RESEARCH ON QUALITY TESTING AND EVALUATION TECHNOLOGY OF FUNDAMENTAL GEOGRAPHIC INFORMATION DATABASE SYSTEM

He Zhang[1,2,*], Yaoling Shang[2], Xiaohua Tong[1], Jie Chen[3], Wei Ma[2], Miao Li[2], Yao Lu[1,2], Haipeng Chen[2]

[1] School of Surveying and Geo-Informatics, Tongji University, Shanghai, China - 3885987@qq.com, xhtong@tongji.edu.cn, 156291029@qq.com

[2] National Quality Inspection and Testing Center For Surveying and Mapping Products, Beijing, China - 1805163052 @ qq.com, 381280239@qq.com, 9105513@qq.com, 113349587@qq.com

[3] Zhejiang Institute of Hydraulics and Estuary, Zhejiang, China - 15690191@qq.com

Commission III, WG III/IVb

KEY WORDS: Geographic information, Database, Quality model, Quality Test, Quality Evaluation

ABSTRACT:

The extensive application of fundamental geographic information data in various industries makes it play an increasingly important role in national economic construction and social development. As a fundamental geographic information database construction project with data and database management system as its core, quality testing and evaluation is the key link and means for measuring database suitability and system application success or failure. This paper proposes a set of fundamental geographic information database system quality model, conducts research on quality testing techniques and methods, and establishes a quality evaluation system, which has been applied in national and provincial basic geographic information database systems.

1. INTRODUCTION

The fundamental geographic information database system is a large spatial database system established with fundamental spatial data 4D products as the main data source. The purpose of the system is to provide fundamental geographic framework information for planning, designing and decision-making support for the sustainable social and economic development of the country, provinces and counties. At the same time, with the rapid development of the national economy and society and the continuous acceleration of digital transformation, the application level of fundamental geographic information data in various industries and fields has continued to deepen, and the demand has continued to increase. The quality of the fundamental geographic information database system directly affects the accuracy and correctness of various analysis results. Doing a good job in quality control can provide a unified positioning basis for the construction of digital cities, provide scientific data support for government decision-making, provide authoritative fundamental geographic information for the society, and provide important guarantee service for the public.

With the continuous advancement of the construction of China's fundamental geographic information database system, its quality control is no longer limited to the "two-level inspection and one-level acceptance" of conventional fundamental geographic information data, but extends to the quality control of the database system. In terms of quality control technology of spatial data, although a large number of related technical and theoretical studies have been carried out at home and abroad, a technical system that can be directly applied to the quality control of fundamental geographic information database systems has not yet been formed. It is urgent to establish an advanced, efficient and targeted quality testing and evaluation technology system to ensure the quality of the fundamental geographic information database system.[1]

In view of the quality control requirements of China's fundamental geographic information database system, this paper fully analyzes the quality characteristics of the fundamental geographic information database system, applies quality control theories and methods, geographic information technology, spatial database technology, etc., and proposes a set of fundamental geographic information database system quality model, conducts research on quality testing techniques and methods, and establishes a quality evaluation system, which has been applied in national and provincial basic geographic information database systems.

2. CONSTRUCTION OF FUNDAMENTAL GEOGRAPHIC INFORMATION DATABASE SYSTEM

Fundamental geographic information data mainly describes the morphological features of the surface and the natural historical and cultural characteristics of the land. It is mainly composed of buildings, landforms, place names, water systems, roads, and a series of auxiliary facilities and socio-economic elements in natural geographic information, It can be used in all walks of life to achieve resource sharing. In general, all industries and fields related to geographic information data use fundamental geographic information data for unified geospatial positioning and spatial analysis.

The map scales in the fundamental geographic information database are roughly divided into three levels: large-scale maps greater than 1:10000, medium-scale maps from 1:10000 to 1:100000, and small-scale maps less than 1:100000. The datasets in the database generally include control survey results, digital line drawings (DLG), digital elevation models (DEM), digital orthophotos (DOM), place names and addresses, cartographic data and metadata, etc.

The fundamental geographic information database system is an information system that takes fundamental geographic

information as the management object and realizes the collection, input, processing, storage, query, analysis, expression, output, update and sharing of basic geographic data. The construction of China's national fundamental geographic information database has been carried out comprehensively, at present, databases of 1:4 million, 1:1 million, 1: 250,000, and 1:50,000 have been established nationwide. The provinces have also actively carried out the construction of fundamental geographic information database system, basically forming the fundamental geographic information data resources with multi-elements, multi-scale and multi-temporal, and the digital surveying and mapping results are increasingly rich.

3. QUALITY CHARACTERISTICS AND QUALITY MODELS OF FUNDAMENTAL GEOGRAPHIC INFORMATION DATABASE SYSTEMS

3.1 Analysis of quality characteristics

Aiming at the requirements of fundamental geographic information database system construction in China, we analyze the quality characteristics of fundamental geographic information database from micro and macro levels by analyzing the content, structure and basic characteristics of fundamental geographic information database, as well as the sources of errors in the process of database construction, the basic structure and functional performance requirements of database system construction.

Micro quality characteristics mainly refer to the location accuracy, attribute accuracy, logical consistency and thematic accuracy of multi-source and multi-scale mass data. The quality problems mainly come from the errors generated in the process of data collection and input.

Macroscopic quality characteristics, mainly investigate the database space reference system, integrity, timeliness and expression form of rationality, as well as the management system function, performance and security. On the one one hand, the quality problems are caused by errors in data processing, such as projection transformation error, graphics data editing error, graphics simplification error, data format conversion error, computer truncation error, etc. On the other hand, when the data is called and used for spatial analysis in the system platform and database, the quality of data use has errors due to the data itself, mainly including spatial analysis error, data integrity error and data current situation error.

3.2 Establishment of quality model

Quality model generally refers to the defined feature set and the relation set between them, which provides a framework for defining quality requirements and evaluating quality.

According to the quality characteristics of the fundamental geographic information database system, a three-level quality model including quality elements, quality sub-elements and test items is established. Specifically, it includes three quality elements of data set quality, management system quality, and documents quality and 11 related quality sub-elements. The quality model framework for the fundamental geographic information database system is shown in figure 1.
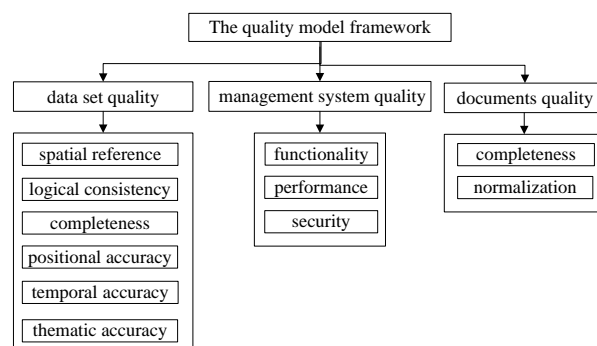

Figure 1. The quality model framework for the fundamental geographic information database system

3.2.1 Data Set Quality

As a fundamental geographic information database system with data as the core, data set is the most basic and important part. The quality control of data set is the key to determine the applicability of database and the success or failure of system application. The quality element of the dataset consists of six sub-elements: spatial reference frame, logical consistency, completeness, location accuracy, time accuracy and thematic accuracy.

(1)The test items of spatial reference frame include geodetic datum, elevation datum and map projection.

(2)The test items of logical consistency include conceptual consistency, format consistency, topological consistency, domain consistency, association consistency, etc. The quality test of the logical consistency is performed on each data set itself in terms of database definition; data file organization, topology relationship, and range of values, as well as the degree of correlation between each data set.

(3) The test items of completeness include both commission and omission level of quality testing.

(4)The test items of positional accuracy include plane accuracy and elevation accuracy.

(5)The temporal accuracy is mainly tested from the aspect of the present situation of the data set.

(6)The test items of thematic accuracy include attribute correctness, image characteristics, resolution, grid parameters, geometric anomalies, etc., and special tests are carried out for the quality characteristics of various datasets in the fundamental geographic information database system.

Quality element	Quality sub-element	Test item
Data set quality	Spatial reference system	Geodetic datum
		Elevation datum
		Map projection
	Logical consistency	Conceptual consistency
		Format consistency
		Topological consistency

		Domain consistency
		Association consistency
	Completeness	Commission
		Omission
	Positional accuracy	Plane accuracy
		Elevation accuracy
	Temporal accuracy	Present situation
	Thematic accuracy	Attribute correctness
		Image characteristics
		Resolution
		Grid parameters
		Geometric anomalies

Table 1. List of quality elements, quality sub-elements and test items for data set quality

3.2.2 Management System Quality

The management system is the operating center of the fundamental geographic information database system, and it undertakes important functions such as input and output, query statistics, and update maintenance of the fundamental geographic information database. Under the quality element of management system, three quality sub-elements are designed from the perspectives of functionality, performance and security respectively.

(1)The test items of the functionality of the management system include completeness and correctness, and the quality test is performed from the most basic and strictest function availability and function realization of the fundamental geographic information system.

(2)The test items of the performance of the management system include reliability and efficiency, and the quality test is carried out from the stability and operation of the management system.

(3)The test items of management system security include physical security, permission control, data security, backup and recovery.

Quality element	Quality sub-element	Test item
Management system quality	Functionality	Completeness
		Correctness
	performance	Reliability
		Efficiency
	Security	Physical security
		Permission control
		Data security
		Backup and recovery

Table 2. List of quality elements, quality sub-elements and test items for management system quality

3.2.3 Documents Quality

A high-quality fundamental geographic information database management system, in addition to high-quality data set and management system, should also have a series of complete and normative technical documents to support the system's operation and maintenance. The documents mainly includes the implementation plan of the fundamental geographic information database system, technical design book, data dictionary, system requirements, detailed system design, user documentation, technical summary and other technical documents, as well as quality control records, operation and maintenance records, backup systems, and security and confidentiality systems. The specific content can be implemented according to the actual construction requirements of the fundamental geographic information database system.

Quality element	Quality sub-element	Test item
Documents quality	Completeness	Completeness of technical documents, operation and maintenance records, quality control records
	Normalization	Normalization of technical documents, operation and maintenance records, quality control records

Table 3. List of quality elements, quality sub-elements and test items for documents quality

4. QUALITY TESTING TECHNICAL METHODS

On the basis of the established quality model of the fundamental geographic information database system, for each quality element and its test items, the quality test is carried out by means of automatic program inspection, human-computer interaction inspection, and manual verification inspection.

4.1 Data quality testing technology based on "rule-scheme"

A quality inspection rule is an abstraction and encapsulation of test objects, test parameters and test methods, and it performs a specific element-level test operation. Integrate quality inspection rules into a rule base, which mainly includes basic requirements for building basic geographic information data, concept consistency, format consistency, topology consistency, attribute correctness, image features, resolution, grid parameters, geometric anomalies, etc. In actual implementation, quality testers can use these rules to flexibly establish a logical combination of a set of quality inspection rules, and formulate test plans with specific requirements to realize automatic program inspections that meet different requirements.

4.2 Multi-source data comparison test technology based on spatial operation

It mainly adopts the automatic reference and comparison algorithm of multi-source data based on spatial operation. The algorithm is divided into three levels. One is to use authoritative thematic data, reference data and existing old data in the database to establish graphics, attributes and logic between the data to be inspected and the existing data through complex

spatial operations. And then through the comparison and analysis of multi-source data with the same name feature elements, data errors can be quickly found. The second is to use different scales and different types of data in the fundamental geographic information database to test the consistency of the association between the data, mainly to test the consistency of the geographic expression of elements, the primary and secondary selection, and the location accuracy between the similar results in the adjacent scales in the database; Continuity of element attributes and consistency of image features among results data in different production responsibility areas; accuracy of association between metadata and database data, etc. The third is to test the integrity of the database by comparing the data before and after warehousing, including testing the overall coverage of the database, whether the content exceeds or fails to meet the design requirements of the database system, and whether the amount of elements of the dataset in the database exceeds or is less than the amount of elements of the production results before warehousing, etc.

4.3 Testing technology based on quality information database

Traditional quality information is usually stored through paper record forms or individual files, which cannot provide the statistical and analytical functions required in the quality testing process. By establishing a quality information database based on error information storage and positioning technology, the recorded information in the quality testing process is stored and managed uniformly in the form of graphs or relational tables, so as to quickly locate and trace the error data in space and facilitate interactive checking, editing and revision, and to achieve quality analysis and quality evaluation.

4.4 Function and performance test of database management system

For functionality, test the coverage of system functions and the correctness of the output of the management system functions to meet design requirements.

For fault tolerance and recoverability testing of management system, including providing effective prompts for important data entry and prompting for illegal input such as incorrect attribute values; it can shield common misoperations and warn you of the deletion of important data and confirmation prompts; the ability to recover the system and data in the event of a failure; the ability to recover from abnormal exits or damage caused by operating errors or illegal inputs; in the case of failure, the ability to remedy the last saved or automatically saved data.

For the time characteristic and resource utilization testing of the management system, including the time required for the system to complete the specified task; the average time required for the system to execute several parallel tasks; the limit of the time required for the system to complete a task under the maximum load; the system throughput can successfully complete the specified task within a given time period Quantity; the degree to which a central processor, memory, network, and other resources are used when the system runs a prescribed task or concurrent access in a prescribed operating environment.

4.5 Security testing of database management system

For security, test communication safety, and system equipment safety to meet design requirements. Test whether the operating system, application software system, network system and equipment related to the system use the identity authentication and access control mechanism, the management system is not illegally used and accessed by unauthorized users; whether there is a password policy of management system, etc. Test whether the relevant technology is used to ensure the integrity of the transmitted data. Test whether the system has backup and recovery capabilities, whether the system backup meets design requirements; whether the system is portable; whether the system has log management capabilities.

4.6 Documentation test

The documents quality mainly tests the completeness and normalization of related technical documents and database system operation and maintenance records, quality control records, backup systems, and security and confidentiality systems.

5. QUALITY EVALUATION SYSTEM

5.1 Design Evaluation Indicators

For each quality element and its test items in the quality model of the fundamental geographic information database system, design compliance determination, error rate, function coverage rate, test case pass rate, etc. as quality evaluation indicators.

Compliance determination is to judge whether the execution results of the test items meet the design requirements of the fundamental geographic information database system. The test results are only in compliance or non-compliance.

The error rate is the ratio of the number of errors to the total number of elements (items) in the inspection range. For the fundamental geographic information database, this indicator is mainly used to evaluate the thematic accuracy of the data set quality. The corresponding test items that use the error rate as the evaluation indicator can be sampled. Generally, samples are taken at a rate of 3% -5% The tester will determine in the test design according to the actual situation of the basic geographic information database under test.

The function coverage rate is the degree of ratio of the function points that have been implemented in the management system to all the function points that the design requirements should have in accordance with the design requirements of the fundamental geographic information database system.

The test case pass rate is the test pass rate of a test item for all its related test cases.

5.2 Set Passing Conditions

For each test item of each quality element and quality sub-element, a specific quality evaluation indicator is set.

(1)For the test items that set conformity determination, the passing condition is that the test result is in accordance with the design requirements of the database system.

(2)For the test items that set error rate, function coverage, and test case pass rate as evaluation indicators, specific tolerance values will be given, and the corresponding test item's pass condition is that the test results do not exceed the tolerance value. For example, for the attribute correctness test item of the data set quality, the quality evaluation indicator is the error rate and the tolerance is set to 0.3%; for the functional correctness test item of the management system quality, the quality evaluation index is the test case pass rate and the limit value set the important functional pass rate to 100% and the general functional item pass rate to not less than 95%.

5.3 Judgment Evaluation Conclusion

The quality evaluation conclusions of the fundamental geographic information database system are set to pass or fail, and the method of gradual evaluation and overall judgment is adopted. According to the test item-quality sub-element-quality element-database system, the evaluation is carried out at four levels, specifically: when all the test items pass, the corresponding quality sub-element evaluation results are passed; all quality sub-elements pass, the corresponding quality element evaluation results are passed; all quality elements pass, the overall fundamental geographic information database system quality evaluation conclusion is passed.
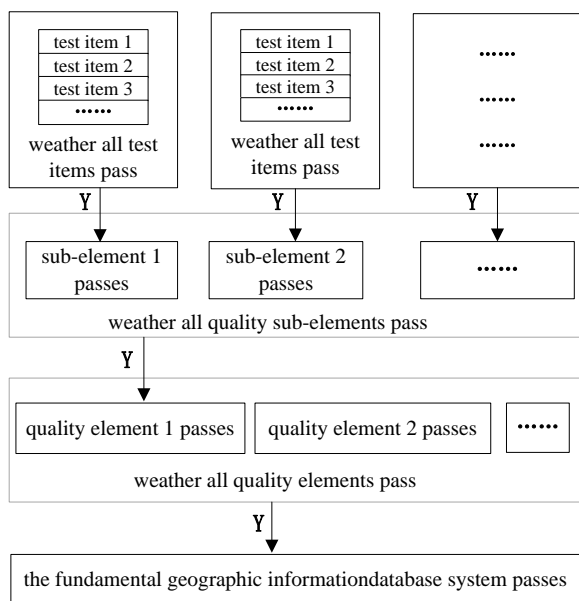

Figure 2. Step by step quality evaluation

6. APPLICATION AND ANALYSIS

6.1 Application

Using the quality model, quality testing technology and quality evaluation indicators proposed in this paper, the relevant national standards have been formulated and applied to the quality testing and evaluation of national 1:50,000 fundamental geographic information database and Shaanxi, Jiangsu and other provinces 1:10,000 fundamental geographic information database, which effectively guarantees the quality of database system. The process is shown in Figure 3.

Through the application, the feasibility and scientificity were verified, and the emphasis of quality evaluation is further focused, and relevant test items were adjusted and improved. mainly include:

(1)In terms of the quality of the data set, the consistency of the data before and after the storage is mainly checked, and the test items related to the quality requirements of the product in the production process are no longer tested, such as the classification correctness test item; Topological consistency test emphasizes the consistency between dataset topology relation and before storage.

(2)The quality of the management system revolves around the users and service objects of the system, focusing on testing and evaluation of functions, performance, and security, and does not involve testing content such as ease of use, compatibility, and satisfaction of commercial software.
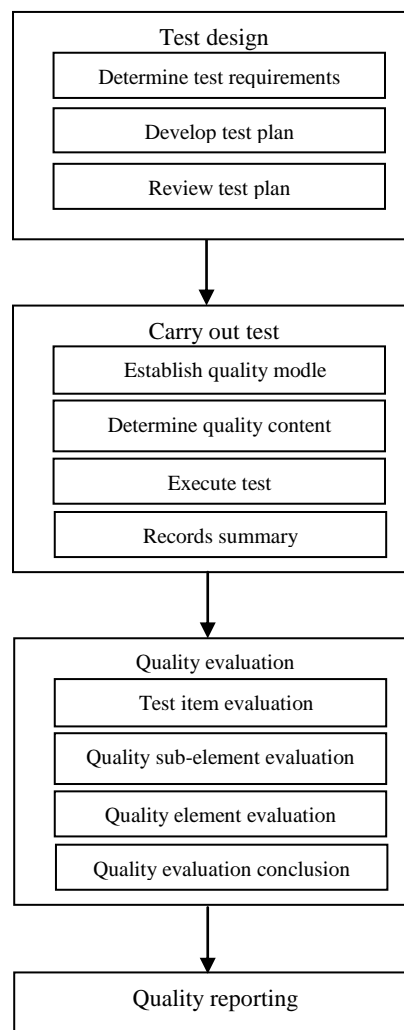

Figure 3. Quality testing and evaluation process of fundamental geographic information database system

6.2 Quality Problem Analysis

During the construction and update of the fundamental geographic information database system, there are many different production links and processes, and each link and process will generate certain errors. According to the law of error propagation, all these artificial or non-human errors will directly affect the quality of the final data results, which in turn

affects the quality of the database system. The reasons for these quality problems include the following:

 (1)The contradictions between data sources. Due to the complexity and variety of various basic geographic information data collection methods, operation processes, and mapping and utilization materials, the content and accuracy of the data are inconsistent. Contradictions will inevitably occur between various data sources, so quality problems are prone to occur.

(2)Differences in data processing procedures. In the process of production, processing and storage of basic geographic information data, it is necessary to go through multiple layers of complex procedures. The operation methods and types of errors generated by different processes and different data processing of the same process are also different, and then they are propagated and accumulated into the final database results.

(3)Differences in data processing platforms. Due to different data operators and processing environments, the processing platforms they use for fundamental geographic information data processing, data exchange, and conversion are also different, which affects the quality of database system results.

(4)Differences in operator capabilities. The limitation of the data operator's own ability also has a great impact on the quality of the data, so it will cause some quality problems.

7. CONCLUSION

According to the requirements of fundamental geographic information database system construction, this paper establishes a set of quality model and evaluation system, determines the content, technical methods and evaluation indicators of quality testing and evaluation, which can cope with the massive database system with large amount of data, multiple types of data and complex data structure. As the service scope of basic geographic information data is more extensive and the service level is deeper, the quality requirements of its database system are more strict. The technical methods and evaluation system proposed in this paper are suitable for quality testing and evaluation of the construction units, management departments and quality assessment institutions of the fundamental geographic information database management system, and will play a positive role in guiding the production, quality control, distribution and use of the results of the fundamental geographic information database in China.

ACKNOWLEDGEMENTS

REFERENCES

Shang Yaoling, Wang Donghua, Liu Jianjun, Zhang Yuanjie. Study on Quality Control of Fundamental Geographic Information Data [J]. Geomatics World,2012.2:13-17.

Ma Wei, Zhang Jixian, Zhao Yousong, Zhang Pengcheng, Dang Yu, Zhao Tingting. Design and Establishment of Quality Model of Fundamental Geographic Information Database. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-3, 2018:1249-1252

Yu Qian, Su Weibo. Research on quality inspection of urban basic geographic information database[J]. Journal of Geomatics, 2012.12:46-47.

Zhenhua Jiang, 2019:Discussion on quality management and system control of surveying and mapping engineering[J]. *Residential and Real Estate*,Jun,2019,272

Song Xiaoling. Discussion on quality control of provincial fundamental geographic information database[J]. Gansu Science and Technology,2006.10:123-124.

Chen Jun, Li Zhi-Lin, Jiang Jie, Zhao Ren-Liang. Key Issues of Continuous Updating of Geo-spatial Databases[J]. Geomatics World, Vol .02,NO .5, Oct ., 2004:1-5

Wang, D. H. , Liu, J. J., Shang, Y. L., Ji, J. P., & Wang, Z. X. (2001). Quality Control in Building 1:50000 Digital Elevation Model Database. Remote Sensing Information.

ISO 19157-2013(E),Geographic information — Data quality[G].

General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Specifications for inspection and acceptance of quality of digital surveying and mapping achievements (GB/T 18316-2008)[G].Beijing, 2008.

General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Specifications of geographic information system software testing(GB/T 33447-2016) [G].Beijing, 2008.

General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Specifications for urban database construction of fundamental geographic information(GB/T 21740-2008) [G].Beijing, 2008.

General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Specification of computer software testing(GB/T 15532-2008)[G].Beijing, 2008.

General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Specifications for quality inspection and acceptance of surveying and mapping products(GB/T 24356-2009)[G].Beijing, 2009.

General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Basic specifications for fundamental geographic information database(GB/T 30319-2013)[G].Beijing, 2013.

General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Specifications for database construction of fundamental geographic information(GB/T 33453-2016)[G].Beijing, 2016.

General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Systems and software engineering—Systems and software Quality Requirements and Evaluation (SQuaRE)—Part 10: System and software quality models(GB/T 25000.10-2016) [G].Beijing, 2016.