# Measuring the wildfire risk of California real estate with spatiotemporal convolutional neural networks*

Paulo Issler† Richard Stanton‡ Nancy Wallace§ Yao Zhao¶

March 6, 2025

**Abstract**

This paper uses spatiotemporal convolutional neural networks (ST-CNNs) to forecast wildfire risk across the state of California. ST-CNNs capture spatial and temporal dependencies interactively in spatiotemporal panel data like wildfires. We find that ST-CNN significantly outperforms both logistic regression and LSTM, a baseline ML model for time-series forecasting. ST-CNN successfully recalls the largest wildfires and corresponds well to the aggregate zip-code risks of California wildfires. Our data include 6,999 wildfires in California from 2001 through 2021 and their effects on the universe of over 7 million single-family residences. Out-of-sample estimates of the probability of wildfire and aggregate losses to the residential tax basis provide disturbing evidence concerning wildfire risks to California homeowners, lenders, and California's largest Property and Casualty (P&C) insurer, State Farm General. Finally, we discuss the urgent need for continued research on improving forward-looking wildfire-occurrence models since the California Department of Insurance approved their use in December 2024.

JEL classification: G21, G22, G52, Q54.

Keywords: Housing, mortgages, climate risk, household finance.

*We are grateful for financial support from the Fisher Center for Real Estate and Urban Economics. We acknowledge helpful comments from Ali Kakhbod, Maryam Farboodi (discussant), and seminar and conference participants at the National AREUEA conference, May 2024; Energy Transition and Climate Change Conference, Redondo Beach, September, 2024; Consumer Finance: New Perspectives on Policies and Products, MIT Golub Center, October 2024; The Philadelphia Federal Reserve Bank, October 2024; Economic Summit on Climate Risk, Federal Housing Finance Agency, October 2024; Financial Stability: Energy Risks in a Time of Interconnectedness, Cleveland Federal Reserve Bank and the Office of Financial Research, November 2024; Advanced Financial Technologies Laboratory, Stanford Engineering, December 2024.

†Haas School of Business, U.C. Berkeley, 545 Student Services Building, Berkeley, CA 94720-1900. Email: pauloissler@berkeley.edu.

‡Haas School of Business, U.C. Berkeley, 545 Student Services Building, Berkeley, CA 94720-1900. Email: rhstanton@berkeley.edu.

§Haas School of Business, U.C. Berkeley, 545 Student Services Building, Berkeley, CA 94720-1900 Email: newallace@berkeley.edu.

¶Haas School of Business, U.C. Berkeley, 545 Student Services Building, Berkeley, CA 94720-1900 Email: yaozhao@haas.berkeley.edu.

# 1 Introduction

Wildfire risk threatens the California economy through increased greenhouse-gas emissions, loss of human life (both directly due to the fires and indirectly due to increased air pollution), and losses to real estate and infrastructure.[1] Paci et al. (2023) estimate that wildfires caused economic losses to the state averaging $117.4 billion per year between 2012 and 2021. Wildfires led to the destruction of more than 60,000 structures and to 302 civilian and firefighter fatalities in California between 2002 and 2021 (Safford et al., 2022). By itself, the 2018 Camp Fire, which burned Paradise, California, caused $27.7 billion in capital losses, $32.2 billion in health costs, $88.6 billion in indirect losses (Wang et al., 2021a), 85 deaths, and the destruction of 18,804 structures.[2]

In this paper, we measure vegetative wildfires using data from the California Department of Forestry and Fire Protection's (CAL FIRE) Fire and Resource Assessment Program (FRAP).[3] Figure 1 presents annual wildfire counts and area burned from 1979 to 2022. The average annual number of wildfires was 400 from 2020 to 2022, compared with 248 between 1979 and 2019.[4] Similarly, the annual average of 2,339,158 acres burned in 2020 and 2022 (Panel b) was 4.9 times the annual average during the prior two decades.[5] Finally, the distribution of fire burn areas from 1979 to 2022 (Panel c) shows a right skew, similar to that found by Diaz (2022).

To further underscore the distributional characteristics of the fire sizes, Figure 2 shows a QQ-plot of the quantiles of the burned areas against i) a Weibull distribution (Panel a); and ii) an exponential distribution (Panel b). For both distributions, their parameters are estimated by minimizing the mean squared error between the data and theoretical quantiles. The figure shows that the wildfire-size quantiles almost perfectly match those of a Weibull distribution with shape parameter of 0.233 — a heavy-tailed distribution. In contrast, the plot of the same data against the (thin-tailed) exponential distribution lies far from the 45-degree line. Despite the usual caveats of interpreting the graphical results of QQ-plots on finite samples (Thode, 2002), Figure 2 suggests that the California historical wildfire size

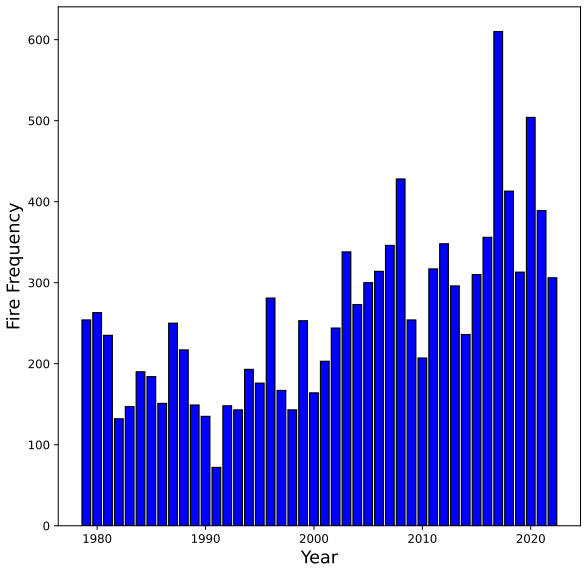
[1] See MacDonald et al. (2023); Safford et al. (2022).

[2] See https://placesjournal.org/article/paradise-redux-five-years-after-camp-fire/.

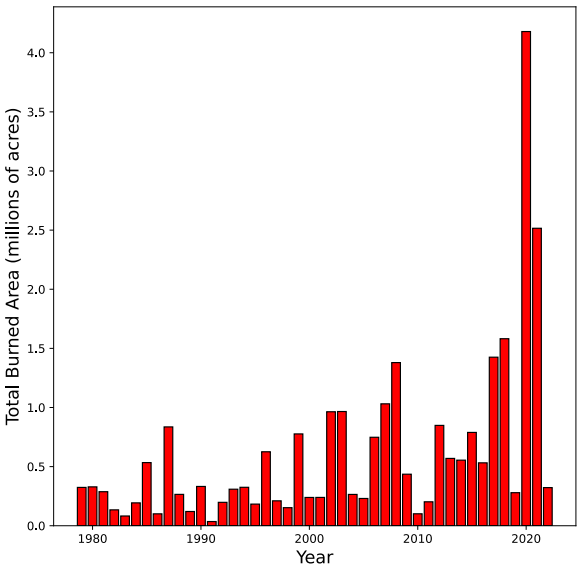
[3] The Fire and Resource Assessment Program (FRAP) distributes annual wildfire perimeter data sets for all public and private lands in California. CAL FIRE defines vegetative wildfires as burning a minimum area of 10 acres for timber fires, 30 acres for brush fires, and 300 acres for grass fires. The GIS data is developed with the cooperation of the United States Forest Service Region 5, the Bureau of Land Management, the National Park Service and the Fish and Wildlife Service (see https://www.fire.ca.gov/what-we-do/fire-resource-assessment-program/fire-perimeters).

[4] Buechi et al. (2021) found that the number of fires over the decade from 2009 to 2018 was 1.4 times the per-decade average between 1979 and 2009.

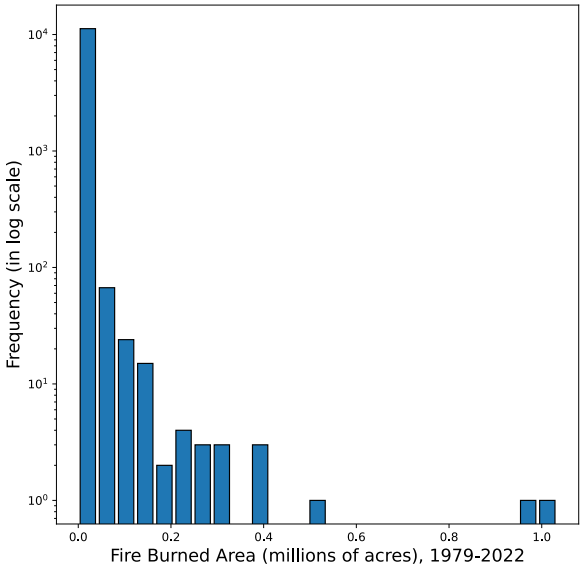
[5] Buechi et al. (2021) also found that the 7.08 million acres burned in the decade from 2009 to 2018 was 1.6 times larger than the per-decade average since 1979, and more than twice that from 1979 to 1988.




(a) Annual wildfire counts



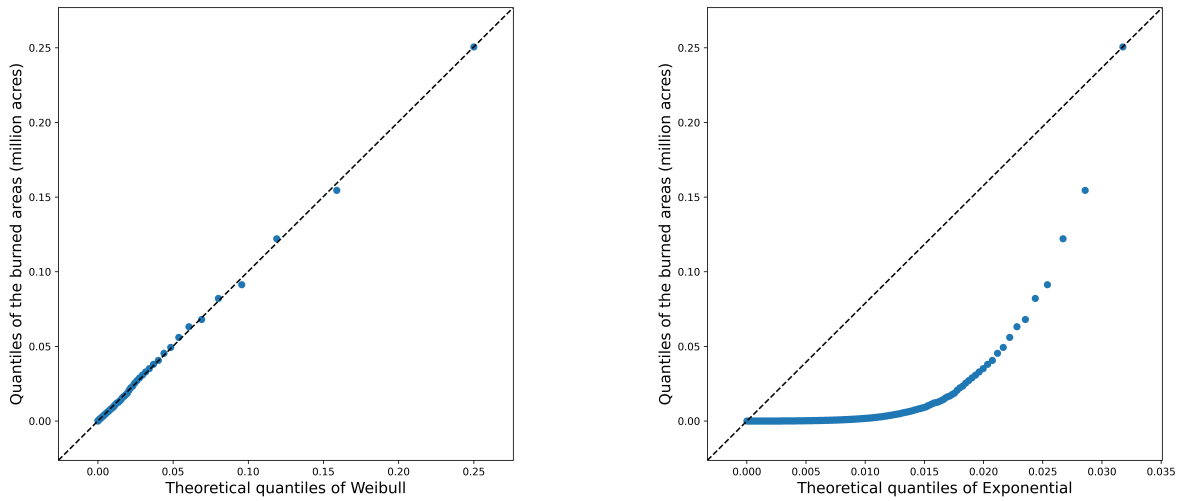

(b) Annual area burned (millions of acres)




(c) Distribution of burn areas

Figure 1: **Frequency and size of California wildfires, 1979–2022.** The wildfire incidence and burn area data are sourced from the California Department of Forestry and Fire Protection (CAL FIRE) (see https://www.frontlinewildfire.com/wildfire-news-and-resources/california-wildfires-history-statistics/)

distribution is likely to be characterized by a heavy-tailed distribution such as the Weibull. For heavy-tailed distribution, the historical averages are unreliable for prediction (Cooke et al., 2014). Thus, reliance on the backward looking historical simple average of wildfire incidence and losses, which has been the policy of the California Department of Insurance under Proposition 103 since 1989, would not be sufficiently informative for future predictions. In fact, a Weibull distribution with shape 0.25 is even more obese than a Pareto distribution with tail index 1, which has an infinite mean (Cooke et al., 2014).




(a) Weibull Distribution

(b) Exponential Distribution

Figure 2: **Historical fire burned area shows a heavy-tailed distribution.** We construct two QQ plots of all CAL FIRE identified fire burned areas between 1979 and 2022. The left figure is plotted against the theoretical quantiles of a Weibull distribution with shape parameter 0.233, which is a heavy-tailed distribution. The right one is plotted against an exponential distribution with rate parameter 217.46, which is a thin-tailed distribution. Note that a Weibull distribution with shape parameter 1 is equivalent to an exponential distribution, and the tails are fatter with smaller parameters. Additionally, the shape of the QQ-plot will not change with different parameters for the exponential distribution.

These results also suggest potential challenges to accurate statistical forecasting of California wildfires, for methods to diversify and securitize wildfire risks, for reserve strategies under Value-at-Risk management regimes, and for the design of risk management strategies due to spatial dependencies that affect many people, properties, and insurance lines simultaneously (see Abatzoglou and Williams, 2016; Joseph et al., 2019; Kousky, 2019; Kousky and Cooke, 2009; Li and Banerjee, 2021). As Kousky and Cooke (2009) and Cooke et al. (2014) point out, climate-change-related risk distributions such as these often have significant correlation

between variables (see also Flannigan et al., 2016; Voosen, 2024), they have heavier tails in which the probabilities of ever more serious damage decrease slowly relative to the extent of the damage, and tail dependencies in which bad outcomes are more likely to occur together. Moreover, recent research suggests that climate change may be directly fattening the tails of the distributions of many extreme events (Cardil et al., 2021; Koh et al., 2023).[6]

High temperatures and low precipitation play an important role in enhancing the flammability of vegetative fuels. Figure 3 shows the maximum annual temperature for the West Climate Region in the U.S. from 1895 to 2024. A pronounced overall increase can be seen, with a dramatic increase of 0.3 °C in 2023 alone. The consensus view among climate scientists is that global and regional temperatures increases of about 1.1 °C (2 °F) since 1980 are largely attributable to anthropogenic greenhouse gas emissions (Burke et al., 2021; MacDonald et al., 2023; Safford et al., 2022; Williams and Abatzoglu, 2016). Recent contributions to this debate have argued that the 2023 temperature increase is the result of decreases in the earth's albedo (the percentage of solar radiation reflected back into space)[7] in addition to ongoing anthropogenic effects and a three-year long la Niña followed by an especially strong el Niño in 2023 (see Goessling et al., 2025; Hodnebrog et al., 2024; Voosen, 2024).

Although there is a large statistical literature investigating these and other factors affecting wildfire risk, much of this literature uses logistic regression or related extensions (Xi et al., 2019). While machine-learning models have been used, their application is mainly limited to cross-sectional applications (see Casolaro et al., 2023; Chen et al., 2023; Makridakis et al., 2023). In this paper we estimate wildfire risk using spatiotemporal Convolutional Neural Networks (ST-CNNs), a technique that has been used in other fields such as image recognition and traffic flow forecasting, but has not previously been applied to climate modeling. ST-CNNs capture spatial and temporal dependencies interactively in spatiotemporal panel data like wildfires. We find that ST-CNN significantly outperforms logistic regression in estimating the likelihood of wildfire at a given point in space and time. Combining our fire-likelihood estimates with measures of the assessed value of houses in each area, we also estimate expected annual fire-related property-tax losses for each area.

The paper is organized as follows: Section 2 presents a discussion of the wildfire-related risks currently facing property and casualty insurance companies in California. Section 3 discusses the current state of wildfire modeling and presents our ST-CNN model. Section 4 discusses the data used in our analysis. Section 5 compares estimation results using ST-

[6]See also the research goals of *The U.S. Global Change Research Program 2022–2031 Strategic Plan* (https://downloads.globalchange.gov/strategic-plan/2022/USGCRP_2022-2031_Decadal_Strategic_Plan.pdf).

[7]The decline in earth's albedo appears to caused by reduced low-cloud cover in the northern mid-latitudes and tropics.

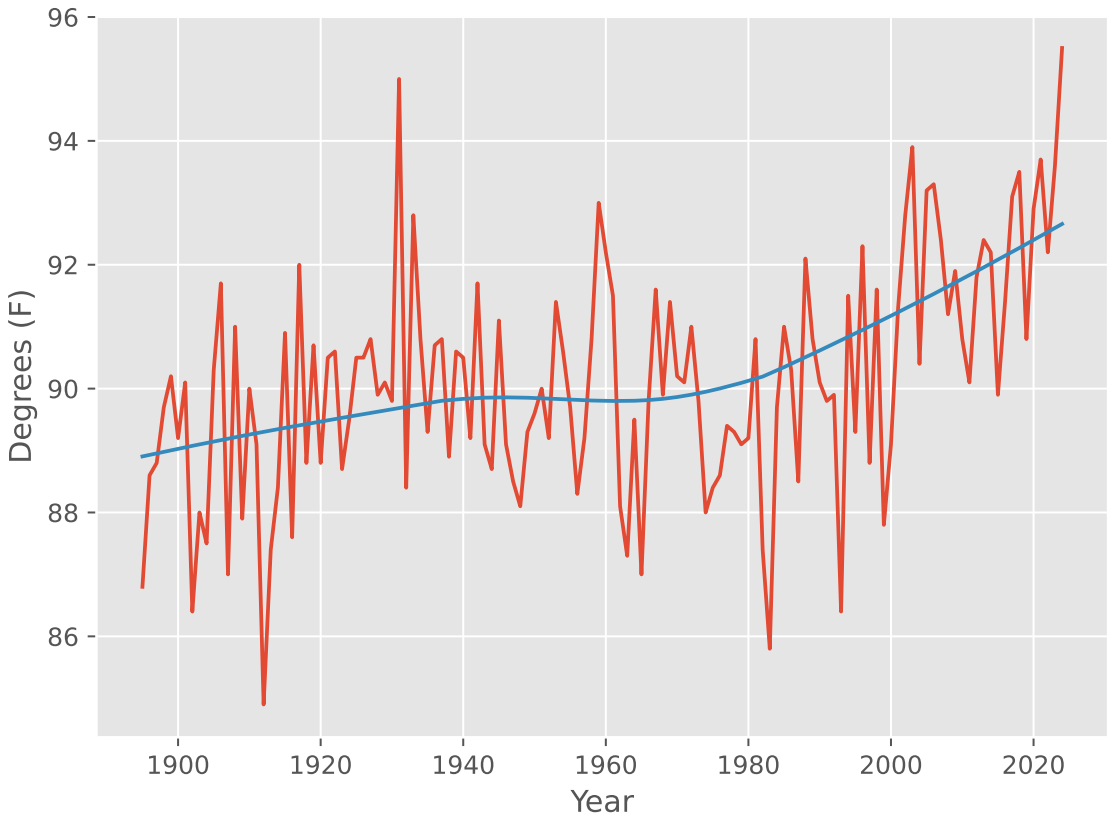



Figure 3: **Maximum annual temperature.** The red line shows the annual maximum temperature for the West Climate Region between 1895 and 2024 from the National Oceanographic and Atmospheric Administration (see https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/regional/time-series), with LOWESS trend line (Cleveland, 1979) in blue.

CNN with those from logistic regression and calculates expected property losses in the state. Section 6 discusses recent problems faced by the insurance company State Farm in light of our estimation results, and Section 7 concludes.

# 2 Property and casualty insurance and wildfire risk

The California property and casualty insurance industry is facing significant wildfire-related challenges. For an insurance company to stay solvent, it must have access to enough capital to pay losses even in catastrophic years. For non-disaster lines of insurance, such as automobile insurance, the premiums in any given year are enough to cover claims from that year. For fire insurance, however, claims may greatly exceed annual revenue, due to the heavily skewed burn-area distribution shown in Figure 1.[8] Property and casualty insurance firms underwriting wildfire risk in California must solve an intertemporal smoothing problem to cover their catastrophic loss years (see Jaffee and Russell, 2013). The likelihood of very significant losses associated with California wildfires, especially in the last five years, has required property and casualty insurance companies to build up reserves, purchase reinsurance, and use other insurance-linked securities to be able to pay their claims in high damage years (see Goss et al., 2020; Opitz, 2023). As shown in Figure 4, loss ratios for fire peril were significantly impacted by the heavy wildfire-loss years of 2017 and 2018. Only at the end of 2019, after two straight years of insurers paying out $1.85 in losses for every $1 of premium earned, did the California Department of Insurance approve 71 rate-increase requests from 50 different companies.[9]

There are also other internal capital market challenges for property casualty insurance companies that hinder solutions to their intertemporal smoothing problem. U.S. accounting requirements preclude earmarking capital surplus to a specific risk and current tax provision require that retained earnings are taxed as corporate income at set aside (Jaffee and Russell, 2013). Additionally, the accumulation of capital holdings to preclude losses in particularly catastrophic future years make these companies very susceptible to takeover, after which there is significant risk of policy non-renewal.[10] Another challenge has been the historical

[8]For example, in 2018 the California insurance industry collected $939,112,586 in property insurance premia and paid out $1,534,083,985 in incurred losses, a 164.22% loss ratio (https://www.insurance.ca.gov/01-consumers/120-company/04-mrktshare/2021/upload/PrmLssChartHistorical2021wa.pdf).

[9]See https://www.insurance.ca.gov/0250-insurers/0800-rate-filings/0100-rate-filing-lists/rate-filing-approvals/.

[10]Stein (1988, 1989) argues that managers of publicly traded companies may face incentives to reduce their company's reported earnings in the short run when they undertake policies that maximize the firm's long-term value. This may give well-informed traders the opportunity to take over the company at a low price, reflecting the low short-run earnings. This mechanism reinforces the takeover risk for an insurance company holding liquid assets.

refusal of the California Department of Insurance to allow the inclusion of reinsurance costs in the calculation of premia for wildfire insurance that could smooth these intertemporal risks.[11] As a result of these ongoing challenges, between 2012 and 2022 California homeowners insurance companies performed significantly worse than the national average on key risk metrics such as the direct incurred loss ratio[12] (73.9% compared with a U.S. average of 59.7%) and the average direct underwriting profit (−13.1% compared with a U.S. average of 3.6%).[13]

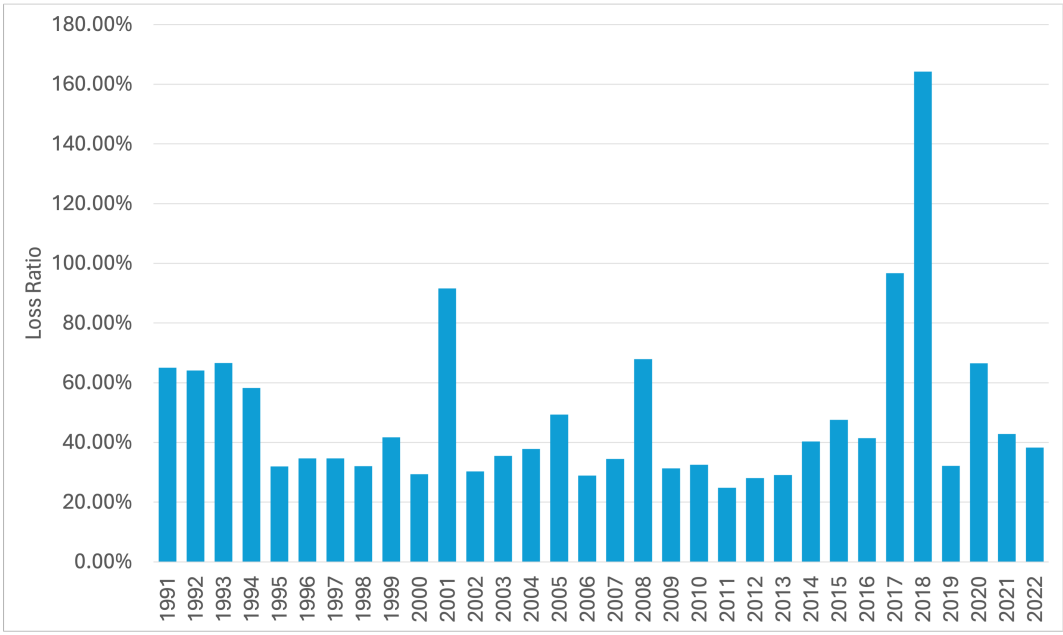



Figure 4: **Realized loss rates (fire peril) for California Property and Casualty insurance companies.** Source: https://www.insurance.ca.gov/01-consumers/120-company/04-mrktshare/2022/upload/PrmLssChartHistorical2022.pdf.

Rate-setting limitations are another major challenge for the long-term viability of homeowner fire-casualty insurance in California. In 1988 California voters passed Proposition 103, which required insurance companies to receive "prior approval" from the California Department of Insurance (CDI) before implementing property and casualty insurance rate increases. As shown by Oh et al. (2024), casualty insurance rates in states like California with high regulatory frictions have not adequately adjusted in response to the growth in losses. Until December 2024, California state insurance regulations required wildfire insurers to set rates

[11] See California Code of Regulations Title 10 § 2644.25, Reinsurance (https://govt.westlaw.com/calregs/Document/I29B16510D99911EFB9A3C8E71143034D.)

[12] The loss ratio is equal to the insurance claims paid plus adjustments, divided by the total earned premiums.

[13] http://insurance.ca.gov/0400-news/0100-press-releases/2023/upload/California-s-Sustainable-Insurance-Strategy-slides.pdf.

for future annual catastrophic coverage as the 20-year historical average fraction of realized damages, rather than based on statistical or actuarial models. Additionally, the CDI did not allow for the costs, or changes in the costs, of reinsurance risk to be included in insurer rate requests. As a result, California's annual rates now rank next to the lowest in the U.S. (see Oh et al., 2024), perhaps threatening the future ability of California homeowners to successfully rebuild and continue to make their mortgage payments after large and destructive wildfires.[14]

On September 21, 2023 California Governor Gavin Newsom issued an Executive Order to authorize the State Insurance Commissioner, Ricardo Lara, to exercise his authority to stabilize California's property insurance markets.[15] At the same time, the CDI introduced the California Sustainable Insurance Strategy, which will allow insurance carriers in the future to apply forward-looking catastrophe models to more accurately assess and price climate-related risks after December 2024 in exchange for expanded property insurance coverage in risky areas.[16] The insurance commissioner's office is currently developing regulations for how exactly the new models can be used for rate setting in the future and who will vet them.[17]

Boomhower et al. (2024) provide a rough summary measure of the relative granularity or complexity of the pricing algorithms used by California homeowners insurance companies. Given available data on the models, they build a measure of model complexity by counting the number of risk-rating variables that each insurer uses to assess the likelihood of wildfire damages for specific home locations. They find that some California insurers price wildfire risk using zip-code-level territory factors and others use parcel-level categorical wildfire risk scores based on qualitative factors such as slope, vegetation, fuel load, and road access. The larger insurers use more granular measures generated by using probabilistic catastrophe models. Overall, it is the firms with the largest market share in high-hazard zip codes that use the most granular risk segmentation. Boomhower et al. (2024) argue that the observed heterogeneity in California insurance price/risk modeling primarily reflects the direct costs of licensing or developing state-of-the-art wildfire models. These costs, together with indirect costs of adoption such as adapting the firm's internal systems and employing professional staff, can run to millions of dollars annually (see Jergler, 2021).

[14]Since 2022, AIG and Chubb have left the high-value home insurance market. State Farm, Farmers, Allstate, USAA, Travelers, and Nationwide have all either limited or paused writing new policies (https://www.insurance.ca.gov/0400-news/0100-press-releases/2023/upload/California-s-Sustainable-Insurance-Strategy-slides.pdf).

[15]https://www.gov.ca.gov/wp-content/uploads/2023/09/9.21.23-Homeowners-Insurance-EO.pdf.

[16]https://www.insurance.ca.gov/01-consumers/180-climate-change/SustainableInsuranceStrategy.cfm.

[17]https://www.politico.com/news/2023/09/21/newsom-orders-action-on-wildfire-insurance-00117488.

The effects of Proposition 103 and the California Department of Insurance's historical reluctance to allow probabilistic models to justify firm-level requests for rate increases, have led to the significant heterogeneity in the current state of statistical rate-setting technology found by Boomhower et al. (2024). The new policies of the California Department of Insurance should allow more active competition among firms to develop state-of-the-art wildfire modeling technology for wildfire rate pricing at the property level. The current challenge is that firm-level development of these newly allowed probabilistic wildfire risk models and the regulatory vetting of the models has not been a transparent process.

# 3 Wildfire modeling

The study of wildfire occurrence has led to large and growing statistical and machine-learning literatures focused on identifying risk factors and producing risk maps or indices.[18] These modeling strategies are best suited to the investigation of general trends across wildfires as a function of features that are predictive of when and where wildfire ignitions occur.

The biophysical variables found in occurrence models are intended to causally explain why wildfire ignitions vary across space and time as the result of temporal and spatial variations in weather, climate, vegetative land coverage, and topography (see Brinkmann et al., 2022; First Street Foundation, 2022; Kearns et al., 2022; Prestemon et al., 2013). Other common variables used in statistical occurrence modeling include anthropogenic features that affect the rising risks of wildfire due to the interaction of human activity and climate (see Abatzoglou et al., 2018; Abatzoglou and Williams, 2016; Apt et al., 2023; Williams and Abatzoglou, 2020; Williams and Abatzoglu, 2016) as well as the effects of the increasing encroachment of urban development into the wildland urban interface (see Alexandre et al., 2016; Kestelman, 2024; Price and Bradstock, 2014; Radeloff et al., 2018).

Historically, the most commonly used form of wildfire occurrence modeling has been logistic regression models or related extensions such as logistic generalized additive models (Xi et al., 2019). More recently, machine learning models such as gradient boosted random forest, neural networks, and deep neural network with multi-layer perceptrons have been applied to wildfire occurrence modeling (see Alkhatib et al., 2023; Ismail and Amarasoma, 2023; Jain et al., 2020; Tong and Gernay, 2023). Convolutional neural networks (CNNs)

[18]For an overview of the statistical literature, see Oliveira et al. (2021); Prestemon et al. (2013); Xi et al. (2019). For machine-learning, see Cruciata et al. (2024); First Street Foundation (2022); Koh et al. (2023); Opitz (2023). An even more recent literature, focused on physics-informed machine learning, has provided accurate and efficient ways of recognizing complex patterns and predicting spatiotemporal weather and climate processes that obey fundamental laws governing physical systems (Kashinath et al., 2020; Seydi et al., 2024).

have been very successfully used with drone-sourced-imagery technology (usually focused on color, motion, and geometric features) for efficient early-warning wildfire detection (see Bulatov and Leidinger, 2021; Georgiev et al., 2020; Jain et al., 2020; Jiao et al., 2020, 2019; Park et al., 2020; Shamsoshoara et al., 2021; Wang et al., 2021b; Zhao et al., 2018; Zheng et al., 2017). Zhang et al. (2019) found that the spatial correlation controls of a CNN model led to superior performance for forestry wildfire susceptibility mapping in Yunnan Province in China, however, their CNN specification was not a forward-looking model. Other than the physics-informed machine learning models, the current generation of machine learning methodologies are primarily cross-sectional and non-forward-looking models (see Casolaro et al., 2023; Chen et al., 2023; Makridakis et al., 2023).

## 3.1 Standardization of the wildfire event space

Among climate scientists, the presence or absence of wildfire occurrence is measured over discrete space-time cells, grids, that are projected onto the earth's surface. The raw weather and climate data are measured as satellite-data projection to the latitude and longitude of each grid point. Other measures such as vegetative land cover, urban density, or infrastructure measurement are also standardized to the grid-cell-level representations so as to harmonize the different spatial-temporal scales of wildfire and predictor data such as weather conditions, land cover and land use (Abatzoglou, 2013; Koh et al., 2023). Currently available climate space-time grid cells are sized at 4 km $\times$ 4 km or less for the dynamic hourly or daily climate and weather measures. The static predictors related to vegetative land coverage, topography, and housing density and electricity transmission lines among other measure are available at raster, or pixel levels, so that they can be easily merged to the grid-cell data.

Given the ready availability of economic and demographic indicators at the census tract or zip-code level, several recent papers in the economics literature use measures of wildfire occurrence at the zip-code or census-tract level (see Biswas et al., 2023; Kahn et al., 2024). A first concern with the use of census tracts or zip codes for wildfire occurrence measurement is that zip codes and census tracts are based on population not geography. The use of zip codes is especially a problem in California, where the average zip code is 60.12 square miles, the largest zip code is 1,773 square miles, and the smallest is 0.01 square miles. Additionally, zip codes are not spatially or temporally consistent with climatological, topographical, or vegetative measurement (see Abatzoglou, 2013). Secondly, any county-level or zip-code-level predictions may suffer from inaccuracy, because they will either incorrectly assume the entire area is burned, or rely on historical fire sizes that as shown in Section 1 are likely to be heavy-tailed. This again suggests that we should conduct cell-level predictions, because predicting

cell by cell with granular local data can avoid relying purely on historical average fire sizes.

## 3.2 Spatiotemporal Convolutional Neural Networks

Spatiotemporal Convolutional Neural Networks (ST-CNNs) are well suited to short-term forecasting problems due to their ability to automatically extract important spatial and temporal features from data without relying on hand-crafted features (Casolaro et al., 2023). Common applications of these models include still-image recognition and action recognition in videos (see Tran et al., 2018). The models have also been successfully applied to address the spatial correlations, temporal correlations, and heterogeneity of the traffic flow data used to forecast urban traffic congestion (Guo et al., 2019) and to the satellite detection of canopy-scale tree mortality and survival from California wildfires (see Dixon et al., 2023). However, to our knowledge they have not previously been applied to modeling the occurrence of wildfires from a forward-looking perspective.

Spatiotemporal CNN (ST-CNN) models are designed to capture spatial patterns. They are also effective in modeling temporal dependencies and identifying correlations between neighboring data points in spatiotemporal panel data such as wildfires, just as they can recognize objects in images by analyzing patterns in pixel values. By applying 3-dimensional convolutional filters across both space and time, ST-CNNs can learn the motion patterns in spatiotemporal panel data and fully use those patterns to account for how past values influence future predictions, for each grid cell. Factoring 3-dimensional convolutions into spatial and temporal components can further enhance accuracy and efficiency, leading to lower training and testing errors (Sra, 2019).

One specific strength of ST-CNNs for wildfire occurrence modeling is that they easily handle the cell adjacency correlation structure of wildfire — if there is a wildfire in one cell location, nearby cell locations are likely to also have wildfires. The ST-CNNs also allow for the temporal aggregation of some wildfire features by accounting for the cumulative effects of phenomena such as maximum temperature and vegetative dryness on the cell-by-cell occurrence of wildfires. Another important strength is the benefit of regularization. The fitted values from ST-CNNs are by nature correlated in space and time, which helps to prevent over-fitting even with a high-dimensional, nonlinear parameter space. These models are also parsimonious as neural network models, because the weight parameters are shared across space and across time, thus significantly reducing model complexity. As will be explained below, the weights in ST-CNNs are comparable to those in classical econometric models, and are therefore highly interpretable. Additionally, they can introduce nonlinearities into the network, thus allowing for the complex functions that are needed to accurately model

the joint spatial correlations and temporal dynamics of wildfire prediction.

For our application of ST-CNNs, we are forecasting one year ahead, since the typical maturity of an homeowners insurance contract is annual. However, longer-run out-of-sample forecasting remains a limitation for applications of ST-CNN to longer term contracting. Another potential drawbacks of using ST-CNN is that the model needs to be slightly customized to prohibit the use of future temporal features to forecast future wildfire occurrences. However, as will be discussed below, overall ST-CNNs offer a powerful approach to time-series forecasting, particularly for tasks where capturing both spatial and temporal dependencies is crucial. Their ability to automatically learn relevant features from raw data, combined with the flexibility offered by different spatiotemporal convolutional designs, makes them a valuable tool for predicting future values in various time-dependent domains.

## 3.3 ST-CNN and spatiotemporal panel data

Suppose we are given a panel dataset $\{(X_{it}, y_{it})\}_{i\in I, t\in T}$, where $y_{it}$ is the dependent variable observed for location $i$ at time $t$ and $X_{it} = (x_{it}^1, x_{it}^2, \cdots, x_{it}^k)'$ is the corresponding $k$-dimensional vector of explanatory variables. $I$ and $T$ denote the sets of location and timestamps, respectively. Assume we have $N$ locations in total, i.e., $|I| = N$. This is a common data structure in many location-based studies, such as real estate, climate, transportation, etc. For generalization purposes, we do not make specific assumptions about the data types of $y_{it}$. It can be either binary if it is a classification problem or continuous if it is a regression problem. Generally, we are trying to model the target variable $y_{it}$ as a function of $\{X_{is}\}_{i\in I, s\leq t}$, which contains all historical information across space up to time $t$.

**Notation** Stack the data across locations and define

- $\boldsymbol{y}_t = (y_{1t}, \cdots, y_{Nt})'$, an $N \times 1$ vector whose $i$-th entry is $y_{it}$ for location $i$,
- $\boldsymbol{X}_t = (X_{1t}, X_{2t}, \cdots, X_{Nt})'$, an $N \times K$ matrix, whose $i$-th row is a $k$-dimensional vector of characteristics for location $i$.

Now further stack the data across time and define

- $\boldsymbol{y} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_T)'$, a $T \times N$ matrix,
- $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_T)$, a tensor with dimension $T \times N \times K$.

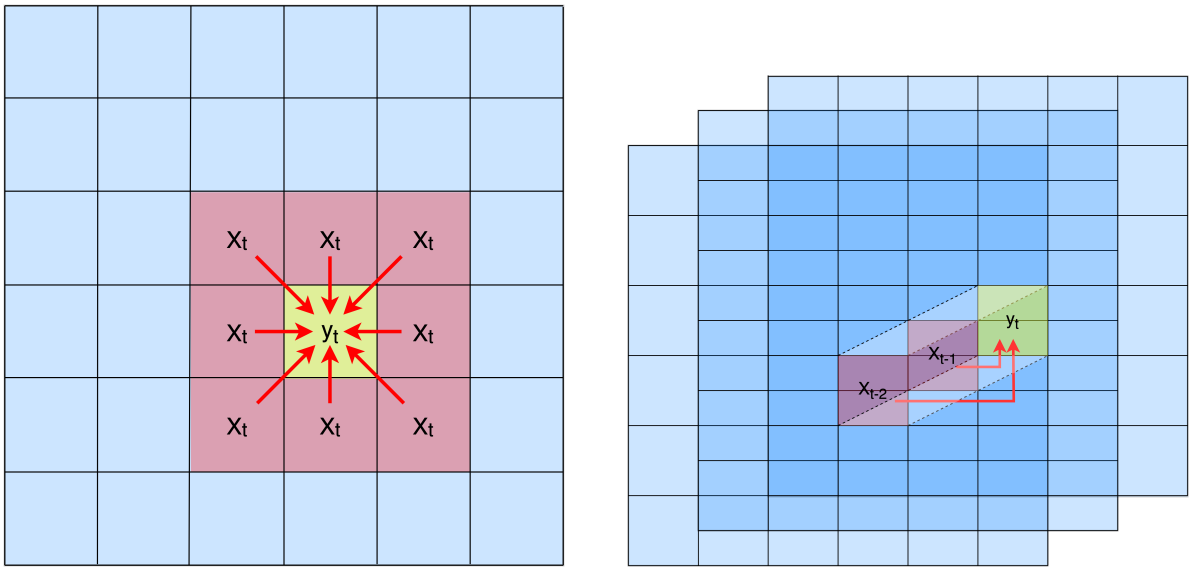
## 3.4 Spatial and temporal dependence

Intuitively, the simplest model to start with is OLS for regression or logistic model for binary classification, where we assume all observations across space and time are independent and

simply use the cross-sectional variations to fit $\boldsymbol{y}_t$ for any $t \in T$.

$$\boldsymbol{y}_t = \begin{cases} \sigma(\beta_0 \mathbf{1} + \boldsymbol{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t) & \text{if classification,} \\ \beta_0 \mathbf{1} + \boldsymbol{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t & \text{if regression,} \end{cases}$$

where $\sigma$ is the logistic function applied element-wise to the vector input, $\mathbf{1}$ is a vector of ones, $\boldsymbol{\beta}$ is the vector of coefficients, and $\boldsymbol{\varepsilon}_t$ is a vector of independent errors.




(a) Spatial dependence (b) Temporal dependence

Figure 5: **Visualizing the potential dependence structure in a spatiotemporal dataset.** The yellow cell refers to a target location-time, whose $y_{it}$ is our modeling goal. In addition to the yellow cell's own observations $X_{it}$, the characteristics of its nearby red cells are assumed to have an effect on the modeling goal as well, which will lead to spatial and temporal dependence. The red cells in Figure 5a are called the "spatial lags," while those in Figure 5b are the usual lags in time series. The remaining blue cells are assumed to have no impact on the yellow cell. In Figure 5b, each big square represents a slice of the panel data at a particular time.

Figure 5 demonstrates the potential dependence structure in space and time for a spatiotemporal panel dataset. Taking the wildfire as an example, at each time $t$ the maximum temperatures in location $i$'s nearby cells could also contribute to the wildfire occurrence in location $i$ through the flow of air, which is the spatial dependence plotted in Figure 5a. In spatial econometrics, the nearby influencing cells all called "spatial lags." In comparison, Figure 5b plots the usual "lags" to show temporal dependence, which means in location $i$ the maximum temperatures of the past few days or weeks could impact the wildfire occurrence at time $t$ as well, because of the accumulated heat over time. One difference between the

spatial and temporal dependence is that the spatial dependence does not have any specific direction, while the temporal dependence must be one-directional, meaning that only past values can affect the future, but not the other way round. However, in space any direction is allowed. There could exist the interaction of spatial and temporal lags too, as demonstrated in Figure 6. In short, this means the spatial lags of the temporal lags, or equivalently, the temporal lags of the spatial lags, might also affect our target location-time.

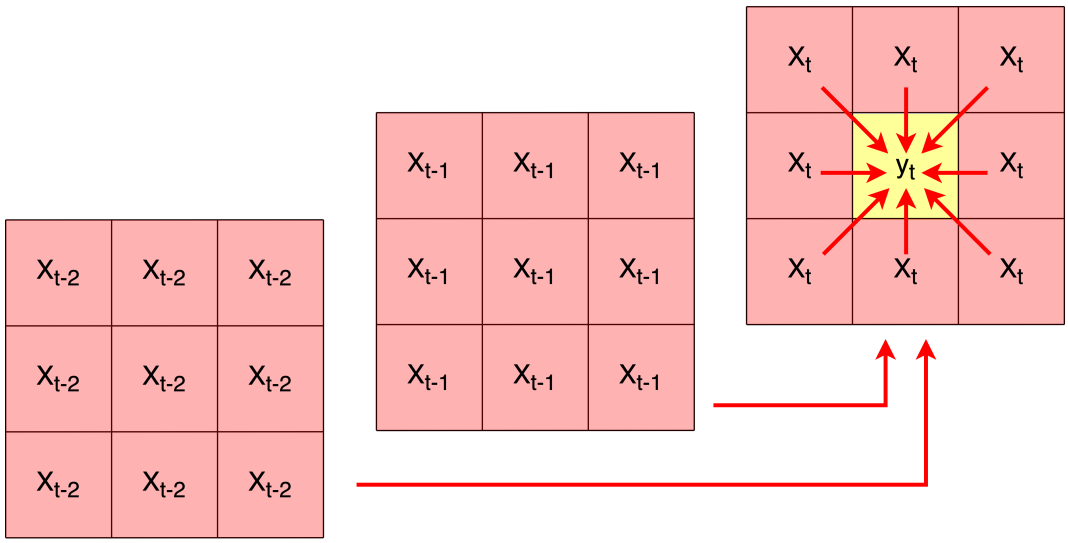



Figure 6: **Interactive spatial and temporal dependence.** Again, the yellow cell represents our target location-time, and the characteristics of red cells affect the target location's $y_{it}$ as well. For convenience, we plot only the cells which matter for the target location. Obviously, they are either spatial lags (as in time $t$) or the spatial lags of the temporal lags (as in time $t-2$ and $t-1$).

To sum up, the analysis above suggests that when dealing with spatiotemporal datasets, it is essential for the model to capture the dependence across both space and time. In the next two subsections, we will review several classical econometric models that are designed to deal with these potential dependence, and then explain how ST-CNN corresponds to better versions of these models.

## 3.5 Finite distributed lag model and 1-d convolution

In time series econometrics, the finite distributed lag (FDL) model is designed to include lagged explanatory variables to fully account for delays in the explanatory variables.

$$\boldsymbol{y}_t = \begin{cases} \sigma(\beta_0 \mathbf{1} + \boldsymbol{X}_t \boldsymbol{\beta}_t + \boldsymbol{X}_{t-1} \boldsymbol{\beta}_{t-1} + \cdots + \boldsymbol{X}_{t-q} \boldsymbol{\beta}_{t-q} + \boldsymbol{\varepsilon}_t) & \text{if classification,} \\ \beta_0 \mathbf{1} + \boldsymbol{X}_t \boldsymbol{\beta}_t + \boldsymbol{X}_{t-1} \boldsymbol{\beta}_{t-1} + \cdots + \boldsymbol{X}_{t-q} \boldsymbol{\beta}_{t-q} + \boldsymbol{\varepsilon}_t & \text{if regression,} \end{cases}$$

where $q < \infty$ is the order of lags, $\boldsymbol{X}_{t-1}, \cdots, \boldsymbol{X}_{t-q}$ are lagged explanatory variables that are added to capture the temporal dependence, $\boldsymbol{\beta}_{t-1}, \cdots, \boldsymbol{\beta}_{t-q}$ are the corresponding coefficient vectors for different lags, and $\boldsymbol{\varepsilon}_t$ is the independent error term vector.

The idea of including lagged explanatory variables, as in the FDL model, can be perfectly replicated by an 1-d convolutional neural network (CNN). Figure 7 graphically shows how a filter in the 1-d convolutional layer works. We fix an arbitrary location cell $i$, and describe its observed $K$ features as a $K$-dimensional time series. For each feature, or "channel" in CNN's terminology, applying the 1-d convolution can be considered as calculating weighted moving averages over a constant window. This moving window is called a "kernel," whose weights are the CNN parameters that will be learned by training. The kernel size is a hyper-parameter that can be tuned via validation. Different kernel sizes are equivalent to different order of lags $q$ in the FDL model. For convenience, we usually choose an odd number as the kernel size, so that for every moving average we will have a unique median cell as the target output cell, and the number of cells around the median cell will be the same on both sides.

The 1-d convolution is conducted channel by channel for all features. The convolutional outputs will then be linearly combined. The only difference between the 1-d convolutional filter and the FDL model is that a non-linear activation function ReLU (rectified linear unit) will be applied to the final output to enhance model complexity. Mathematically, a filter in the 1-d convolutional layer can be described as below. Assuming that the kernel size is $p$, then for each cell $i$,

$$y_{it} = \text{relu}\left(\beta_0 + \sum_{k=1}^{K} \beta_k \sum_{l=1}^{p} w_l^k x_{it-L+l}^k\right),$$
$$\text{where } \text{relu}(x) = \max(0, x).$$

In the equation above, $L = (p+1)/2$ is a constant given $p$, $x_{it}^k$ represents the $k$-th explanatory feature in location $i$ at time $t$, and $\{w_l^k\}_{l=1,\ldots,p}$ are the kernel weights for the $k$-th channel.

With the default convolutional setting shown above, one might be concerned about a *looking-ahead problem* in the convolution, meaning that future values of the explanatory variables are involved in the convolution. To make sure that the temporal dependence is one-directional, we shift the filter output backward, as Figure 8 demonstrates. If we shift the output backward by $L-1$ cells, we can guarantee that for each moving window, the last time cell in the convolution happens to be the target cell. However, note that Figure 8 is just

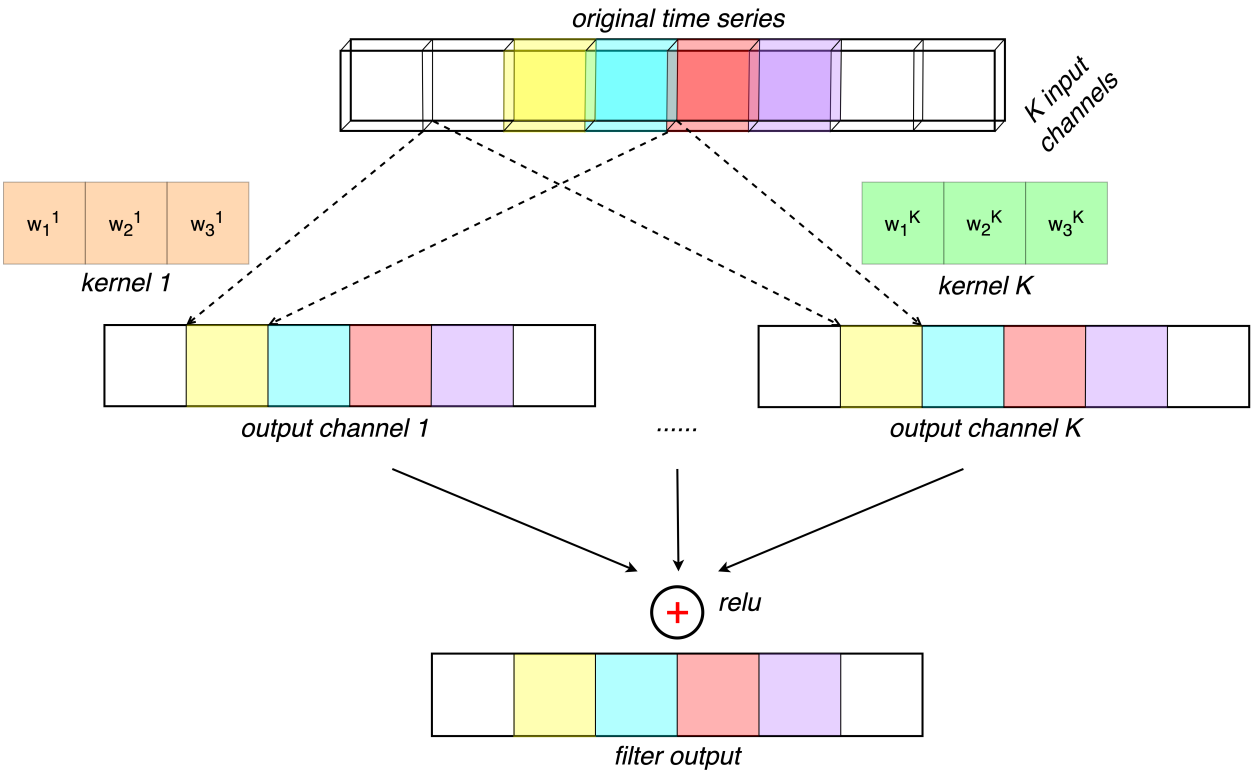



Figure 7: **An example of a filter in the 1-d convolutional layer.** In this figure, we use kernel size 3, which is arbitrarily chosen, as an example to show how a filter works in the 1-d convolutional layer. Note that each cubic or square cell represents a timestamp $t$, and we use the same color to keep track of the same timestamp. At the top we plot the original $k$-dimensional time series, where each dimension corresponds to a feature or an input "channel" in the CNN. Each channel is associated with a specific kernel, which can be understood as the convolutional weights applied to that channel. Specifically, the 1-d convolution is conducted by taking the inner product of the $3 \times 1$ kernel and the channel value of three consecutive cells, on a rolling basis across the entire channel, from left to right. For example, the value of the yellow cell in output channel 1 is the inner product of the kernel 1 vector and the closest 3 cells to the yellow cell, including itself, in the input channel 1 (marked by dashed lines). In other words, the central cell in the convolution is always the target cell. For simplicity, this figure plots only 1 filter, so the filter output has only 1 dimension. But in practice there could be a number of filters. The output of all filters will be stacked to form a multi-dimensional time series, which then becomes the input of the next layer in the neural network.

a conceptual visualization. In practice, one needs to shift the time series of the dependent variable to match the time horizon. With the filter output being shifted, we can then rewrite the 1-d convolution formula as

$$y_{it} = \text{relu}\left(\beta_0 + \sum_{k=1}^{K} \beta_k \sum_{l=1}^{p} w_l^k x_{it-p+l}^k\right).$$

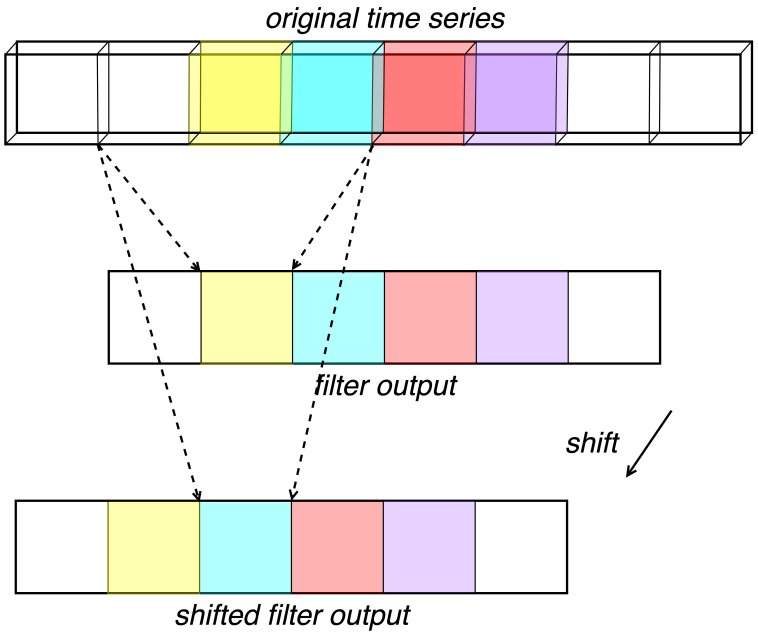



Figure 8: **Customize the 1-d convolution to avoid looking ahead.** In this figure, we conceptually demonstrate how to shift the convolution output so that we can avoid using future values of explanatory variables for estimation. Each cubic or square denotes a timestamp, and we use the same color to keep track of the same timestamp. Take the yellow cell as an example. Before shifting the output, the convolution will involve the blue cell, which is a future time. But after we shift the output backward, the target cell/timestamp of the same convolution changes to the blue one. The number of periods to shift backward will depend on the kernel size.

The 1-d convolution applies to the time dimension. Because it keeps different locations independent, we can simply stack the outputs across different locations.

$$\boldsymbol{y}_t = \text{relu}\left(\beta_0 \mathbf{1} + \boldsymbol{X}_t \boldsymbol{W}_p \boldsymbol{\beta} + \boldsymbol{X}_{t-1} \boldsymbol{W}_{p-1} \boldsymbol{\beta} + \cdots + \boldsymbol{X}_{t-p+1} \boldsymbol{W}_1 \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t\right),$$

where $\boldsymbol{W}_l = \text{diag}(w_l^1, w_l^2, \cdots, w_l^K)$ is a diagonal weight matrix, and $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_K)'$ is the coefficient vector. Note that the above formula is exactly in the same form as the FDL model in time series econometrics, except for two small differences. First, the order of lags

in the FDL model $q$ differs from the kernel size $p$ by 1, but this is simply because of different definitions. Second, in the 1-d CNN model, an additional non-linear activation function relu is added to the output.

## 3.6 Spatial econometric models and 2-d convolution

Several spatial econometric models are proposed to deal with spatial dependence concerns, such as the spatial cross-regressive model (SLX), spatial lag model (SLM), and spatial error model (SEM). In this subsection we will briefly review these classical models and we will explain how the 2-d CNN model could capture these models.

For all spatial econometrics, we will start with a $N \times N$ weight matrix $\boldsymbol{W} = \{w_{ij}\}_{i,j \in I}$, where $w_{ij}$ represents the weight that we impose on location $j$ when target location is $i$. Hence, the $i$-th row of $\boldsymbol{W}\boldsymbol{X}_t$ corresponds to the weighted characteristics of the nearby cells for location $i$, i.e., the spatial lags of location $i$. Note that here $\boldsymbol{W}$ needs to be determined ex-ante.

- **Spatial cross-regressive model (SLX)**
  In SLX, we introduce the spatially lagged exogenous regressors to the model, assuming that the spatial dependence can be captured by the spatial lags in the explanatory variables, namely $\boldsymbol{W}\boldsymbol{X}_t$.

$$\boldsymbol{y}_t = \begin{cases} \sigma(\beta_0 \boldsymbol{1} + \boldsymbol{X}_t\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}_t\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_t) & \text{if classification,} \\ \beta_0 \boldsymbol{1} + \boldsymbol{X}_t\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}_t\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_t & \text{if regression,} \end{cases}$$

  where the additional term $\boldsymbol{W}\boldsymbol{X}_t\boldsymbol{\gamma}$ will control for the effects of spatial lags, and $\boldsymbol{\gamma}$ are the corresponding coefficients. When $\boldsymbol{\gamma} = \boldsymbol{0}$, SLX will degenerate to an OLS model. Obviously, this method has drawbacks because the weight matrix $\boldsymbol{W}$ needs to be determined ex-ante. Although in most cases we can assume closer places should have higher weights, we are uncertain about the rate of spatial decay. On the other hand, even after controlling for the observed characteristics of spatial lags, we still have to conduct additional test to examine whether the spatial dependence is fully captured by $\boldsymbol{W}$. The advantage of SLX lies in that it is both conceptually and computationally easy, because the spatial lag variables can simply be treated as additional regressors. Hence, at least for the dependent variables with continuous values (i.e., for regression problems), this model can easily be estimated by OLS.

- **Spatial lag model (SLM)**
  Similar to SLX, we also use spatial lags to capture the spatial dependence in SLM. However, we switch from $\boldsymbol{W}\boldsymbol{X}_t$ to $\boldsymbol{W}\boldsymbol{y}_t$, assuming that the dependence manifests directly in the dependent variable. Geometrically, adding an autoregressive covariate corresponds to a ripple effect. In contrary to SLX where we usually assume only nearby cells matter, in SLM even if we start with a few nearby cells, the spatial dependence could end up spreading much more widely in space, through cascades of lag effects. The rate of spatial decay depends on the parameter $\rho$.

$$\boldsymbol{y}_t = \begin{cases} \sigma(\beta_0\mathbf{1} + \boldsymbol{X}_t\boldsymbol{\beta} + \rho\boldsymbol{W}\boldsymbol{y}_t + \boldsymbol{\varepsilon}_t) & \text{if classification,} \\ \beta_0\mathbf{1} + \boldsymbol{X}_t\boldsymbol{\beta} + \rho\boldsymbol{W}\boldsymbol{y}_t + \boldsymbol{\varepsilon}_t & \text{if regression.} \end{cases}$$

- **Spatial error model (SEM)**

$$\begin{aligned} \boldsymbol{y}_t &= \begin{cases} \sigma(\beta_0\mathbf{1} + \boldsymbol{X}_t\boldsymbol{\beta} + \boldsymbol{u}_t) & \text{if classification,} \\ \beta_0\mathbf{1} + \boldsymbol{X}_t\boldsymbol{\beta} + \boldsymbol{u}_t & \text{if regression,} \end{cases} \\ \boldsymbol{u}_t &= \lambda\boldsymbol{W}\boldsymbol{u}_t + \boldsymbol{\varepsilon}_t. \end{aligned}$$

  If we assume the spatial dependence is mainly due to some spatially correlated omitted variables, we can model it through the error terms. Specifically, we can apply a spatial version of autoregressive model on $u_t$, by assuming the error terms depend on their spatial lags $\boldsymbol{W}\boldsymbol{u}_t$. Hence, the parameter $\lambda$ controls the spatial decay rate, and $\boldsymbol{\varepsilon}_t$ is still the $i.i.d$ noise.

In both SLM and SEM, there exists a spatial autoregressive component, which will cause fairly high computational cost. For example, when $\boldsymbol{y}_t$ has continuous values, then to obtain the closed form solution for estimators we will have to invert a large matrix $(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}$ in SLM or $(\boldsymbol{I} - \lambda\boldsymbol{W})^{-1}$ in SEM, both of which have an $N \times N$ dimension. When $\boldsymbol{y}_t$ is a binary variable, this would be even more complicated because of the additional logistic function. In this sense, SLX seems computationally much cheaper.

There is a clear tie between SLX and the 2-d CNN model, in the sense that CNN is one instance of SLX with 1) a particular form of weight matrix and 2) additional nonlinearity. Figure 9 demonstrates a specific example of a filter in the 2-d convolutional layer. First, a filter consists of $K$ "kernels," where $K$ equals the number of covariates, or "channels," in the original data. For each kernel of size $k$, it represents a specific weight matrix, assuming that only the $k \times k$ nearest cells matter and farther away cells will have zero weights. Clearly, the size of spatial lags is determined by the kernel size $k$, which is a hyperparameter in this

model. Second, the weights are shared across space. For each target location, we compute the convolution by taking the inner product of the characteristics observed in its $k^2$ closest cells, which are equivalently the spatial lags in SLX, and the kernel. Since we use the same kernel for all locations, this will remarkably decrease the number of parameters from $N^2$ to $k^2$. Third, different kernels work independently, and the convolutional results are combined linearly, followed by a nonlinear activation function relu. Last, Figure 9 demonstrates only one filter. Actually one convolutional layers could have several independent filters, and all of the convolutional results will be passed to the next layer, which then work jointly for the estimation.

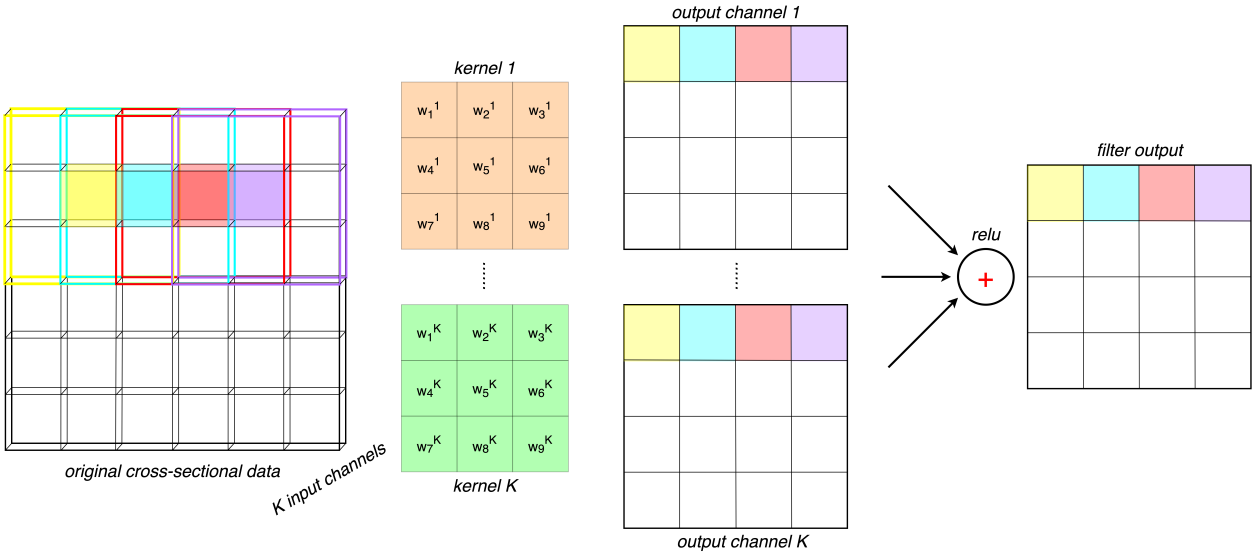



Figure 9: **An example of a filter in the 2-d convolutional layer.** This figure shows an example of a filter in the 2-d convolutional layer, consisting of $K$ kernels of size $3 \times 3$. The original cross-sectional data can be viewed as a map, where each cell represents a specific location. The data has $K$ features, called *input channels*. From left to right, we plot the original map, the kernels, the resulting maps (i.e., the *output channels*) and the final output of this filter. A kernel is associated with a specific weight matrix and its corresponding channel. The colors in the initial map and the channels match with each other. For example, the value in yellow cells are calculate by taking the inner product of the $3 \times 3$ yellow area in the initial map and the weight matrix. Note that the central cell of the $3 \times 3$ yellow area is our target location.

Mathematically, a 2-d convolutional filter in CNN can be written as

$$y_{it} = \text{relu}\left(\beta_0 + \sum_{k=1}^{K} \beta_k \sum_{j=1}^{N} w_{ij}^k x_{jt}^k\right).$$

Besides, $x^k$ represents the $k$-th covariate/channel, and $w_{ij}^k$ represents the kernel weight on

location $j$ when the target location is $i$. Note that $w_{ij}^k = 0$ when cell $j$ is out of the kernel coverage when we position the kernel to be centered at cell $i$. In other words, cell $j$ will receive a non-zero weight only when it is close enough to the target cell $i$ so that it could be covered within the kernel.

To see the connection between SLX and the CNN more clearly, we can re-write the equation above as a spatial cross-regressive model with some nonlinearities. That is,

$$\begin{aligned}\boldsymbol{y}_t &= \text{relu}\left(\beta_0\mathbf{1} + \boldsymbol{W^f} \circ \boldsymbol{X}_t\boldsymbol{\beta}\right) \\ &= \text{relu}\left(\beta_0\mathbf{1} + \boldsymbol{X}_t\boldsymbol{\beta} + (\boldsymbol{W^f} - \boldsymbol{I}) \circ \boldsymbol{X}_t\boldsymbol{\beta}\right),\end{aligned}$$

where $\boldsymbol{W^f} = [\boldsymbol{W}^1, \cdots, \boldsymbol{W}^K]$, which is a 3-d tensor of size $K \times N \times N$ that represents a collection of $K$ kernels for filter $f$, one for each covariate/channel. relu is applied element-wise to inputs of size $N$, and $(\boldsymbol{W^f} - \boldsymbol{I}) \circ \boldsymbol{X}_t$ is defined as

$$(\boldsymbol{W^f} - \boldsymbol{I}) \circ \boldsymbol{X}_t = \left[(\boldsymbol{W}^1 - \boldsymbol{I})\boldsymbol{X}_t\boldsymbol{e}_1, (\boldsymbol{W}^2 - \boldsymbol{I})\boldsymbol{X}_t\boldsymbol{e}_2, \ldots, (\boldsymbol{W}^K - \boldsymbol{I})\boldsymbol{X}_t\boldsymbol{e}_K\right],$$

where $\boldsymbol{e}_k$ is a one-hot vector that has value 1 in its $k$-th entry and 0 anywhere else. Therefore, $\boldsymbol{X}_t\boldsymbol{e}_k$ simply extracts the $k$-th column of $\boldsymbol{X}_t$, which is the collection of the $k$-th covariate for all $N$ locations. As demonstrated in Figure 9, we use the $k$-th kernel $\boldsymbol{W}^k$ to compute the spatial lags for the $k$-th explanatory variable. We deduct the identity matrix to separate out the original covariates and the spatial lags. For each weight matrix $\boldsymbol{W}^k$, although its size is $N \times N$ in this representation, its degree of freedom is in fact the kernel size $k \times k$, because the weights are shared across space.

Clearly, the 2-d convolutional filter is in essence an advanced version of SLX. First, compared with a fixed weight matrix in SLX, a convolutional filter allows for different spatial weights for different covariates. Second, weights are shared across space in the filter, which greatly reduces the degrees of freedom from $N \times N$ to $k \times k$, and hence saves a lot of computational costs. Third, in SLX the weights are determined ex ante, while in the convolutional filter the weights are learned from the data, jointly with other parameters. Additionally, a convolutional layer is much better than SLX in terms of the model expressivity. On one hand, a nonlinear component, namely the relu, is added to the filter. On the other hand, a convolutional layer could have several independently working filters, which could capture different characteristics of the original data and jointly work for the estimation.

## 3.7 Wildfire forecast model

In the previous two subsections, we explained how 1-d and 2-d CNN models can capture the temporal and spatial dependence respectively. Intuitively, for a spatiotemporal prediction task, we will need a 3-d model to perform convolution on the spatiotemporal lags. In this subsection, we will describe the setup details of our spatiotemporal wildfire forecast model.

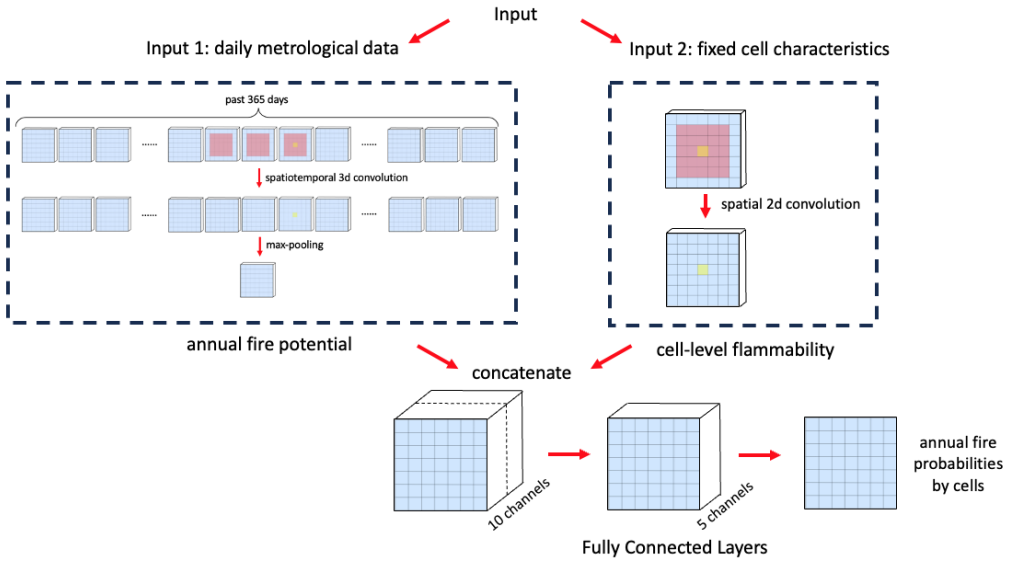



Figure 10: **ST-CNN: Branching structure for dynamic and static measures:** Input 1: includes the daily meteorological measures at each cell; Input 2 includes the fixed cell features such as topography and vegetative coverage

One main challenge in this forecast problem is how to make annual predictions based on daily meteorological data. Because most homeowner insurance contracts are annual, to price them on a daily basis, on each day we have to dynamically forecast the fire probability of the next 365 days. To solve this problem, we aggregate the fire occurrence over the next 365 days, and choose the daily meteorological data of the previous 365 days, i.e., the subsequence $x_{it-364:t}$, as the predictors for location $i$ at day $t$. We define our prediction target as below.

$$y_{it} = \begin{cases} 1 & \text{if a wildfire occurs in location } j \text{ between day } t+1 \text{ and } t+365, \\ 0 & \text{otherwise.} \end{cases}$$

In terms of the predictors, as Figure 10 shows, we divide them into a time-varying group and a time-invariant group. The time-varying group includes all daily meteorological features, such as the max temperature, specific humidity, dry lightning and wind events, based on which we could compute the fire potential dynamically for each location. On the other

hand, the time-invariant group consists of fixed effects that are associated with the general flammability of each location, including its topological features, vegetation conditions, electrical lines, utility providers, etc. We will discuss these predictors in detail in Section 4 below. Generally, this branching structure reflects the fact that the wildfire occurrence is the joint effects of dynamic weather conditions and static location flammability.

We apply CNN of different dimensions to the two predictor groups. For each day $t$, the daily meteorological features from the past 365 days are 3-d spatiotemporal panel data of dimension $365 \times N \times K_1$, where $N$ is the number of locations and $K_1$ is the number of time-varying predictors. In comparison, the static features are simple cross-sectional data of dimension $N \times K_2$, where $K_2$ is the number of time-invariant predictors. To account for both spatial and temporal dependence, a 3-d ST-CNN is designed for the dynamic daily meteorological data. Figure 11 plots an example of the 3-d convolutional structure. For each location-day, the figure describes how we use the weather data of the past few days from its surrounding cells, including itself, to forecast the fire potential. For one filter in the 3-d convolutional layer, we estimate

$$z_{it} = \text{relu}\left(\beta_0 + \sum_{k=1}^{K_1} \beta_k \sum_{l=1}^{p} \sum_{j=1}^{N} w_{ijl}^k x_{jt-p+l}^k\right),$$

where $z_{it}$ is the fire potential of cell $i$ at day $t$, $p$ is the size of temporal convolution, and $w_{ijl}^k$ is the target location $i$'s kernel weight on location $j$, at day $t-p+l$, for the $k$-th channel. Note that $w_{ijl} \neq 0$ only when location $j$ is within the spatial convolution range for location $i$. For time-invariant features, we use the common 2-d CNN to measure the general flammability of cell $i$,

$$m_i = \text{relu}\left(\alpha_0 + \sum_{k=K_1+1}^{K_1+K_2} \beta_k \sum_{j=1}^{N} w_{ij}^k x_j^k\right).$$

We use the max pooling structure to obtain annual fire potential measures. The idea of max pooling is simply taking the maximum value within a given time or spatial window. In our model, we apply a time window version of max-pooling to each location. As Figure 12 shows, after we perform the 3-d CNN and obtain daily measures of fire potential, we use the max pooling algorithm to get the maximum value from the previous 365 days, which we believe should be predictive for the wildfire occurrence of the next 365 days. One major advantage of this method is that we do not need to compute any statistics to aggregate those daily meteorological data into annual measures. Specifically, if we were to use a logistic regression, we would have to first calculate the mean, quantile or extreme values of the daily temperature, humidity, wind, etc., and then perform the annual forecast. This

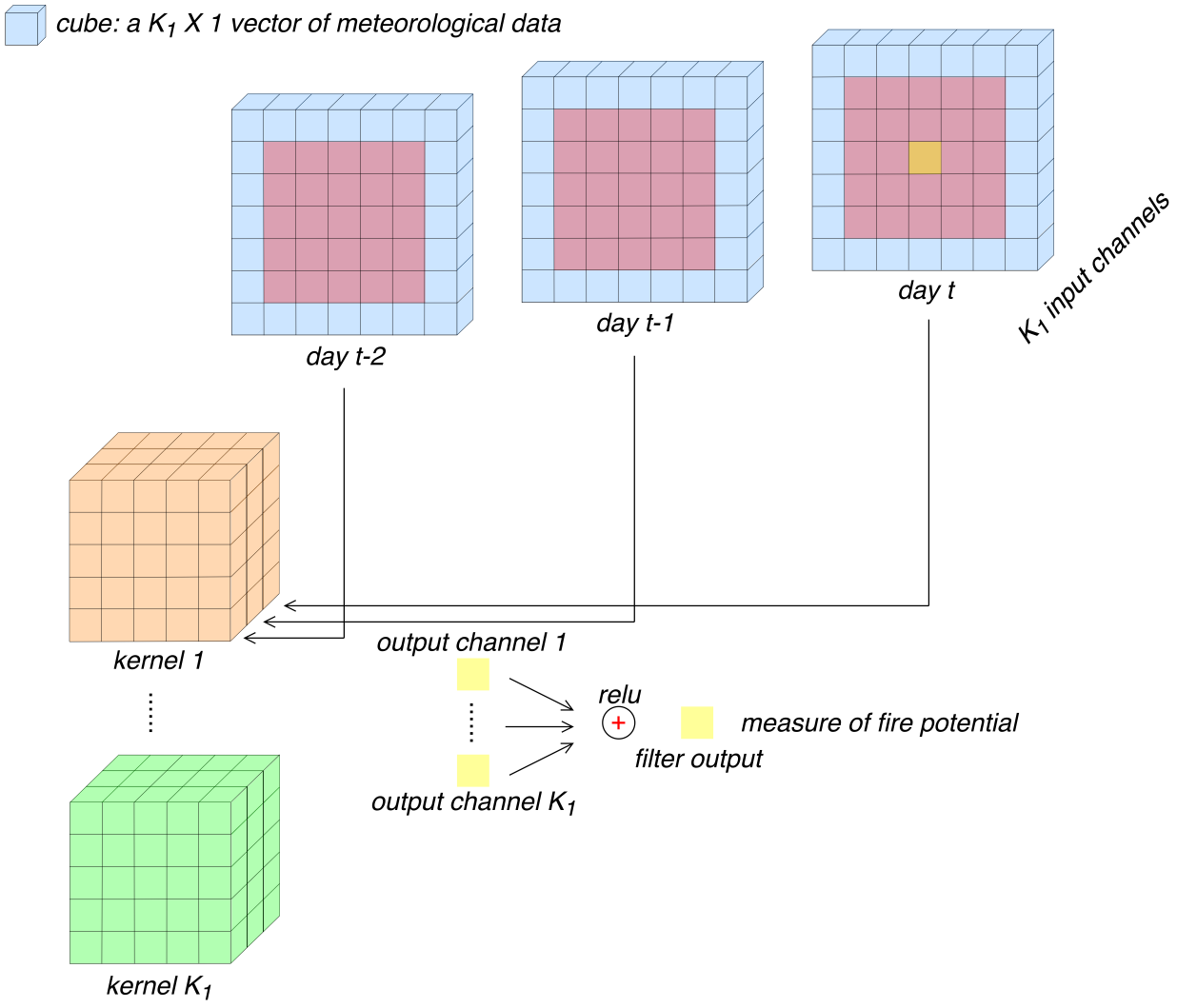



Figure 11: **Example of a filter in a 3-d convolutional layer with kernel size $(3, 5)$**: 3 is the temporal convolution size (i.e., the last three 3 days) and 5 is the spatial convolution size (i.e., the surrounding $5 \times 5$ cells). This kernel size is arbitrarily chosen as an example. Each cube represents a $K_1 \times 1$ vector of meteorological data, corresponding to $K_1$ input channels. Each channel is associated with a particular 3-d kernel of dimension $(3, 5, 5)$. The yellow cube is our target location-time. Its surrounding red cubes plus the yellow one itself are involved in the 3-d convolution, while the blue ones are out of the convolution range and irrelevant. We only show three days of data because other days are irrelevant. Just like the lower-dimensional cases, the convolution is performed by taking the inner products between the original data and the kernel, and the convolutional outputs of all channels are combined to produce the final output.

might lead to over-estimation because the warmest day may not necessarily be the day with strong wind events. In comparison, with this 3-d CNN structure we work directly on the daily data, compute daily fire potentials, and then take the maximum value, avoiding the time-mismatch problem.

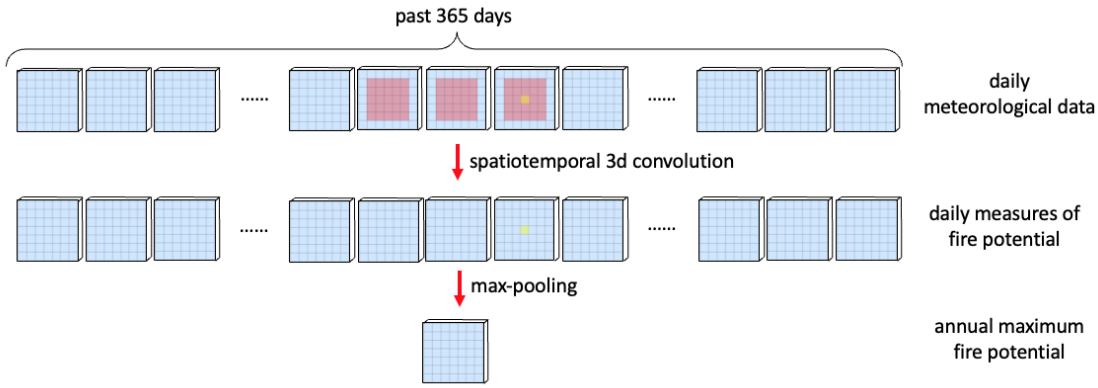



Figure 12: **Example of obtaining annual forecast by max-pooling.** In this graph, we show how we obtain the annual fire potential for each location $i$ on day $t$. Each big blue cube represents a $K_1$-dimensional cross-sectional map on a specific day, while each small cube stands for a location. At the top layer, we collect the daily meteorological data from day $t-364$ to day $t$, and then apply the 3-d CNN to to obtain daily measures of fire potential. Then we apply max pooling, which takes the maximum value from the past 365 days for each location. The 365 maps of fire potential are aggregated into one map after max pooling.

The max pooling helps to reduce the time dimension of time-varying features from 365 to 1. This means we could now easily concatenate measures of the annual fire potential and the time-invariant features, as we show in Figure 10. Then we add several fully connected layers, which are simply linear functions with relu, to allow different features to interact with each other. We use the logistic function in the last layer to produce the annual fire probability for each location. In sum, we could summarize our model setup as follows: for $s = t - 365 + p, t - 364 + p, \cdots, t - 1, t,$

$$
\begin{aligned}
z_{is} &= \text{relu}\left(\beta_0 + \sum_{k=1}^{K_1} \beta_k \sum_{l=1}^{p} \sum_{j=1}^{N} w_{ijl}^k x_{js-p+l}^k\right), \\
z_{it}^a &= \max\left(z_{it-365+p}, z_{it-364+p}, \cdots, z_{it}\right), \\
m_i &= \text{relu}\left(\alpha_0 + \sum_{k=K_1+1}^{K_1+K_2} \beta_k \sum_{j=1}^{N} w_{ij}^k x_j^k\right),
\end{aligned}
$$

where $p$ is the size of temporal convolution. For simplicity, here we assume we only use 1 filter in each layer. In practice, we have many filters in each layer, and the outputs are

concatenated to be vectors with different channels. Then,

$$\begin{aligned}
\boldsymbol{u}_{it} &= \text{relu}\left(\boldsymbol{\beta}_0^u + \boldsymbol{\beta}^u \begin{bmatrix} \boldsymbol{z}_{it}^a \\ \boldsymbol{m}_i \end{bmatrix}\right), \\
\boldsymbol{v}_{it} &= \text{relu}\left(\boldsymbol{\beta}_0^v + \boldsymbol{\beta}^v \boldsymbol{u}_{it}\right), \\
\hat{y}_{it} &= \sigma\left(\beta_0^y + (\boldsymbol{\beta}^y)' \boldsymbol{v}_{it}\right),
\end{aligned}$$

where $\boldsymbol{z}_{it}^a$ and $\boldsymbol{m}_i$ are concatenated features that reflect annual fire potential and location flammability, $\boldsymbol{u}_{it}$ and $\boldsymbol{v}_{it}$ are results from the intermediate fully connected layers, and $\hat{y}_{it}$ is a scalar that predicts the wildfire probability. Note that $\boldsymbol{\beta}^u$ and $\boldsymbol{\beta}^v$ are matrix that collect the coefficients for different channels, while $\boldsymbol{\beta}^y$ in the output layer is a vector.

## 3.8 Training and evaluation

Another challenge in the wildfire forecast problem is the severe class imbalance. Because wildfires are rare events, in any year the majority of grid cells in California do not burn. This means the wildfire datasets include a massive number of *zeros* and only a tiny proportion of *ones*. Furthermore, the minority-to-majority ratio keeps changing across time, because the wildfire occurrence is stochastic even conditional on meteorological and vegetative conditions. To properly handle the imbalanced data with time-varying class distribution, we must use a suitable loss function and evaluation metrics for the prediction task.

As summarized in Esposito et al. (2021) and Leevy et al. (2023), various approaches are developed at the data level and at the algorithm level to alleviate the potential harm of class imbalance to model training, and each of them has its own advantages and disadvantages. In general, data-level approaches are resampling methods, which aim at re-balancing the class distribution in the training data, such as up-sampling the minority class, down-sampling the majority class, or a hybrid of the two. These rebalancing techniques are quite straightforward and easy to implement, but are criticized for wildly bumping up the data size, for increasing the overfitting risk if up-sampling, and for losing important instances if down-sampling. On the contrary, algorithm-level approaches, including class weighting and output thresholding, will keep the original dataset unchanged, but modify the training algorithm to reduce bias towards the majority class. For example, a weighted loss function can be applied to punish more for incorrectly classifying an observation from the minority class. Alternatively, one could simply use a common loss function but adjust the decision threshold as an additional tuning step. Broadly speaking, these algorithm-level techniques are much more memory-efficient, and can be integrated into the training process. However, they do require some customization or extra tuning for the classifier.

The methods summarized above and their variants have been implemented in many previous risk modeling studies, including wildfire prediction. Zhang et al. (2019) apply over-sampling spatially by including 5km buffer zones around all existing fires as the fire samples too, and use a proportional stratified sampling on fire and non-fire samples for training. Dube and Verster (2024) and Akinjole et al. (2024) investigate a number of mainstream resampling methods on the task of credit default prediction. Both papers conclude that the SMOTE-ENN (Synthetic Minority Over-sampling Technique with the Edited Nearest Neighbors) method consistently achieves better classification performance. The SMOTE method, proposed by Chawla et al. (2002), is a widely-used oversampling technique that interpolates synthetic samples between existing minority-class observations and their neighbors, and the SMOTE-ENN further prunes them by removing unreliable samples to prevent overfitting (Batista et al., 2004). Different from the resampling-driven studies, Leevy et al. (2023) propose a constrained threshold optimization approach for credit-card-fraud detection, and find that the best overall results are achieved without up-sampling.

Among these choices, we conclude that the threshold optimization approach is the best suited for wildfire prediction, for several reasons. First, given that we will train our 3-d ST-CNN prediction model on a large panel dataset with 365 days of data for all grid cells in California across many years, a memory-efficient method is highly preferred. Second, we aim at a forward-looking prediction of wildfires, but it is impossible to conduct the resampling ex-ante without knowing the ratio between two classes. Another concern about the aforementioned resampling-based studies is that they mostly test their results on the up-sampled datasets, rather than the original imbalanced data. Third, besides the classification task itself, we will use the predicted probabilities from the model output for follow-up tasks such as expected-loss calculations. Last, Buda et al. (2018) have verified the effectiveness of threshold optimization methods on the CNN model. As a result, we train our spatiotemporal wildfire-prediction CNN model with the common Cross-Entropy Loss Function, and then apply threshold optimization methods from Leevy et al. (2023) to tune the decision threshold as a separate task. Accordingly, the model evaluations will also follow a two-step procedure. In the first step, we will evaluate the trained classifiers with different hyper-parameters using threshold-independent metrics, such as AUC (Area under the curve), based on which we will choose the best hyper-parameters. After we determine the optimal decision threshold, we will further evaluate our best-performing classifier using threshold-dependent metrics, including Precision, Recall, F1 score and so on.

To sum up, the stages of our training, evaluation and threshold optimization procedures are:

- First, we train the models based on the Cross-Entropy loss function.

$$\mathcal{L} = -\frac{1}{|I||T|}\sum_{i\in I}\sum_{t\in T}(1-y_{it})\log(1-\hat{y}_{it}) + y_{it}\log(\hat{y}_{it}),$$

  where $y_{it}$ is the true class label for cell $i$ at day $t$, and $\hat{y}_{it}$ is the predicted probability.
- Second, we use AUC, a threshold-independent metric, to evaluate the overall classification performance for models with different hyper-parameters on the validation set, and choose the best hyper-parameters.
- Third, we sample the last day of each year $t$, and follow Leevy et al. (2023)'s approach to calculate the optimal threshold $\theta_t^*$ for each year $t$ on a rolling basis, using the true class labels and predicted probabilities up to year $t-1$.

$$\begin{aligned} \theta_t^* &= \underset{\theta\in(0,1)}{\operatorname{argmin}}\ \text{G-mean}(\theta; y_{0:t-1}; \hat{y}_{0:t-1}) \\ \text{s.t.} \quad &\text{TPR}(\theta; y_{0:t-1}; \hat{y}_{0:t-1}) \geq \text{TNR}(\theta; y_{0:t-1}; \hat{y}_{0:t-1}), \end{aligned}$$

  where TPR stands for the true positive rate, TNR stands for the true negative rate, and G-mean is the geometric mean of TPR and TNR. All three metrics involved in this step are threshold-dependent.
- Last, with the determined decision thresholds, we further evaluate our model on the validation and test samples.

# 4 Data

The three key environmental conditions that determine vegetative wildfire behavior after ignition events are topography, fuel availability, and climatic conditions. CAL FIRE data include information on the alarm date, the precise location of the wildfire, when available the ignition cause, and the wildfire perimeter for each incident in California from 2001 through 2021. We follow Abatzoglou (2013) and carry out all of our wildfire prediction modeling using $2 \times 2$ kilometer grid cells projected over the entire state of California. We pre-process the dataset by correcting several typos in the fire dates and dropping any fire whose alarm date is missing or after its containment date. We then select all fires whose alarm dates are between 01/01/2001 and 12/31/2021, and merge them with our geocoded $2 \times 2$ kilometer cell network. Overall in our data, we have 6,999 wildfires in total and 42,012 fire day-cells.

| Measures for each grid cell | Specific features |
|---|---|
| **1. Topography** | |
| http://apps.nationalmap.gov/downloader/ | |
| | Elevation |
| | Aspect |
| | Slope |
| **2. Large utility districts** | |
| https://hub.arcgis.com/datasets/CalEMA::california-electric-utility-service-territory/about | |
| | Pacific Gas & Electric |
| | Southern California Edison |
| | Pacific Corporation |
| **3. Transmission lines** | |
| https://www.arcgis.com/home/item.html?id=d4090758322c4d32a4cd002ffaa0aa12 | |
| | Count of transmission lines |
| **4. Daily meteorology** | |
| https://www.climatologylab.org/gridmet.html | |
| | Maximum air temperature |
| | Specific humidity |
| **5. Hourly meteorology** | |
| https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=overview | |
| | Wind speed |
| | Wind direction |
| | Humidity |
| **6. Dry lightning** | |
| www.ncei.noaa.gov/pub/data/swdi/database-csv/v2/ | |
| | Lightning flash counts |
| **7. Vegetative types** | |
| https://gis.data.ca.gov/maps/CALFIRE-Forestry::california-vegetation-whrtype/about | |
| | Indicators for thirteen vegetative types |
| **8. Vegetative canopy** | |
| https://lpdaac.usgs.gov/products/mod44bv061/ | |
| | Percent covered by tree canopy |
| | Percent covered by non-tree canopy |

Table 1: **Table of grid-cell features and their data sources:** The table reports the eight classes of grid-cell measurement for topography, utility district service provider, transmission line counts, hourly meteorological data, daily meteorological data, vegetative types and tree/nontree canopy grid-cell percentage coverage. For each measurement class we also report the download information for the data sources as well as the specific features that we are using from each source.

## 4.1 Topography

Topography is a key factor in vegetative wildfire behavior. It influences the spatial variability of fuels and the biophysical conditions that determine wildfire ignition, the direction of spread, the intensity and the duration of wildfire. In California, topography and air pressure systems play a key role in the direction and speed of the hot dry northeasterly Diablo winds of Northern California and the Santa Ana winds of Southern California. The steepness and southwest aspect of the slopes of these ranges and the high elevation of their ridge tops all induce relatively drier vegetative fuel conditions especially in the high ambient temperature months of July through October. Topography also has an impact on a variety of other features of fire behavior such as fire-line width, flame length, and the direction of spread. Another important aspect of fire behavior that is affected by topography is the rate of spread since many fires accelerate dramatically uphill, thus placing fire fighters, reservoir access, and utility infrastructure at risk (Linn et al., 2007).

As shown in Table 1, our slope, elevation and aspect measure are computed using topographical raster data from the U.S. Geological Services and geoprocessing this information using QGIS software to derive the slope and aspect at each grid. Aspect is the compass direction that each grid cell faces and it is measured in degrees between 0 and 360 in the raw data. We use a cosine transformation of aspect, ranging from -1 to 1, to provide a measure for how closely a grid is facing a southwestern orientation (i.e., the 225-degree direction) (Kumar et al., 1997). The southwestern orientation is important because locations with this orientations have higher solar-radiation exposure. Under the cosine transformation the southwestern aspect is defined as,

$$\text{aspect\_sw} = \cos\left(\frac{\text{aspect} - 225}{180}\pi\right)$$

For simplicity, we use the term aspect to refer to the cosine transformed southwestern aspect from now on.

## 4.2 Utility districts and transmission lines

California utilities have struggled with wildfire related liability associated with inadequate vegetation management around their transmission lines, deferred maintenance of their transmission power pylons, and catastrophic wildfire ignition events associated with damage to lines and pylons within their service areas. California Assembly Bill 1054, passed in July, 2019, funded a $5 billion fund for utility wildfire safety investments that required utilities to file Wildfire Mitigation Plans with the California Public Utilities Commission in exchange

for access to the mitigation plan funds for investment reimbursement. There are currently three utilities that participate in the California Wildfire Fund – San Diego Gas & Electric Company, Southern California Edison, and Pacific Gas & Electric Company. However, there remains considerable controversy concerning the most cost effective utility mitigation strategies: undergrounding of lines or covered conductors.[19]

Again, following Table 1, we focus on the three largest utilities: Pacific Gas & Electric, Southern California Edison, and Pacific Corporation. Two of these utilities currently participate in the California Wildfire Fund under AB1054 and the other, Pacific Corporation, does not. We download the map of California Electric Utility Service Territory[20] and assign the utility provider to each of our 2 km $\times$ 2 km cells. In addition, we obtain the map of U.S. Electric Power Transmission Lines.[21] For each cell, we use ArcGIS to count the number of electrical lines passing through its area.

- **Pacific Gas & Electric (PG&E)** is the largest investor-owned California utility and serves approximately 16 million people throughout a 70,000-square-mile service area in northern and central California.[22] Five of the ten most destructive fires in California since 2015 have been linked to PG&E's electrical network, with regulators finding that in many fires, PG&E violated state law or could have done more to make its equipment safer.[23]
- **Southern California Edison (SCE)** is one of the nation's largest electric utilities and provide electric service to approximately 15 million people through 5 million customer accounts. SCE's service area includes portions of 15 counties and hundreds of cities and communities in a 50,000-square-mile service area within Central, Coastal and Southern California.[24] Public Utilities Commission investigators found SCE liable for damages from SCE power lines associated that ignited the 2017 Thomas fire burning more than 280,000 acres, damaging more than 1,300 structures and causing two deaths in Santa Barbara and Ventura counties. Another Public Utilities Commission again found liability related to Southern California Edison equipment that was likely "associated" with 2018's deadly Woolsey fire, which burned more than 1,600 structures

[19] See "This utility's undergrounding plan is causing sticker shock" by Wes Venteicher and Blanca Begert, *Politico*, October 5, 2023. https://www.politico.com/newsletters/california-climate/2023/10/05/pg-es-undergrounding-plan-is-causing-sticker-shock-00120290.

[20] https://hub.arcgis.com/datasets/CalEMA::california-electric-utility-service-territory/explore.

[21] https://www.arcgis.com/home/item.html?id=d4090758322c4d32a4cd002ffaa0aa12.

[22] https://www.pge.com/en/about/company-information/company-profile.html.

[23] https://www.nytimes.com/interactive/2019/03/18/business/pge-california-wildfires.html.

[24] https://download.newsroom.edison.com/create_memory_file/?f_id=5cc32d492cfac24d21aecf4c&content_verified=True.

in Los Angeles and Ventura counties and killed three people.[25]

- **Pacific Corporation** is the largest transmission-line grid operator in the West, with 2.1 million customers in a service area of 141,500 square miles, including parts of Oregon, Washington, California, Utah, Idaho and Wyoming.[26] S&P Global reported in August 2024 that Pacific Corporation faces at least $46 billion in claims related to Western U.S. wildfires following recent lawsuits in Oregon for fires in Oregon and California.[27]

Given the risks associated with density of California transmission lines and transmission power pylons, we also measure the underlying vegetative ground cover due to their differing susceptibility to combustion events when temperatures, wind, and relative humidity sufficiently lower moisture levels. To measure these risks, as shown in Table 1, we count the number of transmission lines that cross each grid cell and then take the log of that count as a measure to risks of failures of the lines and pylons themselves.

## 4.3 Meteorology

### 4.3.1 Daily measures

Following Table 1, our daily measure of climate data are obtained from gridMET, which is a publicly available dataset of high-spatial resolution (4 km $\times$ 4 km) surface meteorological data covering the contiguous US from 1979 till yesterday (see Chegini et al., 2021). We then use interpolation to map from the 4 km $\times$ 4 km grid cells to a denser 2 km $\times$ 2 km spatial coverage of the state. The gridMET data is updated daily. We focus on the gridMET primary climate variables including maximum air temperature and specific humidity.[28]

### 4.3.2 Hourly measures

In California, there are two types of fire-associated wind: the Diablo Winds of northern California and the Santa Ana winds of southern California. Santa Ana winds have been the driving force behind many of southern California's most devastating fires (see Billmire et al.,

[25] https://www.latimes.com/california/story/2021-12-17/southern-california-edison-faces-550m-penalty-for-wildfires.

[26] https://www.pacificorp.com/about.html.

[27] https://www.spglobal.com/commodityinsights/en/market-insights/latest-news/electric-power/080524-wildfire-claims-against-pacificorp-surge-to-46b-on-oregon-mass-complaints.

[28] We find that the gridMET variables that are identified as derived variables are highly correlated with the gridMET primary climate variables. The gridMET derived variables include: Reference evapotranspiration (ASCE Penman-Montieth), Energy Release Component, Burning Index, 100-hour and 1000-hour dead fuel moisture, mean vapor pressure deficit, 10-day Palmer Drought Severity Index are all highly correlated with the gridMET primary climate variables (https://www.climatologylab.org/gridmet.html).

2014; Jin et al., 2013; Kochanski et al., 2013). More recently, the Diablo winds of northern California have become more associated with wildfire occurrence with their similarly low relative humidity, high temperatures, and very high wind speeds (see Bowers, 2018; Diaz, 2022; Keeley and Syphard, 2018; Liu, 2022; Liu et al., 2021; McClung and Mass, 2020).

Although both of these winds are shaped by atmospheric conditions and driven by the topography of the region, they also have important differences. The Diablo winds (see Abatzoglou et al., 2018, 2021; Diaz, 2022; MacDonald et al., 2023) are strong northeasterly winds that flow over the western slopes of the Sierra Nevada range in eastern California where they heat up and lose humidity before passing through the California Central Valley. High pressure systems in the Central Valley then drive the Diablo winds over the California Coastal Range where the compression and the loss of moisture produces intense, dry, downslope winds.

In contrast, the Santa Ana winds of southern California are gravity driven winds that occur when high pressure builds over the Great Basin to the west of the Sierra Nevada mountains,[29] and low pressure systems develop over the California coast (see Cardil et al., 2021; Gershunov et al., 2021; Guzman-Morales, 2018; Keeley et al., 2021). Cold air from the Great Basin then sinks and dry air from the desert is pushed toward the low-lying coastal areas through the Sierra and coastal mountain canyons, where compression causes the winds to warm by tens of degrees Fahrenheit per mile as it travels.

Hourly climate measurement is required to measure the occurrence of Diablo and Santa Ana winds. Diablo winds are characterized by their intensity, long duration, low moisture content, and northeasterly direction. Following Bowers (2018), Diaz (2022), and Liu et al. (2021), we focus on the area of North Coast of California and the Sierra Nevada Mountains, and identify the occurrence of Diablo winds if the winds are northeasterly in direction, have speeds exceeding 5 m/s, have relative humidity that is below 25%, and have a duration of at least six hours. Following Guzman-Morales (2018), Keeley et al. (2024) and Bowers (2018), the Santa Ana winds are defined as northeasterly winds in the South Coast of California, with relative humidity of below 25%, that last at least 12 hours and during which the wind speed ever reaches 5 m/s[30] and the relative humidity ever drops to below 10%. Following Table 1, our hourly climate measurement data are obtained from ERA5-Land, published by ECMWF (The European Centre for Medium-Range Weather Forecasts).[31] We identify wind events using hourly two meter dewpoint temperature, two meter temperature and ten meter direc-

[29]The Great Basin includes most of Nevada, half of Utah, and sections of Idaho, Wyoming, Oregon, and California.

[30]As discussed in Diaz (2022) and Liu et al. (2021), the wind speed thresholds should vary across datasets with different spatial resolution. Following Diaz (2022), we choose 5 m/s for the ERA5-Land data.

[31]See https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=overview.

tional components of wind (u-component and v-components). The geographical locations used in the wind events identification are based on the map of Level III Ecoregions of North America, downloaded from the EPA (U.S. Environmental Protection Agency) website.[32]

## 4.4 Dry lightning

Dry lightning, occurring when there has been less than 2.5 mm of rainfall, is a major source of wildfire ignition in central and northern California (see Kalashnikov et al., 2022). While human-caused wildfire ignitions predominate in southern California, lightning-caused fires are more common in the northern half of the state, particularly over mountainous terrain (see Balch et al., 2017; Brey et al., 2018; Chen and Jin, 2022; Keeley and Syphard, 2018). Summertime lightning outbreaks in northern California from July through August, unlike the predominantly human-caused fires that originate in a single location, can strike multiple locations and start numerous simultaneous wildfires (Miller et al., 2012). Widespread thunderstorms with dry lightning have produced some of the largest and longest-lasting wildfires in recent decades, including the 1987 wildfire season (see Duclos et al., 1990), the 2008 wildfire season (Wallmann et al., 2010), and the disastrous 2020 wildfire season (Keeley and Syphard, 2021). Based on our CAL FIRE ignition-cause data we find that more than 30% of wildfires in California between 2001 and 2021 with an identifiable cause arised from dry lightning.

## 4.5 Vegetative types

Vegetation has significant effects on wildfire behavior and is thus an important focus of wildfire prediction modeling (see Kearns et al., 2022; Price and Bradstock, 2014). Topographic features such as elevation, aspect, latitude, and slope also influence microclimate conditions, such as temperature, precipitation, direct solar radiation, wind exposure, among others, which together influence the moisture content of fuel (Flannigan et al., 2016, 2009; Westerling, 2014). As noted above, topography can also affect ignition probabilities because steep slopes, ridge tops, and southwest facing slopes are all characterized by drier fuel conditions. Other important vegetative features that affect wildfire ignition probabilities and behavior include the spread of invasive non-native or non-conifer species (Brooks and Matchett, 2006; Calhoun et al., 2022; Holmes et al., 2008) and the degree of nearby urbanization (see Alexandre et al., 2016; Kestelman, 2024; Price and Bradstock, 2014).

As shown in Table 1, we focus on thirteen types of vegetative exposure for each grid cell: agriculture, barren/other, conifer forest, conifer woodland, desert, desert shrub, hardwood

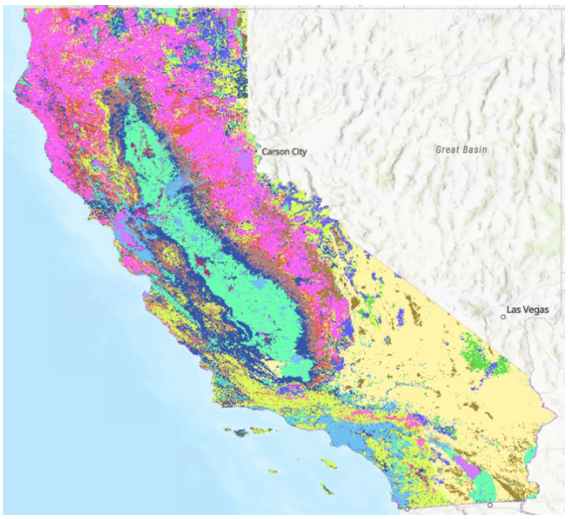
[32] https://www.epa.gov/eco-research/ecoregions-north-america.

forest, hardwood woodland, herbaceous, shrub, urban, water, and wetland following the CAL FIRE's WHR (Wildlife Habitat Relationship) Type indexes discussed in Table 1. Figure 13a presents the geographic locations of the vegetative ground coverage types. As shown, conifer woodlands/forests dominate the Sierra Nevada mountain range and the northern section of the coastal range. The hardwood forest and hardwood woodlands are found in the foothills of the Sierra Nevada and along the coastal range north of San Francisco. Herbaceous, shrub, and woodlands are the dominant vegetative ground covers in the coastal range areas around San Francisco and extend south to Santa Barbara County. The Los Angeles Basin including San Diego is dominated by herbaceous, desert shrub, and desert woodland with small areas of conifer forest at higher elevations. The central valley of California is dominated by irrigated agricultural vegetation. As shown in Subfigure 13c of Figure 13, the historical locations of wildfires in California are found on the western facing slopes of the Sierra Nevada and along the coastal range in both northern and southern California. The northern California wildfires are dominated by hardwood/conifer woodland and forest vegetation. The southern California wildfires have primarily occurred in areas dominated by herbaceous, desert/conifer woodland, and shrub/chaparral vegetation. The irrigated areas of the central valley, as well as the Mojave Desert, have the lowest historical wildfire incidence.

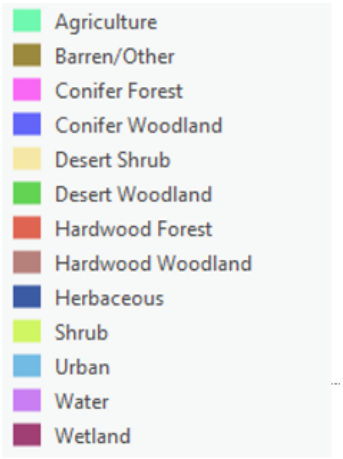
## 4.6 Vegetative canopy density

Finally, as shown in Table 1, in addition to our measures for the vegetative types found within grid cells, we also account for the canopy coverage of the vegetation within the cells. Dense canopy coverage is importantly associated with the effects of incoming solar energy and the live fuel content of the vegetation. The live fuel moisture content is a measure of the water content of live fresh foliage relative to its dry mass (Yebra et al., 2018) and it is an important determinant of the potential for fire ignitions to propagation. California has significant heterogeneity in its canopy coverage and the types of vegetation associated with the coverage. Northern California has dense canopies of conifer and hardwood in the foothills of the Sierra Nevada and the northern Coastal Range that are significantly prone to wildfire (see Chen et al., 2021). The shrubland canopies of the southern California chaparral areas from Monterey County south to the Los Angeles Basin and San Diego pose equally severe threat of wildfire occurrence and rapid propagation (see Dennison and Moritz, 2009; Dennison et al., 2008).

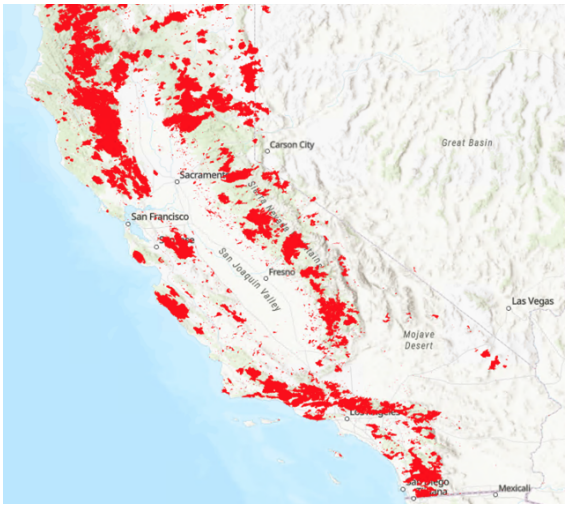
We obtain two measures of canopy coverage from the Moderate Resolution Imaging Spectroradiometer's (MODIS) Vegetation Continuous Fields database from the LP DAAC (Land Process Distributed Active Archive Center). The first measure is the "percent covered




(a) Geographic location of vegetative types in California, 2021




(b) Vegetative types 2021




(c) Historical location of wildfire incidents in California (2000–2021)

Figure 13: **Relationship between vegetative distributions and wildfire incidents.** The vegetative coverage data are sourced from CAL FIRE Forestry (https://gis.data.ca.gov/maps/CALFIRE-Forestry::california-vegetation-whrtype/about). The wildfire incidence and burn area data are sourced from California Department of Forestry and Fire Protection (CAL FIRE) (https://www.fire.ca.gov/what-we-do/fire-resource-assessment-program/fire-perimeters).

by tree canopy," defined as the canopy coverage by woody plants that are greater than or equal to five meters tall (see Chen et al., 2021; DiMiceli et al., 2021). The second measure is the "percent covered by non-tree canopy," defined as the canopy coverage associated with small trees (less than 2.5 meters), grass, or shrubs (see Lai et al., 2022; Mallinis et al., 2019).

# 5 Wildfire forecasting: logistic regression vs. ST-CNN

As discussed in Section 3, logistic generalized additive models have been a common approach to the modeling of wildfire occurrence in the climate literature (Xi et al., 2019) and logistic regression is the most common method used in recent analyses of wildfire incidence in the economics literature (see Biswas et al., 2023; Kahn et al., 2024). This section compares the forecasting performance of logistic regressions and ST-CNN.

## 5.1 Logistic regression

Because wildfires tend to occur on very hot and/or very dry days, we use annual aggregates at the 0.05 quantile for specific humidity and at the 0.95 quantile for maximum air temperature from May to October. For each grid cell, we create a daily indicator for Diablo or Santa Ana wind exposure using hourly measurements of wind speed, direction, and relative humidity. Similarly, a daily indicator for dry lightning is constructed based on flashes and precipitation. We aggregate these daily indicators to annual indicators by assigning a value of one if such an event occurred on at least one day in the past year.

As shown in Table 2, most coefficients are statistically significant and align with expectations. Key meteorological variables, such as specific humidity and maximum air temperature, demonstrate significant relationships with wildfire occurrence. Wildfire occurrence is also shown to be positively associated with dry lightning and Diablo/Santa Ana winds. Additionally, consistent with recent wildfire events, higher counts of transmission lines within grid cells and the interaction of higher counts with indicator variables for the location of those lines in the PG&E and Southern California Edison service provision districts are also positively associated with higher levels of wildfire occurrence, whereas higher transmission line counts in the Pacific Corporation service area are negatively associated with wildfire incidence.

Consistent with the climate literature, the logistic regression results also show that both tree and small-tree/shrubland canopy coverage is positively associated with the occurrence of wildfire. Similarly the fixed-effects controls for other twelve vegetative types, other than desert shrub, are shown to be statistically significantly and positively associated with wildfire

| | coef | std err | z | P> \|z\| |
|---|---|---|---|---|
| Constant | −5.721 | 0.011 | −537.299 | 0.000 |
| Specific humidity (q5: May–Oct) | −0.039 | 0.002 | −16.839 | 0.000 |
| Maximum air temperature (q95: May–Oct) | 0.235 | 0.003 | 81.509 | 0.000 |
| Indicator: Diablo or Santa Ana wind events | 0.284 | 0.010 | 27.489 | 0.000 |
| Indicator: dry lightning | 0.072 | 0.006 | 12.594 | 0.000 |
| Transmision line count | 0.066 | 0.006 | 11.868 | 0.000 |
| Transmision line count × PG & E Indicator | 0.035 | 0.006 | 5.701 | 0.000 |
| Transmision line count × Southern California Edison Indicator | 0.145 | 0.007 | 21.879 | 0.000 |
| Transmision line count × Pacific Corporation Indicator | −0.348 | 0.012 | −28.597 | 0.000 |
| Grid-cell percentage tree canopy | 0.179 | 0.004 | 51.090 | 0.000 |
| Grid-cell percentage non-tree canopy | 0.387 | 0.004 | 99.186 | 0.000 |
| Slope | 0.456 | 0.004 | 113.185 | 0.000 |
| Aspect | 0.005 | 0.002 | 2.904 | 0.004 |
| Vegetative Type Fixed Effects | Yes | | | |
| Agriculture | . | . | . | . |
| Barren/Other | 0.889 | 0.021 | 41.958 | 0.000 |
| Conifer forest | 1.615 | 0.013 | 125.046 | 0.000 |
| Conifer woodland | 1.154 | 0.016 | 71.079 | 0.000 |
| Desert shrub | 0.017 | 0.017 | 0.995 | 0.320 |
| Desert woodland | 0.656 | 0.034 | 19.257 | 0.000 |
| Hardwood forest | 1.698 | 0.013 | 129.062 | 0.000 |
| Hardwood woodland | 1.333 | 0.012 | 107.575 | 0.000 |
| Herbaceous | 1.486 | 0.011 | 131.585 | 0.000 |
| Shrub | 2.029 | 0.011 | 181.326 | 0.000 |
| Urban | 1.209 | 0.014 | 87.196 | 0.000 |
| Water | 1.507 | 0.020 | 73.496 | 0.000 |
| Wetland | 0.781 | 0.028 | 27.644 | 0.000 |
| No. Observations: | 20,864,333 | | | |
| Pseudo R-squ.: | 0.072 | | | |
| Log-Likelihood: | $-1.619 \times 10^6$ | | | |
| LL-Null: | $-1.744 \times 10^6$ | | | |
| LLR p-value: | 0.000 | | | |

Table 2: **Logistic Regression Results.** This table presents logistic regression estimates of annual wildfire incidence using data from 2001 to 2019. Observations are grid cells aggregated from daily to monthly data, with several meteorological, topographical, and utility exposure predictors

occurrence, where the hold-out vegetative type is irrigated agriculture. The two topographical measures for aspect and slope are also positively associated with the occurrence of wildfire, consistent with the discussion found in Section 4.1.

Overall, logistic regression identifies feature associations with wildfire occurrence that are consistent with the climate, topographical, and vegetative coverage literature surveyed in Section 4 and the results of the cross-sectional econometric techniques that mostly characterize this literature, as discussed in Section 3. However, it lacks the ability to handle spatial/temporal correlations, nonlinear relationships, or more general interaction effects. Moreover, logistic regression models are inherently cross-sectional, so are not designed to handle the temporal aspects of forecasting.

## 5.2 ST-CNN model performance

Following the model presentation in Section 3, we apply spatiotemporal CNNs to analyze wildfire occurrence using the same predictors as in Table 2.

Given the class imbalance in California wildfire data, where wildfire instances are rare, we evaluate model performance using the area under the ROC curve (AUC), because this metric is is not sensitive to threshold choices. We divide the dataset into training, validation, and test sets, in time order. The training set covers the period between Dec. 31, 2000 and Dec. 31, 2018. Note that on each day we predict one-year ahead, so the training set includes all wildfires between Jan 2001 and Dec 2019. Given the potential overlapping in the prediction horizon across the days within a year, for computational simplicity we sample the last day of every month in the training set to train the model, which includes altogether 217 days. On each day we cover 96,149 grid cells and for each cell we use a subsequence data of the past 365 days as the predictors. In other words, we train the ST-CNN on a rolling basis of the next-year fire events using the past-year data in order to obtain a dynamic wildfire prediction model. We then use the 2019 subsequence of data observed on Dec 31, 2019 and the 2020 wildfires as the validation, and the 2020 subsequence of data observed on Dec 31, 2020 and the 2021 wildfires as the test. We carefully divide the dataset to avoid any look-ahead problem — the last day in the training set is one year before the validation day. We also leave a one-year gap between validation and test days.

Table 3 displays AUC values for various ST-CNN configurations, demonstrating that larger convolutional kernel sizes enhance ST-CNN's ability to capture spatiotemporal dynamics. The rows and columns of Table 3 represent differing kernel sizes for the 3-d convolutional layer, where, for example, row $p = 7$ corresponds to the temporal convolution size of the last 7 days and the column $k = 9$ corresponds to the spatial convolution size of the

| Training | $k=1$ | $k=3$ | $k=5$ | $k=7$ | $k=9$ |
|---|---|---|---|---|---|
| $p=1$ | 0.7466 | 0.7879 | 0.7925 | 0.7693 | 0.7672 |
| $p=3$ | 0.7608 | 0.7924 | 0.7864 | 0.7807 | 0.7970 |
| $p=7$ | 0.7690 | 0.7718 | 0.7969 | 0.7993 | 0.8310 |
| $p=30$ | 0.7986 | 0.7704 | 0.8125 | 0.7781 | 0.7539 |

| Validation | $k=1$ | $k=3$ | $k=5$ | $k=7$ | $k=9$ |
|---|---|---|---|---|---|
| $p=1$ | 0.7301 | 0.7613 | 0.7897 | 0.7776 | 0.7734 |
| $p=3$ | 0.7315 | 0.7730 | 0.7746 | 0.7811 | 0.8029 |
| $p=7$ | 0.7350 | 0.7519 | 0.7604 | 0.7708 | 0.8130 |
| $p=30$ | 0.7448 | 0.7562 | 0.7473 | 0.7781 | 0.7904 |

| Test | $k=1$ | $k=3$ | $k=5$ | $k=7$ | $k=9$ |
|---|---|---|---|---|---|
| $p=1$ | 0.7033 | 0.7189 | 0.7335 | 0.7513 | 0.7526 |
| $p=3$ | 0.7216 | 0.7278 | 0.7037 | 0.7101 | 0.7274 |
| $p=7$ | 0.6787 | 0.7288 | 0.7513 | 0.6996 | 0.7717 |
| $p=30$ | 0.6943 | 0.7295 | 0.7003 | 0.7238 | 0.7226 |

Table 3: **AUC performance.** The area under the ROC curve (AUC) on training, validation and test samples, for spatiotemporal CNN models with different kernel sizes.

surrounding $9 \times 9$ cells. Thus, for the $p = 7$ row and the $k = 9$ column in Table 3, each of the AUC values reported for the training, validation, and out-of-sample ST-CNN is associated with a channel structure comprised of 3-d kernels of dimension $(7, 9)$. Table 3 presents the AUC results for a range of hyperparameter structures from $p = 1$ to $p = 30$ for the temporal convolution sizes and $k = 1$ to $k = 9$ for the spatial convolution sizes.

As expected, given the inherent randomness of wildfire events, as well as the challenges of imbalanced classification, the AUC values for the training estimates are uniformly low, ranging from 0.7466 for the 3-d kernels of size $(1, 1)$ to a high of 0.8310 for the 3-d kernels of size $(7, 9)$, which are all far from 1. The highest AUC reported in Table 3 is consistently obtained by the 3-d kernels of size $(7, 9)$ across the training, validation and test sets. Hence, we will choose the ST-CNN model with kernel sizes $(7, 9)$ as our prediction model. As shown, the AUC is 0.8130 for the validation data and the AUC is 0.7717 for the test data, indicating that the ST-CNN performs only modestly well in forecasting the annual occurrence of wildfire while minimizing false positive and false negative forecasts.

### 5.2.1 Feature importance

The ST-CNN is a complex non-linear model with spatial and temporal dependence. To figure out how different features contribute to model predictions, we estimate the SHAP (SHapley Additive exPlanations) value, which decomposes the prediction output for any instance as a sum of marginal contributions from its individual features.

Assuming that the complex model is linear, at least locally, we follow Lundberg and Lee (2017) in approximating the SHAP values with the Local Interpretable Model-agnostic Explanations (LIME) method. Specifically, for an arbitrary prediction of the grid cell $i$ on the test set, we randomly sample 1,000 binary inputs $z = (z_1, \cdots, z_K) \in \{0,1\}^K$, where $K$ is the number of features. These simplified binary inputs are then mapped back to the original feature space to construct a dataset of 1,000 perturbed observations $\tilde{x}_i$. Essentially, with the binary values of $z$, in each perturbation we randomly block certain features from functioning. For example, $z_k = 1$ is mapped to the actual value of the $k$-th feature and hence keeps it working, while $z_k = 0$ is mapped to an appropriate non-informative value to block the $k$-th feature. We use a random sample from another grid cell as a non-information value for continuous features. For dummy features, we simply mute them by zeros. Based on the 1,000 perturbed observations $\tilde{x}$, we generate 1,000 counter-factual predictions $\tilde{y}$ made by the ST-CNN model. Then we build the following interpretable linear surrogate model around each prediction, and train it on the newly constructed dataset that includes the 1,000 samples of counterfactual model outputs $\tilde{y}$, and the simplified binary inputs $z$:

$$\tilde{y} = f(\tilde{x}) = f(h_x(z)) = \phi_0 + \sum_{k=1}^{K} \phi_k z_k,$$

where $h_x$ maps the simplified binary inputs to the original features for a particular single input $x$, and $f$ represents our trained ST-CNN model. Clearly, $\phi_k$ measures the marginal impact of the $k$-th feature on the model output of ST-CNN, which is the wildfire probability. The coefficient $\phi_k$ will be different for local models around different observations, and the overall magnitude of these $|\phi_k|$ values can be considered as the feature importance of the ST-CNN model. We build 96,149 such local surrogate models with 1,000 counter-factual predictions for each of the grid cells in the test set, and use Lasso to select the most important features that contribute to the probability change. Figure 14 presents the SHAP value of each model features of the preferred ST-CNN with 3-d kernels of size $(7, 9)$. Similar to the logistic regression results, the slope, daily maximum air temperature, Diablo and Santa Ana wind events, as well as the vegetative conditions, all have large impacts on the wildfire probabilities.

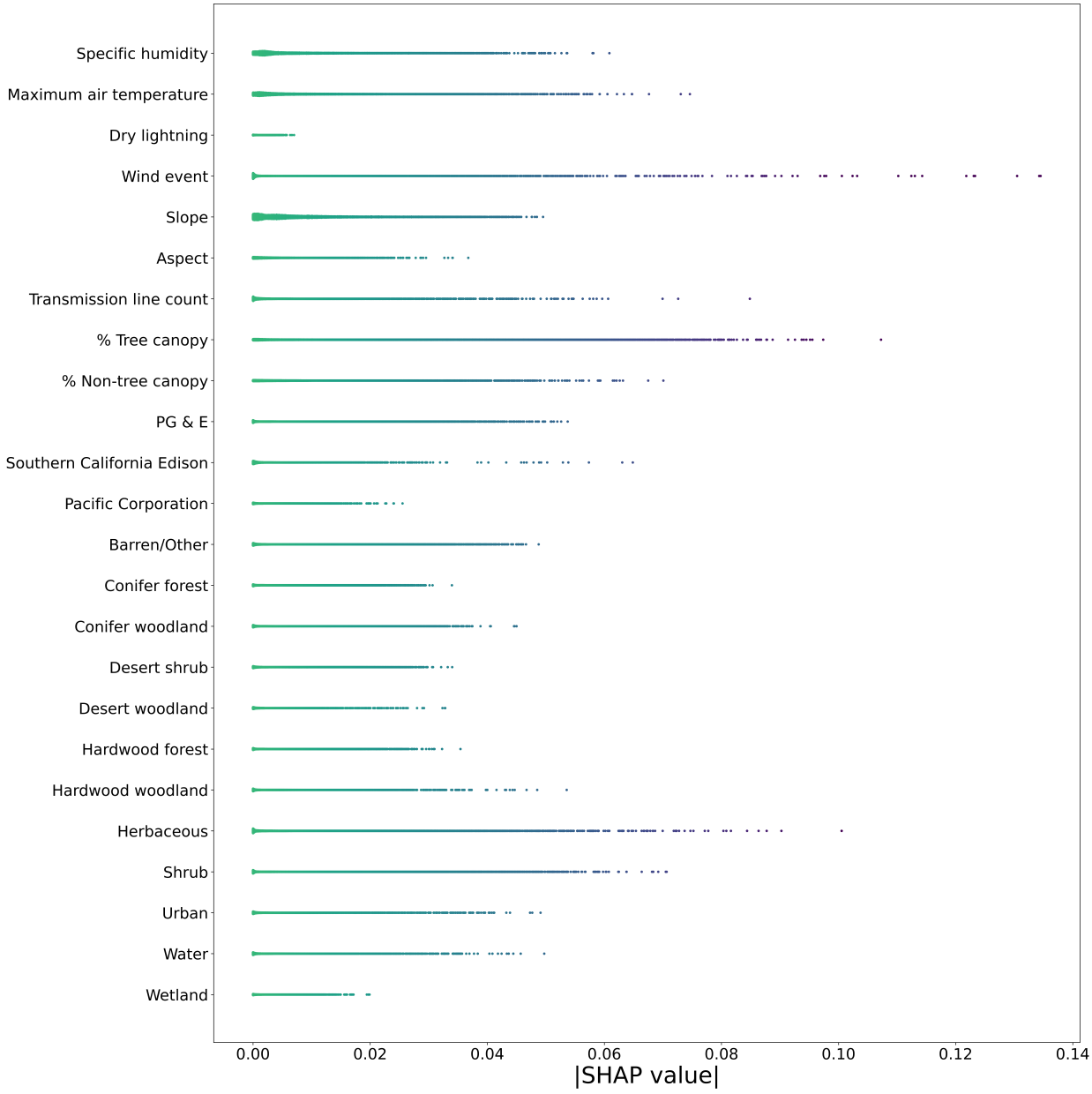



Figure 14: **SHAP feature importance values of the ST-CNN model.** Each dot in this figure represents the magnitude of marginal contribution of the corresponding feature estimated from a specific local surrogate model. The estimated coefficients are carefully standardized to reflect the impact of one standard deviation change for continuous variables, and from zero to one for dummy variables.

## 5.3 Model comparison

We further evaluate the performance of the ST-CNN model by comparing its Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves with the baseline logistic regression. We then discuss our threshold optimization approach and present the results on more evaluation metrics. Last, we discuss the tradeoffs between the two types of potential errors when determining a specific threshold on these curves.

### 5.3.1 ST-CNN vs. logistic regression

Table 4 presents metrics to assess the relative prediction performance of the logistic regression model compared to performance of the ST-CNNs with several representative kernel sizes $(1, 1)$, $(3, 3)$ and $(7, 9)$. The table reports results for the area under the ROC curve (ROC-AUC); the area under the PR curve (PR-AUC); the Cross Entropy, measuring the differences between the predicted probability and the true class assignments of the four models; and the mean-squared error measuring the relative accuracy of probabilistic predictions of the four models. The advantage of these metrics is that they are all threshold-independent, which means they all accurately measure overall model performance without a specific threshold or probability value used to determine whether a prediction is classified as positive or negative.

The results in Table 4 indicate that the spatial and temporal lags are the critical modeling characteristic that can improve the wildfire prediction performance, whereas non-linearity has limited effects. Since we simply use the cross entropy loss function to train the model, the ST-CNN of kernel size $(1, 1)$ is essentially equivalent to a logistic regression with additional nonlinearity and model-learned features. On the validation set, the ROC-AUC improves only a little from 0.7287 for the logistic model to 0.7301 for ST-CNN $(1, 1)$, and the PR-AUC almost remains constant. On the test set, the ROC-AUC improves a bit more from 0.6773 to 0.7033, while the PR-AUC still has very limited improvement. On the contrary, when comparing ST-CNN $(1, 1)$ with the best-performing kernel size $(7, 9)$, we can find a large growth in ROC-AUC from 0.7301 to 0.8130, and an obvious increase in PR-AUC from 0.1180 to 0.1766 on the validation set. On the test set, we can also observe significant improvement from ST-CNN $(1, 1)$ to $(7, 9)$, with ROC-AUC rising from 0.7033 to 0.7717 and PR-AUC rising from 0.0672 to 0.1071. This indicates the non-linearity will not contribute much to the classifier's ability to distinguish the two classes. Instead, the spatial and temporal lags, which corresponds to the cumulative and surrounding weather and vegetative conditions, matter more for the wildfire prediction than just nonlinearity. However, even the simplest ST-CNN $(1, 1)$ still has advantages over the baseline logistic model, since no hand-crafted features are needed to achieve the same level of performance.

| Hyperparameters $(p, k)$ | Validation | | | | Test | | | |
|---|---|---|---|---|---|---|---|---|
| Metric | logistic | $(1,1)$ | $(3,3)$ | $(7,9)$ | logistic | $(1,1)$ | $(3,3)$ | $(7,9)$ |
| ROC-AUC | 0.7287 | 0.7301 | 0.7730 | 0.8130 | 0.6773 | 0.7033 | 0.7278 | 0.7717 |
| PR-AUC | 0.1189 | 0.1180 | 0.1376 | 0.1766 | 0.0600 | 0.0672 | 0.0747 | 0.1071 |
| Cross Entropy | 0.2586 | 0.2604 | 0.2509 | 0.2169 | 0.1772 | 0.1766 | 0.1762 | 0.1626 |
| Mean Squared Error | 0.0615 | 0.0617 | 0.0612 | 0.0582 | 0.0403 | 0.0402 | 0.0402 | 0.0391 |

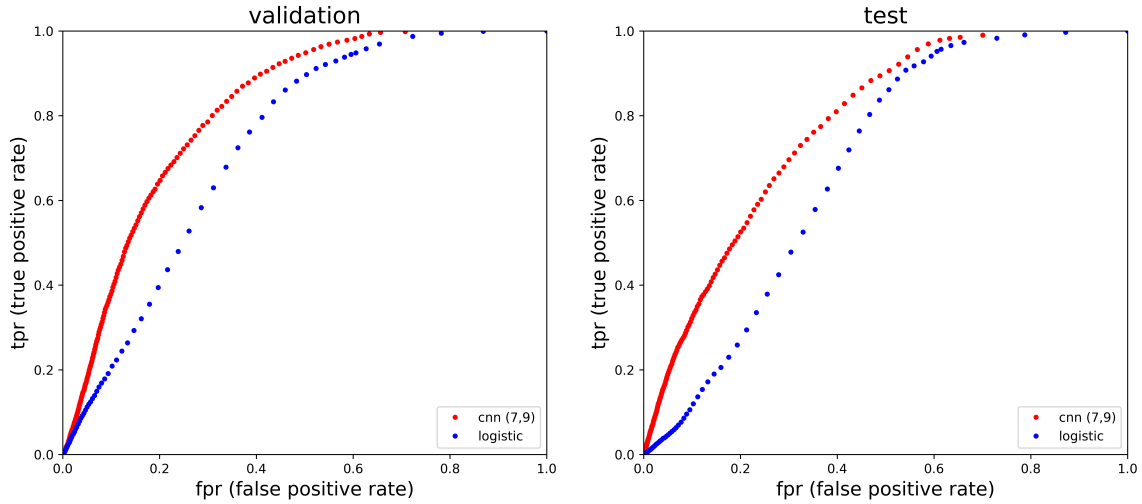
Table 4: **Comparing the logistic regression with ST-CNNs.** The hyperparameter $p$ refers to order of the temporal lags, and $k$ refers to the order of spatial lags.

Figure 15 graphically presents the superior performance of the ST-CNN for the preferred $(7, 9)$ model versus the logistic regression for both the validation and the test samples. As shown, both ROC and PR curves for the ST-CNN $(7, 9)$ model are almost always above those of the logistic model. In other words, conditioning on the same value of false positive rate/Recall, the ST-CNN model can perform better in true positive rate and Precision.

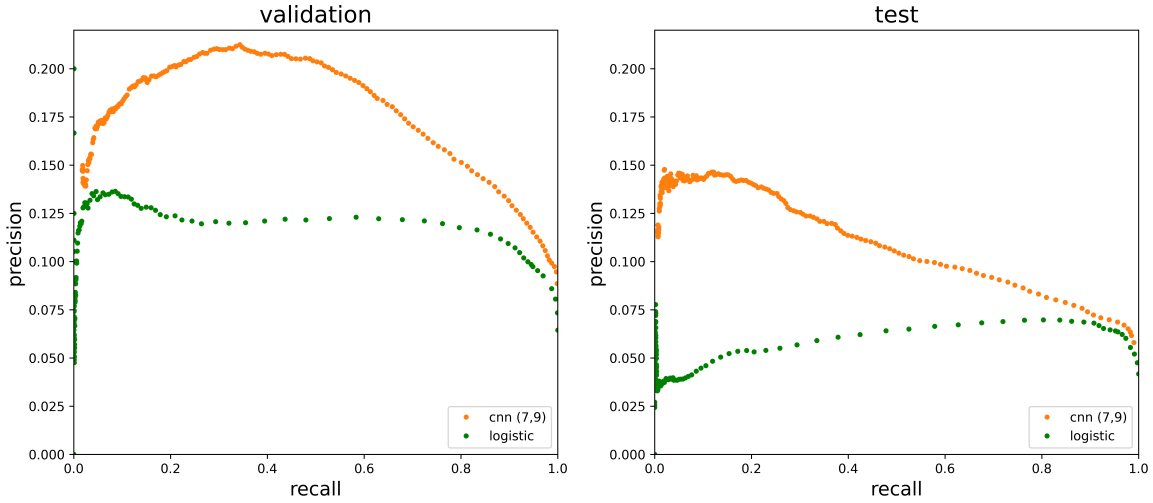
We also consider a further extension of our ST-CNN by conducting experiments from adding a long-short-term memory component to the CNN. Long short-term memory (LSTM) is a type of recurrent neural network (RNN) model with well-designed memory and gating structure for sequence prediction that could, at least potentially, be well suited to the time-series evolution of wildfire incidence over our time period. However, as discussed in detail in Appendix B, we find no evidence that CNN-LSTM outperforms the ST-CNN with max-pooling. Rather, the ST-CNN model with kernel size $(7, 9)$ consistently achieves higher AUC values on training, validation and test sets. Since the convolutional models are more interpretable, are computationally cheaper, and outperform LSTM, we conclude that the ST-CNN model is preferred.

### 5.3.2 Threshold optimization

Another important objective of our analysis is to achieve higher classification accuracy. Since the distribution of wildfire is extremely imbalanced, it is challenging to determine an appropriate threshold ex ante. As explained in Section 3.8, we follow Leevy et al. (2023) and optimize the threshold by maximizing the geometric mean (G-mean) of the true positive rate (TPR) and the true negative rate (TNR) on the training data, while satisfying the constraint TPR $\geq$ TNR. We consider both a constant threshold as well as moving-average thresholds with different window sizes. We find that the model performance, such as the G-mean and F1 score, does not change much with various methods. Therefore, for each year $t$, we simply use all data up until year $t$ to determine the optimal threshold. We report details for different




(a) Receiver Operating Characteristic (ROC) Curve




(b) Precision-Recall (PR) Curve

Figure 15: **Comparing the performance of ST-CNN (7, 9) model with the logistic model on both validation and test samples.** This figure presents the ROC and PR curves for the logistic model and the preferred ST-CNN model with kernel size $(7, 9)$. In the ROC curve, we plot the true positive rate (TPR) against the false positive rate (FPR), while in the PR curve, we plot the Precision against Recall. Each dot in these graph represents a specific threshold of probability for the positive class.

thresholds in Appendix C.

For each prediction model, we calculate the ex-ante optimal thresholds which maximize the G-mean and satisfy TPR $\geq$ TNR on all historical data, for both the validation and out-of-sample test sets. Given these thresholds, Table 5 allows for a more nuanced evaluation and interpretation of the differing hyperparameter structures by comparing a range of threshold-dependent performance statistics and comparing the added spatiotemporal lags of the ST-CNNs to the benchmark logistic regression forecasting model. As shown in Table 5, the ST-CNNs uniformly outperform the logistic regression in terms of classification accuracy, F1 score, and G-mean for both validation and test sets. Similar to Table 4, the primary modeling feature that drives the improvement of ST-CNN compared with logistic regression is the spatial and temporal lags. Besides, given the inherent randomness of wildfire occurrence across time, it is rather difficult to elevate Precision as opposed to Recall. Therefore, the increasing Precision of the spatiotemporal CNN models for both validation and test sets is particularly noteworthy. Assuming that we naively model wildfire as Poisson events, the high Recall and low Precision would correspond to the fact that it is easier to identify the conditions that are associated with high arrival rates, but there is no guarantee that these events will necessarily happen over the next year. We further discuss the Precision-Recall tradeoff in Appendix D.

| Hyperparameters $(p, k)$ | Validation | | | | Test | | | |
|---|---|---|---|---|---|---|---|---|
| Metric | logistic | $(1,1)$ | $(3,3)$ | $(7,9)$ | logistic | $(1,1)$ | $(3,3)$ | $(7,9)$ |
| Accuracy | 0.6236 | 0.6578 | 0.6802 | 0.6636 | 0.5820 | 0.6151 | 0.6860 | 0.7091 |
| F1 score | 0.2068 | 0.2061 | 0.2347 | 0.2447 | 0.1257 | 0.1312 | 0.1462 | 0.1633 |
| G-mean | 0.6839 | 0.6722 | 0.7166 | 0.7420 | 0.6438 | 0.6523 | 0.6653 | 0.6947 |
| True positive | 4718 | 4270 | 4716 | 5238 | 2890 | 2795 | 2585 | 2729 |
| False positive | 34709 | 30976 | 29268 | 31385 | 39060 | 35789 | 28760 | 26683 |
| True negative | 55245 | 58978 | 60686 | 58569 | 53072 | 56343 | 63372 | 65449 |
| False negative | 1477 | 1925 | 1479 | 957 | 1127 | 1222 | 1432 | 1288 |
| Precision | 0.1197 | 0.1211 | 0.1388 | 0.1430 | 0.0689 | 0.0724 | 0.0825 | 0.0928 |
| Recall/True positive rate | 0.7616 | 0.6893 | 0.7613 | 0.8455 | 0.7194 | 0.6958 | 0.6435 | 0.6794 |
| True negative rate | 0.6141 | 0.6556 | 0.6746 | 0.6511 | 0.5760 | 0.6115 | 0.6878 | 0.7104 |

Table 5: **Comparing the logistic regression with ST-CNNs with ex-ante optimized thresholds.**

## 5.4 Out-of-sample wildfire predictions

Figure 16 presents the results of an out-of-sample forecasting exercise for annual wildfire occurrence in 2021, a year with very high wildfire incidence. For comparison, we download the most recent fire perimeters from CAL FIRE and extend our prediction to 2022, a very low actual wildfire incidence year, by simply applying the trained model to the 2021 subsequence of data observed on Dec 31, 2021.

Based upon our ST-CNN with max-pooling strategy to achieve annual out-of-sample forecasts, we present our estimates of the annual wildfire occurrence for the 2021 fire season in the left-hand side of Figure 16. Over the four years preceding 2021, California experienced severe drought conditions, termed "ecological drought conditions" that included significant periods of precipitation anomalies in which Northern California, showed running deficits of more than 6 inches of rain[33] and by September of 2020, these drought conditions led to nearly 14,000 lightning strikes that ignited 900 fires that burned more than 1.5 million acres.[34] As shown in Figure 16, our max-pooling ST-CNN strategy supports quite accurate forecasts of the grid-cell pre-conditions for 2021 and the occurrence and locations of some of the largest fires such as the Mcfarland Fire, the McCash-River Complex-Monument Fires, the Antelope Fire, the Windy-KNP Complex Fires, and the devastating Dixie and Caldor wildfires. Of course randomness remains a key driver of wildfire ignition events. The Dixie Wildfire ignition was caused by 65-foot Douglas fir tree that fell on a Pacific Gas and Electric (PG&E) transmission line and the Caldor Fire ignition event was arson. The accuracy of the ST-CNN in forecasting locational pre-conditions, based on past temporal and cross-sectional dynamics is, however, an important justification for its use in annual wildfire prediction. Notably, the ST-CNN accurately accounted for the cumulative effects of the prior drought conditions as well as the effect of these conditions on the differing vegetative types and canopy coverage of grid cells in the forecast.

The right-hand side of Figure 16 presents our annual out-of-sample wildfire occurrence forecasts for the 2022 wildfire season. Compared with 2021, the fire probability is much lower for Northern California, mainly due to the trend of the Diablo events being weaker compared with the previous years. Note that the wind events stand out in our SHAP value estimates, which means it is one of the most informative predictors.

Figure 17 presents the 2021 ST-CNN out-of-sample predictions for annual wildfire occurrence probabilities, aggregated by zip code from the average grid-cell forecasts.[35] In

[33] https://wrcc.dri.edu/anom/cal_anom.html.

[34] https://earthobservatory.nasa.gov/images/147215/california-continues-to-burn.

[35] The map of California zip codes is downloaded from the California State Geoportal (https://gis.data.ca.gov/datasets/CDEGIS::california-zip-codes/about).

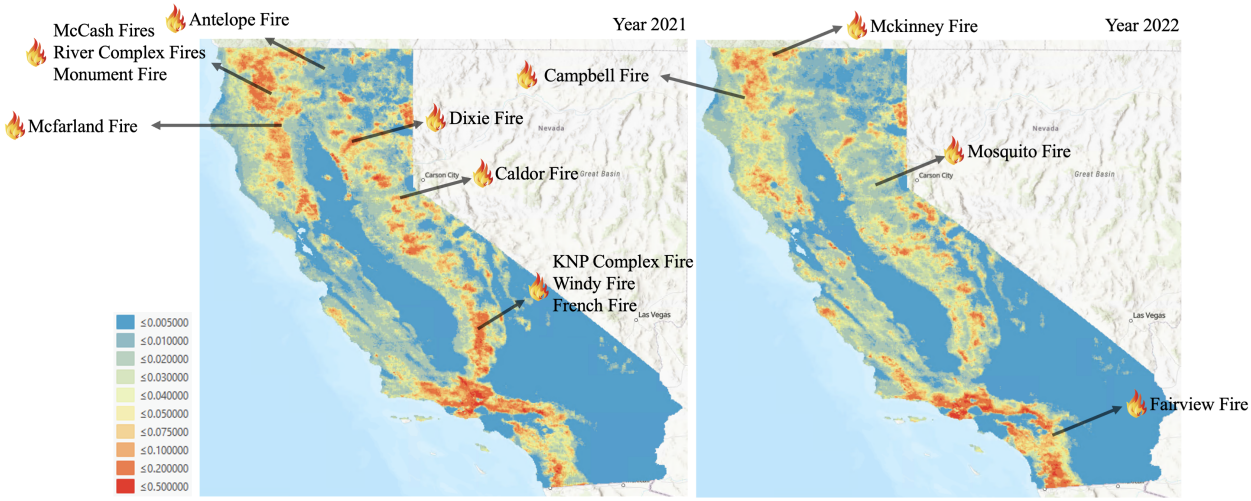



Figure 16: **CNN one-year-ahead out-of-sample wildfire prediction.**

Northern California, Trinity and Siskiyou counties have elevated out-of-sample wildfire forecasts in 2021. Among the top 25 zip codes with the highest predicted annual probability of wildfire in Trinity County, the average is 5.3% (maximum 9.4%, zip 00055, Trinity National Forest) and in Siskiyou the average is 3.9% (maximum 9.0%, zip 96031, Forks of Salmon). The vegetation in these zip codes is primarily forest and the topography is very steep, with the Trinity zip codes often exposed to strong Diablo winds and the Siskiyou zip codes often exposed to dry lightning.

On the northeastern side of the state in the foothills of the Siera Nevada, as shown in Figure 17, the ST-CNN out-of-sample forecasts for the annual probability of wildfire are also elevated in Plumas, Butte, and Mariposa counties. Among the top 25 riskiest zip codes in Plumas County, the average predicted annual probability is 3.7% (maximum 10.4%, zip 95984, Twain). In Butte County the average predicted probability is 2.7% (maximum 5.5%, zip 95969, Paradise).[36] Lastly, in Mariposa County the average predicted probability is 5.7% (maximum 12.1%, zip 95345, Midpines). The vegetation coverage in all three of these counties is characterized as conifer and hardwood woodland or forest, and they are all exposed to Diablo winds, experience frequent occurrences of dry lightning in very steep topographic areas, and have significant exposure to the often-wind-driven failures of the Pacific Gas & Electric grid. Finally, Napa County, just north of the Bay Area on the Pacific coast, also has zip codes with elevated annual predicted probabilities of wildfire, with an average of 4.9% (maximum 9.5%, zip 94576, Deer Park). The vegetation in Napa is primarily shrub and hardwood woodlands, and the Diablo winds and relatively steep topography are additional

[36]This zip code was involved in the devastating 2018 Paradise wildfire.

risk features of this county.


Figure 17: **ST-CNN one-year-ahead (2021) out-of-sample wildfire prediction at the zip-code level, labeled with county names.**

Figure 17 clearly indicates that the ST-CNN forecasts for the annual occurrence of wildfire are significantly higher in the Southern California counties of Venture, Los Angeles, and San Diego. The average annual wildfire probability for the top 25 zip codes in Los Angeles County, aggregated from the ST-CNN cell forecasts, is 12.4%. Additionally, three zip code aggregates in Los Angeles County are the highest in California: zip 91321, Newhall, with a predicted annual probability of wildfire of 27.6%, zip 91307, West Hills, with a predicted annual probability of 21.9%, and zip 91381, Stevenson Ranch, with a predicted annual probability of 21.7%. Ventura County has a zip-code level average of 5.5% annual occurrence of wildfire and has one zip code aggregate with a ST-CNN annual wildfire forecast of 19.2%. Finally in San Diego County, among the top 25 zip codes with the highest wildfire predicted probability, the average is 5.5% and the highest zip code out-of-sample forecast is 11.2%. The Southern California vegetation is primarily shrub, chapparal, and conifer woodland, the terrain in the coastal range is often very steep and the whole area is subject to the continuing hazards of the drying effects of the Santa Ana winds. The actual number of California wildfires in 2021 was 7,396 wildfires, which burned a total of 2,569,386 acres.[37]

[37]See https://www.fire.ca.gov/incidents/2021.

## 5.5 ST-CNN out-of-sample wildfire estimated losses

The ST-CNN generates a wildfire predicted probability for each of the 96,149 grid cells in California. However, being located in a burned grid cell does not mean a residential house will necessarily be burned, since the entire 2 km $\times$ 2 km cell may not burn evenly. Furthermore, even if a house is affected by wildfire, its value loss may range from 0% to 100% depending on the structural conditions of the house. To bridge the gap between cell-level fire probabilities and the house-level fire losses, we use the historical wildfire and property assessed value data to estimate the percentage of value loss for each house in the burned grid cells.

To obtain the percentage of value loss for residential structures post wildfires, we construct a panel dataset of the assessed value of the improvement for the fiscal year pre-wildfire, and the year post-wildfire, using ATTOM Assessor Files, for all residential single family properties found within the grid cells that intersect with the CAL FIRE wildfire perimeters each year. The residential single family house data are obtained from the ATTOM Assessor Files that have been merged to the burned grid cells. The Wildland Urban Interface (WUI) shapefile is obtained from the Silvis Lab for Spatial Analysis for Conservation and Sustainability at the University of Wisconsin.[38] The slope, elevation and aspect of the house are obtained are obtained from the U.S. Geological Survey based upon the latitude and longitude of the property.[39]

We report the results of the regression of percentage loss in pre-wildfire assessed value of the improvement on standard normal transformations for all of the non-dummy features. As shown in Table 6, we find that the baseline average loss is 1.38% of the assessed value of the improvement for wildfire grid-cell exposed residential single family properties. Note that this figure is not the value loss conditional on fire, but conditional on the burned grid cell. The loss percentage increases to 2.52% for properties built before California strengthened its wildfire related building codes in 2008. Percentage losses also increase with the standardized year of the building's age. Single family residential properties at higher slopes, elevations and southwesterly aspects also experience higher losses. Lower housing density with Vegetation WUI classes is also associated with higher losses. The loss percentage will be on average 3.30% higher if a house is in Vegetation land with low or medium housing density. This makes sense given that a house within such WUI classes are more likely to burn. Overall, these results confirm the merits of mandated building codes in reducing residential single family property losses from wildfire (see Baylis and Boomhower, 2021).

Figure 18 presents the out-of-sample forecast for the per-house loss to the assessed value of improvement from wildfire for 2021. The key to the figure reports the dollar amount of loss

[38] https://www.arcgis.com/home/item.html?id=a4985d64969743db8feddf01c96c9435.
[39] http://apps.nationalmap.gov/downloader/.

| | coef | std err | z | P> \|z\| |
|---|---|---|---|---|
| Constant | 0.0138 | 0.002 | 8.056 | 0.000 |
| Indicator: Built before 2008 codes | 0.0114 | 0.002 | 6.898 | 0.000 |
| Building Age | 0.0054 | 0.000 | 22.252 | 0.000 |
| Slope | 0.0048 | 0.000 | 19.531 | 0.000 |
| Elevation | 0.0086 | 0.000 | 34.875 | 0.000 |
| Aspect | 0.0091 | 0.000 | 38.444 | 0.000 |
| WUI Class Fixed Effects | Yes | | | |
| No Vegetation | . | . | . | . |
| Vegetation, Low Housing Density | 0.0330 | 0.001 | 28.222 | 0.000 |
| Vegetation, Medium Housing Density | 0.0340 | 0.001 | 47.396 | 0.000 |
| Vegetation, High Housing Density | −0.0155 | 0.001 | −22.186 | 0.000 |
| No. Observations: | 538,221 | | | |
| Adjusted R-squared | 0.0270 | | | |
| F-Statistic | 1,899 | | | |
| Prob (F-statistic): | 0.0000 | | | |
| Log-Likelihood: | $1.8940 \times 10^5$ | | | |

Table 6: **Regression of the percentage of the pre-wildfire residential single family value (the assessed value of the improvement) that is lost due to wildfire grid-cell exposure.** We estimate the percentage loss of the pre-wildfire assessed value of the improvement one fiscal year after a wildfire for all residential single family structures located within the burned grid cells 2010 through 2019. The residential single family house data are obtained from the ATTOM Assessor Files that have been merged to the CAL FIRE burn area shape files. The WUI class data are obtained from the Silvis Lab for Spatial Analysis for Conservation and Sustainability at the University of Wisconsin (https://www.arcgis.com/home/item.html?id=a4985d64969743db8feddf01c96c9435). The slope, elevation of the house and its aspect are obtained from the U.S. Geological Survey (see http://apps.nationalmap.gov/downloader/), given the latitude and longitude of the property.

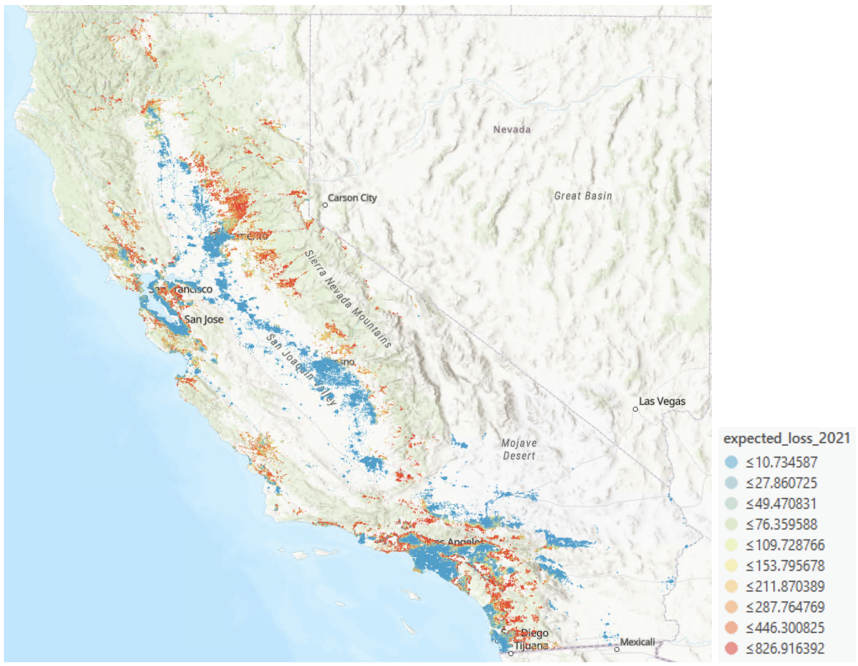



Figure 18: **ST-CNN one-year-ahead (2021) out-of-sample expected annual loss of the assessed value of the improvement for all single family houses in California.**

using quintile cutoffs for visualization. The losses were calculated as the grid-cell predicted probability times the percentage of assessed value loss given by house characteristics times the 2020 assessed value of the improvement house by house. As shown, the areas with the highest expected structural wildfire losses are found in the densely populated Bay Area and the Los Angeles Basin from Santa Barbara south to San Diego as well as the grid cells found in the increasingly populated areas of Sacramento, Placer and El Dorado counties in the foothills of the Sierra Nevada. Compared with Figure 17, the vast forest areas of Siskiyou and Trinity are not as risky in terms of the expected losses to residential assessment because of the low population and housing density, but the fast-spreading wildfire in those dense forest area could pose huge threats to the ecological environment in Northern California. Overall, the expected residential single family structural losses from our ST-CNN wildfire predictions and our empirical estimates of the expected losses to residential assessment from these fires are $0.9 billion for 2021.

# 6 ST-CNN and insurance risk

Given the Section 2 discussion concerning the recent intertemporal loss-smoothing problems of the U.S. fire peril insurance industry, the 2021 California wildfire season was especially damaging to State Farm General, the largest P&C insurance carrier in California. As re-

ported by the National Association of Insurance Commissioners (NAIC), State Farm General experienced a 144.8% fire-peril loss ratio in 2021[40] and AM Best downgraded its Financial Strength Rating (FSR) to B (Fair) from A (Excellent) and the Long-Term Issuer Credit Rating (Long-Term ICR) to "bb+" (Fair) from "a" (Excellent).[41] In March, 2024 State Farm General announced that they would not renew about 30,000 residential P&C policies[42] and then in June, 2024 State Farm General announced that they would renew some of those policies but only if policy holders obtained their fire peril insurance from the California Fair Plan.[43] In September of 2024, State Farm General, the California subsidiary of the State Farm group, requested a 30% rate increase for homeowners insurance in California, citing concerns about financial solvency and the need to protect itself from potential insolvency due to rising costs and risks in the state. As part of State Farm's 30% rate hike request to the California Department of Insurance, State Farm's Exhibit 1 documented its prior insurance rate increases, its use of vendor wildfire-risk models, and the zip codes where Start Farm had not renewed at least one policy.[44]

The models that State Farm identified for its 2021 eligibility filings included CoreLogic Brushfire; the CoreLogic RQE; AIR Touchstone; and the GRID Fire Model. Of course, all of these models are proprietary and State Farm argued, "Public disclosure to competitors of eligibility criteria that constitute confidential trade secret information is bad public policy and impairs competition."[45] State Farm also provided a table of the impacted zip codes for the 28,758 P&C policy non-renewals that the firm had scheduled for 317 zip codes in California.

Table 7 presents the ST-CNN predicted 2020–2022 out-of-sample wildfire probabilities for the 317 zip codes[46] where State Farm did not renew at least one P&C policy and the average forecast probabilities over the same period for the zip codes not included in the State Farm non-renewal actions. At least at the empirical mean, our ST-CNN predicted wildfire probabilities support the State Farm zip code classifications. As shown, the non-renewal zip codes have a higher wildfire probability mean of 3.73%, standard deviation of 3.82%, whereas the zip codes not involved in State Farm's non-renewals have a lower wildfire probability mean of 1.59%, standard deviation of 2.30%. However, the non-renewed

[40] https://www.insurance.ca.gov/01-consumers/120-company/04-mrktshare/2021/upload/Top25grps2021wa_Revised.pdf.

[41] See https://news.ambest.com/pr/PressContent.aspx?refnum=34559&altsrc=2.

[42] https://www.insurancejournal.com/news/west/2024/03/20/765883.htm.

[43] https://news.ambest.com/newscontent.aspx?refnum=258346&altsrc=23.

[44] https://srp-prod-public-pdfs.s3-us-west-2.amazonaws.com/33f3f200-083c-560f-8b03-f1cc1f939577.

[45] https://srp-prod-public-pdfs.s3-us-west-2.amazonaws.com/33f3f200-083c-560f-8b03-f1cc1f939577.

[46] Among them, 8 of the very small zip codes are not geocoded in the California zip code shapefile.

classification also includes grid-cell aggregates with quite low ST-CNN wildfire probabilities, such as the minimum of 0.06% suggesting the possible misclassification of zip codes as risky when our ST-CNN models finds that they are not. Similarly, the zip codes not included in State Farm's non-renewal actions include an upper tail of very high predicted out-of-sample wildfire occurrence probabilities, such as the maximum of 16.99%, again suggesting misclassification of zip codes as riskless where the ST-CNN model finds that they are not.

| | ST-CNN average 2020–2022 predicted wildfire annual probability | |
|---|---|---|
| | Zip codes with at least one non-renewed policy | Zip codes not included in State Farm non-renewals |
| count | 309 | 1411 |
| mean | 0.0373 | 0.0159 |
| std | 0.0382 | 0.0230 |
| min | 0.0006 | 0.0000 |
| 25% | 0.0102 | 0.0008 |
| 50% | 0.0276 | 0.0054 |
| 75% | 0.0514 | 0.0225 |
| max | 0.2554 | 0.1699 |

Table 7: **Summary statistics for the ST-CNN average 2020–2022 predicted out-of-sample zip code aggregate grid-cell wildfire probabilities for zip codes with at least one non-renewed State Farm P&C policy and zip codes not included in State Farm's non-renewal announcement.** We calculate wildfire probabilities at the zip code level by averaging the cell-level forecasts.

On January 9, 2025, in response to the devastating Los Angeles wildfires, the California Insurance Commissioner, Ricardo Lara, implemented a mandatory one-year moratorium on insurance non-renewals and cancellations for residents within the Los Angeles County fire perimeter and adjacent zip codes. This moratorium prevents insurance companies from dropping policies in and adjacent to zip codes of the Palisades and Eaton fires in Los Angeles County, regardless of whether they suffered direct property loss.[47] On January 15, the moratorium was expanded to also cover the zip codes affected by the Hurst, Lidia, Sunset, and Woodley fires.[48]

As an indication of the accuracy of the ST-CNN 2020–2022 average aggregate zip code forecasts for wildfire occurrence, we extend the analysis presented in Table 7 to consider how

[47] https://www.insurance.ca.gov/0400-news/0100-press-releases/2025/release005-2025.cfm.

[48] https://www.insurance.ca.gov/0400-news/0100-press-releases/2025/release008-2025.cfm. The moratorium zip codes were published by the California Insurance Commissioner in Bulletin 2025-1, https://www.insurance.ca.gov/0250-insurers/0300-insurers/0200-bulletins/bulletin-notices-commiss-opinion/upload/Bulletin-2025-1-One-Year-Moratorium-Updated-ZIP-Codes-Amended-17-January-2025.pdf.

the ST-CNN predicted probabilities compare with State Farm's recent (April 2024) non-renewal activity in California. We also compare the ST-CNN results to the actual wildfire experience of the zip codes most seriously involved in the Los Angeles wildfires of January 2025. As was made clear in Section 5.4, the ST-CNN forecasts vary from year to year depending on meteorological conditions, vegetative coverage, and other features used in the model. We are thus using the 2020–2022 averages to infer expected outcomes in a period of three years beyond our one-year out-of-sample forecast.

Table 8 presents the top ST-CNN model aggregate zip code estimates for all zip codes in which State Farm chose not to renew at least one P&C policy in April 2024.[49] Column 3 of Table 8 shows the 2020–2022 average aggregate ST-CNN wildfire probabilities, and Column 4 presents the top twenty-five ST-CNN annual expected aggregate tax-basis losses for 2020–2022 (losses to the assessed value of the single family improvements). Column 6 presents the zip-code percentage of State Farm's P&C policies not renewed and Column 7 indicates whether the zip code was impacted by the California Department of Insurance's 2025 mandatory moratorium on cancellations and new-renewals for the Los Angeles County wildfires. As shown in Table 8, 40% of zip codes with the largest predicted tax basis losses were designated as "moratoria zip codes" due to their exposure to the January 2025 Eaton, Palisades, and Hurst wildfires. As shown, three of the top twenty zip codes by aggregate tax-basis losses were also the three zip codes that experienced the brunt of the direct wildfire losses in the 2025 wildfires including Altadena (91001) for the Eaton fire and Malibu (90265) and Pacific Palisades (90272) for the Palisades fire.

As shown in Table 8, State Farm's average zip-code April 2024 cancellation rates were 14.44% for the Eaton fire, 18.71% for Hurst fire, and 10.11% for the Palisades fire. The State Farm 2024 non-renewal rate for the highest-actual-burn-count zip code in the 2025 Eaton fire (Altadema) was 3.05% and the average non-renewal rate was 58.18% for the two highest-actual-burn-count zip codes in the Palisades fire (Pacific Palisades and Malibu). The 2020–2022 ST-CNN's expected loss to the tax-basis was about $7.6 million for the Altadena assessed values and was about $19.3 million for the structure assessments of Malibu and about $7.7 million for the structure assessments of Pacific Palisades. This result suggests that State Farm's renewal decisions were more focused on high valuations than on the absolute risk of locations, since Altadena has a relatively high predicted annual average 2020–2022

[49]For comparison, Table 11 in Appendix E presents the 120 top rankings by zip code for the ST-CNN estimates of average aggregate tax basis losses (2020–2022) for the entire state of California. As shown, the Table 11 zip-code rankings closely correspond to the top 25 rankings reported in Table 8. The exceptions include Beaumont (92223), San Bernardino (92407), Rancho Cucamonga (91737) and Tujunga (91042) that experienced no State Farm non-renewals in 2024, and the zip codes of Murrieta (92562), Lake Arrowhead (92352), Fremont (94539), and Newhall (92321) that were classified in the top 35 overall rankings in Table 11.

| Ranking | Zip code | ST-CNN Aver. 2020–2022 expected wildfire annual prob. | Loss to residential assessment | City | State Farm zip code nonrenewal rate | CA mandatory moratorium on cancellations and non-renewals January 2025 |
|---|---|---|---|---|---|---|
| | | % | $ | | % | |
| 1 | 91011 | 13.59 | 26,590,949 | La Canada Flintridge | 8.29 | Eaton |
| 2 | 93108 | 5.87 | 20,793,133 | Santa Barbara | 28.65 | |
| 3 | 90265 | 12.15 | 19,267,521 | Malibu | 46.92 | Palisades* |
| 4 | 91302 | 9.92 | 15,065,146 | Calabasas | 60.39 | Palisades |
| 5 | 91214 | 20.20 | 14,761,026 | La Crescenta | 1.17 | Eaton |
| 6 | 92399 | 7.45 | 14,672,026 | Yucaipa | 0.67 | |
| 7 | 91387 | 15.01 | 14,008,202 | Canyon Country | 0.72 | Hurst |
| 8 | 91362 | 10.11 | 13,658,757 | Thousand Oaks | 16.97 | |
| 9 | 92592 | 5.65 | 11,694,243 | Temecula | 0.04 | |
| 10 | 96161 | 1.92 | 10,914,561 | Truckee | 1.32 | |
| 11 | 91390 | 18.17 | 10,321,079 | Santa Clarita | 0.37 | Hurst |
| 12 | 92065 | 7.59 | 10,255,392 | Ramona | 2.40 | |
| 13 | 93063 | 11.91 | 9,934,509 | Simi Valley | 2.65 | |
| 14 | 92679 | 6.03 | 9,686,712 | Trabuco Canyon | 4.26 | |
| 15 | 91381 | 19.24 | 8,580,735 | Stevenson Ranch | 0.61 | Hurst |
| 16 | 91901 | 8.26 | 8,387,484 | Alpine | 6.63 | |
| 17 | 91361 | 13.39 | 8,302,593 | Westlake Village | 19.07 | Palisades |
| 18 | 93065 | 9.92 | 8,237,446 | Simi Valley | 0.82 | |
| 19 | 90272 | 4.24 | 7,668,455 | Pacific Palisades | 69.43 | Palisades* |
| 20 | 91001 | 9.54 | 7,612,629 | Altadena | 3.05 | Eaton* |
| 21 | 91320 | 10.86 | 7,372,269 | Newbury Park | 1.75 | Palisades |
| 22 | 92562 | 5.14 | 7,356,871 | Murrieta | 2.98 | |
| 23 | 92352 | 3.14 | 7,061,444 | Lake Arrowhead | 31.60 | |
| 24 | 94539 | 2.77 | 7,052,555 | Fremont | 0.06 | |
| 25 | 91321 | 25.54 | 6,963,251 | Newhall | 2.27 | Hurst |

Table 8: **Top twenty-five rankings of the average 2020–2022 ST-CNN forecasts for losses to the residential improvement assessment for zip codes in which State Farm General did not renew at least one P&C policy.** This table presents the top 25 zip codes ranked by the ST-CNN expected loss to the assessed value of the improvement (the tax basis) given the average 2020–2022 ST-CNN predicted probability of wildfire occurrence. For each zip code, the table reports the percentage of State Farm's P&C policies that were not renewed in April 2024. We also report whether single family residential houses in each zip code were impacted by the California mandatory moratorium on cancellations and new-renewals after the Los Angeles January 2025 wildfire. Starred wildfires indicate that the zip code is both under the 2025 moratorium and also was a zip code where one or more houses actually burned in the January 2025 wildfires.

wildfire probability of 9.54% whereas Malibu and Pacific Palisades have an average average 2020–2022 probability of wildfire of 8.20%. Although the Hurst Fire is within the California Department of Insurance's designated moratorium area, the fire was fully contained on Jan. 16 and did not result in any structures being destroyed or cause any deaths. The Hurst wildfire outcome is strikingly different from the Eaton and the Palisades fires, that together killed a combined 29 people, destroyed more than 9,400 homes and burned nearly 58 square miles.

Overall, these results suggest that on average State Farm's non-renewal decisions correspond well to the ST-CNN 2020–2022 average zip-code level predicted probabilities of wildfire for the Palisades zip codes. However, State Farm's renewal decisions appear to have left it exposed to substantial remaining risk for the 2025 Eaton wildfire that had very low non-renewal rates given the high ST-CNN probability of wildfire in those zip codes. This apparent mismatch of actual non-renewal rates to predicted wildfire likelihood has clearly exposed California State Farm General to capital losses given its second quarter 2024 loss ratio of 94.8%.[50] The firm is expected to have a 2025 wildfire loss of $7.6 billion before re-insurance payouts from its parent company, State Farm Mutual Automobile Insurance Company. Smoothing these losses over time, as discussed in Section 2, will present a major challenge to the continued viability of State Farm General in the State of California and to State Farm's continued P&C coverage for its 800,000 current California policy holders.

# 7 Conclusions

Wildfire risk is escalating rapidly in California, driven in part by climatological factors such as rising maximum temperatures from May to October in the Western States, increasing urbanization in risk-prone locations, and global forces like El Niño and La Niña that affect drought cycles and snow fall during winter months. Given the very large economic costs associated with wildfires, property and casualty (P&C) insurance companies in California are increasingly questioning their financial survivability. In part, their financial problems have arisen from inherent moral-hazard problems associated with insurance, as well as from legal and institutional frictions that make it difficult for insurance companies to spread wildfire risk over time.

More uniquely, California has a self-inflicted regulatory problem that required wildfire insurers to set rates for future annual catastrophic coverage as the fraction of damages accrued from the 20-year historical mean rather than based on forward-looking statistical or actuarial models. Additionally, only since December 2024 has the California Department

[50] See https://www.insurancejournal.com/news/national/2024/10/25/798514.htm.

of Insurance allowed the inclusion of cost, or changes in the cost, of reinsurance risk to be included in insurer rate requests.[51] As a result of these regulatory issues, California's annual P&C rates now rank next to the lowest in the U.S. (see Oh et al., 2024), leading to insurance company fragility that threatens the future ability of California homeowners to successfully rebuild after fires or even to have access to the mortgage markets.

Since the prohibition on probabilistic models was only lifted in December 2024, this paper proposes a new class of model for wildfire occurrence risk based on spatiotemporal Convolutional Neural Networks (ST-CNNs). These models are uniquely suited to forecast wildfires across the state of California based on highly imbalanced data and numerous important causal features that are characterized by heavily right-skewed distributions. The ST-CNNs capture spatial and temporal dependencies interactively, and can identify correlations between neighboring data points in spatiotemporal panel data such as wildfires. We also find that ST-CNNs significantly outperform logistic regression and LSTM, a baseline machine learning model for time series forecasting, in estimating the likelihood of wildfires.

Using out-of-sample forecasts of wildfire occurrence probability from the ST-CNN model, we estimate expected annual fire-related property losses for thousands of grid cells across the state. We find wide variation in the pre-conditions of wildfire and the estimated probability of wildfire occurrence across the northern and southern areas of the state. Overall, we find that the estimated out-of-sample predicted probability of wildfire at specific geographic locations corresponds quite closely to the largest wildfires that occurred out-of-sample in 2021, and 2022. The ST-CNN model results also perform well in accurately identifying the high risk zip codes that burned in the January 2025 Los Angeles wildfires. Comparing our highest risk zip code estimates to the State Farm General's decisions not to renew about 28,758 P&C policies in 317 California zip codes in April 2024 raises some questions about the accuracy of their models to prioritize risk.

Finally, our results raise concerns about the future financial well-being of the U.S. and California P&C insurance industry and the likely harms to homeowners from modeled-based eligibility classifications that are being applied by the largest carrier in the state. With the partial lifting of actuarial-model prohibitions in California, future work should be able to apply our ST-CNN wildfire estimates to develop economically justifiable premium reductions in exchange for homeowner mitigation investments. These investments should significantly reduce the risk of property losses from wildfire despite the growing external pre-conditions of these risks and property-level mitigation should better align the incentives of P&C insurers, banks, and homeowners, thus reducing moral hazard.

[51] https://www.insurance.ca.gov/0400-news/0100-press-releases/2024/release065-2024.cfm.

# Appendices

## A Variable definitions for the loss-given-fire regression

| Variable | Count | Mean | Std | Min | Max |
|---|---|---|---|---|---|
| Indicator: Built before 2008 codes | 538,221 | 0.9784 | 0.1453 | 0.000 | 1.000 |
| Building Age | 538,221 | 38.9871 | 20.9298 | 1.000 | 216.000 |
| Slope | 538,221 | 3.8836 | 4.1755 | 0.000 | 52.125 |
| Elevation | 538,221 | 307.6384 | 274.9268 | −55.139 | 2538.623 |
| Aspect | 538,221 | 0.2347 | 0.6887 | −1.000 | 1.000 |
| WUI Class Fixed Effects | | | | | |
| No Vegetation | 538,221 | 0.1639 | 0.3702 | 0.000 | 1.000 |
| Vegetation, Low Housing Density | 538,221 | 0.0618 | 0.2408 | 0.000 | 1.000 |
| Vegetation, Medium Housing Density | 538,221 | 0.4012 | 0.4901 | 0.000 | 1.000 |
| Vegetation, High Housing Density | 538,221 | 0.3731 | 0.4836 | 0.000 | 1.000 |

Table 9: **Variable definitions and summary statistics.** These samples are sourced from ATTOM Data Solutions assessor data and include all houses located within the burned grid cells from 2010 through 2019. We report summary statistics for properties that burned and those that did not burn within cells identified by CAL FIRE as wildfire burn areas.

## B Extended CNN-LSTM model results

In our wildfire prediction model, we use a 3-d convolutional layer, with spatial and temporal lags, to capture the daily fire potential for each grid cell. Additionally, as explained in Section 3.7, we use max-pooling to aggregate those daily measures to annual representations, which simply picks the maximum daily fire potential measures in the past 365 days.

For comparison, we also conduct experiments with the CNN-LSTM model, but we find no evidence that the CNN-LSTM could outperform the ST-CNN model with maxpooling. Instead, the ST-CNN with the best kernel size $(7, 9)$ consistently performs better than CNN-LSTM on training, validation and test sets. The Long short-term memory (LSTM) is a type of recurrent neural network (RNN) model with well-designed memory and gating structure for sequence prediction. Essentially, a RNN model is a state-machine model, where a multi-dimensional vector of real values is used as the state $s_t$, and the model generates a sequence of outputs $y_t$. At each step, the hidden state $s_t$ is updated based on both the exogenous

| | $k=1$ | $k=3$ | $k=5$ | $k=7$ | $k=9$ |
|---|---|---|---|---|---|
| Training | 0.7340 | 0.7707 | 0.7837 | 0.7938 | 0.7965 |
| Validation | 0.7287 | 0.7705 | 0.7931 | 0.7876 | 0.7913 |
| Test | 0.6489 | 0.7186 | 0.7472 | 0.7388 | 0.7421 |

Table 10: **The CNN-LSTM Model.** The Area under the curve (AUC) performance on training, validation and test samples, for CNN-LSTM models with different hyper-parameters.

observation $x_t$ and the previous hidden state $s_{t-1}$, and the model produces output $y_t$ according to the updated state $s_t$. Built upon RNN, the LSTM model features a memory cell, which functions as the state vector, and three gating networks, including the input gate, the forget gate and the output gate. Because these gates can work jointly to optimize what information to keep and what to throw away in the state vector, the LSTM model is able to capture both long-term and short-term dependencies in the data, and hence has amazing performance on time series prediction. Given its sophisticated design, we try to extend our ST-CNN to CNN-LSTM by replacing the spatiotemporal convolutional layer with a LSTM layer, so that we will rely on the LSTM structure to learn the next-year fire probability from daily meteorological data, instead of the 3-d convolution with max-pooling. Specifically, we change the model on the left for Input 1 in Figure 10 to an LSTM model.

Table 10 presents the AUC values of the CNN-LSTM model with different kernel sizes. Compared with Table 3, clearly the CNN-LSTM model does not significantly improve the prediction results. This suggests the short-term convolution with max-pooling might be enough to predict the annual wildfire incidence. On one hand, the use of spatiotemporal convolution has guaranteed that we have taken care of the cumulative heat and the surrounding weather conditions when measuring the daily fire potential. According to Table 3, the best temporal lags are all small in training, validation and test, which implies that for daily incidence the short-term meteorological information seems more predictive, while the longer-term memory is not very relevant. On the other hand, because the annual fire occurrence will be labeled positive as long as a fire happens in at least one day, the fire probability at the annual level should by nature be strongly correlated with the maximum daily fire potential, especially when these daily measures have been smoothed by the spatiotemporal convolution.

Given that convolutional models are more interpretable and computationally much cheaper than LSTM, we will choose the ST-CNN model as our preferred model.

# C Threshold optimization

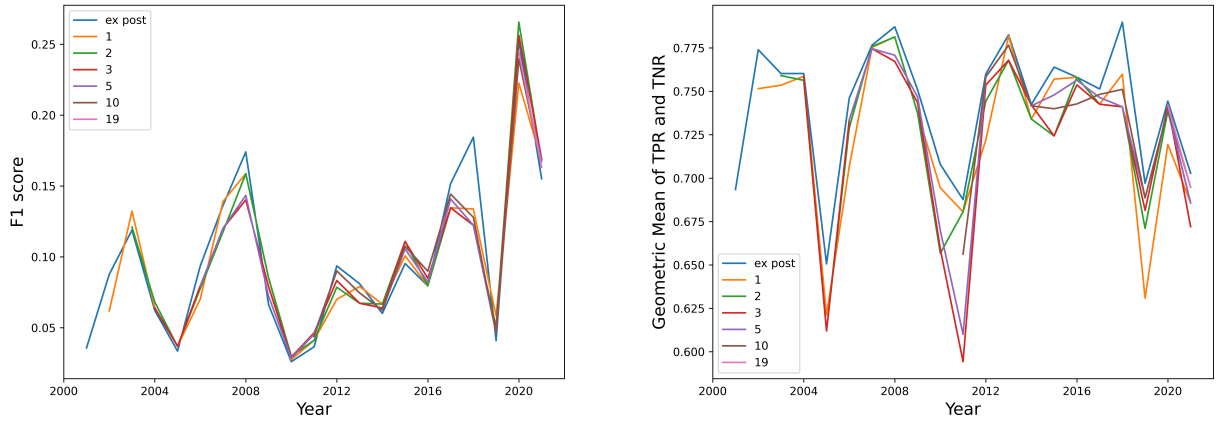
Built upon a good-performing prediction model, the next challenging step in the wildfire prediction task is to choose an appropriate decision threshold, beyond which a probability will be predicted as a positive case. For binary classification tasks with balanced data, the threshold could simply be 0.5. However, if the data is highly imbalanced, 0.5 will no longer be ideal since it would over-predict the majority class. As explained in Section 3.8, we mainly follow Leevy et al. (2023) to determine the optimal threshold by evaluating the performance of various thresholds on the training data, and choosing the one that yields the highest G-mean, the geometric mean of true positive rate (TPR) and true negative rate (TNR), while satisfying the constraint TPR $\geq$ TNR.

In fact, although Leevy et al. (2023)'s approach is validated by extremely imbalanced data, it is not designed in particular to handle non-stationary data imbalance. However, we conduct several experiments and find this approach is surprisingly stable even in the context of wildfire prediction, where the class imbalance is not only severe but also highly non-stationary. Specifically, the proportion of the positive class, i.e., burned grid cells, is time-varying, which suggests that the optimal threshold may vary over time as well. Therefore, we further extend Leevy et al. (2023)'s method by adding a time window and evaluating the threshold only on more recent data samples within the time window. However, we find that the results do not change much with different window sizes. Surprisingly, the results on G-mean and F1 score are even close enough to the ex-post optimal threshold. The extended threshold optimization method is outlined below. For window size $w$ and year $t$, we calculate

$$\theta_t^* = \underset{\theta \in (0,1)}{\operatorname{argmin}} \text{G-mean}(\theta; y_{t-w:t-1}; \hat{y}_{t-w:t-1})$$
$$\text{s.t.} \quad \text{TPR}(\theta; y_{t-w:t-1}; \hat{y}_{t-w:t-1}) \geq \text{TNR}(\theta; y_{t-w:t-1}; \hat{y}_{t-w:t-1})$$

Figure 19 shows the details of our experiment, where we sample the last day of each year between 2000 and 2020 from our wildfire prediction database, and apply different threshold optimization methods to validate the CNN-predicted next-year wildfire probabilities against the realized wildfire events. Specifically, we apply the extended Leevy et al. (2023)'s approach to each year on a rolling basis, with a bunch of window sizes ranging from 1 to 19. For example, for each year $t$, when the window size equals $w$, we will compute the optimal threshold $\theta_t^*$ on the sample between year $t-w$ and $t-1$, and then evaluate the F1 Score and G-Mean on the sample of year $t$. Note that these optimal thresholds are all calculated ex ante from historical data. As a comparison, we further calculate the ex-post optimal threshold assuming that we already know the realized wildfire events of the next year, and evaluate the

F1 score as well as G-mean. As Figure 19 shows, the F1 score and G-mean are fairly stable across different window sizes, and are even close enough to the performance of the ex-post optimal threshold. We believe the stability is mainly attributed to the constraint TPR $\geq$ TNR, which guarantees that positive minority class would never be completely ignored by the classifier in any data sample. Given these results, we optimize the threshold simply using all historical data for each year for simplicity.




(a) F1 score

(b) G-mean

Figure 19: **Classification performance for various threshold optimization methods across 21 years.** This figure presents the F1 score (on the left) and G-mean (on the right) of the optimal threshold generated from several different methods, each of which is marked by a different color. We sample the last day of each year between 2000 and 2020 for this experiment. To compare different methods, we calculate both the ex-post and several ex-ante optimal thresholds for each year on a rolling basis. The ex-post optimal threshold is calculated based on the realized wildfire of the next year, while the ex-ante optimal thresholds are calculated using the extended Leevy et al. (2023)'s method with several different window sizes.

# D Precision-Recall tradeoff

The Precision-Recall tradeoff is closely associated with threshold choice. With a low threshold, the forecast model would be more likely to make Type I errors, leading to a high Recall but low Precision. On the contrary, with a high threshold, more Type II errors would be made, leading to a high Precision but low Recall. In Figure 20, we use the validation year 2020 as an example to show how Precision and Recall respond to various thresholds in the wildfire prediction task. As we can see, Recall decreases monotonically with the threshold,

while Precision shows an inverted U shape, with the peak value only around 0.2.

The low Precision exactly reflects the challenge of wildfire prediction. That is, the dynamic local wildfire prediction task should by nature feature a large amount of false positives, i.e., the Type I errors, because the fire occurrence is extremely stochastic, even conditional on similar meteorological and vegetative conditions. In fact, the Type I errors could arise across both space and time. On the one hand, within the same year the grid cells with high fire potential may not all burn. On the other hand, for the same grid with high wildfire risk, it may not necessarily burn every year. Consequently, a dynamic predictive model that is trained to have overall good performance across years would inevitably have low Precision. In contrast, it is relatively easier to have high Recall since the model can easily identify the high risk features with many years of training data. As we have shown in Table 5, for the validation year 2020, we achieve a Precision of 0.1430 and a Recall of 0.8455. For the test year 2021, we achieve a Precision of 0.0928 and a Recall of 0.6794. In both years, Recall is much higher than Precision, which is in line with our argument.

Given this asymmetry, when varying the threshold the F1 score will closely follow the response of Precision, as Figure 19 shows. Therefore, a good-performing model must strive to improve Precision, although it is extremely difficult, as explained above. As Table 5 shows, the spatiotemporal CNN models can indeed raise Precision, and the improvement becomes larger with larger kernel sizes. This makes sense since the introduction of spatial and temporal lags allows the model to discover more rigorous high risk patterns from a larger feature space.

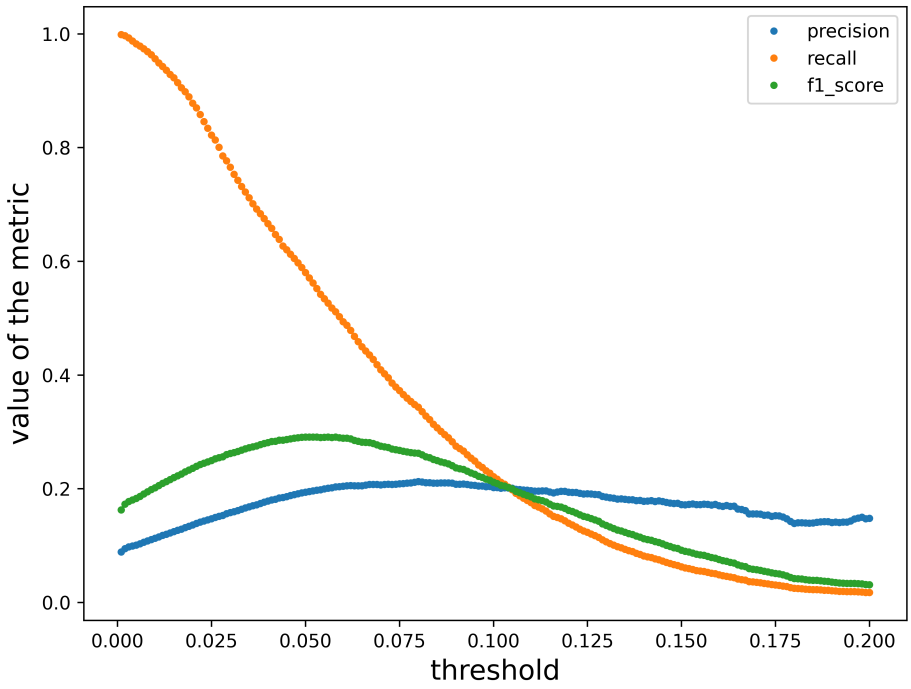



Figure 20: **The Precision-Recall tradeoff.** We use the validation set as an example to show the Precision-Recall tradeoff when picking up the optimal threshold. As this figure clearly shows, the f1 score in the wildfire prediction task depends mainly on the Precision, while Recall plays a relatively less significant role.

# E Top 120 zip codes for expected fire loss

Table 11: **Top 120 zip codes among all zip codes in California with the highest ST-CNN expected loss to the assessed value of the residential real estate structures (a key component, in addition to the assessed value of land, of the property tax basis) given the average 2020–2022 ST-CNN predicted probability of wildfire occurrence.**

| Ranking | Zip code | ST-CNN wildfire prob. 2020–2022 | Loss to residential assessment | City |
|---|---|---|---|---|
| | | % | $ | |
| 1 | 91011 | 13.59 | 26,590,949 | La Canada Flintridge |
| 2 | 93108 | 5.87 | 20,793,133 | Santa Barbara |
| 3 | 90265 | 12.15 | 19,267,521 | Malibu |
| 4 | 91302 | 9.92 | 15,065,146 | Calabasas |
| 5 | 91214 | 20.20 | 14,761,026 | La Crescenta |
| 6 | 92399 | 7.45 | 14,672,026 | Yucaipa |
| 7 | 91387 | 15.01 | 14,008,202 | Canyon Country |
| 8 | 91362 | 10.11 | 13,658,757 | Thousand Oaks |
| 9 | 92592 | 5.65 | 11,694,243 | Temecula |
| 10 | 92223 | 8.35 | 11,379,127 | Beaumont |
| 11 | 96161 | 1.92 | 10,914,561 | Truckee |
| 12 | 91390 | 18.17 | 10,321,079 | Santa Clarita |
| 13 | 92065 | 7.59 | 10,255,392 | Ramona |
| 14 | 93063 | 11.91 | 9,934,509 | Simi Valley |
| 15 | 92679 | 6.03 | 9,686,712 | Trabuco Canyon |
| 16 | 92407 | 9.35 | 8,665,163 | San Bernardino |
| 17 | 91737 | 6.09 | 8,648,978 | Rancho Cucamonga |
| 18 | 91381 | 19.24 | 8,580,735 | Stevenson Ranch |
| 19 | 91042 | 14.92 | 8,451,197 | Tujunga |
| 20 | 91901 | 8.26 | 8,387,484 | Alpine |
| 21 | 91361 | 13.39 | 8,302,593 | Westlake Village |
| 22 | 93065 | 9.92 | 8,237,446 | Simi Valley |
| 23 | 90272 | 4.24 | 7,668,455 | Pacific Palisades |
| 24 | 91001 | 9.54 | 7,612,629 | Altadena |
| 25 | 91320 | 10.86 | 7,372,269 | Newbury Park |
| 26 | 92562 | 5.14 | 7,356,871 | Murrieta |
| 27 | 93551 | 10.00 | 7,247,473 | Palmdale |
| 28 | 91739 | 3.92 | 7,099,801 | Rancho Cucamonga |

Continued on next page

**Table 11 — continued from previous page**

| Ranking | Zip code | ST-CNN wildfire prob. 2020–2022 | Loss to residential assessment | City |
|---|---|---|---|---|
| | | % | $ | |
| 29 | 92352 | 3.14 | 7,061,444 | Lake Arrowhead |
| 30 | 94539 | 2.77 | 7,052,555 | Fremont |
| 31 | 91784 | 4.90 | 6,965,130 | Upland |
| 32 | 91321 | 25.54 | 6,963,251 | Newhall |
| 33 | 92886 | 5.03 | 6,779,260 | Yorba Linda |
| 34 | 93561 | 5.95 | 6,699,938 | Tehachapi |
| 35 | 92346 | 6.08 | 6,277,609 | Highland |
| 36 | 91307 | 17.03 | 6,118,837 | West Hills |
| 37 | 92883 | 7.45 | 6,080,526 | Corona |
| 38 | 91377 | 16.21 | 5,883,997 | Oak Park |
| 39 | 94506 | 2.84 | 5,811,607 | Danville |
| 40 | 91342 | 12.40 | 5,774,785 | Sylmar |
| 41 | 95949 | 3.35 | 5,763,205 | Grass Valley |
| 42 | 93023 | 8.33 | 5,740,550 | Ojai |
| 43 | 91301 | 12.35 | 5,626,243 | Agoura Hills |
| 44 | 95747 | 5.54 | 5,556,152 | Roseville |
| 45 | 93021 | 12.66 | 5,529,633 | Moorpark |
| 46 | 95223 | 3.45 | 5,464,392 | Arnold |
| 47 | 92315 | 3.77 | 5,326,834 | Big Bear Lake |
| 48 | 91208 | 11.44 | 4,929,039 | Glendale |
| 49 | 91701 | 3.77 | 4,920,157 | Rancho Cucamonga |
| 50 | 91741 | 6.79 | 4,918,636 | Glendora |
| 51 | 96150 | 1.42 | 4,892,643 | South Lake Tahoe |
| 52 | 91367 | 4.10 | 4,843,695 | Woodland Hills |
| 53 | 92887 | 7.00 | 4,667,842 | Yorba Linda |
| 54 | 91040 | 14.52 | 4,644,896 | Sunland |
| 55 | 92028 | 4.79 | 4,596,697 | Fallbrook |
| 56 | 91384 | 16.99 | 4,587,457 | Castaic |
| 57 | 95409 | 5.86 | 4,487,038 | Santa Rosa |
| 58 | 93105 | 7.66 | 4,477,009 | Santa Barbara |
| 59 | 91711 | 4.47 | 4,404,047 | Claremont |
| 60 | 95762 | 1.62 | 4,292,404 | El Dorado Hills |
| 61 | 92220 | 9.00 | 4,285,629 | Banning |
| 62 | 95945 | 4.97 | 4,247,209 | Grass Valley |
| 63 | 92064 | 3.65 | 4,010,168 | Poway |

Continued on next page

**Table 11 — continued from previous page**

| Ranking | Zip code | ST-CNN wildfire prob. 2020–2022 | Loss to residential assessment | City |
|---|---|---|---|---|
| | | % | $ | |
| 64 | 91709 | 4.93 | 3,938,504 | Chino Hills |
| 65 | 91311 | 10.49 | 3,909,324 | Chatsworth |
| 66 | 92557 | 4.97 | 3,782,642 | Moreno Valley |
| 67 | 93510 | 10.75 | 3,685,512 | Acton |
| 68 | 95370 | 4.03 | 3,630,736 | Sonora |
| 69 | 92336 | 2.67 | 3,460,268 | Fontana |
| 70 | 92404 | 9.31 | 3,443,585 | San Bernardino |
| 71 | 91360 | 3.92 | 3,438,501 | Thousand Oaks |
| 72 | 92382 | 8.48 | 3,437,465 | Running Springs |
| 73 | 92563 | 2.41 | 3,387,926 | Murrieta |
| 74 | 95667 | 3.31 | 3,343,415 | Placerville |
| 75 | 92651 | 3.74 | 3,333,477 | Laguna Beach |
| 76 | 91750 | 5.23 | 3,317,619 | La Verne |
| 77 | 92595 | 4.36 | 3,306,150 | Wildomar |
| 78 | 92019 | 5.99 | 3,271,905 | El Cajon |
| 79 | 92082 | 5.51 | 3,250,634 | Valley Center |
| 80 | 95404 | 4.08 | 3,242,233 | Santa Rosa |
| 81 | 92325 | 7.14 | 3,188,849 | Crestline |
| 82 | 95247 | 8.62 | 3,153,403 | Murphys |
| 83 | 91935 | 6.85 | 3,075,070 | Jamul |
| 84 | 92657 | 2.28 | 3,010,803 | Newport Coast |
| 85 | 92021 | 5.80 | 3,004,867 | El Cajon |
| 86 | 96146 | 4.43 | 2,919,875 | Olympic Valley |
| 87 | 92373 | 6.66 | 2,912,513 | Redlands |
| 88 | 95959 | 3.14 | 2,885,350 | Nevada City |
| 89 | 92530 | 5.52 | 2,867,598 | Lake Elsinore |
| 90 | 95713 | 5.61 | 2,859,012 | Colfax |
| 91 | 93103 | 3.91 | 2,698,918 | Santa Barbara |
| 92 | 91107 | 4.15 | 2,696,046 | Pasadena |
| 93 | 94582 | 1.48 | 2,691,892 | San Ramon |
| 94 | 94563 | 1.35 | 2,664,708 | Orinda |
| 95 | 94941 | 1.98 | 2,595,985 | Mill Valley |
| 96 | 94028 | 2.36 | 2,566,909 | Portola Valley |
| 97 | 92882 | 5.75 | 2,556,543 | Corona |
| 98 | 90049 | 1.40 | 2,505,946 | Los Angeles |

Continued on next page

**Table 11 — continued from previous page**

| Ranking | Zip code | ST-CNN wildfire prob. 2020–2022 | Loss to residential assessment | City |
|---|---|---|---|---|
| | | % | $ | |
| 99 | 91765 | 2.46 | 2,473,866 | Diamond Bar |
| 100 | 91304 | 5.11 | 2,442,509 | Canoga Park |
| 101 | 94549 | 1.33 | 2,387,995 | Lafayette |
| 102 | 91501 | 10.66 | 2,378,141 | Burbank |
| 103 | 94062 | 1.67 | 2,370,476 | Redwood City |
| 104 | 93111 | 5.24 | 2,363,196 | Santa Barbara |
| 105 | 92026 | 2.76 | 2,346,498 | Escondido |
| 106 | 92591 | 2.86 | 2,346,065 | Temecula |
| 107 | 90290 | 5.79 | 2,325,667 | Topanga |
| 108 | 92532 | 5.58 | 2,301,132 | Lake Elsinore |
| 109 | 92555 | 4.80 | 2,275,147 | Moreno Valley |
| 110 | 94507 | 1.30 | 2,221,651 | Alamo |
| 111 | 93536 | 5.46 | 2,212,870 | Lancaster |
| 112 | 95476 | 1.43 | 2,205,572 | Sonoma |
| 113 | 91006 | 1.62 | 2,161,517 | Arcadia |
| 114 | 92673 | 1.95 | 2,155,128 | San Clemente |
| 115 | 93110 | 3.40 | 2,003,821 | Santa Barbara |
| 116 | 93422 | 2.19 | 1,994,986 | Atascadero |
| 117 | 93117 | 5.26 | 1,992,093 | Goleta |
| 118 | 94566 | 1.13 | 1,988,225 | Pleasanton |
| 119 | 95037 | 2.15 | 1,979,770 | Morgan Hill |
| 120 | 91207 | 9.72 | 1,974,397 | Glendale |

# References


Abatzoglou, John T., 2013, Development of gridded surface meteorological data for ecological applications and modelling, *International Journal of Climatology* 33, 121–131.

Abatzoglou, John T., Jennifer K. Balch, Bethany A. Bradley, and Crystal A. Kolden, 2018, Human-related ignitions concurrent with high winds promote large wildfires across the USA, *International Journal of Wildland Fire* 27, 277–386.

Abatzoglou, John T., Benjamin J. Hatchett, Paul Fox-Hughes, Alexander Gershunov, and Nicholas J. Nausla, 2021, Global climatology of synoptically-forced downslope winds, *International Journal of Wildland Fire* 41, 31–50.

Abatzoglou, John T., and A. Park Williams, 2016, Impact of anthropogenic climate change on wildfire across western US forests, *Proceedings of the National Academy of Sciences* 113, 505–546.

Akinjole, Abisola, Olamilekan Shobayo, Jumoke Popoola, Obinna Okoyeigbo, and Bayode Ogunleye, 2024, Ensemble-based Machine Learning algorithm for loan default risk prediction, *Mathematics* 12, 3423.

Alexandre, Patricia M., Susan I. Stewart, Miranda H. Mockrin, Nicholas S. Keuler, Alexandra D. Syphard, Avi Bar-Massada, Murray K. Clayton, and Volker C. Radeloff, 2016, The relative impacts of vegetation, topography and spatial arrangement on building loss to wildfires in case studies of California and Colorado, *Landscape Ecology* 31, 415–430.

Alkhatib, Ramez, Wahib Sahwan, Anas Alkhatieb, and Brigitta Schütt, 2023, A brief review of machine learning algorithms in forest fires science, *Applied Sciences* 13, 2–15.

Apt, Jerome, Dennis Epple, and Fallaw Sowell, 2023, Forest fires: Why the large year-to-year variation in forests burned?, Working Paper 31738, NBER.

Balch, Jennifer K., Bethany A. Bradley, John T. Abatzoglou, R. Chelsea Nagy, Emily J. Fusco, and Adam L. Mahood, 2017, Human-started wildfires expand the fire niche across the United States, *Proceedings of the National Academy of Science* 114, 2946–2951.

Batista, Gustavo E. A. P. A., Ronaldo C. Prati, and Maria Carolina Monard, 2004, A study of the behavior of several methods for balancing machine learning training data, *ACM SIGKDD Explorations Newsletter* 6, 20–29.

Baylis, Patrick W., and Judson Boomhower, 2021, Mandated vs. voluntary adaptation to natural disasters: The case of U.S. wildfires, Working Paper 29621, NBER.

Billmire, Michael, Nancy H. F. French, Tatiana Loboda, R. Chris Owen, and Marlene Tyner, 2014, Santa Ana winds and predictors of wildfire progression in Southern California, *International Journal of Wildland Fire* 23, 1119–1129.

Biswas, Siddhartha, Mallick Hossain, and David Zink, 2023, California wildfires, property damage, and mortgage repayment, Working Paper 23-05, Federal Reserve Bank of Philadelphia.

Boomhower, Judson, Meredith Fowlie, Jacob Gellman, and Andrew Plantinga, 2024, How are insurance markets adapting to climate change? Risk selection and regulation in the market for homeowners insurance, Working Paper 32625, NBER.

Bowers, Carrie Lynn, 2018, *The Diablo Winds of Northern California: Climatology and Numerical Simulations*, Master's thesis, San Jose State University.

Brey, Steven J., Elizabeth A. Barnes, Jeffrey R. Pierce, Christine Wiedinmyer, and Emily V. Fischer, 2018, Environmental conditions, ignition type, and air quality impacts of wildfires in the southeastern and western United States, *Earth's Future* 6, 1442–1456.

Brinkmann, Peggy, Nancy Watkins, Cody Webb, Dave Evans, Gabriele Usan, Michael Glavan, Lillian Zhang, Carolyn Prescott, Tom Larsen, and Grace Lee, 2022, Catastrophe models for wildfire mitigation: Quantifying credits and benefits to homeowners and communities, Research Paper, Casualty Actuarial Society.

Brooks, Matthew L., and John R. Matchett, 2006, Spatial and temporal patterns of wildfires in the Mojave Desert, 1980–2004, *Journal of Arid Environments* 67, 148–164.

Buda, Mateusz, Atsuto Maki, and Maciej A. Mazurowski, 2018, A systematic study of the class imbalance problem in convolutional neural networks, *Neural Networks* 106, 249–259.

Buechi, Hanna, Paige Weber, Sarah Heard, Dick Cameron, and Andrew J. Plantinga, 2021, Long-term trends in wildfire damages in California, *International Journal of Wildland Fire* 30, 757–762.

Bulatov, Dimitri, and Felix Leidinger, 2021, Instance segmentation of deadwood objects in combined optical and elevation data using convolutional neural networks, in *Earth Resources and Environmental Remote Sensing/GIS Applications XII*, volume 11863, 299–308 (SPIE).

Burke, Marshall, Anne Discoll, Sam Heft-Neal, Jiani Xue, Jennifer Burney, and Michael Wara, 2021, The changing risk and burden of wildfire in the United States, *PNAS* 118, 1–6.

Calhoun, Kendall L., Melissa Chapman, Carmen Tubbesing, Alex McInturff, Kaitlyn M. Gaynor, Amy Van Scoyoc, Christine E. Wilkinson, Phoebe Parker-Shames, David Kurz, and Justin Brashares, 2022, Spatial overlap of wildfire and biodiversity in California highlights gap in non-conifer fire research and management, *Diversity and Distributions* 28, 529–541.

Cardil, Adrián, Marcos Rodrigues, Joaquin Ramierz, Sergio de-Miguel, Carles A. Silva, Machela Mariani, and Davide Ascoli, 2021, Coupled effects of climate teleconnections on drought, Santa Ana winds and wildfires in southern California, *Science of the Total Environment* 756, 1–8.

Casolaro, Angelo, Vincenzo Capone, Gennaro Iannuzzo, and Francesco Camastra, 2023, Deep learning for time series forecasting: Advances and open problems, *Information* 14, 598.

Chawla, Nitesh V., Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer, 2002, SMOTE: Synthetic minority over-sampling technique, *Journal of Artificial Intelligence Research* 16, 321–357.

Chegini, Taher, Hong-Yi Li, and L. Ruby Leung, 2021, HyRiver: Hydroclimate data retriever, *Journal of Open Source Software* 6, 1–3.

Chen, Bin, and Yufang Jin, 2022, Spatial patterns and drivers for wildfire ignitions in California, *Environmental Research Letters* 17, 055004.

Chen, Bin, Yufang Jin, Erica Scaduto, Max A. Moritz, Michael L. Goulden, and James T. Randerson, 2021, Climate, fuel, and land use shaped the spatial pattern of wildfire in California's Sierra Nevada, *Journal of Geophysical Research: Biogeosciences* 126, 1–18.

Chen, Liuyi, Bocheng Han, Xuesong Wang, Jiazhen Zhao, Wenke Yang, and Zhengyi Yang, 2023, Machine learning methods in weather and climate applications: A survey, *Applied Sciences* 13, 12019.

Cleveland, William S., 1979, Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association* 74, 829–836.

Cooke, Roger M., Daan Nieboer, and Jolanta Misiewicz, 2014, *Fat-Tailed Distributions: Data, Diagnostics and Dependence*, volume 1 (John Wiley & Sons).

Cruciata, Giorgio, Liliana Lo Presti, Gabriele Ajello, Paolo Cicero, Giacomo Corvisieri, and Marco La Cascia, 2024, Wildfires classification: A comparative study, in *Image*

*Analysis and Processing — ICIAP 2023 Workshops: Udine, Italy, September 11–15, 2023, Proceedings, Part I*, 62–73 (Springer-Verlag, Berlin, Heidelberg).

Dennison, Philip E., and Max A. Moritz, 2009, Critical live fuel moisture in chaparral ecosystems: A threshold for fire activity and its relationship to antecedent precipitation, *International Journal of Wildland Fire* 18, 1021–1027.

Dennison, Philip E., Max A. Moritz, and Robert S. Taylor, 2008, Evaluating predictive models of critical live fuel moisture in the Santa Monica mountains, California, *International Journal of Wildland Fire* 17, 18–27.

Diaz, Adam, 2022, *A Contribution to the Statistical Analysis of Climate-Wildfire Interaction in Northern California*, Ph.D. thesis, Clemson University.

DiMiceli, Charlene, John Townshend, Mark Carroll, and Robert Sohlberg, 2021, Evolution of the representation of global vegetation by vegetation continuous fields, *Remote Sensing of Environment* 254, 112271.

Dixon, Dan J., Yunzhe Zhu, Christopher F. Brown, and Yufang Jin, 2023, Satellite detection of canopy-scale tree mortality and survival from California wildfires with spatio-temporal deep learning, *Remote Sensing of Environment* 298, 113842.

Dube, Lindani, and Tanja De la Rey Verster, 2024, Assessing the performance of machine learning models for default prediction under missing data and class imbalance: A simulation study, *ORION* 40, 1–24.

Duclos, Philippe, Lee M. Sanderson, and Michael Lipsett, 1990, The 1987 forest fire disaster in California: Assessment of emergency room visits, *Archives of Environmental Health: An International Journal* 45, 53–58.

Esposito, Carmen, Gregory A. Landrum, Nadine Schneider, Nikolaus Stiefl, and Sereina Riniker, 2021, GHOST: Adjusting the decision threshold to handle imbalanced data in machine learning, *Journal of Chemical Information and Modeling* 61, 2623–2640.

First Street Foundation, 2022, The First Street Foundation wildfire model.

Flannigan, M. D., B. M. Wotton, G. A. Marshall, W. J. DeGroot, J. Johnston, N. Jurko, and A. S. Cantin, 2016, Fuel moisture sensitivity to temperature and precipitation: Climate change implications, *Climatic Change* 134, 59–71.

Flannigan, Mike D., Meg A. Krawchuk, William J. de Groot, B. Mike Wotton, and Lynn M. Gowman, 2009, Implications of changing climate for global wildland fire, *International Journal of Wildland Fire* 18, 483–507.

Georgiev, Georgi, G.V. Hristov, Plamen Zahariev, and Diyana Kyuchukova, 2020, Forest monitoring system for early fire detection based on convolutional neural network and UAV imagery, in *2020 28th National Conference with International Participation*.

Gershunov, Alexander, Janin Guzman Morales, Benjamin Hatchett, Kristen Guirguis, Rosana Aguilera, Tamara Shulgina, John T. Abatzoglou, Daniel Cayan, David Pierce, Park Williams, Ivory Small, Rachel Clemesha, Lara Schwarz, Tarik Benmarhnia, and Alex Tardy, 2021, Hot and cold flavors of Southern California's Santa Ana winds: Their causes, trends, and links with wildfire, *Climate Dynamics* 57, 2233–2248.

Goessling, Helge F., Thomas Rackow, and Thomas Jung, 2025, Recent global temperature surge intensified by record-low planetary albedo, *Science* 387, 68–73.

Goss, Michael, Daniel L. Swain, John T. Abatzoglou, Ali Sarhadi, Crystal A. Kolden, A. Park Williams, and Noah S. Diffenbaugh, 2020, Climate change is increasing the likelihood of extreme autumn wildfire conditions across California, *Environmental Research Letters* 15, 505–546.

Guo, Shengnan, Youfang Lin, Shijie Li, Zhaoming Chen, and Huaiyu Wan, 2019, Deep spatial-temporal 3D convolutional neural networks for traffic data forecasting, *IEEE Transactions on Intelligent Transportation Systems* 20, 3913–3926.

Guzman-Morales, Janin, 2018, *Santa Ana Winds of Southern California: Historical Variability and Future Climate Projections*, Ph.D. thesis, University of California San Diego.

Hodnebrog, Øivind, Gunnar Myhre, Caroline Jouan, Timothy Andrews, Piers M. Forster, Hailing Jia, Norman G. Loeb, Dirk J. L. Olivié, David Paynter, Johannes Quaas, Shiv Priyam Raghuraman, and Michael Schulz, 2024, Recent reductions in aerosol emissions have increased Earth's energy imbalance, *Communications Earth & Environment* 5, 1–9.

Holmes, Thomas P., Jr. Huggett, Robert J., and Anthony L. Westerling, 2008, Statistical analysis of large wildfires, in T. P. Holmes, ed., *The Economics of Forest Disturbances: Wildfires, Storms, and Invasive Species*, 59–77 (Springer Science).

Ismail, Fathima Nuzla, and Shanika Amarasoma, 2023, One-class classification-based machine learning model for estimating the probability of wildfire risk, *Procedia Computer Science* 222, 341–352.

Jaffee, Dwight, and Thomas Russell, 2013, Catastrophe insurance, capital markets, and uninsurable risks, *Journal of Risk and Insurance* 64, 205–230.

Jain, Piyush, Sean C. P. Coogan, Sriram Ganapathi Subramanian, Mark Crowley, Steve Taylor, and Mike D. Flannigan, 2020, A review of machine learning applications in wildfire science, *Environmental Review* 28, 478–505.

Jergler, Don, 2021, Grim California wildfire outlook has insurers forking over big bucks for modeling, *The Insurance Journal* June 18.

Jiao, Zhentian, Youmin Zhang, Lingxia Mu, Jing Xin, Shangbin Jiao, Han Liu, and Ding Liu, 2020, A YOLOv3-based learning strategy for real-time UAV-based forest fire detection, in *2020 Chinese Control And Decision Conference (CCDC)*, 4963–4967.

Jiao, Zhentian, Youmin Zhang, Jing Xin, Lingxia Mu, Yingmin Yi, Han Liu, and Ding Liu, 2019, A deep learning based forest fire detection approach using UAV and YOLOv3, *1st International Conference on Industrial Artificial Intelligence (IAI)* 1–5.

Jin, Yufang, James T. Randerson, Nicolas Faivre, Scott Capps, Alex Hall, and Michael L. Goulden, 2013, Contrasting controls on wildland fires in Southern California during periods with and without Santa Ana winds, *Journal of Geophysical Research: Biogeosciences* 119, 432–450.

Joseph, Maxwell B., Matthew W. Rossi, Nathan P. Mietkiewicz, Adam L. Mahood, Megan E. Cattau, Lise Ann St. Denis, R. Chelsea Nagy, Virinia Iglesias, John T. Abatzoglou, and Jennifer K. Balch, 2019, Spatiotemporal prediction of wildfire size extremes with Bayesian finite sample maxima, *Ecological Applications* 29, 1266–1281.

Kahn, Mathew, Amine Ouazad, and Erkan Yönder, 2024, Adaptation using financial markets: Climate risk diversification through securitization, Working Paper 32244, NBER.

Kalashnikov, Dmitri A., John T. Abatzoblou, Nicholas J. Nauslar, Daniel L. Swain, Danielle Touma, and Deepti Singh, 2022, Meteorological and geographical factors associated with dry lightning in central and northern California, *Environmental Research: Climate* 1, 025001.

Kashinath, K., M. Mustafa, A. Albert, J-L. Wu, C. Jiang, S. Esmaeilzadeh, K. Asissadenesheli, R. Wang, A. Chattopadhyay, A. Singh, A. Manepalli, D. Chirila, R. Yu, R. Walters, B. White, H. Xiao, H. A. Tchelepi, P. Marcus, A. Anandkumar, P. Hassanzadeh, and Prabhat, 2020, Physics-informed machine learning: Case studies for weather and climate modelling, *Philosophical Transactions A, Royal Society* 379, 1–34.

Kearns, Edward J., David Saah, Carrie R. Levine, Chris Lautenberger, Owen M. Doherty, Jeremy R. Porter, Michael Amodeo, Carl Rudeen, Kyle D. Woodward, Gary W. Johnson, Kel Markert, Evelyn Shu, Neil Freeman, Mark Bauer, Kelvin Lai, Ho Hsieh, Bradley Wilson, Beth McClenny, Andrea McMahon, and Farrukh Chishtie, 2022, The construction of probabilistic wildfire risk estimates for individual real estate parcels for the contiguous United States, *Fire* 5, 377–386.

Keeley, Jon E., Michael Flannigan, Tim J. Brown, Tom Rolinski, Daniel Cayan, Alexandra D. Syphard, Janin Guzman-Morales, and Alexander Gershunov, 2024, Climate and weather drivers in Southern California Santa Ana wind and non-Santa Ana wind fires, *International Journal of Wildland Fire* 33, 1–16.

Keeley, Jon E., Janin Guzman-Morales, Alexander Gershunov, Alexandra D. Syphard, Daniel Cayan, David W. Pierce, Michael Flannigan, and Tim J. Brown, 2021, Ignitions explain more than temperature or precipitation in driving Santa Ana wind fires, *Science Advances* 7, 1–9.

Keeley, Jon E., and Alexandra D. Syphard, 2018, Historical patterns of wildfire ignition sources in California ecosystems, *International Journal of Wildland Fire* 27, 781–799.

Keeley, Jon E., and Alexandra D. Syphard, 2021, Large California wildfires: 2020 fires in historical context, *Fire Ecology* .

Kestelman, Stephanie, 2024, Environmental externalities of urban growth: Evidence from the California wildfires, Working paper, Harvard University.

Kochanski, Adam K., Mary Ann Jenkins, Jan Mandel, Jonathan D. Beezley, and Steven K. Krueger, 2013, Real time simulation of 2007 Santa Ana fires, *Forest Ecology and Management* 294, 136–149.

Koh, Jonathan, François Pimont, Jean-Luc Dupuy, and Thomas Opitz, 2023, Sptatiotemporal wildfire modeling through point processes with moderate and extreme marks, *Annals of Applied Statistics* 17, 560–582.

Kousky, Carolyn, 2019, The role of natural disaster insurance in recovery and risk reduction, *Annual Review of Resource Economics* 11, 399–418.

Kousky, Carolyn, and Roger M. Cooke, 2009, Climate change and risk management, Technical Report RFF DP 0903-REV, Resources for the Future.

Kumar, Lalit, Andrew K. Skidmore, and Edmund Knowles, 1997, Modelling topographic variation in solar radiation in a GIS environment, *International Journal of Geographical Information Science* 11, 475–497.

Lai, Gengke, Xingwen Quan, Marta Yebra, and Binbin He, 2022, Model-driven estimation of closed and open shrublands live fuel moisture content, *GIScience & Remote Sensing* 59, 1837–1856.

Leevy, Joffrey L., Justin M. Johnson, John Hancock, and Taghi M. Khoshgoftaar, 2023, Threshold optimization and random undersampling for imbalanced credit card data, *Journal of Big Data* 10, 58.

Li, Shu, and Tirtha Banerjee, 2021, Spatial and temporal pattern of wildfires in California from 2000 to 2019, *Scientific Reports* 11, 8779.

Linn, Rodman, Judith Winterkamp, Carleton Edminster, Jonah J. Colman, and William S. Smith, 2007, Coupled influences of topography and wind on wildland fire behaviour, *International Journal of Wildland Fire* 16, 183–195.

Liu, Lu, 2022, The demand for long-term mortgage contracts and the role of collateral, Working Paper, Wharton.

Liu, Yi-Chin, Pingkuan Di, Shu-Hua Chen, ZueMeng Chen, Jiwen Fan, John DaMassa, and Jeremy Avise, 2021, Climatology of Diablo winds in Northern California and their relationships with large-scale climate variables, *Climate Dynamics* 56, 1335–1356.

Lundberg, Scott M., and Su-In Lee, 2017, A unified approach to interpreting model predictions, in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, 4768–4777 (Curran Associates Inc., Red Hook, NY, USA).

MacDonald, Glen, Tamara Wall, Carolyn A. F. Enquist, Sarah R. LeRoy, John B. Bradford, David D. Breshears, Timothy Brown, Daniel Cayan, Chunyu Dong, Donald A. Falk, Erica Fleishman, Alexander Gershunov, Molly Hunter, Rachel A. Loehman, Phillip J. van Mantgem, Beth Rose Middleton, Hugh D. Safford, Mark W. Schwartz, and Valerie

Trouet, 2023, Drivers of California's changing wildfires: A state-of-the-knowledge synthesis, *International Journal of Wildland Fire* 32, 1039–1058.

Makridakis, Spyros, Evangelos Spiliotis, Vassilios Assimakopoulos, Artemios-Anargyros Semenoglou, Gary Mulder, and Konstantinos Nikolopoulos, 2023, Statistical, machine learning and deep learning forecasting methods: Comparisons and ways forward, *Journal of the Operational Research Society* 74, 840–859.

Mallinis, Giorgos, Marius Petrila, Ioannis Mitsopoulos, Adrien Lorent, Stefan Neagu, Bogdan Apostol, Vladimir Gancz, Popa Ionel, and Johann Georg Goldammer, 2019, Geospatial patterns and drivers of forest fire occurrence in Romania, *Applied Spatial Analysis and Policy* 12, 773–795.

McClung, Brandon, and Clifford F. Mass, 2020, The strong, dry winds of central and northern California: Climatology and synoptic evolution, *Weather and Forecasting* 316, 2163–2178.

Miller, J. D., C. N. Skinner, H. D. Safford, E. E. Knapp, and C. M. Ramirez, 2012, Trends and causes of severity, size, and number of fires in northwestern California, USA, *Ecological Applications* 22, 184–203.

Oh, Sangmin, Ishita Sen, and Ana-Maria Tenekedjieva, 2024, Pricing of climate risk insurance: Regulation and cross-subsidies, Working Paper, Columbia Business School.

Oliveira, Sandra, Jorge Rocha, and Ana Sá, 2021, Wildfire risk modeling, *Current Opinion in Environmental Science & Health* 23, 1–6.

Opitz, Thomas, 2023, Editorial: EVA 2021 data challenge on spatiotemporal prediction of wildfire extremes in the USA, *Extremes* 26, 241–250.

Paci, James, Matthew Newman, and Tim Gage, 2023, The economic, fiscal, and environmental costs of wildfire in California, Technical Report, Gordon and Betty Moore Foundation.

Park, Minsoo, Dai Quoc Tran, Daekyo Jung, and Seunghee Park, 2020, Wildfire-detection method using DenseNet and CycleGAN data augmentation-based remote camera imagery, *Remote Sensing* 12, 3715.

Prestemon, Jeffrey P., Todd J. Hawbaker, Michael Bowden, John Carpenter, Maureen T. Brooks, Karen L. Abt, Ronda Sutphen, and Samuel Scranton, 2013, Wildfire ignitions: A review of the science and recommendations for empirical modeling, General Technical

Report SRS-171, Southern Research Station, Forest Service, United States Department of Agriculture.

Price, Owen, and Ross Bradstock, 2014, Countervailing effects of urbanization and vegetation extent on fire frequency on the wildland urban interface: Disentangling fuel and ignition effects, *Landscape and Urban Planning* 130, 81–88.

Radeloff, Volker C., David P. Helmers, H. Anu Kramer, Miranda H. Mockrin, Patricia M. Alexandre, Avi Bar-Massada, Van Butsic, Todd J. Hawbaker, Sebastián Martinuzzi, Alexandra D. Syphard, and Susan I. Stewart, 2018, Rapid growth of the US wildland-urban interface raises wildfire risk, *Proceedings of the National Academy of Sciences* 115, 3314–3319.

Safford, Hugh D., Alison K. Paulson, Zachary L. Steel, Derek J. N. Young, and Rebecca B. Wayman, 2022, The 2020 California fire season: A year like no other, a return to the past, or a harbinger of the future?, *Global Ecology and Biogeography* 31, 2005–2025.

Seydi, Seyd Teymoor, John T. Abatzoglou, Amir AghaKouchak, Yavar Pourmohamad, Ashok Mishra, and Mojtaba Sadegh, 2024, Predictive understanding of links between vegetative soil burn severities using physics-informed machine learning, *Earth's Future* 12, e2024EF004873.

Shamsoshoara, Alireza, Fatemeh Afghah, Abolfazl Razi, Liming Zheng, Peter Z. Fulé, and Erik Blasch, 2021, Aerial imagery pile burn detection using deep learning: The FLAME dataset, *Computer Networks* 193, 108001.

Sra, Survit, 2019, CNN, Lecture note, 6.687 Fall 2019, Massachusetts Institute of Technology.

Stein, Jeremy C., 1988, Takeover threats and managerial myopia, *Journal of Political Economy* 96, 61–80.

Stein, Jeremy C., 1989, Efficient capital markets, inefficient firms: A model of myopic corporate behavior, *Quarterly Journal of Economics* 104, 655–669.

Thode, Henry C., 2002, *Testing for Normality* (CRC Press, Boca Raton).

Tong, Qi, and Thomas Gernay, 2023, Mapping wildfire ignition probability and predictor sensitivity with ensemble-based machine learning, *Natural Hazards* 119, 1551–1582.

Tran, Du, Heng Wang, Lorenzo Torresani, Jamie Ray, Yann LeCun, and Manohar Paluri, 2018, A closer look at spatiotemporal convolutions for action recognition, in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6450–6459.

Voosen, Pau, 2024, El Niño fingered as likely culprit in record 2023 temperatures, *Science* 386, 137.

Wallmann, James, Rhett Milne, Christopher Smallcomb, and Matthew Mehle, 2010, Using the 21 June 2008 California lightning outbreak to improve dry lightning forecast procedures, *Weather and Forecasting* 25, 1447–1462.

Wang, Daoping, Dabo Guan, Shupeng Zhu, Michael Mac Kinnon, Guannan Geng, Qiang Zhang, Heran Zheng, Tianyang Lei, Shuai Shao, Peng Gong, and Steven J. Davis, 2021a, Economic footprint of California wildfires in 2018, *Nature Sustainability* 4, 252–260.

Wang, Shengying, Jing Zhao, Na Ta, Xiaoye Zhao, Mingxia Xiao, and Haicheng Wei, 2021b, A real-time deep learning forest fire monitoring algorithm based on an improved Pruned + KD model, *Journal of Real-Time Image Processing* 18, 2319–2329.

Westerling, Jaap F., 2014, "The global multi-asset market portfolio, 1959–2012": A comment, *Financial Analysts Journal* 70, 9.

Williams, A. Park, and John T. Abatzoglou, 2020, Warmer and drier fire seasons contribute to increases in area burned at high severity in Western US forests from 1985 to 2017, *Geophysical Research Letters* 47, e2020GL089858.

Williams, A. Park, and John T. Abatzoglu, 2016, Recent advances and remaining uncertainties in resolving past and future effects on global fire activity, *Current Climate Change Reports* 2, 1–14.

Xi, Dexen D. Z., Stephen W. Taylor, Douglas G. Woolford, and C. B. Dean, 2019, Statistical models of key components of wildfire risk, *Annual Review of Statistics and Its Applications* 6, 197–222.

Yebra, Marta, Xingwen Quan, David Riaño, Pablo Rozas Larraondo, Albert I. J. M. van Dijk, and Geoffrey J. Cary, 2018, A fuel moisture content and flammability monitoring methodology for continental Australia based on optical remote sensing, *Remote Sensing of Environment* 212, 260–272.

Zhang, Guoli, Ming Wang, and Kai Liu, 2019, Forest fire susceptibility modeling using a convolutional neural network for Yunnan Province of China, *International Journal of Disaster Risk Science* 10, 386–403.

Zhao, Yi, Jiale Ma, Xiaohui Li, and Jie Zhang, 2018, Saliency detection and deep learning-based wildfire identification in UAV imagery, *Sensors* 18, 712.

Zheng, Jiewan, Xianbin Cao, Baochang Zhang, Yuanjun Huang, and Yutao Hu, 2017, Bi-heterogeneous Convolutional Neural Network for UAV-based dynamic scene classification, *2017 Integrated Communications, Navigation and Surveillance Conference (ICNS)* 5B4–1–5B4–12.