THIS IS A POST
PRINT

Published in journal
Modelling Earth
System & Environment

Deriving Forest Fire Probability Maps From the Fusion of Visible/Infrared Satellite Data and Geospatial Data Mining

Prashant K. Srivastava^{1*}, George P. Petropoulos^{2,3}, Manika Gupta⁴, Sudhir K. Singh⁵, Tanvir Islam⁶ and Dimitra Loka³

¹ Institute of Environment and Sustainable Development, Banaras Hindu University, India, prashant.iesd@bhu.ac.in

² University of Aberystwyth, Department of Geography and Earth Sciences, Wales, UK, gep9@aber.ac.uk

³ Department of Soil Water Resources, Institute of Industrial & Forage Crops, Hellenic Agricultural Organization "Demeter" (former NAGREF), Directorate General of Agricultural Research, 1 Theofrastou St., 41335, Larisa, Greece, petropoulos.george@gmail.com; dimloka@gmail.com

⁴ Department of Geology, University of Delhi, Delhi, India, manikagup@gmail.com

⁵ K. Banerjee Centre of Atmospheric and Ocean Studies, IIDS, Nehru Science Centre, University of Allahabad, Allahabad-211002, India, sudhirinju@gmail.com

⁶ NASA Jet Propulsion Laboratory, Pasadena, CA, USA, Tanvir.Islam@jpl.nasa.gov

*Correspondence: prashant.just@gmail.com; prashant.iesd@bhu.ac.in

Abstract: Information on fire probability is of vital importance to environmental and ecological studies as well as to fire management. This study aimed at comparing two forest fire probability mapping techniques, one based primarily on freely distributed EO (Earth observation) data from Landsat imagery, and another one based purely on GIS modeling. The Normalized Burn Ratio (NBR) computed from Landsat data was used to detect the high fire severity and probability area based on the NBR difference between pre- and post- fire conditions. The GIS-based modeling was based on a Multi Criterion Evaluation (MCE) technique, into which other attributes like anthropogenic and natural sources were also incorporated. The ability of both techniques to map forest fire probability was evaluated for a region in India, for which suitable ancillary data had been previously acquired to support a rigorous validation. Subsequently, a conceptual framework for the prediction of high fire probability zones in an area based on a newly introduced herein data fusion technique was constructed. Overall, the EO-based technique was found to be the most suitable option, since it required less computational time and resources in comparison to the GIS-based modeling approach. Furthermore, the fusion approach offered an appropriate path for developing a forest fire probability identification model for long-term pragmatic conservation of forests. The potential fusion of these two modeling approaches may provide information that can be useful to forest fire mitigation policy makers, and assist at conservation and resilience practices.

Keywords: Multi Criteria Evaluation, Forest Fire, Data Fusion, Normalize Burn Ratio, Remote Sensing, Geographical Information Systems, India

1. Introduction

Wildfires act as a triggering factor for ecosystems changes, by affecting ecosystems' structure and vegetation patterns and consequently their overall functions and processes (Knorr et al. 2011; Lhermitte et al. 2011; Karamesouti et al. 2016). On a global scale, about 350 million hectares of land are annually affected by fire events (van der Werf et al. 2006) and wildfires play an important role in the evolution, organization and distribution of ecosystems (Ireland et al. 2015). Therefore, characterization of the fire probability, especially on an operational basis, can provide substantial information related to ecology and biodiversity which can significantly assist in understanding fire behavior of a potentially affected area (Bowman et al. 2010; Chuvieco et al. 2010). To this end, Earth Observation (EO) is a very well-suited technology, since it is able to provide information that can be used in diverse applications related to fires at a range of spatial and temporal scales (Alexandridis et al. 2009; Turner et al. 2003; Vadrevu et al. 2006; Kalivas et al. 2013). Copernicus is the umbrella name for a number of satellite missions that include optical instruments, altimetry systems, radiometers and spectrometers (Borgeaud et al. 2015). The wide area coverage, which is repetitively provided by satellite sensors, as well as the information on non-visible spectral regions, makes them a very valuable tool at different stages of fire management (Pereira et al. 2011; Vadrevu et al. 2009). Additionally, since EO products have become increasingly accessible, wildfire probability and affected areas can be mapped with relatively low labor-intensive costs over large areas (Vhengani et al. 2015).

In recent years, several studies have focused on fire severity mapping and burned area differentiation, mostly in the Mediterranean area, by combining spectral information acquired in the red, Near Infrared (NIR) and thermal (TIR) bands, with successful results (Petropoulos et al. 2014; Mallinis and Koutsias 2012). The most common indices used have been based on Red and NIR bands, for both pre- and post- fire mapping using the Normalized Difference Vegetation Index (NDVI) (Illera et al. 1996; Chuvieco et al. 2004). Several studies have also proposed combining the NIR and Mid-InfraRed (MIR) observations for developing indices such as the Normalized Burn Ratio NBR (calculated before a fire) (Key and Benson 1999), post-NBR (after fire - a unitemporal approach) and the dNBR technique (a bi-temporal approach, namely the difference between pre and post fire) (Roy et al. 2006; Escuin et al. 2008). The results of such studies have highlighted the potential of this index for identification of burned areas, with successful application over various part of the world (Escuin et al. 2008; van Wagtendonk et al. 2004; Epting et al. 2005; Fernandez-Manso et al. 2016; Imperatore et al. 2017; Maselli et al. 2000).

Spatial data, often collected, managed, analyzed and presented from EO sensors, as well as outputs of geospatial modeling, provide useful results for understanding and modeling the probability potential of a given area at different observation scales (Pamdey et al. 2012; Srivastava et al. 2012a, b). In particular, multi-dimensional decision and evaluation models provide tools for analyzing the complex choice alternatives with different environmental impacts. In this framework, multi-criteria evaluation (MCE) has shown to be a very promising and widely used technique in spatial data analysis, which can be employed to investigate a number of choice possibilities and evaluate suitability according to the associated weight of each factor (Jaiswal et al. 2002; Gupta and Srivastava, 2010; Vadrevu et al. 2010).

Nevertheless, to our knowledge, until now only a small number of studies has been concerned with performing comparisons of fire probability predictions between different modeling approaches and data fusion performances. Data fusion often occurs in remote sensing and geospatial modeling, where many images of the same scene exist at different resolutions and bear different properties (Pohl and Genderen, 1998; Whyte et al. 2018). In this context, this study aims to evaluate the use of two forest fire probability mapping techniques, whose operation has been based on totally different principles. The first technique is based on using primarily freely distributed EO Landsat imagery, whilst the other one is based purely on GIS modeling. The added value of the spatial data fusion between the two approaches, mainly based on weighted overlay technique on a GIS platform by utilizing the EO based severity and GIS based fire probability datasets, is also explored herein. As a case study has been used the Jambughoda Wildlife Sanctuary, located in India and characterized by a large number of fires each summer (Joseph et al. 2009; Saha 2002; Kiran et al. 2006).

2. Materials and Methods

2.1. Study area

The Jambughoda Wildlife Sanctuary (22.3667 N, 73.7167 E) is located in India and covers an area of 130.38 Km². It is situated between the [Halol](#) and Jambughoda Taluk of [Panchmahals District](#) and Sankheda Taluk of [Vadodara District](#) in [Gujarat](#) at an average altitude of approximately 145 m (**Figure 1**). Established in 1990, it is a magnificent forest of teak (*Tectona grandis*), bamboos (*Bambusa vulgaris*) and other miscellaneous fauna and flora species of significant ecological importance. Numerous rare and endangered species are found in the area and the sanctuary is home to sloths (*Folivora spp.*), bears (*Ursidae spp.*), leopards (*Panthera pardus*), hyenas (*Hyeonidae spp.*), wolves (*Canis lupus spp.*) and barking deers (*Muntiacus muntjak*). Summer is hot with temperatures often reaching 45°C. The monsoon period arrives in June and the place turns lush green during this season.

No past severe forest fires have been reported in the study area, but in the last few years, due to continuous clearing of tree growth for cultivation and heavy grazing, some forest fires have been reported which have resulted in reductions of natural regeneration of natural forest tree species. During the survey, we found that most of the fire events in the sanctuary are mainly anthropogenic, followed by natural causes such as high temperatures and lightning. Since this is the first time that this sanctuary becomes a study area, no previous references exist in order to support the case. This study area provides an ideal location to evaluate the capability of the two different fire probability modeling approaches and to obtain valuable results for its environmental management practices, due to its highly ecological, cultural and economic significance. Indeed, outputs of this study can be of important practical value to the forest department and policy makers in order to prevent or minimize forest fire probability areas that occur in and around the forest, and contribute to the protection of this important habitat in the future.

2.2. Datasets

LANDSAT satellite images of Jambughoda wildlife sanctuary were utilized in this study to map the fire probability at our experimental site. In our study, two LANDSAT images (path: 148, row: 44) with acquisition dates 23rd October 2009 and 01st April 2010 were used, registered to a UTM 43N WGS84 ellipsoid. Images were obtained atmospherically corrected, from the United States Geological Survey (USGS) archive (<http://glovis.usgs.gov/>). All images were also geometrically corrected and resampled with elevation correction (Nioti et al. 2011; Singh et al. 2014).

2.3. Methodology

All analysis was carried out using ENVI (v. 5.1, ITT Visual Solutions) and ArcGIS (v. 10.2, ESRI) software platforms. An overview of the methodology adopted to satisfy the study objectives is depicted in **Figure 2**.

2.4. GIS-based fire probability modeling

The GIS based *multi-criteria evaluation* (MCE) was used for spatial resource allocation (Charabi and Gastli, 2011). Generally, in the MCE approach standardized factors are taken into account by combining weighted linear parameters, which involves the factor being multiplied by a weight, while at the last summation all weighted linear parameters are utilized to produce a multi-criteria solution (Gorsevski et al. 2012). Saaty (1980) suggested the pair wise comparisons for the development of such weights. In this process, a 9-point reciprocal scale is used for each factor based on its relative importance to every other factor which results into a $n \times n$ matrix of ratings (where n is the number of factors being considered) (Saaty and Vargas, 2013). The evaluations of weights are mostly judged by the principal eigenvector of the matrix. Eastman (1993) had implemented this procedure on a GIS platform with a modification that allowed the degree of consistency to be evaluated as well as the location of inconsistencies; thus,

allowing for an orderly reevaluation based on an iterative solution which converges on a consistent set of consensus weights. The studies, amongst others, concluded that in order to judge the suitability of ranking and corresponding weight of the parameters and deciding whether it is feasible or not, a rigorous analysis using MCE is very much required before finalizing the final weights (Srivastava et al. 2012b). The MCE approach used in this study is summarized as follows (Saaty 1980):

The priorities of the pair wise criteria can be obtained by using Eq. 1:

$$Me = \lambda_{\max} e \quad (1)$$

where $\lambda_{\max}$ = largest eigenvalue of the matrix; M = eigenvector “ e ”.

The precision of the weighting analysis is expressed by computing the Consistency Ratio (CR) and Consistency Index (CI). CR is calculated to determine whether the evaluation is successful or not. A CR less than 0.1 indicates good consistency (Saaty 1980). If it is more than 0.1 the comparison is inconsistent and needs to be re-assessed. The consistency ratio (CR), can be calculated as:

$$CR = CI / RI \quad (2)$$

where CI is the consistency index defined as,

$$CI = (\lambda_{\max} - n) / (n - 1) \quad (3)$$

where n is the number of selected parameters; and RI is the Random Inconsistency Index. The mean RI can be obtained by averaging the CI s of many randomly generated pair-wise comparison matrices.

ArcGIS (v9.3) software was used for the preparation of primary thematic layers. The Survey of India (SOI) topographical sheet, in a 1:50,000 scale was utilized to extract drainage and the contour layer. The contour map was used to generate DEM, which in turn was exploited in deriving the slope map which allowed the understanding of the region's terrain properties. All input datasets were reclassified by assigning suitable ranking to each thematic feature after considering their relative contribution characteristics towards the occurrence of forest fire probability area (Gupta and Srivastava 2010). Knowledge-based weighting assignment was performed on the basis of influence of each governing feature on forest fire occurrence. After weighting assignment, all the reclassified thematic layers were integrated with one another through GIS, using the weighting overlay analysis in Arc GIS 9.1 platform to generate forest fire probability areas.

2.5. Pre-processing of satellite data

Sensor spectral bands, except for the thermal infrared band (i.e. band 6), were layer stacked to form a single image file corresponding to the acquisition date of the dataset. In the next step, atmospheric correction was performed to both images using the 6S radiative transfer code (Ju et al. 2012). 6S requires water, ozone and the aerosol optical depth as input datasets to obtain the conversion equation from the observed radiance to actual reflectance (acr). In this code, acr is calculated according to Vermote et al. (1997) as :

$$y = xa * (R) - xb \quad (4)$$

$$acr = y / (1 + xc * y) \quad (5)$$

where R is the observed radiance ($W/(m^2 \text{ sr Am})$), and xa , xb , and xc are the coefficients obtained from the calculation, where, xa , xb and xc represent inverse of the transmittance, scattering and the reflectance of the atmosphere respectively (Tachiiri 2005; Vermote and Kotchenova 2008).

Herein, data from AErosol Robotic NETwork (AERONET) is used for aerosol optical depth (<http://aeronet.gsfc.nasa.gov>) and the status of ozone (expressed in Dobson Units) in the atmosphere was obtained from NASA website (http://toms.gsfc.nasa.gov/ozone/ozone_v8.html). Water vapour data was retrieved from the global ECMWF ERA interim reanalysis data centered over the Jambughoda National Park and can be downloaded from <http://www.ecmwf.int/>. The two bands were radiometrically normalized before the application of 6S code on a FORTRAN platform. This was done by converting the digital count (DC) of LANDSAT-5 TM data to satellite radiance. The recorded digital values of sensors were converted to reflectance (R_i) at the level of Top Of Atmosphere (TOA) using the following equations:

$$R_i = (L_i * \pi * d^2) / (E_{si_b} * \cos(z_s)) \quad (6)$$

$$L_i = DN_i * G_b + B_b; \quad (7)$$

DN_i = per-pixel raw brightness value; R_i = unitless TOA or planetary reflectance; G_b and B_b = per-band gain and bias from scene header; L_i = Spectral radiance at sensor's aperture; d^2 = earth-sun eccentricity from lookup Table; E_{si_b} = mean solar exo-atmospheric spectral irradiance; z_s = solar zenith angle in degrees.

The xa , xb and xc calculated using 6S are summarised in Table 1. The bidirectional reflectance distribution function (BRDF) coupling correction was used with surface target elevation at 1013 mb pressure level. All the three 6S output terms are provided to the equation 1 and 2 to derive the atmospherically corrected reflectance pixel by pixel using the ENVI band math. The values of xa and xb showed an increasing trend from DOY296 (winter season) to DOY91 (summer season). A significant variation in values can be attributed to the aerosol optical thickness and alterations in water vapour content. However, a small variation is observed in the xc value. As expected, the summer season indicates a higher radiance compared to the winter season. Following the atmospheric correction, image to image co-registration between the two LANDSAT images was performed in ENVI (Banerjee and Srivastava 2013; Srivastava et al. 2012c) by taking the 2009 as reference year. Approximately 40 ground reference points were selected from easily detectable corner points (e.g. road junctions). Finally, an image subset was performed extracting from the acquired image a smaller area covering the study region.

2.6. EO-based Forest fire severity and probability assessment

In order to achieve effective conservation of the forest and its related ecosystem, it is essential to analyze the fire severity over the ecosystem (Escuin et al. 2008) and denote the changes, caused by wildfire, on soil and vegetation (Epting et al. 2005). Forest fire severity and probability estimation provides valuable information on the spatial dimension of the fire-affected area. Additionally, it facilitates the process of decision making focused on restoration or resilience of the fire affected areas (Escuin et al. 2008; Spies et al. 2012). The SWIR and NIR bands of LANDSAT TM have been used by several researchers for forest high forest fire hazards like wildfire (Epting et al.

2005). The NBR index has been shown to be negatively correlated with the burn severity and vice versa (De Santis et al. 2010). Furthermore, fire severity analysis provides information about the post-fire vegetation succession growth pattern. Hence, herein, the NIR-MIR based indices are taken also into account for forest fire severity assessment. The NBR based indices were computed following the pre-processing and atmospheric correction of the satellite images. The Normalized Burn Ratio (NBR) was computed from each LANDSAT image using Band4 (Near Infrared: 0.76-0.90 microns) and Band 7 (Short-Wave Infrared: 2.08-2.35 microns), as described in equation 8 below according to Epting et al. (2005):

$$NBR = 1000[(Band4 - Band7)/(Band4 + Band7)] \quad (8)$$

The dNBR was subsequently derived by subtracting postNBR from preNBR:

$$dNBR = NBR_{pre} - NBR_{post} \quad (9)$$

Unaffected areas typically have $dNBR$ values near zero, indicating no change, while forest fire areas have positive $dNBR$ values.

2.7. Data fusion

Data fusion is used in many GIS models, for example to perform merging and other raster calculations. It combines multiple data sources so that the resulting product makes the best use of each source's strength (Whyte et al. 2018; Wald et al. 1997). For instance, GIS images can be produced from topographical sheets or natural resources maps, where datasets are made up of different types and information. Image fusion is the fusion of two or more images into a single, more complete or more useful image that bears the properties of all the used images (Pohl and Gendereen 1998). In GIS, generally, the data are stored either in vector (a scale-invariant collection of edges and vertices) or raster (array of cells corresponds to area-support) form (Neteler and Mitasova 2008). Several mathematical operations between the raster's datasets with different supports, such as union, intersection, zonal-averaging, and pixel-by-pixel computations, can be used implicitly for fusion operations that are based mainly on two separate processes: interpolation and combination (Nguyen et al. 2012). This method has the advantage that it is very fast and scalable.

In weighted overlay technique for data fusion the users have to define the problem first followed by breaking the model into sub-models with identification of the input layers. For weighted overlay analysis a ranking between 1 and 9 was used. Lastly, the multiplication of input criteria and the weights were performed following the AHP-MCE analysis and then added together to predict the forest fire zones. In our study, the following equation was used for weighted overlay method for data fusion (S):

$$S = \sum_{i=1}^n w_i x_i \quad (10)$$

where w_i = The weight of i_{th} factor map; x_i = criteria score of class of factor i ; S = fused index for each pixel in the map.

2.8. Descriptive and performance statistics

A set of descriptive statistics for the EO-based approach was generated from the variables pre-NBR, post-NBR and dNBR. Those included the mean, median, lower quartile (25%), upper quartile (75%), standard error and the standard deviation. Skewness and Kurtosis measurements were used for symmetry and to check the distribution

in the dataset. A Kolmogorov-Smirnov (K-S) distribution value was also generated and a statistical significance for the difference within groups was established at probability level of < 0.001. Pearson and Spearman's correlation statistics were calculated to understand the relationship between the variables. To quantify the performance of methodology for forest fire zone prediction, a simple expression was used and expressed in terms of percentage between predicted and observed datasets, as shown in equation 11:

$$\% Accuracy = \left[1 - \left(\frac{(total\ obs - total\ true\ pred)}{total\ obs} \right) \right] \times 100 \quad (11)$$

2.9. Natural and anthropogenic factors and GIS integration

The forest fire probability of any area depends on natural and anthropogenic factors, such as vegetation type/density, climate, distances from roads, topography, proximity to settlements/agriculture, drainage density, aspect and slope of the area (Albert-Green et al. 2012; Castedo-Dorado et al. 2011). The rating factor varies from 1 to 9, in which a higher rating indicates that the factor has a high degree of influence on the forest fire probability in an area, whereas influence factors indicate the overall weight of the parameters for forest fire probability assessment (**Table 2**).

A total of 25 villages are found in the area with mostly tribal population in five blocks and two ranges. For their livelihood, they collect minor forest products and use wood as fuel. Road development through this closed forest areas has also imposed an increased probability of forest fire pockets creation by opening access to settlers or gatherers. Consequently, human activities on forest roads and settlements near forested areas create chances for accidental forest fire in the region. In order to delineate the possible impact of these features, the digitized map of the road and railway was processed into three buffer zones of 100m, 200m and 300m and the classes obtained were on the basis of the forest's proximity from the roads/railways (**Fig. 3**). A 100m buffer distance is nearer to the road but farthest from the forest, 200m is farther from the road but at an intermediate distance from the forest and 300m distance is farthest from the road but nearest to the forest. This way, three categories were formed named as far from the forest, at intermediate distance from forest and near to the forest. Similarly, corridors of 1000m, 2000m and 3000m perimeter were created around the settlement locations and agricultural area, and the categories were classified as far from the forest, at intermediate distance from the forest and near to the forest (**Fig. 3**). The weight of the data calculated and CR are shown in a separate pair wise comparison matrix (**Table 2**). All the CR values are well below 0.1 and hence indicate a specific estimation of normalized weights.

In the study area, the bamboo dominated area, which covers almost 40% of the forest, showed extremely high sensitivity to generation of forest fire pockets. Hence, higher rating and weights were assigned followed by other dominant vegetation forms, as shown in the pair-wise matrix (**Table 2**). The other high oil-containing species like rosewood (*Dalbergia spp.*), butter tree (*Vitellaria paradoxa*) etc are well distributed in the area and these species when dried can create a forest fire area even with very small ignition. On the other hand, moist vegetation has lesser forest fire probability than dry vegetation. The thematic map of forest vegetation category map is illustrated in **Fig. 3**. The individual pair wise comparison matrix (**Table 2**) indicates highest chance of fire in the bamboo dominated part of the forest, with highest normalized weight of 50.43%, and least in the densely mixed forest (3.46 %). Water bodies' areas were found to have no role in forest fire pockets generation; hence no weights were assigned to water body areas in determination of forest fire probability area.

2.10. Drainage density and Topography

Drainage density is determined by dividing the total length of streams within a drainage basin by the drainage area (Horton 1945). Low densities allow for long residence times; therefore, abstraction mechanisms have more time to remove water (Patel et al. 2011). Thus, low drainage density was given the minimum weight since these areas are supposed to contain more soil moisture than high drainage density area. The drainage density map (**Fig. 4**) indicates that only smaller parts of area (South, South-East and North East) fall under the low drainage density category. The majority of the area falls under moderate drainage density category. Higher drainage density was assigned a higher weight followed by moderate and low drainage density and the normalized weights of the three categories are 71.32%, 22.00% and 6.69%, respectively, with CR ratio of 0.020 (**Table 2**).

Wind behavior depends on topography and affects the forest fire of the area. For calculation of topography and slope, the Digital Elevation Model (DEM) of the study area was used, as indicated in **Fig. 4**. Further aspects of the area derived from DEM, were classified into eight classes namely north, north-west, west, south-west, north-east, east, south-east and south (**Fig. 4**). The highest weights were given to north, north-west, west and south-west whereas lowest weights were assigned south, east, south-east and north-east. The consistency ratio for the pair comparison matrix was 0.045 (**Table 2**). Steep slopes have a higher forest fire potential in comparison to gentle slopes, which are in turn higher than level ground. An east aspect, generally, gets more direct sunlight than a west aspect. It also specified that in this area north, north-west, west and south-west slopes are generally subject to a greater threat of forest fire probability as the slopes contain drier fuel more readily available for combustion. On the other hand, north-east, east, south-east and south slopes are subject to a lower threat since the slopes contain moist fuel, due to hill shadow and moist wind conditions from the Arabic Sea. As a result, there are two classes generated for slope thematic layer, for example low slope area (<5%) and high elevation area (> 5%) (**Fig. 4**). The weights deduced for high slope area are 88.9% and low slope area it is 11.1%. The consistency ratio for the pair comparison matrix is 0.000 (**Table 2**).

2.11. Forest fire probability micro-zonation

A combination of AHP-MCE based weighting assignment and GIS based weighted overlay technique is a reliable way to identify the vulnerability of specific areas to fire and create a fire probability model. Such a model is reported in this section, in which the forest fire probability is a micro-zonation probability map that is constructed according to the criteria mentioned and described in the subsections 3.1-3.2. This Forest Fire Probability Zone Model (FFPZM) is a useful tool for minimizing wildfire probability and damage, and detecting most probable areas where forest fires may occur; thus, allowing forest fire managers to identify the locations easily and manage these areas effectively. For the forest fire probability zone model, vegetation was given the highest normalized weight (38.6%), followed by proximity to roads/railways (24.9%), settlements (15.55%), slope (9.5%), aspect (5.65%), agriculture (3.44%) and drainage density (2.35%). Furthermore, consistency ratio was calculated for the rank and weighting assignment for all the criteria; the reported value of 0.061 indicates that the weights taken for analysis are precise for estimation.

The above-mentioned criteria were based on the probability of fire spreading and the likelihood of its occurrence. The vegetation factor was provided with the highest weight as it provides source and fuel in the event of forest fire (Jaiswal et al. 2002). The second highest weight was assigned to road/railways, since tourists may intentionally or unintentionally provide ignition materials for forest fire during their travels or recreational activities. Similarly, nearby settlements can also cause forest fire due to their day-to-day activities; hence, they were assigned the third highest weight in the criteria. The slope influences the probability of spreading, therefore, it ranked fourth in the factors. Additionally, aspect plays an important role in forest fire, since fires occur mostly on the brighter side of the hill than its shadowed portion, which can be decided based on aspect map. Consequently, the fifth highest weighting in the category was allotted to the aspect. Since the agricultural areas are located near forests in Jambughoda, and farmers occasionally create fires in forests in order to increase their cropping area (Jaiswal et al. 2002), human

activities are the main factors related to forest fire incidence at locations closest to agricultural area. Subsequently, this factor was assigned to the 6th rank in the category. Lastly, the drainage density provided the least weight because of its minimal role in forest fire occurrence. A summary of all the attributes is provided in Table 2. The following equation was used for the generation of the Forest forest fire Probability Zone Model (FFPZM):

$$FFPZM = \sum_{i=1}^n W_i P_i \quad (12)$$

where, W_i denotes the weightage of the given layer and P_i denotes the individual rasterized pixel value of the input maps.

So the relation of *FFPZM* can be expressed as:

$$\mathbf{FFPZM} = 38.6 * \text{Vegetation} + 24.9 * \text{Proximity to roads} + 15.5 * \text{Proximity to settlements} + 9.5 * \text{Slope} + 5.65 * \text{Aspect} + 3.44 * \text{Proximity to Agriculture} + 2.35 * \text{Drainage density} \quad (13)$$

Final forest fire probability map (**Fig. 5**) was categorized into three classes i.e., low, moderate, and high. The thematic map of FFPZM indicates a higher forest fire probability in West and North-West part of the wildlife sanctuary, which consists mainly of bamboo dominated forest. The areas classified as high and moderate ‘forest fire probability’ zones are those areas where forest fire pockets can be unintentionally caused by anthropogenic activities.

3. Results and Discussion

3.1. EO-Based Identification of Fire Severity and Probability

The NIR-MIR based indices were calculated such as pre-NBR, post-NBR and dNBR after 6S atmospheric correction. The statistical analysis of data suggested a highest mean value of dNBR and pre-NBR than post-NBR, indicating a shifting of forest fire areas in-between the time period analyzed (Table 3). The analysis of lower quartile (25%), median and upper quartile (75%) showed a very high variability in the data, which could be due to phenological and soil moisture changes that occurred from 2009 to 2010. The results showed that the NBR in 2010 (post-NBR) had a small range, but in 2009 (pre-NBR) a broader range was observed in the data values. The standard error estimate in post-NBR was minimum as compared to pre-NBR and dNBR, indicating a higher deviation of post-NBR data than other two datasets. The analysis of post-NBR and dNBR datasets showed a leptokurtic distribution; by contrast pre-NBR showed a platykurtic distribution. Further, the datasets of pre-NBR and dNBR showed a skewed left distribution opposite to post-NBR, which was observed as skewed right. The result of K-S distribution showed that significant differences existed in the population at <.001 level of significance (**Table 3**).

In **Fig. 6 (a-b)**, Band 4 and band 7 reflectance exhibits greatest magnitude and variation of change compared to other bands. The result is that NBR images appear relatively ‘flat’ in the cases of pre-NBR and dNBR, while spectral differences between the bands are enhanced. In the grayscale NBR images, living vegetation had high reflectance in band 4, while band 7 is largely absorbed, so NBR is showing a strongly positive and bright appearance. In **Fig. 6 (c-e)** the spectral profile of dNBR, pre-NBR and post-NBR are illustrated. The spectral analysis of the dataset shows that pre-NBR is closer to dNBR, indicating the similarities between the two datasets. The overall analysis of dNBR

dataset shows a higher frequency of moderate severity class in the area. In this study, temporal difference of NBR is hypothesised to represent optimal separation of high forest fire pockets from lower ones. Some pixels have less dNBR corresponding to the unburned class while the intermediate and higher values belong to the categories moderate and extreme, respectively. The water bodies are chosen as reference level since very low heat is expected over this area and hence can be classified as a non- probability area.

In order to better understand the results, a correlation matrix plot between the pre-NBR, post-NBR and dNBR is generated (**Fig. 7**). Both Pearson (r) and Spearman (r_s) correlation statistics are computed to see the linearity and non-linearity between the datasets respectively. The analysis indicates that the correlations between dNBR versus pre-NBR ($r=0.86$; $r_s=0.87$) are very high as compared to dNBR versus post-NBR ($r=-0.27$; $r_s=-0.15$). The dNBR and post-NBR are negatively correlated to each other. However, the distribution and r_s statistics suggest that these two are following a more non-linear behavior as a higher value of r_s correlation coefficient is obtained in comparison to r . Similar information was also obtained by Escuin et al. (2008). The analysis suggests that a comparable r and r_s relation exists between the variables; hence the simplest r statistics could be enough for defining the relationship. On the other, both pre- and post-NBR are showing feeble positive correlation coefficients with each other; this indicates that their relationships are governed more by seasonality. The results of the analyses of the pre-, post- and dNBR behavior suggest that bi-temporal index dNBR is the most suitable for carrying out the fires severity mapping as it is a combined index from pre and post NBR. Thus, this can be used to discriminate the unburned and fire-burned pixels. Based on the observation of the dNBR behavior and to assess the fire severity levels, the dNBR is categorized into 3 classes, unburned, moderate and high, using the K mean unsupervised classification scheme (**Fig. 8**). The datasets distribution suggested that the unburned class occupied 27.91% of the area, whereas the moderate and high category indicated coverage of 37.14% and 34.95%, respectively. The forest fire pockets in the areas belong mainly to the moderate and high forest fire pockets category. The situation in the north is reflected by the initial decrease of NBR. It is remarkable that the pre-NBR values are consistently higher in the South-East part, corresponding to the high burn severity class, compared to the area subjected to the prescribed burn in the northern part. The average dNBR within the area indicates a gradual increase from low towards higher severity class. This demonstrates that the severity maps are the most suitable for expressing the discrimination between high and low energy. The overall analysis indicated that the EO-based method requires considerably less computational time than the GIS-based one and can be promptly used for prevention from forest fire probability disaster.

3.2. Data fusion and Outputs Validation

The EO based methodology presented herein is based on seasonality and represents the influence of climate on forest fire severity. GIS based delineation of areas is based on natural and human induced factors describing, as a result, more realistic situations. These two assessments i.e. the AHP-MCE based GIS FFPZM and dNBR (a bi-temporal index) could be potentially combined together and provide results which can be used by the forest officials for long term pragmatic conservation of forest and reducing the uncertainty involved in these types of studies. For results validation, the high forest fire probability areas (burnt areas) were identified with the help of field surveys and their geographical locations were recorded with the help of a Garmin made GPS. Approximately four points, depicting some events were collected from the study site, as it is shown in **Fig. 9**. The observed burnt areas showed a close relationship with simulated results. Specifically, in the case of satellite based approach, three out of four points were within high dNBR areas, designated as high forest fire severity and probability, indicating an efficiency of 75% ,while for the GIS based modeling approach, two points were within the high probability area and two within the low, indicating an efficiency of 50%.

For data fusion map, tone scale should not include more than seven tone values in cartographic representation (Erden and Coskun 2010). As a result, in this study a total of 3 classes are established for all assessment. Following criteria value ranges, equal interval breaks procedure is used for the ranking of the class values from 1 to 3. Equal interval classification is taken into account to create a user friendly legend and generally appropriate for

continuously distributed data. The layer generated by GIS maps are raster-based data structures, therefore after criterion map layers generation, vector-based polygon data structures are converted to raster. After all pre-processing, a GIS-grid-based fire hazard model is developed, following the weighted data fusion technique, to determine the level of severity of fire hazard zones, in terms of vulnerability to fire by assessing the relative importance between forest fire probability factors and the location of fire severity. In the next step, the thematic layers obtained from the GIS and EO-based methodology are fused together in order to generate the maximum risk potential area, which needs intense conservation and restoration practices. After all procedures are completed, a fusion map for the determination of optimal sites for new fire station constructions is generated. The thematic fused maps, obtained from all the above procedures, indicate the locations of forest fire probability area, as well as, the possible severity of the burned area.

The overall analysis of the fused maps, indicates that the western region of the Jambughoda sanctuary needs more attention with respect to forest fire and severity mitigation practices. The construction of new fire station is proposed at the intersection of number 1 in **Fig. 9**. Construction at site number 1 is required because it is very close to a residential area and a large fire can endanger both human lives and property, along with forest wildlife. Construction of one more fire station is suggested at site number 2 (**Fig. 9**), since this area is also highly populated and bears similar characteristics to station number 1. Future plans should also include the construction of two more fire stations at sites number 3 and 4.

4. Conclusions

In this study, an investigation of the capability of two forest fire probability mapping methods, one based on the analysis of Landsat imagery and another based on GIS modeling, was performed using the Jambughoda Wildlife Sanctuary in India, as case study. A conceptual framework was also constructed in order to predict the zones of high fire probability areas which were presented based on a data fusion technique. A participatory multi criteria decision making approach was used for identification of the important causative factors for forest fire. The final fused image, generated from LANDSAT TM freely distributed satellite images and GIS, showed a consistent ability to predict and estimate the micro-zonation of forest fire pockets in a typical forest dominated by tropical species.

Our results clearly indicated a highly dynamic and spatial nature of forest fire risk pockets in the study area, depicted by the fusion technique. Such information is important for addressing broad-scale forest fire probability hazard and also as a tactical guide that can help managers design effective forest fire probability control measures. In addition, our case study revealed the complex connections with environmental variables that a forest fire involves. Furthermore, our study findings provided significant insight on the ways that some environmental and social factors can affect the generation processes of forest fires, and can assist the preservation of the forest ecosystem of the region used here as a case study. Notably, the fusion approach employed here is robust and flexible enough to expand further or be implemented in other regions, after necessary adjustments are made. The latter is of considerable scientific and practical value to the wider scientific community, given the present and future open access to EO data from various sensors around the world.

References

- Albert-Green A, Dean C, Martell DL, Woolford DG (2012) A methodology for investigating trends in changes in the timing of the fire season with applications to lightning-caused forest fires in Alberta and Ontario, Canada. *Can J For Res* 43:39-45.
- Alexandridis TK, Lazaridou E, Tsirika A, Zalidis GC (2009) Using Earth Observation to update a Natura 2000 habitat map for a wetland in Greece. *J Env Manag* 90:2243-2251.

- Banerjee R, Srivastava PK (2013) Reconstruction of contested landscape: detecting land cover transformation hosting cultural heritage sites from central India using remote sensing. *Land Use Policy* 34:193-203.
- Borgeaud M, Drinkwater M, Sivestrin P, Rast M (2015) Status of the ESA earth explorer missions and the new ESA earth observation science strategy. *IEEE Int. Geosci Remote Sens Symp* 4189–4192.
- Bowman DM, Murphy BP (2010) Fire and biodiversity. *Conservation Biology for All* 163.
- Castedo-Dorado F, Rodríguez-Pérez JR, Marcos-Menéndez JL, Alvarez-Taboada MF (2011) Modelling the probability of lightning-induced forest fire occurrence in the province of León (NW Spain). *For Syst* 20:95-107.
- Charabi Y, Gastli A (2011) PV site suitability analysis using GIS-based spatial fuzzy multi-criteria evaluation. *Renew Energy* 36:2554-2561.
- Chuvieco E, Cocero D, Riano D, Martin P, Martinez-Vega J, De La Riva J, Pérez F (2004) Combining NDVI and surface temperature for the estimation of live fuel moisture content in forest fire danger rating. *Remote Sens Environ* 92:322-331.
- Chuvieco E, Aguado I, Yebra M, Nieto H, Salas J, Martín MP, Vilar L, Martínez J, Martín S, Ibarra P (2010) Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. *Ecol Modell* 221, 46-58.
- De Santis A, Asner GP, Vaughan PJ, Knapp DE (2010) Mapping burn severity and burning efficiency in California using simulation models and Landsat imagery. *Remote Sens Environ* 114:1535-1545.
- Eastman JR (1993) Idrisi Version 4.1. Worcester (USA), Clark University.
- Epting J, Verbyla D, Sorbel B (2005) Evaluation of remotely sensed indices for assessing burn severity in interior Alaska using Landsat TM and ETM+. *Remote Sens Environ* 96:328-339.
- Erden T, Coşkun, M (2010) Multi-criteria site selection for fire services: the interaction with analytic hierarchy process and geographic information systems. *Nat Hazards and Earth Syst Sci* 10:2127-2134.
- Escuin S, Navarro R, Fernandez P (2008) Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from LANDSAT TM/ETM images. *Int J Remote Sens* 29:1053-1073.
- Fernández-Manso A, Fernández-Manso O, Quintano C (2016) SENTINEL-2A red-edge spectral indices suitability for discriminating burn severity. *Int J Appl Earth Obs Geoinf* 50:170–175.
- Gorsevski PV, Donevska KR, Mitrovski CD, Frizado JP (2012) Integrating multi-criteria evaluation techniques with geographic information systems for landfill site selection: a case study using ordered weighted average. *Waste manag* 32:287-296.
- Gupta M, Srivastava PK (2010) Integrating GIS and remote sensing for identification of groundwater potential zones in the hilly terrain of Pavagarh, Gujarat, India. *Water Int* 35:233-245.
- Horton RE (1945) Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. *Geol Soc Am Bull* 56:275-370.
- Illera P, Fernandez A, Delgado J (1996) Temporal evolution of the NDVI as an indicator of forest fire danger. *Int J Remote Sens* 17: 1093-1105.

- Imperatore P, Azar R, Calò F (2017) Effect of the vegetation on backscattering: an investigation based on Sentinel-1 observations. *IEEE J Sel Topics Appl Earth Observ Remote Sens* 10: 4478–4492.
- Ireland G, Petropoulos GP, Carlson TN, Purdy S (2015): Addressing the ability of a land biosphere model to predict key biophysical vegetation characterisation parameters with Global Sensitivity Analysis. *Environ. Modell Software* 65:94-107.
- Jaiswal RK, Mukherjee S, Raju KD, Saxena R (2002) Forest fire risk zone mapping from satellite imagery and GIS. *Int J Appl Earth Observ Geoinfo* 4:1-10.
- Joseph S, Anitha K, Murthy M (2009) Forest fire in India: a review of the knowledge base. *J Forest Res* 14:127-134.
- Ju J, Roy DP, Vermote E, Masek J, Kovalsky V (2012) Continental-scale validation of MODIS-based and LEDAPS Landsat ETM+ atmospheric correction methods. *Remote Sens Environ* 122:175-184.
- Kalivas D, Petropoulos GP, Athanasiou I, Kollias V (2013) An intercomparison of burnt area estimates derived from key operational products: analysis of Greek wildland fires 2005-2007. *Nonlinear Process Geophys* 20:1–13.
- Karamesouti M, Petropoulos GP, Papanikolaou ID, Kairis O, Kosmas K (2016) Erosion rate predictions from pesera and rusle at a mediterranean site before and after a wild- fire: comparison and implications. *Geoderma* 261:44–58.
- Key CH, Benson NC (1999) The Normalized Burn Ratio (NBR): A Landsat TM radiometric measure of burn severity. USDA (Bozeman, Mont).(Available at <http://nrmssc.usgs.gov/research/ndbr.htm>).
- Kiran Chand T, Badarinath K, Krishna Prasad V, Murthy M, Elvidge CD, Tuttle BT (2006) Monitoring forest fires over the Indian region using Defense Meteorological Satellite Program-Operational Linescan System nighttime satellite data. *Remote Sens Environ* 103:165-178.
- Knorr W, Pytharoulis I, Petropoulos GP, Gobron N (2011) Combined use of weather forecasting and satellite remote sensing information for fire risk, fire and fire impact monitoring. *Comput Ecol Software* 1:112-120.
- Lhermitte S, Verbesselt J, Verstraeten WW, Veraverbeke S, Coppin P (2011) Assessing intra-annual vegetation regrowth after fire using the pixel based regeneration index. *ISPRS J Photogram Remote Sens* 66:17-27.
- Mallinis G, Koutsias N (2012) Comparing ten classification methods for burned area mapping in a Mediterranean environment using Landsat TM satellite data. *Int J Remote Sens* 33 :4408–4433.
- Maselli F, Rodolfi A, Bottai L, Romanelli S, Conese S Classification of Mediterranean Vegetation by TM and Ancillary Data for the Evaluation of Fire Risk. *Int J Remote Sens* 21:3303–3313.
- Neteler M, Mitasova H (2008) Open Source GIS: A GRASS GIS Approach, Number 773 in The International Series in Engineering and Computer Science. Springer, New York.
- Nguyen H, Cressie N, Braverman A (2012) Spatial statistical data fusion for remote sensing applications. *J Am Stat Assoc* 107:1004-1018.
- Nioti F, Dimopoulos P, Koutsias N (2011) Correcting the Fire Scar Perimeter of a 1983 Wildfire Using USGS-Archived Landsat Satellite Data. *GIScience Remote Sens* 48:600-613.
- Pandey PC, Sharma LK, Nathawat MS (2012) Geospatial strategy for sustainable management of municipal solid waste for growing urban environment. *Environ Monit Assess* 184:2419-2431.

- Patel DP, Dholakia MB, Naresh N, Srivastava PK (2011) Water Harvesting Structure Positioning by Using Geo-Visualization Concept and Prioritization of Mini-Watersheds Through Morphometric Analysis in the Lower Tapi Basin. *J Indian Soc Remote Sens* 40:299-312.
- Pereira P, Úbeda X, Martin D, Mataix-Solera J, Guerrero C (2011) Effects of a low severity prescribed fire on water-soluble elements in ash from a cork oak forest located in the northeast of the Iberian Peninsula. *Environ Res* 111:237-247.
- Petropoulos GP, Griffiths HM, Kalivas DP (2014) Quantifying spatial and temporal vegetation recovery dynamics following a wildfire event in a Mediterranean landscape using EO data and GIS. *Appl Geogr* 50: 120-131.
- Pohl C, Van Genderen J (1998) Review article multisensor image fusion in remote sensing: concepts, methods and applications. *Int J Remote Sens* 19:823-854.
- Roy DP, Boschetti L, Trigg SN (2006) Remote sensing of fire severity: assessing the performance of the normalized burn ratio. *Geosci Remote Sens Lett, IEEE* 3:112-116.
- Saaty TL (1980) Analytic hierarchy process, Wiley Online Library.
- Saaty TL, Vargas LG (2013) Criteria for Evaluating Group Decision-Making Methods. *Decision Making with the Analytic Network Process*. Springer.
- Saha S (2002) Anthropogenic fire regime in a deciduous forest of central India. *Curr Sci* 82
- Singh SK, Srivastava PK, Gupta M, Thakur JK, Mukherjee S (2014) Appraisal of land use/land cover of mangrove forest ecosystem using support vector machine. *Environ Earth Sci* 71:2245-2255.
- Spies TA, Lindenmayer DB, Gill AM, Stephens SL, Agee JK (2012) Challenges and a checklist for biodiversity conservation in fire-prone forests: Perspectives from the Pacific Northwest of USA and Southeastern Australia. *Biol Conserv* 145:5-14.
- Srivastava PK, Gupta M, Mukherjee S (2012a) Mapping spatial distribution of pollutants in groundwater of a tropical area of India using remote sensing and GIS. *Appl Geom* 4:21-32.
- Srivastava PK, Han D, Gupta M, Mukherjee S (2012b) Integrated framework for monitoring groundwater pollution using a geographical information system and multivariate analysis. *Hydrolog Sci J* 57:1453-1472.
- Srivastava PK, Han D, Rico-Ramirez MA, Bray M, Islam T (2012c) Selection of classification techniques for land use/land cover change investigation. *Adv Space Res* 50:1250–1265.
- Tachiiri K (2005) Calculating NDVI for NOAA/AVHRR data after atmospheric correction for extensive images using 6S code: A case study in the Marsabit District, Kenya. *ISPRS J Photogramm Remote Sens* 59: 103-114.
- Turner W, Spector S, Gardiner N, Fladeland M, Sterling E, Steininger M (2003) Remote sensing for biodiversity science and conservation. *Trends Ecol Evol* 18:306-314.
- Vadrevu KP, Eaturu A, Badarinath K (2010) Fire risk evaluation using multicriteria analysis—a case study. *Environ Monitor Assess* 166:223-239.
- Vadrevu KP, Badarinath K (2009) Spatial pattern analysis of fire events in Central India—A case study. *Geocarto Int* 24:115-131.

Vadrevu KP, Eaturu A, Badarinath K (2006) Spatial distribution of forest fires and controlling factors in Andhra Pradesh, India using spot satellite datasets. *Environ Monitor Assess* 123:75-96.

van der Werf G, Randerson J, Giglio L, Collatz G, Kasibhatla P, Arellano A (2006) Interannual variability of global biomass burning emissions from 1997 to 2004. *Atmosph Chem Phys* 6:3175–3226.

Van Wagtendonk JW, Root RR, Key CH (2004) Comparison of AVIRIS and Landsat ETM+ detection capabilities for burn severity. *Remote Sens Environ* 92:397-408.

Vermote EF, Tanré D, Deuze JL, Herman M, Morcette JJ (1997) Second simulation of the satellite signal in the solar spectrum, 6S: An overview. *Geosci Remote Sens IEEE Transactions* 35:675-686.

Vermote EF, Kotchenova S (2008) Atmospheric correction for the monitoring of land surfaces. *J Geophys Res* 113, D23S90.

Vhengani L, Frost P, Lai C, Booi N, van den Dool R, Raath W (2015) Multitemporal burnt area mapping using Landsat 8: Merging multiple burnt area indices to highlight burnt areas. *IEEE Int Geosci Remote Sens. Symp* 4153–4156.

Wald L, Ranchin T, Mangolini M (1997) Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images. *Photogram Engin Remote Sens* 63:691-699.

Whyte A, Ferentinos K, Petropoulos GP (2018) A new synergistic approach for monitoring wetlands using sentinels -1 and 2 data with object-based machine learning algorithms. *Environ Modell Software* 104: 40-54.

Author Contributions: Conceptualization, PKS; Methodology, , PKS, KS, MG, GP & SKS.; Software, KS, MG & SKS.; Formal Analysis & Validation, MG, SKS & TI.; X.X.; Writing-Original Draft Preparation, all co-authors.; Writing-Review & Editing, all co-authors.

Funding: This research received funding from the Commonwealth Scholarship Commission, British Council, United Kingdom and Ministry of Human Resource Development, Government of India.

Acknowledgments: The authors would like to thank the Commonwealth Scholarship Commission, British Council, United Kingdom and Ministry of Human Resource Development, Government of India for providing the necessary support and funding for this research. Authors gratefully acknowledge as well the United States Geological Survey (USGS) for the free access to the Landsat satellite images. Dr. Petropoulos's contribution to this work has been supported by the EU Marie Curie Project ENViSION-EO (project contract ID 752094) and the NERC's Newton Fund RCUK project Towards a Fire Early Warning System for Indonesia (ToFEWSI).

Conflicts of Interest: The authors declare no conflict of interest.