

High-fidelity CNN-based surrogate modeling for wildfire spread forecasting

Received: 26 February 2026

Accepted: 28 May 2026

Published online: 06 June 2026

Cite this article as: Nematshahi S., Fan R., Khodaei A. *et al.* High-fidelity CNN-based surrogate modeling for wildfire spread forecasting. *Sci Rep* (2026). <https://doi.org/10.1038/s41598-026-56080-w>

Saeed Nematshahi, Rui Fan, Amin Khodaei, Erin Belval & Ali Arabnya

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

High-Fidelity CNN-Based Surrogate Modeling for Wildfire Spread Forecasting

Saeed Nematshahi¹, Rui Fan^{1*}, Amin Khodaei¹, Erin Belval², and Ali Arabnya^{1,3,4}

Abstract

The rising frequency of large-scale, destructive wildfires has significantly affected not only natural ecosystems but also critical infrastructure, human lives, and properties. Given the devastating and often irreversible environmental and financial consequences, researchers from diverse fields are actively seeking solutions to improve resilience against wildfires. Fire suppression, one of the most effective strategies to mitigate wildfire damage, relies heavily on rapid and high-fidelity forecasting of fire spread. These predictions are essential for planning evacuations by state or local emergency management agencies, implementing preemptive de-energization strategies for electric utilities, and coordinating fire containment efforts by firefighting teams. However, a significant bottleneck across all these planning processes is the significant computational burden imposed by high-resolution wildfire modeling, the demand for improved predictive accuracy, and the need to integrate diverse and large-scale datasets. Since response time is crucial for wildfire risk management, this paper proposes a deep learning-based surrogate model to predict fire spread in just a fraction of a second. We developed and trained a Convolutional Neural Network (CNN) model that efficiently predicts wildfire propagation. The proposed model demonstrates high efficiency, achieving an F1 score of 0.92. The contributions of this paper are twofold: (1) a fast, high-resolution CNN model that can support wildfire-related Public Safety Power Shutoff (PSPS) planning for electric utilities, and (2) a practical tool for firefighting and evacuation teams to support rapid and data-informed risk assessment.

Keywords: emergency risk assessment, fire suppression, power system resilience, wildfire management, wildfire spread forecasting

¹Department of Electrical and Computer Engineering, University of Denver, Denver, CO, USA. ²Rocky Mountain Research Station, U.S. Forest Service, Fort Collins, CO, USA. ³Quanta Technology, Raleigh, NC, USA. ⁴Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, USA. E-mail: rui.fan@du.edu

Introduction

While wildfires have long been a natural ecological process, recent years have seen a marked escalation in their frequency, intensity, and spatial extent—driven primarily by climate change and the accelerating expansion of the Wildland-Urban Interface (WUI) [1].

According to data from the U.S. National Centers for Environmental Information, the average temperature in the United States during the 21st century has increased by approximately 1.64°F compared to the 20th-century baseline. Notably, all of the nine hottest years on record have occurred after 2006. Figure 1 illustrates the trend in the annual average temperature increase in the United States over the period of 1895–2024 [2]. Elevated temperatures, coupled with low humidity and prolonged drought conditions, significantly increase the susceptibility of wildland areas to wildfire ignition and rapid spread.

Concurrently, the continued expansion of the WUI across the United States exacerbates the impacts of wildfires, including increased risk to human life [3], property loss [4], and damage to critical infrastructure [5–7]. Figure 2 presents the distribution of WUI areas by state and county within the United States [8]. The left figures depict the WUI status in 2020, while the right ones highlight the growth in WUI areas over the preceding 30 years. Overlaying the WUI map with the drought map identifies areas more susceptible to wildfire ignition and spread, highlighting regions at risk of costly wildfires.

Wildfire behavior modeling approaches are generally classified into two categories: physics-based fire dynamics simulators and data-driven models. Notable tools include IFTDSS (Interagency Fuels Treatment Decision Support System), MTT (Minimum Travel Time), FlamMap, FSim, FARSITE (Fire Area Simulator), FireSim, FIRETEC, and WFDS (Wildland–Urban Interface Fire Dynamics Simulator). These models are widely used for wildfire behavior analysis, and each offers specific advantages. However, they require the integration of heterogeneous datasets, including landscape characteristics, fuel properties, and meteorological conditions. While this integration is feasible for long-term planning and risk assessment, it can pose significant challenges during emergency situations due to the time-intensive nature of data preparation and computationally intensive model execution. While conventional models remain important, their predictive performance can be limited without dynamic data updating and hybridization [9]. Authors in [10] stated that future developments may involve machine learning algorithms to improve the wildfire models. Interpretable boosting models have achieved high accuracy in predicting wildfire spread scale while identifying key drivers such as humidity, wind speed, and suppression activity [11]. Deep learning frameworks have also enabled cross-scale forecasting of wildfire fronts with lead times up to 72 hours [12]. In addition, specialized models such as TabNet outperformed general-purpose large language models in Fire Radiative Power prediction, highlighting the importance of domain-specific AI for operational wildfire forecasting [13]. In [14], applied machine learning has been used to classify wildfire sizes in Siberia, with XGBoost achieving 88.8% accuracy and highlighting fuel dryness and urban proximity as key drivers. A recent study in [15], developed a hybrid CNN–LSTM framework for short-term wildfire susceptibility forecasting, demonstrating improved

predictive accuracy over conventional machine learning models by integrating meteorological and geospatial drivers.

The authors in [16] introduced the Next Day Wildfire Spread dataset—a publicly available resource encompassing thousands of wildfire events across the contiguous United States from 2012 to 2020. This dataset is designed for tasks such as statistical analysis for land management and disaster preparedness. It captures the spatial progression of wildfires by providing daily snapshots, enabling comparative analysis of fire spread over two consecutive days. Building on this resource, [17] employed the dataset to develop the Attention Swin U-Net with Focal Modulation (ASUFM) for next-day wildfire spread prediction. Similarly, [18] applied three machine learning algorithms—Decision Tree Regressor, XGBoost Regressor, and Artificial Neural Network (ANN)—to model fire spread in the next day, while [19] introduced an explainable Convolutional Neural Network (CNN) architecture with Atrous Spatial Pyramid Pooling (CNN-ASPP) for the same purpose. While these studies demonstrate promising results, they are constrained by the historical dataset’s inherent limitations—namely, a prediction horizon restricted to a single day and a spatial resolution of 1 km. In [20], the authors proposed a Deep Convolutional Inverse Graphics Network (DCIGN) for modeling wildland fire spread at a spatial resolution of 1 km. The model requires approximately 18 to 51 hours for training and takes around 185 seconds to generate a prediction for a wildfire scenario. It achieves an F1 score of 0.78 but is limited to forecasting fire spread over a two-day horizon. In contrast, real-world wildfires often span multiple days to be fully contained, necessitating faster, more accurate, and higher-resolution predictive capabilities to inform tactical decision-making by emergency response teams across multiple sectors. Authors in [21] developed a neural network framework based on the MA-Net architecture for three-day-ahead prediction, reporting an F1 score of 0.67. Large benchmark datasets integrating remote sensing imagery, weather, topography, land cover, and fuel information now support tasks such as active fire detection, burned area mapping, and next-day progression prediction [22]. To the best of our knowledge, no existing study has yet presented a machine learning framework that can accurately predict large-scale wildfire spread while remaining fully independent of data sources that are often costly, delayed, or operationally restrictive. Most current wildfire prediction models rely heavily on one or more of three external inputs: unmanned aerial vehicles (UAVs) for localized real-time observations, historical wildfire databases for training on past fire behavior, or frequently updated satellite imagery for continuous monitoring of active events. Although these inputs can improve forecasting performance, they also introduce important limitations related to coverage, scalability, response latency, weather interference, data availability, and deployment cost. UAV-based systems are typically constrained to small operational areas and require field coordination, while satellite products may suffer from revisit delays, cloud obstruction, smoke contamination, or insufficient spatial resolution during fast-moving events. Similarly, historical wildfire datasets may not generalize well to unprecedented fire behavior driven by changing climate conditions, novel fuel patterns, or region-specific topography. Therefore, a robust machine learning approach that can forecast large-scale wildfire spread using

readily available static or low-frequency environmental data would represent a significant advancement for rapid-response operations and geographically transferable wildfire management.

In this work, we present a CNN-powered surrogate model that achieves an F1 score of 0.92, operating at a spatial resolution of 120 meters with a prediction time of fractions of a second. The model is adaptable to different geographic regions and was fully trained in under 4 hours on a standard personal computer equipped with a 13th Gen Intel® Core™ i7-13700 CPU (2.10 GHz) and 16 GB of RAM, without the need for GPU acceleration. The model was trained on a dataset comprising 1,217 wildfire scenarios and demonstrates robust performance across both small- and large-scale wildfires, with no observable degradation in predictive accuracy based on fire size.

Moreover, it effectively captures a wide variety of wildfire shapes and spatial configurations, highlighting its generalizability and spatial adaptability. This model offers valuable contributions to wildfire emergency risk assessment and mitigation strategies. It can support Public Safety Power Shutoff (PSPS) planning by enabling electric utilities to proactively de-energize power lines in areas projected to be affected by wildfire spread. Additionally, it aids in the development of evacuation plans, particularly when wildfires are forecasted to encroach upon WUI zones within the coming hours or days. Most critically, it provides firefighters with timely and high-resolution predictions of fire progression, facilitating more informed and effective fire suppression planning and resource deployment.

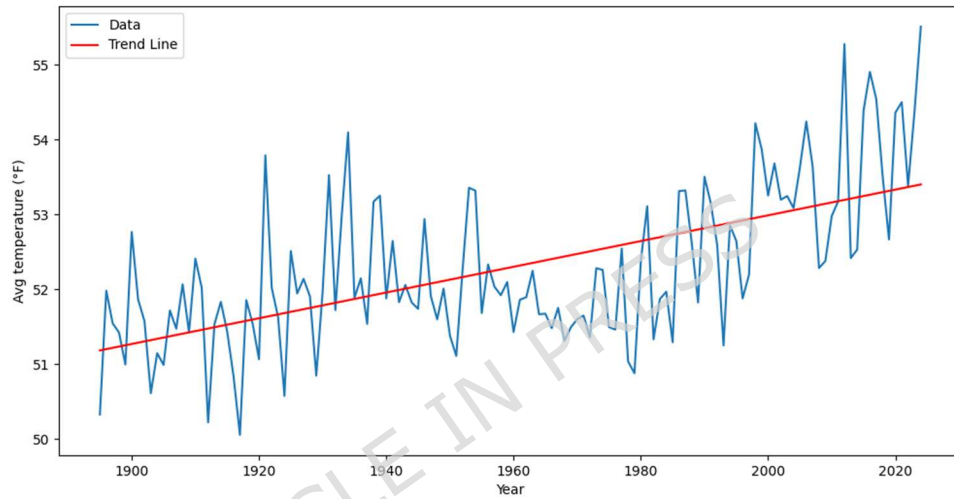
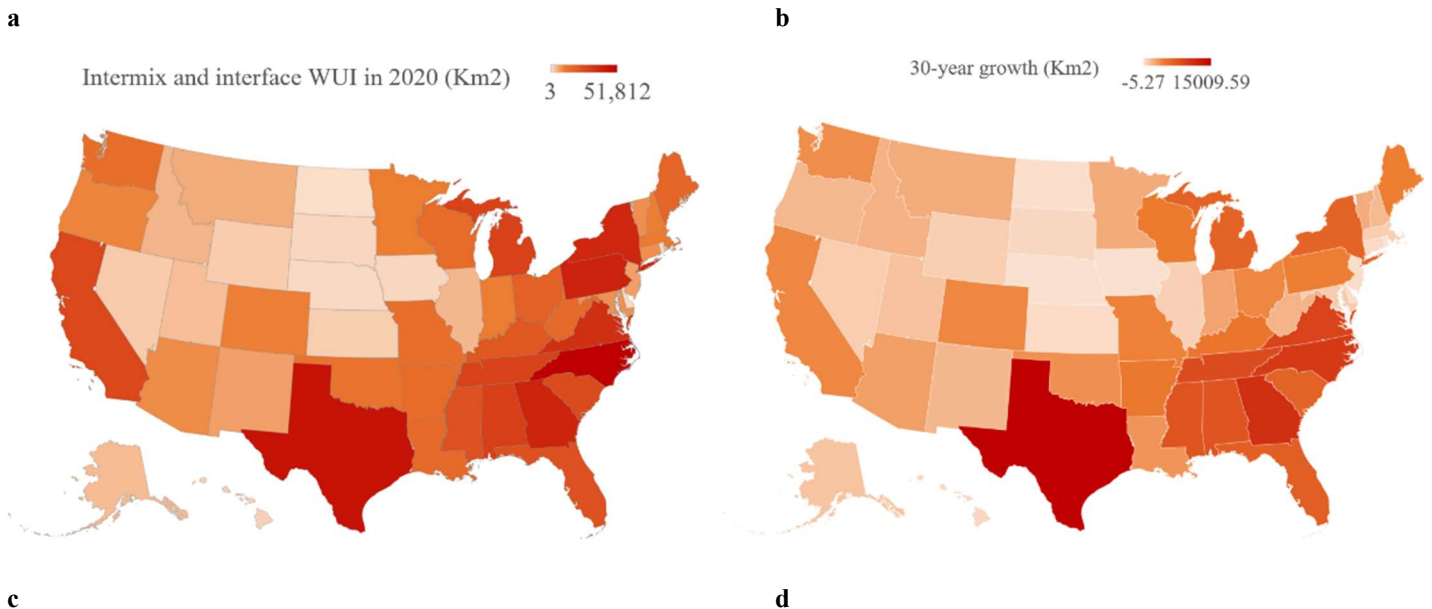


Fig. 1 | US average temperature of the year from 1985 to 2020.



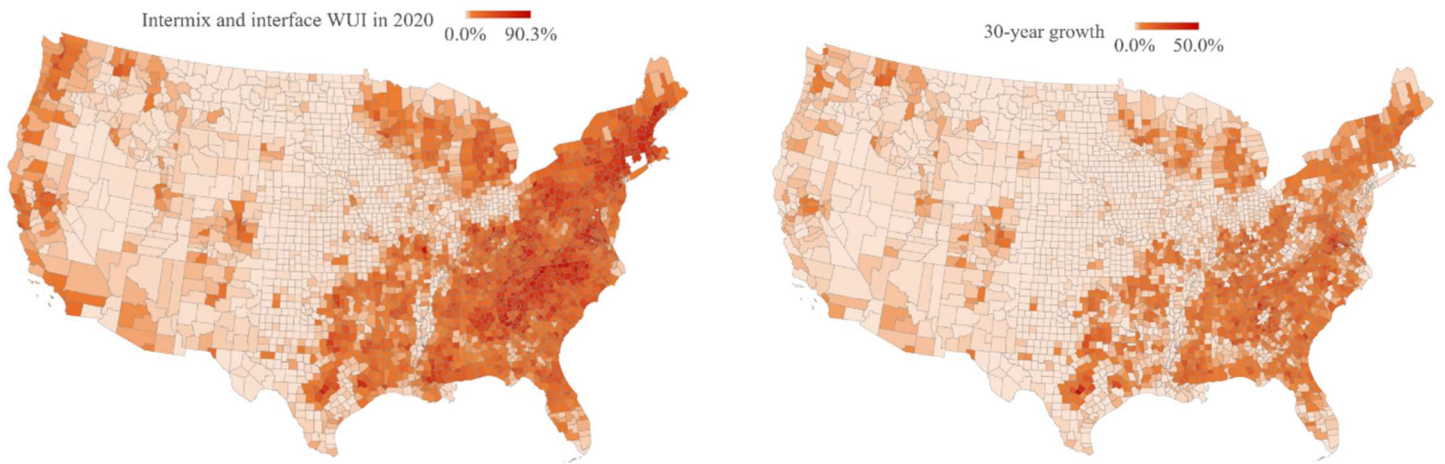


Fig. 2 | Wildland-Urban Interface distribution. a, The 2020 WUI status in different states by Km². **b,** The growth over the last thirty years in different states by Km². **c,** The 2020 WUI status in different counties by percentage. **d,** The growth over the last thirty years in different counties by percentage.

Note: The dataset did not include county-level data for Alaska and Hawaii.

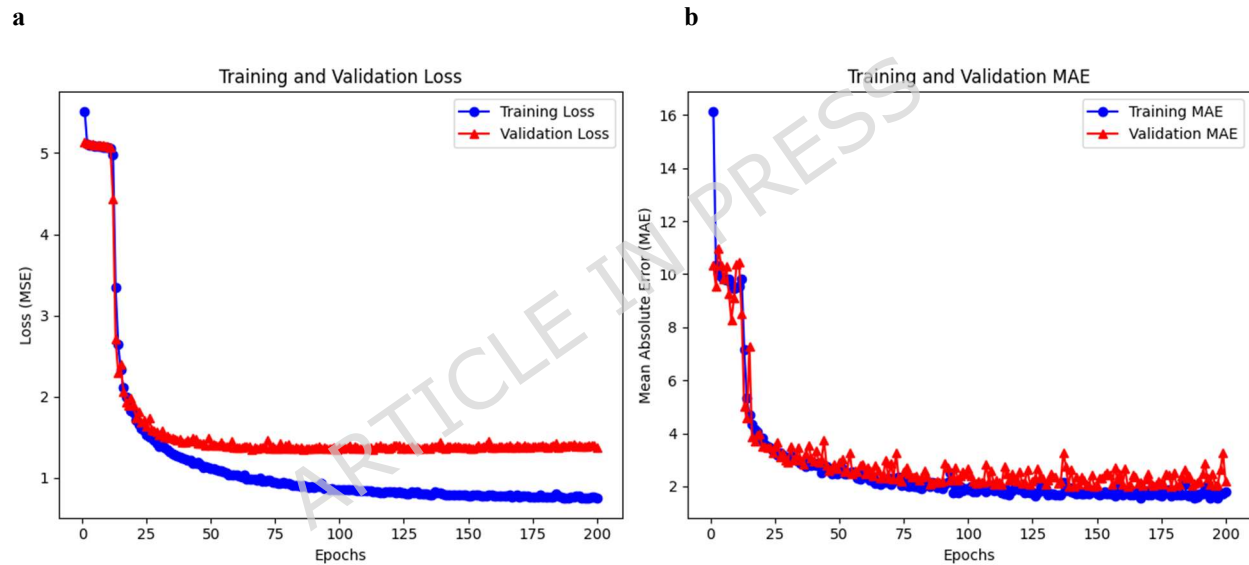


Fig. 3 | Model MSE and MAE through the epochs. a, Training and validation MSE loss. **b,** Training and validation MAE.

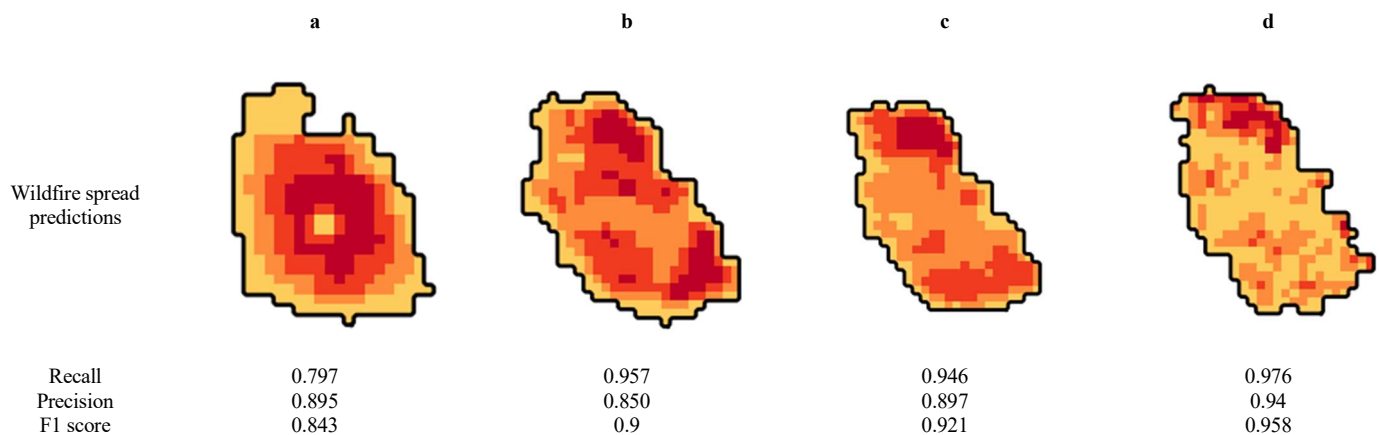


Fig. 4 | Efficiency progress during epochs. a, Prediction of burned area after 15 epochs. b, Prediction of burned area after 30 epochs. c, Prediction of burned area after 45 epochs. d, Prediction of burned area after 200 epochs.

Note: The different colors represent the regression output and do not correspond to the temporal progression of the fire. All pixels within the figure indicate burned areas. As training progresses, the model produces increasingly sharper burn perimeters, reflecting improved accuracy.

Results

Model Establishment

The model begins by extracting features most relevant to wildfire propagation, learning their spatial and contextual influence on fire spread across 1,217 simulated scenarios. During training, it aims to minimize the loss function, which is computed over the entire prediction domain of 464×516 pixels (239,424 total pixels). Figure 3 illustrates the training and validation Mean Squared Error (MSE) across 200 epochs, demonstrating progressive improvement in model accuracy. Notably, the validation curve indicates that acceptable prediction performance can be achieved even after approximately 100 epochs, suggesting efficient convergence. The output of the model would be a continuous variable for every single pixel, indicating the probability of fire occurrence. A subsequent post-processing step thresholds the continuous model output using a cutoff value of 0.5, where pixels with values greater than 0.5 are classified as burned and those below are considered unburned, thereby generating a binary fire spread map. The threshold value was selected based on empirical evaluation across thousands of wildfire scenarios by maximizing the overall F1 score. Since the network output represents normalized fire occurrence likelihood values between 0 and 1, the selected threshold also provides a balanced and interpretable decision boundary between burned and unburned classes.

To provide a clearer illustration of the model's performance progression, Figure 4 visualizes the predicted wildfire spread at selected training epochs—15, 30, 45, and 200. During the training process, the number of true positives exhibits an increasing trend from 1181 at epoch 15 to 1,444 at epoch 200, while the number of false negatives shows a decreasing trend from 299 to 36 over the same period. The results demonstrate a progressive refinement in prediction accuracy, with the F1 score improving from 0.843 at epoch 15 to 0.958 at epoch 200. This highlights the model's ability to achieve reasonably accurate predictions at early training stages, offering significant advantages for time-sensitive applications such as real-time risk assessment. For instance, the model can be trained from scratch using up-to-date weather data, and after only 15 epochs (approximately 15 minutes of training time), it can provide a practical forecast of fire spread. When computational resources and time permit, training can be extended up to 200 epochs to reach maximum performance. While the model has also been trained for 500 epochs, results indicate diminishing returns, with no significant improvements in predictive performance beyond 200 epochs. In addition, for a long-term wildfire mitigation plan, it can be executed through different seasons with different weather streams to receive a thorough wildfire risk analysis in the area of study [23].

Performance

Our surrogate model is a fast-response yet accurate wildfire

spread predictor. To evaluate its generalization performance, we have selected five random instances to depict the accuracy of our approach. The five wildfire scenarios presented in Table 1 were randomly selected from the held-out validation subset following model training to demonstrate the predictive performance of the proposed framework across different wildfire sizes and spatial propagation patterns. Each scenario was initially simulated using FARSITE, and the corresponding fire spread was then predicted by our CNN model. Table I presents side-by-side visualizations of the ground-truth FARSITE outputs and the CNN-generated predictions, allowing for direct visual and spatial comparison. The size of the wildfire spread for different scenarios is diverse, so we present the spread shapes in the same size, but the actual size is mentioned beneath the burned area. The smallest case is 2,775 acres, while the largest case among the five selected scenarios is 12,188 acres. The model's accuracy in predicting the wildfire spread is not dependent on the wildfire size or even shape, which shows its capability to handle both small and large-scale wildfires without losing accuracy. Notably, Ignition 23 on line 3 includes an internal unburned pocket (a "hole") within the fire perimeter—typically a challenging feature to predict. Nevertheless, the model successfully identifies this complex pattern, further illustrating its capability to capture nuanced fire behaviors.

Next to every instance, the confusion matrix and efficiency matrix are mentioned. Since the number of True Negatives (TNs) is very high compared to other values in the confusion matrix, the accuracy for the scenarios always remains over 99%, rendering it an uninformative metric in this context. Instead, Recall and F1 score are particularly useful metrics when evaluating the efficiency of wildfire spread prediction models, especially in scenarios where the dataset is imbalanced and the consequences of missing a fire-affected area are severe. Recall is crucial when the priority is to detect all areas that the wildfire might reach, minimizing false negatives, which is important for emergency response and evacuation planning. However, focusing solely on Recall can lead to many false positives. The F1 score addresses this by providing a balanced metric that combines both Precision and Recall, making it valuable when both false alarms and missed detections carry significant consequences. These metrics are especially relevant in grid-based or spatiotemporal prediction tasks where each unit (e.g., pixel or zone) is classified as fire-affected or unaffected over time.

Across all scenarios, the model achieved an average Recall of 0.925 and an average F1 score of 0.92, demonstrating high reliability in identifying burned areas at a spatial resolution of 120 meters.

Discussion

Our study aims to predict wildfire spread in a fast yet accurate

manner. We developed a CNN-powered surrogate model that estimates the burned area upon receiving the ignition point. Wildfire behavior modeling typically requires the integration of diverse datasets—such as topography, vegetation, and weather—and involves complex physical formulations, making the process computationally intensive and time-consuming. In emergency response scenarios, however, rapid and precise fire spread predictions are critical for effective decision-making. While traditional physics-based models embedded in wildfire simulation tools provide accuracy, the data preparation and simulation process often take several hours. In contrast, the proposed surrogate model delivers wildfire spread predictions in a fraction of a second, achieving high performance with an average F1 score of 0.92 and a Recall of 0.93, making it a practical and scalable tool for real-time emergency planning.

The predictive high fidelity and rapid inference capability of the proposed model make it particularly well-suited for real-time applications in emergency response planning during wildfire events. Its utility extends across multiple facets of wildfire resilience, supporting proactive decision-making for fire suppression strategies, evacuation planning, and infrastructure protection.

First, in the context of fire suppression strategy development, the model facilitates anticipatory planning for strategic coordination of grounded and aerial suppression resources (e.g., water and retardant drops). Also, in case of a large-scale wildfire, they can immediately ask for resources across municipal, state, and federal entities.

Second, it can improve evacuation planning by highlighting areas that are likely to be affected and helping determine safe and efficient evacuation routes. This includes issuing timely alerts, managing traffic flow, and ensuring that shelters are prepared to receive displaced people, animals, and vulnerable populations.

Third, the model can inform power system operational decisions, such as implementing Public Safety Power Shutoffs (PSPS) or dynamically adjusting line ratings based on fire intensity profiles beneath the power lines [24]. Such decisions are vital to minimizing ignition risks from electrical infrastructure and preserving grid stability during evolving wildfire threats.

Fourth, road access control and traffic management are also enhanced through the model's predictive analytics. By identifying intra-zoon road networks and potential choke points, emergency planners can implement preemptive road closures and diversions to both expedite evacuation and facilitate unimpeded access for emergency responders.

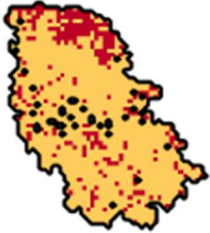
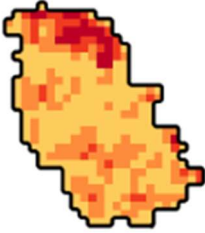
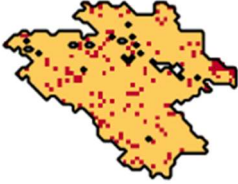
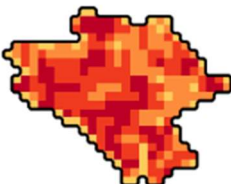
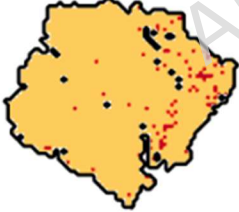
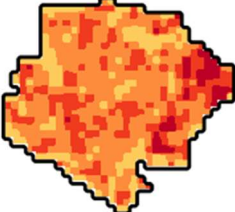
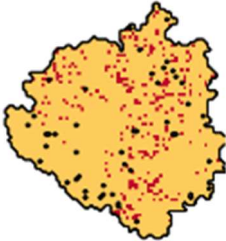
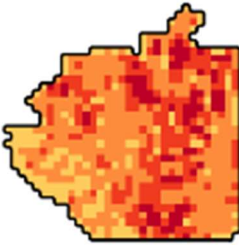
Collectively, the model's performance characteristics align with the operational needs of contemporary wildfire response systems, underscoring its potential to be embedded within decision-support tools for resilience planning under wildfire hazard scenarios.

The predicted binary wildfire spread map is intended to function as a foundational spatial hazard layer within broader operational decision-support systems rather than as a standalone autonomous decision-making tool. For PSPS applications, the predicted fire perimeter can be spatially integrated with utility infrastructure maps, transmission corridors, load criticality information, and operational reliability constraints to support risk-informed de-energization strategies. Similarly, for evacuation planning, the forecasted wildfire spread can be combined with transportation networks, population distribution, shelter availability, and traffic management models to identify vulnerable communities, estimate evacuation bottlenecks, and dynamically support route planning. Therefore, the proposed framework provides the rapid wildfire hazard forecasting component upon which higher-level operational and optimization models can subsequently operate.

The reported Recall and F1 score values quantify the agreement between the proposed CNN surrogate model and the FARSITE-generated wildfire simulations used as reference data. Therefore, these metrics should be interpreted as indicators of surrogate-model fidelity relative to the underlying physics-based simulator rather than direct validation against observational wildfire spread data.

A limitation of the proposed framework is that the CNN surrogate model is trained on wildfire simulations generated by the FARSITE physics-based simulator. Consequently, the model inherently learns and reproduces the assumptions, simplifications, and potential biases embedded within the underlying simulator. Although FARSITE is a widely validated wildfire modeling tool, real-world wildfire behavior may involve additional complexities—such as abrupt atmospheric changes, ember spotting, plume-dominated fire dynamics, suppression interventions, and fine-scale fuel heterogeneity—that are not fully represented in the training data. For example, it relies on semi-empirical correlations of fire behavior and is thus incapable of predicting extreme wildfire behavior, such as fire whirls, fire merging, and rapid fire spread on steep slopes [25-27]. Nevertheless, the proposed deep learning framework is fully adaptable and can be trained using alternative wildfire datasets or operational fire observations as such datasets become available. In this work, FARSITE-generated wildfire scenarios were employed because publicly available datasets containing large numbers of high-resolution and temporally indexed wildfire spread samples suitable for deep learning training remain extremely limited.

Table 1 Wildfire spread predictor analysis

Instance	Burned area		Efficiency metrics		
	Actual	Prediction	Recall	Precision	F1 Score
Ignition 3 on line 41			0.976	0.94	0.958
	Size: 5,266 acres	Size: 5,465 acres			
Ignition 27 on line 10			0.940	0.931	0.936
	Size: 4,145 acres	Size: 4,452 acres			
Ignition 23 on line 3			0.965	0.872	0.916
	Size: 2,775 acres	Size: 3,073 acres			
Ignition 17 on line 18			0.972	0.966	0.969
	Size: 8,160 acres	Size: 8,211 acres			
Ignition 6 on line 40			0.884	0.978	0.929
	Size: 12,188 acres	Size: 11,002 acres			

Note: This entire area encompasses approximately 842,000 acres, bounded by latitudes 37.6° to 38.1° and longitudes -120.7° to -120.0°.

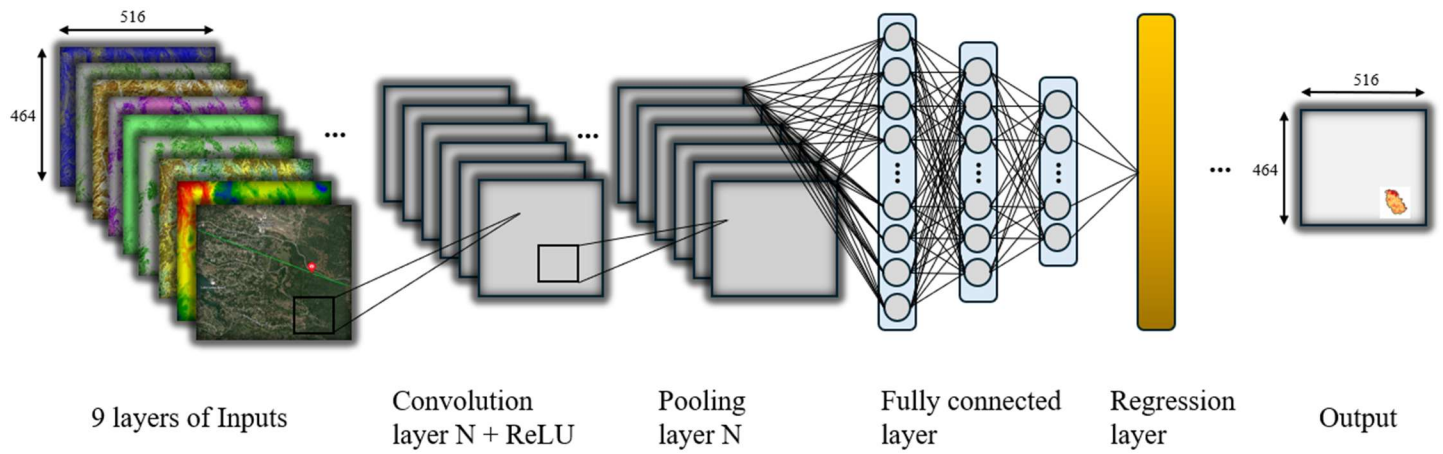


Fig. 5 | The architecture of Convolutional Neural Networks (CNNs). Two convolutional layers and two pooling layers are adopted in the wildfire spread model.

Methods

Dataset

To develop and train our CNN-powered wildfire spread prediction model, we constructed a diverse and high-resolution wildfire simulation dataset. The dataset was generated using FARSITE (Fire Area Simulator)—a widely adopted wildfire behavior model developed by the U.S. Forest Service. It integrates topography, fuel characteristics, and weather conditions to simulate fire spread and produces two-dimensional, GIS-compatible outputs for fire management and planning applications.

Our dataset comprises 1,217 wildfire scenarios covering approximately 842,000 acres in central California, specifically between latitudes 37.6° N to 38.1° N and longitudes -120.7° W to -120.0° W. This region includes parts of the Sierra Nevada foothills and the San Joaquin Valley, encompassing a diverse landscape with oak woodlands, grasslands, and coniferous forests—offering a wide range of fuel types for modeling fire behavior. The area's Mediterranean climate, characterized by hot, dry summers and wet winters, along with strong seasonal winds and prolonged drought conditions, makes it especially susceptible to wildfires. The proximity to rural communities, agricultural zones, and Yosemite National Park further underscores its importance as a case study for wildfire risk assessment and mitigation planning.

We overlaid the IEEE 30-bus transmission network onto the designated study region and introduced random ignition points along the transmission lines, under the assumption that each wildfire event would be fully contained within 96 hours [28]. Each ignition point represents a potential grid-ignited wildfire event. The geographic distribution of transmission infrastructure across diverse terrain within the study area ensures that the dataset captures a wide range of fire spread dynamics, enhancing the generalizability and adaptability of the trained model to different regions within the study zone.

CNN model

The proposed Convolutional Neural Network (CNN) is architected to predict wildfire spread, with a particular focus on crown fire activity, which reflects the intensity and spatial extent of fire propagation through forest canopies. Figure 5

depicts the architecture of the proposed CNN, indicating the input layers, convolutional layers, pooling layers, fully connected layers, regression layer, and the output.

The model ingests two distinct categories of input data: fixed input layers and variable input layers.

The eight fixed geospatial layers are derived from the LANDFIRE database. These layers represent the standard landscape (LCP) variables required for high-fidelity wildfire behavior modeling: elevation, slope, aspect, fuel model, canopy cover, canopy height, crown base height, and crown bulk density. These layers collectively provide the static environmental context for the 842,000-acre study area in central California. To ensure consistent feature weighting and numerical stability during the training of the CNN, each geospatial layer was normalized using Min-Max scaling to a range of [0, 1]. For the concatenation process, these eight normalized geospatial layers were stacked channel-wise alongside the one-hot encoded ignition map, which indicates the spatial origin of the fire. This resulted in a composite input tensor of dimensions 464*516*9* for each simulation instance. While physics-based simulators (such as FARSITE) require the integration and preprocessing of heterogeneous datasets—including landscape, fuel, and dynamic meteorological conditions—for every simulation run, our surrogate model performs this integration only once. By utilizing these layers as a fixed environmental backbone, the model requires only a new ignition map to perform inference, allowing it to generate high-resolution forecasts (120-meter resolution) in fractions of a second.

The model architecture is specifically designed to capture both fine-grained and broad-scale spatial patterns. Following the input concatenation, the network utilizes a sequence of dilated convolutional layers with a dilation rate of 10 and ReLU activation functions, beginning with 32 filters and increasing to 64 filters. Figure 6 shows how dilated convolution captures the information from the surrounding pixels. This use of dilation expands the receptive field, allowing the model to incorporate long-range spatial dependencies (e.g., terrain-driven fire spread) without a proportional increase in computational cost. Each convolutional block is followed by max-pooling layers, which reduce the spatial dimensions (e.g., from 464*516 to

232×258, and then to 116×129), effectively extracting dominant features such as fuel continuity and slope gradients.

To reconstruct high-resolution predictions, the architecture employs transposed convolutional layers (Conv2DTranspose) with 64 and 32 filters, along with upsampling layers, to

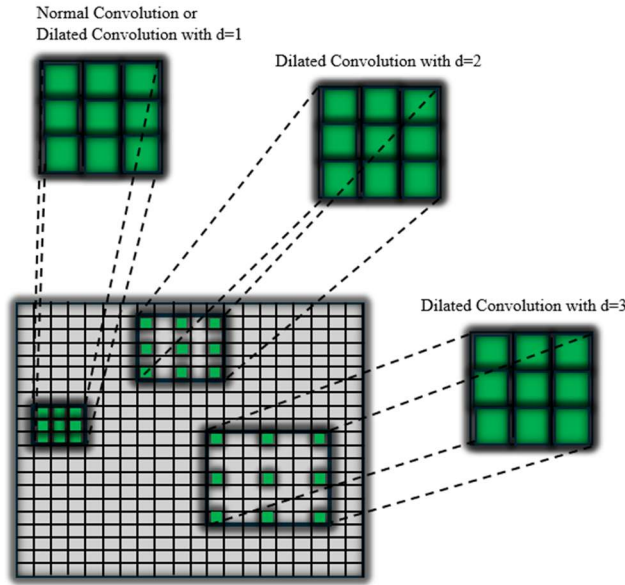


Fig. 6 | The dilated convolutional architecture. A higher dilated convolution reflects more information from the surrounding pixels, which makes the CNN model sharper yet more reliable.

The model is compiled using the Adam optimizer with a learning rate of 0.0005, and trained with mean squared error (MSE) as the primary loss function. Mean absolute error (MAE) is employed as an additional performance metric to assess prediction accuracy during both training and validation phases.

Training and Prediction Process

The model is trained on a dataset architecture in which fixed environmental inputs—such as topography, fuel characteristics, and weather conditions—are replicated across all instances, while variable ignition points and their corresponding crown fire outputs are loaded from separate directories. Each pair constitutes a single training instance. To promote model generalization, prior to training, the 1,217 wildfire scenarios were randomly shuffled, and 20% of the dataset was reserved as a validation subset. The reported performance metrics, including the average Recall and F1 score, were computed on this held-out validation dataset, which was excluded from the model parameter update process during training. The model is trained over 200 epochs with a batch size of 8, using a 20% validation split. This configuration enables the model to effectively learn how ignition locations interact with static geospatial features to influence wildfire spread dynamics.

During training, model performance is evaluated on the held-out validation set using Mean Squared Error (MSE) and Mean Absolute Error (MAE) as loss metrics. In this case, the

progressively restore the spatial dimensions to the original input size. A final 1×1 convolutional layer with linear activation produces a single-channel output representing the predicted crown fire activity as a continuous variable across the spatial grid.

loss function per pixel is:

$$L = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (1)$$

where y_i is the ground truth value for pixel i , $\hat{y}_i$ is the predicted value from the network output, and N is 464×516 pixels per sample × number of samples in batch.

Once trained, the model generates spatial crown fire prediction maps, accurately estimating the progression of wildfire from given ignition points under specific environmental conditions. The final trained model is saved for future inference tasks, enabling real-time wildfire forecasting that supports emergency operations such as fire suppression planning, evacuation routing, and Public Safety Power Shutoff (PSPS) decision-making.

Efficiency metrics

The study area comprises 239,424 pixels, with output for each pixel labeled as 1 to indicate burned area and 0 for unburned. For each wildfire scenario, a confusion matrix was generated by comparing the model's predicted burn map to the ground truth data from the simulation. Due to the overwhelming prevalence of TNs—reflecting the relatively small proportion of burned pixels relative to the total area—the overall accuracy consistently exceeds 0.99, making it an insufficient standalone metric for model evaluation in this context. To more effectively assess model performance, we focused on Recall and the F1 score as primary efficiency metrics.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (2)$$

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

where,

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (4)$$

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

Funding

This work was supported in part by the U.S. Forest Service under grant no 22-CS-11221636-121.

Data Availability

The landscape data is retrieved from LandFire, and the weather stream is accessible from the National Solar Radiation Database

(NSRDB) at <https://nsrdb.nrel.gov/> with hourly intervals. The dataset of 1217 wildfire simulations is available at Zenodo at <https://zenodo.org/records/15707374> [29].

Code Availability

The developed codes are available in the following repository. <https://zenodo.org/records/15707691> [30].

References

- [1] Arabnya, A., & Khodaei, A. The Economics of Climate Resilience in Power Infrastructure. *John Wiley & Sons* (2024).
- [2] Climate change at a Glance National Time Series. National Centers for Environmental Information <https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/national/time-series/> (2025).
- [3] Bowman, D., Williamson, G., Abatzoglou, J. et al. Human exposure and sensitivity to globally extreme wildfire events. *Nat Ecol Evol* **1**, 0058 (2017).
- [4] McConnell, K., Fussell, E., DeWaard, J. et al. Rare and highly destructive wildfires drive human migration in the U.S.. *Nat Commun* **15**, 6631 (2024).
- [5] Wang, X., Bocchini, P. Predicting wildfire ignition induced by dynamic conductor swaying under strong winds. *Sci Rep* **13**, 3998 (2023).
- [6] Zhao, A.P., Alhazmi, M., Li, S. et al. Enhancing Los Angeles' resilient energy systems amid wildfires. *Sci Rep* **15**, 20813 (2025).
- [7] Wang, Z., Wara, M., Majumdar, A. et al. Local and utility-wide cost allocations for a more equitable wildfire-resilient distribution grid. *Nat Energy* **8**, 1097–1108 (2023).
- [8] Radeloff, V. C., et al. "Rapid growth of the US wildland-urban interface raises wildfire risk." *Proceedings of the National Academy of Sciences* **115**.13 (2018): 3314-3319.
- [9] Singh, H., Ang, L.M., Lewis, T. et al. Trending and emerging prospects of physics-based and ML-based wildfire spread models: a comprehensive review. *J. For. Res.* **35**, 135 (2024).
- [10] Vahedi, S., Zhao, J., Pierre, B. et al. Wildfire and power grid nexus in a changing climate. *Nat Rev Electr Eng* **2**, 225–243 (2025).
- [11] Liao, B., Zhou, T., Yu, J. et al. Enhancing wildfire spread scale prediction with explainable AI: a hist gradient boosting and SHAP-based approach. *Nat Hazards* **122**, 378 (2026).
- [12] Li, Y., Zeng, Y., Pan, Z. et al. Real-time forecast of high-resolution wildfire spread via Fast Cross-Scale Deep Learning. *ENG. Environ.* **20**, 65 (2026).
- [13] Ramesh, M., Sun, Z., Li, Y. et al. Assessing WildfireGPT: a comparative analysis of AI models for quantitative wildfire spread prediction. *Nat Hazards* **121**, 13117–13130 (2025).
- [14] Malashin, I.P., Masich, I., Nelyub, V. et al. Assessing wildfire extents in Siberian forests using machine learning. *Sci Rep* **15**, 32834 (2025).
- [15] Mousa, M.H., Algamdi, A.M., Fouad, Y. et al. CNN-MLP framework for forest burned areas prediction using PSO-WOA algorithm. *Sci Rep* **16**, 4982 (2026).
- [16] Huot, F., Hu, R. L., Goyal, N., Sankar, T., Ihme, M., and Chen, Y. F. Next Day Wildfire Spread: A Machine Learning Dataset to Predict Wildfire Spreading From Remote-Sensing Data. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-13, 2022, Art no. 4412513.
- [17] Li, B. S. and Rad, R. Wildfire Spread Prediction in North America Using Satellite Imagery and Vision Transformer. *2024 IEEE Conference on Artificial Intelligence (CAI)*, Singapore, Singapore, 2024, pp. 1536-1541.
- [18] Singh, A., Yadav, R., Sudhamshu, G., Basnet, A., and Ali, R. "Wildfire Spread Prediction using Machine Learning Algorithms. *14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Delhi, India, 2023, pp. 1-5.
- [19] Marjani, M., Mahdianpari, M., Ali Ahmadi, S., Hemmati, E., Mohammadimanesh, F., and Saadi Mesgari, M. Application of Explainable Artificial Intelligence in Predicting Wildfire Spread: An ASPP-Enabled CNN Approach. *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024, Art no. 2504005.
- [20] Hodges, J.L., Lattimer, B.Y. Wildland Fire Spread Modeling Using Convolutional Neural Networks. *Fire Technol* **55**, 2115–2142 (2019).
- [21] Shadrin, D., Illarionova, S., Gubanov, F. et al. Wildfire spreading prediction using multimodal data and deep neural network approach. *Sci Rep* **14**, 2606 (2024).
- [22] Zhao, Y., Gerard, S. & Ban, Y. TS-SatFire: A Multi-Task Satellite Image Time-Series Dataset for Wildfire Detection and Prediction. *Sci Data* **12**, 1817 (2025).
- [23] Nematshahi, S., Khodaei, A., Arabnya, A. An investment prioritization model for wildfire risk mitigation through power line undergrounding. *14th Mediterranean Conference on Power Generation Transmission, Distribution and Energy Conversion (MEDPOWER 2024)*, Athens, Greece, 2024, pp. 394-399.
- [24] Nematshahi, S., Sohrabi, A., Khodaei, B., Arabnya, A., and Fan, R. Wildfire impact analysis on power system operation: a framework leveraging physics-based wildfire simulation and dynamic line rating. *Electric Power Systems Research*, vol 252, January 2026, 112467.
- [25] Liu, N., Lei, J., Gao, W., Chen, H., & Xie, X. Combustion dynamics of large-scale wildfires. *Proceedings of the Combustion Institute*, **38**(1), 157-198. (2021).
- [26] Liu N, Jinmo Wu, Chen H, Xie X, Zhang L, Yao B, Zhu J, Shan Y. Effect of slope on spread of a linear flame front over a pine needle fuel bed: experiments and modelling. *International Journal of Wildland Fire* **23**, 1087–1096. (2014).
- [27] Zhou K, Wang Y, Simeoni A, Dong R. Fire whirls induced by a line fire on a windward slope: a laboratory-scale study. *International Journal of Wildland Fire* **33**, WF23048. (2024).
- [28] Nematshahi, S., Sohrabi, B., Arabnya, A., Khodaei, A., and Belval, E. A Catastrophe Bond Design for the Financial Resilience of Electric Utilities Against Wildfires. *IEEE Transactions on Energy Markets, Policy and Regulation*, vol. 3, no. 1, pp. 133-143, March 2025.
- [29] Nematshahi, S. (2025). CA Wildfire Scenarios [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.15707374>.
- [30] Nematshahi, S. (2025). Wildfire Spread Forecaster: A CNN model. Zenodo. <https://doi.org/10.5281/zenodo.15707691>.