Big Geospatial Data for Environmental and Agricultural Applications

Adhikaram Karuna, Angelle Tonneau, Kinnarika Karimanne

1 Introduction

The recent generation of geoprocessing systems are generating vast continuous streams of massive geospatial data (GGD) systems. Initially, high-resolution geospatial images are generated at a very high rate and are stored in various datasets per week. In addition, 3D data models and simulation are available from national and international organizations in the US, China, and EU. Spatial industry and standards are increasing adoption of geospatial data to optimize the management of natural resources. With the use of geospatial data, new datasets and applications are created. Current data management systems for data-intensive scientific and agricultural applications are becoming a critical research area. We developed a comprehensive tool for processing large geospatial datasets. This study highlights various aspects of geospatial data in the domain of geospatial processing, geospatial processing, geospatial data management, and geospatial data analysis.

2 Big Geospatial Data

Big data (BD) presents various types of challenges in terms of storage, processing, and analysis. Big data also offers many opportunities in terms of storage, processing, and analysis. Big data includes various types of data, such as sensors, social media, and transactional databases. Big data also includes various types of data, such as sensors, social media, and transactional databases. Big data includes various types of data, such as sensors, social media, and transactional databases. Big data includes various types of data, such as sensors, social media, and transactional databases. Big data includes various types of data, such as sensors, social media, and transactional databases. Big data includes various types of data, such as sensors, social media, and transactional databases. Big data includes various types of data, such as sensors, social media, and transactional databases. Big data includes various types of data, such as sensors, social media, and transactional databases. Big data includes various types of data, such as sensors, social media, and transactional databases. Big data includes various types of data, such as sensors, social media, and transactional databases. Big data includes various types of data, such as sensors, social media, and transactional databases. Big data includes various types of data, such as sensors, social media, and transactional databases.
Big Geospatial Data for Environmental and Agricultural Applications

2.2 Vector Data

Many geospatial data are available in the form of vector data representations. Vector data structures are constructed from simple geometric primitives, such as points, lines, and areas. These primitives are used to represent geographic features of the real world, such as geographic objects (e.g., roads, rivers, buildings). Vector data can be used to represent geographic features of any scale and are often used in applications that require precise location information.

Vector data is stored in a format that allows for efficient storage and retrieval. Each geographic feature is represented as a set of points, lines, or areas. This information is organized in a hierarchical structure, where each level represents a different level of detail. Vector data can be queried and analyzed using various techniques, such as spatial queries and network analysis.

2.3 Raster Data

Rasters are a type of database that is used to store geographic data in a grid-like format. Each cell in a raster represents a small area of the Earth's surface and contains a value that represents a particular attribute, such as elevation, temperature, or land use.

Rasters can be used to represent a wide range of phenomena, such as vegetation cover, soil types, and water bodies. Rasters can be analyzed using various techniques, such as image processing and spatial analysis.

2.4 Open Geospatial Standards

The open geospatial community has recognized the importance of standards for interoperability and data exchange. A number of open geospatial standards have been developed to enable the exchange of geospatial data between different systems and applications.

2.4.1 Open Geospatial Consortium (OGC)

OGC (Open Geospatial Consortium) is an international organization that develops open geospatial standards for the exchange of geospatial data. OGC is a membership-based organization that is supported by a wide range of industry stakeholders.

OGC standards are developed through a process of consensus, where stakeholders from various industries and geographic areas contribute to the development of the standards. OGC standards are intended to be open, non-proprietary, and interoperable.

OGC standards include:

- WGS 84: A standard for defining the geographic coordinates of points on the Earth's surface.

OGC standards are used by a wide range of industries, including government, academia, and the private sector, to enable the exchange of geospatial data and to facilitate the development of geospatial applications.

2.4.2 WPS

WPS (Web Processing Service) is an OGC standard for the execution of geospatial processing services over the Internet. WPS is designed to enable the execution of geospatial processing services over the Internet, allowing users to access and use these services from any location.

WPS supports the execution of processing services that operate on geographic data, such as the transformation of geographic data from one format to another or the application of spatial analysis algorithms to geographic data. WPS services are defined using a simple metadata format that describes the service and the data that it operates on.

WPS services can be accessed using a standard HTTP interface, allowing users to submit requests for processing services and receive results. WPS services can be used in a wide variety of applications, including geographic information systems and web mapping services.
3 Big Data and Big Data Technologies

With the increasing growth of data and the complexity of processing it, Big Data technologies have become essential in various fields. These technologies enable organizations to extract insights from massive amounts of data, which can drive decision-making processes and innovation.

3.1 Overview of Big Data Technologies

Big Data technologies include data storage, data processing, and data analysis. Data storage technologies provide the infrastructure to store large volumes of data efficiently. Data processing technologies enable the extraction of useful information from data, while data analysis technologies help in interpreting the insights derived from the processed data.

3.2 Technologies for Big Data Storage

Big data storage technologies support high-speed in-memory systems, scale-out NAS file systems, and distributed file systems. These technologies provide scalable storage solutions that can handle the growing volume of data.

3.3 Technologies for Big Data Processing

Big data processing technologies include batch and stream processing, real-time data processing, and data integration. These technologies enable organizations to process data in real-time or near-real-time, which is crucial for applications like financial transactions and social media monitoring.

3.4 Technologies for Big Data Analysis

Big data analysis technologies include data mining, machine learning, and predictive analytics. These technologies help in discovering patterns, relationships, and insights from data, which can be used to improve decision-making processes.

3.5 Big Data Technologies and Applications

Big data technologies have various applications across industries, such as healthcare, finance, retail, and social media. These technologies enable organizations to make data-driven decisions, improve customer experiences, and optimize operations.

In conclusion, Big Data technologies are crucial in today’s data-driven world. They provide the means to store, process, and analyze large volumes of data, which can help organizations gain a competitive edge by making data-driven decisions.
networking include Radiaxion, Movado (DBS-DAC), Ruckus, Oracle GlassRoom and Nextera fiber optic networks.

This is in contrast to more traditional and distributed data management and storage network designs. A NIDSL database is a collection of data that is modeled in more ways than the single relation used in relational database. NIDSL systems are more complex to implement and may not allow the same degree of freedom in storage as traditional database structures.

This is in contrast to traditional data management and distributed data management, where data is modeled in more ways than the single relation used in relational database. NIDSL systems are more complex to implement and may not allow the same degree of freedom in storage as traditional database structures.

This is in contrast to traditional data management and distributed data management, where data is modeled in more ways than the single relation used in relational database. NIDSL systems are more complex to implement and may not allow the same degree of freedom in storage as traditional database structures.

This is in contrast to traditional data management and distributed data management, where data is modeled in more ways than the single relation used in relational database. NIDSL systems are more complex to implement and may not allow the same degree of freedom in storage as traditional database structures.

This is in contrast to traditional data management and distributed data management, where data is modeled in more ways than the single relation used in relational database. NIDSL systems are more complex to implement and may not allow the same degree of freedom in storage as traditional database structures.

This is in contrast to traditional data management and distributed data management, where data is modeled in more ways than the single relation used in relational database. NIDSL systems are more complex to implement and may not allow the same degree of freedom in storage as traditional database structures.

This is in contrast to traditional data management and distributed data management, where data is modeled in more ways than the single relation used in relational database. NIDSL systems are more complex to implement and may not allow the same degree of freedom in storage as traditional database structures.

This is in contrast to traditional data management and distributed data management, where data is modeled in more ways than the single relation used in relational database. NIDSL systems are more complex to implement and may not allow the same degree of freedom in storage as traditional database structures.

This is in contrast to traditional data management and distributed data management, where data is modeled in more ways than the single relation used in relational database. NIDSL systems are more complex to implement and may not allow the same degree of freedom in storage as traditional database structures.
3.5 GeoDB

The National Geospatial-Intelligence Agency (NGA) [14] is in collaboration with GeoDB Solutions, Inc., recently released an open source version of its software for data sharing and processing purposes. This software, called GeoDB, is designed to provide relational geospatial data, real-time geospatial data, and geospatial data analysis tools.

GeoDB provides an open-source platform for managing and processing geospatial data. It supports various data formats including vector, raster, and 3D data. GeoDB also offers a web-based interface for data management and analysis.

GeoDB is designed to be scalable and flexible, allowing for easy integration with other systems and applications. It supports a wide range of spatial analysis tools, including spatial queries, spatial joins, and spatial statistics.

GeoDB is open-source software, which means that anyone can use, modify, and redistribute it. This makes it an attractive option for organizations that want to have more control over their geospatial data management and analysis processes.

GeoDB is used by various organizations, including government agencies and private companies, for managing and analyzing geospatial data. It is used in various fields such as urban planning, transportation, and disaster management.

4 Geospatial Environmental and Agricultural Services

Following the discussion about Big Geospatial Data and Big Geospatial Data Frameworks in the previous section, we will try to present the necessary information about the research goals and the direction of the developed technologies for geospatial data.

This is important because geospatial data is used in various fields such as environmental monitoring, agricultural management, and urban planning. The data collected from these fields is crucial for making informed decisions and managing natural resources effectively.

In this section, we will discuss the current status of geospatial data and the technologies used to manage it. We will also present some of the emerging technologies that are being developed to improve the efficiency and effectiveness of geospatial data management.

This includes technologies such as cloud computing, big data analytics, and artificial intelligence. These technologies are being used to develop new approaches for managing and analyzing geospatial data, which can help improve decision-making processes in various fields.
4.2 Global Forest Watch

Another important initiative is the Global Forest Watch (GFW) [30], which is a collaborative global effort led by the World Resources Institute to make the quantitative and spatial forest data that supports decision making. GFW is a forest monitoring service that uses a combination of satellite data and ground-truthing to provide near-real-time forest cover and change data for the world. GFW tracks deforestation and reforestation, provides warnings when deforestation is detected, and tracks restoration efforts. GFW is available to anyone and partners with governments, NGOs, and businesses to ensure that the data is accurate and reliable.

4.3 Remaneralysti

Remaneralysti is a software tool developed by the University of Maryland to analyze and visualize satellite data. It is used to detect changes in forest cover and to monitor the effectiveness of conservation efforts. Remaneralysti is a free, open-source software tool that can be used by anyone with access to satellite data. It includes a variety of tools for data visualization, data analysis, and data modeling. Remaneralysti is available on the University of Maryland’s website.

Fig. 4 Remaneralysti is a software tool developed by the University of Maryland to analyze and visualize satellite data.

Fig. 5 The software is available on the University of Maryland’s website and can be used by anyone with access to satellite data.

Fig. 6 Remaneralysti uses a variety of tools for data visualization, data analysis, and data modeling.
By zero mean changes in the accurate estimation of regional vegetation cover. The depth of the data and the spatial resolution of the data. Many standard chlorophyll estimation algorithms are typically optimized by phytoplankton and their biodiversity products. In the water, it is essential to maintain the chlorophyll and other dissolved organic matter (DOM) and the surrounding environment. Based on the new framework for the chlorophyll algorithm, a general purpose algorithm or method was developed. More specifically, the calculation of chlorophyll or phytoplankton using DOM is the first component of the RemoteChlor. Based on the calculation of the DOM, the detection of water bodies can be performed upon request. In particular, the query performs water detection based on the calculation of DOM over the selected domain under a threshold value. The result is delivered in a binary image format. Moreover, the detection of the phytoplankton biomass index is calculated using a threshold of the abundance and biomass of marine phytoplankton. Chlorophyll data are used to estimate the spatial scales for monitoring the water quality and microstructure of coastal water bodies. Both the fine and coarse spatial scales can be used to estimate the aquatic vegetation cover. For example, the detection of phytoplankton using DOM is performed using the chlorophyll index.

Biological results from 4 different receptors are demonstrated for the RemoteChlor estimation method (Fig. 1). In particular, Figure 2 compares the spatial distribution of phytoplankton in coastal waters. The two charts show that the distribution of phytoplankton is better estimated using a finer pixel size. The charts also highlight the importance of using a finer pixel size to accurately estimate the distribution of phytoplankton in coastal waters.

Environmental applications of the RemoteChlor estimation method are demonstrated for the regional scale. In particular, Figure 3 demonstrates the distribution of phytoplankton in coastal waters using the RemoteChlor estimation method. The charts show that the distribution of phytoplankton is better estimated using a finer pixel size. The charts also highlight the importance of using a finer pixel size to accurately estimate the distribution of phytoplankton in coastal waters.

In summary, the RemoteChlor estimation method provides a powerful tool for estimating the distribution of phytoplankton in coastal waters. The method is able to accurately estimate the distribution of phytoplankton using both coarse and fine spatial scales. The results demonstrate the potential of the RemoteChlor estimation method for environmental applications.
Big Data for Environmental and Agricultural Applications

In the era of Big Data, the volume and complexity of data generated from various environmental and agricultural sources have significantly grown. To effectively manage and extract valuable insights from this data, advanced data processing and analytics techniques are crucial. This paper presents an overview of the data processing methodologies and their applications in the context of environmental and agricultural sciences.

Key aspects covered in the paper include:

- **Data Processing:** Techniques for handling large volumes of data, including data cleaning, preprocessing, and feature extraction.
- **Data Analysis:** Methods for extracting meaningful insights from the processed data, such as pattern recognition, predictive modeling, and statistical analysis.
- **Applications:** Case studies illustrating the impact of Big Data on environmental monitoring, crop yield prediction, and sustainable agriculture.

The paper concludes with a discussion on the future prospects of Big Data in environmental and agricultural sciences, highlighting potential areas for further research and development.

**References**


---

**Future Perspectives**

As the volume of data continues to grow, the importance of data processing and analytics in environmental and agricultural sciences will only increase. Future research should focus on developing more efficient and effective data processing algorithms, as well as on the integration of advanced machine learning techniques to enhance predictive modeling capabilities.

Moreover, there is a need to develop user-friendly tools and platforms for data visualization and analysis, making complex data insights accessible to non-experts. Collaboration between academia and industry is essential to translate research findings into practical solutions that can be implemented in real-world scenarios.

In conclusion, the future of Big Data in environmental and agricultural sciences holds tremendous potential for advancing our understanding of natural systems and enabling more sustainable practices.

---

**Author Bio:**

[Name], [Affiliation], [Bio information]

[Date of Publication]