A VARI-based relative greenness from MODIS data for computing the Fire Potential Index

P. Schneider⁎, D.A. Roberts, P.C. Kyriakidis

Department of Geography, University of California, Santa Barbara, CA, 93106, USA

Received 18 September 2006; received in revised form 25 July 2007; accepted 26 July 2007

Abstract

The Fire Potential Index (FPI) relies on relative greenness (RG) estimates from remote sensing data. The Normalized Difference Vegetation Index (NDVI), derived from NOAA Advanced Very High Resolution Radiometer (AVHRR) imagery is currently used to calculate RG operationally. Here we evaluated an alternate measure of RG using the Visible Atmospheric Resistant Index (VARI) derived from Moderate Resolution Imaging Spectrometer (MODIS) data. VARI was chosen because it has previously been shown to have the strongest relationship with Live Fuel Moisture (LFM) out of a wide selection of MODIS-derived indices in southern California shrublands. To compare MODIS-based NDVI-FPI and VARI-FPI, RG was calculated from a 6-year time series of MODIS composites and validated against in-situ observations of LFM as a surrogate for vegetation greenness. RG from both indices was then compared in terms of its performance for computing the FPI using historical wildfire data. Computed RG values were regressed against ground-sampled LFM at 14 sites within Los Angeles County. The results indicate that VARI-based RG consistently shows a stronger relationship with observed LFM than NDVI-based RG. With an average $R^2$ of 0.727 compared to a value of only 0.622 for NDVI-RG, VARI-RG showed stronger relationships at 13 out of 14 sites. Based on these results, daily FPI maps were computed for the years 2001 through 2005 using both NDVI-RG and VARI-RG. These were then validated against 12,490 fire detections from the MODIS active fire product using logistic regression. Deviance of the logistic regression model was 408.8 for NDVI-FPI and 176.2 for VARI-FPI. The c-index was found to be 0.69 and 0.78, respectively. The results show that VARI-FPI outperforms NDVI-FPI in distinguishing between fire and no-fire events for historical wildfire data in southern California for the given time period.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Fire Potential Index; MODIS; VARI; Wildfire risk; Wildfire susceptibility

1. Introduction

Wildfires are a major natural disturbance mechanism in Mediterranean regions and have been a threat to human society for centuries (Pyne et al., 1996). As the wildland–urban interface intrudes further into wilderness areas, wildfires are becoming an increasingly important issue not only affecting natural ecosystems but also imposing a risk of high economic damage to society. An example of this growing vulnerability is the October 2003 fire complex in southern California. 15 fires burned 300,000 ha, destroyed 3300 structures and cost the lives of 24 people, thus making it one of the most destructive fire events in US history (Blackwell & Tutle, 2005; Keeley et al., 2004). In a landscape such as southern California, wildfires directly threaten human lives and properties to a major extent. Indirect consequences include soil erosion and the resulting extreme surface runoff, increased emission of CO$_2$, degradation of wildlife habitat, alteration of species composition, and decrease of air quality (Barro & Conard, 1991; DeBano et al., 1998). It is therefore important to understand the physical and environmental processes that lead to increased wildfire risk, and to use this knowledge to develop accurate models for estimating the spatial and temporal patterns of risk. A spatial awareness of wildfire risk can help mitigate the environmental and economic impact of fire events especially at the wildland–urban interface. This can be accomplished by issuing warnings about spatial and temporal patterns of fire potential to population and fire management authorities. Wildfire managers see great value in
fire risk maps for allocation and planning of firefighting resources. Wildfire risk assessment is further useful for tasks such as evacuation planning or insurance estimates.

Systems for rating wildfire susceptibility that utilize weather station data have existed for many decades (San-Miguel-Ayanz et al., 2003). Examples include the National Fire Danger Rating System (NFDRS) in the United States (Burgan, 1988; Deeming, 1972) and the Canadian Forest Fire Danger Rating System (CFFDRS) (Forestry Canada Fire Danger Group, 1992; Stocks et al., 1989; Van Wagner, 1987). However, all of these systems require the use of spatial interpolation algorithms for generating spatially exhaustive maps of pertinent variables. This introduces considerable uncertainty as oftentimes the station network is not as dense as would be desirable for interpolation. Further stations located in inhomogeneous areas might not be representative of their surroundings, in particular with respect to elevation and fuels. Remote sensing offers comprehensive spatial information about fuel type, fuel properties, and fuel condition (Chuvieco, 2003; Dennison, 2003). However, while remote sensing is a valuable tool for estimating the quantity and state of live fuels (e.g. Bowyer & Danson, 2004; Ceccato et al., 2001; Keane et al., 2001; Roberts et al., 2006; Verbesselt et al., 2007; Viegas et al., 2001), meaningful fire susceptibility assessment also requires information about dead fuel moisture. Although some attempts have been made to utilize remote sensing methods for this task (Camia et al., 2003), it remains challenging as dead fuels are often obscured by a canopy of live fuels and exclusively controlled by atmospheric conditions. Meteorological data will thus always be required for estimating fire susceptibility (Aguado et al., 2003). The state of live fuels is much less dynamic in nature as it is affected by plant physiology and soil moisture. Observing live fuel state using weather data is possible using long-term moisture codes (Viegas et al., 2001), but can be more easily performed using remote sensing data (Chuvieco et al., 2003). Meteorological and remote sensing data therefore complement each other in providing a comprehensive assessment of overall fuel status.

Only a few attempts have been made to date to integrate remote sensing data with information from weather stations for wildfire susceptibility assessment. Examples include Vidal et al. (1994) and more recently the introduction of the Fire Potential Index (FPI) by Burgan et al. (1998). The FPI was originally formulated for and tested on NOAA Advanced Very High Resolution Radiometer (AVHRR) data. It has been validated with historical wildfire data by several authors (Sebastián López et al., 2002; Sudiana et al., 2003) and exhibits a strong relationship with fire frequency ($R^2 \approx 0.72$ according to Burgan et al. (1998)). It is currently offered as a nationwide experimental product through the Wildland Fire Assessment System (WFAS) at http://www.wfas.net/. Over more recent years new remote sensing technology such as the Moderate Resolution Imaging Spectroradiometer (MODIS) has emerged, which has the potential of providing better characterization of fuels than AVHRR imagery. In addition to a finer spatial resolution of 250/500 m, the MODIS sensor also samples more wavelengths than AVHRR, thus allowing the application of potentially more appropriate vegetation indices than the Normalized Difference Vegetation Index (NDVI) (Goward et al., 1991). The FPI product as offered by WFAS employs the NDVI for deriving the fuel condition using Relative Greenness (RG) (Burgan & Hartford, 1993). Several authors have shown, however, that other vegetation indices, such as the Normalized Difference Water Index (NDWI) (Gao, 1996) or the Visible Atmospherically Resistant Index (VARI) (Gitelson et al., 2002), exhibit stronger relationships with fuel condition and LFM than NDVI (Dennison et al., 2005; Roberts et al., 2006; Stow et al., 2005). It appears therefore valuable to investigate the effect of using such an index in the calculation of the FPI for a study site in southern California. The objectives of this study are 1) to evaluate the performance of VARI-based RG against NDVI-based RG as a reference using in situ data and 2) to investigate if using VARI-based RG can improve the performance of FPI in assessing wildfire susceptibility in southern California.

2. Background

2.1. Terminology

A variety of terms are used in the literature for characterizing the likelihood of wildfire events. Wildfire danger, wildfire risk, wildfire hazard, wildfire susceptibility, and wildfire potential are all used to describe similar concepts. The most widely used terms are wildfire danger and wildfire risk. The term fire danger has been criticized as being subjective since it implies a negative connotation regardless of fires also having positive effects in many regions. Adopting concepts of risk analysis in engineering, wildfire risk is generally defined as the product of the occurrence probability of a fire and its expected outcome (Bachmann & Allgöwer, 2001). In this paper we use the term fire risk only when we refer to a general context and when we explicitly intend to include a connotation of fire outcome. Since we neither study ignition probability, which is a requirement for determining occurrence probability, nor perform an analysis of the expected fire outcome in this research, we avoid the term risk otherwise and instead adopt the term wildfire susceptibility as recently introduced by Parisien et al. (2005) and Dasgupta et al. (2006). The term is objective and focuses on the vulnerability of fuels to fire at the pre-ignition stage, which is exactly what the FPI was designed for.

2.2. Vegetation indices

A vegetation index is a measure that reduces spectral information from several bands of remote sensing imagery into one number. It maximizes sensitivity to biophysical parameters of vegetation preferably with a linear response, normalizing external (solar zenith, viewing angle, etc.) and internal effects (background, topography, soil, etc.), and is coupled with a measurable biophysical parameter such as biomass, Leaf Area Index (LAI) or fuel moisture for validation purposes (Jensen, 2005). The most widely used index is the Normalized Difference Vegetation Index (NDVI) suggested by Rouse et al. (1973), which is given for MODIS bands as

$$NDVI = \frac{\rho_{857} - \rho_{645}}{\rho_{857} + \rho_{645}}.$$  

(1)
In Eqs. (1) and (2), $\rho_x$ represents the reflectance for the wavelength $x$ given in nanometers. Despite being widely used, the NDVI has been shown to be associated with several problems (Huete et al., 2002): as all ratio-based indices it is inherently non-linear and it is further impacted by substrate brightness and atmospheric contamination. The NDVI also shows scaling problems saturating in high-biomass conditions. Other vegetation indices such as the Normalized Difference Water Index (NDWI) introduced by Gao (1996), or the Normalized Difference Infrared Index (NDII) (Hardisky et al., 1983; Hunt & Rock, 1989), which uses NIR and short-wave infrared (SWIR) bands, have been developed. More recently, indices based entirely on the visible part of the spectrum have been suggested, such as the green vegetation index VI$_G$ and the Visible Atmospherically Resistant Index (VARI) (Gitelson et al., 2002). VARI is calculated as

$$\text{VARI} = \frac{\rho_{555} - \rho_{469}}{\rho_{555} + \rho_{469} - \rho_{645}}$$

(2)

and has been shown to provide the best overall results for fuel moisture estimation in Mediterranean shrublands (Roberts et al., 2006; Stow et al., 2005). This study is based on these results and adopts VARI as a promising greenness index for southern California. VARI is minimally sensitive to atmospheric effects and was developed for estimating green vegetation fraction, which is precisely the purpose of using relative greenness in the FPI algorithm. The bands used for VARI have been specifically selected for their sensitivity to vegetation fraction (Gitelson et al., 2002). In contrast to NDVI, which is sensitive to changes for small vegetation fractions and insensitive for changes at moderate and high vegetation fractions, VARI shows a linear response to vegetation fraction throughout the entire range. Based on work by Kaufman and Tanré (1992), VARI reduces atmospheric effects by subtracting the blue channel in the denominator. Gitelson et al. (2002) showed that VARI allowed the estimation of green vegetation fraction with an error of less than 10%.

2.3. The Fire Potential Index

The Fire Potential Index (FPI) was first introduced by Burgan et al. (1998) and revised by Burgan et al. (2000). A modified version of the FPI adapted to a European context was also suggested by Sebastián López et al. (2002). The inputs to the FPI consist of three raster layers, namely a fuel model map, a map of 10-hour dead fuel moisture, and a RG map. In the original paper by Burgan et al. (1998), the fuel model map was derived through a combination of NOAA-AVHRR data and expert knowledge. This map was used for deriving the dead fuel moisture of extinction $\text{MX}_d$ for each pixel. The map of 10-hour dead fuel moisture is generated by interpolating the NFDRS estimates at weather stations using weighted inverse distance squared interpolation. Finally, a RG map is used to partition each fuel model into live and dead fractions. RG was introduced by Eidenshink et al. (1990) and Burgan and Hartford (1993). It puts any current NDVI value for a pixel into perspective by normalizing it against the historical range of NDVI values for that particular pixel. It is given as

$$\text{RG} = \frac{\text{NDVI}_{\text{curr}} - \text{NDVI}_{\text{min}}}{\text{NDVI}_{\text{max}} - \text{NDVI}_{\text{min}}} \times 100\%,$$

(3)

where RG is the relative greenness for a particular pixel, NDVI$_{\text{curr}}$ is the current NDVI value and NDVI$_{\text{min}}$ and NDVI$_{\text{max}}$ are the overall historical minimum and maximum values for this pixel computed over the entire period of the time series. A similar technique based on inter-annual variability of NDVI has recently been used by Gabban et al. (2006) in their introduction of the Dynamic Relative Greenness Index (DRGI). Furthermore Chéret and Denux (2007) proposed an annual relative greenness. While a comparison of RG with DRGI and annual relative greenness would be interesting for future research, in this study we focus solely on the impact of VARI on the computation of RG and thus the FPI. In the revised version of Burgan’s paper (Burgan et al., 2000), a maximum live ratio map was also derived from the historical maximum greenness by scaling it between 35% and 75%, in order to avoid similar live ratios for fuel models representing very different vegetation types as it occurred for the original formulation of the FPI algorithm. Finally, the Fire Potential Index was computed as

$$\text{FPI} = \left(1 - \frac{\text{FM}_{10} - 2}{\text{MX}_d - 2}\right) \times \left(1 - \frac{\text{RG} \times \text{LR}_{\text{max}}}{10^5}\right) \times 100$$

(4)

where FM$_{10}$ is the 10-hour time lag fuel moisture, MX$_d$ is the dead fuel moisture of extinction derived from the fuel model parameters, RG is the relative greenness, and LR$_{\text{max}}$ is the maximum live ratio. All variables have units of percent. The resulting FPI values range from 0 to 100, where 0 is equivalent to very low fire susceptibility and 100 represents extremely high fire susceptibility.

3. Data and methods

3.1. Study area

Southern California was selected as a study site. Fig. 1 shows its geographic location and extent. The northern border was arbitrarily chosen as the northern boundary of the San Luis Obispo, Kern and San Bernadino counties, forming an approximately straight line from west to east at 35 47’ N. All other borders follow the political boundaries of California. The area of the test site is 148,079 km$^2$. Southern California is an ideal test area for this study: its Mediterranean climate and vegetation make it very vulnerable to wildfires. In addition, with many parts of southern California densely populated, wildfire risk is a serious concern especially at the wildland–urban interface. The size of the test site is small enough to provide a homogenous climate, but large enough for research on a regional scale. A large amount of high-quality geospatial data is available for southern California. Most data sets required for fire susceptibility assessment are available at high spatial resolutions and the network of weather stations is very dense, particularly in areas of high fire risk. The southern California climate provides
generally cloud-free skies during the fire season, therefore enabling extensive use of optical remote sensing data. The vegetation of the study site, especially in the fire-prone coastal mountain ranges, consists mostly of evergreen shrub species such as chamise \((\text{Adenostoma fasciculatum})\) and ceanothus \((\text{Ceanothus megacarpus} \text{ and } \text{Ceanothus crassifolius})\) as well as drought deciduous shrub species such as sage \((\text{Salvia mellifera}, \text{Salvia leucophylla})\) and sage brush \((\text{Artemisia californica})\).

### 3.2. Data

The primary data set for this study was a 6-year time series of MODIS Terra 16-day surface reflectance composites for the years 2000 to 2005. These composites were generated as described by Dennison et al. (2005). Median reflectance for each band during a 16-day period was computed using the MODIS 500 m spatial resolution daily surface reflectance product (MOD09GHK version 4). The masking of clouds, cloud shadow and snow was performed using the 1 km resolution MOD09GST product resampled to 500 m. The number of low-quality pixels included was minimized by applying a screening process for removing extreme off-nadir views using the look angle data layer. At the stage of data acquisition, each composite was already georectified and atmospherically corrected. Only the image tile h08v05 approximately covering the areas of southern California, parts of Nevada, Arizona and Northern Mexico, was considered here.

Ideally the reference data for validating RG would consist of in situ measurements of live and dead vegetation fraction. Unfortunately such a data set was not available to us for our study area. However, vegetation greenness is closely related to LFM and Roberts et al. (2006) have demonstrated a strong relationship between green vegetation fraction and LFM. We therefore validated relative greenness against a data set of LFM within Los Angeles County provided by the Los Angeles County Fire Department (LACFD). The samples were taken approximately every two weeks at 14 sites following protocols by Countryman and Dean (1979). Fig. 2 shows the locations of the sites and Table 1 lists their geographic coordinates as well as sampled species. The sites are spatially homogeneous and representative of southern California shrubland as they include all dominant species such as \(A. \text{fasciculatum, S. leucophylla, S. mellifera, C. crassifolius, and C. megacarpus}\). According to the sampling protocol, leaves and small stems (3.2 mm or less) from old and new foliage were collected from each site. All dead plant material and reproductive plant parts were removed from the sample before weighing. Furthermore, only branches that included live foliage and stems were sampled. After weighing the samples, drying them at 104 °C for 15 hours, and re-weighing them, LFM was subsequently calculated as a percentage of dry mass as

\[
\text{LFM} = \frac{m_w - m_d}{m_d} \times 100
\]

where \(m_w\) is the original wet mass of a fuel sample and \(m_d\) is the dry mass. Table 1 lists the number of samples per year for each site. Two sites burned during the study period, namely Pico Canyon in 2003 and Sycamore Canyon in 2002. They were replaced by Glendora Ridge in 2002 and Peach Motorway in 2003. Only samples for the first half of 2005 were available.

A fuel model data set of southern California was available for computing the FPI. It was produced by the Fire Resource Assessment Program (FRAP) of the California Department of Forestry and Fire Protection and has been derived from a variety of sources including a Landsat Thematic Mapper classification of primary and secondary vegetation types in conjunction with vegetation size and density, topographic information, and succession pathways based on fire history. While the fuel model data set was delivered at a 30 m spatial resolution, it was

Fig. 1. Location and extent of the southern California test site with fuel model data set from FRAP using the Albini (1976) fuel models and additional classes of non-vegetated surfaces.
resampled at a 500 m resolution by assigning to a coarser pixel the fuel model of the majority of fine-resolution pixels. The fuel model map is based on the classification scheme described by Albini (1976) and includes additional classes for non-vegetated surfaces (classes 15 through 99). Table 2 gives an overview of the fuel models and land cover types occurring in the study area. Values for dead fuel moisture of extinction were taken from Burgan et al. (1998) after crosswalking the fuel model classes to the NFDRS classification (Bradshaw et al., 1983; Deeming et al., 1978) as described in Anderson (1982).

In order to generate a raster of 10-hour dead fuel moisture, daily weather data from time series of all Remote Automated Weather Stations (RAWS) in southern California were used. These data were acquired from the Western Regional Climate Center in Reno, Nevada. A total of 178 stations were available, with the majority located in California and 9 stations situated in Nevada and Arizona. 126 stations recorded data during the entire validation period from 2001 through 2005. 48 additional stations were added between 2001 and 2002 and were included in the interpolation procedure as soon as the data became available. While the interpolation results for 2003 through 2005 are thus slightly more reliable than for the first two years of the study period, this does not affect the conclusions drawn from a relative comparison between NDVI-based FPI and VARI-based FPI as both models use identical weather inputs.

Validation of a fire index requires data on historical fire events. For this study, MODIS-derived annual cumulative fire detection data sets for 2001 through 2005 were acquired from the MODIS Active Fire Mapping Program, which is run by the Remote Sensing Applications Center (RSAC). The fire detections are based on daily Terra and Aqua MODIS imagery that is collected and processed as a joint effort between RSAC, NASA Goddard Space Flight Center and the University of Maryland. The fire detections are collected using MODIS thermal bands at a spatial resolution of 1 km. The minimum detectable fire size is approximately 100 m² for a fire temperature of 600 K or 1 m² at 1100 K. The algorithm used for creating this data set is given by Justice et al. (2002). Since the exact sub-pixel location and the size of the fire within the pixel are not known, the data set used for this study records the

Fig. 2. Locations of the live fuel moisture sample sites in northern Los Angeles County.
centred of the pixel for which a fire has been detected. A total number of 12,804 fire detections were available within the study area for the entire 5-year validation period. The majority of fire detections occurred during summer and fall, so daily FPI mages were only computed between April 1 and November 31 for each year and only fire detections within this period were used. Some fire detections also were located within data gaps in the FPI maps. Thus the usable number of MODIS fire detections was slightly reduced to 12,490.

### 3.3. Computing and validating VARI-based RG

While current applications of the FPI use the NDVI to calculate RG, this is not necessarily a requirement. RG can be computed using any remote sensing index that shows a relationship with vegetation greenness. For this paper, RG was computed using VARI such that Eq. (3) simply becomes

$$\text{RG}_{\text{VARI}} = \frac{\text{VARI}_{\text{curr}} - \text{VARI}_{\min}}{\text{VARI}_{\max} - \text{VARI}_{\min}} \times 100\%.$$  \hfill (6)

where $\text{RG}_{\text{VARI}}$ is the VARI-based relative greenness for a particular pixel, $\text{VARI}_{\text{curr}}$ is the current VARI value at that location and $\text{VARI}_{\min}$ and $\text{VARI}_{\max}$ are the actual overall historical minimum and maximum VARI values for the given pixels as they occurred over the entire span of the time series (2000 through 2005).

The 129 available MODIS surface reflectance composites were subset to cover the southern California study site. For each of the composites, both NDVI and VARI were then computed. Finally, RG was derived from both indices for all 6 years according to Eqs. (3) and (6) and stored as an image stack consisting of 129 bands. For the statistical analysis, values of RG computed for the pixels at each LFM sampling location were extracted from the stack of RG images. A variety of different techniques exist for performing this task. The most intuitive concept would be to extract exactly that 1 pixel in which the sample location falls. While this is a valid approach (Roberts et al., 2006), many authors have either used the average value of $3 \times 3$ pixel windows around each site (Dennison et al., 2005) or manually selected individual pixels in the vicinity of the site, which were located in homogeneous areas identified from aerial photography (Stow et al., 2005). For this project we chose to use both a single pixel and a $3 \times 3$ average approach and compared their respective performances. Several sites are situated close to urban areas making the use of a $3 \times 3$ pixel window potentially problematic. However, with the exception of Laurel Canyon and Schueren Road (Table 1, Fig. 2) all sites are located within sufficiently large and homogeneous patches of vegetation to render the impact of urban land cover negligible for a $3 \times 3$ pixel window. After extracting the NDVI-based RG and VARI-based RG using both strategies for all test sites, the resulting four data sets were regressed against LFM data sampled by LACFD, in order to evaluate their respective relationships. The pairing between the LFM samples and the MODIS composites was performed using temporal proximity to the midpoint of the compositing period as a measure. Relationships at all sites were linear, so only linear models were tested in the regression. The significance of the increase in $R^2$ was evaluated following the methodology of Dennison et al. (2005) by normalizing the corresponding correlation coefficients using a Fisher z-transform (Fisher, 1915; Papoulis, 1990) given as

$$z_f = \arctanh(r) = \frac{1}{2} \ln \left( \frac{1 + r}{1 - r} \right)$$  \hfill (7)

where $r = \sqrt{R^2}$ is Pearson’s correlation coefficient. After computing the difference between $z_f$ for VARI-r and NDVI-r according to

$$Z_{\text{diff}} = \frac{z_{\text{VARI}} - z_{\text{NDVI}}}{\sqrt{\frac{1}{n_{\text{VARI}} - 3} + \frac{1}{n_{\text{NDVI}} - 3}}}$$  \hfill (8)

(Papoulis, 1990), where $n_{\text{VARI}}$ and $n_{\text{NDVI}}$ indicate the respective number of available samples, the significance of the increase in correlation strength was evaluated using a one-tailed test.

### 3.4. MODIS-based FPI for Southern California

In order to test the applicability of MODIS imagery and VARI for computing the FPI, a system for generating FPI maps
using MODIS data was developed. For validation purposes, daily FPI maps of southern California were generated using both traditional NDVI-RG and VARI-RG for the summer and fall seasons (April 1 through November 30) for 5 years (2001 through 2005) using daily weather data and the latest available surface reflectance composite. Active fire data for the year 2000 were not available. Meteorological data were utilized to compute spatially exhaustive estimates of dead fuel moisture for the entire study site. Measurements of daily maximum air temperature and daily minimum relative humidity were used to represent the generally high fire susceptibility in the early afternoon. In order to work with a clean meteorological data set, outliers of more than three standard deviations from the mean were marked and the associated stations removed from the pool of stations for the particular day. Kriging with external drift (KED) was applied for spatially interpolating the station data to rasters (Kyriakidis et al., 2001). KED is a variant of traditional geostatistical interpolation that incorporates auxiliary information within the interpolation process. For air temperature this approach resembles topographically informed interpolation which has been used by several authors (e.g. Dodson & Marks, 1997; Willmott & Matsuura, 1995), but distinguishes itself by providing a stochastic link between the variables and accounting for uncertainty in the trend estimation. In contrast to simple DEM-assisted interpolation, which reduces air temperature at each station to sea level using an average environmental lapse rate, then interpolates the de-trended data and finally adjusts the result by elevation using a DEM, KED can be used independently of the nature of the variable. In this study we therefore used KED for interpolating relative humidity using temperature and elevation as the auxiliary information. Using the interpolated raster data sets of air temperature and relative humidity, 10-hour timelag fuel moisture was then computed for each pixel following Fosberg and Deeming (1971) and Sebastián López et al. (2002) as

$$FM_{10hr} = 1.28 \times \text{emc}$$

where emc is the equilibrium moisture content (i.e. the moisture content at which the net exchange of moisture between fuel and environment is equal to zero), given as

$$\text{emc} = \begin{cases} 
2.22749 + 0.160107 \times RH & \text{if } 10\% \leq RH \leq 50\% \\
-0.014784 \times T_{\text{air}}^2 & \text{if } RH > 50\% \\
21.0606 + 0.005565 \times RH \times T_{\text{air}} & \text{if } RH \leq 10\% \\
-0.00035 \times RH \times T_{\text{air}} & \\
-0.483199 \times RH & \\
0.03229 + 0.281073 \times RH & \\
-0.000578 \times RH \times T_{\text{air}} & 
\end{cases}$$

where RH is the relative humidity in % and $T_{\text{air}} = \frac{9}{5}T_c + 32$ with $T_c$ being the air temperature in degrees Celsius. These empirical equations were developed for non-metric units, thus SI units need to be converted accordingly prior to application.

The FRAP fuel model data set for California was used to produce a raster of the Dead Fuel Moisture of Extinction as it is required for the FPI. This was accomplished by assigning values of extinction moisture to each fuel model type according to the data given in Burgan et al. (1998) (see Table 2). FPI was computed following Eq. (4) using the input rasters of 10-hour fuel moisture, dead fuel moisture of extinction, NDVI-RG and VARI-RG derived from MODIS composite time series, and maximum live ratio.

### 3.5. Validation of FPI using logistic regression

Validation of a fire susceptibility index is challenging as it is impossible to directly measure the quantity in question at a given location and time directly in the field. An indirect alternative approach adopted by most researchers thus relies on an analysis of the statistical relationship between the index and data on historical wildfire events. Such an approach has been adopted by Burgan et al. (1998) in the original suggestion of the FPI, as well as for validating the FPI in Mediterranean areas of Europe (Sebastián López et al., 2002) and in Indonesia (Sudiana et al., 2003). A variety of other indices based on remote sensing and weather data have been validated with ground data for savanna ecosystems by Verbesselt et al. (2006). Viegas et al. (2000) investigated the performance of several fire susceptibility indices for France, Italy, and Portugal. Lasaponara (2005) studied the performance of AVHRR-derived indices with fire activity data in southern Italy. Various components of the NFDRS have been validated with ground data by several authors including Haines et al. (1983), Mees and Chase (1991), Andrews et al. (2003), and Peng et al. (2005).

When using fire activity data for regression-based validation of fire susceptibility indices, the response variable is generally dichotomous (i.e. the outcome is binary). Logistic regression (Collett, 2003; Harrell, 2001; Hosmer & Lemeshow, 2000) is most often used in such cases. It has been successfully applied in validating wildfire susceptibility indices in the past. Examples include the evaluation of the performance of NFDRS fire danger indices (Andrews & Loftsgaarden, 1992; Andrews et al., 2003), the comparison of fire risk indicators in savanna ecosystems (Verbesselt et al., 2006), and fire occurrence probability modeling (Lozano et al., 2007). Logistic regression was also applied in this study for evaluating the performance of VARI-based FPI. The specific form of the logistic regression model used here is

$$\pi(x) = E(Y = 1|x) = \frac{e^{\beta_0 + \beta_1x}}{1 + e^{\beta_0 + \beta_1x}} = \frac{1}{1 + e^{-\beta_0 - \beta_1x}}$$

where $\pi(x)$ expresses the conditional probability of a fire occurring (i.e. the binary response variable $Y$ is equal to 1) for a given FPI value $x$ and $\beta_0$ and $\beta_1$ are the regression coefficients. The logit transformation $g(x)$ is defined as

$$g(x) = \ln \left[ \frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1x$$

and thus has many of the desirable properties of a linear regression model (Hosmer & Lemeshow, 2000).
A data set of compiled MODIS Active Fire detections (see Section 3.2) was used as the response variable with the predictor variable being the FPI. In contrast to the validation of RG with LFM observations, which was limited to shrub vegetation, the validation of FPI using historical fire data was carried out using fire detections for all vegetated areas, in order to evaluate the overall model performance. Daily FPI maps were generated at a 500 m spatial resolution between April 1 and November 31 for the years 2001 through 2005 for both the NDVI-based and the VARI-based method. Otherwise identical inputs were used for both methods, ensuring that all differences can be attributed to the impact of the vegetation index used. For each of the 12,490 MODIS fire pixel centroids the spatially closest FPI value was extracted for the same day as the fire detection. To generate the no-fire data set, an equal number of unburned pixels were randomly sampled from the entire study area and all 5 years and their respective FPI value was recorded. In contrast to previous applications of logistic regression within the context of fire potential validation, the explicitly spatial nature of the FPI makes the size of the no-fire sample somewhat arbitrary. The vast majority of pixels are not associated with actual fire, thus the potential no-fire sample size is only limited by the overall number of pixels generated for the entire validation period (roughly 10^9 in this case). While it would have been possible to sample a much larger number of no-fire pixels, we chose the no-fire sample size to be equal to the number of available fire pixels. The resulting total sample size (fire and no-fire) of 24,980 is high enough to be representative while preserving reasonable fire/no-fire proportions for the logistic regression and drastically limiting the computational expense. However, it is important to note that the fire/no-fire proportions presented here are therefore only applicable for statistical analysis but do not allow any meaningful physical interpretation in the sense of a fire probability.

For this study logistic regression was performed using the following steps:

1. A total of 100 FPI classes corresponding to floating-point FPI values between 0–1, 1–2,..., 99–100 were defined for the analysis.

2. For each class the proportion of fire pixels was computed. Only classes containing a minimum number of 25 pixels were considered for the subsequent analysis in order to derive a representative sample.

3. A logistic model was fitted to the data using a maximum likelihood approach (Hosmer & Lemeshow, 2000). The regression coefficients were obtained iteratively by maximizing the likelihood function \( L(\beta_0, \beta_1) = \prod_{i=1}^{n} \pi \left( x_i \right)^{y_i} \left( 1 - \pi \left( x_i \right) \right)^{1-y_i} \) where \( x_i \) denotes the independent variable and \( y_i \) the dichotomous outcome variable for \( i=1, ..., n \) and \( \pi \) denotes the conditional probability obtained from Eq. (11).

4. The regression was weighted by the total number of available fire/no-fire pixels in each class.

In contrast to ordinary linear regression, there is no widely accepted equivalent to \( R^2 \) as a measure of the overall model fit for logistic regression. It is possible to compute several different pseudo-\( R^2 \) values, but since low values are the norm for them, Hosmer and Lemeshow (2000) suggest limiting their use to comparing competing models and to avoid interpretation in an absolute way as in linear regression. The low pseudo-\( R^2 \) values for logistic regression are easily misinterpreted by those who are accustomed to \( R^2 \) values obtained from ordinary linear regression (Andrews et al., 2003; Hosmer & Lemeshow, 2000). Out of 12 studied pseudo-\( R^2 \) measures for logistic regression Mittlböck and Schomper (1996) recommend using either the squared Pearson correlation coefficient or a linear regression-like sum-of-squares \( R^2_{SS} \) computed as

\[
R^2_{SS} = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{\pi}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}
\]

(13)

where \( y_i \) are the binary values of dependent variable, \( \hat{\pi}_i \) denote the estimates from the logistic regression, and \( \bar{y} \) denotes the overall proportion.

The c-index, which is identical to the area under the ROC (Receiver Operating Characteristic) curve (Harrell, 2001; Hosmer & Lemeshow, 2000), was used to evaluate the
The discrimination power of the logistic model. It is computed by pairing all no-fire samples with all fire samples and comparing the fitted probabilities for each pair. The c-index results from dividing the sum of all cases for which the fitted probability of the fire sample is greater than the fitted probability of the no-fire sample by the total number of pairs. A c-index value of 0.5 indicates no discrimination of the model while a c-index value of 1 indicates perfect discrimination. The model likelihood ratio $\chi^2$ statistic and its associated $p$-value was used to evaluate the significance of the predictor for each individual model.

4. Results and discussion

In this section we first present the results obtained from a regression analysis of NDVI-RG and VARI-RG with LFM at 14 sites in Los Angeles County. We then present a case study applying MODIS-derived VARI-RG for computing FPI in...
southern California and compare its performance to NDVI-FPI validating both products with historical fire detections using logistic regression.

4.1. Performance of VARI-based RG

Fig. 3 shows scatter plots of NDVI-based RG against measured LFM at all 14 sample sites for the $3 \times 3$ pixel extraction window case. With $R^2$ values greater than 0.7, four sites showed a very strong correlation with LFM, namely Bitter Canyon 1, Clark Mountainway, Peach Motorway, and Glendora Ridge. Several other sites have acceptable $R^2$ values between 0.5 and 0.7. Two sites, namely Laurel Canyon and Schuern Road have weak correlations. Caution needs to be applied in interpreting the high correlation coefficients for Peach Motorway and Glendora Ridge. These sites were established recently due to other sites being burned, and thus have a smaller number of data points than the other sites (see Table 1). The relationships for these two sites are therefore not necessarily representative.

Fig. 4 illustrates the corresponding results for VARI-based RG. The scatter plots show stronger relationships between RG values.
and measured LFM than for the NDVI-based case. The $R^2$ values are higher for all sites, with the exception of Bitter Canyon 2 and Peach Motorway. Eight out of 14 sites have very strong correlations with $R^2$ values greater than 0.7. No site had an $R^2$ value of less than 0.5. This indicates that VARI-based RG has a stronger relationship with LFM and might be a good substitute for NDVI-RG in calculations of the FPI for southern California.

Table 3 confirms this conclusion and in addition shows comparable results for the 1 pixel extraction window. In nearly all cases, VARI-based RG shows a stronger relationship with observed LFM than NDVI-based RG. This is true independent of the size of the extraction window. However, we can also observe a trend showing an increase in $R^2$ with an increasing size of the extraction window if we look at each index individually. This trend can be attributed to an elimination of georectification issues and to averaging of potential site inhomogeneities. For the majority of sites (11 out of 14), VARI-based RG for a $3 \times 3$ pixel window showed the strongest correlation with LFM. Only 3 sites showed a different behavior, 2 of which still showed the optimum $R^2$ for VARI-RG but for a $1 \times 1$ pixel window. Only the Peach Motorway site had its maximum $R^2$ for the $3 \times 3$ pixel NDVI-RG case, but as mentioned above needs to be treated with caution due to a smaller sample size. Table 3 also shows the difference in Fisher $z$-scores and the corresponding $p$-values. For the $1 \times 1$ pixel window, 9 out of 14 sites showed improvements that were statistically significant at the 0.95 level. Seven sites showed significant improvements for the $3 \times 3$ pixel window. The two sites that showed slightly lower $R^2$ values for the VARI case (Bitter Canyon 2 and Peach Motorway, each for the $3 \times 3$ pixel window) were found to have an insignificant $R^2$ decrease at the 0.99 level.

The relationships between RG and LFM are obviously site specific. One of the major factors contributing to the differences is the number of available samples. The newly established sites (Peach Motorway and Glendora Ridge) with fewer available samples tend to have very high $R^2$ values due to exclusion of the 2002 data, which decreased the strength of the correlations in most cases. The sites with the weakest relationships overall are Laurel Canyon, Schueren Road, and Trippet Ranch. The first two sites are influenced by the presence of urban land cover, in particular for the $3 \times 3$ extraction window. While only samples of C. crassifolius were used for Trippet Ranch, the site is co-dominated by A. fasciculatum, thus resulting in a relatively weak overall relationship compared to the other sites.

Figs. 3 and 4 show two general classes of relationships. The linear models for Bitter Canyon 1, Bouquet Canyon, and Trippet Ranch have substantially steeper slopes than those of the other sites. This difference can be attributed to the impact of two different plant functional types. Most sites are dominated by evergreen shrub species such as chamise (A. fasciculatum) and ceanothus (C. megacarpus and C. crassifolius). Bitter Canyon 1 is dominated by a drought deciduous shrub (S. leucophylla) while Bouquet Canyon and Trippet Ranch include S. mellifera as a co-dominant with an evergreen shrub (Table 1). The observed importance of plant functional type is consistent with the findings by Roberts et al. (2006).

Fig. 5 illustrates the relative change in $R^2$ values when VARI is used for the computation of RG. It shows the results for both the $1 \times 1$ pixel window case as well as the $3 \times 3$ pixel window case. All percentage values presented here can be interpreted as the relative change in the proportion of LFM variance explained by using VARI-RG in lieu of NDVI-RG as the predictor variable, e.g. a 100% relative increase is equivalent to a doubling of the percentage of variance explained by the independent variable. It is obvious from Fig. 5 that the majority of sites show a relative increase in the strength of the correlation. Only two sites, namely Bitter Canyon 2 and Peach Motorway, showed a slight weakening of the relationship for the $3 \times 3$ pixel window case, with values of $-3.8\%$ and $-0.6\%$, respectively. The greatest relative improvements overall were found for those sites that showed weak correlations for NDVI-RG, indicating that VARI-RG can substantially improve the relationships for some problematic NDVI-RG cases. This is true for both Laurel Canyon and Schueren Road with increases of $49\%$–$175\%$ (Laurel Canyon) and $47\%$–$60\%$ (Schueren Road) over the NDVI-RG $R^2$. Trippet Ranch also showed a substantial strengthening of the correlation by $33\%$ to $73\%$. The MODIS footprint for Schueren Road and Laurel Canyon is to some extent affected by urbanization. The strong improvements for these sites with VARI-RG thus indicate that VARI is potentially less sensitive to the impact of urban surfaces than NDVI. This agrees with the hypothesis by Stow et al. (2005), who suggest that VARI might be less sensitive than other vegetation indices to background material reflectance and spatial inhomogeneities at the subpixel level.

Table 3 $R^2$ values between RG based on NDVI and VARI for all 14 LFM sample sites for two different extraction window sizes

<table>
<thead>
<tr>
<th></th>
<th>1 pixel window</th>
<th>3 × 3 pixel window</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NDVI</td>
<td>VARI</td>
</tr>
<tr>
<td>Bitter Canyon 1</td>
<td>0.775</td>
<td>0.848</td>
</tr>
<tr>
<td>Bitter Canyon 2</td>
<td>0.631</td>
<td>0.642</td>
</tr>
<tr>
<td>Bouquet Canyon</td>
<td>0.641</td>
<td>0.711</td>
</tr>
<tr>
<td>Clark Mountainway</td>
<td>0.606</td>
<td>0.737</td>
</tr>
<tr>
<td>La Tuna Canyon</td>
<td>0.490</td>
<td>0.657</td>
</tr>
<tr>
<td>Laurel Canyon</td>
<td>0.175</td>
<td>0.483</td>
</tr>
<tr>
<td>Pico Canyon</td>
<td>0.592</td>
<td>0.771</td>
</tr>
<tr>
<td>Placerita Canyon</td>
<td>0.512</td>
<td>0.661</td>
</tr>
<tr>
<td>Schueren Road</td>
<td>0.315</td>
<td>0.462</td>
</tr>
<tr>
<td>Sycamore Canyon</td>
<td>0.565</td>
<td>0.762</td>
</tr>
<tr>
<td>Trippet Ranch</td>
<td>0.338</td>
<td>0.587</td>
</tr>
<tr>
<td>Woolsey Canyon</td>
<td>0.577</td>
<td>0.721</td>
</tr>
<tr>
<td>Peach Motorway</td>
<td>0.732</td>
<td>0.803</td>
</tr>
<tr>
<td>Glendora Ridge</td>
<td>0.689</td>
<td>0.702</td>
</tr>
<tr>
<td>Average $R^2$</td>
<td>0.546</td>
<td>0.682</td>
</tr>
</tbody>
</table>

The best overall $R^2$ value for each site is highlighted in bold. For all correlations $p<10^{-11}$. The regressions included all available data sets from 2000 to 2005. See Table 1 for detailed number of data points. The table also shows the difference in Fisher $z$-scores $Z_{diff}$ (see Section 3.3) and the corresponding significance level ($p$). $p$-values indicating a significance of 0.95 or more are shown in italics.
The remaining sites (i.e. mostly those for which NDVI-based RG already achieved fairly high $R^2$ values) showed smaller relative improvements between 2% (Glendora Ridge) and 34% (La Tuna Canyon). Fig. 5 furthermore indicates that the magnitude of relative $R^2$ increase is higher for the 1×1 pixel window case for 11 sites. Only a few sites, namely Placerita Canyon, Schueren Road, and Glendora Ridge showed stronger improvements for the 3×3 pixel window. The unusual behavior of Laurel Canyon with extreme differences between the NDVI- and VARI-based approaches can again be attributed to the size and location of the site. The Laurel Canyon site is only around 250 m × 600 m in size and is surrounded by a densely populated residential area. It is thus severely affected by subpixel mixing effects. In this case the extraction approach using 1 pixel is preferable over the 3×3 pixel window case, for which the signal is contaminated by urban land cover surrounding the site.

Overall, while RG is generally derived using NDVI (Aguado et al., 2003; Burgan & Hartford, 1993; Burgan et al., 1998),
VARI-based RG appears to show a stronger correlation with measured LFM for shrubland ecosystems. It might thus be valuable to integrate VARI into remote sensing-supported wildfire susceptibility assessment in such areas. The physical reason for the better performance of VARI-based RG compared to NDVI-based RG is not yet fully understood. VARI operates entirely in the visible part of the spectrum and several authors (e.g. Boyer and Danson (2004)) have shown that visible reflectance is not sensitive to changes in moisture status at the leaf-scale (Roberts et al., 2006). At canopy scales, however, visible reflectance is correlated with vegetation cover and LAI as shown by Gitelson et al. (2002) and Davis and Roberts (1999). The good correlation of VARI with LFM appears to be associated with green vegetation fraction. VARI was developed for estimation of vegetation cover and Gitelson et al. (2002) have shown for wheat that VARI exhibits a linear relationship with green vegetation fraction while the same relationship for NDVI becomes nonlinear above a fraction of 50% (Roberts et al., 2006). While not tested within the scope of this study, a potential equally linear relationship between VARI and green vegetation fraction for Mediterranean shrublands could explain the higher correlations obtained for VARI-based RG and LFM. These findings are consistent with those by Stow et al. (2005) who hypothesized that the temporal variability of green vegetation fraction co-varies with LFM making VARI a useful predictor of LFM.

4.2. Validation of VARI-based FPI using historical fire data

Fig. 6 shows box plots comparing the summary statistics of fire pixels and no-fire pixels for VARI-based FPI with NDVI-based FPI as a reference. As would be anticipated, the median for fire pixels is greater than the median for no-fire pixels in both cases. There is a substantial overlap between fire and no-fire classes, which is due to the fact that high fire susceptibility does not necessarily result in fires if there is a lack of an ignition source. The VARI-based FPI produces slightly higher FPI values overall. The median of the VARI no-fire class (62.5) lies 5.3 units above that of the NDVI no-fire class (57.2). While for the NDVI case the difference between the fire/no-fire medians is only 10.2 units, the same difference is 16.0 units for the VARI case, indicating a better separability between the two classes for VARI-based FPI. It is also evident that the extreme values for the fire pixels show a smaller range for the VARI-FPI case (47 units) than for the NDVI-FPI case (68 units). Active fire pixels are therefore less likely to be associated with low index values for the VARI case.

The better separability between the fire/no-fire classes is further supported by histograms for fire and no-fire classes of the NDVI-based and the VARI-based FPI (Fig. 7). While the histograms for no-fire pixels have similar shapes in both cases, the fire pixel histograms show obvious differences. The VARI-based FPI confines the actual fire pixels to a smaller range of higher FPI values.

Fig. 8 shows logistic regression models fitted to the proportion of MODIS active fire pixels for each FPI class for both NDVI-FPI and VARI-FPI. The proportion describes the fraction of fire pixels out of all pixels for an FPI class, i.e. a value of 1 indicates that all pixels of a given class were detected as burning. The fitted models are given for the NDVI-FPI case as

$$\hat{\pi}_{\text{NDVI}} = \frac{e^{-3.1779 + 0.0485 \times \text{FPI (NDVI)}}}{1 + e^{-3.1779 + 0.0485 \times \text{FPI (NDVI)}}}$$

and as

$$\hat{\pi}_{\text{VARI}} = \frac{e^{-5.4157 + 0.0757 \times \text{FPI (VARI)}}}{1 + e^{-5.4157 + 0.0757 \times \text{FPI (VARI)}}}$$

for the VARI-FPI case. A visual inspection of Fig. 8 suggests an overall better fit for the VARI-FPI case. The model probability range is an important indicator of overall model performance. Ideally, the logistic regression curve would start at 0 for FPI=0 and reach 1 for FPI=100. While neither model achieves this

![Fig. 7. Histograms for fire and no-fire pixels for NDVI-based FPI (left) and VARI-based FPI (right).](image-url)
goal, the model fitted for VARI-FPI comes closer to the ideal case. For FPI=0 the y-axis offset for the VARI-FPI model is only 0.0044 while the traditional NDVI-FPI case shows an offset of 0.0400. For FPI=100 the NDVI-FPI model reaches only 0.8430 while the VARI-FPI model has a slightly higher value of 0.8964.

The higher y-axis intercept of the NDVI-FPI model indicates a greater number of fire detections for low FPI values than for the VARI-FPI case. For example, the NDVI-FPI range between 0 and 50 corresponded to 34% of the burned pixels. Considering that a low FPI should be an indicator of low fire probability, the fact that over a third of the fires were associated with a low NDVI-FPI is an indicator of poor performance. In contrast, a VARI-FPI range between 0 and 50 translated to 17% burned pixels, thus indicating that a low VARI-FPI is less likely to be erroneously associated with a fire event than NDVI-FPI. The model fitted to VARI-FPI also has a steeper slope than that for NDVI-FPI, in particular for a range of index values between 50 and 90. VARI-FPI thus appears to have a greater sensitivity to fire events for moderately high index values. Fig. 8 also shows that high VARI-FPI values are more likely to be associated with an actual fire event than NDVI-FPI values. The five highest NDVI-FPI classes present in the data account for only 63% of the fire pixels on average, while the five highest VARI-FPI classes can explain 78% of the fire pixels on average.

The regression statistics (see Table 4) confirm the better model fit for the VARI-FPI. The most commonly used measure of model fit in logistic regression is deviance, which is similar to the residual sum of squares in ordinary linear regression. Essentially, a smaller deviance value suggest a better fit of the model. For the given data set, deviance in the NDVI-FPI case is 408.8 and the deviance for the VARI-FPI case is 176.2, indicating that the VARI-based FPI performs better at distinguishing between fire and no-fire events for the actual historical fire data.

The $R^2_{SS}$ for the NDVI-based FPI was estimated as 0.166 while the one for VARI-based FPI resulted in a value of 0.267. Just as for deviance, this measure also confirms that VARI-FPI agrees better with historical fire data. It is important to note once more that the given $R^2_{SS}$ values are helpful as relative measures for evaluation of competing models but are not comparable to $R^2$ values from ordinary linear regression.

The c-index for the NDVI-FPI model reached a value of 0.69, while the VARI-FPI model achieved a c-index of 0.78 and thus displayed greater discriminatory power. c-index values of 0.7 or greater are generally considered to offer acceptable discrimination, while a value of 0.8 or greater is considered to have excellent discriminatory power (Hosmer & Lemeshow, 2000). While both NDVI-FPI and VARI-FPI offer acceptable discrimination of fire susceptibility in southern California (with NDVI-FPI being at the lower end of the range of acceptable values), VARI-FPI is more reliable and is even close to offering ‘excellent’ discriminatory power. The model likelihood ratio $\chi^2$ test and its associated p-value of $p<0.001$ indicate that FPI is a significant predictor variable for each individual model.

While a validation of a fire susceptibility index such as the FPI with historical fire event data is certainly valuable it is also somewhat problematic since even extremely high index values cannot guarantee a fire outbreak because an ignition source might be lacking. This fact is evident in the relatively large overlap between the fire and no-fire classes in Figs. 6 and 7.

The c-index for the NDVI-FPI model reached a value of 0.69, while the VARI-FPI model achieved a c-index of 0.78 and thus displayed greater discriminatory power. c-index values of 0.7 or greater are generally considered to offer acceptable discrimination, while a value of 0.8 or greater is considered to have excellent discriminatory power (Hosmer & Lemeshow, 2000). While both NDVI-FPI and VARI-FPI offer acceptable discrimination of fire susceptibility in southern California (with NDVI-FPI being at the lower end of the range of acceptable values), VARI-FPI is more reliable and is even close to offering ‘excellent’ discriminatory power. The model likelihood ratio $\chi^2$ test and its associated p-value of $p<0.001$ indicate that FPI is a significant predictor variable for each individual model.

While a validation of a fire susceptibility index such as the FPI with historical fire event data is certainly valuable it is also somewhat problematic since even extremely high index values cannot guarantee a fire outbreak because an ignition source might be lacking. This fact is evident in the relatively large overlap between the fire and no-fire classes in Figs. 6 and 7.

The c-index for the NDVI-FPI model reached a value of 0.69, while the VARI-FPI model achieved a c-index of 0.78 and thus displayed greater discriminatory power. c-index values of 0.7 or greater are generally considered to offer acceptable discrimination, while a value of 0.8 or greater is considered to have excellent discriminatory power (Hosmer & Lemeshow, 2000). While both NDVI-FPI and VARI-FPI offer acceptable discrimination of fire susceptibility in southern California (with NDVI-FPI being at the lower end of the range of acceptable values), VARI-FPI is more reliable and is even close to offering ‘excellent’ discriminatory power. The model likelihood ratio $\chi^2$ test and its associated p-value of $p<0.001$ indicate that FPI is a significant predictor variable for each individual model.

While a validation of a fire susceptibility index such as the FPI with historical fire event data is certainly valuable it is also somewhat problematic since even extremely high index values cannot guarantee a fire outbreak because an ignition source might be lacking. This fact is evident in the relatively large overlap between the fire and no-fire classes in Figs. 6 and 7.
The MODIS Active Fire product is a valuable data set for validating fire susceptibility indices, but there is a potential issue related to its use. Larger fires are comprised of several fire detections, which are likely to have similar environmental conditions and thus are spatially and temporally correlated. Treating these independently can possibly cause over- or underestimation of model performance. However, this issue should not affect the relative comparison of model performance of NDVI-FPI versus VARI-FPI assuming that the effect is identical for both explanatory variables. Further research will be necessary to understand the impact of this issue on model performance.

The VARI-based approach to computing the FPI is promising overall, but also has limitations. While VARI is a suitable greenness index for the Mediterranean ecosystems of southern California, it is unlikely to perform as well on a nationwide scale. However, the use of MODIS imagery and its finer spectral resolution as the remote sensing input to the FPI allows for the application of a greater variety of different vegetation indices than NOAA-AVHRR imagery as it is used by WFAS for computing the FPI. Thus a user interested in computing regional FPI maps has the possibility to select the vegetation index that has been shown to perform best for the particular area.

Another drawback of our approach is the limited length of the MODIS time series for computing relative greenness. MODIS data have only been available since the year 2000, while the time series of NOAA-AVHRR data currently used by WFAS ranges from 1979 to 2000 and can thus give more reliable estimates of relative greenness. However, more MODIS data will be integrated within the current system as they become available, so the RG estimates will improve over time. Alternatively, computation of RG could be replaced by fuel condition estimates from spectral mixture analysis (SMA) as described by Roberts et al. (1993) or Roberts et al. (2003). By applying SMA with endmembers for green vegetation and non-photosynthetic vegetation the live ratio of the vegetation could be computed directly and used as an input for the FPI algorithm. Such an approach would eliminate the need for calculating RG using a long time series of remote sensing imagery.

5. Conclusions

This study investigated the use of MODIS remote sensing data for assessing the Fire Potential Index in southern California. Compared to the NOAA-AVHRR data traditionally used for this task, the MODIS sensor offers a finer spectral resolution and therefore facilitates the use of other vegetation indices besides NDVI. The first part of this paper thus focuses on using the vegetation index VARI for the computation of relative greenness (RG), which is a fundamental component of the FPI algorithm. Utilizing LFM measurements sampled at various locations throughout Los Angeles County for validation, it was shown that for the typical shrub vegetation of southern California VARI-derived RG corresponds more closely to changes in live fuel moisture than NDVI-derived RG. It has therefore the potential to provide better results than the latter in the computation of the FPI in such areas.

Based on the conclusions from the first part, the second part of the paper focuses on the implementation of a MODIS-based system for computing the FPI in southern California. Using historical wildfire data from the MODIS active fire product as a reference, FPI based on both the traditional NDVI-RG and VARI-RG was validated over a 5 year period. Analysis of fire and no-fire classes for both approaches indicate that VARI-based FPI performs better at distinguishing between fire and no-fire events. It shows a better separability between the medians of the classes and in addition the FPI values for the fire class are limited to a narrower range of high FPI values. Furthermore a logistic regression was carried out and showed that the model fitted to VARI-FPI was superior to that fitted to NDVI-FPI. The model deviance was 176.2 for the former and 408.8 for the latter. The c-index was computed as 0.78 and 0.69, respectively. Overall MODIS imagery appears to have the potential for improvements in existing wildfire susceptibility algorithms by providing access to locally more appropriate vegetation indices than NDVI for computing relative greenness. Using the VARI vegetation index can improve maps of Fire Potential Index for southern California.

Acknowledgements

This work was supported by NASA Headquarters under the Earth System Science Fellowship Grant NNG-04GG77H. The first author would like to thank UCSB’s VIPER Lab for providing the MODIS composites as well as advice and support. The authors would like to express their gratitude to Bob Burgan (retired from Missoula Fire Sciences Lab) for the FPI source code and valuable advice. Furthermore, the authors also acknowledge the contribution of four anonymous reviewers, whose comments substantially improved the quality of the manuscript.

References


