

基于分层分区法的中国历史耕地数据的网格化重建

魏希文¹, 缪丽娟², 江 源¹, Nath Reshmita³, 崔雪锋⁴

(1. 北京师范大学资源学院, 北京 100875; 2. 南京信息工程大学地理与遥感学院, 南京 210044;

3. 中国科学院大气物理研究所季风系统研究中心, 北京 100029;

4. 北京师范大学系统科学学院, 北京 100875)

摘要: 网格化历史耕地数据集能为历史时期耕地变化研究提供更精确的支持, 并且为全球环境气候变化研究模型模拟提供驱动数据。本文综合考虑了中国历代土地利用开发的特点及自然人文因子对耕地的影响, 设计了一套对中国耕地先分区再分层分配的网格化方法。基于国内3个主流区域耕地数据研究成果, 采用上述方法建立了1820年(清仁宗嘉庆二十五年)和1936年(民国二十五年)中国10 km×10 km分辨率的耕地数据集, 并绘制了分布图。本文还利用国内具有代表性的区域数据集对重建结果进行对比验证。结果表明, 该方法可以保证耕地数量的权威性, 并且建立具有区域性的高精度历史耕地数据集。

关键词: 中国; 历史耕地; 分区分层; 网格化重建

DOI: 10.11821/dlxb201607005

1 引言

自工业革命以来, 全球环境变化主要是由于人为和自然两方面因素所引起的^[1]。现今已经在海洋变暖、全球水循环变化、海平面上升和一些极端气候事件中检测到人类的影响^[2], 其中最直接的就是人类对土地利用和土地覆盖的作用^[3], 大概1/3到1/2的陆地表面被人类活动所改变^[4], 并且土地覆盖的改变引起了对气候系统一系列的反馈作用^[5]。正因为土地利用/覆盖情况的变化对于全球环境变化的显著影响, 全球变化与陆地生态系统计划(Global Change and Terrestrial Ecosystems, GCTE)、国际地圈生物圈计划(International Geosphere-Biosphere Program, IGBP)和国际全球环境变化人文因素计划(International Human Dimensions Program on Global Environmental Change, IHDP)联合发起的土地利用/覆盖变化(Land Use and Land Cover Change, LUCC)计划被提出, 直到现在, 土地利用/覆盖变化研究还是全球环境变化研究与可持续发展科学的重要组成部分^[6-8]。土地利用变化不是间断的, 而是随着历史的发展而持续演变, 所以历史土地利用信息得到了相当的重视^[9-10], 在各个相关方面的应用也愈发重要, 比如目前在生物群落演替等方向的应用^[11]、国家间历史土地发展模式的比较与探究^[12]或者是对历史时期气候的影响^[13]。近年来全球环境变化研究的进一步发展, 历史土地利用/覆盖的研究也要求越发的深入和精细, 带有空间属性的历史土地利用资料也为目前对未来气候情景的预测和发

收稿日期: 2015-12-07; 修订日期: 2016-02-28

基金项目: 国家基础研究发展计划(973) (2011CB952001); 国家自然科学基金项目(41271542) [Foundation: National Basic Research Program of China (973 Program), No.2011CB952001; National Natural Science Foundation of China, No.41271542]

作者简介: 魏希文(1990-), 男, 湖南长沙人, 硕士, 研究方向为历史土地利用重建。E-mail: weixiwen@mail.bnu.edu.cn

通讯作者: 崔雪锋(1976-), 男, 山西人, 博士, 副教授, 研究方向为气候变化、土地利用、粮食安全和地球工程。

E-mail: xuefeng.cui@bnu.edu.cn

展提供了出发点和参考^[14]。

国内外学者在重建具有空间属性的历史土地利用/覆盖数据集已经展开了大量的研究工作^[15]，主要的重建方法分为三类：历史资料、模型反演以及两者的结合使用^[14, 16]。在全球尺度上，最具代表性的是Goldewijk等^[17]、Ramankutty等^[18]和Pongratz等^[19]的研究，其中Goldewijk等^[17]采用遥感数据和自然要素结合的网格化方法，Ramankutty等^[18]采用遥感数据和历史耕地数据项结合综合处理的网格化方法，各自通过一定的算法建立了过去300年全球土地利用数据集，并且随着数据方法的进步而不断的更新数据集。在区域尺度上，葛全胜等^[20]、何凡能等^[21]、林珊珊等^[22]重建了中国传统农区的历史土地数据集，并被广泛使用；叶瑜等^[23-25]重点关注东北地区耕地、草地和林地的土地利用变化及其驱动机制研究；李柯等^[26]通过结合遥感数据与历史耕地资料，重建了中国清代云南省的耕地空间分布；Long等^[27]首次使用约束性CA模型重建江苏省历史土地利用数据；李士成等^[28]通过量化自然影响因素结合耕地的史册和统计数据，重建了过去百年青海和西藏的耕地空间格局。从上可以看出，中国区域范围的历史土地利用重建研究已经相对成熟，但是全国尺度的历史重建还少有开展，其中代表性的研究只有Liu等^[29]重建的中国1700-2005年的耕地数据集和Li等^[30]重建的中国1661-1996年的耕地数据集。

历史耕地网格化就是旨在将史料耕地数据经过数理步奏转化为带有空间信息的耕地数据。充分利用中国丰富的历史文献，建立更高精度的中国区域历史土地利用数据集^[31]，能为历史时期耕地变化研究提供有力的支持，并且为全球环境气候变化研究模型提供模拟输入数据。考虑到如果忽略区域之间的差异性、采用统一方法进行重建，必然会影响到重建数据的精度^[14]。所以文本借鉴分区重建的思路，将中国分为传统农耕区和游牧区两个区域，通过对区域内部各个因素进行分层处理，探讨了全国尺度耕地数据集重建的方法，并以此重建了1820年和1936年这两个时间断面的全国耕地数据来验证精度，期冀此法能对未来全国历史土地利用数据集的重建提供借鉴。

2 数据来源与方法介绍

本文采用分区重建加逐级分配的方法，并借鉴其他学者的区域耕地数据来重建全国的省级耕地数据，然后通过结合人文因子和自然因子，对全国耕地数据进行网格化。最终成图分辨率为10 km×10 km，每个像元的值为0到1，代表其垦殖率（0代表没有耕地存在，1代表是耕地）。

2.1 分区定义：

历史时期的耕地资料，都是以行政区划进行统计^[22]。传统意义上，中国分为四大区域：传统农业区是中国农耕水平发展最为成熟的地区，东北地区是近百年土地变化最为显著的地区，西北地区是中国干旱半干旱地区的代表，而青藏地区以其独有的高原地势区别于其他地区。中国学者在对区域研究划分上也基本遵循这个划分方法^[20, 22, 24, 32]。但在本文中，区域的划分略有改变：青藏地区由于其独特的高耸地势，有“世界屋脊”之称，气候为高原山地气候，全年干旱少雨。受综合条件的影响，该地区在2005年耕地面积仅占土地总面积的4.79%^[33]，所以其主要供给方式为游牧业。而中国历史上，直到明朝时期，游牧文化与农耕文化都是以长城为界^[34]，长城以南是农耕文化，长城以北是游牧文化，西北地区因为其地理气候因素，基本是牧区，所以，本文将西北地区和青藏地区合并，统一称为游牧地区，这个地区供给主要以游牧业为主，农耕水平欠发达；甘肃省虽然位于西北地区，但是其自汉至今，一直是中国养马业的重地，并且自古受农耕文化

的影响,而甘肃省东南大部分地区降水充足,另外黄河穿过其中部,也为农耕提供了有力的条件,所以,本文将甘肃与传统农耕区划分到一起,统称为农耕地区;东北地区近百年土地利用变化最为显著,优越的自然条件让该地区拥有快速发展农业的潜力,故也将其划分到农耕地区范围内。按照此种方法,与用400 mm等降水线做为农牧交错带界线的说法基本一致^[35]。综上所述,本文通过将中国分为两部分,分别为游牧区和农耕区,来对其进行耕地重建。

2.2 数据介绍

2.2.1 耕地数据 耕地数据来源于葛全胜等^[20]对过去内地18省耕地数量重建、叶瑜等^[24]对东北地区过去300年耕地数量重建的结果和李士成等^[28]对过去百年青海和西藏耕地数量的整理。

葛全胜等^[20]结合了清代、民国时期以及新中国的史册数据和调查数据,通过分析订正,得出了一套内地18省的长时间序列耕地重建数据,这里值得注意的是,由于在过去300年内中国政区范围有较大变化,所以选取了政区较为稳定的清代“内地18省”为研究区域,同时作为数据重建的空间范围。

叶瑜等^[24]采用历史数据订正与多源耕地数据关系模型校核相结合的方法,重建了东北地区过去300年耕地面积时间序列和耕地空间格局变化,重建了过去300年东北地区11个时间断面的空间格局。

李士成等^[28]进行了大量的数据收集整理工作,通过引用曹树基^[36]所著《中国人口史:清时期》估算了1910年耕地数据,并订正了1950年到2000年青海和西藏的省域耕地面积数据,由此得出了青海西藏地区过去百年里4个时间断面的耕地数据。

2.2.2 自然数据 自然条件数据来源于魏希文等^[37]结合土壤和气候两个方面影响因子得出的耕地适宜性指数。其中土壤数据为第二次全国土地调查南京土壤所提供的1:100万中国土壤数据,分别选取了土壤30 cm表层有机碳含量和30 cm表层土壤PH值这两个土壤属性值作为土壤参数参与计算耕地适宜性指数。气候数据来自于中国农业科学院农业自然资源与农业区划研究所建立的生态环境背景数据,分别选取了大于零度的有效积温和湿润指数来参与计算,并用数字高程模型(Digital Elevation Model, DEM)进行了校正。

2.2.3 人口数据 人口数据的利用,主要有两类:第一类是利用通过整理史料得出来的第一手成果,第二类是利用其他学者已经订正的数据集。本文的历史人口数据是来源于复旦大学历史地理研究中心创建的《中国人口地理信息系统》^①,该系统现在发布的内容有:1820年中国分府人口地理信息系统、1910年中国分府人口地理信息系统、1936年中国分县人口地理信息系统和2000年中国分县人口地理信息系统。这些内容分别来源于《嘉庆一统志》、北平社会调查所出版的王士达《民政部户口调查及各家估计》、民国内政部统计处1938年发布的数据和国家统计局的2000年人口普查数据。

2.3 重建方法

本文重建方法的基本思路为在分区重建的基础上,对每个区域分层进行分配,最大化的保证耕地数据的权威性。具体来说就是通过整合其他学者具有代表性的区域研究成果,得出全国各省份耕地数量,然后根据不同省份中各县人口比例来分配相应比例耕地,最后利用自然因子指数,将耕地数据进行网格化。

2.3.1 全国省级耕地面积求解 将全国分为游牧区和农耕区进行处理。农耕区依赖的数据来自葛全胜等^[20]过去300年内地18省耕地数据和叶瑜等^[24]过去300年东北地区耕地数据;游牧区依赖的数据来自李士成等^[28]过去百年青海西藏地区耕地格局数据。

① 侯杨方. 中国人口地理信息系统: <http://cpgis.fudan.edu.cn/cpgis/>.

农耕区包括“中国内地18省地区”和“东北地区”,就相当于现今中国除了内蒙古、新疆、青海、西藏以及台湾之外的地区,本文研究区域中的游牧区就是相对农耕区之外的地区(不包括台湾)。在“中国内地18省地区”中,直隶省包括河北、热河以及察哈尔地区;山西含绥远;江苏包含上海;甘肃省包含宁夏^[20]。

农耕区的耕地数据通过整合中国内地18省农区和东北地区的数据得到。本文所考虑的年代为1936年和1820年。对于1936年的耕地数据,在葛全胜等^[20]的数据集中,最贴近的年份为1933年,在叶瑜等^[24]的数据集中最贴近的年份为1940年。中国传统农区在民国时期各地的耕地面积变化起伏较小,数量相对稳定^[20],故使用1933年的数据来对1936年耕地数据直接进行相似计算;而对于东北地区,则采用简单插值法来求算。最后通过整合计算后的数据得出农耕区省级耕地。对于1820年来说,可直接使用传统农区耕地记录数据;而东北地区还是利用插值法求算。对于有些地区人口密度远远小于0.1人/km²的情况,将这些地区归为没有耕地,这是借助Goldewijk等^[17]的假设来处理的:人口密度低于0.1人/km²的没有农业需求。

游牧区在历史上的耕地数据十分缺乏。从中国历史上看,明代游牧文明与农耕文明以长城为明显的界线,长城以南是农耕文明,以北是游牧文明^[34],直到清代,中国政治形式发生前所未有的新变化,使长城失去了向北防御的意义,农耕与游牧被置于统一的中央政权管制下,才开始发生融合^[38],但这种融合发展十分缓慢,直至今天,游牧区大部分居民都保持着其独特的饮食结构和生活习性。以此可以看出:游牧区有着完全不同于农耕区的农耕结构。正是因为游牧地区耕地数据的缺乏和其农耕结构的不同,需要有一个合理的方法来估算这部分耕地的数据来用于空间分配。总体来说,游牧地区因为气候、地势地貌、文化传统等原因,都保持着“食肉饮酪”的饮食传统,这就意味着游牧地区对于耕种这一概念的接受只存在于很小的人群中,加之受当时全国耕种水平的限制,游牧地区的农业需求和人均耕地水平也就保持在一个很接近的水平;就区域来说,青海、西藏地区作为全球变化的敏感带,很有研究意义与目的,所以也引起了一些学者对它的研究^[28, 39],研究结果很有借鉴意义,也比较准确合理。因此,在本文中,先假设游牧区的人均耕地一样,然后通过计算青海、西藏地区的人均耕地,结合游牧区其他省人口数据,从而推断出游牧区的省级耕地数量,而李士成的数据集没有1936年数据,所以这里也用插值法得出1936年数据。对于牧区而言,由于饮食结构的不同,农耕需求也不同于农耕区,于是本文在Goldewijk等^[17]的假设基础上提出进一步假设:牧区人口密度低于0.2人/km²没有农业耕种需求。对于1820年,游牧区除了丝绸之路新疆段之外都没有人口,而这新疆部分人口具有流动、少量的特点,人口密度达不到0.2人/km²,并且游牧区农作物的获取主要通过与中原地区交换所得^[40]。所以说在1820年的游牧区,没有耕种需求,耕地设为0。

数据处理过程中,由于耕地数据与人口数据的数据源不同,造成了数据地图的不匹配。葛全胜等^[20]数据集选取的政区稳定的康熙初年地图作为数据重建地图;李士成等^[28]研究区域为张德铨等^②绘制的现代青藏高原地区;而复旦大学的人口地图均为根据相应时间断面匹配的地图。所以在本文中,耕地数据只使用耕地数量部分,相互校正来反应到人口数据的地图上来。

2.3.2 耕地数据县级分配 古代农业社会中人口是影响耕地分布的主导社会因素^[22]。人作为生产者,与耕地联系最为密切,历史上的战乱与政策,也是先直接作用于生产者然后

② 张德铨,李炳元,郑度. 青藏高原范围与界线地理信息系统数据: <http://www.geodoi.ac.cn/WebCn/doi.aspx?DOI=10.3974/geodb.2014.01.12.V1>

间接影响耕地。所以对于某一区域而言,人口的增长导致农业需求的上涨,这意味着更多耕地的开垦利用。而古代交通不便,各地基本处于自给自足的状态,并且历史时期的中国生产力水平从铁器时代以来就没有显著地提高,于是小区域内对粮食需求的提升就要求更多的劳动力投入。正因为人口因素是耕地分布的向导,故在区域尺度上,人口数量与耕地数量成正比;人口的流向代表了农业需求的转变,相对应的耕地分布也会发生变化。

本文基于全国省级耕地数据,按照每个县(府)人口占对应省份总人口的比例来分配耕地面积,分配公式如下:

$$C_j = \frac{Pop_j}{Pop_p} \times C_p \quad (1)$$

式中: C_j 代表 j 县(府)的耕地面积; Pop_j 代表 j 县(府)的人口总数; Pop_p 代表 p 省的人口总数; C_p 代表 p 省的耕地总数。本文中,因为耕地是按照人口占全省的比例来分配的,所以如果某县出现人口大量集中,给这个区域分配的耕地面积就会大于其本身面积,于是就要对有耕地溢出的县进行平滑处理:当某县的垦殖率大于 0.9 时^[22],就将其溢出的耕地数量平均分配到其他县中。

在分配的过程中,还需要规定限制性条件,来确保分配过程的合理与数据整体的一致性:① 对于全国重要的行政地区(如济南市、长沙市),经济行政职能远远大于农业功能,不予分配耕地。这些区域的共同特点就是人口密度大。于是规定这些行政区的面积为小于最终结果图一个像元大小(100 km²)并且人口密度大于 1000 人/km² 的区域。② 当溢出县位于省边界时,平滑处理只针对与溢出县同处一个省的邻县,以保证省级总耕地面积不变。

2.3.3 耕地网格化重建 在全国县级耕地确定的情况下来进一步分配小区域的耕地很大程度上依赖于县(府)内的自然环境状况如水热条件、地势等。本文使用的分配指标“耕地适宜性指数”由土壤酸碱度、土壤有机碳含量、有效积温和湿润指数 4 个参数构成。该指标并非是某一短时间尺度的耕地分布指示指标,而是在长时间尺度下根据自然状况对耕地分配具有分布趋势指示的一个指标,它是根据耕地适宜性的大小来分配县(府)级条件下的耕地。而在网格化过程中,为了使分配简明,故将归一化的“耕地适宜性指数”分为 0 到 9 这十个级别,其中 0 代表不适宜耕种。这个 0 级别的确认,是对构成指数的 4 个方面进行了详细的探究决定的:中国农业主要作物为稻、小麦和玉米,综合文献中对主要作物的研究记载,统一规定 2000 °C 为积温阈值^[41-43];土壤酸碱度,在综合考虑了中国主要农作物的耐性范围后,取 6~8 为合适区间^[44];而土壤有机碳和湿润指数,因为在过去农耕不发达的情况下,土壤水分和营养很大程度依靠这两个指标反映,所以本文中设定这两个指标没有阈值,数值越高表示越适宜耕种。

在县级耕地数据确定的情况下,耕地网格化的公式如下:

$$\eta(i, j) = \frac{\left[\frac{V(i, j)}{\sum V(i, j)} \times C_j \right]}{area(i)} \quad (2)$$

式中: $\eta(i, j)$ 代表 j 县 i 像元的垦殖率; $V(i, j)$ 代表 j 县 i 像元的值; $\sum V(i, j)$ 为 j 县所有像元值的总和; C_j 为 j 县的总耕地面积; $area(i)$ 代表像元 i 代表的面积,这里为 100 km²。

按照上述分配方法,会出现人口集中的县内耕适因子相差较大的情况,这时耕适因子大的像元就会出现溢出:在成果图中,1936 年的溢出率为 0.81%、1820 年的溢出率为

0.49%，经分析认为本文中最终网格化的溢出率很小并且在可接受范围。从溢出的形成机理上来说，96.2%的溢出像元都处于上一级县级耕地分配中垦殖率大于0.6的县，而超过50%的溢出像元位于上一步分配方案的平滑结果中垦殖率为0.9的县，基本呈集中式分布；并且这些溢出像元在县级范围内，都拥有较高的“耕地适宜性”因子等级。所以总结来看，个别像元的溢出，反应出的是相关区域的人文自然环境优越，对于大尺度的分析来说，基本可以忽略其影响。

3 重建结果对比

按照上述重建方法，本文重建了1820年和1936年分辨率为 $10\text{ km}\times 10\text{ km}$ 的耕地数据。这两个时期在历史上有着重要的研究意义：1820年道光皇帝继位，当时的GDP占全球总量的33%，而因为“康乾盛世”的影响，人口持续增长，这也为农业的发展带来了需求和可能；1936年是民国二十五年，共产党长征后的成功会师加上西安事变，结束了长达10年的内战，奠定了全民族抗战的基础，成为由国内战争走向抗日民族战争的转折点。从两期数据（图1）可以看出，1820-1936年耕地数量有明显提升，并且呈现出由集中聚集逐渐扩散的趋势。东北地区耕地开发最为明显，也是农业发展最快的区域；华北平原自古以来都是丰裕之地，一直都是农耕集中的地区；西南地区农耕有明显的发展，在清民时期随着“改土归流”的实施和移民潮的涌现，土地垦殖迅猛发展。

3.1 区域尺度验证

小区域尺度的历史土地利用覆盖数据集的重建结果对更大一级的区域历史土地利用覆盖数据集具有补充和验证的作用^[45]。所以，本文采用之前学者在中国所做的精确小区域重建数据集，对本文结果进行对比验证。

对于中国小区域，林珊珊等^[22]根据中国历史时期土地利用开发的特点，通过量化自然人文条件以及耕地分布关系，通过一定算法重建了中国传统农区1820年的耕地网格化数据，分辨率为 $60\text{ km}\times 60\text{ km}$ ；Yang等^[46]考虑自然人文等多种因素，利用基于约束性CA模型建立的“自下而上”的分配模型，重建了中国传统农区7个时间断面（1661年、

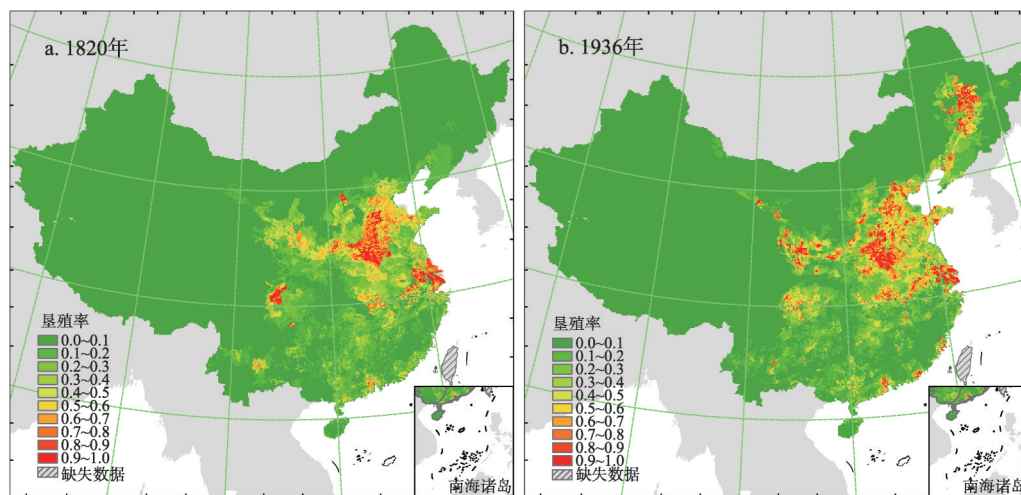


图1 1820年和1936年中国耕地垦殖率($10\text{ km}\times 10\text{ km}$)

Fig. 1 Reclamation ratio of the cropland in China in 1820 and 1936 (at a resolution of $10\text{ km}\times 10\text{ km}$)

1724年、1820年、1887年、1887年、1933年、1952年)的耕地分布,分辨率为 $1\text{ km}\times 1\text{ km}$;李士成等^[15]以现代耕地分布格局为基础,通过量化多种自然人文要素与耕地空间分布的关系,重建了清代西南地区6个时间断面(1661年、1724年、1784年、1820年、1873年、1911年)分辨率为 $10\text{ km}\times 10\text{ km}$ 的耕地数据;叶瑜等^[24]采用历史数据订正与多源耕地数据模型校核相结合的方法,重建了东北地区过去300年耕地面积时间序列,分辨率为县级尺度。本文利用以上数据作为验证数据,来对本研究产生的农耕区数据进行校验与评估,而1936年部分在以上学者的数据中并无匹配的耕地数据,所以,采用相近年代进行验证:传统农区采用1933年数据,东北地区采用1940年数据。

图2a、2c、2e、2g分别对应1820年传统农区 $60\text{ km}\times 60\text{ km}$ 耕地分布图、1933年传统农区 $1\text{ km}\times 1\text{ km}$ 耕地分布图、1820年西南地区 $10\text{ km}\times 10\text{ km}$ 耕地分布图和1940年东北地区县级尺度耕地分布图,简称为对比数据;而图2b、2d、2f、2h是对应对比数据而截取的本文成果图,简称为农耕区数据。总体而言,农耕区数据与对比数据的分布格局大致相同,具有较高的匹配度。具体来看,图2a和图2b的总体分布趋势基本相同,农耕区数量在黄土高原北部、云贵高原中部和珠江三角洲地区有所偏高,但因其分辨率具有差异,故将两者分辨率统一后基本一致;从图2c和图2d来看,图2c数据因为是布尔型,所以只能显示耕地和非耕地,图2d与图2c对应的分布范围和趋势一致,在华北平原略有偏高;图2e和图2f对比的是西南地区的耕地重建数据,总体看农耕区重建数据对比数据有所偏高,特别是四川中部地区和云南中东部地区,这里是如今成都和昆明的所在地,长久以来人口聚集,地理自然环境优越,对农耕的需求很大;图2g和图2h是东北地区对比情况,图2g是县级尺度,图2h分辨率为 $10\text{ km}\times 10\text{ km}$,所以在这,农耕区数据加上对比数据的县级边界,以此来粗略对比,可以看出,农耕区数据在东北地区西南部耕地数量偏少,其他地区基本吻合。

对于游牧区的验证,由于其历来饮食结构与农耕区不同,农业需求不大,相关历史统计资料缺少,故少有国内外专家研究此区域的精确耕地分布数据。本文游牧区的人均耕地参考了李士成等^[28]过去百年青海西藏耕地统计数量而得,但由于在分配过程中,人口数据缺少这块区域的县级数据,分配过程完全依靠自然因子数据,也就缺失了本文方法中人文自然因素中的人文因素,这会对青海西藏地区的分配产生影响,同时本文所选取自然参数与李士成等选取的参数不同,也就更加导致两个结果的差异性,所以本文在游牧区不做精度对比。但是通过游牧区的人均耕地一致来推断其他省耕地数量这一假设是合理的,即游牧区各省的省级耕地总面积合理有效,在拥有相关数据的情况下,网格化结果可以达到农耕区的精度效果,所以说,本文在游牧区的方法是正确的,希望在以后数据充足的条件下,作出更精确的游牧区耕地网格化图。

3.2 全国尺度验证

具有代表性的全国尺度历史土地利用数据集有Liu等^[29]重建的中国1700-2005年耕地数据集和Li等^[30]重建的中国1661-1996年的耕地数据集。因为数据匹配原因,本文只利用Li等的数据集的相关年份来进行对比验证(图3)。Li等^[30]根据相关历史耕地研究数据、人口数据和政府统计数据总结出了中国省级耕地数量数据集,然后基于提出的耕地适宜性指数将耕地分配到网格,得到全国8个时间断面(1661年、1724年、1784年、1820年、1873年、1933年、1980年、1996年),分辨率为 $10\text{ km}\times 10\text{ km}$ 的耕地分布数据。本文取其相应年份数据作为验证数据,来对本研究产生的结果进行验证和评估,因考虑到1936年无匹配数据,故采用1933年作为相似年份来进行验证。

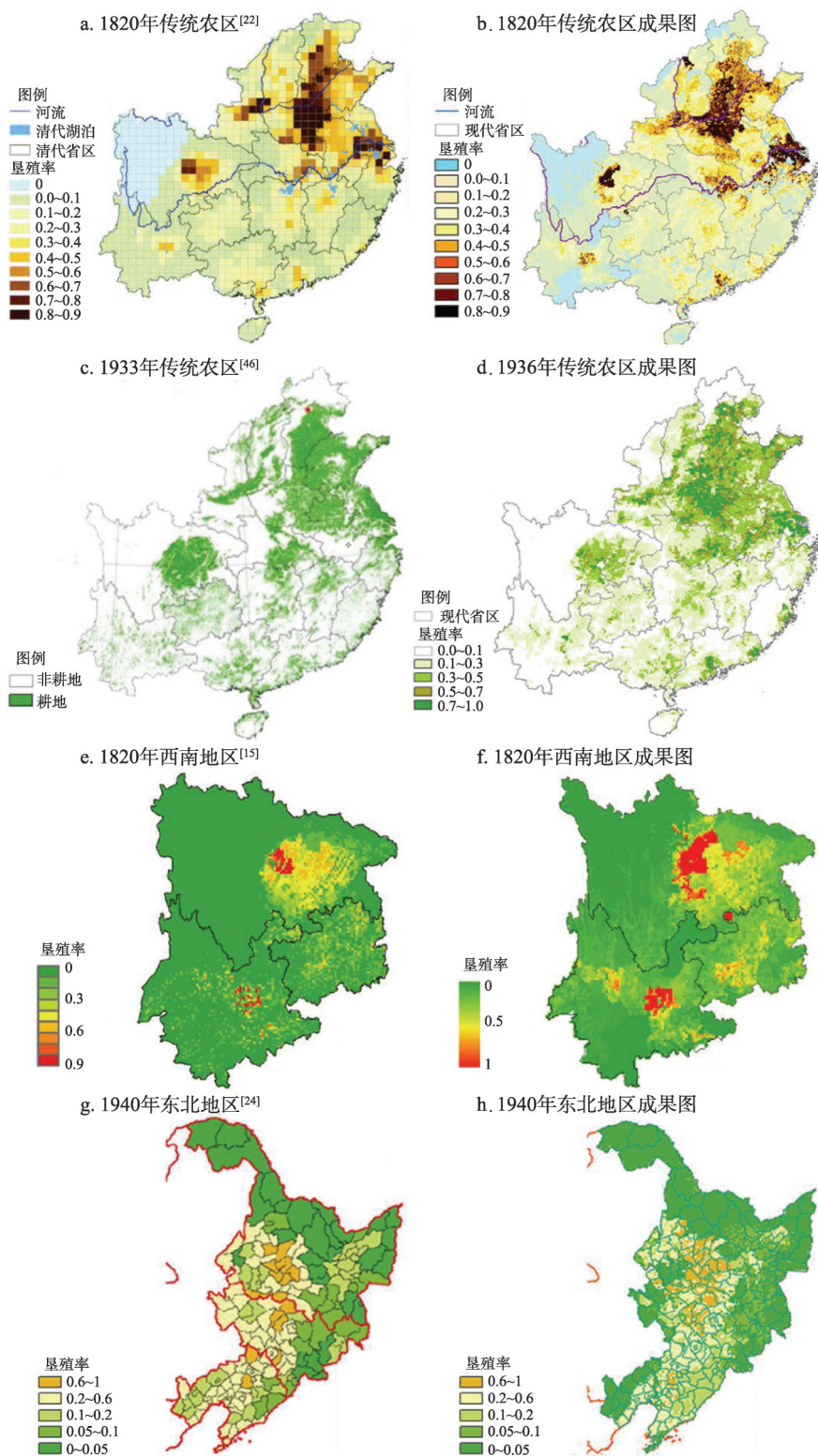


图2 本文成果与其他研究成果^[15, 22, 24, 46]的区域尺度对比验证

Fig. 2 Comparison of the reconstructed map with different results on regional scale

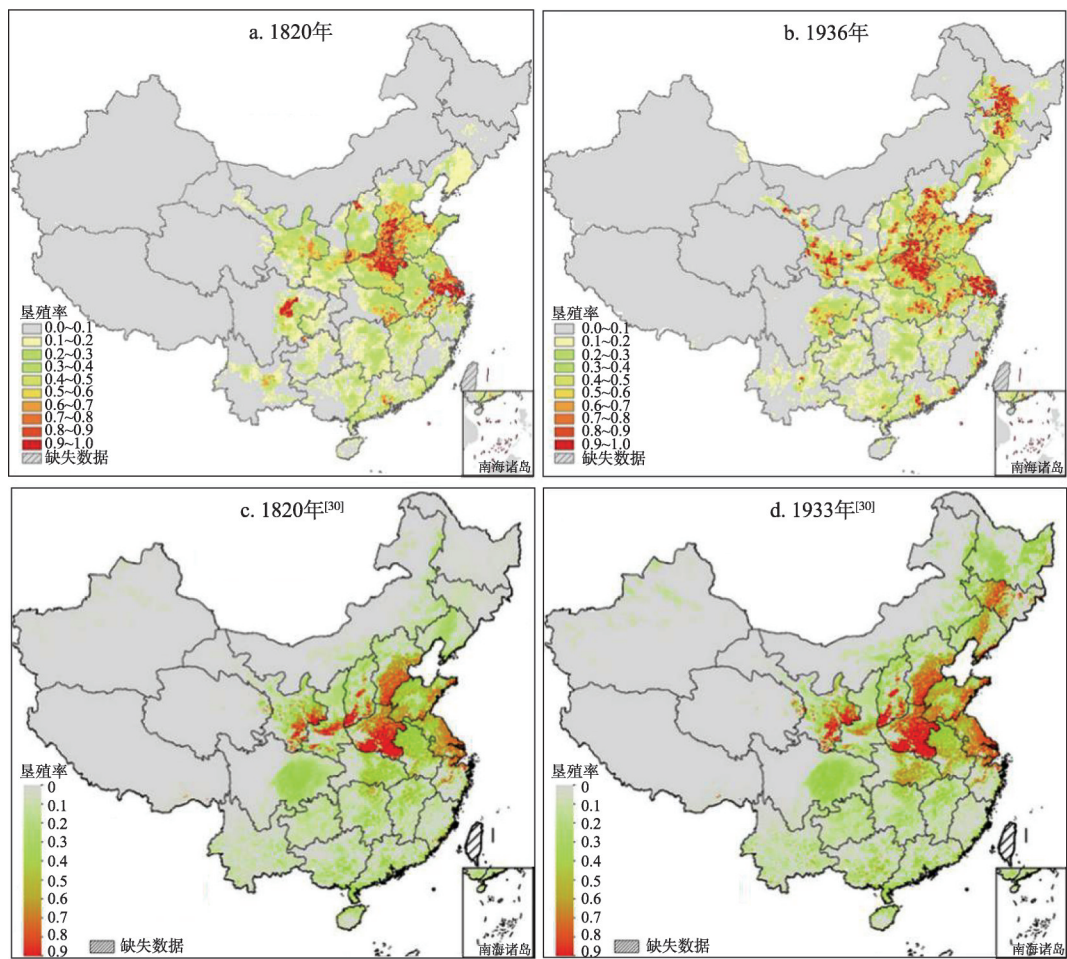


图3 本文成果与其他研究成果^[30]的全国尺度对比验证

Fig. 3 Comparison of the reconstructed map with different results on national scale

图3a和图3b分别为本文1820年和1936年耕地重建数据（以下称为全国数据），图3c和3d分别为Li等^[30]1820年和1933年的耕地重建数据，以下称为对比数据。从图3a与3c、图3b与3d相同年份的的总体对比来看，全国数据中除了甘肃中部以外，其他耕地分布地区都和对比数据保持着较高的匹配度；而从局部上来看，1820年的全国数据中云贵高原中部和甘肃中部耕地分布范围较对比数据广，而华北平原中部分布数量较对比数据少，并且长江入海口处分布数量对比数据多；而1936年的全国数据同样在甘肃中部分布范围相对对比数据广，并且华北平原在分布数量上较之对比数据有所减少，东北平原北部的耕地分布数量有所偏高。

综上所述，对比结果可以总结为：总体趋势基本保持一致，但是局部数量有所差别。究其原因，要归结到人口因子上，全国数据耕地分布较广的地方，是因其有人口的分布，而Li等^[30]的数据是以1980年的耕地格局为依据，他认为1980年的耕地数量分布已经到达最大值，这样就忽略了耕地随历史的变化而变化，这种变化是增减不定的。而数量上的差异，是因为人口的聚集导致了耕地的聚集。在全国的大尺度上，数量的局部差异对于整体的重建或者后期的模拟输入不会造成太大的影响，这样可以认为本文的重建数据是合理有效的。

4 结论

本文提出了“分层分区”的网格化方法来重建历史时期耕地格局,借助于其他学者的耕地数据,得出了全国范围内1820年和1936年分辨率为10 km×10 km的耕地分布图。重建结果表明,从道光年间到民国期间,全国耕地开发显著,特别是东北地区,并且西南地区因为清民时期“改土归流”的实施和移民潮的涌现使农耕有了明显的发展,而耕地的分布模式也从之前的聚集式分布逐渐扩散。从本文重建数据与其他学者数据的对比验证中可以看出:重建结果在全国整体趋势上匹配得很好,而在小区域上会有一定的波动,这是由于分配中因子选择和模型实施的差异性所导致。

总的来看,本文针对全国历史时期耕地分布设计的历史耕地网格化方法,不仅在总体上保证了耕地数量与其他学者研究成果的一致性,又考虑了区域差异对分配结果的影响;从细节上来看,分配过程分为人口和自然要素两步,人口因子主导县级以上尺度的分配,而自然因子主导网格化分配,因为人口主要决定耕种需求,在大尺度上可以作为耕地分布的区域指示,而在下一级的分配中考虑到耕种效率,就要根据区域范围内自然环境要素的优劣情况来耕种了。本文提出的方法,在保证精度与现今主流数据集差距不大的基础上,简化了分配过程,也从一定程度上解决了历史耕地重建过程中对史册资料过渡依赖的问题。但是在分配过程中由于与人口数据相匹配的耕地数据没有相关年份,从而利用其它年份来推断其数值,这样是有误差的。在以后的研究工作中,若能充分利用相关年份的史册、统计资料来对省级耕地数量进行校正,相信可以得到更加准确的耕地数据。并且,也可以在本文方法的基础上,借鉴其他方法来使其在验证与数据处理上更具准确性和科学性,比如Wang等^[47]所使用的Sandwich方法可以对于本文的验证部分具有很大的意义,其利用恰当的知识层生成多个报告层,用于交叉验证;Liao等^[48]利用的3G方法可以为本文以后的研究提供更加宽广的思路,包括GIS(地理信息系统)、GP(遗传编程)和GA(遗传算法)来构成主要结构,GP和GA用于将调查数据转化为栅格数据,而GIS用来鉴定和量化相关的影响因子。

本文重建数据与其他成果匹配度相当高,只存在局部部分区域的偏差,而这些偏差在省级准确耕地数据的保证下,对总体耕地的趋势分布影响很小,这体现了不同影响因子选择的差异性。利用此方法,可以较快并且准确地重建其他历史年代的耕地数据,为未来耕地预测或者相关模型提供基础数据支持。当重建成为一个连续的时间序列后,能更准确的反映历史土地变化/覆盖的变化情况,为研究古气候、古生物等各个领域提供有力的支持。

参考文献(Referances)

- [1] IGBP Report No.53/IHDP Report No.19. Global Land Project: Science Plan and Implementation Strategy. S Secreteriat, 2005: 1-64.
- [2] IPCC. Climate Change 2013: The Physical Science Basis. Switzerland, 2013: 1-30.
- [3] Ramankutty N, Foley J A. Characterizing patterns of global land use: An analysis of global croplands data. Global Biogeochemical Cycles, 1998, 12(4): 667-685.
- [4] Vitousek P M, Mooney H A, Lubchenco J, et al. Human domination of earth's ecosystems. Science, 2008: 3-13.
- [5] Brovkin V, Sitch S, Von Bloh W, et al. Role of land cover changes for atmospheric CO₂ increase and climate change during the last 150 years. Global Change Biology, 2004, 10(8): 1253-1266.
- [6] Li Xiubin. A review of the international researches on land use/land cover change. Acta Geographica Sinica, 1996, 51(6): 553-558. [李秀彬. 全球环境变化研究的核心领域: 土地利用/土地覆被变化的国际研究动向. 地理学报, 1996, 51(6): 553-558.]

- [7] Zhang Hua, Zhang Bo, Shi Huichun. Studies on land use and land cover change in arid regions: Taking Heihe River basin as an example. *Journal of Arid Land Resources and Environment*, 2003, 17(2): 49-54. [张华, 张勃, 石惠春. 干旱地区土地利用/土地覆盖变化研究: 以黑河流域为例. *干旱区资源与环境*, 2003, 17(2): 49-54.]
- [8] Rindfuss R R, Walsh S J, Turner B L, et al. Developing a science of land change: Challenges and methodological issues. *Proceedings of the National Academy of Science of the United States of America*, 2004, 101(39): 13976-13981.
- [9] Ge Quansheng, He Fanneng, Zhang Jingyun, et al. Some ideals related to the development of Chinese historical geography in the 21st century. *Geographical Research*, 2004, 23(3): 374-384. [葛全胜, 何凡能, 郑景云, 等. 21 世纪中国历史地理学发展的思考. *地理研究*, 2004, 23(3): 374-384.]
- [10] Voltaire A, Eickhout B, Schaeffer M, et al. Climate simulation of the twenty-first century with interactive land-use changes. *Climate Dynamics*, 2007, 29(2/3): 177-193.
- [11] Ellis E C, Klein Goldewijk K, Siebert S, et al. Anthropogenic transformation of the biomes, 1700 to 2000. *Global Ecology and Biogeography*, 2010, 19(5): 589-606.
- [12] He Fanneng, Li Meijiao, Xiao Ran. Comparisons of the land use and cover change in China and the United States over the past 300 years. *Acta Geographica Sinica*, 2015, 70(2): 287-297. [何凡能, 李美娇, 肖冉. 中美过去 300 年土地利用变化比较. *地理学报*, 2015, 70(2): 287-297.]
- [13] Zheng, Lin, He. Recent progress in studies on land cover change and its regional climatic effects over China during historical times. *Advances in Atmospheric Sciences*, 2009, 26(4): 793-802.
- [14] Zhu Feng, Cui Xuefeng, Miao Lijuan. China's spatially-explicit historical land-use data and its reconstruction methodology. *Progress in Geography*, 2012, 31(12): 1563-1573. [朱枫, 崔雪峰, 缪丽娟. 中国历史时期土地利用数据集的重建方法述评. *地理科学进展*, 2012, 31(12): 1563-1573.]
- [15] Li Shicheng, He Fanneng, Chen Yisong. Gridding reconstruction of cropland spatial patterns in Southwest China in the Qing Dynasty. *Progress in Geography*, 2012, 31(9): 1196-1203. [李士成, 何凡能, 陈屹松. 清代西南地区耕地空间格局网格化重建. *地理科学进展*, 2012, 31(9): 1196-1203.]
- [16] Miao L J, Zhu F, He B, et al. Synthesis of China's land use in the past 300 years. *Global and Planetary Change*, 2013, 100: 224-233.
- [17] Goldewijk K K, Beusen A, Dreht G V, et al. The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. *Global Ecology and Biogeography*, 2011, 20(1): 73-86.
- [18] Ramankutty N, Foley J A. Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles*, 1999, 13(4): 997-1027.
- [19] Pongratz J, Reick C, Raddatz T, et al. A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles*, 2008, 22: GB3018. doi: 10.1029/2007GB003153.
- [20] Ge Quansheng, Dai Junhu, He Fanneng, et al. Spatiotemporal dynamics of reclamation and cultivation and its driving factors in parts of China during the last three centuries. *Progress in Natural Science*, 2003, 13(8): 825-832. [葛全胜, 戴君虎, 何凡能, 等. 过去 300 年中国部分省区耕地资源数量变化及驱动因素分析. *自然科学进展*, 2003, 13(8): 825-832.]
- [21] He Fanneng, Ge Quansheng, Dai Junhu, et al. Quantitative analysis on forest dynamics of china in recent 300 years. *Acta Geographica Sinica*, 2007, 62(1): 30-40. [何凡能, 葛全胜, 戴君虎, 等. 近 300 年来中国森林的变迁. *地理学报*, 2007, 62(1): 30-40.]
- [22] Lin Shanshan, Zheng Jingyun, He Fanneng. The approach for gridding data derived from historical cropland records of the traditional cultivated region in China. *Acta Geographica Sinica*, 2008, 63(1): 83-92. [林珊珊, 郑景云, 何凡能. 中国传统农区历史耕地数据网格化方法. *地理学报*, 2008, 63(1): 83-92.]
- [23] Ye Yu, Fang Xiuqi, Dai Yujuan, et al. Date assimilation and cultivation ratio reconstruction of three provinces of Northeast China. *Progress in Natural Science*, 2006, 16(11): 1419-1427. [叶瑜, 方修琦, 戴玉娟, 等. 东北 3 省民国时期耕地数据的同化与垦殖率重建. *自然科学进展*, 2006, 16(11): 1419-1427.]
- [24] Ye Yu, Fang Xiuqi, Ren Yuyu, et al. Coverage changes of cropland in northeastern China during the past 300 years. *Science in China: Earth Sciences*, 2009, 39(3): 340-350. [叶瑜, 方修琦, 任玉玉, 等. 东北地区过去 300 年耕地覆盖变化. *中国科学: 地球科学*, 2009, 39(3): 340-350.]
- [25] Ye Yu, Fang Xiuqi, Zhang Xuezheng, et al. Coverage changes of forestland and grassland in northeastern China during the past 300 years. *Journal of Beijing Forestry University*, 2009, 31(5): 137-144. [叶瑜, 方修琦, 张学珍, 等. 过去 300 年东北地区林地和草地覆盖变化. *北京林业大学学报*, 2009, 31(5): 137-144.]
- [26] Li Ke, He Fanneng, Zhang Xuezheng. An approach of reconstructing historical cropland spatial distribution with grid-boxes by utilizing MODIS land cover dataset: A case study for Yunnan Province in the Qing Dynasty. *Geographical*

- Research, 2011, 30(12): 2281-2288. [李柯, 何凡能, 张学珍. 基于MODIS数据网格化重建历史耕地空间分布的方法: 以清代云南省为例. 地理研究, 2011, 30(12): 2281-2288.]
- [27] Long Y, Jin X B, Yang X H, et al. Reconstruction of historical arable land use patterns using constrained cellular automata: A case study of Jiangsu, China. *Applied Geography*, 2014, 52: 67-77.
- [28] Li Shicheng, Zhang Yili, He Fanneng. Reconstruction of cropland distribution in Qinghai and Tibet for the past one hundred years and its spatiotemporal changes. *Progress in Geography*, 2015, 34(2): 197-206. [李士成, 张懿铨, 何凡能. 过去百年青海和西藏耕地空间格局重建及其时空变化. 地理科学进展, 2015, 34(2): 197-206.]
- [29] Liu M, Tian H. China's land cover and land use change from 1700 to 2005: Estimations from high-resolution satellite data and historical archives. *Global Biogeochemical Cycles*, 2010, 24: GB3003. doi: 10.1029/2009GB003687.
- [30] Li S C, He F N, Zhang X Z. A spatially explicit reconstruction of cropland cover in China from 1661 to 1996. *Regional Environmental Change*, 2015: 1-12.
- [31] He Fanneng, Li Shicheng, Zhang Xuezheng, et al. Comparisons of reconstructed cropland area from multiple datasets for the traditional cultivated region of China in the last 300 years. *Acta Geographica Sinica*, 2012, 31(9): 1190-1200. [何凡能, 李士成, 张学珍, 等. 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, 2012, 31(9): 1190-1200.]
- [32] Feng Yongheng, Zhang Shihuang, He Fanneng, et al. Separate reconstruction of Chinese cropland grid data in the 20th century. *Progress in Geography*, 2014, 33(11): 1546-1555. [冯永恒, 张时煌, 何凡能, 等. 20世纪中国耕地格网化数据分区重建. 地理科学进展, 2014, 33(11): 1546-1555.]
- [33] Deng Ai. *The Eco-economy Research of Qinghai-Tibet Plateau Pastoral Areas*. Beijing, China: Nationalities Publishing House, 2005: 101-145. [邓艾. 青藏高原草原牧区生态经济研究. 北京: 民族出版社, 2005: 101-145.]
- [34] Bao Yushan. Nomadism culture and farming culture: Collision, result, rethink: The meaning of cultural survival and equality. *Social Science Front*, 2007(4): 241-246. [包玉山. 游牧文化与农耕文化: 碰撞·结果·反思: 文化生存与文化平等的意义. 社会科学战线, 2007(4): 241-246.]
- [35] Wang Ling, Xie Xianqun, Li Yunsheng, et al. Changes of humid index and border line of wet and dry climate zone in northern China over the past 40 years. *Geographical Research*, 2004, 23(1): 45-54. [王菱, 谢贤群, 李运生, 等. 中国北方地区40年来湿润指数和气候干湿带界线的变化. 地理研究, 2004, 23(1): 45-54.]
- [36] Cao Shuji. *Population History of China (Vol. 5): Qing Dynasty Period*. Shanghai: Fudan University Press, 2001. [曹树基. 中国人口史(第五卷): 清时期. 上海: 复旦大学出版社, 2001.]
- [37] Wei Xiwen, Cui Xuefeng. Construction of cultivation land suitability index for China based on climate and soil properties. *Journal of Beijing Normal University (Natural Science)*, 2016, 52(1): 63-69. [魏希文, 崔雪锋. 基于气候和土壤要素的中国耕地适宜性指数构建. 北京师范大学学报(自然科学版), 2016, 52(1): 63-69.]
- [38] Gao Xiang, Liu Fengyun. The discussion of farming and nomadic: Relationship and development tendency. *Journal of Renmin University of China*, 1995(1): 78-85. [高翔, 刘凤云. 论古代中国北方的农耕与游牧: 对两种文化关系及发展趋势的探讨. 中国人民大学学报, 1995(1): 78-85.]
- [39] Luo Jing, Chen Qiong, Liu Fenggui, et al. Methods for reconstructing historical cropland spatial distribution of the Yellow River-Huangshui river valley in Tibetan Plateau. *Progress in Geography*, 2015, 34(2): 207-216. [罗静, 陈琼, 刘峰贵, 等. 青藏高原河谷地区历史时期耕地格局重建方法探讨: 以河湟谷地为例. 地理科学进展, 2015, 34(2): 207-216.]
- [40] Zhang Jingming. The catering culture research of the nomadic people in north China [D]. Beijing: Renmin University of China, 2004. [张景明. 中国北方游牧民族饮食文化研究[D]. 北京: 中央民族大学, 2004.]
- [41] Liang Yuetian. Wheat variety and accumulated temperature. *Experiment of Agricultural Science*, 1980(7): 10-12. [梁月天. 小麦品种与积温. 农业科学实验, 1980(7): 10-12.]
- [42] Xue Dawei, Fang Maoting, Qian Qian. Application of effective accumulative temperature in rice production. *China Rice*, 2004(4): 47-48. [薛大伟, 方茂庭, 钱前. 有效积温在水稻生产中的应用. 中国稻米, 2004(4): 47-48.]
- [43] Xu Chengzhong, Dong Xingyu, Yang Hongbin, et al. The impact of changes of accumulated temperature to summer maize-winter wheat rotation system. *Shandong Agricultural Sciences*, 2009(2): 34-37. [徐成忠, 董兴玉, 杨洪宾, 等. 积温变迁对夏玉米冬小麦两熟制播期的影响. 山东农业科学, 2009(2): 34-37.]
- [44] Zheng Yingze. The impact of soil pH to crop growth. *Journal of Yuxi Normal University*, 1994, 10(3-4): 64-67. [郑应泽. 土壤酸碱度对农作物生长的影响. 玉溪师专学报, 1994, 10(3-4): 64-67.]
- [45] Li B B, Fang X Q, Ye Y, et al. Accuracy assessment of global historical cropland datasets based on regional reconstructed historical data: A case study in Northeast China. *Science China Earth Sciences*, 2010, 53(11): 1689-1699.
- [46] Yang X H, Guo B B, Jin X B, et al. Research on reconstructing spatial distribution of historical cropland over 300 years

in traditional cultivated regions of China. *Global and Planetary Change*, 2015, 128: 90-102.

- [47] Wang J F, Hai N R, Liu T J, et al. Sandwich estimation for multi-unit reporting on a stratified heterogeneous surface. *Environment and Planning*, 2013, 45(10): 2515-2534.
- [48] Liao Y L, Wang J F, Meng B, et al. Integration of GP and GA for mapping population distribution. *International Journal of Geographical Information Science*, 2010, 24(1): 47-67.

Reconstruction of the historical crop land data in China using the partition and layering-based gridded method

WEI Xiwen¹, MIAO Lijuan², JIANG Yuan¹, NATH Reshmita³, CUI Xuefeng⁴

(1. College of Resources Science and Technology, Beijing Normal University, Beijing 100875, China;
2. School of Geography and Remote Sensing, Nanjing University of Information Science & Technology, Nanjing 210044, China; 3. Center for Monsoon System Research, Institute of Atmospheric Physics, CAS, Beijing 100029, China; 4. School of System Science, Beijing Normal University, Beijing 100875, China)

Abstract: Reconstruction of the historical cropland database could provide an accurate support for the cropland researches on historical periods. Moreover, it will also act as the driving data for the global climate models. In China, there are several ways to rebuild the historical cropland on the regional scale, concerning the traditional cultivation area and the northeastern China. But there are few studies on the national scale, which is primarily due to lacking of accurate historical cropland inventory data. Therefore, it is necessary to simulate the climate and ecosystem scenarios in a wider spatial scale. In the next step, an improvement in the research on reconstruction of the historical cropland database is an efficient and quick method to meet the demand. The present paper has adopted a different approach, which divides the entire landmass of China into cropland area and livestock area according to the climate conditions and cultural diversities of the country. Several representative historical cropland databases are used as the background of our study to ensure the authenticity of the outcomes. Considering the historical land use exploitation, and the impact of the natural and anthropogenic factors on the cropland dynamics, we have developed a partition and layering-based gridded reconstruction method. This is an empirical method for spatial allocation of the historical cropland inventory data in the individual grid cells. Then we use this method to rebuild the historical cropland database maps for China during 1820 (the 25th year in the reign of Emperor Renzong Jiaqing in the Qing Dynasty) and 1936 (the 25th year of the Republic of China), at a resolution of 10 km×10 km. A precision validation is applied on the rebuilt maps by comparing them with other existing regional databases. The results show that this method could ensure the authenticity of the cropland data, with higher accuracy in the regional cropland reconstruction.

Keywords: China; historical cropland; partition and layering; gridded method