

21世纪逐日无缝数据立方体构建方法及逐年逐季节土地覆盖和土地利用动态制图—中国智慧遥感制图 iMap (China) 1.0

刘涵¹, 官鹏^{1,2,3}

1. 清华大学 地球系统科学系 教育部地球系统模拟重点实验室, 北京 100084;
2. 清华大学 中国城市研究院, 北京 100084;
3. 东亚迁徙鸟类与栖息地生态学野外观测研究站, 北京 100084

摘要: 粮食安全、高质量人居环境建设、生物多样性保护、星球健康等社会可持续发展目标和对地球系统的理解、模拟与管理都迫切需要多尺度、长时序、辐射和几何精度高且一致性强的遥感观测数据集和针对用户需求的、信息主题灵活的制图产品。但是由于技术和成本限制, 传统的遥感卫星难以提供同时具有高空间分辨率、高时间频率和高质量的观测数据。现有的制图和反演方案多是针对单一传感器系列, 难以充分挖掘和联合利用多源异构遥感大数据的信息潜力, 造成观测时段和分辨率有限、时空一致性和可比性较差。因此, 遥感领域迫切需要新的技术范式。本文基于前沿的云计算、人工智能、虚拟星座、时空融合重建等技术, 针对现有遥感大数据特别是国产卫星数据, 提出一套智慧遥感制图 (iMap) 框架。该框架从用户需求出发、问题驱动, 能够大大改善当前遥感数据产品难以满足农林管理、国情监测、生态环境保护、防灾减灾、城市建设等用户的多样化、高精度地表监测需求的现状。在该框架的指导下, 基于亚马逊云计算 (AWS) 高性能、高弹性、可扩展的分布式计算资源, 搭建了在线实时、自动化、无服务器、端到端的遥感大数据生产链和并行制图系统, 并生产了首套21世纪中国全境逐日无缝数据立方体 (SDC) 及逐年逐季节土地覆盖和土地利用制图产品。逐日 SDC 综合利用 Landsat 和 MODIS 卫星数据构建虚拟星座, 并通过多源时空数据融合重建技术研制得到无云无缝、高精度的反射率产品, 作为分析就绪数据 (ARD), 为高精度定量遥感反演和制图打下根基。基于这一 SDC, 完成了逐年逐季节地表制图。逐年平均精度超过 80%。在制图过程中, 基于多套多层土地覆盖和土地利用分类体系, 运用有限样本稳定分类理论, 迁移使用全季节普适样本库, 采用自动机器学习 (AutoML) 策略集成优化多种分类器, 并结合时空一致性变化检测和后处理技术。这两套制图产品证明了本文提出的智慧遥感制图框架的可行性和有效性。未来将进一步完善和发展该框架, 以开放和灵活的理念, 为促进中国遥感进一步发展提供新的思路。**关键词:** 无缝数据立方体, 逐日数据, 逐季节制图, 云计算, 智慧遥感制图, 长时序, 变化监测, 人工智能, 时空大数据

引用格式: 刘涵, 官鹏. 2021. 21世纪逐日无缝数据立方体构建方法及逐年逐季节土地覆盖和土地利用动态制图—中国智慧遥感制图 iMap (China) 1.0. 遥感学报, 25(1): 126–147

Liu H and Gong P. 2021. 21st century daily seamless data cube reconstruction and seasonal to annual land cover and land use dynamics mapping—iMap (China) 1.0. National Remote Sensing Bulletin, 25(1): 126–147 [DOI: 10.11834/jrs.20210580]

1 引言

土地覆盖、土地利用是自然和人类力量共同塑造的结果, 也是人类活动对地表自然生态系统

影响的直接表现形式 (Foley 等, 2005; Turner 等, 2007; Song 等, 2018)。土地覆盖和土地利用变化, 与陆地表层物质循环和生命过程密切相关, 影响着生物圈与大气交互作用、生物多样性和生

收稿日期: 2020-12-12; 预印本: 2020-12-31

基金项目: 国家高技术研究发展计划 (编号: 2016YFA0600104)

第一作者简介: 刘涵, 1995年生, 男, 博士研究生, 研究方向为多尺度长时序全球地表覆盖智慧制图、虚拟星座无缝数据立方体重建。E-mail: liuhan18@mails.tsinghua.edu.cn

通信作者简介: 官鹏, 1965年生, 男, 教授, 研究方向为地表覆盖制图、生态遥感。E-mail: penggong@tsinghua.edu.cn

态系统服务功能、地表辐射能量平衡、生物地球化学循环以及资源环境的可持续利用 (Yang 等, 2013; Qiu 等, 2020; Sy 和 Quesada, 2020)。作为地球系统模式的关键输入参数之一, 土地覆盖、土地利用数据对于全球环境变化及其模拟研究具有重要价值 (Feddemma 等, 2005; Beusch 等, 2020)。土地覆盖、土地利用数据也是农林生产、土地政策、城市建设、抵御洪涝和火灾防范的基础数据, 可服务于生态资源环境和城市发展的评估、管理和决策, 促进可持续发展 (刘纪远 等, 2014; 刘芳 等, 2016)。

21世纪以来, 随着社会经济持续快速发展以及工业化、城市化进程的加快, 中国耕地经受巨大压力, 影响到中国粮食安全; 同时一系列环境破坏问题, 也对中国自然生态系统的可持续性构成威胁 (Bai 等, 2018)。为了紧守耕地红线和生态红线, 需要高精度的土地覆盖、土地利用数据来监控土地变化。同时, 解决城市快速扩张引起的高质量农地和其他生态系统损失和环境污染问题, 以及制定城市规划、评估建设效果等都需要在较高的分辨率水平上对土地覆盖、土地利用进行识别 (Gong 等, 2019; Gong 等, 2020)。此外, 退耕还林、保护耕地等土地政策和生态保护工程的目标制定与实施效果评估, 均需要多年份高精度土地覆盖、土地利用数据作为量化的依据 (Mao 等, 2019; Peng 等, 2019)。为了满足上述需求, 至少需要逐年的空间分辨率在 30 m 甚至米级的土地制图更新数据; 除了对空间分辨率有较高要求外, 农业、生物多样性保护和能量平衡计算对土地覆盖、土地利用信息的时间频率要求要高于月甚至需要逐日水平, 对地表覆盖类型的要求到作物或其他植被的种类水平。

国家统计局的统计数据, 通常来自地方的调查数据, 无法体现细致的空间差异性, 且来自不同区域的数据受主观性影响, 准确性难以评估。遥感卫星影像由于具备长时间监测能力、较高空间分辨率等特点, 已成为土地覆盖、土地利用变化监测的重要数据源之一。为了满足各种应用需求, 人们基于遥感数据开展了土地覆盖、土地利用制图的大量努力。然而任何一种产品都无法满足上述应用对空间分辨率、时间频率、土地覆盖/土地利用类型和分类精度的要求。现有粗分辨率制图产品包括全球 5 km 地表覆盖制图产品 GLASS-GLC (Liu 等, 2020)、欧洲航天局气候变化倡议 (ESA-CCI) 300 m 全球土地覆盖数据 ([\[maps.elie.ucl.ac.be/CCI/viewer/index.php\]\(http://maps.elie.ucl.ac.be/CCI/viewer/index.php\) \[2020-12-12\]\)、MODIS 500 m 全球土地覆盖产品 \(Friedl 等, 2010; Sulla-Menashe 等, 2019\)、1 km 国际地圈生物圈计划和信息系统覆盖图 \(IGBP-DISCover\) \(Loveland 等, 2000\)、1 km 马里兰大学土地覆盖图 \(Hansen 等, 2000\)、1 km 全球土地覆盖 2000 \(GLC2000\) 图 \(Bartholome 和 Belward, 2005\) 等。高分辨率产品, 如 30 m 和 10 m 精细分辨率的全球土地覆盖图 \(FROM-GLC, FROM-GLC10\) \(Gong 等, 2013, Gong 等, 2019\)、30 m 全球土地覆盖数据产品 \(GlobeLand30\) \(Chen 等, 2015\)、中国多期土地利用遥感监测数据集 \(CNLUCC\) \(Liu 等, 2005; Zhang 等, 2014; Xu 等, 2018\)、美国国家土地覆被数据库 \(NLCD\) \(Vogelmann 等, 2001; Homer 等, 2004, 2007, 2015, 2020; Xian 等, 2009; Yang 等, 2018\) 和美国土地变化监测、评估和预测产品 \(LCMAP\) \(Brown 等, 2020; Pengra 等, 2020\)。在全类别土地制图产品中, 部分类别动态变化显著、制图精度低, 出现了一些针对单类的土地制图产品, 如水体 \(Wood 等, 2011; Pekel 等, 2016; Ji 等, 2018; Pickens 等, 2020\)、不透水面 \(Pesaresi 等, 2013; He 等, 2018; Liu, Hu 等, 2018; Xu 等, 2018; Gong 等, 2020; Liu 等, 2020\), 森林 \(Hansen 等, 2013; Shimada 等, 2014; Crowther 等, 2015; Hamilton 和 Casey, 2016\), 湿地 \(Hu 等, 2017; Murray 等, 2019\) 等, 但不能满足反映多种土地类型变化的需求。](http://</p></div><div data-bbox=)

总体来看, 已有的土地覆盖产品最高的时间频率通常到年, 更高的时间频率, 如季节尺度土地覆盖数据, 很少 (付薇 等, 2014)。时序制图一致性、可比性差是一大难点 (Sulla-Menashe 等, 2019; Zhu 等, 2020), 降低了时序制图产品的可信度, 阻碍了时序制图产品的应用 (Gómez 等, 2016)。已有大量研究开发了变化检测方法应对一致性问题, 如 (Kennedy 等, 2010; Verbesselt 等, 2012; Brooks 等, 2014; Zhu 和 Woodcock, 2014; Bullock 等, 2020; Zhu 等, 2020), 然而缺少全类别的一致性变化探测技术。另一方面, 常用的传统遥感数据源时间频率不够密集, 因而难以有效反映时间频率更高的地表覆盖动态 (Woodcock 等, 2020)。数据融合是一种有效应对方式 (Zhang 2010; Wulder 等, 2018), 通过融合多种数据, 提高数据的时空分辨率, 然而当前缺少将数据融合实际应用于大范围土地制图的应用。

样本问题是当前土地覆盖制图, 尤其是大范围、高精度、高频率时序制图的阻碍之一 (刘涵等, 2017; Liu 等, 2019; Woodcock 等, 2020)。土地覆盖、土地利用复杂多样, 样本, 尤其是时序样本, 采集耗费大量人力、物力, 成本高、难度大 (Foody, 2010; Zhao 等, 2014; Liu 等, 2017; Pengra 等, 2020)。然而通常建立的样本集往往只服务于当前制图任务和需要, 样本的积累和共享不足 (Li 等, 2017)。利用当前样本库进行迁移和复用方面的尝试偏少。目前, 仅有 FROM-GLC10 一套全球土地覆盖产品对样本进行了跨传感器和时空尺度的迁移复用 (Gong 等, 2019)。有限样本稳定分类等样本迁移理论的提出, 为样本迁移复用、跨传感器、跨尺度快速制图提供了一种理论依据 (Gong 等, 2019; Huang 等, 2020)。

遥感大数据是土地时序制图面临的另一大瓶颈 (Baumann 等, 2016)。遥感卫星数据量大、异构而复杂, 存储、处理、分析需要大量工作和成本 (Le 等, 2016; Baumann 等, 2018)。传统的本地处理 (例如台式计算机) 和数据分发方法 (例如基于场景的文件下载) 无法适应遥感大数据 (Baumann 等, 2016)。为应对这些挑战, 降低数据准备和处理的复杂性, 出现了以数据立方体 DC (Data Cube) 为代表的新的遥感数据处理范式 (Giuliani 等, 2017)。数据立方体旨在通过解决体积、速度和多样性挑战, 以分析就绪的形式提供对大型时空数据的访问 (Killough 2018)。当前, 基于开放数据立方体等标准 (Open Data Cube - <https://www.opendatacube.org/> [2020-12-12], Earth Observation Data Cube - <http://eodatacube.eu> [2020-12-12], Earth System Data Cube - <http://earthsystemdatacube.net> [2020-12-12], Earth on Amazon Web Services-EAWS - <https://aws.amazon.com/earth/> [2020-12-12] 和 Google Earth Engine (Gorelick, Hancher et al. 2017) -GEE, <https://earthengine.google.com> [2020-12-12]), 已开发了一些 DC。例如 Digital Earth Australia (Lewis 等, 2017), Swiss Data Cube (Giuliani 等, 2017) 和 Africa Regional Data Cube (Killough 2019)。这些应用往往依靠高性能计算机和云计算环境为数据基础架构提供支持 (Lewis 等, 2016; Rizvi 等, 2018)。然而这些数据立方体往往只是将已有遥感卫星数据进行基本预处理, 切片形式存储组织, 并未对数据的缺失观测和异常值进行进一步的时

空补齐, 以便于时间序列分析。

Landsat 是当前时间序列分析和制图中使用最广泛的卫星数据源 (Woodcoc 等, 2008; Wulder 等, 2019)。这与 Landsat 数据开放获取、易于访问, 校准和预处理相对成熟等特点有很大关系 (Roy 等, 2010; Loveland 和 Dwyer, 2012; Wulder 等, 2012; Dwyer 等, 2018)。Landsat 卫星数据 30 m 的空间分辨率也相对较高, 然而与高空间分辨率相伴的是低时间频率。此外, Landsat 卫星经历多次更新换代, 传感器之间差异大。同时 Landsat 7 具有明显的条带现象, 时序 Landsat 数据的时间一致性和数据质量难以保证。恶劣的大气条件 (如厚云层), 也常常使数据不完整, 降低了数据的可用性。为应对这些缺陷和问题, 减缓异常值和缺失值的影响, 重建数据是一种有效的应对策略。据重建缺失信息时补充信息的不同来源, 重建算法通常可以分为基于空间的方法、基于光谱的方法、基于时间的方法以及混合方法 (Shen 等, 2015)。基于空间的方法使用数据的其余部分填充缺少的数据区域, 代表性方法如内插方法 (Zhang 等, 2007)。由于缺少足够的信息, 基于空间的方法通常无法重建大面积丢失值。这一缺陷可以通过从光谱影像中的大量冗余光谱信息中重建特定频段中的丢失数据来克服 (Wang 等, 2006; Rakwatin 等, 2008; Shen 等, 2010)。但是, 基于光谱的方法前提是损坏的光谱数据既具有不完整的又具有完整的光谱带。由于云的移动性和传感器的扫描偏差, 在同一地理区域且在不同时期获取的数据可以提供补充信息。对于基于时间的方法, 如果时间间隔太短, 则两个连续数据集集中的云将大部分重叠, 并且时间相关性可能无效。但是, 如果时间间隔太长, 则土地覆盖范围可能会变化很大, 从而破坏了相关性。常见有时间替代法 (Melgani 2006; Chen 等, 2011; Zeng 等, 2013; Cheng 等, 2014), 时间滤波法 (Roerink 等, 2000; Chen 等, 2004; Julien 和 Sobrino, 2010)。将这些方法相结合, 利用在空间, 光谱和时间域中的相关性可以取得更好效果, 如 (Benabdelkader 和 Melgani, 2008; Zeng 等, 2013; Cheng 等, 2014)。除重建方法外, 已提出将 MODIS 与 Landsat 数据融合的多种算法 (Gao 等, 2006; Zhu 等, 2010), 在填补缺失值的同时, 还可以提高数据的时间频率。然而已有研究往往停留在算法研究领域, 缺少大规模制图生产和使用

的实验研究。

此外,已有的土地制图数据,常常针对土地覆盖或土地利用,无法满足其他应用需求。分类系统简单且灵活性差,无法根据不同应用需求,提供不同类别粒度的层次化产品 (Yu等, 2014; 宫鹏等, 2016)。

考虑到已有数据和产品存在的缺陷和中国土地评估的重要性,为了把握和揭示中国土地覆盖、土地利用变化的时空特征,我们基于AWS云计算环境,生产了中国21世纪以来逐年多季节、多分类系统的30 m空间分辨率的土地覆盖、土地利用数据产品。为了提高Landsat数据一致性、时间频率和数据质量,提出基于虚拟星座的时空融合和重建技术,首次建成了21世纪中国逐日30 m的无缝数据立方体SDC (Seamless Data Cube)。我们开发了基于变化检测和时空一致性的长时序高分辨率动态制图框架。为了满足不同的应用需求,该产品中逐年数据的分类系统有3种类型、两级结构,土地覆盖分类系统有FROM-GLC分类系统、FAO Land Cover Classification System (LCCS),土地利用分类系统来自CNLUCC分类系统。针对FROM-GLC土地覆盖分类系统,生产了逐季节产品。

2 数据与方法

如前所述,气候变化研究、地球系统模式发展、及农林管理、碳氮循环和水循环模拟和生物多样性保护及防灾减灾等迫切需要遥感制图创新。这既是世界性的需求也是科学发展的必然趋势。信息技术发展到今天,使人类知识数字化,使人类智慧逐步在虚拟空间得以积累和应用,给既模仿学习人工遥感制图的灵活性、又利用计算机处理超高维多样化的海量数据的准确性和高效性创造了条件。人工智能技术中计算机的人脸识别精度超过人类代表着机器智慧达到了新高度。在遥感制图领域,我们把既能像专业制图人员一样综合利用色彩、纹理、相对位置、图像要素上下文、和其他辅助地理知识,又能充分发挥计算机在制图人员难以做到的快速、大规模处理超过三维空间的海量数据方面的优势的计算机技术称为智慧遥感制图 iMap (intelligent mapping of land cover and land use)。作为实现智慧遥感制图的初步尝试,本文提出一套新的智慧遥感制图框架 (图1),将在这一框架下研制的第一版中国土地覆盖和土地利用逐年逐季节数据集作为智慧遥感制图1.0版本。

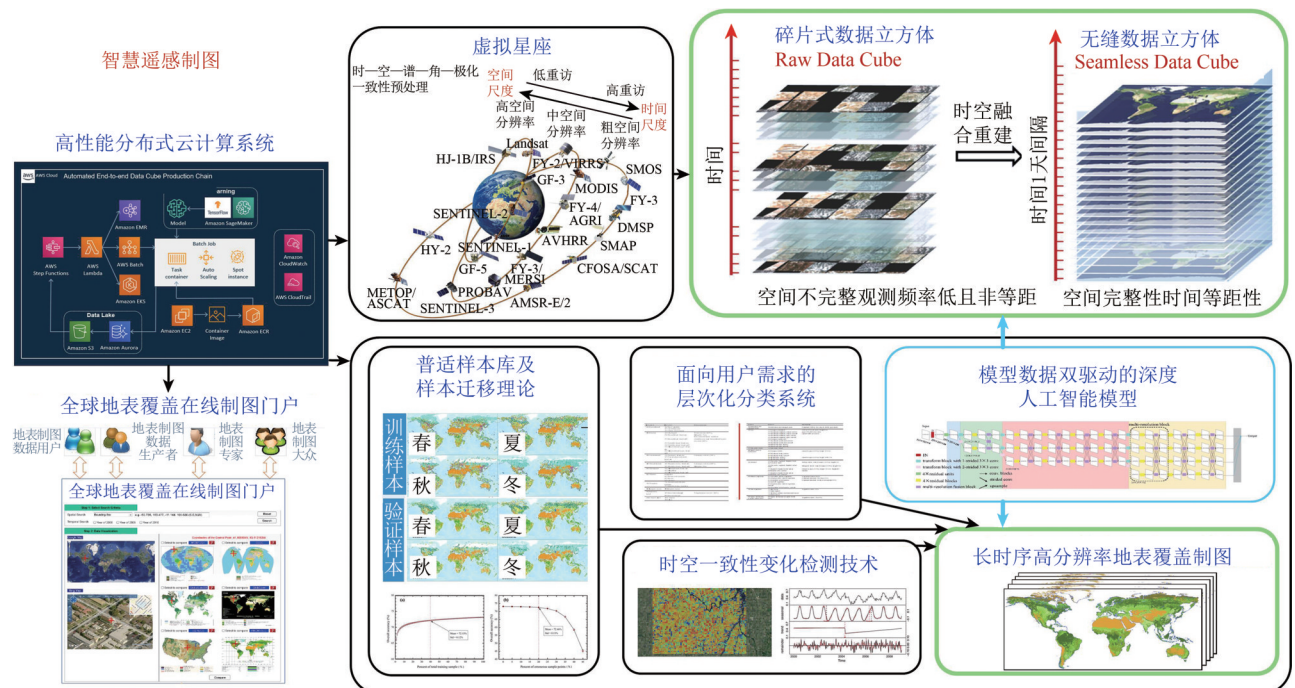


图1 智慧遥感制图概念示意图

Fig. 1 Intelligent remote sensing mapping

新智慧遥感制图框架旨在以高性能云计算平台为支撑,在综合利用国内外多源遥感数据、定

量遥感产品和社交媒体等非传统遥感数据的框架下,构建虚拟星座,结合大数据驱动的机器学习

等人工智能方法, 探索 and 解决无缝智能数据重建、面向用户需求的层次化分类系统设计、普适样本库构建及迁移、时空一致性变化检测等基础性遥感难题, 生产一系列高精度、具有国际先进水平的多尺度长时序全球地表覆盖制图产品, 并提供基于云计算的在线制图平台, 为遥感、地球系统科学以及社会经济发展提供关键性基础支撑, 实现多尺度、多要素、精准化的智慧遥感制图服务。

在智慧遥感制图理念指导下, 本文以 21 世纪中国为研究对象, 提出了基于虚拟星座的无缝数据立方体融合重建的技术方案, 和长时序高分辨率一致性动态制图框架 (图 2)。依据 ARD 数据标准对 Landsat 和 MODIS 影像进行一致性预处理, 组建虚拟星座, 构成原始的数据立方体 RDC (Raw

Data Cube)。提出时空多源引导对抗修复网络等模型, 对 RDC 进行影像质量修复与补充、时空融合和数据平滑, 构建逐日 30 m 无缝数据立方体 SDC (Seamless Data Cube)。基于 SDC, 进一步计算特征立方体 (feature cube)。RDC、SDC 和 feature cube, 共同构成多种层次的 (multi-level) ARD 产品, 用户可根据需求直接应用。在制图方面, 设计自适应用户需求的多种、多层次的土地覆盖、土地利用分类体系, 以有限样本稳定分类理论为指导跨时间迁移使用普适样本库, 基于自动机器学习进行模型调优和集成, 然后以变化检测保证制图的时空一致性, 最终构成长时序高分辨率动态制图框架。

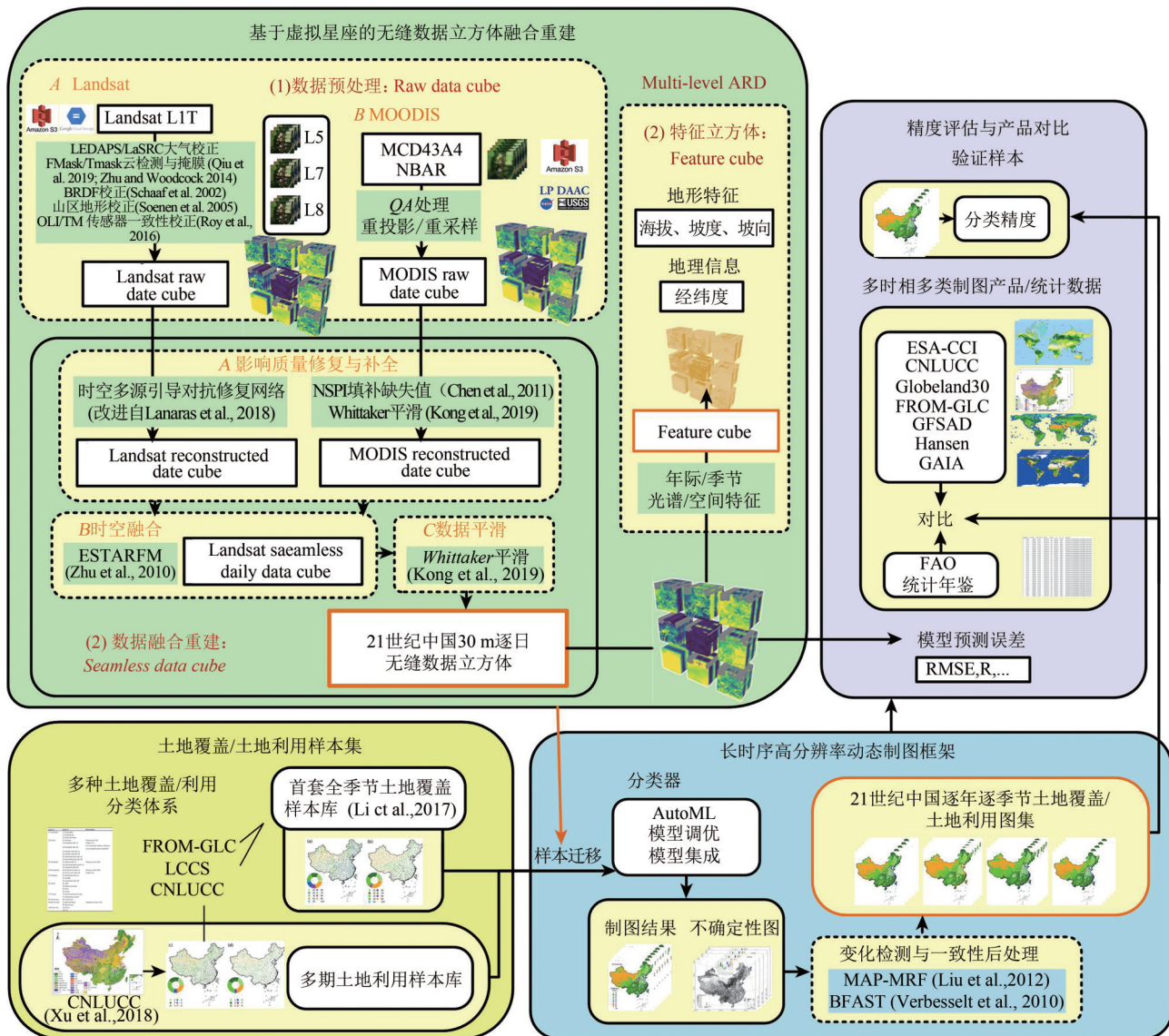


图 2 中国无缝数据立方体和地表覆盖制图流程

Fig. 2 The flowchart of seamless data cube reconstruction and land cover mapping for China

本文基于亚马逊云计算 AWS (Amazon web Services) 平台, 搭建了在线实时、自动化、无服务器、端到端的遥感大数据生产链和并行制图系统, 以完成全部生产处理流程。AWS 公开数据集 (AWS Open Data) 提供了所需的 Landsat、MODIS 等遥感数据源。AWS 高性能、高弹性、可扩展的分布式算力, 满足了快速处理 PB 级数据的计算资源需求。借助 AWS Step Functions, 将 AWS Lambda、Amazon ECS (Amazon Elastic Container Service) 等云服务组合, 搭建了自动化、无服务器、端到端的流式生产链。AWS 提供了完善的人工智能和机器学习服务, 采用 Amazon SageMaker 工具, 高效设计和训练了影像修复模型和制图分类模型, 并借助 AutoML 对模型结构和参数进行深度调优, 最终完成分布式高性能推理。

2.1 基于虚拟星座的无缝数据立方体融合重建

2.1.1 数据预处理

本文使用了来自亚马逊公开数据集 (AWS open data) 和谷歌云存储 (Google cloud storage) 的 2000 年—2019 年所有可用的 Landsat L1T (Level 1 terrain-corrected) 影像。中国范围内一共包括 579 个图幅 (Path/Row)。采用与 Landsat 分析就绪数据 ARD (Analysis Ready Data) 相同的范式

对所有 Landsat 数据进行大气校正、生成表面反射率 (Masek 等, 2006; Vermote 等, 2016)、云及云阴影检测与掩膜 (Zhu 和 Woodcock, 2014; Qiu 等, 2019) 等处理流程。在 ARD 处理范式外, 我们还对影像进行了双向反射分布函数 BRDF (Bidirectional Reflectance Distribution Function) 校正 (Schaaf 等, 2002), 生成 BRDF 校正反射率 NBAR (Nadir BRDF-Adjusted Reflectance), 以及对山区影像进行了地形校正 (Soenenaf 等, 2005)。考虑到传感器之间的差异, 我们对 OLI 和 TM 传感器进行一致性校正 (Royaf 等, 2016)。所有数据组织和存储成空间规则、有时间标记、波段集中的网格 (tile), 得到原始的数据立方体 RDC (Raw Data Cube)。Data cube 组织范式优化了遥感数据的访问性能, 具有时空高互操作性; 时空可变网格, 能以多种网格尺度分发; ARD 处理标准使所有数据可直接用于分析。

为提高数据的时间频率和空间完整性, 本文采用了来自 AWS open data 和美国地质调查局 (USGS) 同期的 500 m 逐日 MODIS MCD43A4 第六代 NBAR 数据产品与 Landsat 组成虚拟星座。用质量信息波段对 MODIS 进行预处理。所用波段如表 1 所示。按照同样范式, 生成了 MODIS raw data cube。

表 1 Landsat 与 MODIS 对应波段
Table 1 The corresponding bands of Landsat and MODIS

波段	Landsat NBAR			MODIS NBAR
	Landsat 5 TM	Landsat 7 ETM+	Landsat 8 OLI	MCD43A4
Blue	0.450—0.520	0.450—0.520	0.450—0.510	0.459—0.479
Green	0.520—0.600	0.520—0.600	0.530—0.590	0.545—0.565
Red	0.630—0.690	0.630—0.690	0.640—0.670	0.620—0.670
NIR	0.760—0.900	0.770—0.900	0.850—0.880	0.841—0.876
SWIR 1	1.550—1.750	1.550—1.750	1.570—1.650	1.628—1.652
SWIR 2	2.080—2.350	2.090—2.350	2.110—2.290	2.105—2.155

2.1.2 影像质量修复与补全

MODIS 数据观测频率高, 互补信息相对多, 采用经典的算法流程对 MODIS 影像进行修复。首先, 本文利用邻近相似像元插补算法 NSPI (Neighborhood Similar Pixel Interpolator) 方法 (Chenaf 等, 2011) 对 MODIS 缺失进行填补。

NSPI 方法原理简单, 同时又较高的计算效率 (Zhuaf 等, 2012)。NSPI 方法基于同类地物光谱信息的空间连续性特征, 以及这种特征在时间尺度上的不变性的合理假设, 以邻近时间影像作为参考, 使用加权线性模型预测目标像素 $(x_{w/2}, y_{w/2}, t_p)$

$$F(x_{w/2}, y_{w/2}, t_p, B) = T_1 \times \sum_{i=1}^N W_i^1 \times F(x_i, y_i, t_p, B) + T_2 \times \left(F(x_{w/2}, y_{w/2}, t_r, B) + \sum_{i=1}^N W_i^2 \times (F(x_i, y_i, t_p, B) - F(x_i, y_i, t_r, B)) \right) \quad (1)$$

式中, F 是反射率, B 是波段, t_p 、 t_r 分别是预测和参考时间, N 是相似像素数量, w 是窗口大小。 $(x_{w/2}, y_{w/2})$ 为预测中心像素, (x_i, y_i) 是第 i 个相似像素位置。 W_i^1 、 W_i^2 是分别基于空间、光谱距离计算的相似像素的权重, T_1 、 T_2 是将基于光谱—空间、光谱—时间的两种预测结合起来的权重。

进一步采用加权重特克平滑算法 (weighted Whittaker smoothing) (Eilers, 2003; Kong 等, 2019) 平滑 MODIS 时间序列获得 MODIS 重建的数据立方体 (reconstructed data cube)。相比其他降噪和平滑算法, 这种算法能够很好地平衡保真度和粗糙度。其本质是一种惩罚最小二乘算法, 可最小化拟合误差并惩罚平滑曲线的粗糙度, 目标是找到以下 Q 最小的最优平滑时间序列 z :

$$Q = S_{wt} + \lambda R_{wt} = (FC - z)^T W_{wt} (FC - z) + \lambda \|D_{wt} z\|^2 \quad (2)$$

$$z = (W_{wt} + \lambda D^T D) / W_{wt} FC$$

式中, S_{wt} 是保真度, R_{wt} 是粗糙度, FC 是原始时间序列, λ 是粗糙度参数, D_{wt} 是差分矩阵, W_{wt} 是权重的对角矩阵, 由对应点的质量决定。

Landsat 数据纹理信息相对丰富, 但观测频率低, 影像修复相对 MODIS 要求高。为此, 提出时空多源引导对抗修复网络 (图 3), 修复 Landsat 观测的条带、云和阴影等缺失和异常值。生成网络结构改进自 (Lanarasaf 等, 2018) 提出的用于超分辨率 Sentinel-2 影像的 DSen2, 派生自 EDSR (Lim 等, 2017), 骨干是 ResNet (He 等, 2016)。主要包括卷积层、ReLU 和过连接。采用长、加性的跳过连接 (skip connection), 将输入待修复影像与模型输出相加, 这意味着网络学习预测的是对输入待修复影像的残差校正图。网络主体由 D (网络深度) 个 ResBlock 组成, 每个 ResBlock 又由两个卷积层、一个 ReLU 层、一个残差缩放层 (Res. Scal.) 以及一个残差连接的加性层构成。网络输入包括待修复影像 (当日 Landsat 影像)、缺失掩膜 (当日 Landsat 影像条带、云和阴影掩膜)、上采样 (upsample) 的多源引导影像 (当日 MODIS 影像)、季节时相剖面

(待修复日期所在季节所有的影像分位数、物候等特征)。网络最终输出是当日完整的修复影像。判别网络部分, 输入是修复影像与真实影像, 主要包括卷积层、ReLU、全连接层和 Sigmoid 函数。为训练网络, 构建了遵循全局随机分布的无云样本数据集, 并在 Landsat 影像上模拟条带、云和阴影等噪声。假设 \hat{y} 是修复影像, y 是真实影像, 网络的损失函数 \mathcal{L} 如下:

$$\mathcal{L} = \mathcal{L}_{MSE} + \alpha \mathcal{L}_{PER} + \beta \mathcal{L}_{ADV} \quad (3)$$

式中, $\mathcal{L}_{MSE} = \sum_i \|\hat{y}_i - y_i\|_1$, 是基于影像 L1 距离的像素级 MSE 损失, $\mathcal{L}_{PER} = \sum_i \|\phi(\hat{y}_i) - \phi(y_i)\|_2^2$ 是测量特征空间中的距离的感知损失, $\phi(\cdot)$ 来自 ImageNet 数据集上预训练的 VGG 网络 (Simonyan 和 Zisserman 2014; Ren 等, 2016), $\mathcal{L}_{ADV} = \sum_i \log(1 - D_i(\hat{y}_i))$ 则是对抗学习模块的对抗损失, 有利于提高图像视觉效果 (Ma 等, 2020) (He 等, 2017), D_i 表示判别网络。

时空多源引导对抗修复网络

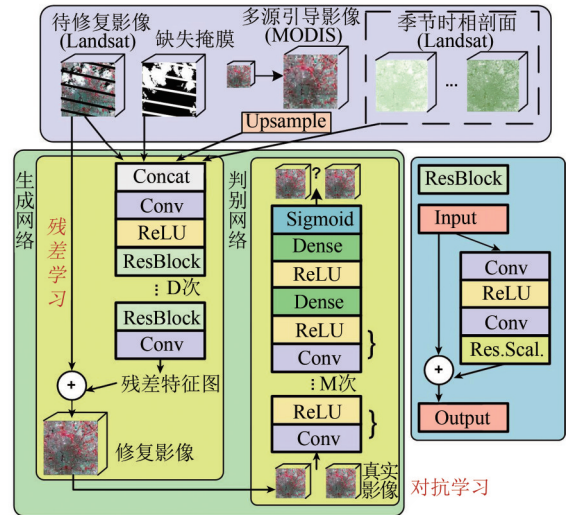


图 3 基于时空多源引导对抗修复网络的影像修复框架

Fig. 3 The framework for image reconstruction based on spatio-temporal multi-source guided adversarial reconstruction network

值得指出的是, 所提出的时空多源引导对抗修复网络, 可以结合多源的遥感数据进行互补修复引导, 包括 MODIS、SAR 等。通过结合时空信息、对抗学习、残差学习等范式, 能够高精度逼真修复影像任何条带、云或阴影导致的缺失。

2.1.3 时空融合

基于改进的时空自适应反射率融合算法 ESTARFM (Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model) (Zhu 等, 2010), 融合有观测日期重建的 Landsat 和 MODIS 数据。相比 STARFM 算法, ESTARFM 通过考虑混合像元效应, 极大提高了在异质区域的反演精度 (Zhu 等, 2010), 具有更好的适用性。对于给定区域, 算法假设在同一时间获取的来自不同卫星传感器的遥感数据具有可比性, 并且经过校正后彼此相关。同时考虑到具有相似反射率变化的相邻相同类别像素, 利用移动窗口方法来充分利用来自相邻像素的信息。对每天, 目标像素 $(x_{w/2}, y_{w/2}, t_p)$ Landsat 反射率 $F(x_{w/2}, y_{w/2}, t_p, B)$ 是两个就近参考观测日期 t_{r1} 、 t_{r2} 分别得到的预测的加权:

$$F(x_{w/2}, y_{w/2}, t_p, B) = T_{t_{r1}} \times F_{t_{r1}}(x_{w/2}, y_{w/2}, t_p, B) + T_{t_{r2}} \times F_{t_{r2}}(x_{w/2}, y_{w/2}, t_p, B) \quad (4)$$

式中, $T_{t_{r1}}$ 、 $T_{t_{r2}}$ 为时间权重, 依据 MODIS 对应参考时间和预测时间的反射率变化计算。 $F_k(x_{w/2}, y_{w/2}, t_p, B)$ 是基于参考时间 t_k 的预测值, $k = r1, r2$,

$$F_k(x_{w/2}, y_{w/2}, t_p, B) = F(x_{w/2}, y_{w/2}, t_k, B) + \sum_{i=1}^N W_i \times V_i \times (C(x_i, y_i, t_p, B) - C(x_i, y_i, t_k, B)) \quad (5)$$

式中, V_i 是第 i 个相似像素的转化系数, 根据粗分辨率像素内相似像素处的 Landsat 和 MODIS 的线性回归估计得到。 W_i 是第 i 个相似像素的权重, 由空间距离和光谱距离计算得到。

对于融合结果进一步经过 weighted Whittaker smoothing 处理, 得到长时序高一致性的高空间分辨率、高时间频率 (30 m、逐日) 的无缝数据立方体 SDC (Seamless Data Cube)。

2.1.4 特征立方体

基于 SDC, 本文进一步构建了特征立方体 (feature cube), 计算了归一化植被指数 NDVI (Normalized Difference Vegetation Index) (Townshend 和 Justice, 1986), 改进的归一化水体指数 MNDWI (Modified Normalized Difference Water Index) (Xu 2006), 归一化建筑物指数 NDBI (Normalized Difference Built-up Index) (Zha 等, 2003), 和归

一化燃烧比 NBR (Normalized Burn Ratio) (Miller 和 Thode, 2007) 4 个常用的代表性指数。取各个波段每年的 0, 25, 50, 75, 100 分位数作为年际物候特征。本文按春季 (3 月、4 月、5 月)、夏季 (6 月、7 月、8 月)、秋季 (9 月、10 月、11 月)、冬季 (12 月、1 月、2 月) 的划分, 计算了各个波段每个季节的分位数作为季节物候特征。除光谱特征外, 还采用多套窗口大小计算了空间特征, 如灰度共生矩阵 GLCM (Gray Level Co-occurrence Matrix)、拓展形态学属性剖面 (EMAPs) (Song 等, 2014) 等。其他辅助特征包括地理特征, 经纬度, 以及地形特征海拔、坡度、坡向。其中海拔数据来自美国航天局的 30 m 海拔数据 SRTM (Shuttle Radar Topography Mission)。

2.2 长时序高分辨率地表动态制图框架

2.2.1 分类系统

采用的分类系统涵盖土地覆盖、土地利用, 共 3 种类型。土地覆盖分类系统来自 FROM-GLC 分类系统 (Gong 等, 2013) 和 LCCS (Bartholome 和 Belward, 2005), 土地利用分类系统来自中国土地利用制图 (CNLUCC) 产品 (Xu 等, 2018) 所用的土地利用分类系统。这些分类体系均有两级结构、超过 20 种类别。FROM-GLC 分类系统有 10 个一级类型, 29 个二级类型 (L 等, 2017)。LCCS 分类系统有 8 个一级类型, 33 个二级类型 (L 等, 2017)。CNLUCC 分类系统有 6 个一级类型, 25 个二级类型 (Xu 等, 2018)。

2.2.2 样本集和样本迁移

土地覆盖样本来自首套全球 30 m 分辨率多季节样本集 (Li 等, 2017), 具有 FROM-GLC 和 LCCS 两种分类系统标签。该样本集是多名有丰富经验的影像解译者从 2015 年左右获取的 Landsat 影像中收集, 具有严格的解译和质量检查方案。其训练样本具有代表性和典型性, 验证样本具有广泛性和系统性 (Gong 等, 2013; Zhao 等, 2014)。全球首套 10 m 土地覆盖产品 FROM-GLC10 (Gong 等, 2019) 基于该样本集生产。中国范围内, 训练样本有 5500 个样本位置, 验证样本有 2100 个样本位置, 样本空间和类别分布分别如图 4 (a) 图 4 (b) 所示 (审图号: GS (2020) 7429 号)。

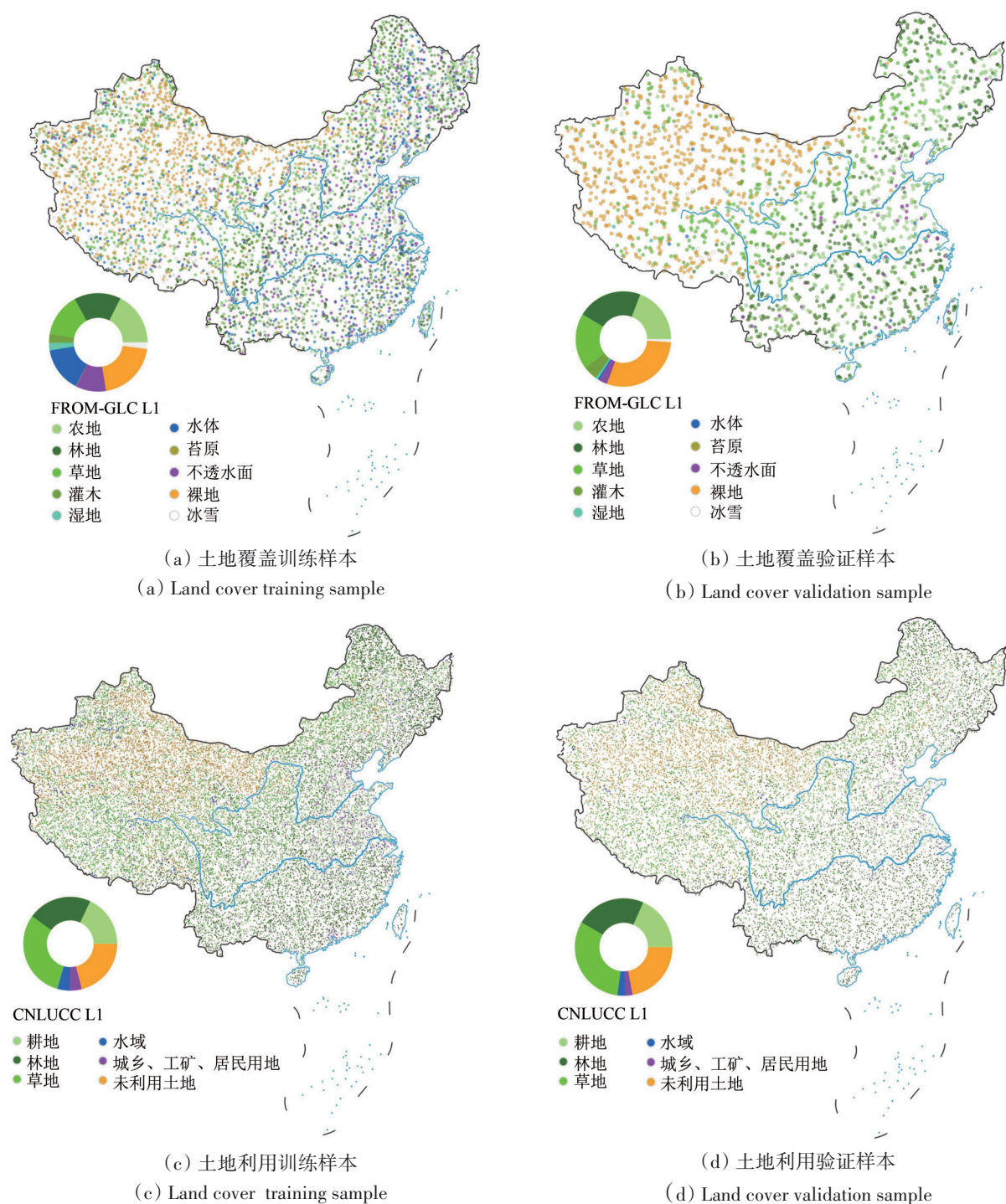


图4 中国制图样本空间分布(土地覆盖:不同颜色代表不同的FROM-GLC一级类,各类比例如饼图;土地利用:不同颜色代表不同的CNLUCC一级类,各类比例如饼图)

Fig. 4 Spatial distribution of mapping samples (Land cover: different colors represent different FROM-GLC level 1 classes, the proportions of each type are shown in the pie chart; Land use: different colors represent different CNLUCC level 1 classes, the proportions of each class are shown in the pie chart)

土地利用样本从CNLUCC产品(Xu等, 2018)中采集, 21世纪以来包括2000年、2005年、2010年、2015年4期。先前研究已经证明了使用现有制图数据可作为分类训练的有效来源(Zhang和Roy 2017; Song等, 2018; Brown等, 2020; Liu等,

2020)。为了保证样本的可靠性和可信性, 进行了筛选。首先, 样本必须来自类别的同质区域, 避免类别边缘, 以减少可能的类别混淆(Brown等, 2020)。其次, 当一个像素周围的8个像素相同时, 中心的单元格才被认为是更可靠的样本(Wang

等, 2019; Brown 等, 2020)。为了使类别比例符合实际情况, 本文从这些位置中, 进行了分层随机采样。最终一共保留了46万随机点做训练样本, 11万随机点做验证样本, 训练和验证样本独立且符合分布。样本数量级大, 且具有多期标签, 空间分布如图4(c)—图4(d)。

对于土地覆盖分类系统, 将2015年分类器, 用于所有年份分类。而土地利用分类系统, 对每年, 采用最邻近的1期分类器(2000年、2005年、2010年、2015年共4期)。尽管样本的迁移使用会引入一定错误, 但有限样本稳定分类理论(Gong 等, 2019; Huang 等, 2020)指出了样本迁移使用的有效性。该理论显示, 使用经全球制图反复试验证明表现突出的随机森林算法, 在采用训练样本量仅占整体的40%, 或训练样本的错误比例达到20%时, 分类精度的损失仅为1%。

2.2.4 分类与后处理

Landsat 数据空间分辨率是30 m, 对空间上下文信息的依赖相对不高, 此处没有采用基于卷积网络的语义分割类算法进行分类。考虑到重建的SDC像素层面的质量较高且信息丰富, 我们采用基于像素的分类算法, 从而使得结果能够保持更多丰富的纹理细节。基于自动机器学习 AutoML (Automated machine learning) 的思想, 采用 AutoGluon 工具 (Erickson 等, 2020) 训练分类器, 自动完成模型选择、超参数选择和调整等过程。采用的分类器包括随机森林 (Random Forest)、超随机树 (Extremely Randomized Trees)、k 最近邻 (k-Nearest Neighbors)、LightGBM 提升树 (LightGBM boosted trees) (Ke 等, 2017)、CatBoost 提升树

(CatBoost boosted trees) (Dorogush 等, 2018)、神经网络 (neural networks), 并通过模型集成, 进一步地提高分类精度。为减少分类结果中噪声干扰, 我们对分类结果进行了后处理。采用最大后验概率马尔科夫随机场 MAP-MRF (Maximum a Posteriori Markov Random Fields) (Liu 和 Cai, 2012; Wang 等, 2015) 对逐年分类标签和概率进行改进, 减少空间噪声。同时采用基于季节和趋势分解的断点识别算法 BFAST (Breaks For Additive Seasonal and Trend) (Verbessell 等, 2010) 进行变化检测, 每一时间段内, 将更新为该段最高置信度的类别, 提高时序一致性和可比性。

3 结果

3.1 首套21世纪中国30 m逐日无缝数据立方体

基于提出的SDC重建技术, 我们生产了首套21世纪以来中国逐日30 m分辨率SDC。结果展示如图5所示(审图号: GS(2020)7429号), 其中图5(a)是2019年中值合成的假彩色影像, 图5(b)—图5(e)是分别从2019年4个季节中随机抽取的一天的中国影像进行镶嵌。整体来说, SDC空间覆盖完整, 具有空间异质性且较为合理, 能很好地反映地表的真实状况。同时, SDC通过提供逐日影像, 可以充分展现任一地点随时间变化的物候等动态信息, 首次实现了高分辨率无缝遥感。此外, 在SDC光谱波段外, 为每幅影像配置了详细的质量控制(QA)波段, 记录影像处理过程的所有信息, 包括原始质量信息, 如是否为云、阴影、缺失值等, 以及重建处理过程和可信度等, 具有高透明度和高可操作性。

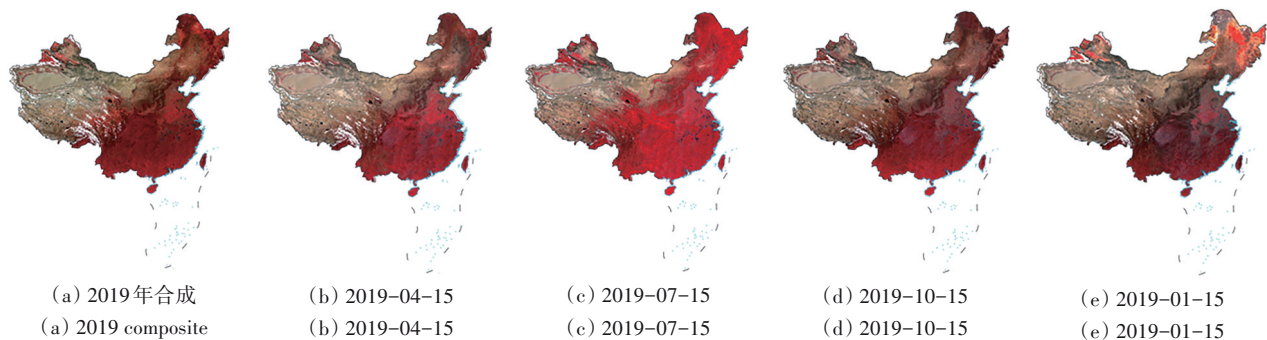


图5 中国无缝数据立方体假彩色合成示例

Fig. 5 Example of false color mosaic for China seamless data cube

为了检验重建模型的精度和SDC的质量, 本文从中国随机选择了500个tile进行检验。参考(Yan和Roy, 2020)的方法, 本文选择有效值 $\geq 20\%$ 的用于重建, 从剩下有效值 $< 20\%$ 的数据中选择有效值 $\geq 2\%$ 的用于检验。因此测试数据是独立的。此外, 对测试数据进行了进一步滤波。考虑可能残留的部分云或阴影(Egorov等, 2019; Qiu等, 2019)影响, 本文对测试数据的云和阴影mask进行20像素的膨胀, 并舍弃这部分观测值, 以尽可能减小云或阴影的发生率, 保证测试数据质量。如图6所示, SDC重建预测的反射率值与实际的观测值能很好的贴合, 其中SWIR1、SWIR2

波段相关系数R高, 分别为0.91、0.90; Blue波段相对较低, 为0.84, 可能与蓝光波段受大气影响严重、质量相对差有关。以RMSE来说, 平均水平为0.025左右, 显示预测与实际观测的差距很小, 具有较高的精度。就波段来说, 近红外、SWIR1、SWIR2的RMSE相对其他波段高, 与这些波段反射率水平偏高有关。就地区来看, 东南地区RMSE水平最低, 东北、华中地区较高。总体来看, 本文提出影像重建技术方案具有高可行性, 生产的高分辨率、高时间频率的SDC具有高可靠性、高精度的特点。

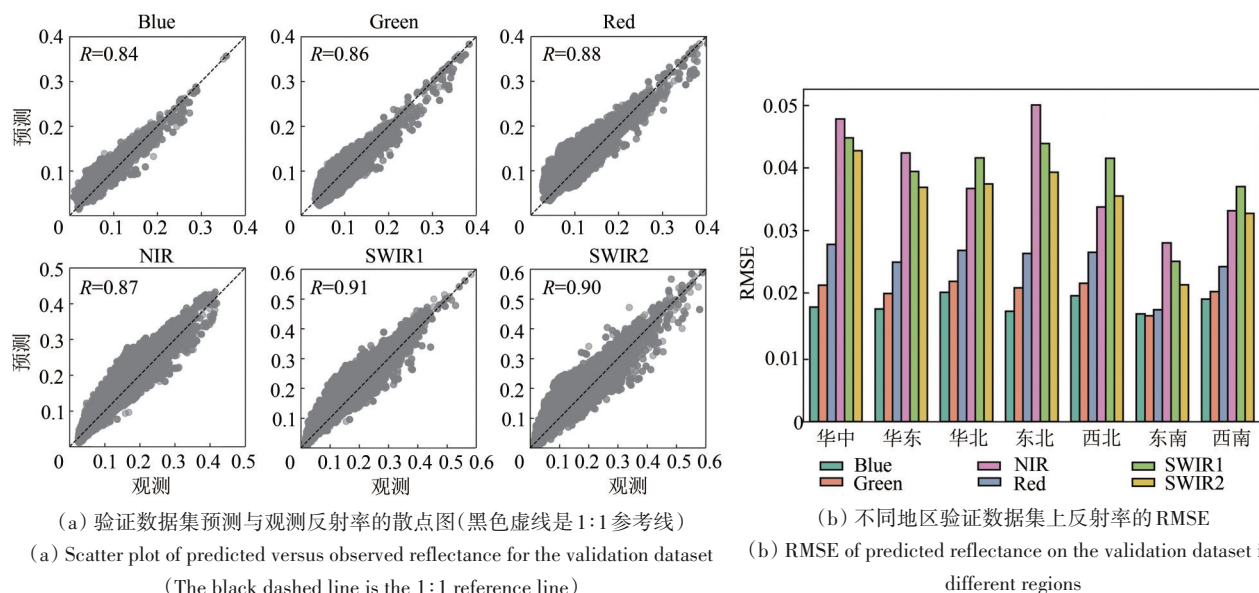


图6 无缝数据立方体验证结果

Fig. 6 Seamless data cube validation results

3.2 首套21世纪中国30 m逐年逐季节土地覆盖、土地利用制图集

基于提出的长时序高分辨率地表覆盖制图框架, 生产了首套21世纪以来中国30 m分辨率逐年土地利用和逐季节土地覆盖图集。2019年FROM-GLC、LCCS和CNLUCC等3种分类体系下制图结果依次显示在图7(a)、图7(b)、图7(c)(审图号: GS(2020)7429号)。从图7中可以看出, 3种分类体系下反映的土地覆盖土地利用空间格局基本一致。西部、西北部地区的沙漠和草地地区, 土地覆盖与土地利用分类体系下的制图结果有较大的差别。在西藏、内蒙古中部地区。土地覆盖结果中分为裸地, 土地利用结果中分为草地。这种差异来自

于土地覆盖与土地利用分类体系中对裸地、草地的定义、判别标准不同有关。2019年中国季节土地覆盖制图结果如图8所示(审图号: GS(2020)7429号)。可以看到在FROM-GLC和LCCS两种分类体系下, 季节动态变化都是显著的。春夏秋冬四个季节制图结果相比较, 夏季的土地格局与年度制图结果更为相似。这与夏季植被生长态势旺盛, 与其他地类更容易区分有关。同时, 夏季受季节性冰雪影响小, 而秋冬季节土地制图中, 冰雪表征多。

局部区域结果显示了我们的制图结果可有效而细致反映年际土地变化动态。图9显示的分别是重庆、西安的结果。重庆位于多云且复杂的山区, 原始影像质量差, 制图难度大。图9显示我

们的制图在重庆取得了准确而一致的结果，且对城市扩张、水体等有良好的刻画，而ESA CCI的结果缺少细节且不够准确。在西安，制图结果显示

了细致的地表细节，尤其在村庄的识别上相比ESA CCI具有更高的准确性。

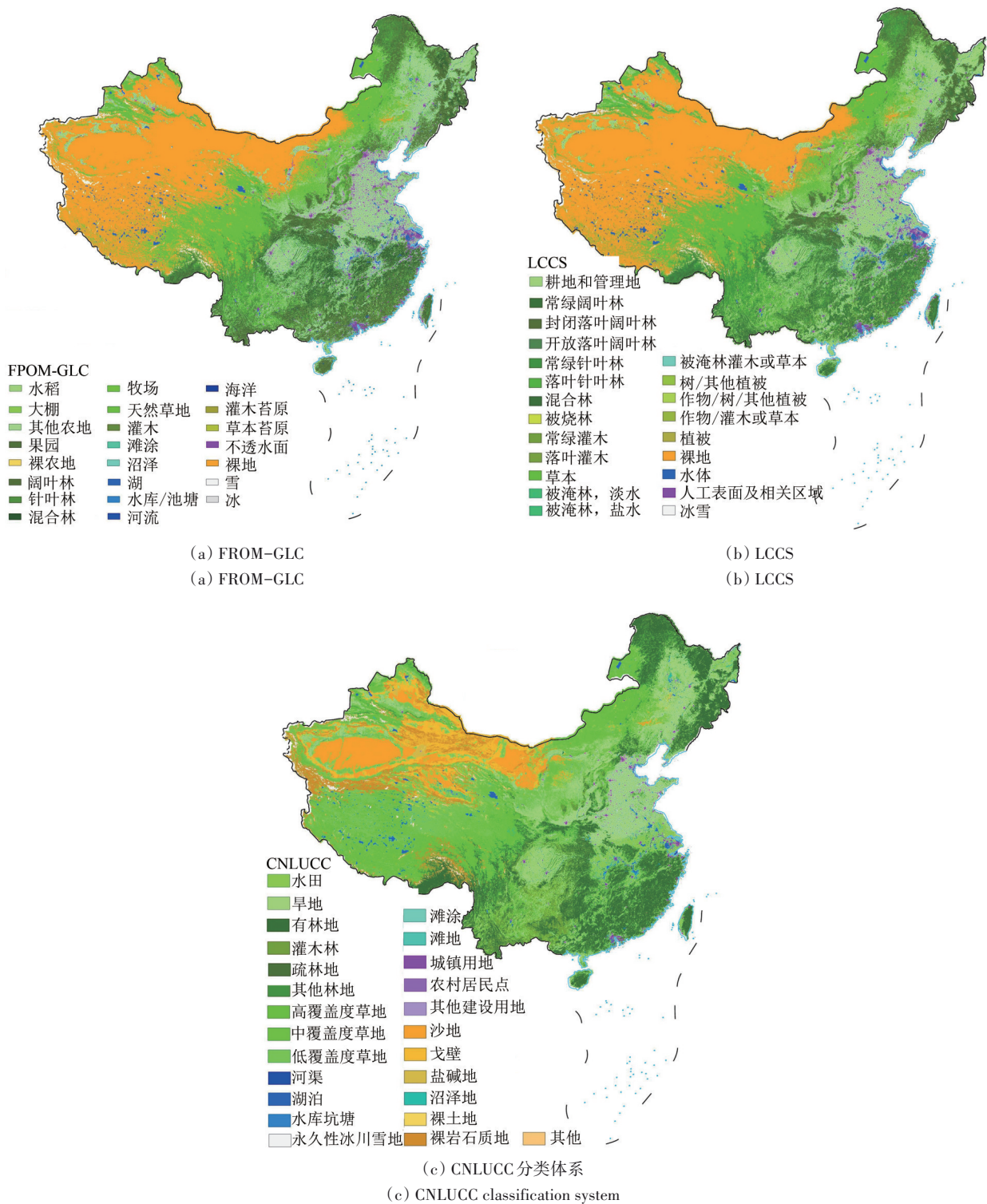


图 7 2019年中国土地覆盖和土地利用图
Fig. 7 Land cover and land use map of China in 2019

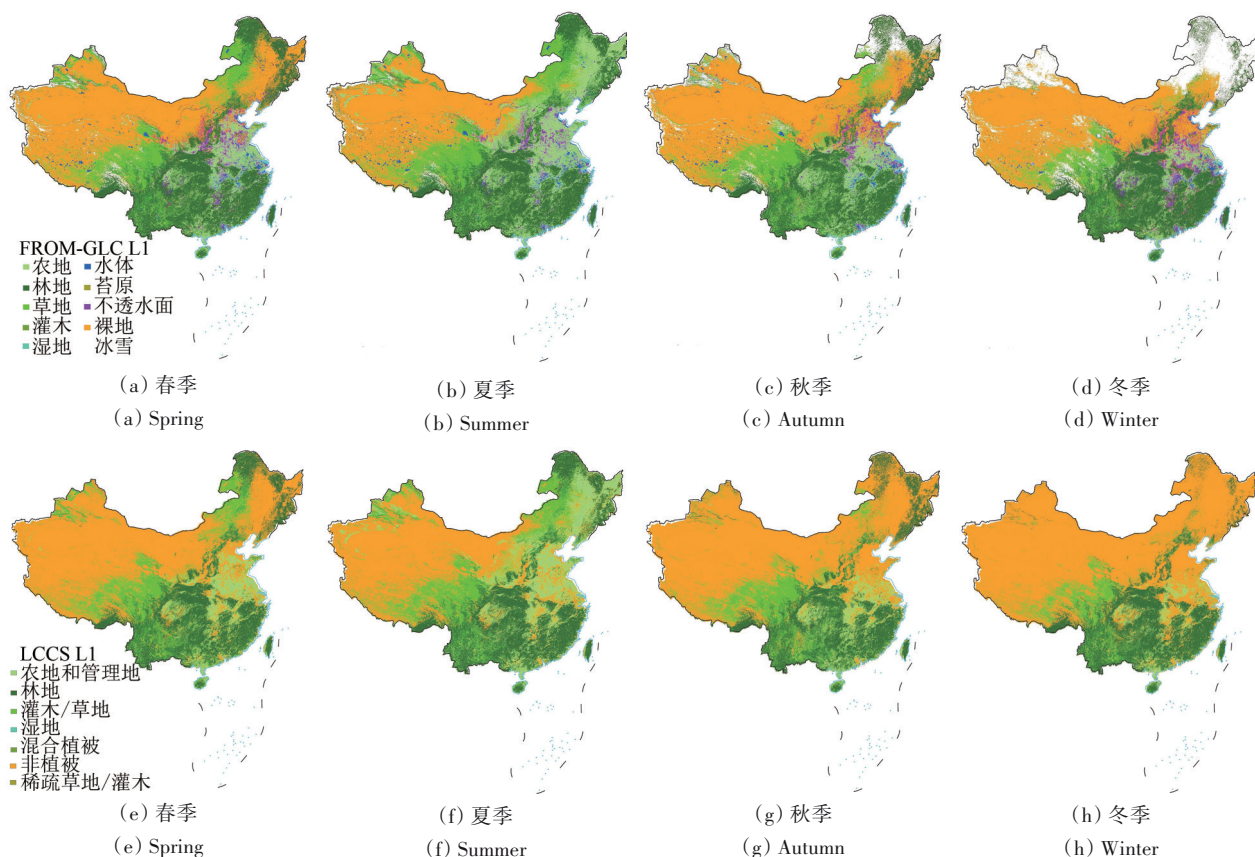


图 8 2019年中国季节土地覆盖图

Fig. 8 Seasonal land cover map of China in 2019

利用制图结果,统计中国 21 世纪以来土地变化情况如图 10 所示。从曲线来看,本文的制图结果年际波动小,时间一致性强。根据统计结果,耕地面积在 2000 年—2019 年间呈显著线性下降趋势 ($p < 0.001$),下降速率为 $7684 \pm 814 \text{ km}^2/\text{a}$ (95% 置信区间)。2019 年中国耕地面积约为 $1.67 \times 10^6 \text{ km}^2$,未低于中国的耕地红线 (约 $1.20 \times 10^6 \text{ km}^2$)。森林在 19 年间从 $2.29 \times 10^6 \text{ km}^2$ 增加到 $2.31 \times 10^6 \text{ km}^2$,增加速率为 $1670 \pm 176 \text{ km}^2/\text{a}$ (95% 置信区间)。森林面积的显著增加 ($p < 0.001$) 可能反映了中国退耕还林等生态工程的效果。不透水面是显著变化的另一大土地覆盖类别 ($p < 0.001$),2000 年—2019 年间约增加了 134755 km^2 ,相比 2000 年增加了约 139%。

总体而言,本文的制图产品在全国范围内的各类曲线与其他不同来源制图产品和统计数据较为一致 (图 10),得到的农地面积与 5 年间隔的 CNLUCC 产品接近,而 300 m 的 ESACCI 和 2000 年、2010 年两期的 Globeland30 产品农地面积估计相对高。从趋势来看,除 ESACCI 显示农地增

加,其他的时序制图产品农地趋势均为减少。此外,对农地在 2015 年的估计与 30 m 的 GFSAD 和 FROM-GLC2015 的结果也很接近。来自 FAO 和中国统计年鉴的统计数字,相对我们的产品以及其他制图产品,不连续,且对农地面积估计偏低,反映了统计数字与遥感制图产品的差异。这可能一定程度上导致了我国对耕地红线的过分担忧。对于森林,本文数据与 ESACCI 和 CNLUCC 估计接近,Hansen 2000 年和 Globeland30 2000 年、2010 年两期的估计相对低。中国统计年鉴 2007 年前数据与估计接近并略高于本文结果,而 FAO 估计整体偏低,两种统计数据来源均显示森林面积增加,与本文产品增加趋势一致。裸地和草地两大类,由于各项产品间定义差别大,本文的估计与其他制图产品估计有差异。裸地和草地与具有同样分类体系的 FROM-GLC2015 年产品在 2015 年的面积更为接近。另一大主要类别,城市类别,可以看到中国统计年鉴的统计数字远远高于我们的估计和其他制图产品。城市类别面积与其他各种来源制图产品估计接近,略低于 CNLUCC 和

Globeland30, 而略高于 GAIA 和 ESACCI。其中 CNLUCC 和 Globeland30 城市面积略高, 主要是因为采用了人工解译方案, 以及其对该类别的定义倾向于居民区。而 ESACCI 是所有制图产品中对城市类别估计最少的, 这可能与分辨率粗, 识别出来的城市相对更少有关。尽管不同统计数字调查来源和方法不同, 不同制图产品的分类体系、制图方法、数据来源各不相同, 但总体来看, 我们的产品各类的面积和趋势都处于合理的范围内。

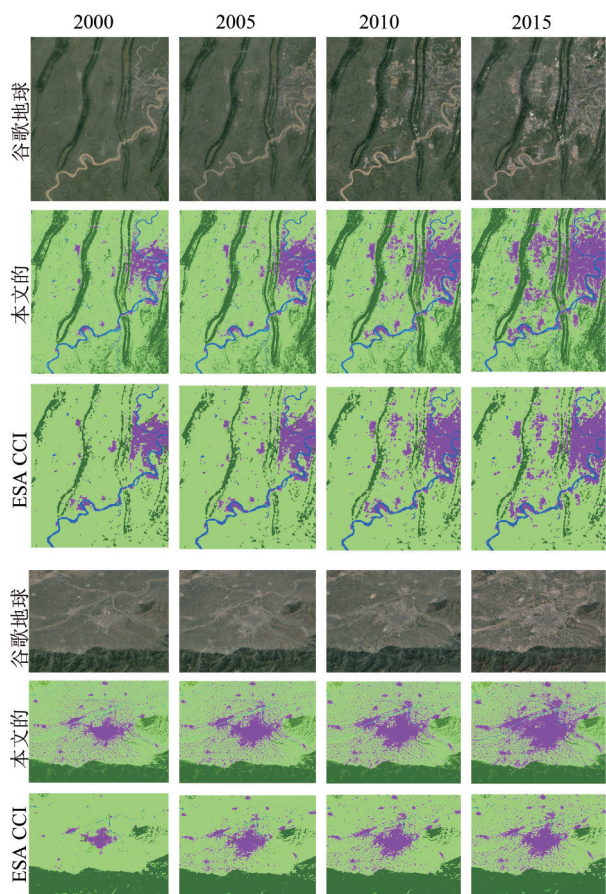


图9 中国重庆(西南)和西安(华中)制图结果比较(本文的: 2000年—2019年, 30 m; ESACCI: 1992年—2015年, 300 m)

Fig. 9 Comparison of mapping results for Chongqing (southwest China) and Xi'an (central China) (this paper: 2000—2019, 30 m; ESACCI: 1992—2015, 300 m)

图 11 显示的是 2019 年分类不确定性图 (审图号: GS (2020) 7429 号), 可以看出, 置信度高的地区往往是同质区域, 如西北裸地、华东农地地区。西南地区置信度相对较低, 可能与云雾条件有关, 此外, 黄土高原地区不确定性也相对较高, 可能与位于裸地、草地、林地等交错边界有关。

从类别来看, 水体、裸地、冰雪、林地是置信度高的几类, 而灌木、湿地两类相对较低, 也反映了这两类制图难度大的特点。

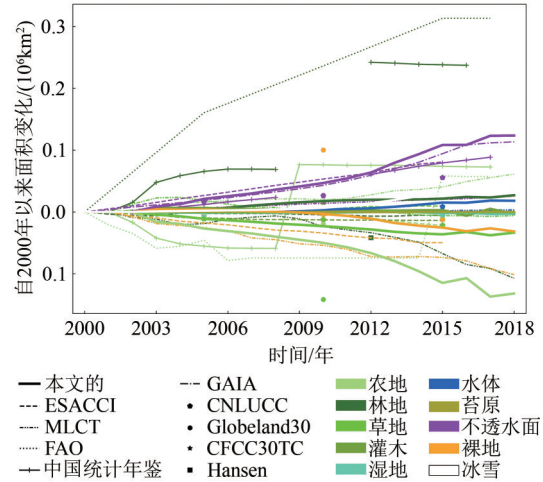


图 10 21 世纪以来中国各类地表覆盖面积变化

Fig. 10 Changes in the land cover area in China since the 21st century

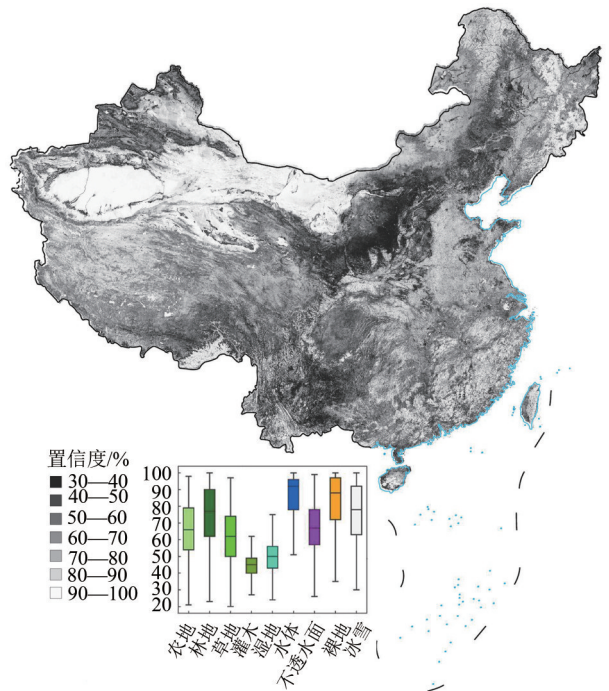


图 11 2019 年中国土地覆盖分类置信度图 (箱型图是每个类的置信度汇总)

Fig. 11 Classification confidence map in 2019 (The box plot is a summary of the confidence for each class)

使用验证样本集对制图结果进行精度评价。包括 2015 年土地覆盖验证数据, 4 期土地利用验证样本。以及随机选择并解译的 500 个逐年土地覆盖、土地利用验证样本。评价指标包括整体精

度 OA (Overall Accuracy), 卡帕系数 (Kappa coefficient), 各类用户精度 UA (user's accuracy) 和生产者精度 PA (producer's accuracy) 等。本文使用土地覆盖验证样本集对土地覆盖制图结果进行评估。

表 2 结果显示 2015 年 FROM-GLC 一级类土地覆盖制图结果整体精度为 80.8%, Kappa 系数为 0.75。其中林地、裸地两类 PA、UA 均超过了 80%。农地类别 PA、UA 分别达到 78.6% 和 84.6%。不透水面类别 PA 达到了 75.0%。灌木、湿地两类

由于样本比例较小, 精度相对较低。灌木一类主要跟林地、草地两类混分, 这 3 类定义根据植被高度和盖度进行划分, 通过光谱信息划分难度通常较高。最新的 LCMAP 产品将草地、灌木两类合为一类。苔原一类, 样本中数量极少, 导致分类精度低。严格说来, 分布在北纬 60° 以上地区, 中国的范围实际并不属于苔原分布区。FROM-GLC 二级类整体精度为 73.3%, Kappa 为 0.67, 除稀有类别外, 各类别精度令人满意。

表 2 2015 年 FROM-GLC 一级分类体系验证结果

Table 2 Validation results of the 2015 FROM-GLC level 1 classification system

类别	10	20	30	40	50	60	70	80	90	100	UA/%
10	308	26	16	8	0	0	0	5	1	0	84.6
20	34	450	9	19	0	0	0	1	0	0	87.7
30	23	25	281	23	1	1	0	3	56	0	68.0
40	6	47	22	28	0	0	0	0	5	0	25.9
50	1	2	1	0	3	8	0	1	1	0	17.7
60	0	1	0	0	1	11	0	0	0	0	84.6
70	0	0	1	0	0	0	0	0	1	0	0.0
80	16	1	2	1	0	0	0	36	1	0	63.2
90	4	1	23	0	1	2	0	2	614	5	94.2
100	0	0	2	0	0	1	0	0	5	15	65.2
PA/%	78.6	81.4	78.7	35.4	50.0	47.8	0.0	75.0	89.8	75.0	80.8

注: PA: 生产者精度, UA: 用户精度。

2015 年 LCCS 一级类、二级类分类结果 OA 为 79.0%、70.4%, kappa 为 0.72、0.64。耕地的 PA、UA 分别为 77.3%、80.5%。林地、非植被的 UA 分别达到 89.5%、94.3%。非植被的二级类人工表面及相关区域、水体两类 PA、UA 分别达到 86.00% 和 84.6%、71.67% 和 95.5%。其中灌木/草本的 UA 相对较低, 为 65.2%, 这再次说明灌木和草地分类难度大, 分类效果不理想。

对 CNLUCC 土地利用分类系统数据评估结果显示, 一级类 2000 年、2005 年、2010 年、2015 年的 OA 分别为 76.8%、77.0%、76.8%、76.0%。其中城乡、工矿、居民用地一类的 UA 偏低, 这与 CNLUCC 人工解译方式和对该类定义倾向于居民区有一定关系。几类与湿地相关的类别的 UA 相对较低, 与湿地类别分类难度大有关。湿地地区光谱时空变异性强, 地表状态变化大, 易与水体、草地、灌木等多类混分。

为了更好地掌握逐年制图精度, 本文利用逐

年样本对制图结果进行了检验。图 12 中显示的是制图结果的 OA。总体来看, 分类总体精度在年际间有所波动。就土地覆盖结果来说, 由于训练样本来自 2015 年, 因而分类精度总体来看 2015 年相对最高。From-GLC、LCCS 一级类的精度为 80.0%±0.55%、78.4%±0.37%, level-2 精度为 72.0%±0.94%、70.0%±0.47%。CNLUCC 土地利用分类结果 2015 年 OA 为 77.2%, 所有年份平均为 77.2%。逐年精度结果总体令人满意。此外, 为了检验基于变化检测的时间一致性处理的效果, 我们同时检验了 FROM-GLC 一级类未经过该处理时的精度, 结果如图 12 中红色线显示。除 2003 年、2011 年外, 经过时间一致性调整的制图结果的精度均高于未进行处理的结果。这揭示了变化检测的有效性。

2015 年 FROM-GLC 一级类分类体系下季节土地覆盖分类结果, 春夏秋冬的 OA 分别为 68.7%、70.5%、65.9%、66.9%。夏季分类精度最高, 这可

能与夏季为生长季,植物状态更好分辨有关。总体来看,季节制图结果精度低于年际制图结果,这与年际制图每年可用观测量多于一个季节的观测量有关。观测量少,云、大气状况、噪声等因素相对影响大,将严重干扰分类结果。

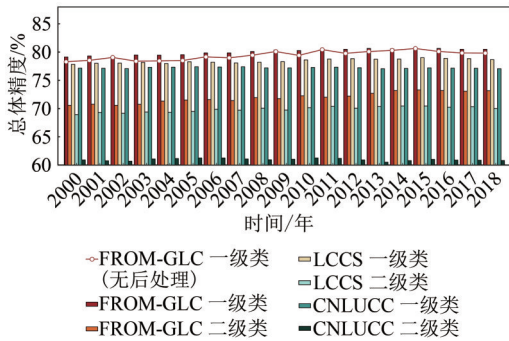


图 12 逐年不同分类系统下制图结果精度

Fig. 12 Annual accuracy of mapping results under different classification systems

4 结 论

本文提出了智慧遥感制图的概念和一套智慧遥感制图的框架,该框架从用户需求出发、问题驱动,能够极大地改善当前数据产品难以满足农林管理、国情监测、生态环境保护、防灾减灾、城市建设等用户多样化、高精度地表监测需求的现状。

在该框架的指导下,本文基于AWS高性能、高弹性、可扩展的分布式计算资源,搭建了在线实时、自动化、无服务器、端到端的遥感大数据生产链和并行制图系统,并生产了首套21世纪中国全境逐日SDC及逐年逐季节土地覆盖和土地利用制图产品。逐日SDC,填补了高时空分辨率遥感观测的空白,作为ARD,对于高精度定量遥感反演和制图具有较大价值。逐年逐季节地表制图产品,平均精度超过80%,具有高时空一致性。这两套产品能够以前所未有的高空间分辨率和时间频率反映地表格局与动态变化,为科学研究、国民经济、社会发展提供了重要的基础地理数据,同时证明了智慧制图框架的可行性和有效性。未来,我们将进一步完善和发展智慧遥感制图理论和框架,以开放和灵活的理念,为促进中国遥感数据的易用和好用提供新思路。

志 谢 亚马逊云服务(AWS)、青岛光环启云数据科技有限公司、清华海峡研究院(厦门)

AI for Earth 联合实验室为此研究中两套产品的研发,提供了大量协助和支持,在此表示衷心的感谢!

参考文献(References)

- Bai Y, Wong C P, Jiang B, Hughes A C, Wang M and Wang Q. 2018. Developing China's Ecological Redline Policy using ecosystem services assessments for land use planning. *Nature Communications*, 9(1): 3034
- Bartholome E and Belward A S. 2005. GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing*, 26(9): 1959-1977
- Baumann P, Mazzetti P, Ungar J, Barbera R, Barboni D, Beccati A, Bigagli L, Boldrini E, Bruno R, Calanducci A, Campalani P, Clements O, Dumitru A, Grant M, Herzig P, Kakalettris G, Laxton J, Koltsida P, Lipskoch K, Mahdiraji A R, Mantovani S, Merticariu V, Messina A, Misev D, Natali S, Nativi S, Oosthoek J, Pappalardo M, Passmore J, Rossi A P, Rundo F, Sen M, Sorbera V, Sullivan D, Torrisi M, Trovato L, Veratelli M G and Wagner S. 2016. Big Data Analytics for Earth Sciences: the EarthServer approach. *International Journal of Digital Earth*, 9(1): 3-29
- Baumann P, Rossi A P, Bell B, Clements O, Evans B, Hoenig H, Hogan P, Kakalettris G, Koltsida P and Mantovani S. 2018. *Fostering cross-disciplinary earth science through datacube analytics*. Earth Observation Open Science and Innovation, Springer, Cham: 91-119
- Benabdelkader S and Melgani F. 2008. Contextual spatio-spectral postreconstruction of cloud-contaminated images. *IEEE Geoscience and remote sensing letters*, 5(2): 204-208
- Beusch L, Gudmundsson L and Seneviratne S I. 2020. Crossbreeding CMIP6 Earth System Models With an Emulator for Regionally Optimized Land Temperature Projections. *Geophysical Research Letters*, 47(15): e2019GL086812
- Brooks E B, Wynne R H, Thomas V A, Blinn C E and Coulston J W. 2014. On-the-Fly Massively Multitemporal Change Detection Using Statistical Quality Control Charts and Landsat Data. *IEEE Transactions on Geoscience and Remote Sensing*, 52(6): 3316-3332.
- Brown J F, Tollerud H J, Barber C P, Zhou Q, Dwyer J L, Vogelmann J E, Loveland T R, Woodcock C E, Stehman S V, Zhu Z, Pengra B W, Smith K, Horton J A, Xian G, Auch R F, Sohl T L, Saylor K L, Gallant A L, Zelenak D, Reker R R and Rover J. 2020. Lessons learned implementing an operational continuous United States national land change monitoring capability: The Land Change Monitoring, Assessment, and Projection (LCMAP) approach. *Remote Sensing of Environment*, 238: 111356
- Bullock E L, Woodcock C E and Holden C E. 2020. Improved change monitoring using an ensemble of time series algorithms. *Remote*

- Sensing of Environment, 238: 111165
- Chen J, Chen J, Liao A, Cao X, Chen L, Chen X, He C, Han G, Peng S and Lu M. 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103: 7-27
- Chen J, Jönsson P, Tamura M, Gu Z, Matsushita B and Eklundh L. 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote sensing of Environment*, 91(3-4): 332-344
- Chen J, Zhu X, Vogelmann J E, Gao F and Jin S. 2011. A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. *Remote sensing of environment*, 115(4): 1053-1064
- Cheng Q, Shen H, Zhang L, Yuan Q and Zeng C. 2014. Cloud removal for remotely sensed images by similar pixel replacement guided with a spatio-temporal MRF model. *ISPRS journal of photogrammetry and remote sensing*, 92: 54-68
- Crowther T W, Glick H B, Covey K R, Bettigole C, Maynard D S, Thomas S M, Smith J R, Hintler G, Duguid M C, Amatulli G, Tuanmu M N, Jetz W, Salas C, Stam C, Piotto D, Tavani R, Green S, Bruce G, Williams S J, Wisser S K, Huber M O, Hengeveld G M, Nabuurs G J, Tikhonova E, Borchardt P, Li C F, Powrie L W, Fischer M, Hemp A, Homeier J, Cho P, Vibrans A C, Umunay P M, Piao S L, Rowe C W, Ashton M S, Crane P R and Bradford M A. 2015. Mapping tree density at a global scale. *Nature*, 525 (7568): 201-205
- Dorogush A V, Ershov V and Gulin A. 2018. CatBoost: gradient boosting with categorical features support. *arXiv preprint arXiv:1810.11363*
- Dwyer J L, Roy D P, Sauer B, Jenkerson C B, Zhang H K and Lymburner L. 2018. Analysis ready data: Enabling analysis of the Landsat archive. *Remote Sensing*, 10(9): 1363
- Egorov A V, Roy D P, Zhang H K, Li Z, Yan L and Huang H. 2019. Landsat 4, 5 and 7 (1982 to 2017) Analysis Ready Data (ARD) Observation Coverage over the Conterminous United States and Implications for Terrestrial Monitoring. *Remote Sensing*, 11(4)
- Eilers P H C. 2003. A Perfect Smoother. *Analytical Chemistry*, 75(14): 3631-3636
- Erickson N, Mueller J, Shirkov A, Zhang H, Larroy P, Li M and Smola A. 2020. AutoGluon-Tabular: Robust and Accurate AutoML for Structured Data. *arXiv preprint arXiv:2003.06505*
- Feddema J J, Oleson K W, Bonan G B, Mearns L O, Buja L E, Meehl G A and Washington W M. 2005. The importance of land-cover change in simulating future climates. *Science*, 310(5754): 1674-1678
- Foley J A, DeFries R, Asner G P, Barford C, Bonan G, Carpenter S R, Chapin F S, Coe M T, Daily G C and Gibbs H K. 2005. Global consequences of land use. *science*, 309(5734): 570-574
- Footy G M. 2010. Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sensing of Environment*, 114(10): 2271-2285
- Friedl M. A, Sulla-Menashe D, Tan B, Schneider A, Ramankutty N, Sibley A and Huang X. 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114(1): 168-182
- Gao F, Masek J, Schwaller M and Hall F. 2006. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote sensing*, 44(8): 2207-2218
- Giuliani G, Chatenoux B, Bono A De, Rodila D, Richard J P, Allenbach K, Dao H and Peduzzi P. 2017. Building an Earth Observations Data Cube: lessons learned from the Swiss Data Cube (SDC) on generating Analysis Ready Data (ARD). *Big Earth Data*, 1(1-2): 100-117
- Gómez C, White J C and Wulder M A. 2016. Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*. 116: 55-72
- Gong P, Li X, Wang J, Bai Y, Chen B, Hu T, Liu X, Xu B, Yang J, Zhang W and Zhou Y. 2020. Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. *Remote Sensing of Environment*, 236: 111510
- Gong P, Li X and Zhang W. 2019. 40-Year (1978—2017) human settlement changes in China reflected by impervious surfaces from satellite remote sensing. *Science Bulletin*, 64(11): 756-763
- Gong P, Liu H, Zhang M, Li C, Wang J, Huang H, Clinton N, Ji L, Li W, Bai Y, Chen B, Xu B, Zhu Z, Yuan C, Suen H P, Guo J, Xu N, Li W, Zhao Y, Yang J, Yu C, Wang X, Fu H, Yu L, Dronova I, Hui F, Cheng X, Shi X, Xiao F, Liu Q and Song L. 2019. Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017. *Science Bulletin*, 64(6): 370-373
- Gong P, Wang J, Yu L, Zhao Y, Zhao Y, Liang L, Niu Z, Huang X, Fu H and Liu S. 2013. Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing*, 34(7): 2607-2654
- Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D and Moore R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*. 202: 18-27
- Hamilton S E and Casey D. 2016. Creation of a high spatio-temporal resolution global database of continuous mangrove forest cover for the 21st century (CGMFC-21). *Global Ecology and Biogeography*, 25(6): 729-738
- Hansen M C, DeFries R S, Townshend J R G and Sohlberg R. 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. *International journal of remote sensing*, 21(6-7): 1331-1364
- Hansen M C and Loveland T R. 2012. A review of large area monitoring of land cover change using Landsat data. *Remote sensing of Environment*, 122: 66-74
- Hansen M C, Potapov P V, Moore R, Hancher M, Turubanova S A,

- Tyukavina A, Thau D, Stehman S V, Goetz S J, Loveland T R, Kommareddy A, Egorov A, Chini L, Justice C O and Townshend J R G. 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 342(6160): 850
- He C, Liu Z, Gou S, Zhang Q, Zhang J and Xu L. 2018. Detecting global urban expansion over the last three decades using a fully convolutional network. *Environmental Research Letters*
- He K, Zhang X, Ren S and Sun J. 2016. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*
- He Z, Liu H, Wang Y and Hu J. 2017. Generative Adversarial Networks-Based Semi-Supervised Learning for Hyperspectral Image Classification. *Remote Sensing*, 9(10)
- Homer C, Dewitz J, Fry J, Coan M, Hossain N, Larson C, Herold N, McKerrow A, VanDriel J N and Wickham J. 2007. Completion of the 2001 national land cover database for the conterminous United States. *Photogrammetric engineering and remote sensing*, 73(4): 337
- Homer C, Dewitz J, Jin S, Xian G, Costello C, Danielson P, Gass L, Funk M, Wickham J, Stehman S, Auch R and Riitters K. 2020. Conterminous United States land cover change patterns 2001—2016 from the 2016 National Land Cover Database. *ISPRS Journal of Photogrammetry and Remote Sensing*, 162: 184-199
- Homer C, Dewitz J, Yang L, Jin S, Danielson P, Xian G, Coulston J, Herold N, Wickham J and Megown K. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States—representing a decade of land cover change information. *Photogrammetric Engineering & Remote Sensing*, 81(5): 345-354
- Homer C, Huang C, Yang L, Wylie B and Coan M. 2004. Development of a 2001 national land-cover database for the United States. *Photogrammetric Engineering & Remote Sensing*, 70(7): 829-840
- Hu S, Niu Z, Chen Y, Li L and Zhang H. 2017. Global wetlands: Potential distribution, wetland loss, and status. *Science of The Total Environment*, 586: 319-327
- Huang H, Wang J, Liu C, Liang L, Li C and Gong P. 2020. The migration of training samples towards dynamic global land cover mapping. *ISPRS Journal of Photogrammetry and Remote Sensing*, 161: 27-36
- Ji L, Gong P, Wang J, Shi J and Zhu Z. 2018. Construction of the 500 m Resolution Daily Global Surface Water Change Database (2001—2016). *Water Resources Research*, 54(12): 10, 270-210, 292
- Julien Y and Sobrino J A. 2010. Comparison of cloud-reconstruction methods for time series of composite NDVI data. *Remote Sensing of Environment*, 114(3): 618-625
- Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q and Liu T Y. 2017. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*
- Kennedy R E, Yang Z and Cohen W B. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. *Remote Sensing of Environment*, 114(12): 2897-2910
- Killough B. 2018. Overview of the open data cube initiative. *IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE*
- Killough B. 2019. The Impact of Analysis Ready Data in the Africa Regional Data Cube. *IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, IEEE*
- Kong D, Zhang Y, Gu X and Wang D. 2019. A robust method for reconstructing global MODIS EVI time series on the Google Earth Engine. *ISPRS Journal of Photogrammetry and Remote Sensing*, 155: 13-24
- Lanaras C, Bioucas-Dias J, Galliani S, Baltsavias E and Schindler K. 2018. Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146: 305-319
- Lewis A, Lymburner L, Purss M B J, Brooke B, Evans B, Ip A, Dekker A G, Irons J R, Minchin S, Mueller N, Oliver S, Roberts D, Ryan B, Thankappan M, Woodcock R and Wyborn L. 2016. Rapid, high-resolution detection of environmental change over continental scales from satellite data - the Earth Observation Data Cube. *International Journal of Digital Earth*, 9(1): 106-111
- Lewis A, Oliver S, Lymburner L, Evans B, Wyborn L, Mueller N, Raevksi G, Hooke J, Woodcock R and Sixsmith J. 2017. The Australian geoscience data cube—Foundations and lessons learned. *Remote Sensing of Environment*, 202: 276-292
- Li C, Gong P, Wang J, Zhu Z, Biging G S, Yuan C, Hu T, Zhang H, Wang Q, Li X, Liu X, Xu Y, Guo J, Liu C, Hackman K O, Zhang M, Cheng Y, Yu L, Yang J, Huang H and Clinton N. 2017. The first all-season sample set for mapping global land cover with Landsat-8 data. *Science Bulletin*, 62(7): 508-515
- Lim B, Son S, Kim H, Nah S and Lee K M. 2017. Enhanced deep residual networks for single image super-resolution. *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*
- Liu D and Cai S. 2012. A spatial-temporal modeling approach to reconstructing land-cover change trajectories from multi-temporal satellite imagery. *Annals of the Association of American Geographers* 102(6): 1329-1347
- Liu H, Gong P, Wang J, Clinton N, Bai Y and Liang S. 2020. Annual Dynamics of Global Land Cover and its Long-term Changes from 1982 to 2015. *Earth System Science Data*, 12(2): 1217-1243
- Liu H, He L and Li J. 2017. Remote sensing image classification based on convolutional neural networks with two-fold sparse regularization. *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*
- Liu H, Li J, He L and Wang Y. 2019. Superpixel-Guided Layer-Wise Embedding CNN for Remote Sensing Image Classification. *Remote Sensing*, 11(2)
- Liu J, Liu M, Tian H, Zhuang D, Zhang Z, Zhang W, Tang X and Deng X. 2005. Spatial and temporal patterns of China's cropland during

- 1990—2000: an analysis based on Landsat TM data. *Remote Sensing of Environment*, 98(4): 442-456
- Liu X, Hu G, Chen Y, Li X, Xu X, Li S, Pei F and Wang S. 2018. High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform. *Remote Sensing of Environment*, 209: 227-239
- Liu X, Huang Y, Xu X, Li X, Li X, Ciais P, Lin P, Gong K, Ziegler A D, Chen A, Gong P, Chen J, Hu G, Chen Y, Wang S, Wu Q, Huang K, Estes L and Zeng Z. 2020. High-spatiotemporal-resolution mapping of global urban change from 1985 to 2015. *Nature Sustainability*, 3(7): 564-570
- Loveland T R and Dwyer J L. 2012. Landsat: Building a strong future. *Remote Sensing of Environment*, 122: 22-29
- Loveland T R, Reed B C, Brown J F, Ohlen D O, Zhu Z, Yang L and Merchant J W. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, 21(6-7): 1303-1330
- Ma C, Rao Y, Cheng Y, Chen C, Lu J and Zhou J. 2020. Structure-Preserving Super Resolution with Gradient Guidance. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*
- Mao D, He X, Wang Z, Tian Y, Xiang H, Yu H, Man W, Jia M, Ren C and Zheng H. 2019. Diverse policies leading to contrasting impacts on land cover and ecosystem services in Northeast China. *Journal of Cleaner Production*, 240: 117961
- Masek J G, Vermote E F, Saleous N E, Wolfe R, Hall F G, Huemmrich K F, Feng G, Kutler J and Teng-Kui L. 2006. A Landsat surface reflectance dataset for North America, 1990—2000. *IEEE Geoscience and Remote Sensing Letters*, 3(1): 68-72
- Melgani F. 2006. Contextual reconstruction of cloud-contaminated multitemporal multispectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 44(2): 442-455
- Miller J D and Thode A E. 2007. Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). *Remote Sensing of Environment*, 109(1): 66-80
- Murray N J, Phinn S R, DeWitt M, Ferrari R, Johnston R, Lyons M B, Clinton N, Thau D and Fuller R A. 2019. The global distribution and trajectory of tidal flats. *Nature*, 565(7738): 222
- Pekel J F, Cottam A, Gorelick N and Belward A S. 2016. High-resolution mapping of global surface water and its long-term changes. *Nature*, 540(7633): 418
- Peng J, Hu X, Wang X, Meersmans J, Liu Y and Qiu S. 2019. Simulating the impact of Grain-for-Green Programme on ecosystem services trade-offs in Northwestern Yunnan, China. *Ecosystem Services*, 39: 100998
- Pengra B W, Stehman S V, Horton J A, Dockter D J, Schroeder T A, Yang Z, Cohen W B, Healey S P and Loveland T R. 2020. Quality control and assessment of interpreter consistency of annual land cover reference data in an operational national monitoring program. *Remote Sensing of Environment*, 238: 111261
- Pesaresi M, Huadong G, Blaes X, Ehrlich D, Ferri S, Gueguen L, Halkia M, Kauffmann M, Kemper T, Lu L, Marin-Herrera M A, Ouzounis G K, Scavazon M, Soille P, Syrris V and Zanchetta L. 2013. A Global Human Settlement Layer From Optical HR/VHR RS Data: Concept and First Results. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(5): 2102-2131
- Pickens A H, Hansen M C, Hancher M, Stehman S V, Tyukavina A, Potapov P, Marroquin B and Sherani Z. 2020. Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series. *Remote Sensing of Environment*, 243: 111792
- Qiu S, Lin Y, Shang R, Zhang J, Ma L and Zhu Z. 2019. Making Landsat Time Series Consistent: Evaluating and Improving Landsat Analysis Ready Data. *Remote Sensing*, 11(1)
- Qiu S, Zhu Z and He B. 2019. Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4—8 and Sentinel-2 imagery. *Remote Sensing of Environment*, 231: 111205
- Qiu T, Song C, Zhang Y, Liu H and Vose J M. 2020. Urbanization and climate change jointly shift land surface phenology in the northern mid-latitude large cities. *Remote Sensing of Environment*, 236: 111477
- Rakwatin P, Takeuchi W and Yasuoka Y. 2008. Restoration of Aqua MODIS band 6 using histogram matching and local least squares fitting. *IEEE Transactions on Geoscience and Remote Sensing*, 47(2): 613-627
- Ren S, He K, Girshick R and Sun J. 2016. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6): 1137-1149
- Rizvi S R, Killough B, Cherry A and Gowda S. 2018. Lessons learned and cost analysis of hosting a full stack Open data cube (ODC) application on the amazon web services (AWS). *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, IEEE*
- Roerink G, Menenti M and Verhoef W. 2000. Reconstructing cloudfree NDVI composites using Fourier analysis of time series. *International Journal of Remote Sensing*, 21(9): 1911-1917
- Roy D P, Ju J, Kline K, Scaramuzza P L, Kovalsky V, Hansen M, Loveland T R, Vermote E and Zhang C. 2010. Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment*, 114(1): 35-49
- Roy D P, Kovalsky V, Zhang H K, Vermote E F, Yan L, Kumar S S and Egorov A. 2016. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote Sensing of Environment*, 185: 57-70
- Schaaf C B, Gao F, Strahler A H, Lucht W, Li X, Tsang T, Strugnell N

- C, Zhang X, Jin Y, Muller J P, Lewis P, Barnsley M, Hobson P, Disney M, Roberts G, Dunderdale M, Doll C, d'Entremont R P, Hu B, Liang S, Privette J L and Roy D. 2002. First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment*, 83(1): 135-148
- Shen H, Li X, Cheng Q, Zeng C, Yang G, Li H and Zhang L. 2015. Missing Information Reconstruction of Remote Sensing Data: A Technical Review. *IEEE Geoscience and Remote Sensing Magazine*, 3(3): 61-85
- Shen H, Zeng C and Zhang L. 2010. Recovering reflectance of AQUA MODIS band 6 based on within-class local fitting. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 4(1): 185-192
- Shimada M, Itoh T, Motooka T, Watanabe M, Shiraishi T, Thapa R and Lucas R. 2014. New global forest/non-forest maps from ALOS PALSAR data (2007—2010). *Remote Sensing of environment*, 155: 13-31
- Simonyan K and Zisserman A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv: 1409.1556*
- Soenen S A, Peddle D R and Coburn C A. 2005. SCS+C: a modified Sun-canopy-sensor topographic correction in forested terrain. *IEEE Transactions on Geoscience and Remote Sensing*, 43(9): 2148-2159
- Song B, Li J, Mura M D, Li P, Plaza A, Bioucas-Dias J M, Benediktsson J A and Chanussot J. 2014. Remotely Sensed Image Classification Using Sparse Representations of Morphological Attribute Profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 52(8): 5122-5136
- Song X P, Hansen M C, Stehman S V, Potapov P V, Tyukavina A, Vermote E F and Townshend J R. 2018. Global land change from 1982 to 2016. *Nature*, 560(7720): 639-643
- Sulla-Menashe D, Gray J M, Abercrombie S P and Friedl M A. 2019. Hierarchical mapping of annual global land cover 2001 to present: The MODIS Collection 6 Land Cover product. *Remote sensing of environment*, 222: 183-194
- Sy S and Quesada B. 2020. Anthropogenic land cover change impact on climate extremes during the 21st century. *Environmental Research Letters*, 15(3): 034002
- Townshend J R and Justice C. 1986. Analysis of the dynamics of African vegetation using the normalized difference vegetation index. *International Journal of Remote Sensing*, 7(11): 1435-1445
- Turner B L, Lambin E F and Reenberg A. 2007. The emergence of land change science for global environmental change and sustainability. *Proceedings of the National Academy of Sciences*, 104(52): 20666-20671
- Verbesselt J, Hyndman R, Newnham G and Culvenor D. 2010. Detecting trend and seasonal changes in satellite image time series. *Remote sensing of Environment*, 114(1): 106-115
- Verbesselt J, Zeileis A and Herold M. 2012. Near real-time disturbance detection using satellite image time series. *Remote Sensing of Environment*, 123: 98-108
- Vermote E, Justice C, Claverie M and Franch B. 2016. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sensing of Environment*, 185: 46-56
- Vogelmann J E, Howard S M, Yang L, Larson C R, Wylie B K and Van Driel N. 2001. Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogrammetric Engineering and Remote Sensing*, 67(6)
- Wang H, Zhao X, Zhang X, Wu D and Du X. 2019. Long Time Series Land Cover Classification in China from 1982 to 2015 Based on Bi-LSTM Deep Learning. *Remote Sensing*, 11(14)
- Wang J, Zhao Y, Li C, Yu L, Liu D and Gong P. 2015. Mapping global land cover in 2001 and 2010 with spatial-temporal consistency at 250 m resolution. *ISPRS Journal of Photogrammetry and Remote Sensing*, 103: 38-47
- Wang L, Qu J J, Xiong X, Hao X, Xie Y and Che N. 2006. A new method for retrieving band 6 of Aqua MODIS. *IEEE Geoscience and Remote Sensing Letters*, 3(2): 267-270
- Wood E F, Roundy J K, Troy T J, Van Beek L P H, Bierkens M F P, Blyth E, de Roo A, Döll P, Ek M and Famiglietti J. 2011. Hyper-resolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resources Research*, 47(5)
- Woodcock C E, Allen R, Anderson M, Belward A, Bindschadler R, Cohen W, Gao F, Goward S N, Helder D and Helmer E. 2008. Free access to Landsat imagery. *Science*, 320(5879): 1011-1011
- Woodcock C E, Loveland T R, Herold M and Bauer M E. 2020. Transitioning from change detection to monitoring with remote sensing: A paradigm shift. *Remote Sensing of Environment*, 238: 111558
- Wulder M A, Coops N C, Roy D P, White J C and Hermosilla T. 2018. Land cover 2.0. *International Journal of Remote Sensing*, 39(12): 4254-4284
- Wulder M A, Loveland T R, Roy D P, Crawford C J, Masek J G, Woodcock C E, Allen R G, Anderson M C, Belward A S, Cohen W B, Dwyer J, Erb A, Gao F, Griffiths P, Helder D, Hermosilla T, Hipple J D, Hostert P, Hughes M J, Huntington J, Johnson D M, Kennedy R, Kilic A, Li Z, Lyburner L, McCorkel J, Pahlevan N, Scambos T A, Schaaf C, Schott J R, Sheng Y, Storey J, Vermote E, Vogelmann J, White J C, Wynne R H and Zhu Z. 2019. Current status of Landsat program, science, and applications. *Remote Sensing of Environment*, 225: 127-147
- Wulder M A, Masek J G, Cohen W B, Loveland T R and Woodcock C E. 2012. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, 122: 2-10
- Xian G, Homer C and Fry J. 2009. Updating the 2001 National Land

- Cover Database land cover classification to 2006 by using Landsat imagery change detection methods. *Remote Sensing of Environment*, 113(6): 1133-1147
- Xu H. 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International journal of remote sensing*, 27(14): 3025-3033
- Xu X, Liu J, Zhang S, Li R, Yan C and Wu S. 2018. China's Multi-Period Land Use Land Cover Remote Sensing Monitoring Data Set (CNLUCC). Resource and Environment Data Cloud Platform: Beijing, China
- Yan L and Roy D P. 2020. Spatially and temporally complete Landsat reflectance time series modelling: The fill-and-fit approach. *Remote Sensing of Environment*, 241: 111718
- Yang J, Gong P, Fu R, Zhang M, Chen J, Liang S, Xu B, Shi J and Dickinson R. 2013. The role of satellite remote sensing in climate change studies. *Nature Climate Change*, 3(10): 875-883
- Yang L, Jin S, Danielson P, Homer C, Gass L, Bender S M, Case A, Costello C, Dewitz J and Fry J. 2018. A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. *ISPRS journal of photogrammetry and remote sensing*, 146: 108-123
- Yu L, Liang L, Wang J, Zhao Y, Cheng Q, Hu L, Liu S, Yu L, Wang X, Zhu P, Li X, Xu Y, Li C, Fu W, Li X, Li W, Liu C, Cong N, Zhang H, Sun F, Bi X, Xin Q, Li D, Yan D, Zhu Z, Goodchild M F and Gong P. 2014. Meta-discoveries from a synthesis of satellite-based land-cover mapping research. *International Journal of Remote Sensing*, 35(13): 4573-4588
- Zeng C, Shen H and Zhang L. 2013. Recovering missing pixels for Landsat ETM+ SLC-off imagery using multi-temporal regression analysis and a regularization method. *Remote Sensing of Environment*, 131: 182-194
- Zha Y, Gao J and Ni S. 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International journal of remote sensing*, 24(3): 583-594
- Zhang C, Li W and Travis D. 2007. Gaps - fill of SLC - off Landsat ETM+ satellite image using a geostatistical approach. *International Journal of Remote Sensing*, 28(22): 5103-5122
- Zhang H K and Roy D P. 2017. Using the 500m MODIS land cover product to derive a consistent continental scale 30 m Landsat land cover classification. *Remote Sensing of Environment*, 197: 15-34
- Zhang J. 2010. Multi-source remote sensing data fusion: status and trends. *International Journal of Image and Data Fusion* 1(1): 5-24
- Zhang Z, Wang X, Zhao X, Liu B, Yi L, Zuo L, Wen Q, Liu F, Xu J and Hu S. 2014. A 2010 update of National Land Use/Cover Database of China at 1:100000 scale using medium spatial resolution satellite images. *Remote Sensing of Environment*, 149: 142-154
- Zhao Y, Gong P, Yu L, Hu L, Li X, Li C, Zhang H, Zheng Y, Wang J, Zhao Y, Cheng Q, Liu C, Liu S and Wang X. 2014. Towards a common validation sample set for global land-cover mapping. *International Journal of Remote Sensing*, 35(13): 4795-4814
- Zhu X, Chen J, Gao F, Chen X and Masek J G. 2010. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment*, 114(11): 2610-2623
- Zhu X, Gao F, Liu D and Chen J. 2012. A Modified Neighborhood Similar Pixel Interpolator Approach for Removing Thick Clouds in Landsat Images. *IEEE Geoscience and Remote Sensing Letters*, 9(3): 521-525
- Zhu Z and Woodcock C E. 2014. Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment*, 152: 217-234
- Zhu Z and Woodcock C E. 2014. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, 144: 152-171
- Zhu Z, Zhang J, Yang Z, Aljaddani A H, Cohen W B, Qiu S and Zhou C. 2020. Continuous monitoring of land disturbance based on Landsat time series. *Remote Sensing of Environment*, 238: 111116
- Fu W, Chen Y L, Shi M W, Zhang X D, Xiao H and Gong P. 2014. The spatio temporal changes of surface cover color in China revealed by satellite observations. *Journal of Remote Sensing*, 18(1): 154-179 (付薇, 陈焱磊, 施楣梧, 张旭东, 肖红和宫鹏. 2014. 卫星动态观测数据揭示中国地表色调分布及时序变化特征. *遥感学报*, 18(1): 154-179) [DOI:10.11834/jrs.20132323]
- Gong P, Zhang W, Yu L, Li C C, Wang J, Liang L, Li X C, Ji L Y and Bai Y Q. 2016. New research paradigm for global land cover mapping. *Journal of Remote Sensing*, 20(5): 1002-1016 (宫鹏, 张伟, 俞乐, 李丛丛, 王杰, 梁璐, 李雪草, 计璐艳和白玉琪. 2016. 全球地表覆盖制图研究新范式. *遥感学报*, 20(5): 1002-1016) [DOI: 10.11834/jrs.20166138]
- Liu F, Yan H M, Liu J Y, Xiao X M, Qin Y W. 2016. Spatial pattern of land use intensity in China in 2000. *Acta geographica sciences*, 2016, 71(7): 1130-1143 (刘芳, 闫慧敏, 刘纪远, 肖向明和秦元伟. 2016. 21世纪初中国土地利用强度的空间分布格局. *地理学报*, 71(7): 1130-1143.) [DOI:10.11821/dlxb201607004]
- Liu H, He L and Li J. 2017. Advances in Deep learning and its applications in image processing. *ZTE Technology Journal*, 23(04): 1-6. (刘涵, 贺霖 和 李军. 2017. 深度学习进展及其在图像处理领域的应用. *中兴通讯技术*, 23(04): 1-6)
- Liu J Y, Kuang W H, Zhang Z X, Xu X L, Qin Y W, Ning J, Zhou W C, Zhang S W, Li R D, Yan C Z, Wu S X, Shi X Z, Jiang N, Yu D S, Pan X Z, Chi W F. Spatiotemporal characteristics, patterns and causes of land use changes in China since the late 1980. 2014. *Acta Geographica Sinica*, 69(1): 3-14 (刘纪远, 匡文慧, 张增祥, 徐新良, 秦元伟, 宁佳, 周万村, 张树文, 李仁东, 颜长珍, 吴世新, 史学正, 江南, 于东升, 潘贤章和迟文峰. 2014. “20世纪80年代末以来中国土地利用变化的基本特征与空间格局.” *地理学报* 69(1): 3-14)

21st century daily seamless data cube reconstruction and seasonal to annual land cover and land use dynamics mapping—iMap (China) 1.0

LIU Han¹, GONG Peng^{1,2,3}

1. Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China;

2. Tsinghua Urban Institute, Tsinghua University, Beijing 100084, China;

3. East Asian Migratory Birds and Habitat Ecology Field Observation and Research Station, Beijing 100084, China

Abstract: Sustainable development goals such as food security, high-quality habitat construction, biodiversity conservation, planetary health, and the understanding, modeling, and management of the Earth system urgently require multi-scale, long time series, high-accuracy, and consistent remote sensing observation datasets and mapping products with flexible classification systems to meet user needs. However, due to technical and cost constraints, it is difficult for conventional remote sensing satellites to provide observations with high spatial resolution, high temporal frequency, and high quality at the same time. The existing mapping and inversion schemes are mostly for a single sensor, making it difficult to fully exploit and jointly utilize the information potential of multi-source heterogeneous remote sensing big data, resulting in limited observation periods and resolutions, low spatial and temporal consistency and comparability. Therefore, new technical paradigms are urgently needed in the field of remote sensing. In this paper, based on advanced technologies, including cloud computing, artificial intelligence, virtual constellation, spatio-temporal fusion reconstruction, an intelligent mapping framework is proposed for remote sensing big data. The framework is user-driven and problem-driven, which can significantly improve the current situation that remote sensing data products can hardly meet users' diversified and high-precision surface monitoring needs in agriculture and forestry management, national monitoring, ecological environment protection, disaster prevention and mitigation, urban planning, etc. Under this framework's guidance, we built an online real-time, automated, serverless, end-to-end remote sensing big data production chain and parallel mapping system based on Amazon Web Services (AWS) high-performance, elastic, and scalable distributed computing resources. We produced the first set of 21st century daily Seamless Data Cube (SDC) and seasonal to annual land cover and land use mapping products of China. Integrating Landsat and MODIS satellite as a virtual constellation, through multi-source spatio-temporal data fusion and reconstruction technology, the daily SDC, cloud-free, high-precision reflectance product, is developed. As Analysis Ready Data (ARD), it lays the foundation for high-precision quantitative remote sensing inversion and mapping. Based on SDC, we developed the seasonal to annual mapping product with multiple multi-level land cover and land use classification systems, whose mean annual accuracy exceeds 80%. The main mapping pipeline includes migrating the all-season sample set based on stable classification theory with limited samples, optimizing and ensembling multiple classifiers by Automatic Machine Learning (AutoML) strategies, and using change detection and post-processing techniques to achieve consistency. The two sets of products demonstrate the feasibility and effectiveness of the intelligent remote sensing mapping framework proposed in this paper. We will continue to improve and develop the framework with an open and flexible concept to provide new ideas to promote remote sensing development in China.

Key words: seamless data cube, daily, seasonal mapping, cloud computing, intelligent remote sensing mapping, long time series, change monitoring, artificial intelligence, spatio-temporal big data