MINI REVIEW

Collecting large geodata in current situations

Chan Xi, Lee Fan, Fuyan Law

Xi C, Fan L, Law F. Collecting large geodata in current situations. J Environmental Geol. 2022; 6(4); 42-44

ABSTRACT

The manifestation of geographical phenomena is the result of the synthesis of several elements. Increased calculation and analysis techniques are demonstrated in collaborative multi-process calculation and coupling, ranging from single-point calculation to distributed coupling calculation, from independent variable mining to large-scale text mining, from spatial analysis to big data, and complex network mining characteristic analysis, as well as the transition from single geographic process calculation to multiple geographic process analysis.

The research on geographical applications supported by big geodata reflects its fundamental characteristics: continuously refined scales and scopes, gradually extended research objects, and the goal of comprehensive law discovery, for example, using over one million sample coverage and high-density sampling to support the inference of statistical characteristics of group behaviors, using hundreds of thousands of sampling data and multi-factor driving.

Key Words: *Geodata; Spatial And temporal distribution; Comprehensive data evaluation*

INTRODUCTION

The development of sensing technologies such as mobile location, Earth observation, and other sensor networks, as well as the continuous expansion of their application fields, has aided the rapid growth of big geodata [1-3]. In essence, big geodata is a large-scale coverage sample set for geographical phenomena that includes time, space, and attribute dimensions [2]. Large sampling data on surface factors and human behaviour observed across a broad spatial and temporal range by ubiquitous sensors and social sensing form an extension of big geodata, allowing the study of geographic objects, factors, and evolution from multiple spatiotemporal scales, richer dimensions, and diverse perspectives. However, big geodata faces significant challenges in data organization, management, modelling, expression, and analysis, as well as issues arising from its diverse sources, multiple observation scales, fluctuating sampling frequency, and accuracy variations. In this context, a thorough understanding and mastery of the characteristics of various types of big geodata is a necessary prerequisite for expanding its geoscientific applications.

The emergence of new sources and types of big Geodata continues to broaden the scope of geographic information science in four fundamental ways [3,4]: The integration of multiple spaces manifests the expansion of the spatiotemporal scope of observation and research from outdoor space to indoor space, from macro space to micro space, from physical space to cyberspace, and from natural space to social space. Extending data's spatiotemporal dimension from survey data to various types of perceptual big data, from multi-dimensional information aggregation manifests itself as a transition from discrete sampling data to continuous data, from data representing a region to data representing an individual object, from independent variable data to high-dimensional covariate data, and from simple relational data to complex network data. The broadening of the scope of geographical law discovery from multivariable correlations to pan-variable analysis connections ranging from multi-factor effects to an understanding of full-factor effects, multi-process interpretations, whole-process analysis of geographical events, and cognition of integrated geographical phenomena. The manifestation of geographical phenomena is the synthesis of multiple elements. (4) Increased calculation and analysis techniques ranging from single-point calculation to distributed coupling calculation, from independent variable mining to large-scale text mining from spatial analysis to big data, and complex network mining characteristic analysis, as well as the transition from single geographic process calculation to multiple geographic process analysis—is demonstrated in collaborative multi-process calculation and coupling. The research on geographical applications supported by big geodata reflects its

Editorial office, Journal of Environmental Geology, South Hampton, United Kingdom

Correspondence: Chan Xi, Editorial office, Journal of Environmental Geology, South Hampton, United Kingdom

Received: 14-July-2022, Manuscript No. PULJEG-22-5678; Editor assigned: 19- July-2022, PreQC No: PULJEG-22-5678 (PQ); Reviewed: 30- July-2022, QC No. PULJEG-22-5678 (Q); Revised: 03-August-2022, Manuscript No. PULJEG-22-5678 (R); Published: 11-August-2022, DOI:10.3037532/2591-7641.2022.6(4); 42-44

basic characteristics: the continuously refined scales and scopes, the gradually extended research objects, and the goal of comprehensive law discovery, for example, utilizing over one million sample coverage and high-density sampling to support the inference of statistical characteristics of group behaviors [5], using more than hundreds of thousands of sampling data and multi-factor driving variables to support global environmental parameter distribution mapping, adopting long time-series remote-sensing images with global coverage to support high precision dynamic mapping of surface elements and comprehensive correlation impact assessment, and using multi-dimensional observation data and multi-model data to jointly support the attribution interpretation of dynamics of large-scale elements [6]. The research mentioned above represents a revolutionary expansion of the scope of geographic information science under the support of big geodata; it also continuously advances the fourth paradigm of geographical research.

This paper summaries current research on big geodata research themes, analytical technologies, and methods in the field of geographic information science and discusses current problems and future directions in this field.

Themes of research are expanded. The growth of big geodata has accelerated the expansion of geographical research topics.

On the basis of an analysis of the spatial and temporal distribution and trend, as well as the factor correlation of traditional elements, a series of different research themes for geographical phenomena and processes have been developed. Typical themes from the multi-domain perspective include the following:

1) Individual behavior data is used in big data analysis. Human mobility data, credit card data, and mobile phone data are just a few examples of individual behaviour data. In this regard, specific themes include:
    (a) the exploration of overall temporal and spatial characteristics of phenomena using individual behaviour data, such as multi-scale patterns of population migration or human mobility, the uniqueness of group consumption behavior, the group spatiotemporal characteristics of ocean fishing behaviour [5,] and the resulting optimization strategy and personalized recommendation strategy.
    (b) The social perception of traditional statistical and survey indicators derived from individual behavior data, such as influenza prediction via network searches, regional and individual economic status assessment, landscape value assessment, and general election voting trend prediction.
    (c) Analysis of a certain phenomenon based on individual behavior data, such as the malaria network transmission pattern based on human mobility and fake news in geographic and virtual space.

2) Multi-source data analysis for big data. Combining data from multiple sources to analyze geoscience phenomena and processes has long been a fundamental method of geographical research. Important themes in this context include:
    (a) the discovery of patterns in historical events and phenomena through multi-source data, such as the reshaping of Asia's Silk Road, the geographical shifts of cultural centers, and the link between climate change and human conflict [7].
    (b) Inference and mapping of traditional socioeconomic indicators at fine spatial resolutions, such as education, health, and emotion, using big geodata correlation.
    (c) Analysis of emergency public events combined with multi-source data, such as emergency measures assessment and spatiotemporal prediction of COVID-19 [8].
    (d) Correlation evaluation of the ecological and environmental consequences of human activities utilizing multi-source data, such as roads with ecological reserves and free-flowing rivers.

Big data analysis with high spatiotemporal resolution. This type of data allows for a more in-depth examination of geographical objects in both temporal and spatial dimensions. The following are examples of typical themes: Global scale mapping of single-surface parameters with high spatial resolution using literature data and correlated big data from spatiotemporal observation, such as tree distribution and arsenic content in the groundwater system.

Global spatiotemporal distribution mapping and surface element evolution pattern discovery using remote sensing data with long time series and high spatial resolution, including land use, vegetation dynamics, and river factor evolution.

According to the research objectives, the extended research themes mentioned above cover five types of typical geographic information analysis and application issues.

1) Spatiotemporal mapping: Spatial and temporal mapping of geographical factors is a fundamental and critical issue. It is essential not only for understanding the evolution of geographical elements, as well as the temporal and spatial pattern of human activities and sustainable development, but also for decision-making. Big geodata development will drive spatiotemporal mapping toward more geographic elements and processes, larger spatial and temporal extents, and finer scales. It is gradually becoming an important foundation for supporting other interdisciplinary geographic information research [1]. The pan information map proposal exemplifies cartography's self-renovation in this context in terms of theory, technology, products, and services.
2) Pattern discovery: one of the primary goals of big geodata mining is to uncover the patterns of geographical phenomena hidden behind massive amounts of sampled and observed data [5]. Individual behaviors or events occur

at random and only on rare occasions in small sample data. However, using long time series and massive sample data to discover the collective characteristics and patterns of geographic processes or events has become a hot direction. Numerous studies show that big data plays an important supporting role in this regard, with significant implications for a wide range of sectors, including the environment, ecology, economy, culture, public health, and management planning.

3) Comprehensive evaluation: the quantitative analysis and inference of the effects of many factors on the spatiotemporal evolution of a specific geographic element, which has traditionally been the primary research material in the field of geographic information. In terms of element content, impact dimension, correlation mode, spatiotemporal scale, and extent, the emergence of big geodata has significantly expanded the scope of comprehensive evaluation research. Thus, with the help of multidimensional observational data and multi-model data combined, attribution interpretation of element dynamics becomes possible, which is also the primary.
4) Spatiotemporal prediction: Spatiotemporal trends and change levels can be speculated, reconstructed, and predicted based on the specific pattern presented by the target object through sampling and observation by mining and analyzing the pattern's evolution and driving mechanism. This will foster a thorough understanding of complex geographic problems and processes, as well as an understanding of future development trends. The scope and range of traditional spatiotemporal prediction problems, impact comprehensiveness, and prediction accuracy and extent have all changed significantly as a result of big geodata.
5) Model development: model development is the analysis and simulation of the spatiotemporal evolution process of geographical phenomena and geographical elements, and it serves as an important auxiliary decision support tool in the field of geographic information. The emergence of spatiotemporal data allows for the incorporation of geo-knowledge into classical machine learning and big data-driven deep learning, broadens the scope of the experience-oriented model, extracts spatiotemporal patterns from big geodata, and builds unique geo-analysis models.

CONCLUSION

With the advent of big geodata, the scope, dimensions, themes, and perspectives of geographic information science research have expanded. This is critical for encouraging the investigation of complex geographical phenomena and problems, as well as the discovery of their regularities. It will also be the frontier and core direction of the discipline's future development. As a result, the ways and contents of the intersection of geographic information science and other fields will be expanded and enriched. In the acquisition and processing of spatiotemporal data, it is necessary to develop comprehensive integration and multidimensional fusion methods of multiple spatiotemporal scales, observation methods, and measurement perspectives and to explore the quality and precision control of big geodata to further support the exploration of complex spatiotemporal evolution models and the discovery of basic geographical patterns. In terms of analysis and modeling methods, it is necessary to extend the granularities and scales of analysis objects, develop joint computing and mining models considering multi-factor and multi geographic processes, improve spatiotemporal inference models to support the transformation of spatiotemporal scales and granularities, build comprehensive attribution models to support quantitatively driven interpretation of high-dimensional variables, improve the geoscientific mechanism models to describe the overall evolution mechanisms and characteristics of individual spatiotemporal changes, and develop coupling models of deep learning and geoscientific mechanisms to improve the applicability and interpretability of results in all aspects. Furthermore, a number of key issues affecting big data analysis and modelling, such as data quality, high-dimensional dynamics, statistical confounding, and multiple coupling, must be addressed in order to develop a better methodology system for big geodata research and to support the development and application innovation of geographic information science.

REFERENCES

1. Guo H, Chen F, Sun Z, et al. Big Earth Data: A practice of sustainability science to achieve the Sustainable Development Goals. Science Bulletin. 2021;66(11):1050-3.
2. Pei T, Liu YX, Guo SH, et al. Principle of big geodata mining. Acta Geographica Sinica. 2019;74(3):586-98.
3. Zhou C, Wang H, Wang C, et al. Geoscience knowledge graph in the big data era. Science China Earth Sciences. 2021;64(7):1105-14.
4. Li X, Zheng D, Feng M, et al . Information geography: The information revolution reshapes geography. Science China Earth Sciences. 2022;65(2):379-82.
5. Song C, Qu Z, Blumm N, et al. Limits of predictability in human mobility. Science. 2010;327(5968):1018-21.
6. Pekel JF, Cottam A, Gorelick N, et al. High-resolution mapping of global surface water and its long-term changes. Nature. 2016;540(7633):418-22.
7. Frachetti MD, Smith CE, Traub CM, et al . Nomadic ecology shaped the highland geography of Asia's Silk Roads. Nature. 2017;543(7644):193-8.
8. Chinazzi M, Davis JT, Ajelli M, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science. 2020 ;368(6489):395-400.