Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Case-based Reasoning for Fire Accident Analysis and Loss Prevention: Evidence from Greenhouse Facilities

Bing Zhang
Yangzhou University
Xin Wan
mz120251087@stu.yzu.edu.cn

Yangzhou University
Yizhou Sun
Yangzhou University
Martin Skitmore
Bond University
Xiaoqian Zu
Yangzhou University
Zhijiang Wu
Yangzhou University
Shuang Dong
East China University of Science and Technology

Research Article

Keywords:

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11395661/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Case-based Reasoning for Fire Accident Analysis and Loss Prevention: Evidence from Greenhouse Facilities

A B S T R A C T

From an accident analysis perspective, greenhouse fires are low-frequency, high-uncertainty events where conventional statistical and physics-based models offer limited support. Greenhouse facilities face increasing fire risk due to electrification, combustible materials, and complex operations. To support systematic analysis of fire accidents in process-like production environments, this study proposes a Case-Based Reasoning (CBR) framework with a multidimensional case representation integrating environmental, facility, and incident attributes.

Using 101 validated greenhouse fire cases (2005–2024), the entropy weight method determines indicator importance across 13 variables. Local similarity combines discrete and continuous approaches, while global aggregation retrieves highly similar cases. Results show that ignition source, personnel presence, and accident severity dominate similarity. Recurring vulnerabilities include electrical ignitions, lack of firefighting infrastructure, and noncompliant facility use. Reverse validation confirms framework stability and reproducibility.

Rather than predicting fire dynamics, the framework enables structured learning from historical accidents under data constraints, transforming empirical data into actionable knowledge for accident prevention and loss mitigation.

Keywords: Case-Based Reasoning； Greenhouse Fires； Fire Risk Assessment； Entropy Weighting； Accident analysis； loss prevention

1. Introduction

Greenhouse-based production facilities constitute an integral component of modern food production systems, contributing to yield stability, resource efficiency, and climate resilience [1, 2]. Concurrently, increasing electrification, the widespread use of combustible envelope materials, dense vegetation, and informal operational practices have heightened the vulnerability of greenhouse facilities to fire incidents, which may lead to substantial economic losses, production disruption, and adverse safety outcomes. From a loss prevention perspective, greenhouse fires have similarities to incidents in small-scale process facilities, where electrical energy, combustible materials, and nonstandard operational practices interact.

Fire incidents in greenhouse facilities have consequences that extend beyond direct structural and crop losses. Severe events may disrupt agricultural production chains, increase material and energy consumption, and expose deficiencies in facility management and safety governance [3]. In recent years, the intensification and electrification of greenhouse operations have been accompanied by increasing fire risk, particularly in regions where safety standards and monitoring systems lag behind technological deployment [4]. These trends highlight the need for systematic analysis of fire accidents and learning approaches tailored to process-like production environments.

Compared with conventional buildings, greenhouses have distinctive fire risk characteristics. Lightweight structural frames and combustible envelope materials, such as plastic films, facilitate rapid fire spread [5]. Dense vegetation contributes to elevated fire loads, while extensive electrical installations for heating, irrigation, and environmental control introduce dominant ignition sources, often operating under high-humidity conditions that accelerate

equipment degradation [6, 7]. In addition, informal modifications and multifunctional uses of greenhouse facilities result in heterogeneous configurations that complicate the application of standardized safety regulations [8]. These characteristics give rise to complex, context-dependent fire incident patterns that are not readily captured using static or rule-based assessment methods.

Research into accident prevention and loss mitigation has predominantly focused on residential, industrial, or commercial buildings, emphasizing fire dynamics simulation, probabilistic analysis, or code compliance [9, 10]. While these approaches provide valuable insights into fire behavior, their applicability to greenhouse systems remains constrained. Model-driven methods rely on predefined structural parameters that are often unavailable in agricultural facilities, whereas data-driven approaches are constrained by the relative scarcity and heterogeneity of greenhouse fire incidents [11, 12]. Furthermore, many studies prioritize physical fire development rather than decision support for prevention and risk management in facility agriculture [13].

Historical fire accident cases represent a valuable yet underutilized source of knowledge for agricultural facility safety analysis. Case-Based Reasoning (CBR) enables new problems to be addressed by retrieving and adapting solutions from similar prior cases, making it well-suited to domains characterized by uncertainty, mixed-type data, and limited sample sizes [14]. Although CBR has been successfully applied in other safety-critical domains [15–17], its application to greenhouse fire accident analysis and learning remains limited.

To address this gap, this study proposes a CBR-based analytical framework for analyzing greenhouse fire accidents. A multidimensional case representation system is developed by

integrating environmental conditions, facility characteristics, and incident attributes. The entropy weight method (EWM) is used to determine indicator importance objectively, and a hybrid similarity computation strategy is introduced to enhance case retrieval accuracy. By transforming historical accident data into actionable system knowledge, the proposed framework supports data-informed decision-making and contributes to improved accident prevention and loss mitigation.

2. Literature review

2.1 *Fire accidents and safety challenges in greenhouse systems*

Fire risk in greenhouse facilities has attracted increasing attention as facility agriculture becomes more intensive and technologically complex. In contrast to conventional buildings, greenhouses are characterized by lightweight structural systems, extensive use of combustible envelope materials, dense vegetation, and highly integrated electrical and mechanical equipment. These features collectively contribute to rapid fire propagation, elevated fire loads, and challenging suppression conditions [5, 7]. From a safety science perspective, greenhouse fires have characteristics of systemic accidents, in which technical conditions, organizational practices, and human factors jointly influence accident evolution and loss severity [10, 11].

From a risk assessment perspective, a range of methodological approaches has been applied to fire safety analysis in agricultural and building systems. These include fire dynamics modeling, probabilistic risk assessment, and code-based evaluation frameworks, which provide insights into ignition mechanisms, fire development, and evacuation performance. However,

despite these advances, studies tend to examine individual factors in isolation and rarely capture the coupled effects of environmental conditions, facility configuration, and operational behavior that shape greenhouse fire risk. In addition, many assessment approaches emphasize static risk classification rather than supporting adaptive decision-making and early warning in facility agricultural systems.

2.2 Case-based reasoning (CBR) and similarity measurement in safety engineering

CBR is a problem-solving model in which new problems are addressed by retrieving and adapting solutions from similar historical cases [14]. It is particularly suited to safety-critical domains characterized by uncertainty, heterogeneous data, and limited sample sizes. Its transparent reasoning process enables assessment outcomes to be traced to specific prior events, thereby enhancing interpretability and practical applicability.

In recent years, CBR has been applied across a range of safety-related domains, including construction safety management, industrial fault diagnosis, and emergency decision-making [15–18]. These applications demonstrate CBR's capacity to integrate structured and unstructured data, facilitating the systematic reuse of experiential knowledge embedded in historical accident records.

Similarity measurement constitutes a core component of CBR performance. Conventional approaches commonly use k-nearest neighbor (kNN) algorithms combined with predefined distance metrics. To enhance retrieval accuracy, hybrid similarity schemes accommodating mixed-type variables have been developed, alongside objective weighting methods that reduce subjectivity in determining feature importance [19, 20]. Entropy-based weighting has been

widely recognized as an effective technique for quantifying indicator importance based on data dispersion, particularly in safety risk assessment contexts [21, 22].

Despite these developments, CBR applications in fire safety have largely focused on industrial or conventional building environments, with limited attention to agricultural facilities. Furthermore, few studies have systematically integrated entropy-based weighting with mixed-variable similarity computation for the analysis of greenhouse fire incidents. This limits the applicability of current CBR approaches for decision support in facility agriculture, where data heterogeneity and operational variability are pronounced.

2.3 Multi-criteria weighting approaches in fire risk assessment

Multi-criteria decision-making (MCDM) methods play an important role in fire risk assessment by aggregating diverse indicators that represent environmental, structural, and operational factors. Of these methods, the Analytic Hierarchy Process (AHP) has been widely applied to structure complex decision problems based on expert judgment [23, 24], whereas entropy weighting provides an objective alternative by deriving indicator importance directly from data distributions [25].

However, the application of these weighting approaches to greenhouse fire risk assessment remains limited. Studies typically apply weighting techniques within static evaluation frameworks and do not integrate them into dynamic reasoning systems capable of learning from historical cases. This limitation constrains the ability to identify recurring risk patterns and to support adaptive safety management in facility agricultural systems.

2.4 Critical review and research gap

Notwithstanding progress in understanding fire dynamics, material behavior, and isolated risk factors in greenhouse systems, the literature remains constrained from an accident-learning and safety-science perspective. Prevailing approaches are characterized by reductionist analyses that overlook systemic interactions, reliance on static indicator-based models that lack the capacity to learn from historical accidents, and limited application of CBR tailored to greenhouse fire risk assessment, particularly approaches integrating objective weighting with similarity-based reasoning for heterogeneous accident data.

These limitations indicate the need for an analytical framework that advances beyond static, factor-isolated evaluations toward system-level reasoning, structured reuse of empirical fire cases, and adaptive decision support under conditions of data scarcity. The proposed CBR framework addresses these requirements by conceptualizing greenhouse fire risk as a systemic problem, incorporating entropy-based objective weighting, and enabling robust similarity-based retrieval of historical cases, thereby establishing the methodological foundation for this study.

3. Methodology

3.1 Framework overview

Greenhouse fire accidents arise from coupled interactions among environmental conditions, facility configurations, operational practices, and human behavior. Accident-oriented reasoning approaches remain largely static and factor-isolated, limiting their ability to

capture the systemic nature of greenhouse facilities or to explain divergent fire outcomes under similar physical conditions. In addition, conventional indicator-based models lack mechanisms for learning from historical accidents and for supporting adaptive, experience-informed decision-making.

To address these limitations, greenhouse fire risk is conceptualized as a system-level problem, and a CBR framework is developed to enable the structured reuse of historical fire knowledge. CBR is particularly appropriate in this context due to the relative infrequency of greenhouse fire events, the heterogeneity and incompleteness of accident data, and the importance of experiential knowledge in understanding risk evolution. Rather than relying on predefined causal rules or large training datasets, the framework evaluates new scenarios by retrieving and comparing similar historical cases, thereby providing transparent, interpretable decision support. The framework is not designed to infer causality or predict accident occurrence, but to facilitate structured comparison and learning from observed accident patterns.

A multidimensional case representation system is constructed across three complementary dimensions: environmental, facility, and incident characteristics. This structure reflects the coupled nature of greenhouse systems and supports systematic cross-case comparison. To reduce subjectivity in indicator weighting, the EWM is applied to quantify each indicator's information contribution based on data dispersion, thereby embedding objective weighting within a data-driven reasoning process.

Similarity computation constitutes the core retrieval mechanism. Given the presence of both categorical and numerical variables, a hybrid local similarity strategy is adopted,

combining discrete scoring for qualitative indicators with continuous interpolation for quantitative variables. Global similarity is then calculated as the entropy-weighted aggregation of local similarities, enabling the identification of historical cases with comparable risk profiles and outcomes.

The overall CBR workflow follows the classical retrieve–reuse–revise–retain (4R) model and comprises seven stages: case collection, indicator system construction, entropy-based weighting, local similarity calculation, global similarity aggregation, case retrieval, and validation (Fig. 1). Through this process, empirical accident records are transformed into actionable system knowledge, enabling the identification of recurrent risk patterns and supporting adaptive safety management in facility agriculture. In this manner, the framework operationalizes the research gaps identified in the literature by integrating objective weighting and similarity reasoning within a system-oriented CBR process.

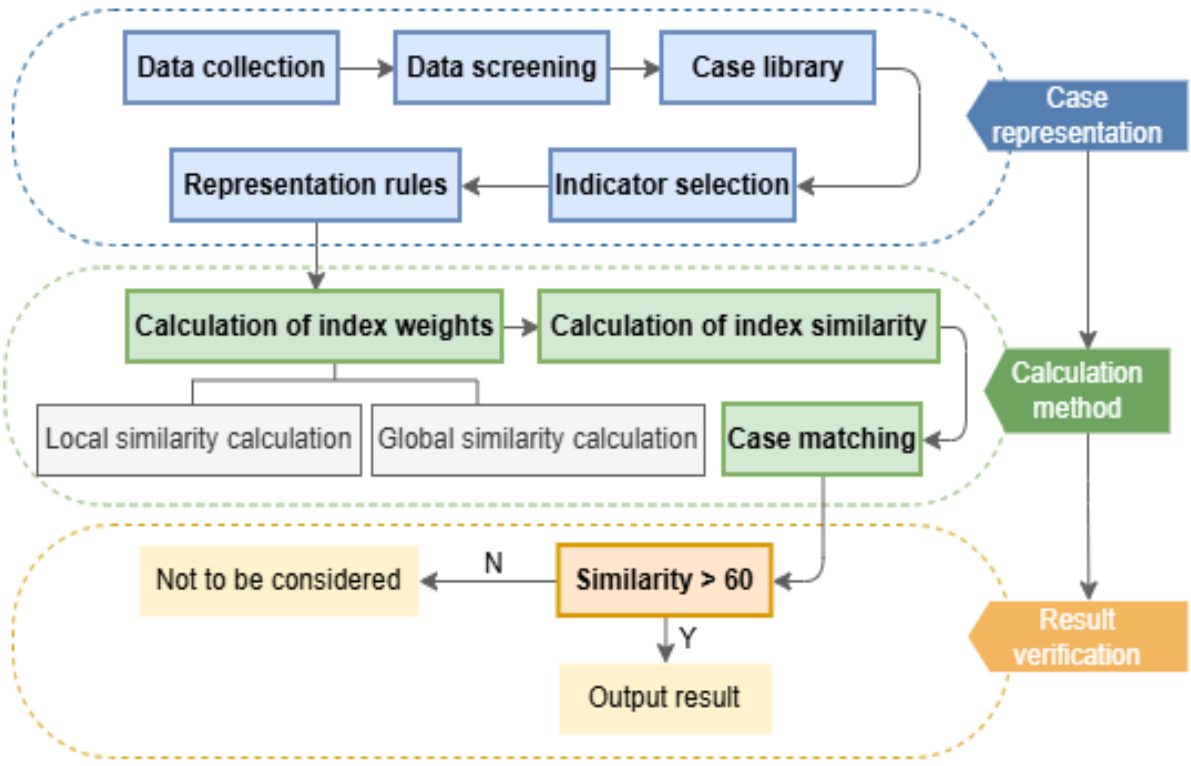

Fig. 1. CBR model framework

3.2 Data collection and cleaning rules

The construction of a reliable case database is important for CBR, particularly in domains where accident data are infrequent, heterogeneous, and inconsistently documented, as is typical of greenhouse fire incidents. To ensure analytical robustness, a structured multi-stage data collection and screening procedure is implemented.

An initial dataset of 200 greenhouse fire reports covering the period 2005–2024 is compiled from authoritative public sources, including officially documented accident investigation reports and widely used media sources in retrospective safety research, such as fire rescue records and major national media platforms (e.g., Weibo, Sohu, NetEase). These sources collectively provide the most comprehensive accessible documentation of agricultural facility fires in the absence of a unified national accident database. A three-stage screening procedure is subsequently applied to improve data consistency and reliability (Fig. 2).

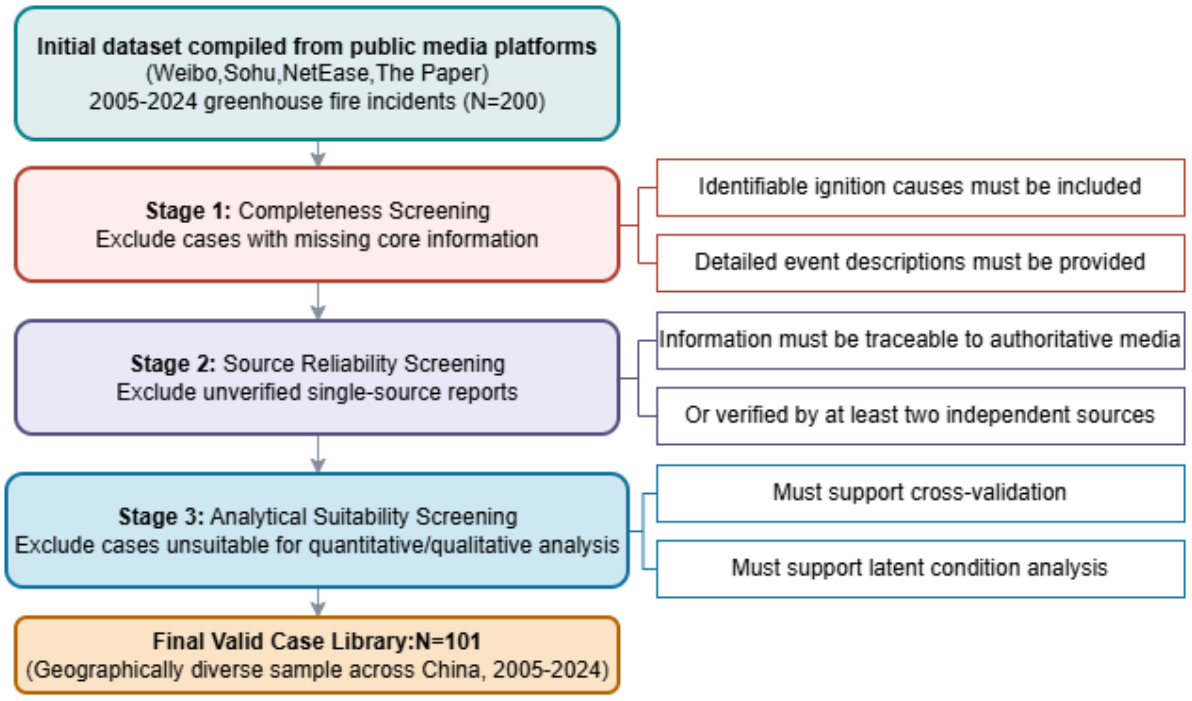

Fig. 2. Flowchart of greenhouse fire case screening

Following this protocol, 101 validated greenhouse fire cases are retained. The resulting dataset spans two decades and includes diverse geographic regions, facility types, and accident outcomes, thereby providing a representative basis for identifying recurrent fire accident patterns under data-constrained conditions.

To support structured reasoning, influencing factors are categorized into two functional levels: fire occurrence-related factors and consequence-related factors. These are subsequently mapped onto three macro-dimensions—environmental, facility, and incident characteristics—forming a standardized case representation framework for CBR analysis (Fig. 3). This taxonomy operationalizes the system-level perspective identified in the literature and ensures methodological consistency from data preparation through to case retrieval.

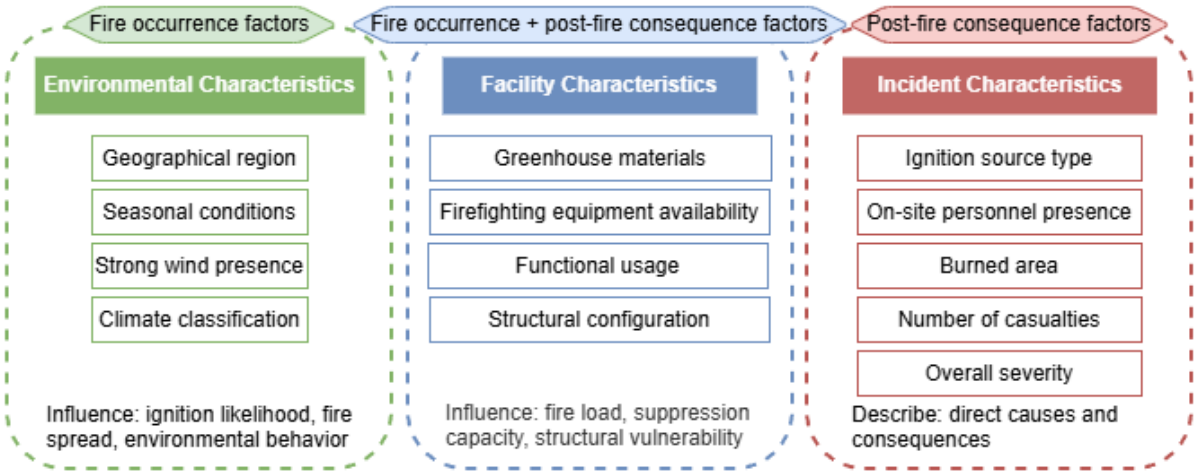

Fig. 3. Taxonomy of greenhouse fire risk factors

3.3 Indicator system

A robust indicator system is central to case representation and similarity reasoning in CBR. Greenhouse fire risk emerges from the combined influence of environmental conditions, facility attributes, and incident-specific factors, rather than from isolated variables [26, 27].

The indicator system is therefore designed to reflect this systemic nature while remaining compatible with heterogeneous historical accident data.

Following CBR principles and prior fire risk studies in agricultural facilities [28, 29], 13 indicators are selected and organized into three macro-dimensions: environmental characteristics, facility characteristics, and incident characteristics. This tripartite structure provides a comprehensive yet parsimonious representation, enabling systematic cross-case comparison without introducing unnecessary model complexity.

Environmental characteristics are external conditions that influence fire exposure and propagation, including geographic region, season, climate type, and strong winds. These factors influence ignition likelihood and fire development indirectly through their effects on microclimate and emergency response conditions [27, 30].

Facility characteristics describe physical and operational attributes, including construction materials, structural form, functional use, coverage area, and the availability of firefighting equipment [29, 31]. These variables determine fire load, structural vulnerability, and initial suppression capability. Although some attributes have limited variability, their inclusion establishes baseline safety conditions and supports the interpretation of outcome variability when combined with other dimensions.

Incident characteristics capture direct causes and consequences, including ignition source type, presence of on-site personnel, number of casualties, and overall accident severity [28, 32]. These indicators reflect system performance under fire conditions and are important for distinguishing cases with similar antecedent conditions but different outcomes, thereby enhancing the relevance of decision support.

By structuring the indicator system around system-relevant dimensions and outcome-oriented variables, the framework establishes a standardized case representation that directly supports entropy-based weighting and hybrid similarity computation in subsequent analysis.

3.4 Entropy weight calculation

Indicator weights play a critical role in similarity computation and case retrieval within the CBR framework. Conventional fire risk assessment methods frequently rely on subjective approaches such as the AHP, which are less appropriate for heterogeneous and incomplete accident datasets characteristic of greenhouse fire incidents. To address this limitation, the EWM is used to objectively determine indicator importance based on inherent information dispersion within the dataset. Indicators with greater variability across cases exhibit stronger discriminatory power and are assigned higher weights, whereas those with limited variation contribute proportionally less to the similarity evaluation.

For a dataset of m fire cases and n indicators, let x_{ij} denote the value of the j-th indicator for the i-th case. For benefit-type indicators, the entropy weight is calculated as follows, with larger values indicating higher risk contribution $\varepsilon = 10^{-10}$. Let , then

$$p_{ij} = \frac{x_{ij} - \min\left(x_j\right)}{\max\left(x_j\right) - \min\left(x_j\right)} + \varepsilon \tag{1}$$

For cost-type indicators, where smaller values correspond to higher risk contribution, Equation (2) is used.

$$p_{ij} = \frac{\max\left(x_j\right) - x_{ij}}{\max\left(x_j\right) - \min\left(x_j\right)} + \varepsilon \tag{2}$$

Moreover, the entropy e_j of the j-th indicator is Equation (3):

$$e_j = -k\sum_{i=1}^{m} p_{ij} \ln\left(p_{ij}\right) \tag{3}$$

The weight coefficient is Equation (4):

$$w_j = \frac{d_j}{\sum_{j=1}^{n} d_j} \tag{4}$$

Unlike static weighting methods, EWM is embedded directly into the CBR reasoning loop; the derived weights modulate each indicator's contribution to the global similarity score, ensuring that empirically discriminative factors dominate case retrieval.

The resulting weights (Table 1) show that accident-related indicators, particularly on-site personnel presence and accident severity, receive the highest coefficients, confirming their dominant role in distinguishing greenhouse fire patterns. Conversely, indicators with limited variability contribute negligibly. This objective weighting scheme provides a transparent, data-driven basis for integrating heterogeneous indicators into CBR similarity assessment.

Table 1

Characterization system for greenhouse fire features.

Indicator code	Indicator name	Information entropy value	Information utility value	Weight coefficient
X1	Urban region	0.9233	0.0767	4.33%
X2	Season	0.8511	0.1489	8.41%
X3	Greenhouse materials	0.9946	0.0054	0.31%
X4	Firefighting equipment	0.8124	0.1876	10.59%
X5	Type of fire source	0.8296	0.1704	9.62%
X6	Strong wind	0.9856	0.0144	0.81%
X7	On-site personnel	0.6332	0.3668	20.71%
X8	Coverage area	0.9971	0.0029	0.17%
X9	The uses of greenhouses	0.988	0.012	0.68%
X10	Shape type	0.7776	0.2224	12.56%
X11	The number of deaths	0.8144	0.1856	10.48%
X12	Climate characteristics	0.9777	0.0223	1.26%
X13	The severity of the accident	0.6444	0.3556	20.08%

3.5 Similarity computation

Similarity computation is the core retrieval mechanism of the CBR framework. Given the heterogeneous nature of greenhouse fire data, which comprises categorical, ordinal, and continuous variables, a two-level strategy is adopted: local similarity for individual indicators and global similarity for overall case matching.

Similarity computation serves as the core retrieval mechanism within the CBR framework. Given the heterogeneous nature of greenhouse fire data, which includes categorical, ordinal, and continuous variables, a two-level strategy is adopted: local similarity is calculated for individual indicators. In contrast, global similarity is used for overall case matching.

3.5.1 Local similarity

For categorical and ordinal indicators, a three-level discrete scoring scheme is used in Equation (5), where ($SimX_{ij}$) represents the similarity score of the (j)-th indicator for the (i)-th historical case, (X_{oj}) denotes the (j)-th indicator of the target case, and (X_{ij}) refers to the (j)-th indicator of the (i)-th historical case.

$$SimX_{ij} = \begin{cases} 100, & X_{oj} - X_{ij} = 0 \\ 50, & |X_{oj} - X_{ij}| = 1 \ (0 \le i \le n,\ 1 \le j \le m) \\ 0, & otherwise \end{cases} \tag{5}$$

For continuous variables with exact numerical outcomes, the scoring method is formulated in Equations (6) and (7).

$$D_{Va} = |X_{oj} - X_{ij}| \tag{6}$$

$$SimX_{ij} = (1 - \frac{D_{Va} - min(D_{Va})}{max(D_{Va}) - min(D_{Va})}) \times 100 \tag{7}$$

3.5.2 Global similarity

Global similarity is derived by aggregating local similarities through entropy-derived

weighting. The formulation is given in Equation (8), where ($Sim(X_o, X_i)$) denotes the similarity between the target case and the (i)-th historical case, and (ω_{ij}) represents the weight coefficient of the (j)-th indicator in the (i)-th historical case.

$$Sim(X_o, X_i) = \sum_{i=1}^{n} SimX_{ij}\omega_{ij} \tag{8}$$

This weighted sum ensures that indicators with higher information contribution dominate the overall similarity, reducing subjectivity and avoiding equal-weight assumptions.

3.5.3 Role in system-level reasoning

By integrating hybrid local similarity measures with entropy-weighted global aggregation, the proposed scheme enables consistent and meaningful comparison of cases characterized by mixed attribute types. The approach prioritizes indicators that are most relevant to fire outcomes and facilitates transparent retrieval of analogous fire scenarios. In this manner, similarity-based reasoning is aligned with both the empirical structure of historical accident data and the systemic characteristics of greenhouse fire risk.

3.6 Validation design

Conventional statistical validation is constrained in greenhouse fire risk assessment due to limited accident data and pronounced scenario heterogeneity. Accordingly, a reverse validation strategy, commonly adopted in CBR-based safety research, is used to evaluate the reliability and validity of similarity-based case retrieval.

3.6.1 Reliability assessment

Reliability refers to the stability and reproducibility of similarity computation. Two

representative greenhouse fire cases from distinct regions are selected as target cases. For each target case, the CBR framework retrieves highly similar historical cases from the remaining database. The degree of overlap among top-ranked retrievals and the consistency of their key attributes are examined. A high proportion of overlapping cases across independent retrievals indicates that similarity measurement remains stable and is not unduly sensitive to minor variations in input data. The repeated identification of structurally similar cases across different target scenarios further demonstrates that similarity rankings are reproducible rather than arbitrary.

3.6.2 Validity assessment

Validity concerns whether the retrieved cases are meaningful and consistent with domain knowledge. The characteristics of top-ranked similar cases—particularly ignition sources, facility attributes, availability of firefighting equipment, and accident severity—are compared with those of the target cases. Coherent and interpretable similarities, especially in outcome-related and system-critical indicators, support the validity of the framework for decision-making applications. Consistency with established risk mechanisms reported in prior studies further indicates that the framework captures critical features of fire accidents rather than superficial associations.

3.6.3 Implications for decision support and adaptive learning

The reverse validation results indicate that the proposed CBR framework consistently identifies recurrent fire accident patterns across heterogeneous greenhouse facilities. The demonstration of stable retrieval performance and meaningful case similarity supports

experience-informed decision-making. Each validated retrieval contributes to enriching the case base, enabling incremental improvement in reasoning capability as additional accident cases are incorporated. Overall, the validation outcomes establish the proposed approach as a robust and interpretable tool for accident similarity analysis, supporting loss prevention under conditions of limited and heterogeneous data.

4. Case analysis

Based on the established CBR framework, similarity between the target case and all candidate cases was computed. This section presents the similarity ranking results, reverse validation, and indicator contribution analysis.

4.1 Similarity results

The target case is a greenhouse fire that occurred on 23 February 2016, in Xitian Gezhuang Village, Shunyi District, Beijing. This case is selected for its representativeness and the severity of the associated losses. Using entropy-weighted global similarity, all candidate cases are ranked accordingly. Following Bai et al. [19], cases with a similarity score exceeding 60 are considered sufficiently similar for CBR analysis.

Of the 101 validated cases, 53 meet this threshold. Fig. 4 illustrates the top 30 similarity rankings for the Beijing (Shunyi) target case, with scores ranging from 0 to 100 and a threshold line at 60.

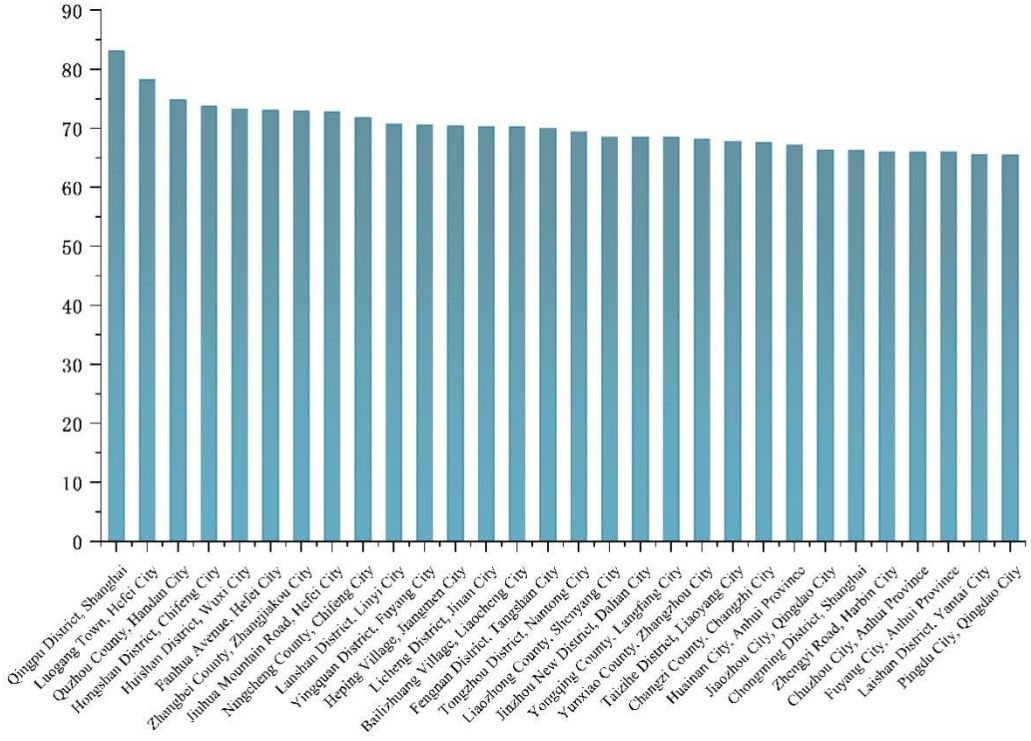

Fig. 4. Top-30 similarity ranking for the Beijing (Shunyi) target case

Table 2 presents the metadata for the top 10 most similar cases. The highest-ranked cases originate from Qingpu District (Shanghai), Luogang Town (Hefei), Quzhou County (Handan), Hongshan District (Chifeng), and Huishan District (Wuxi). The top three cases achieve similarity scores above 75, indicating strong structural comparability. These cases share two dominant characteristics: the use of plastic greenhouses for crop cultivation and the absence of firefighting equipment. Despite these similarities, differences in greenhouse structural types and casualty outcomes are observed, indicating variability in fire dynamics under comparable vulnerability conditions. These findings suggest that highly similar cases reflect consistent hazard patterns, indicating that greenhouse fires across different regions are shaped by recurrent risk profiles rather than isolated, context-specific conditions.

Table 2

Detailed metadata for the top-10 similar cases retrieved for the Beijing (Shunyi) target case.

Rank	Location (City/County, Province)	Date	Greenhouse Use	Materials	Firefighting Equipment (Y/N)	Ignition Source	Similarity Score
1	Qingpu District, Shanghai	2019-01-14	Crop cultivation	Plastic	N	External ignition source	83.06
2	Luogang Town, Hefei City	2005-07-14	Crop cultivation	Plastic	N	External ignition source	78.25
3	Quzhou County, Handan City	2016-03-03	Crop cultivation	Plastic	N	External ignition source	74.79
4	Hongshan District, Chifeng City	2023-06-06	Crop cultivation	Plastic	N	External ignition source	73.73
5	Huishan District, Wuxi City	2015-12-18	Crop cultivation	Plastic	N	External ignition source	73.25
6	Fanhua Avenue, Hefei City	2022-03-24	Crop cultivation	Plastic	N	External ignition source	73.04
7	Zhangbei County, Zhangjiakou City	2019-02-04	Crop cultivation	Plastic	N	External ignition source	72.86
8	Jiuhua Mountain Road, Hefei City	2015-06-10	Commerce and Trade	Plastic	N	Improper storage of flammable materials	72.76
9	Ningcheng County, Chifeng City	2017-01-27	Crop cultivation	Plastic	N	External ignition source	71.79
10	Lanshan District, Linyi City	2017-02-11	Crop cultivation	Plastic	N	External ignition source	70.66

4.2 Reverse testing and stability verification

To evaluate retrieval robustness, a second target case—a greenhouse fire on Huafang Avenue, Hefei City, Anhui Province—is used for reverse testing. The CBR framework again retrieves cases with high similarity scores. Fig. 5 presents the top 15 similarity rankings for the Hefei target case.

Table 3 provides metadata for the top 10 most similar cases associated with this second target. Across both validation exercises, the highest-ranked cases exhibit strong consistency in key attributes, including the absence of firefighting equipment, similar ignition sources, and comparable structural and functional characteristics. Several cases, such as the Quzhou County

(Handan) incident, appear in both retrieval sets, further supporting the stability of similarity evaluation. Cases from geographically dispersed regions, including Guangdong, Hebei, and Inner Mongolia, demonstrate similar accident characteristics despite regional variation.

Table 3

Detailed metadata for similar cases retrieved for the Hefei target case.

Rank	Location (City/County, Province)	Date	Greenhouse Use	Materials	Firefighting Equipment (Y/N)	Ignition Source	Similarity Score
1	Jiuhua Mountain Road, Hefei	2015-06-10	Commerce and Trade	Plastic	N	Improper storage of flammables	94.72
2	Heping Village, Jiangmen	2020-06-09	Livestock and poultry farming	Sheet iron	N	Improper storage of flammables	90.98
3	Luogang Town, Hefei	2005-07-14	Crop cultivation	Plastic	N	External ignition source	90.59
4	Fengnan District, Tangshan	2019-03-26	Crop cultivation	Plastic	N	Improper storage of flammables	90.52
5	Quzhou County, Handan	2016-03-03	Crop cultivation	Plastic	N	External ignition source	85.70
6	Caolaoji Town, Bengbu	2021-03-16	Livestock and poultry farming	Plastic	N	Electrical equipment failure	83.77
7	Zhang 'er Village, Jinan	2017-03-14	Crop cultivation	Plastic	N	Improper storage of flammables	82.74
8	Hongshan District, Chifeng	2023-06-06	Crop cultivation	Plastic	N	External ignition source	81.29
9	Huishan District, Wuxi	2015-12-18	Crop cultivation	Plastic	N	External ignition source	81.24
10	Mentougou, Beijing	2019-03-06	Crop cultivation	Plastic	N	Improper storage of flammables	81.21

Table 4 summarizes the overlap between the Beijing and Hefei retrieval results, highlighting shared cases and common attributes. The high degree of overlap across independent retrieval scenarios indicates that the CBR framework consistently identifies recurrent fire-causing patterns, demonstrating strong reliability and reproducibility.

Table 4

Overlap between Beijing and Hefei retrievals, highlighting shared cases and shared attributes.

Case ID/Location	Appears in Set A (Y/N)	Appears in Set B (Y/N)	Shared Attributes (Ignition, Equipment, Use)
Jiuhua Mountain Road, Hefei	N	Y	Improper storage of flammable materials

Case ID/Location	Appears in Set A (Y/N)	Appears in Set B (Y/N)	Shared Attributes (Ignition, Equipment, Use)
Heping Village, Jiangmen	N	Y	Improper storage of flammable materials
Luogang Town, Hefei	N	Y	External ignition source
Fengnan District, Tangshan	N	Y	Improper storage of flammable materials
Quzhou County, Handan	N	Y	External ignition source
Caolaoji Town, Bengbu	N	Y	Electrical equipment failure
Zhang 'er Village, Jinan	N	Y	Improper storage of flammable materials
Hongshan District, Chifeng	N	Y	External ignition source
Huishan District, Wuxi	N	Y	External ignition source
Mentougou, Beijing	N	Y	Improper storage of flammable materials
Suixian County, Suizhou	N	Y	Improper storage of flammable materials
Huainan City, Anhui Province	Y	Y	Human factors
Zhengyi Road, Harbin	N	Y	External ignition source
Yongqing County, Langfang	N	Y	External ignition source
Chongming District, Shanghai	Y	Y	External ignition source

4.3 Indicator contribution analysis

To explain the ranking of highly similar cases, the contributions of the three indicator dimensions are analyzed for the top five most similar cases relative to the Beijing target case. Table 5 reports the indicator weights, mean local similarity values, and weighted contributions averaged across these cases, while Table 6 presents normalized similarity contributions by dimension.

Table 5

Per-indicator weights, mean local similarity, and weighted contributions averaged across the top-five cases.

Code	Indicator Name	Weight (%)	Mean Local Similarity (0–100)
X7	On-site personnel	20.71	18.64
X13	The severity of the accident	20.08	20.08
X10	Shape type	12.56	8.79
X4	Fire-fighting equipment	10.59	0.31
X11	The number of deaths	10.48	9.78

Table 6

Normalized contributions by dimension for the top-five similar cases.

Analyzed Case	Environmental (%)	Facility (%)	Accident (%)
Caijia Road, Qingpu District, Shanghai	36.91	25.71	37.38
Luogang Town, Hefei City, Anhui Province	29.46	30.85	39.69

Analyzed Case	Environmental (%)	Facility (%)	Accident (%)
Quzhou County, Handan City, Hebei Province	33.72	23.88	42.40
Hongshan District, Chifeng City, Inner Mongolia	33.66	24.22	42.12
Huishan District, Wuxi City, Jiangsu Province	45.36	32.95	21.69

The similarity decomposition reveals three consistent patterns governing case retrieval.

First, incident-related indicators dominate local similarity across all top-ranked cases, consistent with entropy-derived weights. In particular, ignition source type, on-site personnel presence (X7), and accident severity (X13) have the highest contribution coefficients.

Second, facility characteristics exert a moderate yet discriminative influence on similarity outcomes. Variations in cladding materials, structural configuration (X10), and the presence of firefighting equipment (X4) account for differences in case similarity scores. For example, the Wuxi case demonstrates substantially greater facility similarity than the Shanghai Qingpu case, explaining discrepancies among otherwise comparable profiles.

Third, environmental characteristics contribute to regional differentiation. The Wuxi case has high environmental and facility similarity but relatively lower incident similarity, which may be associated with more standardized greenhouse construction and improved safety practices that mitigate fire severity. In contrast, cases from less developed regions, such as Chifeng (incident similarity of 42.12%), exhibit higher incident similarity, likely reflecting recurring deficiencies in electrical maintenance, fire safety management, and combustible material control.

5. Discussion

5.1 Primacy of accident characteristics

Electrical faults, identified as the most prevalent ignition source, often originate in concealed locations, delaying detection until fire escalates [4, 31]. The presence of on-site personnel has a dual effect: absence is associated with increased severity, whereas the presence of untrained individuals does not necessarily improve outcomes. Accident severity receives the highest weight, indicating that the CBR framework prioritizes outcome relevance in similarity evaluation [33].

The observed dominance of incident-related characteristics differs from findings in conventional building fire studies, where structural factors often have comparable influence [26, 34]. Two primary explanations are identified. First, greenhouse facilities have relatively homogeneous structural features, including plastic envelopes, lightweight frames, and similar combustible loads, which reduces the discriminatory capacity of facility-related indicators; accordingly, greenhouse materials (X3) receive the lowest weight. Second, regulatory frameworks governing greenhouse safety are less developed than those for conventional buildings, resulting in greater variability in accident-related variables. The high weights assigned to on-site personnel presence and accident severity reflect gaps in regulatory oversight, informal operational practices, the absence of standardized safeguards, and limited supervisory control, in which human and organizational factors compensate for insufficient engineering controls [13].

Regional variation is also observed, with cases from less developed regions showing greater

similarity in accident characteristics than those from more developed areas. These differences are associated with disparities in electrical maintenance practices, safety oversight, and operator awareness. Such variation emerges through correlations with incident characteristics rather than direct geographic weighting (X1), thereby capturing regional risk profiles without introducing geographic determinism [27].

5.2 Implications for accident prevention and loss mitigation

The proposed framework integrates entropy-based weighting, mixed-type similarity measurement, and combined global–local matching strategies. In contrast to traditional approaches, it accommodates complex and interdependent greenhouse fire scenarios by leveraging historical accident cases rather than relying on predefined causal assumptions [14]. By extending CBR to partially documented agricultural contexts, the study demonstrates reduced subjectivity in indicator weighting, improved handling of continuous and categorical variables, and robust performance under reverse validation.

These characteristics position the CBR framework as a scalable and adaptive tool for supporting intelligent early warning in facility agriculture. The study contributes to fire safety research by illustrating how experience-based reasoning can enhance accident learning and decision support in hazard-prone systems where data scarcity and heterogeneity constrain conventional risk modeling approaches.

5.3 Policy recommendations

Three key operational and regulatory implications for accident prevention and loss

mitigation are identified. First, a substantial proportion of high-similarity cases lack basic firefighting infrastructure, indicating the need for mandatory installation of extinguishers, alarm systems, and electrical cutoff devices within greenhouse facilities. Targeted incentives in resource-constrained regions may help meet these requirements. Second, cases without on-site personnel consistently have higher accident severity, suggesting the importance of minimum staffing requirements or the implementation of remote monitoring systems. Third, the widespread use of combustible plastic coverings highlights the need to promote fire-retardant materials and to strengthen electrical design and installation standards.

These recommendations, grounded in quantified patterns derived from historical cases, provide actionable guidance for reducing loss severity and informing inspection practices and regulatory revisions.

5.4 Limitations and future research

Several limitations are acknowledged. The case database, derived from publicly reported incidents, may exclude unreported events and is therefore subject to potential selection bias. Although the dataset encompasses multiple provinces, its concentration within China limits the generalizability of findings. The selected set of 13 indicators captures key influencing factors but does not incorporate detailed technical parameters such as electrical load characteristics or material ignition thresholds.

In addition, the CBR approach is inherently retrospective, relying on historical cases to inform analysis. The integration of Internet of Things (IoT)-enabled real-time data could enhance the ability to identify dynamic risks and issue early warnings. Future research may

expand the dataset to include international cases, incorporate sensor-based monitoring data, explore hybrid CBR–machine learning approaches, and develop automated early warning systems tailored to facility agriculture.

6. Conclusions

This study develops a CBR framework for greenhouse fire analysis based on 101 validated cases, 13 indicators, entropy-based weighting, and hybrid similarity computation. The results indicate that incident-related characteristics dominate similarity evaluation, while environmental and facility indicators reveal regional variation in fire risk patterns. Reverse validation confirms the model's stability and reproducibility.

From a practical perspective, the framework enables rapid retrieval of analogous cases, supporting data-informed decision-making for accident prevention and loss mitigation. The findings highlight the importance of measures such as providing firefighting equipment, strengthening personnel oversight, and adopting improved materials and electrical safety practices.

This study contributes to the field of loss prevention by demonstrating how experience-based reasoning can facilitate systematic learning from fire accidents in process-like facilities under conditions of limited and heterogeneous data.

References

1. Benke K, Tomkins B (2017) Future food production systems: Vertical farming and controlled environment agriculture. Sustain Sci Pract Policy 13(1):13–26

2. Kozai T (2018) Plant factories with artificial lighting (PFALs): Benefits, problems, and challenges. In: Kozai T (ed) Smart plant factory: The next generation indoor vertical farms. Springer, pp 15–29

3. Goucher L, Bruce R, Cameron D et al (2017) The environmental impact of fertilizer embodied in a wheat to bread supply chain. Nat Plants 3:17012. https://doi.org/10.1038/nplants.2017.12

4. Zhou CJ (2019) Dr. Zhou's investigation notes (Part 91): Reflections on the March 26 solar greenhouse vegetable base fire in Tangshan, Hebei Province. Agric Eng Technol 39(10):46–56

5. Sun P, Rodriguez A, Kim WI, Huang X, Fernandez Pello C (2021) Effect of external and internal heating on the flame spread and phase change of thin polyethylene tubes. Int J Therm Sci 168:107054

6. Babrauskas V (2016) Ignition handbook. Fire Science Publishers

7. Dahanayake KC, Chow CL (2018) Moisture content, ignitability, and fire risk of vegetation in vertical greenery systems. Fire Ecol 14:125–142

8. Twigg J, Christie N, Haworth J, Osuteye E, Skarlatidou A (2017) Improved methods for fire risk assessment in low income and informal settlements. Int J Environ Res Public Health 14(2):139

9. Kim D, Lee M, Koo K, Lee S (2006) Forest fire risk assessment through analyzing ignition characteristics of forest fuel bed. For Ecol Manage 234

10. Kobes M, Helsloot I, de Vries B, Post JG (2010) Building safety and human behaviour in fire: A literature review. Fire Saf J 45(1):1–11

11. Hollnagel E (2018) Functional resonance analysis method and manual (Version 2)

12. Paik JK (2020) Quantitative fire risk assessment and management. In: Advanced Structural Safety Studies, vol. 37. Springer, Singapore

13. Zhou L, Wu X, Xu Z, Fujita H (2018) Emergency decision making for natural disasters: An overview. Int J Disaster Risk Reduct 27:567–576

14. Aamodt A, Plaza E (1994) Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI Commun 7(1):39–59

15. Qin SS, Lü W, Yu HQ, Song LL, Zhou T (2018) Research on similarity and countermeasures of fire accidents based on case-based reasoning. Saf Environ Eng 25(5):150–155

16. Chen Y, Zhang L, Hu J, Liu Z, Xu K (2022) Emergency response recommendation for long-distance oil and gas pipeline based on an accident case representation model. J Loss Prev Process Ind 77:104779. https://doi.org/10.1016/j.jlp.2022.104779

17. Lu Y, Yin L, Deng Y, Wu G, Li C (2023) Using case based reasoning for automated safety risk management in construction industry. Saf Sci 163:106113

18. Pan MZ, Guo CG, Yuan YF (2020) Research on emergency decision-making for fire accidents in renovation projects of old industrial buildings. Fire Sci Technol 39(6):840–844

19. Bai J, Li YK, Lu YJ, Le Y (2016) An approach to forming the top management team of mega infrastructure projects based on improved case based reasoning. J Syst Manage 25(2):272–281

20. Schultheis A, Hoffmann M, Malburg L, Bergmann R (2023) Explanation of similarities in process-oriented case-based reasoning by visualization. In: Case-Based Reasoning Research and Development: 31st International Conference, ICCBR 2023, Aberdeen, Scotland, July 17–20, 2023. Springer, Lect Notes Comput Sci 14141, pp 53–68

21. Huang XJ, Wang Y, Xie SB, Zhang L (2021) Development and review of building fire risk assessment methods. Fire Sci Technol 40(8):1174–1177

22. Lei Y, Shen Z, Tian F, Yang X, Wang F, Pan R et al (2023) Fire risk level prediction of timber heritage buildings based on entropy and XGBoost. J Cult Herit 63:11–22

23. Vidal LA, Marle F, Bocquet JC (2011) Measuring project complexity using the analytic hierarchy process. Int J Proj Manage 29(6):718–727

24. Liang Z, Yang Y, Wang Y, Zhang M, Zhuang Y (2025) Multidimensional quantitative modeling fusion analysis of safety risks in hydrogen refueling stations: A case study of a station in Beijing. J Loss Prev Process Ind, 105865. https://doi.org/10.1016/j.jlp.2025.105865

25. He J, Zhao P, Ling Y, Cai Y, Gao H, Luo X (2025) Comprehensive performance evaluation framework for deep learning models in Natech events extraction and classification. J Loss Prev Process Ind, 105831. https://doi.org/10.1016/j.jlp.2025.105831

26. Lei P, Chen C, Zhao D, Zhang Y, Xu T, Jiao W (2022) Study on heat allocation and temperature profile in a T shaped branched tunnel fire with different branch slopes under natural ventilation. Tunn Undergr Space Technol 126:104508

27. Zhang A, Wang Y, Zou Y (2023) Positive feedback to regional climate enhances African wildfires. iScience 26(12)

28. Caliendo C, Genovese G, Russo I (2020) Risk analysis of road tunnels: A computational fluid dynamic model for assessing the effects of natural ventilation. Appl Sci 11(1):32

29. Resource Innovation Institute (2025) Fire safety in high-tech production greenhouses

30. Cawson JG, Collins L, Parks SA, Nolan RH, Penman TD (2024) Atmospheric dryness removes barriers to the development of large forest fires. Agric For Meteorol 350:109990

31. Plato J (2019) Fighting fire hazards. Greenhouse Canada, November 2019

32. Li L, Ni B, Zhang S, Qiang Y, Zhang Z, Zhou L et al (2022) Tunnel collapse risk assessment based on improved quantitative theory III and EW-AHP coupling weight. Sci Rep 12(1):16054

33. Yu L, Gabbay D (2022) Case based reasoning via comparing the strength order of features. In: Explainable and Transparent AI and Multi Agent Systems: 4th International Workshop (EXTRAAMAS 2022), Revised Selected Papers

34. Li G, Yuan H, Shan Y, Lin G, Xie G, Giordano A (2023) Architectural cultural heritage conservation: fire risk assessment of ancient vernacular residences based on FAHP and EWM. Appl Sci 13(22):12368