

Auxiliary Sampling Technology and System for Mass Surveying and Mapping Results Oriented to Quality Inspection Goals

MA Wei ¹, ZHANG Pengcheng ¹, ZHANG Jixian ¹, SHANG Yaoling ¹, ZHAO Yousong ¹, DONG Shuai ¹, WANG Jindong ¹, ZHAO Tingting ²

¹National Quality Inspection and Testing Center for Surveying and Mapping Products, Beijing 100830, China

²National Geomatics Center of China, Beijing 100830, China

KEY WORDS: Surveying and mapping results, Quality inspection, Auxiliary system, Stratified random sampling

ABSTRACT:

Based on the requirements of quality supervision inspection for national surveying and mapping, this paper has carried out research on sampling and auxiliary sampling technology and sampling system for quality inspection targets for different objects and different types of surveying results. The study designed the quantitative random sampling scheme of the random inspection objects considering the principle of supervision and spot check of surveying and mapping quality. It also designed the stratification scheme of surveying and mapping results under different quality inspection targets and object conditions, and the stratified adjustment scheme with prior quality information. Based on the above scheme, the national surveying and mapping supervision inspection sampling system based on the project rolling pool and the stratified random auxiliary sampling prototype system for the quality inspection target of surveying and mapping results are implemented. In recent years, they have been widely used in the national surveying and mapping geographic information supervision inspection and surveying and mapping results quality inspection, which guaranteed the scientific and reasonable determination of surveying and mapping quality supervision inspection objects, and solved problems in the quality inspection of surveying results, such as unreasonable stratification in manual sampling, and inevitable error in random sampling. This is of great significance for further improving sampling efficiency, reducing sampling error, reducing sampling risk, and promoting informatization of quality inspection.

1. INTRODUCTION

While the sampling of surveying and mapping results based on statistical techniques is an important part of quality control and inspection of surveying and mapping results, statistical sampling survey is a necessary means to obtain, monitor and measure the quality information of surveying and mapping results. Compared with advanced inspection means and accurate test results, the scientific sampling method is a prerequisite for accurately and reasonably inferring the overall quality of surveying and mapping results. The relevant national standards for the quality inspection of surveying and mapping results also stipulate the statistical sampling test method for inferring the quality of surveying and mapping results by inspection samples and using a small sample size to judge the overall quality as accurately as possible. Stratified random sampling provides an effective method for the efficient organization and implementation of the quality inspection of surveying and mapping results and the improvement of the accuracy of quality estimation under the conditions that the quality inspection target is clear, the result type characteristics and distribution characteristics are significant, the quality characteristics of the results and the prior quality information are available. For different types of surveying and mapping results and sampling units, this study carried out research on the design of an auxiliary sampling prototype system for quality inspection of surveying and mapping results, which could effectively improve the science of stratification and sampling implementation, and solve problems in the quality inspection of surveying results, such as unreasonable stratification in manual sampling, and inevitable error in random sampling. This is of great significance for further improving sampling efficiency, reducing sampling error, reducing sampling risk, and promoting informatization of quality inspection.

2. TECHNICAL CONTENT

2.1 Technical Route

The study and design on the stratification scheme of surveying and mapping results under different quality inspection targets and object conditions, the study on the stratification adjustment scheme under the condition of prior quality information, and the design of the Stratified Random Auxiliary Sampling System for Quality Inspection Targets of Surveying and Mapping Results based on the above schemes.

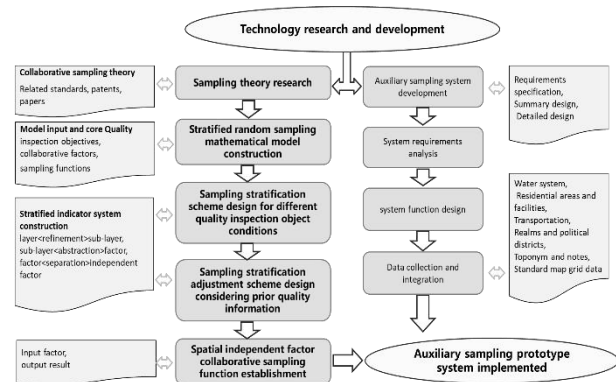


Figure 1. Technical Route

2.2 Technical Method

Sampling theory has been researched in terms of sampling standard, research papers, and patents via literature review and technical research. Established a theoretical basis for implementing auxiliary sampling plan for surveying and mapping results oriented to quality inspection targets.

In terms of researching auxiliary sampling, carrying out project design based on research and analysis. Proposal for layering of mapping result under different quality inspection goals and proposal for adjusting the layering with procession of a priori conditions of quality information have been researched and designed. They are in accordance with current quality inspection standard and stratified random sampling theory.

Work on technology selection, requirements analysis, overall design, and detailed design of the auxiliary sampling prototype system for surveying and mapping results oriented to quality inspection targets have been conducted. These tasks are accomplished with reference to theory and methodology of software engineering. The foregoing research aims to clarify the major function point and function flow of the software. Based on the instruction manual of software requirement and specification, the overall structure of the function design software system has been decided according to requirements analysis. Software proposal is composed through demarcating function modules as well as deciding the algorithm and compiling codes for each module. It is the basis of accomplishing the coding and software system's functions.

In the matters of system integration and testing, methods of black-box testing and white-box testing have been employed. The software passed through integration testing has been combined with other parts of the system in order to conduct a series of strict and effective testing under actual runtime environment. These testing are designed to detect the potential problem so as to ensure the normal operation of the system. They aim to test and verify whether the final software system can meet the client's requirements. Major testing features include function test that examines correctness of the software system's functions as well as the integrity and correctness of its functions with regard to the requirements. Another major test feature is to test the software's robustness by analyzing its fault tolerance and recovery capability.

The developed system has carried out a sampling experiment on actual quality inspection project of mapping in order to verify the technology. This aims to determine whether the auxiliary sampling technology and system for surveying and mapping results oriented to quality inspection targets have satisfied actual demands. Necessary adjustments have been applied to the sampling proposal based the experiment result.

2.3 Data Preparation

Collaborative stratified sampling proposal has been decided based on the preliminary sampling theory research. According to this proposal, the sampling process should integrate with related a priori knowledge and incline on particular impact factors. Factors relating to the sample need to be considered during sampling process include rationality of its spatial distribution, the proportion of its topographical category, the systematicness of its result category, the degree of mapping units being comprehensive, the focus of important elements, and the essentiality of inspection items. Concerning the operability of the sampling system and the randomness of the samples are also required. A set of scientific and reasonable auxiliary sampling system is needed to be developed under on the conditions of these impact factors of sampling process. Apart from researching auxiliary system, collection and coordination of basic data that requires by the system are also conducted simultaneously. Administrative divisions, standard scale grid, and basic geospatial statistics have been collected and coordinated to structure the auxiliary sampling system.

The mapping result will usually be produced and stored according to administrative unit (province, municipality, and county, etc.) or grid unit (different scales including 1:10 million,

1:50,000, and 1:500). Therefore, auxiliary sampling needs to be conducted in the form of organization of particular unit. The spatial statistics of each administrative unit and grid unit serve as the foundation of sampling.

Basic geospatial statistics required by auxiliary sampling also include highway, railway, administrative division, important geographical name, and network of rivers. Compiling and restructuring these statistics while keeping their important attributes including name, ranking, and category are needed to integrate them as impact factors of the random sampling mathematical model in the process of operating the sampling model.

3. MATHEMATICAL MODEL

3.1 Stratified Random Sampling Model

The stratified random sampling has high sampling efficiency and high estimation accuracy. It can not only calculate the overall index but also calculate the indicators at each level. This facilitates the organization of the sampling work.

$$W_h = \frac{N_h}{N}$$

$$f_h = \frac{n_h}{N_h}$$

$$\bar{Y}_h = \frac{1}{N_h} \sum_{i=1}^{N_h} Y_{hi}$$

$$\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}$$

$$\sum_{i=1}^{N_h} Y_{hi} = N_h \bar{Y}_h = Y_h$$

$$\sum_{i=1}^{n_h} y_{hi} = n_h \bar{y}_h = y_h$$

$$S_h^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (Y_{hi} - \bar{Y}_h)^2$$

$$S_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2$$

W_h - Weight of layer h ,	f_h - Sampling ratio,
\bar{Y}_h - Overall average,	\bar{y}_h - Sample average,
Y_h - Overall amount,	y_h - Total sample,
S_h^2 - Overall variance,	s_h^2 - Sample variance.

3.2 Auxiliary Sampling Model

Mathematical model :

$$R = \text{SAMPLING} \{ G | C, D, K, W, S \}$$

The model parameters are designed as follows:

Quality inspection goal (G) – batch, sample volume, sampling principle and method, sampling unit results unit (based on the determined sample volume, design a scientific and reasonable sampling mathematical model to reduce sampling error, improve quality estimation accuracy and reduce sampling risk)

Parameter of Product Characteristics (C, Parameter of Product Characteristics) – Project Type (geographical monitoring of national conditions, 4D products), results type, results distribution, spatial distribution, results organizing unit and spatial complexity of results elements

Parameter of Geo-spatial Distribution (D, Parameter of Geo-spatial Distribution) – transportation, topography, landform, land type, place name, type of difficulty, distribution of urban and rural areas, economic development

Parameter of Prior Quality Knowledge (K, Parameter of Prior Quality Knowledge) – estimation of the quality status and problem distribution of the results based on the characteristics of the results and the actual production and quality control situation

Collaboration Weight (W, Collaboration Weight) - Influence the weight of collaborative sampling factors in sampling results

Random Seed (S, Random Seed) - random number, stratified random sampling, taking into account various parameters

4. DESIGNED AND DEVELOPED SOFTWARE

Four software function modules determined by requirement analysis ① Data Load Module ② Sampling Basic Information Configuration Module ③ Sampling Scheme Configuration Module ④ Sampling and results export module, it completes the design and development of software system functions based on the ArcGIS platform.

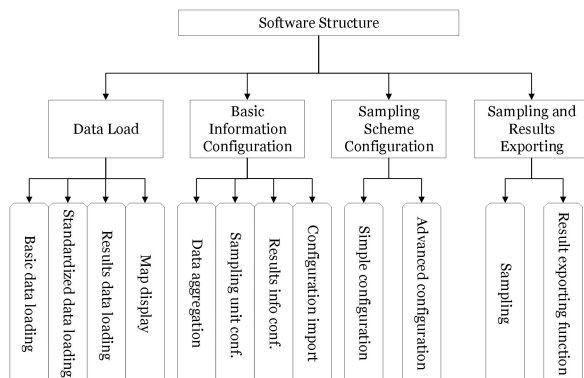


Figure 2. Software Structure

4.1 Data Load Module

The data load module includes four functions: basic data loading, standardized joint table loading, result joint table loading and switching map display status.

(1)Basic data loading functions: the spatialized data such as water system and path collected in this project can be loaded into the system through the interaction of buttons and dialog boxes.

(2)Standardized joint table load function: the current collection and collation of administrative areas (provincial, municipal and county level 3) or grid data (1:1 million, 1:500,000, 1:250,000, 1:100,000, 1:50,000, 1:25,000, 1:10,000, 1:5,000, 1:2000, 1:1000, 1:15, 11 categories) can be loaded into the system according to the sampling requirements through the interaction of buttons and dialog boxes.

(3)Load function of results joint table: the standardized results joint table is loaded into the system through the interaction of buttons and dialog boxes.

(4) Function of switching map display status: Interact with the users through the checkbox to display the status switching of true or false, which is suitable for avoiding the false death of the

ArcMap system and blocking the subsequent operations when a large volume of data is loaded.

4.2 Sampling Basic Information Configuration Module

The sampling basic information configuration module consists of four functions, namely, joint table aggregation, sampling unit display (configuration), results information display (configuration), and sampling configuration file loading.

(1) Joint table aggregation function: the connection between standardized joint table and result joint table is implemented through button interaction, and data preparation for subsequent sampling is made.

(2) Sampling unit display (configuration) function: the system automatically reads the basic data, standardized joint tables and result joint tables imported in the preceding steps, and displays the information of the sampling units in this sampling on the system interface, including sampling units (sampling by administrative units, sampling by map sheets), district grade or map scale.

(3) Results information display (configuration) function: the system automatically reads the basic data, standardized joint tables and result joint tables imported in the preceding steps, and displays the information of the results of this sampling on the system interface, including the total amount of results, batch situation, total sample size, number of production units, topographic conditions, etc.

(4) Sampling configuration file load function: users are allowed to import sample profiles to customize the number or rules of this sample by interacting with buttons and dialogs.

4.3 Sampling Scheme Configuration Module

Sampling scheme configuration module includes two functions: simple random configuration and conditional random configuration.

(1)Simple random configuration: By checking the box, the sampling scheme is configured as simple random sampling.

(2)Conditional random arrangement: By checking the sampling factors (including the production unit factor, achievement distribution factor, achievement edge factor, terrain factor, water system factor, traffic factor, resident factor), and setting the buffer zone of the specific factor (water system traffic) (the default is 0 meters, i.e. passing through), the sampling scheme is configured as conditional random sampling.

4.4 Sampling and Results Exporting Module

Sampling and result exporting module includes two functions, one is sampling function and the other is result exporting function.

(1) Sampling function: Interact through the button and perform sampling according to the sampling configuration mentioned above, and the progress bar is displayed during the sampling process. After the sampling is completed, the “Sampling success” or the “Sampling failure (the sampling requirement cannot be met)” is prompted.

(2) Result exporting function: after successful sampling, the sampling results can be exported to spatial data (shp format) through button interaction.

4.5 Software Development and Testing

The software has been developed by using C# language according to the requirement specification and detailed design of the software based on ArcGIS platform. Unit testing and integrated testing were carried out in time in the process of

system development and after the completion of development. Problems were found and positioned in time through multi-round testing of various real and simulated data, and improvements were implemented. It effectively improved the stability, robustness and operational efficiency of the system. The subsequent system test showed that the function and efficiency of the software system have reached the design expectation.

Padilla, M., Stehman, S. V., Ramo, R., et al, 2015: Comparing the accuracies of remote sensing global burned area products using stratified random sampling and estimation. *Remote sensing of environment*, 160, 114-121.

Wei, D. S., Zhou, X. G., 2019: Automatic sampling of remote sensing image change detection samples based on prior information of vector data. *Journal of Remote Sensing*, 23(3), 464-475.



Figure 3. Software interface

5. CONCLUSION

5.1 Promotion and Application

The technology and software system developed in this paper have been effectively verified and applied in the third national land survey satellite remote sensing orthophoto quality inspection, the third national land survey aviation orthophoto quality inspection, the entire 1:2000 digital orthophoto and digital elevation of Foshan City.

5.2 Analysis of Effectiveness

Testing experiments and practical application show that the technology and software system developed in this paper are universal to all kinds of surveying and mapping results. It can effectively improve the scientific nature of stratification and sampling implementation, solve the problem of unreasonable stratification and unavoidable random sampling error in the current sampling and quality inspection manual sampling, further improve sampling efficiency, reduce sampling error and reduce sampling risk. It has effectively promoted the informatization of quality inspection to a new level.

ACKNOWLEDGEMENTS

This study is sponsored by National Natural Science Foundation of China (41671440).

REFERENCES

Li, H., Xu, H., Peng, B., 2007: Comparison of Two Statistical Models in Contingent Valuation Method. *Scientia Geographica Sinica*, 27(1), 115.