Holdrege *et al. Fire Ecology* (2024) 20:22
https://doi.org/10.1186/s42408-024-00252-4

Fire Ecology

ORIGINAL RESEARCH

Open Access

Wildfire probability estimated from recent climate and fine fuels across the big sagebrush region

Martin C. Holdrege[1,2*], Daniel R. Schlaepfer[1,2], Kyle A. Palmquist[3], Michele Crist[4], Kevin E. Doherty[5], William K. Lauenroth[6], Thomas E. Remington[7], Karin Riley[8], Karen C. Short[8], John C. Tull[9], Lief A. Wiechman[10] and John B. Bradford[1,11]

Abstract

Background Wildfire is a major proximate cause of historical and ongoing losses of intact big sagebrush (*Artemisia tridentata* Nutt.) plant communities and declines in sagebrush obligate wildlife species. In recent decades, fire return intervals have shortened and area burned has increased in some areas, and habitat degradation is occurring where post-fire re-establishment of sagebrush is hindered by invasive annual grasses. In coming decades, the changing climate may accelerate these wildfire and invasive feedbacks, although projecting future wildfire dynamics requires a better understanding of long-term wildfire drivers across the big sagebrush region. Here, we integrated wildfire observations with climate and vegetation data to derive a statistical model for the entire big sagebrush region that represents how annual wildfire probability is influenced by climate and fine fuel characteristics.

Results Wildfire frequency varied significantly across the sagebrush region, and our statistical model represented much of that variation. Biomass of annual and perennial grasses and forbs, which we used as proxies for fine fuels, influenced wildfire probability. Wildfire probability was highest in areas with high annual forb and grass biomass, which is consistent with the well-documented phenomenon of increased wildfire following annual grass invasion. The effects of annuals on wildfire probability were strongest in places with dry summers. Wildfire probability varied with the biomass of perennial grasses and forbs and was highest at intermediate biomass levels. Climate, which varies substantially across the sagebrush region, was also predictive of wildfire probability, and predictions were highest in areas with a low proportion of precipitation received in summer, intermediate precipitation, and high temperature.

Conclusions We developed a carefully validated model that contains relatively simple and biologically plausible relationships, with the goal of adequate performance under novel conditions so that useful projections of average annual wildfire probability can be made given general changes in conditions. Previous studies on the impacts of vegetation and climate on wildfire probability in sagebrush ecosystems have generally used more complex machine learning approaches and have usually been applicable to only portions of the sagebrush region. Therefore, our model complements existing work and forms an additional tool for understanding future wildfire and ecological dynamics across the sagebrush region.

Keywords *Artemisia tridentata*, Climate, Fine fuels, Sagebrush, Statistical model, Wildfire probability

*Correspondence:
Martin C. Holdrege
martinholdrege@gmail.com
Full list of author information is available at the end of the article

Resumen

Antecedentes Los incendios de vegetación son la causa histórica más próxima y en curso, de las pérdidas en las comunidades intactas de artemisa grande (Artemisia tridentata Nutt.) y de las declinaciones en la fauna consumidora obligada de esa especie. En décadas recientes, los intervalos de retorno del fuego se han acortado, el área quemada se incrementó en algunos sitios, y la degradación del hábitat está ocurriendo donde el restablecimiento post-fuego está siendo obstaculizado por pastos anuales invasores. En las décadas por venir, el cambio climático puede acelerar la ocurrencia de esos incendios y sus retroalimentaciones invasivas, aunque el proyectar la dinámica futura de estos incendios requiere de un mejor conocimiento de los factores conducentes de los incendios en el largo plazo a lo largo de la región dominada por la artemisa grande. Integramos acá las observaciones de incendios con datos del clima y vegetación para derivar un modelo estadístico para toda la región de artemisa grande que representa cómo la probabilidad anual de incendios está influenciada por el clima y las características del combustible fino.

Resultados La frecuencia de incendios de vegetación varió significativamente a lo largo de la región de artemisa grande, y nuestro modelo representó mucho de esa variación. La biomasa de los pastos perennes y hierbas, que nosotros usamos como substituto para los combustibles finos, influenciaron la probabilidad de ocurrencia de incendios. La probabilidad de ocurrencia de incendios fue más alta en áreas con alta biomasa de hierbas anuales y pastos, lo que es consistente con el bien documentado fenómeno de un incremento de incendios luego de una invasión de pastos anuales. Los efectos de las anuales en la probabilidad de incendios fueron más fuertes en lugares con veranos secos. La probabilidad de incendios varió con la biomasa de los pastos y hierbas perennes y fue más alta con niveles de biomasa intermedios. El clima, que varía sustancialmente a lo largo de la región de la artemisa grande, fue también predictor de la probabilidad de incendios, y las predicciones fueron más grandes en áreas con una baja proporción de la precipitación recibida durante el verano, con precipitación intermedia y alta temperatura.

Conclusiones Desarrollamos un modelo cuidadosamente validado que contiene relaciones relativamente simples y biológicamente plausibles, con el objetivo de obtener una adecuada performance bajo condiciones noveles, de manera que proyecciones útiles sobre promedios anuales de probabilidad de incendios puedan ser hechas dados los cambios en las condiciones generales. Estudios previos sobre los impactos de la vegetación y clima en la probabilidad de incendios en ecosistemas de artemisa grande han generalmente usado aproximaciones más complejas basados en el aprendizaje automático (machine learning) y han sido generalmente aplicables sólo a porciones de la región dominada por artemisa grande. Por lo tanto, nuestro modelo complementa trabajos existentes y conforma una herramienta adicional para entender futuros incendios de vegetación y la dinámica ecológica a lo largo de la región de artemisa grande.

Background

Long-term shifts in climate, and associated changes in weather and vegetation, are anticipated to affect the frequency, size, and severity of wildfires in many parts of the globe (Abatzoglou and Kolden 2013; Barbero et al. 2015; Parks et al. 2016; Pausas and Keeley 2021). Climate change is expected to increase the prevalence of extreme fire weather (Bowman et al. 2020; Coop et al. 2022), but substantial uncertainty exists around future trends in wildfires that are driven by interactions among climate, vegetation, and fuels (Kloster and Lasslop 2017; Wu et al. 2021). Warming combined with reduced precipitation may, for example, promote fire in tropical rainforests by increasing flammability but reduce fire in arid ecosystems because of fuel limitation. Predicting future changes in wildfire is further complicated by uncertainty about future human population sizes and activities such as fire suppression, conversion to cropland, and deforestation (Knorr et al. 2016; Riley et al. 2019).

More focus has generally been put on understanding drivers of wildfire in forests, woodlands, and savannas than in non-forested drylands such as arid and semiarid shrublands, e.g., sagebrush-dominated ecosystems in the Western US (Crist 2023; Shinneman et al. 2023a, b). However, understanding the drivers of wildfire is critical in sagebrush-dominated ecosystems because fire is a major proximate cause of historical and ongoing losses of sagebrush ecosystems and declines in sagebrush obligate species and the ecosystem services derived from it (Doherty et al. 2022; Remington et al. 2021). Sagebrush ecosystems are being lost or degraded by the combination of annual grass invasion, altered wildfire regimes, conifer expansion, land use change, and climate change (Balch et al. 2013; Remington et al. 2021). Wildfires are a natural part of these ecosystems, and historically, fire return intervals likely varied among sagebrush plant community types, ranging between 35 and 450 years (Baker 2006), but large fires may have been rarer than they are today (Bukowski

and Baker 2013). Anthropogenic ignition was likely common prior to the arrival of European settlers, with indigenous people managing the Great Basin for reasons such as maintaining habitats for animals used as food and driving game during hunts (McAdoo et al. 2013). With both lightning and indigenous sources of ignitions, the landscape likely was a patchwork of areas with lower fuel loads than today, lacking invasive plant infestations (McAdoo et al. 2013). Fire suppression policies and removal of indigenous burning practices led to increased fire return intervals, and subsequent conifer encroachment contributed to higher woody fuel loads. However, in recent decades, fire return intervals have shortened significantly in some areas, and the area burned has increased, largely due to invasion by annual grasses (Baker 2013; Balch et al. 2013; Shinneman et al. n.d.). Subsequent habitat degradation occurs where post-fire re-establishment of sagebrush is hindered by the invasive annual grasses (Coates et al. 2016). Therefore, understanding the factors that drive wildfire in this region and understanding how that may change in the future is of scientific and management interest.

Sagebrush plant communities that retain high ecological integrity and function (e.g., wildlife habitat and ecosystem services) have been defined as those having a sagebrush overstory (primarily big sagebrush, *Artemisia tridentata* Nutt.), ecologically appropriate cover of native perennial grass and forbs in the understory, and low levels of invasive annual grasses and coniferous trees (Doherty et al. 2022). In addition to directly competing with native plants, cheatgrass (*Bromus tectorum* L.), an invasive annual grass, increases the risk of wildfire because it matures early in the growing season and then dries out, forming a continuous fine fuel (Davies and Nafus 2013). Additionally, invasive grasses such as North Africa grass (*Ventanata dubia* [Leers] Cross.) can increase the probability of wildfire due to increased fire spread not only in invaded patches but can facilitate fire spread into adjacent forests (Tortorelli et al. 2023). Big sagebrush does not re-sprout following fire and can be slow to re-establish due to short-distance seed dispersal, low germination rates, and seedling mortality (Schlaepfer et al. 2014). Natural sagebrush regeneration occurs exclusively from seeds in the seed bank and seeds originating from existing plants in unburned islands or from nearby plants outside the fire perimeter (Longland and Bateman 2002; Schlaepfer et al. 2014). In contrast, cheatgrass frequently becomes more abundant post-fire, which can inhibit the re-establishment of sagebrush, native perennial grasses, and native forbs, thereby creating an invasive grass-fire cycle (D'Antonio and Vitousek 1992; Shinneman and Baker 2009). Recurrent fires deplete sagebrush seed in the seed bank, leading to permanent type conversion to an annual grass state in areas where fire return intervals are short (Remington et al. 2021).

A consistent perspective on wildfires across the sagebrush region is essential for informing long-term climate vulnerability and adaptation efforts (Doherty et al. 2022). The prevalence of wildfire varies substantially across the region (Pastick et al. 2021), driven by large variations in climate and vegetation (Remington et al. 2021). Previous studies on the impacts of vegetation and climate on wildfire probability in sagebrush ecosystems have mostly focused on portions of the sagebrush region that are currently more heavily invaded by cheatgrass, such as the Great Basin (Bradley et al. 2018; Pilliod et al. 2017; Smith et al. 2022). While useful for the area they were trained on, such models may be less useful when applied across the entire sagebrush region because of environmental, anthropogenic, climatic, and vegetation differences.

Because long-term projections of ecological dynamics are fraught with uncertainty, simpler statistical models that represent well-understood fundamental relationships may generate more reliable long-term forecasts than more complex, empirically derived predictive algorithms. Recent studies have successfully used machine learning approaches for estimating wildfire probability across landscapes largely dominated by sagebrush (Pastick et al. 2021; Smith et al. 2022). These models are very flexible and can therefore closely fit the observed data, but they also represent complex non-linear relationships and interactions among independent variables that can be difficult to interpret and may create unrealistic predictions when applied under novel conditions, like those projected under climate change. This issue can be especially problematic with machine learning models that are prone to over-fitting (Wenger and Olden 2012). In contrast, simpler models that capture primary ecological relationships can potentially perform well under novel future conditions, even if their relative simplicity causes them to underperform under current conditions (Bell and Schlaepfer 2016), and they may be well suited for integration into process-based ecological simulation models used for assessing long-term ecosystem dynamics in the context of climate change (e.g., Palmquist et al. 2021).

For these reasons, we developed a relatively simple closed-form statistical model for estimating the average wildfire probability in sagebrush ecosystems. Previous work modeling wildfire probability in the sagebrush region has often focused on short time scales (e.g., using variables such as daily fire weather and fuel moisture), due to the strong correlation of area burned with these predictors, and some studies have included multiple ecosystem types instead of just focusing on sagebrush ecosystems (Finney et al. 2011; Short et al. 2020; Smith

et al. 2022; Tortorelli et al. 2023). Although these efforts are essential for informing decisions over the near term, most of them do not currently provide perspectives on potential multi-decadal shifts in wildfire probability that may be induced by climate change, and none yet provides insight at a national scale, though some provide projections for individual areas on the scale of ecoregions (Dye et al. 2023; Riley and Loehman 2016). The computational time required for these analyses can be extensive. Understanding long-term drivers of annual wildfire probability is important for assessing how these ecosystems may change at regional scales and how those changes may inform long-term management decisions.

Our goal was to understand how wildfire probability relates to climate and fuel conditions across the entire sagebrush region. To do this, we developed a statistical model that represents the relationship between annual wildfire probability and a small number of climate and fuel variables. Because we wanted to understand potential fire responses to broad-scale changes in conditions, we related observed fire occurrence to general climatic and average vegetation characteristics. Specifically, we fit a closed-form biologically plausible logistic regression model that captures broad dynamics such that projections of average annual wildfire probability may be made given general changes in conditions. We then examined the sensitivity of modeled fire probabilities to simple shifts in climate and vegetation.

Methods

Study area

Our study area consists of the big sagebrush region (where big sagebrush, *A. tridentata*, is abundant), which is a subset of the global shrubland biome (Whittaker 1975). This region is referred to elsewhere as the sagebrush biome (Jeffries and Finn 2019) and occurs across 13 states in the Western US (Fig. 1). Hereafter, we will refer to our study area as the sagebrush region. For all variables (described below), we acquired data (aggregated to a 1 km × 1 km resolution) for the 782,321 km^2 that comprises this area (we excluded non-sagebrush pixels using the mask described in Doherty et al. 2022). On average, across our study area, the mean annual temperature is 8.8 °C (range across pixels – 2.0–24.2 °C), the mean annual precipitation is 328 mm (54–2282 mm), and the mean proportion of precipitation that falls in summer (June to August) is 0.21 (0.03–0.50) (Fig. 1; Thornton et al. 2020).

Predictor variables

We used two vegetation and three climate variables (Fig. 1) to model annual wildfire probability across sagebrush ecosystems in the Western US. The two vegetation variables were the mean aboveground biomass of annual forbs and grasses (hereafter annual biomass) and the mean aboveground biomass of perennial forbs and grasses (hereafter perennial biomass). We focused on annual and perennial biomass as they determine fine fuel availability in these ecosystems. We acquired this biomass data from version 3 of the Rangeland Analysis Platform (RAP; Jones et al. 2021) over 34 years (1986–2019) and aggregated spatially by calculating the means from the native resolution of ~ 30-m to 1-km pixels to match the resolution of the climate data. The aboveground biomass dataset from RAP is based on a process-based model that estimates annual net primary productivity of annuals and perennials from Landsat normalized difference vegetation index (NDVI) estimates that are collected every 16 days. The model can separately estimate production for annuals and perennials because the measured NDVI for each pixel is disaggregated into separate estimates of NDVI for annuals and perennials based on their fractional cover (from the RAP cover dataset, which is trained on plot-level cover observations) and phenology (Robinson et al. 2019).

Our three climate predictor variables were mean temperature, annual precipitation, and proportion summer precipitation (PSP), which we defined as the proportion of annual precipitation that falls in June through August (Fig. 1). PSP ranges from 0 to 1, where a value of 0 would mean that no precipitation falls in June through August (interestingly, PSP had a trimodal distribution across our study area, Fig. 1c). The three climate indices were calculated using the Daymet daily weather dataset (1-km resolution; Thornton et al. 2020).

For both climate and vegetation variables, we calculated 3-year running averages, by taking the mean of values from the current year and previous 2 years. For example, to calculate precipitation for 2015, we calculated the mean of annual precipitation from 2013, 2014, and 2015. We calculated these 3-year averages for 1988–2019 using climate and vegetation data from 1986 to 2019. Therefore, we had 32 climate and vegetation observations per pixel, resulting in a final dataset that contained 25,034,272 observations (782,321 pixels × 32 years). We chose not to use a single value of predictor variables for each pixel (i.e., means across the entire study period) for two reasons. First is that wildfire influences vegetation, and in sagebrush ecosystems, this is most apparent with the positive feedback between cheatgrass (and other invasive annual grasses) and wildfire. Abundance of annuals generally increases after wildfire (Smith et al. 2023), and we wanted the model to correctly incorporate the effect of annuals on wildfire, while not conflating it with the effect of wildfire on annuals. Second, the values of some of our predictor variables have changed over time, regardless of

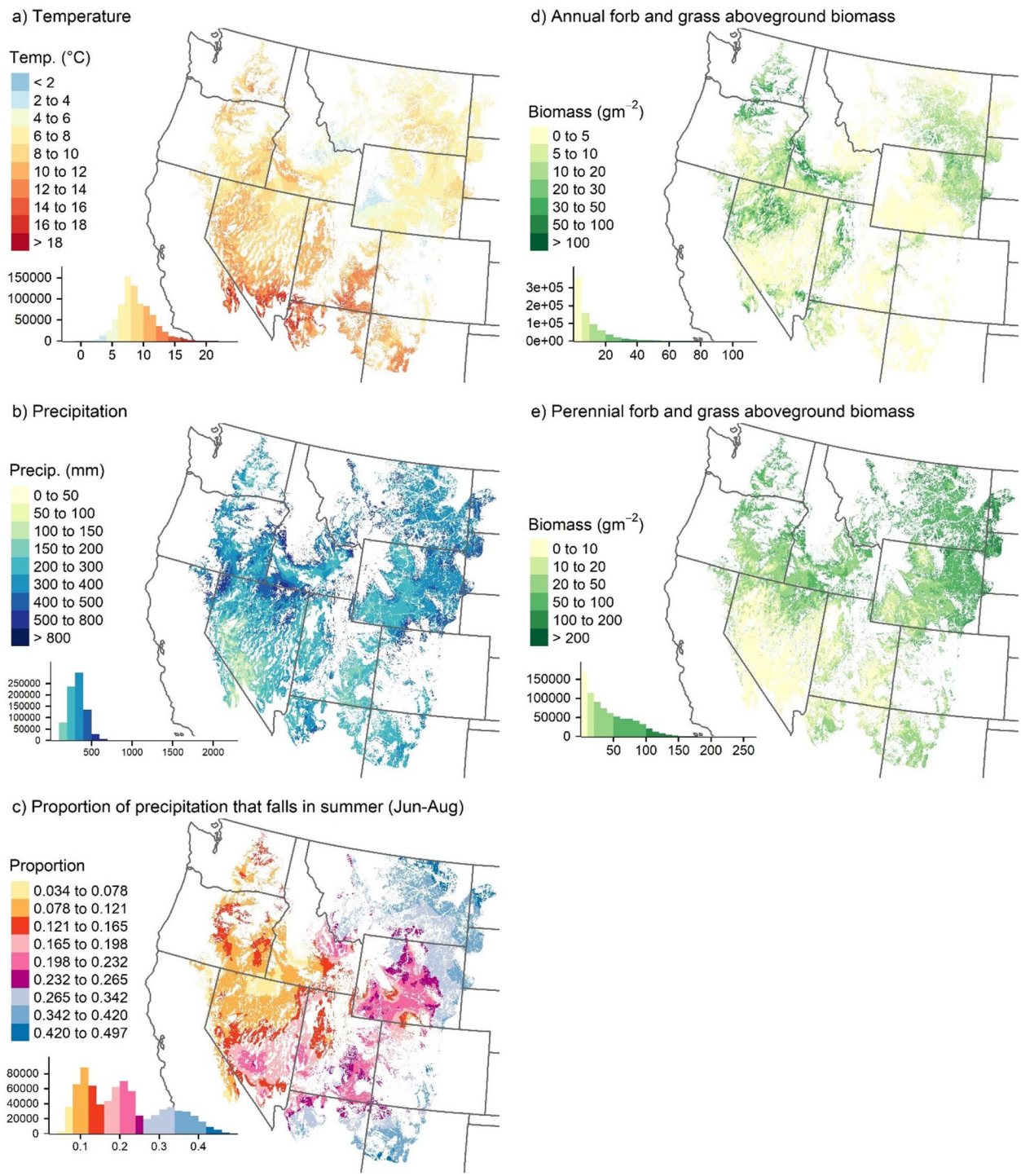

Fig. 1 Maps of **a** mean annual temperature, **b** mean annual precipitation, and **c** mean proportion of precipitation that is received in summer (June to August) and the mean aboveground biomass of **d** annual forbs and grasses and **e** perennial forbs and grasses. Data was aggregated to 1-km resolution and masked to the extent of the sagebrush region (as in Doherty et. al. 2022). Histogram insets present the distribution of values shown on the maps. The means were calculated using climate and vegetation data from 1986 to 2019. Three-year running averages (instead of averages over the entire study period) of these variables were used as predictors in our model

the occurrence of wildfire. For example, the abundance of annuals has increased over the last decades even in unburned locations (Appendix S1; Smith et al. 2023), and similarly, temperatures have been increasing across the region. Additionally, the abundance of annuals fluctuates substantially through time (Additional file 1: Appendix 1; Dahal et al. 2022), largely driven by fluctuations in precipitation (Pilliod et al. 2017). We chose to use 3-year averages to capture general antecedent climate and vegetation conditions that contribute to fine-fuel availability in a given year. This approach was also informed by previous work which found that vegetation and weather conditions from the current year, the previous year, and 2 years previous were predictive of wildfire probability in the Great Basin (Smith et al. 2022).

Response variable

We used the US Geological Survey (USGS) combined wildland fire dataset, which combines fire perimeter data from many sources, including the National Interagency Fire Center and Monitoring Trends in Burn Severity (Welty and Jeffries 2021). We used fire perimeter data from wildfires in the sagebrush region from 1988 to 2019, filtering out prescribed fires. While high-quality fire perimeter data is available back to 1984, annual RAP biomass data was available starting in 1986, which only allowed us to calculate 3-year running averages of biomass for fires starting in 1988. The response variable we used for modeling (described below) was fire occurrence (i.e., 0 or 1 fire) in a given pixel and year. Figure 2a shows the total number of times each pixel burned from 1988 to 2019. We considered a 1-km pixel to have been burned in a given year if > 47% of the pixel area was burned. We chose the 47% threshold, because by using that threshold, the total area of 1-km pixels classified as "burned" most closely matched the true total burned area. This approach limited the bias of predicted average fire probability made by our model.

Model fitting

We fit a logistic regression model relating fire occurrence to climate and vegetation. The response variable was observed wildfire occurrence in a given pixel in a year (i.e., 0 or 1 fire). Logistic regression allowed for the estimation of the annual probability of wildfire. We used a logit link function. Our model implicitly incorporated spatial and temporal effects of vegetation and climate on wildfire probability because we used data from across the sagebrush region (782,321 pixels) over 32 years. For simplicity and because of computational constraints, we did not account for spatial or temporal autocorrelation. By ignoring this autocorrelation, we likely underestimated uncertainty in our parameter estimates. However, given the very large sample size, uncertainty from sampling variability is very low (additionally, because we are using values from all pixels in the sagebrush region, our dataset could also be viewed as a complete census). Therefore, other factors are more likely to cause uncertainty in our parameter estimates, such as error in remotely sensed biomass estimates, collinearity among predictor variables making individual parameter estimates unstable, not correctly representing the true data generating process (e.g., missing variables and interactions, incorrect variable transformations), and other limitations that we highlight

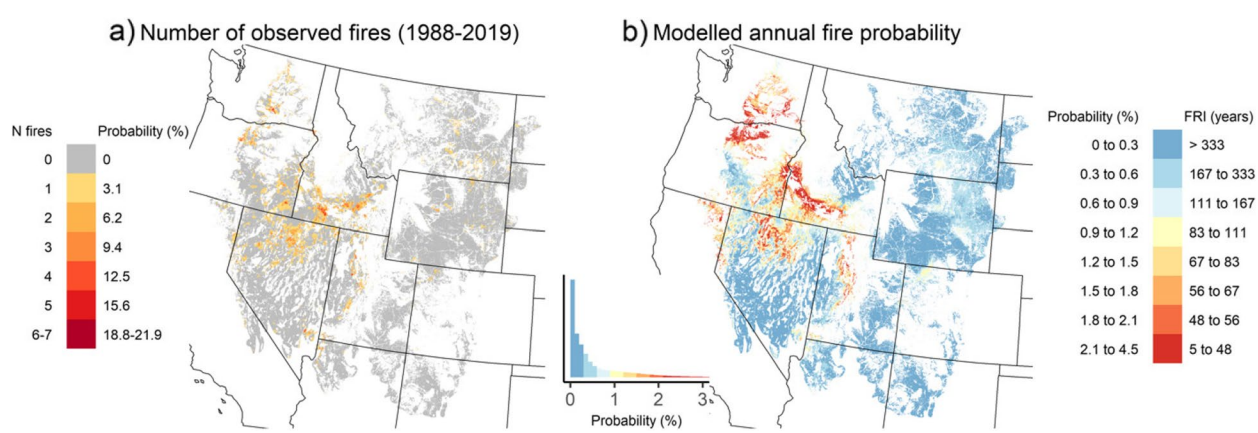

Fig. 2 **a** The number of years in which each pixel burned from 1988 to 2019 (USGS combined wildland fire dataset). The corresponding observed annual fire probability calculated from those fire frequencies is also shown in the legend. **b** The mean annual wildfire probability predicted by our model, based on vegetation and climate conditions. These values represent the mean probability (%) of fire occurring in a given year, and the corresponding fire return intervals (FRI) are also shown in the legend. The Histogram inset shows the distribution of values shown on the map (*x*-axis limits were restricted, 0.9% of data not shown). To more directly compare the average observed and modeled wildfire probability, see Figs. 4 and 5

below in the "Discussion" section. To address some of these concerns, we evaluated the bias in the model across vegetation and climatic variables and their interactions (see the "Visualizing model fit" section), conducted automated selection of predictor variable transformations (described below), assessed whether relationships represented by the model were biologically reasonable, conducted cross-validation (see the "Cross-validation" section), and calculated bias in the final model for five regions within our study area (Additional file 1: Table S2).

Because many of the relationships between wildfire probability and predictor variables were non-linear, we tested applying different functions to the predictor variables to transform them. First, we applied transformations that changed the variable without adding additional terms, these were as follows: x (i.e., no change), $\sqrt{x}$, and $\log_{10}(x)$. If we applied one of these transformations to a variable in the model, we also applied it to that variable where it appeared in an interaction term. Secondly, we also applied a second-order polynomial transformation to each of these already transformed main terms, so in total, we tested 6 possible functions for transforming a given variable: x (i.e., no change), $\sqrt{x}$, $\log_{10}(x)$, $x+x^2$, $\sqrt{x}+x$, $\log_{10}(x)+(\log_{10}(x))^2$. Additionally, before applying a $\log_{10}$ transform, we added 1 to the value so that the transformed variable would not be undefined when the original value was 0, and values of the transformed variable would all have the same sign (e.g., otherwise a change in sign of $\log_{10}$ biomass would occur when going from below to above 1 g/m^2). An exception was PSP (which is constrained between 0 and 1), in which case we added 0.001 prior to the $\log_{10}$ transformation. We did not have a priori hypotheses of the shapes of the relationships, so we chose simple functions that would allow the model to represent plausible non-linear relationships that are monotonically increasing or decreasing or that are parabolic. We used a multi-step approach to find the overall "best" logistic regression model. First, we individually transformed each of the five predictor variables using each of the 6 functions and compared Akaike's Information Criterion (AIC) of the models (5 variables × 6 transformations = 30 models). The best model (lowest AIC) was then selected; this was the model with the best single transformation. In step 2, we repeated the process again for the remaining untransformed variables and chose the best resulting model (4 variables × 6 transformations = 24 models; these models have 2 variables transformed). We repeated this until all variables were transformed or until further transformations no longer substantially improved the model (ΔAIC < 10).

To determine what interactions to include in the model, we visualized all two-way interactions between variables. The best model included no interactions and over-predicted wildfire probability in wet (high annual precipitation) areas with high biomass of annuals but underpredicted in dry (low annual precipitation) areas with high biomass of annuals. Lastly, we repeated the iterative steps of finding the best transformations of the main effects of the five variables, as described above, but this time also included this interaction in the model. We considered adding shrub cover (from the RAP dataset) as an additional predictor variable but did not include it in the final model because it increased model complexity without improving model fit substantially. Similarly, Smith et al. (2022) found that in comparison with herbaceous fuels, shrub cover was a relatively unimportant predictor of wildfire probability in Great Basin rangelands. For computational reasons, we conducted the iterative steps to find the best transformations using a random subset of 5 million observations, before then fitting the model with those transforms to the entire dataset. We compiled data using Google Earth Engine (Gorelick et al. 2017), and statistical analyses were done using R version 4.2.3 (R Core Team 2023). Models were fit using the "glm" function (stats package).

Visualizing model fit

We used partial dependence plots to visualize the relationships represented by the model. However, a limitation of this approach is that if predictor variables are correlated or interactions are present, which is typical for climate and vegetation variables, the relationships do not fully represent the true relationships one would observe in the underlying data (Biecek and Burzykowski 2021). Therefore, we also constructed a "quantile" plot to visualize observed and predicted wildfire probability across each percentile of a predictor variable. To achieve this, predictor data were binned by percentile (i.e., 100 bins), and the mean observed and predicted (modeled) wildfire probabilities were calculated across observations belonging to each bin. Each point shown in the quantile plots is an average of ~ 250,000 observations (1% of the data). The quantile plots also allow for comparison between average observed and predicted wildfire probabilities across the range of a given predictor variable. Observed wildfire frequency (hereafter "observed wildfire probability") was calculated by dividing the number of fire occurrences by the total number of observations.

We created additional "filtered quantile" plots, to assess how relationships between wildfire probability and biomass varied with climate. First, observations were only kept if they fell within the two lowest or highest deciles of a given climate variable (i.e., below the 20th percentile or above the 80th percentile), and then "quantile" plots showing the relationships between wildfire probability and biomass were constructed from this filtered dataset.

This allowed for the comparison of, for example, the relationship between annual biomass and wildfire probability in areas with high PSP versus areas with low PSP. Note that the model predicts annual wildfire probability (which has a range from 0 to 1), but for the sake of readability in figures, we present annual wildfire probability as a percentage (i.e., % wildfire probability per year). In the figures, we also present secondary axes with wildfire probability converted to fire return interval (the mean number of years between fire events).

Cross-validation

We conducted cross-validation to better understand model performance using an environmental blocking approach (blockCV package, Valavi et al. 2019). We used this approach because, due to the underlying spatial autocorrelation in the data, using a random set of pixels as the test dataset would underestimate our out-of-sample prediction error (Roberts et al. 2017). Using three climate variables (mean annual temperature, mean annual precipitation, and mean PSP), we grouped each cell into one of five blocks or "folds" (Additional file 1: Fig. S2). The folds represent regions that are somewhat climatically distinct and therefore can help validate model performance under new climatic conditions (Valavi et al. 2019). Folds were identified using *k*means, which is an unsupervised clustering algorithm (Hartigan and Wong 1979). The data associated with the pixels in a given fold were used as a test dataset, and the data from the remaining folds were used as a training dataset. As a result, we had five training datasets that we fit models to and then created predictions for the five respective test datasets. Each model we fit had the same variable transformations (i.e., the transformations of the best model fit to the entire dataset).

Model sensitivity to changes climate and vegetation

As the final step, we conducted a sensitivity analysis to understand how model predictions would be affected by simple changes in climate and vegetation variables. The goal was not, for example, to create projections of wildfire probability in response to a real climate scenario, but instead to get a general sense of how sensitive the model is to simple changes in predictor variables and to what degree responses vary. To achieve this, we calculated predicted wildfire probability across the study area in response to six different changes of a single climate variable and four different changes of a single vegetation variable. The changes were 2 °C and 5 °C increases in temperature and 20% decreases and 20% increases in precipitation, PSP, annual biomass, and perennial biomass. We also calculated the expected annual area burned for each predictor variable perturbation by multiplying the predicted fire probability of each pixel by the area of the pixel and then summing across pixels. For this sensitivity analysis, we changed predictor variables individually, but we acknowledge that vegetation is expected to change with climate, potentially in complex ways.

Results

Model description

Across the study area, the observed mean annual wildfire frequency was 0.50% (a 200-year fire return interval), and individual pixels experienced between zero and seven fires from 1988 to 2019 (Fig. 2a). The observed mean annual burned area across the study area from 1988 to 2019 was 3941 km^2 (calculated as the area of 30 m × 30 m pixels whose centroid fell within fire perimeters) while the mean expected annual burned area based on the modeled wildfire probabilities was 3938 km^2. Spatial patterns of observed and predicted annual fire probabilities agreed quite well, with the higher fire probabilities occurring in the northern Great Basin and lower probabilities toward the east (Fig. 2). Partial dependence plots illustrated that the predicted wildfire probability was greatest for high annual biomass (> 75 g/m^2), intermediate perennial biomass (46 g/m^2), high temperature (> 15 °C), intermediate precipitation (487 mm), and a low PSP (0.06) (Fig. 3).

We found a modestly sloped positive relationship between predicted wildfire probability and temperature. A strong negative relationship between wildfire probability and PSP was evident (except for a positive relationship at the lowest few percentiles of PSP), which reflects the fact that the eastern portion of the sagebrush region receives a large proportion of precipitation in summer (Fig. 1) and has fewer fires (Fig. 2) than the Great Basin to the west. We observed a positive, saturating relationship between annuals and wildfire probability (Fig. 3d). Predicted wildfire probability was most sensitive to changes in annuals (i.e., steepest slope) when PSP was low (dry summers) and least sensitive when PSP was high (wet summers) (Fig. 3d). Across the ranges of temperature, precipitation, PSP, and perennials, the predicted wildfire probability was low if biomass of annuals was low (20th percentile) (Fig. 3a–c, e).

The shapes of predicted wildfire probability shown in the "quantile" plots (Fig. 4) differ somewhat from those represented by the partial dependence plots (Fig. 3), and this is presumably because of the non-independence of climate and biomass predictor variables as well as interactions between variables. The partial dependence plots (Fig. 3) illustrate how the mean model predictions shift when the value of a given variable is changed, while other variables remain unchanged. The quantile plots are different, in that they allow for the comparison of wildfire probability between actual areas that, for example, have high (e.g.,

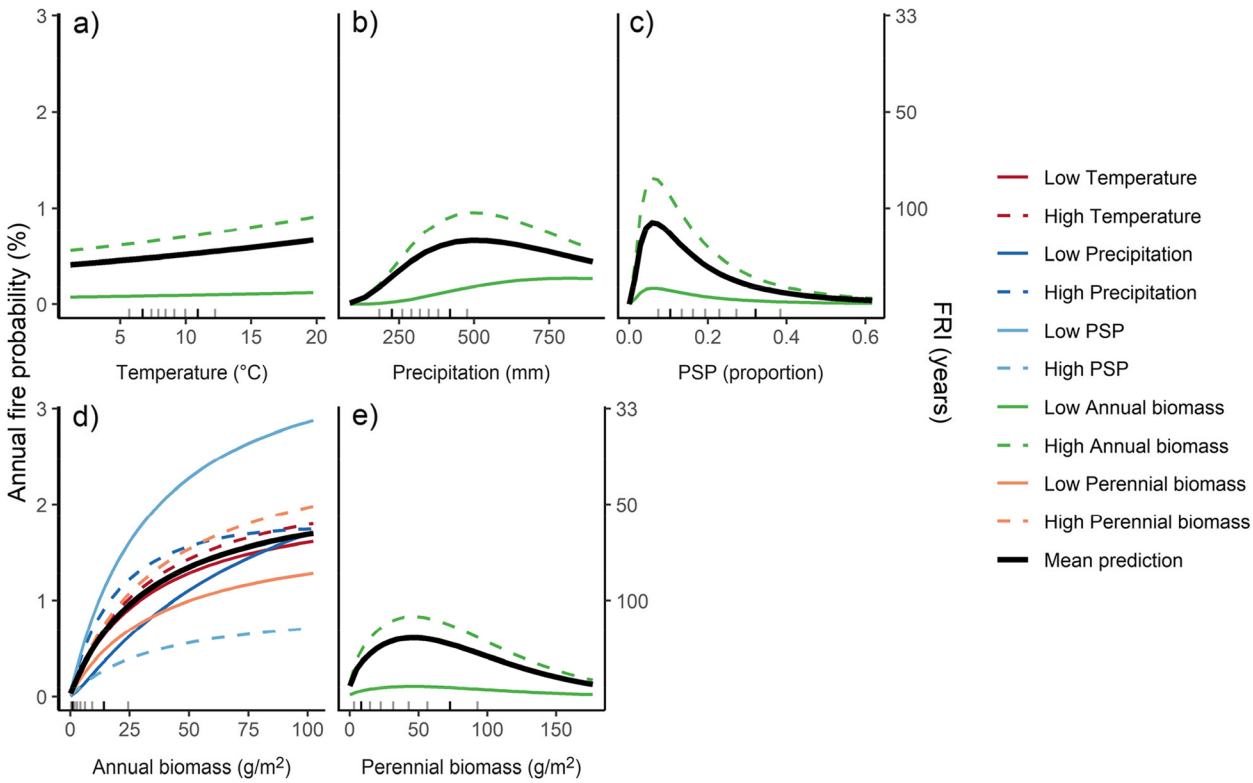

Fig. 3 Partial dependence plots depicting the effect of the five predictor variables on modeled annual wildfire probability. The primary (left) *y*-axis shows the mean predicted wildfire probability for a fixed level of a given predictor variable, and the secondary (right) *y*-axis shows the corresponding fire return interval (FRI). The black line shows the mean predicted fire probability across all combinations of values of the other predictor variables. The colored dashed (solid) lines show the mean predicted wildfire probability when one of the other predictor variables is held at its 20th (80th) percentile. The 9 tick marks above the *x*-axis show the 10th to 90th percentiles (in increments of 10); the darker tick marks are the 20th and 80th percentiles. The *x*-axes show **a** mean temperature, **b** annual precipitation, **c** proportion summer precipitation, **d** aboveground biomass of annual forbs and grasses, and **e** aboveground biomass of perennial forbs and grasses

95th percentile) versus low (e.g., 5th percentile) biomass of annuals. Because of non-independence between climate and vegetation variables, such areas will also differ in other ways, for example, areas with few annuals tend to be cooler than those with abundant annuals (Fig. 5a).

There was high agreement in quantile plots between the average observed and predicted wildfire probability across the ranges of all predictor variables (Fig. 4). Both observed and predicted wildfire probability increased with annual biomass, decreased with PSP, and was highest at intermediate levels of perennial biomass, temperature, and precipitation (Fig. 4). Notably, relationships between wildfire probability and annual and perennial biomass varied with climate (Fig. 5). The relationship between annual biomass and wildfire probability was much stronger in areas that have dry summers (< 20 percentile PSP) compared to areas with wet summers (> 80th percentile PSP) (Fig. 5e). Additionally, for a given amount of annual biomass, the mean observed and predicted wildfire probability tended to be higher in areas with high (> 80th percentile) annual precipitation (Fig. 5e).

The equation of our final model (Eq. 1) is not supported by data outside of the range of the data we used for model fitting (Additional file 1: Appendix 2). For instance, the maximum value of annual biomass in our dataset was 190 g/m^2, and predictions of fire probability at higher biomass values are not supported. Most variables in the model have coefficients such that very high values of the variable would cause wildfire probability to approach zero (i.e., downward facing parabolic shapes; Fig. 3).

$$\begin{aligned} y = &-92.86 + 14.01\log(AFG+1) - 0.8117\log(AFG+1)^2 + 0.5054PFG^{0.5} - 0.03730PFG \\ &+0.02672T + 49.25\log(P+1) - 8.236\log(P+1)^2 \\ &-7.505\log(PSP+0.001) - 3.118\log(PSP+0.001)^2 \\ &-4.047\log(AFG+1) * \log(P+1) \end{aligned} \quad (1)$$

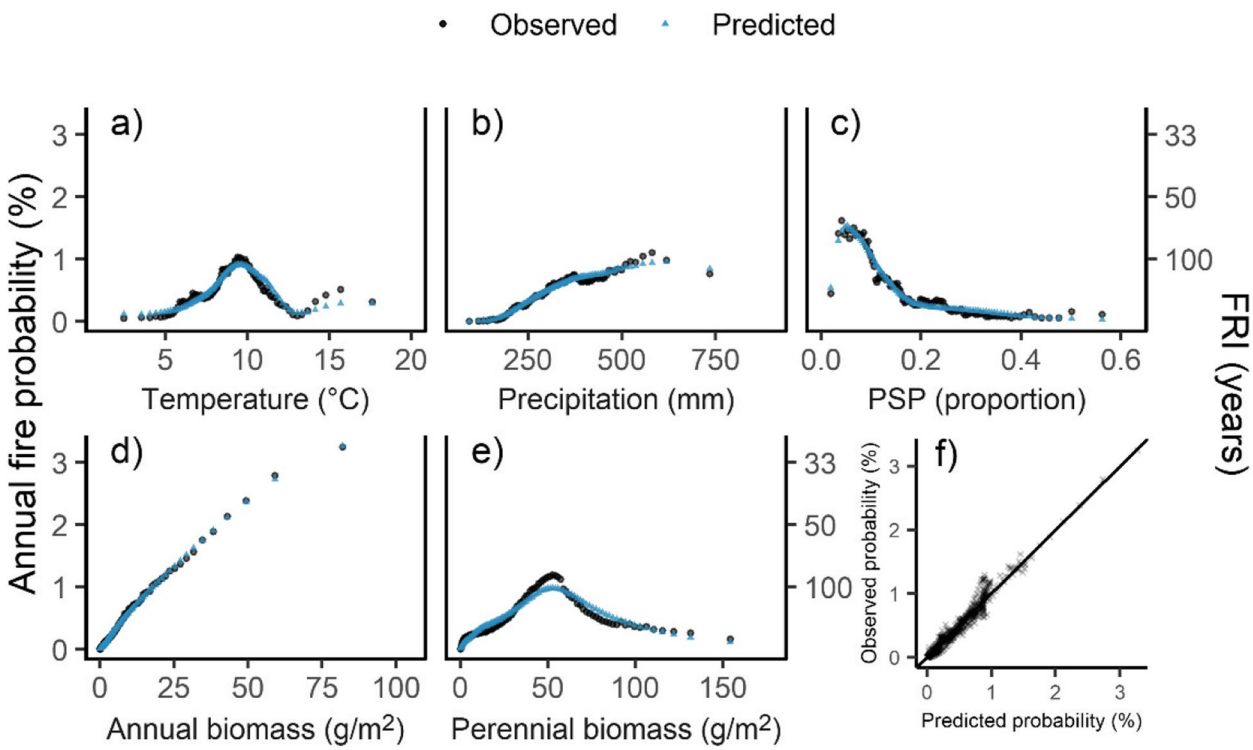

Fig. 4 Comparison of mean observed and predicted annual wildfire probability. Panels show the mean observed (black circles) and predicted (blue triangles) annual wildfire probability for each percentile of **a** mean temperature, **b** annual precipitation, **c** proportion summer precipitation, **d** aboveground biomass of annual forbs and grasses, and **e** aboveground biomass of perennial forbs and grasses. Corresponding fire return intervals (FRI) are shown on the secondary (right) *y*-axis. Data were binned by percentile (i.e., 100 bins) of a given predictor variable, and the *x*-axis (**a**–**e**) shows the mean value of each percentile of that variable. To illustrate, the right-most black circle in **d** shows the mean observed wildfire probability across all pixels where biomass of annuals is between the 99th and 100th percentile, and the rightmost blue triangle in that panel shows the mean predicted wildfire probability for those same pixels. Each point represents the mean of ~ 250,000 observations (i.e., 1% of the entire dataset). In **f**, the mean observed and predicted annual wildfire probability values shown in **a**–**e** are plotted against each other (1:1 line shown for reference)

where

$$y = \ln\left(\frac{p}{1-p}\right)$$

Or if written in terms of p:

$$p = \frac{1}{1+e^{-y}}$$

Abbreviations

log = base 10 logarithm

p = annual wildfire probability

AFG = annual forb and grass aboveground biomass (g/m^2)*

PFG = perennial forb and grass aboveground biomass (g/m^2)*

T = mean temperature (K)*

P = annual precipitation (mm)*

PSP = proportion summer precipitation (proportion of annual precipitation received in June to August)*

*These variables were calculated as 3-year running averages (mean of the current and previous 2 years)

Cross-validation

Cross-validation results suggested that overall, the fit of our final model was fairly robust (Additional file 1: Appendix 3). Partial dependence plots showed that most of the relationships represented in our model remained stable when five different training datasets, each representing ~ 80% of the entire dataset, were used (Additional file 1: Fig. S3). These five models also reproduced observed patterns in wildfire probability quite well for associated test datasets (Additional file 1: Appendix 3). The slope of the relationship between predicted wildfire probability and annual biomass varied the most between the five cross-validation models, suggesting it was the term in the model with the greatest uncertainty. For example, when we withheld the northeast most portion of the study area from the dataset, which has less frequent wildfire than much of the rest of the region, the relationship between annual biomass and wildfire probability was stronger (Additional file 1: Fig. S3), and, consequently, that model then substantially over-predicted mean

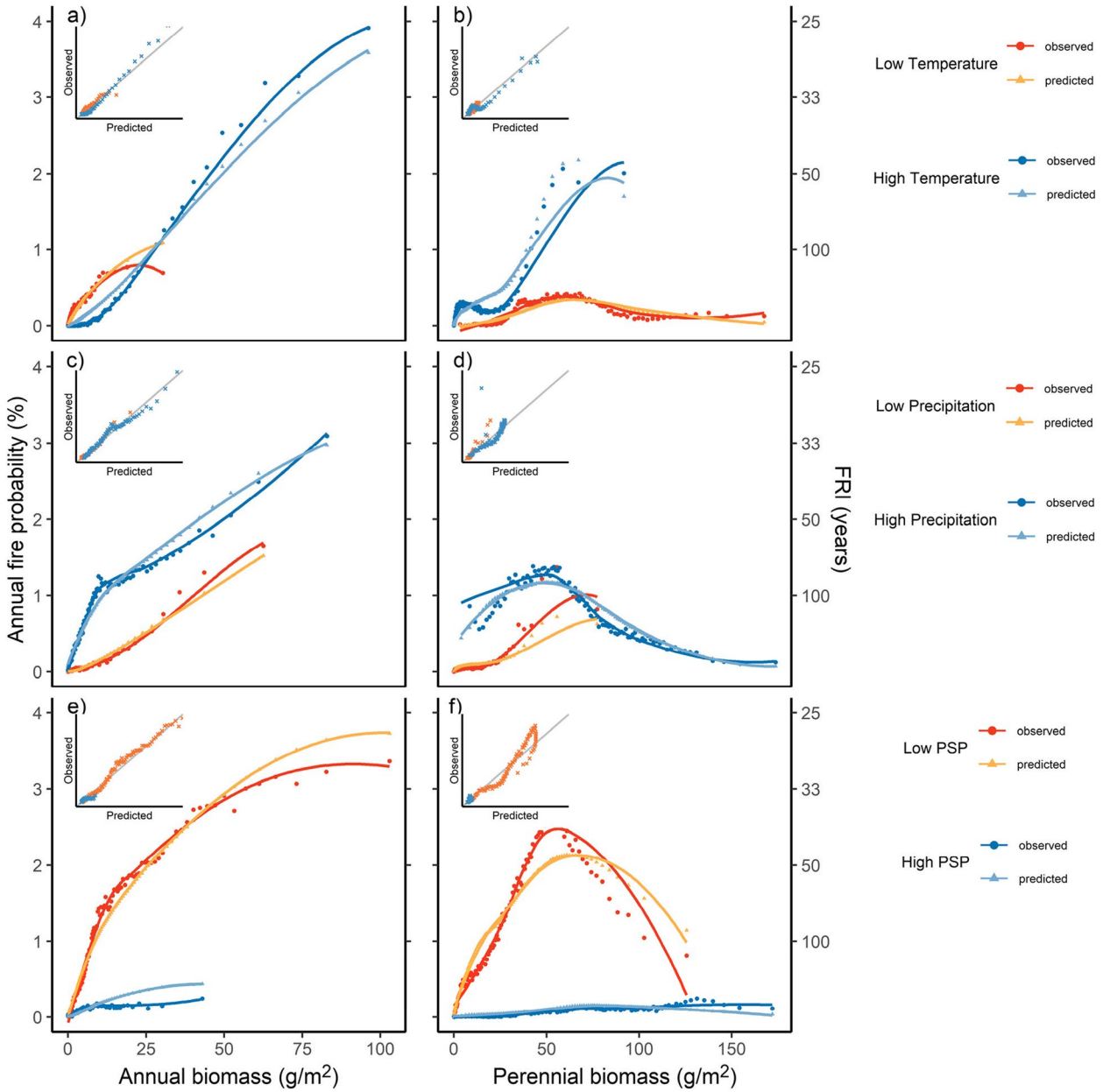

Fig. 5 Assessment of model performance at extreme climate conditions, showing low (blue) and high (red) values of the three climate predictors. Main panels show observed (circles) and predicted (triangles) mean annual wildfire probability across a range of aboveground biomass of (**a**, **c**, **e**) annual grasses and forbs and (**b**, **d**, **f**) perennial grasses and forbs. Corresponding fire return intervals (FRI) are shown on the secondary (right) *y*-axis. Data are shown separately for pixels having low (< 20th percentile; reds) or high (> 80th percentile; blues) levels of **a**, **b** mean temperature; **c**, **d** annual precipitation; and **e**, **f** proportion summer precipitation (PSP). Biomass values were binned by percentile, and the mean observed and predicted annual wildfire probability is shown for each percentile of biomass. To illustrate, the right-most red circle in panel a shows the mean observed wildfire probability across pixels with low precipitation (< 20th percentile) where biomass of annuals is between the 99th and 100th percentile, and the right-most yellow triangle in that panel shows the mean predicted wildfire probability for those same pixels. Each point on the figure represents the mean of ~ 50,000 observations. Best fit lines in the main panels were generated using locally estimated scatterplot smoothing and are included to help visualize the trends in the data. In the insets, the mean observed and predicted annual wildfire probability (%) values shown in the main panels are plotted against each other (1:1 line shown for reference), with colors representing data from areas with low (red) and high (blue) levels of the respective climate variable. The 20th (low) and 80th (high) percentiles of the climate variables were 6.7 °C and 10.9 °C temperature, 226 mm and 420 mm precipitation, and 0.104 and 0.321 PSP, respectively

wildfire probability in the northeast (Additional file 1: Table S2). The relationship with temperature also varied between models fit to the five training datasets (Additional file 1: Fig. S3).

Model sensitivity to changes in climate and vegetation

The model generated plausible estimates of wildfire probability under altered climate and vegetation conditions (Fig. 6). Overall, changes in predicted wildfire probability (both increases and decreases) were most common in the northern Great Basin, where wildfire probability is currently high. The model was more sensitive to changes in the biomass of annuals than perennials. A 20% increase in annuals caused an 11% increase in expected burned area, with wildfire probability increases occurring almost everywhere (Fig. 6h), but increases were very small in areas that currently have low wildfire probability (Additional file 1: Appendix 4). By comparison, a 20% increase in perennials caused a 3% decrease in the expected burned area, which reflects increases in predicted wildfire probability in areas that currently have few perennials and decreases where they are currently more abundant. Warmer temperatures caused consistent, but fairly small increases in predicted wildfire probability across the study area, with 5 °C warming translating to a 14% increase in the expected annual burned area. Increasing annual precipitation by 20% generally increased wildfire probability, except for in a few very wet locations where decreases were predicted (11% increase in expected burned area overall). Lowering PSP (i.e., drier summers) increased predicted wildfire probability (13% increase in expected burned area), while increased PSP reduced wildfire probability (12% decrease in expected burned area).

Discussion

Model overview

We built a closed-form statistical model under the principles of parsimony, which was able to successfully reproduce much of the substantial variation in wildfire patterns across the sagebrush region. The model includes only five variables: three that represent key aspects of climate and two vegetation variables that represent the availability of fine fuels. A benefit of our model is that it represents fairly simple ecologically plausible relationships (Fig. 3) and therefore may be more likely to perform reasonably under novel conditions (Bell and Schlaepfer 2016). Unlike some previous efforts, our model was fit using data from across the entire sagebrush region and complements existing, more complex regional models. Wildfire frequency varies substantially across the sagebrush region which spans a wide range in climate and fuels conditions and consists of plant communities dominated by one of several big sagebrush subspecies that form both sagebrush semidesert shrubland and sagebrush-steppe ecosystems. Our model may prove useful for those trying to understand the general effects that changes in climate and vegetation have on mean annual wildfire probabilities across this region. Because the model forms relatively straightforward links to underlying drivers of wildfire probability, it should be well suited for integration into ecological simulation models used for assessing long-term ecosystem dynamics in the context of climate change (e.g., Palmquist et al. 2021).

Our results suggest that both annual and perennial grass and forb biomass influence wildfire probability, although the shape of those relationships differ. Observed wildfire probability was highest in areas with high annual biomass. Cheatgrass, an invasive species, is the dominant annual grass in sagebrush ecosystems, and the increase in wildfire probability with annual biomass is consistent with the well-documented phenomenon of increased wildfire following cheatgrass invasion (Bradley et al. 2018; Pastick et al. 2021; Smith et al. 2022). Our model suggests that wildfire probability increases most sharply when the abundance of annuals goes from low to moderate implying that the ability of annuals to carry fire increases quickly, even before they become highly abundant. These results are similar to Bradley et al. (2018) who found that areas with even fairly small amounts of cheatgrass were associated with increased wildfire frequency. Both our model (Fig. 3) and the observational trends (Fig. 5) indicate that when there are few annuals, wildfire probability is low, regardless of the climatic conditions. This is consistent with the notion that wildfire "needs" annual grasses in this region (Smith et al. 2023). However, it is

(See figure on next page.)

Fig. 6 Evaluation of model sensitivity to key climate and vegetation variables. The panels show the distribution, across pixels, of the change in the predicted number of times a location will burn per 100 years in response to **a** 2 °C and **b** 5 °C increases in the mean temperature and 20% decreases and 20% increases in **c**, **d** annual precipitation; **e**, **f** proportion summer precipitation; **g**, **h** biomass of annuals; and **i**, **j** biomass of perennials. The dotted lines show the minimum and maximum changes. Individual values shown in the histogram are based on the mean change in predicted wildfire probability across years at a pixel in response to the change in a given predictor variable. The numbers on the panels present the mean change in the expected burned area per year across the study area, relative to the annual burned area predicted under observed (ambient) conditions. We would like to underscore that these are examinations of model sensitivity and are not climate change projections. Ongoing work will integrate this model into a plant community model that will simulate climate and vegetation changes under climate change

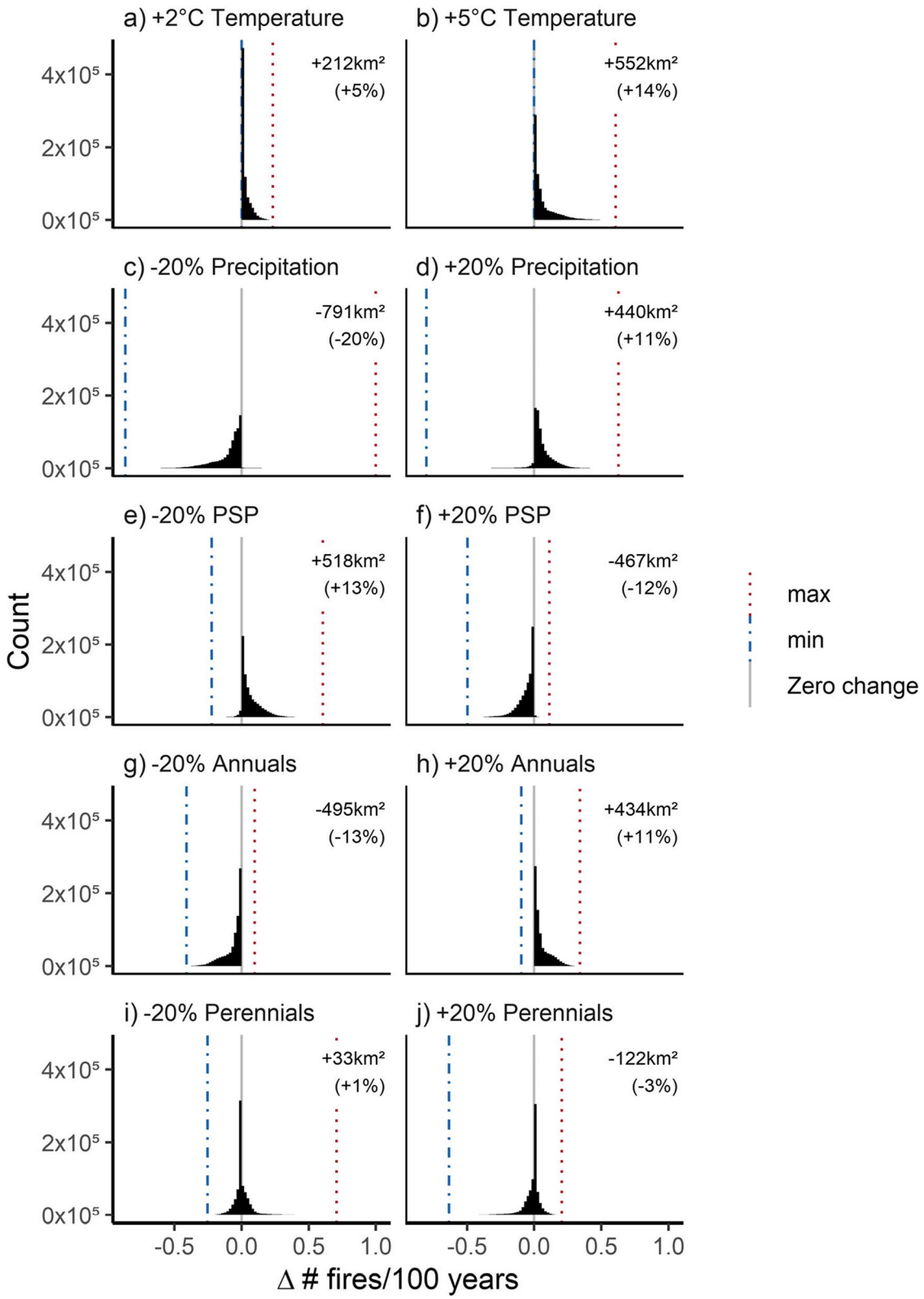

Fig. 6 (See legend on previous page.)

important to note that our cross-validation results indicate uncertainty in the magnitude of the effect of annual biomass on modeled wildfire probability. Additionally, the spatial dataset we used did not distinguish between native and invasive annuals, and while cheatgrass and other invasive annual grasses do generally represent the majority of the annual herbaceous plant biomass in this system (Dahal et al. 2022), native forbs are also an important component of the plant community that do not have the same effect on the wildfire as cheatgrass.

In contrast to annuals, our model indicates that wildfire probability peaks at intermediate levels of perennial grass and forb biomass. Smith et al. (2022) found that wildfire probability peaked at a perennial biomass (~300 g/m^2) that was higher than suggested by our model. This difference could be because their model was trained using vegetation data at a higher spatial and temporal resolution than ours, which resulted in a wider range of biomass values, and their study area was restricted to the Great Basin, which receives less summer precipitation than other parts of the sagebrush region that we included.

In addition to annual and perennial herbaceous biomass, we identified three climatic variables, which were related to wildfire probability: mean temperature, annual precipitation, and proportion summer precipitation (PSP), which we defined as the mean proportion of precipitation that falls in June through August. These variables were meant to capture recent climatic conditions and were calculated as a 3-year running average (e.g., the mean of annual precipitation in the current and preceding 2 years). PSP helped distinguish the fire regimes of the Great Basin in the west and the Great Plains in the east. Wildfire probability was much higher in areas with a low PSP. Much of the sagebrush region has a cool season-dominated precipitation regime, but the north-eastern portion (western Great Plains, northeastern Wyoming) of the region, which has the highest perennial biomass, has a more summer-dominated precipitation regime (Fig. 1, Additional file 1: Table S2). Despite the additional available fine fuels in this area, the wetter summers may be the cause of the lower burn frequency we observed (Additional file 1: Table S2). The abundance of grasses relative to shrubs is higher in summer-dominated precipitation regimes (Paruelo and Lauenroth 1996; Renne et al. 2019), and the pulses of water throughout the summer wets vegetation directly and allows grasses to retain greenness longer, thereby likely maintaining fine fuel moisture later in the growing season.

The frequency of observed wildfires varied with temperature and precipitation, with the highest wildfire frequency occurring at intermediate levels of both temperature and precipitation. High wildfire probability at intermediate precipitation may occur because grass and forb growth (and thereby fine fuels) increases with precipitation, but in general, fuel moisture is higher when there is abundant precipitation (Flannigan et al. 2016), creating a trade-off between fuel quantity and fuel moisture and flammability. Similarly, more wildfires occurred in areas with intermediate temperature, which may also reflect a trade-off between fuel moisture and quantity, where under the hottest conditions, fuels will dry easily, but plant growth and thereby fuels are more limited. However, in our model, predicted wildfire probability increased with temperature across the entire range of temperature, suggesting that after accounting for the availability of fine fuels, warmer conditions increase wildfire probability.

Because we used running 3-year averages of climate and vegetation data to fit our model, it is more useful for understanding the potential impacts of general shifts in climate and vegetation, rather than short-term changes such as the effect an exceptionally warm spring might have on wildfire probability that summer. Using average climate and vegetation data is more computationally tractable, and the data are more readily available; however, using climate variables tends to produce weaker correlations with burned area than daily or monthly weather metrics (Riley et al. 2013). In this way, our model complements previous models—often more complex ones based on machine learning—that incorporate more near-term antecedent conditions such as, for example, total precipitation the preceding season or month (Abatzoglou and Kolden 2013; Pastick et al. 2021; Smith et al. 2022). Such models have sometimes been created with the explicit goal of helping predict fire risk in the upcoming year (Maestas et al. 2022; Smith et al. 2022).

Despite the different modeling approaches we used, the spatial patterns in wildfire probability we predict are broadly similar to a previous statistical model of wildfire probability that was fit to a large portion of the sagebrush region ($r = 0.80$; Pastick et al. 2021) as well as to burn probabilities for the sagebrush region developed via Monte Carlo simulation ($r = 0.68$; Short et al. 2023; Additional file 1: Appendix 5). In line with the observational data, our model predicts the highest wildfire probability in the northern Great Basin (northern Nevada and southern Oregon and Idaho) and lower wildfire probability in the southern and eastern portions of the sagebrush region (Fig. 2). Across the Great Basin, our results also correlate well ($r = 0.82$) with the modeled estimates from Smith et al. (2022) (Additional file 1: Appendix 5).

Model sensitivity to climate and vegetation

To ensure that our model is appropriate for use in studies on the impacts of shifts in climate and vegetation, we evaluated the sensitivity of our model to

univariate modifications of three climate variables (+2 °C and +5 °C warming and ± 20% changes in precipitation and PSP) and two vegetation variables (± 20% changes in biomass of annuals and perennials). For context on the magnitude of these changes, by the end of the century under a CMIP6 intermediate greenhouse gas emissions scenario (SSP2-4.5), 2.5 to 5.2 °C warming, and 0.6 to 15% increases in the mean annual precipitation are projected for western North America (Gutiérrez et al. 2021). We did not evaluate full climate change projections here, although these analyses are planned in future studies that will include associated changes in vegetation and fuels. Overall, our sensitivity analysis of climate variables showed that predicted wildfire probability responses to 5 °C warming and to ± 20% changes in precipitation and PSP were all of fairly similar magnitudes (11–20% changes in expected burned area), and 2 °C warming caused the smallest change (5% increase in expected burned area). The univariate modification of + 5 °C increased annual wildfire probability by less than 0.1 percentage points in most places, with maximum increases of about 0.5 percentage points (Additional file 1: Appendix 4). For comparison, using a fire spread simulation approach, Riley and Loehman (2016) estimated a 0.35 percentage point increase in wildfire probability in northern Idaho shrublands by the mid-twenty first century in response to climate change under a medium–high emissions scenario. Similarly, Gao et al. (2021) used a simpler physically based model that relies on climate to predict wildfire probability and found that, due to warming, wildfire probability is likely to increase in most regions of the USA. However, the model from Gao et al. (2021) did not incorporate vegetation and may have limitations in the sagebrush region where, even under fixed climate conditions, annual grass invasion can strongly impact the fire cycle.

We found that our model was very sensitive to changes in the abundance of annuals, with an 11% increase in expected burned area in response to a 20% increase in annuals. This is a relatively large change in the expected burned area in response to a fairly small increase in annuals, especially given that portions of the sagebrush region have experienced roughly a doubling in biomass of annuals over the last three decades (Additional file 1: Appendix 1). The increase in wildfire probability that drove this change in burn area was mostly concentrated in the northern Great Basin (Additional file 1: Appendix 4), but we expect it would be more widespread with larger increases in annuals. By comparison, predicted wildfire probability responses to changes in the biomass of perennials were more limited and varied geographically (Additional file 1: Appendix 4).

Limitations

Our overall approach and data used create some limitations on the conclusions that can be made using our model. The wildfire, climate, and biomass datasets we relied on all contain errors in their estimates. For example, total herbaceous aboveground biomass estimated by RAP had a correlation of 0.63 with plot-level estimates (Jones et al. 2021), and errors such as these likely affected our model coefficients to varying degrees. The climate data we used was available at a 1-km resolution, and the biomass data was therefore averaged to that scale. Consequently, we missed the wider range in biomass that can occur at finer spatial scales, and using vegetation data collected at fine scales as input into our equation may provide misleading estimates of wildfire probability. Additionally, since our model does not use fire weather predictors, it may be conservative under future climate scenarios in cases where extreme fire weather increases (e.g., due to increased variance of daily temperature) in a way that is not fully captured by changes in average conditions (e.g., increased mean temperature).

Our model relies on observed spatial relationships. While we included only a few variables in our model, they were not fully independent of each other. The effect of multi-collinearity on individual parameter estimates in a model can be substantial, but fortunately, the effect on predictive accuracy may be relatively small. We chose to develop a simple model that could be incorporated into other climate modeling efforts and have adequate functionality under novel conditions. However, even with a simple model such as ours, predictions under future conditions could become unrealistic if the model's predictions rely more on correlative rather than causal relationships and if existing relationships between predictor variables change in the future. Our approach hedges against this problem by representing relationships that are biologically reasonable.

In much of the study area, infrequent fires (fire return intervals > 100 years) are predicted by our model. Due to the relatively short period (1988–2019) of wildfire data used, we, and other researchers using similar datasets, rely on space-for-time substitution for these kinds of estimates. Additionally, our model implicitly incorporates both spatial and temporal relationships, but our study region is very large and spans wide climate gradients, so in the dataset we used, the spatial variability in climate is generally larger than temporal variability in climate, and therefore, our model is heavily influenced by those spatial patterns. As a result, projections using our model assume, for example, that if a cool site becomes hotter, it will eventually have similar fire regime characteristics to a site that has a hot climate today. This assumption is problematic if or when

climate and vegetation are in disequilibrium leading to a mismatch between new and legacy conditions (Felton et al. 2022; Parks et al. 2016). For example, if a site has accumulated fuel because it historically experienced high plant productivity, but then becomes hot and dry (and less productive), it may have elevated fire risk while legacy fuel loads are still present. In comparison with forests, such a mismatch between climate and fuels may be less strong in sagebrush ecosystems over the long run due to the shorter-lived nature and lower overall biomass contributions of the vegetation.

In addition to climate and vegetation, factors such as proximity to roads, likelihood of human ignitions, ability of fire to spread from adjacent land, storm tracks, fuel treatments, and other disturbances are likely also important for understanding wildfire probability in sagebrush communities. In addition, fire extent is heavily influenced by fire suppression efforts. While these influences are broadly reflected in the wildfire occurrence data we used, we did not directly capture them in our model because we wanted to keep it focused on vegetation and climate dynamics. More complex spatial modeling approaches such as FSim can directly incorporate some of these additional factors (Finney et al. 2011); however, substantial uncertainty exists in their future extent, making them challenging to include in estimates of climate change effects on fire. The source of wildfire ignition is, for example, an important factor affecting wildfire trends, and there is substantial spatial variation in the fraction of fires that are human-caused (Balch et al. 2017).

Over the past few decades, human-caused fires have become more frequent within the sagebrush region (National Interagency Fire Center). Human ignitions have lengthened the fire season because they can occur in wetter fuels and also in areas with few lightning strikes (Balch et al. 2017). To try and at least partially assess whether our results were sensitive to the decision of ignoring ignition sources, we fit another model that included the same vegetation and climate predictor variables, but with a measure of human modification on the landscape (Theobald et al. 2020) added as another predictor variable (Additional file 1: Appendix 6). We included this additional variable to act as a rough proxy for the level of human activity and thereby the potential for human-caused ignitions. Including this additional variable did not change the influence of the other variables in the model (i.e., little change in coefficients of the other variables; Additional file 1: Table S3). This suggests that the climate and vegetation relationships represented in our model are not unduly affected by the direct effects of human activity (at least based on the available data used).

Next steps

Feedbacks between climate, vegetation, and wildfire are complex to model, and further research is needed to better understand these interactions. For our sensitivity analysis, we calculated the change in predicted wildfire probability in response to a fixed change in one variable at a time, which clearly does not reflect a real climate change scenario. To address this, we are in the process of incorporating our wildfire model into an individual plant-based model (STEPWAT2; Palmquist et al. 2018), which will enable us to explore fire-vegetation-climate interactions in the sagebrush region. Since our wildfire model can be expressed in a closed-form equation that relies on only five variables, incorporating it into such a plant dynamics model can, at least in some cases, be relatively straightforward. Prior to utilizing it for such purposes, we advise researchers to consider the robustness of the relationships the model depicts and whether its incorporation of average antecedent conditions (as opposed to fire weather) is appropriate for their specific needs.

Conclusion

Understanding the drivers of fire in sagebrush ecosystems is important because these ecosystems are undergoing rapid change largely driven by the invasion of highly flammable annual grasses and subsequent wildfire-induced habitat degradation. We found that annual wildfire probability varied greatly across the sagebrush region and that these observed patterns could be represented quite well by a logistic regression model that included two vegetation (biomass of annual and perennial grasses and forbs) and three climate (temperature, precipitation, PSP) predictor variables. Our model was fit using 3-year averages of antecedent climate and fine fuels from the entire sagebrush region and should be reasonably robust under novel conditions because it consists of ecologically plausible relationships that capture a wide range in climate and fine fuel conditions. It thereby complements existing more complex models that were fit using annual (or sub-annual) data and forms an additional tool for understanding and modeling wildfire and global change impacts on vegetation in the ecologically and economically important sagebrush region.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s42408-024-00252-4.

Additional file 1: Appendix 1. Time-series of biomass of annuals. **Fig. S1.** (Left) Time-series of aboveground biomass of annual forbs and grasses. Different colors represent five climatically different regions or 'folds' of the study area (Fig. S2). The solid lines show the mean biomass across pixels in the fold that have burned at least once between 1988-2019 and dashed lines show the mean biomass across pixels that never burned during that time period. (Right) Ordinary least squares regression lines fit to the values shown in the left panel. **Appendix 2.** Predictor variable summary

statistics. **Table S1.** Summary statistics describing the central tendency and range of predictor variables used to fit the wildfire probability model. Using the model to make predictions with input values of predictor variables that are below the minimums or above the maximums provided here would not be supported by the data we used. Predictor variables were three-year running averages of mean temperature, annual precipitation, proportion summer precipitation, and aboveground biomass of annuals and perennials. For example, the minimum temperature presented here represents the lowest three year mean temperature that occurred in a 1 km pixel in the sagebrush region between 1988 and 2019. **Appendix 3.** Model bias assessment and cross validation. **Fig. S2.** Environmental blocking was used to categorize individual pixels into each of 5 blocks or 'folds' based on similarity of mean annual temperature, mean annual precipitation, and the mean proportion of precipitation received in summer. Observed and predicted wildfire probability for each fold is provided in table S2, to summarize regional bias in the model. Data from pixels in a given fold was also used to create a test dataset, while the remaining four folds were used as the training dataset. Separate models were fit to each of these five training datasets (Fig. S3). **Table S2.** Mean values of each predictor variable, and observed and predicted wildfire probability, across pixels belonging to each of five environmental blocks or 'folds' shown in Fig. S2. Also provided are the mean region-wide values (i.e., means across all pixels). **Fig. S3.** Partial dependence plots depicting the effect of the five predictor variables on modeled annual wildfire probability. Separate lines show results for models fit to each of five training datasets (where one fold was left out) as well as the final model fit to the complete ('biome-wide') dataset. For a given model, the y-axis shows the average predicted wildfire probability (percent wildfire probability per year) for a fixed level of a given predictor variable, across observations of the other predictor variables. Along the x-axis panels show a) mean temperature, b) annual precipitation, c) proportion summer precipitation, d) aboveground biomass of annual forbs and grasses, and e) aboveground biomass of perennial forbs and grasses. **Fig. S4.** Mean observed and predicted annual wildfire probability values for fold 1 (model fit to data from the other four folds). See Fig. 4 in the manuscript for details on interpretation. **Fig. S5.** Mean observed and predicted annual wildfire probability values for fold 2 (model fit to data from the other four folds). See Fig. 4 in the manuscript for details on interpretation. **Fig. S6.** Mean observed and predicted annual wildfire probability values for fold 3 (model fit to data from the other four folds). See Fig. 4 in the manuscript for details on interpretation. **Fig. S7.** Mean observed and predicted annual wildfire probability values for fold 4 (model fit to data from the other four folds). See Fig. 4 in the manuscript for details on interpretation. **Fig. S8.** Mean observed and predicted annual wildfire probability values for fold 5 (model fit to data from the other four folds). See Fig. 4 in the manuscript for details on interpretation. **Appendix 4.** Sensitivity analysis. Fig. S9. Evaluation of model sensitivity to changes in climate and vegetation variables. Panels show the projected change in the number of fires per 100 years in response to a) 2 °C and b) 5 °C increases in mean temperature, and 20% decreases and 20% increases in (c, d) annual precipitation, (e, f) proportion summer precipitation, (g, h), biomass of annuals and (i, j) biomass of perennials. Gray on the maps denotes areas with negligible changes in wildfire frequency (less than 0.1 additional or 0.1 fewer fires per 100 years). Histogram insets present the distribution of values shown on the maps (these are the same histograms as shown in Fig. 6 in the manuscript, except here x-axis limits were restricted to allow for easier comparison of distributions). **Appendix 5.** Dataset comparison. **Fig. S10.** Comparison between wildfire probability estimated by the model presented in this manuscript and long-term wildfire probability modeled by Pastick et al. (2021). The values from Pastick et al. represent the probability of wildfire occurring in a given location over a long period of time (1988-2019). Values from Pastick et al., were aggregated from a 30m x 30m resolution to a 1km x 1km resolution. Points on the figure represent a random sample of pixels in the western and central part of the sagebrush region where the two datasets overlap. The Pearson correlation between the datasets is 0.80. **Fig. S11.** Comparison between wildfire probability estimated by the model presented in this manuscript and burn probability modeled by Short et al. (2023). Values from Short et al. (2023) were aggregated from 250 m x 250 m to a 1km x 1km resolution and converted to % (i.e., multiplied by 100). A random sample of pixels is shown (both datasets cover the entire sagebrush biome). The Pearson correlation between the datasets is 0.68. An ordinary least squares regression line is shown in blue, and the 1:1 line is in black. **Fig. S12.** Comparison between wildfire probability estimated by the model presented in this manuscript and relative wildfire probability modeled by Smith et al. (2022). Annual predicted values (1988-2019) from Smith et al. (2022) were averaged across years and aggregated to a 1km x 1km resolution. Points on the figure represent a random sample of pixels in the Great Basin which is where the two datasets overlap. The Pearson correlation between the datasets is 0.82. **Appendix 6.** Including human modification as a predictor of annual wildfire probability. **Table S3.** Comparison of model coefficients between our main model (described in the manuscript), and a model with human modification (HMod) added as an additional predictor variable. Note that including HMod only caused small changes in the other coefficients. Both models were logistic regression models so that the probability of 'success' (fire) in a given year could be modeled. A logit link function was used. **Fig. S13.** Partial dependence plots depicting the effect of the predictor variables on modeled annual wildfire probability and showing the similarity between the main model (which is described in manuscript; blue) and the model where human modification was included as an additional predictor variable (black). The y-axis shows the average predicted wildfire probability for a fixed level of a given predictor variable, across all combinations of values of the other predictor variables. Rugs on the x-axis show the deciles (10th - 90th percentiles) of the predictor variable. Along the x-axis panels show a) mean temperature, b) annual precipitation, c) the proportion of precipitation that falls in summer (June-Aug), d) aboveground biomass of annual forbs and grasses, e) aboveground biomass of perennial forbs and grasses, and f) human modification. **Fig. S14.** Comparison of mean observed and predicted annual wildfire probability for the model where human modification was included as an additional predictor variable. Panels show mean observed (black circles) and predicted (blue triangles) annual wildfire probability for each percentile of a) mean temperature, b) annual precipitation, c) the proportion of precipitation that falls in summer (June-Aug), d) aboveground biomass of annual forbs and grasses, e) aboveground biomass of perennial forbs and grasses, and f) human modification. Data were binned by percentile (i.e., 100 bins) of a given predictor variable, and the x-axis shows the mean value of each percentile of that variable. To illustrate, the right most black circle in panel d) shows the mean observed wildfire probability across all pixels where biomass of annuals is between the 99th and 100th percentile, and the rightmost blue triangle in that panel shows the mean predicted wildfire probability for those same pixels. Each point on the figures represents the mean of ~250,000 values (i.e., 1% of the entire dataset). **Fig. S15.** (a) Annual wildfire probability predicted by the model that used human modification in addition to historical average vegetation, and climate conditions as predictors. These modeled values represent the probability (%) of fire occurring in a given year. (b) The change in predicted wildfire probability shown in panel (a) and the wildfire probability predicted by the main model (which does not include human modification as a predictor; shown in Fig. 2a). Blues indicate that the model that included human modification predicted lower wildfire probability, whereas reds indicate that this model predicted higher wildfire probability. **Appendix 7.** Variable importance. **Fig. S16.** Variable importance of all variables in the logistic regression model presented in the manuscript. Here variable importance was defined as the absolute value of the test statistic (z) of the respective coefficient. Both climate and vegetation variables are three year running averages (i.e., mean of the current year and previous two years). Abbreviations: T, temperature; P, annual precipitation; PSP, proportion summer precipitation; AFG, aboveground biomass of annual forbs and grasses; PFG, aboveground biomass of perennial forbs and grasses; log, base 10 logarithm. Interactions between variables are denoted by ':'.

Acknowledgements
We are grateful to Megan Creutzberg and Chad Boyd for the advice on our analysis approach and Joseph Smith for the comments on a draft of the manuscript. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the US Government.

Authors' contributions
All authors contributed to the conception of the research and contributed critically to the manuscript writing. JBB, DRS, and MCH developed the analysis methods. MCH conducted the analyses and wrote the first draft of the manuscript.

Funding
Support for this research was provided by the US Fish and Wildlife Service, the USGS Ecosystems Mission Area, and the USGS Climate Adaptation Science Centers.

Availability of data and materials
The data used in this study are available in the ScienceBase repository, https://doi.org/10.5066/P9EFC6YC. The scripts used for the analyses are available in the Zenodo repository, https://zenodo.org/doi/10.5281/zenodo.10637598.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Southwest Biological Science Center, U.S. Geological Survey, 2255 N Gemini Road, Flagstaff, AZ 86001, USA. [2]Northern Arizona University, Center for Adaptable Western Landscapes, S San Francisco Street, Flagstaff, AZ 86011, USA. [3]Department of Biological Sciences, Marshall University, 1 John Marshall Drive, Huntington, WV 25755, USA. [4]Bureau of Land Management, National Interagency Fire Center, 3833 Development Avenue, Boise, ID 83705, USA. [5]U.S. Fish and Wildlife Service, Science Applications Program, Lakewood Colorado 80228, 134 Union Boulevard, Lakewood, CO 80228, USA. [6]School of the Environment, Yale University, 195 Prospect Street, New Haven, CT 06511, USA. [7]Western Association of Fish and Wildlife Agencies, P.O. Box 190150, Boise, ID 83719, USA. [8]U.S. Forest Service, Rocky Mountain Research Station, Missoula Fire Sciences Laboratory, 5775 W Broadway Street, Missoula, MT 59808, USA. [9]U.S. Fish and Wildlife Service, Science Applications Program, 1340 Financial Boulevard, Reno, NV 89502, USA. [10]Ecosystems Mission Area, U.S Geological Survey, 12201 Sunrise Valley Drive, Reston, VA 20192, USA. [11]U.S. Geological Survey, Northwest Climate Adaptation Science Center, Flagstaff, AZ 86001, USA.

Received: 4 May 2023 Accepted: 10 January 2024
Published online: 28 February 2024

References

Abatzoglou, John T., and Crystal A. Kolden. 2013. Relationships between climate and macroscale area burned in the Western United States. *International Journal of Wildland Fire* 22 (7): 1003–1020. https://doi.org/10.1071/WF13019.

Baker, W.L. 2006. Fire and restoration of sagebrush ecosystems. *Wildlife Society Bulletin (1973-2006)* 34 (1): 177–85.

Baker, W.L. 2013. Is wildland fire increasing in sagebrush landscapes of the Western United States? *Annals of the Association of American Geographers* 103 (1): 5–19.

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 Sciences of the United States of America* 114 (11): 2946–2951. https://doi.org/10.1073/pnas.1617394114.

Balch, Jennifer K., Bethany A. Bradley, Carla M. D'Antonio, and José Gómez-Dans. 2013. Introduced annual grass increases regional fire activity across the arid Western USA (1980–2009). *Global Change Biology* 19 (1): 173–183. https://doi.org/10.1111/gcb.12046.

Barbero, R., J.T. Abatzoglou, N.K. Larkin, C.A. Kolden, and B. Stocks. 2015. Climate change presents increased potential for very large fires in the contiguous United States. *International Journal of Wildland Fire* 24 (7): 892. https://doi.org/10.1071/WF15083.

Bell, David M., and Daniel R. Schlaepfer. 2016. On the dangers of model complexity without ecological justification in species distribution modeling. *Ecological Modelling* 330: 50–59. https://doi.org/10.1016/j.ecolmodel.2016.03.012.

Biecek, Przemyslaw, and Tomasz Burzykowski. 2021. *Explanatory model analysis*. Chapman and Hall/CRC, New York. https://ema.drwhy.ai/.

Bowman, David M.J.S.., Crystal A. Kolden, John T. Abatzoglou, Fay H. Johnston, Guido R. van der Werf, and Mike Flannigan. 2020. Vegetation fires in the anthropocene. *Nature Reviews Earth and Environment* 1 (10): 500–515. https://doi.org/10.1038/s43017-020-0085-3.

Bradley, Bethany A., Caroline A. Curtis, Emily J. Fusco, John T. Abatzoglou, Jennifer K. Balch, Sepideh Dadashi, and Mao Ning Tuanmu. 2018. Cheatgrass (Bromus tectorum) distribution in the intermountain Western United States and its relationship to fire frequency, seasonality, and ignitions. *Biological Invasions* 20 (6): 1493–1506. https://doi.org/10.1007/s10530-017-1641-8.

Bukowski, Beth E., and William L. Baker. 2013. Historical fire regimes, reconstructed from land-survey data, led to complexity and fluctuation in sagebrush landscapes. *Ecological Applications* 23 (3): 546–564. https://doi.org/10.1890/12-0844.1.

Coates, Peter S., Mark A. Ricca, Brian G. Prochazka, Matthew L. Brooks, Kevin E. Doherty, Travis Kroger, Erik J. Blomberg, Christian A. Hagen, and Michael L. Casazza. 2016. Wildfire, climate, and invasive grass interactions negatively impact an indicator species by reshaping sagebrush ecosystems. *Proceedings of the National Academy of Sciences* 113 (45): 12745–12750. https://doi.org/10.1073/pnas.1606898113.

Coop JD, Parks SA, Stevens-Rumann CS, Ritter SM, Hoffman CM. 2022. Extreme fire spread events and area burned under recent and future climate in the Western USA. *Global Ecology and Biogeography*, no. March: 1949–59. https://doi.org/10.1111/geb.13496.

Crist, M.R. 2023. Rethinking the focus on forest fires in federal wildland fire management: Landscape patterns and trends of non-forest and forest burned area. *Journal of Environmental Management* 327: 116718. https://doi.org/10.1016/j.jenvman.2022.116718.

Dahal, Devendra, Neal J. Pastick, Stephen P. Boyte, Sujan Parajuli, Michael J. Oimoen, and Logan J. Megard. 2022. Multi-species inference of exotic annual and native perennial grasses in rangelands of the Western United States using harmonized Landsat and Sentinel-2 data. *Remote Sensing* 14 (4): 807. https://doi.org/10.3390/rs14040807.

D'Antonio, Carla M., and Peter M. Vitousek. 1992. Biological invasions by exotic grasses, the grass/fire cycle, and global change. *Annual Review of Ecology and Systematics* 23 (1): 63–87. https://doi.org/10.1146/annurev.es.23.110192.000431.

Davies, Kirk W., and Aleta M. Nafus. 2013. Exotic annual grass invasion alters fuel amounts, continuity and moisture content. *International Journal of Wildland Fire* 22 (3): 353–358. https://doi.org/10.1071/WF11161.

Doherty, K., Theobald, D.M., Bradford, J.B., Wiechman, L.A., Bedrosian, G., Boyd, C.S., Cahill, M., Coates, P.S., Creutzburg, M.K., Crist, M.R., Finn, S.P., Kumar, A.V., Littlefield, C.E., Maestas, J.D., Prentice, K.L., Prochazka, B.G., Remington, T.E., Sparklin, W.D., Tull, J.C., Wurtzebach, Z., and Zeller, K.A. 2022. A sagebrush conservation design to proactively restore America's sagebrush biome: U.S. *Geological Survey Open-File Report 2022–1081 38*: 1–38. https://doi.org/10.3133/ofr20221081.

Dye, Alex W., Peng Gao, John B. Kim, Ting Lei, Karin L. Riley, and Larissa Yocom. 2023. High-resolution wildfire simulations reveal complexity of climate change impacts on projected burn probability for Southern California. *Fire Ecology* 19 (1): 20. https://doi.org/10.1186/s42408-023-00179-2.

Felton, Andrew J., Robert K. Shriver, Michael Stemkovski, John B. Bradford, Katharine N. Suding, and Peter B. Adler. 2022. Climate disequilibrium dominates uncertainty in long-term projections of primary productivity. *Ecology Letters* 25 (12): 2688–2698. https://doi.org/10.1111/ele.14132.

Finney, Mark A., Charles W. McHugh, Isaac C. Grenfell, Karin L. Riley, and Karen C. Short. 2011. A simulation of probabilistic wildfire risk components for the continental United States. *Stochastic Environmental Research and Risk Assessment* 25 (7): 973–1000. https://doi.org/10.1007/s00477-011-0462-z.

Flannigan, M.D., B.M. Wotton, G.A. Marshall, W.J. de Groot, J. Johnston, N. Jurko, and A.S. Cantin. 2016. Fuel moisture sensitivity to temperature and precipitation: Climate change implications. *Climatic Change* 134 (1–2): 59–71. https://doi.org/10.1007/s10584-015-1521-0.

Gao, Peng, Adam J. Terando, John A. Kupfer, J. Morgan Varner, Michael C. Stambaugh, Ting L. Lei, and J. Kevin Hiers. 2021. Robust projections of future fire probability for the conterminous United States. *Science of the Total Environment* 789: 147872. https://doi.org/10.1016/j.scitotenv.2021.147872.

Gorelick, Noel, Matt Hancher, Mike Dixon, Simon Ilyushchenko, David Thau, and Rebecca Moore. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* 202: 18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Gutiérrez, J.M., R.G. Jones, G.T. Narisma, L.M. Alves, M. Amjad, I.V. Gorodetskaya, M. Grose, et al. 2021. "2021: atlas." In *Climate Change 2021: The physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, et al. Cambridge University Press. https://doi.org/10.1017/9781009157896.021.

Hartigan, J. A., and M. A. Wong. 1979. Algorithm AS 136: A K-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28 (1): 100–108. https://doi.org/10.2307/2346830.

Jeffries, Michelle I., and Sean P. Finn. 2019. The sagebrush biome range extent, as derived from classified Landsat imagery. U.S. Geological Survey. https://doi.org/10.5066/P950H8HS.

Jones, Matthew O., Nathaniel P. Robinson, David E. Naugle, Jeremy D. Maestas, Matthew C. Reeves, Robert W. Lankston, and Brady W. Allred. 2021. Annual and 16-day rangeland production estimates for the Western United States. *Rangeland Ecology & Management* 77: 112–117. https://doi.org/10.1016/j.rama.2021.04.003.

Kloster, Silvia, and Gitta Lasslop. 2017. Historical and future fire occurrence (1850 to 2100) simulated in CMIP5 earth system models. *Global and Planetary Change* 150: 58–69. https://doi.org/10.1016/j.gloplacha.2016.12.017.

Knorr, W., L. Jiang, and A. Arneth. 2016. Climate, CO_2 and human population impacts on global wildfire emissions. *Biogeosciences* 13 (1): 267–282. https://doi.org/10.5194/bg-13-267-2016.

Longland, William S., and Sheryl L. Bateman. 2002. Viewpoint: The ecological value of shrub islands on disturbed sagebrush rangelands. *Journal of Range Management* 55 (6): 571. https://doi.org/10.2307/4004000.

Maestas, Jeremy D., Joseph T. Smith, Brady W. Allred, David E. Naugle, Matthew O. Jones, Casey O'Connor, Chad S. Boyd, Kirk W. Davies, Michele R. Crist, and Andrew C. Olsen. 2022. Using dynamic, fuels-based fire probability maps to reduce large wildfires in the Great Basin. *Rangeland Ecology and Management*, no. September. https://doi.org/10.1016/j.rama.2022.08.002.

McAdoo, J. Kent., Brad W. Schultz, and Sherman R. Swanson. 2013. Aboriginal precedent for active management of sagebrush-perennial grass communities in the Great Basin. *Rangeland Ecology & Management* 66 (3): 241–253. https://doi.org/10.2111/REM-D-11-00231.1.

Palmquist, Kyle A., Daniel R. Schlaepfer, Rachel R. Renne, Stephen C. Torbit, Kevin E. Doherty, Thomas E. Remington, Greg Watson, John B. Bradford, and William K. Lauenroth. 2021. Divergent climate change effects on widespread dryland plant communities driven by climatic and ecohydrological gradients. *Global Change Biology* 27 (20): 5169–5185. https://doi.org/10.1111/gcb.15776.

Palmquist, Kyle A., John B. Bradford, Trace E. Martyn, Daniel R. Schlaepfer, and William K. Lauenroth. 2018. STEPWAT2: An Individual-Based Model for Exploring the Impact of Climate and Disturbance on Dryland Plant Communities. *Ecosphere 9* (8). https://doi.org/10.1002/ecs2.2394.

Parks, Sean A., Carol Miller, John T. Abatzoglou, Lisa M. Holsinger, Marc-André. Parisien, and Solomon Z. Dobrowski. 2016. How will climate change affect wildland fire severity in the Western US? *Environmental Research Letters* 11 (3): 035002. https://doi.org/10.1088/1748-9326/11/3/035002.

Paruelo, Jose M., and W.K. Lauenroth. 1996. Relative abundance of plant functional types in grasslands and shrublands of North America. *Ecological Applications* 6 (4): 1212–1224. https://doi.org/10.2307/2269602.

Pastick, Neal J., Bruce K. Wylie, Matthew B. Rigge, Devendra Dahal, Stephen P. Boyte, Matthew O. Jones, Brady W. Allred, Sujan Parajuli, and Zhuoting Wu. 2021. Rapid monitoring of the abundance and spread of exotic annual grasses in the Western United States using remote sensing and machine learning. *AGU Advances* 2 (2). https://doi.org/10.1029/2020av000298.

Pausas, Juli G., and Jon E. Keeley. 2021. Wildfires and global change. *Frontiers in Ecology and the Environment* 19 (7): 387–395. https://doi.org/10.1002/fee.2359.

Pilliod, David S., Justin L. Welty, and Robert S. Arkle. 2017. Refining the cheatgrass–fire cycle in the Great Basin: Precipitation timing and fine fuel composition predict wildfire trends. *Ecology and Evolution* 7 (19): 8126–8151. https://doi.org/10.1002/ece3.3414.

R Core Team. 2023. R: a language and environment for statistical computing. Vienna, Austria: R Foundation For Statistical Computing. https://www.r-project.org/.

Remington, Thomas E., Pat A. Deibert, Steven E. Hanser, Dawn M. Davis, Leslie A. Robb, and Justin L. Welty. 2021. Sagebrush conservation strategy–challenges to sagebrush conservation. U.S. Geological Survey Open-File Report 2020–1125. https://doi.org/10.3133/ofr20201125.

Renne, Rachel R., John B. Bradford, Ingrid C. Burke, and William K. Lauenroth. 2019. Soil texture and precipitation seasonality influence plant community structure in North American temperate shrub steppe. *Ecology* 100 (11): 1–12. https://doi.org/10.1002/ecy.2824.

Riley, Karin L., John T. Abatzoglou, Isaac C. Grenfell, Anna E. Klene, and Faith Ann Heinsch. 2013. The Relationship of Large Fire Occurrence with Drought and Fire Danger Indices in the Western USA, 1984–2008: The Role of Temporal Scale. *International Journal of Wildland Fire 22* (7): 894–909. https://doi.org/10.1071/WF12149.

Riley, Karin L., and Rachel A. Loehman. 2016. Mid-21st-century climate changes increase predicted fire occurrence and fire season length, Northern Rocky Mountains. *United States. Ecosphere* 7 (11): 1–19. https://doi.org/10.1002/ecs2.1543.

Riley, Karin L., A. Park Williams, Shawn P. Urbanski, David E. Calkin, Karen C. Short, and Christopher D. O'Connor. 2019. Will landscape fire increase in the future? A systems approach to climate, fire, fuel, and human drivers. *Current Pollution Reports* 5 (2): 9–24. https://doi.org/10.1007/s40726-019-0103-6.

Roberts, David R., Volker Bahn, Simone Ciuti, Mark S. Boyce, Jane Elith, Gurutzeta Guillera-Arroita, Severin Hauenstein, et al. 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* 40 (8): 913–929. https://doi.org/10.1111/ecog.02881.

Robinson, Nathaniel P., Matthew O. Jones, Alvaro Moreno, Tyler A. Erickson, David E. Naugle, and Brady W. Allred. 2019. Rangeland productivity partitioned to sub-pixel plant functional types. *Remote Sensing* 11 (12): 1427. https://doi.org/10.3390/rs11121427.

Schlaepfer, Daniel R., William K. Lauenroth, and John B. Bradford. 2014. Natural regeneration processes in big sagebrush (Artemisia Tridentata). *Rangeland Ecology and Management* 67 (4): 344–357. https://doi.org/10.2111/REM-D-13-00079.1.

Shinneman, Douglas J., and William L. Baker. 2009. Environmental and climatic variables as potential drivers of post-fire cover of cheatgrass (Bromus tectorum) in seeded and unseeded semiarid ecosystems. *International Journal of Wildland Fire* 18 (2): 191. https://doi.org/10.1071/WF07043.

Shinneman, Douglas J, Michelle I. Jeffries, Justin L. Welty, and David S. Pilliod. n.d. Wildland fire trends tool. Accessed March 21, 2023. https://www.usgs.gov/centers/forest-and-rangeland-ecosystem-science-center/science/wildland-fire-trends-tool.

Shinneman, Douglas J., Eva K. Strand, Mike Pellant, John T. Abatzoglou, Mark W. Brunson, Nancy F. Glenn, Julie A. Heinrichs, Mojtaba Sadegh, and Nicole M. Vaillant. 2023b. Future direction of fuels management in sagebrush rangelands. *Rangeland Ecology & Management* 86: 50–63. https://doi.org/10.1016/j.rama.2022.10.009.

Short, Karen C., Greg K. Dillon, Joe H. Scott, Kevin C. Vogler, Melissa R. Jaffe, Olszewski, Julia H., Mark A. Finney, et al. 2023. Spatial datasets of probabilistic wildfire risk components for the sagebrush biome (270m). *Fort Collins, CO: Forest Service Research Data Archive*. https://doi.org/10.2737/RDS-2023-0050.

Short, Karen C., Mark A. Finney, Kevin C. Vogler, Joe H. Scott, Julie W. Gilbertson-Day, and Isaac C. Grenfell. 2020. Spatial datasets of probabilistic wildfire risk components for the United States (270m). 2nd Edition. *Fort*

Collins, CO: Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2016-0034-2.

Smith, Joseph T., Brady W. Allred, Chad S. Boyd, Kirk W. Davies, Matthew O. Jones, Andrew R. Kleinhesselink, Jeremy D. Maestas, and David E. Naugle. 2022. Where there's smoke, there's fuel: dynamic vegetation data improve predictions of wildfire hazard in the Great Basin. *Rangeland Ecology & Management*, no. September (September). https://doi.org/10.1016/j.rama.2022.07.005.

Smith, Joseph T., Brady W. Allred, Chad S. Boyd, Kirk W. Davies, Andrew R. Kleinhesselink, Scott L. Morford, and David E. Naugle. 2023. Fire needs annual grasses more than annual grasses need fire. *Biological Conservation* 286: 110299. https://doi.org/10.1016/j.biocon.2023.110299.

Theobald, David M., Christina Kennedy, Bin Chen, James Oakleaf, Sharon Baruch-Mordo, and Joe Kiesecker. 2020. Earth transformed: Detailed mapping of global human modification from 1990 to 2017. *Earth System Science Data* 12 (3): 1953–1972. https://doi.org/10.5194/essd-12-1953-2020.

Thornton, M. M., R Shrestha, Wei, Y., Thornton, P.E., Kao, S., and Wilson, B.E. 2020. Daymet: daily surface weather data on a 1-km grid for North America, version 4. netCDF, 0 MB. https://doi.org/10.3334/ORNLDAAC/1840.

Tortorelli, Claire M., John B. Kim, Nicole M. Vaillant, Karin Riley, Alex Dye, Ty. C. Nietupski, Kevin C. Vogler, et al. 2023. Feeding the fire: Annual grass invasion facilitates modeled fire spread across inland northwest forest-mosaic landscapes. *Ecosphere* 14 (2): e4413. https://doi.org/10.1002/ecs2.4413.

Valavi, Roozbeh, Jane Elith, José J. Lahoz-Monfort, and Gurutzeta Guillera-Arroita. 2019. blockCV: an R package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. Edited by David Warton. *Methods in Ecology and Evolution* 10(2);225–32. https://doi.org/10.1111/2041-210X.13107.

Welty, J.L., and M.I. Jeffries. 2021. Combined wildland fire datasets for the United States and certain territories, 1800s-Present. *U.S. Geological Survey Data Release*. https://doi.org/10.5066/P9ZXGFY3.

Wenger, Seth J., and Julian D. Olden. 2012. Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods in Ecology and Evolution* 3 (2): 260–267. https://doi.org/10.1111/j.2041-210X.2011.00170.x.

Whittaker, Robert Harding. 1975. *Communities and ecosystems*, 2d ed. New York: Macmillan.

Wu, Chao, Sergey Venevsky, Stephen Sitch, Lina M. Mercado, Chris Huntingford, and A. Carla Staver. 2021. Historical and future global burned area with changing climate and human demography. *One Earth* 4 (4): 517–530. https://doi.org/10.1016/j.oneear.2021.03.002.

Publisher's Note