

GIS 空间数据的质量探讨

谭红霞

(山东省国土测绘院, 山东 济南 250013)

摘要:空间数据直接影响着 GIS 应用分析结果的可靠程度和应用目标的真正实现,因此,有效地控制空间数据的质量是非常必要的。该文结合山东省1:1万基础地理信息数据采集与建库实践,分析了空间数据质量问题的来源及影响,根据空间数据质量标准,采用软件检查和人机交互检查2种方法,制定各工序相应的检查方案,以减少误差,防止粗差。

关键词:GIS 应用;空间数据;数据质量

中图分类号:P208

文献标识码:A

空间数据是 GIS 最基本和最重要的组成部分,是 GIS 操作的对象,也是 GIS 的核心。数据质量的好坏是基础地理信息数据库建设成败的关键,它可直接影响着 GIS 应用分析结果的可靠程度和应用目标的真正实现^[1]。因此,探讨如何控制空间数据的质量是非常必要的。该文主要对空间数据质量的标准以及数据的检查内容、误差源、检查方法方面进行研究,并结合正在进行的山东省基础地理信息1:1万数据采集、更新与建库项目的实施,对基础地理信息数据生产过程中的质量控制进行探讨。

1 空间数据质量的概念及研究目的

(1)空间数据是指用来表示空间实体的位置、形状、大小及其分布特征诸多方面信息的数据,也可以说是对实体与属性的抽象表达。它可以用来描述来自现实世界的目标,具有定位、定性、时间和空间关系等特性^[1]。定位是指在已知的坐标系里空间目标都具有唯一的空间位置;定性是指有关空间目标的自然属性,它伴随着目标的地理位置;时间是指空间目标随时间的变化而变化;空间关系通常用拓扑关系表示。把空间数据用点、线、面以及实体基本信息等作为基本空间数据结构,来表示人们赖以生存的自然世界。

空间数据质量是空间数据在表达空间位置及其

属性时,所能够达到的准确性、一致性、完整性以及他们之间的统一程度。通常是用空间数据的误差来度量的,误差是指数据与其真值的偏离。由于人类认识和表达能力的局限性及现实世界的复杂性与模糊性,这种抽象表达不可能完全达到真值,只能在一定程度上接近真值。

(2)GIS 数据质量研究的目的是建立一套相对完整的空间数据分析和处理的体系,包括误差源的确定、误差的鉴别、误差的度量及分析方法、制定控制和削弱误差的方案等;从应用的角度讲,研究 GIS 空间数据质量可直接关系到地理信息系统最终运行效果,进而关系着基于地理信息系统的各种宏观决策的正确制定及信息化进程的推进^[1]。

2 空间数据质量的标准

空间数据质量标准是生产、使用和评价空间数据的依据,数据质量是数据整体性能的综合体现。山东省基础地理信息1:1万数据采集、更新与建库项目根据国家标准,结合该地区 GIS 空间数据的需求及数据生产的特点,研究并制定了空间数据质量标准,以此作为对空间数据进行检查的依据。

(1)数据档案:即空间数据的元数据,它是对空间数据情况的一种说明,要求对空间数据的来源、处理及相关数据信息等方面进行较全面、详尽、正确的

* 收稿日期:2009-02-24;修订日期:2009-06-04;编辑:陶卫卫

作者简介:谭红霞(1956—),女,山东济南人,工程师,主要从事技术和质量管理工作。

描述。元数据项主要包括:元数据名称、地理坐标系、高程基准、数学投影方式、比例尺、数据格式、质量评价等等。

(2)位置精度:即空间实体的坐标信息与真实地理位置的几何误差。规定以具有三维地理坐标的点、线、面作为研究对象以表达空间实体的几何形状,其定位精度表现为空间三维坐标数据的精度。其中包括:数学基础精度、平面精度、高程精度、接边精度、几何图形精度、像元定位精度(分辨率)等。

(3)属性精度:是指空间实体的属性值与真实值的相符程度。属性精度通常取决于数据的类型,且与位置精度有关。主要包括要素分类和编码的准确性、要素注记名称的准确性、属性字段和属性值的准确性等。

(4)逻辑一致性:是指空间数据元素之间逻辑关系的一致性,包括数据结构、拓扑关系、空间特征、时间特征以及属性特征的一致性。

(5)数据的完整性:是指空间数据在范围、内容及结构等方面的完整程度,包括数据范围、数据分层、数据类型、属性表结构、实体空间关系、属性分类、空间实体描述、属性字段描述、实体与属性对应关系的完整性等等。

(6)数据的现势性:即数据的时间精度,是指空间数据时间信息的可靠性。因空间数据的更新周期较长,历史数据和现势数据存在较大的差异,这直接影响了空间数据的有效应用,可以通过记录数据更新的时间和频率等来表示。

(7)栅格数据质量标准:该文所指栅格数据为DOM和DEM,其质量标准除包含以上内容外还包括像元值反映相应区域的光谱信息(即影像)和高程信息的正确性。

(8)接边要求:指图幅间的接边。采用任何方式采集的数据,在图幅的接边处均不允许出现逻辑裂隙或重叠。

3 空间数据质量问题的来源及影响

空间数据质量问题,实际上是伴随着数据的采集、处理、转换过程而产生并表现出来的。在数据库的建设过程中,有许多生产环节,每个环节均产生一定的误差,按误差传播理论,每项误差的传播直接影响到最终数据的质量。

3.1 数据源的质量问题及影响

数据源一般包括原始遥感影像、控制点数据、外业实地测量数据、实地调绘资料及信息调查成果等。这些数据源在获取的同时会产生不同程度的误差,如:遥感的系统误差及干扰误差会带来影像变形、清晰程度、地面分辨率的质量问题;外业测量时测量的方法、GPS的信号干扰、仪器及人员操作误差会给后面工序带来量测精度的问题;调绘及属性调查时产生的错误会产生数据使用可信度的问题等等。另外,数据源在时间精度(现势性)不能满足数据内容需求时,也会影响系统数据和应用结果的质量。数据源的质量问题有些是不可逆的,这些误差或质量问题会影响GIS数据的质量。

3.2 数据采集、编辑质量问题及影响

由于技术、设备条件、人员操作水平的限制,在数据获取、数字化采集、属性录入及格式转换时会带来不同程度的误差。这些误差来自数字化设备的精度、采集的方法、数字化的操作误差、属性数据录入中的差错等,会造成数据精度的质量问题。相比数据源的质量问题,这部分数据质量问题相对比较简单,影响因素容易发现,只要引起足够的重视,可控制程度相对较高。

3.3 数据格式转换及处理产生的质量问题及影响

要使基础数据满足数据库的要求,就要把数据在不同的文件格式之间进行转换,在数据转换过程中由于各系统内部的数据结构不同和功能差异,往往会造成信息的损失,包括数据精度上的损失。

数据在多边形关系、拓扑关系的处理中,也会产生位置和属性方面的质量问题。以上质量问题是由于软件系统内部引起的,问题较为复杂,影响因素较隐蔽,控制起来比较困难。

4 空间数据质量的控制

空间数据质量的控制要从数据质量产生和扩散的所有过程和环节入手,根据数据生产的程序制定各工序相应的检查方案,以减少误差,防止粗差。

山东省基础地理信息1:1万数据的采集、更新与建库,是基于DMC数码航摄资料和常规的胶片航摄资料,采用全数字摄影测量方法。

4.1 数字线划图(DLG)的质量检查

(1)数学基础检查:采用检查软件自动检查、人

机交互检查的方法检测内图廓点、公里网线交点坐标是否正确,属性内容是否正确。

(2)接边检查:采用检查软件自动检查和人机交互检查,相邻图幅接边要素在逻辑上应保证无缝接边;相邻图幅接边要素属性应保持一致;相邻图幅接边要素拓扑关系应保持一致。

(3)属性精度检查:采用检查软件和人机交互检查的方法,按设计相关要求对采集的各类要素属性进行检查,检查属性值是否漏赋或赋值错误;检查各项要素所赋的属性字段是否正确;检查各元素所属类别是否发生归类错误,确保元素分类分层的正确。

(4)数据的完整性与正确性的检查:采用检查软件检查文件命名的正确性,检查数据格式、数据组织、数据有效覆盖范围以及要素内容是否符合设计要求。

(5)逻辑一致性检查:采用检查软件和人机交互检查的方法,检查拓扑关系的正确性及各要素间关系是否合理协调;检查点、线、面要素拓扑关系建立是否正确;线划相交情况是否被错误打断;同一要素有无重复数字化的情况;是否出现不合理的悬挂结点、伪节点以及其他错误。

(6)现势性检查:采用人机交互检查的方法检查数据的现势性,通过矢量数据与最新的卫星正射影像叠加对发生变化区域进行更新的情况评定其现势性。

(7)文档质量检查:以元数据样本为依据,利用检查软件检查元数据文件数据项的完整性和正确性。对文档簿采用人工检查的方法,对其内容进行逐项检查,检查文档簿中必填项是否完备,是否符合文档簿的填写要求。

4.2 数字正射影像图(DOM)的质量检查

(1)数学基础检查:利用检查软件检查影像数据文件地理定位的栅格坐标与地理坐标、像元地面尺寸、行列数等是否正确。

(2)接边检查:利用检查软件及目视检查的方法,检查相邻正射影像图的影像接边,检查图幅之间DOM影像图重叠区域同名地物点或线状地物影像的重合情况,其误差是否符合要求。

(3)影像质量的检查:用人工目视的检查方法,检查影像是否清晰易读、反差是否适中、色调是否均匀一致,影像上是否存在因作业不当而造成的重影。

(4)数据的完备性与正确性的检查:利用检查软件检查文件命名的正确性,检查数据格式、数据组织、影像数据有效覆盖范围是否符合设计要求。

(5)现势性检查:检查数据的现势性,通过影像的获取(摄影)日期评定其现势性。

(6)文档资料的检查:以元数据样本为依据,利用检查软件对元数据文件数据项的完整性、正确性做出检查。对文档簿采用人工检查的方法,对其内容进行逐项检查,检查文档簿中必填项是否完备,是否符合文档簿的填写要求。

4.3 数字高程模型(DEM)的质量检查

(1)数学基础检查:利用检查软件检查DEM图幅的图廓数学基础、格网间距、格网行列数等是否正确。

(2)接边检查:利用检查软件及目视检查的方法,检查相邻图幅的地貌形态及重叠区域内同名点(X, Y 坐标相等)高程较差是否符合要求。

(3)高程信息精度检查:采用人机交互检查的方法,利用高程分带图(分层设色)高程与色表的对应关系,使用软件的“查询工具”直接读取当前位置的高程值,通过与周围高程注记点的比较,分析DEM的高程精度。

(4)外业散点检查:采用实地检测的方法检验其高程精度。一般每幅图需随机抽查50个点。将每个检查点的高程值与DEM文件中对应平面位置的内插值比较,计算出误差值,分析DEM的高程精度。

4.4 叠合检查

将具有相同坐标系统和坐标值的DLG, DOM, DEM叠合在一起,来实现检查3D数字产品的几何精度、数据完整性及一致性的目的。

5 结论

由于GIS空间数据类型多种多样,如何在空间数据生产过程中将存在的质量问题迅速地查找出来,并及时地进行编辑修改是质量控制的关键。本文对控制空间数据质量方面的探讨及提出的检查方法,在山东省1:1万数据采集、更新与建库项目建设的实践中,能有效的控制空间数据质量的大多数问题,但由于所探讨的问题本身固有的复杂性、随机性、动态变化性,因此仍然存在一些问题需要研究、

探讨,如:建库数据与一般数字地形图的根本区别、寻找出一个更有效的解决方法,使空间数据的质量进一步提高。要保证空间数据的质量,在实施检验前,首先要考虑以下几个方面,然后确立相应的方法、步骤,以便确保质量,提高工作效率。

(1)首先要了解数字产品的需求及特性,了解数据的标准,了解要素所要表达的形式,制定一个切实可行的检查方案。

(2)检查员要了解数据制作所采用的软件及制

作的基本过程,从而了解作业中容易出错的地方,以引起注意,找出简便易行的查错方法。

(3)充分利用、开发软件的检查功能,结合逻辑推理排除错误存在的可能,以最大限度地减少差错,提高检验效率。

参考文献:

- [1] 胡鹏,吴艳兰,李圣权,代琪. GIS 远程教育[M]. 北京:高等教育出版社,2003.

Study on Quantity of GIS Spatial Datas

TAN Hong - xia

(Shandong Surveying and Mapping Institute, Shandong Jinan 250013, China)

Abstract: Spatial datas will affect the reliability of GIS application results and the realization of application goals. Therefore, it is very necessary to control the spatial datas quality effectively. Combining with the collection of basic GIS datas with the scale of 1:10000, the origins which will effect the quality of spatial datas are analyzed in this paper. According to quality standards of spacial datas, by using software checking method and human and computer cross - examination method, appropriate countermeasures should be made to reduce the error and prevent gross errors.

Key words: Application of GIS technology; spatial datas; datas quantity