Microwave Retrievals of Soil Moisture Improve Grassland Wildfire Predictions

Angela J. Rigden1, Robert S. Powell1, Aleyda Trevino1, Kaighin A. McColl1, and Peter Huybers1

1Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA

Abstract  Statistical analyses of wildfires demonstrate that vapor pressure deficit (VPD) allows for skillful predictions, likely because it reflects fuel moisture content. Soil moisture provides a potentially complimentary measure of water availability but has been less explored because of sparse measurements. Using measurements from the Soil Moisture Active Passive satellite, the predictive skill afforded by using VPD and soil moisture together is explored across the Western United States. Receiver operating characteristic curves estimated from 1,907 fires indicate that inclusion of soil moisture in addition to VPD observations permits for more skillful prediction ($p < 0.05$). When VPD already signals high risk, the addition of soil moisture reduces the false positive rate grasslands (from 75% to 62%), shrublands (76% to 67%), and forests (74% to 68%) for a true positive rate of 75%. These results show potential to improve daily fire risk models with the addition of remotely sensed soil moisture.

Plain Language Summary  Identifying environmental conditions that lead to fire ignition is critical for risk management, fire mitigation, and resource allocation. Vapor pressure deficit (VPD) has been shown to be an effective predictor because it reflects water availability near the land surface. More direct measures of water availability, such as soil moisture, have yet to be thoroughly explored because measurements have historically been sparse. Using recently available soil moisture observations from satellites, we explore the ability to retrospectively predict 1,907 fires across the Western United States in grasslands, shrublands, and evergreen needleleaf forests during June–August. Our results show that soil moisture can be used in conjunction with VPD to improve predictions of fires, particularly in grasslands when fire risk is already high.

1. Introduction

California experienced its most costly, deadly, and largest wildfires to date in 2018 (NOAA, 2020). Over 8.7 million acres burned across the United States that year, well exceeding the 10-year national average (2009–2018) of 6.8 million acres (NOAA, 2020). The Western United States has, more generally, experienced increases in wildfires since the 1980s (Abatzoglou & Williams, 2016; Dennison et al., 2014; Williams et al., 2019). This increase in fire activity was partially attributed to increases in aridity associated with anthropogenic climate change (Abatzoglou & Williams, 2016). More specifically, it was found that increases in atmospheric vapor pressure deficit (VPD) induced by rising temperatures are linked to the increases in fire activity (Seager et al., 2015; Williams et al., 2015, 2019).

The predictive power of VPD lies in its connection to the land surface water balance, particularly the water content of fuel (Seager et al., 2015; Williams et al., 2015, 2019). VPD can be conceptually thought of as the air’s drying potential; all else being equal, evaporation from fuel increases with increasing VPD. When fuels become drier, fuel combustion increases because less heat is required to evaporate the remaining water in the fuel. When the land surface becomes water limited, feedbacks between the land and atmosphere can also lead to extremes in temperature (Hirschi et al., 2010) and VPD (Zhou et al., 2019), serving as a positive feedback that reinforces fuel drying. Although VPD is an indirect measure of land-surface water availability, particularly at seasonal to annual timescales, previous studies have found it to be at least as correlated with annual burned area as compared to other proxies of fuel moisture content (Williams et al., 2015).

More direct measures of water availability, such as soil moisture, are less frequently used to quantify fire risk because in situ observations have been sparse (Levi et al., 2019). There is growing evidence, though,
that remotely sensed soil moisture is effective for assessing wildfire risk in areas without extensive soil moisture measurement networks (Bartsch et al., 2009; Chaparro et al., 2016; Farahmand et al., 2020; Jensen et al., 2018; Sungmin et al., 2020; Thomas Ambadan et al., 2020; Westerling et al., 2006), but its predictive power still requires further assessment (Levi et al., 2019). In particular, it is unclear if remotely sensed soil moisture observations add predictive skill beyond that obtainable from VPD. Most risk assessment models, including the National Fuel Fire Rating Danger System in the United States, rely almost entirely on observations of temperature and humidity to drive daily variations in risk (NWCG, Fire Danger Working Team, 2002). Reanalysis-based fire danger models, such as the European Centre for Medium-range Weather Forecasts (ECMWF) Fire Forecast Model, also rely entirely on atmospheric forcing to model daily variations in fuel moisture (Vitolo et al., 2020).

In this study, we assess the utility of using daily remotely sensed soil moisture in combination with VPD to retrospectively predict fire ignition across the Western United States. We first compare the predictive power of soil moisture and VPD and, then, evaluate the extent to which the combination of these two variables allows for improved fire ignition predictions.

2. Data and Methods

We analyze 1,907 fire ignition points across the Western United States (defined as west of −100°) during June–August from 2015 to 2018 (Figure 1) (GeoMac, 2019). Fires are included in the analysis if they originated in grasslands, shrublands, and evergreen needleleaf forests, as these plant functional types comprise over 80% of the reported fire ignitions in this region. At each ignition point, we extract a 4-year (2015–2018) time series of soil moisture and VPD during the fire season (June–August) from the nearest grid-box center. For purposes of equal evaluation, only days having available soil moisture observations (approximately every 3 days) are evaluated, with missing days filled for both soil moisture and VPD using the most recent previous observation. To avoid influences associated with the fire itself, the day prior to fire ignition is analyzed relative to the other days in the fire season.

Maximum daily VPD estimates are from the Parameter-elevation Relationships on Independent Slopes Model (PRISM) (Daly et al., 2008). This data set is based on weather station observations and gridded to a 4-km resolution using an interpolation algorithm that incorporates spatially resolved covariates, such as topography and other physiographical factors (Daly et al., 2008).

Soil moisture observations are from the Soil Moisture Active Passive (SMAP) satellite mission (Entekhabi et al., 2010), which was launched in 2015. We estimate soil moisture from morning (6:00 a.m.) retrievals (Level 3) using the multitemporal dual channel retrieval algorithm (MT-DCA) (Konings et al., 2016). The MT-DCA algorithm improves upon the baseline SMAP algorithm by refining estimates of vegetation optical depth (VOD) (Konings et al., 2016, 2017), which, in turn, should improve estimates of soil moisture. It is difficult to quantitatively assess the differences between the MT-DCA and baseline algorithms, though, due to sparse in situ measurements. Soil moisture estimates from the MT-DCA algorithm are statistically indistinguishable from those estimated by the baseline SMAP algorithm (Konings et al., 2017). The soil moisture retrievals are gridded to a 9-km Equal-Area scalable Earth-2 (EASE-2) grid and have an approximate 3-day return period, depending on latitude. Soil moisture observations from SMAP are preferentially chosen over other satellite products because baseline SMAP estimates have been shown to most accurately correlate with in situ soil moisture measurements (Cui et al., 2018; Ma et al., 2019), particularly in arid climates (Ma et al., 2019). Soil moisture observations from SMAP represent the moisture content in the first 5 cm of soil, and root-zone soil moisture is estimated from these near-surface SMAP observations using an exponential decay filter (Albergel et al., 2008; Ford et al., 2014). We assume a spatially uniform timescale of soil moisture variation equal to 10 days, though we obtain similar results at timescales between 7 and 13 days.

To quantitatively evaluate the predictive power of soil moisture and VPD, we use receiver operating characteristic (ROC) curves. ROC curves summarize the ability of a binary warning signal to anticipate an event: in this case, the presence or absence of a wildfire. They were originally developed in the field of radar signal detection (Peterson et al., 1954) but have recently been used in the wildfire sciences (Gudmundsson et al., 2014; Krueger et al., 2017). When applied to wildfires, the warning signal is often based on an environmental variable, and the event is the fire ignition. Predictive variables are discretized using a threshold, such that when the predictive variable is less than or greater than the specified threshold, it indicates a warning. If the warning signal correctly identifies the event, it is called a “true positive,” and if it misidentifies the
Figure 1. Wildfire ignition locations and seasonality across the Western United States. Map (a) displays fire ignition points occurring in June–August from 2015–2018 for three plant functional types: grasslands (GRA, green), shrublands (SHR, orange), and evergreen needleleaf forests (ENF, blue). The darker shaded symbols represent fires that occur when the VPD is classified as high ("high vapor pressure deficit," HV). The lighter shaded symbols represent the remaining fires that occur when the VPD is low ("low vapor pressure deficit," LV). Histogram (b) represents the relative frequency of fire occurrence (HV + LV) in each month for each plant functional type. The "fire season" is highlighted in blue, though the results are qualitatively insensitive to the chosen fire season.

event, it is called a “false positive” or, more colloquially, a “false alarm.” The ROC curve involves plotting true positive rates against false positive rates for a range of warning signal thresholds. The area under the ROC curve (AUC) is used to summarize the predictive power of the variable. An area of 0.5 results when the true positive rate equals the false positive rate; this implies that the variable has no predictive power. In contrast, an area of 1 indicates perfect predictive power because there would then be only true positives for all thresholds. ROC curves are particularly suited for our application because they require no parametric assumptions. Implicitly, though, we assume that environmental conditions prefire and postfire are similar, which should be explored in future studies.

ROC curves are estimated for each plant functional type. Aggregation across locations is required because there is only one ignition (or event) per site. For each land cover type, we translate the soil moisture observations into binary warning signals by varying the threshold from the minimum to the maximum observed values (see Table S1 in the supporting information). For each threshold, true positive rates are estimated by calculating the number of fires that occurred when soil moisture was below the threshold divided by the total number of fires. False positive rates are calculated similarly as the number of days in which soil moisture was below the threshold but no fire occurred divided by the total number of days not having a fire. Uncertainty in the ROC curves and associated AUC values are estimated by bootstrapping ignition locations using a thousand realizations. In Table 1, we display the median AUC, as well as the 5th and 95th percentiles as confidence intervals. We apply an analogous methodology for VPD (see Table S1 for VPD thresholds limits).
In our initial application, we construct ROC curves for soil moisture and VPD separately. Assessing both predictors is interesting because of the potential for disequilibrium between the supply of water in the soils and the demand from the atmosphere. Summertime VPD in the Western United States is driven by precipitation, heating, and transport from the surface and within the atmosphere (Seager et al., 2015). Variations in soil moisture are influenced by precipitation and atmospheric conditions as well but also depend on land surface properties such as vegetation, soil properties, and topography (Seneviratne et al., 2010) that tend to increase spatial heterogeneity. Furthermore, soil moisture exhibits longer memory than the atmosphere, particularly in arid regions and deeper in the soil column (McColl et al., 2017).

Given the distinctions between the drivers of soil moisture and VPD and interest in controlling for false positive rates in fire prediction, we also explore whether soil moisture provides additional predictive skill under conditions of high VPD (Figure S1). Rather than using all of the observations, ROC curves are constructed using a subset of data that includes only days with high VPD, with the subscript “HV” denoting this subset of days. The high VPD threshold is defined for each plant functional type based on when the AUC curve fit to VPD approaches 0.5 (Figure S2; 44.5 hPa in grasslands, 42.7 hPa in shrublands, and 27.2 hPa in evergreen needleleaf forests). The results are insensitive to the high VPD threshold (Table S2). Focusing on high-VPD days reduces the number of ignitions by approximately 50%, from 1,907 to 886 (Figure 1; 450 to 156 in grasslands, 633 to 269 in shrublands, and 824 to 461 in evergreen needleleaf forests). To estimate if the AUC in SMHV is statistically larger than VPDHV, we calculate their difference in each bootstrap sample and estimate the 95th and 99th quantiles (Figure S3).

3. Results

As expected, we find that fires are most probable when soil moisture is low and VPD high relative to typical conditions observed during the fire season (Figure 2). In both fire and nonfire conditions, it is clear that soil moisture and VPD are negatively correlated, as also expected from land-atmosphere feedbacks (Seneviratne et al., 2010). The cross-correlation between VPD and soil moisture, however, is modest on nonfire days ($r = -0.50$) and even lower on fire days (i.e., the day before ignition, $r = -0.36$), indicating that these variables are not interchangeable.

Results from the ROC curve analysis show that, when considered individually, VPD outperforms soil moisture for all three plant functional types (Table 1; Figure S4). Specifically, high VPD is a strong indication of fire occurrence, with AUC values ranging from 0.63 in shrublands to 0.67 in grasslands. Soil moisture demonstrates predictive power, especially in grassland ecosystems (AUC = 0.65), but in all cases its AUC is lower than for VPD. Both metrics give a high false positive rate relative to true positives (Figure S4), reflecting that fire incidence is uncommon relative to the occurrence of fluctuations toward high VPD and low soil moisture.

When ROC curves are constructed from days with high VPD, we find that soil moisture permits for better predicting wildfires. For example, for a true positive rate of 75%, high VPD alone would give a false positive
Figure 2. Fire conditions are distinct from fire season conditions. Probability of daily soil moisture and maximum VPD (a) on the day before the fire and (b) during the fire season, excluding the day before and day of the fire. Subplot (c) represents the difference between (a) and (b), with the blue representing conditions that are more likely 1 day before the fire and red representing conditions that are less likely 1 day before the fire. In each subplot, observations are aggregated temporally and spatially across all three plant functional types.

rate of approximately 75% in grasslands, 76% shrublands, and 74% evergreen needleleaf forests. Using soil moisture in conjunction with high VPD, however, reduced the false positive rate to 63% in grasslands, 67% in evergreen needleleaf forests, and 68% in shrublands (Figure 3a). More generally, the fractional AUC for grassland wildfires is 0.57 (90% CI of 0.55–0.59; Figure 3b). The corresponding area for evergreen needleleaf forests is 0.53 (90% CI of 0.52–0.55), and for shrublands is 0.55 (90% CI of 0.53–0.57). Across all ecosystems, the increase is statistically significant ($p < 0.05$; Figure S3, where $p$ values are obtained using a one-sided test for a higher AUC). This increase is particularly notable because the VPD threshold was chosen such that VPD is no longer informative. Based on these results, we conclude that operational fire risk assessments will likely be improved by incorporating remotely sensed soil moisture.

Remotely sensed soil moisture outperforms various soil moisture indicators that are more indirect or regional. In particular, soil moisture outperforms the Keetch-Byram Drought Index (KBDI), an index used in

Figure 3. Soil moisture improves wildfire predictability when VPD is high. (a) Receiver operating characteristic curves for soil moisture estimated on days with high vapor pressure deficit (“HV”), which is defined as over 44.5 hPa in grasslands (GRA), 42.7 hPa in shrublands (SHR), and 27.2 hPa in evergreen needleleaf forests (ENF; Figure S2). Confidence intervals are estimated from bootstrapping sites with 1,000 iterations and represent the 5th and 95th percentiles of the true positive rate. (b) The associated area under the ROC curve (AUC) soil moisture and VPD on HV days. The box and whiskers represent AUC estimates from bootstrapping sites with 1,000 iterations, with the boxes indicating the 25th and 75th percentiles and the whiskers extending to the most extreme data points not considered as outliers, corresponding approximately 2.7 times the standard deviation. Outliers are represented with red dots and are defined as points greater than $q_3 + w \times (q_3 - q_1)$ or less than $q_1 - w \times (q_3 - q_1)$, where $w$ is the maximum whisker length and $q_1$ and $q_3$ are the 25th and 75th percentiles of the sample data.
some fire danger rating models that is based on meteorological observations alone (Alexander, 1990; Keetch & Byram, 1968), in grasslands and shrublands when VPD is high (Figure S5; see the supporting information for methods). KBDI was originally developed and parameterized for forest fires (Keetch & Byram, 1968), which may explain its better performance in forests relative to shrublands or grasslands. In grasslands, these results are consistent with results from Oklahoma, where there exists an extensive measurement network, showing greater skill in predicting grassland fires using direct monitoring of soil moisture relative to the Keetch-Byram Drought Index (Krueger et al., 2017).

4. Discussion

Our results demonstrate that using VPD and soil moisture in conjunction allows for improved predictions of fire risk (Table 1). Improved skill indicates that remotely sensed soil moisture captures distinct hydrologic information not reflected in the atmospheric VPD, especially in grasslands. Previous statistical analyses have shown that there is lower correspondence between interannual variations in burned area and VPD in grasslands (correlation coefficient, $r = 0.62$) than in forests ($r = 0.80$) (Seager et al., 2015), highlighting the potential to use remotely sensed soil moisture to also explain annual fire extent in grasslands.

The stronger predictive power of soil moisture in grasslands relative to forests and shrublands could be physiological or due to the limitations associated with L-band satellite-based measurements of soil moisture. Whereas forests tend to conserve water when VPD is high, grasslands instead generally continue to transpire (Teuling et al., 2010). When water becomes limiting, transpiration declines and, eventually, the grasses dry, lowering the live fuel moisture content. In situ soil moisture observations in Oklahoma grasslands also indicate that soil moisture is a driver of wildfires during the May–October growing season when it influences live fuel moisture content, but not during the dormant season when there is no influence on live moisture content (Krueger et al., 2015). Forests, of course, also eventually respond to soil moisture limitations (Teuling et al., 2010), but greater water storage in forests relative to grasslands suggests a longer and more variable response time.

Also possible is that remotely sensed soil moisture observations more accurately capture plant-available water in grasslands because the rooting depth of grasses is on average less than that for shrubs or forest (Fan et al., 2017). Because soil moisture observations from SMAP represent the near surface (~0–5 cm), we use estimates of deeper soil moisture that are based on these near-surface observations (Albergel et al., 2008; Ford et al., 2014). Soil moisture dynamics in the root zone may not be fully captured in these estimates, though, because near-surface soil moisture and deeper soil moisture are known to be more decoupled in arid regions (Hirschi et al., 2014), such as the Western United States. Furthermore, L-band penetration depths are shallower over forests compared to grasslands because of masking associated with signal propagation through canopy. Observations of deeper soil moisture, for example, from airborne, P-band sensors, may allow for improved wildfire prediction in deeper rooted ecosystems.

Observations of VPD and soil moisture are distinct, and their availability differs according to time and place. Soil moisture can be remotely sensed, including through cloud or smoke, from microwave satellite missions with relatively high accuracy (Ma et al., 2019), allowing for spatially resolved fire-risk assessment across the globe. In contrast, a network of weather stations is typically required for spatial mapping of VPD at the daily timescale. Remotely sensed near-surface VPD, on the other hand, is based on hyperspectral inversions that typically lack the spatial coverage, resolution, and all-weather capabilities of microwave soil moisture retrievals.

In addition to directly relating soil moisture to fire activity, as done in this study, there is also potential to predict live fuel moisture content as an intermediary explanatory variable (Jia et al., 2019; Wang et al., 2019). Predicting fuel moisture content over such a large spatial scale, such as the Western United States, is complicated due to the diversity of fuel types and inherent climate controls on vegetation phenology but should be explored in future work. Longer timescale analyses could also be explored using remotely sensed soil moisture in point of previous studies indicating that early snowmelt leads to decreased water availability and increased fire activity over monthly timescales (Jensen et al., 2018; Westerling et al., 2006). Furthermore, in arid regions, wet anomalies in soil moisture have also been found to increase fire activity later in the year by increasing biomass and fuel loads (Sungmin et al., 2020).
5. Conclusions
Our results show that remotely sensed soil moisture aids in predicting fire when VPD already indicates that risk is high (Figure 1). Dependence on soil moisture, as opposed to VPD alone, has the important implication of greater predictability of fire risk because soil moisture anomalies are more persistent than those of VPD (Seneviratne et al., 2006). Greater predictability of fire risk may be useful for optimizing resource allocation, such as moving aerial assets to more vulnerable regions prior to fire ignition (Fiorucci et al., 2005). This analysis serves as an initial demonstration that remotely sensed soil moisture can lend additional predictive skill with respect to wildfire occurrence.

Data Availability Statement
All data from this analysis are freely available. SMAP data are described in Entekhabi et al. (2010) and can be accessed from the National Snow and Ice Data Center (NSIDC) (https://nsidc.org/data/smap/smap-data.html). PRISM data are described in Daly et al. (2008) and can be accessed online (at https://prism.oregonstate.edu/). GeoMac fire origin data can be accessed from The National Interagency Fire Center (NIFC) (https://data-nifc.opendata.arcgis.com).

Acknowledgments
Financial support was provided by The Rockefeller Foundation Planetary Health Fellows program at Harvard University (A. R.) and the Harvard College Research Program (R. P.). We thank Andrew Feldman for providing the SMAP soil moisture data and Kelly Martin for her insightful comments.

References