

矿产资源潜力评价成果图空间数据库中属性数据质量检查探讨

刘福魁, 马莉, 蔡青

(山东省地质调查院, 山东 济南 250013)

摘要:属性数据是矿产资源潜力评价成果的重要组成部分,从某种意义上说,属性数据质量直接影响矿产资源潜力评价的效果。基于此认识,通过分析属性数据填写的原则、技术要求,提出了采取软件和人工两种方式检查属性数据质量的方法。

关键词:数据库;地理信息系统;矿产资源潜力评价;数据质量

中图分类号:P208

文献标识码:B

0 引言

《山东省矿产资源潜力评价》工作项目隶属于《全国矿产资源潜力评价》计划项目,是山东矿产资源潜力的一次重要省情调查。项目以大陆动力学理论为指导,全面总结基础地质调查和矿产勘查工作的成果和资料,应用现代矿产资源预测评价的理论与方法,全过程使用地理信息系统(GIS)技术,完成了山东省煤、铁、金等 23 个矿种的成矿规律、预测研究工作,圈定出成矿远景区和找矿靶区,并逐一评价其资源潜力^[1,2]。

项目中应用 GIS 技术处理的不是一般的数据,而是空间数据,即具有空间分布位置特征的空间数据以及与空间位置有关的属性数据。GIS 把空间数据抽象为可以用图形表达的点、线、面(多边形),以图层文件的方式进行分层组织;并将有关的属性数据与其关联起来,视二者为一体进行管理。属性数据主要用于存放空间数据的特征属性信息,是查询、更新、检索等操作的基础^[3-6]。

鉴于属性数据在矿产资源潜力评价项目中所处的重要地位,该文在分析属性数据填写要求的基础上,探讨了软件和人工对其质量进行检查的特殊性,

提出了相应的方法。

1 与空间图层文件关联的属性数据

山东省矿产资源潜力评价项目成果,主要体现在应用 GIS 技术所编制的各种专题图件中。按照全国矿产资源潜力评价技术要求,最终的成果图要应用 GIS 技术建立空间数据库。在项目实施过程中,主要使用武汉中地公司的 MapGIS 软件和项目办专门开发的 GEOMAG 软件。图 1 就是在 MapGIS 软件中放大显示的“山东省 1:25 万青岛市幅实际材料图”的部分图面内容。

一幅“山东省 1:25 万青岛市幅实际材料图”,由图 1 软件主窗口左侧图层文件列表栏中“岩性(面)”、“地质界线(线)”、“同位素年龄(点)”等 49 个图层文件构成。

“岩性(面)”中的“莱阳群曲格庄组”地层单元在图幅中共有 86 处出露区,每一处出露区都是“岩性(面)”图层文件中的一个图元,每个图元都要按照全国矿产资源潜力评价数据模型^①中规定好的属性数据项填写专业内容。

为了能够全面完整地描述图元所表达的地质专业内涵,数据模型对“岩性(面)”图层文件中的图元,

* 收稿日期:2012-06-08;修订日期:2012-11-28;编辑:曹丽丽

作者简介:刘福魁(1971—),男,山东栖霞人,高级工程师,主要从事信息技术和地质矿产研究工作;E-mail:lfk74@126.com。

①中国地质科学院矿产资源研究所、中国地质调查局发展研究中心,全国矿产资源潜力评价数据模型,2009年。

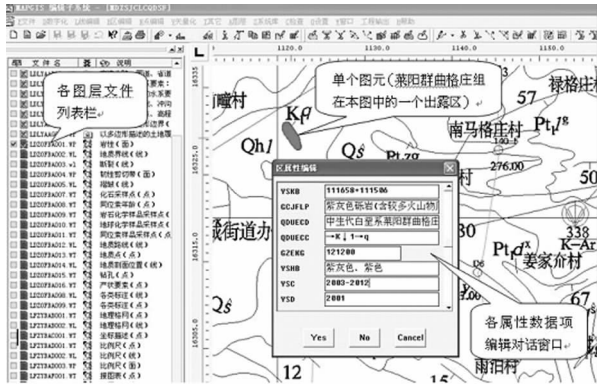


图 1 图层文件中单个图元属性数据项填写实例

规定了“岩石组合(YSKB)、岩性特征(GCJFLP)、填图单位名称(QDUECD)、填图单位代号(QDUECC)、形成时代(GZEKG)、岩石颜色(YSHB)、岩石结构(YSC)、岩石构造(YSD)”等 16 个属性数据项。其中“岩石组合、岩石结构、岩石构造、形成时代”要求填写规定的下属词代码,其他的“岩性特征”等则要求按规定好的长度填写描述性的文字内容。地质专业人员填写每个属性数据项的内容时,要注意完整理解数据模型的规定。例如:前面的“岩石组合”中如果填写了多个岩性,那么后面的“岩石颜色”、“岩石结构”和“岩石构造”填写时,要与“岩石组合”中岩性的顺序对应,避免出现交叉,引发误解。

2 软件检查属性数据质量

全国矿产资源潜力评价项目办除对每幅专题图的编图质量高度重视外;对每幅图内含的各属性数据项填写的专业内容质量,也给予同样的关注。为此,项目办专门委托“综合信息集成课题”开发了 GEOMAG 软件,既用于辅助编图,也用于图件数据质量检查。软件界面、菜单和功能等的设计与 MAPGIS 软件类似,极大方便了软件的推广应用。

图 2 显示了在 GEOMAG 软件主窗口中打开的“山东省莒县下石城宋官疃式复合内生型重晶石矿预测工作区蚀变矿物遥感异常分布图”。检查时,运行菜单栏“图件辅助工具”中的“检查图件数据”功能。软件会弹出“检查图件数据”对话框,对话框中“检查项选择”栏共有 6 个复选项,列出了软件可以检查的 6 个方面。选择“检查属性值域”,就会对那些规定了下属词代码的属性数据项,进行下属词代

码一致性检查。

GEOMAG 软件内部集成了一个数据模型下属词代码字典库,检查时,软件会将填写的代码与字典库中对应数据项下的代码进行比对,凡是不一致的,都会被检测出来。对构成图件的所有图层文件的检查结果,将由软件自动导出,生成如图 3 所示的 EXCEL 格式的检查报告。

由图 3 可知“LYGREDC001.WP”(蚀变矿物遥感异常面)图层文件,其数据项代码为“KCCG”(预测矿种类型)的属性数据项,有 641 个图元所填内容与数据模型规定的下属词不一致。在最后一列详细列举了出错图元的 ID 号,据此可以准确定位出错图元在图中的位置。

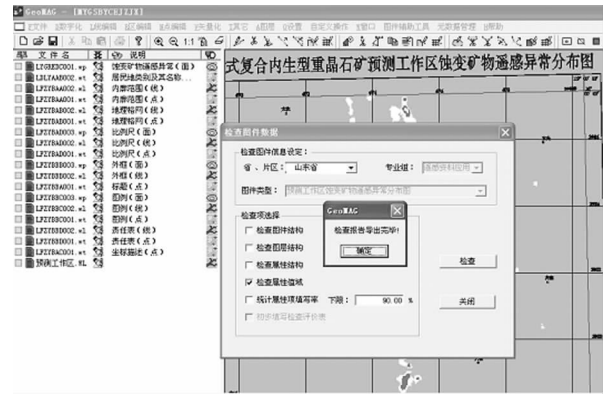


图 2 GEOMAG 图件数据质量检查软件

3 人工检查属性数据质量

GEOMAG 软件只能自动进行逻辑方面的检查,虽然能够指出哪里出了错误,但无法指出错误原因,如何修改就更无能为力。另外,还有很多专业错误是软件无法检查的。解决此类问题,只有借助地质专业人员的人工检查。

3.1 人工辅助核实软件查出的错误

根据图 3 检查报告中的提示,回到 GEOMAG 软件主窗口中,运行菜单栏“图件辅助工具”菜单下的“浏览图元属性”功能,会出现如图 4 所示的次级窗口。次级窗口上方有 3 个选择项,分别是“当前点文件属性”、“当前线文件属性”和“当前区文件属性”;这 3 个选择项与主窗口左侧图层文件列表栏中选择的图层文件有关,软件设定在图层文件列表栏中,每次只能分别勾选一个“点文件”、“线文件”和“区(面)文件”浏览图元属性。因此,在次级窗口上

方点击选中“当前区文件属性”后,次级窗口会显示填写的“LYGREDC001.WP”图层文件的所有图元的属性数据项内容。

如图 4 所示的“LYGREDC001.WP”图层文件共有 13 个属性数据项,其中“预测矿种类型”属性数据项,规定要填写下属词代码。在次级窗口下方有一个“下属词切换”选项,当未勾选该选项时,“预测矿种类型”属性数据项下对应各图元,显示的都是如“1025”和“3530”的代码;当勾选该选项时,对应各图元,有的还是显示代码“1025”,有的则显示汉字“重晶石”。即勾选“下属词切换”选项后,“预测矿种类型”属性数据项下,凡是数据模型规定的代码都能转换为对应的汉字,凡不属于规定代码的都无法实现转换。由软件打开的图名可知,“预测矿种类型”填写下属词代码“3530(重晶石)”是正确的。

利用 GEOMAG 软件的“下属词切换”功能,对那些数据模型规定了下属词的属性数据项,人工可以进行初步的正误判断:凡是错填、误填了非数据模型规定下属词的图元,将无法进行代码和对应汉字间的切换。

这里需要注意,对那些填写了规定下属词代码,而代码内容与地质实际不符的,GEOMAG 软件也会实现正常的转换。这种错误,只有地质专业人员才能鉴别出来。

序号	名称	数据类型	是否必填	备注
1	项目代号	字符串	否	
2	图元名称	字符串	否	
3	图件名称	字符串	否	
4	图件名称	字符串	否	
5	图件名称	字符串	否	
6	图件名称	字符串	否	
7	图件名称	字符串	否	
8	图件名称	字符串	否	
9	图件名称	字符串	否	
10	图件名称	字符串	否	
11	图件名称	字符串	否	
12	图件名称	字符串	否	
13	图件名称	字符串	否	
14	图件名称	字符串	否	
15	图件名称	字符串	否	
16	图件名称	字符串	否	
17	图件名称	字符串	否	
18	图件名称	字符串	否	
19	图件名称	字符串	否	
20	图件名称	字符串	否	
21	图件名称	字符串	否	
22	图件名称	字符串	否	
23	图件名称	字符串	否	
24	图件名称	字符串	否	
25	图件名称	字符串	否	
26	图件名称	字符串	否	
27	图件名称	字符串	否	
28	图件名称	字符串	否	
29	图件名称	字符串	否	
30	图件名称	字符串	否	

图 3 GEOMAG 属性属性项值域检查报告

3.2 人工依靠编码规则检查属性数据质量

数据模型规定的下属词代码都遵循一定的编码规则,代码中的首位数码代表一定的意义。如属性数据项“岩石名称”中的代码“121105”表示的是“砾屑灰岩”,“岩石结构”中的代码“2003”表示的是“砾状结构”,“岩石构造”中的代码“2001”表示的是“层

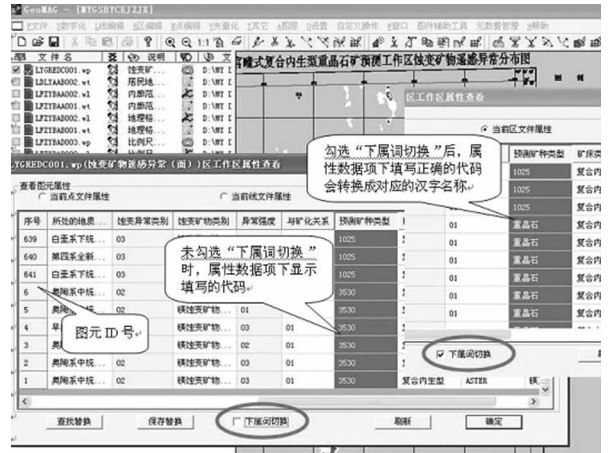


图 4 在 GEOMAG 软件中定位、检查出错的属性数据项值域

状构造”。由此可见,如果“岩石名称”中填的是以“1”开头的 6 位沉积岩代码,相应的“岩石结构”和“岩石构造”则填写以“2”开头的 4 位的代码。

遵循上述的编码规则,可以人工对成果图空间数据库中图层文件所涉及各图元属性数据项中填写的内容,进行初步的正误判断。

3.3 人工对代码性属性数据和描述性属性数据质量的对照检查

由于数据模型规定的属性数据项,有的要求填写下属词,有的要求填写描述性的文字内容。地质专业人员检查这两类属性数据项所填写内容时,关键把握两点:一是看其填写内容的正确性;二是在内容正确的前提下,看其内容是否能满足矿产资源潜力评价技术要求。

为了便于地质专业人员检查成果图空间数据库中填写的属性数据项内容,GEOMAG 软件提供了转出属性数据项内容的功能。转出时,那些规定了下属词代码的属性数据项,软件会自动将下属词代码转换成对应的汉字,这项功能提高了转出后属性数据项内容的可阅读性。

如图 5 上方所示,在 GEOMAG 软件主窗口中打开“山东省韩旺式沉积变质型铁矿预测工作区变质建造构造图”,运行菜单栏中“图件辅助工具”下的“转出图元属性”功能,会出现“转出图层”对话框。对话框左侧的“当前图件图层列表”栏列出了软件主窗口左侧“图层文件列表栏”中的所有图层文件,可以从其中挑选要进行检查的图层文件,选中的图层

