

# Analysis of the Impact Factors of Cropland Distribution and Model Building in Jiangsu and Anhui for Historical Land Use Research

Cun Yuan<sup>1</sup>, Yu Ye<sup>1,2</sup>, Beibei Li<sup>3</sup>, Xueqiong Wei<sup>1</sup>, Xiuqi Fang<sup>1,2</sup>

<sup>1</sup>School of Geography, Beijing Normal University, Beijing, China

<sup>2</sup>Key Laboratory of Environment Change and Natural Disaster, Ministry of Education, Beijing Normal University, Beijing, China

<sup>3</sup>School of Languages and Cultures, Nanjing University of Information Science & Technology, Nanjing, China

Email: [anguoyuancun@sohu.com](mailto:anguoyuancun@sohu.com), [yeyuleaffish@bnu.edu.cn](mailto:yeyuleaffish@bnu.edu.cn), [libeibei-sd@163.com](mailto:libeibei-sd@163.com),  
[xueqiong.wei@mail.bnu.edu.cn](mailto:xueqiong.wei@mail.bnu.edu.cn), [xfang@bnu.edu.cn](mailto:xfang@bnu.edu.cn)

Received: May 3<sup>rd</sup>, 2014; revised: Jun. 2<sup>nd</sup>, 2014; accepted: Jun. 13<sup>th</sup>, 2014

Copyright © 2014 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Accurate historical gridded data of land use are the base for the environmental effect simulation of the land use/land cover change. Because of lacking of sufficient credible historical map data, historical grid data couldn't access directly. It is usually assigned downscaling based on the modern distribution pattern of cropland and its driving factors. This paper analyzed the relationship between spatial distribution of cropland and climate, slope, elevation, soil, river and lake, population based on modern remote sensing data in Jiangsu and Anhui. The most key factors in Jiangsu and Anhui are river and elevation. The area that river flows through doesn't distribute cropland. The correlation coefficient of slope and reclamation rate is as high as  $-0.628$ ; the highest elevation in Jiangsu and Anhui is under the upper limit elevation that can be cultivated because of climate. The paper also constructed the cropland spatial distribution model with spatial resolution of  $10\text{ km} \times 10\text{ km}$ , and assessed the model. It is suggested that the reconstruction data and the remote sensing data are not different obviously from the correlation coefficient or standard deviation. The correlation coefficient is  $0.821$  and the standard deviation is  $20.6$ . This model can be used in cropland reconstruction in Jiangsu and Anhui, and provide references for historical cropland gridded data.

## Keywords

Factors Analysis, Grid Reconstruction, Cropland, Model Building, Jiangsu and Anhui

---

# 面向历史土地利用研究的苏皖地区耕地分布的影响因子分析及宜垦性模型构建

袁存<sup>1</sup>, 叶瑜<sup>1,2</sup>, 李蓓蓓<sup>3</sup>, 魏学琼<sup>1</sup>, 方修琦<sup>1</sup>

<sup>1</sup>北京师范大学地理学与遥感科学学院, 北京, 中国

<sup>2</sup>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, 中国

<sup>3</sup>南京信息工程大学语言与文化学院, 南京, 中国

Email: [anguoyuancun@sohu.com](mailto:anguoyuancun@sohu.com), [yeyuleaffish@bnu.edu.cn](mailto:yeyuleaffish@bnu.edu.cn), [libeibei-sd@163.com](mailto:libeibei-sd@163.com),

[xueqiong.wei@mail.bnu.edu.cn](mailto:xueqiong.wei@mail.bnu.edu.cn), [xfang@bnu.edu.cn](mailto:xfang@bnu.edu.cn)

收稿日期: 2014年5月3日; 修回日期: 2014年6月2日; 录用日期: 2014年6月13日

## 摘要

精确的格点化区域历史土地利用数据是进行历史土地利用/土地覆被变化环境效应模拟的基础。由于缺少充足可信的历史图件等资料, 历史网格化数据无法直接获取, 通常是依据现代耕地分布格局及其形成的驱动因素, 对某一区域的历史耕地总和数据进行降尺度的格点化分配。本文以苏皖地区为例, 基于现代遥感数据分析了耕地空间分布与气候、坡度、海拔高度、土壤、河流湖泊、人口等因子的关系, 发现影响苏皖地区耕地分布最为关键的因素是河湖及坡度。河流湖泊地区为非耕地的剔除区域, 坡度与垦殖率的相关系数高达-0.628, 苏皖地区海拔未达到气候影响下的耕地垦殖上限。在此基础上构建了耕地空间分布模型, 对现代的耕地格局进行10 km × 10 km的模型重建, 并进行了模型评估检验。检验结果显示, 无论是从均值还是从相关系数或标准偏差来看, 宜垦性模型的空间分配结果与现代遥感数据的空间格局差异均不大, 两者的相关系数 $R$ 为0.798, 标准偏差 $S$ 为22.2, 两者差值的均值为0.09。该现代宜垦性模型能够用于耕地格局重建, 可为该区历史时期耕地数据网格化处理提供参考。

## 关键词

因子分析, 网格化重建, 耕地, 模型构建, 苏皖

## 1. 引言

在全球变化研究中, 历史土地利用/土地覆被研究受到广泛的关注。在定量评估模拟历史土地利用/土地覆被的气候效应时, 往往需要将基于行政单元的历史重建数据转化为具有空间属性的网格化数据, 以满足模式模拟的国际化需求。

国内外面向全球变化研究的历史耕地覆盖变化数据集成果, 采用的耕地网格化方法及因子选择不同, 给区域尺度耕地覆盖重建结果带来一定误差。SAGE 数据集通过收集历史耕地数据, 并假设历史耕地分布格局与现代耕地(1992年)分布格局相同, 根据1992年遥感土地覆盖分类数据集将历史耕地垦殖率赋值分布在地图中, 获得具有地理空间属性的历史耕地数据[1] [2]。PJ数据集对SAGE数据集1700~1992年耕地和草地数据进行了调整和补充, 重建了800~1992年土地覆盖数据, 并做出两个方面的假设: 第一, 1700年以前各国人均耕地、牧草地均不变; 第二, 土地利用空间分布格局不变。在此基础上利用历史人口数据重建了800~1700年的耕地和牧草地数据, 进而通过与Ramankutty and Foley的植被分类图进行叠

加分析,最终得到空间分辨率为 $0.5^{\circ} \times 0.5^{\circ}$ 的土地利用数据[3]。HYDE数据集是全球尺度上最具代表性的土地数据集之一,自发布以来不断更新。最初的HYDE2.0认为人口是土地覆盖分配的主要驱动力,人口的变化通过影响对粮食、服务等的需求而影响到森林砍伐等土地覆盖变化[4]。因此选取的主要影响因子为人口密度。最新版本HYDE3.1在其基础上又做了改进[5],提出时间依赖的空间分配算法,该算法选取了包括城镇面积、人口密度、土壤宜耕性、河流、坡度及年平均温度等社会、自然环境六大因子进行耕地面积的空间分配[6]。这些数据集均具有明确的地理属性,可以作为基础变量输入土地利用变化及其环境效应的模型中,目前已广泛应用于全球或区域尺度土地利用环境效应的定量评估模型中[7]-[10]。但李蓓蓓等[11] [12]、何凡能等[13]、Zhang等[14]的基于中国不同区域的历史耕地重建研究表明, HYDE数据集在区域尺度上仍存在一定的误差。

在国内,林珊珊等[15]探讨了中国传统农区历史耕地数据格网化方法,选取了地形和人口作为耕地分布的影响因子,并进行了量化处理。此后,该网格化方法得到了广泛的应用和改进,如何凡能等[16]采用该方法重建了北宋中期境内耕地的空间分布格局;李柯等[17]以地表高程和坡度为影响土地宜垦程度的主导因子,评估了MODIS土地覆盖产品中现代耕地分布区域的宜垦程度,将历史文献资料订正的云南省1671年和1827年县域耕地统计数据分配至 $90\text{ m} \times 90\text{ m}$ 的网格中;李士成等[18]以现代耕地分布格局为基础,量化地形(海拔、坡度)、气候生产潜力(光照、温度、水分)和人口密度等因子,构建土地宜垦性算法模型,重建了清代西南地区6个时点分辨率为 $10\text{ km} \times 10\text{ km}$ 的耕地数据。Liu和Tian[19]采用二分迭代法重建了中国历史耕地、林地和建设用地的分布格局。颜耀文等[20]根据黑河流域历史时期土地开发利用的特点,深入分析了海拔、坡度、土壤、人口等因子与耕地分布的关系,结合古遗址、遥感影像判别的结果,进行了明代以前耕地的空间分布重建。然而,不同地区的耕地开垦过程、性质是不同的,影响耕地开垦的因素并非完全相同。即使是同一地区,在不同的空间尺度上,影响因素也不尽相同。

本文拟基于现代遥感耕地数据,从改进网格化重建方法、深入分析中小尺度区域耕地空间分布格局变化的角度,探讨苏皖地区耕地分布和拟遴选因子的关系,构建土地宜垦性模型进行现代耕地格局重建,并对模型模拟误差进行检验和分析,为重建历史较高空间分辨率和精度的土地覆被数据提供参考。

## 2. 数据来源与研究方法

### 2.1. 研究区概况

苏皖地区是中国的传统农区,是受自然因素和人文驱动因素共同影响较为活跃的热点地区。曾合称江南省,是著名的“鱼米之乡”,清朝康熙六年(公元1667年)拆江南省为江苏、安徽两省而正式建省。位于长江三角洲,以平原丘陵为主,地势低平,海拔多在 $50\text{ m}$ 以下。水系发达,河湖较多(图1)。气候属于暖温带向亚热带的过渡气候,四季分明,春暖、夏炎、秋爽、冬寒。人口众多,据2010年第六次全国人口普查数据显示苏皖共有人口13816万人[21]。

江苏省水田占60.06%,主要集中在太湖、里下河、镇扬丘陵地区的老稻区,约占全省水田总面积的55%左右,旱地主要分布在徐州、淮阴、沿海和沿江地区,约占全省旱地总面积的90%左右[22]。安徽省水田主要遍布于沿江和江淮丘陵一带,旱地主要分布于淮北平原,约占全省旱地面积的80%,其次在江淮丘陵一带。从行政区域分布来看:阜阳、六安、宿县、滁州的耕地面积约占全省耕地面积的55.26%[23]。

江苏省面积 $10.26$ 万平方公里,现有1130个乡镇;安徽省面积 $13.96$ 万平方公里,乡镇级区划单位共计1399个[24]。苏皖地区乡镇平均规模约为 $100\text{ km}^2$ ,因此本文在进行研究时选取的网格大小为 $10\text{ km} \times 10\text{ km}$ ,与苏皖地区乡镇平均规模大体一致。

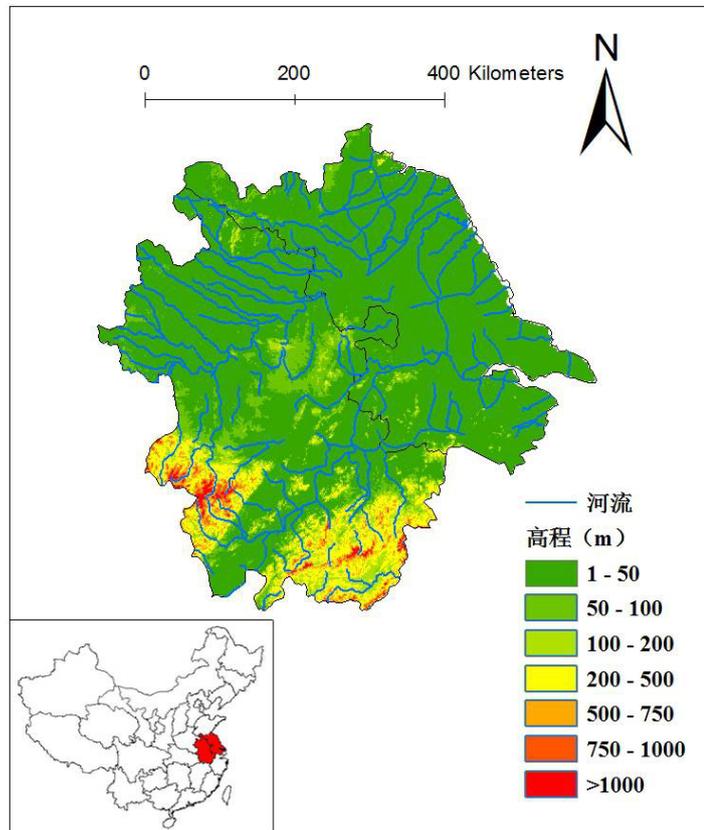


Figure 1. The location and topography of Jiangsu and Anhui  
图 1. 苏皖地区位置及其地形

## 2.2. 数据来源

本文采用的河流湖泊数据来自于中国 1: 400 万全要素基础数据(1970s~1990s), 90 m DEM 数据和 90 m 坡度数据来自于地理空间数据云(<http://www.gscloud.cn/>), 土地利用数据、土壤碳储量数据、土壤氮储量数据以及人口数据等均来自于地球系统科学数据共享平台(<http://www.geodata.cn/Portal/index.jsp>)。这些数据具有不同的空间分辨率, 在使用之前将其在 Arcgis 中进行重采样, 转化为 10 km 网格大小的数据。

## 2.3. 研究方法

苏皖地区遥感数据显示, 1980 年代耕地分布数量最多, 此后耕地数量逐渐减少。因此本文假设 1980 年代耕地的分布范围最为广泛。选取一些对耕地分布影响较大, 较为稳定, 较易量化的因子, 采用统计法、相关分析法探讨其与 1980 年代耕地空间分布的关系; 以 1980 年代苏皖的耕地分布为例采用数学模型方法进行网格化模型重建; 以 2000 年的耕地遥感数据为例, 采用相关分析、方差分析、 $t$  检验分析进行模型的验证。

### 2.3.1. 影响耕地分布的因子

土地资源是指已经被人类所利用和可预见的未来能被人类利用的土地, 具有自然属性和社会属性。耕地是受人类活动影响最大的土地类型之一, 其分布必然会受到自然因素和社会因素的共同影响。其中, 自然因素是影响耕地分布的基础, 也是较易量化的。在众多影响耕地分布的因子中, 选取的因子越多, 工作量越大, 网格化模型越复杂。如果能用较少的因子作为代表是最为理想的。因此, 在选取因子时,

首先要选择主导因子，即对耕地的分布具有较大影响的因子；其次，由于耕地资源的分布具有历史延续性和继承性[20]，影响因子需在长时间序列中具有相对稳定性；第三，最好采用国家标准或者行业标准的已有数据资料；第四，因子数据简单、易量化。

基于上述提到的几个原则，本文选取气候条件、河流湖泊、地形、土壤四方面的自然因子和人口等社会因子进行分析，遴选影响现代垦殖率的主要因子进行土地宜垦性模型构建。

#### 1) 自然因素

##### a) 非陆地剔除

在现行的土地利用现状分类(GB/T 21010-2007)中，耕地和水域及水利设施用地是并列的两大不同类型的土地利用类型，也就是说，水域及水利设施用地(包括河流、湖泊、水库和滩涂等)是不分配耕地的。苏皖地区河湖密布，尤其是江苏省，水面积达 1.67 万平方公里，占全省总面积的 16.3% [25]，在进行耕地网格化分配时，尤其要剔除该区域的水域面积。

##### b) 气候适宜性

气候是决定区域农业发展的最基本条件，对土地气候资源的利用起着促进或抑制的作用[26]。通常情况下，较好的光照、温度、降水空间组合，有利于农作物的生长，土地优先得到开垦。因此，气候条件对耕地的分配具有直接的作用。但是，在苏皖区域尺度下，气候因子是否限制作物的种植进而影响耕地的开垦，或者气候仅仅影响农作物的种植品种、熟制、产量等，需要进一步的分析。本研究基于农作物的种植条件来分析该区域气候适宜性。

##### c) 高程和坡度

在适合农业开垦的气候适宜区域内，高程和坡度也是限制耕地分配的重要影响因子。耕地最先分布于坡度较缓的平原地区，坡度较陡的山区分布耕地较少。中国科学院自然资源综合考察委员会调查资料[27]中提到，耕地按照地面坡度大小可以分为平耕地(坡度 $\leq 2^\circ$ )和坡耕地(坡度 $> 2^\circ$ )，坡耕地又可以分为缓坡耕地( $2^\circ \sim 6^\circ$ )、坡耕地( $6^\circ \sim 15^\circ$ )、陡坡耕地( $15^\circ \sim 25^\circ$ )和很陡坡耕地( $> 25^\circ$ )。一般坡度小于  $3^\circ$  的土地基本无水土流失现象，适宜开垦； $3^\circ \sim 15^\circ$  的土地水土流失现象较普遍，但采取措施后，仍适宜开垦； $25^\circ$  是开垦的极限，超过开垦，会引起严重的水土流失，很少用于开垦农用地。一般来说，在地形平坦的地区，越有利于耕地连片耕作，便于农田水利化和机械化，土地会被优先开垦。地形起伏较大，坡度较陡的地区，不仅交通不便，而且灌溉困难，严重者还会造成水土流失，不利于农业生产活动。

高程和坡度等地形因素不仅直接影响到耕地的分布，还通过区域的水热分布状况间接影响农业生产。海拔高度每上升 100 m，气温降低  $0.5^\circ\text{C} \sim 0.6^\circ\text{C}$  [28]。地形还能影响降水的分布：迎风坡降水量多，背风坡降水量少；山地降水量在一定高度内随海拔的升高而增大到最大降水量，超过此高度，山地降水不再随高度递增大。这种水热组合的差异造成的垂直地带性一方面在自然景观上产生差异，另一方面在一定条件下会限制作物的生长，因此高程和坡度因素是必须要考虑的。

##### d) 土壤适宜性

即使具备了适宜的气候条件，地形条件外，如果土壤条件不适宜，耕地开垦也不能顺利进行。土壤是地球陆地表面具有肥力能够生长植物的疏松层[29]，是绿色植物生长的自然环境，为光合作用提供并协调营养因素(温度、空气)和环境条件(水分、养分)。虽然农作物生长所必需的基本因素——光能和热量主要来自太阳辐射，空气来自大气，但是，水分和养分却主要通过根系从土壤中吸收。在适宜的土壤中种植作物，作物才能生存，才能实现耕地的开垦。土壤碳储量和氮储量作为土壤肥力的一个重要标志，能够反映土壤有机质的含量，是土壤质量的关键与核心，对农业生产和农业生态环境有着重要的指示作用。本研究选用土壤碳氮储量分布图，其所用数据是从中国土种志(1~6册)按省份、植被类型提取土壤层次、剖面数、厚度、碳氮含量，从 134 篇中文期刊文选中按地点、植被类型提取土层深度数据和土壤有

机质或土壤养分数据[30]。这些数据计算的土壤碳储量和土壤氮储量两个指标能够反映土壤的适宜性。

## 2) 社会因素

社会因子主要包括人口、技术、政策、战乱等。诚如林珊珊[15]所言,人口与耕地面积及分布格局的关系最为密切:历史时期的中国农业社会,增加耕地面积一直是传统农业社会中解决粮食问题的一个重要方式[31],耕地面积的大小主要取决于劳动力的多少,人口密度更能代表耕地开垦过程和空间分布;政策和战乱一方面难以量化,另一方面,从内在机制上讲人口数量的变化也间接反映了移民等国家政策的变化。

### 2.3.2. 因子选取与宜垦性模型构建

分析该区域气候因素对农作物耕种的适宜性影响,将坡度与高程、土壤适宜性各要素、人口要素分别与耕地垦殖率进行相关分析,并统计不同坡度不同海拔高度下耕地的分布情况和垦殖率的变化情况,对苏皖地区非陆地区域的垦殖率进行统计分析,剔除河流湖泊等对垦殖率的影响。得到影响耕地分布的关键因子。

分别量化各影响因子,对坡度进行标准化处理,剔除非陆地因子,建立宜垦性数学模型。

## 3. 影响耕地分布的因子分析与选取

### 1) 自然因素

#### a) 非耕地剔除

本文对苏皖地区河湖分布区的垦殖率进行了统计,结果显示:河湖分布区共计网格 106 个,其中,垦殖率为 0 的有 75 个,占比 70.8%,垦殖率大于 50%的有 16 个网格,占比 15.1%。由此可以看出,河流湖泊是影响耕地的主要因子之一。但是在计算的过程中,可能出现湖泊并未布满整个网格的情况,如果按照全部布满网格计算会带来较大的误差。在实际进行耕地分配的过程中,需要考虑河流湖泊占网格的比例关系,减小误差。

HYDE 数据集认为,越靠近水源的地区,越有利于农业的发展。但是,在苏皖地区,多数地区年均降水量在 800 mm 以上[25] [32],即使远离河流湖泊的地方,也可以靠降水补给水分。因此,水分条件不会成为该区农业发展的限制因子。苏皖地区河流湖泊对耕地的限制,仅表现在不同土地利用类型的差异上。

#### b) 气候适宜性

苏皖地区主要的粮食作物有水稻、小麦、玉米、大豆等[33]-[35]。其中,水稻生长对光热水的要求比较高,华北平原水稻种植 $\geq 10^{\circ}\text{C}$ 积温为  $3500^{\circ}\text{C}\sim 4500^{\circ}\text{C}$ ,日平均气温稳定在  $10^{\circ}\text{C}\sim 12^{\circ}\text{C}$ ,水稻才能安全播种(4月中旬),日平均气温稳定在  $20^{\circ}\text{C}\sim 22^{\circ}\text{C}$ 以上才能安全齐穗(8月下旬~9月下旬),日平均气温大于等于  $15^{\circ}\text{C}$ ,水稻才能通过安全成熟期(10月上中旬)[36],安徽省  $10^{\circ}\text{C}$ 积温在  $4600^{\circ}\text{C}\sim 5300^{\circ}\text{C}$ ,4月平均气温  $14^{\circ}\text{C}\sim 16^{\circ}\text{C}$ ,8月平均气温在  $26^{\circ}\text{C}\sim 28^{\circ}\text{C}$ 左右,9月平均气温在  $21^{\circ}\text{C}\sim 24^{\circ}\text{C}$ 左右,10月平均气温  $15^{\circ}\text{C}\sim 18^{\circ}\text{C}$  [23];江苏省  $10^{\circ}\text{C}$ 积温为  $4400^{\circ}\text{C}\sim 5000^{\circ}\text{C}$ ,春季(3月下旬起)月平均温度保持在  $10^{\circ}\text{C}\sim 22^{\circ}\text{C}$ 之间,夏季月均温为  $30^{\circ}\text{C}$ 左右,9月中下旬以后,日均温逐渐降至  $22^{\circ}\text{C}$ 以下[22]。该区夏季多雨,刚好与水稻生长所需大量的水分相吻合,再加上该区水资源丰富,农田灌溉条件优越,水分条件能够满足水稻的种植。由此,苏皖广大平原地区能够满足水稻种植的光热水条件,气候因素并不能限制平原地区耕地的开垦。

#### c) 高程与坡度

将土地垦殖率分别与海拔高度和坡度作相关分析(表 1),发现土地垦殖率与海拔高度和坡度均具有很好的相关性,在 0.01 水平上是显著相关的,且垦殖率与坡度的相关性要高于与海拔高度的相关性。而海拔高度和坡度两个之间本身就具有很好的相关性(相关系数为 0.869),为了更好的说明哪个因子对垦殖率

的影响更大,有必要进行偏相关分析(表 1)。在控制坡度的情况下,垦殖率和海拔高度的相关系数仅为-0.006,显著性水平为 0.764,可靠性大大降低;在固定海拔高度的情况下,垦殖率和坡度的相关系数仅为-0.367,所以坡度是影响垦殖率的最主要地形因素。

尽管如此,海拔对农作物的分布上限是有限制的。玉米由于其适应性强,产量高,分布较为广泛。种植所需的热量条件为 $\geq 10$ 度积温 1800 $^{\circ}\text{C}$ ~2700 $^{\circ}\text{C}$ ,生长期的降水量为 500~800 mm。其种植的垂直分布受海拔高度和地形的影响。在我国玉米栽培的山地和云贵高原等山区,其种植上限可达 2700 m;华北北部和西部山区 $\geq 10$ 度积温递减率为 130 $^{\circ}\text{C}$ ~150 $^{\circ}\text{C}$ ,适宜种植高度为 300~900 m;亚热带西部山区玉米种植区 $\geq 10$ 度积温递减率为 150 $^{\circ}\text{C}$ ~250 $^{\circ}\text{C}$ ,种植高度为 1300~2300 m[36]。苏皖地区 $\geq 10$ 度积温 4500 $^{\circ}\text{C}$ [37],按照 150 $^{\circ}\text{C}$ 积温递减率粗略估算,在海拔为 1500 m 的地方, $\geq 10$ 度积温在 2200 $^{\circ}\text{C}$ 以上,仍能满足玉米的种植条件。而苏皖地区最高海拔不足 1500 m,因此该区域的海拔不能对农作物的种植机耕地的开垦产生限制。

苏皖地区,最高坡度为 31 $^{\circ}$ 。为了揭示坡度和耕地分布的关系,按照第二次全国土地调查技术规程中坡度分级方法,对耕地数量和耕地垦殖率进行了统计分析(表 2),发现耕地主要分布于平坦地区,占到了 95%,平坦地区垦殖率也非常高,达到了 74.21%,当然这和该区地势平坦的特点是密不可分的。并对很陡坡耕地中垦殖率的分布进行了统计,其中大于 25 $^{\circ}$ 的网格数有 15 个,垦殖率均小于 20%,垦殖率为 0 的网格数有 10 个,占比 66.7%;垦殖率为 0%~5%的网格数为 12 个,占比 80%。由此可以看出坡度越小,越适宜耕地耕种,在坡度陡的地方分布耕地很少,坡度是影响耕地分配的重要因子。

在剔除河流湖泊因子的影响下,垦殖率和坡度的相关系数为-0.702,在 0.01 水平上显著相关,并对不同垦殖率下的坡度和海拔高度进行了统计(表 3),说明耕地优先分配在坡度较小,海拔较低的地段,随

**Table 1. Correlation between reclamation rate, elevation and slope and partial correlation between elevation and slope**  
**表 1. 垦殖率与海拔高度、坡度的相关性及偏相关性**

	海拔高度	坡度	海拔高度(固定坡度)	坡度(固定海拔高度)
垦殖率	-0.548**	-0.628**	-0.006	-0.367
显著性水平(双侧)	0.000	0.000	0.764	0.000

\*\*在.01 水平(双侧)上显著相关。

**Table 2. Cropland distribution on different slopes**  
**表 2. 不同地面坡度与耕地分布的关系**

坡度分级	耕地数量(km <sup>2</sup> )	坡度( $^{\circ}$ )	占总耕地的百分比(%)	平均垦殖率(%)
平耕地( $\leq 2^{\circ}$ )	144,033	95.00	-0.006	74.21
缓坡耕地( $2^{\circ}$ ~ $6^{\circ}$ )	3203	2.11	32.68	3203
坡耕地( $6^{\circ}$ ~ $15^{\circ}$ )	3496	2.31	18.90	3496
陡坡耕地( $15^{\circ}$ ~ $25^{\circ}$ )	828	0.55	5.48	828
很陡坡耕地( $> 25^{\circ}$ )	46	0.03	3.07	46

**Table 3. Changes of the max, min and average value of elevation and slope at different level of reclamation rate**  
**表 3. 不同垦殖率下坡度、海拔高度变化情况**

垦殖率(%)	坡度( $^{\circ}$ )			海拔高度(m)		
	Max	Min	Average	Max	Min	Average
>70	12	0	0	392	-3	22
40~70	20	0	1	774	-1	45
20~40	23	0	5	829	-1	133
$\leq 20$	31	0	10	1313	-1	240

着坡度的变陡, 海拔的增高, 垦殖率随之下降。证明河流湖泊和坡度是影响耕地分布的两大因子。

#### d) 土壤适宜性

将耕地垦殖率分别与土壤碳储量和土壤氮储量进行相关分析, 结果表明: 垦殖率与土壤碳储量、土壤氮储量的相关系数分别为 0.027、0.120, 显著性水平分别为 0.183、0.000, 垦殖率与土壤碳氮储量仅存在弱相关性。苏皖地区是传统的农业区, 耕作历史悠久, 各类土壤已熟化为宜垦农业土壤, 土壤因素与垦殖率的弱相关性也是合理的, 土壤因素几乎不限制耕地的分布状况。

#### 2) 社会因素

通过对 1995 年的耕地垦殖率和人口数据进行分析, 发现两者存在弱相关, 相关系数仅为 0.015(显著性水平(双侧)为 0.466)。究其原因, 可能由于在人口快速增长的现代社会, 耕地开垦程度已达饱和, 宜垦土地几乎早已开垦殆尽, 人口的增长并不能引起耕地的增垦, 姑且认为人口对现代耕地的分布并不显著。但是, 在历史时期, 人口的变化导致对粮食需求的变化, 而农耕时期粮食主要依赖于耕地的开垦, 人口的分布可能在一定程度上反映出历史耕地垦殖强度的空间差异性。

综上, 对现代耕地分布的主导因子为河流湖泊和地形。

## 4. 模型构建与检验

### 4.1. 模型构建

苏皖地区河流湖泊量化时, 通过计算每个网格内河流湖泊所占的比重得到网格  $i$  内陆地的面积比  $\varepsilon(i)$ 。坡度和垦殖率成负相关, 坡度越小的地区优先被开垦, 对坡度进行标准化处理:

$$S'(i) = \frac{\max(S) - S(i)}{\max(S) - \min(S)}$$

其中,  $S'(i)$  表示苏皖地区网格  $i$  坡度的标准化值,  $\max(S)$  表示苏皖地区坡度的最大值,  $\min(S)$  表示苏皖地区坡度的最小值,  $S(i)$  表示网格  $i$  的坡度值。

在林珊珊等[15]、李士成等[18]重建清代传统农区和西南地区网格化耕地数据的基础上, 构建苏皖地区耕地宜垦性算法模型:

$$\alpha(i) = \varepsilon(i) \times S'(i) \times N(i) \quad (1)$$

其中,  $\alpha(i)$  为网格  $i$  的宜垦程度,  $\varepsilon(i)$  表示网格  $i$  内陆地面积所占比重,  $S'(i)$  表示网格  $i$  坡度的标准化值,  $N(i)$  表示现代的垦殖强度。

网格  $i$  的耕地面积比重为:  $\beta(i) = \alpha(i) / \sum_i \alpha(i)$ 。

网格  $i$  的耕地面积为:  $\delta(i) = \beta(i) \times A$  ( $A$  为苏皖地区耕地总面积)。

网格  $i$  的垦殖率为:  $K(i) = \delta / \text{area}(i)$  ( $\text{area}$  为网格面积, 此处值为  $100 \text{ km}^2$ )。

在宜垦程度较高的地区, 垦殖率可能会出现大于 1 的情况, 对于现代的耕地分布, 垦殖率最大赋值 1, 超过的耕地按照上述模型再次分配, 直到所有的网格满足条件即可。

将 1980s 的耕地数据按照模型 1 进行分配, 并按照(模型数据 - 遥感数据)/遥感数据的方法分别计算两者的相对误差, 得到结果如图 2。结果显示: 模型 1 的结果误差较小, 相对误差主要集中在  $-0.07 \sim 0.07$  范围内, 只有零星的几个网格误差值较大。

### 4.2. 模型检验

将 2000 年的数据按照重建模型进行重新分配, 由于模型数据和遥感数据具有一致的空间对应关系, 模型数据与遥感数据在空间上可以进行比对。

一方面，在进行模型重建时，是以苏皖地区作为一个整体进行耕地的网格分配，因此不论是遥感数据还是重建数据在苏皖地区大尺度下耕地总量是一致的，两者的均值相同，即模型数据和遥感数据的绝对误差平均值理论上为 0，可以采用对模型数据和遥感数据的绝对误差进行单组样本 T 检验来验证模型的可靠性。T 检验的前提条件是需服从正态分布。首先，将绝对误差做直方图(图 3)，基本符合正态分布，但是对其进行正态分布 K-S 检验，发现 Z 值为 4.513，概率为 0，小于 0.05，并不服从正态分布，需要进行正态转换。转换后 Z 值为 0.494，概率为 0.968，大于 0.05，服从正态分布。

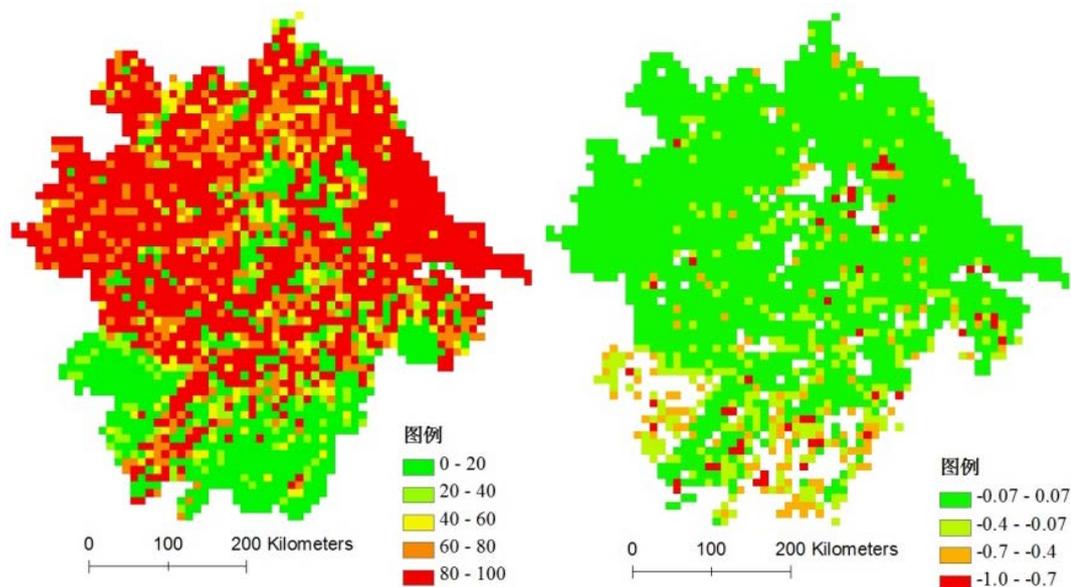


Figure 2. The pattern of cropland distribution and the relative error in Jiangsu and Anhui  
图 2. 耕地分配格局及相对误差

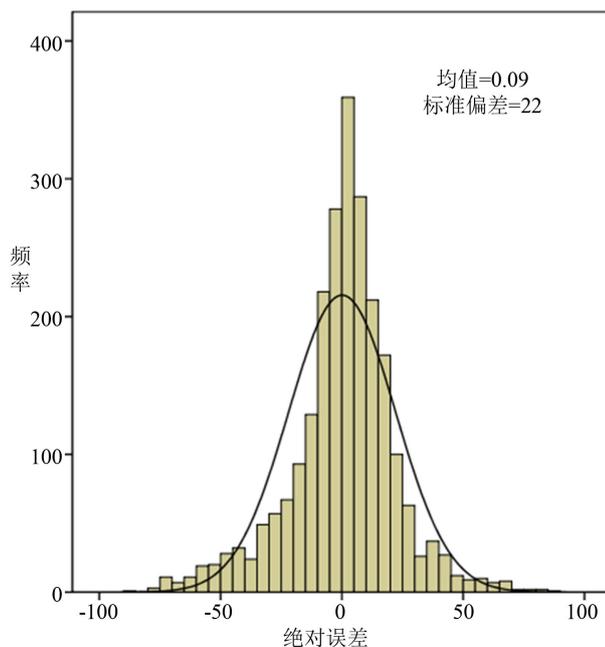


Figure 3. Histogram of absolute error  
图 3. 绝对误差直方图

对绝对误差进行单组样本  $t$  检验(表 4), 假设绝对误差的平均值为 0, 即模型数据与遥感数据并无显著差异。结果显示,  $t = 0.002$ , 概率为 0.999, 大于显著性水平 0.05, 接受假设, 即模型数据与遥感数据不存在显著差异是有统计学意义的。

另一方面, 分别从苏皖地区、安徽省和江苏省的空间尺度下, 计算模型数据和遥感数据的相关系数  $R$  和标准偏差  $S$ :

$$R_{XY} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} S_{XY} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - y_i)^2}$$

检验结果如图 4, 显示不论是苏皖地区, 还是安徽省或者江苏省, 重建数据和遥感数据的相关性都非常好, 相关系数分别为 0.798, 0.828, 0.714; 标准偏差分别为 22.2, 21.5, 23.1。因此本文所选取的影响耕地分布的因子是较为合理的, 构建的耕地网格化方法具有可行性, 本模型基本能够用于历史时期的耕地网格化重建。

## 5. 结论

本文通过分析影响耕地分布的各个因子及构建耕地分配的网格化模型, 得到以下结论:

1) 影响耕地分布的最主要因素是自然因素, 包括河流湖泊、地形等, 坡度与垦殖率的相关系数高达  $-0.628$ , 河流湖泊地区为非耕地的剔除区域。气候因素仅仅影响作物的熟制, 土壤因素影响作物种类。在现代由于人口急剧增长, 土地垦殖早已达到饱和, 人口的持续增长对垦殖率的影响已微乎其微。但是, 在历史时期, 人口的变化会导致对粮食需求的变化, 而农耕时期粮食主要依赖于耕地的开垦, 人口的分布可能在一定程度上反映出历史耕地垦殖强度的空间差异性。

Table 4. One sample  $t$ -test of absolute error of reconstruction data and remote sensing data

表 4. 模型数据与遥感数据绝对误差的单组样本  $t$  检验

检验值 = 0						
	T	Df	Sig.(双侧)	均值差值	差分的 95% 置信区间	
					下限	上限
Normal Score of 绝对误差 Using Blom's Formula	0.002	2382	0.999	-0.0000361	-0.040099	0.040171

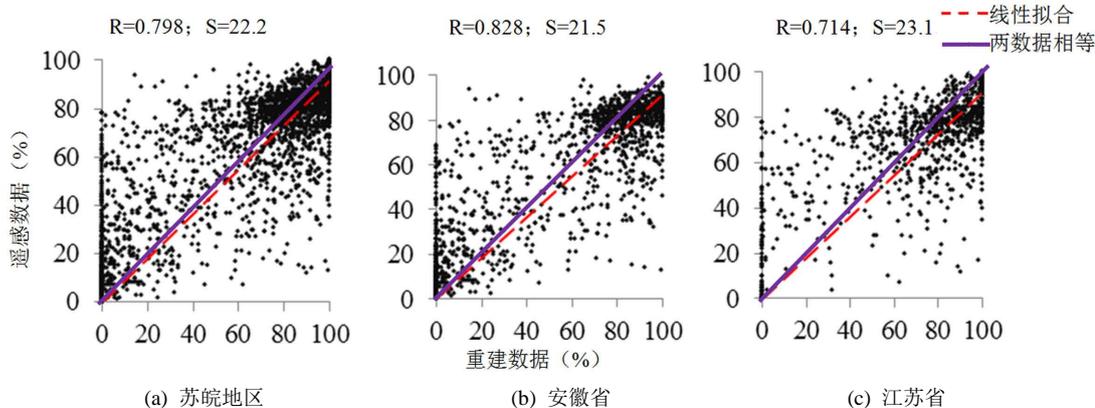


Figure 4. Correlation of reconstruction data with remote sensing data in Jiangsu and Anhui

图 4. 苏皖地区重建数据与遥感数据相关性

2) 通过对模型的评估检验,发现不论是从均值,还是从相关系数或标准偏差来看,本文由河流、坡度构建的耕地网格化模型数据和遥感数据的差异不大,两者的相关系数  $R$  为 0.798,标准偏差  $S$  为 22.2。模型是合理的、可行的,能够用于耕地格局重建,若考虑特定历史时期的人口增长和分布,则能够用它重建历史耕地分布格局。

## 基金项目

国家自然科学基金(41371201)资助。

## 参考文献 (References)

- [1] Ramankutty, N. and Foley, J.A. (1999) Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycle*, **13**, 997-1027.
- [2] Ramankutty, N. and Foley, J.A. (1999) Estimating historical changes in land cover: North American croplands from 1850 to 1992. *Global Ecology and Biogeography*, **8**, 381-396.
- [3] Pongratz, J., Reick, C., Raddatz, T. and Claussen, M. (2008) A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles*, **22**, Article ID: GB3018.
- [4] Goldewijk, K.K. (2001) Estimating global land use change over the past 300 years: The HYDE database. *Global Biogeochemical Cycles*, **15**, 417-433.
- [5] Goldewijk, K.K., Beusen, A., Dreht, G.V. and Vos, M.D. (2011) The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. *Global Ecology and Biogeography*, **20**, 73-86.
- [6] 朱枫, 崔雪锋, 缪丽娟 (2012) 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*, **12**, 1563-1573.
- [7] 李巧萍, 丁一汇, 董文杰 (2006) 中国近代土地利用变化对区域气候影响的数值模拟. *气象学报*, **3**, 257-270.
- [8] IPCC (2007) *Climate change 2007: The physical science basis: Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge and New York.
- [9] Zheng, J.Y., Lin, S.S. and He, F.N. (2009) Recent progress in studies on land cover change and its regional climatic effects over China during historical times. *Advances in Atmospheric Sciences*, **26**, 793-802.
- [10] 尹永飞, 陈星, 张洁, 汤剑平 (2009) 中国过去 300 年土地利用变化及其气候效应. *第四纪研究*, **6**, 1162-1169.
- [11] 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*, **8**, 1048-1059.
- [12] Li, B.B., Fang, X.Q., Ye, Y. and Zhang, X.Z. (2010) Accuracy assessment of global historical cropland datasets based on regional reconstructed historical data—A case study in Northeast China. *Science China Earth Sciences*, **53**, 1689-1699.
- [13] 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*, **9**, 1190-1200.
- [14] Zhang, X.Z., He, F.N. and Li, S.C. (2013) Reconstructed cropland in the mid-eleventh century in the traditional agricultural area of China: Implications of comparisons among datasets. *Regional Environmental Change*, **13**, 969-977.
- [15] 林珊珊, 郑景云, 何凡能 (2008) 中国传统农区历史耕地数据网格化方法. *地理学报*, **1**, 83-92.
- [16] 何凡能, 李士成, 张学珍 (2011) 北宋中期耕地面积及其空间分布格局重建. *地理学报*, **11**, 1531-1539.
- [17] 李柯, 何凡能, 张学珍 (2011) 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*, **12**, 2281-2288.
- [18] 李士成, 何凡能, 陈屹松 (2012) 清代西南地区耕地空间格局网格化重建. *地理科学进展*, **9**, 1196-1203.
- [19] Liu, M.L. and Tian, H.Q. (2010) China's land cover and land use change from 1700 to 2005: Estimations from high-resolution satellite data and historical archives. *Global Biogeochemical Cycles*, **24**, Article ID: GB3003.
- [20] 颀耀文, 王学强, 汪桂生, 余林 (2013) 基于网格化模型的黑河流域中游历史时期耕地分布模拟. *地球科学进展*, **1**, 71-78.
- [21] 国家统计局 2010 年第六次全国人口普查主要数据公报(第 2 号)  
<http://www.stats.gov.cn/tjsj/pcsj/rkpc/6rp/indexch.htm>

- [22] 《中国农业全书\*江苏卷》总编辑委员会 (1998) 中国农业全书\*江苏卷. 中国农业出版社, 北京.
- [23] 《中国农业全书\*安徽卷》总编辑委员会 (1998) 中国农业全书\*安徽卷. 中国农业出版社, 北京.
- [24] 国务院第二次全国农业普查领导小组, 中华人民共和国国家统计局 (2009) 中国第二次全国农业普查资料汇编农村卷. 中国统计出版社, 北京.
- [25] 江苏省水文水资源勘测局 (2002) 江苏省水文志. 江苏古籍出版社, 南京.
- [26] 李世奎, 侯光良, 欧阳海, 崔读昌, 郑剑非 (1988) 中国农业气候资源和农业气候区划. 科学出版社, 北京.
- [27] 孙颢, 石玉林 (2003) 中国农业土地利用. 江苏科学技术出版社, 南京.
- [28] 张家诚 (1991) 中国气候总论. 气象出版社, 北京.
- [29] 李天杰 (2005) 土壤地理学. 第三版, 高等教育出版社, 北京.
- [30] 朱华忠 (2006) 基于生态过程参数的中国森林遥感分类及碳密度变化格局. 博士学位论文, 中国科学院地理科学与资源研究所, 北京.
- [31] 德希珀金斯 (1984) 中国农业的发展(1368-1968年). 上海译文出版社, 上海.
- [32] 安徽省志水利志编委会 (1994) 安徽省水利志水文志. 黄山书社, 合肥.
- [33] 国家统计局农村社会经济调查司 (2012) 中国农村统计年鉴. 中国统计出版社, 北京.
- [34] 安徽省统计局, 国家统计局江苏调查总队 (2013) 安徽统计年鉴 2013. 中国统计出版社, 北京.
- [35] 江苏省统计局, 国家统计局江苏调查总队 (2013) 江苏统计年鉴 2013. 中国统计出版社, 北京.
- [36] 程纯枢 (1991) 中国的气候与农业. 气象出版社, 北京.
- [37] 王静爱, 左伟 (2010) 中国地理图集. 中国地图出版社, 北京.