

August 2018

## A Project on Big Earth Data Science Engineering

GUO Huadong

*Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China*

### Recommended Citation

Huadong, GUO (2018) "A Project on Big Earth Data Science Engineering," *Bulletin of Chinese Academy of Sciences (Chinese Version)*: Vol. 33 : Iss. 8 , Article 8.

DOI: <https://doi.org/10.16418/j.issn.1000-3045.2018.08.008>

Available at: <https://bulletinofcas.researchcommons.org/journal/vol33/iss8/8>

This Article is brought to you for free and open access by Bulletin of Chinese Academy of Sciences (Chinese Version). It has been accepted for inclusion in Bulletin of Chinese Academy of Sciences (Chinese Version) by an authorized editor of Bulletin of Chinese Academy of Sciences (Chinese Version). For more information, please contact [lcyang@cashq.ac.cn](mailto:lcyang@cashq.ac.cn), [yjwen@cashq.ac.cn](mailto:yjwen@cashq.ac.cn).



# 地球大数据科学与工程

郭华东

中国科学院遥感与数字地球研究所 北京 100094

**摘要** 大数据是知识经济时代的战略高地，是国家和全球的新型战略资源。作为大数据重要组成部分的地球大数据，正成为地球科学的一个新的前沿领域，在推动地球科学的深度发展以及重大科学发现上意义重大。文章在分析地球大数据特点的基础上，介绍了正在开展的中国科学院战略性先导科技专项“地球大数据科学与工程”，剖析了专项目标、科学内涵，并对地球大数据促进联合国可持续发展目标实现进行了分析。地球大数据科学与工程集成地球科学、信息科学、空间科技等领域交叉融合前沿科学技术，为地球大数据服务地球科学发展带来新的前景。

**关键词** 地球大数据，对地观测，科学发现，决策支持，可持续发展

**DOI** 10.16418/j.issn.1000-3045.2018.08.008

大数据是知识经济时代的战略高地，是人类的新型战略资源。作为大数据重要组成部分的地球大数据，正成为地球科学的一个新的领域前沿，为推动地球科学的深度发展以及重大科学现象的认知发现作出贡献。

近年来，国际上推出了一系列与地球大数据相关的研究计划。美国的“地球立方体”项目，寻求以整体视角审视地球系统并创建管理地球科学知识的基础设施；欧盟启动的“活地球模拟器”研究项目，依赖于开放的标准，为所有地球科学领域提供多维数据和可扩展的服务；澳大利亚在启动的数据立方体项目基础上，最近又推出了澳大利亚“数字地球”项目；2017年，俄罗斯正式启动“数字地球”计划，包括发射一系列的地球观测卫星，为该计划提供源源不断的数据源。

我国在数字地球和全球地球观测系统方面，进行地球观测数据共享，整合卫星、地面监测和建模系统，评估环境条件和预测，以应对气候变化的挑战，推动地球系统科学发展。可见，地球大数据的发展在全球范围内方兴未艾，开展地球大数据的研究十分重要。

## 1 认识地球大数据

地球科学研究包括对大气、陆地和海洋等圈层的研究，通过空天对地观测、陆地传感器网络及其他方式的观测生成了海量的地球科学数据集，这些数据来源于但不限于空间对地观测数据，还包括陆地、海洋、大气及与人类活动相关的数据，这些与地球相关联的大数据统称为地球大数据。地球大数据是大数据的重要组成部分。

资助项目：中国科学院战略性先导科技专项（A类）（XDA19000000）

修改稿收到日期：2018年8月13日

分, 但有其独特的特征<sup>[1-3]</sup>。地球大数据具有规模大、来源广、多样化、多时相、多尺度、高维度、高复杂性和非结构化等特点, 是针对地球科学形成的新的数据密集型研究方向。

地球大数据通过多种对地观测方式、地球勘测方法及地面传感网产生。其数据集包含海量、多样、多来源、多时相、多标量、高空间、极复杂、非结构性的信息, 这极有利于促进地球科学研究的深层次发展<sup>[4]</sup>。

地球大数据具有大数据的特性, 不仅具有体量大、来源广、多时相、高价值等特点, 同时它也具有高瞬时性、任意空间性、物理相关性等特点。就其“海量规模”而言, 地球大数据具有高分辨率、高动态性以及多波段的特点, 数据的获取速度高且更新周期快。“来源广”是因为其数据来源和采集方法多种多样, 归功于成像原理和模型变化的多样。“多时相”指其在极短的采样间隔和高频信息中的采集能力。“高价值”因为其对生态环境、陆地资源、自然灾害及其他地球科学研究的重要意义。

与地球大数据相关的技术有对地观测技术、通信技术、计算技术、网络技术等。地球大数据可以加深人类了解地球的能力, 同时在数据传输、存储、处理、分析、管理和共享等方面, 地球大数据又面临全新的挑战<sup>[5-7]</sup>。例如, 海量对地观测数据的实时获取和多源数据融合, 从大量、长期、低成本的传感器网络中获得的数据增加了储存、处理和计算的复杂性, 集成的数据储存环境需要考虑物理存储设备和统一存储平台的交互技术和可视化方法等。

作为地球大数据的关键技术之一, 经过近半个世纪的发展, 对地观测技术为地球科学研究提供了新视角和新方法, 尤其在地球系统的宏观认知方面发挥了巨大的作用<sup>[8-12]</sup>。在对地观测技术帮助下, 人类可以更为方便、系统地观测地球, 对地观测技术已成为衡量一个国家科技成就、经济实力、国家安全的重要指标<sup>[13]</sup>。目前, 地球观测数据量正成倍增长, 据国际卫星对地观测委员会

所做的一项全球卫星研究数据显示, 过去的半个世纪, 全球已发射了超过 514 颗地球观测卫星对地球系统进行包括大气系统、海洋系统和陆地系统的综合观测。

地球大数据是具有空间属性的地球领域大数据, 它具有大数据的一般性质, 同时具有很强的时空关联和物理关联<sup>[14]</sup>。这些特征对地学学科的发展可以起到重大的推动作用, 在环境、资源、灾害等领域有重要作用和经济社会价值。地球大数据为地球科学的深入研究带来了重要的发展机遇, 可为创建空间地球信息科学奠定基础、并进一步推动地球科学的发展。

## 2 地球大数据目标

中国科学院充分认识到地球大数据的重要性, 2018 年初设立了战略性先导科技专项 (A 类) “地球大数据科学工程” (CASEarth), 系统开展地球大数据研究。CASEarth 的宗旨即利用地球大数据驱动跨学科、跨尺度宏观科学发现, 以系统性和整体性的理念去研究一系列重大科学问题, 在对地球系统科学认知上有重大突破, 同时在决策支持上实现新的跨越, 持续地产出在科学发现、宏观决策、技术创新和知识传播等方面的重大成果。

CASEarth 的总体目标是建成国际地球大数据科学中心, 其主要内容包括 3 部分: ① 建成地球大数据基础设施。突破数据开放共享的瓶颈问题, 形成多学科融合的地球大数据与云服务平台, 成为支撑国家宏观决策与重大科学发现的国家大数据重大科技基础设施。② 形成地球大数据学科驱动平台。探索大数据驱动、多学科融合、全球协作的科学发现新范式, 示范带动地球系统科学、生命科学及相关学科的重大突破。③ 构建服务政府高层的决策支持系统。具备多问题、多视角全景式可视化分析、模拟与推演能力, 展示和动态推演“一带一路”可持续发展过程与态势, 实现对全景美丽中国可持续发展和面向人类命运共同体的国家全球化战略的精准评价与决策支持。

CASEarth将在以下3个方面显具特色并取得重要产出：**① 实现系列科学发现。**形成大数据驱动科学发现的新方法、新范式，通过再现陆地、海洋、大气、人类社会要素各参数的空间分布和时间动态，揭示全球和区域尺度不同要素之间复杂耦合的相互作用，揭示不同分辨率下要素的细节和不同层次以及耦合关联。**② 产生系列技术创新。**构建高精度地球大数据云服务平台和新型数字地球系统，通过精确的地理关联和物理关联，将海量数据和信息产品集成展示。建成“全国布局，统分结合，透明服务”的地球大数据与云服务平台，形成多学科交叉、可演进、可服务的地球大数据重大基础设施。**③ 服务于政府决策支持。**建成大数据驱动、可视化、可交互、可动态演进的决策支持环境，实现多源空间信息的集成化数字再现与多要素交叉集成评估，提供宏观实时的地球大数据决策支持系统。

### 3 地球大数据专项内涵

CASEarth专项分解为以下8个部分开展研究，以实现理论和技术突破，取得创新成果。专项尤为重视数据的共享，并鼓励国内外学者依托本平台共同开展研究（图1）。

**(1) CASEarth小卫星研制。**研制服务于CASEarth的小卫星，建立CASEarth卫星运行管理与评估系统，完成CASEarth卫星数据接收与产品服务工作。研究面向观

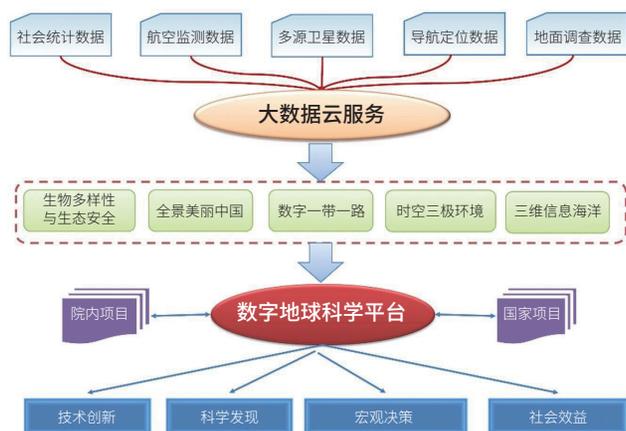


图1 “地球大数据科学工程” (CASEarth) 框架

测任务总体设计技术，高集成、小型化载荷总体技术，红外、多光谱载荷技术，以及海量数据压缩存储与传输技术。通过对地观测卫星总体设计及载荷研制、卫星工程研制、卫星运行管理与评估、卫星数据的接收与产品服务研究4个方面的研究，形成一套由卫星需求到数据产品的全系统流程。

**(2) 大数据与云服务平台。**建设具有一体化服务能力的地球大数据云服务平台，提供统一的计算与存储服务；研究多源异构海量数据接入、汇聚、存储管理与统一访问的标准规范、协议、工具和系统，集成整合海量多源的科学数据资源及空天地一体地球大数据，建成有特色的地球大数据资源库。突破分布式计算资源统一调度和聚合服务的技术、格网数据计算技术、大数据计算处理技术与分析挖掘的新方法，实现大数据驱动的科学发现与决策支持。

**(3) 数字“一带一路”。**构建“一带一路”地球大数据集成技术和评价体系，以及多要素科学数据库，实现“一带一路”地球大数据的综合集成，包括近50年49个大类要素的时空数据；开展“一带一路”地球大数据科学分析，认知“一带一路”全区域环境资源空间分布、可开发潜力、变化趋势科学规律；建立面向可持续发展目标的区域空间评估指标体系，实现“一带一路”可持续发展目标关键指标的科学监测；建立“一带一路”地球大数据分析决策支持系统。

**(4) 全景“美丽中国”。**在多角度、多维度、多环节、多因素、多层次的视角下，以地球系统科学和人-地关系理论为指导，开展基于大数据的资源环境本底分布与格局演变、清洁空气与环境健康、生态文明建设、区域发展与智慧城市以及“全景‘美丽中国’”评价与决策支持系统的研究和开发，全面展示“美丽中国”的本底特征、清洁空气与环境健康、生态文明建设和城市发展等专题情景，对“美丽中国”建设现状及未来情景进行评价与预测，为“美丽中国”建设提供政策建议。

**(5) 生物多样性与生态安全。**研究支持数据整合

和共享的标准以及数据集成应用方法；有机整合生物资源数据与生态、环境、气象、国民经济等数据，形成完整的数据图层；利用分析模型和可视化技术实现对生物多样性资源数据功能挖掘和利用，构建开放开源的生物多样性与生态安全大数据处理利用的通用接口，建立一个以生物多样性与生态安全信息为核心的综合大数据平台，在不同层次上实现不同形式的数据个性化服务与决策支撑。

(6) **三维信息海洋**。形成“两点一面”的海洋信息资源池，其中“一面”是指全球尺度的海洋信息资源池和数据产品，建立全球海洋基础数据服务系统；“两点”是指聚焦优势的方向和区域，在“中国近海”和“两洋一海”两个战略要点开展信息集成和科学研究。在“两洋一海”关键区域，研究南海岛礁多源数据及变化数据库与结构模型、西太平洋深海生物地理信息系统与深海极端生境的多维演示系统、印度洋及其重点港口区海洋灾害同化数据产品及海洋灾害预测预警系统，最终实现海洋信息的实时展示、动态模拟和情景分析。

(7) **时空三极环境**。通过三极专题大数据共享与集成、三极遥感对比研究、三极大数据分析方法和多圈层相互作用模型等基础研究，以及三极生态时空动态与预估、三极水环境与未来水安全、三极气候变化及对我国影响、北极航道监测与精细化预测、冰冻圈变化及其服务功能、极地重大工程冻土等专题研究，实现占领地球系统科学制高点，为极地治理和北极开发提供决策支持等目标。

(8) **数字地球科学平台**。该平台是“地球大数据科学工程”专项的综合展示平台，重点建设地球大数据综合展示与决策支持系统及网络化信息服务系统，主要为面向多学科融合与大数据驱动的科学发现和技术创新提供数据、服务、计算等多类资源及其关系的可视分析支持，兼顾资源、环境、生物、生态等领域的科学传播与公众服务。研制弹性可扩展、多模态数字地球科学平台，保障系统的安全可靠运行，为科学发现与技术创新

提供环境和工具支持。

CASEarth将突破数据开放共享的瓶颈问题，实现资源、环境、生物、生态等领域分散的数据、模型与服务等的全面集成，形成多学科融合、全球先进水平的地球大数据与云服务平台，构建大数据驱动的、具有全球影响力的数字地球科学平台，全景展示和动态推演“一带一路”“美丽中国”可持续发展目标实现进展，全面提升地球大数据在科学发现、宏观决策、技术创新和社会公众知识传播服务等方面的重大成果产出。

## 4 CASEarth促进联合国可持续发展目标实现

2015年9月，在联合国成立70周年之际，各国元首和代表相聚纽约联合国总部，通过了《变革我们的世界：2030年可持续发展议程》。该议程的核心是实现全球可持续发展目标，旨在所有国家和利益攸关方携手合作，阻止地球的退化，以可持续的方式进行消费和生产，管理地球环境和自然资源，使地球能够满足今世后代的需求，让全球走上可持续且具恢复力的道路，形成一个人与大自然和谐共处的世界<sup>[15]</sup>。

联合国《2030年可持续发展议程》是一项为人类、地球与繁荣制定的宏伟战略行动计划，包含经济、社会和环境3个方面，由17个可持续发展目标、169个具体目标及230个指标组成。目前，联合国、各国政府、国际组织等正在开展联合国可持续发展目标（SDGs）指标体系构建以及指标监测评估研究<sup>[16-18]</sup>。但由于SDGs在具体实施过程中面临许多挑战，其中数据缺失是监测SDGs各个目标最艰巨的挑战，数据统计体系不完善、不一致以及指标体系缺失是造成数据缺乏和质量不高的主要原因<sup>[19,20]</sup>；SDGs监测的评价指标模型化问题复杂，受限于数据的可获取性，在进行综合评价时，并不是所选指标均能模型化实现<sup>[21]</sup>。因此，如何科学建立综合、交叉、多要素相互作用评价模型库是一个难点问题。

为应对上述问题和挑战，CASEarth专项提出基于地球大数据的科技创新以促进可持续发展目标实现。以地

球大数据平台为基础, 综合集成资源、环境、生态和生物领域的数据库、模型库和决策方法库, 构建可持续发展评价指标体系和决策支持平台, 对资源、环境、生态等方面的可持续性进行有效的监测和评估, 将地球大数据纳入联合国和我国的支撑可持续发展评价体系中, 同时也服务于我国的生态安全和资源安全保护工作。

CASEarth 专项围绕 SDGs 的研究内容包括: ① 通过构建地球大数据共享服务平台, 努力成为联合国可持续发展目标实施数据提供者、生产者和指标的官方联络者; ② 围绕 SDG2、SDG6、SDG11、SDG13、SDG14 和 SDG15 等目标开展全球、区域、国家以及地区 4 个不同尺度的 SDGs 重点指标选择与空间评估指标体系构建以及指标监测评估; ③ 面向可持续发展目标监测的本土化问题, 尤其是发展中国家, 努力将我国的应用示范成果推动成为普适性的 SDGs 官方应用示范典型案例; ④ 在数据收集分析的基础上, 定期监测和评估可持续发展目标的进展, 形成“地球大数据支撑联合国可持续发展系列报告”; ⑤ 建设面向服务于 SDGs 目标评估和实现的开放型高端智库, 引领我国和国际 SDGs 目标的相关工作, 增强各层次上的科学决策能力。

## 5 结语

地球大数据为地球科学研究提供了全新的方法论, 正在成为认识地球的新钥匙和地球科学研究的新引擎, 有可能给地球科学研究带来重大变革。利用地球大数据, 结合地球系统科学模型, 发展地球大数据知识发现的理论与方法是地球科学中应解决的重大科学问题。地球大数据作为一个新的学科方向, 应开展持续研究, 注重地球科学、信息科学、空间科技等领域的交叉研究, 发展地球大数据研究方向, 促进地球系统科学研究迈上新的高度。作为代表国家水平和能力的国家创新平台, CASEarth 将通过地球大数据研究带来科学方法论和研究视角的创新, 带来可持续发展宏观决策支持的革命。形成地球系统科学发现的新引擎。

**致谢** 感谢“地球大数据科学工程”专项领导小组、协调小组、总体组、科研人员和专项办对本专项立项和研究的支持与贡献, 感谢陈方、梁栋对本文的付出。

## 参考文献

- 1 郭华东, 陈润生, 徐志伟, 等. 自然科学与人文科学大数据——第六届中德前沿探索圆桌会议综述. 中国科学院院刊, 2016, 31(6): 707-716.
- 2 郭华东. 大数据 大科学 大发现——大数据与科学发现国际研讨会综述. 中国科学院院刊, 2014, 29(4): 500-506.
- 3 Guo H D, Wang L, Liang D. Big earth data from space: A new engine for earth science. Chinese Science Bulletin, 2016, 61(7): 505-513.
- 4 Guo H D, Liu Z, Jiang H, et al. Big earth data: A new challenge and opportunity for digital earth's development. International Journal of Digital Earth, 2017, 10(1): 1-12.
- 5 Vatsavai R R, Ganguly A, Chandola V, et al. Spatio temporal data mining in the era of big spatial data: algorithms and applications. ACM Sig spatial International Workshop on Analytics for Big Geospatial Data, 2012.
- 6 Reinsel D, Gantz J, Rydning J. Data age 2025: The evolution of data to life-critical don't focus on big data. Framingham: IDC Analyze the Future, 2017.
- 7 Guo H D. Steps to the digital Silk Road. Nature, 2018, 554: 25-27.
- 8 He G, Wang L, Ma Y, et al. Processing of earth observation big data: Challenges and countermeasures. Chinese Science Bulletin, 2015, 60(5-6): 470-478.
- 9 Ramapriyan H, Brennan J, Waler J, et al. Managing Big Data: NASA Tackles Complex Data Challenges. Earth Imaging Journal. [2013-10-18]. <http://eijournal.com/print/articles/managing-bigdata>.
- 10 Pekel J, Cottam A, Gorelick N, et al. High-resolution mapping of global surface water and its long-term changes. Nature, 2016, 540: 418-422.

- 11 Hansen M C, Potapov P V, Moore R, et al. High-resolution global maps of 21st-century forest cover change. *Science*, 2013, 342: 850-853.
- 12 Guo H D, Fu W X, Li X W, et al. Research on global change scientific satellites. *Science China Earth Sciences*, 2014, 57(2): 204-215.
- 13 郭华东, 陈方, 邱玉宝. 全球空间对地观测五十年及中国的发展. *中国科学院院刊*, 2013, 28(Z1): 7-16.
- 14 Guo H D. Big Earth data: A new frontier in Earth and information sciences. *Big Earth Data*, 2017, 1: 4-20.
- 15 United Nations. Transforming our World: The 2030 Agenda for Sustainable Development. [2018-06-30]. <https://sustainabledevelopment.un.org/post2015/transformingourworld/publication>.
- 16 Anderson K, Ryan B, Sonntag W, et al. Earth observation in service of the 2030 Agenda for Sustainable Development. *Geo-spatial Information Science*, 2017, 20: 77-96.
- 17 World Health Organization. Global Burden of Disease. [2018-06-30]. [http://www.who.int/topics/global\\_burden\\_of\\_disease/en/](http://www.who.int/topics/global_burden_of_disease/en/).
- 18 Getting started with the Sustainable Development Goals: A guide for stakeholders. [2018-06-30]. <http://unsdsn.org/wpcontent/uploads/2015/12/151211-getting-started-guide-FINALPDF-.pdf>.
- 19 United Nations Statistics Division. The Sustainable Development Goals Report 2018. [2018-06-30]. <https://unstats.un.org/sdgs/report/2018/>.
- 20 United Nations. Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators. New York: United Nations, 2016.
- 21 Mark E, Bengtsson M, Akenji L. An optimistic analysis of the means of implementation for Sustainable Development Goals: Thinking about goals as means. *Sustainability*, 2016, 8(9): 962-985.

## A Project on Big Earth Data Science Engineering

GUO Huadong

( Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China )

**Abstract** Recent improvements in capability and ability of storing and utilizing vast quantity of data has enabled revolutionary innovation of big data analytics. Big data in just a span of few years have occupied strategic importance in many aspects of human society, making data an important commodity and a valuable resource in the era of knowledge based economies. Scientific research into big data collection, storage, analysis, and exploitation have developed rapidly and continues to progress at a rapid pace. At the same time, the amount of historical Earth observation data generated over the past five decades and the continued human and capital resource investments of many countries, public and private corporations ensures improved generation of Earth observation well into the future with exponentially increasing volumes of information on Earth systems and science. This has given rise to a new class of big data termed as “Big Earth Data”. Big Earth Data has macro-level capabilities that enable rapid and accurate monitoring of Earth, and is increasingly gaining importance in Earth sciences, adding value to its utilization in problem driven science and innovation. This paper introduces the characteristics of Big Earth Data and analyzes its great potential for development, particularly in regards to the role that Big Earth Data can play in transforming Earth science. With this context the paper outlines the Project on Big Earth Data Science Engineering (CASEarth) of the Chinese Academy of Sciences Strategic Priority Research Program and highlights how the CASEarth would contribute to ensure the actual use of Big Earth Data in support of the achievement of Sustainable Development Goals (SDGs) as articulated in the 2030 Agenda document. The potential prospects of developing Big Earth

Data and its ability to integrate geosciences, information sciences, and space science and technology makes it a strategic future endeavor for revolutionizing Earth Science as a whole.

**Keywords** Big Earth Data, Earth observation, scientific discoveries, decision support, sustainable development



**郭华东** 中国科学院遥感与数字地球研究所研究员，博士生导师。中国科学院院士、俄罗斯科学院外籍院士、芬兰科学与人文院外籍院士、发展中国家科学院院士。主要从事遥感信息科学、雷达对地观测、数字地球等领域研究。现任国际数字地球学会主席、“联合国可持续发展目标技术促进机制10人组”成员、国际环境遥感委员会主席、联合国教科文组织国际自然与文化遗产空间技术中心主任、“数字丝路”国际科学计划主席、国家大数据专家委员会顾问、《国际数字地球学报》和《地球大数据》主编等职。现为中国科学院战略性先导科技专项（A类）“地球大数据科学与工程”负责人。E-mail: hdguo@radi.ac.cn

**Guo Huadong** Professor of Institute of Remote Sensing and Digital Earth (RADI), Chinese Academy of Sciences (CAS), Academician of CAS, Foreign Member of the Russian Academy of Sciences, Foreign Member of the Finnish Society of Sciences and Letters, and Fellow of the World Academy of Sciences for the advancement of science in developing countries (TWAS). He presently serves as President of the International Society for Digital Earth (ISDE), Member of UN 10-Member Group to Support the Technology Facilitation Mechanism, Chairman of the International Committee on Remote Sensing of Environment (ICORSE), Director of the International Centre on Space Technologies for Natural and Cultural Heritage (HIST) under the Auspices of UNESCO, Chair of Science Committee of Digital Belt and Road Program (DBAR), Editor-in-Chief of the *International Journal of Digital Earth* and *Big Earth Data*. He served as President of ICSU Committee on Data for Science and Technology (CODATA). He specializes in remote sensing science and its applications, and has a series of achievements in remote sensing information mechanisms, radar for Earth observation, and Digital Earth science. He has published more than 600 papers and sixteen books, and is the principal awardee of sixteen domestic and international prizes. E-mail: hdguo@radi.ac.cn

■ 责任编辑：张帆