

中国农业资源区划数据集

陈佑启¹, 邹金秋^{1*}

1. 中国农业科学院, 农业资源与农业区划研究所, 北京 100081

摘要: 农业资源调查与农业区划是农业发展的基础性、综合性的长期工作, 从中华人民共和国成立以来始终被列为国家长期研究的重要任务。我国先后开展了三次全国性大规模的农业区划工作, 获得了全国共 2 108 个县(按 80 年代初的行政区划)的农业资源调查和农业区划资料, 取得了全国农业现状区划、中华人民共和国自然地理图集等一批重要成果。并在此基础上进行了全国以及不同地区(县级, 部分地区到乡级)的各种专业性的农业区划与区域性的综合区划, 取得了成套系统的有关全国以及不同规模尺度上的农业资源及其区划成果图件、统计数据与研究报告, 共包括 8 万多项成果和报告, 以及 4 万多幅图件。这套数据是目前我国有关农业资源信息与区划成果的唯一、完整和系统的数据, 在查清农业资源家底、合理利用资源、监测资源动态、研究农业发展战略、建立农产品商品基地及农业地区开发、调整农业结构、因地制宜合理布局生产力等方面提供了大量科学基础资料, 对拟定与落实农业发展规划, 保护农业资源, 挖掘农业生产潜力, 合理布局与配置农业资源, 提高农业综合生产力, 促进农业和农村持续发展等起到了极大作用。

关键词: 农业资源; 农业区划; 科学数据; 数据共享

数据库(集)基本信息简介

数据库(集)名称	中国农业资源区划数据集
数据作者	陈佑启、邹金秋
数据通信作者	邹金秋 (zoujinqiu@caas.cn)
数据时间范围	1949~2016年
地理区域	北纬3°52'~53°33', 东经73°40'~135°2', 地理区域范围涉及到的国家包括中国。
数据量	1 263 748 MB
数据格式	*.doc, *.pdf, *.xls, *.shp等
数据服务系统网址	http://region.agridata.cn/ ; http://agri.ckcest.cn/appcenter/proapp/29.html ; http://www.sciencedb.cn/dataSet/handle/539
基金项目	科技基础性工作专项项目“科技基础性工作数据资料集成与规范化整编”(2013FY110900), 国家国际科技合作专项项目“天空地一体化精准农业物联网平台联合研发”(2014DFE10220), 农业区划科学数据挖掘利用研究(720-36)。

文献 DOI:

10.11922/csdata.2017.0008.zh

数据 DOI:

10.11922/sciencedb.539

文献分类: 地球科学

收稿日期: 2017-12-02

开放同评: 2018-02-27

发表日期: 2018-06-20

* 论文通信作者

邹金秋: zoujinqiu@caas.cn

数据库（集）组成	中国农业资源区划数据集由有关农业区划的文本、统计数据及相关地图数据组成，为了科学管理，总共分成综合农业区划数据库、农业自然区划数据库、农业技术措施区划数据库、农业部门区划数据库、农村经济区划数据库、其他专题区划数据库、农业区划理论与方法信息数据库、其他农业区划信息数据库，共 8 个数据库。
----------	---

引言

农业资源调查与农业区划是农业发展的基础性、综合性的长期工作，从中华人民共和国成立以来始终被列为国家长期研究的重要任务。我国先后开展了三次全国性大规模的农业区划工作。1953～1957 年，当时的中央农业部计划经济局直接负责全国农业区划工作，各大区的农林部及农业厅局的计划部门均建立了相应的农业区划管理部门，具体组织农业资源调查和农业区划研究。这一期间，主要工作集中在农业自然区划草案的编制与自然区划方法的研究等方面。1963 年，中央国务院召开的“全国农业科学技术工作会议”把农业自然资源调查和农业区划列为十年全国农业科学技术发展规划的第一项重点项目。1978 年，中央召开的“全国科学技术大会”，把“全国农业自然资源调查和农业区划研究”列为 1978～1985 年期间全国科学技术规划纲要重点科学技术研究项目中的第一项。1979 年国务院国发（1979）36 号文指出：开展农业自然资源研究和农业区划的研究是合理利用农业自然资源，发展农业生产，建立科学管理，实现农业现代化的重要的基础工作；各有关部、各省、市、自治区必须积极地、有计划地、长期地进行下去；并决定成立全国农业资源调查和农业区划委员会及办公室。在委员会主任委员、国务院副总理王任重的亲自领导下，国家农委、国家科委、农业部等有关部委组织省、地、县各级区划部门与科研单位参加，分期分批地开展了全国农业资源调查和农业区划工作。这项工作共动员了近百万人参与，获得了全国共 2108 个县（按 80 年代初的行政区划）的农业资源调查和农业区划资料；取得了“全国农业现状区划”“中华人民共和国自然地理图集”等一批重要成果，并在此基础上进行了全国以及不同地区（县级，部分地区到乡级）的各种专业性的农业区划与区域性的综合区划，取得了成套系统的有关全国以及不同规模尺度上的农业资源及其区划成果图件、统计数据与研究报告，共包括 8 万多项成果和报告，以及 4 万多幅图件。这套成果经全国农业区划委员会科学技术成果评审组评定，共评出一等奖 12 项、二等奖 52 项、三等奖 229 项。其中，“中国综合农业区划”获国家科技进步一等奖^[1]。

这套数据是目前我国唯一、完整和系统的有关农业资源信息与区划成果的数据。这些农业资源信息与区划成果在查清农业资源家底、合理利用资源、监测资源动态、研究农业发展战略、建立农产品商品基地及农业地区开发、调整农业结构、因地制宜合理布局生产力等方面提供了大量科学基础资料，对拟定与落实农业发展规划，保护农业资源，挖掘农业生产潜力，合理布局与配置农业资源，提高农业综合生产力，促进农业和农村持续发展等起到了极大作用。

在空间信息技术、计算机网络技术等新技术的推动下，资源信息化建设及共享利用早已成为共识，各单位都开展了信息化工作。2002 年，国家信息化领导小组发布了《关于我国电子政务建设指导意见》（中办发〔2002〕17 号文件），提出了推进我国电子政务建设

的指导意见,开启了信息化发展的新纪元。目前,我国的信息化基础设施发生了巨大变化,各“金字号”工程建设、“国家科学数据共享平台”等项目建设等相继完成,搭建了国内最大的“国家农业科学数据共享中心”,大大提高了我国农业信息化水平。

“全国农业资源区划数据集”从2002年开始建设,分别获得了农业部、科技部的资助。其中,国家农业科学数据共享中心—区划科学数据分中心完成数据采集与共享平台建设,是国内最大的农业区划数据资源平台。目前,共包括164 658个真实数据,分成五大数据库39个数据集,长期开展对外免费公益共享,实名制注册用户超7 000人,共享平台日访问量超3 000次以上,产生了良好的社会效益。

1 数据采集和处理方法

全国农业资源区划科学数据大部分是纸质图书资料,保存在中国农业科学院农业资源与农业区划研究所,同时有大量的图书存放在省级农业资源区划管理部门。图1为中国农业科学院农业资源与农业区划研究所农业区划图书馆馆藏图书情况。



图1 原始资料保存状况

针对这些纸质图书数据,采用人工电子化录入的方法进行数据采集。在国家科技部基础性工作及平台项目的资助下,开展资源数据库建设。国家农业科学数据共享中心区划科学分中心根据需要采集的数据类型及现状,制定了科学的数据采集方法与技术流程^[2]。图2详细说明了工作方法 with 处理流程。

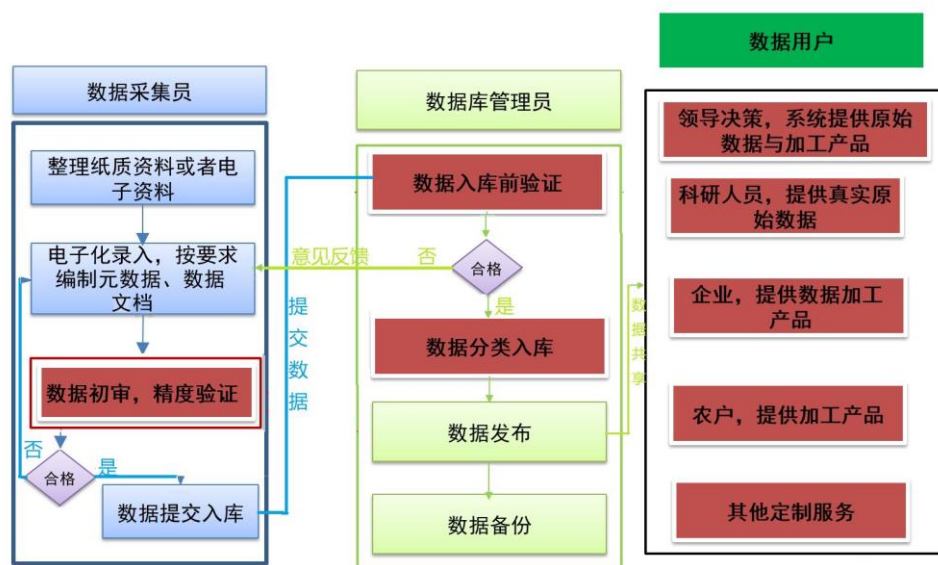


图2 数据采集与处理流程

数据采集方法如下：1) 制定科学的数据录入与校验标准，将相关数据进行科学分类，并由数据采集员分类录入。2) 针对书本中文字资料，采用人工录入的方式电子化后，安排不同人员进行一审和二审校对，保证文字录入正确率 99%以上，并保证文字排版和原书本排版一致，录成电子版本后将文件保存成 WORD 格式。3) 针对其中的表格数据，除在 WORD 文件中以表格形式保存外，还单独提取至 EXCEL 软件中保存成*.xls 文件，以便进行查询统计分析。4) 针对其中的地图数据，除将图片扫描后保存到 WORD 文件外，还利用 GIS 软件对图片进行矢量化，赋予正确的投影和坐标，并保存成通用的 SHP 文件格式。5) 针对照片或者多媒体数据，单独保存成 JPG 格式或者流媒体格式。6) 数据采集员完成数据采集后，将数据提交到数据库管理员。数据库管理员对数据进行精度验证以后，利用开发的管理软件建立每个真实数据的元数据信息，详细说明数据的来源、质量及使用情况。并将数据入库，在数据共享平台 <http://region.agridata.cn/>发布并共享。7) 用户通过实名制注册模式访问共享平台，可以查询和浏览本数据库的全部数据信息（包括各文件对应的元数据及真实文件数据）。

数据库管理与共享平台利用 Oracle 数据库软件作为基础数据管理软件平台，基于 ArcGIS 软件进行系统集成与开发。管理系统为单机版本的数据分析利用软件，共享平台为基于 WebGIS 模式开发的网路共享平台，利用该平台实现数据的免费共享，见图 3。用户除查询和浏览数据外，如确需真实数据时，可以在线申请获取真实数据。数据管理员通过对数据密级、申请用户信息进行审核后，批准是否可以通过数据下载。一旦通过，用户再次登录时即可在线下载对应的真实文件数据。



图3 数据共享平台

2 数据样本描述

在数据库中, 共有文本数据 28 151 个、统计数据 31 371 个、矢量地图数据 2 196 个、栅格数据 10 914 个、多媒体数据 64 个, 总数据 72 696 个。每个数据的组织与管理模式都相同, 针对每个科学数据, 实体文件数据保存成对应的文件格式。其中文本数据的格式为 doc、txt、pdf 等; 统计数据为 dbf、xls; 地图数据为 shp; 栅格数据为 tif、img 等。

针对每个真实文件, 分别建立该数据完整的元数据说明信息。用户首先浏览的是该数据的元数据信息, 通过查阅该数据的时间、摘要、关键词、精度、数据格式等信息后判断该数据是否符合自身需求。如符合, 可以直接申请下载。

用户登录, 通过输入查询条件, 数据查询结果如图 4。点击数据名称, 可以查看该数据的元数据。

文本数据查询条件

行政区名称:

数据标题: 河南省农业区域综合开发总体规划

关键词:

数据日期:

数据单位:

数据类别: 全部

数据形式: 全部

数据集名: 全部

查询

重置

使用说明

用户在阅读使用说明后, 可以根据自己的需要进行行政区、数据标题、关键词、数据生产日期、数据类别和形式等字段的任意简单查询和组合查询。例如输入“河北”, 所有和查询条件完全符合的农业资源区划数据的元数据将分页显示, 用户可以点击“前进”或“后退”进行页面的选择, 也可以通过直接输入页码完成。

数据查询结果

合计: 4个 文本: 3个 统计: 0个 矢量: 0个 栅格: 1个 遥感: 0个 多媒体: 0个

行政区名称	数据名称	关键词	数据日期	数
1 河南省	河南省农业区域综合开发总体规划-文献篇	总体规划, 区域农业, 综合开发, 开发文件汇编, 河南省	1992-01-01	文
2 河南省	河南省农业区域综合开发总体规划	总体规划, 区域农业, 综合开发, 河南省	1992-01-01	文
3 河南省	河南省农业区域综合开发总体规划分区图	河南省, 农业区域, 综合开发总体规划, 分区图	1992-11-01	栅
4 河南省	河南省农业区域综合开发总体规划-规划篇	总体规划, 区域农业, 综合开发, 开发规划, 河南省	1992-01-01	文

图 4 数据查询结果

以“河南省农业区域综合开发总体规划”数据文件为例, 点击该数据名称, 会弹出下图 5 对应的该文件对应的元数据说明信息, 主要包括数据中文名称、英文名称、关键词、数据摘要、数据产生日期、数据版本信息、数据文件语种、数据生产目的、数据所有人及元数据所有人信息等。

查阅元数据后, 可以看到图 5 下方“真实文件名”处有个“数据请求”的命令按钮, 如果需要该数据, 点击此命令申请下载。数据管理员会在 1 个工作日内审理此数据是否许可下载。如通过, 用户再次登录时, 会有提示在线下载该文件的信息, 见图 6。

行政区：	河南省
数据中文名：	河南省农业区域综合开发总体规划
关键词：	总体规划，区域农业，综合开发，河南省
摘要：	本书主要分析河南省概况及战略地位，农业发展现状与开
数据日期：	1992-01-01
版本：	第一版
语种：	中文
目的：	挽救历年区划资源数据，为领导决策，实现资源数据共享
现状：	1
起始时间：	1992-01-01
终止时间：	2012-12-31
数据覆盖范围：	河南省
数据类型：	文本报告
数据类别：	其他数据
数据集：	农业区域开发与规划数据库
发行单位：	中国农科院资源区划所
发行时间：	1992-01-01
发行介质：	电子书
项目名称：	农业资源与区划调查
项目类型：	部级
负责单位名称：	农业资源与区划调查
负责人姓名：	部级
元数据日期：	2014-11-13
元数据更新日期：	2014-11-13
元数据负责人：	陈佑启
元数据负责单位名称：	中国农科院资源区划所
通讯地址：	北京市海淀区中关村南大街12号
邮政编码：	100081
电子邮件：	zoujinqiu@caas.cn
电话号码：	010-82109614
传真号码：	010-82109614
真实文件名：	023060.doc 数据请求

图 5 元数据实例

未读消息		历史请求数据			
	请求数据标题	下载	是否同意	消息回复	标
1	河南省农业区域综合开发总体规划		同意		取消
2	中共中央国务院文件,中央领导同志对农业自然资源调查和农		同意		取消
3	河南省12316三农热线工作简报总第51期		拒绝	你好	取消

图 6 真实文件下载信息

3 数据质量控制和评估

在数据采集与数据库建设过程中，采用了严格的数据质量管理标准体系。按照图 7 所示的数据质量控制流程，每个数据采集过程都经过严格的一审和二审程序。对每一个文件都进行错误率统计，尽可能将误差降低到最小程度，以保证数据录入过程中误差控制在 0.1%

以下。误差率计算方法为：千字中错别字（包括标点符号和数字错误）个数占比；同时还需要综合考虑排版错误、表格数据中的数字错误、地图数据考虑投影坐标和边界走势错误、文件命名误差等。每本纸质图书由数据采集员 1 进行电子化录入，然后交数据采集员 2 进行排版与误差率校核计算；初步符合要求后交二审采集员 3，二审采集员 3 再次对数据质量进行校核，确认误差在 0.1% 以下后，建立数据的元数据信息，元数据的建立应符合农业科学数据核心元数据中规定的要求^[3]，然后提交数据管理员。

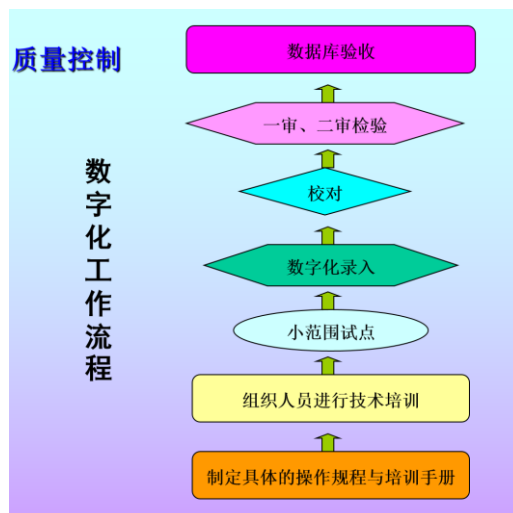


图 7 数据质量控制流程

对于地图数据，在赋予投影与坐标时也充分注意到误差控制。在矢量化过程中尤其要求数据的精细化操作，电子地图图片扫描时要求不变形、大于 500 点以上像素扫描；导入 GIS 软件后，进行数据投影与坐标配准，坐标纠正点不少于 40 个；数据电子化跟踪过程要求与图片地图点、线要素咬合度大于 98%；录入和建立矢量属性时，要求字符错误率在 1% 以下。

总之，在数据采集过程中，这套数据有严格的数据精度保证措施。但是鉴于这套科学数据产生时间相对较早，90% 以上数据产生于 2000 年前，75% 左右的数据形成于 20 世纪 80 年代。随着国家经济建设发展，部分农业资源统计数据已经与现状不符，比如土地利用统计数据，后备耕地资源数据、农业种植和受灾面积、农业养殖统计数据等，因此这些数据为历史存底数据，不能作为现时数据进行科学决策。

4 数据价值

本套数据在现代农业园区规划、两区划定、优势农产品布局、农业产业结构调整及后备资源分析等农业生产管理决策过程中，具有巨大的使用价值。本套数据是我国农业科学工作者在国家极其困难的情况下，花费大量人力、财力收集和整理的珍贵历史资料，摸清了我国农业资源家底，并结合专家知识生产的成套农业区划资源数据集。虽然当时受资源调查手段、图件制作技术及定量评价模型等缺乏的因素制约，但是总体数据精度是非常高。本套数据是国内外唯一的符合我国国情的数据资源，应该受到足够的重视和保护。

5 数据使用方法和建议

本套数据可以作为历史资料分析国内农业资源状况。同时这套数据凝聚了老一辈科学家大量的智慧，科学合理地划分了国内各种作物与资源的分布区域，在自然规则没有发生大改变的前提下，这些数据知识是完全可以利用的。

但是由于数据采集时间相对较早，随着我国社会经济发展，有些现实情况已经有所改变，因此对于这套资源数据中各类生产和加工的统计数据，建议仅作为历史资料进行对比分析；对于其中各类规划分布图等数据，完全可以直接使用或者适当修改后使用，大部分数据使用者反馈表明，这些科学数据依然适用于现代农业布局与区划。

本数据集也在不断地收集和更新相关数据，尤其是随着现阶段“天空地”一体化的更新技术及手段的出现，信息在膨胀和扩张。用户完全可以综合多方面的信息来源，进行数据融合、同化和挖掘后提取最有价值的信息。

数据作者分工职责

这套数据的原始作者包括许多区划专家和农业管理人员，数据分别以书本和内部资料的形式发表。本套数据的建立与加工处理仅仅是在农业部和科技部的资助下，将这些宝贵的科学资源数据进行电子化，实现价值回归。在保持原样的情况下进行电子化处理与共享，没有对数据进行分析处理。数据质量由原始作者负责，数据共享原则与责任由中国农业科学院农业资源与农业区划研究所陈佑启、邹金秋负责。

陈佑启（1962—），男，湖南省岳阳市人，博士，研究员，研究方向为土地利用。主要承担工作：农业大数据，负责国家农业科学数据中心—区划科学数据分中心建设与维护。

邹金秋（1975—），男，湖北省崇阳人，博士，助理研究员，研究方向为农业遥感。主要承担工作：农业系统集成，承担国家农业科学数据中心—区划科学数据分中心建设与维护。

参考文献

- [1] 全国农业资源区划办公室，中国农业资源与区划学会，中国农业科学院农业资源与农业区划研究所. 中国农业资源区划 30 年[M]. 北京：中国农业科学技术出版社，2011：1–30.
- [2] 马海龙，陈佑启，邹金秋. “国家农业科学数据共享中心农业区划数据分中心”系统设计与建设[J]. 中国农业资源与区划，2010，31(1): 75–80.
- [3] 孟宪学. 农业科学数据核心元数据[M]. 北京：中国农业科学技术出版社，2008：221–243.

论文引用格式

陈佑启，邹金秋. 中国农业资源区划数据集[J/OL]. 中国科学数据，2018，3(2). (2018-06-19). DOI: 10.11922/csdata.2017.0008.zh.

数据引用格式

陈佑启, 邹金秋. 中国农业资源区划数据集. [DB/OL]. Science Data Bank, 2017. (2017-12-02). DOI: 10.11922/sciencedb.539.

The dataset of agricultural resources and regional mapping in China

Chen Youqi¹, Zou Jinqiu^{1*}

1. Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, P. R. China

* Email: zoujinqiu@caas.cn

Abstract: Survey and regional mapping of agricultural resources is a long-term basic and comprehensive work of agricultural development. It has been listed as an important task in long-term national research plans since the founding of P. R. China. Agricultural mapping at regional scale has been conducted nationwide three times afterwards, which collected data of agricultural resources and their regional mapping in more than 2108 counties (as per the administrative divisions in the early 1980s), including a huge set of maps, statistical data and reports. This dataset contains the most complete and systematic information about agricultural resources and their regional mapping. It is vital for agricultural resource monitoring and utilization, agricultural regional planning, the formulation of regional agricultural policies, the adjustment of agricultural structure and measurement, the promotion of agricultural productivity, and so on.

Keywords: agricultural resources; agricultural regionalization; scientific data; data sharing

Dataset Profile

Title	The data set of agricultural resources and regional mapping in China
Data corresponding author	Zou Jinqiu (zoujinqiu@caas.cn)
Data authors	Chen Youqi; Zou Jinqiu
Time range	1949 – 2016
Geographical scope	Longitude & latitude: 3°52' N – 53°33' N, 73°40'E – 135°2' E; specific areas include China.
Data volume	1 263 748 MB

Data format	*.doc, *.pdf, *.xls, *.shp, etc
Data service system	< http://region.agridata.cn/ >; < http://agri.ckcest.cn/appcenter/proapp/29.html >; < http://www.sciencedb.cn/dataSet/handle/539 >
Sources of funding	The special project of basic work of science and technology, "data integration and standardization of basic work of science and technology" (2013FY110900); The special project of international scientific and technological cooperation, "Joint research and development of the Internet of things platform on precision agriculture with an integration of aerospace-aviation-ground observation" (2014DFE10220); Research on scientific data mining and utilization of Agricultural Resources and Regionalization in China (720-36).
Dataset composition	This dataset consists of data in text, statistics and map formats. For the convenience of scientific management, the dataset is made up of eight databases: ① Comprehensive agricultural zoning database; ② Agricultural natural zoning database; ③ Agricultural technical measures zoning database; ④ Agricultural sector division database; ⑤ Rural economic division database; ⑥ Other thematic zoning databases; ⑦ Agricultural zoning theory and method database; ⑧ Other agricultural zoning information databases.