SAR-enhanced mapping of live fuel moisture content

Krishna Rao[a,*], A. Park Williams[b], Jacqueline Fortin Flefil[c], Alexandra G. Konings[a]

[a] Department of Earth System Science, Stanford University, United States of America
[b] Lamont-Doherty Earth Observatory, Columbia University, United States of America
[c] Department of Civil and Environmental Engineering, Stanford University, United States of America

ARTICLE INFO

ABSTRACT

Assessing wildfire risk presents several challenges due to uncertainty in fuel flammability and ignition potential. Live fuel moisture content (LFMC) - the mass of water per unit dry biomass in vegetation - exerts a direct control on fuel ignitability, fuel availability and fire spread, and is thus an important parameter in assessing wildfire risk. Current estimates of LFMC from optical remote sensing or meteorological indices are insufficient to accurately map LFMC at landscape scale. In this study, we present a physics-assisted recurrent neural network model for mapping LFMC every 15 days at 250 m resolution over the western US using microwave backscatter (from Sentinel-1) and optical reflectance (from Landsat-8). The model was physics-assisted by adding handcrafted inputs representing radiative transfer processes and related parameters. Field data from the National Fuel Moisture Database was used to train the deep learning model. The model had a satisfactory overall predictability of $R^2 = 0.63$, RMSE = 25.0% and bias = 1.9% when cross-validated at 125 sites. The model was substantially better at predicting site-means ($R^2 = 0.71$ and RMSE = 14.3%) as compared to site-anomalies ($R^2 = 0.55$ and RMSE = 21.3%). Of the 6 land cover types tested, the model was most accurate in predicting LFMC in shrublands ($R^2 = 0.69$), while closed broadleaf deciduous sites had the least accuracy ($R^2 = 0.49$). Incorporating microwave backscatter as a model input significantly enhanced performance (R^2 increasing from 0.44 to 0.63, and RMSE decreasing from 31.8% to 25.0%). Microwave backscatter therefore presents a complementary source of information to optical remote sensing metrics for LFMC mapping. The high resolution dynamic LFMC maps presented here may assist in improved wildfire risk characterization and also have potential applications in quantifying plant drought stress at large scales.

1. Introduction

Wildfires play a critical role in Earth's water and carbon cycles (Bowman et al., 2009; Shakesby and Doerr, 2006). By releasing carbon stored in aboveground vegetation and soil, they emit CO_2 and other particulate matter into the atmosphere, that may adversely affect climate (Randerson et al., 2006), air quality (Crutzen and Andreae, 1990), and human health (Lelieveld et al., 2015). Large wildfires can cause catastrophic damage to property and human lives (Cal Fire, 2018). With climate change, probability of large wildfires has grown in much of the western United States (US) and is expected to continue to do so due to increases in concomitant hot and dry summer time climate that could lead to increases in fuel aridity (Abatzoglou and Williams, 2016; Williams et al., 2019) and associated wildfires (Holden et al., 2018). In the western US, large wildfires have become nearly five times more frequent and consume more than ten times the area now than five decades ago (Westerling, 2016). It is thus vitally important to estimate

and forecast forest wildfire risk.

Wildfire risk depends on three quantities - the availability of ignition sources, the propensity of vegetation and litter fuel to ignite, and the ease of fire spread once the fuel has ignited (Miller and Ager, 2013; Preisler et al., 2004). Live fuel moisture content (LFMC), defined as the mass of vegetation water per unit dry biomass is a key determinant of all the three components - wildfire ignition (Chuvieco et al., 2004a), fuel availability (Kelley et al., 2019) and fire spread (Rossa, 2017). Previous studies of large wildfires have further observed clear thresholds between fire size and LFMC (Chuvieco et al., 2009; Dennison et al., 2008; Nolan et al., 2016).

Even though LFMC strongly influences wildfire risk, its spatio-temporal variability remains poorly understood since LFMC has high spatial heterogeneity due to its dependence on plant species and physiological traits (Jolly and Johnson, 2018; Karavani et al., 2018). Thus, wall-to-wall dynamic maps of LFMC are urgently needed to understand its spatio-temporal characteristics, and ultimately to improve fire

danger assessments.

Current LFMC estimation methods predominantly rely on empirical relationships between LFMC and hydro-meteorological variables (e.g., Jia et al., 2019) or on radiative transfer models using multi-spectral satellite observations (e.g., Yebra et al., 2018). The operational model used by the United States Forest Service (USFS) for fire danger assessment (Bradshaw et al., 1984) is one example of such an empirical model. In this model, LFMC is estimated using lagged 1000-h duff fuel moisture content, which is itself empirically estimated from past precipitation, temperature and relative humidity.

Other models also exist that determine LFMC from meteorological variables but they are often unable to accurately predict LFMC ($R^2 < 0.3$; Ruffault et al., 2018) or constrained to just a few sites (Pellizzaro et al., 2007). In part, this is because LFMC estimates from climate-related variables cannot capture spatial variability in plant adaptation and responses to drought (Nolan et al., 2018). Thus, internal plant hydraulic state (LFMC) does not always mimic external climatic controls as plants have varied drought response strategies (Blackman et al., 2019).

On the other hand, several studies exist that use multi-spectral remote sensing to estimate LFMC either empirically (Argañaraz et al., 2016; Chuvieco et al., 2004b; Peterson et al., 2008) or from simulations using radiative transfer models (Barraza et al., 2014; Bowyer and Danson, 2004; Yebra and Chuvieco, 2009). These multi-spectral remote sensing-based approaches rely on absorptive properties of leaf water in near-infrared (NIR) or shortwave-infrared (SWIR) wavelengths (~900–2500 nm; Yebra et al., 2013). However, plant water content information derived from multi-spectral observations is representative of the top of the canopy only (Newton and Blackman, 1970; Sims and Gamon, 2003). Moreover, optical and IR reflectances are highly sensitive to vegetation characteristics such as canopy structure (Ceccato et al., 2001; Song, 2013) and leaf area index (Carlson and Ripley, 1997), due to which models tend to be extremely site-specific. Lastly, the dual relationship of optical data to both biomass as well as vegetation water has been identified as problematic in common situations where both the water status and biomass vary (Danson and Bowyer, 2004). Thus, very few studies have produced periodically updated large scale LFMC maps using multi-spectral remote sensing alone, with the notable exception of the LFMC maps for Australia created by Yebra et al. (2018).

Microwave remote sensing has the potential to overcome some of the drawbacks of optical remote sensing. Since their wavelengths ($\approx$ 5 cm) are at least four orders of magnitude higher, they penetrate much deeper into the canopy. As a result, they saturate less at high biomass levels (Tian et al., 2016). Additionally, unlike multi-spectral remote sensing, they usually penetrate through clouds (Ulaby and Long, 2014). Furthermore, owing to the high dielectric property of water (Pampaloni and Paloscia, 1986), they can capture vegetation water dynamics better than lower wavelength optical metrics. Indeed, studies have explored the predictive power of both passive (Fan et al., 2018) and active microwave remote sensing (Wang et al., 2019) for fire-related variables, although they were confined to site-scale analysis. The relevance of microwave remote sensing for large scale LFMC mapping across different land cover types, however, remains unexplored.

Owing to microwave and optical remote sensing's complementary sensitivities to both vegetation water and biomass, they may be combined to estimate LFMC, provided sufficient information related to canopy structure exists to help parametrize the relationship between remote observations and LFMC. This idea is conceptually presented in Fig. 1. The conceptual model links the microwave and optical inputs to the components of LFMC - vegetation water and dry biomass - through physical processes. By leveraging microwave and optical data's different levels of sensitivities to vegetation water and biomass, it may be possible to better estimate LFMC.

Much like for optical remote sensing, LFMC estimation efforts using microwave remote sensing can rely either on electromagnetic models or on empirical relationships. Although electromagnetic scattering models can simulate a parametric relationship between LFMC and remotely sensed observables, they need a large number of field data related to plant allometry (Ulaby et al., 1988), that is impractical to gather at large scales. Alternatively, a non-linear, empirical model like an artificial neural network can help capture the parametric relationship without explicitly assuming an analytical form a priori. Doing so allows us to rely on the data to understand the required complexity of the relationship between the input remote observations and LFMC output. Furthermore, using carefully handcrafted input variables chosen based on our understanding of radiative transfer processes, neural networks can be assisted by physics (Karpatne et al., 2017). Here, we investigated whether, by using a large number of LFMC field data to train a non-linear deep learning model with inputs from remote sensing (as well as terrain features that are known to influence the LFMC - remote sensing relationship), it is possible to produce dynamic large scale LFMC maps.

In this study, we present a physics-assisted empirical method to produce high resolution (250 m) LFMC maps across western US every 15 days. To estimate LFMC, we combine Sentinel-1 synthetic aperture radar (SAR) backscatter and Landsat-8 surface reflectance with static variables related to canopy structure in a recurrent neural network (RNN) deep learning model. The western US spans a large climatological range, with a diverse mix of 6 land covers, making it an excellent location to test the robustness of our algorithm. The LFMC maps may be used for wildfire danger assessments and other applications related to plant drought stress.

2. Methods

Our modeling approach stems from the need for large scale LFMC mapping without assuming LFMC's link to meteorological drivers. To overcome this, we employed an observationally-driven approach to directly estimate LFMC from remotely sensed optical and microwave inputs (explained in Section 2.4). We trained and validated the empirical model in the western US (see Section 2.1) on in-situ LFMC data collected by the USFS (detailed in Section 2.2).

We adapted the model to be physics-assisted through the use of carefully designed inputs based on our conceptual understanding of the physical processes linking the inputs to LFMC (Fig. 1). This was partially done through use of the microwave and optical ratios, which were chosen based on the form of radiative transfer models (as further discussed in Section 2.3.3). It was also achieved through the choice of ancillary static inputs (see Section 2.3.4), which provide information about the effect of soil type, vegetation structure, and topography on the remote sensing observations.

2.1. Study area

We chose the western US (consisting of 12 states) as our study area (Fig. 2a) because its fire regimes are undergoing rapid change. The area spans over 3.7 million km^2 and contains all states west of (and including) Texas, Colorado, Wyoming and Montana. Apart from being a regional hot-spot for increased fire activity (Westerling, 2016), the western US also covers diverse climatological, topographical, and ecological regions, making it an ideal choice for developing a robust LFMC mapping algorithm.

The area contains several types of vegetation, including broadleaf deciduous forests, needleleaf evergreen forests, shrublands, grasslands, and sparsely vegetated regions (displayed in Fig. S1 with data from the GLOBCOVER dataset; (Arino et al., 2010). Seasonal temperatures and precipitation range widely. Mean annual temperature (MAT) ranges from 0 °C to 24.5 °C, while mean annual precipitation (MAP) ranges from an arid 50 mm/yr to a very wet 5990 mm/yr (Fig. S2). Here, climate data was obtained from the Parameter-elevation Regression on Independent Slopes Model (PRISM Climate Group Oregon State University, 2004). The study area's terrain is also diverse, with both the
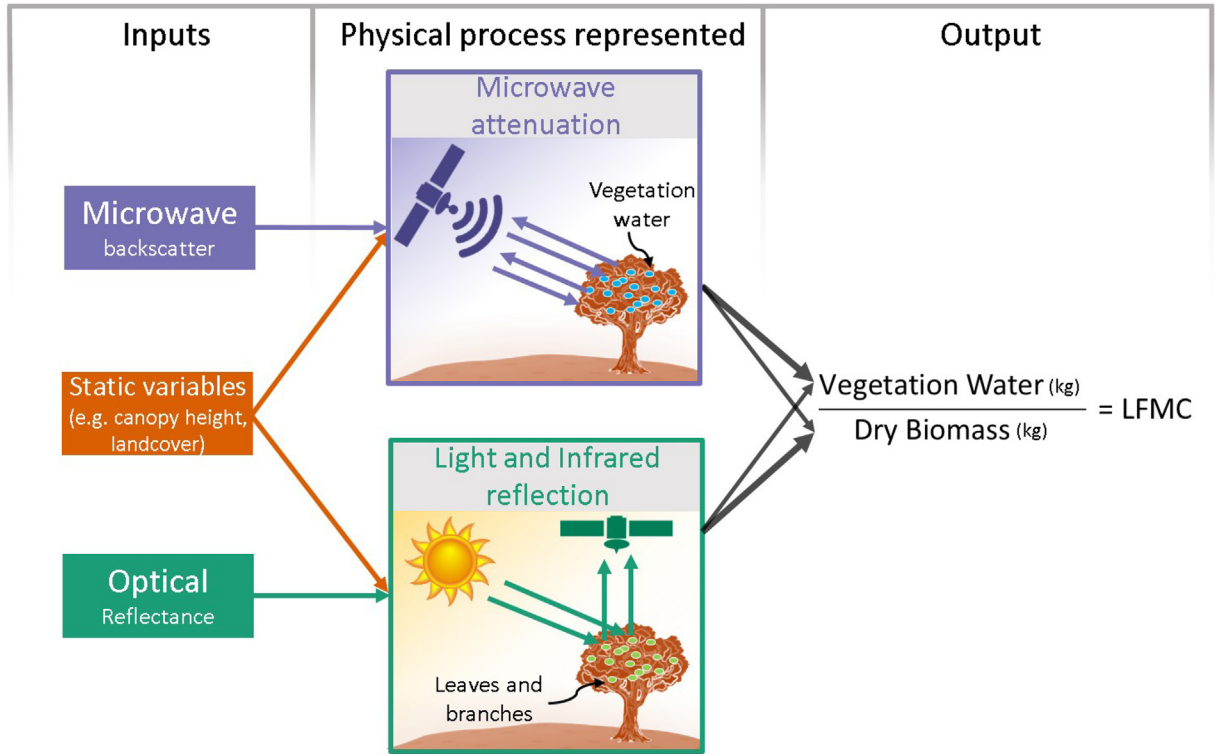
Fig. 1. Conceptual model linking LFMC to inputs variables. The physical process representation is for illustrative purposes only; the empirical model estimates LFMC directly from the inputs. Thicker arrows from the physical process representation to the output represent relatively greater sensitivity. For example, microwave backscatter, due to microwave attenuation, has relatively higher sensitivity to vegetation water as compared to optical and IR reflection.

lowest (−86 m) and highest points (4417 m) in the contiguous US located within it. The elevation data was obtained from the National Elevation Dataset (USGS, 2011) and is mapped in Fig. S2.

2.2. LFMC field data

Live fuel moisture content (LFMC) is defined as

$$LFMC(\%) = \frac{Fresh\,mass - Dry\,mass}{Dry\,mass} \times 100 \qquad (1)$$
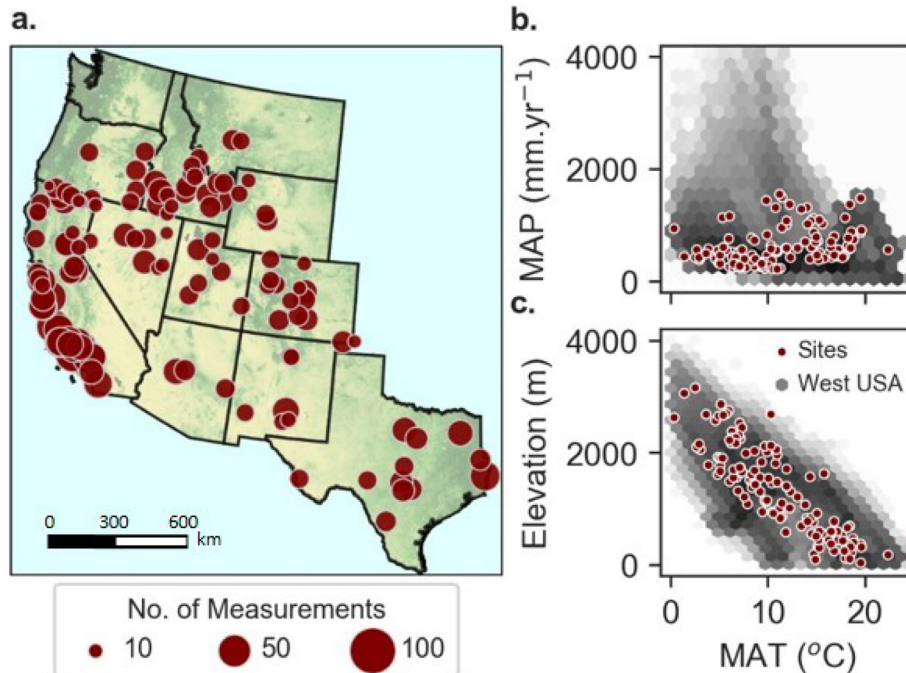

Fig. 2. Study area with LFMC sampling sites denoted as per number of measurements available since 2015 (a) with the same sites overlaid on the climatologic (b) and topographic range (c). Grey hexagons in (b) and (c) represent data for the entire study area, with darkness denoting increased point density.

where fresh mass is the mass of a vegetation element (leaf, twig, etc.) just after being cut, and dry mass is its mass after drying in an oven.

LFMC field data from the National Fuel Moisture Database (NFMD; United States Forest Sevices, 2018) was used for training and validation of the model. The NFMD is a free web-based query system containing more than 200,000 fuel moisture content measurements since 1977. The database spans 976 sites mostly over the western US, although only 239 sites contain data from 2015 onwards (the study period used here). According to the NFMD protocol, the field measurements were taken for representative species within 5 acre (20,000 m^2) plots. The measurements were made during mid-afternoon and on dry days when dew or precipitation were absent. We used LFMC measurements from January 1st, 2015 to March 1st, 2019 only, since Sentinel-1 observations used as an input to the model were not available earlier (see Section 2.3.1). LFMC values ranged from 16% to 320% during the study period.

At many sites, LFMC of multiple species were recorded. Since no information on the relative abundance of different species was available, scaling the individual species' LFMC to a pixel-representative LFMC was infeasible. Spatial heterogeneity in the LFMC of a mixed-species pixel can introduce severe errors in model predictions because of degradation of data quality (Jia et al., 2006). Hence, sites with multiple species were ignored except in cases where LFMC of multiple species closely followed each other throughout the study period (Pearson r between any two species $\geq$ 0.5). For sites where the seasonality of individual species matched, a simple average was assumed to determine the representative LFMC for those sites. Although this was prone to induce errors into the model, it allowed the inclusion of 66 more sites. In total, 114 sites were removed from the study (Pearson r between any two species < 0.5), 66 mixed-species were included (Pearson r between any two species $\geq$ 0.5), and 59 single-species sites were included in our study. Thus, a total of 125 sites were included (shown in Fig. 2a).

The sampling sites covered the diverse range of geography, elevation, and MAT of the western US (Fig. 2a and c). However, sampling sites were relatively dry. While MAP in the western US $\in$ [50–5990] mm/yr, the wettest NFMD site studied had an MAP of only 1560 mm/yr. This was consistent with expectations, since drier sites - being more prone to wildfires - tend to be more closely scrutinized by the USFS. Even though the sampling sites did not span the complete MAP range of the study area, no negative bias was observed. In fact, the mean MAP of sites (610 mm/year) exceeded that of the study area (560 mm/year). The effect of the diminished range of MAP of the sites is studied in detail in Section 4.3.

2.3. Model inputs

2.3.1. Microwave data

Microwave backscatter (σ) was obtained from Sentinel-1's synthetic aperture radar (SAR) owing to its high spatial resolution. The 2-constellation system of satellites (Sentinel-1A and 1B) contains a C-band (5.4 GHz) SAR which most commonly collects data in interferometric-wide swath mode in vertical-vertical (VV) and vertical-horizontal (VH) polarizations over land (Torres et al., 2012). The satellites also collect data in other acquisition modes like stripmap, extra-wide swath, and wave modes depending on the location and mission objectives. The radar's C-band frequency makes it ideal for LFMC estimation as the electromagnetic waves at that frequency are sensitive to foliage and thin woody components - the elements that were sampled for LFMC in the field and that contribute to fire spread substantially (Keane, 2015). Note, however, that C-band backscatter also retains sensitivity to soil moisture (Brocca et al., 2017). Because high-resolution soil moisture information was not available, we instead relied on the hypothesis that the RNN model can isolate vegetation from soil contributions to backscatter based on the relative variations of σ_{VV} and σ_{VH}. This hypothesis was based on the recognition that the two polarizations have different relative sensitivities to soil moisture and vegetation (Dubois et al.,

1995).

The native spatial resolution in interferometric wide swath mode is 5 m × 20 m. However, at that resolution, the image is noisy due to speckle formation. To reduce speckle noise and be consistent with the NFMD plot size, we rescaled the data to 250 m × 250 m pixels using bilinear interpolation.

Although the Sentinel-1 constellation has a revisit frequency (for the same overpass) of 6 days over Europe, it is 12 days for western US. We used ascending passes with a local overpass time of 18:00 only. Descending overpasses were excluded. Because vegetation water content exhibits a strong diurnal cycle, which is reflected in the radar backscatter (Konings et al., 2017a), only a single overpass time was used. The 18:00 overpass was used because vegetation is likely to be more water-stressed then than at pre-dawn, and because of availability of more images in 2015 and early 2016.

We used σ_{VV}, σ_{VH}, and $\sigma_{VV} - \sigma_{VH}$ in decibels (dB) as microwave inputs for the model. We used $\sigma_{VV} - \sigma_{VH}$ as it could be less sensitive to soil moisture than the raw backscatter (Dubois et al., 1995).

Even though Sentinel-1 has a constant look angle of 40°, the incidence angle on the ground can vary depending on location of a site within the swath. Thus, all backscatter values were normalized to the look angle of 40° using cosine square correction (Topouzelis and Singha, 2016).

2.3.2. Optical data

Landsat-8 16-day composites of surface reflectance (Roy et al., 2014; Vermote et al., 2016) were used to create optical inputs. The raw band reflectances for red, green, blue, near-infrared (NIR), and short-wave infrared (SWIR) channels at 30 m resolution were used. The NIR and SWIR were used because water absorbs energy in those spectral regions (Colombo et al., 2008). Apart from the raw band reflectances, three additional vegetation indices were used: the normalized difference vegetation index (NDVI; Tucker, 1979), the normalized difference water index (NDWI; Gao, 1996), and near infrared vegetation index (NIR$_v$; Badgley et al., 2017). Previous studies have shown a strong correlation between NDWI and LFMC (Dennison et al., 2005; Roberts et al., 2006), and a strong predictive capacity of NDVI for LFMC (Chuvieco et al., 2004b; Hardy and Burgan, 1999) (though note that NDVI is a measure of greenness, not water content). Although other metrics like visible atmospheric resistance index (VARI; Gitelson et al., 2002) and normalized difference infrared index (NDII; Hardisky et al., 1983) also exist, we chose to include NDVI and NDWI only, as they already contain NIR and SWIR (the bands most sensitive to vegetation water in the optical domain; Colombo et al., 2008). Not adding more indices with the same bands helps reduce co-linearity among the inputs. Lastly, NIR$_v$ was used as an indicator of vegetation biomass levels as it is related to the carbon assimilation from photosynthesis (Badgley et al., 2017), and may therefore be useful for disentangling biomass and LFMC effects on the Sentinel-1 backscatter.

Surface reflectance values were filtered for snow and clouds. To do so, pixels with a quality assessment index indicating "snow", "cloud", or "cloud shadow" were masked.

2.3.3. Microwave and optical ratios

Apart from including pure microwave and optical data as inputs, some mixed inputs containing the ratios of microwave backscatter and optical reflectance were also used. We pursued a physics-guided approach to enhance the learning capabilities of the RNN model (Karpatne et al., 2017; Reichstein et al., 2019), reasoning that these ratios should improve isolation of LFMC. With higher sensitivities of microwave backscatter to vegetation water and optical reflectance to dry biomass, the ratio of microwave backscatter and optical reflectance mimics the structure of the definition of LFMC (Eq. (1)). We thus included ratios of microwave backscatter of both polarizations and optical reflectances and indices.

2.3.4. Ancillary inputs

We assisted the RNN model by inputting several static inputs drawn from our understanding of physical principles. Since these inputs do not vary over time, they exist merely to assist the model in learning the radiative transfer process between temporally varying inputs and LFMC. We chose inputs only that were easily available on a large scale to preserve our model's scalability.

Static inputs such as silt, sand, and clay content were used as they govern microwave backscatter's sensitivity to soil moisture (Dobson et al., 1985). By adding soil type as a static parameter, we assisted the RNN model in isolating the vegetation-related information from microwave backscatter. Data on soil type was obtained from the Unified North American Soil Map by Liu et al. (2014) which combines soil maps from the US General Soil Map (USDA, 2019), Harmonized World Soil Database (Fischer et al., 2008) and other sources. The map's resolution is ~25 km.

Static inputs related to vegetation structure were also included in the model. Canopy height based on Global Laser Altimetry System (GLAS) lidar measurements aboard the IceSat satellite obtained from Simard et al. (2011) was included (spatial resolution = 1 km). Land cover at 300 m spatial resolution obtained from GLOBCOVER was included as well.

Finally, elevation and slope obtained from the National Elevation Dataset (USGS, 2011) was used to help the model calibrate to local terrain, as local incidence angle can affect the backscatter's parametrization to vegetation water (Attema and Ulaby, 1978). Elevation and slope data was obtained at 30 m spatial resolution. All static and dynamic inputs used in the model are summarized in Table 1.

2.4. Recurrent neural network

The relationship between remote sensing observables and LFMC can be highly non-linear, as shown by radiative transfer simulations of leaf optical properties (Jacquemoud and Baret, 1990) and microwave scattering simulations of vegetation components (Ulaby et al., 1988). We therefore used a neural network to model LFMC. A neural network consists of several layers, each in turn composed of many neurons that combine the inputs linearly with scalar weights and pass it through a non-linear activation function. LFMC is a time-dependent variable for which only certain successive sets of values are realistic. The possible values depend on previous LFMC values and their associated remote sensing observables. We therefore used an RNN based on long-short term memory architecture (LSTM; Hochreiter and Schmidhuber, 1997), which allows propagating memory of essential features in time.

The overpass days for Sentinel-1 (12-day revisit) and Landsat-8 (16-day composite) at a location often differ by a few days. We thus linearly

Table 1

Input variables passed to the recurrent neural network. All variables except Static variables (last column) are temporally varying. The NIR refers to near-infrared, NIR_v refers to near-infrared vegetation, NDVI refers to normalized difference vegetation index, and NDWI refers to normalized difference water index.

Microwave (dB)	Optical (−)	Mixed (dB)	Static
σ_{VV}	r_{red}	$\frac{\sigma_{VV}}{r_{red}}, \frac{\sigma_{VH}}{r_{red}}$	Canopy height (m)
σ_{VH}	r_{green}	$\frac{\sigma_{VV}}{r_{green}}, \frac{\sigma_{VH}}{r_{green}}$	Elevation (m)
$\sigma_{VH} - \sigma_{VV}$	r_{blue}	$\frac{\sigma_{VV}}{r_{blue}}, \frac{\sigma_{VH}}{r_{blue}}$	Terrain Slope (°)
	r_{NIR}	$\frac{\sigma_{VV}}{r_{NIR}}, \frac{\sigma_{VH}}{r_{NIR}}$	Silt fraction (−)
	r_{SWIR}	$\frac{\sigma_{VV}}{r_{SWIR}}, \frac{\sigma_{VH}}{r_{SWIR}}$	Sand fraction (−)
	NDVI	$\frac{\sigma_{VV}}{NDVI}, \frac{\sigma_{VH}}{NDVI}$	Clay fraction (−)
	NIR_v	$\frac{\sigma_{VV}}{NIR_v}, \frac{\sigma_{VH}}{NIR_v}$	Land cover (−)
	NDWI	$\frac{\sigma_{VV}}{NDWI}, \frac{\sigma_{VH}}{NDWI}$	

interpolated the time series inputs to the 1st and 15th day of each month. Having a consistent time step (15 days) is a requirement for the RNN model used in our study. We analysed the amount by which the different time series were modified due to interpolating them to the nearest 1st and 15th day of each month. We found that a maximum change of 6.7% and 3.0% was introduced in the optical and microwave data, respectively. Thus, the interpolation itself did not change the input data by much.

The network topology contained time series of length seven. That is, it estimated LFMC for any 15-day interval (t to $t + 15$) using input variables for the period $t-90$ to $t + 15$ in 15-day increments (six previous time steps and the current one). The estimated LFMC for any 15-day interval was then compared to the field-sampled LFMC measured on one of the 15 days to calculate model error. Although estimating a 15-day averaged LFMC value can be error-prone to rapid fluctuations in LFMC, we are limited by the coarse temporal resolution of the remotely sensed inputs. The errors associated with the temporal averaging are discussed in Section 4.1.

For cases with missing Landsat-8 reflectance (due to snow cover, cloud cover, or missing swaths) or Sentinel-1 backscatter (due to missing swaths), the last valid observation was used. If no valid observations were available for the entire 3 months prior to an LFMC measurement (even for one of the input time series), the LFMC measurement was discarded. The 3-month period was chosen as the RNN model relies on information 3 months prior to LFMC observation. Although optical reflectance may vary within the 3-month window, if the RNN model is trained to convergence, it is expected to rely less on potentially outdated input variables.

The classical RNN model was modified to accept time series as well as static inputs. This was achieved by passing the static inputs directly to the activation function within each neuron. This way, we let the model learn the parametrization of the relationship to local vegetation structure and topography by itself.

We used a 4-layered RNN with each layer containing 10 units with hyperbolic tangent activation function. The model architecture is summarized in Fig. S4. A dropout of 5% was used for each layer to regularize against overfitting. Additionally, a bias regularization of 0.001 was used after initial training led to biased results. The model was fit to minimize mean squared error between the estimated and observed LFMC. To quicken the training process, a Nesterov-Adam optimizer was used which dynamically adjusted the learning steps by incorporating a momentum function (Dozat, 2016).

2.5. Validation and prediction

Empirical approaches such as neural networks are known to over-fit to the training data, rendering the training data unusable for calculating the model's performance. We therefore used 3-fold cross-validation to validate the model. That is, the model was trained iteratively on two-thirds of the sites (83 sites) and validated on the remaining one-third (42 sites) until out-of-sample LFMC estimates were available at all the sites (125 sites). Splitting the data by sites allowed testing the model's performance on sites that were completely different from the training sites. This provides a realistic estimate of the model performance at locations where no validation data was available, which make up the overwhelming majority of the study region. The data was split into training and validation sets using stratified random sampling to ensure that the distribution of land cover types remained equal between the two sets. Model accuracy was calculated on the out-of-sample validation estimates only. The model was validated on 3,020 LFMC measurements.

We analysed model errors by decomposing the predictions into site-mean LFMC ($\overline{LFMC}$, which varies spatially only) and site-anomalies ($LFMC'$, which varies spatially and temporally).

$$\overline{LFMC} = \frac{\sum\limits_{t=1}^{T} LFMC(s,t)}{T(s)} \qquad (2)$$

$$LFMC' = LFMC(s,t) - \overline{LFMC(s)} \qquad (3)$$

where s denotes the spatial index of a site, t denotes the LFMC measurement/estimation date, and $T(s)$ is the record length of site s.

The trained model was forward propagated to estimate LFMC across the entire western US. The predictions were restricted to locations with land cover classes among those that were already present in the training data and to places where valid remote observables were available at least once in the 3 months prior to the estimation. The modeling workflow is summarized in Fig. S3. All modeling was performed on Python v.3.6 using Keras library (Chollet et al., 2015) and the LFMC maps were created using Google Earth Engine (Gorelick et al., 2017).

2.6. Microwave and static inputs' contribution to model performance

The contribution of microwave data towards model performance was tested by excluding all microwave-related variables from the inputs - raw backscattering coefficients and mixed ratios (columns 1 and 3 in Table 1). The model was then retrained on the remaining optical data and static inputs, keeping all other model parameters constant. The degradation in model's accuracy was then used to quantify the contribution of microwave data to the original model performance. A similar approach was also followed for static inputs (column 4 in Table 1) by removing each variable with replacement.

3. Results

During 3-fold cross-validation, the model displayed a strong ability to estimate LFMC with coefficient of determination (R^2) = 0.63, root mean squared error (RMSE) = 25.0%, and bias = 1.9% (Fig. 3a). Although on an absolute scale, the RMSE of the model (25.0%) was moderately high, it was much lower than the standard deviation (SD) of LFMC (41.6%), indicating that the model performed comfortably better than a baseline of predicting average LFMC everywhere at all times.

The model predicted $\overline{LFMC}$ with R^2 = 0.71, RMSE = 14.3% (Fig. 3b), while $LFMC'$ performance was somewhat poorer with R^2 = 0.55, RMSE = 21.3% (Fig. 3c). That is, the model captured spatial variations in LFMC better than temporal variations.

The loss in temporal predictability can largely be attributed to three factors. First, 13% of sites had a bias larger than 20% of the site-mean (points away from grey line in Fig. 3b). Second, even though the model estimated the general pattern in site coefficient of variation (CV) well (Pearson r = 0.74; Fig. 4a), it consistently under-predicted CV (bias = −0.1). This implied that the model's LFMC estimates routinely varied less temporally than observations. Lastly, at sites with low CV (low relative temporal variability), the model was unable to capture
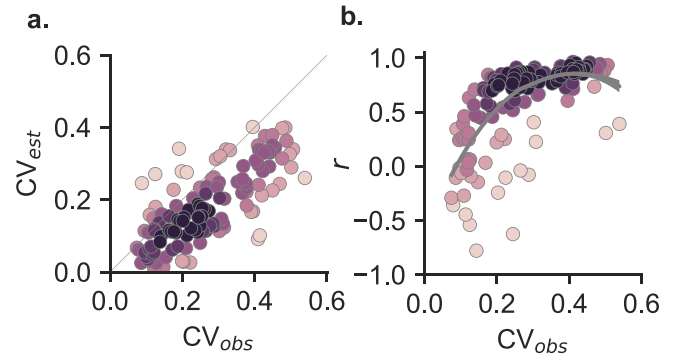

Fig. 4. Influence of site-scale variations on model performance. (a) Estimated and observed coefficient of variation (CV) at each site. The CV is the standard deviation divided by mean LFMC at the site scale. (b) Pearson correlation coefficient (r) between estimated and observed LFMC at each site versus observed CV. The curved grey line indicates best fit with shading of 1 standard deviation on each side. Darkness of points indicates higher point density.

temporal fluctuations altogether, resulting in low or negative Pearson r (Fig. 4b). This was likely caused by difficulty in inferring minute LFMC variations from satellite observations.

Skewed distribution of LFMC field sampling can introduce systematic errors in model predictions. As expected, the LFMC field samples were found to be concentrated towards June–October which overlaps with the fire season in western US (Fig. 5a). Model errors were higher in months with lower training data and in months immediately following low-data months (Fig. 5b). The latter was due to the fact that the RNN model relies on information up to 3 months prior to a measurement.

Fig. 6a shows the distribution of the site-specific RMSE across land cover. Also shown are sample prediction time series across sites with a range of performance quality levels (relative to other sites): the sites shown are the 5th percentile, median and 95th percentile site in terms of their site RMSE (Fig. 6b, c, and d, respectively). Individual site RMSEs ranged from 6.94% at a closed needleleaf site to 72.0% at a grassland site. Mixed forests had the lowest overall RMSE (20.0%). On comparing the RMSE at each site with its unbiased-RMSE (ubRMSE), we found that the model had high biases in many sites, especially in shrub/grassland and grasslands where 17% and 31% of the sites had bias larger than 20% of site-means, respectively (Fig. 6a).

Fig. 6b-d show that the model successfully captured LFMC's temporal dynamics. The sample time series show that at the 5th percentile site based on site RMSE, the model accurately estimated LFMC dynamics (Fig. 6b), while LFMC estimates were increasingly biased at the poorer sites (Fig. 6c and d).

Table 2 summarizes the model's performance across land covers. Shrub/grassland land cover had the highest mean bias of 5.4%, likely
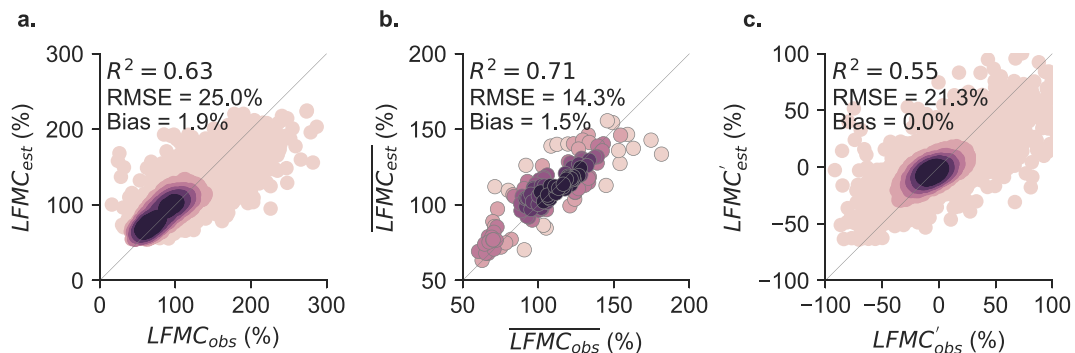

Fig. 3. Model Performance on the validation data. Model performance is shown for raw data (a), site means (b), and site anomalies (c). The y-axis represents out-of-sample cross-validated estimates, and x-axis represents field measurements. Grey line indicates the 1:1 line. Darkness of points indicates higher point density.
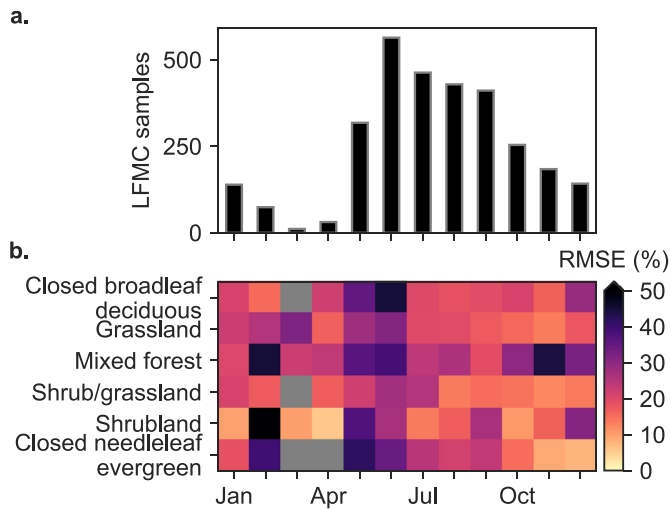
Fig. 5. LFMC field sampling bias and its effect on model performance. Grey indicates no data available.

due to the very low number of sites with this land cover (6) and low associated number of field measurements (178) available for training the model. Model performance in grasslands had the highest RMSE (31.0%) and a moderate R^2 (0.56), possibly due to the high inherent temporal variance in grasslands (temporal SD = 33.1%). This was consistent with previous LFMC modeling studies that observed very high temporal variance in LFMC in grass, making it difficult to model its rapid fluctuations (García et al., 2008).

In terms of variance explained on the validation set (R^2), the model estimated LFMC with high accuracy in shrublands, closed needleleaf evergreen, and mixed forest land cover types. Even though shrublands constituted only 22% of the sites, they comprise 45% of the western US (Fig. S5) making it a crucial land cover for this study. The model performed most accurately in shrublands ($R^2 = 0.69$).

To evaluate the importance of microwave data for LFMC estimations, all microwave-related variables were removed and an alternate model was trained on the remaining inputs. Model performance degraded substantially. The R^2 decreased from 0.63 to 0.44 (a drop of

Table 2
Model performance across vegetation types. Number of sites and number of field samples is denoted by N_{sites} and N_{obs}, respectively. Temporal SD refers to mean of field-sampled LFMC's standard deviation at each site within the respective land cover class.

Land cover	N_{sites}	N_{obs}	Temporal SD (%)	RMSE (%)	R^2	Bias (%)
Closed broadleaf deciduous	9	218	24.7	27.7	0.49	3.0
Shrub/grassland	6	178	29.1	24.9	0.56	5.4
Grassland	16	362	33.1	31.0	0.56	2.8
Mixed forest	9	247	19.8	20.0	0.59	1.0
Closed needleleaf evergreen	58	1,540	24.2	22.4	0.61	1.1
Shrubland	27	475	48.4	28.9	0.69	2.4
Overall	125	3,020	29.9	25.0	0.63	1.9

Table 3
Value of microwave-related time series and static variables for LFMC's predictability. "w/o" denotes "without". For e.g., the row for "All w/o elevation" contains model performance when elevation was removed from the inputs and the model was retrained on the remaining variables as in Table 1.

Input variables	RMSE(%)	R^2
All w/o microwave	31.8	0.44
All w/o elevation	29.4	0.49
All w/o terrain slope	27.2	0.56
All w/o land cover	26.5	0.57
All w/o canopy height	26.0	0.60
All w/o silt fraction	25.2	0.62
All w/o sand fraction	25.1	0.62
All w/o clay fraction	25.1	0.62
All	25.0	0.63

30%) and RMSE increased from 25.0% to 31.8% (an increase of 27%; Table 3). To check whether the reduction in accuracy was solely due to the decreased number of features, we separately trained a model on optical features and 17 time series of random numbers (replacing an equal number of microwave-related variables) and found a similar drop in accuracy ($R^2 = 0.41$, RMSE = 32.6%). This drop in performance confirmed that microwave data greatly improved LFMC estimation,
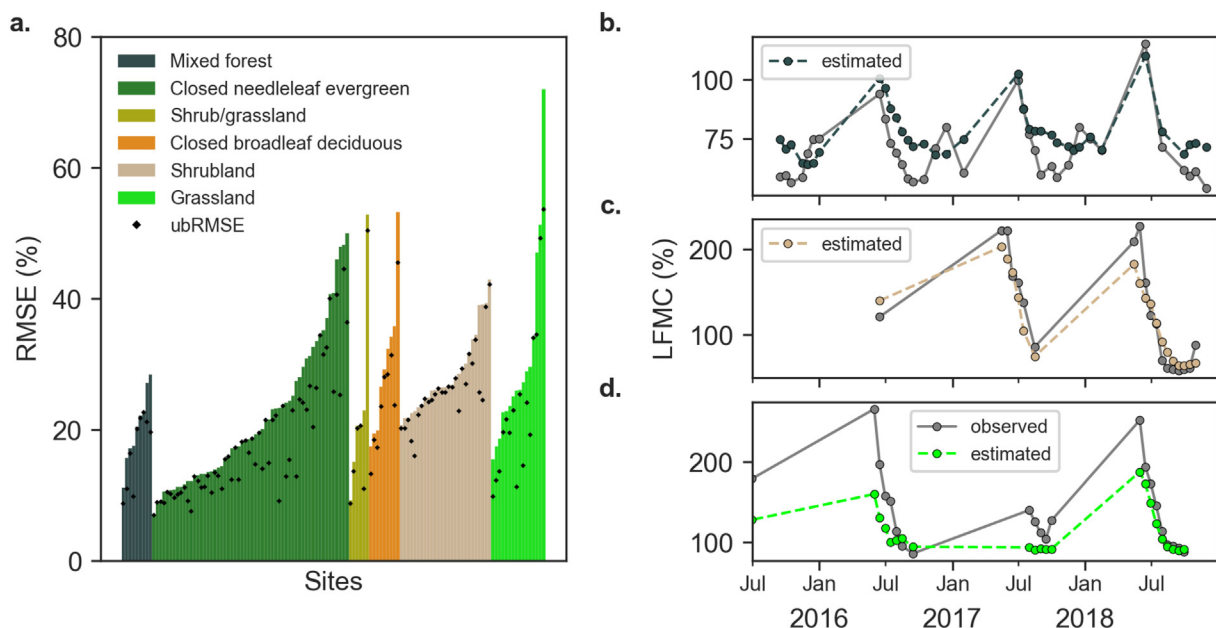

Fig. 6. Model performance at the site scale. (a) Model RMSE (vertical bars) compared to ubRMSE (points) at each site. Observed and estimated time series of LFMC at the 5th (b), median (c), 95th (d) percentile site, respectively, according to model RMSE rank at the site-scale as seen from (a).

Table 4

Model performance compared against other observationally-driven LFMC estimation techniques in literature. Values for Qi et al. (2012) retrieved for the pooled model across all 10 sites with best remote sensing proxies. Values for Chuvieco et al. (2002) obtained for single regression for shrub and grass species. Values for Peterson et al. (2008) obtained for regression for all chapparal and all coastal sage scrub species separately (combined regression not available). Values for Yebra et al. (2018) obtained for validation sites.

Study	Species	Sites	R^2	RMSE(%)
Qi et al., 2012	2	10	0.27	28.1
Chuvieco et al., 2002	5	6	0.63	5.1–128.0
Peterson et al., 2008	6	14	0.74–0.85	13.5–22.4
Yebra et al., 2018	17	32	0.58	40.0
Our approach	56	125	0.63	25.0

even in the presence of traditionally used optical metrics.

Static variables that assist the model parametrize the relationship between remote observables and LFMC also contributed to model performance. Elevation and terrain slope were among the most important static parameters for LFMC predictability, without which, model RMSE deteriorated to 29.4% and 27.2%, respectively. This could be due to the key role that topography plays in determining local incidence angle that affects backscatter's parametrization to vegetation water (Attema and Ulaby, 1978). A similar result was also found by Saatchi et al. (2007) while estimating forest fuel load from radars. Soil properties had the least effect on model performance.

Overall, our model's estimation accuracy was on par with previous LFMC mapping approaches (Table 4), in spite of being validated at much more sites (125) with much diverse species (56). When we compared our model performance to that of the next best study in terms of the number of validation sites (Yebra et al., 2018), we found that our model outperformed it, although a direct comparison was not possible due to differences in geographies and modeling strategies. Furthermore, for a more direct comparison, we compared our method to Qi et al.'s (2012) empirical LFMC estimation study in Utah which used several single input variables such as NDVI, NDWI, etc. Our model had both an improved R^2 (0.49 relative to 0.27 in the Qi et al.'s (2012) study) and a lower RMSE (25% instead of 28%).

Once the model was trained and validated, we produced 15-day LFMC maps from 2015 to 2019 for the western US at 250 m resolution. A series of estimated LFMC maps across the progression of the fire season of 2019 is shown in Fig. 7. The spatial patterns in Fig. 7 such as

the dryness in vegetation in Southern California, high wetness in the Pacific Northwest, etc. were consistent with expectations, suggesting that the model captured the spatio-temporal dynamics across the western US well.

4. Discussion

4.1. The cost of large-scale LFMC estimation

The physics-assisted RNN model was successful in estimating LFMC at landscape-scale over the entire western US while maintaining a cross-validated accuracy ($R^2 = 0.63$, RMSE = 25.0%; Fig. 3a) that matched previous models' accuracy (Table 4). While Myoung et al. (2018) demonstrated an $R^2 = 0.72$ across 7 sites, their study did not include cross validation. Furthermore, most previous predictive LFMC mapping efforts have focused on ecosystem-scale analysis (6–12 sites spanning a single region). It remains to be seen how robust such models are when expanded to larger areas such as our study region (125 sites spanning several ecosystems), especially since empirical models can be extremely site specific (Dennison et al., 2005; García et al., 2008). To the best of our knowledge, the only previous regional-scale analysis was Yebra et al.'s (2018) study. Their study was validated at 32 locations across Australia and had an $R^2 = 0.58$ and RMSE = 40%. Their mapping effort differed from the one in this study in two ways: it used a radiative transfer model rather than an empirical modeling approach, and used only optical frequency observations as dynamic inputs. But given the fact that our estimation errors were lower and our model was validated at approximately 4 times more sites with more diverse vegetation types, empirical approaches such as the one presented here exhibit considerable potential for accurate, large scale LFMC mapping. This is mainly due to the nature of physics-guided neural networks that can robustly scale across diverse landscapes if sufficient (and representative) training samples exist.

Our model did not capture temporal variability as well as spatial variability. Although $\overline{LFMC}$ was predicted with an accuracy of $R^2 = 0.71$, $LFMC'$ was predicted with lower accuracy of $R^2 = 0.55$ (Fig. 3b and c). The CV of LFMC estimates were consistently lower than those of observations (bias in CV = -0.1). However, the general pattern in site CVs was captured (Pearson $r = 0.74$, Fig. 4a) indicating that the systematic bias was likely due to the model setup (and not model fitting). In our model setup, remote sensing observations are interpolated to 15-day windows. Any LFMC variations at finer temporal scales (within
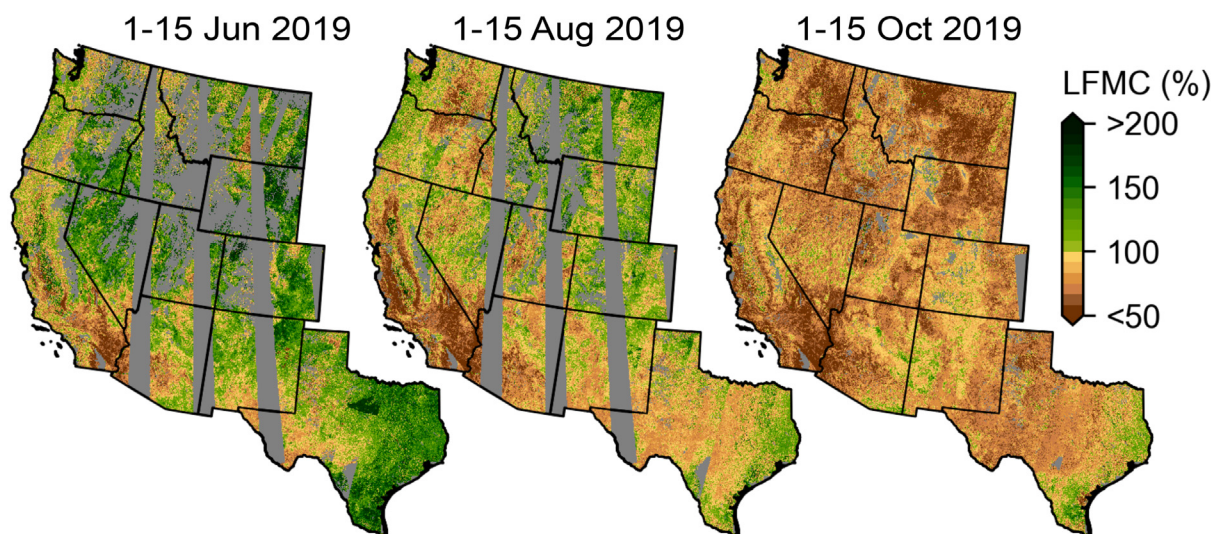

Fig. 7. Estimated LFMC for the first 15-day periods of June, August, and October of 2019. Grey pixels indicate LFMC estimates were unavailable. LFMC estimates were unavailable when Sentinel-1 or Landsat-8 cloud- and snow-free surface reflectance were unavailable in the 3 months prior to time of estimation or when the land cover class of a pixel was absent from the training data.

15 days) are thus expected to be smoothed out, resulting in consistently lower variability in the estimates.

For sites where local fluctuations in LFMC were relatively small (e.g. CV ≤ 0.2), the estimated LFMC was often weakly or negatively correlated to observations (Fig. 4b). While predicting temporal dynamics at sites with low LFMC variability is less relevant to fire danger assessments (locations with no variability have constant fuel characteristics and fire risk depends on ignition availability), future large scale LFMC mapping efforts could benefit from explicit site-based calibrations. Nonetheless, training an empirical model under a single loss metric to obtain an overall high accuracy at large spatial scale (Fig. 3a) comes at the expense of mis-representing site-scale fluctuations at sites with relatively little LFMC variability.

Mixed-species pixels presented another cost of large-scale LFMC mapping that is a weakness in this study - uncertainty of predictions in heterogeneous pixels. We trained and validated the model on both single-species sites and mixed-species sites with similar LFMC seasonal patterns (Fig. 2). However, the actual study area also contains many locations with a diversity of species that have very different LFMC patterns (indeed, 48% of all LFMC sites were excluded for this reason, because the LFMC patterns of different species at the site had an average Pearson $r < 0.5$, as discussed above in Section 2.2). Due to the lack of any high resolution species coverage information, cross-validation at such sites is infeasible. The model performance in pixels with heterogeneous species is therefore unknown. However, when we compared model performance between single-species sites (59 sites) and mixed-species sites (66 sites), we found that model performance remained largely unchanged (for single-species sites, RMSE = 27.3%, R^2 = 0.66 and for mixed-species sites RMSE = 22.9%, R^2 = 0.58). The similarity in model errors between single and mixed-species sites suggests that overall model predictability may not be very sensitive to pixel-scale species diversity. Additional validation in areas where detailed species coverage information is known is still needed. Nevertheless, given the similar performance between the two types of validation sites, we suggest here that the resultant LFMC maps be interpreted as an area-weighted representation of the LFMC within each pixel. Furthermore, since LFMC is mapped at a relatively high resolution of 250 m, errors due to heterogeneity in species cover are expected to be lower than for kilometer-scale models (Myoung et al., 2018).

4.2. Microwave backscatter improves LFMC estimation

After excluding all microwave-related inputs, the model performance degraded markedly with R^2 decreasing from 0.63 to 0.44 and RMSE increasing from 25.0% to 31.8% (Table 3). Our results indicate that by combining microwave data with optical data, and handcrafted ratios of microwave and optical data, LFMC mapping can be enhanced significantly. This provides direct evidence of the complementary nature of the two datasets.

Previous efforts to model LFMC from satellite observations have predominantly focused on optical remote sensing. Optical reflectance metrics such as NDVI are primarily related to photosynthetically active chlorophyll content. The LFMC estimations using NDVI thus rely on the relationship between chlorophyll content and leaf water content for their predictive ability (Tucker, 1979). But changes in vegetation wetness are not always captured in their color. For instance, when Dennison et al. (2005) compared the LFMC predictive strengths of NDVI with NDWI, they found that NDWI performed substantially better, because NDWI is constructed using longer wavelength SWIR and has higher sensitivity to vegetation water. In this work, we showed that a similar principle is at play for the use of microwave observations (which penetrate more deeply than SWIR).

Wang et al. (2019) showed that Sentinel-1 microwave backscatter can be used for LFMC estimations using a coupled soil backscatter and vegetation water cloud model. Although their study was limited to 3 sites, they presented process-based evidence for microwave's sensitivity

to LFMC. Their site-scale findings combined with the landscape-scale empirical evidence presented here confirm the importance of microwaves for LFMC estimation.

Our study maps LFMC only for foliar, herbaceous and fine woody fuels, as the training data from NFMD covers thin vegetation elements only. Sentinel-1's C-band radar backscatter observations are expected to be sensitive to these vegetation elements. However, LFMC of denser woody fuels can be estimated using longer wavelengths such as L-Band (1–2 GHz; Leblon et al., 2002; Tanase et al., 2015). Even though there are no current publicly-available L-Band SAR satellites with frequent revisit, the L-Band NASA-ISRO Synthetic Aperture Radar (NISAR; Rosen et al., 2015) - which is expected to be launched in 2022 and have a revisit of 12 days - could open up new opportunities for large scale mapping of woody LFMC.

4.3. The potential of LFMC maps for fire danger assessment

Land cover plays a dominant role in determining fire risk since it directly governs fuel characteristics. The selectivity of fires to some land cover types warrants the need for a robust model with high accuracy across all land covers (Barros and Pereira, 2014). Our model performance ranged from $R^2 = 0.49$ in closed broadleaf deciduous vegetation to $R^2 = 0.69$ in shrublands (Table 2). Shrublands are some of the most fire prone vegetation in Mediterranean ecosystems, largely due to their high ignition potential (Moreira et al., 2009). Moreover, the western US consists of over 45% of shrublands (Figs. S1, S5) making it a vital fuel class to model. The model's high performance in shrublands makes our LFMC maps even more relevant to wildfire studies.

A novel feature of our RNN model is the absence of any assumptions regarding how different species respond differently to meteorological controls. Rather, using an observation-driven approach, LFMC is estimated from microwave and optical remote sensing's direct sensitivity to vegetation water and biomass, respectively. Hence, the LFMC maps generated by the empirical model are entirely dependent on the field data used to train it, and their representativeness. For instance, since LFMC field samples were predominantly collected during the fire season only (Fig. 5a), model errors were larger during winter months when sufficient training data was unavailable. However, understanding LFMC's spatio-temporal dynamics during the fire season is more important and model errors were consistently low for all land covers during then (Fig. 5b).

Since the field data was derived from sites spanning the whole topographical and climatological range of the western US (Fig. 2b and c), we expect model errors over the study area to be close to our validation error. We investigated whether the sampling sites' clustering towards dry regions (Fig. 2b) could introduce systematic errors in LFMC estimation over the entire study region by comparing site-scale RMSE to MAP, MAT and elevation (Fig. S6). We observed that at the sampling sites, no statistically significant relationship existed between site errors and MAP (*p*-value = .55) suggesting that the true model error over the entire study area (i.e. every pixel in western US) is not expected to be affected by the concentration of the training and validation sites in relatively drier regions. Thus, due to the climatological and topographical representativeness of the NFMD sites, the LFMC maps produced here may be illustrative of LFMC dynamics across the entire western US.

Despite the fact that the sites were evenly distributed in climatologic and topographic space, we did find that shrublands were under-represented in the NFMD (22% of sites were shrublands, but 45% of the study area is; Fig. S5). This under-representation combined with high model accuracy at shrubland sites ($R^2 = 0.69$) may lead to a greater accuracy across the entire study region, making the resultant LFMC maps suitable for studying wildfire risk.

4.4. Alternate applications: drought stress

Although LFMC is primarily used as a fuel wetness indicator in the context of wildfires, it fundamentally represents the hydraulic state of vegetation. If independent estimates of dry biomass are available, it can be related to many other plant hydraulic metrics such as canopy water content, relative water content and plant water potential (see Table 1 in Konings et al., 2019).

Since vegetation hydraulic state can depend on many local, highly heterogeneous vegetation traits (for e.g., soil water distribution factors, roots, xylem and stomata; Anderegg et al., 2016), it cannot be estimated from hydro-meteorological conditions alone. Models rarely have enough information about spatial variability in vegetation traits to correctly estimate it. Dynamic maps of vegetation hydraulic state can enable large scale modeling of vegetation water stress (Brodrick et al., 2019), vegetation drought tolerance (Konings et al., 2017b) and drought-driven tree mortality (Rao et al., 2019).

Microwave remote sensing has undeniably helped in inferring vegetation properties such as forest biomass (Saatchi et al., 2011), isohydricity (Konings and Gentine, 2017), canopy water content (Saatchi et al., 2013), and crop parameters (Steele-Dunne et al., 2017). However, most of these applications have been hampered by the low (tens of kilometers) spatial resolutions of spaceborne radiometers and scatterometers.

SAR offers a unique opportunity to dramatically increase the spatial resolution of observations, but there has been limited success in large-scale efforts to isolate vegetation water information from backscatter. This is due to the difficulty in parametrizing its complex, highly vegetation-dependent relationship with canopy structure and soil moisture (Ulaby et al., 1988). Even simplified versions of the model require a host of vegetation structure information which is impractical to gather at landscape-scales (Graham and Harris, 2003). Empirical approaches such as the one presented here overcome the need for such information. Thanks to its direct relevance to wildfires, LFMC has the advantage of being sampled intensively in the field (Figs. 2, 5a), presenting large volumes of training data globally for empirical models (Yebra et al., 2019). If properly interpreted, LFMC maps produced here could therefore have applications beyond wildfire risk assessment, particularly in vegetation drought stress studies.

5. Conclusion

We presented a physics-assisted, deep learning model to dynamically estimate LFMC, and for the first time ever, produced 15-day LFMC maps at 250 m spatial resolution for the western US. Sentinel-1 backscatter and Landsat-8 optical reflectances were combined with ancillary static variables to train the model on LFMC field samples. Despite covering an area that is highly diverse in terms of vegetation type, topography, and climate, the model accuracy (R^2 = 0.63, RMSE = 25.0%, bias = 1.9%) matched or exceeded previous site-scale empirical as well as process-based methods. This study also showed that microwave backscatter directly enhances LFMC predictability, since estimation accuracy dipped substantially (R^2 from 0.63 to 0.44 and RMSE from 25.0% to 31.8%) on excluding microwave-related variables from the inputs. The LFMC maps produced here can be used for wildfire risk assessment and landscape-scale studies of vegetation drought stress in the western US. If sufficient training data can be acquired, the methodology described here can be applied elsewhere in the world.

Authorship

AGK conceived of the study. KR, AGK, and APW designed the analysis. JFF and KR retrieved data and JFF performed exploratory investigations. KR performed the analysis. KR wrote the first draft of the manuscript, and all authors contributed to interpretation of the data and writing of the manuscript.

Data accessibility statement

The live fuel moisture content maps produced in this paper can be viewed and downloaded from https://kkraoj.users.earthengine.app/view/live-fuel-moisture. The scripts supporting the analysis can be found at https://github.com/kkraoj/lfmc_from_sar.

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2020.111797.

References

Abatzoglou, J.T., Williams, A.P., 2016. Impact of anthropogenic climate change on wildfire across western US forests. Proc. Natl. Acad. Sci. 113 (42), 11,770–11,775. https://doi.org/10.1073/pnas.1607171113.

Anderegg, W.R.L., Klein, T., Bartlett, M., Sack, L., Pellegrini, A.F.A., Choat, B., Jansen, S., 2016. Meta-analysis reveals that hydraulic traits explain cross-species patterns of drought-induced tree mortality across the globe. Proc. Natl. Acad. Sci. 113 (18), 5024–5029. https://doi.org/10.1073/pnas.1525678113.

Argañaraz, J.P., Landi, M.A., Bravo, S.J., Gavier-Pizarro, G.I., Scavuzzo, C.M., Bellis, L.M., 2016. Estimation of live fuel moisture content from MODIS images for fire danger assessment in southern Gran Chaco. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 9 (12), 5339–5349. https://doi.org/10.1109/JSTARS.2016.2575366.

Arino, O., Perez, J.R., Kalogirou, V., Defourny, P., Achard, F., 2010. Globcover 2009. In: ESA Living Planet Symposium, pp. 1–3.

Attema, E.P.W., Ulaby, F.T., 1978. Vegetation modeled as a water cloud. Radio Sci. 13 (2), 357–364. https://doi.org/10.1029/RS013i002p00357.

Badgley, G., Field, C.B., Berry, J.A., 2017. Canopy near-infrared reflectance and terrestrial photosynthesis. Sci. Adv. 3 (3), 1–6. https://doi.org/10.1126/sciadv.1602244.

Barraza, V., Grings, F., Ferrazzoli, P., Salvia, M., Maas, M., Rahmoune, R., Vittucci, C., Karszenbaum, H., 2014. Monitoring vegetation moisture using passive microwave and optical indices in the dry Chaco forest, Argentina. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7 (2), 421–430. https://doi.org/10.1109/JSTARS.2013.2268011.

Barros, A.M., Pereira, J.M., 2014. Wildfire selectivity for land cover type: does size matter? PLoS One 9 (1). https://doi.org/10.1371/journal.pone.0084760.

Blackman, C.J., et al., 2019. Drought response strategies and hydraulic traits contribute to mechanistic understanding of plant dry-down to hydraulic failure. Tree Physiol. 39 (6), 910–924. https://doi.org/10.1093/treephys/tpz016.

Bowman, D.M.J.S., et al., 2009. Fire in the earth system. Science 324 (April), 481–484. https://doi.org/10.1126/science.1163886.

Bowyer, P., Danson, F.M., 2004. Sensitivity of spectral reflectance to variation in live fuel moisture content at leaf and canopy level. Remote Sens. Environ. 92 (3), 297–308. https://doi.org/10.1016/j.rse.2004.05.020.

Bradshaw, L., J. Deeming, R. Burgan, and J. Cohen (1984), The 1978 National Fire-Danger Rating System: technical documentation. General Technical Report INT-169, *Tech. Rep. July*, United States Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station Ogden, doi:https://doi.org/10.2737/INT-GTR-169 .

Brocca, L., et al., 2017. A review of the applications of ASCAT soil moisture products. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 10 (5), 2285–2306. https://doi.org/10.1109/JSTARS.2017.2651140.

Brodrick, P.G., Anderegg, L.D., Asner, G.P., 2019. Forest drought resistance at large

geographic scales. Geophys. Res. Lett. 46 (5), 2752–2760. https://doi.org/10.1029/2018GL081108.

Cal Fire, 2018. Top 20 Most Destructive California Wildfires.

Carlson, T.N., Ripley, D.A., 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sens. Environ. 62 (3), 241–252. https://doi.org/10.1016/S0034-4257(97)00104-1.

Ceccato, P., Flasse, S., Tarantola, S., Jacquemoud, S., Grégoire, J.M., 2001. Detecting vegetation leaf water content using reflectance in the optical domain. Remote Sens. Environ. 77 (1), 22–33. https://doi.org/10.1016/S0034-4257(01)00191-2.

Chollet, F., et al., 2015. Keras.

Chuvieco, E., Riaño, D., Aguado, I., Cocero, D., 2002. Estimation of fuel moisture content from multitemporal analysis of Landsat Thematic Mapper reflectance data: applications in fire danger assessment. Int. J. Remote Sens. 23 (11), 2145–2162. https://doi.org/10.1080/01431160110069818.

Chuvieco, E., Aguado, I., Dimitrakopoulos, A.P., 2004a. Conversion of fuel moisture content values to ignition potential for integrated fire danger assessment. Can. J. For. Res. 34 (11), 2284–2293. https://doi.org/10.1139/x04-101.

Chuvieco, E., Cocero, D., Riaño, D., Martin, P., Martínez-Vega, J., De La Riva, J., Pérez, F., 2004b. Combining NDVI and surface temperature for the estimation of live fuel moisture content in forest fire danger rating. Remote Sens. Environ. 92 (3), 322–331. https://doi.org/10.1016/j.rse.2004.01.019.

Chuvieco, E., González, I., Verdú, F., Aguado, I., Yebra, M., 2009. Prediction of fire occurrence from live fuel moisture content measurements in a Mediterranean ecosystem. Int. J. Wildland Fire 18 (4), 430. https://doi.org/10.1071/wf08020.

Colombo, R., Meroni, M., Marchesi, A., Busetto, L., Rossini, M., Giardino, C., Panigada, C., 2008. Estimation of leaf and canopy water content in poplar plantations by means of hyperspectral indices and inverse modeling. Remote Sens. Environ. 112 (4), 1820–1834. https://doi.org/10.1016/J.RSE.2007.09.005.

Crutzen, P.J., Andreae, M.O., 1990. Biomass burning in the tropics: impact on atmospheric chemistry and biogeochemical cycles. Science 250 (4988), 1669–1678. https://doi.org/10.1126/science.250.4988.1669.

Danson, F., and P. Bowyer (2004), Estimating live fuel moisture content from remotely sensed reflectance, Remote Evans, 92(3), 309 – 321, doi:https://doi.org/https://doi.org/10.1016/j.rse.2004.03.017 , (forest Fire Prevention and Assessment).

Dennison, P.E., Roberts, D.A., Peterson, S.H., Rechel, J., 2005. Use of normalized difference water index for monitoring live fuel moisture. Int. J. Remote Sens. 26 (5), 1035–1042. https://doi.org/10.1080/01431160212331002806.

Dennison, P.E., Moritz, M.A., Taylor, R.S., 2008. Evaluating predictive models of critical live fuel moisture in the Santa Monica Mountains , California. Int. J. Wildland Fire 18–27.

Dobson, M.M., Ulaby, F.F., Hallikainen, M.M., El-rayes, M.M., 1985. Microwave dielectric behavior of wet soil-part II: dielectric mixing models. IEEE Trans. Geosci. Remote Sens. GE-23 (1), 35–46.

Dozat, T., 2016. Incorporating nesterov momentum into adam. In: Proceedings of 4th International Conference on Learning Representations, Workshop Track.

Dubois, P., Syl, J., Engman, T., 1995. Measuring soil moisture with imaging radars. IEEE Trans. Geosci. Remote Sens. 33 (4), 915–926.

Fan, L., et al., 2018. Evaluation of microwave remote sensing for monitoring live fuel moisture content in the Mediterranean region. Remote Sens. Environ. 205 (January 2017), 210–223. https://doi.org/10.1016/j.rse.2017.11.020.

Fischer, G., Nachtergaele, F., Prieler, S., Van Velthuizen, H., Verelst, L., Wiberg, D., 2008. Global agro-ecological zones assessment for agriculture (gaez 2008). In: IIASA, Laxenburg, Austria and FAO, Rome, Italy, pp. 10.

Gao, B.C., 1996. NDWI - a normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58 (3), 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3.

García, M., Chuvieco, E., Nieto, H., Aguado, I., 2008. Combining AVHRR and meteorological data for estimating live fuel moisture content. Remote Sens. Environ. 112 (9), 3618–3627. https://doi.org/10.1016/j.rse.2008.05.002.

Gitelson, A.A., Kaufman, Y.J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetation fraction. Remote Sens. Environ. 80 (1), 76–87. https://doi.org/10.1016/S0034-4257(01)00289-9.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Graham, A.J., Harris, R., 2003. Extracting biophysical parameters from remotely sensed radar data: a review of the water cloud model. Prog. Phys. Geogr. 27 (2), 217–229. https://doi.org/10.1191/0309133303pp378ra.

Hardisky, M., Klemas, V., Smart, M., 1983. The influence of soil salinity, growth form, and leaf moisture on the spectral radiance of. In: Spartina alterniflora. 49. pp. 77–83.

Hardy, C.C., Burgan, R.E., 1999. Evaluation of NDVI for monitoring live moisture in three vegetation types of the western U.S. Photogramm. Eng. Remote. Sens. 65 (5), 603–610.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.

Holden, Z.A., Swanson, A., Luce, C.H., Jolly, W.M., Maneta, M., Oyler, J.W., Warren, D.A., Parsons, R., Affleck, D., 2018. Decreasing fire season precipitation increased recent western US forest wildfire activity. Proc. Natl. Acad. Sci. U. S. A. 115 (36), E8349–E8357. https://doi.org/10.1073/pnas.1802316115.

Jacquemoud, S., Baret, F., 1990. PROSPECT: a model of leaf optical properties spectra. Remote Sens. Environ. 34 (2), 75–91. https://doi.org/10.1016/0034-4257(90)90100-Z.

Jia, G.J., Burke, I.C., Goetz, A.F., Kaufmann, M.R., Kindel, B.C., 2006. Assessing spatial patterns of forest fuel using AVIRIS data. Remote Sens. Environ. 102 (3–4), 318–327. https://doi.org/10.1016/j.rse.2006.02.025.

Jia, S., Kim, S.H., Nghiem, S.V., Kafatos, M., 2019. Estimating live fuel moisture using SMAP L-band radiometer soil moisture for Southern California, USA. Remote Sens. 11 (13). https://doi.org/10.3390/rs11131575.

Jolly, W., Johnson, D., 2018. Pyro-ecophysiology: shifting the paradigm of live wildland fuel research. Fire 1 (1), 8. https://doi.org/10.3390/fire1010008.

Karavani, A., Boer, M.M., Baudena, M., Colinas, C., Díaz-Sierra, R., Pemán, J., de Luis, M., Enríquez-de Salamanca, Á., Resco de Dios, V., 2018. Fire-induced deforestation in drought-prone Mediterranean forests: drivers and unknowns from leaves to communities. Ecol. Monogr. 88 (2), 141–169. https://doi.org/10.1002/ecm.1285.

Karpatne, A., Atluri, G., Faghmous, J.H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N., Kumar, V., 2017. Theory-guided data science: a new paradigm for scientific discovery from data. IEEE Trans. Knowl. Data Eng. 29 (10), 2318–2331. https://doi.org/10.1109/TKDE.2017.2720168.

Keane, R.E., 2015. Fuel Concepts. Springer International Publishing, Cham, pp. 175–184. https://doi.org/10.1007/978-3-319-09015-3_10.

Kelley, D.I., Bistinas, I., Whitley, R., Burton, C., Marthews, T.R., Dong, N., 2019. How contemporary bioclimatic and human controls change global fire regimes. Nat. Clim. Chang. 9 (September), 690–696. https://doi.org/10.1038/s41558-019-0540-7.

Konings, A.G., Gentine, P., 2017. Global variations in ecosystem-scale isohydricity. Glob. Chang. Biol. 23 (2), 891–905. https://doi.org/10.1111/gcb.13389.

Konings, A.G., Yu, Y., Xu, L., Yang, Y., Schimel, D.S., Saatchi, S.S., 2017a. Active microwave observations of diurnal and seasonal variations of canopy water content across the humid African tropical forests. Geophys. Res. Lett. 44 (5), 2290–2299. https://doi.org/10.1002/2016GL072388.

Konings, A.G., Williams, A.P., Gentine, P., 2017b. Sensitivity of grassland productivity to aridity controlled by stomatal and xylem regulation. Nat. Geosci. 10 (4), 284–288. https://doi.org/10.1038/ngeo2903.

Konings, A.G., Rao, K., Steele-Dunne, S.C., 2019. Macro to micro: microwave remote sensing of plant water content for physiology and ecology. New Phytol. https://doi.org/10.1111/nph.15808. p. nph.15808.

Leblon, B., Kasischke, E., Alexander, M., Doyle, M., Abbott, M., 2002. Fire danger monitoring using ERS-1 SAR images in the case of northern boreal forests. Nat. Hazards 27 (3), 231–255. https://doi.org/10.1023/A:1020375721520.

Lelieveld, J., Evans, J.S., Fnais, M., Giannadaki, D., Pozzer, A., 2015. The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature 525 (7569), 367–371. https://doi.org/10.1038/nature15371.

Liu, S., Wei, Y., Post, W.M., Cook, R.B., Schaefer, K., Thorton, M.M., 2014. NACP MsTMIP: Unified North American Soil Map. https://doi.org/10.3334/ornldaac/1242.

Miller, C., Ager, A.A., 2013. A review of recent advances in risk analysis for wildfire management. Int. J. Wildland Fire 22 (1), 1–14. https://doi.org/10.1071/WF11114.

Moreira, F., Vaz, P., Catry, F., Silva, J.S., 2009. Regional variations in wildfire susceptibility of land-cover types in Portugal: implications for landscape management to minimize fire hazard. Int. J. Wildland Fire 18 (5), 563–574. https://doi.org/10.1071/WF07098.

Myoung, B., Kim, S.H., Nghiem, S.V., Jia, S., Whitney, K., Kafatos, M.C., 2018. Estimating live fuel moisture from MODIS satellite data for wildfire danger assessment in Southern California USA. Remote Sens. 10 (1), 87. https://doi.org/10.3390/rs10010087.

Newton, J.E., Blackman, G.E., 1970. The penetration of solar radiation through leaf canopies of different structure. Ann. Bot. 34 (2), 329–348. https://doi.org/10.1093/oxfordjournals.aob.a084373.

Nolan, R.H., Boer, M.M., Resco De Dios, V., Caccamo, G., Bradstock, R.A., 2016. Large-scale, dynamic transformations in fuel moisture drive wildfire activity across southeastern Australia. Geophys. Res. Lett. 43 (9), 4229–4238. https://doi.org/10.1002/2016GL068614.

Nolan, R.H., Hedo, J., Arteaga, C., Sugai, T., Resco de Dios, V., 2018. Physiological drought responses improve predictions of live fuel moisture dynamics in a Mediterranean forest. Agric. For. Meteorol. 263 (August), 417–427. https://doi.org/10.1016/j.agrformet.2018.09.011.

Pampaloni, P., Paloscia, S., 1986. Microwave emission and plant water content: a comparison between Field measurements and theory. IEEE Trans. Geosci. Remote Sens. GE-24 (6), 900–905. https://doi.org/10.1109/TGRS.1986.289705.

Pellizzaro, G., Cesaraccio, C., Duce, P., Ventura, A., Zara, P., 2007. Relationships between seasonal patterns of live fuel moisture and meteorological drought indices for Mediterranean shrubland species. Int. J. Wildland Fire 16 (2), 232. https://doi.org/10.1071/wf06081.

Peterson, S.H., Roberts, D.A., Dennison, P.E., 2008. Mapping live fuel moisture with MODIS data: a multiple regression approach. Remote Sens. Environ. 112 (12), 4272–4284. https://doi.org/10.1016/j.rse.2008.07.012.

Preisler, H.K., Brillinger, D.R., Burgan, R.E., Benoit, J.W., 2004. Probability based models for estimation of wildfire risk. Int. J. Wildland Fire 13 (2), 133–142.

PRISM Climate Group Oregon State University, 2004. PRISM Climate Data. (doi:20 May 2016).

Qi, Y., Dennison, P.E., Spencer, J., Riaño, D., 2012. Monitoring live fuel moisture using soil moisture and remote sensing proxies. Fire Ecology 8 (3), 71–87. https://doi.org/10.4996/fireecology.0803071.

Randerson, J.T., et al., 2006. The impact of boreal forest fire on climate warming. Science 314 (5802), 1130–1132. https://doi.org/10.1126/science.1132075.

Rao, K., Anderegg, W.R., Sala, A., Martínez-Vilalta, J., Konings, A.G., 2019. Satellite-based vegetation optical depth as an indicator of drought-driven tree mortality. Remote Sens. Environ. 227 (March), 125–136. https://doi.org/10.1016/j.rse.2019.03.026.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven earth system science. Nature 566 (7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1.

Roberts, D.A., Dennison, P.E., Peterson, S., Sweeney, S., Rechel, J., 2006. Evaluation of Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and Moderate Resolution

Imaging Spectrometer (MODIS) measures of live fuel moisture and fuel condition in a shrubland ecosystem in southern California. Journal of Geophysical Research: Biogeosciences 111 (4), 1–16. https://doi.org/10.1029/2005JG000113.

Rosen, P.A., Hensley, S., Shaffer, S., Veilleux, L., Chakraborty, M., Misra, T., Bhan, R., Raju Sagi, V., Satish, R., 2015. The NASA-ISRO SAR mission - an international space partnership for science and societal benefit. IEEE National Radar Conference - Proceedings, 2015-June(June) 1610–1613. https://doi.org/10.1109/RADAR.2015.7131255.

Rossa, C.G., 2017. The effect of fuel moisture content on the spread rate of forest fires in the absence of wind or slope. Int. J. Wildland Fire 26 (1), 24–31. https://doi.org/10.1071/WF16049.

Roy, D.P., et al., 2014. Landsat-8: science and product vision for terrestrial global change research. Remote Sens. Environ. 145, 154–172. https://doi.org/10.1016/j.rse.2014.02.001.

Ruffault, J., Martin-StPaul, N., Pimont, F., Dupuy, J.L., 2018. How well do meteorological drought indices predict live fuel moisture content (LFMC)? An assessment for wildfire research and operations in Mediterranean ecosystems. Agric. For. Meteorol. 262 (July), 391–401. https://doi.org/10.1016/j.agrformet.2018.07.031.

Saatchi, S., Halligan, K., Despain, D.G., Crabtree, R.L., 2007. Estimation of forest fuel load from radar remote sensing. IEEE Trans. Geosci. Remote Sens. 45 (6), 1726–1740. https://doi.org/10.1109/TGRS.2006.887002.

Saatchi, S.S., et al., 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. Proc. Natl. Acad. Sci. U. S. A. 108 (24), 9899–9904. https://doi.org/10.1073/pnas.1019576108.

Saatchi, S., Asefi-Najafabady, S., Malhi, Y., Aragão, L.E., Anderson, L.O., Myneni, R.B., Nemani, R., 2013. Persistent effects of a severe drought on Amazonian forest canopy. Proc. Natl. Acad. Sci. U. S. A. 110 (2), 565–570. https://doi.org/10.1073/pnas.1204651110.

Shakesby, R.A., Doerr, S.H., 2006. Wildfire as a hydrological and geomorphological agent. Earth Sci. Rev. 74 (3–4), 269–307. https://doi.org/10.1016/j.earscirev.2005.10.006.

Simard, M., Pinto, N., Fisher, J.B., Baccini, A., 2011. Mapping forest canopy height globally with spaceborne lidar. Journal of Geophysical Research: Biogeosciences 116 (4), 1–12. https://doi.org/10.1029/2011JG001708.

Sims, D.A., Gamon, J.A., 2003. Estimation of vegetation water content and photosynthetic tissue area from spectral reflectance: a comparison of indices based on liquid water and chlorophyll absorption features. Remote Sens. Environ. 84 (4), 526–537. https://doi.org/10.1016/S0034-4257(02)00151-7.

Song, C., 2013. Optical remote sensing of forest leaf area index and biomass. Prog. Phys. Geogr. 37 (1), 98–113. https://doi.org/10.1177/0309133312471367.

Steele-Dunne, S.C., McNairn, H., Monsivais-Huertero, A., Judge, J., Liu, P.W., Papathanassiou, K., 2017. Radar remote sensing of agricultural canopies: a review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 10 (5), 2249–2273. https://doi.org/10.1109/JSTARS.2016.2639043.

Tanase, M.A., Panciera, R., Lowell, K., Aponte, C., 2015. Monitoring live fuel moisture in semiarid environments using L-band radar data. Int. J. Wildland Fire 24 (4), 560–572. https://doi.org/10.1071/WF14149.

Tian, F., et al., 2016. Remote sensing of vegetation dynamics in drylands: evaluating vegetation optical depth (VOD) using AVHRR NDVI and in situ green biomass data over west African Sahel. Remote Sens. Environ. 177, 265–276. https://doi.org/10.1016/j.rse.2016.02.056.

Topouzelis, K., Singha, S., 2016. Incidence angle normalization of wide swath SAR data for oceanographic applications. Open Geosciences 8 (1), 450–464. https://doi.org/10.1515/geo-2016-0029.

Torres, R., et al., 2012. GMES Sentinel-1 mission. Remote Sens. Environ. 120, 9–24. https://doi.org/10.1016/j.rse.2011.05.028.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8 (2), 127–150. https://doi.org/10.1016/0034-4257(79)90013-0.

Ulaby, F.T., Long, D.G., 2014. Microwave Radar and Radiometric Remote Sensing. 984 University of Michigan Press, Ann Arbor. https://doi.org/10.1016/J.JAG.2014.08.003.

Ulaby, F., McDonald, K., Sarabandi, K., Dobson, M., 1988. Michigan microwave canopy scattering models (MIMICS). In: International Geoscience and Remote Sensing Symposium, 'Remote Sensing: Moving toward the 21st Century'. 2. pp. 1009. https://doi.org/10.1109/IGARSS.1988.570506. (August 1988).

United States Forest Sevices, 2018. National Fuel Moisture Database.

USDA, 2019. Description of STATSGO2 Database. NRCS Soils.

USGS, 2011. National Elevation Dataset.

Vermote, E., Justice, C., Claverie, M., Franch, B., 2016. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. Remote Sens. Environ. 185, 46–56. https://doi.org/10.1016/j.rse.2016.04.008.

Wang, L., Quan, X., He, B., Yebra, M., Xing, M., Liu, X., 2019. Assessment of the dual polarimetric sentinel-1A data for forest fuel moisture content estimation. Remote Sens. 11 (13), 1568. https://doi.org/10.3390/rs11131568.

Westerling, A.L.R., 2016. Increasing western US forest wildfire activity: sensitivity to changes in the timing of spring. Philosophical Transactions of the Royal Society B: Biological Sciences 371 (1696). https://doi.org/10.1098/rstb.2015.0178.

Williams, A.P., Abatzoglou, J.T., Gershunov, A., Guzman-Morales, J., Bishop, D.A., Balch, J.K., Lettenmaier, D.P., 2019. Observed impacts of anthropogenic climate change on wildfire in California. Earth's Future 7 (8), 892–910. https://doi.org/10.1029/2019ef001210.

Yebra, M., Chuvieco, E., 2009. Linking ecological information and radiative transfer models to estimate fuel moisture content in the Mediterranean region of Spain: solving the ill-posed inverse problem. Remote Sens. Environ. 113 (11), 2403–2411. https://doi.org/10.1016/j.rse.2009.07.001.

Yebra, M., Dennison, P.E., Chuvieco, E., Riaño, D., Zylstra, P., Hunt, E.R., Danson, F.M., Qi, Y., Jurdao, S., 2013. A Global Review of Remote Sensing of Live Fuel Moisture Content for Fire Danger Assessment: Moving towards Operational Products. https://doi.org/10.1016/j.rse.2013.05.029.

Yebra, M., Quan, X., Riaño, D., Rozas Larraondo, P., van Dijk, A.I., Cary, G.J., 2018. A fuel moisture content and flammability monitoring methodology for continental Australia based on optical remote sensing. Remote Sens. Environ. 212 (April), 260–272. https://doi.org/10.1016/j.rse.2018.04.053.

Yebra, M., et al., 2019. Globe-LFMC, a global plant water status database for vegetation ecophysiology and wildfire applications. Scientific Data 6 (1), 155. https://doi.org/10.1038/s41597-019-0164-9.