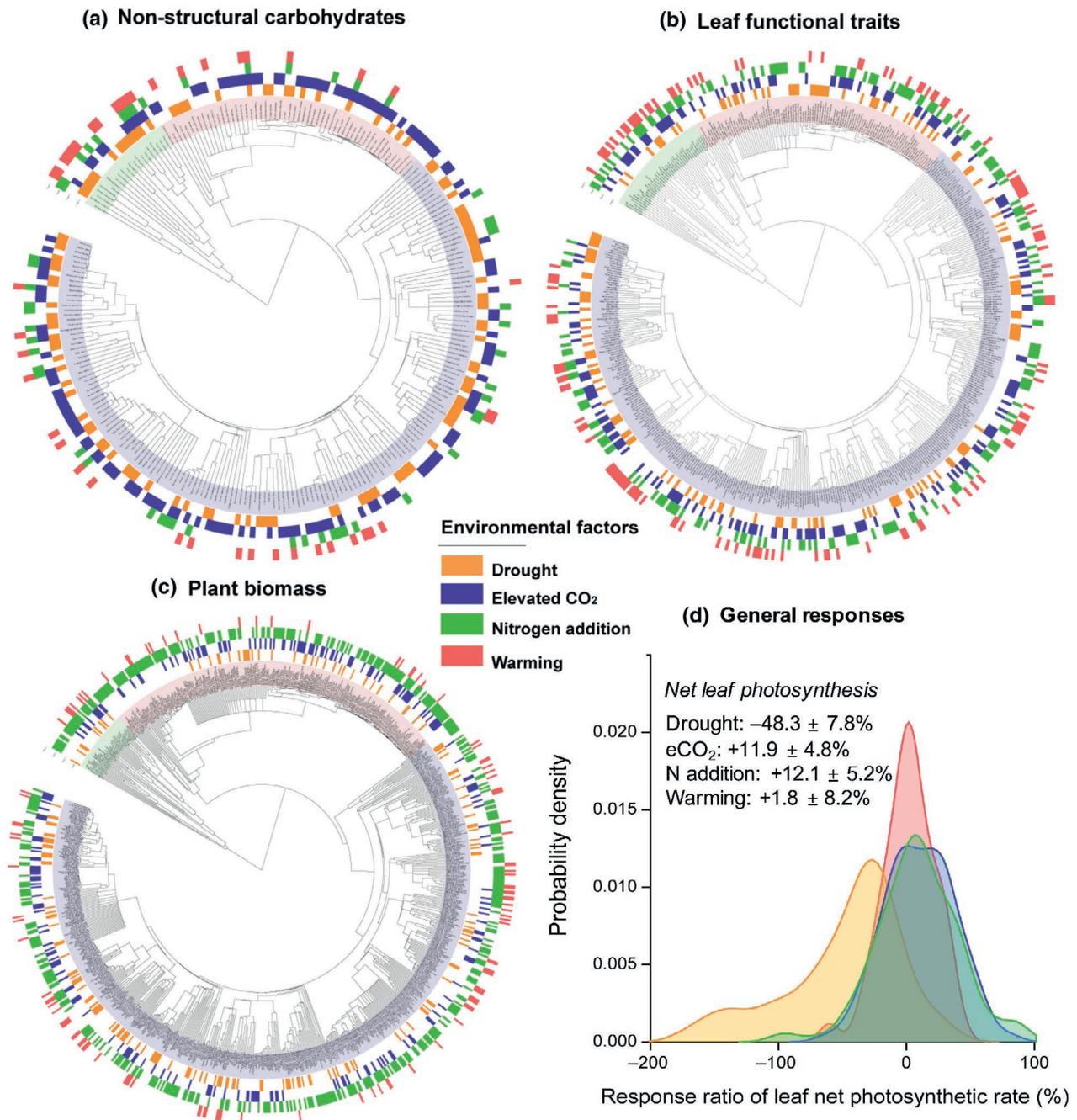


FIGURE 2 The phylogenetic trees of plant species in manipulative experiments that include observations of non-structural carbohydrates (a), leaf functional traits (b) and plant biomass (c). Rings surrounding the phylogenetic tree from inside outwards indicate species from drought (orange bars), elevated CO₂ (blue bars), nitrogen addition (green bars) and warming (red bars) experiments. (d) Probability density of experimental responses of net leaf photosynthesis to different global change factors

(Walker et al., 2014). As shown in Figure 2, hundreds of plant species have been studied in different global change experiments (Figure 2a–c). Responses of plant biomass and leaf functional traits are mostly measured under nitrogen addition and elevated CO₂, while changes in non-structural carbohydrates are more often studied under elevated CO₂ and drought stress (Figure 2a–c). As shown in Figure 2d, net leaf photosynthesis of terrestrial plants is significantly enhanced by elevated CO₂ and nitrogen addition but reduced by drought. Although the response directions of leaf traits are divergent among species, a recent analysis on 515 species has shown robust leaf trait relationships across species under global environmental changes (Cui, Bian, et al., 2020; Ciu, Weng, et al., 2020). Another analysis on root traits from 369 plant species has revealed that root diameter has the strongest impact on root trait variation across species (Ma et al., 2018). The analyses on root traits also have emphasized the important role of mycorrhizal pathways in regulating ecosystem functions under environmental changes (Ma et al., 2018; Soudzilovskaia et al., 2015). Overall, the data mining on species-level observations has facilitated the process-understanding of biological responses to global changes, such as plant non-structural carbohydrates (Adams et al., 2017; Du, Lu, & Xia, 2020), leaf photosynthetic capacity (Kattge, Knorr, Raddatz, & Wirth, 2009) and respiration (Atkin et al., 2015), plant biomass (Xia & Wan, 2008; Xu, Lu, et al., 2019b), and soil microbial community composition (De Vries et al., 2012).

2.3 | Changes in community structures and ecosystem functions

Ecosystems are complex systems including not only biotic and abiotic components but also their interactions (Blew, 1996). It means community structures and ecosystem functions are not the mere summations of individual organisms or processes (Tansley, 1935). Thus, ecosystem responses to global change cannot be characterized or defined by data from a single process, location, or time (LaDeau, Han, Rosi-Marshall, & Weathers, 2017). The data volume has been increasing dramatically in ecosystem-level global change experiments during the past three decades. However, there are some challenges in analyzing the big data from different ecosystems for global response patterns, such as the strong dependence of response patterns upon the sample size of data (van Gestel et al., 2018), the limited data in semi-arid ecosystems, tropical and subtropical forests, and Arctic tundra (Song et al., 2019), and the under-representation of extreme events in global change experiments (Knapp et al., 2015, 2017). Nevertheless, analyzing the data from those ecosystem-level experiments has revealed some general response patterns of community structure and ecosystem functions to global changes. For example, climate warming stimulates biomass accumulation (Lin, Xia, & Wan, 2010) but leads to soil carbon loss (Crowther et al., 2016) without altering the temperature sensitivity of soil respiration (Carey et al., 2016). Elevated CO₂ enhances net accumulations of carbon in terrestrial ecosystems (Luo, Hui, & Zhang, 2006) but could trigger nitrogen limitation on ecosystem productivity (Liang, Qi, Souza, & Luo, 2016; Luo et al., 2004). There

are also many general response patterns of community structures and ecosystem functions to other global change factors, such as precipitation change (Wu, Dijkstra, Koch, Penuelas, & Hungate, 2011), nutrient addition (Midolo et al., 2018; Zheng, Zhou, Luo, Zhao, & Mo, 2019), land use change (Beckmann et al., 2019; Zhou et al., 2019), and biological invasions (Davidson, Cott, Devaney, & Simkanin, 2018; Liao et al., 2008). A recent meta-analysis of 1,119 global change experiments (Song et al., 2019) has shown that the treatment magnitudes of warming and elevated CO₂ in those experiments are consistent with the future projections by Earth system models, but those of nitrogen and precipitation manipulations are often different from model projections. Also, most of these meta-analyses have detected additive rather than non-additive impacts between global change factors on ecosystem functions (e.g., Song et al., 2019; Yue et al., 2017). However, it should be noted that the lack of interactive effects between global change factors could partially result from the limited number of multi-factor experiments and the small range of treatment magnitude, such as similar temperature increase in warming experiments.

Associated with the rapid increase in global meta-analyses or syntheses on general response patterns, data also become rich within a site or experiment. For example, the SPRUCE experiment (<https://mnspruce.ornl.gov>) measures a large number of ecosystem processes or state variables in a northern peatland ecosystem, such as vegetation phenology (Richardson et al., 2018) and productivity (Walker et al., 2017), soil mycorrhizal dynamics (Fernandez, Heckman, Kolka, & Kennedy, 2019), greenhouse gases (Hanson et al., 2016), and element isotopes (Hobbie et al., 2017). Another example is a long-term field warming experiment in the tallgrass prairie in the U.S. Great Plains, which has discovered numerous novel phenomena of carbon cycling in response to climate warming based on measurements of microbial functional genes (Xue et al., 2016), soil respiration (Luo, Wan, Hui, & Wallace, 2001), plant phenology (Sherry et al., 2007) and species composition (Shi et al., 2016), and ecosystem CO₂ exchanges (Niu, Sherry, Zhou, & Luo, 2013). In each of the seventy forest sites in the Forest Global Earth Observatory (ForestGEO; Anderson-Teixeira et al., 2014), growth and survival of all individuals have been monitored. Figure 3 shows the complexity of carbon cycling in a forest ecosystem and the ground-based measurements on multiple layers of biological and environmental factors in a 20 ha ForestGEO site. Spatial analyses on these data sets have shown that topographical factors influence storages of aboveground litter mass (Wang et al., 2019) and soil phosphorus (Qiao et al., 2020), while soil acidity is a key factor regulating soil carbon and nitrogen stocks (Qiao et al., 2020). In aquatic ecosystems, a large-scale comparative study of 69 boreal lakes has shown that surface water CO₂ concentration enhances phytoplankton production but not community composition (Vogt, St-Gelais, Bogard, Beisner, & del Giorgio, 2017). Thus, the big database on the site level can improve our understanding of not only community structure and species diversity, but also environmental impacts on ecosystem processes. It also provides new opportunities to link ground-based community data to ecosystem-level measurements of carbon and water fluxes and even space-based monitoring of ecosystems.

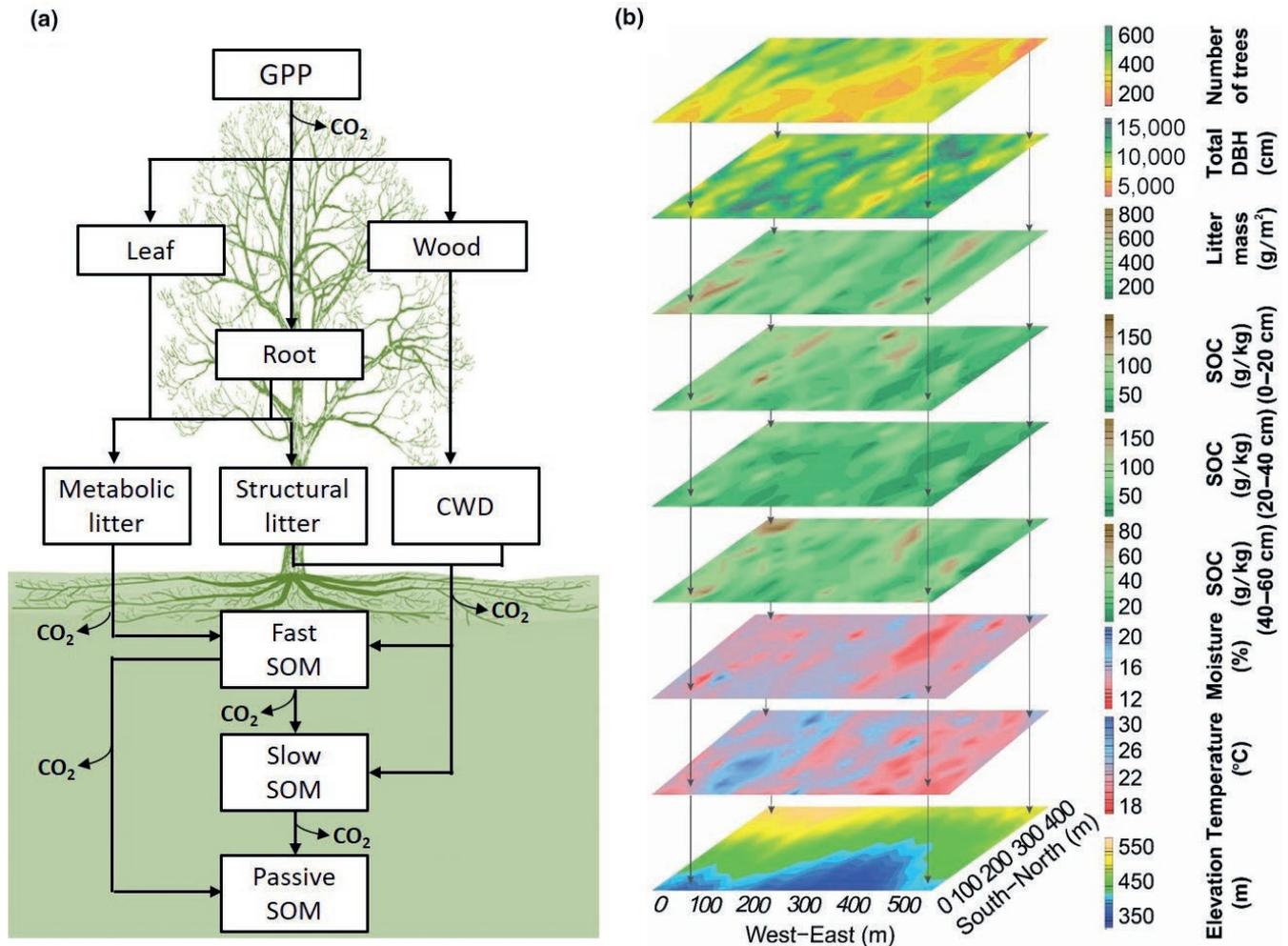


FIGURE 3 Using big data on the site level to improve process-based understanding of ecosystem functions. (a) A conceptual illustration of the carbon cycle in a forest ecosystem. The conceptual model consists of different carbon pools and the carbon in those pools interact via fluxes of carbon. (b) The hierarchical structure of variables aboveground and belowground including tree number, diameter breast at 1.3 m height (DBH), litter mass, soil organic carbon (SOC) in different layers, topography (i.e., elevation), soil microclimate (i.e., soil moisture and temperature) in the 20 ha Tiantongshan forest dynamics plot (29.8°N, 121.8°E). Specifically, the DBH is the sum of all trees in each subplot. Soil temperature was detected every half hour by the temperature data loggers

3 | EMERGENT BIOLOGICAL MECHANISMS AND PHENOMENA BASED ON REGIONAL AND GLOBAL OBSERVATIONS

3.1 | Discovery of biological properties and mechanisms based on observational networks

The rapid increases in ecosystem data volume and velocity are associated with rapidly expanding observational networks like FLUXNET (<http://fluxnet.ornl.gov>). Since the late 1990s, the FLUXNET database has greatly improved our understanding of global change biology (Baldocchi, 2020). In the most recent public version of the FLUXNET database (i.e., FLUXNET2015), there are about 2000 site-years of data at 30 min intervals from >200 sites (Pastorello et al., 2017). As shown by the clustering analysis in Figure 4, measurements of eddy-flux towers have been used to explore the

responses of various terrestrial carbon and water processes to different global change factors, such as climate warming (Anderson-Teixeira et al., 2011), precipitation change (Schwalm et al., 2012; van der Molen et al., 2011), and extreme climate events (Gu et al., 2008; Reichstein et al., 2013). The high veracity of FLUXNET data has enabled the quantifications of the factors contributing to the interannual variability of CO₂ fluxes from site to global scales (Mahecha et al., 2010; Niu et al., 2017; Shao et al., 2015). Cross-site analyses of flux data have revealed some important emergent properties of ecosystem CO₂ exchange between the land and atmosphere. For example, Churkina, Schimel, Braswell, and Xiao (2005) have discovered the convergent ratio between mean daily CO₂ exchange over carbon uptake and release periods as 2.73 ± 1.08 across 28 sites (Churkina et al., 2005). This ratio is relatively invariant even with a larger sample size of flux sites (2.71 ± 1.61 ; Cui, Bian, et al., 2020; Ciu, Weng, et al., 2020). It is also interesting that the total annual net ecosystem CO₂ exchange at each site can be described

Nutrient Network (<https://nutnet.org/>) have revealed some new mechanisms which mediate the response of grassland community to nutrient addition, such as weakened stabilizing effects of diversity by eutrophication (Hautier et al., 2014) and the reduced grassland diversity by adding multiple limiting nutrients (Harpole et al., 2016). As shown in Table S1, there are many observational networks which provide big databases for community structures and ecosystem functions in some specific regions or across the globe. For example, the PhenoCam dataset provides time series of vegetation phenological observations for 393 sites across the globe (Seyednasrollah et al., 2019). Such camera-based big data has also promoted aquatic ecology by providing thousands of seabed photographs (Morris et al., 2014) and plankton images (Luo et al., 2018). Some other databases collect global measurements of specific ecosystem processes or states and lead to discoveries of new global trends. For example, the global data of soil respiration has shown a global increase in soil respiration over recent decades (Bond-Lamberty, 2018; Bond-Lamberty & Thomson, 2010). Soil nitrogen-related datasets have shown regional or global patterns of inorganic nitrogen stock (Wei et al., 2019), nitrogen mineralization (Li, Tian, et al., 2019), and nitrification rate (Li et al., 2020). Global tree-ring data has allowed the analyses on long-term dynamics of tree growth under global changes (e.g., Zhu, Cui, & Xia, 2020). There are also some top-down organized continental-scale observatory networks, such as the National Ecological Observatory Network in the continental United States (Kampe, Johnson, Kuester, & Keller, 2010), the Terrestrial Ecosystem Research Network in Australia (Cleverly et al., 2019), and the Chinese Ecosystem Research Network (Yu & Yu, 2013). These networks offer a rich set of measurement data with improved data standardization. Overall, the regional and global data sets are useful for promoting the discoveries of new emergent mechanisms and ecosystem properties in the biosphere.

3.2 | Global phenomena of biological changes revealed by big remote-sensing data

Space-based observations of the Earth's surface have led to a steady increase in global data sets for terrestrial (Schimel et al., 2019) and marine (Muller-Karger et al., 2018) biospheres. In comparison with the ground-based observations, remote sensors provide a massive volume of data with high velocity. The vast data volume results from the high resolutions of remote measurements at both temporal and spatial scales. For example, many satellites produce data at a spatial resolution of 10 m or even finer, such as SPOT 6/7 (Astrium Services, 2013), IKONOS (Dial, Bowen, Gerlach, Grodecki, & Oleszczuk, 2003), QuickBird (Leboeuf et al., 2007), and ZY-3 (Huang et al., 2020). Some environmental and biological processes can be monitored on the daily time step, such as fire occurrence by Terra/Aqua MODIS (Langner & Siegert, 2009) and vegetation greenness by NOAA-AVHRR (Pu et al., 2007). The role of space-based observations in facilitating studies of global ecology has recently been

systematically reviewed by Schimel et al. (2019). Schimel et al. (2019) also have pointed out three key challenges in using space-based data, including estimation uncertainty from errors of the individual satellite, algorithm uncertainty, and difficulties to integrate biological measurements from genes to ecosystems. These challenges have been existing for the last three decades, and we could learn from the recently successful big data mining on remote-sensing data in some other fields, such as water and energy budgets (Liang, Wang, He, & Yu, 2019; Sheffield, Ferguson, Troy, Wood, & McCabe, 2009) and precision agriculture (Maes & Steppe, 2019). Here, we highlight that data variety and veracity are still limiting the use of space-based observations to answer more questions in global change biology. Take terrestrial GPP (Gross primary production) as an example, we show that the estimate of total global GPP has a large range, varying from 112.1 ± 0.87 Pg C/year in MODIS to 135.11 ± 2.19 Pg C/year in GOSIF over 2001–2010. Such disagreement could stem from different types of quantitative approaches or different algorithms within each approach. For example, the FLUXCOM and MTT databases are derived from machine learning approaches (Jung et al., 2011, 2017), but the VPM and MODIS data are based on the framework of light use efficiency (Running, 2015; Ryu, Berry, & Baldocchi, 2019; Zhang et al., 2017).

Analyses on the space-based global data sets have discovered many new patterns or shifting trends in the biosphere under global changes. These global patterns or trends can be divided into three types, including long-term gradual trends, cyclic or seasonal changes, and abrupt changes in environmental and biological states. On the global scale, the space-based data sets of vegetation index have revealed the global land greening since 1980s (Piao et al., 2020; Zhu et al., 2016) but browning in many regions since 1990s (Pan et al., 2018; Treharne, Bjerke, Tømmervik, Stendardi, & Phoenix, 2019). The big remote-sensing data has also led to the detection of many other global long-term patterns, such as enhancing peak vegetation growth (Huang, Xia, et al., 2018), weakening temperature controls on vegetation activity (Piao et al., 2014, 2017), asymmetric impacts of day and night warming on vegetation and ecosystem productivity (Peng et al., 2013; Xia et al., 2014), and forest mortality and loss (Curtis, Slay, Harris, Tyukavina, & Hansen, 2018; Yang et al., 2018). On the seasonal scale, global remote-sensing data sets have been widely used to study the cyclic changes in different biological processes, such as seasonal area use by lions (Hemson et al., 2005), phytoplankton blooms (Gittings et al., 2017), the timing of peak photosynthetic activity (Park et al., 2019), and plant phenology (Li et al., 2018; Piao et al., 2019). The high velocity of space-based data enables using satellite imagery to detect the abrupt occurrence of global change events and their biological impacts over large areas (Verbesselt, Zeileis, & Herold, 2012). Recently, some near real-time monitoring systems based on space-based data have been applied to forest conservation (Musinsky et al., 2018; Pratihast et al., 2016), flood event (Van Ackere et al., 2019), fire mapping (Pulvirenti et al., 2020), and tree mortality due to insect outbreak (He, Chen, Potter, & Meentemeyer, 2019; Olsson,

Lindström, & Eklundh, 2016). Furthermore, global analyses of the big remote-sensing data have revealed many emergent properties of ecosystems, such as the average optimum air temperature for ecosystem gross primary productivity (Huang, Piao, et al., 2019), high stability of evergreen broadleaf forests (Huang & Xia, 2019) and collapse of rain-use efficiency in semi-arid ecosystems (Du et al., 2018) under extreme droughts, diminishment of vegetation seasonality over northern lands (Xu et al., 2013), and constrained tropical photosynthetic seasonality by hydroclimate (Guan et al., 2015).

4 | PREDICTIONS OF BIOLOGICAL RESPONSES TO GLOBAL CHANGES BY BIG MODELING DATA

In the past decades, Earth system models have been incorporating hundreds or thousands of biological processes and their interactions with environmental changes. The Coupled Model Intercomparison Project (CMIP), for instance, was launched in 1995 and is currently in the sixth phase (i.e., CMIP6; Eyring et al., 2016). Its data volume of modeling outputs was about 36 TB in CMIP3, then quickly

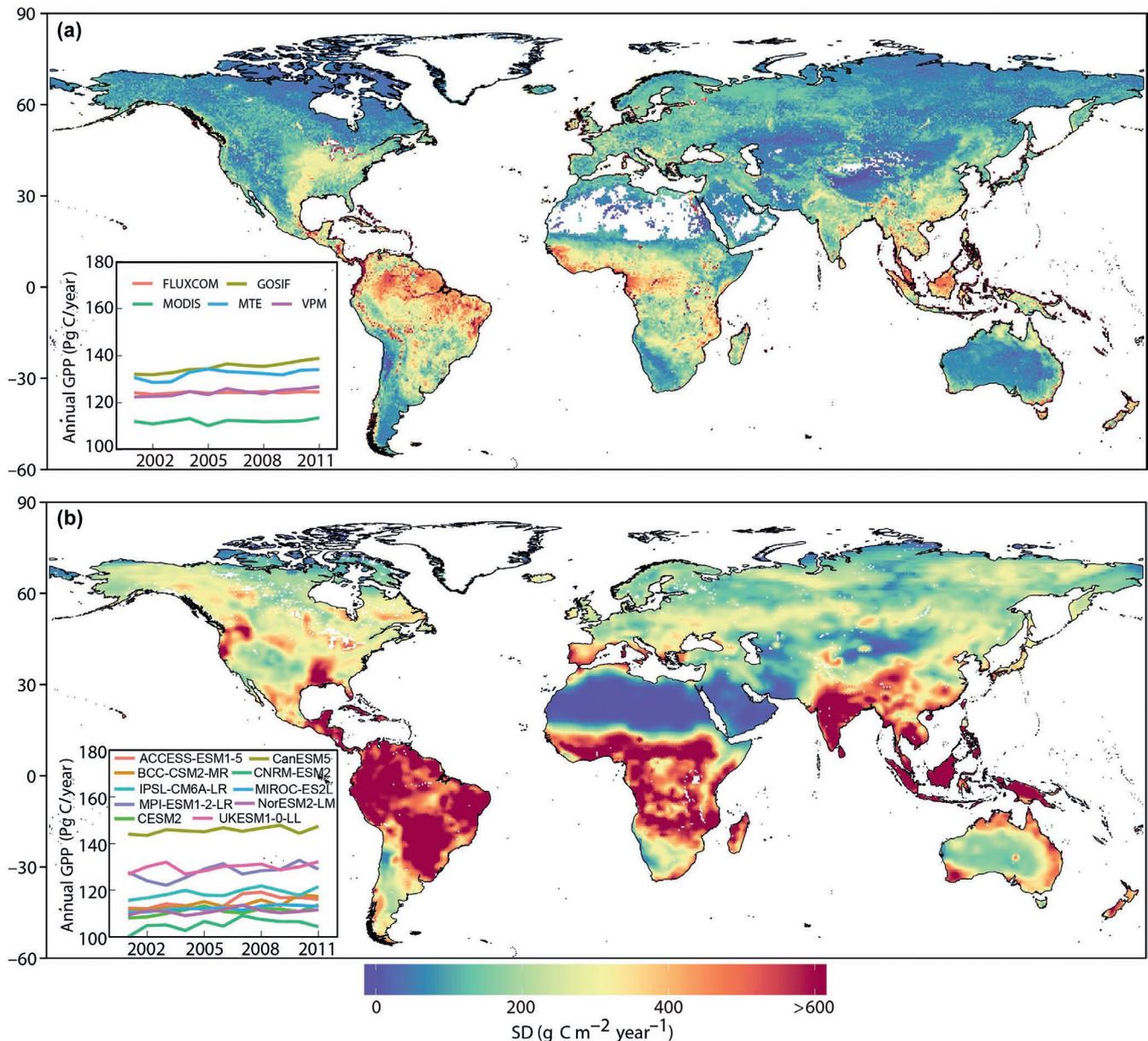


FIGURE 5 Large variations in annual GPP estimations from observation-based datasets (a) and CMIP6 model ensemble (b) over the period 2001–2011. For the upper panel, map shows the standard deviation (SD) of mean annual GPP (2001–2011) among five GPP datasets. The five observation-based GPP datasets are FLUXCOM, GOSIF, MODIS, MTE and VPM. Insert panels show the differences of total global GPP among the five GPP datasets. For the lower panel, map represents the SD of mean annual GPP (2001–2011) among the 10 individual CMIP6 models. The CMIP6 models used in this study are ACCESS-ESM1-5, BCC-CSM2-MR, CanESM5, CESM2, CNRM-ESM2, IPSL-CM6A-LR, MIROC-ES2L, MPI-ESM1-2-LR, NorESM2-LM and UKESM1-0-LL. Insert panels show the differences of total global GPP simulations among the 10 models. All the GPP datasets are calculated at a spatial resolution of $0.5^\circ \times 0.5^\circ$

increased to 3 PB in CMIP5 and 30 PB in CMIP6 (Stockhause & Lautenschlager, 2017), and is expected to reach about 350 PB in 2030 (Overpeck, Meehl, Bony, & Easterling, 2011). Some other projects concentrate on modeling data of terrestrial ecosystems and global changes, such as the MsTMIP (Multi-scale Synthesis and Terrestrial Model Intercomparison Project; <https://nacp.ornl.gov/MsTMIP.shtml>), TRENDY (Trends in net land-atmosphere carbon exchanges; <http://dgvm.ceh.ac.uk/node/21/>), and model integration group of Permafrost Carbon Network (<http://www.permafrostcarbon.org>).

A well-known challenge for using the modeling data is the high disagreement among different models. For example, the disagreement in the estimated global annual GPP is much larger between process-based models (Figure 5b) than data-driven approaches (Figure 5a). As shown in Figure 6, over the period of 1850–2005, CMIP6 model ensemble simulated smaller and more converged global CO₂ fluxes, in comparison with these from CMIP5. At the end of the simulated period, global GPP in CMIP5 was 153.77 ± 41.90 Pg C/year, but decreased to 125.93 ± 13.76 Pg C/year in CMIP6 (Figure 6a). Similar patterns were also found for ecosystem respiration, which

was less uncertain in CMIP6 (118.53 ± 16.21 Pg C/year) than CMIP5 (144.22 ± 38.44 Pg C/year; Figure 6b). However, it should be noted that net ecosystem production is highly uncertain in both CMIP5 (9.44 ± 12.21 Pg C/year) and CMIP6 (7.39 ± 5.82 Pg C/year; Figure 6c), especially the uncertainty of NEP in evergreen broadleaf forests has been amplified from CMIP5 (0.12 ± 0.17 kg C m⁻² year⁻¹) to CMIP6 (0.15 ± 0.23 kg C m⁻² year⁻¹; Figure 6f). The uncertainty of modeling data on the terrestrial carbon cycle could stem from process parameterizations (Luo & Schuur, 2020), model structure (Rafique et al., 2016), or climate forcings (Ahlström, Schurgers, & Smith, 2017; Bonan et al., 2019). As the model uncertainty contributes to about 80% of the total variance of the land carbon cycle in Earth system models (Bonan & Doney, 2018), the divergence among modeling data has promoted research interests in how to trace and reduce model uncertainty (Bonan et al., 2019; Xia, Luo, Wang, & Hararuk, 2013).

The rapid expansion in modeling data size largely results from the increasing model complexity, which is characterized by adding more and more processes into the Earth system models. For example, due to the widespread nitrogen limitation on terrestrial carbon

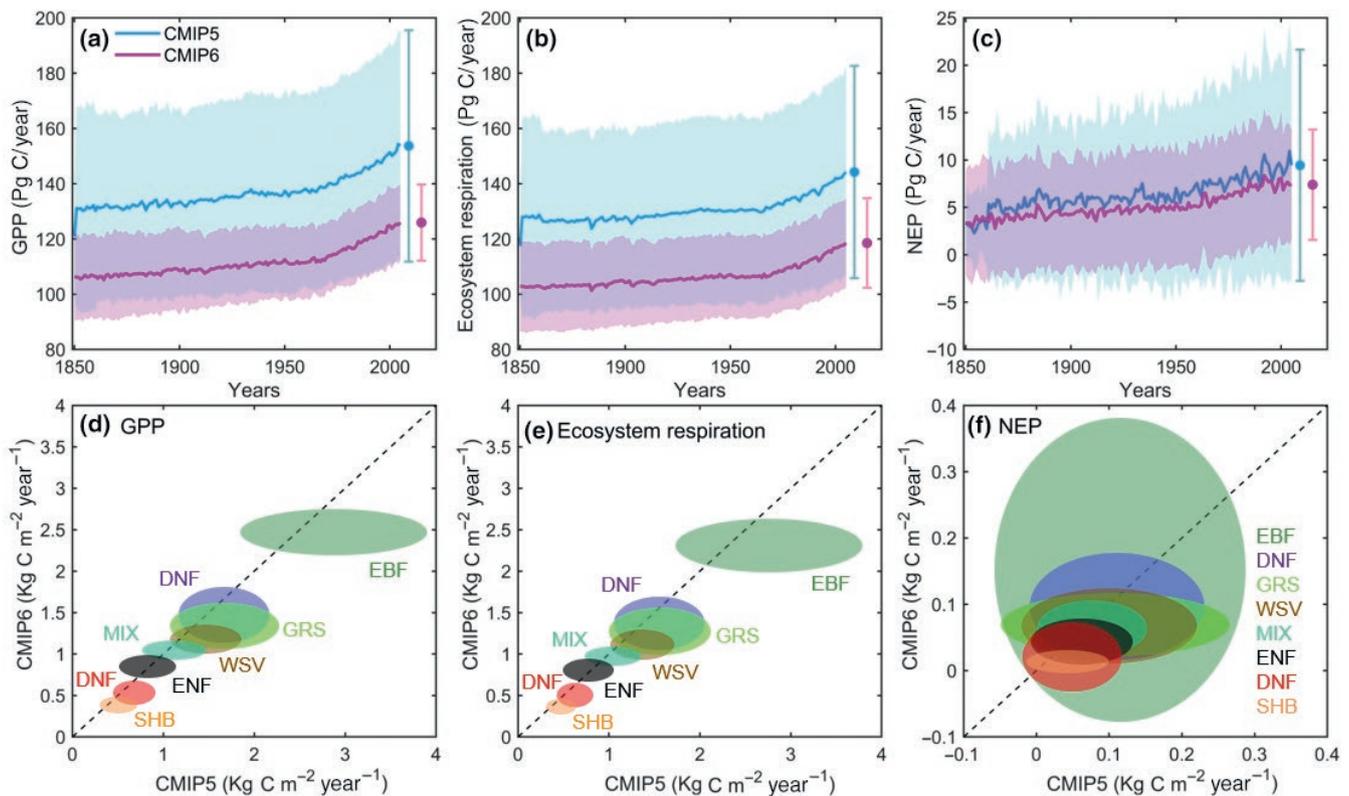


FIGURE 6 Global (a) Gross primary production (GPP), (b) ecosystem respiration, and (c) Net ecosystem production (NEP) modeled from CMIP5 model ensemble in comparison with CMIP6 results over the historical period (1850–2005). In these panels, solid line represents model ensemble mean and shaded polygon represents standard deviation. Error bar corresponds to the model ensemble mean and the standard deviation at the end of the period. Biome-level comparisons between CMIP5 and CMIP6 in modeling (d) GPP, (e) ecosystem respiration, and (f) NEP are shown in the second row. Different shapes represent model ensemble means, and error bar shows standard deviation. DBF, deciduous broadleaf forest; DNF, deciduous needleleaf forest; EBF, evergreen broadleaf forest; ENF, evergreen needleleaf forest; GRS, grassland; MIX, mixed forest; SHB, shrublands; SVN, savannas; WSV, woody savannas. The list of the Earth system models is shown in Table S2. The models were selected based on whether key variables (including GPP, heterotrophic respiration, autotrophic respiration, and NEP) over 1850–2005 were provided. For the models from the same modeling center, we only selected one as the representative because Earth system models (ESMs) within a center produced very similar results. Totally, there were 11 models selected from both of CMIP5 and CMIP6 (Table S2)

storage (Hungate, Dukes, Shaw, Luo, & Field, 2003; Luo et al., 2004; Meyerholt, Sickel, & Zaehle, 2020; Wieder, Cleveland, Smith, & Todd-Brown, 2015), many Earth system models in CMIP6 have implemented nitrogen cycles (Arora et al., 2019; Davies-Barnard et al., 2020). Representation of the nitrogen limitation could be helpful to reduce the deviations of modeling data on the terrestrial carbon cycle, but increases the difficulty to understand the model uncertainty. For example, terrestrial biosphere models in the CMIP6 (e.g., Figure 6) can be separated into two categories as carbon-only and carbon–nitrogen coupled models. However, it is difficult to evaluate whether the difference between carbon–nitrogen coupled models is caused by the nitrogen cycle or propagated from the carbon cycle. An earlier modeling analysis has shown that the implementation of the nitrogen cycle has contrasting effects on ecosystem carbon residence time between woody and nonwoody biomes (Xia et al., 2013). Different representations of nitrogen processes also can lead to heterogeneous modeling data for a given carbon process (Du et al., 2018; Meyerholt et al., 2020). Thus, we can expect a fast increase in the volume of modeling data but a reduction in the traceability of model uncertainty on different biological processes.

Despite the high uncertainty among models, the analyses on these modeling datasets have greatly improved our understanding of biogeochemical and biophysical feedbacks between the terrestrial biosphere and global environmental changes. For example, the modeling data from the TRENDY project have contributed to the annual reports of the global carbon budget (Friedlingstein et al., 2019; Le Quéré et al., 2018), which further provide datasets for carbon sources and sinks on a global level. Another important application of these modeling data is the attributions of various global biological phenomena to their driving environmental factors. For example, the modeling data in the TRENDY project has revealed that CO₂ fertilization effects explain 70% of the global land greening trend during the past decades (Zhu et al., 2016). An analysis based on modeling data from the MsTMIP project has shown the important contributions of rising atmospheric CO₂ concentration and climate change to the enhancing peak rate of global GPP since the 1980s (Huang, Xia, et al., 2018). Furthermore, the big modeling data have provided quantitative predictions of key biological processes or functions under future climate scenarios, such as CO₂ fertilization effects on net land carbon sink (Haverd et al., 2020; Liu et al., 2019).

5 | METHODS TO EXTRACT KNOWLEDGE FROM BIG DATA

The explosion of observational and modeling data challenges traditional methods of data analysis. In global change biology, experimental designs or field surveys in traditional studies are commonly driven by specific biological questions (Luo et al., 2011), so data are collected with certain statistical assumptions such as normality, randomization, and independence (Queen, Quinn, & Keough, 2002). Big data approaches, however, extract emerging knowledge from

unstructured data mainly based on machine learning techniques. The difference between traditional and big data approaches has been compared in many fields, such as geoscience (Reichstein et al., 2019), social economics (Blazquez & Domenech, 2018), and healthcare (Wang et al., 2018). Here, we only briefly review some widely used or newly emerged methods in the field of global change biology, especially on their challenges and opportunities for facilitating the integration between observed and simulated big data in the future.

5.1 | Meta-analyses on experimental data

The term “meta-analysis” was first introduced in 1976 (Glass, 1976), and has been identified as the “grandmother of the big data” in a recent review (Gurevitch, Koricheva, Nakagawa, & Stewart, 2018). As a statistical method, meta-analysis is inherently different from a big data approach, which is developed from computer science (Fan, Han, & Liu, 2014; Gurevitch et al., 2018). In global change biology, meta-analyses are usually designed to examine the hypothesized effectiveness of global-change factors (Luo et al., 2006; Song et al., 2019; Xia & Wan, 2008) or to identify sources of a general phenomenon from highly heterogeneous observations (e.g., Cardinale et al., 2007). As shown in Figure 7, there are a total of 192 meta-analysis papers published in the journal of *Global Change Biology* over 1995–2020. Besides the exponential increase in the number of publications, at least three trends have emerged in the meta-analytical studies. First, the number of meta-analysis exploring the interactive effects of different global change factors has increased fast from 1995–2005 (9% in total cases) to 2015–2020 (17%). Second, among the single global change factors, many meta-analyses have focused on the impacts of land use and land cover changes while nutrient availability has become a hot topic in recent years. Third, the sample size of meta-analytical analysis has been growing fast from <100 (Guo & Gifford, 2002; Peterson et al., 1999) to >1,000 in a recent case (Song et al., 2019). However, cautions should be exercised in using meta-analysis to derive quantitative conclusions (Whittaker, 2010), even recently they are combined with other statistical approaches such as structural equation models. Unlike meta-analysis, big data approaches have different goals as discovering subtle patterns and heterogeneities from a massive sample size of data (Fan et al., 2014). Thus, it is promising to apply big data techniques to these existing meta-analytical datasets for extracting new patterns of biological response to global changes.

5.2 | Machine learning for observational data

Machine learning, particularly deep learning, has been increasingly used to discover new patterns from data. For example, machine learning has benefited the areas of plant specialized metabolism (Moore et al., 2019), plant–pathogen interactions (Sperschneider, 2019), biomass estimation (Anderson et al., 2018; Zhao, Popescu, Meng,

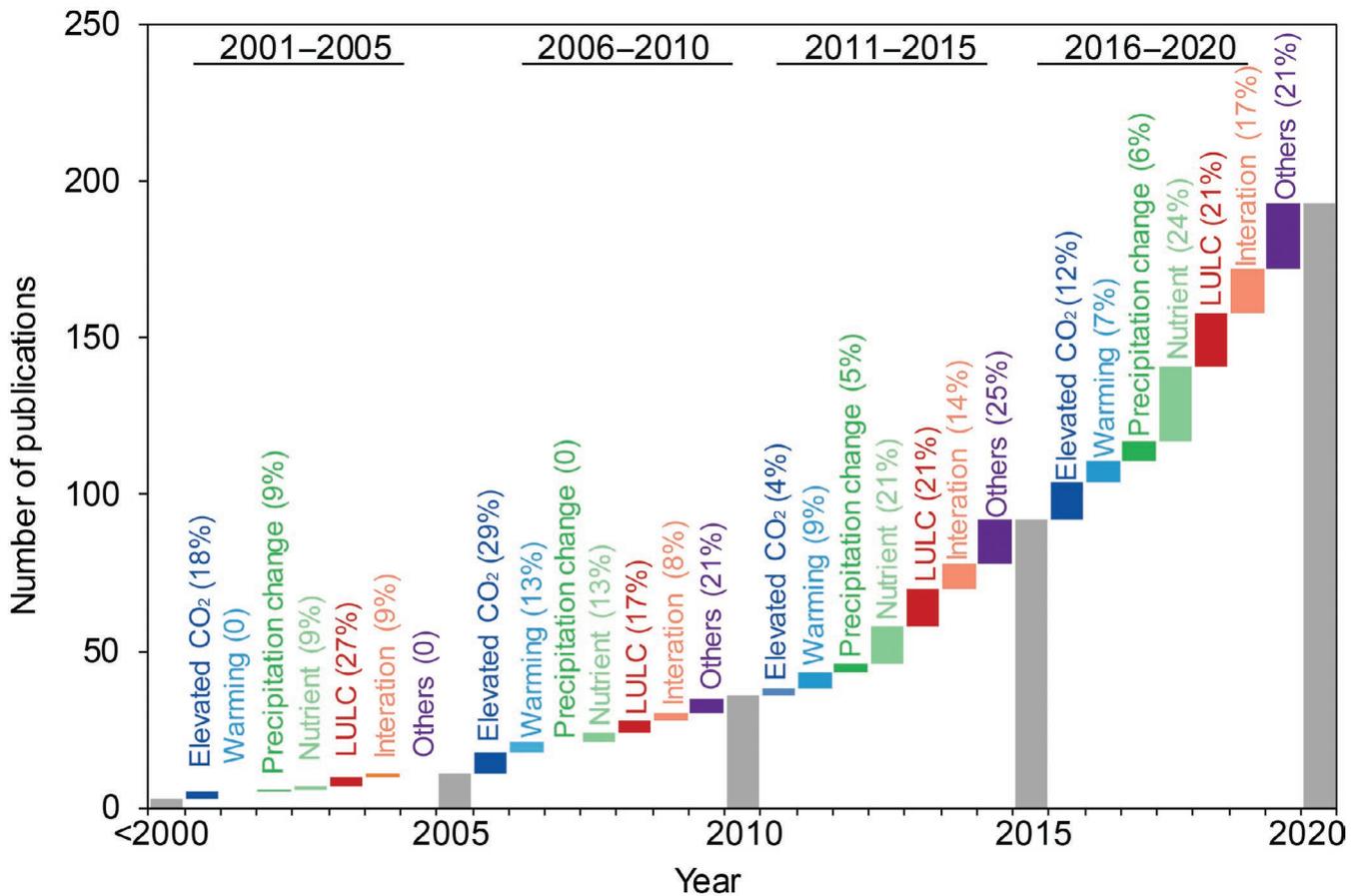


FIGURE 7 The studies published in *Global Change Biology* based on meta-analysis ($n = 192$) during 1995–2020. The results were searched with the keyword of “meta” or “meta-analysis”. The percentage following the labeled texts indicates the relative proportions of different global change factors. The grey bars indicate the sum of publications in different time periods. The first gray bar indicates the number of meta-analytical studies before 2000 ($n = 3$). The interaction ($n = 17$) represents that multiple global change factors are included in the meta-analysis. LULC ($n = 44$) indicates the Land use land cover, and others ($n = 40$) contain the studies on ocean acidification, tropospheric ozone change, and ultraviolet radiation

Pang, & Agca, 2011), wetland mapping (Pashaei, Kamangir, Starek, & Tissot, 2020), and geoscience (Reichstein et al., 2019). The challenges and opportunities of applying machine learning in those research areas offer important lessons for global change biology. For instance, Reichstein et al. (2019) has systematically reviewed the development of machine learning in Earth system science and provided five major challenges for deep learning applications in geoscience. Those challenges include low interpretability of predictive results, high physical consistency among predictions, complex and uncertain data, limited labels for observations, and high computational demand. More perceptively, they have highlighted five opportunities to integrate machine learning with physical modeling, including improving model parameterizations, replacing physical sub-models with machine learning models, constraining sub-models with machine learning algorithms, analyzing model-observation mismatch, surrogate modeling for enhancing computational efficiency. Some of these opportunities are driven by the assumption of a weak theoretical basis in biological processes (Reichstein et al., 2019). The lack of general mathematical equations to describe the biosphere has also been identified as a key challenge to predict life in current

Earth system models (Bonan & Doney, 2018). A recent study has shown that the combination of deep learning and data assimilation can significantly improve the simulations of soil carbon storage in a global land model (Tao et al., 2020). Because machine learning requires no assumption of the probability distribution of data errors, stationarity of time series, or correlations, it is well suited to study nonlinear responses and interactions of biological processes under global changes. However, an existing bottleneck is how to trace the observation-oriented patterns from machine learning and simulated results from Earth system models back to biologically meaningful parameters or processes, which is discussed next.

5.3 | Traceability analysis on modeling data

Uncertainties of modeling data collectively result from three sources, including external forcings, model structure, and process parameterizations (Luo et al., 2016). As shown in Figure 6, the divergence among models in the terrestrial carbon cycle calls for new approaches to trace the difference in modeling data back to its sources. A traceability

analysis has been developed for tracing terrestrial carbon storage to net primary productivity, carbon residence time, and their determinant processes and/or parameters (Figure 8; Luo et al., 2017; Xia et al., 2013). The theoretical basis of the traceability analysis is a differential matrix equation that describes the dynamic disequilibrium of the terrestrial carbon cycle (Luo & Weng, 2010). Such a matrix equation is also applicable to most state-of-the-art terrestrial biosphere models (Huang, Lu, et al., 2018; Xia, Luo, Wang, Weng, & Hararuk, 2012), making the model divergence traceable back to their sources. The traceability analysis has been applied to evaluate influences of climate forcings on carbon cycle (Ahlström, Xia, Arneeth, Luo, & Smith, 2015), compare inter-model differences in global carbon storage (Rafique et al., 2016; Zhou et al., 2018), and explain difference in terrestrial carbon cycle within a model (Jiang et al., 2017; Xia et al., 2013). An online traceability analysis system (i.e., TraceME) has been built to evaluate the terrestrial carbon cycles among CMIP6 models in a traceable, automatic, and visual way (Zhou et al., 2020). Further development of such traceability analysis to include more biological processes or applications of similar approaches in other components of Earth system models could be helpful to increase the interpretability of modeling results. Although models represent our imperfect understanding on the

Earth system, traceability frameworks like those shown in Figure 8 are important to improve the models through analyses of additional data and applications of new big-data techniques.

5.4 | Model-data fusion and ecological forecasting platforms

Model-data fusion is powerful for obtaining information about a phenomenon or a research question by iteratively integrating observational and modeling data. The applications and pitfalls of model-data fusion have been widely reviewed in different fields of ecology, such as integrated population models (Abadi, Gimenez, Arlettaz, & Schaub, 2010; Sun, Royle, & Fuller, 2019), species distribution (Isaac et al., 2020; Pacifici et al., 2017), and ecosystem carbon cycling (Keenan, Carbone, Reichstein, & Richardson, 2011; Luo et al., 2009; Raupach et al., 2005; Wang, Trudinger, & Enting, 2009). Due to the fast developments of data assimilation techniques and cyberinfrastructure in recent years, some model-data fusion platforms have emerged in the field of global change biology. For example, the Data Assimilation Research Testbed applies an Ensemble Kalman Filter to

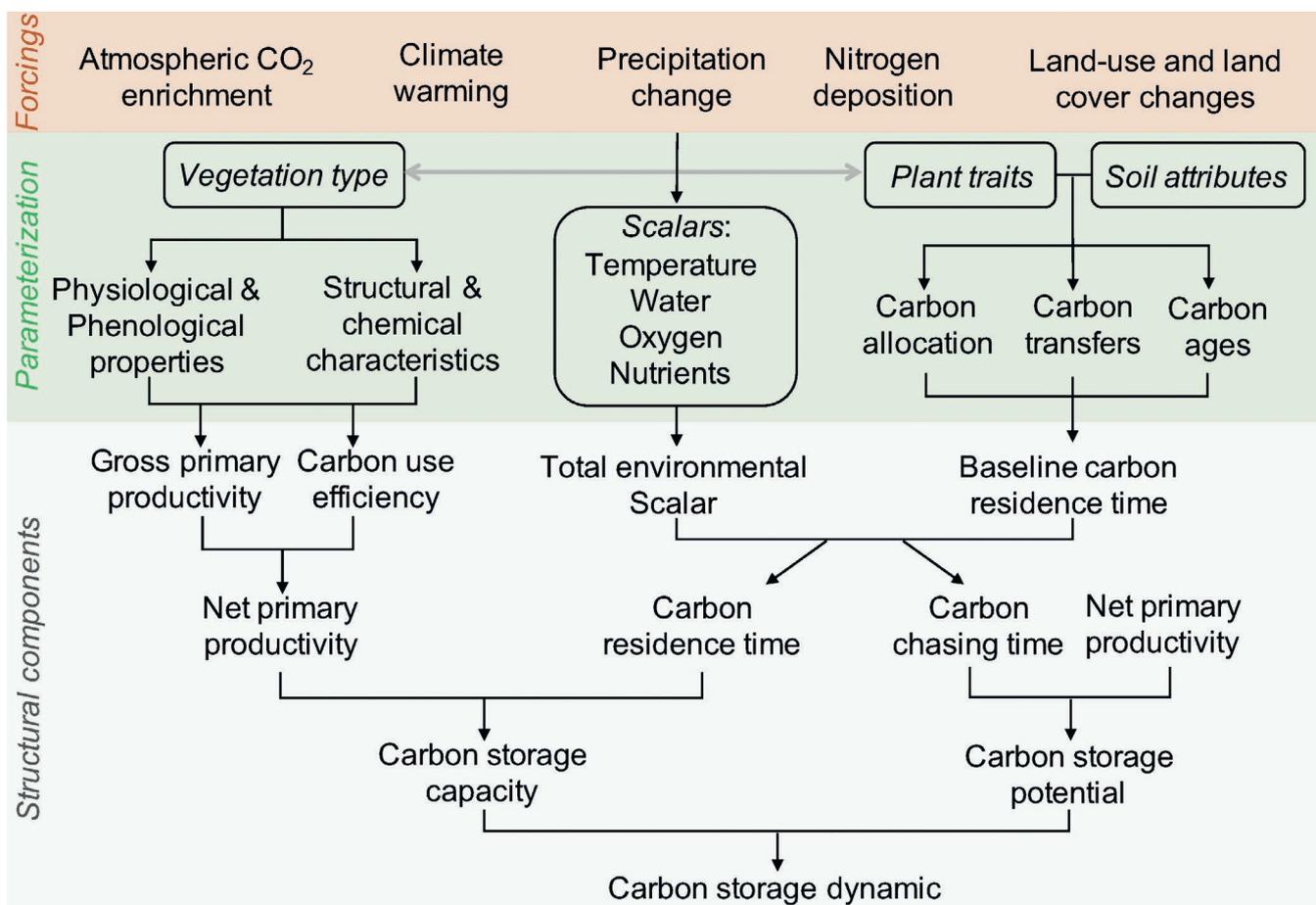


FIGURE 8 A traceability framework of terrestrial carbon cycle. More details about the traceable components have been introduced by Xia et al. (2013), Luo et al. (2017), and Cui et al. (2019). The theoretical derivation of the traceability framework can be found in Luo et al. (2017)

integrate models with a set of observations (Anderson et al., 2009). The Carbon Cycle Data Assimilation System assimilates different data streams into global land-surface models with a stepwise approach (Peylin et al., 2016). These efforts are mainly focusing on large spatial scales, therefore are difficult to assimilate new observations from local and site scales (Dietze et al., 2018). Recently, some new quantitative platforms have emerged for near real-time ecological forecasting under global changes on the site level. For example, the Ecological Platform for Assimilating Data has been applied to improve the forecasting of peatland responses to climate and environmental changes by integrating observations from the SPRUCE experiment and simulated data by the TECO model (Huang, Stacy, et al., 2019; Ma et al., 2017). Another example is the Predictive Ecosystem Analyzer (PEcAn, pecanproject.org), which has combined plant trait observations with terrestrial biosphere models for forecasting ecosystem CO₂ fluxes (Raczka, Dietze, Serbin, & Davis, 2018). The challenges for model-data fusion and ecological forecasting include different scientific assumptions to apply data assimilation techniques, large structural uncertainties among ecological models, errors from external forcings, and high computational demands (Dietze et al., 2018; Huang, Stacy, et al., 2019; Wang et al., 2009).

5.5 | Question or assumption-driven analyses on big data

The “assumption-centred” approach (Medlyn et al., 2015) is helpful to integrate different sources of data to improve our understanding of global change biology. One good example is the FACE Model-Data Synthesis project, which have used field data from FACE experiments to identify the key ecological assumptions causing differences among terrestrial ecosystem models (De Kauwe et al., 2013; Medlyn et al., 2015; Walker et al., 2019). Another example is the Permafrost Carbon Network, which uses multiple data sources and methods to quantify the role of permafrost carbon in affecting global climate change (Schuur & Abbott, 2011). The interactions between research communities of measurements (Plaza et al., 2019), modeling (McGuire et al., 2016), and remote sensing (Nitze, Grosse, Jones, Romanovsky, & Boike, 2018) have benefited our understanding of the permafrost carbon dynamics. Thus, to better extract knowledge from the big data, it is important to distinguish between which research questions can be answered by the increasing data availability and which questions require new techniques.

6 | CONCLUDING REMARKS

Big data science has been challenging the traditional research paradigm in global change biology in myriad ways. As exemplified in this review, the rapid increase in big data sets has facilitated global change biology in identifying global patterns, discovering emergent biological phenomena and mechanisms, and predicting future biosphere changes. However, there are also some challenges for better

using the big data in global change biology. First, most advancements have been done in increasing the data volume and velocity of data accumulation, but limited progress has been made to deal with the data variety and veracity. Second, our ability to assimilate data to improve near real-time forecasting and long-term projections of biological systems lags behind data collections and model developments. Third, the observed big data is neither random nor complete samples of all biota on the globe, while the modeling data include many unrecognized uncertainty sources. Fourth, both observed and modeling big data have paid major attention to responses rather than adaptations of living organisms to global changes. Fifth, there is no effective way to integrate biological data from gene to biosphere as well as environmental data from microenvironment to macro-ecosystems.

There is no doubt that big data provides many new research opportunities for global change biology. Many biological assumptions in current Earth system models are inconsistent with field observations (e.g., Bradford et al., 2014), calling for the discovery of new theories based on the observed big data. Also, tracing the impacts of global changes across organismic scales is another important opportunity, such as which genes determine a key adaptive trait (Gienapp, 2020), which traits regulate an ecosystem function (Cui et al., 2019; Fraser, 2020), and which ecosystem processes shape the terrestrial feedback to global climate change (Arora et al., 2019). New technologies such as the 3D monitoring of the Earth, modern modeling systems, and iterative model-data fusions have together promoted global change biology into high spatiotemporal resolutions. Furthermore, the fast development of cyberinfrastructure tools has been accelerating the pace of discovery by integrating observed and simulated data (Fer et al., 2020; Hoffman et al., 2017). Overall, the explosion of big data and related technologies can improve not only the fundamental understanding of global change biology but also the decision-making processes for sustaining the biosphere under future global changes.

ACKNOWLEDGEMENTS

This work was financially supported by the National Natural Science Foundation of China (31722009) and National Key R&D Program of China (2017YFA0604600). The authors appreciate Dr. Yiqi Luo and three anonymous reviewers for their constructive comments and suggestions on the manuscript. We also thank Rong Wang, Erqian Cui, Kun Huang, Xiaoni Xu, Chen Zhu, and Ning Wei for their help on earlier versions of the manuscript.

AUTHOR CONTRIBUTION

J.X. conceived the analyses and wrote the initial draft of the manuscript, J.W. did the literature survey on *Global Change Biology* and created the related figures, and S.N. contributed to subsequent revisions. All authors agreed to the submission of this manuscript.

ORCID

Jiayang Xia  <https://orcid.org/0000-0001-5923-6665>

Jing Wang  <https://orcid.org/0000-0002-2330-2787>

Shuli Niu  <https://orcid.org/0000-0002-2394-2864>

REFERENCES

- Abadi, F., Gimenez, O., Arlettaz, R., & Schaub, M. (2010). An assessment of integrated population models: Bias, accuracy, and violation of the assumption of independence. *Ecology*, *91*(1), 7–14. <https://doi.org/10.1890/08-2235.1>
- Adams, H. D., Zeppel, M. J. B., Anderegg, W. R. L., Hartmann, H., Landhäusser, S. M., Tissue, D. T., ... McDowell, N. G. (2017). A multi-species synthesis of physiological mechanisms in drought-induced tree mortality. *Nature Ecology & Evolution*, *1*(9), 1285–1291. <https://doi.org/10.1038/s41559-017-0248-x>
- Agapiou, A. (2017). Remote sensing heritage in a petabyte-scale: Satellite data and heritage Earth Engine applications. *International Journal of Digital Earth*, *10*(1), 85–102. <https://doi.org/10.1080/17538947.2016.1250829>
- Ahlström, A., Schurgers, G., & Smith, B. (2017). The large influence of climate model bias on terrestrial carbon cycle simulations. *Environmental Research Letters*, *12*(1), 014004. <https://doi.org/10.1088/1748-9326/12/1/014004>
- Ahlström, A., Xia, J., Arneth, A., Luo, Y., & Smith, B. (2015). Importance of vegetation dynamics for future terrestrial carbon cycling. *Environmental Research Letters*, *10*(5), 054019. <https://doi.org/10.1088/1748-9326/10/5/054019>
- Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., & Avellano, A. (2009). The data assimilation research testbed: A community facility. *Bulletin of the American Meteorological Society*, *90*(9), 1283–1296. <https://doi.org/10.1175/2009BAMS2618.1>
- Anderson, K. E., Glenn, N. F., Spaete, L. P., Shinneman, D. J., Pilliod, D. S., Arkle, R. S., ... Derryberry, D. R. (2018). Estimating vegetation biomass and cover across large plots in shrub and grass dominated drylands using terrestrial lidar and machine learning. *Ecological Indicators*, *84*, 793–802. <https://doi.org/10.1016/j.ecolind.2017.09.034>
- Anderson-Teixeira, K. J., Davies, S. J., Bennett, A. C., Gonzalez-Akre, E. B., Muller-Landau, H. C., Joseph Wright, S., ... Zimmerman, J. (2014). CTF5-ForestGEO: A worldwide network monitoring forests in an era of global change. *Global Change Biology*, *21*, 528–545. <https://doi.org/10.1111/gcb.12712>
- Anderson-Teixeira, K. J., Delong, J. P., Fox, A. M., Brese, D. A., & Litvak, M. E. (2011). Differential responses of production and respiration to temperature and moisture drive the carbon balance across a climatic gradient in New Mexico. *Global Change Biology*, *17*(1), 410–424. <https://doi.org/10.1111/j.1365-2486.2010.02269.x>
- Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., ... Ziehn, T. (2019). Carbon-concentration and carbon-climate feedbacks in CMIP6 models, and their comparison to CMIP5 models. *Biogeosciences Discussions*, 1–124. <https://doi.org/10.5194/bg-2019-473>
- Astrium Services. (2013). SPOT 6 & SPOT 7 imagery user guide, France. Retrieved from <http://www.astrium-geo.com/na/4589-spot-6-completes-technical-commissioning>
- Atkin, O. K., Bloomfield, K. J., Reich, P. B., Tjoelker, M. G., Asner, G. P., Bonal, D., ... Zaragoza-Castells, J. (2015). Global variability in leaf respiration in relation to climate, plant functional types and leaf traits. *New Phytologist*, *206*, 614–636. <https://doi.org/10.1111/nph.13253>
- Baldocchi, D., & Penuelas, J. (2019). The physics and ecology of mining carbon dioxide from the atmosphere by ecosystems. *Global Change Biology*, *25*(4), 1191–1197. <http://dx.doi.org/10.1111/gcb.14559>
- Baldocchi, D. D. (2020). How eddy covariance flux measurements have contributed to our understanding of *Global Change Biology*. *Global Change Biology*, *26*(1), 242–260. <https://doi.org/10.1111/gcb.14807>
- Bay, R. A., Harrigan, R. J., Underwood, V. L., Gibbs, H. L., Smith, T. B., & Rugg, K. (2018). Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science*, *359*(6371), 83–86. <http://dx.doi.org/10.1126/science.aan4380>
- Beckmann, M., Gerstner, K., Akin-Fajjiye, M., Ceaşu, S., Kambach, S., Kinlock, N. L., ... Seppelt, R. (2019). Conventional land-use intensification reduces species richness and increases production: A global meta-analysis. *Global Change Biology*, *25*, 1941–1956. <https://doi.org/10.1111/gcb.14606>
- Blazquez, D., & Domenech, J. (2018). Big Data sources and methods for social and economic analyses. *Technological Forecasting and Social Change*, *130*, 99–113. <https://doi.org/10.1016/j.techfore.2017.07.027>
- Blew, R. D. (1996). On the definition of ecosystem. *Bulletin of the Ecological Society of America*, *77*(3), 171–173. Retrieved from <https://www.jstor.org/stable/20168066>
- Bock, D. G., Caseys, C., Cousens, R. D., Hahn, M. A., Heredia, S. M., Hübner, S., ... Rieseberg, L. H. (2015). What we still don't know about invasion genetics. *Molecular Ecology*, *24*(9), 2277–2297. <http://dx.doi.org/10.1111/mec.13032>
- Bonan, G. B., & Doney, S. C. (2018). Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models. *Science*, *359*(6375), eaam8328. <https://doi.org/10.1126/science.aam8328>
- Bonan, G. B., Lombardozzi, D. L., Wieder, W. R., Oleson, K. W., Lawrence, D. M., Hoffman, F. M., & Collier, N. (2019). Model structure and climate data uncertainty in historical simulations of the terrestrial carbon cycle (1850–2014). *Global Biogeochemical Cycles*, *33*(10), 1310–1326. <https://doi.org/10.1029/2019GB006175>
- Bond-Lamberty, B. (2018). New techniques and data for understanding the global soil respiration flux. *Earth's Future*, *6*(9), 1176–1180. <http://dx.doi.org/10.1029/2018ef000866>
- Bond-Lamberty, B., & Thomson, A. (2010). A global database of soil respiration data. *Biogeosciences*, *7*(6), 1915–1926. <http://dx.doi.org/10.5194/bg-7-1915-2010>
- Bradford, M. A., Warren II, R. J., Baldrian, P., Crowther, T. W., Maynard, D. S., Oldfield, E. E., ... King, J. R. (2014). Climate fails to predict wood decomposition at regional scales. *Nature Climate Change*, *4*(7), 625–630. <https://doi.org/10.1038/nclimate2251>
- Caley, M. J., Fisher, R., & Mengersen, K. (2014). Global species richness estimates have not converged. *Trends in Ecology & Evolution*, *29*(4), 187–188. <https://doi.org/10.1016/j.tree.2014.02.002>
- Cardinale, B. J., Wright, J. P., Cadotte, M. W., Carroll, I. T., Hector, A., Srivastava, D. S., ... Weis, J. J. (2007). Impacts of plant diversity on biomass production increase through time because of species complementarity. *Proceedings of the National Academy of Sciences*, *104*(46), 18123–18128. <http://dx.doi.org/10.1073/pnas.0709069104>
- Carey, J. C., Tang, J., Templer, P. H., Kroeger, K. D., Crowther, T. W., Burton, A. J., ... Tietema, A. (2016). Temperature response of soil respiration largely unaltered with experimental warming. *Proceedings of the National Academy of Sciences*, *113*(48), 13797–13802. <http://dx.doi.org/10.1073/pnas.1605365113>
- Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, *19*, 171–209. <https://doi.org/10.1007/s11036-013-0489-0>
- Churkina, G., Schimel, D., Braswell, B. H., & Xiao, X. (2005). Spatial analysis of growing season length control over net ecosystem exchange. *Global Change Biology*, *11*(10), 1777–1787. <https://doi.org/10.1111/j.1365-2486.2005.001012.x>
- Cleverly, J., Eamus, D., Edwards, W., Grant, M., Grundy, M. J., Held, A., ... Morris, B. (2019). TERN, Australia's land observatory: Addressing the global challenge of forecasting ecosystem responses to climate variability and change. *Environmental Research Letters*, *14*, 095004. <https://doi.org/10.1088/1748-9326/ab33cb>
- Crimmins, T. M., Crimmins, M. A., Gerst, K. L., Rosemartin, A. H., & Weltzin, J. F. (2017). USA National Phenology Network's volunteer-contributed observations yield predictive models of phenological transitions. *PLoS One*, *12*(8), e0182919. <https://doi.org/10.1371/journal.pone.0182919>

- Crowther, T. W., Todd-Brown, K. E. O., Rowe, C. W., Wieder, W. R., Carey, J. C., Machmuller, M. B., ... Bradford, M. A. (2016). Quantifying global soil carbon losses in response to warming. *Nature*, 540(7631), 104–108. <http://dx.doi.org/10.1038/nature20150>
- Cui, E., Bian, C., Luo, Y., Niu, S., Wang, Y., & Xia, J. (2020). Spatially asynchronous changes in strength and stability of terrestrial net ecosystem productivity. *Biogeosciences Discussions*, 1–24. <https://doi.org/10.5194/bg-2020-26>
- Cui, E., Huang, K., Arain, M. A., Fisher, J. B., Huntzinger, D. N., Ito, A., ... Xia, J. (2019). Vegetation functional properties determine uncertainty of simulated ecosystem productivity: A traceability analysis in the East Asian monsoon region. *Global Biogeochemical Cycles*, 33(6), 668–689. <https://doi.org/10.1029/2018GB005909>
- Cui, E., Weng, E., Yan, E., & Xia, J. (2020). Robust leaf trait relationships across species under global environmental changes. *Nature Communications*, 11, 2999. <https://doi.org/10.1038/s41467-020-16839-9>
- Cummins, D., Kennington, W. J., Rudin-Bitterli, T., & Mitchell, N. J. (2019). A genome-wide search for local adaptation in a terrestrial-breeding frog reveals vulnerability to climate change. *Global Change Biology*, 25(9), 3151–3162. <https://doi.org/10.1111/gcb.14703>
- Curtis, P. G., Slay, C. M., Harris, N. L., Tyukavina, A., & Hansen, M. C. (2018). Classifying drivers of global forest loss. *Science*, 361(6407), 1108–1111. <https://doi.org/10.1126/science.aau3445>
- Davidson, I. C., Cott, G. M., Devaney, J. L., & Simkanin, C. (2018). Differential effects of biological invasions on coastal blue carbon: A global review and meta-analysis. *Global Change Biology*, 24(11), 5218–5230. <http://dx.doi.org/10.1111/gcb.14426>
- Davies-Barnard, T., Meyerholt, J., Zaehle, S., Friedlingstein, P., Brovkin, V., Fan, Y., Fisher, R. A., ... Wiltshire, A. (2020). Nitrogen cycling in CMIP6 land surface models: Progress and limitations. *Biogeosciences Discussions*, 1–32. <https://doi.org/10.5194/bg-2019-513>
- De Kauwe, M. G., Lin, Y.-S., Wright, I. J., Medlyn, B. E., Crous, K. Y., Ellsworth, D. S., ... Domingues, T. F. (2016). A test of the 'one-point method' for estimating maximum carboxylation capacity from field-measured, light-saturated photosynthesis. *New Phytologist*, 210(3), 1130–1144. <https://doi.org/10.1111/nph.13815>
- De Kauwe, M. G., Medlyn, B. E., Zaehle, S., Walker, A. P., Dietze, M. C., ... Norby, R. J. (2013). Forest water use and water use efficiency at elevated CO₂: A model-data intercomparison at two contrasting temperate forest FACE sites. *Global Change Biology*, 19, 1759–1779. <https://doi.org/10.1111/gcb.12164>
- de Vries, F. T., Manning, P., Tallowin, J. R. B., Mortimer, S. R., Pilgrim, E. S., Harrison, K. A., ... Bardgett, R. D. (2012). Abiotic drivers and plant traits explain landscape-scale patterns in soil microbial communities. *Ecology Letters*, 15, 1230–1239. <https://doi.org/10.1111/j.1461-0248.2012.01844.x>
- Dial, G., Bowen, H., Gerlach, F., Grodecki, J., & Oleszczuk, R. (2003). IKONOS satellite, imagery, and products. *Remote Sensing of Environment*, 88(1), 23–36. <https://doi.org/10.1016/j.rse.2003.08.014>
- Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., ... White, E. P. (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences of the United States of America*, 115(7), 1424–1432. <https://doi.org/10.1073/pnas.1710231115>
- Du, Y., Lu, R., & Xia, J. (2020). Impacts of global environmental change drivers on non-structural carbohydrates in terrestrial plants. *Functional Ecology*, 34(8), 1525–1536. <http://dx.doi.org/10.1111/1365-2435.13577>
- Du, Z., Weng, E., Jiang, L., Luo, Y., Xia, J., & Zhou, X. (2018). Carbon-nitrogen coupling under three schemes of model representation: A traceability analysis. *Geoscientific Model Development*, 11(11), 4399–4416. <https://doi.org/10.5194/gmd-11-4399-2018>
- Durden, J. M., Luo, J. Y., Alexander, H., Flanagan, A. M., & Grossmann, J. (2017). Integrating "Big Data" into aquatic ecology: Challenges and opportunities. *Limnology and Oceanography Bulletin*, 26, 101–108. <https://doi.org/10.1002/lob.10213>
- Enquist, B. J., Condit, R., Peet, R. K., Schildhauer, M., & Thiers, B. M. (2016). Cyberinfrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity. *PeerJ Preprints*, 4, e2615v2. <https://doi.org/10.7287/peerj.preprints.2615v2>
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5), 1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>
- Fan, J., Han, F., & Liu, H. (2014). Challenges of big data analysis. *National Science Review*, 1(2), 293–314. <https://doi.org/10.1093/nsr/nwt032>
- Fer, I., Gardella, A. K., Shiklomanov, A. N., Serbin, S. P., Kauwe, M. G. D., Raiho, A., ... Dietze, M. C. (2020). Beyond modeling: A roadmap to community cyberinfrastructure for ecological data-model integration. *Preprints*, 2020, 2020010176. <https://doi.org/10.20944/preprints202001.0176.v1>
- Fernandez, C. W., Heckman, K., Kolka, R., & Kennedy, P. G. (2019). Melanin mitigates the accelerated decay of mycorrhizal necromass with peatland warming. *Ecology Letters*, 22(3), 498–505. <https://doi.org/10.1111/ele.13209>
- Franks, S. J. (2011). Plasticity and evolution in drought avoidance and escape in the annual plant *Brassica rapa*. *New Phytologist*, 190(1), 249–257. <https://doi.org/10.1111/j.1469-8137.2010.03603.x>
- Franks, S. J., Sim, S., & Weis, A. E. (2007). Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proceedings of the National Academy of Sciences of the United States of America*, 104(4), 1278–1282. <https://doi.org/10.1073/pnas.0608379104>
- Franzosa, E. A., Hsu, T., Sirota-Madi, A., Shafquat, A., Abu-Ali, G., Morgan, X. C., & Huttenhower, C. (2015). Sequencing and beyond: Integrating molecular 'omics' for microbial community profiling. *Nature Reviews Microbiology*, 13(6), 360–372. <http://dx.doi.org/10.1038/nrmicro3451>
- Fraser, L. H. (2020). TRY – A plant trait database of databases. *Global Change Biology*, 26(1), 189–190. <https://doi.org/10.1111/gcb.14869>
- Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., ... Zaehle, S. (2019). Global carbon budget 2019. *Earth System Science Data*, 11(4), 1783–1838. <https://doi.org/10.5194/essd-11-1783-2019>
- Fu, Z., Stoy, P. C., Poulter, B., Gerken, T., Zhang, Z., Wakbulcho, G., & Niu, S. (2019). Maximum carbon uptake rate dominates the interannual variability of global net ecosystem exchange. *Global Change Biology*, 25(10), 3381–3394. <http://dx.doi.org/10.1111/gcb.14731>
- Gienapp, P. (2020). Opinion: Is gene mapping in wild populations useful for understanding and predicting adaptation to global change? *Global Change Biology*, 26(5), 2737–2749. <https://doi.org/10.1111/gcb.15058>
- Gittings, J. A., Raitsos, D. E., Racault, M.-F., Brewin, R. J. W., Pradhan, Y., Sathyendranath, S., & Platt, T. (2017). Seasonal phytoplankton blooms in the Gulf of Aden revealed by remote sensing. *Remote Sensing of Environment*, 189, 56–66. <https://doi.org/10.1016/j.rse.2016.10.043>
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3–8. <https://doi.org/10.3102/0013189X005010003>
- Gu, L., Hanson, P. J., Post, W. M., Kaiser, D. P., Yang, B., Nemani, R., ... Meyers, T. (2008). The 2007 eastern US spring freeze: Increased cold damage in a warming world. *BioScience*, 58(3), 253–262. <https://doi.org/10.1641/B580311>
- Guan, K., Pan, M., Li, H., Wolf, A., Wu, J., Medvigy, D., ... Lyapustin, A. I. (2015). Photosynthetic seasonality of global tropical forests

- constrained by hydroclimate. *Nature Geoscience*, 8(4), 284–289. <http://dx.doi.org/10.1038/ngeo2382>
- Guo, L. B., & Gifford, R. M. (2002). Soil carbon stocks and land use change: A meta-analysis. *Global Change Biology*, 8(4), 345–360. <https://doi.org/10.1046/j.1354-1013.2002.00486.x>
- Gurevitch, J., Koricheva, J., Nakagawa, S., & Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature*, 555(7695), 175–182. <https://doi.org/10.1038/nature25753>
- Hanson, P. J., Gill, A. L., Xu, X., Phillips, J. R., Weston, D. J., Kolka, R. K., ... Hook, L. A. (2016). Intermediate-scale community-level flux of CO₂ and CH₄ in a Minnesota peatland: Putting the SPRUCE project in a global context. *Biogeochemistry*, 129(3), 255–272. <https://doi.org/10.1007/s10533-016-0230-8>
- Harpole, W. S., Sullivan, L. L., Lind, E. M., Firn, J., Adler, P. B., Borer, E. T., ... Wrang, P. D. (2016). Addition of multiple limiting resources reduces grassland diversity. *Nature*, 537(7618), 93–96. <http://dx.doi.org/10.1038/nature19324>
- Harrison, S., Spasojevic, M. J., & Li, D. (2020). Climate and plant community diversity in space and time. *Proceedings of the National Academy of Sciences of the United States of America*, 117(9), 4464–4470. <https://doi.org/10.1073/pnas.1921724117>
- Harrison, T., Gibbs, J., & Winfree, R. (2019). Anthropogenic landscapes support fewer rare bee species. *Landscape Ecology*, 34(5), 967–978. <https://doi.org/10.1007/s10980-017-0592-x>
- Hautier, Y., Seabloom, E. W., Borer, E. T., Adler, P. B., Harpole, W. S., Hillebrand, H., ... Hector, A. (2014). Eutrophication weakens stabilizing effects of diversity in natural grasslands. *Nature*, 508(7497), 521–525. <http://dx.doi.org/10.1038/nature13014>
- Haverd, V., Smith, B., Canadell, J. G., Cuntz, M., Mikaloff-Fletcher, S., Farquhar, G., ... Trudinger, C. M. (2020). Higher than expected CO₂ fertilization inferred from leaf to global observations. *Global Change Biology*, 26(4), 2390–2402. <http://dx.doi.org/10.1111/gcb.14950>
- Hawes, N. A., Amadoru, A., Tremblay, L. A., Pochon, X., Dunphy, B., Fidler, A. E., & Smith, K. F. (2019). Epigenetic patterns associated with an ascidian invasion: A comparison of closely related clades in their native and introduced ranges. *Scientific Reports*, 9(1). <http://dx.doi.org/10.1038/s41598-019-49813-7>
- He, Y., Chen, G., Potter, C., & Meentemeyer, R. K. (2019). Integrating multi-sensor remote sensing and species distribution modeling to map the spread of emerging forest disease and tree mortality. *Remote Sensing of Environment*, 231, 111238. <https://doi.org/10.1016/j.rse.2019.111238>
- Hemson, G., Johnson, P., South, A., Kenward, R., Ripley, R., & Macdonald, D. (2005). Are kernels the mustard? Data from global positioning system (GPS) collars suggests problems for kernel home-range analyses with least-squares cross-validation. *Journal of Animal Ecology*, 74(3), 455–463. <https://doi.org/10.1111/j.1365-2656.2005.00944.x>
- Hendry, G. A., & Grime, J. P. (1993). *Methods in comparative plant ecology: A laboratory manual*. Dordrecht, the Netherlands: Springer. <https://doi.org/10.1007/978-94-011-1494-3>
- Hobbie, E. A., Chen, J., Hanson, P. J., Iversen, C. M., McFarlane, K. J., Thorp, N. R., & Hofmockel, K. S. (2017). Long-term carbon and nitrogen dynamics at SPRUCE revealed through stable isotopes in peat profiles. *Biogeosciences*, 14(9), 2481–2494. <https://doi.org/10.5194/bg-14-2481-2017>
- Hoffman, F. M., Koven, C. D., Keppel-Aleks, G., Lawrence, D. M., Riley, W. J., Randerson, J. T., Ahlström, A., ... Koch, D. (2017). *International Land Model Benchmarking (ILAMB) 2016 workshop report*, DOE/SC-0186. Germantown, MD: U.S. Department of Energy. <https://doi.org/10.2172/1330803>
- Huang, K., & Xia, J. (2019). High ecosystem stability of evergreen broad-leaf forests under severe droughts. *Global Change Biology*, 25(10), 3494–3503. <https://doi.org/10.1111/gcb.14748>
- Huang, K., Xia, J., Wang, Y., Ahlström, A., Chen, J., Cook, R. B., ... Luo, Y. (2018). Enhanced peak growth of global vegetation and its key mechanisms. *Nature Ecology & Evolution*, 2(12), 1897. <https://doi.org/10.1038/s41559-018-0714-0>
- Huang, M., Piao, S., Ciais, P., Peñuelas, J., Wang, X., Keenan, T. F., ... Janssens, I. A. (2019). Air temperature optima of vegetation productivity across global biomes. *Nature Ecology Evolution*, 3(5), 772–779. <https://doi.org/10.1038/s41559-019-0838-x>
- Huang, X., Wang, Y., Li, J., Chang, X., Cao, Y., Xie, J., & Gong, J. (2020). High-resolution urban land-cover mapping and landscape analysis of the 42 major cities in China using ZY-3 satellite images. *Science Bulletin*, 65(12), 1039–1048. <https://doi.org/10.1016/j.scib.2020.03.003>
- Huang, Y., Lu, X., Shi, Z., Lawrence, D., Koven, C. D., Xia, J., ... Luo, Y. (2018). Matrix approach to land carbon cycle modeling: A case study with the Community Land Model. *Global Change Biology*, 24(3), 1394–1404. <https://doi.org/10.1111/gcb.13948>
- Huang, Y., Stacy, M., Jiang, J., Sundi, N., Ma, S., Saruta, V., ... Luo, Y. (2019). Realized ecological forecast through an interactive Ecological Platform for Assimilating Data (EcoPAD, v1.0) into models. *Geoscientific Model Development*, 12(3), 1119–1137. <https://doi.org/10.5194/gmd-12-1119-2019>
- Hungate, B. A., Dukes, J. S., Shaw, M. R., Luo, Y., & Field, C. B. (2003). Atmospheric science: Nitrogen and climate change. *Science*, 302(5650), 1512–1513. <https://doi.org/10.1126/science.1091390>
- Isaac, N. J. B., Jarzyna, M. A., Keil, P., Dambly, L. I., Boersch-Supan, P. H., Browning, E., ... O'Hara, R. B. (2020). Data integration for large-scale models of species distributions. *Trends in Ecology & Evolution*, 35(1), 56–67. <https://doi.org/10.1016/j.tree.2019.08.006>
- Jiang, L., Shi, Z., Xia, J., Liang, J., Lu, X., Wang, Y., & Luo, Y. (2017). Transient traceability analysis of land carbon storage dynamics: Procedures and its application to two forest ecosystems. *Journal of Advances in Modeling Earth Systems*, 9(8), 2822–2835. <https://doi.org/10.1002/2017MS001004>
- Jiang, Z. (2016). How many species are there on Earth? *Chinese Science Bulletin*, 61(21), 2337–2343. <https://doi.org/10.1360/N972015-01320>
- Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., ... Williams, C. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research Biogeosciences*, 116(G3), G00J07. <https://doi.org/10.1029/2010JG001566>
- Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlström, A., ... Zeng, N. (2017). Compensatory water effects link yearly global land CO₂ sink changes to temperature. *Nature*, 541, 516–520. <https://doi.org/10.1038/nature20780>
- Kampe, T. U., Johnson, B. R., Kuester, M. A., & Keller, M. (2010). NEON: The first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. *Journal of Applied Remote Sensing*, 4(1), 043510. <https://doi.org/10.1117/1.3361375>
- Kattge, J., Bönsch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., ... Wirth, C. (2020). TRY plant trait database – Enhanced coverage and open access. *Global Change Biology*, 26(1), 119–188. <https://doi.org/10.1111/gcb.14904>
- Kattge, J., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönsch, G., ... Wirth, C. (2011). TRY – A global database of plant traits. *Global Change Biology*, 17(9), 2905–2935. <https://doi.org/10.1111/j.1365-2486.2011.02451.x>
- Kattge, J., Knorr, W., Raddatz, T., & Wirth, C. (2009). Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Global Change Biology*, 15, 976–991. <https://doi.org/10.1111/j.1365-2486.2008.01744.x>
- Keenan, T. F., Carbone, M. S., Reichstein, M., & Richardson, A. D. (2011). The model–data fusion pitfall: Assuming certainty in an uncertain world. *Oecologia*, 167(3), 587–597. <https://doi.org/10.1007/s00442-011-2106-x>

- Keenan, T. F., & Niinemets, Ü. (2016). Global leaf trait estimates biased due to plasticity in the shade. *Nature Plants*, 3(1), 1–6. <https://doi.org/10.1038/nplants.2016.201>
- Knapp, A. K., Avolio, M. L., Beier, C., Carroll, C. J. W., Collins, S. L., Dukes, J. S., ... Smith, M. D. (2017). Pushing precipitation to the extremes in distributed experiments: Recommendations for simulating wet and dry years. *Global Change Biology*, 23(5), 1774–1782. <http://dx.doi.org/10.1111/gcb.13504>
- Knapp, A. K., Hoover, D. L., Wilcox, K. R., Avolio, M. L., Koerner, S. E., La Pierre, K. J., ... Smith, M. D. (2015). Characterizing differences in precipitation regimes of extreme wet and dry years: Implications for climate change experiments. *Global Change Biology*, 21(7), 2624–2633. <http://dx.doi.org/10.1111/gcb.12888>
- König, P., Tautenhahn, S., Cornelissen, J. H. C., Kattge, J., Bönisch, G., & Römermann, C. (2018). Advances in flowering phenology across the Northern Hemisphere are explained by functional traits. *Global Ecology & Biogeography*, 27(3), 310–321. <https://doi.org/10.1111/geb.12696>
- LaDeau, S. L., Han, B. A., Rosi-Marshall, E. J., & Weathers, K. C. (2017). The next decade of big data in ecosystem science. *Ecosystems*, 20(2), 274–283. <https://doi.org/10.1007/s10021-016-0075-y>
- Langner, A., & Siegert, F. (2009). Spatiotemporal fire occurrence in Borneo over a period of 10 years. *Global Change Biology*, 15(1), 48–62. <https://doi.org/10.1111/j.1365-2486.2008.01828.x>
- Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., ... Zheng, B. O. (2018). Global carbon budget 2018. *Earth System Science Data*, 10(4), 2141–2194. <https://doi.org/10.5194/essd-10-2141-2018>
- Leboeuf, A., Beaudoin, A., Fournier, R. A., Guindon, L., Luther, J. E., & Lambert, M.-C. (2007). A shadow fraction method for mapping biomass of northern boreal black spruce forests using QuickBird imagery. *Remote Sensing of Environment*, 110(4), 488–500. <https://doi.org/10.1016/j.rse.2006.05.025>
- Li, D., Stucky, B. J., Deck, J., Baiser, B., & Guralnick, R. P. (2019). The effect of urbanization on plant phenology depends on regional temperature. *Nature Ecology & Evolution*, 3(12), 1661–1667. <https://doi.org/10.1038/s41559-019-1004-1>
- Li, L., Zheng, Z., Wang, W., Biederman, J. A., Xu, X., Ran, Q., ... Wang, Y. (2020). Terrestrial N₂O emissions and related functional genes under climate change: A global meta-analysis. *Global Change Biology*, 26(2), 931–943. <https://doi.org/10.1111/gcb.14847>
- Li, Y., Huang, C., Ding, L., Li, Z., Pan, Y., & Gao, X. (2019). Deep learning in bioinformatics: Introduction, application, and perspective in the big data era. *Methods*, 166, 4–21. <https://doi.org/10.1016/j.jymeth.2019.04.008>
- Li, Z., Tian, D., Wang, B., Wang, J., Wang, S., Chen, H. Y. H., ... Niu, S. (2019). Microbes drive global soil nitrogen mineralization and availability. *Global Change Biology*, 25(3), 1078–1088. <http://dx.doi.org/10.1111/gcb.14557>
- Li, Z., Xia, J., Ahlström, A., Rinke, A., Koven, C., Hayes, D. J., ... Luo, Y. (2018). Non-uniform seasonal warming regulates vegetation greening and atmospheric CO₂ amplification over northern lands. *Environmental Research Letters*, 13(12), 124008. <https://doi.org/10.1088/1748-9326/aae9ad>
- Liang, J., Qi, X., Souza, L., & Luo, Y. (2016). Processes regulating progressive nitrogen limitation under elevated carbon dioxide: A meta-analysis. *Biogeosciences*, 13(9), 2689–2699. <http://dx.doi.org/10.5194/bg-13-2689-2016>
- Liang, S., Wang, D., He, T., & Yu, Y. (2019). Remote sensing of earth's energy budget: Synthesis and review. *International Journal of Digital Earth*, 12, 737–780. <https://doi.org/10.1080/17538947.2019.1597189>
- Liao, C., Peng, R., Luo, Y., Zhou, X., Wu, X., Fang, C., ... Li, B. (2008). Altered ecosystem carbon and nitrogen cycles by plant invasion: A meta-analysis. *New Phytologist*, 177(3), 706–714. <https://doi.org/10.1111/j.1469-8137.2007.02290.x>
- Lin, D., Xia, J., & Wan, S. (2010). Climate warming and biomass accumulation of terrestrial plants: A meta-analysis. *New Phytologist*, 188(1), 187–198. <http://dx.doi.org/10.1111/j.1469-8137.2010.03347.x>
- Liu, L. (2013). Computing infrastructure for big data processing. *Frontiers of Computer Science*, 7, 165–170. <https://doi.org/10.1007/s11704-013-3900-x>
- Liu, Y., Piao, S., Gasser, T., Ciais, P., Yang, H., Wang, H., ... Wang, T. (2019). Field-experiment constraints on the enhancement of the terrestrial carbon sink by CO₂ fertilization. *Nature Geoscience*, 12, 809–814. <https://doi.org/10.1038/s41561-019-0436-1>
- Luo, J. Y., Irisson, J.-O., Graham, B., Guigand, C., Sarafraz, A., Mader, C., & Cowen, R. K. (2018). Automated plankton image analysis using convolutional neural networks. *Limnology and Oceanography Methods*, 16, 814–827. <https://doi.org/10.1002/lom3.10285>
- Luo, Y., Ahlström, A., Allison, S. D., Batjes, N. H., Brovkin, V., Carvalhais, N., ... Zhou, T. (2016). Toward more realistic projections of soil carbon dynamics by Earth system models. *Global Biogeochemical Cycles*, 30(1), 40–56. <https://doi.org/10.1002/2015GB005239>
- Luo, Y. Q., Hui, D. F., & Zhang, D. Q. (2006). Elevated CO₂ stimulates net accumulations of carbon and nitrogen in land ecosystems: A meta-analysis. *Ecology*, 87(1), 53–63. <https://doi.org/10.1890/04-1724>
- Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., ... Schimel, D. (2011). Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications*, 21(5), 1429–1442. <https://doi.org/10.1890/09-1275.1>
- Luo, Y., & Schuur, E. A. G. (2020). Model parameterization to represent processes at unresolved scales and changing properties of evolving systems. *Global Change Biology*, 26(3), 1109–1117. <https://doi.org/10.1111/gcb.14939>
- Luo, Y., Shi, Z., Lu, X., Xia, J., Liang, J., Jiang, J., ... Wang, Y.-P. (2017). Transient dynamics of terrestrial carbon storage: Mathematical foundation and its applications. *Biogeosciences*, 14(1), 145–161. <https://doi.org/10.5194/bg-14-145-2017>
- Luo, Y., Su, B., Currie, W. S., Dukes, J. S., Finzi, A., Hartwig, U., ... Field, C. B. (2004). Progressive nitrogen limitation of ecosystem responses to rising atmospheric carbon dioxide. *BioScience*, 54(8), 731–739. [https://doi.org/10.1641/0006-3568\(2004\)054\[0731:PNLOE R\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2004)054[0731:PNLOE R]2.0.CO;2)
- Luo, Y., Wan, S., Hui, D., & Wallace, L. L. (2001). Acclimatization of soil respiration to warming in a tall grass prairie. *Nature*, 413(6856), 622–625. <https://doi.org/10.1038/35098065>
- Luo, Y., & Weng, E. (2010). Dynamic disequilibrium of the terrestrial carbon cycle under global change. *Trends in Ecology & Evolution*, 26(2), 96–104. <https://doi.org/10.1016/j.tree.2010.11.003>
- Luo, Y., Weng, E., Wu, X., Gao, C., Zhou, X., & Zhang, L. (2009). Parameter identifiability, constraint, and equifinality in data assimilation with ecosystem models. *Ecological Applications*, 19(3), 571–574. <https://doi.org/10.1890/08-0561.1>
- Ma, S., Jiang, J., Huang, Y., Shi, Z., Wilson, R. M., Ricciuto, D., ... Luo, Y. (2017). Data-constrained projections of methane fluxes in a northern minnesota peatland in response to elevated CO₂ and warming. *Journal of Geophysical Research: Biogeosciences*, 122(11), 2841–2861. <https://doi.org/10.1002/2017JG003932>
- Ma, Z., Guo, D., Xu, X., Lu, M., Bardgett, R. D., Eissenstat, D. M., ... Hedin, L. O. (2018). Evolutionary history resolves global organization of root functional traits. *Nature*, 555, 94–97. <https://doi.org/10.1038/nature25783>
- Maes, W. H., & Steppe, K. (2019). Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. *Trends in Plant Science*, 24, 152–164. <https://doi.org/10.1016/j.tplan ts.2018.11.007>
- Mahecha, M. D., Reichstein, M., Jung, M., Seneviratne, S. I., Zaehle, S., Beer, C., ... Moors, E. (2010). Comparing observations and process-based simulations of biosphere-atmosphere exchanges on

- multiple timescales. *Journal of Geophysical Research: Biogeosciences*, 115(G2), G02003. <https://doi.org/10.1029/2009JG001016>
- McCall, A. C., & Irwin, R. E. (2006). Florivory: The intersection of pollination and herbivory. *Ecology Letters*, 9(12), 1351–1365. <http://dx.doi.org/10.1111/j.1461-0248.2006.00975.x>
- McGuire, A. D., Koven, C., Lawrence, D. M., Clein, J. S., Xia, J., Beer, C., ... Zhuang, Q. (2016). Variability in the sensitivity among model simulations of permafrost and carbon dynamics in the permafrost region between 1960 and 2009. *Global Biogeochemical Cycles*, 30, 1015–1037. <https://doi.org/10.1002/2016GB005405>
- Medlyn, B. E., Zaehle, S., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hanson, P. J., ... Norby, R. J. (2015). Using ecosystem experiments to improve vegetation models. *Nature Climate Change*, 5, 528–534. <https://doi.org/10.1038/nclimate2621>
- Meyerholt, J., Sickel, K., & Zaehle, S. (2020). Ensemble projections elucidate effects of uncertainty in terrestrial nitrogen limitation on future carbon uptake. *Global Change Biology*, 26(7), 3978–3996. <https://doi.org/10.1111/gcb.15114>
- Midolo, G., Alkemade, R., Schipper, A. M., Benítez-López, A., Perring, M. P., & De Vries, W. (2018). Impacts of nitrogen addition on plant species richness and abundance: A global meta-analysis. *Global Ecology and Biogeography*, 28(3), 398–413. <http://dx.doi.org/10.1111/geb.12856>
- Moore, B. M., Wang, P., Fan, P., Leong, B., Schenck, C. A., Lloyd, J. P., ... Shiu, S.-H. (2019). Robust predictions of specialized metabolism genes through machine learning. *Proceedings of the National Academy of Sciences of the United States of America*, 116(6), 2344–2353. <https://doi.org/10.1073/pnas.1817074116>
- Morris, K. J., Bett, B. J., Durden, J. M., Huvenne, V. A., Milligan, R., Jones, D. O. B., ... Ruhl, H. A. (2014). A new method for ecological surveying of the abyss using autonomous underwater vehicle photography. *Limnology and Oceanography Methods*, 12, 795–809. <https://doi.org/10.4319/lom.2014.12.795>
- Mousavi-Derazmahalleh, M., Bayer, P. E., Hane, J. K., Valliyodan, B., Nguyen, H. T., Nelson, M. N., ... Edwards, D. (2019). Adapting legume crops to climate change using genomic approaches. *Plant, Cell & Environment*, 42(1), 6–19. <https://doi.org/10.1111/pce.13203>
- Muller-Karger, F. E., Hestir, E., Ade, C., Turpie, K., Roberts, D. A., Siegel, D., ... Jetz, W. (2018). Satellite sensor requirements for monitoring essential biodiversity variables of coastal ecosystems. *Ecological Applications*, 28(3), 749–760. <https://doi.org/10.1002/eap.1682>
- Musavi, T., Migliavacca, M., Reichstein, M., Kattge, J., Wirth, C., Black, T. A., ... Mahecha, M. D. (2017). Stand age and species richness dampen interannual variation of ecosystem-level photosynthetic capacity. *Nature Ecology & Evolution*, 1(2), 0048. <http://dx.doi.org/10.1038/s41559-016-0048>
- Musinsky, J., Tabor, K., Cano, C. A., Ledezma, J. C., Mendoza, E., Rasolohery, A., & Sajudin, E. R. (2018). Conservation impacts of a near real-time forest monitoring and alert system for the tropics. *Remote Sensing in Ecology & Conservation*, 4(3), 189–196. <https://doi.org/10.1002/rse2.78>
- Nallu, S., Hill, J. A., Don, K., Sahagun, C., Zhang, W., Meslin, C., ... Kronforst, M. R. (2018). The molecular genetic basis of herbivory between butterflies and their host plants. *Nature Ecology & Evolution*, 2(9), 1418–1427. <http://dx.doi.org/10.1038/s41559-018-0629-9>
- Nativi, S., Mazzetti, P., Santoro, M., Papeschi, F., Craglia, M., & Ochiai, O. (2015). Big Data challenges in building the Global Earth Observation System of Systems. *Environmental Modelling Software*, 68, 1–26. <https://doi.org/10.1016/j.envsoft.2015.01.017>
- Nitze, I., Grosse, G., Jones, B. M., Romanovsky, V. E., & Boike, J. (2018). Remote sensing quantifies widespread abundance of permafrost region disturbances across the Arctic and Subarctic. *Nature Communications*, 9, 5423. <https://doi.org/10.1038/s41467-018-07663-3>
- Niu, S., Fu, Z., Luo, Y., Stoy, P. C., Keenan, T. F., Poulter, B., ... Yu, G. (2017). Interannual variability of ecosystem carbon exchange: From observation to prediction. *Global Ecology and Biogeography*, 26(11), 1225–1237. <https://doi.org/10.1111/geb.12633>
- Niu, S., Luo, Y., Fei, S., Yuan, W., Schimel, D., Law, B. E., ... Zhou, X. (2012). Thermal optimality of net ecosystem exchange of carbon dioxide and underlying mechanisms. *New Phytologist*, 194(3), 775–783. <https://doi.org/10.1111/j.1469-8137.2012.04095.x>
- Niu, S., Sherry, R. A., Zhou, X., & Luo, Y. (2013). Ecosystem carbon fluxes in response to warming and clipping in a tallgrass prairie. *Ecosystems*, 16(6), 948–961. <https://doi.org/10.1007/s10021-013-9661-4>
- Niu, S., Wang, S., Wang, J., Xia, J., & Yu, G. (2020). Integrative ecology in the era of big data – From observation to prediction. *Science China-Earth Sciences*. <https://doi.org/10.1007/s11430-020-9664-6>
- Olsson, P.-O., Lindström, J., & Eklundh, L. (2016). Near real-time monitoring of insect induced defoliation in subalpine birch forests with MODIS derived NDVI. *Remote Sensing of Environment*, 181, 42–53. <https://doi.org/10.1016/j.rse.2016.03.040>
- Overpeck, J. T., Meehl, G. A., Bony, S., & Easterling, D. R. (2011). Climate data challenges in the 21st century. *Science*, 331(6018), 700–702. <https://doi.org/10.1126/science.1197869>
- Pacifici, K., Reich, B. J., Miller, D. A. W., Gardner, B., Stauffer, G., Singh, S., ... Collazo, J. A. (2017). Integrating multiple data sources in species distribution modeling: A framework for data fusion. *Ecology*, 98(3), 840–850. <https://doi.org/10.1002/ecy.1710>
- Pan, N., Feng, X., Fu, B., Wang, S., Ji, F., & Pan, S. (2018). Increasing global vegetation browning hidden in overall vegetation greening: Insights from time-varying trends. *Remote Sensing of Environment*, 214, 59–72. <https://doi.org/10.1016/j.rse.2018.05.018>
- Park, T., Chen, C., Macias-Fauria, M., Tømmervik, H., Choi, S., Winkler, A., ... Myneni, R. B. (2019). Changes in timing of seasonal peak photosynthetic activity in northern ecosystems. *Global Change Biology*, 25(7), 2382–2395. <https://doi.org/10.1111/gcb.14638>
- Pashaie, M., Kamangir, H., Starek, M. J., & Tissot, P. (2020). Review and evaluation of deep learning architectures for efficient land cover mapping with UAS hyper-spatial imagery: A case study over a wetland. *Remote Sensing*, 12, 959. <https://doi.org/10.3390/rs12060959>
- Pastorello, G., Papale, D., Chu, H., Trotta, C., Agarwal, D., Canfora, E., ... Torn, M. (2017). A new data set to keep a sharper eye on land-air exchanges. *Eos, Transactions American Geophysical Union*, 98(8). <https://doi.org/10.1029/2017EO071597>
- Peng, S., Piao, S., Ciais, P., Myneni, R. B., Chen, A., Chevallier, F., ... Zeng, H. (2013). Asymmetric effects of daytime and night-time warming on Northern Hemisphere vegetation. *Nature*, 501(7465), 88–92. <https://doi.org/10.1038/nature12434>
- Peterson, A. G., Ball, J. T., Luo, Y. I., Field, C. B., Reich, P. B., Curtis, P. S., ... Cmeal participants. (1999). The photosynthesis–leaf nitrogen relationship at ambient and elevated atmospheric carbon dioxide: A meta-analysis. *Global Change Biology*, 5(3), 331–346. <https://doi.org/10.1046/j.1365-2486.1999.00234.x>
- Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., ... Prunet, P. (2016). A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle. *Geoscientific Model Development*, 9(9), 3321–3346. <https://doi.org/10.5194/gmd-9-3321-2016>
- Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., ... Zhu, X. (2019). Plant phenology and global climate change: Current progresses and challenges. *Global Change Biology*, 25(6), 1922–1940. <https://doi.org/10.1111/gcb.14619>
- Piao, S., Liu, Z., Wang, T., Peng, S., Ciais, P., Huang, M., ... Tans, P. P. (2017). Weakening temperature control on the interannual variations of spring carbon uptake across northern lands. *Nature Climate Change*, 7, 359–363. <https://doi.org/10.1038/nclimate3277>
- Piao, S., Nan, H., Huntingford, C., Ciais, P., Friedlingstein, P., Sitch, S., ... Chen, A. (2014). Evidence for a weakening relationship between interannual temperature variability and northern vegetation activity. *Nature Communications*, 5, 5018. <https://doi.org/10.1038/ncomms6018>

- Piao, S., Wang, X., Park, T., Chen, C., Lian, X. U., He, Y., ... Myneni, R. B. (2020). Characteristics, drivers and feedbacks of global greening. *Nature Reviews Earth & Environment*, 1, 14–27. <https://doi.org/10.1038/s43017-019-0001-x>
- Plaza, C., Pegoraro, E., Bracho, R., Celis, G., Crummer, K. G., Hutchings, J. A., ... Schuur, E. A. G. (2019). Direct observation of permafrost degradation and rapid soil carbon loss in tundra. *Nature Geoscience*, 12, 627–631. <https://doi.org/10.1038/s41561-019-0387-6>
- Poorter, H., Fiorani, F., Pieruschka, R., Wojciechowski, T., van der Putten, W. H., Kleyer, M., ... Postma, J. (2016). Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. *New Phytologist*, 212(4), 838–855. <https://doi.org/10.1111/nph.14243>
- Pratihast, A. K., DeVries, B., Avitabile, V., de Bruin, S., Herold, M., & Bergsma, A. (2016). Design and implementation of an interactive web-based near real-time forest monitoring system. *PLoS One*, 11(3), e0150935. <https://doi.org/10.1371/journal.pone.0150935>
- Pu, R., Li, Z., Gong, P., Csiszar, I., Fraser, R., Hao, W.-M., ... Weng, F. (2007). Development and analysis of a 12-year daily 1-km forest fire dataset across North America from NOAA/AVHRR data. *Remote Sensing of Environment*, 108(2), 198–208. <https://doi.org/10.1016/j.rse.2006.02.027>
- Pulvirenti, L., Squicciarino, G., Fiori, E., Fiorucci, P., Ferraris, L., Negro, D., ... Puca, S. (2020). An automatic processing chain for near real-time mapping of burned forest areas using sentinel-2 data. *Remote Sensing*, 12(4), 674. <https://doi.org/10.3390/rs12040674>
- Qiao, Y., Wang, J., Liu, H., Huang, K., Yang, Q., Lu, R., ... Xia, J. (2020). Depth-dependent soil C-N-P stoichiometry in a mature subtropical broadleaf forest. *Geoderma*, 370(1), 114357. <https://doi.org/10.1016/j.geoderma.2020.114357>
- Queen, J. P., Quinn, G. P., & Keough, M. J. (2002). *Experimental design and data analysis for biologists*. Cambridge: Cambridge University Press.
- Raczka, B., Dietze, M. C., Serbin, S. P., & Davis, K. J. (2018). What limits predictive certainty of long-term carbon uptake? *Journal of Geophysical Research: Biogeosciences*, 123(12), 3570–3588. <https://doi.org/10.1029/2018JG004504>
- Rafique, R., Xia, J., Hararuk, O., Asrar, G. R., Leng, G., Wang, Y., & Luo, Y. (2016). Divergent predictions of carbon storage between two global land models: Attribution of the causes through traceability analysis. *Earth System Dynamics*, 7(3), 649–658. <https://doi.org/10.5194/esd-7-649-2016>
- Raupach, M. R., Rayner, P. J., Barrett, D. J., DeFries, R. S., Heimann, M., Ojima, D. S., ... Schimmlus, C. C. (2005). Model-data synthesis in terrestrial carbon observation: Methods, data requirements and data uncertainty specifications. *Global Change Biology*, 11(3), 378–397. <https://doi.org/10.1111/j.1365-2486.2005.00917.x>
- Razgour, O., Forester, B., Taggart, J. B., Bekaert, M., Juste, J., Ibáñez, C., & Manel, S. (2019). Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proceedings of the National Academy of Sciences*, 116(21), 10418–10423. <http://dx.doi.org/10.1073/pnas.1820663116>
- Reichstein, M., Bahn, M., Ciais, P., Frank, D., Mahecha, M. D., Seneviratne, S. I., ... Wattenbach, M. (2013). Climate extremes and the carbon cycle. *Nature*, 500(7462), 287–295. <https://doi.org/10.1038/nature12350>
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743), 195. <https://doi.org/10.1038/s41586-019-0912-1>
- Richardson, A. D., Hufkens, K., Milliman, T., Aubrecht, D. M., Furze, M. E., Seyednasrollah, B., ... Hanson, P. J. (2018). Ecosystem warming extends vegetation activity but heightens vulnerability to cold temperatures. *Nature*, 560(7718), 368–371. <https://doi.org/10.1038/s41586-018-0399-1>
- Rodríguez-Trelles, F., Tarrío, R., & Santos, M. (2013). Genome-wide evolutionary response to a heat wave in *Drosophila*. *Biology Letters*, 9, 20130228. <https://doi.org/10.1098/rsbl.2013.0228>
- Running, S. (2015). MOD17A2H MODIS/Terra gross primary productivity 8-day L4 global 500m SIN grid V006. <https://doi.org/10.5067/MODIS/MOD17A2H.006>
- Ryu, Y., Berry, J. A., & Baldocchi, D. D. (2019). What is global photosynthesis? History, uncertainties and opportunities. *Remote Sensing of Environment*, 223, 95–114. <https://doi.org/10.1016/j.rse.2019.01.016>
- Sage, R. F. (2020). Global change biology: A primer. *Global Change Biology*, 26(1), 3–30. <https://doi.org/10.1111/gcb.14893>
- Sayers, E. W., Cavanaugh, M., Clark, K., Ostell, J., Pruitt, K. D., & Karsch-Mizrachi, I. (2019). GenBank. *Nucleic Acids Research*, 47(D1), D94–D99. <https://doi.org/10.1093/nar/gky989>
- Schimel, D., Schneider, F. D., & JPL Carbon and Ecosystem Participants. (2019). Flux towers in the sky: Global ecology from space. *New Phytologist*, 224(2), 570–584. <https://doi.org/10.1111/nph.15934>
- Schuur, E. A. G., & Abbott, B. (2011). High risk of permafrost thaw. *Nature*, 480(7375), 32–33. <https://doi.org/10.1038/480032a>
- Schwalm, C. R., Williams, C. A., Schaefer, K., Baldocchi, D., Black, T. A., Goldstein, A. H., ... Scott, R. L. (2012). Reduction in carbon uptake during turn of the century drought in western North America. *Nature Geoscience*, 5(8), 551–556. <https://doi.org/10.1038/ngeo1529>
- Seyednasrollah, B., Young, A. M., Hufkens, K., Milliman, T., Friedl, M. A., Frohling, S., & Richardson, A. D. (2019). Tracking vegetation phenology across diverse biomes using Version 2.0 of the PhenoCam Dataset. *Scientific Data*, 6(1), 1–11. <http://dx.doi.org/10.1038/s41597-019-0229-9>
- Shao, J., Yuan, T., Li, Z., Li, N., Liu, H., Bai, S. H., ... Zhou, X. (2019). Plant evolutionary history mainly explains the variance in biomass responses to climate warming at a global scale. *New Phytologist*, 222(3), 1338–1351. <https://doi.org/10.1111/nph.15695>
- Shao, J., Zhou, X., Luo, Y., Li, B. O., Aurela, M., Billesbach, D., ... Zhang, J. (2015). Biotic and climatic controls on interannual variability in carbon fluxes across terrestrial ecosystems. *Agricultural & Forest Meteorology*, 205, 11–22. <https://doi.org/10.1016/j.agrfor.2015.02.007>
- Sheffield, J., Ferguson, C. R., Troy, T. J., Wood, E. F., & McCabe, M. F. (2009). Closing the terrestrial water budget from satellite remote sensing. *Geographical Research Letters*, 36, L07403. <https://doi.org/10.1029/2009GL037338>
- Sherry, R. A., Zhou, X., Gu, S., Arnone, J. A., Schimel, D. S., Verburg, P. S., ... Luo, Y. (2007). Divergence of reproductive phenology under climate warming. *Proceedings of the National Academy of Sciences of the United States of America*, 104(1), 198–202. <https://doi.org/10.1073/pnas.0605642104>
- Shi, Z., Xu, X., Souza, L., Wilcox, K., Jiang, L., Liang, J., ... Luo, Y. (2016). Dual mechanisms regulate ecosystem stability under decade-long warming and hay harvest. *Nature Communications*, 7, 11973. <https://doi.org/10.1038/ncomms11973>
- Shugart, H. H., & Woodward, F. I. (2011). *Global change and the terrestrial biosphere: Achievements and challenges*. Hoboken, NJ: John Wiley & Sons.
- Siepielski, A. M., Morrissey, M. B., Buoro, M., Carlson, S. M., Caruso, C. M., Clegg, S. M., ... MacColl, A. D. C. (2017). Precipitation drives global variation in natural selection. *Science*, 355(6328), 959–962. <http://dx.doi.org/10.1126/science.aag2773>
- Song, J., Wan, S., Piao, S., Knapp, A. K., Classen, A. T., Vicca, S., ... Zheng, M. (2019). A meta-analysis of 1,119 manipulative experiments on terrestrial carbon-cycling responses to global change. *Nature Ecology & Evolution*, 3(9), 1309–1320. <https://doi.org/10.1038/s41559-019-0958-3>
- Soudzilovskaia, N. A., Douma, J. C., Akhmetzhanova, A. A., van Bodegom, P. M., Cornwell, W. K., Moens, E. J., ... Cornelissen, H. C. (2015). Global patterns of plant root colonization intensify by mycorrhizal fungi explained by climate and soil chemistry. *Global Ecology & Biogeography*, 24, 371–382. <https://doi.org/10.1111/geb.12272>

- Sperschneider, J. (2019). Machine learning in plant-pathogen interactions: Empowering biological predictions from field scale to genome scale. *New Phytologist*, in press. <https://doi.org/10.1111/nph.15771>
- Stockhause, M., & Lautenschlager, M. (2017). CMIP6 data citation of evolving data. *Data Science Journal*, 16, 30. <https://doi.org/10.5334/dsj-2017-030>
- Sun, C. C., Royle, J. A., & Fuller, A. K. (2019). Incorporating citizen science data in spatially explicit integrated population models. *Ecology*, 100(9), e02777. <https://doi.org/10.1002/ecy.2777>
- Tansley, A. G. (1935). The use and abuse of vegetational concepts and terms. *Ecology*, 16(3), 284–307. <https://doi.org/10.2307/1930070>
- Tao, F., Zhou, Z., Huang, Y., Li, Q., Lu, X., Ma, S., ... Luo, Y. (2020). Deep learning optimizes data-driven representation of soil organic carbon in Earth system model over the conterminous United States. *Frontiers in Big Data*, 3, 17. <https://doi.org/10.3389/fdata.2020.00017>
- Templ, B., Koch, E., Bolmgren, K., Ungersböck, M., Paul, A., Scheifinger, H., ... Zust, A. (2018). Pan European Phenological database (PEP725): A single point of access for European data. *International Journal of Biometeorology*, 62(6), 1109–1113. <https://doi.org/10.1007/s00484-018-1512-8>
- Thums, M., Fernández-Gracia, J., Sequeira, A. M. M., Eguíluz, V. M., Duarte, C. M., & Meekan, M. G. (2018). How big data fast tracked human mobility research and the lessons for animal movement ecology. *Frontiers in Marine Science*, 5, 21. <https://doi.org/10.3389/fmars.2018.00021>
- Treharne, R., Bjerke, J. W., Tømmervik, H., Stendardi, L., & Phoenix, G. K. (2019). Arctic browning: Impacts of extreme climatic events on heathland ecosystem CO₂ fluxes. *Global Change Biology*, 25, 489–503. <https://doi.org/10.1111/gcb.14500>
- Van Ackere, S., Verbeurgt, J., De Sloover, L., Gautama, S., De Wulf, A., & De Maeyer, P. (2019). A review of the internet of floods: Near real-time detection of a flood event and its impact. *Water*, 11(11), 2275. <https://doi.org/10.3390/w11112275>
- van der Molen, M. K., Dolman, A. J., Ciais, P., Eglin, T., Gobron, N., Law, B. E., ... Wang, G. (2011). Drought and ecosystem carbon cycling. *Agricultural and Forest Meteorology*, 151(7), 765–773. <https://doi.org/10.1016/j.agrformet.2011.01.018>
- van Gestel, N., Shi, Z., van Groenigen, K. J., Osenberg, C. W., Andresen, L. C., Dukes, J. S., & Hungate, B. A. (2018). Predicting soil carbon loss with warming. *Nature*, 554(7693), E4–E5. <http://dx.doi.org/10.1038/nature25745>
- Verbesselt, J., Zeileis, A., & Herold, M. (2012). Near real-time disturbance detection using satellite image time series. *Remote Sensing of Environment*, 123, 98–108. <https://doi.org/10.1016/j.rse.2012.02.022>
- Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., & Garnier, E. (2007). Let the concept of trait be functional! *Oikos*, 116(5), 882–892. <https://doi.org/10.1111/j.0030-1299.2007.15559.x>
- Vogt, R. J., St-Gelais, N. F., Bogard, M. J., Beisner, B. E., & del Giorgio, P. A. (2017). Surface water CO₂ concentration influences phytoplankton production but not community composition across boreal lakes. *Ecology Letters*, 20, 1395–1404. <https://doi.org/10.1111/ele.12835>
- Walker, A. P., Aranda, I., Beckerman, A. P., Bown, H., Cernusak, L. A., Dang, Q. L., ... Wullschlegel, S. D. (2014). A global data set of leaf photosynthetic rates, leaf N and P, and specific leaf area. Oak Ridge, TN: ORNL DAAC. <https://doi.org/10.3334/ORNLDAAC/1224>
- Walker, A. P., Carter, K. R., Gu, L., Hanson, P. J., Malhotra, A., Norby, R. J., ... Weston, D. J. (2017). Biophysical drivers of seasonal variability in Sphagnum gross primary production in a northern temperate bog. *Journal of Geophysical Research: Biogeosciences*, 122(5), 1078–1097. <https://doi.org/10.1002/2016JG003711>
- Walker, A. P., De Kauwe, M. G., Medlyn, B. E., Zaehe, S., Iversen, C. M., Asao, S., ... Norby, R. J. (2019). Decadal biomass increment in early secondary succession woody ecosystems is increased by CO₂ enrichment. *Nature Communications*, 10, 454. <https://doi.org/10.1038/s41467-019-08348-1>
- Wang, J., Yang, Q., Qiao, Y., Zhai, D., Jiang, L., Liang, G., ... Xia, J. (2019). Relative contributions of biotic and abiotic factors to the spatial variation of litter stock in a mature subtropical forest. *Journal of Plant Ecology*, 12(4), 769–780. <https://doi.org/10.1093/jpe/rtz018>
- Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, 126, 3–13. <https://doi.org/10.1016/j.techfore.2015.12.019>
- Wang, Y., Ma, A., Liu, G., Ma, J., Wei, J., Zhou, H., ... Zhuang, G. (2020). Potential feedback mediated by soil microbiome response to warming in a glacier forefield. *Global Change Biology*, 26(2), 697–708. <https://doi.org/10.1111/gcb.14936>
- Wang, Y. P., Trudinger, C. M., & Enting, I. G. (2009). A review of applications of model-data fusion to studies of terrestrial carbon fluxes at different scales. *Agricultural and Forest Meteorology*, 149(11), 1829–1842. <https://doi.org/10.1016/j.agrformet.2009.07.009>
- Wegner, K. M., Kalbe, M., Milinski, M., & Reusch, T. B. H. (2008). Mortality selection during the 2003 European heat wave in three-spined sticklebacks: Effects of parasites and MHC genotype. *BMC Evolutionary Biology*, 8, 124. <https://doi.org/10.1186/1471-2148-8-124>
- Wei, N., Cui, E., Huang, K., Du, Z., Zhou, J., Xu, X., & Xia, J. (2019). Decadal stabilization of soil Inorganic nitrogen as a benchmark for global land models. *Journal of Advances in Modeling Earth Systems*, 11(4), 1088–1099. <http://dx.doi.org/10.1029/2019ms001633>
- Whittaker, R. J. (2010). Meta-analyses and mega-mistakes: Calling time on meta-analysis of the species richness-productivity relationship. *Ecology*, 91, 2522–2533. <https://doi.org/10.1890/08-0968.1>
- Wieder, W. R., Cleveland, C. C., Smith, W. K., & Todd-Brown, K. (2015). Future productivity and carbon storage limited by terrestrial nutrient availability. *Nature Geoscience*, 8(6), 441–444. <https://doi.org/10.1038/ngeo2413>
- Wilczek, A. M., Cooper, M. D., Korves, T. M., & Schmitt, J. (2014). Lagging adaptation to warming climate in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences*, 111(22), 7906–7913. <http://dx.doi.org/10.1073/pnas.1406314111>
- Wohlfahrt, G., Tomelleri, E., & Hammerle, A. (2019). The urban imprint on plant phenology. *Nature Ecology & Evolution*, 3(12), 1668–1674. <https://doi.org/10.1038/s41559-019-1017-9>
- Wu, Z., Dijkstra, P., Koch, G. W., Penuelas, J., & Hungate, B. A. (2011). Responses of terrestrial ecosystems to temperature and precipitation change: A meta-analysis of experimental manipulation. *Global Change Biology*, 17(2), 927–942. <http://dx.doi.org/10.1111/j.1365-2486.2010.02302.x>
- Wüest, R. O., Zimmermann, N. E., Zurell, D., Alexander, J. M., Fritz, S. A., Hof, C., ... Karger, D. N. (2019). Macroecology in the age of Big Data – Where to go from here? *Journal of Biogeography*, 47(1), 1–12. <https://doi.org/10.1111/jbi.13633>
- Xia, J., Chen, J., Piao, S., Ciais, P., Luo, Y., & Wan, S. (2014). Terrestrial carbon cycle affected by non-uniform climate warming. *Nature Geoscience*, 7, 173–180. <https://doi.org/10.1038/ngeo2093>
- Xia, J., Luo, Y., Wang, Y.-P., & Hararuk, O. (2013). Traceable components of terrestrial carbon storage capacity in biogeochemical models. *Global Change Biology*, 19(7), 2104–2116. <https://doi.org/10.1111/gcb.12172>
- Xia, J. Y., Luo, Y. Q., Wang, Y.-P., Weng, E. S., & Hararuk, O. (2012). A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state. *Geoscientific Model Development*, 5(5), 1259–1271. <https://doi.org/10.5194/gmd-5-1259-2012>
- Xia, J., McGuire, A. D., Lawrence, D., Burke, E., Chen, G., Chen, X., ... Luo, Y. (2017). Terrestrial ecosystem model performance in simulating productivity and its vulnerability to climate change in the northern permafrost region. *Journal of Geophysical Research: Biogeosciences*, 122(2), 430–446. <https://doi.org/10.1002/2016JG003384>

- Xia, J., Niu, S., Ciais, P., Janssens, I. A., Chen, J., Ammann, C., ... Luo, Y. (2015). Joint control of terrestrial gross primary productivity by plant phenology and physiology. *Proceedings of the National Academy of Sciences of the United States of America*, 112(9), 2788–2793. <https://doi.org/10.1073/pnas.1413090112>
- Xia, J., & Wan, S. (2008). Global response patterns of terrestrial plant species to nitrogen addition. *New Phytologist*, 179(2), 428–439. <https://doi.org/10.1111/j.1469-8137.2008.02488.x>
- Xu, L., Myneni, R. B., Chapin III, F. S., Callaghan, T. V., Pinzon, J. E., Tucker, C. J., & Stroeve, J. C. (2013). Temperature and vegetation seasonality diminishment over northern lands. *Nature Climate Change*, 3(6), 581–586. <http://dx.doi.org/10.1038/nclimate1836>
- Xu, X., Lu, R., Yan, L., & Xia, J. (2019). PlantNE: A global database of plant biomass from nitrogen-addition experiments. *Ecology*, 100, e02840. <https://doi.org/10.1002/ecy.2840>
- Xu, X., Yan, L., & Xia, J. (2019). A threefold difference in plant growth response to nitrogen addition between the laboratory and field experiments. *Ecosphere*, 10(1), e02572. <https://doi.org/10.1002/ecs2.2572>
- Xue, K., Xie, J., Zhou, A., Liu, F., Li, D., Wu, L., ... Zhou, J. (2016). Warming alters expressions of microbial functional genes important to ecosystem functioning. *Frontiers in Microbiology*, 7, 668. <https://doi.org/10.3389/fmicb.2016.00668>
- Yang, J., Tian, H., Pan, S., Chen, G., Zhang, B., & Dangal, S. (2018). Amazon drought and forest response: Largely reduced forest photosynthesis but slightly increased canopy greenness during the extreme drought of 2015/2016. *Global Change Biology*, 24(5), 1919–1934. <https://doi.org/10.1111/gcb.14056>
- Yau, C., & Campbell, K. (2019). Bayesian statistical learning for big data biology. *Biophysical Reviews*, 11(1), 95–102. <https://doi.org/10.1007/s12551-019-00499-1>
- You, M., Yue, Z., He, W., Yang, X., Yang, G., Xie, M., ... Wang, J. (2013). A heterozygous moth genome provides insights into herbivory and detoxification. *Nature Genetics*, 45(2), 220–225. <http://dx.doi.org/10.1038/ng.2524>
- Yu, G., & Yu, X. (2013). Chinese ecosystem research network (CERN) and natural ecosystem protection. *Bulletin of Chinese Academy of Sciences*, 28(2), 275–283.
- Yue, K., Fornara, D. A., Yang, W., Peng, Y., Peng, C., Liu, Z., & Wu, F. (2017). Influence of multiple global change drivers on terrestrial carbon storage: Additive effects are common. *Ecology Letters*, 20(5), 663–672. <https://doi.org/10.1111/ele.12767>
- Zhang, Y., Xiao, X., Wu, X., Zhou, S., Zhang, G., Qin, Y., & Dong, J. (2017). A global moderate resolution dataset of gross primary production of vegetation for 2000–2016. *Scientific Data*, 4(1), 170165. <https://doi.org/10.1038/sdata.2017.165>
- Zhao, K., Popescu, S., Meng, X., Pang, Y., & Agca, M. (2011). Characterizing forest canopy structure with lidar composite metrics and machine learning. *Remote Sensing of Environment*, 115, 1978–1996. <https://doi.org/10.1016/j.rse.2011.04.001>
- Zheng, M., Zhou, Z., Luo, Y., Zhao, P., & Mo, J. (2019). Global pattern and controls of biological nitrogen fixation under nutrient enrichment: A meta-analysis. *Global Change Biology*, 25(9), 3018–3030. <http://dx.doi.org/10.1111/gcb.14705>
- Zhou, G., Luo, Q., Chen, Y., He, M., Zhou, L., Frank, D., ... Zhou, X. (2019). Effects of livestock grazing on grassland carbon storage and release override impacts associated with global climate change. *Global Change Biology*, 25(3), 1119–1132. <http://dx.doi.org/10.1111/gcb.14533>
- Zhou, J., Deng, Y. E., Shen, L., Wen, C., Yan, Q., Ning, D., ... Brown, J. H. (2016). Temperature mediates continental-scale diversity of microbes in forest soils. *Nature Communications*, 7(1), 12083. <https://doi.org/10.1038/ncomms12083>
- Zhou, J., Xia, J., Wei, N., Liu, Y., Bian, C., Bai, Y., & Luo, Y. (2020). TraceME (v1.0) – An online Traceability analysis system for Model Evaluation on land carbon dynamics. *Geoscientific Model Development Discussions*, 1–32. <https://doi.org/10.5194/gmd-2020-76>
- Zhou, S., Liang, J., Lu, X., Li, Q., Jiang, L., Zhang, Y., ... Luo, Y. (2018). Sources of uncertainty in modeled land carbon storage within and across three MIPs: Diagnosis with three new techniques. *Journal of Climate*, 31(7), 2833–2851. <https://doi.org/10.1175/JCLI-D-17-0357.1>
- Zhu, C., Cui, E., & Xia, J. (2020). Both day and night warming reduce tree growth in extremely dry soils. *Environmental Research Letters*, in press. <https://doi.org/10.1088/1748-9326/aba65e>
- Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., Canadell, J. G., ... Zeng, N. (2016). Greening of the Earth and its drivers. *Nature Climate Change*, 6(8), 791–795. <https://doi.org/10.1038/nclimate3004>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Xia J, Wang J, Niu S. Research challenges and opportunities for using big data in global change biology. *Glob Change Biol*. 2020;26:6040–6061. <https://doi.org/10.1111/gcb.15317>