

中国公里网格GDP分布数据集

黄耀欢¹, 江东¹, 付晶莹^{1,2}

(1. 中国科学院地理科学与资源研究所, 北京 100101;

2. 中国科学院大学, 北京 100049)

摘要: GDP空间化即是以一定尺寸的地理网格单元代替行政单元, 便于与土地利用、生态环境背景数据等自然要素数据进行分析整合, 为促进多领域之间数据共享, 实现空间统计综合分析提供极有力的支持。中国公里网格GDP空间分布数据集是在综合分析了人类活动形成的土地利用格局与GDP的空间互动规律的基础上, 建立分产业(一、二、三产业)GDP数量与土地利用类型的空间相关性模型。在区位理论、空间统计学的支持下, 综合分析我国社会经济发展的空间特征和区域差异, 进行定量表达和分层分区。在每个子区内, 选择一定数量的具有代表性的县, 作为建模的样本县, 在样本县各项数据的基础上建立GDP空间分布模型, 利用该模型对GDP分县统计数据空间化, 从而生成1 km网格的GDP空间分布数据。为了对GDP数据的精度进行校验, 在我国东部、中部及西部共选取了40个具有完整的GDP数据的乡(镇)进行校验, 结果显示, 空间化GDP数据的相对误差在6%到17%之间。该数据集实现了我国GDP数据的空间定量模拟, 建立了统一空间坐标参数、统一数据格式、统一的数据和元数据标准的全国1 km网格GDP分布数据集。

关键词: 中国; GDP; 公里网格; 空间分布

DOI: 10.11821/dlxb2014S007

数据引用格式: 黄耀欢, 江东, 付晶莹. 中国公里网格GDP分布数据集. 全球变化科学研究数据出版系统, 2014. DOI:10.3974/geodb.2014.01.07.v1, <http://www.geodoi.ac.cn/doi.aspx?doi=10.3974/geodb.2014.01.07.v1>

1 前言

中国公里网格GDP分布数据集(GDPGrid_China)以土地利用空间格局为社会经济要素空间化的主要影响因子, 以遥感数据和GDP统计数据为主要信息源, 将以行政区域为单位的统计型GDP数据进行了空间化处理, 实现了1公里网格GDP的空间分布数据集。

2 数据集简介

中国公里网格GDP分布数据集(GDPGrid_China)的数据集名称、短名名称、通讯作者、数据作者、地理区域、数据年代、数据空间分辨率、数据出版单位、数据共享网络服务平台、数据集组成、出版及责任编辑等信息一并列于表1。

3 数据研发方法

中国公里网格GDP空间分布数据集研发方法的关键是建立构成GDP产值的要素(第

收稿日期: 2014-03-17; 修订日期: 2014-06-20

基金项目: 中国科学院重点部署项目(2012)(KZZD-EW-08-03) [Foundation: Key Programs of the Chinese Academy of Sciences, No.KZZD-EW-08-03]

作者简介: 黄耀欢(1982-), 男, 安徽人, 副研究员, 主要从事地理信息科学的应用研究。E-mail: huangyh@lreis.ac.cn

表 1 中国公里网格 GDP 分布数据集 (GDPGrid_China) 元数据简表

| | | | |
|----------|---|--------|--------------------------|
| 数据库(集)名称 | 中国公里网格 GDP 分布数据集 | | |
| 数据库(集)短名 | GDPGrid_China | | |
| 通讯作者 | 黄耀欢 (huangyh@lreis.ac.cn) | | |
| 数据作者 | 黄耀欢, 中国科学院地理科学与资源研究所, huangyh@lreis.ac.cn 江东, 中国科学院地理科学与资源研究所, jiangd@igsnr.ac.cn 付晶莹, 中国科学院地理科学与资源研究所, fuji@lreis.ac.cn | | |
| 地理区域 | 地理区域覆盖中国 31 个省级行政区 (香港、台湾、澳门无数据), 经度范围大致位于 73°33'E~135°05'E, 纬度范围大致位于 3°51'N~53°33'N。 | | |
| 数据年代 | 2005 年、2010 年 | | |
| 数据空间分辨率 | 1 公里 | 数据时间频率 | 5 年 |
| 数据格式 | TIF, ZIP | 数据量 | 原始数据 171MB; 压缩数据 15MB |
| 数据出版单位 | 中国科学院地理科学与资源研究所 DOI: 10.3974/ | | |
| 数据共享服务平台 | 中国科学院地理科学与资源研究所全球变化科学研究数据出版系统, http://www.geodoi.ac.cn 国家地球系统科学数据共享平台, http://www.geodata.cn | | |
| 责任编辑 | 刘闯, 石瑞香, 周翔, 何书金 | | |
| 数据共享政策 | 本数据著作权所有者同意依据《全球变化科学研究数据出版系统》管理政策中“全社会公益性共享政策”实施共享。 | | |

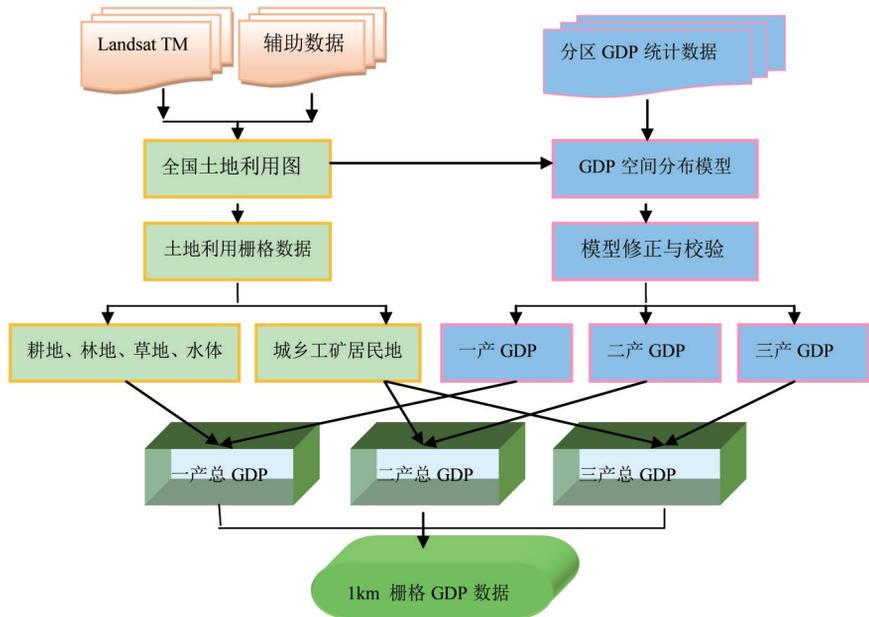


图 1 GDP 栅格数据生成的流程图

一、二、三产业结构)与土地利用类型的关系模型。在综合分析了人类活动形成的土地利用格局与GDP的空间互动规律的基础上,建立分产业(一、二、三产业)GDP数量与土地利用类型的空间相关性模型。利用该模型对GDP分县统计数据空间化,从而生成1 km网格的GDP空间分布数据,该数据集反映了GDP数据在全国范围内的详细空间分布状况。该数据为栅格数据类型,每个栅格代表每平方公里范围内的GDP总产值,单位为万元,数据格式为gird,数据以WGS-84椭球为基准,投影方式为Albers投影,具体流程见图1。

GDP空间化实现的具体技术流程如下:首先,将全国按方案划分为若干个人口/经济区,区内的自然、人口、资源、经济条件应基本一致,在建模时,假设每个区是均一的、各向同性的区域。然后,在每个子区内,选择一定数量的具有代表性的县,作为建模的样

本县构建GDP空间分布模型,样本县应按县级GDP的高、中、低分别选取,并且兼顾各种土地利用类型^[1-4]。

GDP空间分布模型的关键是确定构成GDP产值的要素(第一、二、三产业结构)与土地利用要素(土地利用格局)的关系。表2给出了GDP-土地利用格局的关系矩阵。

表2 GDP-土地利用格局的关系矩阵

| 土地利用类型 | G1 (第一产业) | G2 (第二产业) | G3 (第三产业) |
|-------------|--------------|--------------|--------------|
| 耕地、林地、草地、水体 | 1 | 0 | 0 |
| 城乡工矿居民地 | 0 | 1 | 1 |

关系矩阵中的1表示密切相关,0表示无关或相关性极弱。在某个子区内,分别建立第一、二、三产业模型,实现该区域的GDP空间化过程,获取全国范围的公里网格GDP空间分布。其中,分县统计数据来源于[5, 6],行政边界数据来源于[7]。

4 数据组成

中国公里网格GDP分布数据集是GRID格式的中国公里网格GDP分布数据集,由2个数据集组成,分别为(数据集名称2005年的数据和2010年的数据)压缩后数据量48.1 MB。数据可视化图(ARCGIS TIF格式,单位:万元)见图2、图3。

5 数据可靠性检验

为了对GDP数据的精度进行校验,在我国东部、中部及西部共选取了40个具有完整的GDP数据的乡(镇)(北京:东城区、朝阳区、房山区、大兴区、延庆县;上海:黄浦区、徐汇区、闸北区、松江区、崇明县;吉林:南关区、永吉县、梨树县、东辽县、通化县、抚松县、通榆县、长岭县、汪清县;河南:中牟县、通许县、汝阳县、郟县、内黄县、淇县、封丘县、温县、台前县;新疆:乌鲁木齐县、独山子区、鄯善县、伊吾县、呼图壁县、温泉县、且末县、温宿县;重庆:万州区、黔江区、涪陵区、渝中区、大渡口区、江北区)作为检验样本,对模型结果进行校验:首先将乡(镇)边界数字化为矢量图,将其转换为统一的投影、坐标系统后,利用GIS的空间分析功能,叠加到GDP栅格数据上,统计出各乡(镇)的GDP总和,与统计的GDP数据进行对比,检

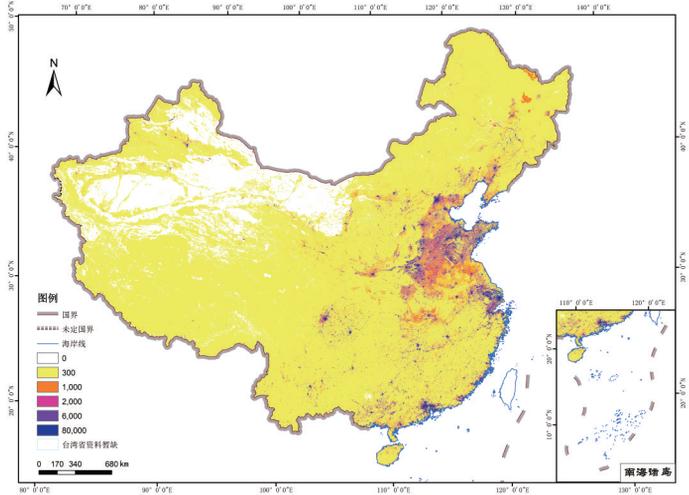


图2 中国公里网格2005年GDP分布图 (ARCGIS TIF数据格式)

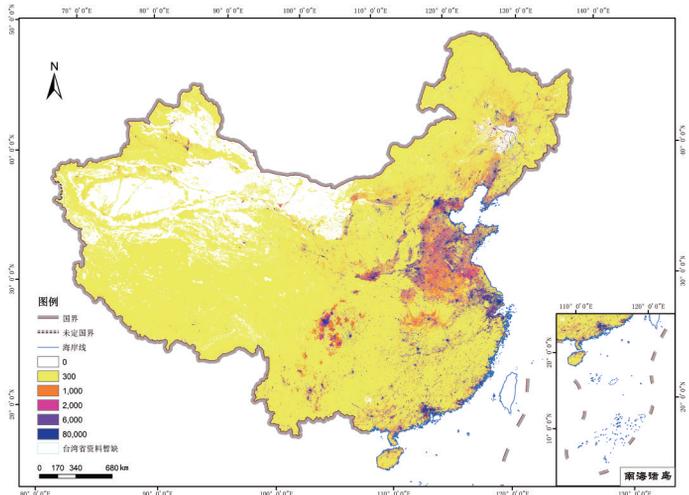


图3 中国公里网格2010年GDP分布图 (ARCGIS TIF数据格式)

验到GDP数据的精度。首先将乡(镇)边界数字化为矢量图,将其转换为统一的投影、坐标系统后,利用GIS的空间分析功能,叠加到GDP栅格数据上,统计出各乡(镇)的GDP总和,与统计的GDP数据进行对比,检

验表明本次工作获得了很好的结果：以统计数据为真值，空间化GDP数据的相对误差在6%~17%之间。

6 结论

定量、定位的获得社会经济信息是分析人文因素与生态环境的响应机理与互动模式的基础，GDP空间分布数据对于区域分析模型建立具有重要的科学价值。该数据集是以1 km空间分辨率为基本单元的中国公里网格GDP分布数据集，提供了将基于社会单元(省、市、县、乡镇行政区划单元等)的统计信息转化为基于公里网格的空间信息的技术方法，为促进多领域之间数据共享，实现空间统计综合分析提供极有力的支持。

致谢：本数据集中用到的中国1:10万土地利用数据来自中国科学院资源环境科学数据中心，数据提供者刘纪远研究员等，在此表示感谢。

参考文献

- [1] 王建华, 江东. 黄河流域二元水循环要素反演研究. 北京: 科学出版社, 2006.
- [2] 江东, 杨小唤, 王乃斌 等. 基于RS、GIS的人口空间分布研究. 地理科学进展, 2002, 17(5): 734-738.
- [3] 江东. 人文要素空间化研究进展. 甘肃科学学报, 2007, 19(2): 91-94.
- [4] 刘红辉, 江东, 杨小唤 等. 基于遥感的全国GDP 1 km网格的空间化表达, 地球信息科学, 2005, 7(2): 120-123.
- [5] 国家统计局. 2005中国统计年鉴. 北京: 中国统计出版社, 2005.
- [6] 国家统计局. 2010中国统计年鉴. 北京: 中国统计出版社, 2010.
- [7] 国家基础地理信息中心. 中华人民共和国国界地理信息系统数据 (1:100万比例尺). 2008.