

# RESEARCH ON NATURAL RESOURCES SPATIO-TEMPORAL BIG DATA ANALYSIS PLATFORM FOR HIGH PERFORMANCE COMPUTING

YIN GAO<sup>1</sup>, JIANWEI LUI<sup>1</sup>, YANG LIU<sup>2</sup>, ZHAOKUN ZHAI<sup>1</sup>, JIAN CHE<sup>1</sup>, HAO LI<sup>1</sup>, RU WANG<sup>1</sup>, JIANJUN LIU<sup>1</sup>\*

<sup>1</sup> National Geomatics Center of China, 100830, Beijing, China

<sup>2</sup> Changjiang Waterway Bureau Survey Center, 430010, Wuhan, China

**KEY WORDS:** Natural Resources, Cloud Architecture, Spatial-temporal Big Data, Platform Framework, Spatial Analysis, High Performance Computing.

## ABSTRACT:

In the era of earth observation big data, a new paradigm for the unified management of natural resource conservation and utilization has been established in China. The current focus of natural resources monitoring is on fully leveraging the value of big data in geospatial and temporal dimensions to support new work on natural resources management. An overall framework for a spatial-temporal big data analysis platform is proposed in this paper, which explores innovative technologies such as heterogeneous cloud collaborative services, business capabilities, data resources, and open sharing. Utilizing private cloud computing resources, the platform establishes collaborative services for storage cloud, computing cloud, and database cloud, demonstrating its ability to perform large-scale online computation of spatial-temporal data. Mainstream spatial data analysis computing methods such as vector computing, grid computing, and other grid computing methods are also established, and an open service model for interdisciplinary applications is explored. The feasibility of the proposed platform is verified using nationwide land cover data of 260 million surface area classifications, indicating its great potential for application.

## 1. INTRODUCTION

Natural resource monitoring data has the typical characteristics of rich information, authenticity and accuracy, global coverage, high spatial resolution and strong timeliness. It is the authoritative spatio-temporal information data for governments at all levels to find out the resources and carry out planning and decision-making (Ferreira and Vale 2022; Li 2014). It has unique advantages in solving the spatial data support in resource investigation, ecological protection, spatial planning, urban management and macro-decision. After years of continuous monitoring, China has obtained full-coverage geospatial data, identified the status and spatial distribution of natural resources and built a large-scale database (Liu et al. 2019). Additionally, the country has also mapped the distribution of artificial and public service facilities, established a private cloud space computing infrastructure, and achieved efficient management and service of China's spatio-temporal data on geographical conditions. At present, the monitoring of geographical conditions has been carried out regularly. The strategy of "covering the whole country and highlighting key points" is adopted to carry out full-coverage monitoring once a year on the land territory, forming a mass of geographical conditions data achievements with full coverage, seamless, high precision and time sequence, greatly promoting the leap-forward development of the ability to obtain land cover information (Li et al. 2016).

Traditional spatial data management based on single-node relational databases has limitations in terms of data management, concurrent reading and writing, efficient calculation, and scalability (Yang et al. 2011). The emergence of cloud computing has presented both opportunities and challenges for the management and analysis of massive spatial data. Combining cloud computing technology with spatial database technology to achieve elastic management and online analysis of massive spatial data has become a new research field in spatial information technology. In recent years, domestic and international scholars and spatial information enterprises have conducted extensive research on spatial data cloud storage and cloud computing. For example, ESRI, a representative international company in the United States, has developed the

Esri Geometry API for JAVA and Spatial Framework for Hadoop to integrate ArcGIS and Hadoop (<http://esri.github.io/gis-tools-for-hadoop/>) and launched the GeoAnalytics Server distributed spatial computing platform embedded with Spark technology in the newly released ArcGIS Pro software, which has initial distributed analysis and computing capabilities for massive spatial data (Allen and Coffey 2019). Additionally, Oracle Corporation in the United States has launched a package solution (<http://www.oracle.com/>) that horizontally expands the Exadata database cloud platform, Oracle database, and Oracle Spatial spatial components with OLTP and OLAP for managing and computing massive spatial data. Oracle Spatial's vector and raster data distributed computing capabilities have been significantly enhanced in the latest database version of Oracle 12c (Saygili 2020).

Domestically, Chen et al. (2013) achieved distributed storage of vector and raster data based on the HDFS distributed file system and HBase NoSQL database. Jin et al. (2016) of Zhejiang University conducted in-depth research on the Spark distributed memory computing framework and developed customized spatial computing operators to realize the distributed spatial computing and analysis of large-scale land use vector data. Li (2014) committed to the research and development of the new generation of SuperMap 8C cloud architecture domestic GIS platform software, which has been practically applied in smart city big data and other fields (<http://www.supermap.com/>). The rapid development of these new technologies provides powerful technical support for the management and analysis of mass result data of geographical national condition monitoring.

The natural resource monitoring data has been extensively utilized at national, provincial, and local levels, with significant positive outcomes observed in China. Specifically, the "multiple planning integration" initiative in Hainan province utilized natural resource data to accurately identify cultivated land, forest land, and natural coastline. This enabled support for overall urban and rural development, optimized industry and facility layouts, and facilitated the realization of a comprehensive plan (Liu et al. 2018). The National Audit Office has leveraged monitoring data of natural resources to quickly ascertain the distribution status of natural resources such as forests, lakes, and wetlands, as well as natural protected

areas in 11 provinces and cities of the Yangtze River Economic Belt spanning approximately 2.05 million square kilometers. Moreover, the data on heaping and digging surfaces, industrial and mining enterprises, power station, and relevant special topics have facilitated accurate auditing of illegal activities (China 2018).

In many parts of the country, natural resource data has aided comparative analysis of changes in natural resource assets such as land, forest, water conservancy, and minerals before and after taking office during the pilot work of outgoing audit of natural resource assets of leading cadres. This provides objective, accurate, and intuitive data support for the outgoing audit. The national land supervision department employs natural resources data to carry out land supervision in four provinces in southwest China, supporting analysis of suspected problems in the office, and field verification through monitoring data. It effectively promotes the routine land supervision work of the state (Jiao and Wang 2020).

Beijing City applies natural resources data in carrying out "urban physical examinations," realizes scientific and refined urban management, and explores new ideas to solve big city diseases (Guan et al. 2020). Shanghai City uses natural resources data to prepare the Shanghai 2040 Master Plan, providing a solid guarantee for building a happy Shanghai with innovation, humanities, and ecology (Yu et al. 2017). These substantial practical achievements fully demonstrate the great application potential of natural resources monitoring data in critical fields such as the spatial planning system, ecological civilization construction, resources and environment audit, and urban governance and management in the new era.

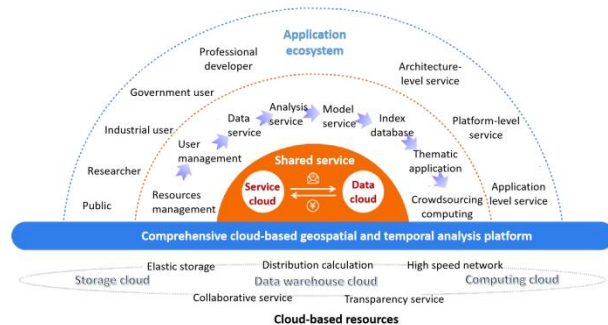
With the recent round of institutional reforms in China, a new pattern of unified management of natural resources has been established, which imposes new requirements on the management and service of natural resources in the new era. The rapid development of big data technologies, such as Earth observation, satellite remote sensing, and intelligent cities, provides an efficient means of obtaining large-scale, objective, accurate, and real-time surface information throughout the country (Thatcher et al. 2018). Future analysis and decision-making of natural resources will therefore be transformed by big data, characterized by diversification, depth, specialization, and real-time capabilities. As the material basis and basic element of the construction of ecological civilization, geographic information will be actively integrated into the overall situation of natural resources management, exploiting the advantages of massive geographic and spatial data. The goal is to accelerate the emergence of data-driven natural resources planning, management, and scientific decision-making in the era of big data.

The proposed platform in this study is a high-performance computing system for the analysis of spatio-temporal big data related to natural resources. The system is based on a national-level geographic data spatio-temporal database and incorporates big data analysis services for geographic conditions and natural resources. The overall platform structure and technical framework are presented in detail, along with the latest system construction utilizing large-scale cloud computing resources. Additionally, the study performs high-performance computing experiments using national-scale geographic data, providing a valuable reference for supporting spatial natural resources analysis services.

## 2. GENERAL FRAMEWORK

The linkage between big data and the analysis of geographical conditions is closely tied to cloud computing. As cloud computing technology has evolved, it has transitioned from a

single centralized architecture to a distributed application architecture. The most recent advancements in open and transparent cloud architecture technology have been proposed by large internet enterprises based on general cloud resource services and shared services, enabling an open and comprehensive cloud architecture that can accommodate the rapidly growing data and business needs. The traditional centralized system architecture is not able to effectively expand and is inadequate in addressing the computing service requirements of big data in the geographical conditions, particularly in support of natural resources businesses, public welfare industry data services, and other major needs (Jiping et al. 2019).



**Figure 1.** Architecture of the comprehensive cloud-based geospatial and temporal analysis platform

To achieve comprehensive cloud, on-demand access, on-demand response, on-demand service, and open sharing of big data, virtualization is used in the cloud computing environment based on distributed computing and storage, optimized hardware, clustered operation, and maintenance management system. This enables dynamic allocation and deployment of computing, storage, network, and other resources, and allows shared services according to the business center and data center. To establish an agile, scalable, shared, and open geographical big data analysis service framework, a cloud resource collaboration service, a business data sharing service, and a special application service ecology are constructed (Figure 1). This leads to a new design of infrastructure cloud, data platform cloud, business platform cloud, and application model service for geography data analysis service technology. These services provide open, efficient, comprehensive, and personalized solutions for natural resources and industry services.

### 2.1 Cloud Resource Collaboration Service

The aim of this study is to develop a supportive environment for the geographic national condition service big data by innovatively utilizing the overall design of cloud architecture. This is achieved by offering heterogeneous cloud collaboration and service coupling through the establishment of a cross-platform comprehensive cloud infrastructure that comprises storage cloud, database cloud, and computing cloud. The objective is to accomplish physical isolation and logical coupling of geographical application data, as well as business isolation and application coupling under the support of distributed cloud computing. The service realizes online collaborative and efficient service of distributed storage cloud, database cloud, and distributed computing cloud with service-oriented loose coupling. This approach comprehensively enhances computing power and expands performance, thereby providing comprehensive performance

support for professional, diversified, and in-depth big data analysis services.

## 2.2 Data Resource Sharing Service

The co-construction, sharing, and open provision of data resources and business applications are necessary for big data in the presence of complex and heterogeneous geographical conditions. This requires the use of state-of-the-art technologies and systematic approaches to reconstruct diverse geographical application data and professional capabilities. Cloud services for data resources and business applications can be realized by coupling them on a unified cloud resource platform and standard standards based on government affairs network, cloud computing, big data, security, and other technologies.

The process of integrated management applications includes resource management, user management, data services, analysis services, model services, index management, subject application, crowdsourcing computing, and other capabilities. These capabilities are designed to address the challenges of the big data era, characterized by large systems, big projects, big science, big research, big management, and big control. The employment of these capabilities can resolve traditional issues such as professional isolation, data isolated islands, and application segmentation.

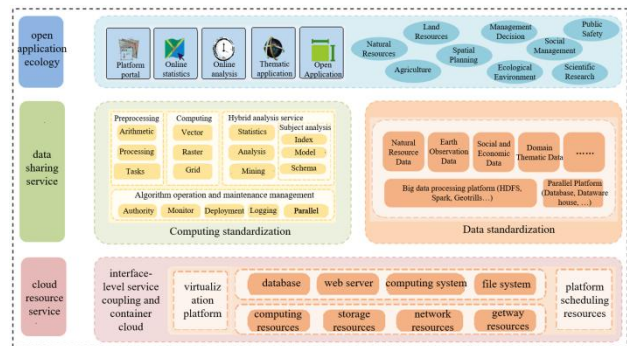
## 2.3 Open and Shared Application Ecosystem

A unified authentication and different networks are adopted to provide unified infrastructure-level services, platform-level services, and application-level services for natural resource service subjects, government departments, industry users, scientific research, and the public. Physical computing resources on the physical server hardware platform are abstracted by infrastructure services to provide users with a unified, high-performance, and secure cloud computing environment. Data processing, image interpretation, data analysis, application service management, and other services are offered by platform-level services. Customized services for specific applications are provided by application-level services. Through the construction of a large-scale, data-driven, and application-oriented service ecology, a multi-level, multi-user, multi-theme, flexible, safe, open, and shared geographical analysis service mechanism is established, creating a positive ecosystem for "big demand, big data, big service, and big application."

## 3. TECHNICAL ROUTE

The advancement of cloud computing and big data technologies has led to the establishment of business cloud service platforms in various professional fields such as telecommunications, petroleum, land, geology, and oceanography. This has facilitated the transition from the traditional centralized application mode to an open and sharing mode (Reddy et al. 2022). This has significantly improved the professional application service capability and provides valuable insights for the development of a geographical analysis platform.

Geographical condition information serves as a public welfare service that caters to both natural resource management and public industry applications, leading to a wider range of service objects, open application fields, and diverse business models. This, in turn, requires higher demands on the flexibility, scalability, and openness of the geographical national condition information analysis platform.



**Figure 2.** Elastic hybrid cloud computing ecological and distributed spatial analysis framework.

To achieve this goal, it is necessary to adopt new technologies such as heterogeneous cloud collaborative services, business cloud-based construction, and data cloud resource virtualization. These technologies can create a transparent, open, and agile cloud computing support architecture that can provide high-performance and high-availability business service capabilities (Figure 2).

### 3.1 Resource Layer

At the resource layer, which serves as the infrastructure service layer, a completely heterogeneous cloud resource pool is constructed based on interface-level service coupling, as well as container cloud and other new virtualization technologies. The database server, web server, file server, and distributed computing cluster are logically coupled and coordinated to achieve unified computing power allocation, dynamic memory management, flexible storage maintenance, and service capability migration. This enables the online collaborative and efficient services of the database cloud, distributed storage cloud, and distributed computing cloud, while providing cluster management services to achieve the service capability of "integrated management, transparency, efficiency, on-demand distribution, and on-demand expansion".

To address the traditional problem of data-intensive and computing-intensive spatial big data analysis, a lightweight computing model of micro-service agile design is adopted, with "agile service and elastic expansion" as the primary goal. Computing resources are "fine-grained disassembled, intensive services, and loosely coupled management" to achieve high reliable operation and flexible maintenance of massive supporting computing resources. The computing mode of "centralized service, distributed computing, and flexible expansion" with distributed computing and dynamic load is adopted to alleviate computing pressure, make the user experience agile, improve the robustness and reliability of database computing power, and enhance the overall service user experience.

### 3.2 Service Layer

At the shared service layer, which serves as the platform service layer, the latest business capabilities and data services are utilized in the design of the middle platform. The traditional server background and application foreground architecture are utilized to move the background business forward and the foreground business backward, thus creating a three-tier service architecture of background, middle, and foreground. This enables the integrated management and sharing of background capability, foreground capability, background data, and foreground data, effectively addressing expansion problems



encountered in traditional architecture, such as business model segmentation, service capability segmentation, business data islands, and business migration difficulties (Figure 3).

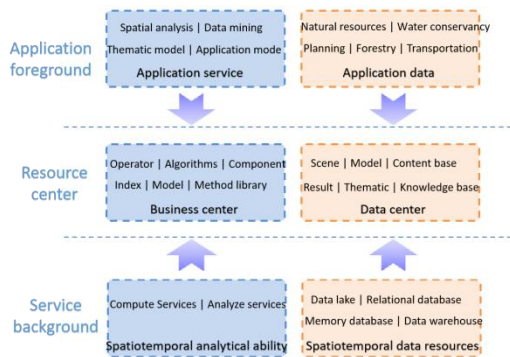


Figure 3. Shared service center in the platform mode.

To enhance the professional support and expansion capabilities of natural resource professional services and industry professional applications, and to ensure a set of cloud platform functions that support diversified application services, the cloud-based businesses adopt a forward sharing approach for computing services, analysis services, and other background capabilities, as well as a backward sharing approach for thematic models, application models, and other business capabilities. This approach maximizes resource sharing, resource reuse, and service co-construction, while promoting and optimizing thematic service capability, and reducing the cost of system maintenance and application migration and development. Moreover, it enhances the iterative ability and reuse vitality of development resources, ultimately improving the application efficiency of the entire service platform.

An open and shared ecosystem for thematic applications is created by opening and sharing the application capabilities of the cluster through data center, forming the index base, model base, method base, and other business capabilities of open services. To manage the challenges of multi-type data related to geographical conditions, a full-link technical approach consisting of data lake, relational database, in-memory database, and data warehouse is adopted for data resource management. Distributed storage architecture supports the unified storage of data with different structures in the data lake, which addresses data integration and distributed computing issues. Structured data is stored in a distributed database to achieve standardized storage, online transactional and analytical service capabilities. Business storage and online analysis of application data are enabled through the use of a distributed data warehouse, facilitating thematic application of geographical conditions and multi-dimensional analysis mining. Centralized management of data resources allows for open sharing of diversified thematic application results, leading to the development of a geographical conditions application result base, special topic database, and knowledge base. The ultimate goal is to establish a self-iterative geography and national conditions application ecology.

### 3.2 Application Layer

The cloud services provided by the application layer, including portals, visualization services, task monitoring, real-time query, and high-performance analysis services, are investigated in this study. Thematic clusters for natural resources, land, water conservancy, planning, forestry, grassland, and other types are constructed through front-end business topics, supported by the self-resource layer and shared service layer. The aim is to

provide active, intelligent, thematic, comprehensive, and personalized services for government agencies, professional departments, scientific research institutions, and the public. Customized models, such as iterative reuse, are available to support stable application patterns and facilitate migration and expansion to new application scenarios or thematic areas. Furthermore, application patterns can be mapped and reconstructed to create a series of reusable service clusters, such as an application scene library, analysis pattern base, and thematic content base. This establishes an interconnected, open, and shared geographical application content ecology, allowing for horizontal promotion and vertical upgrade of application requirements in different industries and regions.

## 4. SYSTEM EXPERIMENT

This study investigated the construction of an open service analysis system for geographical conditions in a private cloud software and hardware environment based on the national geographic condition monitoring spatio-temporal database and employing large-scale cloud storage facilities, all-in-one database cloud platform, and distributed memory computing cloud facilities. The goal was to address the challenges of ultra-large-scale spatio-temporal data distribution scheduling and online computing core needs, and large-scale data computing experiments were conducted on the scale of 100 million to lay the foundation for the construction of a big data analysis platform for geographical conditions.

### 4.1 System Structure

A hybrid cloud collaborative model was utilized in the latest system architecture, which integrated storage cloud, computing cloud, database cloud, online thermal coupling, and cloud collaborative load. Key technologies were developed for the interface-level communication and collaborative application of cloud resources. A heterogeneous cloud resource pool, which consisted of elastic storage cloud, distributed memory cloud, and integrated database cloud, was successfully established. Furthermore, the technical framework for very large-scale distributed collaborative parallel computing and cloud collaborative load was implemented. An online analysis system of geographical conditions, based on the collaboration of database cloud and distributed computing cloud, was constructed. This system enabled the parallel transformation of typical spatial operators and model algorithms. The spatial computing was dynamically analyzed and distributed based on bottleneck indexes, such as variable dimension, operator complexity, and data magnitude. As a result, the computing-intensive capability of the database cloud and the data-intensive capability of the distributed computing cloud were fully leveraged. The system achieved large-scale collaborative parallel computing on the magnitude of hundreds of millions of spatial data points, with domestic leading geospatial big data analysis and computing capabilities. The technical level of geospatial data online analysis was significantly enhanced.

### 4.2 Cloud Resource Pool

A cloud resource pool specifically designed to meet the demands of geographical conditions was established by utilizing advanced software and hardware resources. This pool integrated high-performance storage, computing, and service equipment, ensuring superior performance, availability, and sustainability for big data analysis and calculation in the field. The topology

and composition of the cloud resource pool were depicted in Figure 4.

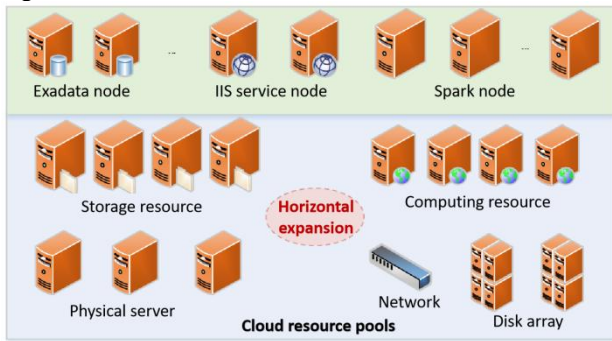


Figure 4. Topology diagram of the cloud resource pool.

### 4.3 Spatial Calculation Method

Focusing on spatial geographic analysis services, this study constructed three main spatial computing systems of vector computing, grid computing, and point-based computing based on a cloud platform infrastructure and distributed spatial computing environment. The platform enabled large-scale spatial analysis service capabilities and provided third-party development interfaces, as well as customized development capabilities for computing operators and models. The spatial computing architecture of the platform was presented in Table 1.

Computing system	Typical computing content
Vector calculation	The quantitative expression and scientific calculation of spatial attributes, including spatial location, spatial distribution, spatial form, spatial distance, and spatial relationships, are fundamental to supporting complex geographic analysis, such as data validation, extraction, conversion, statistical analysis, modeling, and mining.
Raster computation	Raster data is a crucial representation format for remote sensing images of natural resources, digital elevation models, derived data, and thematic raster data. It provides essential raster computing capabilities for spatial overlay analysis, resource assessment, model computation, and terrain analysis.
Grid computing	Spatial data is recursively encoded through geospatial and scale inclusion relations. With the support of a grid algorithm library and distributed computing engine, traditional spatial relationship and association analyses are transformed into associated queries and recursive statistics of attribute data, enabling accelerated computation of complex spatial calculations and association analyses.

Table 1. Cloud platform spatial computing system.

### 4.4 Open Service Model

Open service models have been developed in response to diverse analysis scenarios, including geographic analysis service and natural resources. These models are based on storage resource sharing, computing resource sharing, data

resource sharing, and model resource sharing. A geographic and national knowledge support base has been preliminarily constructed based on metadata, and an open analysis portal platform has been designed and built to enable open customization and tagged topic management. The open analysis service framework of cluster-user-data-operator-model-index-theme has been realized, effectively supporting the open computing and analysis of geographical data.

### 4.5 Computational Performance Verification

A computing capability test is conducted on the cloud platform to evaluate the spatial analysis capability of the big data analysis platform, with computing power being considered as a core performance metric. The test employs typical geographic data and spatial computing scenarios to assess the platform's ability to support large-scale data.



Figure 5. The national wide land cover datasets consist of 260 million vector polygon patches, with a total data amount of 300GB.

The entire land area of the country is covered by a crucial vector data set in geography, known as the land cover data, which comprises 260 million fine vector polygons. These polygons, representing complex polygons such as woodland, grassland, water surface, and road surface, consist of up to 100,000 single-element nodes. As a result, the calculation, analysis, and application of geographical information face bottlenecks due to the complexity of this datasets. To address this issue, the experimental data utilized in this study is the national land cover survey (Figure 5), and the experimental scene is the township area statistics. The data was analyzed and calculated through performance tests. The surface cover of each township spatial unit was determined by using the spatial unit data of township administrative division and summarizing the value of the third-class area of surface cover.

To simplify calculations, the authors converted the national land cover data into WKT format using a 0.2-degree grid and stored it in a distributed file system on the cloud platform. Table 2 presents statistical data on various aspects of the data. Custom-developed space cutting and area statistics operators were utilized on a Spark computing cluster for distributed computing experiments leveraging the existing cloud resource pool. The computing nodes used for the experiment are listed in Table 3.

Data type	Data format	Number of patches	Data volume/GB
Land cover (original)	FileGDB	260 million	/
Township Unit (original)	FileGDB	55,362	/
Land cover	WKT	260 million	1228

(text)				
Township Unit	WKT	55,362	2.78	
(text)				

**Table 2.** Land cover and township data statistics

Request test memory(GB)	Parallel	Nodes	Threads per node	Time
210	24	8	12	32min

**Table 3.** Results of land cover intersection of township unit

Excellent computing power is demonstrated by our analysis and calculations, with a total application memory of 256GB, which is highly advantageous for the Spark memory computing framework. A significant reduction in statistical time for national land cover analysis to an hourly level is achieved through our experiments, providing superior computational efficiency and time cost compared to the traditional spatial computing model of "data split, standalone calculation, result collection". Furthermore, an increase in the number of computing nodes can further improve the computational efficiency.

Our findings suggest that cloud-based big data analysis and calculation offers significant advantages and is well-suited for the demands of large-scale geographical analysis. The number of computing nodes in the cloud resource pool can be increased on a large scale as application demand and user concurrency increase, considering the horizontal expansion mechanism of distributed clusters. This approach enhances the computing performance and load capacity of the system, making it better suited to practical applications. Furthermore, for more complex computing-intensive scenarios such as buffer computing, overlay analysis, and exponential computing, the all-in-one cloud platform's rich interfaces and computing power can provide more flexible and efficient computing services.

## 5. CONCLUSIONS

This paper proposes a big data analysis service platform specifically designed for natural resource management in the contemporary era, which faces new challenges such as the increasing diversification, depth, and real-time requirements of data applications. The proposed platform is built on a new architecture that incorporates features such as comprehensive cloud services, on-demand access, on-demand response, on-demand services, and open data sharing. To realize this architecture, we introduce innovative technologies including heterogeneous cloud collaborative services, business capability, and virtualization of data resources. With the application of geographical conditions as a starting point, we emphasize the role of geographical services in supporting natural resource management and industry applications. This platform offers a broader vision for geographical services, promotes innovative thinking, and presents an opportunity for exploration and support of new work in the field of natural resources.

The analysis of big spatial-oriented data is characterized by its data-intensive and computing-intensive nature, which poses significant challenges in terms of complexity and computational requirements. To overcome these challenges, the authors developed a scalable solution leveraging extensive software and hardware resources. Specifically, they established an elastic cloud-based resource pool comprising storage, computing, and database services, which collectively enable online computational capabilities for large-scale geospatial data analysis.

The conventional methods for spatial data analysis and computation, namely vector computation, grid computing, and lattice computing, have been developed while exploring open service modes, such as resource sharing of storage, computation, data, and model. To assess the performance of the open analysis service framework of cluster-user-data-operator-model-index-theme, a performance test was conducted using 260 million national land cover vector data. The findings suggest that cloud-based big data analysis and computation of geographical conditions offer significant advantages, such as agility, scalability, sharing, and openness, for big data analysis of geographical conditions.

In the forthcoming phase, it is planned to enhance the expansion of the cloud resource pool, collaborative scheduling of cloud resources, upgrade of the business platform, refinement of data extraction, enhancement of the spatial computing interface, development of advanced spatial analysis models, deployment of thematic application models, and provision of multi-user concurrent support. The focus will be on enhancing the design and technology, in tandem with persistent advancement and construction of theme applications, such as ecological value assessment and resource ring environmental audit. The ultimate objective is to establish a robust and effective big data analysis service platform for geographical conditions.

## ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the National Geographic Conditions Monitoring Project for providing strong support for this study.

## REFERENCES

- Allen, D.W., & Coffey, J.M., 2019: *Focus on Geodatabases in ArcGIS Pro*. Esri Press.
- Chen, C., Lin, J., Wu, X., Wu, J., & Lian, H., 2013. Massive Geo-spatial Data Cloud Storage and Services Based on NoSQL Database Technique. *Journal of Geo-Information Science* 15(2), 166-174.
- China, N.A.O.o.t.P.s.R.o., 2018. *Zhonghua Renmin Gongheguo Shenjishu Shenji Jiegou Gonggao*. In: Beijing: National Audit Office of the People's Republic of China.
- Ferreira, D., & Vale, M., 2022. Geography in the big data age: an overview of the historical resonance of current debates. *Geographical Review* 112(2), 250-266.
- Guan, L., Ding, Y., Liu, H., Feng, X., Chen, P., & Ma, H., 2020. Construction and application of health examination evaluation system in new smart city. *Science of Surveying and Mapping* 45(3), 135-142.
- Jiao, S., & Wang, Y., 2020. The Aggregation Effect of "Inspector + Mapping". In: *China Natural Resources News* (pp. 7-8). Beijing
- Jin, F., Zhang, F., Du, Z., Liu, R., & Li, R., 2016. Spatial overlay analysis of land use vector data based on Spark. *Journal of Zhejiang University(Science Edition)* 43(1), 40-44.
- Jiping, L., Chun, D., Xiaochen, K., Shike, Q., Rong, Z., Bin, L., & Lijian, S., 2019. National Geographical Conditions Statistical

Analysis in the Era of Big Data. *Geomatics and Information Science of Wuhan University* 44(1), 68-76.

Li, D., Ding, L., & Shao, Z., 2016. Reflections on Issues in National Geographical Conditions Monitoring. *Geomatics and Information Science of Wuhan University* 41(2), 143-147.

Li, S., 2014. Research on Key technologies of data management and parallel processing under cloud GIS framework. In, *College of Earth and Planetary Sciences: University of Chinese Academy of Sciences*.

Liu, J., Dong, C., Kang, X., Qiu, S., Zhao, R., Li, B., & Sun, L., 2019. National Geographical Conditions Statistical Analysis in the Era of Big Data. *Geomatics and Information Science of Wuhan University* 44(1), 68-76, 83.

Liu, Y., Wang, P., & Huang, C., 2018. A Practical Exploration of the Multi-plan Integration Technical Method in Hainan Province. *NATURAL RESOURCE ECONOMICS OF CHINA* 5(1), 30-34.

Reddy, G.O., Raval, M.S., Adinarayana, J., & Chaudhary, S., 2022: *Data Science in Agriculture and Natural Resource Management*. Springer.

Saygili, O.Y., 2020: *The Introduction to Private Cloud using Oracle Exadata and Oracle Database*. CRC Press.

Thatcher, J., Shears, A., & Eckert, J., 2018: *Thinking Big data in geography: New regimes, new research*. University of Nebraska Press.

Yang, C., Goodchild, M., Huang, Q., Nebert, D., Raskin, R., Xu, Y., Bambacus, M., & Fay, D., 2011. Spatial cloud computing: how can the geospatial sciences use and help shape cloud computing? *International Journal of Digital Earth* 4(4), 305-329.

Yu, J., Chen, Y., & Gu, X., 2017. Spatial Analysis and Management Guideline in the General Urban Design of Shanghai. *Urban Planning Forum* 7(1), 101-107.