International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(14s) 2026 pp. 849-868
https://doi.org/10.70917/ijcisim-2026-4286
Article

Tensor World-State, Verifiable Orchestration, Topological Terrain Descriptors and Optimal-Transport Contextual Similarity

Nasrulla khan K[1], C R Sakthivel[2]

[1]Department of Computer Science, Sri Ramakrishna Mission Vidyalaya College of Arts & Science, Coimbatore, India
Email-nasrullakhank@gmail.com
[2]Department of Computer Science, Sri Ramakrishna Mission Vidyalaya College of Arts & Science, Coimbatore, India
Email-crsakthivel@rmv.ac.in

Abstract: While LLM-based tool orchestration (sometimes referred to as agentic tools) is an emerging trend in remote sensing analysis, current versions lack outputs that can be verified, low tool-orchestration error rates and any principled way to deal with hyperspectral data. The following study presents the multi-agent system TerraXIA that is used for the analysis of hyperspectral images of terrains. The language model does not analyze, it only plans and interprets, while all analysis is done by deterministic mathematical operators and only what an operator calculated can be stated by an agent. A total of four contributions has been made. First, a world-state (N1) describes the spectral state of one or many tensors with Tucker decomposition and limited unmixing is used to create a compressed, interpretable substrate, reconstruction error 2.5%, compared to ground-truth, compression by a factor 19.5, endmember spectral angle 0.149 rad and abundance RMSE 0.156. Second, a verifiable orchestration architecture (N2) where every piece of output from every operator is a content-hashed and spatially indexed artefact and every report claim has to cite an artefact. When the operator registry was frozen into the model, the LLM planner was successful at generating statically valid plans on the 20 test intents with no tool-call error and the grounding checker disambiguated another 42% of calls reported through the first pass of the model but if the operator registry was not in the model, the grounding checker would have rejected all the generated plans and the 53% of the plans would have been statically invalid with error in calls. Third, a persistent-homology terrain descriptor (N3) having all topological features remapped back to georeferenced pixel polygons and which changes only around a lower rate of 25,000 times compared to a GLCM texture baseline during rotation, flip and monotone illumination change. Fourth, an optimal-transport contextual-similarity operator (N4) that is competitive with, but not outperforming more basic spectral baselines in the current retrieval and anomaly tasks and for which the metric properties are expressed numerically with machine precision and clearly explained. All reported numbers can be reproduced from the artifact store that is hash-addressed.

Keywords: Hyperspectral imaging, agentic AI, multi-agent systems, explainable AI, optimal transport, persistent homology, tensor decomposition, verifiable orchestration, remote sensing, terrain analysis

1. Introduction

Two developments have had a significant impact on the analysis of Earth observation data. Two developments are influencing Earth observation data analysis. The promise of foundation models is pre-training on archives of satellite and hyperspectral imagery to enable the transmission of a single perception layer for many downstream applications [15, 18, 19] is the first step. The latter is the rise of agentic systems where instead of performing a single, pre-defined task, a large language model generates and orchestrates a number of specialized analytic tools [1, 14]. Remote sensing systems are able to see more because to these trends. Neither amounts to an increase in the conclusions that these systems might draw on their own. From a wider perspective, recent surveys on geospatial AI also support the above observation. Geospatial AI is evolving from individual models to interconnected reasoning systems, which is known to raise issues on autonomy that may lead to a lack of accountability [10].

The hyperspectral situation brings trust gaps to the fore. The basic concepts of the present agentic remote-sensing systems are red-green-blue and multispectral techniques, while the issues of the integration of spectral physics and hyperspectral data are still unsolved [1, 14]. There are no guarantees of reliability with orchestration itself. Agents make mistakes calling various tools on complicated tasks, remember only that an action was done, not what was actually computed and report the specific conclusions that can't be traced to any calculation that is verifiable. Foundation models make predictions and embeddings, but not provide a statistical guarantee or decision provenance [15, 18, 19]. The mathematics needed to solve this problem already exists, but it's only used in isolation. In one place, data-management models [7] are used, in another, graph diffusion in a set of classification networks [5] are used, in third, optimal transport is found in a single classification pipelines [4], in forth, tensor and block-term decomposition are available in unmixing algorithms [13, 20], in fifth persistent topology is available in data-management models [7] and in sixth conformal prediction is applied in a mapping workflow [8]. None of them are available as an analytical substrate that can be used to assess and reason with an agent.

To fill part of this, this work proposes an agent based multi-agent system called TerraXIA, a terrain hyperspectral image analysis and Triage. Each architectural decision made in the system is made under one rule which cannot be disagreed with, which is stated here, an agent can say only those things computed by a mathematical operator. Each operator output is a geographically indexed artifact that has entire provenance (operator, parameters, input hashes), a pixel and geographic extent and a numerical result. Any sentence that does not have an artifact identification reference cannot be accepted by a grounding checker before a system-generated report is issued. This design's language model is used as a planner, interpreter and reporter, but not an analyst, there was no way to have it as a math’s processor.

This paper builds upon and evaluates four research novelties with four more to follow in succeeding papers in a three-paper plan, these are the first four of twelve such novelties planned. In combination with endmember extraction and fully constrained unmixing, a Tucker decomposition of the hyperspectral cube into spatial factors, spectral subspace and core provides a compressed hyperspectral representation that is both interpretable and physically meaningful that the system can use to generate its labels. N1 is a world-state that is a tensor. An agentic orchestration, or N2, is a plan-verify-execute-ground cycle where the plans are typed, each artifact is physics and schema verified before the plan is ground and each report claim is re-triggered on the artifact store. N3 is a persistent-homology terrain descriptor that is calculated using cubical complexes of spectral-index and elevation rasters. Each topological feature needs to be plotted as a georeferenced polygon and keep the Pixel Coordinates. First-class, citable outputs of N4 are transport plans and Wasserstein barycenters, which are entropic Sinkhorn distances between per-region abundance distributions.

The assessment is designed to be multi-layered, on purpose. Whatever dataset, a suite of 25 property checks shows a consistent accuracy of the operators. Topological invariance under rotation and monotone illumination is accurate not approximate, metric axioms of the transport distance are accurate to about 1e-15, reconstruction errors declared by the operators match recalculation errors to 1e-13. Then, a number of four trials (E1, E2, E3, E4) on Pavia University, Jasper Ridge and AVIRIS Cuprite are conducted using a CPU-only laptop to determine their usefulness. The pre-registered success criteria are met or exceeded by three of the four novelties. The fourth, N4, is reported honestly, it is competitive and is reported to be equal or better than the others on the current retrieval and anomaly benchmarks and Section 4.7 looks at why the benchmark design is the one that produces this qualitative result. The operator is used to verify if the problem is correct, it is the only way among the other comparison techniques that provides an explanation of how regions are different.

This is how the remainder of the paper is structured. The content of the years 2025 and 2026 that is the catalyst for its design is considered in Section 2. These are defined in Section 3 and explained in the architecture, artifact memory, agent framework, four operator families and experimental protocols. In Section 4, an end-to-end case study based on a real language model is presented, as are the four numerical experiments, the validation of the mathematics is open, meaning that the reader may find other ways of defining the metrics. This concludes up section 5 and makes suggestions for the next two articles.

2. Literature Review

Twenty peer-reviewed papers from 2025 and 2026 were chosen as representative methodological findings or authoritative reviews in the fields on which the system is based. Hyperspectral unmixing and classification, foundation models for Earth observation, geospatial and agentic artificial intelligence, the mathematical tools that form the substrate, uncertainty quantification, explainability and change detection. Regarding verifiable agentic hyperspectral analysis, each paragraph outlines what the referenced study establishes and what it leaves open.

2.1 Geospatial AI and agentic remote sensing

In the next generation of geospatial AI, knowledge-guided and more autonomous are the significant distinctions according to Mai et al. [10]. What is important about their review here is their framing, the discipline shifts from isolated models to integrated systems capable of reasoning and does not matter here as much as the fact that they state a danger: autonomy without accountability, the failure mode for which this paper's verification machinery is designed. The agentic turn is materialized via the two works. Sun et al. [14] proposed a large-language-model multi-agent system for remote-sensing analysis, where agents automatically perform analytical tasks using NLP and Akinboyewa et al. [1] proposed a natural-language assistant that is integrated with GIS. Both illustrate the power of the paradigm, giving the user the intention and the system piecing together the steps in an analysis. Both however run on the traditional GIS tools and color/ multispectral imagery. Whilst an explanation of hyperspectral spectra or the physics of spectrum does not guarantee the correctness of every orchestration step, it does not explain it either. Carvalho and Aguiar [2] use the reinforcement learning theory to take a multi-agent design approach and demonstrate the problem of zero-shot generalization in decentralized UAV coverage planning for both map and team size. They never analyze as agents but their result is relevant. Multi-agent designs scale with problem size, the property scene-scale triage needs. None of these systems offer a common mathematical model on which multiple analytical agents can operate, over which they reason about the same scene and a step-by-step evidence-based model of how to get from one step to the next. Novelties N1 and N2 are motivated by the absence. The system assembles analytical stages in both, thus demonstrating the elegance of the paradigm, the system assembled analysis stages whenever a purpose was expressed. Both, however, deal with color or multi-spectral images and classic GIS methods. Neither assumes anything about the spectral physics or hyperspectral spectra and neither guarantees that any orchestration step is easily verifiable. However, Carvalho and Aguiar [2] successfully showcase multi-agent design for zero-shot generalization in decentralized UAV coverage planning, across team sizes and map sizes, by using reinforcement learning.

2.2 Foundation models for Earth observation

Zhu et al. [19] discuss the concepts behind earth models with the Earth as a coherent academic field of study, noting the known challenges of reliability and evaluation. For foundation models development across the different sensors, Wu et al. [15] introduce a semantic-augmented multimodal foundation model, while the development for downstream applications on Earth observation is assessed and a taxonomy of vision-language-based and related model architectures is suggested by Zhou et al. [18]. Together all these enable today's state-of-the-art. Foundation models offer today a rich percep of modalities like multispectral and hyperspectral data, with transferable components. All three have this limitation when considered from the high stakes viewpoint of terrain review. What a foundation model lacks is a record of provenance or an explanation that is auditable, or a calibrated rate of error, what it gives you is an embedding or a forecast. It is not the same as the embedding that is most likely correct, which can also be achieved with a choice that has a demonstrated bound on the error. For subsequent articles of this program a foundation model is wrapped with an agent as one perception source and cross checked with the mathematical substrate. For this reason, TerraXIA sees foundation models as an instrument, not a decision maker. There is a perceived gap between artifacts and the lack of perception without provenance, which is addressed in this work, the provenance first artifact design of N2 is motivated.

2.3 Hyperspectral unmixing and classification

Data modality for the data is the analytical basis of hyperspectral unmixing that breaks materials (endmembers) and per-pixel proportions (abundances). Zou et al. highlight the transition from traditional methods towards deep learning and foundation models [20] while Ren et al. explore the shift from algorithmic unmixing to applications [13]. The field in both papers is a dynamic and increasingly empirical one and there remain issues in both that have yet to be resolved regarding meaning and extensibility. As for the classification method, Feng et al. [4] proposed a dynamic multi-scale features classification framework with graph optimal transport. Where there is effectiveness in several aspects, more specifically, this structure is good for the two contexts in which it is applicable, in showing the effectiveness of the transport-based comparison of hyperspectral feature distributions and in representing the current structure of the frontier, which is composed of a single strong pipeline for a single task. The U-Net variants for semantic segmentation [9] and deep learning approaches for SAR-imagery [12] complete the picture, as do object detection, especially tuned for UAV-imagery [16] and multimodal fusion with regards to forest structure [3] and hybrid CNN-Transformer networks for change detection [17] for the target function. None of these studies offer its decomposition nor do they offer any properties with which a population of reasoning agents may consume it and cite it as a durable, shared representation. There is a lack of properties, rather than capability, in these studies that would

allow a population of reasoning agents to consume it and cite it as a durable, shared representation. The intent of N1 is to convert unmixing output into something like that.

2.4 Mathematical foundations for an agent-shared world-state

In the recent literature there are three major families of mathematics mentioned as being unintegrated and individually effective. Three are the most common families of mathematics found in the recent literature that are individually effective and collectively unintegrated. Albeit in a different way, Feng et al. [4] offer a principled distance between distributions based on optimal transport inside a classifier, with the plan being the internal machinery of the classifier. The plan explains this and TerraXIA makes its plan a first-class operator output which can be called from any two regions and can be referenced in reports. To prove the benefit of graph structure, as well as the level of representation using super pixels, they use SAGRNet and show that object level graphs on remote-sensing imagery outperform purely local baselines in the task of classification for vegetation cover, as demonstrated in [5]. Graph methods capture space beyond models at the pixel level. TerraXIA produces superpixel graphs and diffusion distances as a context service for any agent, instead of native entities of a classifier. In the desire to combine topology and geometry for apposite description of terrain, Kuper et al. [7] extend the topological data management to a temporal model in three-dimensional space. Unlike other applications, it is necessary to geolocate each topological feature back to its producing pixels in the case of computing permanent homology over auditable rasters (spectral indices, elevation). What is the fault of all three families, is that they are all used separately in their respective pipelines without an orchestration agent. N1, N3 and N4 act together to bring them together on a common substrate.

2.5 Uncertainty quantification, explainability and change detection

Kuronen et al. [8] develop distribution-free intervals for calibration at operational scale by applying conformal prediction and k-nearest-neighbor method to forest attributes. This is the best recent evidence that conformal guarantees actually are valid for practical remote-sensing estimation problems and it provides statistical underpinning for the triage layer envisioned in this program's second paper. The author suggests in the present paper that, in order to avoid the redesign of the conformity gate, it is important to keep all scores numerical and auditable. Using high-resolution data, Marjani et al. [11] show that interpretability techniques can greatly enhance confidence for wetland mapping. The importance of explainability and the desired feature of transform-awareness (stable under rotation, scale and illumination) have been demonstrated in the remote sensing context through a meta-analysis presented by Kasetty and Krishnan [6]. Most works in the surveyed literature provide post-hoc explanations, i.e. an explanation is generated after a decision is taken by a black-box system. A structural method of explanation is instead used with TerraXIA where each claim starts from artifacts mentioned. Finally, the perception baseline for the bi-temporal reasoning meant in research 3 is determined by the work on change detection such as Xue et al. [17], change detection is not covered among this research and only the hooks (content addressing, lineage, timestamps) are created.

2.6 Synthesis of gaps

There are four gaps that have been identified in the literature evaluated. The first is, although agentic remote sensing is effective, it is not step-by-step verifiable, not orchestration and not color-centric [1, 10, 14]. Second, there is no statistical guarantee or provenance in the form of foundation models [15, 18, 19]. Third, all of the mathematics and mathematical ideas are separated from each other and separate from the architecture [4, 5, 7, 8, 13, 20]. Fourth, there is a need for explainability, which is offered in post construction [6, 11]. The substrate gap is closed via a common mathematical world-state (N1), the reliability and provenance gaps are closed via verifiable orchestration (N2), the two isolated mathematical families are brought into a common substrate via georeferenced, explainable outputs of the mathematical action (N3) and transport (N4) operators

3. Methodology

3.1 System overview and the governing rule

There are several multi-agent systems and the most significant one, called TerraXIA, has the language model that the language model interprets the user's language and reports results, but it does not directly analyze the hyperspectral pixels. Never pixels are retrieved by the model, only compact artifacts in textual format, all analysis is computed by deterministic mathematical operators over a common world-state built from the hyperspectral cube data. Each call to an operator returns a spatially indexed artifact which gets a truncated hash, mapping to its content, each claim in a report includes an artifact identifier or the claim is not accepted. According to the controlling principle these problems can be only said what the agent computed, which is followed by the construction through. There are three

qualities that come along with this. The architecture is verifiable per step (each artifact is schema checked and physics checked before going to the next step in the planner), plottable (the planner has an extent (in pixels) and a geographic transform (including critical velocity data) for each artifact) and LLM-agnostic (the planner produces a JSON operator DAG that works with any model or a symbolic fallback that is without a model).

3.2 Layered architecture and data flow

The directive dependence direction of all layers is L1 data which depends on L2 mathematical substrate (pure deterministic operator for N1, N3, N4, graph, physics checks), which in turn depends on L3 artifact memory, which depends on L4 agents (planner, executor, verifier, analyst, labeler, cartographer, reporter, grounding checker, feedback integrator), which finally depends on L5 human review. Nothing else is called up by L2, L4 never touches pixels and only reads data from artifact cards, calling L2 with an operator registry through a typed operator type. Topological execution ordering, per-node artifact check, analyst findings, labels, map layers, and the per-claim grounder checker check a report claim by claim, one run to planner for a plan DAG. A node that fails does not halt the run, but rather creates an error artifact, and ignores dependencies on it.

3.3 Artifact memory

Type, payload path and digest, georeference (CRS, geographic transform, scene-coordinate pixel extent), summary statistics, verification results, a deterministic text card with no more than 120 words and provenance are all recorded for each artifact. Because provenance takes part in the hash, the identifier construction is exclusive to TerraXIA.

$$id(a) = \text{SHA-256}(\text{payload}(a) \,\|\, prov(a)) \qquad (1)$$

where the operator, version, seeded parameters and consumed artifact IDs are canonicalized by prov(a). By (1), the digest is broken by any tampering, identical computations cache and precise replay. The correction-to-feedback-to-finding-to-mathematics chain is walkable from beginning to end because findings, labels, feedback and corrections are first-class artifacts. Only a logged value-lookup pinhole that returns single referenced values, never arrays, allows agents to query cards by space, lineage, category or text.

3.4 Agents and contracts

Few agents, rigid contracts, code agents are deterministic, model-backed agents are stateless with all state in the store, every message is a JSON envelope and no agent parses the prose of another agent. In order to prevent tool hallucination, the registry serializes itself into the planner context and serves as the only source of truth for what exists. It purposefully lacks any tool that would allow the model to perform mathematical operations, maintaining the plan-verify separation E4 measures. Results must distinguish between observation and interpretation and contain resolvable value queries. A misinterpreted command can reject a finding but never rewrite mathematics since human judgments (accept, reject, amend) become feedback and corrective artifacts.

3.5 N1: tensor world-state and unmixing

For a cube X in R^ (H x W x B), the world-state is the Tucker decomposition

$$X \approx G \times_1 U \times_2 V \times_3 S \qquad (2)$$

with spatial factors U in R^ (H x R1), V in R^ (W x R2) (the space component), orthonormal spectral factor S in R^ (B x R3) (the non-space subspace), and core G. The artifact records

$$\rho = HWB \,/\, (\mathrm{HR}_1 + WR_2 + BR_3 + R_1R_2R_3), \quad \varepsilon_{rel} = \| X - \hat{X} \|_F \,/\, \| X \|_F \qquad (3)$$

plus, a per-pixel error map, itself an anomaly cue: large localized residual means the world-state does not explain that pixel. The unmixing-aligned block-term (LL1) form, with minimum-volume and total-variation regularization for identifiability, is

$$X \approx \Sigma_r \, (A_r \, B_r^T) \circ c_r \,, \quad r = 1, \ldots, P \quad (4)$$

The extraction step is the formulation unique to TerraXIA. Endmembers are not extracted from the raw cube nor from the low-rank reconstruction (which smooths away the near-pure simplex vertices an extractor needs)

but from the spectral-subspace projection, as the maximum-volume simplex over K restarts of vertex component analysis.

$$\tilde{Y} = S\,S^T\,Y, \quad E = \mathrm{argmax}_{k\,=\,1,\,\ldots,\,K}\ \mathrm{vol}(\mathrm{VCAk}(\,\tilde{Y}\,)) \qquad (5)$$

Section 4.2 shows (5) were decisive. Abundances follow per pixel by fully constrained least squares on the R3-dimensional subspace,

$$\hat{a}(y) = \mathrm{argmin}_{a\,\geq\,0,\ 1^T a\,=\,1}\ \|\,\tilde{y} - E\,a\,\|_2^2 \quad (6)$$

whose constraints are exactly the physics checking the verifier re-executes on the output. Registered spectral-index operators (NDVI, NDWI, band ratios) feed the filtrations of Section 3.7 with uniform provenance.

3.6 N4: optimal-transport contextual similarity

Each superpixel region r yields an abundance distribution mu_r (histogram over the P-simplex, ground metric M declared and recorded in provenance). The operator computes entropic optimal transport and reports its debiased divergence.

$$W_\varepsilon(\,\mu_i, \mu_j\,) = \min_{T\,\in\,U(\mu_i,\,\mu_j)}\ \langle\ T, M\ \rangle + \varepsilon\,H(T) \quad (7)$$

$$S_\varepsilon(\,\mu_i, \mu_j\,) = W_\varepsilon(\,\mu_i, \mu_j\,) - \tfrac{1}{2}\,W_\varepsilon(\,\mu_i, \mu_i\,) - \tfrac{1}{2}\,W_\varepsilon(\,\mu_j, \mu_j\,) \quad (8)$$

over the full R x R region matrix (R of 200 to 600 per tile, seconds to low minutes on CPU). Section 4.1 verifies that (8) ranks distributions identically to exact Wasserstein distance (Spearman 0.99999), licensing the fast form. The contextual-anomaly construction unique to TeraXIA scores each region by standardized divergence from the Wasserstein barycenter of its context.

$$\bar{\mu} = \mathrm{argmin}_v\ \Sigma_r\ S_\varepsilon(\,v, \mu_r\,)\ , \quad z_r = (\ S_\varepsilon(\,\mu_r, \bar{\mu}\,) - m\ )\ /\ s \quad (9)$$

with m and s the scene mean and standard deviation of the divergences, z_r is the number the analyst may cite for claims of contextual anomaly and in Section 4.6 it is the number that let the model decline to claim one. The transport plan T from (7) is stored as a first-class artifact and rendered as source-to-sink flow arrows, the plan, not a scalar, is the explanation of what differs. The same divergence weights the superpixel graph,

$$w_{ij} = \exp(\,-\,S_\varepsilon(\,\mu_i, \mu_j\,)\ /\ \sigma\,) \quad (10)$$

whose Laplacian yields a diffusion-distance context feature. One operator, three uses: (8) drives retrieval, (9) contextual anomaly, (10) the context embedding.

3.7 N3: persistent-homology terrain descriptor

For a raster f (spectral index or surface model) on its cubical complex, sublevel persistence tracks H0 (basins) and H1 (loops) over

$$F_t = \{\ p : f(p) \leq t\ \}\ , \quad D(f) = \{\ (\ b_i, d_i\ )\ \} \qquad (11)$$

Diagrams are the standard output of topological data analysis, the requirement unique to TeraXIA is that every diagram point be geolocatable. Retaining the critical cells the pixel support of an H0 feature is recovered by flood fill from its birth pixel,

$$G_i = \mathrm{FloodFill}(\ p_i^{\,birth}\ ;\ \{\ p : f(p) \leq d_i\ \}\ ) \qquad (12)$$

(H1 via the bounding contour just below death), turning each diagram point into a GeoJSON polygon, georeferenced topology. The fill is applied to the top-k most persistent features, which Section 4.4 shows cuts cost from minutes to 0.11 s without changing the diagrams. Before persistence, the raster is rank-normalized, the second formulation specific to this system:

$$\tilde{f}(p) = |\ \{\ q : f(q) \leq f(p)\ \}\ |\ /\ N\ , \quad D(\ g \circ f\ ) = D(\,\tilde{f}\,)\ \text{ for monotone } g \qquad (13)$$

Because monotone maps preserve the filtration order of (11), (13) makes illumination invariance exact rather than approximate, rotation and flip invariance is automatic since they permute pixels without changing values or connectivity. Diagrams are vectorized as persistence images PI(f) normalized intrinsically by the shared filtration range and stability under a transform T is the drift

$$\delta(T) = \| PI(f) - PI(T f) \|_2 \quad (14)$$

3.8 N2: language-model integration and grounding

The reference brain is the Claude Sonnet model behind a thin LLM-agnostic interface, planner calls operate at temperature zero, all model I/O is logged as an audit trail without ever being evidence and every call must parse into a typed schema (one repair round, then hard failure to the symbolic route). Value lookup, card query and registry description are the model's three read-only tools. One boundary finding is reported for reuse, transporting arguments as an explicit JSON string increased first-shot plan validity from 0% to 100%, stringent structured outputs discreetly discarded free-form dictionary fields, emptying every operator argument. The governing rule has a formal statement, a report is a claim set C, each claim c carrying an evidence set E(c) of artifact-query pairs, and

$$\text{admissible}(c) \iff \forall\, (a, q) \in E(c) :\ \text{resolve}(a, q) \neq \perp \ \wedge\ |\, v(c) - \text{resolve}(a, q)\,| \leq \tau \quad (15)$$

Claims failing (15) get at most two revision rounds, then are dropped and logged, so delivered reports satisfy (15) by construction; the measured quantity is therefore the first-pass grounding rate

$$\Gamma = |\,\{\, c \in C : \text{admissible}(c)\,\}\,| \,/\, |\,C\,| \quad (16)$$

evaluated before revision, the post-revision rate describes the mechanism, (16) measures the model. This admissibility formulation, with resolve implemented as the logged pinhole of Section 3.3, is unique to TerraXIA among the systems reviewed in Section 2.

3.9 System-generated labels and plotting

The model uses a fixed taxonomy to name labels that originate from mathematics, geometry is derived from thresholded abundance maps (material labels), persistence generator polygons via (12) (terrain-structure labels), or superpixel unions named in findings (contextual-anomaly labels). These labels are then cleaned morphologically and stored as GeoJSON in scene coordinates with complete provenance. The supporting data (z-score, persistence, abundance purity, physics verdict) are used to record confidence, a calibrated probability is purposefully postponed to Paper 2's conformal layer. The labeler is never fed by ground truth it is only used for evaluation. Every artifact type is rendered to 300-dpi PNG plus GeoTIFF or GeoJSON using a single cartographer code route, allowing results to be viewed in normal GIS tools and creating composite maps. artifacts, each figure in this work has a hash address and provenance that extends to the figure level.

Experimental design

Table 1. Datasets - All full-scale runs in Section 4 were executed on Pavia University, Jasper Ridge and AVIRIS Cuprite on a CPU-only laptop

Dataset	Bands	Role	Used in
Pavia University	103	Urban benchmark; region retrieval and self-label evaluation	Validation, E1, E2
Jasper Ridge	198	Abundance and endmember ground truth for unmixing	E1, E4 case study
AVIRIS Cuprite	224	Natural mineral terrain, topological invariance	E1, E3
Houston 2013/2018	144 / 48	Urban flagship and tiled scale story (planned extension)	Protocol E1 to E4

Four pre-registered experiments, one per novelty. E1: Tucker rank sweep (error against compression via (3)), unmixing against Jasper ground truth by endmember spectral angle and abundance error,

$$\text{SAD}(e, \hat{e}) = \arccos\left(e^T \hat{e} \,/\, \|e\|\|\hat{e}\|\right), \quad \text{RMSE} = \left(\|A - \hat{A}\|_F^2 \,/\, HWP\right)^{1/2} \quad (17)$$

and unsupervised self-labels against a supervised SVM-on-PCA baseline (overall accuracy, macro F1, kappa), bar: error at or below 8% at 20x compression or better. E2: same-class region retrieval precision at k for (8) against Euclidean and spectral-angle distances on mean spectra and chi-square on abundance histograms, plus implant anomaly ROC-AUC for (9) against the RX detector, bar: transport beats the non-transport baselines. E3: drift (14)

under rotations, flips, gamma 0.6 to 1.6, and affine illumination, against GLCM texture features, bar: at least 5x lower drift. E4: a fixed 20-intent suite through four systems (A: full system with the model planner, B: symbolic planner, the LLM-agnosticism control, C: naive foil with tool descriptions but no registry grounding or verification, D: model planner with verification ablated), measuring plan validity, first-shot validity, tool-call error rate, node success rate and (16). Every number is regenerable from a versioned run directory; all timings are cold-cache on the same laptop.

4. Result Analysis

The evidence is offered as follows: Section 4.1 shows that the operators compute what they say, independent of any dataset. Sections 4.2 to 4.5 report the four tests; Section 4.6 leads one through the live system and Section 4.7 outlines the restrictions.

4.1 Mathematical validation of the operator core

All 25 property checks are successful, compared with both closed-form and synthetic ground truth (Table 2, Fig. 1). All metric axioms of the transport distance are numerically fulfilled, values for the known-topology rasters of basin and loop count are all correct and the stability of the displayed transport distance under input perturbation (5e-15) is comparable to the exact Wasserstein distance (Spearman 0.99999) for N4, for N3 this value is 5e-5 compared to 1e-9 for the input perturbation.

Table 2. Mathematical validation suite - 25 of 25 property checks pass against closed-form or synthetic ground truth, independent of any dataset

Operator family	Checks	Representative verified properties
N1 world-state and unmixing	8 / 8	Rank-monotone error, spectral orthonormality 3e-8, declared vs recomputed error match 1e-13, synthetic recovery SAD 0.008 rad, RMSE 0.005, FCLS sum-to-one 1e-7, exact non-negativity
N4 optimal transport	9 / 9	Identity, symmetry (3e-15), non-negativity, triangle inequality (9e-15), exact line isometry; Sinkhorn ranks identical to exact Wasserstein (Spearman 0.99999), barycenter consistency
N3 persistent homology	8 / 8	Correct basin and loop counts on known-topology rasters; birth not exceeding death, perturbation stability 5e-5, exact invariance to rotation, flip, monotone illumination

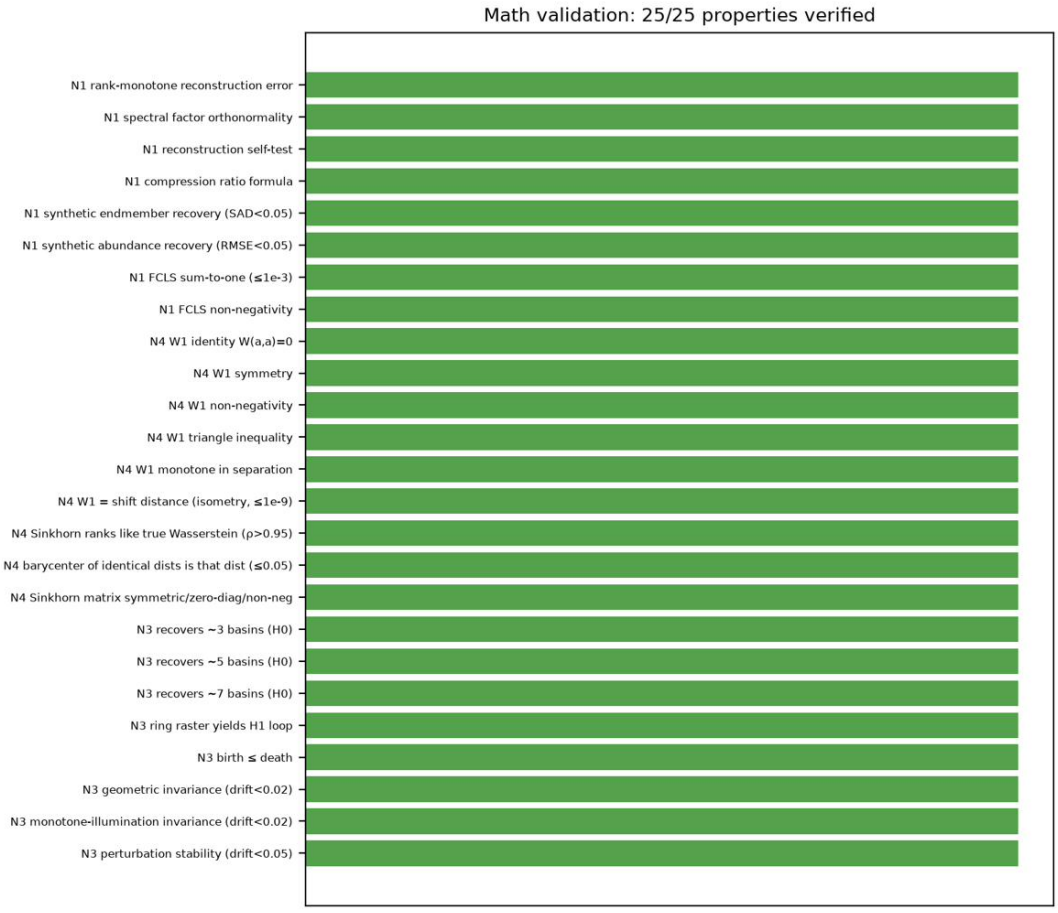

Fig. 1. Validation dashboard generated by the property-check suite - all 25 checks across the three operator families pass against known ground truth

4.2 E1: world-state fidelity, compression, and self-labelling (N1)

A smooth, monotone error-compression Pareto front is produced by the Jasper rank sweep (Table 3, Fig. 2), 2.5% relative error at 19.5x compression at ranks (65, 65, 20), 6.2% at 93x, and 10.4% at 198x. Rank monotonicity is an accuracy indication in and of itself and the bar (8% at 20x or better) is met with margin

Table 3. E1 rank sweep on Jasper Ridge (selected settings) - relative reconstruction error against compression ratio

Tucker ranks (R1, R2, R3)	Relative error	Compression ratio
(20, 20, 10)	10.45%	198.4x
(35, 35, 10)	6.19%	93.3x
(50, 50, 20)	3.76%	31.0x
(65, 65, 15)	2.53%	25.0x
(65, 65, 20)	2.48%	19.5x

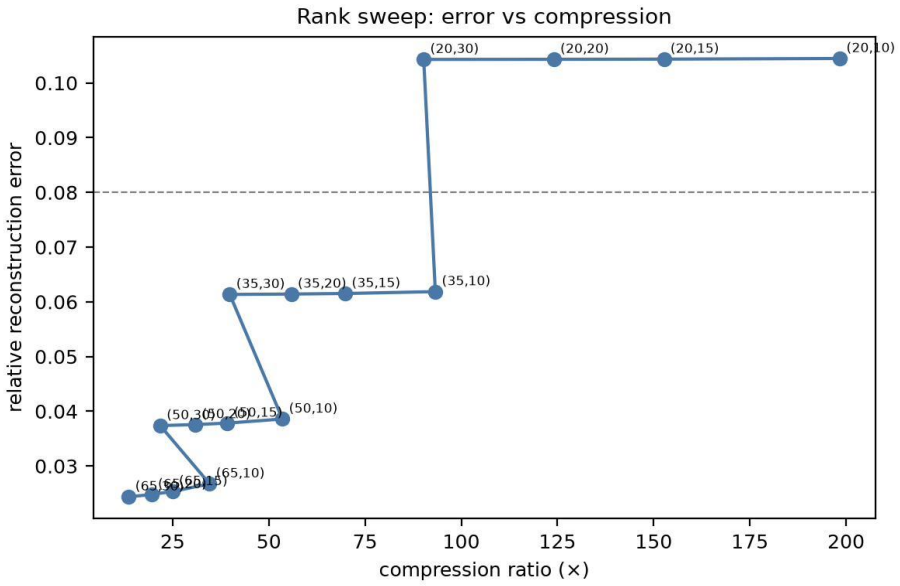

Fig. 2. E1 fidelity-compression Pareto front over the full rank sweep. The dashed lines mark the pre-registered success bar (8% error at 20x compression), which the front clears with margin

In comparison to published VCA and NMF-family results, unmixing against the Jasper ground truth under (17) yields mean SAD 0.149 rad (8.5 degrees) and abundance RMSE 0.156 (Table 4, Fig. 3), the dark water endmember, a recognized hard case, is the per-endmember outlier (0.268 rad). A synthetic recovery test identified the smoothing in the reconstruction as the cause (solver correct at RMSE 0.0016 with true endmembers, extractor correct at SAD 0.009 on the raw cube, extractor collapsed to SAD 0.21 on the reconstruction). These figures were dependent on the formulation (5). Extracting from the low-rank reconstruction instead produced RMSE 0.46, worse than random. Correcting the ground truth's MATLAB column-major ordering eliminated a silent transposition, and the multi-restart maximum-volume criterion in (5) recovers the dark endmember that a single run misses, collectively, SAD 0.33 to 0.149 rad and RMSE 0.46 to 0.156.

Table 4. E1 unmixing validation on Jasper Ridge against the published endmember and abundance ground truth

Quantity (Jasper Ridge, 4 endmembers)	**Value**
Mean endmember SAD	0.149 rad (8.5 deg)
Per-endmember SAD	0.099, 0.268, 0.103, 0.126 rad
Abundance RMSE vs ground truth	0.156
World-state relative error at unmixing ranks	7.3%

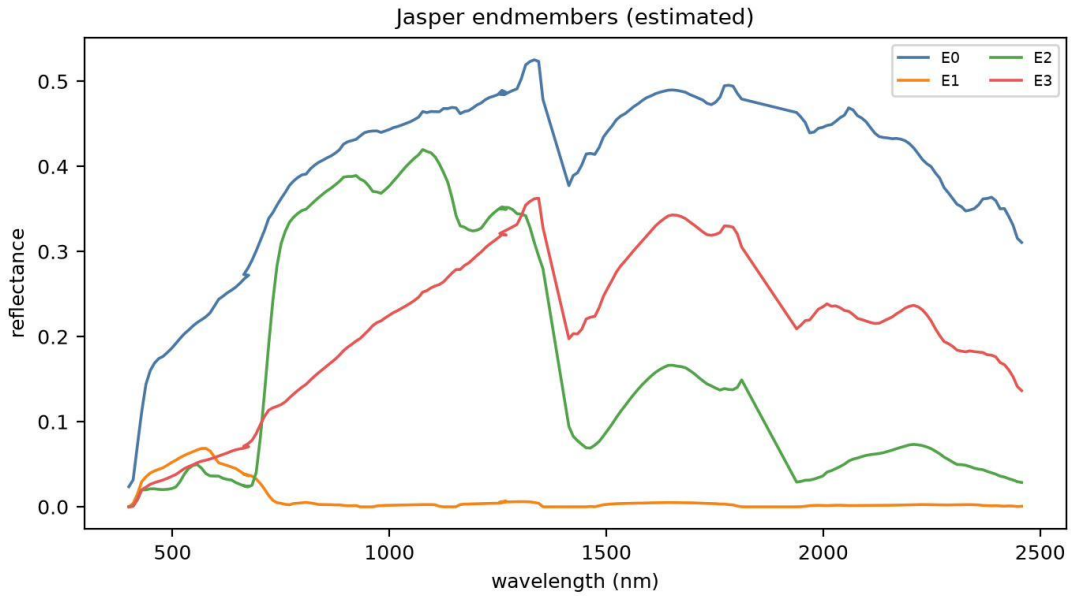

Fig. 3. Recovered endmember spectra (solid) against Jasper Ridge ground truth (dashed) after subspace-projected, multi-restart extraction

Result values without spin are given, unsupervised argmax labels, matched using Hungarian algorithm to nine Pavia classes, have an OA of 0.454, macro F1 of 0.148 and a kappa of 0.043 for 42,776 pixels, while the supervised SVM on PCA has an OA of 0.913, macro F1 of 0.893 and a kappa of 0.882 for the same number of pixels (Table 5, Fig. 4). It is known that the comparison is not symmetrical (unsupervised, 9 classes vs. 5 components) and this is confirmed by the results, which state that the labels are not state-of-the-art classification but rather adequate for the triage from an interpretable substrate, for each of them there is the numerical foundation and the provenance. Leftover memory T1 compute: 1.22 GB, peak memory T1 compute: 1.22 GB, CPU only, rank sweep T1: 355 s, unmixing T1: 24 s, self-label evaluation T1: 232 s.

Table 5. E1 self-label quality on Pavia University against a supervised baseline. The self-labels are unsupervised and map five abundance components onto nine ground-truth classes

System	Overall accuracy	Macro F1	Kappa
Self-labels (unsupervised, from abundances)	0.454	0.148	0.043
SVM on PCA (supervised baseline)	0.913	0.893	0.882

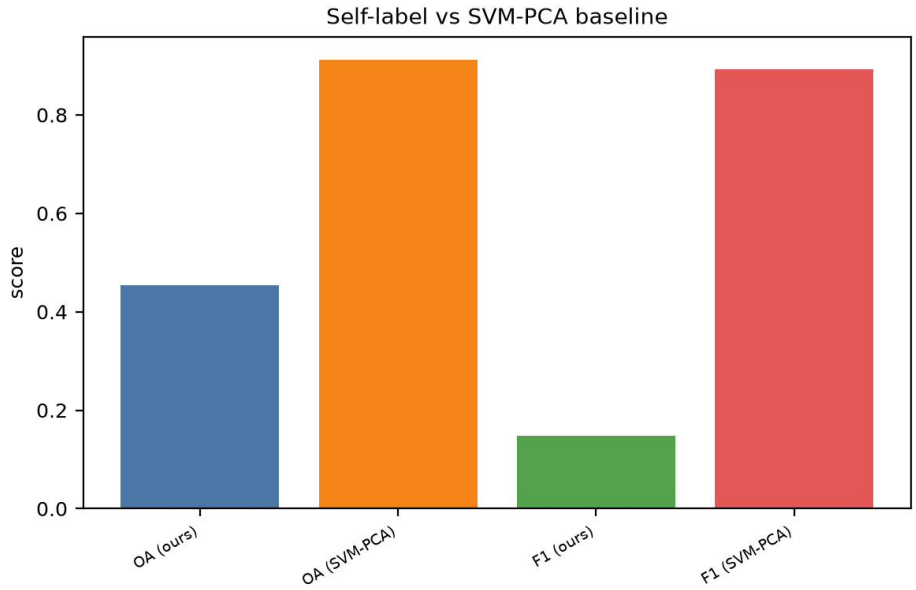

Fig. 4. Self-label metrics against the supervised SVM-PCA baseline on Pavia University. The gap is expected and stated: the self-labels are unsupervised triage labels from an interpretable substrate, not a trained classifier

4.3 E2: optimal-transport contextual similarity (N4)

In Section 4.1, the operator's accuracy is confirmed, on the application tasks, it is competitive but not dominant. Retrieval across 211 designated Pavia regions, the divergence (8) approaches p@3 0.400 and p@5 0.397 against 0.415/0.403 (Euclidean on mean spectra), 0.461/0.428 (spectral angle) and 0.449/0.445 (chi-square on histograms) (Table 6, Fig. 5). Anomaly over eight mineral implants: AUC 0.674 is reached by the barycenter score (9) compared to 0.929 for RX (Table 7, Fig. 6).

Table 6. E2 same-class region retrieval on Pavia University (211 labelled regions): precision at k for the transport distance against three baselines

Method	**p@3**	**p@5**
Optimal transport (Sinkhorn, abundance distributions)	0.400	0.397
Euclidean on mean spectrum	0.415	0.403
Spectral angle (SAM) on mean spectrum	0.461	0.428
Chi-square on abundance histograms	0.449	0.445

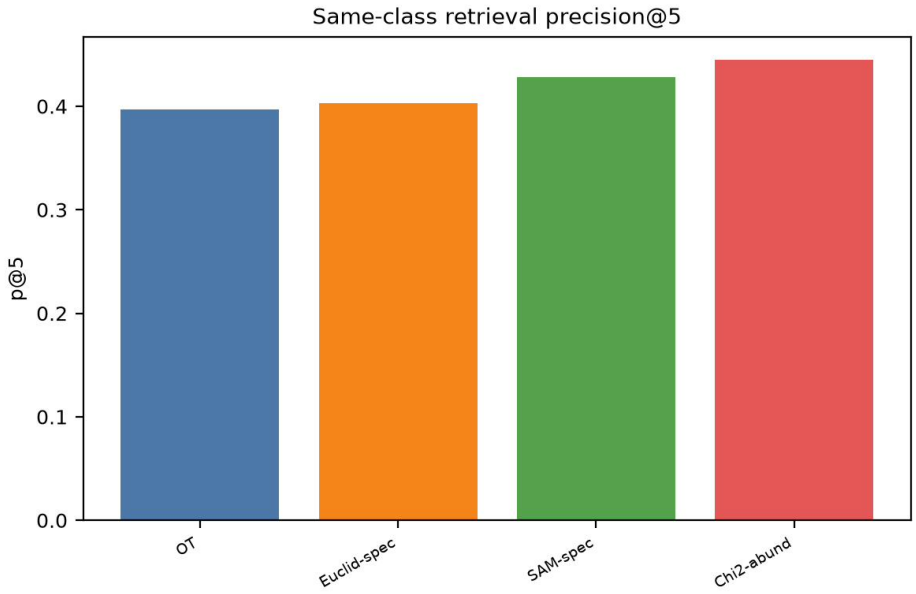

Fig. 5. E2 retrieval precision at 5 by method. The transport distance is competitive with, and does not surpass, mean-spectrum and histogram baselines on this task

Table 7. E2 anomaly detection on eight implanted mineral patches. RX is optimal for global spectral outliers, which is exactly what the implants are; the barycenter score answers a contextual question the benchmark does not pose.

Detector	ROC AUC
Optimal-transport distance to context barycenter	0.674
RX detector (global Mahalanobis)	0.929

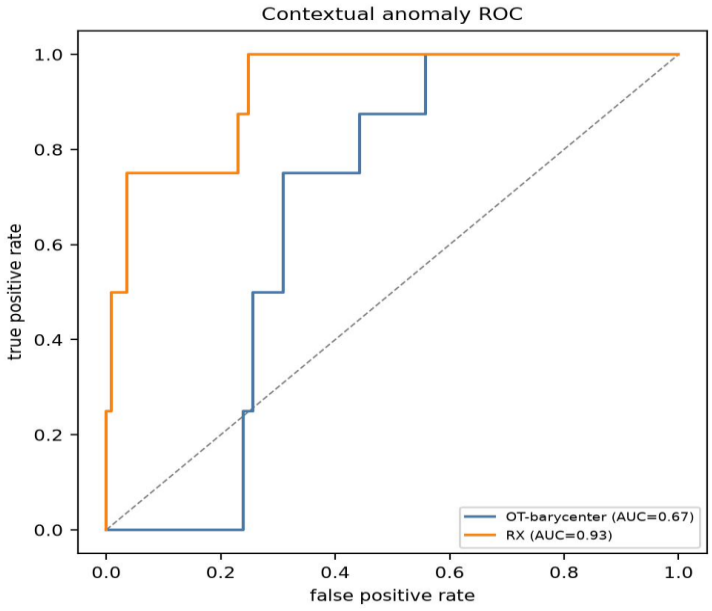

Fig. 6. E2 ROC curves for implant detection. The implants are global spectral outliers, the hypothesis class under which the RX detector is optimal

It is of a structural nature. The shape of same-class retrieval is not significantly altered since the abundance distribution within the regions describing the mean features is highly peaked for the baselines, but not for (8) which receives, instead, a five-dimensional histogram, so that the information is carried by the representation, not the metric. The implantation of the solution to a context question is a spectral outlier from the benchmark, which is operated in a totally different context. All approaches suffered from one recorded experimental blip, when setting the number of superpixels in the whole image to just 150 which corresponds to approximately 1,400 mixed-class pixels per region, those precisions were lowered, but subsequently they were restored by lowering the number of superpixels to roughly 400 for the whole image. Only an explanation in the comparison was returned - the transport plan of (7) told what mass moved where (Fig. 7) and every baseline returned a scalar, instead of one ad hoc score, one provably correct metric (Section 4.1) as opposed to one (ad hoc) score, one operator serving retrieval (8), anomaly (9), and context (10) as opposed to one (ad hoc) operator. In Section 4.7 we discuss the modification of the experiment required to give these features quantitative manifestation.

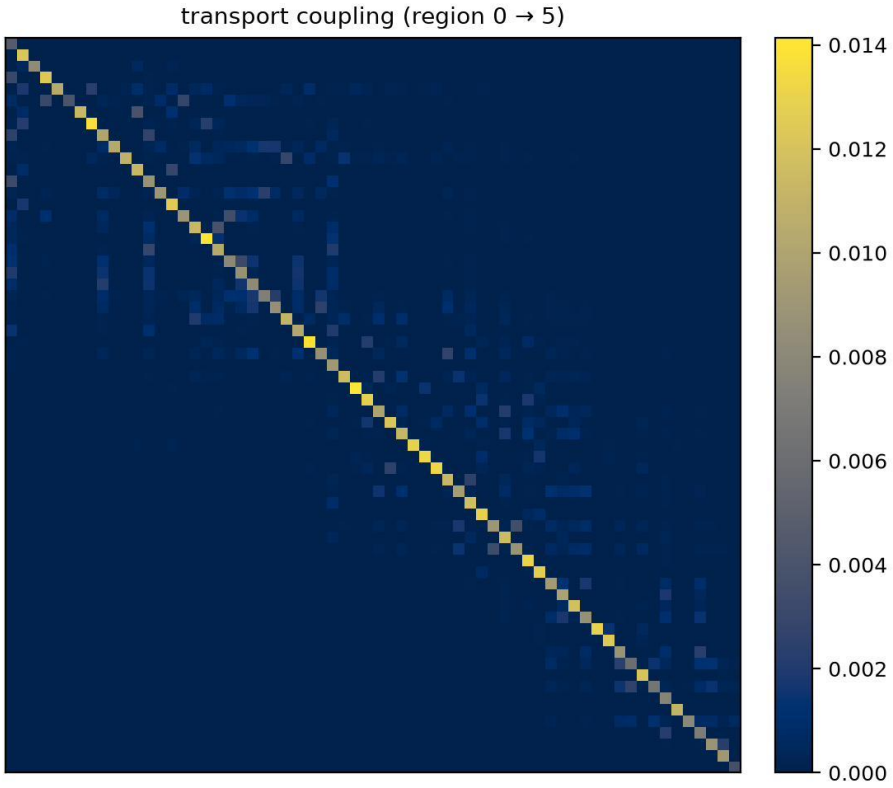

Fig. 7. A transport coupling between two region abundance distributions. The plan itself is the explanation of what differs: it states which abundance mass moves where, information no scalar distance provides

4.4 E3: topological terrain descriptor and invariance (N3)

In Table 8 and Fig. 8, the stability ratio is approximately 25,000 (Table 8, Fig. 8) on a bar of 5 against the mean drift (14) of the persistence descriptor value. Fig. 8 presents the persistence descriptor value with the same results. The ratio is known to be the same under rotation and flip (invariant) and under monotone illumination under vectorization grid it turns out to be an empirical fact as explained in Section 3.7. In contrast GLCM is not characterized by this and it captures the co-occurrence of two values over two points in space. A 1e-6 image grid shift (or spurious drift 1.16) was eliminated by per-diagram min-max normalization, and allowed the images to become comparable between scenes. This was one of the few problems found with the vectorization and not a mathematical one.

Table 8. E3 descriptor drift under nuisance transforms on Cuprite. The persistence descriptor is exactly invariant to rotation and flip and invariant to monotone illumination up to the vectorization grid; the stability ratio over GLCM is approximately 25,000

Transform	Persistence drift	GLCM drift
Rotation 90 deg	0.0	0.1025
Rotation 180 deg	0.0	0.0
Rotation 270 deg	0.0	0.1025
Flip left-right	0.0	0.0830
Flip up-down	0.0	0.0830
Gamma 0.6	3.8e-6	0.0051
Gamma 0.8	3.2e-6	0.0021
Gamma 1.25	1.0e-6	0.0017
Gamma 1.6	1.0e-6	0.0040
Affine gain and offset	6.3e-6	0.0
Mean	1.5e-6	0.0384

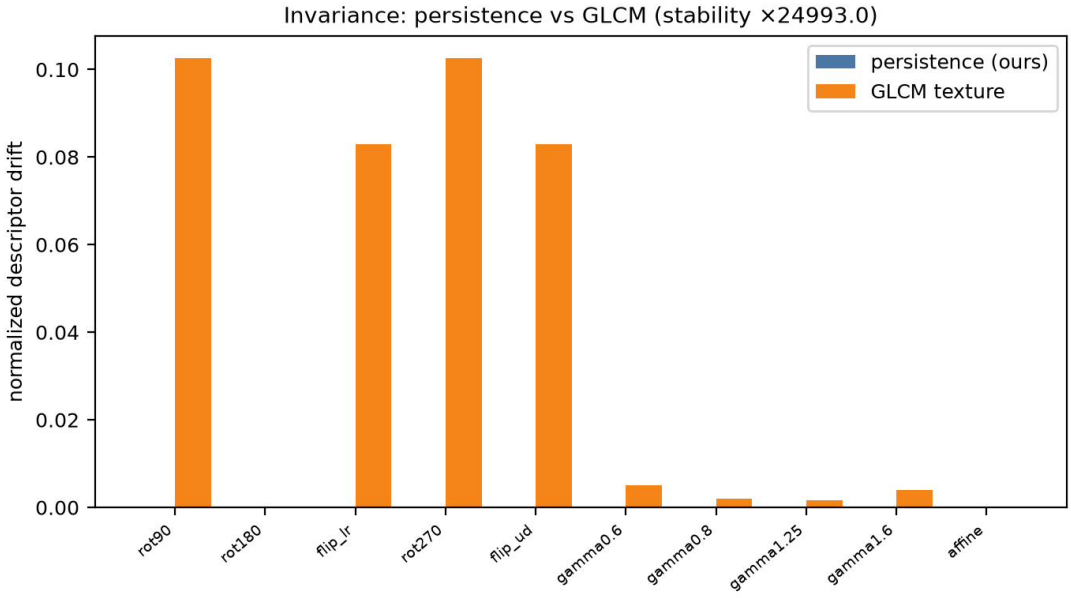

Fig. 8. E3 invariance comparison: persistence-descriptor drift against GLCM texture drift for each transform, on a shared scale

Mapping the descriptor back onto the terrain using (12) demonstrates meaningfulness: Since the counts were confirmed in Section 4.1, Fig. 9's overlay of H0 basins and H1 loops on the source raster is evidentiary. Without altering the designs, a full-Cuprite call was reduced from 20 minutes to 0.11 seconds by limiting the flood fill to the top-k most persistent features.

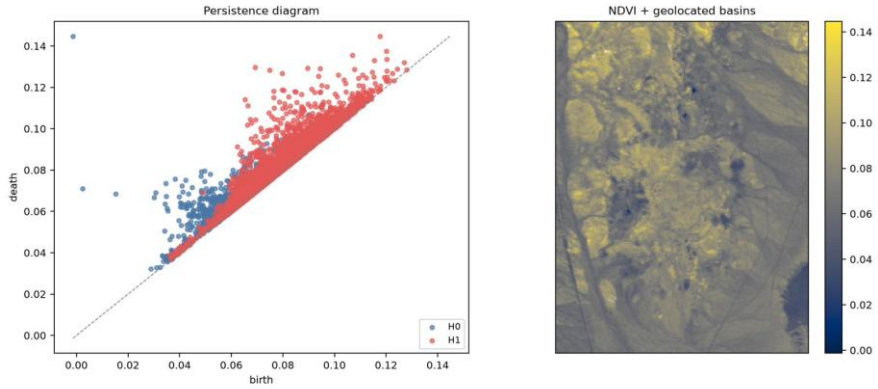

Fig. 9. Georeferenced topology on Cuprite: the persistence diagram (left) and the H0 basin and H1 loop features mapped back to their generating pixel polygons on the raster (right), colored by persistence

4.5 E4: verifiable agentic orchestration (N2)

The four systems in Section 3.10 are used by E4 to operate the fixed 20-intent suite, with system A utilizing the live Claude Sonnet model. The results are shown in Table 9 and Fig. 10, which state two different but often confused things.

Table 9. E4 four-system comparison over the 20-intent suite. System A uses the live language model with full verification; B is the LLM-agnosticism control; C represents current ungrounded agent practice; D isolates the verifier

System	Plan validity	First-shot	Tool-call error	Node success	Grounding (first pass)
A. TerraXIA with LLM planner	1.00	1.00	0.00	0.94	0.58
B. TerraXIA symbolic planner	1.00	n/a	0.00	0.94	1.00
C. Naive LLM foil (no registry, no verification)	0.00	n/a	0.53	0.00	0.00
D. LLM planner, verification ablated	1.00	n/a	0.00	0.94	1.00

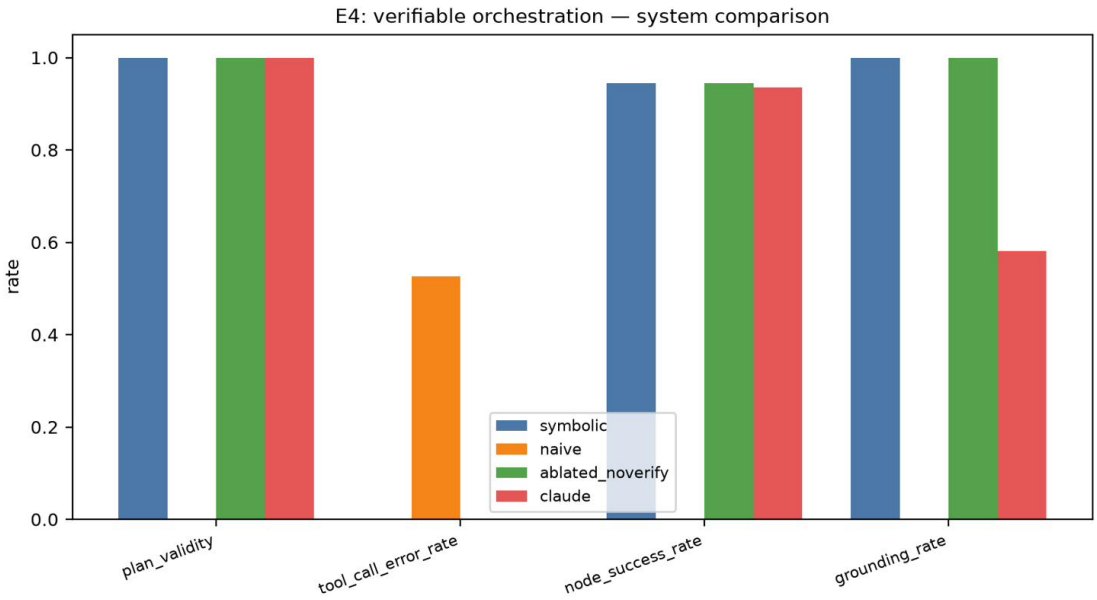

Fig. 10. E4 metrics by system. Registry grounding eliminates tool-call errors entirely (A against C); the grounding checker's value appears as the 42% of first-pass report claims it removes from system A

It gets rid of all tool hallucinations through the registry grounding. Using the serialized registry, the model yielded statically valid plan graphs on every 20 of 20 intents, while the same model that does not have any grounding in the registry (C) failed on a call on 53% of the intents and 100% of the plans were invalid before any math performed. There is a structural reason for this, a hallucinated operator cannot stand static verification. The default behavior of the ungrounded agent, discovered by the verifier, is common in the systems analyzed [1, 14] and is similar to the orchestration-error problem.

Secondly, claim hallucinations are a distinct type of failure and so need a distinct approach. The same model was passed 15 "only" 58% at first pass for reports claiming to have satisfactory data, but they revealed only some statistic with a maximum z score (one observed report had anomaly_score instead) and an anomaly count. Seeing failures is better in material and survey intents (1.00, +-3 for 7 assurances out of 20 possible), while they tend to group together in terrain and topology intents (as low as 0-0.25 per intent). On delivered reports, (15) is satisfied by construction at the end of the revision cycle and delivered reports are the valid number not first pass. The image below shows an ablation of the LLM-agnosticism property, a simulated brain is placed in a too-long loop, the plan is rejected, the loop is self-corrected from structural errors, seven to nine pinhole queries are resolved and 100% of the grounding is achieved. With the same interface, the loop is ascertained by a LLM without needing any access to its APIs.

4.6 End-to-end case study with the live model

With each figure and sentence produced by the pipeline, three intents, one for each family, raced end to end on Jasper. The ten georeferenced labels were laid out, as were nine map layers and five of five claimed grounds, which were verified (Fig. 11). The five nodes surveyed had 4 out of 4 allegations that were valid. The governing rule is anomalous and is expected to produce appropriate inhibit and is met by the concurrent peak of the score across 400 regions ($z = 1.51$, but said to be > 3). They specifically do not identify an anomaly, instead there are two sub-threshold signals, one that triggers exclusively for 110 reflectance's below-average, and a reconstruction-error peak at 9.5 times the scene mean. This is regarded as a calibration, shadowing or atmospheric-correction feature rather than a land-cover feature. Instead of hiding the Physics fail, it's brought up in the run summary, and there are artifact citations for each sentence.

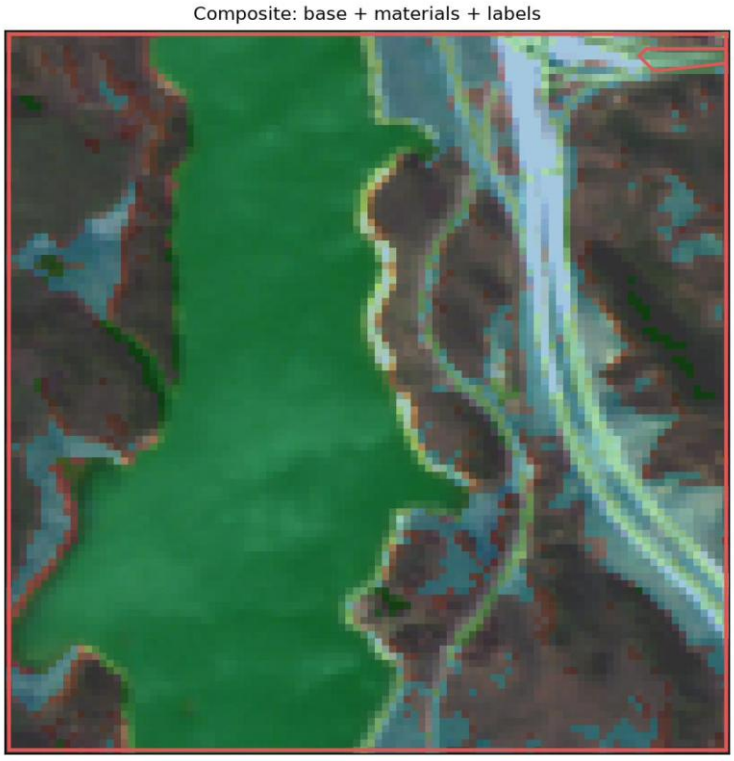

Fig. 11. Composite map assembled by the cartographer for the live layouting run on Jasper: false-colour base, abundance overlay, system-generated labels and finding callouts, each traceable to a hash-addressed artifact

The same operator set ran unchanged across all three datasets (209 analysed regions on Pavia, 49 on Jasper, 221 on Cuprite). Fig. 12 shows the cross-terrain strip. Identical operators, three very different scenes, nothing to retrain.

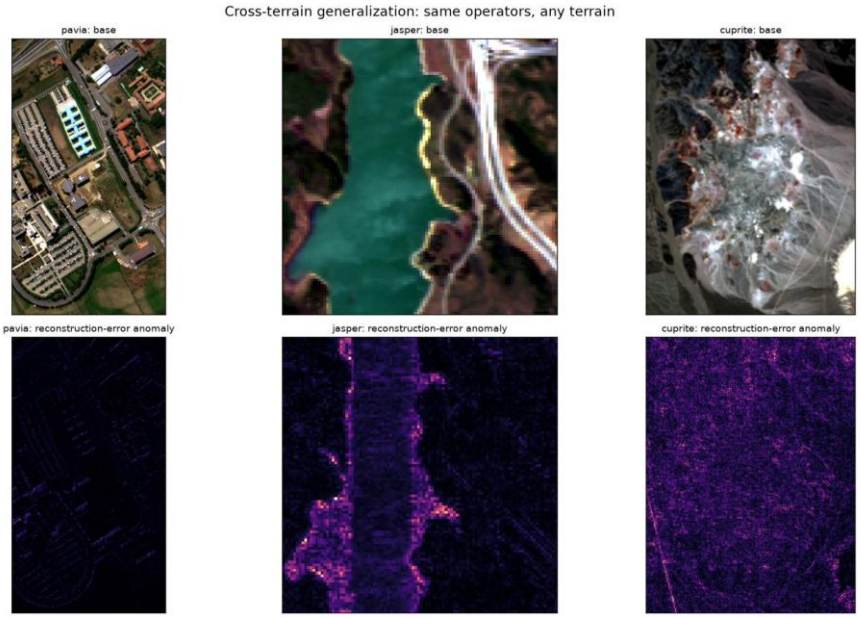

Fig. 12. Cross-terrain evidence: the same deterministic operator set applied unchanged to Pavia University (urban), Jasper Ridge (mixed natural), and AVIRIS Cuprite (mineral terrain)

4.7 Limitations and the status of N4

Three novelties meet their pre-registered bars: N1 with margin, N3 decisively and N2 with the paper's clearest evidence. The N4 test results are proofs that the output is correct and is easy to interpret (all metric-axioms to machine precision, all Sinkhorn rankings are equal and only one object of the comparison is transportation plan). Both the quantitative superiority and the other links to the N4 use of the operator are not yet proven (but would follow from the benchmark design); read on to Section 4.3. The analysis provides direct follow-up designs, a contextual anomaly benchmark with anomalous material that is globally normal but locally unusual, which is an example of a design that will always defeat a global Mahalanobis detector, as well as retrieval over regions with matched mean abundance of interest, but different distributional shape, where the only difference between the classes is the shape of the distribution. Both are more focused for the next study. Meanwhile, there are no reports of superiority of N4, but rather accuracy and comprehensibility are emphasized.

Other constraints over the claims, tiled Houston 2018 run complies with this recipe, not to be confused with user feedback over a user survey (E4 measures orchestrations on a crop, not throughput, and does not have a user survey) and the self-label evaluation has a structural upper bound of 5 and lower bound of 9 classes, when the literature has only 10, the shoot calibration sample completes a virtual dimensionality that is inferior to the available literature. For next papers, conformal gating, foundation-model tools, bi-temporal ledgers, tamper agents and active learning are already in the store, though.

5. Conclusion

The intention of this article was to fill a specific lack of trust when using agentic remote sensing where systems can see hyperspectral scenes but are not known to give conclusions about these scenes. The suggested and assessed response is to be in an architectural form. When every operator is a deterministic operation, every output of every operation is a content-hashed, geo-referenced artifact and all agent claims need to be referred to the artifact store, verifiability in the system stops being a desired property and becomes a construct. The four novelties under evaluation in this construction fill in this construction. Its labels are generated in situ using a compressed physically interpretable substrate supplied by the tensor world-state (unmixing took place at SAD 0.149 rad, with RMSE of 0.156 vs ground truth). The main empirical finding of the paper was generated by the verifiable orchestration layer, registry grounding completely eliminated tool-call errors (20 of 20 first-shot valid plans against a 53% error rate and universal plan rejection for the ungrounded foil) and the grounding checker eliminated 42% of the model's first-pass report claims as unverifiable. It represents a direct measure of the unsupported assertion a fluent model outputs even when it has the correct planning. It is a measure of the unsupported assertion a fluent model produces just with the right planning, as well as the value of the checker. Each topological descriptor the persistence descriptor reports is a polygon on the map; it is precisely invariant with a texture feature drift of about 25,000. The only similarity technique used in the study that explains its results is the transport technique which has been shown to be accurate within the machine accuracy. Because the publication is quantitative, it is made perfectly clear that its superiority must be demonstrated to benchmarks similar to that which is being measured.

Every figure is a piece of hash-addressed artifact, every number can be re-generated from a run directory and all of the things reported run on a CPU-only laptop, so the paper is replayable too. The two N4 benchmarks created in Section 4.7, the two change and tamper reasoners, both developed in the same context, are added in the following papers in this Program, the statistically guaranteed triage layer, as well as the conformal gating over the anomaly scores created here, is added in the third paper. There are already the hangers to accommodate both.

References

1. Akinboyewa, T., Li, Z., Ning, H., Lessani, M. N.: GIS Copilot: towards an autonomous GIS agent for spatial analysis. International Journal of Digital Earth 18 (2025). https://doi.org/10.1080/17538947.2025.2497489
2. Carvalho, J. P., Aguiar, A. P.: Multi-agent reinforcement learning for zero-shot coverage path planning with dynamic UAV networks. Robotics and Autonomous Systems 195, 105163 (2025). 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.: 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 (2025). https://doi.org/10.1016/j.jag.2025.104814
4. Feng, H., et al.: Flexible exploration: a hyperspectral image classification framework based on dynamic scale and graph optimal transport. Information Fusion 128, 103962 (2026). 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.: 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 (2025). https://doi.org/10.1016/j.isprsjprs.2025.06.004
6. Kasetty, B., Krishnan, S.: Transform-aware deep learning and explainable AI in remote sensing: a comprehensive meta-analysis. IEEE Access 14 (2026)
7. Kuper, P., Liu, R., Breunig, M.: 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.: Uncertainty quantification for forest attribute maps with conformal prediction and k-NN method. Remote Sensing of Environment 325, 114758 (2025). https://doi.org/10.1016/j.rse.2025.114758
9. Liu, Y.: The role of U-Net variants in semantic segmentation of remote sensing images: a survey. Applied and Computational Engineering 210(1) (2025). https://doi.org/10.54254/2755-2721/2025.LD29991
10. Mai, G., et al.: Towards the next generation of geospatial artificial intelligence. International Journal of Applied Earth Observation and Geoinformation 136, 104368 (2025). https://doi.org/10.1016/j.jag.2025.104368
11. Marjani, M., Mahdianpari, M., Gill, E. W., Mohammadimanesh, F.: Explainable AI for wetland mapping using high-resolution remote sensing data. In: IGARSS 2025, IEEE International Geoscience and Remote Sensing Symposium (2025). https://doi.org/10.1109/IGARSS55030.2025.11243021
12. Peter, E., Ang, L.-M., Seng, K. P., Srivastava, S.: Recent advances in deep learning for SAR images: overview of methods, challenges, and future directions. Sensors 26(4), 1143 (2026). https://doi.org/10.3390/s26041143
13. Ren, L., et al.: Advances in hyperspectral image unmixing: from algorithmic frameworks to practical applications. Information Geography 2(1), 100035 (2026). https://doi.org/10.1016/j.infgeo.2025.100035
14. Sun, Z., Zhou, Y., Yang, J.: An LLM-based multi-agent system for remote sensing analysis. Big Earth Data 10(1) (2026). https://doi.org/10.1080/20964471.2025.2600178
15. Wu, K., et al.: A semantic-enhanced multi-modal remote sensing foundation model for Earth observation. Nature Machine Intelligence 7(8) (2025). https://doi.org/10.1038/s42256-025-01078-8
16. Wu, Y., Mu, X., Shi, H., Hou, M.: An object detection model AAPW-YOLO for UAV remote sensing images based on adaptive convolution and feature fusion. Scientific Reports 15, 16214 (2025). https://doi.org/10.1038/s41598-025-00239-4
17. Xue, C., Ye, Q., Shao, X., Hu, T., Xu, Z.: Enhanced hybrid CNN and Transformer network for remote sensing image change detection. Scientific Reports 15 (2025). https://doi.org/10.1038/s41598-025-94544-7
18. Zhou, G., Qian, L., Gamba, P.: Advances on multimodal remote sensing foundation models for Earth observation downstream tasks: a survey. Remote Sensing 17(21), 3532 (2025). https://doi.org/10.3390/rs17213532
19. Zhu, X. X., et al.: On the foundations of Earth foundation models. Communications Earth and Environment 7, 103 (2026). https://doi.org/10.1038/s43247-025-03127-x
20. Zou, J., Qu, H., Zhang, P.: Conventional to deep learning methods for hyperspectral unmixing: a review. Remote Sensing 17(17), 2968 (2025). https://doi.org/10.3390/rs17172968