Spatial Cloud Data Management Framework for Emergency Management

Tesfamariam Damte
HiLCoE Software Engineering Programee, Ethiopia
tesfadamte@gmail.com

Fekade Getahun
HiLCoE, Ethiopia Department of Computer Science**,** Addis Abab**a** University, Ethiopia
fekadegetahun@gmail.com

Abstract

Effective management of spatial data is critical for emergency management, and the availability of spatial data through cloud computing demands an efficient framework. This research aims to develop a spatial cloud data management framework for emergency management, capable of handling and analyzing large volumes of spatial data. Objectives include exploring spatial data types, designing the framework with key features, and evaluating its performance during emergencies.

Following a design science approach, the framework comprises four layers: infrastructure, data, application, and client. It leverages cloud storage, PostgreSQL databases, open-source spatial data sources, Apache Spark, PostGIS, and forecasting models. Supporting thick and thin cloud computing clients ensures accessibility for authorized users.

The significance of this study is evident in its transformative impact on emergency management through the integration of Spatial Big data and advanced analytics. By enhancing preparedness, response, and recovery, it contributes to minimizing the impact of disasters and protecting lives. The evaluation findings underscore its practical effectiveness, rendering it a valuable asset for both practitioners and stakeholders in the field.

In conclusion, the spatial cloud data management framework offers a viable solution for efficient spatial data management during emergencies. Its multi-layered architecture and support for cloud clients improve emergency management operations, with practical implications for the field and future advancements.

Keywords: Cloud storage, PostgreSQL, PostGIS, Apache Spark,

1. Introduction

Most of the things in this world are in dramatic change. Emergency events happen everywhere in every corner of the world. Emergency events imply all events that endanger normal functioning of services and companies, endanger lives or resources (living environment) as well as events that are threatening stability of state [1]. To handle such problems emergency management is required.

Emergency events are spatial in nature and geo spatial technology better suited to address them. However, in emergency management system Geo-spatial science faces great challenge in terms of data intensity, computing intensity, concurrency intensity and spatiotemporal intensity [2].This challenge cannot be address by a traditional desktop GIS and the use of new technologies is required.

The emerging cloud computing paradigm brings remarkable solution in challenges of spatial technology with elastic and on-demand access to massively pooled, insatiable, and affordable computing resources [2].

This research tries to design the architectural framework of spatial cloud data management in emergency management. Also address challenges of big spatial data.

2. Related Literature

In the study by X. Ge and H. Wang [3], the potential benefits of cloud-based services for managing extensive geographic data in emergency management are explored. Their approach involves a methodology to handle and analyze vast spatial data volumes during disaster management using cloud computing. Within this context, Aly et al [1] propose a GIS-based cloud computing model for emergency systems, envisioning its potential to enhance the efficiency of disaster management. This model suggests merging Geographic Information Systems (GIS) with cloud computing to create an effective and adaptable framework. Meanwhile, W.W. Song et al [4] propose a spatio-temporal cloud platform to support GIS applications, addressing challenges related to managing large spatio-temporal data and providing computational resources. This platform integrates distributed processing and cloud computing technologies, presenting a scalable and flexible solution. In the context of spatial data analytics, D. Jiang et al [5] emphasize the advantages of employing spatial data analytics for emergency response. They highlight interdisciplinary cooperation, real-time data processing, and visualization using advanced tools as key practices for implementing spatial big data analytics. Lastly, S. Gao et al [6] discuss the use of cloud computing for spatial data management, showcasing cloud-based GIS platforms, spatial data analytics tools, and visualization tools as existing solutions. Their paper stresses the significance of creating scalable, secure, and well-governed cloud-based systems while underscoring research gaps in the evaluation of cloud-based spatial data management solutions.

Overall, these studies collectively shed light on the potential of cloud-based services, GIS-cloud integration, spatio-temporal platforms, and spatial data analytics in enhancing emergency management. They underscore the need for scalable, secure, and adaptable systems that effectively handle diverse data sources and promote real-time decision-making during emergencies. However, gaps persist in efficient data compression, scalable algorithms, interdisciplinary cooperation, and comprehensive evaluation, leaving room for further research and innovation.

After reviewing related literature spatial cloud data management framework for emergency management typically should consists of components, including data acquisition, data storage, data processing, and data visualization.

Different studies proposed different solution for forecasting drought. Demisse, G. B et al [7] introduces a novel spatial drought index (SDI) utilizing satellite-derived data, emphasizing its significance in drought- prone regions' decision-making. The proposed SDI combines the Standardized Precipitation Index (SPI) and the Normalized Difference Vegetation Index (NDVI) through a weighted sum, enabling a comprehensive assessment of drought's spatial variability. Similarly, the study Demisse, G. B et al [8] underscores the importance of sub-national drought monitoring for improved management strategies.

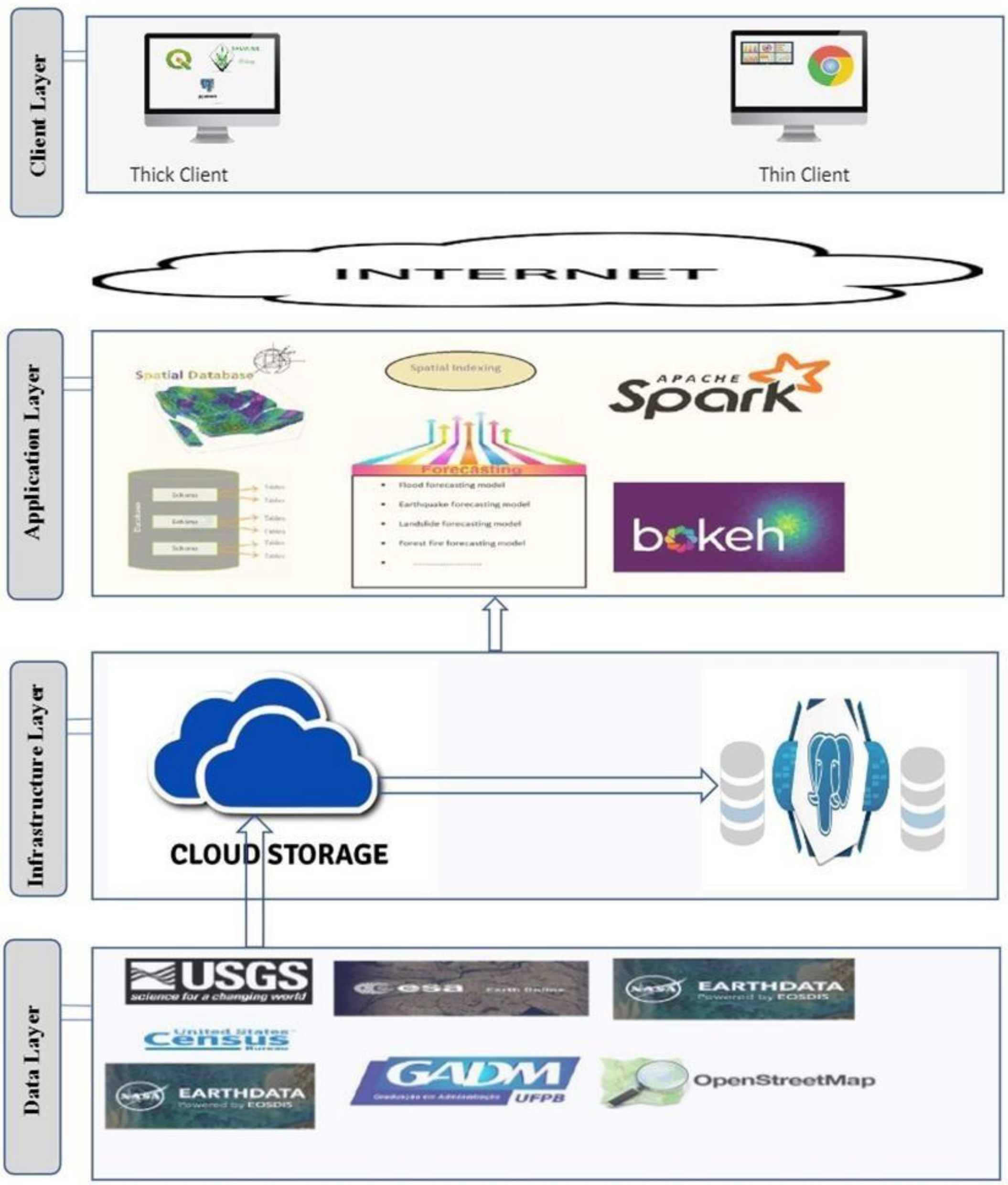
In the context of vegetation health and drought conditions, Huete et al [9] demonstrates NDVI's utility in Australia. Employing NDVI to gauge vegetation cover through satellite imagery, it contributes to understanding the impact of drought on vegetation. Additionally, Nigussie, A et al [10] reveals the enhanced accuracy of combining SPI, NDVI, and rainfall anomaly data for drought risk assessment. Collectively, the reviewed literature highlights the significance of integrating SPI, NDVI, and rainfall anomaly for a holistic understanding of drought conditions.

In conclusion, the literature highlights the significance of cloud-based spatial data management in emergency management, offering best practices and identifying research gaps. The use of SPI and NDVI for drought forecasting models contributes to more accurate and comprehensive assessments of drought conditions.

3. FRAMEWORK OF SPATIAL CLOUD DATA MANAGEMENT FOR EMERGENCY MANAGEMENT

The proposed spatial cloud data management framework for emergency management consists of four layers: Data layer, Infrastructure layer, Application layer and Client layer.

Figure 1: A framework for spatial cloud data management for Emergency management.



3.1 Data Layer

The data layer serves as the foundation of the framework and incorporates various open-source spatial data providers, including USGS Earth Explorer, NASA Earth data, ESA Earth Online, USGS National Map Viewer, OpenStreetMap, GADM, US Census Bureau, CDO (climate data online), and Natural Earth. These sources offer access to a diverse range of satellite imagery, GIS data, demographic information, climatedata, and more. To effectively integrate this data into a PostgreSQL database within the spatial cloud data management framework, a systematic approach is undertaken. The process involves identifying data sources, selecting a suitable cloud store service, creating storage containers, downloading data, transferring it to the cloud, and then configuring the PostgreSQL database to accept and organize the incoming data.

3.2 Infrastructure Layer

The infrastructure layer comprises two essential components: cloud storage and PostgreSQL database. Cloud storage options, including object storage, file storage, and block storage, offer scalability and flexibility for managing spatial data. Object storage is suitable for handling large volumes of unstructured data, while file storage is utilized for smaller datasets requiring file system access. Block storage provides persistent block-level storage volumes for spatial analysis on virtual machines. PostgreSQL, combined with the PostGIS spatial extension, emerges as a robust choice for spatial data management. It supports storage, indexing, and querying of vector and raster data, allowing for effective data organization and retrieval. The hybrid cloud deployment model is considered, incorporating both private and public cloud infrastructures for enhanced scalability and security.

3.3 Application Layer

The application layer encompasses key components, including spatial database, schema and table structures, spatial indexing, forecasting models, Apache Spark, and a web server. PostgreSQL with PostGIS supports data storage, retrieval, and manipulation, while schema and tables are defined to accommodate forecasting models. Spatial indexing enhances data querying efficiency, and Apache Spark facilitates real- time data streaming between sources and the PostgreSQL database. A drought forecasting model is demonstrated as a detailed example, incorporating NDVI calculations, rainfall analysis, SPI computation, and weighted averaging to assess drought risk. Finally, a web server, specifically Bokeh Server, is employed to create interactive web-based data visualizations and dashboards for users to analyze and comprehend the spatial data effectively.

3.4 Client Layer

The client layer encompasses both thin and thick clients. Thin clients, represented by web-based dashboards, offer lightweight access to the framework, enabling users to interact with the data easily. Thick clients, exemplified by QGIS software, provide more extensive capabilities for data analysis, manipulation, and visualization. QGIS is connected to the PostgreSQL database in the cloud, allowing users to create, edit, and analyze geospatial data in a secure and customized environment. This layered approach ensures accessibility, scalability, and adaptability for emergency management scenarios, empowering users to make informed decisions based on real-time spatial data.

The proposed spatial cloud data management framework integrates diverse data sources, cloud-based infrastructure, advanced analytics, and user-friendly interfaces to enhance emergency management processes. By leveraging cutting-edge technologies and methodologies, this framework offers a comprehensive solution for spatial data management, analysis, and visualization in time-sensitive situations.

4. Prototype

Emergency management teams benefit from cloud computing's ability to provide quick and secure access to data, aiding in rapid crisis response. The hybrid cloud deployment model is utilized in this prototype, allowing spatial-based emergency management analysis through QGIS software. Direct connection to the cloud-based PostgreSQL database, facilitated by PostGIS, enables data analysis. Expertise in basic SQL queries and PostGIS functions is required for analysis on the private cloud. Flood forecasting involves processing Landsat ETM+ data, rainfall data, and SPI data from USGS and CDO, respectively. Spark streaming transports this data into the PostgreSQL database. Custom functions in PostgreSQL, such as NDVI and rainfall anomaly, transform Landsat ETM+ and rainfall data, enabling creation of NDVI and rainfall anomaly data. A customized drought risk function utilizes these data layers to generate a drought risk map, aiding in decision-making.

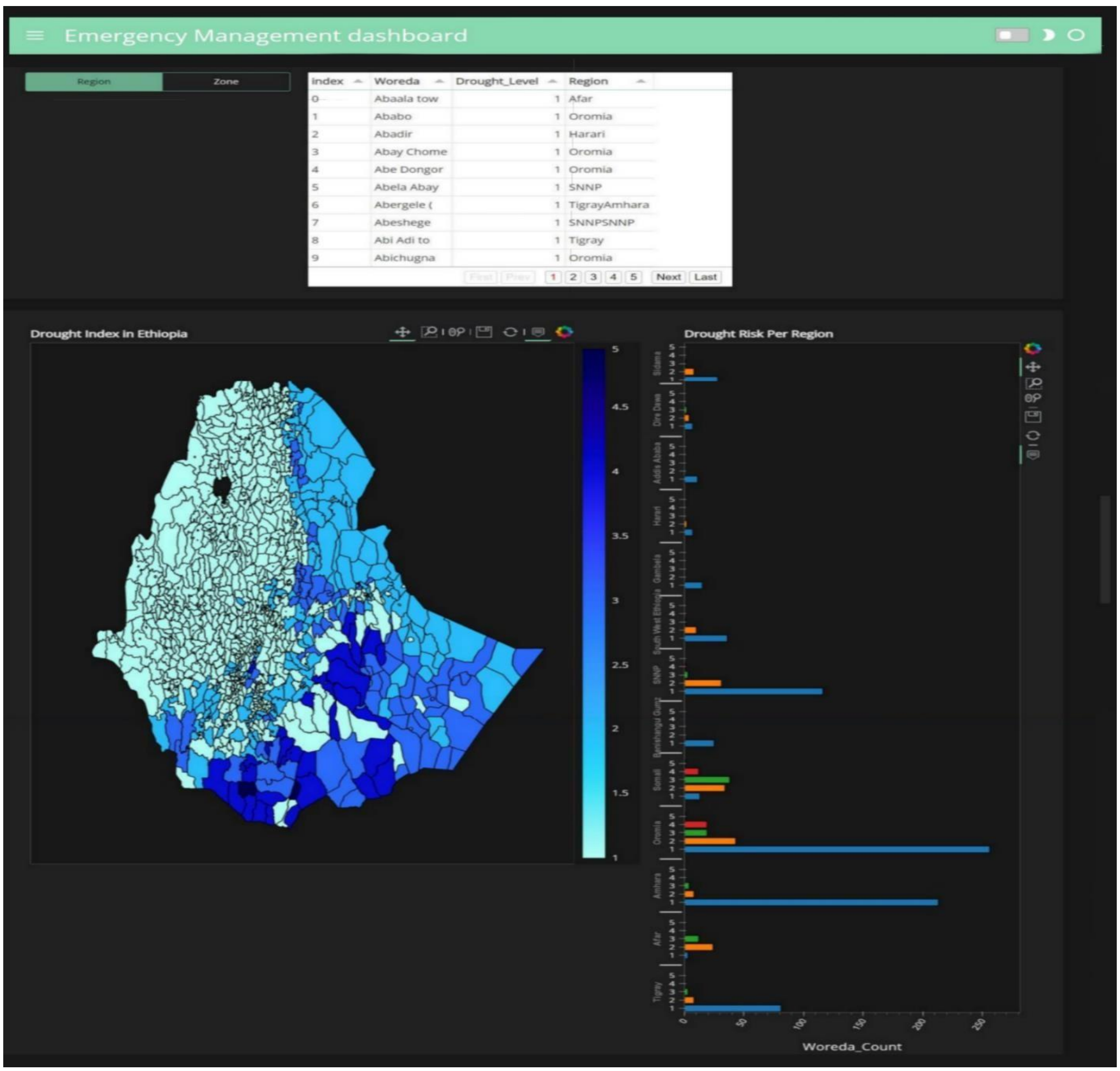
The custom function facilitates the calculation of drought risk based on the integration of various raster datasets. The function calculate_drought_risk is designed to process three essential input raster's: the Standard Precipitation Index (SPI) raster, rainfall anomaly raster, and Normalized Difference Vegetation Index (NDVI) raster. These input raster's are weighted according to their significance in drought risk assessment: SPI with a weight of 40%, rainfall anomaly with 30%, and NDVI with 30%.

The algorithm dynamically creates or replaces a table named drought_map in the PostgreSQL database to store the resulting drought risk map. It begins by checking if the drought_map table already exists; if so, it clears the table, and if not, it creates the table with the necessary columns. The core of the algorithm involves performing arithmetic operations on the input rasters, applying the specified weights, and then combining them into an output raster representing the drought risk. This calculated output raster is then inserted into the drought_map table for visualization.

```
CREATE OR REPLACE FUNCTION calculate_drought_risk(spi_raster raster, rainfall_anomaly_raster raster,
ndvi_raster raster, output_table_name text)

RETURNS void AS $$

DECLARE

   spi_weight float := 0.4;

   rainfall_anomaly_weight float := 0.3;

   ndvi_weight float := 0.3;

   output_raster raster;

BEGIN

   IF EXISTS (SELECT 1 FROM pg_tables WHERE tablename = output_table_name) THEN

      EXECUTE 'DELETE FROM ' || output_table_name;

   ELSE

      EXECUTE 'CREATE TABLE ' || output_table_name || ' (rid serial primary key, rast raster)';

   END IF;


   output_raster := (spi_raster * spi_weight + rainfall_anomaly_raster * rainfall_anomaly_weight + ndvi_raster
* ndvi_weight)::raster;

   EXECUTE 'INSERT INTO ' || output_table_name || ' (rast) VALUES (' || quote_literal(output_raster) || ')';

END;

$$ LANGUAGE plpgsql;
```

To enhance data visualization, a web-based dashboard is crafted, utilizing Python libraries including panel, Hvplot, Geoview, Geopandas, and pandas. These open-source libraries create interactive plots, geospatial maps, and enable data retrieval from the PostgreSQL database. The dashboard serves as a comprehensive tool for decision-makers and the public, providing insights into emergency management. Features encompass risk maps, tables showcasing risk levels, graphs, charts indicating risk status, and a slider for area selection and querying. The interactive dashboard exemplified in Figure 2 displays drought risk levels in Ethiopia. These levels are derived from a weighted average of NDVI, SPI, and Rainfall Anomaly, with higher values denoting severe drought and lower values indicating minimal or no drought. Moreover, the dashboard offers insights into drought risk at the Ethiopian Woredas level. This visualization tool enhances accessibility and empowers users to make well-informed decisions regarding emergency management.

Figure 2: Dashboard for Drought Risk Mapping



5. Evaluation

The evaluation was conducted to assess the effectiveness and usability of the prototype, and to identify areas for improvement.

A total of 10 GIS experts completed the survey. All ten experts were chosen based on their experience and expertise in satellite image analysis, disaster risk mapping, PostgreSQL, PostGIS, and cloud computing. The findings are presented on the table below.

Table 1: Evaluation Criteria and feedback from the respondant

1	Data handling storage efficiency	90% of agreed or strongly agreed the prototype handle and store large volume of spatial data.
2	scalability	100% agreed or strongly agreed prototype can scle up or scale down to accommodate the growing volume of data
3	Security and privacy	60% disagree or strongly disagree the prototype have robust security feature and need to implement database role.
4	Interoperability	100% of respondents agreed or strongly agreed that the prototype can able to integrate with other system or software tools
5	Performance	100%) agreed or strongly agreed that the prototype have fast response times and minimal downtime.
6	User interface and usability	60% of the respondents agree or strongly agree that the prototype is user friendly but 40% disagree because rather than using using SQL script it is better to develop QGIS plugin. Also suggest to use machine learning model
7	Accuracy and reliability	80 % of respondents agree or strongly agree that the prototype able to provide accurate and reliable information. And 20% neutral because all the data are acquired from third party source and they agree that it is better to use data from real time sensors.
8	Cost effectiveness	100% of the respondents agreed or strongly agreed that the prototype is cost-effective, both in terms of development and maintenance costs

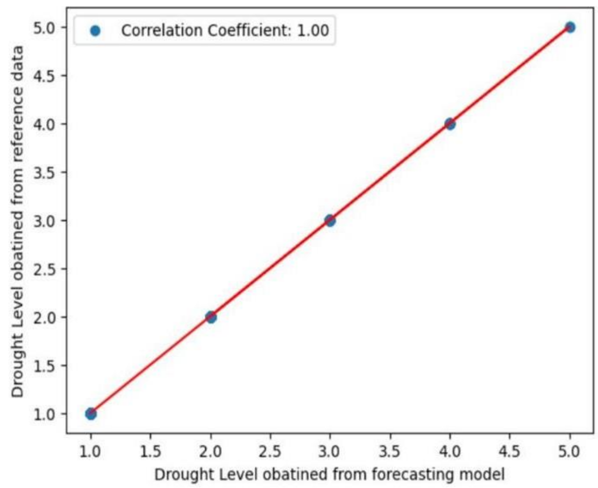
To assess the accuracy of a drought forecasting analysis evaluation metrics like correlation coefficient is used.

The correlation coefficient measures the strength and direction of the linear relationship between two variables (in this case, the drought forecasts and the reference data). It ranges from -1 to 1, where -1 indicates a perfect negative linear relationship, 1 indicates a perfect positive linear relationship, and 0 indicates no linear relationship.

The reference data for this research, to verify the accuracy of forecasted drought result, international databases like https://data.humdata.org is used. This websites have humanitarian data across the word and owned by UN OCHA including drought indexing of Ethiopia for year 2022.

By calculating correlation coefficient of the forecasted drought result of the research framework and reference data from the website sated above, using popular python on Jupiter notebook, the correlation result is 1 for all woredas in Ethiopia.

Figure 3: Correlation of drought by forecasting model and reference data using correlation coefficient



6. Conclusion and Future work

In conclusion, this research highlights the pivotal role of cloud computing in reshaping how we manage emergency situations. Cloud computing offers a secure, adaptable, and cost-efficient platform for responding quickly to unexpected and potentially dangerous events. The main goal of this study was to create a practical framework for handling emergencies using modern technology.

Through an extensive review of relevant literature, we've developed a four-layer framework that encompasses data gathering, storage, processing, and presentation. Each layer corresponds to a specific phase of emergency management: data collection, data storage, data analysis, and data visualization.

Our prototype testing validates the effectiveness of this spatial cloud data management framework. It efficiently handles large volumes of spatial data, scales to accommodate increasing data needs, provides easy access from different locations and devices, integrates smoothly with other systems, and does so in a cost-effective manner. However, there's room for improvement, including enhancing security and privacy measures, making the user interface more intuitive using a dedicated QGIS plugin, and exploring machine learning models to enhance accuracy in predicting droughts. Also, investigating more dependable sources of real-time data is a promising avenue for future exploration.

In essence, this research underscores how cloud computing can transform emergency management while offering a practical framework that aligns different data-related tasks. While the prototype performs well, ongoing refinements and advanced features are crucial to meeting the evolving demands of emergency response. This ensures a more secure, efficient, and robust system for safeguarding lives and resources during emergencies.

Reference

[1] Aly, A. G., El-Sharkawy, M. H., El-Bahnasawy, E. H., & El-Hadad, S. A. (2013). Proposed model of GIS-based cloud computing architecture for emergency system. International Journal of Computer Science and Mobile Applications, 1(1), 17-28

[2]Yang,C., Goodchild,M., Huang,Q., et al., 2011.Spatial cloud computing: how can the geospatial sciences use and help shape cloud computing. International Journal of Digital Earth, 4(4), pp. 305-329

[3] X. Ge and H. Wang, "Cloud-based Service for Big Spatial Data Technology in Emergence Management," in Proceedings of the 2011 IEEE International Conference on Computer Science and

Automation Engineering, Shanghai, China, 2011, pp. 324-328.

[4] W.W. Song, B.X. Sin, S.H. Li, et al., "Building Spatio-temporal Cloud Platform for Supporting GIS Application," in Proceedings of the 2015 IEEE International Conference on Big Data (Big Data), 2015, pp. 3112-3117. DOI: 10.1109/BigData.2015.7364094.

[5] D. Jiang and C. Yang, "Spatial Big Data Analytics for Emergency Management," in IEEE Access, vol. 3, pp. 890-898, 2015, doi: 10.1109/ACCESS.2015.2451843.

[6] S. Gao, X. Yuan and C. Zhang, "Spatial Data Management in the Cloud," in IEEE Transactions on Services Computing, vol. 6, no. 3, pp. 373-387, July-Sept. 2013, doi: 10.1109/TSC.2012.7.

[7] Demisse, G. B., & Rowland, J. D. (2014). "A spatial drought index for the Horn of Africa." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(8), 3374-3384.

[8] Demisse, G. B., & Soliman, A. S. (2017). "Spatial analysis of drought in Ethiopia using the standardized precipitation index." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10(3), 966-972.

Publication 500-292." Available: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication500-292.pdf

[9] Huete, A.R., Didan, K., Miura, T., Rodriguez, E.P., Gao, X. and Ferreira, L.G. (2002). "Overview of the radiometric and biophysical performance of the MODIS vegetation indices." IEEE Transactions on Geoscience and Remote Sensing, 41(6), pp. 1260-1266.

[10]Nigussie, A., & Tsunekawa, A. (2015). "Drought risk mapping using remote sensing and GIS: A case study of the Upper Awash River Basin, Ethiopia." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(6), 2925-2938