

DOI: 10.11766/trxb201402080061

土壤专题图中采样点点位标识模型的构建*

张怀志 张维理[†] 徐爱国 冀宏杰

(中国农业科学院农业资源与农业区划研究所,北京 100081)

摘要 采样点图在直观展示已采集样点分布同时,也为确定补充采样点的位置提供了参考。制图时,通常以采样点编码在地图图面标识各个采样点,但由于受空间地理位置影响,采样点标识常呈无序排列,当地图中采样点密度及地图幅面较大时,读图时难以查找目标采样点。为实现制图后的采样点标识呈有序排列,构建了“土壤专题图中采样点点位标识模型(Soil Sampling Point Labeling Model for Thematic Soil Maps: SAMPLA)”,该模型主要有读图视区划分模型;土壤采样点归属读图视区判定模型;土壤采样点顺序标识模型等3个子模型组成。基于 ArcGIS 10.0,采用 C#语言进行计算机编程,实现了 SAMPLA。用 1:50 000 国家标准分幅和县级地图的土壤采样点对模型进行验证,验证结果表明,模型可适用于不同类型分幅地图、不同比例尺,有助于实现规范化、批量化制图,提高了读图效率;但是读图视区划分方案对标识结果有一定影响;其他行业专业领域,如环境科学、水科学、地质科学等在制作类似于土壤采样点的点位图层时也可采用此模型。

关键词 数字制图;土壤采样点;排序;模型

中图分类号 S159.9 文献标识码 A

土壤调查是土壤科学、农业科学与环境科学的主要研究方法之一^[1-3]。土壤调查主要目的之一是为了了解调查区域内土壤类型、土壤肥力与土壤环境质量状况及分布特征,通过在调查区域内多个点位上进行地面采样和观测,分析和测定采样点土壤理化性状,采用地统计方法生成土壤图、土壤养分图、土壤污染状况分布图等专题图^[4-6]。在土壤调查中,无论是进行专题图的地统计学检验还是进行土壤调查采样点设计均需要绘制土壤采样点图。通过采样点图可以直观地了解采样点点位空间分布现状,为确定补充采样点的分布位置提供参考。在土壤图制图中,各土壤类型理化性状特征可与剖面采样点图层土壤剖面点理化性状描述表相关联,含土壤调查结果的采样点图层能提供带有空间位置信息的土壤理化性状信息,是进行农业与环境主题科学研究的重要基础信息。

用土壤采样点坐标生成土壤样点分布图时,通常以采样点编码在地图图面上标识各个采样点。在土壤采样点图中,每个采样点均具有一个唯一的特定标识码,采样点标识码是查找、了解各采样点

点位特征的依据。当地图中采样点数量较少时,即使对各个采样点标识码不做特殊处理,仍易于在地图中根据采样点标识码找到某一特定采样点;但当地图中采样点数量较多而地图幅面又较大时,采样点标识码若未能呈有序排列,读图时很难找到所需的采样点。例如:在一个县域的土壤调查中,土壤采样点通常可达到上千个,当对调查结果进行汇总和生成土壤采样点图层时,若以土壤采样点原编码标识各采样点,受土壤采样点空间位置影响,采样点标识码在地图图面将呈现无规则分布,即编码相邻的采样点可能位于不同图区,要在数以千计的标识码中找到某个特定的代码则很困难。在专题制图中对空间要素注记算法研究已较多^[7-8],但对土壤样点分布图中采样点进行有序标识的方法、模型未见报道,本文研究目标是根据仿生学原理^[9]和读者阅读习惯^[10],构建采样点标识模型(SAMPLA),对土壤样点分布图中的采样点进行有序标识,以解决由于土壤样点分布图中的采样点标识码无规律,读图时难以找到特定采样点所在位置的问题。

* 科技部科技基础性工作专项(2012FY112100,2006FY120200)资助

[†] 通讯作者, E-mail: zhangweili@caas.cn

作者简介:张怀志(1968—),男,博士,副研究员,主要研究方向为土壤学、农业信息技术。E-mail: zhang_njau@163.com

收稿日期:2014-02-08;收到修改稿日期:2014-08-11

1 SAMPLA 模型方法原理

为便于读图时能根据采样点标识码在样点分布图中找到某一特定采样点的位置, SAMPLA 模型主要采用了两个原理:

(1) 仿生学原理: 对人类视觉的科学研究表明, 人的双眼辨别字的视线角度为 $10^{\circ} \sim 20^{\circ}$, 辨别字母的视线角度为 $5^{\circ} \sim 30^{\circ}$, 而认读视距范围在 38 ~ 76 cm 之间, 以此推算进行读图时, 视野为直径约 6.4 ~ 25.8 cm 的椭圆形。即人在读图时, 是按 6.4 ~ 25.8 cm 的椭圆形区域在土壤样点分布图的图面上逐区阅读、查找。

(2) 阅读习惯: 现代图书常见排版方式养成了读者横向从左至右、纵向从上至下顺序的阅读习惯, 查询图面上点位分布时也是如此, 本模型依据此阅读习惯确定读图视区的命名规则和采样点排序规则。

2 SAMPLA 模型的构建

SAMPLA 由三个子模型组成。一是地图的读图视区划分模型, 所谓读图视区就是地图幅面中与双眼视野范围 6.4 ~ 25.8 cm 大小相当的矩形网格, 是土壤专题图幅多层次网格划分完成的最低层级网格, 是土壤专题图上采样点规则排序的基本单元; 二是土壤采样点位归属读图视区判定模型; 三是土壤采样点顺序标识模型。

2.1 地图读图视区划分模型

该模型功能是将整个地图划分为若干个与人的双眼读图视野大小相似的读图视区, 包括两个步骤。

第一步以多层次网格分割完成地图图面读图视区划分。对于 1:50 000 国家标准图幅及其他比例尺标准图幅的多层次网格分割均采用四等分法; 而对于行政区域类地图图面如县级图、地市级图、省级图、跨行政区划图、流域图等, 第一步将地图图面分为若干个矩形(相关参数及其设置见表 1), 此后的矩形再依次采用四等分法, 直至接近于读图视区; 其中介于地图幅面与读图视区之间的各层级矩形网格, 定义为图区。读图视区划分模型如下:

$$SqN = \begin{cases} 4^{n-1} & n=2,3,4 \text{ 标准图幅} \\ a \times b \times 4^{n-1} & n=2,3,4 \text{ 非标准图幅} \\ & (\text{行政区域地图}) \end{cases} \quad (1)$$

$$a = \begin{cases} (TopLoc - downLoc) / LocDis & LocS < LocSD \\ (TopLoc - downLoc) / LocDis + 1 & LocS \geq LocSD \end{cases} \quad (2)$$

$$b = \begin{cases} (TopLat - downLat) / LatDis & LatS < LatSD \\ (TopLat - downLat) / LatDis + 1 & LatS \geq LatSD \end{cases} \quad (3)$$

式(1)~式(3)中 SqN 为最终读图视区数, n 为多层次网格分割的层级数, 由人工设定, 最多为 4 级, 1 级为图幅的包络矩形, 2 级为包络矩形再分区, 3 级是 2 级分区的四等分, 4 级是 3 级分区的四等分。 $TopLoc$ 、 $TopLat$ 、 $DownLoc$ 、 $DownLat$ 分别为图面左下角、右上角的经纬度坐标。 $LocDis$ 、 $LatDis$ 、 $LocSD$ 、 $LatSD$ 分别为第二等级网格分区内经向间距、纬向间距、最右区扩距、最下区扩距, 也由人工设定。 $LocS$ 、 $LatS$ 分别为经向、纬向分区后的剩余经纬度边距值, 以解决网格矩形划分过程中剩余区域的归属, 由程序自动生成; 当 $LocS < LocSD$ 或 $LatS < LatSD$ 时, 剩余的部分分别归属于最右边区域或者最下边的网格矩形; 当 $LocS \geq LocSD$ 或 $LatS \geq LatSD$ 时, 剩余部分则单独构成新的网格矩形。 a 、 b 分别为经向和纬向的矩形数量。 $TopLoc$ 、 $TopLat$ 、 $DownLoc$ 、 $DownLat$ 的值, 由海量分幅图角点坐标生成模型^[11]生成。

对地图进行读图视区的分割受到地图幅面大小、分幅类型、地图比例尺等影响, 在 SAMPLA 模型中引入了人机交互式的读图视区智能化分割方案设定表, 便于用户根据地图幅面特点, 为每一张地图幅面设定读图视区分割方案(见表 1)。

第二步计算各读图视区的角点坐标。国家标准分幅地图的读图视区划分均采用四等分法, 其读图角点坐标计算相对简单; 而行政区域地图 2 级图区是根据设定参数进行, 其后各级别图区、读图视区采用四等分法, 故行政区域地图图面的读图视区角点坐标计算主要是计算二级图区的角点坐标。国家标准分幅图的读图视区角点坐标 $P(X_A, Y_B)$ 计算方法如下:

$$X_A = DownLoc + A \times (TopLoc - DownLoc) / 2^{n-1} \\ A = 0, 1, 2 \dots 2^{n-1} \quad n = 1, 2, 3, 4 \quad (4)$$

$$Y_B = TopLat - B \times (TopLat - DownLat) / 2^{n-1} \\ B = 0, 1, 2 \dots 2^{n-1} \quad n = 1, 2, 3, 4 \quad (5)$$

式(4)、式(5)中的 X_A 、 Y_B 分别为读图视区角点经纬度坐标; A 、 B 分别为左上角开始计数时的经向、纬向 2 级图区个数, 其中 $A = 0$ 、 $B = 0$ 的点为图幅左上角点。

<http://pedologica.issas.ac.cn>

表 1 地图图面的读图视区分割方案设定表
Table 1 Scheme for map reading vision zone division of maps

分幅类型 Map types	Z2 区经向间距 Meridional spacing between Z2 zones	最右区扩距 ¹⁾ Enlarged distance in bottom of the map	Z2 区纬向间距 Zonal spacing between Z2 zones	最下区扩距 ²⁾ Enlarged distance on the right side of the map	Z2 区下再分区层级数 Number of levels of subdivision under Z2
行政界分幅 ³⁾					0
标准分幅 ⁴⁾					2

- 1) 最下区扩距是指经向按照经向间距进行分割后,为判定剩余的经向空间归属分区而设置的一个参数;
- 2) 右区扩距是指纬向按照纬向间距进行分割后,为判定剩余的纬向空间归属分区而设置的一个参数;
- 3) 行政界分幅“Z2 区经向间距”等 4 个参数由用户根据图幅幅面等设定,无举例;0 表示分级数为 Z0;
- 4) 标准分幅读图视区采用四等分法,仅需设置分级数,故相关区域置灰,2 表示分级数为 Z2

行政区域地图图面的读图视区角点坐标计算的重点是 2 级网格矩形(图区)角点坐标计算,方法如下:

$$X_A = \begin{cases} \text{DownLoc} + A * (\text{TopLoc} - \text{DownLoc} - \text{LocS}) / a & A = 0, 1, 2 \dots a - 1 \quad \text{LocS} < \text{LocSD} \\ \text{DownLoc} + A * (\text{TopLoc} - \text{DownLoc} - \text{LocS}) / a & A = 0, 1, 2 \dots a \quad \text{LocS} \geq \text{LocSD} \\ \text{TopLoc} & \text{其他} \end{cases} \quad (6)$$

$$Y_B = \begin{cases} \text{TopLat} - B * (\text{TopLat} - \text{DownLat} - \text{LatS}) / b & B = 0, 1, 2 \dots b - 1 \quad \text{Lat} < \text{LatSD} \\ \text{TopLat} - B * (\text{TopLat} - \text{DownLat} - \text{LatS}) / b & B = 0, 1, 2 \dots b \quad \text{LatS} \geq \text{LatSD} \\ \text{DownLat} & \text{其他} \end{cases} \quad (7)$$

式(6)、式(7)中的 X_A 、 Y_B 、 A 、 B 意义同上文。式(6)中“其他”指的是最右侧图区的右边两角的经度,式(7)中“其他”指的是最下边各图区的下边两角的纬度。

根据读图视区的划分规则,可知完整的 4 级网格分割中 3 级图区的角点坐标分别为 2 级图区左下角和右上角的经纬度差的平均值;读图视区的角点坐标分别为 3 级图区左下角和右上角的经纬度差的平均值。其角点坐标计算方法可参照式(4)和式(5)进行。

2.2 土壤采样点位归属读图视区判定模型

该模型功能是根据采样点空间地理位置及读图视区信息,确定土壤采样点归属的读图视区以及图区,为顺序标识做准备。土壤采样点归属读图视区及图区是通过图层叠加的方式实现的,同时赋予土壤采样点所在的各级图区码、读图视区码。

2.3 采样点顺序标识模型

该模型功能是在上两个模型的基础上,完成采样点顺序标识。根据读图视区及图区情况,采样点顺序标识分为采样点视区码标识、采样点标识码标识两个步骤。

采样点视区码标识是第一步,该步骤以读图视区为单元,顺次完成其内的所有土壤采样点视区码标识。如果一个读图视区内仅有一个采样点,则该采样点直接获得采样点视区码,但是如果一个读图视区内有不少于 2 个采样点,则用式(8)分别计算采样点与该级别读图视区左上角的距离($SptDisn$),依据距离长短依次赋予采样点视区码, $SptDisn$ 最小的采样点首先获得采样点视区码, $SptDisn$ 最大值的采样点最后获得采样点视区码。

$$SptDisn = \sqrt{(\text{Spt}X - X_A)^2 + (\text{Spt}Y - Y_B)^2} \quad (8)$$

式(8)中, $SptDisn$ 为采样点到读图视区左上角的距离。

第二步是进行采样点标识码标识,按照模型确定的图区码、读图视区码和采样点视区码顺序,依次进行采样点的顺序标识。

3 案例研究

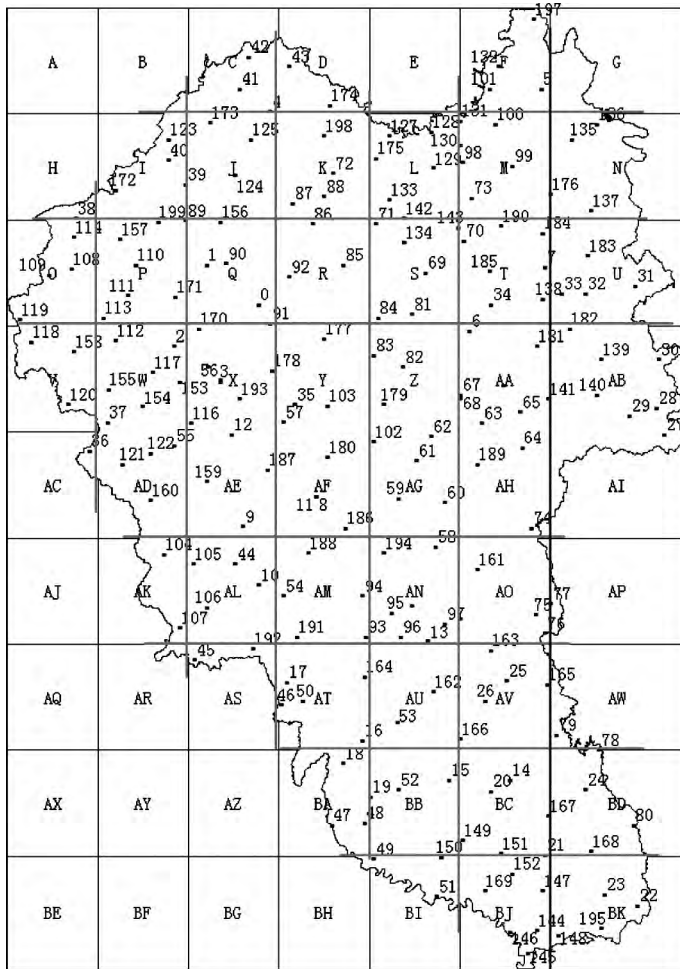
3.1 行政区域地图图面上的土壤采样点顺序标识

图 1 是 1:5 万比例尺四川省三台县土壤采样点图层的采样点标识结果,该图共有采样点 201 个,由于剖面点为无序排列(图 1 左,外框为模型计算用边界,内边界为实际边界),当通过采样点标识码查询到某一个采样点点位时,需要对全图所有采样点均查询一遍,花费时间较多。应用 SAMPLA 模型,全图采用二级网格划分方式,划为 63 个读图视区(约 18cm × 18cm),进行有序标识后,便于读者按上

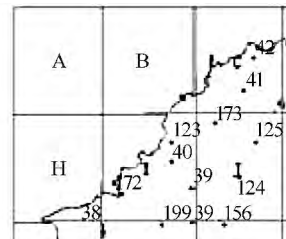
左下右两个层级排列的读图视区和视区内采样点的顺序标识码查找到采样点位置。

模型应用还表明: 同样是三台县的土壤采样点, 由于成图比例尺不同导致的地图大小发生变化, 进而影响到读图视区划分方案的选用, 因此会导致最终的土壤采样点标识码出现一定变化; 此

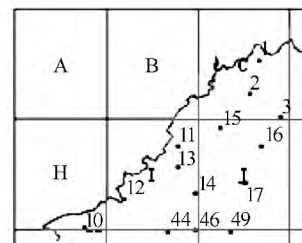
外, 即使是同一张地图, 虽然总读图视区个数相同, 如直接采用 2 级网格矩形作为读图视区和 3 级网格矩形作为读图视区, 由于读图视区码上一级的图区数量不同, 导致了采样点归属的网格矩形发生了变化, 最终导致采样点顺序标识结果出现微小差异。



三台县土壤采样点图(全图)
Santai County soil sampling point distribution map (Whole map)



SAMPLA分析之前(部分图区): 图区中采样点标识码为无序排列
Soil sampling point distribution map, random in labeling arrangement prior to application of the SAMPLA model (partial map)



SAMPLA分析之后(部分图区): 图区中采样点标识码按全图多层次网格划分和在读图视区内上左下右位置顺序排列
Soil sampling point distribution map with the soil sampling points labeled in order after application of the SAMPLA model (partial map)

图1 三台县土壤采样点图采样点顺序标识之前与 SAMPLA 顺序标识后比较

Fig. 1 Comparison between the Santai County soil sampling point distribution maps before and after the application of the SAMPLA model

我国省域、县域面积差别较大, 如新疆维吾尔自治区纬度差 $19^{\circ}10'$, 经度差 26° , 面积为 $1.66 \times 10^6 \text{ km}^2$; 而上海市纬度差 $1^{\circ}13'$, 经度差 $1^{\circ}13'$, 面积为 $6.3 \times 10^3 \text{ km}^2$; 县域中如新疆维吾尔自治区的若羌县纬度差 $5^{\circ}23'$, 经度差 7° , 面积为 $2.223 \times 10^5 \text{ km}^2$, 江苏省扬中县为 300 km^2 ; 且无论是省级、地市级还是县级行政区域, 再加上实际制图时可能出现的同一张地图跨行政区域情况, 区域形状千差万别; 因此确定行政区域地图网格分割方案时, 应根据每一张地图的具体情况, 兼顾地图纵横双向,

合理设置参数进行地图网格分割。值得指出的是, 读图视区的划分方案与地图大小、形状等特性有关, 采样点的密疏分布并不影响地图层级网格的划分。

3.2 标准分幅图上的土壤采样点顺序标识

图 2 是 1:50 000 标准分幅地图(图幅号: H50E006003)的土壤剖面点顺序标识结果, 该图幅共有 22 个剖面点, 含 4 个读图视区(约 $24 \text{ cm} \times 18 \text{ cm}$)。采样点理化性状以表的方式与采样点进行链接, 在此比例尺下, 能清晰分辨采样点图上的采

样点标识,虽然图幅不大,采样点数量较少,但由于剖面点为无序排列(图2左,边框为图幅实际边框),当要通过读图查询到某一个采样点位置需要对所有视区均查询一遍,而用 SAMPLA 标识后(图2右,外框为模型计算用区域,内框为图幅实际边框),读图视区内的剖面点实现了顺序标识,读者根据日常阅读习惯,能够比较容易地在样点分布图中查找到目标采样点的位置。

前文已经提及,读图视区是进行土壤采样点顺序标识的基本单元,是采样点顺序标识的开始,同样地,采样点视区码也是中间过程码,在最终成图

中均可以不出现,图2(右)展示读图视区图层是为了展示模型完整的运行结果。

1:50 000 标准分幅大小通常为 40~50 cm × 40~50 cm 根据人的双眼视野大小测算,多采用 2 级层级网格划分方案(第二级即为读图视区),个别地图采用 3 级层级网格划分方案就可有效地完成地图内的土壤采样点标识。实际制图例证还表明:1:10 000、1:250 000 等国家标准图幅,制图时的图幅尺寸与 1:50 000 标准图幅尺寸相当,故也多采用 2 级层级网格划分方案,然后进行土壤采样点顺序标识。

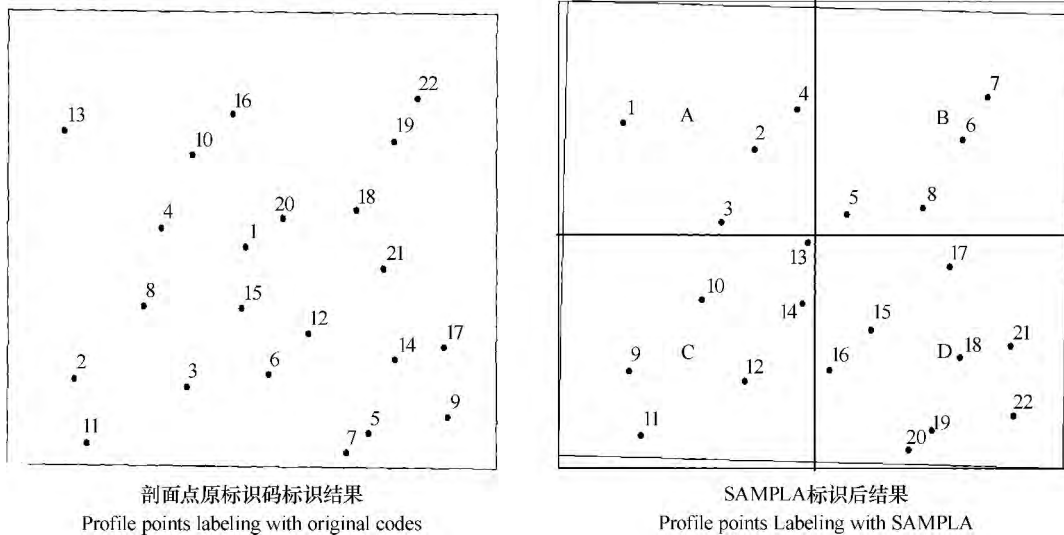


图2 对1个标准分幅的1:5万土壤剖面点分布图的剖面点进行SAMPLA标识的分析结果以H50E006003为例
 Fig.2 SAMPLA labeling of a 1:50 000 standard soil profile site distribution map-with H50E006003 as an example

4 结论

为进行采样点图中采样点的有序标识,SAMPLA模型采用了仿生学原理,按照人的视野对采样点地图进行读图视区划分,读图视区是进行地图采样点顺序标识的基本单元,其建立为在图面每个读图视区内相邻的土壤采样点获得顺序标识提供了基本条件。在对采样点图进行读图视区分割时,SAMPLA模型采用了多层次网格分割方法和上左下右的网格编码,以便在进行超大幅面、非标准幅面地图的读图时,便于读者按现代阅读习惯,在全图按从上至下、从左至右的顺序迅速找到特定采样点所处的读图视区。SAMPLA模型采用了人机交互方式进行图幅多层次网格划分参数设置,便于对不同类型幅面土壤采样点图样点进行顺序标识,采用可视化人机交互界面,用户既可采用系统默认参数由系统

自动完成采样点图中全部采样点的顺序标识;也可以根据地图幅面大小、图面采样点密度自行进行读图视区的分割。本文用县域土壤采样点图和1:50 000土壤剖面点图对模型进行了实测和验证,同时也对省域以及跨行政区的不同比例尺的采样点图进行了测试,表明模型适应于不同幅面类型、不同比例尺采样点图的采样点顺序标识。

参考文献

[1] 中国农业百科全书总编辑委员会土壤卷编辑委员会,中国农业百科全书编辑部. 中国农业百科全书—土壤卷. 北京:农业出版社,1996. China's agricultural encyclopedia editor committee soil science editorial board, China's agricultural encyclopedia newsroom. Encyclopedia of Chinese agriculture—Soil (In Chinese). Beijing: Agriculture Press 1996

[2] 章明奎. 土壤地理学与土壤调查技术. 北京:中国农业科学技术出版社,2011:280—333. Zhang M K. Soil geography and soil investigation techniques (In Chinese). Beijing: China Agri-

- cultural Science and Technology Press ,2011: 280—333
- [3] United States Department of Agriculture Natural Resource Conservation Service. Soil survey laboratory information manual. Soil Survey Investigation Report ,1995
- [4] Boettinger J L ,Howell D W ,Moore A C ,et al. Digital soil mapping: Bridging research environmental application and operation , Springer Science + Business Media B. V ,2010
- [5] Hartemink A E ,Alex M ,Maria L M S. Digital soil mapping with limited data. Springer Science + Business Media B. V ,2008
- [6] 杨琳,朱阿兴,秦承志,等. 一种基于样点代表性等级的土壤采样设计方法. 土壤学报,2011,48(5): 938—946. Yang L ,Zhu A X ,Qin C Z ,et al. A soil sampling method based on representativeness grade of sampling points (In Chinese) . Acta Pedologica Sinica ,2011 ,48(5) :938—946
- [7] 王立 郑昊. 粒子群遗传混合算法在点状注记配置中的应用. 计算机与现代化 2012(10) :30—33. Wang L ,Zheng H. A hybrid algorithm of PSO and GA for automatic placement of point annotation (In Chinese) . Computer and Modernization ,2012(10) :30—33
- [8] 樊红 刘开军 张祖勋. 基于遗传算法的点状要素注记的整体最优配置. 武汉大学学报: 信息科学版,2002,27(6) :560—565. Fan H ,Liu K J ,Zhang Z X. A robust genetic algorithm for automated map name placement(In Chinese) . Geomatics and Information Science of Wuhan University ,2002 ,27(6) :560—565
- [9] 中国科学院生物物理研究所. 生物的启示: 仿生学四十年研究纪实. 北京: 科学出版社,2008. Institute of Biophysics , Chinese Academy of Sciences. Apocalypse of biology: Bionics research forty years (In Chinese) . Beijing: Science Press ,2008
- [10] 张志强. 中国书刊文字排式的演变. 编辑学刊,1998,2:68—70. Zhang Z Q. Development of word array style in Chinese book and magazine (In Chinese) . Editors Monthly ,1998 ,2:68—70
- [11] 张维理. 智能化海量空间信息分析与地图制图软件包 IMAT 设计与构建. 中国农业科学,2014,47(16) :3250—3263. Zhang W L. Design and development of software package Intelligent Mapping Tools (IMAT) (In Chinese) . Scientia Agricultura Sinica ,2014 ,47(16) :3250—3263

DEVELOPMENT OF SOIL SAMPLING POINT LABELING MODEL FOR THE MATIC SOIL MAPS

Zhang Huaizhi Zhang Weili[†] Xu Aiguo Ji Hongjie

(Institute of Agricultural Resources and Regional Planning , Chinese Academy of Agricultural Sciences , Beijing 100081 , China)

Abstract Sampling site maps not only show intuitively distribution of the sampling sites where soil samples have been collected , but also provide reference for researchers to determine where supplementary soil samples should be taken. Traditionally , in soil mapping , codes were used to label each sampling point in the map for identification. However , as affected by their spatial geographic locations , it is impossible to label the sampling sites in sequence , thus making it very hard to locate a required sampling site in the map , especially when it has too many sampling sites located densely or is too large in format. To realize orderly labeling of sampling points , a “Soil Sampling Point Labeling Model for thematic soil maps (SAMPLA) ” was developed , consisting of three sub-models , i. e. map reading vision zone division model , soil sampling site attribution map reading vision zone determination model and , sampling point sequential labeling model. With the aid of ArcGIS 10. 0 and Human Machine Interaction(HMI) design , SAMPLA was translated into a computer program with the C# programming language , and tested with soil sampling points in the 1: 50 000 national standard map sheets and county-level maps. Results show that the model is applicable to maps different in type and scale , helps realize standardized mass mapping , and improves map reading efficiency. But labeling was slightly affected by the map reading vision zone division program. However , besides in soil sciences , SAMPLA can also be used in other scientific fields , such as environmental science , water science , and geological science , for making sampling point distribution maps.

Key words Digital mapping; Soil sampling point; Labeling; Model

(责任编辑: 檀满枝)

<http://pedologica.issas.ac.cn>