

我国 1 : 5 万土壤图数据库的质量控制

徐爱国¹, 张认连¹, 田有国², 冀宏杰¹, 张怀志¹, 龙怀玉¹

(¹ 中国农业科学院农业资源与农业区划研究所/农业部作物营养与肥料重点实验室, 北京 100081; ² 全国农业技术推广服务中心, 北京 100025)

摘要: 【背景】具有时空维度的土壤大数据可为农业、环境、国土等部门了解、研究和管理土壤与环境提供数据支撑。历史上完成的纸质土壤调查图件和资料则是构建土壤时空数据库的重要基础。由于不同地区、不同时期土壤图成图标准差异较大, 构建高质量土壤时空大数据, 对异源、异质的非标准化土壤图资料的整理必须进行质量控制。【目的】本研究的目的是解析我国各地异源土壤普查原始图件状况, 弄清其数字化整合建库中的质量控制关键问题, 探索相应的技术指标与规范, 为构建大比例尺土壤图数据库提供质量保障。【方法】通过分析我国土壤普查图件资料状况, 结合我国数字地形图质量控制方法, 制定了精度指标, 对土壤图数据库实现地理精度质量控制; 对比分县土壤图中的土壤名称和现有国家标准土壤分类, 实现全库土壤类型名称的修编、编码; 通过拓扑检查完成土壤图空间数据库结构统一, 以统一境界实现境界标准化。【结果】利用异源大比例尺土壤图构建大区域土壤图数据库的质量控制主要包括数字化地理精度控制、土壤要素提取标准化、全图整合赋码三大部分。其中数字化地理精度控制包括纸质扫描精度、定位坐标系、不同地形图幅的几何校正精度、几何校正控制点的选取和点数规定。土壤要素提取标准化主要包括土壤要素与非土壤要素采集与核准、空间要素分类码、图层与图幅的标准化。全图整合赋码包括土壤类型名称的修编、编码和行政边界标准化, 针对原始分县土壤图土壤名称的不规范命名, 特别是土类的不规范命名, 遵照国家标准中的土类名称, 借鉴分县土壤志剖面记载、省级和国家级分类汇总文献和著作, 将 300 余个不规范土类命名修编为 60 个国家标准命名; 采用层次编码方法, 按照土纲、土类至土种的五级编码实现全库每个土壤类型的唯一编码; 采用国家标准境界实现行政边界的无缝连接, 并在此基础上实现了土壤图空间库存储单元由分县转换为 1 : 5 万标准分幅。【结论】利用异源大比例尺非标准土壤图资料构建大区域的土壤图数据库, 其质量控制的关键为数字化地理精度控制、土壤要素标准化、全图整合赋码。针对这三个问题构建的质量控制方法和相应指标, 符合我国土壤普查图件数字化建库实际情况, 为建设覆盖我国全域的土壤图数据库提供了质量保障。采用这一方法, 建立了 1 : 5 万土壤图数据库共计 13 240 个, 1 : 5 万标准分幅, 涉及 1 688 个县市和国营农场。本文还讨论了本研究现状并展望其拓展方向, 期望通过结合土壤理化性状信息, 为进行土壤长效性质数据挖掘研究提供数据基础。

关键词: 1 : 5 万土壤图; 数字化; 数据库; 质量控制; 中国

Quality Control of Soil Map Database at 1:50 000 Scale in China

XU AiGuo¹, ZHANG RenLian¹, TIAN YouGuo¹, JI HongJie¹, ZHANG HuaiZhi¹, LONG HuaiYu¹

(¹ *Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081*; ² *National Agricultural Technology Extension and Service Center, Beijing 100025*)

Abstract: 【Background】 Soil large data with time and space dimension, could provide data support for research work and policy decision in agriculture, environment, and land management. The soil survey maps and reports are the base on building soil time and space database. Because of the large differences of soil mapping standards between different areas and different periods, quality control is necessary for developing high quality soil database from non-standard soil maps and soil survey reports.

收稿日期: 2018-04-08; 接受日期: 2018-09-21

基金项目: 国家重点研发计划 (2017YFD0200607)、科技部基础性工作专项 (2012FY112100、2006FY120200)、科技部公益性研究专项 (2001DIA20024)、国家公益性行业 (农业) 科研专项 (201503121-13)

联系方式: 徐爱国, E-mail: xuaiguo@caas.cn

【Objective】 The purpose of this study was to analyze the status of paper soil maps, to identify the key issues of quality control in its digital integration database, and to explore the corresponding technical standards and specifications, so as to provide quality guarantee for large scale soil map database. 【Method】 By analyzing the situation of soil maps and soil survey reports and combining related research for precision control of digital topographic map in China, precision requirements suitable for digitizing soil map were developed. After comparing soil type names from county level soil maps and soil standard names in national standard, soil names, especially soil great group names, were revised in soil map database and every soil name was encoded. 【Result】 For building soil map database from different sources large scale soil maps in large area, the quality control included the geographic precision, standardization of digital soil map elements, integration of soil map database, and coding elements type for the whole database. The contents of geographic precision included scanning precision of paper maps, coordinate system, geometric correction precision, as well as collection and numbers of control points. Standardization of map elements extraction mainly included the collection of soil elements and non-soil elements, classification and code of digital element types, and standard of layer name and map sheets. The integration for whole database included revision of soil types, code of soil types and non-soil types, and standard of boundary by using uniform boundary. In the process of soil data integration and harmonization, revising soil names, especially revising soil great group names and coding, was the mainly quality control method. According to national standard and soil profile records in soil survey reports of county, province and national level for reference, more than 300 non-standard soil great group names were revised to 60 national standard names. By hierarchical code method, a unique code for each soil type in the entire database was achieved with five-level coding of soil order, soil great group to soil species. Using uniform boundary, the boundaries of adjacent counties were seamless. Based on the integration of the entire soil special database, the conversion were achieved from counties map sheet to international standard map sheet at 1 : 50 000. 【Conclusion】 The quality control is very important in building soil map database of large area by using non-standard soil maps from different region and period. The key of quality control were geographic precision, standardization of digital soil map elements, integration of soil map database and coding soil types for the entire database. The quality control methods and corresponding indicators for these three issues conformed to the actual situation of soil survey maps in China. Using the above methods, the 1 : 50 000 soil map database had a total 13 240 standard map sheets of 1 : 50 000 scale, and involving 1 688 counties and county farms. In addition, we discussed the status of this database and prospected the future direction for expansion, in the hope that to provide a data basis for the study of the long-term soil properties data mining by combining the physical and chemical properties of the soil.

Key words: 1 : 50 000 scale soil map; digital; database; quality control; China

0 引言

【研究意义】土壤数据作为国家基础资源数据, 可为农业、环境、国土等部门的管理和研究提供数据支撑, 而高质量的土壤数据库是实现这些应用的基本保证。【前人研究进展】自 20 世纪 70 年代, 欧美发达国家在历史土壤调查图件和资料的基础上, 陆续建成或基本建成了本国大比例尺土壤数据库, 如美国、荷兰、英国、法国及澳大利亚等国^[1-8]。其中美国 SSURGO (县级) 土壤数据库基于的比例尺为 1 : 12 000 至 1 : 63 360, 澳大利亚为 1 : 50 000。我国的土壤数据库建设起步较晚, 20 世纪 90 年代后期, 一些科研院所陆续建立了全国小比例尺^[9-11]、个别省的大、中比例尺土壤数据库^[12-16]。这些数据库几乎均采用历史土壤调查资料建立。可以说, 历史完成的土壤调查图件和资料至今是最主要的数据来源。由于各地土壤调查人员不同、各地技术条件不同, 导致土壤图件成图比例尺、坐标系、图例与注记表达有较大差异。当利

用这些源于各地的土壤图件构建覆盖大区域的、大比例尺土壤图时, 不仅工程量大, 还会遇到异源、异质、异构、非标准化土壤图件存在的各类问题。这些问题不仅是各国在将源于各地的历史数据进行数字化整合建库时都会面临的问题, 也是我国未较早完成全国大比例尺土壤图数据库的原因之一。我国在 20 世纪 80 年代即进行了全国第二次土壤普查 (以下简称二普), 从 1979 年试点县开始到 1994 年成果汇总结束^[17], 调查了 2 444 个县、312 个国营农 (牧、林) 场和 44 个林业区^[9] (二普时期行政区划单位, 下同), 汇总了丰富翔实的土壤调查资料, 包括县级以上各级土壤图、土壤养分系列图及土壤志 (县级图比例尺以 1 : 5 万至 1 : 20 万为主), 这些成果基本代表了 20 世纪 80 年代我国土壤资源状况。但因上述客观的技术和专业问题, 大比例尺建设相对迟缓。【本研究切入点】在土壤调查资料的数字化整合中需要解决的主要问题包括: 统一地理坐标、提高定位精度, 统一制图要素、边界, 规范土壤分类名称, 以及如何处理专业

理解差异导致的土壤分类差异等。此类情况在各国土壤数据整合建库中也普遍存在,使得国家间、地区间难以进行数据整合^[18]。【拟解决的关键问题】本研究在土壤图空间数据建库中,针对上述质量控制关键点,介绍了数字化整合建库质量控制的方法和精度指标,包括以扫描精度、坐标系、控制点选取、偏移精度控制地理定位精度;土壤图层要素的分类提取,采用统一边界,以及对土壤类型进行修编和制定可补码的编码规则等,为 1 : 50 000 土壤图数据库建设的精度质量和属性完整提供了保证,为大比例尺土壤数据在相关部门和领域的应用提供了保障。

1 资料来源

本研究的资料来源为全国第二次土壤普查(1979—1994 年^[18])成果一分县纸质土壤图(含国营农场)和分县土壤(种)志(或土壤调查报告)。本研究共收集分县土壤图合计 1 688 个县市和国营农场(为二普调查时的行政区划单位),基本覆盖我国中东部地区和西北部分地区;与这些县市和农场土壤图对应的分县土壤志主要用来补充、校验土壤图例中各级土壤分类名称,并在土壤类型修编时作为参考。分县土壤图标注以及土壤志中的记载显示,多数省份的外野调查于 1979—1986 年间完成,新疆、西藏等边疆省份于 1988 年左右结束,因此这些土壤普查资料基本代表了 20 世纪 80 年代我国的土壤资源状况。

本研究收集的 1 688 个(1 556 个县、市和 132 个国营农场)土壤图中,农区县、农场的成图比例尺以 1 : 5 万为主,个别郊区、农场为 1 : 2.5 万至 1 : 3 万;林区、农牧交错带的县市一般为 1 : 10 万,部分为 1 : 5 万;牧区县比例尺一般为 1 : 20 万,与《中国土壤普查技术》^[19]中的介绍基本一致。

采用国家基础地理信息中心“1 : 5 万矢量地形要素数据”(以下简称 1 : 5 万 DLG, 2008 版)的公路、铁路、线状水系、境界线 4 个图层,用于扫描土壤图的几何校正(地理配准),境界线还用于矢量土壤图行政边界的标准化。

2 土壤图矢量化的质量控制与要素采集

2.1 地理精度质量控制

2.1.1 数字化方法与扫描分辨率 2000 年以后,计算机技术的快速发展促使扫描矢量化成为采集地图要

素的主要数字化方法,因此本研究采用“扫描矢量化法”进行土壤图的要素采集。

二普期间形成的县级纸质土壤图,主要通过对挖取土壤剖面的土体构造、土壤形态的观测、记载,以及周边环境的观察记载,判断土壤类型,在相同或更大比例尺地形图上勾绘(边远、高寒和不便进入的区域,利用航片或卫片判读解译^[20]),经转绘、清绘而成。每县的图纸厚薄不一、大小不一,有单幅,有多幅。由于年代久远,加之保存条件有限,图件普遍存在折痕、变形和污损。因此,本研究对每份图件均采取低温熨平后用 A0 幅面扫描仪扫描,个别超 A0 幅面的图纸,为保留原图未予裁切,而是折叠后扫描。扫描分辨率设定为 300 dpi^[21]。

2.1.2 栅格土壤图的定位坐标系 扫描土壤图(无投影信息)经几何校正后形成的栅格土壤图(有投影信息)作为矢量化工作底图。在我国,大于 1 : 50 万(含 1 : 50 万)国家基本比例尺地形图的平面坐标系采用高斯-克里格投影(等角横切圆柱投影)。几何校正时以原始图件的坐标系统为定位坐标系,以保证地理配准的精确。二普时期我国采用 1954 北京坐标系,因此,本研究对 1 : 5 万至 1 : 20 万比例尺的分县扫描土壤图采用该坐标系的高斯-克里格投影进行几何校正。

2.1.3 栅格土壤图的平面位置精度 在工程测量中,地形图的平面位置精度,是由主要地物点相对于邻近图根控制点的点位中误差来衡量的。中误差为带权残差平方和平均数的平方根,是作为在一定条件下衡量测量精度的一种数值指标。专业地图和地形图一样,定位精度与地形类型有关,地形越复杂,定位精度越低。因此,本研究参照数字地形图国家标准^[22],对不同的地形类型设置不同的精度限值,即山区的精度要低于平原地区;并参考该标准,以点位中误差来评定栅格土壤图(即扫描土壤图经配准后栅格图)的平面位置精度值,作为对配准精度的限定:即校正后土壤图上地物点对地形图同一个地物点的点位中误差不得大于表 1 中的规定。

地形图的定位精度在所有地图中最为严格。因此,作为专题图,平面位置精度指标比数字地形图^[25]放宽了 0.5 倍(表 1)。以 1 : 5 万比例尺为例,平地 and 山地平面位置精度分别相当于实地 37.5 m 和 50 m,而地形图是 25 m 和 37.5 m。精度放宽的原因主要有两个方面:

第一,用同名点为控制点进行地理配准的地图,精度略低。由于二普分县土壤图原图上没有公里格网,

缺少地理信息，只能选取与地形图中位置相同的地物点，如河流交汇处、道路交叉点、河流和道路的交叉点以及境界线的尖角处作为控制点，进行几何校正，这种方法比用图廓点或公里格网点校正的偏差略大。同时土壤图上较小的河流、湖泊、道路等地物比同比例尺地形图有较多缩减，增加了纠正难度和误差。第二，受调查人员对土壤分类的理解等主观因素的影响，以及土壤诊断剖面的选定位置和数量的影响，人为因素对土壤类型界线的影响要大于几何校正的精度偏差。鉴于此，本研究对栅格土壤图的定位精度比地形图略有放宽，既兼顾了纸质土壤图件的现实状况，又最大限度地保证土壤空间数据库的空间位置准确。

表 1 不同地形栅格土壤图的平面位置精度^[21]
Table 1 Permitted lanimetric accuracy for raster soil map of different topography

比例尺 Scale	地形 Landforms	
	平地、丘陵 Plains and hills mm (pixel, m)	山地、高山地 Mountains mm (pixel, m)
1 : 50000	≤0.75 (9, 37.5)	≤1 (12, 50)
1 : 100000	≤0.75 (9, 75)	≤1 (12, 100)

单位 mm 为等比例尺图面距离，单位像元为分辨率为 300 dpi 时的中误差距离所折算的像元数
Unit mm is the distance on map as the same scale. Unit pixel is the number of pixels converted by the middle error distance when the scanning resolution is 300dpi

表 2 两种地形的部分图幅控制点数及其均方差统计值
Table 2 Statistical value of the numbers of tics and RMS per map sheet for two types of topographies

地形 Topography	图幅数 Frame numbers	平均控制点数 Mean numbers of tics (frame)	最少控制点数 Minimum numbers of tics (frame)	最多控制点数 Maximum numbers of tics (frame)	总均方差 Total RMS of tics (m)	最小均方差 Minimum Total RMS of tics (m)	最大均方差 Maximum Total RMS of tics (m)
平地、丘陵 Plains and hills	355	31.5	21	81	27.5	3.6	41.6
山地、高山地 Mountains	269	31.4	21	99	34.5	4.9	54.1

2.2 图层内容与要素采集

本研究将土壤图数据库图层定名为 Soipy，该图层包括土壤类型和非土壤要素两类内容。前者是主体，后者则包括面状的河、湖、水库、冰川雪被及居民地等。两类要素均在土壤层采集，既保证了土壤层面状要素的连续完整，又指示了土壤类型周边地理环境，其中河流湖泊、冰川雪被等自然要素，可间接反映土壤的成土条件。

要素采集时，对未进行土壤调查，且土壤原始图中标注为林场、草场、矿山的区域（图斑），采集其

2.1.4 栅格土壤图的控制点数 二普中每个原始分县土壤图通常由 1—4 个大小不一的分幅组成。本研究通过对比例尺分别为 1 : 5 万和 1 : 10 万的 298 个县共计 826 幅分幅土壤图几何校正分析来看，对于普通大小的县域，选取河流交汇处、道路交叉点、河流和道路的交叉点以及境界线的尖角处作为控制点，单幅图控制点（tic）数一般不少于 30 个（山地图幅控制点数多于平原图幅），且每点及全部控制点平均均方差符合表 1 相应地形的指标时，抽取校正后栅格土壤图与地形图同名点的位置偏差基本在表 1 规定的范围内。当一幅图的控制点达到适当数量后，位置偏差并不随控制点增多而有显著降低，例如，江苏地处平原的如东县共计 4 幅图，其中一幅选取 48 个控制点，控制点总均方差（RMS）为 37.3 m；而该县另一幅图选取 30 个点位，RMS 则为 33.6 m。表 2 列出了 1 : 5 万比例尺、两种地形 624 个图幅的控制点数和控制点均方差统计值，表中可以看出，平原丘陵地区控制点的 RMS 小于山区，即平原图幅的定位精度要高于山区。因此本研究对 1 688 个县市及农场的单幅土壤图选取控制点数一般不少于 30 个。对分幅太小，在地形 4 要素图层上基本无同名地物点的土壤图，与本县邻幅土壤图图像拼接后再行几何校正。

土地利用属性并赋相应代码，如林地、草地等，以便今后补充土壤调查时，提供历史植被信息以帮助土壤类型的判别。对未进行土壤调查，且土壤原始图中标注为“国营农场”等行政单位的图斑，则按“无资料地区”采集属性。对注记缺失、无法从周边图斑判断属性的，也以“无资料地区”采集。

非土壤地物要素侧面说明了所在地区土壤的成土母质、地形等成土环境。为保持与基础地理信息要素的一致但又不过分突出非土壤要素的详细类别，本研究采用了测绘部门国标^[23]中对非土壤要素分为水系、

居民地、地貌、植被与土质四大类的分类，在非土壤要素属性赋值时，取较上级分类，并统一名称和代码。要素分类取到哪一级取决于对周边土壤类型的形成是否起到说明作用，如对“水系”和“植被与土质”类要素，采用第二级分类，其中“水系”第二级有“常年河、时令河、干河床”，而季节性河流周边土壤受水浸润的时间短于常年河，两类河边的土壤类型会有不同特点，如采用第一级分类“水系”则过于笼统。“植被与土质”第二级有“林地”、“草地”分级，这可以帮助今后诊断草原土壤和森林土壤类型。红树林是酸性硫酸盐土的指示植物，而海洋要素大类中，滩涂的第三级分类“红树林滩”可说明周边存在酸性硫酸盐土，因此这类图斑要素应赋属性“红树林滩”。对于无注记图斑要素，本研究将要素名称定为“无资料地区”。非土壤要素名称及代码详见国标^[24]。

2.3 属性检查和拓扑完整性检查

属性检查包括检查图斑属性赋码的正确性、缺失或相邻图斑属性是否相同，可通过 GIS 软件的专有模块查出并修改。拓扑完整性检查包括要素检查和拓扑检查。要素检查时将矢量图与几何校正后的栅格工作底图套合，全图检查错画、漏画图斑及其属性赋码，补充修正。拓扑检查是检查矢量图的拓扑错误，补充修正，并完成拓扑关系的构建。

2.4 已有矢量土壤图的地理精度控制

对其他单位已经完成的矢量土壤图，本研究通过与 1 : 5 万 DLG 套合，选取同名地物点来校验、修正地理位置精度，尽量减少重复建设。

地理位置精度校验采用以下方法：对有扫描栅格底图的，将扫描栅格图按 2.1.3 的精度要求进行几何校正后，与矢量土壤图套合，提取全图 30—40 个土壤图斑边界点坐标，平均位置中误差符合表 1 要求的，对矢量图重新纠正，直至检验吻合。对无扫描底图的矢量土壤图，将其与 1 : 5 万 DLG 中的公路、铁路、水系、居民地等图层套合，取与水面、居民地吻合的 30—40 个点进行位置中误差检验，不吻合的矢量图重新进行几何校正，直至符合表 1 的要求。

3 全国土壤图数据库整合质量控制

完成分县图形要素、图斑属性采集与拓扑构建后，进行全国土壤空间库整合，包括以下 4 方面内容。

3.1 土壤类型名称修编

本研究对现有全国土壤空间库汇总的土类名称为 312 个，而国家标准^[24]为 60 个。这主要是因二次

土壤普查中，省级、国家级汇总时，土类修正归并结果未反馈，分县土壤图没有修改所致。因此本研究在进行全国土壤空间库整合统一编码之前，先对土壤类型名称进行修编，在高级分类——土类水平上与国家标准^[24]一致。土类名称不规范主要有以下 4 种情况：

(1) 实际分类为亚类。分县土壤调查时暂定为土类，但在省级、国家级逐级汇总时，已统一修正为亚类。通过查阅、研究各省的省级和地级土壤^[25-26]等，逐省逐条进行修编。如调查时命名的土类有冲积土、塿土、滨海风沙土、黑色石灰土、黄红壤等均为亚类，部分修编对照结果见表 3。

(2) 土类更名。如“砖红壤性红壤”是调查初期部分县的土类命名，在调查后期和汇总时，统一更名为“赤红壤”。盐土和盐碱土，也是调查初期部分县的土类命名，在调查后期和汇总时土壤普查办公室已经根据盐碱特性重新分类命名。本研究在土类名称整合时，有的根据下级亚类名称进行上级归类，如亚类为“滨海滩地盐土”“草甸盐土”的，土类分别为“滨海盐土”和“草甸盐土”；有的根据亚类名称和所在地域进行上级归类，如亚类为“残积盐土”的，土类为“漠境盐土”。

(3) 习惯用名。这类问题是修编的主要方面，查阅这些土类所在县的土壤志剖面记载，根据成土因素及剖面特征修正其土类名称。如泛滥土因河流泛滥冲积成土，根据国家标准^[24]，修正土类名称为新积土，亚类名称为冲积土。

(4) 别字。如风砂土、塿土、砂浆黑土等，统一采用国标规范用字修编^[24]。

表 3 列举了土类归并主要涉及的 4 种情况及部分土类修正前后的对比。

在汇总全国分县土壤图和土种志土壤名称的基础上，由曾负责过全国土壤图汇总的老专家指导，对土类、亚类、土属、土种名称进行修编，重点修正明显的分类学错误和不规范用语^[18]。其中，对高级分类名称——土类，实行分省修编，并对个别区域不该出现的土壤名称，查阅原县土壤（种）志剖面描述后进行修改；对亚类名称作为土类出现的，均按国标修改^[24]为亚类。最终土类修编为 60 个；对低级分类——土属、土种名称，尽量保持其原始记载，以便为今后再进行土壤调查时提供基础资料。关于各分县土壤图低级土壤类型名称的修编原则、方法及修编结果，详见文献^[18]。

表 3 源于分县土壤调查资料的部分土类名修正对比表

Table 3 Examples of soil great group names from county soil survey report before and after revision

土类名称不规范类型	原土类名	原亚类名	修编土类名	修编亚类名
Not standard types of great group	Great group before revision	Sub group no- revised	Great group revised	Sub group after revision
原土类应为亚类	冲积土 Alluvial soil		新积土 Neo-alluvial soils	冲积土 Alluvial soil
Great group revising as sub group	瘠土 Cululic cinnamon soils		褐土 Cinnamon soils	瘠土 Cululic cinnamon soils
	滨海风沙土 Coastal Aeolian soils		风沙土 Aeolian soils	滨海风沙土 Coastal Aeolian soils
	黄红壤 Yellowish red earths		红壤 Red earths	黄红壤 Yellowish red earths
	黑色石灰土 Rendzina		石灰（岩）土 Limestone soils	黑色石灰土 Rendzina
原土类更名	砖红壤性红壤 lateritic red earth		赤红壤 Latosols	
Great group renamed	盐土/盐碱土 Solonchaks	草甸盐土	草甸盐土 Meadow solonchaks	草甸盐土 Meadow solonchaks
		Meadow solonchaks		
		残积盐土	漠境盐土 Desert solonchaks	残余盐土 Relic solonchaks
		Residual saline soil		
习惯用名	泛滥土 Flooding soils		新积土 Neo-alluvial soils	冲积土 Alluvial soil
Customary names	绿洲灌耕土 Oasis irrigated soils		灌漠土 Irrigated desert soils	
别字 Misused characters	风砂土 Aeolian soils		风沙土 Aeolian soils	

3.2 土壤类型统一编码与补码

进行全国土壤空间库整合建库，需要进行土壤类型名称的统一编码。为减少编码位数，本研究采用层次编码法，按照土壤分类系统的层级进行分组后编码。即：将土纲、土类、亚类、土属、土种五级分为三组统一编排，三组之间由横线相隔。其中，土类和亚类代码为一组，前两位英文大写字母代表土类，土类第一位码同国标土纲码，第二位码为同土纲下的土类按国标顺序赋字母码，土类码后以 1—2 位英文小写字母代表亚类，亚类按各类型出现县数的多少排序赋码；土属、土种各为一组，均为数字码，也按出现的县数排序赋码，这样使现有收集的土壤类型中，出现区域多、覆盖范围广的土壤类型编码排序靠前。全库土壤统一编码通过“智能化海量空间信息分析与地图制图软件包 IMAT”^[27]成为土壤空间库属性字段，从而保证空间分析时土壤类型的唯一标识。

目前本库包括 1 600 余个县市和农场。为保持已有土壤类型名称及其编码的一一对应关系，对于今后补充到 1：5 万土壤图数据库的县市中新增的土壤类型，按上述编码规则，即亚类以下级别均以出现的县数多少，在现有该土类的最终组位码之后按顺序续码，赋予新土壤类型以统一代码。

3.3 行政界两侧的土壤类型处理

在土壤图数据库整合时，时常出现行政边界两侧土壤类型不一致的情况，根据土壤调查资料和相关文献的记载^[16]，原因主要有两个：第一，二普期间，土

壤调查是以县为单位进行，各县调查人员对土壤分类理解不同，造成行政边界两侧土壤类型不一致。第二，河流、分水岭（山脊）、沟谷等常常构成母质、地形的自然分界，以这类自然分界为行政边界的，两侧因成土条件有异造成土壤类型不同。而前一种原因是造成行政边界两侧土壤类型不一致的主因。

进行土壤类型接边时，结合已有周边剖面数据，借助卫片、航片影像资料进行判读接边，必要时需实地挖取剖面观测验证。鉴于土壤分类问题的复杂性以及本研究的主体目标是全国土壤图的数字化整合建库，而土壤类型接边修正的工程量浩大，为最大限度保持历史资料原貌，为相关研究提供原始记载，本研究重点进行了土壤名称高级分类的修编（详见 3.1 节），并依此对处于县界两侧的土壤类型实现了部分接边；对分类名称仍不一致的情况保留原貌，以便为今后再进行土壤调查时提供基础资料。

3.4 行政边界标准化与图幅转换

3.4.1 边界标准化 二普时期分县土壤图分别由各县手工绘制，相邻两县交界处的县界常常不完全吻合，造成相邻县界有缝隙。为解决这一问题，本研究统一采用 1：5 万 DLG 行政边界裁切的方式替换原始边界，替换标准县界产生的缝隙，由最邻近图斑界线延伸闭合，保证了土壤类型在空间拓扑上连续无缝。行政界两侧的土壤类型按照 3.3 的方法进行处理。行政区划有变化的县市，在矢量化编辑时，以其中一个县界为共用界，与相邻县市土壤图拼接后，用 1：5 万

DLG 县界裁切替换。采用此方法,本研究对 1 688 个县市、国营农场土壤图实现了行政边界标准化,并完成拓扑重建。

3.4.2 图幅变换 空间数据库的关键结构之一是存储单元。由于分县土壤图以县为单位,考虑到以经常变化的县级行政区划为土壤图数据库的存储单元,会影响数据库的准确调用和分析。为保证数据结构的稳定,按照中国 1 : 5 万比例尺土壤图数据库数据模型(CDSM)设计^[28],以 1 : 5 万地形图分幅为全国 1 : 5 万土壤图数据库的数据单元。采用 IMAT 软件^[27]的分幅变换功能,根据分幅与分县对应关系,对每幅涉及的分县土壤图拼接后裁切,实现土壤图数据库从分县到分幅的批量转换,以 1 : 5 万分幅为单位存储,并完成同一分幅内相邻县界两侧图斑的同码融合和拓扑重建。

图幅变换后的 1 : 5 万土壤图数据库,共计 13 240 个 1 : 5 万地形图分幅,涉及 1 688 个县市、国营农场,西安 80 坐标系,高斯投影。此外,对缺失的县市所在分幅,则暂采用中、低精度比例尺图件补充,并以元数据记录。以低顶高图幅共计 10 090 个 1 : 5 万标准分幅。

4 结论与展望

对二普完成的分县大比例尺土壤图的分析显示:不同地区分县土壤图中坐标系、图例与注记表达均有较大差异。利用异源土壤图构建覆盖大区域土壤图数据库的质量控制主要包括:数字化地理精度控制、土壤要素图层标准化、全图整合赋码三项关键任务。其中,数字化地理精度控制包括扫描精度、定位坐标系规定、几何校正(定位)精度;土壤要素图层标准化包括土壤与非土壤要素采集要求、空间要素分类码、图层与图幅的标准化;全图整合赋码质量控制主要包括土壤类型修编、编码,以及行政边界标准化。针对这三项任务而构建的质量控制方法和相应指标,符合我国土壤普查图件数字化建库实际情况,有利于完善我国大比例尺土壤图数据库的数据质量。基于上述质量控制方法,本研究建立的 1 : 5 万土壤图数据库,目前含 13 240 个 1 : 5 万标准分幅,涉及 1 688 个县市和国营农场,对缺失的县市所在分幅,则暂采用中、低精度比例尺图件补充,涉及图幅共计 10 090 个 1 : 5 万标准分幅。

全国 1 : 5 万土壤图数据库作为土壤资源的基础数据库,除了可直接作为研究和决策的数据支撑外,

笔者认为,可开展数据挖掘研究,通过数字土壤制图方法,补充我国土壤长效指标数据的不足。

数字土壤制图是土壤调查与制图学科的研究范畴,是采用新方法对学科研究内容进行延伸,是对土壤基本性质空间分布定量化研究的拓展,而大比例尺土壤图数据库的建设为这类研究提供了很好的基础。例如基于现有图件和剖面数据,结合野外调查,利用数字土壤成图的模型方法,可获得区域的或全国范围土层厚度、土壤质地、土体构造等土壤最基础性质的空间分布数据(图),以补充初次调查时此类数据仅有剖面记录而未成图的不足。全国第二次土壤普查形成的土壤图和土壤(种)志等土壤调查资料,包含了土地资源状况、土壤肥力特征、土地利用、土壤改良方向和措施等信息,是一次全面的、综合的土壤资源调查,这些资料为数据挖掘研究和土壤调查成果的广泛应用提供了很好的数据基础。目前基于数据挖掘的数字土壤制图模型、方法^[29-32]日趋成熟,数字土壤制图的精度与不确定性^[33]定量分析也可为衍生数据的使用提供可靠性支撑。充分利用已有历史数据,对土壤长效性状进行模拟制图,生成衍生属性数据,则更有利于在农业、环境、国土及其他研究领域和部门的广泛应用。

致谢: 张维理研究员对本文给予了指导;黄鸿翔研究员对土类修编给予了指导和建议;雷秋良、武淑霞参与本研究。在资料收集过程中,以下单位和个人给予了大力帮助和支持: 山东、内蒙古、福建、湖北、湖南、河北、河南、陕西、山西、北京、天津、上海、甘肃、宁夏、吉林、江西、江苏、安徽、云南等省、市、自治区土肥站和广西河池市、辽宁沈阳市土肥站,四川省农业厅,广东、贵州、广西、青海、辽宁等省土肥所、黑龙江八一农垦大学张之一教授及相关县市区土肥工作者,在此一并致以诚挚的感谢!

References

- [1] Natural Resources Conservation Service, National Soil Survey Center, USDA. *Soil Survey Geographic (SSURGO) Data Base: Data Use Information*. USDA Miscellaneous Publication Number 1527, 1995.
- [2] DOBOS E, CARRÉ F, HENGL T, REUTER H I, TÓTH G. *Digital Soil Mapping as a Support to Production of Functional Maps*. Luxemburg: Office for Official Publications of the European Communities, EUR 22123 EN, 2006: 68.
- [3] LAGACHERIE P, MCBRATNEY A B, VOLTZ M. *Digital Soil Mapping, an Introductory Perspective*. Amsterdam: Elsevier, 2007:

- 659.
- [4] LAMBERT J J, DAROUSSIN J, EIMBERCK M. *Soil Geographical Database for Europe & the Mediterranean: Instructions Guide for Elaboration at Scale 1:1 000000*. Version 4.0. Luxembourg: Office for Official Publications of the European Communities, EUR 20422 EN, 2003: 64.
- [5] European Soil Data Centre. <http://eusoils.jrc.ec.europa.eu/>.
- [6] CSIRO land and water. Australian soil resource information system (ASRIS). <http://www.asris.csiro.au/asris>.
- [7] Cranfield Soil and AgriFood Institute, Cranfield University. LandIS -land information system: a soils-focused information system for England and Wales. <http://www.landis.org.uk/index.cfm>.
- [8] YOSHITAKE K. A computerized soil information system for arable land in Japan. *Soil Science and Plant Nutrition*, 1984, 30(3): 299-309.
- [9] 张定祥, 史学正, 于东升, 潘贤章, 孙维侠. 中国 1 : 100 万土壤数据库建设的基础. *地理学报*, 2002, 57(增刊): 82-86.
- ZHANG D X, SHI X Z, YU D S, PAN X Z, SUN W X. The basis for establishing China's 1:1000000 soil database. *Acta Geographica Sinica*, 2002, 57(Suppl.): 82-86. (in Chinese)
- [10] 张定祥, 潘贤章, 史学正, 杨金玲, 林杰. 中国 1 : 100 万土壤数据库建设中的几个问题. *土壤通报*, 2003, 34(2): 81-84.
- ZHANG D X, PAN X Z, SHI X Z, YANG J L, LIN J. Several remarks on construction of Chinese soil database at the scale of 1 : 1M. *Chinese Journal of Soil Science*, 2003, 34(2): 81-84. (in Chinese)
- [11] 周慧珍. 中国土壤信息共享研究—1 : 400 万中国土壤分布式查询数据库. *土壤学报*, 2002, 39 (4): 483-489.
- ZHOU H Z. Sharing of soil information data distributed inquiry data base of 1:4M soil information of China. *Acta Pedologica Sinica*, 2002, 39(4): 483-489 . (in Chinese)
- [12] 吴克宁, 张雷, 吕巧灵, 李玲, 杨阳. 基于 MAPGIS 建立中比例尺土壤数据库的研究—以河南省为例. *土壤通报*, 2008, 39(3): 475-479.
- WU K N, ZHANG L, LV Q L, LI L, YANG Y. Establishment of the medium scale soil database based on MAPGIS—a case in Henan province. *Chinese Journal of Soil Science*, 2008, 39(3): 475-479. (in Chinese)
- [13] 刘京, 常庆瑞, 岳庆玲, 陈勇, 王德彩, 陶文芳. 陕西省土壤数据库的设计研究. *干旱农业研究*, 2008, 26(5): 105-108.
- LIU J, CHANG Q R, YUE Q L, CHEN Y, WANG D C, TAO W F. Established of soil information system of Shaanxi province. *Agricultural Research in the Arid Area*, 2008, 26(5): 105-108. (in Chinese)
- [14] 荆长伟, 支俊俊, 张操, 林声盼, 肖锐, 李丹, 吴嘉平, 单英杰, 陈红金, 徐进, 倪志华. 浙江省中小比例尺土壤数据库的构建. *科技通报*, 2012, 28(11): 99-105.
- JI C W, ZHI J J, ZHANG C, LIN S P, XIAO R, LI D, WU J P, SHAN Y J, CHEN H J, XU J, NI Z H. Construction of medium and small scale soil geographic data base, Zhejiang province, China. *Bulletin Science and Technology*, 2012, 28(11): 99-105. (in Chinese)
- [15] 吕成文, 陈志诚, 陈鸿昭, 吴文君. 海南岛 1: 20 万 SOTER 数据库的组织与设计研究. *水土保持学报*, 2003, 17(6): 110-113.
- LU C W, CHEN Z C, CHEN H Z, WU W J. Design of Hainan island's SOTER database at 1:200000 scale. *Journal of Soil and Water Conservation*, 2003, 17(6): 110-113. (in Chinese)
- [16] 吴嘉平, 胡义镰, 支俊俊, 荆长伟, 陈红金, 徐进, 林声盼, 李丹, 张操, 肖锐, 黄慧青. 浙江省 1: 5 万大比例尺土壤数据库. *土壤学报*, 2013, 50(1): 31-40.
- WU J P, HU Y L, ZHI J J, JI C W, CHEN H J, XU J, LIN S P, LI D, ZHANG C, XIAO R, HUANG H Q. A 1:50 000 scale soil database of Zhejiang province, China. *Acta Pedologica Sinica*, 2013, 50(1):31-40. (in Chinese)
- [17] 席承藩, 章士炎. 全国土壤普查科研项目成果简介. *土壤学报*, 1994, 31(3): 330-335.
- XI C F, ZHANG S Y. Brief introduction on achievements in national soil survey project since 1979. *Acta Pedologica Sinica*, 1994, 31(3): 330-335. (in Chinese)
- [18] 张维理, 徐爱国, 张认连, 冀宏杰. 土壤分类研究回顾与我国土壤分类系统的修编. *中国农业科学*, 2014, 47(16): 3214-3230.
- ZHANG W L, XU A G, ZHANG R L, JI H J. Review of soil classification and revision of China soil classification system. *Scientia Agricultura Sinica*, 2014, 47(16): 3214-3230. (in Chinese)
- [19] 全国土壤普查办公室. 中国土壤普查技术. 北京: 农业出版社, 1992.
- National Soil Survey Office. *The Chinese Soil Survey Technology*. Beijing: Agricultural Press, 1992. (in Chinese)
- [20] 陈玉舟, 高正华. 我国土壤调查制图事业的发展与成就. *土壤肥料*, 1989(4): 7-12.
- CHEN Y Z, GAO Z H. Development and achievement of soil investigation cartography in China. *Soil and Fertilizer Sciences in China*, 1989(4): 7-12. (in Chinese)
- [21] 中华人民共和国农业部. 土壤制图: 1: 50 000 和 1: 100 000 土壤图数字化规范: GB/T 32738-2016. 北京: 中国标准出版社, 2016.
- Ministry of Agriculture, the People's Republic of China. Soil Mapping-specifications for Digitizing Soil Maps at the Scale of 1:50,000 and 1:100,000: GB/T 32738-2016. Beijing: Standards Press of China, 2016. (in Chinese)

- [22] 国家测绘局. 数字地形图产品基本要求: GB/T 17278-2009. 北京: 中国标准出版社, 2009.
- State Bureau of Surveying and Mapping. Basic Requirements for Products of Digital Topographic Map: GB/T 17278-2009. Beijing: Standards Press of China, 2009. (in Chinese)
- [23] 国家测绘局. 基础地理信息要素分类与代码: GB/T 13923-2006. 北京: 中国标准出版社, 2006.
- State Bureau of Surveying and Mapping. Specifications for Feature Classification and Codes of Fundamental Geographic Information: GB/T 13923-2006. Beijing: Standards Press of China, 2006. (in Chinese)
- [24] 中国标准化研究院. 中国土壤分类与代码: GB/T 17296-2009. 北京: 中国标准出版社, 2009.
- China National Institute of Standardization. Classification and Codes for Chinese Soil: GB/T 17296-2009. Beijing: Standards Press of China, 2009. (in Chinese)
- [25] 陕西省土壤普查办公室. 陕西土壤. 北京: 科学出版社, 1990.
- Soil Survey Office of Shaanxi Province. *Shaanxi Province Soil*. Beijing: Science Press, 1992. (in Chinese)
- [26] 曲靖地区土壤普查办公室, 云南省土壤普查办公室. 曲靖地区土壤, 1986.
- Soil Survey Office of Qujing District, Soil Survey Office of Yunnan Province. *Qujing District Soil*, 1986. (in Chinese)
- [27] 张维理. 智能化海量空间信息分析与地图制图软件包 IMAT 设计及构建. 中国农业科学, 2014, 47(16): 3250-3263.
- ZHANG W L. Design and development of software package intelligent mapping tools (IMAT). *Scientia Agricultura Sinica*, 2014, 47(16): 3250-3263. (in Chinese)
- [28] 张维理, 张认连, 徐爱国, 田有国, 姚政, 段宗颜. 中国: 1:5 万比例尺数字土壤的构建. 中国农业科学, 2014, 47(16): 3195-3213.
- ZHANG W L, ZHANG R L, XU A G, TIAN Y G, YAO Z, DUAN Z Y. Development of China digital soil maps (CDSM) at 1:50 000 scale. *Scientia Agricultura Sinica*, 2014, 47(16): 3195-3213. (in Chinese)
- [29] ROSSITER D G. Digital soil mapping as a component of data renewal for areas with sparse soil data infrastructures// Hartemink A E, McBratney A B, Mendonca-Santos M L. *Digital Soil Mapping with Limited Data*. Springer: Dordrecht, 2008: 69-80.
- [30] YIYI S, MINASNY B, MCBRATNEY A B, SARWANI M, SUTANDI A. Harmonizing legacy soil data for digital soil mapping in Indonesia. *Geoderma*, 2013, 192: 77-85.
- [31] DOBOS E, BIALKÓ T, MICHELI E, KOBZA J. Legacy soil data harmonization and database development// BOETTINGER J L, HOWELL D W, MOORE A C, HARTEMINK A E, BROWN K. *Digital Soil Mapping. Progress in Soil Science, Vol 2*. Springer: Dordrecht, 2010: 309-323.
- [32] MALONE B P, MINASNY B, ODGERS N P, MCBRATNEY A B. Using model average to combine soil property rasters from legacy soil maps and from point data. *Geoderma*, 2014, 232/234: 34-44.
- [33] TRUONG P N, HEUVELINK G B M. Uncertainty quantification of soil property maps with statistical expert elicitation. *Geodama*, 2013, 202/203: 142-152.

(责任编辑 李云霞)