doi:10.3969/j.issn.1001 - 358X.2016.04.009

基于基础地理空间数据库数据质检子系统的数据检查方案研究

翟雪宁
(甘肃省基础地理信息中心,甘肃 兰州　730000)

摘要:文中以基础测绘 1∶10 000 矢量数据为例,利用甘肃省基础地理空间数据库中的数据质检子系统,研究了基础测绘数据成果的数据模型设计、质量检查项设计和批量检查方法。结果表明:数据质检子系统的应用在确保了数据检查质量的同时,较大提高了基础地理信息数据入库检查的效率。

关键词: 空间数据; 质量检查; 数据模型设计; 质量检查项设计
中图分类号:P208　　**文献标识码**:A　　**文章编号**:1001 - 358X(2016)04 - 0032 - 03

Research on data checking scheme of data quality inspection subsystem based on fundamental Geo - spatial database

Zhai Xuening
(Geomatics Center of Gansu Province, Lanzhou 730000, China)

Abstract: Taking 1∶10000 vector data of basic surveying and mapping as an example, the paper studies data model design, quality check item design and batch checking method of basic surveying and mapping data with the data quality inspection subsystem in the fundamental Geo - spatial database of Gansu province. The results show that data quality control subsystem ensures the quality of the data checking, and greatly improves the inspection efficiency of the basic geographic information data stored into database.

Key words: spatial data; quality check; data model design; quality check item design

为提高测绘服务能力和水平,各省均构建了基础地理空间数据库。作为基础地理空间数据库的核心,测绘成果数据的质量直接决定着基础地理信息的服务能力。为保证测绘成果数据质量,基础测绘数据不但需要在生产、验收过程中进行质量检查,还需执行成果入库前的质量检查。由于测绘成果数据质量的提高是一个动态持续的过程,因此其质量标准也会随之变化。甘肃省基础地理空间数据库中的数据质检子系统提供了将各种质检规则通过自由灵活的配置模块,可将检查项固化成方案,每个方案针对不同的数据类型与不同的比例尺,可以方便的进行调用和共享。同时底层的质检规则库可以支持拓展,能够极大的提高系统质检模块的开放性与拓展性。

以甘肃省基础地理空间数据库中数据质检子系统为例,介绍测绘成果的批量检查方案设计思路。批量数据检查方案设计主要分为两部分:数据模型设计和质量检查项设计。

1　数据模型设计

1.1　目录模型设计

目录模型是根据相关标准创建数据组织标准,用以检查待检数据。根据待检查数据的类型,参照数据组织存储路径,利用目录模型设计模块定义数据目录形式。不同数据成果可根据各自数据内容、特点组织存储。

以基础测绘数据为例,按照统一标准组织数据模版。首先,创建“基础测绘成果”根目录;其次,在

引用格式:翟雪宁. 基于基础地理空间数据库数据质检子系统的数据检查方案研究[J]. 矿山测量,2016,44(4):32 - 34.

根目录下“DLG”、“DOM”、“DEM”、“图历簿”四个文件夹;第三,在“DLG”、“DOM”、“DEM”文件夹下分别创建以标准图号命名的文件夹,“图历簿”文件夹下创建以标准图号命名的.doc 文件;最后,各自图号下创建相应文件,如标准图号命名的 DLG 文件夹下应放以标准图号命名的.mdb 和.mat 文件。

1.2　数据模板

利用数据模版设计模块,可以灵活设置各种类型数据模版。根据待检查数据类型和格式,定义数据命名规则和数据格式。现在已有的命名规则有三种:标准图幅号,或标准图幅号加前、后缀;包含当前数据文件名,或当前数据文件名加前、后缀;以占位符约束数据命名。已有的数据格式:矢量数据为.dlg,高程数据为.dem,.asc,.grd 和.bil,影像数据为.img 和.tiff 以及元数据为.mat。

以大比例尺矢量数据为例,数据模板设计流程为使用“占位符约束命名”规则定义数据命名,并且数据类型为.dlg,其次定义图层的名称及类型,最后定义属性项及顺序。

2　数据检查方案设计

空间数据质量是指数据适用于不同应用的能力,检查项是为了检查数据中的某类错误而制定的规则,根据检查内容从软件内含的规则库中选择相应的检查规则,组成当前所需要的质检规则库。甘肃省基础地理空间数据库中数据质检子系统检查项均按照《GBT 17941 - 2008 数字测绘成果质量要求》和《GBT 17941.1 - 2000 数字测绘产品质量要求 第 1 部分数字线划地形图、数字高程模型质量要求》的规定设计。空间数据质量一般包括成果规格、空间参考系、位置精度、属性精度、逻辑一致性、要素完整性和现势性等,所涉及检查规则软件已基本覆盖。

根据基础地理信息数据的特点和相关标准规范的要求,甘肃省基础地理空间数据库中数据质检子系统检查项设计主要涵盖:数据交汇录、空间参考、位置精度、完整性、属性精度、逻辑一致性、表征质量、栅格/栅格质量、元数据质量、接边检查。已有检查规则集如表 1 所示。

3　数据质检子系统的检查方法应用

以甘肃省基础测绘 1∶10 000 矢量数据为例,介绍详细的矢量数据入库检查规则应用。矢量数据入库检查主要包括数学基础、属性精度、位置精度、接边精度和逻辑一致性的检查。

表 1　检查规则集内容列举

检查规则集	内　　容
数据交汇录	数据目录格式;文件的缺失、冗余。
空间参考系	大地基准;高程基准;地图投影位置。
位置精度	平面精度;高程精度;DOM 接边检查;等高线高程中误差检查;高程中误差检查等。
完整性	数据层完整性,数据内部文件的完整性等。
属性精度	属性项完整性,字段顺序,分类、编码正确,非空字段检查等。
逻辑一致性	空间要素拓扑关系的检查,主要针对空间实体的点、线、面之间的 6 种组合,13 种拓扑关系的应用检查,如重叠关系、分离关系、包含关系、相邻关系等。
表征质量	线形折刺、回头线、粘连、自相交、抖动、变形扭曲,有向点、有向线等。
栅格/影像质量	切范围、高程精度、地面分辨率等。
元数据质量	元数据完整性和属性准确性。
接边检查	相邻图幅接边无遗漏、栅格数据同名格网高程值或坐标取值符合要求。

3.1　数据整理

待检查数据应按照目录格式要求,确保数据命名正确、格式统一。

3.2　数据模型设计

定义以标准图号命名数据,根据要素分类与编码规定,设计数据分层、属性项字段定义及顺序。

3.3　数据质检方案设计

结合 1∶10 000 基础测绘数据 DLG 数据分层及 DEM、DOM 数据特点,制定合理的数据质检方案,确保数据质检内容全面。以下列举涉及的部分检查规则:

(1)空间参考正确性:坐标系统;分带方式、带号、图幅分幅,图廓精度检查。

(2)数据集和图层一致性:数据集结构符合规范;矢量数据数据集精度和容差正确;图层类型、命名正确;不存在多余图层、遗漏图层和空图层。

(3)属性项一致性:属性项类型、精度、命名定义正确;不存在多余、遗漏属性项,属性项;属性值不能为空;要素编码值与属性值逻辑上必须匹配。

(4)等高线检查:根据等高距,检查首曲线、计曲线赋值的正确性;检查高程是否在正确的等高线范围内。

(5)图幅接边检查:要素与内图廓线是否接边;图幅间是否有要素遗漏

(6)数据套合检查:与 DOM 数据套合检查平面精度;与 DEM 数据套合检查平面精度。

(7)拓扑一致性检查:具体检查细则见下表。

表 2　拓扑检查项细则

检查项	要求	检查项	要求
异常要素	不存在点、线、面要素中有属性没有图形的情况	自相交	所有线、面要素不能存在自相交;不能存在面打折
线噪声	不存在 0 长度和极小长度线要素	相交	面要素之间不能相交
面噪声	不能存在 0 面积和极小面积	复合线	不存在复合线
伪节点	线要素均不存在伪节点	复合面	不存在复合面
悬挂点	线要素均不存在悬挂点	缝隙	面与面之间不能存在缝隙
折角	所有线要素不能存在折角过小	面空洞	面层中不能存在空洞
自重叠	层内不存在点、线、面要素自重叠	线是否在面的边界上	境界线必须与境界面边线重合…
检查要素是否在图廓范围内	检查所有要素是否超出理论图廓范围	不同面层的多边形之间不能重叠	湖泊面不能与居民地面重叠…
线在(或不在)面内	等高线不应穿越房屋;道路中心线必须在路面内…	不同面层多边形相互叠加不能有空隙	土质层和植被层多边形叠加后不能有空隙…
判断有向点的方向	根据等水系线判断水闸等地物的方向…	两线相交处必有其他地物	道路与水系相交应有桥梁、涵洞…
判断有向线的方向	根据水系线方向判断水系流向;根据等高线判断水系流向…	线段必须被(或不能被)另一要素类中的线段覆盖	铁路桥必须落在铁路上;公路桥必须落在公路上
…	…	…	…
点必须被(或不能被)线覆盖	输水隧道必须落在水系线上交通附属设施必须落在道路上;	点要素必须(或不能)在多边形内	高程点不能落入河流、房屋面内

3.3　数据质检

选择质检数据源及相应的质检方案,进行批成果数据检查。首先加载质检方案,可按需要对方案中的质检项进行再次选择,执行单幅或批量质检任务。其次运行检查任务,质检过程由计算机自动完成,不需要人工参与。最后查看质检结果,对部分检查结果进行人工排查,导出错误图形,并自动生成质检报告。

4　结　语

通过对数据检查方案的合理设计,甘肃省基础地理信息数据库的数据质检子系统针对各种类型数据配置质检方案并保存到数据库中,以实现对 1∶200、1∶500、1∶1 000、1∶2 000、1∶5 000、1∶10 000 等常用比例尺的矢量数据、栅格数据及元数据的批量检查。实践表明,数据质检子系统的应用不但确保了数据检查的质量,避免了人为因素可引起的误差;而且提高了基础地理信息数据入库检查的效率,节省了各类成本。

参考文献:

[1]　王佑武,马增翼,刘旺生. 数字地形图质量评定方法研究[J]. 甘肃科学学报,2015,27(4):20－23.

[2]　王晓莉. 面向矿山应急服务的空间数据集成平台构建[J]. 矿山测量,2014,42(3):86－89.

[3]　陈国海. 数字测绘产品的质量检查与质量控制[J]. 科技资讯,2013,11(14):28,30.

[4]　GBT 17941－2008 数字测绘成果质量要求[S].

[5]　刘冬晖. 水印信息隐藏在地理空间图像数据传输中的应用[J]. 矿山测量,2015,43(4):27－28.

[6]　GBT 17941－2000 数字测绘产品质量要求 第 1 部分数字线划地形图、数字高程模型质量要求[S].

[7]　康义锋. 基于 SuperMap Objects 的大比例尺基础地理信息数据库管理系统的建立与实现[J]. 矿山测量,2014,42(3):83－85,89.

作者简介:翟雪宁(1987－),女,汉族,甘肃兰州人,助理工程师,本科,从事数据质检、空间数据建库相关工作。

(2016－05－17)