International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Mathematically-Grounded, Explainable Multi-Agent Systems for Terrain Hyperspectral Analysis and Triage

Nasrulla khan k[1], Dr. C R Sakthivel[2]

Research Scholar, Department of Computer Science, Sri Ramakrishna Mission Vidyalaya college of Arts & Science, Coimbatore, India
Research Guide, Department of Computer Science, Sri Ramakrishna Mission Vidyalaya college of Arts & Science, Coimbatore, India

Abstract

Agentic systems, in which massive language models manage analytical tools and large foundation models that are trained on data from earth observation systems are two concurrent trends that are transforming the field of remote sensing. A literature review of 2025-2026 shows that both of them have more powerful and independent analysis but neither of them can meet the needs of high-stakes terrain analysis. The agentic remote-sensing systems of today are characterized by high tool-orchestration error, shallow non-replayable memory, hyperspectral and spectral-physics integration is an open challenge and reasoning about red-green-blue and multispectral tools. Foundation models generate strong embeddings but do not have statistical guarantees and decision provenance. While optimal transport is already a widely used and established mathematical tool for solving the problem, there is also rapid development of other tools such as tensor and block-term decomposition, topological data analysis, graph diffusion and conformal prediction, which are each emerging as a stand-alone tool rather than as a common analytical substrate over which an agent can reason. In this survey, 20 representative papers from the year 2025-2026 related to geospatial and agentic artificial intelligence, Earth foundation models, hyperspectral unmixing and classification, topological and optimal-transport techniques, uncertainty quantification, explainable AI and change and tamper detection are discussed. It organizes them by the research topic that the theoretically inspired, explainable, LLM-independent multi-agent framework TerraXIA tackles. Unlike in other systems, each agent in TerraXIA generates uncertainty-bounded, verifiable triage based on a common mathematical representation of the world-state of a terrain hyperspectral scene. The review provides a consensus scientific basis for the recommended work program by identifying a consistent set of gaps and the relationship of each gap to the four research objectives and 12 innovations in the framework

Keywords

Hyperspectral Imaging, AgenticAI, Multi-Agent Systems, Explainable AI, Optimal Transport, Topological Data Analysis, Conformal Prediction, Tensor Decomposition, Remote Sensing, Terrain Analysis, Triage

1. Introduction

1.1. Motivation

Rapid methodological transformation has started in the field of Earth observation. Geographic AI is rapidly being incorporated into the remote-sensing pipeline, as demonstrated in comprehensive evaluations. Geographic AI is increasingly becoming part of the remote-sensing pipeline, as illustrated by comprehensive assessments. In addition to this, there are two special movements to highlight. The first is the advent of hyper spectral and satellite data pre-trained foundation models to provide a single flexible model capable of handling multiple downstream applications Zhu et al. [19], Wu et al. [15,16]. The latter is the advent of agentic systems, where a large language model creates and orchestrates specialized analytical tools rather than executing a single set task (Sun et al. [14], Akinboyewa et al. [1]). In the field of Earth observation, a momentum of rapid methodological evolution has started. Full-fledged geographic artificial intelligence representations show how the remote-sensing pipeline is increasingly being filled with transformers, multimodal learning, foundation models and multi-agent systems Mai et al. [10]. In addition to this, there are two particular movements. The first is the rise of hyperspectral and satellite pretraining or pretrained foundation models that are expected to provide a single common representation that can be applied to various downstream applications (Zhu et al. [19], Wu et al. [15,16]). The second is the presence of a series of “agentic systems”, where a large language model trains and orchestrates specialized analytical tools (Sun et al. [14], Akinboyewa et al. [1]).

1.2. The Problem: A Gap Between Capability and Trust

This momentum has left a gap between what these systems can see and what they can be trusted to determine, but a deeper look at the research shows that is the case. This survey is in response to four observations: First, the agents are still red-green-blue-centric. Today, copilot and agentic systems effectively synchronize colour and multi-spectral tools, but spectral physics and hyperspectral data have been presented as an open problem, not a solved one (Sun et al. [14], Akinboyewa et al. [1]). Second, agent systems tend to be less reliable and less auditable because they often perform their tools on difficult tasks in a mis-orchestrated way and only remember the action performed but not what has been done on the task. Thirdly, the right mathematics exists but it is used in isolation. Some examples include optimal-transport unmixing (Feng et al. [4]), tensor and block-term decomposition (Ren et al. [13], Zou et al. [20]), topological and spatio-temporal modeling Kuper et al. [7], graph diffusion Gui et al. [5], and conformal prediction Kuronen et al. [8]. Fourth, foundation models lack decision provenance and statistical guarantees when supporting perception: Earth and hyperspectral foundation models provide strong embeddings but not necessarily decision provenance or statistical guarantees (Zhu et al. [19], Wu et al. [15,16], Zhou et al. [18]). Because of this, there is a high level of perception in the field, but not a high level of responsible decision-making. The biggest hurdle for high stakes terrain review is having an inaccurate or unexplained flag and serious repercussions.

1.3. Contribution and Scope of This Survey

In this survey, no new algorithm is proposed. It aims to apply the state of the art to a set of goals to justify a research program. In particular, it contributes three things. For the first step, 20 sample works are first sorted by the thematic pillars that are relevant to the reliable agentic hyperspectral analysis. Second, it identifies exactly which gap did the object of the study not address for each pillar. Third, it maps the identified gaps to the four objectives and twelve innovations of the TerraXIA framework, showing the explicit objective-justification, which establishes the basis of the proposed work in and difference from the current research. The scope is purposefully limited. The review highlights papers published in 2025 and 2026 that directly contribute to one or more of the pillars of the framework: mathematical substrate (optimal transport, tensor, graph and topology), agentic and geospatial AI, foundation models for Earth observation, hyperspectral unmixing and classification, uncertainty quantification, explainability and change and tamper detection. It does not aim to be an exhaustive list of all deep-learning architectures for remote sensing, just a few architecture-related papers have been selected to illustrate the high level of perceptual baseline that the framework must be situated.

1.4. Method of Review

Twenty peer-reviewed publications were selected as authoritative reviews or representative methodological breakthroughs, they were recent (2025-2026), and they directly related to the pillars of the framework. Each work was assessed on three criteria: novelty contribution, specific gap left regarding reliable agentic hyperspectral triage and goal contribution. The rest of the survey is organized thematically, each section stating an argument. The research context is reiterated in Section 2. Pillar by pillar the literature is reviewed in Sections 3-10. The gaps are synthesized in Section 11 then completed in Section 12.

2. Research Context: The TerraXIA Framework

This section provides a quick restatement of the framework the survey is meant to support in order to make the review's argument more readable. TerraXIA is a multi-agent framework for terrain hyperspectral analysis and triage that is explainable, theoretically based and independent of LLM. Every agent reads from and writes to a shared mathematical representation of the scene instead of reading raw pixels as a black box and every agent action returns a verifiable numeric artifact along with its provenance. This shared mathematical world-state is its defining concept.

2.1. The Mathematical Substrate

The five mathematical tools that constitute the world-state were chosen as examples because they provide a transformation of unprocessed spectra to a structured format that can be understood and audited. Together they create a clear separation between the spectral (non-space) and spatial (space) elements of an image. The basis is provided by tensor algebra. Tucker and block-term decomposition separates the hyperspectral cube into a spatial-structure tensor and a low rank spectral subspace, the actual space/non-space distinction and provides a compression that allows for scene-scale analysis. Added to these the contextual-similarity and change operator is offered by optimum transport: The former is the transport plan itself and the latter is Wasserstein and Sinkhorn distances between spectral-abundance distributions that quantify the relationships between regions. Furthermore, topological data analysis also introduces intrinsic structures from persistent homology at various spectral indices and elevation to give rise to auditable basins, ridges, voids and terrain shape that are unaffected by rotation and illumination. The representation is scaled up to a scene level and the two remaining instruments make good decisions. Superpixel graphs, graph-Laplacian and diffusion distances encode spatial context and correlate results across many superpixels, enabling scene-scale triage with these techniques. Finally, conformal and physics restrictions give rise to non-negativity, sum-to-one and bounds on radiometric outputs, which exclude physically implausible outputs and conformal prediction with a distribution-free error rate gives a calibrated error rate and triage.

2.2. The Four Objectives

The framework is organised around four research objectives, which structure the justification developed in Section 11.

2.2.1 A mathematical world-state for spatial and non-spatial analysis

Decompose a massive hyperspectral cube into an interpretable spatial-structure component and a spectral-subspace component, the shared analytic base every downstream agent reads.

2.2.2 An explainable, LLM-agnostic multi-agent architecture (XIA)

Decompose a natural-language terrain intent into a directed graph of mathematical operators over the world-state, where every agent action returns a checkable numeric artifact plus provenance.

2.2.3 A statistically guaranteed triage and decision layer

Combine distribution-free conformal prediction with topological and optimal-transport anomaly scoring to rank regions for human review at a bounded, user-specified error rate.

2.2.4 Validation across layouting, survey and tamper/change

Validate on public terrain hyperspectral benchmarks across all three applications, demonstrating cross-scene, cross-sensor generalisation and scalability to massive scenes. The originality claimed by the framework is not any single operator but the agent-shared mathematical world-state combined with verifiable, uncertainty-bounded triage over massive terrain hyperspectral scenes. The literature review that follows is designed to test and support that claim.

3. Geospatial AI and the Rise of Agentic Remote Sensing

Geospatial AI offers the widest platform to support this. Mai et al., [10] review the topic and quantify literature trends, that the future generation of geospatial intelligence will be knowledge-guided, ethical and increasingly autonomous. In this case, their review is useful, since it is not about a particular method, but about the frontier of the discipline, from isolated models to integrated systems that are capable of reasoning. This framing both promotes the agentic approach of the framework and cautions against the autonomy without accountability which is a well-known danger. The agentic is materialized through two works. Sun et al. [14] introduce a large-language-model-based multi-agent system for remote sensing analysis, in which agents execute analytical tasks in an automated manner. Akinboyewa et al. [1]) present a natural-language query-to-geographical-action assistant, called GIS Copilot, that is incorporated into a geographic information system. Both show the attraction of the paradigm: the system assembles the analytical steps, the user states the intention in a simple way. Both, however, employ conventional GIS tools and imagery. Reasons are not given for hyperspectral spectra or for physics of spectra and neither does guarantee independent verifiability of each orchestration step. They prove that agentic orchestration does work, but do not leave hyperspectral integration and per-step verifiability open. An alternative perspective is multi-agent reinforcement learning. Carvalho and Aguiar [2] propose a decentralized approach for coverage path planning of unmanned aerial vehicles. This framework extends generalizes zero-shot in the sense that it is able to deal with larger maps and different number of agents. This is a special example of an agentic sense (not analyze) but is an important theme to understand in the context of the framework, multi-agent designs can scale and generalize across problem sizes which is exactly what property scene-scale triage would need. It is not a mathematical model that can be shared over which different analytical agents can reason about the same scene: A gap was found. Agentic remote sensing is true and effective, but has no common mathematical world-state, memorized step-by-step, verifiable artifact and is color and tool oriented. This is addressed by Objectives 1 and 2.

4. Foundation Models for Earth Observation: Perception Without Provenance

The agentic turn is concurrent with the foundation-model turn. The conceptual underpinnings of Earth foundation models were outlined by Zhu et al. [19], who looked at issues related to data size, generalization and downstream transfer. By conceptualizing Earth foundation models as a systematic object of research, rather than just a set of technical results, the following are clarified: the promise, one but multiple tasks and the unsolved issues of evaluation and dependability. From the methodology point of view, Zhou et al. [18] review the broader area of multimodal foundation models for Earth observation and propose a taxonomy for vision-language and related architectures, outlining the issues and limitations they face, while Wu et al. [15,16] introduce a multimodal foundation model for Earth observation that leverages a semantic-enhanced approach to learn across multiple sensors. Taken together, these three illustrations give a clear sense of what the state of the art is: The foundation models that are available today provide rich transferable perception across modalities and sensors, including hyperspectral and multispectral data. It should be noted that the restriction is the same for all three in the context of this framework. A model is a foundation model that does not return an auditable reason, a provenance record, or a calibrated error rate, but returns an embedding or a prediction. This will be counted in high stakes terrain review. A provably bounded error decision is not the same as an embedding which is the most likely correct embedding. The method ignores foundation models and demotes them from decision makers to tools: one perception source is one foundation model, it is wrapped by an agent and is cross-checked with the mathematical substrate and conflicting outputs down-weighted. A gap was found. While foundation models enable perception at scale, they do not offer statistical guarantees and decision provenance. This pushes the guarantee layer of Objective 3 and the consistency-checked application of foundation-model perception of Objective 2.

5. Hyperspectral Imaging: Unmixing, Classification, and the Spectral Substrate

The main data modality of the framework is hyperspectral imaging and the most distinguishing analytical challenge is unmixing or separating the materials (endmembers) and their proportions (abundances) at the pixel level. This area is assessed in two recent maps of the terrain. Ren et al. [13] review the algorithmic frameworks and applications of hyperspectral unmixing, from traditional methods to deep-learning techniques. An alternative study by Zou et al. [20] reveals the transition from traditional methods to deep learning and especially, the emergence of foundation models as a potential avenue for the future. They collectively demonstrate unmixing is well underway, an active field of research and increasingly empirical, benefiting from deep learning, while also having issues to be addressed around generalization and interpretability. The other basic hyperspectral application, classification, is moving in the same direction. Feng et al. [4] propose a hyperspectral classification method using dynamic multi-scale features and graph optimum transport. This work is doubly relevant, to this point. More important for the scope of this study, it demonstrates that optimal transport and graph structure can be directly used for hyperspectral abundance and feature distributions. It is a classification method and is a kind of the present frontier. It's designed to turn a one-off classification task into an operator consumable by agents, using the very same mathematical equipment. These studies are vital to the framework because they give and validate the world state's spectral half. Unmixing yields abundances and subspace of the spectra, the space/non-space partition of the cube is obtained, and the research on classification, e.g., Feng et al. [4] has shown that operators defined at distribution level over spectra are effective. Each is still a self-contained algorithm, thus the gap they leave is one of role rather than capabilities. None offers a breakdown as a shared persistent representation that can be consumed by a population of reasoning agents. A gap was found. Hyperspectral unmixing and classification is good and getting better, but cannot be served as a world-state, agent-consumable. This is the main idea behind Objective 1.

6. Mathematical Foundations for an Agent-Shared World-State

The mathematical foundation of the framework can be defended. This section examines the three families of techniques optimal transport, graph and diffusion and topology that the framework suggests combining and that the examined literature demonstrates to be effective on their own.

6.1. Optimal Transport as a Similarity and Change Operator

Optimal transport is a principled distance between distributions and it prescribes how one distribution is transported to the other, through the transport plan. Feng et al. [4] show how the multi-scale spatial features, in addition to the hyperspectral spectrum information comparison via transport, can be useful in a hyperspectral classifier based on graph optimum transport. The method works well, albeit it is just a part of an optimal transport in a single classification pipeline. The framework aims to be a first-class transport plan operator, a user friendly description of what and where, moved and to be reusable across epochs and regions, as a function call. The idea of the framework is to make the transport plan itself into a first class user-friendly transport operator and to be able to call it as a function between epochs and regions, so it can be reused.

6.2. Graph and Diffusion Methods for Spatial Context

In contrast to pixel-wise algorithms, graph-based learning takes into account spatial context. The sampling and multi-scale features-based object-based graph convolutional network with attention (SAGRNet) proposed by Gui et al. [5] outperforms existing graph-based baselines and U-Net++ for vegetation-cover classification. The work validates two of the assumptions of the framework: (1) the graph structure outperforms purely local models and (2) superpixel/object level graphs are a good spatial representation of remote sensing. The question remains, where does the framework go wrong? SAGRNet constructs a graph for classification; the framework constructs superpixel graphs and diffusion distances to be correlated across thousands of tiles and provides a diffusion-geometry correlation at the scene scale as a service that can be called by any agent.

6.3. Topological and Spatio-Temporal Structure

Topological data analysis is one method that is illumination and deformation independent and preserves the geometry of the data, the associated components, loops and voids. The authors Kuper et al. [7] present a four-dimensional model of data in which the geometry is also incorporated and the topology is extended to a temporal model for geometric and topological data in three dimensions. Although it's not analytical, their contribution is directly applicable on two fronts. Its temporal aspect is similar to the vision of the framework with respect to a bi-temporal reasoning ledger with a replayable data domain and links topology and geometry to the appropriate language to describe a temporal terrain. The framework is an extension of the concept of persistent homology and it allows calculation of basins and ridges as well as calculation of voids from an auditable input (e.g. the spectral indices and elevation), rather than just storing the geometry. A gap was found. While all of them are effective, none of them is exposed to an orchestrating agent and all are used separately, which is topology, graph diffusion and optimal transport. The overall goal of Objectives 1 and 3 is to integrate them in a common substrate

7. Uncertainty Quantification and Conformal Prediction for Trustworthy Triage

If its certainty can be trusted, then the triage is important and conformal prediction is the answer. Conformal methods can be applied without distributional modeling and under simple assumptions and they produce prediction sets that are distribution-free: If the error rate is specified, the true value lies in the predicted set at the aforementioned probability. Kuronen et al. [8] produce calibrated uncertainty intervals for attribute maps produced from satellite data by applying conformal prediction in conjunction with a k-nearest-neighbor approach to forest attribute mapping. Their work is the most convincing evidence in this review that conformal guarantees do apply to real-life remote-sensing estimate problems at operational size. This would be the foundation of the triage layer's statistics for the framework. It is a gap Kuronen et al. [8] leave open; a setting, not a soundness. Its conformal machinery is employed in an estimate and mapping operation, instead of in an agentic decision loop. The framework proposes spatial-aware conformal prediction as a layer on top of each agent decision to properly rank regions for human review at a user-specified coverage level, e.g., a 95% guarantee, as well as to gate outputs behind an explicit needs-human-review threshold. Conformal prediction thus turns a soft anomaly score into a conclusion that has a provable error rate, as required by survey-scale and national-security reviews. A gap was found. Conformal prediction has not yet been used (as far as we know) for gate agentic triage with coverage guarantees, but it has already been shown for remote sensing mapping. At the heart of Objective 3 is this.

8. Explainability in Remote Sensing

In high-stakes reviews, explainability is a must. In a paper published in 2025, Marjani et al. [11] show that interpretability techniques can be used for land-cover mapping and that they can substantially boost trust in the resulting product by employing explainable-AI (E-AI) systems to classify wetlands based on high-resolution imagery. As a broader overview, Kasetty and Krishnan [6] provide a meta-analysis of explainable AI and remote-sensing deep learning from a large body of literature, which discusses progress in the area of interpretability and transform-invariant modeling. In this situation, the value of the meta-analysis is that it provides an overview confirmation that explainability is a popular topic in this field and that being transform-aware (i.e., robust to rotation, scale, illumination) is an established target. These are examples of the need for explanation and the lack of the same in the current method. Most of the explainable AI that is applied to remote sensing is post-hoc, which involves adding an explanation (typically via attribution or saliency) to a black-box model after the model has determined the classification. The framework takes an alternative stance. From its inception it works it out mathematically and reason is correct-by-checking instead of explaining-in-the-afterthoughts, as opposed to explaining a black box after it has made a decision. In this way, a transform-invariance, which Kasetty and Krishnan [6] claim is an objective is achieved structurally by using the topological descriptors, which are intrinsically rotation and illumination invariant. A gap was found. Explainability in Remote Sensing is primarily often done post-hoc and linked with black-box models. Objective 2's unique purpose is inbuilt, verifiable, operation-level explanation.

9. Change Detection, Tamper, and Forgery Localization

The application of the system that is the most difficult is tamper and change review, which is technically based on change detection. Xue et al. [17] have introduced EHCTNet, a hybrid CNN + Transformer architecture designed for multi-scale sensitivity and memory to facilitate remote-sensing change detection while retaining structure in the changed areas. It's a robust example of what is considered state of the art today: Hybrid architectures have the ability to reliably identify change over scales. Most of what these detectors produce, however, is a difference mask indicating where change occurred but not semantically describing what changed nor why it is to be believed. The deficiency is even more serious in the case of tamper and forgery review, a black-box detector may calculate a plausibility score without giving an auditable reason. The mathematical underpinning studied in parts 5 and 6 is explicitly cited in the suggested solution in the framework. Optimal transport offers a transport-plan explanation of what is happening between epochs and a basis for assessing changes by significance, while topological descriptors offer illumination-invariant structure that makes tamper localization interpretable and physically justified and not opaque. Further, a physics and constraint verifier eliminates detections which are physically impossible (those that violate the radiometric, sum-to-one or non-negativity constraints), before they are passed to a human by vetoing outputs. A gap was found. Modern modification and tamper detection is accurate but black box, giving scores or masks without a physically based and understandable explanation. The essence of Objective 4 is change and tamper reasoning that is interpretable and guarantees.

10. Deep-Learning Architectures as Perception Baselines

The survey should be honest about the strength of the perception baselines it needs to stand against and against it relies upon. A number of reviewed papers describe this baseline. Liu [9]compares and contrasts different variations of U-Net for semantic segmentation of remote-sensing imagery and confirms that the encoder-decoder segmentation is still a very strong workhorse. Peter et al. [12] point out that the toolbox of deep-learning methods now includes sense modalities other than optical, describing techniques and datasets for imaging from classification to despeckling in synthetic-aperture-radar images. Two more pieces provide examples of specialized perception. To build wall-to-wall forest canopy-height maps, more accurate than the traditional baselines, Chen et al. [3] develop a multimodal model used in combination with spaceborne LiDAR, radar and optical data. Wu et al. [15,16] propose AAPW-YOLO, a YOLO-based detector adapted for detecting small objects in UAV images, which enhances detection accuracy and minimizes parameter numbers. The graph-based boundary of perception is also depicted by Gui et al. [5], discussed previously in Section 6. Together these architectures have eliminated the perceptual bottleneck. They are very precise in detecting, segmenting, fusing and mapping between different sensors. In terms of the framework, they share this common issue that has been referred to when reading this review: they are prediction engines of the perception that they manage and not responsible judgements. There is no common representation that an agent can reason about, no provenance, and no calibrated error rate. As such, it doesn't view them as end-to-end decision systems, but as tools worthy of organization, cross-verification with the mathematical substrate and a guarantee layer. A gap was found. Deep learning perception is an advanced multi-sensor and sophisticated task which produces predictions rather than conclusions that can be independently verified. This shows the need for the orchestration and guarantee levels of Objectives 2 and 3.

11. Synthesis: Gap Analysis and Objective Justification

When analysing the pillar on a pillar basis, there is a single clear conclusion. The field is now very good at perception, is getting better at orchestration across agentic systems, hyperspectral analysis, the mathematical substrate, uncertainty quantification, explainability, and change detection. It does not, however, provide in one system a shared mathematical representation, verifiable reasoning per step and a statistically guaranteed decision over hyperspectral terrain. None of the works reviewed are the whole. They add a piece. The framework takes up this area.

Table 1: Gaps and Objectives Mapping

Gap in the reviewed literature	Addressed by	Principal supporting works
Agentic RS is colour-centric; hyperspectral and spectral-physics integration remains an open problem, and there is no shared mathematical representation of the scene.	O1, O2	Mai et al. [10], Sun et al. [14], Akinboyewa et al., [1]
Hyperspectral unmixing and classification are strong but self-contained; their decompositions are not exposed as a persistent, agent-consumable world-state.	O1	Ren et al. [13], Zou et al. [20], Feng et al. [4]
Optimal transport, graph diffusion, and topology are each powerful but used in isolation, never fused under an orchestrating agent.	O1, O3	Feng et al. [4], Gui et al. [5], Kuper et al. [7]
Tool orchestration is error-prone and memory is shallow and non-replayable, agent steps are not individually checkable.	O2	Sun et al. [14], Akinboyewa et al., [1]; Carvalho and Aguiar [2]
Foundation models and deep perception yield embeddings and predictions but no decision provenance or guarantees.	O2, O3	Zhu et al. [19], Wu et al. [15,16], Zhou et al. [18]
Explainability is largely post-hoc and attached to black-box models rather than built into the reasoning	O2	Marjani et al. [11], Kasetty and Krishnan [6]
Uncertainty quantification via conformal prediction is proven for mapping but not used to gate agentic triage with coverage guarantees	O3	Kuronen et al. [8]
Change and tamper detection is accurate but opaque, producing masks or scores without interpretable, physically grounded justification	O4	Xue et al. [17]
Perception is mature and multi-sensor, but validation across layouting, survey, and tamper on open hyperspectral data, with cross-sensor generalisation, is not demonstrated for an integrated agentic system.	O4	Liu [9], Peter et al. [12], Chen et al. [3], Wu et al. [16]

12. Conclusion

This survey was intended to assist in a research program by making a comparison of the present literature with a given set of goals. Of the 20 works examined, all from 2025 to 2026, the majority are in a stage of resolving perception and rapidly moving towards the development of orchestration, but require systems to make decisions with statistical guarantees, to verify each step and to debate the shared mathematical representation. The best mathematical tools for the job are developed independently, conformal guarantees are established for mapping, but not for agentic triage, explainability is mostly post hoc, change and tamper detection works, but is opaque; foundation models see without being able to trace back to their origins; agentic systems coordinate colour tools without hyperspectral physics or auditability. The absence of an accountable decision layer on top of great perception are qualities of one gap and not discrete defects. The TerraXIA framework lies in that space. It has four goals, addressing one specific deficiency each, ranging from a common mathematical world-state to an explainable and LLM-agnostic multi-agent architecture, a statistically guaranteed triage layer and validation across layouting, survey and tamper on open data. This is why the evaluation concludes that the agent-shared mathematical world-state combined with verifiable, uncertainty-bounded triage over large terrain hyperspectral scenes is the key linchpin of the framework and not any one particular operator. Based on this, the research aims are justified and the suggested program of study is both firmly rooted in and distinct from the existing state of the art.

References

1. Akinboyewa, T., Li, Z., Ning, H., & Lessani, M. N. (2025). GIS Copilot: Towards an autonomous GIS agent for spatial analysis. International Journal of Digital Earth, 18. https://doi.org/10.1080/17538947.2025.2497489
2. Carvalho, J. P., & Aguiar, A. P. (2025). Multi-agent reinforcement learning for zero-shot coverage path planning with dynamic UAV networks. Robotics and Autonomous Systems, 195, 105163. https://doi.org/10.1016/j.robot.2025.105163
3. Chen, M., Dong, W., Yu, H., Woodhouse, I. H., Ryan, C. M., Liu, H., Georgiou, S., & Mitchard, E. T. A. (2025). Multimodal deep learning enables forest height mapping from patchy spaceborne LiDAR using SAR and optical data. International Journal of Applied Earth Observation and Geoinformation, 143, 104814. https://doi.org/10.1016/j.jag.2025.104814
4. Feng, H., et al. (2026). Flexible exploration: A hyperspectral image classification framework based on dynamic scale and graph optimal transport. Information Fusion, 128, 103962. https://doi.org/10.1016/j.inffus.2025.103962
5. Gui, B., Sam, L., Bhardwaj, A., Soto Gomez, D., Gonzalez Penaloza, F., Buchroithner, M. F., & Green, D. R. (2025). SAGRNet: A novel object-based graph convolutional neural network for diverse vegetation cover classification in remote-sensing imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 227. https://doi.org/10.1016/j.isprsjprs.2025.06.004
6. Kasetty, B., & Krishnan, S. (2026). Transform-aware deep learning and explainable AI in remote sensing: A comprehensive meta-analysis. IEEE Access, 14.
7. Kuper, P., Liu, R., & Breunig, M. (2025). A comprehensive temporal model for geometric and topological data management in 3D space. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, X-4/W6, 121-2025. https://doi.org/10.5194/isprs-annals-X-4-W6-2025-121-2025
8. Kuronen, M., Raty, J., Packalen, P., & Myllymaki, M. (2025). Uncertainty quantification for forest attribute maps with conformal prediction and k-NN method. Remote Sensing of Environment, 325, 114758. https://doi.org/10.1016/j.rse.2025.114758
9. Liu, Y. (2025). The role of U-Net variants in semantic segmentation of remote sensing images: A survey. Applied and Computational Engineering, 210(1). https://doi.org/10.54254/2755-2721/2025.LD29991

10. Mai, G., et al. (2025). Towards the next generation of Geospatial Artificial Intelligence. International Journal of Applied Earth Observation and Geoinformation, 136, 104368. https://doi.org/10.1016/j.jag.2025.104368
11. Marjani, M., Mahdianpari, M., Gill, E. W., & Mohammadimanesh, F. (2025). Explainable AI for wetland mapping using high-resolution remote sensing data. In IGARSS 2025 IEEE International Geoscience and Remote Sensing Symposium. https://doi.org/10.1109/IGARSS55030.2025.11243021
12. Peter, E., Ang, L.-M., Seng, K. P., & Srivastava, S. (2026). Recent advances in deep learning for SAR images: Overview of methods, challenges, and future directions. Sensors, 26(4), 1143. https://doi.org/10.3390/s26041143
13. Ren, L., et al. (2026). Advances in hyperspectral image unmixing: From algorithmic frameworks to practical applications. Information Geography, 2(1), 100035. https://doi.org/10.1016/j.infgeo.2025.100035
14. Sun, Z., Zhou, Y., & Yang, J. (2026). An LLM-based multi-agent system for remote sensing analysis. Big Earth Data, 10(1). https://doi.org/10.1080/20964471.2025.2600178
15. Wu, K., et al. (2025). A semantic-enhanced multi-modal remote sensing foundation model for Earth observation. Nature Machine Intelligence, 7(8). https://doi.org/10.1038/s42256-025-01078-8
16. Wu, Y., Mu, X., Shi, H., & Hou, M. (2025). An object detection model AAPW-YOLO for UAV remote sensing images based on adaptive convolution and feature fusion. Scientific Reports, 15, 16214. https://doi.org/10.1038/s41598-025-00239-4
17. Xue, C., Ye, Q., Shao, X., Hu, T., & Xu, Z. (2025). Enhanced hybrid CNN and Transformer network for remote sensing image change detection. Scientific Reports, 15. https://doi.org/10.1038/s41598-025-94544-7
18. Zhou, G., Qian, L., & Gamba, P. (2025). Advances on multimodal remote sensing foundation models for Earth observation downstream tasks: A survey. Remote Sensing, 17(21), 3532. https://doi.org/10.3390/rs17213532
19. Zhu, X. X., et al. (2026). On the foundations of Earth foundation models. Communications Earth & Environment, 7, 103. https://doi.org/10.1038/s43247-025-03127-x
20. Zou, J., Qu, H., & Zhang, P. (2025). Conventional to deep learning methods for hyperspectral unmixing: A review. Remote Sensing, 17(17), 2968. https://doi.org/10.3390/rs17172968