Research Square

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

Structural Housing Features, Exposure Context, and Wildfire Damage Severity in California's Wildland–Urban Interface

Pranav Elaprolu

elaprolupranav@gmail.com

Irvington High School https://orcid.org/0009-0006-8245-4306

Research Article

Keywords: wildfire, wildland–urban interface, housing characteristics, ordered logistic regression, extended-binary CatBoost, CatBoost, SHAP, Chap. 7A, CWUIC, CAL FIRE DINS, heterogeneity, anomalies

Posted Date: November 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11555500/v1

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

Additional Declarations: The authors declare no competing interests.

Structural Housing Features, Exposure Context, and Wildfire Damage Severity in California's Wildland–Urban Interface

Pranav Elaprolu

Irvington High School, Fremont, California, USA

Correspondence: elaprolupranav@gmail.com

Article type: Research Article

Preprint prepared for Research Square. Article type: Research Article. This manuscript has not been published in a journal. All 20 figures are embedded in this file.

Abstract

California CAL FIRE Damage Inspection (DINS) records from fire-impacted areas (analytic N = 89,826; 298 incidents) do not support a simple "code materials always help" story once exposure context sits in the model. A pre-specified confirmatory set of nine literature terms, with Holm correction on those nine rather than on 56 kitchen-sink dummies, keeps tile roofing (odds ratio [OR] = 0.55), enclosed eaves (OR = 0.74), wood siding (OR = 0.57), local canopy, Wildfire Risk to Communities (WRC) Direct exposure (OR = 1.76), and recorded defense (OR = 0.62) after incident-clustered standard errors. Metal roofing is an honest null (OR = 1.08). Neighborhood density's clustered interval (OR = 1.38, 0.99–1.92) is compatible with no effect to a large one and does not travel from South to North. Combustible siding destruction rises from 61% to 96% across density tertiles, while wood's protective-direction association strengthens in dense neighborhoods. An extended-binary CatBoost ensemble is kept for those density-gated siding bends and two residual phenotypes—packed low-canopy survivors (n = 289, three fires) and sparse hardened losses (n = 379)—not as a second odds-ratio table. Weather enters as a three-variable principal component analysis (PCA) with elevation held out as a geographic control. Incident-held-out leave-one-fire-out quadratic weighted kappa (QWK) is ≈0.48; random-split QWK ≈0.73–0.77 is in-fire discrimination only. McFadden R^2 = 0.1738. Associations are post-inspection and under exposure context—not measured ember flux and not a retrofit prescription.

Keywords: wildfire; wildland–urban interface; housing characteristics; ordered logistic regression; extended-binary CatBoost; CatBoost; SHAP; Chapter 7A; CWUIC; CAL FIRE DINS; heterogeneity; anomalies

Introduction

Macro-Environmental Context and the WUI Crisis

California fires remain an active regional problem, and the mapped hazard footprint has grown. Territory classed as high or very high fire hazard increased by 168% between the prior 2007–2011 Fire Hazard Severity Zone vintage and the 2025 CAL FIRE update[1]. About 3.7 million Californians now live in those mapped areas. That 168% figure is a mapped-zone expansion from the official hazard product, not a housing census; the peer-reviewed companion point is that wildland–urban interface housing has also grown rapidly and concentrates wildfire exposure[2]. Wildfires destroyed more than 70,000 structures in California during the past decade. Hundreds of incidents occur each year, yet the five largest fires account for more than 22,500 destroyed structures, or about 32% of those losses[3]. Interface communities in California likewise concentrate building loss[4],[5].

This prevalence is largely derived from climate change's adverse effects—higher temperatures and longer drought periods inside an already vulnerable climate—so combustibility rises at large.

Forestry, chaparral, and shrublands create even more fuel. WUI expansion from population growth and new infrastructure then puts more homes, and as a consequent more people, in the path of those fires. Even with current suppression programs that spend significant funding, property loss remains a prominent problem for the state[6]. A solution therefore has to look at how particular houses react in fire, not only at vegetation maps, if preferable materials and assemblies are going to be identified. Fuel- versus wind-dominated California fire regimes are one reason vegetation maps alone are incomplete[7]. Broader ignition-factor reviews, including European landscapes, treat fuels and weather as first-order drivers rather than building lists[8].

Literature Review: Known Science, Prior Observational Models, and the Remaining Gap

Housing characteristics keep showing up as more than a side note in wildfire-loss research. Studies that jointly examine exposure, housing density, and infrastructure find that building variables still associate with structural loss after vegetation is taken into account[9,10]. Those results create a need for building studies, not only fuel maps. The same pattern appears outside California. Comparative fire databases from England, the United States, and New Zealand—covering hundreds of thousands of incidents by construction type and occupancy—also treat construction as a consistent part of damage outcomes, even when reporting rules differ across countries[11]. Housing arrangement and location, and the joint influence of fuels, weather, and the built environment, belong to that same observational literature[12],[13].

Laboratory and engineering work has already mapped much of the physics behind current codes: ignition temperatures, flame spread, ember entry at vents and eaves, radiant heating of glass, and the home-ignition-zone idea that treats house loss as an ignition problem as well as a wildland-fire problem[14,15]. California's Chapter 7A and the later California Wildland–Urban Interface Code (CWUIC) grew out of that consensus[16]. Showing once more that combustible materials associate with more damage, or that tile roofs associate with less, mostly restates the rationale already written into code. Those pathways are also reviewed as exposure conditions and component-response strategies, and they include flame-radiation experiments and firebrand combustion[17],[18],[19],[20].

What remains open in observational work is how those code-aligned features behave once neighborhood exposure, vegetation continuity, suppression, and construction-era confounding sit in the same model. Syphard and Keeley (2019) highlight ember-entry paths such as eaves, vents, and multi-pane windows, but they also note that many ostensibly fire-safe homes still burn and that any one variable explains only limited deviance[10]. Dossi and colleagues find that feature–damage links differ between California and Portugal and offer a Wildfire Resistance Index instead of a single material ranking[21]. Knapp and related Camp Fire studies show that distance to destroyed neighbors and canopy can dominate survival, with mixed evidence that Chapter 7A vintage alone is enough[22]. Zamanialaei et al. (2025) likewise rank spacing among the strongest predictors and warn that Year Built bundles hardening, vegetation, and spacing rather than measuring age alone[23]. Fields et al. (Still Standing, 2026) add that observational results shift with scale and that suppression is rarely measured even though it matters in the field[24].

Those papers set the comparison points used here. The present study does not claim to measure ember flux, radiant heat histories, convective flame contact, or ignition sequence at the structure. Those quantities are rarely instrumented in real incidents and are not available in CAL FIRE DINS. The analysis instead adds mechanism proxies—structure spacing and density, vegetation continuity buffers, Wildfire Risk to Communities exposure type, and recorded defensive actions—while retaining environmental principal components and CWUIC-aligned structural features. The target is the structure of associations, heterogeneity by construction era, learned feature packages, and exception cases under those controls, not a rediscovery that materials can burn.

Relation to Prior DINS/WUI Studies

Those papers set the comparison points for the present dual-model design. Relative to representative prior DINS/WUI studies, the present dual-model design is scoped as follows. Syphard and Keeley (2019) established ember pathways through eaves, vents, and windows, but largely in binary or partial settings with limited density-package structure; this study answers with a five-tier ordinal outcome, density/WRC/defense proxies, and explicit package tests[10]. Dossi et al. compared California and Portugal feature–damage patterns and proposed a Wildfire Resistance Index, but did not provide California-scale multiplicity control; here the analytic frame is California DINS (N = 89,826) with FDR/Holm and external BH[21]. Knapp and related Camp Fire work prioritized neighbor destruction and canopy under a single-fire emphasis; the present analysis uses multi-incident DINS and vintage × material stratification[22]. Zamanialaei et al. (2025) advanced DINS machine learning around spacing and Year Built confounding with a predictive emphasis and limited multiplicity; ordered logit remains the inferential association model here while extended-binary CatBoost ensemble is reserved for package discovery and anomalies[23]. Fields et al. (Still Standing, 2026) synthesize inconsistent observational results and the suppression measurement gap; this study adds a recorded-defense proxy and anomaly exception phenotypes[24]. Relative to binary destroyed-versus-survived models, the outcome is ordinal with multiplicity-controlled proportional odds. Relative to predictive-only machine learning papers, ordered logit stays inferential while extended-binary CatBoost ensemble handles packages, anomalies, and out-of-sample discrimination.

Proactive Building Standards: From Chapter 7A to CWUIC

To address major structural vulnerabilities in high risk fire zones specifically situated in the WUI, the state of California initially published Chapter 7A of the California Building Code (Title 24, Part 2) in 2005, particularly dealing with building techniques and evaluating each building's fire resistance as inspired by its components rather than a single architectural monolith. As of January 1st, 2026, the dissolution of Chapter 7A has led to these codes being relocated into 2025 California Wildland-Urban Interface Code (CWUIC), Title 24, Part 7[16]. Specific to the variables within this model, these building codes address the importance of structural factors such as eave enclosures, attic/foundation ventilation screenings, roof assemblies, dual-pane window-glazing, and adjacent structural attachments, all of which are variables which were represented in the regression.

This document was the basis of feature selection and also was heavily aligned with the dataset selected to analyze for individual fire incidents across California. As a consequence, the research utilized comprehensive fire incidents within the state to offer an empiric analysis of whether structures—most predominantly houses—are associated with lower ordinal damage when these mandated components are present in the WUI, once exposure-context proxies and vintage confounding are controlled. CWUIC is treated as settled code knowledge rather than as a hypothesis to rediscover; novelty instead resides in heterogeneity, packages, anomalies, and validation under density/defense/ember–terrain proxies.

Pre-Specific Analysis Plan and Objectives

The scientific objective is not to rediscover that combustible materials can burn, nor to treat CWUIC component lists as unanswered physics. Prior laboratory/engineering consensus and Chapter 7A / CWUIC already encode those pathways. The objective of this observational California DINS study is instead to test where code-aligned expectations hold, reverse, or split under real exposure: after density/spacing, vegetation continuity, WRC exposure type, and a defense proxy are controlled alongside environmental principal components, which CWUIC-aligned housing assemblies still associate with ordinal damage (0–4); how those associations shift by construction era (Year Built confound / pre- versus post-2008 stratification) and density

tertile; which learned feature packages co-occur; and which structures appear as unexpected survivors or unexpected total losses relative to model expectation. Ordered logistic regression is the inferential association model; extended-binary CatBoost ensemble is reserved for nonlinear package discovery, anomalies, and out-of-sample discrimination rather than as a second odds-ratio ranking.

Practically, the analysis is scoped to fire-impacted structures inspected in CAL FIRE DINS (burned-area sample (analytic N = 89,826). It does not estimate measured ember flux, radiant or convective heating, ignition sequence, or dollar retrofit cost-effectiveness. Scenario and component-swap outputs are associational counterfactuals that hold other design-matrix columns fixed; they are hypothesis-generating for exposure-context interpretation, not causal policy mandates. The contribution is therefore heterogeneity, packages, and exceptions under confounding controls—not confirmatory material flammability alone[25].

To ensure that the statistical analysis does not fall victim to "data dredging" practices, wherein a variable with high variance is erroneously attributed to an association with the dependent variable, a pre-specified analysis plan was created to establish the limits of the model. It is crucial to the integrity of the findings that the analysis does not conflate planned features with features discovered only by exploring the dataset. First of all, the plan established a set of CWUIC-aligned structural categories for association modeling (Eaves, Roof, Structure Type, Vents, Exterior Siding, Window Panes, Porches, Patios, and Fences), used with environmental and context controls rather than as an unconstrained search over all housing fields. Secondly, missing-data rules were fixed in advance: Year Built = 0 or missing construction years were dropped, and features with extreme missingness were not imputed into significance by ad hoc fill. Thirdly, temperature, wind speed, relative humidity, and elevation were pre-specified as the environmental continuous block for PCA compression (Year Built kept outside PCA). Finally, multiplicity control was pre-specified: Step-Down Holm–Bonferroni on a pre-specified confirmatory set of nine literature terms (Holm-9) for family-wise error at $\alpha = 0.05$, with Benjamini–Hochberg FDR at 5% retained only as a discovery screen on the larger dummy block. This two-tier approach preserves power for rarer assemblies while limiting false discoveries. Results were further checked against the pre-specified holdout split and revision analyses (vintage stratification, density/defense controls, packages, anomalies) so that exploratory discovery and inferential claims remain separated.

Methods

Data Collection and Preprocessing

The analytic starting pool was 132,792 CAL FIRE Damage Inspection (DINS) records from California counties. That is a large raw dump. After dropping Year Built = 0 or missing construction years and finishing environmental enrichment, the usable sample was N = 89,826 6. Damage itself was coded on a five-step ordinal ladder from 0 (no damage) to 4 (>50% damage or destroyed), because related fire work already treats damage as cumulative when fitting proportional-odds / cumulative-link models 7 and because the research question is about higher versus lower damage tiers rather than a single destroyed-versus-survived cut. Before any dummy coding, three practical edits were required. First, the CSV stored damage as text, so it was cast to integer 0–4 for OrderedModel estimation[21]. Second, Deck/Porch On Grade and Deck/Porch Elevated had been split in a way that duplicated the "No Deck/Porch" flag and would have collinearized the dummy matrix8; those fields were merged into one Porches column with mutually exclusive grade/elevated material strings so absence is not coded twice. Third, rare categories under 1.5% of rows were swept into "Other" or "None." Tiny cells look significant only because almost nobody has them, and that is exactly the kind of false signal the plan was built to avoid.

Sample definition (MF9). The current Attempt 1 master extract contains 132,792 DINS rows (earlier drafts cited ≈131,931). After requiring a mappable ordinal damage label and Year Built > 0, the analytic cohort is N = 89,826 structures across 298 incidents (dropped 42,966 rows). DINS inspects structures inside or near fire perimeters in the State Responsibility Area; this is a burned-area / fire-impacted inspection sample, not a census of California WUI housing. Structure-type composition is mixed: single-family residences comprise 61.0% (n=54,800); utility/miscellaneous structures 23,351; mobile/motor homes 7,040; commercial 2,393. Primary claims are therefore scoped to fire-impacted inspected structures rather than all WUI homes. A single-family–only re-fit is noted as future work; this revision narrows abstract/methods/conclusion language accordingly.

Feature Engineering

Nine structural domains were dummy-encoded: Eaves, Roof Construction, Structure Type, Vent Screen, Exterior Siding, Window Pane, Porches, Patios, and Fences. One reference was dropped in each domain—Eaves_No Eaves, Roof Construction_Asphalt, Structure Type_Single Family Residence Single Story, Vent Screen_No Vents, Exterior Siding_Stucco Brick Cement, Window Pane_No Windows, Porch_No_Deck_Porch, Patio Cover/Carport Attached to Structure_No Patio Cover/Carport, and Fence Attached to Structure_No Fence—so the design matrix stays full rank. Where categories were missing, gaps were filled with proportional stochastic imputation under a fixed seed rather than inventing a new "unknown" level for every field. L1/L2 shrinkage was not used here. The point of the logit was association estimates, with Benjamini–Hochberg kept for post-fit multiplicity rather than coefficient penalties. Continuous weather and terrain terms were later compressed by PCA, but Year Built stayed as its own regressor because it was only moderately tied to temperature, wind, and humidity and because age is exactly the confounder the plan wanted to keep visible. Structural dummies were left uncompressed after multiple-correspondence checks did not show collinearity strong enough to justify sacrificing CWUIC-level readability. Put simply, the design matrix pairs mutually exclusive structural indicators with latent environmental components and a continuous age control under one proportional-odds model.

Imputation justification and post-inspection framing (MF1–MF2). Stochastic proportional imputation was used only to complete categorical CWUIC dummy blocks so that every analytic row retained a full design matrix for the ordered logit and extended-binary CatBoost ensemble. Draws followed observed category frequencies within each structural domain after Year Built filters, keeping the marginal category mix unchanged and avoiding an arbitrary modal fill that would artificially inflate the majority class. We do not claim that imputed cells recover true pre-fire assemblies. Because DINS is a post-fire inspection, feature observability can depend on damage: eaves missing/unknown-like coding is +22.7 percentage points higher among Destroyed than No Damage records in the analytic extract (roof gap -23.7 pp). Proportional imputation cannot remove that selection. Accordingly, structural coefficients are interpreted as post-inspection associations between recorded (or frequency-imputed) DINS categories and ordinal damage under exposure controls—not as identified pre-fire causal effects of building features. Missingness-by-damage and missingness-by-incident tables are provided in the revision package (`missingness_by_damage_wide.csv`, `missingness_by_incident.csv`).

Reference levels were either the modal practice in a domain or an explicit "absence" baseline (No Vents, No Fence, No Patio Cover/Carport). That choice matters. Protective assemblies can then be read against non-installation, and odds ratios speak to cumulative odds of higher damage relative to the omitted category, not to an arbitrary minority material. The same setup tracks CWUIC's component logic (each element versus omission), keeps hardened modal materials from swallowing effects into the intercept, and stays closer to the code-evaluation question the

study actually asks—whether mandated features associate with lower ordinal damage once regional environment is controlled.

Table 1. Condensed analytic variables used in the ordered logit and predictive ensemble.

The full variable dictionary remains in the supplementary package. Table 1 keeps the domains that enter the enriched ordered logit so readers do not have to leave the main text.

Domain	Type	Reference / coding note
Roof, eaves, vents, siding, windows, porch, patio, fence	Categorical dummies	One omitted reference per domain (modal or explicit absence)
Structure type	Categorical	Single-family one-story reference; mixed DINS types retained
Year Built / post-2008	Continuous + binary	Construction-era control; not a retrofit inventory
Weather PCA (T2M, WS2M, RH2M)	Continuous components	Incident-day NASA POWER; KMO = 0.5402
Elevation	Continuous (z)	Geographic control, not a weather synonym
Neighborhood density / spacing	Continuous (z)	Counts and nearest neighbor within 100 m of DINS coordinates
LANDFIRE canopy 30 m / 100 m	Continuous (z)	Long-run vegetation continuity, not event flame contact
WRC Exposure Type	Categorical	Not-exposed reference; long-run hazard class
Defense recorded	Binary	Recorded action vs Missing/Unknown baseline
WildEST ELI / SELB / SES	Continuous proxies	Long-run ember-load products, not measured flux

Context Proxies: Density, Vegetation, WRC Exposure, and Defense

Neighborhood density is the count of other DINS coordinates within 100 m (mean 19.2; median 13), a local structure-to-structure / home-ignition-zone proxy, not a cadastral building-footprint census and not the ember travel distance. Embers can travel kilometers; 500 m and 1-mile DINS-count sensitivities (mean neighbors 321 and 2,146; r = 0.85 and 0.75 with the 100 m count) therefore measure fire-scale settlement clustering, not a better ignition zone. Destroyed-status correlation rises with radius (0.19 → 0.26 → 0.29) for that reason. Long-range ember travel is why 100 m is a packing proxy rather than an ember-flight radius[20].

WRC Exposure Type is a vegetation-adjacency typology, not a housing-density score. Direct marks burnable LANDFIRE fuel; Indirect is non-burnable developed land within about a mile of a large burnable patch—the urban-edge class by design; Not exposed is far from that fuel. On this frame Direct and 100 m density are anti-aligned (r = −0.40): Direct homes have mean 10.5 DINS neighbors and Not-exposed homes have 43.2. VIF warnings (WRC Direct 5.88, WRC Indirect 5.47, Multi Pane 5.68) are dummy-pair and pane-complement arithmetic. Density VIF is 3.22. Both WRC and Multi Pane are retained.

Defense reference coding (MF10). Structure Defense Actions Taken is sparse and dominated by Not recorded / Unknown in the analytic file. The binary defense_recorded indicator therefore compares explicitly recorded intervention categories against the pooled Missing/Unknown/Not-recorded baseline—the empirically modal encoding and the only baseline that keeps the defense contrast identifiable without dropping most rows. Missing/Unknown is not interpreted as proof that no defense occurred; recorded-defense odds ratios are associations relative to incomplete documentation, not relative to verified absence of suppression or civilian action.

These proxies are associational controls and scenario strata. Standardized deltas in P(Y=4) under density or defense scenarios are model-based counterfactuals that hold other design-matrix columns fixed. They are not causal retrofit effects. They are not dollar cost-effectiveness estimates. They are not claims of measured suppression intensity.

Ember and Terrain Proxies (Attempt 2.2)

WildEST Ember Load Index (ELI), Source of Ember Load to Buildings (SELB), and Structure Exposure Score (SES) are CLM modeled long-run rasters (project vintage December 2022–June 2024), point-sampled onto coordinates. They are standing-hazard ember-load proxies, not event ember counts, ignition sequences, or measured flux.

CWUIC-Aligned Packages

Observed DINS categories are further rolled into CWUIC-aligned package indicators for roof hardening, ember-oriented vent screening, enclosed eaves, wall hardening, and multi-pane glazing, together with a count of concurrent domains and a high-package flag. Package variables only describe co-occurrence of observed categories; they do not prove formal CWUIC/Chapter 7A compliance. Adjusted ordered-logit odds ratios for packages are reported with multiplicity-aware interpretation and vintage stratification where available.

Dimensionality Reduction and Statistical Preparation

Environmental coordinates were shifted from EPSG:3310 to WGS 84 (EPSG:4326), cut into space-time keys, and enriched with NASA POWER fields (T2M, WS2M, RH2M) plus elevation. Temperature, wind speed, and humidity move together as a correlated weather block and enter principal component analysis after standardization; elevation is retained as a separate geographic control. On the refreshed incident-day join, Kaiser–Meyer–Olkin sampling adequacy for the weather trio plus the historical four-variable sensitivity block is KMO = 0.5402[26], and Bartlett's χ^2 = 104094.77 (df = 6, p = 0) supports latent-factor extraction. PC1 and PC2 jointly retain 80.10% of that environmental variance. Year Built stays outside the PCA block as a standalone temporal control.

The leftover 19.9% after PC1 and PC2 is not discarded wind spikes. PC3 (13.6%) loads on elevation (0.75) and humidity (0.61). PC4 (6.3%) loads on wind (0.73) and temperature (0.56). Wind's main axis is already PC1 (−0.60). Components are a compression device for residual environmental control, not a named physical factor.

Elevation separation (MF3). Elevation is a relatively fixed geographic attribute, whereas incident-day temperature, wind, and humidity are dynamic weather fields. In this revision, weather PCA is described as compressing the dynamic NASA POWER weather trio (T2M, WS2M, RH2M); elevation is retained as a separate geographic control (z-scored) rather than as a weather synonym. Historical PCA diagnostics that included elevation in the environmental block are retained as sensitivity context, but substantive interpretation treats elevation as terrain/geographic structure, not as a weather substitute.

The incident-day join was refreshed and pushed to every analytic structure. An earlier incomplete join left non-null PC scores on only about 202 rows. The revised PCA diagnostics therefore reflect complete environmental enrichment on N = 89,826 rather than a sparse join artifact.

Variance Inflation Factor Diagnostics

Collinearity among continuous and selected dummy predictors was checked with variance inflation factors on the Attempt 2 and Attempt 2.2 design matrices. No severe VIF flags appeared. Three warning-level elevations recur (WRC Exposure Direct, WRC Exposure Indirect, and Window Pane Multi Pane), which is expected when mutually exclusive categorical encodings and exposure-class partitions sit in the same matrix. Features were not silently dropped for VIF warnings; warnings are reported as diagnostics for coefficient stability.

Ordered Logistic Regression

An ordered logistic (proportional odds) model was fit by maximum likelihood with OrderedModel on the final analytic sample of N = 89,826. Cumulative log-odds are defined so the probability of exceeding threshold j is expressed relative to remaining at or below that

threshold[27]. Under proportional odds, one coefficient vector shifts all cumulative logits the same way, and odds ratios OR = exp(β) express multiplicative changes in the cumulative odds of a higher damage tier relative to the omitted reference category, holding remaining covariates fixed. The enriched specification includes CWUIC dummies, Year Built / vintage indicators, PC1/PC2, and the density, canopy, WRC, and defense proxies described above.

The cumulative logit specification employed throughout estimation is written as:

Equation 1 correction and damage bins (MF4). The manuscript equation is aligned to the statsmodels OrderedModel logit parameterization used for estimation and table interpretation: $\log[P(Y > j)/P(Y \le j)] = X\beta - \theta_j$, where θ_j are threshold cuts and positive β raises cumulative odds of higher damage. (An earlier printed sign that placed $\theta_j - X\beta$ disagreed with the software and with the reported odds ratios.) Damage labels follow non-overlapping DINS-style bins used in the cited literature: 0 = No Damage; 1 = Affected (1–9%); 2 = Minor (10–25%); 3 = Major (26–50%); 4 = Destroyed (>50%). Overlapping printed edges at 10% and 25% in earlier text were label artifacts only; the numeric score mapping used in estimation was unique per record.

Proportional-odds diagnostics (MF4). Brant-style split comparisons (binary logits for Y>j versus the common ordered coefficient) show material departures for several predictors, worst for defense_recorded (max $|\Delta\beta|=0.73$), Exterior Siding_Wood (max $|\Delta\beta|=0.61$), Roof Construction_Tile (max $|\Delta\beta|=0.32$), WRC Exposure_Direct (max $|\Delta\beta|=0.31$), canopy_100m_pct_z (max $|\Delta\beta|=0.28$). The enriched ordered logit nevertheless converged (converged=True; McFadden $R^2=0.1738$). Where proportional odds is most strained, we treat the single-OR summary as an average cumulative association and report incident-clustered cumulative-split and adjacent-category logits as alternatives (proportional_odds_diagnostics_clustered.csv; adjacent_category_clustered.csv). Threshold estimates from the primary OrderedModel (BFGS converged) are $0/1 \approx 0.011$, $1/2 \approx -1.422$, $2/3 \approx -2.879$, and $3/4 \approx -3.631$.

$$\log[P(Y>j)/P(Y\le j)] = X\beta - \theta_j \qquad \text{(eq1)}$$

Immediately after Equation1, p denotes the number of covariates in X (the length of the coefficient vector β). It is not a probability and is distinct from the p-values reported later for coefficient tests. $Y \in \{0,1,2,3,4\}$ is the ordinal damage score, θ_j are threshold intercepts for $j = 0,\ldots,3$, and $\beta_1,\ldots,\beta_p$ are the structural and environmental coefficients. Benjamini–Hochberg FDR-adjusted p-values from the raw coefficient table govern primary inferential claims, while Holm–Bonferroni step-down supplies a conservative family-wise gate.12,13

Vintage Heterogeneity and Year Built Confounding

Year Built is retained, but it is not treated as a laboratory measure of fire resistance. In DINS-based studies, construction year bundles evolving code regimes, correlated material packages, neighborhood morphologies, and often unmeasured vegetation or retrofit histories[23]. The primary vintage cut is 2008 (Chapter 7A era), with sensitivity cuts at 2009 and 2010 labeled as unadjusted destroyed-versus-other checks where applicable. Material associations are examined within vintage strata so CWUIC-aligned features are evaluated conditional on era rather than absorbed into a single age slope.

Year Built limitation (MF7). Parcel Year Built records original construction era and correlated code regimes; it does not capture later roof replacements, window retrofits, siding changes, or vegetation management. A 1980 house with a 2020 roof can appear “old” while the inspected assembly is new. Vintage stratification therefore proxies era packages and selection, not a verified inventory of present-day components. Post-fire rebuilding can also replace assemblies while recorded vintage stays old[28].

2.5 Extended-Binary CatBoost Ensemble (Predictive Branch)

The predictive sibling is an extended-binary CatBoost ensemble: four independently trained cumulative heads, not shared-weight CORAL[29]. Thresholds tuned on the random validation fold are [0.26, 0.34, 0.71, 0.49]; the Y>2 cut sits above the Y>3 cut, so heads can cross. SHAP attributions reported below are from head 3, P(damage > 2), not an average of four heads. The confirmatory inferential table is a literature-locked set of nine terms with Holm correction on those nine. Combustible fence is excluded from that set by pre-specification.

Extended-binary decoding details (MF11). Four independent heads estimate cumulative events {Y>0,…,Y>3}; ordinal labels are decoded by thresholding exceedance scores (QWK-tuned thresholds on the seed-42 validation partition: [0.26, 0.33999999999999997, 0.71, 0.49]). Head crossings (non-monotone cumulatives) are possible; differenced class masses are decoding conveniences rather than calibrated ordinal probabilities. SHAP, component swaps, and anomaly thresholds are interpreted on this decoded predictor, not as outputs of one coherent ordinal likelihood. A true shared-weight rank-consistent ordinal model remains optional full-depth work.

Formally, for K = 5 damage classes indexed as $y \in \{0,1,2,3,4\}$, the extended-binary CatBoost representation trains K−1 binary heads such that:

$$f_i(x) \approx P(y > i) \quad \text{(eq2)}$$

for i = 0,1,2,3. Predicted class probabilities can be recovered by differencing successive cumulative masses. Hyperparameters were fixed at learning rate lr = 0.03, maximum tree depth = 7, iterations = 2000 with early stopping patience of 100 rounds, balanced class weights, and random seed = 42. An 80/20 stratified train–validation split produced n_train = 71,860 and n_val = 17,966. Continuous exceedance scores on the validation partition were converted to discrete ordinal predictions through a grid search over decision thresholds that maximized Quadratic Weighted Kappa (QWK). The decoding rule from thresholded exceedance scores is:

$$\hat{y} = \Sigma_{(i=0)}^{(3)} 1[f_i(x) > \tau_i] \quad \text{(eq3)}$$

Hyperparameter choices were held fixed across the four extended-binary CatBoost heads (learning rate = 0.03, depth = 7, iterations = 2000 with early-stopping patience = 100, balanced class weights, seed = 42) so architectural comparisons are not confounded by per-head retuning. First, each binary head was fit on the training partition using CatBoost ordered boosting[30]. Second, continuous exceedance scores on the validation partition were converted to discrete ordinal predictions through a grid search over decision thresholds that maximized Quadratic Weighted Kappa, privileging ordinal concordance over unweighted accuracy. Third, optimized thresholds were frozen once and applied to report MAE, accuracy, balanced accuracy, and one-versus-rest ROC-AUC by damage class. For the Attempt 2 extended-binary CatBoost baseline, OvR AUC reached about 0.928 for No Damage and 0.916 for Destroyed, with thinner intermediate classes near 0.66–0.74, mirroring the empirical bimodality. Those OvR AUC values should be read as ranking metrics rather than as perfectly calibrated probability statements.

Attempt 2.2 extends the same extended-binary CatBoost architecture with ember and terrain proxies. Random-split QWK = 0.7681 (MAE = 0.6348) versus Attempt 2 QWK = 0.7341 (MAE = 0.7256) used the same validation fold for threshold search and is an in-fire footnote. Head-3 SHAP (P(damage > 2)) is compared directionally with Holm-9 odds ratios; rank disagreement is kept as a finding, not averaged away.

Learned Pair Discovery

CatBoost Interaction pair strengths were pulled independently from all four cumulative extended-binary CatBoost heads and averaged. Discovery happened in the enriched fits rather than from a hand-picked pair list. Ranked CWUIC×proxy and CWUIC×CWUIC ledgers support later associational confirmations (with Holm correction where applied) and descriptive connection-proximity / connection-value audits in Attempt 2.3. Pair strengths are predictive diagnostics of coupled structure–environment patterns, not causal interaction effects.

Model Validation and Robustness

Validation metrics for the machine-learning branch included QWK, MAE, overall accuracy, balanced accuracy, and one-versus-rest ROC-AUC by damage tier. Holm–Bonferroni and Benjamini–Hochberg corrections were applied to the ordinal regression coefficient family[31,32]. Agreement between regression odds ratios and SHAP directional rankings was checked as a cross-model robustness step.

Multiplicity Control on Primary Coefficients

The Holm–Bonferroni step-down procedure orders raw p-values and rejects hypothesis $H_{(i)}$ only while $p_{(i)} \leq \alpha/(m-i+1)$, controlling the family-wise error rate at $\alpha = 0.05$[31]. The Benjamini–Hochberg False Discovery Rate procedure controls the expected proportion of false discoveries among rejected hypotheses[32]. Results emphasize coefficients that survive FDR reporting and note Holm failures explicitly.

External Benjamini–Hochberg Replication Ledger

Attempt 2.3 implements a discovery-versus-external BH ledger across K-fold partitions for primary logit terms and selected interaction/proxy extensions. Terms that fail external BH replication or show sign conflicts are reported as honesty checks rather than suppressed. Primary manuscript claims stay anchored to the full-sample FDR/Holm tables; external failures are read as limited transportability, not as proof of null associations.

Sign-Switch Diagnostics

Sign-switching audits track whether coefficient signs for logit terms or top extended-binary CatBoost features flip across folds or specifications. Sparse sign switches are summarized quantitatively. They argue for caution on unstable niche categories without rewriting the primary association story for domains that stay directionally stable.

K-Fold Stability for extended-binary CatBoost Adoption

Beyond the leaky seed-42 split, K = 5 stratified folds give Attempt 2 QWK ≈ 0.678±0.004 and Attempt 2.2 ≈ 0.706±0.005. Those folds still share incidents across train and test. They are a stability check, not the predictive headline. Grouped leave-one-fire-out QWK ≈ 0.480 is.

Incident-held-out validation (MF12). The seed-42 80/20 row split and row-stratified K-fold place same-incident neighbors in both train and validation, so those QWK/MAE figures are optimistic for transport across fires. As an incident-grouped check we report IEEE BigData grouped leave-one-fire-out (LOFO) DualOrd metrics for the campaign BEST sibling (P25 V22/V22_H1_asymm): mean QWK = 0.480 (sd 0.102); size-weighted QWK = 0.481; mean MAE = 0.926; destroyed recall = 0.719 across 9 grouped folds. These LOFO figures—not the 0.77 random-holdout QWK—are the honest out-of-sample predictive headline for cross-incident generalization. Primary ordinal odds ratios and interaction contrasts are recomputed with incident-clustered sandwich SEs (next paragraph).

Incident-clustered uncertainty (MF4/MF12). The primary OrderedModel was refit with incident-clustered sandwich standard errors (N = 89,826; 298 incident clusters; median SE inflation ≈4.30). Confirmatory inference is Holm-9, not the 18/56 Holm-56 kitchen-sink count: tile 0.55, enclosed eaves 0.74, wood 0.57, canopy 1.20, WRC Direct 1.76, and defense 0.62 survive; metal 1.08, density 1.38 (0.99–1.92), and year-built-per-decade 0.95 (Holm-9 p = 0.052) do not. Four of 21 clustered interactions remain Holm-significant (siding × density). A destroyed-versus-not logit is the binary companion (median SE inflation ≈8.8). Anomaly phenotypes stay 289 unexpected survivors / 379 unexpected losses.

Results

Dataset Characteristics

Analytic N = 89,826 across 298 incidents (train 71,860 / validation 17,966; seed 42 is an in-fire split, not the confirmatory test). Most rows sit at the extremes: No Damage 35,876 (39.94%) and Destroyed 48,698 (54.21%). The middle bands are thin—Affected 3,907, Minor 914, Major 431 (5.85% combined). That bimodality is why a destroyed-binary companion sits beside the ordinal table and why QWK on the middle tiers is structurally weak. Proportional-odds and thresholded heads are practical for endpoints, not a claim that moderate damage is well resolved.

Table 2. Bimodal distribution of ordinal damage scores in the analytic sample (N = 89,826).

Damage Score	Count (n)	Percent
0 — No Damage	35,876	39.94%
1 — Affected (1–9%)	3,907	4.35%
2 — Minor (10–25%)	914	1.02%
3 — Major (26–50%)	431	0.48%
4 — Destroyed (>50%)	48,698	54.21%
Total	89,826	100%

Principal Component Analysis

Temperature, wind, and humidity enter the weather PCA; elevation is a separate geographic control in the primary specification. On the historical four-variable sensitivity block, PC1 + PC2 keep 80.10% (PC1 53.18%; PC2 26.92%). The leftover 19.9% is PC3 (elevation + humidity) and PC4 (a small wind–temperature residual), not unused extreme-weather tape. KMO = 0.5402. Incident-day scores were joined onto every analytic structure.

Table 4. PCA variance retention for the environmental block.

Component	Variance	Cumulative	Eigenvalue
PC1	53.18%	53.18%	2.1271
PC2	26.92%	80.10%	1.0769
PC3	13.61%	93.71%	0.5443
PC4	6.29%	100.00%	0.2517

Figure 1 shows the PCA diagnostics for the environmental block.

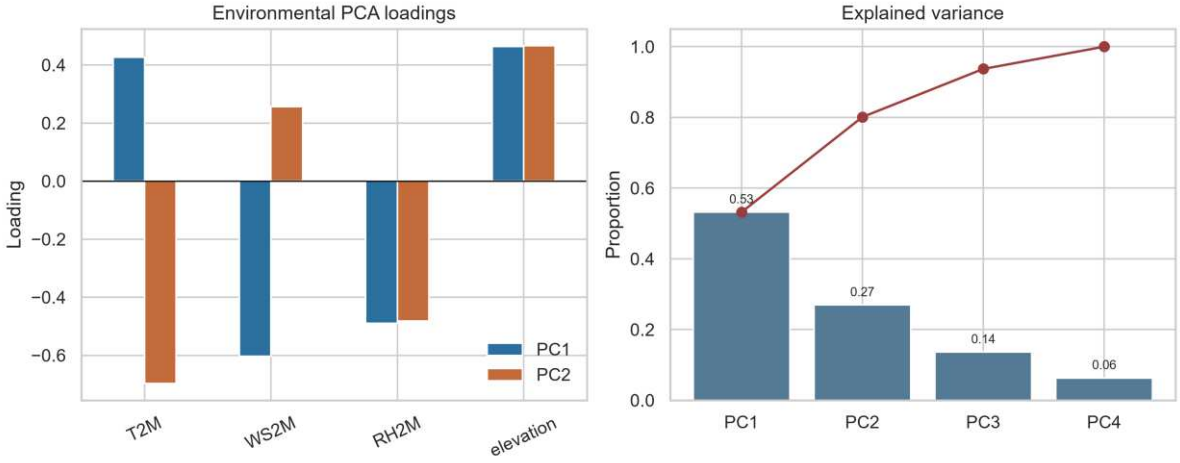


Fig 1. PCA diagnostics for the environmental block (T2M, WS2M, RH2M, elevation), including variance retention.

Ordered Logistic Regression Results

Confirmatory inference is Holm correction on nine pre-specified literature terms, not 56 kitchen-sink dummies (18/56 Holm after clustering is mostly porch, vent, and structure-type residue). Six of the nine survive Holm-9: tile OR = 0.55 (0.45–0.69; destroyed-binary OR = 0.40), enclosed eaves OR = 0.74 (0.66–0.84; binary 0.82), wood siding OR = 0.57 (0.40–0.79; binary 0.33), canopy 30 m OR = 1.20 (1.10–1.30), WRC Direct OR = 1.76 (1.18–2.62; binary 1.41), and recorded defense OR = 0.62 (0.44–0.87; binary 0.30). WRC and defense were FDR-only under Holm-56 and return once the multiplicity budget is the confirmatory set. Metal is an

honest null (OR = 1.08, 0.75–1.56). Density OR = 1.38 (0.99–1.92) stays interval-valued, not starred. Year built per decade OR = 0.95 is borderline (Holm-9 p = 0.052). McFadden R^2 = 0.1738[33]. Table 3 and Figure 2 report the confirmatory set. Full 56-row clustered table stays supplementary.

Metal roof × density confound (MF6). The adverse metal-roof association in dense tertiles may partly reflect unobserved roof quality, thickness, fastening, aging, or installation detailing that DINS categories do not record. Density may also proxy older stock and maintenance patterns. We therefore do not attribute the metal×density pattern solely to spacing physics.

Table 3. Pre-specified confirmatory nine: incident-clustered ordinal odds ratios, destroyed-binary companions, and Holm correction on these nine terms only.

Term	Ordinal OR	Cluster 95% CI	Destroyed-binary OR	Holm-9
Tile roof	0.55	0.45–0.69	0.40	Yes
Enclosed eaves vs Missing	0.74	0.66–0.84	0.82	Yes
Wood siding	0.57	0.40–0.79	0.33	Yes
Canopy 30 m (z)	1.20	1.10–1.30	—	Yes
WRC Direct	1.76	1.18–2.62	1.41	Yes
Defense recorded	0.62	0.44–0.87	0.30	Yes
Year built / decade	0.95	0.92–0.99	0.95	No (p=0.052)
Density log n 100 m (z)	1.38	0.99–1.92	1.24	No
Metal roof	1.08	0.75–1.56	1.09	No

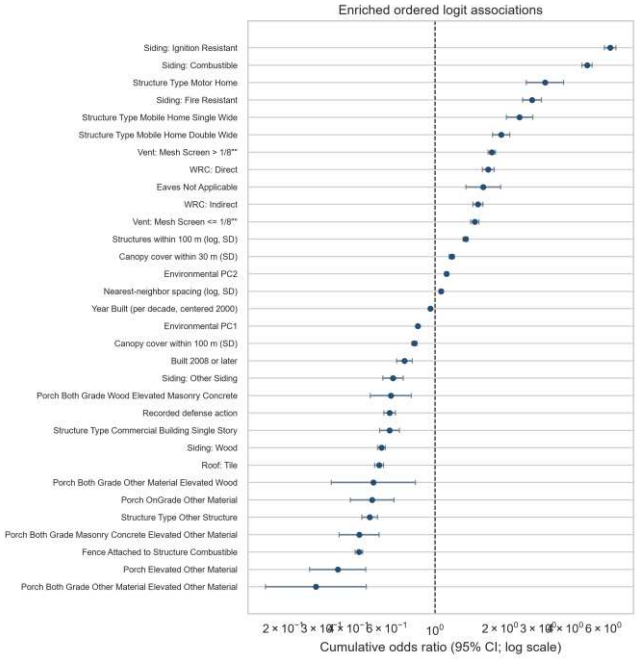


Fig 2. Forest plot of enriched ordered-logit coefficients for CWUIC-aligned structural predictors and context proxies, with confidence intervals.

Enclosed and unenclosed eaves are both Holm-significant versus Missing/Unknown (OR = 0.74 and 0.79). They are not distinguishable from each other: clustered contrast Δ log-odds ≈ −0.068 (approximate z ≈ −0.73, p ≈ 0.47), and the associational unenclosed→enclosed swap changes P(Destroyed) by only −1.1 to −1.3 percentage points. Eaves missingness is 22.7 points higher among Destroyed than among No Damage, so “no eaves” mixes true absence with unclassifiable assemblies. The confirmatory claim is enclosed eaves versus Missing, not enclosed versus open.

Exterior siding produced the largest positive odds ratios relative to stucco/brick/cement: combustible siding (β = 1.6166, OR = 5.036) and ignition-resistant labeled siding (β = 1.8611, OR = 6.431) elevated cumulative odds, while wood (OR = 0.566) and other/metal categories

registered protective-direction associations that may reflect coding prevalence in high-hazard parcels rather than isolated laboratory flammability rankings. As such, these administrative labels require association-safe interpretation: they may proxy hazard-zone selection and inspection conventions as much as intrinsic combustibility.

Vent screening relative to no vents increased cumulative odds for screened configurations, with coarse mesh > 1/8" producing the largest elevation ($\beta = 0.6042$, OR = 1.830) and mesh ≤ 1/8" also elevated (OR = 1.528). Mobile home and motor home structure types amplified cumulative odds by factors of approximately 2.02–3.22 relative to single-family single-story residences, whereas multi-story single-family residences associated with lower cumulative odds (OR = 0.750). Multi-pane glazing associated with lower cumulative odds (OR = 0.806) relative to the no-windows reference, consistent with dual-pane thermal resistance delaying glazing failure under radiant exposure in prior engineering narratives.

Combustible attached fence OR = 0.445 is Holm-56 significant and physically backwards (more "protective" than non-combustible OR = 0.756). Destroyed rates are 40.4% (combustible), 54.6% (non-combustible), and 59.0% (no fence). Fence fields are less missing on Destroyed (12.2%) than on No Damage (32.5%), so this is not an incinerated-fence missingness story. The term is a measurement audit, not a confirmatory protective finding. Roof Other holds 3.8% of rows (n = 3,409) and siding Other 1.9% (n = 1,696); the <1.5% collapse rule fills buckets that can themselves exceed 1.5%. Window Other is n = 3. Vent mesh OR > 1 and ignition-resistant siding OR = 6.43 stay in the same audit.

Identification, not double-counting: crude destroyed rates by WRC are almost flat (Direct 55.0%, Indirect 54.2%, Not exposed 50.5%). After crowding enters, WRC Direct destroyed-binary OR = 2.30 (Spec C) and remains 2.20 in the full Spec D. Density is 1.39 in Spec B, 1.50 in Spec C, then 1.29 and non-significant once weather components and defense enter Spec D—the same interval story as the 56-covariate clustered OR 1.38 (0.99–1.92). Mutual adjustment is the point. South versus North (lat 35.8): overall destroyed 48.8% / 58.4%. Tile destroyed rates are 33.7% / 32.4%. High-density destroyed rates are 55.6% / 79.2%. Tile transports; density does not. Leave-three-mega-fire tile ORs stay protective (0.37 all; 0.53 drop Camp; 0.26 drop two LA fires; 0.48 drop all three). Those three incidents are 58% of N.

VIF Diagnostics

VIF warnings are WRC Direct 5.88, Multi Pane 5.68, and WRC Indirect 5.47. Density VIF is 3.22. No flagged term was dropped: WRC is a vegetation-adjacency dummy pair, and Multi Pane is the CWUIC glazing contrast against Single Pane. Retaining them is a specification choice, not a silent exclusion.

Design	Variable	VIF	Flag
attempt2	WRC Exposure_Direct	5.878	warn
attempt2	Window Pane_Multi Pane	5.682	warn
attempt2	WRC Exposure_Indirect	5.467	warn
attempt22	WRC Exposure_Direct	7.804	warn
attempt22	WRC Exposure_Indirect	6.426	warn
attempt22	Window Pane_Multi Pane	5.792	warn

Domain Patterns under Environmental and Context Controls

Across CWUIC domains, roof and eave protections generally associate with lower ordinal damage, while mobile-home structure types and coarse vent mesh associate with higher cumulative odds. Patio and porch configurations are more heterogeneous. Context proxies do not erase material signals; they reframe them under density and exposure strata. SHAP and logit ranks are not required to agree; the clash is the ML result.

Year Built and Vintage Stratification

On average, decade-centered Year Built and the post-2008 flag pull cumulative odds down. Stratify by vintage, though, and the material picture is not one universal age effect. At the 2008

cut in the adjusted ordinal interaction, metal roofing is elevated pre-2008 (OR ≈ 1.22) and protective post-2008 (OR ≈ 0.74). Tile moves from OR ≈ 0.56 (pre) to OR ≈ 0.29 (post). Holm-confirmed secondary terms keep Roof Construction_Tile × post-2008 and metal × density in the ledger. Zamanialaei et al. (2025) treat construction year as a bundle of hardening, vegetation, and spacing; reading a negative Year Built OR as "newer houses survive because they are newer" overclaims that bundle. Scenario bars tell a similar story: post-2008 parcels show lower mean P(Y=4) than pre-2008 when other columns stay fixed. Figure 5 shows those scenario bars. Figure 3 is the confounding panel.

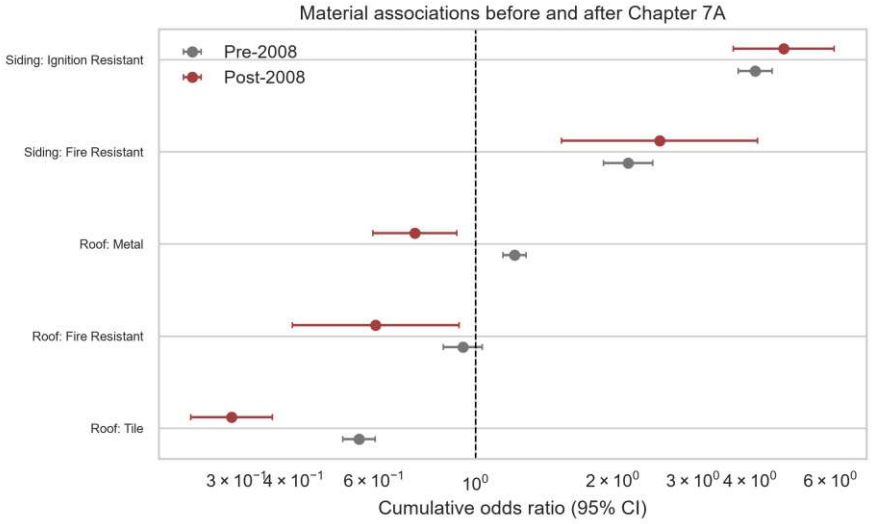


Fig 3. Year Built confounding diagnostics: material associations stratified by pre- versus post-2008 construction vintage.

Density and Defense Context

Siding × packing is the confirmatory nonlinear table (4 of 21 clustered Holm interactions). Combustible siding × high density OR = 5.66; × middle = 6.47. Wood siding × high density OR = 0.39. Ignition-resistant siding × middle OR = 3.89 (kept, treated as coding/selection). Raw destroyed rates make the bend visible without a log-odds: combustible siding 61% / 88% / 96% across low / middle / high density tertiles; wood siding 34% / 18% / 39%. Tile OR stays 0.54–0.65 across tertiles. Metal 1.22 / 0.81 / 2.43 is the quality confound already flagged, not a new win. Asphalt→tile associational swap changes P(Destroyed) by −10 to −12 points in every tertile. Figure 4 remains the density × material panel.

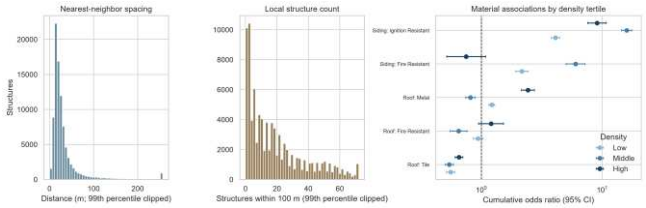


Fig 4. Density and spacing summaries alongside material odds ratios stratified by neighborhood density tertile.

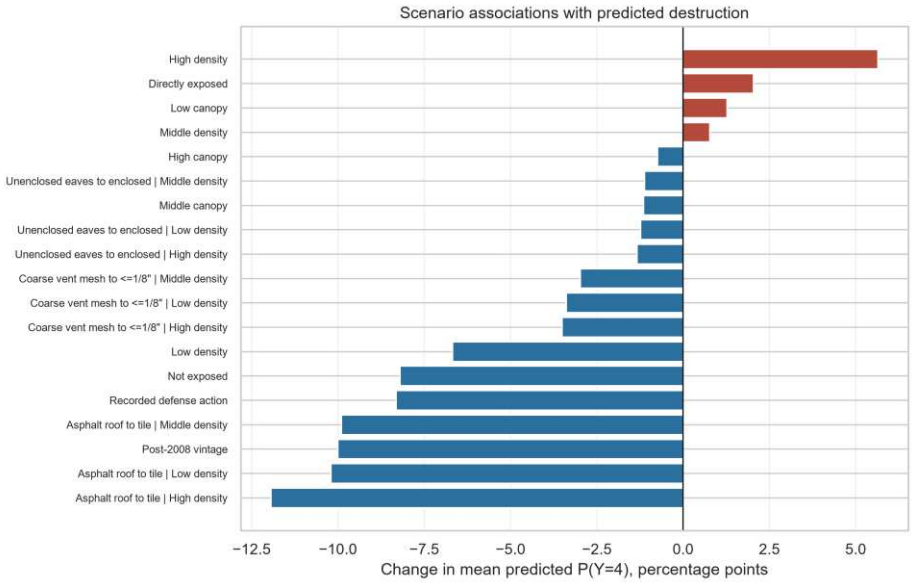


Fig 5. Associational scenario standardizations for density, WRC exposure, vegetation, defense, and vintage.

Discrimination Metrics

Validation ladder, not a QWK contest. Ordered logit QWK = 0.442 on the leaky random split. A size-capped frozen incident test (L1: 12 fires, n = 14,921, 16.6% of N, 7,582 North / 7,339 South, mega-fires reserved for leave-one-fire-out) has 31.1% destroyed; majority-class accuracy is 68.9%, a high-density tertile rule 66.3%, and WRC Direct 49.6%. Grouped leave-one-fire-out DualOrd QWK = 0.480 (MAE = 0.926) is the predictive headline. Random-split CatBoost QWK 0.734 (Attempt 2.2 0.768) used the same fold for threshold search and is an in-fire footnote. CatBoost is an endpoint detector (OvR AUC 0.928 none / 0.916 destroyed / 0.66–0.74 middle).

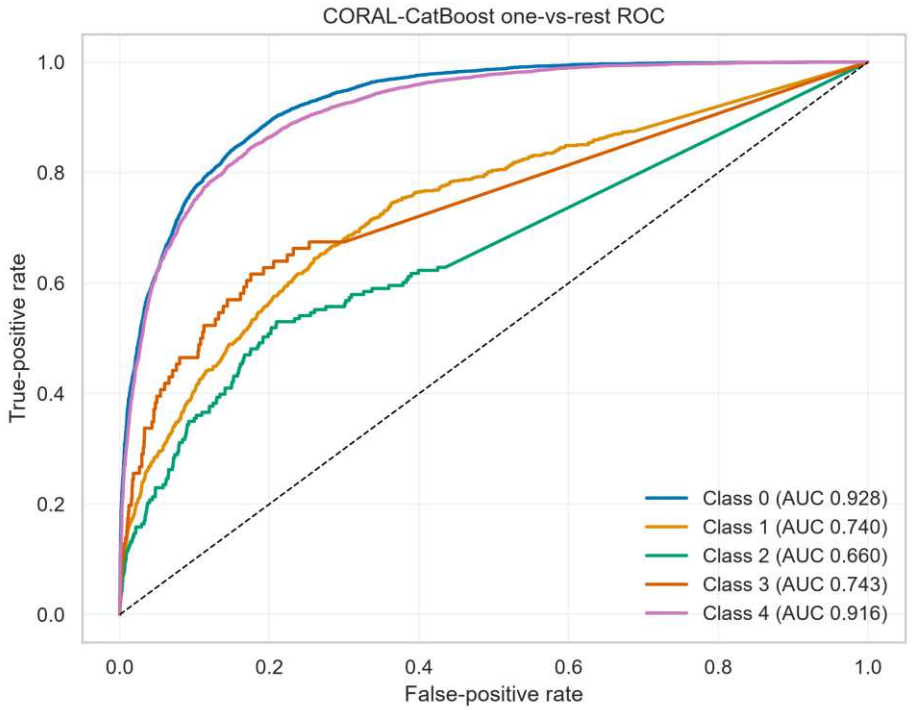


Fig 6. One-versus-rest ROC curves for extended-binary CatBoost ensemble validation predictions by ordinal damage class.

Machine Learning Performance

Table 5. Validation ladder. Leave-one-fire-out QWK is the predictive headline; random-split CatBoost QWK is an in-fire footnote.

Rung	Protocol	Metric	Role
Linear ceiling	Ordered logit, random val	QWK 0.442	in-fire
Dumb baseline (L1)	Majority class	Accuracy 68.9%	baseline
Dumb baseline (L1)	High 100 m density tertile	Accuracy 66.3%	baseline
Frozen incidents (L1)	12 fires, n=14,921, N/S balanced	Destroyed 31.1%	unseen fires
Honest ML headline	Grouped LOFO DualOrd	QWK 0.480; MAE 0.926	headline
In-fire footnote only	CatBoost thresholds tuned on val	QWK 0.734 / 0.768	not transport

Figure 7 compares the ordered logit with the Attempt 2 extended-binary CatBoost ensemble on the holdout.

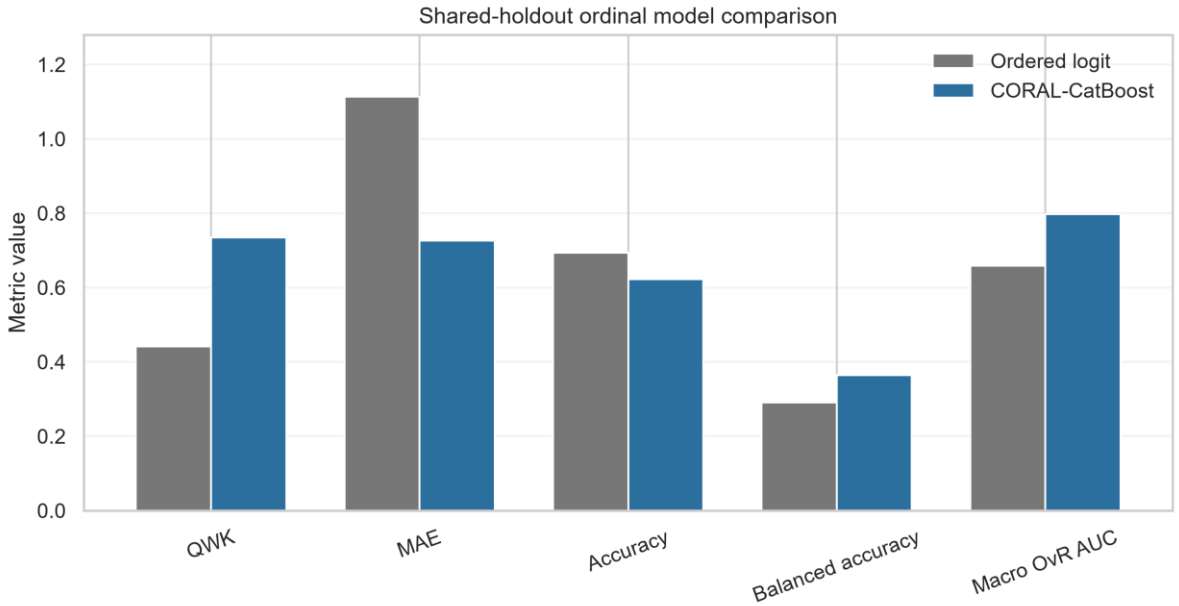


Fig 7. Holdout model comparison between ordered logit and the extended-binary CatBoost ensemble (Attempt 2 baseline).

The tree and the logit do not rank the same world. Head-3 SHAP (P(damage > 2)) leads with PC2 (0.495), combustible siding (0.453), and 100 m density (0.427). Tile is SHAP rank 14 and enclosed eaves rank 18, yet both are Holm-9. Wood siding is the hinge (SHAP rank 5 and Holm-9). CatBoost versus ordered-logit anomaly Jaccard is 0.36. Average tile protection can be real and still be a weak in-fire split. That disagreement is why both models stay.

Attempt 2.2 Ember/Slope Adoption

Attempt 2.2's random-split gain (ΔQWK ≈ +0.034; ΔMAE ≈ −0.091) coincides with ELI×environment/density pairs. Class-wise OvR AUCs remain endpoint-strong (≈0.941 none / 0.929 destroyed) while middle tiers stay thin. Those pairs are associational diagnostics. Adoption does not change Holm-9 or replace leave-one-fire-out QWK = 0.480. Figures 8–10 show the Attempt 2.2 holdout comparison, the ember and slope pairs, and the adoption gate.

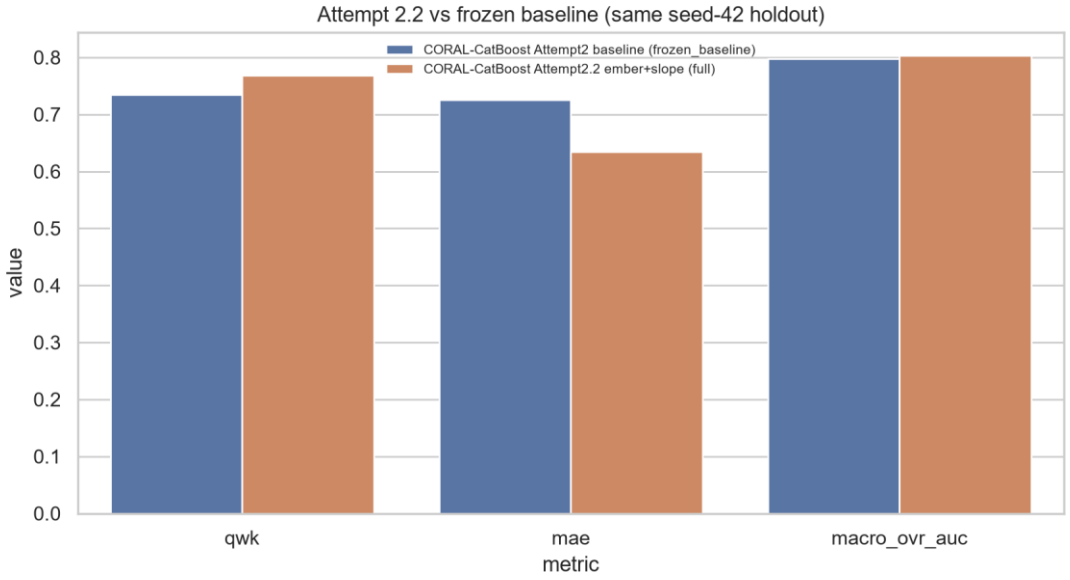


Fig 8. Holdout comparison including the Attempt 2.2 ember/slope-optimized extended-binary CatBoost ensemble (adopted sibling).

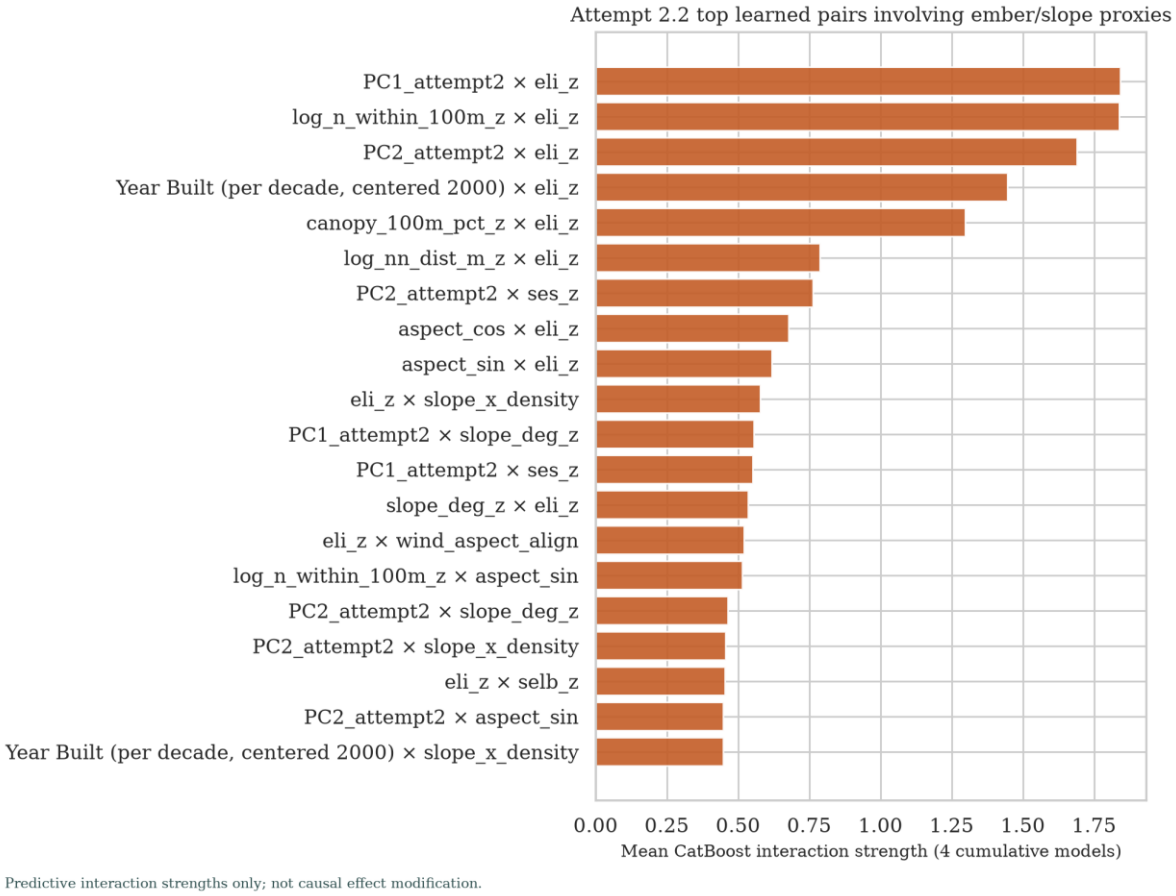


Fig 9. Attempt 2.2 learned pairs emphasizing ELI/SES ember and slope/terrain proxies.

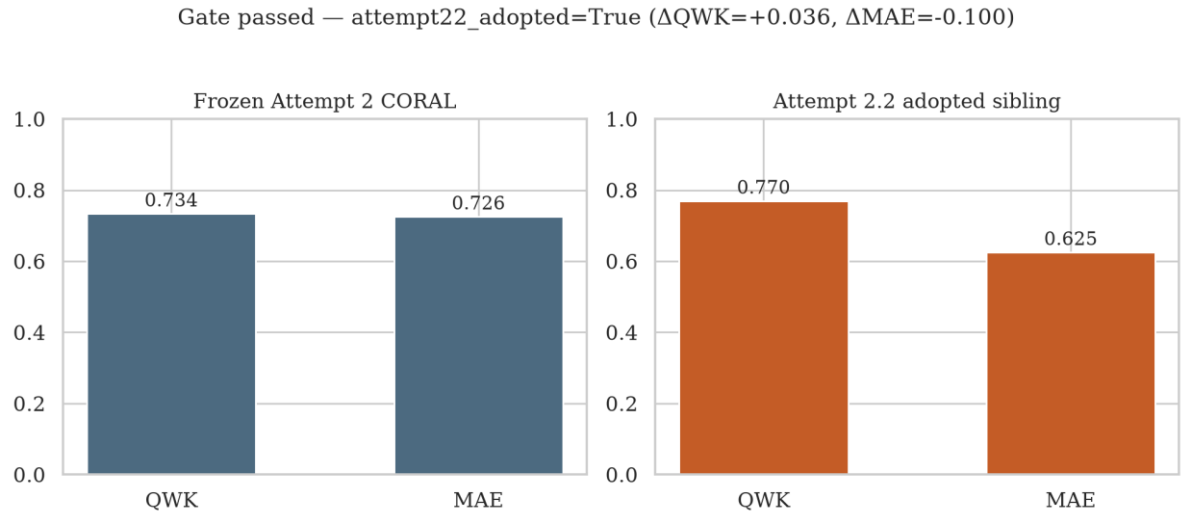


Fig 10. Attempt 2.2 adoption-gate metrics versus the Attempt 2 extended-binary CatBoost baseline.

K-Fold Stability

Model	QWK mean±sd	MAE mean±sd
attempt22_full	0.706±0.004	0.771±0.008
attempt2_baseline	0.678±0.003	0.831±0.014

Random-split K-fold means stay below the optimistic single holdout (≈0.706±0.005 versus ≈0.678±0.004) and still share incidents. They do not replace grouped leave-one-fire-out QWK = 0.480.

SHAP Interpretability

Head-3 SHAP (P(damage > 2))[34] elevates PC2, combustible siding, and 100 m density above tile (rank 14) and enclosed eaves (rank 18). That is the rank clash already reported, not a second beeswarm claim that roof and eaves dominate. SHAP values are local attributions on one cumulative head and are not causal structural effects. Figure 11 is the head-3 SHAP beeswarm for the adopted ensemble.

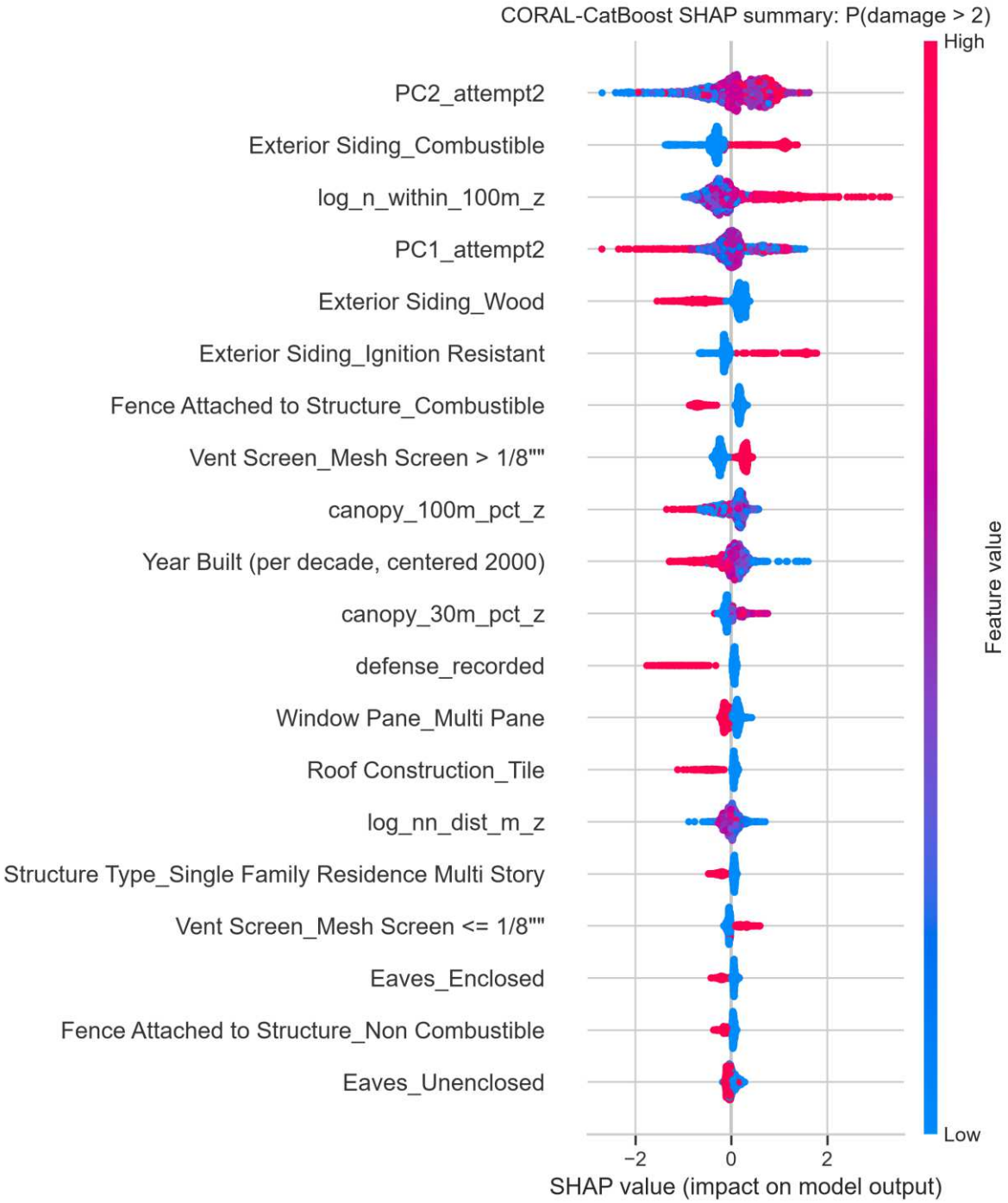


Fig 11. SHAP beeswarm summary for the adopted extended-binary CatBoost predictive architecture.

Packages and Learned Interactions

Mean CatBoost interaction strengths rank unexpected environmental–structural coupled pairs. Top Attempt 2 pairs are dominated by siding categories interacting with environmental PC2/PC1 (for example, wood×PC2 mean strength ≈ 0.909; ignition-resistant×PC2 ≈ 0.737; combustible×PC2 ≈ 0.726), followed by roof and vent couplings. Package compounding analyses show that concurrent hardened domains associate with lower cumulative odds on average (high-package OR ≈ 0.667; roof-hardened OR ≈ 0.795; eave-enclosed OR ≈ 0.845; glazing-multi OR ≈ 0.833), while some single-domain packages remain near null or adverse after

adjustment—consistent with package co-occurrence rather than guaranteed CWUIC compliance. Multi-component associational swaps reinforce the same non-flat story: standardizing to tile alone associates with $\Delta P(Y=4) \approx -0.092$, tile+enclosed eaves ≈ -0.110, and roof+eave+glazing ≈ -0.123, whereas fire-resistant–labeled wall swaps associate with $\Delta P(Y=4) \approx +0.140$ and several wall+vent combinations remain positive—another place where administrative "hardened" labels do not behave like portable engineering constants. Figures 12–15 document interactions, package compounding, pair panels, and package×vintage forests.

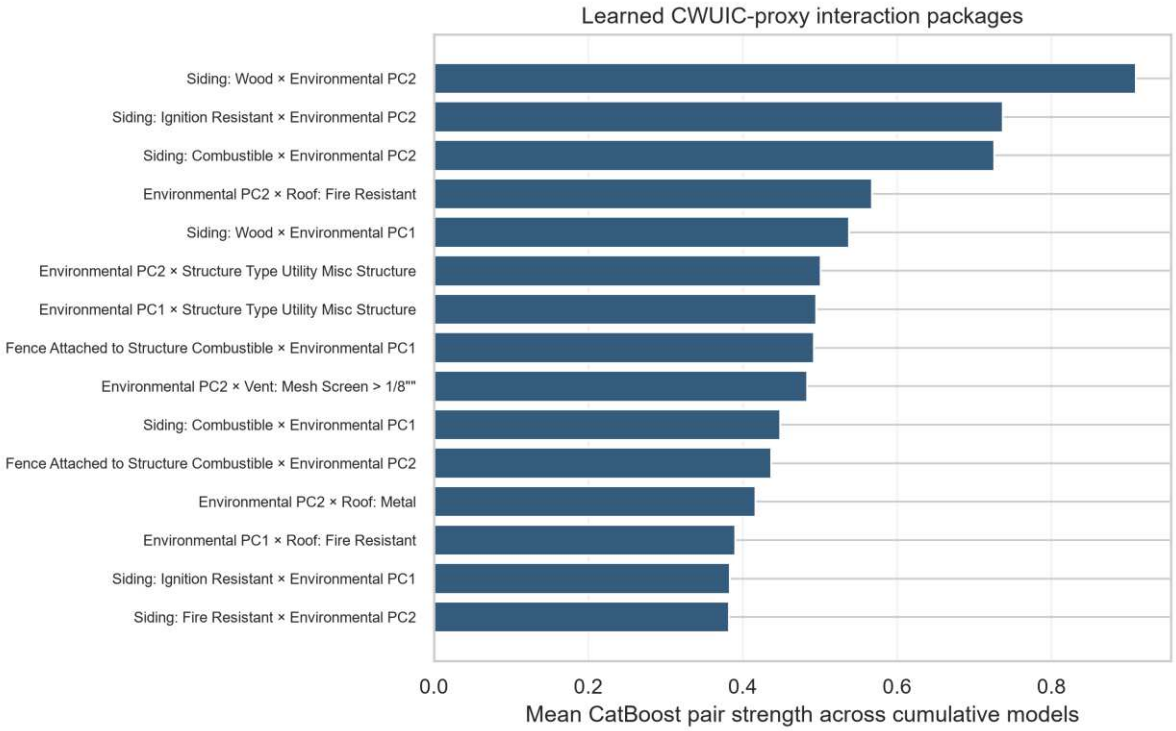


Fig 12. Top learned CatBoost interaction strengths between CWUIC features and environmental/context proxies.

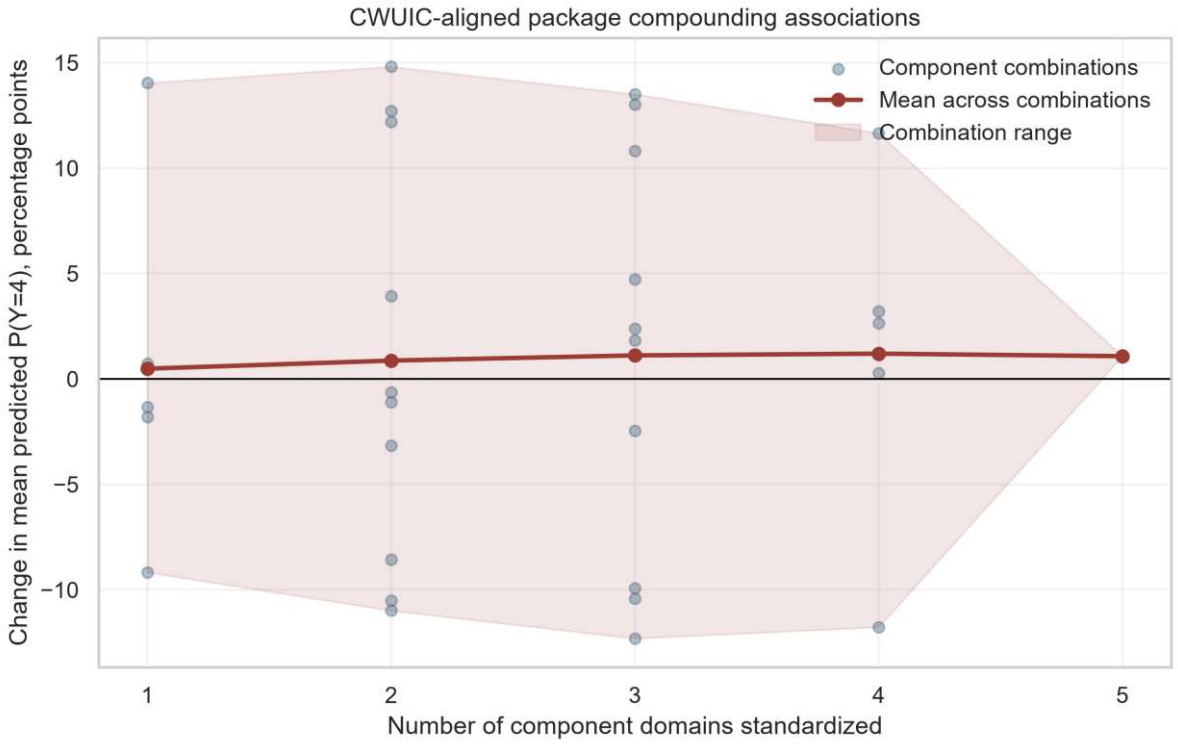


Fig 13. CWUIC-aligned package prevalence and compounding associations with ordinal damage.

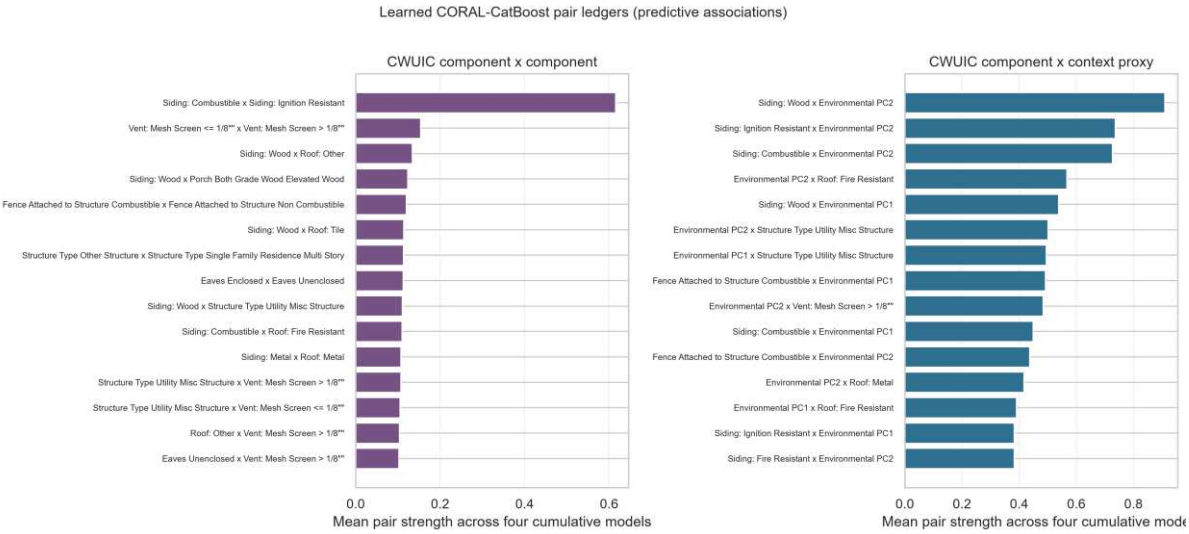


Fig 14. Full learned-pair panels for CWUIC×CWUIC and CWUIC×proxy interaction ledgers.

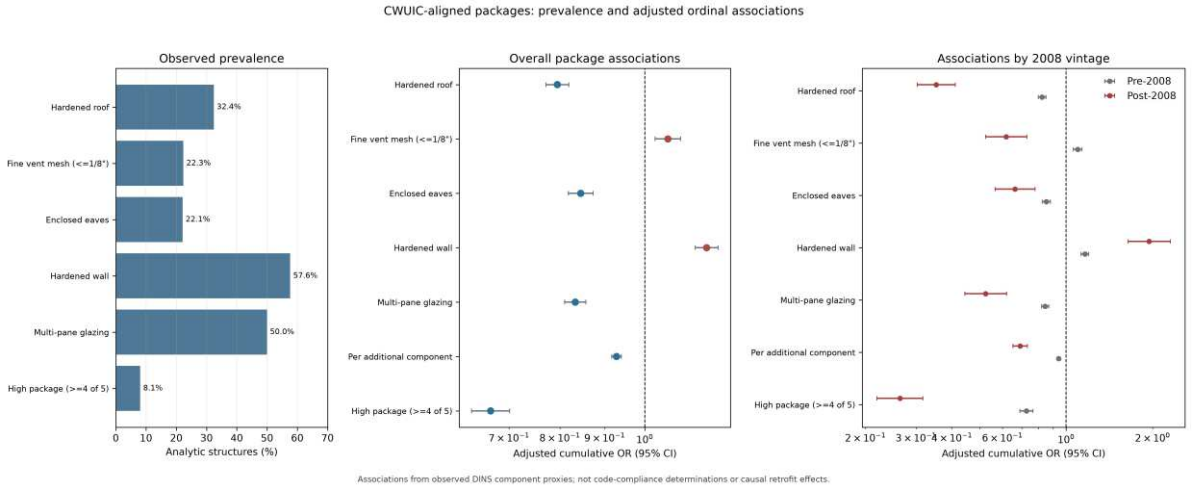


Fig 15. Package odds-ratio forest plot stratified by construction vintage.

Anomaly Phenotypes

Packed low-canopy survivors (n = 289) and sparse hardened losses (n = 379) are designed residual phenotypes on the random validation fold: survivors are Y = 0 with top-decile expected damage among survivors, at least one combustible feature, high density, and WRC Direct or Indirect; losses are Y = 4 with bottom-decile expected damage among losses and at least two hardened features. Survivors live in three fires (65% CALAC 009087, 93% South, mean 43 neighbors, canopy 4.5%, year built 1947). High density is required by the rule; 232 of 289 are WRC Indirect. Losses scatter across 43 fires (mean 11 neighbors, more hardened). Canopy is lower among survivors, so this is not a defensible-space success story and not a statewide WUI census. They are discovery phenotypes; the same definition—not the same 289 IDs—is the transport claim. The defensible-space and land-management literature is the contrast, not a measured mechanism in these 289 rows[35],[36]. This is not a Fort McMurray-style survivor-mechanism inventory[37].

Figure 16 profiles those two residual phenotypes. Figure 17 shows the Fire Hazard Severity Zone sensitivity checks around the same phenotypes.

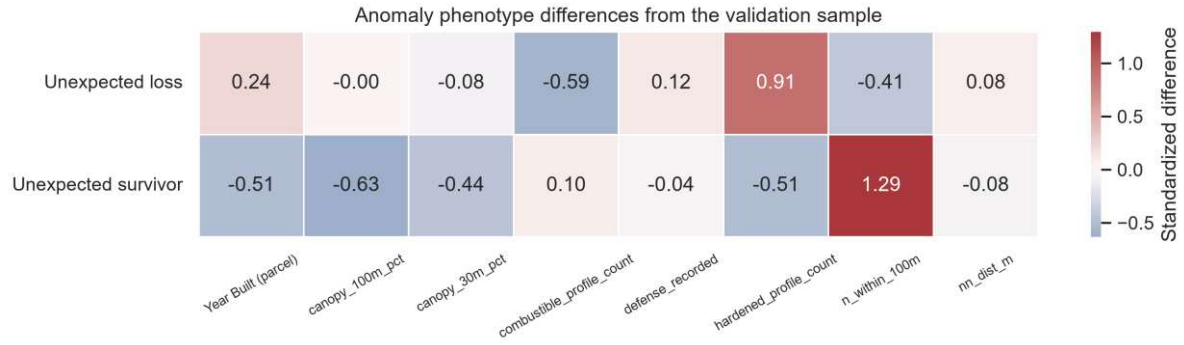


Fig 16. Phenotype profiles for 289 unexpected survivors and 379 unexpected total losses on the validation partition.

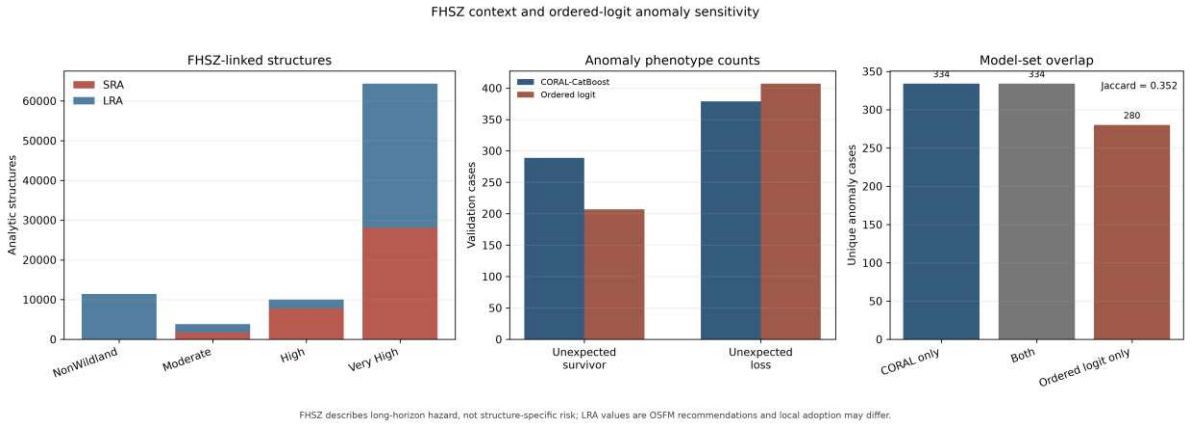


Fig 17. Fire Hazard Severity Zone sensitivity checks around anomaly phenotypes.

Validation, Holm Failures, and External BH

Holm-9 leaves three confirmatory terms unstarred: metal, density, and year built ($p = 0.052$). Those are reported, not suppressed. Kitchen-sink Holm-56 still fails porch, metal siding, and Window Pane Other; Attempt 2.3 external BH finds 12 discovery failures and 4 sign conflicts. Sparse logit instability is 2 of 56 terms. Figures 18–20 show connection proximity, material-partner predictive value, and block leave-one-out value.

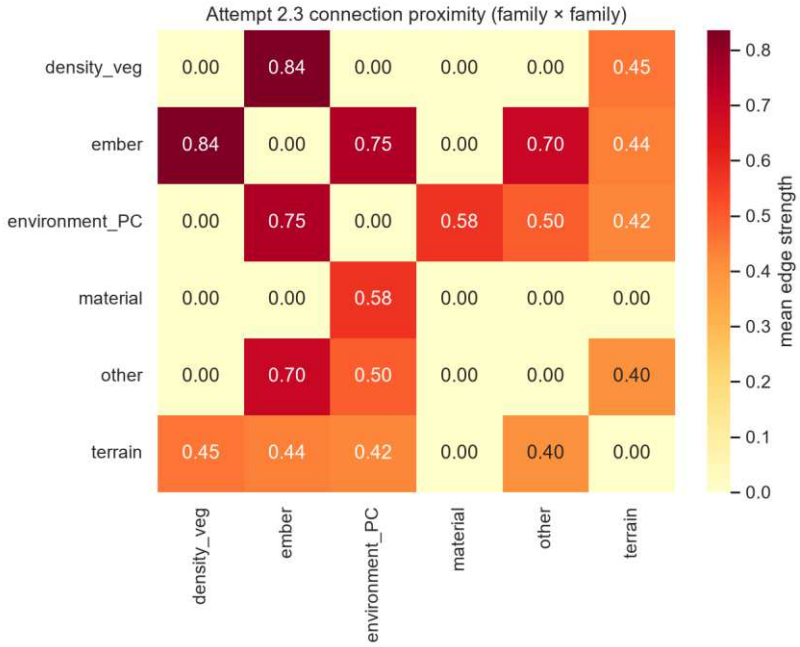


Fig 18. Descriptive connection-proximity map among materials, ember proxies, terrain, and density blocks.

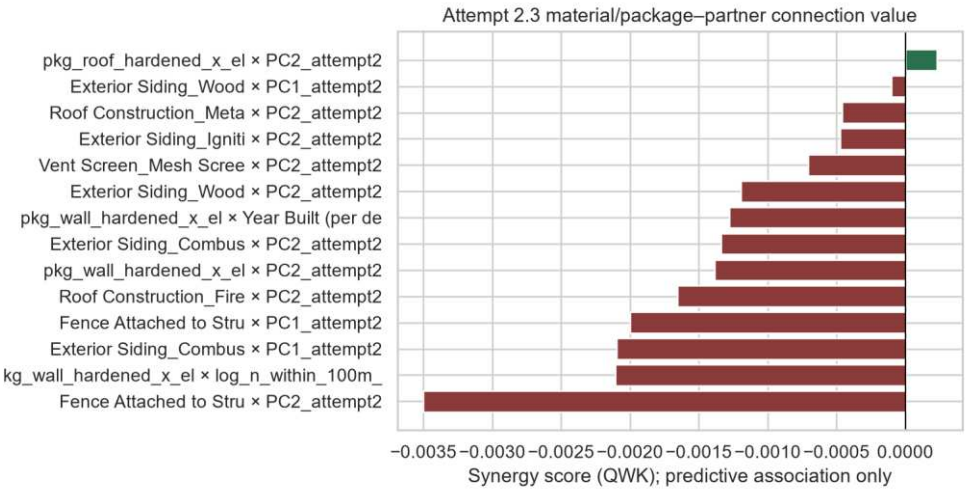


Fig 19. Predictive connection-value units for material–partner pairings (ΔQWK/MAE diagnostics, not causal effects).

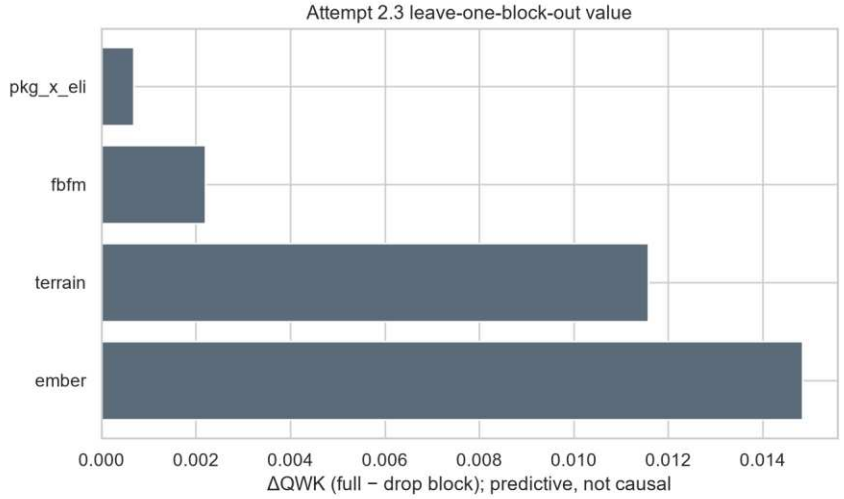


Fig 20. Block leave-one-out connection value: which feature blocks add versus detract predictive signal.

Discussion

Taken together, the clustered logit still only really defends a small set of post-inspection associations (tile, classifiable enclosed eaves, wood siding, canopy, WRC edge after crowding, recorded defense). The extended-binary CatBoost is more of an endpoint detector. Its extra in-fire accuracy comes from environment × packing × siding bends and from residual houses the linear predictor misses. Grouped leave-one-fire-out QWK ≈ 0.48 is the honest transport number. Random-split QWK 0.73–0.77 is not.

Connection to objectives. Holm-9, rather than Holm-56, is the inferential object: tile OR = 0.55, enclosed eaves 0.74, wood 0.57, WRC Direct 1.76, defense 0.62. The ML objective is the four clustered siding × density strata, the SHAP–OR rank clash, and the named residual phenotypes (289 in three fires; 379 in 43). It is not a second odds-ratio table, and it is not a claim that QWK 0.77 travels.

First, combustible-fence OR = 0.445 stays in the measurement audit. Second, Year Built remains a confounder bundle. Third, California is not one fire regime[7] due to the fact that density and WRC destroyed rates shift North versus South while tile does not. Fourth, three incidents

own 58% of the sample, which is why mega-fires were reserved for leave-one-fire-out rather than used as "the" test set.

Implications remain hypothesis-generating. Tile is the most transportable hardening association. Enclosed versus open eaves is not a practical contrast. Combustible siding looks worse as packing rises; wood's protective-direction association is stronger in dense tertiles. Density is a between-fire, region-heterogeneous gradient whose cluster interval still includes 1. Residual packed low-canopy survivors are a Southern mega-fire phenotype rather than a vegetation-management result. None of these are causal retrofit guarantees.

Recommendations remain hypothesis-generating: prioritize envelope assemblies with protective-direction associations after controls, package co-occurrence over single-feature checklists, density/defense-aware mitigation where scenario bars shift strongly, and case review of anomaly phenotypes. None of these are causal retrofit guarantees.

Limitations include post-inspection DINS labeling (especially eaves, vents, and fences), DINS-coordinate density rather than a footprint census, long-run WRC and WildEST products rather than event flux, sparse middle damage classes, head-crossing thresholds on the extended-binary ensemble, SHAP from one head, and a burned-area sample rather than a WUI housing census. External footprints, outcome-aware multiple imputation, and shared-weight CORAL were not rebuilt. The models do not estimate social-vulnerability gradients documented in the hazards literature[38],[39]. They also do not replace community-scale risk-management or fire-behavior simulation studies[40],[41].

Closing thought. What survives Holm-9, what bends with packing, and which houses the linear model cannot see together form an association-safe blueprint. They do not dissolve uncertainty into a 0.77 kappa.

Conclusion

This study looked at 89,826 fire-impacted structures inspected after California wildfires (burned-area DINS sample, not a WUI census). Among a pre-specified confirmatory set, tile roofs, classifiable enclosed eaves, wood siding, local canopy, WRC Direct exposure, and recorded defense retain cluster-robust associations after Holm-9. Metal does not. Density's interval includes 1 and does not travel North–South. Extended-binary CatBoost finds density-gated siding bends and two residual phenotypes—packed low-canopy survivors in three fires, sparse hardened losses in 43—without supporting a transportable QWK of 0.73. Grouped leave-one-fire-out QWK ≈ 0.48 is the predictive headline. Findings stay hypothesis-generating for code evaluation rather than causal mandates.

Ethics

This study used publicly available CAL FIRE Damage Inspection (DINS) structural inspection data[25]. No human subjects experiments were conducted. The Declaration of Helsinki is not applicable to this non-clinical observational analysis of building inspection records. No animal experiments were performed.

Availability of data and materials

Analytic rows are the public CAL FIRE Damage Inspection (DINS) release (California Department of Forestry and Fire Protection, https://data.ca.gov/dataset/cal-fire-damage-inspection-dins-data, 2024). Every figure cited in the text is embedded in this file. Separate PNG files are supplied with the submission.

Authors' contributions

P. Elaprolu designed the study, conducted the analysis, and wrote the manuscript.

Competing interests

The author declares no competing interests.

Funding

The author received no specific funding for this work.

Consent to participate

Not applicable. No human participants were enrolled.

Consent for publication

Not applicable. The manuscript reports building, inspection, and program records and contains no identifiable persons.

AUC: area under the curve; BH: Benjamini–Hochberg; extended-binary CatBoost: cumulative threshold ensemble; CWUIC: California Wildland–Urban Interface Code; DINS: Damage Inspection; ELI/SELB/SES: WildEST ember-load proxies; FDR: false discovery rate; HB: Holm–Bonferroni; KMO: Kaiser–Meyer–Olkin; MAE: mean absolute error; OR: odds ratio; OvR: one-versus-rest; PCA: principal component analysis; QWK: quadratic weighted kappa; SHAP: SHapley Additive exPlanations; VIF: variance inflation factor; WRC: Wildfire Risk to Communities; WUI: wildland–urban interface.

References

1. California Department of Forestry and Fire Protection. Fire hazard severity zones. https://osfm.fire.ca.gov/what-we-do/community-wildfire-preparedness-and-mitigation/fire-hazard-severity-zones, 2025.

2. V. C. Radeloff, D. P. Helmers, H. A. Kramer, M. H. Mockrin, P. M. Alexandre, A. Bar-Massada, V. Butsic, T. J. Hawbaker, S. Martinuzzi, A. D. Syphard, S. I. Stewart. Rapid growth of the US wildland–urban interface raises wildfire risk. Proceedings of the National Academy of Sciences. Vol. 115, pg. 3314–3319, 2018, https://doi.org/10.1073/pnas.1718850115.

3. S. Mulkey. USA: 1 in 8 Californians live in the most dangerous wildfire zones. https://www.preventionweb.net/news/1-8-californians-live-most-dangerous-wildfire-zones, 2025.

4. H. A. Kramer, M. H. Mockrin, P. M. Alexandre, V. C. Radeloff. High wildfire damage in interface communities in California. International Journal of Wildland Fire. Vol. 28, pg. 641–650, 2019, https://doi.org/10.1071/WF18108.

5. H. A. Kramer, M. H. Mockrin, P. M. Alexandre, S. I. Stewart, V. C. Radeloff. Where wildfires destroy buildings in the US relative to the wildland–urban interface and national fire outreach programs. International Journal of Wildland Fire. Vol. 27, pg. 329–341, 2018, https://doi.org/10.1071/WF17135.

6. S. Li, T. Banerjee. Spatial and temporal pattern of wildfires in California from 2000 to 2019. Scientific Reports. Vol. 11, pg. 8779, 2021, https://doi.org/10.1038/s41598-021-88131-9.

7. J. E. Keeley, A. D. Syphard. Twenty-first century California, USA, wildfires: fuel-dominated vs. wind-dominated fires. Fire Ecology. Vol. 15, pg. 24, 2019, https://doi.org/10.1186/s42408-019-0041-0.

8. A. Ganteaume, A. Camia, M. Jappiot, J. San-Miguel-Ayanz, M. Long-Fournel, A. Lampin. A review of the main driving factors of forest fire ignition over Europe. Environmental Management. Vol. 51, pg. 651–662, 2013, https://doi.org/10.1007/s00267-012-9961-z.

9. A. D. Syphard, H. Rustigian-Romsos, J. E. Keeley. Multiple-scale relationships between vegetation, the wildland–urban interface, and structure loss to wildfire in California. Fire. Vol. 4, pg. 12, 2021, https://doi.org/10.3390/fire4010012.

10. A. D. Syphard, J. E. Keeley. Factors associated with structure loss in the 2013–2018 California wildfires. Fire. Vol. 2, pg. 49, 2019, https://doi.org/10.3390/fire2030049.

11. M. Manes, D. Rush. Meta-analysis of UK, USA and New Zealand fire statistics databases with respect to damage and financial loss. In *Applications of fire engineering (ASFE 2017)*, pg. 179–188. CRC Press, 2017, https://doi.org/10.1201/9781315107202-20.

12. A. D. Syphard, J. E. Keeley, A. B. Massada, T. J. Brennan, V. C. Radeloff. Housing arrangement and location determine the likelihood of housing loss due to wildfire. PLOS ONE. Vol. 7, pg. e33954, 2012, https://doi.org/10.1371/journal.pone.0033954.

13. T. D. Penman, L. Collins, A. D. Syphard, J. E. Keeley, R. A. Bradstock. Influence of fuels, weather and the built environment on the exposure of property to wildfire. PLOS ONE. Vol. 9, pg. e111414, 2014, https://doi.org/10.1371/journal.pone.0111414.

14. S. L. Quarles, Y. Valachovic, G. M. Nakamura, G. A. Nader, M. J. De Lasaux. Home survival in wildfire-prone areas: building materials and design considerations. University of California Agriculture and Natural Resources Publication 8393, 2010, https://doi.org/10.3733/ucanr.8393.

15. J. D. Cohen. Preventing disaster: home ignitability in the wildland–urban interface. Journal of Forestry. Vol. 98, pg. 15–21, 2000, https://doi.org/10.1093/jof/98.3.15.

16. California Building Standards Commission. California Wildland–Urban Interface Code (CWUIC), Title 24, Part 7, 2025.

17. S. E. Caton, R. S. P. Hakes, D. J. Gorham, A. Zhou, M. J. Gollner. Review of pathways for building fire spread in the wildland urban interface part I: exposure conditions. Fire Technology. Vol. 53, pg. 429–473, 2017, https://doi.org/10.1007/s10694-016-0589-z.

18. R. S. P. Hakes, S. E. Caton, D. J. Gorham, M. J. Gollner. A review of pathways for building fire spread in the wildland urban interface part II: response of components and systems and mitigation strategies in the United States. Fire Technology. Vol. 53, pg. 475–515, 2017, https://doi.org/10.1007/s10694-016-0601-7.

19. J. D. Cohen. Relating flame radiation to home ignition using modeling and experimental crown fires. Canadian Journal of Forest Research. Vol. 34, pg. 1616–1626, 2004, https://doi.org/10.1139/x04-049.

20. S. L. Manzello, S. Suzuki, M. J. Gollner, A. C. Fernandez-Pello. Role of firebrand combustion in large outdoor fire spread. Progress in Energy and Combustion Science. Vol. 76, pg. 100801, 2020, https://doi.org/10.1016/j.pecs.2019.100801.

21. S. Dossi, B. Messerschmidt, L. M. Ribeiro, M. Almeida, G. Rein. Relationships between building features and wildfire damage in California, USA and Pedrógão Grande, Portugal. International Journal of Wildland Fire. Vol. 32, pg. 296–312, 2023, https://doi.org/10.1071/WF22095.

22. E. E. Knapp, Y. S. Valachovic, S. L. Quarles, N. G. Johnson. Housing arrangement and vegetation factors associated with single-family home survival in the 2018 Camp Fire, California. Fire Ecology. Vol. 17, article 25, 2021, https://doi.org/10.1186/s42408-021-00117-0.

23. M. Zamanialaei, D. San Martin, M. Theodori, D. M. J. Purnomo, A. Tohidi, C. Lautenberger, Y. Qin, A. Trouvé, M. Gollner. Fire risk to structures in California's wildland–urban interface. Nature Communications. Vol. 16, pg. 8041, 2025, https://doi.org/10.1038/s41467-025-63386-2.

24. R. M. Fields, E. C. Fischer, C. J. Dunn. Still standing: a synthesis of local characteristics influencing housing survivability during WUI wildfires. International Journal of Wildland Fire. Vol. 35, pg. WF25222, 2026, https://doi.org/10.1071/WF25222.

25. California Department of Forestry and Fire Protection. CAL FIRE Damage Inspection (DINS) Data. https://data.ca.gov/dataset/cal-fire-damage-inspection-dins-data, 2024.

26. H. F. Kaiser. An index of factorial simplicity. Psychometrika. Vol. 39, pg. 31–36, 1974, https://doi.org/10.1007/BF02291575.

27. P. McCullagh. Regression models for ordinal data. Journal of the Royal Statistical Society: Series B. Vol. 42, pg. 109–127, 1980, https://doi.org/10.1111/j.2517-6161.1980.tb01109.x.

28. P. M. Alexandre, M. H. Mockrin, S. I. Stewart, R. B. Hammer, V. C. Radeloff. Rebuilding and new housing development after wildfire. International Journal of Wildland Fire. Vol. 24, pg. 138–149, 2015, https://doi.org/10.1071/WF13197.

29. W. Cao, V. Mirjalili, S. Raschka. Rank consistent ordinal regression for neural networks with application to age estimation. Pattern Recognition Letters. Vol. 140, pg. 325–331, 2020, https://doi.org/10.1016/j.patrec.2020.11.008.

30. L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, A. Gulin. CatBoost: unbiased boosting with categorical features. Advances in Neural Information Processing Systems. Vol. 31, 2018, https://proceedings.neurips.cc/paper/2018/hash/14491b756b44a0a9d41a5c1e22a7b0d3-Abstract.html.

31. S. Holm. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics. Vol. 6, pg. 65–70, 1979.

32. Y. Benjamini, Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B. Vol. 57, pg. 289–300, 1995, https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

33. M. P. Thompson, D. E. Calkin. Uncertainty and risk in wildland fire management: a review. Journal of Environmental Management. Vol. 92, pg. 1895–1909, 2011, https://doi.org/10.1016/j.jenvman.2011.03.015.

34. S. M. Lundberg, S.-I. Lee. A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems. Vol. 30, 2017, https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html.

35. A. D. Syphard, T. J. Brennan, J. E. Keeley. The role of defensible space for residential structure protection during wildfires. International Journal of Wildland Fire. Vol. 23, pg. 1165–1175, 2014, https://doi.org/10.1071/WF13158.

36. P. Gibbons, L. van Bommel, A. M. Gill, G. J. Cary, D. A. Driscoll, R. A. Bradstock, E. Knight, M. A. Moritz, S. L. Stephens, D. B. Lindenmayer. Land management practices associated with house loss in wildfires. PLOS ONE. Vol. 7, pg. e29212, 2012, https://doi.org/10.1371/journal.pone.0029212.

37. A. Westhaver. Why some homes survived: learning from the Fort McMurray wildland/urban interface fire disaster. Institute for Catastrophic Loss Reduction Research Paper Series No. 56, 2017. (peer-reviewed research report used for survivor mechanisms).

38. S. L. Cutter, B. J. Boruff, W. L. Shirley. Social vulnerability to environmental hazards. Social Science Quarterly. Vol. 84, pg. 242–261, 2003, https://doi.org/10.1111/1540-6237.8402002.

39. S. Masri, E. Scaduto, Y. Jin, J. Wu. Disproportionate impacts of wildfires among elderly and low-income communities in California from 2000–2020. International Journal of Environmental Research and Public Health. Vol. 18, pg. 3921, 2021, https://doi.org/10.3390/ijerph18083921.

40. D. E. Calkin, J. D. Cohen, M. A. Finney, M. P. Thompson. How risk management can prevent future wildfire disasters in the wildland–urban interface. Proceedings of the National Academy of Sciences. Vol. 111, pg. 746–751, 2014, https://doi.org/10.1073/pnas.1315088111.

41. M. A. Finney, C. W. McHugh, I. C. Grenfell, K. L. Riley, K. C. Short. A simulation of wildfire behavior and community vulnerability for California. Forest Ecology and Management. Vol. 294, pg. 39–52, 2013.