

Estimation of the relationship between satellite-derived vegetation indices and live fuel moisture towards wildfire risk in Southern California

Kristen Whitney

Center of Excellence in Earth
System Modeling and
Observations
Chapman University
Orange, CA USA
kwhitney@chapman.edu

Seung Hee Kim

Center of Excellence in Earth
System Modeling and
Observations
Chapman University
Orange, CA USA
sekim@chapman.edu

Shenyue Jia

Center of Excellence in Earth
System Modeling and
Observations
Chapman University
Orange, CA USA
sjia@chapman.edu

Menas Kafatos

Center of Excellence in Earth
System Modeling and
Observations
Chapman University
Orange, CA USA
kafatos@chapman.edu

Abstract—Southern California possesses a Mediterranean climate having semi-arid to arid characteristics and contains shrubland areas at high risk to wildfire. To assess wildfire danger, fire agencies have been monitoring the moisture of vegetation, called live fuel moisture (LFM), using field-based sampling. Unfortunately, spatial and temporal resolution of live fuel moisture data are significantly limited because sampling is labor intensive. Remote sensing satellite data has been used to monitor vegetation moisture content and health of shrublands. Therefore, a potential approach to overcome the limitations of manual measurements of live fuel moisture is to use vegetation indices (VIs) derived from satellite data. The objective of this study is to understand the link between vegetation indices derived from a Moderate Resolution Imaging Spectroradiometer (MODIS) aboard both Terra and Aqua satellites and in-situ live fuel moisture data. In this study, five vegetation indices were calculated using 6 bands of MODIS data within the visible and infrared spectrum collectively with the focus on the three best performing: enhanced vegetation index (EVI), normalized difference water index (NDWI), and visible atmospherically resistant index (VARI). Six sites with multi-year live fuel moisture data collection type were each represented with one pixel of MODIS data with a 500m by 500m spatial resolution covering the time period of February 2000 through December 2017 acquired aboard Terra and June 2002 through December 2017 acquired aboard Aqua. Simple linear regression was then applied to measure the coefficient of determination (R^2) between the vegetation indices and live fuel moisture data. The results show a great variance of R^2 between the sites as well as a variance of best performing VI. The two strongest coefficients of determination, $R^2=0.74$ and $R^2=0.72$, were calculated at one site for enhanced vegetation index vs. live fuel moisture over a 15-year time period of data collected on Aqua and a 17-year time period of data collected on Terra respectively. The relationship was also affected by annual climate conditions including precipitation. Our preliminary results indicate that the satellite data reasonably well-represents the live fuel moisture with higher temporal resolutions over a large area. Utilizing the remote sensing data in wildfire danger assessment will support fire agencies by saving resources for collecting ground data and providing better dataset in both time and space. This will also be beneficial for land management and planning, stakeholders and local governments.

Keywords—remote sensing, live fuel moisture, vegetation index, wildfire, Southern California, chaparral ecosystem

I. INTRODUCTION

Wildfires are a natural part of the Mediterranean climate of Southern California consisting of semi-arid to arid characteristics with hot and dry summers and wet winters. A shrubland ecosystem in this region is particularly sensitive to climate variability such as heat wave and drought. Long term and regional scale monitoring is essential to assess the condition of the ecosystem and evaluate its susceptibility to wildfire.

Over the years, monitoring the condition of a shrubland ecosystem has relied on ground observation. Unfortunately, the sampling sites for natural vegetation conditions are sparse and the sampling times are not coordinated between sites and agencies. Multi-temporal remote sensing satellite data collects radiation that is emitted or reflected off the earth's surface in the visible, and infrared spectrums and is useful in observing the state of its changes in land surface vegetation coverage. Compared to the ground observation, the satellite data provides higher temporal resolutions over a large area. Utilizing remote sensing data will save time and resources to collect ground data and will be beneficial for land management and planning [1, 2].

Moisture content in live vegetation is affected by a seasonal cycle of precipitation and soil moisture. The characteristics of regional climate in Southern California is dry summer and fall thus moisture content of live chaparral vegetation decreases, resulting in an increase of wildfire risk. Fire agencies typically monitor the moisture of vegetation using field-based sampling of live fuel moisture (LFM). The objective of this study is to understand the link between vegetation indices derived from remote sensing and LFM and develop a formulation for estimating live fuel moisture using MODIS data collected from two satellites. This study focuses on six sites in Los Angeles (LA) County, Southern California in shrubland areas at high risk to wildfire with relatively consistent biweekly LFM collection managed by the Los Angeles County Fire Department (Fig.1). Chamise (*Adenostoma fasciculatum*) is the most common shrub in the

chaparral in LA County and LFM values in this study are measured from leaves collected from this chamise vegetation.

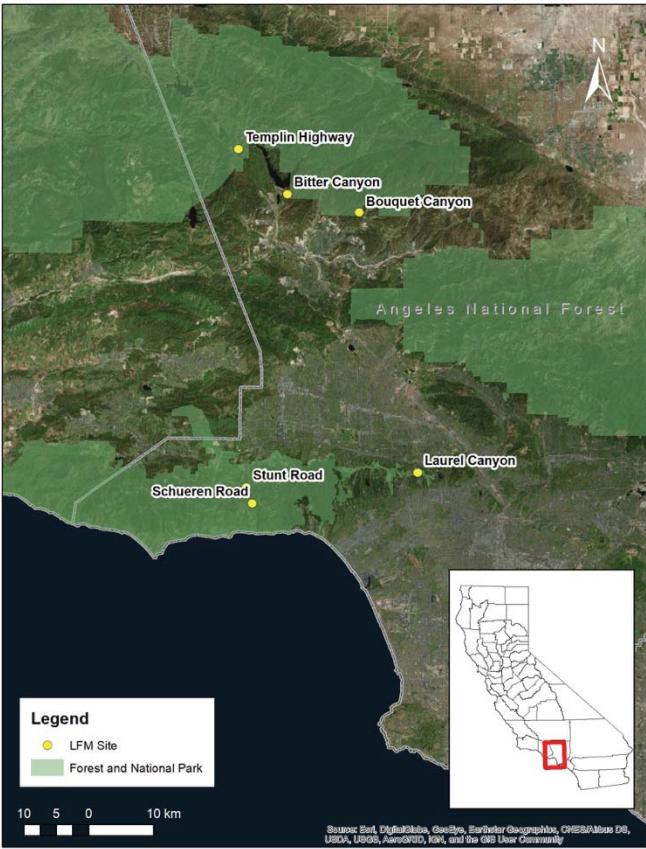


Fig. 1. Map showing locations of six sites. **Templin Highway**, **Bitter Canyon** and **Bouquet Canyon** are considered inland sites. **Laurel Canyon** is in the Hollywood Hills and also considered inland. **Stunt Road** and **Schueren** are coastal site Malibu Canyon. Bold indicates the name representing each site.

II. MATERIALS AND METHODOLOGY

A. Wildfire risk and Live Fuel Moisture

Live fuel moisture (LFM) is collected from clipping live foliage, weighing when wet (W_w), then re-weighing when dry (W_d):

$$LFM = \frac{W_w - W_d}{W_d} * 100 \quad (1)$$

where a low fire risk is greater than 120% a moderate fire risk between 80% and 120%, a high risk between 60% and 80% and an extremely high risk less than 60% (Fig.2). Previous studies have shown a strong correlation between in-situ LFM and vegetation index derived from satellite images [3, 4]. LFM data sampling intervals are typically every two weeks; however, the intervals can be longer during wet seasons, when leaves and twigs remain wet after rainfall events. This tends to force fire agencies to postpone their LFM sampling by a few days, since wet vegetation induces significant errors in LFM. During the fire season, LFM sampling may also be cancelled due to the breakout of wildfire and other factors that may lead to difficulty accessing the LFM sites and or reduction of manpower available to collect data. In this study, we applied linear

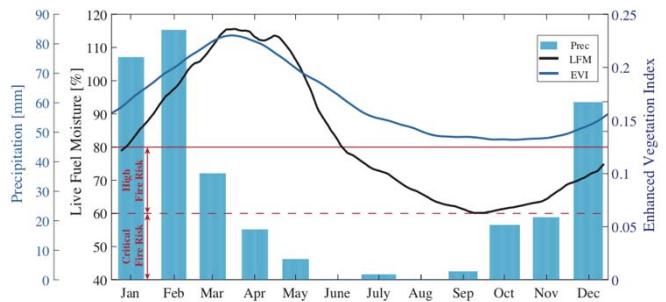


Fig.2. Daily mean over 17 years of LFM and Enhanced Vegetation Index (EVI) and monthly mean of precipitation at Bitter Canyon. LFM less than 80% is considered high fire risk.

interpolation to generate a daily dataset of LFM based on six selected LFM sites with biweekly sampling data with a goal of approximately 23 data points per year. The LFM observation data were obtained from the National Fuel Moisture Database, USFS Wildland Assessment System and was accessed at: <https://www.wfas.net/index.php/national-fuel-moisture-database-moisture-drought-103>.

B. MODIS remote sensing datasets

Land surface reflectance Moderate Resolution Imaging Spectroradiometer (MODIS) data MOD09A1 Version 6 and MYD09A1 Version 6 are level-3 product collections containing 13 layers. MOD09A1 is acquired on satellite Terra which passes over the equator from north to south in the morning, and MYD09A1 is acquired on satellite Aqua which passes from south to north over the equator in the afternoon. At a specific pixel location on the ground, the satellites obtain data at different times of the day with sensors oriented at different angles with respect to the earth's surface. For both datasets, pixel values in a tile are collected in an 8-day time composition at sinusoidal projection. The composite allows surface reflection at the ground level to be estimated without atmospheric scattering observation. The focus on land surface reflectance includes bands 1-7 in the visible and infrared spectrums. Each pixel is considered to have a spatial resolution of roughly 500m by 500m. Each pixel contains data collected during days with the lowest cloud coverage and minimum aerosol noise within the 8 days. The date in which the lowest intensity of blue band (the band with minimum wavelength) is also used when applicable. The pixel value is of type 16-bit unsigned integer. The bands used in this study to calculate vegetation indices are band 1: red (620-670)nm, band 2: near infrared (NIR) (841-875)nm, band 3: blue (459-479)nm, band 4: green, (545-565)nm, and band 5: short wave infrared (SWIR) (1230-1250)nm. Two additional layers provide quality information that applies to each pixel individually as well as the entire tile. The tiles are re-projected to Universal Trans Mercator (UTM) system in the ellipsoid World Geodetic System 1984 (WGS 84). In this study, MODIS data was accessed at EarthExplorer: <https://earthexplorer.usgs.gov/> and the tool used to re-project is available to download at: https://lpdaac.usgs.gov/tools/modis_reprojection_tool.

C. Vegetation indices

Vegetation index (VI) algorithms can be calculated using reflectance data and provide a quantitative evaluation of health and other factors related to land cover. The most common indices are the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI). Normalized Difference Water Index (NDWI) and Normalized Difference Infrared Index (NDII) are used to monitor changes in water content stored in leaves. Visible Atmospherically Resistant Index (VARI) uses all visible bands to evaluate the fraction of vegetation cover. The five vegetation indices above were the most commonly used factors to estimate LFM [5]. Either EVI, NDWI, or VARI produced the best results at each of the six sites and are the focus in this study.

EVI focuses on the relation of the reflectance of near infrared and red, and in addition includes coefficients and reflectance in the blue portion of the visible spectrum to remove aerosol noise. The gain factor, G=2.5, C1=6 and C2=7.5 are coefficient terms that use the blue light band to correct for aerosol noise in the red band; L=1 is the canopy background adjustment term [6]. The value of the coefficients can vary depending on land surface conditions. The ones chosen in this study are commonly used and are also used in the previous studies.

$$EVI = G * \frac{NIR - RED}{NIR + C1*RED - C2*BLUE + L} \quad (2)$$

NDWI uses near infrared and shortwave infrared bands [7]. Compared to EVI, NDWI is more directly sensitive to changes in moisture content and less to chlorophyll. The difference in near infrared and shortwave infrared intensities is expected to show a difference in absorption and emission intensities with vegetation due to interaction with the physiology of the vegetation.

$$NDWI = \frac{NIR - SWIR}{NIR + SWIR} \quad (3)$$

VARI includes bands only in the visible spectrum where land cover with a greater intensity of reflection in the green spectrum is considered healthier [8].

$$VARI = \frac{GREEN - RED}{GREEN + RED - BLUE} \quad (4)$$

The ‘500 m State Flag’ layer of the MODIS data was used to throw out pixels that are marked with cloud coverage. Linear interpolation was then used to fill in the gaps left by the pixel values that are thrown out. A moving mean with a window of 31 days was applied to the VI time series.

D. Linear regression and coefficient of determination

Linear regression was applied to measure the correlation between the vegetation index and the ground truth data for the selected six sites. Univariate linear regression was applied to LFM versus each of the vegetation indices. R^2 was used as the key measurement of model performance. Linear regression

applied to LFM versus VI calculated from dataset MOD09A1 over 17 years produced an empirical model function:

$$\text{Estimated LFM} = \beta_1 * VI + \beta_0 \quad (5)$$

We calculated VIs and regression using MODIS Terra and Aqua data separately. All the analysis was conducted in MATLAB®.

III. RESULTS

Comparisons of the linear regression correlation of LFM and three vegetation indices, EVI, NDWI, and VARI, calculated from MODIS satellites Terra and Aqua at six sites in LA County, California showed that a single VI does not dominate. Furthermore, comparison of results of calendar years at a single sight also showed a variance of the strength of correlation.

A. Comparison of vegetation indices

The time period continuing through 2017 of data collection differs site to site. Bitter, Laurel and Schueren have LFM data concurrent with all MODIS data, Bouquet has 16 years of concurrent data collection, Stunt has a 12-year period beginning in 2006 and Templin a 5-year period beginning in 2013. The vegetation index EVI performed best at three sites, VARI at two, and NDWI at one site. VIs derived from MOD09A1 collected on Terra tend to perform better than VIs derived from MYD09A1 data collected on Aqua. The average coefficient of determination at all sites show that EVI and VARI calculated from MOD09A1 produce the greatest R^2 . From here on vegetation indices calculated from the MOD09A1 dataset and three sites with concurrent collected data covering the greatest time period is the focus in this study.

TABLE 1: COEFFICIENT OF DETERMINATION BETWEEN LFM AND VIS AT SIX SITES USING MODIS TERRA (MOD09A1) AND AQUA (MYD09A1) REFLECTANCE PRODUCTS

Site	EVI		NDWI		VARI		Per Site Mean R^2
	Terra	Aqua	Terra	Aqua	Terra	Aqua	
Bitter	0.73	0.74	0.70	0.67	0.68	0.67	0.70
Templin	0.70	0.66	0.54	0.67	0.63	0.60	0.64
Bouquet *	0.68	0.66	0.72	0.51	0.59	0.41	0.60
Stunt	0.62	0.61	0.59	0.46	0.59	0.58	0.58
Laurel	0.60	0.60	0.57	0.36	0.62	0.55	0.55
Schueren	0.44	0.42	0.52	0.46	0.65	0.58	0.51
Per VI Mean R^2	0.63	0.62	0.61	0.52	0.63	0.57	

*Bouquet Canyon is missing LFM data during 2004

The sites are ordered in Table 1 by greatest to least mean R^2 and the greatest value at each site is in bold. All p-values in this study are less than 0.05.

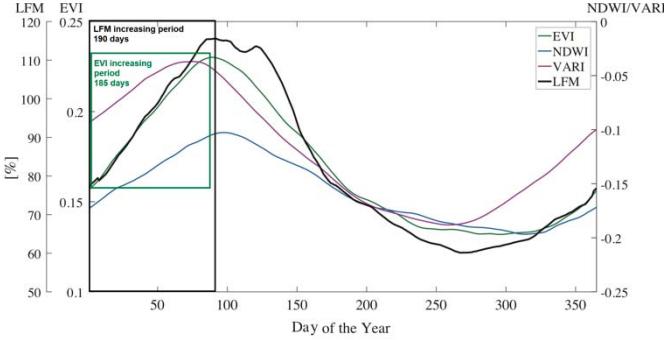


Fig.3. Multi-year average of LFM, EVI, NDWI and VARI at site Bitter canyon from 2001 to 2017.

There was an inconsistency of the days with the mean maximum and minimum values of the VIs in comparison to each other as well as the LFM daily mean time series (Fig.3). Of the three VI daily mean time series, EVI showed the least temporal difference in comparison with LFM maximum, minimum and growth period. The maximum EVI is three days ahead of the maximum LFM and the minimum is two days behind LFM minimum. The growth period represented by days with an increasing LFM is 190 days, and with an increasing EVI is 185 days. VARI, instead, tended to peak earlier than LFM.

B. Site comparison of LFM estimation

Comparison of the multi-year mean at the three sites in this study with the greatest number of LFM samples over concurrent time periods from 2001 to 2017 showed a site to site variation of top performing VI as well as a substantial difference in the magnitude of an EVI time series. One site is considered coastal (Schueren) and two are considered inland (Bitter and Laurel).

TABLE 2: ESTIMATED LFM EMPIRICAL MODELS AT THREE SITES WITH LFM DATA COLLECTION DURING 2001-2017

Site	VI	β_1	β_0	# of LFM samples
Bitter	EVI	460	4.74	394
Laurel	VARI	581	93.8	386
Schueren	VARI	418	94.1	391

EVI daily mean at the coastal site is annually consistently greater than the inland sites and the LFM daily mean at the coastal site is seasonally greater than the inland sites indicating greater vegetation moisture at the coastal site during the first half of the year (Fig. 4). The LFM ranges (maximum value minus minimum value) at the three sites differed, however LFM means over the 17-year period were similar with a max difference of approximately 5. Alternatively, EVI ranges are similar with a value of approximately 0.10 at all sites, however the mean values varied by 0.13. Comparisons of the mean temporal LFM collected at different sites and the mean temporal EVI illustrated that one univariate linear regression combining multiple sites, including those relatively close in proximity, resulted in weaker correlation than applying linear regression to each site separately.

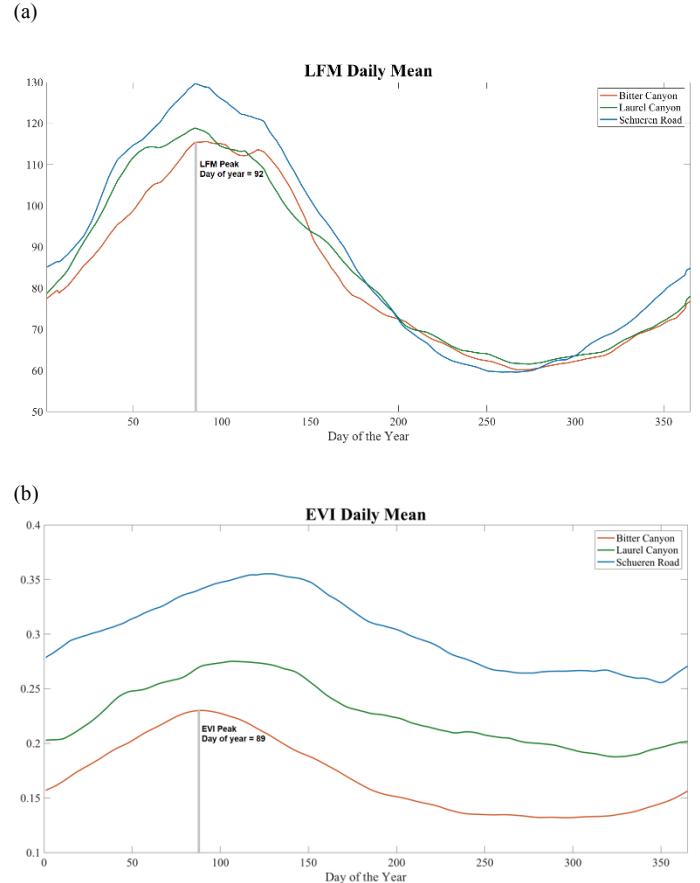


Fig.4. Comparison of three sites (a) LFM multi-year average and (b) EVI multi-year average, both calculated using the daily data from 2001 to 2017.

C. Comaprison of interannual calendar years

As the site having the best correlation between LFM and VI, we selected Bitter Canyon to investigate inter-annual variation of LFM and explore the relationship with precipitation. Precipitation data in relation to the Bitter Canyon set was collected at the Del Valle Remote Automatic Weather Station (RAWS network) and is one climatological factor to represent a monthly seasonal pattern averaged over the 17-year period (Fig. 2). Both the daily mean maximum of LFM (day 92) and EVI (day 89) had a lag time following the max monthly mean of precipitation which occurs in February.

Not only is there a site to site variation of VI performance, there is also a year to year variation at a single site. The year showing the strongest LFM versus EVI correlation at Bitter was 2017 with an R^2 of 0.94 and the year showing the weakest correlation was 2007 with an R^2 of 0.08. Over the entire time period, both the LFM and EVI maximums occurred in 2005, the year with greatest annual precipitation (Fig. 5). In addition, the LFM and EVI minimums both occurred in 2016, towards the end of the most extreme drought in California's recorded history.

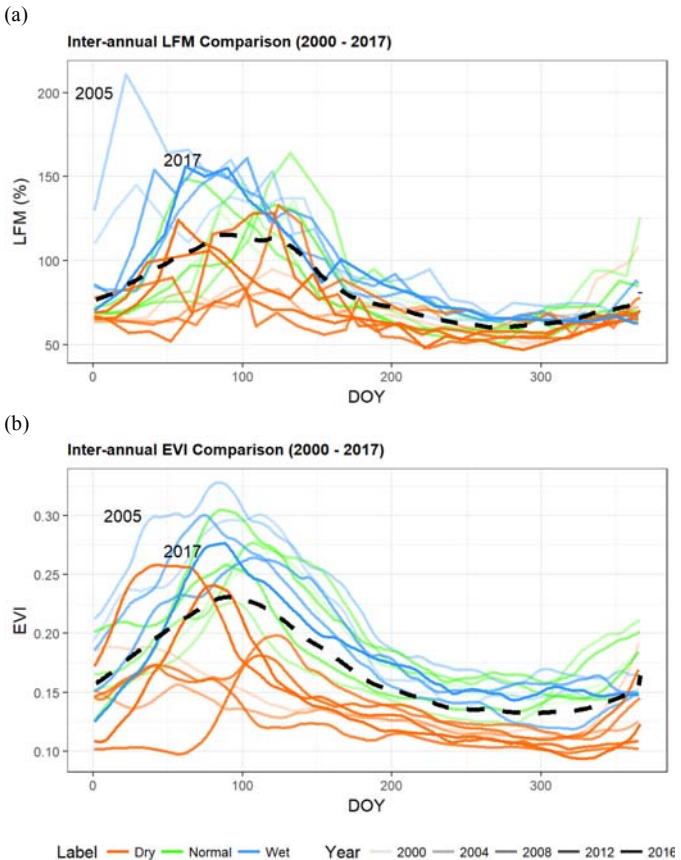


Fig. 5. Time series for each calendar year time period show the interannual variations of (a) LFM and (b) EVI at Bitter Canyon. Blue represents the five years with greatest annual precipitation (2001, 2003, 2004, 2005, and 2010), orange represents years with lowest annual precipitation (2002, 2007, 2012, 2013 and 2015). Green represents other years that had a normal level of precipitation between 80 mm and 100 mm. Dark and shallow color indicates short to long gap from present. Black dashed line represents multi-annual average of LFM and EVI.

We also investigated whether the correlation between LFM and VI was determined by the inter-annual difference of LFM or precipitation. The five years with the weakest correlation occur when the annual mean LFM and annual precipitation are both less than the mean over the time period.

Although R^2 was much higher in years with a moderate fire risk ($LFM > 80\%$), the correlation decreased as annual average LFM or precipitation rose up (Fig. 6). Among the two extremes, wettest year (2005, $R^2=0.73$) and driest year (2013, $R^2=0.72$), the R^2 was at the similar level, indicating very dry and very wet moisture condition can both weaken the strength of relationship between LFM and EVI. For years with the best correlation (above median), moisture condition was at the moderate level.

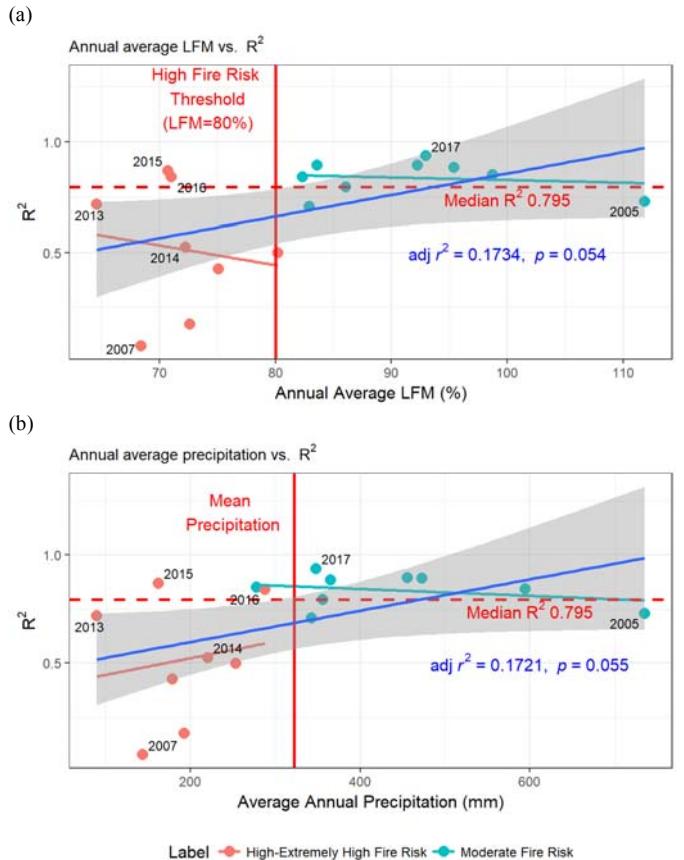


Fig. 6. Relationships between LFM vs. EVI coefficient of determinations (R^2) for each year from 2001 to 2017 and (a) annual LFM mean, (b) annual precipitation at Bitter Canyon. Years with LFM below/above 80% were colored as moderate fire risk years (cyan) and high to extremely high fire risk years (pink). Years with extremely high and low annual precipitation were highlighted with labels.

IV. DISCUSSION AND CONCLUSION

The difference across years and LFM sites can be well explained by the different condition of moisture. For sites with a lower moisture level, EVI was less sensitive to the dynamics of LFM, especially when the moisture continued to decrease after the peak of growing season. The coastal site tended to reflect a greater magnitude of EVI throughout the year, but a greater magnitude of LFM only seasonally compared to an inland site. An inter-seasonal assessment of the strength of correlation may address this. As the site to site variations of LFM were not similar to the type of variations of EVI, combining data collection from all did not provide a single empirical model to be utilized at multiple locations (Fig. 4). Unknowns such as the comparison of type of vegetation and spatial coverage of vegetation my attribute to this.

In the inter-annual comparison of correlation, very dry years (e.g. 2007) not only had an overall lower but also less dynamic LFM (Fig. 5), which was also reflected in model performance (Fig. 6). Drought, especially the prolonged ones, can change the pigment and plant structure and the leaf water absorption, and thus weaken the reflectance in NIR and SWIR bands, which made EVI difficult to capture the dynamics of LFM. The increase in vegetation moisture can reach a

saturation point, thus weakening the relationship of the strength of correlation with the inclusion of precipitation during relatively wet years (Fig. 6).

Although the moisture condition had a great impact on the correlation between VIs and LFM, adding climatological variables may not help to improve the predictability of LFM from remotely sensed images, mainly due to the collinearity between climatological variables and VIs. Instead, due to the inconsistency of correlation between LFM and VIs in time and space, LFM estimation using the remote sensing data should include variables that can address the inter-annual difference and inter-location difference [3], such as inter-annual statistics, cross-site statistics, and even topographical variables. Investigating a non-linear model can also be helpful, due to the saturation and insensitivity of VIs among the extremely low and high LFM values. Other than that, retrieving LFM from the radiance transfer models (RTMs) should be considered to replace the traditional empirical models. Compared with empirical relationship built upon in-situ measurements, RTMs are physically meaningful and directly address the response in reflectance when vegetation water content changes. A recent study focused on Australia has shown promising outcome of deriving LFM using an inversing technique of RTM [9]. In the future study, we will apply RTMs to achieve a better estimation of LFM.

This study demonstrates the benefits freely accessible MODIS data with a relatively high temporal resolution can contribute to measuring LFM data by improving cost and time efficiency. Results indicate that a strong linear relationship is regionally site specific. Applying a multivariate analysis by introducing additional climatological variables may decrease interannual variation and strengthen single site correlation. Further investigation into site characteristics has the possibility to expand applications for using remote sensing satellite data to assess wildfire risk globally in shrubland ecosystems within a Mediterranean climate.

ACKNOWLEDGEMENT

We thank USFS Wildland Fire Assessment System (WFAS) for hosting and providing in situ LFM observation data. We also thank Dr. Son V. Nghiem from NASA Jet Propulsion Laboratory for research idea, discussion and experiment design.

REFERENCES

- [1] D. A. Roberts, P. E. Dennison, S. Peterson, S. Sweeney, and J. Rechel, "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*, vol. 111, Aug 30 2006.
- [2] P. Schneider, D. A. Roberts, and P. C. Kyriakidis, "A VARI-based relative greenness from MODIS data for computing the fire potential index," *Remote Sensing of Environment*, vol. 112, pp. 1151-1167, Mar 18 2008.
- [3] S. H. Peterson, D. A. Roberts, and P. E. Dennison, "Mapping live fuel moisture with MODIS data: A multiple regression approach," *Remote Sensing of Environment*, vol. 112, pp. 4272-4284, Dec 15 2008.
- [4] B. Myoung, S. H. Kim, S. V. Nghiem, S. Jia, K. Whitney, and M. C. Kafatos, "Estimating Live Fuel Moisture from MODIS Satellite Data for Wildfire Danger Assessment in Southern California USA," *Remote Sensing*, vol. 10, p. 87, 2018.
- [5] M. Yebra, P. E. Dennison, E. Chuvieco, D. Riano, P. Zylstra, E. R. Hunt, et al., "A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products," *Remote Sensing of Environment*, vol. 136, pp. 455-468, Sep 2013.
- [6] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote sensing of environment*, vol. 83, pp. 195-213, 2002.
- [7] B. C. Gao, "NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sensing of Environment*, vol. 58, pp. 257-266, Dec 1996.
- [8] A. A. Gitelson, Y. J. Kaufman, R. Stark, and D. Rundquist, "Novel algorithms for remote estimation of vegetation fraction," *Remote Sensing of Environment*, vol. 80, pp. 76-87, 2002/04/01/2002.
- [9] M. Yebra, X. W. Quan, D. Riano, P. R. Larraondo, A. I. J. M. van Dijk, and G. J. Cary, "A fuel moisture content and flammability monitoring methodology for continental Australia based on optical remote sensing," *Remote Sensing of Environment*, vol. 212, pp. 260-272, Jun 2018.