Design and Implementation of Yunnan Province Soil Environmental Quality Database Based on ArcSDE for Oracle

Yi Zhao¹, Hong Liang ¹*, Ningshan Zhou¹, Juan Zhao², Yuxia Zhang² and Yinxia Peng²

¹College of Information Science & Engineering, Yunnan University, 650500 Kunming, China
²Yunnan Environmental Monitoring Center, 650034, Kunming, China
Email: ynliangh@126.com

Abstract. The environmental quality monitoring work in Yunnan Province began in the 1970s. After 40 years of development, the Yunnan Province Environmental Monitoring System was initially established. However, the soil environmental monitoring data of the province lacks a unified warehousing standard, and a spatial database of soil environmental quality has not been established to integrate and standardize the data. Based on the research and reference to the relevant standards of soil database construction method and management at home and abroad, GeoDatabase is used as the spatial data model, and the conceptual model, logical organization model and storage model of Yunnan Province soil environmental quality database are designed by using Oracle and ArcSDE technologies. In order to accurately reflect the spatial location of the soil sampling points and the characteristic properties of the samples, and to ensure the uniqueness and rationality of the primary key in the data table, this paper designs the coding of the soil environmental quality database. The construction of soil environmental quality database in Yunnan Province realized the seamless integration of spatial data and attribute data, and provided comprehensive and accurate basic data and scientific decision support for soil pollution prevention and soil ecological restoration in Yunnan Province.

1. Introduction
In recent years, with the national attention to soil environmental protection work and the development of soil pollution prevention and control work in various departments of Yunnan Province, the trend of soil pollution aggravation in the province has been initially controlled. However, the lack of uniform storage standards for soil environment-related data in the province has resulted in massive fragmentation and decentralization of soil environmental quality data. Data accuracy and reliability are poor, and it is impossible to provide accurate and comprehensive basic data for management departments at all levels, especially Lack of effective comparative analysis of large amounts of historical and recent data. Therefore, the establishment of soil environmental quality database is the basis for soil research and soil environmental quality warning. Foreign work on establishing a database of soil environmental quality has been carried out for many years. In 1966, the US Department of Soil Conservation (SCS) established a new and comprehensive National Soil Information System (NASIS)[1]. Canada began construction of the Soil Information System (CanSIS) in 1963 and was put into operation in 1971 [2]. The United Nations Environmental Protection Agency hosted the Global Soil Degradation Assessment (GLASOD) in 1987 to demonstrate the establishment and application of the SOTER database [3]. The research and
application of China's soil database started late. After the introduction of SOTER method into China, various regions and related departments have carried out related research. The key point of establishing soil database by SOTER method is the mapping of SOTER units, while the database of soil environmental quality focuses on the management and application of the data of various types of soil sampling points, requiring that the data of soil profile sampling points be stored as spatial data [4,5]. The particularity of soil environmental quality survey data determines that the soil environmental quality database is not applicable to the existing SOTER methodology. Therefore, it is particularly important to select a reasonable data organization method to design the soil environmental quality database for the analysis and evaluation of soil environmental quality.

2. Design of Soil Environmental Quality Database in Yunnan Province

2.1. Design of database conceptual model
The soil environmental quality data of Yunnan Province covers a wide range of types, and the soil environmental quality data can be divided into basic geographic data, basic background data, and soil quality survey data and metadata. The basic geographic data includes remote sensing image maps of various cities and counties in Yunnan Province, as well as vector data and raster data such as administrative maps, water system maps, and traffic maps. The basic background data includes the distribution data of multi-target geochemical elements in Yunnan Province, land cover type data, agricultural land type data, and enterprise data of key polluting industries. The soil quality survey data refers to the multi-scale soil environmental survey data carried out by the land and resources department, the agricultural department and the environmental protection department since the 1980s, including the province's soil background value survey data, soil pollution status survey data, routine monitoring data of soil environmental quality and agricultural product survey data. Among them, the soil quality survey data is divided into 68 selected and measured indicators according to the national environmental quality standards and the characteristics of ecological environment conditions in Yunnan Province according to the needs of different survey items in different departments. The metadata of the database is the data dictionary content such as the basic description information of the recorded data, the field code, and the soil evaluation standard. The conceptual model design of the database is shown in Figure 1.

![Figure 1. Conceptual Model Design of Soil Environmental Quality Database](image)

2.2. Design of Database Logical Organization Model
The spatial data model GeoDatabase uses object-oriented technology to abstract the real world into feature classes (points, lines, polygons) containing spatial information and object classes (feature properties) that do not contain spatial information. Each feature class and object class has properties, behaviors, and rules, and each feature establishes an association between objects such as relationship classes and geometric networks. [6]. In this paper, the GeoDatabase spatial data model is used to hierarchically design, process and manage data. The data layer is divided into three types: raster data layer, vector data layer and attribute data layer. The vector data is mainly the soil survey point data of each department and the vector data in the basic geographic data and the basic background data and is organized and stored by the Feature Dataset. The raster data is mainly remote sensing image data, and
the image catalog is organized and managed by Raster catalog. The attribute data mainly includes data of soil environmental monitoring indicators and metadata information, etc., which are mainly stored in Table and the relationship between spatial data and attribute data is well solved by the Relationships model in the GeoDatabase model. Geodatabase realizes the seamless integration of spatial data and attribute data. All spatial data layers should be set with a unified coordinate system CGCS2000, so that different data layers can be superimposed and visualized in the same space [7]. The logical organization model is designed as shown in Figure 2.

![Figure 2. Design of Database Logical Organization Model](image)

### 2.3. Physical storage model and data coding design

The soil environmental quality database of Yunnan Province needs to design the table structure to reasonably organize the soil quality survey data. The soil quality survey data is the core data of the database. Different departments must monitor the necessary and selected indicators according to the different needs of the survey objects and survey purposes, but they can all be summarized as basic survey data (only collected from the surface 0 ~ 20 cm soil surface mixed sample data), including soil quality survey field record table, soil sample physical and chemical properties table, soil sample inorganic matter monitoring table, soil sample organic matter monitoring table. For the soil background value survey data table, the soil sample profile record table was added to sample the soil profile in the surface, middle and deep layers. The soil sample profile record table and the soil sample physical and chemical properties table, the inorganic matter, and the organic matter monitoring table establish a one-to-many relationship through the sample code. For the risk point monitoring data table, on the basis of the basic survey data table, the agricultural product sample field record table, the agricultural product sample inorganic material test table, and the agricultural product sample organic matter test table are added. This paper takes the design of the soil background value survey data table as an example and shows it as shown in Figure 3.
In order to accurately reflect the spatial location of the soil survey sample and the characteristic properties of the sample, and to ensure the uniqueness and rationality of the primary key in the data table, it is necessary to encode the data in the database [8]. The design of data coding should conform to the basic principles of clear hierarchy, simple and clear, and easy to expand. The sample code uses 22-bit code, and the first to sixth code are administrative division codes above the county level and county level. The coding scheme refers to the Administrative Division Code of the People's Republic of China (GB/T2260-2007), for example: The administrative division code of Fumin County, Kunming City, Yunnan Province is 530124. The seventh digit to the tenth digit code are the investigation department and the investigation item number, and the eleventh to thirteenth digit code is the total number of sampling points in the county space calculation unit, for example, the coding number of the sampling point No. 60 is 060. The fourteenth to sixteenth digit code is the depth lower limit value (in centimeters) of the soil sampling section. For example, when the lower limit of the sampling depth is 20 cm, the code is 020, and the seventeenth to twenty-second code is the time element code. The main investigation time is recorded. For example, in January 2019, it is recorded as code 201901. The sample code uses a 9-bit code and uses the first 12 bits of the above sample code to display the geographical location and number of the sample point. The soil sample code is shown in Figure 4.

3. Implementation of Yunnan Province Soil Environmental Quality Database System
This paper uses the "database management system (Oracle) + spatial data engine (ArcSDE)" model to design and organize soil environmental quality data, and achieve integrated management of spatial data and business data. The Geodatabase model provided by ArcSDE is used to organize complex spatial
data, and the spatial and non-spatial attribute data are organized in a two-dimensional relational table and stored in the commercial database Oracle, and the two kinds of data are correlated by an internal association code. ArcSDE, as a middleware between the user client and the Oracle server, does not store data. Instead, it stores the spatial data in oracle in multiple associated tables[9,10]. The client application makes a spatial data request to the ArcSDE server through the ArcSDE application interface (ArcSDEAPI). ArcSDE uses the characteristics of the spatial data object to complete the search results locally and return the result to the client application. This paper takes the soil-related data display of Kunming City, Yunnan Province as an example (Figure 5). The soil environmental quality database system can realize the management and update of data such as input, deletion and modification of vector feature data and attribute data. The system can dynamically customize thematic maps and implement spatial data analysis according to actual needs.

**Figure 5.** Spatial data of soil environmental quality in Kunming City, Yunnan Province

**Conclusions**

This paper uses Esri’s spatial database engine (ArcSDE) and large commercial relational database software (Oracle) combined with Geodatabase spatial data model to design and construct a spatial database of soil environmental quality in Yunnan Province. The database realizes hierarchical storage and management of various forms of data such as raster data, vector data, and non-spatial attribute data. The spatial database of soil environmental quality fully considers the characteristics of soil profile data in one-dimensional and multi-layered space. The sampling point code and soil sample code of soil survey data are designed to realize the correlation between soil spatial data and soil attribute data. The establishment of the database enhances the spatial visualization capability of soil environmental quality data in Yunnan Province and enhances the management efficiency and management level of the environmental monitoring and management department in Yunnan Province. It provides powerful
technical support for the subsequent spatial analysis of soil environmental quality and soil environmental quality assessment and early warning work.

References
[1] Zhao Q. G, Shi X. Z., 2007 *Introduction to soil resources* (Science Press) p 123-138.
[2] Chen H. W, Zhang F. R., et al. 1997 *Soil resource survey*. (China Agricultural University Press) p 68-90
[3] Zhang G L, Wu Y J, Zhao Y G. Physical suitability evaluation of reserve resources of cultivated land in China based on SOTER 2010 *Transactions of the Chinese Society of Agricultural Engineering* 26(4):1-8.
[4] Yao L, Zhang L P, Zhang S J, et al. Design and Implementation of Soil Environmental Quality Database 2011 *Soils*
[5] Arrouays D, Mckenzie N, Hempel J, et al. GlobalSoilMap: basis of the global spatial soil information system. 2014 *Soil Science Society of America Journal* 79(5):1519.
[6] JIN X., Dong S. C., et al. Design and Implementation of Spatial Database of Liangzhu Ancient City Site in Zhejiang Based on ArcGIS Geodatabase.2018 *Journal of Nanjing University* 54(1).
[7] Jiang Z, Zhang P, Jinzhong B I. The Model of Crustal Horizontal Movement Based on CGCS2000 Frame. 2009 *Acta Geodaetica Et Cartographica Sinica* 38(6):471-476.
[8] Yang X H, Huang Y H, Dong P L, et al. An updating system for the gridded population database of China based on remote sensing, GIS and spatial database technologies 2009 *Sensors* 9(2):1128-1140.
[9] Xiong L H, Yang F. Application Study of Spatial Database Technology Based on ArcSDE. 2004, *Computer Applications* 24(3):90-92.
[10] Xiaomei Y , Yunyan D. U., et al. Discussion on Methods of Building Image Database Using ArcSDE --A Case Study of Building RS Image Database of Chinese Coastal Zone 2012 *Geo-information Science* 4(4):16-23.