International Journal of Disaster Risk Reduction 132 (2026) 105967

Contents lists available at ScienceDirect

International Journal of Disaster Risk Reduction

journal homepage: www.elsevier.com/locate/ijdrr

A historical analysis of factors driving the daily prioritization of wildland fires in California

Dung Nguyen [a,*], Erin J. Belval [b], Yu Wei [a], David E. Calkin [c]

[a] Department of Forest and Rangeland Stewardship, Colorado State University, Fort Collins, CO, 80523, USA
[b] USDA Forest Service, Rocky Mountain Research Station, Fort Collins, CO, 80526, USA
[c] USDA Forest Service, Rocky Mountain Research Station, Missoula, MT, 59801, USA

ARTICLE INFO

Keywords:
Wildfire management
Incident prioritization matrix
Wildfire ranking
Machine learning
Random forest

ABSTRACT

During periods of heightened wildland fire activity in the United States, multiagency coordinating groups must prioritize among multiple on-going fires to allocate scarce suppression resources. While many studies have explored factors that influence wildfire suppression expenditures and personnel allocation, understanding the specific factors that affect daily wildfire prioritization has remained unexplored. In this study, we first examine wildfire reporting and ranking processes across different regions of the United States to provide insight into criteria used for fire ranking. We then focus on examining the 12 criteria used for ranking fires daily by California's multiagency coordination group. We developed a computer program to replicate the California prioritization process and found that fire rankings generated by this program align well with the historical rankings, indicating close adherence of California's fire managers to their ranking rules. A correlation analysis revealed weak correlations among the 12 criteria, suggesting that no criterion should serve as a proxy for another during fire priority evaluations. We further applied a Random Forest machine learning model, which identified threats and damage to structures, fire size, and evacuations as the most impactful criteria in determining fire priority. Our findings can benefit wildfire decision makers by providing clear insights into the existing wildfire priority assessment process, so that adjustments to the process can be made for better management outcomes. Policymakers can also leverage these insights to develop evidence-based fire management policies, regulations, and practices that promote more efficient responses to fire risks while fostering greater public trust in fire management efforts.

1. Introduction

Fire suppression plays a crucial role in managing wildland fires across the United States (US). According to reports from the US Forest Service, between 97 % and 99 % of all US wildfires that occurred in the last three decades have been suppressed or effectively controlled by the initial suppression response [1,2]. When unwanted wildfires are not contained during the initial response phase (the phase of fire management occurring during the first operational period after the fire is discovered [3,4]), they often become large wildfires that cause the majority of the wildfire damage in the country [5,6] and require substantial amount of resources to manage (e.g., firefighters and equipment such as aircraft, helicopters, bulldozers, or engines). Fires allocated more resources (i.e., larger number

* Corresponding author.
E-mail address: dung.nguyen@colostate.edu (D. Nguyen).

https://doi.org/10.1016/j.ijdrr.2025.105967
Received 23 May 2025; Received in revised form 23 November 2025; Accepted 13 December 2025
Available online 15 December 2025
2212-4209/

of resources per hectare of fire) are more likely to be contained [7]. However, when multiple fires are managed simultaneously, particularly if they are in relatively close spatial proximity, the limited suppression resources that are available must be shared among the fires. Prioritizing resource allocation among those fires is often a challenge requiring consideration of multiple factors such as the ecological and socio-economic impacts of fires [8], uncertainty around future fire behavior, and the completeness and accuracy of fire situation data [9]. As the effects of climate change become more pronounced, both the number of days with simultaneous fires and the number of fires burning on those days are expected to increase [10,11], increasing the necessity of daily resources prioritization. This is especially important when there are often multiple fires on a day that may simultaneously demand resources.

In the US, the management of wildland fire occurs in a multiagency environment, with multiple federal and state agencies having responsibility for wildfire response for lands under their jurisdiction. The allocation of fire resources is managed by the ten Geographic Area Coordination Centers (GACCs) in the US (Fig. 1). Coordination between agencies occurs within each GACC and is facilitated by a Multi-Agency Coordination (MAC) Group. This coordination allows agencies to share resources with each other. While the sharing of fire suppression resources is efficient, allowing one agencies' resource to act as surge resource for another agency, it also requires the agencies to jointly prioritize the use of those resources when many fires are competing. Because the agencies have different missions and thus different priorities, a process is required that allows agencies to quickly and efficiently reach agreement on where to send scarce suppression resources.

To facilitate the daily prioritization of wildfires across multiple agencies, a GACC's MAC Group often uses a pre-identified multi criteria decision analysis process that involves quantitively scoring ongoing fires according to multiple criteria. In this process, each fire is scored based on multiple criteria weighted by their importance, and then ranked by the total weighted score. The two GACCs that comprise the state of California (the Northern and Southern California GACCs), for example, use an "incident prioritization matrix" [12] to evaluate each fire daily based on four broad categories, including life and safety threats, property damage or risk, resource impact or loss, and incident complexity and duration. Each category contains a set of criteria with associated guidelines to rate each fire. For example, "evacuations" is one criterion of the "life and safety threats" category, and a fire leading to either potential, pre-cautionary, or in-progress evacuations would respectively receive 1–2, 3–4, or 5 rating points according to the California prioritization matrix. The importance of each fire is measured by the sum of its rating points across all criteria. While the points-based prioritization process helps guide decision makers of the MAC Group, the decision makers are also given the flexibility to deviate from the initial rankings due to the highly complex and dynamic nature of wildland fire incidents. Within a GACC, a fire with the highest daily rank based on points may often, but not always, receive the top priority for suppression resource allocation. For example, fires that are emerging threats, but not yet high priority, where additional resources can provide a potentially high marginal return may be prioritized over the others. Previous research has also found that regardless of ranking, fires with greater public attention, media coverage, and political pressure may receive higher levels of suppression resources [13].

The multi-criteria decision process for ranking wildland fires in California was established at least a decade ago, as documented in the California Statewide Multi-Agency Coordination System Guide dating back to 2013 [12], if not earlier. As fire activity, behavior, and outcomes change through time, reassessing the criteria for wildfire prioritization is critical. Because the agency representatives on the MAC Groups are experts who evaluate fire importance in real time with access to additional information and context, examining

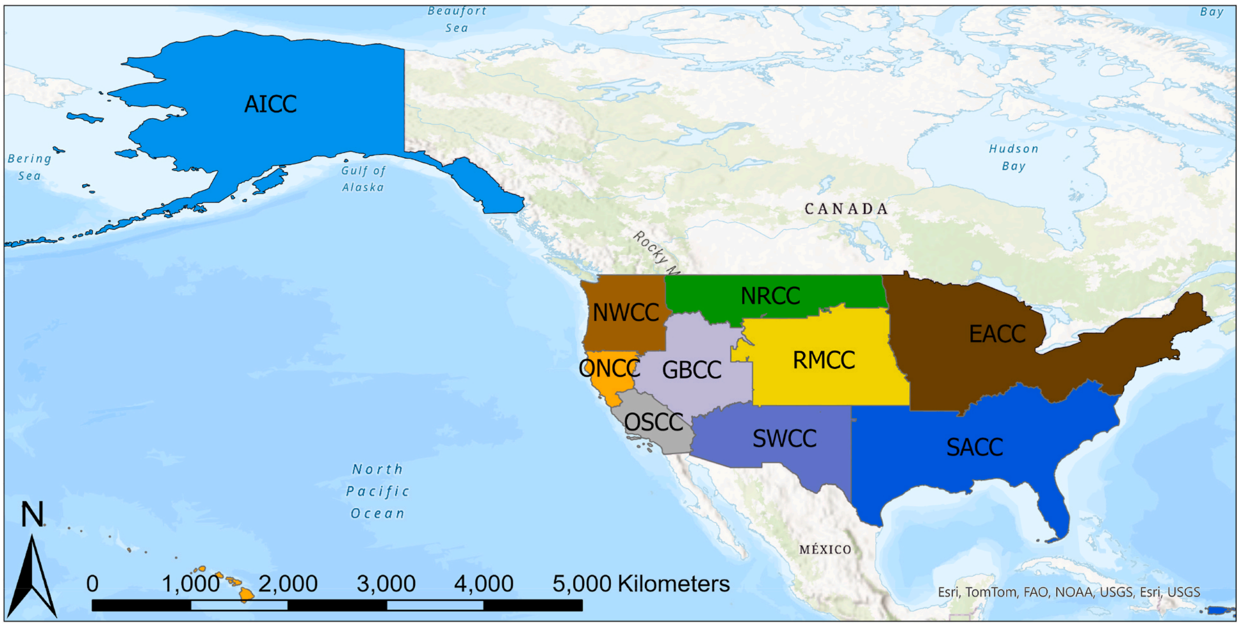

Fig. 1. Locations and spatial boundaries of the ten Geographic Area Coordination Centers across the United States. Full GACC names and their corresponding abbreviations are available at https://gacc.nifc.gov/.

their fidelity to the established system for ranking allows us to examine if the current set of criteria (also called factors in this paper) are complete and sufficient for ranking fires. Decision makers should regularly re-evaluate existing ranking factors to ensure that they continue to be relevant and can sufficiently support the fire priority evaluation process. Understanding the influences of these factors on fire priority ranking can be insightful for future fire management decision making. However, this is a gap in the literature as we found no study that has addressed this research topic.

Past research has focused on explaining the factors associated with suppression outcomes for individual fires, including suppression expenditures and other suppression resources usage. Many studies have found that suppression expenditures are strongly influenced by fire size [14–17] and fire intensity [15,17]. Other factors also drive resource allocation including weather [14], topographic and landscape conditions such as forest cover and slope steepness [17,18], homes of proximity to fire risk [15,19], private land ownership [16,17], and sociopolitical factors such as media coverage and political pressure [13,20]. Studies have also found that personnel, equipment, and other resources allocated to fires may be driven by factors such as preparedness level (https://www.nifc.gov/fire-information), presence of homes, number of fire days and fire growth potential [21–23].

Factors influencing suppression expenditures and resources usage may not directly imply wildfire priority ranking. This is because a clear relationship between fire priority and final funding and resource levels may not exist. For example, a long-lasting fire may hold a more substantial amount of resources in total, but it may not be prioritized more on specific days when the fire has approached its natural spread limit. Wildfire prioritization is a process repeated daily for on-going fires as factors influencing the prioritization of those fires change over time. Many prior studies used final wildfire outcomes (suppression expenditure or resources usage) that are jointly measured and reported in total (across multiple days instead of daily) for their analyses, which make it difficult to identify the causal mechanism behind influencing factors. Several past studies, including Bayham and Yoder [22] and Plantinga et al. [24], have addressed this issue by tracking outcomes from individual fires as they progress over time, such as the loss of homes and damage to watersheds, endangered species, or recreation sites. However, like many other studies, they were designed to understand factors that impact wildfire's final outcomes rather than examining the impact of these factors on daily wildfire ranking decisions directly, which is more important in determining allocations of scarce resources.

In the US, researchers have lacked access to historical wildfire ranking data. This likely explains the absence of studies examining wildfire prioritization processes in existing literature. To our knowledge, the Fire Situation Reports dataset (SIT-209, available at https://www.wildfire.gov/application/sit209) includes a field tracking fire priority, but the data are often incomplete with missing entries for many days within a fire's duration, making it unsuitable for analyzing daily fire prioritization. In 2024, a comprehensive dataset of the Incident Management Situation Reports (IMSR) was published [25], which provides historical daily wildfire rankings across the US that facilitate our study. We selected California, a region with 105 million acres of land and a warm Mediterranean climate [26,27], for our study as it has been a focal point for wildfire management in the US. Wildfire management in California has been a national top priority for decades, with pressing challenges due to climate change [28], excessive fuel buildup [29], and rapid population growth in the wildland urban interface (WUI) areas [30,31]. Over the past four decades, declining summer precipitation has led to increased fuel aridity and drought in California [32,33], resulting in longer fire seasons [34] and more intense fire activity in the state [35–39]. Additionally, accelerated WUI development in California has also led to more fire risk to both residents and residential structures [30,31]. Thus, daily wildfire ranking has become critical in supporting resource allocation decisions in California's two GACCs, especially during the fire season when multiple large fires may occur daily.

This study examines current wildfire ranking systems across US GACCs, explores how the 12 criteria officially documented in California's prioritization process explain the state's historical fire rankings, and identifies the most influential criteria that determine fire prioritization on days with multiple simultaneous fires. Our paper is organized as follows. We begin with introducing the daily wildfire reporting process and prioritization schemes across GACCs in the US. We then describe the development of a rule-based program that simulates California's process for ranking simultaneous wildfires daily, from high to low priorities, in each of the state's two GACCs. The rules were designed to align with the MAC Group's scoring matrix employed by both GACCs. More specifically, the matrix assigns points to individual fires based upon information reported daily in the Fire Situation Reports (SIT-209). Using an automated method to assign points to each fire by parsing several fields from the SIT-209s, we generated daily fire rankings and compared them to the historical fire rankings as reported by the IMSR. This comparison allowed us to assess whether the pre-determined multi-criteria decision analysis matrix resulted in rankings that were consistent with the GACCs' fire managers' decision during actual fire prioritization. Additionally, we conducted a correlation analysis to examine the relationships among ranking criteria, and developed a Random Forest [40] model to identify the importance of these criteria (or factors) within California's scoring matrix. Unveiling the key factors behind wildfire prioritization and their interrelationships offers insights into the current fire ranking assessment process, enabling opportunities to refine future wildfire response protocols. Fire management agencies can use these findings to re-evaluate the current wildfire prioritization process, and target their focus on the most influential factors to tailor their response strategies for optimal allocations of suppression resources. Furthermore, policymakers can leverage these insights to formulate evidence-based fire management policies, regulations, and practices, thus mitigating fire risks and bolstering community resilience in fire prone areas.

2. Wildfire reporting process and ranking schemes across the US

In the US, large wildfire incidents are officially reported using the Incident Status Summary form, i.e. the ICS-209. A copy of this form along with its purpose and detailed description of reported data is available at https://www.nwcg.gov/ics-forms. Generally, every wildfire incident exceeding specific sizes (i.e. 100 acres in timber or 300 acres in grass) or satisfying complexity criteria (i.e. having a Type 1 or 2 incident management team assignment) needs to file a ICS-209 report (https://www.nifc.gov/sites/default/files/

document-media/2023_ICS-209_User_Guide.pdf). ICS-209 report is typically completed daily by fire mangers. While some ICS-209 data elements (referred to as “blocks” in the form) are required, many other elements are optional. The archived version of ICS-209 is called SIT-209.

Each GACC in the US employs a region-specific prioritization matrix with criteria for ranking the most relevant wildfires within its geographical area. Across GACCs, these criteria can be grouped into five broad categories, 1) safety threats to firefighters and the public, 2) damage and threats to properties, 3) potential threats and losses to infrastructure and historical, natural, or commercial resources, 4) likelihood that resources will allow managers to meet the incident management objectives (such as time and firefighting resources for containment), and 5) other social and political considerations (Fig. 2). For each GACC, the prioritization matrix is often accompanied by a standardized process to guide the assignment of points to each fire incident using information from the ICS-209 (or its archived version, SIT-209). Of the eight GACCs that have published their prioritization matrices online (Fig. 2), we found five with detailed processes for linking matrix criteria to ICS-209 information, including California (with two GACCs covering the north and south portions of the state), Northern Rockies, the Great Basin, and the Northwest.

While the criteria used to rank fires is thematically similar across different GACC matrices, the specific use of ICS-209 elements varies by region. For instance, the process outlined for the California GACCs recommends using information from only one or two ICS-209 blocks for each ranking criterion, while the Great Basin’s matrix requires between two and 11 different blocks for each criterion (Fig. 2). Certain blocks, such as the significant fire event summary (ICS-209 Block 28) and fire threats in the next 12 to 72 h (ICS-209 Block 38), are more frequently used to support fire ranking across multiple GACCs (Table 1). Using a greater number ICS-209 blocks allows for a more comprehensive ranking assessment, but also requires more time with a potentially higher level of subjectivity in both data processing and analysis.

3. California case study: material and methods

We collected and used information from both IMSR [25] and SIT-209 for all fires within the state of California occurring between 2016 and 2020. We obtained the daily rankings of historical fires from IMSR and extracted a comprehensive set of daily information associated with each fire throughout its multiple-day duration from SIT-209, including weather, fire behavior, suppression resources, among the others. We then used Structured Query Language (SQL) queries to join daily wildfire records that exist in both IMSR and SIT-209, as illustrated by Appendix Fig. A1. A more detailed process and additional results from this work were published in Ref. [25]. In California, we found 5132 daily fire records that were reported by IMSR, but only 4671 of them (91 %) have corresponding SIT-209 entries (Table 2). Unsuccessful connections between fire observations in the IMSR and SIT-209 datasets include 461 records (9 %); these failures were primarily due to 1) incident names found in IMSR but not in SIT-209, and 2) discrepancies in resource assignment information, such as when a fire may appear in both datasets with the same name, but differ in the recorded number of crews, engines, and helicopters allocated to it. Since there was no absolutely clean approach to connect the mismatched cases, we decided to exclude those 461 IMSR fire records from our analyses.

Both GACCs in California use the same prioritization matrix (Table 3) with a detailed process for rating fires (Table 4). This fire rating process assigns points to each fire based on 12 different criteria. Each criterion has a maximum value of 5 points, and a fire can

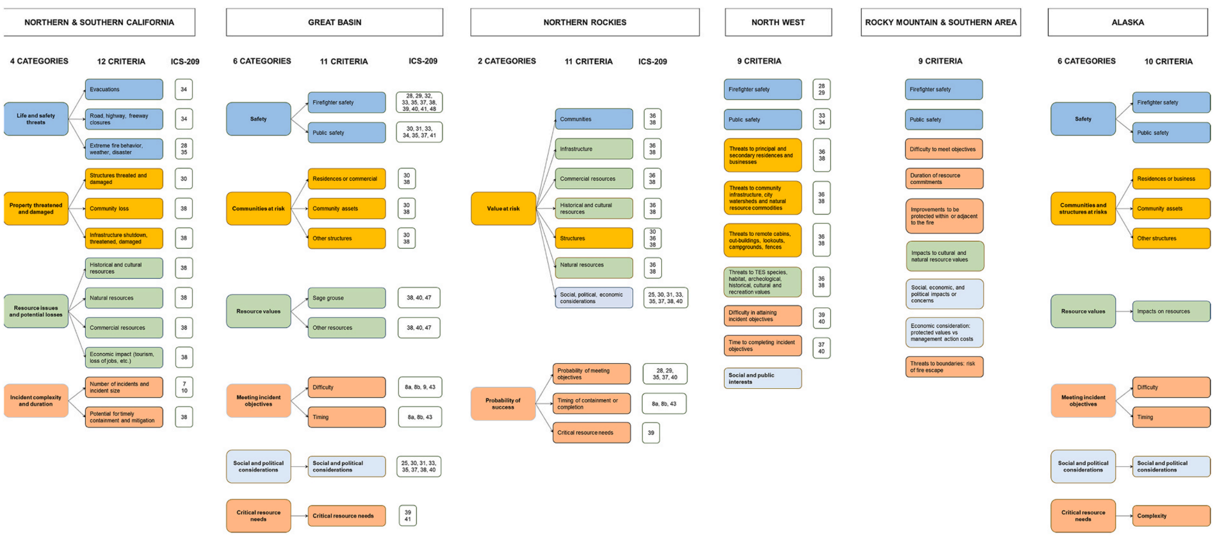

Fig. 2. Criteria for daily wildfire ranking that are included in the incident prioritization matrices of different GACCs in the US [12]. Similar criteria across different GACCs are highlighted with the same color. The two GACCs in California use the same matrix. The Rocky Mountain and the Southern Area also use the same matrix. Five GACCs posted online instructions for using the ICS-209 information to rank wildfires daily based on their matrix’s criteria. For these five GACCs, ICS-209 block identification numbers are presented next to the ranking criteria.

Table 1
Incident priority ratings implemented by different GACCs. ONCC: Northern California, OSCC: Southern California, GBCC: Great Basin, NRCC: Northern Region, NWCC: Northwest Region. More detailed information about the SIT-209 data fields can be found at https://www.wildfire.gov/application/sit209.

ICS-209 Block	SIT-209 Data Field Description (marked by "x" if used by a GACC)	OSCC-ONCC	GBCC	NRCC	NWCC
7	Current fire size	x			
8	Percentage of containment		x	x	
9	Description of fire incident		x		
10	Complexity level of fire incident	x			
25	Short description of fire ignition point		x	x	
28	Summary of significant events	x	x	x	x
29	Fuel model identification		x	x	x
	Narratives of hazardous material involvement				
30	Fire damage assessment	x	x	x	
	Information about affected structures				
31	Casualties and illnesses due to fire incident		x	x	
32	Casualties and illnesses due to fire incident		x		
33	Narratives of life safety and health status		x	x	x
34	Information about life safety management	x	x		x
35	Narratives of weather concerns	x	x	x	
36	Projected fire activity within the next 12 h			x	x
	Projected fire activity within the next 24 h				
	Projected fire activity within the next 48 h				
	Projected fire activity within the next 72 h				
	Projected fire activity beyond the next 72 h				
37	Information about strategies for incident management		x	x	x
38	Projected fire threat within the next 12 h	x	x	x	x
	Projected fire threat within the next 24 h				
	Projected fire threat within the next 48 h				
	Projected fire threat within the next 72 h				
	Projected fire threat beyond the next 72 h				
39	Projected critical fire management resources needed within the next 12 h		x	x	x
	Projected critical fire management resources needed within the next 12 h				
	Projected critical fire management resources needed within the next 12 h				
	Projected critical fire management resources needed within the next 12 h				
	Projected critical fire management resources needed beyond the next 72 h				
40	Strategic discussions		x	x	x
41	Planned actions		x		
43	Anticipated date for completing fire management objectives	x	x	x	
	Expected date for fire containment				
47	Remarks on the fire incident		x		
48	Identification of the management organization for the fire incident		x		

Table 2
Summary of the 2016–2020 California wildfire data connection between IMSR and SIT-209.

	2016	2017	2018	2019	2020	2016–2020
Total number of daily fire records reported by IMSR	596	1024	773	564	2175	5132
Total number of IMSR records connected to SIT-209	528	967	699	559	1918	4671
Total number of IMSR records unconnected to SIT-209	68	57	74	5	257	461
Successful connection rate (%)	89	94	90	99	88	91

accumulate up to 60 points from all criteria. Fires occurring on the same day in each GACC are ranked based on their total points, with higher values associated with higher rankings.

Based on the ICS-209 information recommended for California's fire ranking (Fig. 2), we collected associated data fields in the SIT-209 (Table 3) to build a computer program that automates California's process of ranking fires (Table 4). This computer program (https://github.com/thumit/SITrank) implements SQL matching queries, LUCENE proximity search, and JAVA wildcard search based on a set of predefined keywords, and allocates points to each criterion based on results of these matching and searching processes. For each criterion, a discrete value from 0 to 5 would be assigned when a keyword is found within the associated SIT-209 data fields. When one or multiple keywords were found for a criterion by multiple searching and matching techniques (SQL, LUCENE, JAVA wildcard), the maximum value would be assigned as the final score for that criterion (more details of the rule-based scoring for each criterion can be found in Appendix Tables A2-A13). The total point for each fire was calculated by summing its scores across all 12 criteria. Following the California ranking rules, we assigned a higher daily rank to a fire with a larger total point value.

Our first research objective is to evaluate the alignment between fire rankings determined by our rule-base program and historical fire rankings reported by IMSR. Comparisons were done using all 4671 fire records, with fire rankings can be between 1 and 16 on each single day. A high degree of consistency between the program's rankings and the historical rankings would indicate a high likelihood

Table 3
California prioritization matrix (reproduced from Ref. [12]) and the corresponding rule-based ranking. Each ICS-209 block may be associated one or multiple SIT-209 data fields, with details presented in Table 1. In this table, X represents three additional SIT-209 fields including "road closure flag", "area closure flag", and "trail closure flag", while Y represents one additional SIT-209 field "first date when the fire incident is reported", which we use to support the rule-based ranking process.

Wildfire ranking criteria (i.e. factors)	ICS-209 blocks recommended by the matrix	SIT-209 fields utilized by the rule-based program	Maximum points allocated to a ranking criterion
A. Life and safety threats			
A.1. Evacuations	34	33, 34	5
A.2. Road closures	34	33, X	5
A.3. Fire behavior and weather	28, 35	36	5
B. Property threatened or damaged			
B.1. Structures	30	30	5
B.2. Community	38	38	5
B.3. Infrastructure	38	38	5
C. Resource impact or loss			
C.1. Cultural resources	38	38	5
C.2. Natural resources	38	38	5
C.3. Commercial resources	38	38	5
C.4. Economic impact	38	38	5
D. Incident complexity and duration			
D.1. Number of fire incident and fire size	7, 10	7	5
D.2. Fire containment potential	43	43, Y	5
Maximum Total Priority Points			**60**

Table 4
California's wildfire incident rating process (reproduced from Ref. [12]). Note that for D1, the rule-based program only used total fire size to allocate points. This is because of the difficulty to identify the number of incidents, especially for the case of "complex" incident which contains a set of smaller incidents being grouped into a single large one.

California priority ranking	Priority Points
A.1. Evacuations in progress	5
A.1. Precautionary evacuations	3–4
A.1. Potential evacuations within 48–72 h or evacuations completed	1–2
A.2. Major highway or freeway closed	4–5
A.2. State routes or improved roadways closed	2–3
A.2. Potential for closures within 48–72 h	1
A.3. Extreme weather or disaster occurring or predicted to continue within 24 h	4–5
A.3. Extreme weather or disaster occurring or predicted to continue within 24–72 h	3–4
A.3. Extreme weather or disaster occurring but predicted to diminish	1–2
B.1. Greater than 200 structures threatened	4–5
B.1. Between 25 and 200 structures threatened	3–4
B.1. Less than 25 structures threatened	1–2
B.2. Potential for greater than 75 % community loss	5
B.2. Potential for 50–75 % community loss	4
B.2. Potential for 25–50 % community loss	3
B.2. Potential for less than 25 % community loss	1–2
B.3. Infrastructure systems shutdown and/or damaged	5
B.3. Potential threat to infrastructure within 24–48 h	3–4
B.3. Potential threat to infrastructure within 72+ hours	1–2
C.1. Historical and significant cultural resources	1–5
C.2. Natural resources (T&E species habitat, watershed, forest health, soils, air shed, etc.)	1–5
C.3. Commercial resources (grazing, timber, agricultural crops, etc.)	1–5
C.4. Potential for economic impact (tourism such as fishing or hunting, loss of jobs, etc.)	1–5
D.1. 5+ incidents or total fire size is greater than 25,000 acres	4–5
D.1. 3–4 incidents or total fire size is between 5000–25,000 acres	2–3
D.1. 1–2 incidents or total fire size is less than 5000 acres	1
D.2. Expected containment in less than 3 days	5
D.2. Expected containment in 3–7 days	4
D.2. Expected containment in 8–14 days	3
D.2. Expected containment in 15–21 days	2
D.2. Unknown time to containment or long-term management	1

that our interpretation of the published California ranking rules is accurate, and California's fire officials are adhering to their published ranking rules. On the other hand, inconsistencies could imply either a discrepancy in our scoring system, a lack of adherence of fire managers to their published rules, or influences of other factors beyond the official ranking rules that may drive the final daily ranking decisions.

Our second research objective is to explore the relationship between ranking criteria used by California, and examine the importance of each individual criterion to the state's fire ranking process. Two analyses were conducted to meet this objective, which utilized only fire records with historical rankings in the top N of each day (i.e. N = 2, 3, 4, or 5), as those are wildfires that hold greater significance for daily resource prioritization decisions. Fig. 3 illustrates the distribution of 4671 fire records in California where the linkage between IMSR and SIT-209 has been established, grouped by the number of fires per day (a single day may contain one to 16 simultaneous fires). We excluded days where there is only one fire (red column in Fig. 3) and retained only the top N fires from each day for investigation (blue columns in Fig. 3). We identified 1858, 2100, 2092, and 1865 fire records associated with the top 2, 3, 4, and 5 fires daily. For these top N fire records on each day, we first extracted points for every criterion determined by the rule-based program and summing points across all criteria to calculate the program's fire rankings, which may be different from historical rankings. We then employed a min-max normalization [41] for each separate set of N daily fires using those programmatic points. This normalization technique transforms the original modeling points assigned to each criterion (between 0 and 5) into a numerical value between 0 and 1, while the relative importance of each criterion is equally weighted across different days. Because normalization was applied separately for each day's set of N fires, the relative importance is evaluated among those N fires, independent of the other fires excluded from that day's selection. The normalized points were used in two analyses (both were conducted by using the R software).

For the first analysis, we used the two R libraries (hmisc and corrplot [42,43]) to create boxplots that visualize the relationship among different factors (or criteria) in the California matrix, and constructed a Spearman correlation matrix along with its correlogram to evaluate the correlation between each pair of factors. We followed a conventional approach to interpret the correlation coefficient values of less than 0.4 as "weak", between 0.4 and 0.7 as "moderate", and more than 0.7 as "strong" positive relationship between factors [44]. Understanding correlations among factors may benefit the fire ranking process, because if factors are strongly correlated, they may lead to unanticipated weighting (i.e., double counting) in the prioritization process.

For the second analysis, we employed a Random Forest (RF) classification model [40] in R (using 4 libraries: randomForest, dplyr, caTools and caret [45–48]) to examine the importance of each factor and identify the key factors that drive wildfire prioritization ranking in California. The RF technique is well-suited for wildfire management studies [49], such as its application in analyzing the influence of suppression resource decisions [7]. To implement our RF model, we split the top N daily fire records into two datasets: one for training the model with 80 % of all samples, and the other for testing the model accuracy with the remaining 20 % data. The RF model was trained using all 12 ranking factors from the California matrix as predictor variables. It is consisted of 500 decision trees, with each split in a tree considering a random subset of three variables, and a default terminal node size of one. Empirical studies have shown that using 500 decision trees is sufficient to ensure stable and accurate model predictions, with only marginal improvement in prediction outcomes when the number of trees exceeds 500 [50]. The number of variables at each node split was determined by using the *tuneRF* function within *R's randomForest* library, which identified 3 as the optimal value to achieve the best balance between model accuracy and computational efficiency. The default node size setting allowed each tree to grown to its maximum depth, and potential overfitting was mitigated by the ensemble nature of the RF algorithm, which aggregates results across 500 trees rather than constraining the depth of individual trees. To evaluate the importance of each variable, we reported and used the Mean Decrease Accuracy (MDA) as it is widely used for measuring variable importance [51–53]. The MDA measures the decrease in prediction accuracy when a

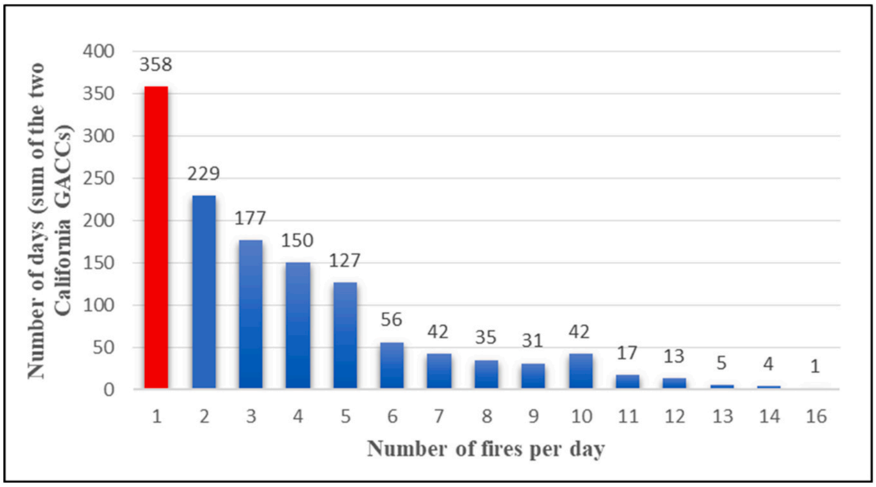

Fig. 3. Frequency in the two California's GACCs. A total of 4671 IMSR fire records were distributed into 1287 days (sum of the two GACCs) where there is at least one fire per day. Among these, days with only one fire (red column) were excluded, and only fire records in the top N (between 2 and 5) daily were included in our Random Forest analysis.

variable is removed. In our RF model, variables with higher values of MDA would indicate that they are more important factors, as the accuracy of fire ranking prediction would drop more significantly when these factors are excluded.

4. California case study: results

4.1. Comparing fire ranking by IMSR and by the rule-based modelling approach

Comparisons between the historical fire ranks reported by IMSR and the fire ranks generated by the rule-based program may be biased by daily observations that have very few fires on a single day. In particular, if there is just one fire on a day, both ranking approaches will place that fire at the top rank, resulting in a perfect ranking alignment and inflating the estimation of how well the two approaches agree. To mitigate this bias, in addition to fully examining 4671 fire records of all ranks (the E0 and %E0 columns in Table 5), we also explored two other cases with 4313 and 3855 fire records, respectively by excluding days with only one fire (the E1 and %E1 columns in Table 5) or excluding days with one or two simultaneous fires (the E2 and %E2 columns in Table 5). The outcomes revealed a consistent alignment in fire ranking between IMSR and the rule-based program. This consistency is evident by a larger number of fire records associated with a smaller difference in fire rank between the two approaches (Table 5), and a notable concentration of larger numbers along the diagonal of the fire ranking heatmap (red, orange, and yellow cells in Fig. 4).

We closely examined cases where there was substantial fire ranking misalignment (the fire counts in the green cells in Fig. 4). For instance, seven fire records were historically ranked first, but beyond tenth (10+) by the rule-based program. Conversely, there was one fire record that would be ranked either first or second (1–1.5) by the rule-based program but was ranked eighth historically. To gain a better understanding of the underlying causes for these ranking deviations, we examined 107 records with a fire ranking difference of five or more, which contains 2.3 % of the total 4671 fire records. Our analysis revealed that the rule-based method encountered difficulties in assigning points for several ranking criteria (B2, B3, C1, C2, C3, C4), as illustrated in Fig. 5. This issue mainly arises from infrequent scenarios where the values at risk cannot be easily identified by keyword search. For example, with information like “The Karuk tribe ancestral lands are currently threatened with numerous sites in the area”, our rule-based program did not assign the appropriate points as none of the words in the sentence are covered by the predefined set of keywords used by the program. Consequently, the program underestimated the fire threat.

4.2. Examining factors that influence wildfire prioritization ranking in California

California fires with historical ranks in the top five daily generally had a higher average score determined by the rule-based program for each scoring criterion (Fig. 6). Because the rules reflect actual management decision process to give more points to a criterion (i.e. “factor”) when it poses a higher risk, our finding indicates a consistence tendency for fires with higher historical ranks to be associated with increased fire risks simultaneously to human, natural resources, among other factors. This tendency is illustrated by the correlogram in Fig. 7, which shows positive linear relationships between each pair of factors, and negative linear relationships between each factor and fire rank. An exceptional case was observed for D2 (fire containment potential) which, as shown in both Figs. 6 and 7, has negative correlations with all other factors and has a slightly positive correlation with fire rank. This contrast between D2 and the other factors is expected, as fire mangers often prefer assigning a higher priority to a fire when its containment can be achieved more quickly (note that D2 receives a higher score when less time is anticipated for containment). However, fires associated with greater risks (reflected by higher scores for factors other than D2, such as larger fire size), are typically more difficult to contain and often require longer suppression time (lower D2 score) and efforts. Most of the correlation coefficient values as shown in Fig. 7 are close to zero, indicating generally weak linear relationships. The lack of strong correlations among all factors means that they are relatively independent of each other in terms of their influence on fire ranking. As it is less likely that one factor could be used as a

Table 5
The difference between fire rank reported in IMSR and fire rank generated by the rule-based approach. For the rule-based approach, fires with equal total points were assigned with the same average rank (e.g., two fires with the same 50 points will be ranked 1.5 instead of 1 or 2). Column abbreviations: **E0** for all 4671 fire records; **E1** for records from days with more than 1 fire/day; **E2** for records from days with more than 2 simultaneous fires/day; **%E0**, **%E1**, and **%E2** were calculated by dividing the values of **E0**, **E1**, and **E2** by the **SUM** within each respective column.

Difference in fire rank (IMSR vs rule-based)	Number of fire records within the difference in fire rank					
	E0	E1	E2	%E0	%E1	%E2
0–0.5	2029	1671	1349	43.4	38.7	35.0
1–1.5	1527	1527	1391	32.7	35.4	36.1
2–2.5	652	652	652	14.0	15.1	16.9
3–3.5	216	216	216	4.6	5.0	5.6
4–4.5	140	140	140	3.0	3.2	3.6
5–5.5	72	72	72	1.5	1.7	1.9
6–6.5	22	22	22	0.5	0.5	0.6
7–7.5	7	7	7	0.1	0.2	0.2
8–8.5	4	4	4	0.1	0.1	0.1
9–9.5	2	2	2	0.0	0.0	0.1
SUM (total number of records)	**4671**	**4313**	**3855**	**100**	**100**	**100**

RULE-BASED FIRE RANK	IMSR FIRE RANK									
	1	2	3	4	5	6	7	8	9	10
1-1.5	890	257	101	56	18	6	3	1	0	0
2-2.5	252	378	165	85	33	8	9	1	4	1
3-3.5	97	169	209	136	49	14	16	9	4	0
4-4.5	20	72	128	126	56	35	18	10	7	1
5-5.5	12	16	49	50	99	54	24	7	10	3
6-6.5	6	11	17	28	59	45	38	29	3	2
7-7.5	3	8	9	7	19	27	28	34	20	3
8-8.5	0	2	6	6	6	19	14	33	17	10
9-9.5	0	0	3	8	10	6	6	4	23	37
10+	7	16	13	21	24	32	34	20	25	105

Fig. 4. Heatmaps of 4671 fire ranking records based on the historical IMSR data and based on the rule-based ranking program. When multiple fires have the same total point determined by the program, they will receive an equivalent non-integer rank. For instance, a value of 1.5 represents the case when the top two fires have the same total point given by the program, and both fires are assigned a rank of 1.5 instead of 1 or 2. In contrast, historical IMSR fire ranks are always discrete integer values.

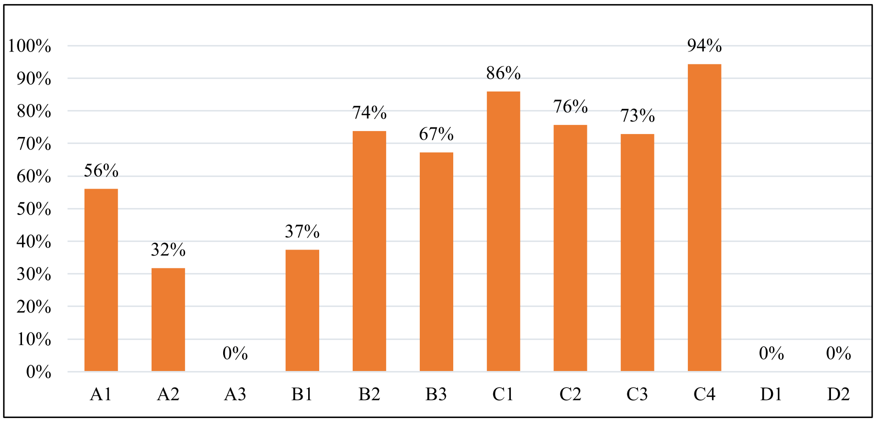

Fig. 5. Percentage of fire records with zero points allocated by the rule-based program for at least one ranking criterion. This was calculated based on the total of 107 fire records, where the difference between fire ranks determined by the rule-based program and historical fire ranks reported by IMSR is greater than or equal to 5.

proxy for another, fire managers should analyze data and interpret findings for each factor separately to understand its individual impact on the overall fire ranking.

Results from RF models trained with different numbers of fires per day (N) are presented in Table 6. As expected, the overall model's accuracy for fire rank prediction declines as more fires are ranked, dropping from 75 % to 45 % when N increases from 2 to 5. Across all models, the prediction accuracy (i.e. "Accuracy" values in Table 6) significantly surpasses the classification accuracy expected from random guessing (i.e., "NIR" values in Table 6), as also illustrated by Fig. 8. This indicates that RF models can capture meaningful prediction patterns based on the available data, despite inherent data uncertainties and the limited instructions for fire ranking evaluation. It is important to note that the reported accuracy reflects the discrepancy between model-predicted rankings and historical rankings. The true predictive ceiling is unknown, because historical rankings are influenced not only by the rule-based scoring criteria but also by additional contextual, operational, or other considerations that are not officially represented in the current ranking rules (or decision matrix). Therefore, imperfect accuracy does not imply the modeling approach is inadequate; rather, it reflects the challenge of predicting management decisions based solely on documented inputs.

Strong correlations among predicting variables can lead to multicollinearity, which may affect the interpretability of variable importance in a RF model. However, correlations among California's 12 fire ranking factors are generally weak, which benefits our interpretation of RF model results. Fig. 9 presents the importance of fire ranking factors measured by Mean Decrease Accuracy (MDA). Different RF models consistently identify the same set of three most influential factors: B1 (threats and damage to structures), D1 (fire size), and A1 (evacuations). However, in the Top-5-Fires RF model, D1 is identified as a stronger overall predictor than B1 (Fig. 9). This shift may be due to the increased sample size making it clearer that some larger-size fires may not necessarily cause greater threats and damages. Although the correlation coefficients between B1 and D1 indicate their positive relationship (from 0.32 to 0.39 as in Fig. 7),

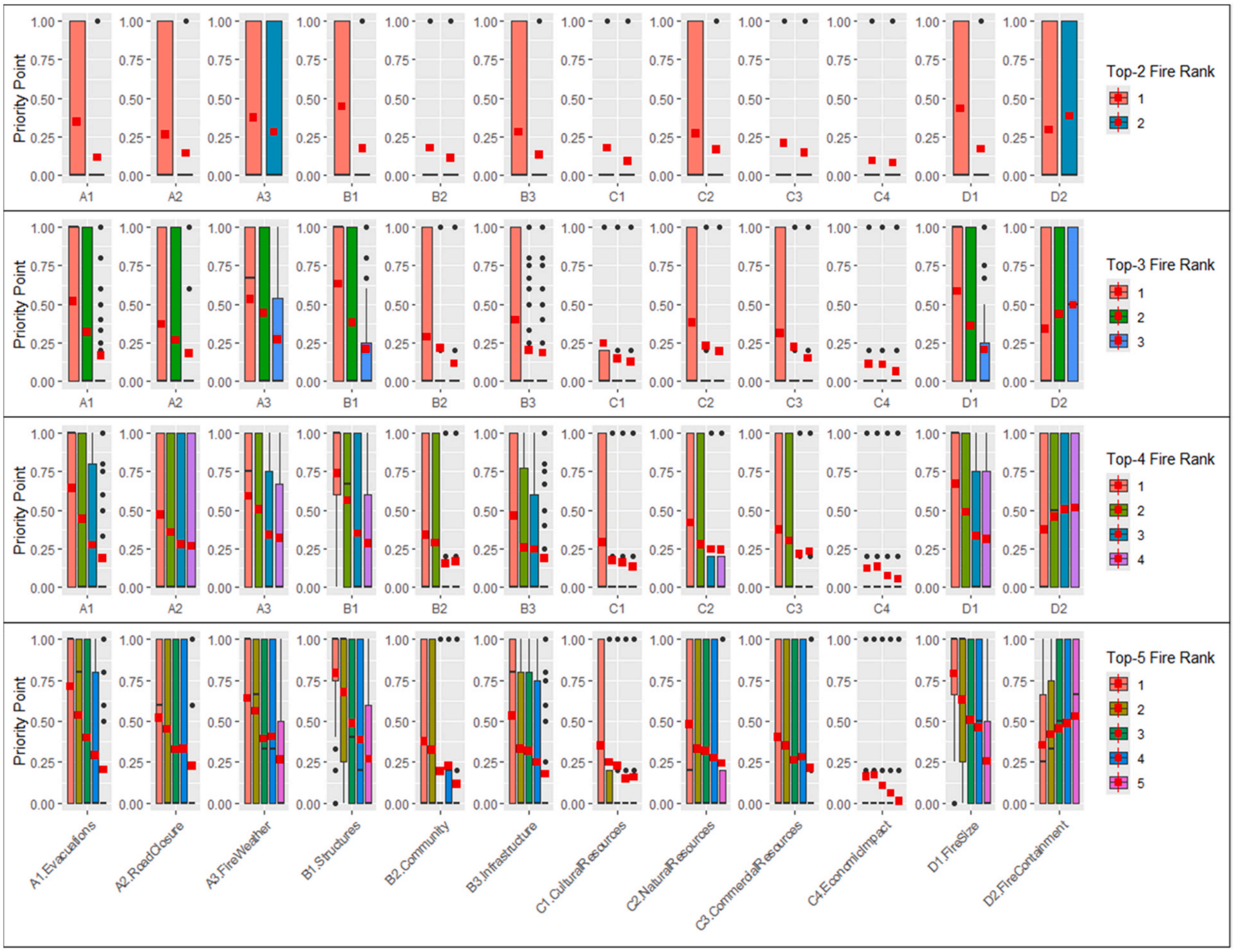

Fig. 6. Visualization of boxplots depicting normalized fire ranking points determined by the rule-based program. Columns with different colors represent different historical fire ranks. In each column, a red dot represents the average value of normalized points for a criterion, calculated from all fire records with the same historical fire rank.

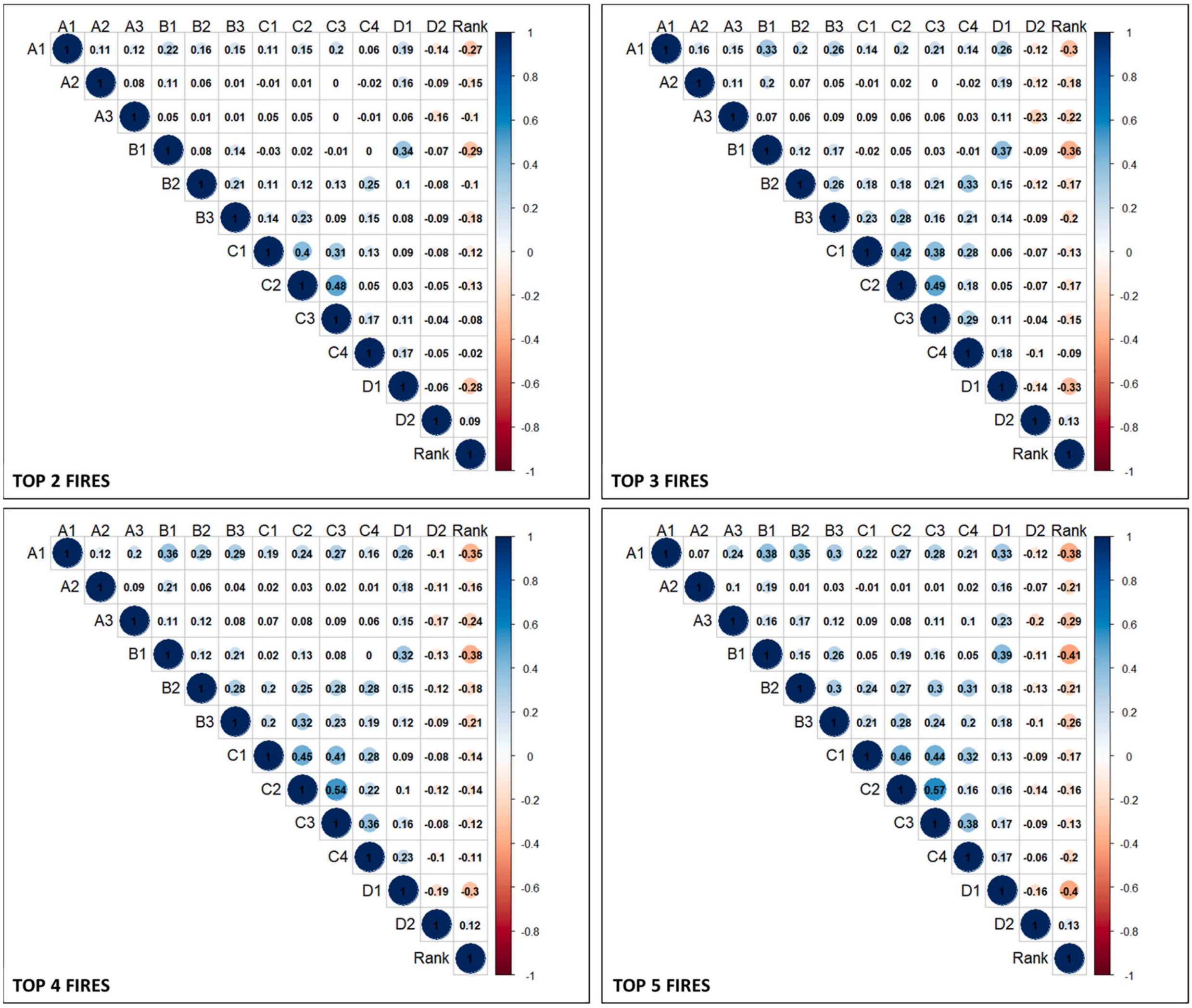

Fig. 7. Correlogram with Spearman correlation coefficient values presented within the correlogram (instead of those values being presented in a separate correlation matrix). Positive correlations are shown in blue, and negative correlations are shown in red. Color intensity and circle size are proportional to the correlation coefficient values. Circles are absent when correlations are not significant at 0.05 level.

the correlation is not strong enough to ensure the consistent dominance of each of these two factors across different RF models with varying number of ranking fires.

The nine remaining factors (excluding B1, D1, A1) have lower importance with inconsistent MDA orders across RF models (Fig. 9). Nonetheless, several patterns emerge. For example, B3 (threat and damage to infrastructure) and A3 (fire behavior and weather) have MDA values increased consistently when N increases, suggesting that they play a greater role when ranking a larger number of fires. In contrast, C4 (economic impact) and C2 (natural resource impact) display their decreased importance when more fires are ranked. The remaining 5 factors, including A2 (road closures), B2 (community impact), C1 (cultural resource impact), C3 (commercial resource impact), and D2 (fire containment potential), have varied MDA values when N increases, implying their inconsistency for fire ranking prediction. This inconsistency may stem from uncertainties in the information upon which the ranking evaluation process relies. For instance, the MDA importance orders of A3 and D2 fluctuate across models with different number of fires per day (Fig. 9), as fire behavior and weather conditions (A3) can change rapidly during a day, while the potential for fire containment (D2) is also highly uncertain, especially for large fires where the time required for containment is often difficult to estimate.

A notable observation from our sample fires is that the same set of N fires may preserve their rankings (or their relative ranking orders) for several days, and all such instances are used to train the RF models. Including these stable-ranking fires objectively reflects the real-world fire progression and persistence. However, there is a potential that the models could become overly dependent on these repeated samples. To assess whether RF models could be biased by this temporal dependency, we conducted additional tests to compare models with and without a temporal variable "PR", which represents the "previous rank" of the same fire on the closest preceding day. Results show that models with PR have significantly higher prediction accuracy than those without PR

Table 6
Results of the Random Forest models developed for California fire ranking. Classes from R1 to R5 respectively represent the first to the fifth priority in daily fire ranking.

Top N fires	N = 2			N = 3				N = 4					N = 5					
Confusion matrix																		
Observed (→)		R1	R2		R1	R2	R3		R1	R2	R3	R4		R1	R2	R3	R4	R5
Predicted (↓)	R1	130	38	R1	104	34	10	R1	73	24	5	11	R1	42	16	8	4	3
	R2	56	148	R2	17	55	25	R2	15	48	17	12	R2	11	30	11	8	3
				R3	19	51	105	R3	6	17	43	26	R3	9	16	18	11	8
								R4	11	16	40	56	R4	5	4	17	32	14
													R5	8	9	21	20	47
Overall statistics																		
Accuracy	0.75			0.63				0.52					0.45					
95 % Confidence Interval	0.70–0.79			0.58–0.67				0.47–0.57					0.40–0.50					
No Information Rate (NIR)	0.50			0.33				0.25					0.20					
P-Value (Accuracy > NIR)	2.0E-16			2.2E-16				2.0E-16					2.00E-16					
Kappa	0.4946			0.4429				0.3651					0.3133					
Mcnemar's Test P-Value	0.0795			0.0006				0.4566					0.0789					
Statistics by Class																		
	R1				R1	R2	R3		R1	R2	R3	R4		R1	R2	R3	R4	R5
Sensitivity	0.70				0.74	0.39	0.75		0.70	0.46	0.41	0.53		0.56	0.40	0.24	0.43	0.63
Specificity	0.80				0.84	0.85	0.75		0.87	0.86	0.84	0.79		0.90	0.89	0.85	0.87	0.81
Positive Predictive Value	0.77				0.70	0.57	0.60		0.65	0.52	0.47	0.46		0.58	0.48	0.29	0.44	0.45
Negative Predictive Value	0.73				0.87	0.74	0.86		0.90	0.83	0.81	0.84		0.89	0.86	0.82	0.86	0.90
Prevalence	0.50				0.33	0.33	0.33		0.25	0.25	0.25	0.25		0.20	0.20	0.20	0.20	0.20
Detection Rate	0.35				0.25	0.13	0.25		0.17	0.11	0.10	0.13		0.11	0.08	0.05	0.09	0.13
Detection Prevalence	0.45				0.35	0.23	0.42		0.27	0.22	0.22	0.29		0.19	0.17	0.17	0.19	0.28
Balanced Accuracy	0.75				0.79	0.62	0.75		0.78	0.66	0.63	0.66		0.73	0.65	0.55	0.65	0.72

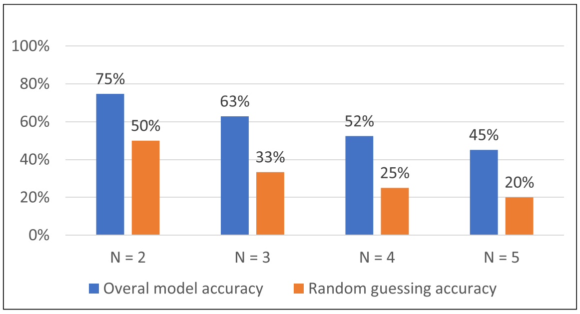

Fig. 8. Compare the accuracy of fire ranking by the Random Forest model and by random selection (i.e., equal probability for each ranking class).

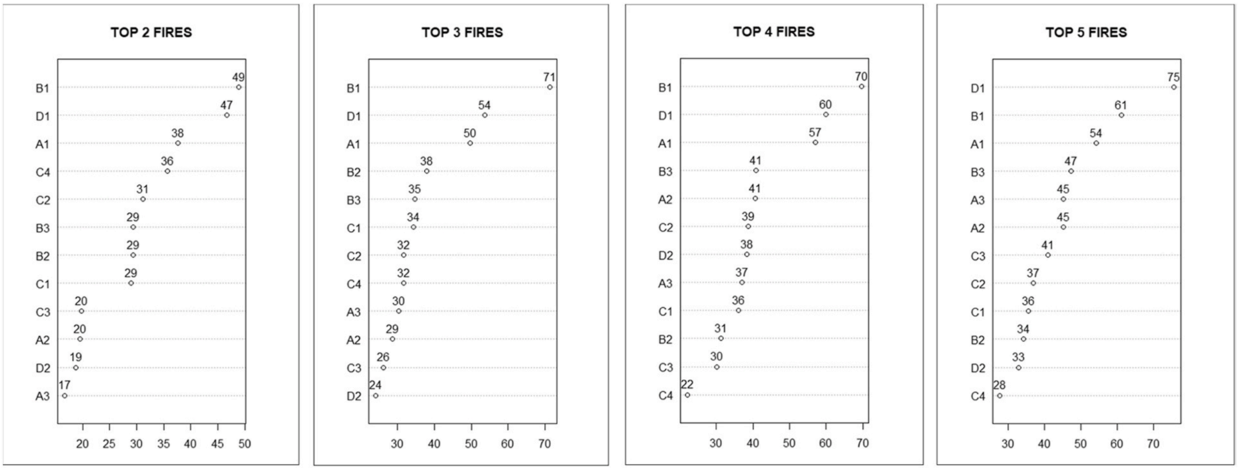

Fig. 9. Mean Decrease Accuracy (MDA) reported for different Random Forest models.

(Appendix Table A14 vs Table 6). However, the top three most influential factors (excluding PR) and their importance rankings remain the same between models with or without PR (Appendix Fig. A15), indicating that fire rank prediction is not overly reliant on past rankings. This consistency suggests that there are consistent drivers for fire ranking beyond the temporal ranking dependency, which can mitigate our concern about bias.

5. Conclusions and discussion

The daily ranking of wildland fires is a practice in most GACCs of the US to facilitate effective allocations of suppression resources to ongoing large fires. In this study, we first examined publicly available wildfire ranking systems across GACCs. We then developed a rule-based program to mimic the fire ranking process in California using its 12 ranking criteria. Further random-forest-based analyses were conducted to understand the relationship among these criteria and identify key criteria for fire ranking in the state.

Each GACC of the US may implement a different fire ranking system, and we found that these systems generally agree on using five broad groups of criteria to guide ranking decisions (Fig. 2). Within each GACC, specific criteria may reflect regional priorities, such as the impact to the sage grouse habitat which is used as a ranking criterion in the Great Basin but not everywhere else. While many GACCs use similar criteria for fire ranking, their specific instructions for assessing each criterion may differ. For some GACCs, detailed instructions for fire ranking based on their criteria are difficult to find, as this information is not always posted publicly. California was selected as our case study not only due to its escalated wildfire activity, but also because of its well-documented and publicly accessible fire ranking process.

Our rule-based program was designed to assess the California's fire ranking process based on matching its 12 fire ranking criteria to the historical wildfire data. One challenge in implementing the rule-based approach lies in the fact that historical fire dataset (SIT-209) contains both quantitative and qualitative data fields. Rating fires using quantitative data (e.g., fire size, fire cost, number of damaged structures, etc.) is much easier for a computer program than qualitative data (e.g., description of fire behavior in multiple sentences). Rating fires based on qualitative data (e.g., narrative descriptions of fire behavior) is more difficult because it requires interpreting unstructured text and extracting relevant but non-standard details into discrete options to support further evaluations. The California matrix guidelines provide limited guidance on qualitative data fields, making it challenging to establish consistent rules for point assignment. To address this, we used our best judgment to select keywords that aligned with the matrix guidelines and assigned rating points accordingly. However, some scenarios were still not well captured, which impaired our program's effectiveness in allocating fire ranking points (as illustrated in Fig. 5). While fire ranking predictions by the program are still significantly better than random guessing, its accuracy could be further improved by expanding the set of keywords used to assess qualitative criteria and finetuning the current point allocation rules.

Several important insights can be derived from our case study for California. First, the rule-based program was able to generate fire rankings that align well with the historical fire rankings, which implies a high likelihood that GACC fire managers priorities do match the process documented in the California prioritization matrix when determining fire rankings, and our interpretation of the California ranking rules was generally accurate. Second, we discovered that fires with higher ranks often have higher scores across all 12 factors of the California's fire ranking system. Generally, historical fires with higher daily ranks were reported with more comprehensive information, which helps explain why our program's prediction is more accurate for these higher-rank historical fires. Third, we found generally weak correlations among many factors influencing fire rating scores. This finding underscores the need for fire managers to carefully evaluate the contribution of each individual factor to the overall fire ranking. Including all factors in the ranking process is important for assessing each individual factor's contribution because they are less correlated than we expected. Finally, the Random Forest machine learning models provided insights into the relative importance of these factors. These models have consistently identified threats and damage to structures (B1), fire size (D1), and evacuations (A1) as the three most influential factors for daily fire ranking in California. Closer examinations on these three factors indicate that they have higher correlations with fire rankings, and the inter-correlations among those three factors are also higher compared to the other factors. This is potentially because these three factors are easier to quantify, or they may be reliably present on all high priority fires whereas the other factors are more hit or miss. Our results also suggest that when more fires are ranked simultaneously, the identified top three factors (B1, D1, and A1) still remain dominant, while several secondary factors such as threats to infrastructure (B3) and fire behavior/weather (A3) gain greater relative importance. This likely reflects a consistent and systematic process in which managers consider additional factors to differentiate among multiple high-severity fires that share similar characteristics.

Most GACCs use factors in their ranking systems that are largely backward looking, which focus on existing fire risks rather than the potential to achieve specific fire risk mitigation objectives. This approach may overlook scenarios where resources can be more effective on fires at lower risk. For example, significant usage of firefighters may not help mitigate the damage of higher-ranked fires to structures during extreme weather, while personnel could yield greater marginal efficiency on protecting structures from the emerging but still lower-rank fires. Considerations such as resource efficiency and the likelihood of achieving mitigation objectives are not explicitly documented in California's fire ranking process but are officially incorporated in the fire ranking systems of some other GACC regions (Fig. 2). While these considerations are absent from California's formal guidelines, it is likely that fire managers will account for them informally before making the final wildfire prioritization and resource allocation decisions.

The GACC wildland fire response relies on coordination between agencies to allocate shared resources for containing fires or minimizing fire damage within a GACC. This multiagency coordination process is, theoretically, an economically efficient way to allocate resources, but it may complicate the assessment of wildfire readiness at the agency level. While hiring decisions and equipment acquisition is done at the agency level, the availability of those resources has implications for the broader interagency effort. An individual agency is typically looking to ensure that it has adequate resources for initial fire attack, but when a large fire occurs, it may require resource mobilization across the GACC's multiple agencies. The incident prioritization process, particularly the use of a GACC decision matrix, is a method used to get multiple agencies in agreement when it comes to sending resources to a fire. While a GACC matrix provides a structured framework to allocate resources efficiently, assessment processes relying on a matrix can also unintentionally skew discussions. Pitfalls include potential domination by strong personalities, groupthink, and over-reliance on rigid criteria that may not fully account for resource effectiveness. Revisiting and refining the decision matrix to reflect contemporary challenges and priorities will be essential for its continued relevance and efficacy. Ultimately, enhancing transparency, integrating forward-looking factors, and strengthening collaboration with fire managers are critical steps toward improving wildfire prioritization and achieving balanced, efficient, and effective multi-agency resource allocation.

While our study focuses on fire ranking, future studies can build on the existing knowledge of fire ranking systems by exploring how fire ranking influence the daily and cumulative allocation of suppression resources, such as engines, hand crews, dozers, and aircraft. A deeper understanding of the relationship between fire ranking and resource distribution can provide valuable insights into optimizing firefighting strategies and resource management. For example, a recent study has identified daily resource joint ordering and assignment patterns for large fire incidents [23]. One potential direction of our further research is to connect fire ranking with the joint allocation of different resource types. While a potential correlation between fire ranking and resource allocation or funding may exist, we have not seen any effort that specifically addresses this topic. Further research expansions could also evaluate temporal dynamics and resource shift challenges to better support real-world prioritization decisions. For instance, temporal factors such as the number of days a fire remains a high-priority event, or operational constraints such as relocating resources from ongoing but lower-ranked fires to newly emerging higher-ranked fires, may influence both daily fire ranking and resource dispatch decisions. Advancing studies in these areas will contribute to understanding and improving wildfire management and response efforts.

As a final remark, achieving sustainable wildfire management may ultimately require coexisting with wildfires [54] rather than relying solely on fire suppression. While fire managers and researchers are working to find viable and sound solutions, suppression remains a critical tool for addressing immediate fire threats. This study assesses the fire prioritization process in California, a region with some of the most critical wildfire challenges in the US. It provides insights into how fire managers evaluate and rank wildfires, which can be useful for decision making in fire assessment and fire response. Findings from this research also offer evidence for potential improvements of fire management policies, regulations, and practices to mitigate fire risks, enhance community resilience in fire-prone areas, and build public trust in fire management efforts.

CRediT authorship contribution statement

Dung Nguyen: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Erin J. Belval:** Writing – review & editing, Validation, Project administration, Methodology, Investigation, Conceptualization. **Yu Wei:** Writing – review & editing, Writing – original draft, Validation, Project administration, Methodology, Investigation, Conceptualization. **David E. Calkin:** Writing – review & editing, Validation, Supervision, Conceptualization.

Declaration of competing interest

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

Acknowledgements

USDA Forest Service - Rocky Mountain Research Station and Colorado State University for jointly supporting this work through a Challenge Cost Share Agreement #22-CS-11221636-189. The findings and conclusions in this paper are those of the author(s) and should not be construed to represent any official USDA or U.S. Government determination or policy.

Appendix A. Supplementary materials

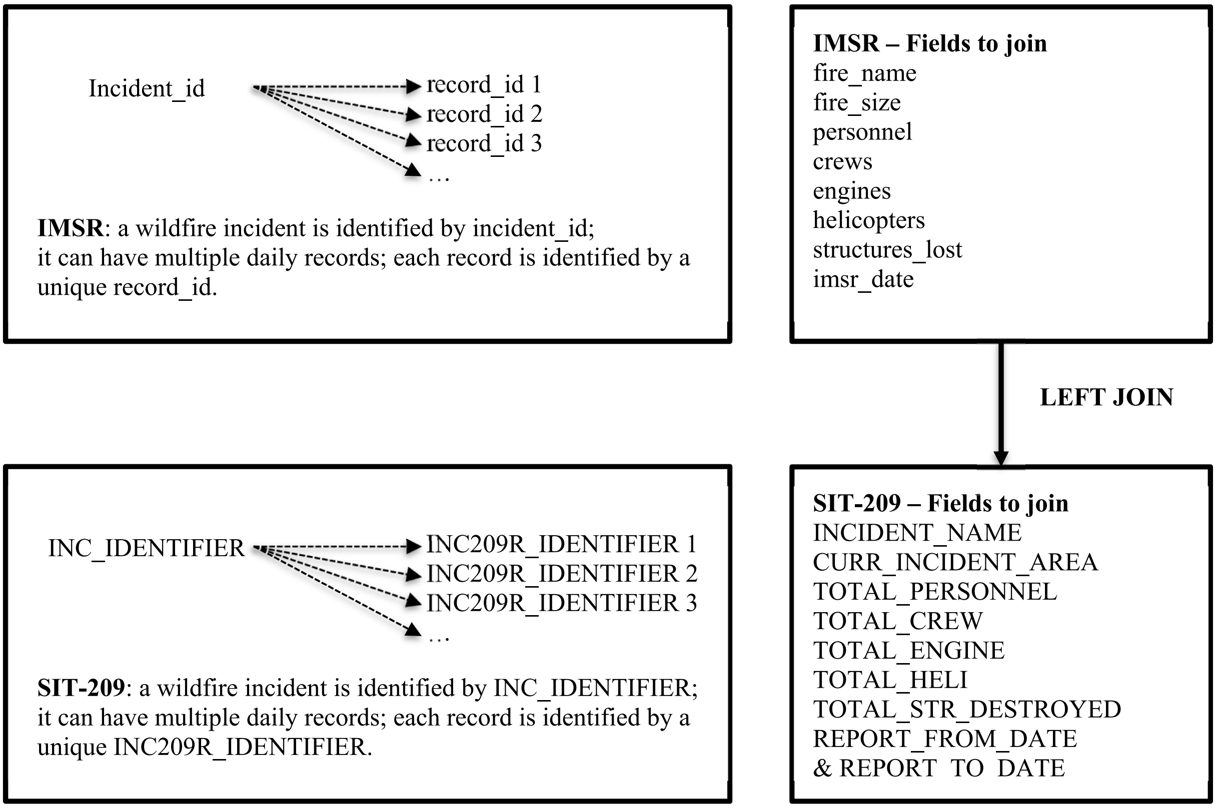

Fig. A1. Matching daily fire records between SIT-209 and IMSR datasets.

Table A2
Calculate A1. The final A1 point is the maximum between Box 33 point and Box 34 point.

Box 34 SQL Queries	Box 34 Point
No Evacuation(s) Imminent	1
Planning for Evacuation	3
Evacuation(s) in Progress	5
Others	0

(*continued on next page*)

Table A2 (*continued*)

Box 34 SQL Queries			Box 34 Point
BOX 33 LUCENE Search		**BOX 33 JAVA Wildcard Searches**	Box 33 **Point**
evac*	NO		0
AND NOT no evac*~2,	YES	evac*lifted	1
evac* center*		potential*evac, evac*expected	2
		advisor*evac, evac*advisor, voluntary*evac, evac*notice, evacuation not	3
		evac*warning	4
		Null results for all the above searches	5

Table A3
Calculate A2. The final A2 point is the maximum between X point and Box 33 point.

X Fields SQL Queries			X Point
EITHER ROAD_CLOSURE_FLAG, AREA_CLOSURE_FLAG, TRAIL_CLOSURE_FLAG			3
Others			0
BOX 33 LUCENE Search		**BOX 33 JAVA Wildcard Search**	Box 33 **Point**
EITHER highway*, hwy* motorway*, area*, road*, rd, route*, trail*	NO		0
AND clos*	YES	potential*clos, possible*clos, clos*developed, clos*assessed	1
AND NOT discontinu*, lift*, remove*, *open*		area*, road*, rd, route*, trail*	3
AND NOT no clos*~4, clos* none~4, allow* public~1		motorway*, highway*, hwy*	5

X includes: ROAD_CLOSURE_FLAG, AREA_CLOSURE_FLAG, TRAIL_CLOSURE_FLAG

Table A4
Calculate A3. The final A3 point is the Box 36 point.

BOX 36 LUCENE Search		BOX 36 JAVA Wildcard Search	Box 36 Point
NOT none, contained, controlled, no spread*~1, no grow*~1, declare* out~1, burn* out~1, no activit*~1, no threat*~1	NO		1
	YES	little, minimal, small, low, confined, limited, no substantial	2
		to move, continu*grow, continu*spread	3
		moderate, varying rates, increase*activit, active	4
		rapid, continu*priority, more priority	5

Table A5
Calculate B1. The final B1 point is the Y point.

Y Field SQL Queries	Y Point
Number of destroyed or damaged or threatened structures: 200+	5
Number of destroyed or damaged or threatened structures: 100-199	4
Number of destroyed or damaged or threatened structures: 25-99	3
Number of destroyed or damaged or threatened structures: 5-24	2
Number of destroyed or damaged or threatened structures: 1-4	1
Number of destroyed or damaged or threatened structures: 0	0

Y includes SIT209_HISTORY_INCIDENT_209_AFFECTED_STRUCTS.

Table A6
Calculate B2. The final B2 point is the Box 38 point.

BOX 38 LUCENE Search		BOX 38 JAVA Wildcard Search	Box 38 Point
communit*	NO		0
AND EITHER threat*, impact*, risk* **AND NOT** no threat*~4, no risk~4, no impact~4	YES	potential*threat, potential*risk, potential*impact, possible*threat, possible*risk, possible*impact, could be*threat, could be*risk, could be*impact, minimal*threat, minimal*risk, minimal*impact, low*threat, low*risk, low*impact, limited*threat, limited*risk, limited*impact, slight*threat, slight*risk, slight*impact, threat*decrease	1
		moderate*threat, moderate*risk	3
		Null results for all the above searches	5

Use: CURRENT_THREAT_12, 24, 48.

Table A7
Calculate B3. The final B3 point is the Box 38 point.

BOX 38 LUCENE Search	Field for JAVA Wildcard Search		Box 38 Point
EITHER power*, energy*, water*, line*, corridor*, system*	CURRENT_THREAT_GT72	YES	1
AND NOT no threat*~4, no risk~4, no impact~4, not damage~4	CURRENT_THREAT_72	YES	2

(*continued on next page*)

Table A7 (*continued*)

BOX 38 LUCENE Search	Field for JAVA Wildcard Search		Box 38 Point
	CURRENT_THREAT_48	YES	3
	CURRENT_THREAT_24	YES	4
	CURRENT_THREAT_12	YES	5
	And also search for either: shut*down, damage*		
	Null results for all the above searches		0

Use: CURRENT_THREAT_12, 24, 48, 72, GT72.

Table A8
Calculate C1. The final C1 point is the Box 38 point.

BOX 38 LUCENE Search		BOX 38 JAVA Wildcard Search	Box 38 Point
EITHER histor*, cultur* **AND EITHER** threat*, impact*, risk* **AND NOT** no threat*~4, no risk~4, no impact~4	NO		0
	YES	potential*threat, potential*risk, potential*impact, possible*threat, possible*risk, possible*impact, could be*threat, could be*risk, could be*impact, minimal*threat, minimal*risk, minimal*impact, low*threat, low*risk, low*impact, limited*threat, limited*risk, limited*impact, slight*threat, slight*risk, slight*impact, threat*decrease	1
		moderate*threat, moderate*risk	3
		Null results for all the above searches	5

Only use: CURRENT_THREAT_12.

Table A9
Calculate C2. The final C2 point is the Box 38 point.

BOX 38 LUCENE Search		BOX 38 JAVA Wildcard Search	Box 38 Point
EITHER natural resource*, specie*, habitat*, watershed*, forest*, soil*, air*shed* **AND EITHER** threat*, impact*, risk* **AND NOT** no threat*~4, no risk~4, no impact~4	NO		0
	YES	potential*threat, potential*risk, potential*impact, possible*threat, possible*risk, possible*impact, could be*threat, could be*risk, could be*impact, minimal*threat, minimal*risk, minimal*impact, low*threat, low*risk, low*impact, limited*threat, limited*risk, limited*impact, slight*threat, slight*risk, slight*impact, threat*decrease	1
		moderate*threat, moderate*risk	3
		Null results for all the above searches	5

Only use: CURRENT_THREAT_12.

Table A10
Calculate C3. The final C3 point is the Box 38 point.

BOX 38 LUCENE Search		BOX 38 JAVA Wildcard Search	Box 38 Point
EITHER commercial resource*, graz*, timber*, agricultur*, crop* **AND EITHER** threat*, impact*, risk* **AND NOT** no threat*~4, no risk~4, no impact~4	NO		0
	YES	potential*threat, potential*risk, potential*impact, possible*threat, possible*risk, possible*impact, could be*threat, could be*risk, could be*impact, minimal*threat, minimal*risk, minimal*impact, low*threat, low*risk, low*impact, limited*threat, limited*risk, limited*impact, slight*threat, slight*risk, slight*impact, threat*decrease	1
		moderate*threat, moderate*risk	3
		Null results for all the above searches	5

Only use: CURRENT_THREAT_12.

Table A11
Calculate C4. The final C4 point is the Box 38 point.

BOX 38 LUCENE Search		BOX 38 JAVA Wildcard Search	Box 38 Point
EITHER economic*, tourism*, fishing*, hunting*, job* **AND EITHER** threat*, impact*, risk* **AND NOT** no threat*~4, no risk~4, no impact~4	NO		0
	YES	potential*threat, potential*risk, potential*impact, possible*threat, possible*risk, possible*impact, could be*threat, could be*risk, could be*impact, minimal*threat, minimal*risk, minimal*impact, low*threat, low*risk, low*impact, limited*threat, limited*risk, limited*impact, slight*threat, slight*risk, slight*impact, threat*decrease	1
		moderate*threat, moderate*risk	3
		Null results for all the above searches	5

Only use: CURRENT_THREAT_12.

Table A12
Calculate D1. The final D1 point is the Box 43-Z point.

Box 7 SQL Queries	Box 7 Point
Incident area (acre): 25000+	5
Incident area (acre): 15000-24999	4
Incident area (acre): 5000-14999	3
Incident area (acre): 2500-4999	2
Incident area (acre): 1-2499	1
Incident area (acre): 0	0

Only use: CURR_INCIDENT_AREA

Table A13
Calculate D2. The final D2 point is the Box 7 point.

Box 43-Z SQL Queries	Box 43-Z Point
Anticipated days to contain the fire: <3	5
Anticipated days to contain the fire: 3-7	4
Anticipated days to contain the fire: 8-14	3
Anticipated days to contain the fire: 15-21	2
Anticipated days to contain the fire: >22 or unknown	1

Only use: REPORT_FROM_DATE, ANTICIPATED_COMPLETION_DATE

Table A14
Results of the Random Forest models developed for California fire ranking with the additional inclusion of PR variable, which is the previous rank of the same fire on the closest preceding day. Classes from R1 to R5 respectively represent the first to the fifth priority in daily fire ranking.

Top N Fires	N = 2			N = 3				N = 4					N = 5					
Confusion matrix																		
Observed (→)		R1	R2		R1	R2	R3		R1	R2	R3	R4		R1	R2	R3	R4	R5
Predicted (↓)	R1	150	35	R1	112	20	7	R1	82	18	4	2	R1	52	10	7	1	1
	R2	36	151	R2	20	97	35	R2	13	68	18	9	R2	14	45	10	7	5
				R3	8	23	98	R3	5	14	59	28	R3	6	12	31	11	6
								R4	5	5	24	66	R4	1	5	13	34	16
													R5	2	3	14	22	47
Overall statistics																		
Accuracy	0.81			0.73				0.65					0.56					
95 % Confidence Interval	0.77–0.85			0.69–0.77				0.61–0.70					0.51–0.61					
No Information Rate (NIR)	0.50			0.33				0.25					0.20					
P-Value (Accuracy > NIR)	2.0E-16			2.2E-16				2.0E-16					2.00E-16					
Kappa	0.6183			0.5964				0.5397					0.4467					
Mcnemar's Test P-Value	1.0000			0.4664				0.6559					0.7801					
Statistics by Class																		
	R 1				R1	R2	R3		R1	R2	R3	R4		R1	R2	R3	R4	R5
Sensitivity	0.81				0.80	0.69	0.70		0.78	0.65	0.56	0.63		0.69	0.60	0.41	0.45	0.63
Specificity	0.81				0.90	0.80	0.89		0.92	0.87	0.85	0.89		0.94	0.88	0.88	0.88	0.86
Positive Predictive Value	0.81				0.81	0.64	0.76		0.77	0.63	0.56	0.66		0.73	0.56	0.47	0.49	0.53
Negative Predictive Value	0.81				0.90	0.84	0.86		0.93	0.88	0.85	0.88		0.92	0.90	0.86	0.87	0.90
Prevalence	0.50				0.33	0.33	0.33		0.25	0.25	0.25	0.25		0.20	0.20	0.20	0.20	0.20
Detection Rate	0.40				0.27	0.23	0.23		0.20	0.16	0.14	0.16		0.14	0.12	0.08	0.09	0.13
Detection Prevalence	0.50				0.33	0.36	0.31		0.25	0.26	0.25	0.24		0.19	0.22	0.18	0.18	0.23
Balanced Accuracy	0.81				0.85	0.75	0.79		0.85	0.76	0.71	0.76		0.82	0.74	0.65	0.67	0.75

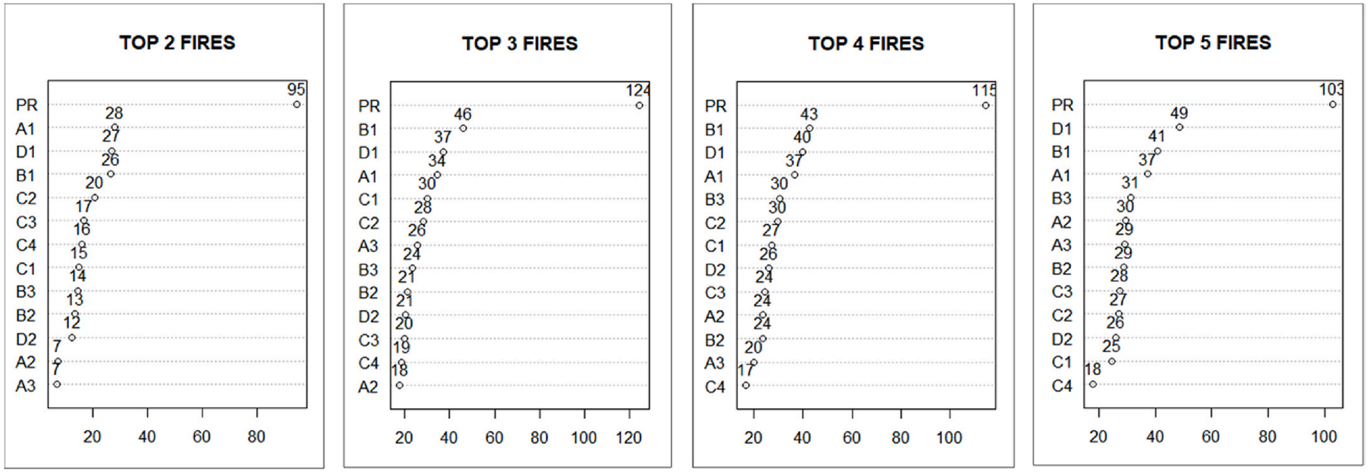

Fig. A15. Mean Decrease Accuracy (MDA) reported for different Random Forest models. The only difference, compared to Fig. 9, is the inclusion of the PR variable, which is the previous rank of the same fire on the closest preceding day.

Data availability

I have shared links to my data and source code

References

[1] S.L. Stephens, L.W. Ruth, Federal forest-fire policy in the United States, Ecol. Appl. 15 (2) (2005) 532–542.
[2] USDA, The Rising Cost of Fire Operations: Effects on the Forest Service's Non-fire Work, 2015.
[3] D.L. Martell, A review of operational research studies in forest fire management, Can. J. For. Res. 12 (2) (1982) 119–140.
[4] USDA, Interagency standards for fire and fire aviation operations. National interagency fire center: boise, ID. 2025, Available from: https://www.nifc.gov/sites/default/files/redbook-files/RedBook_Final.pdf.
[5] K. McConnell, E. Fussell, J. DeWaard, S. Whitaker, K.J. Curtis, L.St. Denis, et al., Rare and highly destructive wildfires drive human migration in the US, Nat. Commun. 15 (1) (2024) 6631.
[6] J.K. Balch, V. Iglesias, A.L. Mahood, M.C. Cook, C. Amaral, A. DeCastro, et al., The fastest-growing and most destructive fires in the US (2001 to 2020), Science 386 (6720) (2024) 425–431.
[7] K.M. Collins, O.F. Price, T.D. Penman, Suppression resource decisions are the dominant influence on containment of Australian forest and grass fires, J. Environ. Manag. 228 (2018) 373–382.
[8] M. Dimopoulou, I. Giannikos, Towards an integrated framework for forest fire control, Eur. J. Oper. Res. 152 (2) (2004) 476–486.
[9] M.P. Thompson, D.E. Calkin, Uncertainty and risk in wildland fire management: a review, J. Environ. Manag. 92 (8) (2011) 1895–1909.
[10] J.T. Abatzoglou, C.S. Juang, A.P. Williams, C.A. Kolden, A.L. Westerling, Increasing synchronous fire danger in forests of the western United States, Geophys. Res. Lett. 48 (2) (2021) e2020GL091377.
[11] S. McGinnis, L. Kessenich, L. Mearns, A. Cullen, H. Podschwit, M. Bukovsky, Future regional increases in simultaneous large Western USA wildfires, Int. J. Wildland Fire 32 (9) (2023) 1304–1314.
[12] D. Nguyen, E.J. Belval, Y. Wei, D.E. Calkin, A collection of incident prioritization matrices from different regions in the United States. https://doi.org/10.6084/m9.figshare.28890398.v1, 2025.
[13] G.H. Donovan, J.P. Prestemon, K. Gebert, The effect of newspaper coverage and political pressure on wildfire suppression costs, Soc. Nat. Resour. 24 (8) (2011) 785–798.
[14] D.E. Calkin, K.M. Gebert, J.G. Jones, R.P. Neilson, Forest service large fire area burned and suppression expenditure trends, 1970–2002, J. For. 103 (4) (2005) 179–183.
[15] K.M. Gebert, D.E. Calkin, J. Yoder, Estimating suppression expenditures for individual large wildland fires, West. J. Appl. For. 22 (3) (2007) 188–196.
[16] J. Liang, D.E. Calkin, K.M. Gebert, T.J. Venn, R.P. Silverstein, Factors influencing large wildland fire suppression expenditures, Int. J. Wildland Fire 17 (5) (2008) 650–659.
[17] R. MacMillan, L. Sun, S.W. Taylor, Modeling individual extended attack wildfire suppression expenditures in British Columbia, For. Sci. 68 (4) (2022) 376–388.
[18] M.S. Hand, M.P. Thompson, D.E. Calkin, Examining heterogeneity and wildfire management expenditures using spatially and temporally descriptive data, J. For. Econ. 22 (2016) 80–102.
[19] P. Baylis, J. Boomhower, Moral Hazard, Wildfires, and the Economic Incidence of Natural Disasters, National Bureau of Economic Research, 2019.
[20] D.E. Calkin, T. Venn, M. Wibbenmeyer, M.P. Thompson, Estimating US federal wildland fire managers' preferences toward competing strategic suppression objectives, Int. J. Wildland Fire 22 (2) (2012) 212–222.
[21] M. Hand, H. Katuwal, D.E. Calkin, M.P. Thompson, The influence of incident management teams on the deployment of wildfire suppression resources, Int. J. Wildland Fire 26 (7) (2017) 615–629.
[22] J. Bayham, J.K. Yoder, Resource allocation under fire, Land Econ. 96 (1) (2020) 92–110.
[23] Y. Wei, M.P. Thompson, E.J. Belval, D.E. Calkin, J. Bayham, Understand daily fire suppression resource ordering and assignment patterns by unsupervised learning, Machine Learning and Knowledge Extraction 3 (1) (2020) 14–33.
[24] A.J. Plantinga, R. Walsh, M. Wibbenmeyer, Priorities and effectiveness in wildfire management: evidence from fire spread in the western United States, Journal of the Association of Environmental and Resource Economists 9 (4) (2022) 603–639.
[25] D. Nguyen, E.J. Belval, Y. Wei, K.C. Short, D.E. Calkin, Dataset of United States Incident Management Situation Reports from 2007 to 2021, figshare, 2023, https://doi.org/10.6084/m9.figshare.24243184.v3.
[26] E. Kauffman, Climate and Topography, 12, Atlas of the Biodiversity of California, 2003, p. 15.
[27] J.E. Keeley, A.D. Syphard, Historical patterns of wildfire ignition sources in California ecosystems, Int. J. Wildland Fire 27 (12) (2018) 781–799.
[28] A.L. Westerling, B. Bryant, Climate change and wildfire in California, Clim. Change 87 (2008) 231–249.
[29] R. Rasker, Resolving the increasing risk from wildfires in the American West, Solutions 6 (2) (2015) 55–62.
[30] A. Miranda, J. Carrasco, M. González, C. Pais, A. Lara, A. Altamirano, et al., Evidence-based mapping of the wildland-urban interface to better identify human communities threatened by wildfires, Environ. Res. Lett. 15 (9) (2020) 094069.
[31] V.C. Radeloff, D.P. Helmers, H.A. Kramer, M.H. Mockrin, P.M. Alexandre, A. Bar-Massada, et al., Rapid growth of the US wildland-urban interface raises wildfire risk, Proc. Natl. Acad. Sci. 115 (13) (2018) 3314–3319.
[32] N.S. Diffenbaugh, D.L. Swain, D. Touma, Anthropogenic warming has increased drought risk in California, Proc. Natl. Acad. Sci. 112 (13) (2015) 3931–3936.
[33] A.P. Williams, R. Seager, J.T. Abatzoglou, B.I. Cook, J.E. Smerdon, E.R. Cook, Contribution of anthropogenic warming to California drought during 2012–2014, Geophys. Res. Lett. 42 (16) (2015) 6819–6828.
[34] S. Li, T. Banerjee, Spatial and temporal pattern of wildfires in California from 2000 to 2019, Sci. Rep. 11 (1) (2021) 8779.
[35] T. Brown, S. Leach, B. Wachter, B. Gardunio, The Northern California 2018 extreme fire season, Bull. Am. Meteorol. Soc. 101 (1) (2020) S1–S4.
[36] H.D. Safford, A.K. Paulson, Z.L. Steel, D.J. Young, R.B. Wayman, The 2020 California fire season: a year like no other, a return to the past or a harbinger of the future? Global Ecol. Biogeogr. 31 (10) (2022) 2005–2025.
[37] A.L. Westerling, Wildfire Simulations for California's Fourth Climate Change Assessment: Projecting Changes in Extreme Wildfire Events with a Warming Climate: a Report for California's Fourth Climate Change Assessment, California Energy Commission, Sacramento, CA, 2018.
[38] A.P. Williams, J.T. Abatzoglou, A. Gershunov, J. Guzman-Morales, D.A. Bishop, J.K. Balch, et al., Observed impacts of anthropogenic climate change on wildfire in California, Earths Future 7 (8) (2019) 892–910.
[39] J.H. Yoon, S.Y.S. Wang, R.R. Gillies, L. Hipps, B. Kravitz, P.J. Rasch, 2. Extreme fire season in California: a glimpse into the future? Bull. Am. Meteorol. Soc. 96 (12) (2015) S5–S9.
[40] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32.
[41] D. Singh, B. Singh, Investigating the impact of data normalization on classification performance, Appl. Soft Comput. 97 (2020) 105524.
[42] Jr FE. Harrell, Jr MFE. Harrell, Package 'Hmisc, 2019, ' CRAN2018, 2019, pp. 235–236.
[43] T. Wei, V. Simko, M. Levy, Y. Xie, Y. Jin, J. Zemla, Package 'corrplot, ' Statistician 56 (316) (2017) e24.
[44] P. Schober, C. Boer, L.A. Schwarte, Correlation coefficients: appropriate use and interpretation, Anesth. Analg. 126 (5) (2018) 1763–1768.
[45] S. RColorBrewer, M.A. Liaw, Package 'Randomforest.' University of California, Berkeley: Berkeley, CA, USA, 2018.
[46] M. Kuhn, Building predictive models in R using the caret package, J. Stat. Software 28 (2008) 1–26.
[47] J. Tuszynski, M.H. Khachatryan, Package 'Catools.', 2024.
[48] W. Yarberry, Dplyr, in: CRAN Recipes: DPLYR, Stringr, Lubridate, and Regex in R, Springer, 2021, pp. 1–58.

[49] P. Jain, S.C. Coogan, S.G. Subramanian, M. Crowley, S. Taylor, M.D. Flannigan, A review of machine learning applications in wildfire science and management, Environ. Rev. 28 (4) (2020) 478–505.
[50] P. Probst, A.L. Boulesteix, To tune or not to tune the number of trees in random forest, J. Mach. Learn. Res. 18 (181) (2018) 1–18.
[51] H. Wang, F. Yang, Z. Luo, An experimental study of the intrinsic stability of random forest variable importance measures, BMC Bioinf. 17 (2016) 1–18.
[52] H. Han, X. Guo, H. Yu, Variable selection using mean decrease accuracy and mean decrease gini based on random forest, IEEE (2016) 219–224.
[53] G. Louppe, L. Wehenkel, A. Sutera, P. Geurts, Understanding variable importances in forests of randomized trees, Adv. Neural Inf. Process. Syst. 26 (2013).
[54] M.A. Moritz, E. Batllori, R.A. Bradstock, A.M. Gill, J. Handmer, P.F. Hessburg, et al., Learning to coexist with wildfire, Nature 515 (7525) (2014) 58–66.