Machine Learning Based Forecasting for Building Fires

Wai Cheong Tam*, Hongqiang Fang*, Yifei Ding

Fire Research Division, Engineering Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland, USA

Department of Building Environment and Energy Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

* corresponding authors: waicheong.tam@nist.gov and hongqiang.fang@nist.gov

1 Abstract

The fast-evolving conditions of rapid fire progressions demand swift and informed decision-making from firefighters. This paper presents a series of research efforts to develop artificial intelligent-driven technologies that provide real-time, actionable information during fire emergencies. By leveraging synthetic data and machine learning, these technologies aim to enhance hazard recognition, reduce firefighter risk, and improve operational effectiveness, ultimately strengthening life safety and mitigating property loss.

2 Introduction and motivation

Although the annual number of residential fires in the United States has decreased by over 50 % since 1980 [1], the severity of fire hazards has intensified. Modern residential homes increasingly contain synthetic furnishings causing room conditions to deteriorate much faster than in past decades. For example, one study [2] reported that the transition to flashover (i.e., the point at which nearly all combustible items ignite simultaneously) can occur in as early as 5 minutes in modern residential homes, compared with approximately 30 minutes for home with legacy furniture. When flashover occurs, the temperature can exceed more than 800 °C. Despite these hazards, current fireground practice relies heavily on firefighter experience and visual cues, which are often unreliable in smoke-obscured conditions [3]. If these early indicators are missed or misinterpreted, the rapid escalation to flashover can place firefighters at severe and potentially fatal risk.

Traditional approaches for hazard prediction, including empirical correlations [4], inverse modeling [5], and computational fluid dynamics (CFD) [6], face trade-offs between speed, fidelity, and practicality. Although CFD provides detailed representations of fire dynamics in complex geometries, its computational demands are prohibitive for real-time use. For example, reference [6] reports that a single timestep of simulation can take about five minutes of computation on high-performance computers. This makes deployment infeasible for real-time operational use. In addition, these methods often assume idealized sensor inputs, fixed boundary conditions, and known ventilation configurations which are assumptions that rarely hold on the fireground [7].

Recent advances in machine learning (ML) have proven highly effective for real-time predictions and pattern recognitions across numerous engineering applications [8-10]. The use of ML opens the opportunity to use sensor streams to learn nonlinear relationships between early fire signatures to forecast critical events before they occur. Motivated by these opportunities, the Artificial Intelligence–Enabled Smart Fire Fighting project initiative at the National Institute of Standards and Technology launched a multi-year research effort aimed at establishing a robust ML-based framework to provide firefighters actional information to enhance safety and situational awareness. The project focuses on three key research thrusts: 1) development of fire data generators, 2) real-time prediction of flashover under realistic constraints, and 3) interpretable ML techniques that support trust, transparency, and operational adoption.

3 Development of Fire Data Generators

A barrier to developing reliable ML models for hazard predictions is the scarcity of relevant, high-quality data. Full-scale fire experiments are expensive and so it is difficult to capture the full range of realistic fire scenarios and building configurations needed for robust model training. To address this gap, two physics-based synthetic data generation tools, namely the CFAST Fire Data Generator [11] and the FDS Data Generator [12], have been developed to systematically and automatically produce large numbers of simulated fire cases. These generators enable the collection of diverse and customizable datasets that support a wide range of ML-enabled tasks, including multi-compartment temperature recovery, flashover forecasting, and firefighter safety assessment.

3.1 CFAST Fire Data Generator (CData) [11]

The CFAST Fire Data Generator (CData)[1] was first introduced by Tam et al. (2023) [13] to support ML model development for predicting onset of rapid fire progression, such as flashover, in full-scale multi-compartment building structures. CData is built on the Consolidated Fire and Smoke Transport (CFAST) model [14], a widely used two-zone fire model capable of simulating temperature and smoke conditions in multi-compartment structures.

CData automates the end-to-end workflow of generating thousands of statistically varied CFAST simulations. The data generator tool consists of three major components, including 1) preprocessor module, 2) executor module, and 3) accumulator module, which can be used together to systematically sample fire scenarios with different material properties, fuel characteristics, fire growths, and venting conditions based on user-specified probability density functions. The sampling is specified using a four-letter namelist format that mirrors the coding structure of both FDS and CFAST which is intended to facilitate the ease of adoption for existing users. Once the sampling is completed, CData executes the CFAST simulation cases in parallel which significantly accelerate the collection of comprehensive multivariate data (i.e., temperature, gas concentrations, visibility, and/or heat flux). CData has been used extensively in recent research efforts [15-20], providing the foundation for training ML classifiers and regressors that can generalize more effectively to realistic fire scenarios.

3.2 FDS Data Generator (FD-Gen) [12]

[1] https://pages.nist.gov/cfast/

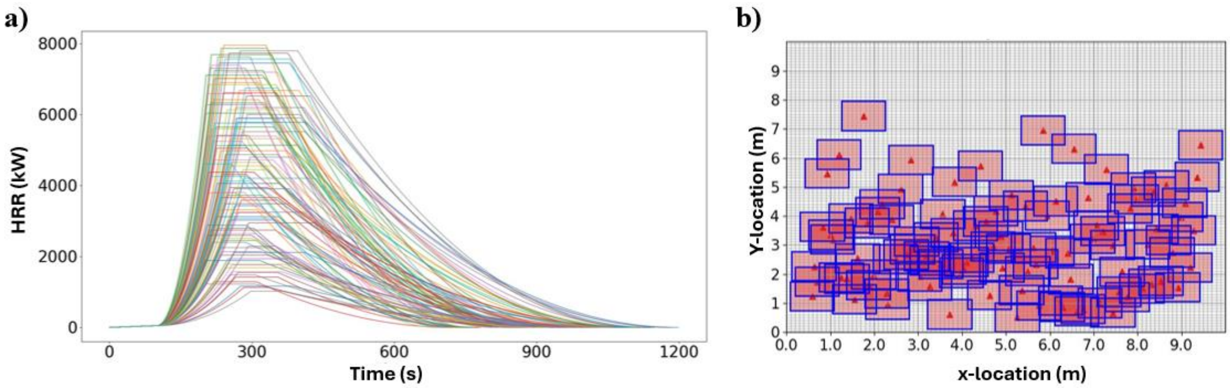

Figure 1. a) HHR profiles and b) the corresponding locations of the 50 generated fires.

Building on the synthetic data generation paradigm established by CData, FD-Gen v1.0.0[2], developed by Fang and Tam (2025) [12], extends automated data generation into the computational fluid dynamics (CFD) domain. Whereas CData relies on the two-zone model (CFAST), FD-Gen automates the creation of input files for the Fire Dynamics Simulator (FDS) [21] which enables the collection of high-fidelity data that capture the detailed physics of heat and mass transfer, smoke movement, and ventilation-driven flows at fine spatial resolution.

FD-Gen employs Latin Hypercube sampling [12] to statistically generate diverse fire and environment parameters, enabling a more uniform and efficient exploration of the parameter space compared with simple random sampling. Key randomized parameters include fire source coordinates, heat release rate (HRR), vent openings, door and window timing conditions, obstruction sizes and locations. Similar to CData, FD-Gen also uses a four-letter namelist to specify the details of the input files.

The FD-Gen workflow comprises four core steps: 1) preparing a base FDS case; 2) constructing an FD-Gen script; 3) executing FD-Gen; and 4) outputting the data. Using this automated process, FD-Gen can generate hundreds to thousands of FDS inputs with various fire scenarios within minutes which eliminates days of manual editing that would otherwise be required. Figure 1 shows the representative HRR curves and spatial distributions of ignition locations produced by FD-Gen. Additional details, including example applications in commercial buildings, road tunnels, and residential dwellings, are provided in reference [12].

4 Machine Learning Based Flashover Prediction Models

Three sequential studies were conducted to address the critical challenges limiting real-life flashover prediction. The first study focuses on sensor failure if there is a loss of temperature data when heat detectors exceed their operational limits well before flashover. The second study addresses dynamic fire and venting conditions in a realistic residential home with scenarios involving arbitrary fire locations, window breakage, and door openings. The third study tackles the challenge of structural variability. Each study introduces a progressively more advanced

[2] https://github.com/usnistgov/FD-Gen

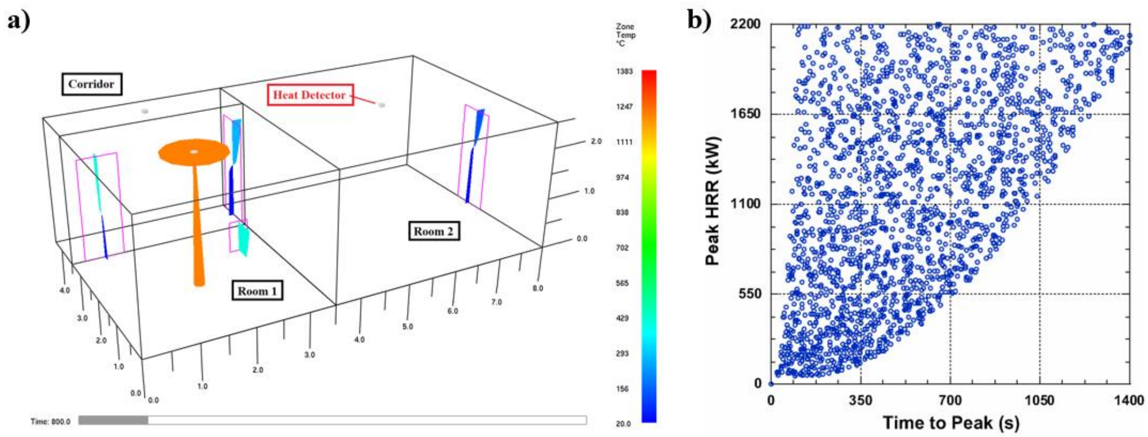

Figure 2. a) A schematic of the three-compartment structure and b) scatter plot of fire size.

modeling strategy to overcome the constraints and advance toward deployable, real-time, AI-enabled fireground decision-support systems.

4.1 Prediction model for Flashover occurrence (P-Flash) [15]

A major obstacle to reliable flashover prediction is the loss of temperature data caused by heat detector failure at elevated temperatures. In this study, heat detectors are assumed to fail at around 150 °C [15] which is well below the onset of flashover condition (i.e., upper layer gas temperature reaching 600 °C). As a result, temperature information becomes unavailable when predictive models need it most. Traditional fire models rely on continuous temperature input. However, when the temperature is missing due to heat detector failure, these models cannot provide any predictions. To address this limitation, the first study aims to develop a ML-based model capable of reconstructing the missing temperature information and using the recovered profiles to determine whether flashover conditions are met.

To train the P-Flash, synthetic temperature data are generated using CData. The study considers a three-compartment structure (Fig. 2a) with a heat detector located at the ceiling center of each compartment. A total of 1000 fire scenarios are simulated across a wide range of heat release rates and growth behaviors from slow to ultra-fast development (Fig. 2b).

P-Flash reconstructs missing temperature information through a multi-stage modeling framework. First, each temperature time series is segmented into phases based on detector failure timestamps which enables the model to focus on available temperature data (i.e., ≤ 150 °C). Next, statistical and trend-based features, such as mean temperature, maximum temperature, and first-derivative of temperature, are extracted using rolling windows. These features are then passed to two support vector regression models: the first uses early-phase data from all compartments and the second uses later-phase data where at least two detectors remain functional. Together, these regressors estimate the temperature evolution in the room of fire origin. When all detectors have failed, P-Flash employs a "learning-from-fitting" strategy to extrapolate the temperature growth using the earlier temperature patterns.

Results show that P-Flash accurately reconstructs missing temperature. The mean absolute errors for recovered temperatures are about 11 °C when one detector has failed, about 13 °C when two detectors have failed, and about 30 °C when all detectors are unavailable. The model correctly

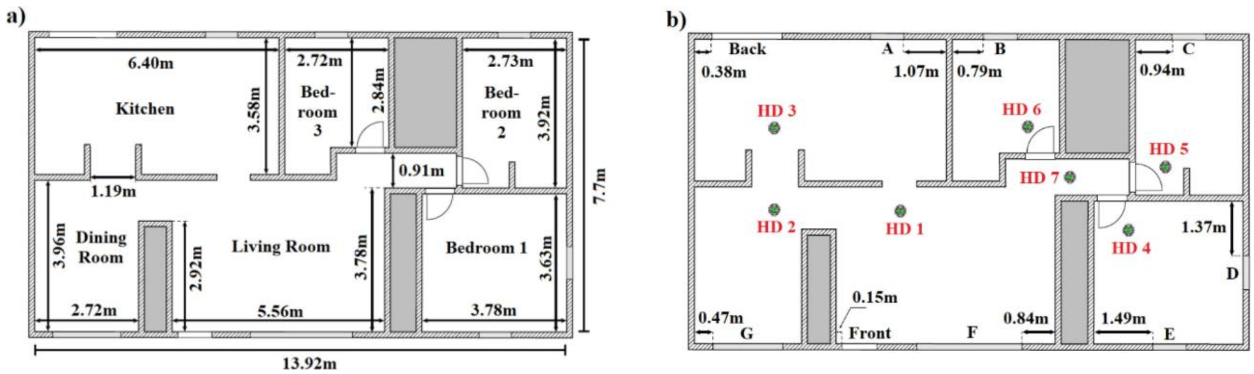

Figure 3. Plan view dimensioned drawing of a) the single-story structure [21] and b) vent openings with heat detectors (HD).

identifies flashover occurrence within ± 20 s in approximately 83 % of cases and even when forecasting 150 s ahead, P-Flash achieves an accuracy of about 81 % which demonstrates strong early-warning capability.

4.2 P-Flash Version 2 [16,17]

Another challenge in flashover prediction is that real fireground conditions involve rapidly changing venting conditions (i.e., unpredictable window or door openings) in addition to the loss of temperature data caused by heat detector failure. These dynamic changing events can dramatically alter oxygen availability and heat-flow patterns, creating complex and nonlinear temperature behaviors that are difficult for traditional models to capture. The objective of the second study is therefore to develop a ML model that provides forecasting for a more complex residential structure accounting for detector operational temperature limits and arbitrary venting and fire conditions.

To train such a model, a learning-by-synthesis approach is used to collect a large and diverse dataset using CData. A total of 110,000 fire simulations are performed in a seven-compartment, single-story residential structure (see Fig. 3). The dataset covers peak heat release rates ranging from 100 kW to 5100 kW with varying fire locations, arbitrary door-opening time of 1 s to 900 s, and window breakage thresholds from 100 °C to 200 °C. These variations are designed to capture the complex temperature behavior due to window breakage and door opening.

In this study, P-Flashv2 is developed using an attention-based bidirectional gated recurrent unit architecture. The bidirectional method enables the model to learn long range temporal dependencies while the sensor-wise self-attention mechanism identifies which compartment temperature data are the most informative. P-Flashv2 also uses dropout regularization and early stopping to ensure model generalization. The model inputs are temperature segments extracted using sliding windows and the outputs are probabilities of flashover occurrence in future time.

P-Flashv2 demonstrates strong predictive performance. For 30-s-ahead predictions, the model achieves about 87.7 % accuracy, outperforming eight different state-of-the-art recurrent neural network baselines. For 60-s-ahead predictions, accuracy is about 89.5 %, indicating robust early-warning capability. Real-world applicability is assessed by validating the model against full-scale fire experiments and the results show that the model has about 82.7 % and 85.6 % accuracy for 30-s-ahead and 60-s-ahead cases. Additional analyses on the attention maps are conducted and it

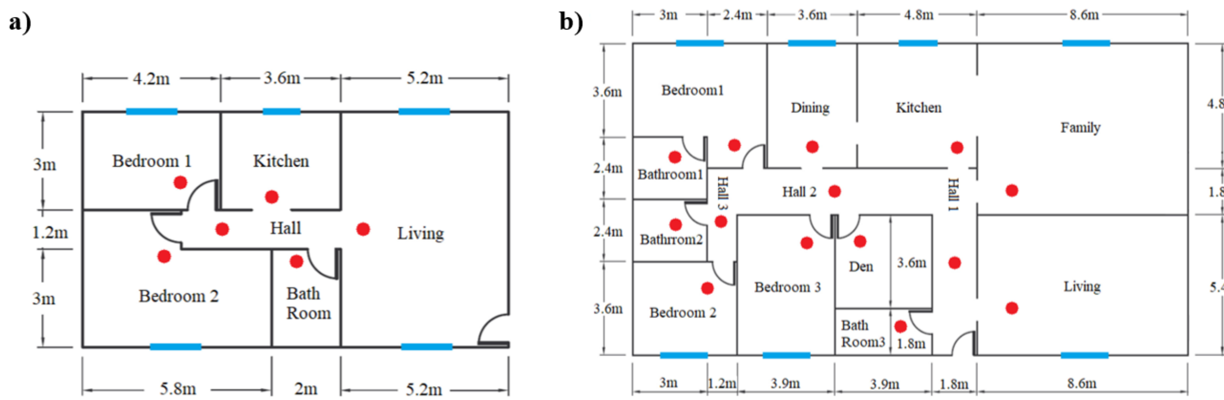

Figure 4. Layout of a) 6-compartment and b) 14-compartment homes with HD shown in red dots.

is observed that temperature data from compartments near the room of fire origin contribute most strongly to the model's predictions. This finding provides interpretable insight into the model's decision-making process.

4.3 Flashover Prediction Neural Network (FlashNet) [18]

While the first two studies focus on prediction under sensor failure and dynamic fire and venting conditions within a single building layout, the third study tackles challenges associated with structure variability. In real-life scenarios, homes differ substantially in size, geometry, compartment connectivity, and venting configurations. These structural differences influence how heat and smoke propagate which makes it difficult for a single model to perform reliably across different layouts. Existing state-of-the-art approaches typically rely on layout-specific assumptions. This means that if there are twenty different residential floorplans, twenty separate models are needed which is impractical for real-life applications. For that, this study addresses this challenge by developing a scene-agnostic model capable of predicting flashover accurately across diverse residential structures.

A total number of 78,000 simulated fire scenarios are generated using CData. These cases include 17 representative U.S. residential floorplans, ranging from compact multi-room apartments to larger single-family homes with 3 to 14 compartments. Figure 4 shows the 6-compartment and 14-compartment residents building structures. Similar to the previous studies, a wide range of fire locations, heat release rates (i.e., 50 kW to 4600 kW), and temperature-triggered window breakage conditions are considered. Because of the operational limitations of heat detectors, all temperature readings above 150 °C are truncated. Rolling window is used to extracted temperature segments.

The Flashover Prediction Neural Network (FlashNet) achieves layout generalization by converting each home into graph representations. The model captures the complex temperature variation using a spatiotemporal graph convolutional neural network. In this formulation, compartments are treated as nodes while door openings define edges that represent potential pathways for heat and smoke movement. A key innovation is the geometric average adjacency matrix, which aggregates structural connectivity across all 17 floorplans into a unified representation. This allows the model to learn common structural patterns and temperature behaviors without requiring separate architectures for each layout. Temporal features are extracted using 1-D temporal convolutions and

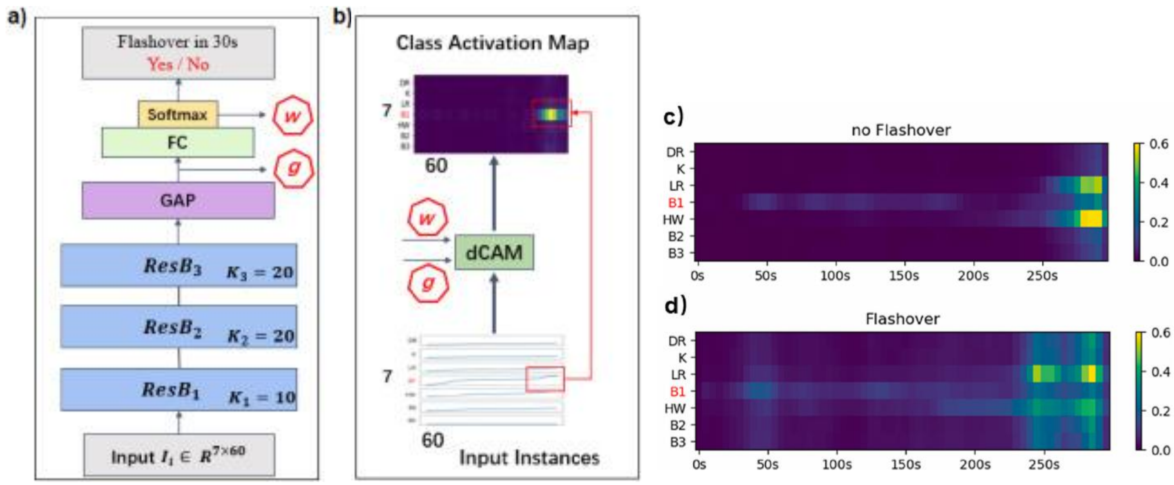

Figure 5. a) Overall model structure of xFlashNet, b) information flow to obtain the dimension-wise class activation map, and dCAM for c) a no flashover case and d) a flashover case.

spatial dependencies are learned using Chebyshev-based graph convolutions. FlashNet outputs flashover likelihood in 10-s and 30-s lead times.

Benchmarking results demonstrate that FlashNet outperforms five state-of-the-art baselines using Support Vector Machine, Multi-layer Perceptron, Long Short-Term Memory (LSTM), Bidirectional LSTM, and simple Convolutional Neural Networks. With a 30-s lead time, FlashNet achieves about 92.1 % accuracy with balanced precision and recall. Ablation studies show that both temporal and graph convolution layers contribute significantly to enhance model robustness. Importantly, FlashNet maintains strong, consistent performance across nearly all 17 residential structures with about 94 % accuracy.

5 Model Interpretability

Although ML models have demonstrated strong capability for predicting rapid fire events such as flashover, heir adoption in fire training or emergency response has been limited by a fundamental barrier: a lack of interpretability. Because ML-derived predictions may influence life-critical decisions, incident commanders and firefighters must understand why a model reaches a particular conclusion and whether its reasoning is consistent with known knowledge in fire dynamics. For that, addressing this interpretability challenge is essential for building trust, validating the physical relevance of learned patterns, and ensuring safe deployment of ML-driven decision-support tools.

The first interpretability study [19] aims to address the black-box nature of deep neural networks by developing a visualization-based framework for explaining flashover predictions. xFlashNet (explainable FlashNet) is introduced and it is based on a residual convolutional neural network trained on multi-compartment temperature data and paired with dimension-wise Class Activation Maps (dCAM). dCAM revealed which portions of the temperature information most strongly contributed to the model decision. Fig. 5a shows the learned features (*g*) and the weights from Softmax layer (*w*) provided to dCAM and Fig. 5b shows the resulting attribution map, where brighter regions indicate more important temperature segments. Using this method, the study

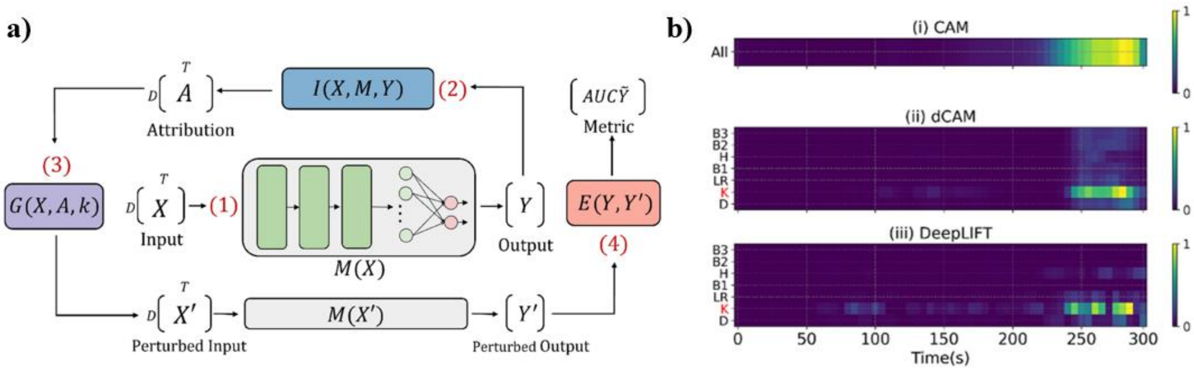

Figure 6. a) Schematic of the interpretability analysis framework and b) attribution results for a fire case with i) CAM, ii) dCAM, and iii) DeepLIFT.

demonstrates that the model focuses on physically meaningful cues. For example, it focuses the rise of temperatures in the room of origin and the temperature increase in adjacent compartments (see Fig, 5c and 5d). These findings suggest that deep learning models can learn interpretable and domain-consistent features which provide a step-forward for transparent ML-based fire forecasting.

However, the initial approach also highlights key methodological limitations. Most notably, dCAM suffers from a dispersion effect where attribution scores appear smeared across multiple timestamps and compartments. For example, dCAM highlights broad regions of importance. In addition, dCAM lacks a mechanism to evaluate whether the highlighted regions are in fact influential or only simply artifacts. These shortcomings motivate the need for a more robust and quantitatively validated interpretability approach.

The second study [20] builds on these gaps by introducing a more robust, systematic interpretability framework based on DeepLIFT, a backpropagation-based attribution method that produces element-level importance scores which overcomes the temporal smearing seen in dCAM. Fig. 6a outlines the workflow of the interpretability framework, including 1) model execution, 2) attribution score computation, 3) targeted perturbation testing, and 4) quantitative evaluation of attribution relevance. Fig. 9b shows the attributions obtained from three different interpretability methods: a) CAM, b) dCAM, and c) DeepLIFT. As shown in Fig. 9b, DeepLIFT produces more localized attributions compared to CAM and dCAM. Through parametric analysis, this study further identifies optimal reference values for DeepLIFT and this finding highlights that the quality of the interpretability results depends on careful selection of the reference value.

Together, the studies mark progress in interpretable ML and form a comprehensive framework for deploying reliable, transparent, and physically grounded ML systems for flashover prediction and broader smart firefighting applications.

6 Conclusions and Future Work

This research establishes a comprehensive ML framework for real-time flashover prediction that addresses the major limitations of existing models: sensor failure, dynamic ventilation, and structural variability. The three studies, including P-Flash, P-Flashv2, and FlashNet, demonstrate that ML can provide accurate, early hazard forecasts under increasingly realistic fireground constraints. Together, they form the foundation for next-generation AI-enabled smart firefighting systems that enhance situational awareness, reduce risk, and support more informed tactical decision-making.

Future work will extend these hazard-prediction capabilities into a full AI-driven recommendation system. Building on the tenability-based reinforcement learning (RL) framework developed in [23], we will integrate ML hazard forecasts with RL-based path planning to generate real-time evacuation guidance. This system will dynamically update safe routes as conditions evolve, enabling both firefighters and occupants to make safer, faster decisions. Ultimately, this research aims to advance toward intelligent, adaptive, and trustworthy fireground decision-support tools that significantly improve life safety outcomes.

7 References

[1] Ahrens, M., & Maheshwari, R. (2021). *Home structure fires*. National Fire Protection Association, Quincy, MA.

[2] Kerber, S. (2012). Analysis of changing residential fire dynamics and its implications on firefighter operational timeframes. *Fire Technology*, 48(4), 865–891.

[3] Kim, H. J., & Lilley, D. (1999, January). Comparison of theories for room flashover. In *37th Aerospace Sciences Meeting and Exhibit* (p. 343).

[4] Babrauskas, V. (1980). Estimating room flashover potential. *Fire Technology*, 16(2), 94–103.

[5] Overholt, K. J., & Ezekoye, O. A. (2012). Characterizing heat release rates using an inverse fire modeling technique. *Fire Technology*, 48(4), 893–909.

[6] Jahn, W. (2017). Using suppression and detection devices to steer CFD fire forecast simulations. *Fire Safety Journal*, 91, 284–290.

[7] Koo, S. H., Fraser-Mitchell, J., & Welch, S. (2010). Sensor-steered fire simulation. *Fire Safety Journal*, 45(3), 193–205.

[8] Tam, W. C., Yuen, W. W., & Chow, W. K. (2016). Numerical study on the importance of radiative heat transfer in building energy simulation. *Numerical Heat Transfer, Part A: Applications*, 69(7), 694–709.

[9] Tam, W. C., & Yuen, W. W. (2019). *OpenSC: An open-source calculation tool for combustion mixture emissivity/absorptivity*. National Institute of Standards and Technology, Gaithersburg, MD, USA.

[10] Li, J., Brown, C., Dzikowicz, D. J., Carey, M. G., Tam, W. C., & Huang, M. X. (2023). Towards real-time heart health monitoring in firefighting using convolutional neural networks. *Fire Safety Journal*, 140, 103852.

[11] Reneke, P. A., Peacock, R. D., Gilbert, S. W., & Cleary, T. (2021). *CFAST – Consolidated Fire and Smoke Transport (Version 7) Volume 5: CFAST Fire Data Generator (CData)*. NIST Technical Note 1889v5, National Institute of Standards and Technology, Gaithersburg, MD.

[12] Fang, H., & Tam, W. C. (2025). *Fire Data Generator (FD-Gen) v1.0.0*. NIST Technical Note 2332, National Institute of Standards and Technology, Gaithersburg, MD.

[13] Tam, W. C., Fu, E. Y., Peacock, R., Reneke, P., Wang, J., Li, J., & Cleary, T. (2020). Generating synthetic sensor data to facilitate machine learning paradigm for prediction of building fire hazard. *Fire Technology*, 56, 1–22.

[14] Peacock, R. D., Reneke, P. A., & Forney, G. P. (2017). *CFAST – Consolidated Model of Fire Growth and Smoke Transport (Version 7) Volume 2: User's Guide*. NIST Technical Note 1889v2, National Institute of Standards and Technology, Gaithersburg, MD.

[15] Wang, J., Tam, W. C., Jia, Y., Peacock, R., Reneke, P., Fu, E. Y., & Cleary, T. (2021). P-Flash - A machine learning-based model for flashover prediction using recovered temperature data. *Fire Safety Journal*, 122, 103341.

[16] Fu, E. Y., Tam, W. C., Wang, J., Peacock, R., Reneke, P. A., Ngai, G., Leong, H. V., & Cleary, T. (2021, May). Predicting flashover occurrence using surrogate temperature data. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(17), 14785–14794.

[17] Tam, W. C., Fu, E. Y., Li, J., Peacock, R., Reneke, P., Ngai, G., Leong, H. V., Cleary, T., & Huang, M. X. (2023). Real-time flashover prediction model for multi-compartment building structures using attention-based recurrent neural networks. *Expert Systems with Applications*, 223, 119899.

[18] Tam, W. C., Fu, E. Y., Li, J., Huang, X., Chen, J., & Huang, M. X. (2022). A spatial temporal graph neural network model for predicting flashover in arbitrary building floorplans. *Engineering Applications of Artificial Intelligence*, 115, 105258.

[19] Fan, L., Tam, W. C., Tong, Q., Fu, E. Y., & Liang, T. (2023). An explainable machine learning-based flashover prediction model using dimension-wise class activation map. *Fire Safety Journal*, 136, 103849.

[20] Fan, L., Tong, Q., Fang, H., Zhong, W., Tam, W. C., & Liang, T. (2025). An interpretability analysis framework to enhance deep learning model transparency: A case study on flashover prediction using time-series sensor data. *Fire Technology*, 61, 1–26.

[21] McGrattan, K., Hostikka, S., McDermott, R., Floyd, J., Weinschenk, C., & Overholt, K. (2013). *Fire Dynamics Simulator: Technical Reference Guide, Volume 1 – Mathematical Model*. NIST Special Publication 1018-1.

[22] Madrzykowski, D., & Weinschenk, C. (2019). *Impact of fixed ventilation on fire damage patterns in full-scale structures*. UL Firefighter Safety Research Institute, Columbia, MD.

[23] Fang, H., Tam, W. C., Lovreglio, R., Emon, M. I., & Huang, M. X. (2025). Development of a tenability-based path planning model for building fire evacuations using reinforcement learning. *Journal of Building Engineering*, 80, 108557.