Sci. Bull. (2016) 61(7):505–513
DOI 10.1007/s11434-016-1041-y

CrossMark

www.scibull.com
www.springer.com/scp

Progress

Earth Sciences

Big Earth Data from space: a new engine for Earth science

Huadong Guo · Lizhe Wang · Dong Liang

Received: 4 December 2015 / Revised: 23 February 2016 / Accepted: 24 February 2016 / Published online: 16 March 2016

Abstract Big data is a strategic highland in the era of knowledge-driven economies, and it is also a new type of strategic resource for all nations. Big data collected from space for Earth observation—so-called Big Earth Data—is creating new opportunities for the Earth sciences and revolutionizing the innovation of methodologies and thought patterns. It has potential to advance in-depth development of Earth sciences and bring more exciting scientific discoveries. The Academic Divisions of the Chinese Academy of Sciences Forum on Frontiers of Science and Technology for Big Earth Data from Space was held in Beijing in June of 2015. The forum analyzed the development of Earth observation technology and big data, explored the concepts and scientific connotations of Big Earth Data from space, discussed the correlation between Big Earth Data and Digital Earth, and dissected the potential of Big Earth Data from space to promote scientific discovery in the Earth sciences, especially concerning global changes.

Keywords Big data · Big Earth Data from space · Digital Earth · Earth sciences · Earth observation · Scientific big data · Data-intensive science

1 Introduction

A data explosion is occurring, caused by a mix of social activities and a new information technology revolution. Data volume has increased rapidly, and data types are varied and complex, exceeding the capabilities of traditional data management systems and processing modes. Hence, the concept of big data has arisen at this moment in history. Big data is a term describing large datasets which exceed conventional processing capabilities. It also comprises a methodology and wave of techniques that use data to solve problems and make new discoveries. According to an article by the International Data Corporation (IDC), all the digital data created, replicated, and consumed will be more than double every 2 years. There were about 4.4 zettabytes (ZB) of data created, replicated, and consumed worldwide in 2013. The IDC estimates that by the year 2020, this figure will reach 44 ZB [1]. China's global share will increase from 13 % in 2012 to 21 % in 2020 [2]. In the current digital era, national competitiveness will be reflected in the size, quality, and applicability of a country's data. Big data has become a manifestation of informational sovereignty; it will be the next topic of international debate and will play a significant role in border, coastal, and air defenses [3].

Big data has begun to significantly influence global production, circulation, distribution, and consumption patterns. It is changing humankind's production methods, lifestyles, mechanisms of economic operation, and country governance models. Big data is a strategic highland in the era of knowledge-driven economies, and it is a new type of strategic resource for all nations. In the near future, the competition around big data will not only determine international patterns in the information industry, but it will also profoundly impact economic development, national security, scientific and technical progress, and comprehensive competitiveness. As a branch of big data, scientific big data is the typical representative of data-intensive science. It has changed the scientific world, evolving from experimental science, to theoretical science, and then to

H. Guo (✉) · L. Wang · D. Liang
Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China
e-mail: hdguo@radi.ac.cn

computational science. It is transforming science from model-driven to data-driven, and it is bringing innovation to scientific research methodologies.

Scientific big data is the new key to understanding nature and the new engine of scientific discovery, and it is also driving the development of disciplinary innovation. At present, the top ten countries giving the most attention to big data research globally are the United States, China, Germany, the United Kingdom, Japan, India, Canada, France, Australia, and South Korea. Heretofore, the United States has had the largest percentage of researchers in big data at about 34%. China is not far behind at 23%. Regarding the top 25 research institutes with the highest number of research publications about big data, there are 16 universities from the United States and 6 institutes or universities from China. The Chinese Academy of Sciences is the top producer of papers related to big data research [4].

Scientific big data is primarily produced by large experimental facilities, detection equipment, sensors, and computer simulation processes. After nearly half a century of development, Earth observation technology has provided a new vision and new methods for Earth science research. It has further deepened humanity's understanding of Earth, especially its macro-knowledge. Big Earth Data from space can be defined as big data obtained via the methodologies of Earth observation, such as data from spaceborne, airborne, and ground sensors. It comprises a massive amount of data detailing the processes of Earth's environment and change, and the interaction between humans and the Earth. The data are acquired using space technology, usually satellites, accompanied by ground receiving systems. The data exert a profound influence on the development of Earth sciences.

With the increased use of remote sensing satellites, navigation satellites, geophysical satellites, and various platforms, as well as diversified observation instruments and sensors, Big Earth Data from space is becoming the mainstream of research in big data. It assumes the scientific features of big data, such as massive size, multiple sources, heterogeneity, multi-temporality, multiple scales, and non-stationarity. Big Earth Data from space comprises Earth observation technology, communication technology, and computer technology, bringing new opportunities to Earth science research. It has the potential to promote deep development of the Earth sciences and contribute to important scientific discoveries.

The Academic Divisions of the Chinese Academy of Sciences Forum on Frontiers of Science and Technology for Big Earth Data from Space was held in Beijing on 17–18 June, 2015, with a number of goals: studying the frontier of geospatial information science in the big data era, exploring the concepts and scientific connotations of Big Earth Data from space, discussing the relationship between Big Earth Data from space and the next-generation of Digital Earth, and promoting the development of Earth sciences. It was sponsored by the Academic Divisions of the Chinese Academy of Sciences (CASAD); co-sponsored by the Division of Earth Sciences of CASAD and the Academic Works and Publications Committee of CASAD; and co-organized by the Chinese Academy of Sciences Institute of Remote Sensing and Digital Earth and Science China Press. The forum set up four themes: (1) observation of Big Earth Data from space, (2) methods for processing and knowledge discovery in Big Earth Data from space, (3) Big Earth Data from space in the geoscience disciplines, and (4) Big Earth Data from space in Earth sciences. Also at the forum, twenty-six researchers from the fields of Earth observation, geography, geology, atmospheric sciences, marine science, geophysics, geochemistry, space physics and computer science delivered academic presentations on Big Earth Data from space.

Summarizing the results of the workshop, this paper provides a review of big data topics in Earth observation from space. The remainder of this paper is organized as follows. Section 2 gives an overview of Earth observation big data. Section 3 introduces methods for processing and knowledge discovery using Big Earth Data from space. Section 4 presents the point that Big Earth Data from space drives the development of Earth sciences. Finally, in Section 5 gives a summary as well as the conclusion.

2 Earth Observation Big Data

With the help of spaceborne and airborne Earth observation platforms, humans can uninterruptedly observe the Earth. We can rapidly reproduce and objectively reflect status, phenomena, processes, spatial distribution, and locations within the geosphere through information processing in the service of economic construction and social development. In the twenty-first century, Earth observation technology, as the core of geospatial information science and technology, has become a comprehensive national embodiment of the capacity for scientific and technological achievement, economic strength, and national security. The demand for Earth observation applications and the development of satellite and sensor technologies has grown, substantially increasing the number of Earth observation satellites, as well as performance indexes. The data volume of Earth observation has doubled and redoubled. According to statistics from the global satellite mission generated by the Committee on Earth Observation Satellites (CEOS), from 1962 to 2012, more than 514 Earth observation satellites were launched globally for comprehensive observation of the Earth system, covering the atmosphere, ocean, and land.