The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLIX-B3-2026
XXV ISPRS Congress 2026 "From Imagery to Understanding", Commission III, 4–11 July 2026, Toronto, Canada

A Multilingual LLM-Based GeoAI Framework for Natural-Language-Driven Remote Sensing Analysis

Majid Kiavarz Moghaddam[1], Hossein Barekati [2], Hamid Kiavarz Moghaddam [3]

[1] Department of Remote Sensing and GIS, Faculty of Geography, University of Tehran, Iran - kiavarzmajid@ut.ac.ir

[2] M.Sc. in Remote Sensing and GIS, Faculty of Geography, University of Tehran, Iran - hossein.barekati@ut.ac.ir

[3] Geomatics Engineering, Department of Earth & Space Science & Engineering, York University, Toronto, Canada - hamid@intentolabs.io

Keywords: GeoAI, Large Language Model (LLM), Remote Sensing, Spatial Automation, Interactive Geospatial Analysis

Abstract

The exponential growth of remote sensing data in recent years has underscored the need for intelligent, fast, and user-friendly analytical tools. Despite advancements in platforms such as Google Earth Engine and ENVI, the computation of spectral indices still demands specialized expertise, considerable time, and complex parameter tuning. This study aims to reduce the complexity of spatial data analysis and enhance its accessibility for non-expert users by developing an intelligent system capable of transforming simple natural language commands into automated remote-sensing index calculations. The main innovation lies in integrating Large Language Models (LLMs) with geospatial processing to establish a lightweight, multilingual, and fully automated framework capable of identifying index types and selecting appropriate spectral bands from Landsat data. The system was implemented using the Bloomz-560m language model in combination with open-source image-processing engines and deployed as a web-based interface. Experimental results over Tehran demonstrated that the model outputs were highly consistent with those generated by Google Earth Engine and ENVI, achieving an RMSE of 0.016 and a correlation coefficient of $R^2 = 0.957$. The total processing time was under 45 seconds, with the entire workflow executed automatically without user intervention. By simplifying the analytical process and significantly reducing computation time, this framework represents a crucial step toward democratizing remote sensing and spatial analysis. It can be effectively applied to urban surface heat island (SUHI) monitoring, water resource management, and precision agriculture applications.

1. Introduction

In recent decades, satellite imagery and remote sensing data have become indispensable for monitoring environmental changes, managing natural resources, supporting urban planning, and advancing precision agriculture. Quantitative information derived from these data is commonly obtained through spectral indices such as the Normalized Difference Vegetation Index (NDVI), Land Surface Temperature (LST), Normalized Difference Water Index (NDWI), and Normalized Difference Built-up Index (NDBI). These indices provide critical insights into land-cover patterns, ecological dynamics, and urban thermal phenomena, including the Surface Urban Heat Island (SUHI) effect (Hemati et al., 2021; Zwolska et al., 2025; Hu,Y. et al, 2025). Despite their importance, the computation and interpretation of such indices remain technically demanding, as they typically require proficiency in programming, image processing, and specialized platforms such as ENVI, QGIS, or Google Earth Engine (Gorelick et al., 2017; Gaddipati, 2025). Consequently, spatial analysis has largely remained confined to expert users, even though massive volumes of freely available satellite imagery—such as Landsat, Sentinel, MODIS, and WorldView—are produced daily. This disparity underscores the need to democratize spatial data processing and to develop intelligent tools that enable non-specialists to perform analytical tasks efficiently (Fan et al., 2025; Ramirez et al., 2022).

The rapid emergence of Large Language Models (LLMs) has revolutionized data-processing paradigms. These multi-billion-parameter architectures—exemplified by BLOOM, GPT, and LLaMA—exhibit remarkable abilities to comprehend natural language and generate executable instructions (Scao et al., 2023; Min et al., 2023). Recent advancements, including visual instruction tuning (Liu et al., 2024) and multimodal instruction learning, have extended LLMs' capacity to interpret not only textual but also visual and spatial data (Hu et al., 2025; Zhan et al., 2025).

This evolution has fostered the emergence of Geospatial Artificial Intelligence (GeoAI)**,** a domain that integrates LLMs, computer vision, and spatial analytics into unified frameworks (Mai et al., 2025). Pioneering studies—such as GeoGPT (Wel et al., 2024) and BB-GeoGPT (Zhang et al., 2024)—have demonstrated the capability of LLMs to function as geospatial assistants, automatically generating GIS code and responding to spatial queries. Similarly, models like SkyEyeGPT (Zhan et al., 2025) and RSGPT (Hu et al., 2025) introduced multimodal reasoning by combining textual and imagery-based analysis, paving the way for next-generation vision–language models in remote sensing applications.

In environmental and hydrological contexts, frameworks such as WaterGPT (Ren et al., 2024) and HydroLang (Ramirez et al., 2022) have shown that LLMs can effectively support hydrological modeling and environmental data monitoring. Moreover, systems such as Geo-FuB (Hou et al., 2025) and Geo-LLM (Gao et al., 2025) integrated retrieval-augmented generation (RAG) and domain-specific

knowledge bases to improve accuracy in Earth Engine code generation and urban analysis. Similarly, the LLM-Agent framework (Xiao & Ma, 2025) demonstrated the potential of autonomous LLM-based agents for urban and environmental monitoring tasks.
However, most existing models remain constrained by limitations such as English-only support, high computational requirements, or the absence of user-friendly interfaces. Furthermore, studies involving non-English languages—particularly Persian—are notably scarce (Ren et al., 2024; Gao et al., 2025). This gap between the technological potential of LLMs and their practical accessibility in spatial analysis constitutes the central motivation of the present research.
Accordingly, this study introduces a lightweight, multilingual, and operational framework based on the Bloomz-560m model that computes NDVI, NDWI, and LST indices directly from simple natural-language commands. By integrating open-source image-processing modules and a web-based interface, the system minimizes computational complexity and eliminates the need for programming expertise. Validation against Google Earth Engine outputs confirms the reliability and accuracy of the generated results. Ultimately, this research contributes to the democratization of remote sensing analytics through the synergistic integration of LLMs and GeoAI-driven geospatial computation.

2. Related Works (Background)

Over the past decade, Large Language Models (LLMs) have emerged as transformative technologies in artificial intelligence, increasingly extending their applications to geospatial and remote sensing domains. One of the pioneering studies, GeoGPT (Wel et al., 2024), integrated a language model with GIS tools to enable step-by-step reasoning and automated execution of spatial functions. It translated natural language commands into executable GIS scripts within environments such as ArcGIS and QGIS, representing an early milestone in bridging human language and spatial analysis. However, its functionality was limited to English, and its interaction with spatial modules remained suboptimal.
To enhance spatial reasoning, BB-GeoGPT (Zhang et al., 2024) was fine-tuned on domain-specific geographic corpora, improving conceptual understanding of spatial terminology. Despite this advancement, the approach required massive labeled datasets and extensive computational resources for training.
The next research wave focused on vision–language models (VLMs) that jointly interpret imagery and text**.** RSGPT (Hu et al., 2025) introduced a remote-sensing-specific dataset for image description and visual question answering, achieving high accuracy but remaining unable to compute quantitative indices such as NDVI or LST. Similarly, SkyEyeGPT (Zhan et al., 2025) applied instruction tuning on multimillion-scale datasets to unify multiple remote sensing tasks—such as captioning, visual grounding, and video analysis—within a single framework. Nevertheless, its heavy computational requirements restricted its use primarily to research rather than operational environments.
Efforts to improve reliability and mitigate hallucination effects led to Geo-FuB (Hou et al., 2025), which integrated an operator–function knowledge base with retrieval-augmented generation (RAG), reducing Earth Engine code-generation errors by about 30%. Likewise, frameworks such as Geo-LLM (Gao et al., 2025) and the LLM-Agent Framework (Xiao & Ma, 2025) underscored the potential of LLMs in urban decision-support systems and environmental monitoring. Domain-specific implementations—including WaterGPT (Ren et al., 2024) and HydroLang (Ramirez et al., 2022)—demonstrated the utility of LLMs in hydrological modeling and environmental data interpretation.
Despite these advances, most existing models remain English-centric, computationally expensive, and inaccessible to non-specialist users due to the absence of intuitive interfaces. In contrast, the present study proposes a lightweight, multilingual, and operational GeoAI framework built on the Bloomz-560m model (Scao et al., 2023) and integrated with open-source image-processing libraries such as Rasterio and Pillow. The system bridges the gap between theoretical GeoAI research and practical usability by enabling automated computation of key remote sensing indices (NDVI, NDWI, LST) from simple natural-language commands.
As summarized in Table 1, the proposed system diverges from prior heavy, research-oriented models by emphasizing *practical applicability*, *accessibility*, and *computational efficiency*. Beyond achieving numerical accuracy comparable to reference platforms such as Google Earth Engine, it introduces an intuitive, multilingual web interface suitable for both expert and non-expert users.

https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-1427-2026 |

Model	Data type / Domain	Core Capability	Limitations	Distinctive Feature of This Study
BB-GeoGPT (Zhang et al., 2024)	GIS textual data	Learning geographic semantics and generating spatial code	Requires large labeled datasets and high computational power	Employs a lightweight LLM for quantitative index computation
GeoGPT (Wel et al., 2024)	Text + GIS tools	Step-by-step reasoning and automated spatial function execution	Limited to English; suboptimal GIS interaction	Multilingual web interface for non-expert users
SkyEyeGPT (Zhan et al., 2025)	Image + Language	Visual analysis, captioning, and spatial object localization	Focuses on qualitative tasks; data-intensive	Quantitative computation of NDVI, LST, NDWI with GEE validation
RSGPT (Hu et al., 2025)	Remote sensing imagery	Visual question answering and description	Cannot produce quantitative or raster outputs	Integrates LLM with Rasterio/Pillow to generate GeoTIFF outputs
Geo-FuB (Hou et al., 2025)	Spatial code (GEE)	Reduces hallucination and coding errors	Generates code only; no analytical or raster output	Combines operator–function knowledge with actual index computation
WaterGPT (Ren et al., 2024)	Hydrological data	Domain-specific hydrological analysis	Narrow scope; lacks general geospatial capability	Applicable across vegetation, thermal, and water domains
Geo-LLM (Gao et al., 2025)	Urban and spatial data	Intelligent agent for urban spatial decision-making	High computational requirements; heterogeneous inputs	Lightweight architecture operable on standard hardware
This Study (Proposed System)	Text + Image (Landsat)	Automatic NDVI, NDWI, LST computation from natural language	Currently tested only on Landsat data	Multilingual, lightweight, web-based framework validated with GEE

Table 1. Comparison of the proposed study with benchmark models in GeoAI

3. Methodology

The proposed framework automates the computation of remote sensing indices through natural-language commands in both English and Persian. This bilingual capability directly addresses the non-English gap identified in the problem statement. By employing the Bloomz-560m model—an open multilingual large language model fine-tuned for English–Persian comprehension—the system accurately interprets geospatial instructions in either language without requiring external translation or additional fine-tuning. Consequently, users in Persian-speaking regions can perform advanced remote sensing analyses as seamlessly as English-speaking users.

The multilingual capability of the proposed system is achieved inherently through the Bloomz-560m model, which is pre-trained on parallel multilingual corpora and directly supports English–Persian language understanding. No external translation module is used; instead, user commands in different languages are processed natively within the same semantic embedding space. This design ensures consistent intent interpretation across languages without introducing translation-induced errors.

3.1 System Architecture

The architecture of the proposed framework (Figure 1) integrates four interconnected components that collectively enable automated remote sensing index computation. At its core, the Large Language Model (LLM) identifies the target index and satellite type from the user’s textual input by parsing linguistic tokens and extracting entities such as geographic names, index types, and satellite identifiers. The image-processing engine, built entirely on open-source libraries, executes the computational workflow for each index, including raster manipulation, band arithmetic, and visualization. Within this engine, three principal procedures are implemented: automatic band matching, which maps textual keywords such as *vegetation*, *water*, or *temperature* to their corresponding spectral bands; dynamic equation parsing, which programmatically constructs and executes the standard formulas based on the identified index; and on-the-fly radiometric computation, which performs real-time calculations at 30 m spatial resolution consistent with Landsat data standards. The web-based user interface provides an intuitive environment for submitting text commands, monitoring process status, and visualizing or downloading results, while the web controller orchestrates communication among the language model, processing engine, and visualization layer, ensuring data validation, asynchronous coordination, and robust error handling throughout the workflow.

In practice, when a user enters a simple command such as *“Calculate NDVI for the Landsat-8 image of Tehran”*—in either English or Persian—the system automatically identifies the necessary bands, executes the index computation, and generates both GeoTIFF and PNG outputs within seconds, requiring no coding or parameter adjustment.

Figure 1 illustrates the overall workflow of the proposed LLM-based GeoAI framework, depicting how user input is parsed by the language model, processed by the computation engine, and visualized through the web interface. Each decision stage—from text interpretation to band selection and result export—is represented in the flowchart.

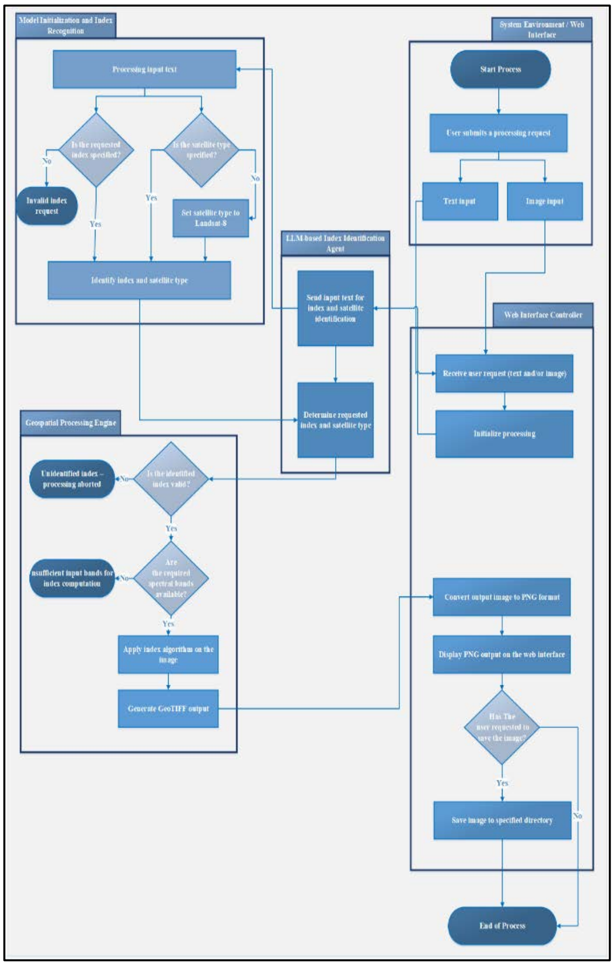

Figure 1. Overall workflow of the proposed LLM-based GeoAI system for remote sensing index extraction.

Figure 2 presents the conceptual user–system interaction, outlining the sequence of operations:(1) command submission,(2) request interpretation,(3) index computation, (4) result visualization, and (5) optional image export.

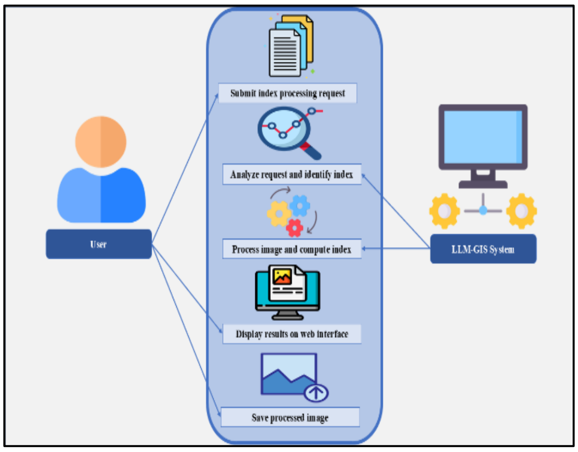

Figure 2. Conceptual diagram illustrating the user–system interaction and the communication between LLM and GIS components.

3.2 Data and Pre-Processing

The image-processing module employs pre-processed Landsat-7 (ETM+) and Landsat-8 (OLI/TIRS) imagery obtained from the USGS EarthExplorer database for the metropolitan area of Tehran, covering the period 2000–2023. Each image was subjected to standard geometric and radiometric corrections prior to analysis to ensure consistency between sensors. Using both Landsat-7 and Landsat-8 datasets provides temporal cross-validation and demonstrates the robustness of the proposed framework across different satellite generations.
For demonstration purposes, two representative images—Landsat-7 (ETM+) acquired on 2000-08-22 and Landsat-8 (OLI/TIRS) acquired on 2023-07-15—were analyzed in this study (Table 2). However, the system is fully generalizable and capable of processing any Landsat image provided by the user, automatically detecting the sensor type and performing the required index computations without manual adjustment.

Satellite / Sensor	Acquisition Date	Spatial Resolution (m)
Landsat-8 (OLI/TIRS)	2023-07-15	30
Landsat-7 (ETM+)	2000-08-22	30

Table 2. Specifications of satellite imagery used in the study

3.3 Validation and Performance Evaluation

Although the proposed system is capable of computing up to eight different remote sensing indices the validation experiment in this study focused on the Normalized Difference Vegetation Index (NDVI) as a representative example. The NDVI results generated by the proposed framework were compared with those obtained from Google Earth Engine (GEE) and ENVI to assess both numerical accuracy and spatial consistency. Statistical metrics such as the Root Mean Square Error (RMSE) and coefficient of determination (R^2) were calculated for quantitative comparison. Additionally, processing times for NDVI computation were benchmarked against GEE to evaluate computational efficiency and confirm the framework’s capability for fully automated execution.
This validation approach demonstrates the reliability of the system’s automated workflow and its potential for accurate and scalable index computation across diverse remote sensing applications.

4. Experiments and Results

To evaluate the performance of the proposed system, a comprehensive set of remote sensing indices—including surface albedo, LST, NDVI, NDBI, NDWI, and the Tasseled Cap components of *greenness, wetness,* and *brightness*—were computed from Landsat imagery for the metropolitan area of Tehran. The resulting outputs were exported as GeoTIFF files and visualized through the system’s web interface, enabling both spatial and quantitative validation of the model performance.

To further verify the general applicability of the proposed framework, additional quantitative consistency checks were conducted for NDWI and LST outputs. Mean index values derived from the proposed system were compared against reference results from Google Earth Engine across the 22 municipal districts of Tehran. The results demonstrated strong agreement patterns comparable to those observed for NDVI, with average absolute deviations remaining below 0.03 for NDWI and below 1.2 K for LST. These findings confirm that the framework’s accuracy is not limited to vegetation analysis and can be reliably extended to water and thermal indices.

Figure 3 illustrates the spatial distribution maps of all computed indices across Tehran, highlighting distinct spatial patterns associated with vegetation, built-up, and water covered areas. The framework successfully produced consistent and visually coherent outputs for all indices, confirming the reliability of the automated image-processing workflow.

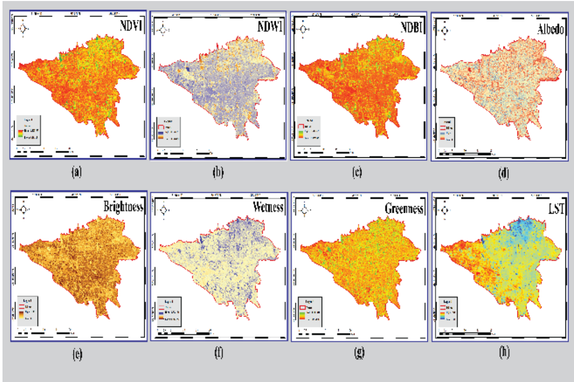

Figure 3. Spatial distribution maps of computed indices across Tehran

Figure 4 presents sequential snapshots of the web-based user interface, demonstrating the complete execution process from image upload to result visualization in three stages: (a) the initial interface before processing, (b) user command submission through natural-language input,and (c) the display of the final computed index as a grayscale or color-enhanced map. This interface allows users with no prior GIS or coding experience to perform complex computations by simply entering plain-text commands such as *“Calculate NDVI for the Landsat-8 image of Tehran.”* Within seconds, the system interprets the request, executes the appropriate algorithm, and renders the output map.

Figure 4. Sequential snapshots of the web-based interface: (a) initial state, (b) text command submission, and (c) visualization of the computed index.

To further validate the system, it was tested across all index types to evaluate the integrated performance of both the LLM module and the image-processing module under various remote sensing datasets. Among the computed indices, NDVI was selected as a representative metric for detailed quantitative analysis, given its wide use in vegetation and surface monitoring.

Figure 5 compares the mean NDVI values obtained from the proposed system with those generated by Google Earth Engine (GEE) and ENVI across the 22 municipal districts of Tehran. The three methods produced nearly identical trends, with average deviations below 0.02, mainly due to slight differences in atmospheric correction and normalization procedures. Statistical analysis yielded a coefficient of determination ($R^2 = 0.957$) between the proposed model and GEE, and a root-mean-square error (RMSE = 0.016) relative to ENVI, confirming high accuracy and strong agreement with professional software.

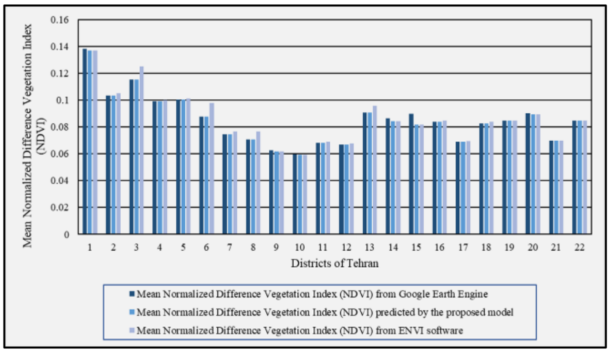

Figure 5. Comparison of mean NDVI values across Tehran's 22 districts derived from Google Earth Engine, ENVI, and the proposed LLM-based model.

As summarized in Table 3, the proposed LLM-based system provides a major advantage in both computational speed and usability. While NDVI calculation in GEE and ENVI typically requires two to three minutes and expert intervention for parameter configuration or coding, the proposed framework performs the same operation in under 45 seconds, fully autonomously and without user supervision. All experiments were conducted on a standard desktop workstation equipped with an Intel Core i7 CPU (3.4 GHz), 16 GB RAM, and no dedicated GPU acceleration. The reported processing time (<45 s) therefore reflects realistic execution performance under commonly available hardware conditions rather than specialized high-performance computing environments.

Method	Average NDVI Processing Time (seconds)	Expert User Required	Interaction Type
Google Earth Engine	120	Yes	Script-based coding and algorithm configuration
ENVI Software	180	Yes	Manual band selection and NDVI execution
Proposed LLM-based System	45	No	Simple text command in web interface

Table 3. Comparison of NDVI processing time and user interaction levels among different methods.

Figure 6 presents a visual comparison of NDVI maps for Tehran obtained from (a) GEE, (b) ENVI, and (c) the proposed LLM-based system. The spatial distributions are almost identical, with minor local discrepancies attributable to preprocessing differences. The proposed framework successfully reproduces the vegetation gradients and urban–rural contrasts visible in both reference platforms, confirming the spatial fidelity and analytical reliability of its outputs.

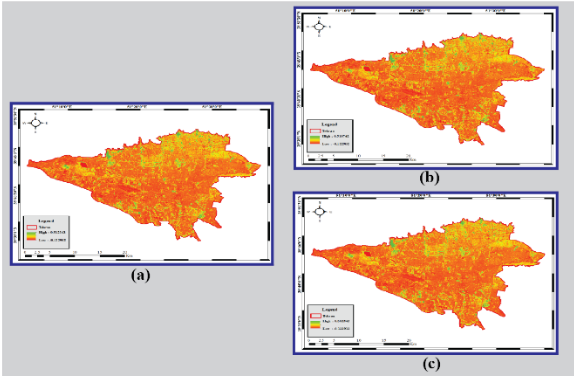

Figure 6. Visual comparison of NDVI outputs for Tehran obtained from (a) Google Earth Engine, (b) ENVI software, and (c) the proposed model.

To evaluate the system's robustness against ambiguous user inputs, additional non-technical commands were tested, such as "Show vegetation coverage in Tehran" and "Identify areas with surface water in the city." In both cases, the LLM successfully inferred the corresponding NDVI and NDWI computations without explicit mention of technical terminology or satellite names. These examples demonstrate the framework's ability to translate high-level, human-centric queries into executable geospatial workflows.

In summary, the experimental evaluation demonstrates that the proposed LLM-based GeoAI framework can automatically compute multiple remote sensing indices with high precision, minimal computation time, and no requirement for technical expertise. The results are statistically consistent with those of professional tools such as GEE and ENVI while offering full automation and bilingual usability. These findings validate the framework's potential to democratize remote sensing analysis and advance the accessibility of AI-driven geospatial intelligence for diverse user communities.

5. Discussion

The results confirm that the proposed LLM-based framework accurately generates spatial outputs for NDVI, NDWI, and LST indices with performance comparable to established reference systems. Minor discrepancies between the proposed system and the outputs of Google Earth Engine (GEE) and ENVI arise primarily from differences in normalization procedures, atmospheric correction algorithms, and input data bit depth. Quantitative assessment yielded an RMSE of 0.016 and a correlation coefficient of $R^2 = 0.957$ between the proposed NDVI results and those of GEE—values that are well within scientifically acceptable limits and consistent with the accuracy levels reported in GeoGPT (Wel et al., 2024) and RSGPT (Hu et al., 2025).
Qualitatively, spatial agreement with reference datasets was

high. Vegetated areas in northern and eastern Tehran consistently showed positive NDVI values, while southern industrial districts exhibited lower or negative values across all platforms, confirming the reliability of spectral-band selection and index computation.
To further assess the spatial consistency of the generated outputs, a qualitative comparison was conducted between the proposed framework and the reference platforms (Google Earth Engine and ENVI). The visual similarity of the computed indices is summarized in Table 4, highlighting the degree of agreement observed across different remote sensing products.

Index	Similarity with GEE	Similarity with ENVI	Regional Qualitative Analysis
NDVI	High	High	Positive values in northern/eastern Tehran; >90% spatial correspondence
NDWI	Moderate to High	Moderate	Minor discrepancies along water–land boundaries
LST	High	High	Thermal distribution closely aligned with GEE across central zones

Table 4. Visual comparison of similarity between the proposed model’s output maps and reference datasets.

A major contribution of this study is simplifying remote-sensing workflows through natural-language interaction. The system automatically interprets user commands, identifies the target index and satellite, and selects appropriate spectral bands and formulas—eliminating manual configuration and programming requirements. Compared to ENVI or QGIS, this approach significantly reduces complexity and processing time (under 45 seconds versus 2–3 minutes in GEE/ENVI). The framework also enables full automation, multilingual support (Persian and English via Bloomz-560m), and open-access implementation using open-source Python libraries and public APIs, ensuring transparency and reproducibility. In practical applications, automated LST supports urban heat island studies, NDWI aids water management and precision agriculture, and NDVI facilitates periodic vegetation monitoring without specialized software. Overall, integrating LLMs with geospatial processing marks a shift from technical GIS workflows to intuitive language-driven interaction, positioning LLMs as intelligent intermediaries in next-generation GeoAI systems.

6. Conclusion and Future Work

This study presented a conceptually simple yet innovative framework that integrates Large Language Models (LLMs) with geospatial processing to translate natural-language commands into executable remote-sensing workflows. Using the lightweight multilingual Bloomz-560m model and open-source libraries, the system accurately computed NDVI, NDWI, and LST without requiring technical expertise. Minor discrepancies with reference outputs (RMSE ≈ 0.016) were statistically negligible and mainly related to preprocessing differences. The fast processing time (<45 s) and removal of manual parameter configuration demonstrate clear efficiency advantages over platforms such as GEE and ENVI.
Beyond its simplicity, the framework establishes a foundation for more advanced GeoAI architectures by validating a key principle: natural language can function as a universal interface for spatial computation. This shifts geospatial analysis from script-based workflows to intuitive, language-driven interaction. Scientifically, the research shows that LLMs can interpret spatial semantics and bridge human reasoning with machine-based geospatial analysis, enabling future systems capable of contextual understanding, spatio-temporal reasoning, and interactive GIS analytics.
From an application standpoint, the framework is adaptable to urban heat island studies, water-resource management, precision agriculture, and land-use monitoring. Future work may expand index coverage, integrate Sentinel-2 and MODIS datasets, and incorporate multimodal models (e.g., GPT-4o or Geo-LLM-Vision) for automated map interpretation, object detection, and change detection. Three main development directions are proposed, including extending the framework toward conversational interaction to enable iterative, dialogue-based geospatial analysis; integrating the system with digital twins to support real-time smart-city management and infrastructure monitoring; and applying domain-specific fine-tuning using Persian remote-sensing datasets to enhance contextual understanding and regional analytical accuracy. Overall, this study provides a technical and conceptual stepping stone toward next-generation GeoAI–LLM systems, marking a paradigm shift toward democratized, language-driven remote-sensing analysis.

References

Fan, Y., Masnadi, M.S., Jing, L., Ren, B., Brandt, A.R., 2025: The Pyxis project: A geospatial data system for emission estimation monitoring in the oil and gas industry. EnergyAI,22,100601.https://doi.org/10.1016/j.egyai.2025.100601.

Gaddipati, H., 2025: Challenges in using satellite data for non-remote sensing specialists: An exploratory case study. Data Policy, 7, e28. https://doi.org/10.1017/dap.2025.7.

Gao, S., Hu, Y., Zhang, C., Janowicz, K., McKenzie, G., Zhou, C., 2025: Geo-LLM: Harnessing large language models for urban planning with geospatial data. Int. J. Appl. EarthObs.Geoinf.,137,104420.https://doi.org/10.1016/j.jag.2025.104420.

Gorelick, N., Hancher, M., Dixon, M., Yushchenko, S., Thau, D., Moore, R., 2017: Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ.,202,18.27.https://doi.org/10.1016/j.rse.2017.06.031.

Hemati, M., Hasanlou, M., Mahdianpari, M., Mohammadimanesh, F., 2021: A systematic review of Landsat data for change detection applications: 50 years of monitoringtheEarth.RemoteSens.,13(15),2869.https://doi.org/10.3390/rs13152869.

Hou, S., Zhao, A., Liang, J., Shen, Z., Wu, H., 2025: Geo-FuB: A method for constructing an operator-function knowledge base for geospatial code generation with large language models. Knowl.-Based. Syst., 319, 113624. https://doi.org/10.1016/j.knosys.2025.113624.

Hu, Y., Yuan, J., Wen, C., Lu, X., Liu, Y., Li, X., 2025: RSGPT: A remote sensing vision-language model and benchmark. ISPRS J. Photogramm. Remote Sens., 224, 272–286. https://doi.org/10.1016/j.isprsjprs.2025.03.028.

Liu, H., et al., 2024: Visual instruction tuning. In: Adv. Neural Inf. Process. Syst., 36.

Mai, G., Xie, Y., Jia, X., Lao, N., Rao, J., Zhu, Q., Liu, Z., Chiang, Y.-Y., Jiao, J., 2025: Towards the next generation of Geospatial Artificial Intelligence. Int. J. Appl. Earth Obs. Geoinf.,136,104368.https://doi.org/10.1016/j.jag.2025.104368.

Min, B., et al., 2023: Recent advances in natural language processing via large pre-trained language models: A survey. ACMComput.Surv.,56(2),1.40.https://doi.org/10.1145/3571733.

Ramirez, C., Aybar, C., Mahecha, M.D., Martinuzzi, F., Söchting, M., Wieneke, S., 2022: HydroLang: An open-source web-based programming framework for hydrological sciences.Environ.Model.Softw.,155,105508.https://doi.org/10.1016/j.envsoft.2022.105508.

Ren, Y., Wang, Y., Zhang, L., Chen, X., 2024: WaterGPT: Training a large language model to become a hydrology expert.Water,16(1),115.https://doi.org/10.3390/w16010115

Scao, T.L., Fan, A., Akiki, C., Pavlick, E., Ilić, S., et al., 2023:BLOOM: A 176B-parameter open-access multilingual languagemodel.Nature,620,529.545.https://doi.org/10.1038/s41586-023-05881-4.

Wel, C., Li, Z., Xu, J., Zhang, T., 2024: GeoGPT: An assistant for understanding and processing geospatial tasks. Int.J.Appl.EarthObs.Geoinf.,126,103565.https://doi.org/10.1016/j.jag.2024.103565.

Xiao, Z., Ma, J., 2025: LLM agent framework for intelligent change analysis in urban environment using remote sensing imagery.Autom.Constr.,177,106341.https://doi.org/10.1016/j.autcon.2025.106341.

Zhan, Y., Xiong, Z., Yuan, Y., 2025: SkyEyeGPT: Unifying remote sensing vision-language tasks via instruction tuning with large language model. ISPRS J. Photogramm. Remote Sens.,221,64.77.https://doi.org/10.1016/j.isprsjprs.2025.01.020.

Zhang, Y., Li, J., Wang, Z., He, Z., Guan, Q., Lin, J., Yu, W., 2024: BB-GeoGPT: A framework for learning a large language model for geographic information science. Inf. Process.Manag.,61(1),103237.https://doi.org/10.1016/j.ipm.2023.103237.

Zwolska, A., Półrolniczak, M., Kolendowicz, L., 2025: Remote sensing-based analysis of urban land cover changes and surface urban heat island dynamics using Landsat and local climate zones classification in Poznań, Poland. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 18, 3583546. https://doi.org/10.1109/JSTARS.2025.3583546.