Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 38, 2026, Sofia

Reviews

DOI: https://doi.org/10.3897/arb.v38.e25

HYBRID INTELLGENT MODELLING OF ATMOSPHERIC POLLUTION FROM EARTH OBSERVATION DATA: A CRITICAL REVIEW OF DIGITAL TWINS, NEURAL SURROGATES AND GROUNDED LLM INTERFACES

Dimitar Makariev, Plamen Trenchev

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: dmakariev@space.bas.bg; ptrenchev@space.bas.bg

Key words: *Earth Observation, Atmospheric Pollution, Hybrid Modeling, Earth System Digital Twins, Neural Surrogates, Symbolic Regression, Uncertainty Quantification, Foundation Models, Grounded LLMs, Copernicus Observia AI*

Abstract

Atmospheric processes governing air pollution are inherently nonlinear, multi-scale, and characterised by substantial uncertainties arising from emissions, meteorology, chemical transformations, and sensor noise. This critical narrative review synthesises contemporary intelligent modelling paradigms for atmospheric pollution analysis using satellite (Sentinel-5P TROPOMI, Sentinel-3) and ground-based observations, identifying the operational role of each method family and their integration in a unified hybrid architecture. Building on our SES-2025 review [1], we synthesise individual method families into a modular five-layer pipeline. The framework delivers physical consistency via Earth System Digital Twins, predictive power via neural and Extreme Learning Machine surrogates, interpretability via symbolic regression, calibrated uncertainty via Monte Carlo ensembles, and natural-language accessibility via grounded LLM interfaces. A benchmarking strategy targeting Balkan-region NO_2 and $PM_{2.5}$ episodes is outlined. The novelty lies in synthesising the complementary roles of these methods within an auditable modular architecture, with explicit separation between quantitative modelling backends and LLM-based interaction layers.

1. Introduction

Atmospheric composition and pollution dynamics constitute one of the most pressing environmental and public-health challenges of the Anthropocene. Accurate, timely, and uncertainty-aware modeling is indispensable for air-quality forecasting, source apportionment, population exposure assessment, health-impact studies, and evidence-based policy design at urban, regional, and continental scales [2, 3]. The Copernicus programme delivers petabyte-scale, heterogeneous, multi-temporal datasets most notably from Sentinel-5 Precursor (TROPOMI) with daily global observations of NO_2, CO, SO_2, HCHO, tropospheric ozone, and aerosol optical depth (AOD); Sentinel-3 SLSTR for fire radiative power and smoke plume tracking. Extracting physically consistent, actionable intelligence from these streams demands methods that transcend traditional numerical weather prediction or chemical transport models (CTMs) alone. Physics-based CTMs remain essential for process understanding and long-term projections, yet they are computationally expensive and subject to structural uncertainties in emission inventories, boundary conditions, and sub-grid parameterizations. Data-driven and intelligent methods machine learning, deep learning, symbolic regression, uncertainty quantification, and more recently foundation models and LLM interfaces offer complementary strengths in speed, downscaling, pattern recognition, and accessibility.

2. Review Methodology

This work constitutes a critical narrative review and conceptual framework paper. Literature was identified through searches of Scopus, Web of Science, IEEE Xplore, ScienceDirect, and arXiv (2018–2026). Inclusion criteria: peer-reviewed publications or accepted conference papers on method families in atmospheric remote sensing or pollution modeling; preprints only for systems accepted at major venues (ICCV, Nature) or with official institutional documentation. Of approximately 85 sources reviewed, 30 are cited. Method families were categorized thematically: physics-based backbones (ESDTs, CTMs, system dynamics), ML surrogates (NN, ELM), uncertainty and explainability (Monte Carlo, Bayesian networks, symbolic regression), foundation and multimodal models, LLM/RAG interfaces, and complementary nonlinear methods (phase-space reconstruction, Markov chains, fuzzy logic, reinforcement learning).

3. Identified Research Gaps

3.1. Absence of hybrid benchmarks. Existing EO benchmarks focus on single tasks (segmentation, regression, change detection) and rarely combine physics simulators, ML surrogates, policy scenarios, and human AI interaction with explicit metrics for physical consistency, uncertainty coverage, and decision quality. This is the primary structural gap the proposed framework and benchmarking strategy (Section 12) address.

3.2. Physical consistency of foundation models. Most multimodal pre-training objectives do not enforce conservation laws or atmospheric chemistry constraints. Hybrid pipelines embedding symbolic constraints inside fine-tuning loops constitute a promising remedy, prototyped in this framework using Copernicus-FM and Aurora backbones [11, 16].

3.3. LLM-model integration gap. Grounded LLM interfaces (e.g., Observia AI) are not yet coupled to validated atmospheric simulation backends in any peer-reviewed operational deployment. The proposed Layer 5 architecture provides a blueprint for this integration.

3.4. Regional data sparsity. The Balkans and South-Eastern Europe are under-represented in atmospheric AI benchmarks relative to Western European urban corridors. The case studies proposed in Section 12 explicitly target this gap.

3.5. Uncertainty propagation across interfaces. End-to-end propagation of uncertainty from satellite retrieval through CTM coupling to policy indicator is rarely performed; each layer interface that does not propagate uncertainty produces overconfident downstream outputs.

4. Atmospheric-Pollution-Specific Modeling Challenges

4.1. Satellite Column-to-Surface Conversion

TROPOMI delivers tropospheric vertical column densities (VCDs), not surface concentrations. Conversion requires knowledge of the vertical profile shape, which depends on planetary boundary layer height (PBLH), temperature inversion strength, and chemical gradients. PBLH uncertainties from ERA5 alone contribute 15–30% relative error in surface NO_2 estimates over complex terrain.

4.2. Spatio-Temporal Mismatch

TROPOMI has a single daily overpass (~ 13:30 local solar time) with a 3.5×5.5 km² pixel at nadir. Ground stations report hourly data at point locations. Simple co-location introduces representativeness errors of up to 40% for $PM_{2.5}$ in complex urban environments [28].

4.3. Emission Inventory Uncertainty

National emission inventories carry relative uncertainties of 20–50% for road transport NOx and up to 100% for agricultural NH_3. These propagate directly into CTM forecasts and data-assimilation backgrounds.

5. Earth System Digital Twins and System Dynamics

5.1. Earth System Digital Twins (ESDTs)

Earth System Digital Twins constitute dynamic, multi-scale virtual replicas that fuse real-time Earth-observation streams, numerical simulation ensembles, and AI surrogate models to enable "what-now", "what-next", and "what-if" analyses [2, 3]. In the atmospheric domain they support high-resolution pollutant transport modeling, urban air-quality digital twins, wildfire-smoke plume forecasting, and policy-scenario comparison under varying emission trajectories and meteorological regimes [2].

5.2. System Dynamics

System-dynamics (SD) modeling formalizes delayed feedback loops, nonlinear couplings, and policy levers that govern emission generation, atmospheric dispersion, population exposure, health impacts, and mitigation effectiveness [4, 5, 29]. SD complements ESDTs by making the causal skeleton explicit and communicable to non-technical stakeholders policy makers, local authorities, and the public. When embedded inside an ESDT, SD supplies the structural backbone for transparent "what-if" simulations.

6. Machine Learning Surrogates: Neural Networks and Extreme Learning Machines

6.1. Neural Networks and Physics-Informed Variants

Deep neural architectures excel at cloud classification, AOD retrieval, pollutant-plume segmentation and tracking, chemical-transform emulation, and fast surrogate modeling of expensive CTM components [10, 11]. Convolutional networks extract spatial features from multispectral imagery; recurrent and transformer models capture temporal dynamics; graph neural networks handle pollutant transport on irregular grids. Physics-informed neural networks (PINNs) embed governing equations (continuity of mass, advection-diffusion) as soft constraints in the loss function, improving generalization under distribution shift relative to unconstrained architectures.

6.2. Extreme Learning Machines (ELMs)

Extreme Learning Machines [23] employ random hidden-layer projections with analytically solved output weights, offering training speeds two to three orders of magnitude faster than gradient-based deep networks while retaining competitive accuracy on streaming time-series, anomaly-detection, and now casting tasks. Online ELM variants support incremental learning as new observations arrive, directly addressing temporal distribution shift in operational monitoring.

7. Uncertainty Quantification and Equation Discovery

7.1. Monte Carlo Methods

Monte Carlo (MC) methods remain the reference technique for propagating uncertainties from emission inventories, meteorological fields, model-parameter ensembles, sensor noise, and structural model error into probabilistic risk metrics and forecast envelopes [6, 7]. In the hybrid context, MC is applied both to ESDT cores and to AI surrogate outputs, delivering calibrated confidence intervals indispensable for risk-sensitive decisions. Surrogate-based UQ methods (polynomial chaos expansion, Gaussian process emulation) can reduce the required model evaluations by one to two orders of magnitude relative to naive MC while preserving coverage properties.

7.2. Genetic Programming and Symbolic Regression

Genetic programming and symbolic regression extract compact, human-readable mathematical expressions directly from data, optionally embedding physical priors, conservation constraints, or known chemical reaction pathways [8, 9]. These techniques are valuable for reduced-order modeling of atmospheric processes, discovery of governing equations for sub-grid phenomena, and generation of interpretable relationships that complement black-box neural predictions. Recent benchmarks demonstrate that SR can recover known atmospheric scaling relations (e.g., simplified NO_x–O_3 titration relationships) with relative errors below 5% from synthetic data generated by full CTM runs [9].

8. Foundation Models for Earth Observation and Atmospheric Applications

The period 2024–2026 witnessed rapid growth in foundation-model research tailored to Earth observation [14, 17, 18, 21]. TerraMind [15] is a large-scale generative multimodal model designed for Earth observation, accepted at ICCV 2025, supporting any-modality generation and retrieval across optical, SAR, and auxiliary streams. Copernicus-FM [16] is a unified foundation model trained on the Copernicus data ecosystem for cross-mission, cross-sensor generalization. Aurora [11, 30] is an Earth-system foundation model pre-trained on more than one million hours of geophysical reanalysis data, achieving competitive skill in medium-range weather forecasting at a fraction of operational NWP inference cost. SkySense++ [19] delivers strong performance on scene classification and change detection. Panopticon [20] demonstrates zero-shot generalization across heterogeneous EO sensors.

9. Grounded LLM Interfaces: Case Study on Copernicus Observia AI

Large language models must not be treated as atmospheric models. Their scientific role in the proposed framework is precisely delimited: retrieval, orchestration, explanation, and human–model interaction. Quantitative prediction must remain grounded in observations, CTMs, ESDTs, validated ML surrogates, or statistically calibrated models.

9.1. Copernicus Observia AI: Documented Current Functionality

In June 2025, the European Environment Agency (EEA) launched the beta version of Copernicus Observia AI ("Ask Copernicus") [12, 13], a conversational assistant for finding, understanding, and using Copernicus products. Based on official documentation, the system implements a retrieval-augmented generation (RAG) pipeline on local EEA servers, with a vector database over CLMS and CAMS documentation and source-grounded responses; on-premise execution is stated as a GDPR/EUDPR design goal [13]. Independent peer-reviewed evaluation of its atmospheric-domain accuracy or TROPOMI-specific coverage is not yet available.

9.2 Proposed Integration in the Hybrid Workflow

Table 1 distinguishes documented current functionality from proposed future integration targets. In the hybrid pipeline (Section 11), Observia AI would function in two complementary roles: (a) front-end conversational interface translating user intent into pipeline actions (data retrieval, scenario launch, report generation) consistent with the documented beta capability; and (b) automated retrieval agent supplying fresh Copernicus scenes, metadata, or model diagnostics to upstream layers on demand requiring future API or agentic integration not yet available.

Table 1. Copernicus Observia AI – documented current capability vs. proposed hybrid pipeline integration

Capability	Current Observia (beta) [12, 13]	Proposed hybrid pipeline extension
Copernicus data discovery	Yes – documented	Automated scene staging into Layer 1
Product explanation / QA guidance	Yes – documented	Model-aware metadata validation pre-ingest
Natural-language archive query	Yes – documented	Trigger structured data pulls into Layer 1
Scenario simulation	No direct evidence	Via ESDT/CTM backend (Layer 2); requires future API
Monte Carlo UQ	No direct evidence	Via Layer 4; results surfaced through Layer 5
Policy impact assessment	No direct evidence	Via coupled exposure health module; requires validation
Automated provenance reports	Partial – draft summaries documented	Full pipeline provenance; uncertainty-qualified indicators

Sources: [12] EEA announcement; [13] Ask Copernicus beta documentation. "No direct evidence" denotes a proposed future integration target, not an existing feature.

9.3. LLM/RAG Risk Register

Table 5 provides the full risk register. Key risks specific to the LLM/RAG layer include: (a) hallucination of atmospheric quantities LLMs can confidently generate incorrect numerical values without flagging uncertainty; (b) stale grounding if the vector database is not updated, the system may cite superseded TROPOMI algorithm versions or outdated CAMS forecasts; (c) prompt-injection vulnerability in agentic deployment; (d) attribution laundering users may trust LLM-synthesized text as the primary source rather than the underlying Copernicus datasets.

10. Complementary Nonlinear and Probabilistic Methods

10.1. Phase-Space Reconstruction and Nonlinear Dynamics

Takens' embedding theorem [24] guarantees that a scalar time series can be used to reconstruct the geometry of the underlying attractor via delay embedding, provided sufficient embedding dimension and lag [25]. These invariants serve as physically meaningful features for downstream neural and ELM predictors in Layer 3.

10.2. Bayesian Networks

Dynamic Bayesian Networks provide principled probabilistic graphical models for multi-sensor data fusion and sequential belief updating as new satellite overpasses or ground measurements arrive. They represent dependencies among emissions, meteorology, satellite observations, and ground measurements while naturally propagating uncertainties. Strengths: explicit uncertainty representation; principled fusion of heterogeneous sources.

10.3. Markov Chains and Hidden Markov Models

Markov chains and HMMs capture regime transitions (clean ↔ moderately polluted ↔ smog episode) driven by synoptic meteorology, emission changes, or seasonal cycles. They supply compact state representations that improve sample efficiency for reinforcement learning agents and furnish stochastic scenario generators inside MC ensembles.

10.4. Fuzzy Logic Systems

Fuzzy logic systems provide a mathematically rigorous framework for handling inherent vagueness in air-quality index definitions, expert rules for episode classification, and multi-criteria decision support when combining health-protection and economic objectives.

10.5. Reinforcement Learning

RL agents can learn adaptive strategies for intelligent satellite tasking during detected pollution events, real-time control of urban traffic management, or active learning to identify the most informative Copernicus scenes for uncertainty reduction.

Table 2. Role of complementary nonlinear and probabilistic methods in the hybrid pipeline

Method	Atmospheric role	Strength	Limitation	Pipeline layer
Phase-space / CCM	Attractor geometry; causal inference	Nonlinear causality; no distributional assumptions	Requires long low-noise series	Layers 1, 3
Bayesian Networks	Multi-sensor fusion; UQ propagation	Principled uncertainty; heterogeneous fusion	Structure learning NP-hard	Layer 4
Markov Chains / HMMs	Regime transitions; scenario generation	Interpretable; low cost	Markov assumption may fail	Layers 4, 5
Fuzzy Logic	AQI; episode classification	Handles linguistic vagueness	Manual rules; limited scalability	Layer 5
Reinforcement Learning	Adaptive satellite tasking; control	Learns optimal strategy	Sample-inefficient; reward design	Layer 5

11. Proposed Integrated Hybrid Workflow

The five-layer hybrid pipeline for atmospheric pollution analysis (Fig. 1) is designed to be modular, auditable, and extensible. Each layer has defined inputs, outputs, and validation requirements. No single component is irreplaceable; the architecture supports incremental implementation starting from any subset of layers.

Table 3. Hybrid pipeline layer specifications

Layer	Inputs	Outputs	Key method(s)	Validation metric
1. Data Integration	TROPOMI VCDs; ERA5; EEA stations; AERONET; GIS layers	Harmonized feature space; uncertainty metadata; QA flags	Spatio-temporal alignment; QA filtering; gap detection	Data completeness (%); temporal coverage; QA pass rate
2. ESDT + SD Core	Layer 1 output; boundary conditions; emission inventories	Physically consistent pollutant fields; scenario trajectories	CTM / ESDT; SD feedback models	RMSE vs. station holdout; mass-balance check
3. Neural / ELM Emulation	Layers 1–2 outputs; real-time sensor streams	Surrogate predictions; anomaly scores; real-time alerts	PINN; ELM; foundation model fine-tuning	MAE; RMSE; CRPS; 90% interval coverage
4. Symbolic + UQ Layer	Layers 2–3 outputs; observational data	Interpretable equations; MC ensembles; calibrated uncertainty intervals	Symbolic regression; Monte Carlo; Bayesian networks	Symbolic compactness; calibration curves; CRPS
5. LLM Orchestration + Analytics	All upstream outputs; user queries	Cited natural-language answers; provenance-tracked reports; pipeline triggers	Copernicus Observia AI (beta); RAG; LLM orchestrator	Grounding accuracy; hallucination rate (requires independent evaluation)

CRPS: Continuous Ranked Probability Score. Validation metrics for Layer 5 cannot yet be computed for Observia AI in the absence of independent benchmark datasets.

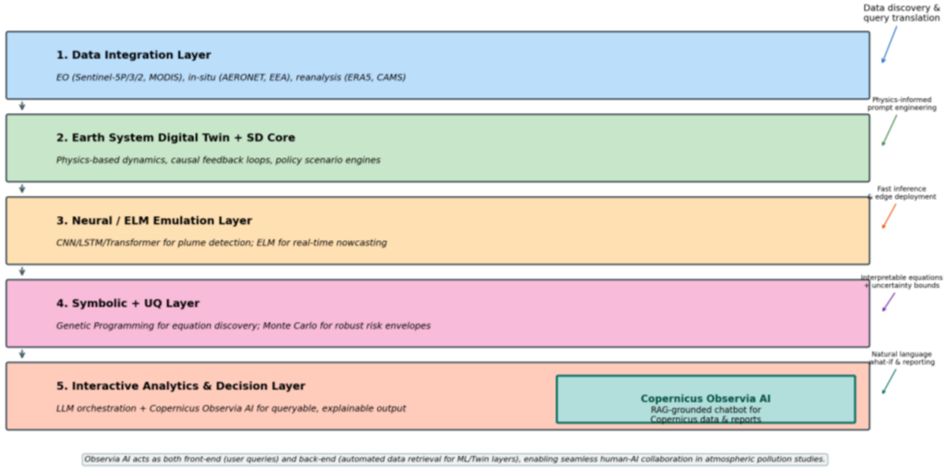

Fig. 1. Hybrid atmospheric modeling workflow integrating Sentinel-5P TROPOMI, Sentinel-3, ERA5, and EEA ground data into five layers: data integration, ESDT/CTM physics core, neural/ELM emulation, symbolic/UQ analysis, and LLM orchestration. Copernicus Observia AI supports user interaction and automated data retrieval.

12. Benchmarking and Validation Strategy

A validated hybrid pipeline requires a structured benchmarking approach that evaluates each layer independently and the integrated system end-to-end. The following strategy targets Balkan-region atmospheric pollution episodes, selected for their relevance to the identified regional data sparsity gap (Section 3) and for the availability of co-located TROPOMI and EEA ground data.

12.1. Case Study Candidates

1. Winter NO_2 inversion episodes, Sofia agglomeration (November–February). Rationale: strong PBLH suppression, topographic trapping, high satellite-to-surface discrepancy. Data: TROPOMI L2 NO_2 (QA $\geq$ 0.75); AQMD and EEA surface NO_2 (hourly); ERA5 PBLH and temperature profiles; AERONET Sofia.

2. Summer ozone episodes, Balkan urban corridor (June–August). Rationale: high photochemical activity; cross-border transport; health relevance. Data: TROPOMI tropospheric ozone and HCHO; EEA O_3 surface (hourly); CAMS regional ensemble forecasts.

3. Wildfire smoke transport, Balkan–Eastern Mediterranean (July–September). Rationale: extreme distribution shift; multi-sensor challenge. Data: Sentinel-3 SLSTR fire radiative power; CAMS fire emissions; AERONET AOD; EEA $PM_{2.5}$.

12.2. Performance Metrics

- **RMSE, MAE, bias:** standard point-prediction accuracy against held-out EEA surface observations
- **CRPS:** probabilistic forecast quality for MC ensemble outputs
- **90% interval coverage:** calibration of uncertainty estimates;
- **Spatial generalization:** leave-station-out cross-validation on withheld EEA sites
- **Symbolic compactness:** number of terms in SR-extracted equations; parsimony vs. accuracy trade-off
- **Inference cost:** wall-clock time per prediction per layer; relevant for real-time deployment
- **Physical consistency:** mass-balance check; monotonicity of pollutant concentration with emission rate in sensitivity experiments.

13. Interrelationships, Synergies, and Comparative Discussion

No single method family suffices across the full spectrum of atmospheric remote-sensing tasks. Fig. 2 visualizes the hierarchical synergies that emerge when the components are combined under a unified hybrid architecture.

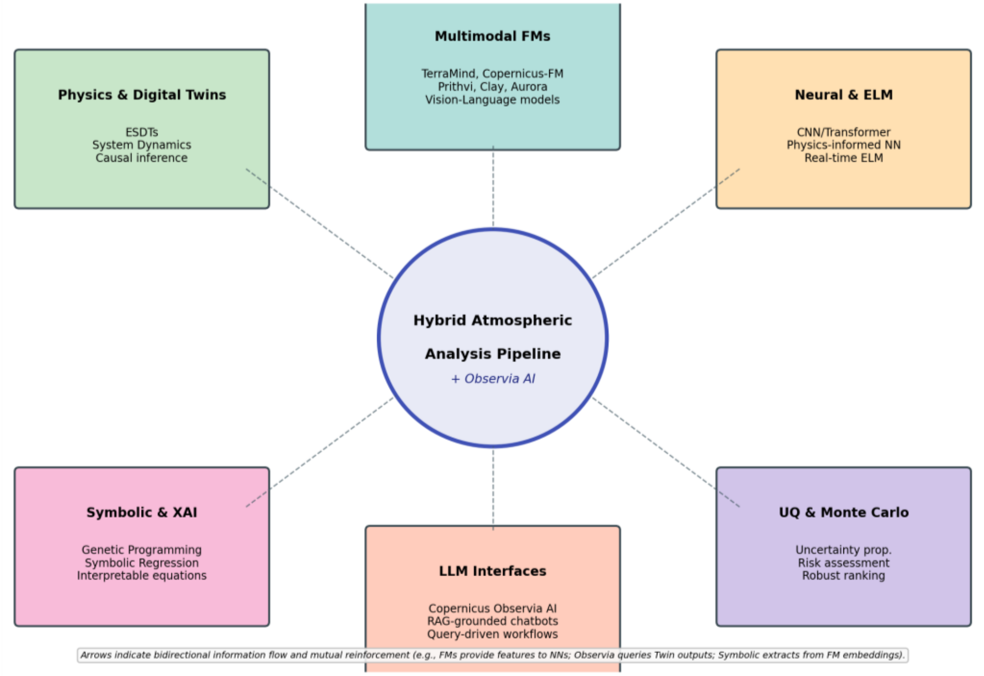

Fig. 2. Synergies among intelligent methods for atmospheric remote sensing. Foundation models and grounded LLM interfaces provide the user-facing integration layer, while neural, symbolic, Monte Carlo, and causal-validation components reinforce the hybrid workflow.

Table 4. Comparative characteristics of primary method families for atmospheric pollution modeling

Method Family	Predictive accuracy	Interpretability	Real-time	Physical consistency	Atm. maturity	Key limitation
Digital Twins + CTM	High	High	Low–Med.	Very High	High	High compute cost; emission inventory dependence
NN / PINN	Very High	Low–Med.	Medium	Med.–High (PINN)	High	Data-hungry; distribution shift without physical constraints
ELM	High	Low	Very High	Medium	Med.–High	Weaker spatial representation than deep nets
Symbolic Regression	Medium	Very High	High	High (w/ priors)	Medium	Instability on noisy high-dim. inputs
Multimodal FMs	High	Low–Med.	Medium	Emerging	Rapid growth	Physical consistency not guaranteed; high inference cost
LLM Interfaces	N/A	High (grounded)	High	Depends on backend	Beta 2025	Not a quantitative model; hallucination risk without grounding

Qualitative assessments based on reviewed literature 2018–2026. "Emerging" indicates rapid development with limited long-term validation in atmospheric pollution applications.

The addition of grounded LLM interfaces lowers the activation energy for domain experts, local authorities, and policy makers to exploit sophisticated pipelines, while mandatory source citation and provenance tracking address the black-box concern in regulated environmental decision contexts. LLM interfaces must remain a communication and retrieval layer; they must not be treated as substitutes for validated quantitative models. The framework enforces this through the layer architecture in Table 3 and the risk register in Table 5.

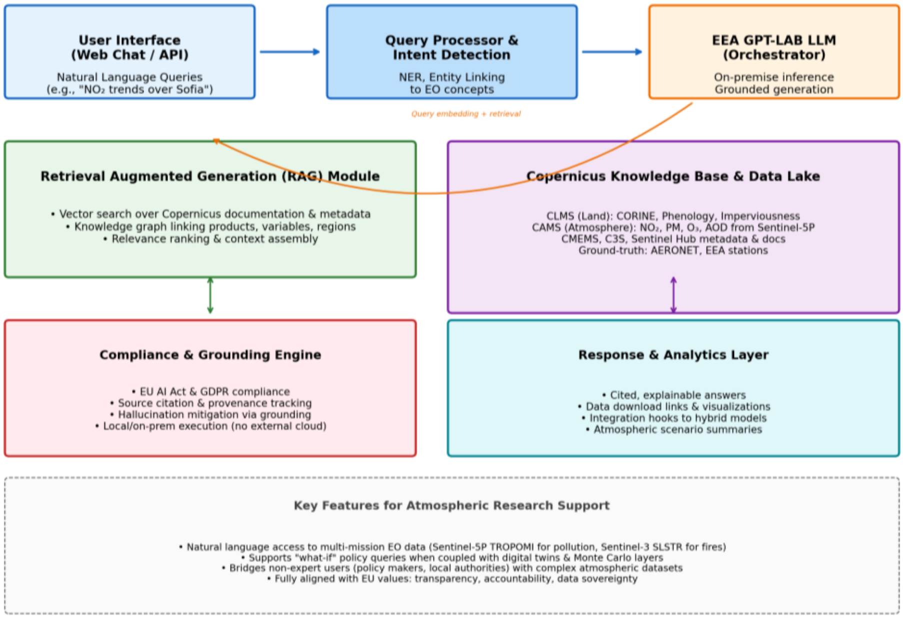

Fig. 3. Current and proposed roles of grounded LLM interfaces in hybrid atmospheric-pollution modeling: data discovery and product explanation vs. scenario launch, MC-UQ queries, and provenance-tracked policy reports.

14. Limitations, Threats to Validity, and Future Research Directions

14.1. Scientific and Engineering Limitations

1. Cross-domain generalization: neural surrogates and foundation models degrade under distribution shift from new sensors, extreme events, or under-sampled regions. Physics-informed regularization, continual adaptation, and meta-learning are active remedies.

2. Physical consistency of foundation models: most pre-training objectives do not enforce conservation laws. Hybrid fine-tuning of Copernicus-FM and Aurora with embedded symbolic constraints is a promising remedy.

3. Absence of hybrid benchmarks: community-standard benchmarks combining EO streams, physics simulators, ML surrogates, and explicit physical-consistency metrics are largely absent. The benchmarking strategy in Section 12 addresses this gap for the Balkan domain.

4. LLM grounding and evaluation: independent evaluation of Observia AI's atmospheric-domain hallucination rate and factual accuracy on TROPOMI-specific queries is a near-term research priority.

Table 5. Threats to validity and proposed mitigations

Threat	Type	Proposed mitigation
Publication bias toward successful AI models	Literature selection bias	Explicit inclusion of failure-mode analyses; preprint coverage
Limited atmospheric-specific validation of EO foundation models	External validity	Restrict performance claims to validated tasks; cite benchmark datasets
Fast obsolescence of LLM tools (Observia beta may be superseded)	Construct validity	Version-controlled integration; modular LLM orchestration layer
Hallucinated metadata or unsupported causal claims from LLMs	Internal validity	Mandatory source citation; human-in-the-loop for policy outputs
Regional data sparsity in Balkan stations	External validity	Leave-station-out CV; spatial extrapolation uncertainty reporting
Satellite–ground mismatch confounds validation	Measurement validity	Explicit representativeness error quantification in Layer 1
SR instability on noisy inputs may yield spurious equations	Internal validity	Physical prior constraints; cross-validation against independent datasets

14.3. Near-Term Research Directions

Priorities for 2026–2027: (a) benchmarked hybrid pipeline prototype for Balkan NO_2 and $PM_{2.5}$ episodes; (b) physics-constrained pre-training experiments using Copernicus-FM, TerraMind, and Aurora; (c) independent evaluation of Observia AI grounding accuracy for TROPOMI-specific queries.

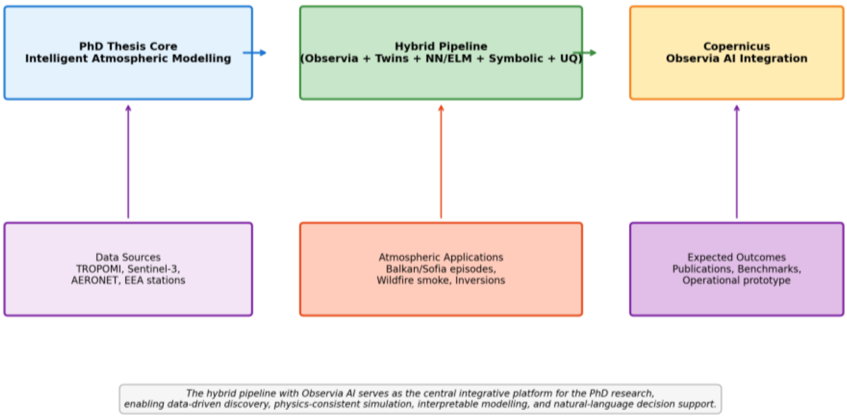

Fig. 4. Research and implementation roadmap for hybrid atmospheric-intelligence systems, organized by pipeline layer and covering near-, mid-, and long-term milestones for benchmarking, federated deployment, and LLM interface evaluation during 2025–2028.

15. Conclusion

This review has synthesized the contemporary landscape of intelligent atmospheric modeling into a modular five-layer hybrid framework integrating Earth System Digital Twins, system dynamics, Monte Carlo uncertainty quantification, neural and Extreme Learning Machine surrogates, symbolic regression, multimodal foundation models, and grounded LLM interfaces. The proposed architecture should be understood as a research agenda and integration blueprint rather than a validated operational system. Key conclusions are: (a) each method family occupies a specific, non-redundant role in the five-layer architecture with defined inputs, outputs, and validation metrics; (b) LLM interfaces are a communication and retrieval layer, not quantitative atmospheric models; (c) Copernicus Observia AI represents a promising but beta-stage grounded data-discovery interface whose atmospheric-domain performance requires independent evaluation; and (d) the primary outstanding gap is not the availability of individual method components but the absence of community-standard hybrid benchmarks combining EO streams, physics simulators, ML surrogates, and explicit physical-consistency and uncertainty-coverage metrics.

References

1. Makariev, D., Trenchev, P. (2025). Advanced Computational Approaches for Data-Driven and Intelligent Modeling of Atmospheric Processes. Proc. 21st International Scientific Conference SES 2025, Sofia, Bulgaria.
2. Hazeleger, W., et al. (2024). Digital twins of the Earth with and for humans. Communications Earth & Environment, 5, 153. https://doi.org/10.1038/s43247-024-01327-5
3. Stevens, B., et al. (2024). Earth Virtualization Engines (EVE). Earth System Science Data, 16(4), 2113–2122. https://doi.org/10.5194/essd-16-2113-2024
4. Peng, J., et al. (2024). System dynamics model applications in urban studies in the big-data era: A review. Sustainable Futures, 6, 100152.
5. Amiri, M., et al. (2024). A system dynamics approach to air pollution management. Proc. International System Dynamics Conference 2024.
6. Trees, V.J.H., et al. (2025). MONKI: A three-dimensional Monte Carlo simulator for remote-sensing retrievals. Geoscientific Model Development, in review.
7. U.S. EPA (2018). Guidance on the Development, Evaluation, and Application of Environmental Models. Report EPA/600/R-09/011A.
8. Pan, Y., Bélanger, S. (2024). Genetic algorithm-based atmospheric correction for coastal remote sensing. Remote Sensing of Environment, 302, 113956.
9. McGovern, J., et al. (2025). Potential for equation discovery with AI in the climate sciences. Earth System Dynamics, 16, 93–109. https://doi.org/10.5194/esd-16-93-2025
10. Benson, V., et al. (2025). Atmospheric transport modeling of CO_2 with neural networks. Journal of Advances in Modeling Earth Systems, 17, e2023MS004194.
11. Kochkov, D., et al. (2024). Neural general circulation models for weather and climate. Nature, 632, 1060–1066. https://doi.org/10.1038/s41586-024-07744-y

12. European Environment Agency (2025). https://land.copernicus.eu/en/news/copernicus-observia-ai-launched-to-guide-users-through-copernicus-data
13. European Environment Agency (2025). Ask Copernicus – Powered by EEA Observia AI. Beta release documentation. https://ask.copernicus.eu/ [Accessed June 2025].
14. Xiao, A., et al. (2024, rev. 2025). Foundation Models for Remote Sensing and Earth Observation: A Survey. arXiv:2410.16602.
15. Jakubik, J., et al. (2025). TerraMind: Large-scale generative multimodality for Earth observation. arXiv:2504.11171. Accepted ICCV 2025.
16. Zhu, Z., et al. (2025). Towards a Unified Copernicus Foundation Model for Earth Vision. arXiv:2503.11849.
17. Lu, S., et al. (2025). AI Foundation Models in Remote Sensing: A Survey. arXiv:2408.03464v1
18. Zhu, X.X., Xiong, Z., Wang, Y. *et al.* On the foundations of Earth foundation models. *Commun Earth Environ* 7, 103 (2026). https://doi.org/10.1038/s43247-025-03127-x.
19. Wu, K., Zhang, Y., Ru, L. *et al.* A semantic-enhanced multi-modal remote sensing foundation model for Earth observation. *Nat Mach Intell* 7, 1235–1249 (2025).
20. Panopticon team (2025). Panopticon: Advancing Any-Sensor Foundation Models for Earth Observation. CVPRW 2025, EarthVision Best Paper Award.
21. Zhou, G., et al. (2025). Advances on Multimodal Remote Sensing Foundation Models for Earth Observation Downstream Tasks: A Survey. Remote Sensing, 17 (21), 3532.
22. AI4Copernicus project (2023–2025). Bridging Artificial Intelligence with Earth Observation. H2020. https://ai4copernicus-project.eu/
23. Huang, G.-B., Zhu, Q.-Y., Siew, C.-K. (2006). Extreme learning machine: Theory and applications. Neurocomputing, 70 (1–3), 489–501. https://doi.org/10.1016/j.neucom.2005.12.126
24. Takens, F. (1981). Detecting strange attractors in turbulence. In Rand, D., Young, L.-S. (Eds.), Dynamical Systems and Turbulence. Lecture Notes in Mathematics, 898, 366–381. Springer.
25. Packard, N.H., et al. (1980). Geometry from a time series. Physical Review Letters, 45, 712–716.
26. Sugihara, G., et al. (2012). Detecting causality in complex ecosystems. Science, 338, 496–500. https://doi.org/10.1126/science.1227079
27. Grange, S.K., Carslaw, D.C. (2019). Using meteorological normalisation to detect interventions in air quality time series. Science of the Total Environment, 653, 578–588. https://doi.org/10.1016/ j.scitotenv.2018.10.344.
28. Verhoelst, T., Compernolle, S., Pinardi, G., et al. (2021). Ground-based validation of the Copernicus Sentinel-5P TROPOMI NO_2 measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global networks. Atmospheric Measurement Techniques, 14 (1), 481–510. https://doi.org/10.5194/amt-14-481-2021.
29. Sterman, J.D. (2000). Business Dynamics: Systems Thinking and Modeling for a Complex World. McGraw-Hill, New York.
30. Bodnar, C., et al. (2024). Aurora: A foundation model of the atmosphere. arXiv:2405.13063.

ХИБРИДНО ИНТЕЛИГЕНТНО МОДЕЛИРАНЕ НА АТМОСФЕРНО ЗАМЪРСЯВАНЕ ЧРЕЗ ДАННИ ОТ ДИСТАНЦИОННИ НАБЛЮДЕНИЯ НА ЗЕМЯТА: КРИТИЧЕН ПРЕГЛЕД НА ДИГИТАЛНИТЕ ДВОЙНИЦИ, НЕВРОННИ СУРОГАТНИ МОДЕЛИ И ГОЛЕМИ ЕЗИКОВИ МОДЕЛИ

Д. Макариев, П. Тренчев

Резюме

Атмосферното замърсяване е нелинеен, многомащабен процес, повлиян от значителни неопределености, свързани с емисии, метеорология, химични трансформации и ограничения на сензорите. Настоящият обзор синтезира съвременни интелигентни подходи за моделиране на атмосферното замърсяване чрез данни от дистанционни изследвания и наземни наблюдения. Въз основа на предходния обзор от SES-2025 е предложена модулна хибридна рамка, интегрираща цифрови близнаци на земната система и химико-транспортни модели, невронни и Extreme Learning Machine емулатори, символна регресия, Монте Карло квантификация на неопределеността и grounded LLM интерфейси. Специално внимание е отделено на преобразуването на сателитни колонни измервания към приземни концентрации, пространствено-времевото несъответствие между сателитни и наземни данни, метеорологичното влияние, неопределеността в емисионните инвентари и устойчивостта на моделите при екстремни епизоди на замърсяване. Разгледани са съвременни фундаментални модели за наблюдение на Земята, включително TerraMind, Copernicus-FM, Aurora, SkySense++ и Panopticon. Copernicus Observia AI е анализирана като нов бета-интерфейс за достъп до данни от Copernicus, като ясно се разграничават документираните възможности от предложените бъдещи интеграции.