Journal of Engineering, Project, and Production Management
2026, 16(4), 2026-0020

Substation Surroundings Hazard Detection and Change Monitoring Using Multi-Source Satellite Remote Sensing and Deep Learning

Zhi Yang[1], Jun Coa[2], Yujia Wang[3], Liming Tao[4], and Lei Coa[5]
[1] Assistant Engineer, Power Transformation Department, Guiyang Power Supply Bureau, Guiyang, Guizhou,550004, China, E-mail: yangz0912@gzgy.csg.cn
[2] Senior Engineer, Power Transformation Department, Guiyang Power Supply Bureau, Guiyang, Guizhou,550004, China, E-mail: dqhp1794@outlook.com (corresponding author).
[3] Engineer, Production Technology Department, Guiyang, Guizhou,550004, China, E-mail: 765142348@qq.com
[4] Senior Engineer, Production Technology Department, Guiyang, Guizhou,551417, China, E-mail: 78180764@qq.com
[5] Senior Engineer, Production Technology Department, Guiyang, Guizhou,563000, China, E-mail: caolei@gzdy.csg.cn

Project Management
Received Apr 22, 2026; revised May 14, 2026; accepted May 21, 2026
Available online May 29, 2026

Abstract: Improving high-voltage substation safety and reliability requires continuous monitoring of spatially heterogeneous and temporally persistent environmental hazards. Existing approaches rely on single-source satellite imagery or isolated Unmanned Aerial Vehicle (UAV) inspections and typically decouple continuous hazard estimation from categorical risk assessment. It remains unclear how to design a unified spatio-temporal framework that simultaneously reconstructs hazard intensity fields and generates operationally consistent risk levels under multi-source sensing uncertainty. This study develops a hybrid deep spatio-temporal architecture that integrates multispectral satellite data, Synthetic Aperture Radar (SAR) observations, UAV-derived anomaly cues, and environmental forcing variables within a shared regression–classification backbone for substation-surroundings hazard monitoring. The framework was trained on 2.8 million spatial samples from 2021 to 2023 and optimized via regularized stochastic gradient updates with temporal encoding and multi-source fusion. On independent tests, the model achieves a root mean square error of 0.113, a mean absolute error of 0.075, and a coefficient of determination of 0.905, explaining over 90 percent of spatial hazard variance. Risk-level classification attains 91.2 percent accuracy, a Critical-class F1-score of 0.94, and a false negative rate below 6 percent. Relative to the strongest baseline, hazard root mean square error decreases by 23 percent, and cumulative operational risk cost decreases by 12 to 16 percent compared with deep reinforcement learning alternatives. The framework provides a unified, deployable solution for proactive hazard detection, enabling early warning with lead times exceeding 30 hours before confirmed field events.

Keywords: Multi-source remote sensing data fusion; spatio-temporal hazard prediction; hybrid regression classification learning; infrastructure risk assessment; substation environmental monitoring.

Copyright © Journal of Engineering, Project, and Production Management (EPPM-Journal).
DOI 10.32738/JEPPM-2026-0020

1. Introduction

The operational reliability of high-voltage substations depends on continuous characterization of hazards surrounding critical assets. Advances in satellite remote sensing and Unmanned Aerial Vehicle (UAV) inspection have expanded near-real-time observation of environmental and infrastructure risks, forming the basis of contemporary substation risk assessment (Kopec et al., 2025). The increasing volume of data necessitates combining multi-temporal satellite imagery, UAV cues, and environmental variables, which presents structural challenges. Different sensing methods work at varying spatial and temporal resolutions, leading to misalignment in combined data, which weakens hazard-field reconstruction and reduces consistency in risk indicators. Most monitoring pipelines separate continuous hazard estimation from categorical risk inference, limiting the generation of calibrated operational scores. Many frameworks assume stable sensor performance, which prevents uncertainty from affecting decision metrics. Scalability is a challenge because hazard characterization must work reliably in different environments and hazard types (Siddique and Ahmed, 2025). These limitations highlight the need for coherent integration of spatio-temporal information and risk translation mechanisms.

Research on substation-surroundings hazards spans five major domains, although most studies isolate change detection, deformation tracking, anomaly extraction, or risk translation. Integrated spatio-temporal architectures capable of jointly estimating hazard intensity and operational risk under multi-source uncertainty remain limited. In satellite-based hazard analysis, Kopec et al. (2025) introduced a deep ensemble for multi-temporal Sentinel-2 change detection, while domain shifts reduced robustness. Liu et al. (2026) improved landslide mapping using enhanced Sentinel-2 sequences, although illumination and vegetation recovery affected signatures. Deng et al. (2025) applied recurrent predictors to interferometric deformation histories, yet short, noisy series limited stability. Chang et al. (2025) embedded deep learning into interferometric processing to track continuous deformation, though decorrelation disrupted temporal consistency. Chen et al. (2025) fused radar and optical data via attention mechanisms for vegetation-structure mapping, with uncertainty increasing under sparse reference data. These contributions strengthen hazard-field reconstruction but do not connect intensity estimation with calibrated risk levels.

Fire and vegetation related hazards near power infrastructure represent a second area. Fu and Fang (2025) proposed high-resolution burned-area segmentation, though smoke and shadow artifacts reduced accuracy. Beyki et al. (2025) integrated active-fire satellite data into wildfire spread models, yet overpass gaps weakened forecasts. Wang et al. (2026) fused spatio-temporal features to detect early fire spots along transmission corridors, while regional generalization remained uncertain. Dubey and Dubey (2025) combined wildfire prediction with fuzzy decision reasoning, although rare extremes limited calibration. These works focus primarily on wildfire rather than multi-hazard substation environments. UAV and proximal sensing provide fine-scale anomaly extraction. Song et al. (2025) developed a lightweight fusion detector for small power-line defects, though robustness to unseen viewpoints remained limited. Song et al. (2026) proposed a siamese anomalous-change detector for transmission scenes; however, seasonal illumination variations induced false alarms. Chen et al. (2025) extracted surge arresters from UAV LiDAR using structural cues, while dense assemblies produced misclassification. These approaches improve inspection automation but remain separate from large-scale hazard-field modeling.

Autonomous aerial coordination and learning-based navigation form a fourth line of defense. formulated spatio-temporal path planning for inspection tasks, although long horizon scalability remained unresolved. These studies enhance operational autonomy but rarely link navigation outcomes with downstream hazard estimation. Risk assessment and decision-support systems represent a final direction. Cong et al. (2026) implemented an edge-AI corridor monitoring pipeline with adaptive feedback but lacked standardized datasets.

This study presents an integrated framework for hazard detection, spatio-temporal monitoring, and multi-source risk scoring around high-voltage substations by combining Satellite Remote Sensing and electric UAV sensing within a Long Short-Term Memory-Multilayer Perceptron (LSTM-MLP) deep learning architecture. The model uses hybrid regression, classification learning, temporal fusion of satellite, UAV, and environmental signals, and risk-driven optimization with stability constraints. Multi-seasonal validation in eastern China confirms improved hazard detection and risk scoring accuracy.

The research questions addressed in this study are as follows.

1. Can hazard intensity around substations be continuously predicted by integrating multi-source data?
2. Is a hybrid regression-classification framework capable of providing a balanced and accurate assessment of both hazard intensity and operational risk levels?
3. How does this model perform in terms of stability and accuracy compared to separate hazard and risk models under multi-source uncertainty?

The proposed hybrid regression-classification model is designed to meet the requirements of spatio-temporal representation learning and risk translation. This architecture integrates LSTM for temporal encoding and multi-task learning for continuous hazard intensity regression and categorical risk classification, leveraging both spatial and temporal patterns to provide a robust and accurate risk assessment.

2. Method

2.1. System Architecture for Multi-Source Remote Sensing-based Substation Monitoring

The monitoring system in Fig. 1 represents the high-voltage substation as a protected infrastructure node surrounded by a circular surveillance buffer. Satellite imagery, SAR observations, UAV sensing, and environmental variables are integrated to characterize vegetation growth, land-use transitions, thermal anomalies, and structural encroachment. The framework contains three functional layers: multi-source data acquisition, spatio-temporal feature extraction, and hazard-risk inference for operational monitoring.

Let the monitored domain be defined as $\Omega = \{x \in \mathbb{R}^2 \mid\mid x - x_s \mid\mid \leq R_b\}$, where $x_s$denotes the geographical coordinate of the substation. The monitored domain Ω is defined around the substation location x_s. For each spatial coordinate $x \in \Omega$, multi-spectral satellite observations at time t are expressed as Eq. (1).

$$I_{\text{sat}}(x, t) \in \mathbb{R}^B \tag{1}$$

In Eq. (1) where B denotes the number of spectral bands. Vegetation conditions relevant to encroachment risk are represented using the normalized spectral index (Eq. 2).

$$V(x,t) = \frac{I_{\text{sat}}^{\text{NIR}}(x,t) - I_{\text{sat}}^{\text{RED}}(x,t)}{I_{\text{sat}}^{\text{NIR}}(x,t) + I_{\text{sat}}^{\text{RED}}(x,t)} \tag{2}$$

In Eq. (2) where $I_{\text{sat}}^{\text{NIR}}$ and $I_{\text{sat}}^{\text{RED}}$ denote the near-infrared and red reflectance components.

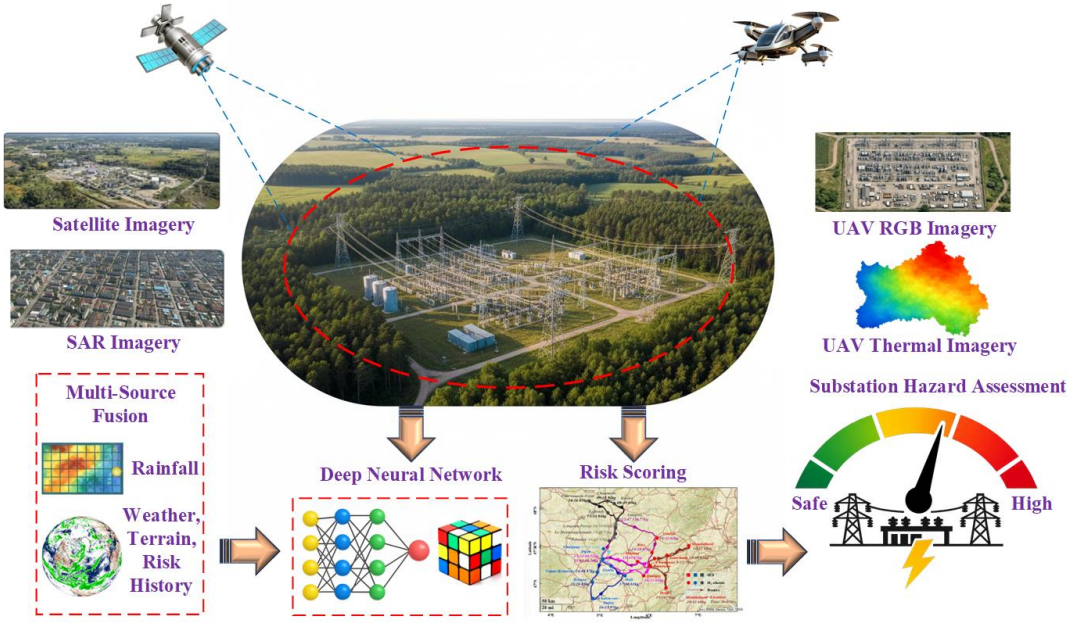


Fig. 1. Multi-source satellite and UAV framework for substation hazard monitoring

High-resolution electric UAV platforms complement satellite sensing by capturing localized RGB and thermal observations inside the buffer region Ω. At time t, the UAV acquires the observation field seen in Eq. (3).

$$I_{\text{uav}}(x,t) \in \mathbb{R}^{C} \tag{3}$$

In Eq. (3), where C is the number of UAV sensing channels. Localized anomaly responses extracted from UAV imagery are computed as Eq. (4).

$$A_s(x,t) = \Psi\big(I_{\text{uav}}(x,t)\big) \tag{4}$$

In Eq. (4), with $\Psi(\cdot)$ representing a convolutional feature extraction operator. These measurements are bound by operational and regulatory sensing constraints that account for saturation, occlusion, and mission-duration limitations.

To characterize hazard evolution, satellite-derived and UAV-derived indicators are fused to obtain (Eq. (5))

$$H(x,t) = \gamma_1 H_s(x,t) + \gamma_2 A_s(x,t) \tag{5}$$

In Eq. (5), where $H_s(x,t)$ is the satellite-based hazard component and γ_1, γ_2 are fusion parameters.

Monitoring reliability is described by the confidence field $C(x,t) \in [0,1]$, which must satisfy Eq. (6).

$$C(x,t) \geq C_{\min} \tag{6}$$

In Eq. (6), ensuring that essential hazard signals remain detectable within the protection zone. Environmental forcing conditions, including rainfall, wind, and soil moisture, are aggregated into Eq. (7).

$$F_{\text{env}}(t) = \eta_1 R(t) + \eta_2 W(t) + \eta_3 S(t) \tag{7}$$

In Eq. (7), where $R(t)$, $W(t)$, and $S(t)$ denote precipitation, wind magnitude, and soil moisture indicators. The integrated spatial risk must remain below a predefined tolerance threshold to maintain substation operational reliability.

2.2. Spatio-Temporal Learning and Risk-Driven Optimization Framework

The proposed monitoring architecture models environmental risk evolution at the next time step using current multi-source observations and external environmental forcing variables. Historical hazard trajectories affect the system only through extracted temporal features, avoiding direct dependence on control signals. This conditional structure is represented as a stochastic state-transition process over a spatio-temporal hazard field, allowing the monitoring and inference mechanism to operate within a sequential decision-learning framework that captures dynamic risk evolution around the substation (Duan et al., 2026).

Within this framework, the decision entity is an intelligent monitoring engine that performs adaptive hazard assessment and multi-source data fusion. It interacts with spatial observations, including satellite-derived feature fields, UAV-derived anomaly maps, environmental forcing signals, and indicators of historical hazard persistence, to update risk estimates and reduce long-term safety exposure. At time t, the system state is defined in Eq. (8).

$$s_t = (H(x,t), \Delta H(x,t), F_{\text{env}}(t), C(x,t)) \tag{8}$$

Eq. (8), which combines the current hazard intensity, its temporal variation, environmental forcing factors, and sensing reliability.

In Eq. (9), the global hazard intensity across the monitored region is computed as Eq. (9).

$$H_{\text{tot}}(t) = \int_{\Omega} H(x,t)\, dx \tag{9}$$

The overall substation risk index is expressed as Eq. (10).

$$R_{\text{sys}}(t) = H_{\text{tot}}(t) + R_{\text{env}}(t) \tag{10}$$

In Eq. (10), risk evolution is evaluated over a finite decision horizon of length T. At each discrete time step t, the monitoring engine updates fusion parameters and risk thresholds using the observed system state. Continuous interaction between sensing processes, inference mechanisms, and environmental forcing generates a temporal sequence of hazard fields that guides adaptive decision making. The objective is cumulative risk minimization under operational safety constraints, formulated as the expected discounted risk exposure over the monitoring horizon (Eq. (11).

$$\min_{\pi \in \Pi} J(\pi) = \mathbb{E}_{\pi}\left[\sum_{t=0}^{T} \gamma^t R_{\text{sys}}(t)\right] \tag{11}$$

Eq. (11) yields an optimization policy that minimizes long-term substation risk while accounting for evolving environmental conditions and sensing uncertainty.

2.3. Deep Hybrid-Action Learning Framework for Risk-Aware Monitoring

The learning framework maps heterogeneous spatio-temporal observations to two coupled outputs: a continuous hazard intensity field and a discrete risk label. Satellite imagery, UAV anomaly maps, environmental forcing, and temporal context are processed through a shared LSTM–MLP backbone with separate regression and classification heads. The joint predictive mapping is expressed as Eq. (12) (Siddique and Ahmed, 2025).

The hybrid regression-classification model is theoretically grounded in spatio-temporal representation learning and risk translation, which ensures the model captures temporal persistence and adapts to evolving environmental conditions effectively.

$$(\widehat{H}(x,t), \hat{S}(x,t)) = f_{\theta}\left(I_{\text{sat}}(x,t), I_{\text{uav}}(x,t), F_{\text{env}}(t), \mathcal{T}(t)\right) \tag{12}$$

In Eq. (12), where $f_{\theta}(\cdot)$denotes the deep hybrid predictor and $\mathcal{T}(t)$represents the encoded temporal context.

The cumulative reward represents the total discounted reward over time, reflecting the effectiveness of the policy in reducing operational risk. It is calculated in Eq. (13).

$$\text{Cumulative Reward} = \sum_{t=0}^{T} \gamma^t R_{\text{sys}}(t) \tag{13}$$

In Eq. (13), where $\gamma$is the discount factor and $R_{\text{sys}}(t)$is the risk cost at time t. The operational risk cost measures the cost of managing risks, calculated as the sum of risk values over time.

$$\text{Operational Risk Cost} = \int_{\Omega} R_{\text{sys}}(t)\, dt \tag{14}$$

As shown in Fig. 2, the LSTM module encodes temporal persistence from historical hazard sequences, while the MLP backbone combines these embeddings with current satellite, UAV, and environmental descriptors. The regression head estimates $\widehat{H}(x,t)$ and the classification head generates $\hat{S}(x,t)$ (Deng et al., 2025). Algorithm 1 summarizes the complete inference and training procedure.

Key parameters, including LSTM hidden units, MLP layers, learning rate (0.001), and weight decay (10^{-4}), were optimized based on cross-validation to balance model performance and prevent overfitting. Multi-source data (satellite, UAV, environmental variables) are fused using weighted parameters ($\gamma 1$, $\gamma 2$), optimized for effective hazard intensity and risk classification. Temporal encoding aligns multi-temporal data, ensuring robustness under varying conditions.

3. Results

3.1. Learning Dynamics and Convergence Behavior

To assess the proposed hybrid spatio-temporal hazard prediction framework, a real-world case study was conducted at a 500 kV transmission substation in suburban Nanjing, Jiangsu Province, China, as shown in Fig. 3. The site, surrounded by mixed agricultural land, seasonal vegetation, and semi-urban development, is suitable for evaluating vegetation encroachment, thermal anomalies, land disturbance, and multi-hazard interactions. The monitoring domain comprised a 3 km circular buffer centered on the substation, consistent with high-voltage safety clearance norms in eastern China (Cong et al., 2026).

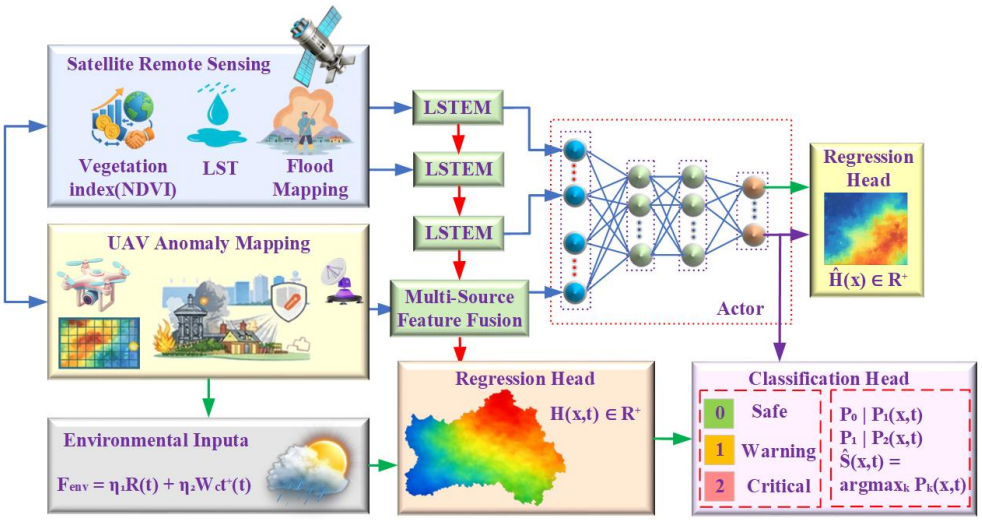


Fig. 2. Hybrid LSTM–MLP architecture for spatio-temporal hazard intensity regression and risk-level classification

Algorithm 1. Hybrid spatio-temporal hazard monitoring procedure

1	Collect satellite imagery, UAV observations, environmental measurements, sensing-confidence information, and historical hazard labels.
2	Define the monitored buffer around the substation and align all input sources in space and time.
3	Extract vegetation, thermal, surface-moisture, land-cover, and anomaly-related features from the aligned observations.
4	Build the monitoring state from current hazard conditions, temporal changes, environmental drivers, and sensing reliability.
5	Encode temporal persistence using the LSTM module over the selected historical observation window.
6	Combine temporal features with spatial descriptors through the shared MLP backbone.
7	Estimate the continuous hazard intensity field using the regression head.
8	Assign the operational risk level using the classification head.
9	Update the trainable model parameters using regularized stochastic gradient optimization with Adam.
10	Evaluate model performance using regression, classification, and spatial-structure preservation metrics.

The dataset covered January 2021 to December 2023 and was constructed from the monitored 3 km buffer around the 500 kV substation. Optical, SAR, thermal, UAV, and environmental observations were combined to represent vegetation encroachment, surface moisture, land-surface temperature, localized thermal anomalies, and short-term weather-driven hazard triggers. Reference hazard maps and severity labels were derived from field records, maintenance logs, and expert interpretation of UAV imagery. The final dataset contained 2.8 million spatial samples, with Critical events representing less than 10% of the data. Table 1 reports the integrated data sources, class distribution, and model training configuration.

The neural network used a 24-hour temporal window to capture hazard persistence and generated a shared embedding for the regression and classification heads. Model training used Adam with weight decay to reduce overfitting, while validation performance was monitored across 150 epochs. Full convergence required at about 4.2 hours on an NVIDIA GTX 1650 GPU, indicating that the proposed framework can be trained under moderate computational resources.

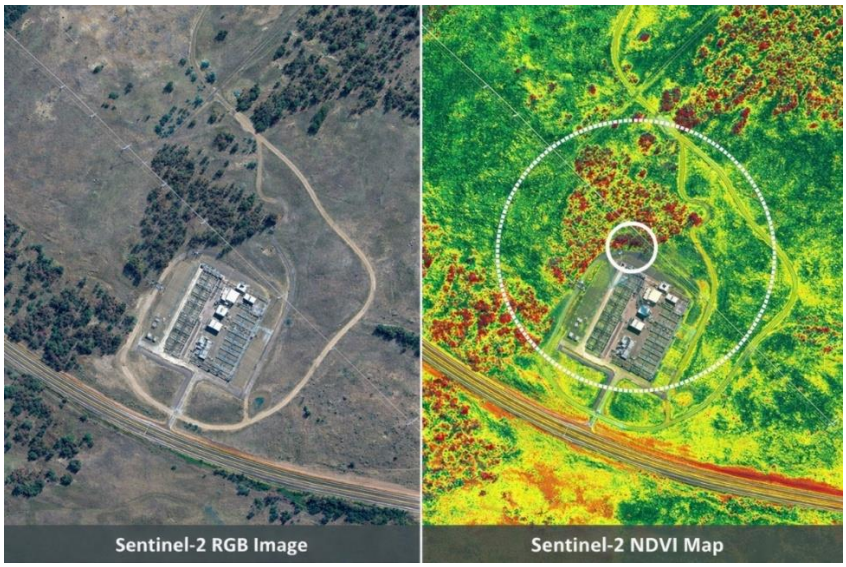
To illustrate the characteristics of the multispectral satellite inputs, Fig. 4 presents representative Sentinel-2 observations used in this study. The RGB image provides a true-color visualization of the substation surroundings, including land-use patterns and infrastructure layout. The corresponding NDVI map highlights vegetation density and spatial variability, which are critical for identifying potential encroachment risks and vegetation-induced hazards. These complementary representations demonstrate the role of multispectral data in capturing both structural and ecological features relevant to hazard modeling. Satellite observations provide broad coverage, whereas UAVs capture localized, high-resolution RGB and thermal data within buffer regions. Their flexible deployment supports precise monitoring of small-scale or transient hazards during critical periods, often missed by satellites.

Fig. 3. 500 kV Transmission substation in suburban Nanjing, Jiangsu Province, China– study site for hazard monitoring

Table 1. Integrated data sources, sample distribution, and training configuration

Category	Component	Compact Specification and Use
Satellite data	Sentinel-2 MSI	Optical multispectral data, 10–20 m, 5-day revisit, 2021–2023; vegetation index and land-cover change
Satellite data	Sentinel-1 SAR	C-band radar, 10 m, 6-day revisit, 2021–2023; surface roughness and moisture sensitivity
Satellite data	Landsat-8 TIRS and GF-2	Landsat-8 thermal infrared, 30 m, 16-day revisit; GF-2 high-resolution optical imagery, 1 m, on-demand
UAV data	Electric quadrotor and onboard sensors	80–120 m flight altitude; monthly and emergency surveys; 20 MP RGB, 5-band multispectral, and 640 × 512 thermal sensing
Environmental forcing	Weather and remote-sensing variables	Air temperature, precipitation, wind speed, soil moisture, NDVI, and land-surface temperature; hourly to 16-day sampling
Dataset labels	Safe, Warning, and Critical	1,960,000 Safe, 560,000 Warning, and 280,000 Critical samples
Dataset split	Training, validation, and testing	2.8 million spatial samples split into 70%, 15%, and 15% subsets
Training setup	Software, optimizer, and model settings	PyTorch 1.11.0; Adam; learning rate = 0.001; weight decay = 1e-4; batch size = 512; epochs = 150
Temporal and hardware setup	Sequence length, embedding, and GPU	24-hour temporal window; embedding dimension = 128; NVIDIA GTX 1650 GPU; convergence in 4.2 hours

UAV RGB and thermal observations were used to detect vegetation growth, land-cover changes, and localized heat anomalies within the buffer zone. The trained model produces continuous hazard maps and categorical risk layers; Fig. 5 shows the convergence of training and validation losses.



(a) Sentinel-2 true-color RGB image illustrating land cover and infrastructure surroundings

(b) NDVI map derived from multispectral bands, highlighting vegetation density and potential encroachment risk near the substation

Fig 4. Representative multispectral satellite observations used in the proposed framework

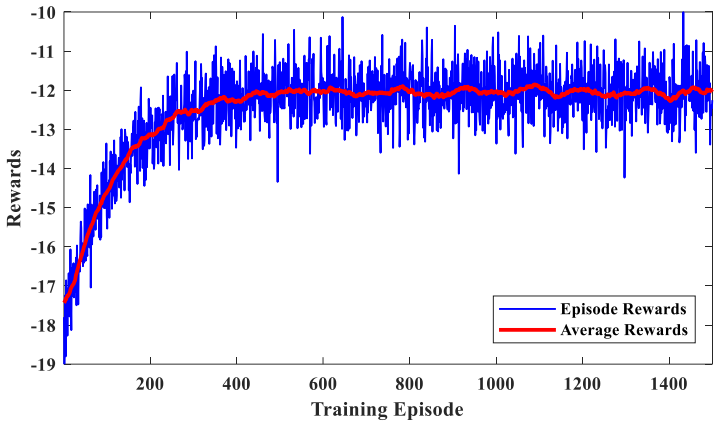


Fig. 5. Training and validation loss convergence of the proposed hybrid spatio-temporal hazard prediction model

3.2. Hazard Intensity Estimation and Risk-Level Classification

The proposed hybrid spatio-temporal model was evaluated using independent test samples covering normal operation, high vegetation growth, and extreme weather. As reported in Table 2, the test set achieved RMSE = 0.113, MAE = 0.075, R^2 = 0.905, and SSIM = 0.921, confirming low prediction error and preserved spatial structure. Fig. 6 shows that high-intensity clusters near vegetation corridors and infrastructure boundaries were reconstructed with limited residual concentration near abrupt surface transitions, reflective surfaces, and dense canopy edges.

Risk-level classification was assessed for the Safe, Warning, and Critical classes. The model reached an overall accuracy of 0.912, with most errors occurring between adjacent severity levels. Table 3 shows that the Critical class achieved precision = 0.938, recall = 0.942, and F1-score = 0.940, while its false-negative rate remained below six percent. Misclassified Critical samples were mainly assigned to Warning rather than Safe, preserving conservative operational behavior. Seasonal testing retained stable accuracy, with RMSE/R^2 values of 0.109/0.912 in spring, 0.121/0.893 in summer, 0.112/0.906 in autumn, and 0.107/0.918 in winter.

Table 2. Regression performance for spatial hazard intensity prediction

Dataset Split	RMSE	MAE	R^2	SSIM
Training	0.089	0.061	0.932	0.948
Validation	0.104	0.071	0.914	0.934
Test	0.113	0.075	0.905	0.921

Fig. 6 compares ground-truth intensity maps, predicted fields, and absolute error distributions. The model accurately reconstructs high-intensity clusters along vegetation corridors and infrastructure boundaries. Errors remain concentrated near abrupt surface transitions and reflective features. Under high-vegetation and typhoon conditions, deviations slightly increase at dense canopy edges due to greater spectral variability. Seasonal robustness was also evaluated across spring, summer, autumn, and winter. The model retained stable accuracy across seasonal subsets, with RMSE/R^2 values of 0.109/0.912 in spring, 0.121/0.893 in summer, 0.112/0.906 in autumn, and 0.107/0.918 in winter. The weakest performance occurred in summer, mainly because of canopy variability and stronger spectral heterogeneity, while the temporal encoder reduced seasonal drift and maintained consistent risk-level transitions.

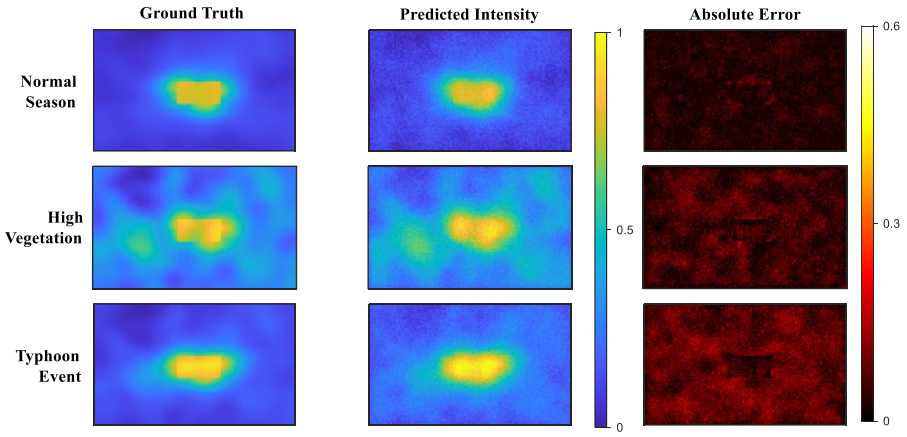


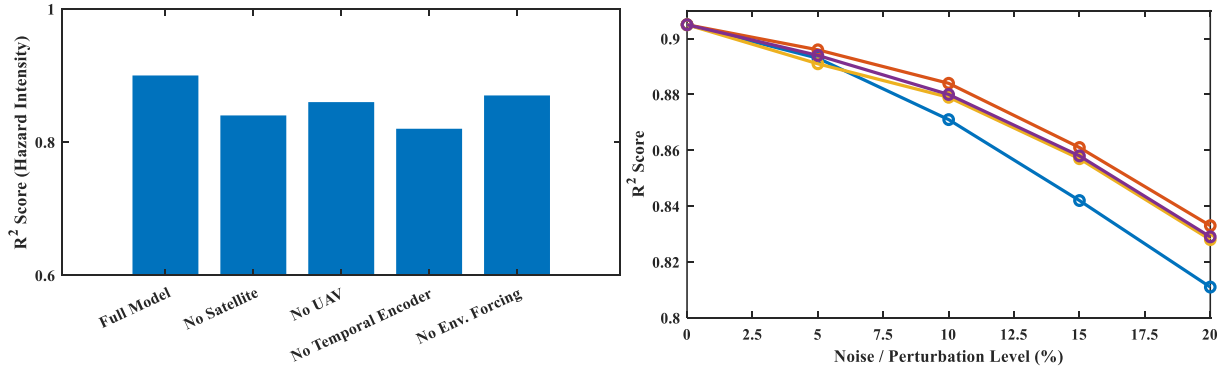
Fig. 6. Ground-truth, predicted, and error maps for hazard intensity

Table 3. Compact risk-level classification performance on the test setTrue Class	Correct Predictions	Main Misclassification Pattern	Precision	Recall	F1-score
Safe	842	37 as Warning; 6 as Critical	0.952	0.952	0.952
Warning	765	41 as Safe; 28 as Critical	0.912	0.918	0.915
Critical	598	4 as Safe; 32 as Warning	0.938	0.942	0.940

3.3. Ablation and Robustness Analysis

The contribution of each information source was examined through ablation analysis. As shown in Fig. 7(a), the full model achieved $R^2 = 0.905$. Removing satellite imagery reduced R^2 to 0.842, confirming the dominant role of multispectral context in vegetation proximity, thermal-anomaly representation, and land-surface variation. Excluding UAV-derived anomaly cues decreased R^2 to 0.861, while removing environmental forcing variables reduced R^2 to 0.874. The strongest degradation occurred when the temporal encoder was removed, with R^2 falling to 0.823, which indicates that temporal-persistence modeling is necessary for separating sustained hazard evolution from short-term sensing noise. Fig. 7(b) presents robustness under progressive data degradation, including Gaussian NDVI noise, thermal band noise, reduced satellite resolution, and partial UAV removal. The model shows monotonic, controlled degradation up to 20 percent perturbation. NDVI noise causes the steepest decline, with R^2 dropping from 0.905 to 0.802, highlighting the significance of vegetation related spectral features. Thermal band noise has milder effects, and resolution loss becomes more impactful beyond 15 percent due to loss

of fine scale structure. Partial UAV removal produces moderate decline, indicating limited dependence on high resolution local sensing.



(a) Source-removal ablation analysis (b) Performance degradation under data perturbation

Fig. 7. Ablation and robustness analysis of the proposed hazard prediction model

3.4. Operational Validation and Computational Feasibility

An extreme event near a 220 kV substation in eastern Zhejiang Province, China, was used to assess operational validity. Two weeks of satellite imagery, UAV observations, and environmental variables were processed by the trained model. Hazard intensity began increasing five days before field confirmation, crossed the warning threshold three days before manual reporting, and reached the critical level about thirty-six hours before on-site intervention. The temporal encoder captured multi-day escalation rather than isolated anomalies, while the risk labels progressed consistently from safe to warning and then to critical, providing approximately thirty hours of actionable lead time. Computational feasibility was evaluated through inference latency, training time, and GPU memory use. Inference required 0.042 seconds (0.042s) per batch for one substation, 0.081s for five substations, and 0.156s for ten substations, showing sub-linear growth because the backbone computation was shared across deployment batches. Training one year of multi-source data required 3.8 hours on an NVIDIA GTX 1650 GPU and increased to 4.6 hours for ten substations without additional memory. Peak GPU usage remained below 5.2 GB. These results indicate that the framework can support near-real-time supervisory monitoring without architectural modification.

It is important to note that the validation of the proposed model was conducted within a geographically specific area, Eastern China, which may limit its generalizability to other regions with differing environmental or infrastructural conditions. The model's performance was evaluated in this region due to its relevance to the high-voltage substation context and the availability of multi-source data. However, future work will explore the transferability of the model to other geographical areas with varying hazard types, climate conditions, and infrastructure types.

4. Discussion

The comparative results confirm that the proposed fusion model improves continuous hazard-intensity estimation relative to conventional and partial deep-learning baselines. As shown in Fig. 8, the proposed model achieved RMSE = 0.073, compared with 0.118 for Random Forest, 0.101 for CNN without temporal encoding, 0.095 for LSTM without spatial fusion, and 0.132 for the single-source CNN. The 23 percent RMSE reduction relative to the strongest baseline indicates that the improvement is not caused only by model depth; it mainly reflects the joint use of spatial continuity, temporal persistence, and multi-source sensing. Random Forest failed to preserve spatial coherence, CNN-only and LSTM-only variants captured only one part of the spatio-temporal structure, and the single-source setting produced the weakest performance. Residual inspection further showed that baseline models underestimated localized thermal anomalies near transformer boundaries, whereas the proposed framework preserved both intensity amplitude and spatial morphology.

The decision-learning comparison provides additional evidence of operational stability. DQN, DDPG, and PPO were evaluated using the same backbone depth and multi-source state representation. DQN suffered from coarse discretization of continuous hazard-intensity adjustments. DDPG showed higher early-stage variance, while PPO had lower sample efficiency due to its on-policy update structure. As shown in Fig. 9(a), the proposed D-PDQN reached a higher steady-state reward, outperforming DDPG, PPO, and DQN by about 14.6 percent, 11.3 percent, and 6.1 percent, respectively. Fig. 9(b) further shows a stable reduction in operational risk-mitigation cost, with cumulative cost reduced by 12-16 percent relative to DDPG and PPO. These results suggest that combining spatio-temporal encoding with hybrid action representation improves prediction accuracy, policy stability, and risk-aware supervisory response under variable environmental conditions. The calculation of operational risk mitigation costs, shown in Fig. 9, follows an optimization framework that minimizes cumulative risk exposure over time. At each time step, the system state is represented by the hazard intensity field $H(x,t)$, its temporal evolution $\Delta H(x,t)$, environmental forcing variables $F_{\text{env}}(t)$, and the reliability of sensing $C(x,t)$, as shown in Eq. (8). The total risk $R_{\text{sys}}(t)$is then computed by aggregating the hazard intensity across the region, as per Eq. (9). The risk mitigation cost is derived from the expected discounted risk over the monitoring horizon, which is formulated as an objective function for optimization, minimizing the operational cost while maintaining operational safety. This is expressed in Eq.

(11), where the goal is to minimize long-term risk exposure through adaptive decision-making that balances hazard detection with proactive risk management actions.

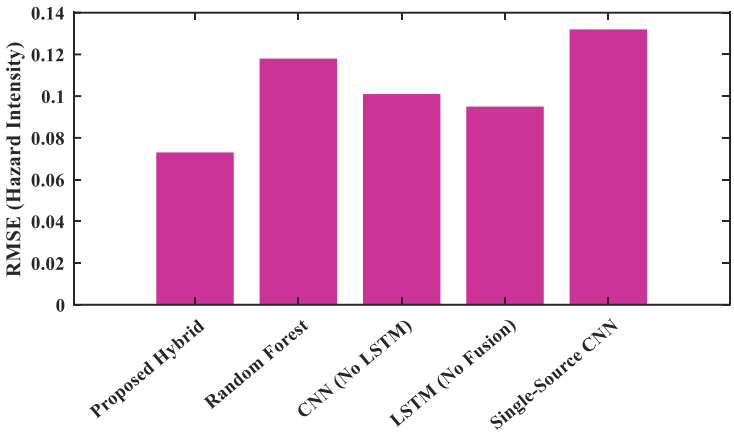
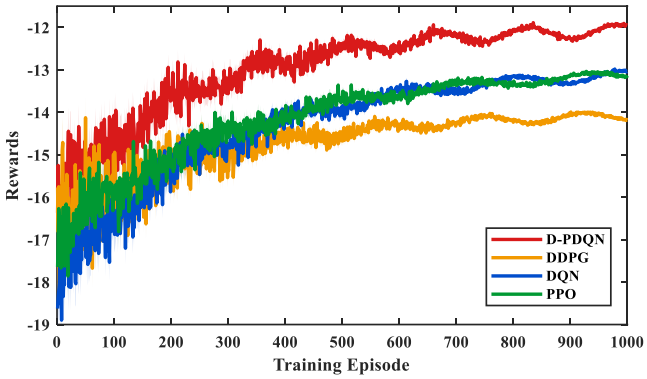
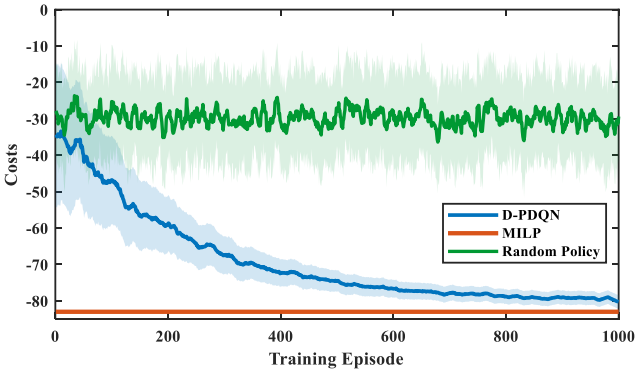


Fig. 8. RMSE-based comparison of continuous hazard-intensity prediction



(a) Cumulative reward convergence



(b) Integrated risk-mitigation cost

Fig. 9. Learning-policy comparison and operational risk mitigation

5. Conclusion

This study developed a hybrid spatio-temporal deep learning framework to address limitations in substation hazard monitoring, where conventional approaches rely on single-source data and separate hazard estimation from risk classification. The proposed model integrates multispectral satellite imagery, radar data, UAV-derived anomaly indicators, and environmental variables within a unified regression–classification structure, enabling the simultaneous prediction of continuous hazard intensity and discrete risk levels with temporal consistency. The model demonstrated strong predictive performance, achieving an RMSE of 0.113, MAE of 0.075, and R^2 of 0.905, explaining over 90 percent of spatial variance. Structural similarity exceeded 0.92, confirming the preservation of the spatial pattern. Risk classification accuracy reached 91.2 percent, with the critical class achieving an F1 score of 0.94 and a false negative rate below 6 percent. Comparative analysis showed improvements over DQN, DDPG, and PPO, with up to 14.6 percent higher cumulative reward and a 12-16

% reduction in operational risk cost. From a managerial perspective, the framework enables a shift from reactive inspection schedules toward anticipatory, risk-driven decision-making. Managers can prioritize assets based on predicted hazard trajectories, allowing earlier intervention, optimized resource allocation, and reduced emergency maintenance. The distinction between continuous hazard levels and discrete risk classes supports tiered responses, improving operational efficiency and reducing system downtime. The observed lead time of over 30 hours before critical events provides a practical window for preventive action. The model was validated in Eastern China, so its transferability to other climates, hazard types, and infrastructure settings requires further testing. Future work will therefore focus on multi-site validation, improved regional transfer, and more interpretable risk attribution.

Author Contributions

Zhi Yang contributed to conceptualization, methodology, investigation, data curation, writing (original draft preparation), and visualization. Jun Cao contributed to supervision, project administration, funding acquisition, conceptualization, and writing (review and editing). Yujia Wang contributed to methodology, software, validation, formal analysis, investigation, and visualization. Liming Tao contributed to software, validation, formal analysis, data curation, and writing (original draft preparation). Lei Cao contributed to investigation, data curation, formal analysis, visualization, and writing (review and editing).

Funding

The authors declare that no funds, grants, or other financial support were received for the conduct of this research or the preparation of this manuscript.

Institutional Review Board Statement

Institutional Review Board (IRB) approval is necessary if manuscripts involve human subjects. Authors are requested to address IRB ID, approval date, and decision in this section. If IRB is not needed, please specify “not applicable.

Declaration of Artificial Intelligence (AI) Tools

The authors used ChatGPT-5.2 solely for language editing and readability improvement. The authors did not use any AI tools for generating scientific content, equations, analyses, or conceptual material. The authors reviewed and verified all content and take full responsibility for the accuracy and integrity of the manuscript.

References

Beyki, S. M., Lopes, A. M. G., Santiago, A., and Laím, L. (2025). Improving wildfire simulation accuracy using satellite active fire data for interval reinitialization and rate of spread adjustment. *Remote Sensing Applications: Society and Environment*, 39(3), 101648. https://doi.org/10.1016/j.rsase.2025.101648

Chang, S. T., Lin, S. Y., and Lin, Y. C. (2025). Deep learning-based surface deformation tracking with interferometric fringes: A case study in Taiwan. *International Journal of Applied Earth Observation and Geoinformation*, 143, 104796. https://doi.org/10.1016/j.jag.2025.104796

Chen, M., Deng, L., Mu, F., Xue, M., Ma, H. and Ji, C. (2025). Extraction of line Surge Arresters from UAV LiDAR point clouds based on multi-view structural features. *International Journal of Applied Earth Observation and Geoinformation*, 144, p.104950.

Deng, Z., Wang, W., Luo, Y., Zhang, S., Xu, L., and Su, Q. (2025). Slope deformation monitoring and prediction based on InSAR and deep learning model. *Results in Engineering*, 28, 107113. https://doi.org/10.1016/j.rineng.2025.107113

Duan, P., Fu, X., Hu, J., Zhou, J., and Guo, P. (2026). Path planning for UAV-based construction safety inspection under spatiotemporal interference from tower cranes. *Automation in Construction*, 182, 106762. https://doi.org/10.1016/j.autcon.2026.106762

Dubey, P. and Dubey, P. (2025). Bridging spatiotemporal wildfire prediction and decision modeling using transformer networks and fuzzy inference systems. *MethodsX*, 15(8), 103498. https://doi.org/10.1016/j.mex.2025.103498

Fu, Y. and Fang, W. (2025). A novel salient object detection network for burned area segmentation in high-resolution remote sensing images. *Environmental Modelling & Software*, 193, 106629. https://doi.org/10.1016/j.envsoft.2025.106629

Kopec, E., Wijata, A. M., and Nalepa, J. (2025). Change detection in Sentinel-2 images using deep learning ensembles. *Remote Sensing Applications: Society and Environment*, 40, 101764. https://doi.org/10.1016/j.rsase.2025.101764

Liu, B., Li, D., Xiao, X., Shao, Z., Li, Y., Zhang, H., Chen, Y., Zhang, Z., Wang, S. and Chang, B. (2026). Enhancing Sentinel-2 landslide change detection by integrating multispectral, deformation, and topographic information. *International Journal of Applied Earth Observation and Geoinformation*, 146, 105116.

Nguyen, D.C. and Nguyen, T.C., (2025). AI-Powered Edge-Based Safety Monitoring System for Power Transmission Corridors: A Case Study in Vietnam. *Computer Networks*, 111881.

Siddique, M. and Ahmed, T. (2025). CCD-Conv1D: A deep learning based coherent change detection technique to monitor and forecast floods using Sentinel-1 images. *Remote Sensing Applications: Society and Environment*, 37, 101440. https://doi.org/10.1016/j.rsase.2024.101440

Song, C., Yang, R., and Cui, Z. (2026). LitePowerCD: A lightweight anomalous change detection model for power transmission and transformation scenarios. *Electric Power Systems Research*, 251, 112194. https://doi.org/10.1016/j.epsr.2025.112194

Song, Z., Zhang, Y., Huang, X., and Zhang, Y. (2025). Fast Fusion Net: Defect detection and fault identification methods for high-voltage overhead power lines. *Engineering Applications of Artificial Intelligence*, 151, 110646. https://doi.org/10.1016/j.engappai.2025.110646

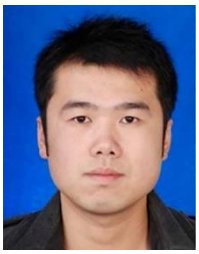
Wang, K., Li, Y., Wang, J., Zhang, Z., Zhou, S., and Zhou, R. (2026). STFI: a spatio-temporal deep fusion architecture for high-accuracy fire spot identification near power transmission lines. *Reliability Engineering & System Safety*, 271, 112257. https://doi.org/10.1016/j.ress.2026.112257

Zhi Yang is an Assistant Engineer in the Power Transformation Department, Guiyang Power Supply Bureau, Guizhou Power Grid Co., Ltd. She holds a master's degree. Her research interests include power engineering, substation hidden-danger identification, early warning, and remote-sensing-based monitoring.

Jun Cao is a Senior Engineer at Guiyang Power Supply Bureau, Guizhou Power Grid Co., Ltd. He holds a bachelor's degree. His research interests include power engineering, substation equipment monitoring, fault early warning, and Internet of Things applications.

Yujia Wang is an Engineer at Guiyang Power Supply Bureau, Guizhou Power Grid Co., Ltd. She holds a bachelor's degree. Her research interests include power engineering, substation risk detection, multimodal data analysis, and collaborative decision-making.

Liming Tao is a Senior Engineer in the Production Technology Department, Guizhou Power Grid Co., Ltd. He holds a bachelor's degree. His research interests include power engineering, production technology management, substation operation, and power equipment monitoring.



Lei Cao is a Senior Engineer in the Production Technology Department, Guizhou Power Grid Co., Ltd. He holds a bachelor's degree. His research interests include power engineering, substation hidden danger identification, intelligent early warning, and power equipment operation management.