scientific reports

OPEN

Benchmarking performance of annual burn probability modeling against subsequent wildfire activity in California

Christopher J. Moran, Matthew P. Thompson✉, Bryce A. Young, Joe H. Scott & Melissa R. Jaffe

Wildfire simulation is deployed extensively to support risk management, and in the US has driven billions in federal investment. Foundational to strategic risk analysis is spatial information on the likelihood of burning in a fire year, typically provided by burn probability (BP) models. The recency of BP maps is a key driver of their accuracy, especially in disturbed landscapes that have experienced changes in fire spread potential. Few published examples exist comparing BP values against subsequent fire activity, and none to our knowledge evaluate annually updated BP maps. Here, we present a novel performance evaluation of the operational wildfire simulation system FSim, confronting updated BP maps with subsequent fire activity across the state of California over a 4-year period (2020–2023). Results show strong predictive ability: across 5 equal-area BP classes, 56.7–79.8% of the burned area occurred in the top 20% of mapped area; mean (median) BP values in burned areas were 238.5–348.8% (551.4–880.7%) greater than in unburned areas; differences in empirical cumulative distribution functions of BP for burned/unburned areas were statistically significant; Logarithmic Skill Scores ranged from – 0.072 to 0.389 against two reference models. Findings indicate reliable forecast performance and useful application of up-to-date BP maps, critical to support ongoing wildfire risk mitigation.

Keywords Wildfire, Hazard, Risk, Simulation, Forecast verification, Decision support

In the US, the risks and complexities of wildfire have magnified dramatically, with changing socioecological fire regimes, expanding development in fire-prone areas, faster rates of fire growth, more destructive impacts to communities and landscapes, and escalating costs[1–5]. This is consistent with global patterns of environmental change and human activity that are interacting to increase likelihood of extreme wildfire events with grave implications for ecosystem services and public health[6–9]. Continued reliance on wildfire suppression as the dominant management action may exacerbate future hazard while increasing strains on an incident response system already challenged by increasing synchronous fire danger and resource scarcity[10–13]. Hence growing recognition of the need for more proactive approaches to wildfire management as well as risk and decision analytic frameworks to support them[14–18].

A foundational need for proactive risk management is information on the likelihood of experiencing wildfire in any given location over a given time horizon. Likelihood can vary by orders of magnitude across landscapes and is therefore a significant driver of hazard and exposure as well as prioritization and efficacy[19]. Emergency managers, ignition prevention programs, hazardous fuels reduction programs, community hardening programs, land use planners, water and power utilities, and insurance providers, among others, all need reliable information on wildfire likelihood.

This information is typically provided through burn probability (BP) modeling, now widely used for operational and research purposes[20–23]. BP models differ by the intended application. Generally, all share the same workflow of iteratively simulating fire spread across landscapes while accounting for variability in factors such as ignition patterns and weather. Developing gridded BP datasets involves simulating multiple fire events under a range of conditions and calculating BP values as the number of times a given cell burns divided by the number of simulations. Some models are intended for near real-time incident support[24], where ignitions are known, and generate conditional BPs corresponding to patterns of potential fire spread under forecasted weather

Pyrologix, Vibrant Planet, Missoula, USA. ✉email: matt.thompson@pyrologix.com

over a defined planning horizon. Some models capture uncertainty in both ignition location and weather, and generate conditional BPs given user-defined scenarios[25]. Some models also account for uncertainty in ignition likelihood and generate annualized BPs that probabilistically simulate ignition location and timing as inputs into fire spread simulations[26].

BP modeling is deployed extensively in the US for a range of decision support applications and products, particularly for land management agencies like the USDA Forest Service[27–31]. Our focus here is on FSim, a comprehensive large fire simulation system that generates tens of thousands of synthetic fire years and models fire occurrence and spread while accounting for stochasticity in ignitions, weather, and spotting[26]. One of its primary outputs is spatially gridded data of annualized BP values (120–270 m) that reflect average large fire potential based on input landscape conditions. Updating fuel conditions to account for large disturbance, as is done annually here, is intended to reflect dynamic spatial patterns of fire likelihood.

Although these results can be used to estimate expected annual area burned—as the sum-product of all burnable pixels and their corresponding BP values—this is not our primary focus as it does not directly lead to targeted, localized intervention; rather our focus is on FSim's ability to spatially discriminate wildfire likelihood. We focus on FSim because its results have been integrated into multiple strategic assessment and planning efforts and informed budgetary allocations on the scale of billions. Notable nationwide examples leveraging FSim outputs include the National Cohesive Wildland Fire Management Strategy[27], the Wildfire Crisis Strategy[32], the Wildfire Risk to Communities project[33], and the FEMA National Risk Index[34]. FSim results have also been widely deployed to support localized fuels management and operational response strategies[35–37].

Despite growing use and sophistication of BP models, the wildfire science community has not widely adopted probabilistic forecast verification as a common practice. Few published examples exist comparing BP values against subsequent fire activity[38–41]. There are many challenges to measuring predictive accuracy of BP models, including insufficient or inaccurate fire observation data, very low BP values in some locations (i.e., < 1 in 1,000), limited ability to model how suppression alters large fire spread patterns, and, critically, changing landscape conditions[29,42,43].

For spatially discriminating areas more or less likely to burn to support targeted intervention, accurately reflecting dynamic landscape conditions is essential. A longer post-simulation evaluation period enables better capture of interannual variability in fire activity, but the original BP map becomes increasingly outdated the longer the evaluation period extends. By contrast, more frequently updating BP maps enables better capture of interannual variability in the landscape's ability to support or inhibit large fire growth, particularly salient in recently disturbed areas where prior fire activity can lead to self-limiting behavior[44].

Recent work evaluating the performance of FSim BP maps across the conterminous US—based on a static 2014 landscape and contrasted against observed wildfires from 2016 to 2022—reported a moderate correlation of mean BP with observed burned area but underprediction of burned areas in some regions[40]. The authors suggest that discrepancies may have stemmed from an outdated fuels layer and argued that frequent updates to BP maps could improve their usefulness. Here, motivated in part by these recent findings, we present a novel performance evaluation of FSim, confronting annually updated BP maps with subsequent fire activity across the state of California over a 4-year period (2020–2023).

California presents a compelling and challenging use case—it contains diverse landscapes with high fuel loads influenced by fire exclusion and tree mortality and evolving fire regimes driven by human ignitions and extreme wind events, and has significant attention and need to proactively address growing wildfire risks[45–54]. California can also have significant interannual variability in wildfire activity. In 2020, the state set historical records for total area burned and in 2021 experienced the largest single fire in state history (the Dixie Fire), with significant impacts to communities and landscapes[15,55–57]. The fire years of 2022 and 2023 by contrast burned only ~ 10% of the amount burned in 2020 and 2021.

Our analysis is built on four pairwise comparisons of annual BP and fire activity, with each subsequent BP map updated in response to the prior year's disturbances. Informed by prior work, we examine how the observed area burned varies with BP values, compare distributions of BP values inside and outside of observed burned areas, and calculate logarithmic skill scores[39,40,58]. We establish performance benchmarks based on percentiles of observed area burned and BP area mapped, discuss interpretations and risk management implications, and offer future directions in probabilistic forecast verification for wildfire modeling.

Results

Observed area burned and burn probabilities

Maps of simulated BP values (Fig. 1) reveal areas of higher fire likelihood throughout the Sierra Nevada, Northern and Southern Coast Ranges, Klamath Mountains, Transverse Ranges, and Peninsular Ranges, with lower likelihood in the Central Valley and Mojave Desert areas. Overlaid observed fire perimeters generally align with hotspots of high BP, particularly in the Northern Coast Range and Klamath Mountains. Interannual variability in spatial patterns of BP is evident in the wake of significant wildfires, particularly notable in the northwest portion of the state with widespread reductions in BP after the 2020 and 2021 fire years (Fig. 2). Within areas burned in 2020, the pre-burn mean BP was 0.054, reduced to 0.024 post-burn, a 56.19% reduction (median BP was reduced by 66.23%). Within areas burned in 2021, the pre-burn mean BP was 0.058, reduced to 0.018, a 69.2% reduction (median BP was reduced by 73.32%).

Proportional comparisons of expected area burned (eAB) and observed area burned (oAB) reveal aligned patterns across five equal area BP classes (Fig. 3). BP class definitions vary by year; exact delineations of the BP bins are provided in supplementary materials. The general interpretation is that the model effectively predicts that most area burns in the higher BP classes, with slightly more area burned in the Medium and Low-Medium than predicted. The proportion of oAB in the Med-High or High BP classes was 90.3% in 2020, 80.9% in 2021, 93.6% in 2022, and 65.7% in 2023, notably higher than the ~ 40% that would result if area burned followed a

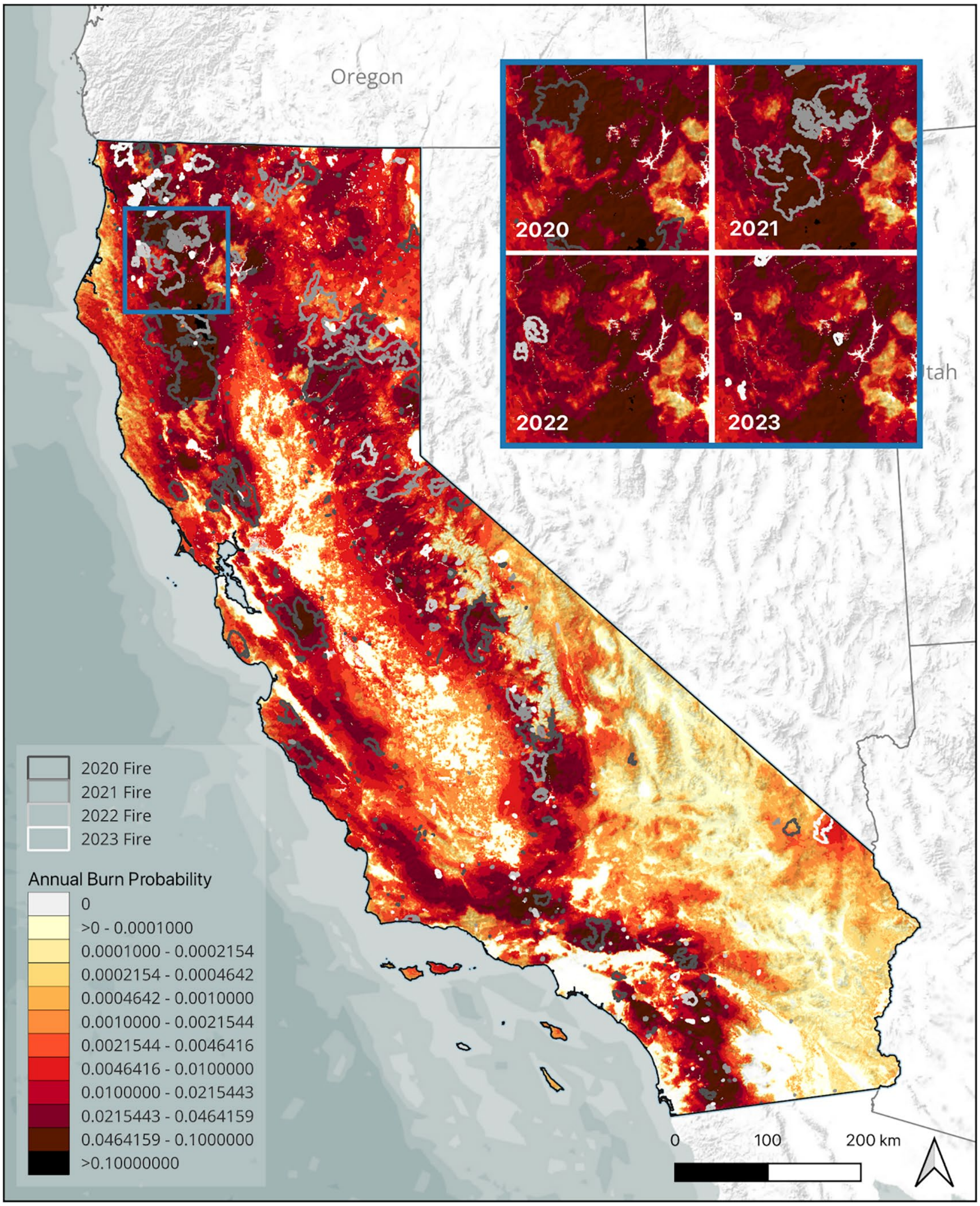

Fig. 1. Burn probability values and observed fire perimeters. The state map shows 2020 BP and 2020–2023 fire perimeters. Inset panels contain BP and fire perimeters for respective years. The chronological inset map shows burned areas that are accounted for in subsequent BP maps. Map created in QGIS 3.34 using Esri World Terrain Base © Esri, USGS, NOAA. State boundaries from U.S. Census Bureau TIGER/Line® Shapefiles (public domain).

random distribution across equal-area BP classes. Predicted proportions of eAB for the Med-High and High BP classes closely matched oAB values (90.1%, 89.7%, 89.7%, and 89.6%), except for 2023 where less area burned in the Med-High and more burned in the Med BP classes. The High BP class alone accounted for 69.2%, 56.8%, 79.8%, and 56.7% of oAB across respective years, varying from eAB (67.1%, 66.4%, 66.4%, and 66.4%). The mean observed area burned for all 4 years in the High BP class was similar to the expected area burned (65.6% vs 66.6%). The proportion of observed area burned in the Low and Low-Med classes was 2.3%, 5.8%, 3.4%, and 9.1% across respective years. These values are greater than expected area burned (2.1%, 2.2%, 2.3%, and 2.3%). The mean observed area burned of all 4 years in Low and Low-Med classes exceeded expected area burned (5.1% vs 2.2%).

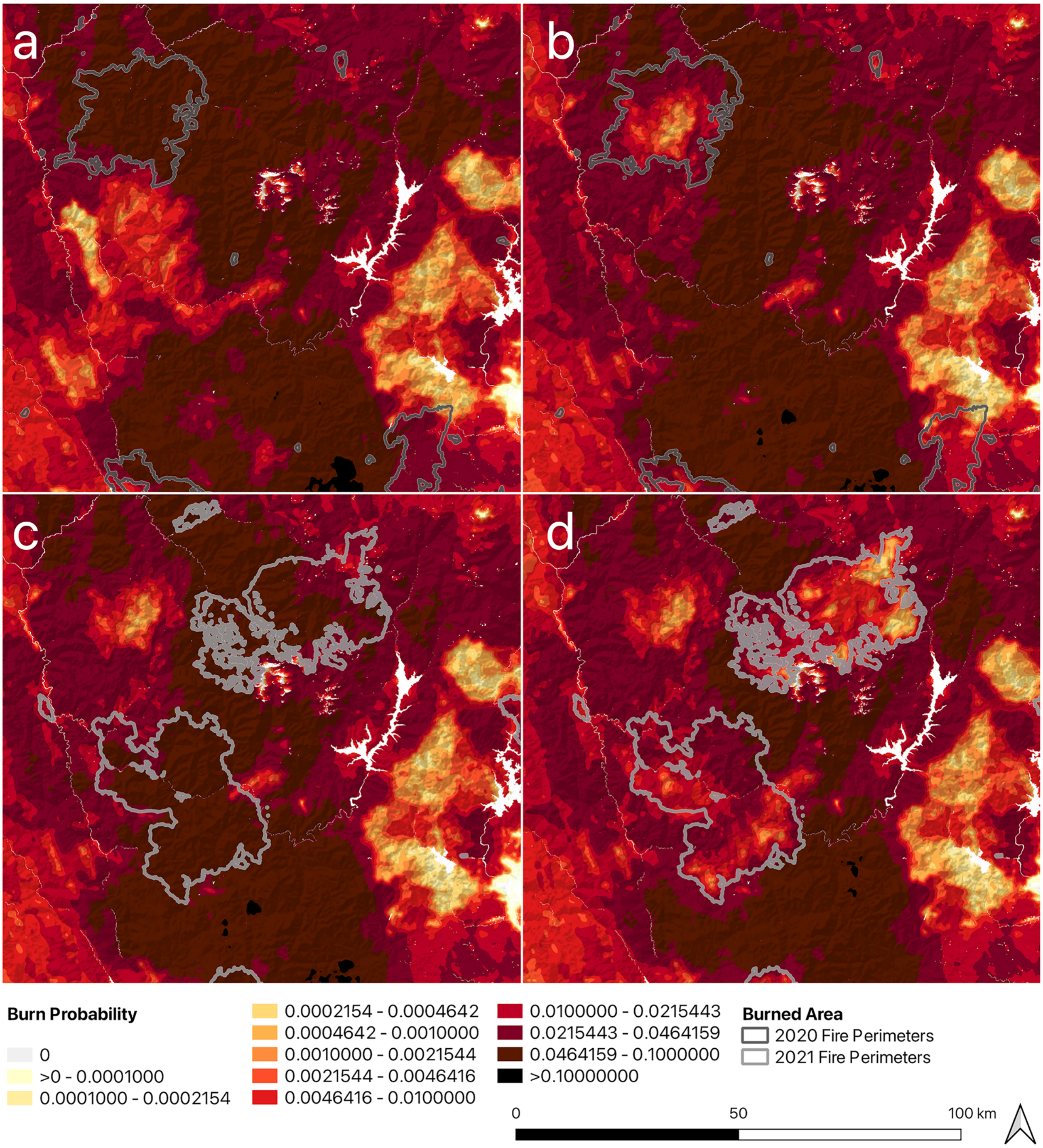

Fig. 2. Interannual variability in simulated BP due to prior large fire disturbance; location matches inset map from Fig. 1. Panels (**a**) and (**b**) show observed 2020 wildfires with pre-and post-simulation BPs (2020 and 2021 BP maps), respectively. Panels (**c**) and (**d**) show observed 2021 wildfires with pre-and post-simulation BPs (2021 and 2022 BP maps), respectively. The map reflects both the impact of wildfires prior to 2020 as well as wildfires in 2020 and 2021 on subsequent fire spread potential and BP values. Map created in QGIS 3.34 using Esri World Terrain Base © Esri, USGS, NOAA.

Burn probabilities inside and outside of burned areas

Mean and median BP values were greater within burned areas than outside (Fig. 4). While distributions of BP values inside burned areas showed variability across years, BP values outside of burned areas were consistently concentrated near zero. In 2020, mean BP values in burned areas were 348.78% greater than unburned and median BP values in burned areas were 880.67% greater. In 2021, mean BP values in burned areas were 292.43% greater than unburned and median BP values in burned areas were 551.44% greater. In 2022, mean BP values in burned areas were 271.70% greater than unburned and median BP values in burned areas were 664.66% greater.

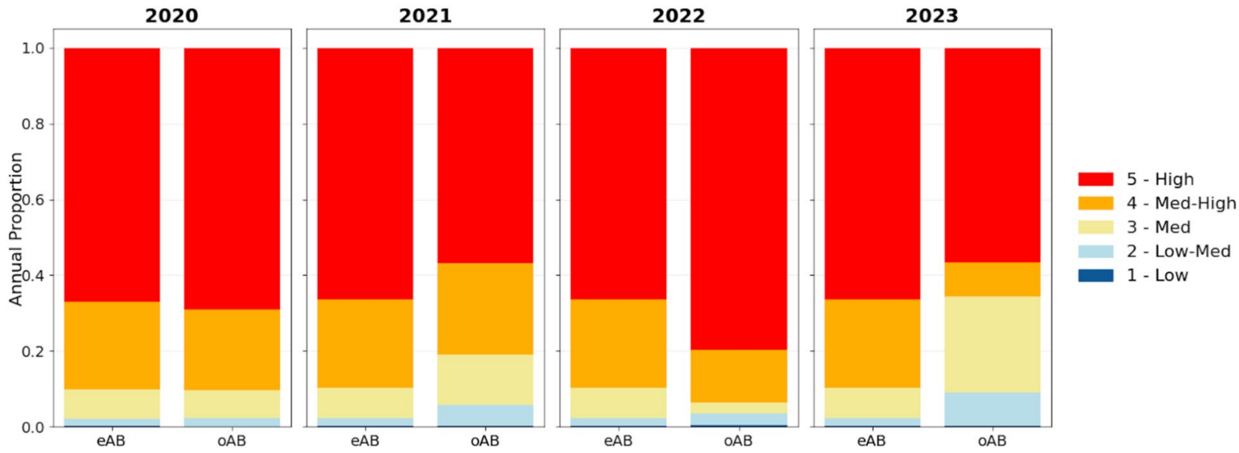

Fig. 3. Proportional comparison of expected area burned (eAB) and observed area burned (oAB) across five equal area BP classes. BP class definitions vary by year; exact delineations of the BP bins are provided in supplementary materials.

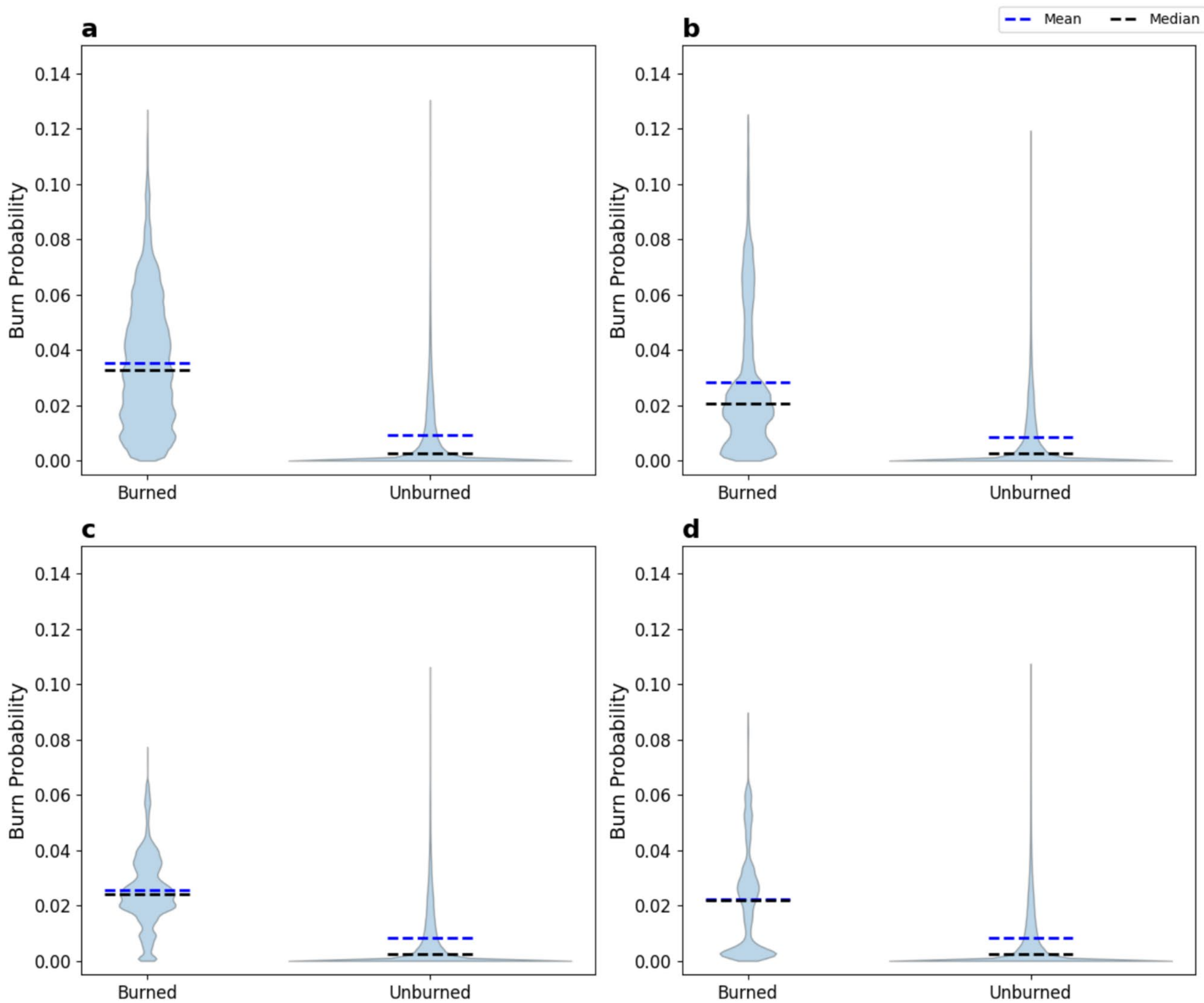

Fig. 4. Violin plots comparing the distribution of burn probability values for burned and unburned areas. The horizontal lines represent the median (black) and mean (blue) within each respective class. Panels (**a**–**d**) correspond to years 2020–2023, respectively.

In 2023, mean BP values in burned areas were 238.53% greater than unburned and median BP values in burned areas were 599.16% greater.

Cumulative distribution plots comparing BP values for burned and unburned areas confirm consistent patterns across years of higher BPs in burned areas (Fig. 5). Kolmogorov-Smirnoff tests on a 0.01% random sample of burned/unburned areas yielded statistically significant differences for all 4 years ($p < 0.01$) with test statistic values of 0.5447 (2020), 0.4248 (2021), 0.6136 (2022), and 0.3738 (2023). In 2020, 95% of burned area occurred at or above the 50th percentile of BP values and 51% occurred above the 90th percentile. In 2021, 88% of burned area occurred at or above the 50th percentile and 33% above the 90th percentile. In 2022, 96% of burned area occurred above the 50th percentile and 38% above the 90th percentile. In 2023, 74% of burned area occurred above the 50th percentile and 36% above the 90th percentile.

Logarithmic skill scores

Logarithmic Skill Scores (LSS; 58) were computed to provide a quantitative measure of forecast quality using two different reference logarithmic scores as the baselines for deriving the skill score, two different sets of data (burned only and all cells), and across all four years totaling 16 LSS estimates (Table 1). Mean LSS were 0.276

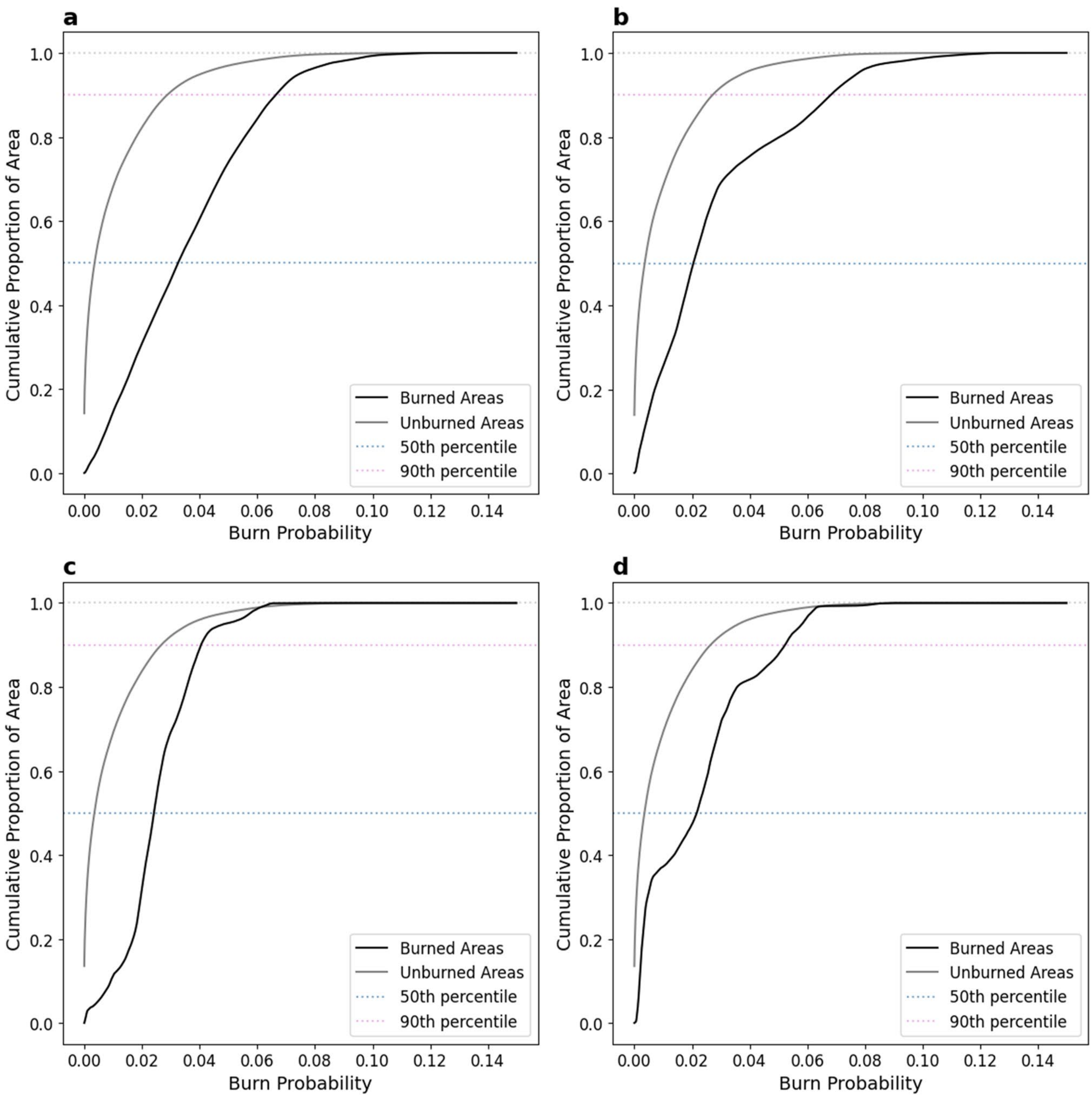

Fig. 5. Cumulative distribution plots comparing burn probability values in burned versus unburned areas. Panels (**a**–**d**) correspond to years 2020–2023, respectively.

	Log Skill Score		Log Skill Score		Mean Log Score		Mean Log Score		Mean Log Score	
	15-year mean BP Reference		Resampled BP Reference		Predicted BP		15-year mean BP Reference		Resampled BP Reference	
Year	Burned Cells	All Cells	Burned Cells	All Cells	Burned Cells	All Cells	Burned Cells	All Cells	Burned Cells	All Cells
2020	0.338	0.308	0.389	0.373	3.643	0.169	5.504	0.245	5.938	0.270
2021	0.264	0.221	0.326	0.310	4.048	0.118	5.504	0.151	6.008	0.171
2022	0.296	−0.003	0.361	0.246	3.872	0.022	5.504	0.022	6.062	0.030
2023	0.203	−0.072	0.278	0.191	4.384	0.024	5.504	0.022	6.070	0.030

Table 1. Logarithmic skill scores of the BP results using two reference models: (1) The 15-year reference model is a spatially-constant, naive model that represents the 2006–2020 mean annual acreage burned and (2) the resampled BP spatially randomizes predicted BP values. The mean logarithmic scores used to derive the skill scores are also reported.

and 0.339 from 2020 to 2023 for burned cells using the naive reference model and the resampled reference model, respectively, which can be interpreted as the percentage improvements over these reference models. The LSS drops when both burned and unburned cells are considered with means of 0.114 and 0.280 against the naive and resampled reference models, respectively (Table 1). The decrease can largely be attributed to two years of negative skill in 2022 and 2023 for the all-cells LSS. The skill score is higher (0.296) for burned cells in 2022 compared to 2021 (0.264) even though the all-cells skill score is negative in 2022, a result from a well-below average year in total area burned but demonstrated skill in predicting the spatial location of the areas that did burn (c.f. Figure 2).

To specifically test the value of annual updates, the LSS (the version derived from the 15-year mean reference model) were calculated using the 2020 BP predictions in comparison to the subsequent years (2021–2023) and compared to the LSS derived from the BP in matched years. LSS improved in five of the six comparisons by a mean of 0.028 (42%)–indicating superior performance of up-to-date BP maps relative to the continued use of the 2020 BP predictions for multiple years.

Discussion

Results generally showed strong predictive performance of annual BP maps as evaluated through multiple comparisons: markedly greater proportions of burned areas in higher BP bins; higher mean and median BPs in burned than unburned areas; statistically significant differences in cumulative distributions of BP between burned areas and unburned areas; and Logarithmic Skill Scores (LSS) for BP maps indicating outperformance of reference models. A general performance benchmark of observing 57–80% of area burned in the top ~ 20% of mapped BP area was established.

In addition to the general performance benchmark, the importance of interannual updates was evident. The largest cause of year-to-year model differences were from previous year disturbances represented in the fuel inputs, the most impactful of those being wildfire scars. We observed a mean percent reduction in BP of 69% in the following years (2021–2023) within the wildfire perimeters occurring in the study time frame (2020–2022). The model represented beyond -perimeter effects with a mean percent reduction in BP of 15% in a 2 km buffer from these wildfires. The BP maps responded on a landscape-level as well with total mean reductions in BP of 10% and 7% following the large fire years of 2020 and 2021, respectively, and only a 1% reduction in BP following the mild fire year in 2022. The average improvement of 42% in the LSS using updated inputs over relying on the increasingly-outdated 2020 BP map also highlights the added value of interannual updating.

Some relevant prior work evaluates the performance of conditional BP maps generated for individual wildfire events with known ignition locations, which significantly shrinks the uncertainty relative to simulating thousands of potential fire years and millions of potential events. Paz et al.[41] established a similar performance benchmark for a single wildfire, reporting that 87% of observed burned area occurred in the top 30% of mapped BP area, which is within the range of results we present. Allaire et al.[38] calculated Brier Skill Scores against a posteriori reference models based on the final footprints of seven observed fires. Although not directly comparable, our LSS were within the ranges they present (−9.986—0.352) and generally closer to their best performing ensemble.

In two cases our LSS were negative—for low fire activity years (2022–2023) and when considering BP forecasts for all cells (not just burned cells) using the 15-year mean BP reference. This is largely due to the large number of higher BP predictions with a relatively low level of positive cases (burned cells). For these years, the analysis still shows model skill in all three analyses. For burned cells only, the LSS was generally similar to other years. The spatially randomized BP reference model intended on evaluating the spatial skill alone given the exact same distribution of BP was present and the corresponding LSS were 0.246 for 2022 and 0.191 for 2023. This highlights the difficulty in predicting burn probability in systems that can exhibit dramatic interannual variability (e.g., ~ 10% of area burned in 2022 and 2023 relative to 2020 and 2021). Analysis on BP modeling skill scores could be expanded to explore how they vary with fire activity or alternative reference models.

More directly relevant work includes evaluation of landscape- scale BP maps across larger regions with multiple years of fire observations, but without annual updates to fuels after disturbance. Evaluating simulation modeling in Alberta, Canada, Beverly and McLaughlin[39] reported that distributions of BP in burned areas were not heavily skewed towards higher BP values, but that most of the burned area (75–80%) occurred in the top 50th percentile of mapped BP. Evaluating simulation modeling for the conterminous US, Carlson et al.[40] established that 68% of the observed burned area occurred in the top 40% of mapped BP area, and for Mediterranean California that 62% of observed area burned occurred in the top 40% of mapped BP area. The performance benchmark discrepancy between[40] and that established here could be attributed to several factors, including using more up-to-date and customized fuel input layers, more recent historical fire weather and occurrence data, differences and improvements in modeling methods and calibration over time, or simply narrowing the scope of analysis to a more fire prone region. The present analysis encompasses landscape scale evaluation of BP maps over multiple years while resolving issues of large disturbance effects through annual updates and generally indicates the strongest BP modeling performance published to date.

Several extensions and improvements are apparent. First, researchers could pursue validation across different geographic areas and across longer time horizons, especially in areas with comparatively less frequent fire activity, or for different BP systems used in operational decision support[27,59]. Second, validation could be performed on other FSim model outputs such as perimeters, fire progression maps, or magnitude of daily spread events[60–63]. Other efforts could focus on continued model improvement. For example, better information on fuel break and suppression effectiveness could improve containment modeling[48,64,65]. While temporal trends in fire occurrence were utilized in calibration, the reliance on historical data results in climate change not being directly considered and previous work has developed approaches for its incorporation in more out-year probability estimates[66].

It is likely that application of machine learning-based approaches for BP prediction will increase[67–70], highlighting a need for comparative validation and use case exploration. Costa-Saura et al.[71] contrasted BP interpretations from fire spread and random forest models and noted that firefighting decision making might be more tightly linked with fire spread information. Applications of process-based fire spread simulation will persist, particularly for accompanying analyses based on the underlying simulated perimeters to analyze transmission into communities or watershed impacts[14,72]. This highlights the broader needs of accounting for user needs and experience with using risk-based information for decision support[31,73].

In summary, the primary implication is that managers can have confidence in using well-calibrated BP maps to support ongoing wildfire risk assessment and planning in California and other fire prone regions, while sharing the sentiments expressed in Beverly and McLaughlin[39] and Carlson et al.[40] that caution is warranted. In fact, with growing wildfire activity and more communities at risk the use of BP modeling will likely increase, and rigorous evaluation of these predictions will increase confidence when prioritizing resources for prevention and mitigation. We hope this effort can catalyze broader collaborative efforts in probabilistic wildfire forecast verification with FSim and other BP models.

Methods

We generated annual burn probability (BP) maps for each year from 2020 to 2023 using a customized version of the Large Fire Simulation Model (FSim; version 1.0.9) in projects for the California Department of Forestry and Fire Protection. FSim is a stochastic , iterative model that simulates plausible fire ignition, growth, intensity, suppression, and containment throughout hypothetical years[26,74]. FSIM is developed by the United States Forest Service Missoula Fire Science Laboratory utilizing the foundational surface spread model[75], crown fire spread and initiation models[76,77] and methods for their spatially-explicit application[78,79]. Each iteration in the model represents an entire year and given enough iterations is intended to capture the full range of potential outcomes for a given set of inputs along with estimates of central tendency and variability.

Simulations were performed individually for unique pyromes, i.e., areas of relatively homogenous contemporary vegetation, climate, and fire regimes built to support landscape fire simulation[80]. Each pyrome was an individual modeling domain with unique calibration parameters and a 30 km buffer to allow ignitions to spread outside pyrome boundaries preventing edge effects. The BP estimates from each domain were then mosaicked by summing the overlapping buffer areas after normalizing by iteration count, which were a minimum of 10,000 per domain. Input parameters and outputs were reviewed in 2020 by fire experts in California and calibrated against historical fire occurrence data using best practices[81]. Input data were updated annually in response to observed disturbances based on fire observation and fuel treatment data and expert-defined fuel calibration rulesets that follow the general framework developed in the LANDFIRE program[82]. Calibration targets for fire size and occurrence distributions were developed using ordinary least squares regression over the prior 15 years of fire activity; more details on calibration are available in[29].

We use a historical fire perimeters package to determine whether an area was burned by a fire in our given time frame (2020–2023). To create this package, we merged Welty and Jeffries[83] perimeters with National Intergency Fire Center perimeters and clipped fires that burned only partially within California to the California boundary. We filtered the resulting geometries for large fires (> 100 acres) and turned them into binary rasters at a 30 m resolution. We excluded non-burnable land cover from the analysis as defined by LANDFIRE.gov, except where BP was greater than zero to retain urban and agricultural areas classified as burnable following methods in the Wildfire Risk to Communities project[84]. This method of ignoring non-burnable features such as water, barren, rock, and ice ensured that we did not give our BP models credit for predicting that non-burnable features would not burn.

We employ three methods to evaluate BP modeling performance results against subsequent annual fire activity. First, we compare the proportional values of expected area burned (eAB) and observed area burned (oAB) across five equal-area BP bins. To quantify eAB, we multiply the total land area mapped in each BP bin by the mean BP of that bin. Second, we compare mean and median BP values within and outside of burned areas using violin plots and compare cumulative distribution functions of burned and unburned areas by BP with Kolmogorov–Smirnov tests[85].

Third, we calculate logarithmic scores (LS; 58) and derive l ogarithmic s kill s cores (LSS) using two reference models: one based on the previous fire activity produced by calculating the constant BP value that would generate the expected area burned that matched the latest 15-year mean area burned (2006–2020) in the Fire Occurrence

Database[86], and another based on spatially randomizing the predicted BP values across the entire domain. The former represents a standard, naive reference analogous to a climatological average in weather forecasting skill scoring and the latter a reference isolating the spatial skill by having the same distribution of BP values across the study area. We iteratively tested the variability in LSS following BP spatial randomization and found minimal variation in the LSS values. The LS followed the binary formulation as follows:

$$LS = \frac{1}{N} \sum\nolimits_{i=1}^{N} - \ln\left(1 - |p_i - o_i|\right) \tag{1}$$

where i is the probability at cell i and i is the binary outcome burned (1) or not burned (0) and N is the number of cells. The LSS is then:

$$LSS = 1 - \frac{LS}{LS_{ref}} \tag{2}$$

where ref is either the 15-year mean BP or the spatially randomized mean BP models. LSS are chosen over the ubiquitous Brier Skill scores (e.g., 38) to account for the rarity of wildfire occurrence, a known limitation[58].

Data availability

Burn probability and perimeter data are hosted at the Open Science Framework: https://osf.io/z6gnt/. Moran, Christopher J, M.P. Thompson, B.A. Young, J.H. Scott, M.R. Jaffe. 2025. "2020-2023 Raster Data for Benchmarking Performance of Annual Burn Probability Modeling against Subsequent Wildfire Activity in California." OSF. May 15.. doi:10.17605/OSF.IO/Z6GNT

Received: 27 February 2025; Accepted: 18 June 2025
Published online: 03 July 2025

References

1. Balch, J. K. et al. The fastest-growing and most destructive fires in the US (2001 to 2020). *Science* **386**(6720), 425–431 (2024).
2. Downing, W. M., Dunn, C. J., Thompson, M. P., Caggiano, M. D. & Short, K. C. Human ignitions on private lands drive USFS cross-boundary wildfire transmission and community impacts in the western US. *Sci. Rep.* **12**(1), 2624 (2022).
3. Higuera, P. E. et al. Shifting social-ecological fire regimes explain increasing structure loss from Western wildfires. *PNAS nexus* **2**(3), pgad005 (2023).
4. Radeloff, V. C. et al. Rapid growth of the US wildland-urban interface raises wildfire risk. *Proc. Natl. Acad. Sci.* **115**(13), 3314–3319 (2018).
5. Turco, M. et al. Anthropogenic climate change impacts exacerbate summer forest fires in California. *Proc. Natl. Acad. Sci.* **120**(25), e2213815120 (2023).
6. Aguilera, R., Corringham, T., Gershunov, A. & Benmarhnia, T. Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California. *Nat. Commun.* **12**(1), 1493 (2021).
7. Bowman, D. M. et al. Human exposure and sensitivity to globally extreme wildfire events. *Nat. Ecol. Evolut.* **1**(3), 0058 (2017).
8. Cunningham, C. X., Williamson, G. J. & Bowman, D. M. Increasing frequency and intensity of the most extreme wildfires on Earth. *Nat. Ecol. Evol.* **8**(8), 1420–1425 (2024).
9. Guo, Y., Wang, J., Ge, Y. & Zhou, C. Global expansion of wildland-urban interface intensifies human exposure to wildfire risk in the 21st century. *Sci. Adv.* **10**(45), eado9587 (2024).
10. Abatzoglou, J. T., Juang, C. S., Williams, A. P., Kolden, C. A. & Westerling, A. L. Increasing synchronous fire danger in forests of the western United States. *Geophys. Res. Lett.* **48**(2), e2020GL091377 (2021).
11. Cullen, A. C., Goldgeier, B. R., Belval, E. & Abatzoglou, J. T. Characterising ignition precursors associated with high levels of deployment of wildland fire personnel. *Int. J. Wildland Fire* https://doi.org/10.1071/WF23182 (2024).
12. Kreider, M. R. et al. Fire suppression makes wildfires more severe and accentuates impacts of climate change and fuel accumulation. *Nat. Commun.* **15**(1), 2412 (2024).
13. Thompson, M. P. et al. Wildfire response: a system on the brink?. *J. Forest.* **121**(2), 121–124 (2023).
14. Ager, A. A. et al. Predicting Paradise: Modeling future wildfire disasters in the western US. *Sci. Total Environ.* **784**, 147057 (2021).
15. Iglesias, V. et al. Fires that matter: reconceptualizing fire risk to include interactions between humans and the natural environment. *Environ. Res. Lett.* **17**(4), 045014 (2022).
16. Hessburg, P. F., Prichard, S. J., Hagmann, R. K., Povak, N. A. & Lake, F. K. Wildfire and climate change adaptation of western North American forests: a case for intentional management. *Ecol. Appl.* **31**(8), e02432 (2021).
17. Thompson, M. P. et al. Potential operational delineations: new horizons for proactive, risk-informed strategic land and fire management. *Fire Ecology* **18**(1), 17 (2022).
18. United Nations Environment Programme. Spreading like Wildfire—The Rising Threat of Extraordinary Landscape Fires. A UNEP Rapid Response Assessment. Nairobi (2022).
19. Barros, A. M., Ager, A. A., Day, M. A. & Palaiologou, P. Improving long-term fuel treatment effectiveness in the National Forest System through quantitative prioritization. *For. Ecol. Manag.* **433**, 514–527 (2019).
20. Alcasena, F., Ager, A. A., Belavenutti, P., Krawchuk, M. & Day, M. A. Contrasting the efficiency of landscape versus community protection fuel treatment strategies to reduce wildfire exposure and risk. *J. Environ. Manag.* **309**, 114650 (2022).
21. Kearns, E. J. et al. The construction of probabilistic wildfire risk estimates for individual real estate parcels for the contiguous United States. *Fire* **5**(4), 117 (2022).
22. Parisien, M. A., Dawe, D. A., Miller, C., Stockdale, C. A. & Armitage, O. B. Applications of simulation-based burn probability modelling: a review. *Int. J. Wildland Fire* **28**(12), 913–926 (2019).
23. Erni, S. et al. Mapping wildfire hazard, vulnerability, and risk to Canadian communities. *Int. J. Disaster Risk Reduct.* **101**, 104221 (2024).
24. Finney, M. A. et al. A method for ensemble wildland fire simulation. *Environ. Model. Assess.* **16**, 153–167 (2011).
25. Finney, M. A. An overview of FlamMap fire modeling capabilities. In In: Andrews, Patricia L.; Butler, Bret W., comps. 2006. Fuels Management-How to Measure Success: Conference Proceedings. 28–30 March 2006; Portland, OR. Proceedings RMRS-P-41. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station, **41**, 213–220 (2006).
26. Finney, M. A., McHugh, C. W., Grenfell, I. C., Riley, K. L. & Short, K. C. A simulation of probabilistic wildfire risk components for the continental United States. *Stoch. Env. Res. Risk Assess.* **25**, 973–1000 (2011).

27. Calkin, D. E., Thompson, M. P., Finney, M. A. & Hyde, K. D. A real-time risk assessment tool supporting wildland fire decisionmaking. *J. Forest.* **109**(5), 274–280 (2011).
28. Calkin, D. E., O'Connor, C. D., Thompson, M. P. & Stratton, R. D. Strategic wildfire response decision support and the risk management assistance program. *Forests* **12**(10), 1407 (2021).
29. Dillon, G. K., Scott, J. H., Jaffe, M. R., Olszewski, J. H., Vogler, K. C., Finney, M. A., Short, K. C., Riley, K. L., Grenfell, I. C., Jolly, W. M., & Brittain, S. Spatial datasets of probabilistic wildfire risk components for the United States (270m). 3rd Edition. Fort Collins, CO: Forest Service Research Data Archive. (2023). https://doi.org/10.2737/RDS-2016-0034-3.
30. Drury, S. A., Rauscher, H. M., Banwell, E. M., Huang, S. & Lavezzo, T. L. The interagency fuels treatment decision support system: functionality for fuels treatment planning. *Fire Ecol.* **12**, 103–123 (2016).
31. O'Mara, T., Meador, A. S., Colavito, M., Waltz, A. & Barton, E. Navigating the evolving landscape of wildfire management: a systematic review of decision support tools. *Trees For. People* **16**, 100575 (2024).
32. USDA. Confronting the Wildfire Crisis. Available at: https://www.fs.usda.gov/managing-land/wildfire-crisis. Accessed 26, 2025 (2024).
33. Jaffe, M. R., Scott, J. H., Callahan, M. N., Dillon, G. K., Karau, E. C., & Lazarz, M. T. Wildfire Risk to Communities: Spatial datasets of wildfire risk for populated areas in the United States. 2nd Edition. Updated 10 September 2024. Fort Collins, CO: Forest Service Research Data Archive. (2024). https://doi.org/10.2737/RDS-2020-0060-2
34. Zuzak, C. et al. The national risk index: establishing a nationwide baseline for natural hazard risk in the US. *Nat. Hazards* **114**(2), 2331–2355 (2022).
35. Ager, A. A., Day, M. A., Aparício, B. A., Houtman, R. & Stinchfield, A. Optimizing the implementation of a forest fuel break network. *PLoS ONE* **18**(12), e0295392 (2023).
36. Ott, J. E., Kilkenny, F. F. & Jain, T. B. Fuel treatment effectiveness at the landscape scale: a systematic review of simulation studies comparing treatment scenarios in North America. *Fire Ecol.* **19**(1), 1–29 (2023).
37. Thompson, M. P., Gannon, B. M. & Caggiano, M. D. Forest roads and operational wildfire response planning. *Forests* **12**(2), 110 (2021).
38. Allaire, F., Filippi, J. B. & Mallet, V. Generation and evaluation of an ensemble of wildland fire simulations. *Int. J. Wildland Fire* **29**(2), 160–173 (2020).
39. Beverly, J. L. & McLoughlin, N. Burn probability simulation and subsequent wildland fire activity in Alberta, Canada-Implications for risk assessment and strategic planning. *For. Ecol. Manag.* **451**, 117490 (2019).
40. Carlson, A. R. et al. Evaluating a simulation-based wildfire burn probability map for the conterminous US. *Int. J. Wildland Fire* https://doi.org/10.1071/WF23196 (2025).
41. Paz, S., Carmel, Y., Jahshan, F. & Shoshany, M. Post-fire analysis of pre-fire mapping of fire-risk: A recent case study from Mt. Carmel (Israel). *For. Ecol. Manag.* **262**(7), 1184–1188 (2011).
42. Parisien, M. A. et al. Commentary on the article "burn probability simulation and subsequent wildland fire activity in Alberta, Canada-implications for risk assessment and strategic planning" by JL Beverly and N. McLoughlin. *For. Ecol. Manag.* **460**, 117698 (2020).
43. Plucinski, M. P. Contain and control: wildfire suppression effectiveness at incidents and across landscapes. *Curr. For. Rep.* **5**, 20–40 (2019).
44. Parks, S. A., Holsinger, L. M., Miller, C. & Nelson, C. R. Wildland fire as a self-regulating mechanism: the role of previous burns and weather in limiting fire progression. *Ecol. Appl.* **25**(6), 1478–1492 (2015).
45. Balch, J. K. et al. Human-started wildfires expand the fire niche across the United States. *Proc. Natl. Acad. Sci.* **114**(11), 2946–2951 (2017).
46. Keeley, J. E. & Syphard, A. D. Historical patterns of wildfire ignition sources in California ecosystems. *Int. J. Wildland Fire* **27**(12), 781–799 (2018).
47. Keeley, J. E. & Syphard, A. D. Twenty-first century California, USA, wildfires: fuel-dominated vs. wind-dominated fires. *Fire Ecol.* **15**(1), 1–15 (2019).
48. Gannon, B. et al. A quantitative analysis of fuel break effectiveness drivers in Southern California National Forests. *Fire* **6**(3), 104 (2023).
49. Miller, R. K., Field, C. B. & Mach, K. J. Barriers and enablers for prescribed burns for wildfire management in California. *Nat. Sustain.* **3**(2), 101–109 (2020).
50. North, M. P. et al. Pyrosilviculture needed for landscape resilience of dry western United States forests. *J. Forest.* **119**(5), 520–544 (2021).
51. North, M. P. et al. Strategic fire zones are essential to wildfire risk reduction in the Western United States. *Fire Ecol.* **20**(1), 50 (2024).
52. Schwartz, M. W. & Syphard, A. D. Fitting the solutions to the problems in managing extreme wildfire in California. *Environ. Res. Commun.* **3**(8), 081005 (2021).
53. Shive, K. L. et al. Thinning with follow-up burning treatments have increased effectiveness at reducing severity in California's largest wildfire. *For. Ecol. Manag.* **572**, 122171 (2024).
54. Steel, Z. L., Safford, H. D. & Viers, J. H. The fire frequency-severity relationship and the legacy of fire suppression in California forests. *Ecosphere* **6**(1), 1–23 (2015).
55. Carreras-Sospedra, M. et al. Air quality and health impacts of the 2020 wildfires in California. *Fire Ecol.* **20**(1), 6 (2024).
56. Cova, G., Kane, V. R., Prichard, S., North, M. & Cansler, C. A. The outsized role of California's largest wildfires in changing forest burn patterns and coarsening ecosystem scale. *For. Ecol. Manag.* **528**, 120620 (2023).
57. Taylor, A. H., Harris, L. B. & Skinner, C. N. Severity patterns of the 2021 Dixie Fire exemplify the need to increase low-severity fire treatments in California's forests. *Environ. Res. Lett.* **17**(7), 071002 (2022).
58. Beneditti, R. Scoring rules for forecast verification. *Mon. Weather Rev.* **138**(1), 203–211 (2010).
59. Noonan-Wright, E. K. et al. Developing the US wildland fire decision support system. *J. Combust.* **2011**(1), 168473 (2011).
60. Duff, T. J., Chong, D. M., Taylor, P. & Tolhurst, K. G. Procrustes based metrics for spatial validation and calibration of two-dimensional perimeter spread models: A case study considering fire. *Agric. For. Meteorol.* **160**, 110–117 (2012).
61. Duff, T. J., Chong, D. M. & Tolhurst, K. G. Indices for the evaluation of wildfire spread simulations using contemporaneous predictions and observations of burnt area. *Environ. Model. Softw.* **83**, 276–285 (2016).
62. Fox-Hughes, P. et al. An evaluation of wildland fire simulators used operationally in Australia. *Int. J. Wildland Fire* https://doi.org/10.1071/WF23028 (2024).
63. Thompson, M. P. et al. Simulating Daily Large Fire Spread Events in the Northern Front Range, Colorado, USA. *Fire* **7**(11), 395 (2024).
64. Gannon, B. M. et al. A geospatial framework to assess fireline effectiveness for large wildfires in the western USA. *Fire* **3**(3), 43 (2020).
65. Young, J. D. et al. The cost of operational complexity: A causal assessment of pre-fire mitigation and wildfire suppression. *For. Policy Econ.* **169**, 103351 (2024).
66. Riley, K. L. & Loehman, R. A. Mid-21st-century climate changes increase predicted fire occurrence and fire season length, Northern Rocky Mountains United States. *Ecosphere* **7**(11), e01543 (2016).
67. Jain, P. et al. A review of machine learning applications in wildfire science and management. *Environ. Rev.* **28**(4), 478–505 (2020).
68. Pang, Y. et al. Forest fire occurrence prediction in China based on machine learning methods. *Remote Sens.* **14**(21), 5546 (2022).

69. Nur, A. S., Kim, Y. J., Lee, J. H. & Lee, C. W. Spatial prediction of wildfire susceptibility using hybrid machine learning models based on support vector regression in Sydney, Australia. *Remote Sens.* **15**(3), 760 (2023).
70. Shang, C., Wulder, M. A., Coops, N. C., White, J. C. & Hermosilla, T. Spatially-explicit prediction of wildfire burn probability using remotely-sensed and ancillary data. *Can. J. Remote. Sens.* **46**(3), 313–329 (2020).
71. Costa-Saura, J. M., Spano, D., Sirca, C. & Bacciu, V. Contrasting patterns and interpretations between a fire spread simulator and a machine learning model when mapping burn probabilities: A case study for Mediterranean areas. *Environ. Model. Softw.* **163**, 105685 (2023).
72. Scott, J., Helmbrecht, D., Thompson, M. P., Calkin, D. E. & Marcille, K. Probabilistic assessment of wildfire hazard and municipal watershed exposure. *Nat. Hazards* **64**, 707–728 (2012).
73. Colavito, M. The human dimensions of spatial, pre-wildfire planning decision support systems: A review of barriers, facilitators, and recommendations. *Forests* **12**(4), 483 (2021).
74. Finney, M. A., Zimmer, S. N., Riley, K. L. & Grenfell, I. C. A generalized wildfire containment algorithm. *Ecol. Model.* **505**, 111–134 (2025).
75. Rothermel, R. C. A mathematical model for predicting fire spread in wildland fuels. Res. Pap. INT-115. Ogden, UT: U.S. Department of Agriculture, Intermountain Forest and Range Experiment Station. 40 (1972).
76. Rothermel, R.C. Predicting behavior and size of crown fires in the northern Rocky Mountains. Res. Pap. INT-438. Ogden, UT: U.S. Department of Agriculture, Intermountain Forest and Range Experiment Station. 46 (1991).
77. Scott, J. H., & Reinhardt, E. D. Assessing crown fire potential by linking models of surface and crown fire behavior. Res. Pap. RMRS-RP-29. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 59 (2001).
78. Finney, M. A. Fire growth using minimum travel time methods. *Can. J. For. Res.* **32**(8), 1420–1424 (2002).
79. Finney, M. A. An Overview of FlamMap Fire Modeling Capabilities. In: Andrews, Patricia L.; Butler, Bret W., comps. 2006. Fuels Management-How to Measure Success: Conference Proceedings. 28–30 March 2006; Portland, OR. Proceedings RMRS-P-41. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 213–220 (2006).
80. Short, K. C., Grenfell, I. C., Riley, K. L., & Vogler, K. C. Pyromes of the conterminous United States. Fort Collins, CO: Forest Service Research Data Archive (2020). https://doi.org/10.2737/RDS-2020-0020
81. Scott, J. H., Short, K. C., & Finney, M. A. FSim: the large-fire simulator: Guide to best practices. (2018) Available at: https://pyrologix.com/wp-content/uploads/2019/11/FSimBestPractices_0.3.1.pdf
82. La Puma, I. P. ed. LANDFIRE technical documentation: U.S. Geological Survey Open-File Report 2023–1045, 103. (2023) https://doi.org/10.3133/ofr20231045
83. Welty, J. L., & Jeffries, M. I., Combined wildland fire datasets for the United States and certain territories, 1800s-Present: U.S. Geological Survey data release, (2021) https://doi.org/10.5066/P9ZXGFY3.
84. Scott, J. H., Dillon, G. K., Callahan, M. N., Jaffe, M. R., Vogler, K. C., Olszewski, J. H., Karau, E. C., Lazarz, M. T., Short, K. C., Riley, K. L., Finney, M. A., & Grenfell, I. C. Wildfire Risk to Communities: Spatial datasets of landscape-wide wildfire risk components for the United States, Second Edition. Fort Collins, CO: Forest Service Research Data Archive. (2024) https://doi.org/10.2737/RDS-2020-0016–2
85. Chakaravarti, I. M., Laha, R. G. & Roy, J. *Handbook of Methods of Applied Statistics,* Vol. I, 392–394 (John Wiley and Sons, 1967).
86. Short, K. C. Spatial wildfire occurrence data for the United States, 1992–2020. 6th Edition. Fort Collins, CO: Forest Service Research Data Archive. (2022) https://doi.org/10.2737/RDS-2013-0009.6

Acknowledgements
We're grateful for the insight and help of staff from the California Department of Forestry and Fire Protection and the USDA Forest Service to develop these products.

Author contributions
C.J.M.: conception, analysis, writing; M.P.T.: conception, analysis, writing; B.A.Y.: analysis, writing; J.H.S.: conception, analysis; M.R.J.: conception, analysis.

Declarations

Conflict of interests
The original work developing the burn probability maps was funded by the California Department of Forestry and Fire Protection. The authors declare no competing interests

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-07968-6.

Correspondence and requests for materials should be addressed to M.P.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.