

Wildfire Susceptibility Scores

Out-of-sample validation of hexagon-level wildfire risk scores for the United States

Property Guardian, Data Science

October 2026

Abstract

Background. Property insurers select and price wildfire risks using commercial scores whose accuracy is rarely tested against wildfires the score had not seen. We evaluate the Property Guardian (PG) Wildfire Susceptibility Scores, which assign four outputs to every H3 resolution-9 hexagon (about 0.1 km², 26 acres) of the United States: Hazard Risk, Recurrence Risk, Local Wildfire Landscape and Proximity to Fuels.

Data. All 81 million hexagons, with no sampling. Outcomes are mapped wildfire perimeters (NIFC/WFIGS, InterAgency Fire Perimeter History, MTBS). The primary test uses 665 wildfires from 2025 against scores frozen before they burned. A walk-forward design re-fits the model as of 2010, 2015 and 2020 and scores the following 1–15 years of wildfires. Results are reported by land area and by properties, using housing units as a proxy to focus the test where residential and commercial property exposure is greatest.

Methods. Discrimination (AUC, class ladders, capture), calibration (observed/expected), a random-placement null, a paired head-to-head against a leading commercial wildfire score and USFS FSim burn probability, and tests of timing after wildfire. Uncertainty comes from a cluster bootstrap that never splits a wildfire, with sizes calibrated by simulation.

Results. Hazard Risk ranked 2025 wildfires at an AUC of 0.85 (95% CI 0.83–0.87), and held that skill 1–15 years ahead (AUC 0.82–0.85). Land it rated Very High burned 218 times as often as land it rated Very Low, and each class burned more than the class below it in every walk-forward window. Where properties are, Hazard Risk outperformed the leading commercial score (AUC 0.961 vs 0.940, $p = 0.003$); across all land the two were statistically indistinguishable. On land that had burned before, properties rated Very High on Recurrence Risk were 21 to 29 times as likely to be inside a wildfire perimeter that year as properties rated Very Low, in every model from 2010 to 2025. After a large wildfire, the same land was about half as likely to burn again in the next two years (0.53 $\times$), and the length of that lull varied by fuel class: in Southern California chaparral it was concentrated in the first two years (0.43 $\times$), while in western conifer forest, where fuels take longest to rebuild, it was deepest in years three to six (0.33–0.34 $\times$). Where wildfire has a long, regular record, Years Remaining anticipated the next event: in Malibu, the model frozen after the 2018 Woolsey Fire put the next wildfire a median of six years out, and the Franklin Fire burned the area in 2024.

Conclusions. Hazard Risk tells insurers where wildfire risk is. It carries durable, calibrated predictive skill at least on par with a leading commercial score, and better where properties

are. Recurrence Risk tells them when. Used with Hazard Risk, it is a skillful temporal signal that separates high-hazard locations where wildfire is likely this year from those where a recent wildfire has temporarily lowered the risk.

Keywords: wildfire; risk score; validation; insurance; calibration; cluster bootstrap; H3.

1 Introduction

Wildfire losses have made the location-level wildfire score a central input to property underwriting. Insurers buy these scores from a small number of commercial vendors. Published evidence of their accuracy is usually descriptive (the share of destroyed properties that a score rated "high") and rarely reports a base rate, an interval, or a test on wildfires the score had not already seen. This paper validates the PG Wildfire Susceptibility Scores with the controls a skeptical actuary would expect (Actuarial Standards Board 2019). Specifically:

1. scores are frozen before the outcome wildfires;
2. outcomes are wildfires the model never saw;
3. uncertainty respects the fact that one wildfire burns thousands of neighboring hexagons;
4. the comparison with a leading commercial score uses the same land and the same wildfires;
5. the analysis plan was written and time-stamped before the confirmatory data were examined.

Land and properties. Insurers do not carry wildfire risk on empty forest, range or desert; they carry it on the residential, commercial and industrial properties built in and around it. Every result in this paper is therefore reported two ways: by land area, which shows how well a score ranks wildfire across the whole landscape, and by properties, which focuses the test on the places where property exposure is greatest. To locate that exposure we weight each hexagon by its housing units (SILVIS 2020). Housing units are a proxy, not the target: they are the most complete and consistently mapped national record of where buildings are, and commercial and industrial properties sit largely in and alongside the same developed areas. Results "by properties" therefore speak to personal and commercial lines alike.

We test the two primary scoring outputs, Hazard Risk and Recurrence Risk, the two supporting outputs, Local Wildfire Landscape and Proximity to Fuels, and the underlying susceptibility index that drives them, and we test them over time. Throughout, each output is evaluated as its continuous 0–1 score unless a class ladder is shown.

2 Study area and data

Study area. The analysis covers all 50 states and the District of Columbia. Every hexagon of the H3 resolution-9 grid (Brodsky 2018) is scored and evaluated: about 81 million hexagons, each about 0.1 km² (26 acres, roughly 300–350 m across). Nothing is sampled; results describe the whole country, not a selection of sites.

Wildfire outcomes. A hexagon is burned in a year if its centre falls inside a mapped wildfire perimeter of that year. The perimeter sources are WFIGS and InterAgency Fire Perimeter History perimeters above 200 acres (National Interagency Fire Center 2026); MTBS perimeters of at least 2,000 acres (Eidenshink et al. 2007); and prescribed burns, included because the model's training record includes them. The 2025 test set comprises 665 wildfires. Hexagons whose 2025 wildfire was already in the model's training record are excluded.

Comparators and covariates. A widely used commercial wildfire score, using its overall score area-weighted to each hexagon, with data predating every wildfire tested; USFS FSim burn probability (LANDFIRE 2020, 270 m; Dillon et al. 2023; Finney et al. 2011); and housing units from the SILVIS 2020 wildland–urban interface data (Radeloff et al. 2023), used as the proxy for property exposure when weighting results by properties (Sections 1 and 4) and to define WUI strata.

3 The model and its outputs

The model estimates for each hexagon a wildfire susceptibility index: the probability that land with its vegetation (LANDFIRE 2025) has burned in the 1900–2025 perimeter record. It is blended over the surrounding neighborhood to give the smoothed susceptibility index. A wildfire frequency model then describes how the chance of wildfire changes with time since the last wildfire for each vegetation type. It gives each hexagon a long-run recurrence rate, an annualized burn probability, and Years Remaining: the number of years until the hexagon is due to burn again, given the years since its last recorded wildfire and the expected return interval of its vegetation. A negative value means the hexagon is overdue.

The product publishes four outputs, each as a 0–1 score and five classes from Very Low to Very High (Table 1). Hazard Risk and Recurrence Risk are the two primary risk scores. The susceptibility index is not published on its own, but it drives every output, so we evaluate it as a component.

Table 1: The four outputs: what they measure, what they are built from, and the key out-of-sample result.

Output	What it measures	Built from	Key result
Hazard Risk	Overall wildfire hazard of the hexagon	The neighborhood-level susceptibility index combined with the long-run recurrence rate	AUC 0.85 [0.83, 0.87] by land, 0.96 by properties (2025); 0.82–0.85 1–15 years ahead
Recurrence Risk	How likely the hexagon is to burn this year	The annualized burn probability plus overdue status from Years Remaining	AUC 0.72 [0.69, 0.75] by land, 0.85 by properties (2025); 0.65–0.72 1–15 years ahead
Local Wildfire Landscape	Wildfire hazard of the surrounding landscape	The spatially smoothed susceptibility index	AUC 0.81 [0.80, 0.83] by land, 0.95 by properties (2025); 0.79–0.84 1–15 years ahead
Proximity to Fuels	How close the hexagon is to burnable fuels	Stepwise distance to the nearest high-susceptibility hexagon, refined by the smoothed index	AUC 0.81 [0.79, 0.83] by land, 0.95 by properties (2025); 0.79–0.84 1–15 years ahead

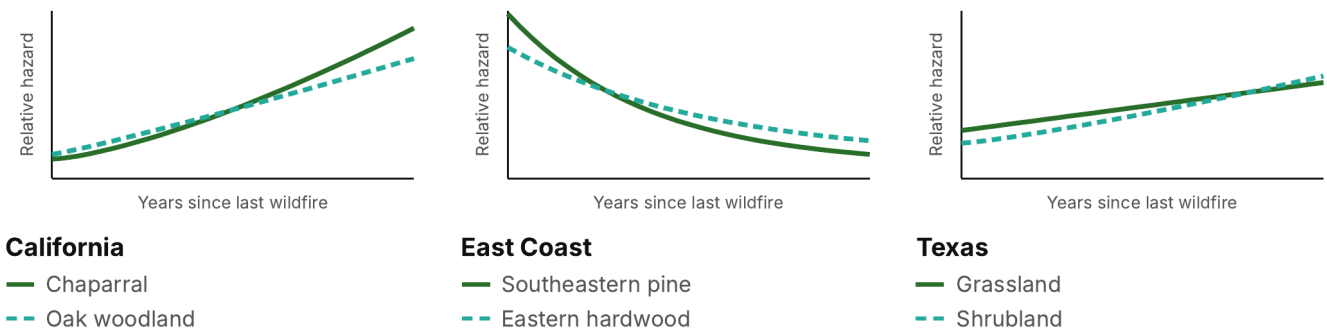


Figure 1: How the timing component behaves: relative hazard of burning by years since the last wildfire, for two vegetation types each in California, the East Coast and Texas. Schematic; curves are shown without scale.

Each vegetation type has its own shape. Chaparral and oak woodland grow back after wildfire, so their hazard rises with every year since the last wildfire. Southeastern pine and eastern hardwood are most likely to burn soon after a wildfire and become less likely as time passes, which is typical of wildfire-maintained systems. Texas grassland and shrubland sit in between, with a gentle rise. Years Remaining reads a hexagon's position on its vegetation's curve, and Recurrence Risk uses it to set overdue status.

4 Methods

Out-of-sample design. The product tested is the frozen file delivered to clients on 25 May 2026. It was scored against 2025 wildfires; the confirmatory 2026 season will be added after the 1 December 2026 data freeze. In the walk-forward design, the full published pipeline was re-fit as of 2010, 2015 and 2020 using only earlier wildfires, then scored against the next 1, 5, 10 and 15 years of wildfires from the same perimeter record. As an acceptance test, a re-fit as of 2026 reproduced the published model: identical training counts and model parameters within 2×10^{-6} ; Hazard Risk class differed on 1 of 80,971,577 hexagons.

Ranking accuracy (AUC). Our main measure is the area under the ROC curve. It answers one question: if we pick one hexagon that burned and one that did not, how often does the score rate the burned one higher? An AUC of 0.5 is a coin flip, and 1.0 means every burned hexagon was rated above every unburned one. We also report the share of burned land captured by the riskiest 20% of land, and burn rates by class. Adjacent-class differences are tested one-sided with Holm correction (Holm 1979).

By land and by properties. *By land*, each hexagon counts equally, so the result describes how well a score ranks wildfire across the whole landscape, most of which has nothing insured on it. *By properties*, each hexagon is weighted by its housing units, our proxy for property exposure (Section 1), so the AUC question becomes: pick one property that was inside a wildfire perimeter and one that was not; how often does the score rate the burned property higher? This measures accuracy where it matters for underwriting: on insured exposure, not on empty land.

Uncertainty. Hexagons are not independent: one wildfire burns thousands of contiguous cells. In simulation, the per-hexagon shortcut (DeLong et al. 1988, or an iid bootstrap) declared a score "better than random" in up to 57% of null worlds at a nominal 5% level. We therefore resample H3 resolution-3 blocks, merged so that no wildfire is split (Lahiri 2003; Roberts et al. 2017), with 2,000 replicates. This held false positives to 4–7%. Paired comparisons use identical resamples, and their standard errors are widened by a simulation-calibrated factor of 1.25.

Random-placement null. Each observed wildfire footprint is moved as a rigid body, randomly rotated, to a random location within 150 km, 1,000 times. The statistic is the mean risk percentile of burned land. Randomizing whole footprints, rather than individual hexagons, preserves the spatial autocorrelation that invalidates naive tests (Dale and Fortin 2002).

Calibration. Observed/expected ratios by group, with wildfire-grouped intervals (Steyerberg et al. 2010), using an equal number of hexagons from every tenth of the susceptibility index re-weighted to the whole country.

Pre-registration. Primary hypotheses, margins and test order were frozen (SHA-256 recorded) before any 2026 comparison. Deviations are logged.

5 Results

5.1 All outputs against random

Hazard Risk, Recurrence Risk, Local Wildfire Landscape and Proximity to Fuels, along with the susceptibility index, rank burned land well above chance, in 2025 and 1–15 years ahead (Figure 2, Table 2). Weighted by properties, every output ranks better still.

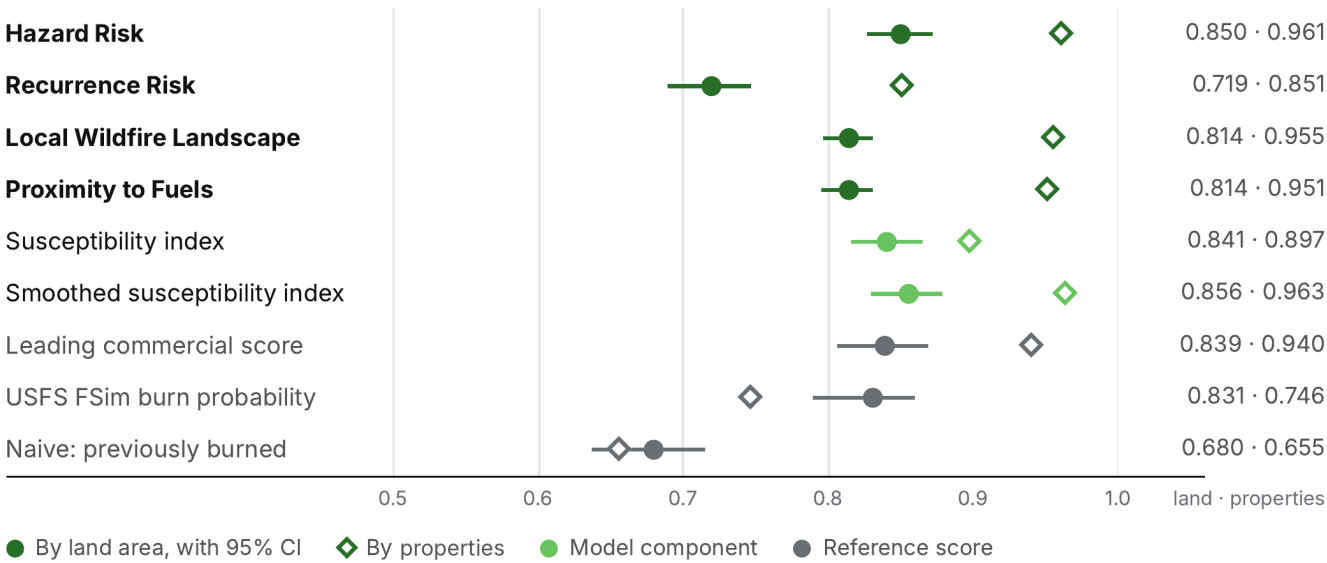


Figure 2: Ranking accuracy (AUC) of each output, component and reference score on 2025 wildfires. Axis runs from 0.5 (a coin flip) to 1.0.

Table 2: Ranking accuracy of the four outputs, the susceptibility index, and reference scores: 2025 wildfires by land area (95% CI) and by properties, and the range over nine walk-forward windows (1–15 years ahead).

Output	AUC 2025 [95% CI]	AUC 2025, properties	AUC 1–15 y ahead
Hazard Risk	0.850 [0.826, 0.872]	0.961	0.82–0.85
Recurrence Risk	0.719 [0.689, 0.747]	0.851	0.65–0.72
Local Wildfire Landscape	0.814 [0.796, 0.831]	0.955	0.79–0.84
Proximity to Fuels	0.814 [0.795, 0.831]	0.951	0.79–0.84
Susceptibility index	0.841 [0.816, 0.865]	0.897	0.80–0.86
Smoothed susceptibility index	0.856 [0.830, 0.879]	0.963	0.81–0.88
Leading commercial score	0.839 [0.806, 0.869]	0.940	—
USFS FSim burn probability	0.831 [0.790, 0.860]	0.746	—
Naive: previously burned	0.680 [0.637, 0.716]	0.655	0.63–0.68

5.2 Class ladders

Every step of the Hazard Risk scale burned more than the step below it (largest Holm-adjusted $p = 0.01$). Land rated Very High burned 218 times as often as land rated Very Low in 2025. Local Wildfire Landscape and Proximity to Fuels also rise class by class (Figure 3).

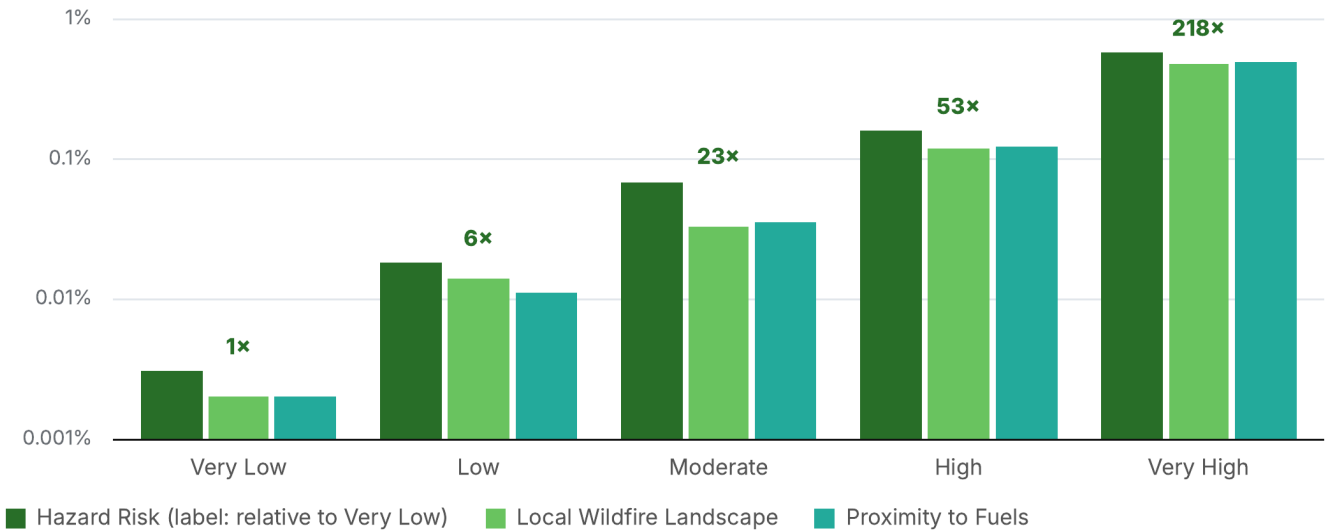


Figure 3: Share of each class's land burned in 2025, land-area weighted. Logarithmic scale.

Table 3 turns the question around: of all the land that burned, how had Hazard Risk rated it beforehand? The answer is the same in every window. Land rated High or Very High made up 38%

of the country, yet it held more than 90% of the land that burned; Very High alone held about three-quarters. Land rated Very Low covered a third of the country and held less than 1%. A score that concentrates nearly all burned land into a minority of the map, and almost none into its lowest class, is a sharp score. Because each column sums to 100%, these shares do not depend on window length, and they are nearly identical from one window to the next.

Table 3: Share of burned land by prior Hazard Risk class, walk-forward windows (as-of year + years ahead). Each column sums to 100%.

Hazard Risk class	Share of all land	2010 + 10 y	2015 + 5 y	2020 + 5 y	2025 + 1 y
 Very Low	34%	0.5%	0.6%	0.4%	0.3%
 Low	14%	1.7%	1.4%	1.2%	1.3%
 Moderate	15%	6.0%	6.0%	5.4%	6.0%
 High	14%	15.0%	16.1%	15.9%	13.9%
 Very High	24%	76.8%	75.9%	77.1%	78.6%
Total	100%	100%	100%	100%	100%

A fair test also asks whether wildfires burned the land the score rated highest, or simply landed in wildfire country. The 2025 wildfires burned land at the Hazard Risk score's 83rd percentile on average; the same wildfires placed at random nearby would average the 69th ($p = 0.001$; Figure 4). The riskiest 20% of land contained 75% of burned land.

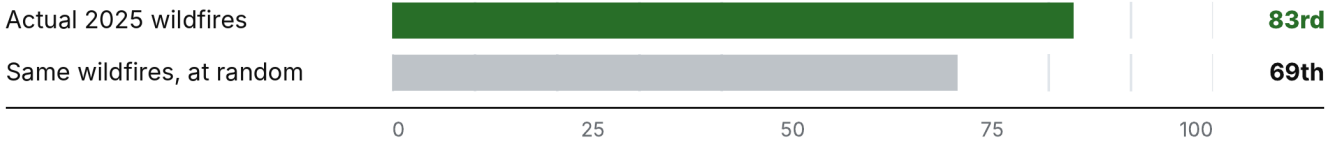


Figure 4: Mean Hazard Risk percentile of burned land: actual 2025 wildfires against the same footprints placed at random within 150 km (1,000 placements).

5.3 Calibration of the susceptibility index

The susceptibility index states the share of land like each hexagon that has burned. Nationally it states 7.1%; the record shows 7.5% (O/E 1.06, 95% CI 0.97–1.14). O/E by tenth of the susceptibility index is 1.02–1.11, and by major vegetation type 1.00–1.17. This is calibration to the index's own 125-year training record, not an out-of-sample forecast; its forward skill is shown in Table 2 (AUC 0.80–0.86 1–15 years ahead).

5.4 Head-to-head with a leading commercial score

Weighted by properties, the PG Hazard Risk score out-ranked the leading commercial score (Figure 5, Table 4). Across all land the two were statistically indistinguishable: Hazard Risk kept $1.03\times$ the leading commercial score's skill, with a lower bound of 0.92. On wildfires that ignited after 1 June 2025, the property-weighted result was a tie ($+0.003$).

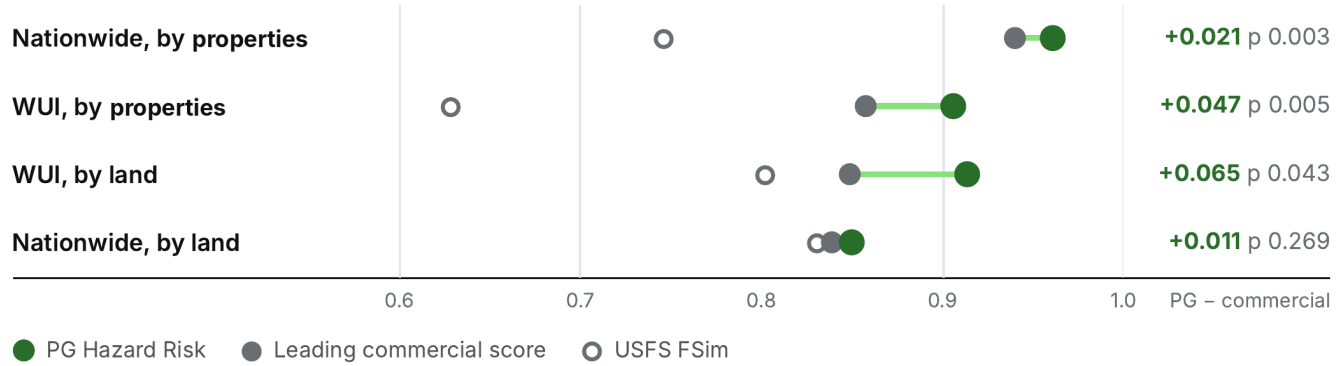


Figure 5: PG Hazard Risk vs the leading commercial score and USFS FSim by segment, 2025 wildfires (AUC, axis 0.6–1.0). The bar marks the gap between PG and the commercial score.

Table 4: Ranking accuracy (AUC), PG Hazard Risk vs the leading commercial score and USFS FSim, 2025 wildfires. p is one-sided (PG better), with wildfire-grouped standard errors widened by 1.25.

Segment	PG Hazard Risk	Leading commercial	USFS FSim	Difference [95% CI]	p
Nationwide, by properties	0.961	0.940	0.746	+0.021 [+0.006, +0.030]	0.003
WUI, by properties	0.906	0.858	0.628	+0.047 [+0.016, +0.077]	0.005
WUI, by land	0.914	0.849	0.802	+0.065 [+0.015, +0.130]	0.043
Nationwide, by land	0.850	0.839	0.831	+0.011 [−0.011, +0.042]	0.269

5.5 Stability over time

Figure 6 scores each year's wildfires from 2010 to 2025 with the most recent model frozen before that year (re-fits as of 2010, 2015 and 2020, and the published model for 2025), on a 1% random sample of land with wildfire-grouped intervals. Hazard Risk ranked each year's wildfires at AUC 0.79–0.85 in every one of the 16 years, and it did not lose skill as a model aged: a model frozen in 2010 ranked 2014's wildfires as well as 2011's. Recurrence Risk ranked them at 0.66–0.74, above the naive "burned before" rule in every year.

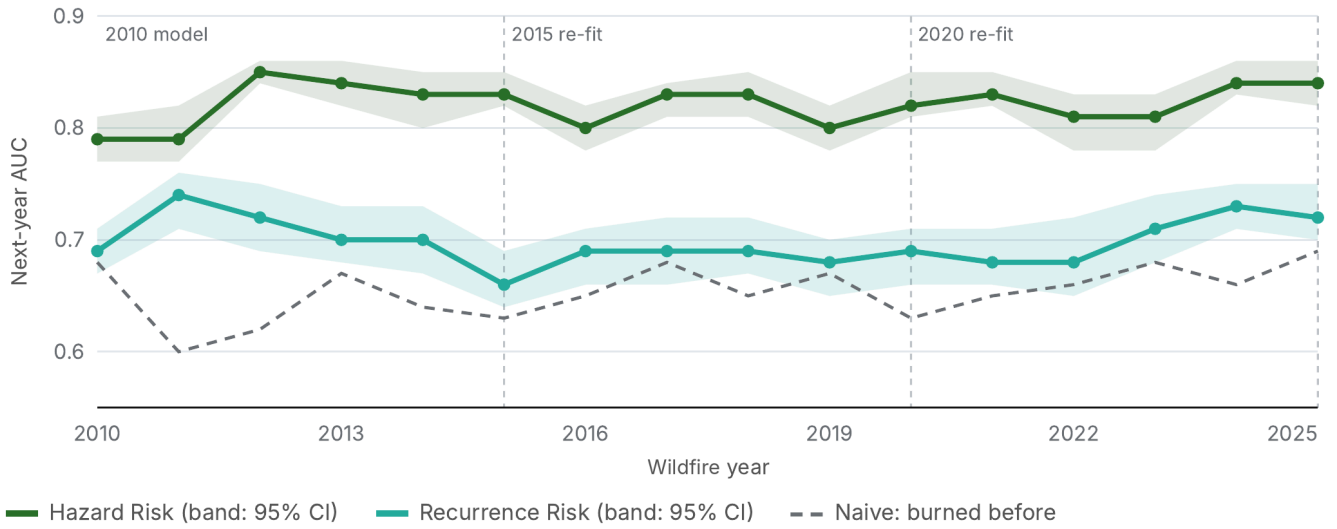


Figure 6: Next-year ranking accuracy (AUC), every wildfire season 2010–2025, scored by the most recent model frozen before that year. 1% random sample of land; wildfire-grouped 95% intervals.

Table 5: Ranking accuracy (AUC, land area) of the four outputs and the susceptibility index in each walk-forward window (as-of year + years ahead).

Output	2010+1	2010+5	2010+10	2010+15	2015+1	2015+5	2015+10	2020+1	2020+5
Hazard Risk	0.82	0.85	0.84	0.84	0.83	0.84	0.84	0.85	0.85
Recurrence Risk	0.69	0.72	0.70	0.68	0.65	0.68	0.67	0.69	0.70
Local Wildfire Landscape	0.79	0.82	0.83	0.83	0.84	0.82	0.83	0.83	0.83
Proximity to Fuels	0.79	0.82	0.83	0.83	0.84	0.82	0.83	0.83	0.83
Susceptibility index	0.80	0.85	0.85	0.86	0.86	0.85	0.85	0.86	0.86
Smoothed susceptibility index	0.81	0.86	0.87	0.87	0.88	0.86	0.87	0.88	0.87
Naive: previously burned	0.68	0.63	0.63	0.63	0.64	0.66	0.65	0.64	0.65

Every output and the susceptibility index hold their skill 1–15 years ahead, and Hazard Risk beats simply knowing where wildfire has burned by +0.14 to +0.22 AUC.

5.6 Recurrence Risk and time since wildfire

Recurrence Risk is the model's timing score. Hazard Risk says where wildfire risk is; Recurrence Risk measures how likely a wildfire is to burn the land *this year*, using two outputs of the wildfire frequency model: the annualized burn probability, and overdue status from Years Remaining. It is

designed to be read with Hazard Risk, and that is where it is most skillful: it separates high-hazard land where wildfire is likely now from high-hazard land where a recent wildfire has temporarily lowered the odds. On its own, it ranked 2025 wildfires at AUC 0.72 by land and 0.85 by properties, and 0.65–0.72 in every walk-forward window. This section shows what it adds beyond Hazard Risk.

Where properties are, a Very Low rating means very few burn. For insurers, the question that matters is what happens to properties. Even on land that has burned before, wildfire in a populated area is a low-frequency, high-severity event: in any given year, only a small share of properties falls inside a wildfire perimeter, but when it happens the losses are large. The test is therefore whether Recurrence Risk separates the few properties that burn from the many that do not. For each frozen model (2010, 2015, 2020 and 2025), we took land that had burned at some point before, read the Recurrence Risk class that model gave it at the time, and measured the share of properties on that land that fell inside a wildfire perimeter in that same year (Figure 7, Table 6). We use the rating as it stood at the time because Recurrence Risk changes as land recovers from a wildfire.

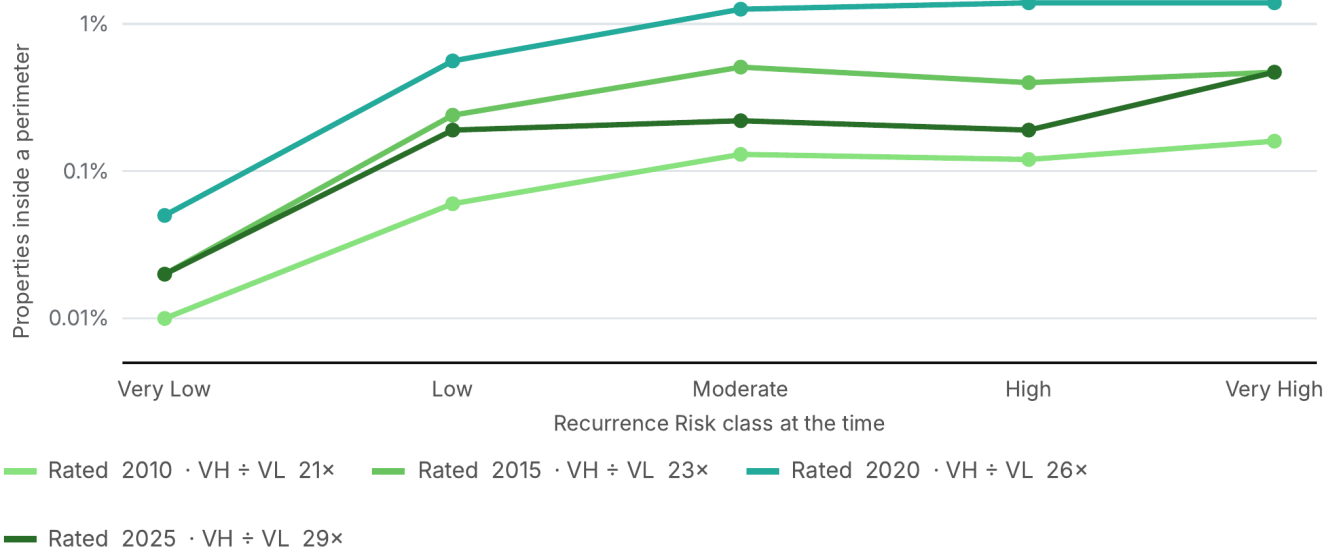


Figure 7: Properties on land that had burned before: share inside a wildfire perimeter that year, by Recurrence Risk class at the time, for each frozen model. Logarithmic scale.

Table 6: Properties on land that had burned before: share inside a wildfire perimeter that year, by the Recurrence Risk class at the time. Every hexagon, weighted by properties.

Rated in → burned in	Very Low	Low	Moderate	High	Very High	VH ÷ VL
2010	0.01%	0.06%	0.13%	0.12%	0.16%	21×
2015	0.02%	0.24%	0.51%	0.40%	0.47%	23×
2020	0.05%	0.56%	1.26%	1.39%	1.39%	26×
2025	0.02%	0.19%	0.22%	0.19%	0.47%	29×
Share of these properties	31%	22%	28%	9%	10%	

The gap between the two ends of the scale is large. Properties rated Very Low were inside a wildfire perimeter at 0.01–0.05% a year, roughly one property in 2,000 to 10,000. Properties rated Very High burned at 0.16–1.39% a year. In every backtest, a property in a Very High Recurrence Risk hexagon was 21 to 29 times as likely to burn as a property in a Very Low hexagon. Very Low is not an edge case: about three in ten properties on previously burned land carry that rating at any given time. The classes in between rise broadly in order, with Moderate and High close together.

Recurrence Risk moves, Hazard Risk holds: after a large wildfire. Recurrence Risk is far more dynamic than Hazard Risk. Hazard Risk reflects where a location sits, and it changes slowly. Recurrence Risk reflects where that location is in its fuel cycle, so it can change from one year to the next as fuels burn off and grow back, and how quickly it changes depends on the fuel class. Because a wildfire consumes the fuel the next one needs, Recurrence Risk deliberately lowers its rating right after one. That is how a hexagon can rate High or Very High on Hazard Risk, because of where it is, yet Low on Recurrence Risk, because it has just burned. To test this, we followed the same land through time: once land burns, how likely is it to burn again in each following year, compared with what is normal for that same land? Figure 8 and Table 7 answer the question first for all susceptible land (susceptibility index above 0.5), regardless of fuel type, and then for three individual fuel classes, using wildfires of 10,000 acres or more from 1984 to 2012.

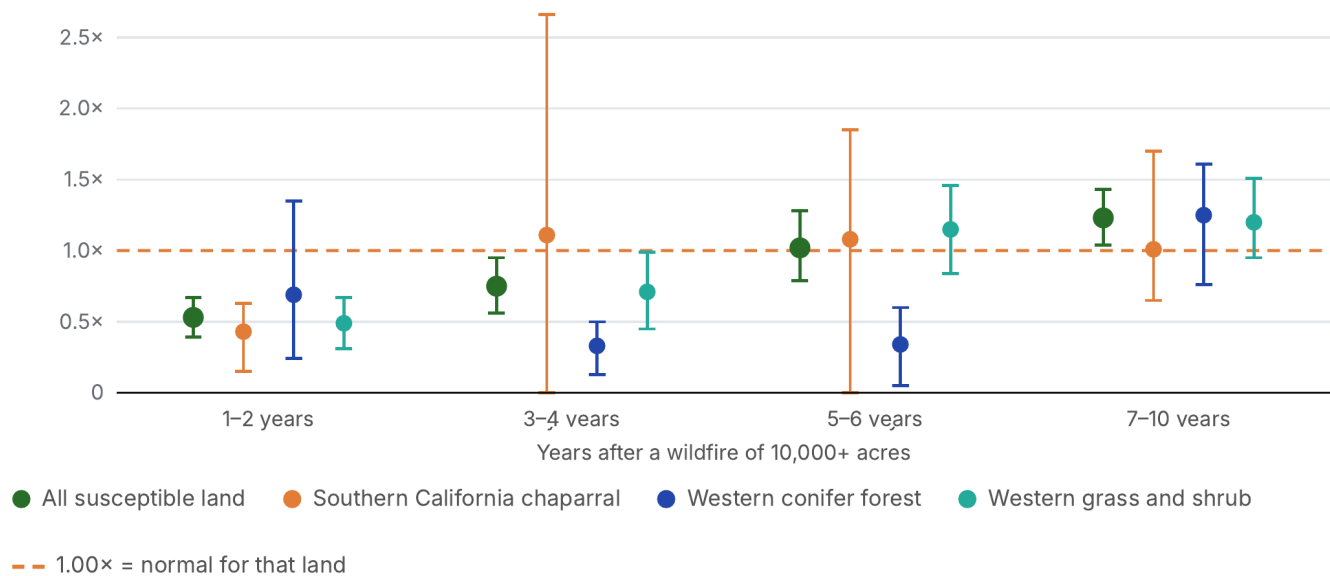


Figure 8: After a large wildfire, the chance that the same land burns again, relative to its own normal rate, by years after the wildfire, for all susceptible land and by fuel class, with wildfire-grouped 95% intervals. Points below the dashed line mean less likely to burn than normal.

Table 7: After a large wildfire (10,000+ acres): the chance the same land burns again, relative to that land's own long-run annual burn rate (1.00x = normal), with wildfire-grouped 95% intervals. Susceptible land, 1% random sample, wildfires 1984-2012.

Land	1–2 years	3–4 years	5–6 years	7–10 years
All susceptible land	0.53× [0.39, 0.67]	0.75× [0.56, 0.95]	1.02× [0.79, 1.28]	1.23× [1.04, 1.43]
Southern California chaparral	0.43× [0.15, 0.63]	1.11× [0.00, 2.66]	1.08× [0.00, 1.85]	1.01× [0.65, 1.70]
Western conifer forest	0.69× [0.24, 1.35]	0.33× [0.13, 0.50]	0.34× [0.05, 0.60]	1.25× [0.76, 1.61]
Western grass and shrub	0.49× [0.31, 0.67]	0.71× [0.45, 0.99]	1.15× [0.84, 1.46]	1.20× [0.95, 1.51]

After a large wildfire, the same land is markedly less likely to burn again for the next few years. Across susceptible land it is about half as likely in the first two years (0.53x) and three-quarters as likely in years three and four (0.75x). In Southern California chaparral it is under half as likely in the first two years (0.43x); in western grass and shrub, about half (0.49x); in western conifer forest, where fuels take longest to rebuild, about a third as likely in years three to six (0.33–0.34x). By years five to ten the chance is back at, or slightly above, the land's normal rate. The fuel class sets the clock, and that is why Recurrence Risk, unlike Hazard Risk, has to be updated as land recovers.

Almost all susceptible land that has burned (99.8%) is rated High or Very High on Hazard Risk, so these results describe exactly the combination of high hazard and low recurrence. After a large wildfire, such land burned at 0.94% a year in the first two years and 1.33% a year in years three and four, against a normal 1.78% a year for the same land. Hazard Risk says where the risk is; a recent

wildfire, which is what lowers Recurrence Risk, says when it is temporarily lower. The window closes, however, while Recurrence Risk can stay low for longer: susceptible land rated Low or Very Low on Recurrence Risk has a median of about eight years since its last wildfire. A low Recurrence Risk on high-hazard land should therefore be read as low for now, not as low for good.

Years Remaining in practice: Malibu. The clearest test of timing is a place with a long, regular wildfire record. The hills above Malibu in the Santa Monica Mountains (37 hexagons around 34.03° N, 118.73° W) burned in 1935, 1943, 1958, 1970, 1982, 1985, 1993, 1996, 2018 (Woolsey) and 2024 (Franklin). Because the full pipeline was re-fit as of 2010, 2015 and 2020 using only earlier wildfires, we can read what each frozen model said about this place before the next wildfire, and compare it with what happened (Table 8, Figure 9).

Table 8: Malibu: what each frozen model said before the next wildfire, and what happened (median over the 37 hexagons).

Model frozen	Hazard / Recurrence	Annual burn prob.	Return interval	Last wildfire known	Years Remaining	Expected next	Actual next
2010	Very High / Very High	7.4%	14 y	1996	−7 (overdue)	~2003	2018
2015	Very High / Very High	8.3%	14 y	1996	−8 (overdue)	~2004	2018
2020	Very High / Very High	3.5%	13 y	2018	+6	~2025	2024
2026 (published)	Very High / Very High	4.7%	11 y	2024	+5	~2031	to come

Every model rated Malibu Very High on both Hazard Risk and Recurrence Risk. The models frozen in 2010 and 2015, which knew of no wildfire since 1996, read the area as seven to eight years overdue; the Woolsey Fire burned it in 2018. The model frozen in 2020, after Woolsey, put the next wildfire about six years out, around 2025; the Franklin Fire burned it in December 2024. The stated annual burn probability, 3.5–8.3%, brackets what actually happened: from 1980 to 2025, 6.1% of this area burned in an average year.

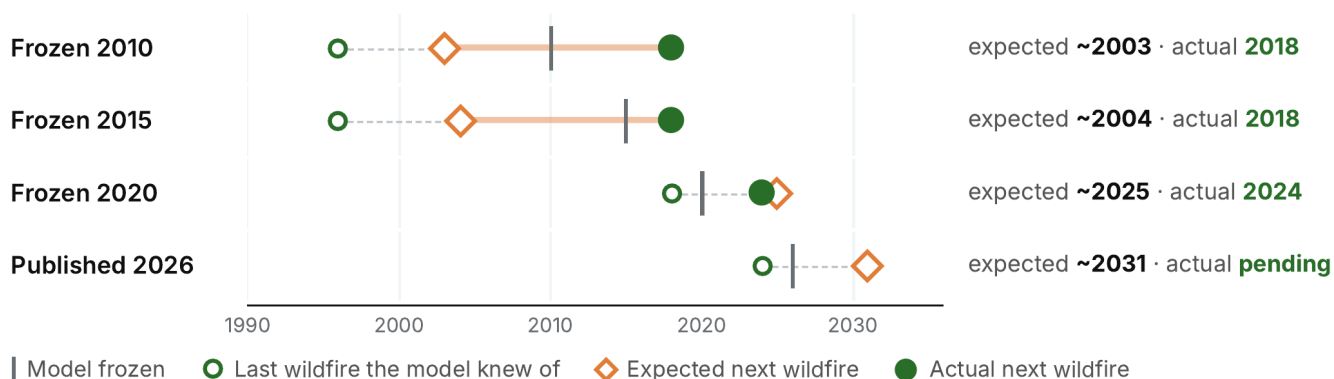


Figure 9: Malibu: the next wildfire each frozen model expected, against the wildfire that came (median over 37 hexagons). The orange bar marks the gap between expected and actual.

The same reading works wherever wildfire has a long, regular record (Figure 10). Near Santa Paula in Ventura County the model expects wildfire about every 5 years, and the record since 1960 shows a median gap of 5 years; on the Snake River Plain in Idaho it expects about every 7 years against a recorded median of 4; above Malibu, about every 11 years against a recorded median of 8. On the Snake River Plain, which the published model showed as due, the Sailor Cap Fire burned part of the area in May 2026.

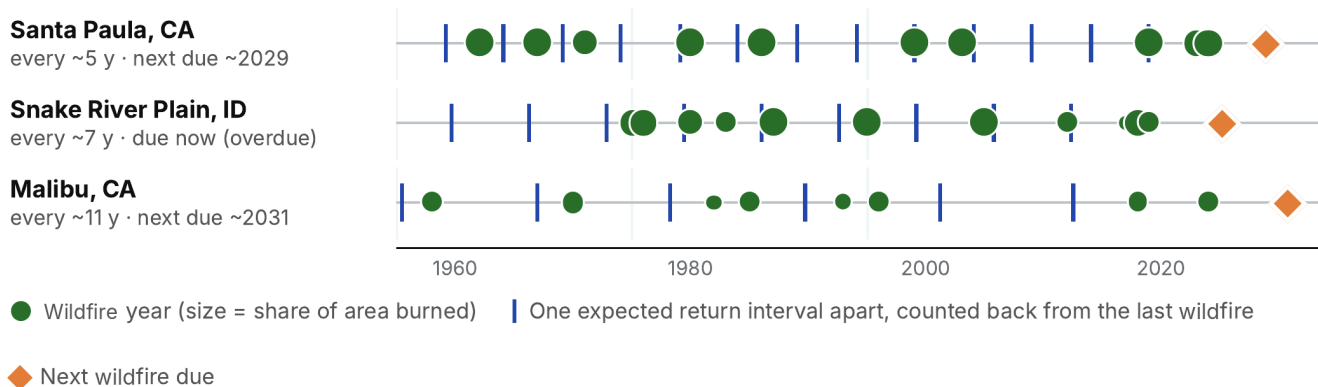


Figure 10: Wildfire history of three repeat-burn places against the published model's expected cadence, 1955–2026. 19 hexagons per site.

Malibu is the case Years Remaining is built for. On land with no recorded wildfire, which has no history to time from, Years Remaining reads as overdue by construction and carries no timing information. It should therefore be read only where wildfire history exists.

5.7 The Spokane Complex: a 2026 lookback

On 1 August 2026 three wildfires started on the edge of Spokane, Washington (Old Trails, Fairview and Autumn Lane), and together burned about 9,932 acres. Local agencies reported several hundred homes and other structures destroyed. The scores below were frozen on 21 May 2026, before the wildfires, and had seen no outcome from them. In line with the analysis plan, 2026 wildfires are not

compared with the leading commercial score until the season's data are frozen in December; this case shows the PG scores alone.

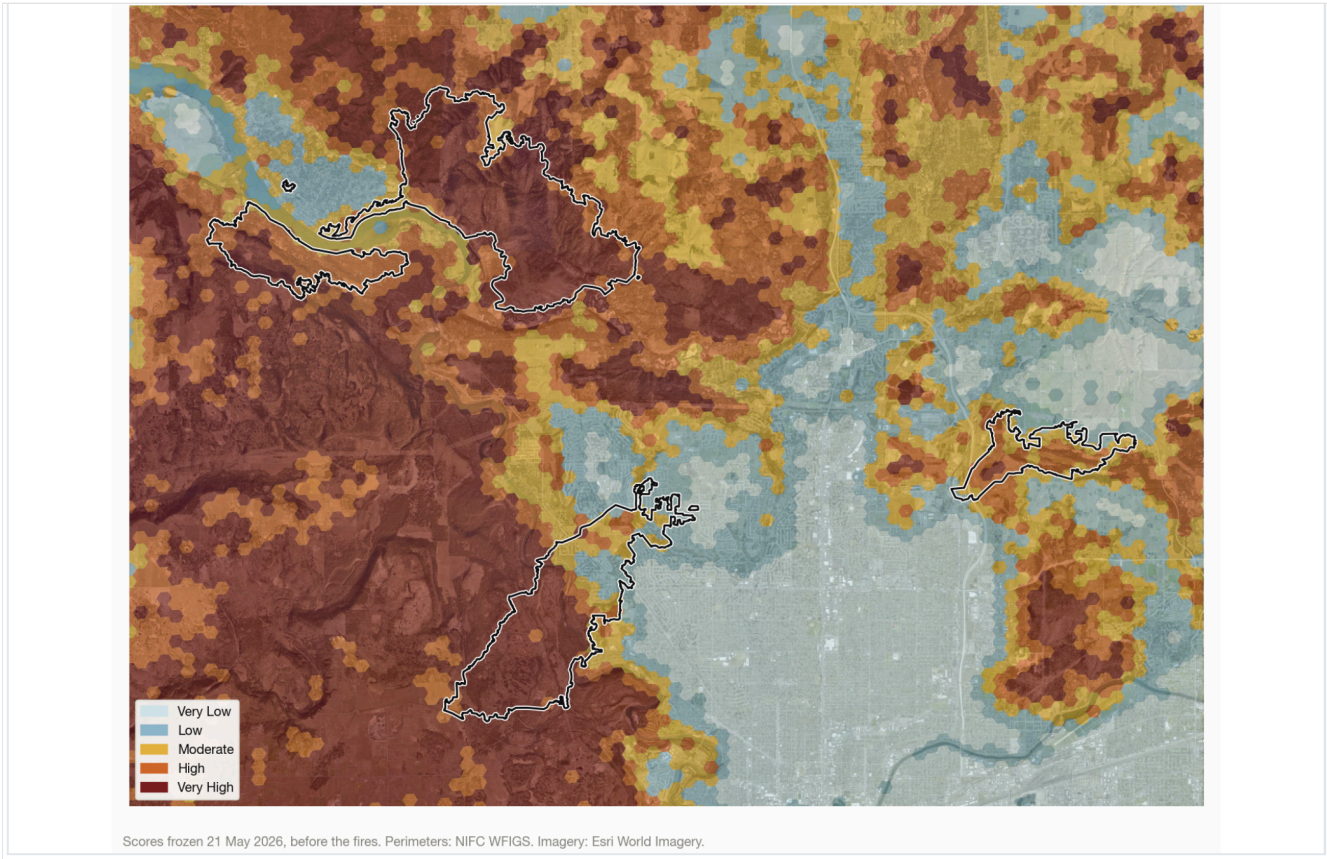


Figure 11: The Spokane Complex: Hazard Risk class of every hexagon (frozen 21 May 2026) with the three wildfire perimeters.

Table 9: The Spokane Complex: share rated High or Very High on the burned footprint and on the unburned land within about 1.7 km of it.

Score	Burned (land)	Burned (properties)	Surrounding (land)
Hazard Risk	83%	19%	55%
Recurrence Risk	85%	25%	59%
Local Wildfire Landscape	90%	34%	65%
Proximity to Fuels	95%	65%	74%

Of the 424 hexagons inside the perimeters, 83% were rated High or Very High by Hazard Risk before the season, against 55% of the unburned land around them. The ranking held locally as well (AUC 0.70 for burned against surrounding land). Properties tell a more mixed story. The perimeters contained about 1,860 properties (housing units in the 2020 data), and only 19% of them sat on land

Hazard Risk rated High or Very High: the wildfires ran into built-up neighborhoods on the city's edge, which a vegetation-based score rates lower (Section 7). Proximity to Fuels, meanwhile, rated 65% of those properties High or Very High, against 21% of properties nearby that did not burn. One wildfire is an illustration, not evidence: it is included because it was blind and because it destroyed properties rather than remote timber.

6 Discussion

The evidence supports all four outputs, each for its own purpose.

Hazard Risk is accurate, calibrated to history, durable over 15 years, and at least as good as a leading commercial score. Where properties are, it is better. Its classes form a clean risk ladder in every window, about 200 times from Very Low to Very High, with three-quarters of burned land in the Very High class. It is the score that says where wildfire risk is.

Recurrence Risk is the timing signal, and it is most skillful read together with Hazard Risk. On land that has burned before, properties it rated Very High were 21 to 29 times as likely to be inside a wildfire perimeter that year as properties it rated Very Low, in every model from 2010 to 2025. It is also the more dynamic of the two: after a large wildfire it lowers its rating while Hazard Risk holds steady, marking a window of reduced risk whose length depends on the fuel class, from about two years in chaparral to as long as six in western conifer forest. Its overdue component carries information only on land with recorded wildfire history.

Local Wildfire Landscape and Proximity to Fuels rank wildfires almost as well as Hazard Risk (0.81 by land, 0.95 by properties) and explain where a hexagon's hazard comes from: the surrounding landscape, and the distance to burnable fuel. In the 2026 Spokane Complex, Proximity to Fuels identified most of the properties that burned on the city's edge.

For underwriting, the practical reading is to select and price on Hazard Risk, read Local Wildfire Landscape and Proximity to Fuels to see where the hazard comes from, and use Recurrence Risk alongside Hazard Risk as the timing signal: is wildfire risk elevated this year, or temporarily lowered by a recent wildfire?

7 Limitations

- **Outcome definition.** Outcomes are mapped perimeters above 200 acres; smaller wildfires and individual structure losses are not scored. Prescribed burns are under-recorded in the 2025+ perimeter source.
- **Comparison scope.** The commercial score is compared as delivered; its structure-level components are not evaluated.
- **Seasons.** One confirmatory-quality season (2025) is reported; the 2026 blind test follows the December freeze.

- **Urban conflagration.** Urban conflagration is outside the scope of vegetation-based wildfire modeling. Property Guardian maintains a separate urban conflagration model; it was not included in this study.
- **Disclosure.** The property-weighted primary endpoint was chosen after the 2025 pilot and before any 2026 data.

8 Conclusion

On wildfires it had not seen, the PG Hazard Risk score ranks where wildfire burns at least as well as a leading commercial score, and better where properties are, with a clean class ladder that held in every year from 2010 to 2025. Recurrence Risk adds the temporal signal that Hazard Risk cannot: on land that has burned before, it separated the properties that burned from those that did not in every model from 2010 to 2025, and after a large wildfire it captured a window of reduced risk lasting from about two years in chaparral to as long as six in western conifer forest. Local Wildfire Landscape and Proximity to Fuels rank wildfires nearly as well as Hazard Risk and show where the hazard comes from. The susceptibility index behind them matches the wildfire record.

References

- Actuarial Standards Board. 2019. Actuarial Standard of Practice No. 56: Modeling.
- Brodsky, I. 2018. H3: Uber’s Hexagonal Hierarchical Spatial Index. Uber Engineering Blog.
- Dale, M. R. T., and M.-J. Fortin. 2002. “Spatial Autocorrelation and Statistical Tests in Ecology.” *Écoscience* 9 (2): 162–67.
- DeLong, E. R., D. M. DeLong, and D. L. Clarke-Pearson. 1988. “Comparing the Areas Under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach.” *Biometrics* 44 (3): 837–45.
- Dillon, G. K., J. H. Scott, M. R. Jaffe, et al. 2023. Spatial Datasets of Probabilistic Wildfire Risk Components for the United States (270 m). 3rd ed. Forest Service Research Data Archive.
- Eidenshink, J., B. Schwind, K. Brewer, Z.-L. Zhu, B. Quayle, and S. Howard. 2007. “A Project for Monitoring Trends in Burn Severity.” *Fire Ecology* 3 (1): 3–21.
- Finney, M. A., C. W. McHugh, I. C. Grenfell, K. L. Riley, and K. C. Short. 2011. “A Simulation of Probabilistic Wildfire Risk Components for the Continental United States.” *Stochastic Environmental Research and Risk Assessment* 25 (7): 973–1000.
- Holm, S. 1979. “A Simple Sequentially Rejective Multiple Test Procedure.” *Scandinavian Journal of Statistics* 6 (2): 65–70.
- Lahiri, S. N. 2003. *Resampling Methods for Dependent Data*. Springer Series in Statistics. Springer.
- LANDFIRE. 2025. LANDFIRE 2024 Fuel Vegetation Type (FVT), CONUS (LF 2.5.0). U.S. Geological Survey; USDA Forest Service.
- National Interagency Fire Center. 2026. WFIGS Interagency Fire Perimeters and InterAgency Fire Perimeter History, All Years. NIFC Open Data.
- Radeloff, V. C., D. P. Helmers, M. H. Mockrin, A. R. Carlson, T. J. Hawbaker, and S. Martinuzzi. 2023. The 1990–2020 Wildland–Urban Interface of the Conterminous United States – Geospatial Data. 4th ed. Forest Service Research Data Archive.
- Radeloff, V. C., M. H. Mockrin, D. Helmers, et al. 2023. “Rising Wildfire Risk to Houses in the United States, Especially in Grasslands and Shrublands.” *Science* 382 (6671): 702–7.

Roberts, D. R., V. Bahn, S. Ciuti, et al. 2017. "Cross-Validation Strategies for Data with Temporal, Spatial, Hierarchical, or Phylogenetic Structure." *Ecography* 40 (8): 913–29.

Steyerberg, E. W., A. J. Vickers, N. R. Cook, et al. 2010. "Assessing the Performance of Prediction Models: A Framework for Traditional and Novel Measures." *Epidemiology* 21 (1): 128–38.

"Leading commercial score" refers to a widely used commercial wildfire score, anonymized.