

技术方法

区域地质图数据库建库中的质量检查方法探讨

蔡青,马瑜宏,马丽新

(山东省地质调查院,山东 济南 250013)

摘要:数据质量检查是地质图空间数据库建设项目中的一项重要内容,也是一项复杂的过程,应该贯穿数据的输入和空间数据库建立的每一个环节,检查技术方法和手段也一直进行更新。该文以1:5万区域地质图空间数据库建设为例,列举了建库过程中不同阶段的多种的检查方法和手段,对于不同内容使用不同的检查方法进行探讨,通过人工检查及软件检查2种方式来确保入库数据的质量。

关键词:区域地质图;空间数据库;数据质量检查;MapGIS

中图分类号:P208

文献标识码:B

0 引言

区域地质调查是一项综合性的基础地质工作,随着区域地质调查的不断深入,地质信息日趋庞大,地质资料急剧增加,纸质地质图和纸介质载体信息存储方法难以满足区调发展需要,为充分、迅速、有效地利用基础地质数据,输出地质成果,开展了1:5万区域地质图空间数据库建设项目,数据库属于回溯型建设,即在原有资料的基础上进行数字矢量化^[1]。该项目以MapGIS软件为平台,自开展以来已经形成了一整套比较完整的作业流程和规范^[2]。

数据的质量是数据库的关键,将直接影响数据库的可用性和精确性,错误的或不准确的数据信息将会产生错误的处理结果,导致错误或不准确的决策^[3]。因此数据的质量控制、数据库的质量检查工作非常重要。数据质量检查一直是地质图空间数据库建设项目的重要工作内容,对数据质量有着严格要求,检查技术方法和手段也一直进行更新,来确保数据质量的精确性。空间数据质量问题是伴随着数据的采集、处理、转换过程而产生的,每个环节产生的误差都直接影响到最终数据的质量^[4]。所以该文对1:5万区域地质图空间数据库建设中不同阶段的多种检查方法进行探讨。

1 资料收集、预处理质量检查

资料的预处理是为了确保入库资料的质量,在入库前对收集到的图形资料进行精度检查,对属性资料进行正确性检查,对图形资料和属性资料进行一致性检查而进行的一项工作。对于质量不高的数据源,通过预处理工作不但可减少误差,还可提高数字化工作的效率。

首先将收集的原始图件,交与地质专家进行检查审核,将其中的明显错误:如明显的标记错误、明显的线划错误及缺失、设色错误以及地质体时代或属性标注错误等非观念认识差异造成的错误等预先进行修改,矢量化时以修改后的为准,可以减少因发现错误时需要重新进行拓扑而进行的大量重复性工作,而重新拓扑可能会造成线弧不套合等更多严重错误。

2 数据建库时的质量检查

空间数据的质量控制是一项复杂的过程,应该贯穿数据的输入和空间数据库建立的每一个环节^[5]。地质图空间数据和属性数据的错误类型主要有:空间数据的不完整或重复,包括点、线、面数据的丢失或重复,数据断线、多边形不封闭等;图层文

收稿日期:2013-04-24;修订日期:2013-07-06;编辑:陶卫卫

作者简介:蔡青(1979—),女,山东定陶人,工程师,主要从事地质矿产信息技术研究工作;E-mail:wrka_qq@sina.com。

件、数据表等不完整;空间数据位置不准确,点位不准确,线段过短,结点不重合;拓扑关系错误;属性与空间数据联接错误;属性结构、字段类型及长度与规定不符,属性值数据错误或不完整;图面符号表达不清晰、压盖关系不合理;图框外整饰不规范等等。

2.1 矢量化过程中数据质量检查

对于1:5万地质图,因图面内容较多,必须手工进行矢量化,才能保证数据精度。在矢量化过程中要尽量减少线段上的重复点、Z字线段、自相交或线段缺失等错误,需要成区的线相交一定要建立结点。

首先需要人工对矢量化完成后的图件与原始地质图进行对比,补齐漏掉的点和线图元,对于字符、子图大小和角度及线型,线宽等进行修改,尽量接近原图。然后使用MapGIS软件中的“图形处理”模块的“其他”工具项,进行“清除微短弧线、清重点坐标及自相交、检查重叠线”等检查。如果发现微短线,应放大图件仔细查看原因,有可能是单独一条微短线,删去就可以,也有可能是线段自相交,剪断后出现的微小线段,这就需要整条线段重新整理。重叠线,重叠坐标等错误,可以直接使用MapGIS编辑系统自有的检查修改功能进行删除。

由于水系的特殊性,其线方向即代表水系的流向,要求矢量化水系从上游到下游。矢量化过程中或后期处理时误操作都可能导致线方向错误。可以在MapGIS编辑系统内选择显示线方向,结合地形地貌判断其方向的正确性并作相应的修改。也可在等高线和高程点图层附属性值后,使用1:5万地质图空间数据库建设总项目组提供的基于MapGIS二次开发的应用程序GISEditTool进行检查。方法是同时打开等高线、高程点和水系线文件,并将水系线文件置于当前编辑状态下,执行辅助工具中“水系线方向初步判别”功能,软件即会提示可能的方向错误,检查后确定是否存在错误。如果出错,利用“改线方向”功能进行改正。项目要求属性库中的等高线必须连续,该软件也可以对等高线是否封闭进行检查。

2.2 拓扑处理后数据质量检查

拓扑类的错误可能是由于矢量化操作不当或拓扑前的检查不仔细导致,也可能是拓扑处理完成并划分了不同的图层文件后操作不当引起。拓扑类的主要错误有:图层套合检查;同一图层弧段和界线的

套合;不同图层间的套合,主要是水系和地层,断层和地层,断层和岩体,地层和岩体界线之间的套合关系;检查水体、海岸线与地质体边界线是否重合,是否有无效弧段和悬挂弧段等。

检查此类错误,可以利用MapGIS自带“区拓扑错误检查”功能,但是有些错误不会被检测出来。利用GISEditTool程序中“辅助工具”,可以基本查出出现的各种错误,并且能够根据查出的错误进行一定程度的修改。主要检查内容包括“检查并更正拓扑套合关系、检查区拓扑一致性、线或弧段Z字形、自相交批量检查、清除重叠坐标、自动删除多余弧段等。检查时需要选择标准文件,一般应选中地质界线图层文件,以它为基准对其他它线、区文件进行一致性检查。

对于查出错误后的修改,需要在MapGIS工程文件中同时打开错误报告文件和所有相关的线、区文件(包括地质界线、水系、断层、水体、沉积岩、变质岩、非正式地层单位、侵入岩、脉岩、构造变形带等参于拓扑的文件),取消还原显示,并选中显示坐标点,将窗口放大到清楚地看到错误为止。对于线弧不一致错误,通过线弧结点平差,移动线段或弧段上坐标点进行微调。如果是出现多个线和区图层不套合时,线和弧段应同时修改,否则会出现更多的线弧不一致错误,或者是重新进行拓扑。

2.3 属性挂接后数据质量检查

属性类错误是在图层分层完成后进行属性录入过程中产生的与属性有关的错误。由于地质图数据库涉及图层文件很多,每个文件均有不同的属性结构,每个属性结构又存在很多字段,每个字段的类型、大小不同,字段名称均为字母,在录入过程中容易出现错漏现象。属性录入时项目要求,同一字段项中汉字之间的连接符用全角符号,英文与数字之间的连接符用半角符号,因而需要在录入属性时仔细录入。

属性挂接时出现的主要问题有图层分层错误,属性与空间数据联接错误;属性结构、字段类型及长度与规定不符,属性值数据错误或不完整;同一地质体,属性内容不同;图元编号不唯一等错误。

对于属性内容的检查,可以利用MapGIS中的检查模块。打开工程文件,在被检查文件前打上“√”,设为当前编辑状态,在检查工具栏中选择“工作区属性检查”,出现检查属性内容对话框,单击字

段名称,右侧属性内容框内显示所有该字段属性值(图1),从这里可以检查出建库要求的必填项中的属性内容是否有空值,如果有空值的话,需要查资料来补充完整。双击属性内容中的任一项,赋该属性内容的图元闪烁,就可以清楚的看到属性值与图面内容,注释是否一致,图元设色及大小设置是否统一等问题。

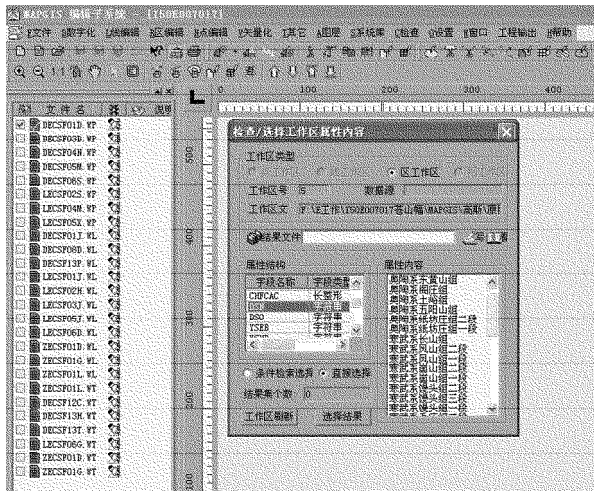


图1 工作区属性检查示意图

检查相同地质体的属性内容是否一致,需要输出属性表来核对。一种方法是利用 MapGIS 自带的“属性库管理”输出属性表,也可以利用工具软件 GISTools 中的“属性文件转换”功能。打开生成的 DBF 属性文件,选择不重复的记录,进行筛选。同一地质体应该只有一条记录,如果出现多条,就证明其中有些图元属性是错误的,需要修改。同时也可以就这些属性项与地质专家填写的属性表来核对,进一步检查属性内容的填写是否准确。

MapGIS 软件自带模块不能对属性结构类数据错误进行检查,必须依靠项目组下发的检查软件。Check_MapGIS 程序可以对属性结构进行检查,它要求在 D 盘建立事先编辑好的“属性结构.TXT”,这个文件中有全部带属性结构的图层文件名,字段名称、类型和大小。选择输出路径,选择文件属性检查,指定被检查文件,计算。在界面右侧方框出现检查结果,根据结果进行修改。GISEditTool 程序可以对属性中的空格进行检查和删除,还可以对图元编码进行自动整理,避免了图元编码重复,不唯一这类错误出现。

2.4 图面内容检查

输出整图,与原始图件进行对照检查,此项工作

只能由人工进行检查,主要是针对区的设色,线段类型、线宽大小及子图、字符的角度等等,是否有错漏,位置偏移等问题与原图进行比对。在屏幕上显示的颜色与输出后纸质上的颜色有差别,必须根据纸质上打印的色进行微调,使图件看起来美观,且尽量与原图保持一致。

3 利用软件对建库后的全面检查

对于数据库的检查和修改,仅应用 MapGIS 主控程序来进行不够方便和完善,在检查过程中还会有人为的错误产生,且工作量非常大。针对这种情况,为了提高数据的质量,开发了针对数据库的质量检查软件,地质数据质量检查与评价系统 Geo-Check。

在引入规则时选择“1:5 万地质矿产图库标准”,即可以对 1:5 万地质图数据库进行包括拓扑关系和属性结构内容的详细的全面的检查。进入界面后,选择“空间数据检查”,“空间拓扑集成检查评价”界面(图2),执行“检查与评价”。结束后,生成 2 个 Excel 文件,一个是检查评价报告,一个是空间数据检查结果。



图2 空间拓扑集成检查评价示意图

检查评价报告是对查出的缺陷数进行统计,得出分数,显示质量检查是否通过。空间数据检查包括“拓扑结点检查、重叠点坐标检查、Z 字线检查、自相交线弧检查、多余弧段检查、重复点线面检查、文件压缩存盘情况检查、套合一致性检查、拓扑一致性检查、等高线封闭性检查、图幅边界矢量数据精度检查、微小线段或弧段检查、水系方向判别检查、属性结构检查、属性代码检查、数据值域检查、图元唯一性检查、接触关系检查、断层属性检查、产状属性检

查、数据填写率检查、成果目录检查、图层文件命名情况检查、多套成果一致性检查、投影参数、误差校正、其他检查”27项内容。从以上罗列的条目就可以看出,这个检查与评价系统对数据库进行了全面细致的检查,基本涵盖了建库中所有数据质量检查的内容,所以说这是建库后必须要做的一次检查,可以将所有错误一次性检查出来。在每个检查结果表内都会显示详细的错误信息,方便找到出现错误的图层或图元来完成修改,可进行多次检查,直到结果符合项目要求为止。通过应用此软件的检查,将属性数据库中的错误率降至最低,从而保证了入库数据的质量。

4 结语

数据质量检查是区域地质图空间数据库项目工作中的一项重要内容,数据的质量直接影响到数据的更新和应用,为了确保数据库的质量,数据库的检查工作是非常必要的。但同时数据质量检查和修改也是一项非常复杂和费时的的工作,对于复杂的地质

图空间数据库的检查,应用GIS技术,研发新的检查程序进行软件检查,可以提高检查速度和精度,减少检查过程中的人为因素产生的错误,更好地保证数据的质量。该文以1:5万地质图空间数据库建设为例,列举了建库过程中不同阶段的多种的检查方法和手段,希望对以后的数据库建设的数据质量检查和质量控制提供一些参考。

参考文献:

- [1] 李定平,胡光道,程咱. MapGIS下空间数据库的建立及其典型问题研究[J]. 武汉大学学报, 2005, 30(11): 1029-1032.
- [2] 陈爱明,柯育珍,周录英. MapGIS地质图空间数据库建设常见错误与分析[J]. 资源环境与工程, 2008, 22(5): 543-546.
- [3] 谭红霞. GIS空间数据的质量探讨[J]. 山东国土资源, 2009, 25(6): 28-31.
- [4] 姜作勤. 数据质量研究与实践的现状 & 空间数据质量标准[J]. 国土资源信息化, 2004, (3): 23-28.
- [5] 陈述彭,鲁学军,周成虎. 地理信息系统导论[M]. 北京: 科学出版社, 1999.

Study on Data Quality Check Method in Regional Geological Map Spatial Database Construction

CAI Qing, MA Yuhong, MA Lixin

(Shandong Geological Surveying Institute, Shandong Jinan 250013, China)

Abstract: Data quality check is an important and complicated part in geological map spatial database construction project. It should run through each aspects of data input and database construction. Methods and means of the check technique have been under continuous updates. In this paper, setting regional geological map spatial database construction with the scale of 1:50000 as an example, different methods and means in different stages of the database construction process have been illustrated. Various methods should be used in data quality check. By using manual check and software check, the quality of inputting data can be guaranteed.

Key words: Regional geological map; spatial database construction; data quality check; MapGIS