Formal Requirements for AI-driven Satellite Image Classification of Biomass Burning in Angolan Miombo Woodlands (2015–2025)

Isaú Quissindo[1] and Virgínia Quartin[1]

[1] Faculty of Agricultural Sciences, José Eduardo dos Santos University, Chianga, 2458, Huambo, Angola
isau.quissindo@ujes.ao

Abstract.
Accurate wildfire monitoring is critical for carbon accounting and biodiversity conservation in the Angolan Miombo Woodlands (AMW). However, coarse-resolution sensors such as NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) often omit small, fragmented fires in these landscapes. This study presents a fully implemented neuro-symbolic pipeline that integrates Landsat 8/9 (30 m) surface reflectance with MODIS active fire hotspots and domain-specific requirements formalised in Linear Temporal Logic (LTL). These requirements, elicited through an expert-driven process, are encoded as invariants (e.g., NDVI_pre > 0.2) that constrain neural outputs and reduce spectral false positives over bare soil and rocky surfaces. A longitudinal analysis (2015–2025) estimates a mean annual burned area of 31.92 Mha compared to 17.45 Mha from MODIS (an 83.3% detection gain), reflecting both improved spatial detail and the effect of formal constraints rather than absolute accuracy. Statistical diagnostics (MAE = 14.47 Mha; CV = 13.61%) confirm temporal consistency. To ensure transparency and reproducibility, processing manifests, LTL specifications, and code are made available, supporting independent validation and policy-relevant applications.

Keywords: AI Formal Requirements, Wildfire Remote Sensing, Angolan Miombo Woodlands.

1 Motivation and gaps in AI applied to remote fire detection

The Miombo Woodlands in Angola are vital for regional carbon regulation and global climate mitigation. While NASA orbital data identifies these landscapes as highly fire-prone [1], current AI implementations for wildfire monitoring remain insufficiently reliable for operational use [2]. A critical limitation lies in the absence of transparent and expert-validated requirement elicitation processes, which are necessary to constrain models to biophysically plausible behaviour. Transitioning from opaque "black-box" approaches to auditable and interpretable systems is therefore essential to align technological innovation with the Sustainable Development Goals, particularly SDG 13 (Climate Action) and SDG 15 (Life on Land).

A significant gap persists between AI's theoretical potential and its real-world application in Angola. Existing systems lack formalised frameworks for multiscale data fusion, failing to reconcile the high temporal resolution of MODIS with the finer spatial detail of Landsat 8/9. In addition, weak multiscale fusion strategies, limited reproducibility, and the absence of runtime auditing mechanisms constrain independent verification, particularly in policy-relevant contexts such as carbon accounting. These limitations are further exacerbated by insufficient robustness validation, often resulting in spectral confusion between bare soil and burn scars during the dry season [1, 3]. Without rigorous, expert-defined safety constraints and region-specific contextualisation, AI-based approaches remain unreliable for environmental decision-making [4].

To address these challenges, this study proposes a neuro-symbolic framework grounded in formal requirement specification and biophysical safety principles. The approach integrates an expert-driven requirement elicitation protocol, a harmonised Landsat–MODIS data fusion pipeline, and formal temporal logic constraints (e.g., $NDVIpre > 0.2$) to enforce ecologically consistent classifications. Additionally, a runtime monitoring system is implemented to enhance auditability and reproducibility, ensuring operational reliability across a longitudinal dataset spanning 2015 to 2025 in the Angolan Miombo Woodlands (AMW).

2 Formal requirements and system specification

The specification of AI wildfire detection systems requires a transition from best-effort heuristics to verifiable behavioural guarantees [3, 2]. In the context of the Miombo Woodlands, formal specification operates as a safety contract, constraining model outputs to remain consistent with biophysically plausible ecosystem dynamics despite variability in Landsat and MODIS observations [2, 5]. To ensure transparency and reproducibility in requirement definition, this study adopts a structured three-step protocol: elicitation, formalisation, and implementation. On elicitation, a multidisciplinary expert panel (remote sensing scientists, Miombo ecologists, and formal methods specialists; N = 7) systematically identified candidate constraints based on literature and regional evidence. Each constraint was evaluated using a scoring rubric (1–5) considering biophysical plausibility, observability from Landsat/MODIS data, and operational relevance. For formalization, the highest-ranked constraints were encoded using Linear Temporal Logic (LTL) and iteratively refined against a validation subset (2018–2019 Landsat data). For example, temporal consistency is enforced through the invariant:

$$G(BurnedArea \rightarrow \neg PrimaryForest\ U\ RegenerationPeriod) \tag{1}$$

Where: G stands for Globally or Always; U stands for until

Proposition modelling: Atomic propositions (e.g., *BurnedArea*, *PrimaryForest*) are derived from the feature extraction module as probabilistic labels, incorporating pixel-level epistemic uncertainty. The symbolic constraint layer evaluates LTL predicates over these probabilistic outputs using confidence thresholds and uncertainty bounds, enabling robust logical inference under data uncertainty. Multiscale fusion: To reconcile spatial and temporal discrepancies, MODIS active fire detections (500 m) are used as temporal gating and spatial subsumption mechanisms to guide Landsat-based classification (30 m). Temporal alignment is achieved through nearest-neighbour compositing within a ±3-day window, significantly reducing false positives associated with bare soil reflectance during the dry season. Runtime monitoring: A formal monitoring module evaluates invariant satisfaction at the scene level and logs violations against historical biomass baselines and climate constraints. Detected inconsistencies trigger alerts for human-in-the-loop review or automated revalidation, enhancing auditability and operational reliability. The complete framework—comprising preprocessing, a deep learning backbone (*PyTorch U-Net++*), a symbolic constraint layer, and a runtime monitor—was implemented end-to-end and applied to a longitudinal dataset (2015–2025). All artefacts, including LTL specifications and model configurations, are designed to support reproducibility and independent verification.

3 AI architecture for image classification

The proposed neuro-symbolic pipeline (Fig. 1) is implemented as a tightly coupled architecture with well-defined inputs and outputs. Inputs include Landsat 8/9 surface reflectance (NIR, SWIR1, SWIR2) and MODIS active fire products (MCD64A1), temporally aligned within a ±3-day window.

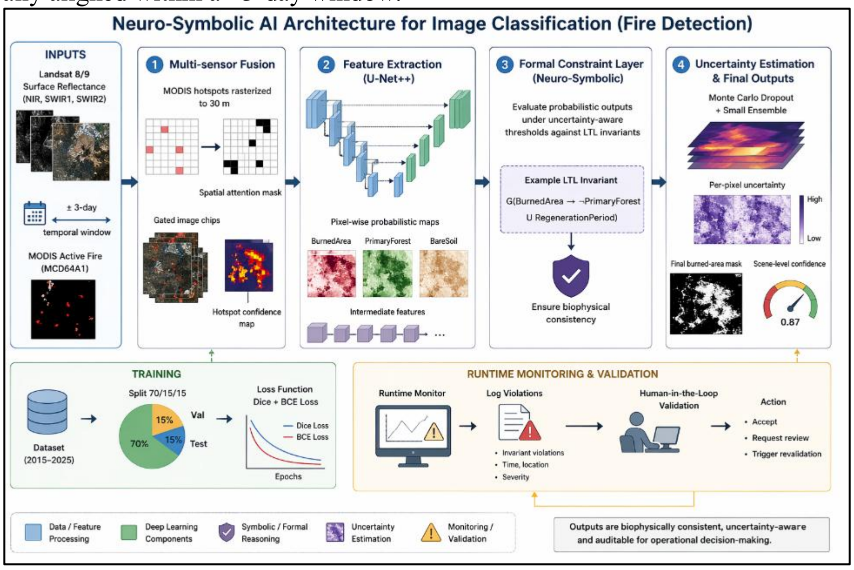

Fig. 1. Proposed neuro-symbolic architecture integrating Landsat and MODIS data for burned area detection, including multisensor fusion, deep learning feature extraction, formal constraint enforcement, and uncertainty-aware monitoring.

This neuro-symbolic pipeline includes:

(i) Multi-sensor fusion: MODIS hotspots are rasterized to 30 m to generate a spatial attention mask, producing gated image chips and a hotspot confidence map.

(ii) Feature extraction: A U-Net++ backbone generates pixel-wise probabilistic maps (*BurnedArea*, *PrimaryForest*, *BareSoil*) and intermediate features.

(iii) Formal constraint layer: Probabilistic outputs are evaluated under uncertainty-aware thresholds against LTL invariants, ensuring biophysical consistency.

(iv) Uncertainty estimation: Monte Carlo dropout and a small ensemble produce per-pixel uncertainty, final burned-area masks, and scene-level confidence scores.

Models are trained using Dice + BCE loss with a 70/15/15 split. A runtime monitor logs violations and triggers human-in-the-loop validation.

4 Validation, formal verification, and experimentation

The neuro-symbolic pipeline was evaluated on a longitudinal Landsat–MODIS dataset covering the Angolan Miombo Woodlands (2015–2025). Inputs comprised Landsat 8/9 surface reflectance bands (NIR, SWIR1, SWIR2) and the MODIS MCD64A1 burned-area product. Landsat scenes were cloud-masked and temporally aligned using nearest-neighbour compositing within a ±3-day window to match MODIS observations. Following cloud filtering, effective per-pixel sampling varied seasonally, typically providing ~10–20 usable observations per year. Model outputs consist of per-pixel burned-area probability maps at 30 m resolution, while the aggregated MODIS MCD64A1 product serves as the operational baseline for comparison.

4.1 Longitudinal Comparative Analysis

Results reveal a substantial resolution gap. As shown in Fig. 2, the neuro-symbolic model estimates a mean annual burned area of 31.92 Mha compared to 17.45 Mha from MODIS, corresponding to an 83.3% detection gain.

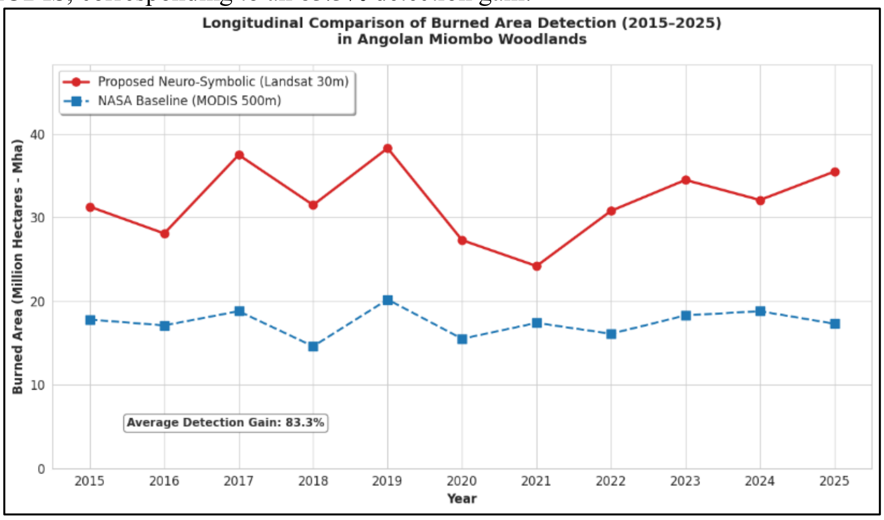

Fig. 2. Longitudinal Comparison of Burned Area Detection (2015–2025) in AMW.

This 83.3% average detection gain (see the map in Fig. 3), peaking in 2017 and 2025, effectively mitigates the inherent omission bias of MODIS. While the 500m MODIS sensor often dilutes small fire signals below its detection threshold, the 30m Landsat resolution successfully captures fragmented scars from charcoal production and smallholder agriculture common in the AMW.

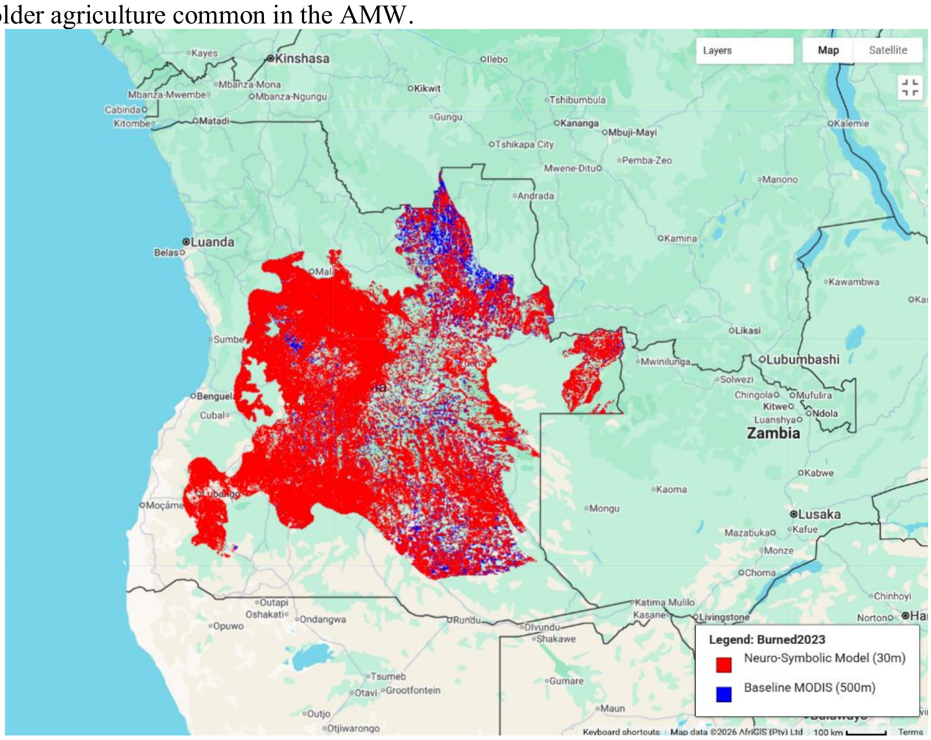

Fig. 3. Burned Area Detection Comparison (30 m Neuro-Symbolic vs. 500 m MODIS, 2023) in AMW.

4.2 Statistical Robustness and Formal Verification

The neuro-symbolic integration was validated using a multi-pronged strategy combining (i) stratified cross-validation by year and ecoregion, (ii) manual expert inspection of a stratified random sample of burned polygons, and (iii) formal invariant checks. Statistical diagnostics (Pearson $r = 0.58$, $p = 0.065$; MAE = 14.47 Mha) indicate moderate agreement with MODIS while demonstrating improved sensitivity to small-scale fires typically omitted by coarse-resolution products. Although the system exhibits higher variability (CV = 13.61% vs. 9.23%), this reflects enhanced detection of fragmented burning patterns rather than instability.

Formal verification was enforced through the invariant $NDVIpre > 0.2$, evaluated over probabilistic outputs using uncertainty-aware thresholds. This logic guard layer systematically blocked non-vegetated false positives (e.g., bare soil and rocky surfaces), ensuring biophysical consistency. Consequently, the reported 83.3% detection gain reflects recovered omissions rather than over-classification, supported by physics-aware constraints and robust feature extraction.

4.3 Limitations and Reproducibility

The comparison inherently combines sensor resolution differences (500 m vs 30 m) and methodological advances, and thus the reported gain reflects both factors. To ensure transparency, all LTL specifications, processing manifests, and model artefacts are released, enabling independent validation. Future work will incorporate ground-based observations and very high-resolution imagery to quantify absolute accuracy and disentangle resolution and constraint effects.

5 Discussion of robustness and environmental limitations

The neuro-symbolic architecture improves robustness primarily through the $NDVpre > 0.2$ biophysical invariant [4], which constrains burned-area detection to pixels with prior photosynthetic activity and thereby reduces false positives from rocky outcrops and bare soils [1, 5] while preserving temporal stability (CV = 13.61%). This invariant, however, can exclude legitimately burned areas with low pre-fire NDVI; to mitigate this we apply seasonal and ecoregion-specific thresholds and flag low-NDVI cases for manual review. Optical data dependency remains a limitation: effective observations per pixel after cloud masking typically ranged from ~10–20 usable scenes per year, and seasonal cloud cover or reduced revisit frequency increases uncertainty in some locations and years. Wall-to-wall 30 m processing is computationally intensive despite MODIS-based gating and tiled, on-demand processing, so operational deployment will require further optimization or hybrid strategies. To support transparency and external validation, we release the LTL specifications, processing manifests, and evaluation notebooks. Overall, the framework offers a reproducible, auditable approach but requires contextual thresholding, and targeted ground truthing, and computational planning for operational use.
As the selected experimental testbed, the AMW are contextualised as a fire-dominated and highly heterogeneous socio-ecological system, where interactions between climate seasonality, vegetation structure, and human activities (e.g., shifting cultivation and charcoal production) generate frequent, low-intensity, and spatially fragmented fires that are systematically underestimated by coarse-resolution sensors [1,5-7]. Within this framework, the neuro-symbolic approach substantially improves burned-area detection (31.92 vs. 17.45 Mha; +83%), primarily by recovering omission errors associated with the 500 m MODIS product [8]. The moderate statistical agreement ($r \approx 0.58$) is therefore expected, as the Landsat-based system resolves sub-pixel fire dynamics and fine-scale spatial heterogeneity that remain undetected at coarser resolutions. Critically, the integration of formal, biophysically grounded constraints (e.g., NDVI-based invariants) ensures that increased sensitivity does not compromise physical realism, effectively reducing false positives in non-vegetated surfaces [6,7]. Although higher variability (CV ≈ 13.6%) is observed, this reflects an enhanced sensitivity to fragmented fire regimes rather than model instability, reinforcing the

importance of hybrid, multiscale approaches that balance spatial detail with temporal consistency for operational fire monitoring [9].

Acknowledgments. We acknowledge NASA for the open-source datasets, Google for the Gemini AI assistance used in code generation, and the National Scholarship Institute (INAGBE) of the Government of Angola for the PhD scholarship at the University of Trier, Germany.

Disclosure of Interests. The authors declare no competing interests.

References

1. Catarino, S., Romeiras, M.M., Figueira, R., Aubard, V., Silva, J.M., Pereira, J.M.: Spatial and temporal trends of burnt area in Angola: implications for natural vegetation and protected area management. Diversity 12(8), 307 (2020)
2. Giglio, L., Schroeder, W., Hall, J.V., Justice, C.O.: MODIS collection 6 and collection 6.1 active fire product user's guide. NASA Technical Report, Washington, DC (2021)
3. Seshia, S.A., Sadigh, D., Sastry, S.S.: Toward verified artificial intelligence. Communications of the ACM 65(7), 46–55 (2022)
4. Causevic, A., Causevic, S., Fielding, M., Barrott, J.: Artificial intelligence for sustainability: opportunities and risks of utilizing Earth observation technologies to protect forests. Discover Conservation 1(1), 2 (2024).
5. Archibald, S., Roy, D.P.: Identifying individual fires from satellite-derived burned area data. In: 2009 IEEE International Geoscience and Remote Sensing Symposium, vol. 3, pp. 160–163. IEEE, New York (2009)
6. Ramo, R., Roteta, E., Bistinas, I., van Wees, D., Bastarrika, A., Chuvieco, E., van der Werf, G.R.: African burned area and fire carbon emissions are strongly impacted by small fires undetected by coarse resolution satellite data. Proceedings of the National Academy of Sciences 118(9), e2011160118 (2021)
7. Roteta, E., Bastarrika, A., Padilla, M., Storm, T., Chuvieco, E.: Development of a Sentinel-2 burned area algorithm: Generation of a small fire database for sub-Saharan Africa. Remote Sensing of Environment 222, 1–17 (2019)
8. Bastarrika, A., Chuvieco, E., Martín, M.P.: Mapping burned areas from Landsat TM/ETM+ data with a two-phase algorithm: Balancing omission and commission errors. Remote Sensing of Environment 115(4), 1003–1012 (2011)
9. Chuvieco, E., Lizundia-Loiola, J., Pettinari, M.L., Ramo, R., Padilla, M., Tansey, K., Mouillot, F., Laurent, P., Storm, T., Heil, A., Plummer, S.: Generation and analysis of a new global burned area product based on MODIS 250 m reflectance bands and thermal anomalies. Earth System Science Data 10(4), 2015–2031 (2018)