ENVIRONMENTAL RESEARCH
LETTERS

LETTER • OPEN ACCESS

Winter and spring climate explains a large portion of interannual variability and trend in western U.S. summer fire burned area

ENVIRONMENTAL RESEARCH
LETTERS

LETTER

Winter and spring climate explains a large portion of interannual variability and trend in western U.S. summer fire burned area

Ronnie Abolafia-Rosenzweig*, Cenlin He and Fei Chen

National Center for Atmospheric Research, Boulder, CO, United States of America
* Author to whom any correspondence should be addressed.

E-mail: abolafia@ucar.edu

Keywords: wildfire, climate, statistical modeling, remote sensing, snow drought, burned area, fire forecasting

Supplementary material for this article is available online

OPEN ACCESS

RECEIVED
22 November 2021

REVISED
5 April 2022

ACCEPTED FOR PUBLICATION
20 April 2022

PUBLISHED
29 April 2022

Abstract

This study predicts summer (June–September) fire burned area across the western United States (U.S.) from 1984 to 2020 using ensembles of statistical models trained with pre-fire season climate conditions. Winter and spring climate conditions alone explain up to 53% of the interannual variability and 58% of the increasing trend of observed summer burned area, which suggests that climate conditions in antecedent seasons have been an important driver to broad-scale changes in summer fire activity in the western U.S. over the recent four decades. Relationships between antecedent climate conditions with summer burned area are found to be strongest over non-forested and middle-to-high elevation areas (1100–3300 m). Statistical models that predict summer burned area using both antecedent and fireseason climate conditions have improved performance, explaining 69% of the interannual variability and 83% of the increasing trend of observed burned area. Among the antecedent climate predictors, vapor pressure deficit averaged over winter and spring plays the most critical role in predicting summer fire burned area. Spring snow drought area is found to be an important antecedent predictor for summer burned area over snow-reliant regions in the nonlinear statistical modeling framework used in this analysis. Namely, spring snow drought memory is realized through dry anomalies in land (soil and fuel) and atmospheric moisture during summer, which favours fire activity. This study highlights the important role of snow drought in subseasonal-to-seasonal forecasts of summer burned area over snow-reliant areas.

1. Introduction

Since the beginning of the 21st century, the Intergovernmental Panel on Climate Change has projected increases in summer wildfire risk in North America associated with decreased snowpack and increased summer drying caused by human-induced global warming [1, 2]. This projection holds true to a large degree in the western United States (U.S.) as observations have revealed persistent increasing trends in temperature, aridity and wildfire burned area since the mid-1980s [3–8]. It is well established that the observed increase in fire activity in the western U.S. is predominately explained by warmer and drier springs and summers caused by natural and human-induced climate changes [3, 8–11]. This is consistent with analyses of global paleo fire and climate records that indicate rapid warming has historically played an important role in modulating broad-scale fire activity over the past two millennia [12, 13]. Climate simulations considering a range of greenhouse gas concentration trajectories have projected that abrupt human-caused climate changes will likely contribute to further increases in fire hazard over the next century in the U.S. [14, 15]. This projection has striking societal and environmental consequences because fires cause thousands of smoke-related deaths per annum which is projected to increase throughout this century [15, 16], increased COVID-19 mortality [17], permanent changes to ecosystems [18], persistent

changes to streamflow [19], and multi-billion-dollar suppression expenditures (www.nifc.gov/fire-information/statistics/suppression-costs). A robust understanding of the relationships between fire activity and climate conditions susceptible to change can enable better preparation for these consequences via implementation of policy and management that addresses these relationships as well as models that more accurately predict fire activity to inform resource allocation.

Since the early 2000s, a suite of statistical analyses has quantified relationships between summer fire activity and various climate conditions in the western U.S. [3, 5, 9–11, 20–22]. Many of these relationships have been reviewed by Littell *et al* [23, 24]. These statistical models primarily differ based on spatial and temporal resolutions and domains, model complexity, and climate predictors. However, they each support that broad-scale fire activity across the western U.S. has been mostly explained by fluctuations in antecedent and fire season climate conditions since the mid-1980s. For instance, Westerling *et al* [3] concluded that wildfire activity in the western U.S. made an abrupt transition in the mid-1980s from infrequent and short-burning wildfires to more frequent longer-burning fires due to unusually warm springs and longer and drier summers. A series of follow-up studies have shown that after this transition, the majority of interannual variability in burned area can be explained by climate conditions [5, 9, 11].

The above body of work provides valuable insights on relationships between climate variability and fire. However, they do not provide in-depth analyses that relate only pre-fire climate conditions to summer fire activity, which are essential for lead-time (e.g. subseasonal-to-seasonal) fire forecasts [25]. Previous studies that have compared antecedent climate conditions with fire season burned area have reported significant relationships between pre-fire climate and fire activity, but have not comprehensively explored how much of the interannual variability or trend in fire season burned area can be explained by pre-fireseason climate alone [9, 10, 26–30]. Multiple of these studies have reported the effects of antecedent soil moisture conditions. Namely, wetter conditions 1–3 years before fire seasons in fuel-limited regions correspond with greater fire activity, and drier antecedent conditions in the months preceding the fire season tend to favour more fire activities [26–28]. A recent study that explored relationships between antecedent climate and fire season activity in a multivariate analysis showed machine learning models that predict burned area using both pre-fire and fire season conditions outperform models that use only fire season conditions [20]. Although previous research underlines that antecedent conditions contain unique information for enhancing fireactivity predictability, it has not yet established a comprehensive quantitative relationship between fire season severity and pre-fire climate alone. Moreover, interannual changes in winter and spring snowpack have important relationships with fires [3, 10], motivating our hypothesis that a large portion of summer burned area variability and trend can be explained by pre-fire climate conditions alone, with an important contribution from pre-fire snowpack conditions in snow-reliant areas over the western U.S.

Indeed, dramatic temperature-driven declines in snowpack across the western U.S. [31] are considered an important link between climate trends and increasing fire hazard [3, 10]. Namely, reductions in winter and spring snowpack (particularly during snow drought) drive earlier spring melt, increase surface temperature and evaporation due to snow-albedo feedback, and more quickly deplete moisture of vegetation and soil, leading to longer and drier summers [32]. The historically observed trend of declining snowpack has been projected to continue through the remainder of this century [33, 34]. Thus, understanding relationships between snow drought and fire over contemporary landscapes is imperative to project how future snowpack changes will impact fire activity, particularly in the western U.S. However, the snow-drought-fire interactions and the predictive ability of snow drought combined with other winter and spring climate conditions in predicting summer fire activity has not been fully understood and evaluated.

In this study, we use ensembles of nonlinear statistical models to predict summer fire burned area across the western U.S. based on pre-fire (winter and spring) climate including snow drought conditions. A unique novelty of this study is that we explore the predictability of burned area with combinations of pre-fire climate conditions and the role of snow-drought-fire interactions in modulating summer fire hazard. The goal of this study is to answer three following science questions. (i) What percent of interannual variability and trend in summer burned area can be explained by pre-fire climate conditions alone? (ii) What predictive information is contained in summer climate conditions that is not provided by winter and spring climate conditions? (iii) To what degree can the inclusion of snow drought enhance the predictability of summer fire activity relative to other traditionally used climate predictors? The overarching objectives of this study are to advance the understanding of seasonal climate-fire relationships, explore a methodology capable of predicting broad-scale fire burned area across the western U.S., and thus better inform policy and resource allocation.

2. Method

2.1. Study domain

The spatial domain considered in this study includes all areas classified by MODIS satellite observations

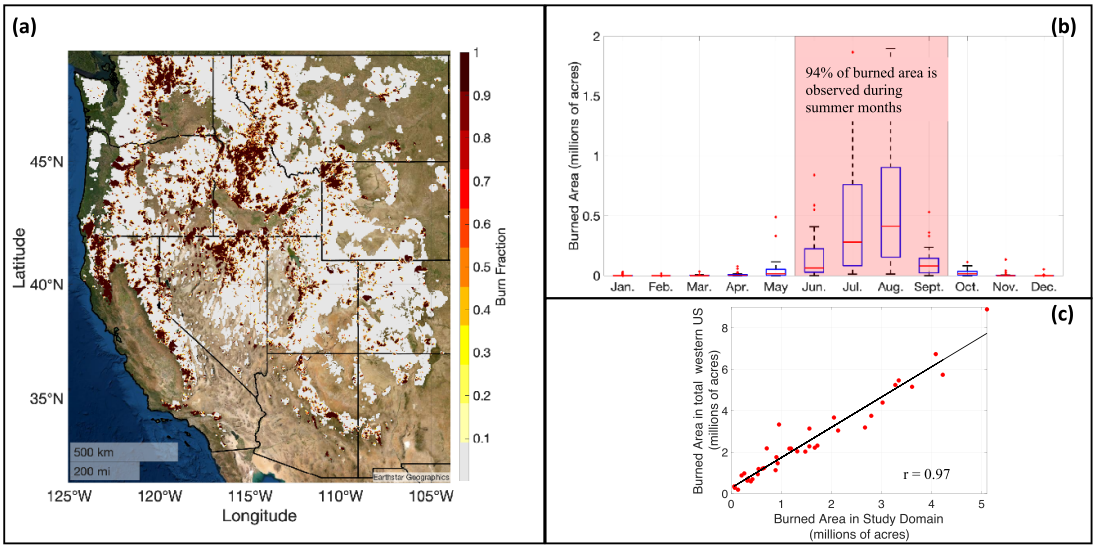

Figure 1. Study domain details. (a) White shaded regions represent the snow-effected area used for the analysis. 500 m resolution observed burned fractions from 1984 to 2020 (in yellow to red shades) are overlain. (b) Box and whisker plots of burned area by month highlight that 94% of the burned area from 1984 to 2020 has occurred during summer months. Variability in boxplots is from the spread in burned area across different years, and whisker lengths are equivalent to the interquartile range. Outliers are plotted as red '+'. (c) Scatter plot of summer burned area for the study domain shown in (a) (horizontal axis) and the entire western U.S. (vertical axis) during 1984–2020 reveals that interannual variability of summer burned area in the study domain is representative of the entire western U.S.

[35] as forest, grassland, or savanna in the western U.S. (north of 32° latitude and west of −104° longitude) below typical tree lines (<3300 m) [36] that have received adequate snowfall in winter and spring (peak snow water equivalent (SWE) > 100 mm) (figure 1(a)). The peak SWE threshold is imposed to limit the study domain to areas potentially affected by snow drought following Livneh and Badger [33] which supports assessing the importance of spring snow drought as a predictor of summer fire activity. The main results and conclusions presented herein are qualitatively insensitive to the choice of this peak SWE threshold (see sensitivity analyses in figures S1–S3 available online at stacks.iop.org/ERL/17/054030/mmedia). The preceding elevation and vegetation screening are performed to confine the domain to areas susceptible to fires, which reduces the domain area by 8% while retaining 99% of the burned area relative to all areas that meet the peak SWE threshold. After applying the aforementioned spatial screening, the study domain contains 43% land area and 61% burned area during 1984–2020 relative to the entire western U.S. (figure 1(a)). The interannual variability of burned area for the study domain is representative of the burned area over the entire western U.S., indicated by a very high correlation ($r = 0.97$) between fire-season burned area across the study domain and the entire western US (figure 1(c)). We select summer months (June–September) as the fire season in this analysis, which contains 94% of the total burned area within the study domain (figure 1(b)). Figure S10 shows that including May in the summer months (i.e. May–September) does not affect the qualitative relationships between pre-summer climate and summer burned area presented herein. Note that the domain has experienced an increasing trend in annual burned area from 1984 to 2020 for each calendar month (figure S4 and table S1).

2.2. Generalized additive models

All models in this study are trained or evaluated over a 37 year period (1984–2020) when satellite-observed fire records from Landsat and MODIS are available (see table S4). We employ the widely-used generalized additive models (GAMs) [37] using a Gaussian distribution to quantify the relationship between summer burned area and pre-fire climate conditions. GAMs are trained in the statistical software R [38] using the mgcv package [39]. GAMs are logistic regression models, similar to normal linear regression models, but replace the linear form of $\sum \beta_j x_j$ with a sum of nonlinear smooth functions $\sum s_j(x_j)$. Initial testing was performed using more complex machine learning algorithms: random forest (RF) and support vector machines (SVM). However, the GAMs provided a better performance (higher r and lower root mean square error (RMSE)) than the more complex RF and SVM models. Thus, this study uses GAMs for prediction and analyses due to their superior performance with high computational efficiency. Further, GAMs have been successfully used to forecast drought [40] and explore climate and fire relationships [26].

Climate predictors are pre-treated with a principal component analysis to remove linear correlations among each predictor set [41]. We use

all principal components as predictors in GAMs. This dramatically increases the computational efficiency of model selection and training by reducing the number of potential model configurations. Further, this allows all information from predictors to be used by GAMs. This choice may result in model overfitting which is penalized in the model ensemble selection procedure (described below) and robust evaluations presented in this study (i.e., leave-one-year-out cross validations and retroactive forecasts). To further limit the influence of nonlinear relationships between predictors (i.e. overlapping information), we only consider GAMs that compute a concurvity for each covariate to be less than 0.4 [42]. Following the methodology of Zhang *et al* [43], we include elevation bins as intercepts, as shown in equation (1), to allow for different baseline burned areas among three bins: (a) $0 \leqslant z < 1100$ m, (b) $1100 \leqslant z < 2200$ m, (c) $2200 \leqslant z \leqslant 3300$. Thus, this binning procedure allows for the inclusion of elevation effects on climate-fire relationships (e.g., [28]) while tripling the number of data points in model fitting (37 years $\times$ 3 elevation bins $=$ 111 data points) which improves model robustness. Binning data by elevation also implicitly accounts for elevation-dependent vegetation distributions. Specifically, lower elevation areas (0–1100 m) are comprised of 43% forested and 57% non-forested areas, whereas middle (1100–2200 m) and high (2200–3300 m) elevation areas are comprised primarily of non-forested vegetation classifications (86% and 96%, respectively) based on 500 m MODIS land type data.

GAMs used in this study can be written as:

$$g(E(\text{Burned Area})) = s(\text{PC}_1) + \ldots + s(\text{PC}_n) + \text{elevation bin} \quad (1)$$

where $g(\cdot)$ is a Gaussian link function, E(Burned Area) is the expected value of burned area given the set of predictors, $s(\cdot)$ is a nonlinear smooth function of climate predictors, PC_i is the ith principal component for a set of climate predictors, and n is the number of climate predictors used in each GAM. To explicitly account for vegetation effects on climate-fire relationships, we run an additional experiment where GAM predictors are binned by vegetation (forested and non-forested areas) instead of elevation.

We use an ensemble approach to ensure that results are insensitive to model-specific characteristics. First, 100 bestperforming model ensemble members (evaluated based on all unique combinations of the antecedent climate predictors described in table 1 and data sources listed in table S4; totally 121,652 ensemble members tested) are chosen based on minimized Akaike Information Criterion. To enforce consistency in model structure across each ensemble, we group each of the 100-member ensembles according to their number of climate predictors (i.e. ten ensembles using two to eleven predictors). We do not consider models exceeding eleven predictors as these are prone to over-parameterization (i.e. overfitting). We select the bestperforming ensemble (i.e. optimal number of predictors) on the basis of a maximized Taylor Score [44] in the leave-one-year-out cross-validation for each ensemble mean to ensure presentation of robust predictions (figure S9) [25].

A novel aspect of these models is the inclusion of snow drought conditions based on a recently developed snow drought index (SWEI; see section 2.4) [45]. This is because, despite possible interactions between early spring snowmelt and fire activity [3, 46], snow drought conditions have not been assessed within multivariate fire prediction frameworks. We also compare the effectiveness of snow drought with that of a traditionally used meteorological drought index (Palmer drought severity index; PDSI) in predicting summer burned area, since PDSI has been a widely used predictor of burned area within statistical modeling frameworks [9, 11, 23, 25]. All results presented herein are from the bestperforming model ensembles using snow drought conditions or PDSI conditions, which use five or seven predictors, respectively (figure S9).

2.3. Remotely sensed burned area

As described in table S4, we use fire perimeters from the Landsat-based Monitoring Trends in Burn Severity (MTBS) dataset [47] to derive 1984–2019 summer burned area. The MTBS dataset considers large fires (>404 ha for the western U.S.), which have been estimated to account for greater than 95% of total burned area in the domain [48]. Following the methodology from Higuera and Abatzoglou [5], we derive the 2020 fire data, which has not been included in the MTBS dataset at the time of conducting this study, by bias correcting MODIS burned area [49] towards MTBS data using a simple linear model. A linear model is adequate for this bias correction because MODIS and MTBS summer burned area are highly correlated ($r^2 = 0.99$). The high correlation between MTBS and MODIS burned area also supports that MTBS accurately depicts interannual variability and trend in summer burned area in spite of its omission of small fires.

2.4. Climate data

Climate conditions evaluated in this study (table 1) are from widely used spatially and temporally continuous data products (table S4). All climate data used in this study are derived entirely or partially from observations with long-term high-resolution gridded records whenever possible (table S4). Given the lack of spatially and temporally continuous observational products for soil moisture, evapotranspiration (ET) and potential evapotranspiration (PET), we use a ~12 km hourly reanalysis dataset (Phase 2 of the North American Land Data

Table 1. Descriptions of climate predictors. Climate predictors are averaged over winter (November through February), spring (March through May), winter–spring (November through May) and summer months (June through September). All spatial averaging is performed over the entire study domain (see figure 1).

Predictor	Description
Spring snow water equivalent drought index (SWEI)	Temporally and spatially averaged SWEI during spring months. SWEI calculated based on Huning and AghaKouchak [45]
Spring snow drought area	Total area considered to be in snow drought conditions (spring SWEI < -0.5)
Winter vapor pressure deficit (VPD)	Temporally and spatially averaged VPD during winter months
Spring vapor pressure deficit	Temporally and spatially averaged VPD during spring months
Winter–spring vapor pressure deficit	Temporally and spatially averaged VPD during winter and spring months
Summer vapor pressure deficit	Temporally and spatially averaged VPD during summer months
Winter temperature	Temporally and spatially averaged temperature during winter months
Spring temperature	Temporally and spatially averaged temperature during spring months
Winter–spring temperature	Temporally and spatially averaged temperature during winter and spring months
Summer temperature	Temporally and spatially averaged temperature during summer months
Winter precipitation	Temporally summed and spatially averaged precipitation during winter months
Spring precipitation	Temporally summed and spatially averaged precipitation during spring months
Winter–spring precipitation	Temporally summed and spatially averaged precipitation during winter and spring months
Summer precipitation	Temporally summed and spatially averaged precipitation during summer months
Winter potential evapotranspiration (PET)	Temporally and spatially averaged PET during winter months
Spring potential evapotranspiration	Temporally and spatially averaged PET during spring months
Winter–spring potential evapotranspiration	Temporally and spatially averaged PET during winter and spring months
Summer potential evapotranspiration	Temporally and spatially averaged PET during summer months
Winter evapotranspiration (ET)	Temporally and spatially averaged ET during winter months
Spring evapotranspiration	Temporally and spatially averaged ET during spring months
Winter–spring evapotranspiration	Temporally and spatially averaged ET during winter and spring months
Summer evapotranspiration	Temporally and spatially averaged ET during summer months
Summer soil moisture	Temporally and spatially averaged total column soil moisture during summer months
Spring Palmer Drought Severity Index (PDSI)	Temporally and spatially averaged PDSI during spring months.
PDSI drought area	Total area considered to be in drought conditions (spring PDSI < -0.5)

Assimilation System) to quantify these conditions [50, 51]. Snow drought is derived using the Huning and AghaKouchak [45] snow water equivalent index (SWEI) formulation throughout spring months (March through May) of each year, where SWE is taken from the observationally-based 4 km daily University of Arizona Snow Water Equivalent (UA-SWE) dataset [52, 53] which assimilates a suite of snow measurements with modeled SWE (table S4). All climate data are interpolated to the MODIS 500 m grid using the nearest neighbour approach as a pre-processing step. We include both spatially averaged SWEI and snow drought area as predictors, where snow drought area is defined as the area experiencing snow drought (SWEI < -0.5) [45] each year. To enforce consistency in the comparison of models using SWEI and those using PDSI, we also consider spatially averaged PDSI and PDSI drought area (areas where PDSI < -0.5) [54] during spring months. Based on previous research [23], we consider five other antecedent climate variables—vapor pressure deficit (VPD), surface temperature, precipitation, PET, and ET—that are spatially aggregated and temporally averaged across winter (November–February), spring (March–May), or winter–spring (November–May) to create three antecedent climate predictors for each of these five variables. Hence, we consider 17 potential antecedent climate predictors in burned area statistical models. Precipitation, temperature and VPD come from the observationally-based 4 km daily PRISM dataset [55]. Summer climate conditions listed in table 1 are used in combination with antecedent conditions in GAMs to explore the added benefit of including fireseason climate conditions in addition to pre-fire conditions in predicting summer burned area. Correlations between antecedent

predictors and summer burned area are presented in figure S5 and correlations between summer predictors and summer burned area are presented in figure S6.

3. Results

3.1. Pre-fire climate conditions alone explain a large portion of interannual variability and trend in summer burned area

To quantify the interannual variability and trend in summer burned area explained by winter and spring climate conditions, we apply a 100-member ensemble of GAMs to predict the extent of summer (June–September) burned area using only pre-fire winter (November–February) and spring (March–May) climate conditions across the western U.S. over a 37 year period (1984–2020) (see section 2.2 for model details and training procedure). All results shown in this section represent robust model evaluations where predictions are made outside of the training record. Relationships between climate and burned area presented herein should be interpreted as a lower bound because optimal GAMs selected in this analysis do not necessarily account for all relevant combinations of climate conditions, and thus there are likely model configurations that present stronger climate-fire relationships (e.g. section 3.3).

The GAM ensemble mean captures 43% of the interannual variability in observed summer burned area using snow drought and pre-fireseason climate conditions (figures 2(a) and (b)) and correctly predicts the sign of burned area anomalies for 26 of 37 years (70%). The ensemble mean has a low average bias (5.9% of observed burned area) but large RMSE (1 million acres or 65% of the mean observed burned area). The 95% ensemble range (2.5–97.5 percentiles) envelopes observed burned area for 21 of 37 years. The ensemble mean is also significantly correlated with the observed total burned area in the entire western US ($r = 0.58$, $p < 0.001$).

To test the forecasting ability of the GAM framework, we also performed a retroactive forecast validation to mimic the process of real-time operational forecasting that does not rely on future data for model training. In this analysis, retroactive forecasts are made over the second half of the study period (2002–2020) one-year at a time while training GAMs with data from preceding years. Retroactive forecasts from GAMs using antecedent snow drought and climate predictors shows skilful agreement with observed burned area ($r = 0.66$, $p < 0.01$), with the ensemble mean predicting the correct anomaly sign 74% of years and the 95% ensemble range containing observed values 63% of the years (figures S7(a) and (b)). However, these forecasts have a larger mean bias (−18%) and RMSE (1.2 million acres) relative to the leave-one-year-out analysis. The extreme 2020 fire season was largely underestimated by the GAM ensemble mean (see paragraph below for detailed discussion).

Accounting for both antecedent (winter and spring) and fire season climate conditions (table 1) improves model performance (relative to GAMs using only antecedent predictors), with the GAM ensemble mean explaining 65% of the variability in observed burned area and a mean bias reduction from 5.9% to 2.6% in the leave-one-year-out cross validation (figure 2). GAM-predicted burned area in the study domain is also highly correlated with the observed burned area in the entire western US ($r = 0.74$). Inclusion of summer conditions results in only a slight bias reduction for the extreme 2020 fire season which is underestimated by GAMs (figure 2). Failure to capture the extreme 2020 burned area extent is likely due to the rarity of the event which is not reflected in the training data [20, 56]. Recent research suggests the extreme 2020 fire season was enabled by anomalously dry conditions [5]; however, burned area underestimation from our climate-based statistical models motivate future research to further understand the respective contributions of large-scale climate and factors such as local fire weather, fuel availability and ignitions to this extreme fire year. Overall, 39% of the remaining burned area variability that is not explained with antecedent climate predictors can be explained by the unique information from summer climate predictors. The rest (35%) of the burned area variability this is not explained by these antecedent and fireseason climate conditions is likely caused by the combined effects of other climate conditions that are not considered in these models (e.g. meteorological drought, see section 3.3), local fire-weather (e.g. wind speed) [57], fuel availability [58, 59], and ignition sources [60, 61]. Table S2 provides a summary of GAM model setups and table S3 summarizes the corresponding evaluation metrics.

Trends in predicted and observed broad-scale burned area from 1984 to 2019 are presented in figure 2. Because the summer of 2020 had the unusually extreme burned area and appears at the end of the considered record, including it in the trend calculation results in a dramatic increase (17%) in the observed trend and therefore is not included in trends reflected in figure 2. However, in this paragraph we report trends including and excluding the 2020 burned area. The GAM ensemble mean using only antecedent predictors simulates an increasing trend of 2.1%/year (from 1984 to 2019) compared to the observed 3.6%/year trend, suggesting that a large portion (58%) of burned area's increasing trend can be explained by pre-fire season climate alone from 1984 to 2019. When including 2020 in this calculation, the modelled trend is 1.9%/year and the observed trend is 4.1% of the year, with GAMs

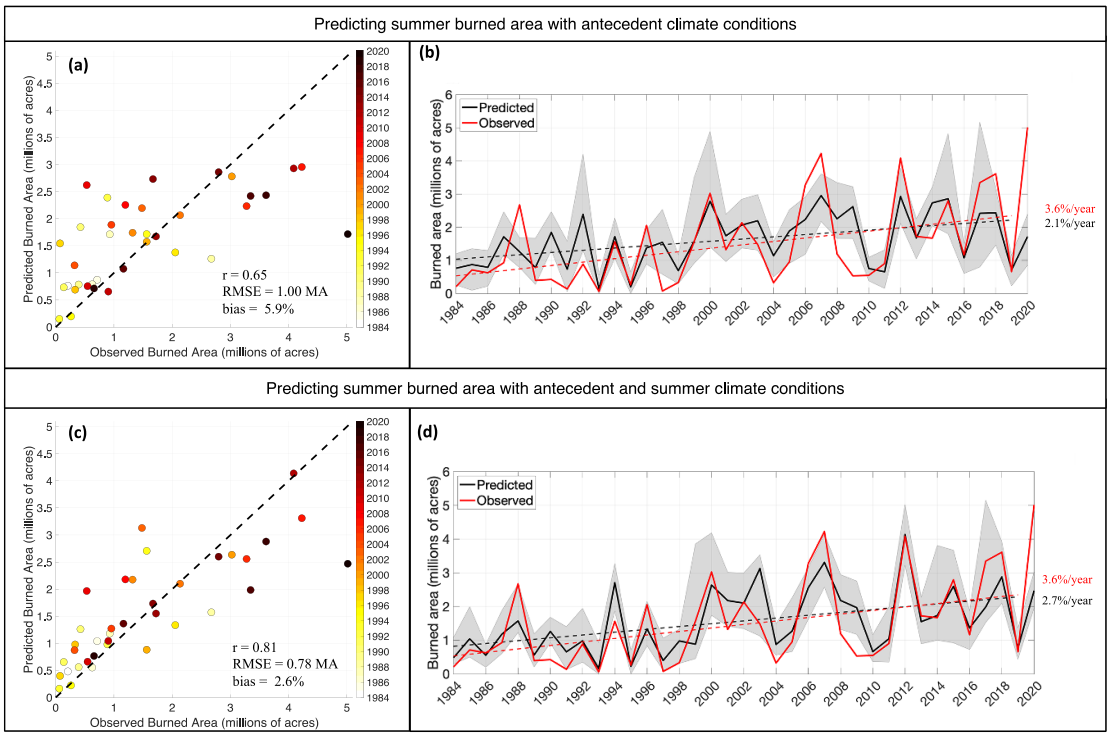

Figure 2. Comparing observed and predicted burned area from a leave-one-year-out cross validation. Predicted burned area represents GAM simulated summer burned area based on climate conditions in leave-one-year-out analyses. (a), (b) Predicted summer burned area is from models that include only antecedent (winter and spring) climate conditions. (c), (d) Predicted summer burned area is from models that include both antecedent and summer climate conditions. Scatter plots in (a), (c) represent observed (horizontal axis) and predicted (vertical axis) summer burned area, with dots colored by year. Time series shown in (b) and (d) present observed and predicted summer burned area for years 1984–2020. Grey shading represents model uncertainty based on the range of the bestperforming 100-member ensemble (2.5–97.5 percentiles). Black and red dashed lines represent the increasing burned area trend from 1984 to 2019 from simulations and observations, respectively. Pearson's correlation coefficient (r), RMSE in millions of acres (MA) and percent bias between the ensemble mean and observed burned area are shown on respective scatter plots.

explaining 46% of the observed trend from 1984 to 2020. When GAMs include summer climate conditions as additional predictors, the simulated trend for 1984–2019 and 1984–2020 increases to 2.7%/year. Hence, the combined antecedent and summer climate explains 75% of the observed trend from 1984 to 2019 and 66% of the observed trend when considering 1984–2020.

Relationships between antecedent climate and summer burned area are strongest in middle-to-high elevation and non-forested areas (figure 3). GAMs using only antecedent climate conditions predict an increasing burned area trend of 14%, 61% and 72% of the observed trend in low, middle and high elevation regions, respectively (0–1100 m, 1100–2200 m and 2200–3300 m, respectively). Antecedent climate conditions alone explain 46% of summer burned area interannual variability in middle elevation areas (figures 3(c) and (d)). The low correlation between predicted and observed burned area in high elevation areas is largely attributable to the underestimation in the 1988 fire season, which was largely affected by the Yellow Stone National Park fires (figures 3(e) and (f)). When this year is removed from the correlation calculation, antecedent climate explains 50% of the observed burned area interannual variability in high elevation areas. In low elevation areas antecedent climate only explains 6% of the observed burned area variability (figures 3(a) and (b)); however, when summer climate conditions are included as additional predictors, GAMs explain 40% of burned area variability. This indicates that wetter regions (typical for these low elevation areas) have weaker coupling between pre-summer climate and summer fire activity; noting that lower elevation areas that are more heavily forested (see section 2.2) receive 73% and 65% more mean winter–summer rainfall than middle and high elevation areas in this domain, respectively. This is consistent with previous research that found fire season precipitation exerts substantially stronger controls on fireseason burned area, relative to antecedent snow conditions, in western U.S. forests [62]. GAMs using antecedent climate predictors simulate an increasing burned area trend of 40% and 83% of the observed trend in forested and non-forested domains, while explaining 42% and 70% of the interannual variability in summer burned area, respectively (figures 3(g)–(j)).

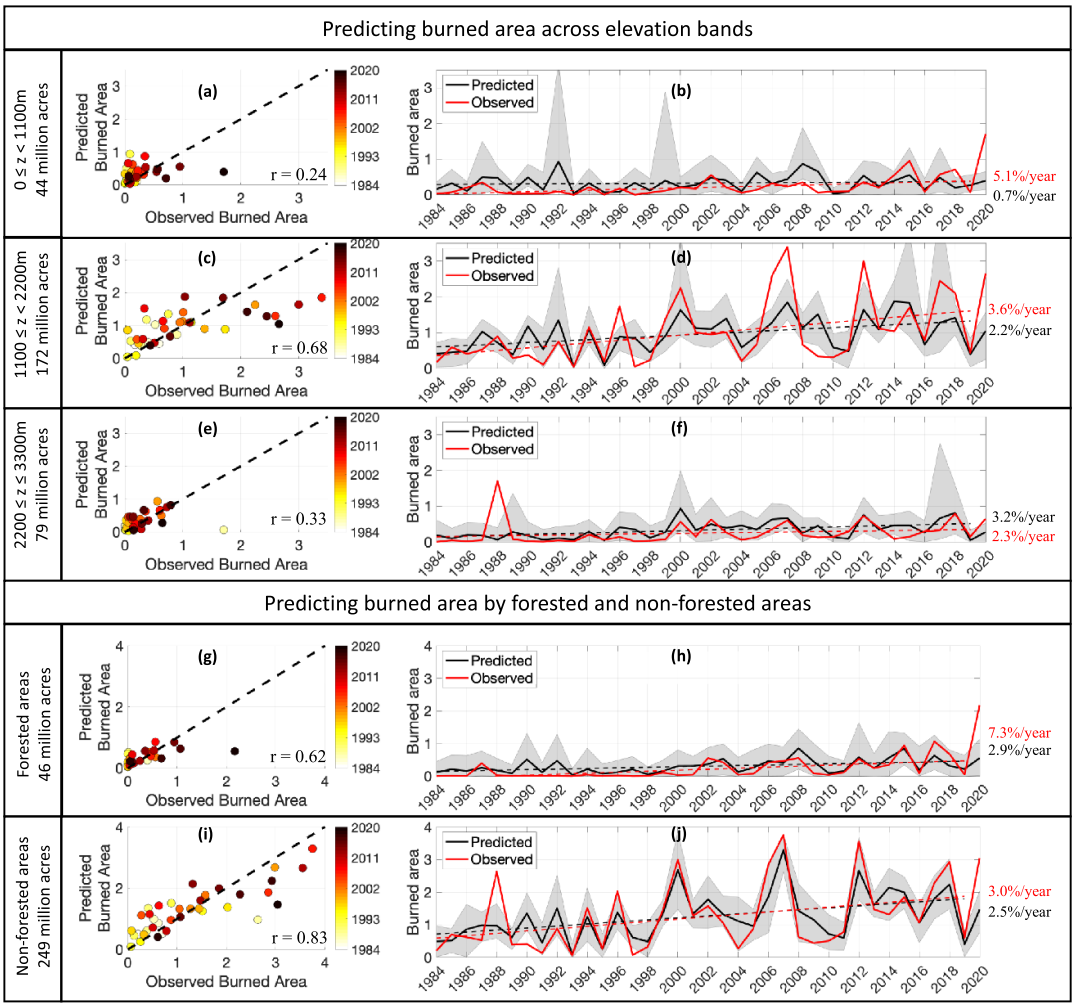

Figure 3. Comparing observed and predicted burned area from leave-one-year-out cross validations based on elevation and vegetation classifications. Predicted burned area represents GAM simulated summer burned area based on antecedent climate conditions in leave-one-year-out analyses. (a)–(f) Are the same predictions presented in figure 2 (a), (b), except predictions are maintained separated by elevation bins rather than aggregated for the domain-wide total burned area. Burned area is presented in millions of acres. (a), (b) Predicted and observed burned area for 500 m pixels between 0 and 1100 m. (c), (d) Predicted and observed burned area for 500 m pixels between 1100 and 2200 m. (e), (f) Predicted and observed burned area for 500 m pixels between 2200 and 3300 m. (g), (h) Predicted and observed burned area for 500 m pixels classified by MODIS as forested. (i), (j) Predicted and observed burned area for 500 m pixels classified by MODIS as non-forested vegetation. Scatter plots represent observed (horizontal axis) and predicted (vertical axis) summer burned area, with dots colored by year. Time series present observed and predicted summer burned area for years 1984–2020. Grey shading represents model uncertainty based on the range of the bestperforming 100-member ensemble (2.5–97.5 percentiles). Black and red dashed lines represent the increasing burned area trend from 1984 to 2019 from simulations and observations, respectively. Pearson's correlation coefficient (r) between the ensemble mean and observed burned area are shown on respective scatter plots.

3.2. The importance of snow drought as a predictor of summer fire burned area

Figure 4 presents the relative importance of all the pre-fire climate predictors based on two measures. First, we consider the number of model ensemble members in which each predictor appears as a measure of importance. As such, we assume that predictors with higher importance will appear more often in the bestperforming 100 model ensemble members. Secondly, we perform a leave-one-column-out (LOCO) analysis [20]. The LOCO importance is measured by repeatedly training statistical models, each time without one particular predictor, and computing the ratio of skill metrics (e.g. RMSE and r^2) for the re-trained model relative to the original model including all of its original predictors.

Analyses show that snow drought area is an important antecedent predictor of summer burned area, as the third most frequently used predictor for the bestperforming 100 ensemble members (57% appearances; figure 4(a)). Only winter–spring VPD and winter VPD appear more often than snow drought area. All the other predictors appear in less than half of the ensemble members (figure 4(a)). This finding expands on previous work that showed concurrent VPD is closely related with fire burned area [4–6, 63] by highlighting the importance of antecedent VPD as a predictor in a multivariate

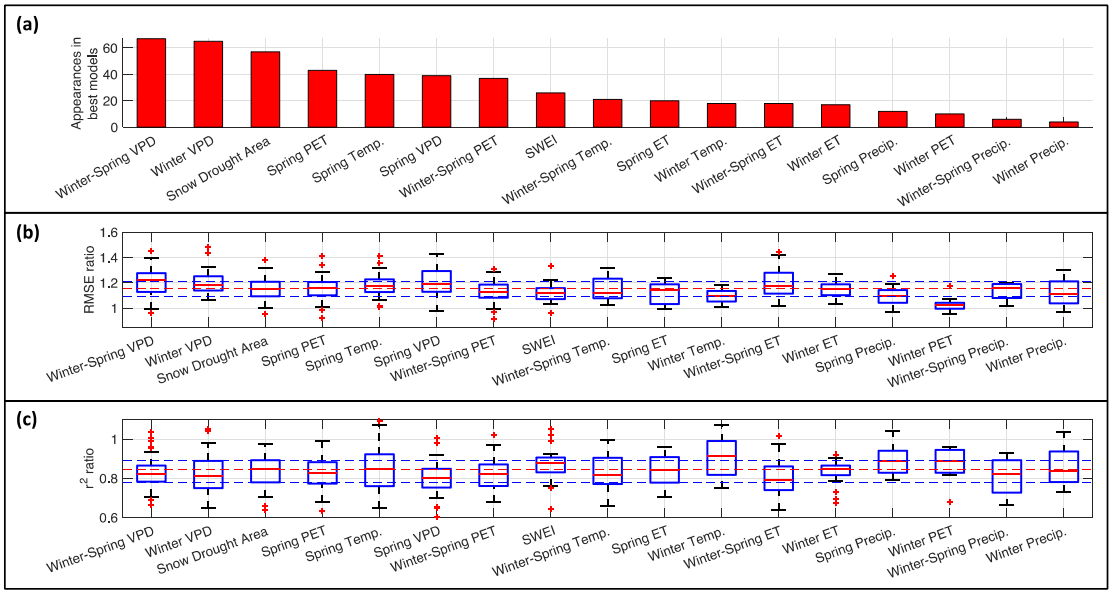

Figure 4. Predictor importance quantified by number of appearances in the 100 best performing model ensemble members and a LOCO analysis. (a) Number of times that each predictor appears in the bestperforming 100 ensemble members. (b) LOCO analysis presented in context of RMSE as the ratio of best-fit RMSE when respective predictors are excluded from model construction relative to the best-fit RMSE using the original set of climate predictors. A ratio greater than unity indicates performance is degraded by removing respective predictors. (c) LOCO analysis results presented in context of r^2 as the ratio of best-fit r^2 when respective predictors are excluded from model construction relative to the best-fit r^2 using the original set of climate predictors. A ratio less than unity indicates performance is degraded by removing respective predictors. Red and blue dashed reference lines in (b), (c) indicate the median and interquartile range for statistics corresponding to snow drought area, respectively. Variability in boxplots is from the ensemble spread, and whisker lengths are equivalent to the interquartile range. Outliers (extreme values) are plotted as red '+'. Predictor descriptions are provided in table 1.

framework. For the LOCO RMSE and r^2 analyses, snow drought area is considered valuable, with increases in RMSE and decreases in r^2 when the snow drought area predictor is left out (figures 4(b) and (c)). However, these LOCO analyses do not find snow drought area to be significantly more important than most other antecedent predictors. Overall, snow drought area is a valuable predictor of summer fire activity, yet other predictors can play a more valuable role in some model subsets.

The importance of snow drought in statistical models is consistent with previous research linking fire activity in the western U.S. with spring snowmelt timing [3, 46]. Spring snow drought is a valuable predictor due to the long memory of snowpack [64] that is realized through summer fire hazard conditions. For instance, SWEI is significantly ($p < 0.05$) correlated with summer soil moisture or VPD over 71% of the study domain (figures 5(a) and (b)), and hence the domain-averaged SWEI is significantly correlated ($r = 0.54$, $p < 0.01$) with the domain-averaged summer soil moisture and VPD. To further understand the underlying physical mechanisms, we employ the widely-used physically-based Noah with multi-parameterization land surface model (LSM) [65, 66] simulations to quantify responses of summer fire-favourable conditions to perturbations of spring snowpack via increasing or decreasing modeled April 1st SWE by 30% (details on LSM simulations are provided in the supplement text S1). The simulation with decreased spring SWE shows a 32 day earlier snow melt-out than the simulation with increased spring SWE. In turn, throughout the duration of the summer, the earlier snowmelt caused by the decreased spring SWE drives lower soil moisture and higher Bowen ratios (figures 5(c) and (d)), which favour enhanced summer VPD [67, 68] and thus increased fire activity [4–6, 22]. The simulations presented here are offline LSM experiments driven by observation-constrained atmospheric conditions to mitigate the associated uncertainty from atmospheric forcings, and hence do not include land surface feedback to the atmosphere, which are expected to further strengthen the relationship between dry springs and dry summers due to drier soils favouring less precipitation in summer months [69]. Therefore, the physical model experiments support the statistical relationships shown in figures 5(a) and (b) and reveal the associated mechanisms as mentioned above.

3.3. Comparing snow drought predictive abilities to the traditionally used Palmer Drought Severity Index

A breadth of research has quantified relationships between the traditionally-used PDSI [70] and fire in the western U.S. [9–11] which has been leveraged to create burned area forecasting frameworks [25, 71]. However, the PDSI does not explicitly consider snowpack processes, which is particularly important

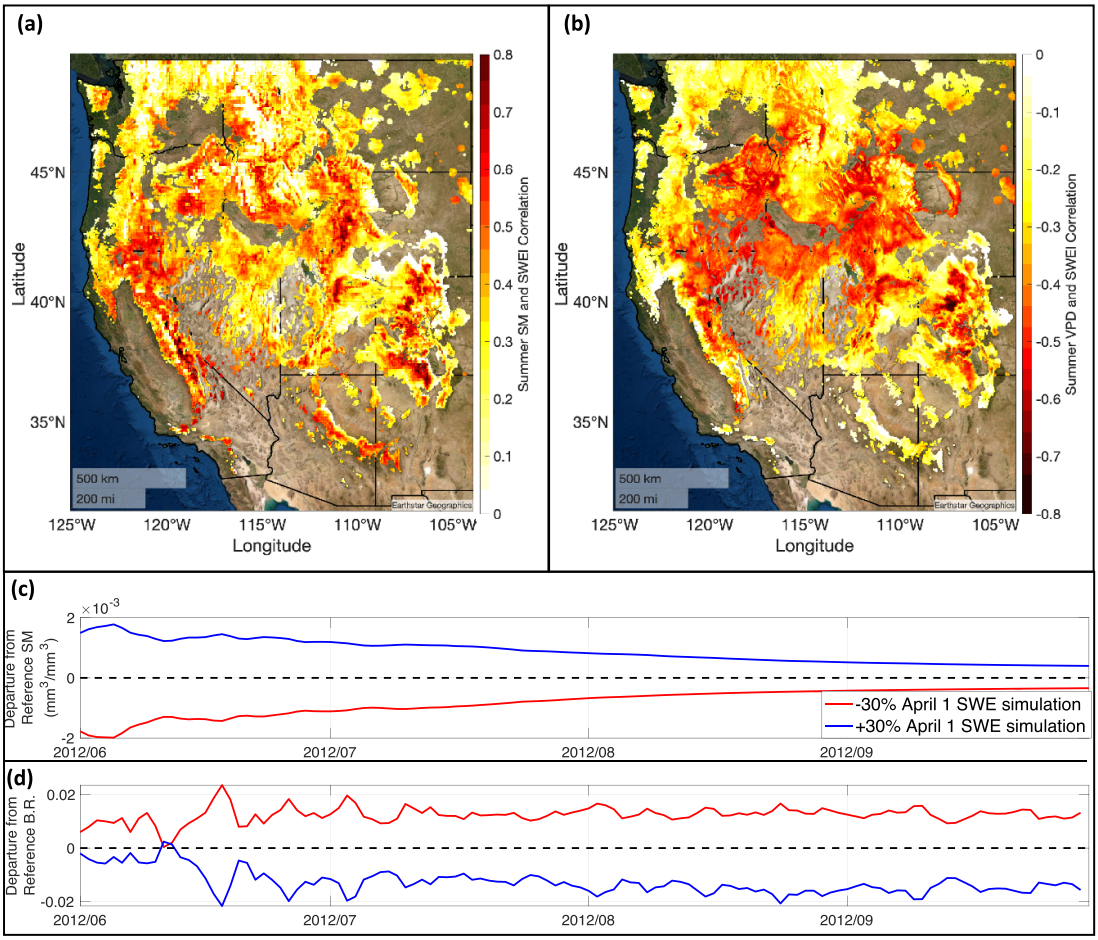

Figure 5. Spring snow drought memory is realized through summer fire hazard conditions. (a) Correlations between spring snow drought (SWEI) and summer soil moisture. (b) Correlations between spring snow drought (SWEI) and summer VPD. Data presented in (a), (b) are from sources listed in table S4. (c) Differences in time series of summer soil moisture from physical land surface model simulations that imposed a 30% decrease (red line) or increase (blue line) in 1 April SWE relative to a baseline simulation. The black dashed line at 0 is shown as reference marking no departure from the baseline simulation. (d) Same as (c) but for the Bowen ratio (B.R; ratio of sensible heat to latent heat flux). (c), (d) Present spatially averaged departures of soil moisture and Bowen ratios from the baseline simulation, where negative (positive) values indicate decreases (increases) relative to the baseline simulation. See supplementary information text S1 for details of model simulations.

over the western U.S. In future climate-based burned area predictions, it is important to resolve if snow drought is a better predictor for fire activity than PDSI over regions affected by snow processes.

To answer this question, we re-train GAMs, as described above, but replace SWEI-related variables (spatially averaged SWEI and SWEI drought area during spring months) with the PDSI equivalent (spatially averaged PDSI and PDSI drought area during spring months; see section 2.4, tables 1 and S4). The best 100-member PDSI model ensemble is re-selected rather than using the same best-fit predictors from SWEI models. The ensemble mean for GAMs using spring PDSI with other antecedent predictors has relatively better performance in the leave-one-year-out cross validation than the SWEI ensemble: $r = 0.73$ for PDSI versus $r = 0.65$ for SWEI models, with PDSI models also having a slightly lower RMSE and mean bias (figure S8 and table S3). We also found that the PDSI ensemble using antecedent climate and summer predictors explains 69% of the summer burned area interannual variability and 83% of the observed trend in a leave-one-year-out cross validation, which is higher than the counterpart SWEI ensemble which explains 65% of the interannual variability and 75% of the increasing trend. In the retroactive validation, ensembles using PDSI instead of SWEI better capture the interannual variability of observed burned area ($r = 0.75$); however, the PDSI ensemble simulates a larger mean bias by 4% (figure S7). However, when considering the spatial domain more dominated by snowfall (with a higher peak SWE threshold of 150 mm instead of 100 mm: reducing the original study area by 27%), GAMs using SWEI and antecedent climate outperform the counterpart PDSI ensemble in the leave-one-year-out cross validations (figures S1 and S8; table S3). These results support the legacy of using PDSI to predict burned area over the

broader western U.S.; however, SWEI is a more valuable predictor than PDSI for forecasting burned area over more snow-reliant western U.S. areas (e.g. peak SWE > 150 mm).

4. Discussion

We further present a simple linear correlation analysis to show relationships between the climate predictors considered in this study (table 1) and summer burned area in figures S5 and S6. Consistent with previous literature [4–6], summer VPD has the strongest linear relationship ($r = 0.76$) with summer burned area compared to other climate conditions considered in this study. Interestingly, four (VPD, PET, ET and temperature) out of the five winter climate conditions have higher correlations than their spring counterparts; whereas spring precipitation shows a stronger relationship with summer burned area relative to winter precipitation. The three antecedent climate conditions that have the highest linear correlation with summer burned area are: winter–spring average VPD ($r = 0.59$), winter VPD ($r = 0.58$) and PDSI drought area ($r = 0.56$). Spatially averaged snow drought and snow drought area have significant linear correlations ($p \leqslant 0.05$) with summer burned area, which however are not anomalously high ($r = -0.32$ and 0.37, respectively), suggesting a complex and nonlinear relationship between snow drought and summer fire as discussed in section 3.2. Summer climate conditions have higher correlations than their antecedent counterparts, as expected. This is particularly true for precipitation which has a strong negative correlation with burned area when considering summer accumulation ($r = -0.62$; $p < 0.01$) but insignificant ($p > 0.1$) relationships when considering antecedent precipitation summed over winter, spring or winter through spring.

The sign of all linear correlations between various climate conditions and burned area are consistent with the assumption that warmer and drier conditions come with more burned area, which is supported by previous research [5, 10, 11]. The physical explanation for this is that fire spread is a series of ignitions where heat source(s) raise a fuel to ignition temperature. When fuels are moist, evaporation consumes a large sum of energy, reducing flammability and thus the rate of fire spread [72]. Hence, over the previous four decades, climate-driven fuel flammability has been the primary factor in governing interannual variability in burned area. However, this result is limited by the context in which fire-climate relationships are established: typically, large areas (e.g. regional-to-continental scale) over time periods of a few decades. At either relatively small spatial scales (e.g. county or state levels) or over long time periods (e.g. paleorecords that are 100s of years), fire-climate relationships may be violated. For instance, at local-to-regional spatial scales, variability in burned area can be dominated by combinations of factors other than climate such as: natural and anthropogenically-caused changes in fuel availability [26, 58, 59], local wind patterns [57], and natural and anthropogenic ignition sources [60, 61]. Furthermore, these fire-climate relationships may be violated when considering paleorecords. For example, from 1 to 1750 AD global cooling was found to drive a long-term decreasing trend in fire activity, whereas the decreasing trend in fire activity from 1750 into the late 20th century has been attributed to population-driven land-cover changes which overwhelmed the affects of climate warming during this period [12].

Thus, statistical relationships between climate and fire activity presented herein are limited to the context in which relationships are established, and interpretations of these relationships over longer time periods or smaller areas require careful consideration of land-use and land-cover changes [24, 73] or local weather conditions [74]. Hence, future long-term projections (e.g. over the next century) based on climate-fire statistical relationships presented in this study should be interpreted as changes in climate's contribution to fire hazard rather than predictions of total burned area because other factors (e.g. reduction of burnable vegetation) may counter affects of climate warming on fire activity.

5. Conclusion

Variations in winter and spring climate conditions alone can explain a large portion of the interannual variability of summer burned area over the recent 37 years across the western U.S. These relationships between burned area and pre-fire climate and snow drought variability reveals that the majority of climate-fire relationships can be explained through climatic preconditioning. Specifically, antecedent climate conditions are able to predict and explain up to 53% of burned area interannual variability, and combining antecedent and fireseason climate conditions increases that predictability to 69%. PDSI was found to be a better predictor for burned area than SWEI over a broader western U.S. domain; however, over a more snow-reliant area, statistical models using SWEI rather than PDSI simulate burned area more accurately. Spring snow drought area is an important predictor for summer burned area, because spring snowpack memory is realized through drier land and atmospheric conditions that are more conducive to summer wildfire spread. Our statistical modeling approach enables the creation of a novel subseasonal-to-seasonal fire forecasting framework, benefiting future fire risk management and warning systems.

Strong coupling between broad-scale climate and fire in the western U.S., as presented in this study, has been interpreted by some studies to suggest that restoration and fuel treatment are likely an ineffective approach towards reversing current increasing

wildfire trends [2]. However, human actions may be capable of weakening climate-fire coupling [26] as well as climate trends conducive to fire [6, 22], and thus may provide mitigation to projected increases in fire risk in the western U.S. Validation of land management strategies aimed towards reducing fire risk can leverage the methodology presented in this study to understand if employed strategies reduce coupling between climate and fire, particularly during hot and dry years.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://data.mendeley.com/datasets/53v5twv757/5.

Code availability

All scripts used to perform analysis are accessible here: https://github.com/RAbolafiaRosenzweig/Fire-Prediction-Western-US.

Acknowledgments

This study was supported by NOAA's Climate Program Office's Modeling, Analysis, Predictions, and Projections Program (MAPP), Grant# NA20OAR4310421. We also acknowledge the support from the NCAR Water System Program. NCAR is sponsored by the National Science Foundation. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation. We thank Dr. J Dudhia, Dr. M LeMone, Dr. Z Zhang, Dr. J Littell and Dr. W Tang for helpful discussions.

Conflict of interest

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

ORCID iDs

Ronnie Abolafia-Rosenzweig https://orcid.org/0000-0002-6169-6430
Cenlin He https://orcid.org/0000-0002-7367-2815
Fei Chen https://orcid.org/0000-0003-2573-3828

References

[1] Watson R T and Team C W 2001 *Climate Change 2001: Synthesis Report. A Contribution of Working Groups I, II, and III to the Third Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge: Cambridge University Press)
[2] Easterling W *et al* 2007 *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* ed M L Parry, O F Canziani, J P Palutikof, P J van der Linden and C E Hanson (Cambridge: Cambridge University Press) p 14
[3] Westerling A L 2006 Warming and earlier spring increase Western U.S. forest wildfire activity *Science* **313** 940–3
[4] Higuera P E, Shuman B N and Wolf K D 2021 Rocky mountain subalpine forests now burning more than any time in recent millennia *Proc. Natl Acad. Sci.* **118** e2103135118
[5] Higuera P E and Abatzoglou J T 2021 Record-setting climate enabled the extraordinary 2020 fire season in the western United States *Glob. Change Biol.* **27** 1–2
[6] Williams A P, Abatzoglou J T, Gershunov A, Guzman-Morales J, Bishop D A, Balch J K and Lettenmaier D P 2019 Observed impacts of anthropogenic climate change on wildfire in California *Earths Future* **7** 892–910
[7] Hu J M and Nolin A W 2020 Widespread warming trends in storm temperatures and snowpack fate across the Western United States *Environ. Res. Lett.* **15** 034059
[8] Abatzoglou J T and Williams A P 2016 Impact of anthropogenic climate change on wildfire across western US forests *Proc. Natl Acad. Sci.* **113** 11770–5
[9] Littell J S, McKenzie D, Peterson D L and Westerling A L 2009 Climate and wildfire area burned in western U.S. ecoprovinces, 1916–2003 *Ecol. Appl.* **19** 1003–21
[10] Abatzoglou J T and Kolden C A 2013 Relationships between climate and macroscale area burned in the western United States *Int. J. Wildland Fire* **22** 1003
[11] Riley K L, Abatzoglou J T, Grenfell I C, Klene A E and Heinsch F A 2013 The relationship of large fire occurrence with drought and fire danger indices in the western USA, 1984–2008: the role of temporal scale *Int. J. Wildland Fire* **22** 894
[12] Marlon J R, Bartlein P J, Carcaillet C, Gavin D G, Harrison S P, Higuera P E, Joos F, Power M J and Prentice I C 2008 Climate and human influences on global biomass burning over the past two millennia *Nat. Geosci.* **1** 697–702
[13] Marlon J R *et al* 2009 Wildfire responses to abrupt climate change in North America *Proc. Natl Acad. Sci.* **106** 2519–24
[14] Intergovernmental Panel on Climate Change (IPCC) 2021 *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* ed V Masson-Delmotte *et al* (Cambridge: Cambridge University Press)
[15] Ford B, Val Martin M, Zelasky S E, Fischer E V, Anenberg S C, Heald C L and Pierce J R 2018 Future fire impacts on smoke concentrations, visibility, and health in the contiguous United States *GeoHealth* **2** 229–47
[16] Johnston F H, Henderson S B, Chen Y, Randerson J T, Marlier M, DeFries R S, Kinney P, Bowman D M J S and Brauer M 2012 Estimated global mortality attributable to smoke from landscape fires *Environ. Health Perspect.* **120** 695–701
[17] Zhou X, Josey K, Kamareddine L, Caine M C, Liu T, Mickley L J, Cooper M and Dominici F 2021 Excess of COVID-19 cases and deaths due to fine particulate matter exposure during the 2020 wildfires in the United States *Sci. Adv.* **7** eabi8789
[18] Coop J D *et al* 2020 Wildfire-driven forest conversion in Western North American landscapes *BioScience* **70** 659–73
[19] Williams A P *et al* 2022 Growing impact of wildfire on western US water supply *Proc. Natl Acad. Sci.* **119** e2114069119
[20] Kuhn-Régnier A, Voulgarakis A, Nowack P, Forkel M, Prentice I C and Harrison S P 2021 The importance of antecedent vegetation and drought conditions as global drivers of burnt area *Biogeosciences* **18** 3861–79

[21] Alizadeh M R, Abatzoglou J T, Luce C H, Adamowski J F, Farid A and Sadegh M 2021 Warming enabled upslope advance in western US forest fires *Proc. Natl Acad. Sci.* **118** e2009717118
[22] Zhuang Y, Fu R, Santer B D, Dickinson R E and Hall A 2021 Quantifying contributions of natural variability and anthropogenic forcings on increased fire weather risk over the western United States *Proc. Natl Acad. Sci.* **118** e2111875118
[23] Littell J S, Peterson D L, Riley K L, Liu Y and Luce C H 2016 A review of the relationships between drought and forest fire in the United States *Glob. Change Biol.* **22** 2353–69
[24] Littell J S 2018 Drought and fire in the Western USA: is climate attribution enough? *Curr. Clim. Change Rep.* **4** 396–406
[25] Westerling A L, Gershunov A and Cayan D R 2003 Statistical forecasts of the 2003 Western wildfire season using canonical correlation analysis *Exp. Long-Lead Forecast Bull.* **12** 49–53
[26] Abatzoglou J T, Williams A P, Boschetti L, Zubkova M and Kolden C A 2018 Global patterns of interannual climate–fire relationships *Glob. Change Biol.* **24** 5164–75
[27] Krueger E S, Ochsner T E, Carlson J D, Engle D M, Twidwell D and Fuhlendorf S D 2016 Concurrent and antecedent soil moisture relate positively or negatively to probability of large wildfires depending on season *Int. J. Wildland Fire* **25** 657
[28] Crimmins M A and Comrie A C 2004 Interactions between antecedent climate and wildfire variability across south-eastern Arizona *Int. J. Wildland Fire* **13** 455
[29] Keeley J E 2004 Impact of antecedent climate on fire regimes in coastal California *Int. J. Wildland Fire* **13** 173
[30] Swetnam T W and Betancourt J L 1998 Mesoscale disturbance and ecological response to decadal climatic variability in the American Southwest *J. Clim.* **11** 20
[31] Mote P W, Li S, Lettenmaier D P, Xiao M and Engel R 2018 Dramatic declines in snowpack in the western US *npj Clim. Atmos. Sci.* **1** 2
[32] Luce C H, Vose J M, Pederson N, Campbell J, Millar C, Kormos P and Woods R 2016 Contributing factors for drought in United States forest ecosystems under projected future climates and their uncertainty *For. Ecol. Manage.* **380** 299–308
[33] Livneh B and Badger A M 2020 Drought less predictable under declining future snowpack *Nat. Clim. Change* **10** 452–8
[34] Marshall A M, Abatzoglou J T, Link T E and Tennant C J 2019 Projected changes in interannual variability of peak snowpack amount and timing in the Western United States *Geophys. Res. Lett.* **46** 8882–92
[35] Friedl M and Sulla-Menashe D 2019 MCD12Q1 MODIS/Terra+Aqua land cover type yearly L3 Global 500m SIN Grid V006 NASA EOSDIS Land Processes DAAC
[36] Berdanier A B 2010 Global treeline position *Nat. Educ. Knowl.* **3** 11
[37] Hastie T and Tibshirani R 1986 Generalized additive models *Stat. Sci.* **1** 297–318
[38] R Core Team 2020 R: a language and environment for statistical computing. R foundation for statistical computing (available at: www.R-project.org/)
[39] Wood S N 2011 Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models *J. R. Stat. Soc.* **73** 3–36
[40] Mathivha F, Sigauke C, Chikoore H and Odiyo J 2020 Short-term and medium-term drought forecasting using generalized additive models *Sustainability* **12** 4006
[41] Dormann C F *et al* 2013 Collinearity: a review of methods to deal with it and a simulation study evaluating their performance *Ecography* **36** 27–46
[42] Lydersen J M, Collins B M, Coppoletta M, Jaffe M R, Northrop H and Stephens S L 2019 Fuel dynamics and reburn severity following high-severity fire in a Sierra Nevada, USA, mixed-conifer forest *Fire Ecol.* **15** 43
[43] Zhang Z, Bortolotti L E, Li Z, Armstrong L M, Bell T W and Li Y 2021 Heterogeneous changes to Wetlands in the Canadian Prairies under future climate *Water Resour. Res.* **57** e2020WR028727
[44] Taylor K E 2001 Summarizing multiple aspects of model performance in a single diagram *J. Geophys. Res. Atmos.* **106** 7183–92
[45] Huning L S and AghaKouchak A 2020 Global snow drought hot spots and characteristics *Proc. Natl Acad. Sci.* **117** 19753–9
[46] Westerling A L 2016 Increasing western US forest wildfire activity: sensitivity to changes in the timing of spring *Phil. Trans. R. Soc.* B **371** 20150178
[47] Eidenshink J, Schwind B, Brewer K, Zhu Z-L, Quayle B and Howard S 2007 A project for monitoring trends in burn severity *Fire Ecol.* **3** 3–21
[48] Kolden C A, Lutz J A, Key C H, Kane J T and van Wagtendonk J W 2012 Mapped versus actual burned area within wildfire perimeters: characterizing the unburned *For. Ecol. Manage.* **286** 38–47
[49] Roy D P, Jin Y, Lewis P E and Justice C O 2005 Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data *Remote Sens. Environ.* **97** 137–62
[50] Xia Y *et al* 2012 Continental-scale water and energy flux analysis and validation for North American land data assimilation system project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow *J. Geophys. Res. Atmos.* **117** D03110
[51] Xia Y, Ek M, Wei H and Meng J 2012 Comparative analysis of relationships between NLDAS-2 forcings and model outputs *Hydrol. Process.* **26** 467–74
[52] Broxton P, Zeng X and Dawson N 2019 Daily 4 km Gridded SWE and Snow Depth from Assimilated In-Situ and Modeled Data over the Conterminous US, Version 1 (Boulder, CO: NASA National Snow and Ice Data Center Distributed Active Archive Center) (https://doi.org/10.5067/0GGPB220EX6A)
[53] Zeng X, Broxton P and Dawson N 2018 Snowpack change from 1982 to 2016 over conterminous United States *Geophys. Res. Lett.* **45** 12940–7
[54] Alley W M 1984 The Palmer Drought Severity Index: limitations and applications *J. Appl. Meteorol.* **23** 1100–9
[55] Daly C, Halbleib M, Smith J I, Gibson W P, Doggett M K, Taylor G H, Curtis J and Pasteris P P 2008 Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States *Int. J. Climatol.* **28** 2031–64
[56] Joshi J and Sukumar R 2021 Improving prediction and assessment of global fires using multilayer neural networks *Sci. Rep.* **11** 3295
[57] Abatzoglou J T, Balch J K, Bradley B A and Kolden C A 2018 Human-related ignitions concurrent with high winds promote large wildfires across the USA *Int. J. Wildland Fire* **27** 377
[58] Taylor A H, Trouet V, Skinner C N and Stephens S 2016 Socioecological transitions trigger fire regime shifts and modulate fire–climate interactions in the Sierra Nevada, USA, 1600–2015 CE *Proc. Natl Acad. Sci.* **113** 13684–9
[59] Wang S S-C and Wang Y 2020 Quantifying the effects of environmental factors on wildfire burned area in the south central US using integrated machine learning techniques *Atmos. Chem. Phys.* **20** 11065–87
[60] Veraverbeke S, Rogers B M, Goulden M L, Jandt R R, Miller C E, Wiggins E B and Randerson J T 2017 Lightning as a major driver of recent large fire years in North American boreal forests *Nat. Clim. Change* **7** 529–34
[61] Nagy R C, Fusco E, Bradley B, Abatzoglou J T and Balch J 2018 Human-related ignitions increase the number of large wildfires across U.S. ecoregions *Fire* **1** 4
[62] Holden Z A, Swanson A, Luce C H, Jolly W M, Maneta M, Oyler J W, Warren D A, Parsons R and Affleck D 2018

Decreasing fire season precipitation increased recent western US forest wildfire activity *Proc. Natl Acad. Sci.* **115** E8349–57
[63] Williams A P *et al* 2015 Correlations between components of the water balance and burned area reveal new insights for predicting forest fire area in the southwest United States *Int. J. Wildland Fire* **24** 14
[64] Xu L and Dirmeyer P 2011 Snow-atmosphere coupling strength in a global atmospheric model: snow-atmosphere coupling *Geophys. Res. Lett.* **38** L13401
[65] Niu G-Y *et al* 2011 The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements *J. Geophys. Res.* **116** D12109
[66] Yang Z-L *et al* 2011 The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins *J. Geophys. Res.* **116** D12110
[67] Seager R, Hooks A, Williams A P, Cook B, Nakamura J and Henderson N 2015 Climatology, variability, and trends in the U.S. vapor pressure deficit, an important fire-related meteorological quantity *J. Appl. Meteorol. Climatol.* **54** 1121–41
[68] Ficklin D L and Novick K A 2017 Historic and projected changes in vapor pressure deficit suggest a continental-scale drying of the United States atmosphere *J. Geophys. Res. Atmos.* **122** 2061–79
[69] Quiring S M and Kluver D B 2009 Relationship between winter/spring snowfall and summer precipitation in the northern great plains of North America *J. Hydrometeorol.* **10** 1203–17
[70] Palmer W C 1965 Meteorological drought. Weather Bureau Research Paper No. 45 (Washington, DC: U. S. Dept. of Commerce) p 58
[71] Preisler H K and Westerling A L 2007 Statistical model for forecasting monthly large wildfire events in Western United States *J. Appl. Meteorol. Climatol.* **46** 1020–30
[72] Brey S J, Barnes E A, Pierce J R, Swann A L S and Fischer E V 2021 Past variance and future projections of the environmental conditions driving Western U.S. summertime wildfire burn area *Earths Future* **9** e2020EF001645
[73] McKenzie D and Littell J S 2017 Climate change and the eco-hydrology of fire: will area burned increase in a warming western USA? *Ecol. Appl.* **27** 26–36
[74] Nelson R M Jr 2002 An effective wind speed for models of fire spread *Int. J. Wildland Fire* **11** 153
[75] Abolafia-Rosenzweig R, He C and Chen F 2022 *Mendeley Data* (available at: https://data.mendeley.com/datasets/53v5twv757/5) (Accessed 24 March 2022)