Ecosystem Observation and Research in the Big Data Era: Postprint

Authors: Yu Guirui, He Honglin, Zhou Yuke

Date: 2023-03-19T00:00:00+00:00

Abstract

Global change, sustainable development, and a series of globalized resource and environmental issues are becoming increasingly severe, attracting international attention and emerging as hot topics in scientific research. The research focus of ecology has gradually shifted from small-scale individuals, communities, and ecosystems to regional, continental, and even global scales. Concurrently, with advancements in sensor technology and information network technology, ecosystem observation is transitioning from short-term to long-term monitoring, and toward coordinated observation of macrostructural dynamics and spatiotemporal patterns of ecological services; from traditional fixed-location and small-scale collaborative observation to large-scale integrated networked observation at regional or global scales, and toward space-air-ground integrated three-dimensional observation. Ecology has entered the big data era. How to leverage big data technology to achieve organic integration of traditional process-based ecological research and big data-driven ecological research, advance the development of grand ecological theory, investigate regional and global ecosystem evolution mechanisms, and support theoretical and applied research on sustainable human society development centered on addressing global climate change and maintaining ecosystem functions, represents major challenges and opportunities for ecological observation research in the big data era. This article reviews the current status of ecosystem observation research, discusses key characteristics of ecosystem observation research in the big data context, and recommends adhering to the "Big Science, Big Data" concept to organize and implement the construction of a national ecosystem observation Big Science facility. This would achieve integration of China's ecological observation research with the global ecological observation research system, and realize scientific and technological objectives of observing changes in Earth's life system at regional, national, and global scales, diagnosing ecosystem functional status, understanding ecosystem process mechanisms, maintaining and protecting ecosystem functions, and serving sustainable development of human society.

Full Text

Ecosystem Observation and Research under the Background of Big Data

Institutions: Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences; Key Laboratory of Ecosystem Network Observation and Modeling, Chinese Academy of Sciences; College of Resources and Environment, University of Chinese Academy of Sciences

Keywords: ecological observation network, big data, macroecology, Internet of Things, big science project
DOI: 10.16418/j.issn.1000-3045.2018.08.010

Global environmental issues such as global change and sustainable development have become increasingly severe, emerging as focal points of international concern and scientific research. The emphasis of ecological research has gradually shifted from small-scale studies of individuals, communities, and ecosystems to regional, continental, and even global scales. Concurrently, with the development of sensor technology and information networks, ecosystem observations have transitioned from short-term monitoring to long-term observations of ecosystem macrostructure and the spatio-temporal dynamics of ecological services. Traditional localized and small-scale collaborative observations are evolving toward large-scale networked observations at regional and global scales, as well as integrated three-dimensional sky-space-ground observations. Ecology has entered the era of big data. A major challenge and opportunity lies in applying big data technologies to achieve organic integration between traditional process-based ecological research and big-data-driven ecological research, thereby advancing the development of grand ecological theories, investigating regional and global ecosystem evolution mechanisms, and supporting theoretical and applied research on sustainable human development centered on addressing global climate change and maintaining ecosystem functions. This paper reviews the current status of ecosystem observation and research, discusses the key characteristics of ecosystem observation and research in the context of big data, and proposes organizing and implementing a national ecosystem observation big science project following the "big science, big data" concept. Such an initiative would integrate China's ecological observation research with the global ecological observation system, enabling the monitoring of Earth's living system changes at regional, national, and global scales, diagnosing ecosystem functional status, understanding ecosystem processes and mechanisms, maintaining and protecting ecosystem functions, and serving the scientific and technological goals of sustainable human development.

As global environmental issues such as the community with a shared future for mankind, sustainable development, and ecological security become practical challenges facing human society in the new era, ecological civilization construction has become a millennium-long plan for the sustainable development of the Chinese nation, placing new demands on ecological research. How to achieve observations from small to large scales (global) and from short-term to long-term, obtain global-scale ecosystem observation data, understand regional and global ecosystem dynamic change patterns, and protect and restore ecosystems are major challenges currently facing ecosystem observation research. The implementation of the International Biological Program (1964-1974) gradually formed the "big science, big data" concept in ecology, which involves implementing large-scale scientific projects to obtain massive observation data and conduct collaborative scientific research. This subsequently gave rise to the Long-Term Ecological Research (LTER) network [1], marking ecology's entry into the "big science, big data" era. Currently, regional/global-scale networked long-term ecosystem observations have become important platforms for obtaining large-scale ecological information, ecosystem and global change have become hot topics in ecological research, and cross-scale integration has become a frontier issue in ecology.

Current Status of Ecosystem Observation and Research

In recent years, ecological observation technologies have developed rapidly, substantially enhancing the capacity to obtain ecological observation data at both micro and macro scales. For example, the introduction of molecular markers into ecology has revolutionized macro-biological research; stable isotope technology enables tracking of biogeochemical cycles of carbon, nitrogen, and water; and high-throughput field observation techniques such as eddy covariance have enabled direct measurement of ecosystem functional changes. A new generation of global multi-scale ecological observation research networks is gradually taking shape. Representative examples include the U.S. National Ecological Observatory Network (NEON) and the Australian Terrestrial Ecosystem Research Network (TERN), which are based on Internet of Things technology, automated observations, and integration of ground-based and remote sensing observations. These new-generation ecological observation networks employ large numbers of ecological sensors, cover diverse observation indicators, and collect massive amounts of multi-source heterogeneous data, laying a solid foundation for ecology to enter the "big data era."

Characteristics of Ecosystem Observation and Research in the Big Data Era

As ecology continues to expand toward both molecular (ultra-micro) and global (ultra-macro) scales, ecological observations are transitioning from traditional, discontinuous plot-based ground surveys to integrated "sky-space-ground" multi-scale, multi-element, and multi-process observations. The mode of participation is also shifting from single ecology researchers to public involvement through "citizen science" models. Multi-scale, multi-element, and multi-process ecological observation systems supported by Internet of Things technology have become key features of new-generation ecological observation systems. Specifically, by leveraging various IoT communication methods, an integrated "sky-space-ground" tower-group collaborative observation system is formed (Fig. 2a [Figure 2: see original paper]), integrating sensors from small-scale individual instruments and flux towers to regional/global-scale drones and satellites. Through IoT technologies such as WiFi and ZigBee, convenient local area observation networks are established, enabling real-time data transmission, remote control, and intelligent alerts for ecological observations, and ensuring stable acquisition of multi-scale (from molecular to global), multi-element (water, soil, atmosphere, organisms, etc.), and multi-process (carbon cycle, water cycle, energy cycle, etc.) ecosystem observation data.

Large-scale ecological research requires participation from researchers with diverse disciplinary backgrounds and cannot succeed without public involvement (Fig. 2b [Figure 2: see original paper]). The development of Internet technology has promoted the advancement of citizen science. In recent years, numerous citizen science project platforms have been established worldwide, utilizing "citizen scientists" to collect more environmental data. For example, the BudBurst project in the United States engages people across the nation to collaboratively collect plant phenological data, helping to discover how plants respond to environmental changes. Citizen science projects have gradually become another important source of ecological big data.

Multi-source, Multi-scale Data Integration, Mining, and Simulation Prediction as Key Approaches in Big Data Era Ecology

The rapid expansion of remote sensing data, long-term positioning observation data, experiments, and model simulation data has strongly promoted the development of ecological data mining and analysis technologies. Multi-source, multi-scale integration, mining, and simulation prediction have become essential approaches for ecological research in the big data era. Big data-supported ecological data integration and mining methods include three major categories: Meta-analysis methods; Data-driven mining methods; and Process-based model-data fusion methods. Meta-analysis primarily involves quantitative synthesis of multiple independent studies with common research objectives from existing literature, analyzing inter-study differences and comprehensively evaluating results. Artificial intelligence data mining technologies, represented by deep learning, can establish deep neural network-style mapping relationships among various ecological elements, potentially generating new insights within existing ecological frameworks and subsequently inferring underlying ecological mechanisms. Since it is impossible to comprehensively obtain or predict regional/global ecosystem conditions through observations and controlled experiments alone, computer simulation of large-scale ecosystem structural and functional changes has become an important pathway for ecological modeling and prediction. Although the volume of ecological observation data is increasing daily, model simulation uncertainty remains significant. Process-based model-data fusion methods provide new avenues for applying massive datasets to model evaluation, benchmarking, and constraint to reduce uncertainty. These methods offer important means for timely and effective integration and mining of observation and experimental data from different sites, regions, time series, ecosystem processes, and elements, revealing universal patterns in ecosystem processes, variation, and responses to environmental factors, and serving ecosystem management and decision-making.

Ecological Big Data Driving the Development of Ecological "Grand Theory" and Emerging Disciplines

Due to the complexity of ecological research objects and the lack of long-term, massive observation data, researchers face great uncertainty in understanding ecological processes and patterns. Ecology has yet to develop a unified theoretical framework that is widely accepted and can maximally explain and accurately predict ecological phenomena and processes across different scales [6]. Influenced by the data deluge, a new paradigm (the fourth paradigm) centered on data—where scientific research is conceived, designed, and implemented through processing and analyzing massive datasets to obtain scientific discoveries—is gradually emerging in ecology. Simultaneously, the big data era demands the development of efficient grand theories that can generate numerous theoretical predictions from fundamental principles with minimal parameters [7]. Due to the diversity of ecosystem components and the complexity of their relationships, regardless of data volume, without the support of grand theory, big data will largely lose its power and effectiveness, forcing reliance on simple inductive analysis based on statistical correlations that can never accurately predict ecosystem changes. Only with the support of massive multi-scale, multi-element, and multi-process ecological data, enabling accurate understanding of ecosystem processes and patterns, can a unified theoretical framework be developed. Integrating multi-source, multi-scale data integration technologies and ecological simulation prediction techniques can improve the predictive capacity of ecosystem properties and processes, and promote the development of new ecological disciplines such as macroecology and predictive ecology.

Data Open Sharing Culture Urgently Needs Cultivation and Strengthening

In the big data era, only by forming an open, shared, and collaborative big data culture can the true value of ecological big data be realized. Observation data is the lifeline of ecological research, and ecologists treasure their observational data. In the big data era, only by breaking the traditional model of data closure and forming an open and shared "data thinking" model can large-scale ecological problems be solved. The deep application of novel data acquisition and intellectual property protection methods such as data crowdsourcing, data publication, and data identification will also promote data open sharing in the ecological field.

Prospects for China's Ecosystem Observation Research in the Context of Big Data

Since the implementation of long-term ecological research programs in the 1980s, ecological research has been closely coupled with the "big science, big data" concept. The Chinese Ecosystem Research Network (CERN), established in 1988, has formed an observation and research network comprising 44 stations covering different ecosystem types in China, accumulating vast amounts of scientific data. In the big data era, on one hand, we should utilize big data technologies to fully integrate and mine existing data resources to advance ecological research; simultaneously, we should follow the "big science, big data" concept, fully apply sensor, IoT, and big data technologies, implement ecological big science project plans, achieve leapfrog development in ecosystem observation research, and effectively support national ecological civilization construction.

(1) Actively promote integrated ecological data mining research based on data open sharing. China has conducted long-term observations and research on ecosystem structure and function, accumulating substantial data. For example, CERN has collected long-term observation and experimental data from different ecosystems across China; the ChinaFLUX network has conducted 15 years of carbon and water flux monitoring; and the implementation of the "Big Earth Data Science Engineering" project has provided advanced big data integration and mining tools, laying the foundation for comprehensive integrated research on ecosystem structure, function, processes, and major scientific questions at the national scale.

(2) Actively promote the construction of the big science project "China Eco-observation and Experiment Network" (CEOBEX) (Fig. 3 [Figure 3: see original paper]). This initiative aims to achieve national-scale integration of sky-space-ground observations, integration of observation-experiment-simulation, and coordinated multi-scale-multi-element-multi-process observations. It will establish distributed, networked observation and experimental scientific facilities covering different ecological regions across China, enabling national-scale networked observations, networked experiments, and simulation and prediction. This will integrate China's ecological observation research with the global ecological observation research system, enabling the monitoring of Earth's living system changes at regional, national, and global scales, diagnosing ecosystem functional status, understanding ecosystem processes and mechanisms, maintaining and protecting ecosystem functions, and serving the scientific and technological goals of sustainable human development.

Fig. 3 [Figure 3: see original paper] Ecological Big Science Project: China Eco-observation and Experiment Network (CEOBEX)

ChinaRxiv 中X

Abstract

Global issues such as global change and sustainable development have become increasingly prominent, emerging as hotspots in academic literature. Consequently, ecological research has evolved from local scales to regional, continental, and even global scales. Simultaneously, advancements in sensor and network technologies have transformed ecosystem observations from short-term monitoring to long-term observations of ecosystem macrostructure and spatio-temporal patterns of ecosystem services, and from small-scale, localized observations to regional and global networked observations with integrated space-ground systems. In summary, ecological research has entered the era of big data. A critical challenge and opportunity lies in applying new technologies, integrating multi-disciplinary knowledge, merging multi-scale and multi-source data, and developing big-data-driven ecological models to achieve organic integration between traditional process-based ecological research and big-data-driven ecological research. Big-data-driven ecology should promote the development of ecological theory, advance mechanistic understanding of regional and global ecosystem evolution, and support sustainable human development under the pressing concerns of global climate change, biodiversity protection, and ecosystem function maintenance. This paper reviews the current status of ecosystem observation and research, discusses key characteristics of ecosystem observation and research in the big data context, and proposes organizing and implementing a national ecosystem observation project utilizing new technologies such as intelligent interconnection of big data and Internet of Things, data fusion-mining, and high-performance computing. This project aims to integrate China's ecological observation research with the global ecological observation system, monitor changes in Earth's living systems at regional, national, and global scales, diagnose ecosystem functional status, understand ecosystem processes and mechanisms, maintain and protect ecosystem functions, and ultimately serve the scientific and technological goals of sustainable human development.

Keywords: ecological observation network, big data, macroecology, Internet of Things, big science project

Author Biography

YU Guirui, Deputy Director of the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences; Director of the CAS Key Laboratory of Ecosystem Network Observation and Modeling. Dr. Yu also serves as Director of the Synthesis Research Center of the Chinese Ecosystem Research Network (CERN), Director of the Long-term Ecological Committee of the Ecological Society of China, and member of the AsiaFlux Network Steering Committee. He has been selected for the CAS "Hundred Talents Program" and has led over ten major projects, including the CAS Knowledge Innovation Program Key Project and National Key Research and Development Program. He has published more than 110 SCI papers and ten monographs, receiving the First Prize of the National Science and Technology Progress Award and the First

Prize of the Science and Technology Award from the Ministry of Environmental Protection. His research focuses on terrestrial ecosystem flux observation theory and methods, ecosystem carbon and water cycles and their management, plant ecophysiology, and water movement in soil-vegetation-atmosphere systems. E-mail: yugr@igsnrr.ac.cn

References

1. Aronova E, Baker K S, Oreskes N. Big science and big data in biology: From the international geophysical year through the international biological program to the long term ecological research (lter) network, 1957—-present. Hist Stud Nat Sci, 2010, 40: 183-224.
2. Hampton S E, Strasser C A, Tewksbury J J, et al. Big data and the future of ecology. Frontiers in Ecology and the Environment, 2013, 11(3), 156–162.
3. Jonathan T O, Gerald A M, Sandrine Bony et al. Special online collection: dealing with data. Science, 2011, 331(6018): 639-806.
4. 戴圣骐, 赵斌. 大数据时代下的生态系统观测发展趋势与挑战. 生物多样性, 2016, 24: 85-94.
5. Millennium Ecosystem Assessment. Ecosystems and human well-being: general synthesis. Washington, DC: Island Press and World Resource Institute, 2005: 77-101.
6. Anonymous. An elegant chaos. Nature, 2014, 507(7491): 139.
7. Marquet P A, Allen A P, Brown J H, et al. On theory in ecology. Bioscience, 2014, 64: 701-710.

Note: Figure translations are in progress. See original paper for figures.