中国科学: 地球科学 2020 年 第 50 卷 第 7 期: 1009 ~ 1020
SCIENTIA SINICA Terrae earthcn.scichina.com
《中国科学》杂志社 SCIENCE CHINA PRESS
论 文

全球历史LUCC数据集数据可靠性的评估方法及评估案例

方修琦[1,2*], 赵琬一[1,2], 张成鹏[1,2], 张頔旸[1,2], 魏学琼[3], 邱维理[1,2], 叶瑜[1,2]

1. 北京师范大学环境演变与自然灾害教育部重点实验室, 北京 100875;
2. 北京师范大学地理科学学部地理学院, 北京 100875;
3. 南京信息工程大学地理科学学院, 南京 210044
* 通讯作者, E-mail: xfang@bnu.edu.cn

收稿日期: 2019-05-04; 收修改稿日期: 2019-10-18; 接受日期: 2019-11-11; 网络版发表日期: 2020-04-01
国家重点研发计划"全球变化与应对"专项项目(编号: 2017YFA0603304)和国家自然科学基金项目(批准号: 41807433)资助

摘要 土地利用/土地覆被变化(LUCC)既是全球变化的表现形式, 也是全球变化的重要驱动力, 准确的全球历史LUCC数据产品是正确认识LUCC对全球变化影响的基础. 但国际上已有的多套全球历史土地利用/覆盖数据集存在彼此矛盾之处, 且均已被许多区域尺度重建结果证明存在很大偏差, 改进全球历史土地覆盖数据产品精度是国际LUCC和过去全球变化研究关注的重点领域. 从时间过程和空间分布两个方面评估已有土地覆盖数据产品的可靠性, 是改进数据集质量的基础和前提. 理论上, 判断全球土地覆盖数据产品可靠性的高低的标准是其与实际土地覆盖数据("真值")的相近程度或偏差大小, 但与现代全球土地覆盖数据产品可提供实地抽样验证不同, 可用于历史产品可靠性评估的历史证据是极其有限的. 有鉴于此, 本文构建了全球LUCC数据集数据可靠性评估框架, 并针对可用于历史土地覆盖数据评估的数据和指标的类型和特点, 提出4种评估方法, 包括: (1) 以定量重建的区域LUCC数据为基准的准确性评估; (2) 基于区域发展史实的合理性评估; (3) 基于自然和人文特征及专家知识的合理性评估; (4) 基于多套数据一致性的可能性评估, 并给出了5个历史耕地可靠性的具体评估案例. 该方法亦适用于林地、草地等其他土地类型的评估.

关键词 LUCC, 全球数据集, 可靠性, 评估方法

1 引言

人类的各种土地利用(land use, LU)在满足发展需求的同时, 也改变了陆地表层的土地覆盖(land cover, LC), 进而通过影响陆地生态系统的生物多样性、改变陆地表面物理过程(地表反照率、辐射强迫、蒸散发、土壤侵蚀)、改变全球碳、氮等生物地球化学循环和水循环, 深刻地影响全球的、区域的气候与环境变化(Ruddiman, 2003; Foley等, 2005; Goldewijk等, 2007; Pongratz等, 2008; Pielke等, 2011; Zhang等, 2012; Ellis等, 2013; 李蓓蓓等, 2014; Gaillard等, 2018). 因此, 土地利用/土地覆盖变化(LUCC或LULC)

中文引用格式: 方修琦, 赵琬一, 张成鹏, 张頔旸, 魏学琼, 邱维理, 叶瑜. 2020. 全球历史LUCC数据集数据可靠性的评估方法及评估案例. 中国科学: 地球科学, 50: 1009–1020, doi: 10.1360/SSTe-2019-0093
英文引用格式: Fang X, Zhao W, Zhang C, Zhang D, Wei X, Qiu W, Ye Y. 2020. Methodology for credibility assessment of historical global LUCC datasets. Science China Earth Sciences, 63: 1013–1025, https://doi.org/10.1007/s11430-019-9555-3

© 2019 《中国科学》杂志社

不仅是全球变化的表现形式, 而且是全球变化的重要驱动力.

在深入了解过去土地利用/覆盖时空演变过程的基础上, 建立面向全球变化研究的全球历史LUCC数据产品, 是定量评估LUCC对全球变化影响的基础. 自20世纪90年代LUCC国际研究计划实施以来, 在全球和区域历史土地利用/土地覆盖重建方面取得了诸多成果, 国际上相继建立了多套高时空分辨率的全球历史土地利用/覆盖数据集. 其中具有代表性的是荷兰环境评价局(Netherlands Environmental Assessment Agency)建立的HYDE数据集(已更新到HYDE3.2版本, 空间分辨率为5′×5′, 重建时段为10000BC~AD2015)(Goldewijk和Battjes, 1997; Goldewijk, 2001; Goldewijk等, 2011, 2017), 美国威斯康星大学全球环境和可持续发展中心(Center for Sustainability and the Global Environment)的Ramankutty和Foley(1999)建立的“全球土地利用数据集”(简称SAGE数据集, 重建时段为AD1700~1992, 分辨率为0.5°×0.5°), Pongratz等(2008)建立的PJ数据集, Kaplan等(2009, 2010)建立的KK10数据集等. 这些数据集一般是以有限的区域历史耕地和人口记录为基础, 综合考量现代土地覆盖、历史人口和气候/土地适宜性等因素进行反演而得, 已被广泛地应用于历史碳排放估算、土地利用/覆盖变化对区域及全球气候变化影响的模拟等全球变化研究当中, 成为定量研究土地覆盖变化对全球变化影响的基本数据集(Ruddiman, 2003; Goldewijk等, 2007; Pongratz等, 2008; Pielke等, 2011; Zhang等, 2012; Ellis等, 2013; Gaillard等, 2018), 对全面认识人类活动对全球变化的贡献做出了重要贡献. 但由于不同数据集用于重建的原始数据源和格网化分配方法的不同, 以及重建方法和空间分配模型的某些局限, 上述全球数据集之间存在彼此矛盾之处, 且均已被众多区域尺度重建结果证明存在很大偏差(Kaplan等, 2009; 李蓓蓓等, 2010; Li等, 2013; Leite等, 2012; He等, 2013; Zumkehr和Campbell, 2013; Yuan等, 2017; 杨绪红等, 2019). 如何改进全球历史土地覆盖数据产品精度, 一直是国际LUCC和过去全球变化研究关注的热点, 也是国际上正在执行的LandCover6k计划的研究目标(Gaillard等, 2015, 2018).

对已有土地覆盖数据产品进行可靠性评估是改进数据集质量的前提, 通过可靠性评估, 可区分哪些区域或时段的数据是相对可靠的, 哪些是通过订正可改进的, 哪些是需要用新数据另行重建的. 现代全球土地覆盖数据产品的可靠性是可以通过实地抽样进行检验的, 并已发展了多种检验和改进可靠性的方法(陈军和陈晋, 2018). 然而, 评价全球土地覆盖数据产品的可靠性的难点在于, 作为评价历史土地覆盖数据产品可靠性参照的土地覆盖真实状况(简称“真值”)就是重建的目标值, 它们已成为过去, 在绝大多数情况下不但无法直接获得, 而且可用于对其重建的历史文献和自然记录亦十分有限, 难以像现代全球土地覆盖数据产品那样通过实地抽样验证, 亦无公认的方法. 已有对全球LUCC数据集数据的可靠性评估, 绝大多数是将基于历史记录的区域尺度重建结果近似地视为历史“真值”, 根据全球数据集的数据相对于区域重建数据偏差的大小判断其可靠性(李蓓蓓等, 2010; 何凡能等, 2012, 2019; He等, 2013; Yuan等, 2017; Wei等, 2019; Li等, 2019), 但区域土地利用定量重建工作量巨大, 且可用作重建证据的资料难以获得, 已有的区域重建结果的空间和时间范围均十分有限, 对全球历史土地覆盖数据产品的可靠性评估在更多情况下需要在没有区域重建结果的情况下开展. 因此, 在缺少可利用的历史数据的情况下, 如何评估历史土地覆盖数据的可靠性是一个亟待解决的难题.

本文尝试构建全球LUCC数据集数据可靠性评估的基本框架, 针对可用于历史土地覆盖数据评估的数据和指标的类型和特点, 提出不同的评估方法, 并以耕地可靠性为例给出具体案例. 这些评估方法亦适用于林地、草地等其他土地类型的评估(何凡能等, 2019; 杨帆等, 2019).

2 全球历史LUCC数据集数据可靠性评估的基本框架

过去全球土地利用/覆盖数据的可靠性评估包括时间过程评估和空间分布评估两个方面. 时间过程的可靠性评估针对给定的空间单元, 评价数据的时间变化过程及在不同时间断面上的可靠性. 空间分布的可靠性评估针对给定的时间断面, 评价数据在不同尺度空间单元上的可靠性(图1).

全球土地覆盖数据产品的可靠性可从多个角度进行刻画. 理论上, 判断全球土地覆盖数据产品可靠性高

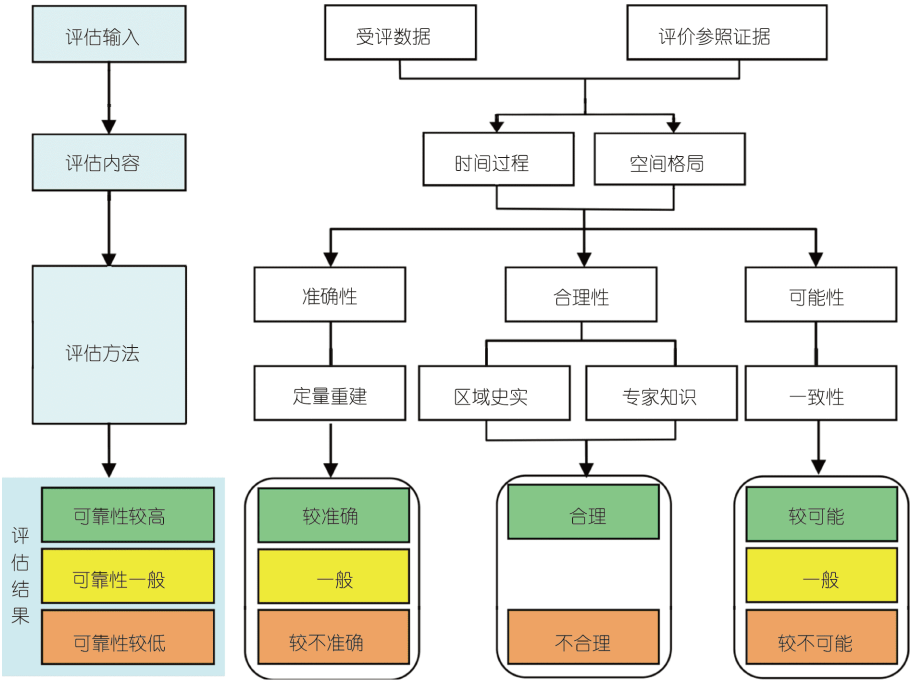

图 1　历史土地利用/覆盖数据可靠性评估框架

低的标准是其相对于实际土地覆盖数据(“真值”)的接近程度或偏差大小, 偏差越小可靠性越高, 偏差越大可靠性越低. 然而, 作为评价全球土地覆盖数据产品可靠性参照值的“真值”已成为过去, 它们需要通过重建才能获得, 而可用于对其重建的历史文献和自然记录亦十分有限. 鉴于待评估的全球土地覆盖产品本身就是一种重建结果, 对其进行可靠性评估所用的证据需要独立于这些全球土地覆盖数据产品、且更接近真值或能够充分体现真值的某些特点. 其中, 用已知更接近真值的定量区域重建数据作为目标值, 可以从准确性的角度对LUCC数据产品的可靠性进行定量评价. 如果存在多套重建值, 但无法确定哪一套更接近真值, 则可根据各套数据之间的一致性程度对评价目标可靠的可能性进行判断. 利用能够体现真值特点的代用指标, 只能从合理性的角度对LUCC数据产品的可靠性做出定性的评价. 针对被评价数据主体及评价参照指标的不同, 具体评估方法分为4类(图1).

全球LUCC数据集数据可靠性评估受数据可得性的限制, 在实际评价中, 具体选择什么评价目标和方法, 依赖于可用作评价目标参照值的证据的特点. 亦可针对具体的评价区域和时段, 将不同评价目标与方法结合起来使用.

3　以区域LUCC定量重建数据为基准的准确性评估

3.1　评估方法

准确性评估是一种定量评估, 它以基于历史文献或自然记录定量重建的区域耕地、林地和草地数据为评估基准, 假设这些重建结果最接近相应类别土地覆盖数据的历史真值, 根据研究区域内全球LUCC数据集中的土地覆盖数据相对于历史重建值偏差的大小, 定量评估数据的准确性. 已有的对全球LUCC数据集的可靠性评估大多是此类型(李蓓蓓等, 2010; Li等, 2013; 何凡能等, 2012, 2019; He等, 2013; Wei等, 2016, 2019; Yuan等, 2017).

区域历史土地覆盖数据可以通过挖掘相关历史文献或自然记录进行重建, 亦可通过收集已发表重建成果获得. 在将多时相、多空间尺度的区域重建数据与全球LUCC数据集两套数据统一到相同的时间断面或空间单元后, 计算给定区域内全球LUCC数据集中的土地覆盖数据相对于历史重建值的偏差, 以全球LUCC数据集的数据相对于区域重建数据绝对或相对偏差的大小作为准确性的判别指标, 偏差越小, 数据的可靠性越高. 根据给定的绝对或相对偏差标准, 可把全

球尺度LUCC数据集中数据的准确性区分为较准确、一般或较不准确等不同类型, 但从已有案例研究看, 对如何确定可靠性的判别标准或阈值尚无一般的标准.

准确性评估的优点是定量, 可定量评估全球LUCC数据集数据的精度, 该评估方法要求研究者在开展评估工作前必须重建出或获取到待评估区域的土地覆盖数据重建值. 但区域土地利用定量重建工作量巨大, 且可用于重建的历史文献和自然记录十分有限, 只能在资料条件或工作基础较好的有限区域进行, 在如今全球多数地区尚未开展区域LUCC数据重建工作的情况下, 运用此方法有一定的局限性.

3.2 案例1: HYDE数据集中过去千年德国耕地总量的准确性

以Bork等(1998)基于区域史实重建的德国AD600~2000垦殖率为参照标准, 从HYDE3.1(Goldewijk等, 2011)、HYDE3.2(Goldewijk等, 2017)数据集中提取德国过去千年垦殖率, 并计算其与Bork垦殖率的相对偏差, 据此从时间过程的角度评估HYDE3.1、HYDE3.2数据集德国过去千年垦殖率数据的准确性.

过去千年HYDE3.1数据集中德国的垦殖率变化在总量和变化趋势上与Bork垦殖率曲线明显不符. HYDE3.1的垦殖率在各个时间断面上均较Bork垦殖率显著偏低, 与Bork垦殖率曲线相比, 1940年以前低20%以上, 1830年以前低50%以上, 1300年以前低75%以上. Bork垦殖率曲线中过去千年德国垦殖率的显著波动在HYDE3.1中几乎没有体现. 因此, HYDE3.1数据集的德国过去千年垦殖率是较不准确的(图2).

HYDE3.2数据集吸收了Bork等(1998)等德国区域土地重建成果, 对历史耕地总量进行了修正, 其变化趋势与Bork垦殖率曲线相近, 在数量上与Bork等(1998)历史重建值的相对偏差缩小到–13~15%, 总体上较为准确. AD1700以前, HYDE3.2的垦殖率与历史重建值偏差较小, 但其时间分辨率为100年, 导致其峰值和谷值的位相与Bork曲线略有偏差, 特别是难以体现三十年战争(AD1618~1648)等持续时间小于百年的历史事件造成的垦殖率波动. AD1700~1940, HYDE3.2的时间分辨率提高至10年, 但其垦殖率以50年为间隔呈阶梯状变化, 在每个50年内取相同数值, 使得其相对于Bork垦殖率的偏差以100年为周期波动于–12~9%, 多数时点上数据的准确性甚至低于1700年前(图2).

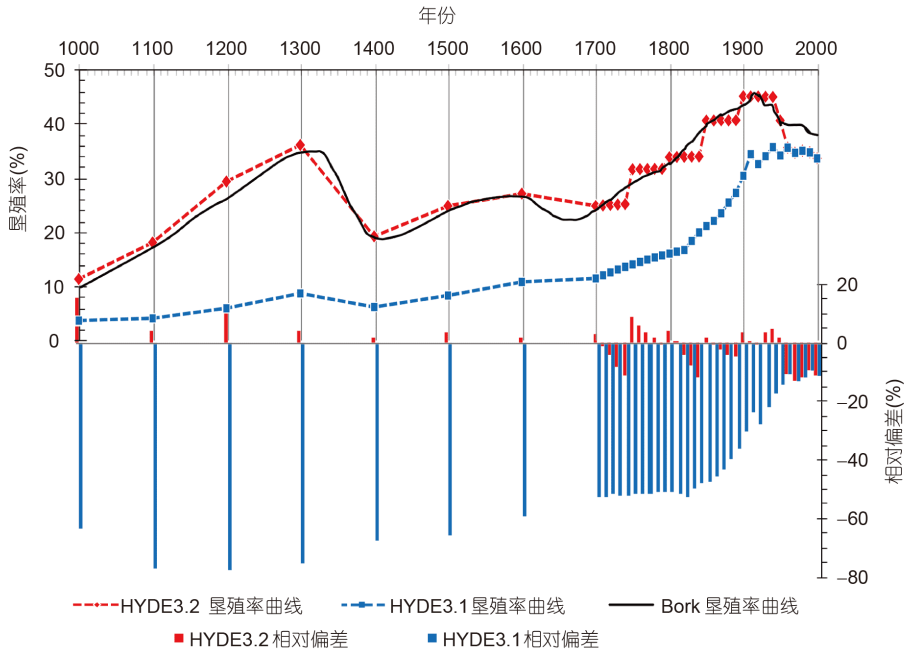

图 2 过去千年德国HYDE数据集中的垦殖率与历史重建值的对比

数据集据Goldewijk等(2011, 2017), 垦殖率与历史重建值的数据据Bork等(1998)

3.3　案例2: HYDE3.2过去300年华北地区耕地空间分布的准确性

Wei等(2019)集成重建的过去300年华北地区历史耕地数据集包括17世纪后半叶、18世纪、19世纪、1916年、1933年、20世纪80年代和2008年7个时间断面上空间分辨率为10km×10km的耕地数据，其所采用的原始数据来源于清代和民国时期地方志、民国以来的统计资料和现代耕地调查资料，历史记录的空间分辨率到县级单元. 对应于华北地区历史耕地数据集的17世纪后半叶、18世纪、19世纪、1916年、1933年、20世纪80年代和2008年7个时间断面, 从HYDE3.2数据集(Goldewijk等, 2017)中选择相近的时间断面, 提取华北地区1700年、1750年、1850年、1910年、1930年、1980年和2008年的耕地数据，计算其与Wei等(2019)华北地区历史耕地数据集对应时间断面的相对偏差，据此评估HYDE3.2数据集10km×10km网格尺度上空间分布的准确性(图3).

HYDE数据集从HYDE3.1版本(Goldewijk等, 2011)开始吸收了葛全胜等(2008)过去300年中国耕地重建成果，HYDE3.2数据集中华北地区的耕地数据总量与Wei等(2019)的重建结果较为相近，7个时间断面中有6个相对差异在−10~13%，只有1980年的相对差异为−16.60%(图3)，对历史耕地重建而言，可以认为总体上较为准确.

从垦殖率的空间分布来看，尽管HYDE3.2数据集和Wei等(2019)两个数据集均显示，过去300年华北地区耕地主要分布在平原区，而周边山区为低垦殖区，但垦殖率相对差异的空间分布显示存在较大差异. 在1700~1930年5个时间断面上, HYDE3.2与Wei等(2019)区域重建结果对比，相对差异的绝对值>50%网格约占网格总量的49~57%，而相对差异在−10~10%的格网仅约占8~12%. 垦殖率相对差异的空间分布显示, HYDE3.2数据集在华北地区存在北部总体偏低、而南部总体偏高系统性偏差，相对差异>50%的格网中有76~86%分布在黄河以南地区, 而相对差异<−50%的格网中有63~67%分布在黄河以北地区. 且存在HYDE3.2数据集的垦殖率在长江、黄河和淮河等大河沿岸明显偏高的现象. 在1980年和2008年2个现代时间断面上, 两个数据集耕地空间分布的差异较历史时期小，相对差异在−10~10%的格网所占比例提高为13~21%, 相对差异的绝对值>50%格网所占比例降低为29~35%. 但在空间分布上，HYDE3.2数据集在平原地区垦殖率整体偏小，而周边山地和高原区偏大，平原区有31~66%的格网相对差异<−10%，而周边山地和高原区有48~70%的格网相对差异>10%.

在造成上述空间差异的诸多原因中，可以肯定的主要原因有Wei等(2019)是对县域单元耕地数据进行空间分配，能够更有效地降低空间分配的误差. HYDE3.2数据集假设河流附近的土地优先被开垦而先将耕地分配至河流附近的格网，与华北地区悠久的土地开发明显不符合，而Wei等(2019)的空间分配模型中未使用河流因子.

4　基于区域发展史实和科学认知的合理性评估

4.1　评估方法

合理性评估是一种定性评估，可靠的LUCC数据首先必需是合理的. 对合理性可以从两个方面进行判断: 一是以区域历史事实为证据判断是否有相关的历史证据支持; 二是根据相关的科学认知推断是否具备相应的自然和社会条件. 因此，全球数据集的数据合理性评估一般采取两种方法，即根据区域土地开发历史，通过比对全球LUCC数据集的数据与区域史实在时空分布上的一致性判断其合理性; 或根据现代农业特点、自然地理环境时空特征或相关自然与人文专家知识，从全球LUCC数据集所用假设及方法与相关自然和人文专家知识的符合程度判别其合理性. 合理性评估的结果可分为合理和不合理两种类型.

4.1.1　基于区域发展史实的合理性评估

该方法采用区域农业发展史实作为合理性评估的参照标准. 这些直接或间接刻画区域发展史实的记录有多种，例如，历史文献中关于区域开垦与农业发展过程的直接文字描述和有关土地政策、技术进步、人口数量、重大社会事件等间接记录，能够反映农业开发历史的考古或自然沉积记录等.

通过系统收集整理与农业相关的历史和自然记录，从中提取农业开发历史上的重大事件、关键转折点等能够反映历史变化过程的信息，或能够反映某个时间断面农业开发空间格局的信息. 从全球土地覆盖

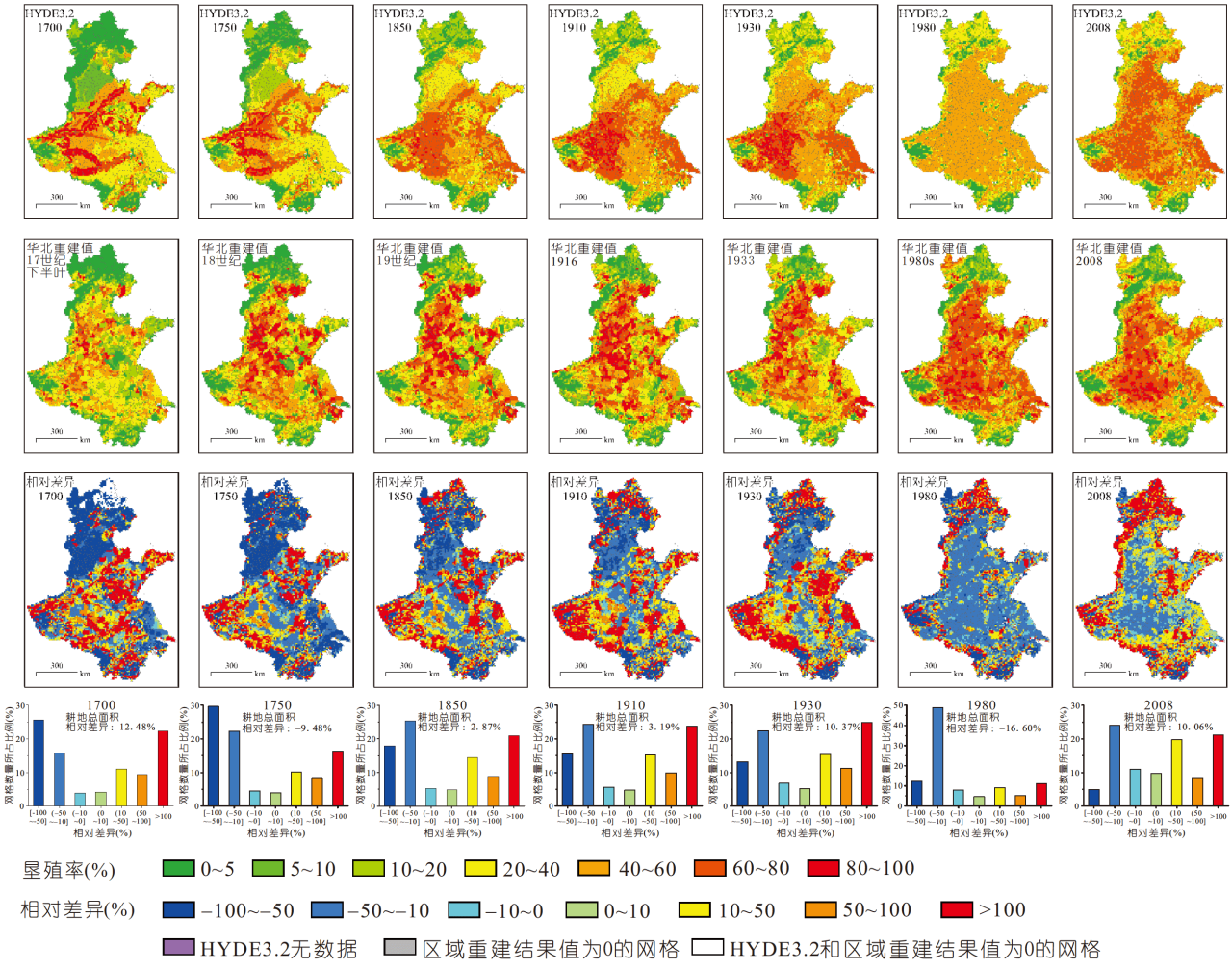

图 3　过去300年华北地区HYDE3.2和基于历史记录重建的格网化耕地及其相对差异分布

据Goldewijk等(2017)和Wei等(2019)

数据集中提取与上述记录耕地变化序列或空间分布信息，将全球LUCC数据集耕地数据的时空分布与上述史实揭示的耕地时空变化特征进行对比，判断其合理性.

4.1.2　基于自然和人文特征及专家知识的合理性评估

基于自然和人文特征及专家知识的合理性评估是一种间接评估方法. 在缺乏定量或定性历史土地覆盖记录的情况下，可用此方法间接地对全球土地覆盖数据集中某些区域或时段的数据的合理性进行评估. 该评估方法将作为区域土地利用/覆盖形成条件的自然要素和社会要素的时空特征及其与土地利用/覆盖关系作为已知条件，分析全球LUCC数据集在重建证据或代用指标的使用、数据转换或空间分配的方法和假设前提等方面是否符合已有对相关原理或事实的认知，全球土地覆盖数据集数据产品的时空分布是否与这些自然和人文现象的时空分配相匹配，进而评估其合理性. 例如:

(1) 对某种土地覆盖类型而言, 在给定的区域或时段内从自然地理分异特征与地理环境演变过程看是否具备该类型存在的自然条件.

(2) 对历史土地覆盖重建或空间分配所使用的方法和假设条件而言，这些方法或假设是否符合相关的原理，是否能够得到现代观测记录或历史记录的支持，是否具有时间和空间上的普适性.

(3) 对全球LUCC数据集中的数据而言, 这些数据的时空分布与影响土地利用与土地覆盖的温度、降水、地形、河流水系等自然地理要素的时空分布特征是否相互匹配, 与社会经济和技术发展的阶段性是否矛盾. 例如, HYDE3.2中德国AD1700~1950的垦殖率表现出每50年一次的阶梯式变化(图1), 从区域发展的一般特点看是不合理的, HYDE3.2数据集中AD1700~1930的耕地分布在华北平原的北部和南部存在明显的差异(图2), 与华北平原相对均一的自然环境特征也是不符的.

4.2　案例3: HYDE3.2数据集东北地区过去千年的耕地分布的合理性

本案例基于辽代以来各朝代古聚落遗址和清代城镇数量的时空变化, 对HYDE3.2数据集(Goldewijk等, 2017)中东北地区过去千年耕地数据的合理性进行评估.

东北地区与定居种植农业密切相关的历朝古聚落遗址的数量及其空间分布能够反映农业发展和农耕区的变化(Jia等, 2018). 过去千年中国东北地区(37°~55° N, 120°~136°E)的古聚落遗址变化显示(Jia等, 2018), 辽(AD916~1125)和金(AD1115~1234)两代聚落数量众多且分布北界达到47°N, 农耕区北界与现代相近, 而元(AD1279~1368)和明(AD1368~1644)时期的古城数量显著减少, 分布北界南移了3~4个纬度, 显示耕地数量和分布范围的显著缩小(图4).

东北地区绝大多数的现代聚落是清代以来发展起来的(曾早早等, 2011), 其中包括为管理从事农业开发汉民而建的州、道、府、县等各级行政单元治所的城市. 由于现代聚落对清代聚落存在继承性, 清代聚落遗

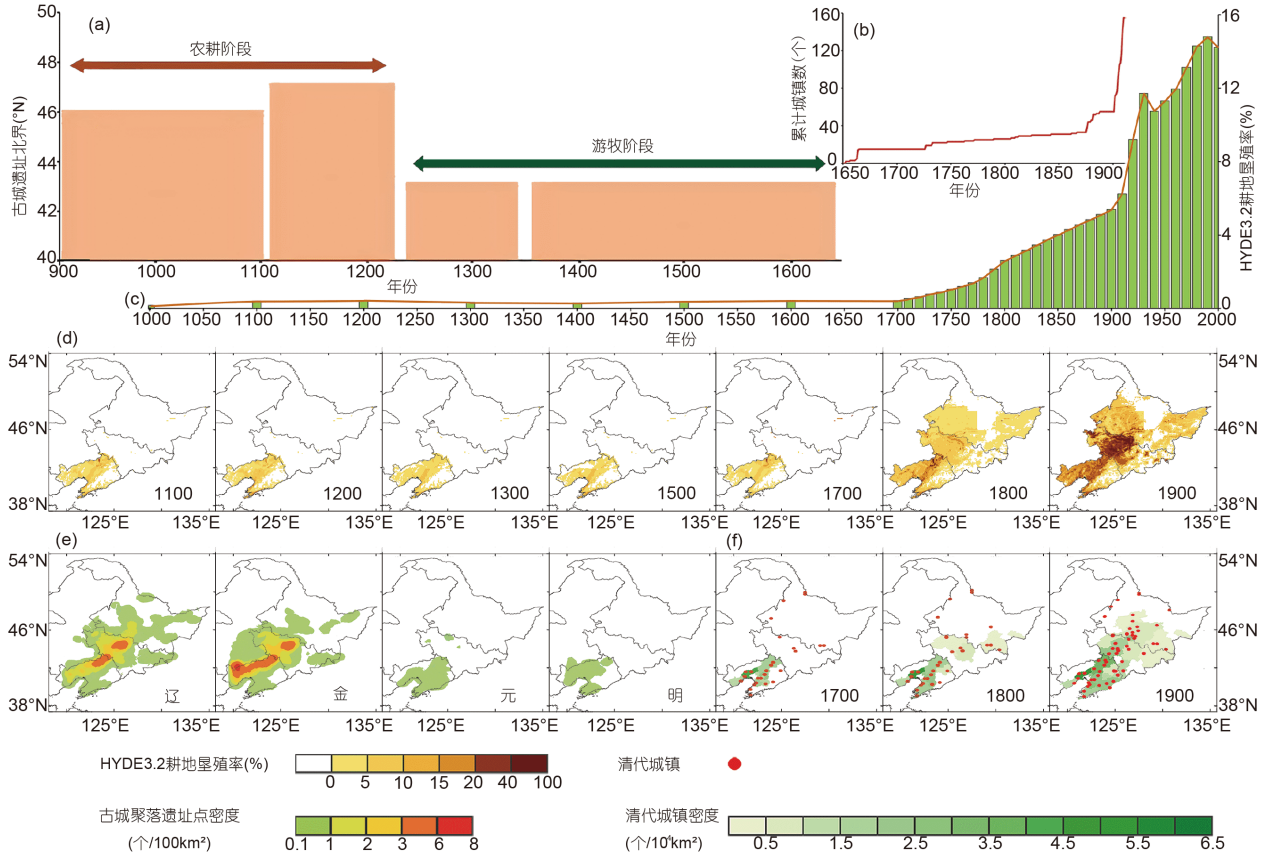

图 4　HYDE3.2数据集中过去千年东北地区垦殖率的时空变化与古聚落遗址和县城设置时空变化的比较

(a) 辽至明代东北地区古遗址北界(Jia等, 2018); (b) 清代东北三省城镇累计个数(方修琦等, 2005); (c) HYDE3.2过去千年东北三省垦殖率(Goldewijk等, 2017); (d) HYDE3.2过去千年东北三省垦殖率空间分布(Goldewijk等, 2017); (e) 辽至明代东北地区古聚落遗址空间分布(Jia等, 2018); (f) 清代东北城镇体系空间演变(方修琦等, 2005)

址数量和空间分布不能够反映农耕区的实际情况，但清代东北地区州、道、府、县等各级行政单元治所建立所反映的城镇体系演变能够指示东北地区农耕区的时空变化过程(方修琦等, 2005). 清代东北三省(辽宁、吉林、黑龙江)的城镇体系演变显示, 城镇设置在1860年以前发展缓慢，1700年以前的城镇仅分布在辽宁省境内，1860年时吉林省和黑龙江省南部仅有少量分布，1860~1895年呈较快速增长，1895~1911年呈快速增长，清代东北地区所设州县的3/4以上是在清末约50年中建立的，其中，吉林省城镇总数的75.5%、黑龙江城镇总数的87.8%建立于1901~1911年(图4)(方修琦等, 2005).

过去千年HYDE3.2东北地区耕地的时空变化与上述变化对比存在明显的差别. 从时间变化过程看，HYDE3.2中东北地区的耕地在过去千年是持续增加的，没有反映出辽金时期的开垦峰值及从辽金到元明之间耕地数量的急剧减少. 清代垦殖率呈近线性增长，没有反映1860年前后开垦速度变化的显著差异. 在空间分布上，HYDE3.2中辽、金时期的耕地仅在辽宁省境内，与同期吉林省和黑龙江省境内存在大范围农耕区的史实明显不符. 而HYDE3.2数据集在AD1800和AD1900的耕地分布北界亦较城镇设置指示的农耕区范围明显偏北(图4).

以上对比结果说明，HYDE3.2数据集中东北地区过去千年的耕地数量变化和空间分布均不合理. 究其原因，HYDE3.2数据集对东北地区耕地数量及其空间分布从现代到历史的回溯没有考虑该区域独特的发展历史过程. 对过去千年而言, HYDE3.2数据集既没有考虑辽金时期汉人移民及少数民族学习汉人耕作技术带来的农耕发展，也没有考虑蒙元游牧民族统治下带来的从农业到牧业生产模式的转变(Jia等, 2018). 在清代，HYDE3.2数据集未考虑1860年东北地区封禁政策解禁后东北地区移民开垦的快速增长及移民开垦区域从南到北逐步推进的过程(方修琦等, 2005; 叶瑜等, 2009).

4.3 案例4: HYDE3.2环渤海地区过去7000年耕地数据的合理性

HYDE3.2数据集在进行历史耕地的格网化分配时, 在现代耕地的空间范围内, 根据各空间单元的垦殖适宜性由大到小的次序，对历史上的耕地数量进行分配(Goldewijk等, 2011, 2017). 上述格网化分配方法是以古今自然环境保持不变为前提的，一旦自然环境发生改变，就可能影响耕地分配模型及其分配结果的合理性. 本案例以渤海海岸线及黄河水系变化史实为依据，通过分析用于耕地空间分配的前提条件或影响因子是否成立，评估HYDE3.2在环渤海地区过去7000年耕地数据的合理性.

受海平面变化与三角洲堆积的共同作用，历史上的沿海平原随海岸线的进退而生消变化，使得古今沿海地带可耕土地的空间范围并不相同. 过去7000年来渤海西部沿海平原随着海岸线的阶段性退缩而不断增长(薛春汀，2009；商志文等，2018)，累计增长约$3.12\times10^4 km^2$. 陆地是耕地存在的最基本自然条件，在这些过去7000年中新增的陆地范围内，即使现代适宜分布耕地，至少在其成为陆地之前的历史时期尚不具备开垦的可能性. 从HYDE3.2数据集中提取渤海西部沿海平原5000BC和AD1200耕地数据，将其与相应时期的海岸线进行匹配(图5)，可以看出HYDE3.2数据没有考虑该区域海岸线的迁移过程，两个时期的耕地数据均被分配到历史时期尚未成陆的海域内，因此是不合理的.

从图5中可以看出，HYDE3.2数据在7000a BP和AD1200的垦殖率分布均存在现代黄河沿线垦殖率显著偏高的现象，这显然与HYDE3.2数据集认为河流附近的土地优先被开垦而将耕地优先分配至河流附近格网有关. 但不论HYDE3.2数据集沿河流优先分配耕地的方法是否合理，在上述两个时间断面上现黄河附近的耕地分布都是不合理的，因为重建的黄河河道演变历史表明，7000a BP(全新世早期)的黄河在现黄河河道之北入渤海(吴忱和何乃华, 1991)，而黄河在1128年夺淮后南流入黄海，直到1855年后才行水于现河道(钮仲勋等, 1994).

5 基于多套数据一致性的可能性评估

5.1 评估方法

一致性评估是一种针对存在多套数据产品，但无客观评估参照标准的可靠性评估方法. 该方法假设每套土地覆盖数据集所采用的数据和方法均有其合理性, 但也各自存在一定的局限性, 各数据产品能够不同程度地逼近土地覆盖的“真值”，即有相当比例的数据与真实情况相吻合或接近真实情况，同时也存在一定

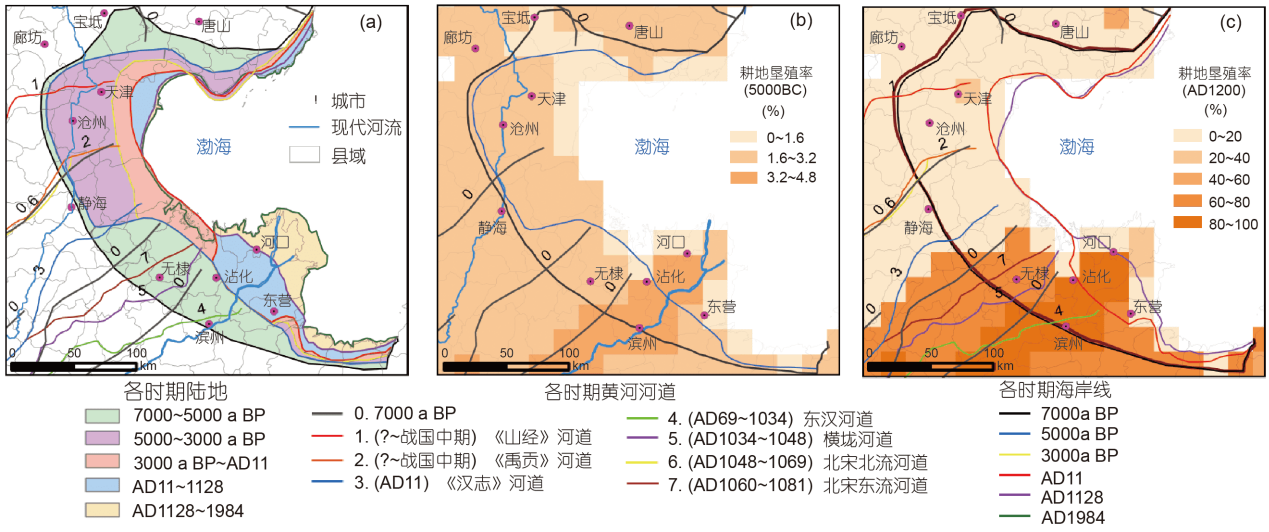

图 5　渤海西部历史时期海岸线和黄河河道位置及HYDE3.2数据集垦殖率分布比较

(a) 渤海海岸线(薛春汀, 2009)与黄河河道演变(钮仲勋等, 1994; 刘国纬, 2011); (b) 5000BC HYDE3.2垦殖率(Goldewijk等, 2017); (c) AD1200 HYDE3.2垦殖率(Goldewijk等, 2017)

比例与真实情况不符的错判. 将每套全球土地覆盖数据产品视为对真实土地覆盖状况的一种专家判断, 用不同数据集对给定空间单元或时段土地覆盖类型或数量特征判断的一致性程度作为指标, 以此评估该空间单元或时段数据可靠的可能性. 根据一致性的高低可将评估结果分为可能性较高、可能性一般和可能性较低三种类型, 一致性越高, 数据可靠的可能性越大.

对多套土地覆盖数据产品进行时间和空间叠加分析, 将各个空间单元或时段为某种土地覆盖类型或该类覆盖达到某一覆盖率的数据集套数或占总套数的比例定义为重合度. 重合度由大到小代表一致性从高到低, 重合度越高意味着不同数据集对该空间单元或时段土地覆盖类型或数量特征判断的一致性程度越高, 该空间单元或时段数据可靠的可能性越大.

5.2　案例5: 10套现代全球土地覆盖数据集中中国耕地数据的一致性

选取10套国际广泛使用的全球现代土地覆盖数据集, 从每套数据集中提取中国1km的耕地像元, 利用ArcGIS中的地图代数工具进行空间叠加, 叠加结果(重合度)取值范围为0~10, 0表示没有一款数据认为该像元上有耕地存在, 10表示所有数据集一致认为该像元有耕地分布, 数值越大表明在该像元有耕地分布的可能性越大. 对如何根据一致性的高低判断可靠性尚无统一的标准, 本案例根据10套数据的重合度将耕地像元划分为3个信度等级: (1) 1~3为一致性较低, 即可靠的可能性较低; (2) 4~5为一致性中等, 即可靠的可能性一般; (3) 6~10为一致性较高, 即可靠的可能性较高(图6).

从全国来看, 一致性较高与较低的比例大致相当, 分别为42%和43%, 一致性中等的比例占15%. 即在所有耕地格网中, 可靠的可能性较高的格网只有42%. 从分省看, 有20个省份一致性较高耕地格网的比例>50%, 全部集中在中国的东部和中部(除内蒙、浙江和福建), 主要分布在华北平原、成都平原、东北平原等地区. 6个省份一致性较高和较低耕地格网的比例均<50%, 包括陕西、甘肃、宁夏、浙江、福建和云南, 主要分布在农牧交错带和南方丘陵山地. 有6个省份一致性较低耕地格网的比例>50%, 包括西藏、青海、新疆、四川、台湾和内蒙古, 主要分布在农牧交错带北部、南方丘陵山地、青藏高原以及西北山区和山麓地带.

一致性较高的耕地格网集中分布在中国主要的农耕区, 这些区域具有适宜农业生产的气候环境和平坦的地形, 耕地开垦强度高, 集中成片分布且无大面积的多年休耕, 种植作物类型稳定, 易于在遥感影像中准确辨识. 一致性较低的耕地格网大多分布在农业气候条件相对恶劣、地形复杂的中低垦殖强度区域, 主要对应布尔型数据中的混合耕地像元, 这些混合耕地

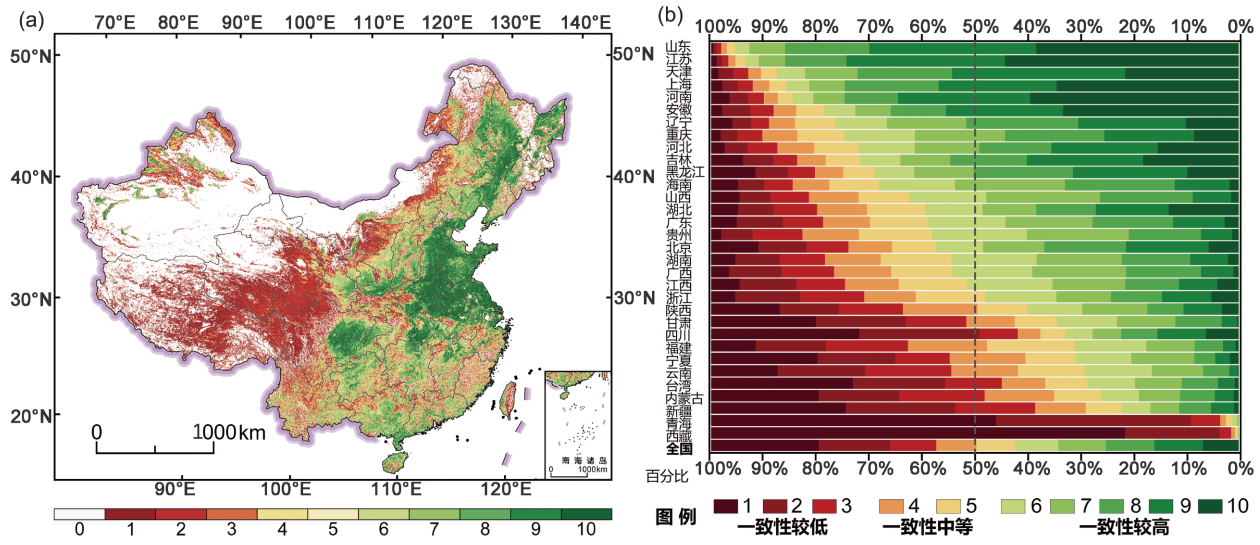

图 6　10套全球耕地数据集空间叠加显示的中国耕地空间分布一致性

根据重合度划分的一致性(可靠的可能性)等级: 1~3为较低; 4~5为中等; 6~10为较高. (a) 10套耕地数据的空间分布一致性; (b) 32个省级单元(不包括香港和澳门)的一致性程度统计. 10套全球土地覆盖数据集包括: (1) 7套布尔型数据集: IGBP-DISCover(Loveland等, 2000), GLC-UMD (Hansen等, 2000), GLC-MODIS(Friedl等, 2002), GLC2000(Bartholomé和Belward, 2005), GLCNMO(Tateishi等, 2001), ESA-CCI_LC(Bontemps等, 2013)和GlobeLand30(Chen等, 2015), 其耕地像元包括耕地类型和耕地与自然植被混合类型两类; (2) 3套融合数据集: GLC-Consensus (Tuanmu和Jetz, 2014), GLC-Share(Latham等, 2014)和HybridCropland(Fritz等, 2015), 其耕地像元为耕地比例≥1%的像元

像元辨识的不确定和各套数据集采用的耕地分类标准的不同, 都可能使得耕地格网的一致性较低.

6　结论

土地利用/土地覆盖变化(LUCC)既是全球变化的表现形式, 也是全球变化的重要驱动力, 在深入了解过去土地利用/覆盖时空演变过程的基础上, 建立面向全球变化研究的全球LUCC数据产品, 是定量评估LUCC对全球变化影响的基础. 自20世纪90年代以来, 国际上相继建立了多套面向全球变化研究的全球历史土地利用/覆盖数据集, 但上述全球数据集彼此之间存在一定的矛盾, 且均已被许多区域尺度重建结果证明存在很大偏差.

评估已有土地覆盖数据产品的可靠性是改进数据集质量的基础和前提. 针对可用于可靠性评估的历史数据有限的现实, 本文构建了全球历史LUCC数据集数据可靠性评估框架, 根据可用于历史土地覆盖数据评估的数据和指标的类型和特点, 提出4种可靠性评估方法, 并以耕地数据为评估对象, 给出了5个具体评估案例. 主要结论如下:

(1) 可靠性评估包括针对给定的空间单元的时间过程评估和针对给定时间断面的空间分布评估两个方面. 可靠性评估的难点在于, 作为评价全球土地覆盖数据产品可靠性参照值的“真值”实际上就是重建的目标值, 除少数可利用接近真值的定量重建值作为参照标准进行准确性评估外, 更多是能够体现真值特点的代用指标, 从合理性和一致性的角度间接地判断全球土地覆盖数据产品的可靠性.

(2) 准确性评估以定量重建的区域LUCC数据为基准, 是一种定量评估. 它假设基于历史文献或自然记录定量重建的区域土地覆盖数据最接近相应类别数据的历史真值, 根据区域内全球LUCC数据集中的数据相对于历史土地覆盖重建值的接近程度, 定量评估数据的准确性.

(3) 合理性评估是一种定性评估, 根据所用证据的不同, 可分为基于区域发展史实的合理性评估和基于自然和人文特征及专家知识的合理性评估两类. 合理性评估通过比对全球LUCC数据集的数据与区域史实在时空分布上的一致性, 或判别全球LUCC数据集所用假设及方法与相关自然和人文专家知识的符合程度, 对全球数据集的数据进行合理性判断.

(4) 可能性评估是根据多套数据产品的一致性判断其可靠的可能性. 该方法将每套全球土地覆盖数据产品视为对真实土地覆盖状况的一种专家判断, 各数据集对给定空间单元或时段土地覆盖类型或数量特征判断的一致性程度越高, 其可靠的可能性越大.

本文提出的历史LUCC评估方法的适用范围不局限为耕地, 亦适用于林地、草地等其他土地类型的评估(何凡能等, 2019; 杨帆等, 2019). 在数据可靠性评估中, 针对定量的准确性和可能性评估, 如何客观、合理地确定定量的判别标准或阈值, 如何将案例研究推广到更大时空尺度数据的可靠性判断, 有待在更大范围的区域评估中讨论.

参考文献

陈军, 陈晋. 2018. GlobeLand30遥感制图创新与大数据分析. 中国科学: 地球科学, 48: 1391–1392

方修琦, 叶瑜, 葛全胜, 郑景云. 2005. 从城镇体系的演变看清代东北地区的土地开发. 地理科学, 25: 129–134

葛全胜, 戴君虎, 何凡能, 潘媛, 王梦麦. 2008. 过去300年中国土地利用、土地覆被变化与碳循环研究. 中国科学D辑: 地球科学, 38: 197–210

何凡能, 李士成, 杨帆, 李美娇. 2019. 全球历史土地利用数据集中国草地数据可靠性评估. 中国科学: 地球科学, 49: 421–433

何凡能, 李士成, 张学珍, 葛全胜, 戴君虎. 2012. 中国传统农区过去300年耕地重建结果的对比分析. 地理学报, 67: 1190–1200

李蓓蓓, 方修琦, 叶瑜, 张学珍. 2014. 中国东北地区过去300年耕地开垦导致的碳收支. 中国科学: 地球科学, 44: 1987–1996

李蓓蓓, 方修琦, 叶瑜, 张学珍. 2010. 全球土地利用数据集精度的区域评估——以中国东北地区为例. 中国科学: 地球科学, 40: 1048–1059

刘国纬. 2011. 黄河下游治理的地学基础. 中国科学:地球科学, 41: 1511–1523

钮仲勋, 等. 1994. 历史时期黄河下游河道变迁图. 北京: 测绘出版社

商志文, 王宏, 李建芬, 施佩歆. 2018. 渤海湾沧桑巨变: 渤海湾1.1万年来的海陆演化过程. 中国矿业, 27: 286–289

吴忱, 何乃华. 1991. 2万年来华北平原主要河流的河道变迁. 见: 吴忱, 等, 主编. 华北平原古河道研究论文集. 北京: 中国科学技术出版社. 132–148

薛春汀. 2009. 7000年来渤海西岸南岸海岸线变迁. 地理科学, 2: 217–222

杨帆, 何凡能, 李美娇, 李士成. 2019. 全球历史森林数据中国区域的可靠性评估. 地理学报, 74: 923–934

杨绪红, 金晓斌, 项晓敏, 范业婷, 刘晶, 单薇, 周寅康. 2019. 近300年中国耕地开垦导致的碳排放估算. 中国科学: 地球科学, 49: 554–568

叶瑜, 方修琦, 任玉玉, 张学珍, 陈莉. 2009. 东北地区过去300年耕地覆盖变化. 中国科学D辑: 地球科学, 39: 340–350

曾早早, 方修琦, 叶瑜. 2011. 基于聚落地名记录的过去300年吉林省土地开垦过程. 地理学报, 66: 986–993

Bartholomé E, Belward A S. 2005. GLC2000: A new approach to global land cover mapping from Earth observation data. Int J Remote Sens, 26: 1959–1977

Bontemps S, Defourny P, Radoux J, Van Bogaert E, Lamarche C, Achard F, Mayaux P, Boettcher M, Brockmann C, Kirches G, Zülkhe M, Kalogirou V, Seifert F M, Arino O. 2013. Consistent global land cover maps for climate modelling communities: Current achievements of the ESA's land cover CCI. Edimburgh: Proceedings of the ESA Living Planet Symposium. 9–13

Bork H R, Bork H, Dalchow C, Faust B, Piorr H P, Schatz T. 1998. Landschaftsentwicklung in Mitteleuropa: Wirkung des Menschen auf Landschaften. Gotham: Klett. 328

Chen J, Chen J, Liao A, Cao X, Chen L, Chen X, He C, Han G, Peng S, Lu M, Zhang W, Tong X, Mills J. 2015. Global land cover mapping at 30m resolution: A POK-based operational approach. ISPRS J Photogramm Remote Sens, 103: 7–27

Ellis E C, Kaplan J O, Fuller D Q, Vavrus S, Klein Goldewijk K, Verburg P H. 2013. Used planet: A global history. Proc Natl Acad Sci USA, 110: 7978–7985

Foley J A, Defries R, Asner G P, Barford C, Bonan G, Carpenter S R, Chapin F S, Coe M T, Daily G C, Gibbs H K, Helkowski J H, Holloway T, Howard E A, Kucharik C J, Monfreda C, Patz J A, Prentice I C, Ramankutty N, Snyder P K. 2005. Global consequences of land use. Science, 309: 570–574

Friedl M A, McIver D K, Hodges J C F, Zhang X Y, Muchoney D, Strahler A H, Woodcock C E, Gopal S, Schneider A, Cooper A, Baccini A, Gao F, Schaaf C. 2002. Global land cover mapping from MODIS: Algorithms and early results. Remote Sens Environ, 83: 287–302

Fritz S, See L, McCallum I, You L, Bun A, Moltchanova E, Duerauer M, Albrecht F, Schill C, Perger C, Havlik P, Mosnier A, Thornton P, Wood-Sichra U, Herrero M, Becker-Reshef I, Justice C, Hansen M, Gong P, Abdel Aziz S, Cipriani A, Cumani R, Cecchi G, Conchedda G, Ferreira S, Gomez A, Haffani M, Kayitakire F, Malanding J, Mueller R, Newby T, Nonguierma A, Olusegun A, Ortner S, Rajak D R, Rocha J, Schepaschenko D, Schepaschenko M, Terekhov A, Tiangwa A, Vancutsem C, Vintrou E, Wenbin W, van der Velde M, Dunwoody A, Kraxner F, Obersteiner M. 2015. Mapping global cropland and field size. Glob Change Biol, 21: 1980–1992

Gaillard M J, LandCover6k Interim Steering Group Members. 2015.

LandCover6k: Global anthropogenic land-cover change and its role in past climate. PAGES Program News, 23: 38–39

Gaillard M J, Morrison K, Madella M, Whitehouse N. 2018. Past land-use and land-cover change: The challenge of quantification at the subcontinental to global scales. Past Land Use and Land Cover, 26: 3

Goldewijk K K. 2001. Estimating global land use change over the past 300 years: The HYDE database. Glob Biogeochem Cycle, 15: 417–433

Goldewijk K K, Battjes J J. 1997. A Hundred Year (1890–1990) Database for Integrated Environmental Assessments (HYDE, version1.1). Bilthoven: National Institute of Public Health and the Environment

Goldewijk K K, Beusen A, Doelman J, Stehfest E. 2017. Anthropogenic land use estimates for the Holocene–HYDE 3.2. Earth Syst Sci Data, 9: 927–953

Goldewijk K K, Beusen A, Drecht G V, Martine D V. 2011. The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12000 years. Glob Ecol Biogeogr, 20: 73–86

Goldewijk K K, Drecht G V, Bouwman A F. 2007. Mapping contemporary global cropland and grassland distributions on a 5 × 5 minute resolution. J Land Use Sci, 2: 167–190

Hansen M C, Defries R S, Townshend J R G, Sohlberg R. 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. Int J Remote Sens, 21: 1331–1364

He F N, Li S C, Zhang X Z, Ge Q S, Dai J H. 2013. Comparisons of cropland area from multiple datasets over the past 300 years in the traditional cultivated region of China. J Geogr Sci, 23: 978–990

Jia D, Li Y, Fang X. 2018. Complexity of factors influencing the spatiotemporal distribution of archaeological settlements in north-east China over the past millennium. Quat Res, 89: 413–424

Kaplan J O, Krumhardt K M, Ellis E C, Ruddiman W F, Lemmen C, Goldewijk K K. 2010. Holocene carbon emissions as a result of anthropogenic land cover change. Holocene, 21: 775–791

Kaplan J O, Krumhardt K M, Zimmermann N. 2009. The prehistoric and preindustrial deforestation of Europe. Quat Sci Rev, 28: 3016–3034

Latham J, Cumani R, Rosati I, Bloise M. 2014. Global land cover share (GLC-SHARE) database beta-release version 1.0-2014

Leite C C, Costa M H, Soares-Filho B S, de Barros Viana Hissa L. 2012. Historical land use change and associated carbon emissions in Brazil from 1940 to 1995. Glob Biogeochem Cycle, 26: 1–13

Li B B, Jansson U, Ye Y, Widgren M. 2013. The spatial and temporal change of cropland in the Scandinavian Peninsula during 1875–1999. Reg Environ Change, 13: 1325–1336

Li Y K, Ye Y, Zhang C C, Li J, Fang X Q. 2019. A spatially explicit reconstruction of cropland based on expansion of polders in the Dongting Plain in China during 1750–1985. Reg Environ Change, 5, doi: 10.1007/s10113-019-01575-8

Loveland T R, Reed B C, Brown J F, Ohlen D O, Zhu Z, Yang L, Merchant J W. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. Int J Remote Sens, 21: 1303–1330

Pielke R A, Pitman A, Niyogi D, Mahmood R, McAlpine C, Hossain F, Goldewijk K K, Nair U, Betts R, Fall S, Reichstein M, Kabat P, Noblet N D. 2011. Land use/land cover changes and climate: Modeling analysis and observational evidence. WIREs Clim Change, 2: 828–850

Pongratz J, Reick C, Raddatz T, Claussen M. 2008. A reconstruction of global agricultural areas and land cover for the last millennium. Glob Biogeochem Cycle, 22: 1–16

Ramankutty N, Foley J A. 1999. Estimating historical changes in global land cover: Croplands from 1700 to 1992. Glob Biogeochem Cycle, 13: 997–1027

Ruddiman W F. 2003. The anthropogenic greenhouse era began thousands of years ago. Clim Change, 61: 261–293

Tateishi R, Bayaer M A, Al-Bilbisi H, Tsendayush J, Shalaby A, Kasimu A, Hoan N T, Kobayashi T, Alsaaideh B, Rahman M M, Tsevengee E, Yamada Y, Kajikawa S. 2001. A new global land cover map, GLCNMO. Forest, 8: 1369–1372

Tuanmu M N, Jetz W. 2014. A global 1-km consensus land-cover product for biodiversity and ecosystem modelling. Glob Ecol Biogeogr, 23: 1031–1045

Wei X Q, Ye Y, Zhang Q, Fang X Q. 2016. Reconstruction of cropland change over the past 300 years in the Jing-Jin-Ji area, China. Reg Environ Change, 16: 2097–2109

Wei X, Ye Y, Zhang Q, Li B, Wei Z. 2019. Reconstruction of cropland change in North China Plain Area over the past 300 years. Glob Planet Change, 176: 60–70

Yuan C, Ye Y, Tang C, Fang X. 2017. Accuracy comparison of gridded historical cultivated land data in Jiangsu and Anhui Provinces. Chin Geogr Sci, 27: 273–285

Zhang X Z, Wang W C, Fang X Q, Ye Y, Zheng J Y. 2012. Agriculture development-induced surface albedo changes and climatic implications across northeastern China. Chin Geogr Sci, 22: 264–277

Zumkehr A, Campbell J E. 2013. Historical U.S. cropland areas and the potential for bioenergy production on abandoned croplands. Environ Sci Technol, 47: 3840–3847

(责任编委: 岳天祥)