

中国土壤系统分类土系数据库的建立、更新与应用

张学雷 张甘霖 龚子同

(中国科学院南京土壤研究所 南京 210008)

摘 要 本文根据美国土壤系统分类 NASIS 数据库系统建立的经验,参考 SOTER 数据库结构的特点和中国土壤系统分类研究的现状,论述了我国土系数据库建立、更新的主要内容和技术规范,展望了土系数据库在不同领域的应用前景。

关键词 系统分类;土系;数据库

1 中国土壤系统分类土系数据库建立的重要性

土壤(土系)数据库是现阶段土壤学基础和应用研究中十分重要的内容,就世界范围来看发展迅速。作为基本数据库之一,土系数据库是联系基础研究(土壤发生和分类)和生产应用的重要桥梁。以美国为例,该国土壤系统分类已拥有 18000 多个土系,是在 100 余年(1899~2000 年)全国土壤调查的基础上逐渐建立起来的,如今已经覆盖全美国 90% 的面积。所有这些土系的剖面记载和理化分析数据都已输入计算机,构成国家土壤信息系统 NASIS(National Soil Information System)数据库的基础,用户可上网(<http://nasis.nrcs.usda.gov>)获取有关土系的信息。1994 年 10 月, NASIS1.0 版本在各州土壤调查数据库 SSSD(State Soil Survey Database)的基础上开始发展,最新的版本 NASIS4.1.1 是 2000 年 3 月 15 日颁布的,它在世界范围内影响广泛。

80 年代到 90 年代末,中国土壤系统分类^[1]在开展高级分类单元研究的同时,在全国范围内进行了 14 个样区土壤基层分类的研究,并相继提出土系的概念。在随后中国科学院特别支持项目、国家自然科学基金重点项目“中国土壤系统分类基层单元研究”中,土系的划分和建立成为基本研究内容。目前,项目研究组努力的目标是在 3 年内建立 1000 余个土系,在研究样区建立土系数据库,通过全国权威土系参比机构(建议筹建中)的审定,逐步建立全国土系数据库,为将来“国家土壤信息系统”的建立打下基础。

土系数据库的建立可提高土系划分研究成果的显示度,并将研究结果直接与不同需求的用户见面,成为科研成果转化为现实生产力的桥梁。

2 土系数据库建立的基本内容与主要技术要点

2.1 基本内容

美国农业部自然资源保护中心(USDA—Natural Resources Conservation Service)建立了 3 类土壤地理数据库,用以表达 3 类土壤图,它们分别是土壤调查图 SSURGO (Soil Survey Geographic)、各州土壤图 STATGO (State Soil Geographic)和国家土壤图 NATSGO (National Soil Geographic)。NASIS 本身无空间数据库,目前正着用与上述 3 类土壤图合并,形成真正

的空间和属性数据一体化信息系统。在 SOTER^[2] 方法中, 这个问题已经解决。因此, 在中国土壤系统分类土系数据库乃至今后的国家土壤信息系统建立中, 开始就要注意这个问题, 避免走不必要的弯路。

在 GIS 环境下, 对图斑空间信息与属性信息分级连接建立统一的数据库的方法已被 SOTER 的理论和实践证明有着许多的优点, 得到国际土壤学会、国际土壤信息参比中心、联合国粮农组织的认可和使用。所以, 在建立中国土壤系统分类土系数据库时, 我们应借鉴 SOTER 的原理和方法。

为了与 SOTER 数据库兼容, 在属性数据库内容的选择上参考了 SOTER 数据库的指标选择。SOTER 工作手册对数据库项目的设定分硬性的和可选择性的, 共 118 项, 土层数据项目有 55 项, 其中硬性必选的有 14 项。论述景观、土壤剖面数据的项目有 62 项, 地形体、土壤剖面和发生层等几个不同级别数据库都含有共同的项目, 如它们都有一些共同的土系图斑编码、土壤剖面编码等, 以便连接。参考上述 SOTER 数据库的内容, 结合土系的划分特点, 这里暂时列出图斑信息、景观环境数据库、典型单个土体数据库及土壤发生层数据库共计 70 余项, 其中除基本含盖了 SOTER 数据库所列的硬性选择项目以外, 还考虑到土系与农业利用密切相关的特点, 增加了土壤养分状况的有关项目, 样区还可根据其特殊情况增加特殊的项目。

2.2 主要技术要点

1 硬、软件环境

NASIS 特殊的操作环境(如 HP 9000/K200 workstations running HP-UX 10.20)在中国普通用户中普及程度较低, 考虑到目前我国有关单位和潜在用户的工作环境条件, 中国土壤系统分类土系数据库的建立可在以下主要硬软环境中完成。硬件环境: 数字化板或扫描仪等输入设备, P II 300 以上、内存 32 兆以上、硬盘 3G 以上微机, 绘图仪、打印机等输出设备。软件环境: ARC/INFO 3.4 以上空间数据库输入软件、ARCVIEW 3.0 以上空间数据库编辑、输出软件, Office Excel 或 Access 属性数据库录入、编辑、输出软件。

2 数据源的要求

(1)Tic 点要求转换为大地真实坐标 大地经纬网坐标是统一的地理坐标参照系, 无论何种形式数字化土系图件, 要将图件定位于大地经纬网坐标系中, 只有这样, 图件才有真实的距离和比例尺。要完成这项工作, 通常有两种做法, 一是在完成数字化输图之后, 在相同 GIS 环境下, 建立一个只有 Tic 点的“空”图层, 将其 Tic 文件改为大地经纬网坐标, 然后在 GIS ARC/INFO 中用 Transform 命令将图上坐标转换为真实的大地经纬网坐标; 二是在输图之前, 在确定 Tic 后, 即将其改为大地经纬网坐标, 接下来数字化输入的空间数据就在真实的大地经纬网坐标系中了, 在 ARCVIEW 3.1 版本, 可直接应用数字化(Digitizer)功能方便地完成这项工作。

在坐标转换过程当中, 需将地理坐标的分和秒的 60 进制换算成 10 进制, 通过换算公式 $00'/60+00''/3600$ 完成。以安徽宣城周王茶场 H-50-46-(27)1:1 万图幅为例, 用于定位的图幅四个对角点坐标由 $118^{\circ}37'30''$ 、 $30^{\circ}52'00''$ 、 $118^{\circ}37'30''$ 、 $30^{\circ}50'00''$ 、 $118^{\circ}37'30''$ 、 $30^{\circ}50'00''$ 、 $118^{\circ}41'15''$ 、 $30^{\circ}52'00''$ 换算后形成的 10 进制的 Tics 地理坐标为 118.62500 、 30.87500 、 118.62500 、 30.83330 、 118.68750 、 30.83330 、 118.68750 、 30.87500 。

(2)土系编码 美国土壤系统分类土系的编码由 6 个字符组成,前两位为州的代码,后面为土系序号,4 位字符的预留可保证土系编码的长期稳定,不至于像我们的电话号码那样,因为开始建库时预留不足常常“升位”改变。如美国 OR0011(Bakeoven 土系)、IA0033(Sharpsburg 土系)等,OR 为 Oregeon 州的代码,是经过注册而被公众认可的,0011 为该州土系的序号。需要说明的是,美国土系的编码与单个土体(Pedon)的编码完全是两回事,如 Bakeove 土系的编码是 OR0011,而 Bakeoven 单个土体的编码是 S89OR-065-001^[3]。

我国行政区域名是按照中国的各个行政区划划分而成,其划分标准依照原国家技术监督局发布的国家标准^[4]C 也可从 <http://www.msn.net.cn/services/area-domain.htm> 查得)而定,包括“行政区域名”34 个,适用于我国的各省、自治区、直辖市(表 1)。

表 1 中国行政区域名

BJ 北京市	SH 上海	TJ 天津市	CQ 重庆市	HE 河北省	SX 山西省	NM 内蒙古自治区	LN 辽宁省	JL 吉林省
HL 黑龙江省	JS 江苏省	ZJ 浙江省	AH 安徽省	FJ 福建省	JX 江西省	SD 山东省	HA 河南省	HB 湖北省
HN 湖南省	GD 广东省	GX 广西壮族自治区	HI 海南省	SC 四川省	GZ 贵州省	YN 云南省	XZ 西藏自治区	SN 陕西省
GS 甘肃省	QH 青海省	NX 宁夏回族自治区	XJ 新疆维吾尔自治区	TW 台湾	HK 香港	MO 澳门		

仿照美国土系编码的做法,我们在标准化行政区域名后面加 4 位字符,将其预留为土系序号,如江苏省第 1 个土系可编号为 JS0001。

(3)比例尺 理论上,在空间数据库的真实大地经纬网坐标确定以后,即可输出任意比例尺的图件,但是考虑到精度问题,要求输入底图的比例尺为 1:1 万至 1:5 万之间,这与常用地形图比例尺为相同,资料收集和操作上较容易实现。

3 数据库的结构

(1)空间数据库(几何数据库) 主要包括土系、水系、道路、居民点和行政区划界线。

(2)属性数据库 土系属性数据库由四个数据表组成:图斑信息表、景观环境数据表、典型单个土体数据表和土壤发生层数据表(表 2)。

数据库数据库管理与基本操作可参考文献[5]。

3 土系数据库的维护与更新

3.1 依据

土系数据库有些要素是相对稳定的,如岩性、地形等,有些则是变化的,如土地利用、土壤养分状况、植被等方面,一个实用的土系数据库应该反映现状,特别是土系的类型、数量随土壤调查的推进,一定会有些变化,如早期美国土壤系统分类中的 Miami 土系如今已经细分出几十个新的土系,因此土系数据库要有相应的更新,从国内外经验看,一些失败的数据库系统的原因之一,也是由于数据过时,无人问津,自然被搁置起来。所以,目前许多数据库系统,包括全球性的大系统,对数据更新都十分重视,在建库初期,就规定了数据更新的周期和途径

3.2 做法

国家级土系数据库的更新是有关根的,要具有权威性。美国农业部自然资源保护中心

NRCS 每年都颁布新版本的 NASIS, 以反映新的土壤调查结果, 再者, 它注重用户意见的反馈, 最新版 NASIS4.1.1 甚至增加了用户的联系信箱 (User email address), 用来倾听用户对数据库使用的想法。我们希望在不久的将来能建立起中国国家级土系信息参比机构, 对来自全国各地的土系起到确认、建立等“终审”作用, 并使之进入国家级土系数据库。特别是当前我国土系数据库刚开始建立, 参比、更新的工作是大量的, 更应慎重行事, 形成一个良好的开端。

表 2 土系属性数据库结构

图斑信息

1 制图图斑编号(见制图规范) 2 土系编码 3 暂定土系编号 4 土系名称 5 土系组成百分比(如果制图图斑中多于一个土系时) 6 表层质地 7 坡度 8 侵蚀度

景观环境数据

9 土系编码 10 暂定土系编号 11 建立日期 12 母质 13 地形 14 植被 15 排水情况 16 地表石质度(%) 17 土地利用状况

典型单个土体数据

18 土系编码 19 暂定土系编号 20 单个土体编码 21 经度 22 纬度 23 海拔高度 24 FAO 分类 25
CST 亚类 26 土族 27 采样日期 28 表层有机碳(g/kg)

土壤发生层数据

29 单个土体编码 30 发生层数目 31 发生层深度 32 诊断层 33 诊断特性 34 润态颜色 35 干态颜色
36 结构类型 37 粗碎屑的丰度(%) 38 过渡状况 39 总砂量(2.0~0.05mm) 40 粉粒(0.05~0.002mm)
41 粘粒(<0.002mm) 42 质地(美制) 43 容重(kg/dm) 44 水提 pH 45 盐提 pH 46 CEC (pH7.0时,
cmol/kg) 47 有机碳(g/kg) 48 全氮(g/kg) 49 全磷(g/kg) 50 盐基总量(cmol/kg) 51 盐基饱和度(%)
52 电导率(Ece, ds/m) 53 交换性 H⁺ (cmol/kg) 54 交换性 Al³⁺ (cmol/kg) 55 交换性 Ca²⁺ (cmol/kg) 56 交
换性 Mg²⁺ (cmol/kg) 57 交换性 K⁺ (cmol/kg) 58 交换性 Na⁺ (cmol/kg) 59 水溶性 Ca²⁺ (cmol/kg) 60 水溶
性 Mg²⁺ (cmol/kg) 61 水溶性 Na⁺ (cmol/kg) 62 水溶性 K⁺ (cmol/kg) 63 水溶性 CO₃²⁻ (cmol/kg) 64
HCO₃⁻ (cmol/kg) 65 SO₄²⁻ (cmol/kg) 66 Cl⁻ (cmol/kg) 67 (g/kg) 68 (g/kg)
69 (AL 水铝英石、CH 绿泥石、IL 伊利石、KA 高岭石、MO 蒙脱石、SE 三氧化二物、VE 蛭石) 70 Fe、Al、Mn
等全量分析数据 71

