Scientific and Social Research, 2024, Volume 6, Issue 7
https://ojs.bbwpublisher.com/index.php/SSR
Online ISSN: 2981-9946
Print ISSN: 2661-4332

An Integrated Quality Inspection Scheme for Updated Basic Geographic Data Results in DWG and SHP Formats

Xinchao Xu[1]*, Dede Zhang[1], Tao Xu[2]

[1]Henan Institute of Metrology, Zhengzhou 450000, Henan Province, China

[2]School of Computer and Information Engineering, Henan University, Zhengzhou 450046, Henan Province, China

Corresponding author: Xinchao Xu, zhzh_xxc@126.com

Abstract: The quality inspection of updated basic geographic data results involves various types of data results, and the conversion effect of data formats significantly impacts the quality of the results. Addressing the issues of low conversion efficiency and high error rates in commonly used DWG and SHP format data in quality inspection operations, this paper proposes the application of DWG rasterization technology in field spot-checking operations. The study designs an integrated quality inspection scheme for updated basic geographic data results based on the ArcGIS Engine technology system. This scheme provides a detailed analysis of the current status of quality inspection operations for updated basic geographic data results and offers technical implementation methods for specific business content, providing feasible data quality control methods and efficient solutions for the smooth conduct of quality inspection work.

Keywords: Quality inspection of updated basic geographic data results; DWG rasterization; ArcGIS engine

Online publication: July 26, 2024

1. Introduction

Updated basic geographic data results refer to new version datasets formed by revising and supplementing existing basic geographic information data through various techniques and methods, such as remote sensing image interpretation, field surveys, GIS data analysis, and so on. According to standards like "Surveying and Mapping Results Quality Inspection and Acceptance" (GB/T 24356—2023) and "Digital Surveying and Mapping Results Quality Inspection and Acceptance" (GB/T 18316—2008), the quality inspection of updated basic geographic data results requires field spot-checks and indoor map quality inspections following standard procedures, ultimately generating quality inspection reports for the updated basic geographic data results [1–2].

During field spot-checks, various updated basic geographic data results are generally overlaid to form a DWG working base map, which is then used for field manual spot-checks and accuracy plotting. Indoor map quality inspection mainly targets vector data (such as DLG, SHP, and others) of updated basic geographic data results and field verification results, including checks for mathematical accuracy, data and structural correctness, geographic accuracy, finishing quality, and attachment quality, ultimately outputting detailed quality inspection reports [3]. This process involves multiple types of data files and multiple data type conversions, making the workflow cumbersome and the data transfer process complex, which can easily affect data quality and, consequently, the quality inspection results [4].

Although DWG is a widely accepted drawing exchange format in the CAD industry, excelling in detailed graphical representation for engineering and architectural design, it is not conducive to the storage and spatial analysis of geographic features [5]. The current common practice is to convert field spot-check data in DWG format to SHP format for further analysis and inspection by indoor quality inspectors. However, as DWG is a proprietary file format of AutoCAD, efficient and lossless DWG to SHP conversion remains a business challenge [6]. Although there are various methods available, such as Zhang using Arcpy to convert urban regulatory DWG data to ArcGIS format, Li et al. using semantic conversion technology in FME to achieve CAD and GIS data sharing, and Zhang proposing a method for establishing 3D models of DWG format topographic map data [7–9]. Chai has implemented a non-destructive conversion method from CAD to GIS data for residential areas [10]. These methods are mostly suitable for specific scenarios and have weak generalization capabilities, failing to fully meet the needs of field quality inspection work.

Based on this, this paper studies the key issues of data format conversion in the quality inspection of updated basic geographic data results. After thoroughly investigating the need for quality inspection work for updated basic geographic data results, the paper designs an integrated quality inspection scheme for updated basic geographic data results in DWG and SHP formats using the ArcGIS Engine software development engine. This scheme cleverly rasterizes DWG data and integrates it with SHP format vector data, preserving the rich data visualization style information of DWG data while providing a lossless precision working base map for field spot-checks. Additionally, by saving the field spot-check results as SHP format files, it achieves seamless integration of field and indoor data, providing a high-precision, high-efficiency, and high-stability technical approach for the quality inspection of updated basic geographic data results.

2. Quality inspection work for updated basic geographic data results

The workflow of the integrated quality inspection scheme for updated basic geographic data results is illustrated in **Figure 1**. The quality inspection tasks are divided into two main parts: field spot-checks and indoor verification.

The field spot-check process includes the following steps. Initially, a cartographic representation of the updated basic geographic data results is performed to create the DWG working base map. Subsequently, the DWG files are rasterized to generate DWG raster images. Using the spatial coordinates of the raster images as a reference, SHP plotting layers are created. Data spot-checks are conducted on the SHP plotting layers, yielding spot-check plotting result data.

The indoor verification process includes the following steps. Accuracy verification of the field-obtained SHP plotting vector results. Examination of the vector and raster data in the updated basic geographic data results based on the "Inspection and Acceptance of Surveying and Mapping Results"

(GB/T 24356—2023). Quantitative scoring of various quality inspection records, generation of quality inspection reports, and evaluation and feedback on the updated basic geographic data results. The critical technologies involved in these processes will be elaborated upon in the following sections.

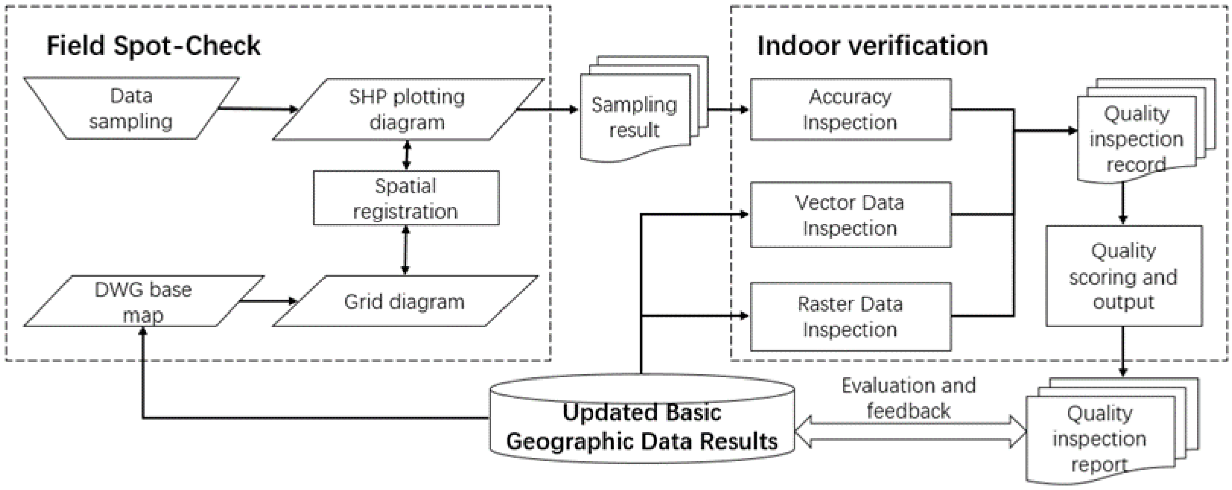

Figure 1. Service flow chart of the integrated quality inspection solution

3. Field quality spot-check based on the integration of DWG and SHP data

The field quality spot-check operations primarily comprise three critical steps: DWG working base map rasterization, spatial registration, and spot-check plotting. Utilizing the operational interfaces provided by the ArcGIS Engine software development platform, such as *ICadDrawingWorkspace*, *ISpatialReference*, and *IFeatureClass*, this study has developed the corresponding technical implementation scheme for these steps.

3.1. Rasterization of DWG working base map

Due to the cumbersome steps and suboptimal conversion quality involved in converting DWG files to SHP vector files, this study employs the *ICadDrawingDataset* interface provided by ArcGIS Engine to rasterize DWG files into raster layers. The specific implementation process is illustrated in Algorithm 1 in **Table 1**.

Initially, in lines 1–2, the CAD workspace factory interface *cadFactory* is instantiated, utilizing its *OpenFromFile* method to specify the directory containing the DWG maps. Subsequently, in lines 3–4, the *drawingCAD* interface for drawing CAD files is defined, employing its *OpenCadDrawingDataset* method to extract the DWG file into *cadDataset*. Following this, in lines 5–6, the raster layer *cadLayer* is defined, utilizing its *CadDrawingDataset* method to rasterize the DWG data *cadDataset*. Finally, in line 7, the raster layer *cadLayer* representing the DWG working base map is returned.

Table 1. Algorithm 1: Rasterization of DWG working base map

Input: DWG base map;
Output: ArcGIS raster layer object;
IWorkspaceFactory cadFactory = new CadWorkspaceFactoryClass();
IFeatureWorkspace cadWorkspace = (IFeatureWorkspace) cadFactory.OpenFromFile("DWG base map directory", 0);
ICadDrawingWorkspace drawingCAD = cadWorkspace as ICadDrawingWorkspace;
ICadDrawingDataset cadDataset = drawingCAD.OpenCadDrawingDataset("DWG base map");
ICadLayer cadLayer = new CadLayerClass();
cadLayer.CadDrawingDataset = cadDataset;
return cadLayer;

3.2. Spatial registration methods for DWG and SHP data

The DWG working base maps used for updating basic geographical data typically employ either a geographic coordinate system or a projected coordinate system. Field-surveyed positions of sample points are generally recorded in latitude and longitude coordinates. In cases where these coordinate systems differ, the GeometryEnvironment class can be utilized to achieve coordinate system transformation of layer data, as outlined in Algorithm 2 in **Table 2**.

Algorithm 2 takes as input the layer data “layer” and the new coordinate system “newReference”, and outputs the layer with updated coordinates. Initially, in lines 1–2, the spatial reference “currentReference” of the layer “layer” is obtained. Subsequently, in lines 3–4, a transformation environment “envi” for the new coordinate system “newReference” is created, and its “Transform” method is applied to convert “currentReference” to “newReference”. Finally, in lines 6–8, the layer object “layer” is assigned with the data transformed to the new coordinate system “newReference”, and the updated layer is returned.

Table 2. Algorithm 2: Layer coordinate system transformation

Input: Layer `layer`, new coordinate system `newReference`;	
Output: Layer `layer` with updated coordinate system	
1:	IFeatureClass featureClass = layer.FeatureClass; // To retrieve a feature class
2:	ISpatialReference currentReference = featureClass.SpatialReference; //To retrieve the current spatial reference
3:	GeometryEnvironment envi = new GeometryEnvironment(); // To create a geometry environment object
4:	envi.SpatialReference = newReference; // To set a new spatial reference
5:	envi.Transform(currentReference); // To transform a spatial reference
6:	featureClass.SpatialReference = newReference;// To update the spatial reference of a feature class
7:	layer.SpatialReference = newReference; // To update the spatial reference of a feature layer
8:	return layer

3.3. Vector plotting of field spot-checks results

Following spatial registration, the vectorized DWG working base map aligns with the spatial coordinate system of the field inspection layer. Field inspection operations can directly annotate on the vector layer. The plotting principle for point features is outlined in Algorithm 3 in Table 3. Initially, in line 1, the vector layer "layer" is opened. Subsequently, in lines 2–3, a point object "point" with coordinates (x, y) is created; these coordinates are typically selected interactively from the map. Next, in lines 4–6, the "point" is added to the "layer". Finally, in line 7, the updated inspection layer "layer" is returned.

Table 3. Algorithm 3: Plotting inspection points on a vector layer

Input: Vector layer `layer`, inspection point coordinates (x, y);	
Output: Updated vector layer `layer`;	
1:	IFeatureClass featureClass = layer.FeatureClass; // To retrieve layer features
2:	IPoint point = new PointClass(); // To create a point feature
3:	point.PutCoords(x, y); // To set coordinates
4:	IFeature feature = featureClass.CreateFeature(); // To create a new feature in `layer`
5:	feature.Shape = point; // To set the coordinates of a feature
6:	feature.Store(); // To store the feature
7:	return layer;

4. Indoor quality verification of basic geographical data updates

In addition to verifying the accuracy of field-sampled data, the quality control of in-house updates for basic geographical data involves several inspection steps. These include checks for completeness, structural integrity, spatial relationships, attribute consistency, and graphical accuracy, as illustrated in **Figure 2**. To streamline operational functions, internal quality control is categorized into accuracy inspection of sampled data, vector data examination, and raster data verification.

4.1. Accuracy inspection of sampled data

The accuracy inspection module for sampled data primarily involves calculating and statistically analyzing the precision of planar points, elevation points, and inter-point distances. Since the sampled data results are recorded in the inspection layer "layer" (refer to Algorithm 3), it is necessary to obtain the corresponding true coordinates from the dataset to conduct the accuracy checks effectively. This is achieved using relevant vector data manipulation methods available in ArcGIS Engine. The inspection data is retrieved from the inspection layer, and each point is compared with the data in the original dataset layer. The nearest spatial point serves as the checkpoint for verification. This approach can be directly implemented using the *INearestNeighbor* interface, simplifying the process significantly. Once the inspection comparison data is obtained, statistical metrics such as average error and root mean square deviation are calculated following the precision requirements specified in the "Quality Inspection and Acceptance of Surveying and Mapping Results" (GB/T 24356 — 2023), to determine compliance with accuracy standards.

Figure 2. Indoor quality verification scheme

4.2. Vector data inspection

The vector data inspection module entails thorough checks for data integrity, spatial topology, attribute coherence, and cartographic fidelity.

Data integrity check: This phase involves validating the completeness and integrity of vector data files and layers. This is achieved by loading map documents and accessing the *Map* object, then meticulously scrutinizing the *MapFrame* and Layer properties to ensure comprehensive coverage.

Spatial topology check: This critical assessment identifies anomalies such as self-intersections, dangling nodes, and pseudo-nodes. Given the extensive scale of basic geographical data updates, spatial topology checks are efficiently executed using established tool interfaces within ArcGIS Engine. For example, detecting self-intersections utilizes the *TopologyException* class to analyze topological irregularities at polygon vertices. The validation of dangling points employs the *IsPointOnLine* method of the *ITopologicalOperator* interface to ascertain the alignment of a point with a line. Similarly, pseudo-node detection is facilitated through the *CheckPseudoNodes* method of the *IPseudoNodeChecker* interface. Detailed implementations for overlap, gap, and other spatial topology errors are documented comprehensively in the ArcGIS Engine official resources.

Attribute data check: This phase encompasses the verification of unique values, mandatory field conditions, and attribute consistency. The identification of unique values is facilitated by the *IUniqueValueRenderer* interface, which validates and renders distinct attribute values. The assessment of mandatory field conditions involves iterating through fields and evaluating the *IsNullable* property to ensure data completeness. Consistency checks utilize diverse methods within the *ITopology* interface to maintain data coherence.

Graphical quality check: This stage focuses on ensuring the accuracy and precision of cartographic outputs. It involves meticulous manual review, encompassing assessments of graphic depiction accuracy, symbol configuration validity, annotation placement correctness, interior and exterior map frame adornment integrity, alignment of map frame lines, and latitude and longitude gridline precision.

4.3. Raster data inspection

The raster data inspection module primarily examines the outcomes of Digital Elevation Models (DEM) and Digital Orthophoto Imagery (DOM). A DEM is a dataset of regularly spaced grid points in a projected plane (such as a Gauss-Krüger projection), consisting of plane coordinates (X, Y) and elevation (Z) data. The grid spacing of a DEM should be compatible with its elevation accuracy, forming a systematic grid series. DOM data involves digitized aerial photographs or remote sensing images processed with DEM correction, cropped according to chart sheet extents.

Both types of raster data products typically adhere to strict production standards and quality control requirements. Their inspection focuses on coordinate projection, spatial reference information, file formats, and naming conventions, image integrity, data resolution, and map sheet extents. This scrutiny does not typically involve graphical content but rather entails thorough verification by personnel.

5. Conclusion

The integrated quality inspection scheme for updated basic geographic data results in DWG and SHP formats provides a comprehensive solution to the challenges of data format conversion and quality inspection. By leveraging DWG rasterization technology and the ArcGIS Engine platform, the scheme ensures high-precision, high-efficiency, and high-stability quality inspections. This approach not only enhances the accuracy and reliability of the quality inspection results but also provides a feasible and efficient method for data quality control in geographic information systems.

Disclosure statement

The authors declare no conflict of interest.

Reference

[1] National Quality Inspection and Testing Center for Surveying and Mapping Products, Ministry of Natural Resources Sichuan Surveying and Mapping Product Quality Supervision and Inspection Station, Ministry of Natural Resources Heilongjiang Surveying and Mapping Product Quality Supervision and Inspection Station, et al., 2023, GB/T 24356-2023 Surveying and Mapping Results Quality Inspection and Acceptance. Standards Press of China, Beijing.

[2] National Surveying and Mapping Bureau of Surveying and Mapping Standardization Institute, National Surveying and Mapping Product Quality Supervision and Inspection Testing Center, Shaanxi Province Surveying and Mapping Product Quality Supervision and Inspection Station, 2008, GB/T18316-2008 Digital Mapping Results Quality Inspection and Acceptance. Standards Press of China, Beijing.

[3] Xu X, Zhang D, Sun H, 2024, Research and Application of the Results Quality Inspection System of Base Geographic Update Data. Resource Guide, 2024(02): 43–46.

[4] Zhang ZM, 2018, Preliminary Study on Quality Control of Basic Geographic Monitoring of National Conditions. Surveying and Spatial Geographic Information, 41(06): 94–96.

[5] Min Y, Sheng BY, Luo RP, 2023, Automatic Extraction Method of DWG Construction Drawing Data. Digital Manufacturing Science, 21(03): 213–218.

[6] Zhao Y, Yu HW, Wang P, 2021, Conversion of DWG Format Data to SHP Format Data Considering Object Layering. Surveying and Mapping Standardization, 37(02): 41–45.

[7] Zhang HL, 2021, Quick Conversion of Urban Planning DWG Data to ArcGIS File Format Supported by Arcpy. Tropical Landforms, 42(01): 41–46.

[8] Li DD, Li YL, 2020, Research on Pipeline Facility Data Conversion Model Method. Low Carbon World, 10(03): 200–201.

[9] Zhang GW, Method for Building 3D Models from DWG Topographic Map Data. North China Natural Resources,

2022(01): 107–109.
[10] Chai XY, 2022, Research on Lossless Conversion Method from CAD to GIS Data — Taking Residential Areas and Their Facilities as an Example. Surveying and Spatial Geographic Information, 45(04): 157–159.